LBNL-40583
UC-000

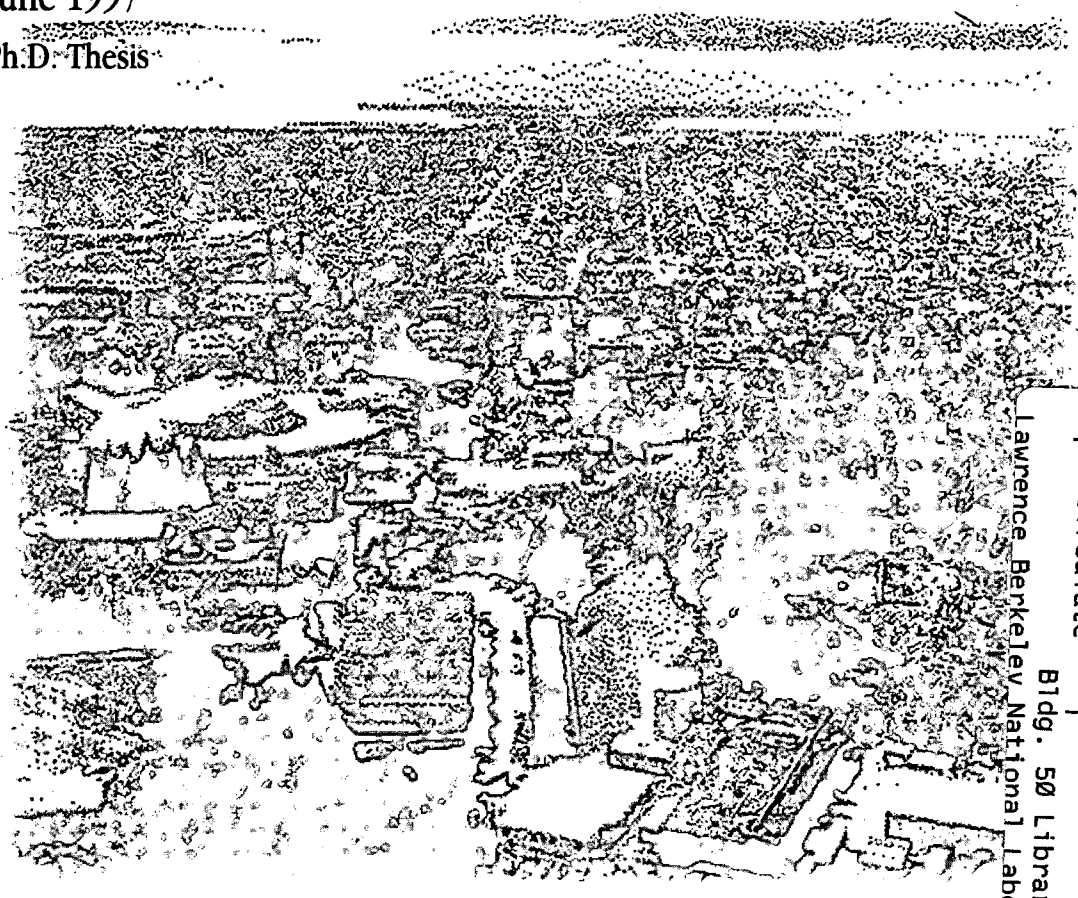
ERNEST ORLANDO LAWRENCE
BERKELEY NATIONAL LABORATORY

Donor Spectroscopy at Large Hydrostatic Pressures and Transport Studies in Compound Semiconductors

Leonardo Hsu

Materials Sciences Division
Center for Advanced Materials

June 1997
Ph.D. Thesis



DISCLAIMER

LBNL-40583
UC-000

Donor Spectroscopy at Large Hydrostatic Pressures and Transport Studies in Compound Semiconductors

Leonardo Hsu

Department of Physics
University of California, Berkeley

and

Center for Advanced Materials
Materials Sciences Division
Ernest Orlando Lawrence Berkeley National Laboratory
University of California
Berkeley, California 94720

Ph.D. Thesis

June 1997

This work was supported in part by US NSF grants DMR-91 15856 and DMR-94 17763 and in part by the Director, Office of Energy Research, Office of Basic Energy Sciences, Materials Sciences Division, of the U.S. Department of Energy under Contract No. DE-AC03-76SF00098.

Donor Spectroscopy at Large Hydrostatic Pressures and Transport Studies in Compound Semiconductors

by

Leonardo Hsu

A. B. (Harvard University) 1991

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Physics

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, BERKELEY

Committee in charge:

Professor Eugene E. Haller, Co-Chair
Professor Peter Y. Yu, Co-Chair
Professor Steven G. Louie
Professor Raymond Jeanloz

Fall 1997

Abstract

Donor Spectroscopy at Large Hydrostatic Pressures and

Transport Studies in Compound Semiconductors

by

Leonardo Hsu

Doctor of Philosophy in Physics

University of California, Berkeley

Professor Eugene E. Haller, Co-Chair

Professor Peter Y. Yu, Co-Chair

In the first part of this work, I describe studies of donors in AlSb and in GaAs at large hydrostatic pressures, two materials in which the conduction band minimum is not parabolic, but has a "camel's back shape." These donors were found to display only one or two absorption lines corresponding to ground to bound excited state transitions. This is in contrast with donors in parabolic-minimum semiconductors such as Si or Ge in which many such transitions can be seen due to the infinite number of bound excited states. It is shown that due to the non-parabolic dispersion, camel's back donors may have as few as one bound excited state and that higher excited states are auto-ionized. Thus, it is possible that transitions to these other states may be lost in the continuum.

In the second part, calculations of mobilities in GaN and other group III-Nitride based structures were performed. GaN is interesting in that the carriers in nominally undoped material are thought to originate from a impurities which have an ionization energy level resonant with the conduction band, rather than located in the forbidden gap. These donors have a short range potential associated with them which can be effective in scattering electrons in certain situations. It was found that effects of these resonant donors can be seen only at high doping levels in III-Nitride materials and in $Al_xGa_{1-x}N$ alloys, where the defect level can be pushed into the forbidden gap.

Calculations were also performed to find intrinsic mobility limits in $Al_xGa_{1-x}N/GaN$ modulation doped heterostructures. Theoretical predictions show that electron mobilities in these devices are capable of rivaling those found in the best $Al_xGa_{1-x}As/GaAs$ heterostructures structures today. However, the currently available nitride heterostructures, while displaying mobilities superior to those in bulk material, have sheet carrier concentrations too large to display true two-dimensional electron gas behavior.

Table of Contents

Donor Spectroscopy at Large Hydrostatic Pressures in Compound Semiconductors

Transport Studies in Group III-Nitride Semiconductors

Acknowledgments

The writing of this dissertation and my completion of graduate school would not have been possible without the help of a great many people. First and foremost, I would like to thank Professor Eugene E. Haller for his generous support and guidance during my years here at Berkeley. Providing not only assistance through numerous discussions of my work, but also help with the repairing of equipment in the lab, his aid has been invaluable at all levels. My thanks are also due to Dr. Wladek Walukiewicz for his help in understanding many of the theoretical aspects of physics and for his guidance in performing the second half of this work. I am further indebted to Professor Peter Yu for acting as my departmental advisor in reading this manuscript and for many educational discussions regarding high pressure techniques and the physics of shallow donors. I would also like to express my gratitude to Dr. Jeff Wolk, who donated much of his time to help me get started with this work.

Jeff Beeman deserves much credit for creating and repairing the detectors without which this work would not have been possible and for his invaluable assistance with all manner of laboratory equipment and techniques. I would like to express my thanks to Bill Hansen for sharing his extensive lab expertise, to Edith Bourret-Courchesne for her help in all matters concerning GaAs, and to Professor Raymond Jeanloz for his help with the high pressure work and allowing me to use his high pressure laboratory, as well as reading this dissertation. I am grateful to Anthony Chen for his help with the high pressure work and to Professor Steven Louie for taking the time to review this manuscript. I also thank Margaret Ragsdale and David Hom for taking care of the great majority of the administrative and bureaucratic tasks for me. I am grateful to Professor Steven Louie for taking the time to read this manuscript.

My graduate school experience has also been enriched by my fellow graduate students, especially Matt McCluskey, who have not only taught me a lot, but have been

fine card playing companions at lunch. I would particularly like to thank Oscar Dubon (now Dr.) for his help with various aspects of spectroscopy.

My parents also deserve much credit for supporting me in this endeavor. I especially thank my father for sparking my initial interest in physics and to my mother for providing constant encouragement.

Finally, I would like to acknowledge the support of the Office of Naval Research through their fellowship program, the National Science Foundation through grants DMR-91 15856 and DMR-94 17763, and the Department of Energy under contract No. DE-AC03-76SF00098.

Donor Spectroscopy at Large Hydrostatic Pressures in Compound Semiconductors

1. Introduction

Semiconductors are traditionally defined as materials which have an energy gap between 0 and 3 eV. The basis of nearly all electronic devices, the usefulness of semiconductors stems from the fact that their resistivity can be made to vary from 10^{-4} Ω·cm to more than 10^{9} Ω·cm with the addition of small amounts of impurities. Today, the concentration and types of impurities present in these materials can be controlled very precisely, enabling the production of a wide variety of solid state devices.

Certain impurities affect the electrical properties of semiconductors by introducing energy levels into the band gap. These electrically active impurities can be divided into two categories, donors and acceptors. Donors are impurities which, by virtue of the number of electrons in their valence shell, are able to give up, or donate, electrons to the conduction band of the semiconductor. In general, the energy levels of donors are found in the upper half of the band gap. Acceptors are impurities which are able to bind, or accept, conduction electrons from the semiconductor and their energy levels are usually found in the lower half of the band gap. Each of these two categories may be further subdivided into shallow and deep impurities. The energy levels introduced by shallow impurities are usually only a few tens of meV from the edges of the bandgap, whereas the energy levels of deep impurities are separated from the band edges by a significant fraction of the band gap. As a general rule, impurity atoms which are found one column to the left or right in the periodic table of the lattice atom they replace form shallow acceptors and donors, respectively. The focus of this work is the behavior of the energy levels of shallow donors and we will henceforth limit the discussion to this type of electrically active impurity.

1.1 Effective Mass Theory

Effective mass theory is a simple, yet accurate way of calculating the electronic properties of shallow donors in semiconductors. In principle, such a calculation would require solving a Schrödinger equation for roughly 10^{23} interacting particles, a problem which is clearly intractable even for today's fastest computers. However, due to the structural properties of semiconductors, a number of simplifying approximations can be made. These approximations and the model with which the electronic structure of shallow impurities in semiconductors is calculated make up effective mass theory. This formalism was created by Kittel and Mitchell (Kittel 1954) and developed by Kohn and Luttinger (Luttinger 1955) and Faulkner (Faulkner 1969) among others. A simple physical picture of the results is presented here. A detailed mathematical description of effective mass theory can be found in appendix A.

A perfect semiconductor crystal at 0 K has a set of fully occupied valence bands and a set of completely unoccupied conduction bands. If an additional electron is introduced, it will occupy the lowest minimum of the conduction band and have an energy of

$$E = E_G + \frac{\hbar^2 k^2}{2 m^*} \tag{1.1}$$

relative to the top of the valence band where E_G is the energy gap of the semiconductor, **k** is the wavevector of the electron, and m^* is the effective mass of the electron in the crystal, which is proportional to the inverse of the band curvature. Comparing this expression with that for the kinetic energy of a free electron, we see that the only difference is the addition of the band gap energy E_G and the use of the electron's crystal mass rather than its mass in free space.

If a lattice atom is replaced with one which is found one column further to the right in the periodic table, a crystal with one extra positive charge at one of the lattice sites is produced. This positive charge generates a Coulomb potential which, at distances further

than a few lattice constants, is screened by the other ions of the lattice. This screened Coulomb potential can be written

$$V'(r) = \frac{e}{\varepsilon r} \tag{1.2}$$

where ε is the static dielectric constant of the semiconductor. The extra valence electron of the donor atom is bound by this weakened Coulomb potential in the same manner that a electron is bound to a proton to produce a hydrogen atom, with bound state energies given by a Schrödinger equation similar to that for a hydrogen atom

$$\left(\frac{-\hbar^2}{2m^*}\nabla^2 - eV'(r)\right)\psi = E\,\psi \tag{1.3}$$

The conduction band states of the crystal are analogous to the continuum of energies available to the electron of an ionized hydrogen atom. Bound states of the electron are found in the gap at energies given by

$$E = E_G - \left(\frac{e^2 m^*}{2\hbar^2 \varepsilon^2}\right)\frac{1}{n^2} \qquad n = 1,2,... \tag{1.4}$$

where E=0 at the top of the valence band. The ground state envelope wavefunction of the extra valence electron is a hydrogenic 1s wavefunction scaled with an effective Bohr radius of

$$a_0^* = \frac{\varepsilon \hbar^2}{m^* e^2} \tag{1.5}$$

In a typical semiconductor, values of the dielectric constant are on the order of 10, while effective masses are roughly one tenth the mass of a free electron. This combination of factors produces binding energies of about one one-thousandth of a free electron Rydberg,

or about 10 meV, with effective Bohr radii of about 50 Å. With such a large Bohr radius, the wavefunctions of the bound electron states extend over tens of thousands of lattice points, giving a result which is self-consistent with the assumptions made in effective mass theory.

In the case of semiconductors which have an anisotropic conduction band minimum, such as Si or Ge, one can find the bound state energy spectrum simply by modifying (1.3) to take the anisotropic effective mass into account as follows:

$$\left(\frac{-\hbar^2}{2\,m_{\perp}^{*}}\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}\right)+\frac{-\hbar^2}{2\,m_{\parallel}^{*}}\frac{\partial^2}{\partial z^2}+V(r)\right)\psi = E\,\psi \qquad (1.6)$$

The results predicted by effective mass theory can be checked experimentally using techniques such as infrared absorption spectroscopy where one measures absorption peaks corresponding to the energies of allowed transitions between these hydrogenic levels or photothermal ionization spectroscopy (PTIS), in which the same transitions are revealed as peaks in a photoconductivity spectrum.

Comparing theory to experiment, one finds that while the energies of odd-parity bound excited states (those corresponding to the p or f states of a hydrogen atom) can be predicted quite well, the theoretical ground state energy of shallow donors is often very different from the experimentally measured value. The reason for this discrepancy can be found in the shape of the wavefunctions. The ground state wavefunction of a hydrogen atom is non-zero at the origin. In contrast, the p and f state wavefunctions have a vanishing probability density at their centers. Although the screened Coulomb potential which we have up to now employed is a good approximation at distances of several lattice constants from the impurity atom, it is not at all a good approximation very close to the impurity atom, where the potential varies very quickly and in a way which depends on the exact species of impurity atom. Thus, the p- and f-like bound state wavefunctions of shallow donors are largely unaffected by the local potential in the immediate vicinity of the

impurity and have energies which are well described by effective mass theory. On the other hand, the ground state of the donor electron, which interacts strongly with the "central-cell" potential of the impurity, has an energy which cannot be accurately calculated by simple effective mass theory, which does not take the chemical species of the impurity into account. This difference between the true binding energy and that obtained from effective mass theory is known as a chemical shift. Experimentally, the effect of a chemical shift between two different impurities is seen as a constant shift in the energies of the transition lines. The spacings between the peaks, however, are unaffected.

1.2 Background and Motivation

The entirety of this work grew out of an attempt to understand the infrared absorption spectrum of the Si shallow donor in indirect-gap GaAs. At atmospheric pressure, GaAs is a direct gap semiconductor with a band gap of approximately 1.4 eV. The minimum energy point of the conduction band occurs at the Γ symmetry point and properties of shallow donors in this material are determined by the conduction band wavefunctions near that point. These Γ-band donors have been thoroughly investigated by various characterization methods and their properties and behavior are by now understood quite well (Narita 1971, Bose 1987, Stillman 1994) . When sufficient hydrostatic pressure is applied, however, GaAs can be transformed into an indirect gap semiconductor (Yu 1978, Wolford 1985).

It has been known for some time that hydrostatic pressure can be used to modify the band structure of semiconductors (Armistead 1985). A general trend observed in all semiconductors is that as higher pressures are applied, the energy of the conduction band near the Γ and L symmetry points increases relative to the top of the valence band while the energy of the conduction band near the X symmetry point decreases (Lee 1985). Thus, if

the conduction band has its minimum energy near the Γ or L points at atmospheric pressure, it is possible to shift this minimum to the X point due to the opposite signs of the pressure coefficients. In GaAs, this transition from direct to indirect gap semiconductor occurs at approximately 40 kbar, as shown in figure 1-1. While in this indirect gap configuration, shallow donors in GaAs possess properties which are influenced by the conduction band wavefunctions near the X point of the Brillouin zone and we refer to these impurities as X-band donors. Although numerous studies of X-band donors in GaAs have been undertaken (Kobayashi 1983, Leroux 1986, Liu 1990, Dmochowski 1994, Leroux 1995), all of them without exception were performed using photoluminescence rather than absorption spectroscopy and thus, no information about the excited states of the donors could be obtained. In contrast, by using broadband infrared absorption spectroscopy as we have, it is, in principle, possible to deduce the entire bound excited state spectrum of these donors.

Initially, it was expected that the Si donor absorption spectrum in indirect GaAs would look similar to the absorption spectra of shallow hydrogenic donors in Si (Jagannath 1981) (see figure 1-2), since Si also has its conduction band minimum near the X symmetry point. From the energies of the absorption lines, effective mass theory could be used to determine the transverse and longitudinal effective masses associated with the X conduction band minimum in GaAs. However, unlike the spectra of donors in Si, only one transition line was observed for Si in indirect GaAs. To determine whether this lack of spectral features is characteristic of all shallow donors in indirect GaAs or peculiar to the Si shallow donor, the remaining elements which form shallow donors in GaAs were also investigated. Before discussing the results of those studies, however, I shall describe the experimental techniques used in this work.

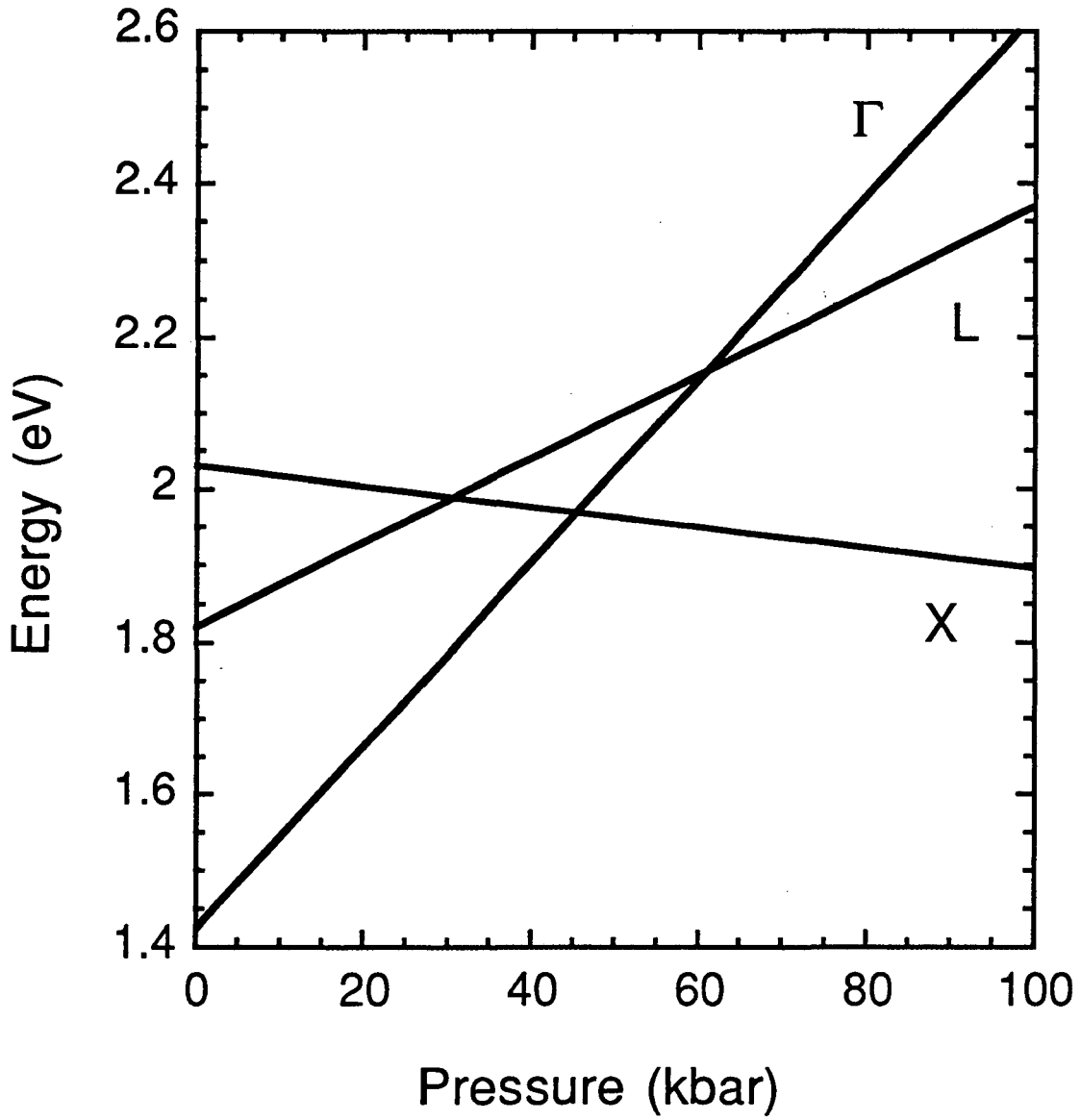


Figure 1-1. Plot of the energies of the local conduction band minima at various symmetry points for GaAs as a function of pressure. Energies are relative to the top of the valence band.

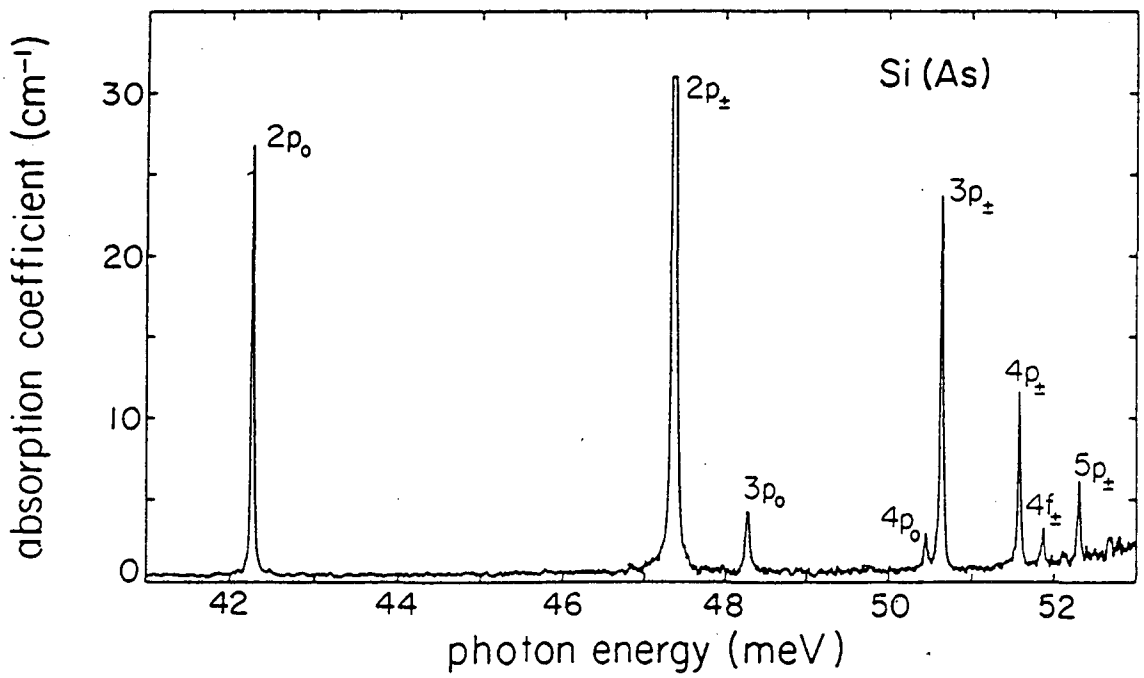


Figure 1-2. Absorption spectrum of arsenic donors in silicon recorded at liquid helium temperatures (Jagannath 1981).

2. Experimental Techniques

The principle characterization method used in this study was infrared absorption spectroscopy. The first section provides an introduction to Fourier transform infrared spectroscopy, describing the experimental apparatus and principle behind its workings. In the second section, details of the diamond anvil cell (DAC) used to apply large hydrostatic pressures to the samples along with some of the specialized methods used to perform infrared spectroscopy through the cell are presented.

2.1 Fourier Transform Spectroscopy

The instrument used to investigate X-band donors in III-V semiconductors was a Digilab FTS-80E Fourier transform spectrometer. Like all Fourier transform spectrometers, the core of this machine is a Michelson interferometer, which is shown in figure 2-1 and consists of a broad band light source, one fixed and one moving mirror, a beamsplitter, and a detector. Light from the source travels to the beamsplitter, where half of the beam is transmitted through to the moving mirror and half is reflected 90° towards the fixed mirror. The mirrors reflect the light back to the beamsplitter where the two beams combine and interfere and are directed towards the detector. As the moving mirror changes its position, the intensity of the light reaching the detector also changes and the detector response is recorded as a function of the position of the moving mirror to produce an interferogram. By performing a Fourier transformation, the interferogram is converted into a spectrum showing the detector response and indirectly, the intensity of the beam, as a function of the frequency of the light.

To see why this is the case, consider a situation in which the interferometer's light source is monochromatic with wavelength λ. If the optical path length from the beamsplitter to each of the two mirrors, moving and fixed (hereafter referred to as M and F), are equal, then the two beams which were in phase when split by the beamsplitter will

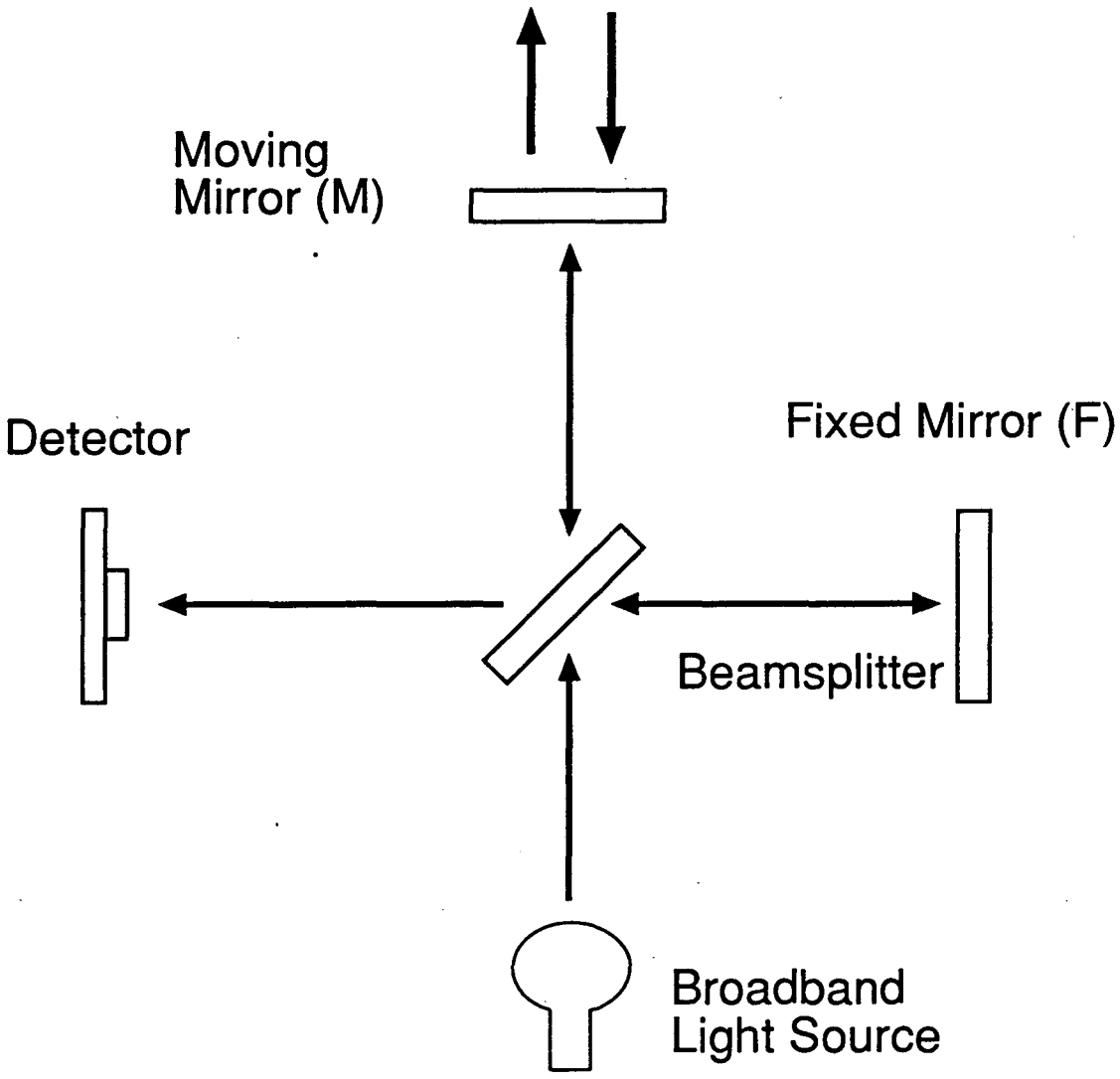


Figure 2-1. Schematic diagram of the essential pieces of a Michelson interferometer.

still be in phase when they recombine. The intensity of the light reaching the detector is then a maximum. This position is known as "zero path." As the moving mirror is translated, the two beams are no longer precisely in phase when they recombine and the resultant intensity reaching the detector is smaller. When mirror M has moved outwards by a distance $\lambda/4$, the optical path length of that arm of the interferometer will have increased by $\lambda/2$. In this case, the two beams will be exactly out of phase when they recombine and interfere destructively, producing a minimum in the intensity of the light reaching the detector. In general, this intensity can be expressed as

$$I(\theta) = I_0 \cos^2 \theta \tag{2.1}$$

where I_0 is the intensity incident on the detector at zero path and θ is the phase angle between the recombining beams, which can also be written as $\theta = 4\pi\Delta/\lambda$ where Δ is the displacement of the moving mirror from the zero path position. If mirror M moves with a known velocity v, the phase angle can also be expressed as $\theta = 4\pi vt/\lambda$.

Assuming that the detector response is proportional to the photon flux incident on the detector, the interferogram will be a periodic function with a frequency of $4v/\lambda$ as can be seen from (2.1). In the idealized case where the moving mirror has an infinite travel, taking the Fourier transform of such an interferogram produces a delta function, a spectrum which is zero everywhere except for a single infinitely narrow peak located at a frequency of $4v/\lambda$. However, since mirrors in real spectrometers can only move a finite distance, the resultant peak is not an infinitely sharp delta function, but has some non-zero width corresponding to the uncertainty in the exact frequency of the cosine wave. An exact calculation (Bell 1972) shows that for the case of a monochromatic light source and an interferometer with a finite maximum mirror displacement, the resulting peak has the form

of a sinc function. The width of the peak given by this function is the fundamental instrument-limited resolution and has a full width at half maximum (FWHM) of 0.61/L wavenumbers where L is the maximum mirror displacement in cm. In practice, one uses a mirror displacement which is large enough so that the instrumental resolution is finer than the linewidths of the peaks under study in order to fully resolve them.

In the spectrometer used for our studies, the source is not monochromatic, but emits radiation over a broad range of frequencies. The intensity incident on the detector is not a simple squared cosine function but the sum of such cosine waves from many different frequencies. At zero path, all of the functions have their maximum value due to constructive interference of all frequencies of the light and an absolute maximum is produced in the interferogram. As the moving mirror is displaced from zero path, the intensities of the various frequency components of the beam are modulated at different rates, producing the complicated interferogram shown in figure 2-2. The corresponding spectrum after a Fourier transformation is a function which shows the relative amplitudes of the various frequency components of the interferogram and is shown in figure 2-3. Due to the electronic filters in the Digilab FTS-80E, the zero path of the interferogram is displayed as a negative-going peak instead of an absolute maximum.

In absorption spectroscopy, one studies the absorptions of the impurities of interest which are present in a sample by placing the sample in front of the detector. These absorptions are often narrow in frequency and appear as dips in the transformed spectrum, occurring when light of a certain frequency is absorbed by the impurities and never reaches the detector. However, as can be seen from figure 2-3, the detector response is usually a very complicated shape with features resulting from particulars of the source and the detector, as well as absorptions from the various filters, cryostat windows, and other optical components of the system in addition to those of the sample. In order to separate the absorptions of interest from the others, the usual procedure is to collect two spectra. The first is obtained with the sample of interest. The second is a reference spectrum, taken

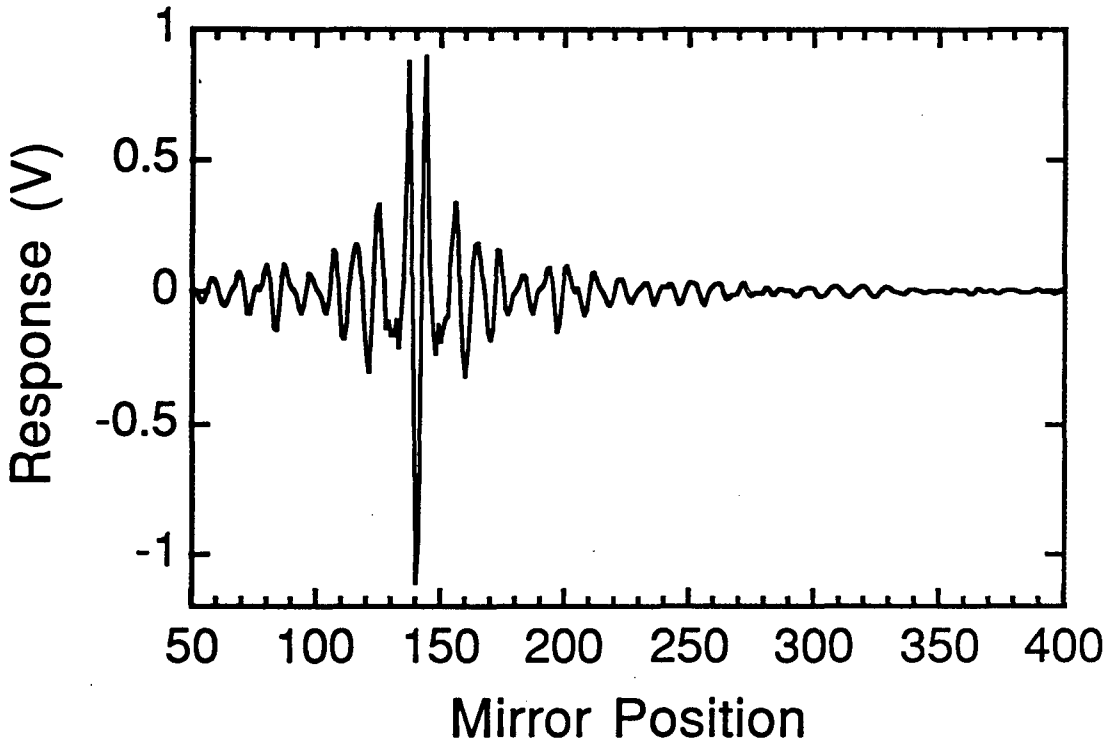


Figure 2-2. Sample interferogram taken by the Biorad FTS 80E Fourier Transform Spectrometer. Each point corresponds to 1.25 μm of mirror displacement.

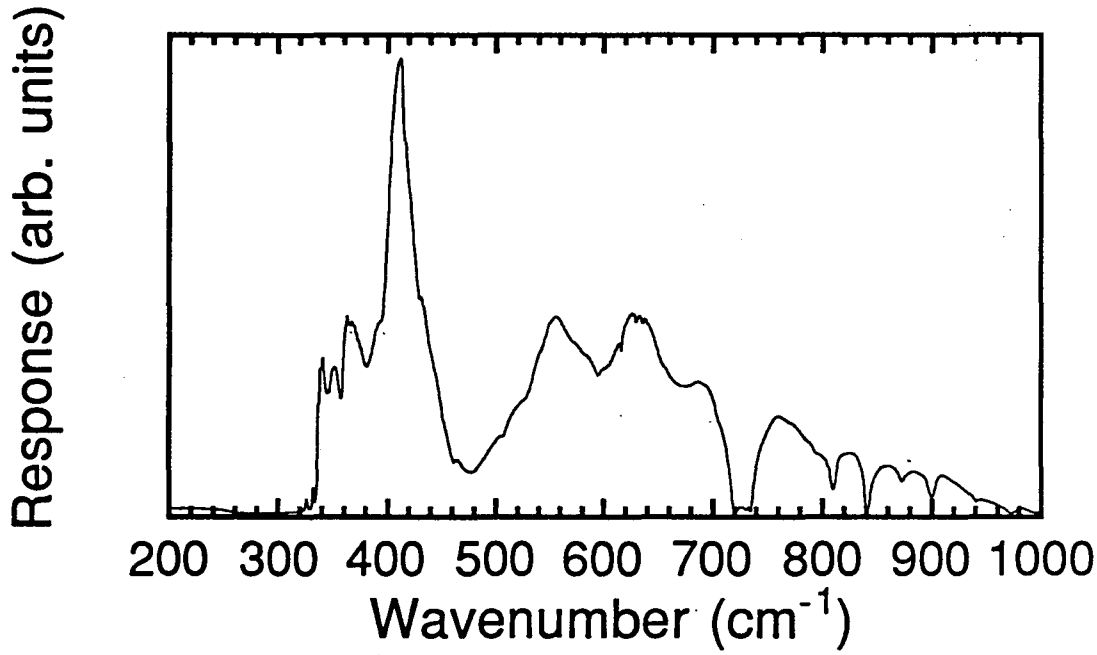


Figure 2-3. Detector response spectrum corresponding to the interferogram in figure 2-2.

using a piece of material which is identical to the sample of interest except that it does not contain any of the impurities under study. Making a direct comparison of the two spectra, one can see which absorptions are due to instrumental features, since these will appear in both, and which absorptions are due to the impurity centers of interest, as these will not be present in the reference spectrum.

The unit used to quantify the absorption of light is the linear absorption coefficient α, which is the product of the defect concentration and the absorption cross section of the defect. If $I_0(\nu)$ is the intensity of light of a particular frequency reaching the detector through a sample which is free of a particular impurity, then $I(\nu) = I_0(\nu)\, e^{-\alpha(\nu)x}$ is the intensity of light reaching the detector after passing through a sample which does contain the impurity where x is the thickness of the sample. Taking the negative log of the ratio of intensities I and I_0, one obtains

$$-\ln\left(\frac{I}{I_0}\right) = -\ln\left(e^{-\alpha(\nu)x}\right) = \alpha(\nu)\, x \qquad (2.2)$$

The frequency dependent quantity $\alpha(\nu)x$ is known as the absorbance. By performing this operation on the two spectra, an absorbance spectrum is formed in which absorption characteristics of the impurities of interest are revealed as peaks and spectroscopic features due to particulars of the instrument are suppressed.

The chamber of the spectrometer used here is evacuated during the measurements to eliminate infrared absorptions due to water vapor and other air molecules such as CO_2. A SiC globar heated to 1200°C was used as a source of blackbody radiation. Three different detectors were employed: a Ge:Ga photoconductor sensitive to radiation between 90 cm^{-1} and 1000 cm^{-1}, a Ge:Be photoconductor with an effective operating range from 190 cm^{-1} to

2000 cm^{-1}, and a Ge:Cu photoconductor which was most effective at detecting photons with energies between 350 cm^{-1} and 3500 cm^{-1}.

These doped Ge detectors must be cooled to liquid helium temperatures (4.5 K to 10 K) where the dopants are frozen out to function properly. Their spectral range is determined by the ionization energy of the impurity at low frequencies and by the decreasing photoionization cross section of the impurities at high frequencies. When a photon of sufficient energy is absorbed by one of the dopants, a free hole is created which drifts in the electric field created by the bias voltage across the detector and is captured by an ionized acceptor after a certain time. These carriers create a current in the external circuit which is converted into a voltage signal by a transimpedance amplifier and the time varying voltage signal is read as an interferogram by the spectrometer electronics.

Two different beamsplitters were used: a mylar foil 3 μm thick which was useful for frequencies between 0 and 1000 cm^{-1} and a beamsplitter made of a thin layer of Ge between two wedged crystals of KBr which was useful for studying absorptions between 450 cm^{-1} and 3500 cm^{-1}. Using different combinations of these optical components, the entire energy range between 90 cm^{-1} and 3500 cm^{-1} could be investigated.

2.2 The Diamond Anvil Cell

2.2.1 Introduction

The attainment of ever increasing pressures available for research has progressed in parallel with the development of harder high strength substances. In the early twentieth century, Percy Bridgman first used hardened steels and then sintered tungsten carbide to produce vessels that could generate up to 100 kbar of pressure (Bridgman 1952). The highest pressures that could be obtained were limited by the compressive strength of the anvils. Although diamond, the hardest substance known, may seem a logical material with which to construct pressure anvils, it was not until 1950 that a brief attempt was made to

generate high pressures using diamond anvils (Lawson 1950). The use of diamond as an anvil material was then forgotten until 1959, when the first diamond anvil cells (DAC) were introduced.

The first cells were of two types, developed independently. One was designed to be used in powder X-ray diffraction experiments at the University of Chicago (Jamieson 1959). The second was developed at the National Bureau of Standards (NBS) for infrared transmission experiments (Weir 1959). Since that time, numerous other designs have appeared for various applications, some able to provide hydrostatic pressures of up to a few megabars.

The basic operating principle of a diamond anvil cell (DAC) is very simple. As can be seen in figure 2-4, two diamonds with flat parallel faces (the anvils) are forced together to apply pressure on a sample. The diamonds themselves are cut in the same way as those for jewelry purposes except that the point normally opposite the table is ground and polished down to a flat surface known as the culet. In a modern cell, the sample and a pressure transmitting medium (usually a fluid) are enclosed in a cylindrical space formed by a hole in a metal gasket with dimensions of up to several hundred microns in height and diameter. When the diamonds are pressed together, they apply pressure to the medium which is confined by the gasket walls. The pressure on the fluid is transmitted to the sample hydrostatically. In the earliest cells, however, the sample was simply squeezed between the diamond faces, producing a uniaxial stress. The diamond anvil cell is particularly convenient for many experiments since the diamonds allow optical access to the sample over a wide range of photon energies.

The most important function of the DAC itself, besides that of applying force to the diamonds, is to keep the stresses on the diamonds as small and homogeneous as possible by keeping the diamond faces parallel as well as translationally aligned. During an experiment, enormous forces act on the diamonds and inhomogeneous stresses will be generated if they are not kept perfectly aligned, possibly causing catastrophic failure.

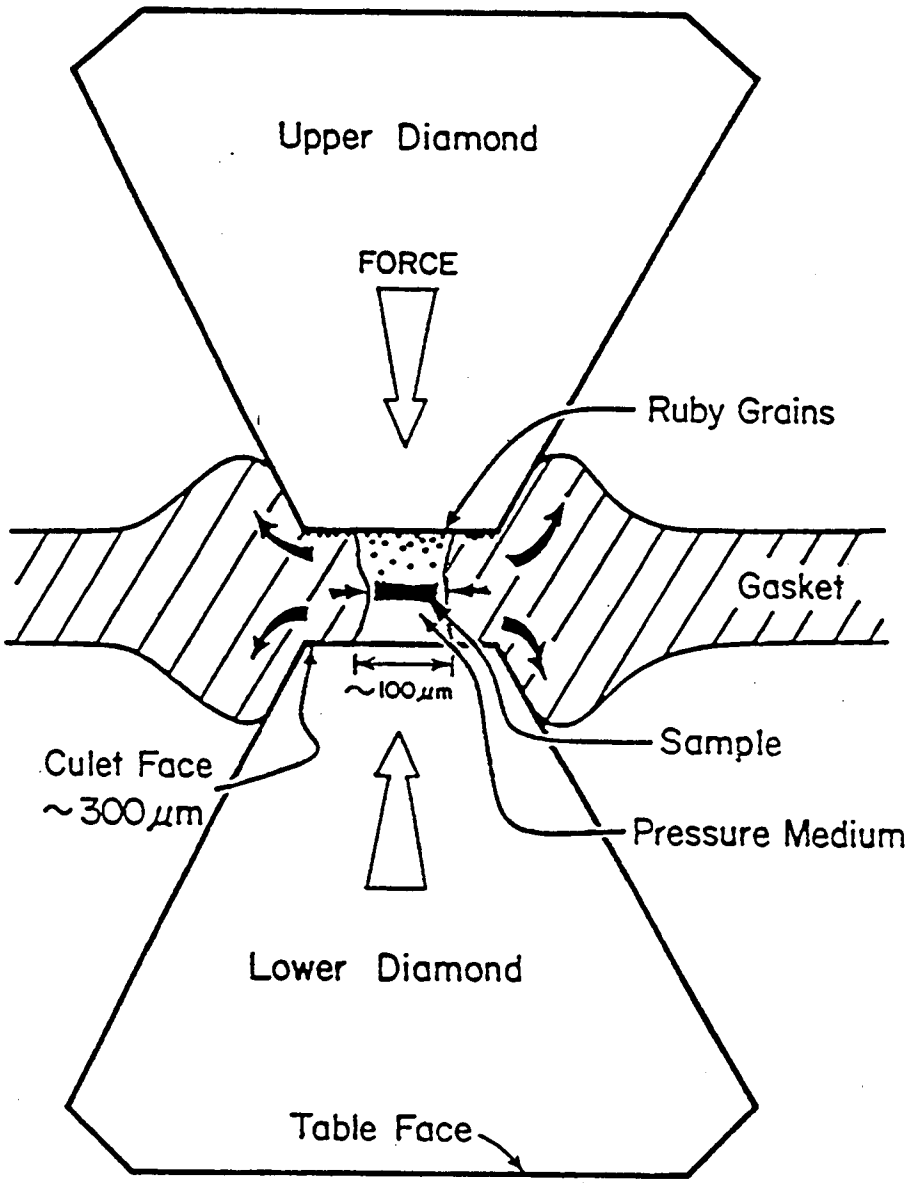


Figure 2-4. Magnified view of the essential parts of the diamond anvil cell. The diamonds are forced together, squeezing the gasket and pressure medium, which in turn applies hydrostatic pressure to the sample (Williams 1991).

The general philosophy of the method used to align the diamonds is shown in figure 2-5. In this cell, which was designed by Piermarini and Block (Piermarini 1975) at NBS, the top diamond is held by a translating diamond mount plate and the bottom one rests in a tilting diamond mount hemisphere. The translating diamond mount plate can be displaced laterally by means of three adjustment screws situated 120° apart around the plate. The hemisphere of the bottom mount allows the culet of the bottom diamond to be tilted until it is parallel to the culet of the top diamond, again by means of three symmetrically placed adjustment screws. During the application of pressure, the extended piston which holds the lower diamond is forced into a close-fitting cylindrical cavity which is part of an assembly holding the upper diamond, ensuring that once the diamonds are aligned, they stay aligned. Pressures of up to 500 kbar have been obtained with this type of cell (Jayaraman 1983).

The cell used in the work presented here and shown in figure 2-6 is a Merrill-Basset type cell designed by Sterer, Pasternak, and Taylor (Sterer 1990). Instead of forcing a piston into a close-fitting cylinder, two circular plates or platens which hold the backing plates on which the diamonds are mounted are held together by six screws spaced evenly about their rim. As in the NBS cell, one of the diamond backing plates (the lower one) can be translated laterally by means of three adjustment screws. To make and keep the culet faces parallel, the six screws around the rim are tightened in such a way as to eliminate the appearance of optical fringes between the diamond culets when they are placed together and the thickness of the cell is then measured at three points around the rim. During the application of pressure, the tightness of the screws is adjusted so as to keep the relative thicknesses around the rim constant. Three pins jutting from the bottom plate mate with three holes in the top plate to help maintain the alignment. Although it is more difficult to maintain alignment of the diamonds, the Merrill-Basset type cell has the advantage in that it can be made smaller than any of the other types of cells. Pressures of up to about 380 kbar have been reached using this type of cell (Sterer 1990).

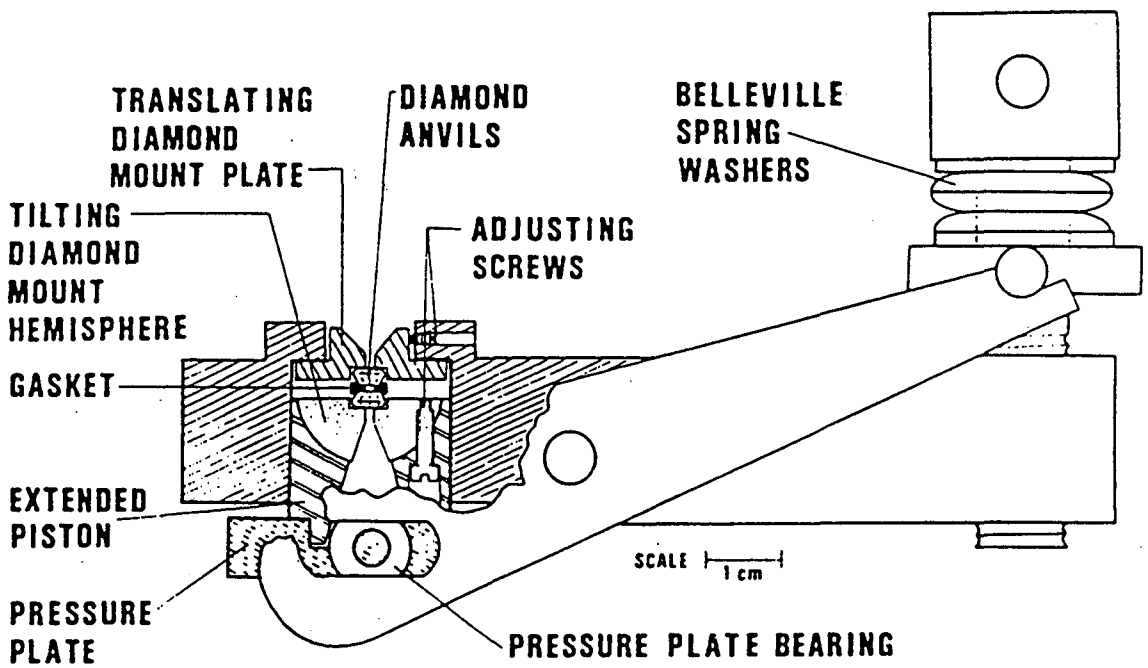


Figure 2-5. Diamond anvil cell designed by Piermarini and Block at the National Bureau of Standards. The parts used in alignment of the diamonds are indicated (see text) (Piermarini 1975).

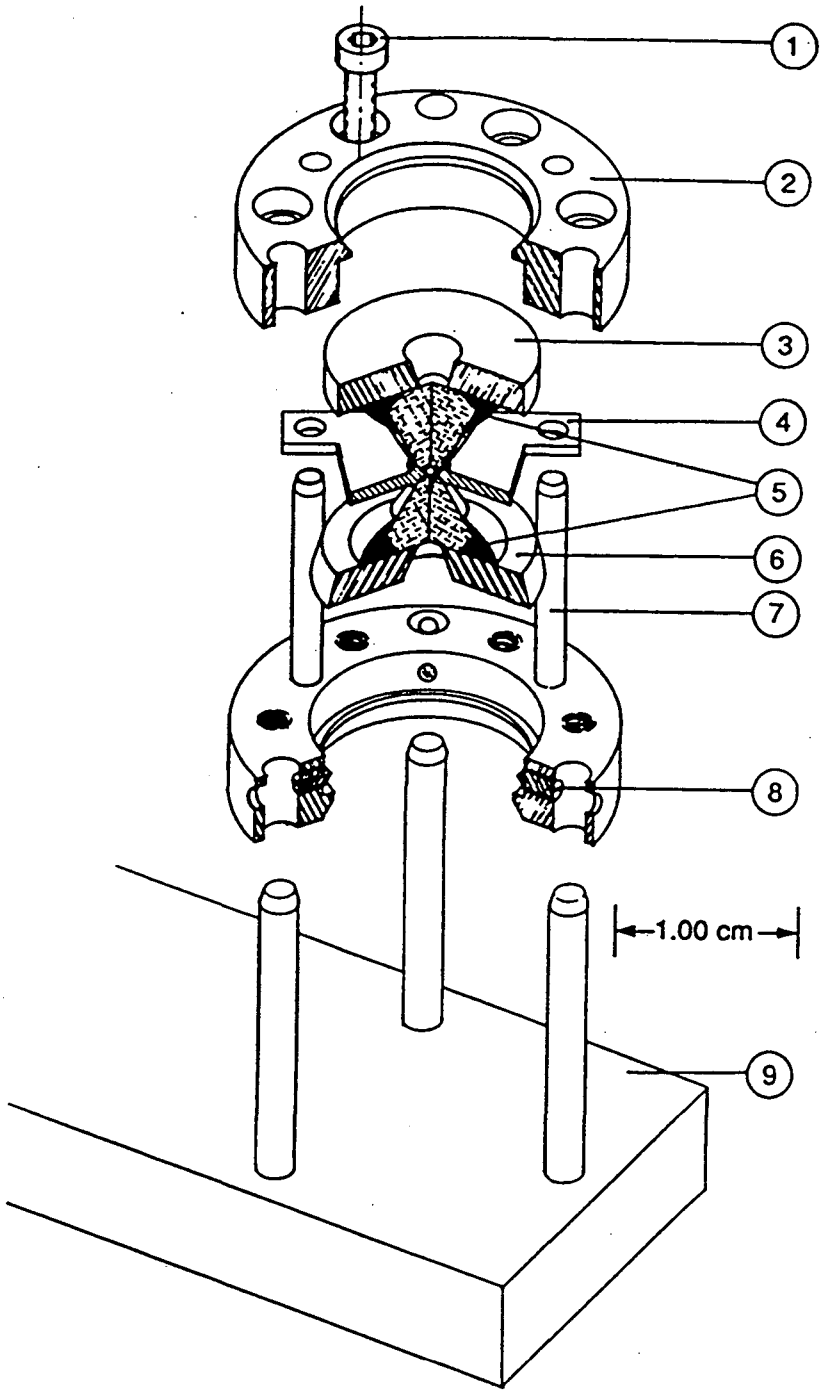


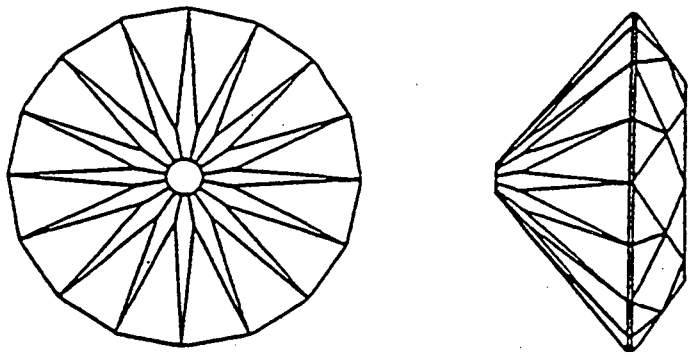
Figure 2-6. Exploded view of the diamond anvil cell used in the work described here. The parts indicated are: (1) allen head screw used to hold the cell together and apply pressure, (2) platens, (3) upper backing plate (fixed), (4) gasket, (5) diamonds, (6) lower backing plate (adjustable), (7) posts to aid in alignment of platens and gasket, (8) lateral adjustment screws for the lower backing plate, and (9) cell holder used when tightening allen screws. (Sterer 1990)

2.2.2 Diamonds

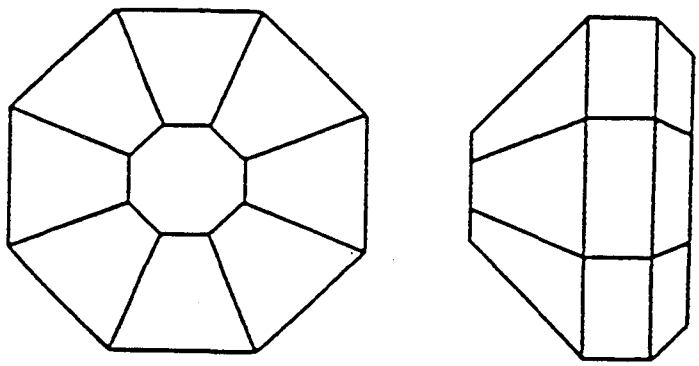
The diamonds used in a DAC range in size from 1/8 carat for the small Merrill-Basset cells up to 1/3 carat in other cells. Culets are either octagonal or 16-sided in shape, and range in size from 0.2 to 0.8 mm in diameter. For experiments requiring very high pressures, it is advantageous to use larger diamonds along with small, 16-sided culets. In addition, the culet edges, which are the most susceptible to failure, may be beveled in order to relieve the stresses in that area (Jayaraman 1986).

There are two main shapes, or cuts, of diamond used in high pressure work. Both are shown in figure 2-7. The first is the modified brilliant design, which is formed simply by taking a diamond which has already been cut in the shape commonly found in diamond jewelry, and grinding off the tip to produce the culet. This was the first design used for diamond anvils due to its great availability. The second shape is known as the standard design, and is shaped in such a way as to place more of the diamond's volume in the critical area below the culet (Seal 1987). This strengthens the anvil for use in very high pressure experiments.

Diamonds can also be classified by the types of impurities present in them and depending on the experiment to be performed, different types of diamonds may be better suited for certain purposes. All diamonds can be divided into two categories, type I and type II. Type I diamonds are characterized by the presence of nitrogen impurities and are also known as "yellow" diamonds due to their slight yellowish coloration. The nitrogen is present in several forms, including single interstitial atoms, pairs, and larger aggregates (Davies 1977). These defects produce strong absorption bands in the 1100 cm^{-1} to 1400 cm^{-1} range, causing the diamonds to be opaque to those frequencies of light. Type II diamonds contain virtually no nitrogen and are transparent in that range. Type II's can be further classified into type IIa diamonds, which contain very few of any impurities and type IIb diamonds, which are synthetic, doped with boron, and very expensive. Both stypes of diamonds have strong absorptions in the region from 1900 to 2300 cm^{-1} due to two-



(a)



(b)

Figure 2-7. Side and top view of diamond anvils. (A) shows the modified brilliant cut and (B) shows the standard cut. (Seal 1987)

phonon excitations. Since type I diamonds are much more common and less expensive than the others, they are used in all situations except those in which one requires optical access in the 1100 to 1400 cm^{-1} region. Diamonds of both types were used in the present work, though the vast majority were of type I.

2.2.3 Gaskets

As mentioned previously, in the first diamond anvil cells, the samples were simply squeezed between the bare diamond faces, generating more of a uniaxial stress rather than a hydrostatic pressure. In 1965, however, Van Valkenburg (Van Valkenburg 1965) introduced the use of gaskets. The gasket itself is a piece of metal roughly 0.2 to 0.25 mm thick and can be made of inconel, spring steel, stainless steel, BeCu, or some other reasonably hard material. In the loading process, the gasket is pre-indented between the diamonds to a thickness from 0.1 to 0.15 mm and a hole of a few hundred microns diameter is drilled through the material in the center of the depression. This hole serves as the sample space, inside which the sample and pressure medium are sealed. Since the sample is completely surrounded by a fluid, the pressure on the sample is hydrostatic. In addition, the gasket acts as a support ring for the culet edges where it has extruded out from between the anvils, helping to prevent diamond failure due to the stresses concentrated there.

2.2.4 Pressure Media

Throughout the almost 40 year history of the DAC, many different pressure transmitting media have been used in diamond anvil cells. The purpose of these fluids is to transform the uniaxial force of the diamonds into a hydrostatic pressure on the sample. A listing of a few of them and the maximum value at which the pressure they transmit remains reasonably hydrostatic is given in table 2-1. The methanol-ethanol mixtures are the easiest to use, since the cell can be loaded at room temperature with no special apparatus.

Medium	Maximum useful pressure (kbar)
Methanol:Ethanol (4:1)	~ 200
N_2	~ 130
H_2	> 600
He	> 600
Ne	160
Ar	90
Xe	300

Table 2-1. Some pressure media and their useful pressure ranges (Jayaraman 1983).

However, this mixture has many absorptions in the infrared and is thus inconvenient for infrared spectroscopy. For the studies described here, N_2 was used as the pressure medium. Although the loading must be performed cryogenically, the procedure itself is not very difficult and is described later. In addition, the use of nitrogen allows one to use the stretch mode of the N_2 and of the CO_2 dissolved in the nitrogen as an *in situ* pressure calibration in some cases. For maximum hydrostaticity at the highest pressures, either He or H_2 must be used, but loading a DAC with either of these fluids requires sophisticated apparatus (Mao 1979, Jayaraman 1986).

2.2.5 Sample Preparation

Although diamond anvil cells are capable of generating tremendous pressures, the size of the sample which can be studied is very limited. The samples used in this work were cut in the shape of small disks, approximately 300 μm in diameter and 50 μm in thickness. The reason for this shape is that in order to obtain the largest possible absorption peaks in infrared spectroscopy, one must minimize the amount of light which does not travel through the sample. Any modulated light which does not pass through the sample cannot be absorbed by impurities and increases the background photon flux incident on the detector. If the background is too large, any change in the photon flux due to absorption by impurities may be lost in the noise. In addition, the sample space drilled into the gasket is round and samples which touch two opposing sides of this space will most likely be subject to some non-uniform stress by the gasket. The shape that best satisfies these two conditions is a circle.

The first step in preparing the samples is to cut or cleave a small piece of material with enough surface area to produce several of the small disks. Using wax, this piece is mounted on a glass slide for ease of handling. If the exposed side of the sample is not already smooth enough for infrared transmission experiments, it is polished using 0.3 μm

Al_2O_3 grit on a felt polishing pad. After polishing, an ultrasonic grinder is used to cut the disks from the material. The grinder works by touching the tip of a shaped cutting tool to the sample surface and, by vibrating the tip up and down rapidly, crushing an outline of the tip into the sample. During the cutting process, a slurry of 3 μm aluminum oxide grit and water is splashed on the sample to aid in the cutting process. The tip used to make DAC samples was made from two hollow cylinders, placed side-by-side. These cylindrical shells were cut from the middle of a No. 23 hypodermic needle. When the grinder was finished, two donut shaped depressions about 120 μm deep were into of the sample surface. The columns left standing in the middle of the rings are 300 μm in diameter and polished on top.

After as many columns as possible are cut into the surface of the material, it is turned over and remounted onto the glass slide. Before turning the piece over however, it is important to ensure that the doughnut shaped depressions in the surface are filled with wax. The samples must now be lapped to the correct thickness. The piece of material, attached to the glass slide is lapped by hand using a 3 μm Al_2O_3 slurry. This process removes the material on which the columns were rooted and leaves free-standing disks of material, surrounded by wax. When these disks have been made thin enough, they are given a final polishing by 0.3 μm grit on a felt polishing pad. The glass slide is placed in acetone to dissolve the wax away and the freed samples are collected with a fine paint brush.

In principle, one could also lap the initial piece of material to the desired thickness before cutting them out with the ultrasonic grinder. However, due to the violent action of the grinder, it was found that many of the samples became demounted from the glass slide and were lost by this method.

2.2.6 Loading the Diamond Anvil Cell

The first step in loading a diamond anvil cell is to center the diamonds over the hole in the middle of the backing plates and to glue them there. For the low temperature experiments performed, it was found that Miller-Stephenson 907 or Stycast 2850 are both suitable epoxies, remaining intact throughout many thermal cyclings of the DAC between liquid helium and room temperatures. As little epoxy as necessary is used for best results. The top backing plate is pressed into the top half of the cell and the bottom backing plate can be roughly centered in the bottom platen using the three set screws. It is important to make sure that the backing plate cavities are clean and free from debris and that the backing plates themselves are completely pushed into their respective halves of the cell. If not, they could shift under the high pressures present during an experiment, changing the relative alignment of the diamond culets. The top and bottom halves of the cell are now brought together gently, so as not to damage the diamonds by hitting them together and three of the six main screws are put in and just barely tightened.

The diamonds must now be aligned, both translationally and tilt-wise. To accomplish the former, the three set screws in the bottom half of the DAC can be manipulated until the two octagonal culets of the diamonds coincide, as seen through a low power microscope. In most cases, the bottom diamond will also have to be rotated to bring the sides of the two culets into coincidence.

At this point, it is very likely that colored interference fringes will be observed between the two diamond faces, an example of Newton's rings. Through a combination of tightening and loosening the three main screws, the fringes can be made to broaden and disappear, indicating that the two faces are now mutually parallel. However, in the process, the lateral alignment will probably be lost. The procedure for the translational alignment is then repeated, checking the parallelism of the culets afterwards and fixing it if necessary. By iterating these two procedures, the culets can be aligned completely. The

thicknesses of the cell at three positions around the rim are now measured and recorded and all three should agree to within 50 μm. If they do not, it is probable that some dirt which should be removed is underneath one of the backing plates.

The next step is to make the gasket for the cell. A triangular piece of metal is cut out which is large enough to slip over the three alignment posts of the cell and three holes are drilled into its corners, corresponding to the positions of the three posts. To pre-indent the gasket, it is slipped over the posts onto the lower diamond and the upper half of the cell is placed on top of it. All six screws are now put into the cell and slowly tightened while maintaining the relative thicknesses around the cell as measured in the previous step. It is critically important to turn each screw through a very small angle each time so as to minimize the amount by which the culets are put out of alignment. These six screws are size 4-48, so a twist of 30° corresponds to compressing one side of the cell by 44 μm. The relative thicknesses of each of the three measured points are checked often to insure that the culet faces remain parallel during the pre-indentation procedure. The purpose of pre-indenting the gasket is two-fold: first, pre-indenting the gasket provides a target for drilling the hole for the sample space hole and second, the gasket becomes work-hardened, enabling it to withstand higher pressures before failing.

After pre-indenting the gasket to the desired thickness (about twice the thickness of the sample), the screws are slowly released, again while maintaining the parallelness of the culets. A hole is drilled in the center of the indentation using a #78 tungsten carbide drill bit. This provides a space which is 400 μm in diameter to contain the 300 μm diameter sample. The hole must initially be larger than the sample as it will decrease in size during the loading process. After deburring the hole using a steel needle, the thickness of the gasket can be ascertained and, if not yet thin enough, further pre-indentation can be

performed and the hole re-drilled. When pre-indentation is completed, the alignment of the diamonds is rechecked, as some slippage may have occurred.

To load the sample into the cell, the gasket is put onto the bottom half of the cell. The sample, along with a few chips of ruby several microns in diameter, is placed inside the hole in the gasket and the top half of the cell is placed on top of the gasket. The six allen head screws are put into their places and screwed down until just before any force is required to turn them further. These screws are then adjusted so that the relative thickness measurements around the cell indicate that the culets are once again aligned. The cell is placed on its tightening stage, which is an aluminum plate with three posts enabling the screws to be tightened without rotating the entire cell, and the entire apparatus is immersed into a liquid nitrogen bath. After the nitrogen has ceased to boil rapidly, helium gas is bubbled into the liquid. This has the effect of virtually eliminating the boiling of the nitrogen, ensuring that no nitrogen gas will be trapped inside the sample space when the cell is tightened. If nitrogen bubbles were to become trapped within the sample space, the gasket hole will simply collapse inwards and crush the sample due to the much greater compressibility of a gas. Once the nitrogen is calm, the six screws can be tightened slowly as for the pre-indentation process an allen head screwdriver. A sudden increase in the force required to turn the screws will be experienced when the gasket seals. After giving all of the screws a few more turns to ensure that the sample space is truly sealed, the cell is removed from the liquid nitrogen. When it has returned to room temperature, the alignment of the cell is again checked and corrected if necessary.

2.2.7 Measuring the Pressure

In order for a diamond anvil cell to be truly useful, one must be able to measure the pressure inside the cell. The ruby fluorescence method for determining this pressure was developed by Forman (Forman 1972) and its invention at that time removed the greatest obstacle against the widespread use of these cells in high pressure research. In this

method, the shift of the R fluorescence lines of the Cr^{3+} ion in ruby is measured to determine the pressure. The lines observed are due to electronic transitions and may be excited by any light which is more energetic than the transition itself, such as the green light from an Ar ion laser or by a reasonably intense broadband source, such as a mercury arc lamp. At room temperature, two lines, R_1 and R_2, can be observed, arising from transitions from the ^2E to the $^4A_2$ crystal field split states. In principle, only one fluorescence line should be observed. However, since the Cr^{3+} ion replaces an Al atom in the lattice and is too big to sit substitutionally on the Al site, it occupies an interstitial spot of lower symmetry. This splits the ^2E state into an E and a ^2A state giving rise to the observed doublet (Ferraro 1984). The lower energy R_1 line is thermally activated and is not seen at low temperatures. The energy of these ruby fluorescence lines is temperature dependent so when exciting the ruby with a laser, one must take care to heat the ruby as little as possible.

The pressure dependence of the ruby lines has been measured for pressures of up to 1 Mbar and was calibrated by comparing the shift of the ruby line to the change in lattice parameter and the equations of state for Cu, Mo, Ag, and Pd (Mao 1978). Based on a fit to the experimental data, the proposed calibration curve is:

$$P\,(Mbar) = 3.808\left(\left(\frac{\Delta\lambda}{6942}+1\right)^5 - 1\right) \qquad (2.3)$$

where $\Delta\lambda$ is the shift of the ruby line in Å from its value at atmospheric pressure. The ruby calibration curve obtained by Mao along with a linear extrapolation are plotted in figure 2-8. The curve is essentially linear for pressures below 200 kbar.

A block diagram of the pressure measurement system used in this work is shown in figure 2-9. The beam from a 25 mW argon ion laser is used to excite the ruby fluorescence. The laser is reflected through the microscope onto the DAC located on the

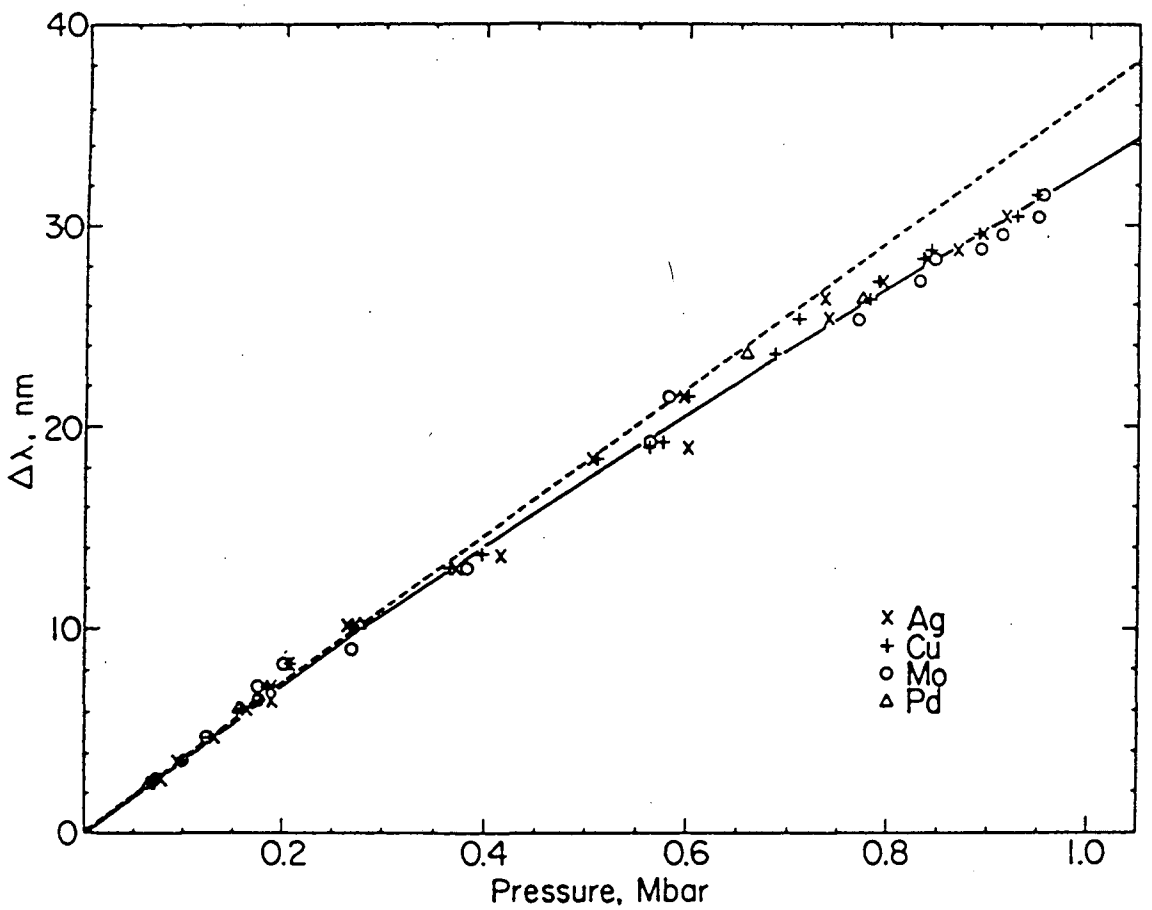


Figure 2-8. The ruby calibration curve of Mao (Mao 1978).

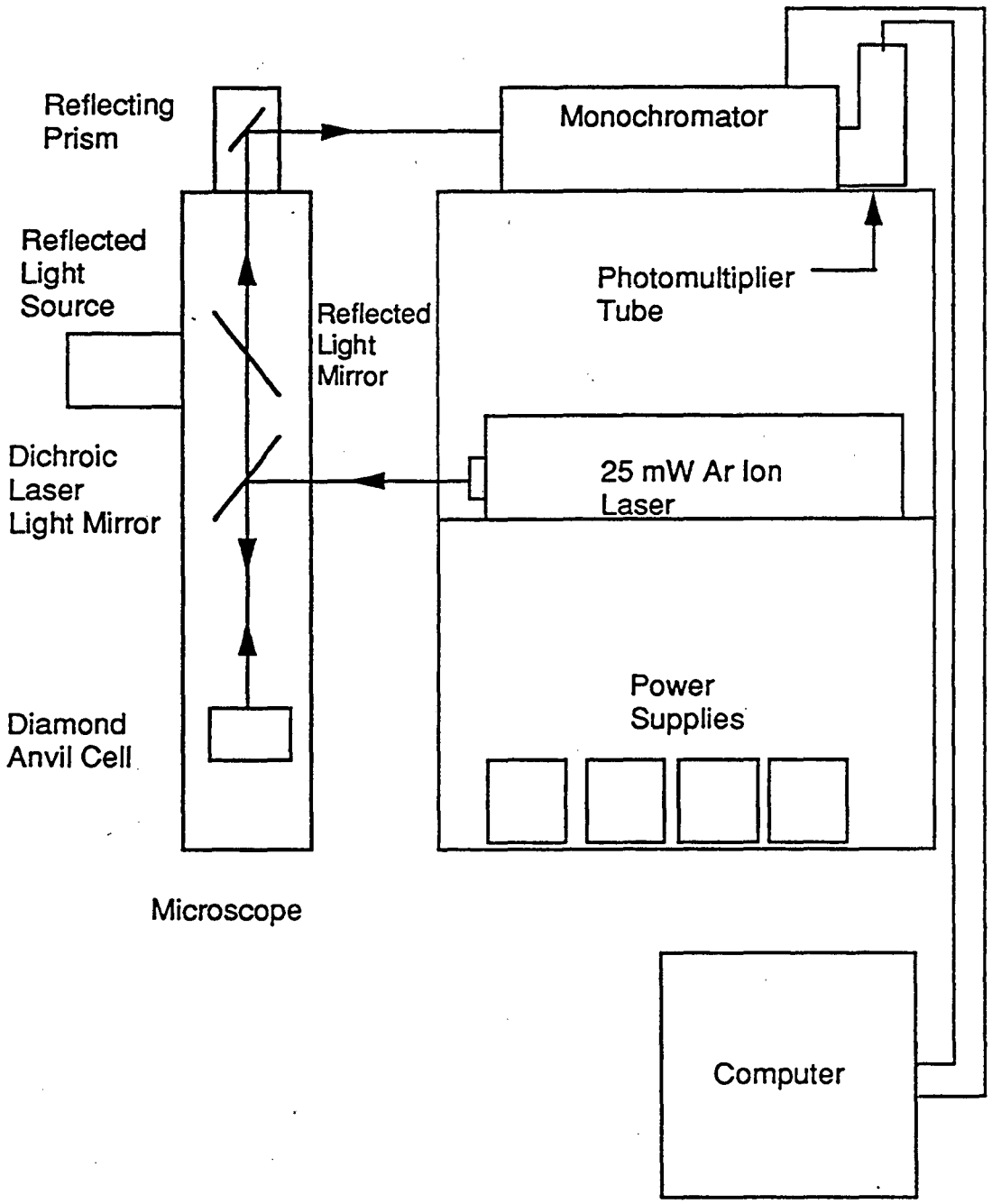


Figure 2-9. Diagram of the ruby fluorescence pressure measurement system (Wolk 1992b).

sample stage by means of a 45° dichroic mirror. The ruby fluorescence is collected by the objective lens and travels back up the microscope where it is directed by a prism into a computer controlled monochromator. A photomultiplier tube at the exit slit of the monochromator detects any light reflected by the diffraction grating and sends a signal through a pre-amplifier to an A/D board inside the computer. Under computer control, the monochromator can scan through a range of wavelengths to produce a spectrum of the ruby lines as shown in figure 2-10. By measuring the shift of the lines from those at room temperature, the pressure inside the cell can be determined. Various filters in the microscope prevent any reflected laser light from entering the monochromator, where it could potentially destroy the photomultiplier tube.

One drawback of this system is that although it is able to measure the pressure in the cell at room temperature, this measurement cannot be performed while the cell is in the cryostat at liquid helium temperature, when the pressure is potentially quite different from its room temperature value. However, using the infrared active vibrational modes of the N_2 pressure medium and of the CO_2 dissolved in the N_2, one is able to measure the *in situ* pressure in certain cases (McCluskey 1996). Although the vibrational mode of N_2 is not normally infrared active, a coupling between this mode and lower frequency phonons in solid nitrogen enables it to be observed in this case. The presence of CO_2 dissolved in the N_2 pressure medium is due to the cell being loaded in the open atmosphere.

Two different methods were used to determine the pressure dependence of the CO_2 and N_2 modes. The first involved comparison of the two with the shift of the C_{Sb} local vibrational mode (LVM) in AlSb, which was assumed to be linear in pressure, as is the case for the Si_{Ga} LVM in GaAs (Wolk 1991). In the second method, the cell was cooled to liquid helium temperatures and the frequencies of the N_2 and CO_2 vibrational modes were measured. The cell was then warmed back to room temperature and cooled again in a different cryostat in which photoluminescence was performed to determine the low temperature pressure using the ruby fluorescence calibration. This method involved the

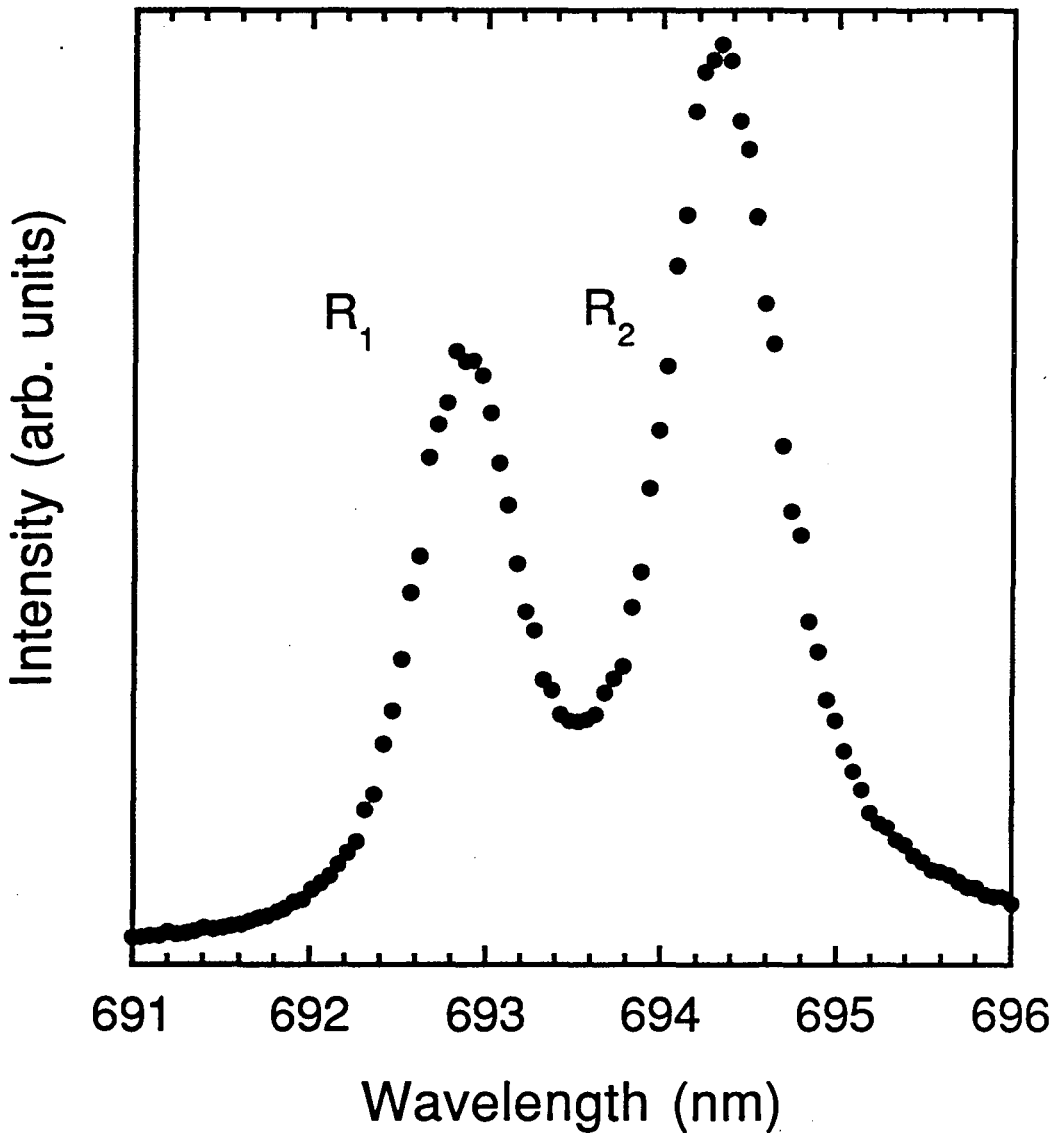


Figure 2-10. Sample ruby fluorescence spectrum at atmospheric pressure and room temperature. At higher pressures, the peaks shift towards longer wavelengths.

assumption that the cell would reach the same low temperature pressure on two successive temperature cyclings, an assumption which was later shown to be accurate by simply using the same measurement technique twice in a row.

The observed pressure dependence of the CO_2 and N_2 lines is plotted in figure 2-11. Good experimental agreement is found between the open symbols, which are derived from the first method of calibration and the filled symbols, which come from the second method. Over the pressure ranges measured, the pressure dependence of the vibrational modes is linear and can be fit to the following equations:

$$\nu(CO_2) = \begin{cases} 2349.3\ cm^{-1} + 12.3\ P, & 0 \le P \le 0.45\ GPa \\ 2347.6\ cm^{-1} + 8.4\ P, & 0.45 \le P \le 1.3\ GPa \\ 2349.5\ cm^{-1} + 6.9\ P, & 1.3 \le P \le 1.9\ GPa \\ 2345.1\ cm^{-1} + 6.6\ P, & 1.9 \le P \le 7.0\ GPa \end{cases} \qquad (2.4)$$

$$\nu(N_2) = \quad 2326.5\ cm^{-1} + 2.7\ P, \qquad 1.9 \le P \le 7.0\ GPa \qquad (2.5)$$

Since nitrogen was used as the pressure medium for all of the work described here, this *in situ* pressure calibration could in principle be used all the time. However, oftentimes, the spectral region containing these lines was not accessible due to beamsplitter limitations. As a partial solution, many spectra were taken in which both the low temperature and room temperature pressures were measured, so that a calibration of the low temperature pressure based on the room temperature pressure could be made. This data is shown in figure 2-12. A linear fit to the data points yields

$$P(LT) = -11.08 + 1.06\ P(RT)\ \ kbar \qquad (2.6)$$

where P(RT) is the room temperature pressure and P(LT) is the low temperature pressure in kbar. This calibration was used to estimate the low temperature pressure for all data in which the *in situ* calibration was not available.

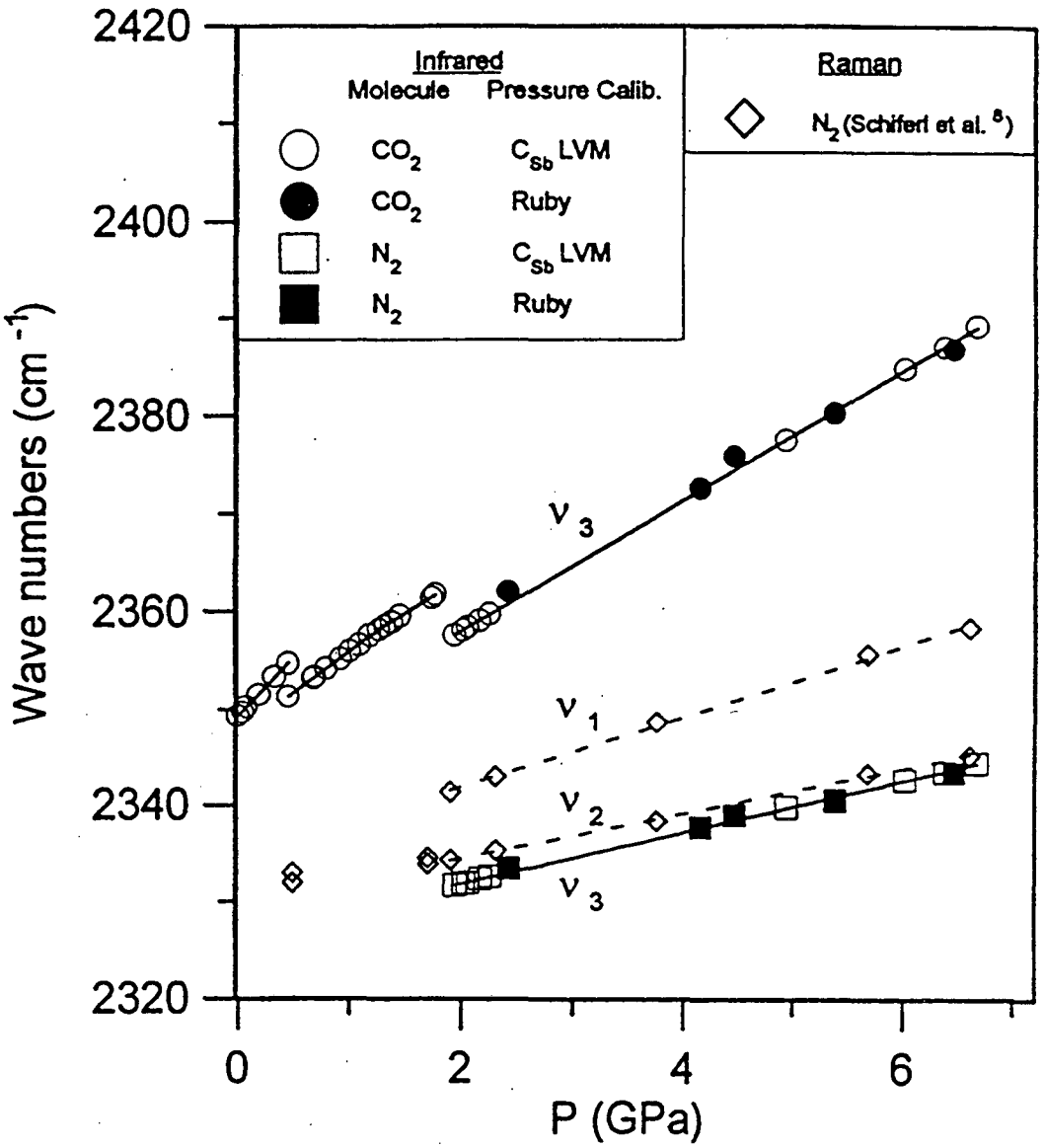


Figure 2-11. Pressure dependence of the CO_2 and N_2 infrared absorption lines in nitrogen at liquid helium temperatures. The position of the AlSb:$^{12}C_{Sb}$ local vibrational mode was used to calibrate the pressure for the open circles and squares, while the ruby fluorescence method was used for the filled shapes. Open diamonds are Raman active modes as measured by Schiferl et al. (Schiferl 1985). The two discontinuities at 4.5 and 19 kbar are due to phase transitions of the solid nitrogen medium. (McCluskey 1996)

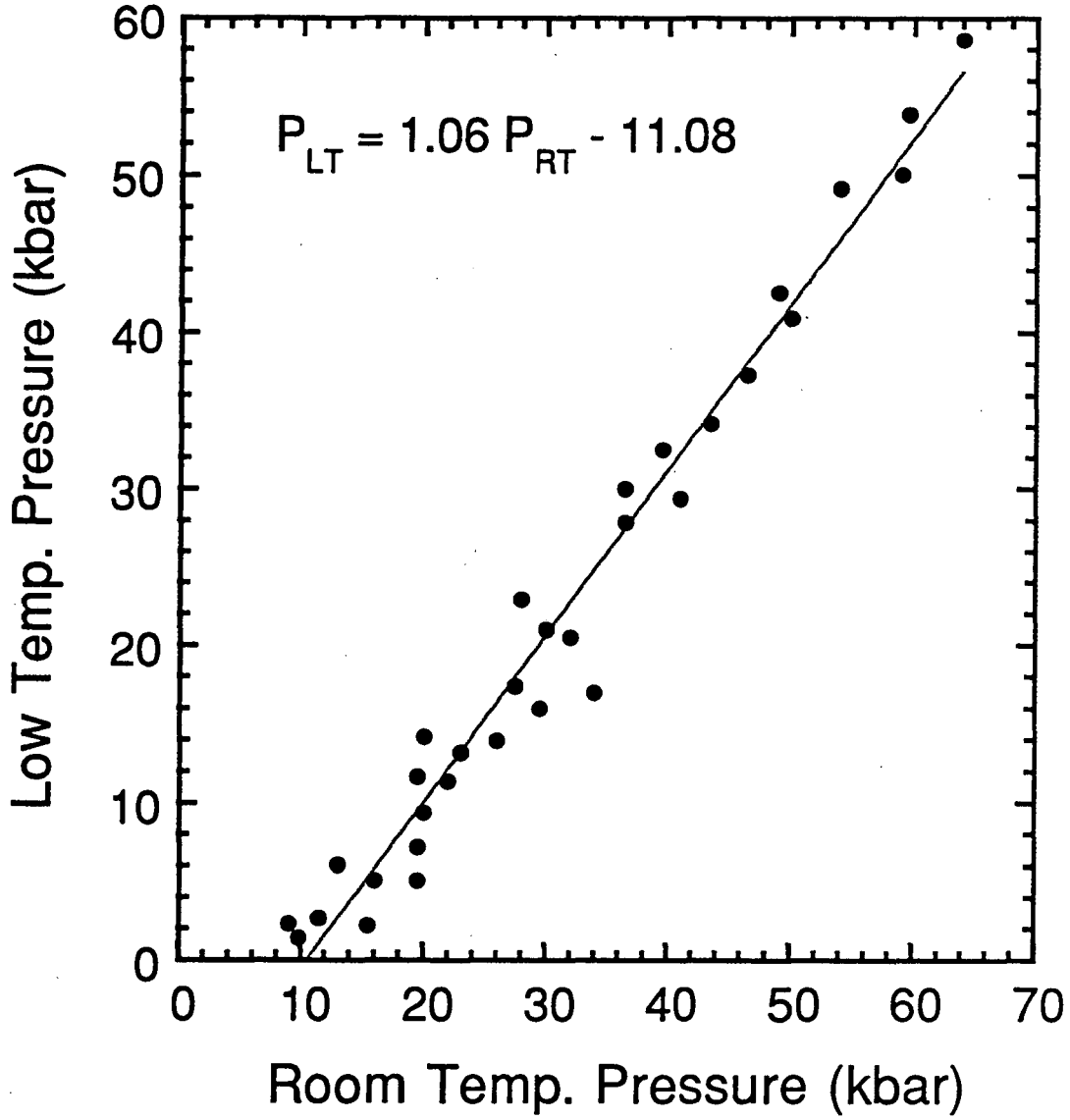


Figure 2-12. Relationship between room temperature and low temperature pressures in the diamond anvil cells used in this work. The validity of this calibration breaks down below roughly 15 kbar.

2.2.8 Optical spectroscopy through a Diamond Anvil Cell

The spectroscopy of impurities through a diamond anvil cell presents certain challenges due to the small thickness of the sample. While it is common to perform studies of the intrinsic properties of materials in a DAC, where the number of atoms contributing to the effect is on the order of 10^{22} cm^{-3}, spectroscopy of impurities, where concentrations may be as low as 10^{15} cm^{-3} requires much greater sensitivity. The root of the problem is getting a large enough photon flux through the sample hole in order to reliably measure the absorption coefficient. For this reason, the vast majority of previous impurity studies using high pressure have been performed using the photoluminescence techinique or an infrared laser as the light source for absorption measurements. However, neither of these two methods is suitable for studying the excited state spectrum of impurities.

A solution to this problem was found by Jeff Wolk (Wolk 1991), a former graduate student in this group. The assembly shown in figure 2-13 was designed for the purpose of transmitting as much light as possible through the diamond anvil cell. The assembly consists of 5 pieces and is hung vertically inside of a cryostat, enabling it to be cooled to liquid helium temperatures. Modulated light from the spectrometer enters horizontally at the bottom, where it is reflected upwards by the 45° mirror (5). The cone (4) serves to concentrate the light and direct it through the diamond anvil cell (3), which has its own small light concentrating cones built into the backing plates. Finally, the detector (1), of a type which was described in section 2.1, is mounted inside an integrating optical cavity and is located directly behind the cell, as close to the diamond as possible. All surfaces on which light might be incident were gold plated to increase their reflectivity and the amount of light being directed through the DAC.

This assembly performs several functions. First, by locating the detector directly behind the cell, as much of the light coming through the cell as possible is collected. Second, by restricting the path through which light can enter the cell and reach the detector, virtually all stray radiation, i.e., light which is not being modulated by the interferometer, is

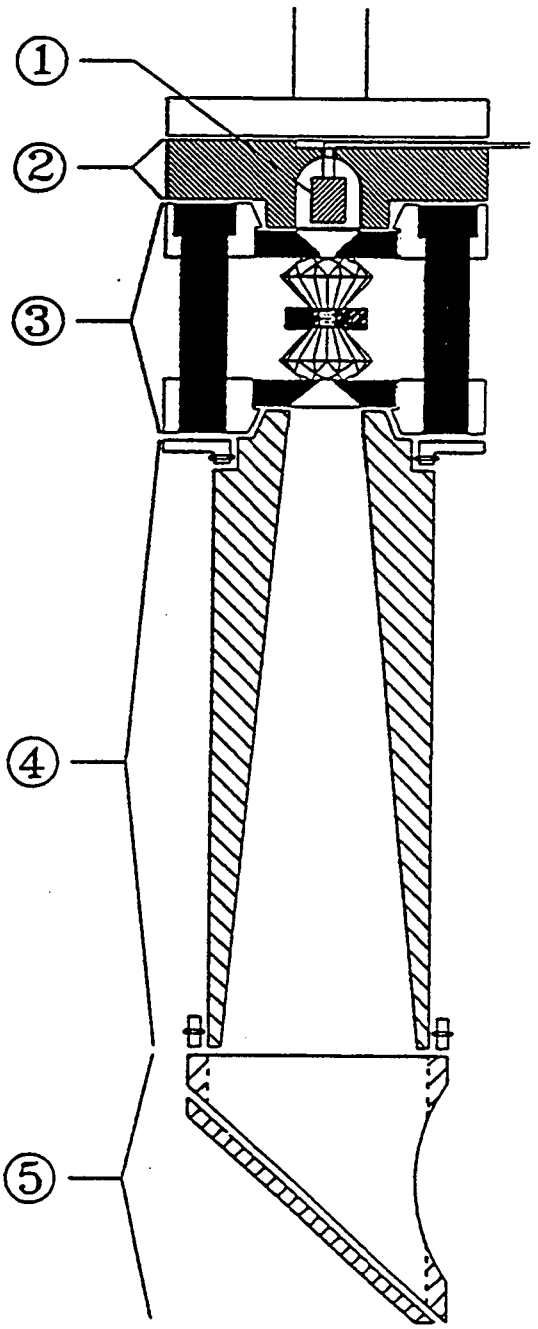


Figure 2-13. Schematic diagram of the assembly used to perform impurity absorption spectroscopy in a diamond anvil cell. The parts are: (1) doped Ge photoconductor, (2) photoconductor housing, (3) diamond anvil cell, (4) light concentrating cone, and (5) 45° mirror (Wolk 1992b).

eliminated. The latter effect is very important, as extra non-modulated photons contribute only to the detector noise and could overwhelm the small number of modulated photons reaching the cell. Finally, in order to maximize the amount of spectrometer light entering the cell, it is only necessary to align the entrance to the mirror assembly, which is 12.7 mm in diameter, with the spectrometer beam, which is 10 mm in diameter, a relatively simple task. Without such a cone, one would have to align the 0.4 mm hole through the DAC with the center of the 10 mm spectrometer beam in order to maximize the signal, a difficult task indeed.

A further difficulty with performing far infrared impurity absorption spectroscopy through a diamond anvil cell is the problem of obtaining a good reference sample. Ideally, a reference sample should be identical to the sample containing the impurities of interest in all respects except for the presence of the impurities. Unfortunately, with diamond anvil cells, obtaining a reference spectrum is not as easy as loading two cells, one containing a reference and one containing the sample. Due to variations in the point at which the gasket actually seals during the loading, no two loads are ever identical. The size of the gasket hole will vary from loading to loading. The distance between the diamonds and perhaps the diamonds themselves will be slightly different. The sample and reference may be of slightly different dimensions due to the violent sample preparation method (ultrasonic cutting). The amount of light leakage around the sample may be different. Each effect is in itself very small, but since the signals that we are interested in are also very small, it is easy for them to become lost among the unwanted features. Fortunately, due to the existence of a deep impurity state known as a DX center in many III-V semiconductors, we are able to use the same piece of material as both reference and sample.

The DX center is an alternate configuration which shallow extrinsic impurities can assume in many III-V semiconductors under certain conditions (Lang 1977, Wolk 1992, Hubík 1993, Becla 1995). For example, upon the application of hydrostatic pressure in excess of 20 kbar, some species of shallow donors in GaAs are known to transform into

DX centers, which are deep states within the band gap at these pressures. In the case of group IV impurities such as Si, this transformation involves the breaking of a bond from the impurity atom to a neighboring atom and a lattice relaxation of the impurity atom into the plane of the remaining three neighboring atoms (Chadi 1991). For DX centers involving group VI impurities such as S, it is one of the host lattice atoms which moves. One of the most striking characteristics of DX centers, and the property which was exploited, is the persistent photoionization they show at low temperatures (Lang 1977). If the DX centers in a crystal are exposed to photons of sufficient energy, they can be transformed back into the normal shallow donor configuration. At low temperatures (< 100 K), an energy barrier prevents their return to the DX state. A configuration coordinate diagram of this process can be seen in figure 2-14.

Since DX centers are deep states which are spectroscopically inactive in the energy range with which we are concerned (except possibly for their local vibrational mode, the position of which is well known (Wolk 1991)), each sample could be used as its own reference. To obtain an absorption spectrum, a sample was cooled in the dark at pressures greater than 20 kbar, allowing the impurities to assume their DX configuration, and a reference spectrum was taken. The light incident on the sample during this measurement did not possess enough energy to perturb the DX centers. Light of much higher energy (about 1 eV) was then shined on the sample using either an AlGaAs diode, an incandescent light bulb, or a quartz-halogen bulb, converting the DX centers back into the normal shallow donor configuration, and a second spectrum was taken. Taking a ratio of the post- and pre-illumination spectra produced the absorbance spectra shown in the next sections.

In order to improve the quality of the resulting spectra, the source of high energy light could be left on during the collection of the second spectrum. Since each of the three light sources was capable of generating light of energy greater than the GaAs band gap, this light could, in addition to ionizing the DX centers, also create electrons and holes within the sample which are captured by ionized donors and acceptors. The impurities then become

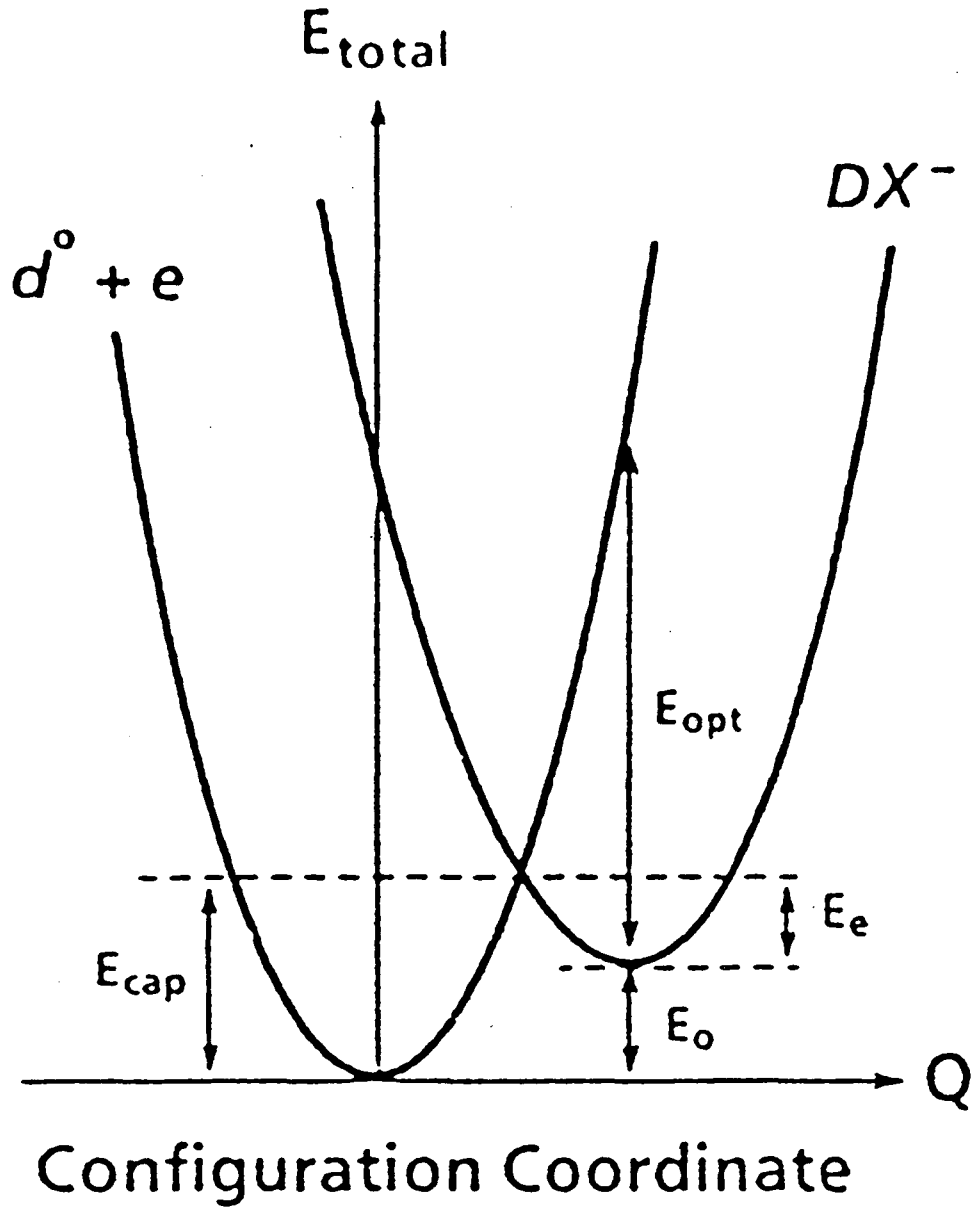


Figure 2-14. Configuration coordinate diagram for DX centers. The energies are: E_{cap}, energy of the barrier from the shallow state to the DX state, E_{opt}, optical ionization energy for transformation from DX to shallow state, and E_e, thermal ionization energy for transformation from DX to shallow state.

neutral, reducing the random electric fields present in the material which contribute to the broadening of the spectral lines through the Stark effect (Armistead 1984). This high energy light was prevented from reaching the detector by a piece of black polypropylene placed between the sample and the detector.

3. Camel's Back Donors in III-V Semiconductors

3.1 X-band donors in GaAs

As stated in the first section, the main motivation for this work was to discover the reason behind the lack of absorption peaks in the Si donor spectrum observed in indirect gap GaAs. To determine whether this behavior is peculiar to Si or characteristic of X-band donors in GaAs in general, the first step was to measure the absorption spectra of other impurities known to form shallow donors in this material. GaAs samples doped with Sn, Ge, S, Se, and Te were obtained and absorption spectroscopy, as discussed in the previous section, was performed. The samples are described below, followed by a brief review of the results obtained for each donor and then finally a discussion of the results.

3.1.1 Samples

The optimum range of the impurity concentration in samples for infrared absorption spectroscopy is limited by two factors. On one hand, the concentration cannot be too high, or else donor wavefunction overlap will broaden the peaks into indistinguishability (Larsen 1976). On the other hand, the concentration cannot be too low or else the absorption signal may be too weak to detect. Unfortunately, the vast majority of GaAs grown today is either very highly doped for use in the manufacture of electronic devices, or very pure for use in x-ray detectors (Sumner 1994). Neither of these is suitable for our experiments and the relative scarcity of appropriate samples was the biggest limiting factor in our studies of GaAs.

Two types of samples were used in this investigation. The first was bulk GaAs, grown either by the Bridgman or the liquid encapsulated Czochralski technique. The net doping concentration of these samples ranged from 4×10^{15} cm^{-3} to 2.2×10^{18} cm^{-3}. Hall effect measurements on the more resistive samples showed a high degree of compensation (as high as 0.8 or 0.9, as is usual for resistive bulk GaAs). However, as noted in the

previous section, band edge light could be used to reduce the Stark broadening of absorption lines due to the compensating charged impurities.

The second type of samples used were epitaxial layers grown by the liquid phase epitaxy (LPE) method. Most of these layers were grown in Elizabeth Bauser's group at the Max-Planck-Institute in Stuttgart and had net carrier concentrations between 10^{15} cm^{-3} and 3 x 10^{16} cm^{-3} with very low compensation ratios. Other epitaxial layers were grown by Dawnelle Wynne at the Lawrence Berkeley National Laboratory and had carrier concentrations around 10^{16} cm^{-3} and compensation ratios of about 0.4.

Overall, it was possible to obtain samples doped with six different dopants, Si, Sn, Ge, S, Se, and Te, though some could only be obtained with one doping concentration.

3.1.2 Group IV dopants: Si, Sn, Ge

Attempts to obtain absorption spectra were the most successful with impurities from group IV of the periodic table and the most extensively studied dopant was Si (Hsu 1997a). In figure 3-1, absorbance spectra from three different Si samples with different doping concentrations are shown. Spectrum (a) was taken using a bulk-grown sample with 10^{17} cm^{-3} impurities, spectrum (b) from a bulk-grown sample doped to 10^{16} cm^{-3}, and (c) from an LPE sample doped with 3.5 x 10^{15} cm^{-3} impurities. The spectra are not shown to scale. Because of light leakage around the sample in the diamond anvil cell, it is impossible to make an accurate determination of the absorption coefficient α of these lines. However, it appears to be roughly equal to that of shallow donors in other materials (~1 cm^{-1} /10^{15} cm^{-3} impurities) (Jagannath 1981, Scott 1976). The absorbance spectrum of the most lightly doped sample shows a single peak with a full width at half maximum (FWHM) of 8 cm^{-1}. This same peak can also be seen in the spectra of the more heavily doped samples, though broadened due to the higher donor concentrations. In spectra (a) and (b), the absorbance is greater on the high energy side of the peak, possibly indicative of a continuum absorption.

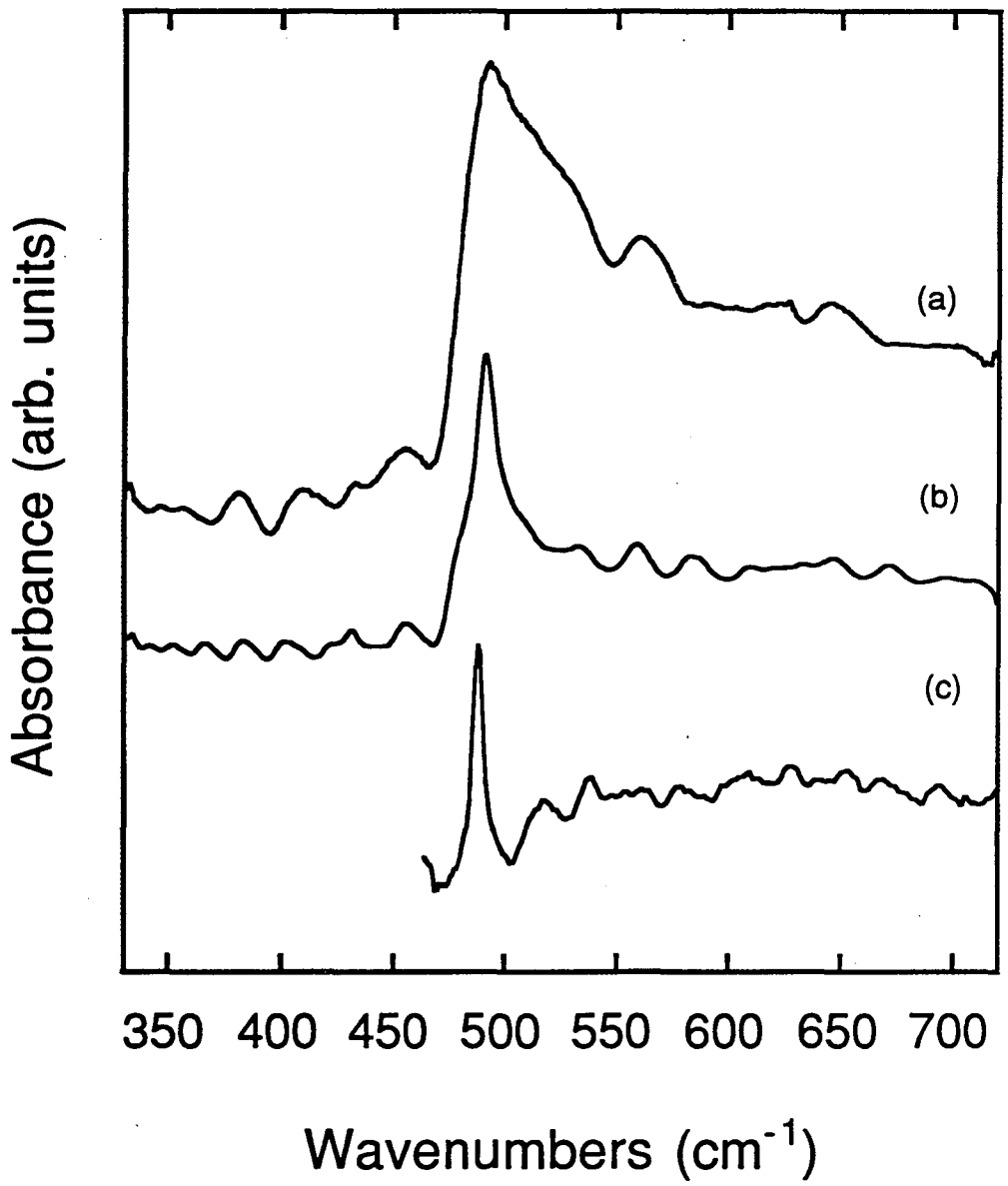


Figure 3-1. Absorption spectra of GaAs:Si for three different doping concentrations: (a) [n] = 10^{17} cm^{-3} at 51 kbar, (b) [n] = 10^{16} cm^{-3} at 50 kbar, (c) [n] = 3.5 x 10^{15} cm^{-3} at 75 kbar.

As the pressure was increased, the energy of this single absorption line shifted downwards, indicating that the energy difference between the donor ground state and that of the corresponding bound excited state was decreasing. This is opposite to the behavior of the normal Γ-band donors in GaAs, whose transition lines shift to higher energy as a function of pressure (Wasilewski 1986). A plot of the peak position as a function of pressure is shown in figures 3-2a and 3-2b. In figure 3-2a, data obtained from the more lightly doped samples is shown while peak positions of the more heavily doped samples are plotted in 3-2b. There is a fair amount of scatter in the data, much of which arises from the difficulty in determining the peak positions of the very broad peaks observed with the most highly doped samples. In addition, the pressures at which the data points shown as open symbols were taken were estimated from equation (2.6) and are thus uncertain to roughly ±3 kbar. In contrast, the frequency of the CO_2 vibrational mode was used to determine the exact pressures at which the data points shown as closed symbols were taken and the uncertainty in pressure for those is smaller than the symbols.

The fit in figure 3-2a was determined from the filled symbols which are the most reliable, as they were measured using the highest quality sample. The pressure coefficient of this transition line was determined to be –0.40 cm^{-1}/kbar. As can be seen from figure 3-2b, the absorption peak occurs at slightly lower energies in the more heavily doped samples (10^{16} to 10^{17} cm^{-3}), but the pressure dependence of the peak is essentially the same to within the experimental uncertainty. This effect has also been seen in the Γ-band donors at atmospheric pressure (Lee 1988).

The second dopant to be studied was Sn. Figure 3-3 shows spectra from two Sn doped samples, one with an impurity concentration of 1.2 x 10^{15} cm^{-3} and one doped to 2.3 x 10^{16} cm^{-3}. Both samples in this case were LPE grown. The absorption coefficient of these peaks was about the same in magnitude as that for Si and the peak from the more lightly doped sample has a FWHM of 4.7 cm^{-1}. Figure 3-4 shows the pressure

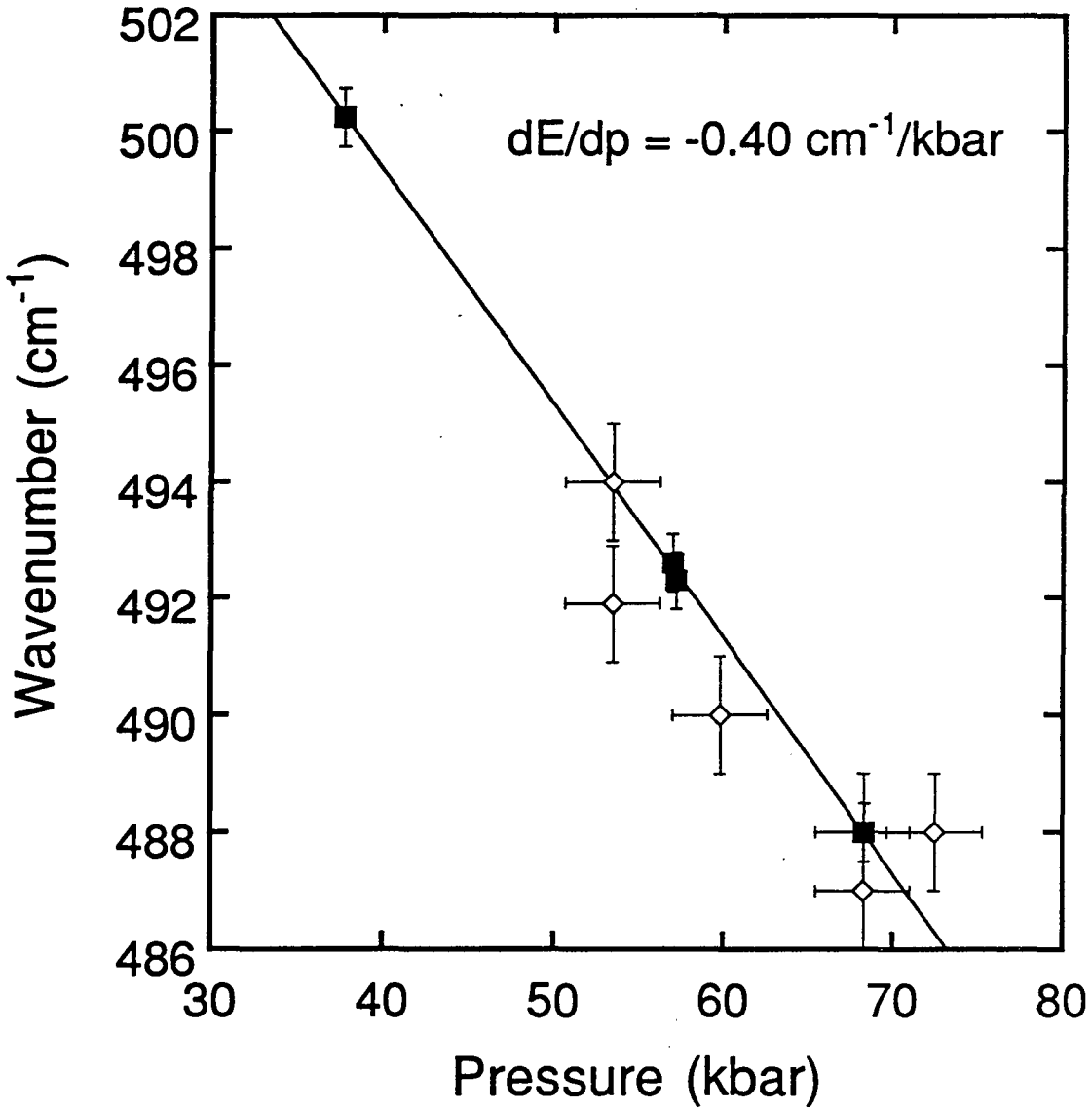


Figure 3-2a. Pressure dependence of the Si peak. Low temperature pressures for the open diamond data points were estimated using eq. (2.6) while pressures for filled symbols were determined by the *in situ* method. Diamonds represent bulk samples with [n] = 4 x 10^{15} cm^{-3}. Filled squares are LPE grown layers with [n] = 3.5 x 10^{15} cm^{-3}. Error bars show pressure and frequency uncertainties in the measurement. Pressure uncertainties of the filled symbols are smaller than the symbols. The line is a fit to the solid points.

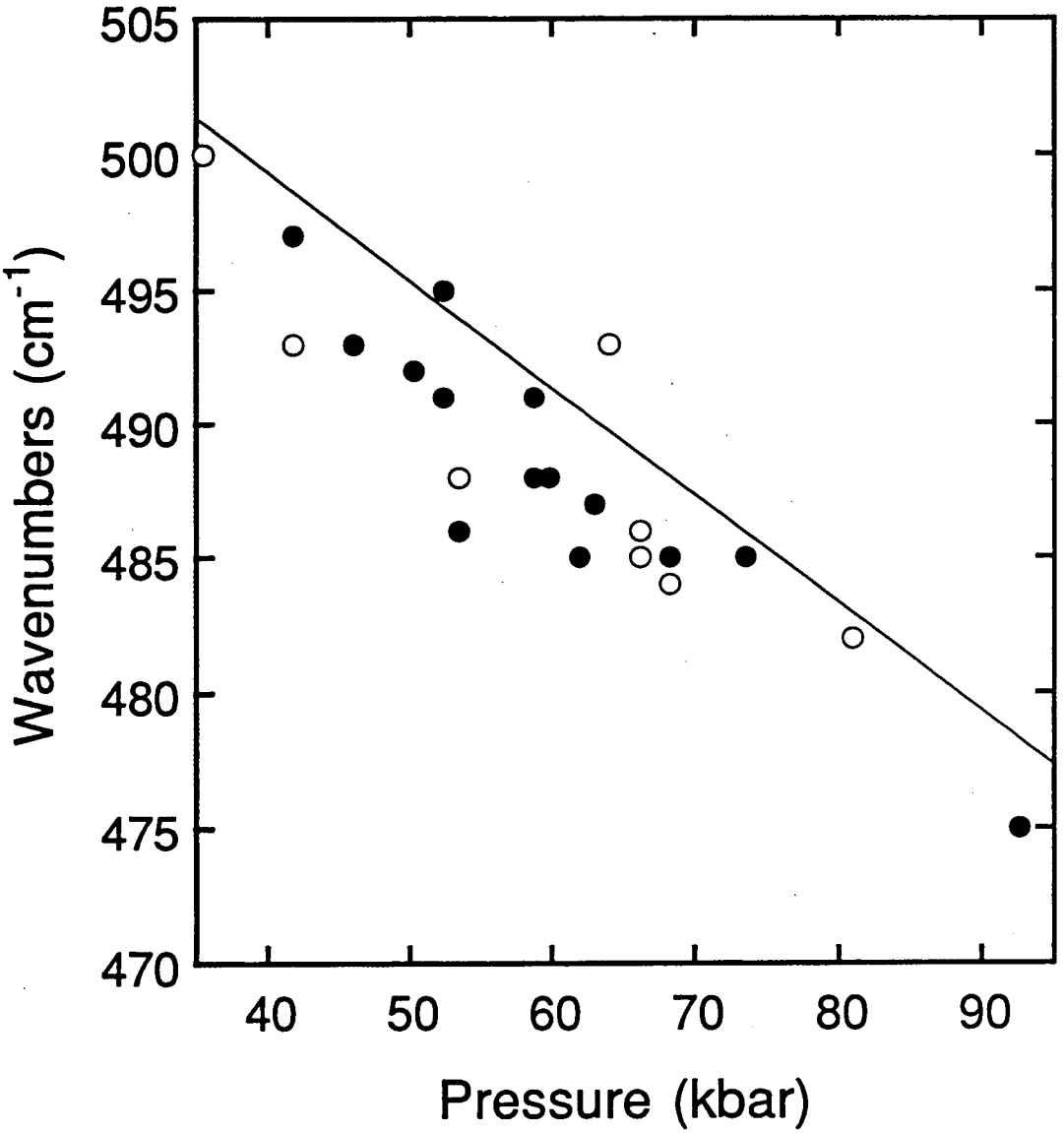


Figure 3-2b. Pressure dependence of the Si peak in heavily doped samples. The filled circles are from bulk samples with [n] = 10^{16} cm^{-3} and open circles are from bulk samples doped to 10^{17} cm^{-3}. The line is the fit from figure 3-2a, showing the decrease in the energy of the absorption peak with greater doping concentration. All pressures were estimated from equation (2.6) and have errors of ±2.8 kbar. The uncertainty in the wavenumber position is roughly 5 cm^{-1}.

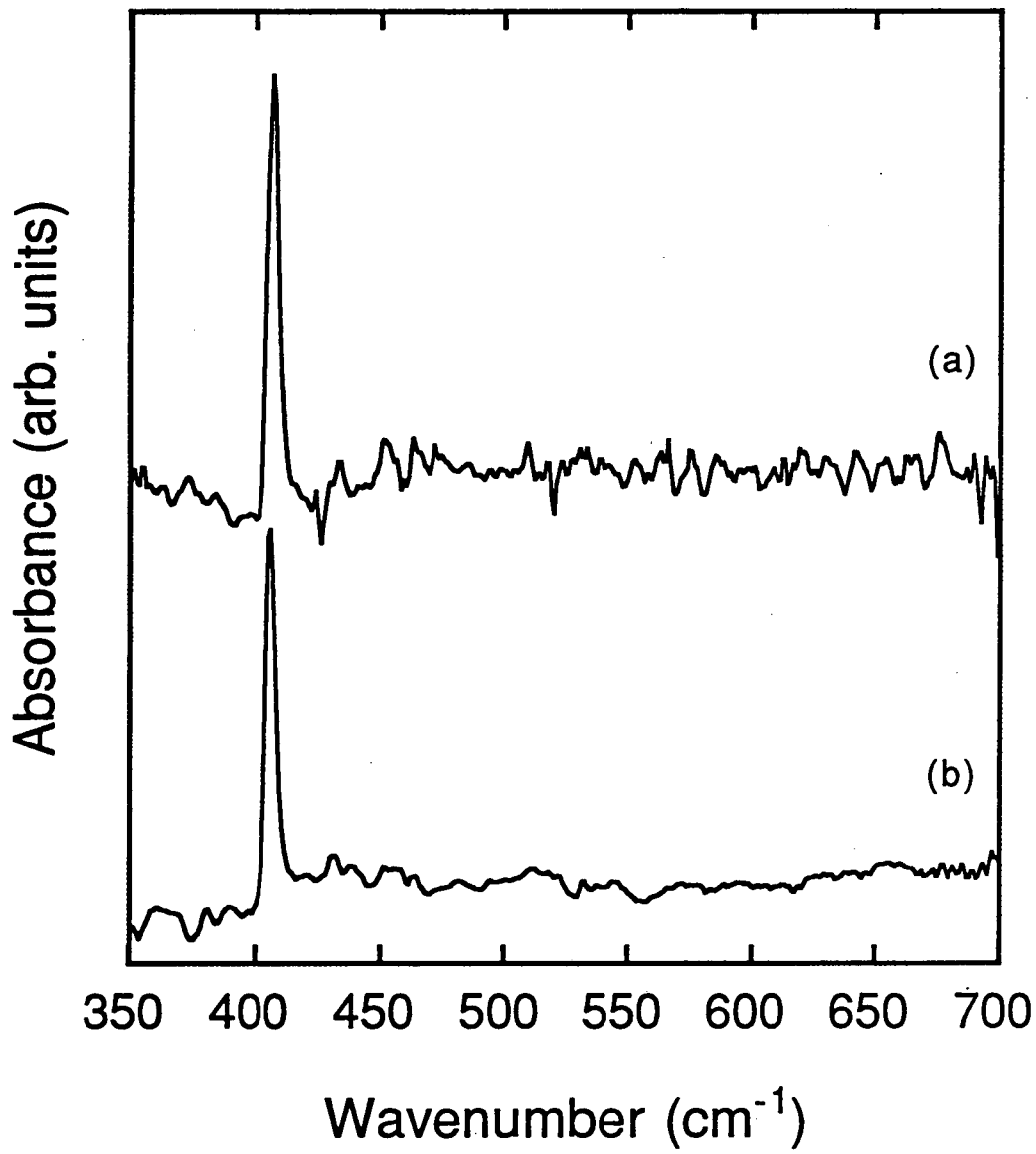


Figure 3-3. Absorption spectra of GaAs:Sn at two doping concentrations: (a) [n] = 2.3 x 10^{16} cm^{-3} at 64 kbar and (b) [n] = 1.2 x 10^{15} cm^{-3} at 50 kbar.

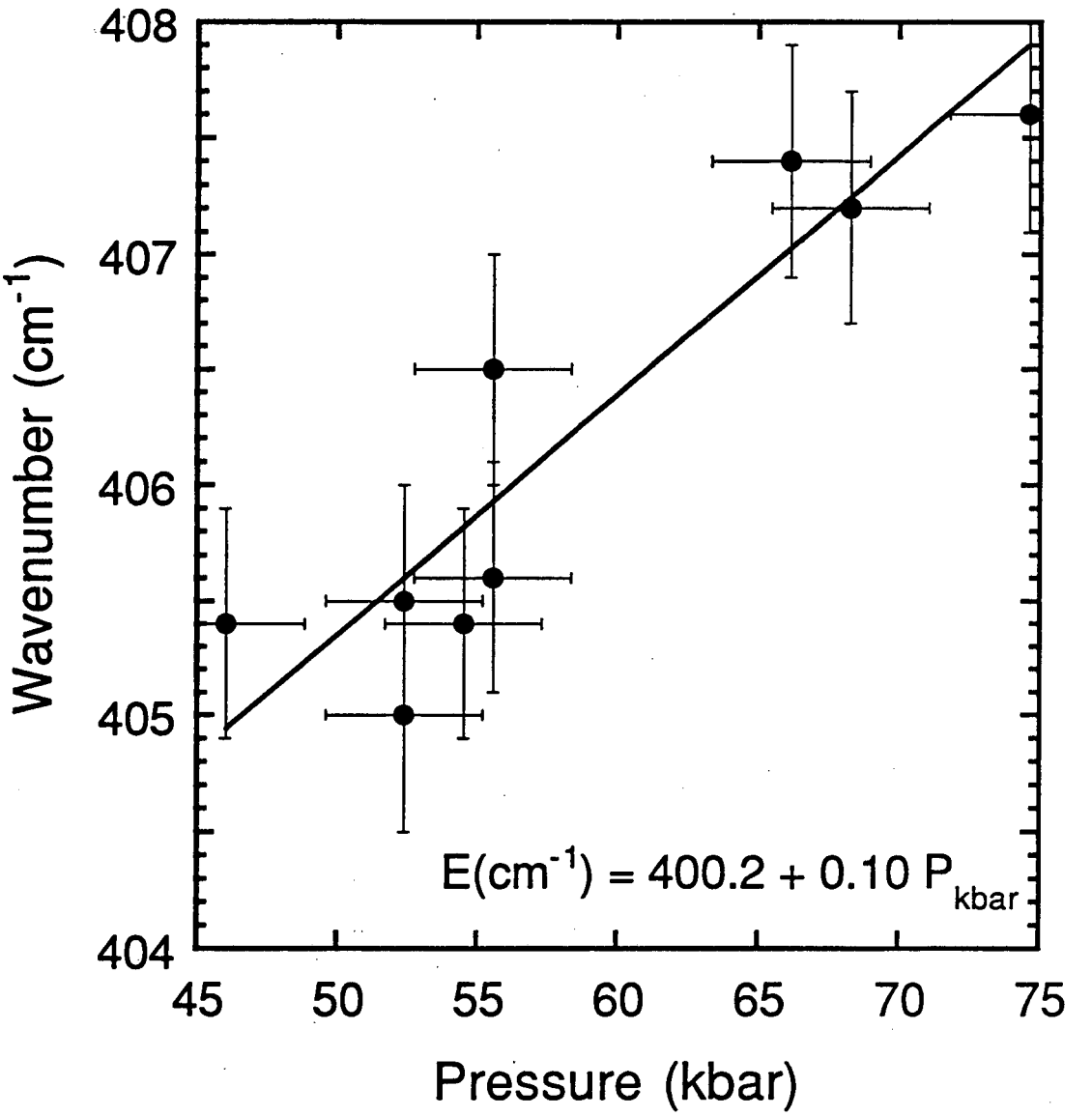


Figure 3-4. Pressure dependence of the Sn donor absorption peak in indirect GaAs. The pressure for each data point was calculated using equation (2.6). All samples have a doping concentration of 1.2×10^{15} cm^{-3}.

dependence of the Sn absorption line. Due to the energy range in which this peak is found, the low temperature pressures at which these spectra were taken could not be measured directly and had to be estimated using equation (2.6), again resulting in pressure uncertainites of roughly ±3 kbar. The pressure dependence of this line is estimated to be +0.1 cm^{-1}/kbar. It is not altogether surprising that the pressure coefficients of Si and Sn have opposite signs as the binding energies of the normal Γ-band donors in GaAs have been found to have different derivatives with respect to magnetic field and pressure. These differences have been attributed to the central cell potentials of the various impurity species (Wasilewski 1986).

Ge, the final group IV donor studied, is amphoteric and crystals doped with Ge during growth are usually highly self-compensated. This indicates that the Ga substitutional sites and the As substitutional sites have approximately the same affinity for Ge impurities. As large compensation ratios are detrimental to obtaining unbroadened donor absorption spectra, the Ge-doped samples used in this study were doped by neutron transmutation doping (NTD) (Alexiev 1993). A more detailed discussion of this technique can be found in Appendix B. An ultra-pure GaAs epitaxial layer ($n \sim 10^{12}$ cm^{-3}) was exposed to a thermal neutron flux at the research reactor at the University of Missouri. The neutron dose was calculated to produce a Ge_{Ga} concentration of roughly 4×10^{14} cm^{-3}. This figure was chosen for two reasons: first, such an impurity concentration should be easily detectable by our system (based on the strength of the Si and Sn absorptions for an impurity concentration of 10^{15} cm^{-3}) and second, this concentration should result in relatively sharp absorption lines, as most broadening mechanisms are absent at low donor concentrations. However, no Ge donor transition lines were ever observed in the entire spectral region which could be studied from pressures of 40 kbar to 70 kbar. Due to the technique used to introduce the Ge impurities, a slightly larger number of Se impurities

were produced as well. However, their presence should not have affected or obscured the Ge absorptions. The implications of this null result are discussed later.

3.1.3 Group VI dopants: S, Se, Te

The S doped GaAs samples which were investigated were LPE layers grown at LBNL by Dawnelle Wynne. These layers had carrier concentrations in the low 10^{16} cm^{-3} and compensation ratios of about 0.4. A few LPE layers were also grown by Elizabeth Bauser's group with doping concentrations from 10^{14} cm^{-3} to 10^{15} cm^{-3}. A bulk sample with a doping concentration of 2.2×10^{18} cm^{-3} was also studied.

In contrast to the results obtained for Si and Sn, no sharp transition lines for S could be found. Instead, only a broad continuum response was seen (Hsu 1997b). Some representative spectra at different pressures are shown in figure 3-5. Furthermore, the pressure dependence of the S donor absorption differs considerably from that of Si and Sn. In the Si and Sn doped samples, the energy of the infrared absorption lines varied little with pressure, changing at the rate of less than one wavenumber per kbar. The S absorption, however, shows a much greater sensitivity to pressure, as can be seen in figure 3-6. Here, the frequency at which the absorption reaches one half of its maximum value is plotted as a function of the applied pressure, showing a pressure derivative of -5.2 cm^{-1}/kbar. Also shown in this figure are data points taken from the very heavily doped (>10^{18} cm^{-3}) sample, which show a similar pressure derivative.

Se doped GaAs samples were obtained in two ways. One GaAs LPE layer was grown at the MPI in Stuttgart by Bauser with a doping concentration of 2.3×10^{16} cm^{-3}. A second layer was obtained through NTD (Alexiev 1993). During the neutron transmutation doping of GaAs, both Ge and Se donors are created. Therefore, the sample previously described, besides being doped with Ge, also contained roughly 7×10^{14} cm^{-3} Se impurities. However, no Se related absorptions were ever detected in either sample.

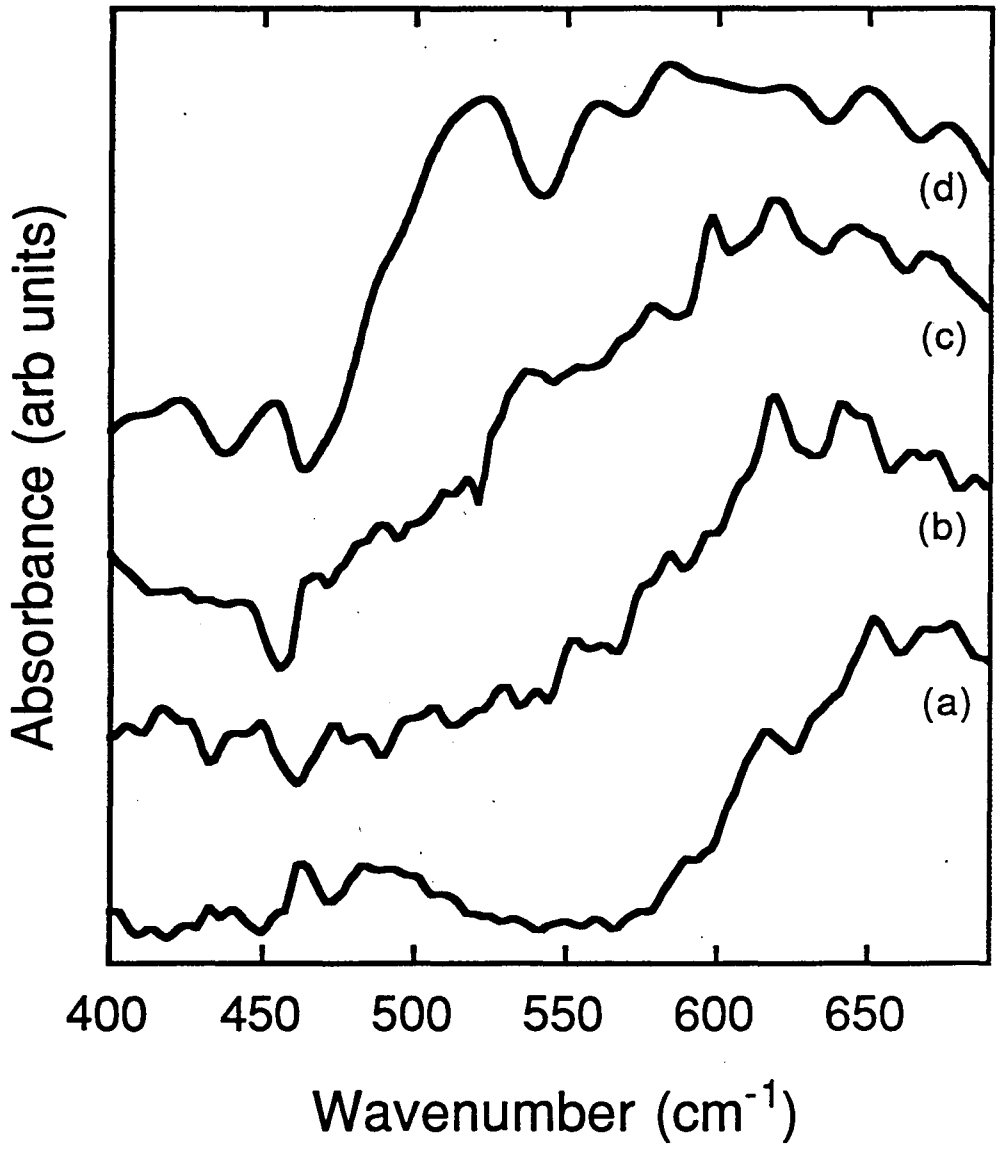


Figure 3-5. Absorption spectra for GaAs:S taken at four different pressures: (a) 52 kbar, (b) 59 kbar, (c) 68 kbar), (d) 76 kbar. The net doping concentration in this sample is 2.5 x 10^{16} cm^{-3}.

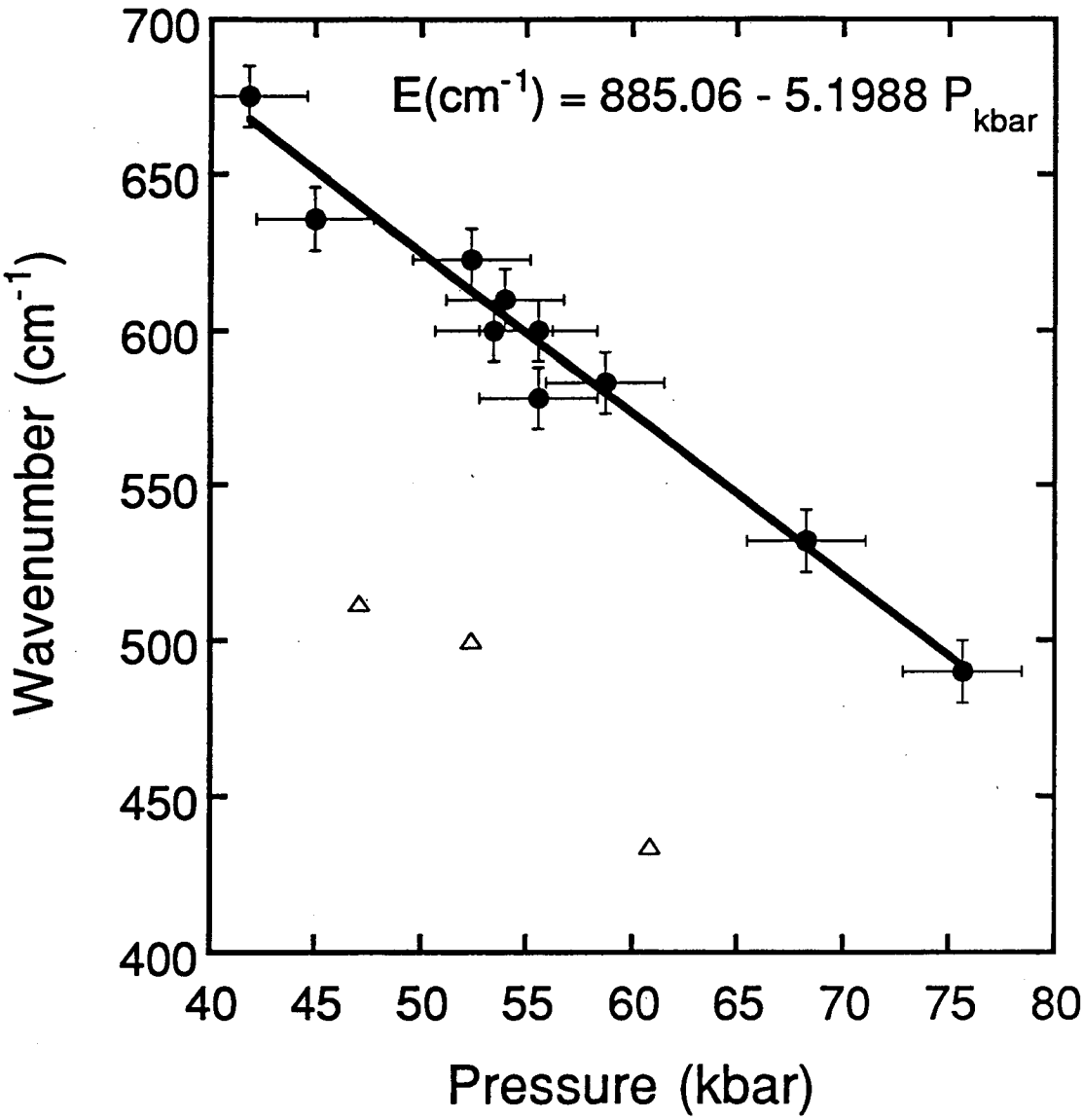


Figure 3-6. Plot of the frequency at which the continuum absorption observed in GaAs:S reaches one half of its maximum value as a function of pressure. A linear fit to the solid data points is also shown. The circles are from samples doped with 2 x 10^{16} cm^{-3} S impurities. The triangles are from a sample with a doping concentratrion of 10^{18} cm^{-3}.

Lastly, a single piece of bulk grown GaAs:Te was obtained. The net doping concentration of this sample was quite high at 1.2 x 10^{17} cm^{-3}. However, as with the Se doped GaAs, no reproducible donor related absorptions could be detected in these samples.

3.1.4 Discussion

As can be seen, a variety of different types of spectra were obtained for X-band donors in GaAs. Besides the fact that Si and Sn have similar spectra and Ge, Se, and Te showed neither sharp nor continuum absorptions, there seems to be little commonality among the behaviors of the six shallow donors. Furthermore, only the spectra of Si and Sn bear any resemblance to that of a normal shallow donor such as is found in Si or Ge. With regard to the lack of response in the latter three however, it should be noted that due to the experimental technique used, two conditions must be met in order for a donor related absorption to be observed: (1) the impurity must initially be in the deep DX configuration and (2) one must be able to convert these DX centers into the shallow donor state by optical excitation. In principle, if condition (1) were not fulfilled, one should still be able to see any sharp donor absorptions as dips in the detector response spectrum. However, due to the extremely small thickness of the samples, no such features that could be positively identified were found.

While the vast quantity of literature regarding DX centers may seem to suggest that all donors in GaAs undergo the shallow-deep transformation above a certain pressure, a careful search of the literature reveals that this is not the case, particularly for Ge, Se and Te.

For example, although it is well established that Te forms a DX center in AlGaAs:Te for an Al fraction greater than 25% (Shan 1989), electrical experiments on GaAs:Te at low pressures (up to 20 kbar) show that the Te DX state lies well above the conduction band (Sallese 1990, Suski 1991) and its pressure coefficient is such that it is probable that the DX center will never be the most stable form of this impurity in GaAs at

any pressure. If this is the case, then condition (1) above is not satisfied and it is not possible to measure the donor spectrum of Te by the technique used. The non-observance of any donor related absorptions in the spectra, whether sharp or continuum-like, over a range of pressures from 40 to 70 kbar may thus be interpreted as confirmation that Te does not form DX centers at these higher pressures.

The Se DX center in pure GaAs under hydrostatic pressure has not been well studied and like Te, there is no indication that it is or is not the most stable form of the impurity at high pressure. Again, the lack of any observed absorption for this donor using this experimental procedure indicates that probably it is not. It has been determined through photoluminescence measurements that the Se donor in indirect gap GaAs has a binding energy of 115 meV (Kobayashi 1983), and so any donor related absorptions should be well within our detection range. Given the doping concentration of the samples which were studied, any absorptions which did exist would surely have been seen if Se does form a DX center.

Also, Ge has been found to be an unusual donor in the sense that both the DX center and the non-photoionizable deep A_1 configuration play important roles in its behavior under hydrostatic pressure (van der Wel 1993). There is no data in the literature to indicate whether or not the DX state is the most stable at high pressures, but once again, the non-observation of Ge donor spectra implies that it is not.

Concerning the donors which did show some photoabsorption, although the absorption spectrum of S does not display any ground to bound excited state transitions, the shape of the absorption is characteristic of transitions from a bound state to a continuum of states such as the conduction band. The fact that the absorption occurs at much lower energy for the more highly doped sample supports this interpretation since at higher impurity concentrations, electron-electron interaction can be expected to push the energy of the bound state higher, closer to the conduction band (Pearson 1949). The reason for the complete absence of any sharp transitions is still unknown. However, it is very unlikely

that Stark broadening is responsible for the absence of any sharp absorptions, as even the Si doped samples with net carrier concentrations near 10^{17} cm^{-3} displayed a recognizable "peak." The energy at which this absorption occurs is in rough agreement with photoluminescence experiments which indicate that the binding energy of S is approximately 109 meV near the crossover pressure of the X and Γ regions of the conduction band (Leymarie 1990).

Finally, regarding the spectra of the Si and Sn donors, one's first impulse may be to dismiss the lack of additional transition peaks as due to the relatively impure nature of the material ([n] > 10^{15} cm^{-3}). However, in GaP, a semiconductor with a band structure very similar to that of GaAs under pressure, impurity spectra typically consist of at least two or three clearly visible transition lines, even at doping concentrations up to 10^{17} cm^{-3} (Kopylov 1977). Furthermore, the transition lines of donors in GaP are of similar width and are found in a similar spectral range as those of donors in indirect GaAs, indicating that the binding energies and Bohr radii for both should be similar. Thus, the suggestion that Stark broadening has caused the disappearance of all other transition lines is untenable.

It is also impossible that the non-observation of a second line in both donors is due to the energy of that line lying within the reststrahl band. The reststrahl band in GaAs is less than 25 cm^{-1} wide whereas the chemical shift between Si and Sn is almost 100 cm^{-1}. Even if a second line of one of the donors is hidden by the reststrahl reflection, the second line of the other donor should be clearly visible. In the framework of effective mass theory, this leaves only the possibility that the intensity of a second absorption line must be much smaller than that of the observed transition line. In order to carry out a more focused search for a small peak, the theoretical energy spectrum for donors in GaAs was calculated. However, in the process, an interesting possibility concerning the conduction band minimum of GaAs near the X-point was discovered. Instead of the normal parabolic band minimum present in Si, Ge, and at the Γ point in GaAs, it has been argued that this band

minimum, which is found along the Δ symmetry line, is in fact strongly *non*-parabolic with what is known as a "camel's back" shape (Lawaetz 1975). Furthermore, it has been suggested that this non-parabolicity may be universal to this particular conduction band minimum in all compound semiconductors due to the lack of inversion symmetry in the zincblende lattice (Kopylov 1985).

3.1.5 Influence of the Camel's Back

In elemental semiconductors such as Si, the two lowest conduction bands are degenerate at the X symmetry point due to the inversion symmetry of the diamond lattice. At the local minimum, which occurs some distance from the X point along the Δ symmetry axis, the lowest band is parabolic to very good approximation as confirmed by the success of effective mass theory in predicting the energy spectrum of donors in Si. The dispersion relation of the two lowest conduction bands in Si can thus be written as

$$E_n(k) = (-1)^n D\,k + A\,k^2, \qquad n = 1,2 \tag{3.1}$$

where D and A are constants near the conduction band minimum and k, which is 0 at the X symmetry point, is the reciprocal space distance in the <100> direction. The conduction bands of Si are illustrated in figure 3-7.

However, as was pointed out by Lawaetz, III-V semiconductors such as GaP have no inversion symmetry and the degeneracy at the X point is lifted. In order to calculate the shape of the conduction band in this case, degenerate perturbation theory is applied to the band structure of Si. Introducing a repulsive interaction between the two bands, the total Hamiltonian for the system becomes non-diagonal and can be written as

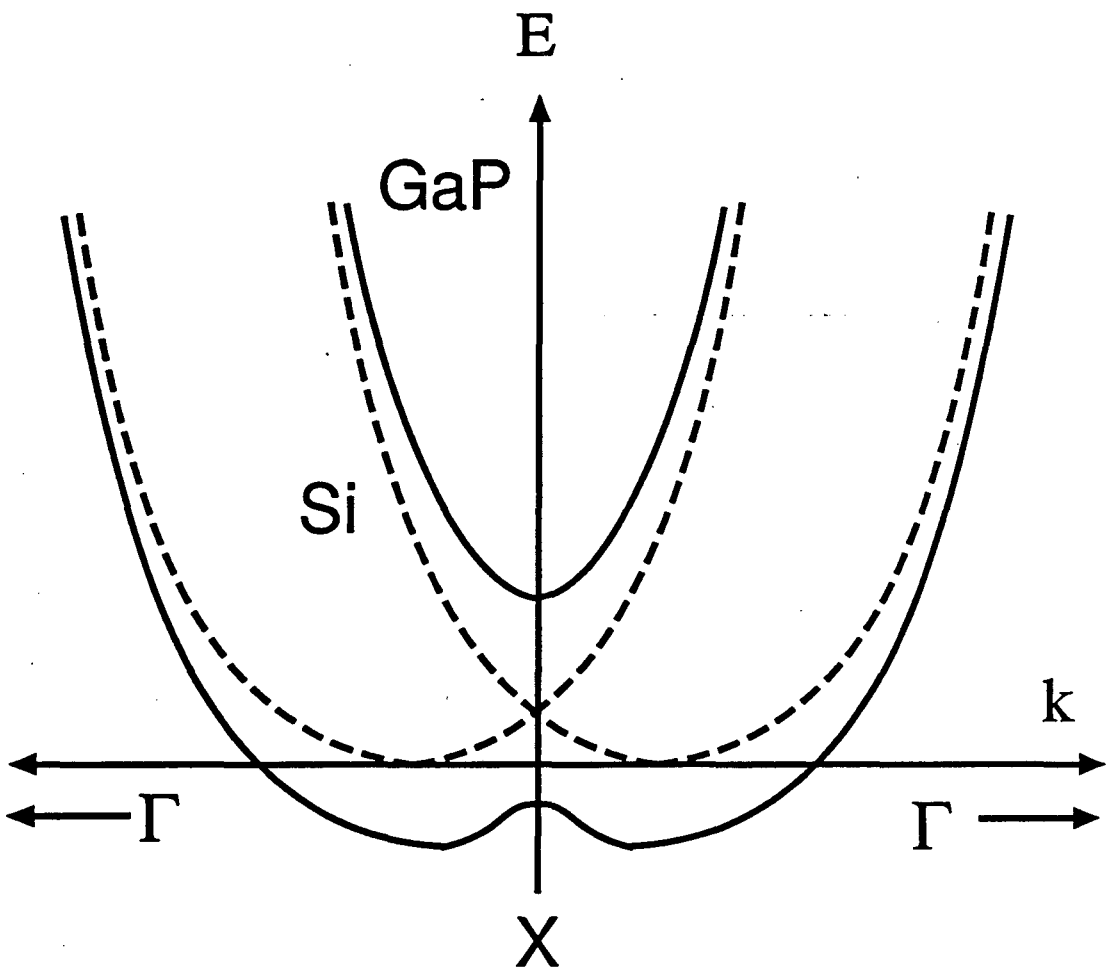


Figure 3-7. Schematic diagram showing conduction bands of Si (dashed line) and GaP (solid line) near the X symmetry point. The center of the graph is the boundary of the Brillouin zone and the horizontal axis is the [100] direction. It has been suggested that all III-V compound semiconductors have conduction bands similar to that of GaP at this symmetry point.

$$H = \begin{vmatrix} E_1(k) & \frac{\delta}{2} \\ \frac{\delta}{2} & E_2(k) \end{vmatrix} \tag{3.2}$$

where the parameter δ is equal to the energy difference between the two bands at the X symmetry point. Diagonalizing this matrix for the new energies, we obtain

$$\begin{aligned} E_{\pm}(k) &= \frac{1}{2}\left(E_1(k) + E_2(k) \pm \sqrt{\left(E_1(k) - E_2(k)\right)^2 + \delta^2} \right) \\ &= Ak^2 - \sqrt{D^2 k^2 + \left(\frac{\delta}{2}\right)^2} \end{aligned} \tag{3.3}$$

As can be seen from (3.3), the band minimum is no longer a simple parabola in the <100> direction. The amount of non-parabolicity is determined by the magnitude of D. Since only the band curvature in the <100> direction is affected in this simple model, the transverse effective masses are unperturbed, while the longitudinal mass, which is inversely proportional to the band curvature in the <100> direction, can no longer be considered a constant, even at relatively short distances from the band minimum (Lawaetz 1975). The band structure of a compound semiconductor such as GaP near the X point is shown superimposed on that of Si in figure 3-7.

Experimental evidence for this so-called "camel's back" structure (so called because the slope of the conduction band at the X point is zero and the effective mass there is negative) has been observed in the donor ionization spectra of GaP (Kopylov 1977) and AlSb (Ahlburn 1968), where it is impossible to fit the observed line spacings to the theoretical spacings calculated by Faulkner (Faulkner 1969) for donors. In addition, what estimates there are of the longitudinal effective mass from donor ionization spectra (Onton 1969, Scott 1976) are completely different from those obtained from cyclotron resonance experiments (Suziki 1976, Miura 1983).

Because of this non-parabolicity, simple effective mass theory as embodied by equations (1.3) and (1.6) are no longer applicable to this problem. The Schrödinger equation for shallow donors in this case is (Kopylov 1977)

$$\left[\frac{\hbar^2\left(k_x^2+k_y^2\right)}{2m_\perp}+\frac{\hbar^2 k_z^2}{2m_0}-\frac{e^2}{\varepsilon r}+\frac{\delta}{2}-\sqrt{D^2 k_z^2+\left(\frac{\delta}{2}\right)^2}\right]\psi = E\psi \tag{3.4}$$

where k_z is in the <100> direction and the energy axis is such that the lower of the two energy bands passes through E=0 at the X symmetry point. Although this equation cannot be solved analytically, one can use simple variational theory to determine upper bounds for the energies of the excited states of shallow donors (Kopylov 1977). Much more sophisticated methods have been employed by other authors (Chang 1980), however, they do not yield significantly better results.

To simplify the calculations, the Hamiltonian in (3.4) was divided into two parts

$$\begin{aligned} H = H^0 + H^1 = &\frac{\hbar^2 k_t^2}{2m_t}+\frac{\hbar^2 k_l^2}{2m_l}-\frac{e^2}{\varepsilon r} \\ &+\frac{\hbar^2 k_l^2}{2m_0}-\frac{\hbar^2 k_l^2}{2m_l}+\frac{\delta}{2}-\sqrt{\frac{\delta^2}{4}+D^2 k_l^2} \end{aligned} \tag{3.5}$$

where

$$H^0 = \frac{\hbar^2 k_t^2}{2m_t}+\frac{\hbar^2 k_l^2}{2m_l}-\frac{e^2}{\varepsilon r} \tag{3.6}$$

and

$$H^1 = \frac{\hbar^2 k_l^2}{2m_0}-\frac{\hbar^2 k_l^2}{2m_l}+\frac{\delta}{2}-\sqrt{\frac{\delta^2}{4}+D^2 k_l^2} \tag{3.7}$$

The first part, H^0, is the usual effective mass theory Hamiltonian for a parabolic conduction band minimum, the solutions of which are well known. What remains then, is to find the expectation values of the non-quadratic part (3.7).

The trial wavefunctions used in this calculation are

$$\begin{aligned}
\psi_{1s}(\mathbf{r}) &= C_{1s}\exp\left(-\sqrt{\alpha_{1s}^2 x^2+\alpha_{1s}^2 y^2+\beta_{1s}^2 z^2}\right)\\
\psi_{2p_0}(\mathbf{r}) &= C_{2p_0}\, z\exp\left(-\sqrt{\alpha_{2p_0}^2 x^2+\alpha_{2p_0}^2 y^2+\beta_{2p_0}^2 z^2}\right)\\
\psi_{2p_\pm}(\mathbf{r}) &= C_{2p_\pm}\, x\exp\left(-\sqrt{\alpha_{2p_\pm}^2 x^2+\alpha_{2p_\pm}^2 y^2+\beta_{2p_\pm}^2 z^2}\right)\\
\psi_{3p_0}(\mathbf{r}) &= C_{3p_0}\, z^3\exp\left(-\beta_{3p_0}|z|\right)
\end{aligned} \tag{3.8}$$

which are simply hydrogenic wavefunctions scaled to take into account the anisotropic effective mass. In order to avoid excessively tedious calculations, the weak dependence of the $3p_0$ wavefunction on x and y was ignored (Kopylov 1977). The C's are normalizing constants while α and β are the variational parameters. The energies of the lowest four states as computed by simple variational theory are

$$\begin{aligned}
E_{1S} &= E_{1s}^0+\frac{|E_{1s}^0|}{3}\left(\frac{m_l}{m_0}-1\right)+\frac{\delta}{2}-\\
&\quad\frac{16}{3\pi}\sqrt{\Delta|E_{1s}^0|\,\delta\,\frac{m_l}{m_0}}\int_0^\infty\frac{dp}{(1+p^2)^3}\sqrt{p^2+\frac{\delta}{4\,\Delta|E_{1s}^0|}\frac{m_0}{m_l}}\\
E_{2p_0} &= E_{2p_0}^0+\frac{3}{5}|E_{2p_0}^0|\left(\frac{m_l}{m_0}-1\right)+\frac{\delta}{2}-\\
&\quad\frac{256}{5\pi}\sqrt{\Delta|E_{2p_0}^0|\,\delta\,\frac{m_l}{m_0}}\int_0^\infty\frac{p^2\,dp}{(1+p^2)^5}\sqrt{p^2+\frac{\delta}{4\,\Delta|E_{2p_0}^0|}\frac{m_0}{m_l}}
\end{aligned} \tag{3.9}$$

$$E_{2p_\pm} = E^0_{2p_\pm} + \frac{\left|E^0_{2p_\pm}\right|}{5}\left(\frac{m_l}{m_0} - 1\right) + \frac{\delta}{2} -$$
$$\frac{32}{5\pi}\sqrt{\Delta\left|E^0_{2p_\pm}\right|\delta\,\frac{m_l}{m_0}}\int_0^\infty \frac{dp}{\left(1+p^2\right)^4}\sqrt{p^2 + \frac{\delta}{4\,\Delta\left|E^0_{2p_\pm}\right|}\frac{m_0}{m_l}}$$

$$E_{3p_0} = E^0_{3p_0} + \frac{\left|E^0_{3p_0}\right|}{5}\left(\frac{m_l}{m_0} - 1\right) + \frac{\delta}{2} -$$
$$\frac{1024}{5\pi}\sqrt{\Delta\left|E^0_{3p_0}\right|\delta\,\frac{m_l}{m_0}}\int_0^\infty \frac{p^2\left(1-p^2\right)^2 dp}{\left(1+p^2\right)^8}\sqrt{p^2 + \frac{\delta}{4\,\Delta\left|E^0_{3p_0}\right|}\frac{m_0}{m_l}}$$

where

$$\Delta = \frac{2m_0}{\hbar^2}\frac{D^2}{\delta} \tag{3.10}$$

The E^0's are the energies of the various states as calculated by the ordinary effective mass theory and m_l is inversely proportional to the curvature of the conduction band at the band minimum near the X-point. Although no experiments have yet been performed to determine the parameters D, δ, and m_l, some theoretical estimates have been made for the values at atmospheric pressure (Kopylov 1985). Using those estimates, numerical values of the energies are

$$\begin{aligned} E_{CB} &= -9.44\ meV \\ E_{1s} &= -49.6\ meV \\ E_{2p_0} &= -26.5\ meV \\ E_{2p_\pm} &= -9.34\ meV \\ E_{3p_0} &= -9.30\ meV \end{aligned} \tag{3.11}$$

As noted previously, all energies are computed with respect to the energy of the lowest conduction band at the X symmetry point. Comparing the binding energies of the excited states with those of a normal hydrogenic donor, one sees that the order is the same,

where the states with larger principal quantum numbers are less tightly bound. Furthermore, the $2p_0$ level is deeper than the $2p_\pm$ level. A simple reason for this fact is that the binding energy of the $2p_\pm$ state is dependent on the curvature of the conduction band perpendicular to the <100> direction while the binding energy of the $2p_0$ state is related to the curvature in the <100> direction. Since the transverse curvatures are larger than the longitudinal curvature, the transverse effective masses are smaller than the longitudinal effective mass, resulting in a smaller binding energy. In addition, the non-parabolicity of the longitudinal curvature causes the energy of the $2p_0$ level to be lowered even further.

Comparing the energies for the $2p_\pm$ and $3p_0$ levels with that of the true conduction band minimum along the Δ axis, one sees that the energies of these two excited states are higher. This indicates that these are not bound states, but autoionized states within the conduction band. In this situation, no discrete transitions from the ground state to these states are possible and only the absorption corresponding to excitations from the 1s to $2p_0$ state would be observed as a sharp transition line. This is in agreement with the experimental observations for Si and Sn.

Using these results, the energy level diagrams for Si and Sn donors associated with the camel's minimum in GaAs are drawn schematically in figure 3-8 along with the level diagram of a theoretical donor. Since the conduction band parameters at high pressure are not known, the atmospheric pressure values have been used. The energy spacing between the ground state and the $2p_0$ state is based on the experimental observations while the energies of the remaining levels are based on the theory. Comparing the estimated ground state energies of Si and Sn with the calculated ground state energy, it is obvious that the chemical shifts are large. This is also the case in GaP, where the theoretical ground state energy is 62 meV but the binding energies of the shallow donors are closer to 100 meV (Kopylov 1977).

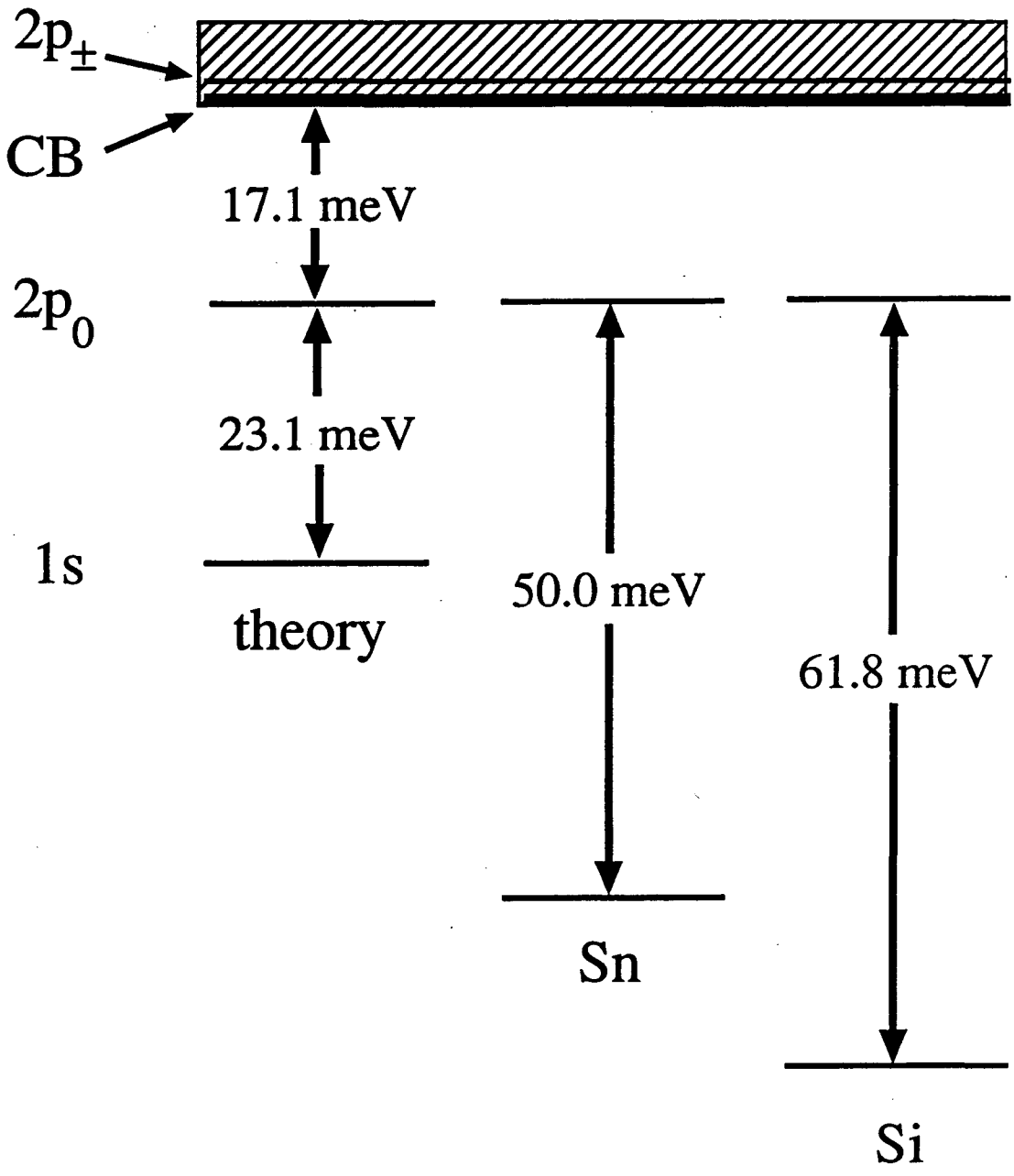


Figure 3-8. Schematic diagram of the ground and bound excited states of Si and Sn donors in indirect gap GaAs. The theoretically calculated levels are also shown. All energies are shown for an applied pressure of 40 kbar.

One effect which has been neglected here is valley-orbit splitting. Unlike direct gap GaAs in which there is only one conduction band minimum in the Brillouin zone, indirect gap GaAs has six equivalent minima located in the <100> directions. Valley-orbit interaction in this case leads to a splitting of the six-fold degenerate ground state into a triply, doubly, and a singly degenerate ground state. For donors in Si, this effect has been shown to lower the ground state energies by a significant fraction of the binding energy calculated by a simple one-band effective mass theory (Aggarwal 1964). In the total absence of information regarding this energy splitting in GaAs however, this effect has not been included in figure 3-8.

It should be noted that since the values obtained through equations (3.9) above are the results of a variational calculation, they are only upper bounds for the true values and without additional knowledge of the actual wavefunctions, it is impossible to estimate the error in the calculation. It is therefore possible that the $2p_{\pm}$ and $3p_0$ excited states are in fact bound states and not resonant states. The GaAs parameters used in the calculation of the donor spectrum are theoretical extrapolations from GaP at a pressure of 1 bar and are almost certainly somewhat different from the true values at 40 kbar. However, based on the signal to noise ratio of the best spectra taken, the peak heights of any additional transitions must be at least a factor of 50 smaller than the height of the observed peak. Such an imbalance would be extremely unusual and has never been observed in any donor for the electronic transitions which are in question here. Although we cannot say with absolute certainty that the model that is presented here is the true reason for the absence of other lines, it is certainly a plausible one.

To further investigate the effects of the camel's back structure on the energy spectrum of donors, it was decided to study the pressure dependence of donors in other indirect III-V semiconductors which should also have a camel's back structure at the conduction band minimum near the X symmetry point. Of the many possibilities, however, only InP, GaP, and AlSb can be grown pure enough to study the donor spectrum

and of these three, only AlSb is suitable for our purposes. InP, although similar in many respects to GaAs, does not become an indirect gap semiconductor until pressures greater than 90 kbar are applied and due to the small size of the diamond anvil cells used in this work, pressures greater than 80 kbar could not be obtained without running a great risk of destroying the diamonds. Although GaP can be grown with doping concentrations appropriate for study by absorption spectroscopy, no evidence of DX center formation by donors in this material has yet been found. Without the ability to obtain such near perfect references as in the case of GaAs, there is little hope of seeing GaP donor spectra by the technique employed here. AlSb, on the other hand, is an indirect gap material with its conduction band minimum near the X point. Furthermore, it is known that donors in this material form DX centers (Becla 1995). Although infrared absorption spectra were investigated in n-type AlSb at atmospheric pressure as early as 1960 (Ahlburn 1968), higher pressures have never been studied.

3.2 X-band donors in AlSb

AlSb is an indirect gap III-V semiconductor with a conduction band minimum located close to the X symmetry point and a bandgap of 1.6 eV. AlSb has been investigated as a possible material for infrared detectors (Brar 1994) and optical holographic memories (McKenna 1996) and is used in combination with InAs to make Heterostructure Field Effect Transistors (HFETs) and High Electron Mobility Transistors (HEMTs) (Bolognesi 1996). However, despite these applications, relatively few efforts have been made to investigate the basic properties of AlSb, principally because of its hygroscopicity. Samples of AlSb left exposed to the atmosphere will disintegrate into powder in just a few weeks, making special precautions necessary for handling and storage.

Because the pressure coefficients of the various conduction band minima, the camel's back minimum near the X-point remains the lowest point for all pressures. Unfortunately, only Se and Te are known to form shallow donors in AlSb. The group IV impurities all preferentially occupy the Sb sublattice to form acceptors and S has not been observed to form donors in this material. Thus, it was impossible to compare the behavior of group IV and group VI donors in this material and to compare them with donors in GaAs.

3.2.1 Samples

A piece of AlSb doped with Se was obtained from Prof. Anant Ramdas at Purdue University. This piece was cut from a wafer which had been grown by Bell and Howell Research Laboratories in the 1960's using the Bridgman technique. The net-impurity concentrations was 10^{16} cm^{-3}. The compensation ratio of this sample appeared to be quite low, as the application of band-edge light made no discernible difference in the donor

spectra. In contrast, the absorption peak in some of the bulk grown Si doped GaAs samples was all but invisible without band edge light, due to severe Stark broadening.

Since AlSb degrades quickly in the presence of water, the final polishing steps in the creation of the samples were performed in methanol.

3.2.2 Selenium

Figure 3-9 shows the donor spectrum of Se in AlSb, taken at atmospheric pressure. Of all the samples surveyed so far, AlSb:Se produces a spectrum which is the most similar to that of the normal hydrogenic donors in Si and Ge. In this spectrum, there are two sharp lines, one of which lies on the shoulder of a very large continuum absorption. Both lines have also been observed in bulk samples and have been identified as caused by the Se donor (Ahlburn 1968). Calculating the energy spectrum for donors in AlSb using eqs. (3.9), we obtain

$$\begin{aligned} E_{CB} &= -7.45\ meV \\ E_{1s} &= -55.9\ meV \\ E_{2p_0} &= -29.7\ meV \\ E_{2p_\pm} &= -10.7\ meV \\ E_{3p_0} &= -10.7\ meV \end{aligned} \tag{3.12}$$

for the binding energies of the first few states. As in the case for GaAs, none of the band parameters of the X conduction band minimum in AlSb have been determined experimentally, so theoretical values extrapolated from GaP have been used. In principal, band parameters can be calculated from the energy spacing of the absorption lines, however, in the present model, there are too many unknown parameters to be able determine from the existing experimental data. Based on the calculated estimates of the excited state energies, two of the observed lines can be assigned to transitions from the

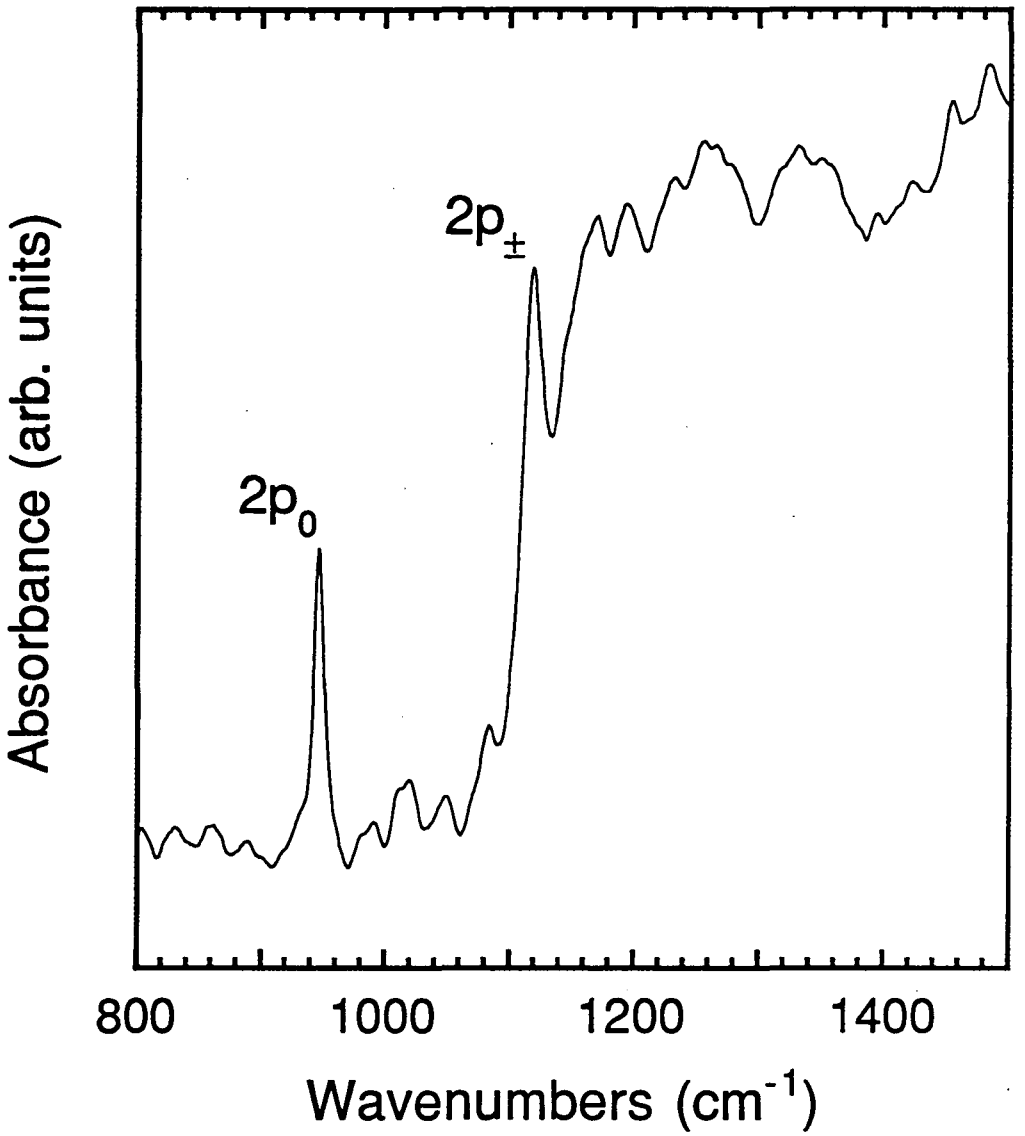


Figure 3-9. Absorption spectrum of AlSb:Se at 1 bar. Two electronic transition peaks can be seen clearly, corresponding to transitions from the ground state to the $2p_0$ and $2p_\pm$ bound excited atates.

ground state to the $2p_0$ and $2p_\pm$ bound excited states. The splitting of these lines under uniaxial stress also confirms this assignment (Ahlburn 1968).

When subjected to hydrostatic pressure, the Se donor spectrum changes in a number of unusual ways (Hsu 1996). The one which is perhaps the most relevant for our purposes is shown by the sequence of spectra in figure 3-10. At very low pressures, the separation of the 1s to $2p_\pm$ peak from the continuum becomes smaller and smaller as the pressure is increased until at approximately 7 kbar, the absorption peak has become virtually indistinguishable. As the pressure continues to increase, the transition peak separates from the continuum and is again clearly visible by 13 kbar. This phenomenon was found to be reproducible over a number of different sample loadings and it is very unlikely that this disappearance is due to uniaxial stress on the sample as the 1s to $2p_0$ transition remains relatively sharp. In addition, the applied pressures are quite low and from visual observation, it did not appear as if the diamonds on gasket were pressing on the sample directly. This unusual behavior is attributed to the ability of the camel's back structure to cause some excited states which are normally bound to become resonant states. As the pressure is increased, the curvature of the conduction band undoubtedly changes. It is possible that in AlSb, it changes in such a way as to force the $2p_\pm$ excited state first out of, then back into, the forbidden gap.

A second way in which the Se donor is unusual is found by plotting the energies of the electronic transitions as a function of pressure. As shown in figure 3-11, the frequencies of these two transition lines decrease quadratically, rather than linearly. Because of the relative deepness of the Se donor, the vibrational mode of the CO_2 dissolved in the pressure medium could be and was used to determine the low temperature pressures at which the measurements were made, resulting in very little scatter of the data. In addition to the surprising quadratic shift, a third peak appears at roughly 30 kbar. Though its position is pressure independent at first, it eventually begins to decrease in energy with applied pressure at higher values.

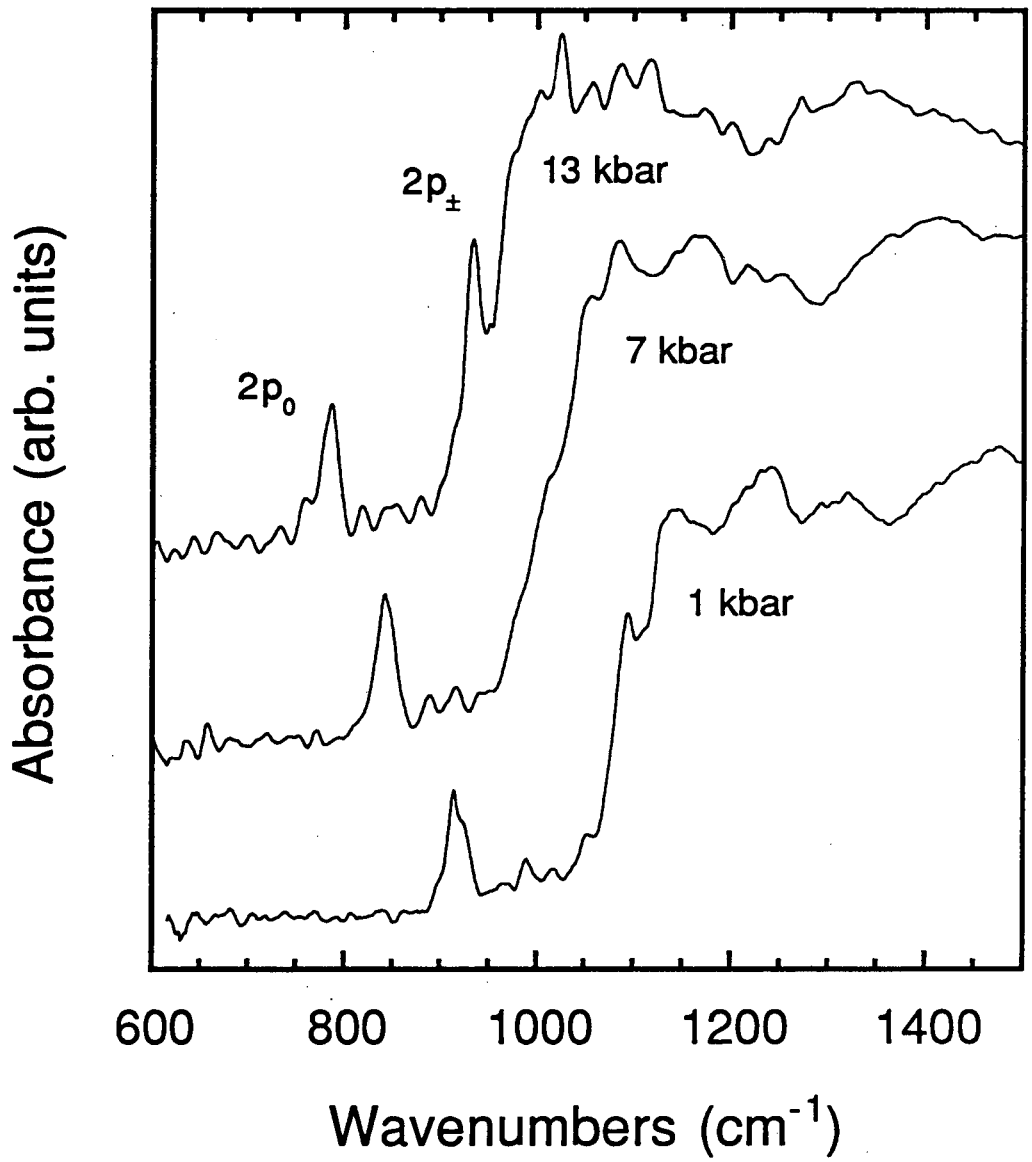


Figure 3-10. AlSb:Se donor spectra taken at low pressure. The second electronic transition peak disappears into the continuum upon increasing the pressure, then reappears as the pressure is increased further.

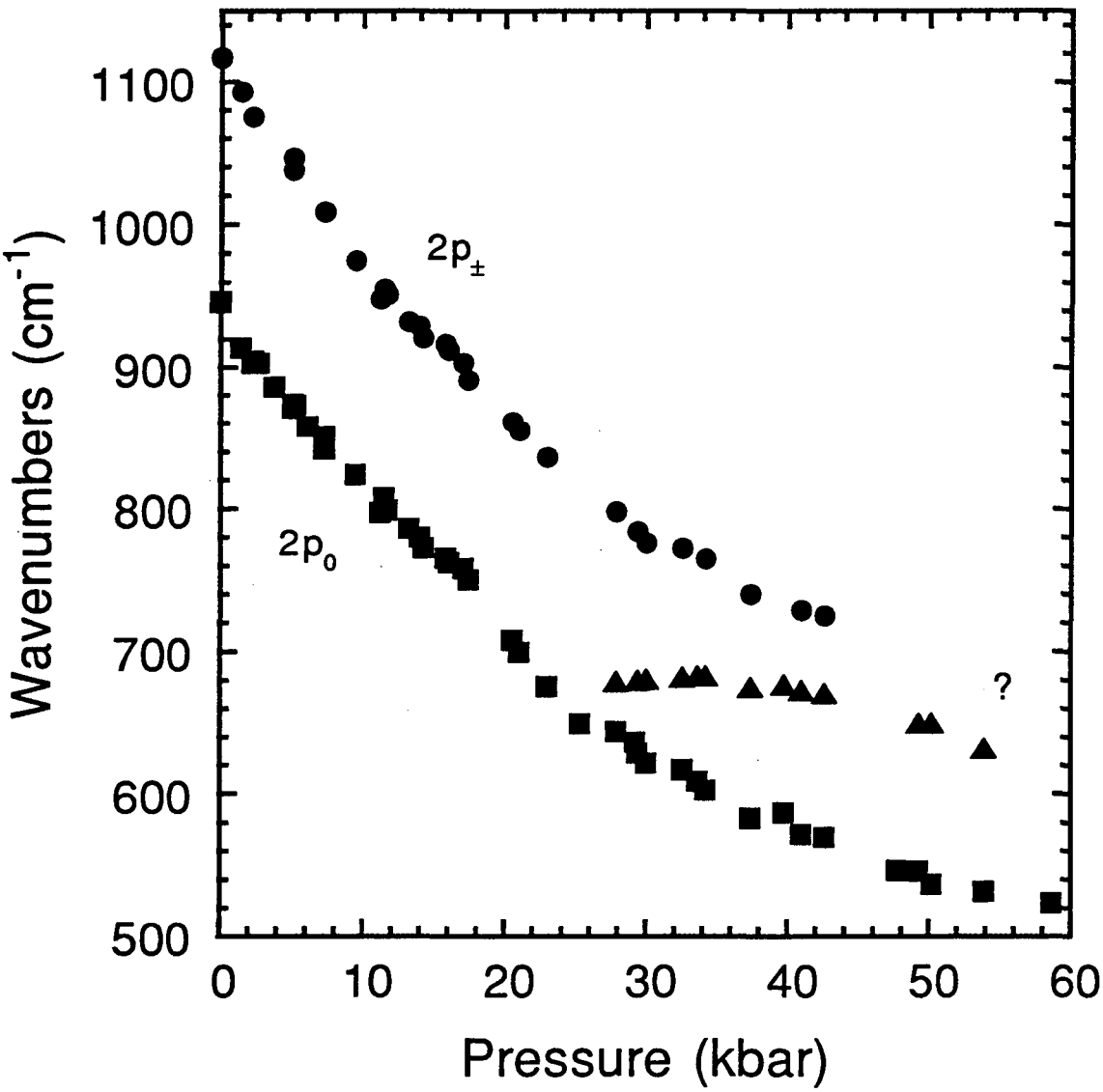


Figure 3-11. Pressure dependence of the electronic transition peaks in AlSb:Se. Also plotted is a third peak (labeled ?) which first appears at 30 kbar, then increases in intensity as the pressure increases.

Coinciding with the appearance of this third peak, the absorption peak corresponding to the 1s to $2p_{\pm}$ transition begins to broaden until it is all but impossible to track. During this process, the new absorption line grows in intensity until it resembles the original electronic transition peak. The progression of these two events is tracked in figure 3-12.

A possible explanation for these observations can be found by considering the energies of the two-phonon modes in AlSb. Although no measurements of the pressure dependence of the two-phonon modes in AlSb have been performed, Ves and co-workers (Ves 1985) have measured the pressure dependence of the one-phonon LO and TO modes at the zone center. By forming various additive combinations of these frequencies, the pressure dependences of the two optical phonon modes may be estimated. The energies are plotted in figure 3-13 along with the experimentally observed energies of the 1s to $2p_{\pm}$ transition and the mysterious third peak.

If we assume a weak repulsive interaction between the 1s to $2p_{\pm}$ transition and one of these multi-phonon modes, we can create a simple two-level model to explain the observations. The total Hamiltonian of this system is given by

$$H = H_{elec} + H_{pho} + H_{\text{int}} \tag{3.13}$$

where H_{elec} and H_{pho} are the Hamiltonians for the electronic transition and the multiple phonons respectively and H_{int} describes the interaction between these two systems. If H_{int} can be treated as a weak perturbation, then the Hamiltonian matrix can be written as

$$H = \begin{vmatrix} \omega_{elec} & A \\ A & \omega_{phon} \end{vmatrix} \tag{3.14}$$

where

$$A = \left\langle n_{pho}, \psi_{2p_{\pm}} \middle| H_{\text{int}} \middle| n+1_{pho}, \psi_{1s} \right\rangle \tag{3.15}$$

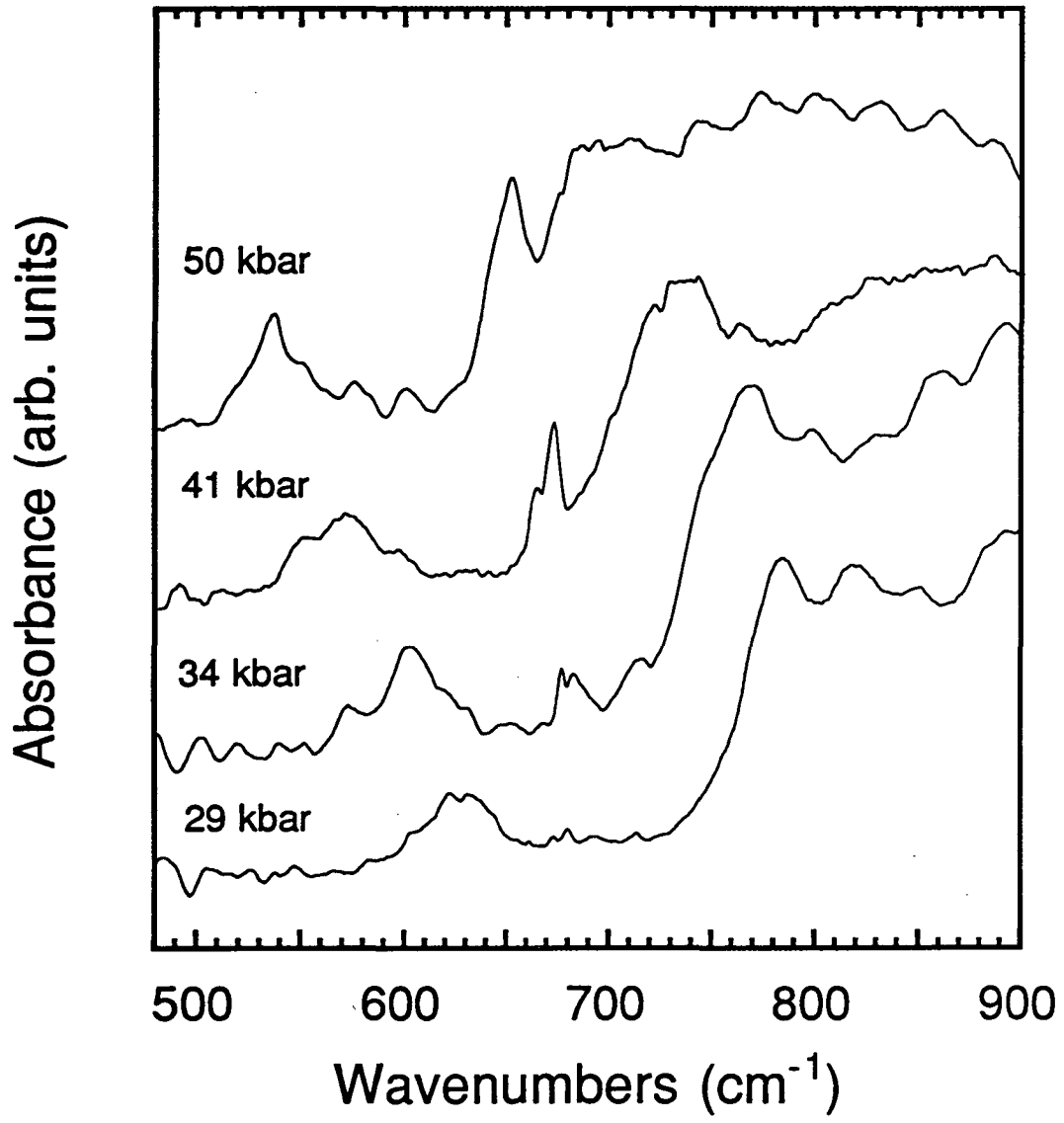


Figure 3-12. Sequence of AlSb:Se absorption spectra showing emergence and growth of the third peak as the pressure is increased.

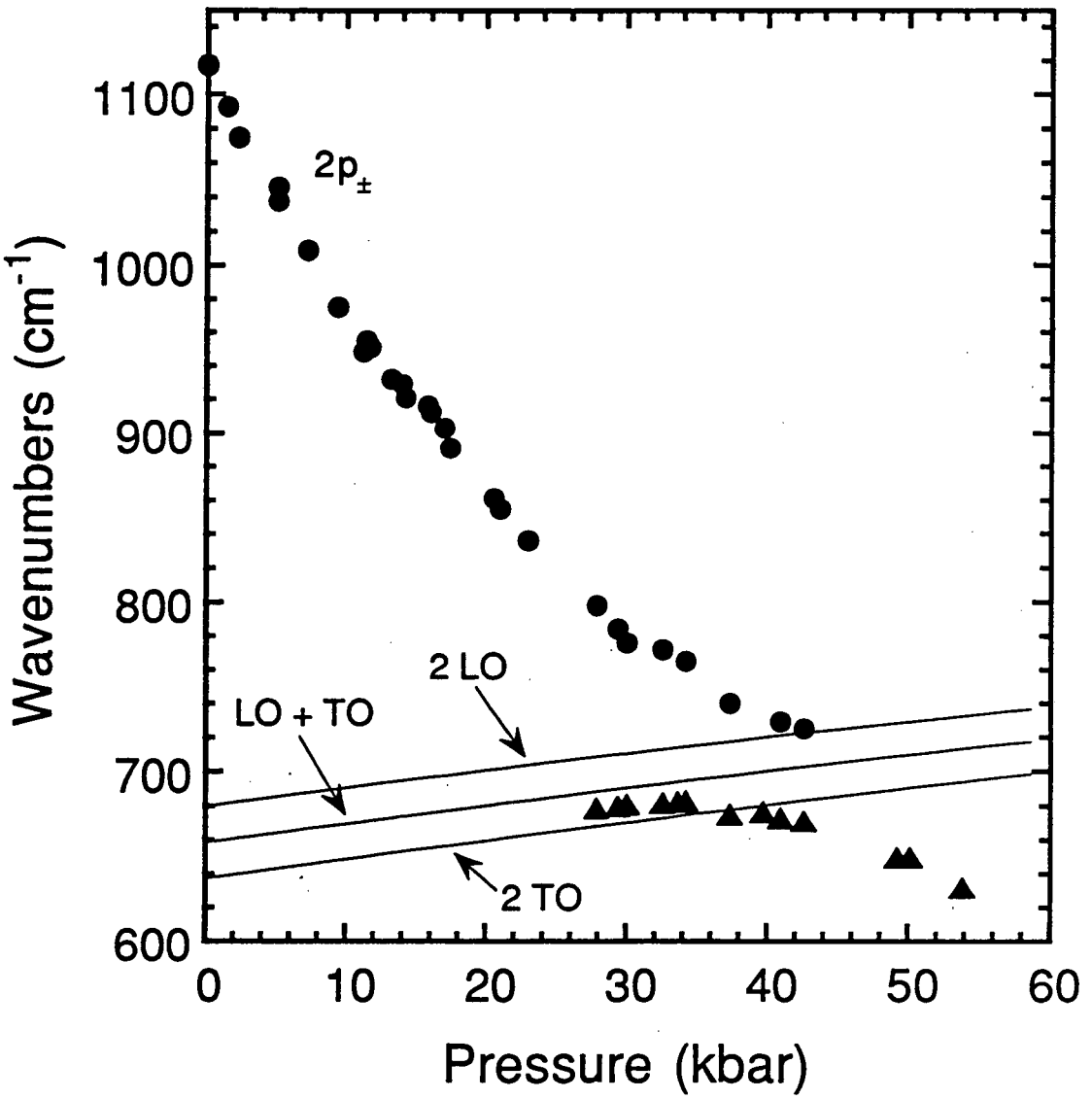


Figure 3-13. Energies of the 1s to $2p_{\pm}$ electronic transition peak and the third peak plotted as a function of pressure. Also plotted are the theoretically extrapolated frequencies of the 2 LO, LO+TO, and 2 TO phonons.

Using simple degenerate perturbation theory to diagonalize this Hamiltonian, we obtain

$$\omega_{\pm} = \frac{1}{2}\left(\omega_{elec} + \omega_{pho} \pm \sqrt{\left(\omega_{elec} - \omega_{pho}\right)^2 + 4A^2}\right) \tag{3.16}$$

for the new eigenenergies of this system. This is exactly the same type of calculation which was performed to obtain the camel's back structure of the conduction band in the previous section. The wavefunctions corresponding to eigenstates of this Hamiltonian are no longer pure electronic transitions or pure phonon excitations, but a mixture of the two given by

$$|\psi\rangle = a|\psi_{pho}\rangle + b|\psi_{elec}\rangle \tag{3.17}$$

where the square modulus of the coefficient a is

$$|a|^2 = \frac{A^2}{\left(\omega_{pho} - \omega_{+}\right)^2 + A^2} \tag{3.18}$$

Experimentally, this quantity is the normalized peak area of the lower energy peak.

The pressure dependence of the electronic transition was estimated to be

$$\omega_{elec\ (cm^{-1})} = 1109.7 - 14.27P_{(kbar)} + 0.1013P^2_{(kbar)} \tag{3.19}$$

by performing a quadratic fit to the energies of the peaks which appeared to correspond to the 1s to $2p_{\pm}$ electronic transition. The interacting phonon was assumed to be the LO+TO mode, with a pressure dependence of

$$\omega_{pho\ (cm^{-1})} = 658.7 + 1.116P_{(kbar)} - 0.0019P^2_{(kbar)} \tag{3.20}$$

There is no special reason for choosing this particular combination except that of the three possible zone center two-phonon combinations shown in figure 3-13, it produces the best

fit. In fact, due to anharmonic terms, eq. (3.20) likely overestimates the frequency of the LO+TO mode and is perhaps closer to the true frequency of the 2 LO mode. Although it is more likely that the observed absorption is due to zone edge rather than zone center two-phonon modes due to the greater density of states at the zone edge, the dispersion curves of optical phonons in AlSb are quite flat. Thus, little error is introduced by using the known zone center frequencies.

The new eigenenergies obtained through (3.16), (3.19) and (3.20) are plotted in Figure 3-14 along with the experimentally observed energies. The interaction parameter A was chosen to produce the best fit between theory and experiment, good agreement being obtained for a value of A = 25 cm^{-1}. The situation shown here is a classic example of an anti-crossing of two weakly interacting levels. As the pressure is increased, the lower branch, which is initially "phonon-like," acquires more of the characteristics of an electronic transition, becoming "electronic-transition like." Similarly, the upper branch which is initially "electronic-transition-like," becomes more and more "phonon-like" at high pressures, the area of its peak decreasing to reflect this change.

The experimentally observed normalized area of the lower peak is plotted along with the theoretical prediction of the simple two-level model (eq. 3-18) in figure 3-15. Since the original 1s to $2p_{\pm}$ electronic transition peak becomes too difficult to trace at the higher pressures, the area of this lower branch peak was normalized to that of the 1s to $2p_0$ transition. Thus, it is implicitly assumed that the areas of the two electronic transition peaks are equal and remain so even as the pressure is increased. At high pressures, the presence of Fabry-Perot fringes in the spectrum and the closeness of this peak to the continuum absorption made the determination of precise peak areas difficult, hence the greater scatter in the data. Naturally the same value of 25 cm^{-1} was used for A in the calculation of the theoretical curve.

It should be noted that only a single free parameter (A) was used in this model. The pressure dependence of the 1s to $2p_{\pm}$ electronic transition was determined by fitting to

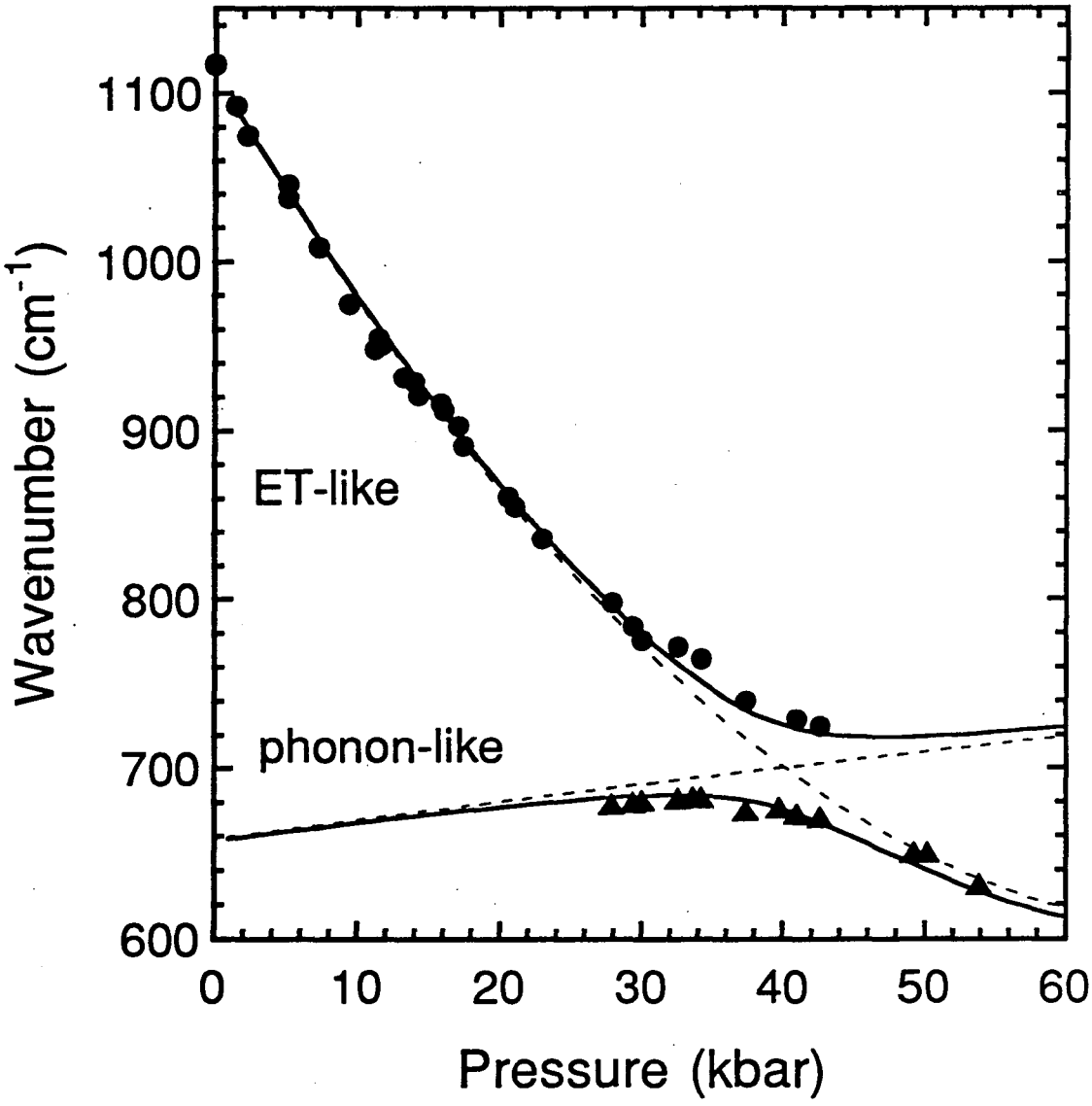


Figure 3-14. Pressure dependence of the frequencies of the upper and lower branches as calculated from the simple two level model. The curves calculated from (3.16) are superimposed on the experimentally observed peak positions. The value of the interaction parameter A is 25 cm^{-1}.

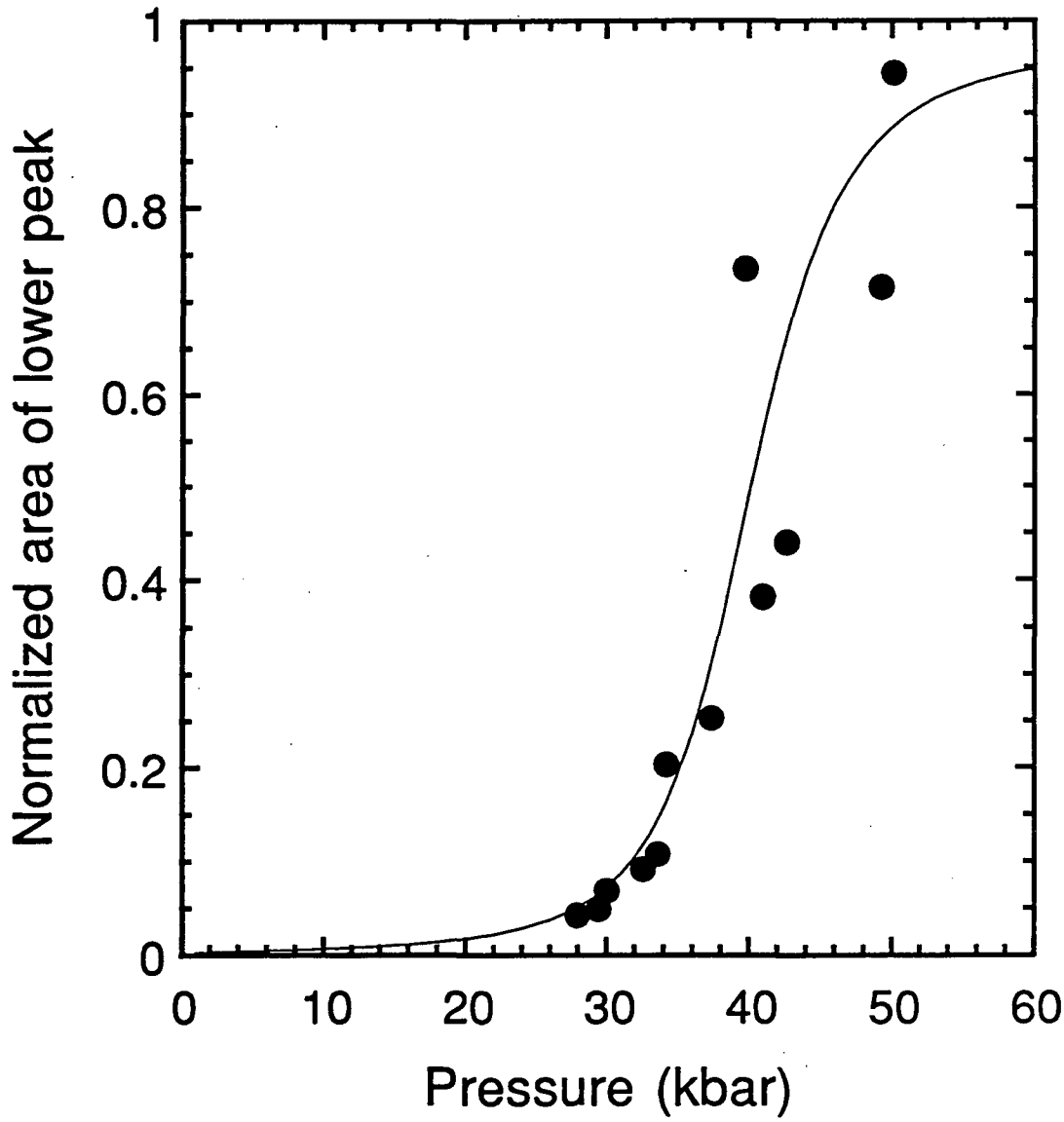


Figure 3-15. Pressure dependence of the area of the third peak divided by the area of the 1s to $2p_0$ electronic transition peak. The curve is calculated from (3.18) assuming that the oscillator strengths of both electronic transitions are equal at all pressures.

experimental data while the energy of the multi-phonon mode was obtained by adding the frequencies of the zone center LO and TO phonon modes as determined by (Ves 1986).

Although the interaction of the phonons with the 1s to $2p_{\pm}$ transition is quite spectacular, it is not known why no interaction of these phonons with the 1s to $2p_0$ transition is observed. The magnitude of the effect is ultimately determined by the interaction parameter A in (3.15) which depends on the initial and final states of the particular electronic transition involved and perhaps by some coincidence, the value corresponding to the interaction between the 1s to $2p_0$ transition with the multi-phonon mode is very small. Alternatively, it is possible that the interaction between the two modes may be zero due to incompatible symmetries.

4. Conclusion

From what was to be a routine measurement of the GaAs:Si X-band donor spectrum, we have discovered the behavior of impurities associated with a camel's back conduction band minimum to be both diverse and puzzling. In GaAs at hydrostatic pressures above 40 kbar, two different donor spectra have been identified: one in which the absorption spectrum consists of a single transition peak, as is the case for Si and Sn, and one in which the spectrum shows only a broad continuum absorption, as is the case for S. Unfortunately, due to the difficulty of performing infrared absorption spectroscopy through a diamond anvil cell, none of the other shallow donors (Ge, Se, Te) could be studied. We were therefore unable to determine whether these two behaviors were characteristic of group IV and group VI impurities.

A variational calculation of the energy spectrum of a camel's back donor in GaAs provided a possible explanation for the behaviors observed for Si and Sn, namely that the non-parabolicity of the band minimum caused some of the normally bound excited states of the donor to become unbound, having energies within the conduction band continuum. Our study of the pressure dependence of the Se donor spectrum in AlSb appears to support this model, showing that for certain pressures, a peak corresponding to a ground to bound excited state transition can be made to disappear into the continuum absorption. Although this model does not explain the total lack of sharp transitions for the S camel's back donor in GaAs, a number of authors (Morgan 1968, Kumagai 1984, Piotrkowski 1990) have suggested that due to the different local environments, that the pressure dependencies and properties of deep levels in compound semiconductors can be expected to depend on the sublattice which the impurity atom occupies. Even though the impurities studied here are nominally "shallow" donors, the relatively large binding energies of 80 meV or greater indicate that the electronic Bohr radii of these centers are small compared to those of the

normal donors in Si, Ge, and direct gap GaAs and thus that the central cell potential can be expected to exert a greater influence on the donor properties.

Another interesting phenomenon discovered during the investigation of donors in AlSb was an anti-crossing of an electronic transition with a two-phonon mode. Although phonons are known to interact in this manner with free electron excitations (such as cyclotron resonance excitations under a varying magnetic field, also known as the magnetopolaron effect), we believe that this is the first observation of a phonon interacting with a impurity bound electron.

Very recently, an effort was made to study the infrared absorption spectrum of the Te donor in AlSb. From the results for AlSb:Se, one might expect to observe two electronic transitions of approximately equal intensity and for the energies of these transitions to shift quadratically as a function of pressure. However, this was not the case. Although two electronic transitions were indeed observed and identified, the peak corresponding to the 1s to $2p_0$ transition was much smaller than the 1s to $2p_\pm$ peak, which was also much greater in intensity than the continuum absorption. Furthermore, the energy of this 1s to $2p_\pm$ peak shifts to lower energies linearly with pressure while the 1s to $2p_0$ peak broadens in such a way as to make a determination of its pressure dependence impossible. It does not appear likely that this broadening is due to uniaxial stress as the width of the 1s to $2p_\pm$ transition remained roughly constant. The study of the AlSb:Te system is still in its preliminary stages however, and it is too early to make more than these qualitative observations.

Due the lack of consistent and comprehensive data, no conclusive picture of the effect of the camel's back can yet be drawn. The following are some suggestions for further studies and work regarding this problem.

First, a study of X-band donors in InP could be made. As mentioned previously, InP does not become an indirect gap semiconductor until pressures greater than 90 kbar are applied. With a different cell design, such as a piston-cylinder type, such pressures could

easily be obtained. In addition, it has been established that at least one donor (S) in InP, preferentially occupies the DX state at pressures greater than 82 kbar (Wolk 1992b). This will enable measurements to be made using our technique.

Second, a different technique altogether, photothermal ionization spectroscopy (PTIS), could be used to investigate the energy spectrum of these donors. In this technique, contacts are made to the sample and the sample itself is used as the detector. Relative to all optical characterization methods, PTIS has by far the greatest sensitivity. One possible problem with this technique however, is the fact that in order to introduce wires into the diamond cell, none of the traditional pressure media listed in Table 2-1 can be used, since the act of sealing the gasket would also cut the wires leading to the sample. Instead, a solid pressure medium such as $CaSO_4$ powder must be used. Unfortunately, the pressures generated by such a material are usually quite inhomogeneous (variations of $\pm$ 4 kbar are not unknown). However, if the pressure dependence of the transition lines is not too large, then the increased sensitivity gained through this technique may still be an advantage.

Although, the behavior of shallow donors in semiconductors is regarded as one of the better understood aspects of these materials, our studies of the camel's back donors shows that there is yet much work to be done.

5. References

Aggarwal, R. L., Solid State Comm. **2**, 1963 (1964).

Ahlburn, B. T. and A. K. Ramdas, Phys. Rev. **167**, 717 (1968).

Alexiev, D. and K. S. A. Butcher, Nuc. Instr. & Meth. B **83**, 430 (1993).

Armistead, C. J., P. Knowles, S. P. Nadja, and R. A. Stradling, J. Phys. C **17**, 6415 (1984).

Armistead, C. J., F. Kuchar, S. P. Nadja, S. Porowski, C. Sotomayor-Torres, R. A. Stradling, and Z. Wasilewski, Proceedings of the 17th International Conference on the Physics of Semiconductors, San Francisco, USA, 1984, edited by J. D. Chadi and W. A. Harrison, (Springer-Verlag, New York, 1985), p. 1047.

Becla, P., A. Witt, J. Lagowski, and W. Walukiewicz, Appl. Phys. Lett. **67**, 395 (1995).

Bell, R. J., *Introductory Fourier Transform Spectroscopy*, (Academic Press, New York, 1972).

Bolognesi, C. R. and D. H. Chow, IEEE Elec. Dev. Lett. 17, 534 (1996).

Bose, S. S., B. Lee, M. H. Kim, and G. E. Stillman, Appl. Phys. Lett. **51**, 937 (1987).

Brar, B., L. Samoska, H. Kroemer, and J. H. English, J. Vac. Sci. and Tech. B **12**, 1242 (1994).

Bridgmann, P. W., *The Physics of High Pressures* (Bell, London, 1952).

Chadi, D. J. and K. J. Chang, Phys. Rev. B **39**, 10063 (1989).

Chang, Y.C. and T. C. McGill, Solid State Comm. **33**, 1035 (1980).

Davies, G., Chem. Phys. Carbon **13**, 1 (1977).

Dmochowski, J. E., R. A. Stradling, D. J. Dunstan, A. D. Prins, A. R. Adams, K. E. Singer, T. Fujisawa, H. Kukimoto, Materials Science Forum **143-147**, 1075 (1994).

Faulkner, R. A., Phys. Rev. **184**, 713 (1969).

Ferraro, J. R., *Vibrational Spectroscopy at High Pressures - The Diamond Anvil Cell*, (Academic Press, New York, 1984).

Forman, R. A., G. J. Piermarini, J. D. Barnett, and S. Block, Science **176**, 284 (1972).

Hsu, L., A. K. Ramdas and E. E. Haller, Proceedings of the 7th International Conference on Shallow Level Centers in Semiconductors, Amsterdam, The Netherlands, 1996(to be published).

Hsu, L., S. Zehender, E. Bauser, and E. E. Haller, Phys. Rev. B **55**, 10515 (1997a).

Hsu, L. and E. E. Haller, in preparation (1997b).

Hubík, P., V. Smíd, J. Kristofik, B. Stepaánek, and V. Sestáková, Solid State Comm. **86**, 19 (1993).

Jagannath, C., Z. W. Grabowski, and A. K. Ramdas, Phys. Rev. **B 23**, 2082 (1981).

Jamieson, J. C., A. W. Lawson, and N. D. Nachtrieb, Rev. Sci. Instrum. **30**, 1016 (1959).

Jayaraman, A., Rev. Mod. Phys. **55**, 65 (1983).

Jayaraman, A., Rev. Sci. Instrum. **57**, 1013 (1986).

Jost W., M. Kunzer, U. Kaufmann, and H. Bender, Phys. Rev. B **50**, 4341 (1994).

Kittel, C. and A. H. Mitchell, Phys. Rev. **96**, 1488 (1954).

Kobayashi, M., T. Yokoyama, S. Narita, Jap. J. Appl. Phys. **22**, L612 (1983).

Kopylov, A. A. and A. N. Pikhtin, Sov. Phys. Semic. **11**, 510 (1977).

Kopylov, A. A., Solid State Comm. **56**, 1 (1985).

Kumagai, O., H. Kawai, Y. Mori, K. Kaneko, Appl. Phys. Lett. **45**, 1322 (1984).

Lang, D. V. and R. A. Logan, Phys. Rev. Lett. **39**, 635 (1977).

Larsen, D. M. Phys. Rev. B **13**, 1681 (1975).

Lawaetz, P., Solid State Comm. **16**, 65 (1975).

Lawson, A. W. and T. Y. Tang, Rev. Sci. Instrum. **21**, 815 (1950).

Lee, M.-W., D. Romero, H. D. Drew, M. Shayegan, and B. S. Elman, Solid State Comm. **66**, 23 (1988).

Lee, S., J. Sanchez-Dehesa, J. D. Dow, Phys. Rev. B **32**, 1152 (1985).

Leroux, M., G. Neu, C. Verie, Solid State Comm. **58**, 289 (1986).

Leroux, M., G. Neu, J. Leymarie, P. Gibart, Semicon. Sci. Tech. **10**, 672 (1995).

Leymarie, J., M. Leroux, and G. Neu, Phys. Rev. B **42**, 1482 (1990).

Liu, X., L. Samuelson, M. E. Pistol, M. Gerling, S. Nilsson, Phys. Rev. Rev. **B 42**, 11791, (1990).

Luttinger, J. M. and W. Kohn, Phys. Rev. **97**, 869 (1955).

Mao, H.K., P. M. Bell, R. M. Hazen, and D. J. Steinberg, J. Appl. Phys. **49**, 3276 (1978).

Mao., H. K. and P. M. Bell, Science **203**, 1004 (1979).

McCluskey, M., L. Hsu, L. Wang, and E. E. Haller, Phys. Rev. B **54**, 8962 (1996).

McKenna, J. M., D. D. Nolte, W. Walukiewicz, and P. Becla, Appl. Phys. Lett. **68**, 735 (1996).

Miura, N., G. Kido, M. Suekane, and S. Chikazumi, J. Phys. Soc. Jpn. **52**, 2838 (1983).

Morgan, T. N., Phys. Rev. Lett. **21**, 819 (1968).

Mott, N.F. and R.W. Gurney, Electronic Processes in Ionic Crystals (Clarendon: Oxford, 1940).

Narita, S. and M. Miyao, Solid State Comm. **9**, 2161 (1971).

Onton, A. Phys. Rev. **186**, 786 (1969).

Pearson, G.L. and J. Bardeen, Phys. Rev. **75**, 865 (1949).

Piermarini G. J., and S. Block, Rev. Sci. Instr. **46**, 973 (1975).

Piotrkowski, R., T. Suski, P. Wisniewski, K. Ploog, and J. Knecht, J. Appl. Phys. **68**, 3377 (1990).

Sallese, J. M., D. Lavielle, J. Singleton, A. Levcuras, J.-C. Grenet, P. Gibart, and J.-C. Portal, Phys. Stat. Sol. A **119**, K41 (1990).

Schiferl, D., S. Buchsbaum, and R. L. Mills, J. Phys. Chem. **89**, 2324 (1985).

Scott, W. and J. R. Onffroy, Phys. Rev. B **13**, 1664 (1976).

Seal, M., High Temp. High Pres. **16**, 573 (1984).

Seal, M. in High Pressure Research in Mineral Physics, ed. M. H. Manghnani and Y. Syono, Terra Scientific, Tokyo, 1987, p. 35.

Shan, W., P. Y. Yu, M. F. Li, W. L. Hansen, and E. Bauser, Phys. Rev. B **40**, 7831 (1989).

Sterer, E., M. P. Pasternak, and R. D. Taylor, Rev. Sci. Instrum. **61**, 1117 (1990).

Stillman, G. E., S. S. Bose, M. H. Kim, B. Lee, T. S. Low, Handbook on Semiconductors, ed. T. S. Moss, vol. 3A, 783 (1994).

Sumner, T. J., S. M. Grant, D. Alexiev, and K. S. A. and Butcher, Nuc. Instr. & Meth. A **348**, 518 (1994).

Suski, T., P. Wisniewski, C. Skierbiszewski, L. H. Dmowski, P. J. van der Wel, J. Singleton, L. J. Giling, and J. J. Harris, J. Appl. Phys. **69**, 3087 (1991).

Suziki, K. and N. Miura, Solid State Comm. **18**, 233 (1976).

Turner, W. J. and W. E. Reese, Phys. Rev. **117**, 1003 (1960).

Van der Wel, P. J., P. Wisniewski, T. Suski, J. Singleton, C. Skierbiszewski, L. J. Giling, R. Warburton, P. J. Walker, N. J. Mason, R. J. Nicholas, and M. Eremets, J. Phys: Condens. Mat. **5**, 5001 (1993).

Van Valkenburg, A., Conference Internationale Sur-les-Hautes Pressions, Le Creusot, Saone-et-Loire, France (1965).

Ves, S., K. Strössner and M. Cardona, Solid State Comm. **57**, 483 (1986).

Vesaghi, M. A., Phys. Rev. B **25**, 5436 (1982).

Wasilewski, Z. and R. A. Stradling, Semicon. Sci. and Tech. **1**, 264 (1986).

Weir, C. E., E. R. Lippincott, A. Van Valkenburg, and E. N. Bunting, J. Res. Natl. Bur. Stand., Sec. A **63**, 55 (1959).

Wolford, D. J. and J. A. Bradley, Solid State Comm. **53**, 1069 (1985).

Wolk, J. A., M. B. Kruger, J. N. Heyman, W. Walukiewicz, R. Jeanloz, and E. E. Haller, Phys. Rev. Lett. **66**, 774 (1991).

Wolk, J. A., W. Walukiewicz, M. L. W. Thewalt, and E. E. Haller, Phys. Rev. Lett. **68**, 3619 (1992a).

Wolk, J. A., DX Centers in III-V Semiconductors under Hydrostatic Pressure, Ph. D. thesis, Physics Dept., Univ. of California--Berkeley and LBL Report No. LBL-33379 (1992b).

Yu, P. Y., B. Welber, Solid State Comm. **25**, 209 (1978).

Transport Studies in Group III-Nitride Semiconductors

1. Introduction

From a technological standpoint, one of the most important parameters of a semiconductor is its mobility. Defined as the proportionality constant between the magnitude of an electric field applied to the semiconductor and the resulting average drift velocity of the carriers, the mobility also describes how quickly the charge carriers in a semiconductor can respond to external fields. In today's semiconductor industry, where devices are made for a variety of high speed applications, a knowledge of the mobilities of carriers in various materials, the factors affecting the mobilities, and the intrinsic mobillity limits for each material is important. In this section, I describe calculations which were performed to investigate the factors limiting the mobilities in n-type III-V nitrides.

In the most general case, to calculate the mobility in a particular material, one must set up the Boltzmann equation for the velocity distribution function f of the electrons and then calculate its rate of change $[df/dt]_{COLL}$ due to collisions of the electrons with other particles and quasi-particles in the lattice (Howarth 1953). In principle, this is done by solving the Boltzmann equation, which can be a difficult mathematical task. In the case where electrons undergo only elastic collisions, however, it is possible to express $[df/dt]_{COLL}$ in the form:

$$\left[\frac{\partial f}{\partial t}\right]_{COLL} = \frac{-(f - f_0)}{\tau} \tag{1.1}$$

where f_0 is the equilibrium velocity distribution of the electrons and τ is a quantity known as the relaxation or scattering time which does not depend on the velocity distribution. To find the mobility in this situation, one can imagine the following picture: In a

semiconductor without any external applied fields, the carriers move around randomly, with zero average velocity, and with a mean time between collisions of τ. If an electric field is now applied, the free electrons will develop an acceleration anti-parallel to the field and the magnitude of the acceleration can be calculated from Newton's second law

$$F = m^* a = eE \tag{1.2}$$

Since the carriers, on average, are accelerated for a time τ before undergoing a collision (which is assumed to bring their velocities back to the equilibrium distribution f_0 with zero average velocity), the average velocity they attain before a collision is

$$v_{AVG} = a\tau = \frac{e\tau}{m^*} E = \mu E \tag{1.3}$$

giving us the mobility in terms of the charge, effective mass, and mean time between collisions of the carriers.

This simplistic expression belies the fact that τ is, in general, dependent on many different factors, making it quite complicated to calculate. In most cases, it is on the order of picoseconds.

In addition, one must be extremely careful when referring to τ as a "mean time between collisions." In transport calculations such as these, the actual parameter of interest is the time constant for the decay of an electron current. Thus, τ involved in the calculations is more precisely the "mean time between collisions which cause an electron to cease its contribution to the current." Since all of the scattering mechanisms described below are capable of scattering an electron into any angle and all but one are elastic, the τ

referring to the mean time between collisions and the τ_c referring to time constant for the decay of the current are related by

$$\tau_c = \tau \langle 1 - \cos\theta \rangle. \tag{1.4}$$

With the additional factor of 1–cos θ, small angle scattering events (in which the net electron current changes by only a little) receive much less weight than near-perfect backscatters (in which the electron now contributes a current opposite to the original direction). Hereafter, all τ's will represent the time constant (or relaxation time) for decay of the current.

Another important caveat is that all calculations in this work were performed in the low field regime. Here, the change of the velocity distribution function of the electrons from the equilibrium distribution is small and the deviations can be taken as linear. Also, the average gain in velocity of the electrons in-between collisions is small compared to their thermal velocity. On the other hand, most semiconducting devices contain space-charge regions where the electric fields may be on the order of 10 kV/cm. In this, the high-field regime, the velocity distribution of the carriers is skewed towards higher velocities and the deviation from the equilibrium distribution is considerable. Furthermore, intervalley scattering, in which more than one conduction band minimum is involved in determining the transport properties, becomes an important consideration--sometimes leading to a saturation or even a decrease in the electron drift velocity with increasing field. Nevertheless, an understanding of the low field mobility in semiconductors is an essential part of an investigation of their basic properties.

1.1 Scattering Mechanisms

In the real world, many factors prevent an electron from being accelerated forever by an electric field. In this section, I discuss the various mechanisms which were considered in this work for calculating the electron mobility. These are scattering by acoustic and optical phonons, charged impurities, the localized potentials of resonant defects, and alloy disorder.

In polar semiconductors, the three most important phonon scattering mechanisms for electrons have been shown to be scattering from polar optical phonons, scattering from acoustic phonons via the deformation potential, and piezoelectric mode scattering by acoustic phonons (Ehrenreich 1960). In principle, deformation potential scattering due to optical phonons is also possible, but this has been shown to vanish for materials whose conduction band minima have Γ_6 symmetry (Harrison 1956), a class which encompasses all zincblende direct gap III-V compound semiconductors. Phonons cause the lattice atoms to move from their equilibrium positions, leading to two effects. The first is the creation of electric dipoles. The interaction of the conduction electrons with these dipoles is the basis for scattering from polar optical phonons and from acoustic phonons via piezoelectric-mode scattering. The second effect of disrupting the periodicity of the lattice is that the Bloch wavefunctions of the electrons are no longer momentum eigenstates of the Schrödinger equation for the crystal. As a result, a Bloch electron with a certain momentum can make a transition to a Bloch state with a different momentum in effect, being scattered. A more physical analogy is that of many balls (the electrons) all rolling in one direction across a level parking lot. If an earthquake suddenly strikes, causing the surface of the parking lot to ripple, the gravitational potential energy of the balls can suddenly be altered, producing a change in their kinetic energy and causing them to move in many different directions. In the same way, phonons alter the local band structure of a crystal, causing changes in the carriers' potential energy. As the temperature increases, the phonon population also increases, leading to an increase in the scattering rate. In all but the most highly doped

crystals, phonon scattering is the dominant mechanism limiting the electron mobility at temperatures greater than roughly 100 K.

Charged impurities are another important scattering mechanism. When an electron is released in a semiconductor, either from an impurity level or from the valence band, it leaves behind a positvely charged center. The Coulomb fields of these positive centers are distributed randomly throughout the crystal, creating random electric fields which can interact with the conduction electrons and scatter them elastically. Although these fields are screened to some extent by the conduction electrons, they still play an important role in limiting the mobility at low temperatures, where phonon scattering becomes less effective.

A third factor which limits the electron mobilities in the III-nitride semiconductors is interaction of the electrons with the short range potentials of defects in these materials. There is an increasing body of evidence that the large electron concentrations typically found in III-nitride samples originate from a native defect or chemical impurity which has its first ionization level resonant with the conduction band (Maruska 1969, Pankove 1990). Experiments have shown that upon the application of hydrostatic pressure, which causes the conduction band of GaN to move upwards with respect to the valence band, this donor level can be pushed into the forbidden gap, as evidenced by a reduction in the number of free carriers (Perlin 1995, Wetzel 1996). Unlike the hydrogenic donors such as Si in GaAs or P in Si, whose electronic wavefunctions extend over many thousands of lattice points, the resonant donors in GaN are expected to have wavefunctions which are localized about the donor center. The potentials associated with these donors are not simple screened Coulomb potentials, but the superposition of a Coulomb potential with another potential which is only a few angstroms in extent. The effects of this type of scattering have been calculated previously in narrow gap semiconductors (Litwin-Staszewska 1971).

The effects of alloy disorder have also been taken into consideration when calculating mobilities in alloys of the III-nitride semiconductors. In elemental and binary semiconductors, the electrons see a more or less perfectly periodic potential and the

wavefunctions are Bloch waves. These solutions are eigenstates of the crystal momentum. However, alloys such as $Al_xGa_{1-x}N$ have Al and Ga atoms arranged randomly on the Ga sublattice and the potential experienced by conduction electron is far from periodic. As GaN and AlN have different band gaps, the band gap of an AlGaN alloy varies on a microscopic scale and this variation can cause scattering of conduction electrons in much the same way as acoustic phonons through the deformation potential.

Finally, although the GaN grown today is known to contain high densities of extended structural defects such as stacking faults and dislocations, no effort is made here to calculate their effect on the electron mobility. There are two reasons for this. The first is that since the size and exact type of such defects can vary over a wide range, it is extremely difficult to estimate the scattering times associated with such structures without having to resort to the introduction of unknown parameters which could be adjusted to fit anything. Second and more important, it has been found that such defects are not the primary reasons for the low electron mobilities generally found in GaN (Hwang 1997). Thus, our results which neglect their contribution should still be reasonably accurate.

1.2 Calculating the Mobility

In this section, the actual expressions used for the relaxation times due to the various scattering mechanisms are given, as well as a brief description of the method used in calculating the mobility.

For acoustic phonon scattering through the deformation potential, the relaxation time has been calculated by Shockley (1950) to be

$$\frac{1}{\tau_{def}} = 4.167\times10^{19}\frac{E_1^2}{c_l}\left(m_{eff}T\right)^{3/2}x^{1/2} \qquad (1.5)$$

where E_1 is the deformation potential, T the temperature, c_l the longitudinal elastic constant, m_{eff} the electron effective mass in units of the free electron mass, and x is the energy in units of k_bT.

The acoustic mode piezoelectric scattering time for zincblende semiconductors has been calculated by Zook (1964) as

$$\frac{1}{\tau_{pie}} = 1.052\times10^{7}\times h_{14}\left(\frac{3}{c_l}+\frac{4}{c_t}\right)\left(m_{eff}\frac{T}{x}\right)^{1/2} \tag{1.6}$$

where h_{14} is the only non-zero element of the piezoelectric tensor and c_l and c_t are longitudinal and transverse elastic constants.

For screened ionized impurities, the characteristic scattering time is (Dingle 1955)

$$\frac{1}{\tau_{Coul}} = 2.415\frac{N_i}{\sqrt{m_{eff}}\,\varepsilon^2}\left(\ln(1+\xi)-\frac{\xi}{1+\xi}\right)(x\,T)^{-3/2} \tag{1.7}$$

where

$$\xi = \frac{8m^*R^2k_bT}{\hbar^2}\,x \tag{1.8}$$

In these expressions, N_i is the concentration of ionized impurities, ε is the dielectric constant, and R is a screening length which is given by

$$\frac{1}{R^2} = \frac{16\pi^2e^2\,m_{eff}^{3/2}(2\pi k_BT)^{1/2}}{\varepsilon h^3}F_{-1/2}\left(\frac{E_F}{k_BT}\right) \tag{1.9}$$

where $F_{-1/2}$ is the standard Fermi integral.

When incorporating the effect of short range potentials into the scattering calculations, two new terms are introduced. The reason for this is that the total potential of an ionized and localized defect can be separated into two parts,

$$V(r) = V_{Coul}(r) + V_{SR}(r) \tag{1.10}$$

where V_{Coul} is the long range screened Coulomb potential and V_{SR} is the short range potential. Both potentials are localized on the same center. The electron scattering rate is proportional to the square of the matrix element of the potential (1.10) and so consists of three terms; the standard Coulomb term proportional to the square of the matrix element $< S \mid V_{Coul} \mid S >$, a short range term proportional to $< S \mid V_{SR} \mid S >^2$, and an interference term proportional to $< S \mid V_{Coul} \mid S> <S \mid V_{SR} \mid S >$, where $|S >$ is the conduction band Bloch amplitude. The inverse relaxation time of the short range term is (Litwin-Staszewska 1971)

$$\frac{1}{\tau_{SR}} = \frac{N\, m_{eff}\, k\, A}{\pi\, \hbar^3} \tag{1.11}$$

and that for the interference term is

$$\frac{1}{\tau_{int}} = \frac{N\, e^2\, m_{eff}\, A\, F_{int}}{\hbar^3\, \varepsilon\, k}. \tag{1.12}$$

Here, A is the matrix element of the short range potential $<S \mid V_{SR} \mid S >$, N is the concentration of resonant defects and k is the electron wavevector.

The screening factor in the interference term is

$$F_{int} = 4\left[1 - \frac{1}{\eta}\ln(1+\eta)\right] \tag{1.13}$$

where $\eta = (2kR)^2$ and the screening length R is the same as the one introduced for Coulomb scattering.

We can now compare the effectiveness of Coulomb scattering with scattering by short range potentials. Noting the dependencies of the scattering rates on electron energy, one sees that for a highly degenerate electron gas with concentration n, $k \propto n^{1/3}$, $R \propto n^{-1/6}$ and to a good approximation, $F_{int} = 4$. Assuming that in unintentionally doped samples the electron concentration is proportional to the defect concentration, one finds from Eqs. (1.10) and (1.11), that $1/\tau_{SR} \propto n^{4/3}$ and $1/\tau_{int} \propto n^{2/3}$. It can also be shown for scattering by ordinary coulomb centers that $1/\tau_{Coul} \propto n^{1/3}$. Hence, short range scattering can be expected to be more effective than Coulomb scattering in samples with a high free carrier concentration.

Alloy disorder is a well established scattering mechanism with an inverse relaxation time of (Makowski 1973)

$$\frac{1}{\tau_{all}} = \frac{N_{sites}\, k\, m_{eff}\, (1-y)\, y\, V^2}{\pi\, \hbar^3} \tag{1.14}$$

where N_{sites} is the number of primitive cells per volume, V is the conduction band offset between the two corresponding binary compounds (AlN and GaN for $Al_yGa_{1-y}N$), and y is the alloy fraction.

Incorporating the effets of optical phonons poses a special problem. If it were possible to calculate individual energy dependent scattering times due to each of the above mechanisms, a net scattering time could be calculated through Mathiessen's rule (Böer 1990)

$$\frac{1}{\tau_{tot}} = \frac{1}{\tau_{def}} + \frac{1}{\tau_{pie}} + \frac{1}{\tau_{Coul}} + \frac{1}{\tau_{all}} + \frac{1}{\tau_{SR}} + \frac{1}{\tau_{int}} + \frac{1}{\tau_{op}} \tag{1.15}$$

and the mobility could be calculated by integrating the relaxation times over all electron energies. However, due to the nature of scattering by polar optical phonons, this procedure is invalid for our purposes. We recall that eq. (1.1) was based on the assumption that the collisions were elastic, or very nearly so. For Coulomb scattering and scattering by acoustic phonons, this is always the case except at exceptionally low temperatures since the energies of acoustic phonons are negligible compared to typical electron energies and scattering by central potentials is elastic. Optical phonons, however, have energies which are of the same order of magnitude as those of typical conduction electrons, making scattering highly inelastic. In this case $[df/dt]_{COLL}$ for any particular energy will depend on the properties of electrons with very different energies and equation (1) will not be correct. Thus, for optical phonons, no simple relaxation time τ can be defined. Instead, the Boltzmann equation must be solved directly to find the mobility. The calculations involved in this task are straightforward, but rather tedious and can be found in (Howarth 1953), the net result being that the mobility can be expressed as

$$\mu = 308.6\left[\left(\frac{1}{\varepsilon_\infty}-\frac{1}{\varepsilon_0}\right) m_{eff}^{\,3/2} T^{1/2}\left(\frac{E_{op}}{k_B T}\right) F_{1/2}\left(\frac{E_F}{k_B T}\right)\right]^{-1}\left(\frac{D_{3/2,3/2}}{D}\right). \tag{1.16}$$

where ε_∞ is the high frequency dielectric constant, E_F is the Fermi energy and E_{op} is the optical phonon energy. The D's in this equation are the determinants

$$D_{3/2,3/2} = \begin{vmatrix} 0 & \beta_0^{(3/2)} & \beta_1^{(3/2)} & \cdots \\ \beta_0^{(3/2)} & d_{00} & d_{01} & \cdots \\ \beta_1^{(3/2)} & d_{10} & d_{11} & \cdots \\ \cdot & \cdot & \cdot & \\ \cdot & \cdot & \cdot & \end{vmatrix} \tag{1.17}$$

and

$$D = \begin{vmatrix} d_{00} & d_{01} & \cdots \\ d_{10} & d_{11} & \cdots \\ \cdot & \cdot & \\ \cdot & \cdot & \end{vmatrix} \tag{1.18}$$

where

$$\beta_r^{(3/2)} = \int_0^\infty E^{3/2}\, \Phi_r \frac{\partial f_0}{\partial E}\, dE \tag{1.19}$$

$$d_{rs} = \int_0^\infty \Phi_r\, L(\Phi_s) \frac{\partial f_0}{\partial E}\, dE \tag{1.20}$$

In order to incorporate all of the scattering mechanisms described so far, including optical phonons, the operator L is taken to be

$$L(\Phi_s) = L_{op}(\Phi_s) + \frac{\Phi_s}{A}\left(\frac{1}{\tau_{def}} + \frac{1}{\tau_{pie}} + \frac{1}{\tau_{Coul}} + \frac{1}{\tau_{SR}} + \frac{1}{\tau_{\text{int}}} + \frac{1}{\tau_{all}}\right) \tag{1.21}$$

where

$$A = 4.768 \times 10^{28} \frac{E_{op}}{k_B T}\left(\frac{m_{eff}}{T}\right)^{1/2}\left(\frac{1}{\varepsilon_\infty} - \frac{1}{\varepsilon_0}\right) x^{-3/2} \tag{1.22}$$

and L_{op} is a complicated expression that can be found in Howarth (1953).

Using this formalism, mobilities in a many III-V semiconductors have been calculated with good agreement with experimental results (Ehrenreich 1960, Lehochky 1974, Walukiewicz 1979). We now apply these results to calculating mobilities in III-nitride materials.

2. Mobilities in the Bulk III-Nitrides

2.1 GaN

Much research is currently focused on the growth of high quality epitaxial films of GaN. Because of its wide direct bandgap, tunable between 1.9 eV and 6.2 eV by alloying with In or Al, GaN is considered a prime candidate for the manufacture of visible to ultraviolet optoelectronic devices (Strite 1992) and III-V compatible dielectrics (Li 1986). In recent years there have been several reports on calculations of electron mobilities in III-V nitrides (Molnar 1993, Chin 1994, Rode 1973). Those calculations were aimed at an assessment of the relative contributions of the standard scattering mechanisms and the establishment of the phonon mobility limits in these materials. Comparison of the calculations with experimental results was rather difficult as undoped GaN and InN typically exhibit large electron concentrations and mobilities are not phonon limited.

Using the method outlined in the previous section to calculate electron moblilities, experimental data obtained for GaN layers which were intentionally doped with either Si or Ge are examined first. Since the short range potential scattering is associated only with localized defects, one does not expect it to play an important role in these materials. In figure 2-1 the concentration dependent mobility calculated for GaN doped with hydrogenic donors is plotted for a few different values of the compensation ratio. The materials parameters used in the calculations are listed in Table 2-1. It is seen that in GaN intentionally doped with either Si or Ge the electron mobility shows a very weak dependence on electron concentration even in the high concentration region (Nakamura 1992, Wickenden 1993, Gaskill 1995, Rowland 1995). As has been discussed above, such a dependence is consistent with scattering by Coulomb potentials only. The experimental data in figure 2-1 can be satisfactorily explained assuming the presence of compensating acceptors with a compensation ratio of roughly $N_A/N_D \approx 0.4$. Similar

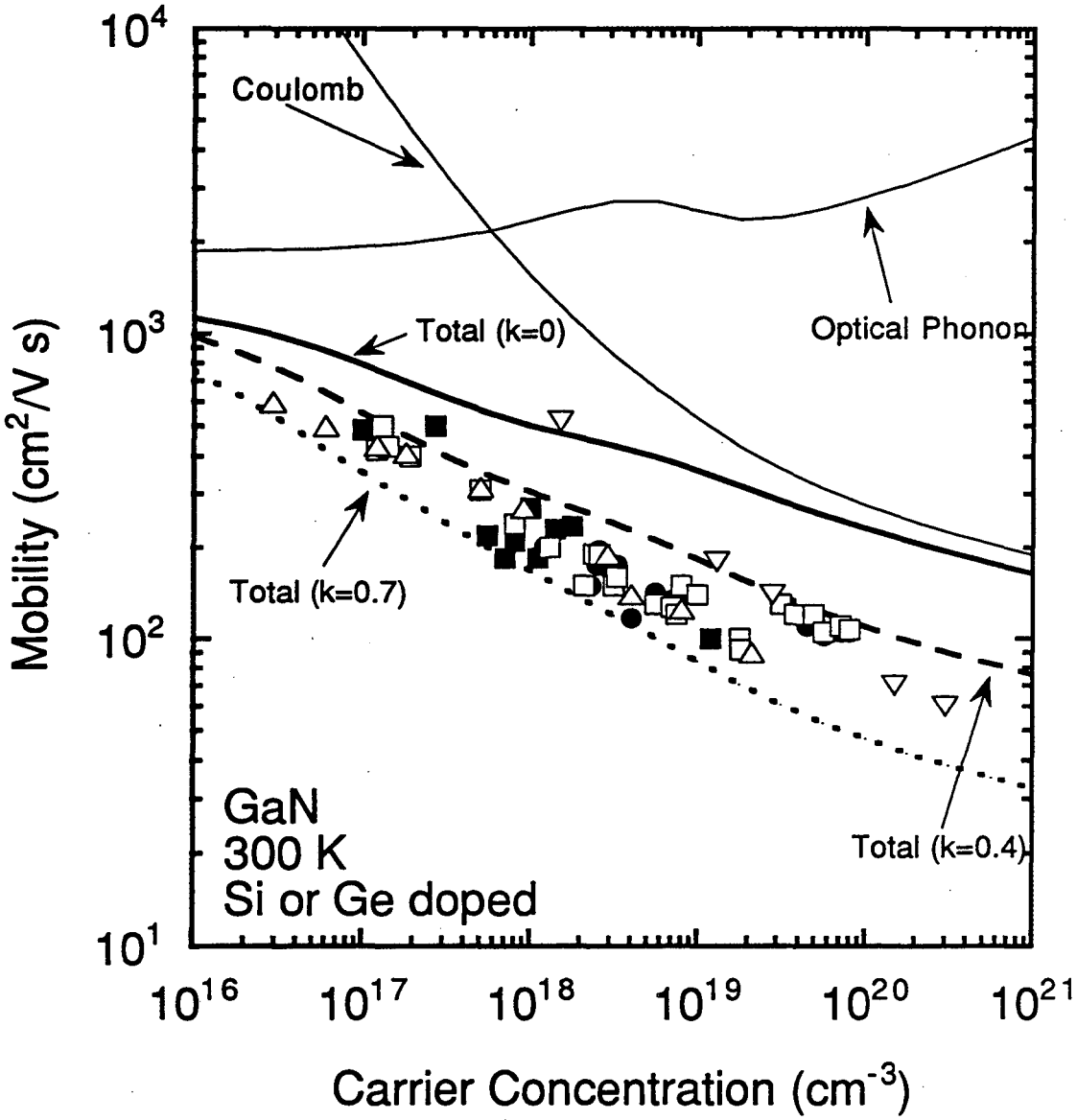


Figure 2-1. Plot of the best mobilities obtained for GaN intentionally doped with Si or Ge. The curves are calculated mobilities for various compensation ratios.

Parameter	GaN	InN
Density (g/cm^3)	6.1	6.81
ε_0	9.5	15.3
ε_∞	5.35	8.4
LO phonon energy (meV)	90.5	76
m^*/m_0 (at band minimum)	0.21	0.11
Lattice parameter a_0 (Å)	4.52	4.98
Acoustic phonon vel. (cm/s)	6.6×10^5	5.07×10^5
Piezoelectric constant (V^2/dyn)	13384	13384
Deformation potential (eV)	8.5	2.5
Elastic constants: c_L (dyn/cm^2)	2.66×10^{12}	1.75×10^{12}
c_T (dyn/cm^2)	6.2×10^{11}	2.2×10^{11}
Matrix element of short range potential A ($erg\ cm^3$)	3.7×10^{-34}	2.5×10^{-34}

Table 2-1. GaN and InN parameters used in the calculations presented here.

values of the compensation ratio are frequently encountered in other intentionally doped III-V semiconductors (Walukiewicz 1979).

In figure 2-2, calculated mobilities are plotted along with the best mobility data found in the literature for layers which were grown without any intentional doping (Sasaki 1987, Nakamura 1991, Sun 1993, Gaskill 1995). Here the conduction electrons are assumed to originate from resonant defects rather than hydrogenic donors. The total mobilities are plotted as a function of concentration at room temperature along with the component mobilities due to the most important individual scattering mechanisms. As can be seen, the electron mobilities of unintentionally doped GaN layers is lower than those of layers which were intentionally doped for carrier concentrations in excess of 10^{19} cm^{-3}. Although it is possible that these layers have a much greater compensation ratio than the ones which were intentionally doped, another possible explanation lies in scattering by the short range potentials associated with the highly localized native defects. At defect concentrations greater than a few times 10^{19} cm^{-3}, scattering from short range potentials due to interference and short range scattering becomes the dominant mechanism in limiting the mobility. The component mobilities in fig. 2-2 for short range and interference scattering were calculated assuming a value of $A = 3.7 \times 10^{-34}$ erg cm^3 for the short range potential as found in eq. (1.11).

So far, no clear case can be made for the responsibility of either compensation or short range scattering effects for the very low mobilities in GaN at high doping concentration. Experiments have been performed showing that n-type GaN is rather heavily compensated [Yi 1996], even at high electron densities, with theoretical calculations indicating that a triply charged Ga vacancy, whose formation energy is quite low in highly n-type material, may be the responsible for the compensation [Neugebauer 1996]. On the other hand, a large number of experiments also indicate the presence of a resonant donor in unintentionally doped n-type GaN [Perlin 1995, Wetzel 1996]. On the application of hydrostatic pressure which causes the conduction band edge of GaN to

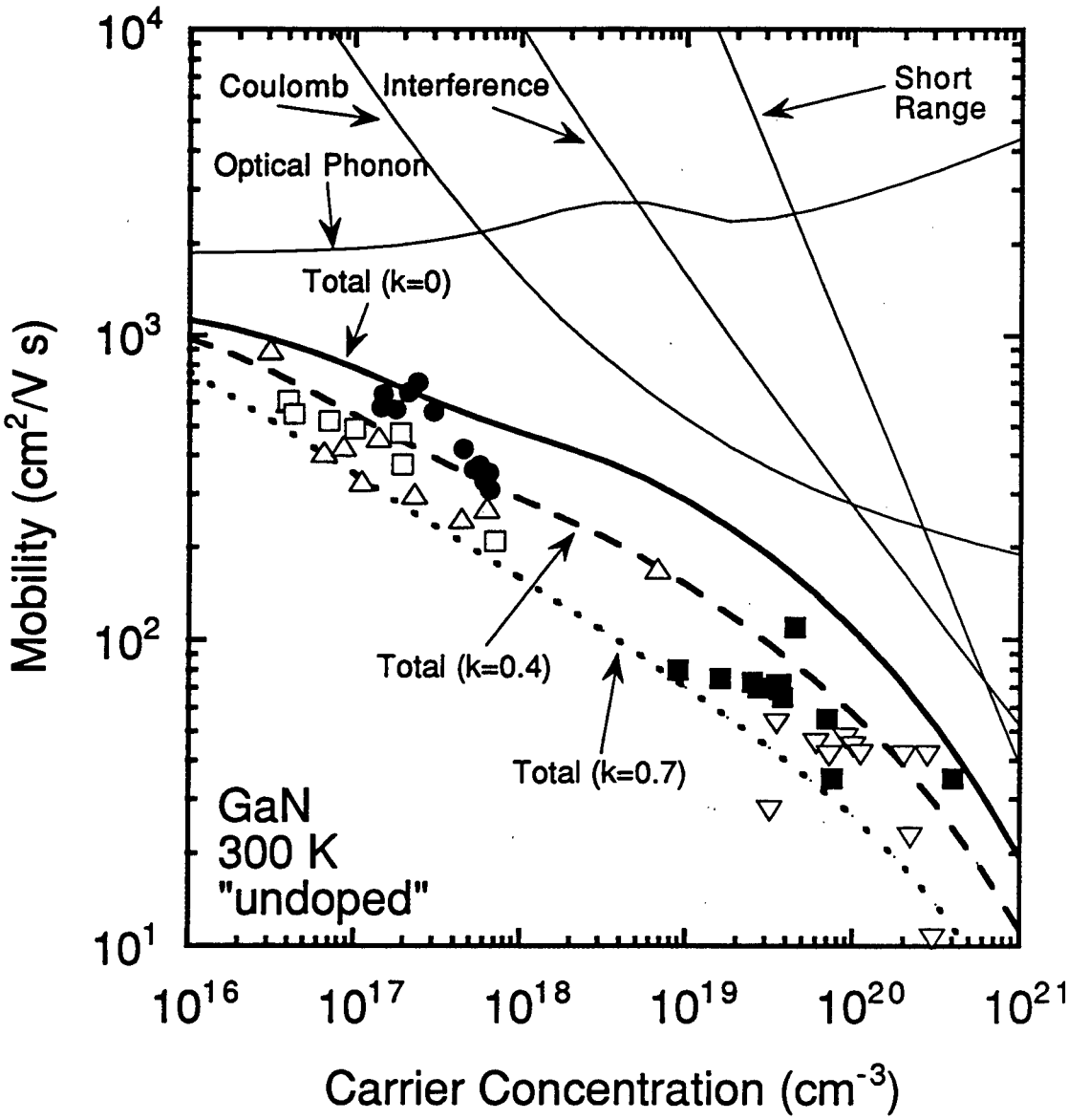


Figure 2-2. Plot of the best mobilities obtained for nominally undoped GaN. The curves are calculated values, showing the total mobilities as well as the individual contributions from optical phonons, Coulomb scattering, and scattering by the short range potential.

increase in energy with respect to the valence band, the electron concentration has been observed to decrease dramatically, indicating that the resonant donor level has been pushed into the forbidden gap. Of course, due to the poor structural quality of many GaN films, arising from lattice mismatch with the substrates, structural defects may also play an important role in the low mobilities which have been measured. Both the quality and consistency of GaN growth results must be improved before a definite conclusion can be made.

2.2 InN

Another III-Nitride semiconductor that exhibits large concentrations of free electrons is InN. In typical as-grown, unintentionally doped InN, electron concentrations are close to 10^{21} cm^{-3} and there is a very limited amount of experimental data for samples with lower electron concentrations. It is quite likely that in this case, the electrons also originate from highly localized donor defects which have an energy level located well above the conduction band edge. One can therefore expect that scattering by short range potentials should play a significant role in this material.

Calculated room temperature electron mobilities in unintentionally doped InN are shown in figure 2-3 along with the best available experimental data (Hovel 1972, Tansley 1984, Wisk 1993, Yamamoto 1994, Abernathy 1995). Because of the relatively small energy gap of E_G =1.9 eV, at large carrier concentrations (> 10^{19} cm^{-3}) the Fermi energy becomes comparable to the energy gap. In such cases it is necessary to incorporate the effects of nonparabolicity on the density of states effective mass. In a simple approximation, the energy dependent effective mass is given by,

$$m^* = m_0^* \left(1 + 2\frac{E}{E_{gap}} \right) \tag{2.1}$$

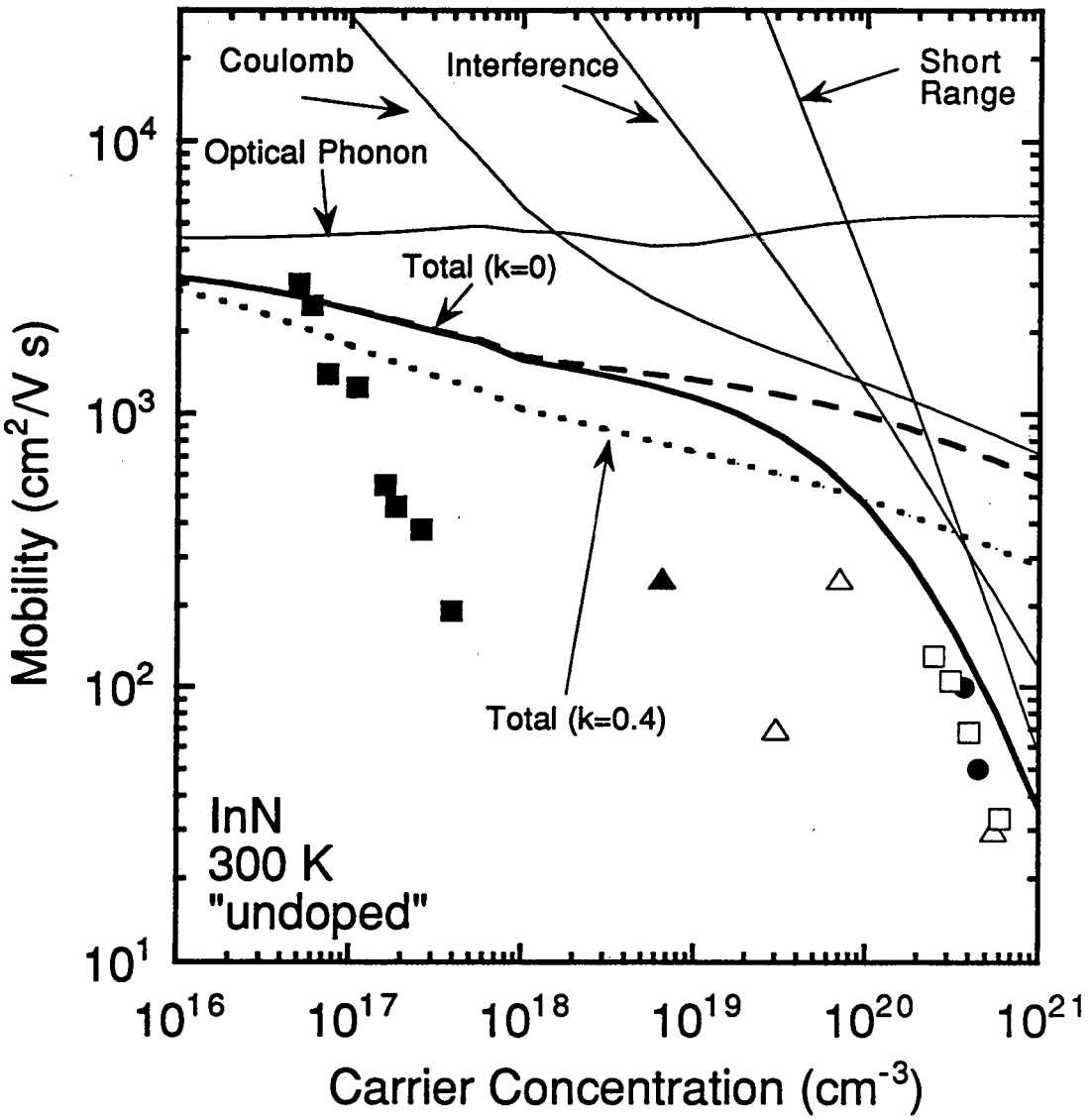


Figure 2-3. Theoretically calculated mobilities for InN showing contributions from optical phonons, Coulomb scattering, and scattering from the short range potentials. The dashed lines show calculated mobilities assuming no scattering by short range potentials. The best experimentally obtained mobility values are also plotted.

where m_0^* is the conduction band edge effective mass and E is the electron energy.

The results in figure 2-3 show that, as in GaN, optical phonon scattering limits the total mobility in InN exhibiting low carrier concentrations at room temperature. With increasing electron and thus defect concentration, the scattering by Coulomb potentials becomes the dominant mechanism for limiting the mobility. Finally, at very high defect concentrations the mobility is entirely determined by short range and interference scattering. Although the amount of experimental data on electron mobilities in InN is small, it is quite clear that the strong concentration dependence observed at very high electron densities cannot be explained by simple Coulomb scattering. The rapid decrease of the mobility for n larger than 10^{20} cm^{-3} can easily be explained, however, by scattering by native defects with a short range parameter $A = 2.5 \times 10^{-34}$ erg cm^3.

Unfortunately, there is no data available on intentionally doped InN. From the theoretical curves shown in figure 2-3, one finds that at high electron concentrations the mobility in InN doped with hydrogenic donors should be at least one order of magnitude higher than in the unintentionally doped material. Therefore it is expected that doping with shallow hydrogenic donors should greatly improve the electrical characteristics of InN.

2.3 $Al_xGa_{1-x}N$ alloys

Due to their importance in the production of GaN-based devices, such as high electron mobility transistors (HEMTs), the alloys of the III-nitrides, especially AlGaN, have also been the focus of much recent research. With a few simple modifications, it is easy to extend our mobility calculations to model AlGaN alloys as well. Experiments have shown that the electron mobilities in AlGaN alloys remain relatively constant upon varying of the alloy concentration (Yoshida 1982, Koide 1986). In a few of the data sets,

a temporary decrease in the mobility is also observed for a certain range of Al fraction. We would like to be able to explain this trend qualitatively.

In addition, AlGaN is a particularly interesting alloy as it also affords us the chance to study the effects of the resonant donor on the mobility and free carrier concentrations as a function of Al content. Since the bandgap of AlN is greater than that of GaN, the conduction band energy increases with respect to the valence band as the alloying concentration increases. In contrast, since the resonant defect is highly localized and does not interact much with the crystal, its energy level remains relatively constant. This situation is illustrated schematically in figure 2-4. As the energy of the conduction band approaches that of the defect level, it becomes energetically less favorable for defects to be ionized and a "freeze out" of carriers originating from these defects is observed (Yoshida 1982, Khan 1983, Koide 1986, Zhang 1995). This is the same reduction in carrier concentration observed in GaN under large hydrostatic pressures (Perlin 1995, Wetzel 1996). In addition, as the defect level falls below the Fermi level due to the aforementioned carrier "freeze out," resonant scattering becomes a possibility. The origin of this mechanism is the scattering of electrons from conduction band states to bound states of the defect and vice versa. Experimentally, this is observed as a decrease in the mobility in those samples in which the Fermi and defect levels are close in energy. Resonant scattering is important only when the defect and Fermi levels are approximately equal in energy since in that situation, the maximum number of conduction electrons are available to be scattered elastically into bound states of the defect and there are many empty conduction band states available for the reverse process, resulting in a decrease of the electron mobility (Raikh 1986). When the energy of the defect level is far above or far below the Fermi level, there are either too few electrons with sufficient energy to be affected by this scattering mode or too few empty conduction band states for defect bound electrons to scatter into and resonant scattering ceases to affect the electron mobility appreciably (see figure 2-5).

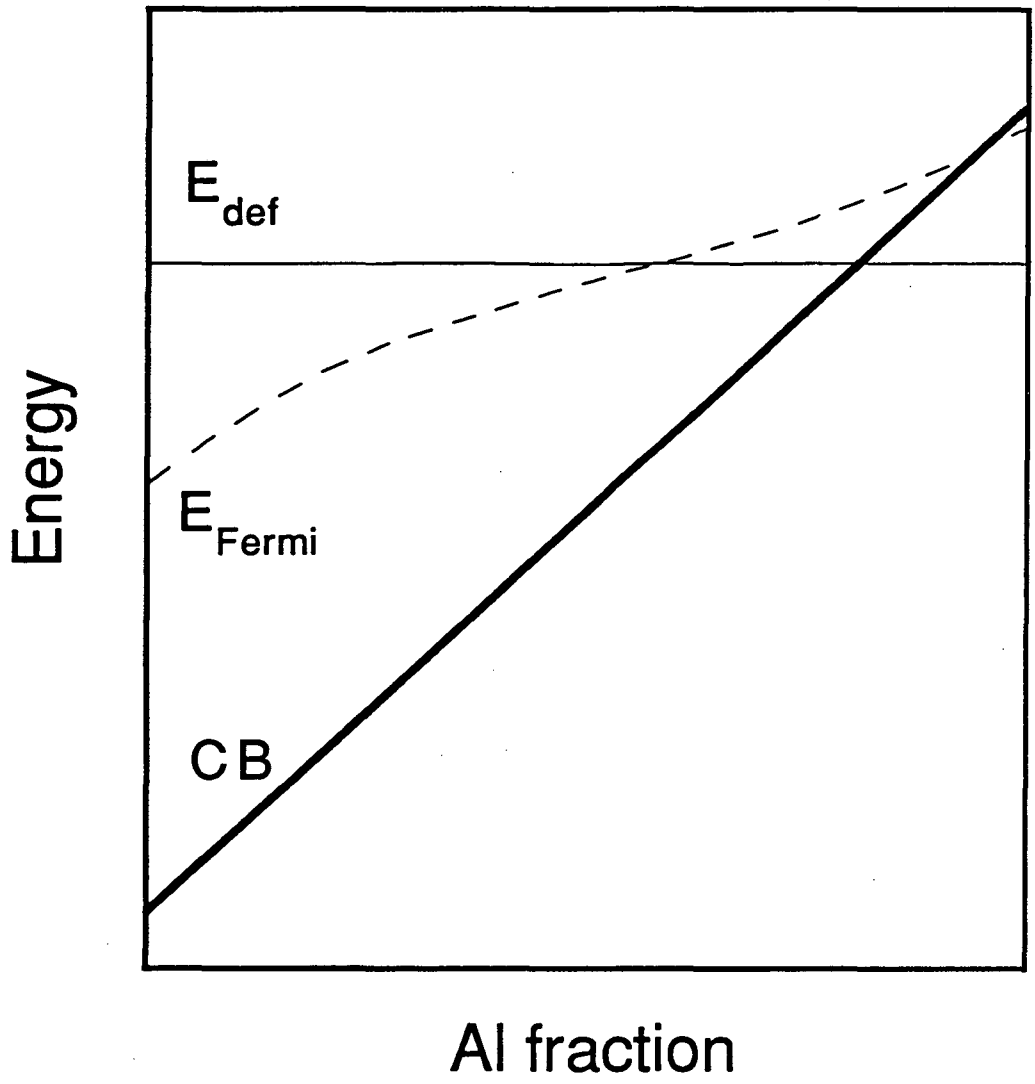


Figure 2-4. Schematic diagram of movement of the conduction band. Fermi energy, and resonant defect level in GaN as a function of alloying with Al.

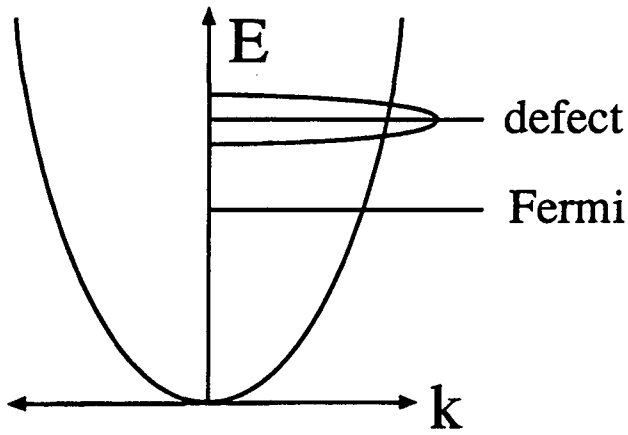


Figure 2-5a. When the defect energy is far above the Fermi level, few electrons have enough energy to be elastically scattered into the defect level. The defect level is shown broadened schematically due to the finite lifetime of an electron in that state.

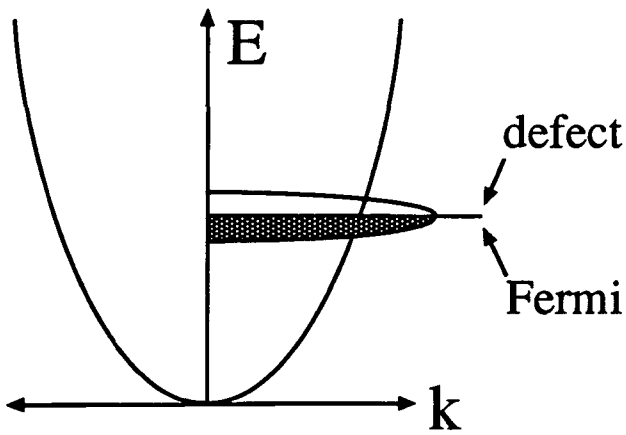


Figure 2-5b. When the defect and Fermi energies coincide, electrons can be freely scattered into and out of defect states.

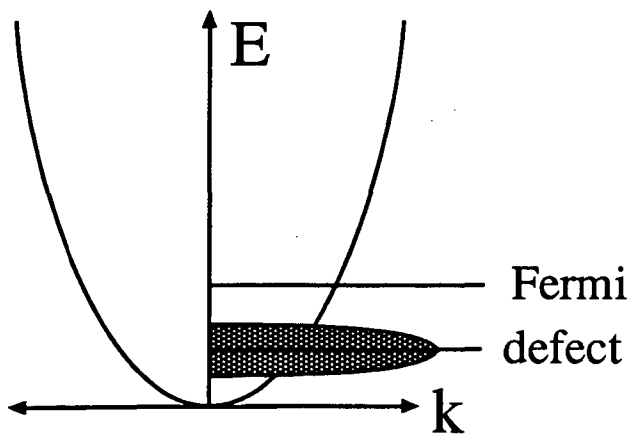


Figure 2-5c. At this point, there are no longer any empty defect states for conduction electrons to scatter into nor any empty conduction band states into which defect bound electrons can scatter elastically.

In order to include effects of the resonant donor in the mobility calculations, Fermi's golden rule can be used to calculate the relaxation time for the elastic scattering of an electron from a free (conduction band) state to a bound state of the resonant donor. Also taken into account the energy broadening of the resonant level due to the finite lifetime of the electron in the bound state. Since the effect of this scattering mechanism is to completely remove one electron's contribution to the current, the τ calculated by Fermi's golden rule in this case is equivalent to the τ representing the decay of electron current and thus no factor of $1-\cos\theta$ is required.

Assuming that the defect density of states has a Lorentzian shape, we can write the characteristic relaxation time for this process can be written as (Wilamowski 1990)

$$\frac{1}{\tau_{res}} = \frac{\pi \hbar n_{def}}{m^* k_{def}} \left[\frac{\Gamma^2}{(E_{def} - E)^2 + (\Gamma/2)^2} \right] \tag{2.2}$$

where E is the electron energy, m* is the effective mass, and k_{def} is the wavevector of an electron which has a kinetic energy equal to the energy of the defect relative to the conducton band. E_{def} and n_{def} are the defect energy and concentration, respectively. The characteristic broadening of the donor level Γ is given by

$$\Gamma = \frac{m^* k_{def}}{\pi \hbar^2} V \langle CB|H|def \rangle^2_{k=k_{def}} \tag{2.3}$$

In principle, the matrix element involving the conduction band and donor state electron wavefunctions and interaction Hamiltonian can be obtained through a first principles calculation. However, to get a general idea of the magnitude of Γ, this quantity was used as a fitting parameter in the plots that follow.

In attempting to fit theoretical curves for the "freeze out" of the carriers to experimentally measured points from four different groups, the energy of the defect level above the conduction band was estimated to be 475 meV. These fits are shown in figures 2-6a through 2-6d. The defect concentration was assumed to be constant for all samples grown by each group and the defect energy and broadening were estimated by fitting the theoretical curves to the experimental points. Since each data point represents an independent sample, there is no *a priori* reason to believe that the total defect concentration for each group's data was necessarily constant as assumed. However, since each set of samples were grown under similar conditions, we feel that our assumption was justified as a starting point. As can be seen, the fits are not perfect and the agreement with experiment varies.

In calculating the "freeze out" curves shown, two additional effects other than the variation of the conduction band energy with alloying content were considered. The first was the effect of electron-electron interaction (Walukiewicz 1990) . Since free electrons have the ability to arrange themselves in ways to minimize their repulsive energy while electrons bound to defect centers do not, this has the effect of lowering the energy of the conduction band by (Berggren 1981)

$$\Delta E_{CB} = \frac{-2.88\, e^2\, n^{1/3}}{\pi^{1/3}\, \varepsilon} \tag{2.4}$$

where ε is the static dielectric constant and n is the number of free carriers, so that the actual energy of the defect level relative to the conduction band is somewhat larger than 475 meV depending on the free carrier concentration. The behavior of the conduction band, Fermi level, and defect level calculated for each case is plotted in figures 2-7a through 2-7d. Experimentally, estimates of the defect energy level range from 400 meV (Wetzel 1996) to 800 meV (Perlin 1995) above the GaN conduction band edge as

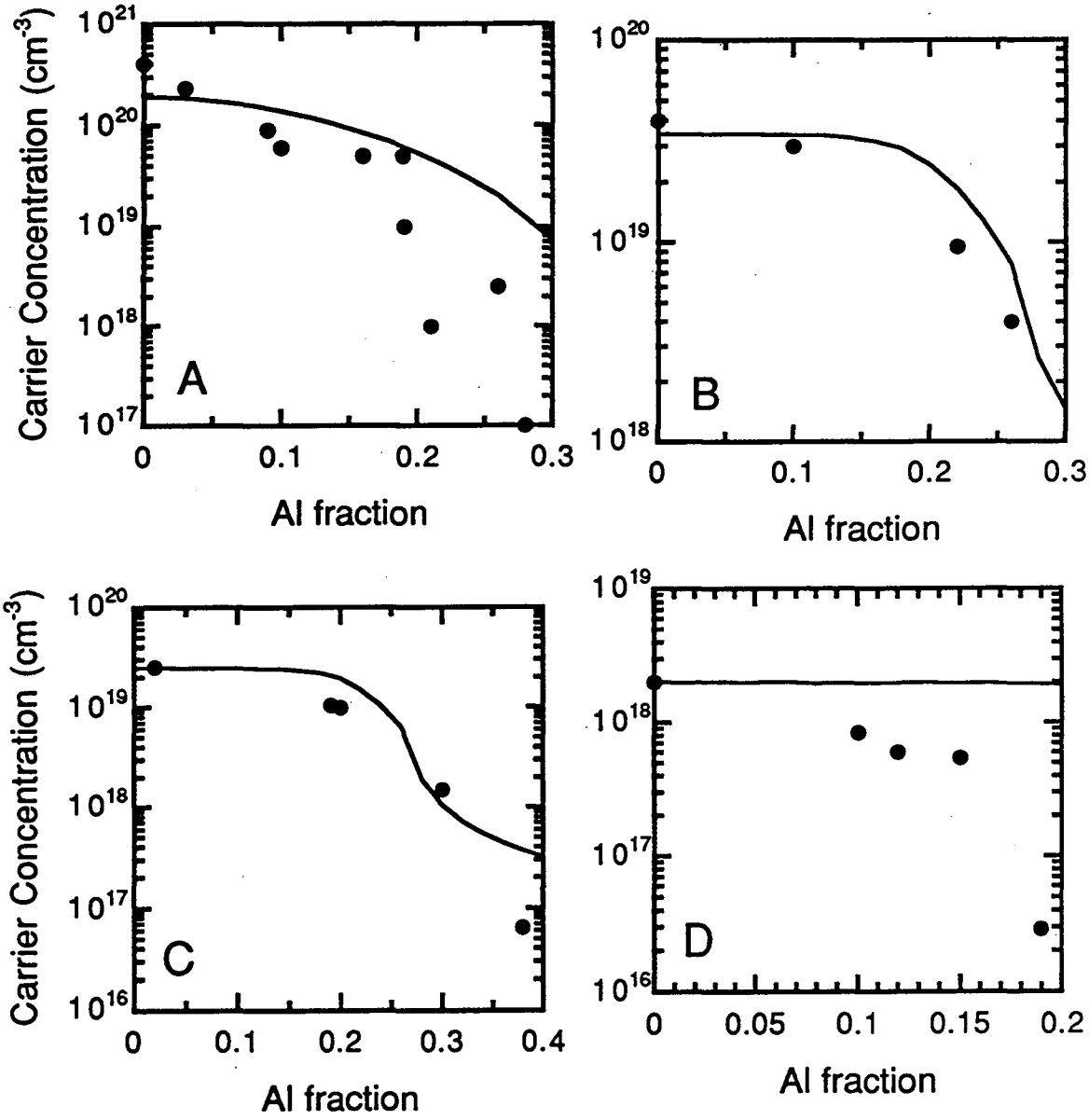


Figure 2-6. Plots of carrier concentration vs. Al fraction in $Al_xGa_{1-x}N$ for four sets of data. Curves are theoretical fits assuming a constant defect concentration and a defect level 475 meV above an unperturbed GaN conduction band. References are A (Yoshida 1982), B (Koide 1986), C (Khan 1983), and D (Zhang 1995).

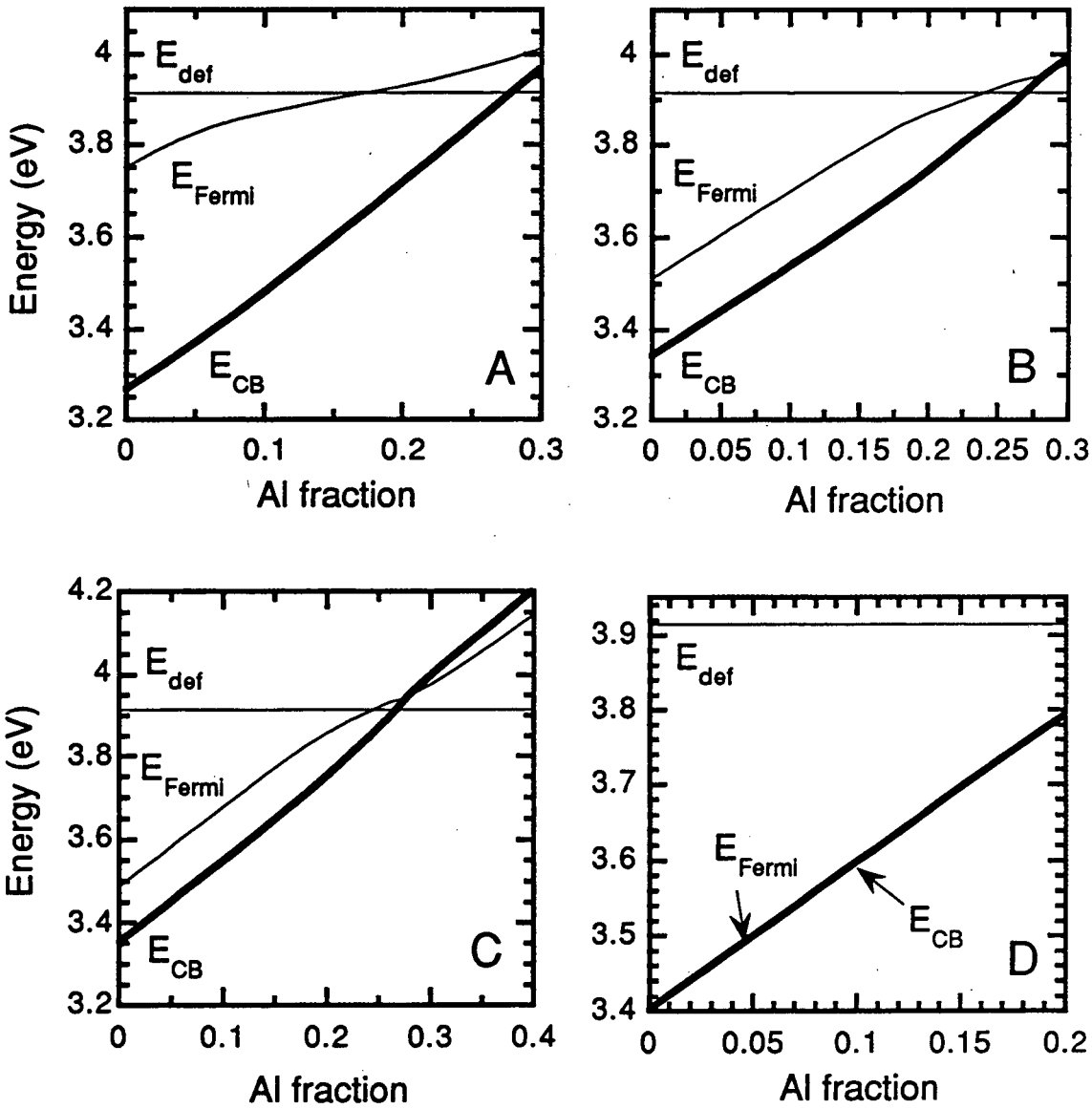


Figure 2-7. Plot of defect, Fermi, and conduction band energies as a function of Al fraction in $Al_xGa_{1-x}N$. Each plot corresponds to the carrier concentrations from the references in figure 2-6.

obtained through measurements which used hydrostatic pressure to vary the GaN energy gap. It is clear that more work needs to be done before a reliable determination can be made.

Finally, in figures 2-8a through 2-8d, the calculated mobilities are plotted together with experimentally obtained data points (Yoshida 1982, Khan 1983, Koide 1986, Zhang 1995). In two out of the four sets of data (Yoshida 1982, Koide 1986), the mobility can be seen to remain roughly constant except for a slight drop near Al alloying fractions of 0.2. Comparing this data with theory, it can be seen that were it not for resonant scattering, the mobilities would remain very nearly constant across the entire measured range of alloying concentrations. Resonant scattering, however, causes a drop in the mobility at the point where the Fermi level and defect level cross. Due to the almost order of magnitude difference in defect concentrations between these two sets of data, the crossing takes place at slightly different alloying concentration. As mentioned above, the defect broadening parameter Γ was estimated from the width of the dip in mobility. The best agreement with experimental data was obtained for a Γ of 25 meV.

The remaining two sets of data (Khan 1986, Zhang 1995) show only a steady decrease in the mobility as the Al fraction is increased. This behavior can be qualitatively explained by adjusting the number of compensating acceptors. As the free carrier concentration decreases due to "freeze out" of the resonant defects, the ability of the remaining conduction electrons to screen the random Coulomb fields of the impurities decreases, resulting in a drop in the mobility.

Based on such limited data, it would be premature to make any solid conclusions regarding the resonant donor in GaN. Our model is certainly plausible and is able to explain the qualitative trends observed. However, since the effect of alloying with Al on the incorporation of defects, whether extrinsic or intrinsic is unknown, it is possible that the AlGaN alloys simply contain fewer impurity atoms than the unalloyed GaN samples.

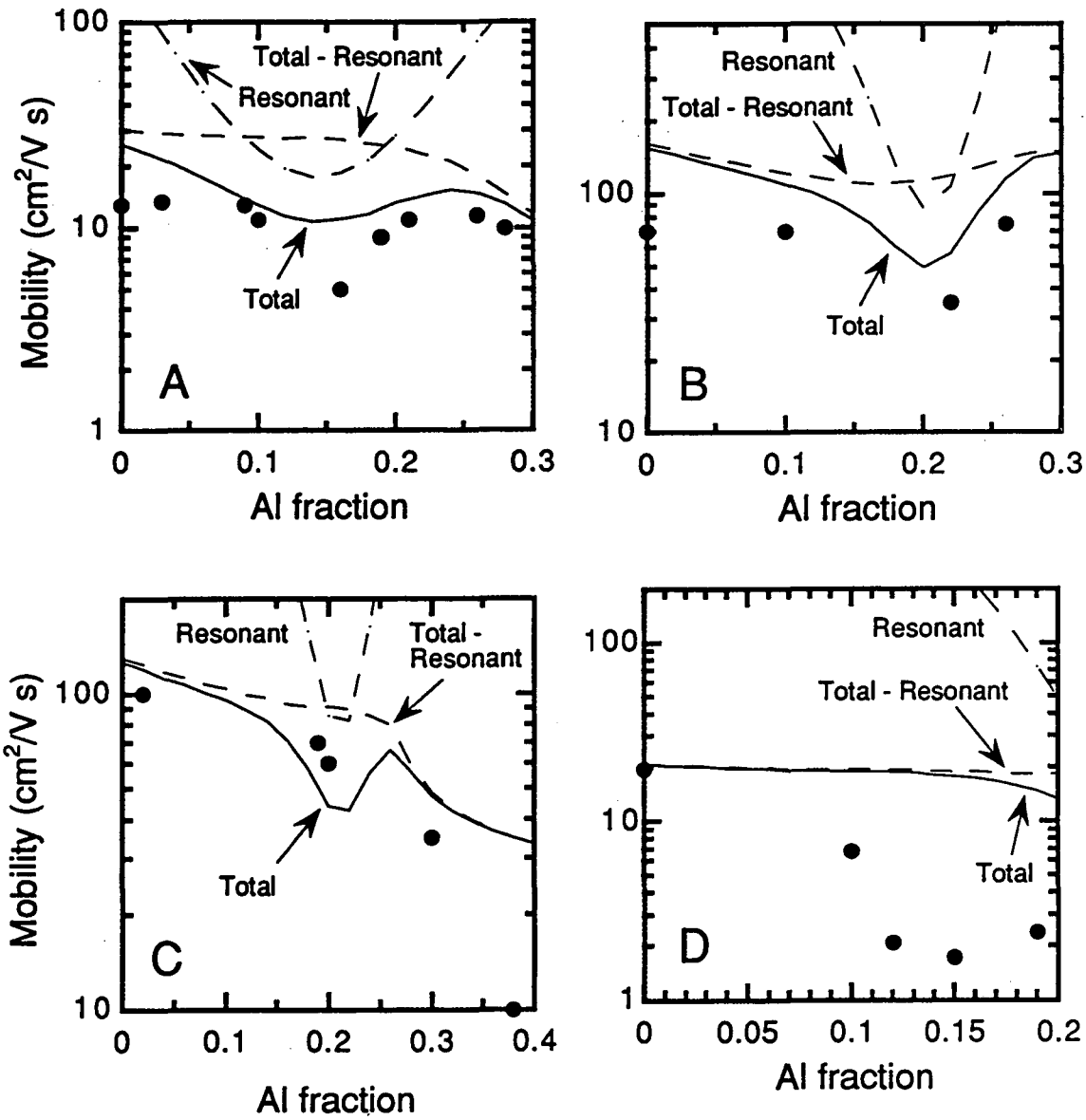


Figure 2-8. Plots of experimental (dots) and theoretical (lines) mobilities for $Al_xGa_{1-x}N$ as a function of Al content. Dashed lines show component mobilities corresponding to resonant scattering and the combination of all other mechanisms excluding resonant scattering.

One may argue that this is actually the cause of the apparent "freeze out" and that the drop in mobility is merely a variation in the quality of the samples. Because of this potential variation in defect concentration between samples when Al fraction is the experimental parameter which is varied, further experiments to study the effects of the resonant donor using hydrostatic pressure should be performed.

If the observed reduction in electron mobility is in fact caused by the crossing of the resonant defect and Fermi levels, one might also expect to see evidence of resonant scattering upon the application of hydrostatic pressure to GaN, InN, or their alloys, since the effect of hydrostatic pressure is to increase the energy of the conduction band minimum relative to the valence band (Perlin 1992). In contrast with alloying, however, the defect density in the sample will be a known and constant quantity for all pressures, leading to results which can be interpreted in a much less ambiguous fashion. No such experiments to measure elecron mobility as a function of hydrostatic pressure have yet been performed, but some predictions regarding the result can be made using this model.

In figure 2-9, calculated results are shown for the electron mobility in GaN under hydrostatic pressures in the range of 0 to 200 kbar for a defect concentration of 10^{19} cm^{-3}. As expected, a sudden decrease in the electron mobility occurs near 100 kbar, where the defect and Fermi levels have approximately the same energy. This effect could be tested experimentally by Hall effect measurements performed on a sample under hydrostatic pressure inside of a diamond anvil cell (Sakai 1982, Chen 1991). Although electrical measurements at such high pressures are difficult, a sample with a suitable Al content could be used to show the resonant effect at lower pressures.

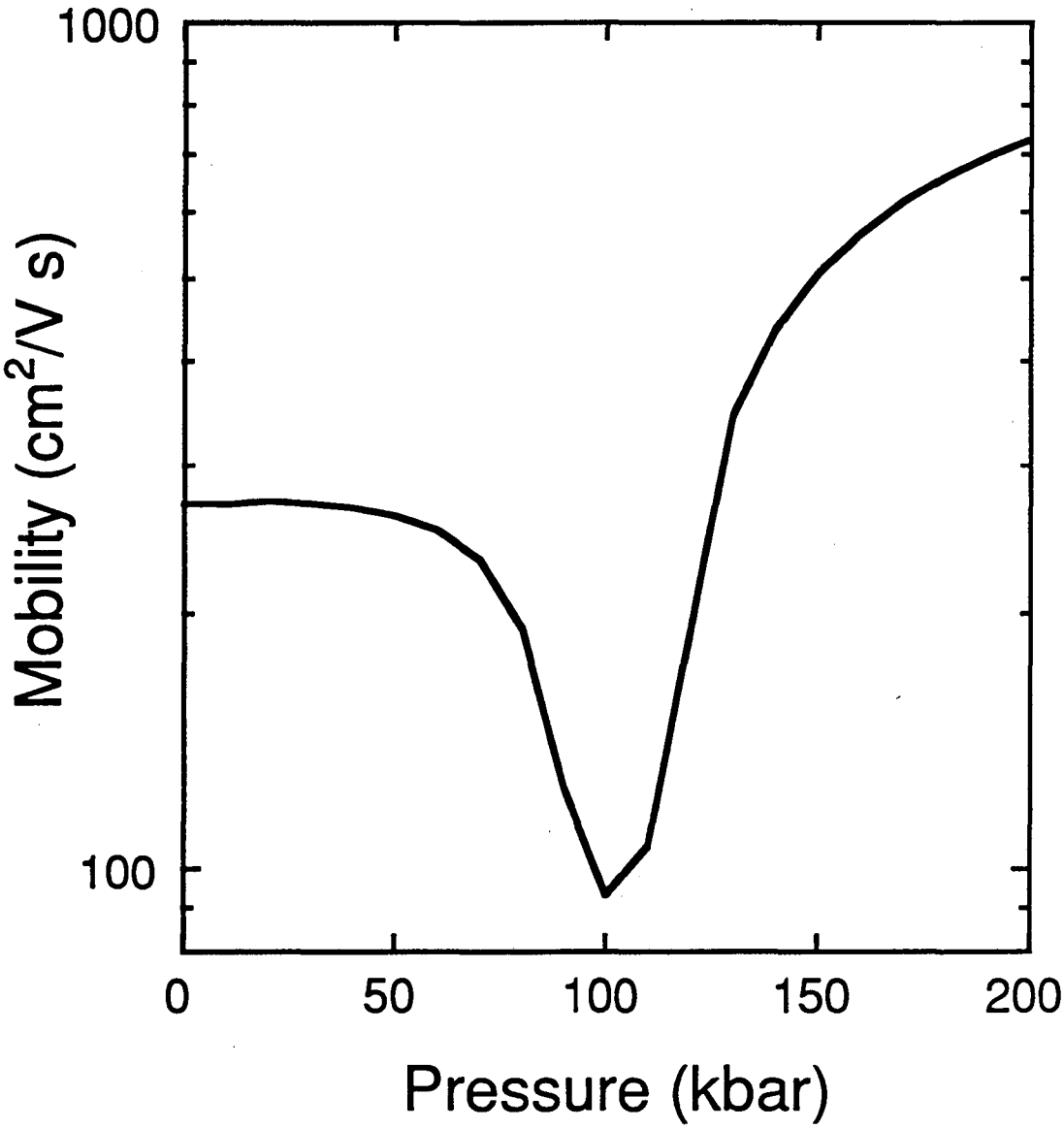


Figure 2-9. Plot of the calculated electron mobility as a function of hydrostatic pressure for GaN doped with 10^{19} cm^{-3} resonant donors.

3. Mobilities in $Al_xGa_{1-x}N$/GaN Modulation Doped Heterostructures

3.1 Introduction

As mentioned previously, GaN and other nitride based semiconductors are considered prime candidates for optoelectronic devices due to their wide bandgap. However, due to the large concentrations of charged scattering centers found even in unintentionally doped GaN, electron mobilities are generally quite poor. Since the response time of electronic devices depends heavily on the carrier mobility, this is a significant disadvantage for GaN as a material for making high speed semiconductor devices. As opposed to scattering from optical phonons at high temperatures, however, Coulomb scattering need not be considered an insurmountable limit on the low temperature mobility. As this type of scattering arises from electric fields which diminish as r^{-2} with distance, it was proposed by Esaki and Tsu in 1969 (Esaki 1969) that in order to reduce the interaction between carrier electrons and these fields and increase the mobility, modulation doped heterostructures (MDHs) could be grown in which the electrons were spatially separated from their parent donors.

The simplest modulation doped heterostructure consists of a doped epitaxial layer of one semiconductor grown on top of a different semiconductor, hence the name heterostructure. By choosing the two semiconductors appropriately, the conduction band of the doped layer can be made higher in energy than the conduction band of the undoped layer. When this occurs, it becomes energetically favorable for a certain amount of the conduction electrons in the doped layer to be transferred over to the undoped layer and occupy the conduction band there. This transfer of charge results in a bending of the conduction bands, as seen in figure 3-1, which prevents more charge from being transferred. Charge conduction perpendicular to the plane of the figure now occurs in the potential well formed by the band bending (Walukiewicz 1992). The most important difference between such a structure and an ordinary bulk semiconductor is that the

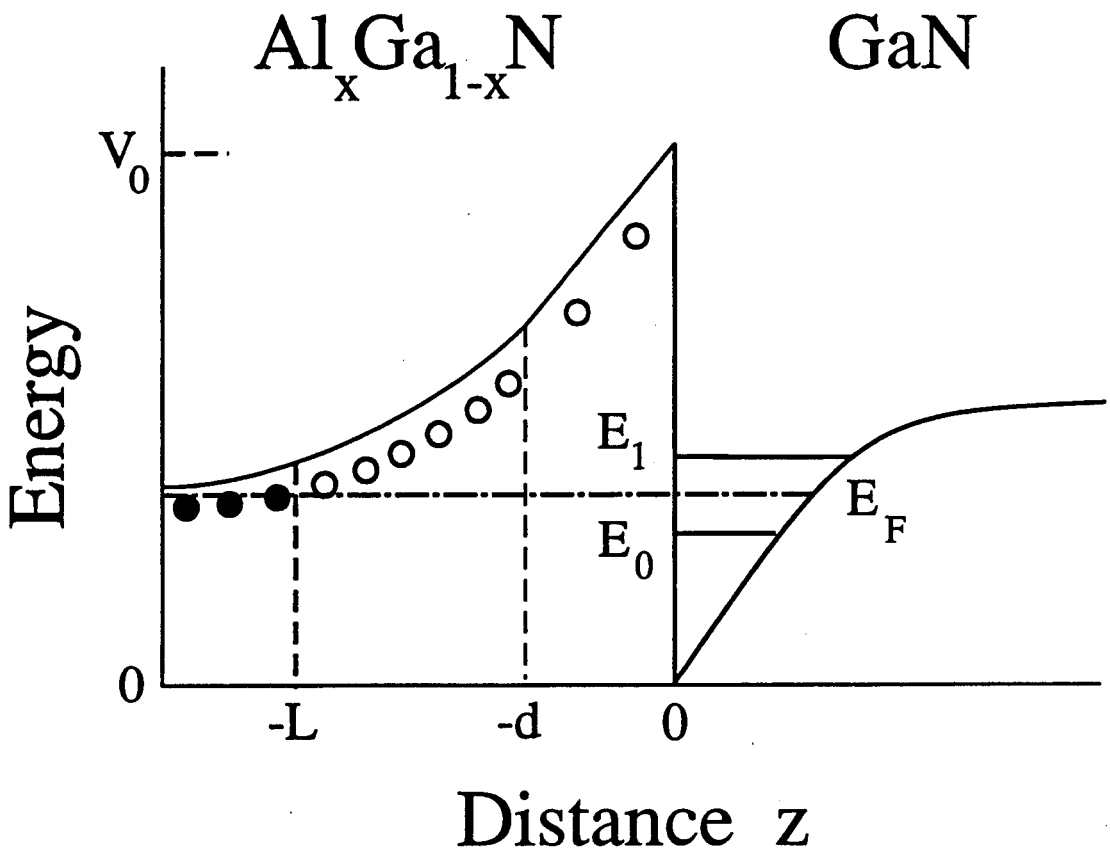


Figure 3-1. Schematic diagram of the energy subbands in a single-quantum-well modulation-doped heterostructure, showing the lowest two energy bands, Fermi energy, and both neutral (closed circles) and ionized (open circles) impurities in the AlGaN layer.

conduction electrons are now spatially separated from the donors from which they originated. The Coulomb fields felt by the conduction electrons are much weaker in this case than in the bulk case, where electrons and their ionized donors occupy the same volume. Thus, scattering of the electrons by the electrostatic interaction is much less effective in a MDH than in bulk semiconductors, resulting in higher mobilities, especially at low temperatures, where optical phonon scattering ceases to be effective.

Naturally, the creation of such a structure requires extremely precise control over the growth of the semiconductor, especially at the heterojunction. Any deviations from an atomically sharp interface will cause fluctuations of the conduction band energy in the undoped material near the interface, leading to scattering of the electrons. It was not until the late seventies that the molecular beam epitaxy (MBE) technique was sufficiently developed to grow such structures (Dingle 1978). Subsequent studies involving the AlGaAs/GaAs system showed that low temperature electron mobilities could be increased up to three orders of magnitude in MDHs as compared to bulk GaAs (Tsui 1981, Störmer 1981, Hiyamizu 1983). This discovery was the foundation for the creation and development of high speed GaAs semiconductor devices. Concurrent with the experimental development of these AlGaAs/GaAs heterostructures, there was a burst of theoretical work aimed at modeling the electron mobilities in these structures (Ando 1982a, Price 1982, Walukiewicz 1984, Hirakawa 1986). By now, the electron transport behavior in these materials is well understood.

Based on the experience with GaAs, it was natural to also grow AlGaN/GaN MDHs to try to maximize the electron mobilities in GaN to optimize the operating characteristics of GaN-based devices as well. Several attempts have been made with a fair degree of success (Khan 1993, Khan 1994, Redwing 1996). Some estimates of theoretical mobility limits in these structures have been made using a three dimensional approximation (Shur 1996) which, although accurate at room temperature, are not suitable for low temperatures and cannot properly describe the scattering from remote impurities, which is

the most important characteristic of modulation doped structures. One difference between the AlGaAs/GaAs MDHs and the AlGaN/GaN MDHs in existence today is the growth method used. AlGaAs/GaAs MDHs are typically grown using molecular beam epitaxy (MBE). However, due to the still poorly understood nature of GaN growth, MBE GaN is of much worse quality than that grown by vapor phase epitaxy (VPE). For this reason, VPE has been used to create all AlGaN/GaN MDHs so far, with surprisingly good results.

In this chapter, we extend the results of previous chapters to calculate mobilities in these structures. Again, we take all major scattering mechanisms, including acoustic and optical phonons, ionized impurities, and alloy disorder, into account.

3.2 Electronic Structure of the 2 DEG

The conduction band structure of a AlGaN/GaN heterostructure near the interface is shown in figure 3-1. In the ideal case, the GaN (at $z \geq 0$) is undoped while the AlGaN ($z \leq 0$) is selectively doped, consisting of an undoped region ($0 \geq z \geq -d$) known as the "spacer" and a doped region ($z \leq -d$). As GaN has a greater affinity for electrons than AlGaN, electrons from the donors in the AlGaN are transferred to the GaN, causing the conduction band to bend as shown. The positively charged donors in the AlGaN produce an electric field which creates a potential well in the GaN, confining the electrons to a narrow sheet at the interface and leading to a quantization of the energy band structure into subbands. At equilibrium, the transfer of electrons from the AlGaN to the GaN is determined by the equation:

$$V_0 - \frac{4\pi e^2}{\varepsilon_S}\left(N_S + N_{depl}\right) d - \frac{4\pi e^2}{2\varepsilon_S}\frac{(N_S + N_{depl})^2}{N_I} - E_b = E_F \qquad (3.1)$$

where V_0, E_F, E_b and the spacer width d are as shown in figure 3-1. N_S is the areal (two-dimensional) electron concentration in the GaN, N_{depl} is the areal concentration of residual charged impurities in the GaN, and N_I is the bulk concentration of so-called "remote"

donors in the AlGaN. E_b is the donor binding energy in the AlGaN and ε_S is the static dielectric constant of GaN. Assuming that only one subband is occupied, the Fermi energy E_F is given by

$$E_F = E_0 + \frac{\pi \hbar^2}{m^*} N_S \tag{3.2}$$

where $\pi\hbar^2/m^*$ is the two dimensional density of states in GaN and E_0 is the energy of the lowest subband. In most realistic cases, N_{depl} is at least two orders of magnitude smaller than N_S and thus can be neglected in this equation. In figure 3-2, the two-dimensional electron gas density is shown as a function of the AlGaN doping level for a $Al_{0.15}Ga_{0.85}N$/GaN structure with a variety of spacer widths.

In order to determine the GaN conduction band structure and quantized energy levels of the 2 DEG, Poisson's equation

$$\nabla^2\phi = -\frac{4\pi\rho}{\varepsilon} = -\frac{4\pi}{\varepsilon}|\psi|^2 \tag{3.3}$$

and Schrödinger's equation with GaN effective mass m*

$$-\frac{\hbar^2}{2m^*}\nabla^2\psi + V(z)\psi = E\psi \tag{3.4}$$

must be solved self-consistently for the electronic wavefunction (Ando 1982b). In the calculations performed for this work, the potential energy term consisted of the electrostatic potential energy as well as the exchange-correlation energy of the electrons. The form of the exchange correlation energy was taken from a paper by Hautmann and Sander (Hautman 1982), who used an expression derived by Gunnarsson and Lundqvist (Gunnarsson 1976).

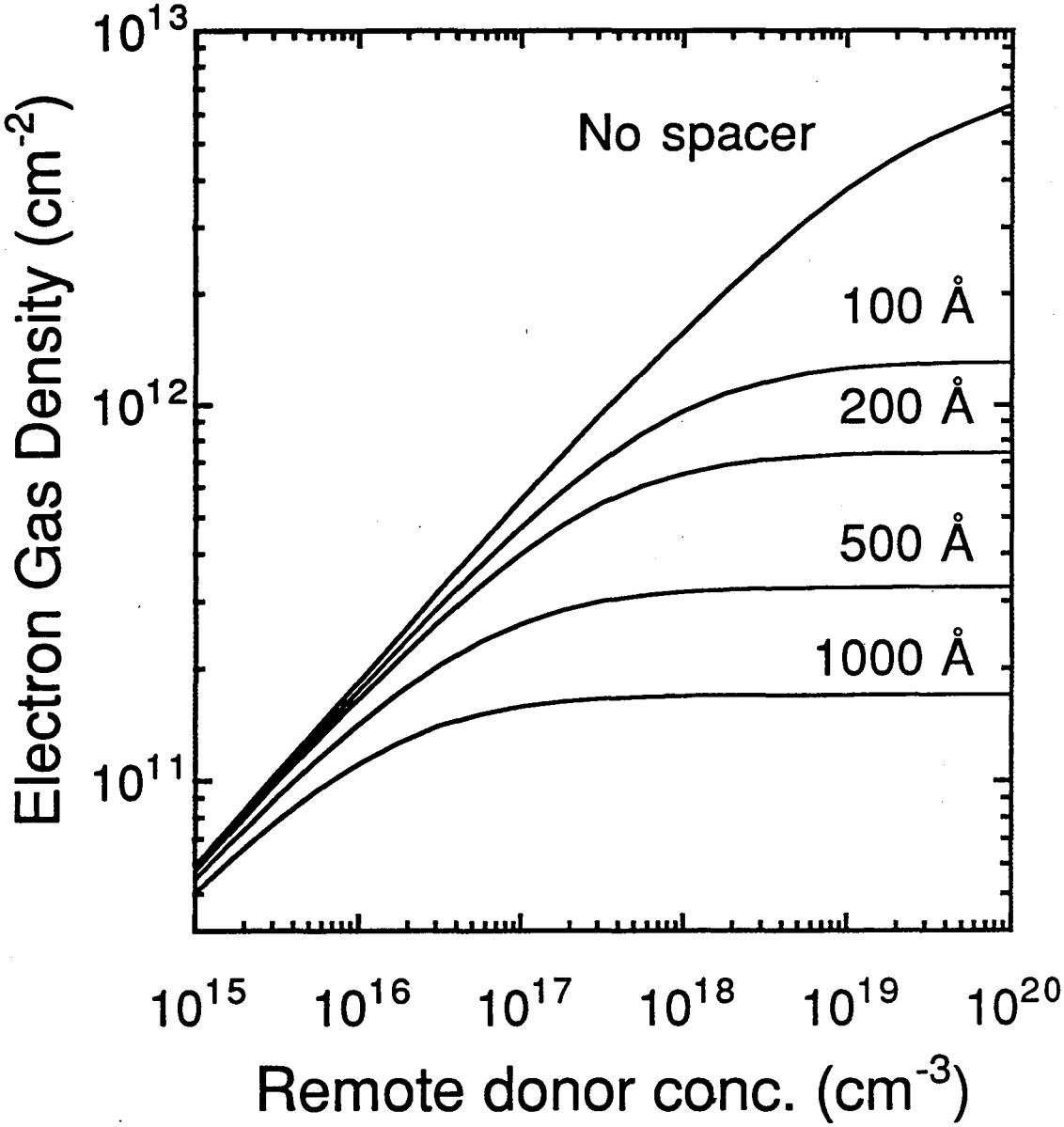


Figure 3-2. Graph of equation (3.1), showing two-dimensional electron gas density in the GaN layer as a function of $Al_{0.15}Ga_{0.85}N$ doping concentration (N_I) for various spacer widths (d).

In order to solve the two coupled equations, a trial wavefunction with one free parameter was devised and, using Poisson's equation, an expression for the potential energy of the electrons was calculated. A simple variational calculation to minimize the total energy of the electrons was then performed to determine the best value for the parameter. Using this value, the resulting expression for the potential energy was inserted into the Schrödinger equation (3.4) to determine the electronic energy levels and a new wavefunction. This wavefunction could then be compared to the original one, which was then modified as necessary. Naturally, an exact solution to both equations can only be found numerically but our intention was to find the closest possible analytic approximation to the true answer.

Previous calculations of a 2 DEG in Si and in the AlGaAs/GaAs system have shown that

$$\psi_0 = \phi_{x,y}\,\chi(z) = \phi_{x,y}\sqrt{\frac{b^3}{2}}\,z\exp(-bz/2) \tag{3.5}$$

where $\phi_{x,y}$ is a two dimensional plane-wave wavefunction and b is a variational parameter, is a good approximation to the electronic wavefunction of the lowest subband. Using this as a starting ground state electron wavefunction, a number of different analytic forms were tried. Of these,

$$\chi(z) = \frac{b^2}{\sqrt{6}}\,z^{3/2}\exp(-bz/2) \tag{3.6}$$

minimized the energy of the electrons. The parameter b, of course, depends on the electron gas density and is a measure of the width of the 2 DEG. As a function of the electron density, b varies roughly as $N_S^{0.315}$.

In figure 3-3, the energy levels of the lowest three subbands are shown as a function of the electron gas density. As can be seen, only the lowest subband is occupied

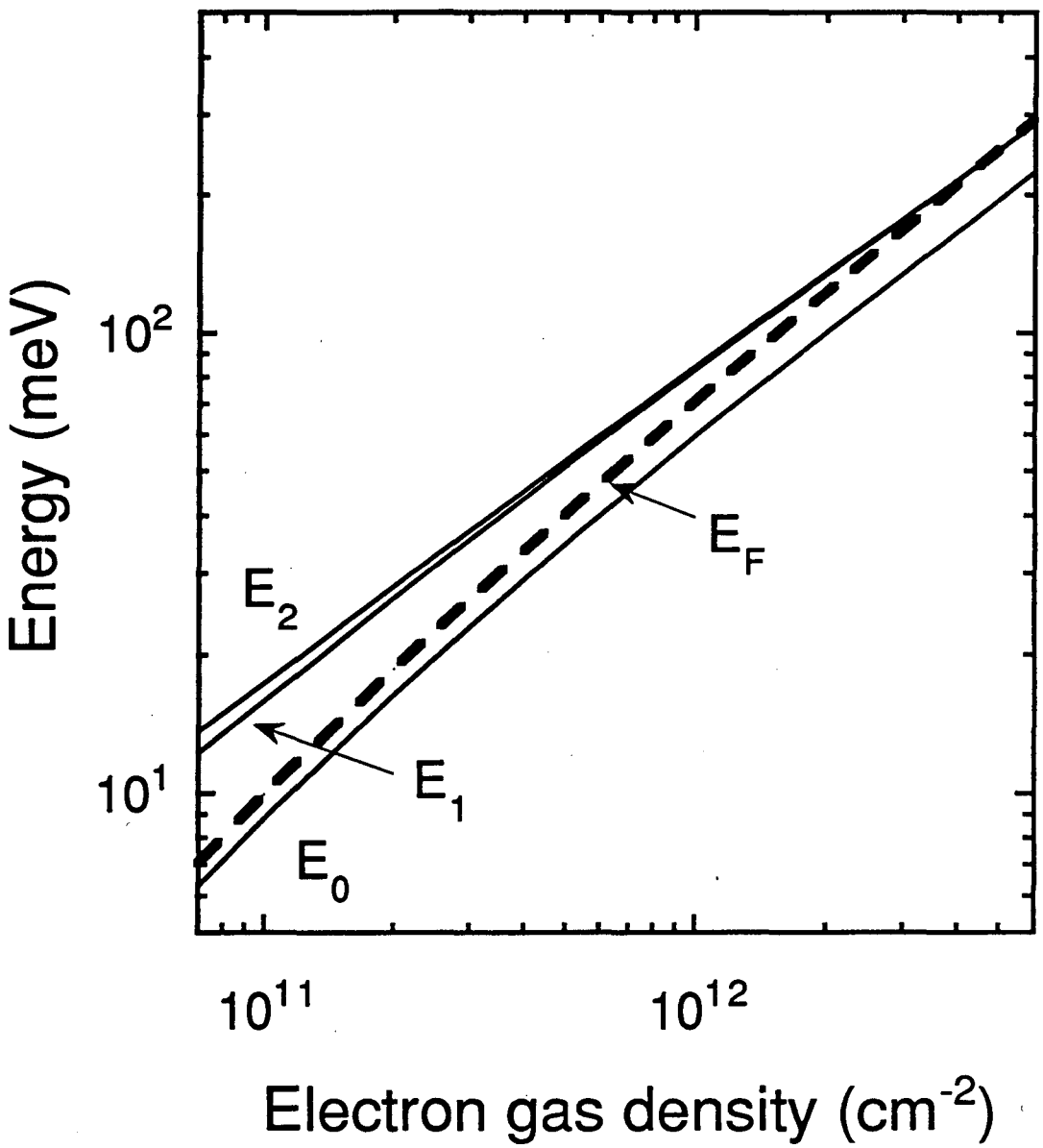


Figure 3-3. The lowest three energy subbands in the quantum well plotted as a function of GaN electron gas density. The heavy dotted line shows the position of the Fermi energy.

for electron concentrations below 4.5 x 10^{12} cm^{-2}. At greater concentrations, the higher subbands become occupied and their actual energies will deviate from those shown, which were calculated assuming that only the lowest subband is occupied. The energy separation between all subbands except for the first and second is very small so that once the electron gas concentration exceeds about 5 x 10^{12} cm^{-2}, many subbands become occupied very quickly, resulting in a loss of true two-dimensional behavior of the gas and a lowering of the mobility due to intersubband scattering. For this reason, in the mobility calculations which follow, only the case in which the lowest subband is the sole occupied level is treated.

3.3 Scattering Mechanisms

As in our study of mobilities in bulk III-Nitride compunds, all of the standard scattering mechanisms have been incorporated into the calculations of the 2D mobilities of these electrons. For typical GaN based MDHs grown today, electron concentrations exceed 10^{12} cm^{-2} which, for well widths of a few tens of Å, results in equivalent bulk electron concentrations of roughly 10^{19} cm^{-3}, a regime in which degenerate statistics can be used and the Fermi energy does not vary appreciably with temperature below 100K. At these low temperatures, scattering by optical phonons is insignificant so the total relaxation time can be approximated as a sum of the relaxation times due to each scattering process by Matthiessen's rule:

$$\frac{1}{\tau_{tot}} = \sum_{i} \frac{1}{\tau_i} \tag{3.7}$$

Since all of the mechanisms which are important at low temperatures are elastic and have well defined relaxation times, the complex variational method employed previously for bulk semiconductors is not necessary. The relaxation times for these scattering processes are

calculated in much the same way as for the three-dimensional cases. Here, the results are simply stated along with references.

At temperatures above 110 K, the validity of (3.7) becomes questionable due to the relaxation-time approximation made for inelastic optical phonon scattering (Walukiewicz 1984) and because of the limited applicability of degenerate statistics. However, at these higher temperatures, the mobility in these heterostructures is dominated by scattering from polar optical phonons. Because optical phonons in GaN have such a large energy (90.5 meV) compared to the energy separation between subbands, one must take a large number of subbands into account when calculating their effects and in so doing, the problem changes from a two-dimensional to a three dimensional one. This reflects the fact that electrons which absorb an optical phonon gain so much energy that they may be scattered completely out of the confining potential and into the bulk. Thus, in calculating the optical phonon component of the mobility, the previous variational principle method (Howarth 1953) and general Fermi-Dirac statistics are used in a three-dimensional approximation. The GaN material parameters which were used in these calculations are the same as those used in the previous section and can be found in Table 2-1.

Phonon scattering

Phonon scattering plays an important role in limiting the electron mobility in III-V semiconductors at temperatures above 100 K. As is the case for bulk material, the three most important phonon scattering processes are deformation potential acoustic, piezoelectric acoustic, and polar optical.

In MDHs, although the movement of the electrons is confined to a thin layer of perhaps 100 Å near the interface, it is usually assumed that acoustic phonons can propagate freely in all three dimensions. The relaxation time for the interaction of confined electrons with three dimensional acoustic phonons due to screened deformation potential scattering is given by (Walukiewicz 1992):

$$\frac{1}{\tau_{DP}} = \frac{3\,m^* a_C^2\, b\, k_B T}{16\,\pi\,\hbar^3 c_L} \int_0^{\pi} S(q)^2 (1-\cos\theta)\, d\theta \tag{3.8}$$

where a_C is the deformation potential, b is the variational parameter in (3.6), c_L is the elastic constant and S(q), the screening factor, is

$$S(q) = \frac{q}{q + q_S H(q)} \tag{3.9}$$

with

$$H(q) = \int_0^{\infty} dz \int_0^{\infty} dz'\ \chi(z)^2 \chi(z')^2 \exp\left(-q|z-z'|\right) \tag{3.10}$$

and $q_S = 2m^*e^2/\varepsilon_S\hbar^2$. The change in electron momentum q during a scattering process is related to the scattering angle θ between $\mathbf{k}$ and $\mathbf{k} + \mathbf{q}$ by $q = 2\,k_F \sin\theta/2$. For a degenerate electron gas, $k = k_F$.

For the zincblende structure, the relaxation time for screened piezoelectric mode scattering is calculated by (Walukiewicz 1992)

$$\frac{1}{\tau_{PE}} = \frac{1}{\tau_L} + \frac{2}{\tau_T} \tag{3.11}$$

where

$$\frac{1}{\tau_{L,T}} = \frac{k_B T\, \alpha_{L,T}}{\pi\hbar k^2} \int q\, S(q)^2 f_{L,T}(q)\, d\theta \tag{3.12}$$

with

$$\alpha_L = (eh_{14})^2 \frac{m^*}{4\hbar^2 c_L} \frac{9}{32}$$

$$\alpha_T = (eh_{14})^2 \frac{m^*}{4\hbar^2 c_T} \frac{13}{64} \tag{3.13}$$

and

$$f_L(q) = \frac{13 + 78(q/b) + 72(q/b)^2 + 82(q/b)^3 + 36(q/b)^4 + 6(q/b)^5}{13(1+(q/b))^6}$$

$$f_T(q) = \frac{1 + 6(q/b) + 12(q/b)^2 + 2(q/b)^3}{(1+(q/b))^6} \tag{3.14}$$

The combination of these two scattering processes (deformation potential and piezoelectric) gives the contribution to the relaxation time from acoustic phonons. As can be seen from (3.8) and (3.12), the acoustic phonon scattering rates are linear functions of temperature. This approximation is true at temperatures at which the thermal energy is greater than the acoustic phonon energy. At very low temperatures, in the Bloch-Grüneisen regime, only phonons with small wavevectors participate in scattering and the relaxation times increase superlinearly with temperature. The above expressions then over-estimate the acoustic phonon contributions to the total scattering rate (Störmer 1990). However, since temperature independent processes, such as Coulombic scattering tend to dominate the low temperature mobility, the deviations of the acoustic phonon scattering rate from linearity will have little effect on the total mobility.

As stated before, because the optical phonon energy is large (90.5 meV) compared to the energy separation of the subbands, the highly inelastic nature of polar optical scattering makes the total scattering rate the sum of many intersubband and intrasubband scattering processes. This results in a smearing out of the characteristic features of a 2 DEG, the most important one being the density of electrons within the potential well. For

this reason, the relaxation time for scattering of electrons in this 2 DEG by optical phonons is approximated by that calculated for the bulk (three dimensional) semiconductor case using the variational method (Howarth 1953).

Coulomb Scattering

The greatly enhanced low temperature mobilities found in modulation doped heterostructures over bulk semiconductors are due to the difference in the scattering of electrons from ionized impurities. In bulk semiconductors, the ionized impurities occupy the same region of space as the conduction electrons, making Coulomb scattering a very efficient process. Of course, the electrostatic interaction between an ionized donor and a conduction electron is somewhat screened by other conduction electrons. However, in order to achieve the high electron concentrations needed for efficient screening, the crystal must itself be highly doped, leading to higher concentrations of ionized impurity centers and effectively negating any beneficial screening effects.

In a AlGaN/GaN MDH, there are two different types of ionized impurity scattering. The first type is scattering by residual ionized impurities in the GaN, which works as described above. The second type is scattering by the ionized donors in the AlGaN barrier left behind by the conduction electrons. Since the electric field of the ionized centers drops off as the distance squared, this type of scattering is much less effective in limiting the electron mobility. A further type of Coulombic scattering which is not considered is scattering by charges at the heterojunction interface.

The standard method for calculating the relaxation times corresponding to the two types of Coulomb scattering employ the following expression (Hirakawa and Sakaki 1986)

$$\frac{1}{\tau_{Coul}} = \int_0^{\pi} \nu\,(\theta)\,d\theta \qquad (3.15)$$

where

$$v(\theta) = \frac{\pi\hbar(1-\cos\theta)}{2m^*}\left(\frac{q_s}{q}\right)^2 \int dz\left[S(q)F(q,z)\right]^2 N(z) \tag{3.16}$$

Here, N(z) is the distribution of Coulombic scattering centers, S(q) is given in (3.9), and

$$F(q,z) = \int dz' \left|\chi(z')\right|^2 \exp(-q|z-z'|). \tag{3.17}$$

The integral in (3.16) over z can be divided into three integrals, corresponding to scattering from remote ionized impurities in the doped AlGaN ($-L \le z \le -d$) and in the AlGaN spacer ($-d \le z \le 0$) and scattering from residual impurities in the GaN ($z > 0$). The contribution due to ionized impurities in the spacer layer can be neglected for concentrations up to 10^{15} cm^{-3} for spacers of less than a few hundred angstroms.

The calculation for the Coulombic scattering rates was performed in a temperature independent approximation, assuming that all scattering events involve electrons at the Fermi level. At temperatures above 100 K, when the Fermi energy starts to shift upwards, the approximation is no longer valid and overestimates the Coulombic scattering rate. However, at such temperatures, the mobility is dominated by phonons, so the effect of this error on the total mobility is negligible.

Alloy Disorder

Although the wavefunction that is used for the 2 DEG disappears at the interface, some electron density will inevitably penetrate into the AlGaN layer due to the finite nature of the potential barrier. Thus, scattering of the electrons due to alloy disorder must be considered. Following the procedure outlined in (Ando 1982a), the relaxation time for alloy disorder scattering is

$$\frac{1}{\tau_{all}} = \frac{m^* x(1-x)\Omega\langle V\rangle^2}{\hbar^3} \int_{-\infty}^{0} |\chi'(z)|^4 \, dz \tag{3.18}$$

where <V> is the conduction band offset between AlN and GaN, Ω is the volume of a unit cell, x is the Al fraction in the AlGaN and $\chi'(z)$ is the the part of the wavefunction which describes the penetration of the electron gas into the alloy

$$\chi'(z)^2 = \frac{4\pi e^2}{\varepsilon_S V_0}\left(\frac{1}{2}N_S + N_{depl}\right)\exp\left(\sqrt{\frac{8m^* V_0}{\hbar^2}}\, z\right) \tag{3.19}$$

Alloy disorder scattering rates are quite sensitive to the electron gas density, varying as the square of N_S. This dependence corresponds to the degree to which the electronic wavefunction penetrates the barrier into the AlGaN. As alloy disorder is a short range interaction, the screening of this potential has been neglected. All alloy disorder scattering times and contributions to the mobility are calculated assuming an Al fraction of 15% in the AlGaN layer.

Other scattering mechanisms

Some effects which have not been considered are interface roughness scattering, scattering due to interface charges, and the effect of the lattice mismatch between AlGaN and GaN (Ando 1982a). These scattering processes could in principle be incorporated into the calculations. However, as the aim is to determine inherent mobility limits in these heterostructures, they have been ignored. The incorporation of these effects requires knowledge of certain parameters such as the interface roughness, charge density, and polarization. Since these quantities are difficult to estimate and are in any case adjustable, inclusion of these mechanisms is largely pointless.

3.4 Electron Mobilities in AlGaN/GaN MDHs

In order to see the relative importance of the various scattering mechanisms described above in determining the total mobility, we first examine the temperature dependence of the 2 DEG mobility. In figure 3-4, the total mobility as well as the component mobilities of the electrons in a $Al_{0.15}Ga_{0.85}N$/GaN structure are shown as a function of temperature in the range from 1 to 300 K for a fairly "typical" heterostructure with a 200 Å spacer. As mentioned before, all Coulomb-type contributions were calculated in a temperature independent approximation, so the contribution to the mobility from all types of ionic impurities appear as straight lines. Since the relaxation time for Coulombic scattering processes is inversely proportional to the impurity concentration, it is easy to recalculate this graph for different impurity concentrations.

At very low temperatures, the electron mobility is dominated mainly by alloy disorder scattering and interactions with the Coulomb field of the remote donors. Starting at about 5 K, acoustic phonon scattering becomes the main mechanism limiting the mobility through both deformation potential and piezoelectric scattering. The strengths of both types are roughly equal. Of course, the exact temperature at which acoustic phonon scattering becomes dominant will depend on the remote donor concentration as well as the spacer width and alloy composition of the AlGaN layer. As one would expect, at temperatures above 170 K, the mobility is limited by polar optical phonon scattering.

As can be seen from figure 3-4, the inherent mobility limit in this particular MDH, which saturates at roughly 8×10^5 cm^2/V s is set by a combination of alloy disorder, remote ionized-impurities and, to a lesser extent, phonon scattering. There are several ways one could increase this inherent limit. Using wider spacers, the remote donor contribution can be made negligible even for concentrations up to 10^{20} cm^{-3}. Alloy disorder can be made less severe either by decreasing the Al fraction of the AlGaN or by reducing the penetrating electron density which can be accomplished by growing wider spacers or reducing the

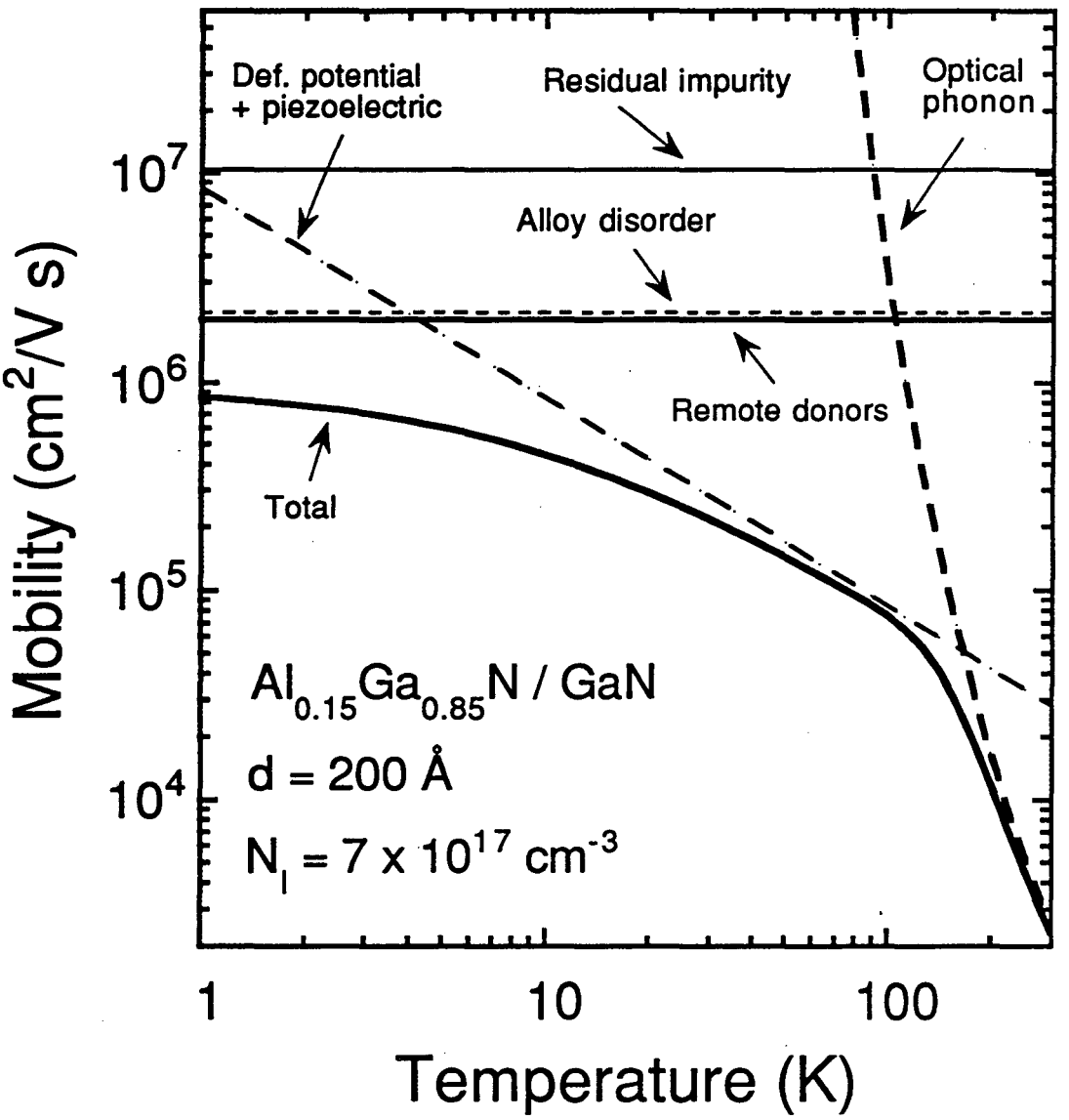


Figure 3-4. Temperature dependence of the electron mobility in a $Al_{0.15}Ga_{0.85}N$/ GaN MDH. All curves are calculated mobilities. The 2 DEG density for this case is 6.2×10^{11} cm^{-2} and the concentration of residual ionized impurities in the GaN is 1×10^{14} cm^{-3}.

remote doping concentration. By optimizing these parameters, the inherent mobility of a AlGaN/GaN MDH can theoretically be increased to almost 6 x 10^6 cm^2/V s, as will be seen. It should be noted however, that obtaining such mobilities puts very stringent requirements on the purity of the GaN layer. The residual charged impurity concentration must be less than 10^{13} cm^{-3}.

Figure 3-5 shows the calculated temperature dependence of the mobility for a MDH structure which better approximates what can be grown today. As one can see, mobilities well in excess of 10^4 cm^2/Vs are expected. However, a search of the literature shows that the highest published mobilities for such a structure are still only about 7500 cm^2/Vs with accompanying electron gas densities of 6 x 10^{12} cm^{-2} or greater (Redwing 1996). Comparing these densities with the results of the calculation in section II.3.2, it is obvious that at such high concentrations, a great number of subbands must be occupied because of the extremely close spacing of the higher lying levels. Although the electrons are still confined to a channel of only about 100 Å wide, true two dimensional behavior will not be observed in these structures due to a virtual continuum of bands being occupied, completely smearing out the effects of the subband quantization. The single period Shubnikov de Haas oscillations observed with these structures is due to the spherical Fermi surface formed by the three dimensional electron gas, the 1/H frequency of the oscillations corresponding to the area of the cross section of the Fermi sphere. It is possible to model these experimental mobilities using the three dimensional framework which was described in section II.2.1 and the results are shown in figure 3-6. Good agreement with experimental data can be obtained using an ionized impurity concentration of 1.2 x 10^{17} cm^{-3}. Of course at such high electron concentrations, alloy disorder scattering should also be taken into account. In order to observe true two dimensional behavior, the doping level of the AlGaN barrier must be reduced so that electron gas concentrations less than 4 x 10^{12} cm^{-2} are achieved.

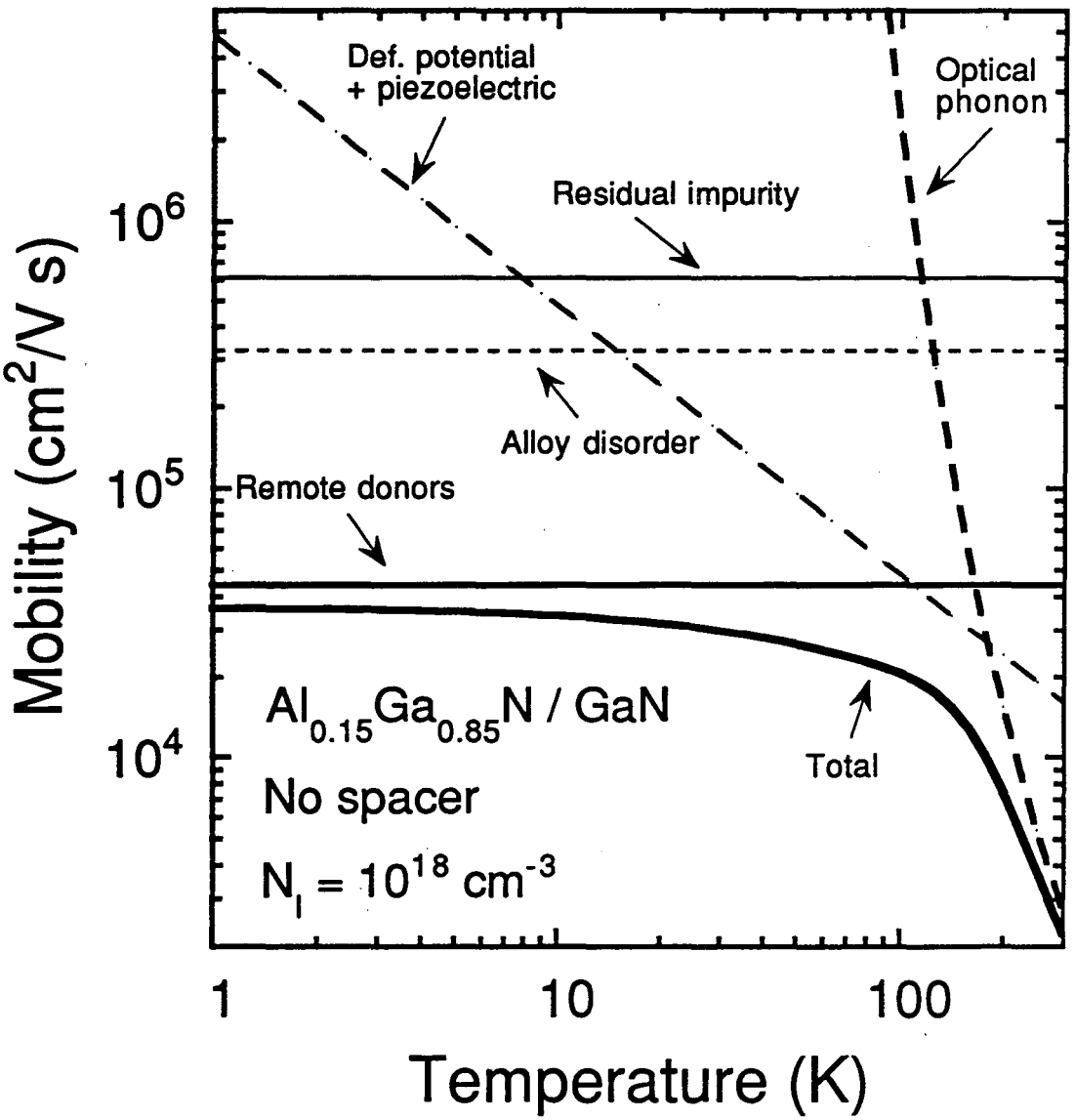


Figure 3-5. Temperature dependence of the mobility in a $Al_{0.15}Ga_{0.85}N$/GaN MDH which more closely approximates what can be achieved today. The residual ionized impurity concentration in this calculation is 4 x 10^{15} cm^{-3} and the doping level corresponds to a electron gas density of 1.59 x 10^{12} cm^{-2}.

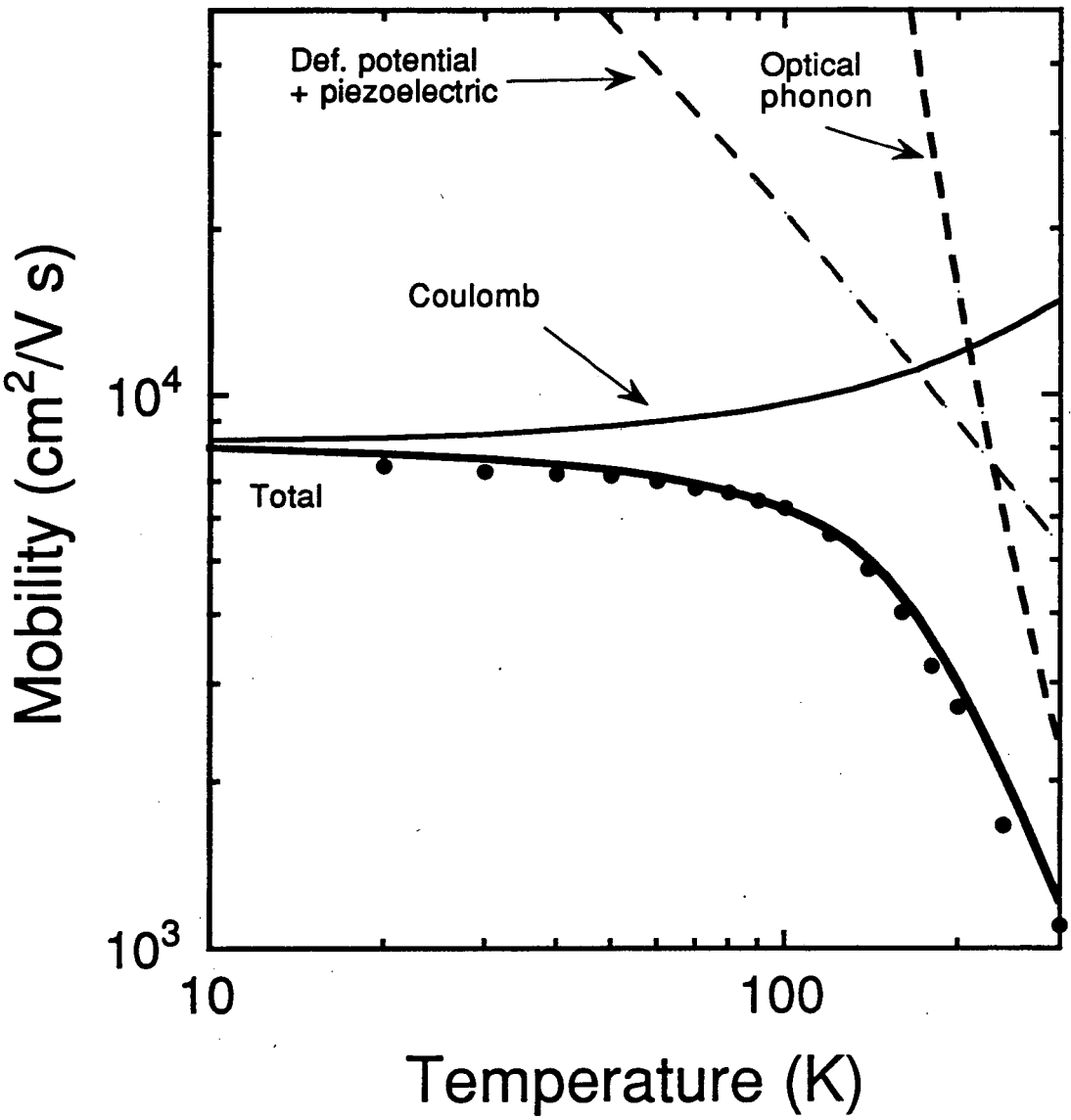


Figure 3-6. Comparison of experimentally obtained mobilities (filled circles) (Redwing 1996) with a three dimensional model for GaN mobilities. The component mobilities from optical and acoustic phonons are shown, along with a component for the contribution of Coulombic scattering from residual impurities. The theoretical curves were calculated for $N_S = 6 \times 10^{12}$ cm^{-2}, assuming an effective width of 500 Å for the electron gas. The Coulomb scattering contribution to the mobility was adjusted in order to fit the experimental data.

Another parameter which affects the mobility and may easily be varied is the electron gas density. There are several reasons for the dependence of the electron mobility on the density. First, as the electron density increases, the Fermi energy increases, as does the magnitude of the Fermi k vector. Since most of the scattering processes considered here are elastic, the vast majority of scattering events occur at the Fermi energy where there are an abundant number of occupied states to scatter and unoccupied states to scatter into and the processes themselves are dependent on the Fermi k vector. Second, as the electron density increases, the distribution of the electrons in the well changes, becoming more confined. This appears as an increase in the parameter b in (3.6). As can be seen from (3.8), acoustic deformation-potential scattering is directly dependent on b and increases as N_S increases. Increased N_S also results in better screening of Coulomb potentials and leads to less efficient scattering by ionized impurities. Finally, as mentioned previously, penetration of the wavefunction into the alloy layer increases with the electron density affecting the scattering rate due to alloy disorder.

The electron gas density in a MDH can be varied by several methods, such as changing the gate voltage of a HEMT device or by changing the remote doping concentration. Since the electron gas density saturates quickly for MDHs with wide spacers, we plot the results of our calculations as a function of remote doping concentration, which is related to the 2 DEG concentration as shown in figure 3-2. Figure 3-7 shows the remote doping dependence of the mobility at 4 K for two MDHs--one with a 200 Å spacer and one without a spacer.

For the MDH with no spacer (figure 3-7a), the dominant process limiting the mobility is Coulomb scattering by remote impurities. At higher doping concentrations, alloy disorder also begins to play a significant role in reducing the mobility. In figure 3-7b, mobilities in a MDH with a 200 Å spacer are shown. As expected, Coulomb scattering from remote donors is now much less important at low doping concentrations--residual ionized impurities in the GaN now account for the bulk of the scattering. If this

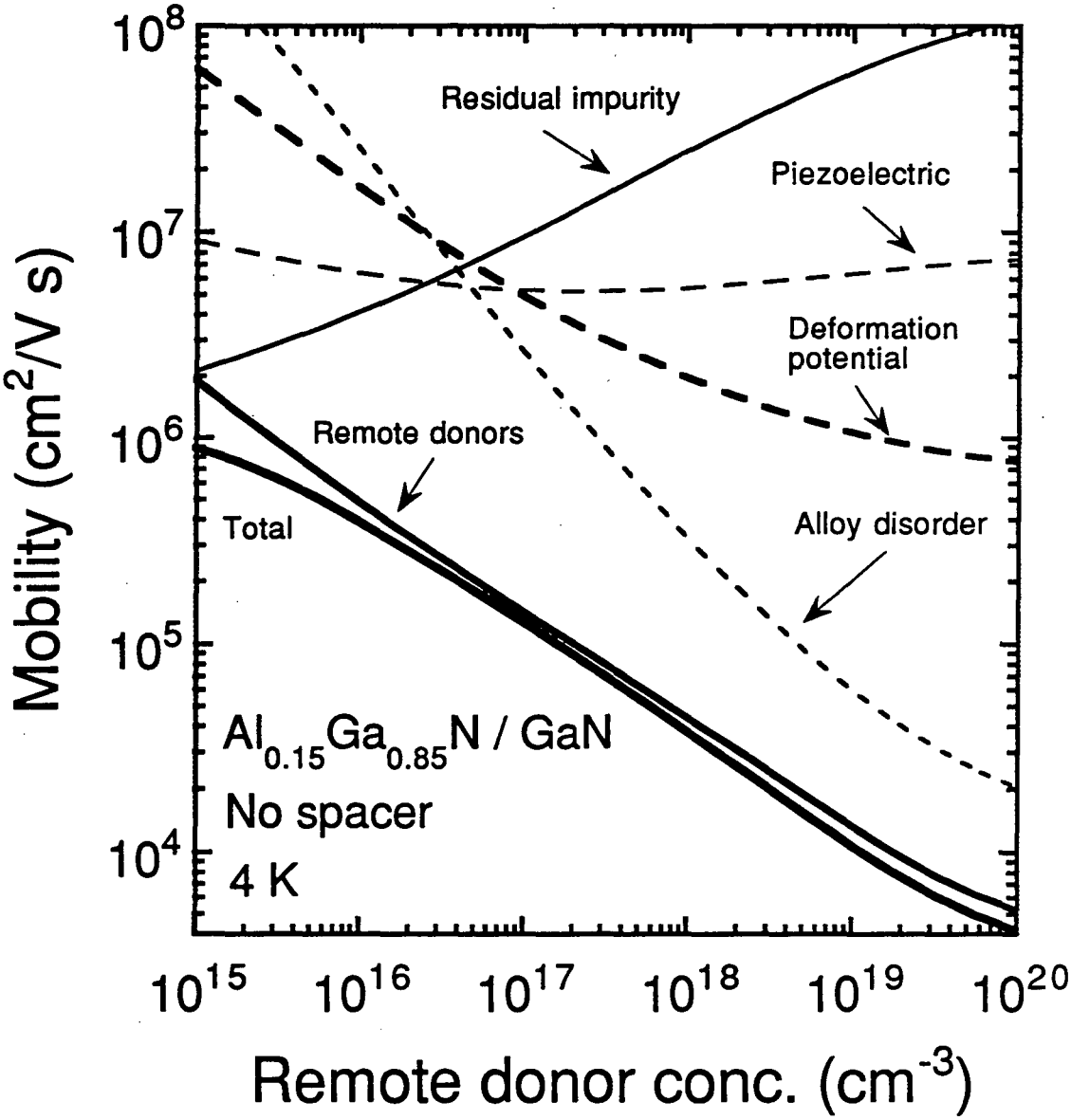


Figure 3-7a. Remote doping dependence of the 4 K electron mobilities in $Al_{0.15}Ga_{0.85}N/GaN$ heterostructures with no spacer. The GaN background impurity concentration is 1×10^{14} cm^{-3}.

background ionized impurity concentration can be reduced further, one may reach a point where piezoelectric mode scattering by acoustic phonons is the dominant effect. As the remote doping concentration (and thus the electron gas density) is increased, the Coulomb fields of the residual impurites are effectively screened, causing that component of the mobility to increase. Although the remote donors are more effectively screened as well, the increase in their number leads to an overall lowering of that contribution to the mobility. Scattering by alloy disorder also becomes important at high electron densities. Another feature of the results shown in figure 3-7b is that above remote doping concentrations of 3 x 10^{18} cm^{-3}, the mobility levels off. The reason is that for heterostructures with a 200 Å spacer, the 2D electron density saturates at this doping concentration due to the potential drop across the spacer and thus there is no further change in the number of remote ionized donors or of the penetration of the electron wavefunction into the alloy for high remote doping concentrations.

Finally, we examine the dependence of the mobility on the spacer width. In figure 3-8a, the total 4 K mobility is plotted as a function of remote doping concentration for a variety of spacer widths, assuming a residual impurity concentration of 10^{13} cm^{-3}. It should be kept in mind that for different spacer widths, the same remote donor concentration results in very different electron concentrations. For the smaller spacer widths, the mobility decreases as the remote doping concentration is increased due to Coulombic scattering from the remote donors as shown in figure 3-7. For larger spacer widths, the electron mobility is governed almost exclusively by piezoelectric mode acoustic phonon scattering and Coulombic scattering from residual charged impurities in the GaN. The remote donors are too far away to be effective and the electron gas density is too low for alloy disorder scattering to play any significant role. However, due to the large spacer width, the electron density is low enough so that the residual charged impurities are not well screened. If the residual ionized impurity concentration is higher, such as 10^{15} cm^{-3}, it is no longer advantageous to have such a large spacer. In that case, as can be seen from

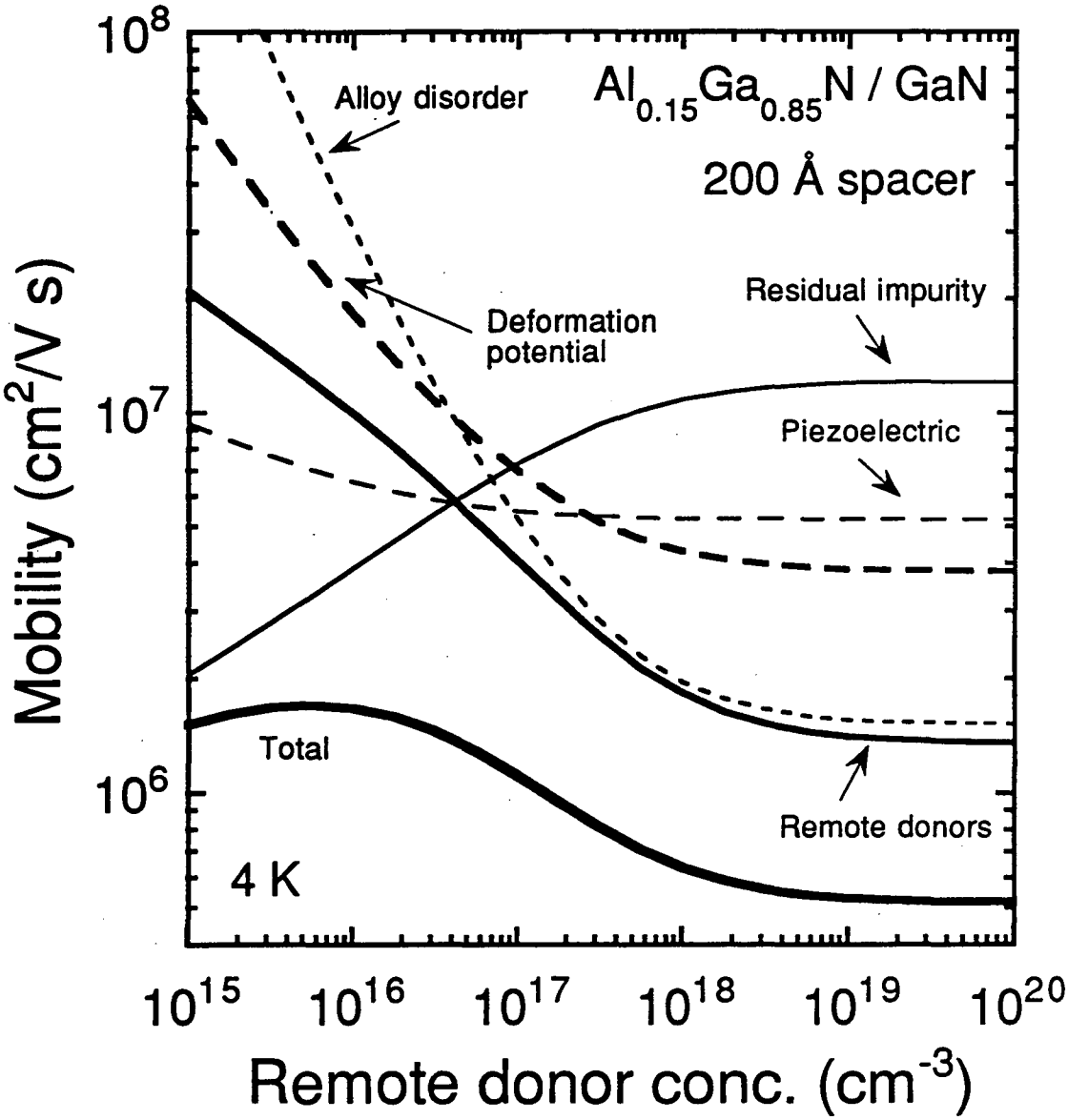


Figure 3-7b. Remote doping dependence of the 4 K electron mobilities in $Al_{0.15}Ga_{0.85}N/GaN$ heterostructures with a 200 Å wide spacer (B). The GaN background impurity concentration is 1 x 10^{14} cm^{-3}.

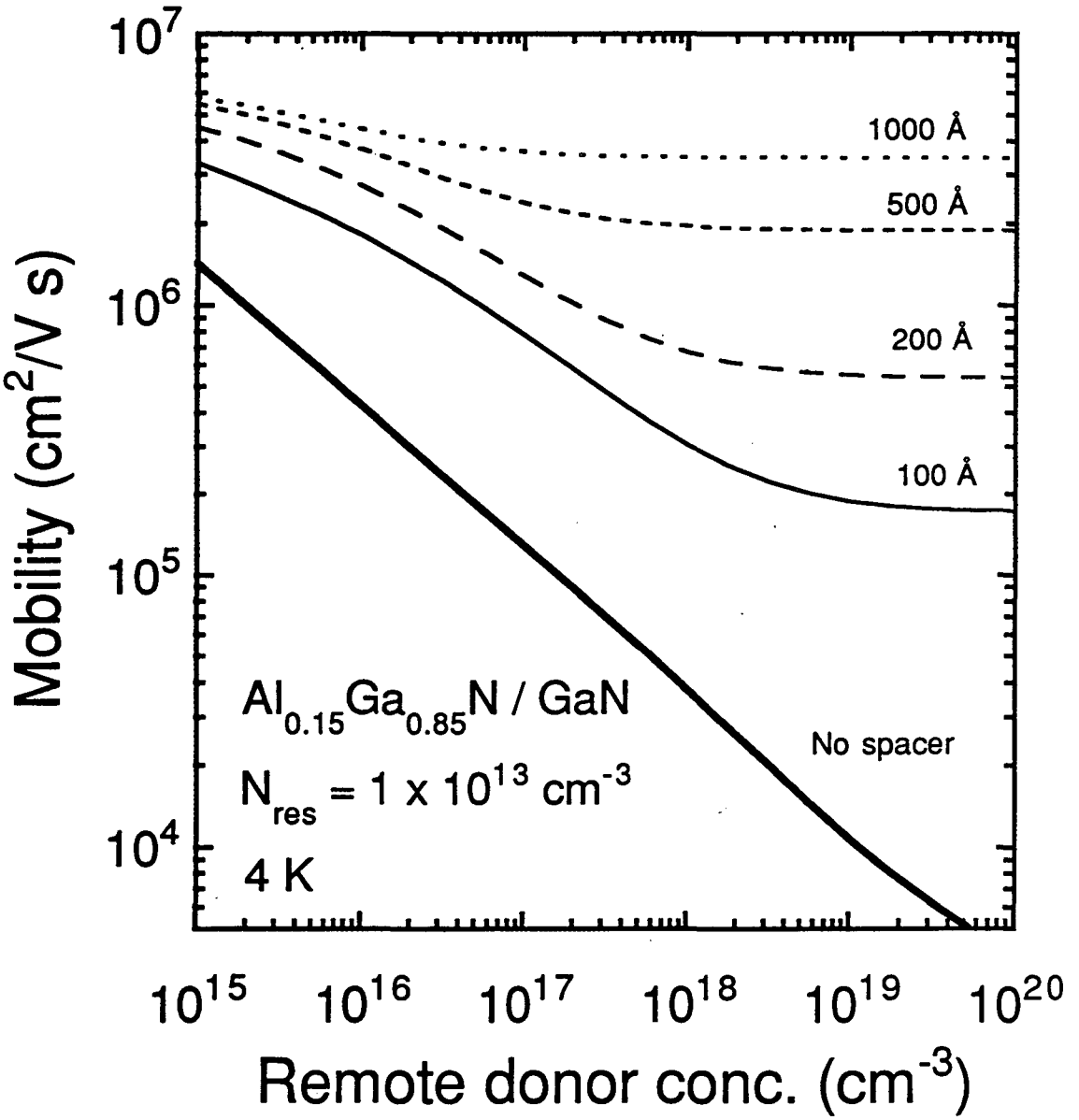


Figure 3-8a. Mobility plotted as a function of remote doping concentration for a variety of spacer widths. For low residual ionized impurity concentrations (10^{13} cm^{-3} in this case), a large spacer produces the highest mobilities.

figure 3-8b, the highest mobilites are actually attained by growing only a 200 Å spacer with a remote doping concentration of 10^{17} cm^{-3}, which from figure 3-2 corresponds to $N_S = 4 \times 10^{11}$ cm^{-2}, as compared to the other spacer widths for which we have made calculations. For larger spacers, the electron density is simply not large enough to effectively screen any residual charged impurities in the GaN.

Another situation in which a larger N_S is desirable is in device applications, in which it is often the conductivity, (the product of the electron charge, the mobility, and N_S) which is the critical parameter which must be maximized. Figure 3-9a shows the conductivity plotted for a number of different spacer widths at 4 K. In each case, the conductivity has been maximized with respect to the remote doping concentration, for concentrations in the range between 10^{15} and 10^{20} cm^{-3}. In figure 3-9b, the doping concentration at which the maximum conductivity is achieved is plotted for the same spacer widths. Again, as with the mobility, the combination of parameters which produces the highest conductivity depends on the residual charged impurity concentration in the GaN layer. If it can be made very small (say, 10^{13} cm^{-3}), then to maximize the conductivity, a 500 Å spacer layer should be grown and the remote doping concentration should be 1.5×10^{17} cm^{-3}. On the other hand, if growing pure GaN is a problem and residual impurities cannot be reduced below 10^{15} cm^{-3}, then the spacer should only be made about 200 Å thick and the AlGaN layer should be much more heavily doped (about 5×10^{18} cm^{-3}) in order to maximize the conductivity.

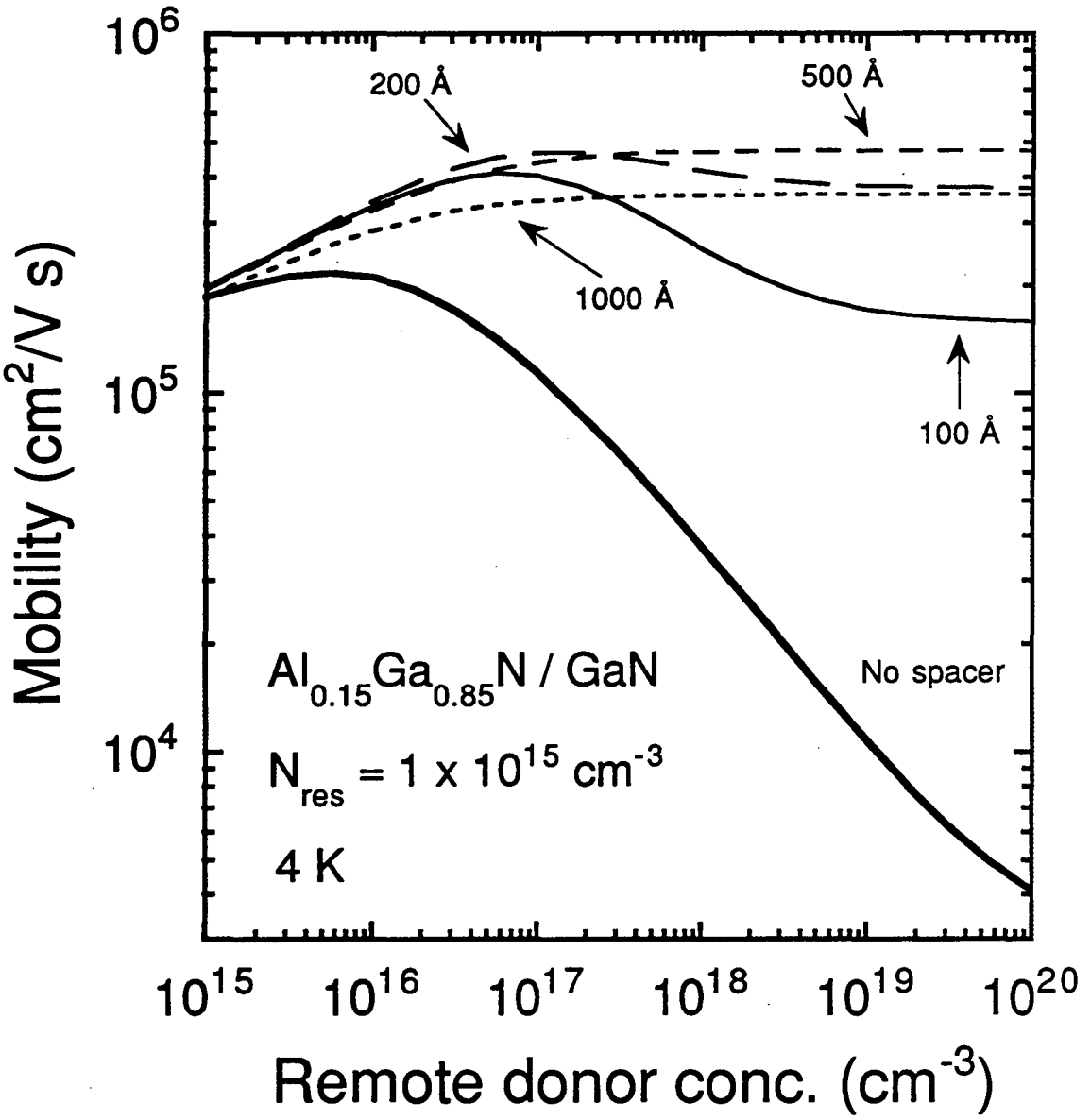


Figure 3-8b. Mobility plotted as a function of remote doping concentration for a variety of spacer widths. For high residual ionized impurity concentrations (10^{15} cm^{-3} here), a small spacer, leading to large electron densities, gives the highest mobilities.

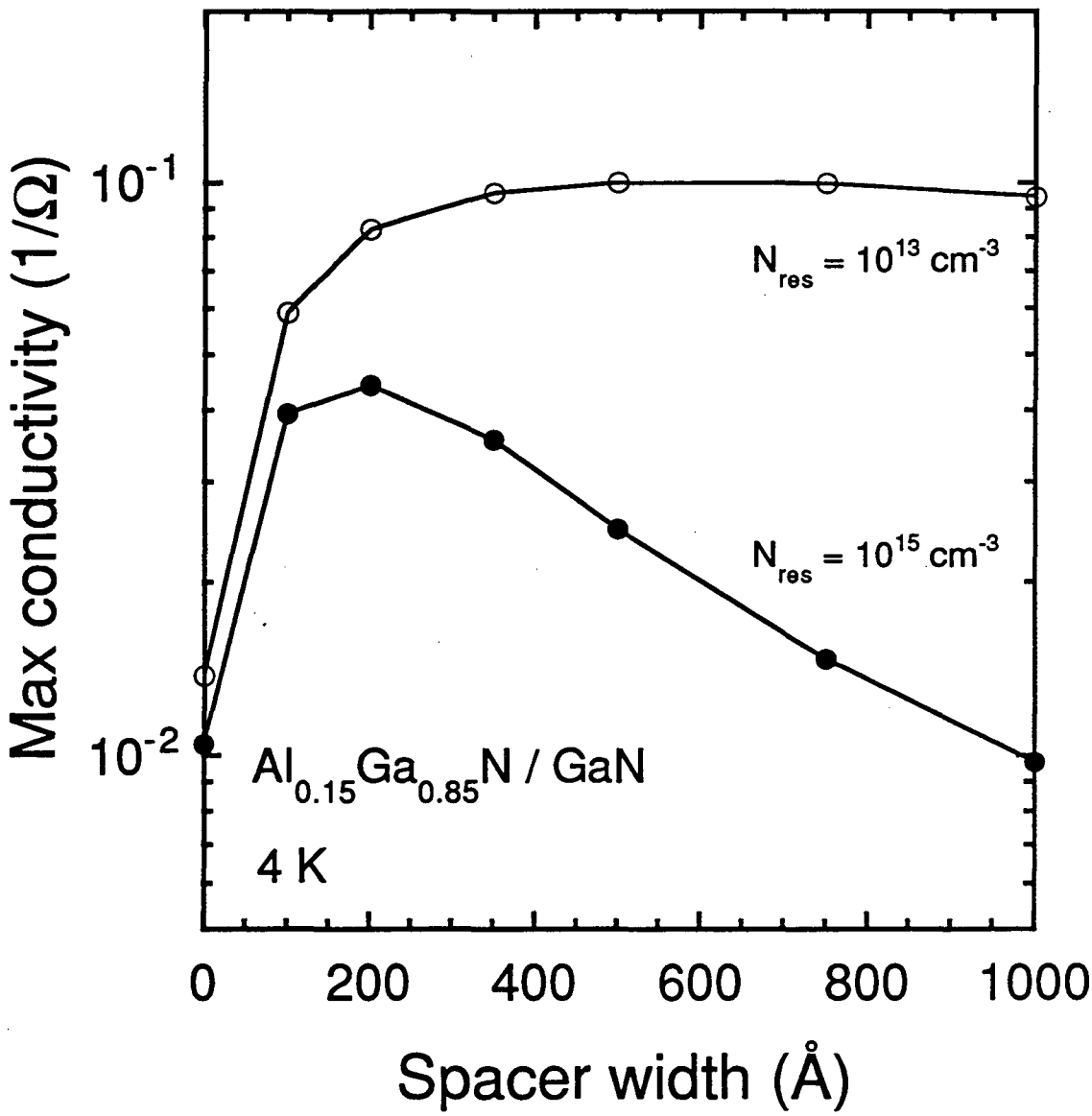


Figure 3-9a. Maximum conductivity plotted for seven different spacer widths for $Al_{0.15}Ga_{0.85}N$/GaN MDHs with two different residual impurity concentrations, 10^{15} and 10^{13} cm^{-3}.

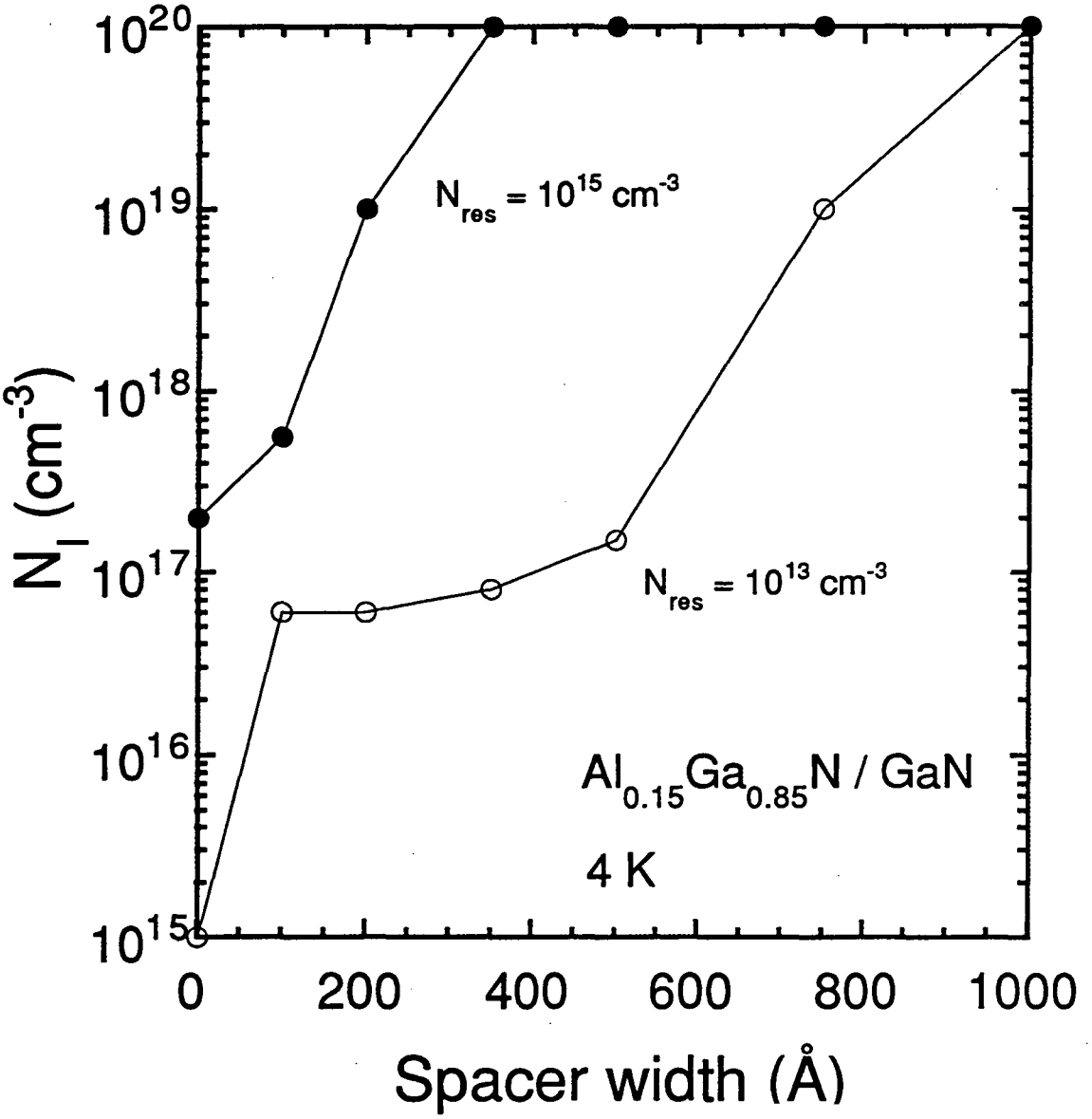


Figure 3-9b. Remote doping concentration at which maximum conductivity is achieved, plotted for the same spacer widths as in figure 3-8a, for the same two values of residual impurity concentration.

4. Conclusion

Some aspects of the electron mobility have been studied in GaN and other group III-Nitride based structures. There is increasing evidence that O may be the dominant donor in nominally undoped GaN and that it is a resonant donor, with its first ionization level within the conduction band. Due to the nature of this donor and its localized potential, the electron mobilities in GaN material which contains large concentrations of this impurity may suffer. In order to achieve the highest possible mobilities, hydrogenic donors should be used as dopants. In addition, the presence of resonant defects can cause the mobilities in AlGaN alloys with a certain Al percentage to become lowered due to resonant scattering. Although no conclusive signs of this effect have yet been seen experimentally, Hall effect measurements on samples under large hydrostatic pressure may reveal it.

The studies of mobilities in AlGaN/GaN MDHs have shown that so far, the purity of the AlGaN layers that can be grown today does not yet permit formation of a true two-dimensional gas at the AlGaN/GaN interface. With just a bit more improvement, low temperature electron mobilities in these structures can be expected to jump from currently observed values of almost 8000 cm^2/V s to several tens of thousands. Furthermore, if impurity concentrations in AlGaN can be reduced to 10^{15} cm^{-3}, enabling effective spacers to be grown, low temperature mobilities of nearly 10^6 cm^2/V s should be able to be obtained. Although current technology is still some distance from achieving high quality nitride films, these calculations clearly demonstrate that much higher mobilities are possible, providing some incentive for the development of improved growth techniques.

5. References

Abernathy, C. R., J. D. MacKenzie, S. R. Bharatan, K. S. Jones, and S. J. Pearton, Appl. Phys. Lett. **66**, 1632 (1995).

Ando, T., J. Phys. Soc. Japan **51**, 3900 (1982a).

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

Berggren, K. F. and B. E. Sernelius, Phys. Rev. B **24**, 1971 (1981).

Böer, K. W., *Survey of Semiconductor Physics* (Van-Nostrand-Reinhold, New York, 1990), p. 779.

Chen, A.L., P. Y. Yu, and M. P. Pasternak, Phys. Rev. B **44**, 2883 (1991).

Chin, V. W. L., T. L. Tansley, and T. Osotchan, J. Appl. Phys. **75**, 7365 (1994).

Dingle, R. B., Philos. Mag. **46**, 831 (1964).

Dingle, R., H. L. Störmer, A. C. Gossard, and W. Wiegmann, Appl. Phys. Lett. **33**, 665 (1978).

Ehrenreich, H., Phys. Rev. **120**, 1951 (1960).

Esaki, L., and R. Tsu, IBM Research Laboratories, Internal Report No. RC 2418, 1969 (unpublished).

Gaskill, D.K., K. Doverspike, L.B. Rowland, and D. L. Rode,.in *Compound Semiconductors 1994. Proceedings of the Twenty-First International Symposium, San Diego, Sept. 1994*, edited by H. Goronkin and U. Mishra (IOP Publishing, Bristol 1995), p. 425.

Gunnarsson O. and B. I. Lundqvist, Phys. Rev. B **13**, 4274 (1976).

Harrison, W. A., Phys. Rev **104**, 1281 (1956).

Hautman J. and L. M. Sander, Phys. Rev. B **32**, 980 (1982).

Hirakawa K. and H. Sakaki, Phys. Rev. B **33**, 8291 (1986).

Hiyamizu, S., J. Saito, K. Nanbu, and T. Ishikawa, Jpn. J. Appl. Phys. **22**, L609 (1983).

Hovel, H. J., and J. J. Cuomo, Appl. Phys. Lett. **20**, 71 (1972).

Howarth, D. and E. H. Sondheimer, Proc. Phys. Soc. London A **219**, 53 (1953).

Hwang, C.-Y., M. J. Schurman, W. E. Mayo, Y.-C. Lu, R. A. Stall, and T. Salagaj, J. Elec. Mat. **26**, 243 (1997).

Khan, M. A., R. A. Skogman, R. G. Schulze, and M. Gershenzon, Appl. Phys. Lett. **43**, 492 (1983).

Khan, M. A., A. Bhattarai, J. N. Kuznia, and D. T. Olson, Appl. Phys. Lett. **63**, 1214 (1993).

Khan, M. A., J. N. Kuznia, D. T. Olson, W. J. Schaff, J. W. Burm, and M. S. Shur, Appl. Phys. Lett. **65**, 1121 (1994).

Koide, Y., H. Itoh, N. Sawaki, I. Akasaki, et al., J. Electrochem. Soc. **133**, 1956 (1986).

Lehochkey, S. L., J. G. Broerman, D. A. Nelson, and C. R. Whitsett, Phys. Rev. B **9**, 1598 (1974).

Li, X. J., Z. Xu, Z. J. He, H. Z. Cao, W. Su, Z. C. Chen, F. Zon, and E. G. Wang, Thin Solid Films **139**, 261 (1986).

Litwin-Staszewska, E., S. Porowski, and A. A. Filipchenko, Phys. Stat. Sol. B **48**, 525 (1971).

Makowski, L. and M. Glicksman, J. Phys. Chem. Sol. **34**, 487 (1973).

Maruska, H. P. and J. J. Tietjen, Appl. Phys. Lett. **15**, 327 (1969).

Molnar, R. J., T. Lei and T. D. Moustakas, Appl. Phys. Lett. **62**, 72 (1993).

Nakamura, S., Jpn. J. Appl. Phys. **30**, L1705 (1991).

Nakamura, S., T. Mukai, M. Senoh, Jpn. J. Appl. Phys. **31**, 2883 (1992).

Neugebauer, J. and C. G. Van de Walle, Appl. Phys. Lett. **69**, 503 (1996).

Pankove, J. I., Mater. Res. Symp. Proc. **162**, 515 (1990).

Perlin, P., I. Gorczyca, N. E. Christiansen, I. Grzegory, H. Teisseyre, T. Suski, Phys. Rev. B **45**, 13307 (1992).

Perlin, P., T. Suski, H. Teisseyre, M. Leszczynski, I. Grzegory, J. Jun, S. Porowski, P. Boguslawski, J. Bernholc, J. C. Chervin, A. Polian, and T. D. Moustakas, Phys. Rev. Lett. **75**, 296 (1995).

Price, P. J., Surf. Sci. **113**, 199 (1982).

Raikh, M. E., A. L. Efros, Sov. Phys. Sol. State **28**, 735 (1986).

Redwing, J. M., M. A. Tischler, J. S. Flynn, S. Elhamri, M. Ahoujja, R. S. Newrock, and W. C. Mitchel, Appl. Phys. Lett. **69**, 963 (1996).

Rode, D. L., Phys. Stat. Sol. B **55**, 687 (1973).

Rode, D. L., in *Semiconductors and Semimetals, Vol. 10*, edited by R. K. Willardson and A. C. Beer (Academic Press, New York 1975), Chap. 1.

Rowland, L. B., K. Doverspike, D. K. Gaskill, Appl. Phys. Lett. **66**, 1495 (1995).

Sakai, N., T. Kajiwara, K. Tsuji, and S. Minomura, Rev. Sci. Instrum. **53**, 499 (1982).

Sasaki T. and S. Zembutsu, J. Appl. Phys. **61**, 2533 (1987).

Shockley, W., *Electrons and Holes in Semiconductors* (Van-Nostrand-Reinhold, New York, 1950), p. 264ff.

Shur, M., B. Gelmont and M. Asif Khan, J. Elect. Mat. **25**, 777 (1996).

Störmer, H. L., A. C. Gossard, W. Wiegmann, and K. Baldwin, Appl. Phys. Lett. **39**, 912 (1981).

Störmer, H. L., L. N. Pfeiffer, K. W. Baldwin, and K. West, Phys. Rev. B **41**, 1278 (1990).

Strite, S., and H. Morkoc, J. Vac. Sci. Technol. B **10**, 1237 (1992).

Sun C. J. and M. Razeghi, Appl. Phys. Lett. **63**, 973 (1993).

Tansley T. L. and C. P. Foley, Electron. Lett. **20**, 1066 (1984).

Tsui, D. C., A. C. Gossard, G. Kaminsky, and W. Wiegmann, Appl. Phys. Lett. **39**, 712 (1981).

Walukiewicz, W., L. Lagowski, aL. Jastrzebski, M. Lichtensteiger, and H. C. Gatos, J. Appl. Phys. **50**, 899 (1979).

Walukiewicz, W., H. E. Ruda, J. Lagowski, and H. C. Gatos, Phys. Rev. B **30**, 4571 (1984).

Walukiewicz, W., Phys. Rev. B **41**, 10218 (1990).

Walukiewicz, W., in *Semiconductor Interfaces and Microstructures*, edited by Z. C. Feng (World Scientific, Singapore, 1992), p.1.

Wetzel, C., W. Walukiewicz, E. E. Haller, J. Ager III, I. Grzegory, S. Porowski, and T. Suski, Phys. Rev. B **53**, 1322 (1996).

Wickenden, D. K., W. A. Bryden, Inst. Phys. Conf. Ser. No. 137, 381 (1993).

Wilamowski, Z., K. Swiatek, T. Dietl, J. Kossut, Solid State Comm. **74**, 833 (1990).

Wisk, P. W., C. R. Abernathy, S. J. Pearton, F. Ren, Mater. Res. Soc. Symp. Proc. **282**, 599 (1993).

Yamamoto, A., M. Tsujino, M. Ohkubo, and A. Hashimoto, Solar Energy Mat. and Solar Cells **35**, 53 (1994).

Yi, G.-C. and B. W. Wessels, Appl. Phys. Lett. **69**, 3028 (1996).

Yoshida, S., S. Misawa, S. Gonda, J. Appl. Phys. **53**, 6844 (1982).

Zhang, X., P. Kung, A. Saxler, D. Walker, T. C. Wang, and M. Razeghi, Appl. Phys. Lett. **67**, 1745 (1995).

Zook, D., Phys. Rev. **136**, 869 (1964).

Appendix A. Effective Mass Theory

Effective Mass Theory (EMT) is a formalism for calculating the electronic structure properties of impurities in semiconducting crystals. First developed by Kittel and Mitchell (Kittel 1954), the numerical results that can be obtained from EMT are surprisingly accurate given its simplicity. The basic mathematical manipulations of this formalism are outlined in this section.

Electrons in Perfect Crystals

Before attempting to describe the electronic states of impurities in crystals, one must first be able to describe the electronic states of a perfect crystal. These states can in principle be calculated from a many-body Hamiltonian which takes into account all of the interactions between all of the particles in a crystal. As the number of interacting bodies is on the order of 10^{22} cm^{-3} however, the exact problem is quite insoluble. In order to make the problem more tractable, certain approximations can be made to simplify its analysis.

One simplification is known as the Born-Oppenheimer approximation. There are two types of particles in a semiconductor crystal, lattice ions and electrons. Since the former are so much more massive than the latter and move much more slowly, the lattice and electronic degrees of freedom can be treated separately in this problem. In effect, the electronic problem is solved assuming that the lattice ions are frozen in a particular position. If desired, the ionic problem can be solved separately and added to the electronic Hamiltonian as a perturbation.

Unfortunately, the number of electrons in a typical crystal is still on the order of Avogadro's number. To simplify the mathematical problem of dealing with so many particles, the many-electron Hamiltonian is replaced by a one-electron Hamiltonian, in which the interactions of one electron with all the other electrons and lattice ions are treated

in an averaged manner. Depending on the level of sophistication, this one-electron Hamiltonian may also take into consideration exchange and correlation effects of the electron gas.

In a perfect crystal, the one-electron Schrödinger equation is

$$\left(T + V^0\right)\psi^0_{n\mathbf{k}} = E^0_{n\mathbf{k}}\,\psi^0_{n\mathbf{k}} \tag{A.1}$$

where T is the usual kinetic energy term and V^0 is the potential felt by an electron in a perfect crystal. The parameter n is a band index and $\mathbf{k}$ is the wavevector of the electron. Plotting the energies as a function of $\mathbf{k}$ for each n produces the familiar semiconductor energy band structure. If the bands are far enough apart energetically so that any interaction between their states can be ignored, then the energies near a band minimum or maximum occurring at $\mathbf{k}_0$ can be found by expanding $E^0_{n\mathbf{k}}$ in a Taylor series about $\mathbf{k}_0$

$$E^0_{n\mathbf{k}} = E^0_{n\mathbf{k}_0} + \frac{\hbar^2}{2}\left(\mathbf{k}-\mathbf{k}_0\right)\mathbf{M}^{-1}\left(\mathbf{k}-\mathbf{k}_0\right) \tag{A.2}$$

where $\mathbf{M}^{-1}$ is the inverse of the effective mass tensor

$$\mathbf{M}^{-1} = \frac{1}{\hbar^2}\nabla_{\mathbf{k}}\cdot\left(\nabla_{\mathbf{k}}\,E^0_{n\mathbf{k}}\right)_{\mathbf{k}=\mathbf{k}_0} \tag{A.3}$$

As can be seen from (A.2), the maxima and minima are parabolic as a function of the wavevector.

A common choice for the electronic wavefunctions $\psi^0_{n\mathbf{k}}$ is the set of Bloch functions

$$\psi^0_{n\mathbf{k}} = e^{i\mathbf{k}\cdot\mathbf{r}}\,u^0_{n\mathbf{k}}(\mathbf{r}) \tag{A.4}$$

where $u^0_{n\mathbf{k}}$ is a function which has the same translational symmetry as the lattice and thus also as V^0. These functions form a complete orthonormal set in the sense that

$$\langle \psi^0_{n'k'} | \psi^0_{nk} \rangle = \delta_{n'n}\delta_{k'k} \tag{A.5}$$

The Bloch functions are extended in real space and have the same amplitude $|\psi^0_{nk}|^2$ in each unit cell of the crystal, as can be seen from the translational symmetry of u^0_{nk}.

These functions are by no means the only choice of basis, however. The Wannier functions, which are linear combinations of Bloch functions and localized in real space are another common choice of basis.

Electrons in Imperfect Crystals

If impurities are now introduced into the perfect crystal, the one-electron Hamiltonian can be modified as follows

$$H = T + V = T + V^0 + U \tag{A.6}$$

where V is the new one-electron potential written as the sum of the old potential of the perfect crystal and U, the impurity potential. Using this Hamiltonian to construct a new Schrödinger equation, two types of solutions can be found: those with an energy which is within the allowed bands of the perfect crystal and those with an energy within the band gap. The states with energies within the original bands are free electron solutions (if the energy is above the gap) or solutions corresponding to electrons involved in the lattice bonds (if the energy is below the gap). The states with energies within the band gap are those of electrons bound to the impurity atoms.

One can verify that these impurity bound states are spatially localized by recalling that any eigenfunction of this Hamiltonian must be expressible as a linear combination of the Bloch functions of the perfect crystal. Since only Bloch functions with an imaginary wavevector **k** can have energies within the band gap, those Bloch solutions either decay or

grow exponentially (eq. A.4). As the latter are impermissible on physical grounds, the true solutions must decay exponentially and are thus localized.

For impurity potentials with a finite spatial range such as so-called "deep centers," the impurity levels may or may not exist within the energy gap, depending on the details of the potential. For an infinite-range potential, such as the Coulomb potential however, it has been proven (Mott 1940) than an infinite number of impurity bound states exist in the gap either just above the lower band edge or just below the upper band, depending on the sign of the potential. These states are hydrogen-like and as such can be assigned a quantum number n. For the impurity potential of a donor, as n approaches infinity, the localized wavefunctions merge smoothly into the propagating solutions of the conduction band.

General Results for Impurity States

To solve for the impurity states, the impurity state wavefunctions are first expanded in terms of the Bloch functions as follows

$$\psi = \sum_{nk} F_{nk}\, \psi^0_{nk}(\mathbf{r}) \tag{A.7}$$

where the sum is over all bands n and all **k** points. The problem now becomes one of determining the coefficients F_{nk}. Combining this wavefunction with the Hamiltonian from (A.6), the resulting Schrödinger equation is

$$\left(T + V^0 + U\right) \sum_{nk} F_{nk}\, \psi^0_{nk}(\mathbf{r}) = E \sum_{nk} F_{nk}\, \psi^0_{nk}(\mathbf{r}) \tag{A.8}$$

This equation can be further manipulated by multiplying both sides by $\psi^0_{n'k'}$ and integrating over all space, using the orthonormality of the Bloch functions to simplify the result. Exchanging the primed and unprimed variables, we obtain

$$E^0_{nk} F_{nk} + \sum_{n'k'} \langle \psi^0_{nk} | U | \psi^0_{n'k'} \rangle F_{n'k'} = E F_{nk} \tag{A.9}$$

This is a set of coupled equations for the coefficients F_{nk}. To develop these equations further, the explicit form of the Bloch functions from (A.4) is substituted in for the $\psi^0_{n'k'}$'s and the product of the u^0_{nk}'s is expanded in terms of plane waves as follows

$$u^{0\,*}_{nk}(r) u^0_{nk}(r) = \sum_{\mathbf{G}} C^{nn'}_{kk'}(\mathbf{G})\, e^{i\mathbf{G}\cdot\mathbf{r}} \tag{A.10}$$

Since the function u^0_{nk} has the periodicity of the lattice, the only vectors G which contribute to the sum are those which are reciprocal lattice vectors. Performing this substitution produces

$$E^0_{nk} F_{nk} + \sum_{n'k'} \sum_{\mathbf{G}} C^{nn'}_{kk'}(\mathbf{G})\, U(\mathbf{k}-\mathbf{k}'-\mathbf{G}) F_{n'k'} = E F_{nk} \tag{A.11}$$

where U(**q**) is defined to be the Fourier transform of U(**r**)

$$U(\mathbf{q}) = \int d^3r\, U(\mathbf{r})\, e^{-i\mathbf{q}\cdot\mathbf{r}}. \tag{A.12}$$

Up to this point, no approximations have been made aside from the initial use of the one-electron Hamiltonian. In order to proceed further, however, some assumptions about the specific form of the impurity potential must be made. For reasons discussed in section 1, the impurity potential will be assumed to be a simple screened Coulomb potential

$$U(\mathbf{r}) = \frac{-e^2}{\varepsilon r} \tag{A.13}$$

We now examine two cases in which EMT is applied to donors: one in which there is an isotropic conduction band minimum at **k**=0, and one in which there are several equivalent but non-isotropic band minima away from the Brillouin zone center.

Results for a Single Minimum

By far the simpler of the two is the case of a single isotropic conduction band minimum at the center of the Brillouin zone. This is the situation in GaAs at atmospheric pressure, where in addition, the second lowest conduction band is far enough separated from the lowest band that its effects may be neglected. Thus, we do not consider the effects of other bands and the index n may be dropped from (A.11). The equation now becomes

$$E_{\mathbf{k}}^{0} F_{\mathbf{k}} + \sum_{\mathbf{k}'} \sum_{\mathbf{G}} C_{\mathbf{k}\mathbf{k}'}(\mathbf{G})\, U(\mathbf{k}-\mathbf{k}'-\mathbf{G}) F_{\mathbf{k}'} = E F_{\mathbf{k}} \tag{A.14}$$

Anticipating that the only coefficients with **k** near 0 (the band minimum) will contribute significantly to the solution, we make the assumption that

$$|\mathbf{k}-\mathbf{k}'| \ll \mathbf{G} \tag{A.15}$$

As a consequence, all terms in the summation in (A.14) for which $\mathbf{G} \neq 0$ may be dropped since

$$|U(\mathbf{k}-\mathbf{k}'-\mathbf{G})| \ll |U(\mathbf{k}-\mathbf{k}')| \tag{A.16}$$

which follows from

$$U(\mathbf{q}) = \int d^3 r \, \frac{-e^2}{\varepsilon r} e^{i\mathbf{q}\cdot\mathbf{r}} = \frac{-4\pi e^2}{\varepsilon q^2} \tag{A.17}$$

In addition, $C_{\mathbf{k}\mathbf{k}'}(0)$ can be now approximated by

$$C_{\mathbf{k}\mathbf{k}'}(0) \approx C_{\mathbf{k}\mathbf{k}}(0) = 1 \tag{A.18}$$

which follows indirectly from (A.10) and the orthonormality of the Bloch functions.

Next, the sum over **k'** is converted to an integral. This does not introduce any appreciable error since for macroscopic samples, the **k** points in the Brillouin zone are so closely spaced as to be essentially continuous.

Finally, we expand E_k^0 about **k**=0 to order k^2 as in equation (A.2). For a crystal with cubic symmetry, this yields

$$E_{\mathbf{k}}^0 \approx E_0^0 + \frac{\hbar^2 k^2}{2m^*} \tag{A.19}$$

where m* is the effective mass. Applying approximations (A.15) through (A.19), equation (A.14) becomes

$$\frac{\hbar^2 k^2}{2m^*} F(\mathbf{k}) + \int d^3k\, U(\mathbf{k}-\mathbf{k}')\, F(\mathbf{k}') = E_B\, F(\mathbf{k}) \tag{A.20}$$

where E_B is the binding energy of the impurity state given by $E - E_0^0$ and F_k has been re-written as F(**k**). This equation is isomorphic to the momentum-space Schrödinger equation for a particle of mass m* in a potential U. By defining

$$F(\mathbf{r}) = \int d^3k\, F(\mathbf{k})\, e^{i\mathbf{k}\cdot\mathbf{r}} \tag{A.21}$$

equation (A.20) can be transformed into the real-space equation

$$\left[\frac{-\hbar^2\nabla^2}{2m^*} + U(\mathbf{r})\right] F(\mathbf{r}) = E_B\, F(\mathbf{r}) \tag{A.22}$$

Since the potential (A.13) which we use is that of a hydrogen atom in a continuous dielectric medium, (A.22) is analogous to the Schrödinger equation for a hydrogen atom, the solutions of which are well known. The only differences are that the electron mass is

now m* and e^2 must be replaced everywhere by e^2/ε in order to take the dielectric response of the material into account. The bound state energies of the impurity are

$$E_{Bn} = E_0^0 - \frac{e^4 m^*}{2\hbar^2 n^2 \varepsilon^2} \tag{A.23}$$

and the ground state coefficient F_{nk} is given by

$$F(\mathbf{r}) = \frac{1}{\sqrt{\pi a^{*3}}} e^{-\frac{r}{a^*}} \tag{A.24}$$

where a* is an effective Bohr radius

$$a^* = \frac{\hbar^2 \varepsilon}{e^2 m^*} = a_0 \varepsilon \frac{m_0}{m^*}. \tag{A.25}$$

The complete ground state wavefunction can thus be approximated to first order as

$$\psi(\mathbf{r}) \approx F(\mathbf{r})\, u_0^0(\mathbf{r}) = F(\mathbf{r})\, \psi_0^0(\mathbf{r}) \tag{A.26}$$

as F is negligibly small for $\mathbf{k} \neq 0$.

At this point, we can now go back and test the validity of the approximation made in (A.15). The Fourier transform of (A.24) is

$$F(\mathbf{k}) = \frac{8\pi^{1/2}}{a^{*5/2}\left(k^2 + \frac{1}{a^{*2}}\right)} \tag{A.27}$$

which is appreciable out to approximately $|\mathbf{k}| = 1/a^*$. Condition (A.15) then becomes

$$|\mathbf{G}| >> \frac{1}{a^*} \tag{A.28}$$

Since reciprocal lattice vectors are on the order of $2\pi/a$ and typical lattice constants are roughly 10 a_0, this condition may also be expressed as

$$\frac{2\pi}{10} >> \frac{m^*}{m_0}\frac{1}{\varepsilon} \qquad \text{or} \qquad \varepsilon\frac{m_0}{m^*} >> 3.18 \tag{A.29}$$

In a typical semiconductor, the dielectric constant is about 10 and effective masses are approximately one tenth of the free electron mass, so in general, this condition is satisfied quite well.

Single Band with Several Equivalent Minima

Although the case with a single conduction band minimum is the easiest to calculate, many common semiconductors have several equivalent minima at or near the L or X symmetry points of the Brillouin zone. For example, the conduction band of Ge has four equivalent minima located at the Brillouin Zone boundary in the <111> directions and Si has six equivalent minima in the <100> directions. While equation (A.14) is still valid, it can no longer be assumed that F(**k**) is localized around any one of these minima. Instead, it must be written as a linear combination of coefficients

$$F_k = \sum_{i=1}^{N} \alpha_i F_k^{(i)} \tag{A.30}$$

where N is the number of band extrema and each $F_k^{(i)}$ is localized about the ith minimum. Substituting this expression into (A.14) produces

$$E_{\mathbf{k}}^0 \sum_i \alpha_i F_k^{(i)} + \sum_{\mathbf{k}'}\sum_{\mathbf{G}} C_{kk'}(\mathbf{G})\, U(\mathbf{k}-\mathbf{k}'-\mathbf{G}) \sum_i \alpha_i F_k^{(i)} = E \sum_i \alpha_i F_k^{(i)} \tag{A.31}$$

This equation is clearly much more complicated to solve than that for the single extremum case because there are now two types of coupling terms--the intravalley terms which

describe the interaction between states of the same minimum (valley) and the intervalley terms which couple states from different minima. The simplest way to deal with this complication is to neglect the intervalley terms, hoping that the $F_k^{(i)}$ at different extrema do not overlap appreciably. Under this approximation, (A.31) becomes a set of independent hydrogenic equations similar to (A.20) or (A.22) except that the kinetic energy is now anisotropic. Defining the z-axis to be along the direction from the band minimum to the center of the Brillouin zone, the energy term can be expressed as

$$E_{\mathbf{k}}^0 = E_0^0 + \frac{\hbar^2}{2}\left(\frac{k_x^2 + k_y^2}{m_t^*} + \frac{k_z^2}{m_l^*}\right) \tag{A.32}$$

There are now two effective masses. The longitudinal mass is inversely proportional to the curvature of the band parallel to the z-axis and the transverse mass corresponds to the curvature perpendicular to the z-axis at the band minimum. The analog of (A.22) in this case is then

$$\frac{-\hbar^2}{2}\left(\frac{1}{m_t^*}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right) + \frac{1}{m_l^*}\frac{\partial^2}{\partial z^2}\right)F^{(i)}(\mathbf{r}) + U(\mathbf{r})F^{(i)}(\mathbf{r}) = E\,F^{(i)}(\mathbf{r}) \tag{A.33}$$

where, as usual

$$F^{(i)}(\mathbf{r}) = \int d^3k\; F^{(i)}(\mathbf{k})\, e^{i(\mathbf{k}-\mathbf{k}_i)\cdot\mathbf{r}} \tag{A.34}$$

The complete wavefunctions are linear combinations of the solutions to the effective-mass Schrödinger equation at each band minimum and to first order in $\mathbf{k}$, they are

$$\psi(\mathbf{r}) = \sum_i \alpha_i\, F^{(i)}(\mathbf{r})\, \psi_{\mathbf{k}_i}^0(\mathbf{r}) \tag{A.35}$$

All valleys must be included in the wavefunctions due to symmetry considerations, even though they have been assumed not to interact.

Solutions to this anisotropic effective mass equation can only be obtained by a variational calculation. Even so, Faulkner (Faulkner 1969) was able to obtain results which agreed quite closely with experimental observations by using linear combinations of anisotropic hydrogenic wavefunctions as the trial solutions.

Appendix B. Neutron Transmutation Doping of GaAs

Neutron transmutation doping (NTD), or the doping of semiconductor materials through irradiation by neutrons, is the best technique available for producing crystals with extremely uniform doping profiles. NTD makes use of the fact that certain stable atomic isotopes can transmute into a different element entirely after absorbing a neutron and decaying via various mechanisms such as beta decay. Since these stable isotopes are distributed uniformly throughout the material, the doping atoms which are the product of the transmutation are also uniformly distributed. Irradiation is achieved by placing the crystal to be doped near the core of a nuclear reactor. Although NTD is most commonly used to dope silicon, this process was used in this work to introduce Ge and Se dopants into GaAs.

Naturally occurring GaAs consists of two isotopes of Ga (^{69}Ga and ^{71}Ga) and a single isotope of As (^{75}As). The capture cross sections of these elements for thermal neutrons (neutrons with kinetic energies of roughly tens of meV) are 1.68 barns, 4.86 barns, and 4.30 barns, respectively. Upon absorbing a neutron, each of these three elements transmutes as follows:

$$\begin{aligned} &^{69}Ga(n,\gamma)\,^{70}Ga \xrightarrow{21\text{ min}} {}^{70}Ge + \beta^- \\ &^{71}Ga(n,\gamma)\,^{72}Ga \xrightarrow{14\text{ h}} {}^{72}Ge + \beta^- \\ &^{75}As\,(n,\gamma)\,^{76}As \xrightarrow{26\text{ h}} {}^{76}Se + \beta^- \end{aligned} \qquad \text{(B.1)}$$

All three of the newly formed isotopes transform into new elements by β decay. If the resultant Ge atoms are found on Ga sites and the Se atoms on As sites, then all of these end products will act as donors. It is in principle possible for the Ge atoms to be found on the As sublattice as well and to act as compensating acceptors. However, previous experiments have shown that a large majority of the Ge atoms produced through NTD

remain on Ga sites (Vesaghi 1982). Following irradiation, the crystals must be stored in a secure environment until the radioactivity decreases to a safe level.

Based on the capture cross sections of the three isotopes listed above, the concentration of donors resulting from NTD can be expressed as

$$\begin{aligned} n(^{70}Ge) &= 0.023\,\phi\,t \\ n(^{72}Ge) &= 0.042\,\phi\,t \\ n(^{76}Se) &= 0.107\,\phi\,t \end{aligned} \qquad \text{(A.2)}$$

where ϕ is the neutron flux and t the time of exposure to that flux.

Up to now, only the effect of thermal neutrons has been considered. In a typical reactor, neutrons with energies ranging from tens of meV up to the MeV regime are present. However, by moderating the neutron flux, the ratio of thermal to fast neutrons can be made as great as a few thousand. For doping by NTD, thermal neutrons are preferable to fast neutrons for two reasons. First, the capture cross section of these elements for neutrons varies inversely with the neutron velocity and so slow neutrons are more efficient at doping than fast neutrons. Second, fast neutrons damage the crystal by the creation of Frenkel pairs. Since the energy to displace a lattice atom is only on the order of a few tens of electron volts, energetic neutrons with hundreds of keV can displace many hundreds of lattice atoms through collisions with their nuclei. These displaced atoms can generate many deep levels within the band gap. In addition, the gamma rays emitted by atoms which have captured a thermal neutron cause the irradiated atom to recoil with a kinetic energy on the order of hundreds of eV and thus these atoms also become displaced from their substitutional position. For this reason, neutron irradiated crystals are found to have very high resistivities and low carrier concentrations.

In order to repair the lattice damage, the NTD crystals are annealed. Studies have shown that the annealing of irradiated epitaxial layers of GaAs at 600°C for 5 minutes is

sufficient to remove all structural defects detectable by Deep Level Transient Spectroscopy (DLTS) (Alexiev 1993). Of course, care must be taken to insure that rapidly diffusing contaminants such as Cu are not introduced into the crystal during the anneal.

In the NTD doping of epitaxial layers such as was performed here, one must also concern oneself with the possible products present in the substrate, as it is much less pure than the layer and unexpected products may be formed through irradiation. For diamond anvil cell work, however, the substrates were always removed before spectroscopy was performed.

Ernest Orlando Lawrence Berkeley National Laboratory
One Cyclotron Road | Berkeley, California 94720

Prepared for the U.S. Department of Energy under Contract No. DE-AC03-76SF00098