UNIVERSITY OF CALIFORNIA
Santa Barbara

Experimental Investigation of Nonlinear Intersubband Absorption and Many-Body Effects in Gallium Arsenide Quantum Wells in the Far-Infrared

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy
in
Physics
by
Keith Allen Craig

Committee in charge:

Professor Mark S. Sherwin, Chairperson
Professor S. James Allen
Professor Douglas J. Scalapino

October 1997

The dissertation of Keith Allen Craig is approved:

_____

_____

_____

Committee Chairperson

December 1997

December 12, 1997

Acknowledgments

When I first came to graduate school, I wanted to do high-energy particle theory, which is now evolving into string theory. However, I decided that, among other things, I did not want to sit in front of a computer (or even more imposing—a blank pad of paper) and spend all day every day creating the theoretical underpinning of physics.

I decided that I wanted to do a variety of things: use a computer sometimes, solder sometimes, design parts sometimes, machine sometimes, analyze data sometimes, read papers, write papers, hunt for screwdrivers etc... Only *some* of this would be done in front of a computer. I managed to get a research position with Prof. Mark Sherwin, when it was *very* hard to find any position. It was a miracle that I had asked everyone working in experimental condensed matter physics at UCSB for a job, although I really wanted to work with Mark, and everybody had said they had neither time nor funding for another student. Mark had said this before everyone else, but things changed unexpectedly.

Mark was always aware of how my research was developing, both in terms of "getting stuff done" and in terms of my scientific development. He was a continual source of support and guidance, while keeping in mind times it was better to let me guide myself. Thank you, Mark!

While working with Mark, I had the honor of working with many other wonderful people. First of all, Bryan Galdrikian, who was a great help in keeping my sense of humor, and was continually supportive during frustrating experimental moments (samples shattering into a million pieces, etc). Bryan made both the highs and the lows a little higher. Thanks and "Woof"!

Mike Crommie came down from Berkeley for a wonderful summer in which we worked on an analog simulation project. Mike is a lively and dynamic individual, and working with him was a huge and enjoyable lesson in how to do scientific research and work in a (small) group.

Chris Felix, Bill Bewley and John Plombon taught me everything they had learned about these experiments. Chris was my long-term mentor and introduced me to FIR measurements, as well as answering a huge number of questions. Thanks to all of you, with special thanks to Chris for all his time and patience!

James Heyman's post-doc stay at UCSB was an exciting time. Teaching him what I knew helped me clarify my own knowledge and its limits. Soon, James and I used that knowledge along with his to push our measurements and thinking forward at a rapid rate. We had a great deal of fun along the way.

During this time, Charley Unterrainer showed up from Austria. Working with James and I (as well as a seemingly countless number of projects of his own!), Charley brought new ideas and vast FIR expertise to our group. He was also a good friend, someone I could trust to be caring and yet brutally honest.

There were many graduate students who worked with me and supported and challenged me in different ways: Andrea Markelz, who, in addition to brightening the lab with her sense of humor, was always a great source of support and sanity. John Černe, I'll always remember our "Gershwin Night" in the FEL, and how much fun tedious setup tasks could be if you were with friends. Thanks to Naggi Asmar, Ester Yuh, and Paul Pinsukanjana, who always took the time to answer questions and solve problems together.

Of course, none of this could have been done without the FEL staff, whose dedication and accomplishments are made physical every time the FEL emits FIR pulses. Thank you all for your work on the FEL, and your general support and assistance around the lab. A special thanks to David Enyeart, for taking time to teach (again and again!) how to make the FEL do my bidding.

There would have been no measurements without the samples and an understanding of what they are and how they work; thanks to the folks who created them and shared their expertise: Art Gossard, Ken Campman, Pete Hopkins, and Kevin Maranowski. I am continually amazed by their skill and creativity in creating *and understanding* samples!

Thanks also to the faculty I've worked with: Jim Allen, Björn Birnir, and David Awshalom, who were not only brilliant and full of ideas, but also brought a sense of perspective and experience to the research.

And thanks to the next generation of researchers in our group: Jon Williams, Mark Su, and Carey Cates. It brightened the end of my graduate career to help introduce you to the world of research at UCSB, as well as helping to clarify many research issues for me.

Thanks to all my friends in Santa Barbara: Jill Katz, Justine Sutton, Jayne Selwa, Edwin Shaw, Alan Lash, Lonnie Lash, Tina Fear, and many others who were true friends. I'm so happy to have met you all.

Finally, thanks to Karin Fisher, who will become my wife in 1998. She has been a source of love, support, and sanity for my entire graduate career.

To my teachers, friends, family, and dogs.

Vita

November 23, 1965	Born: Hinsdale, Illinois.
June 1986 – Aug. 1986	Summer research intern at the National Center for Atmospheric Research, Boulder, Colorado.
Sep. 1983 – June 1987	B.A. in Physics, Carleton College, Northfield, Minnesota.
Sep. 1988 – Dec. 1989	Teaching Assistant, Dept. of Physics, University of California, Santa Barbara.
Jan. 1989 – June 1996	Research Assistant, Dept. of Physics, University of California, Santa Barbara.

PUBLICATIONS

"Effect of Random Noise on a Mode-Locked System," M.F. Crommie, K. Craig, M.S. Sherwin, and A. Zettl, Physical Review B, **43**, 13699 (1991).

"Resonant Harmonic Generation and Dynamic Screening in a Double Quantum Well," J.N. Heyman, K. Craig, B. Galdrikian, M.S. Sherwin, K. Campman, P.F. Hopkins, S. Fafard, and A.C. Gossard, Physical Review Letters, **72**, 2183 (1994).

"Probing Terahertz Electron Dynamics in Semiconductor Nanostructures with the UC Santa Barbara FELs," J.P. Kaminski, S.J. Allen, M.S. Sherwin, B. Keay, J.S. Scott, K. Craig, J.N.Heyman, P. Guimaraes, K. Campman, P.F. Hopkins, A.C. Gossard, D. Chow, M. Lui, and T.Y. Liu, Nuclear Instruments and Methods in Physics Research A, **341**, 169 (1994).

"Far-Infrared Saturation Spectroscopy of a Single Square Well," K. Craig, C.L. Felix, J.N. Heyman, A.G. Markelz, M.S. Sherwin, K. Campman, P.F. Hopkins, and A.C. Gossard, Semiconductor Science and Technology, **9**, 627 (1994).

"Probing Terahertz Dynamics in Semiconductor Nanostructures with UCSB Free-Electron Lasers," S.J Allen, K. Craig, C.L. Felix, P. Guimaraes, J.N. Heyman, J.P. Kaminski, B. Keay, J.S. Scott, M.S. Sherwin, K. Campman, P.F. Hopkins, A.C. Gossard, D. Chow, M. Lui, and T.Y. Liu, Journal of Luminescence, **60**, 250 (1994).

"Temperature and Intensity Dependence of Intersubband Relaxation Rates from Photovoltage and Absorption," J.N. Heyman, K. Unterrainer, K. Craig, B. Galdrikian, M.S. Sherwin, K. Campman, P.F. Hopkins, and A.C. Gossard, Physical Review Letters, **74**, 2682 (1995).

"Materials Science in the Far-IR with Electrostatic Based FELs," S.J. Allen, K. Craig, B. Galdrikian, J.P. Kaminski, J.S. Scott, M.S. Sherwin, K. Unterrainer, M. Wanke, K. Campman, P.F. Hopkins, A.C. Gossard, D.H. Chow, M. Lui, and T.K. Liu, Nuclear Instruments and Methods in Physics Research A, **358**, 536 (1995).

"Nonlinear Quantum Dynamics in Semiconductor Quantum Wells," M.S. Sherwin, K. Craig, B. Galdrikian, J.N. Heyman, A.G. Markelz, K. Campman, S. Fafard, P.F. Hopkins, and A.C. Gossard, Physica D, **83**, 229 (1995).

"Optical Rectification as a Probe of Quantum Dynamics in a Heterostructure," K. Unterrainer, J.N. Heyman, K. Craig, B. Galdrikian, M.S. Sherwin, K. Campman, P.F. Hopkins, and A.C. Gossard, Superlattices and Microstructures, **17**, 159 (1995).

"Undressing a Collective Intersubband Excitation in a Quantum Well," K. Craig, B. Galdrikian, J.N. Heyman, A.G. Markelz, J.B. Williams, M.S. Sherwin, K. Campman, P.F. Hopkins, and A.C. Gossard, Physical Review Letters, **76**, 2382 (1996).

"Far-Infrared Pump-Probe Measurements of the Intersubband Lifetime in an AlGaAs/GaAs Coupled-Quantum Well," J.N. Heyman, K. Unterrainer, K. Craig, J. Williams, M.S. Sherwin, K. Campman, P.F. Hopkins, A.C. Gossard, B.N. Murdin, and C.J.G.M. Langerak, Applied Physics Letters, **68**, 3019 (1996).

"Nonlinear Resonant Optical Rectification in a Coupled Quantum Well," K. Unterrainer, J.N. Heyman, K. Craig, B. Galdrikian, M.S. Sherwin, H. Drexler, K. Campman, P.F. Hopkins, and A.C. Gossard, Surface Science, **362**, 401 (1996).

FIELD OF STUDY

Major Field: Experimental Condensed-Matter Physics

Study of Linear and Nonlinear Intersubband Absorption in GaAs/AlGaAs Semiconductor Quantum Wells

Professor Mark Sherwin

Abstract

Experimental Investigation of Nonlinear Intersubband Absorption and Many-Body Effects in Gallium Arsenide Quantum Wells in the Far-Infrared

by

Keith Allen Craig

I have studied the nonlinear (and linear) intersubband absorption of wide GaAs/-AlGaAs quantum wells. The main sample was a 40 nm single square well, which was modulation-doped to roughly 1.4×10^{11} e/cm^2. I used the UCSB Free-Electron Laser to measure nonlinear absorption, with an edge-coupling technique to couple the radiation in and out of the sample and a gate to control the charge density in the well.

I measured the linear absorption of the sample as a function of charge density, which is impossible in a multiple well sample. It was discovered that the absorption strength of the sample did not vary linearly with the charge density, as was expected. A waveguide theory by Załużny was used to model the nonlinear absorption strength. The external electric field, an input parameter to the theory, was unknown. I was unable to obtain good agreement with experiment regardless of the external electric field used.

Using nonlinear absorption, I studied the dynamics of the depolarization shift, which is a many-body screening effect. In the square well, the depolarization shift shifts the intersubband absorption frequency from roughly 61 cm^{-1} to 81 cm^{-1}. This shift is proportional to the difference in population between the subbands. When the intensity of the incident radiation was large, absorption of the radiation populated the upper subband. As the upper subband was populated, the stimulated emission from the upper subband reduced the net amount of absorption, as is seen in standard saturation measurements. Due to the depolarization shift, the peak also shifted down in frequency as the upper subband population increased. Since the peak shifted in frequency while the sample was illuminated by the laser, the absorption of the laser changed in time, which in turn changed the populations. The rich absorption behavior which resulted was well modelled by theory.

I measured the energy relaxation time for electrons in the upper subband (which is below the LO phonon energy) to vary from 0.7 ns to 50 ps as a function of intensity. This wide variation was explained in terms of LO phonon emission, which depends upon the average electron energy.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The general goal of this work was to investigate the nonlinear intersubband absorption of wide quantum wells in the far-infrared (FIR), with the hope of gaining new theoretical as well as practical understanding.

I had two main theoretical motivations; to investigate and better understand the many-body effects in the wells, and to understand the relaxation mechanisms in our samples.

There were several applied motivations for this work; to gain experience and understanding of FIR GaAs devices, to measure the actual relaxation times in wide wells (below the LO phonon), and to investigate the possibilities of novel optical devices constructed from quantum wells operating at FIR frequencies.

1.1 Theoretical Motivations

The many-body effects in quantum wells are not well understood theoretically, and have not been thoroughly investigated experimentally. The most noticeable (although not the only) many-body effect at FIR frequencies is the depolarization effect, also

called the depolarization shift. The depolarization shift is a dynamical screening effect: the motion of the charged electrons changes the electric fields inside the well in such a way that the absorption frequency for radiation is shifted above the intersubband spacing. (The depolarization shift is discussed more completely in Chapter 2.) In wide quantum wells, the depolarization shift is more easily observed and investigated than in narrower quantum wells. This is because, as the well width is increased, the separation of the subbands decreases, while the interaction energy (and the depolarization shift) of the electrons remains roughly constant. Thus the depolarization shift is a larger fraction of the absorption frequency in wide wells than in narrower wells.

For example, in a 40 nm square quantum well (such as the one studied in this work), the separation of the subbands is approximately 61 cm^{-1}, while the intersubband resonance is observed and predicted at approximately 81 cm^{-1}. Thus, the many-body effects shift the absorption by 20 cm^{-1}, or almost 33% of the subband spacing. In contrast, the many-body effects can be ignored in a 5 nm quantum well, because the subband spacing is so much larger.[1]

Since most experimental investigations in the past have been done on narrow quantum wells, many-body effects such as the depolarization shift were largely ignored because they were so small as to be unobservable. In addition, many researchers may have quite reasonably wanted to avoid the added complexity. However, many-body effects are quite interesting from the viewpoint of fundamental condensed mat-

[1]The many-body effects enter the intersubband resonance frequency as $\widetilde{\omega}_{12}^2 = \omega_{12}^2 + \omega_p^2$, where $\widetilde{\omega}_{12}$ is the intersubband resonance frequency, ω_{12} is the intersubband spacing, and ω_p is the effective plasma frequency due to the many-body effects. Thus, as the subband spacing increases, the many-body effects become unobservable much faster than linear.

ter theory, and wide quantum wells are a excellent system to examine them. (Quantum wells allow much greater control over parameters such as charge density and confining potential than most other condensed matter systems.) The experimental information about the depolarization shift can be compared with theories of many-body interactions. In this work, I have compared my results with a theory of nonlinear intersubband absorption (see Reference [29]) and found generally good agreement. These results are discussed in Chapter 6.

The other theoretical motivation was to understand the mechanisms by which the electrons lose energy. In short, there are several different ways electrons can lose energy in the well, and by measuring the energy relaxation rates, it is possible to draw some conclusions about the relaxation mechanisms. This information is useful in terms of understanding the interactions of the electrons with the fundamental excitations of the solid. My measurements strongly suggest that both acoustic phonon emission and LO phonon emission are important in our samples, and that their relative importance changes as a function of average electron energy. These processes are introduced and discussed in Chapter 7.

1.2 Applied Motivations

GaAs quantum wells have the possibility of being very efficient optical devices. The control over the confining potential allows these devices to operate at FIR frequencies. This is very interesting, because few "good"[2] devices currently exist for creation, modulation, or detection of FIR radiation. Thus, the applied motivations can

[2]By "good" here, I mean that the devices are small, cheap, robust, efficient and relatively easy to use.

all be viewed as taking the first steps toward creating efficient (and even novel!) devices which operate in the FIR. Such devices could be very useful for environmental monitoring, short-range collision-avoidance systems, and scientific research, to name a few applications.

The first step toward the goal of new FIR devices is simply to start working with such devices and learn all that we can about them. This involves learning how to create and construct them, how to couple radiation into them and out of them, how to handle them to not damage them, and their general behavior and possibilities. Our success in creating samples with the desired absorption properties and measuring their absorption demonstrates our achievements in learning to work with these samples.

Another step toward making practical devices is to measure some of their fundamental properties, such as the energy relaxation time for electrons below the LO phonon. (See Chapter 7 for an introduction to energy relaxation times in GaAs.) The relaxation time for electrons below the LO phonon energy has been measured several times in the past, by different techniques. These measurements yielded values in the range from a few ps to half a nanosecond. The reason for this wide range of values was unknown. It was unclear if some of the measurements were flawed, or if some unrecognized parameter was causing the relaxation time to vary. A large amount of heated discussion occurred about which measurements were correct and why, but no clear consensus had emerged when these measurements were started. My measurements were one of the first to suggest that the average energy of the electrons was a crucial factor in determining the energy relaxation time of the electrons. This simple insight allows most of the huge range of measured relaxation times to be understood within a single picture: the differing experimental techniques measured samples with

widely differing electron energies, and thus yielded widely differing relaxation times. As mentioned above, the relaxation times are discussed in Chapter 7.

The last practical motivation was to discover new effects which could be used to make new devices which had not been possible before. I discovered, based on the theory of Reference [29] and my observations, that it is possible to make a mid-range saturable absorber; a device which absorbs very little at high and low intensities, but which absorbs strongly at moderate intensities. Such a device is based on the dynamics of the depolarization shift; as far as I know, no other known process could produce such a device.

1.3 GaAs Modulation-Doped Quantum Wells

The quantum wells studied here are composed of different semiconductor layers which confine electrons in a defined potential. The wells are modulation-doped, meaning that the dopants are placed outside the well, so that the ionized donors do not scatter the charged electrons, since the donors are spatially separated from the electrons. See Chapter 2 for more discussion of the samples.

1.4 Intersubband Absorption

The absorption feature of interest in our samples is the intersubband absorption from the lowest subband (denoted 1) to the first excited subband (denoted 2). In the single electron approximation, the intersubband transition can be thought of like an atomic transition: a photon is absorbed, and an electron from the lower state is promoted to the upper state. The subbands have a parabolic dispersion as a function of in-plane

wavevector:

$$E_n(k) = E_n + \frac{\hbar^2 k^2}{2m^*} \tag{1.1}$$

where $E_n(k)$ is the energy of an electron in the nth subband with wavevector magnitude k, E_n is the energy at the bottom of the nth subband, and m^* is the effective mass of the electrons in the conduction band of GaAs. The lower two subbands as a function of in-plane wavevector is shown in Figure 1.1. Since the momentum of the electrons in the plane of the well is not changed by absorbing a photon, the absorbing electron moves vertically on a plot of energy vs wavevector. Since the subbands are spaced an equal amount of energy apart vertically on this plot, the intersubband transition is not broadened by a variation in the separation of the subbands with wavevector.

The depolarization shift causes the absorption to occur at a higher frequency than the separation of the subbands. This absorption feature is sometimes called the collective excitation, to distinguish it from a single-electron transition which would occur at the subband spacing. The depolarization-shifted frequency is explained more fully in Chapter 2.

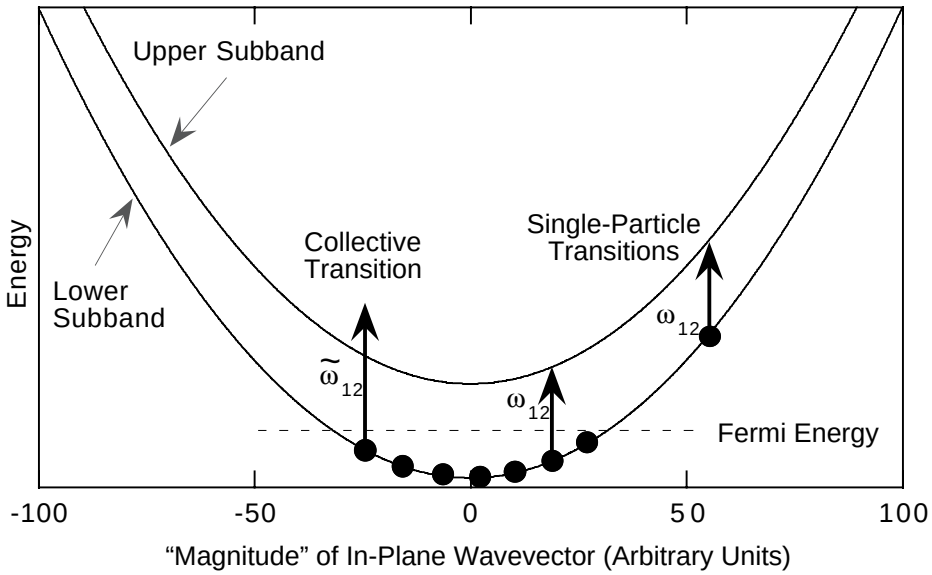

Figure 1.1: The energy as a function of in-plane wavevector for electrons in the lowest two subbands of a quantum well. The subbands are parabolic, following Equation (1.1). The large dots represent electrons. The vertical arrows represent intersubband transitions; the two on the right are the single-particle transitions with frequency ω_{12}, which are not observable using FIR absorption. Note that all these transitions have the same energy. Also shown is the collective excitation with frequency $\tilde{\omega}_{12}$, which is at a higher energy than the single-particle mode. The single-particle picture clearly does not do an adequate job of explaining the collective mode. (The extra energy of the absorbed photon which creates a collective excitation goes into internal excitations of the electron gas.)

Chapter 2

Samples

A great deal of work went into creating samples which would be straightforward to measure and understand. The samples had to be carefully designed to possess the desired charge density and resonance frequencies; self-consistent computer programs were extensively used to simulate possible structures. Once the samples were grown, they were processed to create electrical contacts and metallizations for optical purposes. After processing, and during measurements, the charge density in the well was measured using capacitance-voltage profiling. In addition, care was taken in how the sample was cooled.

Most of the measurements described in this thesis was performed on a 40 nm wide modulation-doped square quantum well, hereafter referred to as the "square well"[1] or the "single square well." For comparison, measurements of a sample consisting of a coupled double square well are also discussed; this sample will be referred to as the "double square well."[2] The double square well was composed of 8.5 nm and

[1]The working title of the square well sample was "KC2."

[2]The working title of the double square well sample was "DSQC."

7.3 nm wells separated by a 2.5 nm barrier. (The center barrier and the outer barriers were $Ga_{0.7}Al_{0.3}As$, and the wells were pure GaAs.) For a full description of the design of the double square well, see Reference [10].

2.1 Sample Design

The samples were grown using molecular beam epitaxy (MBE). The design specifies the aluminum mole fraction as a function of position, and the location and amount of silicon doping. For a description of molecular beam epitaxy and a general background on semiconductor quantum wells, see Reference [25].

2.1.1 Desired Charge Density

With the square well, I wanted to study the dynamics of the depolarization shift by measuring the nonlinear intersubband absorption. To simplify the measurements and analysis, I wanted to dope the well so that only the first subband contained charge. (I used a self-consistent program to calculate the charge density where the second subband began to fill.) Control over the amount of charge in the well was obtained by spatially separating the donors from the well, and then heavily doping the donor regions. See Figure 2.1 for the design, and Figure 2.2 for the self-consistent band diagram. (There was symmetric doping on both sides of the well.) The wavefunctions are shown in Figures 2.3 and 2.4.

A simple mechanism controls how much charge transfers into the well in this sample. At room temperature the electrons are free to move onto and off of the donors; as electrons transfer from the donors into the quantum well, the potential (for electrons) at the well increases as the well becomes more negatively charged. Charge

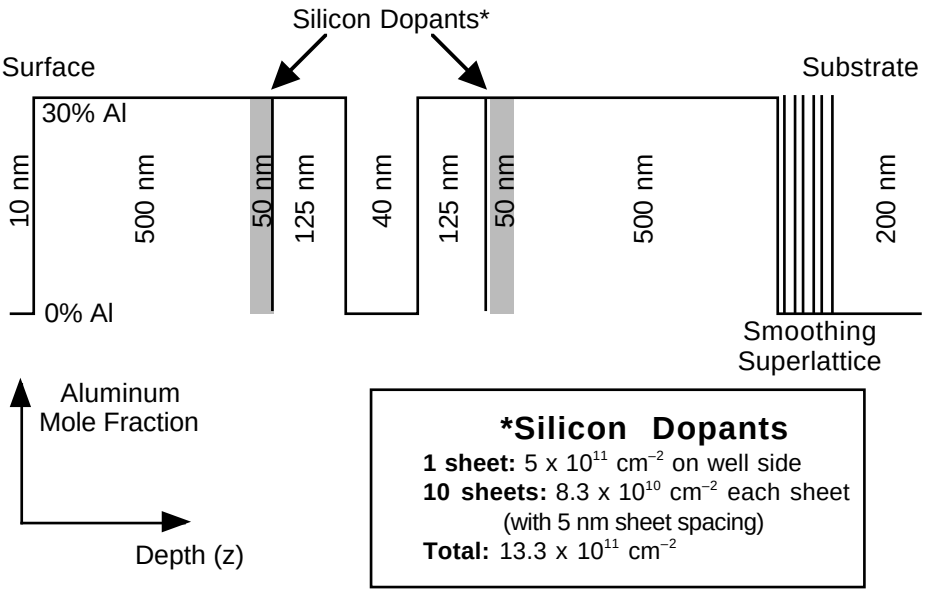

Figure 2.1: The growth design for the square well. It is a symmetrically doped 40 nm square well.

transfer stops when the chemical potential of the well moves above the chemical potential of the donors. As long as there is excess charge remaining at the donors after the well is full (when warm), the only parameter which determines the charge in the well is the distance from the doping to the well. (In the square well sample, this was 125 nm.) As the sample is cooled, the amount of charge in the well remains constant, but the electrons at the donors become trapped by deep traps. Using this method, the amount of charge in the well could be carefully controlled.

This doping method relies upon the intricacies of the deep traps created by the silicon donors (DX centers). For a general review of the field of DX centers, see Reference [17]. The method leaves the sample in an ideal state: the charge on the donors is “frozen” to the donors, and does not affect the ability to deplete the well using the gate. Since the electrons in the well are spatially separated from the silicon

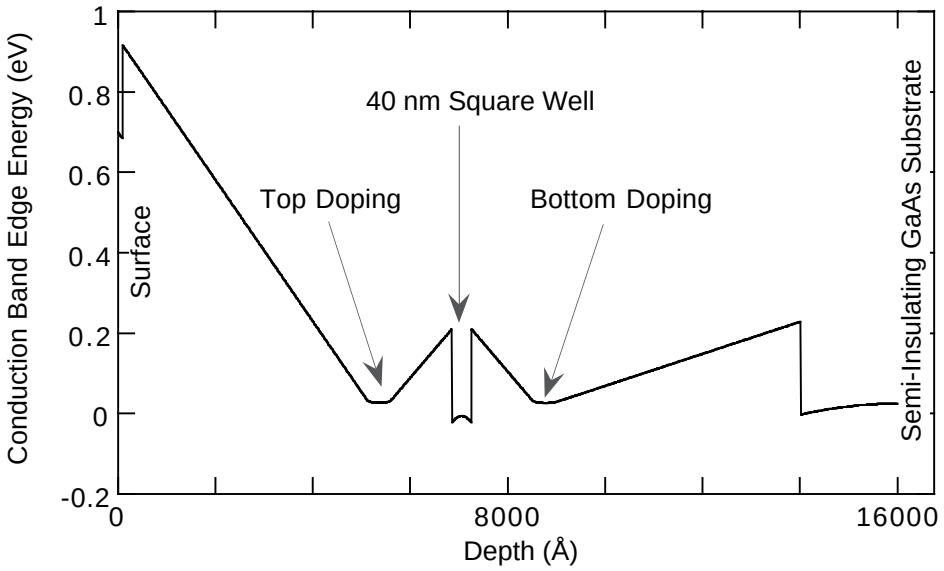

Figure 2.2: The calculated conduction band shape for the square well. The vertical axis is potential energy, and the horizontal axis is depth beneath the surface of the sample. The results for this figure were calculated using a numerical simulation written by Greg Snider; for his address, see Reference [23].

(and therefore separated from the deep traps), they cannot freeze out. However, some of the measurements indicate that some of the charge on the donors may not have been completely frozen out at low temperature. (See, for example, Figure J.1 and the accompanying discussion.) It may be that not all the electrons were frozen onto true DX centers, or it may be that the DX centers in our sample were metastable with a long (but not infinite) time constant.

2.1.2 Built-In Electric Field

A sample parameter which is of crucial importance in determining the intersubband absorption frequency is the built-in electric field across the well in the growth direction, hereafter referred to as $E_{\mathrm{built-in}}$. The doping method used should theoretically

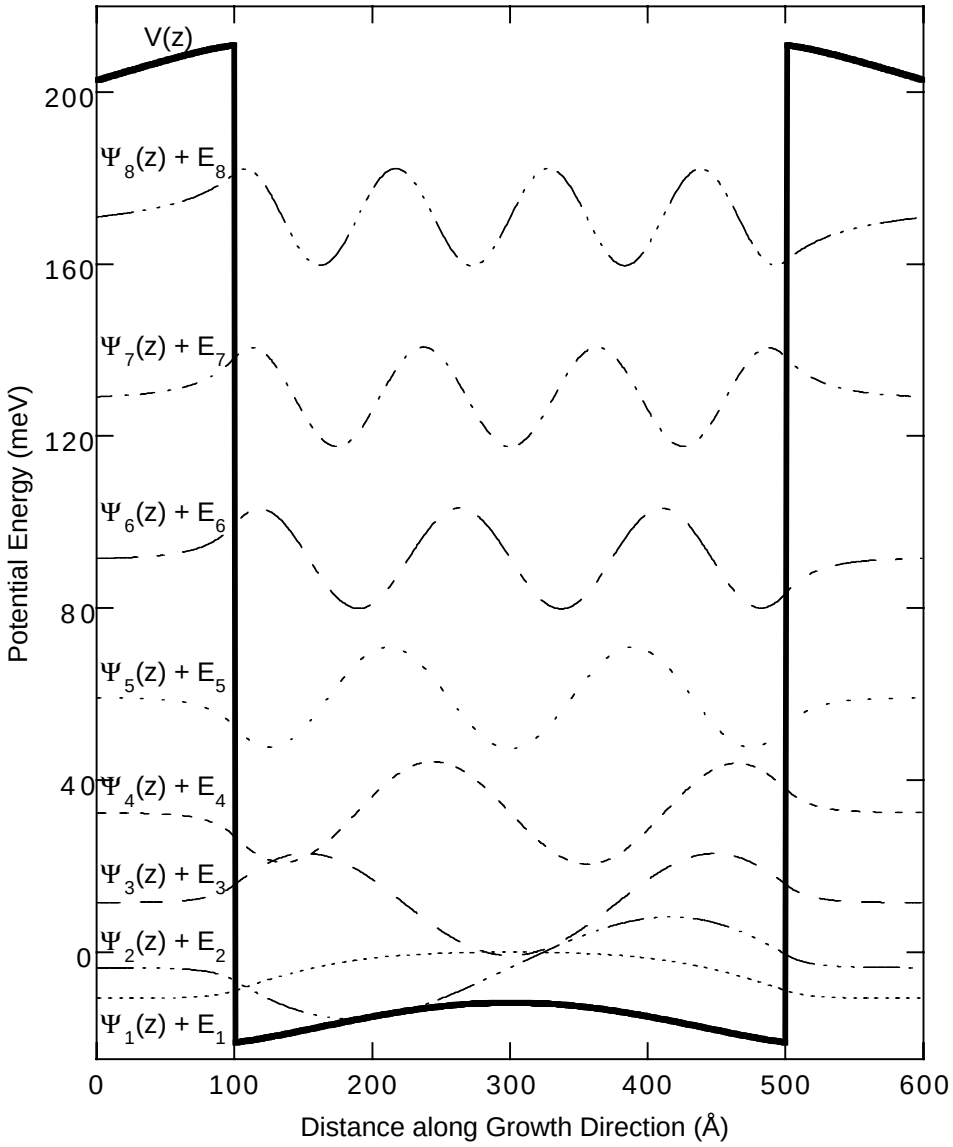

Figure 2.3: The self-consistent well potential and all the wavefunctions in the square well, as calculated by the self-consistent program by Bryan Galdrikian. The wavefunctions are offset by their energy on the potential plot. The vertical axis is potential energy for the well and the wavefunction offsets, and the horizontal axis is distance along the growth direction. Other than the offsets, the wavefunctions are plotted on an arbitrary linear scale. The static electric field across the well is zero.

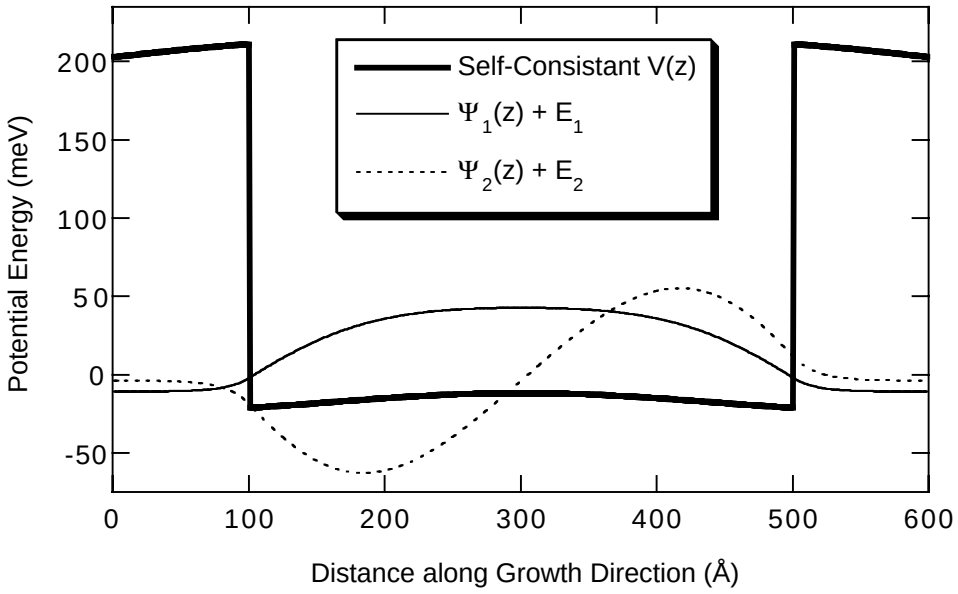

Figure 2.4: The self-consistent well potential and the lowest two wavefunctions in the square well, as calculated by the self-consistent program by Bryan Galdrikian. The wavefunctions are offset by their energy on the potential plot. The vertical axis is potential energy for the well and the wavefunction offsets, and the horizontal axis is distance along the growth direction. Other than the offsets, the wavefunctions are plotted on an arbitrary linear scale. The static electric field across the well is zero.

leave zero built-in electric field in the region of the well at zero volts gate bias. However, the experimental measurements indicated that this was not what happened.

I expected $E_{\mathrm{built-in}} = 0$ because the donors on both sides of the well should have their chemical potential pinned at the ionization energy of the donors: when warm, the heavily-doped regions should act as an (almost) unlimited source of electrons at the ionization energy of the donors. The electrons leave the donors and distribute themselves self-consistently to raise the chemical potential everywhere to at least the ionization energy of the donors. Once this equilibrium is reached, the donor's potential energy should be at the ionization energy of the donors, as long as there are

still electrons at the donors. At this point the electrons will stop moving. Then, as the sample is cooled, the electrons which are still on the donors are trapped in DX centers, and are no longer free to move. Therefore the potential at the donors should not be changed by cooling. This *should* result in the donors being pinned at the same potential energy, leaving no possibility of a static field across the well without the application of a gate bias.

Experimental measurements left this conclusion in doubt, as the best fit to the data was found by assuming $E_{\mathrm{built-in}}$ was not zero. (See Appendix B for more details about the built-in electric field and its relation to the gate voltage.) This could be explained if the long time constants of the deep traps hamper the free motion of the electrons to equalize the chemical potential when the sample is warm. Then the system could be stuck out of equilibrium when warm, and there *could* be a static electric field across the well. Upon cooling, this static field would be frozen in. This could explain the measurements indicating that the built-in field across the well was not zero.

2.1.3 Resonance Frequency

We designed the square well so that the resonance would be in a convenient frequency range for the U.C. Santa Barbara Free-Electron Laser (FEL). This meant that the resonance should be between approximately 60 cm^{-1} and 110 cm^{-1}; in this range the FEL is intense and relatively easy to tune and operate. The main factors determining the resonance frequency are the width of the well, the built-in electric field at the well, and the charge density in the well. Using self-consistent simulations and the doping scheme described in the last section, we were able to design a sample with the resonance in the region around 80 cm^{-1}, which was excellent for these

measurements.

2.2 Depolarization Shift

The depolarization shift is a dynamic screening effect [1]. One of the main goals of this work was to study the depolarization shift under nonlinear conditions where the changing populations of the subbands cause changes to the depolarization shift, and therefore the absorption frequency. A full derivation of the depolarization shift can be found in Reference [2].

Consider a single electron in a quantum well. It will have an intersubband conductivity peaked at the separation of the subbands. This frequency-dependent conductivity can then be used to calculate a dielectric function for a gas of identical electrons. The absorption computed from this dielectric function occurs at a *higher frequency* than the original single electron. The other charged electrons can be thought to have created an additional coulombic restoring force which increases the oscillation frequency; this can be thought of as "stiffening the spring" of a classical oscillator.

An analysis of the depolarization shift with electrons in both subbands shows that the shift depends on the difference in charge density between the two subbands involved in the resonance [30]. Thus, the frequency shift is a maximum when all the electrons are in the lower subband, and the shift vanishes if the populations of the two subbands are equal. The depolarization shift can be written as:

$$\widetilde{\omega}_{12} = \omega_{12}\sqrt{1 + \alpha\frac{(N_1 - N_2)}{N_s}} \tag{2.1}$$

where $\widetilde{\omega}_{12}$ is the depolarization-shifted absorption frequency, ω_{12} is the intersubband spacing frequency, α is the depolarization shift parameter (see below), N_1 (N_2) is

the population density in the first (second) subband, and N_s is the total population density in the well. The depolarization shift parameter α is given by

$$\alpha = \frac{8\pi N_s e^2}{\epsilon E_{12}} \int_{-\infty}^{+\infty} dz \left[\int_{-\infty}^{z} dz' \, \chi_2(z') \chi_1(z') \right]^2 \tag{2.2}$$

where e is the charge of the electron, ϵ is the dielectric constant of the bulk semiconductor, E_{12} is the intersubband spacing energy, and χ_i is the i-th envelope wavefunction in the growth direction.

Since the depolarization shift depends upon population, and the depolarization shift influences the absorption frequency, the absorption frequency is a function of population in these samples. This intriguing concept is developed more fully in Chapter 6.

A common misconception of the depolarization shift is that the many-body effects shift the subbands, so that the subbands are actually separated by $\hbar\widetilde{\omega}_{12}$. This is not true! The subbands are separated by $\hbar\omega_{12}$, which includes the effect of the distortions to the potential due to the electrons. These quantities have different meanings: $\hbar\omega_{12}$ is the energy separation of the subbands, and $\hbar\widetilde{\omega}_{12}$ is the energy of photons which are absorbed by the intersubband transition. It is surprising at first that the intersubband spacing is different from the resonance frequency. However, this is correct; the extra energy goes into internal excitations of the electron gas (essentially a plasmon). The state created when the well (with all electrons in the ground subband) absorbs one photon is *not* simply a state with one electron in the excited subband. It is a more complicated collective excitation, which can be thought of *only approximately* as one electron in the excited state.

The depolarization shift is the dynamical part of the coulombic repulsion of the electrons, which is called the Hartree term in the static case. In the static case, the

Hartree term overestimates the interaction energy, because it assumes the electrons overlap each other fully. The physical system obeys the Pauli exclusion principle, which greatly reduces the spatial overlap of the electrons. The Pauli exclusion principle may be included in a theory by using an exchange and correlation potential, which includes the effects of the electrons avoiding each other, and removes the extra interaction energy from the simple Hartree calculation.

The same problem occurs in the dynamic case. The dynamical part of the Hartree term, the depolarization shift, creates too large a frequency shift because it does not take into account the dynamics of the exchange-correlation potential. The dynamical part of the exchange and correlation potential is called the excitonic correction to the depolarization shift, and can be thought of physically as the lowering of energy of the final state due to the attraction of the excited electron in the upper subband and the hole in the lower subband where the electron came from. Our static self-consistent simulations include both the depolarization shift and the excitonic shift, but the dynamical theory of Załużny does not include the excitonic shift.

2.3 Self-Consistent Simulations

The self-consistent program was used not only to design the sample, but also to fit the experimental linear absorption data. These data were modeled using a self-consistent computer program written by Bryan Galdrikian. (For a discussion of the method used, see Reference [7, Appendix B].) This program calculates the self-consistent subband energies in the self-consistent potential. It also finds the dynamical many-body corrections (such as the depolarization shift) and the dipole matrix element $\langle z_{12} \rangle$. The approximation to the exchange-correlation potential of Vosko, Wilk, and

Nusair was used in these calculations. (For details of this approximation, see Reference [4].) Accurate estimates of the bare subband spacing E_{12} and the dipole matrix element $\langle z_{12} \rangle$ were essential for use in fitting my linear and nonlinear absorption data to theory. The output of the self-consistent program was the main source of these parameters. (Note that, although the self-consistent program also calculated the resonance frequency $\widetilde{E}_{12}$, it usually made more sense to use the experimentally measured values for $\widetilde{E}_{12}$. However, experimental measurements of the dipole matrix element led to physically impossible results (see Chapter 4), and I did not attempt experimental measurement of the subband spacing E_{12}.)

The simulations were performed using as many of the experimentally measured parameters as inputs to the self-consistent program as possible. (See Table 2.1 for a concise listing of the parameters.) The width of the well was varied within a narrow range ($\pm 4\%$) around the design-width of the well, in order to find the best fit to the data. (The sample growers believe their growth calibrations to be accurate to approximately $\pm 4\%$.) The charge density was calculated directly from capacitance-voltage measurements. Surprisingly, the single most important parameter in determining the absorption frequency in the self-consistent simulations was the static electric field across the potential, E_{dc}, which is the sum of the built-in static field ($E_{\mathrm{built-in}}$) discussed above and the static field due to the charges on the gate. (The total static field across the well is sometimes called the tilt, because it causes the potential to look "tilted.") The effects of the static electric field dominated the optical results (especially the resonance frequency), in part because our structures are relatively wide, so that any field across the structure causes a large potential-energy difference from one side of the well to the other.

I have not yet found a direct method to experimentally measure the static electric

Parameter	**Symbol**	**Source**
Well Width	w	Growth Program $\pm$ 4%
Charge Density	N_s	Capacitance Measurements
Static Field across Well	$E_{\rm dc}$	Free Parameter

Table 2.1: The input parameters for the self-consistent program and how they were determined. See text for discussion.

field at the well. This complicated the data analysis because the built-in field seemed to change after each cooldown. (This is discussed in Appendix B.) Therefore, I used the built-in field as a part of the free input parameter for the static field in the self-consistent program.

The self-consistent modeling procedure was as follows: the experimental parameters were input into the self-consistent program, and some arbitrary built-in field was used. Then the results (usually the intersubband resonance energy $\widetilde{E}_{12}$) were compared with experimental data. Depending upon how well they fit, the built-in field (and sometimes the well width) were changed, and the self-consistent program was run again. This cycle continued until an acceptable fit was found. (Note that the waveguide mode modelling used a much more complicated procedure, which is explained in Chapter 4.)

2.4 Processing

The details of the sample processing are described in Appendix A. The sample was approximately 1.0 cm long and 8 mm wide when the processing was complete. On the well-side of the sample, there were four ohmic contacts in the corners which

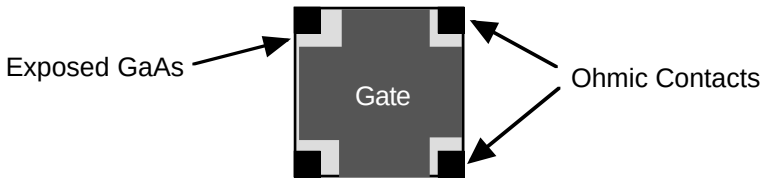

Figure 2.5: The sample after the processing was complete. The dark squares on the corners are ohmic contacts to the well, and the gate in the middle of the sample is evaporated aluminum.

contacted the well, and most of the rest of the top was covered with an aluminum gate. The back-side of the sample was completely covered with aluminum. The gate on the top is used to confine the FIR radiation within the sample and to deplete the sample by applying a voltage to the gate relative to the ohmic contacts. The sample is shown schematically in Figure 2.5.

In addition, the sample used for nonlinear measurements had an anti-reflection coating on the edge of the sample where the FIR radiation exited the sample. This coating greatly reduced the interference fringes due to multiple reflections within the sample. The creation of this coating is described in detail in Appendix A. The transmission of a test piece of GaAs with and without the anti-reflection coating is shown in Figure 2.6.

2.5 Capacitance Measurements

The charge density in the well was measured using capacitance measurements. I used capacitance-voltage (or $C(V)$) to measure the total amount of charge in the well at a given gate voltage, once the sample was cold and ready for optical measurements. I also used it to monitor any changes in the charge density during the measurements.

The general idea of the capacitance measurement is that the measured capaci-

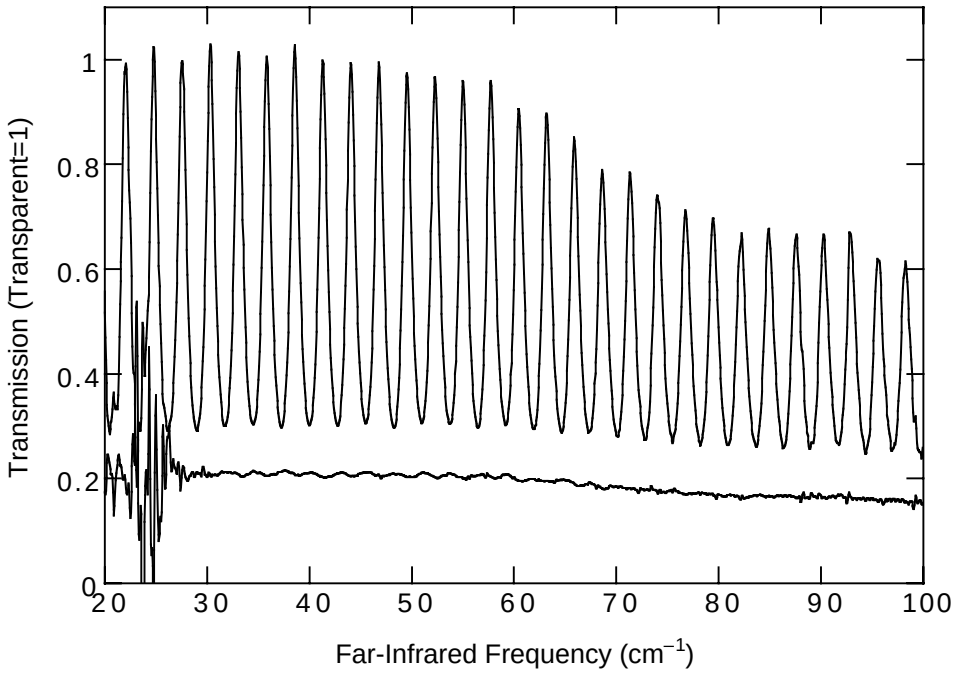

Figure 2.6: Experimental FT-IR data of fringes in the transmission of a piece of GaAs at normal incidence. The upper curve is semi-insulating GaAs, and the lower curve is semi-insulating GaAs with an impedance-matching metallic anti-reflection coating. The coated piece has almost no fringes, although the total transmission is lower. (Note that the period of these fringes is quite wide due to the normal incidence measurement technique, in which the sample approximately 0.05 cm thick. The fringes are spaced much more closely when passing through 1.0 cm of GaAs.)

tance is just due to the separation of the gate and the conducting electron sheet at the well. Clearly this value also depends upon the area of the sheets and the dielectric constant of GaAs. As the gate voltage is changed by dV, an amount of charge dQ is removed from the well according to the relationship $dQ = C(V)\,dV$. Thus, by integrating the measured capacitance as a function of voltage from depletion up to a given gate voltage, the total amount of charge in the sample at that gate voltage can be determined:

$$Q(V_0) = \int_{V_G=V_{\mathrm{depleted}}}^{V_G=V_0} C(V_G)\,dV_G \tag{2.3}$$

where $Q(V_0)$ is the total charge in the sample at a gate voltage V_0, V_{depleted} is the voltage when the sample is depleted, and $C(V_G)$ is the measured capacitance as a function of gate voltage V_G. A small amount of uncertainty was introduced into the capacitance results due to uncertainties in the area of the gate, and the amount of residual charge in the depleted sample. The technique is discussed in more detail in Chris Felix's thesis, Reference [5], on pages 37–41.

2.6 Cooling Procedures

The samples studied here have very long time constants for the charge density to come to equilibrium after being exposed to visible light. Visible light can ionize donors and cause charge to spatially transfer large distances through the sample. In addition, the DX-centers created by the silicon doping tend to exhibit rich behavior including metastablity. To simplify the measurements, care was taken to keep the sample in the dark for as long as possible before measurements. (The samples were kept in a flow cryostat with black poly windows which completely blocked visible light for at least a few hours, and usually several weeks, before measurements.)

The details of the cooling procedure was also suspected of having an influence on the final state of the cold sample, so an effort was made to repeat the same cooling procedure for each measurement. When measuring the linear absorption using FT-IR, the sample was cooled to 77 K by liquid nitrogen for roughly one hour, then cooled to 4.2 K with liquid helium, and then slowly lowered to 1.5 K by pumping on the helium bath. In the nonlinear absorption measurements using the FEL, the sample was cooled over a period of approximately one hour using a low rate of helium flow in the continuous flow cryostat.

Chapter 3

Linear Spectroscopy

Linear spectroscopy was essential for understanding the basic properties of the samples and for planning the experimental aspects of the nonlinear (high intensity) experiments. (By linear spectroscopy, I mean spectroscopy with a low enough incident intensity of radiation that the optical properties of the sample are independent of the incident intensity.) The ability to measure absorption quickly and accurately over a very wide range of frequencies is what made linear spectroscopy so useful.

I used a Bomem Fourier-Transform Infrared (FT-IR) spectrometer with a weak broad-band source to measure the linear absorption of our samples. I performed all the optical measurements (both linear and nonlinear) in the waveguide edge-coupling geometry, in which the FIR is sent into the sample through one edge, and it propagates through the sample and passes out the other edge. (It is a transmission measurement.) I constructed a special-purpose cryostat which contained both the detector and the sample, as well as various optics. The cryostat was connected to the Fourier-Transform Infrared Spectrometer (FT-IR), so that the FIR from the spectrometer entered the cryostat, passed through the sample, and the transmitted FIR was measured

by the detector in the same cryostat. I measured the charge density in the sample using capacitance measurements, and I varied the charge density in the well using the gate voltage. By applying a large negative gate voltage, almost all the electrons were removed from the sample, creating an optical reference. (See Appendix E for a discussion of the residual charge in the "depleted" well.[1])

I observed very strong intersubband absorption from our samples, and observed the absorption to shift in frequency by a relatively large amount as a function of gate voltage. The frequency shifts were mainly caused by changes to the static electric field across the potential well. Most of the observed absorption lines were relatively narrow.

I was not able to fully understand some parts of my observations: the integrated absorption strength in the 40 nm square well sample was not linearly related to the amount of charge in the sample, and there were some asymmetries with the absorption lineshape and the baselines.

3.1 Bomem Fourier-Transform Infrared Spectrometer

I used a Bomem DA3.002 Fourier-Transform Infrared Spectrometer (FT-IR) to measure linear absorption. The spectrometer consists of a broad-band source, an interferometer, a sample chamber, and the necessary electronics to digitize and Fourier-Transform the signal from the detector as a function of time.

The spectrometer source (a mercury-xenon lamp) is essentially a black body;

[1] By "depleted", I really mean quasi-depleted, as discussed in Appendix E. However, henceforth I will refer to the sample as being depleted when a large enough negative gate bias has been applied that the charge density stops changing.

it emits broad-band radiation which extends weakly into the far-infrared. Due to the very dim source, I worked hard to ensure that I would be able to do accurate measurements using what little radiation I had. The interferometer and the rest of the system are unremarkable; see an introductory book on FT-IR for the details (for example, Chamberlain, Reference [3]).

3.2 Waveguide Edge-Coupling Geometry

I coupled FIR into the sample using a waveguide coupling geometry, shown schematically in Figure 3.1. The GaAs of the sample (essentially just a dielectric slab) forms the waveguide, and the FIR is confined inside the GaAs dielectric by thin layers of aluminum on the top and bottom of the sample. (The layer of aluminum on the top is the gate; the layer on the bottom is only for optical confinement.) FIR enters through the edge of the sample and propagates through the GaAs between the aluminum layers, with electric field parallel to the growth direction. As the FIR propagates, it is absorbed by the intersubband resonance of the quantum well. The radiation which is not absorbed passes through the sample and exits through the opposite edge of the sample, and I measure this transmission.

Assuming that the incident FIR illuminates the edge of the sample uniformly, the simplest way to think about the waveguide absorption is to imagine a single waveguide mode propagating down the waveguide. The intersubband transition of the quantum well will absorb energy from the frequencies of the mode which are resonant with the absorption as the mode propagates. Ignoring the quantum well, and assuming the metal on both sides of the dielectric to be infinitely conductive, the lowest mode in the waveguide is the TEM mode, which has a uniform oscillating

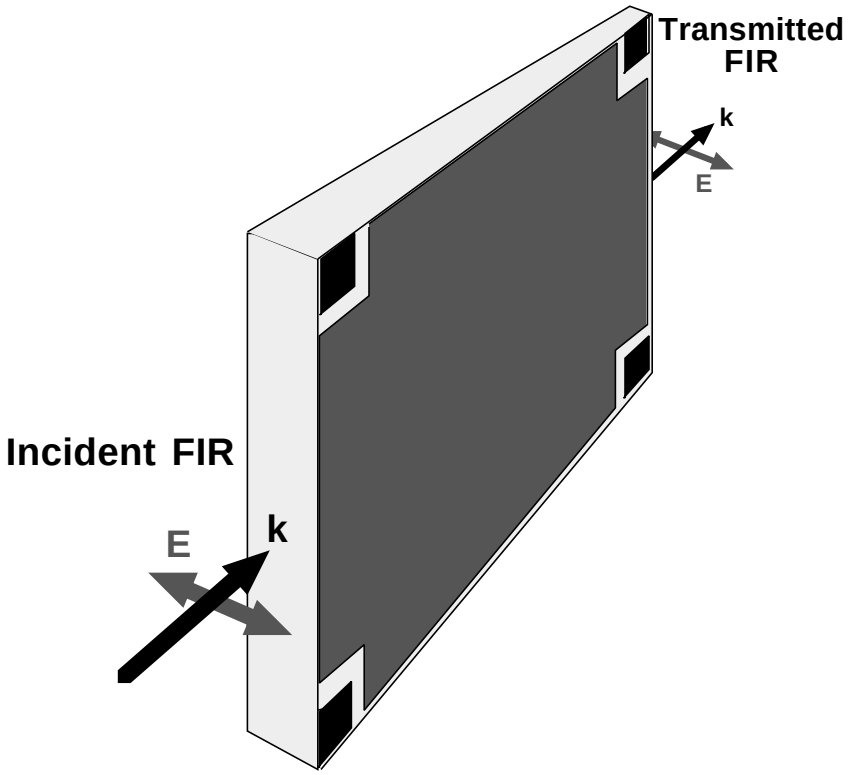

Figure 3.1: Schematic diagram of the edge-coupling technique. The FIR is sent into one edge of the sample, propagates through the waveguide formed by the sample and the aluminum layers on the top and bottom, and the transmitted radiation emerges from the other edge. The intersubband transition can only absorb the electric field which is polarized in the growth direction.

electric field which is perpendicular to the aluminum walls [11, Chapter 8]. (This means the electric field is in the growth direction of the quantum well, so that the oscillating field can couple strongly to the intersubband transition.) This mode is independent of frequency. The linear absorption in this waveguide is

$$\alpha(\omega) = \frac{8\pi e^2 N_s f_{12} \hbar \Gamma}{cam\sqrt{\epsilon}} \left(\frac{(\hbar\omega)^2}{(\widetilde{E}_{12}^2 - (\hbar\omega)^2)^2 + 4(\hbar\omega)^2\Gamma^2} \right), \tag{3.1}$$

where $\alpha(\omega)$ is the frequency-dependent attenuation coefficient (all quantities are in cgs units), e is the electronic charge, N_s is the electronic sheet density in the quantum well in units of cm^{-2}, $f_{12} = 2m\hbar^{-2}E_{12}|\langle z_{12}\rangle|^2$ is the oscillator strength for the transition from the lowest state to the first excited state (where E_{12} is the *bare* intersubband energy spacing, and $\langle z_{12}\rangle$ is the expectation value of the dipole matrix element), Γ is the energy of the half-width at half-maximum of the peak, c is the speed of light, a is the width of the waveguide (approximately 0.045 cm in our samples), m is the effective mass of electrons in GaAs, ϵ is the dielectric constant of the GaAs (approximately 13), ω is the frequency of the radiation, and $\widetilde{E}_{12}$ is the depolarization-shifted intersubband resonance energy. (See Appendix C for a full discussion of this formula.) For the case where the absorption of the quantum well is so strong that it changes the form of the waveguide modes, see Chapter 4.

The measurements were performed inside a single cryostat containing a Winston cone, sample, optics, and detector. (See Figure 3.2.) The details of this setup and its operation are described in Appendix D. The Winston cone was used to concentrate the diffuse FIR from the FT-IR onto the edge of the sample to increase the signal.

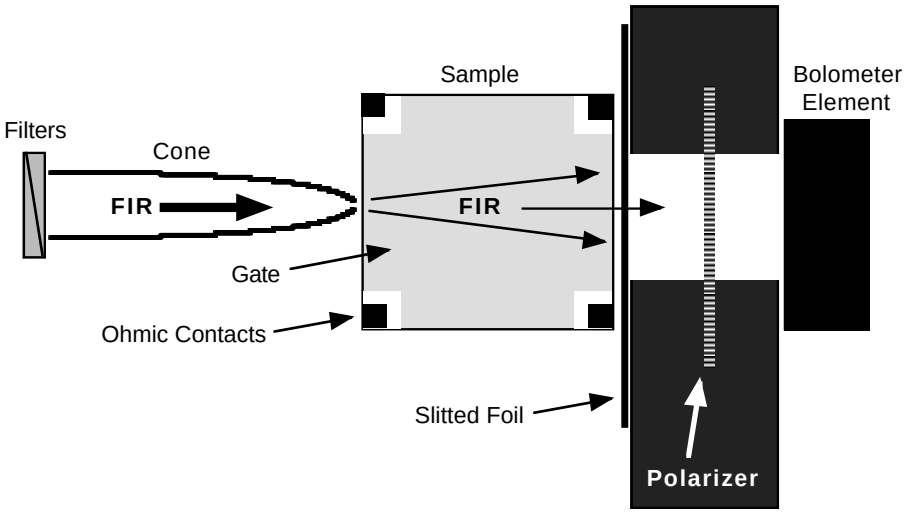

Figure 3.2: The setup inside the FT-IR cryostat. The radiation enters from the spectrometer (from the left), passes through the filters, the cone, the sample, the polarizer, and finally into the detector. Inside the sample the radiation spreads in a wedge. (See Appendix D for details.)

3.3 Absorption Results: Square Well

All data in this section were taken on the 40.0 nm square well. The attenuation coefficient $\alpha(\omega)$ was calculated from the transmission spectra of the sample at each gate voltage normalized to the reference transmission of the depleted sample. The attenuation coefficient is defined by the relation:

$$I(L) = I(0)\exp(-\alpha L) \tag{3.2}$$

where $I(L)$ is the intensity of radiation after passing through a length L of the sample, and $I(0)$ is the initial intensity. Therefore $\alpha(\omega)$ was calculated with

$$\alpha(\omega) = \frac{1}{L}\ln\left(\frac{I_{\mathrm{depleted}}(\omega)}{I_{\mathrm{partially-full}}(\omega)}\right) \tag{3.3}$$

where L is now the length of the sample, $I_{\mathrm{depleted}}(\omega)$ is the reference spectrum of the depleted sample, and $I_{\mathrm{partially-full}}(\omega)$ is the spectrum with electrons in the sam-

ple. The attenuation coefficient (absorption) spectra were then fit to a Lorentzian lineshape, and the various parameters of the fit were extracted.

The raw attenuation coefficient as a function of frequency is shown in Figure 3.3 for many different gate voltages. In this figure, as well as all the other results in this section, I have removed the "dip" in the spectrum due to residual absorption in the reference. (For more information and details, see Appendix E.)

For the data in this section, the well had a charge density of 9.70×10^{10} e/cm^2 at $V_G = 0$ volts, which decreased approximately linearly with gate voltage to a residual charge density of approximately 1.6×10^{10} e/cm^2 at $V_G = -1.3$ volts.

3.3.1 Frequency

The intersubband resonance frequency for the 40 nm square well is shown in Figure 3.4. The dependence on gate voltage is due to two effects: the static field at the well ($E_{\rm dc}$) and the depolarization shift. (For an explanation of the static field across the well, $E_{\rm dc}$, see Appendix B.) Because the square well potential is symmetric, it has no linear Stark shift;[2] the lowest order Stark effect is second-order. Thus, the single-particle square well has a minimum in absorption frequency at zero static field; applying a field in either direction shifts the absorption to higher frequencies, and the shift is proportional to the static field squared. (This is why the frequency looks approximately parabolic as a function of gate voltage.)

The depolarization shift is another source of frequency shift. As more charge is added to the well, the depolarization shift increases the absorption frequency. Hence, as the gate voltage is changed, *both* the charge density and the static field across

[2]The Stark shift is the shift in absorption frequency due to a static electric field.

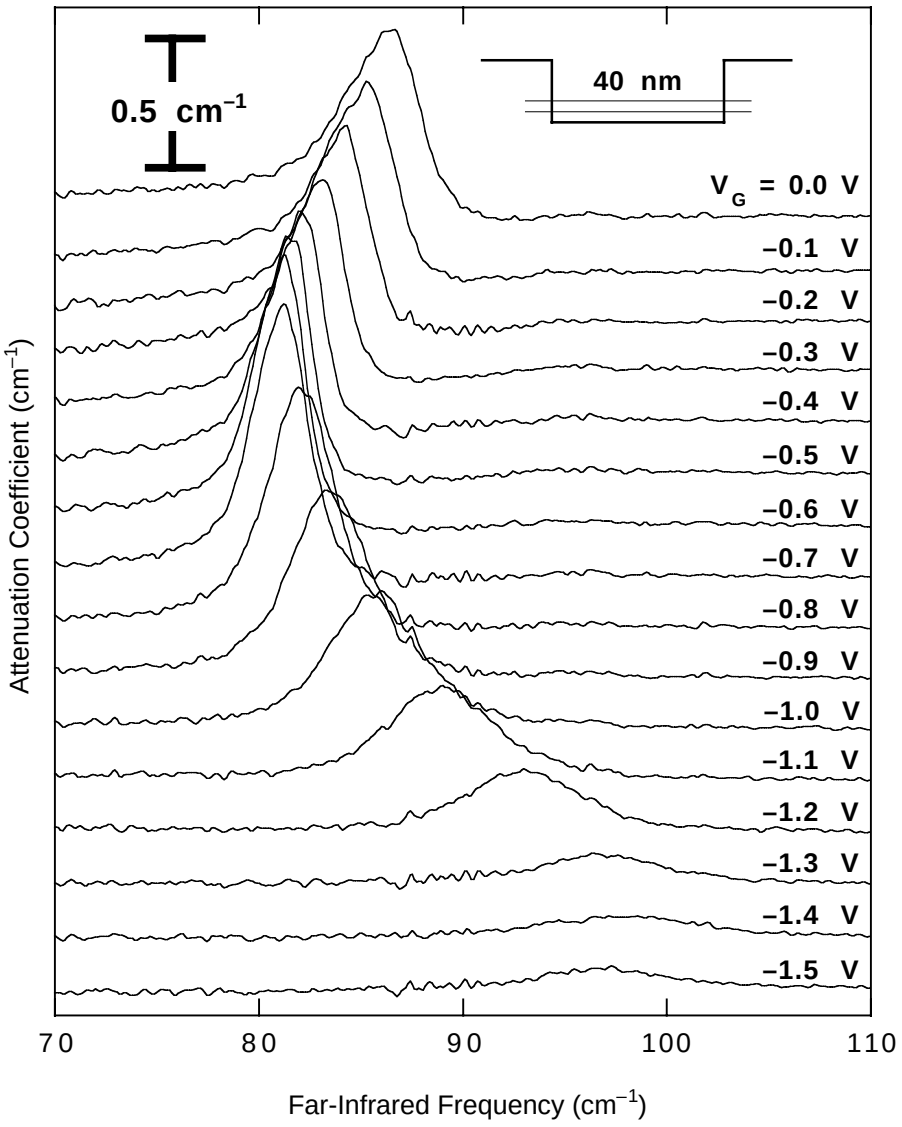

Figure 3.3: The attenuation coefficient as a function of frequency for the square well sample. The absorption for a number of gate voltages are shown, as labeled on the plot. The hole due to residual absorption in the reference has been removed (see Appendix E for details).

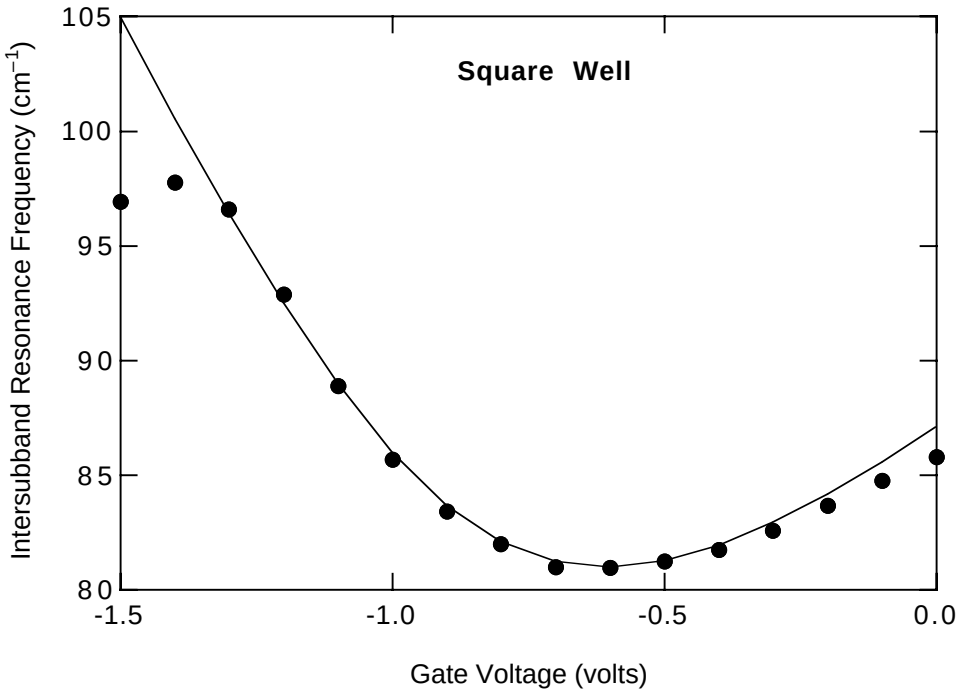

Figure 3.4: The linear intersubband resonance frequencies for the square well sample, as a function of gate voltage. The dots are the experimental data, the solid curve is a fit using the standard perturbative waveguide theory. I believe the deviations from the fit at low charge densities are due to the charge in the well disconnecting from the contacts.

the well change, and the frequency is shifted by *both* a change in the depolarization shift and a change in the Stark shift. Bryan Galdrikian's self-consistent program (see Chapter 2 and Appendix B) was used to calculate these shifts. Figure 3.5 shows the output of the program for the bare subband spacing E_{12} and the resonance frequency $\tilde{E}_{12}$ with and without the exchange-correlation potential.

The fit to the frequency in Figure 3.4 was calculated in several steps. First, I guessed at a gate voltage where the static electric field $E_{\rm dc}$ was zero. (See Appendix B for details.) Then the self-consistent program was used to find the correct well width to fit the absorption frequency at this gate voltage. Once the well width

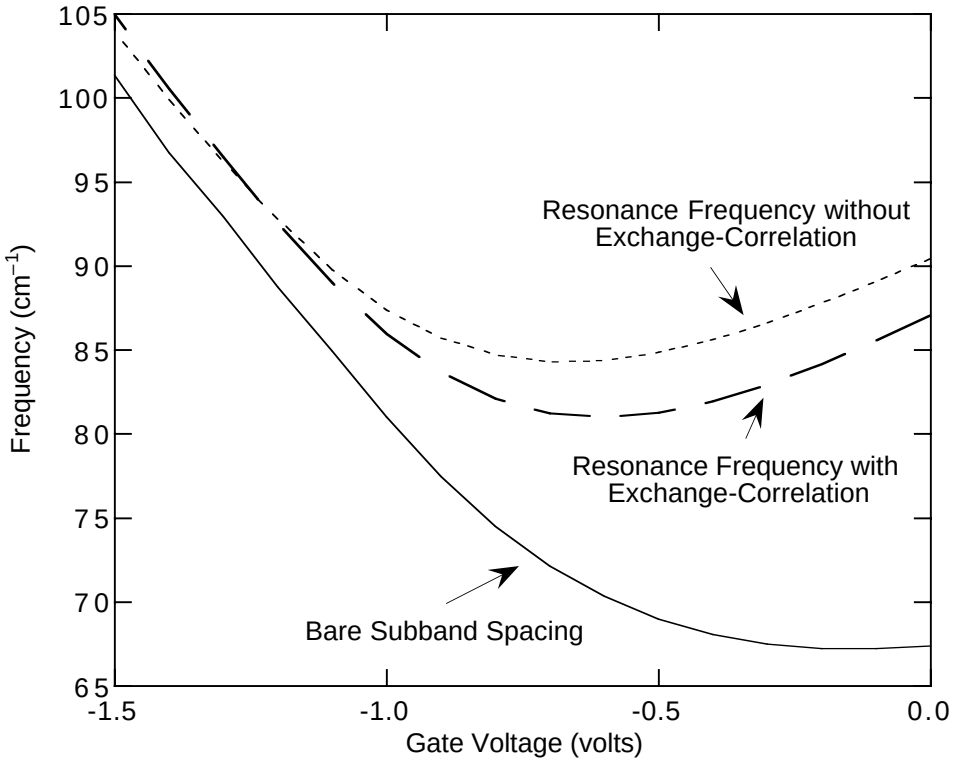

Figure 3.5: The results of Bryan Galdrikian's simulation for the square well with and without exchange and correlation. The curve labeled "Bare Subband Spacing" is the separation of the subbands in the self-consistent potential (E_{12}) without the exchange-correlation potential. The curves labeled "Resonance Frequency" are the intersubband resonance frequencies where radiation is absorbed ($\widetilde{E}_{12}$) with and without the exchange-correlation. The depolarization shift is the difference between the bare subband spacing and the resonance frequencies.

Parameter	**"Expected"**	**Self-Consistent**
Well Width	40.0 nm	38.832 nm
Charge Density at $V_G = 0$	$\sim 1.0 \times 10^{11}$ e/cm^2	9.70×10^{10} e/cm^2
V_G at $E_{\rm dc} = 0$	0 volts	-0.45 volts

Table 3.1: The parameters from the growth of the square well sample, and the input values for the self-consistent program. These values were used to simulate the linear perturbative absorption. ("V_G at $E_{\rm dc} = 0$" means the gate voltage at which there was no static electric field across the well.)

was found, the self-consistent program was run using this well width to determine the well parameters at each gate voltage. Finally, the results of this simulation were compared with the data. If the comparison indicated that the initial guess at the zero-field voltage had been incorrect, then a new guess was made and the entire process was repeated. The input parameters to the self-consistent program used to generate the fit shown are given in Table 3.1.

The fit used in Figure 3.4 uses the standard perturbative approximation for the waveguide formed by the sample. (This means that the absorption was calculated using Equation (C.9).) The fit obtained is quite good; it overestimates the frequency slightly at large charge densities. (Note that below $V_G = -1.3$ volts, I believe the electron gas in the well is disconnected from the ohmic contacts, so the static electric field at the well stops changing. See Appendix E for a discussion.) Considering that the only free parameters in the theory are the gate voltage where the static field is zero and the width of the well, this is quite good. Note that the more complex waveguide theory of Chapter 4 does not improve the fit measurably.

3.3.2 Integrated Absorption

The integrated absorption, defined by:

$$A = \int_{\omega=0}^{\omega=\infty} \alpha(\omega)\, d\omega \tag{3.4}$$

is an accurate measure of the absorption strength which does not depend upon the width of the peak. (Since the total area of the peak is proportional to the strength of the absorption, changes in the linewidth cause changes in the height of the peak.)

The integrated absorption was expected to be linearly proportional to the charge density. However, this was not the case for the absorption of the square well. As shown in Figure 3.6, the experimental data start out at low charge density increasing more rapidly than the linear perturbative theory; then they suddenly turn over and start decreasing around a gate voltage of -0.7 V. The fit was calculated using Equation (C.27); note that there are no free parameters in this fit. The best agreement *appears* to be at low charge density, but this is an artifact of the fitting process. At low charge densities, the ohmic contacts disconnect from the sample, and the residual charge in the sample is not detected by the capacitance measurements. Therefore, it was necessary to calculate the residual charge density using the optical data and linear perturbation theory. However, when this charge density is estimated from the absorption strength, the absorption strength calculated from the calculated charge density duplicates the absorption strength at the lowest charge density. (In other words, the fitting at low charge density is mathematically circular.)

Note that the long time constants which the sample exhibited to stabilize the amount of charge after changing the gate voltage may have effected this data somewhat. (See Appendix J for details about the long time constants.) I was not aware of the long time constants when I did these measurements. The sample was depleted

several times during the measurements, and then partially full measurements were done just after the depletion. The only clear signal of this is that the $V_G = 0$ absorption appears somewhat larger than the trend; this point was measured just seconds after the sample had been depleted. (The long time constants would cause extra charge to be in the well, and therefore make the integrated absorption anomalously large.)

The striking disagreement between theory and experiment for the absorption strength was quite surprising. It is especially strange that, at low charge densities, the sample was measured to absorb *more* radiation than expected. The only plausible explanation we can find is that the coupling to the well in the waveguide, at least at low charge density, is stronger than was predicted by linear perturbation theory. The failure of the linear perturbative theory of absorption led us to question our most basic assumptions, and eventually led to our use of the waveguide theory described in detail in Chapter 4.

3.3.3 Peak Width

The observed absorption peaks for the square well were relatively narrow, as shown in Figure 3.7. It is beyond the scope of this work to predict the linewidths. However, it is probable that the lines are narrow partially because our wide quantum well gives the electrons “plenty of room” to exist in the well, which means that the low probability of an electron being found at the walls of the well make the rough interfaces on those walls less important.

One interesting aspect of the data is that the line starts out broad at low charge density, and then becomes more narrow as the charge density increases. This is very unexpected; one would imagine that as the charge density increased, more scattering between electrons would cause the peaks to become wider. This may suggest

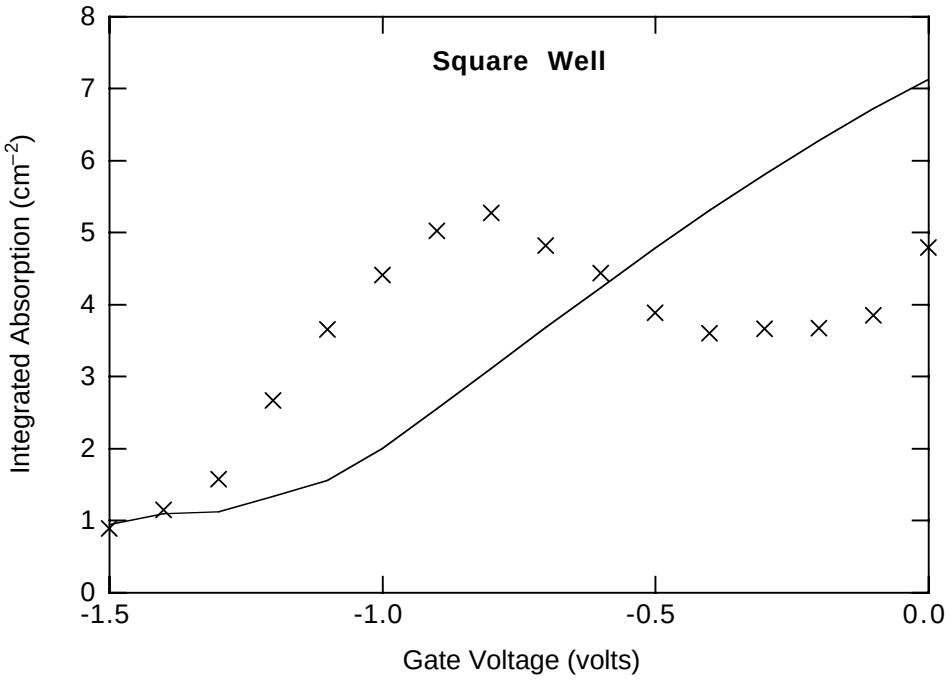

Figure 3.6: The integrated absorption strength for the square well sample, as a function of gate voltage. The "X"s are the experimental data, the solid curve is a fit using the standard perturbative waveguide theory. Note that the excellent agreement at low charge densities is an artifact of the fitting process; see text for details.

some sort of correlation or many-body effect which causes the peak to narrow as the charge density increases. This might also explain the anomalous integrated absorption results.

3.3.4 Lineshape and Baselines

The lineshape was another mystery. At low charge density, the line is fit exactly (to within small experimental fluctuations) by a Lorentzian fit. An example is shown in Figure 3.8 for $V_G = -1.0$ V. However, at larger charge densities, the fit clearly deviates from a Lorentzian; Figure 3.9 shows the fit at $V_G = 0$ V.

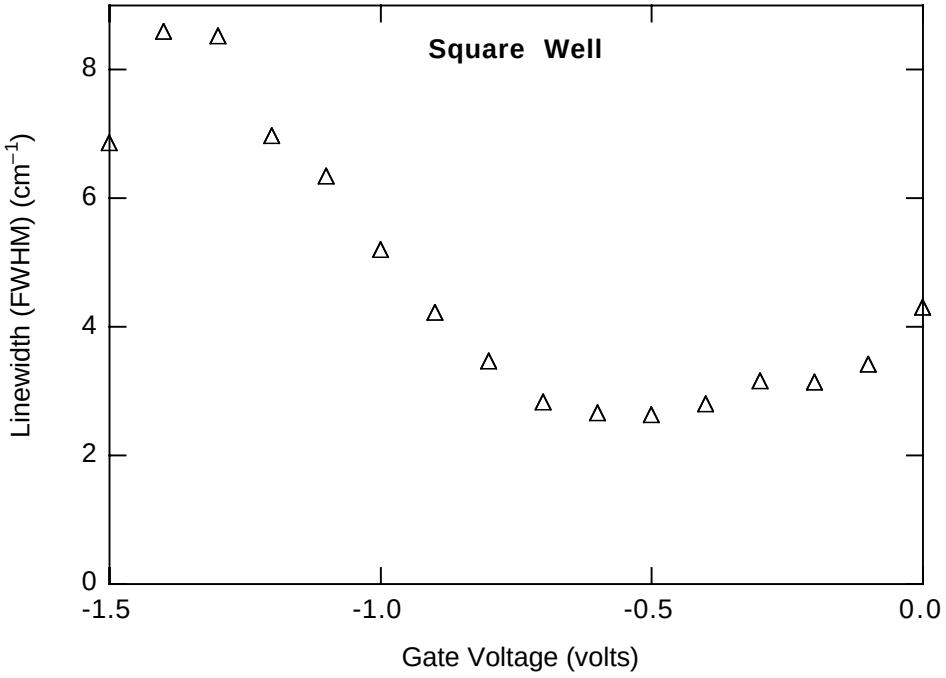

Figure 3.7: The full-width at half-maximum (FWHM) of the absorption peaks for the square well sample, as a function of gate voltage. The open triangles are the experimental data; a prediction of the linewidth is beyond the scope of this work. See Figure 4.8 for comparison.

In addition, the baseline far from the peak seems to be dependent upon the charge density, but not on the overall absorption strength. (By baseline I mean the amount of absorption seen far from the peak, where the absorption is normally very close to zero.) In Figure 3.9, there is a roughly constant amount of absorption at frequencies below the peak, and there is a roughly constant negative amount of absorption at frequencies above the peak. Although the absorption strength is similar at lower charge density in Figure 3.8, the baseline there closely follows the fit as expected.

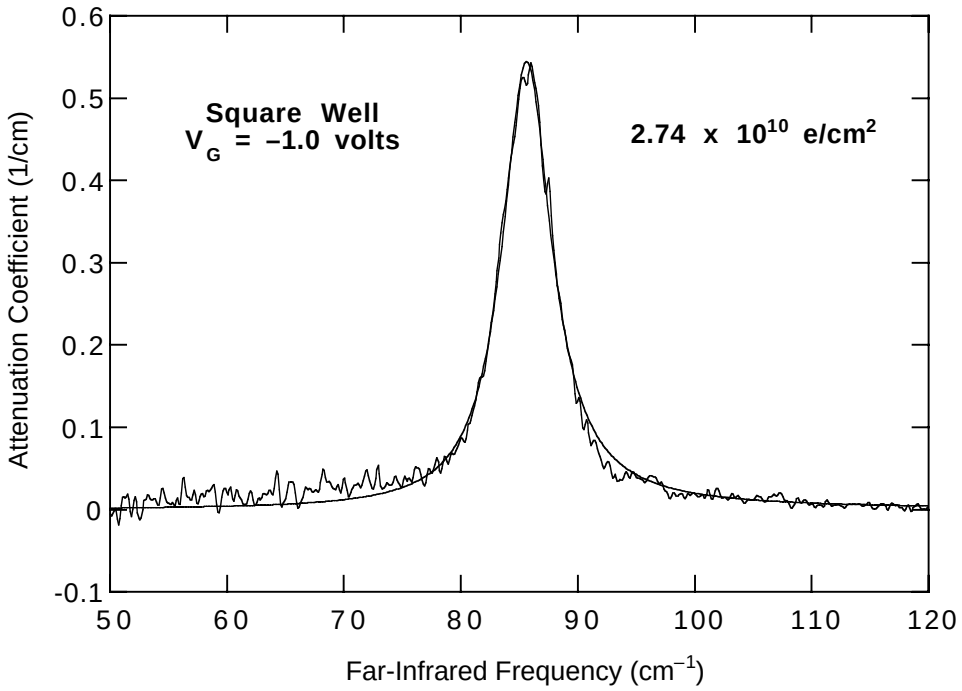

Figure 3.8: The attenuation coefficient and Lorentzian fit as a function of frequency at -1.0 volts gate bias for the square well sample. At these low charge densities, the peak and baseline are well fit by a Lorentzian.

3.3.5 Discussion of Anomalies

There are several unusual aspects of the linear absorption data for the square well which appeared at high charge densities:

- Integrated absorption strength does not depend linearly on charge density
- Non-Lorentzian lineshape
- Roughly constant baseline shifts above and below peak
- Significant narrowing of linewidth.

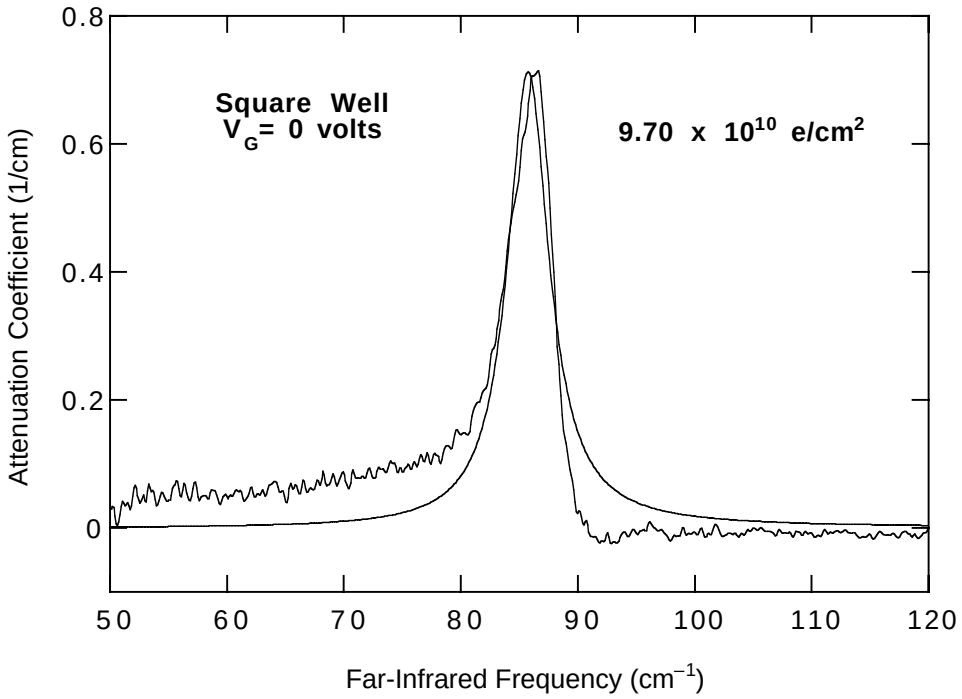

Figure 3.9: The attenuation coefficient and Lorentzian fit as a function of frequency at zero volts gate bias for the square well sample. At these higher charge densities, the peak and baseline both deviate significantly from the fit.

These may be actual effects, or they may be artifacts of the measurement process. An example of an artifact is a problem involving the phase correction spectrum used to symmetrize the interferograms. In a normal sample which does not absorb strongly, it is fine to use the same phase correction spectrum for the absorbing and reference samples. However, our samples absorb so strongly ($\alpha \approx 0.7$ 1/cm in samples roughly 1.0 cm long) that the absorbing and reference interferograms have very different structure. Using the same phase correction (from the depleted sample) to symmetrize these two interferograms may have introduced artifacts into the spectra. If this is true, it could mean that the anomalies have no physical meaning.

However, this phase correction problem should also exist for the double square well sample, which has a comparable absorption strength. None of the anomalies listed above were seen in that sample, which was measured using the same techniques. (In addition, the baseline looks very different in Figures 3.9 and 3.8, which are both the square well sample at different charge densities but comparable absorption strength.) This indicates that the anomalies are physical: the most likely explanation would then be some sort of many-body or waveguide effect. These would both become more important at higher charge densities, as observed, and could also be dependent upon the high mobility (μ = 360,000 cm^2/V s) and narrow linewidth observed in the square well.

A many-body effect would involve an interaction between the electrons which occurs more strongly at higher charge densities. The simulations of the well were calculated in the local density approximation, which replaces the true many-body wavefunctions with an approximate single wavefunction for the entire electron gas. (It is currently an open question how to calculate the true many-body wavefunctions for a realistic system.) Since this theoretical problem has not been solved, we have no idea what the result of a full many-body calculation would be.

Another example of an artifact would be a misalignment of the spectrometer which caused it to incorrectly choose the location of zero-path difference for the Fourier Transform calculation. Jon Williams has recently mentioned that even small errors in the location of ZPD (zero-path difference) can cause the peaks to become asymmetric and the baselines to be shifted. The Bomem spectrometer finds ZPD based on the interference signal of a white light source. The construction of the spectrometer makes it possible for the white-light optics to be out of alignment with the FIR optics. (For example, the FIR beamsplitters mount onto a holder which contains

the white-light beamsplitter. A small particle of metal or dust could have caused the FIR beamsplitter to be in a slightly different position than the white-light beamsplitter.) Therefore, the peak shape and baseline shifts could be caused by a misalignment of the spectrometer. This explanation sounds quite reasonable. However, it does not explain why the baseline and peak shape would look distorted when the well is full and normal when the well is partially depleted on the same day.

Yet another possibility is that the anomalies listed above are due to waveguide effects, where the conductivity of the well at higher charge densities changes the boundary conditions for the waveguide which the radiation is propagating through. Changing the boundary conditions (which now become a function of frequency) means that the modes change (as a function of frequency) and change the coupling of the modes into the well region. This possibility is explored in Chapter 4.

3.4 Comparison with the Double Square Well

In general, the data from the square well differ quite a bit from the observed absorption of the double square well, which is in many regards a similar sample. The double square well sample had a lower mobility (μ = 10^5 cm^2/V s) than the single square well due to the increased scattering of electrons from the barrier between the wells. The linewidth was also wider than the single square well. Interestingly (and again unlike the single square well), nothing strange happens as the charge density increases: the peak does not narrow, the lineshape remains Lorentzian, the baseline remains near zero, and the integrated absorption depends linearly on the charge density.

First consider the absorption data for the double square well, Figure 3.10. (There was no measurable residual absorption in this sample.)

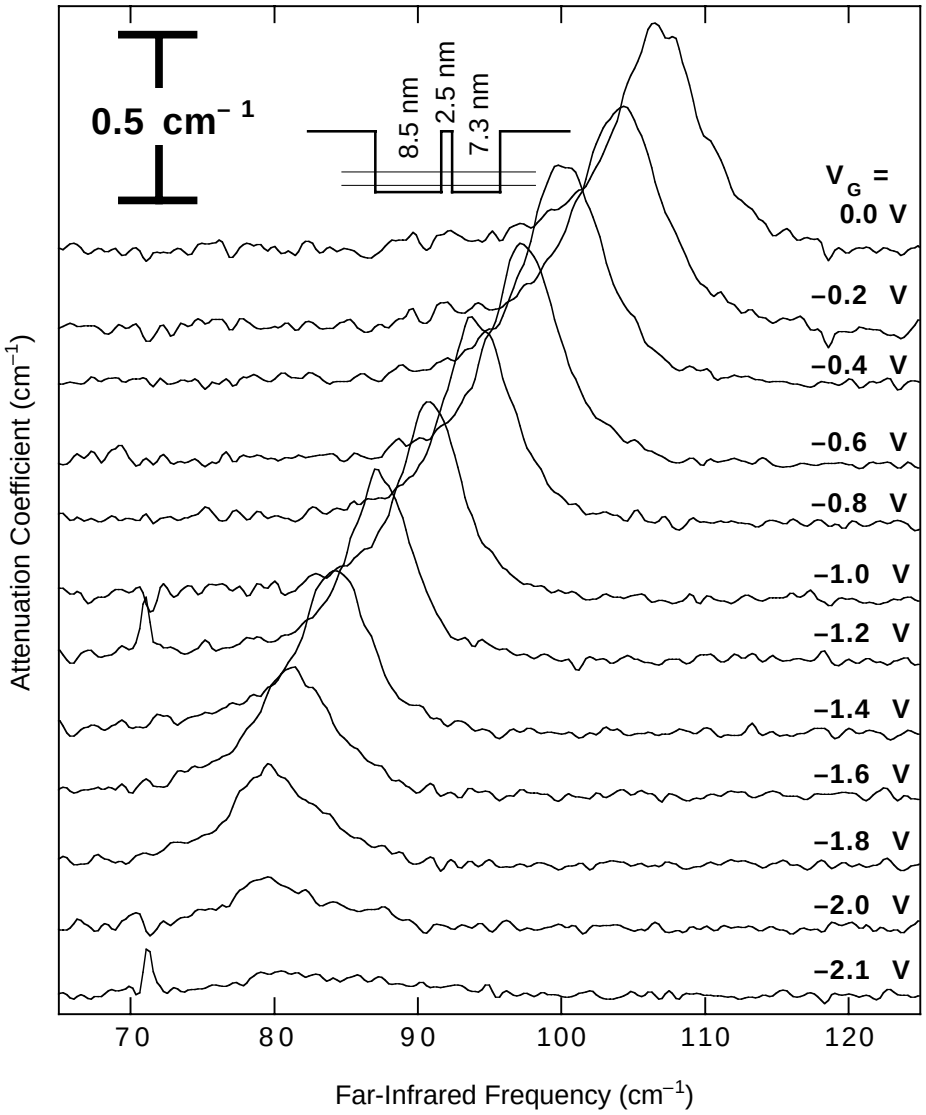

Figure 3.10: The linear attenuation coefficient as a function of frequency for the double square well. The absorption for a number of gate voltages is shown; each curve is labeled with the appropriate voltage.

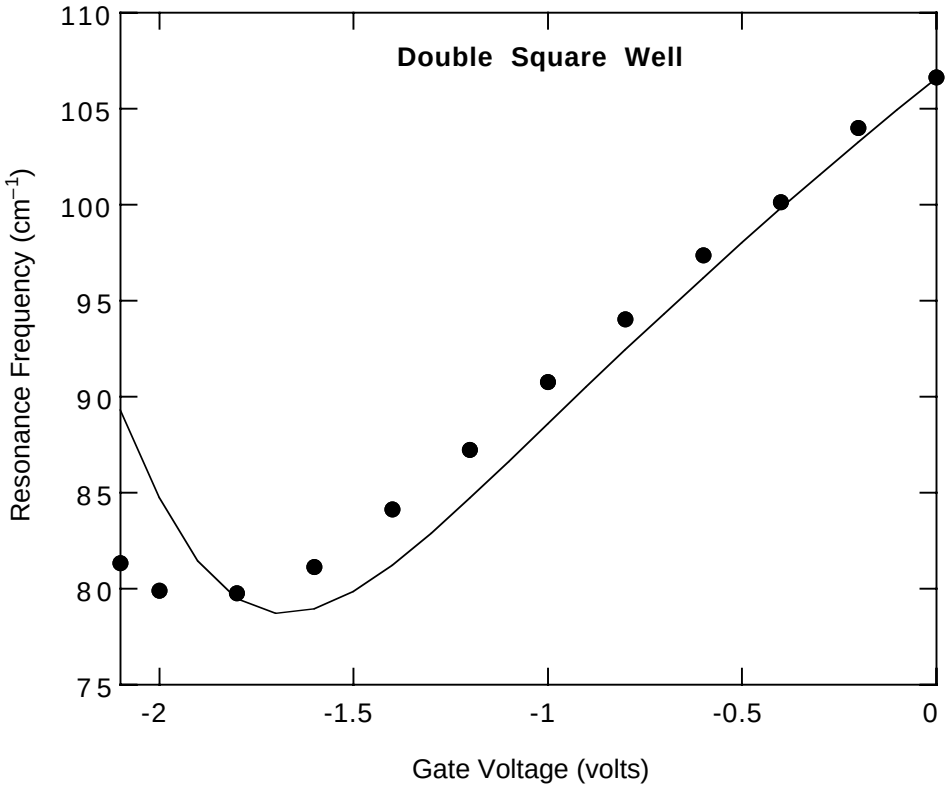

Figure 3.11: The frequency dependence of the intersubband absorption for the double square well as a function of gate voltage. The dots are the experimental data, and the curve is a fit using linear perturbative theory.

The frequency dependence of the resonance for the double square well is shown in Figure 3.11. The parameters used to fit the data are shown in Table 3.2.

The integrated absorption of the double square well is shown in Figure 3.12. Note that the integrated absorption is, within experimental uncertainty, well fit by the linear perturbative theory (which is linear in the charge density). This is in contrast to the single square well, whose integrated absorption was not linear in the charge density.

The linewidth of the absorption peaks for the double square well is shown in

Parameter	**"Expected"**	**Self-Consistent**
Well-Barrier-Well Widths	8.5-2.5-7.3 nm	8.785-2.584-7.549 nm (Growth $\times$ 1.03355)
Charge Density at $V_G = 0$	$\sim 2.0 \times 10^{11}$ e/cm^2	2.58×10^{11} e/cm^2
V_G at $E_{\rm dc} = 0$	0 volts	-0.25 volts

Table 3.2: The parameters from the growth of the double square well, and the input values for the self-consistent program. These values were used to simulate the linear perturbative absorption.

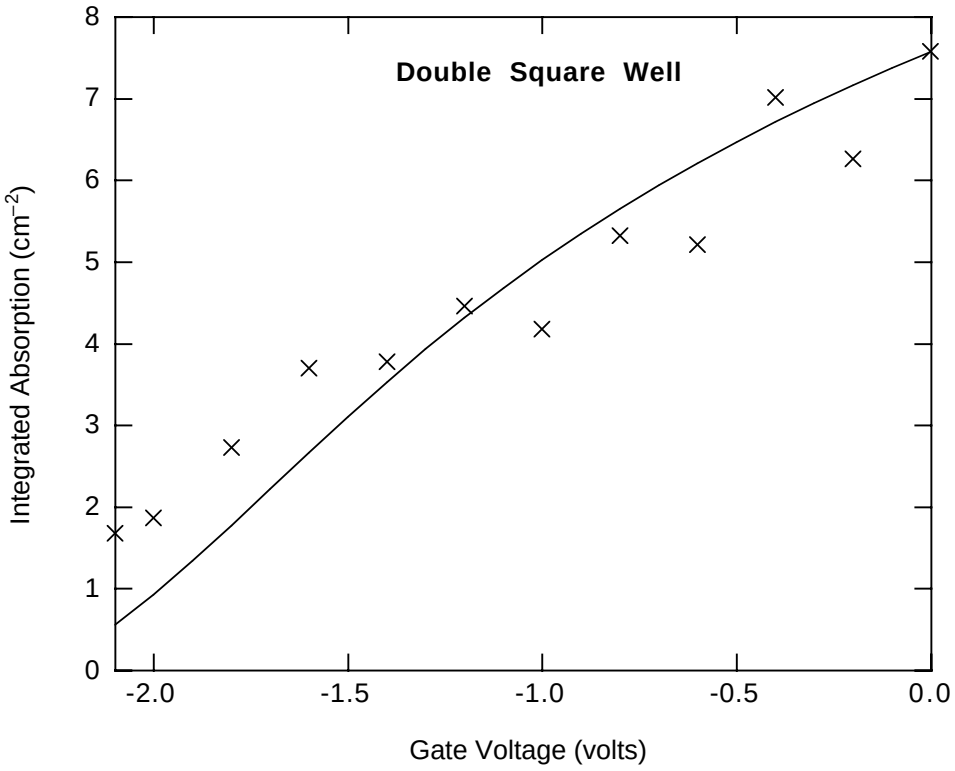

Figure 3.12: The integrated absorption of the double square well as a function of gate voltage. The dots are the experimental data, and the curve is a fit using linear perturbative theory.

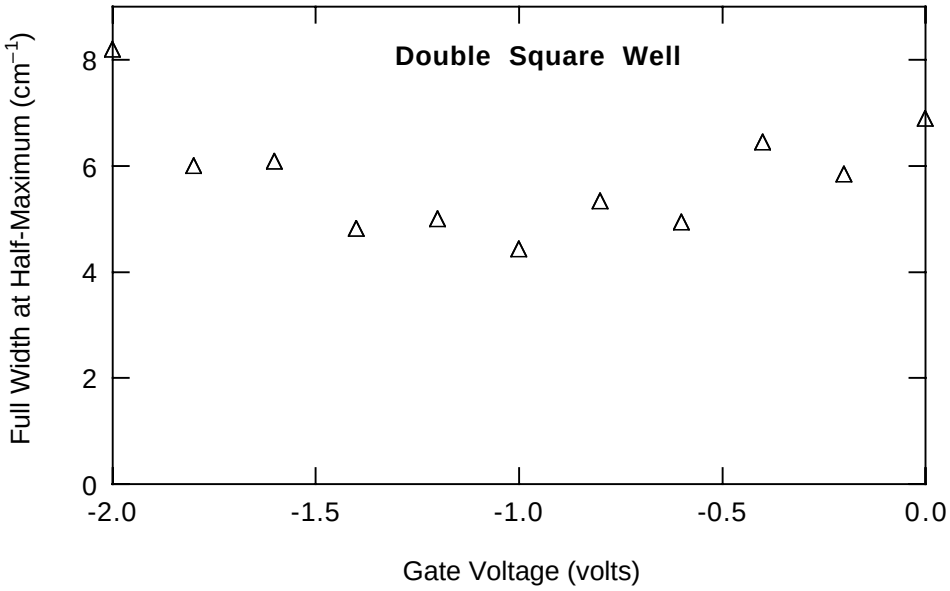

Figure 3.13: The linewidth of the absorption peak of the double square well as a function of gate voltage.

Figure 3.13. Again unlike the single square well, there is not a marked decrease in linewidth as the charge density increases.

Finally, Figure 3.14 clearly shows that even at the highest charge density measured, the attenuation coefficient of the double square well is very well fit by a Lorentzian fit function. Unlike the single square well, both the peak and the baseline are well fit.

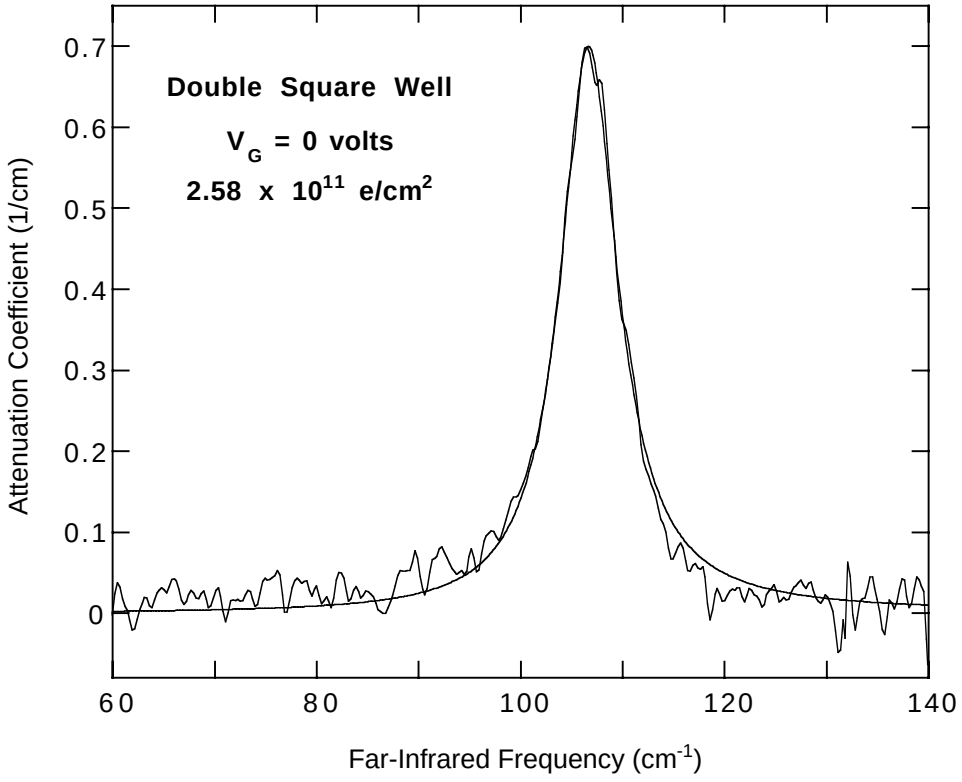

Figure 3.14: The attenuation coefficient as a function of frequency for the double square well at $V_G = 0$ volts. The solid line is the experimental data, and the dashed line is the Lorentzian fit.

Chapter 4

Waveguide Theory

The linear absorption measurements revealed a very surprising result: the integrated absorption of the intersubband resonance was not a linear function of charge density for the square well sample. (See Figure 3.6 in Chapter 3.) The failure of the linear approximation was shocking and completely unexpected: I had always taken it for granted that the absorption strength of the quantum well would be linearly proportional to the number of electrons, as predicted by Equation (3.1).

Furthermore, at low charge densities, the integrated absorption appears to be linear in the charge density. This sounds fine, but when these data are used to calculate the dipole matrix element, the resulting oscillator strength was $f_{12} = 1.4$. This is too large and hence non-physical, but it adds to the mystery of what is really happening with the absorption strength.

I still do not fully understand why our samples have a "nonlinear integrated absorption" as I have come to call it. A possible explanation is that the nonlinear integrated absorption is caused by many-body effects which are not included in the local density approximation (L.D.A.).

Another explanation of the nonlinear integrated absorption is that the very strong absorption changed details of the waveguide modes, which modified the absorption. Consider the radiation propagating down the waveguide formed by the sample, encased on either side with an aluminum film. As the conductivity of the well increases, the waveguide modes travelling down the sample begin to "see" a different boundary condition at the quantum well side of the waveguide. It is an unusual boundary condition: a film with a large peak in the conductivity in the growth direction, σ_{zz}, at the intersubband resonance frequency. As the conductivity increases, the waveguide modes could change in such a way as to "avoid" the side of the waveguide with the quantum well. In this case, as the electric field at the well decreases, the absorption would decrease. This would explain the violation of the simple idea that the absorption is proportional to the number of electrons: adding more electrons would change the waveguide modes, and that could reduce the absorption. The ultimate limit is that the conductivity is so high that the modes completely avoid the quantum well, leading to zero absorption! Compare this to the standard perturbative absorption calculation in a waveguide: there one ignores the quantum well in solving for the modes, and then assumes the well does not change the modes at all, other than the absorption. Certainly this perturbative assumption must break down at some point.

The square well sample had a very narrow linewidth, with a full-width at half-maximum of roughly 2.5 cm^{-1}. This concentrates all the oscillator strength of the transition in a very narrow frequency range, leading to a very high *peak* conductivity. This is a crucial parameter for the type of mode-changing effect described above; the peak conductivity would determine the boundary condition for the mode at that frequency, and the mode would determine how much electric field there was at the well to be absorbed at that frequency. (A wide peak would spread the conductivity

over a wide frequency range, reducing the *peak* conductivity and the effect on the waveguide modes.)

In this chapter I will discuss a theory by M. Załużny [31] which is based on the effects described above, and explain my extensions to it, as well as discussing the successes and failures of the theory at modelling the experiment data.

4.1 Załużny Simulations

After communicating with M. Załużny about my nonlinear integrated absorption results, he was able to create a theory incorporating the ideas outlined above. (See Reference [31].) In short, he started with the expression for the intersubband conductivity, and then used it as a boundary condition to find the waveguide modes in the sample. The attenuation coefficient of these waveguide modes can then be found from an eigenvalue problem.

The theory only calculates the absorption for a single mode traveling down the waveguide. In order to model the experimental data, I adapted the theory to the case of a superposition of modes. This involved two challenges: how to determine the superposition of modes, and how to calculate the absorption due to a superposition of modes. These challenges were greatly complicated by the fact that the eigenvalue problem is non-Hermitian: the eigenvalues have a real part which determines the propagation speed through the sample, and an imaginary part which determines the loss (absorption).

4.1.1 Numerical Solution of Eigenvalue Problem

The theory of Załużny includes several complex wavevectors: q_z is the wavevector in the growth direction, $q_{tot} = (\omega\sqrt{\epsilon}/c)$ is the total wavevector (and is pure real), and q is the wavevector in the direction propagating down the waveguide. The electric field in the waveguide can be written:

$$\vec{E}(\vec{r}, t) = \vec{E}(z)\exp[i(qx - \omega t)] \tag{4.1}$$

where z is in the growth direction, and x is in the propagation direction. The wavevectors are related by the geometrical relation:

$$q_{tot}^2 = q_z^2 + q^2. \tag{4.2}$$

Since q_{tot} is known for a given mode, if q_z can be calculated, then q can be found. Then $\alpha = 2\,\mathbf{Im}(q)$ gives the desired attenuation coefficient for the mode propagating down the waveguide. (The factor of 2 comes from the fact that α is the attenuation coefficient for *intensity*, not electric field.)

Finding the attenuation coefficient reduces to solving for q and q_z in the transcendental eigenvalue equation which Załużny derives:

$$\tan(q_z L) = (-i)\frac{q^2}{q_{tot} q_z}\Lambda(\omega) \tag{4.3}$$

where $\Lambda(\omega)$ is the scaled intersubband conductivity, and L is the thickness of the waveguide (in the growth direction). (See Reference [31], Equations 9 and 16.) Solving Equation (4.3) is challenging: the equation is transcendental, the eigenvalue q_z is complex, and the equation itself has both real and imaginary parts.

Finding the eigenvalues of Equation (4.3) is equivalent to simultaneously finding the roots of two real equations with two real unknowns. Due to the fact that the

equation is transcendental, the root finding must be done numerically. Finding roots in a plane numerically is quite difficult in general, and it is even more challenging in this case because there are *many* possible roots to these equations which correspond to different modes. In addition, as the parameters are smoothly varied (for example, as the charge density is increased), the different roots exhibit a very complicated anti-crossing behavior.

In order to solve this problem numerically, I wrote a computer program which starts at a frequency which is *very* far below the resonance, where the solution to the eigenvalue equation is very close to the known perturbative value (since the conductivity far from the peak is so small). Then, by using a two-dimensional iterative root finder,[1] I was able to slowly follow the desired root into the nonlinear region by using very small steps of the frequency and using the previously found root as the initial guess for finding the next root. As long as the frequency steps were kept small enough, I could flawlessly track the individual roots through all the anti-crossings. This procedure was quite successful, although it was moderately computationally intensive.

4.1.2 Behavior of Modes

In general, the conductivity of the quantum well distorts the electric field shape of the modes. This is caused by the fact that the boundary condition on the $z = 0$ side of the waveguide is no longer simply a perfect metal. (The $z = 0$ side of the waveguide is the one the quantum well is on.)

Figures 4.1 and 4.2 plot the x-component of the Poynting vector (simply de-

[1] I used the Newton-Raphson method for nonlinear systems of equations from *Numerical Recipes in C*, Section 9.6, Reference [20].

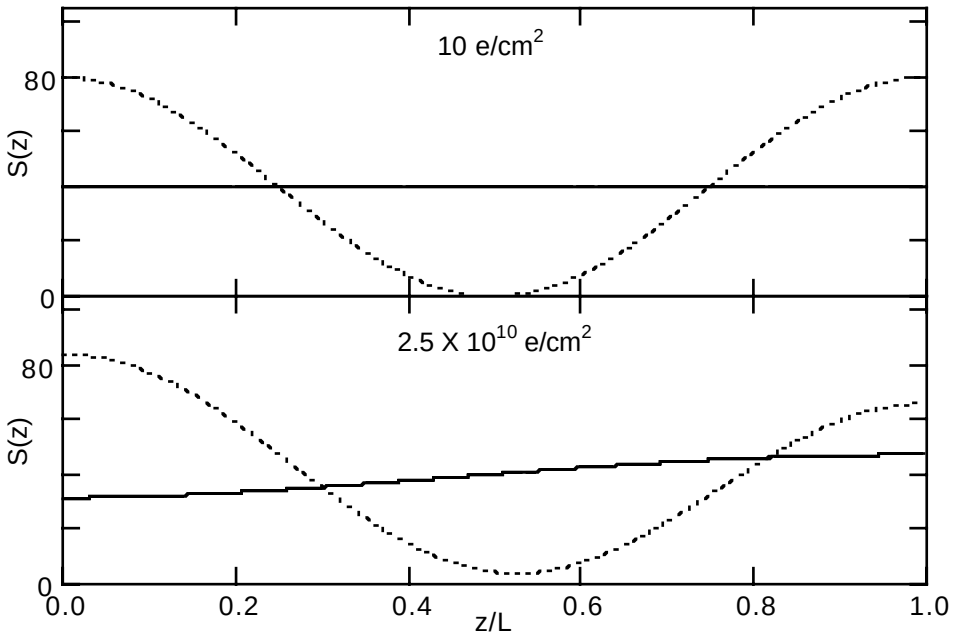

Figure 4.1: The Poynting vector $S(z)$ of the lowest two modes as a function of position in the waveguide, at two low charge densities. The quantum well is at $z = 0$, and the width of the waveguide is L. The $n = 0$ mode is the solid curve, and the $n = 1$ mode is the dashed curve. The modes are calculated for the resonance frequency of the well.

noted $S(z)$) for the lowest two modes at several charge densities. (The x-component of the Poynting vector is the component of energy flow *down* the waveguide.) The parameters used to calculate these modes were $\omega_{12} = 70$ cm^{-1}, $\widetilde{\omega}_{12} = 80$ cm^{-1}, $\Gamma = 1.5$ cm^{-1}, the waveguide width $L = 0.0455$ cm, and the dipole matrix element $\langle z_{12} \rangle = 8.0$ nm.

As expected, the modes distort as the conductivity of the well increases. These modes are calculated at the intersubband resonance frequency for the well; the modes at any other frequency will look different.

The important quantity for absorption is the size of the Poynting vector at the well: since the well can only absorb radiation which is *at the well*, the amount of

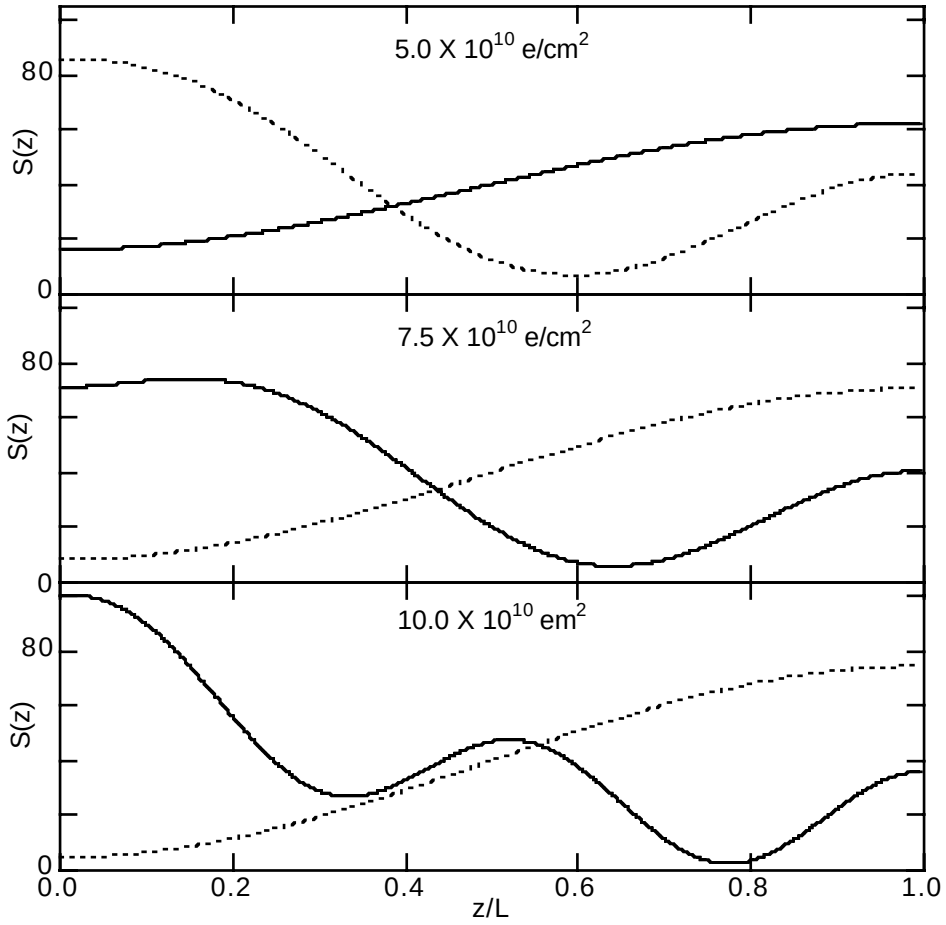

Figure 4.2: The Poynting vector $S(z)$ of the lowest two modes as a function of position in the waveguide, at two charge densities. The $n = 0$ mode is the solid curve and the $n = 1$ mode is the dashed curve. The modes are calculated for the resonance frequency of the well. Modes $n = 0$ and $n = 1$ anti-cross at 6.87×10^{10} e/cm^2 at a lower frequency, so they switch shapes from the top plot to the middle plot. Another anti-crossing has taken place between the $n = 1$ mode and a higher mode in the bottom plot.

absorption of a mode is proportional to $S(z = 0)$. Some of the modes "avoid" the quantum well as the conductivity increases by distorting in such a way that the Poynting vector at the well decreases. However, some modes actually increase the Poynting vector at the well as the conductivity increases. As the conductivity increases in Figures 4.1 and 4.2, the $n = 0$ mode's Poynting vector at the well decreases while the $n = 1$ mode's Poynting vector at the well increases. Therefore, the attenuation coefficient for the $n = 0$ mode will decrease as the charge density increases, while the attenuation coefficient for the $n = 1$ mode will increase.

Figure 4.2 illustrates the anti-crossing behavior mentioned in the last section. From the top plot in Figure 4.2 to the middle plot, the two modes shown anti-cross with each other at a lower frequency. In other words, when tracking these modes from very low frequency in the linear regime, at some frequency below the resonance, the modes anti-cross and switch character. Thus, when plotting the Poynting vector at the resonance frequency, the modes look like they have switched labels. Furthermore, in the bottom plot in Figure 4.2, the mode which was originally $n = 1$ at low charge density (and had became $n = 0$ in the middle plot) has anti-crossed with a higher mode, taking on the character of the higher mode. This anti-crossing behavior is discussed further in Reference [31].

4.1.3 Determination of Superposition of Modes

The superposition of modes at the entrance to the waveguide is determined from the external electric field profile of the radiation which is incident on the sample edge. The coefficients of the superposition are found by matching boundary conditions for the electric field across the dielectric interface at the sample edge. The boundary condition is that the tangential electric field be continuous across the boundary. The

tangential component is the only interesting component of the field, because of the anisotropy of the intersubband conductivity. The inner product for the modes is:

$$\langle E_i(z)|E_j(z)\rangle \equiv \int_{z=0}^{z=L} E_i^*(z)E_j(z)\,dz \tag{4.4}$$

where L is the width of the waveguide, z is the coordinate in the growth direction, $E_i(z)$ and $E_j(z)$ are the complex electric fields for modes i and j, and the superscript "*" denotes complex conjugation. Since this is a non-Hermitian problem, the eigenvectors (modes) *are not orthogonal*, which adds some complications.

The total electric field in the waveguide can be expressed as a superposition in terms of the modes:

$$|E_{\rm total}\rangle = \sum_{n=1}^{n=N} c_n|n\rangle \tag{4.5}$$

where I have adopted the notation that $|n\rangle$ is the electric field of the n-th mode, $(c_1, c_2, \cdots, c_N)$ are the complex superposition coefficients, and that the number of modes is N.

Once the external electric field $E_{\rm ext}(z)$ is specified, the superposition coefficients for the modes just inside the sample may be found by solving:

$$\begin{pmatrix} \langle 1|1\rangle & \langle 1|2\rangle & \cdots & \langle 1|N\rangle \\ \langle 2|1\rangle & \langle 2|2\rangle & \cdots & \langle 2|N\rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle N|1\rangle & \langle N|1\rangle & \cdots & \langle N|N\rangle \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_N \end{pmatrix} = \begin{pmatrix} \langle 1|E_{\rm ext}\rangle \\ \langle 2|E_{\rm ext}\rangle \\ \vdots \\ \langle N|E_{\rm ext}\rangle \end{pmatrix} \tag{4.6}$$

where $|E_{\rm ext}\rangle$ is the external electric field, and there are N modes. This equation simply matches the external electric field to the total field just inside the waveguide. The calculated shape of the total field just inside the sample ($|E_{total}\rangle$) agreed exactly with the shape of the external field, showing that the boundary condition across the

interface was satisfied. (Note that all the quantities in Equation (4.6) are in general complex.)

The shape of the external electric field has now taken on a physical meaning. It determines which waveguide modes will be excited, and since each waveguide mode has a different attenuation coefficient, the details of the external field are now very important. The shape of the external field is discussed further below.

It is important to realize that *both* the modes and the superposition coefficients are functions of the conductivity, and therefore the frequency and charge density. As the conductivity of the well changes (with frequency or charge density) the behavior of the individual modes changes, causing the superposition coefficients $\{c_n\}$ to change. For example, one might think that since a constant external field will couple only into the lowest waveguide mode (which also has constant electric field), the absorption of the sample will just be the absorption of the lowest mode. This is wrong, because the changes in conductivity change the electric field of the lowest mode. Thus, in the nonlinear conductivity regime, additional modes will be excited by a constant external field, and one must also take their absorption into account.

4.1.4 Calculation of Attenuation Coefficient

The theory of Załużny [31] calculates the attenuation coefficient α in a natural way as the imaginary part of the wavevector which propagates the mode through the sample. However, this clearly works only for a single mode. In the case of multiple modes, one must calculate the Poynting vector at the front and the back of the sample from the electric fields[2] at the front and back of the sample. Then, using the length of the

[2]To calculate the electric field at the back of the sample, simply use the wavevector q to propagate the mode's field from the front of the sample: $E_{\rm back} = \exp(qD)E_{\rm front}$, where D is the length of the

sample, the attenuation coefficient can easily be calculated from

$$\alpha(\omega) = \frac{1}{D} \ln \left(\frac{S_{\text{front}}}{S_{\text{back}}} \right) \tag{4.7}$$

where S is the x-component of the Poynting vector integrated across the width of the sample, and D is the length of the sample.

Surprisingly, the calculated attenuation coefficient depends on the length of the sample. This is not a mistake! Consider the simple case of two modes propagating through a sample: $|1\rangle$ has a very large attenuation coefficient, and $|2\rangle$ has a very small attenuation coefficient. If the initial condition is equal energy in the two modes, the first mode will rapidly be absorbed, and the second mode will keep propagating for a very long time (and distance). If you chose a very short sample (shorter than the attenuation length of $|1\rangle$, ie $D \ll 1/\alpha_1$), then the attenuation coefficient will be roughly half the attenuation coefficient of $|1\rangle$, because only half of the radiation is in $|1\rangle$. But as longer and longer samples are considered, the attenuation coefficient of the samples will decrease. All the energy in $|1\rangle$ will be absorbed regardless of length, and not much loss occurs in $|2\rangle$, regardless of length. Thus, the transmission of the sample will always be roughly 50%, and the attenuation coefficient will depend strongly on the length.

4.1.5 External Electric Field Shape and Ray-Tracing

The computer program I wrote calculates the attenuation coefficient as a function of frequency for a given sample with a given external field shape. The external field

sample. Note that this not only decreases the field using the attenuation coefficient of the mode, but it also oscillates the field the correct amount; the relative phase with which the fields recombine at the back of the sample is critical!

shape was an unknown function. My first attempt was to try several qualitatively different field shapes, hoping that they would all give the same (or very similar) results. This would have meant that the shape of the external field had very little effect on the predictions of the theory. Instead, the different external field shapes caused the predictions of the theory to change radically! The frequency of the peak changed by several wavenumbers (a moderate change) and the linewidth changed by almost a factor of two. In addition, the *shape* of the peak changed: it became distinctly non-Lorentzian and rather asymmetric. The external field shape is an important parameter.

In an attempt to gain insight into the problem, I wrote a simple ray-tracing program[3] to see how a believable set of initial rays would traverse the Winston cone. The results were heavily dependent upon several arbitrary assumptions I made during the calculation. The main assumptions attempted to take the wavelength of light into account. One assumption was choosing the number of bins in the histogram which counts the density of rays passing through different spatial regions of the exit. (The bin width was selected to be approximately one-eighth of a wavelength.) In addition, approximately 5.7% of the rays were discarded because they were so close to the center of the cone that the radial weightings[4] used would cause them to contribute an excessive amount to the center of the histogram.

Despite the removal of these rays, the ray-tracing results still had a large peak

---

[3]As mentioned on page 174, the assumption that the radiation is well-described by ray optics is questionable.

[4]In order to simulate a three dimensional cone in a two-dimensional simulation, the rays were weighted by $(R_{\rm in}/R_{\rm out})$, where $R_{\rm in}$ is the distance from the center axis the ray enters, and $R_{\rm out}$ is the distance from the center axis when the ray exits. This accounted for the fact that a differential annulus of radius r and thickness dr has area $A = (2\pi r\, dr)$.

in the ray-density at the center of the exit, in addition to a small, relatively constant background density which had roughly one-third the power of the peak; see Figure 4.3. This created a dilemma. Since the cone has a circular exit, and both the input rays and the cone are rotationally symmetric about the central axis of the cone, it is valid to make the assumption that the field at the exit of the cone is only a function of radius from the central axis. This means that the field which the sample sees varies as a function of vertical distance along the edge of the sample. (Vertical here means perpendicular to the growth direction *and* to the propagation direction.) On the other hand, the external field which is an input to the theory of Załużny is (incorrectly) assumed to be constant in the vertical direction. It was unclear how to proceed. Can the correct result be approximated by doing several waveguide simulations at different vertical positions in the sample with different external fields? This is a question which might be better answered by a full physical optics analysis of the cone and the sample. At the very least, the theory of Załużny would need to be extended for the case of inhomogeneous vertical excitation, which could be exceedingly difficult, if it is even possible.

In the face of these challenges and complexities, I decided to use a constant electric field as the input external field for the simulation. A calculation was performed which showed that the peak in the ray-tracing output only contained approximately 8% of the total energy.[5] Although ignoring the peak is a fairly drastic assumption, it does remove many of the earlier assumptions in the ray-tracing, as well as simplifying the calculations and improving their generality.

In order to honestly and consistently address these problems and do a more be-

[5]This calculation took into account the scaling to convert the one-dimensional ray-tracing output into the correct two-dimensional output of the cone.

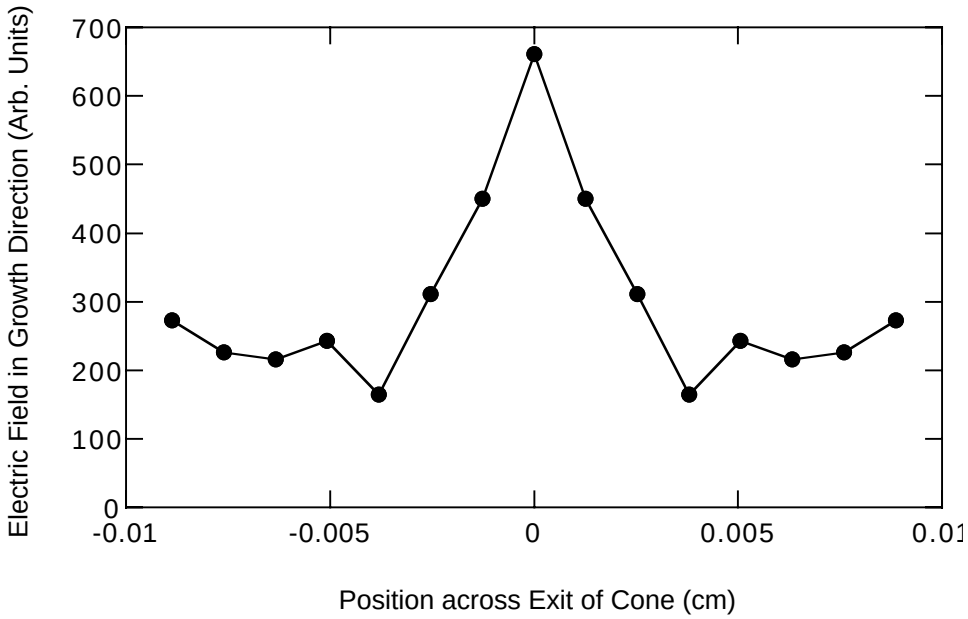

Figure 4.3: The results of the ray-tracing simulation for the electric field at the exit to the cone, along the a line which passes through the central axis of the cone. The calculation used histogram bins which were half of a wavelength wide, and excluded the 5.7% of rays which exited the cone too close to the central axis. (The dots are the calculated values, and the lines are a guide to the eye.) This result was approximated by a constant electric field.

lievable measurement and analysis, the most realistic plan is to modify the experimental apparatus to simplify the external electric field shape into the sample. If the field were a constant across the edge of the sample, then issues about inhomogeneity in the vertical direction, as well as uncertainty about the exact shape of the field would be solved. This field could be achieved by using the radiation from the FT-IR without focusing, in order to create a field which is spatially homogeneous on length scales comparable to the sample. (However, this could lead to problems of low signal due to the lack of concentration of the radiation.)

4.2 Fitting Procedure

The procedure for finding a fit to the linear absorption data was rather complicated. There were two main steps to fitting the data: first the waveguide simulation was inverted to find the well parameters from the measured waveguide data, and then the waveguide data was calculated from the well parameters. (By inverted, I mean that the simulation had to be used in such a way that the input parameters are calculated from the output parameters.) The procedure is summarized visually in Figure 4.4. In this section, I introduce the terminology that "well absorption" refers to the absorption of the well without any effects due to the waveguide, and "waveguide absorption" refers to the absorption of the well including the waveguide effects.

The waveguide simulation takes the parameters of the absorption of the well, and calculates the absorption of the waveguide. Therefore, before I could use the waveguide simulation, I needed to know the parameters of the absorption of the well. The linewidth is the most important and fundamental of these parameters: it determines the effect of the waveguide corrections on all the other parameters. (This is simply because a narrow line will have a relatively higher peak conductivity and have stronger waveguide effects than a wider line.) Thus, the first stage of the fitting process was to calculate the linewidth of the absorption of the well from the experimental waveguide absorption data.

The first step of this stage was to calculate the function at each charge density which converts the linewidth of the well absorption to the linewidth of the waveguide absorption. The first box in Figure 4.4 consists of choosing several well absorption linewidths, and using them as input parameters to the waveguide simulation, using the best available well parameters. The output waveguide absorption plots are then

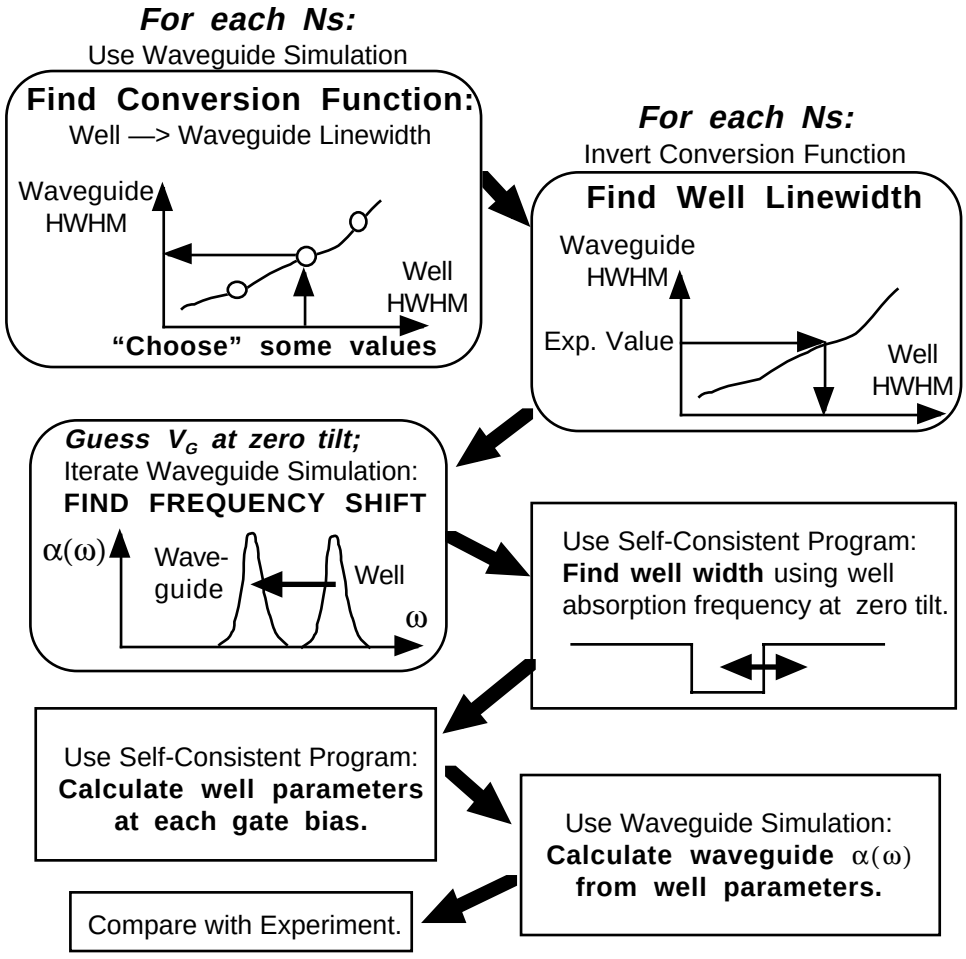

Figure 4.4: A visual representation of the fitting procedure using the waveguide simulation and the self-consistent simulation. The boxes with rounded corners are steps involved in numerically inverting the waveguide simulation to get the quantum well values from the waveguide data. Once the well values have been found, the boxes with sharp corners are the forward steps of actually simulating the sample. (See text for details.)

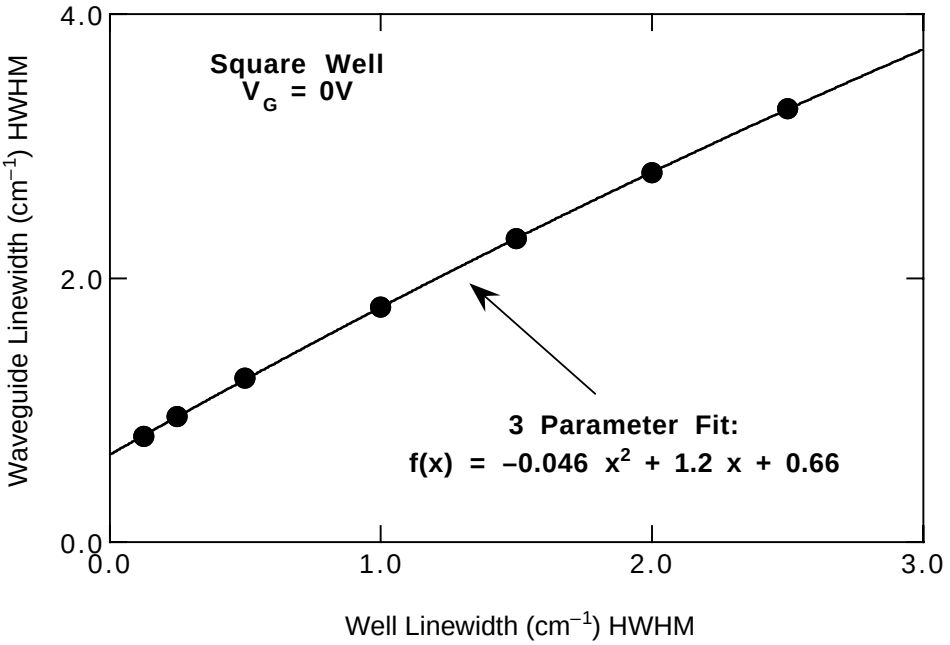

Figure 4.5: A sample linewidth model for the square well sample at $V_G = 0$. The model is generated by choosing arbitrary values for the linewidth of the well, and then using the waveguide simulation to calculate the corresponding linewidth of the waveguide. Then the results are plotted and fitted as shown. The fit is later inverted to calculate the well linewidth from the observed waveguide linewidth.

fit with a lorentzian to extract the waveguide linewidth. Then all these data points are plotted with waveguide linewidth as a function of well linewidth. (See Figure 4.5.) These data points are then fit with a function of the form:

$$y(x) = ax^2 + bx + c \tag{4.8}$$

where y is the waveguide linewidth, x is the well linewidth, and a, b, and c are fit parameters. This is repeated at each gate voltage (charge density and static field across the well) to form a "linewidth model" of the waveguide effects at each charge density. A sample linewidth model is shown in Figure 4.5.

The second box in Figure 4.4 simply inverts each of the linewidth models. This is done simply by solving Equation (4.8) for x as a function of y. Then, the experimentally observed waveguide linewidths at each gate voltage are used with the inverted functions to find the well linewidth.

In the third box, the frequency shift of the peak due to the waveguide effects are calculated for the gate bias where $E_{\rm dc} = 0$. (The gate voltage where $E_{\rm dc} = 0$ is guessed. After the entire process of generating a complete fit to the data is complete, the goodness of fit is examined and the whole process is repeated with a different guess of the gate voltage where $E_{\rm dc} = 0$. This whole process is repeated until the best fit to the data is found. The meaning of the static field across the well, $E_{\rm dc}$, is discussed in Appendix B.) Different well resonance frequencies are input into the waveguide simulation with the well parameters for the given gate voltage, until the calculated waveguide absorption frequency fits the observed waveguide absorption frequency. In practice, this is an easy step, because once the linewidth is determined, the shift of the peak from the well absorption to the waveguide absorption is just a constant shift for a given set of parameters.

At this point, the waveguide simulation has been inverted! Now the process can continue in a more straightforward method. In the fourth box, the self-consistent program (not the waveguide simulation) is used to find the well width which gives the correct depolarization shifted *well* absorption frequency at $E_{\rm dc} = 0$. This is done by varying the well width until the resonance frequency agrees with the well absorption frequency obtained in the previous (third) box.

In the fifth box, the well parameters at each gate voltage are calculated using the self-consistent program. This is done using the well width obtained in the previous (fourth) box, and the static field across the well is calculated from the gate voltage

and assumption of the gate voltage where $E_{\rm dc} = 0$.

In the sixth box, the waveguide absorption is calculated from the well parameters at each gate voltage, using the waveguide simulation. (The input parameters for the waveguide simulation are the results of the previous step.)

The final box is to compare the waveguide absorption results with the experimental data. The comparison is studied to see if it could be improved by choosing a different gate voltage where the well has $E_{\rm dc} = 0$. This can often be done intelligently rather than randomly by inspection of the data and the fits.

4.3 Results of Square Well Waveguide Modelling

The waveguide theory was used to model the linear absorption of the square well, using a constant external electric field. In general, with a constant external field, the waveguide effects broadened the peak and shifted it to lower frequencies. However, the general shape of the peak remained Lorentzian except at very high charge densities. At very high charge densities (larger than in the square well sample), the peak began to develop internal structure and become less Lorentzian. (See Figure F.1.) However, if the external field was not constant, the peak developed internal structure at much lower charge densities.

Figure 4.6 shows the absorption frequency as a function of gate voltage. The goodness of fit is comparable to the fit from linear perturbation theory, Figure 3.4. The parameters used to generate these fits are shown in Table 4.1. Note that, using the constant external field, the waveguide's resonance frequency is lower than the well's resonance frequency. This is caused by the waveguide modes at the well's resonance frequency being excluded more strongly from the well than the modes at slightly

Parameter	**"Expected"**	**Self-Consistent**
Well Width	40.0 nm	38.5 nm
Charge Density at $V_G = 0$	$\sim 1.0 \times 10^{11}$ e/cm^2	9.70×10^{10} e/cm^2
V_G at $E_{\rm dc} = 0$	0 volts	-0.5 volts
External Field Shape	Unknown	Constant (Flat)

Table 4.1: The "expected" parameters of the square well, and the input values for the self-consistent program which was used to generate the input data for the waveguide simulation. These values were chosen to give the best waveguide output fit, not the best linear perturbative fit.

lower frequencies.

The main goal of using the waveguide theory was to improve the fit of the integrated absorption. Although the waveguide theory improves the fit somewhat, it does not fit the data well; see Figure 4.7. The waveguide theory reduces the integrated absorption at higher charge densities, due to the modification of coupling between the well and the perturbed modes. The simulation follows a general trend in the data, but it does not capture the large bump at moderate charge densities, nor does it really match the qualitative shape of the experimental data well. (Note that the excellent fit at low charge densities is not chance; the optical data were used to calculate the residual charge density. See Appendix E.)

One very interesting outcome of the waveguide analysis was the calculated linewidth of the well. Figure 4.8 shows the calculated well linewidth, in addition to the calculated waveguide linewidth and experimentally measured waveguide linewidth. The waveguide simulation (with constant external field) broadens the peak appreciably; therefore, the simulations indicate that the absorption peak of the well is significantly

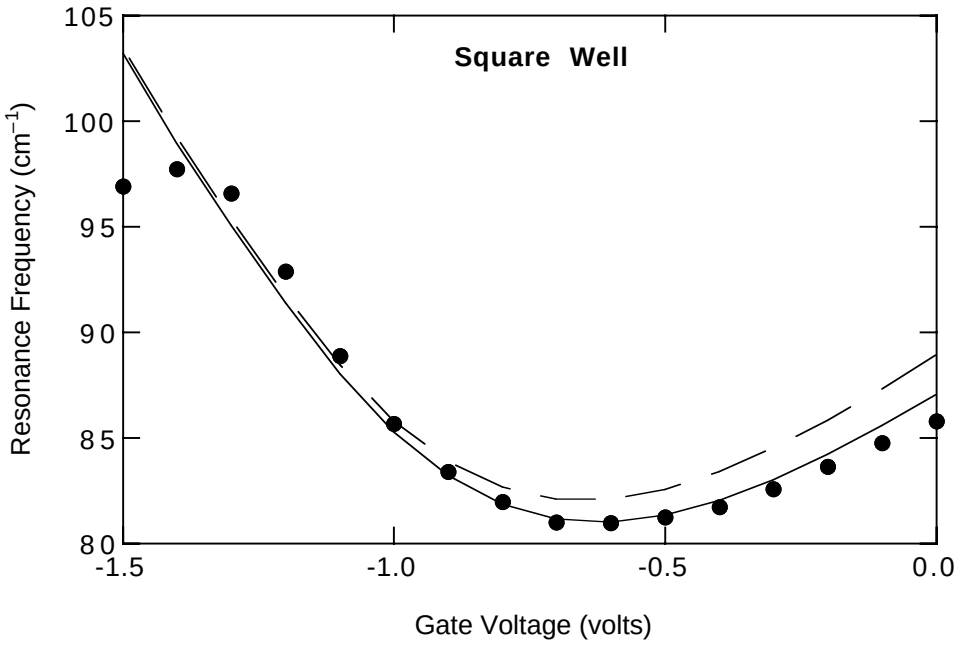

Figure 4.6: The square well resonance frequency as a function of gate voltage, with a fit using the waveguide theory. The dots are experimentally measured values, the solid line is the waveguide theory, and the dashed line is the prediction of linear perturbation theory which, when used as an input in the waveguide theory, gives the best agreement between the waveguide theory and experiment.

narrower than the waveguide absorption.

4.4 Double Square Well Waveguide Modelling

In general, the fits for the double square well using the waveguide theory look very similar to the fits using the linear perturbative theory. The same data can be fit equally well by two different theories because of the uncertainty in the well width and the static field. By independently fitting the data using the two theories, each theory can achieve a moderately good fit using reasonable parameters. The exception is the

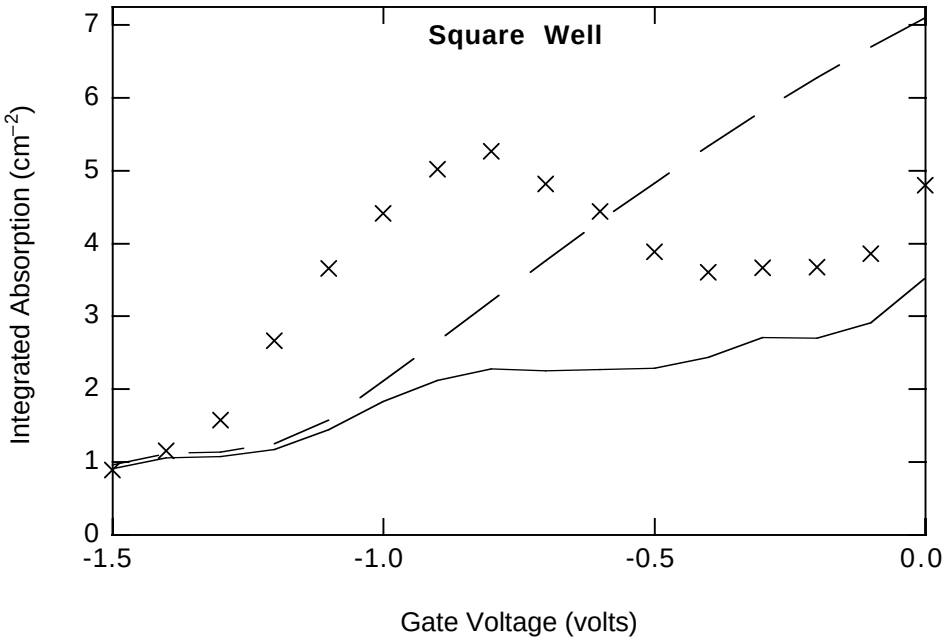

Figure 4.7: The square well integrated absorption as a function of gate voltage. The "X"s are experimentally measured values, the solid line is the waveguide theory, and the dashed line is the prediction of linear perturbation theory which, when used as an input in the waveguide theory, gives the best agreement between the waveguide theory and experiment.

integrated absorption; see below.

Figure 4.9 shows the waveguide theory fit to the resonance frequency. The parameters used in the simulation are shown in Table 4.2. The fit is not greatly improved by the waveguide theory.

Figure 4.10 shows the integrated absorption fit to the double square well. This simulation is interesting, because once the linewidth and frequency have been fit, there are no free parameters left to vary the fit for the integrated absorption. And, interestingly enough, the experimental data are fit better by the linear perturbative theory than by the waveguide theory. This indicates that either the waveguide theory is incorrect or the external electric field shape which was input into the waveguide

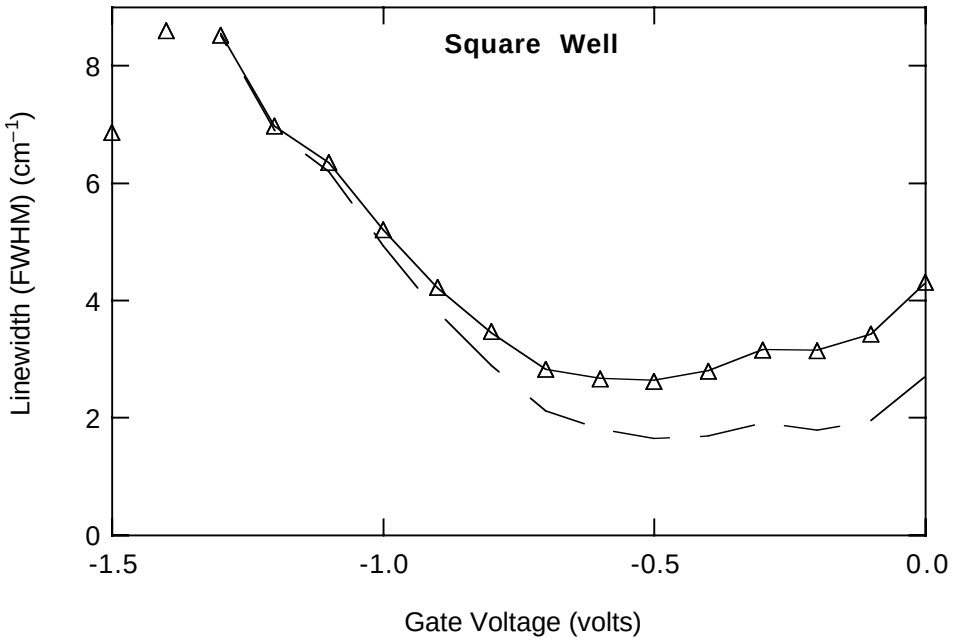

Figure 4.8: The square well absorption linewidth as a function of gate voltage. The open triangles are experimentally measured values, the solid line is the waveguide theory, and the dashed line is the prediction of linear perturbation theory which, when used as an input in the waveguide theory, gives the best agreement between the waveguide theory and experiment. Note that the waveguide theory broadens the line significantly, indicating that the absorption of the well is much narrower than observed.

Parameter	**"Expected"**	**Self-Consistent**
Well-Barrier-Well Width	8.5-2.5-7.3 nm	8.755-2.575-7.519 nm (Growth $\times$ 1.03)
Charge Density at $V_G = 0$	$\sim 2.0 \times 10^{11}$ e/cm^2	2.58×10^{11} e/cm^2
V_G at $E_{\rm dc} = 0$	0 volts	-0.6 volts
External Field Shape	Unknown	Constant (Flat)

Table 4.2: The "expected" parameters of DSQC, and the input values for the self-consistent program which was used to generate the input data for the waveguide simulation. These values were chosen to give the best waveguide output fit, not the best linear perturbative fit.

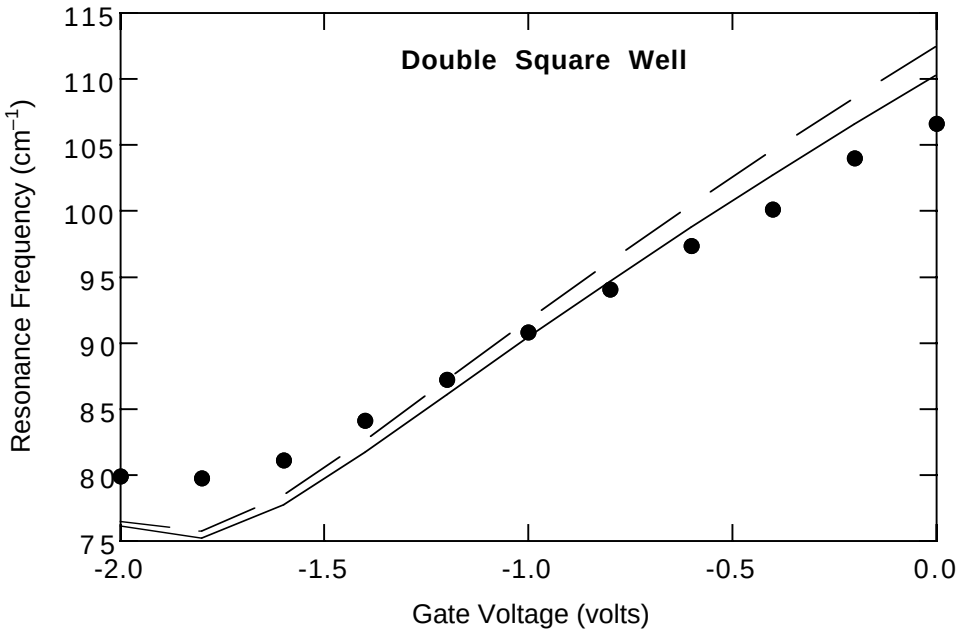

Figure 4.9: Double square well resonance frequency as a function of gate voltage. The dots are experimentally measured values, the solid line is the waveguide theory, and the dashed line is the prediction of linear perturbation theory which, when used as an input in the waveguide theory, gives the best agreement between the waveguide theory and experiment. The waveguide simulation does not significantly improve the goodness of fit compared to the linear perturbative theory.

theory is incorrect. This behavior also indicates that the behavior of the double square well and the square well may be qualitatively different.

Finally, the linewidth fits for the double square well using the waveguide theory are shown in Figure 4.11. The results are similar to the square well (but not quite as striking).

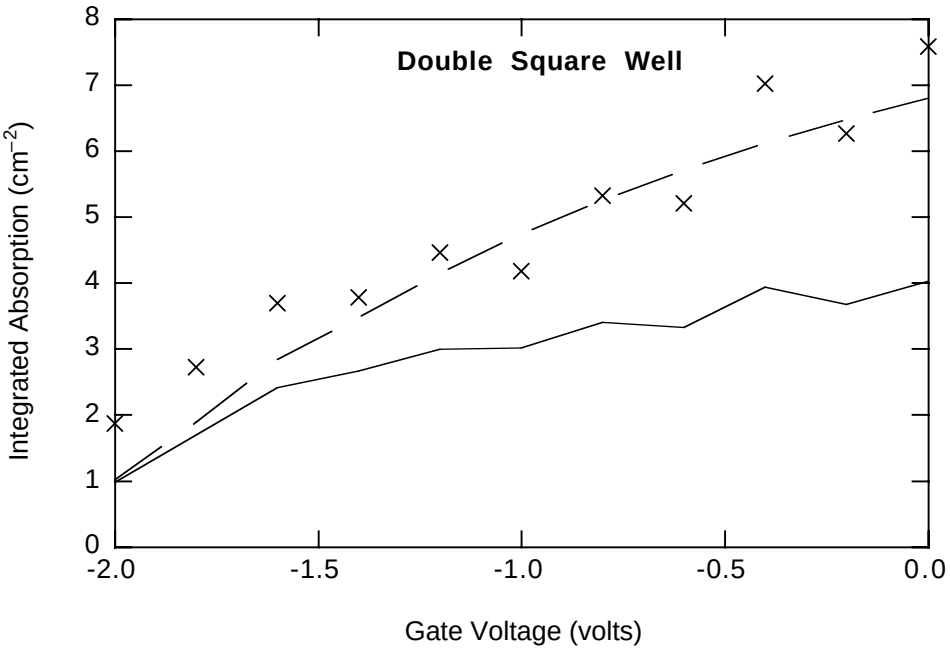

Figure 4.10: Double square well integrated absorption as a function of gate voltage. The "X"s are experimentally measured values, the solid line is the waveguide theory, and the dashed line is the prediction of linear perturbation theory. The linear perturbative theory fits the data much better than the waveguide theory.

4.5 Conclusions on Waveguide Effects

The edge-coupling geometry has strengths and weaknesses. The strengths are that it is experimentally straightforward to set up and understand, and it gives a large absorption due to the long path length the radiation travels through the sample. In addition, by using samples of different lengths, the amount of absorption can be varied. The weaknesses are that it can be difficult to model the modes and estimate the shape of the electric field which enters the sample. In addition, experimentally it can be challenging to prevent light from leaking around the sample.

It is difficult to draw conclusions about the waveguide theory, other than the fact

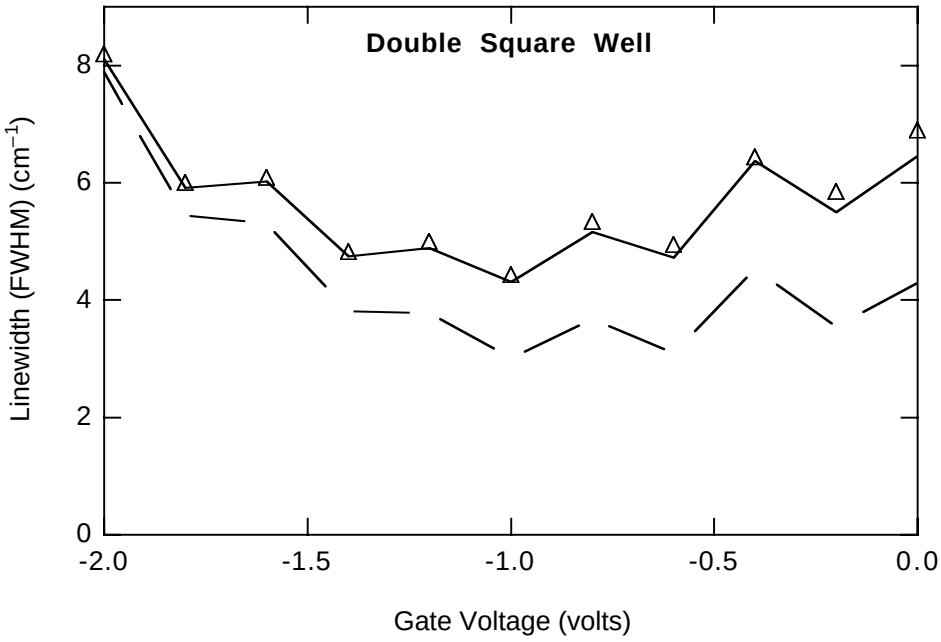

Figure 4.11: Double square well linewidth as a function of gate voltage. The open triangles are experimentally measured values, the solid line is the waveguide theory, and the dashed line is the prediction of linear perturbation theory. As with the square well, this calculation indicates that the actual absorption linewidth of the well is much narrower than the linewidth of the waveguide.

that the external electric field shape is a critical input parameter. My calculations using the waveguide theory were unable to significantly improve the fits to the experimental data. However, my lack of knowledge of the external electric field shape prevents me from supplying the theory with the correct input parameters, and therefore I cannot conclude that the theory is incorrect. Careful measurement or control of the external field shape would be the next step in addressing these issues.

Chapter 5

Nonlinear Measurements

The nonlinear absorption measurements were performed in an edge-coupling transmission geometry, with attenuators to vary the intensity at the sample and a detector to measure the transmission through the sample. In this Chapter I will describe the main parts of this setup and the procedures which were used. Details about the setup and the procedures can be found in Appendix G.

5.1 Optical Setup

The source of intense, tunable far-infrared radiation (FIR) was the University of California Santa Barbara Free-Electron Laser, or FEL. For our purposes, the FEL produces roughly 500 watts near 80 cm^{-1} and is continuously tunable. The radiation consists of pulses which were $2.5\ \mu s$ in duration, with a repetition rate of approximately 1 Hz. For a discussion of the FEL itself, see Chris Felix's thesis, Reference [5].

The entire optical setup was enclosed inside a purged glove bag. The bag was continuously purged with dry nitrogen gas to remove water vapor from the beam

path. The bag eliminated almost *all* water vapor absorption, allowing us to measure the absorption of our samples as a function of frequency, which otherwise would have been impossible.[1] The glove bag is described in detail in Appendix G.

The sample was inside a continuous-flow cryostat with a sample holder designed to prevent most of the FIR from leaking around the sample. A piece of plexiglass was used to prevent the FIR from passing around the sides of the sample, and aluminum foil was used to prevent radiation from going above or below the sample. The setup is shown in Figure 5.1, and is described more fully in Appendix G.

Figure 5.2 shows the nonlinear absorption setup. A small amount of FIR was diverted by a beamsplitter into a reference detector which was used to normalize intensity fluctuations in the FEL output. The FIR passed through a set of attenuators, and then was focused onto the edge of the sample. An FIR polarizer just before the sample ensured that the radiation entering the sample was polarized parallel to the table (and therefore the growth direction of the sample). After the radiation exited the sample and cryostat, it was recollimated by a focusing mirror and passed through a second set of attenuators which was used to keep the intensity at the bolometer roughly constant. Then the radiation was focused into the bolometer, which was used to measure the transmission of the sample at different gate voltages. Additional details about the setup and its operation are discussed in Appendix G.

---

[1]The absorption measurements include frequencies where normal air (with water vapor) is practically opaque. Even if absorption measurements had been possible, the intensity inside the sample would have been unknown without the glove bag.

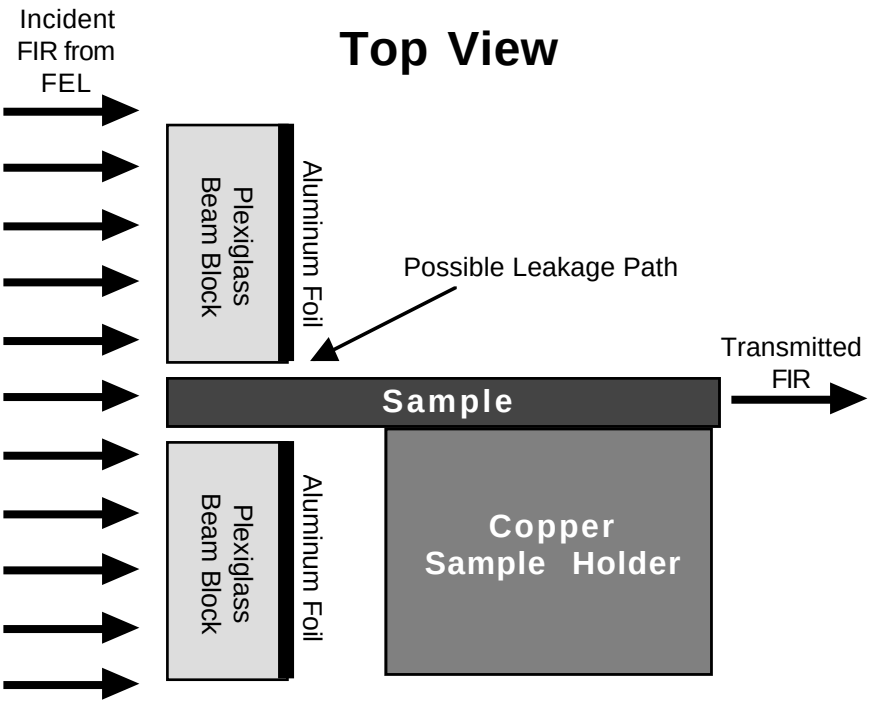

Figure 5.1: A schematic top view of the sample holder used for nonlinear experiments. The sample is inserted into a plexiglass piece which has a slit cut in it which is slightly wider than the size of the sample. The plexiglass acts as a beam block to prevent FIR from going around the sample. Aluminum foil is glued to the back of the plexiglass to ensure that no radiation can pass through the block. The sample mounting piece blocks leakage around the back of the sample, although a small amount of leakage around the front of the sample is possible. The aluminum foil on the *front* of the plexiglass, which blocks FIR from going above or below the sample, is not shown.

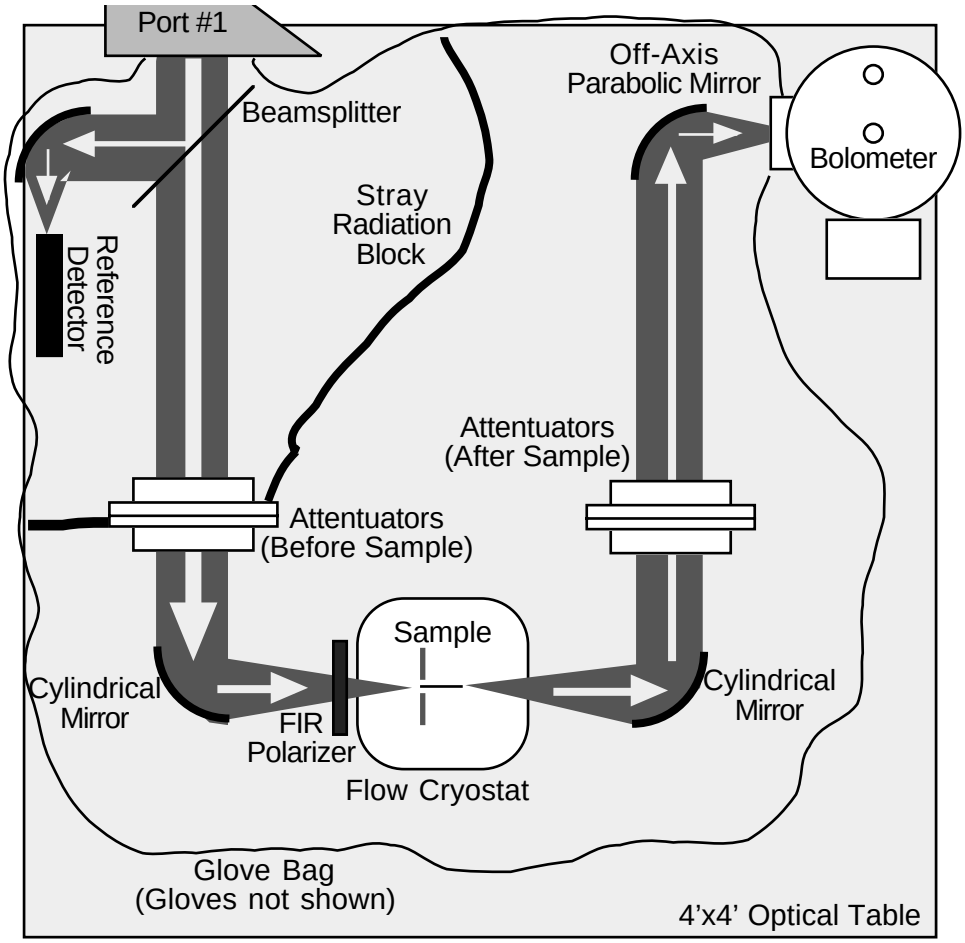

Figure 5.2: The experimental setup used to measure nonlinear absorption. The FIR emerges from the port, passes through a beamsplitter which splits a small amount into a reference detector, then passes through a set of attenuators, and next passes through a polarizer as it is focused onto the sample edge by a cylindrical mirror. After passing through the sample, it is recollimated by another cylindrical mirror, then passes through another set of attenuators before being focused into the bolometer.

5.2 Procedure

The goal was to measure the absorption of the sample over a range of frequencies and intensities. After cooling the sample and sealing the glove bag, the attenuators were set to provide the minimum intensity[2] at the sample, and the transmission of the sample was recorded with electrons in the well and without electrons in the well. Then the attenuators were moved and the absorption measurement was repeated at different intensities. (The sample was allowed to come to equilibrium after the gate voltage was changed; this is described in Appendix J.) After the absorption had been measured at a wide range of intensities, the frequency of the FEL was changed by a small amount, and then the above measurements were repeated until the absorption had been measured at a large number of frequencies and intensities. (The measurements were done at approximately ten to fifteen frequencies and intensities, leading to approximately one to two hundred absorption measurements.) This procedure is described fully in Appendix G.

The intensity was varied while keeping the total attenuation in the beam constant, by removing attenuation from before the sample and adding it after. This kept the total energy at the bolometer roughly constant, while the intensity at the sample could be varied by five orders of magnitude. This ensured that the bolometer was operating in the linear regime at all intensities. The attenuators were moved by hand using gloves built into the glove bag.

---

[2]The minimum intensity was the lowest intensity where the signal to noise ratio of the bolometer was adequate. Due to the measurement technique, *all* the measurements were done at this signal to noise ratio, not just the low-intensity measurements.

5.3 Electronics

The experiment was controlled by a single Macintosh computer, and all the data were acquired and stored by this computer and the supporting electronics. The electronics setup is shown in Figure 5.3 and described in detail in Appendix G. The detector signals were measured using digitizing oscilloscopes, and the gate voltage was controlled by the computer.

5.4 Intensity Measurement

The accurate measurement of the intensity inside the sample was very important to allow combination of data from different frequencies, and to allow estimation of the relaxation times of the electrons. At the end of the measurements, the intensity outside the cryostat was measured directly, and the transmission into the sample was estimated to calculate the intensity inside the sample. These values were used to calibrate the signal from the bolometer, which was then used to calculate the intensities for all the data. The details of this procedure are in Appendix H.

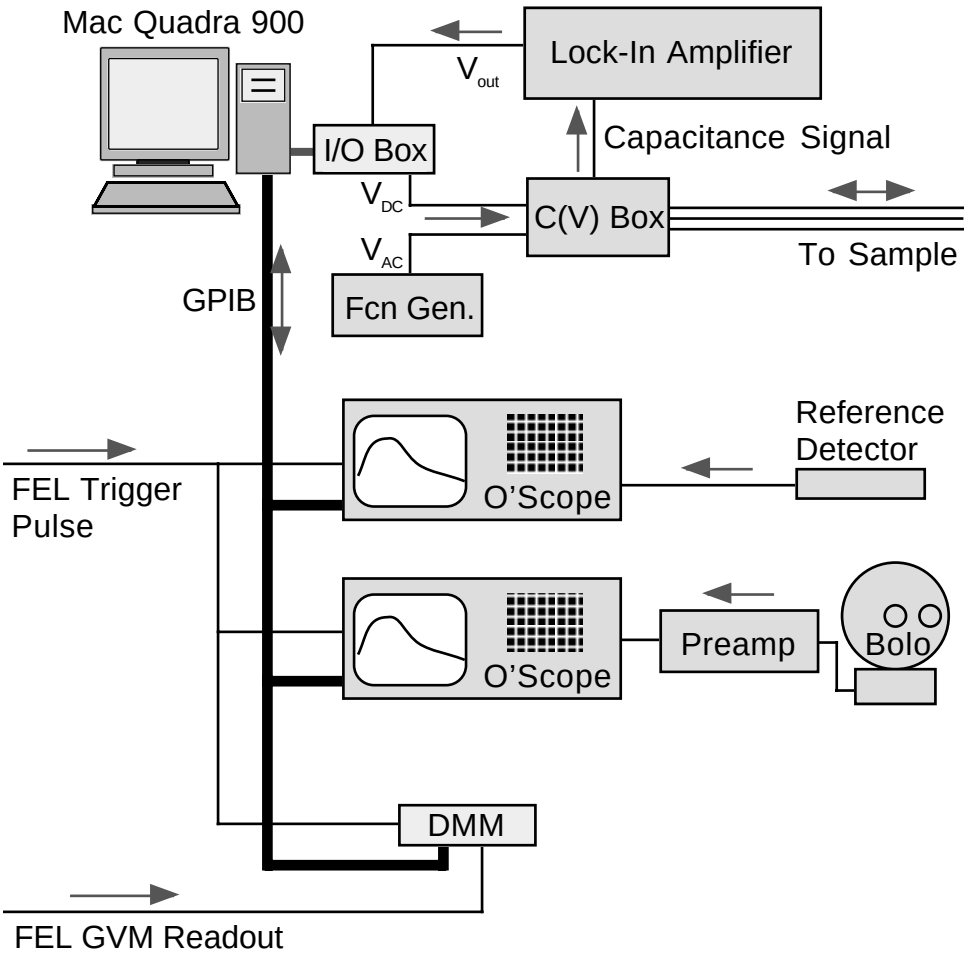

Figure 5.3: A schematic diagram of the electronics used in nonlinear experiments. The arrows show the direction of data flow. The I/O box is connected to the Lab-NB card, which allowed the computer to source and measure voltages. The DMM is the digital multimeter which was used to measure the terminal voltage.

Chapter 6

Nonlinear Results

The main goal of this work was to measure the nonlinear absorption of a wide quantum well system (where the depolarization shift cannot be ignored). Since I was uncertain how the data would look, I simply measured the absorption at every point on a grid of frequencies and intensities.

In this Chapter, I will present the data as they were measured, with absorption as a function of intensity at many different frequencies. Then I will briefly describe how the data were interpolated to be plotted with absorption as a function of frequency, at a constant intensity. Next, I will present and describe a theory by M. Załużny of nonlinear intersubband absorption which includes the depolarization shift. Then I will compare the measured data with this theory, describe in detail what happens physically in the sample, and discuss the similarities and differences between theory and experiment.

6.1 Nonlinear Absorption Data

The absorption was measured while varying the intensity at constant frequency, and then changing to a new frequency and varying the intensity again, and so on... These data are plotted with absorption as a function of intensity at constant frequency for each curve. The data were later interpolated into absorption as a function of frequency, with intensity constant for each curve.

6.1.1 Raw Data

Figures 6.1, 6.2, and 6.3 show the absorption data as a function of intensity for the sample with a gate voltage $V_G = 0$ V, and a charge density of 1.39×10^{11} e/cm^2. These data are tabulated in numerical form in Appendix K.

At the lowest frequencies measured, there is not too much happening. The absorption is rather small (although it does start to increase at the highest intensities), since the measurement frequency is well below the linear intersubband absorption frequency. As the frequency increases, the shape changes, and around 76.9 cm^{-1}, it is clear that the absorption is small at low intensity and then *increases* with intensity, finally decreasing again at the highest intensities. Then, as the frequency is increased still further, the intensity at which the absorption increases moves to the left (decreases). At this point (80.9 cm^{-1}) the absorption is large at low intensity and then decreases at higher intensity. Finally, as the frequency is increased still more, the absorption retains its shape but decreases in magnitude as the FEL frequency is now well above the linear intersubband absorption frequency.

There is some scatter in this data. As one example, in Figure 6.4 at 77.7 cm^{-1}, between 10^{-2} W/cm^2 and 10^{-1} W/cm^2, there is a sharp spike (in a relative valley) in

the observed attenuation coefficient. I believe that this is simply experimental noise; such sharp structures were not observed repeatedly in the same place. This noise could have been caused by many possible effects, some of which are: variations in humidity in the glove bag, frequency drift of the FEL, some sort of short-term drift in the detectors, thermal expansion changing the optical path slightly, electrical noise in the detector signals, changes in the mode-structure of the FEL beam, a scattering of FIR into the detectors from some object moving around the table, or long time constants in the sample. In general, the noise in the data is less noticeable after interpolation, since the noise at different frequencies is not correlated.

Figures 6.4, 6.5, and 6.6 show the absorption data as a function of intensity for the sample with a gate voltage $V_G = -0.5$ V, and a charge density of 5.69×10^{10} e/cm^2. These data are tabulated in Appendix K.

There is slightly more scatter in these data. However, they show essentially the same behavior as for the $V_G = 0$ case. As the frequency increases, the absorption increases at high intensity. As the frequency is increased more, the absorption increase occurs at lower intensities, until around 80.6 cm^{-1} the absorption is a maximum at low intensity and decreases with intensity. (This assumes that the spike in the data for 80.6 cm^{-1} at a few tens of mW/cm^2 is just experimental noise.) Then, increasing the frequency further simply results in a decrease in the magnitude of the absorption.

6.1.2 Interpolation Procedure

For comparison with theory and general ease of examination, I wanted to plot the absorption values as a function of frequency at a constant intensity. However, as can be seen from the data in the previous section, each of the measured points at different frequencies were measured at a different intensity. This is mainly due to the fact

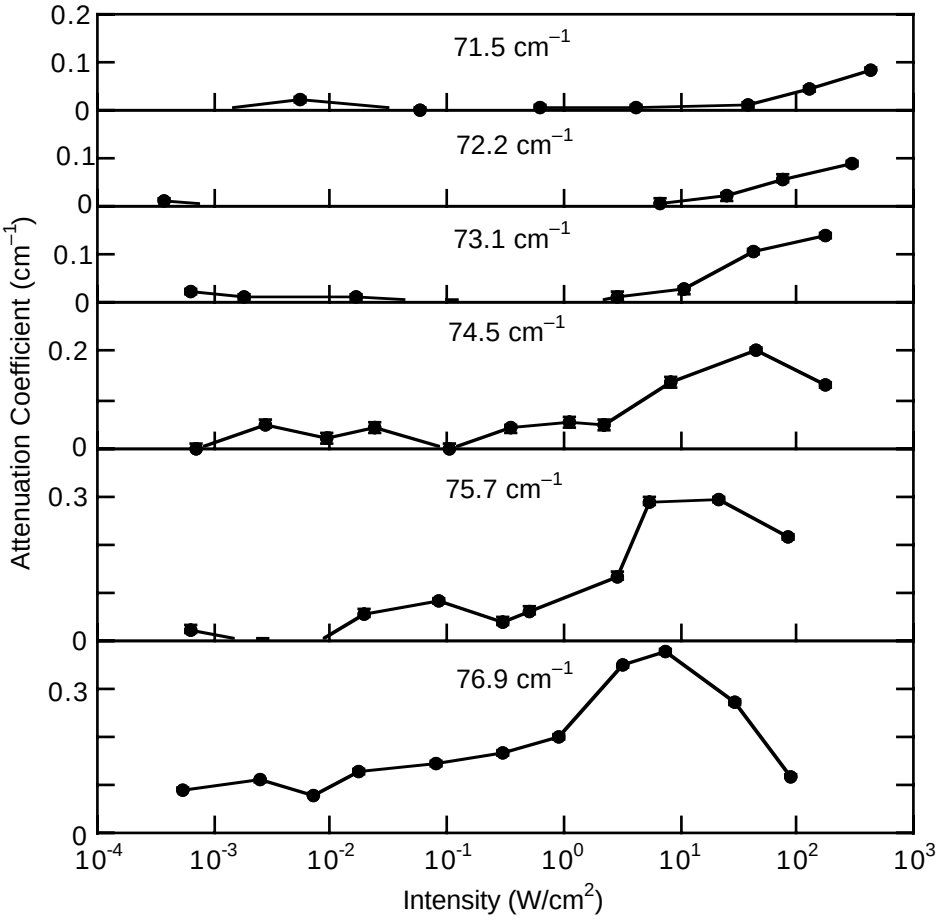

Figure 6.1: The absorption as a function of intensity for the square well sample at $V_G = 0$ V, for the lowest frequencies measured. The lines simply connect the points and provide a guide for the eye. The error bars shown are the standard error calculated from the variation in transmission values.

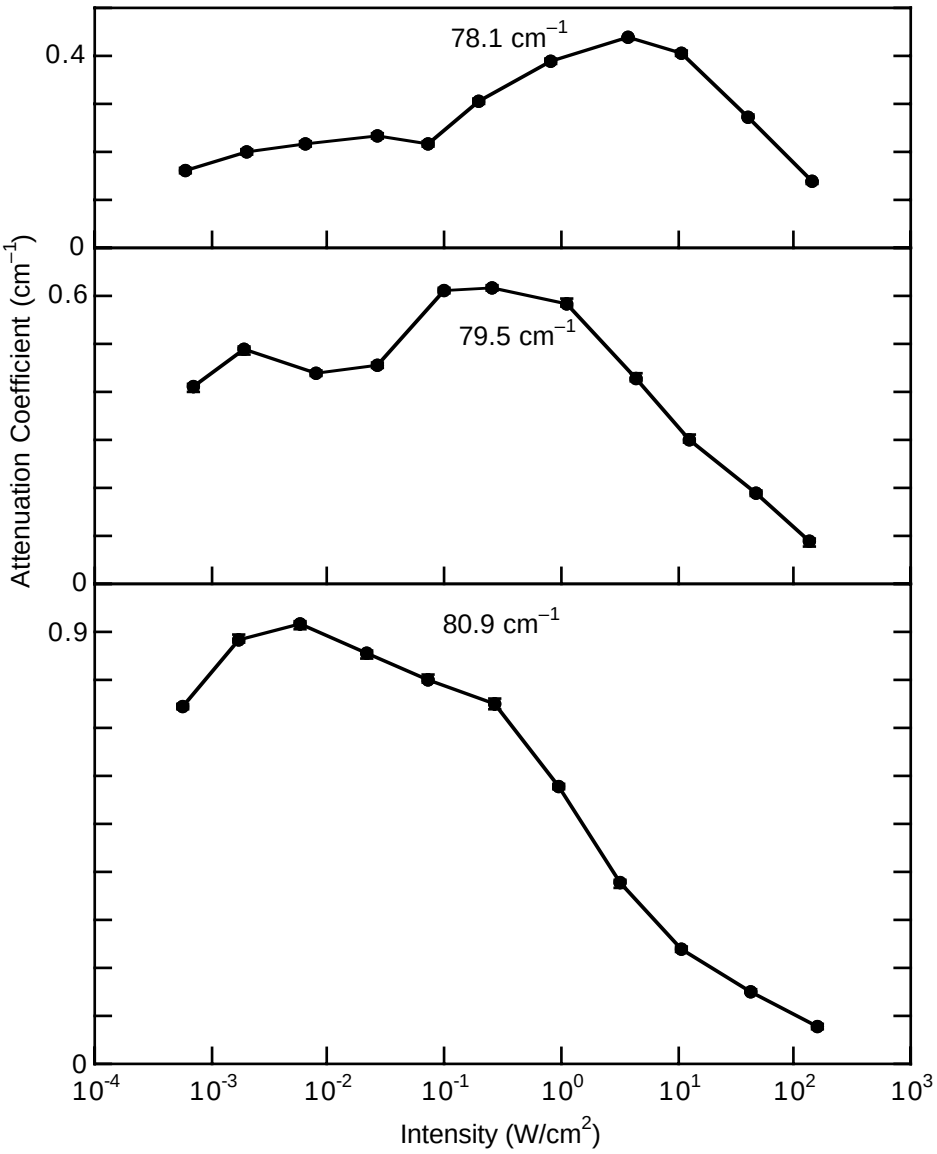

Figure 6.2: The absorption as a function of intensity for the square well sample at $V_G = 0$ V, for the middle frequencies measured.

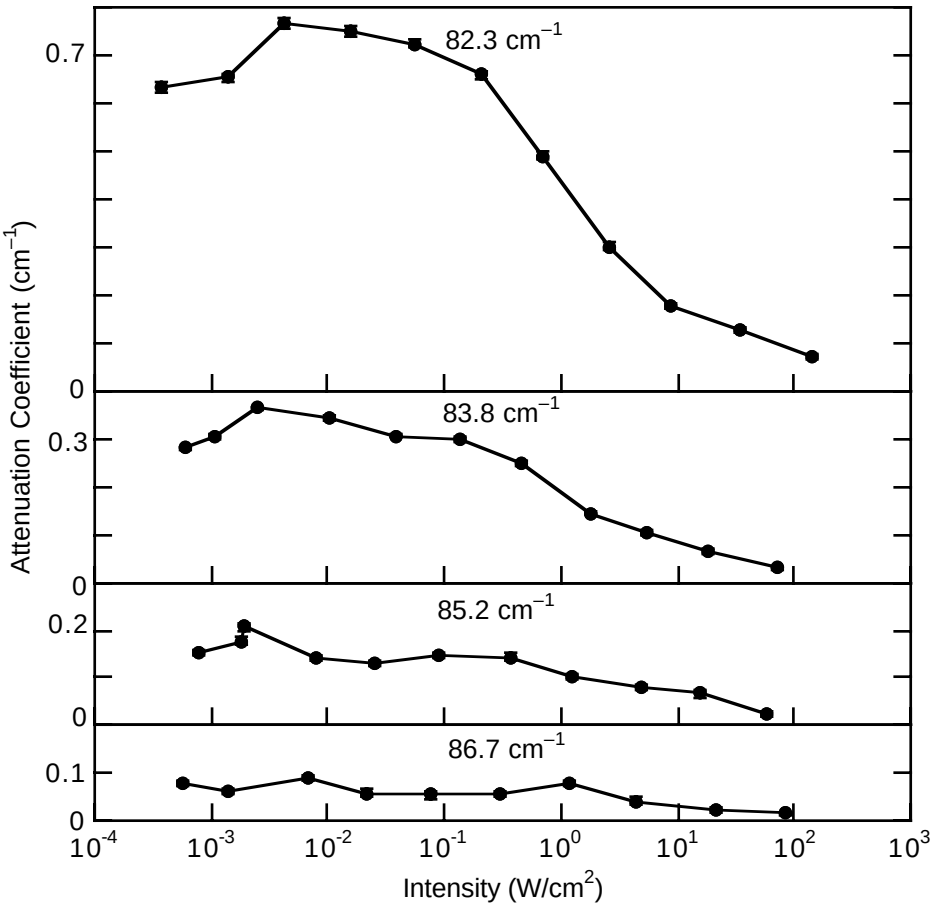

Figure 6.3: The absorption as a function of intensity for the square well sample at $V_G = 0$ V, at the highest frequencies measured.

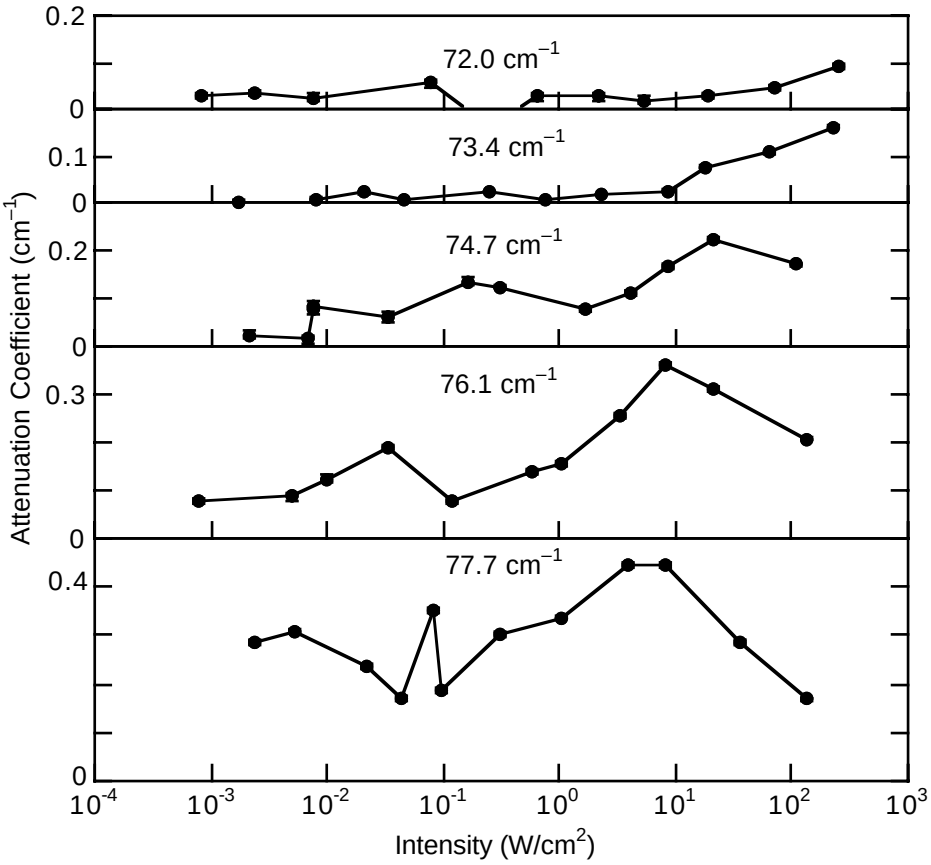

Figure 6.4: The absorption as a function of intensity for the square well sample at $V_G = -0.5$ V, for the lowest frequencies measured.

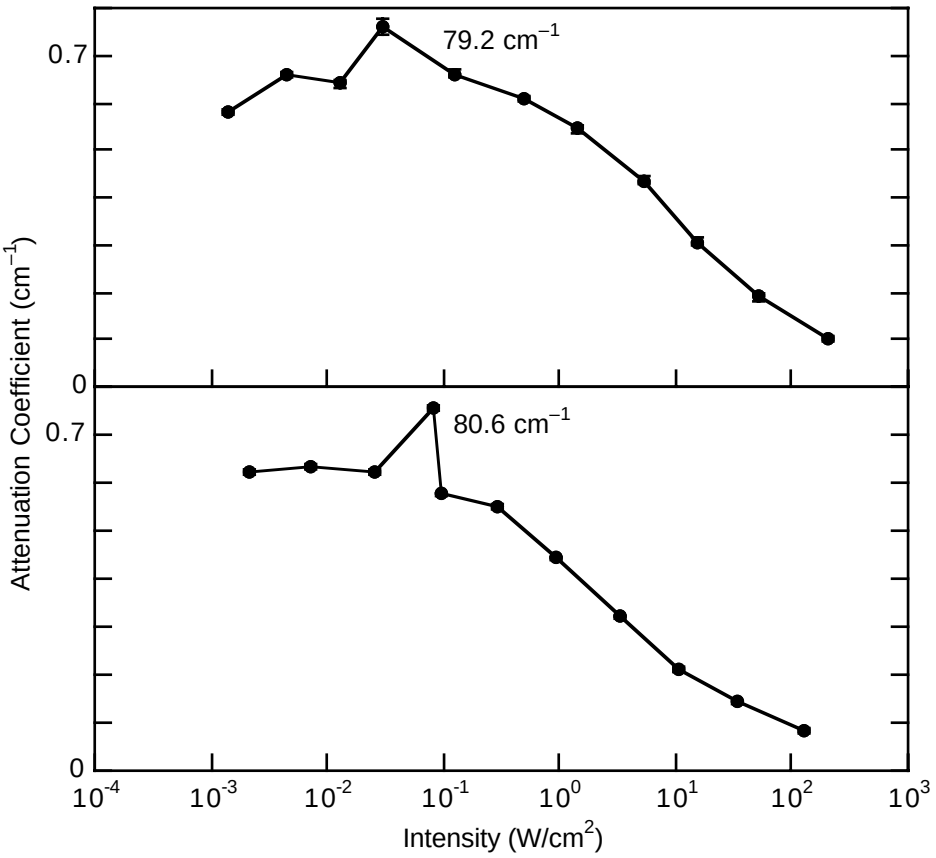

Figure 6.5: The absorption as a function of intensity for the square well sample at $V_G = -0.5$ V, for the middle frequencies measured.

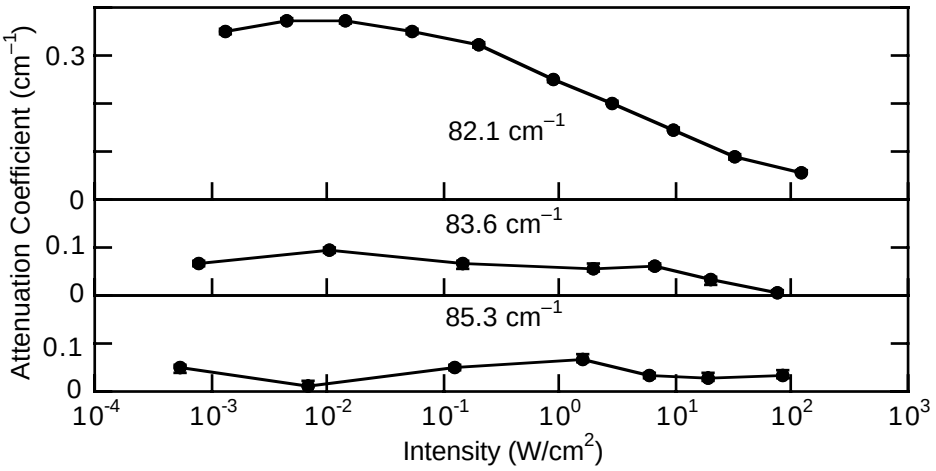

Figure 6.6: The absorption as a function of intensity for the square well sample at $V_G = -0.5$ V, at the highest frequencies measured.

that the attenuators were frequency dependent, which meant that the each attenuator changed the intensity by a different amount at different frequencies. In addition, the FEL needed to be re-tuned to change the frequency, which changed the intensity of the FEL output.

In order to plot the absorption values at different frequencies at the same intensity, the data from the previous section were interpolated. The treatment of the endpoints was found to be most sensible if the requirement was imposed that no values would be interpolated beyond the measured range. This means that, at very low and very high intensities, not every frequency will contribute a point on the absorption vs frequency plots. (The interpolation itself was a simple linear interpolation.)

The interpolation procedure started with the choice of several intensities at which absorption plots would be created. Then the absorption as a function of intensity plots were numerically interpolated to obtain a set of absorption values at each of the chosen intensities. Then the resulting plots were examined to see if the resulting

interpolated plots could be improved by choosing different intensity values.

Note that the choice of the intensity values at which to interpolate and plot was arbitrary. Experimentally, however, it was important that the measured absorption values were not so closely spaced in intensity that multiple points were interpolated between the same two experimentally measured points. In practice, this was not a problem: the usual choice was to choose to display the data at factors of ten in intensity, and each attenuator changed the intensity by a factor of roughly three. The only remaining concern was that the intensity values be chosen such that they are far apart enough to make the different plots look different, but close together enough so that the trends are apparent from one plot to the next.

6.1.3 Absorption as a Function of Frequency

The interpolated results for $V_G = 0$ V are shown in Figure 6.7. The results are interpolated at 0.01 W/cm^2, 1 W/cm^2, 10 W/cm^2, and 80 W/cm^2. As the intensity increases in Figure 6.7, the height of the peak decreases, the peak broadens, and the center of the peak shifts to lower frequency. This behavior cannot be explained by the standard saturation theory. Although the peak might broaden somewhat due to electron heating effects in a standard saturation experiment, the large frequency shift of the center of peak would not occur. I will discuss later that the "broadening" of the peak is not a heating effect.

The interpolated results for $V_G = -0.5$ V are shown in Figure 6.8. The results are interpolated at 0.01 W/cm^2, 1 W/cm^2, 10 W/cm^2, and 100 W/cm^2. The behavior of the peak is similar to Figure 6.7, and the data appear to be slightly less noisy, in that the peak shapes are slightly smoother.

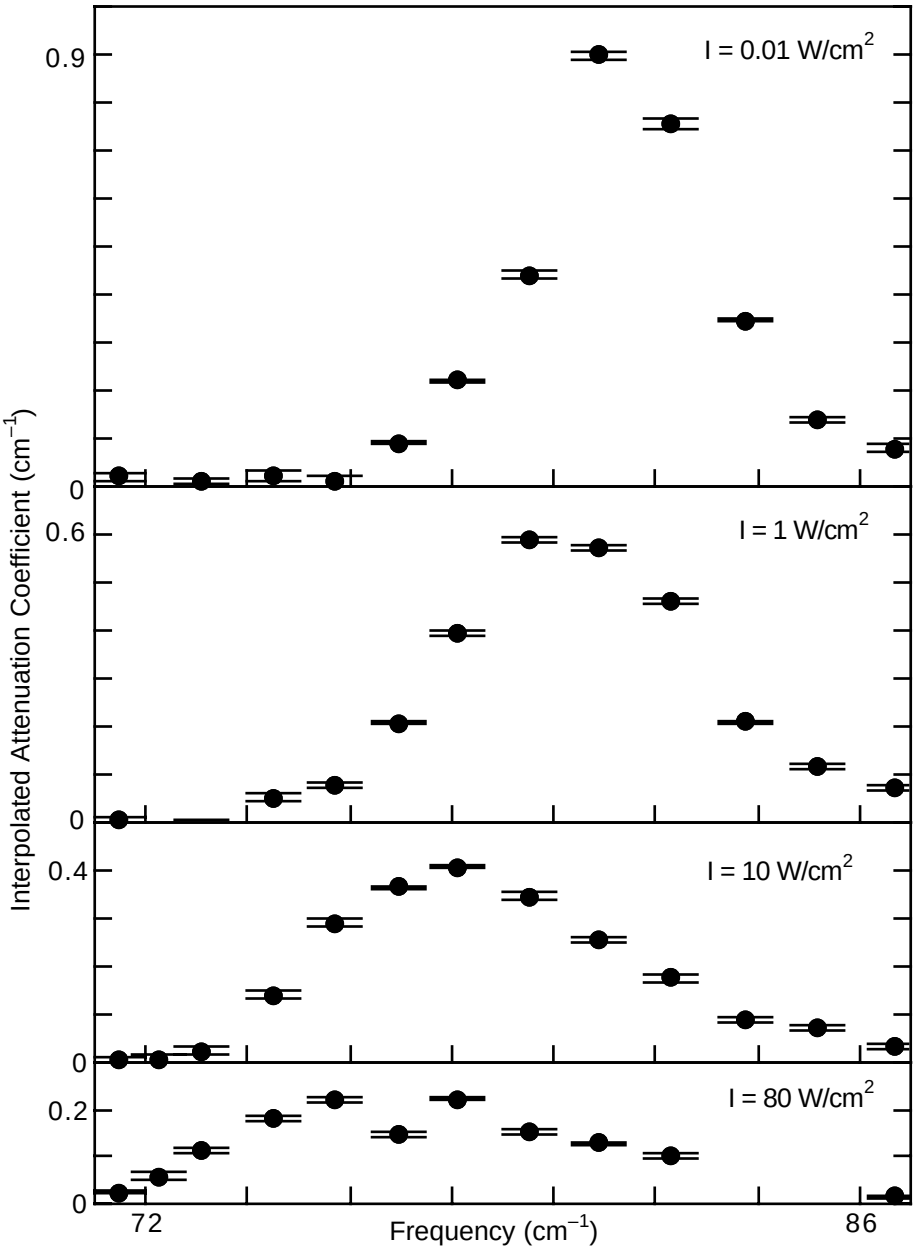

Figure 6.7: The absorption as a function of frequency for the square well sample at $V_G = 0$ V interpolated at 0.01 W/cm^2, 1 W/cm^2, 10 W/cm^2, and 80 W/cm^2.

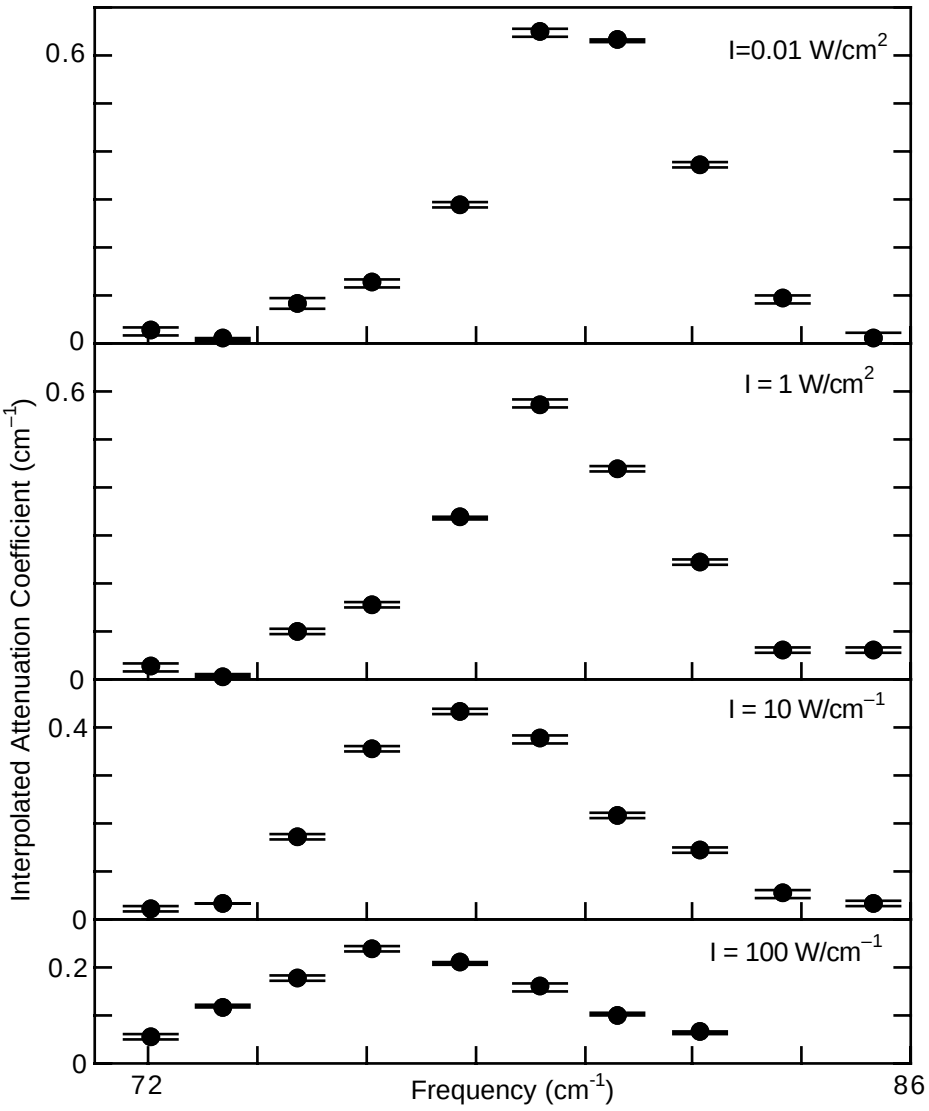

Figure 6.8: The absorption as a function of frequency for the square well sample at $V_G = -0.5$ V interpolated at 0.01 W/cm^2, 1 W/cm^2, 10 W/cm^2, and 100 W/cm^2.

6.2 Theory of Załużny

In Reference [29], M. Załużny created a theory of the conductivity of the two-subband system as a function of incident intensity and frequency, including depolarization and population effects. (This is a completely separate theory from the waveguide mode theory discussed in Chapter 4.) Given a monochromatic excitation of a given frequency and intensity, the theory calculates the conductivity of the intersubband system at that frequency. (Assuming the waveguide is linear, the attenuation coefficient α is linearly related to the real part of the conductivity.)

A key to understanding this theory is that the theory only considers a single frequency at a time: it calculates the conductivity of the system *at the frequency it is being driven at!* (By driving frequency, I mean the frequency of the exciting laser.) Fortunately, this is exactly what I measured in these experiments, but it is important to realize that the theory does not calculate the conductivity at any frequency other than the driving frequency.

Following Reference [29], the equation of motion of the (2×2) density matrix ρ is given by:

$$\frac{d\rho}{dt} = (1/i\hbar)[\mathcal{H}_0 + \mathcal{H}'(t), \rho] - R\rho \tag{6.1}$$

where t is time, $i = \sqrt{-1}$, $\hbar$ is Planck's constant divided by 2π, $\mathcal{H}_0$ is the unperturbed Hamiltonian of the subband system, $\mathcal{H}'(t)$ is the time-dependent perturbing Hamiltonian, and R is a relaxation operator which removes energy from the electrons and dephases them in each subband.

Writing out the components of Equation (6.1),

$$\dot{\rho}_{12} \quad = \quad i(\omega_{21} + i/T_2)\rho_{12} + (i\hbar)\mathcal{H}'_{12}(t)\Delta\rho \tag{6.2}$$

$$\dot{\rho}_{21} = -i(\omega_{21} - i/T_2)\rho_{21} - (i\hbar)\mathcal{H}'_{21}(t)\Delta\rho \tag{6.3}$$

$$(\dot{\rho}_{11} - \dot{\rho}_{22}) = (2i/\hbar)[\mathcal{H}'_{21}(t)\rho_{12} - \rho_{21}\mathcal{H}'_{12}(t)] - (\Delta\rho - \Delta\rho^{(0)})/T_1 \tag{6.4}$$

where T_1 is the energy relaxation time (the intersubband relaxation time), T_2 is the dephasing time (the intrasubband relaxation time), $\mathcal{H}'_{ij}(t) = \langle\chi_i|\mathcal{H}'(t)|\chi_j\rangle$, (where i and j are subband indices which are 1 or 2), χ_i is the envelope wavefunction in the z direction[1] for the i-th subband, $\Delta\rho = \rho_{11} - \rho_{22}$ is the difference in population between the lower and upper states, and $\Delta\rho^{(0)} = \rho_{11}^{(0)} - \rho_{22}^{(0)}$ is the unperturbed population difference. (The unperturbed density matrix is denoted $\rho^{(0)}$.)

The perturbation Hamiltonian consists of the oscillating electric field of the incident radiation, and the changes to the potential due to the motion of the electrons. Thus:

$$\mathcal{H}'(t) = eD(t)z - (4\pi e^2/\epsilon)\int_{-\infty}^{z} dz' \int_{-\infty}^{z'} dz''\, \Delta n(z'', t) \tag{6.5}$$

where $D(t)$ is the external time-dependent electric field, e is the charge of the electron, ϵ is the dielectric constant of the semiconductor, and $\Delta n(z'', t)$ is the change in the electron density due to the external field. Note that it is the field caused by the changes in the electron density which creates the depolarization shift. This perturbation Hamiltonian does not include the effects of the exchange and correlation of the electrons. It only includes the coulombic repulsion of the electrons (the Hartree term).

Assuming a monochromatic field allows the field to be written as:

$$D(t) = \widetilde{D}\exp(-i\omega t) + \widetilde{D}^*\exp(i\omega t) \tag{6.6}$$

where ω is the angular frequency of oscillation of the external field.

[1]The z direction is the growth direction.

Making the rotating wave approximation in this case allows the perturbation Hamiltonian to be explicitly written in the form:

$$\mathcal{H}'_{ij}(t) = \widetilde{\mathcal{H}}'_{ij}(\omega)\exp(-i\omega t) + \widetilde{\mathcal{H}}'_{ij}(-\omega)\exp(i\omega t). \tag{6.7}$$

Similarly, the density matrix elements may be written as:

$$\rho_{ij}(t) = \widetilde{\rho}_{ij}(\omega)\exp(-i\omega t) + \widetilde{\rho}_{ij}(-\omega)\exp(i\omega t). \tag{6.8}$$

After some mathematical manipulation, and considering only terms in ρ_{12} and ρ_{21} which have time-dependence $\exp(-i\omega t)$ or $\exp(i\omega t)$, and more mathematical manipulation, a steady-state solution can be found. The solution is that the nonlinear complex conductivity of the system can be written as:

$$\widetilde{\sigma}_{zz}(\omega, I) = (1/\widetilde{D})\int_{-\infty}^{+\infty} j(\omega, z)\, dz \tag{6.9}$$

$$= \frac{-e^2 i\omega N_s}{m^*} f_{21} \frac{\mathcal{W}(\omega, I) H_{12}}{E_{21}^2 - (\hbar\omega)^2 - i2\hbar\omega\Gamma_2} \tag{6.10}$$

where $I = c\epsilon^{1/2}|\widetilde{D}|^2/2\pi$ is the incident intensity, $j(\omega, z)$ is the current induced by the radiation, N_s is the electronic sheet density (in units of 1/area), $f_{21} = 2m^*|\langle\chi_2|z|\chi_1\rangle|^2\omega_{21}/\hbar$ is the oscillator strength, $\Gamma_i = \hbar/T_i$ are the relaxation *rates*, and E_{21} is the energy of the subband spacing. The two other parameters, $\mathcal{W}(\omega, I)$ and $H_{12} = \widetilde{\mathcal{H}}'_{12}(\omega)/e\widetilde{D}\langle\chi_1|z|\chi_2\rangle$, can be found by solving the following two coupled equations:

$$\mathcal{W}(\omega, I) = \left[1 + \frac{I}{I_0}\frac{|H_{12}|^2 2\Gamma_2^2[E_{21}^2 + (\hbar\omega)^2]}{[E_{21}^2 - (\hbar\omega)^2]^2 + (2\hbar\omega\Gamma_2)^2}\right]^{-1} \tag{6.11}$$

where $I_0 = c\epsilon^{1/2}\Gamma_1\Gamma_2/8\pi e^2|\langle\chi_2|z|\chi_1\rangle|^2$ is the saturation intensity. The other coupled equation is:

$$H_{12}\left[1 + \frac{\alpha_{11}E_{21}^2}{E_{21}^2 - (\hbar\omega)^2 - i2\hbar\omega\Gamma_2}\mathcal{W}(\omega, I)\right] = 1 \tag{6.12}$$

where the depolarization shift parameter α_{11} is given by:

$$\alpha_{11} = \frac{8\pi N_s e^2}{\epsilon E_{21}} \int_{-\infty}^{\infty} dz \left[\int_{-\infty}^{z} dz' \, \chi_2(z') \chi_1(z') \right]^2 . \tag{6.13}$$

The parameter $\mathcal{W}(\omega, I)$ physically represents the fractional net absorption of the system, or equivalently the population difference between the subbands. (The value of $\mathcal{W}(\omega, I)$ is the steady-state value of $\Delta\rho$.) When the subband populations are equal, there is no net absorption and $\mathcal{W}(\omega, I) = 0$. When all the electrons are in the ground state, the absorption is a maximum and $\mathcal{W}(\omega, I) = 1$.

Combining the coupled equations into a single equation by eliminating H_{12} yields the following function[2] in terms of $\mathcal{W}(\omega, I)$:

$$f(\mathcal{W}) = -1 + \mathcal{W} + \tag{6.14}$$
$$\left[\frac{2\mathcal{W}\Gamma_2^2 (I/I_0)[E_{21}^2 + (\hbar\omega)^2]}{E_{21}^4[1 + 2\alpha_{11}\mathcal{W} + \alpha_{11}^2\mathcal{W}^2] + (\hbar\omega)^2[4\Gamma_2^2 - 2E_{21}^2 - 2\alpha_{11}\mathcal{W}E_{21}^2 + (\hbar\omega)^2]} \right] .$$

where I have shortened $\mathcal{W}(\omega, I)$ to $\mathcal{W}$ for notational simplicity. The possible values of the parameter $\mathcal{W}$ lie in the interval from zero to one, and the roots of $f(\mathcal{W}) = 0$ in that interval give the steady-state values of $\mathcal{W}$. There is either one root or three roots of this function, depending on the values of parameters such as ω and I. These roots must be found numerically, due to the complexity of the dependence of Equation (6.14) on $\mathcal{W}$. The roots are discussed in more detail in the section on bistability below, which also includes plots of Equation (6.14).

Once the value of $\mathcal{W}$ has been found from Equation (6.14), it is straightforward to find the corresponding value of H_{12} from Equation (6.12), and then to find the value

[2]There are other possible forms of Equation (6.14) which are simpler, but the one listed here is quite similar to the form used (and checked) in the numerical code.

of the conductivity $\tilde{\sigma}_{zz}(\omega, I)$ from Equation (6.10). The absorption is then assumed to be linearly related to the real part of the conductivity.

The Załużny theory does not take into account the reduction in intensity as the beam propagates through a long sample. Therefore, I modified the numerical simulation to include the fact that the intensity of the radiation decreases as it propagates through the sample, due to the absorption by the well. This can be a major effect; the intensity at the front and the back of the sample can differ by 60% or more. To include this effect in the simulations, the sample was numerically divided into narrow strips which are perpendicular to the direction of propagation. (Usually the sample was divided into one hundred strips; this was enough that using ten fewer strips did not change the results.) Then the incident intensity was used as the input into the theory for the first strip. Given this intensity (which is essentially constant inside the strip due to its short length), the absorption of the strip was calculated. From this, the transmission of this strip was calculated, giving the intensity to be used as the input into the next strip. By doing this, the absorption of each part of a long sample was realistically included.

The effect of breaking the sample into strips is to shift the results around slightly. However, without the strips, there is a discontinuity at lower frequencies, and this feature is modified significantly by the strips. Without the strips, the discontinuity is truly discontinuous; there is an infinite slope there.[3] However, when sliced up into strips, the simulation instead gives a steep (but not vertical) connection between the high and low absorption states. The reason for the smoothing of the discontinuity is that without the slices, there is conceptually a single slice. When that single strip

[3]This discontinuity is the sudden switch between the two stable states; see the discussion on "Bistability" below for a complete explanation.

switches from low to high absorption, it is infinitely sharp. However, with many slices, each strip switches at a slightly different frequency, because each strip "sees" a slightly different intensity. Thus, moving from lower to higher frequencies at a given initial intensity, first the front strip switches (it has a slightly higher intensity than any of the others), and then at a slightly higher frequency, the second strip switches, and so on. Including a large number of slices makes the switches blend together to form a smooth-looking transition. I would expect the sample to exhibit very similar behavior to the sliced simulation.

The Załużny theory omits exchange and correlation, which are a correction to the coulomb repulsion of the electrons. It also omits the excitonic correction to the depolarization shift, which was discussed in Chapter 2. Therefore, the experimentally measured absorption frequency (including both the depolarization shift and excitonic correction) were used in the theory as the absorption frequency. The bare subband spacing E_{12} was calculated with exchange and correlation.

6.3 Comparison with Theory

The theory of Załużny was used to fit the nonlinear absorption data, and to extract a saturation intensity from the data. In order to do this, a computer program was written to find the roots of Equation (6.14). The sample parameters (the depolarization shifted absorption frequency $\widetilde{\omega}_{12}$ from the low intensity absorption measurements, the bare absorption frequency ω_{12} from self-consistent modelling with exchange and correlation, the charge density N_s from capacitance measurements, the sample length, and the linewidth Γ_2 from the low intensity absorption measurements) were entered into the program, as well as the measured incident intensity I on the sample.

Then the program was used to calculate the predicted absorption for a variety of saturation intensities I_0. The output of the Załużny theory program was then compared with the experimental data, and the values of the saturation intensity were adjusted until an acceptable fit was found.

Once the fit was found, this determined the relaxation time for that intensity. At first, it was assumed that a single relaxation time would fit an entire data set for all intensities. However, this did not work. The best fits were clearly obtained using *different relaxation times at different intensities*. This was a major discovery as well as a surprise. The reasons for the variation of relaxation time with intensity are discussed in detail in Chapter 7.

The experimental results of nonlinear absorption for the square well, with fits from the Załużny nonlinear absorption theory, are shown in Figures 6.9 and 6.10. The linear absorption data (top curve) were used to fit the peak width and peak height. Then, using only a single adjustable parameter (the relaxation time), each of the other curves was fit at each intensity. Figure 6.9 shows the absorption for zero volts gate bias, which corresponds to 1.39×10^{11} e/cm^2. The absorption data are referenced to -2.0 V gate voltage, where the capacitance indicates that the sample is depleted.[4] The relaxation times which have been extracted from these fits are noted in the captions of the Figures, and are discussed in Chapter 7. Figure 6.10 shows the absorption for -0.5 V gate bias, again referenced to -2.0 V. This corresponds to a charge density of 5.69×10^{10} e/cm^2.

The theory of Załużny contains bistability: for some values of the excitation frequency and intensity, there are two stable absorption values, and one unstable ab-

[4]I believe that the sample still had a bit of residual charge in it when the capacitance indicated it was depleted. See Appendix E.

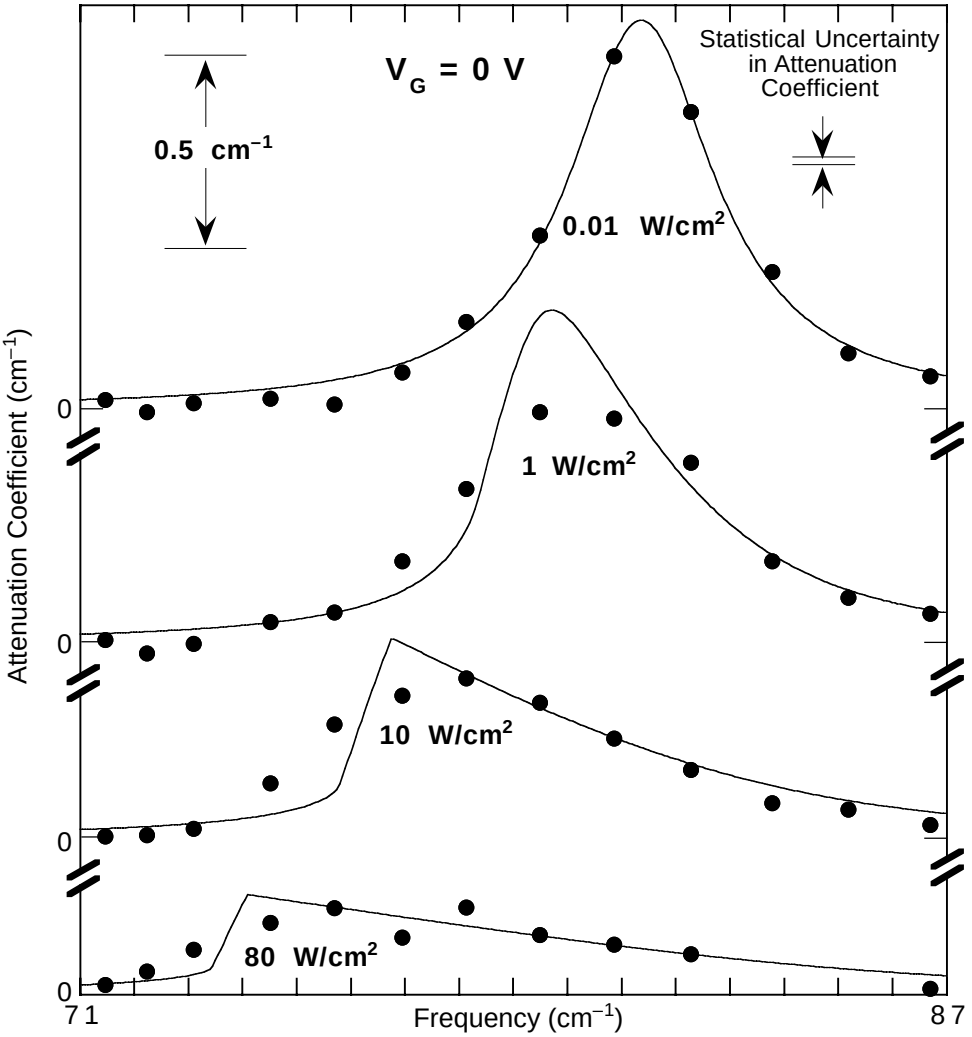

Figure 6.9: The experimental nonlinear absorption results for the square well, at zero volts gate bias. The attenuation coefficient is shown as a function of frequency at several different intensities. The large dots are the same data as in Figure 6.7, and the solid curves are theoretical fits using the nonlinear absorption theory of Załużny. The calculated relaxation times for the lower three curves are 730 ps, 310 ps, and 110 ps, from upper to lower.

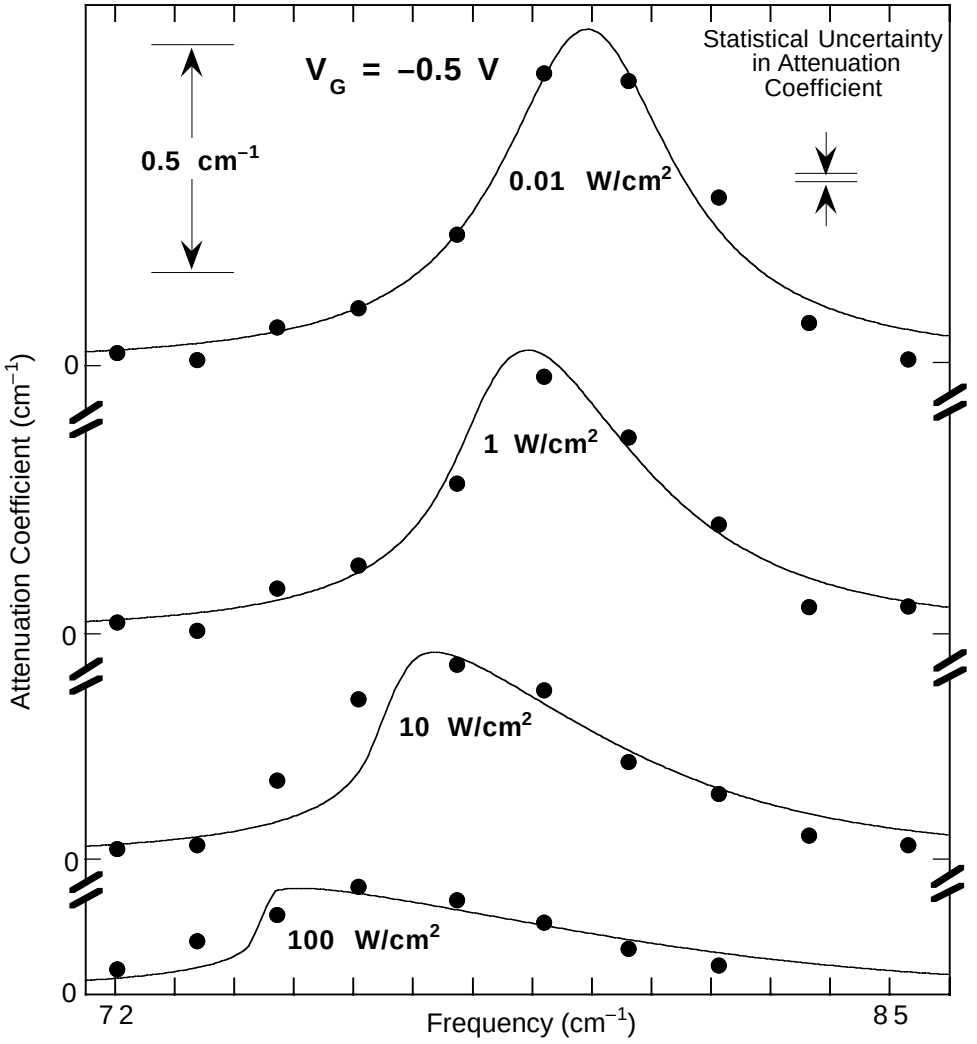

Figure 6.10: The experimental nonlinear absorption results for the square well, with -0.5 volts gate bias. The attenuation coefficient is shown as a function of frequency at several different intensities. The large dots are the same data as in Figure 6.8, and the solid curves are theoretical fits using the nonlinear absorption theory of Załużny. The calculated relaxation times for the lower three curves, from upper to lower, are 480 ps, 150 ps, and 50 ps.

sorption value. Which of the stable absorption values is observed depends partially on the past history of the system. In fitting the data, the stable solution to Equation (6.14) was chosen which corresponds physically to the system starting out with all the electrons in the ground state. The other stable solution, which corresponds to the sample starting out with a strongly excited population, did not fit the data as well. (In my measurements almost all the electrons did start out in the ground state.) This is discussed in detail in the section entitled "Bistability" below.

6.4 Discussion of Depolarization Shift Dynamics

There are several conceptually distinct pieces to understand about the depolarization shift dynamics. First I will discuss a simple saturation experiment without a depolarization shift. Next I will explain how the linear absorption of the system behaves as a function of subband population with the depolarization shift. Then I will describe how the subband populations and absorption evolve in time during intense monochromatic excitation. Finally, I will include the additional complication that the absorption frequency is the same as the excitation frequency, and thus the subband populations are not held constant during the measurement. The timescales on which the dynamics occur is also discussed.

6.4.1 Simple Saturation

The basic foundation for understanding the experimental results is a simple saturation experiment. By simple saturation, I mean the saturation of a system where the absorption frequency does not depend upon the populations of the energy levels.

Imagine such a system with two energy levels. Given energy relaxation processes

with a relaxation time T_1, the system will absorb radiation as shown in Figure 6.11. As the intensity increases, the population of the upper level will increase, and the total amount of absorption will decrease. This is due to the stimulated emission from electrons in the upper level competing against absorption from electrons in the lower level. (If the populations of the upper and lower levels are equal, there is no net absorption, as the stimulated emission completely cancels the absorption.) By analyzing how the absorption decreases with intensity using a rate equation, it is straightforward to calculate the relaxation time T_1.

6.4.2 Fixed Subband Populations

Consider a hypothetical experiment: the population state of the quantum well system with the depolarization shift is "set" to a certain amount in the upper and lower subbands using a continuous-wave laser. Then the linear absorption of the system as a function of frequency is measured using a weak probe.

The theoretical results[5] are shown in Figure 6.12. As the fraction of charge in the upper subband is increased, two effects are apparent: the area (and height) of the peak decreases, and the frequency of the center of the peak decreases. The area of the peak decreases because of the simple saturation effect discussed above.

The peak shifts to lower frequency as the upper subband is populated due to a reduction in the depolarization shift. As electrons are placed in the upper subband, they reduce rather than increase the coulombic restoring force on the other electrons, which reduces the depolarization shift, and therefore the absorption frequency. (This

[5]This was calculated simply by assuming that the width of the peak remained constant, that the height (and therefore area) of the peak was proportional to $(N_1 - N_2)$, and that the center-frequency of the peak is given by Equation (2.1).

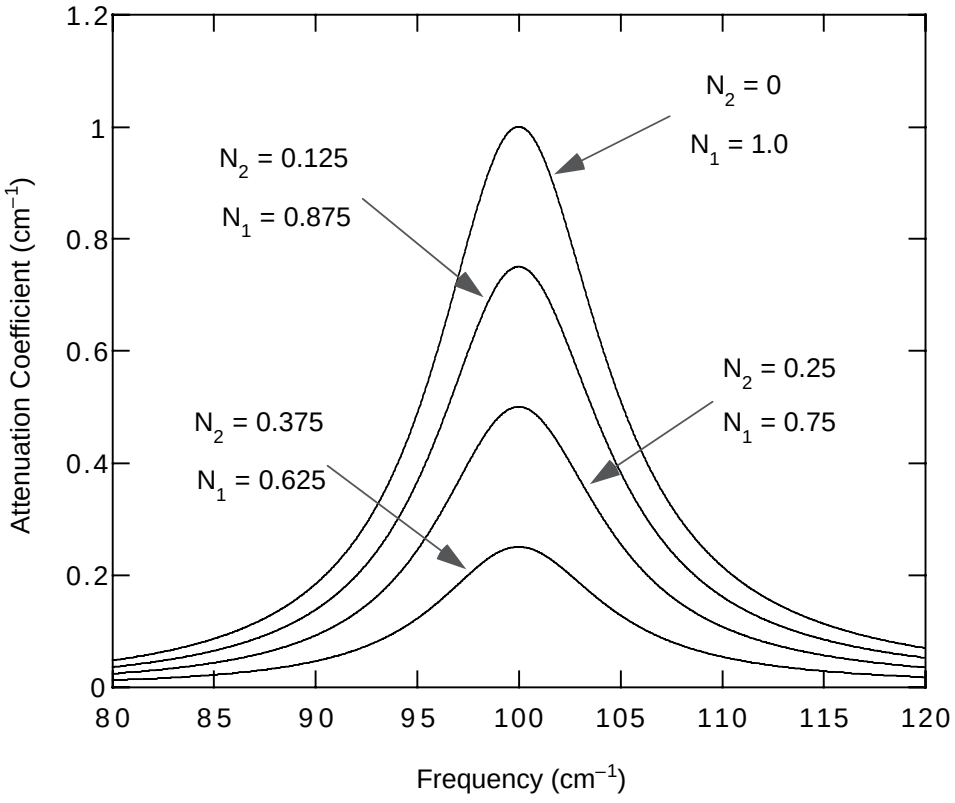

Figure 6.11: The attenuation coefficient as a function of frequency at several different level populations, for a simple system where the absorption frequency does not depend upon the level populations. As the level populations change, stimulated emission from the upper state cancels the absorption from the lower state, and the total amount of absorption decreases. Here N_1 denotes the fraction of the population in the lower state, and N_2 denotes the fractional population in the upper state. ($N_1 + N_2 = 1$)

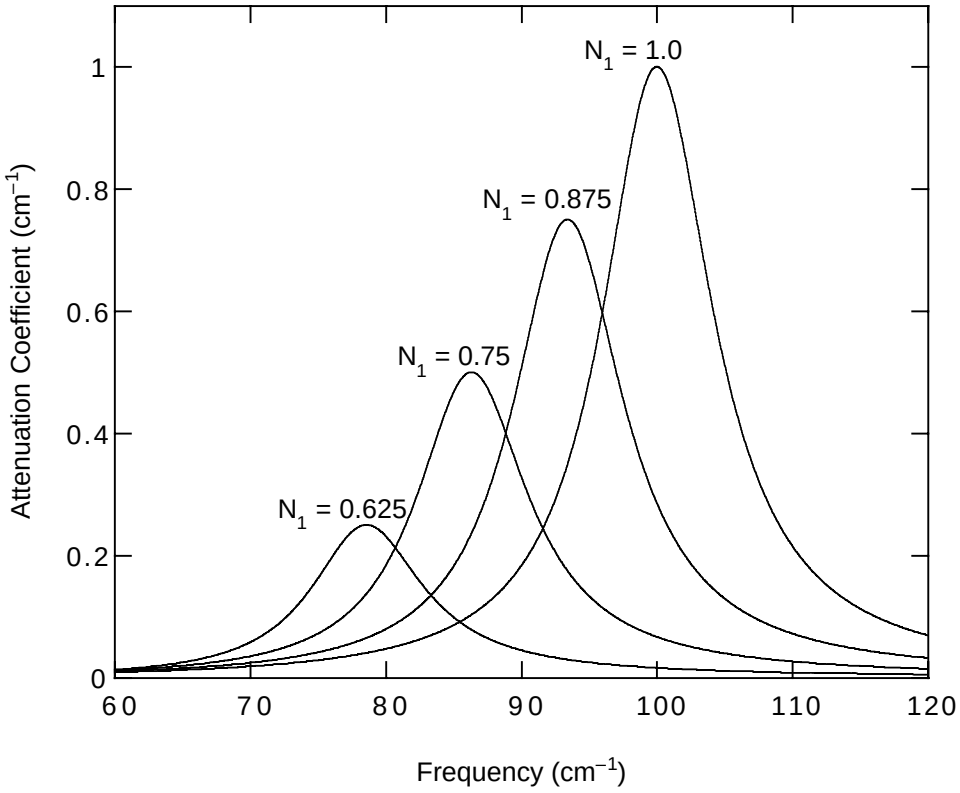

Figure 6.12: The attenuation coefficient as a function of frequency with various fractions of the charge in the lowest subband. The depolarization-shifted linear absorption frequency $\widetilde{\omega}_{12} = 100$ cm^{-1}, the depolarization factor $\alpha = 1.041$, the bare subband spacing $\omega_{12} = 70$ cm^{-1}, and the scaled total charge $N_s = 1.0$. As the population is moved from the lower subband to the upper subband, the absorption strength is reduced and the peak shifts to lower frequency.

occurs according to Equation (2.1).) Note that these changes to the linear absorption of the system are simply due to the population of the subbands; other than creating the excited population in the subbands, the laser radiation has no effect on the absorption of the system.

6.4.3 Time-Evolution of Population

The actual experiment is more complicated than the simple experiment considered above. The reason is that the population is not simply "set" as it was above. The population in the real experiment is allowed to vary according to the relaxation rates and the optical excitation.

Consider the cold ($k_B T \ll E_{12}$) sample before optical excitation. Most of the electrons will be in the ground state. The linear absorption will consist of a strong absorption with a large depolarization shift. When the FIR laser is suddenly turned on at a certain frequency and intensity, the sample will start absorbing the radiation as determined by the absorption of the sample *at that time*.

If the FIR laser frequency is far from the peak in the linear absorption, and/or the laser intensity is weak, then the total amount of FIR absorbed by the sample will be small, and the energy relaxation processes will keep the electrons in the lower subband. In this case, the measured steady-state absorption will simply be the linear absorption at that frequency.

However, if the intensity of the FIR beam is large, and/or the frequency is near the linear resonance so that the initial linear absorption of the sample is large at the FIR beam frequency, then the sample will start absorbing a large number of photons. Assuming this absorption rate is greater than the relaxation rate, the population of the upper state will increase, and the absorption of the sample will change. As the

absorption of the sample changes, there are two possible effects on the number of photons absorbed. If the FIR laser frequency is below the frequency of the center of the absorption peak, as the peak shifts to lower frequency, the absorption at the beam frequency will *increase*: the absorption peak will shift down in frequency into the FIR laser frequency. (This has been observed in the experimental data; see Figures 6.1 and 6.4.) On the other hand, if the FIR frequency is above the frequency of the center of the absorption peak, the increase in population in the upper subband will still shift the peak to lower frequencies. But as the peak shifts, it will move *away* from the FIR laser frequency, and the amount of absorption from the beam will *decrease*. See Figure 6.13 for a visual representation.

Note that the effect of shifting the center of the peak closer to the laser line is partially offset by the reduction in the total amount of absorption due to the saturation of the subband populations. If the peak is narrow enough, then the shifting peak will increase the absorption faster than the saturation reduces it. However, for a wider peak, the shift of the peak in frequency will not create enough additional absorption to compensate for the loss in oscillator strength due to the saturation. For the rest of this discussion I will assume that the peak is narrow. This condition was easily satisfied by the square well sample.

Here is where things start to become very interesting: as the populations change and the absorption starts to change, the amount of radiation absorbed by the sample changes. For radiation with frequency below the center of the peak, the increase in absorption will cause even more electrons to be excited to the upper subband, leading to even more of a downward shift in the absorption peak. In fact, the peak will continue to shift down in frequency until the center of the peak is below the laser line. At this point, as the peak shifts lower, the amount of absorption is reduced (both

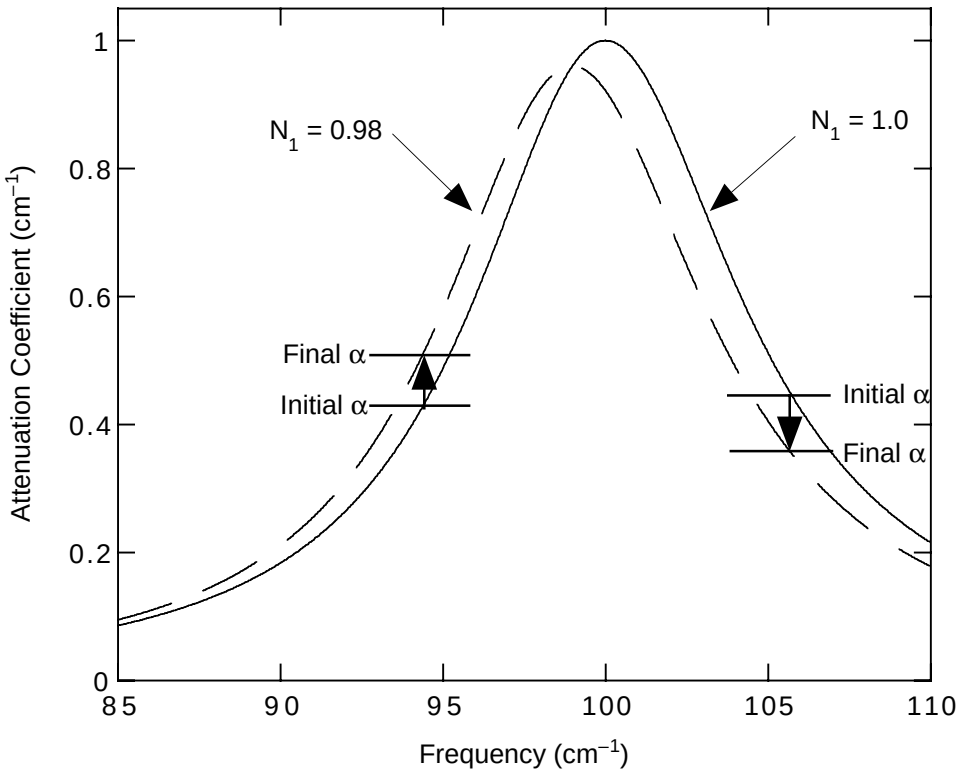

Figure 6.13: A cartoon of the attenuation coefficient of the quantum well with two slightly different subband populations. By absorption of radiation, the population in the ground subband changes from $N_1 = 1.0$ to $N_1 = 0.98$. Due to the change in population, the peak shifts to lower frequency (and loses a small amount of area). If the exciting radiation frequency is below the maximum value of the peak, then the absorption at the FIR beam frequency will *increase*. (As shown for 95 cm^{-1}.) However, if the frequency of the exciting radiation is above the maximum value of the peak, then the absorption will *decrease*. (As shown for 105 cm^{-1}.) This assumes that the peak is narrow; see text for discussion.

by the saturation of the absorption and by the peak shifting away from the laser line). At some point, the peak will reach equilibrium, with the laser frequency somewhere on the right half of the peak. This process is illustrated in Figure 6.14.

This point is stable to fluctuations and noise: a slight decrease in laser intensity (noise) will cause the populations to relax a bit toward a state with more charge in the lower subband. This will cause the peak to shift a bit up in frequency, which will increase the absorption at the laser frequency. This will keep the subband populations roughly constant.

This all assumes that the initial intensity and linear absorption strength is such that the initial absorption was strong enough to overwhelm the relaxation processes. If this is not the case, then the populations will not change, and the peak will not shift. This is a different stable state.

We now have a *complete* picture of what happens as a fixed frequency laser is used to illuminate the sample: if the intensity is large enough, the peak begins to shift downward in frequency and lose area. Once it starts shifting, it will keep shifting until it reaches a point where the laser frequency is on the right side of the peak; at some point on this side of the peak, a stable state will be found. At this stable state, the absorption at the laser frequency and the intensity of the laser beam are exactly enough to maintain the upper subband population, and the system is stable to noise.

6.4.4 Nonlinear Absorption Lineshape

In Figure 6.14 in the previous section, I plotted the linear absorption of the system at all frequencies while the system was being excited with a monochromatic laser beam. In these experiments, the entire linear absorption spectrum was not measured; instead, only the absorption of the laser beam *at the excitation frequency* was mea-

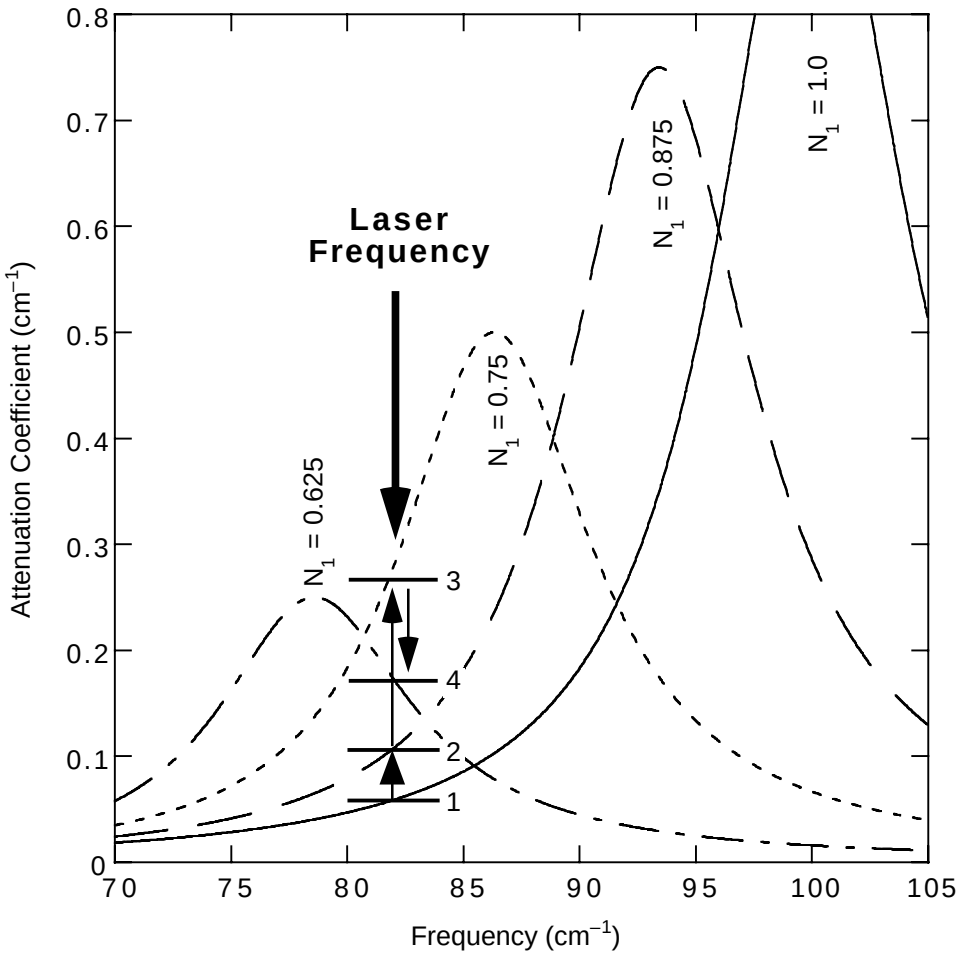

Figure 6.14: A cartoon of the evolution of the subband populations and the absorption. All electrons start in the lower subband, with the absorption labeled $N_1 = 1.0$. The laser is turned on at the frequency shown; the beam is intense enough that the upper subband population increases. Soon, the lower subband population is $N_1 = 0.875$, and the absorption is given by that curve. At the laser frequency, the absorption has increased from the initial value (1) to the value labeled 2. The peak continues to shift down to $N_1 = 0.75$; the absorption at the laser frequency increases to 3. Finally, the center of the peak shifts below the laser line, the absorption begins to decrease, and a stable state is found, with absorption 4 and lower subband population $N_1 = 0.625$.

sured. Thus, each absorption curve of the actual experimental results is at a constant intensity, but *not at a constant subband population.* The populations of the upper and lower subbands vary along the curve as a function of frequency. In comparison, the "cartoons" in Figures 6.11, 6.12, 6.13, and 6.14 above are plotted with constant subband populations.

Therefore, the interpretation of plots like those shown in Figures 6.9 and 6.10 is the following: the absorption measured is the steady-state absorption which the system arrived at when it was excited with the given frequency of radiation at the given intensity. Since different excitation frequencies will cause the system to end up in a different population state, this means that every point along the theoretical curves of Figures 6.9 and 6.10 has a *different subband population*. This means that the linear absorption of the sample in any given population state is similar[6] to the shifting peaks of Figure 6.12. The smooth curves of the Załużny theory are composed of different points on *different* peaks. See Figure 6.15 for a visual representation.

6.4.5 Timescales

The response of the absorption peak to changes in the subband population should be very fast. At most, it should take a single FIR cycle time for the population change to modify the absorption. This is because the depolarization shift is essentially a dynamical screening process; it should therefore adjust in a time which is at most a cycle of FIR (about 0.5 ps). Since the intensity of the FIR from the FEL changes on a timescale of several nanoseconds, the delay between the repopulation of the subbands and the change in absorption is negligible.

[6]The main deviation from Figure 6.12 would be a variation in the linewidth of the peak.

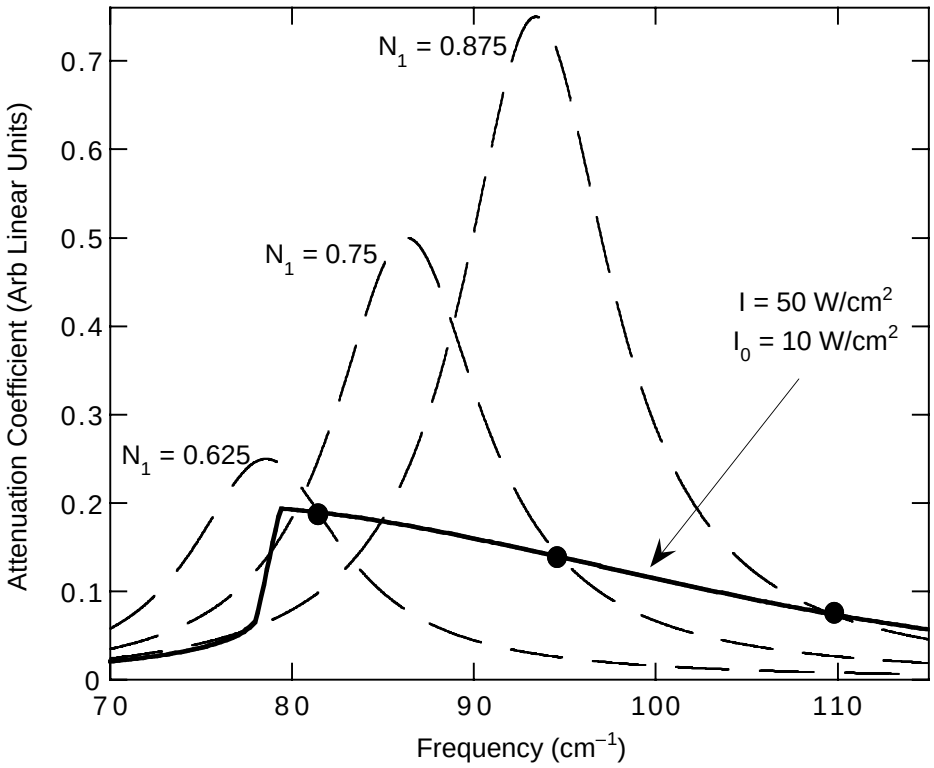

Figure 6.15: Each point on the absorption curve of the theory of Załużny is just a stable point on an absorption peak of the sample in a given intersubband population state. The solid curve is the absorption predicted by the theory of Załużny at the intensities shown. Each of the three dashed curves are the intersubband absorption with a fixed fraction (N_1) of charge in the lower subband, as labeled. (These dashed lines are the same as the ones shown in Figure 6.12.) The large dots show that each point on the Załużny theory absorption curve is just one point from the absorption curve at a fixed population state. Furthermore, this plot graphically finds the steady-state absorption points on the dashed curves, which occur at the intersections of the Załużny curve and the *right* side of the dashed peak, and are marked with large dots. For example, the left dot marks the steady-state point on the $N_1 = 0.625$ absorption peak for the calculated intensities. In addition, this dot marks the point on the Załużny absorption curve where $N_1 = 0.625$. (The absorption parameters used in this plot are identical to the ones used in Figure 6.12.)

The time for the populations to come to steady-state during constant excitation is much longer. This time is determined by the intersubband relaxation time, which has been measured here to be on the order of a fraction of a nanosecond. Since the steady-state corresponds to an equalization of the relaxation rates and the excitation rates, it may require several (or many) intersubband relaxation times for the system to come to steady-state. However, since the relaxation time is roughly 2500 times shorter than the FIR pulse, the sample will be in a steady-state most of the time the laser is on. Therefore the assumption made above that the experimentally measured absorption corresponds to the steady-state absorption is valid.

6.4.6 Bistability

The theory discussed above contains bistability. For some values of the excitation frequency and intensity, there are two stable absorption values. Which absorption value is observed can depend on the past history of the system.

Figure 6.16 shows the theoretically calculated absorption for a hypothetical set of parameters. Between approximately 63.5 cm^{-1} and 69.2 cm^{-1}, the system exhibits two stable solutions. Figure 6.17 plots Equation (6.14) at several frequencies at the same parameters. (The roots of this function determine the steady-state values of $\mathcal{W}$.) As can be clearly seen, the number of roots of the function change as the frequency enters and leaves the bistable region. At frequencies above or below the bistable region, there is only one root, and inside the bistable region there are three roots. The root with the intermediate $\mathcal{W}$ value is unstable. The root with the lower value of $\mathcal{W}$ corresponds to the root where the system starts out in the excited state; the higher-value root corresponds to the system starting out in the ground state.

This bistability can be understood in a very physical way. Consider the sample

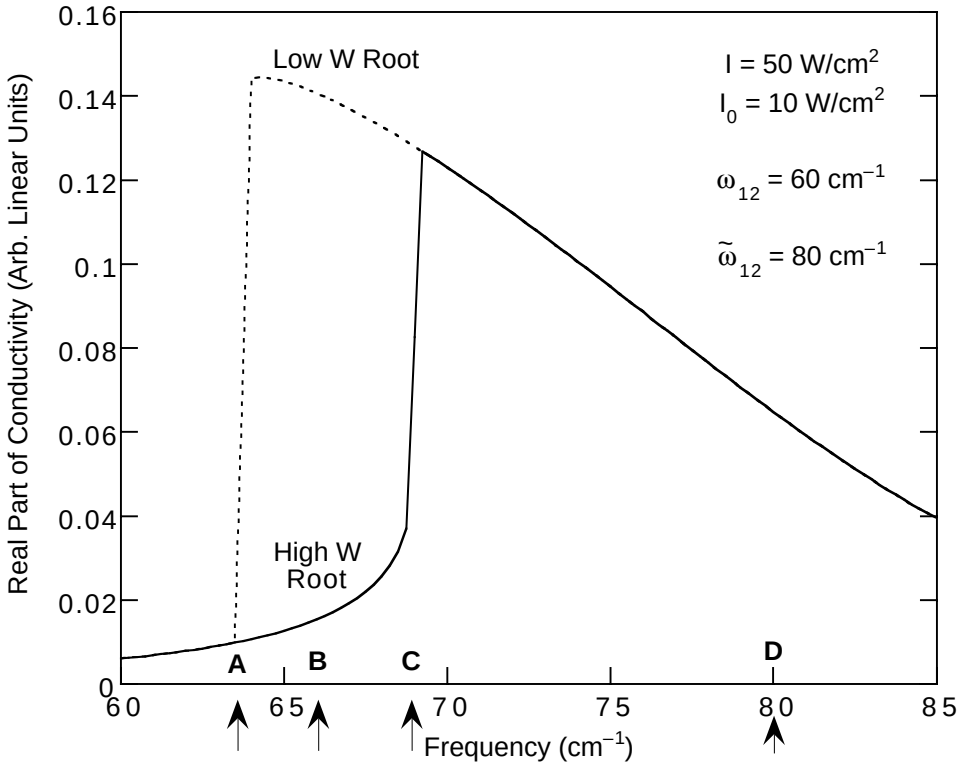

Figure 6.16: The output of the Załużny theory for the absorption as a function of frequency for the parameters shown. The two curves correspond to the (numerically) high and low values of $\mathcal{W}$ which are roots of Equation (6.14). (The middle root corresponds to an unstable solution.) Figure 6.17 plots Equation (6.14) for these parameters at frequencies which are at the lower edge of the bistable region (labeled "A"), the middle of the bistable region ("B"), the upper edge of the bistable region ("C"), and well above the bistable region ("D").

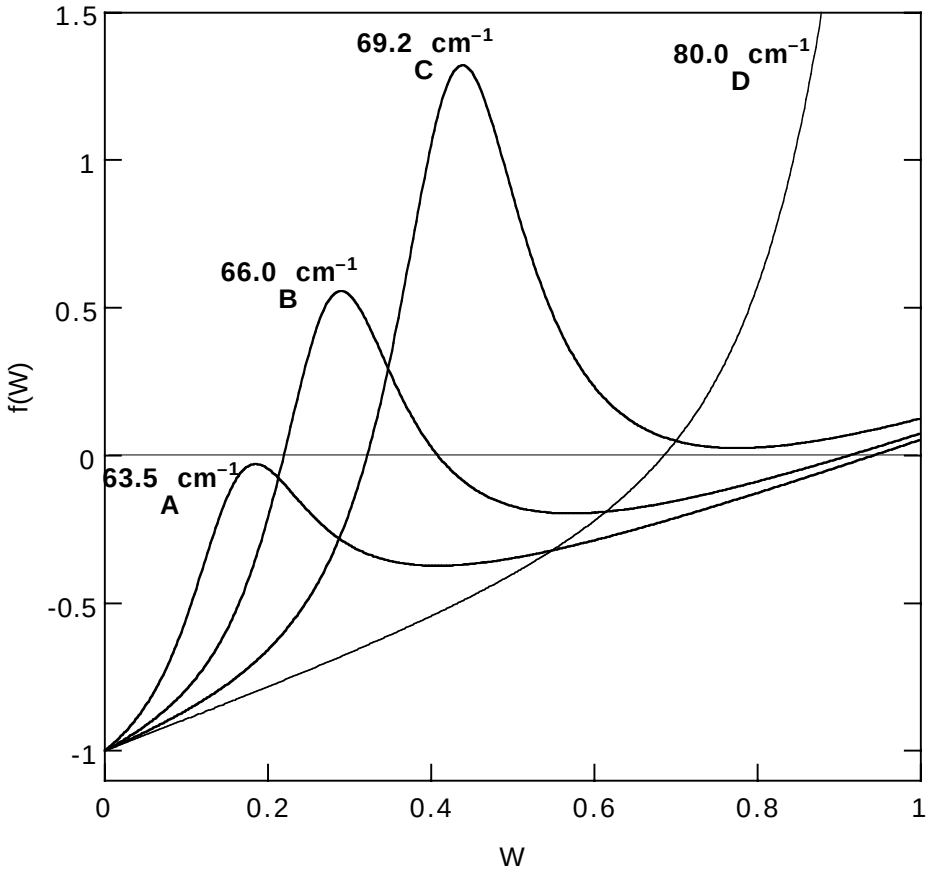

Figure 6.17: A plot of the function in Equation (6.14) at several different frequencies, with the same parameters as in Figure 6.16. The number of roots changes as a function of frequency, as the frequency enters and leaves the bistable region. Compare with Figure 6.16 for the meaning of the frequencies and the letter labels. (The lower frequency edge of the bistable region is labeled “A”, the middle of the bistable region “B”, the upper edge of the bistable region “C”, and well above the bistable region “D.”)

with all the electrons in the lower subband. If I set the laser frequency to be fairly far below the initial peak, then the initial absorption of the laser will be quite small (because the laser frequency is far out on the tail of the linear absorption peak). Even if the intensity is rather large, the small attenuation coefficient value will mean that the amount of radiation absorbed is not enough to overwhelm the relaxation processes, and bring the peak down into the laser frequency. Therefore, a small absorption value is stable.

However, consider starting the laser at the same frequency and intensity when the system already has a significant population in the upper subband. Then the laser will turn on while the absorption peak of the sample is shifted to lower frequency. Since the peak starts out much nearer to the laser frequency, the amount of absorption at the laser frequency is much larger than the linear absorption. Thus, the initial absorption of the excited system will be enough to start changing (or at least maintaining) the subband populations. The system will find a stable absorption value with a large amount of absorption and a significant population in the upper subband. This is also a stable absorption value, but it is a larger absorption than when the sample started with all the electrons in the lower subband. Hence, the sample exhibits two stable absorption values for the same laser frequency and intensity. Which one is actually chosen by the physical system depends upon the past history; in this case, the subband populations when the laser is turned on.

The bistability can be seen in the theory as the sharp cutoff in absorption below the linear absorption peak frequency at moderate to high intensity. The cutoff is the discontinuity (“switching”) between the high-absorption state to the low-absorption state. Actually, there is a hysteresis loop in the neighborhood of the cutoff: this is the region where there are two values of attenuation coefficient which are stable.

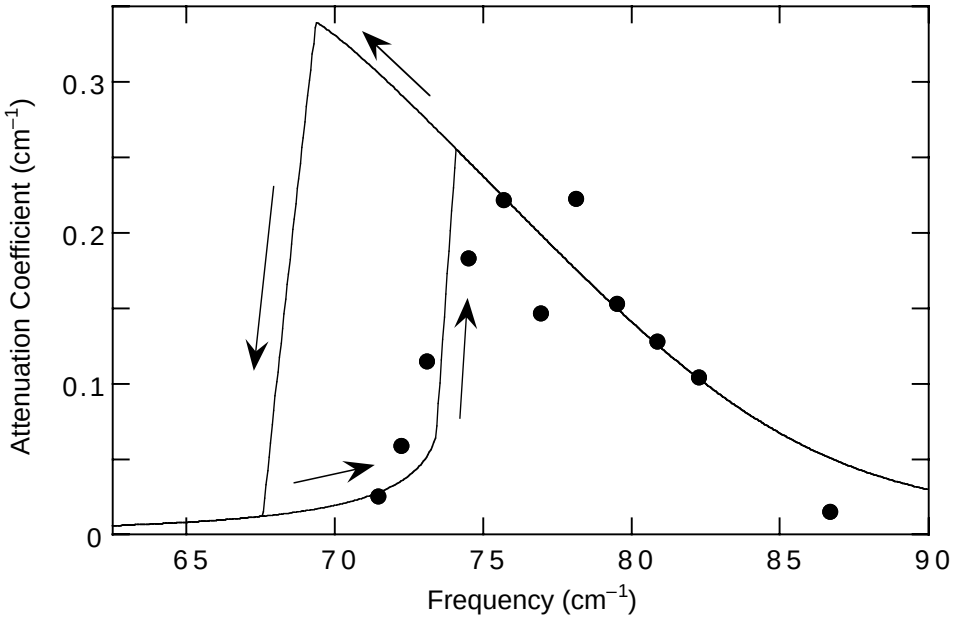

Figure 6.18: The two stable solutions of the Załużny theory for absorption as a function of frequency, at 80 W/cm^2 and a saturation intensity of 32.4 W/cm^2, with $V_G = 0$ V. The arrows depict the direction which the system could be driven around the hysteresis loop. The solution which has larger absorption values in the bistable region corresponds to the system starting out in an excited population state. The lower absorption solution corresponds to the system starting out near the ground state. The dots shown are experimental data at 80 W/cm^2; note that they clearly follow the solution which corresponds to the system starting out near the ground state, as expected.

See Figure 6.18, where both stable states are plotted. Note that, in Figure 6.18, the experimental data are much better fit by the solution which started near the ground state, with most of the electrons in the lower subband. This is expected, since the sample is cold and unexcited before the FEL pulse.

The arrows in Figure 6.18 correspond to the direction the system could be driven around the hysteresis loop by a continuous-wave tunable source. This could not be done with the UCSB FEL, because that FEL is pulsed, not continuous-wave. The FEL could be set to shift in frequency a small amount each pulse, but the sample

would relax to the ground state in between each pulse, completely losing all memory of its state.

6.5 General Discussion

In general, the agreement between theory and experiment shown in Figures 6.9 and 6.10 is strikingly good. The linear absorption data (top curve) were used to fit the peak width and peak height. Then, using only a single adjustable parameter (the relaxation time T_1), each of the other curves was fit at each intensity. The goodness of fit for these curves indicates that the theory worked well: a single fit parameter simultaneously fit the height of the peak, the shape of the peak, and the position of the sharp change in absorption at the low-frequency side.

On the low-frequency side of the high-intensity peaks, the theory predicts a sharper change in absorption (switch between the two bistable states) than is seen in the experiment. It is possible that the time-structure of the FEL pulse smoothed the measured data. The FIR pulse from the FEL can contain several longitudinal cavity modes of the FEL resonance cavity. This causes the FEL output beam to oscillate in intensity at 33 MHz, with an amplitude of roughly 25%–75%. Since the oscillation of the intensity is relatively slow on the timescale of the response of the sample, the sample would see a slowly increasing intensity followed by a slowly decreasing intensity, repeating over and over. Assuming the FEL is at a frequency and intensity within the sample's bistable region, the sample would switch from the low absorption state into the high absorption state as the intensity increased, and as it decreased, it would switch back to the low absorption state. (Note that the high-to-low switch would occur at a different intensity than the low-to-high, due to the hysteresis.) This

would lead to a time-averaging in the measurement process, with the measured absorption being an average of the high-absorption state and the low-absorption state, weighted by the fraction of time spent in each state.

This effect could easily lead to the observed smoothing. The lower the laser frequency is below the absorption frequency, the higher the intensity needed to switch the sample into the high absorption state. Thus, as the frequency of the laser is lowered, the amount of time spent in the high-absorption state is smoothly reduced, and the average absorption observed decreases smoothly. This explanation suggests that, on short time scales, the experimental absorption could be as sharp as the theory of Załużny predicted, and the experimental technique smoothed the data.

6.6 Novel Device Possibilities

The nonlinear absorption observed in the square well sample leads to the clear possibility of a novel device. Just below the frequency of the linear absorption peak, the absorption of the sample is very unusual. It has a small absorption at low intensities; then at moderate intensities the radiation changes the populations of the subbands enough to shift the absorption peak down into the frequency of interest, increasing the absorption. Finally, at higher intensities the populations saturate and the absorption again decreases.

This absorption is unique as far as I know. It is not immediately clear what such a nonlinear absorber could be used for, but it seems probable that there are practical applications. This behavior is completely different from the behavior of most materials, which absorb strongest at low intensity and then saturate at higher intensities.

In addition to the idea of a novel device, the quantum well system is quite in-

teresting in that it is a general nonlinear absorber at relatively low intensities. Most nonlinear materials require high intensities to exhibit nonlinear effects. (Such nonlinear effects are useful for optical modulation and harmonic generation.) However, due to the long relaxation times exhibited by these samples, even a relatively small intensity is sufficient to change the populations of the subbands and change the absorption. (For example, even 1 W/cm^2 is sufficient to significantly change the absorption.) A nonlinear absorber which operates at low intensities could be useful for many technological applications.

Chapter 7

Energy Relaxation Times

One of the goals of this work was to measure the energy relaxation time (T_1) for electrons in the first excited subband of the square well sample, where the subband separation (≈ 8 meV) is much less than the LO phonon energy (36.7 meV). Energy relaxation times are important from the fundamental physics standpoint of understanding the interactions of the electrons with each other and with other excitations of the crystal. In addition, knowledge and understanding of energy relaxation is essential for the creation of new solid-state devices, where such rates determine many device parameters such as maximum switching rates and steady-state populations in lasers.

7.1 Previous Measurements

Many measurements have been made of the energy relaxation time for electrons in GaAs. These measurements can be divided into two groups: electrons with energy

greater than the LO phonon,[1] and electrons with energy less than the LO phonon.

7.1.1 Background

Electrons in GaAs lose energy mainly by emission of phonons. (In most modulation-doped samples grown by molecular beam epitaxy, the impurity concentration is so low as to be negligible.) However, the electrons also scatter off each other, which keeps them in a thermal distribution (a Fermi distribution with a single temperature describing all the subbands). Electron-electron scattering cannot remove energy from the electron gas. At low temperatures, measurements indicate that the time for the electrons to thermalize is extremely fast, faster than 0.01 ps (10 fs) [22, page 329], [13]. Thus, the excited carriers quickly thermalize, and then the thermal distribution cools by emitting phonons.

7.1.2 Relaxation Above the LO Phonon Energy

A great deal of work has been done measuring the relaxation time in narrow quantum structures, where the electrons have energy greater than the GaAs LO phonon energy and are able to relax by emission of LO phonons. The following are a selection of the measurements which have been performed.

Seilmeier et al. [21] measured the relaxation time in a doped 5 nm multiple quantum well using time-resolved intersubband absorption measurements. At room temperature (300 K) they measured a time of approximately 10 ps, at relatively high charge densities ($\approx 5 \times 10^{11}$ e/cm^2).

[1] LO phonons are longitudinal optical phonons, which are the phonons most strongly coupled to the electrons in GaAs. These phonons have a relatively flat dispersion, and therefore a relatively constant energy of 36.7 meV.

Julien et al. [12] measured the relaxation time in two doped multiple quantum well samples with well widths of 8.5 nm and 9.6 nm, and very high charge densities (1.3×10^{12} e/cm^2 and 1.9×10^{12} e/cm^2 respectively.) Using intersubband saturation spectroscopy, they measured the relaxation times to be 10 ps for the 8.5 nm sample and 15 ps for the 9.6 nm sample.

Tatham et al. [24] measured the relaxation time in an undoped 14.6 nm multiple quantum well using time-resolved Raman spectroscopy to measure the time-dependent population of the subband states. Using very low excitation intensities to keep the electron density below 3×10^{10} e/cm^2, they measured a relaxation time of approximately 0.5 ps at 30 K.

Levenson et al. [14] measured the relaxation time in a series of undoped multiple-well samples with well widths 12.0 nm, 22.5 nm, and 24.0 nm. The measurement technique was a time-resolved interband pump and probe. For the two narrower samples (where the subbands were separated by more than an LO phonon energy), the relaxation times at 15 K were 8 ps and 20 ps for the 12.0 nm and 22.5 nm samples respectively. The charge density was estimated to be 3.5×10^{11} e/cm^2 and 7×10^{11} e/cm^2 for the 12.0 nm and 22.5 nm samples respectively.

In these measurements the electrons lose their energy almost entirely by emission of LO phonons. The characteristic theoretical time for electrons to emit LO phonons is roughly 0.1 ps [22, page 56]. However, the energy relaxation time for the electrons is often limited by the time for the LO phonons to decay into acoustic phonons. This is because the LO phonons become trapped inside the well by the mismatch of ionic masses at the interface, and therefore remain inside the well with the electrons. While inside the well, these phonons can reheat the electrons. Therefore, the timescale for the electrons to lose their energy becomes the LO phonon decay time, which is es-

timated to be approximately 7 ps [22, page 55]. This is known as the hot phonon effect.

Much of the variation in measured relaxation times probably occurs due to the variation in charge densities. In the measurements by Tatham et al. [24], the charge density was kept very low, which may have kept the phonon population low enough that the phonons could not effectively reheat the electrons. The other measurements find a relaxation time which is comparable to the time for the LO phonons to decay.

In addition to the charge density variations between these measurements, there are many variations and differences between these measurements which make direct comparisons difficult: several different optical measurement techniques are used, the measurements are done at a variety of different temperatures, and the samples vary in design, quality, and doping (meaning some are undoped and therefore have both electrons and holes, while others are doped and have only electrons). Considering these factors and the difficulties involved in making accurate measurements, a relatively wide range of results is to be expected.

7.1.3 Relaxation Below the LO Phonon Energy

A single electron with energy less than the LO phonon energy cannot emit any LO phonons due to energy conservation. The only way it can relax is to emit acoustic phonons. (Note, however, that a *gas* of electrons with average energy less than the LO phonon energy will be in a thermal distribution, and a fraction of the electrons will have enough energy to emit LO phonons.) The measurements of energy relaxation times below the LO phonon are somewhat rarer, due to the challenges of working at long wavelengths or small energy shifts.

Oberli et al. [19] measured the relaxation time in a 21.5 nm undoped multiple

quantum well at 5 K using time-resolved Raman scattering. Based on the difference between the decay rates of the subband populations, they estimated a relaxation time on the order of 325 ps.

The third sample of Levenson et al. [14] (the one with the 24.0 nm well width from above) had an intersubband separation less than the LO phonon energy. The measured relaxation time using the same methods as described above was 35 ps.

Heyman et al. [10] measured the relaxation time of the double square well sample described in Chapter 2, by measuring a photovoltage which is proportional to the excited population. At 10 K and low intensity, they measured a relaxation time of 1.2 ns (1200 ps). At a temperature of 50 K, they measured 20 ps at low intensity, and at 10 K and high intensity they measured a time of 15 ps.

Recently, Murdin et al. [18] performed time-resolved intersubband absorption measurements of two doped samples in the FIR. The multiple quantum well samples had well widths of 30 nm and 24 nm, with charge densities of 2.0×10^{11} e/cm^2 per well. They observed a strong dependence of relaxation time on intensity. At the lowest intensity they observe a relaxation time of 500 ps for the 30 nm well and 200 ps for the 24 nm well.

Acoustic phonon emission is a very slow process, due to the relatively weak coupling between the electrons and the acoustic phonons. The theoretical time for electrons to lose their energy to acoustic phonons is on the order of hundreds of picoseconds [6], and depends upon the details of the envelope wavefunctions (see below). Since acoustic phonons may have (essentially) arbitrarily small energy, any electron with excess energy can emit an acoustic phonon.

It is now generally understood that the electron temperature plays a key role in intersubband relaxation for subband spacings smaller than the LO phonon energy.

Since the carrier-carrier scattering keeps the electrons in a thermal distribution, there are always a fraction of electrons with energy greater than the LO phonon energy, which can relax via LO phonon emission. The fraction of electrons which are able to emit LO phonons is strongly dependent upon the electron temperature. This is graphically depicted in Figure 7.1.

This explains the results of Levenson et al. [14], as the 24.0 nm sample they studied had an intersubband spacing fairly close to the LO phonon energy, and even a modest amount of excess kinetic energy allowed the electrons to rapidly emit LO phonons. Although the sample was cooled to 15 K, it is possible that the electron temperature was higher due to the intense optical pump. It also explains the variation in relaxation times observed by Heyman et al. [10] and Murdin et al [18]. At higher temperatures and excitation intensities, the average electron energy increases, allowing a larger fraction of electrons to possess enough energy to relax via LO phonon emission. Murdin et al. have created a simple theory of the energy balance between the electrons and the phonons. This theory suggests that the relaxation time is roughly constant for electron temperatures below 35 K, and shortens rapidly as the electron temperature rises above 35 K.

The relaxation time measured by Oberli et al. [19] is surprisingly long, considering the relative proximity of the upper subband (26.8 meV) to the LO phonon energy (36.7 meV). Goodnick et al. [8] have performed a Monte Carlo simulation of these measurements, and found that the long times are due to the rather high charge densities (4×10^{11} e/cm^2) and hot phonon effects which repopulate the excited subband with the high energy tail of the thermal distribution.

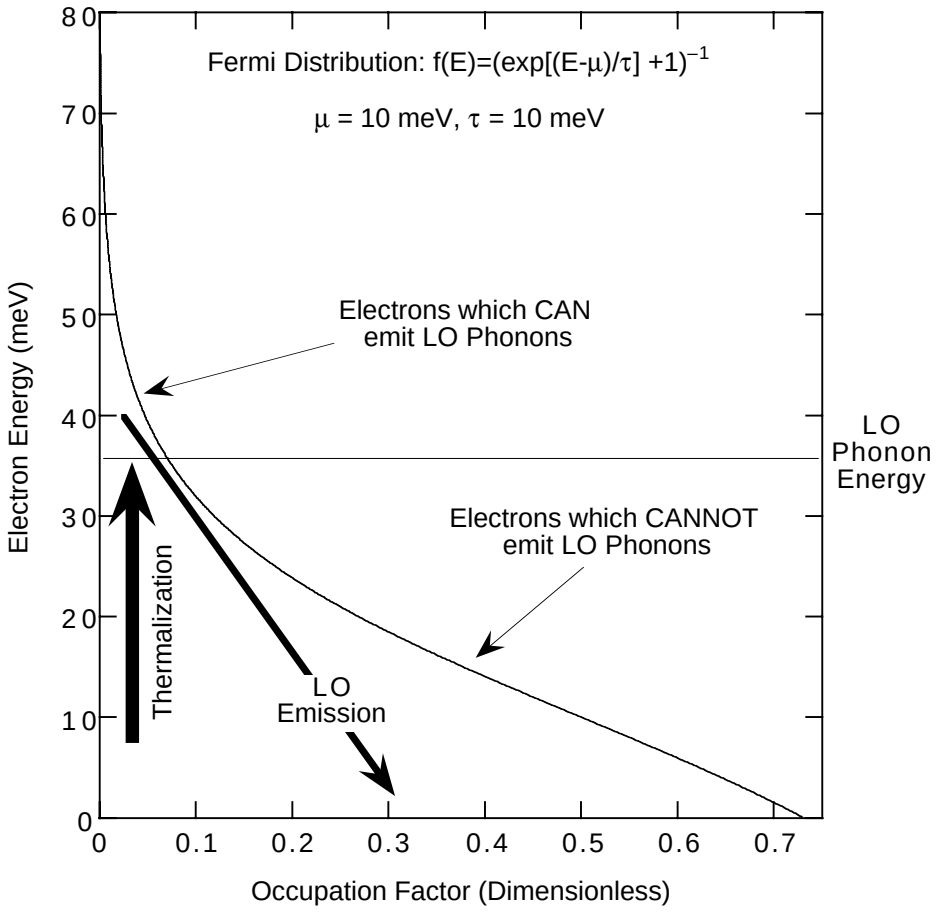

Figure 7.1: The electron occupation function at a given temperature and chemical potential. Only a small fraction of the electrons are above the LO phonon energy and able to emit LO phonons. As these energetic electrons emit LO phonons and relax, the electron-electron scattering keeps the electron gas in a thermal distribution by pushing other electrons above the LO phonon energy.

Gate Voltage	Intensity	Relaxation Time (T_1)
0.0 V	1.0 W/cm^2	730 ps
0.0 V	10.0 W/cm^2	310 ps
0.0 V	80.0 W/cm^2	110 ps
-0.5 V	1.0 W/cm^2	480 ps
-0.5 V	10.0 W/cm^2	150 ps
-0.5 V	100.0 W/cm^2	50 ps

Table 7.1: The measured relaxation times for the square well sample. The uncertainty in the intensity and therefore T_1 is roughly a factor of two.

7.2 Results

The measured energy relaxation times for the square well sample are tabulated in Table 7.1. These times are plotted as a function of intensity in Figure 7.2.

The uncertainty in the relaxation times is mainly due to the uncertainty in the intensity inside the sample, which is discussed in Appendix H. In short, the intensity (and therefore all the relaxation times) are uncertain to roughly a factor of two.

7.3 Discussion

First I will describe a calculation of the acoustic phonon emission time. Then I will discuss the experimental results, and then discuss and speculate about the meaning and limitations of the standard relaxation times T_1 and T_2.

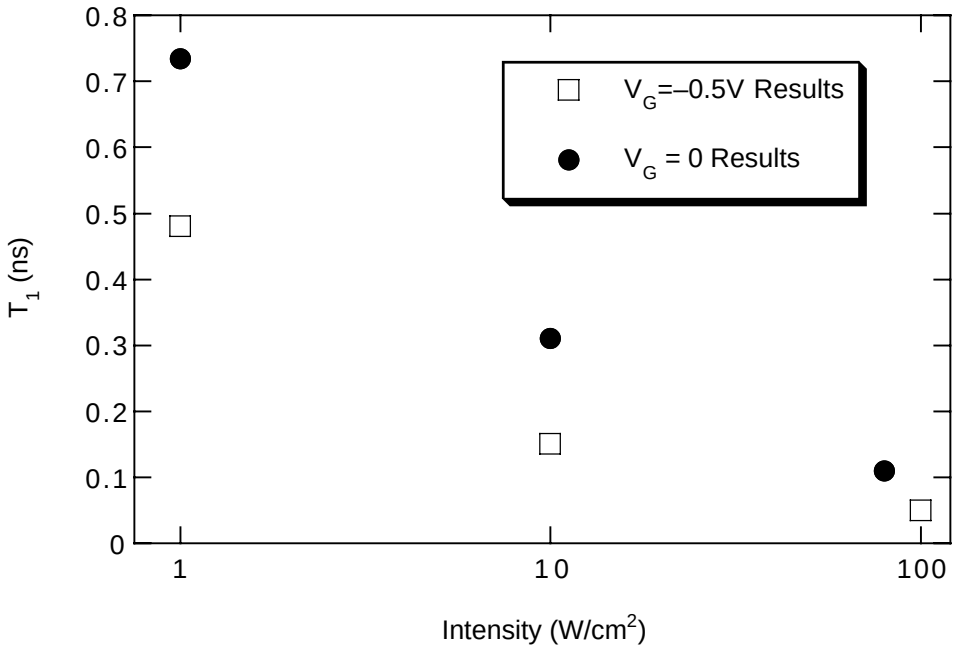

Figure 7.2: The measured relaxation time T_1 as a function of intensity on a logarithmic scale. The solid dots are the $V_G = 0$ data, and the open squares are the $V_G = -0.5$ V data.

7.3.1 Acoustic Phonon Emission Calculation

In Reference [6], Ferriera and Bastard derive an expression for the energy relaxation time due to acoustic phonon emission in GaAs. Using first-order time-dependent perturbation theory, they find:

$$\frac{1}{\tau_{if}} = \frac{2c_0}{8\pi^2\hbar}\int_{-\infty}^{+\infty} dq_{||}\,|f(q_{||})|^2 \int_0^{2\pi} d\theta$$
$$\int_0^{+\infty} dk_f\, k_f\, \hbar\omega(1+n)\delta(E_i - E_f - \hbar\omega + (\hbar^2/2m^*)(k_i^2 - k_f^2)) \quad (7.1)$$

where τ_{if} is the relaxation time from an initial state i to a final state f, $q_{||}$ is the wavevector of the acoustic phonon in the growth direction, k_i and k_f are the in-plane wavevectors of the electron in the initial and final states (respectively), $\omega =$

$c_s(q_{||}^2 + |\mathbf{k}_i - \mathbf{k}_f|^2)^{1/2}$ is the frequency of the phonon, m^* is the effective mass of electrons in GaAs, and E_i and E_f are the energies at the bottom of the initial and final subbands. The coupling constant $c_0 = D^2/2\rho c_s^2$, where D is the deformation potential for electrons (8.6 eV), ρ is the mass density of GaAs (5.3 g/cm^3), and c_s is the longitudinal speed of sound (3700 m/s). The function $f(q_{||})$ is given by

$$f(q_{||}) = \int_{-\infty}^{+\infty} \chi_i(z) e^{-iq_{||}z} \chi_f(z)\, dz \tag{7.2}$$

where $\chi_i(z)$ and $\chi_f(z)$ are the envelope wavefunctions in the initial and final states. The thermal acoustic phonon occupation n is given by

$$n = \left[\exp\left(\frac{\hbar\omega}{kT}\right) - 1\right]^{-1} \tag{7.3}$$

where ω is still the phonon frequency, k is Boltzmann's constant, and T is temperature.

Using a numerical computer program written by James Heyman, the emission rate of acoustic phonons for the square well sample was calculated for zero temperature. (Equation (7.1) was used, with the thermal phonon occupation $n = 0$.) Due to the relatively large width of the structure (40.0 nm), only the phonons with a wavelength roughly equal to the well width had a measurable contribution to $f(q_{||})$ in Equation (7.2). Since these are very low frequency phonons, they do not remove much energy from the electrons. This led to a rather long relaxation time of 740 ps. By comparison, the double square well was calculated to have a relaxation time of approximately 230 ps. This time is shorter because the double square well is narrower, meaning that the electronic wavefunctions can efficiently couple to higher frequency phonons.

7.3.2 Discussion of Measured T_1

The measured relaxation time (T_1) for the square well sample decreases as the intensity of exciting radiation increases. At very low intensities, T_1 was 730 ps, which is strikingly consistent[2] with the estimate of acoustic phonon emission of 740 ps. This T_1 is much longer than the times seen in experiments where the subband spacing is greater than the LO phonon energy, by a factor of approximately one hundred or more. The agreement with a theory involving only acoustic phonon emission suggests that, at the lowest intensities measured, LO phonon emission is relatively unimportant.

As the intensity of the exciting radiation increased, T_1 became shorter. This is consistent with the idea that a larger fraction of the heated electron gas was able to relax by emission of LO phonons. The shortest measured T_1 was on the order of 50 ps. This is still roughly seven times longer than T_1s observed for experiments with subband spacings greater than the LO phonon energy. (Above the LO phonon energy, all electrons in the upper subband are able to relax via LO phonon emission.) Apparently, even at the highest intensities measured, only a modest fraction of the electron gas was able to relax via LO phonon emission.

It is possible that, for intensities lower than were measured here, T_1 for electrons in the square well could be longer, perhaps much longer. To accurately measure T_1 using this technique, I need to be able to observe significant changes in the absorption peak caused by an excited population, in order to fit the data to the theory of Załużny and obtain a T_1. The existence of a large enough excited population in the sample to change the absorption may mean that T_1 is already shorter than it was at linear

[2]Remember that the experimental measurement is uncertain to roughly a factor of two!

absorption intensities. Thus, $T_1 = 730$ ps at the lowest intensities which I have measured (1 W/cm^2) should not be confused with T_1 at very low intensities, which I have not measured.

It is interesting that the relaxation times for the two different charge density cases differ so much. The T_1's at the lower charge density (5.69×10^{10} e/cm^2) are shorter at the same intensity than the higher charge density (1.39×10^{11} e/cm^2) relaxation times. The T_1's differ by roughly a factor of two, while the charge density differs by somewhat more than a factor of two. Perhaps the higher charge density electron gas screens the electron-electron interactions more, causing less electron-electron scattering in the higher density case. (Then the scattering of electrons up above the LO phonon energy might be reduced, leading to slower overall relaxation.) Another possibility is that the high density electron gas screens the electron-phonon interaction, reducing the coupling between the electrons and the phonons, and slowing the relaxation.

To summarize, my experimental results suggest that, at low optical excitation, the electrons were too cold to emit LO phonons, and the emission of acoustic phonons was the main mechanism by which electrons relaxed. However, as the optical excitation increased, the electrons gained energy, and the emission of LO phonons became much more important. In all cases, the electron-electron scattering served to keep the electrons in a thermal distribution. Without the constant fast thermalization of the electron gas, the LO phonon emission would allow only the most energetic electrons above the LO phonon energy to relax, and then completely stop.

7.3.3 Meaning of T_1 and T_2

This section is based on intuitive notions I have about the intersubband system, on discussions with colleagues, and on speculation, but not on any specific calculations.

Usually,[3] T_1 is the time for the population to relax, and T_2 is the time for the polarization to relax. The standard interpretation is that T_1 is the time for the electrons to lose their energy, and T_2 is the time for the electrons to lose their quantum-mechanical phase. Note that changing states (a T_1 process) necessarily involves a change of phase.

In many systems, these are the only two timescales. For example, in an atomic system, the electrons can dephase in their state, or they can change to another state (losing energy), but there are no other options. (They can't change to another state without dephasing, and dephasing *and* changing to another state is just changing to another state.) The key is that (in most systems) the populations cannot relax without relaxing the energy, and vice versa. Thus population relaxation and energy relaxation are equivalent.

However, in the subband system of a quantum well, the population relaxation can be different from the energy relaxation. Due to the additional quantum number of the momentum in the plane of the well, electrons *can* change subbands without changing energy, as depicted in Figure 7.3. Thus, a third timescale enters the problem.

This timescale becomes important at very low electron temperatures (where the average energy of the electrons is much less than the intersubband spacing). It could be possible for an electron in the upper subband to relax almost hori-

[3]This is true in the theory of Załużny, and other theories which define the density matrix and relaxation operators as Załużny did.

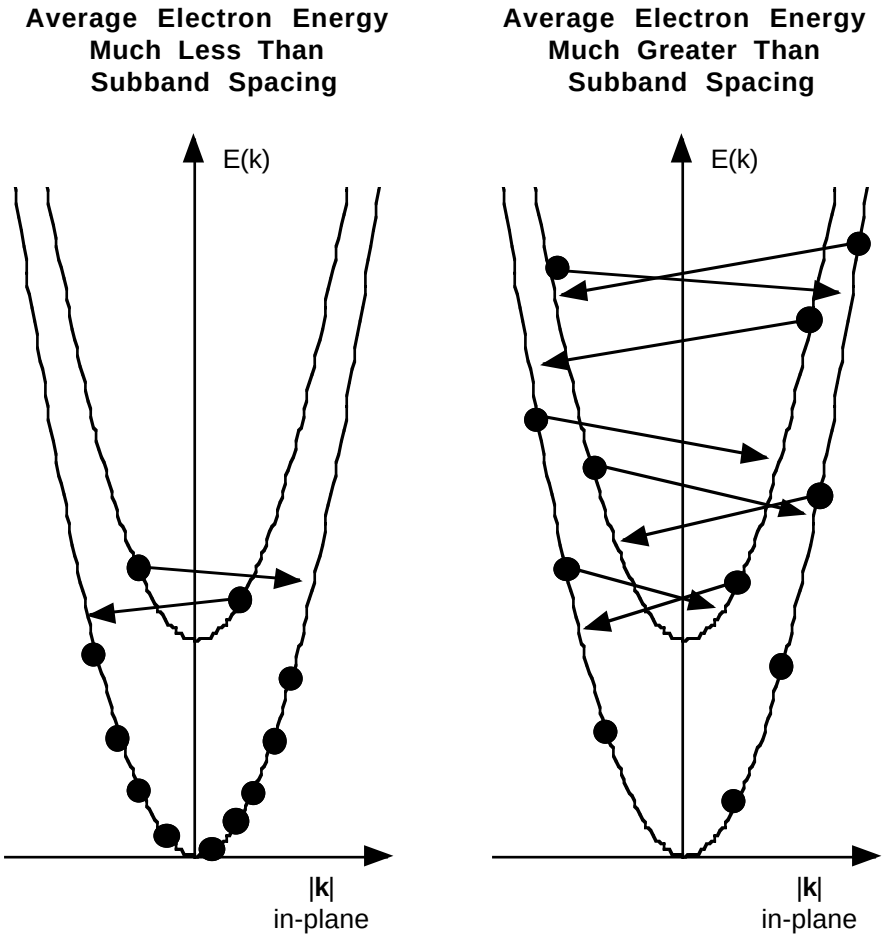

Figure 7.3: The additional timescale between T_1 and T_2. The lowest two subbands are shown, plotted with electron energy as a function of in-plane wavevector. On the left, the low energy of the electrons means that acoustic phonons may move electrons from the upper subband to the lower subband, but not back. Thus, without losing energy, the population relaxes. However, on the right, the average energy of the electrons is large enough that the acoustic phonons will simply move electrons back and forth between the subbands without a significant population relaxation.

zontally (on an energy vs wavevector plot) from the upper subband to the lower one faster than it could lose its energy. An example would be acoustic phonon emission: the phonon carries lots of momentum but rather little energy. An electron emitting an acoustic phonon could get the necessary momentum to change subbands without losing very much of its energy. (See left side of Figure 7.3.) Assuming that this population relaxation process with a single acoustic phonon occurs faster than the energy relaxation process (which could require *many* acoustic phonons), the population would relax in a time faster than T_1. In addition, since it would take longer to emit a phonon than to scatter off other electrons, this population relaxation would take longer than T_2. Thus the new timescale is between T_1 and T_2.

This time would cease to matter at higher electron temperatures (where the average energy of the electrons is significantly greater than the intersubband spacing). In this case, the distribution of electrons would be such that the populations in the two subbands would be roughly equal. Then the acoustic phonon process would have very little effect on the populations: some electrons from the lower subband would be scattered to the upper subband, but a (roughly) equal number of electrons would scatter from the upper subband to the lower one. (See the right side of Figure 7.3.) Thus, the subband populations would remain essentially unchanged by this process, and the only thing which would allow the populations to relax would be a true T_1 process which would remove energy from the electrons. (An example of a true T_1 process would be the emission of multiple acoustic phonons, or the emission of LO phonons by the most energetic electrons in the thermal distribution.)

This suggests that at very low electron energies the relaxation should be faster than T_1, and as the electron energy increases, the relaxation time should increase to T_1. If this is correct, then the measured times from the Załużny theory are really the

population relaxation times, and not the energy relaxation times, at least at low incident intensity. (It is likely that the lowest measured intensities used here are enough to make the energy relaxation time and the population relaxation time equal.) However, in this work, I have not included these effects. I have simply reported the times T_1 and T_2 from the fits of the Załużny theory, and will not discuss the possibility of another timescale further.

A theoretical analysis of this problem would be the next step in understanding these processes in detail.

Chapter 8

Conclusions and Future Directions

We have learned a great deal about wide GaAs/AlGaAs quantum wells: we have learned how to construct a sample with the desired linear intersubband absorption properties, how to verify those linear absorption properties using FTIR, and how to measure the nonlinear absorption of the sample. In addition, we have learned that the system holds promise for optical devices in the FIR; perhaps even an entirely new class of device.

In addition, we now know a lot more about the dynamics of the depolarization shift. The behavior is well described by the theory of Załużny, with some caveats made about the sharpness of the bistable switch.

And we found that even the linear absorption presented surprises, as the integrated absorption strength was not simply related to the charge density in the sample. The explanation of this remains unanswered.

In this Chapter I will summarize the results of this work briefly, and then describe what I see as the future possibilities to move beyond this work.

8.1 Conclusions

The main results of this work are discussed in Chapters 6 and 7. The nonlinear absorption of the samples was indeed complicated and non-intuitive. As the intensity increased, the peak shifted downward in frequency as the depolarization shift was diminished. In general, the results were well described by the theory of Załużny. The experimental data had the unusual characteristic shape predicted by the theory of Załużny.

The relaxation times measured were relatively long, but quite reasonable considering the small intersubband spacing, low excitation intensities, and low temperatures of the sample. The fact that they varied strongly with intensity lends further weight to the premise that the electron temperature is a critical parameter in the relaxation time.

In addition, the linear absorption, which was originally planned only as a diagnostic and planning tool, led to some interesting discoveries. The square well sample exhibited several unusual behaviors, including the integrated absorption strength not varying linearly with charge density, and several unexpected details about the lineshape and baseline. Although waveguide simulations were done, the results did not fit the data well.

Perhaps the most interesting conclusion is that the dynamics of the depolarization shift could provide a basis for a whole new class of devices. Such devices could, among other things, be designed to absorb at moderate intensities. They would also exhibit nonlinear absorption at relatively low intensities.

8.2 Future Possibilities

There are three main experimental possibilities for further work. One is to try to investigate the unexpected behavior of the integrated absorption strength in the linear absorption data. Does it happen for other similar samples? Is it only visible in samples from the wafer I studied? How does it depend upon mobility in the quantum well? How does it depend upon the details of how FIR is coupled into and out of the sample? If the sample also has a backgate (so that the static field across the well can be controlled independently from the charge density), does the same behavior show up when the static field is constant? What effect is causing this? Currently, I am unaware of any theories which explain what I have seen. These measurements could aid in understanding the apparent violation of some basic concepts of physics.

Another possibility for future work is to create the "simpler" experiment mentioned in Chapter 6. As mentioned there, I believe that the linear absorption of the system should always look like a Lorentzian lineshape, even when the subband populations are far from equilibrium. To measure this would involve a strong FIR source to pump the sample and induce a population change, and some sort of weak tunable or broadband source to do linear spectroscopy during the strong pumping. This would be a complicated experiment due to the practical challenges of having two beams passing through the sample simultaneously, and the challenge of measuring the transmission of the weak probe in the presence of scattering from the strong pump. However, this would be a beautiful experiment in that it would be straightforward to understand the results. The payoffs would probably largely be in terms of understanding solid-state physics issues, although it might be a useful starting point to tackle the next experiment below.

A last interesting experiment to do would be to probe the bistability of the nonlinear absorption. There are many ways to do this measurement: most involve some way of monitoring the transmission of the sample while the intensity of incident radiation is modulated. The bistability should appear in a lack of symmetry in where the absorption switches from low to high compared to where it switches from high to low. This would be a challenging experiment due to the required time-resolution and the difficulty in precise time-resolved measurements of extremely low-energy pulses. (Fast pulses with only moderate intensity are necessarily very low-energy.) In addition, any fluctuations in the intensity of the source could well make it essential to not average the data. (Fluctuations in intensity could make the switching occur at slightly different times for each pulse, meaning that averaging would smear all the switches into a big blob.) The payoff of this experiment would not only be a better understanding of the time-dependence of the nonlinear absorption and depolarization shift, but also practical knowledge about how to build and think about new classes of optical devices based upon this bistability.

Finally, it would be useful to have a theory which could explain the unexpected linear absorption data. This would probably best be done after some additional experimental measurements are done, since at this point the source of the abnormalities is very uncertain.

Appendix A

Sample Processing

After a certain sample was grown, I was given part of the wafer by the crystal growers. I then cleaved the wafer into a piece the desired size for a sample. Next, I needed to process the sample to allow it to be used in the edge-coupling geometry with a gate to control the charge density. The samples were processed in the UCSB Engineering II Opto-electronics Laboratory. I put ohmic contacts and gates on the samples using shadowmasks, and an anti-reflection coating using a more involved process. In this Appendix I will describe these processes in detail.

A.1 Ohmic Contacts

The ohmic contacts were placed on the top of the samples at the corners. The contact consists of a metal which is placed or evaporated onto the sample, and then annealled to diffuse a conducting material down into the sample to the well. Once annealled, the ohmic contact should provide electrical contact to the well.

The ohmic contacts had the lowest yield of all the processes. This yield could probably have been improved by etching the contact region so that the well was not

so far from the surface, but this was not done. (The well was approximately 680 nm below the surface, as compared with more traditional samples where the well is 150 to 200 nm below the surface.) Before the ohmic contacts were evaporated, the samples were cleaved into chips of the desired size. Then they were cleaned (acetone, then isopropyl alcohol, then de-ionized water) and dried, and then a clean shadowmask (made from Be-Cu foil) which covered the region which would later be gated was affixed to the sample using photoresist as glue. Gold-germanium-nickel-gold[1] ohmic contacts were evaporated onto the sample, and then the shadowmask was removed in acetone, which dissolved the photoresist. See Figure A.1. The contact evaporation recipe was 5 nm of nickel, then 100 nm of gold-germanium, then 20 nm of nickel, and finally 100 nm of gold. The contacts were annealled in the rapid thermal annealler for 30 s at 420° C, although sometimes more annealling was necessary.

As mentioned above, the annealling was often unsuccessful because the well was very deep: approximately 680 nm below the surface. Although etching the ohmic contact regions would have improved the yield of working ohmic contacts, no etching was performed. The etching was avoided in part because of worries about the possibility of a pinhole in the photoresist which would allow the deposition of a small dot of gold-germanium-nickel-gold in the gate area, which would diffuse down and short the gate to the channel. (This was a worry because the titanium-gold gates which were used when I started studying these samples would short to the well after several measurements.) The worry about holes in the photoresist seems unfounded,

[1]Indium was tested as a replacement contact material for gold-germanium-nickel-gold, but the indium contacts had a much lower yield than the gold-germanium-nickel-gold. This may be due to the combination of the well being so deep below the surface of the sample (≈ 680 nm), and the indium process being more prone to contamination.

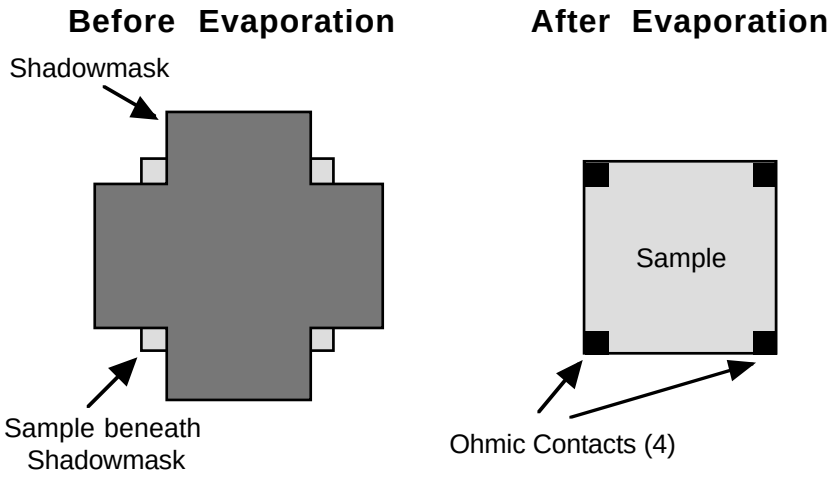

Figure A.1: The shadowmask used to evaporate the ohmic contacts onto the sample. On the left, the sample is shown under the shadowmask; only the corners which will have ohmic contacts are exposed. On the right, the sample is shown with the mask removed and the ohmic contacts in place.

as the aluminum gates have not shorted to the channel, even when the ohmic contacts were deposited using photolithography. It now seems more likely that the problems observed in the past with gates were caused by the use of titanium-gold gates, which do not form stable Schottky contacts. (See below.)

A.2 Gates

The gate is used to control the charge density in the well by applying a negative voltage to the gate, and to confine the FIR fields inside the sample. In addition, the gate provides a boundary condition at the well surface so that electric field couples strongly to the well. The gate was evaporated using a different shadowmask setup: a special sample holder was used, where the sample was glued in a recessed “trench” in the holder using photoresist. Then microscope coverslips were affixed to the surface

of the holder above the "trench" in such a way that they covered the ohmic contacts and a bit of the area around the ohmic contacts. (The coverslips were convenient because it was easy to look *through* them to see where the ohmic contacts were. It essential to prevent the gate from touching the ohmic contacts and shorting to the well.) See Figure A.2. Once the photoresist was baked, the entire unit was covered with 80 nm or more of evaporated aluminum.

Aluminum was chosen as a gate material because it has good adhesion properties on GaAs, and because the titanium-gold gates which are commonly used at UCSB proved unstable in nonlinear optical experiments. In his book on GaAs processing, Williams [26] mentions that Ti-Au gates are unstable, as the gold can (over time) diffuse through both the Ti and the GaAs to form a conducting path to the well. Williams states that a diffusion-preventing layer of platinum, molybdenum or palladium is needed between the titanium and the gold to prevent the gate from eventually shorting to the well.[2] The aluminum gates proved very stable; no degradation of the gates was ever observed.

After the gate evaporation, the sample was flipped over and a similar thick layer of aluminum was evaporated onto the entire back surface of the chip. This layer (like the gate) is used to confine the FIR inside the sample. (*Unlike* the gate, this layer is not used to control the charge density in the sample, and the boundary condition on this side of the sample is unimportant.)

---

[2]Other groups at UCSB using Ti-Au gates have not had problems with their gates diffusing and shorting to the channel. It is possible that these groups complete their measurements before the gates diffuse. Or it may be that the large electric fields and/or thermal energy which is present in the Free-Electron Laser beam accelerated the diffusion process and caused our gates to short to the well prematurely.

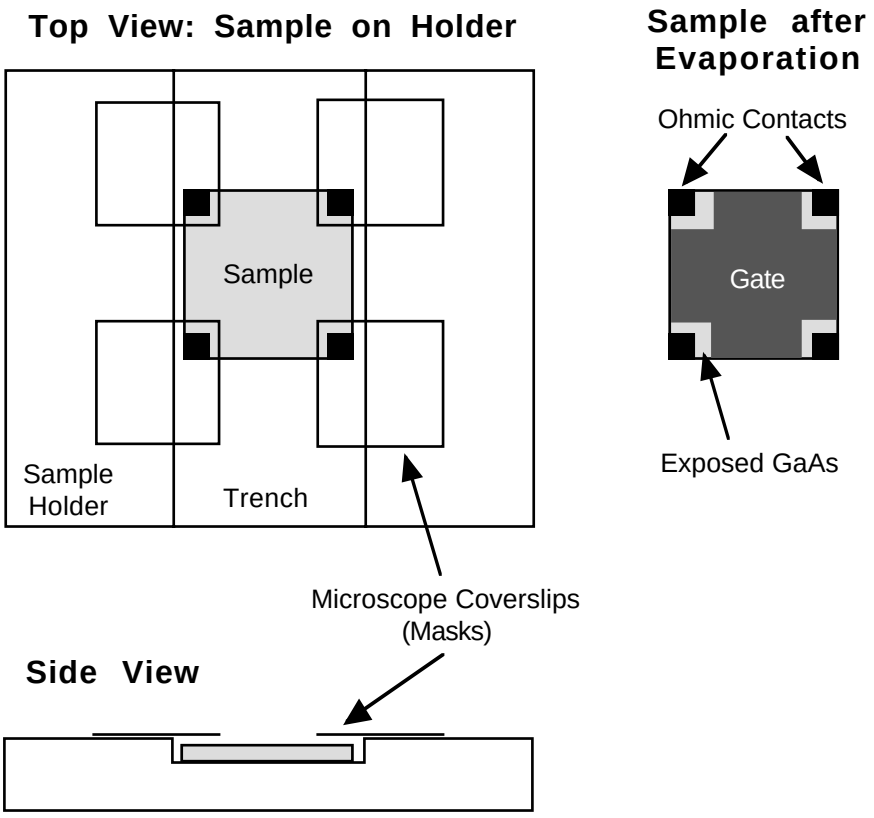

Figure A.2: The shadowmasks and sample holder used to evaporate the gate. The holder consists of a trench into which the sample is glued with photoresist. Then microscope coverslips are glued onto the holder to mask the ohmic contacts.

After all the processing was complete (which might include an anti-reflection coating as described below), the samples would have gold wires connected to the electrical contacts with silver paint, and be mounted on the optical sample holder. This optical sample holder would then be mounted in one of several cryostats, and could be used for linear or nonlinear measurements. The attachment method to the optical sample holder was to use a thin layer of silicone vacuum grease, then a very thin polypropylene sheet slightly larger than the sample, and then another thin layer of vacuum grease to attach the sample. The polypropylene is used to electrically insulate the sample from the holder. The vacuum grease is used as a "glue" because it can be removed easily and does not become brittle or fail at low temperatures.

Later, when samples were taken off their optical sample holders, it was observed that this back aluminum layer had developed holes and flaked off in places. It was possible that the vacuum grease may have reacted with the aluminum, or that mechanical abrasion in mounting, cooling, warming, and dismounting the sample had removed the aluminum in spots. Perhaps a more robust metallization would be appropriate here, although for optical measurements it is very important that the metal be highly conductive. If the metal is not highly conductive, then the metallic layer itself absorbs a significant fraction of the FIR over the length of the sample. Previous calculations estimated that titanium would be rather lossy if evaporated on the sample, and all the gold gate metallizations start with a layer of titanium.

A.3 Anti-Reflection Metallic Coating

The last processing step was the evaporation of an anti-reflection coating onto the edge of the sample, as described in Reference [16]. The purpose of the anti-reflection

coating was to reduce as much as possible the interference fringes which are caused by multiple reflections inside the sample. Due to the long length and high dielectric constant of the sample, the fringes are quite close together; they are spaced by:

$$\omega_{\mathrm{fringes}} = \frac{1}{2nD} \tag{A.1}$$

where n is the index of refraction and D is the length of the sample. For our sample, the spacing is approximately 0.13 cm^{-1}, which is on the order of the pulse-to-pulse variation in frequency of the Free-Electron Laser (FEL). The fringes are a problem because they make the transmission of the sample using the FEL *very* noisy, because the FEL frequency jitter is enough to change the transmission by a factor of two just due to the fringes.

The anti-reflection coating is based on the principle of impedance matching. By evaporating a film of resistance R_{film}, the reflections can be eliminated by matching the impedances as follows:

$$\frac{1}{R_{\mathrm{GaAs}}} = \frac{1}{R_{\mathrm{film}}} + \frac{1}{R_{\mathrm{free-space}}} \tag{A.2}$$

where $R_{\mathrm{GaAs}} = 106\ \Omega$ is the wave impedance of GaAs, and $R_{\mathrm{free-space}} = 377\ \Omega$ is the impedance of free space. The theory of this type of anti-reflection coating is fully described in Reference [16]. If the impedance matching condition is exactly met, then part of the radiation will be absorbed by the metallic film, and part will be transmitted, but nothing will be reflected, thus eliminating the interference fringes. Interestingly enough, the sign of the fringes changes as the resistance of the film varies through the impedance matching condition. Experimental data on test pieces of GaAs at normal incidence are shown in Figure 2.6.

Note (as is clear from Figure 2.6) that the absorptive anti-reflection coating substantially *reduces* the absolute transmission of the sample! In my tests, the fringes

were dramatically reduced, and I could afford to lose transmitted intensity, especially for the nonlinear measurements. The coating was very successful for my purposes because it greatly increased my signal to noise ratio.

One needed to exercise care when evaporating the anti-reflection coating. The desired resistance was $R_{\rm film} = 144\ \Omega$, which is a rather high resistance for a metallic film. Therefore, a poor conductor (NiCr) was chosen, which would allow the most allowance for thickness errors. The final procedure which worked well was to create a calibration GaAs[3] chip with gold contacts, shown in Figure A.3. The chip was designed to leave a 2 mm $\times$ 2 mm GaAs square exposed between the gold contacts. Thus, if the exposed GaAs square were exactly square, the resistance measured between the gold pads would be the sheet resistance (ohms per square) of the NiCr film between the pads. This chip was then placed in the evaporator along with the sample.

The sample was mounted on a special holder which allowed it to be placed on edge, so that the edge to be coated faced the evaporation source. The sample was "painted" with photoresist everywhere except on the edge to be coated, to prevent the NiCr from shorting out the contacts. The calibration chip was wired up in the evaporator using a 4-probe technique so that the resistance across the GaAs calibration chip could be accurately monitored during evaporation.[4] Then the evaporation was done slowly while monitoring the resistance of the chip; as soon as the "target" resistance was reached, the shutter was closed and the evaporation stopped. This

[3]A discussion with Axel Lorke led to the realization that NiCr reacts with GaAs; therefore the calibration chip needed to be GaAs. This way, the reaction on the sample would be similar to the reaction on the calibration chip.

[4]The thermal evaporator in the Engineering II Optoelectronics Lab was used, which had several vacuum feed-throughs for electrical connections.

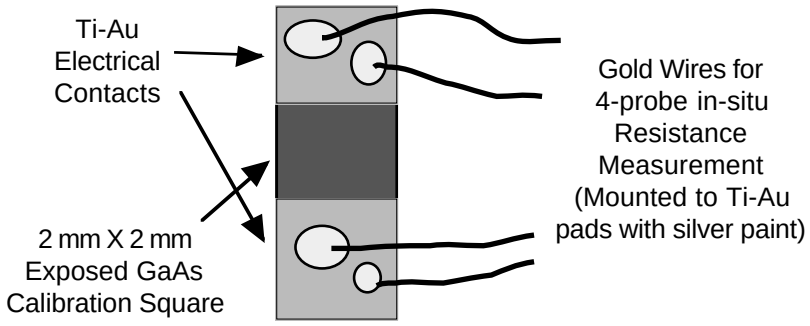

Figure A.3: The GaAs chip used to monitor the resistance of the NiCr film in the evaporator during anti-reflection coating evaporation. The chip had thick gold contacts (evaporated over titanium for adhesion) at both ends, such that a 2 mm $\times$ 2 mm GaAs square was left exposed between the gold contacts. Then the resistance across this square was monitored with a 4-probe measurement during the evaporation; see text for a description of the correction factors used.

worked well; the anti-reflection coatings produced were never perfect, but they were always close enough to reduce the fringes by a factor of five or more, which greatly reduced the noise problem due to frequency jitter.

A.4 Anti-Reflection Coating Calibrations

Several correction factors needed to be taken into account for the changes in resistance during the evaporation compared to the resistance during optical measurements. Some of these factors were found from several test pieces which were evaporated, allowed to oxidize, and then measured for resistance at 4.2 K.

First of all, the distances on the calibration chip were rarely exact (the GaAs "square" was not square), and therefore the ratio of the width of the rectangular GaAs region to the length of the GaAs region was measured under a microscope with a length scale in the optics. The geometrical factor G was the ratio of the length

over the width of the GaAs rectangle. The length was defined as the dimension in the direction which current flows, and the width was measured in the direction perpendicular to the current flow.

The temperature dependence of the film resistance needed to be taken into account. This was done by evaporating several test films, recording the resistance measured at room temperature after the sample had been allowed to oxidize (see below), and then cooling the samples to 4.2 K and measuring the resistance. The change in resistance factor (called f) from room temperature to 4.2 K was found to be $f = 0.9331$. (This means that the resistance at 4.2 K was actually *higher* than the resistance at room temperature!)

The difference in distance from the evaporation source to the sample edge and from the source to the calibration chip needed to be included in the calculations. Assuming the thickness of metal evaporated varies with the inverse square of the distance, a distance correction factor d was calculated. Given the measured distances of 17.5 cm from the evaporation boat to the calibration chip, and 16.9 cm from the boat to the sample edge, the distance correction factor was $d = (17.5^2)/(16.9^2) = 1.072$. This means that the coating was deposited onto the sample 7.2% faster than on the calibration piece.

The oxidation of the metal film after evaporation needed to be considered. The films were exposed to air for several days, and then the final resistance was measured.[5] These data were plotted, and average values for the changes were found. The

[5]The NiCr film was found to oxidize very rapidly at first (for several minutes), and then slow down and stabilize after a day or two. (The oxidation was fairly well fit by an exponential function of time.) Titanium was also tried as a coating material, but the titanium film continued to oxidize for many days with little sign of slowing, and was therefore deemed unsuitable.

correction formula was based in part on the assumption that oxidation would conceptually remove a constant thickness of conducting metal and replace it with non-conducting oxide. The thickness t of a metallic film was assumed to be inversely proportional to the resistance R: $R \propto 1/t$.

For the actual sample evaporation, a "target" resistance measurement was calculated based on the desired sheet resistance $R_{\rm goal}$, and the correction factors above. The evaporation was continued until the measured resistance had reached the calculated "target" value, $R_{\rm target}$. Then, after the sample had oxidized and cooled to 4.2 K, the sheet resistance of the edge would (hopefully!) be the desired value $R_{\rm goal}$.

Considering only oxidation, the target resistance $R_{\rm target}$ would be:

$$\frac{1}{R_{\rm target}} = \frac{1}{R_{\rm goal}} + t_0 \tag{A.3}$$

where t_0 is the thickness of the oxide layer in reciprocal resistance units.

Taking everything into account, the target resistance $R_{\rm target}$ was calculated as:

$$\frac{1}{R_{\rm target}} = \frac{1}{dGfR_{\rm goal}} + t_0 \tag{A.4}$$

$$R_{\rm target} = \left[\frac{1}{dGfR_{\rm goal}} + t_0\right]^{-1} \tag{A.5}$$

where $t_0 = 0.000295$ is now the average value of the measured changes due to oxidation, expressed as a thickness in reciprocal ohms.

Appendix B

Static Electric Field in Simulations

The static electric field across the well ($E_{\rm dc}$) is an important input parameter to a self-consistent simulation program, because the wide quantum structures which were studied are very sensitive to electric field. In this Appendix I will derive an expression to correctly include the fields at the well due to the charge on the gate when using a self-consistent program. (This result was originally derived by Dr. James Heyman using a different method.) I will also discuss the details of how the gate voltage causes the well to deplete.

I consider samples which have a front metallic gate and no back gate. Any electrons near the donors are assumed to be "frozen" into deep traps (DX centers) so that the electron/donor regions contain a fixed amount of positive charge which cannot move or change in any way. (The donors are also assumed to be far enough from the well so that they are not within the simulated region.)

The self-consistent simulation accepts an undoped potential (containing the well) and numerically adds electrons to the well and calculates the electric field due to the electrons, and their effect on the potential. (See Reference [7, Appendix B].) In addi-

tion to the potential, the static fields in the region of the well must be included in the calculation. The static electric fields tilt the potential of the well, changing the inter-subband separation and resonance frequencies. The total self-consistent potential in the region of the well can be written as:

$$V_T(z, V_G) = V_{\text{well}}(z) + \int \{E_{\text{built-in}} + E_{N_s}(z, V_G) + E_G(V_G)\}\, dz \quad \text{(B.1)}$$

$$= V_{\text{well}}(z) + zE_{\text{built-in}} + zE_G(V_G) + \int E_{N_s}(z, V_G)\, dz. \quad \text{(B.2)}$$

In this equation, z is the spatial coordinate in the growth direction, $V_T(z, V_G)$ is the total potential, $V_{\text{well}}(z)$ is the potential of the well (due to the bandgap engineering of the conduction band), and $E_{\text{built-in}}$ is the "built-in" (or intrinsic) electric field due to the ionized donors and other charges which are fixed in the sample. (Since the built-in field is caused by charges which are fixed, it is not dependent on the gate voltage and is a constant. [1]) $E_{N_s}(z, V_G)$ is the field due to the charged electrons in the well, and $E_G(V_G)$ is the field due to the charge on the gate.

The well potential $V_{\text{well}}(z)$ and the field due to the electrons $E_{N_s}(z, V_G)$ are both easily included in the simulation. The well potential is known from the growth program[2] of the sample, and the self-consistent program automatically includes the

[1]The field $E_{\text{built-in}}$ is mostly due to the net effect of the ionized donors near the well. Since the donors are outside the simulation region, their field is constant across the well. However, this field depends on the *difference* in number of ionized donors on either side of the well and seemed to change each time a sample was cooled. (The charge density N_s varied much less on different cooldowns since it depends only on the *total* number of ionized donors.) Because of this variability, the built-in field needed to be fit independently for each cooldown.

[2]The input potential of the well was varied from the growth program within a narrow range of parameters to take into account small calibration errors in the growth rate.

field due to the electrons in the well. Since the built-in field $E_{\rm built-in}$ is just a constant found by fitting the data, the only parameter left to specify is the field due to the charge on the gate, $E_G(V_G)$. In order to calculate this, I will consider in detail how the gate depletes the well.

The gate and the well are similar to a parallel-plate capacitor with gate area A, plate separation d, and filled with dielectric[3] ϵ. Unlike a capacitor, the sample has fixed charges near the well (due to ionized donors) which attract electrons into the well at zero volts bias. (A capacitor has equal and opposite charges on its plates which are proportional to the applied voltage; if the voltage is zero, there is no net charge on either plate.) A normal capacitor has:

$$Q = CV \tag{B.3}$$

where Q is the charge on one plate (the charge on the other plate is $-Q$), $C = (\epsilon A/4\pi d)$ is the capacitance, and V is the voltage across the plates. In the sample, the charge in the well is given by:

$$Q_{\rm well} = -Q_0 - CV_G \tag{B.4}$$

where $Q_{\rm well}$ is the total charge in the well, $-Q_0$ is the charge in the well with zero volts gate bias, and $-CV_G$ is the extra positive charge in the well due to the gate bias ($V_G \leq 0$). The voltage of the gate relative to the grounded well is V_G. Note that the

[3]The electric field in a voltage-biased dielectric-filled capacitor is $E = (V/d)$, where V is the applied voltage and d is the separation of the plates. Although the dielectric increases the capacitance by ϵ, if the capacitor is biased with a voltage, then it will still have the applied voltage V across it (and field (V/d) inside it). However, the voltage-biased capacitor will have a factor of ϵ more *charge* than if the dielectric had been absent.

charge on the gate is simply:

$$Q_{\rm gate} = +CV_G. \tag{B.5}$$

Note the signs of the charges in Equation (B.4) and Equation (B.5). (All symbols in these equations *except* V_G are taken to be positive.) The charge on the gate is negative, and the well has less negative charge in it as the gate voltage becomes more negative.

Now I will calculate the field E_G caused by the charge $Q = CV_G$ on the gate. For a moment ignoring the dielectric on one side of the gate and using Gauss's Law on a Gaussian "pillbox" which passes through the gate,

$$\int_{\rm surface} E_G \, dA = 4\pi Q_{\rm enclosed}. \tag{B.6}$$

Using Equation (B.5), the field through the ends of the "pillbox" is

$$E_G = \frac{4\pi CV_G}{2A}. \tag{B.7}$$

Now, imagine bringing the dielectric slab up to the charged gate; the field inside the dielectric will just be the field in Equation (B.7) divided by ϵ, the dielectric constant of the semiconductor. Thus, the field from the charged gate *inside* the sample is:

$$E_G = \frac{4\pi CV_G}{2\epsilon A} = \frac{V_G}{2d}. \tag{B.8}$$

Equation (B.8) is the value for the field due to the charge on the gate to be used in the self-consistent simulation. The simulation program I used allows the user to enter the conduction band edge potential ($V_{\rm well}(z)$) from a file, and then specify the total static electric field across the well as a separate input parameter. In this case, the total static field parameter ($E_{\rm dc}$) for the self-consistent simulation is the sum of the fields due to the built-in field and the field due to the charges on the gate:

$$E_{\rm dc} = E_{\rm built-in} + E_G(V_G). \tag{B.9}$$

In order to better understand how the sample depletes, consider the electric fields in the sample. Much like a capacitor, the fields due to the changing charge in the well and the gate cancel out everywhere except between them. (This ignores fringing fields outside the gated area.) Using the expression for the total field (the expression in $\{\cdots\}$ in Equation (B.1)), consider the field in the sample at $V_G = 0$:

$$E_{\text{total}}(V_G = 0) = \{E_{\text{built-in}} + E_{N_s}(z, V_G) + E_G(V_G)\} \quad \text{(B.10)}$$

$$= \left\{E_{\text{built-in}} \pm \left(\frac{2\pi Q_0}{\epsilon A}\right) + 0\right\} \quad \text{(B.11)}$$

where the $\pm$ indicates that the field from the electrons in the well points toward the well, and the 0 represents the lack of field from the uncharged gate. If a gate voltage $V_G = V$ is applied, there is then a net field of

$$E_{\text{total}}(V_G = V) = \{E_{\text{built-in}} + E_{N_s}(z, V) + E_G(V)\} \quad \text{(B.12)}$$

$$= \left\{E_{\text{built-in}} \pm \left(\frac{2\pi Q_0}{\epsilon A} - \frac{V}{2d}\right) - \left(\frac{V}{2d}\right)\right\} \quad \text{(B.13)}$$

Looking at the change in field between $V_G = 0$ and $V_G = V$,

$$\triangle E_{\text{total}} = \pm\left(\frac{V}{2d}\right) + \left(\frac{V}{2d}\right) = \begin{cases} (V/d) & \text{between gate and well} \\ 0 & \text{elsewhere.} \end{cases} \quad \text{(B.14)}$$

Other than the offset fields and charges ($E_{\text{built-in}}$ and fields due to $-Q_0$), the fields in the sample behave identically to a capacitor; the fields from the opposite charges in the well and gate add constructively between the gate and well and cancel outside that region. In this picture, the potential at the back of the well (far side from the gate) does not change, as the well is electrically grounded. The front side of the well rises when a negative gate voltage is applied, and the gradient in the field (between the tilting region between the gate and the well and the non-changing region behind

the well) occurs across the charged well. Therefore, as the well depletes, the front end is being lifted while the bottom of the well is being bent, so that the height of the potential at the back of the well never changes. See Figure B.1.

This leads to an intuitive picture of how the well depletes: as a negative gate bias is applied, the well is tilted and bent upward. As the center of the well rises, *and* as the potential becomes tilted, the subband energies will increase, relative to the constant chemical potential of the electrons. (The electron's chemical potential is fixed by grounding the ohmic contacts to the well.) As the subbands rise, the electrons which find themselves above the chemical potential flow out of the sample through the ohmic contacts. This is the mechanism by which the well depletes.

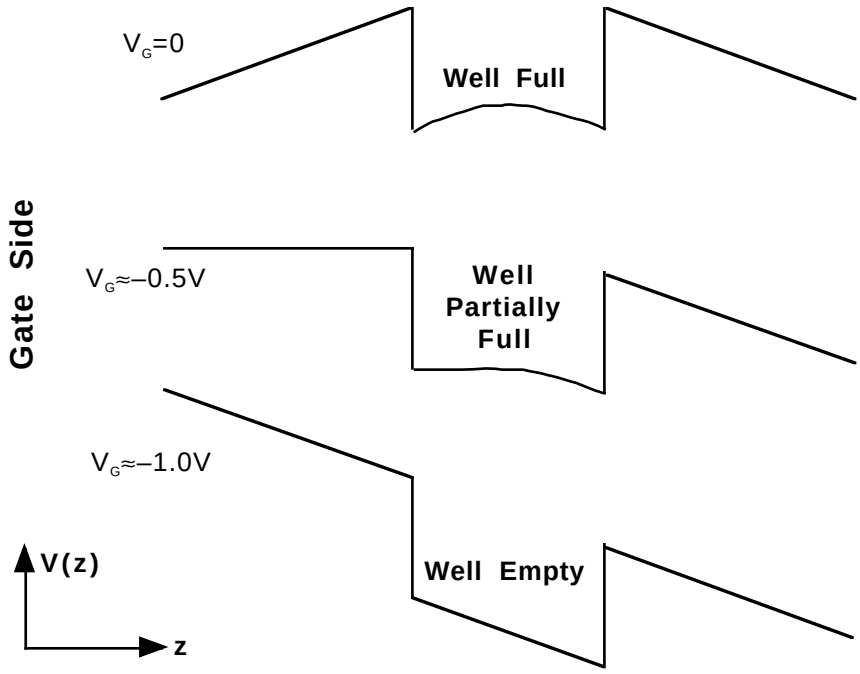

Figure B.1: Cartoon of the well depleting; the vertical axis is potential energy ($V(z)$), and the horizontal axis is the growth direction z. At the top, the well is full, and the static field across the well is due to the electric field from the electrons in the well. ($E_{\mathrm{built-in}}$ is taken to be zero.) As the well depletes, the back of the well stays at a fixed potential, and the front is lifted while the bottom bends. (Gate voltage values are fictitious.) Note that this diagram depicts the potential energy for *electrons*.

Appendix C

Lorentzian Lineshape Functions

This Appendix is designed to clarify the details of frequency dependence of the Lorentzian absorption lineshape. The Lorentzian lineshape is the form of the absorption of a simple-harmonic oscillator, as well as the intersubband resonance I have observed. There is some confusion due to the fact that both the classical and quantum mechanical lineshapes are called “Lorentzian”, although they are different. However, they *are* similar if the linewidth is much less than the resonance frequency, which is easily the case for our samples. In addition, there are Lorentzian lineshape functions in the literature which are approximate forms valid only near the resonance. An example of this is the rotating wave approximation (abbreviated RWA) which is often implicit in the quantum mechanical function. In this Appendix I will present a derivation of both the full and approximated forms of the classical absorption lineshape, calculate the area under the curve, and give the forms of the fitting functions used for data analysis.

C.1 Full Absorption Lineshape

To derive the classical frequency dependence, consider a classical damped, driven simple harmonic oscillator. Given a mass m, a natural frequency ω_0, and a damping time of τ, the equation of motion of the oscillator is:

$$m\frac{d^2x}{dt^2} + \frac{2m}{\tau}\frac{dx}{dt} + m\omega_0^2 x = eE(t) \tag{C.1}$$

where x is the displacement of the oscillator, e is the charge of the electron, and $E(t)$ is the time-dependent electric field which drives the oscillator. Assuming a time-dependence of the electric-field of $E(t) = E_0 \exp(i\omega t)$, I can try a solution of the form $x(t) = x_0 \exp(i\omega t)$. Plugging this in to Equation (C.1), and solving for x_0:

$$x_0 = \frac{eE_0/m}{(\omega_0^2 - \omega^2) + 2i\omega/\tau}. \tag{C.2}$$

The absorption is proportional to the conductivity $\sigma(\omega)$, which is:

$$\sigma(\omega) = \frac{Ne}{E}\left(\frac{dx}{dt}\right) \tag{C.3}$$

where N is the 3-D electronic density, and E is the electric field. Plugging $x(t)$ into the conductivity yields:

$$\sigma(\omega) = \frac{i\omega N e^2/m}{(\omega_0^2 - \omega^2) + 2i\omega/\tau}. \tag{C.4}$$

Equation (C.4) above is the classical conductivity of a simple-harmonic oscillator. One way to convert this to the approximate quantum-mechanical oscillator conductivity is to insert the quantum-mechanical oscillator strength[1] $f_{12} = 2m\hbar^{-2}E_{12}|\langle z_{12}\rangle|^2$.

[1]Note that the E_{12} in the oscillator strength is the energy separation of the subbands, *not* $\widetilde{E}_{12}$ which includes the depolarization shift. This is because the many-body interactions of the electrons increase the absorption *frequency*, but do not increase the absorption *strength*.

(In the classical case $f_{12} = 1$ exactly, while the quantum-mechanical oscillator strength is always less than or equal to 1.) Inserting f_{12}, multiplying top and bottom by $\hbar^2$, and rewriting,

$$\sigma_{zz}(\omega) = \frac{ie^2 N_s f_{12}\hbar}{m}\left(\frac{\hbar\omega}{(\widetilde{E}_{12}^2 - (\hbar\omega)^2) + 2i\omega\hbar^2/\tau}\right) \tag{C.5}$$

where I have changed notation slightly: $\sigma_{zz}(\omega)$ is the zz element in the conductivity tensor. I have also changed the resonant energy to $\widetilde{E}_{12} = \hbar\widetilde{\omega}_{12}$, where the tilde means that the depolarization shift has been included in this resonance frequency, and the subscript "12" denotes the transition between the two lowest subbands. In addition, the mass m is now the effective mass of electrons in the conduction band of GaAs.

There are several ways to denote the linewidth: either in terms of the relaxation times which are phenomenologically associated with the linewidth, or in terms of the actual frequency (or energy) span of the peak. I will follow the notation of Załużny [31] and use Γ as the half-width at half-max energy-span of the peak. Using the definition I have used for τ, the relation between τ and Γ is:

$$\Gamma = \hbar/\tau. \tag{C.6}$$

The conductivity can be written in terms of Γ as:

$$\sigma_{zz}(\omega) = \frac{ie^2 N_s f_{21}\hbar}{m}\left(\frac{\hbar\omega}{(\widetilde{E}_{12}^2 - (\hbar\omega)^2) + 2i\hbar\omega\Gamma}\right) \tag{C.7}$$

which is the same as Załużny has in Equation 10 of Reference [31] (other than two minus signs that cancel out later, and a typographical error in Reference [31] in which the wrong quantity is squared in the denominator).

The real part of the conductivity is proportional to the absorption (at least in the simple perturbative waveguide):

$$\mathbf{Re}[\sigma_{zz}(\omega)] = \frac{2e^2 N_s f_{12}\hbar\Gamma}{m}\left(\frac{(\hbar\omega)^2}{(\widetilde{E}_{12}^2 - (\hbar\omega)^2)^2 + 4(\hbar\omega)^2\Gamma^2}\right). \tag{C.8}$$

The absorption coefficient of the waveguide in the linear perturbative approximation[2] is given by: [31]

$$\alpha(\omega) = \frac{1}{a}\frac{4\pi}{c\sqrt{\epsilon}}\mathbf{Re}[\sigma_{zz}(\omega)] \tag{C.9}$$

where a is the width of the waveguide (in this case, the thickness of the GaAs substrate, approximately 0.45 mm), and $\alpha(\omega)$ is the attenuation coefficient for radiation propagating down the waveguide. Putting this all together gives the final result without approximations:

$$\alpha(\omega) = \frac{8\pi e^2 N_s f_{12}\hbar\Gamma}{cam\sqrt{\epsilon}}\left(\frac{(\hbar\omega)^2}{(\widetilde{E}_{12}^2 - (\hbar\omega)^2)^2 + 4(\hbar\omega)^2\Gamma^2}\right). \tag{C.10}$$

C.2 Rotating Wave Approximation

I will now derive the rotating wave approximation, or RWA form of the Lorentzian often seen in the literature. Consider Equation (C.7) for the complex conductivity $\sigma_{zz}(\omega)$; I can rewrite it as:

$$\sigma_{zz}(\omega) = \frac{ie^2 N_s f_{21}\hbar}{m}\left(\frac{\hbar\omega}{(\widetilde{E}_{12} - \hbar\omega)(\widetilde{E}_{12} + \hbar\omega) + 2i\hbar\omega\Gamma}\right). \tag{C.11}$$

Using the approximation that $\hbar\omega \approx \widetilde{E}_{12}$, I can approximate the second factored term in the denominator as:

$$(\widetilde{E}_{12} + \hbar\omega) \approx 2\hbar\omega. \tag{C.12}$$

Rewriting and canceling,

$$\sigma_{zz}^{\mathrm{RWA}}(\omega) = \frac{ie^2 N_s f_{21}\hbar}{2m\Gamma}\left(\frac{1}{(\widetilde{E}_{12} - \hbar\omega)\Gamma^{-1} + i}\right). \tag{C.13}$$

[2] See Chapter 4 for a more in-depth discussion of waveguides.

Now, taking the real part and using Equation (C.9) to calculate the attenuation coefficient, I find that:

$$\alpha^{\mathrm{RWA}}(\omega) = \frac{2\pi e^2 N_s f_{12} \hbar \Gamma}{cam\sqrt{\epsilon}} \left(\frac{1}{(\widetilde{E}_{12} - \hbar\omega)^2 + \Gamma^2} \right). \tag{C.14}$$

Equation (C.14) is the attenuation coefficient in the rotating wave approximation for the linear perturbative absorption of the waveguide.

Note that this equation is incorrect in Reference [10]; the oscillator strength f_{12} was incorrectly expanded as:

$$f_{12} \neq \frac{2m\hbar\omega|\langle z_{12}\rangle|^2}{\hbar^2} = \frac{2m\omega|\langle z_{12}\rangle|^2}{\hbar} \tag{C.15}$$

instead of the correct form:

$$f_{12} = \frac{2mE_{12}|\langle z_{12}\rangle|^2}{\hbar^2} \tag{C.16}$$

(see Equation (1) in Reference [10], with the understanding that Equation (1) is for the absorption cross-section σ as opposed to the conductivity σ. In addition, Reference [10] uses a linewidth parameter T_2 which is related to Γ by $\Gamma = \hbar/T_2$.) The error is that the incorrect relation $E_{12} = \hbar\omega$ is used instead of $E_{12} = \hbar\omega_{12}$, where ω_{12} is the frequency separating the subbands.

C.3 Area under the Curve

Next, I shall outline the calculation of the area under the curve of Equation (C.10). This can be written as:

$$A = \int_0^\infty \alpha(\omega)\, d\omega \tag{C.17}$$

This integral can most simply be evaluated using contour integration in the complex ω-plane. First I will introduce some simplifying notation and change variables:

$$\beta \equiv \left(\frac{8\pi e^2 N_s f_{12} \hbar \Gamma}{cam\sqrt{\epsilon}} \right) \tag{C.18}$$

$$E \equiv \widetilde{E}_{12} \tag{C.19}$$

$$z \equiv \hbar\omega \tag{C.20}$$

The integral A can now be concisely written:

$$A = \frac{\beta}{\hbar} \int_{z=0}^{z=\infty} \frac{z^2}{(E^2 - z^2)^2 + 4\Gamma^2 z^2} \, dz \tag{C.21}$$

Note that the integrand is only a function of z^2; I can extend the lower limit of integration to $z = -\infty$ and divide the result by 2:

$$A = \frac{\beta}{2\hbar} \int_{z=-\infty}^{z=\infty} \frac{z^2}{(E^2 - z^2)^2 + 4z^2\Gamma^2} \, dz. \tag{C.22}$$

I will close the contour in the upper-half complex z-plane. The only thing left is to find the poles and their residues. Without showing all the details, I will state that the poles in the upper-half plane are at:

$$z_{1,2} = i\Gamma \pm \sqrt{E^2 - \Gamma^2}. \tag{C.23}$$

Each of the two poles are simple poles. Therefore, since the integrand can be written as $f(z) = q(z)/p(z)$, the residue can be found from:

$$a_{-1} = \left. \frac{q(z)}{p'(z)} \right|_{z=z_{1,2}} \tag{C.24}$$

(See any book on complex analysis, for example Mathews and Walker [15, section 3-3].) For this integrand, the function to plug the poles into is:

$$\frac{q(z)}{p'(z)} = \frac{z^2}{z(8\Gamma^2 - 4E^2 + 4z^2)} \tag{C.25}$$

Plugging in the poles, the residues are:

$$a_{-1}^{(1)} = \frac{i\Gamma + \sqrt{E^2 - \Gamma^2}}{8i\Gamma\sqrt{E^2 - \Gamma^2}}$$

$$a_{-1}^{(2)} = \frac{-i\Gamma + \sqrt{E^2 - \Gamma^2}}{8i\Gamma\sqrt{E^2 - \Gamma^2}}.$$

The integral is given by $2\pi i \sum(\text{residues})$:

$$A = \frac{\beta}{2\hbar}\left(2\pi i\left[\frac{1}{4i\Gamma}\right]\right) = \frac{\beta\pi}{4\hbar\Gamma}. \tag{C.26}$$

Putting back the physical units and full notation, the area under the curve is:

$$A = \left(\frac{8\pi e^2 N_s f_{12}\hbar\Gamma}{cam\sqrt{\epsilon}}\right)\left(\frac{\pi}{4\hbar\Gamma}\right) = \frac{2\pi^2 e^2 N_s f_{12}}{cam\sqrt{\epsilon}}. \tag{C.27}$$

Note that this area is in cgs units, so the units are radians/(cm-s). To convert to units of (1/cm^2), multiply A by the conversion factor:

$$\left(\frac{1}{2\pi}\right)\left(\frac{8065.4\ \text{cm}^{-1}}{2.41796 \times 10^{14}\ \text{Hz}}\right) = 5.309 \times 10^{-12}(\text{cm}^{-1}/\text{Hz}) \tag{C.28}$$

Putting Equation (C.27) together with the conversion factor and inserting the physical constant values, yields the following handy rule:

$$A = (1.756 \times 10^{-16}\ \text{cm}/\text{Å}^2) N_s E_{12} |\langle z_{12}\rangle|^2. \tag{C.29}$$

where A is in units of (1/cm^2), N_s is in units of (1/cm^2), E_{12} is in wavenumbers (cm^{-1}), and $\langle z_{12}\rangle$ is in units of Å.

C.4 Fitting Functional Forms

The most common application of the Lorentzian and RWA approximated attenuation coefficient functions is to fit experimental data. The full Lorentzian form of the fit

function (from Equation (C.10)) is:[3]

$$\alpha_{\mathrm{Fit}}(C,H,W,\Omega) = H\left(\frac{\Omega^2 W^2}{(C^2-\Omega^2)^2+\Omega^2 W^2}\right) \tag{C.30}$$

where C (center frequency), H (peak height), and W (full-width at half-max) are free parameters to be determined by the data. (The frequency in wavenumbers is denoted Ω.) The fitting function for the RWA form (Equation (C.14)) is:

$$\alpha_{\mathrm{Fit}}^{\mathrm{RWA}}(C,H,W,\Omega) = \frac{H}{4}\left(\frac{W^2}{(C-\Omega)^2+W^2/4}\right) \tag{C.31}$$

The full and RWA expressions for the attenuation coefficient are compared in Figure C.1.

The area under the curve of the full fit function (Equation (C.30)) is

$$A = \left(\frac{\pi}{2}\right) HW. \tag{C.32}$$

(This result is found from a contour integration similar to the last section.) If H and W are in units of inverse centimeters, then A is in units of (1/cm^2).

The area under the RWA curve is not a useful quantity, since the approximation is only valid near the resonance. However, the RWA fit function can be used to obtain the fit parameters from the data near the resonance, and then these parameters can be used to find the area using the full expression (Equation (C.32)), since the RWA and full expressions share the same set of fit parameters.

---

[3]Note that the factor of 4 in the denominator of Equation (C.10) has vanished because the parameter W is now the *full*-width at half-max, whereas the parameter Γ in Equation (C.10) was the *half*-width at half-max. ($W = 2\Gamma$)

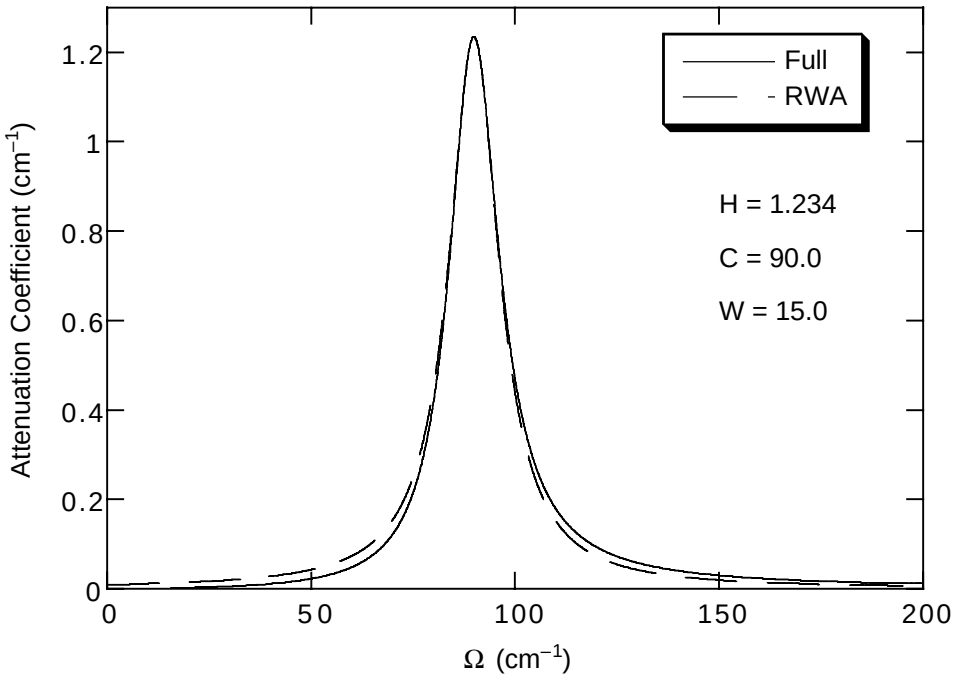

Figure C.1: A comparison of the attenuation coefficient of the full and RWA fit expressions (Equations (C.30) and (C.31)) for the parameter values shown. The RWA approximation overestimates the attenuation coefficient at frequencies below the resonance, and underestimates the attenuation coefficient at frequencies above the resonance.

Appendix D

Linear Absorption Measurements

The linear absorption measurements involved the construction of a custom sample-detector cryostat which was mounted to the Fourier-Transform Infrared Spectrometer (FT-IR). This cryostat contained a Winston cone, the sample, a detector, and filters, a polarizer, and optical coupling between them all. This Appendix describes the Winston cone which was used to concentrate the FIR onto the sample edge, and then describes the rest of the optics in the cryostat and how they worked together. Finally, the techniques used to do linear measurements are described.

D.1 Winston Cone Concentrator

A Winston cone [27, 28] was used to concentrate the diffuse FIR from the FT-IR onto the edge of the sample for the linear absorption measurements. The cone is a rotationally symmetric off-axis parabolic reflector, with the inside coated with gold.

The operation of the cone is easiest to conceptualize in the limit of geometrical optics. However, most measurements were performed at approximately $125\ \mu m$ wavelength, and the exit of the cone is only $480\ \mu m$ in diameter, so only a handful

of wavelengths could fit across the exit of the cone. Hence the geometrical optics approximation is not necessarily valid. However, a physical optics analysis of the complicated shape of the cone and non-trivial input waves from the spectrometer would be very difficult; therefore this discussion will limit itself to geometrical, or "ray," optics.

The cone accepts most of the radiation within a certain acceptance angle, and forces that radiation to exit through the back exit of the cone. A small amount of the radiation within the acceptance angle, and most of the radiation outside, is reflected back out the front of the cone.

Different rays entering the cone undergo a different number of reflections before they exit, causing any optical information entering the cone to be scrambled and destroying the possibility of an image. In fact, although the rays entering the cone are reasonably well collimated (within 14° of each other), the rays which exit can make *any* angle up to 90° to the central axis of the cone. The fact that the rays emerge from the back of the cone into 2π steradians meant that it was important to physically position the cone as close as possible to the edge of the sample, to get the maximum possible FIR into the sample.

The cone used[1] in my setup had an acceptance angle of 6.5° from the central axis, and a length of 0.835 inches. The diameter at the front was 0.174 inches, and the diameter at the exit was 0.019 inches. The acceptance angle was chosen to closely match the optics of the Bomem FT-IR I used, which had a beam in the sample chamber which came to a 1.0 cm diameter waist with f/4 optics,[2] giving a maximum entry

[1] I purchased the cone from I.R. Labs, 1808 E. 17th Street, Tucson, AZ 85719. (602) 622-7074.

[2] f/4 refers to the ratio of the *focal length* to the *diameter* of the beam being 4. By diameter, I mean the diameter of the beam at the focusing element (the "clear aperture").

angle of $\text{Tan}^{-1}(1/8) = 7.1°$. This cone should theoretically therefore be able to catch most of the beam which falls into the entrance. The concentration of FIR achieved by the cone is roughly the ratio of the area of the entrance divided by the area of the exit, or $(0.174/0.019)^2 = 84$, which is quite good. In reality, loss of light at the cone exit due to the wide range of angles emitted by the cone will somewhat reduce the effective improvement. This concentration of the FT-IR beam is important because the energy in the FT-IR beam is distributed inside a circle of diameter 1.0 cm, and the black-body source in the FT-IR is quite dim at long wavelengths. By increasing the amount of energy I was able to funnel into the sample, I was able not only to acquire data more rapidly, but also to acquire data of much higher quality than would otherwise have been possible.

The shape of the particular cone used was described by the curve:

$$y(x) = \frac{-d_2}{2} - x\,\tan\theta + \frac{d_2\left(\sin\theta - \sqrt{1 + 2x\frac{1-\sin\theta}{d_2\cos\theta}}\right)}{\sin\theta - 1} \qquad \text{(D.1)}$$

where θ is the maximum acceptance angle in radians ($6.48° = 0.113$ rad), and d_2 is the diameter of the exit of the cone. The function $y(x)$ gives the radius of the cone at any distance x from the exit of the cone. The three-dimensional shape of the cone is obtained by rotating the function in Equation (D.1) about the x-axis. A cross-section of the cone is shown in Figure D.1.

D.2 Cryostat and Detector

The cryostat containing the Winston cone, sample, and detector was a model HD-3 constructed by I.R. Labs. It contains a liquid nitrogen cooled radiation shield which surrounds a liquid helium cooled cold-plate. Mounted on the cold-plate are, in the

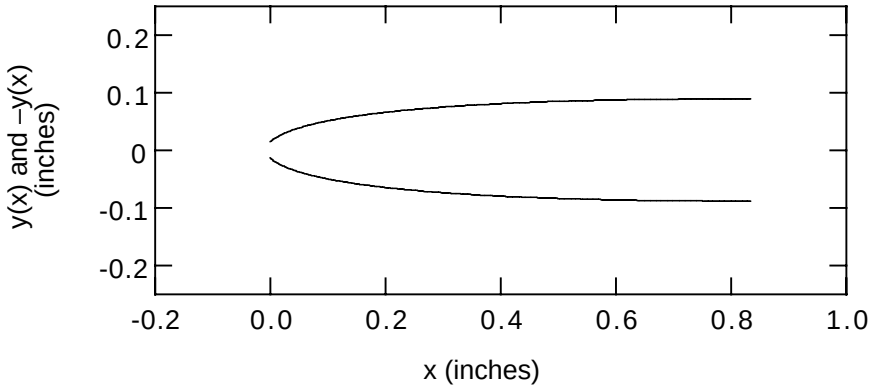

Figure D.1: A cross section of the Winston cone. The plot shows $y(x)$ and $-y(x)$ from Equation (D.1). (The actual cone shape would be obtained by rotating the curve $y(x)$ about the x-axis.) The aspect ratio shown is correct.

order they are encountered by the radiation: filters, the Winston cone, the sample, a collecting optical "cavity" for the detector with a polarizer, and the detector. (See Figure 3.2.)

The liquid helium reservoir and cryostat are designed to facilitate pumping on the space above the helium without excessive helium loss. The cold plate can achieve temperatures as low as 1.5 K.

The detector was a germanium bolometer. Several modifications needed to be made to it, including the use of an external load resistor for the bias circuit. (The detector housing included a load resistor, but this resistor was both far too low in resistance and far too noisy to be useful for precise measurements.)

The cryostat was placed where a detector could normally go on the FT-IR spectrometer, at the side exit-port of the Bomem. (This was a custom port which was built here at UCSB by another research group.) The first object in the optical path

from the outside is the vacuum window. It was made of wedged[3] white polyethylene (white poly), which was strong enough to withstand the force of the vacuum. It was also opaque to visible radiation, reducing the heat load on the interior of the cryostat. Next, inside the cryostat, was the entrance to the Winston cone, with wedged filters on the front. The cone was mounted using baffles so that radiation could not leak between the cone and the nitrogen shield. The two filters were calcium-fluoride (CaF_2) and z-cut crystal quartz. When cooled to low temperatures (as during operation) the filters absorbed (or reflected) all the radiation from 200 cm^{-1} up to high enough frequencies that the white polyethylene vacuum window was opaque. Thus, the only radiation reaching the inside of the nitrogen shield had frequencies below 200 cm^{-1}.

The cone concentrated the FIR onto the sample edge. The diameter of the exit of the cone was equal to the thickness of the sample, and care was taken to mount the sample in the cryostat such that the sample edge was very close to the cone. (Note that the cone would change the polarization of the radiation if it was polarized initially. However, the output of the FT-IR source is unpolarized, and the radiation acquires only a slight polarization due to the optics of the spectrometer.)

Although some of the radiation emerged from the cone at a very large angle, it was refracted very close to the optical axis (the axis defined as going through the center of the cone) as it entered the GaAs. From Snell's Law, $n_1 \sin\theta_1 = n_2 \sin\theta_2$ (where n_i is the index of refraction of medium i and θ_i is the angle from the normal of the ray in medium i) the maximum limiting angle a ray could have after entering

[3]The window and the filters were *wedged*: they were manufactured so that the front and back surfaces were not exactly parallel but instead had a small (several degrees) angle relative to each other. This ensures that the interference fringes from different points on the window (or filter) will have different periods, and will cancel out (on average).

the GaAs was

$$\theta_2 = \text{Sin}^{-1}\left(\frac{\sin 90^\circ}{n_2}\right) = 16.1^\circ$$

where the index of GaAs $n_2 = 3.6$. Thus, as long as the radiation intersected the GaAs, it was refracted fairly close to the normal. In the interest of intersecting the maximum amount of radiation, the exit of the cone was placed as close as possible to the sample.

The radiation then passed through the sample, being absorbed by the intersubband absorption at some frequencies and not as much at others. The transmitted radiation next emerged from the opposite edge of the sample. Although it was refracted close to the normal upon entering the GaAs, the slight angle it retained is enough to substantially refract it away from the normal upon leaving the GaAs. In fact, as long as the edges of the sample were parallel, it was refracted back to its original angle before it entered the GaAs, so it could have emerged from the sample at any angle up to 90°.

The detector mount consisted of a copper piece with a hollow space to which the bolometer was mounted. The hollow space was sealed on the front with a piece of copper-beryllium foil, which had a slit sawed in the middle. The slit was 0.019 inches wide, which was the same width as the sample, and roughly 3 mm long. (The slit was long to allow for the collection of all the radiation which had spread out in the theoretical 16.1° wedge from the cone. The sample was planned to be roughly 5 mm long, and 5 mm high, with at least a 3 mm wide gate, so that the entire "wedge" of radiation would remain in the gated region.) Behind the foil, inside the hollow space was an FIR polarizer situated such that it passed only the electric field parallel to the growth direction. Since the electric field perpendicular to the growth direction cannot couple to the intersubband absorption, by eliminating it I was able to roughly double

the signal/noise ratio. In addition, I painted the inside of the hollow space (only on the side before the polarizer) with a black epoxy known to be fairly absorbent in the FIR. This epoxy ensured that reflections before the polarizer would not rotate the polarization and allow the unwanted polarization to pass. The detector covered the entire back exit of the hollow space.

The polarizer was a 1000 line per inch gold on mylar polarizer.[4] (It consisted of a sheet of mylar on the order of 0.010 inches thick, with gold lines evaporated onto the surface.) It had good transmission of the desired polarization, as well as excellent rejection of the undesired polarization.

The setup I built and used has several advantages over our group's old setup, as described in Chris Felix's thesis [5]. The main advantage is that I have greatly improved the optical efficiency by having the sample and the detector in the same cryostat. (Our old setup used one cryostat for the sample, and a separate cryostat for the bolometer.) By optical efficiency, I mean that most of the radiation which reaches the detector in the new system passes through the sample, and conversely, the radiation reaching the sample has a close to optimum chance of reaching the detector.

[4]The polarizer was purchased from Buckbee-Mears, 245 E. 6th Street, 6th Floor, St. Paul, MN 55101, (612) 228-6581.

D.3 Experimental Procedure

I would begin the measurements with the sample in the dark inside the cryostat, with the ohmic contacts shorted to the gate. Due to the metastability of the donors[5] in the sample, it was suspected that the most repeatable results were obtained when the sample had been in the dark continuously for several weeks before measurements were done. The white polyethylene vacuum window prevented visible light from entering the cryostat once it was assembled. Shorting the ohmic contacts to the gate ensured that the gate was not holding a static charge.

The details of the cooling procedure may affect the final state of the sample. Therefore, the cooling procedure was repeated as consistently as possible: first both the helium and nitrogen reservoirs were filled with liquid nitrogen and allowed to cool for approximately one hour. Then the nitrogen was poured out, and the helium and nitrogen spaces were filled (with helium and nitrogen respectively). Finally, the helium reservoir was connected to a pumping system and slowly pumped down for approximately one hour from atmospheric pressure to several Torr.

Once the sample was cold, the charge density was measured using capacitance-voltage measurements ($C(V)$). The phase angle of the current response of the sample was compared with a reference capacitor, and the capacitance was checked for repeatability. Any deviations from the normal behavior indicated that something (usually an electrical contact to the sample) was not functioning correctly. (It was normal

[5]Visible radiation can ionize the donors, and this charge can become spatially separated from the ionized donors due to the bandgap engineering. Due to the large distances the charges can move, and the behavior of the donors, the time to return to thermodynamic equilibrium can be many hours, or perhaps even days!

for the first $C(V)$ scan to be slightly different from the others, usually having slightly more charge than the later ones. This occurred by the depletion voltage being slightly more negative for the first $C(V)$ scan. It was only considered a problem if the $C(V)$ scans continued to change after the first time.)

Next, a phase correction spectrum [3] was measured with the sample depleted. Measuring the phase correction using the depleted sample may have caused some baseline problems in the spectra, because our samples are highly absorptive, and I symmetrized all the interferograms with the depleted phase correction spectrum. This problem could have been solved by taking separate phase corrections for each gate voltage; some newer spectrometers automatically work this way.

Next I measured the transmission of the depleted sample; then I measured the spectrum of the partially full sample at evenly spaced gate voltages, and then I repeated the measurement of the depleted spectrum. (I used a resolution of 0.5 cm^{-1}, which was much less than the width of the peak.) Finally, I repeated the capacitance measurement. The usual case was that the initial and final $C(V)$ scans (and the first and last depleted spectra) agreed quite well, meaning that the sample had not drifted or changed. Severe long-term drifts in the sample could usually be traced to problems with the electrical contacts to the sample.

Appendix E

Residual Absorption in the Reference

Several samples I investigated had a measurable intersubband absorption under large negative gate bias, when the capacitance indicated that the sample was depleted. This left a residual absorption in the reference spectrum, which appeared as a negative absorption (or emission) in the absorption spectra. In this Appendix I will explain how this issue was treated in the data analysis.

As the gate voltage of the sample was changed in small steps, the intersubband absorption was observed to shift in frequency and strength. However, beyond some negative gate voltage (where the capacitance dropped to zero), the small peak stopped changing. I believe that at this point, the chemical potential of the electrons in the well is comparable to the random fluctuations of the potential in the plane of the well. The sample is “quasi-depleted” in the sense that no more of the electrons can flow out of the contacts, because they have become localized in the two dimensional plane of the well. One source of potential fluctuations in the plane of the well is fluctuations in the well width in the plane.

In addition, once the electrons are isolated from the ohmic contacts, they no longer feel the increased voltage which may be applied to the gate. Thus, as the gate voltage is made more and more negative from zero volts, the static electric field across the well will keep changing until the electrons become isolated, at which point further reductions in gate voltage will have no effect on the static field.

Two steps were taken to analyze these data: first, the number of electrons contributing to the "dip" (negative absorption) was calculated and added to the charge density measured by capacitance. Second, the absorption due to the dip was fit and subtracted from the spectra.

E.1 Extra Charge Density of Dip

Capacitance ($C(V)$) measurements do not include the charge which never depletes, and therefore give a value for N_s which is always too low by an additive constant. This is because the capacitance measures the charge in the well by watching small amounts of charge flow out of the sample, and integrating them. However, the residual charge never leaves the sample and is never seen by the $C(V)$. Hence, it is necessary to add the residual charge density to the $C(V)$ charge density by hand.

The attenuation coefficient for electrons in the sample is given from Equation (C.10) by:

$$\alpha(\omega) = \frac{8\pi e^2 N_s f_{12} \hbar \Gamma}{cam\sqrt{\epsilon}} \left(\frac{(\hbar\omega)^2}{(\widetilde{E}_{12}^2 - (\hbar\omega)^2)^2 + 4(\hbar\omega)^2\Gamma^2} \right), \tag{E.1}$$

where the quantities are as specified in Appendix C. At resonance ($\hbar\omega = \widetilde{E}_{12} = \hbar\widetilde{\omega}_{12}$), this expression simplifies to:

$$\alpha(\widetilde{\omega}_{12}) = \frac{2\pi e^2 N_s f_{12} \hbar}{cam\sqrt{\epsilon}\Gamma}. \tag{E.2}$$

Solving for N_s, I find that:

$$N_s = \frac{\alpha(\widetilde{\omega}_{12})\Gamma cam\sqrt{\epsilon}}{2\pi e^2 f_{12}\hbar}. \tag{E.3}$$

Using this equation, I can convert the experimentally measured peak-value of the attenuation coefficient for the residual absorption (along with some other values from simulations and measurements) to estimate the charge density in the depleted sample. This is the residual charge density which should be added to the charge density measured by $C(V)$ to obtain the actual charge density in the sample at any gate voltage.

Note that this procedure will fail if the absorption peak does not shift in frequency with applied gate voltage enough to separate it from the absorption when the sample has more charge in it. Then the residual peak will be entangled with the absorption at different gate voltages, and it may not be possible to separate them. Fortunately this was not the case for the square well sample.

E.2 Correction of Spectra

The spectra need to be corrected to remove the dip, since it is just an artifact of the optical referencing process and has nothing to do with the absorption of the partially-full sample. The process outlined in this section will not only remove the dip from the partially-full spectra, but will also place an absorption *peak* in the quasi-depleted spectrum, showing that the sample was still absorbing radiation at that point.

The standard calculation of the attenuation coefficient for the electrons in a sample normalized to the transmission of the sample without electrons is:

$$\alpha(\omega) = \frac{1}{L}\ln\left(\frac{I_{\mathrm{fully-depleted}}(\omega)}{I_{\mathrm{partially-full}}(\omega)}\right) \tag{E.4}$$

where $I_{\mathrm{fully-depleted}}(\omega)$ is the transmission spectrum of the *fully*-depleted sample, and $I_{\mathrm{partially-full}}(\omega)$ is the transmission spectrum of the partially-full sample. However, this assumes that the fully-depleted spectrum has no absorptions due to electrons in the quantum well, because it is fully-depleted.

It turns out the simplest "fix" is exactly correct: you don't want the dip in the spectra due to the residual absorption in the reference, so you fit the dip to a Lorentzian, and then subtract this fit from all the spectra. The reason this works is as follows: you really want to calculate the attenuation coefficient using Equation (E.4), but instead you calculated using:

$$\alpha(\omega) = \frac{1}{L} \ln \left(\frac{I_{\mathrm{quasi-depleted}}(\omega)}{I_{\mathrm{partially-full}}(\omega)} \right) \tag{E.5}$$

where $I_{\mathrm{quasi-depleted}}(\omega)$ is the reference containing the residual absorption. The fit of the residual (negative) absorption is

$$\alpha_{\mathrm{Fit}}(\omega) = \frac{1}{L} \ln \left(\frac{I_{\mathrm{quasi-depleted}}(\omega)}{I_{\mathrm{fully-depleted}}(\omega)} \right). \tag{E.6}$$

(Note that the argument of the natural log is inverted from "normal" because $\alpha_{\mathrm{Fit}}(\omega) < 0$.) Since the peak being fit is the absorption of the well relative to no absorption from the well, it actually contains the desired quantity $I_{\mathrm{fully-depleted}}(\omega)$. By subtracting the fit of the residual (negative) absorption, you have:

$$\begin{aligned} \alpha(\omega) &= \frac{1}{L} \ln \left(\frac{I_{\mathrm{quasi-depleted}}(\omega)}{I_{\mathrm{partially-full}}(\omega)} \right) - \alpha_{\mathrm{Fit}}(\omega) \\ &= \frac{1}{L} \ln \left(\frac{I_{\mathrm{fully-depleted}}(\omega)}{I_{\mathrm{partially-full}}(\omega)} \right). \end{aligned} \tag{E.7}$$

Thus, this approach of subtracting "what you don't want" actually works correctly!

Practically, the method I used was to find a nice absorption spectrum where the negative residual absorption from the reference was far from any other absorption

features. Then I fit the negative residual absorption from that spectrum, and then subtracted the fit from *all* the absorption spectra. This yields a set of absorption spectra which show only the absorption of the well at that gate voltage, without artifacts of absorption from different gate voltages entering. Therefore, when it is working well, one should see a single peak in each of the spectra, and as the sample depletes more and more, the peak should get smaller and smaller, but the most depleted spectrum will still have a small absorption peak due to the residual electrons. See Figure E.1.

By making this subtraction to the spectra, *and* by adding the charge density related to the residual electrons into the measured charge density, the effects of the residual charge in the quasi-depleted sample are almost entirely removed.

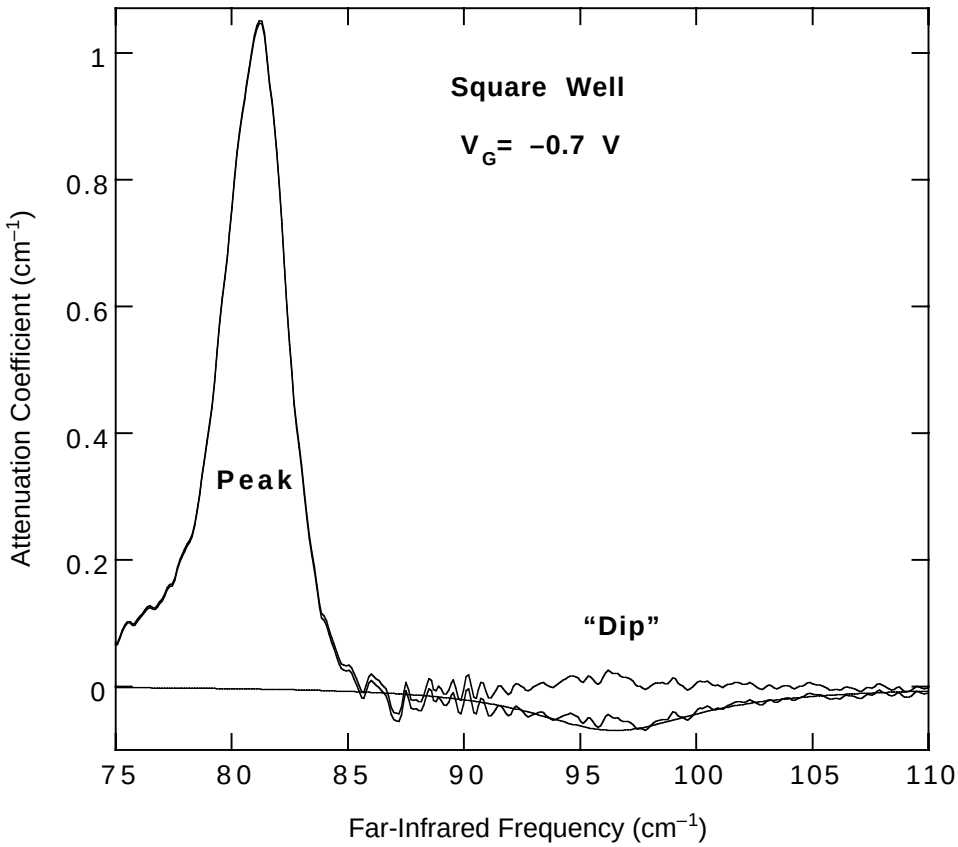

Figure E.1: Actual linear absorption data from the square well, showing the dip due to absorption in the reference, the fit used to remove the dip, and the experimental data after the dip has been removed. The peak (extending off the plot) is around 82 cm^{-1}, and the dip is around 96.5 cm^{-1}. The lower curves are the raw absorption data and the fit to the dip. The top curve is the absorption data with the dip-fit subtracted. Due to experimental noise in this spectra, the fit appears to remove slightly too much absorption.

Appendix F

Onset of Waveguide Effects

The waveguide mode theory created by M. Załużny in Reference [31] is somewhat cumbersome to use for fitting experimental results. The theory calculates the waveguide absorption from the absorption of the quantum well. However, the measurements are performed on the waveguide, not the quantum well. This leads to a lengthy process of numerically inverting the results of the theory to fit the observed data, as described in Chapter 4.

At sufficiently low charge densities (or for sufficiently wide absorption peaks) the waveguide theory will simply reproduce the results of linear perturbation theory, which are much easier to calculate. In this case, the waveguide theory need not be used. It is useful to have an idea of the range of parameters where the waveguide theory deviates strongly from linear perturbative theory, and where the results are similar and the simpler theory may be used with confidence. In this Appendix, I will discuss a criteria created by Załużny, calculate the charge density where the waveguide theory becomes important for the square well according to this criteria, and present some calculations and fits of the absorption lineshape as a function of

charge density.

F.1 Criteria of Załużny

In his analysis (Reference [31]) Załużny states that, when the parameter $\bar{\gamma} \gg 1$, and the value of the mode index n is not too big, then the difference between the results of the waveguide theory and the linear perturbative theory are dependent upon the value of $\Theta = \bar{\gamma}\bar{\Lambda}$. He states further that when $\Theta > 1$, the waveguide theory will diverge from linear perturbation theory, and when $\Theta < 1$ the two theories will agree (at least approximately). The parameters are defined as:

$$\bar{\gamma} = \bar{q}_{\rm tot} L \tag{F.1}$$

$$= \left(\frac{\widetilde{\omega}_{12}\sqrt{\epsilon}}{c} \right) L \tag{F.2}$$

where L is the width of the waveguide, ϵ is the dielectric constant of the semiconductor, $\widetilde{\omega}_{12}$ is the depolarization-shifted resonance frequency of the intersubband absorption, c is the speed of light, and $\bar{q}_{\rm tot} = (\widetilde{\omega}_{12}\sqrt{\epsilon}/c)$ is the total wavevector of the mode on resonance. The parameter $\bar{\Lambda}$ is defined as:

$$\bar{\Lambda} = \frac{4\pi}{c\sqrt{\epsilon}} \frac{N_s e^2 f_{12} \hbar}{2m\Gamma}. \tag{F.3}$$

($\bar{\Lambda}$ is a scaled measure of the peak conductivity.) N_s is the sheet density of the electrons in units of inverse area, e is the charge of the electron, $f_{12} = 2m\hbar^{-2} E_{12} |\langle z_{12} \rangle|^2$ is the oscillator strength, m is the effective mass of electrons in the conduction band, and Γ is the half-width at half-max energy-span of the absorption peak of the well.

For our square well sample, $\widetilde{\omega}_{12} \approx 87.1$ cm$^{-1} = 1.64 \times 10^{13}$ rad/s, $\epsilon = 13$, and $L = 0.0455$ cm. This gives a value of $\bar{\gamma} = 89.8$. Therefore the square well easily

satisfies the assumption $\bar{\gamma} \gg 1$. The assumption about the mode index not being too big is somewhat more uncertain; in this case it depends upon what the initial coefficients are which couple the external field into the waveguide.[1]

Assuming these assumptions are satisfied, the equations above may be solved for the charge density where $\Theta = 1$, which is the approximate charge density where the waveguide effects should become important:

$$N_s = \frac{2mc\Gamma\sqrt{\epsilon}}{4\pi e^2 f_{12}\hbar q_{\rm tot} L}. \tag{F.4}$$

For the square well, $\Gamma \approx 2.0$ cm^{-1} $= 3.97 \times 10^{-16}$ ergs, $f_{12} = 0.85$, and $m = m_e/16 = 5.6 \times 10^{-29}$ gm. Using these values, Equation (F.4) gives $N_s = 2.1 \times 10^{10}$ e/cm^2. This is approximately one-fifth of the maximum charge density; therefore the waveguide theory has effects even at relatively low charge densities. It is interesting to note that this is approximately the point where the square well integrated absorption becomes nonlinear.

F.2 Calculated Absorption Lineshapes

To investigate this further, I have calculated the absorption lineshape predicted by the Załużny waveguide theory at several charge densities. The external electric field in all calculations was taken to be constant. The results are shown in Figure F.1 for the square well sample, and in Figure F.2 for the double square well sample. Note

[1] Załużny only calculated the absorption of a single mode, in which case it is relatively simple to decide what the mode index n is. However, my calculations utilize a superposition of modes; I am assuming that Załużny's statement can be recast in this case to mean that the square modulus of the coefficients of the higher index modes is sufficiently small compared to the square modulus of the lower index modes.

that, in these Figures, the plotted attenuation coefficient is scaled by the value of Θ: if there were no waveguide effects, all the curves would coincide. Thus, although the larger values of Θ appear to be lower, this is a result of the scaling. The charge densities for Figure F.1 (F.2) are shown in Table F.1 (F.2). The other parameters used are shown on the plots. These results suggest that the criteria of Załużny is not sufficiently conservative when used with a superposition of modes; even when $\Theta < 1$, the absorption is significantly modified by waveguide effects. As Θ and the charge density increase, the absorption shifts to lower frequency and loses relative area noticeably from the linear perturbative result (which is $\Theta = 0.0001$).

Note that the absorption curves in Figures F.1 and F.2 look almost identical at the same values of Θ, as predicted by Załużny's criteria. This indicates that, even for this case of a superposition of modes, the parameter Θ is the important parameter for controlling the strength of the waveguide effects. Interestingly, due to the differences between the samples,[2] the charge density in the double square well is approximately 3.5 times greater than the square well at the same values of Θ.

The parameters of the calculated lineshapes for the square well were extracted by fitting the peaks to Lorentzian lineshapes. The center frequency of the peak, the integrated absorption, and the linewidth are plotted as a function of Θ in Figures F.3, F.4, and F.5. All the calculated lineshapes were well-fit by the Lorentzian fits.

Care should be taken in interpreting Figure F.4 (the integrated absorption data): the areas are directly from the fits from Figure F.1, so they have been scaled by

[2]The oscillator strength in the double square well $f_{12} = 0.34$, and the peak width $\Gamma = 3.4$ cm^{-1}, which both lessen the maximum conductivity, and therefore the waveguide effects. Therefore, a higher charge density is needed in the double square well to observe the same magnitude of waveguide effects as the square well.

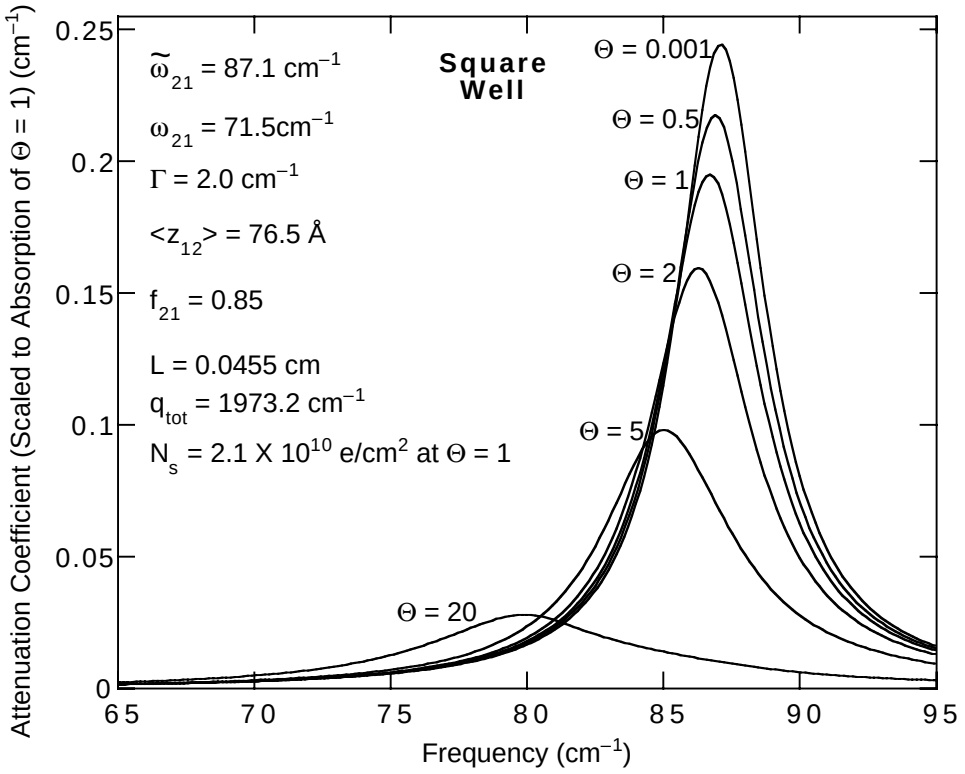

Figure F.1: Calculated absorption as a function of frequency for several values of the parameter Θ using the waveguide theory of Załużny for the square well sample. All the results are scaled to the absorption of $\Theta = 1$; all the peaks would coincide without nonlinear waveguide effects. As the value of Θ increases, the peak shifts to lower frequency and loses area in a relative sense. Note that at large values of Θ, the peak shifts significantly more than the linewidth. This calculation was done using a flat external field.

Θ	N_s
0.001	2.09×10^7 e/cm^2
0.5	1.04×10^{10} e/cm^2
1.0	2.09×10^{10} e/cm^2
2.0	4.18×10^{10} e/cm^2
5.0	1.04×10^{11} e/cm^2
20.0	4.18×10^{11} e/cm^2

Table F.1: The charge density values for the various values of Θ plotted in Figure F.1 for the square well.

Θ	N_s
0.001	7.3×10^7 e/cm^2
0.5	3.6×10^{10} e/cm^2
1.0	7.3×10^{10} e/cm^2
2.0	1.5×10^{11} e/cm^2
5.0	3.6×10^{11} e/cm^2
20.0	1.5×10^{12} e/cm^2
50.0	3.6×10^{12} e/cm^2

Table F.2: The charge density values for the various values of Θ plotted in Figure F.2 for the double square well. Note that the charge density values are approximately 3.5 times greater than those of the square well at the same values of Θ.

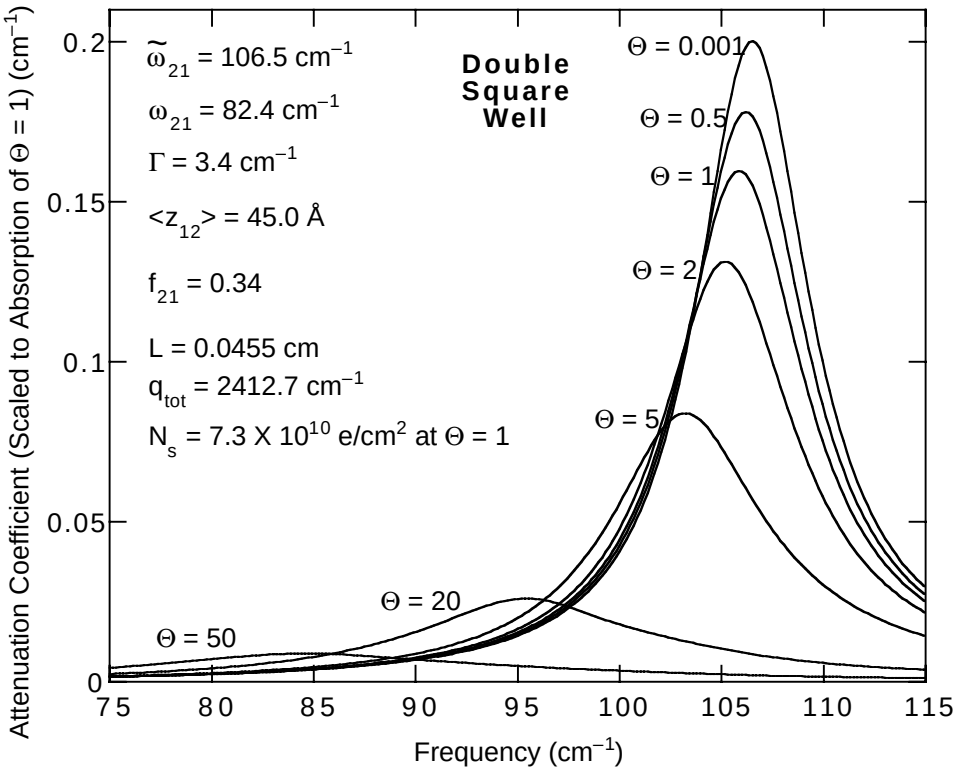

Figure F.2: Calculated absorption as a function of frequency for several values of the parameter Θ using the waveguide theory of Załużny for the double square well sample. All the results are scaled to the absorption of $\Theta = 1$; all the peaks would coincide without nonlinear waveguide effects. These results look qualitatively similar to the results for the square well, but the charge densities are approximately 3.5 times greater than for the square well.

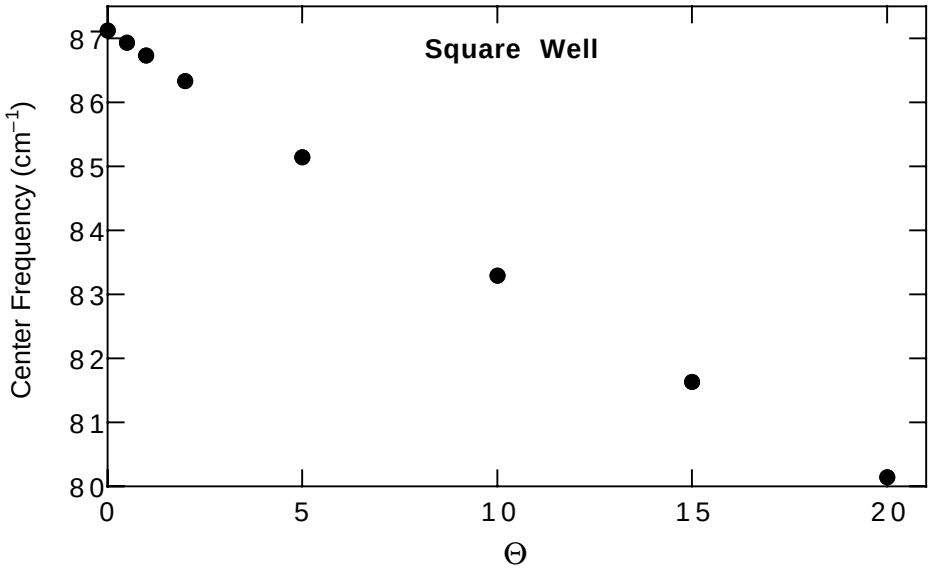

Figure F.3: The center frequency of the calculated lineshapes for the square well sample from Figure F.1, as a function of the parameter Θ.

Θ. Thus, any variation in Figure F.4 is due just to the waveguide effects; without waveguide effects it would just be a horizontal line. In addition, the numerical values of the area are normalized to the linear absorption value. Thus, the values on the vertical axis of Figure F.4 may be interpreted as the multiplicative reduction factor in the integrated absorption due to the waveguide effects. It is interesting to see that the absorption strength can be diminished by such large factors. At a value of $\Theta = 50$ (too large to be shown in the Figure), the reduction factor is 0.14. This means that 86% of the absorption strength has been screened away!

I remind the reader that the waveguide theory of Załużny did not significantly improve the fits to the experimental data from those of linear perturbation theory. (See Chapter 4.) This casts some doubt on the usefulness of this theory; I do not intend to diminish that doubt with this Appendix. On the other hand, it does seem

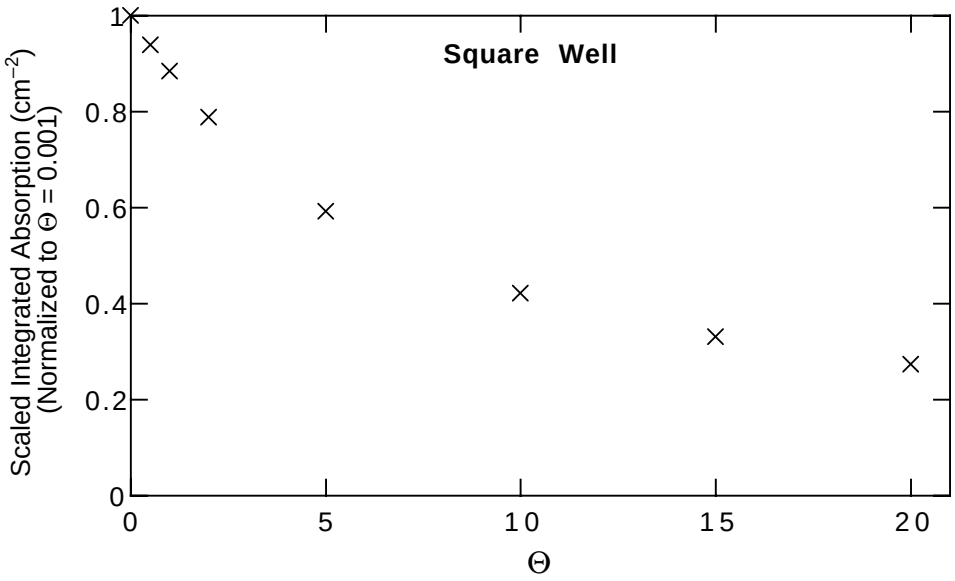

Figure F.4: The integrated absorption of the calculated lineshapes for the square well sample from Figure F.1, as a function of the parameter Θ. Note the vertical scale: the areas have been scaled by Θ, *and* the values has been normalized to the area of the $\Theta = 0.001$ curve. Without the waveguide effects, all the points would fall on a horizontal line with a relative area of one. The vertical axis may be interpreted as the multiplicative factor by which the waveguide effects have reduced the absorption strength. Thus, as the parameter Θ increases, the area under the peak does not increase linearly with Θ (and therefore N_s) as linear perturbation theory predicts.

"safer" to do measurements in a regime where the theory predicts linear perturbation theory will work than in a regime where linear perturbation theory is predicted to fail.

F.3 Avoiding Waveguide Effects

Assuming the criteria of the waveguide theory are correct, and one wishes to experimentally avoid the parameters where the waveguide theory becomes important, there are several options. The most obvious solution is to work at low charge densi-

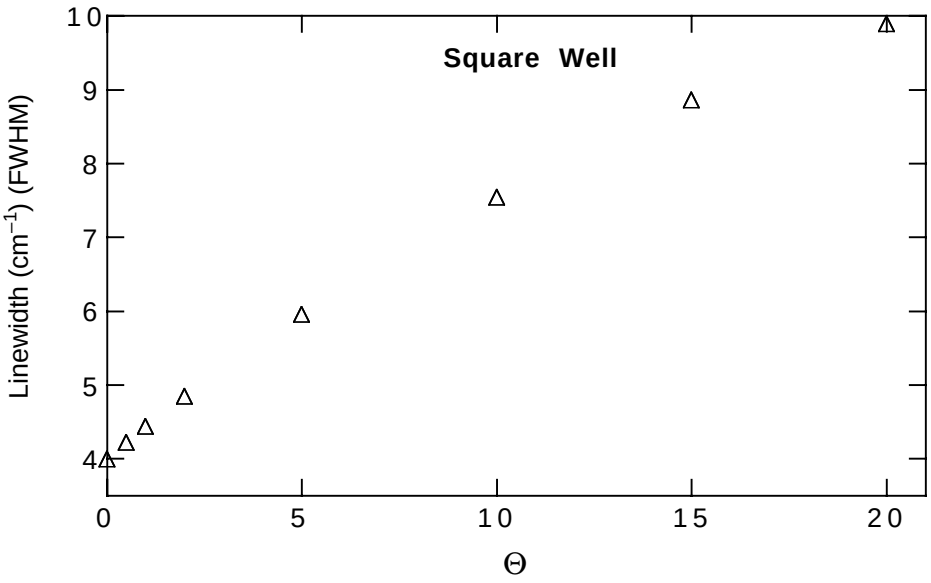

Figure F.5: The linewidths of the calculated lineshapes for the square well sample from Figure F.1, as a function of the parameter Θ. As Θ increases, the line becomes significantly wider.

ties, where the screening effects of the waveguide theory are unimportant. Depending upon the oscillator strength and linewidth of the sample, and the measurements of interest, this may or may not be a practical solution. Related solutions work by reducing the maximum intersubband conductivity, either by growing samples with large linewidths or small oscillator strengths.

Another possibility is to make the waveguide narrower; it is difficult for the electric field to significantly vary on length scales less than the wavelength. Thus, in a narrower waveguide, the mode fields have less freedom to distort in such a way as to significantly change the field at the well. Since $\Theta = \bar{\gamma}\bar{\Lambda}$, reducing the width of the waveguide L will reduce $\bar{\gamma}$ and therefore Θ by the same factor. This could be achieved by growing the samples on thinner wafers, or by physically thinning the

sample after growth. It could be difficult to couple radiation into a narrow waveguide, especially if the sample is dramatically thinned. In addition, by thinning the waveguide, the waveguide theory may become more complicated: the criteria of Załużny requires that $\bar{\gamma} \gg 1$, which will be less satisfied as the waveguide becomes narrower. I have not investigated what happens when this condition is no longer satisfied. Załużny states that, in this limit, the deviations of the waveguide theory from linear perturbation theory are no longer controlled by the product of $\bar{\gamma}\bar{\Lambda}$. This suggests that the deviations from linear perturbation theory are controlled by a more complicated function of these *two* parameters, instead of a single parameter.

Appendix G

Nonlinear Setup and Procedure

There were a large number of important details about the nonlinear absorption setup which allowed it function very efficiently. This Appendix describes those details.

First the glove bag is described, and then the various parts of the optical setup. Then the procedure used to measure the absorption is discussed, and finally the electronics and software are described.

G.1 Glove Bag

In order to greatly reduce water vapor absorption of the FIR, the entire setup was enclosed in a "glove bag." The glove bag consisted of clear plastic sheeting completely covering the table and optics and sealed to Port #1. The bag was sealed (with small holes to allow venting) and continuously purged with dry nitrogen gas. Sturdy rubber gloves taped[1] into the glove bag allowed the manipulation of optical components inside the bag such as attenuators during the measurements. The vast majority

[1]With the exception of the threaded holes in the optical table, all tape used on the glove bag was clear packing tape.

of the threaded holes in the surface of the optical table were covered with transparent office tape in order to reduce the time needed to purge the inside of the glove bag. (There is a great deal of volume and surface area below the holes which would be *very* slow to dry by purging due to the poor air flow below the table surface.) The optics were mounted to the table using magnetic bases or optical posts screwed into the table. When magnetic bases were used, the transparent office tape under the base was removed so that the base would grip the surface more strongly. The glove bag was penetrated in several places: at the port, at the flow cryostat's helium intake, and at the detector's window. At each of these points, a plate[2] was taped into the glove bag, and the hole in the plate was used as a template to cut the bag. Then a mating plate was affixed to the port, cryostat, or detector, and double-stick tape was used to seal the glove bag plate onto the mating plate, creating a quasi-airtight seal. Long optical posts were stategically placed under the glove bag to support it when it was not inflated.

During the initial inflation of the glove bag after it had been constructed, the glove bag would tend to overinflate and become extremely rigid. This was undesirable for two reasons: it indicated that the bag was not venting well (therefore water vapor may not have been able to easily escape), and the gloves on the rigid bag were difficult to use. (In a worst case scenario, the unrelieved pressure might have become great enough to rip the glove bag open.) Therefore, groups of small holes (approximately 1 mm in diameter) were poked in the bag in small groups, and visibly marked with a felt-tipped pen for later sealing if the bag developed additional leaks and failed to inflate. The dry nitrogen gas used was taken from boiling liquid nitrogen. Invari-

[2]The plates used for access points in the glove bag were plexiglass except for the cryostat, which used a set of metal rings with taped holes.

ably, each successive time the glove bag was used, a few of the holes would need to be covered with clear packing tape to inflate the bag as fully as it had inflated the previous time.

The theory of the glove bag is simple: dry air is mixed with wet air inside the bag, and the mixture is expelled through the vents (and leaks). The volume of the glove bag was estimated at several hundred liters, and the flow rate was 5–10 liters per minute. Thus, in some fraction of an hour (the "mixing time"), equal volumes of wet and dry air are mixed, and the amount of water vapor in the bag should be reduced by one-half. Each additional "mixing time" lowers the water vapor by another factor of two. As long as the flow rate of dry nitrogen into the bag and leak rate out of the bag are such that the bag visibly inflates (there is positive pressure inside the bag relative to the pressure outside), then the purging should be effective. The greater the flow rate into the bag (and the leak rate), the faster the purging will proceed.

G.2 Detailed Optical Setup

A four foot by four foot optical table was used as a foundation for the optical setup. The radiation from the FEL emerged from Port #1 of the optical transport system. (The optical transport system is the network of evacuated pipes and movable mirrors and lenses which direct the FIR from the FEL to the different setups.)

A diagram of the optical setup is shown in Figure 5.2. All the measurements reported here were done using this setup. The FIR emerges from the port with the electric field polarized in the plane of the optical table top. After emerging from the port, a small amount of the radiation (approximately 5%) was split off by a beam-splitter into a reference pyroelectric detector. This detector was used to attempt to

normalize pulse-to-pulse variations in beam energy; the nonlinearity of the absorption being measured ensured that this normalization process would not be entirely successful. All the focusing mirrors used in this setup were metal-surface with an off-axis parabolic shape. A two-dimensional[3] off-axis parabolic mirror was used to focus the FIR onto the detector element of the pyroelectric detector. The pyroelectric detector used was quite slow; hence it should have measured the total energy in the beam, as its time constant was much slower than the duration of the pulse. (The bolometer which measures the transmission through the sample was also quite slow. Thus, at least in the linear absorption regime of the sample, the reference detector and the bolometer should both have measured the total energy of the pulse, and therefore have been able to compensate for variations in the energy of the beam.)

Next the FIR passes through a set of mylar attenuators. The attenuators were cut from 10 mil thick mylar sheets into a circular shape with rectangular tabs for handling. The mylar used to make the attenuators is birefringent: it has a set of perpendicular axes, and radiation polarized along the axes travels at different speeds. (For an introductory discussion of birefringence, see Reference [9, pages 288–292].) Hence, to prevent the attenuators from changing the linearly polarized FEL beam into an elliptical polarization, great care was taken to cut the attenuators so that the rectangular handling tab was aligned to be parallel with one of the optical axes. Then, whenever they were used, the tabs were placed vertically so that the optical axes of the mylar were parallel and perpendicular to the electric field of the FEL. In order to make accurate absorption and intensity measurements, it was essential that the FIR passing through the sample be polarized in the growth direction, since the inter-

[3]By “two-dimensional” I mean that the mirror focuses in both the horizontal and vertical directions.

subband resonance only absorbs that polarization. Therefore, an FIR polarizer was placed directly in front of the window of the cryostat, so that the radiation entering the cryostat would be completely polarized in the growth direction. (Note that even if the mylar attenuators had not been used, the focusing mirrors can also change the polarization slightly.) The polarizer was two inches in diameter and made of the same material used in the linear absorption section. (See page 179 for details.)

At 80 cm^{-1} each 10 mil attenuator had a transmission of 0.293. The transmission of the mylar was modeled using the equation:

$$T = \exp\left(-\alpha_0 \left(\frac{\omega}{\omega_0}\right) d\right) \tag{G.1}$$

where T is the transmission of the mylar, d is the thickness of the mylar, ω is the frequency of radiation, ω_0 is the frequency at which the attenuation coefficient was measured, and α_0 is the attenuation coefficient which was measured at ω_0. The measured values are $\alpha_0 = 51.46$ cm^{-1}, at a frequency of $\omega_0 = 83.90$ cm^{-1}. The attenuation of 10 mils of mylar near 80 cm^{-1} was found to be a very practical amount, as a large dynamic range was needed, and a smaller factor would have lengthened the measurement duration painfully. (In addition, it was difficult to determine the optical axes in thinner pieces of mylar for orientation.)

The attenuators in front and in back of the sample were each held by a spring-loaded holder which pressed all of them together. It was suspected that the random spaces between the attenuators might create complicated standing-wave patterns which would make the attenuation for a given number of attenuators less repeatable. Occasionally, unexpected behavior in the amount of attenuation measured was observed, despite the precautions taken.

A wall of conducting black foam was used to block stray radiation which might

not have passed through the attenuators before the sample. The block extended from the table to a height well above the top of the beam. This block separated the table into a "bright" section on the port side, and an "attenuated" section on the other side, which the FIR could enter *only* through the first set of attenuators. This setup was found to be essential due to the large attenuations (sometimes more than a factor of 1,000,000) which were used: even 1% of the stray radiation from the "bright" section would have completely overwhelmed the signal in the "attenuated" section. This block greatly improved the reproducibility of the results, especially when using very large amounts of attenuation in front of the sample.

An aluminum cylindrical mirror was used to focus the FIR through the cryostat window onto the edge of the sample. The cylindrical mirror[4] focused only in the horizontal plane, and did not effect the radiation in the vertical plane. The unusual focal properties of the mirror were desirable because I wanted as much total energy to measure on the bolometer as possible, which meant having as large a focal spot on the sample as possible. (At a constant intensity and pulse duration, the energy of the pulse is determined by the size of the spot.) Since the width of the spot in the horizontal direction was limited to the thickness of the sample, the only way to enlarge the spot was to extend it in the vertical direction. (This way, a wider stripe of sample is being illuminated by the FIR.) In a more perfect world, no focusing would have been necessary; this would have created a spot which was more uniform in the horizontal direction to simplify the simulation of the waveguide effects. However, in order to achieve the high intensities necessary to observe interesting saturation effects, I had to concentrate the radiation. A mirror which focused in the horizontal

[4]The cylindrical mirrors used were custom machined in the UCSB Physics Machine Shop with a five inch focal length.

while retaining the vertical was chosen as the best tradeoff.[5]

The sample mounting scheme and beam blocks are shown schematically in Figure 5.1 in Chapter 5. The sample was mounted in such a way that the front edge was inserted through a slit in a piece of plexiglass. The plexiglass (which has aluminum foil glued to the backside) acted as a block to prevent FIR from leaking around the sample.[6] Any FIR which did not pass through the sample but arrived at the bolometer would cause errors in the measurements of the absorption of the sample. A small amount of light could leak *between* the plexiglass and the front of the sample. The back of the sample was glued to the copper sample mount, which effectively blocked all FIR which might leak around the back of the sample.

In addition to the material to prevent light from going around the sides of the sample, aluminum foil was placed on the *front* of the plexiglass above and below the sample. This foil had a rectangular hole cut in it which was roughly 40 mm wide (to prevent it from perturbing the horizontally polarized FIR electric fields), and exactly 4.0 mm high. This foil not only prevented light from leaking above or below the sample, but also defined the cross-sectional area of the sample exposed to the radiation. The fact that the vertical extent of the spot on the sample was 4.0 m high (and roughly constant intensity) was used extensively in the intensity calibration part of the experiment.

This beam-blocking technique was chosen in part to make it easy to understand the incoming FIR: the dielectric of the sample was mounted so that it was flush with

---

[5]See Appendix I for a discussion of an alternative method to increase the intensity at the sample using a telescope constructed of TPX lenses.

[6]Plexiglass is highly absorptive at FIR frequencies, and what little radiation was not absorbed by the plexiglass was reflected by the aluminum foil.

the dielectric of the plexiglass. This created a relatively flat wall of dielectric which reflected some of the beam and transmitted the rest. It is this setup which allowed us to assume that the FIR hitting the sample could be approximated by simple plane waves. If I had, instead, placed a metallic beam block around the sample, the radiation near the sample would have been a very complicated standing wave pattern due to the vastly different reflectances of the metal and the dielectric. In addition, this setup had the advantage that the total reflection from the plexiglass/aluminum foil plates was relatively low, which reduced the amount of stray radiation which was left to bounce around in unknown ways inside the cryostat and on the table.

The cryostat itself was a liquid helium continuous flow cryostat. The vacuum windows of the cryostat were made of black poly. Since the rest of the cryostat was made of vacuum-tight metal, this ensured that the cryostat was as opaque to visible light as the black poly. (Visible light ionizes DX centers in the sample and generally causes problems. Black poly is a black plastic material which transmits FIR but is quite opaque in the visible and infrared.) I observed no problems we could associate with visible light leakage onto the sample. The sample was in vacuum on the end of a copper cold-finger, which was cooled by a continuous flow of liquid helium from a pressurized storage dewar. The bottom of the cryostat was mounted onto the optical table by a plate which allowed the cryostat to slide left and right in order to align it laterally to the FIR beam. Threaded shafts could be used for fine lateral positioning of the cryostat. (Once properly positioned laterally, the cryostat was secured using heel clamps.) This plate was itself mounted on top of a lab jack, which allowed the entire cryostat to be translated up and down to adjust the height of the sample to the FIR beam.

Note that the samples used for these measurements had an anti-reflection coating

on the back edge (the edge opposite from the edge the FIR entered the sample). Due to the pulse-to-pulse variations in frequency of the FEL (estimated at roughly 0.1% of the frequency), the transmission of the sample without the coating was so noisy (and dependent on very small slow drifts in FEL frequency) that precise measurements were impossible. Although the coating (described on page 148) reduced the total transmission of the sample by a large factor, it reduced the fluctuations in transmission due to small changes in frequency by a much larger factor, and thus greatly improved the signal to noise ratio.

After passing through the sample, the transmitted beam was diverging. A cylindrical focusing mirror was placed after the cryostat to recollimate the FIR. The reason for recollimation was that an additional set of attenuators was used between the sample and the bolometer to control the intensity at the bolometer, and it was suspected that placing variable amounts of dielectric in a converging or diverging section of the beam would change the effective distance traveled by the beam, and therefore change the focal points or amount of divergence. (Placing attenuators in a collimated part of the beam does not have this problem; at most it shifts the entire beam sideways a small distance.)

After passing through the attenuators, the radiation was focused by a two-dimensional off-axis parabolic mirror into the bolometer. The bolometer used was a 4.2 K germanium bolometer with a Winston cone collector.

Note that another setup with lenses to form a telescope to contract the FIR beam was used to increase the intensity at the sample. This setup did not yield useful data for other reasons, but it is described in Appendix I.

G.3 Detailed Experimental Procedure

The night before the measurements, the glove bag was sealed with the cryostat in place, and several pieces of clear packing tape were used to seal the opening where the bolometer would go. Then the dry nitrogen gas was turned on to begin purging and removing the water which was in the air in the bag, as well as on the surfaces of the table and the optics.

The next morning, the bolometer was cooled and thick double-sided tape was applied to the front of the bolometer to stick it to the bolometer plate on the glove bag. Then, as quickly as possible, the clear tape sealing the bag at the bolometer hole was removed, and the bolometer was stuck onto the plate, sealing the bag. (The bolometer could not be stuck onto the bag the night before, because it needed to be filled with liquid nitrogen, allowed to precool, and then inverted and emptied before being filled with liquid helium. It was not possible to invert and empty it while it was attached to the bag, and the cryogens would not keep it cold long enough to allow it to be cooled the night before.) Fortunately, after purging all night, there was very little water left in the bag, and any moist air leaking in from outside would quickly be purged out. In addition, the bolometer could be stuck onto the bag while the bag still had enough positive pressure to keep a strong outward current of air moving out through the bolometer hole. This suggests that very little moist air was able to enter the bag.

After the bolometer was mounted, the electronics were set up; this usually took several hours, allowing the bag to have a final purge.

When everything was ready, the sample was cooled slowly; this usually took approximately one hour. The final temperature of the sample, as measured by a tem-

perature diode on the cold finger, was usually around 10 K. This temperature seemed to vary by a degree or two on different days, but it always corresponded to the lowest temperature which could be achieved using any flow rate of liquid helium. The heater in the cryostat was not used; it was found that the temperature was *extremely* stable without it.

Once the sample was cold, several capacitance scans were taken to measure the total charge in the well, the depletion voltage, and how stable the charge was. When the contacts were fine, the charge was stable, and the repeatability from day to day was quite good; the depletion voltage usually varied only by a few percent.

Next, the cryostat was optimized in position if it or the optics had been moved since the last optimization. The FEL was set to a frequency around 90 cm^{-1}, above the peak of the intersubband resonance. First, as many attenuators as possible were placed in front of the sample, and the attenuators behind the sample were removed. Next, the total transmission with the sample depleted[7] was maximized by moving the cryostat from side to side and observing the transmitted signal on the oscilloscope. After this had been done, the transmitted signal was maximized by moving the cryostat up and down. (The beam block inside the cryostat is designed in such a way that the beam can only pass through the gated region of the sample.)

Finally, the absorption measurements were started. (This would usually occur at around 1:00 P.M.) At the start of each new frequency, the attenuators would all be placed in front of the sample, and would be removed from behind the sample. In

[7]Due to the strong absorption of the sample, if the FEL frequency was near the resonance, the maximum transmission signal could be observed by moving the cryostat such that the leakage around the sample was maximized. Since this was highly undesirable, the sample was depleted to insure that it did not absorb.

addition, at the start of each frequency, the computer would deplete the sample, wait a bit, and then set the gate voltage to the (partially) full value, and wait for 150 s, for the amount of charge in the sample to stabilize; see Appendix J for the reason for this. Once the charge was stable, the transmission of approximately 50 FEL pulses was measured with the sample (partially) full, and then the computer would set the gate voltage to the depleted value. Then the computer would (after a 1 s delay) start recording the transmission of the empty sample for 50 pulses.

Next, the computer would set the gate voltage back to the (partially) full value, start timing the 150 s delay, and tell me to change the attenuators. While the 150 s was elapsing, I would add an attenuator behind the sample, and remove one from the front of the sample. (The attenuators were moved in that order so as to not blast the bolometer with an extra-bright pulse. For whatever reason, the bolometer took tens of seconds to recover from a bright pulse.) When I was done, I entered the total attenuation amount *behind* the sample into the computer, and then the computer waited the remaining part[8] of the 150 s, and started taking transmission data with the sample (partially) full. This would be repeated for each desired intensity; once all the desired intensities had been measured, I would tell the computer program that I was changing the frequency, and then go retune the FEL to a frequency approximately 1.0–1.5 cm^{-1} lower. Then I would start again measuring the absorption with all the attenuators in front of the sample, and proceed to measure the different intensities as above.

An important experimental technique used was to vary the attenuation at the sample without changing it much at the bolometer. This was accomplished by keeping

[8]The experiment would have taken approximately 50% longer had it not been possible to change the attenuation during the 150 s delay. It usually took at least 50 s or so to change the attenuators.

the *total* amount of attenuation in the beam constant, and simply *moving*[9] the attenuators from the front to the back of the sample. It was important to keep the intensity at the bolometer relatively constant because I wanted to vary the intensity at the sample by more than five orders of magnitude, and the bolometer would not respond linearly over that large a range in intensity. (The absolute linearity of the bolometer was crucial for the intensity calibration, discussed below.)

Note that a huge amount of time could have been saved if the transmission of the depleted sample was measured once at each frequency, rather than after each (partially) full measurement. This would have saved not only the time taken to measure the depleted transmission, but also 150 s of delay for each intensity! Theoretically this should have been possible, since the total amount of attenuation in the beam was constant for each frequency, and the transmission of the depleted sample should not be a function of intensity. However, when I tried doing this, it did not work. I observed slow drifts of the transmission (roughly one to five minutes timescale), which may have been caused by slow drifts in the FEL frequency (the attenuators were frequency-dependent), slow drifts in the FEL power, some type of slow drift in the detectors, slow changes in the humidity of glove bag, or mechanical/thermal shifts in the optics. This problem may have also been caused by the combination of many birefringent attenuators oriented approximately correctly with the polarizer in front of the sample; this could have made the total transmission of the setup sensitive to the precise orientation of the attenuators. Each time the attenuators were moved, they were oriented slightly differently, which could have created a different

[9]Experimentally there were actually two sets of attenuators: one attenuator would be removed in front of the sample, and a different attenuator (the same thickness) would be placed behind the sample. This was done because it was difficult to move the attenuators large distances inside the glove bag.

amount of elliptical polarization in the beam, which would have been removed by the polarizer. Thus, it seems reasonable that the depleted transmission would need to be measured each time the attenuators were moved. Either way, these slow drifts and lack of repeatability in attenuator placement made it essential to measure the depleted transmission as close in time as possible to the (partially) full transmission.

It took approximately one hour to measure the absorption at roughly ten to fifteen intensities at a single frequency; then it took from several minutes to many hours to retune the FEL to the next frequency. It usually took approximately fourteen hours to complete the measurements of the absorption.

The above procedure was repeated until a low enough frequency was reached that no significant absorption was measured at *any* intensity. Then the only remaining step was to calibrate the intensity at the sample for that day. For details, see Appendix H.

G.4 Electronics and Software

All the data for the experiment was recorded with a Macintosh computer running LabVIEW 3.1 by National Instruments, with a Lab-NB board and a GPIB board. (The Lab-NB was used for its digital-to-analog and analog-to-digital voltage converters.) In addition to recording the data, the computer also displayed the preliminary results in real-time and controlled parts of the experiment. A schematic diagram of the computer and other electronics and their interconnections is shown in Figure 5.3 in Chapter 5.

The voltage signals from the detectors were recorded on HP54505 digitizing oscilloscopes as functions of time. The scopes were triggered from the FEL electrical trigger pulse. Since the repetition rate of the FEL is approximately 1 Hz, each FEL

pulse could be digitized and the data downloaded to the computer before the next pulse. (The data were transferred to the computer using GPIB.) By downloading and recording data for each FEL pulse, statistical information about the noise in the data and drift over time was recorded for later analysis. (These data would have been lost if the data had been averaged on the oscilloscopes.) The computer recorded either the voltage amplitude of the pulse, or the integrated area under the entire pulse. Since both the reference pyroelectric detector and the bolometer are quite slow (with timescales of approximately 1 ms), the traces were easily digitized in detail. The signal from the bolometer was preamplified using a Stanford Linear Research preamp; it was set to filter out both high-frequency and low-frequency noise, in addition to amplifying the voltage. This preamp allowed a lower overall signal level at the bolometer, leading to the ability to measure the absorption at lower intensities.

The terminal voltage of the FEL was measured using a Fluke 8840A digital multimeter (DMM) to monitor the output of the generating voltmeter, which directly measures the terminal voltage. (The terminal voltage times the charge of the electron is the energy of the electrons in the resonator; it can be used to directly calculate the frequency of radiation emitted by the FEL.) The voltage (exactly a factor of one million smaller than the true terminal voltage) was transferred to the experimental setup using the built-in general-purpose cabling network in the lab. The multimeter was triggered by the FEL trigger pulse, and then the data were transferred to the computer using the GPIB.

The capacitance was measured with a standard setup, using a $C(V)$ box to add the DC voltage from the computer to the oscillating voltage from a function generator, doing some filtering, and converting the current response from the sample into a voltage for measurement on the lock-in amplifier. The analog output of the lock-in

amplifier was read by the Lab-NB board in the computer.

The custom LabVIEW programs used to control the experiment and acquire the data were quite complicated. They evolved and grew over several years. I will describe the final versions, which were used to take the data described here.

First, the program set the gate voltage to the depleted value and waited 5 s for the sample to stabilize. The program then asked the user to enter the amount of attenuation between the sample and the bolometer. Next, the program entered the main loop: the scopes were set to run, the gate voltage was set to the full value, the program waited 150 s for the sample to stabilize, and a set of (approximately 50) full transmission pulses was recorded (which includes measuring the frequency by sampling the GVM voltage). Next, the gate voltage was set to the empty value, and the sample was allowed to stabilize for 1 s. Next a set of empty data was taken, the results were stored, and the data were plotted on the screen. Finally, the gate voltage was set to the full value, and the computer started tracking the time while it asked the user to move the attenuators and enter the next attenuation value after the sample. During this time, the user would change the attenuators to keep the same total attenuation in the beam but change the intensity at the sample. Next, the user entered the thickness of the attenuators after the sample (or told the program to stop, in which case the gate voltage was set to the full value and the program stopped), the remaining amount (if any) of the 150 s wait was waited, and the loop repeated with the scopes being set and the full data being taken.

When everything went well,[10] this program was run once for each frequency, and

[10]When everything did not go well, the program was sometimes started and stopped several times to fix a problem. Since the option was always given not to log the data to the master file, it was relatively easy to keep data out of the master file when the setup was malfunctioning.

stored all the data for all the frequencies in a single data file. The data included all the raw results of the full transmission values and empty transmission values; these results were not averaged so that drifts and standard deviation could be analyzed later. Other programs could then read these data, calibrate the intensity, and do the final analysis.

While the program was running, it displayed the attenuation coefficient as a function of intensity (which was scaled by an unknown constant at that point, since the intensity was always calibrated at the end). In addition, the program displayed the range between the attenuation coefficient plus or minus a standard error, giving me a graphical real-time assessment of the amount of noise and drift.

The program also used a package of programs which smoothly "slide" the gate voltage in many small steps with a controllable delay between each step to change the gate voltage. Among other things, this limits the amount of current which flows through the ohmic contacts when the gate voltage is changed by a large amount suddenly.

The software worked very well. It made the experiment much easier in two main ways: it automatically measured and recorded almost all the data (except the attenuation after the sample, which it prompted for and compared with expected values), which greatly reduced the possibility of human error. (For example, by having the computer directly talk to the digital multimeter which measured the terminal voltage, I never had to enter the frequency of the radiation.) It also allowed me to *see* the data plotted as they were being taken, which allowed me to quickly notice and correct any problems that developed.

Appendix H

Intensity Calibration

A great deal of thought and effort were put into the measurement of the intensity inside the sample for the nonlinear absorption measurements. It was not possible to do a sensitive and direct measurement of the intensity inside the sample; thus an indirect measurement was done, by measuring the intensity outside the sample and estimating the fraction of radiation entering the sample. This Appendix describes the details of how the intensity was measured and calculated, and an estimate is made of the uncertainty in the intensity.

H.1 Calibration Algorithm

The intensity measurement was designed to be as automated as possible. The absolute signal from the bolometer for the depleted sample was stored at each frequency and intensity, along with the amount of attenuation between the sample and the bolometer at each frequency and intensity. After all the absorption measurements, the previous intensities were calibrated by recording the signal on the bolometer at that time, and then removing the cryostat and carefully measuring the intensity at the

point where the front edge of the sample had been.

The thinking behind this method was that the intensity of the FEL may fluctuate both in time at a single frequency, *and* it will fluctuate by a very large amount (as much or more than a factor of three) as it is tuned through different frequencies.[1] Therefore, any assumptions about the FEL intensity being constant during the entire measurement would be completely wrong.

The intensity at each frequency and intensity could be recorded by either the reference detector or the bolometer. However, as mentioned in Appendix G, the precise orientation of the attenuators combined with the polarizer had a significant effect on the intensity at the sample. Therefore, using the reference detector and assuming that the attenuators in front of the sample were oriented correctly would have been error-prone. (The beam does not pass through any attenuators before reaching the reference detector.) On the other hand, recording the bolometer signals and working backwards to the intensity at the sample should have been quite accurate: there were no polarizing elements between the sample and the bolometer, and the bolometer itself should have been polarization-independent. Thus, the orientation of the attenuators after the sample should have had very little (if any) effect on the amount of signal seen on the bolometer.

[1] For reasons no one understands, the FEL intensity seems to be a periodic function of frequency. If the angle of the end-mirrors of the lasing cavity are changed slightly, the position of the minima and maxima shift. The amplitude is so large that there are frequencies where the FEL will not lase at all.

H.2 Calibration Measurements

The duration of the FIR pulse, T_{FIR}, was set to 2.5 μs at each frequency. The high-speed reference pyroelectric detector in the FEL diagnostics box was used to measure the duration. The duration was adjusted by varying the length of the electron beam pulse. (The relationship between the electron and FIR pulse durations is very complicated due to the fact that the FEL starts lasing more quickly at some frequencies than other frequencies. For a constant FIR pulse duration, the electron beam pulse duration had to be varied by as much as 50%.)

Upon completion of the absorption part of the data taking, the total attenuation in the beam was recorded, and the sample was depleted. Then the signal on the bolometer was averaged and recorded. I then proceeded to measure the intensity at the sample position, which could be used to calculate the intensity inside the sample. Finally, the measured bolometer signal at the end was used together with the calculated intensity inside the sample to calibrate the previous data, when the bolometer signal had been measured, but the intensity had been unknown. (Figure H.1 schematically depicts the main parts of a section of the optical setup with the cryostat in place. The transmissions of the various elements and interfaces are listed in Table H.1.)

After the cryostat had been removed, the total energy in a defined fraction of the beam, E_{TK}, was measured using the Thomas-Keating acoustic cell and an iris, which were placed near the sample position.[2] Note that the polarizer was left in place. The

[2]Since the beam near the sample position was not diverging or converging much in the vertical direction, the iris and detector were wide enough to capture the entire beam. Since the Thomas-Keating can be damaged by large local intensities, the Thomas-Keating was placed behind the horizontal focus of the FIR.

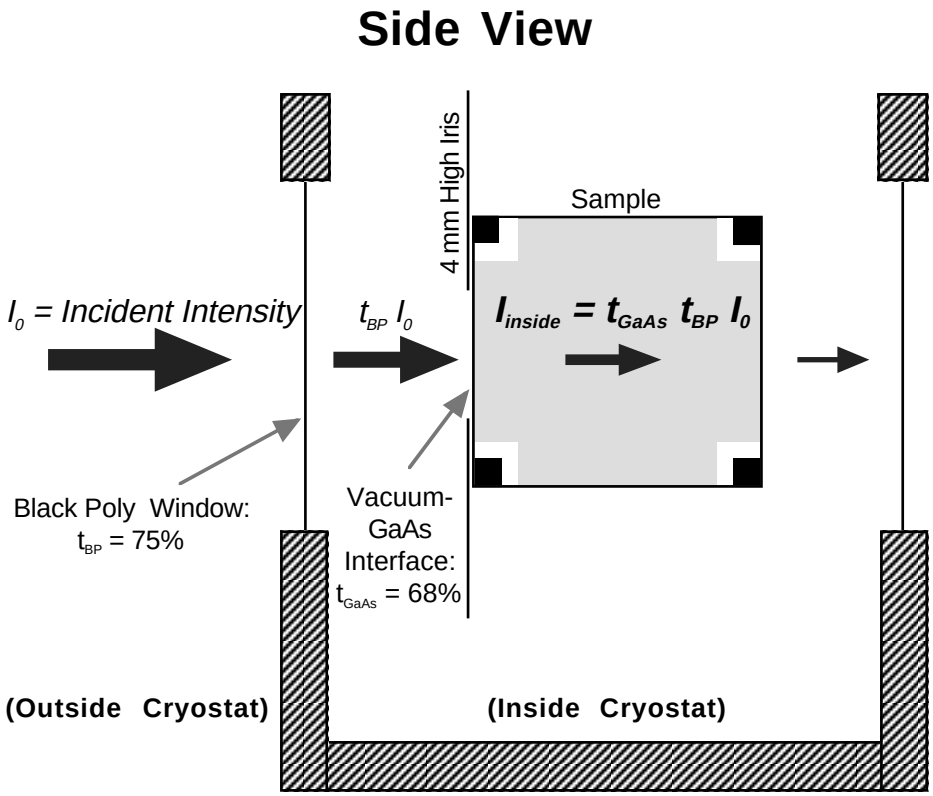

Figure H.1: A schematic of the optics for use in calculating the intensity calibration inside the sample. The radiation outside the cryostat which will impinge upon the sample has intensity I_0; after passing through the black poly window and into the sample, it has intensity $I_{\rm inside} = t_{\rm GaAs} t_{\rm BP} I_0$. Note that this is only considering the radiation which will strike the sample; other radiation will strike different parts of the cryostat and will have very different behavior.

Object	**Transmission**	**Value**
Black Poly Window	t_{BP}	75%
Horizontal Fraction of Beam on Sample	f	$\approx 40\%$
GaAs Front Interface	t_{GaAs}	68%
Attenuators after Sample	t_{after}	(Variable)

Table H.1: The transmission of various optical elements and interfaces, their notation, and actual values. The transmission of the black poly window was measured at 84.07 cm^{-1}.

iris for the Thomas-Keating was constructed of aluminum foil glued to cardboard (for strength), with the foil and cardboard cut out to form an iris with a rectangular aperture which was approximately 60 mm wide (much wider than the beam at the measurement point) and exactly 4.0 mm high. The height of the iris was chosen to be the same as the beam block at the front of the sample in the cryostat, which was also exactly 4.0 mm high. The iris was designed to be taped onto the front of the Thomas-Keating, with the lower edge resting on the base of the Thomas-Keating. When the Thomas-Keating was in position (on a specially machined aluminum block on top of the cryostat mount), the iris was positioned exactly at the same vertical height as the iris on the sample, thus insuring that the Thomas-Keating was measuring the energy in the same part of the beam.

The measurement of the radiation profile at the sample edge was an important part of the calibration. The automated spot-sizer was used: it consists of a pyroelectric detector with a narrow vertical slit iris attached to the front. The entire detector/iris bundle was mounted on a motorized x-y translation stage. Special software was used to step the detector, and measure the beam profile in either the x or y direction. The main challenge of the spatial profile measurement was placing the detector at the cor-

rect height and, more importantly, at the correct z point, by which I mean the point along the optical axis where the front edge of the sample had been.[3] This was essential since the radiation was coming to a focus in the horizontal plane, and a small error in the z position corresponded to a significant error in the size of the profile. Once the detector had been aligned, the important measurement was the horizontal scan of the beam at the sample edge. This profile was approximately gaussian. The profile was wide enough that the intensity varied by only a relatively small amount over the width of the sample. This indicates that using a constant external electric field as the input function for the waveguide simulation would be approximately correct. (Due to the small number of charge densities measured in the nonlinear setup, no waveguide simulations were done for this data. However, the use of a relatively constant external field in the measurement helps to insure that no unexpected waveguide effects were unwittingly observed.)

H.3 Calibration Analysis

From the measurement of the profile, and assuming that the horizontal alignment procedure of the cryostat placed the sample at the most intense part of the profile,[4]

[3]The correct z point along the optical axis was located by placing a precisely-sized aluminum block on the cryostat mount (against an edge of the cryostat mount) so that the back of the block was coplanar with the front edge of the sample. Then the spot-sizer detector was moved forward until it was touching the block, and the block could then be removed.

[4]Although this assumption sounds reasonable, the alignment of the cryostat often hinted at strange behavior: among other things, tilting the cryostat an almost immeasurable amount often changed the transmission by a significant amount. However, the alignment procedure can be assumed to have placed the sample quite close to the most intense part of the beam profile.

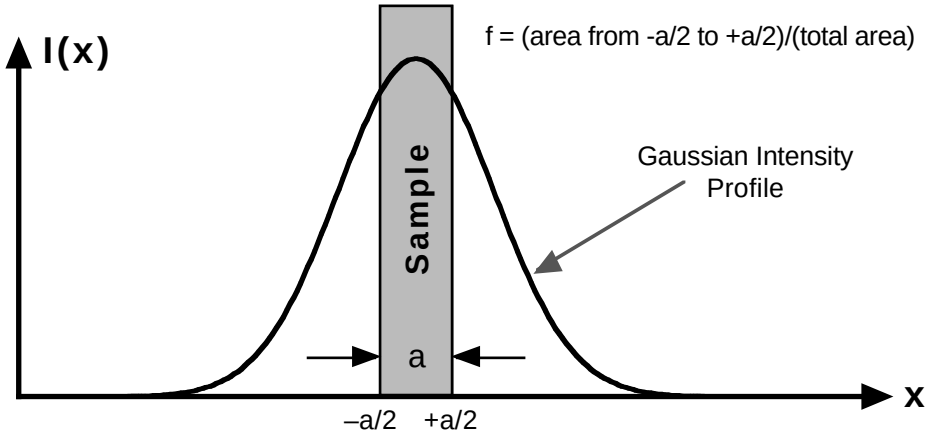

Figure H.2: A cartoon of the horizontal intensity profile near the sample. The shaded region represents the width and position of the sample edge in the incident profile. The gaussian fraction f is the fraction of the total energy which falls onto the sample. It is calculated by numerically integrating a gaussian distribution which is the same width as the experimentally measured profile: $f = (\int_{-a/2}^{+a/2} I(x)\,dx)/(\int_{-\infty}^{+\infty} I(x)\,dx)$.

the horizontal fraction f of radiation which fell onto the sample can be calculated. This fraction is depicted visually in Figure H.2.

Given the above measurements, the (assumed constant) intensity without the cryostat in place at the point where the front edge of the sample would have been can be written as:

$$I_{\rm outside} = \frac{E_{\rm TK} f}{T_{\rm FIR} A_{\rm sample}} \tag{H.1}$$

where $A_{\rm sample}$ is the exposed area of the front edge of the sample (4.0 mm $\times$ 0.455 mm). This expression is just the energy of the part of the beam hitting the sample ($E_{\rm TK} f$) divided by the duration of the beam ($T_{\rm FIR}$) and the area of the sample ($A_{\rm sample}$), according to the basic definition of the intensity.

In order to calculate the intensity *inside* the sample, the transmission through the various elements of the cryostat and sample must be considered. Since the region

of the beam now under consideration is very small (only the part which will strike the sample), and considering the size and shape of the windows and apertures in the cryostat, it is a good approximation that none of the beam being considered here collided with any metal components. In this case, the only elements between the sample and the outside are the black poly vacuum window, and the dielectric interface between the vacuum and the GaAs. Denoting the transmission of the black poly t_{BP}, and the transmission of the dielectric interface t_{GaAs}, the intensity inside the sample is:

$$I_{\mathrm{inside}} = t_{\mathrm{GaAs}} t_{\mathrm{BP}} (I_{\mathrm{outside}}) \tag{H.2}$$

$$= t_{\mathrm{GaAs}} t_{\mathrm{BP}} \left(\frac{E_{\mathrm{TK}} f}{T_{\mathrm{FIR}} A_{\mathrm{sample}}} \right). \tag{H.3}$$

The calibration relies on relating this intensity inside the sample to the bolometer signal. Since the transmission of the cryostat should consist entirely of radiation which passed through the sample, the bolometer signal will be linearly related to I_{inside}. There will be an inverse factor of the transmission of the attenuators after the sample, $1/t_{\mathrm{after}}$, since the lower the transmission of the attenuators between the sample and the bolometer, the higher the intensity was in the sample for a given bolometer signal. All the other factors (transmission of the anti-reflection coating, transmission out of the cryostat, efficiency of the mirrors, fraction of radiation reaching the bolometer, responsivity of the bolometer, etc...) are constant for that day, and can be combined into a single constant C. Thus, the signal measured on the bolometer can be related to the intensity inside the sample by:

$$I_{\mathrm{inside}} = C \frac{S_{\mathrm{bolo}}}{t_{\mathrm{after}}} \tag{H.4}$$

where S_{bolo} is the signal from the bolometer. Note that the bolometer signal can

be measured using any technique which is linearly related to the response of the bolometer (for example, the area under the peak on the oscilloscope *or* the amplitude of the peak) as long as this measurement technique is used consistently during the entire measurement. These factors are also part of the constant C.

The constant C is determined from the calibration data:

$$C = \frac{t^0_{\rm after} I^0_{\rm inside}}{S^0_{\rm bolo}} \tag{H.5}$$

where all the values with superscript "0" are determined from the calibration data at the end of the measurements.

Also note that, in Equation (H.5), the meaning of $t^0_{\rm after}$ is slightly different from the normal meaning. Any attenuators which were *in front* of the sample while measuring the calibration value of $S^0_{\rm bolo}$ but were removed when the energy was measured using the Thomas-Keating should be included (along with the attenuation behind the sample) in $t^0_{\rm after}$. (Attenuators were sometimes placed in front of the depleted sample during the calibration phase measurement of $S^0_{\rm bolo}$ to keep the intensity in the sample moderate during the calibration, in case prolonged high intensities might damage the sample.) This is because, in the calibration stage, the only important value is how much attenuation the bolometer "saw" which the Thomas-Keating did not "see." Once C for the entire data set has been determined from the calibration data, Equation (H.4) can be used to calculate the intensity inside the sample for all the absorption data.

Thus, summarizing and combining the parts, the complete result for determining the intensity inside the sample is:

$$I_{\rm inside} \;=\; C\left(\frac{S_{\rm bolo}}{t_{\rm after}}\right) \tag{H.6}$$

$$= \left\{ t_{\mathrm{GaAs}} t_{\mathrm{BP}} \left(\frac{E^0_{\mathrm{TK}} f^0}{T_{\mathrm{FIR}} A_{\mathrm{sample}}} \right) \left(\frac{t^0_{\mathrm{after}}}{S^0_{\mathrm{bolo}}} \right) \right\} \left(\frac{S_{\mathrm{bolo}}}{t_{\mathrm{after}}} \right) . \qquad \text{(H.7)}$$

H.4 Uncertainty in Intensity

The measurement of the intensity inside the sample was somewhat indirect. The intensity outside the sample was measured, the transmission through the cryostat and into the sample was estimated, and then this value was used to calibrate a single averaged measurement by the bolometer. Then this value was used to calibrate *all* the data sets.

One source of uncertainty in the measurement is the assumption that the sample is located exactly at the peak of the gaussian horizontal beam profile. As noted above, the transmission of the cryostat was very sensitively dependent upon not only the horizontal position of the cryostat, but also the *angle* of the cryostat. Therefore, it is likely that the entire sample was not exactly at the maximum of the horizontal intensity profile. However, it is also likely that the sample was close to the maximum of the intensity profile, since the cryostat was manipulated very carefully to maximize the transmission. It is difficult to calculate or estimate the uncertainty associated with this type of problem; the uncertainty was estimated by comparing the measured transmission of the cryostat with the predicted theoretical value. One of the trials was approximately 46% too low, indicating that the cryostat was not at the maximum of the intensity profile, and that the intensity may have been lower than calculated by that factor.

The estimated uncertainties for each of the quantities in the analysis of the intensity are shown in Table H.2. Where possible, the uncertainties were based upon variations of quantities which were measured several times; other quantities, such

Value	Uncertainty	Source
t_{BP}	$\pm 10\%$	Observed Variation; FT-IR Measurement
t_{GaAs}	(negligible)	Simple Theory
f^0	$+0\%; -50\%$	Variations in Cryostat Transmission
E^0_{TK}	$\pm 10\%$	Estimate
T_{FIR}	$\pm 15\%$	Observed Fluctuations
A_{sample}	(negligible)	Simple, Direct Measurement
$S_{\mathrm{bolo}}, S^0_{\mathrm{bolo}}$	(negligible)	Direct Measurement
$t_{\mathrm{after}}, t^0_{\mathrm{after}}$	$\pm 20\%$	Variation in Measured Values

Table H.2: The uncertainties of the relevant parameters for calculation of the intensity inside the sample. The source gives an indication of how the uncertainty was determined.

as the energy which the Thomas-Keating measures, were simply estimated based on experiences with the detector, and rough estimates of included uncertainties such as the placement of the iris used with the Thomas-Keating.

To estimate the uncertainty in the intensity inside the sample, consider Equation (H.7). By taking a differential, squaring both sides, ignoring uncorrelated cross-terms, and regrouping:

$$(\Delta I_{\mathrm{inside}})^2 = \left[\left(\frac{\Delta t_{\mathrm{BP}}}{t_{\mathrm{BP}}}\right)^2 + \left(\frac{\Delta f}{f}\right)^2 + \left(\frac{\Delta E_{\mathrm{TK}}}{E_{\mathrm{TK}}}\right)^2 + \left(\frac{\Delta T_{\mathrm{FIR}}}{T_{\mathrm{FIR}}}\right)^2 + \left(\frac{\Delta t^0_{\mathrm{after}}}{t^0_{\mathrm{after}}}\right)^2 + \left(\frac{\Delta t_{\mathrm{after}}}{t_{\mathrm{after}}}\right)^2\right] (I_{\mathrm{inside}})^2. \tag{H.8}$$

Using this Equation with the values from Table H.2, the complete uncertainty is calculated at $+35\%; -60\%$. This uncertainty is not symmetric because of the uncertainty in f^0; the gaussian fraction could be much smaller than calculated, but it could

not be any bigger, since the calculation assumes that the sample is positioned at the most intense part of the profile.

Appendix I

Increased Intensity with Lenses

In order to increase the total intensity at the sample, the setup described in Chapter 5 was modified by inserting an FIR telescope. However, due to inconsistencies in the data obtained with this setup, the results are not reported here. This Appendix describes the setup and what was learned during its creation.

A converging TPX[1] lens was placed at the port, with a theoretical focal length of 1.0 m. A diverging TPX lens with theoretical focal length -0.25 m was placed so that the focal points coincided; this created a beam condenser (or telescope) which shrank the FEL beam by a factor of four in diameter. (The theoretical focal lengths turned out to be significantly different from the measured focal lengths for both visible and FIR radiation.) This resulted in a geometrical increase of sixteen in intensity, which was somewhat offset by losses in the TPX, leading to an observed intensity enhancement of approximately thirteen.

[1]TPX is a plastic which has approximately the same index of refraction in the visible and the FIR. In addition, it is not overly lossy in the FIR. Therefore, it is an excellent material for FIR lenses, because a visible laser can be used to align the setup, and the FIR should behave in approximately the same manner as the visible light.

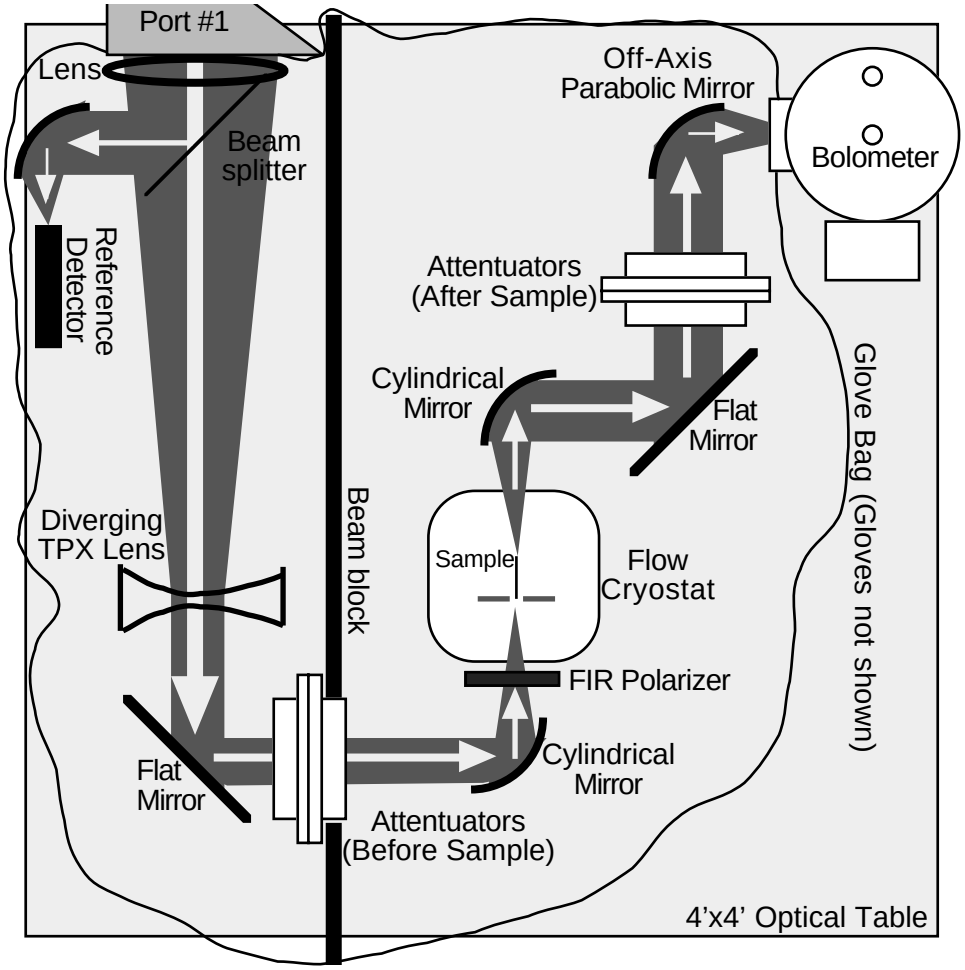

Figure I.1: The experimental setup used to measure nonlinear absorption with lenses. This setup differs from the setup without lenses (Figure 5.2) in that the FIR passes through a telescope consisting of two TPX lenses soon after it emerges from the port. This telescope shrinks the beam but leaves it collimated. In addition, the setup after the lenses is physically different due to spatial limitations of the table.

I first attempted to construct a telescope to condense the beam from off-axis parabolic mirrors;[2] however, it proved so difficult as to be completely impractical. The alignment was critical; any deviation from perfect alignment caused the beam to develop very strong asymmetries. In addition, after many hours of unsuccessful alignment it became clear that the quality of the mirrors was not good enough for this application. For normal, simple focusing applications (collimated beam going to a focus), the reflected light from the mirror needs to be in the correct place for distances on the order of the focal length. (For distances which are more than several focal lengths beyond the focus, the beam quality is rarely important.) However, I needed a short focal length mirror to make the telescope fit on the table and yet condense the FIR beam strongly, and the beam quality degraded seriously when examined more than several focal lengths away from the telescope. Specifically, the short focal length mirror had a focal length of 1 inch, and the beam was found to have very poor quality a distance of more than a few inches past the telescope. Since it was necessary for the beam quality to be quite high for at least several *feet*, these mirrors could not be used to make a functioning telescope. The quality of the lenses was much higher, and the alignment of the lenses was *much* easier and more forgiving. (I believe that the lenses themselves were not of higher quality than the mirrors used, but that the imperfections in the lenses had a smaller negative impact on the quality of the beam.)

After the telescope, the setup had to be changed around a bit due to limits on space on the table. A rigid wall of plexiglass and aluminum foil replaced the black conducting foam of the setup to separate the “bright” section of the table from the “attenuated” section. (The wall was a sandwich of plexiglass with aluminum foil

[2]The telescope constructed from mirrors would have been less lossy and the alignment would have been identical in the visible and the FIR.

filling: the plexiglass absorbs the FIR without reflecting a large amount, and then the aluminum foil assures that the transmission through the wall is exactly zero. The low reflectivity of the wall ensured that stray radiation would not keep bouncing around until it eventually found a way around the wall.)

Although the results from the telescope setup were not useful, the technique itself appears sound. It did increase the measured intensity at the sample focus by a factor of thirteen.

Appendix J

Long Time Scales for Sample to Stabilize

The charge density in the well was expected to stabilize in several RC time constants after the gate voltage was changed. The RC time of the sample was calculated from the resistance and capacitance of the sample to be on the order of several milliseconds. However, it was found from optical measurements (and confirmed with capacitance measurements) that after changing the gate voltage to partially full (after the sample had been depleted), the well would *overfill* and it would take *several minutes* (!!) for the sample to reach equilibrium. (See Figure J.1.)

This can be understood if the assumption is made that the DX centers are metastable but with a time constant on the order of minutes instead of weeks. If the DX centers could accept and release charge in deep traps slowly (on a timescale of minutes), then for the time the sample is depleted, the DX centers will be above the Fermi energy, and electrons will slowly be leaking off. When the gate voltage is suddenly (in a few seconds) changed so that the well is below the Fermi energy, the electrons rush back in and stabilize in a few RC time constants. However, since the electrons

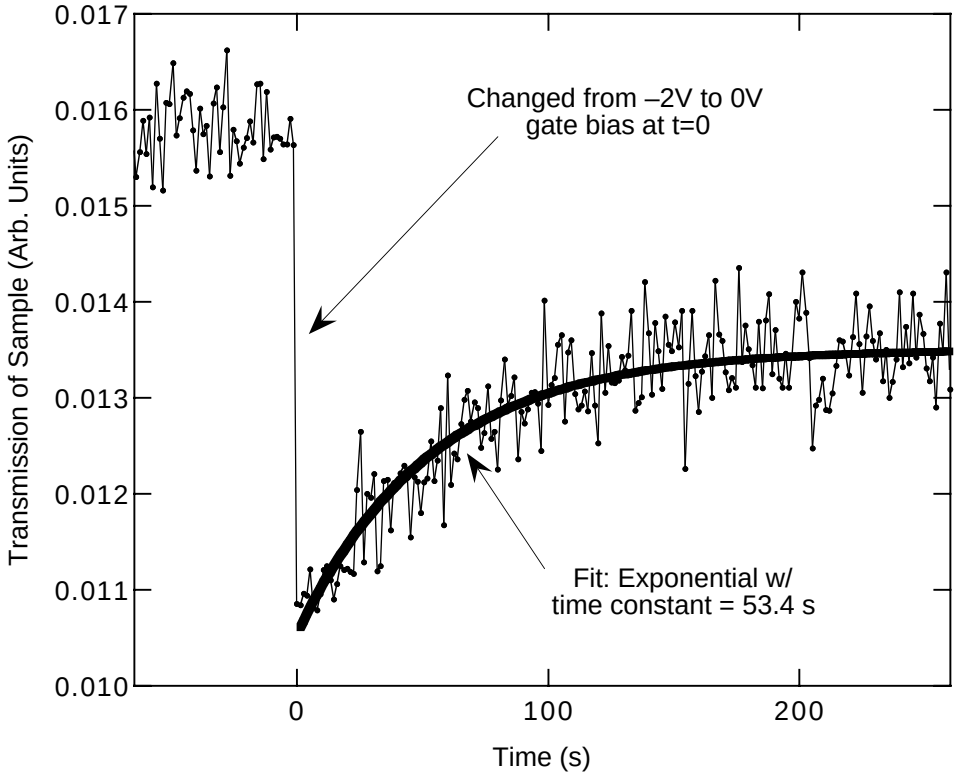

Figure J.1: The measured transmission of the sample as a function of time while the gate voltage is changed from -2.0 V to 0 V. When the gate voltage is changed, the transmission drops as expected, but the transmission then slowly increases as time passes until it reaches a steady state. This is attributed to changes in the occupation of the DX centers (see text). The heavy line is an exponential relaxation fit using the time constant shown.

which leaked off the DX centers while the sample was depleted cannot jump back onto the DX centers in a few milliseconds, the DX centers have a net positive charge that they did not have before (and will only have for a few minutes!). This extra positive charge attracts extra electrons into the well. However, this situation of the well being overly full only lasts for a few minutes, while the electrons slowly trickle back onto the DX centers. Once the DX centers are back to equilibrium, the well will be at equilibrium also.

The problem with this argument is that DX centers are expected to be metastable with a characteristic time of many days at low temperature. One possibility is that not all the silicon atoms in our samples formed DX centers. (This might be related to the details of the growth; I know very little about DX centers.) If this were true, then the very long time constants for the sample to stabilize could be attributed to the large spatial separation of the donors and the well (approximately 125 nm in the square well sample).

Due to these long time constants, the procedure was modified as described in Appendix G to allow the sample to come close to equilibrium before transmission measurements were performed. Due to the exponential nature of the approach to equilibrium, a tradeoff had to be made in terms of the time spent waiting for the sample to stabilize. A good balance was found by waiting approximately three time constants (150 s). This was expected get the sample charge density to within 5% of the final equilibrium value. Any additional waiting time would have made it impractical to complete all the measurements in a single cooldown.

Appendix K

Tabulated Nonlinear Data

This Appendix contains the nonlinear absorption results in tabulated form. "Frequency" means the FIR frequency at which the absorption was measured, "Intensity" is the intensity of FIR inside the sample at which the absorption was measured, "Atten. Coef." is the measured attenuation coefficient α, and "Stnd Error" is the standard error in the attenuation coefficient. The standard error is calculated statistically: it is simply the standard deviation of the multiple different calculated attenuation coefficient values divided by the square root of the number of measurements. For these data, the number of measurements was 50 partially full and 50 depleted.

The following are the tabulated nonlinear data for $V_G = 0$ volts, which are plotted in Figures 6.1, 6.2, and 6.3 of Chapter 6.

Frequency (cm^{-1})	**Intensity** (W/cm^2)	**Atten. Coef.** (cm^{-1})	**Stnd Error** (cm^{-1})
71.466	593.977E-6	-6.670E-3	5.265E-3
71.466	5.402E-3	21.651E-3	6.276E-3
71.466	58.089E-3	1.815E-3	6.357E-3
71.466	613.044E-3	6.597E-3	4.898E-3
71.466	4.166	3.806E-3	5.677E-3

71.466	37.381	10.019E-3	5.008E-3
71.466	128.789	42.910E-3	5.040E-3
71.466	426.351	84.755E-3	3.758E-3
72.237	369.206E-6	10.345E-3	7.995E-3
72.237	13.220E-3	-17.034E-3	8.136E-3
72.237	333.479E-3	-31.694E-3	8.423E-3
72.237	6.777	6.049E-3	8.917E-3
72.237	25.006	19.941E-3	7.120E-3
72.237	74.244	58.021E-3	7.516E-3
72.237	304.268	88.037E-3	6.707E-3
73.094	629.250E-6	23.341E-3	5.071E-3
73.094	1.777E-3	11.350E-3	5.949E-3
73.094	16.385E-3	12.119E-3	5.378E-3
73.094	111.332E-3	-605.231E-6	8.827E-3
73.094	1.275	-4.080E-3	6.885E-3
73.094	2.913	13.256E-3	7.398E-3
73.094	10.973	25.580E-3	7.820E-3
73.094	42.709	105.012E-3	7.410E-3
73.094	172.670	138.408E-3	7.121E-3
74.514	263.106E-6	-13.126E-3	10.508E-3
74.514	716.094E-6	633.331E-6	12.049E-3
74.514	2.729E-3	51.960E-3	10.300E-3
74.514	9.319E-3	23.664E-3	11.473E-3
74.514	23.779E-3	44.188E-3	10.001E-3
74.514	104.634E-3	421.366E-6	9.598E-3
74.514	346.374E-3	41.705E-3	9.859E-3
74.514	1.144	54.966E-3	8.523E-3
74.514	2.210	47.502E-3	10.981E-3
74.514	8.154	136.456E-3	8.827E-3
74.514	44.534	202.584E-3	6.933E-3
74.514	172.044	132.155E-3	6.225E-3
75.696	624.783E-6	24.292E-3	10.039E-3
75.696	2.575E-3	-3.818E-3	10.169E-3
75.696	6.305E-3	-8.462E-3	10.926E-3
75.696	19.643E-3	58.292E-3	7.688E-3

75.696	84.725E-3	83.328E-3	6.328E-3
75.696	300.465E-3	41.484E-3	8.247E-3
75.696	515.992E-3	63.613E-3	8.294E-3
75.696	2.893	135.994E-3	7.373E-3
75.696	5.306	290.715E-3	8.668E-3
75.696	21.341	292.559E-3	5.687E-3
75.696	83.294	217.754E-3	6.151E-3
76.944	537.336E-6	91.078E-3	4.242E-3
76.944	2.538E-3	109.007E-3	4.819E-3
76.944	7.259E-3	80.083E-3	4.646E-3
76.944	17.755E-3	125.550E-3	5.465E-3
76.944	79.001E-3	145.239E-3	4.811E-3
76.944	292.897E-3	166.337E-3	3.991E-3
76.944	891.275E-3	200.912E-3	3.857E-3
76.944	3.228	348.042E-3	4.523E-3
76.944	7.552	375.825E-3	5.129E-3
76.944	29.694	273.690E-3	4.871E-3
76.944	90.913	119.031E-3	4.706E-3
78.130	603.674E-6	160.786E-3	4.644E-3
78.130	1.970E-3	201.982E-3	5.147E-3
78.130	6.306E-3	217.452E-3	4.525E-3
78.130	26.889E-3	233.403E-3	4.415E-3
78.130	72.112E-3	217.840E-3	4.736E-3
78.130	198.807E-3	306.836E-3	5.864E-3
78.130	795.691E-3	390.012E-3	5.033E-3
78.130	3.808	436.650E-3	4.321E-3
78.130	10.778	404.587E-3	4.553E-3
78.130	40.532	275.187E-3	3.628E-3
78.130	144.169	136.664E-3	3.406E-3
79.497	686.159E-6	409.553E-3	6.464E-3
79.497	1.872E-3	488.435E-3	8.179E-3
79.497	7.811E-3	439.879E-3	7.251E-3
79.497	26.287E-3	455.974E-3	6.563E-3
79.497	98.739E-3	611.893E-3	6.433E-3
79.497	262.013E-3	617.667E-3	7.369E-3
79.497	1.094	585.318E-3	7.835E-3

79.497	4.419	430.103E-3	6.379E-3
79.497	12.947	303.092E-3	7.258E-3
79.497	48.007	189.200E-3	7.598E-3
79.497	137.165	87.614E-3	6.805E-3
80.861	555.120E-6	744.035E-3	6.712E-3
80.861	1.693E-3	884.452E-3	8.392E-3
80.861	5.795E-3	916.038E-3	8.875E-3
80.861	21.337E-3	854.026E-3	8.731E-3
80.861	73.739E-3	802.295E-3	8.252E-3
80.861	264.573E-3	748.272E-3	11.040E-3
80.861	936.540E-3	578.474E-3	6.815E-3
80.861	3.292	375.751E-3	6.188E-3
80.861	11.006	237.741E-3	5.102E-3
80.861	42.264	152.275E-3	4.539E-3
80.861	154.882	79.485E-3	4.676E-3
82.281	369.842E-6	633.970E-3	9.652E-3
82.281	1.394E-3	654.324E-3	7.437E-3
82.281	4.245E-3	764.434E-3	10.916E-3
82.281	15.340E-3	750.182E-3	9.723E-3
82.281	56.649E-3	724.092E-3	9.623E-3
82.281	205.148E-3	659.738E-3	8.165E-3
82.281	680.515E-3	491.388E-3	7.617E-3
82.281	2.652	302.695E-3	6.680E-3
82.281	8.637	178.280E-3	7.629E-3
82.281	33.976	126.538E-3	6.191E-3
82.281	139.424	75.006E-3	5.226E-3
83.774	585.387E-6	281.054E-3	5.174E-3
83.774	1.063E-3	306.559E-3	3.484E-3
83.774	2.439E-3	366.316E-3	3.407E-3
83.774	10.227E-3	346.757E-3	5.347E-3
83.774	39.005E-3	307.909E-3	4.538E-3
83.774	133.022E-3	299.738E-3	4.140E-3
83.774	445.636E-3	249.308E-3	5.591E-3
83.774	1.846	147.247E-3	4.496E-3
83.774	5.474	104.798E-3	4.153E-3
83.774	17.896	65.434E-3	4.502E-3

83.774	70.615	31.900E-3	4.165E-3
85.183	768.954E-6	153.504E-3	6.673E-3
85.183	1.827E-3	178.006E-3	6.321E-3
85.183	1.914E-3	207.489E-3	9.168E-3
85.183	7.932E-3	140.849E-3	6.693E-3
85.183	24.776E-3	130.913E-3	6.495E-3
85.183	89.398E-3	147.470E-3	6.722E-3
85.183	379.279E-3	144.341E-3	6.235E-3
85.183	1.247	104.552E-3	5.541E-3
85.183	4.872	81.377E-3	5.668E-3
85.183	15.638	65.865E-3	6.162E-3
85.183	58.706	21.593E-3	5.640E-3
86.693	563.753E-6	77.956E-3	5.430E-3
86.693	1.380E-3	60.497E-3	6.953E-3
86.693	6.886E-3	87.117E-3	5.673E-3
86.693	21.895E-3	57.123E-3	7.815E-3
86.693	75.057E-3	54.790E-3	7.290E-3
86.693	296.553E-3	55.866E-3	6.659E-3
86.693	1.197	79.668E-3	6.238E-3
86.693	4.304	41.687E-3	8.179E-3
86.693	21.232	20.577E-3	5.751E-3
86.693	83.261	14.659E-3	4.359E-3

The following are the tabulated nonlinear data for $V_G = -0.5$ volts, which are plotted in Figures 6.4, 6.5, and 6.6.

Frequency (cm^{-1})	**Intensity** (W/cm^2)	**Atten. Coef.** (cm^{-1})	**Stnd Error** (cm^{-1})
72.029	830.102E-6	29.664E-3	7.192E-3
72.029	2.410E-3	34.511E-3	8.225E-3
72.029	7.491E-3	25.782E-3	8.930E-3
72.029	76.635E-3	56.733E-3	8.343E-3
72.029	268.024E-3	-35.455E-3	11.102E-3
72.029	662.635E-3	27.028E-3	7.188E-3
72.029	2.137	27.536E-3	8.713E-3
72.029	5.394	19.998E-3	7.739E-3
72.029	18.968	28.283E-3	7.915E-3
72.029	71.112	48.332E-3	5.863E-3
72.029	248.097	91.383E-3	5.254E-3
73.373	1.766E-3	2.336E-3	3.174E-3
73.373	1.859E-3	-8.333E-3	2.849E-3
73.373	8.237E-3	8.483E-3	1.855E-3
73.373	20.700E-3	21.996E-3	2.373E-3
73.373	47.117E-3	4.904E-3	3.838E-3
73.373	246.837E-3	23.083E-3	1.965E-3
73.373	743.635E-3	6.155E-3	1.505E-3
73.373	2.299	18.699E-3	1.574E-3
73.373	8.769	25.671E-3	1.624E-3
73.373	17.792	77.419E-3	2.337E-3
73.373	63.174	106.716E-3	2.201E-3
73.373	227.110	162.048E-3	1.734E-3
74.722	2.185E-3	24.692E-3	10.390E-3
74.722	6.974E-3	14.459E-3	9.737E-3
74.722	7.461E-3	77.301E-3	9.140E-3
74.722	7.775E-3	85.421E-3	9.591E-3
74.722	33.579E-3	61.585E-3	11.277E-3
74.722	159.595E-3	135.846E-3	6.718E-3
74.722	306.425E-3	125.290E-3	5.470E-3
74.722	1.703	79.135E-3	5.766E-3

74.722	4.062	110.446E-3	6.603E-3
74.722	8.374	168.522E-3	5.316E-3
74.722	21.623	222.865E-3	6.331E-3
74.722	110.228	170.345E-3	4.378E-3
76.083	786.009E-6	77.144E-3	5.205E-3
76.083	4.978E-3	87.008E-3	7.016E-3
76.083	9.731E-3	124.783E-3	6.964E-3
76.083	32.827E-3	189.884E-3	6.172E-3
76.083	117.691E-3	76.420E-3	5.794E-3
76.083	593.728E-3	139.176E-3	3.339E-3
76.083	1.055	156.645E-3	5.303E-3
76.083	3.399	256.282E-3	5.849E-3
76.083	8.227	363.561E-3	5.574E-3
76.083	21.341	312.762E-3	6.777E-3
76.083	132.430	208.395E-3	4.311E-3
77.735	2.345E-3	287.583E-3	5.581E-3
77.735	5.319E-3	310.445E-3	4.509E-3
77.735	21.783E-3	236.414E-3	5.498E-3
77.735	42.803E-3	168.768E-3	5.539E-3
77.735	82.129E-3	353.918E-3	5.279E-3
77.735	95.558E-3	187.548E-3	7.618E-3
77.735	311.640E-3	301.597E-3	5.633E-3
77.735	1.023	337.711E-3	4.000E-3
77.735	3.911	443.882E-3	4.373E-3
77.735	7.961	445.429E-3	6.105E-3
77.735	36.133	284.994E-3	4.146E-3
77.735	135.129	168.131E-3	4.220E-3
79.200	1.386E-3	580.634E-3	7.174E-3
79.200	4.412E-3	659.486E-3	6.530E-3
79.200	13.212E-3	640.839E-3	9.124E-3
79.200	30.625E-3	762.016E-3	18.328E-3
79.200	123.370E-3	661.550E-3	9.290E-3
79.200	498.677E-3	608.873E-3	7.349E-3
79.200	1.393	544.984E-3	9.030E-3
79.200	5.392	435.180E-3	9.035E-3
79.200	15.583	304.837E-3	8.565E-3

79.200	51.544	189.173E-3	10.085E-3
79.200	199.994	100.386E-3	6.013E-3
80.609	2.186E-3	622.702E-3	4.421E-3
80.609	7.310E-3	632.571E-3	3.936E-3
80.609	25.767E-3	622.406E-3	3.913E-3
80.609	82.092E-3	757.285E-3	5.993E-3
80.609	96.171E-3	576.040E-3	3.057E-3
80.609	299.447E-3	550.735E-3	4.200E-3
80.609	945.639E-3	442.022E-3	3.644E-3
80.609	3.349	319.939E-3	4.582E-3
80.609	10.365	211.353E-3	4.139E-3
80.609	34.829	142.519E-3	3.720E-3
80.609	128.744	83.513E-3	3.294E-3
82.126	1.361E-3	351.065E-3	3.876E-3
82.126	4.600E-3	374.468E-3	4.428E-3
82.126	14.540E-3	370.645E-3	4.399E-3
82.126	54.586E-3	351.210E-3	4.293E-3
82.126	205.062E-3	321.746E-3	4.839E-3
82.126	890.164E-3	247.703E-3	3.851E-3
82.126	2.761	202.374E-3	3.543E-3
82.126	9.629	146.505E-3	3.637E-3
82.126	32.393	89.849E-3	3.886E-3
82.126	123.433	56.579E-3	3.431E-3
83.645	802.823E-6	65.066E-3	6.075E-3
83.645	10.482E-3	95.141E-3	7.718E-3
83.645	144.442E-3	64.288E-3	6.645E-3
83.645	1.948	58.146E-3	7.717E-3
83.645	6.597	60.931E-3	6.362E-3
83.645	20.095	31.554E-3	6.327E-3
83.645	75.777	5.802E-3	7.305E-3
85.306	529.045E-6	50.217E-3	8.350E-3
85.306	6.774E-3	10.637E-3	10.445E-3
85.306	128.359E-3	51.494E-3	7.093E-3
85.306	1.545	68.979E-3	6.874E-3
85.306	6.070	33.140E-3	5.729E-3

85.306	19.439	29.459E-3	8.540E-3
85.306	81.933	35.776E-3	8.874E-3

Bibliography

[1] S.J. Allen, Jr., D.C. Tsui, and B. Vinter. On the absorption of infrared radiation by electrons in semiconductor inversion layers. *Solid State Communications*, 20:425, 1976.

[2] Tsuneya Ando, Alan B. Fowler, and Frank Stern. Electronic properties of two-dimensional systems. *Reviews of Modern Physics*, 54:437, 1982.

[3] John Chamberlain. *The Principals of Interferometric Spectroscopy*. John Wiley & Sons, 1979.

[4] R.M. Dreizler and E.K.U. Gross. *Density Functional Theory*. Springer-Verlag, Berlin, Heidelberg, 1990.

[5] Christopher Lee Felix. *Linear and Nonlinear Far-infrared Studies of Aluminum Gallium Arsenide Quantum Wells*. PhD thesis, University of California, Santa Barbara, 1993.

[6] R. Ferriera and G. Bastard. Evaluation of some scattering times for electrons in unbiased and biased single-quantum-well and multiple-quantum-well structures. *Physical Review B*, 40:1074, 1989.

[7] Bryan Galdrikian. *Nonlinear and Nonperturbative Dynamics in Quantum Wells*. PhD thesis, University of California, Santa Barbara, 1994.

[8] S. M. Goodnick and P. Lugli. Monte Carlo simulation of intersubband relaxation in semiconductor quantum wells. *Superlattices and Microstructures*, 5:561, 1989.

[9] Eugene Hecht. *Optics*. Addison-Wesley, 1990.

[10] J.N. Heyman, K. Unterrainer, K. Craig, B. Galdrikian, M.S. Sherwin, K. Campman, P.F. Hopkins, and A.C. Gossard. Temperature and intensity dependence of intersubband relaxation rates from photovoltage and absorption. *Physical Review Letters*, 74:2682, 1995.

[11] John David Jackson. *Classical Electrodynamics*. John Wiley & Sons, second edition, 1975.

[12] F. H. Julien, J. M. Lourtioz, N. Herschkorn, D. Delacourt, J. P. Pocholle, M. Papuchon, R. Planel, and G. Le Roux. Optical saturation of intersubband absorption in GaAs-$Al_xGa_{1-x}As$ quantum wells. *Applied Physics Letters*, 53:116, 1988.

[13] W. H. Knox, D. S. Chemla, G. Livescu, J. Cunningham, and J. E. Henry. Femtosecond carrier thermalization in dense fermi seas. *Physical Review Letters*, 61:1290, 1988.

[14] J. A. Levenson, G. Dolique, J. L. Oudar, and I. A. Abram. Time resolved intersubband relaxation in GaAs/$Ga_{1-x}Al_xAs$ multiple quantum wells. *Solid State Electronics*, 32:1869, 1989.

[15] Jon Mathews and R.L. Walker. *Mathematical Methods of Physics*. Addison-Wesley, 1970.

[16] S.W. McKnight, K.P. Stewart, H.D. Drew, and K. Moorjani. Wavelength-independent anti-interference coating for the far-infrared. *Infrared Physics*, 27:327, 1987.

[17] P. M. Mooney and T. N. Theis. The DX center: A new picture of substitutional donors in compound semiconductors. *Comments on Condensed Matter Physics*, 16:167, 1992.

[18] B. N. Murdin, W. Heiss, C. J. G. M. Langerak, S.-C. Lee, I. Galbraith, G. Strasser, E. Gornik, M. Helm, and C. R. Pidgeon. Direct observation of the LO phonon bottleneck in wide GaAs/$Al_xGa_{1-x}As$ quantum wells. *Physical Review B*, 55:5171, 1997.

[19] D. Y. Oberli, D. R. Wake, M. V. Klein, R. Henderson, and H. Morkoç. Intersubband relaxation of photoexcited hot carriers in quantum wells. *Solid-State Electronics*, 31:413, 1988.

[20] William H. Press, Saul A. Teukolsky, William T. Vetterling, and Brian P. Flannery. *Numerical Recipes in C*. Cambridge University Press, 1995.

[21] A. Seilmeier, H. J. Hübner, G. Abstreiter, G. Weimann, and W. Schlapp. Intersubband relaxation in GaAs-$Al_xGa_{1-x}As$ quantum well structures observed directly by an infrared bleaching technique. *Physical Review Letters*, 59:1345, 1987.

[22] Jagdeep Shah. *Hot Carriers in Semiconductor Nanostructures*. Academic Press, 1992.

[23] Greg Snider. Department of Electrical Engineering, University of Notre Dame, Notre Dame, Indiana 46556. e-mail: snider.7@nd.edu.

[24] M. C. Tatham, J. F. Ryan, and C. T. Foxon. Time-resolved raman measurements of intersubband relaxation in GaAs quantum wells. *Physical Review Letters*, 63:1637, 1989.

[25] Claude Weisbuch and Borge Vinter. *Quantum Semiconductor Structures*. Academic Press, 1991.

[26] Ralph E. Williams. *Modern GaAs Processing Methods*. Artech House, 1990.

[27] Roland Winston. Nonimaging optics. *Scientific American*, page 76, March 1991.

[28] Roland Winston and W. T. Welford. *High Collection Nonimaging Optics*. Academic Press, 1989.

[29] M. Załużny. Influence of the depolarization effect on the nonlinear intersubband absorption spectra of quantum wells. *Physical Review B*, 47:3995, 1993.

[30] M. Załużny. Saturation of intersubband absorption and optical rectification in asymmetric quantum wells. *Journal of Applied Physics*, 74:4716, 1993.

[31] M. Załużny. Intersubband polariton effects in metal-clad optical waveguides. *Solid State Communications*, 97:809, 1996.

Index