Infra-red Optical Properties of SiGe/Si Heterostructures

by Elizabeth A. Corbin

A theoretical thesis submitted to the University of Newcastle-upon-Tyne for the degree of Doctor of Philosophy.

September 1995

Dedicated to Adam

Declaration

This thesis has not previously been submitted by the candidate for a degree in this or any other university.

E.A. Corbin

September, 1995

Abstract

We present full-scale relativistic pseudopotential calculations of the first-order susceptibility in p-type SiGe/Si multiple quantum well structures with a view to exploring the suitability of such systems for infrared applications in the 3-5μm and 8-15μm ranges. A derivation of an expression for the linear susceptibility, or absorption, is given and the frequency dependence of the linear response due to transitions between the valence minibands is determined.. The microscopic origin of the absorption is demonstrated for both parallel and normal incident light. Comparisons between calculated and experimental results are presented and shown to be in good agreement. The effects of changing well width, temperature, doping concentration and germanium concentration in the well are considered.

We also consider Auger recombination and discuss the possibility of engineering the miniband structure in order to prevent certain Auger processes occuring, Preliminary results from full scale Auger calculations are also presented.

Some of the work presented in this thesis has been published in the following papers:

1. Absorption in p-type Si-SiGe Strained Quantum Well Structures, E. Corbin, K.B. Wong and M. Jaros, Phys. Rev. B., **50**, 2339 (1994)

2. Auger-free Si-SiGe Quantum Well Structures for Infra-red Detection at 10μm, Solid State Electronics, E. Corbin and M. Jaros (to be published.)

3. Absorption and Recombination in p-type SiGe quantum well structure. E. Corbin, M. Cusack, K.B. Wong and M. Jaros, Superlattices and Microstructures, **16**, 349, (1994)

4. GaAs-AlAs and Si-SiGe Quantum Well Structures for Applications in Non-linear Optics, M.J. Shaw, K.B Wong, E. Corbin and M. Jaros, Solid State Electronics, **37**, 1303, (1994)

5. Optical Properties of Mesoscopic Si-Ge Hetrostructures, M. Jaros, J.P. Hagon, R.J. Turton, A. Miloszewski, K.B. Wong and E. Corbin, Physics Scripta, **T45**, 174, (1992)

Acknowledgments

I would particularly like to thank my supervisor Prof. Milan Jaros for his friendly guidance and participation throughout the course of this research. My thanks also go to the rest of the members of the Theory of Condensed Matter group at Newcastle, in particular Dr. Mike Shaw without whom I would never have got this far.

This work was financially supported in by the Defence Research Agency, Malvern, U.K.

Contents

Chapter 1

Infra-red Detectors

1.1 Introduction

Infra-red (IR) radiation was discovered by Herschel (1800) using a mercury glass thermometer to detect sunlight that had been dispersed by a prism. The IR range of wavelengths covers the region $\approx 1 - 1000\mu$m. The basic characteristics of this range of wavelengths which makes it of interest are that it does not penetrate metals but passes through many crystalline and gaseous materials, in particular the earth's atmosphere. There are two windows in the atmosphere in which IR absorption is at a minimum, i.e. $8 - 15\mu$m and $3 - 5\mu$m, which are of particular interest.

In this chapter we review the development of long wavelength IR detectors from the earliest thermal detectors to present day quantum well systems. We discuss various figures of merit which can be used to compare detectors of different types and the detection mechanisms for each type of detector. We also aim to give an indication of the current state-of-the-art for each system.

1.1.1 Uses of Infra-red Detectors

Efforts have been made over many years to create a sensor which can 'see' in the infra-red range of wavelengths, in an analogous manner to the human eye. The present day demand for IR detectors is huge and they are put to a wide variety of uses. Presently, most equipment is used by the military who are undeterred by the high cost and power usage of the current devices. IR detectors can be used for thermal imaging, guidance, reconnaissance, surveillance and communication systems. In civilian life they can be employed for such varied tasks as an improved night sight system for drivers or aircraft, to satellite research mapping, non-destructive testing and heat flow problem analysis. The IR range of wavelengths is particularly appropriate for monitoring the composition of the atmosphere and molecular clouds as the absorption lines of oxygen, water, carbon monoxide and nitrous oxide all lie between 3-18 μm.

There are also many space-related and astronomy uses. IR radiation is particularly good for observing planetary systems as in this range of wavelength the star and its circling planets are of approximately the same brightness. The future requirements for energy efficiency will make infrared imaging a routine technique for discovering potential energy leaks in housing and other structures. Infrared detectors have also been used with success to count deer and seals. The same technique was not effective with elephants, however, due to a tendency of the elephants to stand under trees! (Prinizivalli 1992). The move away from military applications means that the cost and power consumption of the devices must be decreased, and ease of use must be improved. However, lower device performance should be acceptable.

1.1.2 Arrays

To create a useful imaging system large numbers of single detectors must be formed into arrays such that each array element (pixel) contributes to the overall response. Arrays are formed by methods similar to processes used for silicon integrated circuits. A typical sequence for array manufacture would be (Broudy 1981): a slab of high quality material is bonded to a substrate and then polished and etched to optimum thickness (i.e. thick enough to absorb most of the optical radiation and thin enough to minimize bias current). The array is then delineated using a photolithographic process with etching, and external electrical contacts are attached by means of wire bonding.

There are two different approaches that can be used to create imaging systems. Scanned Imagers were developed many years ago and use a system of rapidly rotating precision mirrors to scan the entire target scene sequentially across a very small number of IR sensors, usually a linear array (Wood 1993). The detectors used need to be highly sensitive with a fast response time. At the moment they need to be cryogenically cooled (to less that 100K) in order to obtain useful sensitivity. This is because the dwell time is very short. This method is used because comparatively few sensors are needed, thus reducing the cost. This type of array is currently made from either HgCdTe or InSb operating at 100K. The typical cost of such an array in 1993 was $100,000 (Wood 1993).

An alternative approach is to use staring imagers where a large 2-D array of sensors, typically an array of 256×256 elements, is placed in the focal plane of the lens, with one IR detector for each pixel in the display. No mirrors are involved and thus the response time of the detector can be longer; up to 100 times that required for scanning arrays. The relative sensitivity is therefore increased by approximately an order of magnitude. Scanned imagers were

popular when detectors were very expensive but for high quality focal plane applications staring arrays are required. This produces an additional quality requirement on the detectors in that the pixel-to-pixel variation must be small in order to produce a high quality image. The future of array design is in staring arrays so it is these we shall concentrate on in this chapter.

1.1.3 Noise

All detectors produce noise, which is defined as the electrical voltage or current output from the detector that is not coherent with the radiation signal power. In general, the detector noise is a function of the bias, modulation frequency and the area of the detector. In a practical set up there are also contributions to the total noise from radiation noise and noise produced by the amplifiers in the circuit. For a general discussion of noise in devices see Eismann (1977) and references therein.

In thermal detectors the active element will have a statistical interchange of energy with its surroundings, and hence random temperature fluctuations will occur which set a lower limit to the detectable radiation.

In photoconductive detectors there are several sources of noise, 1/f noise, Johnson noise, noise caused by dark current and noise resulting from the fluctuation of the number of photons (photon noise).

Johnson noise is due to the random thermal motion of the charge carriers in the crystal. This noise is essentially due to small changes in the voltages at the terminals of the device due to the random arrival of charge at the terminals. For normal frequencies of operation of IR detectors, i.e. $h\nu \gg kT$, this noise contribution is independent of frequency and current but increases with temperature.

1/f noise. Many electrical devices exhibit excess noise, large at low fre-

quency and decreasing as the frequency is increased. The exact physical mechanism responsible for this noise varies from device to device and is not fully understood.

Generation-Recombination (g-r) Noise also known as *shot* noise, is noise due to the random generation of free charge carriers by the crystal lattice vibrations and their subsequent random recombination. Because of the randomness of these events there is unlikely to be exactly the same number of charge carriers in the free state at succeeding instances of time. This fluctuation leads to conductivity changes that will be reflected as variations in the current flow through the crystal.

Photon Noise occurs when free charge carriers are generated by external photons rather than lattice vibration, fluctuations in generation rate will be caused by the random arrival of photons at the detector. The detector thus reflects the fluctuations in the photon stream. The main source of this noise is background radiation falling on the detector.

Dark Current, is the current that flows in the circuit when the detector is not illuminated. All photodetectors produce dark current, i.e current generated from sources other than the target. There are three main contributions to the dark current in quantum wells. Thermionic emission where carriers are thermally excited into current carrying states above the top of the well. This contribution varies exponentially with temperature and is the dominant source of dark current at temperatures above 45K. Secondly there is temperature dependent tunnelling, which is the tunnelling of carriers into the neighbouring wells for low lying well states. This contribution only depends weakly on temperature and is observed when the temperature is below that necessary for thermionic emission, thus this term is dominant below 30K. Finally there is thermally assisted tunnelling where a thermal excitation is followed by tun-

nelling from the excited stated. This contribution is very small compared to the thermionic emission except at very high fields.

Ways to reduce noise include using the lowest feasible bias current, keeping the detector temperature below the thermal generation rate, reducing the detector background radiation as much as possible and improving processing methods to reduce impurities, dislocations etc. Contributions from the dark current and photon noise often limit the detectors ultimate performance.

1.1.4 Figures of Merit

In order to fully describe a detector's characteristics a large number of parameters need to be known, namely the detector signal, noise, bias and temperature as well as the radiation signal power, the physical and electrical properties of the detector and the background conditions under which the detector is operated (Eisenman 1977). This obviously becomes a lengthy process so in order to ease comparison between various systems certain figures of merit have been defined, the most commonly used of which are the responsivity, R, the detectivity, D*, and the quantum efficiency, η.

Detectivity, D*, is the signal to noise ratio normalized to unit area and unit bandwidth.

$$D^*_\lambda = \frac{R_\lambda \sqrt{A_p}}{i_n} \qquad cmHz^{\frac{1}{2}}W^{-1} \tag{1.1}$$

where A_p is the photosensitive area of the detector, i_n is the noise current defined as

$$i_n = \sqrt{4egI} \qquad AHz^{\frac{1}{2}} \tag{1.2}$$

where I is the total current in the device and g is the photo-conductive gain. The responsivity R_λ is defined as

$$R_\lambda = \frac{I_s}{P_\lambda} = \frac{e\lambda\eta g}{hc} \approx 0.8\lambda\eta g \tag{1.3}$$

where η is the absorbance of the detector, sometimes known as the quantum efficiency, I_s is the r.m.s. signal current from a monochromatic source, and P_λ is the r.m.s. signal power.

The quantum efficiency is defined as the efficiency of the process whereby a photon stream incident on the detector crystal is converted to free electrons (or holes) within the crystal. Its values range from 0 to 1, or 0% to 100%, corresponding to no photons absorbed in the crystal and all incident photons absorbed respectively. To optimize detector performance it is desirable to maximize the quantum efficiency.

Pixel to pixel variation cannot be fully described by single pixel characteristics and 'spatial' noise caused by such variation must be accounted for. In practical applications NEΔT is a general figure of merit for describing arrays, defined as the noise equivalent temperature difference i.e. the minimum temperature difference across the target that would produce a signal to noise ratio of unity. A general expression was derived by Shepard (1988),

$$NE\Delta T = \frac{N_n}{dN_b/dT_b} \tag{1.4}$$

where T_b is the background temperature, N_n is the total number of noise electrons per pixel, i.e.

$$N_n^2 = N_t^2 + N_b + u^2 N_b^2 \tag{1.5}$$

where N_t is the photoresponse-independent temporal noise and u is the residual non-uniformity after correction by electronics, N_b is the shot noise from IR background radiation. When the uniformity term dominates this expression can be approximated to (Levine 1993)

$$NE\Delta T = \frac{u\bar{\lambda}T_b}{1.44} \tag{1.6}$$

where $\bar{\lambda}$ is the average wavelength. In the spatial noise limit $NE\Delta T$ is proportional to u, the non-uniformity. Thus for high detectivities, i.e $D^* > 10^{10}$cm

$Hz^{\frac{1}{2}}$ W^{-1} the performance of the device is limited by the degree of non-uniformity and NEΔT is the relevant figure of merit.

1.1.5 Background Limited Performance

An ideal background limited performance (BLIP) device is one in which the number of noise carriers, N_n, is dominated by the shot noise of the optically generated carriers. If this condition can be achieved then the maximum possible detectivity will result. The temperature at which this condition is reached, T_{BLIP}, is a sensitive function of the dark current. Thus to operate a device at the highest possible temperature without degrading device performance to an unacceptable degree, the dark current must be minimized.

1.2 Thermal Detectors

We are now in a position to compare various types of detectors and discuss their potential for the creation of practical IR imaging systems.

There are two main types of detector; thermal detectors and photodetectors. One distinguishing feature between thermal and photodetectors is that in thermal detectors the excitations generated by photon absorption relax to a thermal distribution at an elevated temperature in the thermometer before they are detected. In photodetectors, it is the non-thermal distribution of excited carriers that is detected before it relaxes. The output of a photodetector is governed by the rate of absorption of photons not on the photons' energy.

Another important difference between thermal detectors and photodetectors is that thermal detectors accept or 'see' radiation from the entire electromagnetic spectrum whilst photon detectors cut off the longer wavelengths at their sharp absorption edges at λ_c. In photon detectors the speed of response

is governed by the by the free carrier lifetime, typically less that 10^{-6}s. For thermal detectors, however, the response time is related to the thermal time constant which is of the order of 10^{-3}s. In photodetectors the non-equilibrium charge carrier concentration generated by photo-ionization can easily be made to exceed the dark carrier concentration by cooling the device to sufficiently low temperatures.

Thermal detectors typically operate at room temperature but are unselective slow and of low sensitivity. They are, however, cheap and easy to use with no cooling requirement. The lack of sensitivity can be overcome by creating arrays of large numbers of detectors (Wood 1993). A review of thermal detector technology is given by Richards (1994).

All thermal detectors include an absorbing element which converts incident electromagnetic radiation into heat. The absorption of light raises the temperature of the device and thus results in a change of some temperature dependent parameter, e.g. the conductivity. Therefore the output of the thermal detector is proportional to the amount of energy absorbed per unit time and, provided that the absorption efficiency is the same at all wavelength, it is independent of wavelength.

The radiation thermocouple is one of the most widely used thermal detectors, principally for infrared spectroscopy. Typical D*'s range from 10^8-10^9 cm $\mathrm{Hz}^{\frac{1}{2}}\mathrm{W}^{-1}$. Another common type of thermal detectors is a bolometer, developed by Langley (1881), in which the incident radiation heats a fine wire or magnetic strip causing a change in its electrical resistance. Pt and Ni are common materials used for the wires. Carbon resistance bolometers operating at 4.2K have proved very successful for IR astronomy where very sensitive detectors are required (Richards 1995). An alternative approach is the Golay pneumatic detector (Golay 1947) in which heat is absorbed in a thin metal

film then transfered to a small volume of gas. The resulting pressure increase changes the angle of a mirror in an optical amplifier. A third approach is to use a pyroelectric material whose dielectric constant is a sensitive function of temperature. Both these types of devices have been used in IR spectroscopy since the 1960's (Haoni 1963).

Recent developments include microbolometers fabricated from micromachined Si. These detectors can attain thermal isolation (the limiting factor in thermal detector performance) approaching the theoretical maximum and also demonstrate excellent IR sensitivity. NEΔT's of better that $0.05°C$ have been demonstrated (Wood 1993).

The current state of the art is an uncooled array of approximately 80,000 pixels on a 50μm centre consisting of a vanadium oxide resistive material deposited in thin film form on an Si microstructure containing transistor switches at each pixel. A NEΔT of 0.039°C in the 8-12 μm range was observed (Hanson 1993).

1.3 Photodetectors

Photodetectors are semiconductor devices which can detect optical signals through electronic processes resulting in a fast and selective response. Radiation is absorbed within the material by interaction with electrons bound to lattice or impurity atoms or with free electrons. The observed electrical output signal thus results from the changed electron energy distribution. The exact nature of this interaction allows us to further subdivide IR photodetectors into four distinct categories: intrinsic, extrinsic, photoemmisive and quantum well detectors. We shall briefly study these first three categories before concentrating our attention on quantum well systems.

1.3.1 Intrinsic Detectors

Radar technology developed in World War II but IR technology was still in its infancy. Only active IR image converters and single element PbS devices (Cashman 1946) were available. Neither of these options was capable of performing passive IR imaging. The lead salt family were explored, (Cashman 1959), including PbSe and PbTe which were used in missile guidance and 3-5 μm thermal imaging. III-V semiconductors were discovered (Rieke 1959) and InSb-based systems were developed. InSb operates at 77K but only in the 3-5 μm range. What was needed was a device that would operate at this higher temperature but in the 8-12μm range, i.e. an InSb-like material but with half the band gap. Several candidate were considered, grey Sn, HgSe and HgTe all of which turned out to be semi-metallic. In 1959, Lawson (1959) published results that $Hg_{1-x}Cd_xTe$ exhibited semiconducting properties over much of the range of composition, ranging from a wide gap semiconductor at x=1 to a semimetal at x=0. The band gap of this tertiary alloy can be varied by altering the concentrations of HgTe and CdTe. The band gap varies approximately linearly with composition. Thus the band gaps and hence the wavelength response can be tuned to any wavelength between 1-20μm. By 1965 HgCdTe photoconduction IR technology had advanced sufficiently for prototypes of imaging systems and small array detectors were developed. In the 1970's rapid progress was made, detector technology matured and other compositions suitable for applications in the 1-3μm, 3-5μm and 15-30μm ranges were investigated (Riene 1977). This system has been extensively studied, (see, for example, Long 1970, Harman 1974), and a review of the development of HgCdTe can be found in Kruse (1981). Photo-conductive detectors of HgCdTe are the current industry standard for IR applications.

An intrinsic photodetector works by absorption across the fundamental

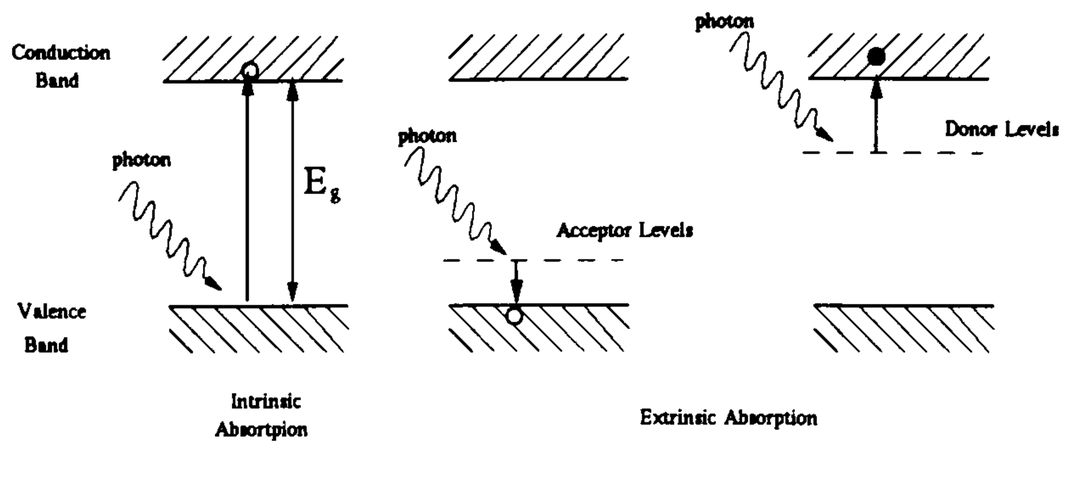


Figure 1.1: *The absorption mechanisms for intrinsic and extrinsic absorption.*

band gap, i.e. an incoming photon excites an electron from the valence band into the conduction band as shown in figure 1.1. Absorption leaves a hole in the valence band. These photoexcited carriers are collected to produce a photocurrent or photovoltage. As the absorption of a photon creates an electron and a hole, there is no need for doping. The advantage of an intrinsic system is its low thermal generation of carriers i.e. there are no dopant levels in the fundamental gap with excitation energy comparable to kT. Thus they can be operated at room temperature and there is no need for expensive cooling devices. $Hg_{1-x}Cd_xTe$ has been used for photo-conductors, junction photo-diodes, charge coupled devices and metal-oxide-semiconductor field effect transistors, MOSFETS. Each of the above can be used for IR detection, the choice depends on the specific application.

A typical detectivity for a HgCdTe system at 80K with a cutoff wavelength of $\approx 10\mu$m is $D^* = 3 \times 10^{12}$ cm $Hz^{\frac{1}{2}}$ W^{-1}. The advantages of HgCdTe are that the system is lattice matched, a large range of band gaps is available and a high operating temperature is permitted. It also exhibits high optical absorption coefficients. The disadvantages are that in the wavelength range of interest this alloy is a narrow gap semiconductor and thus is difficult to grow and the change in wavelengths with alloy composition is rapid when the wavelength exceeds

10μm. Also it is not compatible with Si technology and thus integrating into very large arrays will prove problematic. The reproducibility of devices is poor and thus large arrays have high pixel-pixel variation resulting in a large spatial noise. This system also suffers from large dark tunnelling currents and Auger recombination rates that exceed radiate recombination rates.

Progress in array size is intimately linked to the development of improved crystal growth methods and techniques for interconnecting to Si readout circuitry. Liquid phase epitaxy, LPE, has allowed increasingly large arrays in both medium and long wavelength bands. Array sizes of up to 640×480 have been developed (Kozlowski 1994). Recent developments have been the successful growth of HgCdTe on Si substrates (Wang 1995).

There are other compounds which can be used to fabricate intrinsic detectors. The most well know of these is InSb, for which arrays of size 256×256 have been demonstrated (Hackwell 1993). However, this system has a maximum wavelength range of up to 5μm as the band gap is fixed. There has also been recent studies of other ternary alloy systems, namely InAsSb, HgZnTe and HgMnTe (Rogalski 1989a 1989b and 1991 respectively.)

1.3.2 Extrinsic Detectors

Impurity photo-conductivity was recognized at an early date as a good candidate for sensitive IR detection. An review of the properties of Si and Ge extrinsic detectors is given by Bratt (1977). Impurity ionization levels in Ge range from 0.01eV (Al, Sb and others) to 0.35eV (Fe), and in Si from 0.04 (B) to 0.55eV (Cd). Therefore the cutoff wavelength for extrinsic photo-conductivity ranges from 2-120μm, which is by far the largest wavelength span of any set of devices. It is this great versatility which has been the key reason for interest in these systems. These detectors were first studied in the 1950's using Ge and

Si. Interest was in the 3-5μm and the 8-14μm atmospheric windows, thus research concentrated on Si with dopant energy levels of 0.04-0.06eV. The first published results were by Burstein (1954). Deep level impurity atoms were then discovered and research shifted to Ge doped with Au, Cu and Zn. IR devices consisting of Ge:Zn and Ge:Cu were demonstrated and developed into very sensitive detectors. (Bratt 1961) These detectors needed cooling to 4K, however, which severely limited applications. In 1957 it was discovered that by introducing impurities into SiGe alloys a new range of ionization energy levels was available (Morton 1959). Ge:Hg was discovered in 1962 (Borrello 1962) which has the required 90meV ionization energy and is easily reproducible. This has been used as an IR detector in the 8-14 μm range for many years.

The device mechanism is very simple. IR radiation is absorbed by an impurity atom when an incoming photon has sufficient energy to excite an electron (hole) residing in the ground state of a donor (acceptor)-type impurity into either an excited state or into the conduction (valence) band free of the impurity binding energy (see figure 1.1). There is also the process of photothermal conductivity where the electrons in the excited states interact with phonons and are thus excited into the conduction or valence band.

There has been a relatively recent resurgence of interest in Si extrinsic detectors due to their compatibility with existing Si technology. Fouks (1994) demonstrated 36 and 48 element linear arrays of Si:Ga, and showed that deeply cooled extrinsic Si and Ge IR detectors posses detectivity close to the radiation limit even under extremely low background irradiance, therefore these photodetectors are the main ones used in space-based applications where the background current very low. The wavelength detection range can be extended to 240μm by introducing stress, by means of a screw and piston, to lift the valence band degeneracy (Kazanski 1979, Haller 1979)

1.3.3 Photoemissive Detectors

These can be divided into two types; metal-silicide Schottky barrier devices and semiconductor hetrojunction internal photoemission (HIP) detectors.

1.3.3.1 Schottky Barriers

Rapid progress has been made with silicide Schottky barriers, the most developed of which is PtSi, (Akiyama 1994, Chin 1993) which employs films of platinum approximately 10 Å thick on p-type silicon. Radiation is transmitted through the Si and absorbed in the metal producing hot holes which are then emitted over the potential barrier into the Si which becomes negatively charged. In area arrays this charge is transferred into a charge coupled device, CCD, by the direct charge injection method. The effective quantum efficiency of such a device is approximately 1% in the 3-5 μm range. The dark current in this device is caused by the thermionic emission of holes over the potential barrier, thus the cooling requirements are similar to extrinsic device, i.e. performance is greatly enhanced by lowering the temperature.

The large advantage of this system is its compatibility with existing standard integrated circuit Si processing. Schottky photoemission is independent of semiconductor doping, minority carrier lifetime and alloy concentration leading to spatial uniformity characteristics an order of magnitude greater than that obtained for any other types of detector, and thus the spatial noise, which is a product of pixel-pixel variation in large arrays, is greatly reduced. The signal conditioning electronics required for staring arrays are therefore simplified. Other advantages of this system are that it is cheap to make and the small size of the detector (approximately $17\mu m^2$) which is the smallest among the 2-D focal plane array detectors. Responsivity of image sensors is proportional to the fill factor, i.e. the ratio of the detector area to the pixel

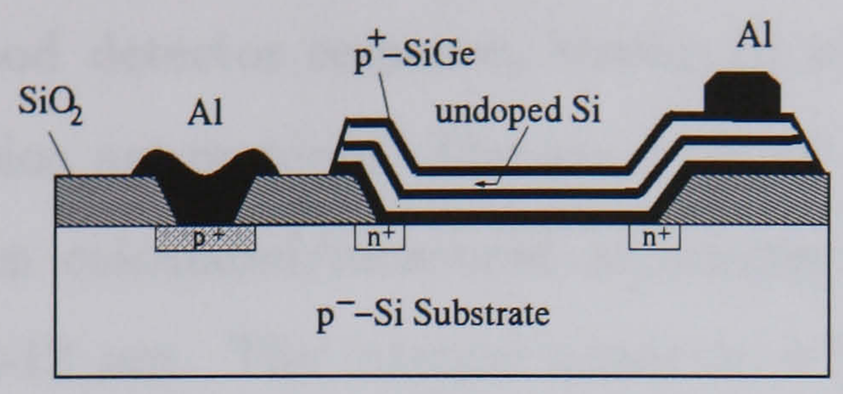
size. Improvement of this ratio is one of the most important developments in Schottky barrier image sensors. The state of the art of PtSi is 1040x1040 element arrays in the 3-5μm IR region (Kimata 1991, Akiyama 1994). Applications are limited to 1-3μm and 3-5μm ranges due to fixed PtSi cutoff at 5.7μm.

Other materials can be used to create this type of device, the most popular of which is IrSi which extends cutoff wavelength to 10μm (Tsaur 1988). However, the uniformity of IrSi arrays is substantially inferior to PtSi due to poor process repeatability (Tsaur 1989).

1.3.3.2 Hetrojunction Internal Photoemission Detectors

These devices were first constructed of a single layer of degenerately doped p^+-SiGe layer as an emitter and a p-type Si substrate as a collector. The detector mechanism is that incident IR radiation is absorbed in the p^+ SiGe emitter layer mainly through free carrier absorption. The photoemitted holes are then sent over the SiGe/Si hetrojunction barrier by internal photoemission. The range of wavelength for the device is determined by the height of the barrier and the Fermi energy and thus can be altered by changing the alloy composition and the doping levels. A device of this type has been demonstrated by Lin (1990 and 1991). These detectors are particularly suited for creating arrays operating in the 8-11 μm range. They are compatible with Si and have good pixel-pixel uniformity.

These detectors extend the wavelength range possible with Schottky barrier devices and have the potential for high quantum efficiency. In Schottky detectors photons can excite carriers from far below the Fermi level such that they do not have sufficient energy to overcome the barrier. If operating near the threshold, only a small percentage of the photoexcited carriers can ex-



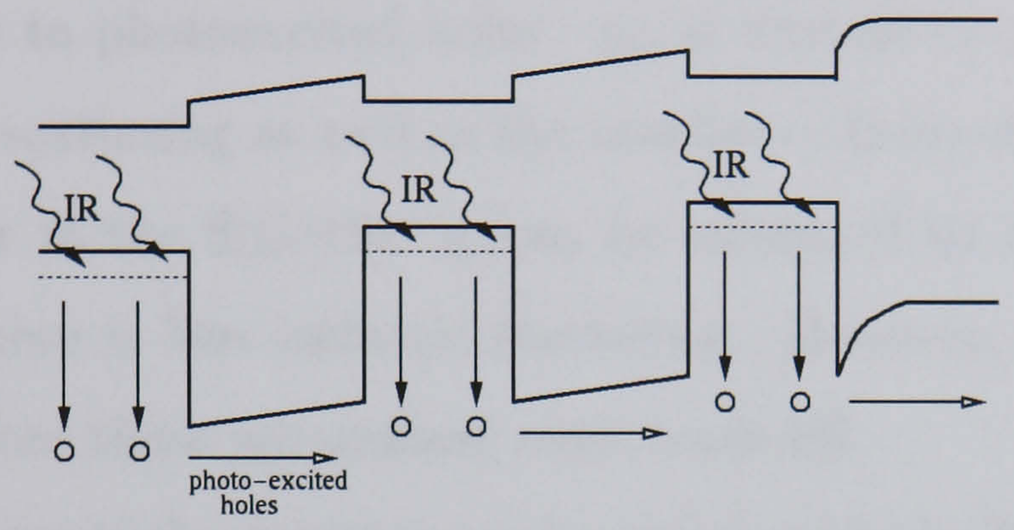


Figure 1.2: *A schematic diagram of a multi-layer HIP detector. The potential profile is given in the lower half showing the absorption mechanism*

ceed the barrier energy therefore the quantum efficiency rises only slowly with energy above the potential barrier. In HIP detectors, the band of occupied hole states is narrow in the $p^+Si_{1-x}Ge_x$ layer due to its semiconductor band structure, thus the device has a much sharper turn on and avoids the need to design a device with a significantly lower barrier than the required generating energy. Thus a correspondingly lower operating temperature is required to reduce dark current to acceptable values. Also photoexcited holes traveling over the potential barrier are less likely to be back scattered because of the more favourable ratio of the effective masses.

The device characteristics of the $Si_{1-x}Ge_x$/Si HIP can be tailored through the selection of layer thickness, doping concentration, (boron was used as the dopant), and composition to optimize the response for a given application. These detectors boast high D^* compared with QWIPs in the long wavelength region. Lin (1990) reported a detectivity of $D^* = 5 \times 10^9$ cm $Hz^{\frac{1}{2}}W^{-1}$

To achieve a good detector response, strong IR absorption and efficient internal photoemission are required. The quantum efficiency for such devices (Lin 1992) has been calculated/measured as ranging from 5% to 0.1% for wavelengths from 2-12 μm. The internal quantum efficiency, η_i, is the ratio of collected holes to photoexcited holes. η_i, is limited by inelastic hole-hole and hole-phonon scattering as well as the number of holes redirected from the SiGe/air interface to the SiGe/Si. η_i can be enhanced by reducing the alloy layer width so there is less inelastic scattering. However, this reduces that absorption therefore these parameters must trade off.

The performance of the detector can be improved by including several alloy layers stacked between Si barriers, see figure 1.2 (Lin 1994). Each layer has a high η_i, and contributes to the overall absorption. If an electric field is applied in the z-direction, the photoexcited holes are redirected to the Si substrate, thus increasing η_i, . These detectors show enhanced long wavelength photoresponse and near ideal thermionic emission limited dark current characteristics. A quantum efficiency of 6% at 5μm, 40% at 10μm and 1.5% at 15μm show significant improvement on single layer devices. Lin (1994) reports a $Si_{0.7}Ge_{0.3}$ HIP with a quantum efficiency of approximately 8% at 5μm and an operating temperature of 30K, with a cutoff wavelength of 23μm.

1.4 Quantum Well Infrared Detectors

A multiple quantum well system is fabricated by growing a sequence of semiconductor layers where a material with a smaller band gap (e.g. GaAs) is sandwiched between layers of a material with a larger band gap (e.g. GaAlAs). The carriers are thus confined to a potential well for motion parallel to the growth direction whereas they are quasi-free in the motion perpendicular to the growth

direction. The height of the potential well is determined by the conduction or valence band offsets. The confining potential leads to the formation of a series of discrete electronic subbands or minibands. There are many different types of quantum well infra-red photodetectors (QWIPs) which can be conveniently divided into those comprised of III-V compounds e.g. GaAs/AlAs, type II structures formed of InAs/GaSb compounds, II-IV compounds e.g. HgTe/CdTe which form type III systems, and group IV systems e.g. SiGe/Si. The main topic of this thesis is the latter systems, so we shall briefly review the progress of the first two sections before concentrating our attention on the SiGe systems.

Wide gap semiconductors such as GaAs, AlAs, Si and Ge are easier to grow than narrow gap systems such as HgCdTe. However, alloys and superlattices that utilize transitions across the fundamental gap are not within the range required for IR detection. An alternative is to engineer small band gaps in wide band gap semiconductors as was first suggested by Esaki (1977). The progress in modern thin film growth techniques such as molecular beam epitaxy (MBE) has enabled process control over composition, doping and thickness of the quantum well layers. Thus the electronic band structure of semiconductor systems can be tailored in an almost continuous way to custom design a system for specific applications. An additional parameter can be introduced by growing lattice mismatched hetrostructures and this altering the position of the electronic states by strain. Therefore, by tailoring the quantum well structure it should be possible to adjust the energy levels so that low-energy infrared photons can induce an intersubband transition between the ground state and the first excited state of the quantum well.

To operate these devices as detectors, electric fields must be applied and the device acts as a photo-conductor. It is possible to design systems that

operate in photo-voltaic mode (Schneider 1995), as discussed in section 1.4.8, but all detectivities quoted here are for devices operated in photo-conductive modes.

It was first suggested in 1979 (Schulman 1979), that superlattices could improve on HgCdTe for IR detection. The advantages of QWIPs are they have a higher degree of uniformity, important in detector arrays, small leakage currents because the superlattices have large effective masses therefore tunneling is suppressed, possible lower Auger recombination rates due to the LH-HH splitting, increased effective mass and utilize well understood device fabrication and materials processing techniques.

The disadvantages are the relatively low quantum efficiency, narrow band spectral characteristics and short carrier lifetimes which means lower operating temperatures are required compared to intrinsic systems.

1.4.1 n-type III-V QWIPs

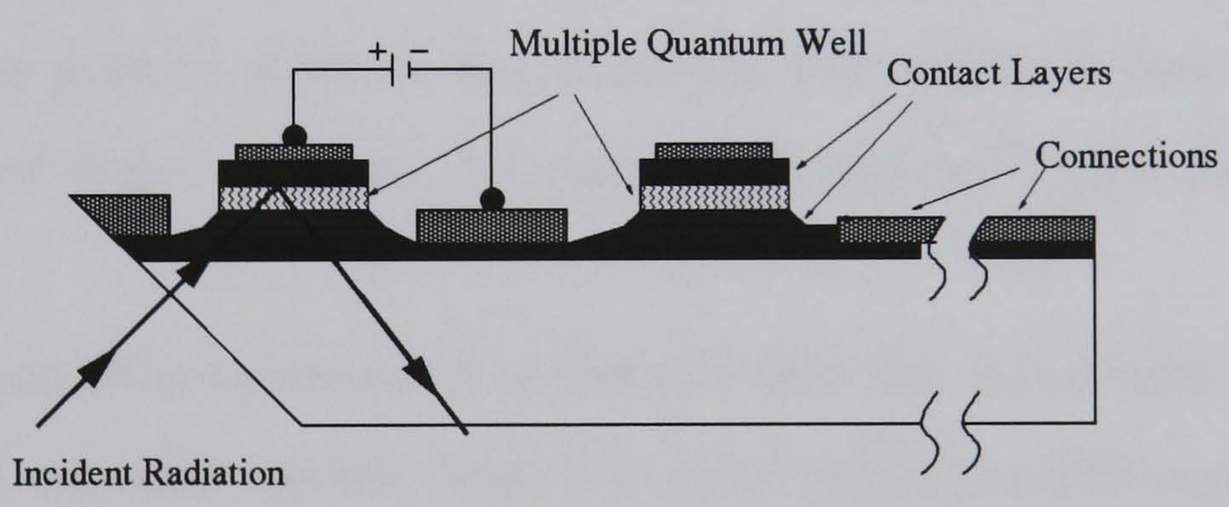


Figure 1.3: *A schematic representation showing the position of the multiple quantum well in a typical experimental setup and the angle of incident light necessary to produce absorption in n-type GaAs/AlGaAs structures.*

The first system to be extensively studied was the GaAs/AlGaAs structure, which has the advantages that it is lattice matched and thus easy to grow.

For a review of GaAs/AlGaAs systems see Levine (1993). The band gap of the AlGaAs alloy is larger than that of GaAs, therefore the electrons tend to accumulate in the GaAs layers. Thus by altering the relative concentrations of Al and Ga in the alloy barriers, i.e. $Al_xGa_{1-x}As$, the height of the barrier can be varied. These systems can be doped with either donors, to create a n-type structure where the transitions occur between conduction minibands, or acceptors to give a p-type system and transitions between valence minibands. At low temperatures these carriers will occupy the ground states of the quantum wells.

IR detection in a multiple n-type quantum well structure was first investigated experimentally by Smith (1983). A typical experimental setup for a multiple quantum well is shown in figure 1.3.

Strong intersubband absorption in the conduction band was observed by West and Eglash (1985). The first working QWIP was demonstrated by Levine in 1987. This device utilized transitions between two bound quantum well states. The mechanism for this device was an absorption of a photon by an electron in the ground state of the quantum well which is then excited into the first excited state, which in this case is still confined to the well (see figure 1.4).

This electron then tunnels out of the well into the continuum where, in the presence of an external electric field, it is collected to produce a photocurrent. By reducing the width of the quantum well it is possible to move the first excited state into the continuum. The major advantage of this is that the electron no longer has to tunnel through the barrier to reach the continuum, thus the voltage needed to collect the photoelectrons is dramatically reduced, which in turn reduces the dark current. As the carriers are no longer required to tunnel through the barrier the width of the AlGaAs barriers can be increased

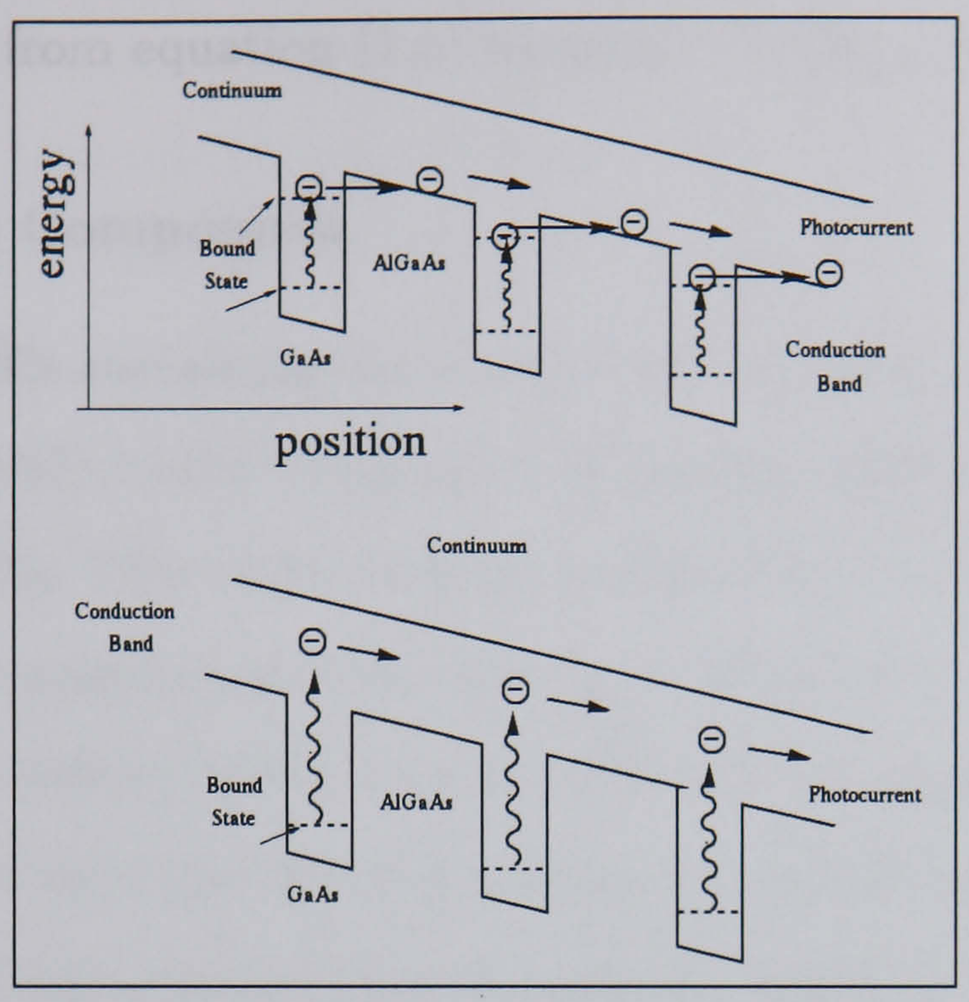


Figure 1.4: *Schematic diagram showing the processes involved in intersubband absorption in a quantum well system. The upper diagram demonstrates bound to bound transitions and the lower half bound to continuum transitions*

thus further reducing dark current by reducing the sequential tunneling current which is the dominant source of dark current at temperatures below 30K. However, D* does not improve indefinitely with increased barrier width as higher bias voltages are needed which leads to a higher dark current.

The first bound-continuum device, which operated at 10 μm, was demonstrated in 1988 by Levine (1988), and showed an order of magnitude improvement in performance over the bound to bound system. For temperatures up to 80K the reported detectivity of these QWIPs is comparable with HgCdTe. The current state of the art is the demonstration of a 128x128 pixel array comprising of n-type GaAs/$Al_{0.27}Ga_{0.73}As$ utilizing bound-continuum transitions within the conduction band, operating at 35K with a cutoff wavelength of $\lambda_c = 9.5\mu m$, with a detectivity of $D^* = 4.5 \times 10^{13}$ cm $Hz^{\frac{1}{2}}$ W^{-1} (which is 75% of BLIP D*) (Kozlowski 1993). This array also demonstrated excellent

uniformity with u from equation (1.6) equal to 4% before correction.

1.4.1.1 Indium Compounds

Other n-type QWIPs containing GaAs have been studied, in particular GaAs/$Ga_{0.5}In_{0.5}P$ (Rao 1987), GaAs/$As_{0.5}In_{0.5}P$ (Watanabe 1987) and $In_{0.15}Ga_{0.85}As$/GaAs (Levine 1993). This latter system is interesting as it uses GaAs as the barriers. Binary compounds such as GaAs have better electron transport properties compared to ternary alloys such as AlGaAs. Therefore if GaAs could be used as the barrier then the device performance would improve. This leaves the problem of finding a compound with a smaller bandgap than GaAs to form the wells. The first attempt at this was by Gunapala et al. (1994) who used InGaAs quantum wells. This system has a 1.2% lattice mismatch, but despite this the peak responsivity and detectivity, at 16.7μm at 40K, were comparable with the traditional GaAs/AlGaAs system. The results also indicated excellent electron transport properties. Other alternatives using In were also explored by Gunapala (1991), including an InGaAs/InP system which showed excellent detectivity of $D^* = 9 \times 10^{10}$ cm $Hz^{\frac{1}{2}}W^{-1}$, a quantum efficiency of 12% and a responsivity of $R = 1.2AW^{-1}$ at an operating temperature of 77K. This detectivity compares favourably with the best GaAs/AlGaAs QWIPs operating at this wavelength, whilst the responsivity and optical gain are five times higher. These excellent hot-electron transport properties are attributed to the high mobility of binary InP barriers.

1.4.2 Normal Incidence

For large 2-D arrays normal incident illumination is important. For absorption involving transitions between conduction subbands at the Γ point the absorption is polarization sensitive and only the component of radiation with electric

field perpendicular to the plane of the quantum well, i.e. the z-direction, is absorbed. This is because in the conduction band the electromagnetic field can only couple conduction bands of different symmetry under reflection in the plane of the well, which requires a component of the radiation electro magnetic field perpendicular to the well. 2-D arrays can be created from n-type QWIPs such as GaAs/AlGaAs to absorb normal incidence but only if gratings and optical coupling waveguides are used (Levine 1993, Hasnain 1989 and Goosen 1988) to provide a component of the normally incident light in the z direction, as shown in figure 1.5.

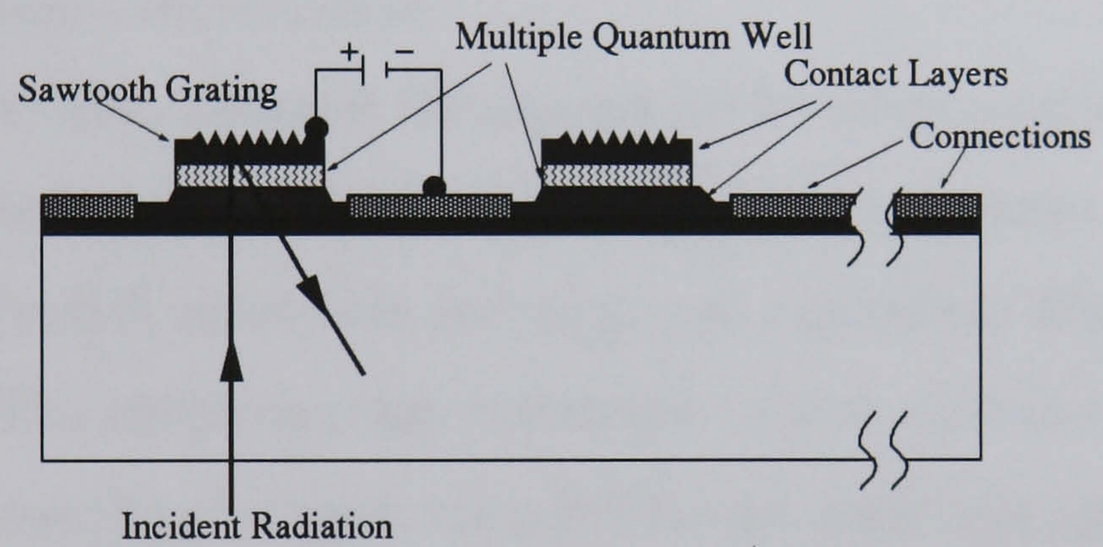


Figure 1.5: *Schematic representation of experimental set up including a saw-tooth grating to obtain normal incidence response.*

A grating coupled waveguide involves etching a grating into the top contact layer of each pixel to scatter incident photons into propagation paths at an angle to multiple quantum well interfaces. The waveguide allows coupled radiation to enter at an angle and pass through the multiple quantum well structure several times. Quantum efficiencies of up to 95% for normal incidence have been obtained using these techniques (Andersson 1991).

A different approach to achieve normal incidence absorption is to use intersubband transitions involving electrons in the ellipsoidal conduction band

minima of GaSb/AlSb (experimental: Samoska 1993, Brown 1992 and Xie 1993, theoretical : Shaw 1994). For certain parameter combinations in these structures the L-valley conduction bands can have lower or comparable energy to the direct gap minima and can have considerable carrier population. These ellipsoidal L-valley states have anisotropic effective masses and thus favourable selection rules for normal incidence absorption. Also the L valleys have high density of states due to the high effective mass. Another potential advantage of using the conduction band is the simpler structure compared to the valence band, thus devices should be more controllable and less susceptible to factors such as the doping concentration.

Zhang (1994) demonstrates NI absorption between conduction minibands of X-valley states in AlAs/AlGaAs quantum well structures. These systems were grown on [113] Si substrates and displayed a quantum efficiency of approximately 13%. The effective mass anisotropy of the electrons in these valleys again provides coupling between the parallel and perpendicular motions of the electrons when the principal axes of one of the ellipsoids are tilted with respect to the growth direction. As a result of this coupling, inter-conduction subband transitions at normal incidence in indirect gap semiconductor quantum wells are allowed when the MQW's growth direction is not along the [001] direction.

Wang (1993) grew AlAs/$Al_{0.5}Ga_{0.5}As$ system on a [110] GaAs substrate, creating a type II superlattice and observed normal incidence absorption due to indirect transitions between the X-band and Γ-band.

1.4.3 p-type III-V QWIPs

Most previous work has been on n-type systems as the lower effective mass in the conduction band was expected to lead to superior transport properties. It is also possible to use transitions within the valence band. Mixing between

the light hole and heavy hole minibands away from the centre of the Brillouin zone relaxes the selection rules and normal incidence light can be absorbed. Band mixing and the selection rules for p-type QWIPs are discussed in more detail in later chapters. The disadvantage of using p-type structures is the mean free path of the excited holes is very small due to the large hole effective mass. This leads to smaller gain and lower detectivity compared with n-type systems.

The first QWIP working with normal incident light inducing transitions between valence subbands was developed by Levine (1991) using $GaAs/Al_{1-x}Ga_xAs$. D*'s of up to 3.1×10^{10} cm $Hz^{\frac{1}{2}}W^{-1}$, at 77K for a cutoff wavelength of 8μm were observed, and a quantum efficiency of up to 28% achieved at low temperature operation.

1.4.4 Type II superlattices

The QWIPs previously discussed have all required doping as they utilize intersubband transitions. An alternative is to use compounds which create type II superlattices which have fundamental gaps smaller than either of the constituents, thus wide gap materials can be used to create small band gap quantum wells.

One problem with these superlattices is that the electrons and holes are localized in different layers, due to the staggered nature of the band edge lineups, therefore the optical matrix elements are rapidly reduced as the layer thicknesses increases. In a simple InAs/GaSb system the layer thickness required to absorb at 10μm is too thick for the optical matrix elements to be large enough to be useful (Campbell 1991). If an alloy of InSb and GaSb is used instead then this problem is overcome and the required wavelength range can be achieved with smaller layer thicknesses. This system has the added advantage

that the superlattices can be grown lattice-matched to GaSb substrates (Chow 1991). Both $InAs/Ga_{1-x}In_xSb$ and $InSb/InAs_xSb_{1-x}$ superlattices have been demonstrated to have gaps in the 8-14 μm range (Miles 1990 and Kurtz 1989).

These superlattices have intrinsic absorption mechanisms, i.e. electrons excited from valence to conduction band. This allows the operating temperature of the detector to be much higher than for intersubband QWIPs. The internal strain of up to 7% due to the introduction of InSb into the barrier lowers the conduction band minima of InAs and raises the top heavy hole state of GaInSb, thus reducing the superlattice band gap at a given layer thickness.

The effective band gap of this system can be tuned for absorption throughout the 3-20μm range. Dissimilar superlattice structures can have nominally similar effective bandgaps. However, other band structure related features such as effective mass, miniband width and optical matrix elements can vary substantially, resulting in very different fundamental absorption properties. It is then necessary not only to engineer the bandgap but also to fully optimize the band structure of this material for a particular infrared detector application (Johnson 1990). The optical properties of this superlattice are as good as HgCdTe but the electronic properties are better. HgCdTe has larger optical matrix elements but lower density of states. Also the effective mass is much larger in the superlattice than the alloy, so diode tunneling currents are greatly reduced, (Smith 1987, Mailhot 1987a,b).

There are problems with the growth of these structures, however. The switching of the anion from Sb to As at the growth interface provides a different set of problems than the growth of common anion structures like GaAs/AlAs (Collins 1993). When this technology is brought to maturity we could expect higher temperature BLIP than from HgCdTe. Above 12μm HgCdTe large area detectors suffer from tunneling noise currents and the sensitivity of these

currents to small variations across a wafer. Ga-based systems offers a marked reduction in these currents and less sensitivity to wafer variation. Also the strain in these systems results in a large LH-HH splitting which could be engineered to suppress the Auger recombination which is the dominant limiting factor in HgCdTe performance.

1.4.5 Type III Superlattices

A variation on the type II systems is given by HgTe/CdTe which are classed as type III as one of the constituents, HgTe, is a semi-metal. Schulman (1979) first suggested using HgTe/CdTe superlattices with the hope that the detection wavelength would be easier to control in the longer wavelength range than HgCdTe, and tunneling currents will be reduced. These superlattices have large IR absorption coefficients therefore high quantum efficiency can be obtained in relatively thin structures. The first successful superlattice of this type was grown by Faurie (1982). Determination of the valence band offset, doping and growth technology all proved problematic in the early development (Boukerche 1986 and 1988). However, the system benefits from the same basic properties that make HgCdTe such a good candidate in the first place, namely high absorption coefficients, high density of states etc. Theoretical calculations (Jones 1985) predict absorption coefficients of up to $10^4 cm^{-1}$, an order of magnitude higher than those of bulk HgCdTe. Jones (1985) demonstrated IR absorption experimentally showing absorption coefficients approaching the theoretical maximum at room temperature. This large absorption is not a peculiarity of the superlattice but a feature of the intrinsically high density of states of HgTe. There have also been reports that the Auger lifetimes may be an order of magnitude greater than for the bulk (Jiang 1991). It has the additional advantage of more control over the cutoff wavelength and the

larger effective masses which reduces undesirable tunneling currents. As the absorption is across the fundamental gap it is not polarization dependent. The problems with the growth technology have been overcome and device quality samples have been demonstrated (Lansari 1993); these systems could become serious candidates for IR detection. However, a large degree of uniformity, which is what holds back HgCdTe arrays needs to be demonstrated.

1.4.6 δ-Doping

δ-doping is a way of creating quantum wells without interfaces. The electrostatic attraction between the free carriers and the ionized dopants results in an approximately V-shaped potential well perpendicular to the dopant plane which confines the carriers into a quasi-2-D sheet and quantizes their energies into a series of subbands. A review of delta-doping is given by Harris (1993). The potential profile strongly resembles a parabolic quantum well due to the finite width of the doped layer, (Zenner 1988). The barrier height and well thickness can be varied by altering the doping density and the width of the doped layer. This method produces very wide absorption peaks with full width half maximum (FWHM) values of up to 175meV, compared with a typical linewidth for absorption in a GaAs/AlGaAs system of approximately 15meV. This is due to the non-parabolicity of hole bands and the high doping levels which means that the bands can be filled up to several hundred meV's below the top of the band.

Hole intersubband absorption in δ-doped multiple Si layers has the advantage that there is no limitation of critical thickness so as many δ-doped layers as required can be grown to improve the quantum efficiency. The peak wavelength of theses devices can be engineered by adjusting the concentrations of dopants in the δ-doped layers. δ-doping has also been demonstrated using

n-type doping in Si (Scolfaro 1995).

1.4.7 SiGe/Si Quantum Wells

Quantum wells constructed of SiGe/Si have the advantage of compatibility with existing silicon technology. Also most of the potential discontinuity occurs in the valence band so p-type structures can be engineered to absorb in the IR range, offering the additional advantage of normal incidence detection without the use of gratings etc. The 4.17% lattice mismatch between bulk Si and Ge results in strained structures which makes the growth process more difficult. Early attempts at SiGe/Si growth using high temperature MBE techniques (Kasper 1975 and 1977) were troubled by islanding and other growth problems. These were eventually solved by reducing the temperature of the growth process so that pseudomorphic growth occured, i.e. the in-plane growth lattice is strained to match that of the substrate and the vertical lattice constant is elastically strained. The first successful SiGe/Si structure was grown by Bean (1984) who used a germanium concentration of 20%. Psuedomorphic growth on Si substrates of alloys up to 80% has since been reported (Wang 1993).

The first IR detector based on a p-type SiGe/Si system was demonstrated by Karunasiri (1990) who observed a broad parallel incidence peak at around 8μm. There have been many other recent reports of absorption in p-type structures (see, for example, People 1992, Zanier 1995, Boucard 1995). In particular Fromherz (1995) produced a systematic study of absorption for a range of well widths and germanium concentrations. These results are studied in more detail in Chapter 6 where we compare the experimental results with our calculated curves. A typical D* for a SiGe/Si system is 3.3×10^9 cm $\mathrm{Hz}^{\frac{1}{2}}\mathrm{W}^{-1}$, (Fromherz 1995), for parallel incidence absorption. This value how-

ever, was obtained with an unoptimized structure and further study could well improve on this. Normal incidence absorption in p-type structures was first observed by Park (1991).

If quantum wells are grown on a relaxed SiGe buffer a larger proportion of the potential discontinuity occurs in the conduction band and intersubband absorption between conduction minibands has been observed in such structures (Hertie 1991). Infrared absorption has also been demonstrated in δ-doped n-type Si (Tempel 1990) δ-doped p-type Si (Park 1991).

Another example of using substrate orientation to change the selection rules of the system is given by Lee (1992) who reports strong electron intersubband IR absorption using Sn δ-doped Si/SiGe multiple quantum well structures grown in [110] substrates. Again normal incidence absorption is allowed because of the titled ellipsoidal of constant energy surfaces.

1.4.8 Photovoltaic QWIPs

Some devices can be operated either in photoconductive or photovoltiac modes. Photovoltaic devices operate under zero bias voltage (open circuit); photoconductive devices a reverse bias voltage is applied (short circuit). A photovoltiac detector is essentially a photo-diode operating with no external bias; the voltage is created optically. The charge separation and therefore the photovoltage is generated by photoelectrons being swept into the n-region and photo-holes being swept into the p-region by the junction electric field. The charge separation of the electrons and holes produces a voltage that forward biases the junction. The forward bias increases the diffusion current. HgCdTe and junction diodes can be operated as photovoltiac devices.

The advantage of photovoltiac operation is that the total current is much lower therefore the dark current is very small. However, much lower respon-

sivities and hence detectivities are obtained in this mode of operation. The noise in photo-voltaic devices is due to Johnson noise only, i.e. the noise floor is reduced. Theoretical analysis (Schneider 1995) shows that the detectivities of photo-conductive and photo-voltaic devices are of the same order of magnitude.

The first photovoltiac QWIP was demonstrated by Goosen (1988) using an asymmetric modulation-doped single quantum well with a graded barrier on one side of the well. The photoexcited carriers that escape the well are trapped in the graded barrier region producing a charge separation. Rosencher (1990) reported another photovoltiac device structure using coupled asymmetric quantum wells, where the electrons are photoexcited from one well and stored in an adjacent well for a short time before tunneling back to their original ground state. These devices produced virtually no dark current but the photoresponse produced was below the noise of practical amplifier circuits. More recently Schneider (1995) reported strong photovoltiac response using barrier layers with an asymmetric potential distribution. A graded potential barrier adjacent to the quantum well is used to induce asymmetric transport behaviour. The excited carriers tunnel out of the quantum well across narrow tunnel barriers into the barrier region. An internal electric field, induced by a shifted doping distribution with respect to the quantum well centers leads to a unidirectional motion of the photoexcited carriers at zero applied voltage. An unoptimized detectivity of 4.4×10^{10}cm $\mathrm{Hz}^{\frac{1}{2}}\mathrm{W}^{-1}$ was observed.

1.5 The Advantages of Si-based Detectors

The ability to couple multi-element arrays of infrared detectors to their respective amplifiers in an integrated circuit is an important consideration in

device design. Detectors and current components can all be made on the same Si chip using standard integrated circuit technology. The ability to fabricate resistors, capacitors, transistors, diodes and CCD's along with the detector would allow a substantial amount of on-chip processing. Such techniques are mandatory if very large multi-element arrays, i.e. thousands of elements, are to be achieved. If the detector is not silicon based the problem of bringing this many leads off the focal plane would be formidable. Future trends in IR detectors are towards larger electronically addressed arrays and higher operating temperatures, with the aim of producing higher performance and smaller, lighter and more affordable systems. The ultimate aim must be a technology in which the IR material is grown and deposited onto the VLSI Si substrate already containing the addressing circuits. This will be essential to achieving the array size (more that 1000x1000 elements) needed for unscanned high performance imagery at low cost. Rapid progress has been made on the silicide Schottky barrier, and these arrays point the way for other detector materials, i.e. must be readily compatible with Si technology to survive in the long term.

1.6 Overview of Thesis

We have discussed in this chapter how QWIPs have great promise as IR detectors. Two major advantages for any potential material are compatibility with existing Si technology thus allowing monolithic integration of array electronics, and the ability to absorb normally incident light vital for array design, without the need for additional complication such as gratings and waveguides.

With this in mind we present a detailed study of p-type SiGe/Si quantum well structures utilizing absorption between valence minibands. In chapters 2 and 3 we review the local empirical pseudopotential method we use to calculate

the band structures and optical matrix elements for bulk and quantum well materials respectively. We include the effects of spin and strain and discuss their effect on the bandstructure.

In Chapter 4 we present a detailed derivation of the first order susceptibility expression we use to calculated the absorption, by means of density matrix theory. This enables us to predict the absorption response for a given system.

The sampling techniques used and the relative merits of barrier and well doping are discussed in Chapter 5. The resulting Fermi levels and the exchange interactions for both barrier and well doping are calculated. We compare our calculations with recent experimental results in Chapter 6. We also discuss the role of the 'damping constant', which appears the denominator of the susceptibility expression , which affects the line width of the calculated absorption spectra.

Further calculated results for parallel incident absorption are presented in Chapter 7 where the microscopic origin of the peaks, the effect of changing the well width, Ge concentration in the well, temperature and doping concentration are discussed.

Normal incidence absorption and a discussion of band mixing are the subject of Chapter 8. Finally, in Chapter 9, we discuss Auger recombination which can be a limiting factor in device performance. Efforts are made to optimize structures in order to minimize this non-radiative recombination and preliminary results for Auger recombination rates calculated using a Monte-Carlo technique are presented.

Chapter 2

The Bulk Calculation

The pseudopotential method first devised by Herring (1940) and refined by Philips (1959) and Bassani (1961). This method enables us to calculate eigenfunctions and eigenvalues for the system and hence the band structures and optical matrix elements for both the bulk and quantum well cases.

We first discuss the symmetry of the bulk crystal, then review the approximations which must be made in order to reduce the problem to a tractable one. We introduce the concept of orthogonal plane waves, and from this formulate an expression for the pseudopotential matrix elements of the secular equation for bulk materials. We describe the model used for the local pseudopotential and calculate the band structures of bulk silicon and germanium.

2.1 Symmetry of the Bulk Crystal

In order to determine the energy eigenvalues and eigenfunctions of our systems we need to solve the Schrödinger equation for the system. If we know the symmetry properties of the Hamiltonian operator the problem can be greatly simplified and expected degeneracies and selection rules can be derived via

group theory methods. The symmetry operations can also be applied to the crystal lattice to obtain the irreducible segment which contains all relevant information and from which the properties of the rest of the zone can be deduced using the symmetry operations.

An infinite three dimensional lattice can be defined in terms of 3 non-coplanar basic lattice vectors, $\mathbf{a_1}, \mathbf{a_2}$ and $\mathbf{a_3}$ as shown in figure 2.1.

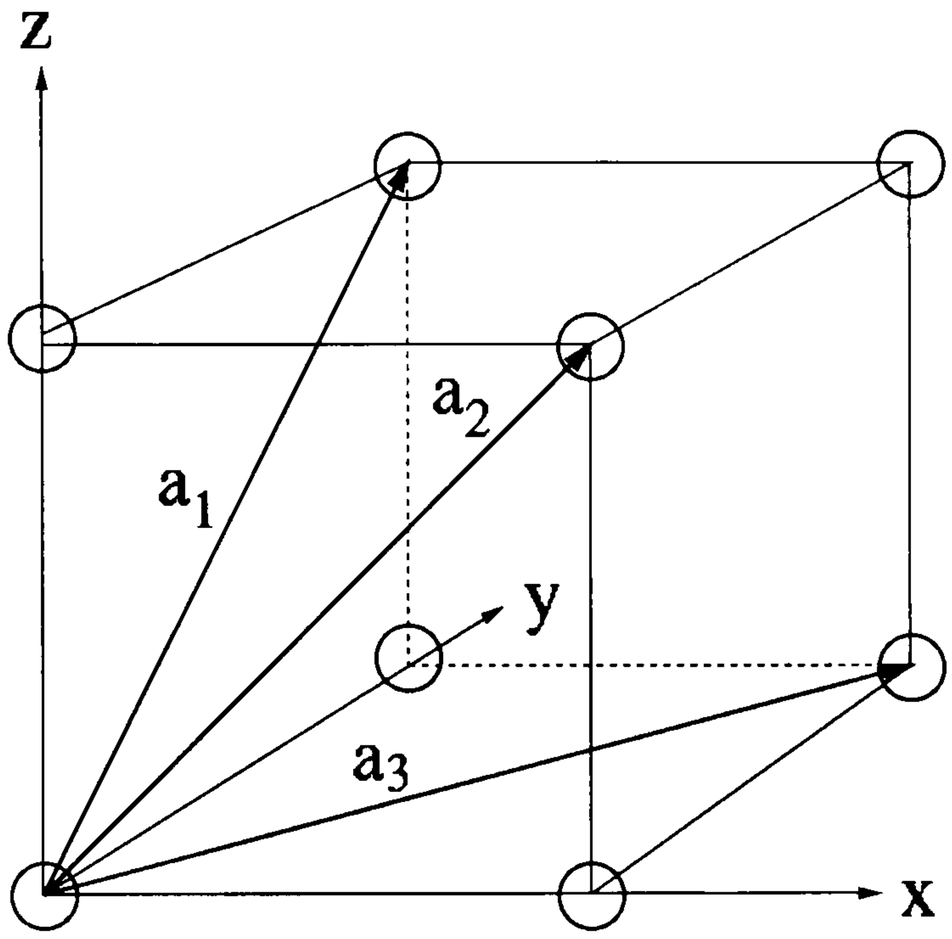


Figure 2.1: *Primitive lattice vectors for a face-centred cubic structure. The x, y and z axes indicate the Cartesian axes used in this Chapter.*

The set of lattice vectors for the crystal is given by

$$\tau_{\mathbf{n}} = n_1\mathbf{a_1} + n_2\mathbf{a_2} + n_3\mathbf{a_3} \tag{2.1}$$

where $\mathbf{n} = (n_1, n_2, n_3)$ where n's are integers. A pure translation through a lattice vector is called a primitive translation and is denoted $\{\mathbf{E}|\tau_{\mathbf{n}}\}$ (Seitz 1940), where $\mathbf{E}$ is the identity operator, and indicates that no rotation operations are included. The rotational parts of the space group symmetry, R, make up the point group of the lattice. i.e. the group of symmetry operations

which leave the highest symmetry point of the lattice unmoved, as distinct from translations in which no point of the lattice remains in the same place. A symorphic space group is defined as a group of symmetry operations $\{\mathbf{R}|\tau_\mathbf{n}\}$ which contains as a subgroup the set of all primitive translations of the lattice but contains no other pure translations, i.e. the set of all operations which transforms the 3-D array of crystal lattice points onto itself. Bravais lattices may be classified in terms of the maximal point group compatible with the lattice.

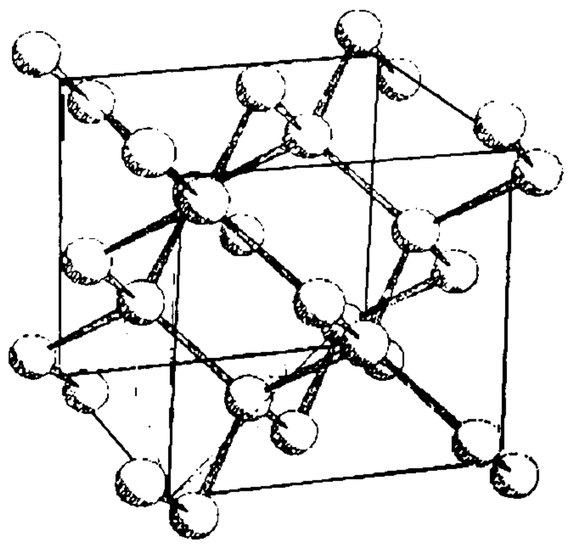


Figure 2.2: *A perspective view of the diamond structure.*

The crystals of both Si and Ge are of diamond structure as shown in figure 2.2. The diamond lattice can be considered to consist of two face-centered cubic lattices displaced with respect to each other by the vector $\mathbf{f} = \frac{a}{4}(1,1,1) = \frac{1}{4}(\tau_\mathbf{1} + \tau_\mathbf{2} + \tau_\mathbf{3})$. Thus the space group of the diamond lattice is non-symmorphic as it contains pure, non-primitive translations. However, at the centre of the Brillouin zone, Γ, $e^{i\mathbf{k}.\mathbf{f}} = 1$ and the characters of the irreducible representations are simply the point group O_h presented in table 2.1.

This group consists of 48 rotations which leave the crystal unchanged. A rotation is denoted C_n where the rotation is through an angle $\frac{2\pi}{n}$ radians.

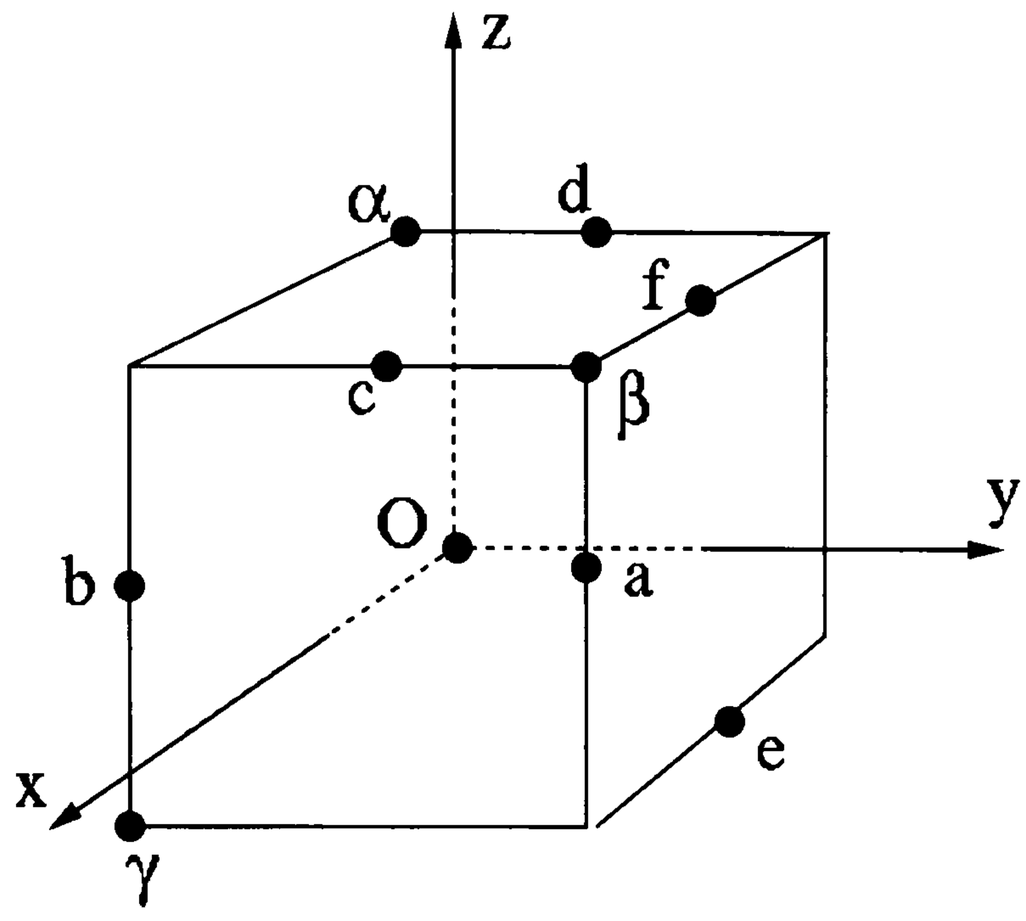


Figure 2.3: *Symmetry axes for the diamond structure.*

From figure 2.3 it can be seen that the crystal is invariant under rotations of π and $\frac{\pi}{2}$, i.e. C_2 and C_4, about the x, y and z axes. Rotations of π are also possible about the axes O_a, O_b, O_c, O_d, O_e and O_f. These operations are denoted C_2'. Rotations of $\frac{2\pi}{3}$ are possible around the axes $O_\alpha, O_\beta, O_\gamma$ and O_δ. These operations can be either in a right or left hand screw sense which are not equivalent as they are for C_2 i.e. $C_2^- = C_2^+$. Similarly the C_4 rotations can be performed in either sense. Note also that C_4^2, that is C_4 applied twice is equivalent to C_2. Thus we have 24 rotations which leave the crystal unchanged. The inclusion of the inversion operator, I, which transforms (x, y, z) to $(-x, -y, -z)$ increases this to the 48 operations that comprise the O_h point group. From these operations we can determine the irreducible segment of the Brillouin zone. The first Brillouin zone with the irreducible segment indicated is shown in figure 2.4.

	E	$8C_3''$	$3C_2$	$6C_4$	$6C_2'$	I	$8IC_3''$	$3IC_2$	$6IC_4$	$6IC_2'$
Γ_1	1	1	1	1	1	1	1	1	1	1
Γ_2	1	1	1	-1	-1	1	1	1	-1	-1
Γ_{12}	2	-1	2	0	0	2	-1	2	0	0
Γ_{15}'	3	0	-1	1	-1	3	0	-1	1	-1
Γ_{25}'	3	0	-1	-1	1	3	0	-1	-1	1
Γ_1'	1	1	1	1	1	-1	-1	-1	-1	-1
Γ_2'	1	1	1	-1	-1	-1	-1	-1	1	1
Γ_{12}	2	-1	2	0	0	-2	1	-2	0	0
Γ_{15}	3	0	-1	1	-1	-3	0	1	-1	1
Γ_{25}	3	0	-1	-1	1	-3	0	1	1	-1

Table 2.1: *Character table for the irreducible representations of the O_h group (from Cornwell 1969). See text for explanation of notation*

2.2 The Crystal Hamiltonian

Before the quantum well calculation can be attempted we need to solve the problem for the bulk case. The bulk band structure of a suitable host material is input into the quantum well calculation. To describe the band structure and optical properties of solids we require the electronic wavefunctions and energies. In order to obtain these we must first consider the crystal as a whole. The Schrödinger Equation for the crystal is

$$H\psi(\mathbf{R},\mathbf{r}) = E\psi(\mathbf{R},\mathbf{r}) \tag{2.2}$$

where $\mathbf{r}$ represents the coordinates of the electrons and $\mathbf{R}$ the nuclei. H is the total Hamiltonian of the system and E represents the eigenvalues. The total Hamiltonian is comprised of

$$H = T_e + T_L + V_{ee} + V_{eL} + V_{LL} \tag{2.3}$$

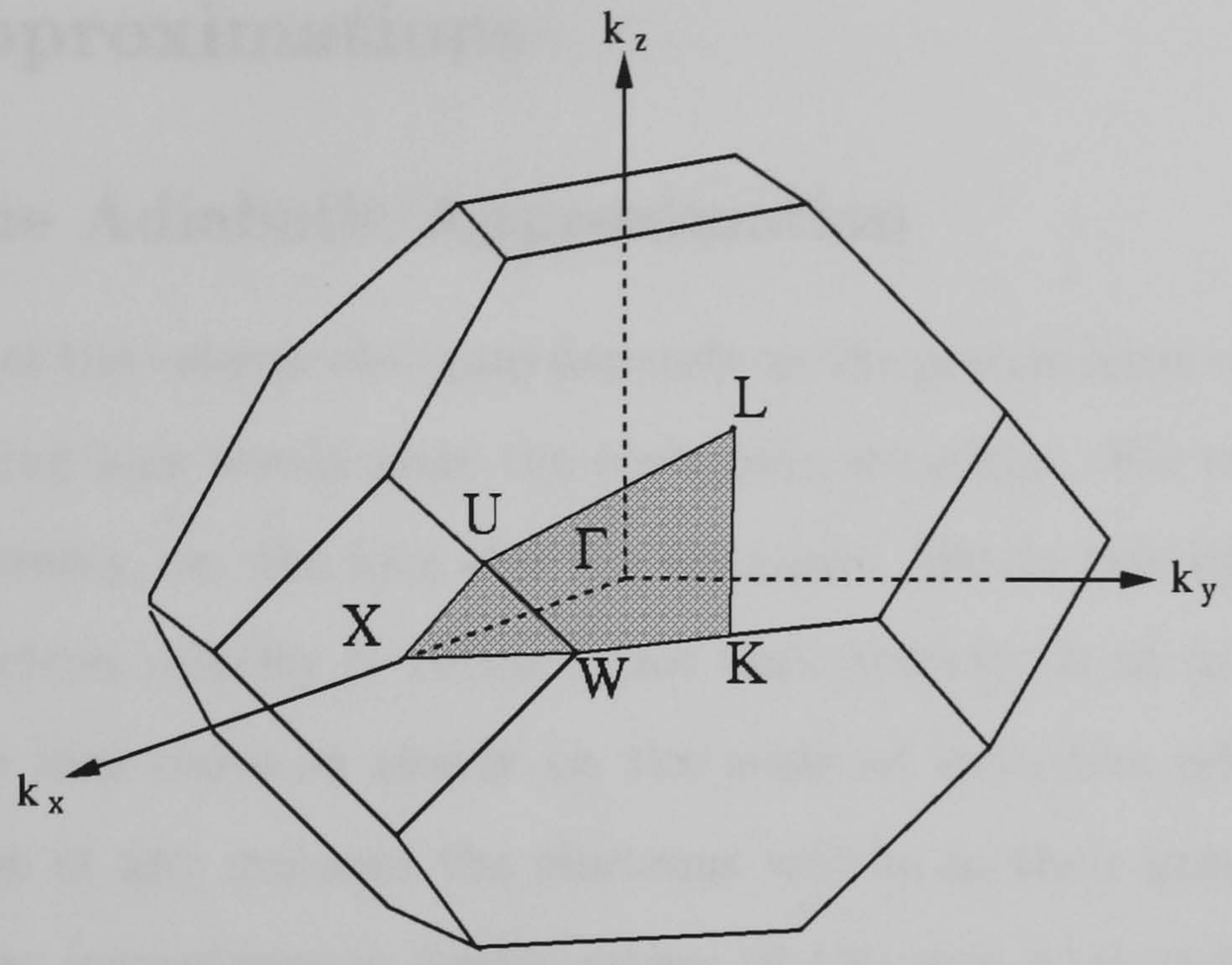


Figure 2.4: *The first Brillouin zone associated with the diamond structure. Points of high symmetry are indicated. Shaded area shows the irreducible segment.*

where T_e and T_L are the kinetic energies of the electrons and nuclei respectively, V_{ee}, V_{eL} and V_{LL} are the electron-electron, electron-nucleon and nucleon-nucleon interactions. These interactions are Coulombic thus the total Hamiltonian can be written as

$$H=-\sum_i \frac{\hbar^2}{2m_i}\nabla_i^2-\sum_I \frac{\hbar^2}{2M_I}\nabla_I^2+\sum_{i<j}\frac{e^2}{|\mathbf{r_i}-\mathbf{r_j}|}-\sum_{I,i}\frac{Z_I e^2}{|\mathbf{R_I}-\mathbf{r_i}|}+\sum_{I<J}\frac{Z_I Z_J e^2}{|\mathbf{R_I}-\mathbf{R_J}|} \tag{2.4}$$

In order to solve this equation we would need knowledge of the interactions between all the constituents of the crystal. This is a many-body problem which is impossible to solve even if the calculation is limited to a very small volume of the crystal. Therefore in order to proceed we must make a series of simplifying approximations.

2.3 Approximations

2.3.1 The Adiabatic Approximation

The energies of the valence electrons depends on the precise ionic configuration. Thus displacing ions would alter the electronic structure. We must separate these two systems, i.e. the ions and the electrons. We do this by noting that a typical electron velocity is $10^6 \mathrm{ms}^{-1}$, an ionic velocity is at most $10^3 \mathrm{ms}^{-1}$, therefore the ions move so slowly on the scale of velocities relevant to the electrons that at any moment the electrons will be in their ground state for any particular instantaneous configuration of the ions (Ashcroft 1981). We can therefore write the eigenfunctions of H in the form

$$\psi_{n,v}(\mathbf{R},\mathbf{r}) = \zeta_{n,v}(\mathbf{R})\phi_n(\mathbf{R},\mathbf{r}) \tag{2.5}$$

where $\zeta_{n,v}$ is the ionic wavefunction and $\phi_n(\mathbf{R},\mathbf{r})$ is the electronic wavefunction, v and n are quantum numbers of the nuclei and electrons respectively. For a given nuclear configuration $\mathbf{R}$ we can write a Hamiltonian in terms of electron variables only. The Schrödinger Equation then becomes

$$H\zeta_{n,v}(\mathbf{R})\phi_n(\mathbf{R},\mathbf{r}) = E_{n,v}\zeta_{n,v}(\mathbf{R})\phi_n(\mathbf{R},\mathbf{r}) \tag{2.6}$$

The electronic Schrödinger Equation is

$$H_e\phi_n(\mathbf{R},\mathbf{r}) = E_n\phi_n(\mathbf{R},\mathbf{r}) \tag{2.7}$$

where H_e is the electronic Hamiltonian which we can solve to give the electronic states of the crystal. The electronic wavefunction is still a function of $\mathbf{R}$, the ionic configuration, but we can now consider the electronic wavefunctions for a given ionic configuration.

$$H_e = T_e + + V_{ee} + V_{eL} \;=\; -\sum_i \frac{\hbar^2}{2m_i}\nabla_i^2 + \sum_{i<j}\frac{e^2}{|\mathbf{r_i}-\mathbf{r_j}|} - \sum_{I,i}\frac{Z_I e^2}{|\mathbf{R_I}-\mathbf{r_i}|} \tag{2.8}$$

2.3.2 The One-electron Approximation

To further simplify the problem we introduce the one-electron approximation. The many electron wavefunctions are written as anti-symmetrised products of one-electron wavefunctions, known as 'determinental wavefunctions'. We consider the 'one-electron' to be moving in a field comprised of the field produced by the ionic cores and the field produced by all other electrons. These wavefunctions automatically satisfy the Pauli exclusion principle. The best choice for one-electron wavefunctions are those which satisfy the Hartree-Fock equations (Ashcroft 1981). We can express the electronic states of the crystal in terms of one-electron wavefunctions, ψ_n, which are related to the Hartree-Fock equation by

$$H_{HF}\psi_n(\mathbf{r}) = E_n\psi_n(\mathbf{r}) \tag{2.9}$$

The Hamiltonian, H_{HF}, can be expressed as

$$H_{HF} = -\frac{\hbar^2}{2m}\nabla^2 - \sum_I \frac{Z_I e^2}{|\mathbf{R_I} - \mathbf{r}|} + V_{coul} + V_{exch} \tag{2.10}$$

where V_{coul} is the coulomb potential and V_{exch} is the exchange potential. Together they replace the electron-electron interaction term in equation (2.8).

2.3.3 The Band Approximation

We have simplified the problem considerably but the solution of equation (2.10) is still not easy, in particular the determination of the exchange potential presents many problems. To circumvent this problem we replace the electron-nuclei and the electron-electron interactions with an effective one-electron potential $V(\mathbf{r})$. The form of this potential is chosen such that it is invariant under the symmetry operations of the crystal; in particular

$$V(\mathbf{r} + \mathbf{R}) = V(\mathbf{r}) \tag{2.11}$$

where $\mathbf{R}$ is the Bravais lattice vector of the lattice. The Schrödinger Equation for a single electron becomes

$$H\psi(\mathbf{r}) = [\frac{-\hbar^2}{2m}\nabla^2 + V(\mathbf{r})]\phi_n(\mathbf{r}) = E_n\phi_n(\mathbf{r}) \tag{2.12}$$

The problem remains of what form $V(\mathbf{r})$ should take. This one electron potential, $V(\mathbf{r})$, consists of the periodic potentials due to the ions and periodic effects due to interaction of the 'one-electron' with all other electrons. Therefore to determine $V(\mathbf{r})$ exactly would require solving the one-electron Schrödinger equation for every electron. Obviously this is not possible as you need to know the potential before you can solve the Schrödinger Equation. One way round this is to take a shrewd guess at the potential, $V_0(\mathbf{r})$, calculate the electron wavefunctions and from these recompute the potential to give $V_1(\mathbf{r})$. This procedure is repeated until $V_0(\mathbf{r}) = V_1(\mathbf{r})$, i.e. a self consistent potential is found that reproduces itself. Alternatively this potential can be fitted empirically to experimental results. The success of the band approximation is due to the fact that the eigenvalues and electronic states depend to a large extent of the symmetry properties of the crystal which are present in the potential.

There are several different approaches that can be used to determine this potential, i.e the tight binding, cellular and augmented plane wave methods have all been used with varying degrees of success. (See Ashcroft 1981, or Bassani 1975 for a discussion of these techniques). We use the pseudopotential method which we shall now develop.

2.4 The Plane Wave Expansion

Let us consider the electrons in the crystal in more detail. They can be divided into two types; the core electrons which are tightly bound to their respective nuclei, and the valence electrons which are less tightly bound and are likely to

be found away from the ionic centres. We are only interested in the characteristics of these valence electrons, but the variational principle says that if we do not correctly model the low energy states then the energies of the higher lying states in which we are interested will be unreliable.

A schematic representation of the core and valence functions is shown in figure 2.5. The core wavefunctions are highly oscillatory in the vicinity of the ions. This is because the electrons have large kinetic energy within the core, and also high negative potential energy due to the ion.

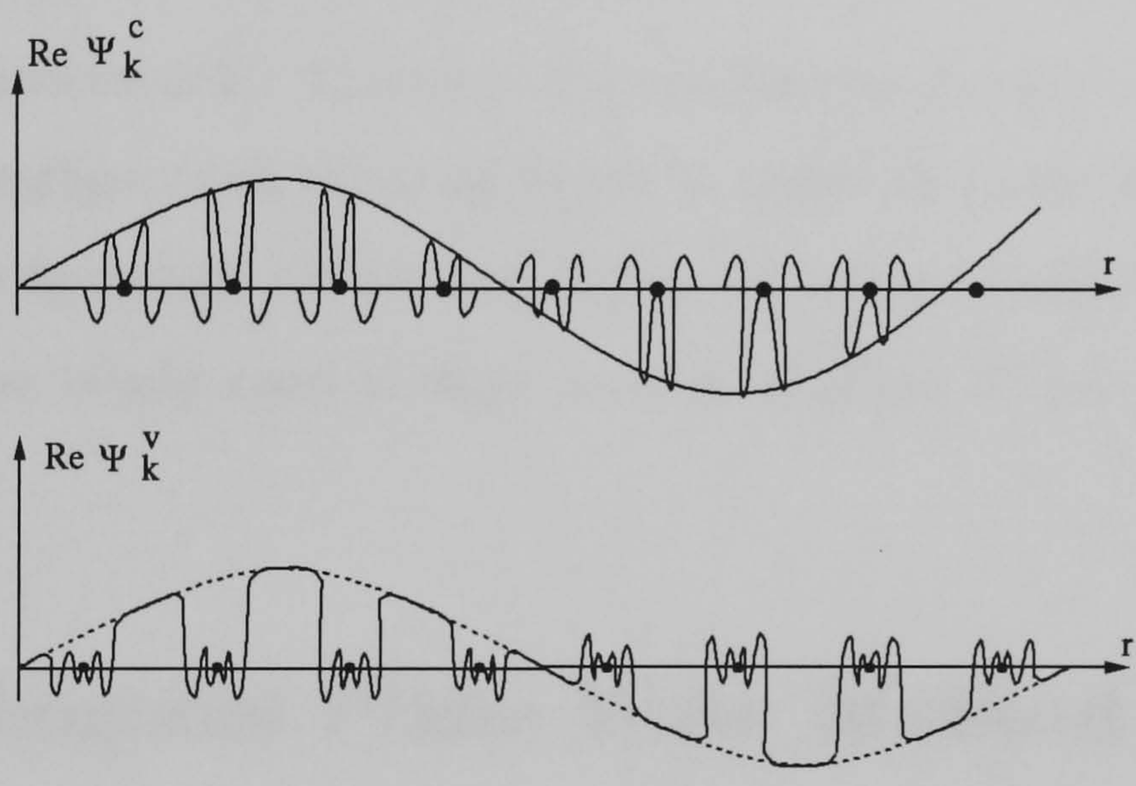


Figure 2.5: *Schematic representation of the form of the core and valence wavefunctions*

In the regions between the nuclei, the wavefunctions of the valence electrons would be expected to be smooth and slowly varying, thus plane waves would seem to be a good basis set for expanding the wavefunctions. However, around the cores, the valence states are also highly oscillatory and a great many plane waves would be needed to reproduce the states. The valence electrons have higher total energies than core electrons whilst experiencing the same negative potential energy, therefore they must have higher kinetic energies and thus

more oscillatory wavefunctions within the core region. Eigenstates of some Hamiltonian with different eigenvalues must be orthogonal, i.e. for any valence wavefunction $\psi^v_{\mathbf{k}}(\mathbf{r})$ and any core wavefunction $\psi^c_{\mathbf{k}}(\mathbf{r})$

$$\int \psi^c_{\mathbf{k}}(\mathbf{r})^* \psi^v_{\mathbf{k}}(\mathbf{r}) = 0 \tag{2.13}$$

Core wavefunctions only have appreciable magnitude around the immediate vicinity of the ion, therefore the main contribution to equation (2.13) must come from around the core region. From Bloch's theorem we know that we need only consider the region round a single ion (as the integrand must be the same from cell to cell). Therefore the oscillations of $\psi^v_{\mathbf{k}}(\mathbf{r})$ within the core region must interlace with those of $\psi^c_{\mathbf{k}}(\mathbf{r})$ in order to cause the integrals in equation (2.13) to vanish for all core levels. In order to expand the valence wavefunction we would need a large number of short λ, i.e. large $\mathbf{k}$ plane waves.

2.5 Orthogonal Plane Wave Method

In order to overcome this problem Herring (1940) suggested that a basis set comprised of plane waves orthogonalized to the core states be used. Thus we create an orthogonal plane wave, OPW_k, by adding a suitable linear combination of core orbitals, $\psi_c(\mathbf{r})$, to the appropriate plane waves. i.e.

$$OPW_k = e^{i\mathbf{k}.\mathbf{r}} + \sum_c b_c \psi_c(\mathbf{r}) \tag{2.14}$$

where the coefficients b_c are chosen such that the OPW_k's are orthogonal to every core state,

$$\int \psi^*_c(\mathbf{r}) OPW_k(\mathbf{r}) d\tau = 0. \tag{2.15}$$

Therefore

$$b_c = -\int \psi^*_c(\mathbf{r}) e^{i\mathbf{k}.\mathbf{r}} d\tau. \tag{2.16}$$

In Dirac notation we can express the OPW as

$$OPW_k = |\mathbf{k}\rangle - \sum_c |c\rangle\langle c|\mathbf{k}\rangle \tag{2.17}$$

where $|\mathbf{k}\rangle$ is the normalized plane wave and $|c\rangle$ is the normalized core function. Thus an OPW behaves like a plane wave far from the nuclei and like a core state in the vicinity of the ion. We can now expand our valence electron wavefunction using these OPW's as a basis set,

$$\phi_k(\mathbf{r}) = \sum_G a_G OPW_{k+G} = \sum_G a_G|\mathbf{k}+\mathbf{G}\rangle - \sum_c |c\rangle\langle c|\mathbf{k}+\mathbf{G}\rangle \tag{2.18}$$

where the $\mathbf{G}$'s are the reciprocal lattice vectors of the system. We can then substitute this expression into the Schrödinger Equation of (2.12) to give:

$$\begin{aligned}[\frac{-\hbar^2}{2m}\nabla^2 + V(\mathbf{r})]\sum_G a_G|\mathbf{k}+\mathbf{G}\rangle - \sum_c E_c\langle|c\rangle\langle c|\mathbf{k}+\mathbf{G}\rangle \\ = E_{\mathbf{k}}\sum_G |a_G\mathbf{k}+\mathbf{G}\rangle - \sum_c |c\rangle\langle c|\mathbf{k}+\mathbf{G}\rangle]\end{aligned} \tag{2.19}$$

To obtain the secular equation we multiply from the left by the plane waves $\langle\mathbf{k}+\mathbf{G}|'$.

$$\begin{aligned}\sum_G a_G \quad [\frac{-\hbar^2}{2m}|\mathbf{k}+\mathbf{G}|^2\delta(\mathbf{G},\mathbf{G}') + \langle\mathbf{k}+\mathbf{G}'|V|\mathbf{k}+\mathbf{G}\rangle - \sum_c E_c\langle\mathbf{k}+\mathbf{G}'|c\rangle\langle c|\mathbf{k}+\mathbf{G}\rangle] \\ = \quad E_{\mathbf{k}}\sum_G a_G[\delta(\mathbf{G},\mathbf{G}') - \sum_c\langle\mathbf{k}+\mathbf{G}'|c\rangle\langle c|\mathbf{k}+\mathbf{G}\rangle]\end{aligned} \tag{2.20}$$

This equation can then be diagonalized to produce the electronic eigenfunctions and eigenvectors. The method has the additional advantage over the plane wave expansion that only the eigenvalues of the valence and conduction bands are produced, as the basis functions are automatically orthogonal to the core states.

2.6 The Pseudopotential Method

The pseudo-potential method is really just an alternative way of writing the secular equation produced by the OPW method. We can simplify equation

(2.20) by collecting all terms involving core states:

$$[\frac{-\hbar^2}{2m}|\mathbf{k}+\mathbf{G}|^2\sum_G a_G|\mathbf{k}+\mathbf{G}\rangle + [E_k - E_c]\sum_G a_G\sum_c |c\rangle\langle c|\mathbf{k}+\mathbf{G}\rangle$$
$$= E_{\mathbf{k}}\sum_G a_G|\mathbf{k}+\mathbf{G}\rangle \qquad (2.21)$$

By comparison with the original Schrödinger Equation in equation (2.12) we can see that the wavefunction can be represented as

$$\phi_{\mathbf{k}}(\mathbf{r}) = \sum_{\mathbf{G}} a_{\mathbf{G}}|\mathbf{k}+\mathbf{G}\rangle \qquad (2.22)$$

This gives us an equation of the form

$$[\frac{-\hbar^2}{2m}\nabla^2 + W(\mathbf{r})]\phi_{\mathbf{k}}(\mathbf{r}) = E_{\mathbf{k}}\phi_{\mathbf{k}}(\mathbf{r}) \qquad (2.23)$$

where $W(\mathbf{r})$ is known as the pseudopotential and is given by

$$W(\mathbf{r}) = V(\mathbf{r}) + \sum_c [E_k - E_c]|c\rangle\langle c| \qquad (2.24)$$

By examining the form of the pseudopotential we can see that the attractive Coulomb potential, $V(\mathbf{r})$, is cancelled by the repulsion of the core electron and valence electrons. Thus the pseudopotential is weak and represents a smooth function over the whole crystal, i.e. the rapid fluctuation in the core regions are smoothed over. To obtain the secular equation, equation (2.21) is again multiplied on the left by $\langle \mathbf{k}+\mathbf{G}'|$

$$\|\sum_G a_G[\frac{-\hbar^2}{2m}|\mathbf{k}+\mathbf{G}|^2\delta(\mathbf{G},\mathbf{G}') + \langle \mathbf{k}+\mathbf{G}'|W|\mathbf{k}+\mathbf{G}\rangle - E_{\mathbf{k}}\delta(\mathbf{G},\mathbf{G}')]\| = 0. \qquad (2.25)$$

As in the OPW case, convergence would be expected to be rapid. It should be noted that the set of eigenfunctions $\psi_k(n)$ is not unique and any linear combination of core functions can be added and it will still be a valid solution of the Schrödinger Equation i.e.

$$\phi'_k(\mathbf{r}) = \phi_k(\mathbf{r}) + \sum_c a_c\psi_c(\mathbf{r}) \qquad (2.26)$$

We can include this non-uniqueness by the inclusion of an arbitrary operator in the definition of the pseudopotential (Bassani 1975).

2.6.1 The Local Pseudopotential

We can further simplify the problem by assuming that the Coulomb potential, $V(\mathbf{r})$ can be expressed as a sum of the individual potentials at each atom.

$$V(\mathbf{r}) = \sum_j v(\mathbf{r} - \mathbf{R_j}) \tag{2.27}$$

where j labels the atomic sites. Similarly we can write the pseudopotential as sum of spherical potentials centered on each lattice site,

$$W = \sum_j w(\mathbf{r} - \mathbf{R_j}) \tag{2.28}$$

We can now write the matrix elements from the secular equation (2.25) as

$$\begin{aligned}
\langle \mathbf{k}+\mathbf{G}'|W|\mathbf{k}+\mathbf{G}\rangle &= \frac{1}{\Omega}\int e^{-i(\mathbf{k}+\mathbf{G}).\mathbf{r}} \sum_j w(\mathbf{r}-\mathbf{R_j}) e^{i\mathbf{k}.\mathbf{r}} d\tau \\
&= \frac{1}{\Omega}\sum_j e^{-i(\mathbf{G}.\mathbf{R_j})} \int e^{-i(\mathbf{k}+\mathbf{G}).(\mathbf{r}-\mathbf{R_j})} w(\mathbf{r}-\mathbf{R_j}) e^{-i(\mathbf{k}.\mathbf{r}-\mathbf{R_j})} d\tau \\
&= \frac{1}{N\Omega_0}\sum_j e^{-i(\mathbf{G}.\mathbf{R_j})} \int e^{-i(\mathbf{k}+\mathbf{G}).\mathbf{r}} w(\mathbf{r}) e^{i\mathbf{k}.\mathbf{r}} d\tau
\end{aligned} \tag{2.29}$$

where Ω is the crystal volume which appear due to the normalization constant $\Omega^{1/2}$ in the pseudo-function; Ω_0 is the atomic volume and N is the number of atoms.

The matrix elements can now be separated into the form

$$\langle \mathbf{k}+\mathbf{G}'|W|\mathbf{k}+\mathbf{G}\rangle = S_{\mathbf{G}} V_{\mathbf{G}} \tag{2.30}$$

where $S_{\mathbf{G}}$ is the structure factor

$$S_{\mathbf{G}} = \frac{1}{N}\sum_j e^{i\mathbf{G}.\mathbf{R_j}} \tag{2.31}$$

which depends only on atomic position, and a form factor

$$V_{\mathbf{G}} = \langle \mathbf{k}+\mathbf{G}|w|\mathbf{k}\rangle = \frac{1}{\Omega_0}\int e^{-i(\mathbf{k}+\mathbf{G}).\mathbf{r}} w(\mathbf{r}) e^{i\mathbf{k}.\mathbf{r}} = \frac{1}{\Omega_0}\int w(\mathbf{r}) e^{-i\mathbf{G}.\mathbf{r}} d\mathbf{r} \quad (2.32)$$

which depends only on individual atomic potentials.

2.6.2 Spin-Orbit Coupling

Up until now no mention has been made of spin-orbit coupling. The summations in the previous sections can be considered to implicitly include a summation over the spin states. To fully describe the band structure of Si and Ge we need to consider the spin-orbit coupling resulting from the interaction of the magnetic moment of the electron spin with the magnetic field experienced by the electron. The spin-orbit interaction lowers the symmetry of the Hamiltonian so that states that were degenerate are split. Other relativistic corrections to the Hamiltonian which do not involve spin may actually be larger in magnitude that the spin-orbit term but they are less noticeable because they do not lower the symmetry of the system. These other effects are automatically included in an empirical approach where parameters are determined from experimental data. The spin-orbit interaction mixes states but does not destroy the underlying polarization selection rules on the spatial wavefunctions (Niles 1992).

In the absence of spin-orbit coupling and strain the top of the valence band at the zone centre has Γ_{15} symmetry i.e. atomic p-like structure, and is 6-fold degenerate. The spin orbit interaction splits this into a four-fold degenerate Γ_8 band and a 2-fold degenerate Γ_7 band known as the spin split-off band. We include this interaction by modifying the form of the Hamiltonian, which becomes

$$H = -\frac{\hbar^2}{2m}\nabla^2 + V_L(\mathbf{r}) + V_{SO}(\mathbf{r}) \quad (2.33)$$

where V_L is the local pseudopotential and V_{SO} is the spin-orbit coupling term. The Schrödinger Equation then becomes

$$H\phi_{n,\mathbf{k}}(\mathbf{r},s) = E_{n,\mathbf{k},s}\phi_{n\mathbf{k}}(\mathbf{r},s) \tag{2.34}$$

where n is the band index, $\mathbf{k}$ is the wavevector in the first B.Z. and s is the spin index. The bulk wavefunction is given by

$$|\phi_{\mathbf{k},s}\rangle = \sum_{\mathbf{G}} a_{\mathbf{G},s}|\mathbf{k}+\mathbf{G},s\rangle \tag{2.35}$$

The matrix elements of this operator H, using equation (2.25) are given by ,

$$\begin{aligned}\langle \mathbf{k}+\mathbf{G}'|H|\mathbf{k}+\mathbf{G}\rangle = \frac{\hbar^2}{2m}|\mathbf{k}+\mathbf{G}|^2\delta(\mathbf{G},\mathbf{G}')\delta(\mathbf{s},\mathbf{s}') \\ +S_{\mathbf{G}}V_{\mathbf{G}}\delta(\mathbf{s},\mathbf{s}') + \langle \mathbf{k}+\mathbf{G}',s'|V_{SO}|\mathbf{k}+\mathbf{G},\mathbf{s}\rangle\end{aligned} \tag{2.36}$$

To evaluate the spin-orbit matrix elements we use the method developed by Weisz (1966) and modified by Bloom (1968) which gives as an expression for the matrix elements of the spin-orbit Hamiltonian

$$\langle \mathbf{k}+\mathbf{G}',s'|V_{SO}|\mathbf{k}+\mathbf{G},s\rangle = (\mathbf{k}+\mathbf{G}')\wedge(\mathbf{k}+\mathbf{G}).\sigma_{s,s'}[-i\lambda^s S_{\mathbf{G}}] \tag{2.37}$$

where $\sigma_{ss'}$ is the Pauli spin matrix. The symmetric contribution to the spin-orbit coupling is denoted λ^s. There is also an antisymmetric contribution λ^a but as we have no anti-symmetric structure factors, (as the anions and cations are the same species in Si and Ge) this will not contribute. λ^s is defined as

$$\lambda^s = \frac{1}{2}\mu(1+\alpha) \tag{2.38}$$

$$\lambda^a = \frac{1}{2}\mu(1-\alpha) \tag{2.39}$$

where the form of λ^a has been included for completeness. The value of α is equal to the ratio of the splittings in the free atoms (Hermann 1963). μ is treated as an adjustable parameter which is fitted empirically to reproduce

the spin splitting in the bulk unstrained materials. Thus our final expression for the pseudopotential matrix is

$$\langle \mathbf{k}+\mathbf{G}'|H|\mathbf{k}+\mathbf{G}\rangle = \frac{\hbar^2}{2m}|\mathbf{k}+\mathbf{G}|^2\delta(\mathbf{G},\mathbf{G}')\delta(\mathbf{s},\mathbf{s}') + S_{\mathbf{G}}V_{\mathbf{G}}\delta(\mathbf{s},\mathbf{s}') + (\mathbf{k}+\mathbf{G}')\wedge(\mathbf{k}+\mathbf{G}).\sigma_{s,s'}[-i\lambda^s S^s\mathbf{G}] \quad (2.40)$$

This equation is again solved by direct diagonalization to give required eigenfunctions and eigenvalues for the bulk material.

2.6.3 Empirical Pseudopotentials

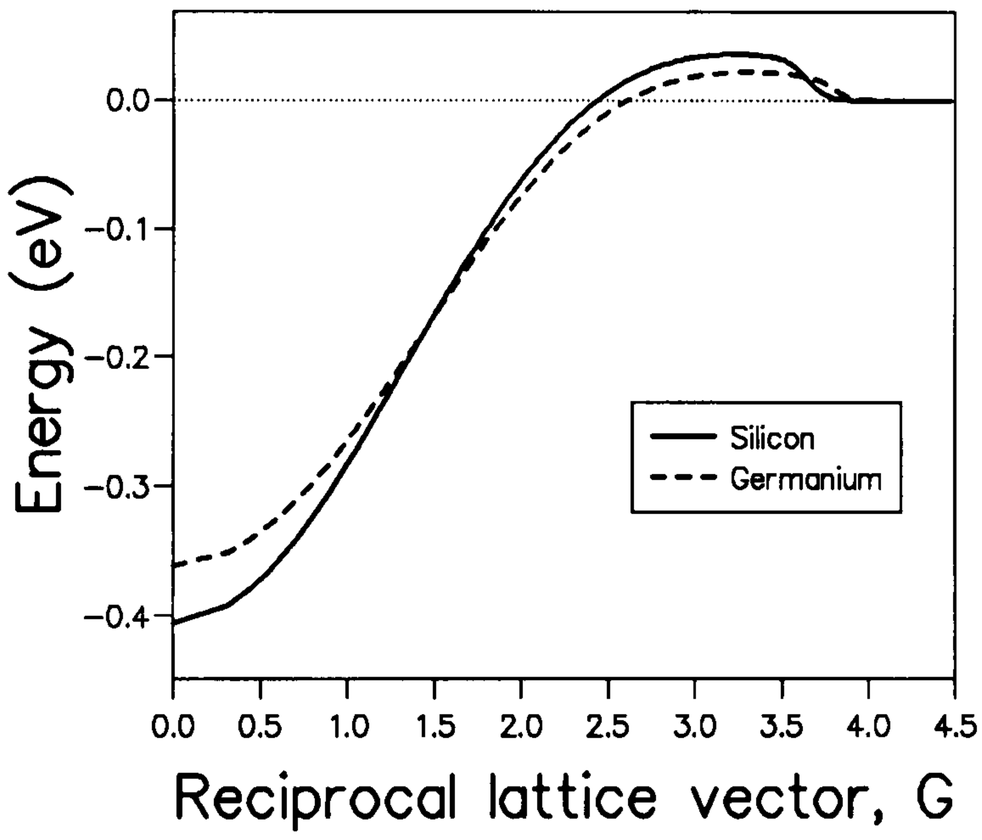


Figure 2.6: *Plot of the form of the empirical pseudopotentials for Si and Ge*

Examination of the structure factor in equation (2.31) shows that this can be easily calculated for any given $\mathbf{R}$ and $\mathbf{G}$. We use a basis set comprising of the first 65 reciprocal lattice vectors. This leaves only the determination of the form factors, which requires us to find a form for the atomic potentials.

We use a model devised by Friedel, Hybertsen and Schlüter (1988). The basic expression used to describe the variation of the pseudopotential with

Material	a_1	a_2	a_3	a_4	a_5	a_6
Si	106.0686	2.2278	0.6060	-1.9720	5.0	0.3
Ge	54.4512	2.3592	0.7400	-0.3800	5.0	0.3

Table 2.2: *Coefficients for the potentials for Silicon and Germanium from Friedel (1986). All values are in atomic units*

reciprocal lattice vector $\mathbf{G}$ is given by

$$V_a(\mathbf{G}) = \frac{a_1(G^2 - a_2)}{e^{a_3(G^2 - a_4)} + 1}\left[\frac{1}{2}tanh\left[\frac{a_5 - G^2}{a_6}\right] + \frac{1}{2}\right] \tag{2.41}$$

where the values for the constants a_i are given for silicon and germanium in table 2.2. The term in square brackets is included to ensure a fast cutoff for large $\mathbf{G}$ and has little effect on the eigenvalues. This expression was obtained by fitting a curve to three points $V_a(\mathbf{G})$ corresponding to the stars of the vectors $(1,1,1),(2,2,0)$ and $(3,1,1)$. 110 plane waves were included. To fit the $\mathbf{G} \to 0$ limit the calculated electron-phonon matrix elements were fitted to experimental values. The maxima of the curve was 'forced' to be between $\mathbf{G} =$ (220) and (311) thus ensuring that the derivative at (311) was negative. Also the valence band offsets obtained using the above expression were compared to those obtained using a microscopic L.D.A. approach. The above three steps were repeated self-consistently until convergence was achieved. The potentials obtained by this method for silicon and germanium are shown in figure 2.6.

2.6.4 Band Structures for Silicon and Germanium

We now apply the above theory to calculate the bulk band structures for Si and Ge which are the constituents of all quantum well structures discussed in this thesis. The first Brillouin zone for a crystal with a diamond lattice in shown in figure 2.4. Points of high symmetry are marked; Γ is the zone

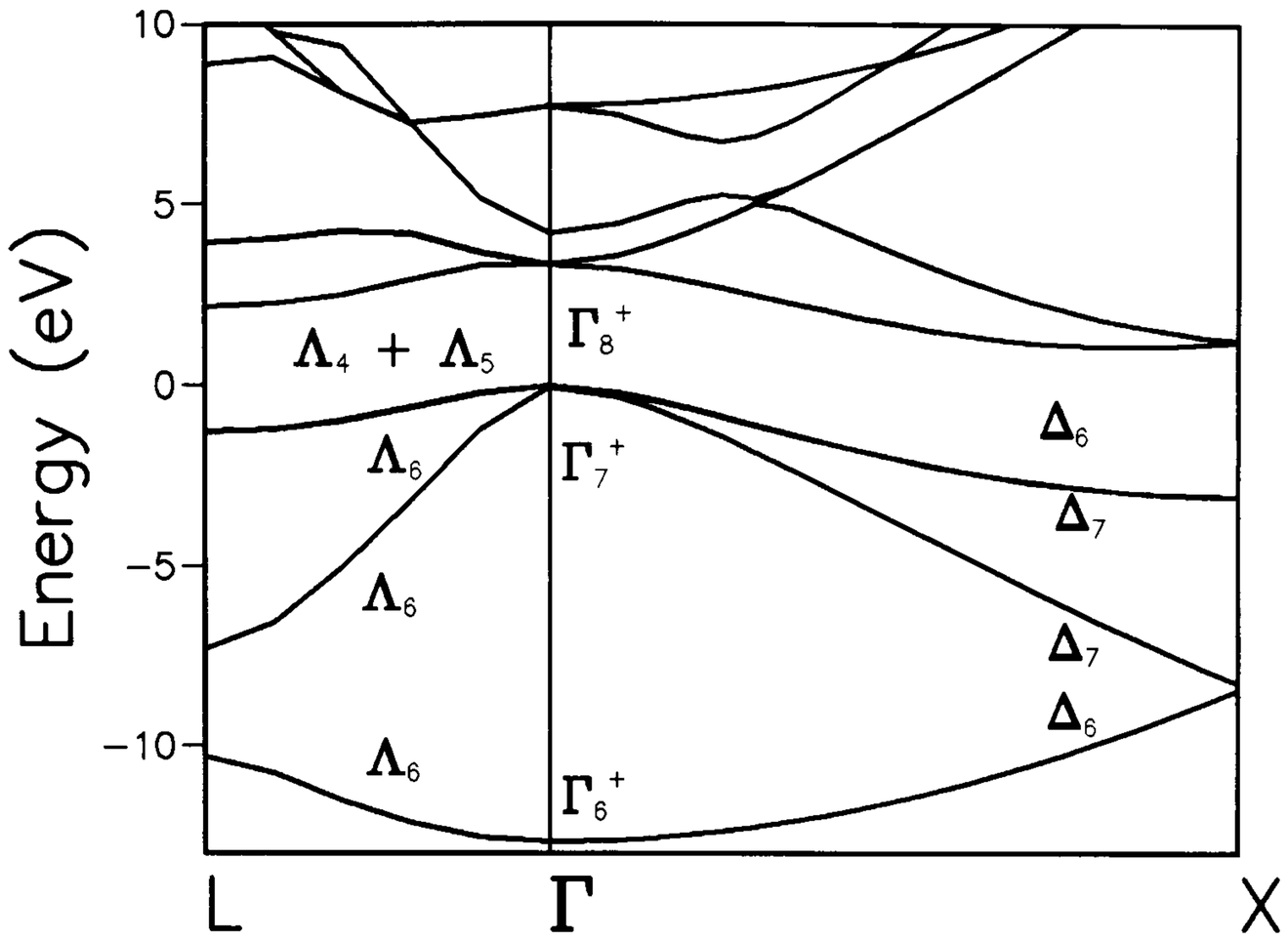


Figure 2.7: *The band structure for bulk silicon along the symmetry lines L-Γ -X.*

centre, X is along the k_x axis and L is along $(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$. The irreducible segment is indicated which is determined by symmetry considerations (Bassani 1975). Calculating the band structure along the symmetry lines $X - \Gamma - L$ produces for Si and Ge the band structure shown in Figure 2.7 and 2.8.

The effect of spin orbit coupling is to lift the six fold degeneracy at the top of the valence band at the zone centre. The heavy hole (HH) and light hole (LH) bands are shifted up in energy by an amount $\frac{1}{3}\Delta_0$ and the spin split off band (SO) lowered by $\frac{2}{3}\Delta_0$. The values for Δ_0 have been determined experimentally using techniques such as electro-reflectance (see, for example, Akasaka 1974 or Daunois 1978) The accepted values for Δ_0 are 44meV for Si and 290meV for Ge. To reproduce these values we require a value of μ in equation (2.38) of $\mu = 0.00029$ atomic units for Si and $\mu = 0.00161$ a.u. for Ge. We can see this splitting in the Ge bandstructure in figure 2.8. The

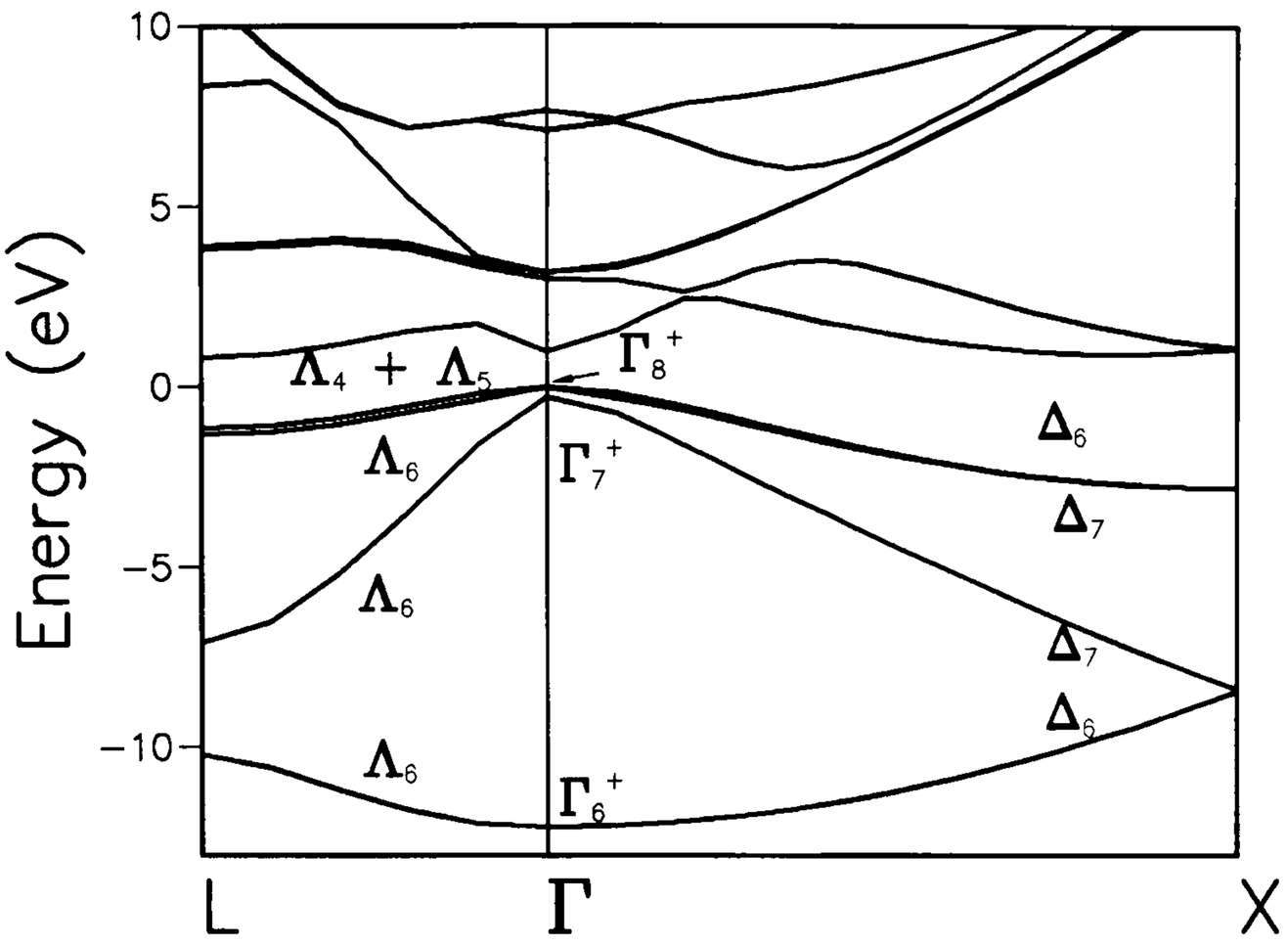


Figure 2.8: *The band structure for bulk germanium along the symmetry lines* L-Γ - X

spitting in the Si structure in figure 2.7 is too small to see on this scale.

Chapter 3

The Quantum Well Calculation

A multiple quantum well consists of alternating layers of different materials, typically in our case $Si_{0.7}Ge_{0.3}$ sandwiched between Si barriers, as shown in figure 3.1. These constituents have different band gaps, thus discontinuities in the potential profile will result from the direct growth of one on the other.

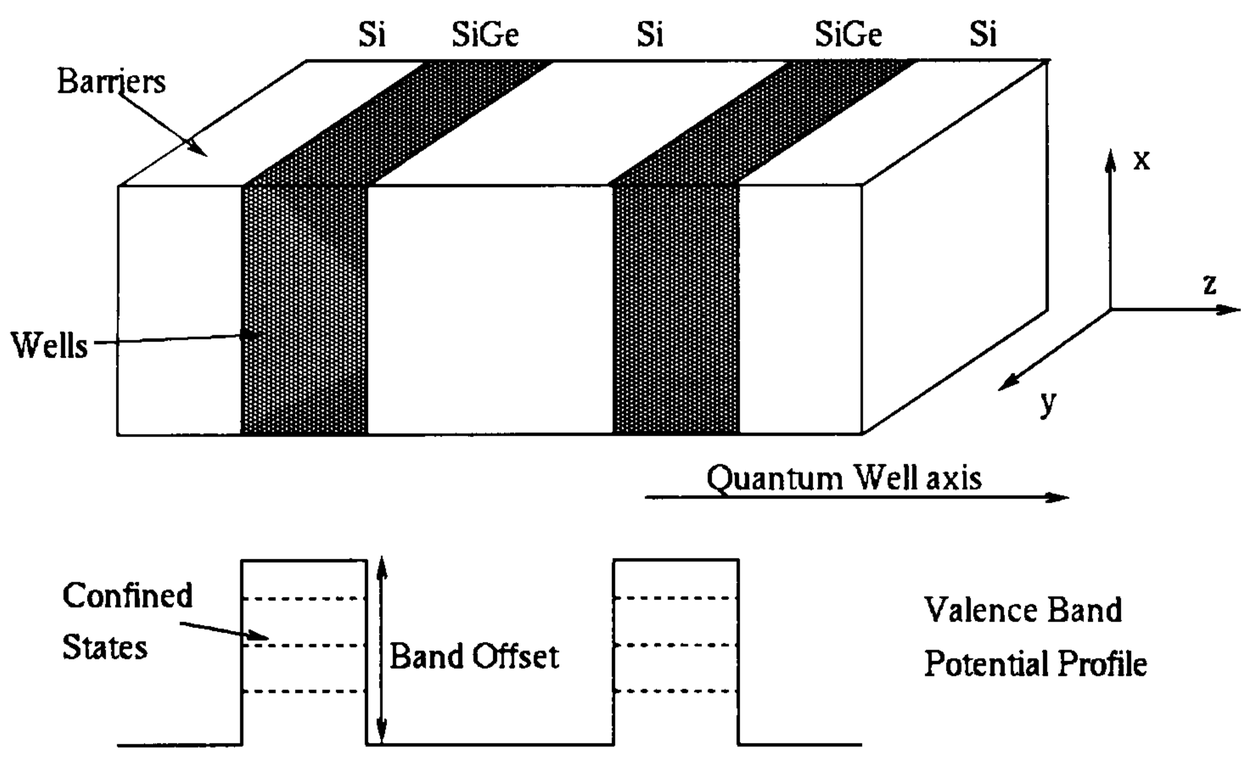


Figure 3.1: *Schematic representation of a quantum well structure. The discontinuities in the valence band potential profile are shown. The z-axis is defined to be the growth axis.*

3.1 Symmetry Properties of the Quantum Well

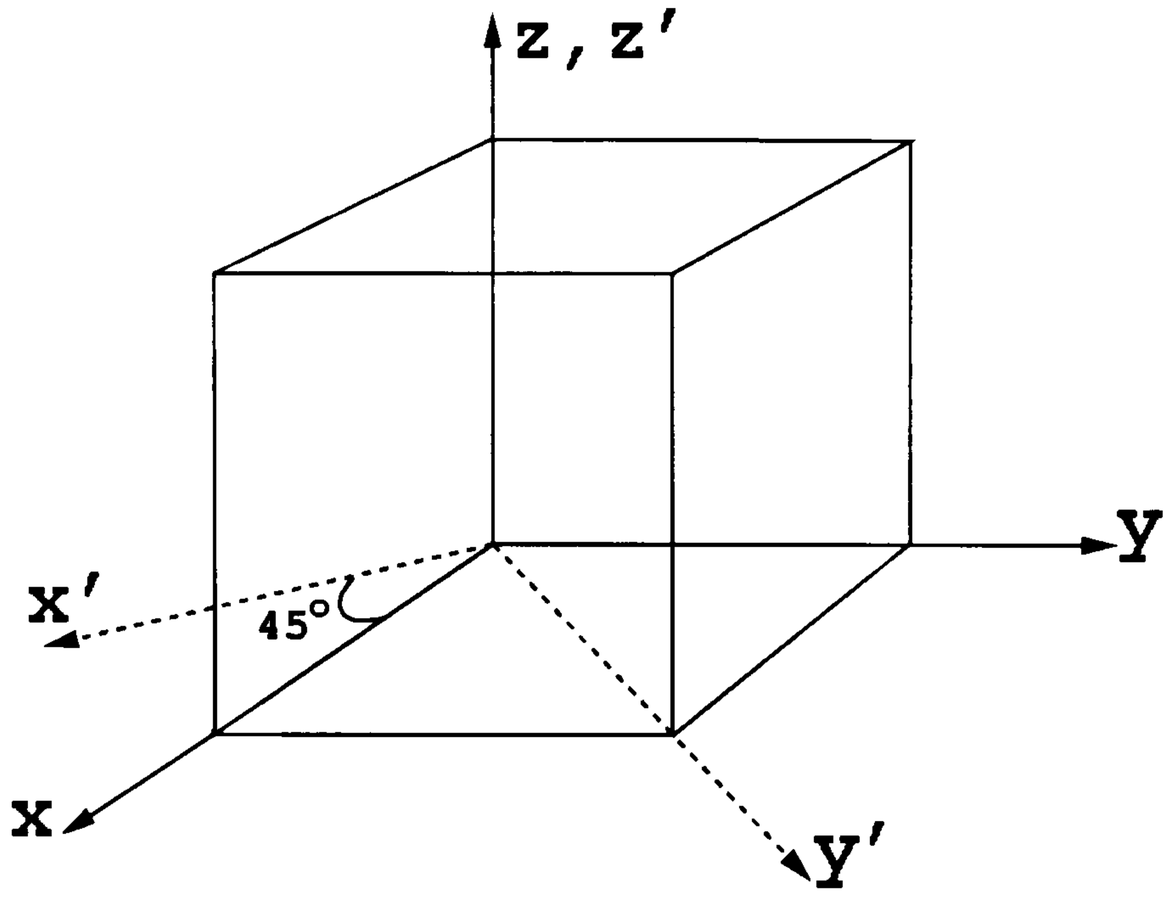


Figure 3.2: *The relation between the two axis frames. The axes labelled* (x, y, z) *are usually used for studying diamond structures. The conventional reference frame for the* C_{2v} *symmetry group is the axes* (x', y', z').

The standard Schönflies notation is used for symmetry operators throughout this section (Bradley 1972). The symmetry operation of a space group $\mathcal{G}$ is denoted $\{\mathbf{R}|\mathbf{t}\}$, where $\mathbf{R}$ is a point symmetry operation, i.e. a rotation, reflection or inversion etc. $\mathbf{t}$ is a translation vector. In general $\mathbf{t} = \mathbf{G} + \tau$ where $\mathbf{G}$ is a reciprocal lattice vector and τ is a fractional translation smaller than any lattice vector. If all τ's in a space group are zero, then the group is said to be symmorphic. In diamond structure, both strained and unstrained, there exists non-primitive τ that translate the crystal structure onto itself, i.e. $\mathbf{t} = (1,1,1)\frac{a_0}{4}$ where a_0 is the lattice constant of the unstrained diamond structure. Introducing strain changes the degree of symmetry and some of the point group operations disappear and the translational vectors may change.

The space group for unstrained diamond is $\mathcal{O}_h^7$ (see section 2.1). For Ge grown on Si substrates some of the symmetry operations of this group are no longer present, namely the improper rotations (i.e. rotations followed by inversions) and the threefold rotations. The space group for the distorted bulk lattice is $\mathcal{D}_{4h}^{19}$ (Hahn 1983).

We consider the z-axis to be the growth direction. The x and y axes were originally defined to be along the edge of the conventional face-centered cubic unit cell. This is the reference frame in which our lattice vectors were defined in terms of in figure 2.1. Our susceptibility tensors are also defined in this frame. For the purpose of considering the symmetry of the quantum well a different set of axes are more convenient. The x' an y' axes are in the plane parallel to the interfaces, and are defined to lie diagonally across the face of the conventional face-centered cubic unit cell. The two sets of axis, shown in figure 3.2, are rotated by 45° to each other. The z-direction remains unchanged.

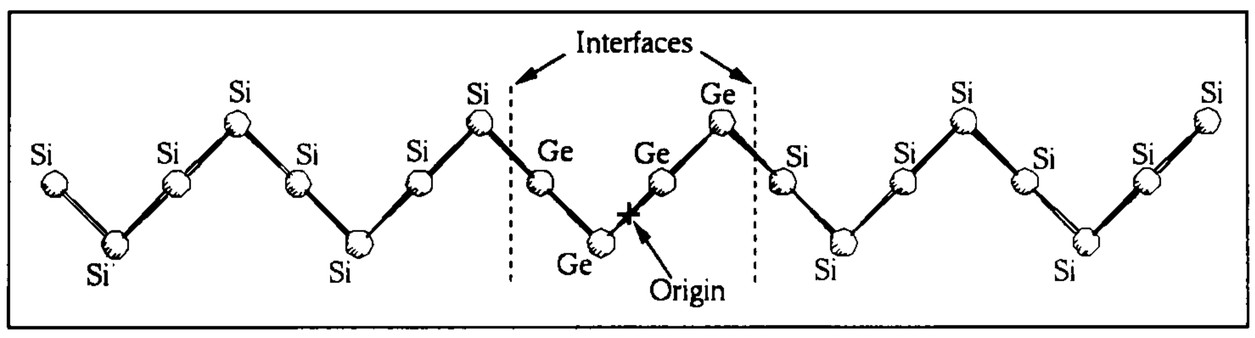


Figure 3.3: *A schematic representation of a single spiral of atoms indicating the position of the interfaces and the origin of the system.*

Consider the spiral structure shown in figure 3.3, consisting of 1 lattice constant of germanium, or alloy 'pseudo-atoms', surrounded by several lattice constants of Si. The interfaces are considered to occur at the mid- bond positions indicated. The point of highest symmetry, taken to be the origin, is the mid-bond in the centre of the well in the z-direction.

The atomic positions of a strained SiGe system are shown in figure 3.4.

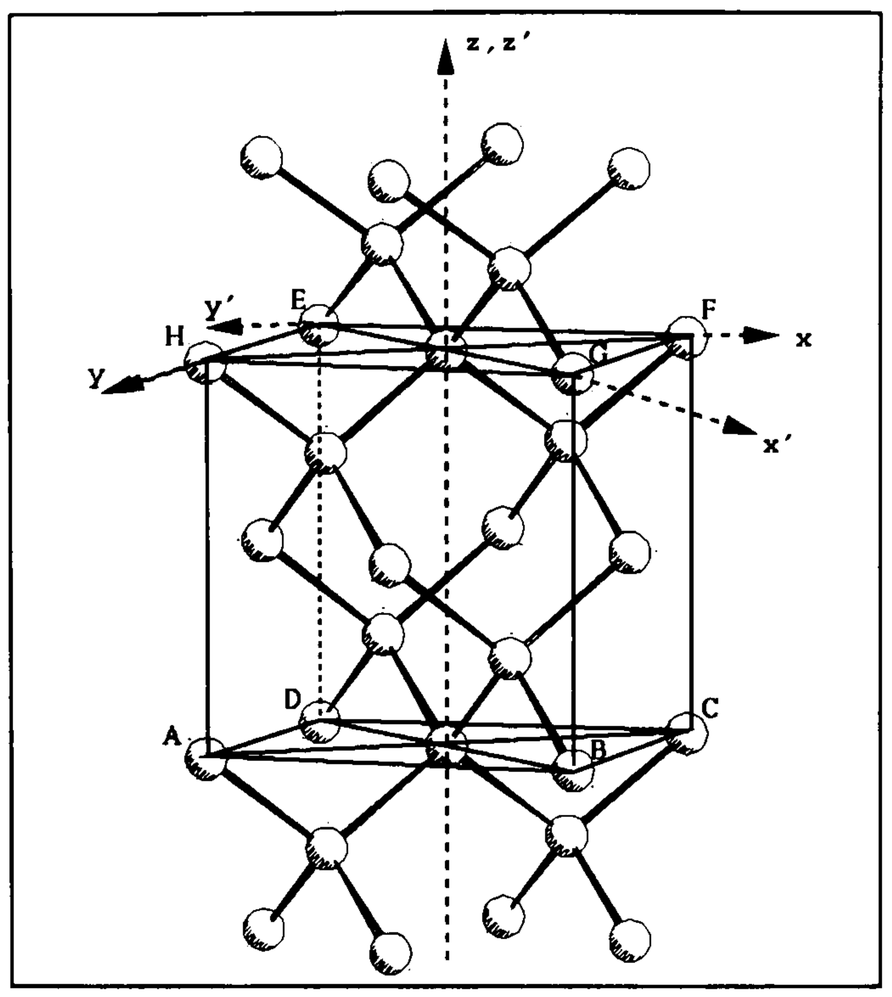


Figure 3.4: *The atomic positions in a strained SiGe/Si quantum well system. The symmetry axes and planes are indicated. The mirror planes $\sigma(y')$ is $BDEG$ and the mirror plane $\sigma(x')$ is $ACFH$*

The system is invariant under the operations of E, the identity operator; C_{2z}, a rotation of π radians about the principal axis, which is taken to be the axis of highest symmetry, in this case the z-direction; $\sigma_{x'}$, a reflection in the $x'z'$ plane and $\sigma_{y'}$, a reflection in the $y'z'$ plane. These operations comprise the symmetry point group C_{2v}. The character table for the point group for C_{2v} symmetry is given in table 3.1. The space group for the quantum well is D_{2h}.

Consider a general point with real space coordinates (x, y, z) and wavevector space coordinates (k_x, k_y, k_z). The effect of the above symmetry operators on such a point are given in table (3.2). The Hamiltonian of the system is by definition invariant under these symmetry operators. Thus for a general symmetry operator S which appears in the point group we have

$$E(\mathbf{k}) = E(S\mathbf{k}) \tag{3.1}$$

	E	C_{2z}	$\sigma_{x'}$	$\sigma_{y'}$
Δ_1	1	1	1	1
Δ_2	1	1	-1	-1
Δ_3	1	-1	1	-1
Δ_4	1	-1	-1	1

Table 3.1: *The character table for the point group associated with the quantum well structure,* C_{2v}.

Operation	x	y	z	k_x	k_y	k_z
E	x	y	z	k_x	k_y	k_z
σ_x	y	x	z	k_y	k_x	k_z
σ_y	$-y$	$-x$	z	$-k_y$	$-k_x$	k_z
C_{2z}	$-x$	$-y$	z	$-k_x$	$-k_y$	k_z

Table 3.2: *The effect of operating on the real space point (x,y,z) and the point in wavevector space* (k_x, k_y, k_z) *with the symmetry operations of the* C_{2v} *point group.*

where $E(\mathbf{k})$ is an energy eigenvalue with the wavevector $\mathbf{k}$.

It is shown in appendix A that for any matrix element P_{ab} calculated in the irreducible segment can be related to matrix elements in other areas of the zone by

$$P^{\mathrm{e}}_{ab}(\mathcal{S}\mathbf{k}) = P^{\mathcal{S}^{-1}\mathrm{e}}_{ab}(\mathbf{k}). \tag{3.2}$$

Therefore, given knowledge of the matrix element at wavevector $\mathbf{k}$ we can deduce the matrix element at $(\mathcal{S}\mathbf{k})$. Thus we need only find the matrix elements for wavevectors present in the irreducible segment of the first Brillouin zone and from there information about the rest of the zone can be obtained by use of the symmetry operators.

3.1.1 Time Reversal Symmetry

The Hamiltonian of the crystal $H(\mathbf{r})$, is real and so are all the eigenvalues. Thus taking the conjugate of the Schrödinger equation we get

$$H(\mathbf{r})\phi_{\mathbf{k}}(\mathbf{r})^* = E(\mathbf{k})\phi_{\mathbf{k}}(\mathbf{r})^* \tag{3.3}$$

$\phi_{\mathbf{k}}(\mathbf{r})^*$ is simply a Bloch function corresponding to a wavevector of $-\mathbf{k}$, therefore

$$E(\mathbf{k}) = E(-\mathbf{k}) \tag{3.4}$$

for all $\mathbf{k}$, therefore $E_{\mathbf{k}}$ has inversion symmetry even when the point group does not contain the inversion operator. If this operator is applied to the time dependent Schrödinger equation we get

$$H(\mathbf{r})\psi^* = i\hbar\frac{\partial\psi^*}{\partial(-t)} \tag{3.5}$$

i.e. the sign of the time is reversed.

For the matrix elements the operation $\mathcal{T}$ does not correspond directly to the inversion operator as P_{ab} is in general a complex number.

$$P^{\mathbf{e}}_{ab}(\mathbf{k}) = -(P^{\mathbf{e}}_{ab}(-\mathbf{k}))^* \tag{3.6}$$

Applying equation (3.1) to the four wavevectors in table (3.2) we obtain eight wavevectors for each miniband energy. Using (3.6) and (3.2) we get corresponding values for the momentum matrix elements which can be generated for each wavevector in the irreducible segment. Table (3.3) explicitly states the effects of all 8 transformations on the wavevectors and momentum matrix elements. Thus we can now construct the irreducible segment of the first Brillouin zone for a structure with C_{2v} point group symmetry, shown in figure 3.1

Operation	k_x	k_y	k_z	P_x	P_y	P_z
E	k_x	k_y	k_z	P_x	P_y	P_z
σ_{ad}	k_y	k_x	k_z	P_y	P_x	P_z
σ_{bd}	$-k_y$	$-k_x$	k_z	$-P_y$	$-P_x$	P_z
C_{2z}	$-k_x$	$-k_y$	k_z	$-P_x$	$-P_y$	P_z
$\mathcal{T}E$	$-k_x$	$-k_y$	$-k_z$	$-P_x^*$	$-P_y^*$	$-P_z^*$
$\mathcal{T}\sigma_{ad}$	$-k_y$	$-k_x$	$-k_z$	$-P_y^*$	$-P_x^*$	$-P_z^*$
$\mathcal{T}\sigma_{bd}$	k_y	k_x	$-k_z$	P_y^*	P_x^*	$-P_z^*$
$\mathcal{T}C_{2z}$	k_x	k_y	$-k_z$	P_x^*	P_y^*	$-P_z^*$

Table 3.3: *The eight related wavevectors under the C_{2v} point group and time reversal symmetry operations, and the Cartesian components of the momentum matrix elements at these wavevectors. $\mathcal{T}$ represents the time reversal operation.*

The width of the Brillouin zone is governed by the period of the quantum well, where the period is equal to the number of lattice constants in the unit cell of the quantum well. A lattice constant is defined as four monolayers (i.e. single atomic planes) of either Si, Ge or alloy atoms.

3.1.2 The Zone Folding Scheme

A first idea of the band structure of the quantum well can be obtained by folding the band structure of the host material. The bulk band structure is folded at values of k_z related to the quantum well reciprocal lattice vectors, e.g. a typical structure consists of 80 monolayer Si barriers and a 20 monolayer $Si_{0.7}Ge_{0.3}$ alloy well. The width of the B.Z. in this structure is 1/period = $1/25 = 0.04(\frac{2\pi}{a})$ where a is the lattice constant of the host material.

Consider the uppermost valence band of our host material. We can see

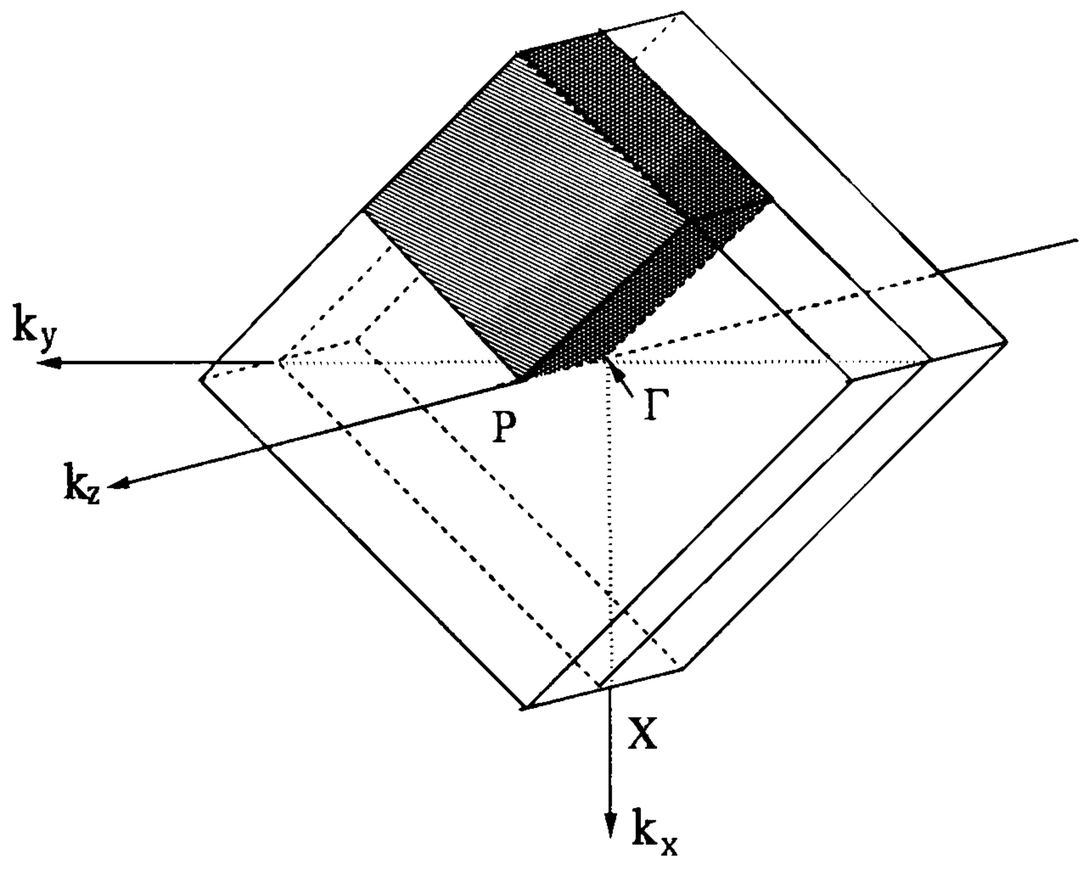


Figure 3.5: *The first Brillouin Zone for a quantum well structure. The width of the zone is governed by the period of the quantum well structure. Points of high symmetry that are referred to in the text are indicated The shaded region represents the irreducible segment*

from figure 3.6 that the band structure is folded at $k_z = (0.02)\frac{2\pi}{a}$ where a is the host lattice constant, in an analogous manner to the transition from extended- to reduced- zone representation. Thus we can obtain a rough idea of the electron-wavevector dispersion curves in the quantum well. This is a very simple model, however, and there are many other factors to take into account in the change from the bulk material to the quantum well. Notably the difference in the potentials and lattice constants of the constituent materials, the discontinuities in the potential profile caused by the differing band gaps of the materials and the interaction between states folded onto the same k_z causing momentum mixing between states of different bulk origin.

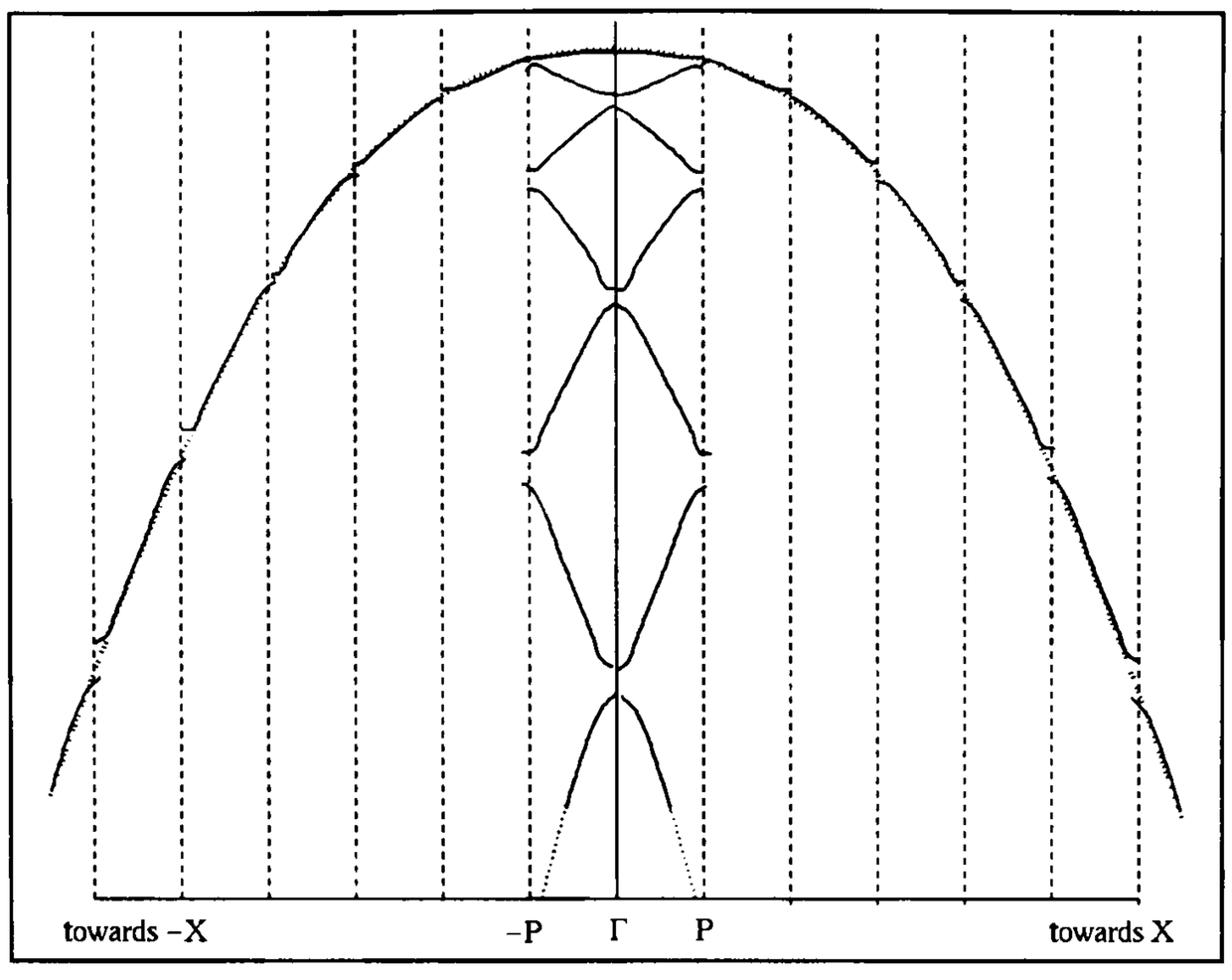


Figure 3.6: *Schematic representation of the band folding scheme showing the folding of the bulk band through the period of the quantum well.*

3.2 The Pseudopotential Quantum well Calculation

In order to model a quantum well we need to find an appropriate basis set with which to expand the quantum well wavefunctions. This is the bulk pseudopotential eigenfunctions for an appropriate 'host' material, often taken to be an average of the constituents of the quantum well stem. In this thesis we always used strained silicon as our host. We begin by considering the Schrödinger Equation for the bulk host material.

$$H_0\phi_{n\mathbf{k}}(\mathbf{r},s) = E_{n\mathbf{k}}\phi_{n\mathbf{k}}(\mathbf{r},s) \tag{3.7}$$

The renormalized bulk wavefunctions can be expressed as

$$\phi_{n\mathbf{k}}(\mathbf{r},s) = \frac{1}{\sqrt{\Omega}}\sum_{\mathbf{G}}\sum_{s} a_{n\mathbf{k}}(\mathbf{G},\mathbf{s})e^{i(\mathbf{k}+\mathbf{G}).\mathbf{r}}|s\rangle \tag{3.8}$$

where Ω is the crystal volume. In order to simplify notation we shall drop the spin indices and consider spin to be implicitly included.

The quantum well Hamiltonian, H_{QW}, is given by

$$H_{QW}\psi = (H_0 + V)\psi = E\psi \tag{3.9}$$

where H_0 is the bulk Hamiltonian and V describes the difference in potential between the quantum well and the host crystal due to the effects mentioned in the previous section. The quantum well eigenfunctions are expanded using the eigenfunctions of the host crystal, $\phi_{n\mathbf{k}}$, as the basis set.

$$\psi = \sum_n \sum_k A_{n\mathbf{k}}\phi_{n\mathbf{k}} \tag{3.10}$$

where n is the number of bulk bands and k is the bulk wavevectors which are folded onto the particular quantum well wavevectors. The coefficients $A_{n\mathbf{k}}$ represent the contribution of various bulk states to the quantum well wavefunction. Substituting 3.8 into 3.9, multiplying from the left by $\phi_{n'\mathbf{k}'}$ and integrating over all space gives a matrix equation of the form

$$A_{n'\mathbf{k}'}(E_{n'\mathbf{k}'} - E) + \sum_n \sum_{\mathbf{k}} A_{n\mathbf{k}}\langle\phi_{n'\mathbf{k}'}|V|\phi_{n\mathbf{k}}\rangle = 0 \tag{3.11}$$

where the orthogonality of the bulk states i.e. $\langle\phi_{n\mathbf{k}}|\phi_{n'\mathbf{k}'}\rangle = \delta_{n,n'}\delta_{\mathbf{k},\mathbf{k}'}$ has been invoked. Substituting for the form of the host eigenfunctions from equation 3.8 gives

$$A_{n'\mathbf{k}'}(E_{n'\mathbf{k}'} - E) + \sum_n \sum_{\mathbf{k}} \sum_{\mathbf{G},\mathbf{G}'} A_{n\mathbf{k}}a_{n'\mathbf{k}'}(\mathbf{G}')a_{n\mathbf{k}}(\mathbf{G})\frac{1}{\Omega}\langle\mathbf{k}' + \mathbf{G}'|V|\mathbf{k} + \mathbf{G}\rangle = 0 \tag{3.12}$$

Incorporating the local approximation reduces the matrix element of the above to

$$\begin{aligned}\frac{1}{\Omega}\langle\mathbf{k}' + \mathbf{G}'|V|\mathbf{k} + \mathbf{G}\rangle &= \frac{1}{\Omega}\int e^{-i(\mathbf{k}'+\mathbf{G}').\mathbf{r}}V(\mathbf{r})e^{i(\mathbf{k}+\mathbf{G}).\mathbf{r}}d\tau \\ &= \frac{1}{\Omega}\int e^{-i\mathbf{g}.\mathbf{r}}V d\tau \end{aligned} \tag{3.13}$$

where $\mathbf{g}$ is the quantum well reciprocal lattice vector defined as

$$\mathbf{g} = \mathbf{k}' + \mathbf{G}' - \mathbf{k} - \mathbf{G} \tag{3.14}$$

We now need an expression for the potential V. The total potential in the local approximation can be written as the sum of individual potentials at each atom, therefore for the quantum well this is

$$\begin{aligned} V(\mathbf{r}) &= V_{QW}(\mathbf{r}) - V_0(\mathbf{r}) \qquad (3.15) \\ &= \sum_j [\sum_{t_s} v_s(\mathbf{r} - \mathbf{t_s} - \mathbf{R_j}) + \sum_{t_a} v_a(\mathbf{r} - \mathbf{t_a} - \mathbf{R_j}) - \sum_{t_0} v_0(\mathbf{r} - \mathbf{t_0} - \mathbf{R_j})] \end{aligned}$$

where $\mathbf{R_j}$ are the quantum well lattice vectors, t_s and t_a are the positions of the silicon and alloy atoms in the strained quantum well and t_0 are the positions of the atoms in the host crystal.

We can write the Fourier Transforms of the potential $v(\mathbf{g})$ as

$$v(\mathbf{g}) = \int v(\mathbf{r}) e^{-i\mathbf{g}.\mathbf{r}} d\tau \tag{3.16}$$

Substituting this form for the potentials in equation 3.13

$$\frac{1}{\Omega}\langle \mathbf{k}' + \mathbf{G}'|V|\mathbf{k} + \mathbf{G}\rangle = \frac{1}{\Omega_{QW}}[v_s(\mathbf{g})\sum_{t_s} e^{i\mathbf{g}.\mathbf{t_s}} + v_a(\mathbf{g})\sum_{t_a} e^{i\mathbf{g}.\mathbf{t_a}} - v_0(\mathbf{g})\sum_{t_0} e^{i\mathbf{g}.\mathbf{t_0}}] \tag{3.17}$$

where we have used

$$\frac{1}{\Omega}\sum_j e^{i\mathbf{g}.\mathbf{R_j}} = \frac{N}{\Omega} = \frac{1}{\Omega_{QW}} \tag{3.18}$$

where Ω_{QW} is the volume of the quantum well unit cell and N is equal to the number of quantum well unit cells in the crystal. Thus we have an expression for the matrix of the Hamiltonian obtained by substituting the above form for the matrix elements into equation (3.12)

$$\begin{aligned} &A_{n'\mathbf{k}'}(E_{n'\mathbf{k}'} - E) + \sum_n \sum_{\mathbf{k}} \sum_{\mathbf{G},\mathbf{G}'} A_{n\mathbf{k}} a_{n'\mathbf{k}'}(\mathbf{G}) a_{n\mathbf{k}} \\ &\times \frac{1}{\Omega_{QW}}[v_s(\mathbf{g})\sum_{t_s} e^{i\mathbf{g}.\mathbf{t_s}} + v_a(\mathbf{g})\sum_{t_a} e^{i\mathbf{g}.\mathbf{t_a}} - v_0(\mathbf{g})\sum_{t_0} e^{i\mathbf{g}.\mathbf{t_0}}] = 0. \quad (3.19) \end{aligned}$$

Direct diagonalization of the above equation will yield the eigenfunctions and eigenvalues of the system for any quantum well system.

3.2.1 Optical Matrix Elements

Given the quantum well eigenfunctions and eigenstates, we need to find the transition probability between the electron states, i.e. the optical or momentum matrix elements, $M_{i,f}$ which will be the input to our susceptibility calculations in later sections. In the presence of an electromagnetic field the kinetic energy of an electron, using the one-electron approximation, is expressed as (Bassini 1975)

$$\frac{1}{2m}\left(\mathbf{p}+\frac{e\mathbf{A}(\mathbf{r},t)}{c}\right)^2 \tag{3.20}$$

where e is the electronic charge, $\mathbf{A}$ is the vector potential associated with the field and $\mathbf{p}$ is the momentum operator. The scalar potential V can be taken to be zero without any loss of generality because of the arbitaryness in the gauge. Expanding this expression and neglecting all terms containing $\mathbf{A}^2$ we can write the interaction Hamiltonian of an electron in a radiation field as

$$H_e = \frac{e}{mc}\mathbf{A}(\mathbf{r},t).\mathbf{p} \tag{3.21}$$

For radiation of a given frequency the vector potential can be written as

$$\mathbf{A}(\mathbf{r},t) = A_0\epsilon e^{i\eta.\mathbf{r}-\omega t} + A_0\epsilon e^{-i\eta.\mathbf{r}-\omega t} \tag{3.22}$$

where ϵ is the polarization vector, η is the wavevector of the photon and ω is the frequency. The first term of equation (3.22) describes the absorption of a photon, the complex conjugate describes emission.

The probability per unit time of the transition of an electron from an initial state $|i\rangle$ to a final state $|f\rangle$ under the influence of the radiation field is given

by first order time dependent perturbation theory.

$$P_{if} = \frac{2\pi}{\hbar}\left(\frac{eA_0}{mc}\right)^2 |\langle f|e^{i\eta.\mathbf{r}}|\epsilon.\mathbf{p}|i\rangle|^2 \delta(E_i - E_f - \hbar\omega) \tag{3.23}$$

where E_i and E_f are the initial and final energies of the states involved in the transition and $\hbar\omega$ is the energy of the photon. Group theory can be applied to show that the probability of a transition is zero unless

$$\mathbf{k_f} = \mathbf{k_i} + \eta \tag{3.24}$$

where $\mathbf{k_i}$ and $\mathbf{k_f}$ are the initial and final wavevectors in the first Brillouin zone. In the dipole approximation we assume that the momentum of the photon is negligible compared to that of the electrons, therefore $\mathbf{k_i} = \mathbf{k_f} = \mathbf{k}$.

$$P_{if} = \frac{2\pi}{\hbar}\left(\frac{eA_0}{mc}|\hat{\epsilon}M_{if}(\mathbf{k})|^2 \delta(E_i(\mathbf{k}) - E_f(\mathbf{k}) - \hbar\nu)\right) \tag{3.25}$$

where A_0 is the magnitude of the vector potential of the applied electromagnetic field, $\hat{\epsilon}$ is the unit polarization vector and $M_{if}(\mathbf{k})$ is defined as

$$\hat{\epsilon}.M_{if}(\mathbf{k}) = \langle f|\hat{\epsilon}.\mathbf{p}|i\rangle \tag{3.26}$$

Substituting for the forms of the initial and final states from equations (3.8) and (3.10) we get

$$\hat{\epsilon}.M_{if}(\mathbf{k}) = \sum_{\mathbf{k}}\sum_{\mathbf{n},\mathbf{n'}}\sum_{\mathbf{G}}\sum_{\mathbf{s}} \mathbf{A^i_{nk,s}} \mathbf{A^{f*}_{n'k',s'}} \mathbf{a_{nk,s}}(\mathbf{G}) \mathbf{a_{n'k',s'}}(\mathbf{G}) \hat{\epsilon}.(\mathbf{k}+\mathbf{G}) \tag{3.27}$$

where there is a double summation over bulk bands, n, n', a sum over plane waves, $\mathbf{G}$, spin s and a sum over bulk k-points folded onto the quantum well wavevector $\mathbf{k}$.

3.3 Strain

The lattice constant of germanium is 4.17% larger than that of silicon. This means that any structure that involves directly growing Ge or SiGe alloy on

silicon is bound to involve straining i.e. the bonds in the layers must be compressed or stretched in order to accommodate the difference in bulk lattice constant. For the quantum well to be part of a useful electronic device its structure must be free of imperfections which would alter the electronic and optical properties of the quantum well. In this thesis we assume that all the silicon barriers of our quantum well structures are unstrained and all the strain is confined to the alloy. This produces a situation as shown in figure 3.7.

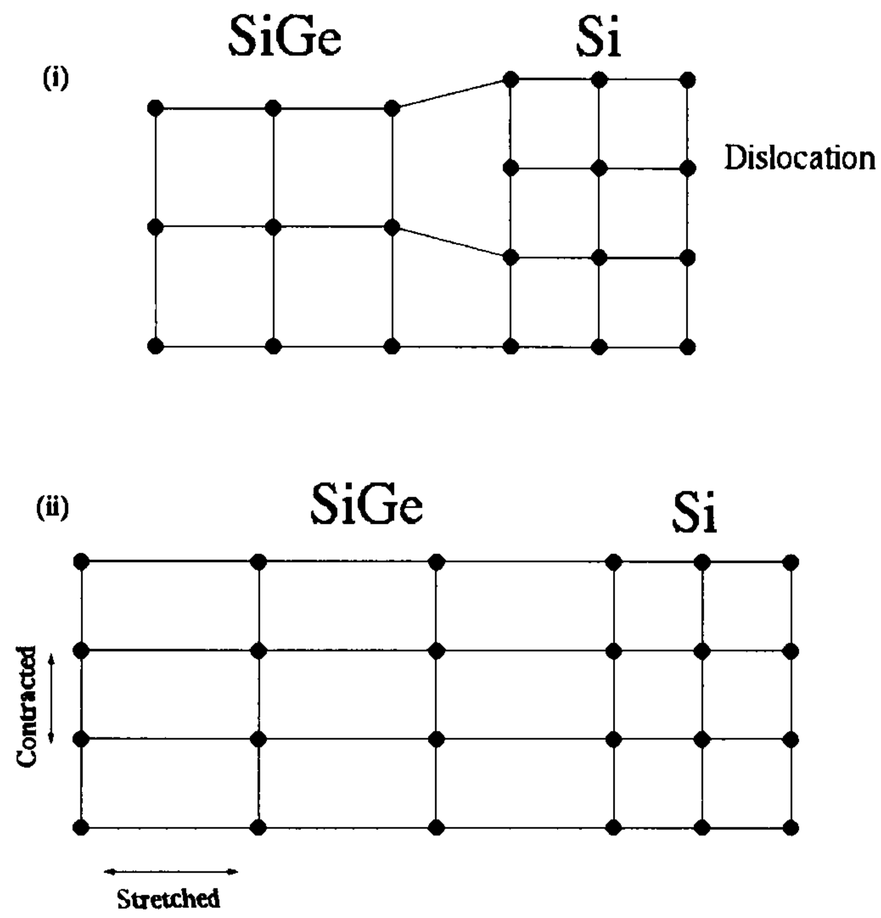


Figure 3.7: *The alteration of atomic positions due to strain. The Si barriers are unstrained so the perpendicular lattice constant of the alloy adjusts so that no dislocations are formed*

Modern day MBE techniques are performed at relatively low temperatures, typically 580^{0}C (Bean 1984) which means that the growth can proceed commensurately i.e. the perpendicular lattice constants adjust independently resulting in a tetragonal distortion of the cubic lattice. The elastic energy for the two materials is minimized and the energy of the interface atoms is lowered. This energy is stored as strain energy within the strained layers. This

stored energy must remain lower than a certain limit or dislocations will form as a more energetically favourable set up. The number of monolayers that can be grown before dislocations form is known as the critical thickness which is dependent on alloy concentration, the width of the well and the number of periods grown. After a generation of misfit dislocations occurs at the interface the lattice resumes growth with the natural lattice constant as determined by Vergard's law. A relatively high number of dislocations within the lattice makes it unsuitable for device applications. The critical thickness, however, is calculated assuming equilibrium conditions which certainly do not exist during the growth process, thus the theoretical calculations and the experimental findings for the critical thickness are not in good agreement. For a $Si_{0.5}Ge_{0.5}$ alloy grown on a silicon barrier the predictions for the critical thickness range from 150 Å (Dodson 1986) to 1000 Å (Chern 1991). Thus all the structures considered in this thesis are well below the critical thickness and we shall assume no dislocations form.

The major effect of strain is to alter the positions of the atoms. The coordinates of the atoms are obviously an important input to the quantum well calculation, we must therefore find the new strained atomic positions. We assume all interfaces to be ideal, i.e. the bulk structure of the material continues up to the interface and then changes abruptly to the other constituent. This is obviously not the situation in the real quantum well, however, extensive studies have been done on the effects of imperfections at the interface of SiGe hetrojunctions (Van de Walle 1986) which show that a macroscopic approach to the interface yields good results.

We assume that the lattice constant parallel to the interface plane is constant throughout the quantum well structure, and is equal to the bulk lattice constant of silicon.

We consider all our quantum well structures to be grown on a cubic Si (001) substrate. We shall first discuss the case of pure Ge grown on cubic Si. The values for various alloys can then be obtained by linear interpolation using the virtual crystal approximation. As we already know our parallel lattice constant, $a_{||}$, to be that of bulk Si, we now need to find the perpendicular lattice constant, $a_{\perp}$, for the alloy. The Si barriers are assumed to maintain their cubic structure, therefore $a_{||}^{barrier} = a_{\perp}^{barrier}$. From Van de Walle (1986) we get an expression for $a_{\perp}$ for the wells.

$$a_{i\perp} = a_i[1 - D^i(\frac{a_{||}}{a_i} - 1)] \tag{3.28}$$

where i denotes the constituent material, a_i is the equilibrium, unstrained lattice constant for material i and the constant D is a function of the elastic constants c_{11} and c_{12}. For a 001 substrate this is given by

$$D_{001} = 2(\frac{c_{12}}{c_{11}}) \tag{3.29}$$

Values of c_{11} and c_{12} are obtained from experiment. Values in the literature can vary by up to 10%. However, as only the ratio of the two constants is used this is unlikely to be a major source of error. Again values for the alloys are obtained by linear interpolation. We also need to know the strain components which depend on the ratio of the strained to unstrained lattice constants for each material, i.e.

$$\epsilon_{xx} = \epsilon_{yy} = (\frac{a_{||}}{a_i} - 1) \tag{3.30}$$

$$\epsilon_{zz} = (\frac{a_{i\perp}}{a_i} - 1)$$

We also need the lattice constants of the host material which for this thesis is always taken to be strained silicon. The parallel lattice constant for the host

Const	Si	Ge	15% Ge	30% Ge	50% Ge
a_i	5.4307	5.6575	5.4647	5.4987	5.5441
$a_{\|\|}$	5.4307	5.4307	5.4307	5.4307	5.4307
c_{11}	1.59	1.30	1.55	.150	1.45
c_{12}	0.61	0.45	0.59	0.56	0.53
$a_{i\perp}$	5.4307	5.8145	5.4906	5.5495	5.6270
$\epsilon_{xx} = \epsilon_{yy}$ (Mbar)	0	-0.04	-0.0062	-0.012	-0.02
ϵ_{zz} (Mbar)	0	0.028	0.0047	0.0092	0.015

Table 3.4: *Values of lattice constants, elastic constants and strain components for Si, Ge and alloys*

is again that of cubic silicon. The perpendicular lattice constant is calculated by

$$a_{\perp}^{host} = \frac{t_{Si} + t_{Ge}}{total \quad no. \quad of \quad atoms} \tag{3.31}$$

This ensures that the volume of the quantum well unit cell is the same as that of the host material containing a similar number of atoms.

3.3.1 Alloy Potentials

To obtain the potentials for the alloys we linearly interpolate between the coefficients for Si and Ge. This procedure does not reproduce the band offset however, so the first coefficient, a_1 is adjusted so that the top of the strained alloy bulk band structure is correctly positioned with respect to the top of the bulk silicon valence band. The band offsets are obtained by linear interpolation of the results of Van de Walle (1983). The coefficients for the alloys used in this thesis are given in table (3.5) The alloy potentials produced by this procedure are presented in figure 3.8.

Material	a_1	a_2	a_3	a_4	a_5	a_6
15%	97.2260	2.2475	0.6261	-1.7332	5.0	0.3
30%	88.5538	2.2672	0.6462	-1.4944	5.0	0.3
50%	77.4869	2.2935	0.6730	-1.1760	5.0	0.3

Table 3.5: *Coefficients for the potentials for alloys interpolated from the coefficients of Silicon and Germanium from Friedel et al. All values are in atomic units*

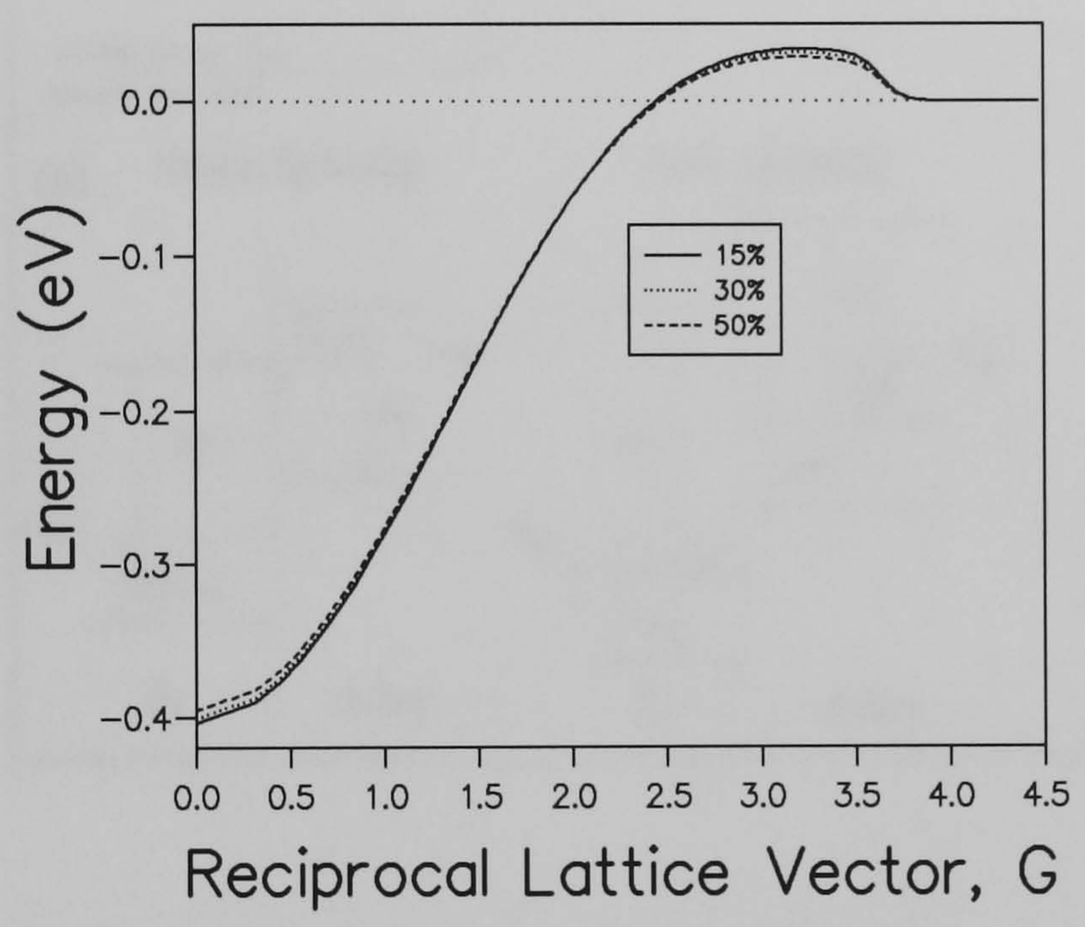


Figure 3.8: *Plot of the alloy potentials produced by the coefficients given in table 3.5.*

3.3.2 The Effects of Strain and Spin on the Band structure

The most noticeable effect of strain on the band structure is the lifting of the degeneracy at the top of the valence band. This degeneracy is also affected by the spin-orbit coupling. Our pseudopotential calculation includes the effects of spin and strain when the appropriate atomic coordinates are used and thus

reproduce the splitting at the top of the valence band explicitly. However, we can only obtain the relative energies of the electron states. In order to relate these to the band edges, i.e. the top and bottom of the quantum well, it is useful to be able to assess the strain and spin splittings qualitatively. We use deformation potentials to describe the splitting at the top of the valence band.

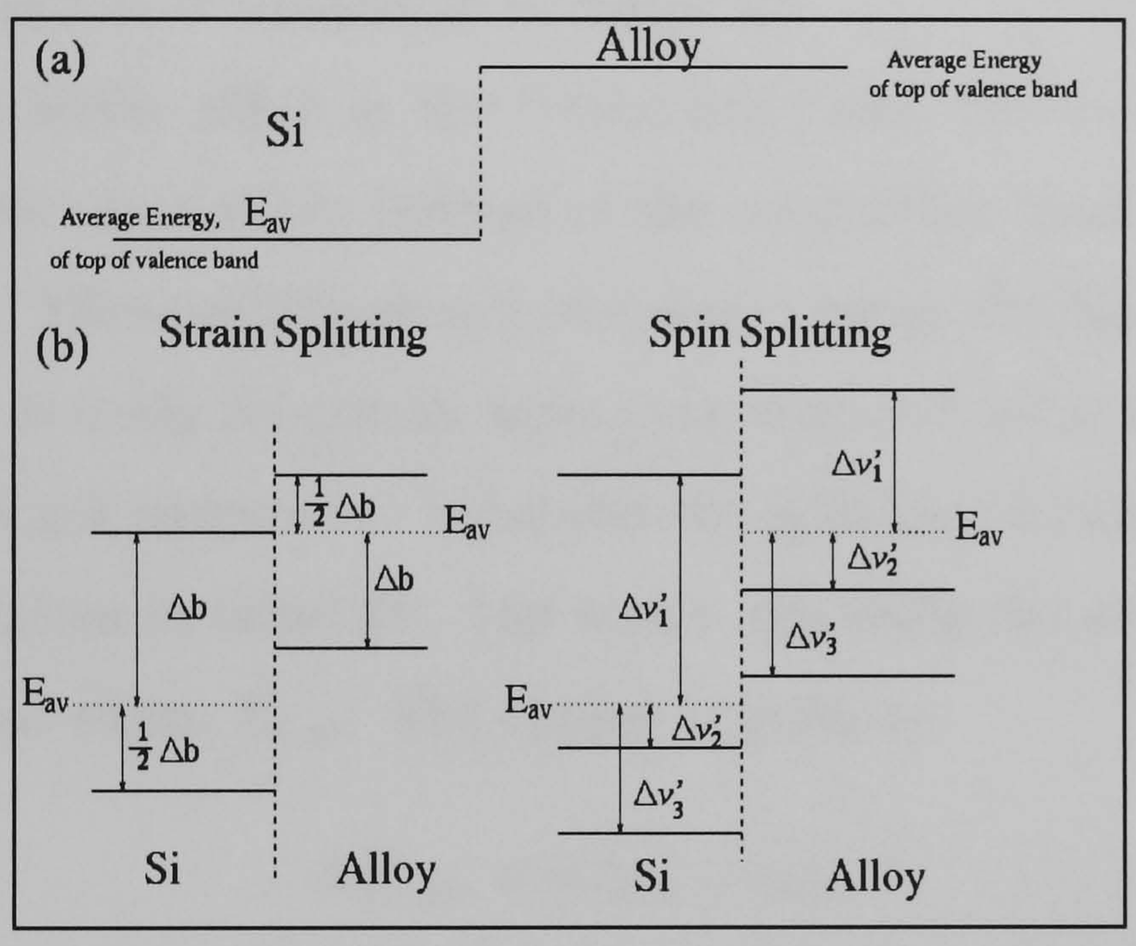


Figure 3.9: *Schematic representation of the effect of spin and strain on the band structure of a quantum well system.*

The strain splits the 6-fold degenerate state at Γ into a doublet state of energy

$$\Delta v_1 = \Delta_b = 2b(\epsilon_{zz} - \epsilon_{xx}) \tag{3.32}$$

and a four-fold multiplet with energy of

$$\Delta v_2 = \Delta v_3 = -\frac{1}{2}\Delta_b = -b(\epsilon_{zz} - \epsilon_{xx}) \tag{3.33}$$

The spin-orbit coupling further splits this four fold state into two doublets,

with the splittings now given by

$$\begin{aligned}
\Delta v_1' &= -\frac{1}{6}\Delta_0 + \frac{1}{4}\Delta_b + \frac{1}{2}[\Delta_0^2 + \Delta_0\Delta_b + \frac{9}{4}\Delta_b^2]^{\frac{1}{2}} \\
\Delta v_2' &= \frac{1}{3}\Delta_0 - \frac{1}{2}\Delta_b \\
\Delta v_3' &= -\frac{1}{6}\Delta_0 + \frac{1}{4}\Delta_b - \frac{1}{2}[\Delta_0^2 + \Delta_0\Delta_b + \frac{9}{4}\Delta_b^2]^{\frac{1}{2}} \qquad (3.34)
\end{aligned}$$

This is demonstrated schematically in figure 3.9.

There is a similar effect in the conduction band where strain splits the 6-fold degenerate state at the bottom of the conduction band into a two and four fold state. These splittings will obviously change the fundamental band gap but as we are only concerned with inter-valence band transitions in this thesis this need not concern us. Values for the splittings for the alloys used in this thesis are given in table 3.6. The strain also shifts the average energy of the valence band edges, $E_{v,av}$. The change is given by

$$\Delta E_{v,av} = a(2\epsilon_{xx} + \epsilon_{zz}) \qquad (3.35)$$

We can compare these calculated values with the splittings produced by our bulk strained alloy calculations shown in table (3.7). The results for Ge are for germanium strained to a silicon buffer. The silicon is unstrained hence the degeneracy of the LH1 and HH1 minibands. The agreement between our calculated results and the ab initio results of Van de Walle is excellent for the HH1-LH1 splittings, agreeing to within a couple of meV. The results for the HH1-SO1 splittings are less close, with our pseudopotential method consistently producing larger splittings. However,the difference between the two sets of results is not more than 10% for any case and considerably less for the lower Ge concentrations.

Const	Si	Ge	15% Ge	30% Ge	50% Ge
b	-2.35	-2.55	-2.38	-2.41	-2.45
$\Delta v_1 = \Delta_b$	0	-350	-52	-102	-170
$\Delta v_2, \Delta v_3$	0	180	26	51	85
Δ_0	40	300	80	120	170
$\Delta v_1'$	13	118	190	340	560
$\Delta v_2'$	13	118	19	34	56
$\Delta v_3'$	27	-393	-72	-125	-198

Table 3.6: *Energy level splittings calculated using equations* 3.34 *All values are in meV.*

Alloy	HH1-LH1 splitting		HH1-SO1 splitting	
Concentration	P.P.	V.D.W.	P.P.	V.D.W
Si	0	0	45	40
Ge	157	157	747	667
15%	35	33	130	124
30%	57	57	230	219
50%	85	86	369	346

Table 3.7: *Comparison of the results of Van de Walle (V.D.W.) with our pseudopotential calculations (P.P.) HH1-LH1 and HH1-SO1 splittings. All values are in meV.*

3.4 Optical Selection Rules

At the zone centre all the minibands are decoupled, i.e. there is no mixing between them. Thus the symmetry properties of the energy minibands near Γ are essentially the same as those in the bulk material (Man 1992). Ma et al. (1992) analyze the band structure and symmetry properties of coherently grown strained $Si_{1-x}Ge_x$ alloys on substrates of various orientations and use group theory arguments to determine the selection rules. All structures in this thesis are considered to be grown in (001) substrates. The strain in the alloy lifts the symmetry of the diamond structure, therefore the space groups and the irreducible representations which denote the new energy bands will be different from those of an unstrained diamond structure. The energy bands are either shifted with no change of irreducible representation or are split and require new irreducible representations to describe them.

As we know the space group of the distorted lattice, the irreducible representations of the space group at any point $\mathbf{k}$ of the distorted Brillouin Zone can be obtained. As we include spin orbit coupling in our calculations, the double space groups must be used. The band structure of strained Ge grown on [001] Si along the symmetry lines of the distorted B.Z. calculated by Ma et al. is reproduced in figure 3.10.

At the Γ point the top valence band state is the heavy hole state, which is represented by the irreducible representation Γ_6^+ with $|J, m_J\rangle = |\frac{3}{2}, \pm\frac{3}{2}\rangle$. The light hole state is represented by the irreducible representation Γ_7^+ with $|\frac{3}{2}, \pm\frac{1}{2}\rangle$. Both these bands originate from the 4-fold degenerate valence band I_8^+ of the unstrained structure. The LH band interacts with the SO band as both have the same $m_\mathbf{J}$. The HH band has a different $m_\mathbf{J}$ and thus to first order does not interact with either other state. For this substrate orientation the symmetry along the [001] direction is not reduced. Therefore the group of $\mathbf{k}$ along the

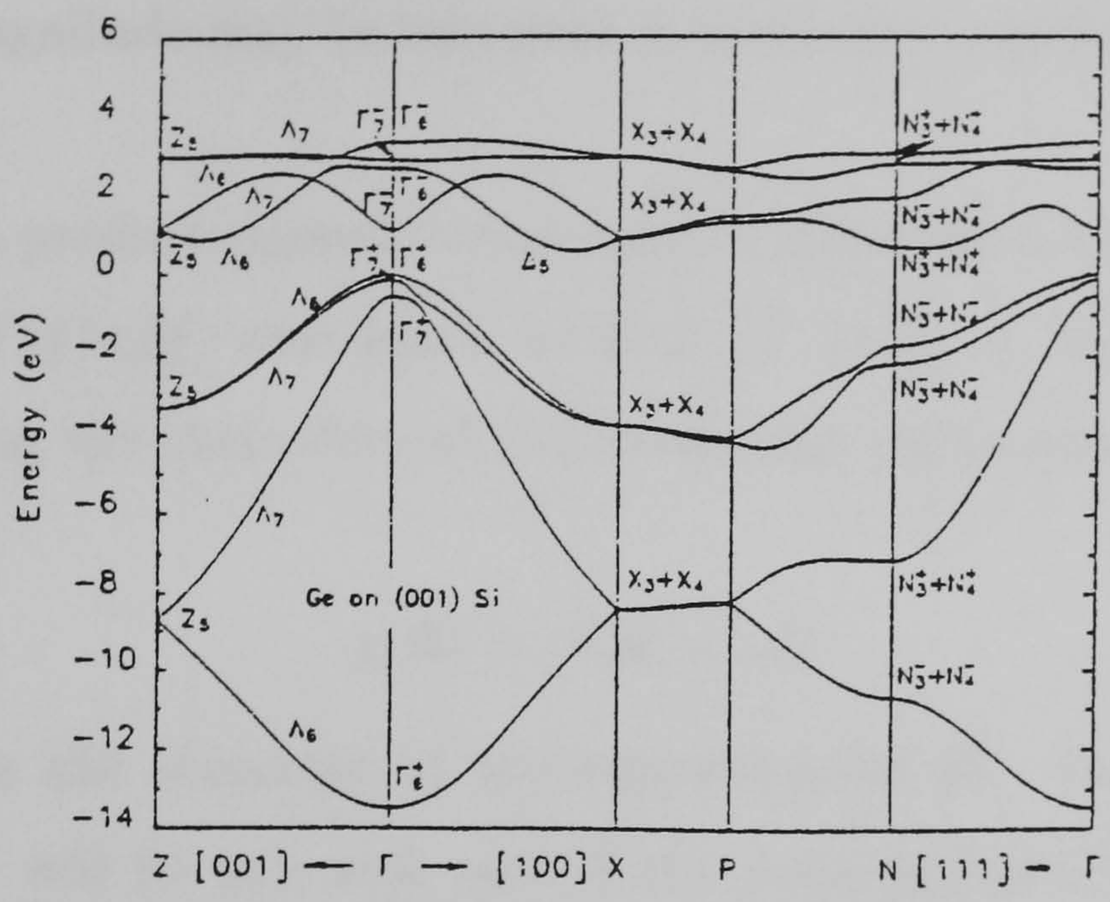


Figure 3.10: *The band structure of coherently strained Ge grown on a (001) Si substrate along different directions of the distorted Brillouin zone. (Reproduced from the work of Ma, Wang and Schulman.)*

[001] direction is unchanged and the relevant irreducible representations for the bands in this direction are $\wedge_6$ and $\wedge_7$. The symmetry in the [100] direction is changed by strain and all bands are now represented by the Δ_5 irreducible representation.

3.4.1 Direct Products of Irreducible Representations

For optical transitions one needs to calculate matrix elements of the form

$$\langle \psi_i | V | \psi_f \rangle \tag{3.36}$$

where V is the perturbation operator, which for the case of photons in the dipole approximation is given as **e**.**r** where **e** is the polarization vector of the radiation field.. If this product is zero then the transition is symmetry forbidden. Group theory cannot predict the magnitude of any allowed transitions,

indeed the magnitude may be zero even if symmetry considerations predict it to be allowed.

The direct product representation used in group theory then tells us that the product $V(\mathbf{r})\psi_i(\mathbf{r})$ transforms as $\Gamma^i \otimes \Gamma^V$ where $\otimes$ indicates the direct product. Using the characters of the irreducible representations this can be expressed as

$$\chi(R) = \chi^i(R)\chi^V(R). \tag{3.37}$$

where $\chi(R)$ is the character of the representation R. The direct product $\Gamma = \Gamma^i \otimes \Gamma^V$ will be non zero only if the reduction of Γ contains Γ^f, the irreducible representation of the final state. The number of times that Γ^f appears in the reduction is given by for the point group R.

$$n_{iV}^f = h^{-1} \sum_R \chi^i(R)\chi^V(R)\chi^f(R)^* \tag{3.38}$$

where h is the order of the group. For non-symmorphic space groups this is extended to

$$n_{iV}^f = (h')^{-1} \sum_R \chi^{\mathbf{i}}(\{\mathbf{R}|\mathbf{t_R}\})\chi^{\mathbf{V}}(\{\mathbf{R}|\mathbf{t_R}\})\chi^{\mathbf{f}}(\{\mathbf{R}|\mathbf{t_R}\})^* \tag{3.39}$$

where the sum is over all elements $\{\mathbf{R}|\mathbf{t_R}\}$ of $\mathcal{G}$, only one transformation of $\{\mathbf{R}|\mathbf{t_R}\}$ being included for each rotation $\mathbf{R}$ of $\mathcal{G}$.

Thus in order to predict which transitions are allowed we need only know which irreducible representations of the point groups represent the initial and final states and the perturbing operator. Fortunately, as in much of group theory, the work has already been done and we need only apply the conclusions to our case. In table 3.8 we show the irreducible representations that are contained in the direct product of irreducible representations at the Γ point.

Thus from table 3.8 we can see that transitions between the valence bands are not allowed at the zone centre. For example, the heavy hole band transforms as Γ_6^+ and the light hole band as Γ_7^+. Transitions between these bands

Γ point	Γ_i	Γ_6^+	Γ_7^+	Γ_6^-	Γ_7^-
Γ_2^- ; z	$\Gamma_i \otimes \Gamma_2^-$	Γ_6^-	Γ_7^-	Γ_6^+	Γ_7^+
Γ_5^-; (x,y)	$\Gamma_i \otimes \Gamma_5^-$	$\Gamma_6^- \oplus \Gamma_7^-$	$\Gamma_6^- \oplus \Gamma_7^-$	$\Gamma_6^+ \oplus \Gamma_7^+$	$\Gamma_6^+ \oplus \Gamma_7^+$

Table 3.8: *Selection rules for direct optical transitions at the zone centre. The first column gives the dipole irreducible representations and the basis functions.*

$\wedge$ line	$\wedge_i$	$\wedge_6$	$\wedge_7$
$\wedge_1$;z	$\wedge_i \otimes \wedge_1$	$\wedge_6$	$\wedge_7$
$\wedge_5$; (x,y)	$\wedge_i \otimes \wedge_5$	$\wedge_6 \otimes \wedge_7$	$\wedge_6 \otimes \wedge_7$

Table 3.9: *Selection rules for direct optical transitions along the symmetry line $\wedge$, as shown in figure 3.10. The first column gives the dipole irreducible representations and the basis functions.*

for z-polarized light gives us the expression

$$\Gamma_6^+ \otimes \Gamma_2^- \otimes \Gamma_7^+ \tag{3.40}$$

This expression is non-zero only if the product representation of $\Gamma_2^- \otimes \Gamma_7^+$ contains the irreducible representation Γ_6^+. From table 3.8 we can see that

$$\Gamma_2^- \otimes \Gamma_7^+ = \Gamma_7^- \tag{3.41}$$

i.e. does not contain the irreducible representation Γ_6^+, therefore this transition is forbidden at the zone centre.

Along the [001] axis the heavy hole band is denoted by the irreducible representations $\wedge_6$ and the light hole and split off bands as $\wedge_7$. The selection rules for theses states are given in table 3.9. Away from the zone centre a transition between heavy hole and light hole bands is allowed for x or y polarized light but not for z as shown by the fact that for z-polarized light the

appropriate irreducible representation is $\wedge_1$, and

$$\wedge_1 \otimes \wedge_7 = \wedge_7 \tag{3.42}$$

which does not contain the required final irreducible representation of the heavy hole band, namely $\wedge_6$. For normal incident light the irreducible representation for the perturbation is $\wedge_5$. Hence for a transition where the initial state is the HH band ($\wedge_6$) and the final state the split off band ($\wedge_7$) we get

$$\wedge_5 \otimes \wedge_7 = \wedge_6 \oplus \wedge_7 \tag{3.43}$$

which does contain the irreducible representation for the final HH state, $\wedge_6$. Thus the selection rules depend on the polarization of the light. We can conclude that transitions between the valence bands of a bulk, strained $Si_{1-x}Ge_x$ alloy grown on a (001) Si substrate are forbidden at the zone centre. However, away from the zone centre where the symmetry properties of the bands are changed, transitions between valence bands are still forbidden for parallel incident light but are allowed for normal incident light. Remember that group theory cannot help us to predict the magnitude of these transitions.

3.4.2 Selection Rules in the Quantum Well

Up until now we have been discussing selection rules in the bulk material. We know that this is a good approximation to the selection rules that we would expect in the quantum well system, for transitions that occur around the zone centre, which at reasonable doping levels and temperatures accounts for the majority of all transitions.

Our quantum well structures possess C_{2v} symmetry. For the C_{2v} point group the photon vector $\epsilon.\mathbf{p}$ transforms as Δ_1 for z-polarized i.e. parallel incident light, whilst $\epsilon.\mathbf{p}$ belongs to Δ_3 or Δ_4 for x and y polarized light

respectively. Δ_3 and Δ_4 are actually equivalent from time reversal symmetry, so the distinction between them is superficial for selection rules for structures with time reversal symmetry. However, the fact that parallel incident and normal incident light transform as different irreducible representations of the point group means that the selection rules will be polarization dependent. This is not the case for the cubic point group of the bulk crystal O_h where $\epsilon.\mathbf{p}$ transforms as Γ_{15} for all optical polarizations (Bassani 1975).

We can also derive additional selection rules for the quantum well system using the envelope function approximation (Bastard 1988). This approximation does not allow for any mixing between the minibands, and thus is valid only at the zone centre.

Wavefunctions of the initial and final states involved in the optical transition can be written in the form

$$\psi_i(\mathbf{r}) = u_i(\mathbf{r}) f_i(\mathbf{r}) \tag{3.44}$$

where $u_i(\mathbf{r})$ is the periodic part of the Bloch function at the zone centre, (assumed to be the same in both types of layers).

$$f_i(\mathbf{r}) = \frac{1}{\sqrt{S}} e^{i\mathbf{k}_\perp.\mathbf{r}_\perp} F_i(z) \tag{3.45}$$

where $F_i(z)$ is the envelope function which describes the z-motion of the electron in the miniband i. $\mathbf{k}$ and $\mathbf{r}$ are the 2-D wave and position vectors respectively with $\mathbf{k_{\perp i}} = (k_x, k_y)$ and $\mathbf{r}_{\perp i} = (x, y)$. Using the envelope function approximation (Bastard 1989) we get an expression for the optical matrix element

$$\begin{aligned} \epsilon.\mathbf{p_{if}} &= \epsilon.\int_\Omega \psi_i^*(\mathbf{r}) \mathbf{p} \psi_f(\mathbf{r}) d^3r \\ &= \epsilon.\langle u_i|\mathbf{p}|u_f\rangle \int_\Omega f_i^* f_f d^3r + \delta_{if}\epsilon. \int_\Omega f_i^* \mathbf{p} f_f d^3r \end{aligned} \tag{3.46}$$

where

$$\langle u_i|\mathbf{p}|u_f\rangle = \int_{\Omega_0} u_i^*(\mathbf{r})\mathbf{p}u_f(\mathbf{r})d^3r \tag{3.47}$$

This term can be removed from the integral as u_i and u_f vary rapidly over the characteristic lengths of variation of the envelope functions. Thus we can write

$$\langle f|\epsilon.\mathbf{p}|i\rangle = (\epsilon_x\hbar k_x + \epsilon_y\hbar k_y)\delta_{n_i n_f}\delta_{\mathbf{k}',\mathbf{k}} + \epsilon_z\delta_{\mathbf{k}',\mathbf{k}}\int dz F_{n_i}^*(z)p_z F_{n_f}(z) \tag{3.48}$$

The overlap integral between the envelope functions selects the quantum numbers of the initial and final minibands. The atomic-like dipole matrix elements $\langle u_i|\mathbf{p}|u_f\rangle$ give rise to selection rules on the polarization of the light wave.

$$\langle u_i|\epsilon.\mathbf{p}|u_f\rangle = \frac{1}{S}\int d^3r F_{n_i}^*(z)e^{-i(\mathbf{k}\perp\mathbf{r}_\perp)} \tag{3.49}$$
$$\times[\epsilon_x p_x + \epsilon_y p_y + \epsilon_z p_z]\delta_{n_i,n_f}\delta_{\mathbf{k}',\mathbf{k}} + \epsilon_z\delta_{\mathbf{k}',\mathbf{k}}\int dz F_{n_i}(z)p_z F_{n_f}$$

We can see from the above expansion that polarizations ϵ_x and ϵ_y give rise to allowed transitions only if the initial and final states coincide, i.e. $\omega = 0$. Polarization ϵ_z corresponds to an electromagnetic wave propagating in the layer plane with an electric field vector parallel to the growth axis of the quantum well. For this polarization optical transitions between subbands are allowed provided $\mathbf{k}' = \mathbf{k}$, i.e. vertical in $\mathbf{k}$-space. The heterostructure Hamiltonian has a definite parity, thus initial and final subbands must be of opposite parity for a transition to be allowed. As in our structures only the HH1 miniband is populated which has an even envelope wavefunction, transitions are allowed for z-polarized light between HH1 and HH2 minibands. The HH1-HH3 transition is parity forbidden. HH1-HH4 is in theory allowed but the large separation in energy between these minibands would mean any such transition would be weak and outside our energy range of interest.

Chapter 4

Derivation of First Order Susceptibility

We begin by deriving some results from density matrix theory which we shall need in our derivation of the linear susceptibility. We then consider the first-order susceptibility in terms of the current operators and current densities. There are alternative approaches to this derivation (see Butcher 1990) but this is the most appropriate for semiconductors.

4.1 Density Matrix Theory

4.1.1 The Trace Operator

The trace of operator O is defined as

$$Tr\{O\} = \sum_i O_{ii} \tag{4.1}$$

i.e. the sum of the diagonal matrix elements. The trace of an operator is independent of the representation used for the operator.

4.1.2 The Projection Operator

The projection operator , $P(\psi)$, for a state ψ is defined as

$$P(\psi_n)\theta = \psi \int d\xi \psi^* \theta \tag{4.2}$$

where θ is an arbitrary wavefunction of the system and $d\xi$ is the volume element. If ψ is a member of a complete orthonormal set then $P(\psi)$ picks out the contribution to the expansion of θ from the state ψ. The $(ji)th$ element of $P(\psi)$ is given by

$$P(\psi)_{ji} = \int d\xi u_j^* P(\psi) u_i = \int d\xi u_j^* \psi \int d\xi \psi^* u_i = a_j a_i^* \tag{4.3}$$

where a_i and a_j are coefficients defined by $a_i = \int d\xi u_i^* \psi$.

In general, the expectation value of the operator can be expressed as

$$< O >= Tr\{P(\psi)O]\}$$

This equation tells us the mean value of the operator, O, when the system is in the state ψ. If we have a state ψ_n we know the probability of the system being in that state is p_n, then the expectation value of the operator can be weighted accordingly. The expectation value of the whole system is then given by summing over all possible states ψ_n

$$< O >= \sum_n p_n Tr\{P(\psi_n)O\} \tag{4.4}$$

The projection operator provides us with a description of the physical state of the system equivalent to that given by the wavefunction. The advantage of the projection operator is that is can be used to describe the time development of the system as well as its current situation. Thus we need an equation of motion for $P(\psi)$. Consider the time dependent Schroedinger Equation for θ,

an arbitrary wavefunction and differentiate with respect to time.

$$\begin{aligned} i\hbar\frac{d\psi}{dt} &= H\psi \\ i\hbar\frac{d[P(\psi)]}{dt} &= i\hbar(\frac{d\psi}{dt})\int d\xi\psi^*\theta + \psi\int d\xi i\hbar(\frac{d\psi^*}{dt})\theta \\ &= H\psi\int d\xi\psi^*\theta - \psi\int d\xi(H\psi)^*\theta \end{aligned} \tag{4.5}$$

The first term in the final equation above can be identified as $HP(\psi)$. As H is hermitian, i.e. $H = H^*$ the second term simplifies to

$$\psi\int d\xi(H\psi)^*\theta = \psi(\int d\xi\theta^* H\psi)^* = \psi\int d\xi\psi^* H\theta = P(\psi)H\theta \tag{4.6}$$

Equation 4.5 then becomes

$$i\hbar\frac{d(P(\psi))\theta}{dt} = HP(\psi)\theta - P(\psi)H\theta \tag{4.7}$$

Cancelling the arbitrary wavefunction θ leaves us with for the equation of motion of the projection operator

$$i\hbar\frac{d(P(\psi))}{dt} = HP(\psi) - P(\psi)H = [H, P(\psi)] \tag{4.8}$$

4.1.3 The Density Operator

Precise knowledge of the state of a large system is seldom, if ever, available. However, we can use what we know about the system to determine the probability p_n that the system is in the state ψ_n. In general, the expectation value of an operator may be written as

$$\begin{aligned} < O > &= \sum_i\sum_j P(\psi)_{ji}O_{ij} && (4.9) \\ &= \sum_j [P(\psi)O]_{ij} && (4.10) \\ &= Tr\{P(\psi)O\} && (4.11) \end{aligned}$$

This can be rewritten as

$$< O >= Tr\{\rho O\} \tag{4.12}$$

where

$$\rho = \sum_n p_n P(\psi_n). \tag{4.13}$$

ρ is defined as the density operator of the system. Knowledge of this operator allows us to determine the expectation value of any dynamical operator. The density operator can be determined from the statistical properties of the system without a detailed knowledge of the wavefunction at a given time and place.

4.1.4 Equation of motion of the Density Operator

The probabilities, p_n depend on the initial state of the system and do not change with time. Therefore differentiating our expression for the density matrix operator with respect to time we get

$$i\hbar \frac{d\rho}{dt} = \sum_n p_n [H, P(\psi_n)] = [H, \sum_n p_n P(\psi_n)] = [H, \rho] \tag{4.14}$$

where we have used the identity

$$[\sum_i a_i O_i, \sum_j b_j O'_j] = \sum_i \sum_j a_i b_j [O_i, O'_j] \tag{4.15}$$

where O_i and O'_j are arbitrary operators and a_i and a_j are arbitrary constants.

4.2 Derivation of $\chi^{(1)}$

We first consider a general ensemble of charged particles, namely electrons and ions. Our aim is to calculate the response of such a system when an external optical field is applied. We shall assume that the spatial variation of

the applied field within the system can be ignored; this is commonly known as the electric dipole approximation. In order to determine the response to the applied field we need to know the the macroscopic polarization current density induced by the applied field. Thus we need to find the form of the Hamiltonian and the current operator for both the unperturbed (no external field) and perturbed cases.

4.2.1 The Current Operator

4.2.1.1 Thermal Equilibrium

The current operator is defined as the product of the charge and velocity summed over all charged particles in a volume V. The macroscopic polarization current density can therefore be considered to be the expectation value of the current operator normalized to the volume V. For the unperturbed system, i.e. no applied field, the current operator, $\mathbf{\Pi}$ is defined as

$$\mathbf{\Pi} = -\frac{e}{m}\sum_j \mathbf{p}_j + e\sum_k \frac{Z_k}{M_k}\mathbf{p}_k, \tag{4.16}$$

where j labels an electron of charge e, mass m and momentum $\mathbf{p}_j$, and similarly k refers to the ionic equivalent.

In general the unperturbed Hamiltonian of the system can be written:

$$H_0 = H_0^{K.E.} + V_0, \tag{4.17}$$

where $H_0^{K.E.}$ is the kinetic energy term, given by

$$H_0^{K.E.} = \frac{1}{2m}\sum_j \mathbf{p}_j^2 + \sum_k \frac{1}{2M_k}\mathbf{p}_k^2, \tag{4.18}$$

and V_0 represents the potential energy of the system.

4.2.1.2 The Perturbation Hamiltonian

When an electro-magnetic field is applied the system is perturbed. This perturbation can be added to the original Hamiltonian in the form of an interaction Hamiltonian so that the Hamiltonian for the whole perturbed system can be written as

$$H(t) = H_0 + H_1(t)$$

where the time dependence is introduced because of the fluctuation of the applied electro-magnetic field. The interaction Hamiltonian is given by

$$H_1(t) = \sum_i \left\{ \frac{Z_i e}{2m_i} \left(\mathbf{A}(\mathbf{r},t).\mathbf{p}_i + \mathbf{p}_i.\mathbf{A}(\mathbf{r},t)\right) + \frac{(Z_i e)^2}{2m_i} |\mathbf{A}(\mathbf{r},t)|^2 + Z_i e \Phi(\mathbf{r},t) \right\}, \tag{4.19}$$

where i labels individual charged particles (electrons or ion-cores) of mass m_i and charge $-Z_i e$. $\mathbf{A}(\mathbf{r},\mathrm{t})$ is the vector potential of the applied field, and $\Phi(\mathbf{r},t)$ is its scalar potential.

For convenience we chose to work in the Coulomb gauge in which $\boldsymbol{\nabla}.\mathbf{A} = 0$ and $\Phi(\mathbf{r},t) = 0$. In this gauge the vector potential is related to the electric field by

$$\begin{aligned} \mathbf{A}(\mathbf{t}) &= -\int_{-\infty}^{t} d\tau \mathbf{E}(\tau) \\ &= \int_{-\infty}^{\infty} d\omega \mathbf{E}(\omega)(i\omega)^{-1} e^{-i\omega t}, \end{aligned} \tag{4.20}$$

The electric field $\mathbf{E}(\tau)$ can be transformed into the frequency domain using the appropriate Fourier transform i.e.

$$\mathbf{E}(\tau) = \int_{-\infty}^{\infty} d\omega \mathbf{E}(\omega) e^{-i\omega\tau}$$

where

$$\mathbf{E}(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\tau \mathbf{E}(\tau) e^{-i\omega\tau}$$

As we are working in the electric dipole approximation, the vector potential $\mathbf{A}(\mathbf{r},t)$ does not vary with position $\mathbf{r}$, therefore we drop this from our notation. We note that in the Coulomb gauge the vector potential and the momentum operators commute,

$$[\mathbf{A}(t),\mathbf{p}] = \mathbf{A}(t).\mathbf{p} - \mathbf{p}.\mathbf{A}(t) = 0. \tag{4.21}$$

The number of electrons contained within volume V is denoted N, i.e. $\sum_j = N$. Thus the interaction Hamiltonian can be written as

$$\begin{aligned} H_1(t) &= \frac{e}{m}\sum_j \mathbf{A}(\mathbf{r},t).\mathbf{p}_j + \frac{Ne^2}{2m}|\mathbf{A}(\mathbf{r},t)|^2 \\ &\quad -\sum_k \frac{Z_k e}{M_k}\mathbf{A}(\mathbf{r},t).\mathbf{p}_k + \frac{1}{2}\sum_k \frac{(Z_k e)^2}{M_k}|\mathbf{A}(\mathbf{r},t)|^2, \end{aligned} \tag{4.22}$$

where j and k label the electrons and ion-cores respectively. To simplify the above expression we introduce the the notation

$$\beta = \frac{e^2}{m}N + \sum_k \frac{(Z_k e)^2}{M_k}, \tag{4.23}$$

thus

$$H_1(t) = -\mathbf{A}(t).\mathbf{\Pi} + \frac{\beta}{2}\mathbf{A}(t)^2. \tag{4.24}$$

The equation of motion of the density operator, equation (4.14), involves the commutator $[H,\rho]$. $\mathbf{A}(t)^2$ appears in the definition of the interaction Hamiltonian, (4.24) and is independent of position in the electric dipole approximation and thus commutes with ρ_0 . Therefore this term will make no contribution to the relationship in equation (4.14) and we can drop this term from the expression for the interaction Hamiltonian which leaves us with

$$H_1(t) = -\mathbf{A}(t).\mathbf{\Pi}. \tag{4.25}$$

4.2.1.3 The perturbed current operator

We now move on to consider the current operator for the perturbed system, $\mathbf{\Pi}'$.

The total Hamiltonian for the system is

$$\begin{aligned} H(t) &= H_0^{K.E.} + V_0 + H_1 \\ &= \left(\frac{1}{2m}\sum_j \mathbf{p}_j^2 + \sum_k \frac{1}{2M_k}\mathbf{p}_k^2\right) + V_0 + \left(\frac{e}{m}\sum_j \mathbf{A}(t).\mathbf{p}_j + \frac{Ne^2}{2m}|\mathbf{A}(t)|^2\right. \\ &\quad \left. - \sum_k \frac{Z_k e}{M_k}\mathbf{A}(t).\mathbf{p}_k + \frac{1}{2}\sum_k \frac{(Z_k e)^2}{M_k}|\mathbf{A}(t)|^2\right), \end{aligned} \tag{4.26}$$

where (4.17), (4.18), and (4.22) have been incorporated. After rearrangement this can be expressed as

$$H(t) = \frac{1}{2m}\sum_j \left[\mathbf{p}_j + e\mathbf{A}(t)\right]^2 + \sum_k \frac{1}{2M_k}\left[\mathbf{p}_k - Z_k e\mathbf{A}(t)\right]^2 + V_0. \tag{4.27}$$

From inspection of the above equation we can deduce that the velocities of the electrons and ions are given by $[\mathbf{p}_j + e\mathbf{A}(t)]$ and $[\mathbf{p}_k - e\mathbf{A}(t)]$ respectively. The equilibrium current operator was defined as the sum of the charge $\times$ velocity for both types of particles. Thus by analogy we can define the perturbed current operator $\mathbf{\Pi}'$ to be

$$\begin{aligned} \mathbf{\Pi}' &= -\frac{e}{m}\sum_j \mathbf{p}_j + e\sum_k \frac{Z_k \mathbf{p}_k}{M_k} - \sum_j \frac{e^2\mathbf{A}(t)}{m} - \sum_k \frac{Z_k e^2 \mathbf{A}(t)}{M_k} \\ \mathbf{\Pi}' &= \mathbf{\Pi} - \beta\mathbf{A}(t), \end{aligned} \tag{4.28}$$

where $\mathbf{\Pi}$ is the unperturbed current operator and β is defined by (4.23).

4.2.2 The Macroscopic Polarization Current Density

The macroscopic polarization current density, $\mathbf{J}$(t), is defined as the expectation value of the current operator per unit volume i.e.

$$\mathbf{J}(t) = V^{-1} < \mathbf{\Pi}' >$$

The volume V must be chosen to be small enough such that our assumption that the field does not vary spatially within the volume still applies, but must be sufficiently large such that we can ignore the fluctuations in the charged particle density. We need to obtain the expectation value in terms of the current density operator. Thus using the density operator approach we can obtain an expression for the expectation value of the current density $\mathbf{J}(t)$

$$\mathbf{J}(t) = V^{-1}Tr\{\rho(t)\mathbf{\Pi}'\}. \tag{4.29}$$

We can now substitute our expression for $\mathbf{\Pi}'$ into this equation giving

$$\mathbf{J}(t) = V^{-1}Tr\{\rho(t)\mathbf{\Pi} - \beta\rho(t)\mathbf{A(t)}\}. \tag{4.30}$$

$\beta\mathbf{A(t)}$ is independent of position $\mathbf{r}$ therefore is unchanged by the operation of $\rho(t)$. The trace of $\rho(t)$ is equal to 1 as can be easily demonstrated. Consider the operator $O = 1$. The expectation value of this operator must be 1. Therefore using equation (4.12) we have

$$< O >= Tr\{\rho(t)O\} = 1 \tag{4.31}$$

Therefore $Tr\{\rho(t)\}$ must equal 1 for any $\rho(t)$. Thus

$$Tr\{\beta\rho(t)\mathbf{A}(t)\} = Tr\{\beta\mathbf{A}(t)\rho(t)\} = \beta\mathbf{A}(t)Tr\{\rho(t)\} = \beta\mathbf{A}(t) \tag{4.32}$$

Equation 4.30 then simplifies to

$$\mathbf{J}(t) = V^{-1}Tr\{\rho(t)\mathbf{\Pi}\} - V^{-1}\beta\mathbf{A(t)}. \tag{4.33}$$

A perturbation series for ρ(t) can be expressed as

$$\rho(t) = \rho_0 + \rho_1(t) + \rho_2(t) + \rho_3(t) + \ldots \rho_n(t) \tag{4.34}$$

where ρ_1 is a linear function of the perturbation Hamiltonian, H_1, ρ_2 is quadratic etc. We can similarly express the polarization current density as a power series,

$$\mathbf{J}(t) = \mathbf{J}^{(1)}(t) + \mathbf{J}^{(2)}(t) + \ldots + \mathbf{J}^{(n)}(t) + \ldots \tag{4.35}$$

We can then equate powers of $\mathbf{A}(t)$. As we are only interested in the first order susceptibility we keep only the first order terms. For a discussion of a general nth order derivation consult Shaw (1994). Thus equating first order terms gives us

$$\mathbf{J}^{(1)}(t) = -V^{-1}\beta\mathbf{A}(t) + V^{-1}Tr\{\rho_1(t)\mathbf{\Pi}\} \tag{4.36}$$

4.2.3 The Interaction Picture

We introduce the concept of the interaction picture. We define an operator in the interaction picture O' such that

$$O'(t) = U_0(-t)O(t)U_0(-t) \tag{4.37}$$

where O is the original operator and U_0 is the time development operator and is defined as

$$U_0 = e^{-iH_0t/\hbar} \tag{4.38}$$

If an operator commutes with U_0 or is independent of t then $O'(t) = O(t)$. Substituting the equation of motion of the density operator, equation (4.14), into the power series, equation (4.34) and equating equal powers of the Hamiltonian, $H_1(t)$ we get

$$i\hbar\frac{d(\rho_0)}{dt} = [H_0, \rho_0] \tag{4.39}$$

$$i\hbar\frac{d(\rho_1(t))}{dt} = [H_0, \rho_1(t)] + [H_1(t), \rho_0] \tag{4.40}$$

Thus, as we know ρ_0 from consideration of the system in thermal equilibrium, we can find ρ_1, the first order perturbation due to the applied field. This approach can be extended to the n-th order. Integrating equation (4.40) with respect to t and incorporating the time development operator, U_0, i.e. move into the interaction picture, we get

$$i\hbar U_0(-t)\rho_1(t)U_0(t) = \int_{-\infty}^{t} d\tau U_0(-\tau_1)[H_1(\tau_1), \rho_0]U_0(\tau_1). \tag{4.41}$$

Expanding the square bracket and inserting $U_0(\tau_1)U_0(-\tau_1) = 1$ between the operators gives, after some rearrangement

$$\rho_1' = U_0(-t)\rho_1(t)U_0(t) = -(i\hbar)^{-1}\int d\tau[H_1'(\tau_1), \rho_0]. \tag{4.42}$$

Multiplying on the right by $U_0(-t)$ and on the left by $U_0(t)$ to give ρ_1

$$\rho_1(t) = (i\hbar)^{-1}U_0(t)\int_{-\infty}^{t} d\tau_1[H_1'(\tau_1), \rho_0]U_0(-t) \tag{4.43}$$

This is a description of the time evolution of a perturbed state in terms of the density operator. Substituting this expression for $\rho_1(t)$ into the first order current density equation (4.36)

$$\mathbf{J}^{(1)}(t) = -V^{-1}\beta\mathbf{A}(t) + V^{-1}Tr\left\{(i\hbar)^{-1}U_0(t)\int_{-\infty}^{t} d\tau[H_1'(\tau_1), \rho_0]U_0(-t)\mathbf{\Pi}\right\}. \tag{4.44}$$

We now change notation in order to introduce α , the polarization of the field **A**, i.e.

$$H_1(t) = -\Pi_\alpha A_\alpha(t) \tag{4.45}$$

and

$$H_1'(t) = -U_0(-t)\Pi_\alpha A_\alpha(t)U_0(t) \tag{4.46}$$

where H_1' is the perturbing Hamiltonian in the interaction picture.

To transfer the current operator into the interaction picture we let

$$\Pi_\alpha'(t) = U_0(-t)\Pi_\alpha U_0(t) \tag{4.47}$$

As **A**(t) commutes with U_0, the interaction Hamiltonian in the interaction picture can be expressed as

$$H_1'(t) = -\Pi_\alpha(t)A_\alpha(t). \tag{4.48}$$

Therefore

$$[H_1'(\tau_1), \rho_0] = [-\Pi_{\alpha_1}(\tau_1)A_{\alpha_1}(\tau_1), \rho_0] \tag{4.49}$$

Substituting this expression for the commutator into 4.44 gives

$$J_\mu^{(1)}(t) = -V^{-1}\beta A_{\alpha_1}\delta_{\mu\alpha} - V^{-1}Tr\left\{(i\hbar)^{-1}U_0(t) \int_{-\infty}^{t} d\tau A_{\alpha 1}(\tau_1)[\Pi_{\alpha_1}(\tau_1), \rho_0]U_0(-t)\Pi_\mu\right\}, \tag{4.50}$$

As $U_0(t)$ commutes with each $\mathbf{A}$(t) it can be removed from the product involving A_α. Rearranging gives

$$\begin{aligned} J_\mu^{(1)}(t) &= -V^{-1}\beta A_{\alpha_1}\delta_{\mu\alpha} - V^{-1}(i\hbar)^{-1} \\ &\times \int_{-\infty}^{t} d\tau A_{\alpha_1}(\tau_1) Tr\left\{U_0(t)[\Pi_{\alpha_1}(\tau_1), \rho_0]U_0(-t)\Pi_\mu\right\}, \end{aligned} \tag{4.51}$$

We expressed the optical field in terms of the electric field in equation (4.20) Substituting this into equation (4.51)

$$\begin{aligned} J_\mu^{(1)}(t) &= -V^{-1}\beta \int_{-\infty}^{\infty} d\omega_1 E_{\alpha_1}(\omega_1)(i\omega_1)^{-1} e^{-i\omega_1\tau_1}\delta_{\mu\alpha} \\ &\quad -V^{-1}(i\hbar)^{-1} \int_{-\infty}^{t} d\tau Tr\left\{U_0(t)[\Pi_{\alpha_1}(\tau_1), \rho_0]U_0(-t)\Pi_\mu\right\} \\ &\quad \times \int_{-\infty}^{\infty} d\omega E_{\alpha_1}(\omega_1)(i\omega_1)^{-1} e^{-i\omega_1\tau_1} \\ &= \int_{-\infty}^{\infty} d\omega[-V^{-1}\beta(i\omega_1)^{-1}\delta_{\mu\alpha} - V^{-1}(i\hbar)^{-1} \\ &\quad \times \int_{-\infty}^{t} d\tau_1 Tr\left\{U_0(t)[\Pi_{\alpha_1}(\tau_1), \rho_0 e^{-i\omega_1(\tau_1 - t)}]U_0(-t)\Pi_\mu\right\}] \\ &\quad \times E_{\alpha_1}(\omega_1)(i\omega_1)^{-1} e^{-i\omega_1 t} d\omega \end{aligned} \tag{4.52}$$

The macroscopic polarization current density in terms of conductivity tensors in the first order is defined as (Butcher 1990)

$$J_\mu^{(1)}(t) = \int_{-\infty}^{\infty} d\omega_1 \sigma_{\mu\alpha_1}^{(1)}(\omega_1) E_{\alpha_1}(\omega_1) e^{(-i\omega_1 t)} \tag{4.53}$$

By comparison of the above two equations we can identify the conductivity tensor to be

$$\begin{aligned} \sigma_{\mu\alpha_1}^{(1)}(\omega_1) &= -V^{-1}\beta(i\omega_1)^{-1}\delta_{\mu\alpha} - V^{-1}(i\hbar)^{-1}\int_{-\infty}^{t} d\tau_1 \\ &\times Tr\left\{U_0(t)[\Pi_{\alpha_1}(\tau_1), \rho_0 e^{-i\omega_1(\tau_1 - t)}]U_0(-t)\Pi_\mu\right\}] \end{aligned} \tag{4.54}$$

We now wish to eliminate t from the above expression to give us the time-independent conductivity. We expand the commutator in equation (4.54), insert $U_0(-t)U_0(t) = 1$ between the operators and rearrange to give

$$U_0(t)[\Pi_{\alpha_1}(\tau_1), \rho_0]U_0(-t) = [\Pi_{\alpha_1}(\tau_1 - t), \rho_0]. \tag{4.55}$$

Now introduce a change of variable $\tau_1 \rightarrow \tau_1' + t$ which changes the limits of integration i.e. when $\tau = t, \tau_1'$=0.

$$\sigma^{(1)}_{\mu\alpha_1}(\omega_1) = -V^{-1}\beta(i\omega_1)^{-1}\delta_{\mu\alpha} - V^{-1}(i\hbar)^{-1}\int_{-\infty}^{0} d\tau_1' Tr\left\{[\Pi_{\alpha_1}(\tau_1'), \rho_0]e^{-i\omega_1\tau_1'}\Pi_\mu\right\} \tag{4.56}$$

The trace of a product of operators is invariant under cyclic permutation of their order, therefore rearranging the operators gives, dropping the primes on the τ for convenience.

$$Tr\{\rho_0[\Pi_\mu, \Pi_{\alpha_1}(\tau_1)]\} \tag{4.57}$$

Thus we can obtain an expression for the first order susceptibility (Butcher 1990)

$$\begin{aligned}\chi^{(1)}_{\mu\alpha}(\omega_1) &= -\frac{1}{i\omega_1^2\epsilon_0}\sigma^{(1)}_{\mu\alpha_1} \\ &= -\frac{\beta\delta_{\mu\alpha_1}}{\omega_1^2 V\epsilon_0} - \frac{1}{V\epsilon_0\omega_1^2\hbar}\int_{-\infty}^{0} d\tau_1' Tr\{[\rho_0[\Pi_\mu, \Pi_{\alpha_1}]e^{-i\omega_1\tau_1'}\}\end{aligned} \tag{4.58}$$

4.3 Application to Semiconductors

Thus far no account has been taken of any specific system of charged particles, we have discussed any medium in the dipole approximation. We must now relate this equation specifically to semiconductors. In order to do this we must make additional assumptions about our model in order to study the current operators further. We assume that the fixed ion cores are surrounded by a sea of electrons free to move throughout the entire volume of the system. We also

assume that the density of electrons is sufficiently low that the electronic dipole interaction is negligible. i.e. each electron experiences the same macroscopic electric field $\mathbf{E}(t)$. We must also make the transition from many- to single-particle operators.

4.3.1 Conversion to Single Particle Operators

We originally defined the current operator to be a sum of single particle momentum operators in equation (4.16). Ignoring the ionic contribution we can write

$$\mathbf{\Pi} = -\frac{e}{m}\sum_{j}^{N}\mathbf{p}_j, \tag{4.59}$$

where $\mathbf{p}_j$ is the single-electron momentum of the jth electron. Similarly

$$H_0 = \sum_{j}^{N} H_j, \tag{4.60}$$

The unperturbed time development operator of a many particle system is

$$U_0(t) = e^{-iH_0 t/\hbar}. \tag{4.61}$$

Combining the above two equations leads to

$$U_0(t) = U_1(t)U_2(t)\ldots U_j(t)\ldots U_N(t), \tag{4.62}$$

where the single-particle unperturbed time development operators are,

$$U_j(t) = e^{-iH_j t/\hbar}. \tag{4.63}$$

In equation (4.54) we defined

$$\Pi_\alpha(t) = U_0(-t)\Pi_\alpha U_0(t) \tag{4.64}$$

Similarly for a single particle operator system we have

$$\Pi_\alpha(t) = -\frac{e}{m}U_0(-t)\sum_{j} p_j^\alpha U_0(t), \tag{4.65}$$

where p_j^α is the component of momentum of the j'th electron in the α direction. The single particle operators of different electrons commute, therefore $U_l(t)$ commutes with p_j^α provided $l \neq j$. We also note that $U_l(t)U_l(-t) = 1$. Therefore

$$\Pi_\alpha(t) = -\frac{e}{m}\sum_j^N p_j^\alpha(t), \tag{4.66}$$

where

$$p_j^\alpha(t) = U_j(-t)p_j^\alpha U_j(t). \tag{4.67}$$

We can now express the multiple commutator of our susceptibility expression in terms of single particle operators.

$$[\Pi_\mu, \Pi_{\alpha_1}(\tau_1)] = [-\frac{e}{m}\sum_j^N p_j^\mu, -\frac{e}{m}\sum_j^N p_j^\alpha(\tau_1)] = \frac{e^2}{m^2}\sum_j^N [p^\mu, p_j^\alpha(\tau_1)] \tag{4.68}$$

Substituting this into the Trace in the expression for χ, equation (4.58) gives

$$Tr\{\rho_0 \sum_j^N [p_j^\mu p_j^\alpha(\tau_1)]\}e^{-i\omega_1\tau_1} \tag{4.69}$$

We must now reduce the operation Trace to single particle operators. The value of a trace is independent of representation so we can use Slater determinants to express the many particle wavefunctions as anti-symmetrised combinations of single particle wavefunctions. The Slater determinant Φ_I, representing the many electron state, I, is (Inkson, 1983),

$$\Phi_I = \frac{1}{\sqrt{N!}}(-1)^{\mathcal{P}}\mathcal{P}\{u_{i_1}(\theta_1)u_{i_2}(\theta_2)\ldots u_{i_N}(\theta_N)\}, \tag{4.70}$$

where $u_{i_j}(\theta_j)$ is the single-electron orthonormal wavefunction in the i_j'th state, and θ_j represents the coordinates of the j'th electron. N is the total number of electrons in the system. $\mathcal{P}$ is a permutation operator which represents a summation of terms obtained by the permutations of the labels i_1, i_2, ..., i_N. $(-1)^{\mathcal{P}}$ takes the value $+1$ or -1 for permutations of odd or even parity

respectively. The many-particle state I can be defined by the particular single-electron states which occur in equation (4.70). These correspond to the single electron states which are occupied when the system is in state I. Our Trace is now of the form

$$T = Tr\{\rho_0 \sum_j^N C_j\}, \tag{4.71}$$

where C_j is a single particle operator and represents the commutator in 4.69. From the definition of the trace this can be written as

$$T = \sum_I (\rho_0 \sum_j^N C_j)_{II}, \tag{4.72}$$

Using the matrix product rule of operators this becomes

$$T = \sum_I \sum_J (\rho_0)_{IJ} (\sum_j^N C_j)_{JI}. \tag{4.73}$$

As previously stated the trace of a function does not depend on the representation used therefore we choose to work in the energy representation (Butcher 1990) which consists of the complete set of orthonormal eigenstates $\{\psi_n\}$ which satisfy the time-independent Schrödinger equation. In the energy representation the density matrix operator ρ_0 is diagonal i.e. $(\rho_0)_{IJ} = (\rho_0)_{II}\delta_{II}$. Therefore the trace may be expressed as

$$T = \sum_I \{(\rho_0)_{II} (\sum_j^N C_j)_{II}\}, \tag{4.74}$$

where ρ^0_{II} are the diagonal elements of the unperturbed many particle density operator. The matrix elements are defined as

$$P(\psi)_{ji} = \int d\xi u_j^* P(\psi) u_i. \tag{4.75}$$

Thus we can write

$$T = \sum_I \rho^0_{II} \int d\xi_1 \int d\xi_2 \ldots \int d\xi_N \Phi_I^* (\sum_j^N C_j) \Phi_I, \tag{4.76}$$

where $d\xi_j$ is the volume element of the j'th electron. Substituting for the wavefunction Φ_I from equation (4.70) gives

$$T = \sum_I \rho^0_{II} \frac{1}{N!} \left[\int d\xi_1 \int d\xi_2 \ldots \int d\xi_N (-1)^{\mathcal{P}} \mathcal{P}\{u^*_{i_1}(\theta_1) u^*_{i_2}(\theta_2) \ldots u^*_{i_N}(\theta_N)\} \times \right.$$
$$\left. (\sum_j^N C_j)(-1)^{\mathcal{P}'} \mathcal{P}'\{u_{i_1}(\theta_1) u_{i_2}(\theta_2) \ldots u_{i_N}(\theta_N)\} \right]. \qquad (4.77)$$

There are N! permutations of each $\mathcal{P}$ and $\mathcal{P}'$ and N operators C_j, therefore the expansion of the square bracket yields $N \times (N!)^2$ terms. Each term for a given C_j and specific permutation $\mathcal{P}$ is denoted by $\alpha_1, \alpha_2, \ldots, \alpha_n$ etc. and for $\mathcal{P}'$ by β_1,β_2,. . . ,β_N, is represented by

$$(-1)^{\mathcal{P}'}(-1)^{\mathcal{P}} \int d\xi_1 u^*_{\alpha_1}(\theta_1) u_{\beta_1}(\theta_1) \ldots \int d\xi_j u^*_{\alpha_j}(\theta_j) C_j u_{\beta_j}(\theta_j) \ldots \int d\xi_N u^*_{\alpha_N}(\theta_N) u_{\beta_N}(\theta_N), \qquad (4.78)$$

where the fact that the wavefunctions and operators of different electrons commute has been used. The u_i's are orthonormal therefore the above equation equals 0 unless $\mathcal{P} = \mathcal{P}'$ when it becomes

$$\int d\xi_j u^*_{\alpha_j}(\theta_j) C_j u_{\alpha_j}(\theta_j). \qquad (4.79)$$

Each C_j yields N! permutations $\mathcal{P}$, each equal to one permutation of $\mathcal{P}$'. There are therefore N! such terms. However not all are independent. Different permutations of α_i's for i≠j leave the above equation unchanged. Since only α_i appears only N independent terms remain of the form

$$(N-1)! \int d\xi_j u^*_{\alpha_j}(\theta_j) C_j u_{\alpha_j}(\theta_j), \qquad (4.80)$$

Each of these N terms α_j represents one of the N occupied single electron states $i_1 \ldots i_n$. Therefore the square bracket becomes

$$(N-1)! \sum_{j=1}^{N} \left[\int d\xi_j u^*_{i_1}(\theta_j) C_j u_{i_1}(\theta_j) + \ldots + \int d\xi_j u^*_{i_N}(\theta_j) C_j u_{i_N}(\theta_j) \right]. \qquad (4.81)$$

All electrons are independent and indistinguishable, therefore all c_j's are identical and we can replace $\sum_j$ in the above equation by N which leaves

$$N! \sum_i \int d\xi u_i^*(\theta_i) C u_i(\theta_i), \tag{4.82}$$

where the sum i runs over all occupied states and C is any single particle operator. Substituting this back into our expression of the trace gives

$$T = \sum_I \rho_{II}^0 \sum_i \int d\xi u_i^* C u_i. \tag{4.83}$$

This can be generalized by extending the sum over occupied states i to the sum over all single electron states i', both occupied and unoccupied. Introducing the notation $n_I^{i'}$ which is the occupation number of the i'th single electron state (either 1 or 0) when the system is in the many particle state I.

$$\begin{aligned} T &= \sum_I \rho_{II}^0 \sum_{i'} n_I^{i'} \int d\xi u_{i'}^* C u_{i'} \\ &= \sum_{i'} \sum_I \rho_{II}^0 n_I^{i'} \int d\xi u_{i'}^* C u_{i'} \end{aligned} \tag{4.84}$$

where $\sum_I \rho_{II}^0 n_I^{i'}$is the average occupation of the i th state when the system is in thermal equilibrium and corresponds to the Fermi Dirac distribution

$$f_{i'} = \frac{1}{1 + e^{-(E_{i'} - \mu)/kT}}. \tag{4.85}$$

Combining the above gives

$$T = \sum_{i'} f_{i'} \int d\xi_{i'} u_{i'}^* C u_{i'} = \sum_{i'} f_{i'} (C)_{i'i'}. \tag{4.86}$$

In a semiconductor system single electron states are Bloch functions and can be described by the wavevector in the Brillouin zone and the band (or miniband) index. Therefore the summation over the single electron states $\sum_i$, can be replaced by a sum over the wavevectors in the Brillouin zone $\sum_k$ and a sum over the minibands $\sum_a$. Therefore

$$T = \sum_k \sum_a f_{a_k} (C)_{a_k a_k}, \tag{4.87}$$

where a_k is state in miniband a with wavevector k. We shall drop the suffix k for clarity. Our expression for the susceptibility now becomes

$$\chi^{(1)}_{\mu\alpha}(\omega_1) = -\frac{e^2 N}{Vm\omega_1^2\epsilon_0}\delta_{\mu\alpha_1} - \frac{e^2}{i\omega^2 Vm^2\epsilon_0\hbar}\sum_k \int_{-\infty}^{0} d\tau_1 \sum_a f_a(C)_{aa} e^{-i\omega_1\tau_1} \quad (4.88)$$

where $C = [p^\mu, p^{\alpha_1}(\tau_1)]$ We now have a single particle operator expression. We must now expand C i.e.

$$C = [p^\mu, p^\alpha(\tau_1)] = p^\mu p^\alpha(\tau_1) - p^\alpha(\tau_1)p^\mu \quad (4.89)$$

We need an expression for the matrix element of C. By the rules of matrix products operator $O = O'O''$ then

$$O_{aa} = \sum_b (O')_{ab}(O'')_{ba} \quad (4.90)$$

where b is summed over all single particle states of the system. Therefore

$$(C)_{aa} = \sum_b [p^\mu_{ab} p^\alpha_{ba}(\tau_1) - p^\alpha_{ba}(\tau_1)p^\mu_{ab}] \quad (4.91)$$

where p^μ_{ab} are the matrix elements of p^μ and $p^\alpha_{ba}(\tau_1)$ are matrix elements of $p^\alpha(\tau_1)$. We now need to know the form of the matrix elements in the interaction picture. By definition in equation (4.67) we have

$$p^\alpha_{ab}(\tau) = [U_0(-\tau)p^\alpha U_0(\tau)]_{ab}. \quad (4.92)$$

Applying the matrix product rule (4.90) to (4.92) we obtain,

$$p^\alpha_{ab}(\tau) = \sum_{uv} [U_0(-\tau)]_{au} p^\alpha_{uv} [U_0(\tau)]_{vb}. \quad (4.93)$$

We can define the time development operators, U_0, as a power series in the unperturbed Hamiltonian H_0. The time development operators are diagonal matrices in the energy representation and thus unperturbed time development operators can be written as

$$[U_0(\tau)]_{uv} = e^{-iE_u t/\hbar}\delta_{uv}. \quad (4.94)$$

This enables us to write

$$\begin{aligned} p^{\alpha}_{uv}(\tau) &= e^{iE_u t/\hbar} p^{\alpha}_{uv} e^{-iE_v t/\hbar} \\ &= e^{i\Omega_{uv} t} p^{\alpha}_{uv}, \end{aligned} \tag{4.95}$$

where we define the energy gap Ω_{uv} by,

$$\Omega_{uv} = (E_u - E_v)/\hbar. \tag{4.96}$$

The summation over b is over all single electron states, i.e. over all wavevectors within the Brillouin Zone and all superlattice minibands. In the electric dipole approximation we can neglect momentum matrix elements between states of different wavevector (Bassini 1975) and need only consider transitions which involve no change of k, i.e. are vertical in wavevector space. Therefore the summations may be restricted to sums over b, with the wavevector fixed at k, the wavevector associated with state a. Combining this into our expression for the first order susceptibility gives

$$\begin{aligned} \chi^{(1)}_{\mu\alpha}(\omega_1) = -\frac{e^2 N}{V m \omega_1^2 \epsilon_0} \delta_{\mu\alpha_1} - \frac{e^2}{i\omega^2 V m^2 \epsilon_0 \hbar} \sum_k \int_{-\infty}^{0} d\tau_1 \sum_a f_a \sum_b \\ \times [p^{\mu}_{ab} p^{\alpha}_{ba}(\tau_1) e^{i(\Omega_{ba} - \omega_1)\tau_1} - p^{\alpha}_{ab}(\tau_1) p^{\mu}_{ba} e^{i(\Omega_{ab} - \omega_1)\tau_1}] \end{aligned} \tag{4.97}$$

We now need to perform the time integrals in the above equation. We assume that the frequencies lie in the upper half of the complex plane, to ensure that these integrals will converge. We drop this condition in the next section when we introduce the relaxation terms in a phenomenological way. The time integrals can be moved inside the summations over a and b

$$\int_{-\infty}^{0} d\tau_1 [p^{\mu}_{ab} p^{\alpha}_{ba}(\tau_1) e^{i(\Omega_{ba} - \omega_1)\tau_1} - p^{\alpha}_{ab}(\tau_1) p^{\mu}_{ba} e^{i(\Omega_{ab} - \omega_1)\tau_1}] \tag{4.98}$$

$$= [\frac{p^{\mu}_{ab} p^{\alpha}_{ba}(\tau_1)}{i(\Omega_{ba} - \omega_1)} - \frac{p^{\alpha}_{ab}(\tau_1) p^{\mu}_{ba}}{i(\Omega_{ab} - \omega_1)}] \tag{4.99}$$

$$= \frac{1}{i}[\frac{p^{\mu}_{ab} p^{\alpha}_{ba}(\tau_1)}{(\Omega_{ba} - \omega_1)} + \frac{p^{\alpha}_{ab}(\tau_1) p^{\mu}_{ba}}{(\Omega_{ba} + \omega_1)}] \tag{4.100}$$

where we have used the fact that $\Omega_{ab} = -\Omega_{ba}$.

Therefore our expression for $\chi^{(1)}$ becomes

$$\chi^{(1)}_{\mu\alpha_1}(\omega_1) = -\frac{e^2 N}{Vm\epsilon_0\omega_1^2}\delta_{\mu\alpha} + \frac{e^2}{Vm^2\epsilon_0\hbar\omega_1^2}\sum_k\sum_a f_a \sum_b \times \left[\frac{p^\mu_{ab}p^\alpha_{ba}}{(\Omega_{ba}-\omega_1)} + \frac{p^\alpha_{ab}p^\mu_{ba}}{(\Omega_{ba}+\omega_1)}\right]. \quad (4.101)$$

4.3.2 Relaxation

Thus far we have taken no account of the relaxation processes that must exist in a system like this, we simply restricted the available frequencies to the upper half of the frequency plane to ensure convergence of the integrals. In a real system the phenomena of relaxation would give the necessary convergence. We include the relaxation by means of a transition line-width introduce by making the substitution

$$\Omega_{uv} \rightarrow \Omega_{uv} \pm i\Gamma_{uv}, \quad (4.102)$$

where the sign is chosen to ensure convergence. This should give a good approximation to the relaxation processes provided that fields remain small and only small population changes are introduced. This is valid as we have already made the assumption that the changes induced by the field can be included as a perturbation to the no-field case. For strong fields this is no longer true and the density matrix equations for the system would have to be solved directly. Thus our final expression for the first order susceptibility is

$$\chi^{(1)}_{\mu\alpha}(\omega_1) = -\frac{e^2 N}{Vm\epsilon_0\omega_1^2}\delta_{\mu\alpha} + \frac{e^2}{Vm^2\epsilon_0\hbar\omega_1^2}\sum_k\sum_a f_a \sum_b \times \left[\frac{p^\mu_{ab}p^\alpha_{ba}}{(\Omega_{ba}-i\Gamma_{ba}-\omega_1)} + \frac{p^\alpha_{ab}p^\mu_{ba}}{(\Omega_{ba}+i\Gamma_{ba}+\omega_1)}\right]. \quad (4.103)$$

4.3.3 Absorption

The real part of the linear susceptibility, $Re[\chi^{(1)}]$ is proportional to the refractive index. We are interested in the imaginary part, i.e. $Im[\chi^{(1)}]$ which represents the losses in the medium, i.e. the absorption. The absorption coefficient, α, is given by

$$\alpha = \frac{4\pi\omega}{n_0 c} \times Im[\chi^{(1)}] \tag{4.104}$$

where n_0 is the refractive index of the system, taken to be the average refractive index for the system. The factor of 4π arises in the change from e.s.u. to S.I. units.

4.4 Effect of Symmetry Transformations on the First Order Susceptibility

The effect of the symmetry operations contained in the C_{2v} point group when performed on the **k**-vectors and the matrix elements is given in table (4.1). Assume that the matrix elements and energies are known throughout the irreducible segment of the zone. Consider the component of the first order susceptibility $\chi^{(1)}_{xx}$. The application of the operator σ_{ad} to a point in the irreducible segment transforms P_x to P_y. The energies will be unaffected by this transformation, thus $\chi^{(1)}_{xx}$ in the new zone is identical to $\chi^{(1)}_{yy}$ in the original segment. Repeating this procedure for the remaining regions for the Brillouin zone and we get for the total contribution to $\chi^{(1)}_{xx}$;

$$\chi^{(1)}_{xx} = 2\chi^{(1)\,[IR]}_{xx} + 2\chi^{(1)\,[IR]}_{yy} + 2\chi^{(1)\,[IR]}_{xx^*} + 2\chi^{(1)\,[IR]}_{yy^*} \tag{4.105}$$

where $[IR]$ represents the contribution from the irreducible segment and xx^* means that the conjugates of the numerators are used. However, $p_{ab} = p^*_{ba}$ so

Operation	k_x	k_y	k_z	P_x	P_y	P_z
E	k_x	k_y	k_z	P_x	P_y	P_z
σ_{ad}	k_y	k_x	k_z	P_y	P_x	P_z
σ_{bd}	$-k_y$	$-k_x$	k_z	$-P_y$	$-P_x$	P_z
C_{2z}	$-k_x$	$-k_y$	k_z	$-P_x$	$-P_y$	P_z
$\mathcal{T}E$	$-k_x$	$-k_y$	$-k_z$	$-P_x^*$	$-P_y^*$	$-P_z^*$
$\mathcal{T}\sigma_{ad}$	$-k_y$	$-k_x$	$-k_z$	$-P_y^*$	$-P_x^*$	$-P_z^*$
$\mathcal{T}\sigma_{bd}$	k_y	k_x	$-k_z$	P_y^*	P_x^*	$-P_z^*$
$\mathcal{T}C_{2z}$	k_x	k_y	$-k_z$	P_x^*	P_y^*	$-P_z^*$

Table 4.1: *The eight related wavevectors under the C_{2v} point group and time reversal symmetry operations, and the Cartesian components of the momentum matrix elements at these wavevectors. $\mathcal{T}$ represents the time reversal operation.*

this expression reduces to

$$\chi_{xx}^{(1)} = 4\chi_{xx}^{(1)\,[IR]} + 4\chi_{yy}^{(1)\,[IR]} \tag{4.106}$$

An identical expression is obtained for $\chi_{yy}^{(1)}$.

For the z-component, $\chi_{zz}^{(1)}$, the numerator has 2 matrix elements in the z direction. Each of the sixteen symmetry operations produces $\chi_{zz}^{(1)}$ in a different segment of the zone, therefore

$$\chi_{zz}^{(1)} = 16\chi_{zz}^{(1)\,[IR]} \tag{4.107}$$

We also have the components $\chi_{xy}^{(1)}$ and $\chi_{yx}^{(1)}$. Using the same procedure we get

$$\chi_{xy}^{(1)} = \chi_{yx}^{(1)} = 4\chi_{xy}^{(1)} + 4\chi_{yx}^{(1)} \tag{4.108}$$

The linear susceptibility calculations were performed in the x, y, z reference frame. In chapter 3 the relative position of the old and new axis, (x, y, z) and

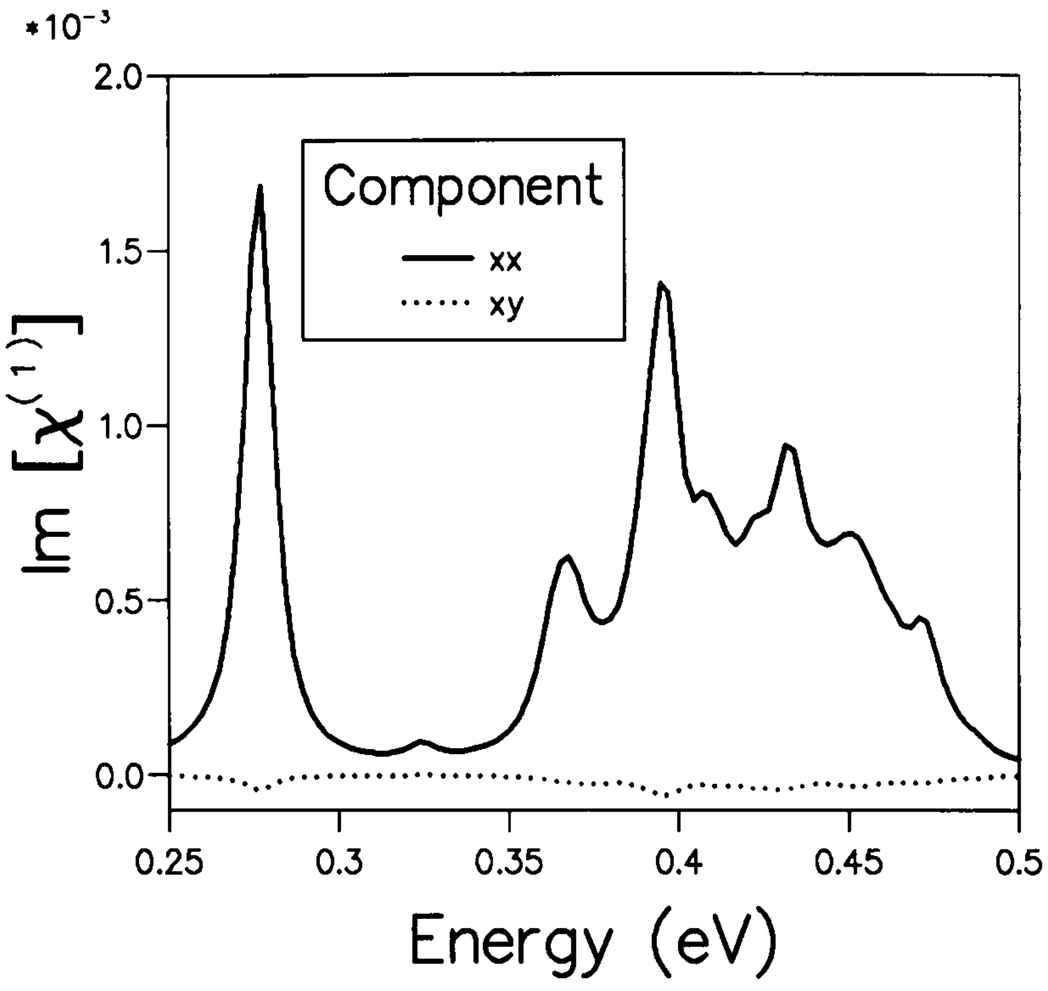


Figure 4.1: *The comparison between components of the first order susceptibility for a 50%, 20 monolayer structure*

(x', y', z') were discussed. The xx referred to in $\chi^{(1)}_{xx}$ is defined with respect to the old set of axis. In order to apply the symmetry operations we must relate the old and new axes.

$$\chi^{(1)}_{x'x'} = \chi^{(1)}_{xx} - \chi^{(1)}_{xy} \tag{4.109}$$

$$\chi^{(1)}_{y'y'} = \chi^{(1)}_{xx} + \chi^{(1)}_{xy} \tag{4.110}$$

$$\chi^{(1)}_{z'z'} = \chi^{(1)}_{zz} \tag{4.111}$$

In figure 4.1 we plot the calculated $\chi^{(1)}$'s for a typical structure, namely a 20 monolayer wide well containing a $Si_{0.5}Ge_{0.5}$ alloy. χ_{xy} is very small can be seen to be very small compared with χ_{xx}, thus for all future calculations the normal incidence response will be taken the be the average of $\chi^{(1)}_{xx}$ and $\chi^{(1)}_{yy}$ and $\chi^{(1)}_{xy}$ will be neglected.

Another point worth making is that in a practical IR detector, the incoming light will be unpolarized, therefore for normal incidence absorption both x and

y polarizations will be present. Thus the total normal incidence output will be the sum of these two components. This is not included in subsequent calculations where all incoming light is considered to be polarized in one direction only.

Chapter 5

Doping and Fermi Levels

In this chapter we consider the remaining criteria which need to be determined before the absorption response can be calculated. We discuss the method used for sampling the Brillouin zone and details of the calculation.

For a quantum well structure utilizing intersubband transitions to be a useful device it must be doped, i.e. extra carriers introduced to populate the valence band. We consider doping of the well and the barrier, and calculate Fermi levels for various temperatures and structures with each method. We also consider the exchange and direct interactions which arise when extra charge carriers are introduced.

5.1 Sampling

Both the expression for the pseudopotential matrix elements and the susceptibility expression contain a sum over $\mathbf{k}$ which should in theory be over every discrete k-point point in the Brillouin zone. For an 'infinite' crystal the k-points tend to a continuum and the sum to an integral. We approximate this integral to a sum over a large number of k-points in the quantum well B.Z.

and normalize appropriately by making the substitution

$$\frac{1}{V}\sum_{\mathbf{k}}^{N_{BZ}} \rightarrow \frac{1}{N_{SZ}}\left(\frac{V_{SZ}}{V_{BZ}}\right)\left(\frac{1}{V_C}\sum_{\mathbf{k}}^{N_{SZ}}\right) \tag{5.1}$$

where N_{SZ} is the total number of sampling points, V_C is the volume of the primitive unit cell, V_{BZ} is the volume of the Brillouin zone and V_{SZ} is the volume of the sampled region of the zone. The zone need only be sampled over regions where there is significant carrier population as where there are no carriers there will be no contribution to the susceptibility. For 0K the extent of carriers is defined by the Fermi surface and thus only the region within this need be sampled. For higher temperatures the carriers move away from the zone centre and thus the sampling must be performed over a larger volume.

In order to determine our sampling zone we make a rough calculation of the Fermi energy. Without knowledge of the Fermi level it is impossible to sensibly chose the area of the zone to sample, as demonstrated by figure 5.1.

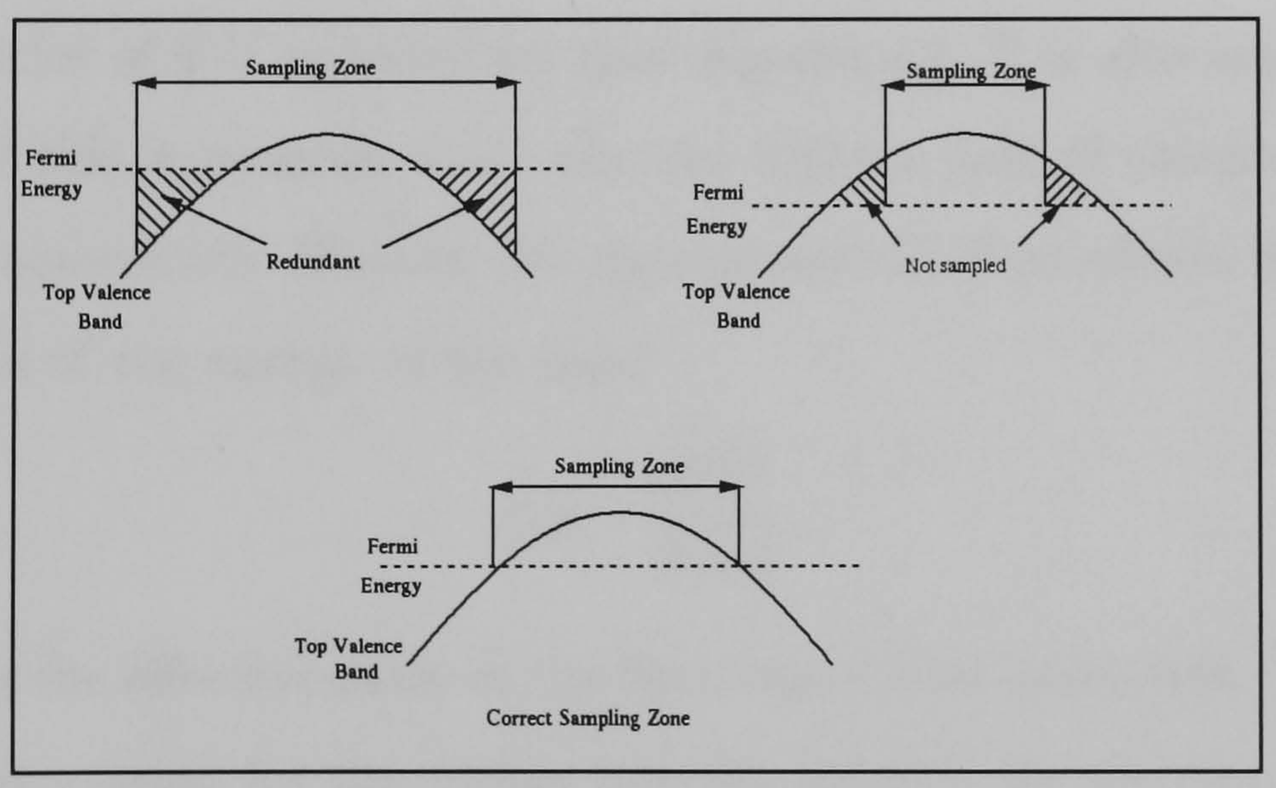


Figure 5.1: *Choosing the correct sampling zone requires knowledge of Fermi level.*

The Fermi levels can be calculated using an effective mass approach and

assuming parabolic bands. The formula for finite temperature is

$$E_F = -kTln(e^{\frac{\pi\hbar^2 N_h}{m_{hh}kT}} - 1) \tag{5.2}$$

$$E_F(T \to 0) = -\frac{N_h \pi \hbar^2}{m_{hh}}. \tag{5.3}$$

The distribution of hole carriers was determined using the familiar Fermi-Dirac distribution

$$f_h = 1 - f(E) = \frac{1}{1 + e^{(E_F - E)/kt}}. \tag{5.4}$$

The total number of carriers may be expressed as

$$N_h = \int_{-\infty}^{0} D(E) f_h(E) dE \tag{5.5}$$

where $D(E)$ is the density of states which in the effective mass approximation is given by (Ando 1982)

$$D(E) = -2\frac{1}{(2\pi)^2} 2\pi k \frac{dk}{dE} \tag{5.6}$$

where the factor of 2 is included for spin degeneracy. It is also assumed in the derivation of this expression that only the highest pair of valence minibands needs to be considered. Making the approximation of parabolic bands we get an expression of the energy of the band

$$E = -\frac{\hbar^2 k^2}{2m_{hh}} \tag{5.7}$$

where m_{hh} is the effective mass of the first heavy hole miniband. Thus substituting this expression for the energy into the formula for the density of states we get

$$D(E) = \begin{cases} \frac{m_{hh}}{\pi\hbar^2} & \text{if } E < 0 \\ 0 & \text{if } E > 0 \end{cases} \tag{5.8}$$

We need to find a suitable value for the effective mass. Values quoted in the literature are for bulk materials. The non-parabolicity of the quantum

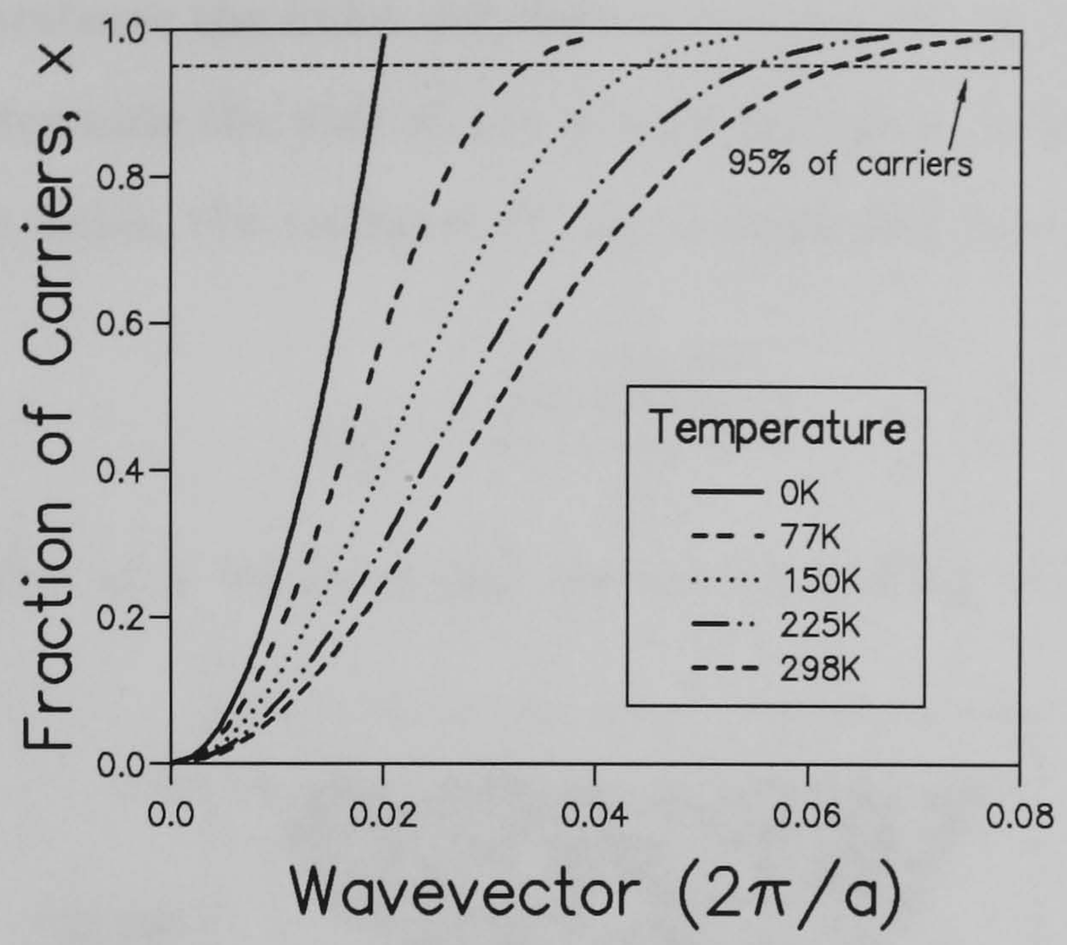


Figure 5.2: *The distribution of carriers with wavevector for a 50%, 20 monolayers structure with a doping concentration of* $4.2\times10^{15}\mathrm{m}^{-2}$.

well minibands means these values will not accurately represent the masses in the quantum well. Therefore we calculate the band structure along the $\Gamma - X$ symmetry line. The best fit parabola to the top valence miniband is determined and thus a value for the effective mass, m^*, is found. Due to the non-parabolicity of the calculated band structure a different value for m^* will be obtained dependent on the distance along the k_x axis considered for the fitting procedure. The most useful values are likely to be obtained by restricting k_x to approximately k_F, the wavevector at the Fermi energy.

We again use effective mass expressions to predict a rough estimate of the Fermi levels and the distribution of carriers. For 0K all the holes occupy states with energy above E_F. Thus, assuming parabolic bands, we can substitute E_F into equation (5.7) to obtain the $k_{\|}$ vector corresponding to this energy, i.e.

$$|k_{\|}| < \frac{\sqrt{2m_h h|E_F|}}{\hbar}. \tag{5.9}$$

For finite temperatures the holes are distributed according to the Fermi-Dirac statistics To determine the size of the sampling region required to include a fraction x of the holes, the energies E_x^T are substituted into equation (5.7) to give

$$|k_x^T| < \frac{\sqrt{2m_h h|E_x^T|}}{\hbar}. \tag{5.10}$$

We find the value of $\mathbf{k}$ which marks the extent of 95% of the carriers. The

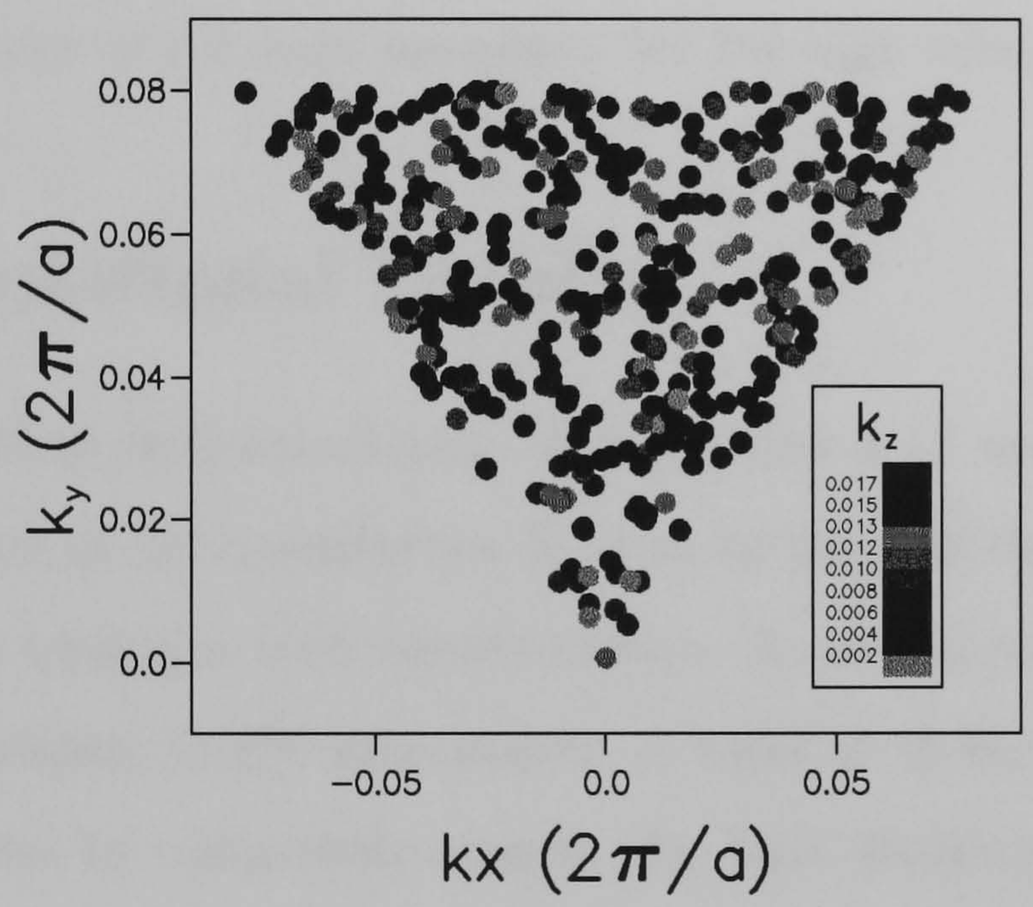


Figure 5.3: *A typical high temperature random sampling set consisting of 400 points within the irreducible segment of the Brillouin zone.*

procedure for fitting m^* is then repeated to this new k limit. The previous steps are then repeated. The zone up to this limit of k is then sampled. This procedure is performed at 0K and 298K. Figure 5.2 shows a typical distribution showing that for a low temperature sampling set is $0.03\left(\frac{2\pi}{a}\right)$ is a sufficient sampling volume, and for high temperatures $0.08\left(\frac{2\pi}{a}\right)$. A random sampling technique was used which has certain computational advantages. A typical high temperature random sampling set is shown in figure 5.3. The number of

sampling points was increased until convergence was achieved. For all structure in this thesis a total of 400 points was sufficient for convergence. As we can see from figure 5.3 two sampling zones are necessary. The high temperature sampling set only contains few points in the low temperature region, insufficient for convergence. To calculate at a sufficiently dense level to be sure of including enough points to make a low temperature calculation valid would require a prohibitively large number of sampling points when extended to cover the volume of the zone necessary for the high temperature situation.

5.2 Calculational Details

Before the quantum well calculation is performed it is necessary to sample the bulk structure at the superlattice **k**-point to provide the basis set for the expansion of the quantum well wavefunctions. A typical period in this thesis is 27 lattice constants, or 108 monolayers. A total of 55 bulk sampling points would be required to completely sample the bulk Brillouin Zone along the z-axis. This results in a prohibitively large matrix to be diagonalized. Thus the bulk sampling set is restricted to the central 15 points about the bulk valence band maxima, i.e. Γ. As we are dealing only with intravalence band transitions, the quantum well states will be constructed from the eigenfunctions from this region. Transitions across the fundamental gap would not be well represented by this approximation as the minima of the conduction band would not be included in the bulk sampling set. However, we do not consider such transitions in this thesis and this restricted bulk sampling set is a good approximation in our situation.

The barrier width is taken throughout to be 80Å. Experimental barrier widths are of the order of 500Å. However, increasing the barrier width increases

the period, which in turn means more bulk sampling points are required. A value of 80Å is sufficient to ensure that there is no significant overlap between states confined within the wells.

A total of 31 bulk bands (62 including spin degeneracy) was required for convergence.

5.3 Exchange Shifts

Our pseudopotential derivation was based on the assumption of intrinsic semiconductors, i.e. the valence band was completely occupied and the conduction band was empty, no doping was present. For a device based on transitions between minibands to be useful it is necessary to introduce impurities in the form of dopants to provide extra carriers. These extra carriers will interact with each other and thus produce many body effects which have not been accounted for in the Hamiltonian of the system.

In quantum well structures which are heavily doped the direct and exchange Coulomb terms must be considered. If the wells are doped then the charge neutrality within the well will mean that the direct interactions are small and the exchange interactions would account for the major part of this correction. For doped barriers the exchange contribution will still be significant but the charge of the acceptors and the carriers are now localized in different layers, thus the direct term may no longer be insignificant. The close proximity of the ionized acceptors and the charge carriers within short period structures may limit the direct interaction to negligible amounts. The main effect of the direct interaction is to move the minibands up in energy. However, there is little change in the separation of the minibands and thus this effect would make little difference to our results which are concerned with transitions between

the minibands (Bandara 1988).

We calculate the exchange shift from the work of Bandara et al. (1988) who obtain two approximate expressions for the exchange shift at the zone centre, $k = 0$, and at the Fermi energy, $k = k_f$ by assuming strong localization of the ground state within the well. Bandara's work was on n-type structures, but the results are equally applicable for our p-type situation. The expressions are

$$E_{exch}(0) \approx \frac{-e^2 k_f}{4\pi\epsilon}\left[1 - 0.25(\frac{k_f}{k_l})\right] \tag{5.11}$$

$$E_{exch}(k_f) \approx \frac{-e^2 k_f}{4\pi\epsilon}\left[\frac{2}{\pi} - 0.25(\frac{k_f}{k_l})\right] \tag{5.12}$$

where $k_f = \sqrt{2\pi\sigma}$ where σ is the 2-D carrier density, and $k_l = \frac{\pi}{L}$. We take an average of these two extremes and include this as a rigid shift applied to the top two states, i.e. HH1, as these are the only significantly occupied states at low temperatures. This is obviously a simplification of the real situation, as the exchange shift is in fact a sensitive function of k. However, a more subtle interpretation would be very complicated particularly as there is no temperature dependence in Bandara's expression, whereas the Fermi vector must be a function of temperature. As only the upper two levels are affected by this correction the result is to shift the absorption peaks to higher energies. The values of the exchange shifts for the structures considered in this thesis assuming the barriers to be doped to the concentrations used by Fromherz (1995) are given in table 5.1.

5.4 Doping

In order to produce extra charge carriers to make the quantum well a useful device, modulation doping is used. There are two alternative approaches to

Ge conc in well	Doping concentration	Well width 27 Å (20 Mls)	38Å (28 Mls)	49Å (36 Mls)	60Å (44 Mls)
15%	$2.2 \times 10^{18} cm^{-3}$	25	24	23	22
30%	$2.2 \times 10^{18} cm^{-3}$	24	23	22	21
50%	$7.0 \times 10^{17} cm^{-3}$	13	13	13	12

Table 5.1: *Exchange shifts calculated for the structures considered in this thesis assuming the barriers to be doped. All values are in meV.*

this, either the barriers or the wells can be doped. Here we discuss both cases and calculate and compare Fermi levels in each situation.

5.4.1 Barrier Doping

With this technique impurities, usually boron in the case of SiGe systems, are introduced into the barrier. These impurities provide empty states in the barriers as they have a valency of one less than the host material (boron has a valency of three compared to Si and Ge which have a valency of four) as shown in figure 5.4. This situation is unstable because the Fermi energy is not equal on both sides of the hetrojunction barrier, i.e. in the barrier layer the Fermi level is at the acceptor level, and on the well side it is negative as there are no free carriers in the well. To restore equilibrium the carriers cross the barrier either by tunneling or by thermionic effects. Once in the well, they quickly loose energy by emitting several phonons to relax to the bottom of the well in a time scale of 10^{-12}s. This process continues until all states with energies less that $(V_b - R)$, where V_b is the barrier height and R is the binding energy of the impurity, are filled. This movement of electrons to empty states in the barrier creates holes in the well. These holes move to the top of the well (or,

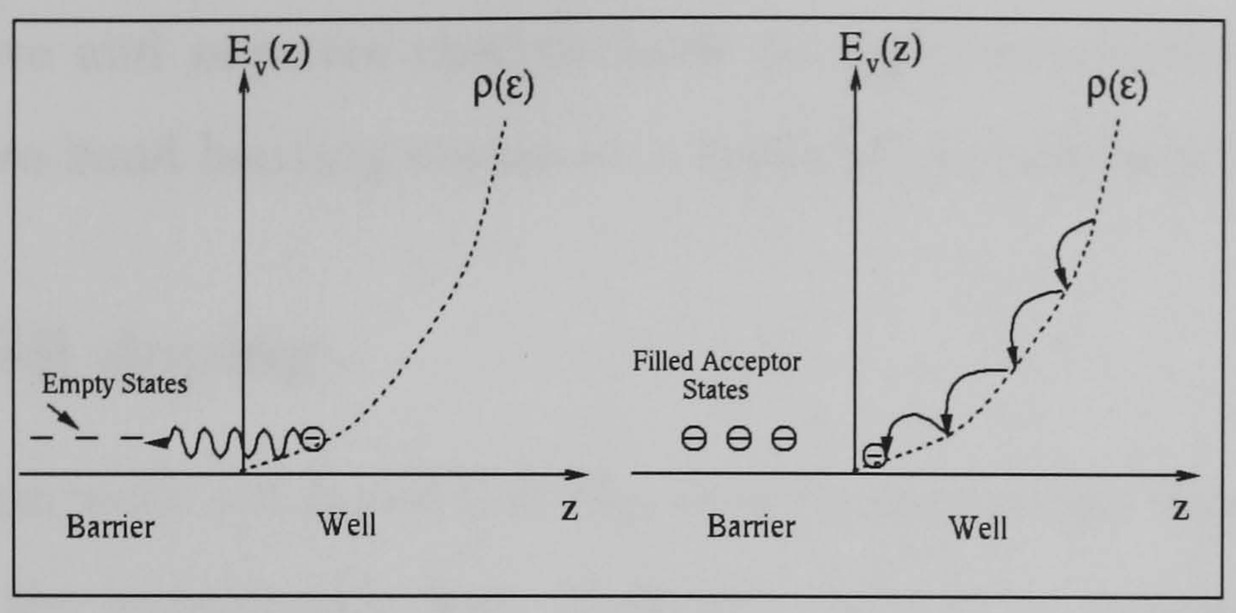


Figure 5.4: *Microscopic picture of barrier doping showing irreversible charge transfer from the acceptors to the well.*

electrons drop down to fill the lower energy states resulting in holes at the top of the well).

The reverse processes, i.e. the capture of electrons by impurities, requires electrons to absorb many phonons to increase its energy. This increase in energy is impossible at $T = 0$ and difficult at room temperature as $V_b - R$ is of the order of hundreds of meV's whereas kT$\approx$ 25meV. Thus by doping only the barrier a spontaneous and irreversible charge transfer to the well acting material is induced. Carriers and parent acceptors are separated which leads to band banding due to dipoles formed by the positive donors negative carriers.

Due to the presence of free carriers, electrical conduction can take place at $T = 0$ in the layer plane, in contrast to the z-direction where current flow is very difficult due to the size quantization of the electron z-motion. There is a great improvement in mobility of carriers in the well due to the spatial separation of carriers and acceptors. This separation is further enhanced by only doping the central part of the barrier, i.e. using spacer layers, thus further increasing the separation of the quasi bi-dimensional gas and the free carriers.

There are disadvantages to doping the barriers, however. The separation

of the positive and negative charges leads to high electric fields across the interface. Also band bending occurs as a result of the non-neutral system.

5.4.2 Well doping

If the quantum wells are doped directly, then the advantage of separating the carriers and the impurities is lost. However, there is no problem with band bending or the high electric fields. This is the method prefered for practical devices.

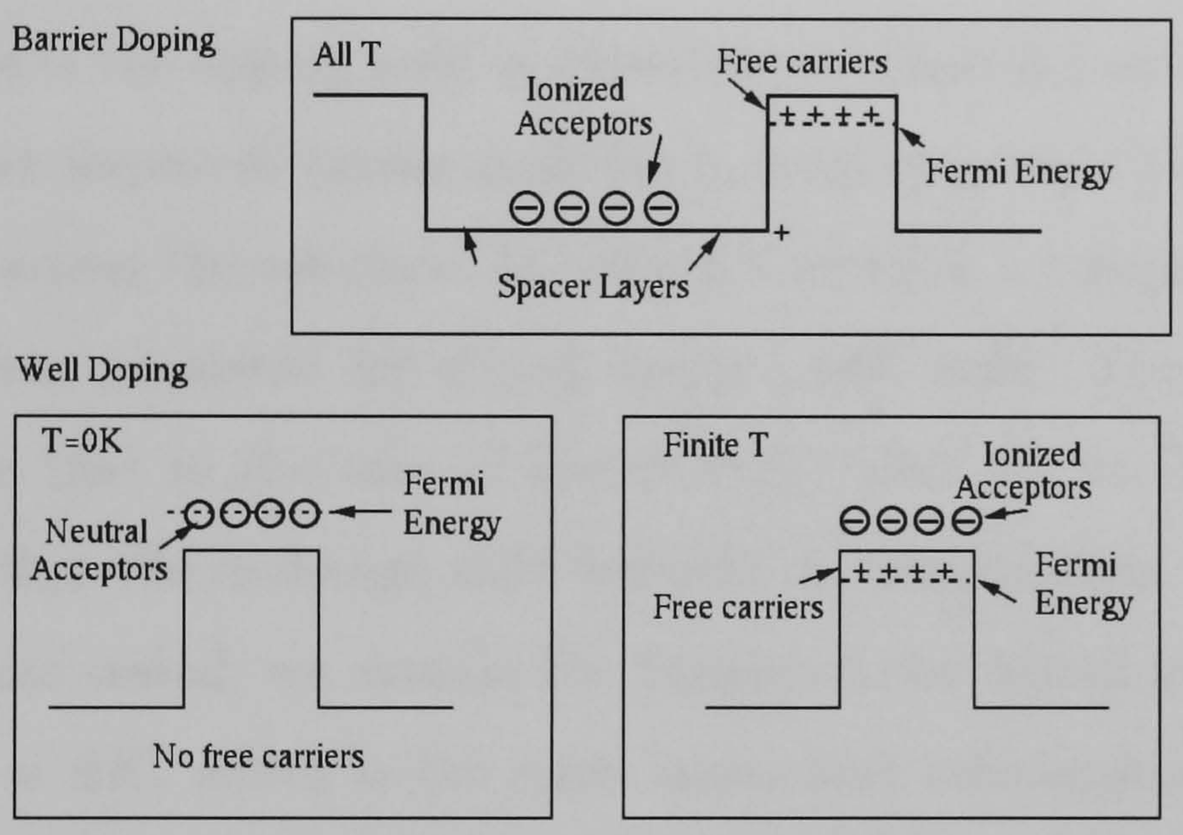


Figure 5.5: *Schematic representation of the potential profile for doped barriers and doped wells at 0K and finite temperature.*

The potential profile for a doped well is shown in figure 5.5. The acceptors provide an energy level in the forbidden gap at approximately 30meV (for boron in 15% Ge alloy). At 0K this level will be filled with its own electrons and no free charge carriers will be available. As the temperature rises, the acceptors be ionized, i.e loose their holes, or be filled by electrons. These holes will 'fall' into the well to become the free carriers. The acceptor level is considered to

be 30meV above the quantum well band edge, rather than the bulk band edge, by analogy with the bulk case, the ionization energy is considered to be the amount of energy taken to release the hole into the semiconductor band, which in this case is the top of the valence band. This means that the exchange shift is now a function of temperature, as the exchange is calculated based on the number of free carriers in the well. For low temperatures the acceptors will not be fully ionized and thus the number of free carriers will be lower than for the barrier doping situation where it is assumed that all carriers are ionized at all temperatures.

The nature of the doping used in experimental cases depends on the application, whether improved carrier mobility is more important that minimizing electric fields across the interface. In table 5.4 we show a comparison between the Fermi levels calculated for doped barriers and wells. The most obvious differences are that in the case of doped wells there are no free carriers at $T = 0K$ and that the exchange shift depends on temperature. In this thesis, unless otherwise stated, we assume the barriers to be doped as this provides carriers at $T = 0K$, which is the most convenient temperature at which to deduce the microscopic contributions to the absorption spectra. However, we continue to ignore the direct interaction which is a less good approximation for barrier than well doping, and band bending effects are not considered.

5.5 Fermi Levels

The effective mass approximation is adequate for determining the sampling zone but can lead to uncertainties in the absolute number of carriers which may not remain constant for all temperatures. Thus a theoretical variation in doping concentration with temperature will result which is not physically

correct. The number of carriers is constant depending only on the level of doping used. To avoid this problem we recalculate E_F after the sampling has been done. We use the relation

$$\int_{-\infty}^{+\infty} f(E)D(E)dE = N_A \tag{5.13}$$

where $D(E)$ is the density of states and N_A is the doping concentration. For the case of doped barriers even at low temperatures all our dopants will be ionized. To calculated the Fermi energy for doped barriers the expression

$$(no.\ of\ acceptors\ -\ no.\ of\ carriers)^2 \Rightarrow (N_A - \sum_{states} \frac{1}{1+e^{(E_F-E)/kT}})^2 \tag{5.14}$$

is minimized with respect to E_F. The Fermi levels calculated by this method are given in table 5.3. We can see that E_F increases with temperature and quickly moves above the top of the valence band. This is to be expected because of a sufficiently high temperature thermally excited carriers will be created and therefore the Fermi level must lie above the bottom of the conduction band.

For doped wells the situation is slightly more complicated. The total number of carriers present in the system must remain constant, but the number of free carriers, i.e. those holes which have been ionized from their parent acceptors will vary with temperature. Thus equation 5.14 becomes

$$(no.\ ofacceptors\ -\ no.\ of\ carriers)^2 \Rightarrow (N_A f(E) - \sum_{states} \frac{1}{1+e^{(E_F-E)/kT}})^2 \tag{5.15}$$

where $f(E)$ is the Fermi function associated with the acceptor levels. The exchange shift is then calculated for the number of ionized carriers, and thus is dependent on temperature.

	Barrier Doping		Well Doping	
Temperature	Fermi Level	Exchange Shift	Fermi Level	Exchange Shift
0K	-8.6	13.1	0	0
77K	13.6	13.1	12.5	6.1
150K	28.8	13.1	22.7	10.3
225K	48.6	13.1	40.5	11.6
298K	69.7	13.1	60.8	12.1

Table 5.2: *Fermi levels and exchange shifts calculated for a 50% Ge, 20 monolayers well for barrier and well doping for a 2-D doping concentration of 4.2 × $10^{15}m^{-2}$. The zero of energy is taken to be the top of the HH1 miniband at* Γ. *All values are in meV.*

	Barrier Doping		Well Doping	
Temperature	Fermi Level	Exchange Shift	Fermi Level	Exchange Shift
0K	-13.1	21.3	0	0
77K	-11.0	21.3	11.8	2.5
150K	-1.3	21.3	22.7	10.3
225K	12.3	21.3	40.5	11.6
298K	28.1	21.3	60.8	12.1

Table 5.3: *Fermi levels and exchange shifts calculated for a 20% Ge, 24 monolayers well for barrier and well doping for a 2-D doping concentration of 1.32 × $10^{16}m^{-2}$. The zero of energy is taken to be the top of the HH1 miniband at* Γ. *All values are in meV.*

Chapter 6

Comparison of Calculated and Experimental Absorption Response

Our pseudo-potential method for calculating band structures is well established and has been applied with success to a wide range of systems (see, for example Shaw 1993, Gell 1986, Turton 1993 and references therein). However, there are a number of uncertainties which could cast doubt on the validity of our predictions. The pseudopotentials have been adjusted to reproduce the bulk characteristics of Si and Ge, but there is no experimental data available to allow us to check that the effective masses, miniband non-parabolicity and energies are correctly represented in our modelling. For these p-type structures the energy subspace of interest is only a few hundred meV. If the band structure and band edges are not accurate, any attempt to engineer band structures to particular requirements is pointless. With this in mind we compare our predictions for the absorption spectra with the experimental results of Fromherz (1994a,b,c and 1995) and Boucaud (1995a,b).

6.1 Parallel Incidence

6.1.1 The Role of the Damping Constant

Another area of uncertainty is a suitable value to use for the damping constant, $\Gamma_{a,b}$ which appears in the denominator of the susceptibility expressions. The damping constant should in theory by a function of temperature, the position in the zone at which the transition occurs and the minibands between which the transition occurs. To include all these features is impossible so we approximate it to a constant characteristic of the system in question.

Fromherz (1995) recently reported a series of experimental results of absorption spectra for a range of p-type structures involving alloys with Ge concentration in the wells of up to 50%. In order to understand the position and width of their optical peaks we recall that the theoretical lineshape consists of two parts. Firstly, the linewidth in the ideal system reflects the band structure. Secondly, there is the contribution due to inhomogeneous broadening due to well width fluctuation and other deviations from ideal systems. As for the band structure contribution, our calculations include contributions from all active areas of the Brillouin zone, thus some broadening will result because of the shape of the miniband dispersion, e.g. the variation in the separation of the uppermost heavy hole miniband, HH1, and HH2, the first excited heavy hole level, across the zone will give rise to transitions between the two at slightly different energies depending on the area of the zone in which the excitation occurs. Where the bands are flat, there is a high density of states and thus a large contribution to the absorption is likely from these regions. Where the bands are more curved the density of states will be lower and thus lower contributions to the absorption will result. Contributions to the line width of this type will be present in even perfect structures and is fully contained in

our pseudopotential band structure calculations.

There is another contribution to the lineshape which is a feature of the imperfect growth of these structures, inevitable in strained layers and particularly with high germanium concentrations.

6.1.2 Determination of a Suitable Damping Constant

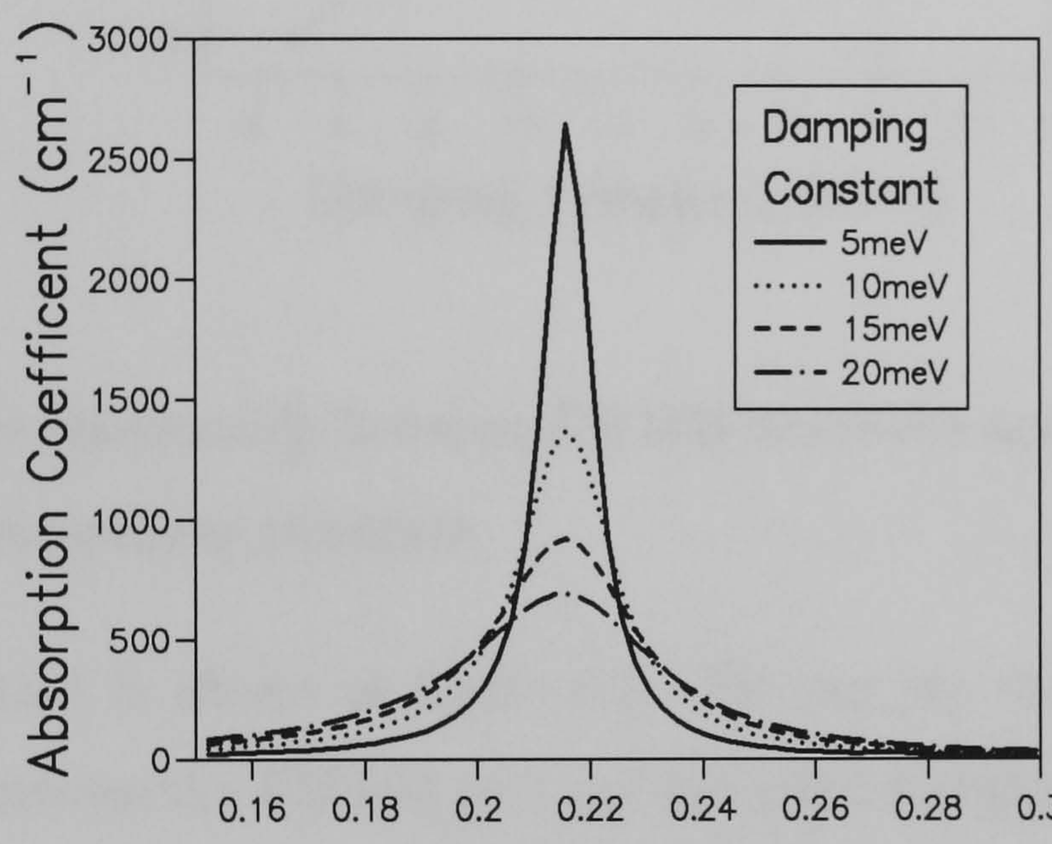


Figure 6.1: *The change in the absorption spectra as the damping constant is increase from 5meV to 20meV for a 50%, 20 monolayer structure.*

To make contact with experiment we choose the structure containing 27 Å wide $Si_{0.5}Ge_{0.5}$ alloy well for which optical lineshapes are available. In figure 6.1 we can see the change in the absorption spectra produced by varying the damping constant over a range of 5-20meV. The basic trend is for the magnitude of the peak to decrease and the width to increase as the damping constant increases. The peak position, however, is unchanged by changing the damping constant. The trend of the FWHM linewidth with increasing

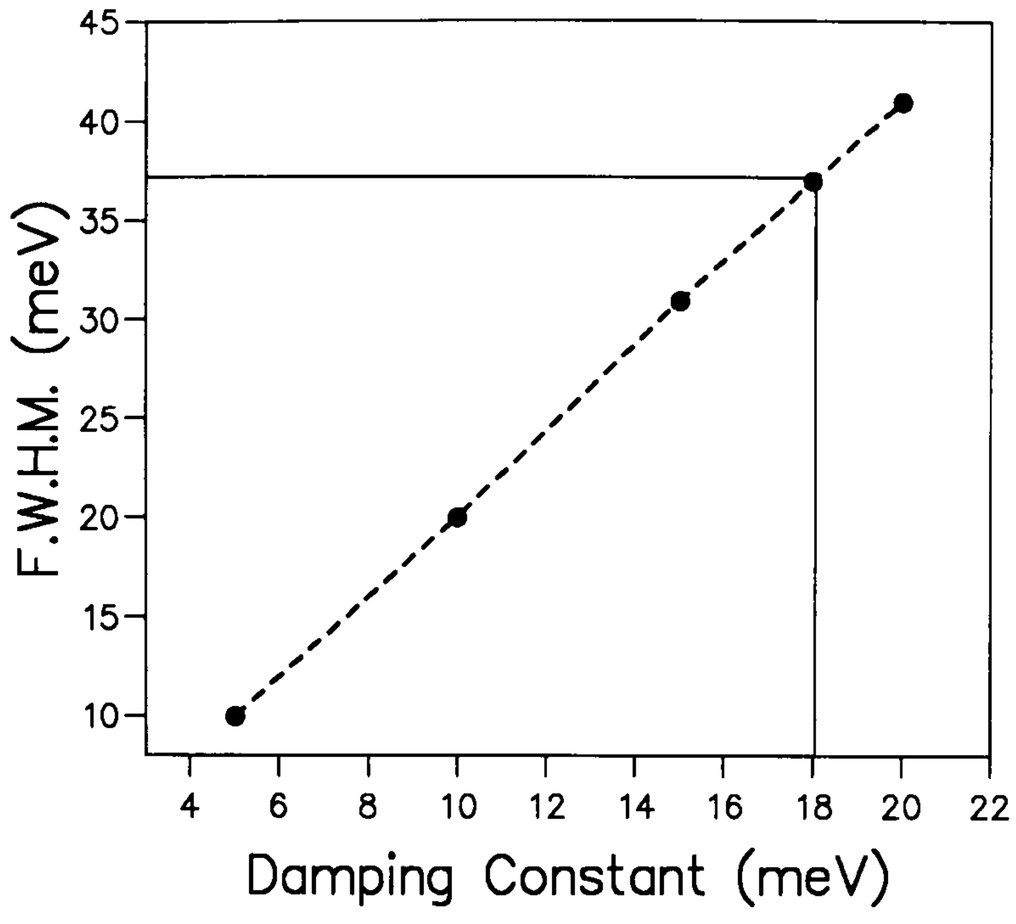


Figure 6.2: *The relationship between FWHM linewidth and damping constant for a 50%, 20 monolayer structure.*

damping constant is shown in figure 6.2. We can see that there is a linear relationship between the FWHM and the damping constant. From this graph we can see that a damping constant of 18meV best reproduces the experimental full width half maximum (FWHM) line width of 37meV. We then compare our calculated absorption spectra with a schematic representation of the results of Fromherz (1995) for the same structural parameters, As the magnitude of the absorption is extremely sensitive to the value used for the damping constant, for ease of comparison we normalize the magnitude of all the peaks to unity. These absorption curves were calculated at 12K, the temperature at which Fromherz (1995) conducted his experiments, with a barrier doping concentration of $7.0 \times 10^{17}\text{cm}^{-3}$ across 60 Å of the barriers. This corresponds to a 2-D doping concentration of $4.2 \times 10^{1}1\ \text{cm}^{-2}$. The experimental linewidth is broad, particularly when compared with that of GaAs structures which typically have linewidths of the order of 3-8meV. This wide peak is clearly a

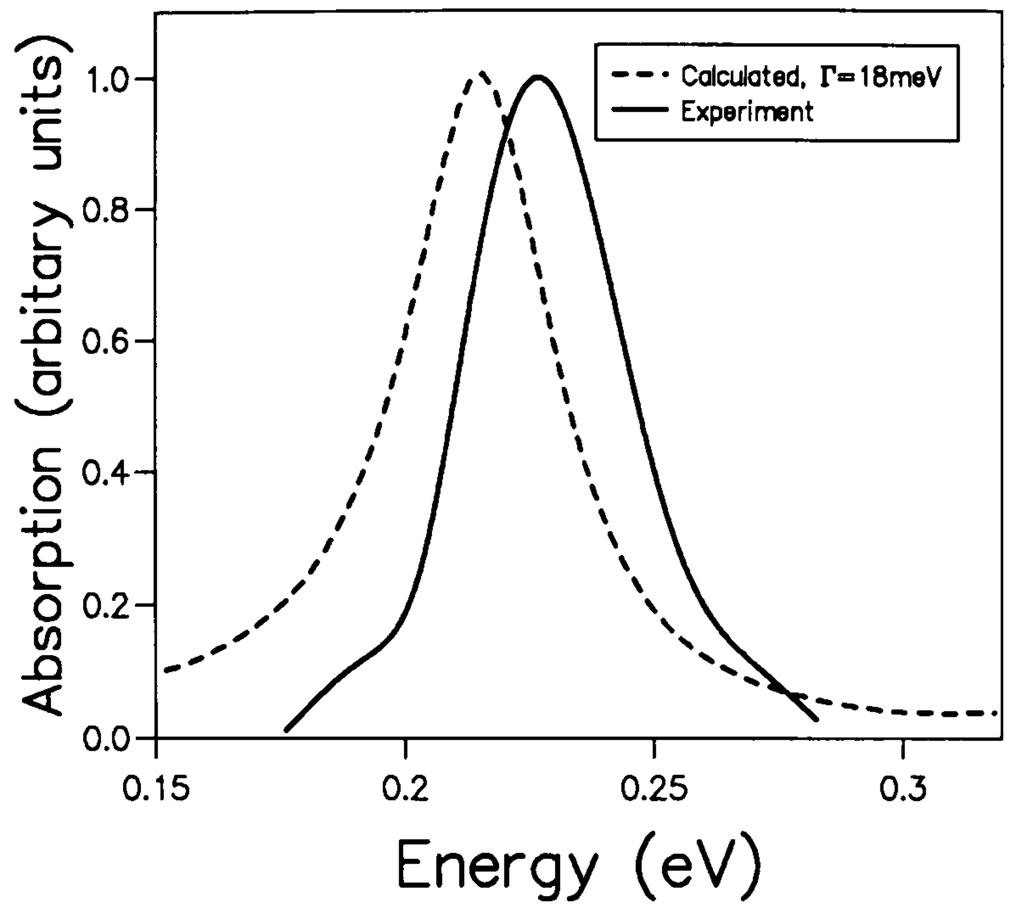


Figure 6.3: *The comparison of calculated curves and experimental results for a 50%, 20 monolayer structure. The magnitudes of all the peaks have been normalized to the same value to ease comparison.*

result of the problems associated with the growth of these strained structures.

6.1.3 Comparison with Experimental Results of Fromherz et al.

Further comparisons were made for a 30% Ge alloy well 44 monolayers wide well and a 21% Ge alloy well 36 monolayers wide well. A doping concentration of 2.2×10^{18} cm^{-3} was used for these structures.

The comparisons using a range of damping constants for the 30% case is shown in figure 6.4. In this case the a damping constant 15meV best reproduces the FWHM line width of 37meV. Figure 6.5 shows the comparison of these two curves more clearly. In figure 6.6 we show the final comparison for the 21% case. For this structure a damping constant of 14meV is needed to mimic

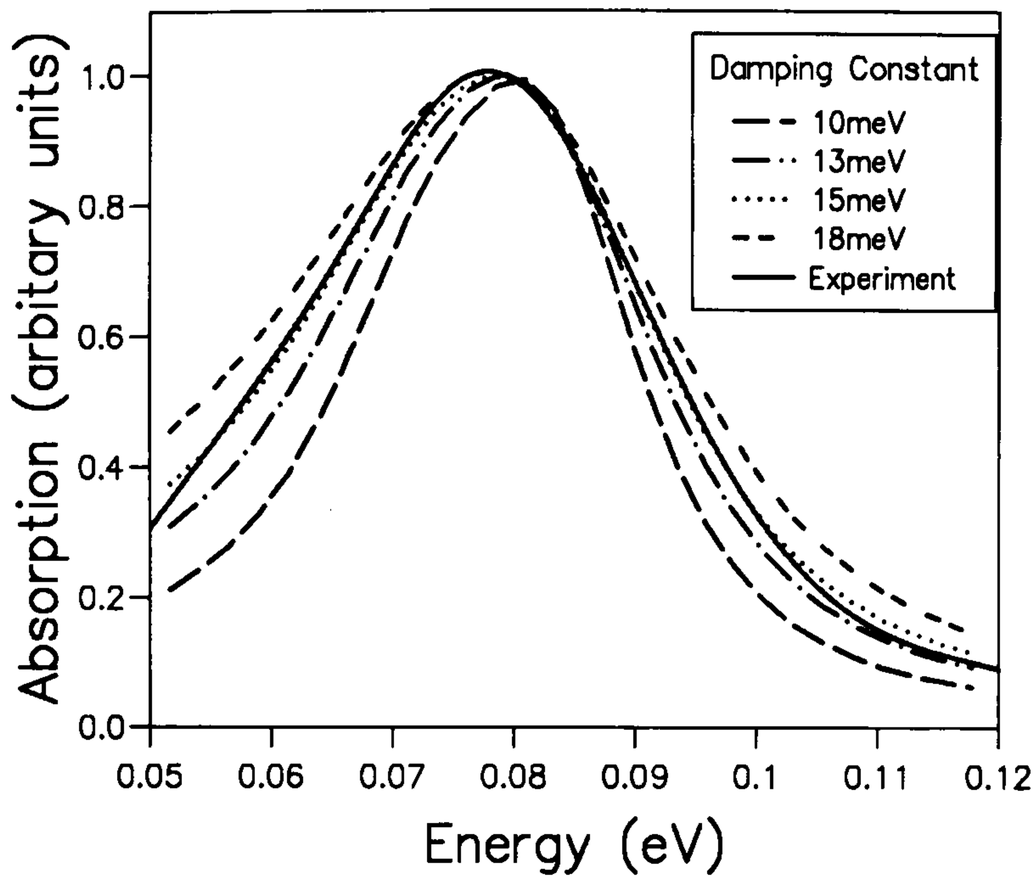


Figure 6.4: *Comparison of calculated curves with a range of damping constants for a 30%, 44 monolayer well, with the experimental results of Fromherz et al. The magnitudes of the peaks have been normalized for ease of comparison*

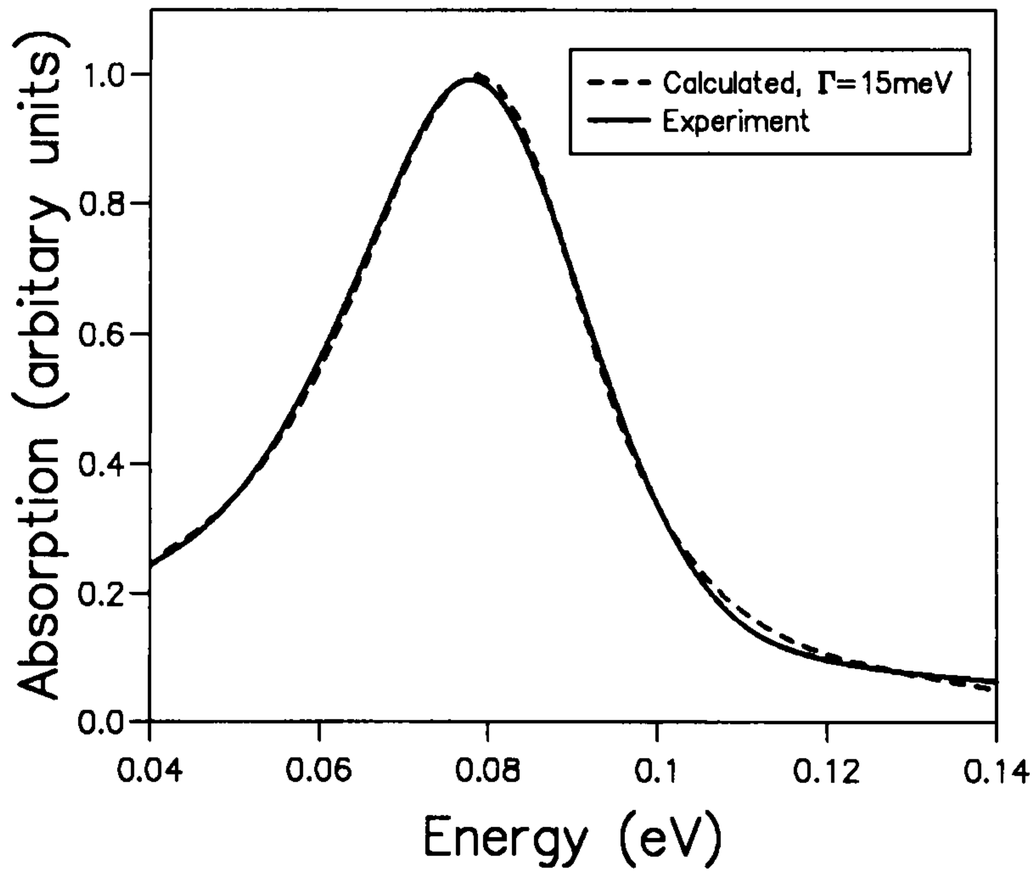


Figure 6.5: *Comparison of the experimental and calculated results for 30%, 44 monolayer well, using the best fit for the damping constant of* Γ*=15meV. The magnitudes of the peak have been normalized for ease of comparison.*

the experimental line width.

The agreement between calculated and experimentally observed peak positions is good in all cases, a difference of 2meV for the 30% case, 10meV for 50% and 9meV for the 21% structure. These results are very encouraging.

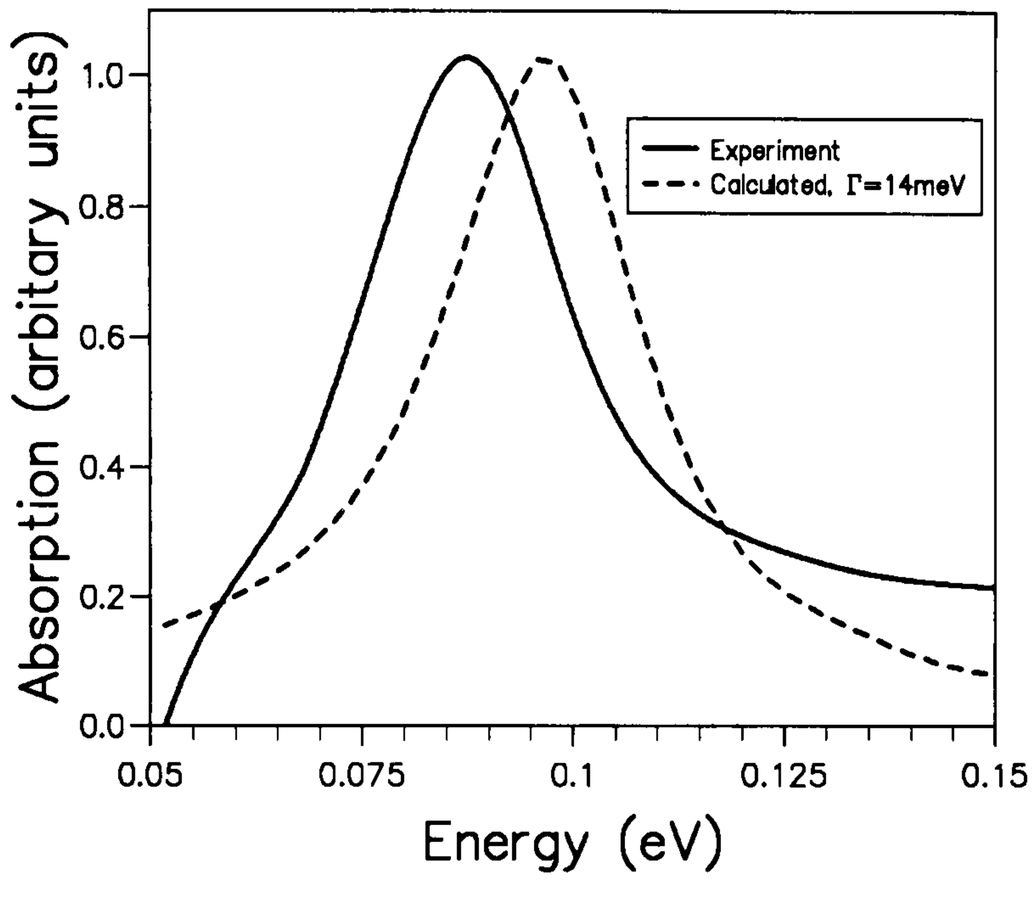


Figure 6.6: *Comparison of calculated and experimental results for a 21%, 36 monolayer well, showing that a damping constant of 15meV best reproduces the experimental line width*

We note that although the experimental FWHM line widths are very similar we needed different Γ's to fit them. This can be attributed to a greater degree of band structure width in the 30% case, i.e. more variation in the separation of the HH1 nd HH2 minibands than in the 50% case, thus the peak is broader for any given Γ. Details of band structure effects will be discussed in the following chapter.

The discrepancy between calculated and experimental results is largest for the 50% structure. This is to be expected as the the high germanium concentration makes this structure most susceptible to growth problems such as grading of the germanium concentration across the width of the well, varia-

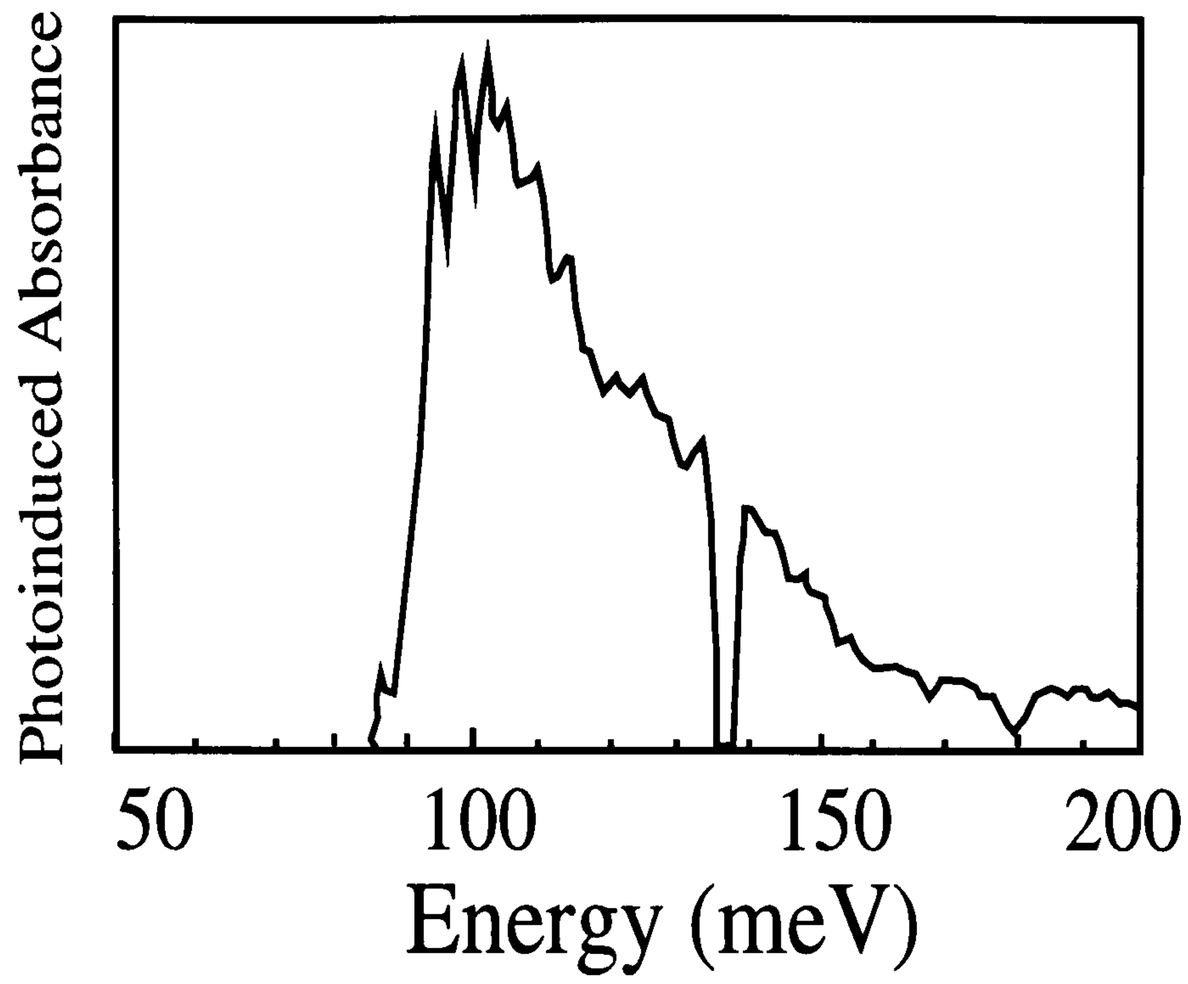
tion from well to well, interface diffusion etc. Also the doping concentration is unlikely to be known with absolute precision. Our crude method of including the exchange shift means that the predicted peak position is very sensitive to the doping level. An uncertainty of a factor of 2 in the doping concentration would change our exchange shifts for the 50% structure by 5meV. For the 21% structure where a higher doping concentration was used a similar uncertainty in the doping concentration i.e. the difference between a doping concentration of $2.2\times10^{18}\mathrm{m}^{-2}$ over 60 Å of the barriers and $1.1\times10^{18}\mathrm{m}^{-2}$ over 60 Å of the barriers would lead to a change in the exchange shift of of 6meV which would account for the majority of the discrepancy between the two peaks.

Also it must be remembered that our calculations are performed for perfect structures with no account taken of imperfections at the interface or non-ideal growth characteristics of any sort. We have not included many body effects such as depolarization or the direct interaction which may be expected to be significant in heavily doped structures.

6.1.4 Comparison with Experimental Results of Boucaud et al.

In figure 6.7 (a) we reproduce experimental results from Boucaud (1995). The experimental results are based on a 20% germanium concentration, with a 30Å well. In these experiments the wells were doped, with a doping concentration of $4.2\times10^{18}\mathrm{cm}^{-3}$. We calculate that this results in a Fermi energy of 18meV above the HH1 miniband for 77K and an exchange shift of 5meV. The 30Å well used by Boucaud (1995) is difficult for us to model as it corresponds to 22 monolayers which is not an integral number of lattice constants. Therefore we performed calculation with a 24 monolayer well. In figure 6.7 we see the change in the absorption spectra with damping constant for a 20%. 24

(a) Experimental (Boucaud et al.)



(b) Calculated

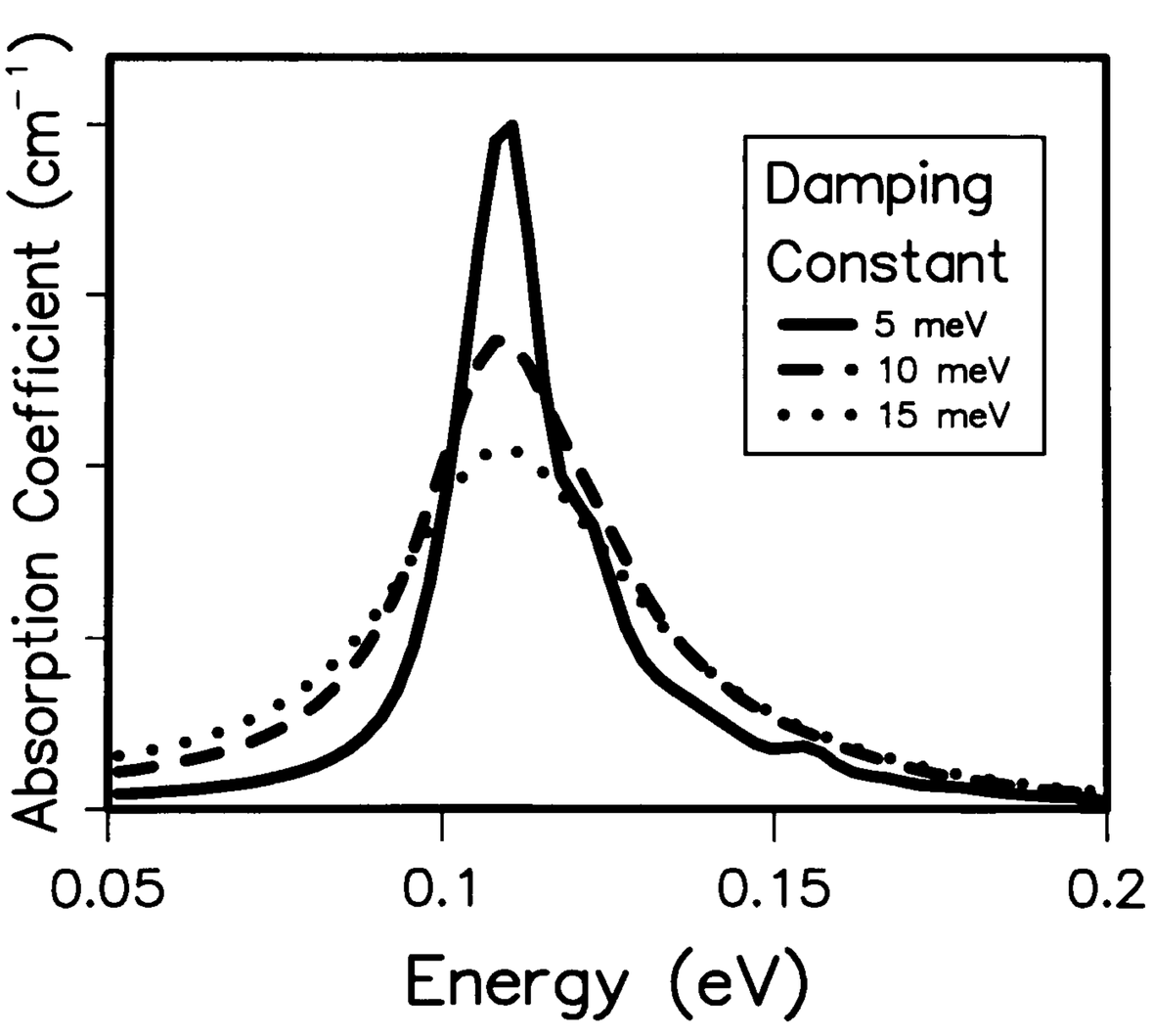


Figure 6.7: *The experimental results of Boucaud et al. for a 20%, 30 Å structure (a) compared with our calculated results for a 20%, 24 monolayer structure with a range of damping constants (b).*

monolayer structure. Again the magnitude of all the curves has been normalized to ease comparison. From this we see that the damping constant which best reproduces the experimental value of 35meV is 10meV. The experimental linewidth for these results is approximately the same as that of Fromherz et al., however, a smaller damping constant is required to reproduce this linewidth in the calculated curve. This is a feature of the band structure. The miniband dispersion for this and other closely related structures is presented in the next chapter where the structures are discussed in depth.

Our calculated absorption coefficients for these well doped structures are extremely low. This is because only a small percentage of the acceptors are ionized in our model for the Fermi levels. At 77K only 4% of the acceptors are ionized leading to a very small free carrier concentration and consequently a very low absorption coefficient. Our assumption that the acceptors all lie at a certain energy above the superlattice valence band edge, with no inclusion of a distribution for the acceptor energy levels may be to blame for this problem. We can again see that our agreement with experimental results is good. We can see that our predicted peak position is is at a slightly higher energy than the experimental results. Also there is more of a tail in the experimental results at higher energy.

6.2 Normal Incidence

Boucaud et al. (1995) also observed normal incidence response in this system and the experimental curves are reproduced in figure 6.8. The normal incidence response of this and other structures is discussed in more detail in chapter 8. In this section we merely compare with the experimental results with the aim of finding a suitable value for the damping constant to use in the other sections.

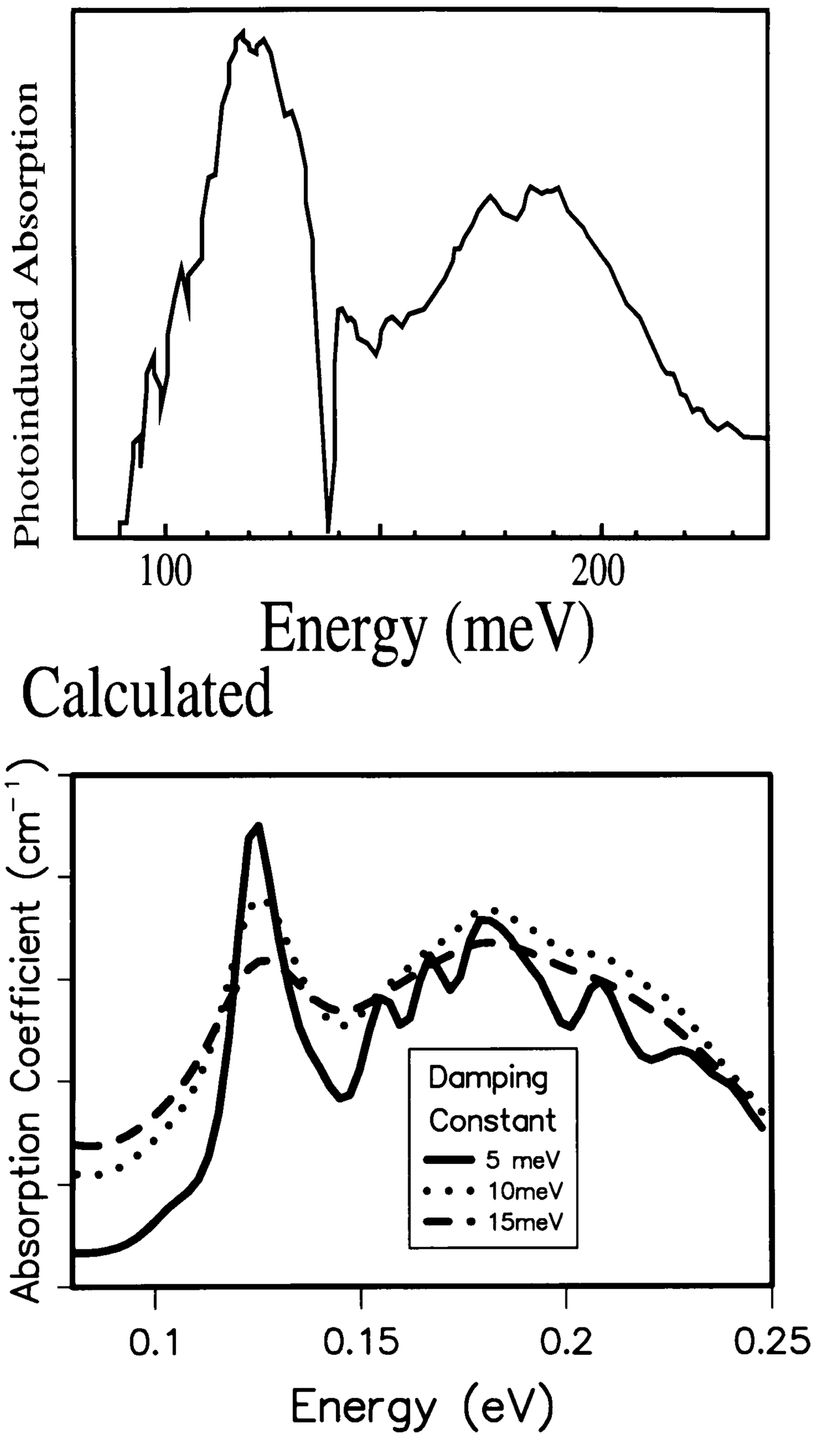


Figure 6.8: *The experimental normal incidence response for a 20%, 30 Å well from Boucaud et al. (a), compared with calculated results for 20%, 24 monolayers (b).*

We can see from this comparison that we predict not only the peak of the absorption to within a few meV, but also the line shape is in good agreement with the higher energy broad multiple peak in evidence in both calculated and experimental curves. Increasing the damping constant for this normal incidence response smoothes over the peaks at higher energy. The presence of these multiple peaks in the experimental data is an indication of the high quality of the samples. The 5meV damping constant curve best reproduces the actual profile of the absorption response, with the high energy peaks of a lower magnitude than the peak at $\approx$ 120meV, whereas the 10meV case best reproduces the experimental line width.

In conclusion, we have shown good agreement between our calculated absorption response and experimental curves for range of germanium concentrations in the well, well widths and for both polarizations of light. In the next chapter we consider the absorption spectra in more detail and discuss the microscopic origin of the peaks.

Chapter 7

Results of Parallel Incidence Absorption Calculations

In this chapter we present the results of first order susceptibility calculations for a range of well widths and germanium concentrations in the alloy well. We first consider the parallel incidence response of structures consisting of $Si_{0.5}Ge_{0.5}$ wells. These are not the best candidates for IR detection as the difficulties in growth increase with high germanium concentration and the excited holes are still confined quite low in the wells which would require high electric fields before they could tunnel out of the wells. However, they are the easiest to analyze theoretically as the important minibands do not cross in the region of interest, therefore attributing the microscopic origin to the produced absorption spectra is straightforward. We then move on to consider structures with lower germanium concentrations and can apply the general principles deduced for the higher Ge concentrations to these more complicated situations.

7.1 $Si_{0.5}Ge_{0.5}$/Si Structures

We begin by considering a structure consisting of a strained 20 monolayer wide well containing $Si_{0.5}Ge_{0.5}$ alloy sandwiched between unstrained 80 monolayer Si barriers, hereafter referred to as 50%, 20 monolayers. The potential profile for this structure is shown schematically in figure 7.1.

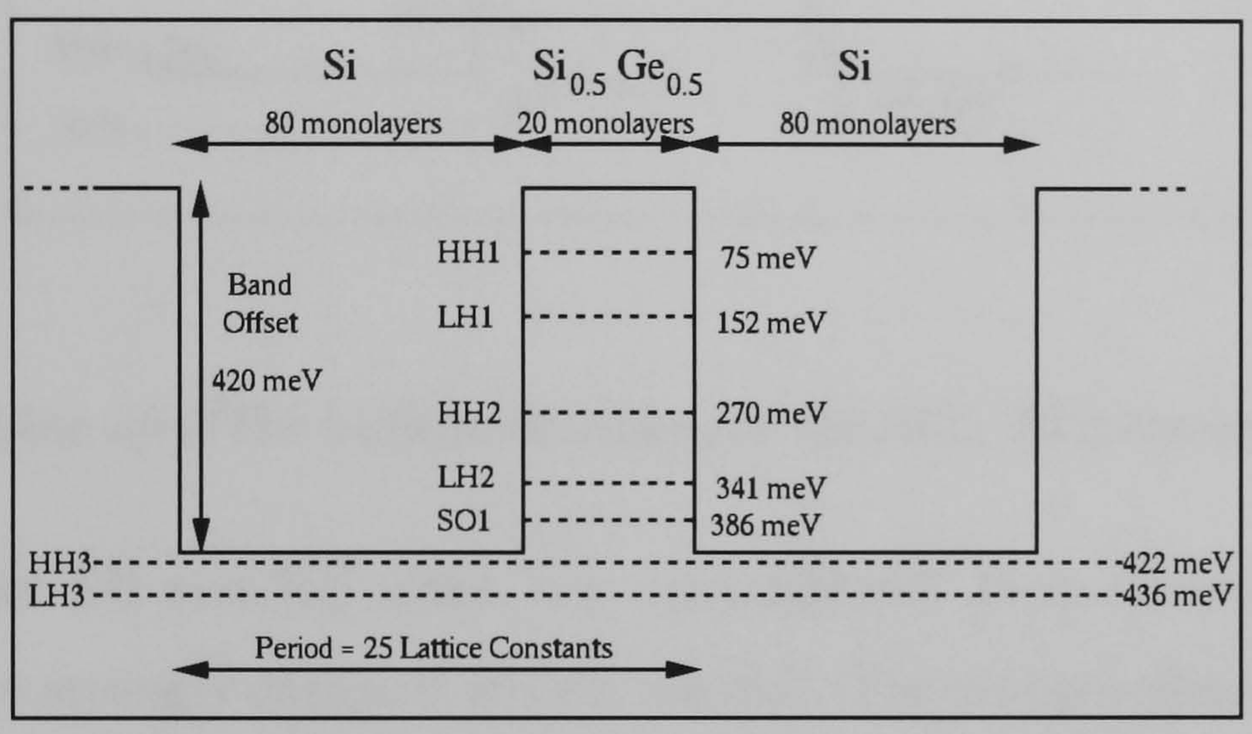


Figure 7.1: *Schematic representation of a 50%, 20 monolayer structure. The position of the top seven bands at the zone centre are indicated. The zero of energy is taken to be the top of the well.*

The valence band offset for this structure is 420 meV obtained from the linear interpolation of the results of Van de Walle (1986). The energies are labelled from the top of the bulk band edges which are considered to form the well. We can see the position of these band edges in figure (7.2). The well as 'seen' by the heavy and light hole bands is 420meV deep. However, the SO1 miniband 'sees' an offset of only 100meV although the SO well is shifted in energy by 363 meV below the HH1 band edge. We calculated the charge densities at the zone centre by simply squaring the eigenfunctions. The charge densities for this structure are presented in figure 7.3.

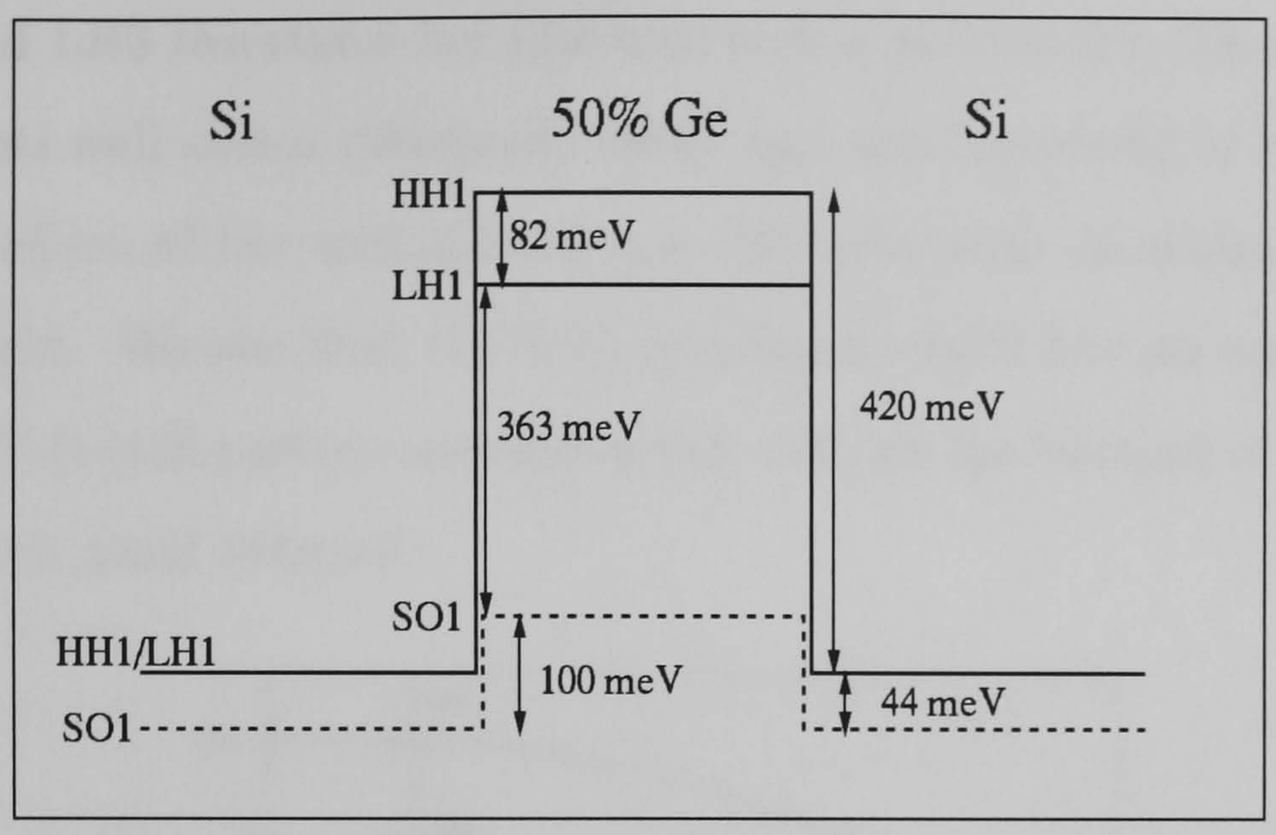


Figure 7.2: *Line up of the bulk band edges for the 50%, 20 monolayer structure.*

We see for LH and HH states less that 420meV from the top of the well the states are strongly confined within the well. For energies greater than this,

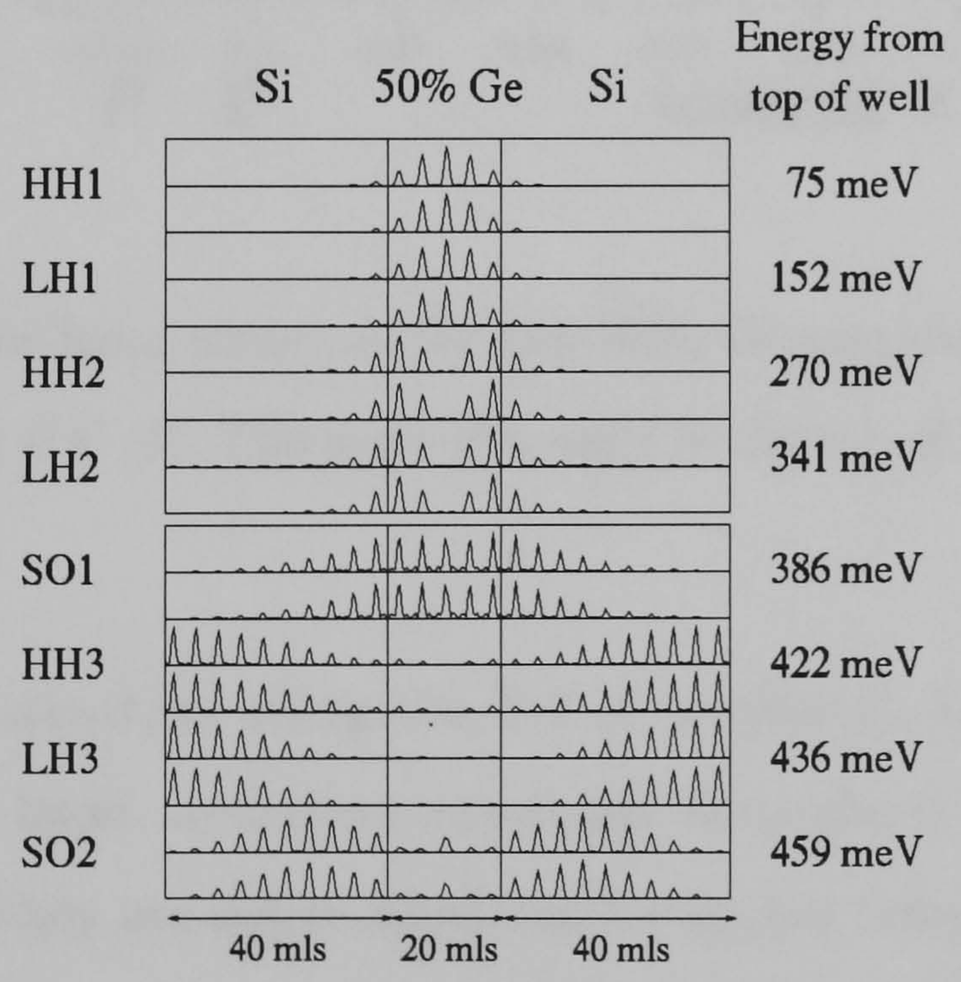


Figure 7.3: *The charge densities for the top eight states of the 50%, 20 monolayer structure. The zero of energy is the top of the well.*

e.g. HH3 and LH3 the states become confined in the barrier. These states are known as anti-well states (Bastard, 1988) and are the result of the quantum mechanical effect of the well influencing the behaviour of states even above the well height. We see that the SO2 miniband which has an energy greater than 420meV is still partly confined in the well, as the bottom of the SO well does not occur until 464meV.

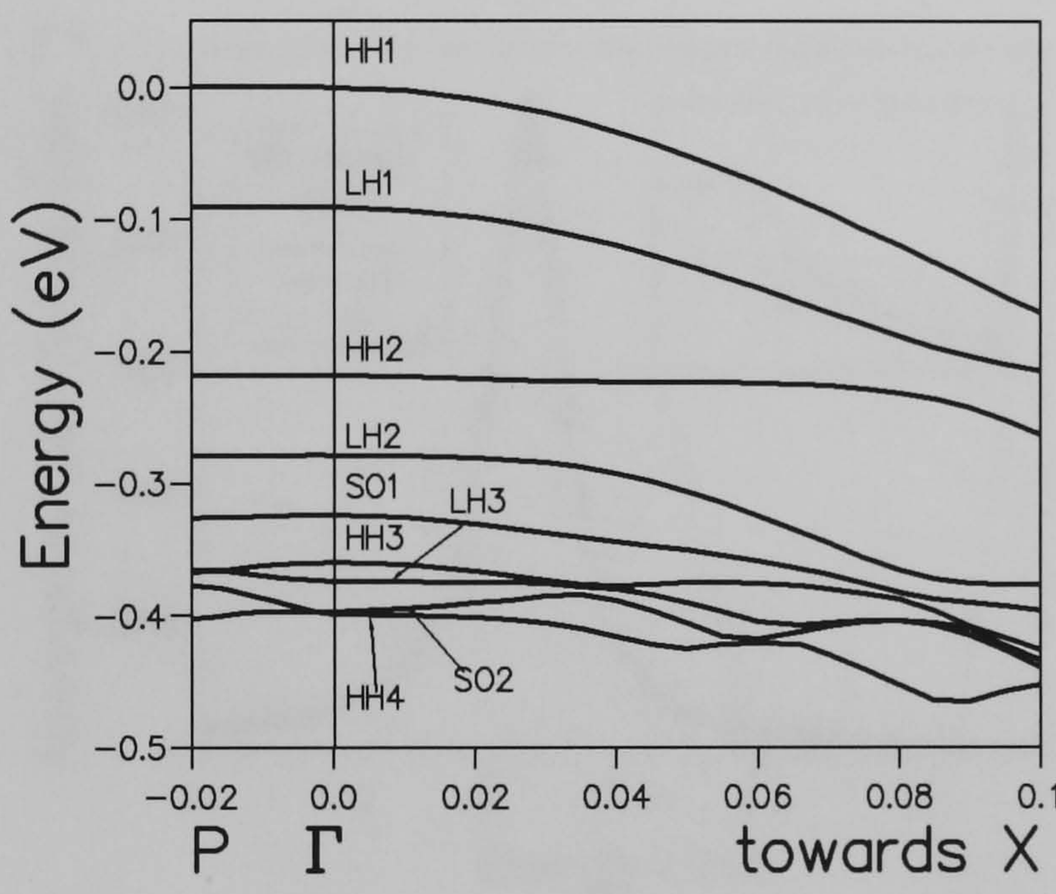


Figure 7.4: *The miniband structure for the 50%, 20 monolayer structure along the symmetry axes P-Γ -X. The zero of energy is the top of the HH1 miniband at the zone centre.*

The miniband structure along the P-Γ-X symmetry lines is presented in figure (7.4). The band structures presented throughout this thesis are all calculated at 0K. They are not recalculated for higher temperature even when absorption calculations are performed at higher temperatures, as this would be prohibitively time consuming. The main effect on increasing temperature is to reduce the fundamental band gap but the separation between the minibands remains relatively unchanged. This diagram includes a 13meV exchange shift

corresponding to a doping concentration of $7.0\times10^{17}\text{cm}^{-3}$ across 60 Å of the barriers which is a 2-D doping concentration of $4.2\times10^{15}\text{m}^{-2}$. This is the doping concentration used by Fromherz (1994) for a similar structure. We can see that the bands along the growth direction are almost flat indicating that there is no dispersion in the well. Along the Γ-X direction we can see distinct non-parabolic behaviour.

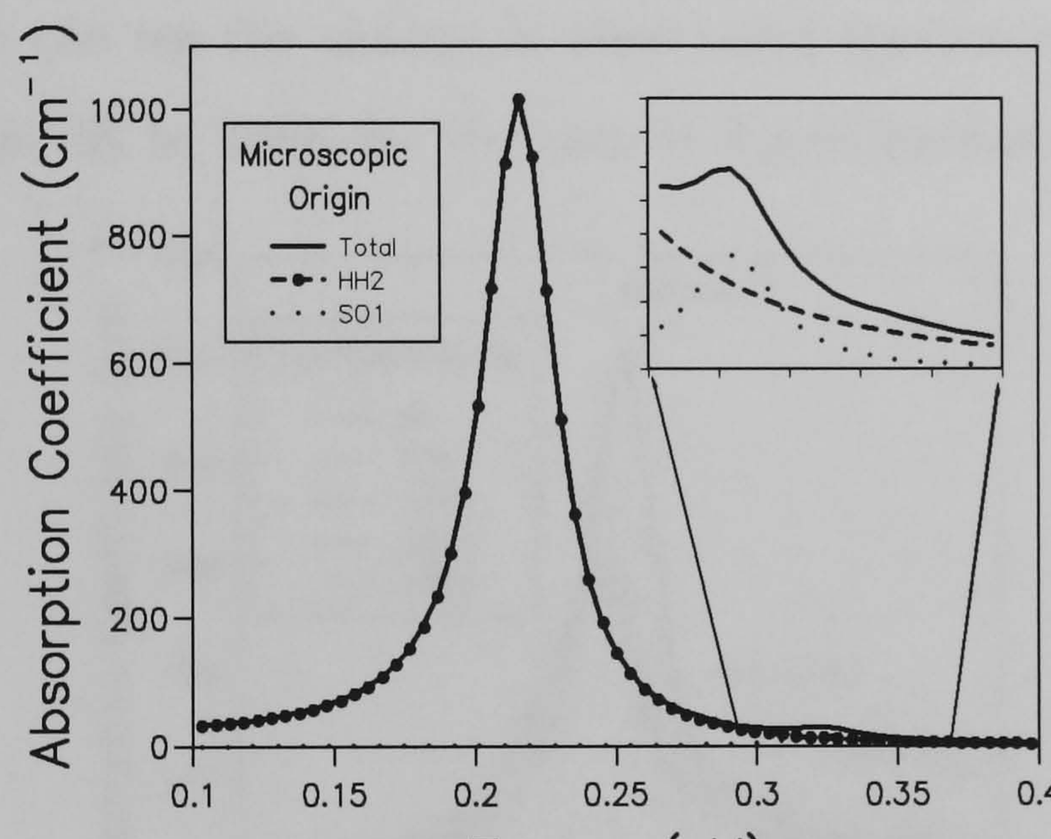


Figure 7.5: *The parallel incidence absorption response for a 50%, 20 monolayer structure calculated at 0K, showing the microscopic origin of the peaks.*

The calculated parallel incidence absorption response for this structure is presented in figure (7.5). This structure is calculated using a damping constant of 18meV. The reason for this choice is explained in Chapter 6. The microscopic origin, i.e. the levels between which the contributing transitions occur is indicated. We can see that the peak is produced by transitions between HH1 and HH2, as would be expected for light incident in the z-direction (see discussion in Chapter 3). We also note a very small peak between 300meV and 350meV. This region is expanded in the inset of the figure and we can

see that it is caused by transitions between SO1 and HH1. This transition is forbidden for parallel incident light in the absence of band mixing. Therefore this transition must be occuring away from the zone centre where the mixing between the minibands is greater.

7.1.1 Changing Temperature

In figure 7.6 we can see the change in absorption spectra as the temperature is increased from 0K to 298K for the case of doped barriers. We can see that

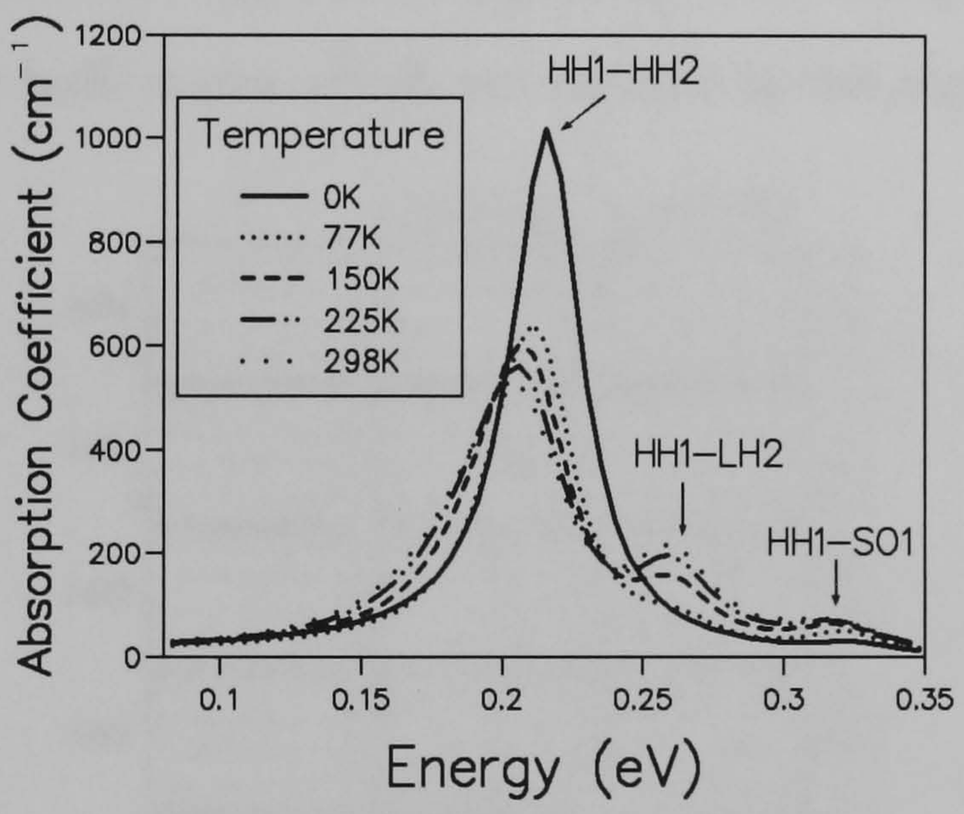


Figure 7.6: *The change in the absorption spectra with temperature for a 50%, 20 monolayer structure assuming the* barriers *to be doped with a 2-D doping concentration of* $4.2\times10^{15}\,cm^{-2}$.

the peak at $\approx$ 225meV due to HH1-HH2 transitions, which dominates at low temperatures is reduced in magnitude and moves to a slightly lower energy as the temperature increases. The peak at 325meV due to excitations between HH1-SO1 increases very slightly in magnitude and a new peak appears at $\approx$ 250meV which was not present on the 0K response which can be attributed to transition between HH1 and LH2 minibands. All these features can be

explained by considering the behaviour of the carriers as the temperature increases. At 0K 95% of the carriers occupy regions up to $\mathbf{k}=0.03\left(\frac{2\pi}{a}\right)$ away from the zone centre. As the temperature increases this volume is increased to $\mathbf{k}=0.08\left(\frac{2\pi}{a}\right)$. The distribution of carriers with temperature was demonstrated in the previous discussion on Fermi levels in Chapter 4. At these sites around the zone centre the band mixing is at minimum. As the temperature increases the carriers move away from the zone centre and start to occupy the more heavily mixed states. In figure (7.7) the coefficients for the contributions from different bulk states, the $A_{n\mathbf{k}}$'s from equation 3.10, are plotted showing the components of the bulk states which are present in the eigenfunctions.

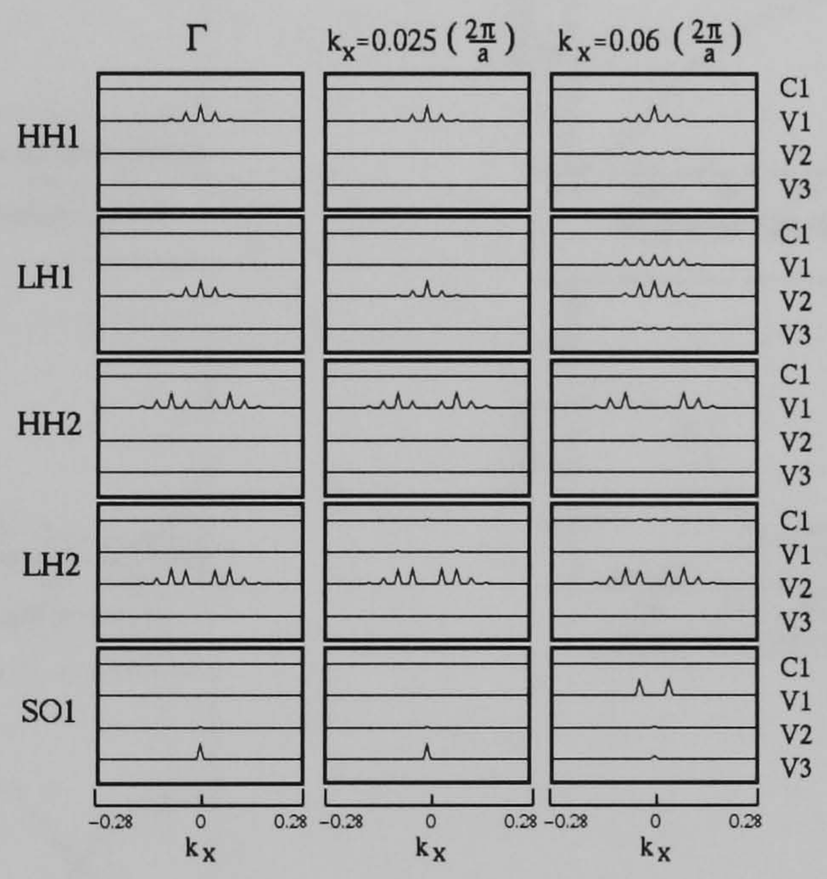


Figure 7.7: *Plot of the coefficients of the eigenfunctions of the top five states in the valence band for the 50%, 20 monolayers structure, showing contributions from states of different bulk momenta at selected points across the zone*

There are two distinct types of band mixing; mixing between states of different bulk wavevector, i.e. mixing due to zone folding in the quantum well; and bulk components of other states present in the quantum well miniband,

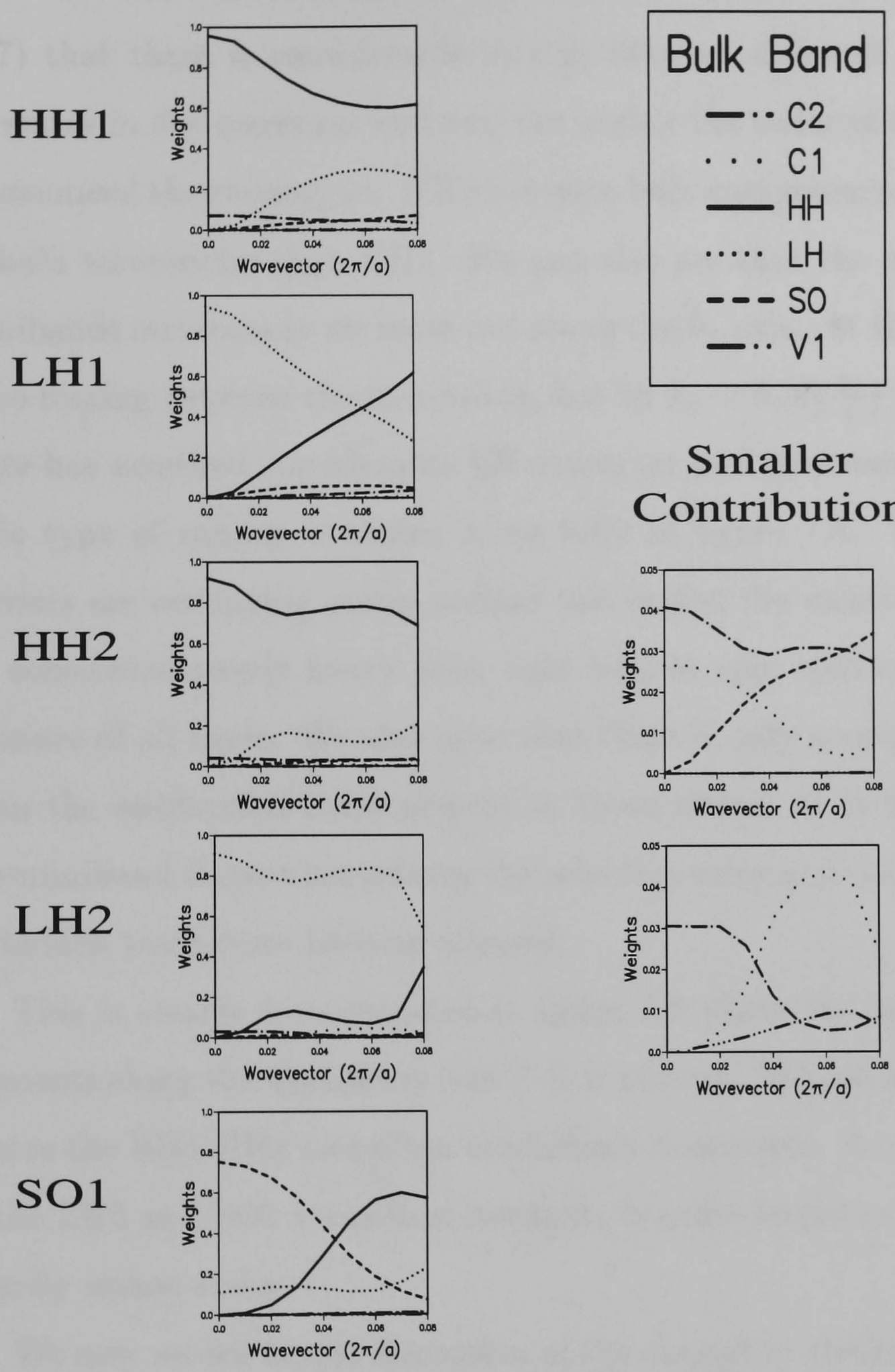


Figure 7.8: *Plot of the coefficients of the eigenfunctions of the top five states in the valence band for the 50%, 20 monolayers structure along the symmetry line* Γ-X, *showing the increase in mixing between the minibands along the* k_x *axis. For the HH2 and LH2 minibands an expanded section of the smaller contributions is shown.*

e.g. bulk LH contributions to the HH1 miniband.. We can see from figure (7.7) that there is considerable mixing between different bulk momenta for all states in the quantum well and the higher the order of the states the more pronounced the mixing, i.e. HH2 contains bulk components with a wider range of bulk wavevector than HH1. We can also see that the mixing between the minibands increases as we move out along the k_x axis. At the zone centre there is no mixing between the minibands, but by $k_x = 0.06(\frac{2\pi}{a})$ the first heavy hole state has acquired considerable LH character and vice versa for the LH state. This type of mixing is shown more fully in figure 7.8. This means that if carriers are occupying states around this region the minibands can no longer be considered purely heavy hole, light hole or spin split off bands, but are a mixture of all three. We also note that there is only a small degree of mixing from the conduction band present in these states. It is the mixing between the miniband states that relaxes the selection rules and means that previously forbidden transitions become allowed.

This is clearly demonstrated in figure 7.9 where the square of the matrix elements along the symmetry line Γ-X is plotted. We can see that at the zone centre the HH1-HH2 transition completely dominates. As we move away from Γ the LH2 and SO1 transition strengths become larger as these are the more heavily mixed states.

We now return to our discussion of the change in absorption spectra as the temperature increases shown in figure 7.6. The reduction in the magnitude of the main peak is due to a lower number of carriers present at the zone centre (we assume that carrier concentration remains constant at all temperatures for doped barrier systems) and more carriers occupying states where the strength of the excitation between HH1 and HH2 is reduced, therefore less contribution to this peak arises. The magnitude of the smaller peak due to LH2-HH1

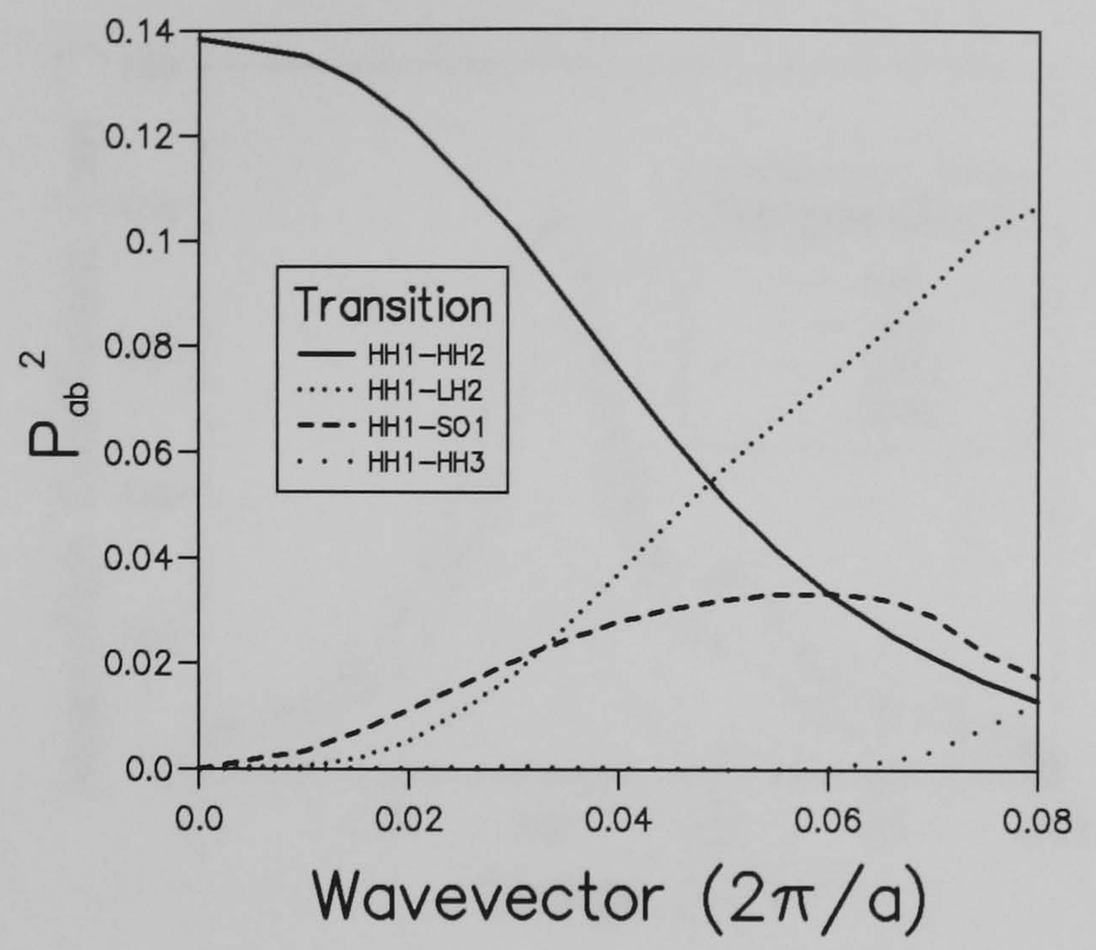


Figure 7.9: *The variation of the strength of the matrix elements for parallel incident light for the 50%, 20 monolayer structure along the symmetry line* Γ-X

transitions increases as the temperature rises and the carriers occupy more mixed states, and also moves to a slightly lower energy because, as can be seen from the band structure diagram 7.4 the separation between the LH2 and HH1 minibands decreases away from the zone centre.

For doped wells the change of the absorption spectra with temperature is shown in figure 7.10. We can see the same trends as mentioned for the barrier doped case, but the magnitude of the absorption increases with temperature as the number of free carriers increases as the temperature rises.

The carriers moving to large $k_{||}$ where the HH1 state has acquired some LH character means that the HH1-LH2 transition is no longer forbidden. This transition will always be weaker than the HH1-HH2 as its magnitude is proportional to the amount of bulk HH components in the LH2 miniband.

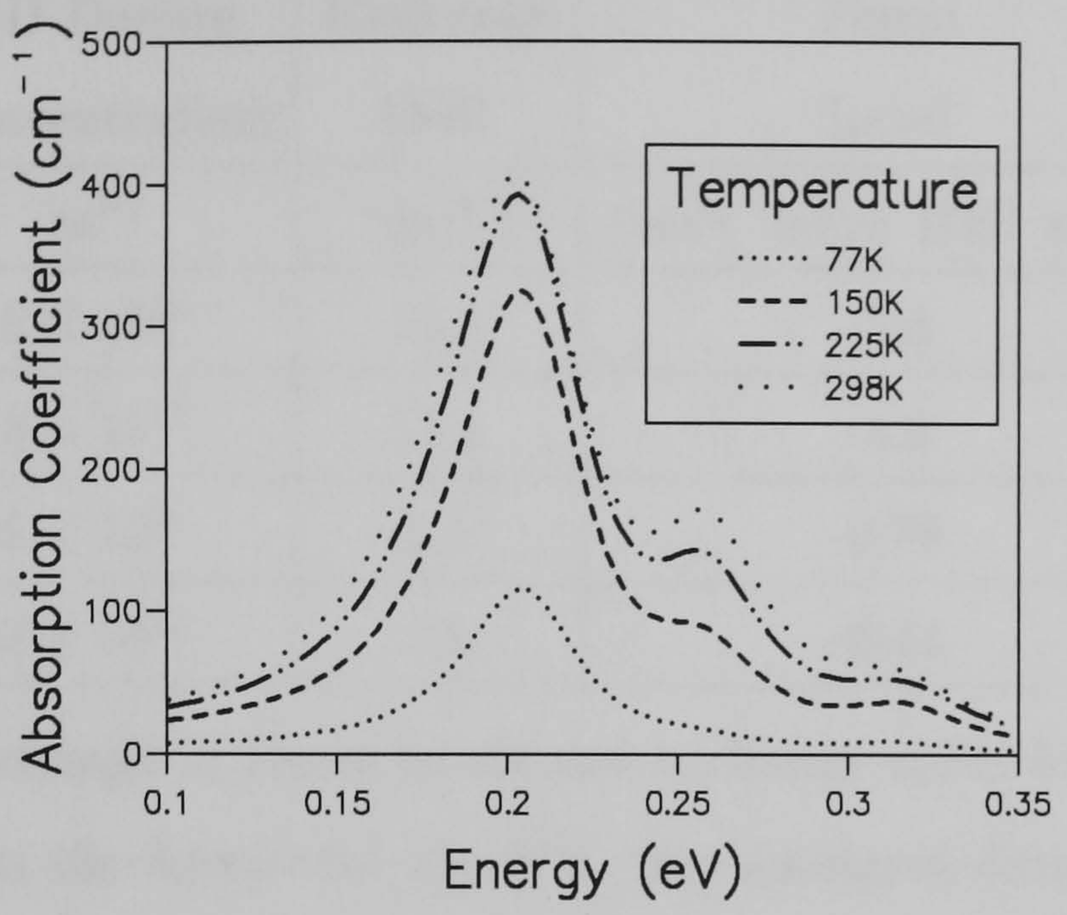


Figure 7.10: *The change in the absorption spectra with temperature for a 50%, 20 monolayer structure assuming the* wells *to be doped with a 2-D doping concentration of* $4.2\times10^{15}\,cm^{-2}$.

7.1.2 Changing Doping Concentration

We now consider the effect of changes in doping concentration on the absorption spectra. The level of doping affects both the Fermi level and the exchange shifts. In table (7.1) we show the Fermi levels at and exchange shifts for a range of doping concentrations and in figure 7.11 we show the affect these changes have on the absorption spectra at 0K.

We can see that as the doping concentration is decreased the Fermi level moves towards the top of the HH1 miniband, as would be expected as less carriers are present. This reduction in the number of carriers also leads to a reduction in the exchange shift. We can see these results in the absorption spectra with the magnitude of the absorption drastically reducing with lower doping concentrations, and shifting to lower energies reflecting the reduction

2-D Doping Concentration	Exchange Shift	Fermi Level
m^{-2}	(meV)	(meV below HH1 at Γ)
6×10^{16}	15.4	-5.8
3×10^{16}	11.1	-3.0
6×10^{15}	5.1	-0.78
3×10^{15}	3.6	-0.44

Table 7.1: The change in Fermi levels and exchange shifts for varying doping concentrations in the *barrier* for the 50%, 20 monolayer structure.

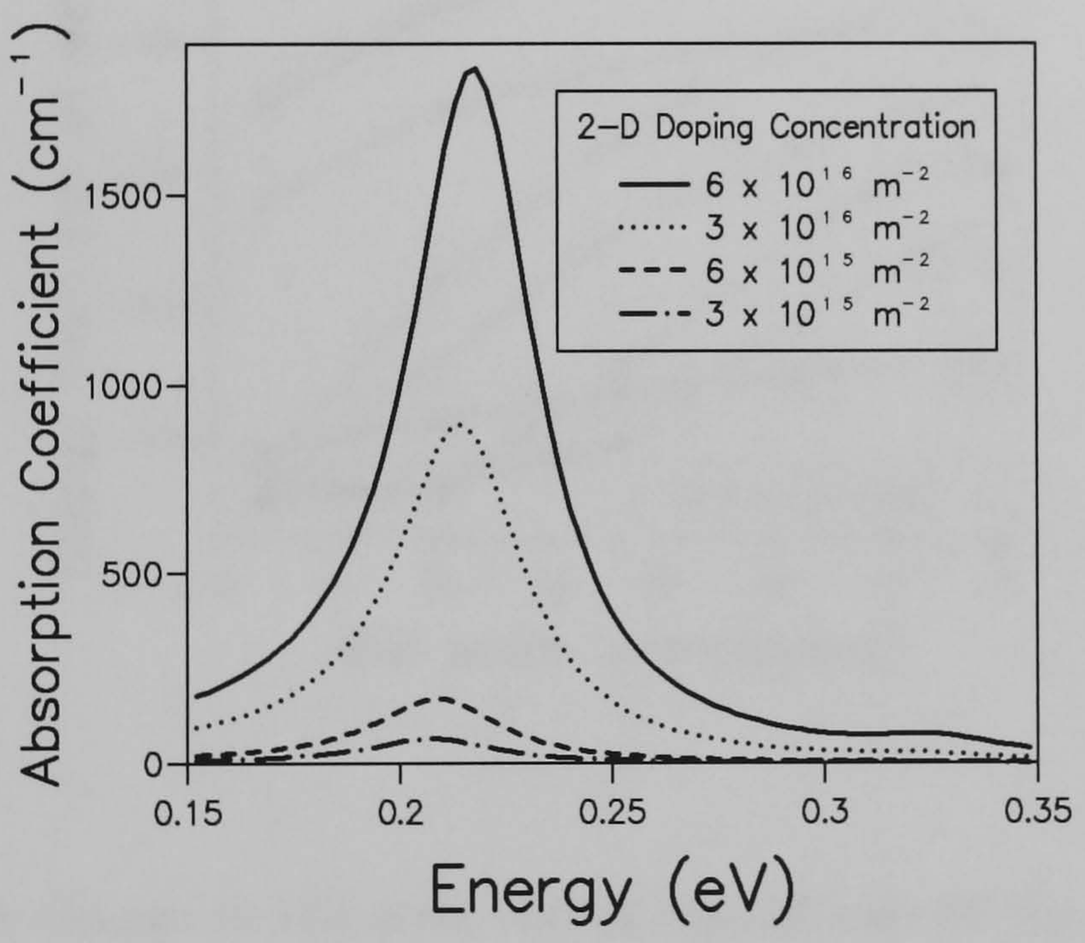


Figure 7.11: *The effect on the absorption spectra of the 50%, 20 monolayer structure of increasing the barrier doping concentration.*

in exchange shift. We can see that decreasing the doping concentration by a factor of two decreases the magnitude of the absorption by a similar amount. Thus high doping results in large absorption coefficients. However, as discussed in Chapter 1, high doping concentrations also lead to high dark currents, hence

the doping must be kept as low as possible for good device performance.

7.1.3 Changing Well Width

We now wish to explore the effect on the band structure and absorption spectra of changing the width of the well. In figure 7.12 we can see the change in the zone centre separations as the well width is increased from 12 monolayers to 44 monolayers. The dashed line indicates the position of the HH1 miniband

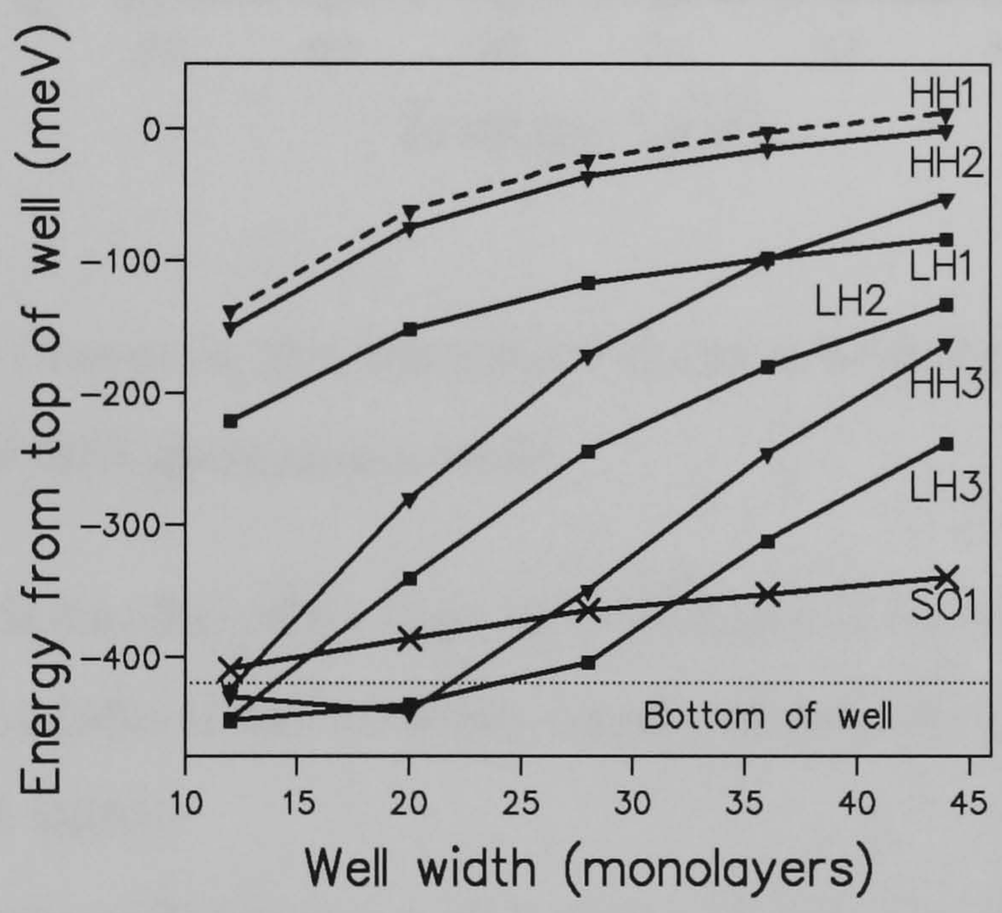


Figure 7.12: *The change in the zone centre separations of the minibands with well width for structures with 50% Ge wells. The dashed line shows the position of the HH1 miniband when an exchange shift of 13meV is included.*

when an exchange shift of 13meV is included. We can see the general trend that would be expected from simple particle in a box predictions, that the separation between minibands decreases as well width increases. It is interesting to note that the energy of the SO1 miniband is only slightly changed over the range of well widths. This is because as shown in figure 7.2 the SO1 miniband

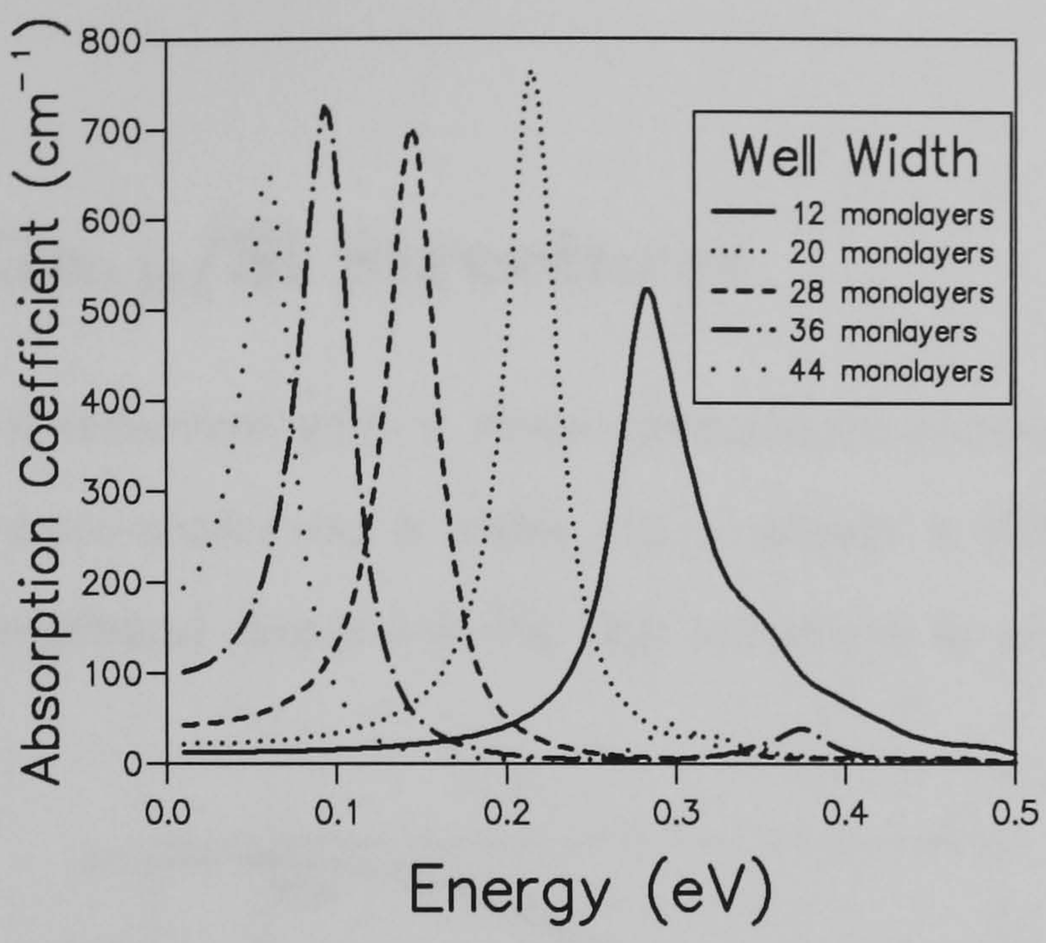


Figure 7.13: *The change in the absorption spectra with well width for structures consisting of 50% germanium wells.*

experiences a much smaller offset than the HH and LH bands. Hence this level is not as strongly confined as the other bands, therefore it is less affected by the change in well width.

This is also a very useful diagram as it shows the change between absorption from bound -bound and bound to continuum transitions. We can see that for a 12 monolayer well the HH2 miniband lies just outside the well, thus will still have a large overlap with the HH1 miniband but photo-excited holes will be easier to collect as they do not have to tunnel out of the wells. When the well width is increased to 20 monolayers the HH2 state has moved down into the well and a higher electric field would be required to collect the holes, resulting in a corresponding and undesirable increase in dark current.

In figure 7.13 we see the change in the absorption spectra as the well width is altered. Notice how the width of the peak increases and the magnitude decreases for the 12 monolayer case when the HH2 miniband is outside the

well.

7.2 $Si_{0.70}Ge_{0.30}$/Si Structures

We now analyze a structure with a lower germanium concentration. In this instance we will concentrate on a wider well, namely a 30%, 44 monolayer structure. The miniband dispersion for this structure is presented in figure 7.14.

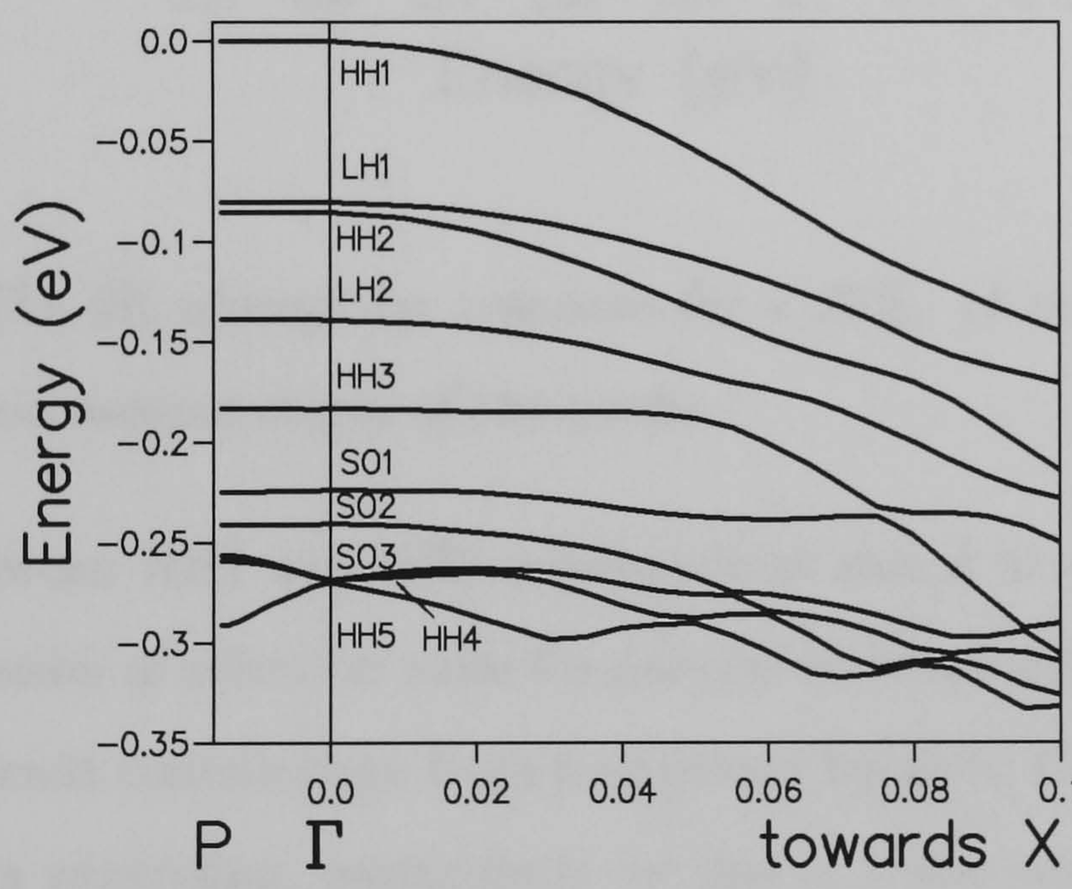


Figure 7.14: *Miniband structure for a 30%, 44 monolayer well along the symmetry axis P-Γ-X.*

We can see that the levels are much closer together for this structure compared with the band structure presented in figure 7.4. This is due to the increased well width. The parallel incidence absorption spectra for this structure, calculated at 0K is presented in figure 7.15. This structure was calculated using a damping constant of 15meV following the discussion in Chapter 6. We can see from this diagram that the majority of the response is due to

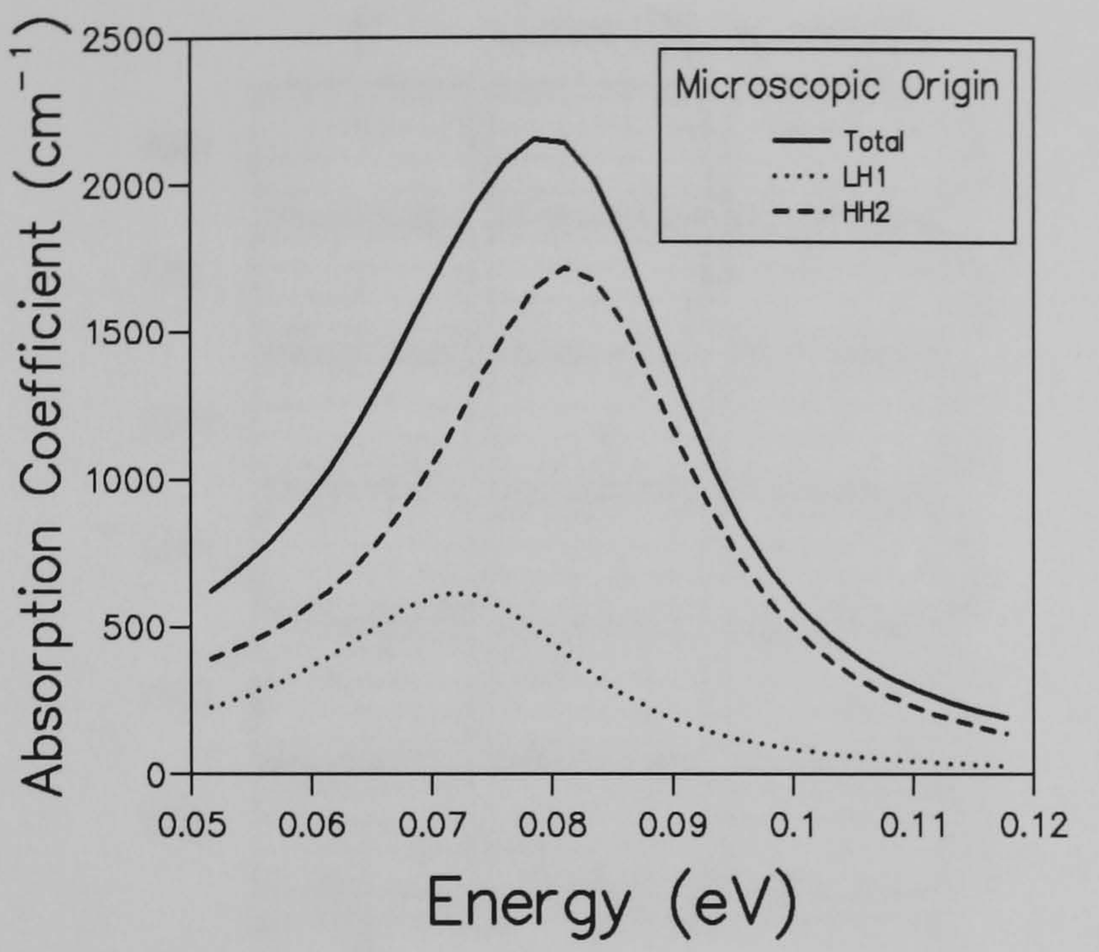


Figure 7.15: *The 0K absorption response for a 30%, 44 monolayer structure showing the microscopic origin of the peaks.*

transitions between HH1 and HH2 minibands as would be expected from our previous discussion of selection rules for parallel incident light. However, there is also a significant contribution from transitions between LH1 and HH1. This at first appears surprising, particularly as this is a low temperature calculation where carriers are confined to the region $\mathbf{k} \approx 0.03\left(\frac{2\pi}{a}\right)$. But if we look at the eigenfunction coefficients, $A_{n\mathbf{k}}$'s, plotted in figure 7.16 we can see that at $\mathbf{k}= 0.025\left(\frac{2\pi}{a}\right)$ the LH1 miniband has already acquired a significant admixture from the HH bulk state. The closer the states are in energy the more strongly they will mix (Man 1994). Hence this transition will become allowed with a transition strength proportional to the degree of admixture of the heavy hole bulk states into the light hole band. We also notice that this transition occurs at a lower energy than the HH1-HH2, even though these states are very close in energy around Γ, again suggesting that it is occuring away from the zone centre where the mixing is stronger.

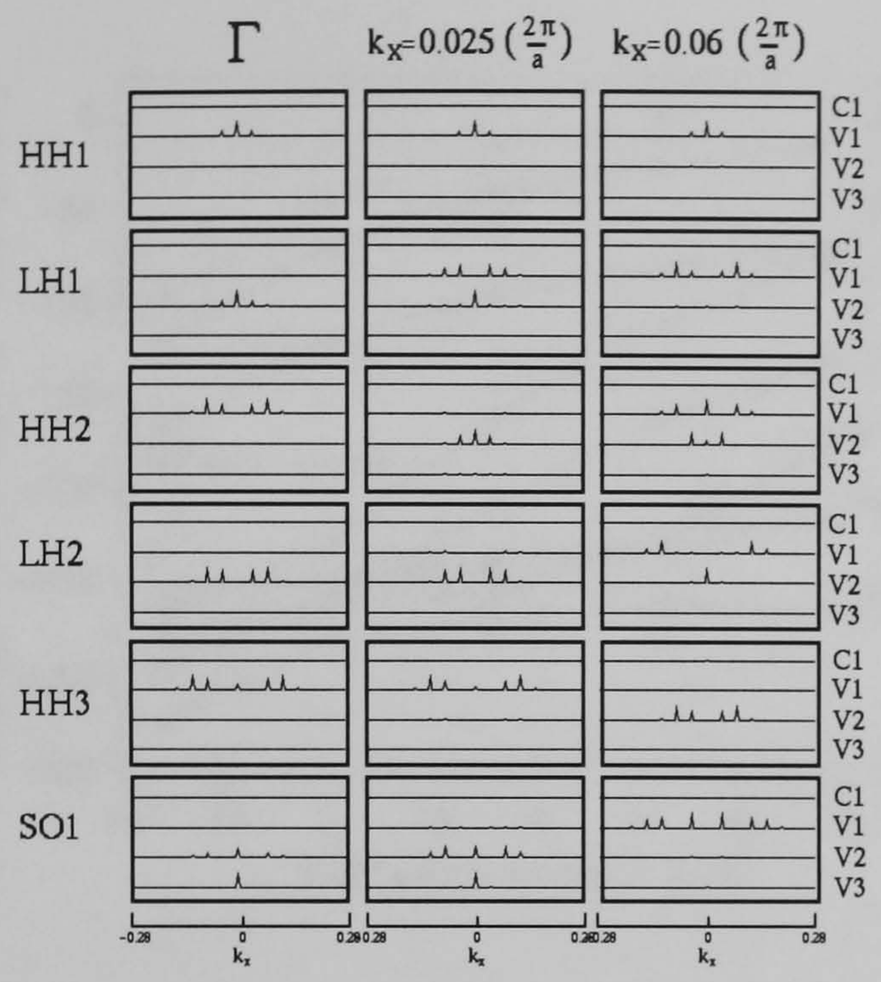


Figure 7.16: *The eigenfunction coefficients, $A_{n\mathbf{k}}$'s for the 30%, 44 monolayer structure showing significant contributions from different bulk states for all values of* $\mathbf{k}$ *away from* Γ. *The bulk conduction state is labelled C1, and HH bulk state V1, the LH bulk state V2 and SO, V3.*

The change in the position of the zone centre energy levels as the well width is decreased is shown in figure 7.17. We can see that the inclusion of the exchange shift moves the HH1 miniband above the top of the bulk band edge which we consider to be the top of the well. This is not a problem as the bulk band edges were calculated without reference to doping or the resultant many body effects, and can therefore only be considered a rough guide to the actual potential profile. From figure 7.17 we can see that structures with well widths of 12 monolayers and 20 monolayers are potentially useful for device applications as the HH2 miniband lies outside or just inside the well.

The change of the absorption response with well width is shown in figure 7.18. All peaks are predominately due to transitions between HH2 and HH1 minibands. The peaks follow the expected trend again moving to lower

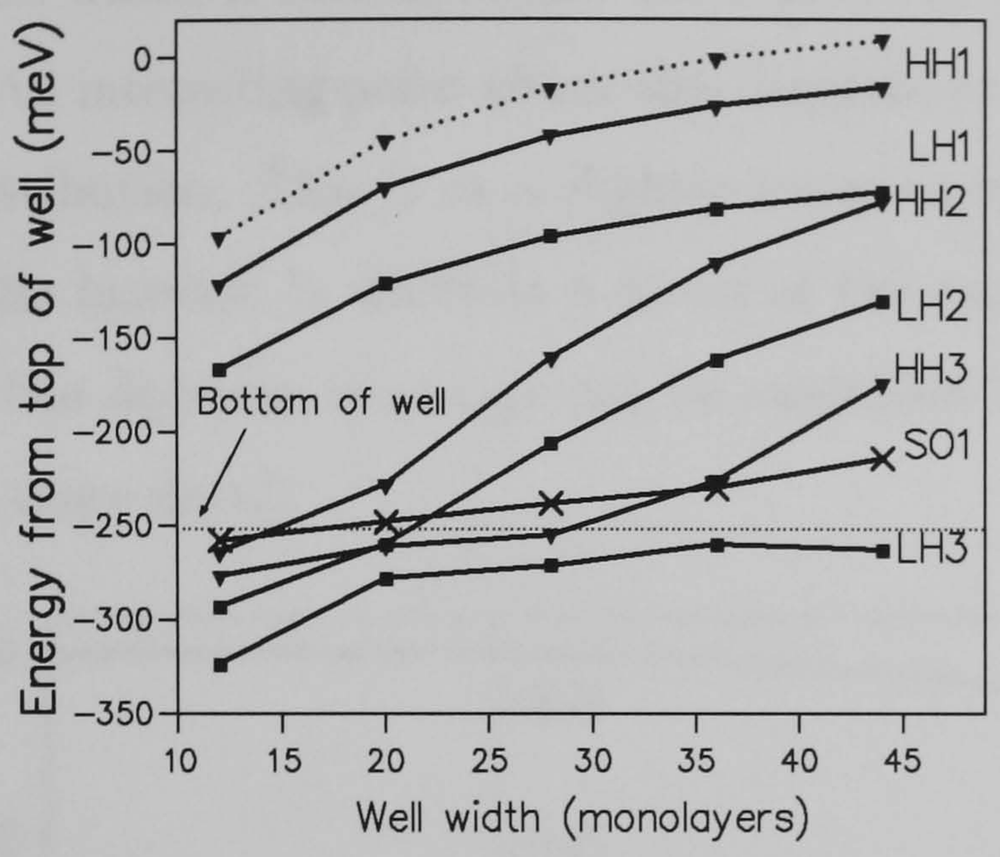


Figure 7.17: *Change in* Γ *separation of the energy levels as the well width is increased for 30% Ge structures. The dashed line indicates the position of the HH1 miniband when an exchange shift of 24meV is included.*

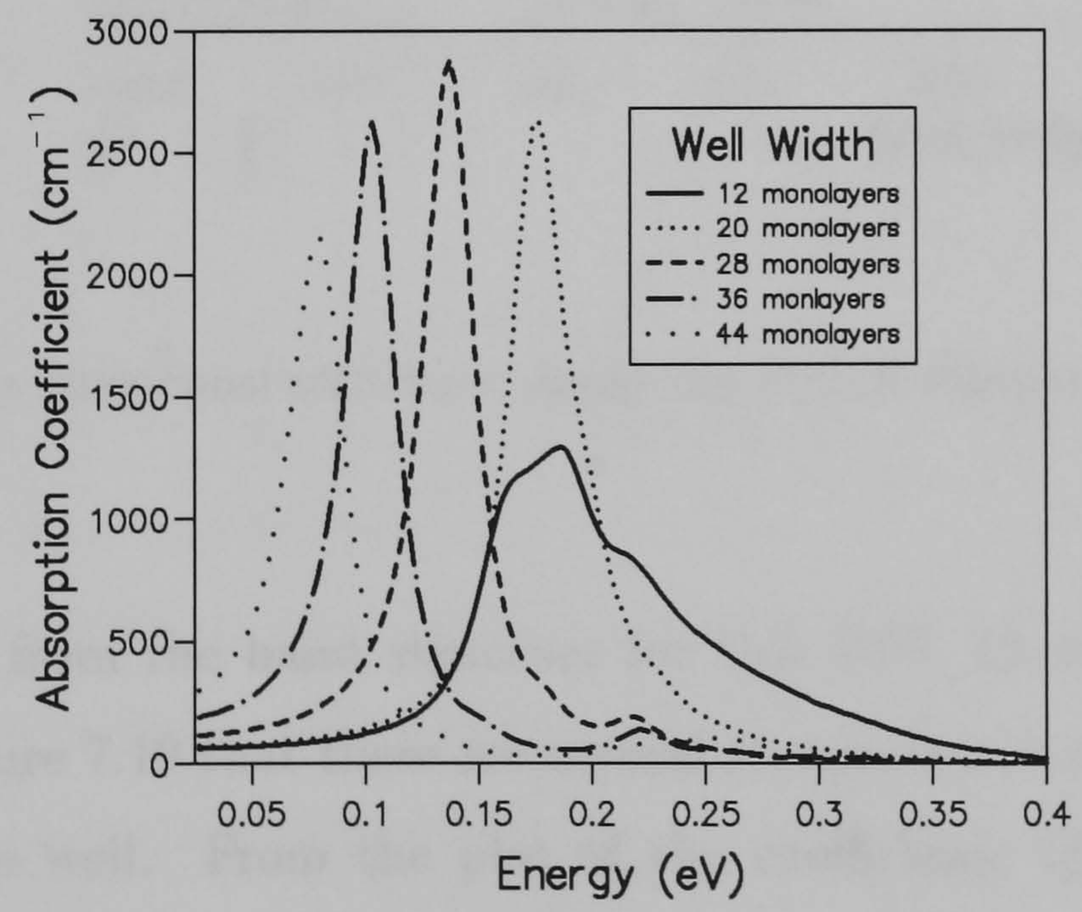


Figure 7.18: *The change in absorption spectra with well width for 30% structures. All curves were calculated at 0K.*

energies as the well width is increased and the separation between the minibands decreases. An interesting point about this diagram is the position of the 12 monolayer contribution. This is at a slightly lower energy than the trend would suggest. The increase in width is a result of the excited state moving into the barrier. The decrease in energy can be explained by considering the band structure in more detail.

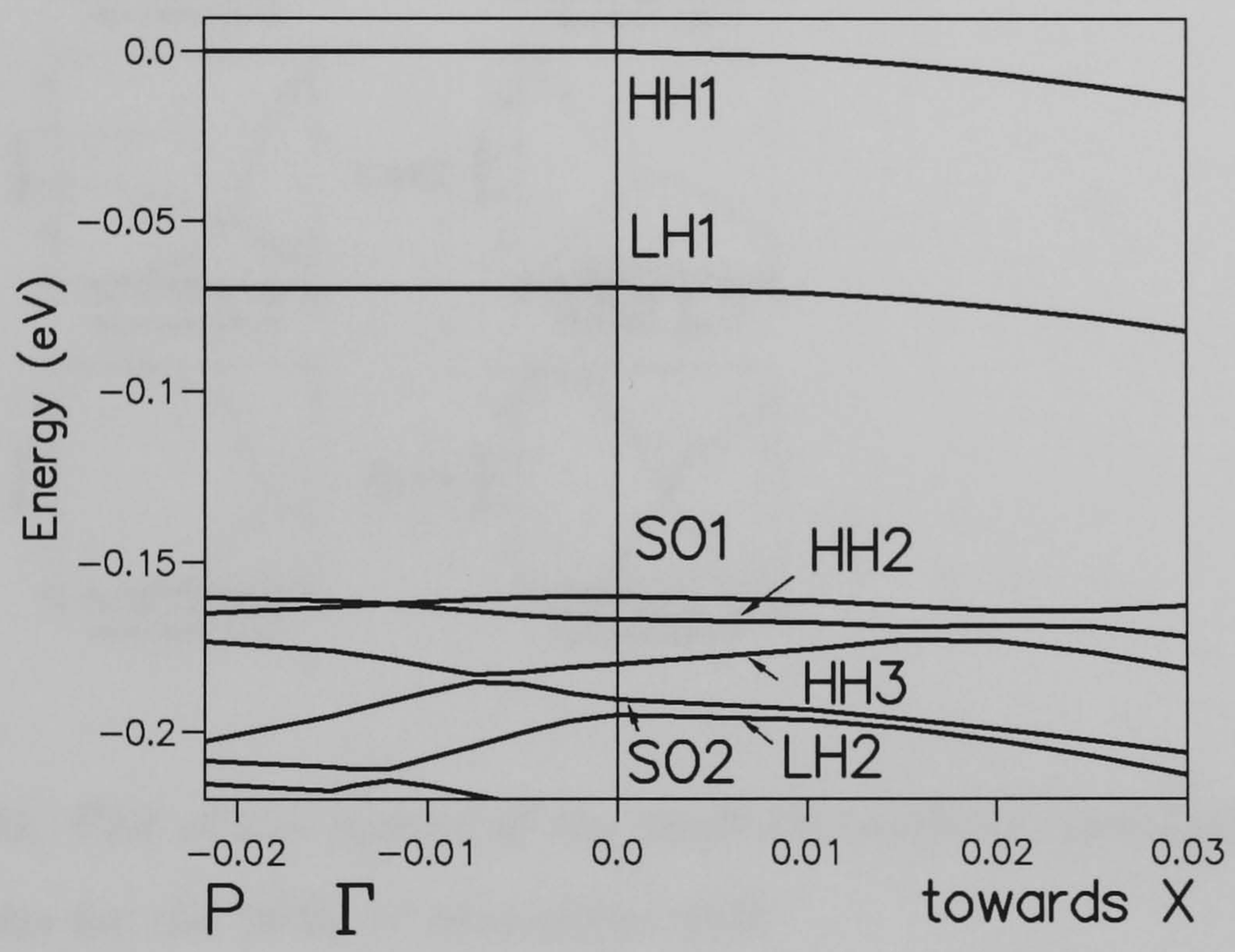


Figure 7.19: *The miniband structure along the P-Γ-X directions for a 30%, 12 monolayer well.*

We can see from the band structure for this 30%, 12 monolayer system presented in figure 7.19 that there are several states very close in energy lying just outside the well. From the plot of the coefficients of the bulk states present in the quantum well eigenfunctions shown in figure 7.20 we can see that all these states have significant contribution from the bulk heavy hole state. The small energy separation between these bands means that the mixing is enhanced. Thus all these states contribute to the parallel incidence response

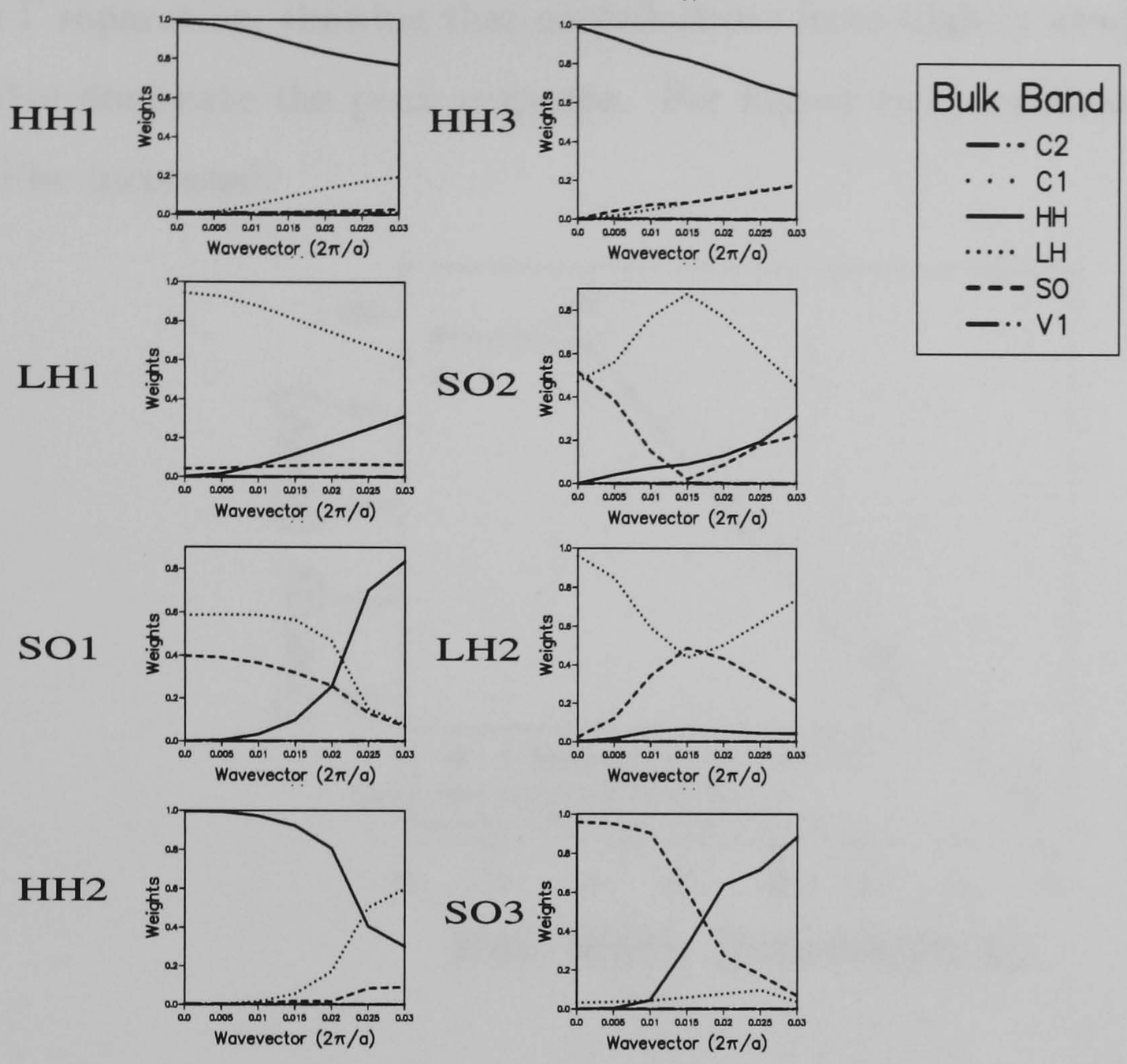


Figure 7.20: *Plot of the square of the coefficients of the eigenfunctions along the* Γ*-X axis for the 30%, 12 monolayer well.*

for this structure hence the very broad peak.

This also demonstrates an important point that major contributions to the absorption occur in areas of the zone away from the zone centre, thus the separation of the relevant minibands at the zone centre is not always a good indication of the resultant absorption peak response. This point is illustrated by figure 7.21 which shows the difference between the zone centre separation of the levels and the energy of the peak absorption response. We can see that for a complicated case like the 12 monolayer situation the peak response could not be predicted from a calculation performed at the zone centre alone. For all the other well widths the peak response occurs at a slightly lower energy that

the Γ separation, showing that contributions from slightly away from the zone centre dominate the peak response. For higher temperatures this difference will be increased.

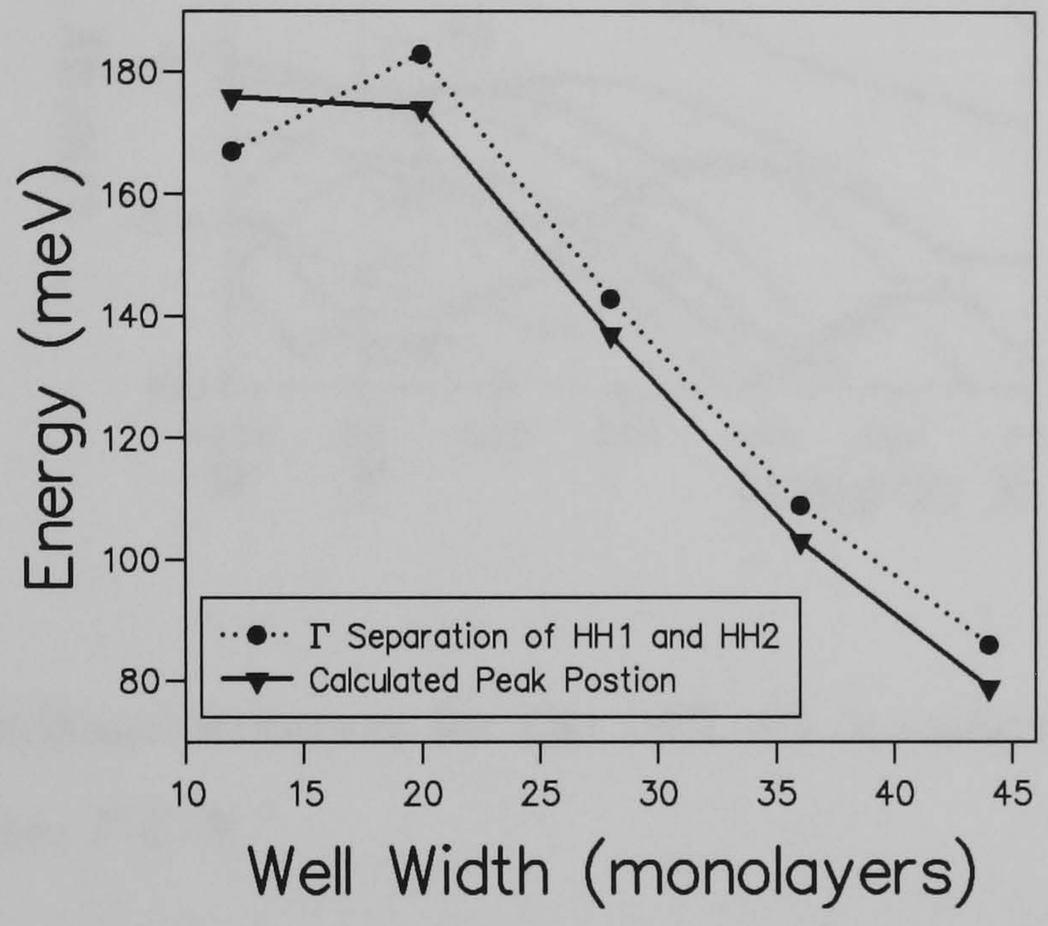


Figure 7.21: *The trend of the peak absorption response as the well width is increased for 30% structures. Also shown is the Γ separation of the HH1 and HH2 minibands.*

7.3 $Si_{0.85}Ge_{0.15}$/Si Structures

We now move on to consider structures with lower germanium concentrations. These are more useful for device applications as the relevant levels are nearer the top of the well, the band offset is such that the absorption occurs around the desired 10μm, i.e. $\approx$ 120meV. However, the difficulty in analyzing these structures is clear if we look at the miniband structure for a 20 monolayer wide $Si_{0.85}Ge_{0.15}$/Si system (15%, 20 monolayers). This diagram includes an exchange shift of 25meV corresponding to a 3-D doping concentration of 2.2

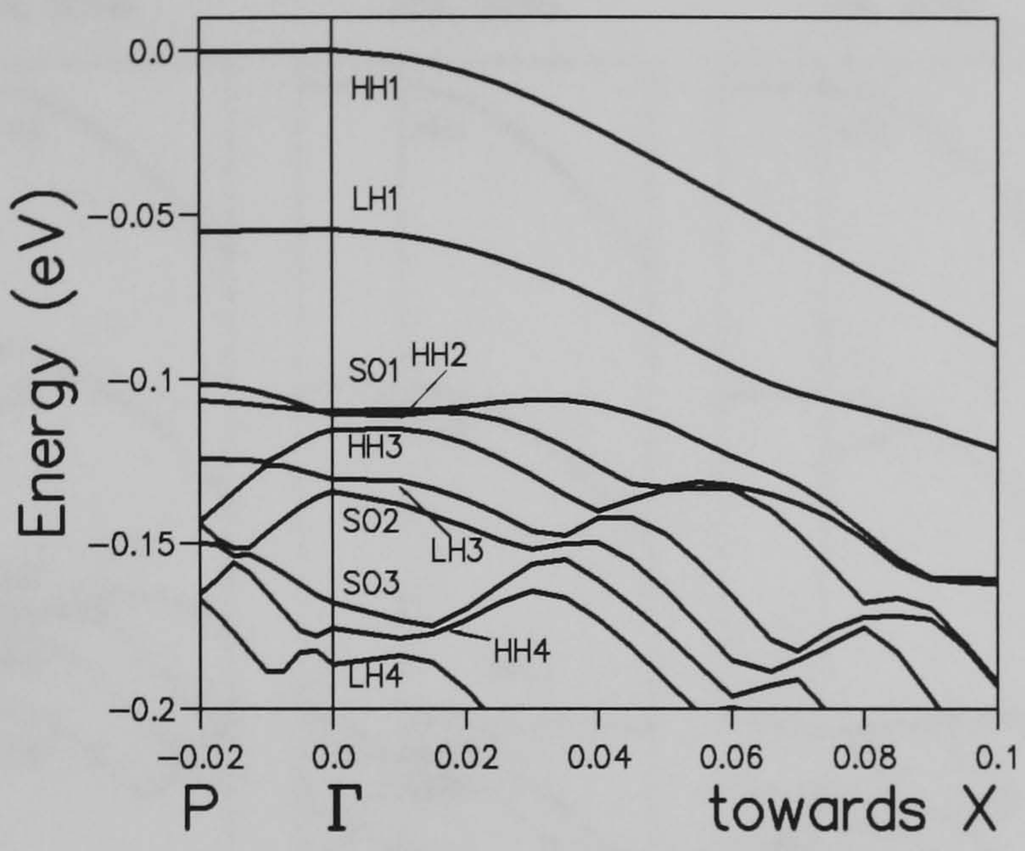


Figure 7.22: *Miniband structure for the 15%, 20 monolayer structure along the symmetry lines P-Γ-X.*

$\times 10^{19}\mathrm{cm}^{-3}$ across 60 Å of the barrier. Again this value is obtained from the work of Fromherz for a similar structure. We can see that the SO1 and HH2 minibands are very close in energy with the SO1 miniband uppermost at the zone centre but the HH2 miniband moves above it both along the growth direction and along the k_x axis. This crossing makes assignment of the microscopic origin of the peaks more difficult as the bands are labelled on the basis of their nature at Γ, and the microscopic origin is determined by the order of the states at the zone centre. Therefore, if crossing of the states occurs, contributions from the minibands after they have crossed will be mislabelled. We can unravel the band crossing and label the bands away from the zone centre by considering the $A_{n\mathbf{k}}$'s and the charge densities. However, this does not help with the problem of attributing the microscopic origin for the absorption. In order to do this we consider similar structures with slightly less complex band structures and examine the contributions to the absorption spectra. From this

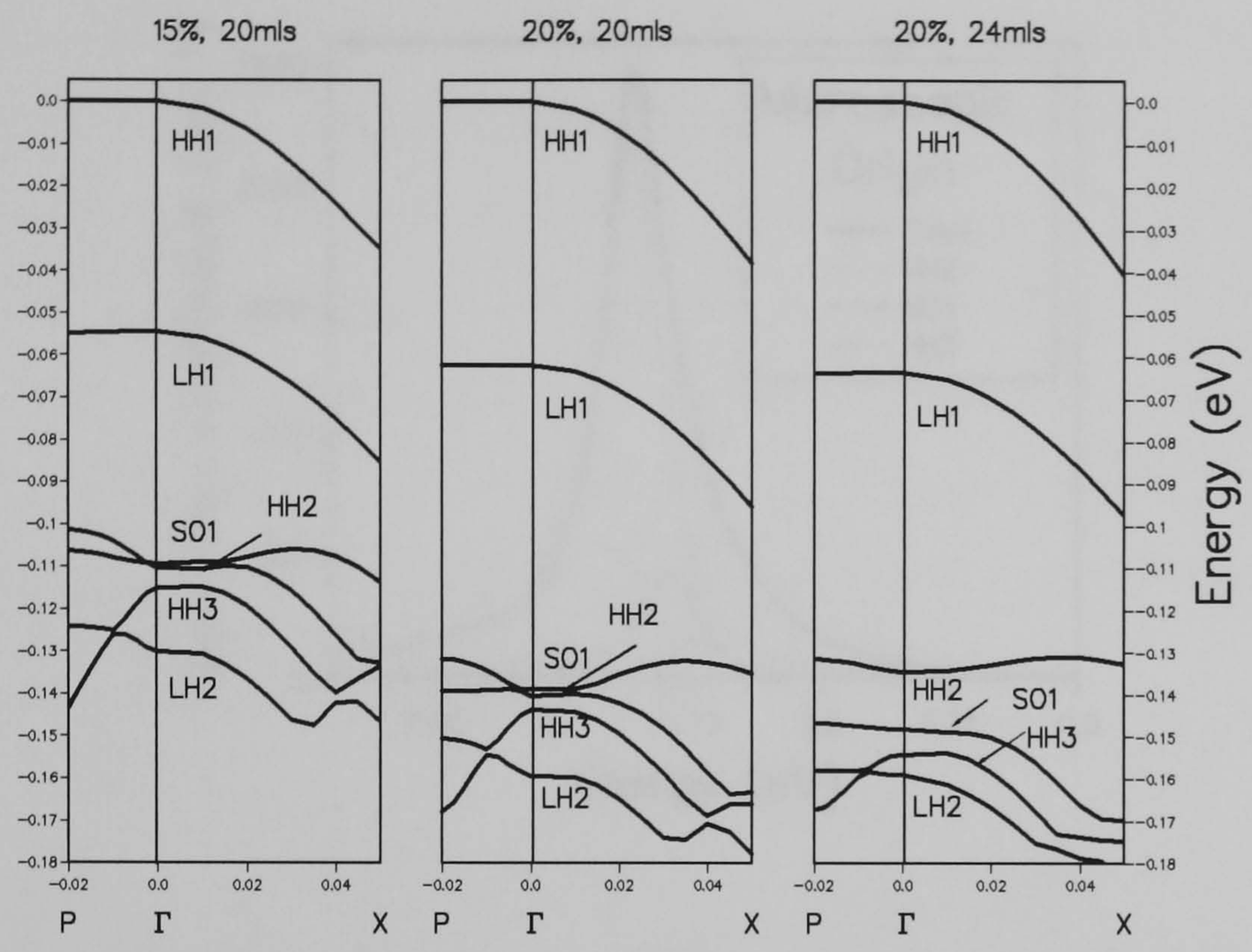


Figure 7.23: *Comparison of the miniband dispersion for three similar structures, namely 15%, 20 Mls, 20%, 20Mls and 20%, 24 Mls along the P-Γ-X directions.*

we can deduce the behaviour of the 15%, 20 monolayer system. This also gives us an opportunity to analyze the effect of small changes in structural parameters on the absorption response. For this purpose we compare three structures, 15%, 20 monolayers, 20%, 20 monolayers and 20%, 24 monolayers. The band structures along the P-Γ-X directions are plotted in figure 7.23. From figure 7.23 we can see that increasing the germanium concentration by 5% increases the separation of the HH1 and HH2/SO1 minibands by $\approx$ 30meV. Hence we would expect the absorption peak to move by a corresponding amount. Increasing the well width of the 20% structure by 4 monolayers, or one lattice constant, which is the smallest, convenient increment possible (increasing the well width by less than this amount alters the symmetry of the system and makes direct comparisons more difficult), lifts the HH2 miniband by $\approx$ 10meV.

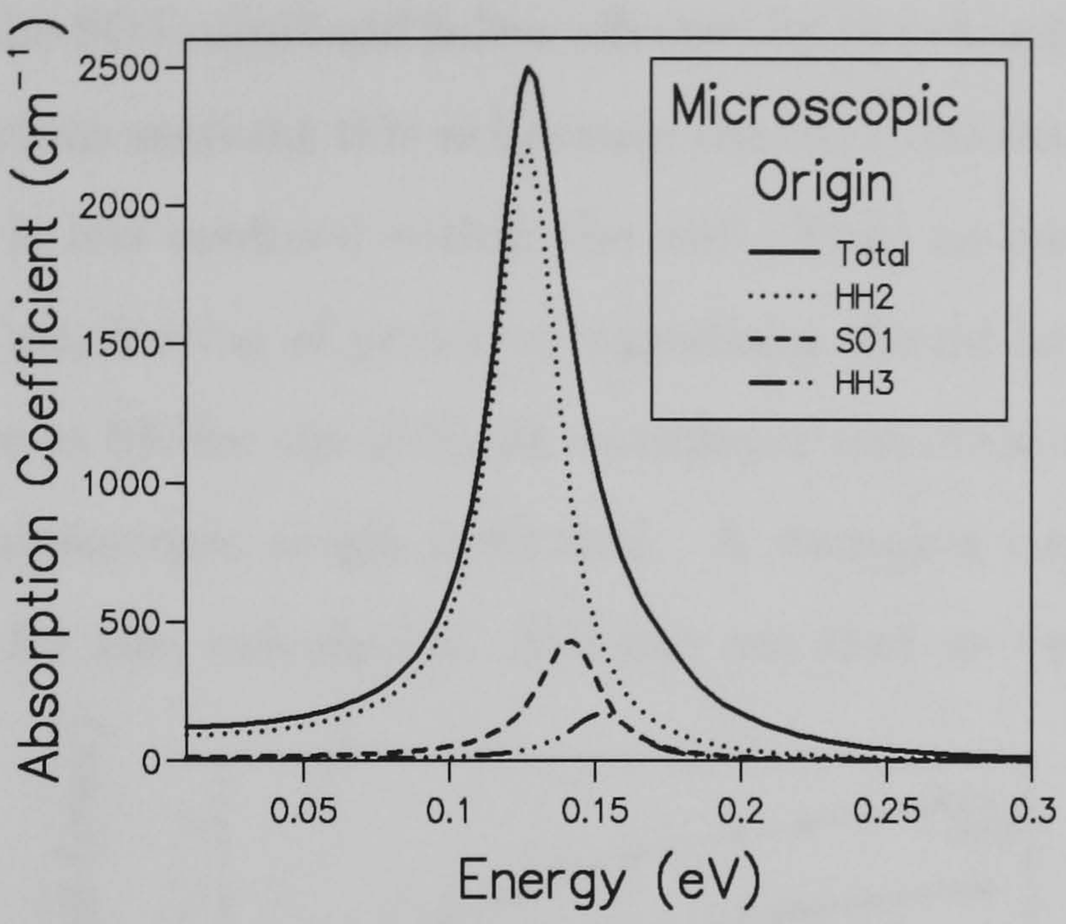


Figure 7.24: *The absorption response calculated at 0K for a 20%, 24 monolayer structure.*

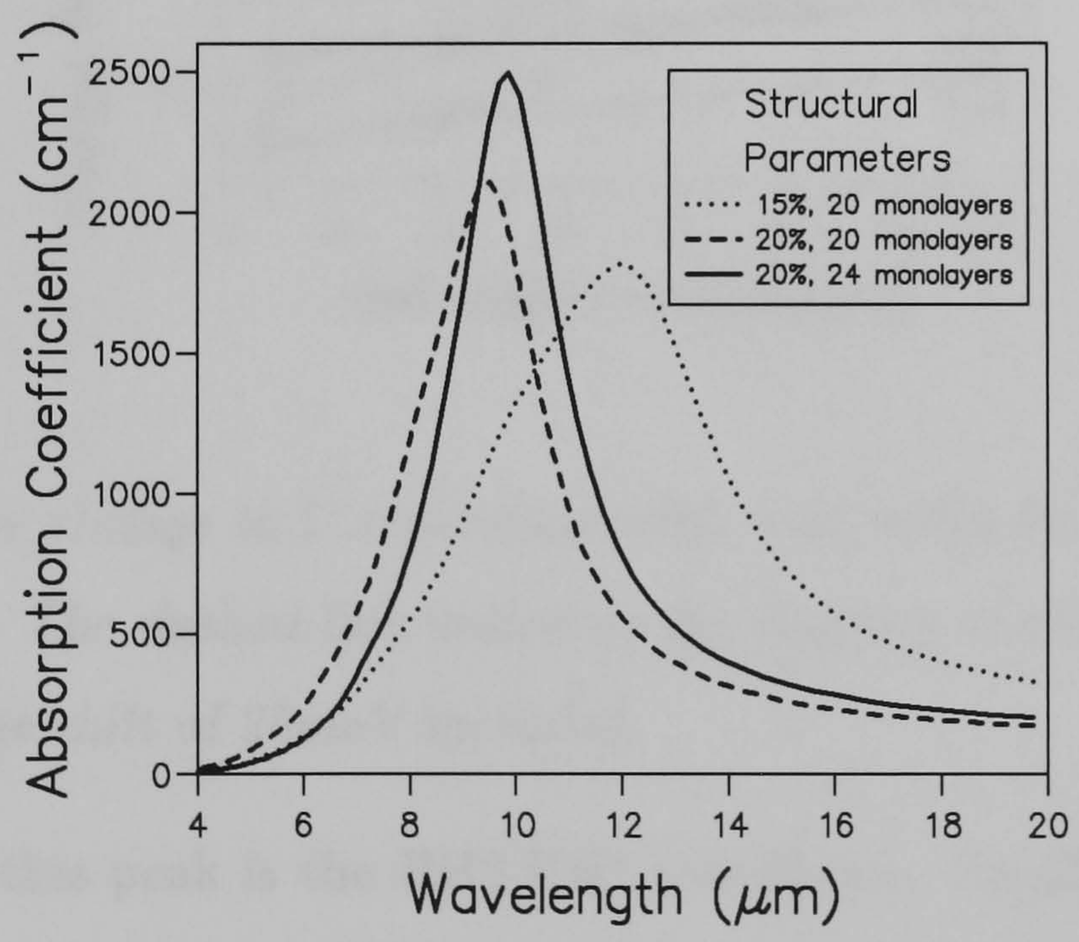


Figure 7.25: *The change in absorption response with structural parameters. All curves calculated at 0K using a damping constant of 15meV.*

The position of the SO1 miniband is less affected by the change in well width as discussed in previous sections this is because the SO1 miniband sees a smaller offset, therefore is less confined within the well. Thus as these minibands no longer cross the attributing of peaks to transitions should be simple. The absorption response at 0K for the 20%, 24 monolayer structure is shown in figure 7.24 with the microscopic origin indicated. A damping constant of 15meV was again used for this calculation. We can see that as expected the main

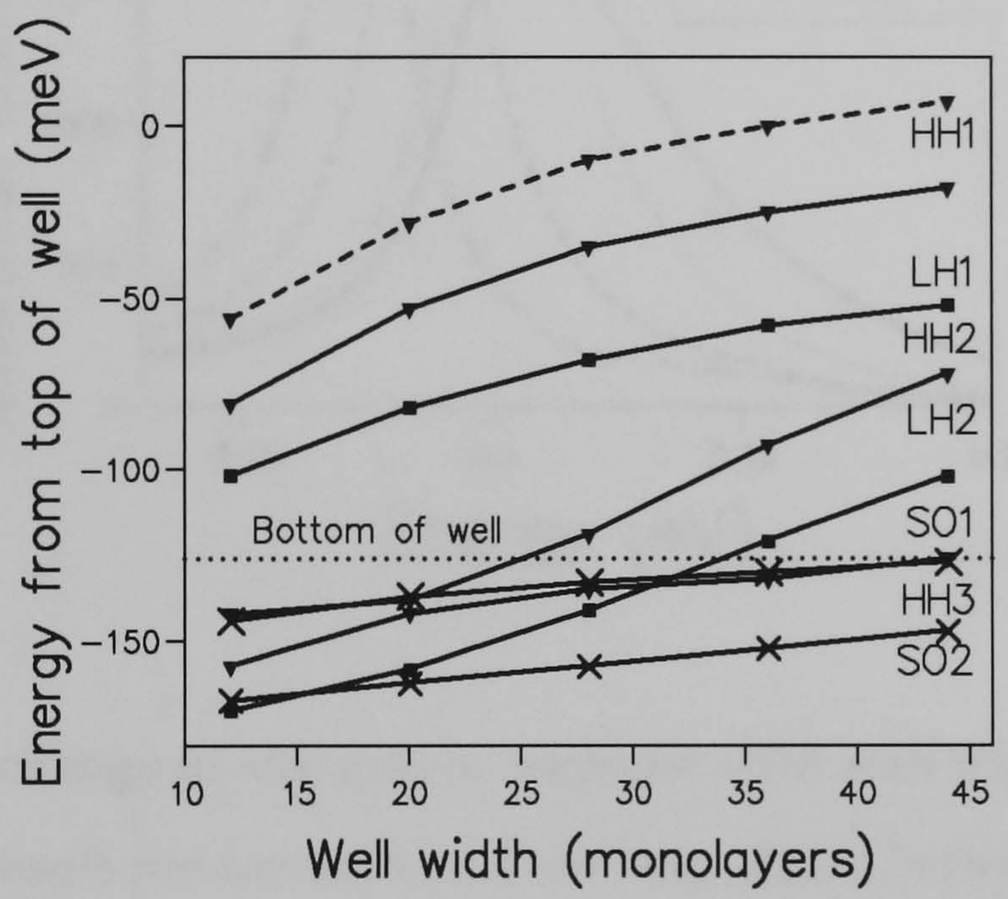


Figure 7.26: *The change in Γ separation with well width for structures with 15% alloy wells. The dashed line indicates the position of the HH1 miniband with an exchange shift of 25meV included.*

contribution to this peak is the HH2-HH1 transitions. Smaller contributions from SO1-HH1 and lower minibands can also be seen. If we now plot the absorption spectra for all three curves at 0K we can see the effect of the small changes. The response for the 15% structure is considerably wider despite all curves having the same damping constant of 15meV. This is again because of the greater degree of non-parabolicity of the minibands in this structure.

This plot is against wavelength and shows that all of these structures ab-

sorb at around 10μm, with the HH2 miniband just outside the well. All three of these structures are therefore good candidates for parallel incidence IR absorption in the 8-15μm range utilizing bound-continuum transitions. The

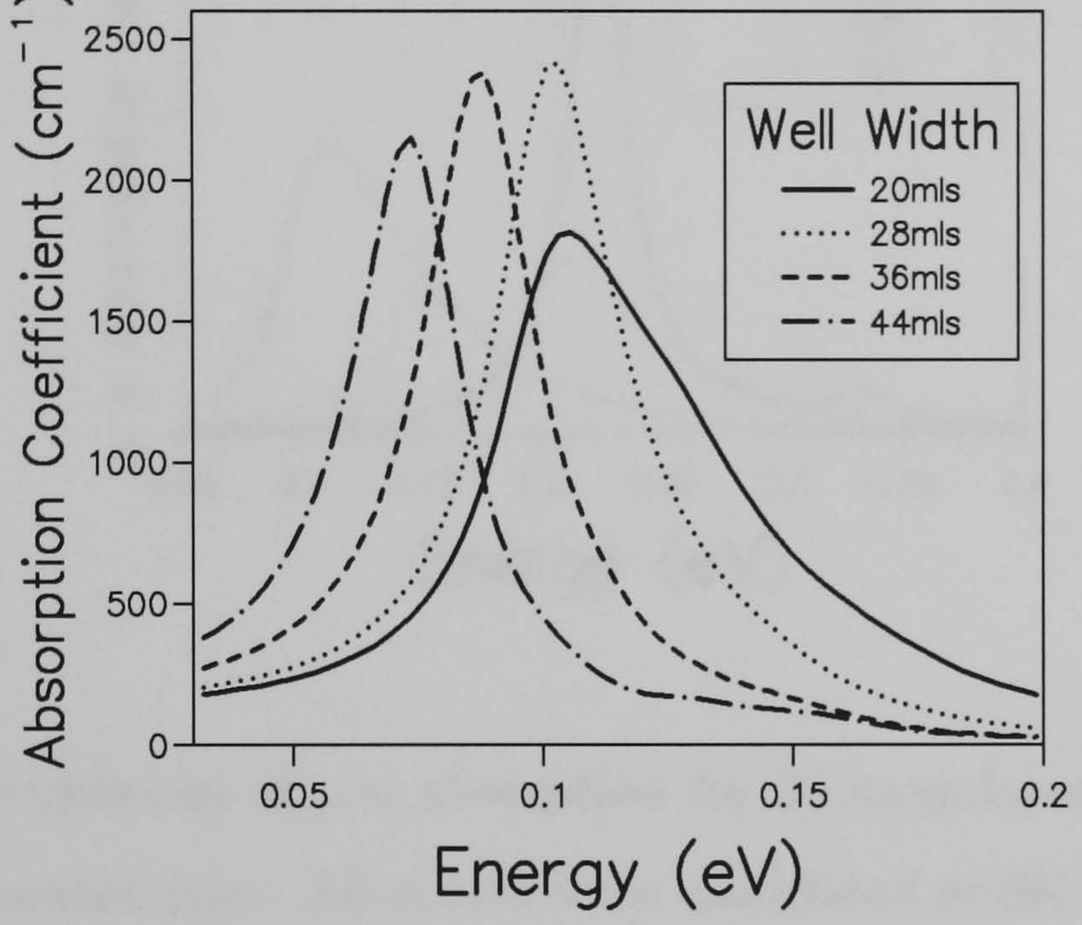


Figure 7.27: *The change in absorption response with well width for 15% structures. All peaks result predominately from transitions between HH1 and HH2 minibands. All curves were calculated at 0K using a damping constant of 15meV.*

change in the zone centre separation of the energy levels with well width for the 15% structures is shown in figure 7.26 and the comparison of the absorption curves for different well widths in figure 7.27. We can see that all parallel incidence absorption between HH1 and HH2 minibands is due to bound-bound transitions for well widths above 28 monolayers. For well widths below this the absorption becomes a bound-continuum processes. This change is reflected in the increased width of the peak for the 20 monolayer case shown in figure 7.27.

Finally, we compare the parallel incidence absorption for 20 monolayer

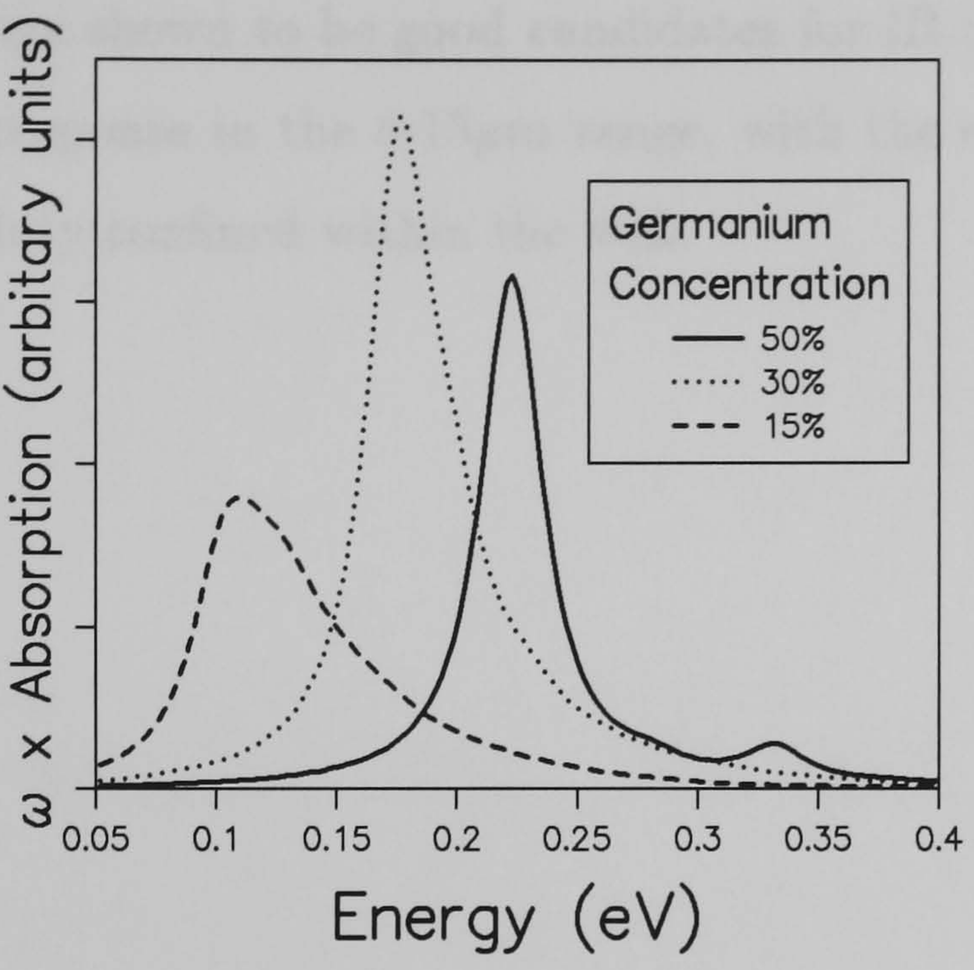


Figure 7.28: *Comparison of $\omega\times$ absorption for 20 monolayer wide wells with varying Ge concentrations. All curves were calculated at 0K using a damping constant of 15meV.*

wells with a range of germanium concentrations in the well. The absorption coefficient is proportional to ω^{-1}, the photon frequency, so to aid comparison we plot $\omega\times$ absorption in figure 7.28. All peaks are predominantly due to transitions between HH1 and HH2 minibands. We can see that magnitude of the 15% cases is considerably lower than the other two again as a consequence of it being a bound-continuum transition thus the overlap of the wavefunctions is reduced compared to the bound-bound transitions of the higher germanium concentrations.

In conclusion, we have shown for parallel incidence absorption dominant absorption mechanism is transitions between HH2 and HH1 minibands. Smaller contributions to the response may arise from transitions with other lower lying minibands, particularly at high temperature due to band mixing effects.

Structures containing 15% or 20% Ge alloys with wells approximately 20

monolayers wide were shown to be good candidates for IR detectors, showing strong absorption response in the 8-15μm range, with the excited state lying just outside or slightly confined within the well.

Chapter 8

Normal Incidence Absorption

We discussed in Chapter 1 why normal incidence (NI) absorption is a great asset in device design. 2-D arrays must be illuminated perpendicular to the plane of the wafer, and if absorption can be achieved without the use of gratings and waveguides etc. the system is cheaper and easier to make. Normal incidence absorption in p-type SiGe/Si structures was first demonstrated by Park (1991) and by several groups since (for example People 1992 and Zanier 1995). We discuss the normal incidence absorption spectra of SiGe/Si quantum wells and show that they are good candidates for normal incidence IR detection in both the 3-5μm and 8-15μm ranges. We begin by discussing band mixing in more detail and then proceed to demonstrate the microscopic origin of the normal incidence absorption spectra of structures with a range of Ge concentrations in the well. The change in absorption response as the well width, doping concentration and temperature are altered is discussed. Again we start the discussion with the structures with higher germanium concentrations which are easier to analyze, before progressing onto the lower germanium concentrations.

8.1 Band Mixing

The bulk band structure can be considered to be composed of sp^3 hybrid states (see Appendix 2) (Heine 1960). At the zone centre of a bulk unstrained crystal this is decoupled to a low lying s-state and 3 degenerate p-like states. Away from the zone centre the coupling between these bands means that they can no longer be considered purely s or p-like. The lifting of the degeneracy of these bands away from the zone centre is an indication of this. The state which hybridises most strongly with the lowest s-like state is called the light hole state because of the greater curvature of the band being interpreted as a lower effective mass. Thus the bulk light hole state always contains significant s-like character away from the zone centre. There is also effects such as the valence band anisotropy and non-parabolicity that can effect the character of the states and the optical transitions between them. All band mixing effects are included implicitly in our pseudopotential calculations which constructs the quantum well states from whatever bulk states are required to reproduce the required form of the quantum well eigenfunctions. This is taken care of in the diagonalization routines which select the appropriate $A_{n\mathbf{k}}$'s for the bulk states.

We showed in Chapter 3 that normal incidence absorption is not permitted in the absence of band mixing. A discussion of band mixing within the framework of the tight-binding method is given by Chang (1985 and 1992) Schulman (1985), and Sanders (1985). There are two distinct types of band mixing. States of different bulk momenta are folded onto the same superlattice k-point in the zone folding picture. These states will mix in the final quantum well wavefunction. We saw in figure 7.7 that there is contributions to the minibands at the zone centre from states with different bulk momenta, i.e. $\mathbf{k_{bulk}} \neq 0$. The number of these states present increases as the quantum

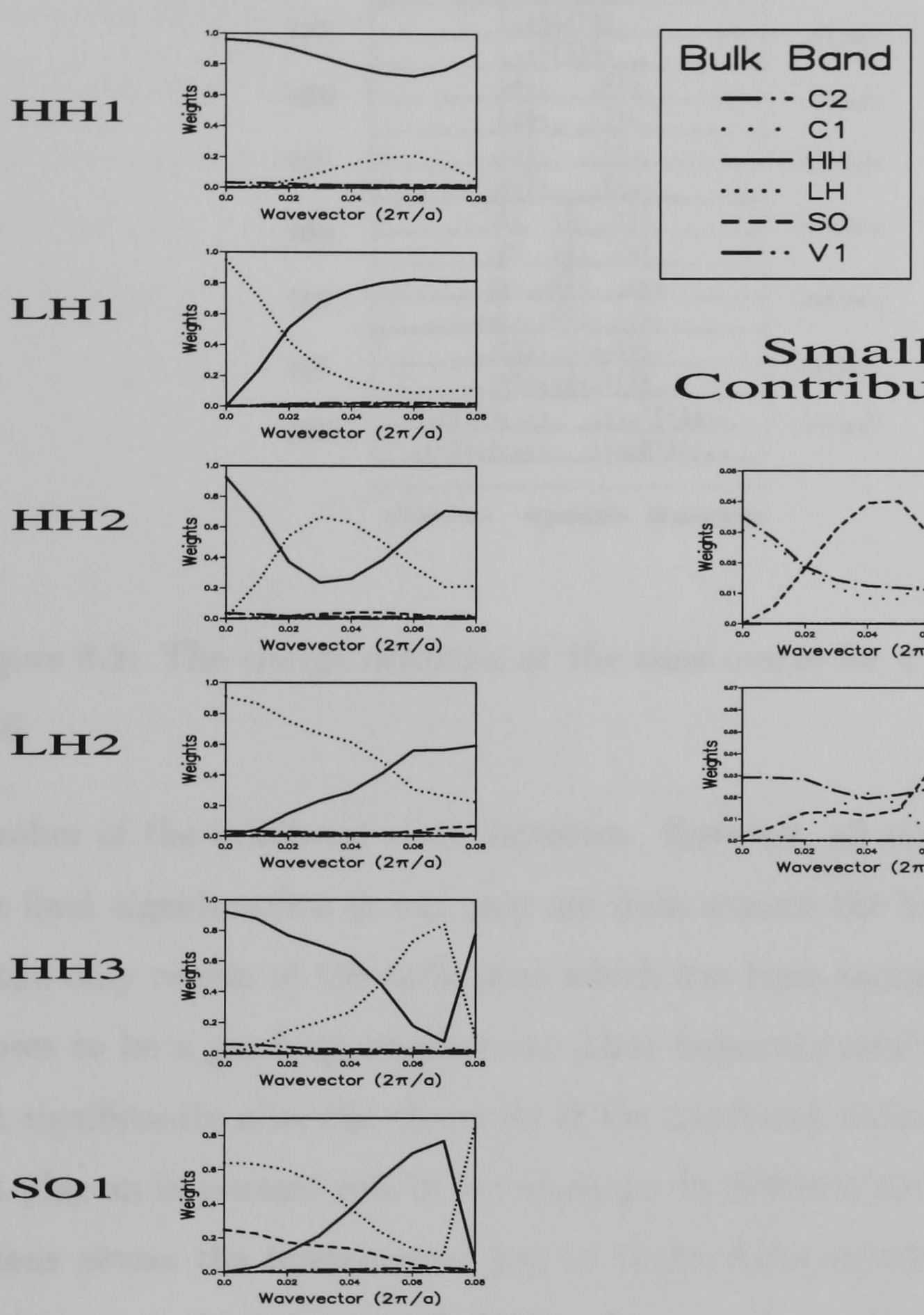


Figure 8.1: *Contributions from the bulk states which make up the quantum well minibands for a 30%, 44 monolayer structure. The x-axis is along the* Γ*-*X *direction.*

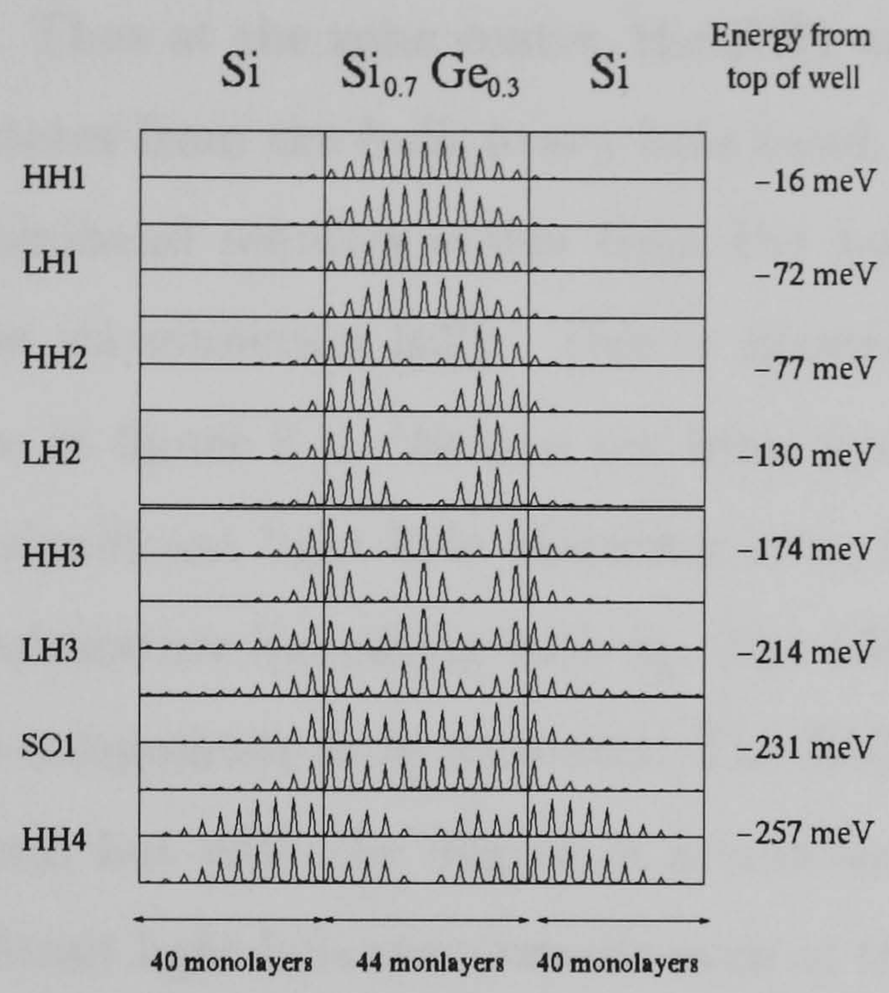


Figure 8.2: *The charge densities at the zone centre for a 30%, 44 monolayer well.*

number of the miniband state increases. However, all states contributing to the final eigenfunction in our case are from around the bulk Γ point, as this is the only region of the bulk zone which has been sampled. This has been shown to be a good approximation. Thus momenta mixing of this type does not significantly alter the character of the miniband states and therefore does not play an important role in our analysis. In different situations, where transitions across the fundamental gap in GaAs/AlAs structures are considered for instance, then states with X-like character from the X conduction band minima mix with states from the Γ point, (see, for example Gell 1987), then this type of mixing is of importance when attempting to explain the optical properties of the system.

The second type of mixing is loosely referred to as mixing between the quantum well minibands. This is inaccurate terminology. The source of this mixing is contributions from different bulk states which make up the quantum

well wavefunctions. Thus at the zone centre, the HH1 miniband is comprised almost entirely of states from the bulk heavy hole band. Away from the zone centre, the HH1 miniband requires states from the bulk light hole valence band to expand the wavefunction fully. This is illustrated for the 30%, 44 monolayer structure in figure 8.1. We can see from figure 8.1 that the HH1 miniband acquires significant light hole character away from the zone centre, with the degree of admixture increasing with $k_{||}$. The LH1 miniband similarly acquires heavy hole components as $k_{||}$ increases. The HH2 and LH2 minibands follow the same trend but with the degree of admixture reduced. The SO1 miniband has significant light hole components even at the zone centre. These states have the same m_J thus are more prone to mixing than the heavy hole states (Man 1994). These states were identified as spin split off states from the charge densities. It can be seen from figure 8.2 that the split off state has a ground state wavefunction, i.e. no nodes along the k_z direction thus must be considered a SO state despite the large admixture of the LH bulk state even at the zone centre.

We can also see in figure 8.1 that the HH2 and LH2 minibands have small but noticeable components of the lowest conduction band present. This is an anti-bonding s-like state, and its presence relaxes the selection rules and allows normal incidence absorption to occur (Bertho 1994).

8.2 $Si_{0.5}Ge_{0.5}$/Si Quantum Well structures

The normal incidence absorption spectra for this structure is presented in figure 8.3. A damping constant of 5meV was used for this structure. This is a somewhat unrealistically low value as discussed in Chapter 6 but using larger values tends to mask the detail in these NI response curves as shown in fig-

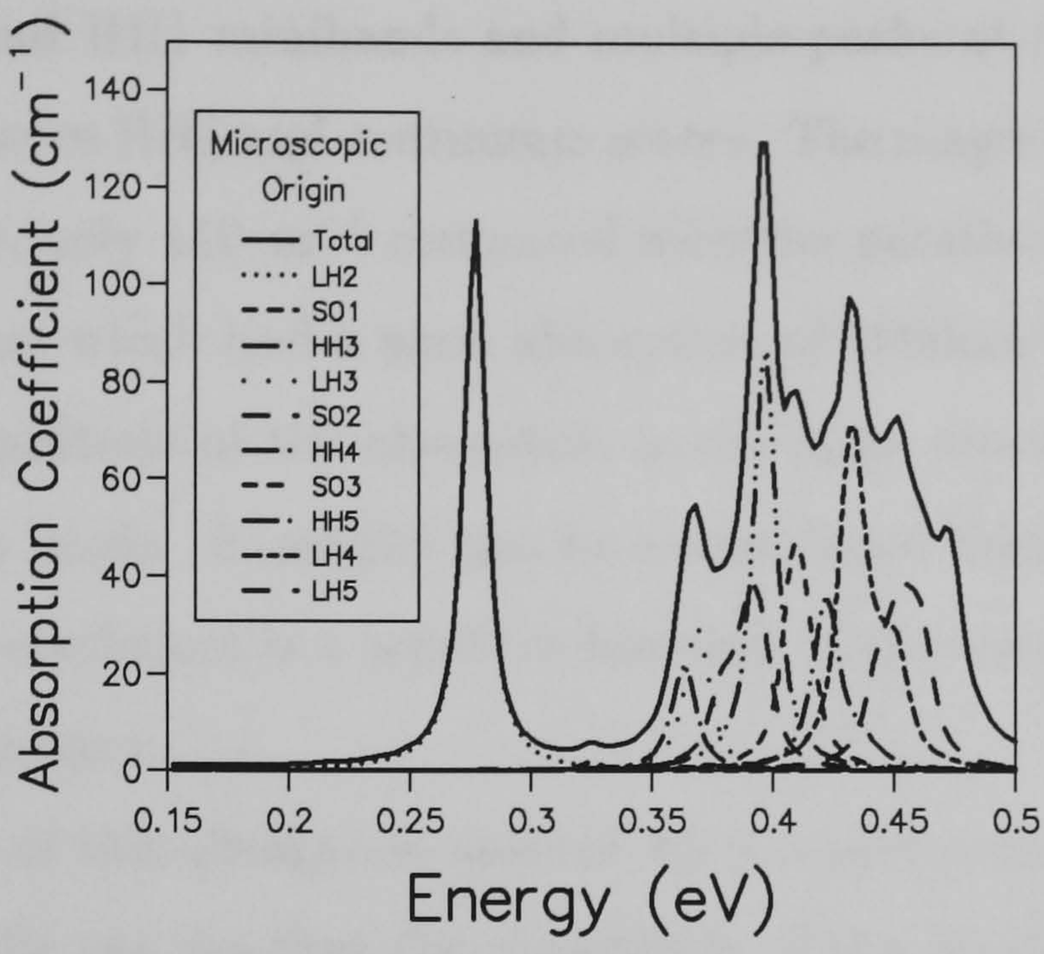


Figure 8.3: *The normal incidence absorption spectra for a 50% ,20 monolayer structure showing the microscopic origin of the peaks at 0K.*

ure 6.8. We can see that the single peak at $\approx$ 275meV is due to transitions

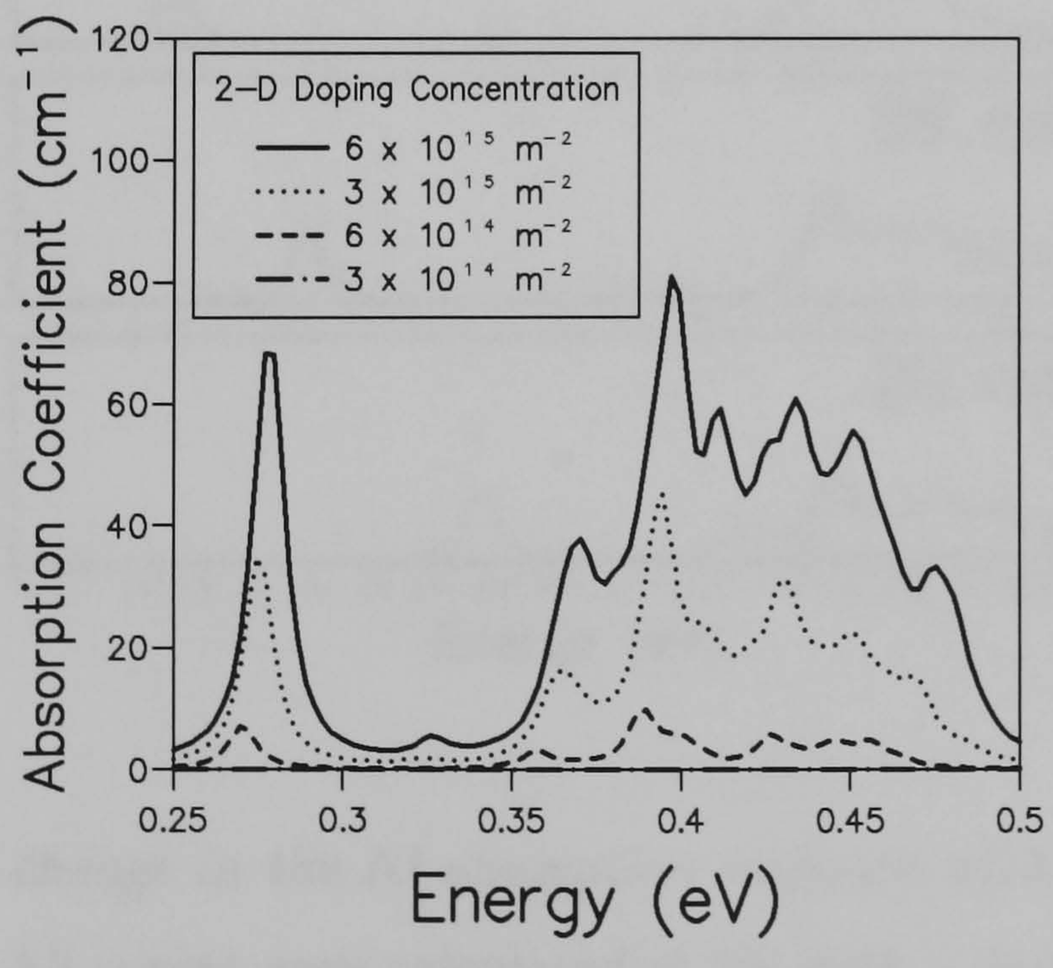


Figure 8.4: *The change of NI absorption spectra with doping concentration for a 50%, 20 monolayer structure.*

between LH2 and HH1 minibands and multiple peaks at higher energy from transitions between HH1 and continuum states. The magnitude of the absorption is very low, only 140cm^{-1} compared with the parallel incidence response for this structure which had a peak absorption of 1000cm^{-1}. Another reason for the low magnitude of the absorption is the large separation between the relevant energy levels. It should also be remembered that the magnitude of the absorption coefficient is a sensitive function of the damping constant and the photon frequency.

The change of this absorption spectra with doping concentration is shown in figure 8.4. We can see that the magnitude of the NI absorption increases

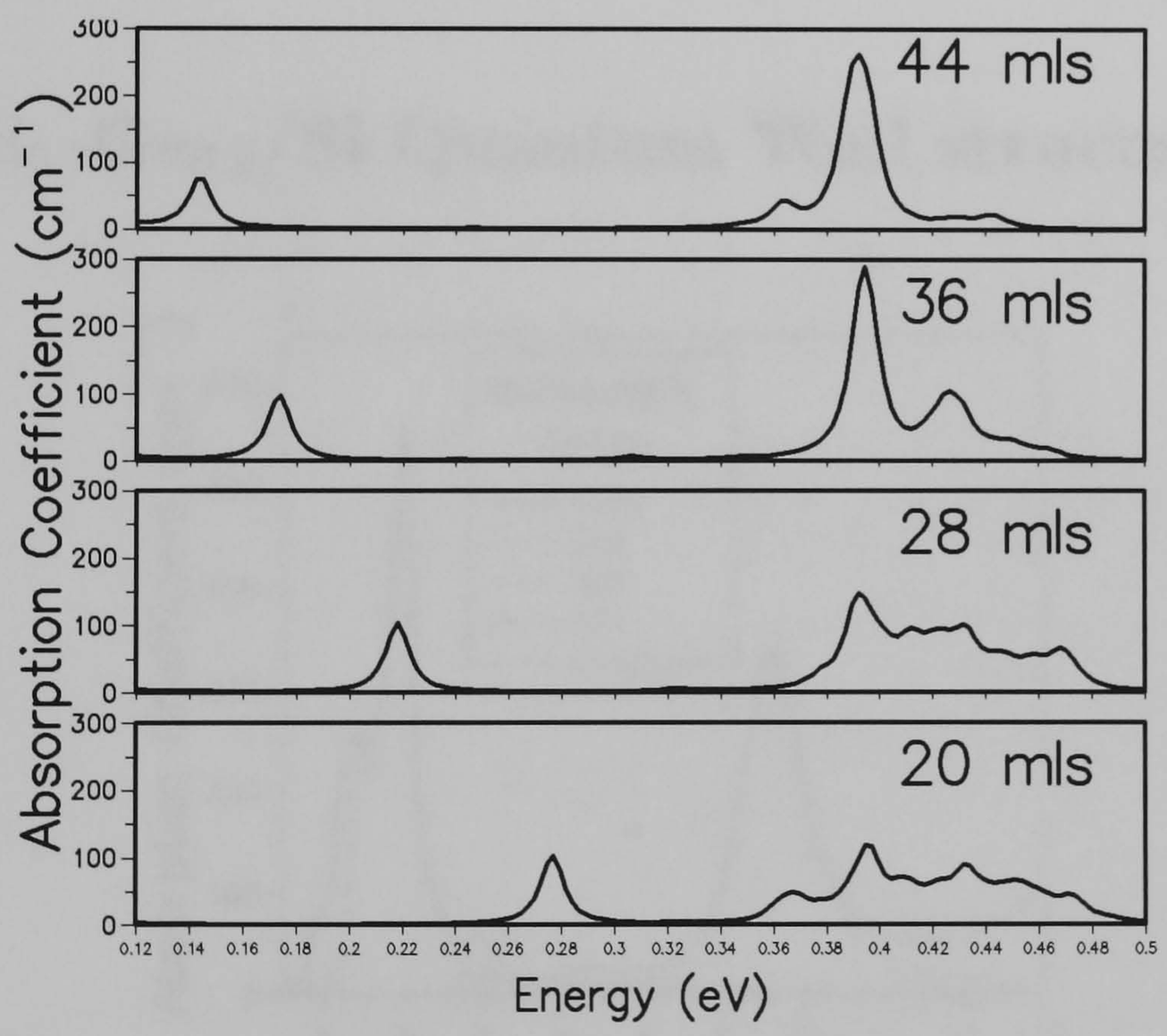


Figure 8.5: *The change in the NI absorption response with well width for a 50% structures. All curves were calculated at 0K with a damping constant of 5meV.*

rapidly with doping concentration. However, as previously mentioned, so does

the dark current, thus the increase in the magnitude of the response must be traded off against this undesirable increase in dark current.

The change in the normal incidence spectra for 50% structures as the well width is increased is shown in figure 8.5. We can see that the low energy peak which is due to transitions between LH2 and HH1 moves in energy as the LH2 miniband moves deeper into the well as the well width increases. The multiple peaks at higher energy, however, are relatively unaffected by the increase in the well width. This is because, as we can see from figure 8.3 they are due to transitions involving states that lie outside the well in the continuum. The exact shape of these peaks is a detailed function of the band structure.

8.3 $Si_{0.7}Ge_{0.3}$/Si Quantum Well structures

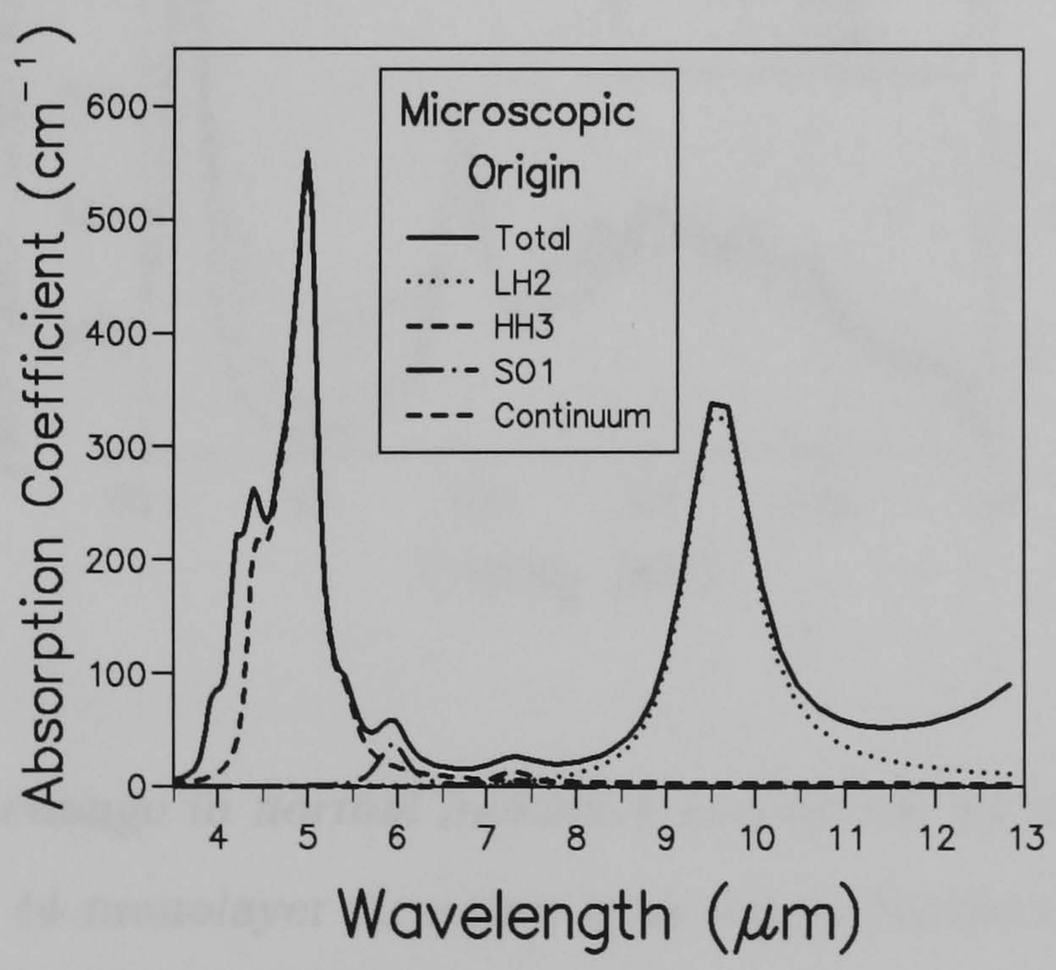


Figure 8.6: *The microscopic origin of the normal incidence response for a 30%, 44 monolayer well calculated at 0K.*

If we look at the NI absorption response for the 30%, 44 monolayer structure we discussed earlier, shown in figure 8.6, plotted against wavelength instead of energy, we see that this system has absorption peaks in both the 3-5μm and 8-12μm ranges. The miniband dispersion for this structure was presented in figure 7.14. The peak at 10μm is due to transitions between HH1 and LH2 and the lower wavelength peak between a variety of levels in the continuum. The LH2 miniband lies about 100meV within the well as shown in figure 7.17. Thus this structure could be used as a 2-colour IR detector.

The change in the absorption response of this structure as the temperature is increased is shown in figure 8.7. The peak at $\approx$ 75meV is due to transi-

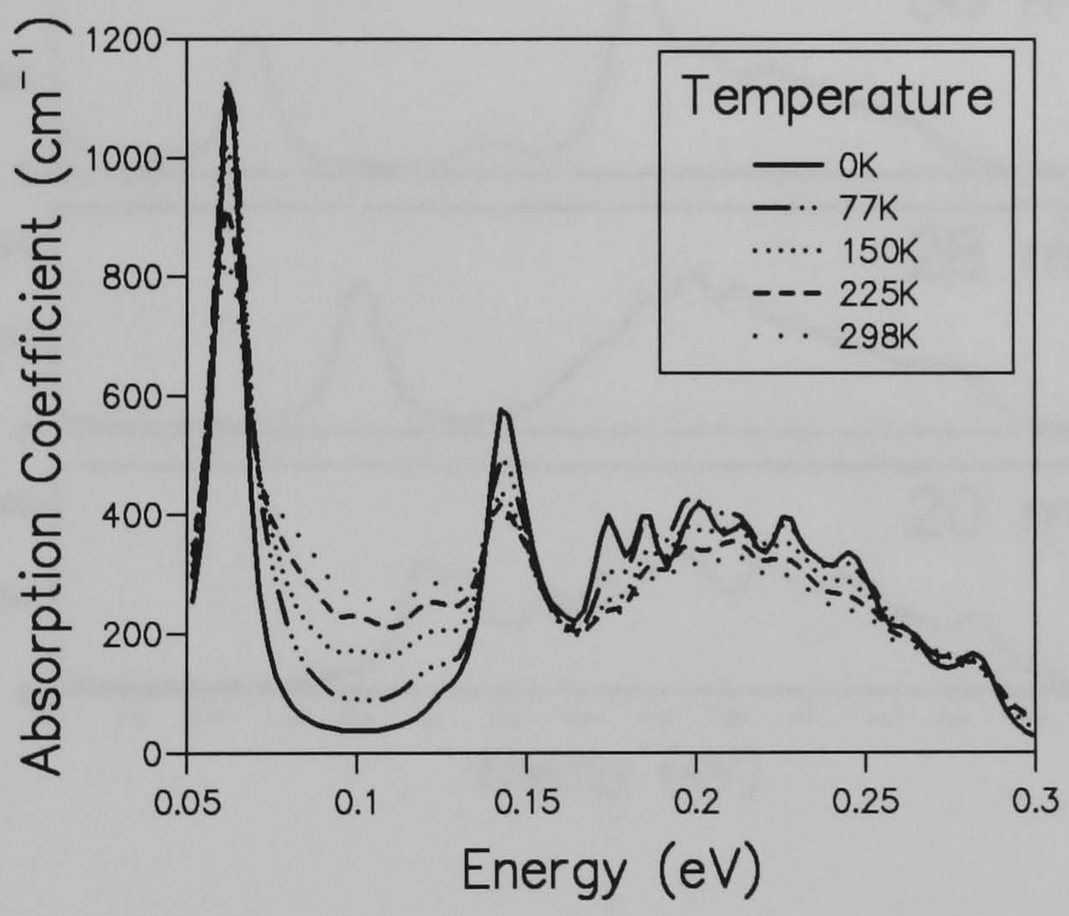


Figure 8.7: *The change in normal incidence absorption spectra with temperature for a 30%, 44 monolayer structure with doped barriers.*

tions between the LH1 and HH1 minibands. This peak is, in general, outside the energy range of interest and thus is not usually plotted on the graphs presented in this thesis. We can see that the magnitude of the main peaks decreases slightly whilst there is a significant increase in absorption between the

main peaks. At higher temperatures the carriers occupy a significantly larger volume of the Brillouin zone, and thus 'see' a much larger range of energy separations between the minibands. In particular, if we look back at the miniband dispersion for this structure presented in figure 7.14, we can see that the LH2-HH1 separation is $\approx$ 140meV at the zone centre, decreasing to $\approx$ 90meV at $k_x = 0.08(\frac{2\pi}{a})$. Thus the width of the peaks increases as the temperature rises.

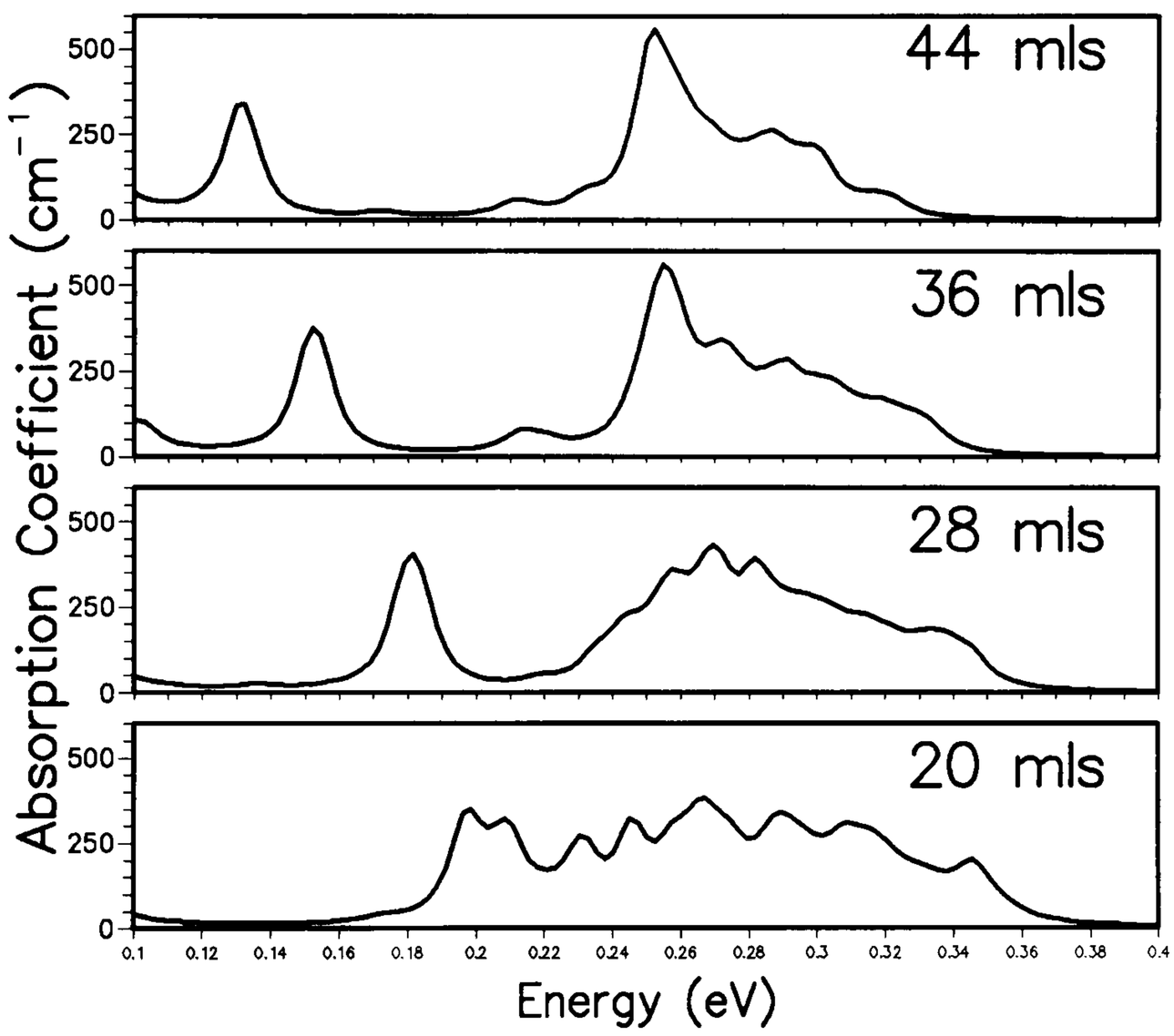


Figure 8.8: *The change in normal incidence absorption spectra with well width for 30% Ge structures, calculated at 0K with a damping constant of 5meV.*

The change in normal incidence absorption spectra with well width for these 30% structures is presented in figure 8.8. Again we see the low energy peak due to transitions between the confined LH2 miniband move to lower energy as the well width increases. The higher energy peaks again stay at approximately the same energy. The exact nature of the higher energy response

is a detailed function of the position and dispersion of the lower lying bands. The continuum bands move closer together in energy as the well width is increased resulting in the narrower high energy response at larger well widths.

8.4 $Si_{0.85}Ge_{0.15}$/Si Quantum Well structures

For the high germanium concentrations the NI absorption occurs at a higher energy than the parallel incidence response, due to the increased separation of the relevant energy levels. For the lower germanium concentrations, in partic-

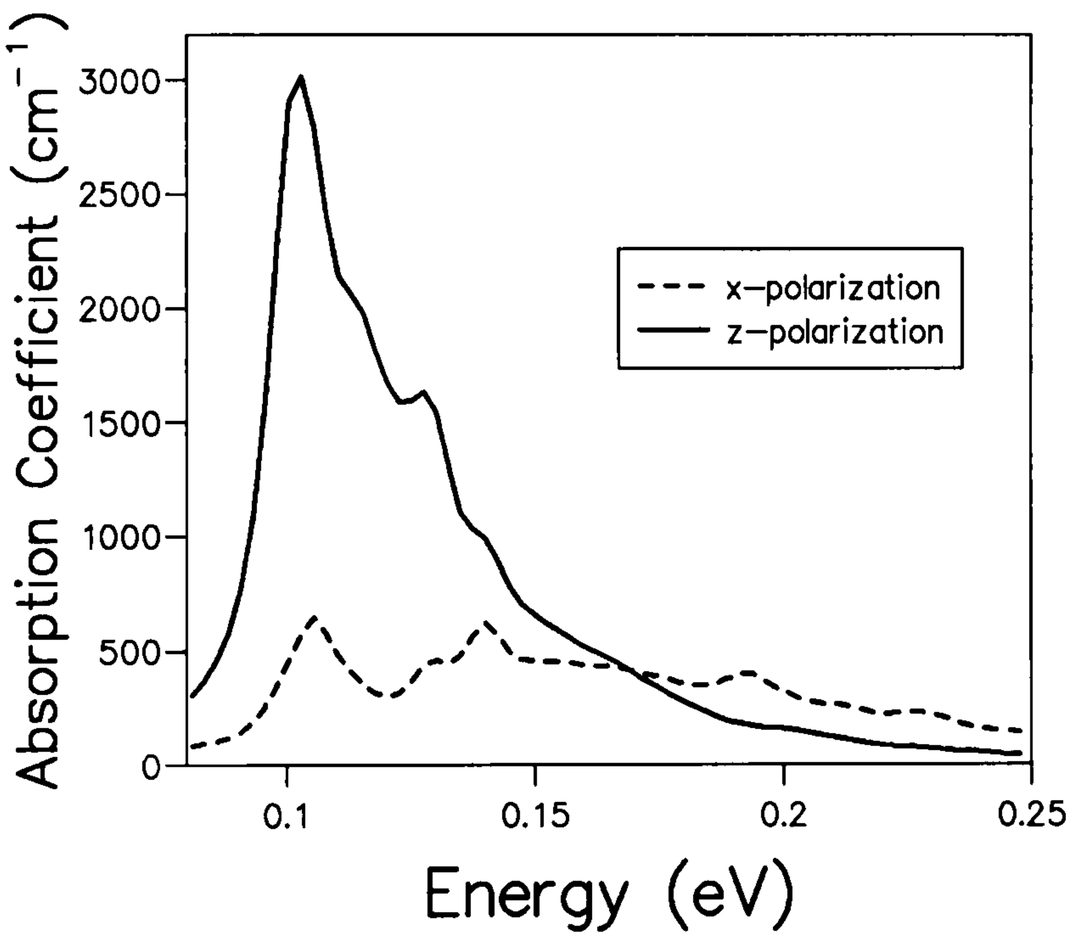


Figure 8.9: *Comparison of normal incidence (x-pol) and parallel incidence (z-pol) for a 15%, 20 monolayer structure using a damping constant of 5meV.*

ular the 15% 20 monolayer case, the normal and parallel incidence absorption occur at approximately the same energy as the SO1 and HH2 minibands lie very close in energy. In figure 8.9 we compare the normal and parallel incidence responses for a 15%, 20 monolayer well. The normal incidence response in this case is due to transitions between the SO1 and HH1 minibands. The

miniband structure for this system was presented in figure 7.22. Zanier (1995) compares the experimental parallel and normal incidence response of a similar structure and shows a 6-fold decrease in magnitude. We see approximately the same reduction in magnitude in this case.

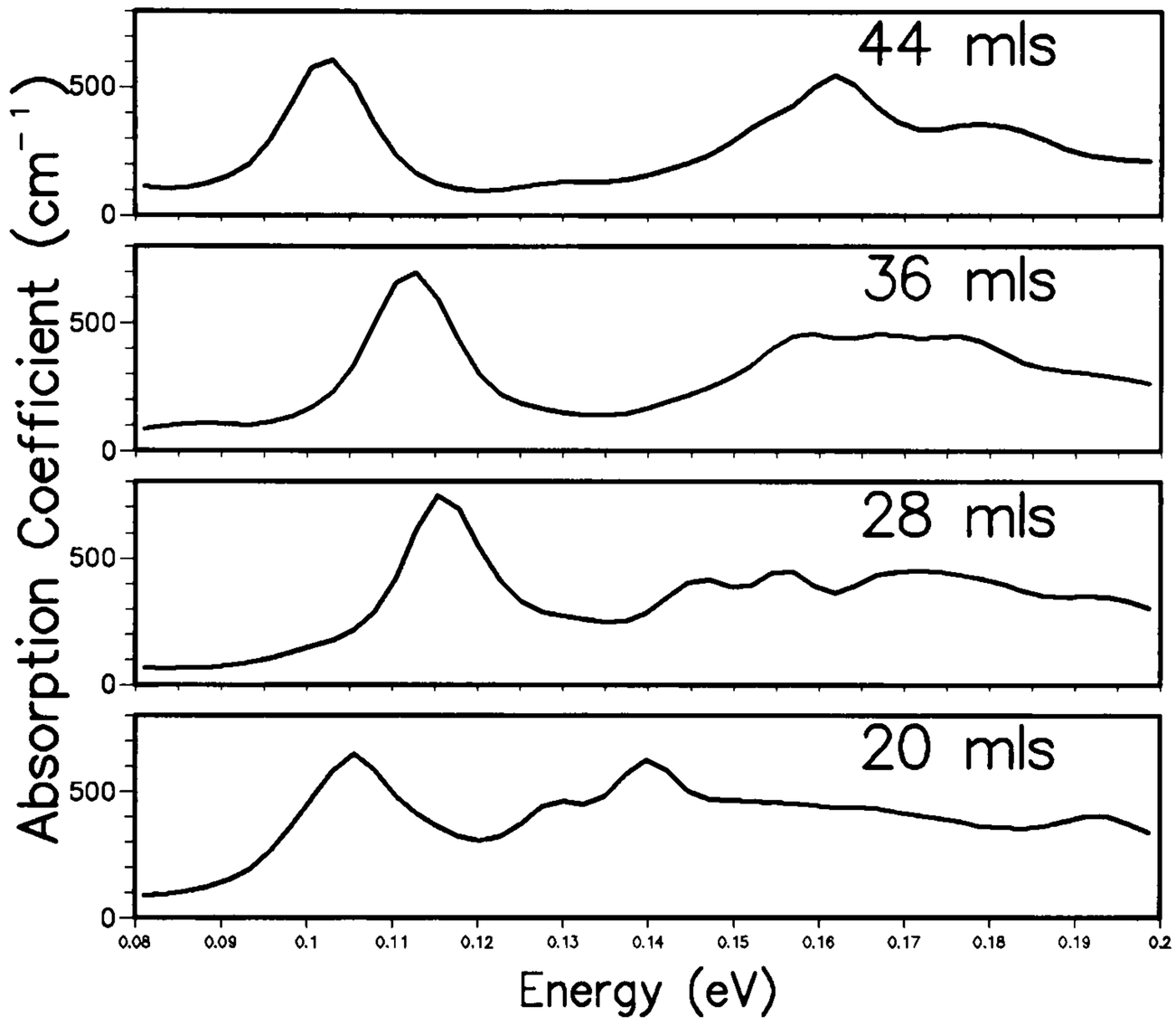


Figure 8.10: *The change in normal incidence absorption spectra with well width for 15% Ge structures, calculated at 0K with a damping constant of 5meV.*

In figure 8.10 we show the change in the normal incidence response with well width. The trend is less clear cut in this case compared with the higher germanium concentrations. For the 20 monolayer structure the peak at ≈ 105meV is due to excitations between the SO1 and HH1 minibands. Similarly for the 28 monolayer structure. For the 36 and 44 monolayer structures, however, the low energy peak is due to transitions between LH2 and HH1. The LH2 miniband moves above the SO1 miniband as shown in figure 7.26 and

once this miniband is confined within the well the peak moves to lower as the well width is increased. The very broad higher energy peaks are again a results of transitions to states outside the well. These peaks are less clearly separated from the higher energy peaks than for the higher germanium concentrations. This is because of the smaller depth of the well (126meV for 15%, c.f. 420meV for 50%). Thus the continuum levels are much closer in energy to the confined states for this lower Ge concentration.

8.5 $Si_{0.80}Ge_{0.20}$/Si Structures

We now consider in more detail the 20%, 24 monolayer structure which we compared with experimental results in Chapter 6. The absorption spectrum for this structure, with the microscopic origin of the peaks indicated, is given in figure 8.11. This diagram was calculated using a barrier doping concentration of $2.2\times10^{18}cm^{-3}$ over 60Å of the barrier. As previously discussed, barrier doping provides carriers at 0K, the most convenient temperature to deduce the microscopic origin of the peaks, whereas well doping does not. We can see that the main peak is caused by transitions between SO1 and HH1. The multiple peaks at higher energy are due to transitions between a variety of continuum states. We can see the change in the normal incidence absorption for this structure for a range of doping levels in the well in figure 8.12. Again we see the large increase in the magnitude of the absorption response as the doping concentration increases and more free carriers are available. We also see the expected shift of the peaks to higher energies as the increased doping concentration leads to a larger exchange shift thus increasing the separation of the HH1 and lower minibands.

The change in absorption spectra with temperature for the 20%, 24 mono-

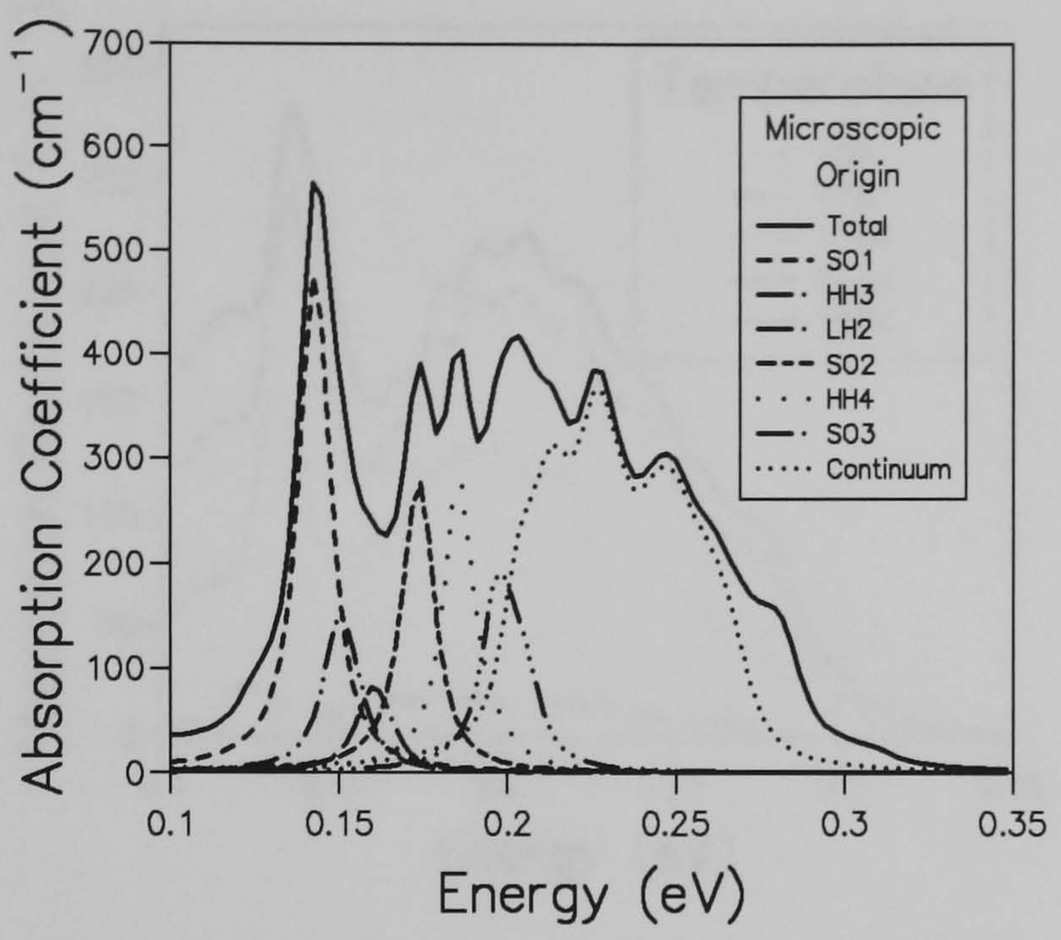


Figure 8.11: *Microscopic origin of the normal incidence response of a 20%, 24 monolayer structure, calculated at 0K with a damping constant of 5meV. Barrier doping of 2.2 $\times 10^{18}$ was assumed.*

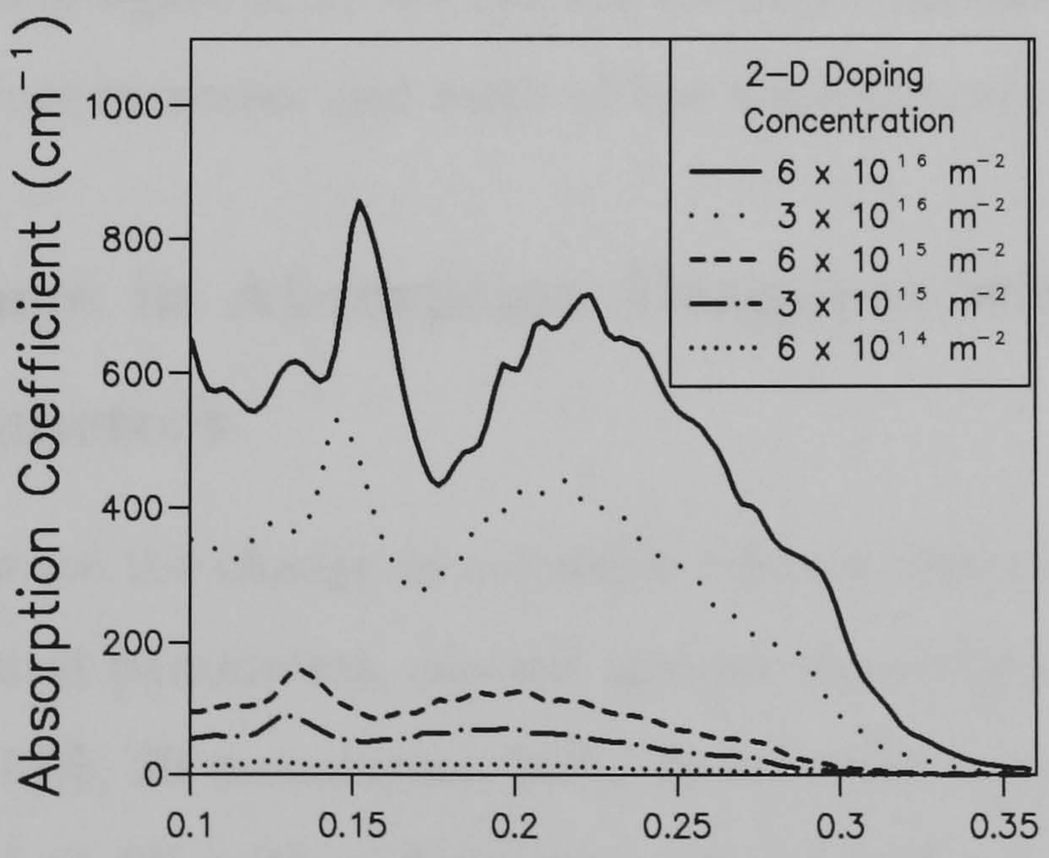


Figure 8.12: *The change in absorption response with doping concentration for a 20%, 24 monolayers structure assuming the well to be doped.*

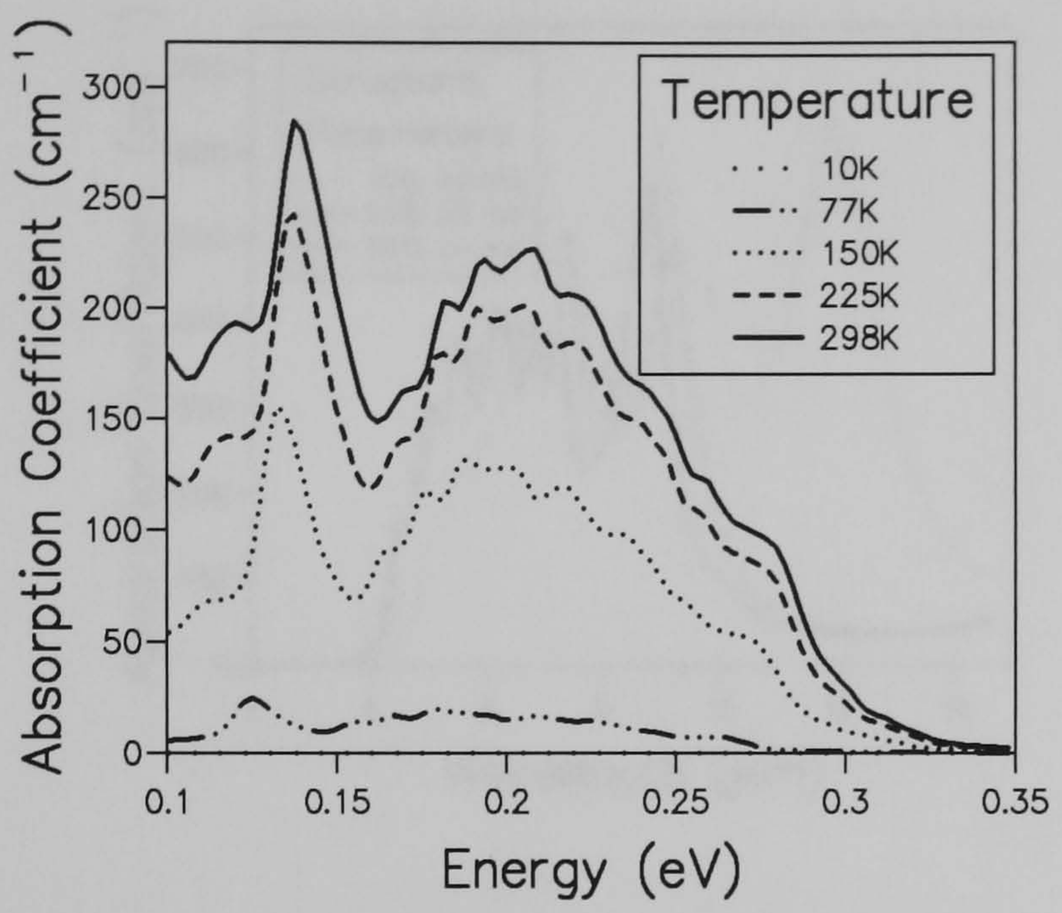


Figure 8.13: *The change in absorption spectra with temperature for the 20%, 24 monolayer structure with doped* wells.

layer structure with wells doped with $4.2\times10^{18}\text{cm}^{-3}$ across the width of the well, is presented in figure 8.12. We can see the expected increase in absorption as the temperature increases and more of the acceptors become ionized.

8.5.1 Change in Absorption Response with Structural Parameters

In figure 8.14 we see the change in normal incidence response with small variations in structural parameters, plotted against wavelength. The absorption spectra for the 15%, 20 monolayers, 20%, 20 monolayers and 20%, 24 monolayers calculated at 0K with a damping constant of 5meV are shown. The miniband dispersion for these structures was presented in figure 7.23 where it was shown that increasing the germanium concentration moves the SO1 miniband up in energy by approximately 30meV, whereas increasing the well width

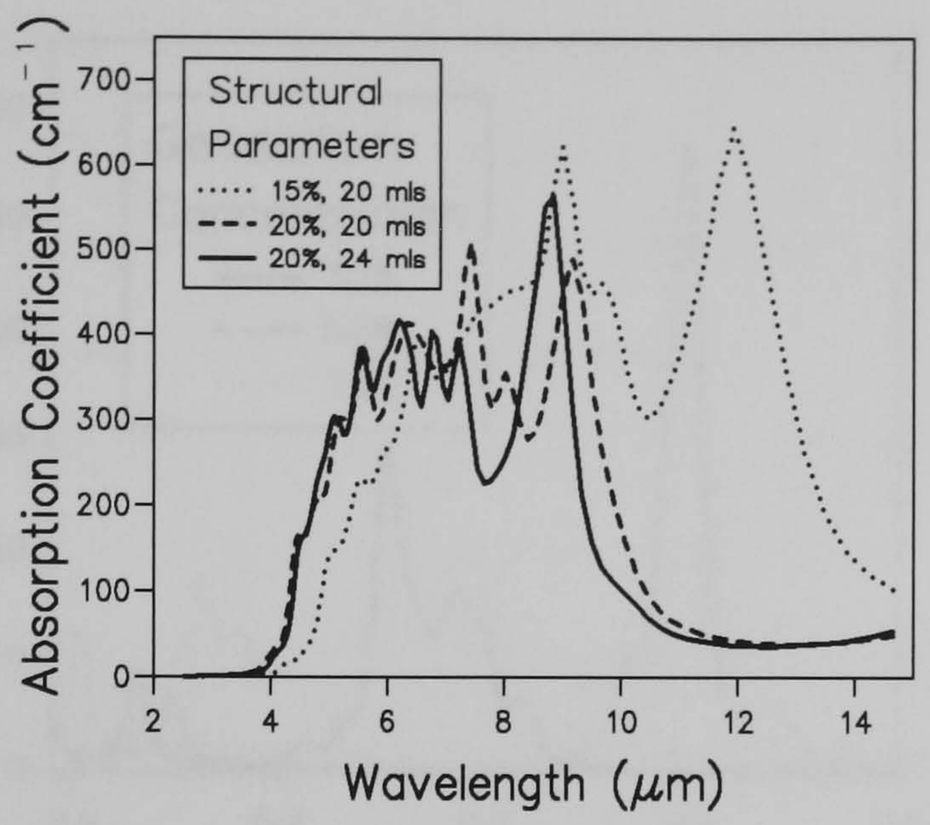


Figure 8.14: *The change in absorption response with structural parameters for 15% and 20% Ge wells plotted against wavelength.*

by four monolayers moves the SO1 miniband by only 10meV. These trends are strongly reflected in the absorption spectra where we see only a very small difference between the 20%, 20 and 24 monolayer structures but the 15%, 20 monolayer spectrum is shifted in energy by approximately 50meV. We can see that all these structures show strong normal incidence response in the 8-15μm range.

Finally we compare the normal incidence response for the full range of germanium concentrations considered. Again we plot $\omega\times$ absorption to provide a meaningful comparison of the magnitudes. The absorption spectra for 44 monolayer wide wells with 15%, 30% and 50% germanium concentrations in the alloy wells can be seen in figure 8.15. We can see that all the curves have the same basic shape with a single peak at lower energy (between 100 and 150meV) due to transitions between HH1 and the confined LH2 miniband (SO1 for 15%), and a broader peak at higher energy from transitions to continuum states. As the direct band gap of germanium is less that that of silicon ($\approx$

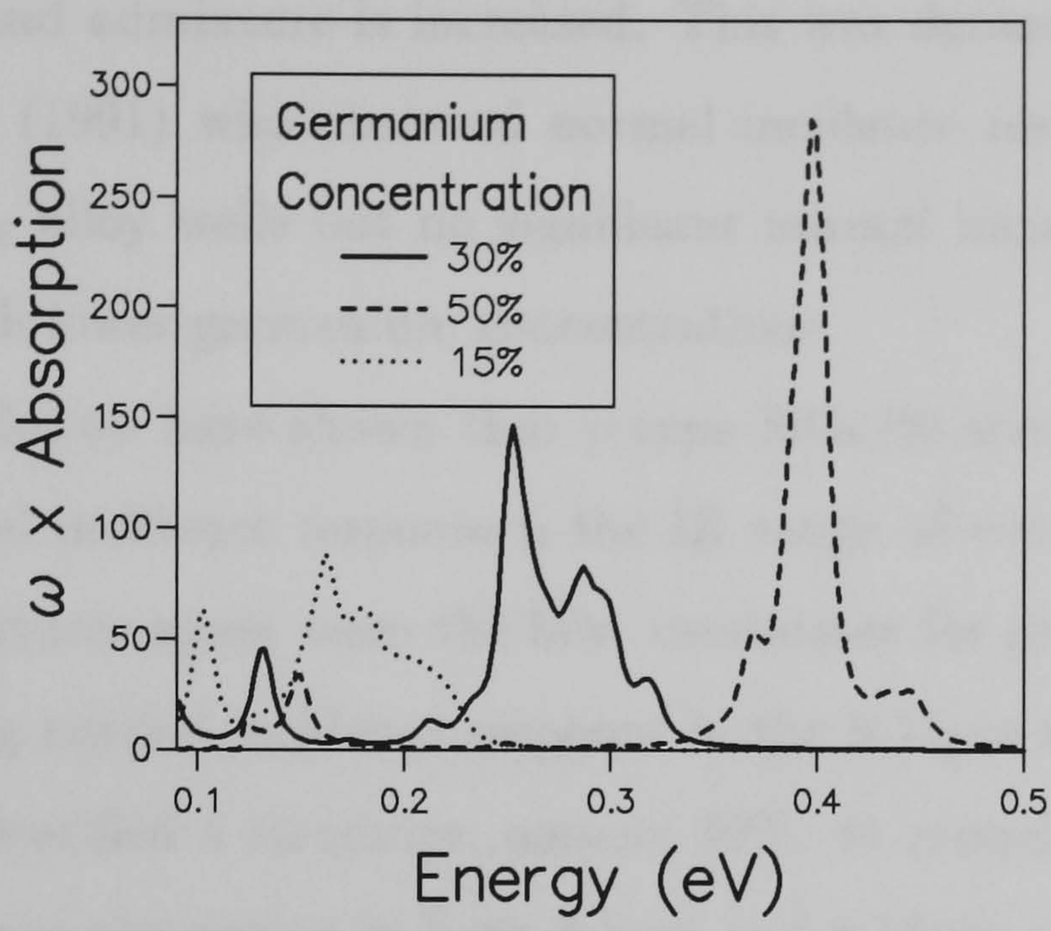


Figure 8.15: *The change in normal incidence absorption spectra with germanium concentration calculated at 0K. All wells are 44 monolayers wide and a damping constant of 5meV was used.*

0.9eV for Ge and 3.43eV for Si), higher concentrations of Ge in the alloy wells leads to smaller fundamental band gaps. Thus the conduction band is closer in energy to the top of the valence band for the 50% case, and there is a larger degree of admixture from the lowest lying conduction band, which, as previously discussed is an s-like state. This was demonstrated in figure 7.8. Contributions from this s-like state are not necessary for normal incidence absorption as the 15% structure show no noticeable conduction band admixture yet significant normal incidence response was predicted. Normal incidence response has also been observed experimentally in these low germanium concentration structure, as shown in Chapter 6. The mixing between the light and heavy hole minibands is sufficient to produce a significant normal incidence response as discussed at the beginning of this Chapter. However, the magnitude of the normal incidence response would be expected to increase as the degree of

conduction band admixture is increased. This was demonstrated experimentally by Park (1991) who observed normal incidence response in structures with $Si_{0.4}Ge_{0.6}$ alloy wells but no significant normal incidence absorption in structures with lower germanium concentrations.

To conclude, we have shown that p-type SiGe/Si structures produce significant normal incidence response n the IR range of wavelengths. The 15% and 20% structures again seem the best candidates for practical applications showing strong normal incidence response in the 8-15μm range.

We also identified a structure, namely 30%, 44 monolayers, which shown normal incidence absorption in both 3-5μm and 8-15μm ranges.

Chapter 9

Auger Recombination

There are many different types of recombination which can occur in solids. Recombination from defects, impurities and vacancies can in principle be avoided in perfect samples. However, several types of recombination are unavoidable in any situation where energy bands are present. These can be divided into two basic types, radiative and non-radiative processes. In this chapter we concentrate on the non-radiative processes.

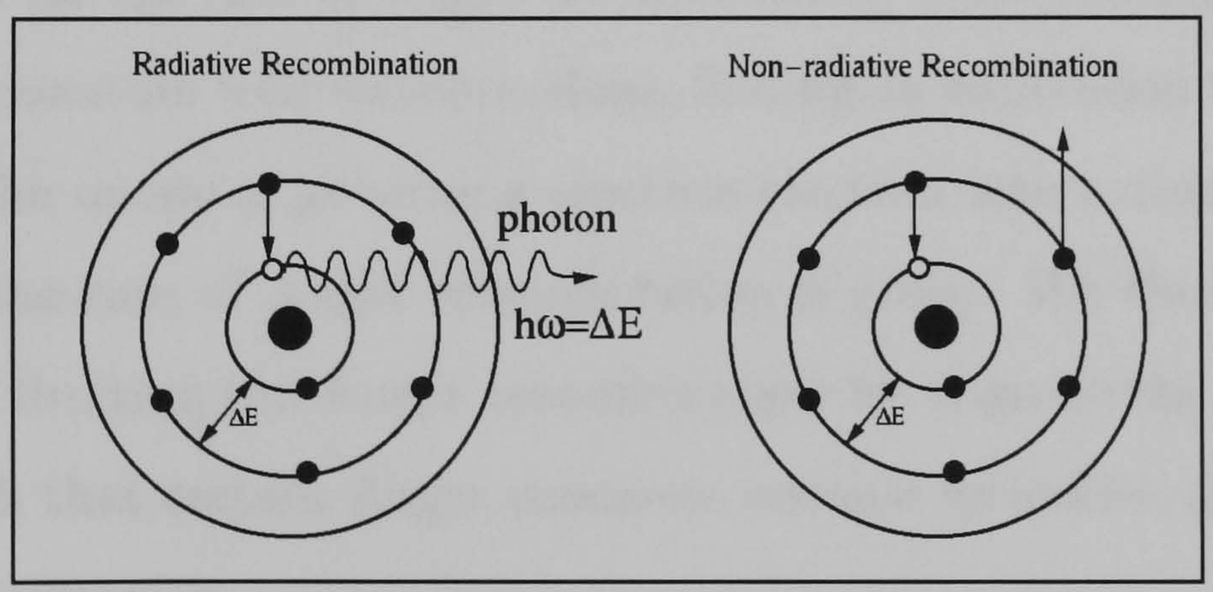


Figure 9.1: *Schematic representation of radiative and non-radiative recombination in an atom.*

Recombination of this type was first observed by Pierre Auger in 1925

whilst studying the behaviour of atoms in gases. An electron drops into a vacancy in a lower orbit and the energy released is given to a second electron which is ejected from the atom as shown in figure 9.1 or excited to a higher orbit. The analogous process in solids involves electron collisions where the electron recombines with a hole in a lower energy band and gives up its energy to a second electron which is then typically excited into a higher energy band. The rate of recombination of this type is roughly proportional to np^2 for p-type materials (assuming Boltzmann statistics) and thus for heavily doped structures will exceed the radiative recombination rate which is approximately proportional to np.

A limiting factor in device performance can be the effect of Auger Recombination, AR, which results in the recombination of the optically excited carriers before they can be removed to a useful part of the device. There have been many studies of Auger Recombination in both bulk and quantum well systems (see, for example Haug, 1983a and b, Takeshima 1984, and for a complete review of the theory of recombination in solids, Landsberg 1991) We derive an expression for the rate of Auger recombination, first considering a suitable form for the quantum well wavefunctions, finding an expression for the matrix elements of the operator governing electron-electron interactions, and an expression for the rate of Auger recombination is given. We then consider the possibility of limiting the Auger recombination by engineering the miniband structure such that certain Auger processes become forbidden on the grounds of energy conservation.

The general form of the equation for the matrix elements is

$$M = \int\int \psi_3^*(\mathbf{r_1})\psi_4^*(\mathbf{r_2})\frac{e^2}{\epsilon|\mathbf{r_1}-\mathbf{r_2}|}\psi_1(\mathbf{r_1})\psi_2(\mathbf{r_2})d\mathbf{r_1}d\mathbf{r_2} \qquad (9.1)$$

where $\psi_m(\mathbf{r})$ are the quantum well wavefunctions. Thus to calculate the matrix elements for Auger recombination we require accurate knowledge of four

electron wavefunctions (the initial and final states of both electrons) and must perform a 12-fold integration in k-space to account for all possible initial and final states of the two electrons. However, the requirement to conserve momentum and energy reduces this integration from 12 to 8-fold.

9.1 Slater Determinants

Consider the Schroedinger equation for the system as a whole

$$\mathbf{H}\Psi = E\Psi \tag{9.2}$$

where Ψ is the system wavefunction. This can be expressed as a product of the one-electron wavefunctions

$$\Psi = \phi_1(\alpha_1)\phi_2(\alpha_2)\phi_3(\alpha_3)\ldots\phi_n(\alpha_n) \tag{9.3}$$

where α represents the space and spin coordinates. However, we are discussing fermions which must obey Pauli's exclusion principle, i.e. two electrons cannot occupy the same state α, therefore an anti-symmetric representation for the electron wavefunction is required so we must again introduce Slater determinants to represent the wavefunction.

$$\Phi = C \begin{vmatrix} \phi_1(1)\phi_1(2)\phi_1(3)\phi_1(N) \\ \phi_2(1)\phi_2(2).....\phi_2(N) \\ \vdots \\ \phi_N(1) ..... \phi_N(N) \end{vmatrix} \tag{9.4}$$

where C is the normalization factor. Expanding the determinant we get

$$\Phi = C\sum_P (-1)^{p_j} P_j \phi_1(\alpha\mathbf{1})\cdots\phi_n(\alpha_\mathbf{n}) \tag{9.5}$$

where P_j is the permutation operator which permutes the assignment of the particles to states in the product expansion of the determinant, p_j is the parity

of P_j and represents the number of pair interchanges from which it can be generated. These wavefunctions form an orthonormal set, thus can be used as a basis set to expend any N-electron wavefunction i.e.

$$\Psi = \sum_j b_j \Phi_j \tag{9.6}$$

where the b_j are suitable defined constants.

9.2 Matrix Elements

We need to evaluate the matrix elements of the operators between the eigenfunctions of our unperturbed Hamiltonian. Consider an initial one-electron state $\theta_1, \theta_2, \ldots \theta_n$ and a final one-electron state $\psi_1, \psi_2, \ldots \psi_n$. The matrix elements of an operator M between the initial and final states using the expression for the wavefunction given in equation (9.3) is

$$\begin{aligned} M_{if} &= \int \theta_i M \psi_f d\tau \\ &= C_i^* C_f \sum_{j,k=1}^{n!} (-1)^{p_j + p_k} \int \cdots \int P_j \left[\phi_1^*(1) \cdots \phi_n^*(n)\right] \\ &\quad \times M P_k \left[\psi_1(1) \cdots \psi_n(n)\right] d\tau_1' \cdots d\tau_n' \end{aligned} \tag{9.7}$$

To simplify this expression we assume that θ and ψ come from the same orthonormal set, i.e. $\theta_m = \psi_m$, that the operator M is symmetrical in the particles, i.e. is unaffected by the interchange of particles on which it operates and that we restrict our calculation to two body forces, i.e. M is a sum of terms which involves at most variable of two particles. As we assume that M is symmetric with respect to the interchange of particles, the operator P_k can be brought out of the integral

$$\begin{aligned} M_{if} &= C_i^* C_f \sum_{k=1}^{n!} P_k [\int \cdots \int \sum_{l=1}^{n!} (-1)^{p_l} P_l \left[\phi_1^*(1) \cdots \phi_n^*(n)\right] \\ &\times M \psi_1(1) \cdots \psi_n(n) d\tau_1' \cdots d\tau_n'] \end{aligned} \tag{9.8}$$

where we have used the fact that $P_k^{-1}P_j$ is just another permutation operator denoted as P_l and the parity of P_l is $p_j + p_k^{-1} = p_j + p_k$. The sum over k produces $n!$ identical terms as it is applied to the expression after the integration over space and spin coordinates has been performed. Equation (9.8) therefore becomes

$$M_{if} = C_i^* C_f n! \sum_{l=1}^{n!} (-1)^{p_l} [\sum_{j=1}^{n-1} \int \theta_{l_j}^*(j)\theta_j(j)d\tau'] \times \int \theta_{ln}^*(n) M \psi_n(n) d\tau_n' \tag{9.9}$$

We introduce the notation $l_1, l_2 ... l_n$ to label states before the operation of P_l i.e. $\psi_\alpha = P_l \psi_{l\alpha}$. Thus we can write

$$P_l[\psi_1^* \cdots \psi_n(n)] = \prod_{j=1}^{m} \psi_{lj}(j) \tag{9.10}$$

as permutations acting on states are equivalent to permutation acting on particles. We can now express our matrix elements in the form

$$M_{if} = C_i^* C_f n!(A + B) \tag{9.11}$$

where

$$A = \prod_{j=1}^{N-1} \left[\int \phi_j^*(j)\psi_j(j)d\tau_j'\right] \int \phi_N^*(N) M \psi_N(N) d\tau_N' \tag{9.12}$$

is the term arising when there is no interchange of electrons, and the term describing the single interchange of electrons is

$$B = -\sum_{r,s=1;r<s}^{N} \prod_{j=1}^{N-1} \left[\int \phi_{[rs]j}^*(j)\psi_j(j)d\tau_j'\right] \int \phi_{[rs]N}^*(N) M \psi_N(N) d\tau_N' \tag{9.13}$$

where r and s label the electrons which are interchanged, thus $lj \rightarrow [rs]_j$, and the restriction $r < s$ must be included to avoid double counting of interchanges. We need only include these two terms as our earlier assumption states that only two-body forces would be considered.

To find an expression for the normalization constants C_i and C_f we consider the case where $M = I$, the identity operator. We also assume that $\theta_i = \psi_i, (i = 1, ...n)$, i.e. there are no electron transitions. Each term in the expression for B will now contain orthogonality integrals, therefore $B = 0$. As the determinental wavefunctions are identical $C_i = C_f$ and as the wavefunctions are already considered to be normalized, $A = 1$. Therefore

$$M_{if} = |C|^2 n! = 1 \tag{9.14}$$

$$|C|^2 = \frac{1}{n!} \tag{9.15}$$

For Auger type transitions we need to consider the case where $\theta_i = \psi_i$ except for $i = a, b; a < b$. The matrix element may therefore be expressed as

$$M_{if} = \int \int [\theta_a^*(1)\theta_b(2)^* - \theta_b^*(1)\theta_a^*(2)] M(1,2)\theta_a(1)\theta_b(2) d\tau_1' d\tau_2' \tag{9.16}$$

where the dummy variables k and l have been replaced by 1 and 2.

9.3 The Coulomb Potential

The operator M represents the two-body perturbation energy of the Coulomb interaction. This can be written as

$$M(1,2) \equiv M(\mathbf{r}_1, \mathbf{r}_2) = \frac{e^2 e^{-\lambda|\mathbf{r}_1 - \mathbf{r}_2|}}{\epsilon|\mathbf{r}_1 - \mathbf{r}_2|} \tag{9.17}$$

where $\mathbf{r_1}$ and $\mathbf{r_2}$ are the coordinates of the two interacting electrons, λ is the screening radius (Landsberg 1991) and ϵ is the dielectric constant.

The Fourier transform of this perturbing potential is

$$M(\mathbf{r}_1, \mathbf{r}_2) = \sum_k M_k e^{ik.|\mathbf{r}_1 - \mathbf{r}_2|} \tag{9.18}$$

To determine the coefficients M_k we consider the expression

$$\frac{1}{\Omega^2}\int_\Omega\int_\Omega M(\mathbf{r}_1,\mathbf{r}_2)e^{-ik.|\mathbf{r}_1-\mathbf{r}_2|}d\tau_1 d\tau_2 =$$
$$\frac{1}{\Omega^2}\sum_{k'} M_{k'}\int_\Omega e^{i(k'-k).\mathbf{r}_1}d\tau_1\int_\Omega e^{i(k-k').\mathbf{r}_2}d\tau_2 \qquad (9.19)$$

where we have used equation (9.18). If we assume periodic boundary conditions we have

$$\int_\Omega e^{i\mathbf{k}.\mathbf{r}}d\tau = \Omega\delta_{\mathbf{k},0} \qquad (9.20)$$

Thus

$$\begin{aligned} M_k &= \frac{1}{V}\int M(\mathbf{r})e^{-i\mathbf{k}.\mathbf{r}}d\tau \\ &= \frac{e^2}{\epsilon V}\int\frac{e^{-(\lambda\mathbf{r}+i\mathbf{k}.\mathbf{r})}}{|\mathbf{r}|}d\tau \end{aligned} \qquad (9.21)$$

where $\mathbf{r}=\mathbf{r_1}-\mathbf{r_2}$. The volume element $d\tau = r^2 dr dx d\phi$ where dx is the cosine of the angle between $\mathbf{k}$ and $\mathbf{r}$ and ϕ is the azimuthal angle. Integrating equation (9.21) we get

$$\begin{aligned} M_k &= \frac{2\pi e^2}{\epsilon\Omega}\int_0^\infty re^{-\lambda\mathbf{r}}\int_{-1}^1 e^{-i\mathbf{k}.\mathbf{r}x}dxdr \\ &= \frac{2\pi e^2}{\epsilon k\Omega}i\int_0^\infty e^{-\lambda\mathbf{r}}\left[e^{-i\mathbf{k}.\mathbf{r}}-e^{i\mathbf{k}.\mathbf{r}}\right]dr \\ &= \frac{2\pi e^2}{\epsilon k\Omega}i\left[\frac{1}{\lambda+ik}-\frac{1}{\lambda-ik}\right] \\ &= \frac{4\pi e^2}{\epsilon\Omega(k^2+\lambda^2)} \end{aligned} \qquad (9.22)$$

Thus we have an expression for the matrix elements of our operator and can now solve equation 9.16.

9.4 Rate of Auger Recombination

The physically observable quantity is the Auger rate, i.e. the number of transitions per unit time. To calculate this we use Fermi's Golden Rule (Landsberg

1991), which provides the probability that an electron will undergo a transition due to the perturbing operator. For an electron in an initial state i, the probability of a transition to the final state f per unit time is

$$T_{if} = \frac{2\pi}{\hbar}|M_{if}|^2\delta(E_f - E_i) \tag{9.23}$$

The Brillouin Zone is sampled using a Monte-Carlo integration method (Hammersley 1964) based on the technique employed by Kane (1967). It is assumed that the dielectric constant, ϵ, is a function of $\mathbf{k}$ but the dependence on frequency is not included. This should be a very good approximation in the range of energies considered here. For full details of the calculation procedure see Williams (1995)

9.5 Energy condition for the prevention of Auger Recombination

It is often stated in the literature that the dominant process of Auger recombination in bulk materials is the CHCC process (Chiu 1982), i.e. the initial electrons are in the conduction band, and one recombines with a hole at the top of the valence band and the resulting energy is given to the second electron which is promoted to a higher energy state within the conduction band. This process is shown in figure 9.2.

In our situation of p-type quantum wells the conduction band is too far away in energy to be involved in such recombination (the energy released by recombination of carriers in two valence minibands is insufficient to promote electrons to the conduction band). Only the HH1 miniband has significant hole population before absorption occurs. The process of intervalence band absorption produces a hole in a lower lying miniband as shown in figure 9.3.

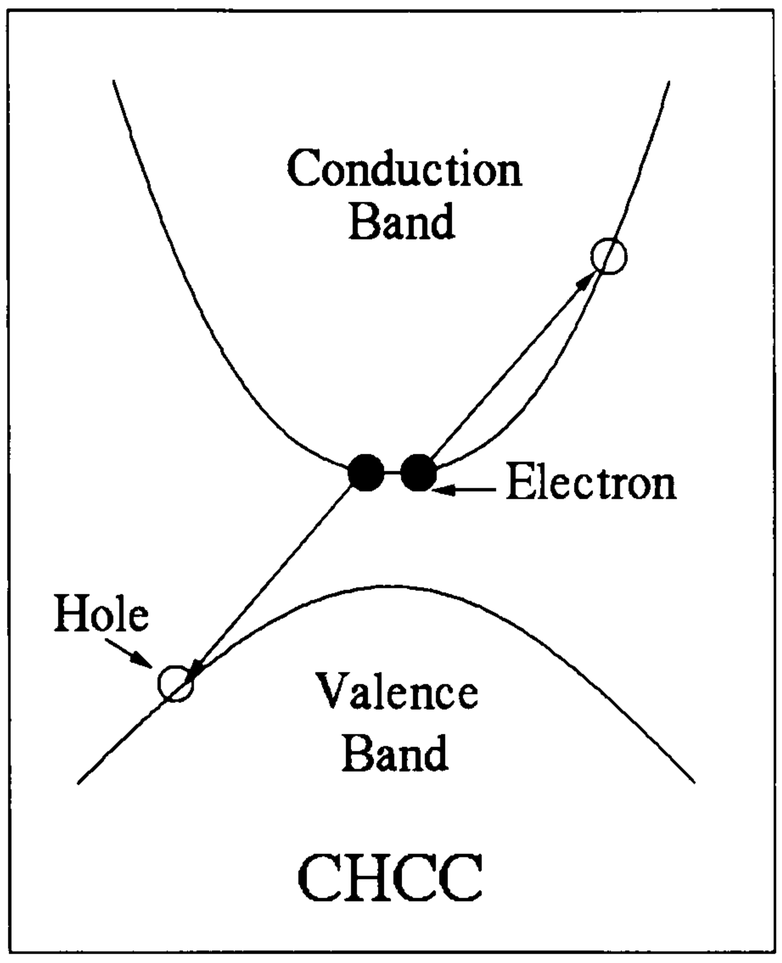


Figure 9.2: *The CHCC process in bulk system.*

In the following discussion we label the minibands as A, B, C and D. The HH1 miniband is always state A and the LH1 similarly state B. The nature of states C and D, however, depends on the structure being discussed. For low germanium concentrations with narrow wells, state C is the SO1 miniband, and state D the HH2 miniband. For higher germanium concentrations and for wider wells the order of these states changes. State C becomes HH2 and state D is, in general, the LH2 miniband.

If we assume that the light is incident parallel to the growth direction, the absorption occurs between the HH2 and HH1 minibands. In figure 9.3 we have a schematic representation of just one of the many auger processes that can occur. This process is assumed to occur in a structure where the HH2 miniband is state C. An electron is excited from HH2 to HH1 by an incoming photon. The hole produced is the charge carrier which we wish to collect. However, the electrons in the populated LH1 miniband can recombine non- radiatively resulting in an electron-hole recombination with the resulting energy released

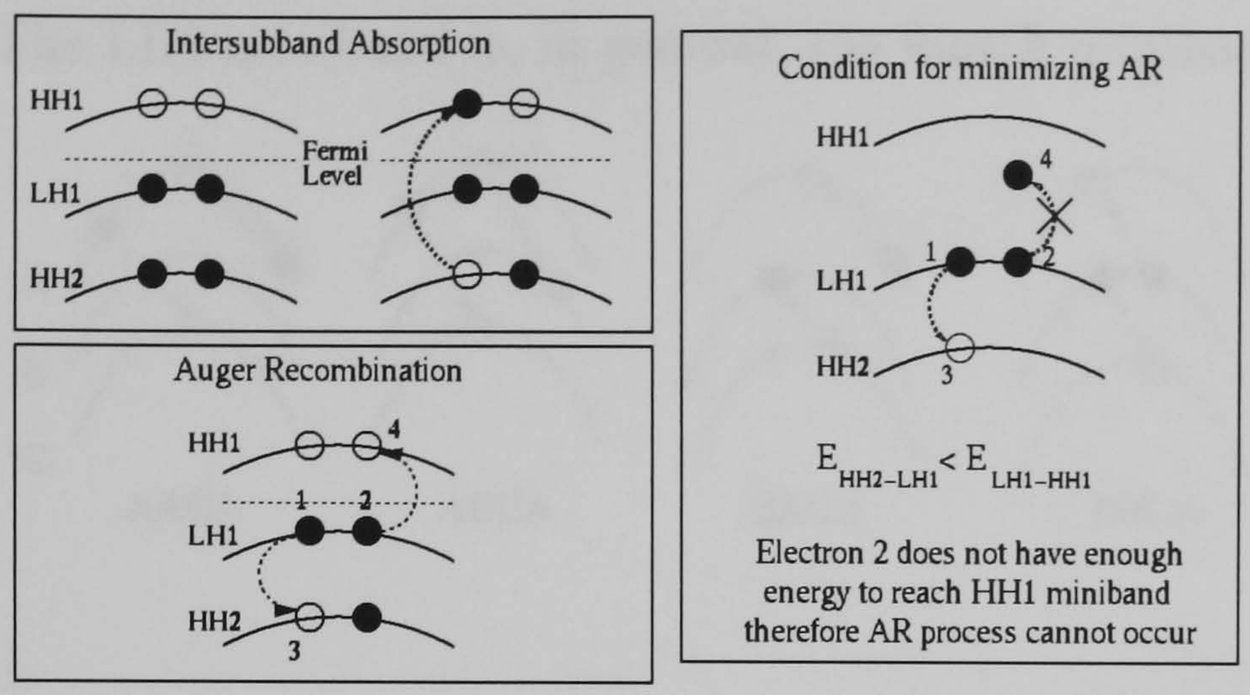


Figure 9.3: *Schematic representation of intervalence band absorption and the subsequent Auger recombination. Also shown is the energy condition which prevents this particular Auger process from occuring.*

$\mathbf{k_1}$	$\mathbf{k_2}$	$\mathbf{k_3}$	$\mathbf{k_4}$
A	A	C	A
A	B	C	A
B	A	C	A
B	B	C	A

Table 9.1: The possible combinations of transitions that can occur when the photo-hole is assumed to be in miniband C initially.

being given to a second electron in the LH1 miniband which is promoted to the HH1 miniband (providing that a suitable state is available to accept it). A list of all possible transitions assuming the photo-excited hole to be in the third miniband is given in table 9.5. The electron with initial wavevector $\mathbf{k_1}$ recombines with a hole with wavevector $\mathbf{k_3}$ and $\mathbf{k_2} \rightarrow \mathbf{k_4}$. These processes are presented schematically in figure 9.4. The labeling of the minibands in this diagram has been kept very general. These transitions are also relevant for normal incident light when the absorption occurs between the LH2 and HH1

minibands. The LH2 miniband is, in general, the fourth miniband.

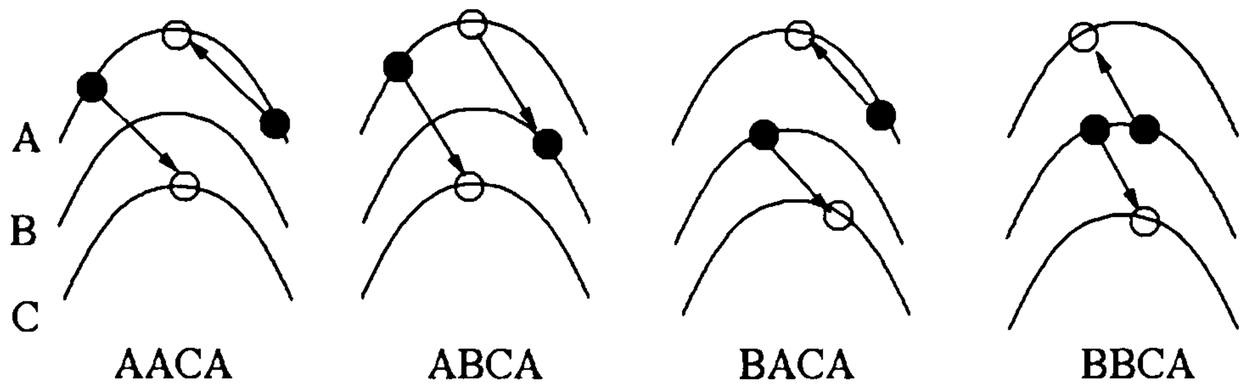


Figure 9.4: *Schematic representation of possible Auger recombination routes that can occur if the photo-hole is considered to be in miniband C initially.*

The transitions outlined in table 9.5 are also relevant for normal incidence absorption in structures with low germanium concentrations when the absorption occurs between the SO1 and HH1 minibands, labelled C and A respectively.

For the situation when the photo-excited carrier is in state D, i.e. parallel incidence light in structures with low germanium concentrations or normal incidence light in higher Ge concentration systems, there are a greater number of possible processes. These are listed in table 9.5 and shown schematically in figure 9.5. Obviously, which of these transitions is possible is governed by the energy separation of the minibands across the Brillouin zone.

In theory, of course, it is also possible for transitions involving holes in the LH1 miniband, which would give rise to further combinations. However, in the structures considered in this thesis, and for practical temperature and doping concentrations, there is unlikely to be significant hole population of the LH1 miniband at any time. Thus these transitions are not considered here.

Let us consider the process BBCA as shown in figure 9.4. This process can be limited if the miniband structure is engineered such that

$$E_B - E_C < E_B - E_A \tag{9.24}$$

k_1	k_2	k_3	k_4
C	C	D	A
C	B	D	A
C	A	D	A
B	C	D	A
B	B	D	A
B	A	D	A
A	C	D	A
A	B	D	A
A	A	D	A

Table 9.2: The possible combinations of transitions that can occur when the photo-hole is assumed to the in the D miniband initially.

or, more instructively

$$\Delta E_{HH1-LH1} < \Delta E_{LH1-HH2} \tag{9.25}$$

and thus the energy generated by the electron-hole recombination is not sufficient for the second electron to reach the HH1 miniband and the recombination process cannot occur. Thus we examine the separation of the minibands at the zone centre in order to find a suitable structure in which this process would be inhibited.

The change of zone centre separation with well width for the top five minibands for 15% and 30% Ge concentrations in the well are presented in 7.17 and 7.26. For structures in which the HH2 miniband lies below the SO1 miniband, the processes should be labelled BBDA as the photo-hole will be generated in band D for parallel incident light. It should be remembered that these are zone centre separations and away from the zone centre the miniband separations may be reduced. Thus these diagrams are the upper limit of the energy

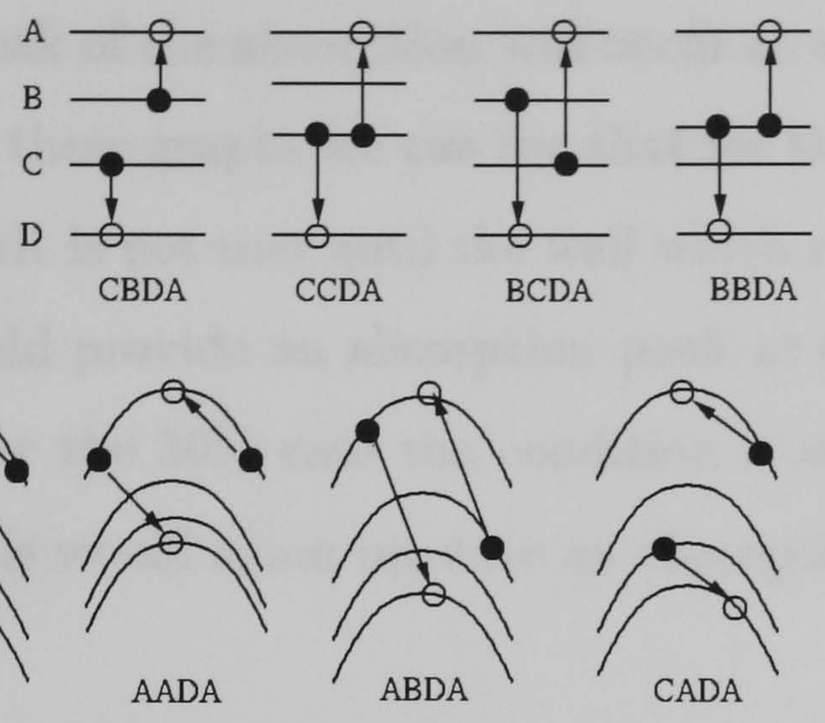


Figure 9.5: *Schematic representations of the possible auger recombination routes that can occur considering the photoexcited hole to be in miniband D initially.*

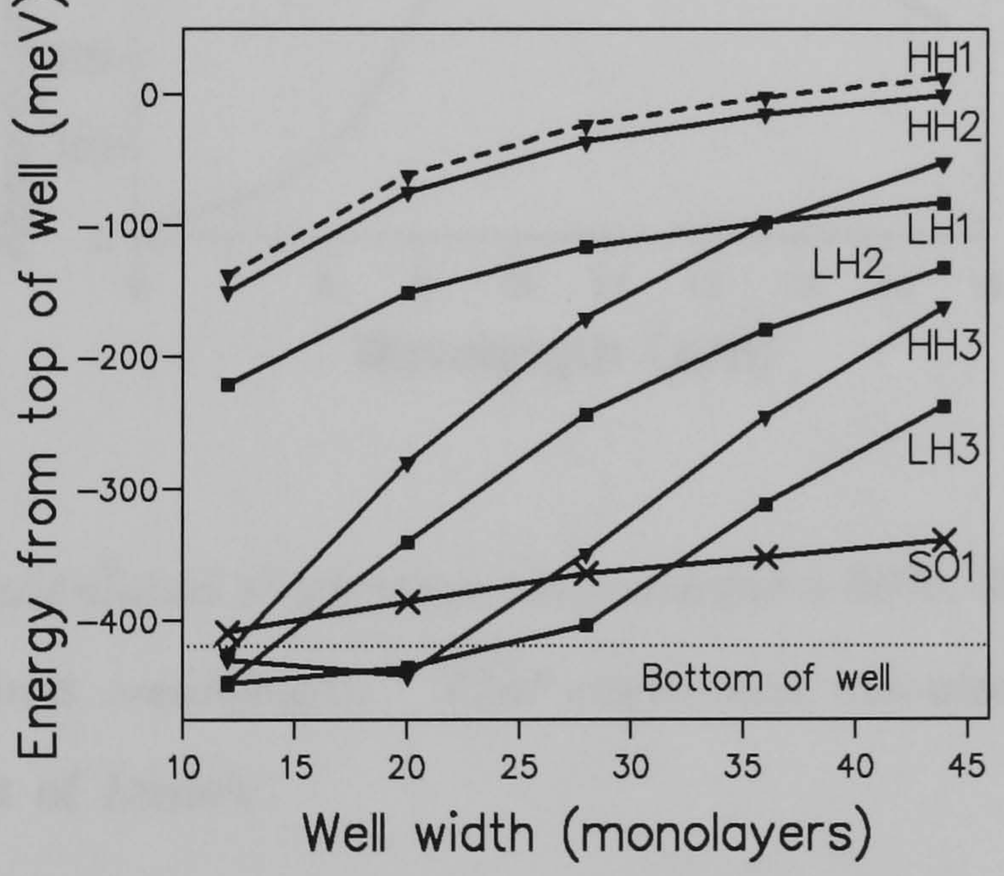


Figure 9.6: *Change in zone centre separation with well width for 50% Ge structures. The dashed line indicates the position of the HH1 miniband with an exchange shift of 13meV included.*

separation and the peak of the absorption will occur at a slightly lower energy. From examination of these graphs we can see that for the 15% case the condition for minimizing AR is not met until the well width reaches 44 monolayers. Such a structure would provide an absorption peak at $\approx$ 50meV, outside the region of interest. For the 30% case the condition is met at a well width of 36 monolayers but this would again produce an absorption peak at too low an energy.

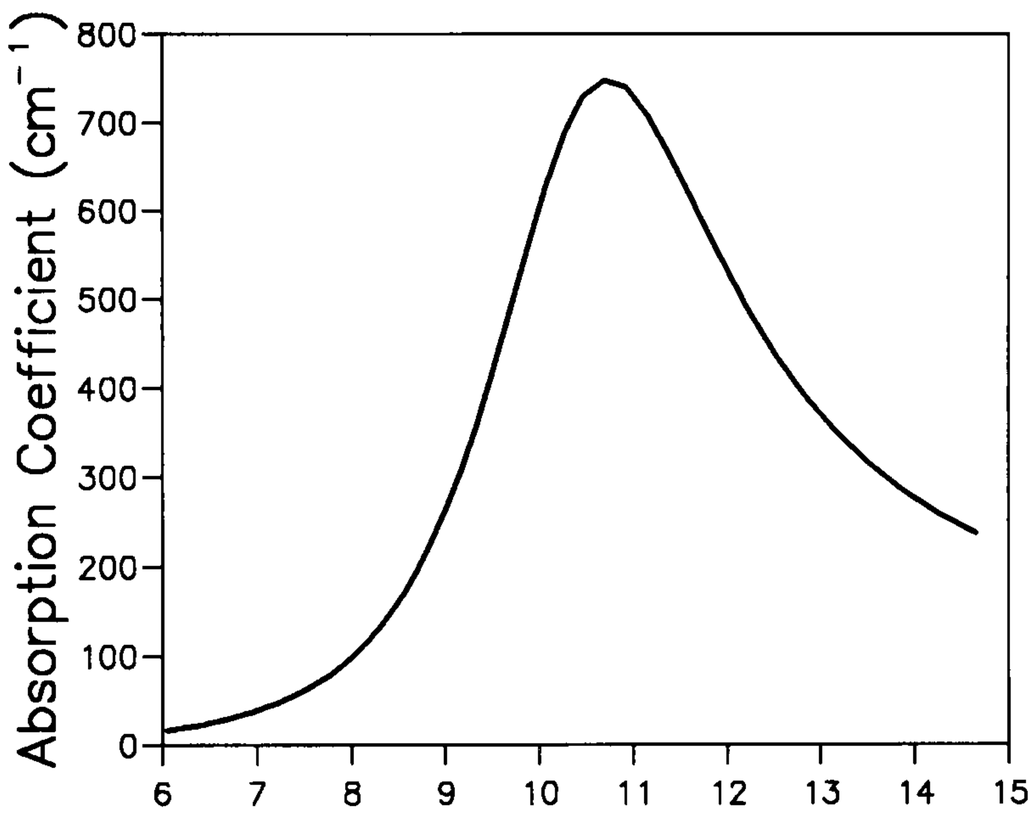


Figure 9.7: *The calculated absorption response for a 50%, 32 monolayer structure plotted against wavelength. This curve was calculated at 0K using a damping constant of 18meV.*

For the 50% case the trend of the zone centre separation with well width is reproduced again for convenience in figure 9.6. The energy condition is met by all well widths above 28 monolayers. We can also see that a well width of 32 monolayers would produce a peak at around 120meV (10μm) whilst having a band structure favourable to minimizing AR. In figure 9.7 we can see the calculated absorption spectra for such a structure and see that indeed we do

achieve a peak at $\approx 10\mu$m, or $\approx$ 120meV. The miniband structure for this system is shown in 9.8. We can see from this diagram that the process BBCA

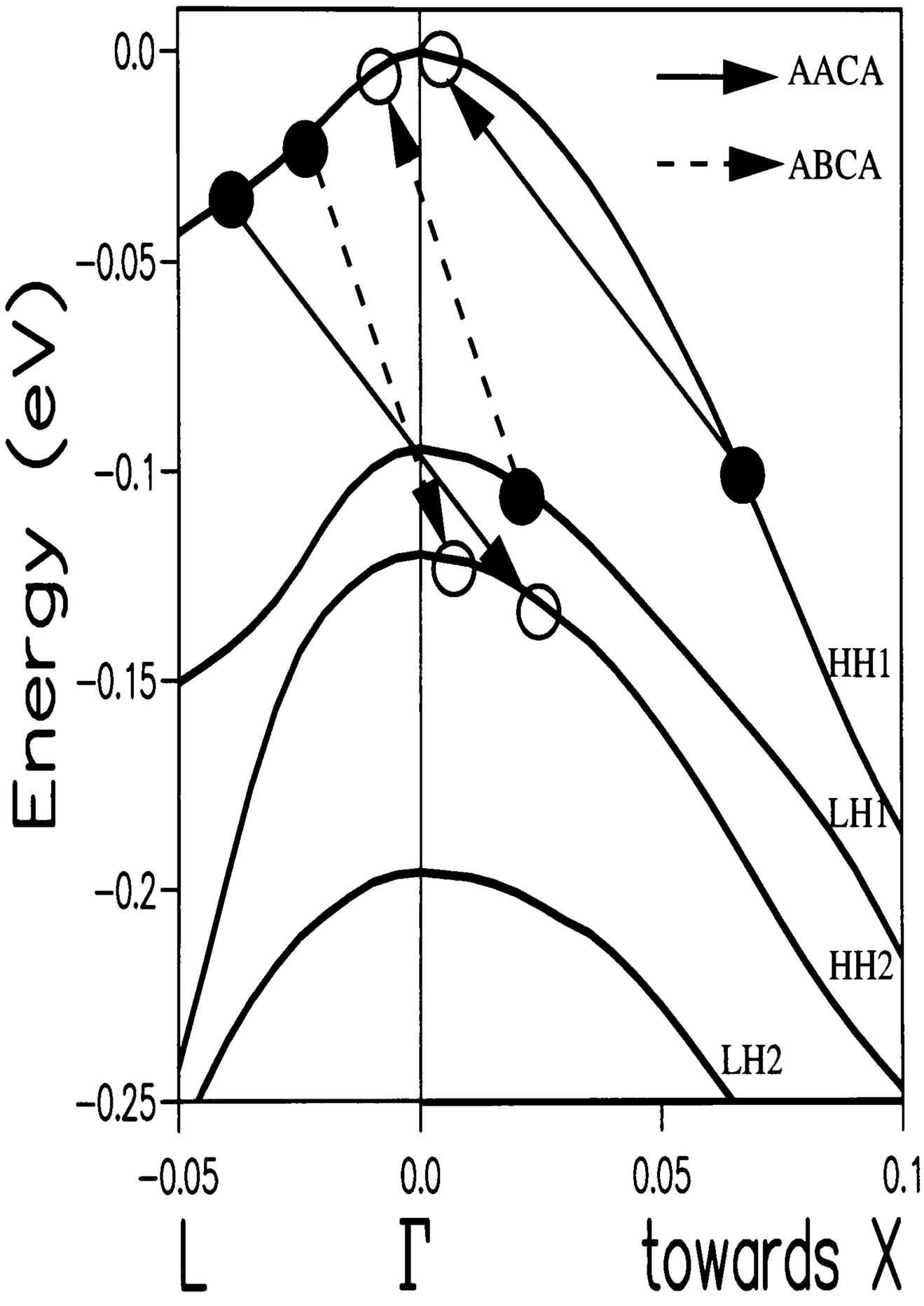


Figure 9.8: *Miniband dispersion along the symmetry lines L-Γ-X for the 50%, 32 monolayers structure. Two of the Auger processes that can still occur are indicated.*

cannot occur anywhere in the zone in the directions shown in figure 9.8 as the condition for the conservation of energy and momentum cannot be met. However, other process, such as the two shown in figure 9.8 can still occur. To reduce the likelihood of the remaining transitions, other band gaps could be engineered, e.g. the CBDA process in figure 9.5 could be minimized in a structure where (assuming state C to be the HH2 miniband and state D to be LH2 and the light to be incident normally, so that the principle absorption mechanism is between the LH2 and HH1 minibands)

$$\Delta E_{HH2-LH2} < \Delta E_{HH1-LH1} \tag{9.26}$$

Processes such as AACA which rely on the curvature of the minibands to meet the energy and momentum conservation requirements are much harder to eliminate that the other processes.

9.6 Results of Auger Recombination Calculations

Preliminary results from full-scale Auger recombination calculations performed by C. Williams (1995) show that no one process is dominant. In fact the rate for all the processes outlines above is very similar. In figure 9.9 we can see the rate of Auger recombination for a 15%, 20 monolayer structure for a total of all the processes. A range of values, from 0 to 1, were assumed for the occupancy of the HH2 or SO1 miniband. This structure has not been optimized to prevent any particular process. We can see that the rates for the normal and parallel incidence case are very similar and that the rate increases linearly with the occupation of the appropriate miniband, as would be expected.

In figure 9.10 we can see the breakdown of the total Auger rate into its

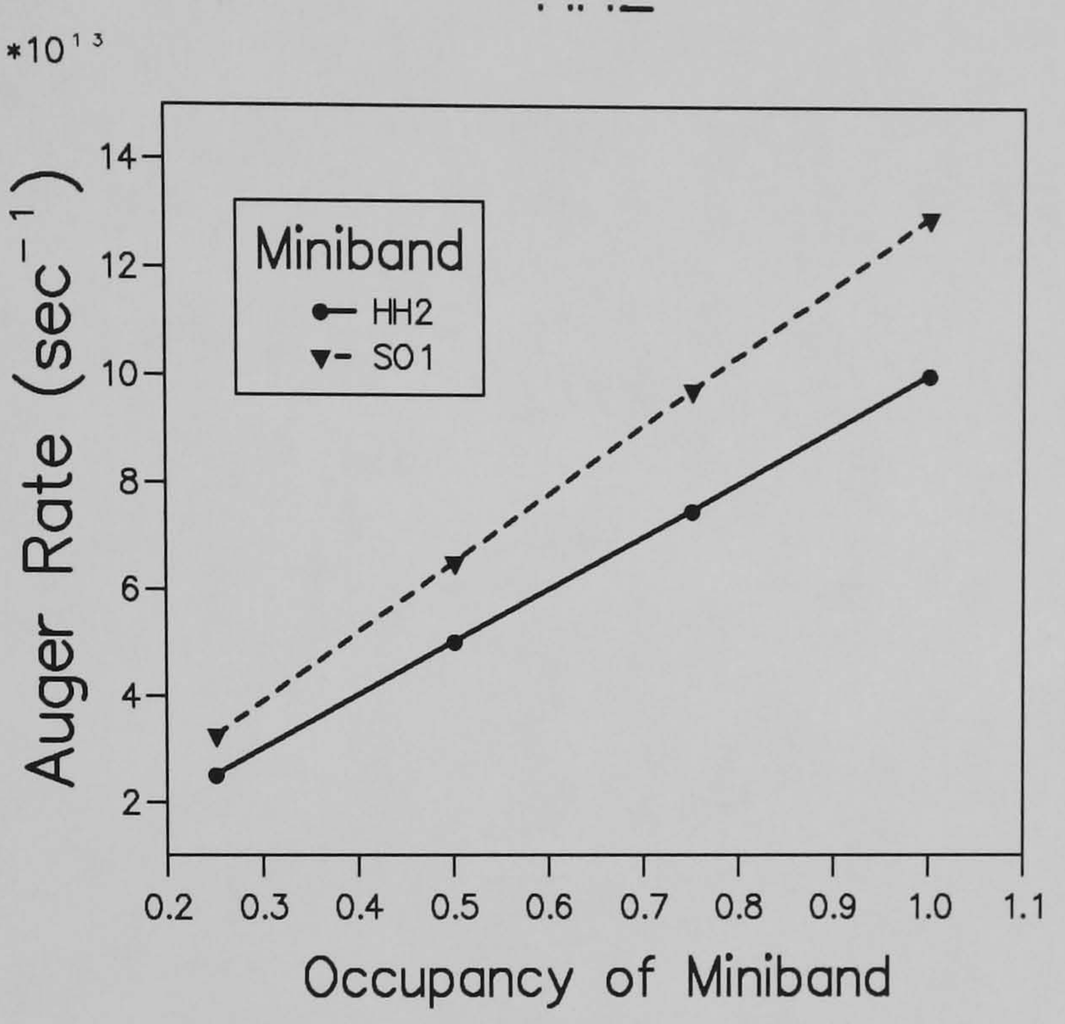


Figure 9.9: *The rate of Auger recombination in a 15%, 20 monolayer structure, with varying values of occupancy for the HH2 and SO1 minibands.*

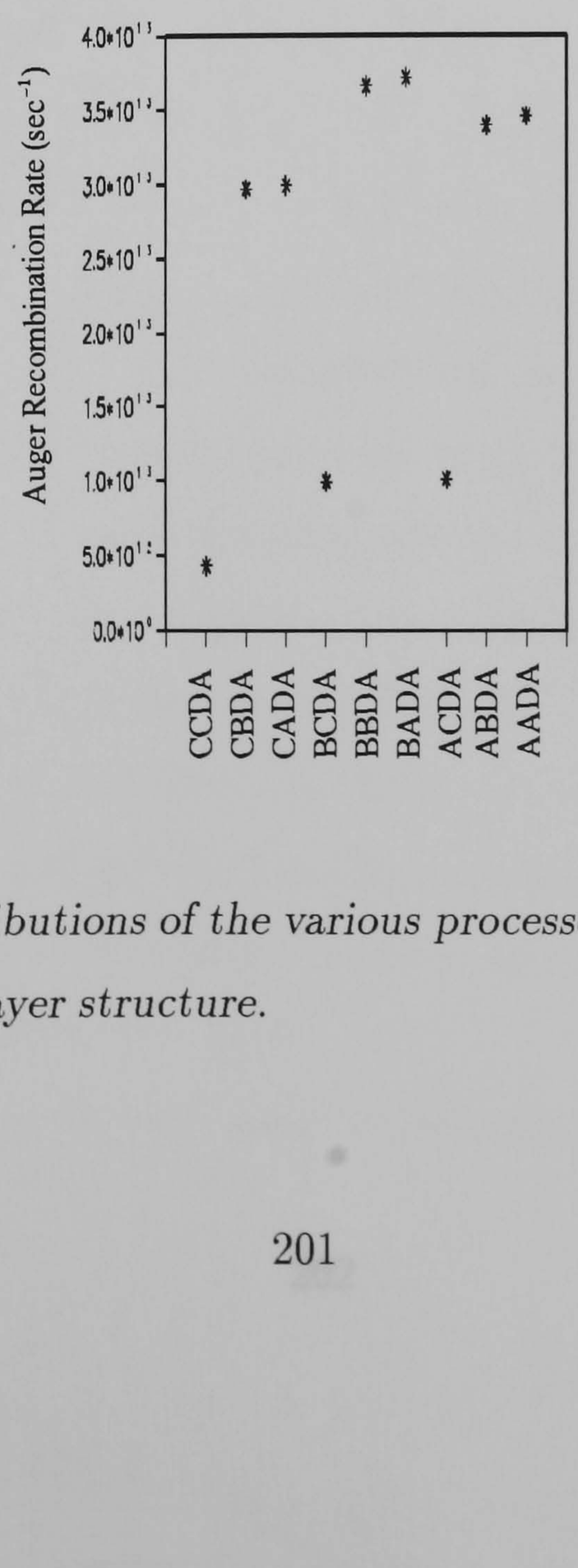


Figure 9.10: *The contributions of the various processes to the total Auger rate for the 15%, 20 monolayer structure.*

contributing components. We can see clearly that no process dominates in this calculation. However, this does not completely invalidate the above discussion. Elimination of just one Auger process must help reduce the overall Auger recombination rate of the system. discussed above for the prevention of the one processes.

Unfortunately, further results for the 50%, 32 monolayer system are not available at this time, thus we are unable to conclude whether the band structure engineering discussed above leads to any significant improvement in the Auger rates. These results will be presented in Williams (1995).

Chapter 10

Conclusion

We have discussed in this thesis the possibility of using intravalence band transitions in p-type SiGe/Si quantum well structures for infra-red detection. We began with a review of the development of IR detectors and the current state of the art of many systems, with particular focus on quantum well structures of various types. The importance of developing a system that is compatible with existing Si technology and the desirability of obtaining normal incidence absorption without the use of gratings and waveguides was discussed.

A detailed description of the calculational methods used to produce the miniband dispersion curves, namely a local empirical pseudopotential method including strain and spin orbit coupling effects, was given and the derivation of the formula used for calculating the first order susceptibility.

We compared our calculated results with the experimental curves of two groups, (Fromherz 1995 and Boucard 1995) and showed good agreement in both cases for a range of germanium concentrations and well widths. This comparison enabled us to determine suitable values of the damping constant to be used for structures of this type.

We further analyzed a wide selection of SiGe/Si systems with a range

of well widths and germanium concentrations for both parallel and normal incidence light and discussed the microscopic origin of the observed peaks. This showed that for parallel incident light the main source of the absorption is for transitions between the HH2 and HH1 minibands. We also noted that the zone centre separation of these levels is not always a good indication of the peak absorption response as many important contributions to the overall absorption come from areas of the Brillouin zone away from the zone centre. This trend is more pronounced in the normal incidence case where the presence of mixing between the minibands relaxes the selection rules and allows normal incident light to be absorbed. This mixing is more pronounced away from the zone centre, thus the peak positions of the normal incidence absorption response cannot be predicted accurately from a calculation based solely on the Γ point. The principle absorption mechanism for normal incidence light in structures with low germanium concentrations is transitions between the HH1 and SO1 minibands. For higher Ge concentrations, the LH2 miniband is usually involved.

We predicted several structures that would appear to be good candidates for infra-red detection. In particular, structures containing 15% or 20% germanium concentrations within the well, with well widths of around 20 monolayers produce parallel incidence absorption response at $\approx 10\mu$m, utilizing bound-continuum transitions, which is ideal for device applications.

For normal incidence absorption the above mentioned structure again appear to be good candidates for detector applications, with the normal incidence response again peaking at $\approx 10\mu$m.

We also proposed a 'two-colour' IR detector comprising a 30%, 44 monolayer well, which produced normal incidence absorption in both the 3-5μm and the 10-15μm range.

The Auger recombination rates for various structures were calculated and the engineering of the miniband structure to produce conditions that would help to eliminate AR were discussed. We proposed that a structure consisting of a 50%, 32 monolayer well would produce absorption at $\approx 10\mu$m with an energy band structure that would minimize the non-radiative recombination.

10.1 Further Study

Our model could be extended to include such things as the effect of electric fields on the optical response (Hagon 1994), interface disorder (Turton 1993) and defects (Jaros 1993).

Second order and higher susceptibility could be considered (Wong 1994). Our susceptibility expression assumed a low intensity limit allowing the relaxation processes to be included in a very simple way. This model of the relaxation processes could be improved (Butcher 1990) and extended to include the high intensity limit.

Many body effects, such as depolarization, the direct interactions and a more subtle interpretation of the exchange shift may help to improve our comparison with experimental data.

In conclusion, we have shown that p-type SiGe/Si based detectors have the potential to be the IR detectors of the future. The figures of merit of these systems will always be inferior to that of III-V's but advantages of the compatibility with the existing technology, the use of mature growth processing techniques and the ability to absorb normal incidence radiation directly. It remains to be seen whether these advantages will outweigh the poorer figures of merit.

Appendix A

The Effect of Symmetry Operators on Momentum Matrix Elements

The momentum matrix element between states a and b at wavevector $\mathbf{k}$ is defined as

$$P^{\mathbf{e}}_{ab}(\mathbf{k}) = \int d^3\mathbf{r}\psi^*_{a\mathbf{k}}(\mathbf{r})\mathbf{e}.(-i\hbar\frac{\partial}{\partial\mathbf{r}})\psi_{b\mathbf{k}}(\mathbf{r}), \tag{A.1}$$

where $\mathbf{e}$ is the polarization direction, $\psi_{n\mathbf{k}}(\mathbf{r})$ represents the Bloch wavefunction of the quantum well miniband of index n at wavevector $\mathbf{k}$. For the point $(\mathcal{S}\mathbf{k})$ we have

$$P^{\mathbf{e}}_{ab}(\mathcal{S}\mathbf{k}) = \int d^3\mathbf{r}\psi^*_{a\mathcal{S}\mathbf{k}}(\mathbf{r})\mathbf{e}.(-i\hbar\frac{\partial}{\partial\mathbf{r}})\psi_{b\mathcal{S}\mathbf{k}}(\mathbf{r}), \tag{A.2}$$

One result of group theory is that for non-degenerate Bloch functions we can write

$$\psi_{n\mathcal{S}\mathbf{k}}(\mathbf{r}) = \psi_{n\mathbf{k}}(\mathcal{S}^{-1}\mathbf{r}). \tag{A.3}$$

Therefore equation $(A.2$ becomes

$$P^{\mathbf{e}}_{ab}(\mathcal{S}\mathbf{k}) = \int d^3\mathbf{r}\psi^*_{a\mathbf{k}}(\mathcal{S}^{-1}\mathbf{r})\mathbf{e}.(-i\hbar\frac{\partial}{\partial\mathbf{r}})\psi_{b\mathbf{k}}(\mathcal{S}^{-1}\mathbf{r}). \tag{A.4}$$

We now introduce a change of notation and let $\mathcal{S}\mathbf{r} \rightarrow r'$ and

$$r'_\alpha = \sum_\beta (\mathcal{S}^{-\infty})_{\beta\alpha} r_\beta \tag{A.5}$$

where $\mathcal{S}^{-\infty}$ are the elements of the matrix representing the transformation $\mathcal{S}^{-\infty}$. Therefore we have

$$\frac{\partial}{\partial r_\alpha} = \sum_\beta \frac{\partial}{\partial r'_\beta} . \frac{\partial r'_\beta}{\partial r_\alpha} . \tag{A.6}$$

and

$$\begin{aligned} \frac{\partial r'_\beta}{\partial r_\alpha} &= \frac{\partial}{\partial r_\alpha} \left[\sum_\gamma (\mathcal{S}^{-1})_{\beta\gamma} r_\gamma \right] \\ &= \sum_\gamma (\mathcal{S}^{-1})_{\beta\alpha} \delta_{\alpha\gamma} \\ &= (\mathcal{S}^{-1})_{\beta\alpha} . \end{aligned} \tag{A.7}$$

We now need to express the operator $\mathbf{e}.\frac{\partial}{\partial \mathbf{r}}$ is the new reference frame;

$$\begin{aligned} \mathbf{e}.\frac{\partial}{\partial \mathbf{r}} &= \sum_\alpha e_\alpha \frac{\partial}{\partial r_\alpha} \\ &= \sum_\alpha e_\alpha \sum_\beta \frac{\partial}{\partial r'_\beta} (\mathcal{S}^{-1})_{\beta\alpha} \\ &= \sum_\alpha \sum_\beta \frac{\partial}{\partial r'_\beta} [(\mathcal{S}^{-1})_{\beta\alpha} e_\alpha] \\ &= \mathcal{S}^{-1} \mathbf{e}.\frac{\partial}{\partial \mathbf{r}'}, \end{aligned} \tag{A.8}$$

The volume element $d^3\mathbf{r}$ is unchanged between reference frames, therefore $d^3\mathbf{r} = d^3\mathbf{r}'$. Our final expression for the matrix element at $\mathcal{S}\mathbf{k}$ is

$$P^{\mathbf{e}}_{ab}(\mathcal{S}\mathbf{k}) = \int d^3\mathbf{r}' \psi^*_{a\mathbf{k}}(\mathbf{r}') \mathcal{S}^{-1} \mathbf{e}.(-i\hbar \frac{\partial}{\partial \mathbf{r}'}) \psi_{b\mathbf{k}}(\mathbf{r}'). \tag{A.9}$$

Comparing this expression with our original definition of the matrix element in equation $(A.1)$ we can recognize the right hand side as $P^{\mathcal{S}^{-1}\mathbf{e}}_{ab}(\mathbf{k})$. We can therefore conclude that under t he symmetry transformation $\mathcal{S}$,

$$P^{\mathbf{e}}_{ab}(\mathcal{S}\mathbf{k}) = P^{\mathcal{S}^{-1}\mathbf{e}}_{ab}(\mathbf{k}). \tag{A.10}$$

Therefore, given knowledge of the matrix element at wavevector $\mathbf{k}$ we can deduce the matrix element at $(\mathcal{S}\mathbf{k})$. Thus we need only find the matrix elements for wavevectors present in the irreducible segment of the first Brillouin zone and from there information about the rest of the zone can be obtained by use of the symmetry operators.

Appendix B

Valence Bond Orbitals

Si and Ge both have a valence of four and form covalent tetrahedral bonds. The one-electron wavefunctions are considered to form valence bond orbitals as shown in figure B.1. These orbitals must have definite transformation prop-

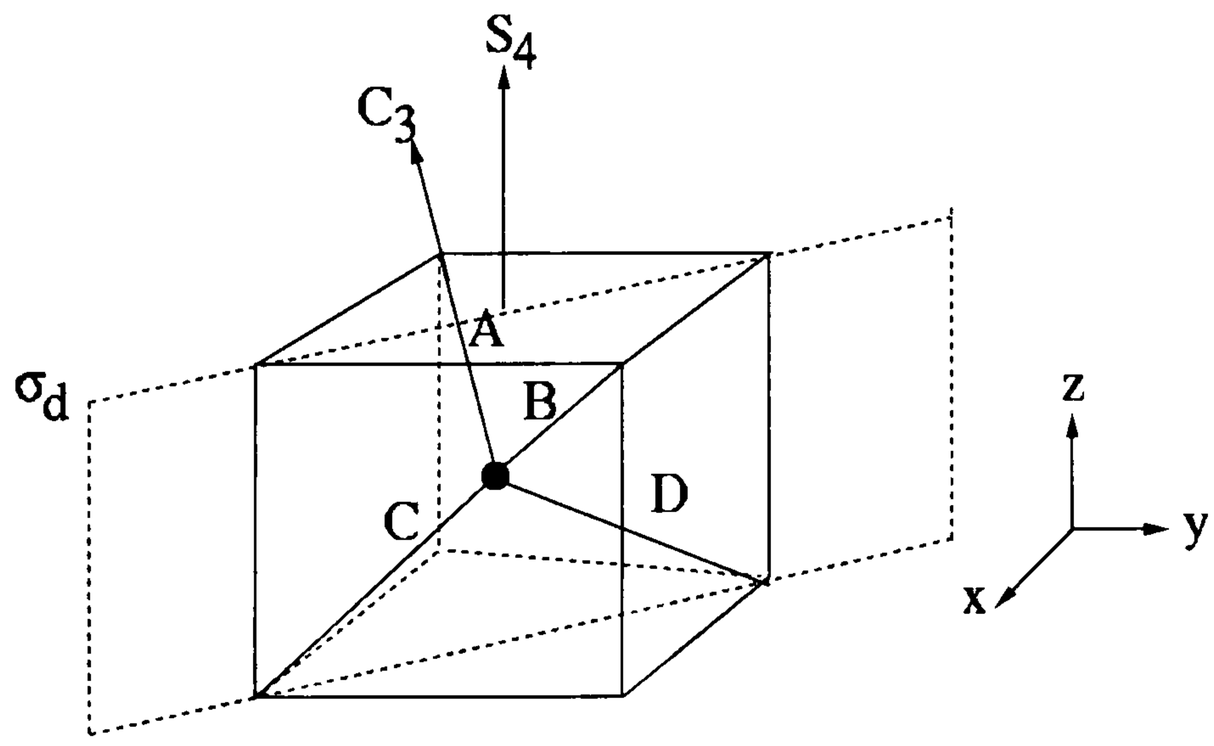


Figure B.1:

erties so that we can use them properties in calculating matrix elements and energy values. We begin by considering the simplest possible case, that of the

hydrogen molecule and apply the results obtained to the more complex case of Si and Ge.

B.1 Hydrogen

The hydrogen molecule consists of two protons considered to be positioned at $\mathbf{R_1} = (0, 0, -\frac{1}{2}a)$ and $\mathbf{R_2} = (0, 0, \frac{1}{2}a)$. The potential energy of one electron in the hydrogen molecule is shown in figure B.2. This potential includes the

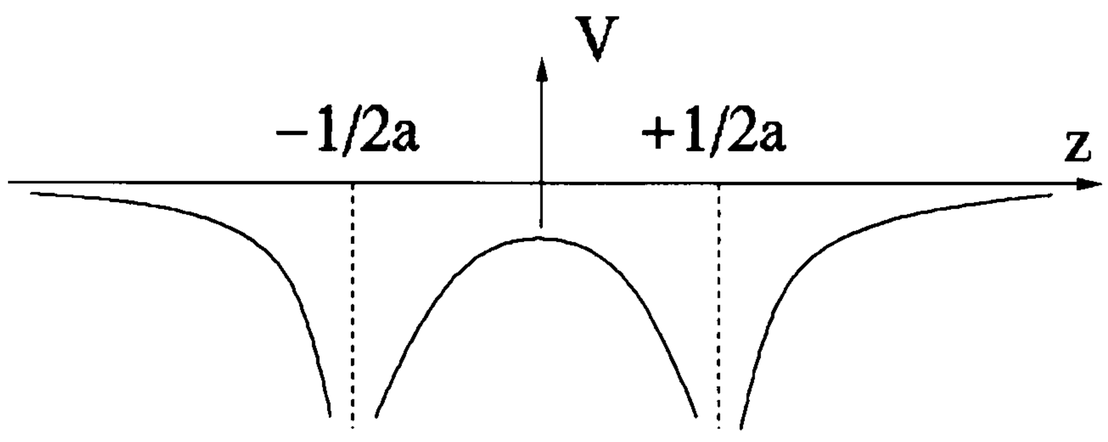


Figure B.2: *Potential energy experienced by one electron in a hydrogen molecule.*

nuclear potential and the self consistent average potential of the second electron. It can be seen that this potential is symmetric about the origin and the electron can easily penetrate from one atom to another. There are two possible orbitals that can from. The lowest orbital is formed from the sum of the atomic orbitals as shown in figure B.3. The low energy of this orbital means

Figure B.3: *Schematic representation of the bonding (top) and anti-bonding (bottom) wavefunctions of the hydrogen molecule.*

it physically represents two atoms bound together in a stable molecule, thus

is termed a bonding orbital. The other possible wavefunction shown in the lower half of figure B.3 represents the next lowest energy state and is termed an anti-bonding orbital. The higher energy detracts from the binding of the molecule and thus this system is less stable. The difference in energy is due to the lower K.E. of the bonding orbital. The K.E. is proportional to $\nabla\psi$, which, as can be seen from figure B.3 is larger for ψ_a than for ψ_b. In the region around each nuclei the attraction from that nuclei will dominate compared with the effect of the neighbouring nuclei. Thus ψ_b and ψ_a will both be approximately atomic $1s$ wavefunctions near the nuclei. Thus we can deduce an appropriate form for the orbitals

$$\begin{aligned} \psi_b &= \psi_{1s}(\mathbf{r} - \mathbf{R_1}) + \psi_{1s}(\mathbf{r} - \mathbf{R_2}) \\ \psi_a &= \psi_{1s}(\mathbf{r} - \mathbf{R_1}) - \psi_{1s}(\mathbf{r} - \mathbf{R_2}) \end{aligned} \tag{B.1}$$

The more overlap there is between functions centred on the two atoms the lower the K.E. and hence stronger bonds are formed by electrons in bonding orbitals. Each type of orbital can be occupied by two electrons, one with spin up and one with spin down.

B.2 Covalent Bonds.

The strength of a covalent bond tends to increase by having lopsided orbitals stretching out towards one another, hence the term directed valence orbitals. The wavefunctions can be expressed in a similar form to equations (B.1) but without the restriction to $1s$ orbitals, i.e.

$$\begin{aligned} \psi_b &= \psi_1(\mathbf{r} - \mathbf{R_1}) + \psi_2(\mathbf{r} - \mathbf{R_2}) \\ \psi_a &= \psi_1(\mathbf{r} - \mathbf{R_1}) - \psi_2(\mathbf{r} - \mathbf{R_2}) \end{aligned} \tag{B.2}$$

Representation	E	$6S_4$	$3C_2$	$8C_3$	$6\sigma_d$	
A_1	1	1	1	1	1	r^2
A_2	1	-1	1	1	1	
E	2	0	2	-1	0	
T_1	3	1	-1	0	-1	
T_2	3	-1	-1	0	1	x,y,z
$D = A_1 + T_2$	4	0	0	1	2	

Table B.1: Character table for the point group T_d.

For Si, near the nucleus the wavefunction must be approximate to a $3s$ or $3p$ orbital, since lower orbitals are already filled. A directed valence orbital can then be formed as a linear combination of a $3s$ and $3p$ atomic orbitals. However, we know silicon forms four valence bands with its neighbors directed towards the alternate corners of a cube as shown in figure B.1. We must therefore show that four appropriately directed, equivalent orbitals, $\psi_A, \psi_B, \psi_C, \psi_D$, can be formed from atomic $3s, 3p_x, 3p_y, 3p_z$ functions. The directed orbitals are represented by the point group T_d. The character table for this group is given in table B.1. From inspection of the tetrahedral structure we can form a representation, D, of the group. There is only a contribution to the character of a representation when an orbital is transformed into itself. Thus $\chi(E) = 4$, $\chi(S_4) = 0$, $\chi(\sigma_d) = 2$, $\chi(C_2) = 0$ and $\chi(C_3) = 1$. As this representation contains a 4, but the character table only contains 3's, this representation must be reducible. From the character table we can see that this representation can be reduced to $A_1 + T_2$. $\psi(3s)$ is a spherically symmetrical orbit and transforms according to A_1. $\psi(3p_x), \psi(3p_y)$ and $\psi(3p_z)$ transform like x, y, z according to representation T_2. Thus $\psi(3s)$, $\psi(3p_x), \psi(3p_y)$ and $\psi(3p_z)$ also transform according to D. Thus linear combinations of these orbitals have required sym-

metry properties of $\psi_A, \psi_B, \psi_C, \psi_D$ The correct linear combinations, which can be found by inspection, or but using projection operators are

$$\begin{aligned}
\psi_A &= \psi(3s) - \psi(3p_x) - \psi(3p_y) + \psi(3p_z) \\
\psi_B &= \psi(3s) + \psi(3p_x) + \psi(3p_y) + \psi(3p_z) \\
\psi_C &= \psi(3s) + \psi(3p_x) - \psi(3p_y) - \psi(3p_z) \\
\psi_D &= \psi(3s) - \psi(3p_x) + \psi(3p_y) - \psi(3p_z)
\end{aligned} \tag{B.3}$$

These wavefunctions are unnormalized. Each of these directed orbitals forms a bonding orbital with the neighbouring atom in the appropriate direction. Combined with the two spin functions they produce eight spin-orbitals. These are all occupied in intrinsic Silicon, four of the electron coming from the central Si atom and one from each of the neighbours. Thus near the Si atom the wavefunction is like a full shell $(3s)^2(3p)^6$.

B.3 Bulk Si

For bulk silicon we can expect to from the electron wave function ψ_k out of directed valence orbitals of the type given in equation (B.3). At the centre of the Brillouin zone, $\mathbf{k} = 0$ the directed valence orbitals are just linear combination of the atomic orbitals. Thus we can form Bloch functions with $\mathbf{k} = 0$

$$\begin{aligned}
\psi_{s,k=0}(\mathbf{r}) &= \sum_m \phi_s(\mathbf{r} - \mathbf{r}_m) \\
\psi_{x,k=0}(\mathbf{r}) &= \sum_m \phi_x(\mathbf{r} - \mathbf{r}_m) \\
\psi_{y,k=0}(\mathbf{r}) &= \sum_m \phi_y(\mathbf{r} - \mathbf{r}_m) \\
\psi_{s,k=0}(\mathbf{r}) &= \sum_m \phi_z(\mathbf{r} - \mathbf{r}_m)
\end{aligned} \tag{B.4}$$

where the summation is over all atoms $\mathbf{r}_m$. At $\mathbf{k} = 0$ symmetry forbids the mixing of ψ_s with ψ_x, ψ_y, ψ_z and the levels approximate to atomic ones, i.e. the

s-level lowest and the triply-degenerate p-level highest with each level having two-fold spin degeneracy.. The inclusion of spin-orbit coupling splits the 6-fold level into a double and a quadruplet. This is exactly analogous to the splitting of an atomic p level into a $j = \frac{1}{2}$ doublet and $j = \frac{3}{2}$ quadruplet.

Away from the zone centre, the symmetry changes and the mixing of the s and p like levels occurs. This leads to changes in the optical properties of the system which are discussed in Chapters 7 and 8.

References

Aakiyama, A. T. Sasaki, T. Seto, A. Mori, R. Ishigaki, S. Itoh, N. Yutani, M. Kimata, N. Tubouchi, Optical Engineering, **33**, 64, (1994)

Akasaka, Y., Horie, K. Nakamure G. pn. J. Appl. Phys, **13** 1846, (1974)

Andersson,J.Y., L. Lundquist and Z.F. Paska, Appl. Phys. Lett, **58**, 2264 (1991)

Ando T., Rev. Mod. Phys., **54**, 437, (1982)

Ashcroft N.W. and N.D. Mermin, "Solid State Physics", Holt-Saunders, Tokyo, (1981)

Bandara K. M. S. V., D.D. Coon., O. Byungsung, Y.F. Lin, and M. Francombe , Appl. Phys. Lett., **53**, 1931 (1988)

Bassini F. and V. Celli, J. Phys. Chem. Solids, **20**, 64 (1961)

Bassani F and G. Pastori Parravicini, Electronic States and Optical Transitions in Solids (Pergamon, New York, (1975)

Bastard G., "Wave Mechanics Applied to Semiconductor Heterostructures", Les Editions de Physique, Les Ulis Cedex (1988).

Bean J.C., L.C. Feldman, A.T. Fiory, S. Nakakara and I.K. Robinson, J. Vac. Sci and Tech. A, **2**, 436 (1984)

Bertho D., J.M. Jancu and C. Jouanin, Phys. Rev. B., **50**, 16957 (1994)

Bloom S. and T.K. Bergstresser, Solid State Commun., **6**, 465, (1968)

Borrello S.R. and H. Levinstein, J Appl. Phys, **33**,, 2947 (1962)

Boucaud P., L. Gao, F. Visocekas, Z- Moussa, J-.M. Lourtioz, F.H. Julien, Appl. Phys. Lett., (to be published) (1995a)

Boucaud P., L. Gao, F. Visocekas, Z- Moussa, J-.M. Lourtioz, F.H. Julien, J. Cryst. Growth, (to be published) (1995b)

Boukerche M, P.S. Wijewarnasuriya, S. Sivanthan, I.K. Sou, Y.J. Yim, K.K. Mahavadi and J.P. Faurie, J. Vac. Sci. Technol, A, **6**, 2830 (1988)

Boukerche M, J. Reno, I.K. Sou, C. Hsu and J.P. Faurie, Apl. Phys. Lett., **48**, 1733 (1986)

Broudy R.M. and V.J. Mazurczyck, Semiconductors and Semimetals, Chapter 5, **18**, 1981

Burstein E,, J.W Davission, E.E. Bell, W.J. Turner, and H.G. Lipson, Phys Rev, **93**, 65, (1954)

Bratt P.R., W. Engeler, H. Levinstein, A. MacRae and J. Pehek, Infrared Physics, **1**,27 (1961)

Bratt P.R., Semiconductors and Semimetals, Vol. 12, 1977

Brown E.R., S.J. Eglash and K.A. McIntosh, Phys. Rev. B, **46**, 7245 (1992)

Butcher P. N., and Cotter D., 'The Elements of Nonlinear Optics', Cambridge University Press, Cambridge (1990).

Campbell I.H., I.Sela, B.K. Laurich and D.L. Smith, Appl. Phys. Lett, **57**, 846 (1991)

Cashman J, Opt. Soc. Am, **36**, 356, (1946)

Cashman Proc. Int. Radio Eng. **47**, 1471, (1959)

Chang Y.C. and J.N. Schulman, Phys. Rev. B., **31**, 2069 (1985)

Chang Y.C. and R.B. James, Phys. Rev. B., **39**, 12672 (1992)

Chern C.H., K.L.. Wang, G. Bai and M.A. Nicolet, Mat. Res. Soc. Symp. Proc, **220**, 175 (1991)

Chin V.W.L. , M.A. Green and J.M.V. Storey, Solid state Electr., **36**, 1107,

(1993)

Chiu L.C. and A. Yariv, I.E.E.E. Journal of Quantum Electronics, **QE-18**, 1406 (1982)

Chow D.H. , R.H. Miles. J.N. Schulman, D.A. Collings and T.C. McGill, Semicond. Sci. Technol., **6**, C47, (1991)

Collins D.A. , M.W. Wang, R.W. Grant and T.C. McGill, J. Vac. Sci. Technol. **B11**, 1418 (1993)

Cornwell J.F. Group Theory and Electronic Energy bands in solids. (1969)

Daunois, A and Aspnes, D.E. Phys. Rev. B, **18** 1824 (1978)

Dodson B.W. and P.A. Taylor, Appl. Phys. Lett, **49**, 642 (1986)

Eisenman W.L., J.D. Merriam and R.F. Potter, Semiconductors and Semimetals, Vol 12 (1977)

Esaki L. and H. Sakaki IBM Tech. Disc. Bull, **20**, 2456 (1977)

Faurie J.P. ,, A. Million, J. Piaquet, Appl. Phys. Lett, **41**, 713, (1982)

Fouks B.I. , V.F. Kocherow,, I..I. Taubkin, Optical Engineering, **33**, No.5, 1485, (1994)

Friedel P., M.S. Hybertsen and M. Schluter, Phys. Rev., B, **39**, 7974 (1988)

Fromherz T., E. Koppensteiner, J.F. Nutzel, M. Helm, G. Bauer and G. Abstreiter, Superlattices and Microstructures, **15**, 229 (1994a)

Fromherz T., E. Koppensteiner, J.F. Nutzel, M. Helm, G. Bauer and G. Abstreiter, Phys. Rev. B., **50**, 15073, (1994b)

Fromherz T., E. Koppensteiner, J.F. Nutzel, M. Helm, G. Bauer and G. Abstreiter, Solid State Electronics, **37**, 941 (1994c)

Fromherz T., J.F. Nutzel, H. Hertle, M. Helm, G. Bauer and G. Abstreiter, Japanese Journal of Appl. Phys. Part 1, **33**, 2361 (1994d)

Fromherz T., J.F. Nutzel, H. Hertle, M. Helm, G. Bauer and G. Abstreiter, Vibrational Spectroscopy, **8**, 109 (1995)

Gell M. A., Ph. D. Thesis, "Electronic and Optical Properties of III-V semiconductor superlattices", University of Newcastle-upon-Tyne (1986a)

Gell M.A., K.B. Wong, D. Ninno and M. Jaros, J. Phys. C, **19**, 3821 (1986)

Golay M.J.E., Rev.Sci. Instrum. **18**, 357 (1947)

Goosen K.W., S.A. Lyon,and K. Alavi, Appl. Phys. Lett, **52**, 1701 (1988)

Goosen K.W., S.A. Lyon and K.Alavi Appl. Phys. Lett, **53**, 1027, (1988)

Gunapala S.D., B.F. Levine, D.Ritter, R. Hamm, and M.B. Panish, Appl. Phys. Lett, **58**, 2024 (1991)

Gunapala S.D. et.al. Appl. Phys. Lett, **64**, 3431 (1994)

Hackwell J.A. and D.W. Warren, Opt. Eng. **32** 3102 (1993)

Haller E.E., M.R. Hueschen and P.L. Richards, Appl. Phys. Lett, **34**, 495, (1979)

Hammersley T., "Monte-Carlo Methods", Methuen (1964)

Hanson C.M. , Proc SPIE **1735**, 17 (1992)

Haoni A., J. Phys, **24**, 294 (1963)

Hagon J.P., R.J. Turton, A.T. Miloszewski, G.S.M. Elfardag and M. Jaros, Superlattices and Microstructures, **16**, 125 (1994)

Hahn T., International Tables for Crystalography, etidth by T. Hahn, Space Group Symmety, Vol A, Reidel, Holland, (1983)

Harman T.C. and I. Melngailis, Appl. Solid State Sci., **4**, (1974)

Harris J.J., J. Material Science, Materiels in Electronics, **4**, 93 (1993)

Hasain G, B.F. Levine, C.G. Bethea, R.A. Logan, J. Walker and R.J. Malik, Appl. Phys. Lett, **54**, 2515, (1989)

Haug A., Appl. Phys. Lett., **42**, 412 (1983a)

Haug A., J. Phys. C. **16**, 4159 (1983b)

Heine V., 'Group Theory in Quantum Mechanics', Pergamon Press, (1960)

Herman F. and S. Skillman, "Atomic Structure Calculations", Prentice Hall,

Engleweood Cliffs, New Jersey, U.S.A., (1963)

Herring C., Phys. Rev., **57** 1163 (1940)

Herschel W., Philos. Trans. R. Soc, **90**, 284, (1800)

Hertie H, G. Schuberth, E. Gornik and G. Abstreiter, Appl. Phys. Lett, **59**, 2977, (1991)

Inkson J. C., "Many Body Theory of Solids", Plenum, London (1983).

Jaros M., A.W. Beavis, E. Corbin, J.P. Hagon, R.J. Turton and K.B. Wong, Journal of Vacumn Sci. and Tech. B., **11**, 1689 (1993)

Jiang Y.,, M.C. Teich and W.I. Wangm J. Appl. Physm **69**, 6869, (1991)

Johnson N.F. , H. Ehrenreich, G.Y. Wul and T.C. McGill, Phys. Rev. B, **41**, 3655 (1990)

Jones C.E. , T.N. Casselman, J.P. Faurie, S. Perkowitz and J.N. Schulman, Appl. Phys. Lett, **47**, 140, (1985)

Kane E.O., Phys. Rev., **159**, 624 (1967)

Kane E.O., J. Phys. Chem. Solids, **1**, 82 (1956)

Karunsari R.P.G., J.S. Park, Y.J. Mii and K.L. Wang, Appl. Phys. Lett, **57**, 2585, (1990)

Kasper G.and H.J. Herzog, Appl. Phys. Lett, **8**, 199, (1975)

Kasper G. and H.J. Herzog, Thin Solid Films, **44**, 357, (1977)

Kazanski A.G. and P.L. Richards, Appl. Phys. Lett, **31**, 496, (1977)

Kimata M., N. Yutani, H. Yagi, J¿ Nakanishi, N. Tsubouchi and T. Seto, Proc. SPIE **1540**, 238, (1991)

Kozlowski L.J. et al. IEEE Trans. Electron. Devices, **ED-38**, 1124, (1991)

Kozlowske L.J., R.B. Bailey, S.A. Cabelli, D.E. Cooper, I.S. Gergis, Optical Engineering **33**, 54, (1994)

Kozlowski L.J. , NATO A.S.I Series Vo. E 270, "Quantum Well Intersubband Transitions, Physics and Devices", H.C. Liu, B.F. Levine and J.Y.

Andersson Eds. Kulwer Academic Publishers, Dordrecht, p. 43 (1994)

Kurtz S.R., L.R. Dawson, R.M. Biefield, I.J. Fritz and T.E. Zipperian, IEEE Electron. Device Lett. **10**, 150, (1989)

Kruse P.W., Semiconductors and Semimetals, Vol 18 (1981)

Lansari Y., J.W. Cook and J.F. Schetzina, J. Electronic Materials, **22**, 809 (1993)

Langley L.P., Nature, **25**,15, (1881)

Landsberg P.T., 'Recombination in Semiconductors', Cambridge University Press, (1991)

Lawson W.D., S. Nielson, E.H. Putley and A.S. Young, J. Phys. Chem. Solids, **9**, 325,(1959)

Long D. and J.L. Schmit, Semiconductors and Semimetals, Vol 5 (1977)

Levine B.F. C.G. Bethea, K.K. Choi, J. Walker and R.J. Mailk,., Appl. Phys. Lett, **50**, 1092, (1987)

Levine B.F. C.G. Bethea, K.K. Choi, J. Walker and R.J. Mailk, J. Appl. Phus., **64** 1591 (1988)

Levine B.F., J Appl. Phys, **74**, R1, (1993)

Lin T.L. and J. Maserjian, Appl. Phys. Lett, **57**, 1425 (1990)

Lin T.L., A. Ksendzoq, S.M. Dejewski and E.W. Jones, IEEE Trans. Electron. Devices, **38**, 1141, (1991)

Lin T.L., T. George, E.W. Jones, A¿ Ksendzov and M.L. Huberman, Appl. Phys. Lett, **60**, 380 (1992)

Lin T.L., J.S. Park, S. Gunapara, E.W. Jones and H.M. Del Castillo, Opt. Eng, **33**, 716, (1994)

Mailhiot C and D.L. Smith, Phys, Rev, **35**, 1242, (1987a)

Mailhiot C and D.L. Smith, J. Vac. Sci and Tech A., 1268, (1987b)

Man P. and D.S. Pan, Appl. Phys. Lett, **64**, 321 (1994)

Man P. and D.S. Pan, Appl. Phys. Lett, **61**, 2799 (1992)

Man P. and D.S. Pan, Phys. Rev. B., **44**, 8745 (1991)

Ma Q.M., K.L. Wang and J.N. Schulman, Phys. Rev. B., **47**, 1937 (1992)

Miles R.H., D.H. Chow, J.N. Schulman and T.C. McGill Appl. Phys. Lett, **57**, 801 (1990)

Morton G.A., M.L. Schultz and W.E. Harty, RCA Rev, **20**, 599, 1959

Niles D.W. , D.Rioux and H. Höchst, Phys., Rev., B, **46**, 12547, (1992)

Park J.S., R.P.G. Karunasiri, Y.J. Mii and K.L Wang, Appl. Phys. Lett, **58**, 1083, (1991)

Park J.S., R.P.G. Karunasiri, and K.L Wang, Appl. Phys. Lett, **60**, 103, (1992)

Park J.S., R.P.G. Karunasiri, and K.L Wang, Appl. Phys. Lett, **61**, 681, (1992)

People R., J.C. Bean, C.G. Bethea, S.K. Sputz and L.J. Peticolas, Thin Solid Films, **222**, 121 (1992)

Piotrowski J., Acta Physica Polonica A, **87**, 303, (1995)

Phillips J.C. and L. Kleinman, Phys. Rev., **116** 287 (1959)

Prinzivalli R.A., African Journal of Ecology, **30**, 349, (1992)

Rao M.A., E.J. Caine, H. Kroemer, S.I. Long and D.I Bab, J. Appl. Phys., **61**, 643 (1987)

Reine M.B., and R.M. Broudy, Proc. SPIE Tech. Symp., 21st, San Diego, California (1977)

Rieke F.F., L.H. DeVaus, and A.A. Tuzzolion, J. Proc Int. Radio Eng. **47**, 1475 (1959)

Richards P.L. , J. Appl. Phys., **76**,1 (1994)

Rogalski A., Prog. Quantum Electronics, **13**, 191 (1989a)

Rogalski A., Prog. Quantum Electronics, **13**, 299 (1989b)

Rogalski A.,, Infrared Phys., **31**, 117 (1991)

Rosencher E., P. Bois, B. Vinter, J. Nagle and D. Kaplan, Appl. Phys. Lett, **56**, 1822 (1990)

Samoska L.A., B. Brar, H. Kroemer, Appl. Phys. Lett, **62**, 2539, (1993)

Sanders G.D. and Y.C. Chang, Phys. Rev. B., **31**, 6892 (1985)

Scolfaro L.M.R., D. Beliaev, R. Enderlein and J.R. Leite, Solid State Coms., **93** 469, (1995)

Schulman J.N. and T.C. McGill Appl. Phys. Lett, **34**, 663 (1979)

Schulman J.N. and Y.C. Chang, Phys. Rev. B., **31**, 2056 (1985)

Scheider H., Ehret S., Larkins E.C., Raiston J.D. and Koidl P., 'Long wavelength IR detectors and Arrays', **2**, 69 (1995)

Scott M.W. Appl. Phys, **40**,, 4077 (1969)

Seitz F., "The Modern Theory of Solids", McGraw-Hill, New York (1940).

Shepherd F.D., IR Detectors and Arrays, SPIE **930**, 2, (1988)

Smith D.L. and C. Mailhiot, J. Vac. Sci and Tech A. **5**, 2060 (1987)

Smith D.L. , T.C. McGill amd J.N. Schulman, Appl. Phys. Lett, **43**, 180 (1983)

Shaw M.J. and M. Jaros, Phys. Rev. B., **50**, 7768 (1994)

Shaw M.J., PhD Thesis, 'Microscopic Theory of Optical Nonlinearity in Semiconductor Superlattices', University of Newcastle Upon Tyne, (1994)

Shaw M.J., K.B. Wong and M. Jaros, Phys. Rev. B., **48**, 2001 (1993)

Takeshima M. J. Appl. Phys., **49**, 6118, (1978)

Tempel G., N. Schwarz, F. Muller, F. Kock, H.P. Zeindl and I. Eisele, Thin Solid Films, **184**, 171 (1990)

Turton R.J. and M. Jaros, Semiconductor Science and Technology **8**, 2003 (1993)

Tsaur B.Y., M.M. Weeks, R. Trubiano, IEEE Electron. Dev. Lett., **9**, 650.

(1988)

Tsaur B.Y., I.E.E.E. Electron. Dev. Lett., **10**, 361 (1989)

Van de Walle C.G. and R.M.Martin, Phys. Rev., B, **34**, 5621 (1986)

Wang K.L. and R.P.G. Karunasiri, Chap.5, Semiconductor Quantum Wells and Superlattices for Long-Wavelength IR Detectors, ed. M.O Manasreh, pub. Artech House, Boston.

Wang Y.H., S.S. Li, P. Ho and M.O. Manasreh, J Appl. Phys, **74**,, 1383 (1993)

Watanabe M.O. and Y. Ohba, Appl. Phys. Lett., **50**, 906 (1987)

Weisz G. , Phys, Rev, **149**,504 (1966)

West L. and S. Eglash, Appl. Phys. Lett, **46**, 1156, (1985)

Williams, C., Ph.D. Thesis, "Impact Ionization and Auger Recombination Rates in SiGe/Si Superlattices", University of Newcastle Upon Tyne, (1995)

Wolf J.,, Opt. Eng., **33**, 1492, (1994)

Wong K.B. and M. Jaros, Phys. Rev. B., **50**, 17328 (1994)

Wood R.A. and N.A. Foss, Laser Focus World, p.101, (1993)

Zenner A.,, F. Koch and K. Ploog, Surf. Sci., **196**, 671 (1988)

Zhang Y., N. Baruch and W.I. Wang, J Appl. Phys, **75**,, 3690 (1994)

Zanier S., J.M. Berrior, Y. Guldner and J.P. Vieren, I. Sagnes, F. Glowacki, Y. Campidelli and P.A. Badoz, Phys. Rev. B, **51** 14311 (1995)