Electronic and Optical Properties of Semiconductor Quantum Wells and Dots

by Mark Alexander Cusack

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

August, 1996

Declaration

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

M. A. Cusack

M. A. Cusack

August, 1996

Abstract

The influence of structural size and geometry on the electronic states, optical transitions and nonlinear optical response in semiconductor quantum wells and dots has been investigated. A quantum mechanical theory of optical nonlinearities in semiconductors has been used in conjunction with the empirical pseudopotential band structure method to determine the structural parameters leading to optimum second harmonic generation in *p*-type asymmetric GaAs/$Al_xGa_{1-x}As$ quantum well structures. Maximum confinement of all participating states is of paramount importance for optimising such responses. The multi-band effective mass approximation was adapted for the calculation of the electronic structure and optical transitions in cubic GaAs/AlAs quantum dots and in the recently discovered InAs/GaAs self-assembled structures. The calculation revealed the importance of factors such as the quantum mechanical mixing between bulk basis states, inhomogeneous strain and symmetry in the construction of the optical signatures of these dots.

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

1. "The role of localisation in determining second-order optical nonlinearities in *p*-type GaAs-AlAs asymmetric quantum wells", M. A. Cusack, M. J. Shaw, and M. Jaros, Physical Review B 49, 16 575 (1994).
2. "Novel GaAs-AlAs and Si-SiGe quantum well structures with large optical nonlinearities", M. Jaros, M. A. Cusack, M. J. Shaw, K. B. Wong, and B. M. Adderley, SPIE Vol. 2212, Linear and Nonlinear Integrated Optics, 184 (1994).
3. "Determining the nonlinear optical properties of semiconductor superlattices using highly parallel algorithms in a distributed computing environment", M. A. Cusack, J. P. Hagon, and M. J. Shaw, In the proceedings of the 1994 Parallel Virtual Machine user's group meeting, Oak Ridge National Laboratory, Tennessee.
4. "Absorption and recombination in *p*-type SiGe quantum-well structures", E. Corbin, M. A. Cusack, K. B. Wong, and M. Jaros, Superlattices and Microstructures **16**, 349 (1994).
5. "The electronic structure of InAs/GaAs self-assembled quantum dots", M. A. Cusack, P. R. Briddon, and M. Jaros, Physical Review B 54, 2300 (1996).

Acknowledgments

I should like to thank Professor Milan Jaros for providing direction and assistance through the course of my studies. To the other members of the Theory of Condensed Matter Group, many thanks for all the support and encouragement. I should also like to thank Professor P. M. Petroff at the University of California, Santa Barbara for providing recent experimental results prior to publication.

This study was made possible by the financial assistance of the Engineering and Physical Sciences Research Council.

Contents

Chapter 1

Introduction

Progress in modern crystal growth methods such as molecular beam epitaxy (MBE) and metal-organic chemical vapour deposition (MOCVD) has enabled the growth of one semiconductor on top of another of a different atomic composition with monolayer precision. Creative uses of these techniques have resulted in so-called semiconductor heterostructures with dimensions in the nanometre range that aid in the study of fundamental physics and can be incorporated into novel electronic or optical devices. This chapter provides an introduction to the physics of semiconductor heterostructures and an overview of the detailed study of their electronic and optical properties presented in the chapters that follow.

1.1 Semiconductor Heterostructures

The conduction electrons in a bulk semiconductor such as GaAs are essentially free to propagate throughout the crystal, virtually unaffected by the presence of the periodic potential due to the regular lattice of atoms. The electrons are distributed over a continuous energy band with the quantum aspect of their character hidden away. For conduction electrons in a manmade semiconductor heterostructure the situation is quite different. During the design phase, an effort is made to limit an electron's degrees of freedom, confining the motion of charge essentially to planes, lines or points, depending on the type of structure.

The most studied semiconductor heterostructure is the GaAs/$Al_xGa_{1-x}As$ quantum well structure, first proposed by Esaki and Tsu in 1970 and constructed by Chang *et al.* in 1973. The fabrication of these structures proceeds with the growth of a layer of $Al_xGa_{1-x}As$ (where the aluminium mole fraction can be between 0 and 1) on a GaAs substrate followed by a thin ($\sim 10^{-9}$ m) layer of GaAs and then capped with another layer of $Al_xGa_{1-x}As$. Quantum well structures possess tailor-made properties conferred on them at the time of fabrication by a judicious choice of material composition and the thickness of the epitaxial layer.

The ability to select the electronic and optical characteristics of quantum well structures is a result of the reduced freedom of motion of the charge carriers within them. The lowering of the degrees of freedom of electrons and holes in quantum wells compared to that afforded to electrons in bulk semiconductors arises from the difference in the band gaps of the two materials that constitute the quantum well. The band gap is an intrinsic property of each semiconductor and is not changed by the interface, but at the interface, there will be a discontinuity between the two conduction bands and a discontinuity between the valence bands. The partition between the conduction and valence band discontinuities is determined by the local chemistry on an atomic scale and generally has to be determined experimentally or by *ab initio* calculations. For an electron in the conduction band or a hole in the valence band of such a heterostructure, the discontinuity results in a step in potential energy. The alternate layering of materials with different band gaps forms a

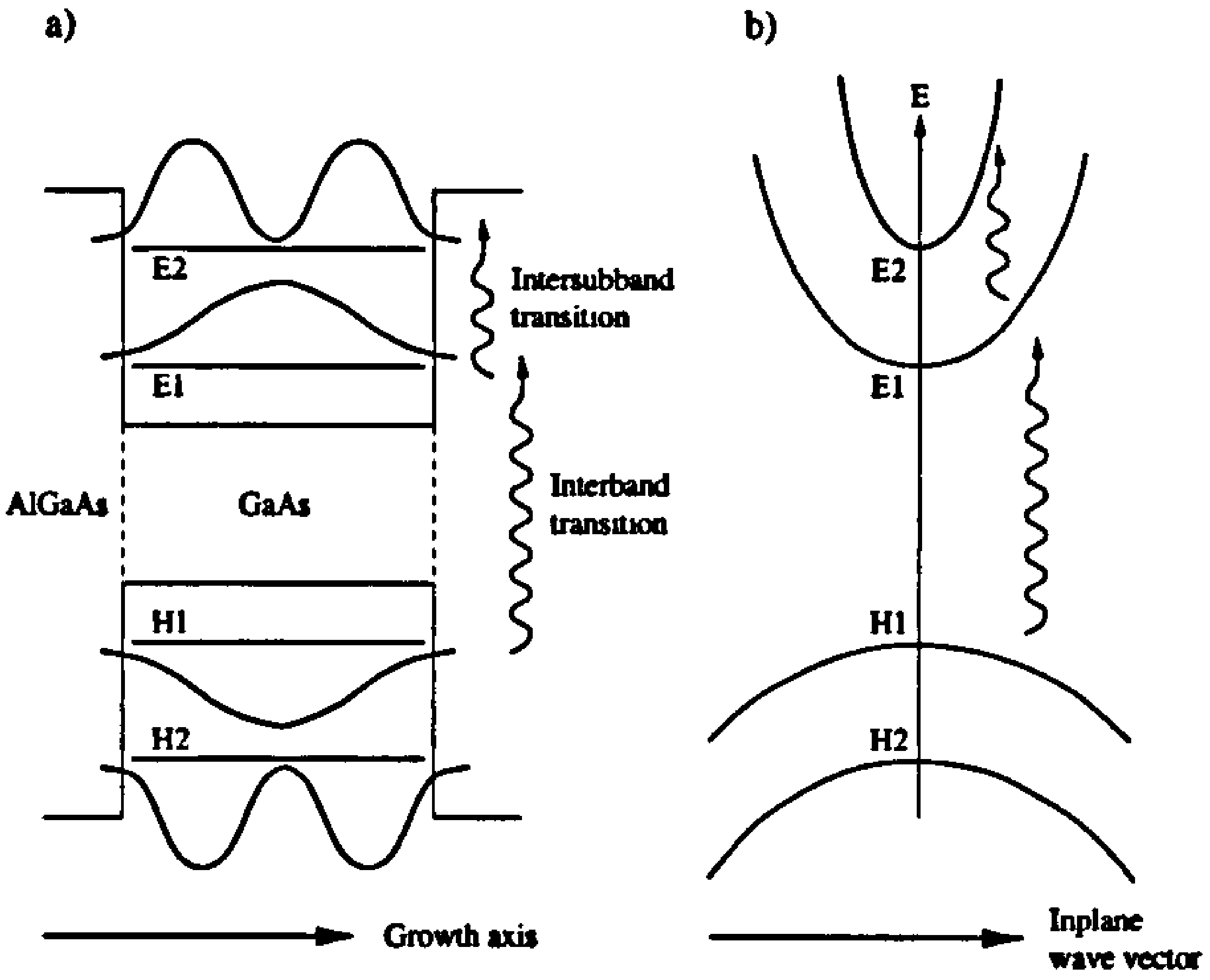

Figure 1.1: (a) A quantum well formed from the differences in band gaps of the two semiconductor constituents. The charge densities for the confined states and allowed transitions are sketched. (b) Subband energy versus inplane wave vector.

potential well and barrier to the electrons and holes [Figure 1.1(a)]. If the width of the potential well formed by the material with the smaller band gap is less than the de Broglie wavelength of the electrons or holes in the material (for example, less than $\approx$ 15 nm for electrons in GaAs), the motion of the carriers can be considered as quantised in the direction parallel to the growth axis, whereas in the plane of the quantum well, their motion is still free. The electrons and holes are distributed over discrete energy subbands rather than over an energy continuum as in bulk [Figure 1.1(b)]. The relative spacing of these subbands depends on the depth and width of the potential well which are controlled by the choice of aluminium mole fraction and epitaxial layer thickness of GaAs respectively. By "engineering" the subband gaps, quantum wells can be fabricated that absorb and emit light at specific frequencies.

Further control over the band structure can be achieved by growing a large number of alternating layers of GaAs and $Al_xGa_{1-x}As$. When the alloy layers are sufficiently thick (> 100 Å) the wave functions associated with the states confined in the individual GaAs wells do not penetrate far enough into the barrier for them to overlap. The energy band structure of these so-called multi-quantum well (MQW) structures is the same as that of a single quantum well. However, if the alloy layers are made more narrow, the wave functions in neighbouring wells interact and the discrete quantum well energy states become broadened into bands of allowed energies. Such a periodic structure of quantum wells is called a superlattice. In comparison with a MQW structure, the superlattice offers an additional degree of flexibility in that it is now possible to choose the superlattice parameters so as to alter the curvature of the energy bands.

Doping quantum well structures with donor or acceptor impurities provides further scope for tailoring their electronic and optical properties. Doping supplies mobile charge carriers to the lowest conduction and valence subbands and thus allows absorption and emission to take place as a result of intersubband transitions (i.e., transitions between conduction or valence subbands) in addition to the interband transitions (i.e., valence subband to conduction subband) that are generally responsible for the generation of optical spectra in these materials. Doped GaAs/AlGaAs quantum well systems are efficient absorbers of infra-red radiation in the 8-12 μm wavelength range because of the narrow energy separation of quantum well states and the large electron transition probabilities for intersubband transitions. The Earth's atmosphere is transparent to electromagnetic radiation with wavelengths in this range. This makes doped GaAs/AlGaAs quantum well systems suitable as infra-red detectors in remote sensing applications. The preferred geometry of devices based on intersubband transitions is that which allows normal incidence absorption to occur. Normal incidence absorption is the result of electron transitions induced by incident light that has a velocity vector parallel to the quantum well growth direction.

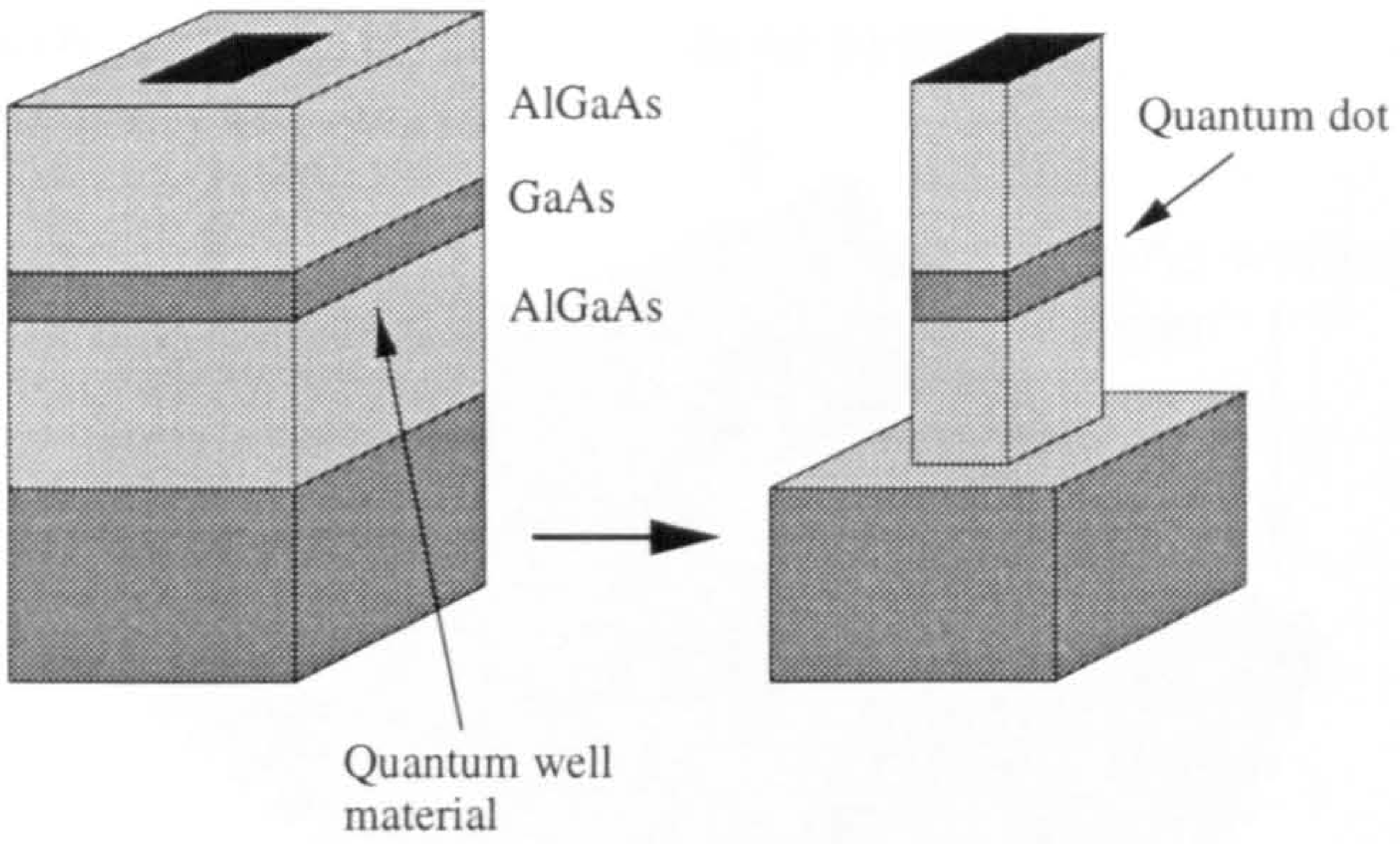

Figure 1.2: The favoured method for fabricating quantum dots employs lithography and etching of quantum well material.

In the case of donor-doped quantum wells, in which transitions occur between conduction subbands, selection rules forbid normal incidence absorption (Coon and Karunasiri, 1984; West and Eglash, 1985). However, in acceptor-doped quantum wells, strongly allowed intravalence band transitions dominate the optical spectrum and normal incidence absorption can take place. As a result, infra-red detectors based on *p*-type GaAs/AlGaAs quantum well structures have become the focus of a great deal of attention recently (e.g. Chang and James, 1989; Szmulowicz and Brown, 1995).

By taking the concept of quantum confinement to its ultimate conclusion one arrives at the quantum dot. In a quantum dot the charge carriers are confined essentially to a point, denied the freedom of motion conceded to electrons in a bulk semiconductor. Because of the complete quantisation of the motion of the charge confined within them, quantum dots exhibit an atomic-like spectrum of energies and a density of states that resembles a series of delta functions. The most popular method of fabricating quantum dots involves the patterning of quantum wells using advanced lithography and etching techniques similar to those used in the fabrication of state-of-the art

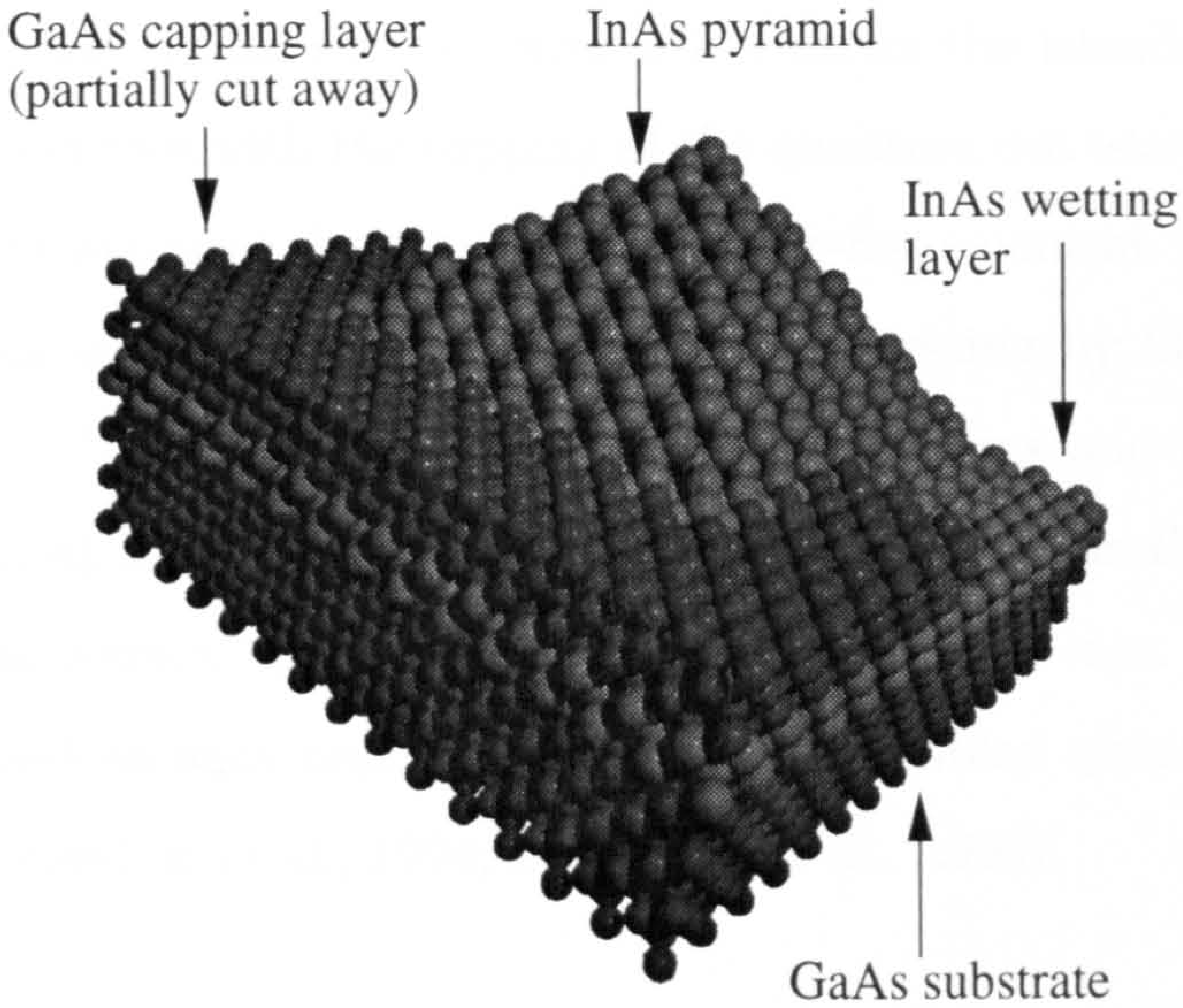

Figure 1.3: InAs/GaAs self-assembled quantum dot.

integrated circuits (Reed, 1993). An electron beam scans the semiconductor surface, which has been coated with a thin polymer layer called a resist. A series of process steps replaces the resist with a thin layer of metal in areas where the beam was scanned at high intensity. A shower of reactive gas then etches away the unprotected quantum well material, leaving pillars as small as 1,000 angstroms across (Figure 1.2). Surface effects repel charge from the outside of the pillar, confining it to a 100 angstrom region of the GaAs well layer in the process.

Recent studies (Moison *et al.*, 1994; Marzin *et al.*, 1994a, Medeiros-Ribeiro *et al.*, 1995; Grundmann *et al.*, 1995a) have shown that it is also possible to attain three-dimensional confinement of charge within the strained islands of InAs that form on the surface of GaAs during the Stranski-Krastanow (Stranski and Krastanow, 1939) growth process. This growth mode begins with an initial MBE layer deposition of InAs on a GaAs substrate. After a certain critical thickness is reached islands of InAs with a pyramidal geometry form spontaneously and a thin wetting layer is left under the islands (Figure 1.3). Fabrication concludes with the capping of the quantum dot island with a layer of the substrate material. By this method defect-free quantum dots with sizes $\approx$120 Å can be constructed with no need for processing by lithography and etching. It has been demonstrated that the island sizes and areal densities can be controlled by varying growth parameters such as the thickness of the initial two-dimensional layer deposition, and the growth rate. Recently, the first lasers based on such nanometre-scale, self-assembled quantum dots were reported (Kirstaedter *et al.*, 1994; Ledentsov *et al.*, 1996).

1.2 Optical Nonlinearities in Heterostructures

The propagation of a harmonic light wave through a medium such as a bulk semiconductor creates what can be thought of as a polarisation wave, that is, an undulating redistribution of charge in response to the field. If we assume that the amplitude of the optical field is negligible in comparison with the cohesive field between electrons and ions ($\sim 10^9$ V/cm) then we can predict a linear relationship between the electric field and the resulting polarisation. If the optical field is comparable to the internal field binding electrons and ions, we might anticipate a significant departure from this linear relationship, causing the medium to exhibit a significant nonlinear response in the form of polarisations involving second-, third-, and higher orders of the field. The polarisation in the low intensity limit is related to the applied field by the linear susceptibility tensor. In the high intensity regime where nonlinear effects are significant, the relationship between the nonlinear polarisations and the applied fields are governed by the nonlinear susceptibilities. The susceptibility tensors depend on the directional properties and the frequency of the applied field(s) in a way that is strongly tied to the microscopic (band structure) properties of the semiconductor.

Because of the high field intensities required, experimental observation of nonlinear optical effects had to await the development of the laser. Soon after, Franken *et al.* (1961) observed second harmonic generation by focusing a 3 kW/cm^2 pulse of red (694.3 nm) ruby laser light onto a quartz crystal. Only one part in 10^8 of this incident wave was converted to the 347.15 nm ultraviolet second harmonic. The observation of many other nonlinear optical effects soon followed, including difference frequency generation and optical rectification by Bass *et al.* (1962a, 1962b) and third harmonic generation by Terhune *et al.* (1962). Second harmonic generation in the compound III-V semiconductors such as GaAs and InAs was first observed by Patel (1965, 1966) using a CO_2 laser beam. The early measurements were soon followed by the publication of what has become the dominant theory relating nonlinear optical phenomena to the microscopic properties of the medium (Butcher and McLean, 1963, 1964). There, contact between the macroscopic nonlinear optical response and the electronic band structure was made via a quantum mechanical derivation of expressions for the susceptibilities.

Apart from the academic interest, nonlinear optical phenomena have a range of practical applications, principally in signal processing, frequency conversion and optical switching. For example, second-order nonlinearities, such as second harmonic generation and difference frequency mixing, are of paramount importance for optical wavelength conversion techniques, which are the basis of new high-resolution spectroscopic tools (Rosencher *et al.*, 1996). Of the third-order nonlinearities perhaps the most important is the intensity-dependent refractive index. This is the process by which the refractive index of the medium is changed by an amount proportional to the optical intensity.

This phenomenon has been exploited in optical switches.

The first experimental evidence of the large optical nonlinearities attainable in semiconductor heterostructures was provided by Fejer *et al.* (1989). They studied the second harmonic generation due to electron transitions between conduction subbands of GaAs/$Al_xGa_{1-x}As$ quantum wells; they measured second harmonic amplitudes 73 times larger than those of the bulk GaAs host material. Much larger enhancements were demonstrated by Rosencher *et al.* (1989) and Boucaud *et al.* (1990) who employed doubly resonant asymmetric step quantum wells to achieve second harmonic signals 1,800 times larger than in GaAs. Second harmonic generation arising from interband (across gap) transitions has also been observed, albeit with some difficulty, by Janz *et al.* (1994) and Qu *et al.* (1994). The measurements proved troublesome as the second harmonic signal strength due to these transitions is comparable to the response of the host material.

Much interest has focused on the large mid-to-far infrared second harmonics associated with transitions between valence subbands in quantum wells, where more favourable optical selection rules exist (Tsang and Chuang, 1992; Qu and Ruda, 1993; Shaw *et al.*, 1993a). Transitions between valence subbands readily find application in practical devices since they tend to be directionally dependent on light polarised in the plane of the interfaces, whereas similarly induced transitions between conduction subbands are usually considered forbidden. This phenomenon is studied in some detail in the chapters that follow.

1.3 Overview of Thesis

This thesis investigates the influence of structural size and geometry on the confined electron states, optical transitions and nonlinear optical processes in semiconductor quantum wells and dots.

In Chapter 2, an outline is given of the empirical pseudopotential method and the effective mass approximation, both used extensively to construct energy bands, wave functions and optical matrix elements in semiconductor heterostructures. The effects of strain on the band structure of bulk semiconductors is reviewed and the necessary correction to the effective Hamiltonian presented. By way of a comparison, a calculation of the band structure and wave functions of a GaAs/AlAs quantum well structure is detailed, utilising both the empirical pseudopotential method and the effective mass approximation.

The theory of optical nonlinearities in semiconductors is outlined in Chapter 3. The relationship between the (non)linear polarisation field and the applied optical field(s) is described, introducing the susceptibility tensors during the process. Using the quantum mechanical density matrix formalism, an expression for the second-order susceptibility governing the extent to which a semiconductor radiates second harmonics is derived, accounting for the microscopic details of the medium.

In Chapter 4, the expression for the second harmonic susceptibility derived in Chapter 3, together with the empirical pseudopotential band structure method given in Chapter 2 are applied to acceptor-doped asymmetric GaAs/$Al_xGa_{1-x}As$ quantum well structures. The role of material parameters such as quantum well depth, quantum well width and degree of quantum well asymmetry in optimising the second harmonic response is investigated.

In Chapter 5, the effective mass approximation outlined in Chapter 2 is used to determine the electronic structure and optical matrix elements in cubic GaAs/AlAs quantum dots. The importance of mixing between bulk valence bands during the construction of the heterostructure states is considered. The effective mass approximation is also applied to InAs/GaAs self-assembled quantum dots to generate electron and hole energy levels, wave functions and optical matrix elements. The effect of dot size and of the inhomogeneous strain present in this structure on the confined electron and hole states, and on transition matrix elements connecting them, is determined. A comparison between the theoretical predictions and recent photoluminescence data is made.

Finally, in Chapter 6, the principal conclusions which may be drawn from the results of these calculations are summarised and discussed.

Chapter 2

Theory of Electronic Band Structures in Semiconductors

To make quantitative statements concerning the optical and electronic properties of semiconductor heterostructures we have to know the electronic band structure, including the energy band and the corresponding wave function. In this chapter, we detail the calculation of the band structures. We focus on two widely employed methods: the empirical pseudopotential method which provides a description of the band structure throughout the Brillouin zone and the $\mathbf{k} \cdot \mathbf{p}$ method, useful for the study of conduction and valence band features near the band edges.

2.1 The Empirical Pseudopotential Method

The empirical pseudopotential (EP) formalism has been applied successfully to bulk materials (Cohen and Bergstresser, 1966) and more recently, to a variety of heterostructures (Jaros and Wong, 1984; Jaros *et al.*, 1985; Mäder and Zunger, 1995). Here, we present an overview of the empirical pseudopotential method for calculating the band structure of III-V compounds and describe how this scheme can be extended to predict the electronic structure of GaAs/AlAs quantum wells.

2.1.1 Representing Bands of Bulk Materials

In the empirical pseudopotential approach, bulk single particle energies and wave functions are solutions to

$$(H_0 + V^{so})\phi_{n,\mathbf{k},s}(\mathbf{r}) = E_{n,\mathbf{k},s}\phi_{n,\mathbf{k},s}(\mathbf{r}) \tag{2.1}$$

$$H_0 = -\frac{\hbar^2}{2m_0}\nabla^2 + V(\mathbf{r}) \tag{2.2}$$

$$V^{so} = \frac{\hbar}{4m_0^2c^2}\sigma \cdot \nabla V(\mathbf{r}) \wedge \mathbf{p} \tag{2.3}$$

in which n is an energy band index, $\mathbf{k}$ is the crystal momentum restricted to the first Brillouin zone and s is a spin index. V^{so} is the spin-orbit potential arising from the interaction of the electron spin magnetic moment with the magnetic field experienced by the electron. $V(\mathbf{r})$ is the total three dimensional *atomistic* potential which, in the energy independent, local approximation, is a superposition of screened atomic potentials, i.e.,

$$V(\mathbf{r}) = \sum_i \sum_{\mathbf{R}} v_i(\mathbf{r} - \mathbf{R} - \boldsymbol{\tau}_i) \tag{2.4}$$

for atom species i at basis site $\boldsymbol{\tau}_i$ in cell $\mathbf{R}$. Since $V(\mathbf{r})$ is periodic, the wave function is of the Bloch form

$$\begin{aligned} \phi_{n,\mathbf{k},s}(\mathbf{r}) &= e^{i\mathbf{k}\cdot\mathbf{r}}u_{n,\mathbf{k},s}(\mathbf{r}) \\ u_{n,\mathbf{k},s}(\mathbf{r}) &= u_{n,\mathbf{k},s}(\mathbf{r} + \mathbf{R}). \end{aligned} \tag{2.5}$$

The wave function $\langle \mathbf{r}|\phi_{n,\mathbf{k},s}\rangle$ is expanded in a set of normalised plane wave states

$$|\phi_{n,\mathbf{k},s}\rangle = \sum_{\mathbf{G},\zeta} a_{n,\mathbf{k},s}(\mathbf{G},\zeta)|\mathbf{k} + \mathbf{G},\zeta\rangle \tag{2.6}$$

where $\langle \mathbf{r}|\mathbf{k}+\mathbf{G},\zeta\rangle = (1/\sqrt{\Omega})\exp[i(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}]|\zeta\rangle$, $\mathbf{G}$ is a reciprocal lattice vector, Ω is the crystal volume, and $|\zeta\rangle = |\uparrow\rangle, |\downarrow\rangle$ is a spin state with a component of angular momentum $\pm(1/2)$ along the z direction. In principle, one could compute matrix elements of H_0 and V^{so} in this basis and obtain the solutions of (2.1) by direct diagonalisation.

Because of the large strength of the crystalline potential close to the lattice atoms, the character of the wave function changes drastically in this region. In the region close to the lattice sites the wave function takes on an atomic character and oscillates rapidly. At large distances from the lattice atoms it takes on the aspect of a slowly modulated free electron. To ensure rapid convergence of the series (2.6) under these conditions the true potential, V, is replaced with a pseudopotential, W. The pseudopotential has the same functional form as the real potential in the regions of the crystal occupied by valence electrons but smoothly approaches zero within the core region of a lattice atom. A consequence of this replacement is that we will be completely unable to describe the highly localised states occupied by the core electrons. However, since it is only the loosely bound valence electrons that participate in the optical and transport processes important for the design of semiconductor devices the substitution of the real crystalline potential by a pseudopotential is a valid one.

The matrix element for the pseudopotential in the plane wave basis is

$$\langle \mathbf{k}+\mathbf{G}',\zeta'|W|\mathbf{k}+\mathbf{G},\zeta\rangle = \delta_{\zeta'\zeta}\cdot\frac{1}{\Omega}\int_{\Omega}\sum_{i}\sum_{\mathbf{R}} w_i(\mathbf{r}-\mathbf{R}-\boldsymbol{\tau}_i)e^{i(\mathbf{G}-\mathbf{G}')\cdot\mathbf{r}}d\mathbf{r}. \quad (2.7)$$

Using a change of origin $\mathbf{r}-\mathbf{R}-\boldsymbol{\tau}_i \rightarrow \mathbf{r}$ and the fact that $\sum_{\mathbf{R}} e^{i(\mathbf{G}-\mathbf{G}')\cdot\mathbf{R}} = \mathcal{N}$, where $\mathcal{N}$ is the number of primitive unit cells in the crystal, the matrix element becomes

$$\langle \mathbf{k}+\mathbf{G}',\zeta'|W|\mathbf{k}+\mathbf{G},\zeta\rangle = \frac{1}{2}\cdot\delta_{\zeta'\zeta}\sum_{\boldsymbol{\tau}_i} e^{i(\mathbf{G}-\mathbf{G}')\cdot\boldsymbol{\tau}_i}w_i(\mathbf{G}-\mathbf{G}') \qquad (2.8)$$

with

$$w_i(\mathbf{G}-\mathbf{G}') = \frac{2}{\Omega_0}\int_\Omega e^{i(\mathbf{G}-\mathbf{G}')\cdot\mathbf{r}} w_i(\mathbf{r})d\mathbf{r}. \tag{2.9}$$

In (2.9) $\Omega_0 = \Omega/\mathcal{N}$ is the primitive cell volume. The term $\sum_{\boldsymbol{\tau}_i} e^{i(\mathbf{G}-\mathbf{G}')\cdot\boldsymbol{\tau}_i}$ is known as the structure factor and depends on the positions of atoms within the primitive cell. The term $w_i(\mathbf{G}-\mathbf{G}')$ is called the form factor and depends on the potential associated with the ith atom of the primitive cell. Group III-V materials such as GaAs are based upon the face centered cubic lattice with a two atom basis. If the lattice point is chosen to be at the bond centre with atoms positioned at $\pm\boldsymbol{\tau}$ then the potential matrix element is

$$\begin{aligned}&\langle \mathbf{k}+\mathbf{G}',\zeta'|W|\mathbf{k}+\mathbf{G},\zeta\rangle = \\ &\frac{1}{2}\cdot\delta_{\zeta'\zeta}\left[w_{Ga}(\mathbf{G}-\mathbf{G}')e^{i(\mathbf{G}-\mathbf{G}')\cdot\boldsymbol{\tau}} + w_{As}(\mathbf{G}-\mathbf{G}')e^{-i(\mathbf{G}-\mathbf{G}')\cdot\boldsymbol{\tau}}\right].\end{aligned} \tag{2.10}$$

Conventionally, linear combinations of the atomic pseudopotentials are taken to generate symmetric and antisymmetric form factors:

$$w^s(\mathbf{G}-\mathbf{G}') = \frac{1}{2}\left[w_{Ga}(\mathbf{G}-\mathbf{G}') + w_{As}(\mathbf{G}-\mathbf{G}')\right] \tag{2.11}$$

$$w^a(\mathbf{G}-\mathbf{G}') = \frac{1}{2}\left[w_{Ga}(\mathbf{G}-\mathbf{G}') - w_{As}(\mathbf{G}-\mathbf{G}')\right]. \tag{2.12}$$

The matrix element becomes

$$\begin{aligned}\langle \mathbf{k}+\mathbf{G}',\zeta'|W|\mathbf{k}+\mathbf{G},\zeta\rangle = \delta_{\zeta'\zeta}\cdot&\left[w^s(\mathbf{G}-\mathbf{G}')\cos(\mathbf{G}-\mathbf{G}')\cdot\boldsymbol{\tau}\right. \\ &\left.+iw^a(\mathbf{G}-\mathbf{G}')\sin(\mathbf{G}-\mathbf{G}')\cdot\boldsymbol{\tau}\right].\end{aligned} \tag{2.13}$$

These form factors are fitted empirically to obtain the experimentally determined band gaps of the semiconductor. The form factors at the first three nonzero values of $(\mathbf{G}-\mathbf{G}')$ at which the structure factor is also nonzero are used. The higher terms are set to zero using the assumption that the pseudopotential is smooth into the core region where the strong repulsive and attractive potentials of the crystalline potential nearly cancel.

	GaAs	AlAs	InAs
a_0	5.6500	5.6500	6.0500
$w^s(3)$	-0.2394	-0.2274	-0.2697
$w^s(8)$	-0.0070	0.0087	-0.0070
$w^s(11)$	0.0880	0.1017	0.0654
$w^a(3)$	0.0500	0.0500	0.0500
$w^a(4)$	0.0343	0.0220	0.0273
$w^a(11)$	0.0250	0.0250	0.0208
μ	0.0013	0.0002	0.0023
α	1.3800	9.9680	0.7952

Table 2.1: The parameters used in the empirical pseudopotential calculations outlined in the text. The (anti)symmetric form factors $w(|\mathbf{G}-\mathbf{G}'|^2)$ and the spin parameter μ are in Rydbergs. The lattice constant, a_0, is given in angstroms.

In the plane wave basis, the spin-orbit correction (2.3) has the form (Weisz, 1966; Bloom and Bergstresser, 1968)

$$\begin{aligned}\langle \mathbf{k}+\mathbf{G}',\zeta'|V^{so}|\mathbf{k}+\mathbf{G},\zeta\rangle &= (\mathbf{k}+\mathbf{G}')\wedge(\mathbf{k}+\mathbf{G})\cdot\sigma_{\zeta'\zeta} \\ &\times[-i\lambda^s\cos(\mathbf{G}-\mathbf{G}')\cdot\boldsymbol{\tau}+\lambda^a\sin(\mathbf{G}-\mathbf{G}')\cdot\boldsymbol{\tau}] \qquad (2.14)\end{aligned}$$

where the $\sigma_{\zeta'\zeta}$ are Pauli spin matrix elements. The symmetric and antisymmetric contributions to the spin-orbit coupling, λ^s and λ^a, are defined by

$$\lambda^s = \frac{1}{2}\mu(1+\alpha) \qquad (2.15)$$

$$\lambda^a = \frac{1}{2}\mu(1-\alpha). \qquad (2.16)$$

Here, α is the ratio of the spin-orbit splitting of the free atom levels tabulated by Hermann and Skillman (1963). The parameter μ is adjusted together with the form factors to reproduce the experimentally determined band structure.

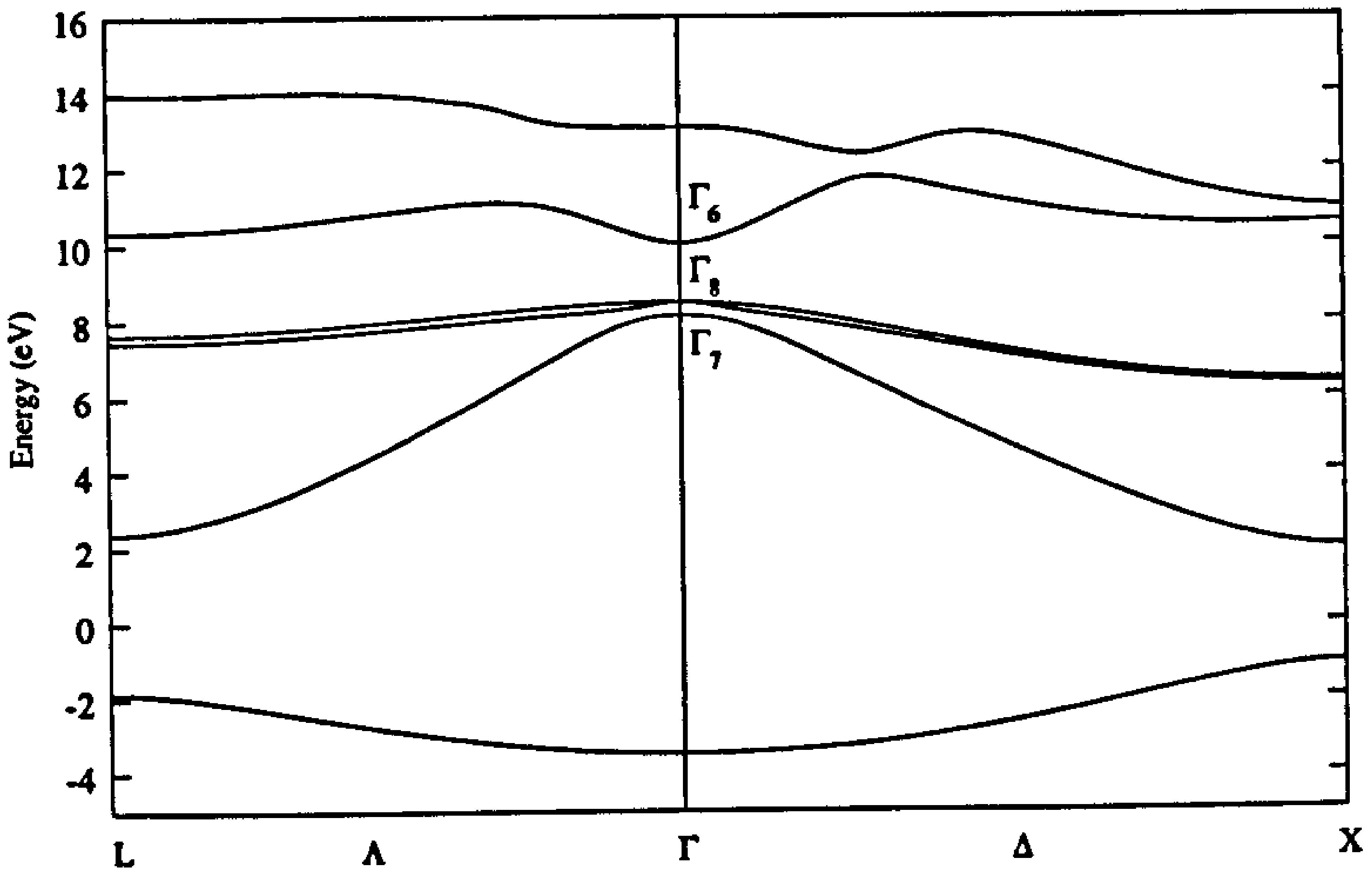

Figure 2.1: Band structure of GaAs.

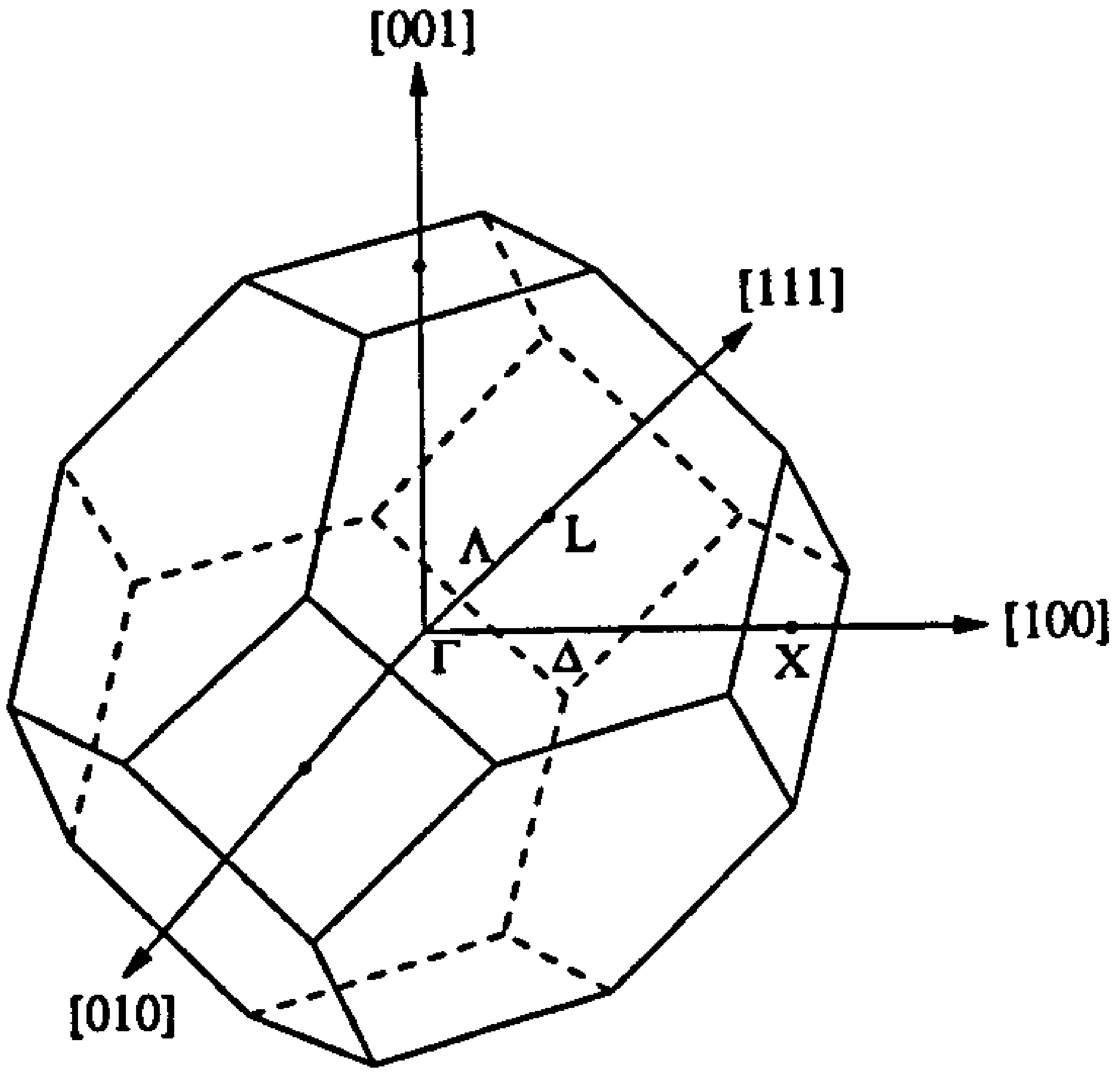

Figure 2.2: The first Brillouin zone for a crystal based on a face-centred cubic lattice.

The parameters required for an empirical pseudopotential calculation for the bulk band structure of several III-V compounds are given in Table 2.1.

The GaAs band structure calculated by the empirical pseudopotential method is shown in Figure 2.1, which represents the energy bands at different **k** points in the first Brillouin zone (shown in Figure 2.2).

2.1.2 Describing Heterostructures

The electronic structure of microstructures comprising of two or more bulk semiconductors can also be determined using the empirical pseudopotential model. Consider the method applied to a GaAs/AlAs quantum well. The empirical pseudopotential method requires that we consider an infinite array of unit cells, each containing a single GaAs/AlAs quantum well. When modelling the electronic structure of an isolated quantum well, the size of the unit cells are chosen to be large enough relative to the actual size of the quantum well that the periodic replicas of this structure do not interact with each other. However, it is often desirable to allow the states in neighbouring quantum well unit cells to interact. In these multi-quantum well structures, often referred to as superlattices, the barrier width is chosen to be small enough to allow this interaction to occur. The unit cell for a quantum well structure in the

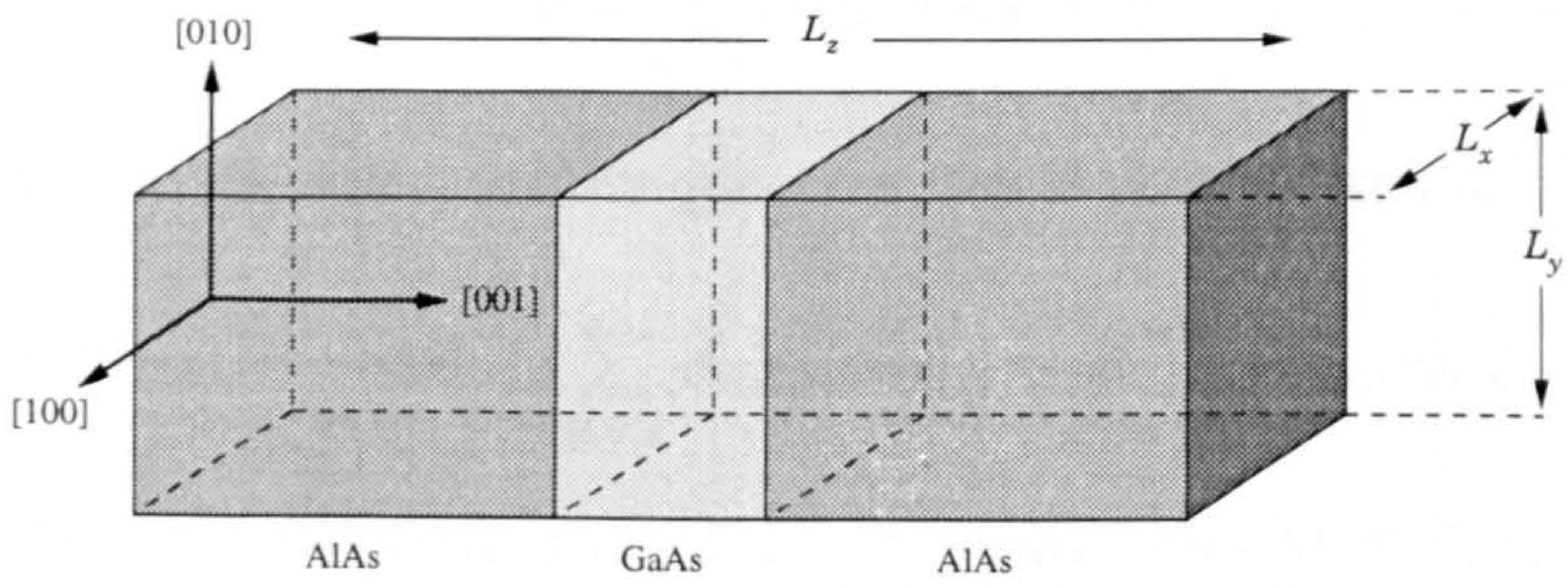

Figure 2.3: The unit cell for a GaAs/AlAs quantum well structure.

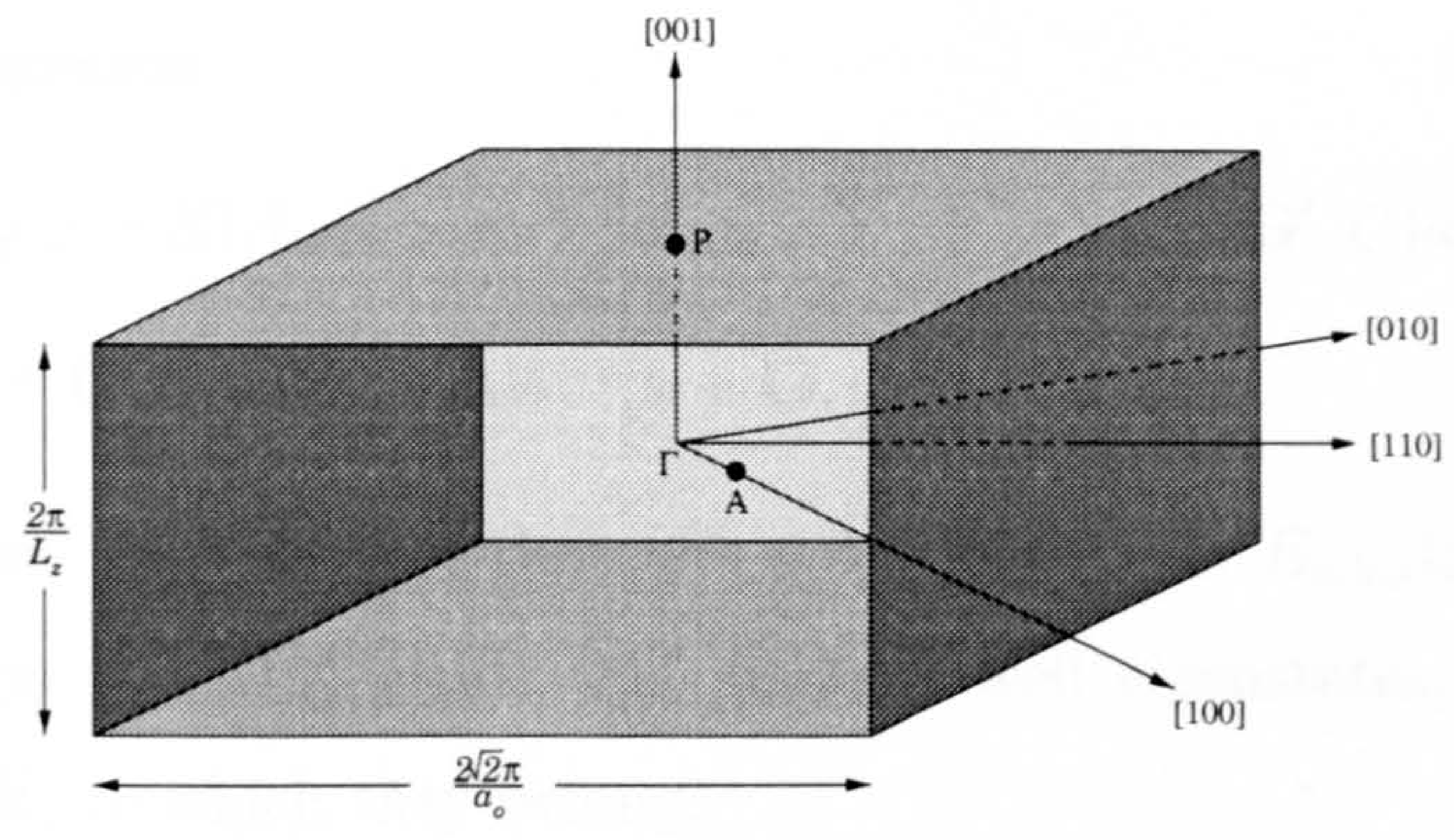

Figure 2.4: The first Brillouin zone for the quantum well unit cell shown in Figure 2.3.

empirical pseudopotential scheme is shown in Figure 2.3. The discontinuities in the band edges of the heterostructure form a potential well for electrons and an inverted well for holes. In Figure 2.4 we show the first Brillouin zone for the quantum well structure.

Following standard perturbation theory, the quantum well Hamiltonian consists of an unperturbed term from the GaAs layer plus a perturbation representing the difference between the microscopic potentials of AlAs and GaAs. Because the solutions to the Schrödinger equation for bulk GaAs form a complete set the quantum well states can be expanded in terms of them:

$$|\psi\rangle = \sum_{n,\mathbf{k},s} A_{n,\mathbf{k},s} |\phi_{n,\mathbf{k},s}\rangle. \tag{2.17}$$

The Schrödinger equation for the quantum well is thus written as

$$(H_0 + V^{so} + \Delta W + \Delta W^{so}) \sum_{n,\mathbf{k},s} A_{n,\mathbf{k},s} |\phi_{n,\mathbf{k},s}\rangle = E \sum_{n,\mathbf{k},s} A_{n,\mathbf{k},s} |\phi_{n,\mathbf{k},s}\rangle. \tag{2.18}$$

in which ΔW and ΔW^{so} are the difference between the AlAs-GaAs local potential and the AlAs-GaAs spin-orbit potential respectively. Substituting (2.6) into (2.18) and multiplying on the left by a dual basis vector $\langle \phi_{n',\mathbf{k}',s'}|$ gives the matrix equation

$$(E_{n',\mathbf{k}',s'} - E)A_{n',\mathbf{k}',s'} + \sum_{n,\mathbf{k},s} A_{n,\mathbf{k},s} \sum_{\mathbf{G},\mathbf{G}'} \sum_{\zeta,\zeta'} a^*_{n',\mathbf{k}',s'}(\mathbf{G}',\zeta')a_{n,\mathbf{k},s}(\mathbf{G},\zeta)$$
$$\times\langle \mathbf{k}'+\mathbf{G}',\zeta'|\Delta W + \Delta W^{so}|\mathbf{k}+\mathbf{G},\zeta\rangle = 0 \tag{2.19}$$

where we have used the fact that $(H_0 + V^{so})|\phi_{n,\mathbf{k},s}\rangle = E_{n,\mathbf{k},s}|\phi_{n,\mathbf{k},s}\rangle$. Direct diagonalisation of (2.19) yields the quantum well eigenstates, $|\psi\rangle$, and the eigenvalues, E, to which they belong.

The quantum well potential matrix element is

$$\langle \mathbf{k}'+\mathbf{G}',\zeta'|\Delta W|\mathbf{k}+\mathbf{G},\zeta\rangle = \delta_{\zeta'\zeta}\cdot\frac{1}{\Omega}\int_\Omega \sum_{\mathbf{R}}\Bigg[\sum_i\{w_i(\mathbf{r}-\tau_i-\mathbf{R})$$
$$-w_{Ga}(\mathbf{r}-\tau_i-\mathbf{R})\} + \sum_j\{w_j(\mathbf{r}-\tau_j-\mathbf{R}) - w_{As}(\mathbf{r}-\tau_j-\mathbf{R})\}\Bigg] e^{i\mathbf{g}\cdot\mathbf{r}}d\mathbf{r} \tag{2.20}$$

where $\mathbf{R}$ is now a lattice vector of the quantum well direct lattice, and $\mathbf{g} = \mathbf{k}+\mathbf{G}-\mathbf{k}'-\mathbf{G}'$ is a quantum well reciprocal lattice vector with the components defined as

$$g_x \equiv \frac{2\pi l}{L_x} \quad g_y \equiv \frac{2\pi m}{L_y} \quad g_z \equiv \frac{2\pi n}{L_z}. \tag{2.21}$$

L_x, L_y and L_z are the dimensions of the quantum well unit cell and l, m and n are integers. In (2.20) the summations involving indices i and j are over all cation and anion sites in the quantum well unit cell respectively. Making a change of origin $\mathbf{r}-\tau-\mathbf{R} \rightarrow \mathbf{r}$ and using the reciprocal space expressions for the atomic pseudopotentials (2.9) we get

$$\langle \mathbf{k}'+\mathbf{G}',\zeta'|\Delta W|\mathbf{k}+\mathbf{G},\zeta\rangle = \delta_{\zeta'\zeta}\cdot\frac{\Omega_0}{2\Omega_{QW}}$$
$$\times\left[\sum_i\{w_i(\mathbf{g}) - w_{Ga}(\mathbf{g})\}e^{i\mathbf{g}\cdot\tau_i} + \sum_j\{w_j(\mathbf{g}) - w_{As}(\mathbf{g})\}e^{i\mathbf{g}\cdot\tau_j}\right] \tag{2.22}$$

where we have taken advantage of the relationship $\sum_{\mathbf{R}} e^{i\mathbf{g}\cdot\mathbf{R}} = \mathcal{N}_{QW} = \Omega/\Omega_{QW}$, where $\mathcal{N}_{QW}$ is the number of quantum well unit cells in the crystal and Ω_{QW}

is the volume of a single quantum well unit cell. In terms of the symmetric and antisymmetric form factors for bulk AlAs and GaAs, a potential matrix element is

$$\begin{aligned}\langle \mathrm{k}' + \mathrm{G}', \zeta' | \Delta W | \mathrm{k} + \mathrm{G}, \zeta \rangle &= \delta_{\zeta'\zeta} \cdot \frac{\Omega_0}{2\Omega_{QW}} \\ \times &\Bigg[\sum_i \{ w^s_{AlAs}(\mathrm{g}) + w^a_{AlAs}(\mathrm{g}) - w^s_{GaAs}(\mathrm{g}) - w^a_{GaAs}(\mathrm{g}) \} e^{i\mathrm{g}\cdot\boldsymbol{T}_i} \\ &+ \sum_j \{ w^s_{AlAs}(\mathrm{g}) - w^a_{AlAs}(\mathrm{g}) - w^s_{GaAs}(\mathrm{g}) + w^a_{GaAs}(\mathrm{g}) \} e^{i\mathrm{g}\cdot\boldsymbol{T}_j} \Bigg] \end{aligned} \tag{2.23}$$

where i and j are summed over barrier cations and barrier anions respectively.

For the spin-orbit term ΔW^{so} we have

$$\begin{aligned}\langle \mathrm{k}' + \mathrm{G}', \zeta' | \Delta W^{so} | \mathrm{k} + \mathrm{G}, \zeta \rangle &= (\mathrm{k}' + \mathrm{G}') \wedge (\mathrm{k} + \mathrm{G}) \cdot \sigma_{\zeta'\zeta} \\ \times &\Bigg[\sum_i \{ \lambda^s_{AlAs} + \lambda^a_{AlAs} - \lambda^s_{GaAs} - \lambda^a_{GaAs} \} e^{i\mathrm{g}\cdot\boldsymbol{T}_i} \\ &+ \sum_j \{ \lambda^s_{AlAs} - \lambda^a_{AlAs} - \lambda^s_{GaAs} + \lambda^a_{GaAs} \} e^{i\mathrm{g}\cdot\boldsymbol{T}_j} \Bigg] \end{aligned} \tag{2.24}$$

where i and j are summed over barrier cations and barrier anions respectively.

To evaluate the matrix element (2.23) knowledge of the bulk symmetric and antisymmetric pseudopotential form factors at quantum well reciprocal lattice vectors is required. Values are known only at the bulk reciprocal lattice vectors. To determine the form factors at the g a curve is fitted through the known values. The symmetric form factors $w^s_{GaAs}(0)$ and $w^s_{AlAs}(0)$ control the relative alignment of the bulk band structures and therefore the band offsets (ΔE_c and ΔE_v). Values for these form factors are chosen such that the valence band offset is 500 meV, taken from the experimental and theoretical comparison of Lambrecht and Segall (1988).

2.1.3 Transition Probabilities in Quantum Wells

The forthcoming investigation into the nonlinear optical properties of quantum well systems requires that we consider what happens when an electron confined to a quantum well interacts with a photon of an applied optical field. Since the photon has negligible momentum in comparison to the electron, we consider the electron's wave vector before and after the interaction to be equal to a vector $\mathbf{k}$. Fermi's Golden Rule gives the transition probability per unit time, $P_{i\rightarrow f}$, that an electron under the action of a perturbation makes a transition from a subband labelled i to a subband labelled f as

$$P_{i\rightarrow f} = \left(\frac{2\pi eA}{\hbar mc}\right) |\hat{\mathrm{e}} \cdot \mathbf{M}_{i\rightarrow f}(\mathbf{k})|^2 \delta(E_f(\mathbf{k}) - E_i(\mathbf{k}) - \hbar\omega) \tag{2.25}$$

where A is the magnitude of the vector potential of the electromagnetic field, ω is the frequency of the field and $\hat{\mathrm{e}}$ is the unit polarisation vector.

The main band structure dependent quantity is the optical matrix element

$$\hat{\mathrm{e}} \cdot \mathbf{M}_{i\rightarrow f} = \langle f|\hat{\mathrm{e}} \cdot \mathbf{p}|i\rangle \equiv p^{\eta}_{fi} \tag{2.26}$$

in which $\mathbf{p} = -i\hbar\nabla$ is the momentum operator and p^{η}_{fi} is the momentum matrix element resulting from the interaction of light polarised in the $\eta (= x, y, z)$ direction. The initial quantum well state is given by (2.17) and (2.6) to be

$$|i\rangle = \sum_{n,\mathbf{k},s} A^{i}_{n,\mathbf{k},s} \sum_{\mathbf{G},\zeta} a_{n,\mathbf{k},s}(\mathbf{G},\zeta)|\mathbf{k}+\mathbf{G},\zeta\rangle. \tag{2.27}$$

The dual vector of the final state is

$$\langle f| = \sum_{n',\mathbf{k}',s'} A^{f*}_{n',\mathbf{k}',s'} \sum_{\mathbf{G}',\zeta'} a^{*}_{n',\mathbf{k}',s'}(\mathbf{G}',\zeta')\langle \mathbf{k}'+\mathbf{G}',\zeta'|. \tag{2.28}$$

We find that the momentum matrix element in a quantum well system is

$$p^{\eta}_{fi} = \hbar \sum_{\mathbf{k}} \sum_{n,n'} \sum_{s,s'} A^{f*}_{n',\mathbf{k},s'} A^{i}_{n,\mathbf{k},s} \sum_{\mathbf{G},\zeta} a^{*}_{n',\mathbf{k},s'}(\mathbf{G},\zeta) a_{n,\mathbf{k},s}(\mathbf{G},\zeta)(k_\eta + G_\eta). \tag{2.29}$$

In the above expression k_η and G_η are the components of the wave vector and bulk reciprocal lattice vector in the direction of the electric field vector.

2.2 The $\mathbf{k} \cdot \mathbf{p}$ Method

For optical devices, the most important regions of the band structure are those most commonly populated by the charge carriers of the crystal: the lowest conduction band valley and the top of the valence band. The $\mathbf{k} \cdot \mathbf{p}$ method (Kane, 1966) provides a limited description of the electronic structure near these band extrema. For III-V direct bandgap materials these extrema occur at the Brillouin zone centre.

2.2.1 Representing Bands of Bulk Materials

By the $\mathbf{k} \cdot \mathbf{p}$ scheme, if the solution of the one electron Schrödinger equation (2.1) for a direct bandgap semiconductor is known at the zone centre it is possible to obtain solutions in the immediate neighbourhood by regarding the scalar product $\mathbf{k} \cdot \mathbf{p}$, where $\mathbf{k}$ is a wave vector measured from Γ, as a perturbation.

Consider the one electron Schrödinger equation (2.1) written in terms of the periodic part of the Bloch function (2.5) using $\mathbf{p} = -i\hbar\boldsymbol{\nabla}$:

$$Hu_{\mathbf{k}}(\mathbf{r}) = \left[H_0 + V^{so} + \frac{\hbar}{m_0}\mathbf{k} \cdot \mathbf{p} + \frac{\hbar^2 k^2}{2m_0}\right] u_{\mathbf{k}}(\mathbf{r}) = E_{\mathbf{k}} u_{\mathbf{k}}(\mathbf{r}). \tag{2.30}$$

The band label and spin label have been dropped for simplicity. It is convenient to divide the unperturbed eigenfunctions and energies into two classes, A and B. We place the states close to the band edges (e.g. $\Gamma_6, \Gamma_8, \Gamma_7$ in GaAs) into class A and all other higher and lower bands into class B. Then, using the perturbation method due to Löwdin (1951), treat the influence of the states in B on those in A as a perturbation.

We expand the function

$$u_{\mathbf{k}}(\mathbf{r}) = \sum_{j'}^{A} a_{j',\mathbf{k}} u_{j',0}(\mathbf{r}) + \sum_{\gamma}^{B} a_{\gamma,\mathbf{k}} u_{\gamma,0}(\mathbf{r}) \tag{2.31}$$

where j' is in class A and γ is in class B. Because the electron wave functions are P-like near the top of the valence band and S-like near the bottom of the conduction band we choose for the unperturbed basis functions $|S\rangle$, $|X\rangle$, $|Y\rangle$, $|Z\rangle$ which transform like atomic S, P_x, P_y, P_z functions under the symmetry operations which map the local tetrahedron onto itself. Conventionally, we take linear combinations of these functions such that the spin-orbit interaction term is diagonal in this basis. Specifically, we have in class A

$$\begin{aligned}
u_{1,0} &\equiv \left|\frac{1}{2},\frac{1}{2}\right\rangle = |iS\uparrow\rangle \\
u_{2,0} &\equiv \left|\frac{1}{2},-\frac{1}{2}\right\rangle = |iS\downarrow\rangle \\
u_{3,0} &\equiv \left|\frac{3}{2},\frac{3}{2}\right\rangle = -\frac{1}{\sqrt{2}}|(X+iY)\uparrow\rangle \\
u_{4,0} &\equiv \left|\frac{3}{2},\frac{1}{2}\right\rangle = -\frac{1}{\sqrt{6}}|(X+iY)\downarrow\rangle + \sqrt{\frac{2}{3}}|Z\uparrow\rangle \\
u_{5,0} &\equiv \left|\frac{3}{2},-\frac{1}{2}\right\rangle = \frac{1}{\sqrt{6}}|(X-iY)\uparrow\rangle + \sqrt{\frac{2}{3}}|Z\downarrow\rangle \\
u_{6,0} &\equiv \left|\frac{3}{2},-\frac{3}{2}\right\rangle = \frac{1}{\sqrt{2}}|(X-iY)\downarrow\rangle \\
u_{7,0} &\equiv \left|\frac{1}{2},\frac{1}{2}\right\rangle = \frac{1}{\sqrt{3}}|(X+iY)\downarrow\rangle + \frac{1}{\sqrt{3}}|Z\uparrow\rangle \\
u_{8,0} &\equiv \left|\frac{1}{2},-\frac{1}{2}\right\rangle = \frac{1}{\sqrt{3}}|(X-iY)\uparrow\rangle - \frac{1}{\sqrt{3}}|Z\downarrow\rangle
\end{aligned} \tag{2.32}$$

representing the Brillouin zone centre wave functions for the spin degenerate conduction band, heavy hole band, light hole band and split-off band respectively. These unperturbed wave functions have the same transformation properties as the eigenfunctions of the total momentum operator $\mathbf{J}(=\mathbf{L}+\mathbf{S})$ and we label them using the total angular momentum and its projection onto the z-axis, i.e., $|J, M_J\rangle$.

With Löwdin's method we need only solve the eigenequation

$$\sum_{j'}^{A}(U^A_{jj'} - E\delta_{jj'})a_{j',\mathbf{k}} = 0 \tag{2.33}$$

instead of

$$\sum_{j'}^{A,B}(H_{jj'} - E\delta_{jj'})a_{j',\mathbf{k}} = 0 \tag{2.34}$$

where to first order

$$\begin{aligned}
U^A_{jj'} &= H_{jj'} + \sum_{\gamma\neq j,j'}^{B} \frac{H_{j\gamma}H_{\gamma j'}}{E_0 - E_\gamma} \\
H_{jj'} &= \langle u_{j,0}|H|u_{j',0}\rangle = \left[E_j(0) + \frac{\hbar^2 k^2}{2m_0}\right]\delta_{jj'} + \sum_{\alpha=x,y,z}\frac{\hbar k_\alpha}{m_0}p^\alpha_{jj'} \quad (j,j'\in A) \\
H_{j\gamma} &= \langle u_{j,0}|\frac{\hbar}{m_0}\mathbf{k}\cdot\mathbf{p}|u_{\gamma,0}\rangle = \sum_{\alpha=x,y,z}\frac{\hbar k_\alpha}{m_0}p^\alpha_{j\gamma} \quad (j\in A, \gamma\notin A)
\end{aligned} \tag{2.35}$$

where E_0 is the average energy of the states in class A. We thus obtain

$$U^A_{jj'} = \left[E_j(0) + \frac{\hbar^2 k^2}{2m_0}\right]\delta_{jj'} + \sum_\alpha \frac{\hbar k_\alpha}{m_0}p^\alpha_{jj'} + \frac{\hbar^2}{m_0^2}\sum_{\gamma\neq j,j'}^{B}\sum_{\alpha,\beta}\frac{k_\alpha k_\beta p^\alpha_{j\gamma}p^\beta_{\gamma j'}}{E_0 - E_\gamma}. \tag{2.36}$$

Let $U^A_{jj'} \equiv D_{jj'}$. We obtain the matrix of the form $D_{jj'}$

$$D_{jj'} = E_{j'}(0)\delta_{jj'} + \sum_\alpha \frac{\hbar k_\alpha}{m_0}p^\alpha_{jj'} + \sum_{\alpha,\beta} D^{\alpha\beta}_{jj'}k_\alpha k_\beta \tag{2.37}$$

where $D^{\alpha\beta}_{jj'}$ is defined as

$$D^{\alpha\beta}_{jj'} = \frac{\hbar^2}{2m_0}\left[\delta_{jj'}\delta_{\alpha\beta} + \sum_\gamma^B \frac{p^\alpha_{j\gamma}p^\beta_{\gamma j'} + p^\beta_{j\gamma}p^\alpha_{\gamma j'}}{m_0(E_0 - E_\gamma)}\right]. \tag{2.38}$$

The effects of $V(\mathbf{r})$ are now encoded in the momentum matrix elements $p_{nn'} \equiv \langle u_{n,0}|\mathbf{p}|u_{n',0}\rangle$.

Suppose we are mainly interested in the top four valence bands (the heavy hole and light hole bands, both doubly degenerate) and we ignore the coupling to the conduction and split-off bands. We put the top four valence states into class A and all other bands into class B. To write out explicitly the elements of the matrix $D_{jj'}$ for this situation, we define

$$\begin{aligned} A_0 &= \frac{\hbar^2}{2m_0} + \frac{\hbar^2}{2m_0^2} \sum_{\gamma}^{B} \frac{p^x_{X\gamma} p^x_{\gamma X}}{E_0 - E_\gamma} \\ B_0 &= \frac{\hbar^2}{2m_0} + \frac{\hbar^2}{m_0^2} \sum_{\gamma}^{B} \frac{p^y_{X\gamma} p^y_{\gamma X}}{E_0 - E_\gamma} \\ C_0 &= \frac{\hbar^2}{m_0^2} \sum_{\gamma}^{B} \frac{p^x_{X\gamma} p^y_{\gamma Y} + p^y_{X\gamma} p^x_{\gamma Y}}{E_0 - E_\gamma} \end{aligned} \tag{2.39}$$

and define the band structure parameters γ_1, γ_2, γ_3 (Luttinger and Kohn, 1955; Luttinger, 1956) as

$$\begin{aligned} \gamma_1 &= -\frac{2m_0}{3\hbar^2}(A_0 + 2B_0) \\ \gamma_2 &= -\frac{m_0}{3\hbar^2}(A_0 - B_0) \\ \gamma_3 &= -\frac{m_0}{3\hbar^2} C_0. \end{aligned} \tag{2.40}$$

The Luttinger parameters $\{\gamma_i\}$ can be written in terms of the effective mass of the heavy hole and light hole bands along different crystal directions (Gershoni *et al.*, 1993):

$$\begin{aligned} \frac{m_0}{m_{hh}^{[001]}} &= \gamma_1 - 2\gamma_2 \\ \frac{m_0}{m_{lh}^{[001]}} &= \gamma_1 + 2\gamma_2 \\ \frac{m_0}{m_{hh}^{[111]}} &= \gamma_1 - 2\gamma_3 \\ \frac{m_0}{m_{lh}^{[111]}} &= \gamma_1 + 2\gamma_3. \end{aligned} \tag{2.41}$$

In most applications of the $\mathbf{k} \cdot \mathbf{p}$ method the $\{\gamma_i\}$ are fitted to the experimental [001] and [111] masses at Γ. We obtain the Luttinger-Kohn Hamiltonian matrix (Luttinger and Kohn, 1955) $\overline{\overline{\mathbf{U}}}^{A} \equiv \overline{\overline{\mathbf{D}}}$, denoted as $\overline{\overline{\mathbf{H}}}^{LK}$, in the basis functions given by (2.32)

$$\overline{\overline{\mathbf{H}}}^{LK} = \begin{array}{c} \begin{bmatrix} -P-Q & S & -R & 0 \\ S^* & -P+Q & 0 & -R \\ -R^* & 0 & -P+Q & -S \\ 0 & -R^* & -S^* & -P-Q \end{bmatrix} \end{array} \begin{array}{l} \left|\frac{3}{2},\frac{3}{2}\right\rangle \\ \left|\frac{3}{2},\frac{1}{2}\right\rangle \\ \left|\frac{3}{2},-\frac{1}{2}\right\rangle \\ \left|\frac{3}{2},-\frac{3}{2}\right\rangle \end{array}$$

$$\begin{aligned} P &= \frac{\hbar^2}{2m_0}\gamma_1(k_x^2+k_y^2+k_z^2) \\ Q &= \frac{\hbar^2}{2m_0}\gamma_2(k_x^2+k_y^2-2k_z^2) \\ R &= \frac{\hbar^2}{2m_0}\sqrt{3}\left[-\gamma_2(k_x^2-k_y^2)+2i\gamma_3k_xk_y\right] \\ S &= \frac{\hbar^2}{2m_0}2\sqrt{3}\gamma_3(k_x-ik_y)k_z \end{aligned} \tag{2.42}$$

in which '$*$' denotes the Hermitian conjugate. The approximate solution to (2.1) can be obtained from the eigenvalues and eigenvectors of the Luttinger-Kohn matrix. The wave function $\phi_{n,\mathbf{k}}(\mathbf{r})$ is given by

$$\phi_{n,\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}}u_{n,\mathbf{k}}(\mathbf{r}) \tag{2.43}$$

$$u_{n,\mathbf{k}}(\mathbf{r}) = \sum_{j=1}^{4} a_{j,\mathbf{k}}u_{j,0}(\mathbf{r}) \tag{2.44}$$

and $E_{n,\mathbf{k}} = E$.

2.2.2 Strain Effects on Band Structures

The introduction of a homogeneous strain in a bulk semiconductor changes the lattice parameter and in some cases, the symmetry of the material. These in turn produce significant changes in the electronic band structure. Homogeneous strained configurations can quite generally be divided into two configurations: the isotropic or hydrostatic component, which gives rise to a volume change without disturbing the crystal symmetry, and the anisotropic component which in general reduces the symmetry present in the strain-free lattice.

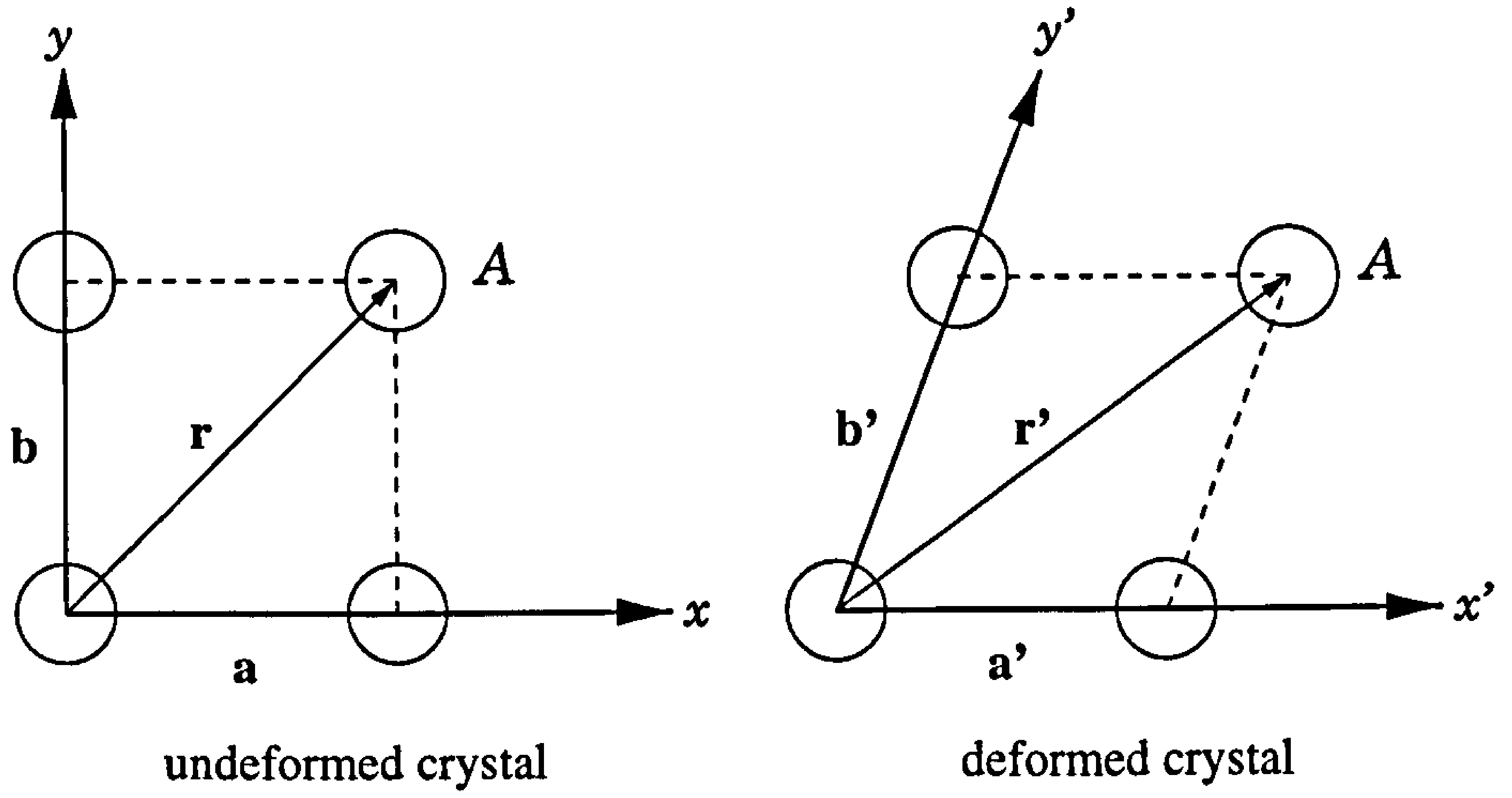

Figure 2.5: Axes of a crystal under uniform deformation.

A homogeneous strain induces change in energy gaps and, when the symmetry of the crystal is reduced, removes degeneracies.

In this section, the effects of a uniform deformation on the conduction and valence band edges of III-V semiconductors at Γ are reviewed. The additional terms to be added to the one electron Schrödinger equation (2.1) to account for strain in bulk semiconductors are discussed using the language of the $\mathbf{k}\cdot\mathbf{p}$ scheme.

As shown in Figure 2.5, the unit vectors $\mathbf{a}$, $\mathbf{b}$, (and $\mathbf{c}$) in the undeformed crystal are related to $\mathbf{a}'$, $\mathbf{b}'$, (and $\mathbf{c}'$) in the uniformly deformed crystal by

$$\begin{aligned}\mathbf{a}' &= (1+\epsilon_{xx})\mathbf{a} + \epsilon_{xy}\mathbf{b} + \epsilon_{xz}\mathbf{c}\\ \mathbf{b}' &= \epsilon_{yx}\mathbf{a} + (1+\epsilon_{yy})\mathbf{b} + \epsilon_{yz}\mathbf{c}\\ \mathbf{c}' &= \epsilon_{zx}\mathbf{a} + \epsilon_{zy}\mathbf{b} + (1+\epsilon_{zz})\mathbf{c}. \end{aligned} \tag{2.45}$$

We assume a homogeneous strain and a symmetric strain tensor i.e. $\epsilon_{xy} = \epsilon_{yx}$. For a cubic crystal under these conditions there are six independent nonzero components of the strain tensor. These are:-

$$\epsilon_{xx},\ \epsilon_{yy},\ \epsilon_{zz},\ \epsilon_{xy}=\epsilon_{yx},\ \epsilon_{yz}=\epsilon_{zy},\ \epsilon_{zx}=\epsilon_{xz}. \tag{2.46}$$

To label a position A (or atom A) in the undeformed crystal, we have

$$\mathbf{r} = x\mathbf{a} + y\mathbf{b} + z\mathbf{c}. \tag{2.47}$$

The same atom in the deformed crystal can be labelled either as

$$\mathbf{r}' = x\mathbf{a}' + y\mathbf{b}' + z\mathbf{c}' \tag{2.48}$$

using the new basis vectors $\mathbf{a}'$, $\mathbf{b}'$, $\mathbf{c}'$ or as

$$\mathbf{r}' = x'\mathbf{a} + y'\mathbf{b} + z'\mathbf{c} \tag{2.49}$$

in the original bases of the deformed crystal. In the linear strain regime, the change of volume becomes

$$\frac{V+\delta V}{V} = \mathbf{a}'\cdot\mathbf{b}'\wedge\mathbf{c}' = 1 + (\epsilon_{xx}+\epsilon_{yy}+\epsilon_{zz}). \tag{2.50}$$

The quantity $\epsilon_{xx}+\epsilon_{yy}+\epsilon_{zz}$ is the trace of the strain matrix $\bar{\bar{\epsilon}}$, or $\mathrm{Tr}(\bar{\bar{\epsilon}})$, which is exactly the fractional change of the volume $\delta V/V$ of the crystal under uniform deformation:

$$\frac{\delta V}{V} = \epsilon_{xx}+\epsilon_{yy}+\epsilon_{zz}. \tag{2.51}$$

The effect of a strain on the conduction band of a III-V semiconductor is to produce a hydrostatic energy shift proportional to the fractional volume change given by

$$H_\epsilon^c = a_c(\epsilon_{xx}+\epsilon_{yy}+\epsilon_{zz}) \tag{2.52}$$

where a_c is the intraband (absolute) hydrostatic deformation potential of the conduction band. In the valence band, the orbital-strain Hamiltonian H_ϵ^v can be written as (Pollak and Cardona, 1968)

$$H_\epsilon^v = -a(\epsilon_{xx}+\epsilon_{yy}+\epsilon_{zz}) - 3b[(L_x^2-\frac{1}{3}\mathbf{L}^2)+\mathrm{c.p.}] - \frac{6d}{\sqrt{3}}[\{L_xL_y\}\epsilon_{xy}+\mathrm{c.p.}] \tag{2.53}$$

where **L** is the angular momentum operator, c.p. denotes cyclic permutations with respect to the indices x, y and z, and the quantities in curly brackets indicate the symmetrised product: $\{L_x L_y\} = \frac{1}{2}(L_x L_y + L_y L_x)$. The parameter a is the hydrostatic deformation potential. It represents the intraband (absolute) shift of the orbital valence bands due to the hydrostatic component of the strain. The quantities b and d are uniaxial deformation potentials appropriate to strains of tetragonal and rhombohedral symmetries, respectively. Values for the deformation potentials a_c, a, b and d are given in Table 2.2 for GaAs and InAs, taken from the summary of experimental and theoretical results by Van der Walle (1989). We will neglect the stress-dependent, spin-orbit contribution to the valence band strain Hamiltonian since it is small compared to the orbital-strain contribution (Pollak, 1973).

In the eight-band $|J, M_J\rangle$ representation the total effective Hamiltonian for a bulk semiconductor, $H = H_0 + V^{so} + H^c_\epsilon + H^v_\epsilon$, can be written at the centre of the Brillouin zone as (Pollak, 1990)

$$\overline{\overline{\mathbf{H}}} = \left[\begin{array}{cccccccc}
A & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & A & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & B & E & F & 0 & G & H \\
0 & 0 & E^* & C & 0 & F & I & J \\
0 & 0 & F^* & 0 & C & E & J & -I \\
0 & 0 & 0 & F^* & E^* & B & -H^* & G \\
0 & 0 & G^* & I^* & J^* & -H & D & 0 \\
0 & 0 & H^* & J^* & -I^* & G^* & 0 & D
\end{array}\right]
\begin{array}{l}
|S\uparrow\rangle \\
|S\downarrow\rangle \\
\left|\frac{3}{2},\frac{3}{2}\right\rangle \\
\left|\frac{3}{2},\frac{1}{2}\right\rangle \\
\left|\frac{3}{2},-\frac{1}{2}\right\rangle \\
\left|\frac{3}{2},-\frac{3}{2}\right\rangle \\
\left|\frac{1}{2},\frac{1}{2}\right\rangle \\
\left|\frac{1}{2},-\frac{1}{2}\right\rangle
\end{array}$$

$$A = E_0 + a_c(\epsilon_{xx} + \epsilon_{yy} + \epsilon_{zz})$$

$$B = -a(\epsilon_{xx} + \epsilon_{yy} + \epsilon_{zz}) - (b/2)(2\epsilon_{zz} - \epsilon_{xx} - \epsilon_{yy})$$

$$C = -a(\epsilon_{xx} + \epsilon_{yy} + \epsilon_{zz}) + (b/2)(2\epsilon_{zz} - \epsilon_{xx} - \epsilon_{yy})$$

$$
\begin{aligned}
D &= -\Delta - a(\epsilon_{xx} + \epsilon_{yy} + \epsilon_{zz}) \\
E &= d(\epsilon_{xz} - \epsilon_{yz}) \\
F &= (\sqrt{3}/2)b(\epsilon_{xx} - \epsilon_{yy}) - id\epsilon_{xy} \\
G &= -d(\epsilon_{xz} - i\epsilon_{yz}) \\
H &= (\sqrt{3/2})b(\epsilon_{xx} - \epsilon_{yy}) + (i/\sqrt{2})d\epsilon_{xy} \\
I &= -(b/\sqrt{2})(2\epsilon_{zz} - \epsilon_{xx} - \epsilon_{yy}) \\
J &= (\sqrt{3/2})d(\epsilon_{xz} - i\epsilon_{yz}).
\end{aligned}
\tag{2.54}
$$

Diagonalisation of (2.54) yields the positions of the band edges in a bulk semiconductor subject to an arbitrary uniform deformation.

Consider the form of the effective Hamiltonian for the case of a strictly hydrostatic strain. Under these conditions the strain matrix is diagonal, the only nonzero elements being $\epsilon_{xx} = \epsilon_{yy} = \epsilon_{zz}$. From (2.54) it is apparent that the total Hamiltonian is also diagonal for this strain configuration. If the stress is compressive, the negative fractional volume change and the negative conduction band deformation potential common to most semiconductors combine to shift the lowest conduction band at Γ up in energy relative to its unstrained position. A compressive strain also causes a positive shift to the heavy hole, light hole and split-off bands relative to their unstrained positions. No splitting of the heavy hole and light hole bands occurs since the cubic symmetry is retained.

A uniaxial stress applied parallel to the [001] axis results in a diagonal

	a_c (eV)	a (eV)	b (eV)	d (eV)	E_0 (eV)	Δ (eV)
GaAs	-7.17	1.16	-1.7	-4.5	1.52	0.34
InAs	-5.08	1.00	-1.8	-3.6	0.41	0.38

Table 2.2: Deformation potentials, band gap and spin splitting of GaAs and InAs.

strain matrix with the nonzero components ϵ_{zz} and $\epsilon_{xx} = \epsilon_{yy}$. In this case additional so-called biaxial terms enter into the effective Hamiltonian through the terms B, C and I. The biaxial strain, which can be thought of as a hydrostatic strain plus a uniaxial strain, is proportional to $2\epsilon_{zz} - \epsilon_{xx} - \epsilon_{yy}$. The uniaxial stress splits the heavy hole and light hole states at Γ due to the reduction of symmetry, and introduces a coupling between the heavy hole and split-off states.

2.2.3 Describing Heterostructures

The $\mathbf{k}\cdot\mathbf{p}$ theory in conjunction with the effective mass approximation (EMA) has been used to describe the electronic energy levels of semiconductor heterostructures (Bastard and Brum, 1986). The effective mass approximation for the heavy hole and light hole bands is stated as follows (Luttinger and Kohn, 1955). If the dispersion relation of the set of degenerate bands satisfying (2.1) is given by

$$\sum_{j'=1}^{4} H_{jj'}^{LK} a_{j',\mathbf{k}} \equiv \sum_{j'=1}^{4} \left[E_{j,0}\delta_{jj'} + \sum_{\alpha,\beta} D_{jj'}^{\alpha\beta} k_\alpha k_\beta \right] a_{j',\mathbf{k}} = E_{\mathbf{k}} a_{j,\mathbf{k}} \tag{2.55}$$

then the solution $\psi(\mathbf{r})$ for the semiconductors in the presence of a perturbation $U(\mathbf{r})$, such as an impurity potential or a quantum well potential

$$\left[H_0 + V^{so} + U(\mathbf{r})\right] \psi(\mathbf{r}) = E\psi(\mathbf{r}) \tag{2.56}$$

is given by

$$\psi(\mathbf{r}) = \sum_{j=1}^{4} F_j(\mathbf{r}) u_{j,0}(\mathbf{r}) \tag{2.57}$$

where $F_j(\mathbf{r})$ satisfies

$$\sum_{j'=1}^{4} \left[E_{j,0}\delta_{jj'} + \sum_{\alpha,\beta} D_{jj'}^{\alpha\beta} \left(-i\frac{\partial}{\partial\alpha}\right)\left(-i\frac{\partial}{\partial\beta}\right) + U(\mathbf{r})\delta_{jj'} \right] F_{j'}(\mathbf{r}) = EF_j(\mathbf{r}). \tag{2.58}$$

The envelope function can be expanded in a set of normalised plane wave states

$$F_j(\mathbf{r}) = \sum_{\mathbf{g}} A_j(\mathbf{g})\langle\mathbf{r}|\mathbf{g}\rangle \tag{2.59}$$

in which g is a reciprocal vector of the lattice of unit cells containing a single heterostructure or impurity potential. Substituting (2.59) into (2.58) and multiplying on the left by $\langle\mathbf{g}'|$ gives a square matrix of order $4 \times n_{\mathbf{g}}$, where $n_{\mathbf{g}}$ is the number of plane waves appearing in the expansion (2.59). The matrix elements representing an arbitrary perturbation in the plane wave basis are determined by approximating $\mathrm{U}(\mathbf{r})$ as a piecewise-continuous function of position, i.e.,

$$\begin{aligned}\langle\mathbf{g}'|U|\mathbf{g}\rangle &\equiv \frac{1}{\Omega}\int_{\Omega} U(\mathbf{r})e^{i\Delta\mathbf{g}\cdot\mathbf{r}}d\mathbf{r} \\ &\simeq \frac{1}{\Omega}\sum_{\mathbf{R}_c}^{\mathcal{N}_c} U(\mathbf{R}_c)\int_{\Omega_c} e^{i\Delta\mathbf{g}\cdot\mathbf{R}_c}d\mathbf{r} \\ &= \frac{\Omega_c}{\Omega}\mathrm{sinc}\left(\frac{\Delta g_x l}{2}\right)\mathrm{sinc}\left(\frac{\Delta g_y l}{2}\right)\mathrm{sinc}\left(\frac{\Delta g_z l}{2}\right)\sum_{\mathbf{R}_c}^{\mathcal{N}_c} U(\mathbf{R}_c)e^{i\Delta\mathbf{g}\cdot\mathbf{R}_c}\end{aligned} \tag{2.60}$$

where Ω_c is the volume of cube of side l in the unit cell over which $U(\mathbf{r})$ can be considered constant. In (2.60) $\Delta\mathbf{g} = \mathbf{g} - \mathbf{g}'$, $\mathbf{R}_c$ is the position vector of cube c, $\mathcal{N}_c$ is the number of cubes in the heterostructure unit cell and Ω is the volume of the heterostructure unit cell. The attraction of this approach is that there is no need to explicitly match wave functions across a boundary between the barrier and well materials; the method is thus easily applicable to an arbitrary confining potential. The boundary conditions are that the states in neighbouring heterostructures do not significantly overlap.

Consider once more the GaAs/AlAs quantum well structure described in Section 2.1.2. Within the effective mass approximation the microscopic character of the perturbation potential is neglected. Thus, for a quantum well perturbation the valence band offset represents the sole remnant of the atomistic potential. Accordingly, we set $U(\mathbf{r})$ in (2.58) equal to the valence band offset (500 meV) in the well layer and zero in the barrier layer and place the heavy hole and light hole bulk states of GaAs at Γ in class A, in the Löwdin sense.

For the case of confinement in one dimension, the only plane waves with nonzero wave vectors to enter the series (2.59) will be those with wave vectors equal to reciprocal lattice vectors parallel to the [001] direction. The number of plane waves is chosen such that convergence of the expansion (2.59) is achieved. The quantum well Hamiltonian matrix is constructed in this basis and diagonalised to give eigenvalues for the valence hole subbands and the expansion coefficients A_j.

A number of simplifying assumptions have been made during the construction of the quantum well Hamiltonian matrix. First, the periodic parts of the Bloch functions of the heavy and light hole states at Γ do not differ very much from the GaAs to the AlAs layers. Second, we assume that the $\{\gamma_i\}$ adopt values appropriate to bulk GaAs and that these values do not change with position in the unit cell. If the $\{\gamma_i\}$ were to depend on position then the placement of the partial derivatives in (2.58) with respect to the matrix elements $D_{jj'}$ would require careful consideration in order to preserve the Hermicity of the effective Hamiltonian and the conservation of current across the interface. The scheme by which the abrupt change of the $\{\gamma_i\}$ across a heterostructure interface can be taken into account when using a plane wave basis has been discussed in detail by Baraff and Gershoni (1991). Throughout this work we assume that the $\{\gamma_i\}$ of the well material are appropriate throughout the heterostructure, justified by the dimensions of the structures involved and our interest in only the lowest energy eigenstates of the confining potential.

2.2.4 Transition Probabilities in the $\mathbf{k}\cdot\mathbf{p}$/EMA Scheme

Within the $\mathbf{k}\cdot\mathbf{p}$/EMA approximation, the wave function for the mth valence band can be written as

$$\psi_m^v(\mathbf{r}) = \sum_{j=1}^{4} F_j^m(\mathbf{r}) u_{j,0}^v(\mathbf{r}) \tag{2.61}$$

and for the nth conduction state as

$$\psi_n^c(\mathbf{r}) = F^n(\mathbf{r}) u_0^c(\mathbf{r}). \tag{2.62}$$

For the momentum matrix element coupling the two states we have (Bastard, 1988)

$$\begin{aligned} \hat{\mathbf{e}}\cdot\mathbf{p}_{nm} &= \sum_{j=1}^{4} \hat{\mathbf{e}}\cdot\left[\int_\Omega F^{n*}F_j^m u_0^{c*}\mathbf{p}u_{j,0}^v d\mathbf{r} + \int_\Omega u_0^{c*}u_{j,0}^v F^{n*}\mathbf{p}F_j^m d\mathbf{r}\right] \\ &\simeq \sum_{j=1}^{4} \hat{\mathbf{e}}\cdot\left[\langle F^n|F_j^m\rangle\langle u_0^c|\mathbf{p}|u_{j,0}^v\rangle + \langle F^n|\mathbf{p}|F_j^m\rangle\langle u_0^c|u_{j,0}^v\rangle\right]. \end{aligned} \tag{2.63}$$

In the second line of (2.63) use has been made of the slowly varying nature of the envelope function F over dimensions comparable to a bulk unit cell. Because the bulk Bloch functions for the conduction and valence band edges are orthogonal the second term in (2.63) vanishes and we get

$$\hat{\mathbf{e}}\cdot\mathbf{p}_{nm} \simeq \sum_{j=1}^{4} \hat{\mathbf{e}}\cdot\langle F^n|F_j^m\rangle\langle u_0^c|\mathbf{p}|u_{j,0}^v\rangle. \tag{2.64}$$

Since the conduction band states are S-type, the only nonzero momentum matrix elements between bulk band edge states are

$$\langle S|p^x|X\rangle = \langle S|p^y|Y\rangle = \langle S|p^z|Z\rangle = P_{cv}. \tag{2.65}$$

The across-gap matrix element P_{cv} is related to E_p, the characteristic energy from Kane's theory, by

$$E_p = \frac{2m_0}{\hbar^2}|P_{cv}|^2. \tag{2.66}$$

The Kane energy for GaAs is given in the final column of Table 2.3.

m_e^*	$m_{hh}^{[001]}$	$m_{lh}^{[001]}$	$m_{hh}^{[111]}$	$m_{lh}^{[111]}$	γ_1	γ_2	γ_3	E_p (eV)
0.0804	0.498	0.107	1.176	0.096	5.677	1.834	2.392	25.7

Table 2.3: Effective masses and Luttinger-Kohn parameters for GaAs derived from an empirical pseudopotential calculation. The characteristic energy E_p from Kane's theory is taken from Landolt-Börnstein (1982).

2.3 Electronic Structure of a GaAs/AlAs Quantum Well System

The band structure of a quantum well system consisting of 9 lattice constants (50.85 Å) of GaAs and 21 lattice constants (118.65 Å) of AlAs has been determined using the empirical pseudopotential method and the $\mathbf{k}\cdot\mathbf{p}$/EMA method. The lattice constants of GaAs and AlAs were taken to be $a_0 = 5.65$Å. In order to facilitate a direct comparison between the $\mathbf{k}\cdot\mathbf{p}$/EMA and empirical pseudopotential schemes, the band structure parameters $\{\gamma_i\}$ were derived from the effective masses predicted by an empirical pseudopotential calculation for GaAs. Values for the effective masses were obtained by sampling the energies of the relevant states close to Γ and assuming parabolic bands. The Luttinger parameters, together with relevant effective masses, obtained using (2.41), are listed in Table 2.3.

In Figure 2.6 the dispersion of the highest four spin degenerate valence bands are plotted along the [110] direction in the quantum well Brillouin zone for the 9GaAs/21AlAs system. The solid lines denote the empirical pseudopotential result and the dashed lines are the result of a four-band $\mathbf{k}\cdot\mathbf{p}$/EMA calculation. The subband labels HHn and LHn refer to heavy hole and light hole respectively and reflect the origin of the dominant component of the quantum well wave function. A good agreement between the two methods has been

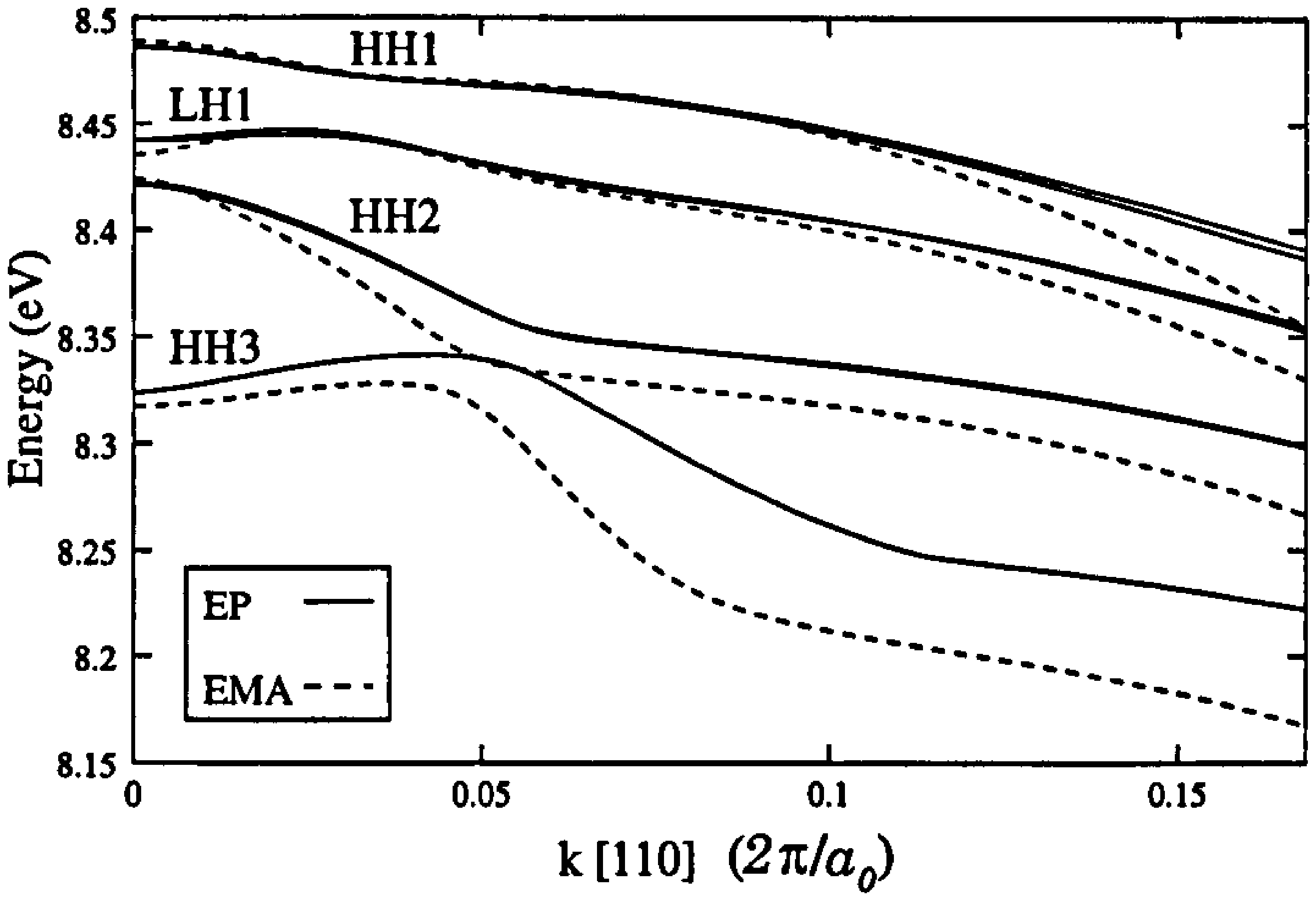

Figure 2.6: The uppermost four valence bands in a 9GaAs/21AlAs quantum well structure plotted along the [110] k-space direction.

achieved, especially for the dispersion in the neighbourhood of Γ. In the [110] direction symmetry permits the lifting of the spin-up vs spin-down degeneracy, resulting in a doubling of the number of EP bands seen in Figure 2.6. The splitting occurs as a result of the difference between the anion and cation atomic potentials. Since the $\mathbf{k}\cdot\mathbf{p}$/EMA calculation does not take into account the individual atomic potentials, the $\mathbf{k}\cdot\mathbf{p}$/EMA bands do not display this lifting of degeneracy. More sophisticated $\mathbf{k}\cdot\mathbf{p}$/EMA treatments account for the difference between the cation and anion potentials by introducing extra terms into the Luttinger-Kohn Hamiltonian (2.42). In these models, the size of the splitting is controlled by the Kane parameter B'. From the EP results, this splitting can be seen to be small. This justifies our Kane parameter choice $B' \equiv 0$.

At Γ the energies predicted by the EP and $\mathbf{k}\cdot\mathbf{p}$/EMA techniques should be the same. The differences occur because of the difficulty in determining accurate zone centre effective masses from the empirical pseudopotential cal-

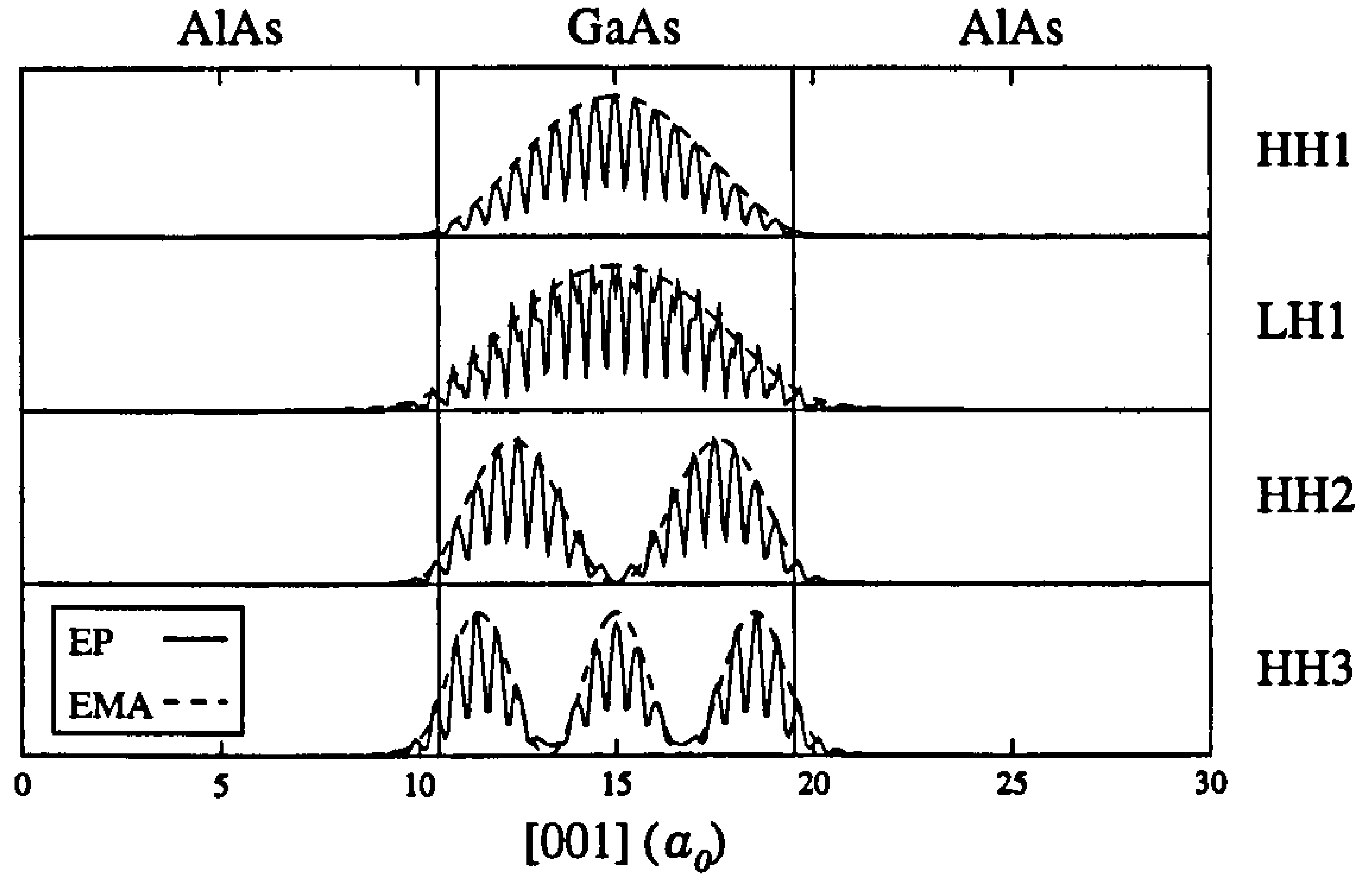

Figure 2.7: Charge density plots for the top four valence bands in a 9GaAs/21AlAs quantum well structure along the [001] real space axis.

culation. Deviations away from Γ are more pronounced in the lower lying subbands. Better agreement between these bands and those calculated by the empirical pseudopotential method would be achieved by including the split-off and the lowest conduction bulk bands in the $\mathbf{k}\cdot\mathbf{p}$ calculation.

The charge densities along the [001] axis in real space are shown in Figure 2.7 for the valence subband wave functions at Γ. For each band, the charge densities for the spin degenerate pair have been summed. The empirical pseudopotential results are denoted by the solid lines and the $\mathbf{k}\cdot\mathbf{p}$/EMA calculation by the dashed line. For the $\mathbf{k}\cdot\mathbf{p}$/EMA calculation the envelope function $|F(\mathbf{r})|^2$ has been plotted. For each state, the envelope of the charge density predicted by the two methods is very similar. Due to the extra complexity of the crystalline potential used in the empirical pseudopotential method, the component of the wave function rapidly varying on an atomic scale is implicitly included. The lack of significant penetration of the wave function into the barrier layer justifies the sole use of GaAs effective masses in the $\mathbf{k}\cdot\mathbf{p}$/EMA calculation.

The empirical pseudopotential scheme clearly provides the more accurate description of the energy bands and wave functions of the III-V semiconductor compounds, however, the choice of band structure method depends on the size of the computational problem at hand. In the empirical pseudopotential approach, the conventional computational effort scales as the cube of the number of atoms. In $\mathbf{k} \cdot \mathbf{p}$/EMA calculations, the computational effort is essentially independent of the number of atoms in the system. In Chapter 4, where quantum well structures containing $\sim$ 100 atoms are the focus of attention, an empirical pseudopotential approach is practical. In Chapter 5, where attention is switched to quantum dot structures containing $\sim 10^6$ atoms, the $\mathbf{k} \cdot \mathbf{p}$/EMA method is appropriate.

Chapter 3

Theory of Optical Nonlinearities in Semiconductors

In this chapter the general scheme for describing the nonlinear optical response of a semiconductor to an electromagnetic field is established. A description of the response is given in terms of the polarisation or polarisation-current density induced in the medium by the electric field associated with the incident light. The response is related to the incident field by the susceptibility or conductivity tensors. The tensors provide a direct link between the band structure and nonlinear optical phenomena. We introduce the quantum mechanical density matrix formalism and use it to derive an expression for the susceptibility tensor governing the extent to which a semiconductor radiates second harmonics.

3.1 Polarisation in the Time Domain

When the amplitude of the optical field applied to the material system is small in comparison with the cohesive field between electrons and ions ($\sim 10^9$

V/cm) then the resulting polarisation can be approximated by a term that is proportional to the applied field. If the wavelength of the incident light is much larger than the interatomic spacings then the spatial variations of the induced and applied fields may be neglected. Within this "long wavelength" or dipole approximation, the relation between the applied field, $\mathbf{E}(t)$, and the induced polarisation, $\mathbf{P}(t)$, is given by

$$\mathbf{P}(t) = \epsilon_0 \int_{-\infty}^{\infty} d\tau \mathbf{T}(t;\tau) \cdot \mathbf{E}(\tau). \tag{3.1}$$

Here $\mathbf{T}(t;\tau)$ is a second-rank tensor called the linear time domain response function. Because of the principles of time invariance and of causality, the tensor depends only on the difference $t-\tau$ and vanishes if this difference is less than zero. In addition, $\mathbf{T}(t;\tau)$ is real since $\mathbf{E}(t)$ and $\mathbf{P}(t)$ are real. The relationship between $\mathbf{P}(t)$ and $\mathbf{E}(t)$ can also be written as (Butcher and Cotter, 1990)

$$\mathbf{P}(t) = \epsilon_0 \int_{-\infty}^{\infty} d\tau \mathbf{T}(t-\tau) \cdot \mathbf{E}(\tau). \tag{3.2}$$

Making the change of variable $\tau \rightarrow t-\tau'$, (3.2) becomes

$$\mathbf{P}(t) = \epsilon_0 \int_{-\infty}^{\infty} d\tau' \mathbf{T}(\tau') \cdot \mathbf{E}(t-\tau'). \tag{3.3}$$

When the applied field is not negligible in comparison to the atomic field which holds the electric charges in equilibrium, the motion of the charges exhibits anharmonicity. Under such circumstances, a purely linear relationship between the polarisation and the applied field is no longer sufficient. It is convenient to expand these quantities as a power series in the fields (Bloembergen, 1965). Equation (3.1) must be replaced by

$$\mathbf{P}(t) = \mathbf{P}^{(1)}(t) + \mathbf{P}^{(2)}(t) + \mathbf{P}^{(3)}(t) + \cdots \tag{3.4}$$

where the superscripts in parentheses indicate the order in the fields. As for the linear case, the form of $\mathbf{P}^{(n)}(t)$ can be written on the basis of time invariance, causality and reality principles

$$\mathbf{P}^{(n)}(t) = \epsilon_0 \int_{-\infty}^{\infty} d\tau_1 \cdots \int_{-\infty}^{\infty} d\tau_n \mathbf{T}^{(n)}(\tau_1, \ldots, \tau_n) | \mathbf{E}(t-\tau_1) \cdots \mathbf{E}(t-\tau_n) \quad (3.5)$$

where $\mathbf{T}^{(n)}$ is a $(n+1)$ rank tensor called the nth-order time domain response function. The vertical bar replaces the column of n dots conventionally required to indicate a contracted tensor product. In suffix notation (3.5) reads

$$P_{\mu}^{(n)}(t) = \epsilon_0 \int_{-\infty}^{\infty} d\tau_1 \cdots \int_{-\infty}^{\infty} d\tau_n T_{\mu\alpha_1 \ldots \alpha_n}^{(n)}(\tau_1, \ldots, \tau_n) E_{\alpha_1}(t-\tau_1) \cdots E_{\alpha_n}(t-\tau_n) \quad (3.6)$$

where the Greek subscripts take the values x, y and z which label the cartesian coordinate axes and we invoke the 'repeated-index summation convention'; that is to say, repeated subscripts on the right hand side of (3.6) are understood to be summed over x, y and z.

Although $\mathbf{P}^{(n)}(t)$ is uniquely defined by (3.5), the response function $\mathbf{T}^{(n)}$ is not unique, because of the $n!$ ways to specify the order of the fields; there exists an intrinsic permutation symmetry. This arbitrariness in the definition of $\mathbf{T}^{(n)}$ is removed by the use of a symmetrisation procedure (Flytzanis, 1975)

$$T_{\mu\alpha_1 \ldots \alpha_n}^{(n)}(\tau_1, \ldots, \tau_n) = \frac{1}{n!} \sum_{p} T_{\mu\alpha_1 \ldots \alpha_n}^{(n)}(\tau_1, \ldots, \tau_n) \quad (3.7)$$

where the symbol $\sum_p$ refers to a sum over all permutations of the n pairs $(\alpha_1, \tau_1), \ldots, (\alpha_n, \tau_n)$.

3.2 Polarisation in the Frequency Domain

An alternative description of the optical properties of the medium are provided by frequency domain response functions, known as the susceptibility tensors. The susceptibility tensors arise when the electric field $\mathbf{E}(t)$ is expressed in terms of its Fourier transform $\mathbf{E}(\omega)$ by means of the Fourier integral identity

$$\mathbf{E}(t) = \int_{-\infty}^{\infty} d\omega \mathbf{E}(\omega) e^{-i\omega t} \quad (3.8)$$

where

$$\mathbf{E}(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} d\tau \mathbf{E}(\tau) e^{i\omega\tau}. \tag{3.9}$$

Substituting (3.8) into (3.5) we obtain

$$\mathbf{P}^{(n)}(t) = \epsilon_0 \int_{-\infty}^{\infty} d\omega_1 \cdots \int_{-\infty}^{\infty} d\omega_n \chi^{(n)}(-\omega_\sigma; \omega_1, \ldots, \omega_n) | \mathbf{E}(\omega_1) \cdots \mathbf{E}(\omega_n) e^{-i\omega_\sigma t} \tag{3.10}$$

where

$$\chi^{(n)}(-\omega_\sigma; \omega_1, \ldots, \omega_n) = \int_{-\infty}^{\infty} d\tau_1 \cdots \int_{-\infty}^{\infty} d\tau_n \mathbf{T}^{(n)}(\tau_1, \ldots, \tau_n) \exp\left[i \sum_{j=1}^{n} \omega_j t_j\right] \tag{3.11}$$

and $\omega_\sigma = \omega_1 + \omega_2 + \cdots + \omega_n$. Equation (3.11) defines the nth-order susceptibility tensor. The sign attached to a frequency is negative if the photon is emitted and positive if it is absorbed. Intrinsic permutation symmetry implies that $\chi^{(n)}(-\omega_\sigma; \omega_1, \ldots, \omega_n)$ is invariant under all $n!$ permutations of the n pairs $(\alpha_1, \omega_1), \ldots, (\alpha_n, \omega_n)$.

We are concerned with fields emitted by lasers, which are narrow superpositions of monochromatic waves. Because the Fourier transforms of such waves involve delta-functions, the polarisation is completely determined by the values of the susceptibility tensors at the discrete frequencies involved. Thus, by expanding $\mathbf{P}(t)$ in the frequency domain,

$$\mathbf{P}^{(n)}(t) = \int_{-\infty}^{\infty} d\omega \mathbf{P}^{(n)}(\omega) e^{-i\omega t} \tag{3.12}$$

where

$$\mathbf{P}^{(n)}(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} dt \mathbf{P}^{(n)}(t) e^{i\omega t} \tag{3.13}$$

we obtain, from (3.10),

$$\mathbf{P}^{(n)}(\omega) = \epsilon_0 \int_{-\infty}^{\infty} d\omega_1 \cdots \int_{-\infty}^{\infty} d\omega_n \chi^{(n)}(-\omega_\sigma; \omega_1, \ldots, \omega_n) | \times \mathbf{E}(\omega_1) \cdots \mathbf{E}(\omega_n) \delta(\omega - \omega_\sigma). \tag{3.14}$$

In arriving at (3.14), use has been made of the identity

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} dt e^{it(\omega-\omega_\sigma)} = \delta(\omega-\omega_\sigma). \tag{3.15}$$

Similar to the expansion of $\mathbf{P}(\mathrm{t})$ as a power series in the fields (3.4), we have expanded the Fourier component of the polarisation at the frequency ω_σ as a power series, so that

$$\mathbf{P}(\omega) = \mathbf{P}^{(1)}(\omega) + \mathbf{P}^{(2)}(\omega) + \mathbf{P}^{(3)}(\omega) + \cdots. \tag{3.16}$$

Notice that in (3.14), for any given ω, there may be several sets of frequencies $\omega_1, \ldots, \omega_n$ (each of which may be positive, negative or zero) which satisfy $\omega - \omega_\sigma = 0$.

3.2.1 Second Harmonic Susceptibility

We now consider the evaluation of the integrals in (3.14) for second-order non-linear processes, in the case when the applied field consists of a superposition of monochromatic waves. Specifically, we will consider the phenomenon of second harmonic generation, for which the polarisation field oscillates at twice the driving field frequency. The applied optical field with frequency $\Omega(\geq 0)$ may be written as

$$\mathbf{E}(t) = \frac{1}{2}\left[\mathcal{E}(\Omega)e^{-i\Omega t} + \mathcal{E}(-\Omega)e^{i\Omega t}\right]. \tag{3.17}$$

It follows immediately from (3.8) and (3.17) that the Fourier transform $\mathbf{E}(\omega)$ of $\mathbf{E}(t)$ is given by

$$\mathbf{E}(\omega) = \frac{1}{2}\left[\mathcal{E}(\Omega)\delta(\omega-\Omega) + \mathcal{E}(-\Omega)\delta(\omega+\Omega)\right]. \tag{3.18}$$

When $\mathbf{E}(t)$ is given by (3.17), we may rewrite (3.10) for the second-order polarisation $\mathbf{P}^{(2)}(t)$ in the form

$$\mathbf{P}^{(2)}(t) = \frac{1}{2}\left[\mathcal{P}^{(2)}(2\Omega)e^{-i2\Omega t} + \mathcal{P}^{(2)}(-2\Omega)e^{i2\Omega t} + 2\mathcal{P}^{(2)}(0)\right]. \tag{3.19}$$

The Fourier transform $\mathbf{P}^{(2)}(\omega)$ of $\mathbf{P}^{(2)}(t)$ is given by

$$\mathbf{P}^{(2)}(\omega) = \frac{1}{2}\left[\mathcal{P}^{(2)}(2\Omega)\delta(\omega-\Omega) + \mathcal{P}^{(2)}(-2\Omega)\delta(\omega+\Omega) + 2\mathcal{P}^{(2)}(0)\delta(\omega)\right]. \tag{3.20}$$

If we consider a specific frequency $\omega = 2\Omega$, then by substituting (3.18) and (3.20) into (3.14) we obtain an expression for $\mathcal{P}^{(2)}(2\Omega)$:

$$\mathcal{P}^{(2)}(2\Omega) = \frac{1}{2}\epsilon_0\chi^{(2)}(-2\Omega;\Omega,\Omega) : \mathcal{E}(\Omega)\mathcal{E}(\Omega). \tag{3.21}$$

The cartesian μ component is given by

$$\mathcal{P}^{(2)}_{\mu}(2\Omega) = \frac{1}{2}\epsilon_0\chi^{(2)}_{\mu\alpha_1\alpha_2}(-2\Omega;\Omega,\Omega)\mathcal{E}_{\alpha_1}(\Omega)\mathcal{E}_{\alpha_2}(\Omega). \tag{3.22}$$

The spectra generated from the frequency dependence of the tensor components of $\chi^{(2)}(-2\Omega;\Omega,\Omega)$ fully characterises the second harmonic response of the material for the applied field considered. Clearly these components should depend on the microscopic properties of the medium.

3.2.2 Second Harmonic Conductivity

The optical response of the medium can be described in terms of the polarisation-current density $\mathbf{J}(t)$ rather than the polarisation $\mathbf{P}(t)$, where

$$\mathbf{J}(t) = \frac{\partial \mathbf{P}(t)}{\partial t}. \tag{3.23}$$

By differentiating (3.4) we obtain the power series

$$\mathbf{J}(t) = \mathbf{J}^{(1)}(t) + \mathbf{J}^{(2)}(t) + \cdots. \tag{3.24}$$

Following a procedure analogous to that detailed in the previous section leads to the relationship between the amplitude of the second harmonic polarisation-current density $\mathcal{J}^{(2)}(2\Omega)$ and the amplitude of the applied field

$$\mathcal{J}^{(2)}(2\Omega) = \sigma^{(2)}(-2\Omega;\Omega,\Omega) : \mathcal{E}(\Omega)\mathcal{E}(\Omega) \tag{3.25}$$

where $\sigma^{(2)}(-2\Omega;\Omega,\Omega)$ is the second-order conductivity tensor, related to the susceptibility by

$$\sigma^{(2)}(-2\Omega;\Omega,\Omega) = -i\Omega\epsilon_0\chi^{(2)}(-2\Omega;\Omega,\Omega). \tag{3.26}$$

When dealing with semiconductor systems it is often more convenient to describe the optical response of the medium in terms of the polarisation-current density rather than the polarisation. The polarisation approach produces expressions for the susceptibility tensors involving matrix elements of the position operator; these are incompatible with the description of the electronic structure of semiconductors in terms of an infinite array of identical unit cells. However, the polarisation-current density description yields expressions for the conductivity tensors containing momentum matrix elements which remain well defined when periodic boundary conditions are in operation. We choose to formulate an expression for the second harmonic conductivity and then convert to one for the second harmonic susceptibility using the relation (3.26).

3.3 A Microscopic Expression for the Second Harmonic Susceptibility

3.3.1 Density Matrix Formalism

The most convenient route to an expression for the second harmonic susceptibility is via the quantum mechanical density matrix formalism. We begin by considering a collection of ν independent electrons under the influence of an electromagnetic field. The Hamiltonian H of each electron in the ensemble consists of two parts:

$$H = H_0 + H'. \tag{3.27}$$

The first term is the unperturbed, one electron Hamiltonian with eigensolutions $\phi_n(\mathbf{r})$ and energies E_n. The second term is the interaction Hamiltonian. It is the correction to the unperturbed Hamiltonian arising from the interaction of light with the electron. The wave function $\psi^i(\mathbf{r},t)$ of the ith electron in the collection can be expanded in the complete set of functions $\{\phi_n\}$:

$$\psi^i(\mathbf{r},t) = \sum_n a_n^i(t)\phi_n(\mathbf{r}). \tag{3.28}$$

The ensemble average of a physical quantity **A** is given by

$$\begin{aligned}\langle \mathbf{A}\rangle &= \frac{1}{\nu}\sum_{i=1}^{\nu}\langle\psi^i|\mathbf{A}|\psi^i\rangle \\ &= \sum_{m,n}\overline{a_n^i(t)a_m^{i\,*}(t)}\mathbf{A}_{mn} \\ &= \sum_{m,n}\rho_{nm}\mathbf{A}_{mn} \\ &= \mathrm{Tr}(\rho\mathbf{A}).\end{aligned} \tag{3.29}$$

The Hermitian operator $\rho(t)$, with matrix elements

$$\rho_{nm} = \frac{1}{\nu}\sum_{i=1}^{\nu}a_n^i(t)a_m^{i\,*}(t) = \overline{a_n^i(t)a_m^{i\,*}(t)} \tag{3.30}$$

in the $\{\phi_n\}$ representation, is the one electron density matrix (Kohn and Luttinger, 1957). The density operator $\rho(t)$ varies with time according to

$$\frac{\partial\rho}{\partial t} = -\frac{i}{\hbar}[H,\rho]. \tag{3.31}$$

Premultiplying (3.31) by $\langle\phi_n|$ and postmultiplying by $|\phi_m\rangle$ gives the equation of motion of a density matrix element:

$$\frac{\partial\rho_{nm}}{\partial t} = -\frac{i}{\hbar}[H,\rho]_{nm}. \tag{3.32}$$

There are further interactions (such as those resulting from incoherent scattering processes) that cannot be conveniently included in the Hamiltonian H.

Such interactions result in a relaxation of the perturbed density matrix element back to thermal equilibrium. These effects are included in the density matrix formalism by adding phenomenological damping terms to the equation of motion. Equation (3.32) becomes (Boyd, 1992)

$$\frac{\partial \rho_{nm}}{\partial t} = -\frac{i}{\hbar}[H, \rho]_{nm} - \Gamma_{nm}(\rho_{nm} - \rho^{eq}_{nm}). \tag{3.33}$$

The damping term indicates that ρ_{nm} relaxes to its equilibrium value ρ^{eq}_{nm} at a rate Γ_{nm}. The diagonal matrix element ρ_{nn} represents the population of the system in the state $|\phi_n\rangle$, while the off-diagonal matrix element ρ_{nm} indicates that the state of the system has a coherent admixture of $|\phi_n\rangle$ and $|\phi_m\rangle$. At thermal equilibrium, the diagonal matrix element ρ^{eq}_{nn} represents the thermal population distribution, the Fermi-Dirac distribution, in the case of a sufficiently high electron density. The off-diagonal elements ρ^{eq}_{nm} are taken to be zero due to ensemble average, as thermal excitation is expected to be an incoherent process and cannot produce any coherent superpositions of states.

3.3.2 The Current Density Operator

In order to obtain an expression for the second harmonic polarisation-current density via the density matrix method, we are required to introduce the current density operator, $\boldsymbol{\pi}$. The one electron current density operator in a volume V is related to the momentum operator via

$$\boldsymbol{\pi} = -\frac{e}{m}\mathbf{p}. \tag{3.34}$$

In the presence of an applied field with a vector potential $\mathbf{A}(t)$ and scalar field $\Phi(t)$ the current density operator becomes

$$\boldsymbol{\pi}' = \boldsymbol{\pi} - \frac{e^2}{2mc}\mathbf{A}(t). \tag{3.35}$$

We choose to work in the velocity gauge, in which $\nabla \cdot \mathbf{A} = 0$ and $\Phi(t) = 0$. Since we are working within the dipole approximation, we have neglected the $\mathbf{r}$ dependence of the vector potential. From (3.29), the polarisation-current density is given by

$$\mathbf{J}(t) = \frac{1}{V}\mathrm{Tr}[\rho(t)\pi']. \tag{3.36}$$

When the amplitude of the applied field is small then it is possible to expand the density operator in powers of the field:

$$\rho(t) = \rho^{(0)} + \rho^{(1)}(t) + \rho^{(2)}(t) + \cdots. \tag{3.37}$$

The substitution of (3.24) and (3.37) into (3.36) allows us to write the second-order polarisation-current density in terms of the second-order density operator

$$\mathbf{J}^{(2)}(t) = \frac{1}{V}\mathrm{Tr}[\rho^{(2)}(t)\pi] \tag{3.38}$$

where use has been made of the fact that the second term in the perturbed polarisation-current density operator (3.35) only contributes to first order in $\mathbf{J}(t)$. For the optical field given by (3.17) we have

$$\mathbf{J}^{(2)}(t) = \frac{1}{2}\left[\mathcal{J}^{(2)}(2\Omega)e^{-i2\Omega t} + \mathcal{J}^{(2)}(-2\Omega)e^{i2\Omega t} + 2\mathcal{J}^{(2)}(0)\right] \tag{3.39}$$

and

$$\rho^{(2)}(t) = \frac{1}{2}\left[\rho^{(2)}(2\Omega)e^{-i2\Omega t} + \rho^{(2)}(-2\Omega)e^{i2\Omega t} + 2\rho^{(2)}(0)\right]. \tag{3.40}$$

Thus the component of the current density oscillating at the frequency 2Ω is given by (3.38), (3.39) and (3.40) to be

$$\mathcal{J}^{(2)}(2\Omega) = \frac{1}{V}\sum_{n,m}[\rho^{(2)}_{nm}(2\Omega)\pi_{mn}]. \tag{3.41}$$

3.3.3 A Perturbation Solution for the Density Operator

Substituting the power series (3.37) into (3.33), expanding the commutator and equating terms of the same order in the field gives the set of equations

$$\frac{\partial \rho_{nm}^{(0)}}{\partial t} = -i\omega_{nm}\rho_{nm}^{(0)} - \Gamma_{nm}(\rho_{nm}^{(0)} - \rho_{nm}^{eq}) \tag{3.42}$$

$$\frac{\partial \rho_{nm}^{(1)}}{\partial t} = -(i\omega_{nm} + \Gamma_{nm})\rho_{nm}^{(1)} - \frac{i}{\hbar}\left[H', \rho^{(0)}\right]_{nm} \tag{3.43}$$

$$\frac{\partial \rho_{nm}^{(2)}}{\partial t} = -(i\omega_{nm} + \Gamma_{nm})\rho_{nm}^{(2)} - \frac{i}{\hbar}\left[H', \rho^{(1)}\right]_{nm} \tag{3.44}$$

etc., where $\omega_{nm} = (E_n - E_m)/\hbar$. For the zeroth order term, $\rho_{nm}^{(0)}$, we simply have the thermal equilibrium condition

$$\rho_{nm}^{(0)} = \rho_{nn}^{eq}\delta_{nm}. \tag{3.45}$$

For an optical field given by (3.17) we have for the first order term

$$\rho_{nm}^{(1)}(t) = \frac{1}{2}\left[\rho_{nm}^{(1)}(\Omega)e^{-i\Omega t} + \rho_{nm}^{(1)}(-\Omega)e^{i\Omega t}\right] \tag{3.46}$$

and for the interaction Hamiltonian

$$H'(t) = \frac{1}{2}\left[\mathcal{H}'(\Omega)e^{-i\Omega t} + \mathcal{H}'(-\Omega)e^{i\Omega t}\right]. \tag{3.47}$$

Substituting (3.46) and (3.47) into (3.43) and matching the $e^{-i\Omega t}$ dependence, one obtains

$$\rho_{nm}^{(1)}(\Omega) = \frac{\mathcal{H}'_{nm}(\Omega)}{\hbar(\Omega - \omega_{nm} + i\Gamma_{nm})}(\rho_{mm}^{(0)} - \rho_{nn}^{(0)}). \tag{3.48}$$

To second-order in the field we have

$$\rho_{nm}^{(2)}(t) = \frac{1}{2}\left[\rho_{nm}^{(2)}(2\Omega)e^{-i2\Omega t} + \rho_{nm}^{(2)}(-2\Omega)e^{i2\Omega t} + 2\rho_{nm}^{(2)}(0)\right]. \tag{3.49}$$

Expanding the commutator in (3.44) gives

$$\begin{aligned}\left[H', \rho^{(1)}\right]_{nm} &= \left[H'\rho^{(1)} - \rho^{(1)}H'\right]_{nm} \\ &= \sum_{v}(H'_{nv}\rho_{vm}^{(1)} - \rho_{nv}^{(1)}H'_{vm}).\end{aligned} \tag{3.50}$$

Substituting (3.47) and (3.48) into (3.50) gives

$$\left[H',\rho^{(1)}\right]_{nm} = \sum_v \left[\frac{\mathcal{H}'_{nv}(\Omega)\mathcal{H}'_{vm}(\Omega)}{4\hbar^2(\Omega-\omega_{vm}+i\Gamma_{vm})}(\rho^{(0)}_{mm}-\rho^{(0)}_{vv}) - \frac{\mathcal{H}'_{nv}(\Omega)\mathcal{H}'_{vm}(\Omega)}{4\hbar^2(\Omega-\omega_{nv}+i\Gamma_{nv})}(\rho^{(0)}_{vv}-\rho^{(0)}_{nn})\right]. \tag{3.51}$$

Using (3.44), (3.49) and (3.51) we obtain the expression for the second harmonic density matrix component

$$\rho^{(2)}_{nm}(2\Omega) = \sum_v \left[\frac{\mathcal{H}'_{nv}(\Omega)\mathcal{H}'_{vm}(\Omega)}{2\hbar^2(2\Omega-\omega_{nm}+i\Gamma_{nm})(\Omega-\omega_{vm}+i\Gamma_{vm})}(\rho^{(0)}_{mm}-\rho^{(0)}_{vv}) - \frac{\mathcal{H}'_{nv}(\Omega)\mathcal{H}'_{vm}(\Omega)}{2\hbar^2(2\Omega-\omega_{nm}+i\Gamma_{nm})(\Omega-\omega_{nv}+i\Gamma_{nv})}(\rho^{(0)}_{vv}-\rho^{(0)}_{nn})\right]. \tag{3.52}$$

Consider now the interaction Hamiltonian $H'(t)$. In the dipole approximation, the interaction Hamiltonian in the velocity gauge is

$$H'(t) = -\pi\cdot \mathrm{A}(t) \tag{3.53}$$

where the vector potential, $\mathrm{A}(t)$, is related to the electric field by

$$\begin{aligned}\mathrm{A}(t) &= -\int_{-\infty}^{t} d\tau \mathrm{E}(\tau)\\ &= \int_{-\infty}^{\infty} d\omega \mathrm{E}(\omega)(i\omega)^{-1}e^{-i\omega t}.\end{aligned} \tag{3.54}$$

Thus, for a field given by (3.17) we may write the amplitude of the interaction Hamiltonian oscillating at a frequency Ω as

$$\mathcal{H}'(\Omega) = -\frac{\pi\cdot\mathcal{E}(\Omega)}{i\Omega}. \tag{3.55}$$

Using (3.41), (3.52) and (3.55) we can write for the second harmonic current density amplitude

$$\mathcal{J}^{(2)}(2\Omega) = \frac{-1}{2\hbar^2\Omega^2 V}\sum_{n,m,v} \left[\frac{[\pi_{nv}\cdot\mathcal{E}(\Omega)][\pi_{vm}\cdot\mathcal{E}(\Omega)]\pi_{mn}}{(2\Omega-\omega_{nm}+i\Gamma_{nm})(\Omega-\omega_{vm}+i\Gamma_{vm})}(\rho^{(0)}_{mm}-\rho^{(0)}_{vv}) - \frac{[\pi_{nv}\cdot\mathcal{E}(\Omega)][\pi_{vm}\cdot\mathcal{E}(\Omega)]\pi_{mn}}{(2\Omega-\omega_{nm}+i\Gamma_{nm})(\Omega-\omega_{nv}+i\Gamma_{nv})}(\rho^{(0)}_{vv}-\rho^{(0)}_{nn})\right]. \tag{3.56}$$

Comparison with (3.25) and (3.26) gives an expression for the second harmonic susceptibility

$$\chi^{(2)}_{\mu\alpha\beta}(-2\Omega;\Omega,\Omega) = \frac{-i}{2\epsilon_0\hbar^2\Omega^3 V}\sum_{n,m,v}$$
$$\left[\frac{\pi^{\mu}_{mn}\pi^{\alpha}_{nv}\pi^{\beta}_{vm}}{(2\Omega-\omega_{nm}+i\Gamma_{nm})(\Omega-\omega_{vm}+i\Gamma_{vm})}(\rho^{(0)}_{mm}-\rho^{(0)}_{vv})\right.$$
$$\left.-\frac{\pi^{\mu}_{mn}\pi^{\alpha}_{nv}\pi^{\beta}_{vm}}{(2\Omega-\omega_{nm}+i\Gamma_{nm})(\Omega-\omega_{nv}+i\Gamma_{nv})}(\rho^{(0)}_{vv}-\rho^{(0)}_{nn})\right]. \tag{3.57}$$

Taking into account the intrinsic permutation symmetry described in Section 3.1, this expression becomes

$$\chi^{(2)}_{\mu\alpha\beta}(-2\Omega;\Omega,\Omega) = \frac{-i}{4\epsilon_0\hbar^2\Omega^3 V}\sum_{n,m,v}$$
$$\left[\frac{\pi^{\mu}_{mn}\pi^{\alpha}_{nv}\pi^{\beta}_{vm}}{(2\Omega-\omega_{nm}+i\Gamma_{nm})(\Omega-\omega_{vm}+i\Gamma_{vm})}(\rho^{(0)}_{mm}-\rho^{(0)}_{vv})\right.$$
$$+\frac{\pi^{\mu}_{mn}\pi^{\beta}_{nv}\pi^{\alpha}_{vm}}{(2\Omega-\omega_{nm}+i\Gamma_{nm})(\Omega-\omega_{vm}+i\Gamma_{vm})}(\rho^{(0)}_{mm}-\rho^{(0)}_{vv})$$
$$-\frac{\pi^{\mu}_{mn}\pi^{\alpha}_{nv}\pi^{\beta}_{vm}}{(2\Omega-\omega_{nm}+i\Gamma_{nm})(\Omega-\omega_{nv}+i\Gamma_{nv})}(\rho^{(0)}_{vv}-\rho^{(0)}_{nn})$$
$$\left.-\frac{\pi^{\mu}_{mn}\pi^{\beta}_{nv}\pi^{\alpha}_{vm}}{(2\Omega-\omega_{nm}+i\Gamma_{nm})(\Omega-\omega_{nv}+i\Gamma_{nv})}(\rho^{(0)}_{vv}-\rho^{(0)}_{nn})\right]. \tag{3.58}$$

As the summation in (3.58) is over all states, we can replace the indices v, n, m in the last two terms by m, v, n respectively, resulting in

$$\chi^{(2)}_{\mu\alpha\beta}(-2\Omega;\Omega,\Omega) = \frac{-i}{4\epsilon_0\hbar^2\Omega^3 V}\sum_{n,m,v}(\rho^{(0)}_{mm}-\rho^{(0)}_{vv})$$
$$\left[\frac{\pi^{\mu}_{mn}\pi^{\alpha}_{nv}\pi^{\beta}_{vm}}{(2\Omega-\omega_{nm}+i\Gamma_{nm})(\Omega-\omega_{vm}+i\Gamma_{vm})}\right.$$
$$+\frac{\pi^{\mu}_{mn}\pi^{\beta}_{nv}\pi^{\alpha}_{vm}}{(2\Omega-\omega_{nm}+i\Gamma_{nm})(\Omega-\omega_{vm}+i\Gamma_{vm})}$$
$$-\frac{\pi^{\mu}_{nv}\pi^{\alpha}_{vm}\pi^{\beta}_{mn}}{(2\Omega-\omega_{vn}+i\Gamma_{vn})(\Omega-\omega_{vm}+i\Gamma_{vm})}$$
$$\left.-\frac{\pi^{\mu}_{nv}\pi^{\beta}_{vm}\pi^{\alpha}_{mn}}{(2\Omega-\omega_{vn}+i\Gamma_{vn})(\Omega-\omega_{vm}+i\Gamma_{vm})}\right]. \tag{3.59}$$

We can recast (3.59) in a form which involves a summation over populations rather than a summation over population differences. This is accomplished by replacing the indices v, n, m in all terms multiplied by $\rho_{vv}^{(0)}$ with m, v, n:

$$\begin{aligned}
\chi^{(2)}_{\mu\alpha\beta}(-2\Omega;\Omega,\Omega) = \frac{-i}{4\epsilon_0\hbar^2\Omega^3 V}\sum_{n,m,v}\rho^{(0)}_{mm}
&\Bigg[\frac{\pi^{\mu}_{mn}\pi^{\alpha}_{nv}\pi^{\beta}_{vm}}{(2\Omega-\omega_{nm}+i\Gamma_{nm})(\Omega-\omega_{vm}+i\Gamma_{vm})}\\
&-\frac{\pi^{\mu}_{nv}\pi^{\alpha}_{vm}\pi^{\beta}_{mn}}{(2\Omega-\omega_{vn}+i\Gamma_{vn})(\Omega-\omega_{mn}+i\Gamma_{mn})}\\
&+\frac{\pi^{\mu}_{mn}\pi^{\beta}_{nv}\pi^{\alpha}_{vm}}{(2\Omega-\omega_{nm}+i\Gamma_{nm})(\Omega-\omega_{vm}+i\Gamma_{vm})}\\
&-\frac{\pi^{\mu}_{nv}\pi^{\beta}_{vm}\pi^{\alpha}_{mn}}{(2\Omega-\omega_{vn}+i\Gamma_{vn})(\Omega-\omega_{mn}+i\Gamma_{mn})}\\
&-\frac{\pi^{\mu}_{nv}\pi^{\alpha}_{vm}\pi^{\beta}_{mn}}{(2\Omega-\omega_{vn}+i\Gamma_{vn})(\Omega-\omega_{vm}+i\Gamma_{vm})}\\
&+\frac{\pi^{\mu}_{vm}\pi^{\alpha}_{mn}\pi^{\beta}_{nv}}{(2\Omega-\omega_{mv}+i\Gamma_{mv})(\Omega-\omega_{mn}+i\Gamma_{mn})}\\
&-\frac{\pi^{\mu}_{nv}\pi^{\beta}_{vm}\pi^{\alpha}_{mn}}{(2\Omega-\omega_{vn}+i\Gamma_{vn})(\Omega-\omega_{vm}+i\Gamma_{vm})}\\
&+\frac{\pi^{\mu}_{vm}\pi^{\beta}_{mn}\pi^{\alpha}_{nv}}{(2\Omega-\omega_{mv}+i\Gamma_{mv})(\Omega-\omega_{mn}+i\Gamma_{mn})}\Bigg].
\end{aligned} \tag{3.60}$$

Finally, replacing indices m, n, v with a, b, c and collecting together like terms gives

$$\begin{aligned}
\chi^{(2)}_{\mu\alpha\beta}(-2\Omega;\Omega,\Omega) = \frac{-i}{4\epsilon_0\hbar^2\Omega^3 V}\sum_{a,b,c}\rho^{(0)}_{aa}
&\Bigg[\frac{\pi^{\mu}_{ab}\pi^{\alpha}_{bc}\pi^{\beta}_{ca}+\pi^{\mu}_{ab}\pi^{\beta}_{bc}\pi^{\alpha}_{ca}}{(\omega_{ba}-2\Omega-i\Gamma_{ba})(\omega_{ca}-\Omega-i\Gamma_{ca})}\\
&+\frac{\pi^{\beta}_{ab}\pi^{\mu}_{bc}\pi^{\alpha}_{ca}+\pi^{\alpha}_{ab}\pi^{\mu}_{bc}\pi^{\beta}_{ca}}{(\omega_{bc}+2\Omega+i\Gamma_{bc})(\omega_{ca}-\Omega-i\Gamma_{ca})}\\
&+\frac{\pi^{\beta}_{ab}\pi^{\mu}_{bc}\pi^{\alpha}_{ca}+\pi^{\alpha}_{ab}\pi^{\mu}_{bc}\pi^{\beta}_{ca}}{(\omega_{cb}-2\Omega-i\Gamma_{cb})(\omega_{ba}+\Omega+i\Gamma_{ba})}\\
&+\frac{\pi^{\alpha}_{ab}\pi^{\beta}_{bc}\pi^{\mu}_{ca}+\pi^{\beta}_{ab}\pi^{\alpha}_{bc}\pi^{\mu}_{ca}}{(\omega_{ca}+2\Omega+i\Gamma_{ca})(\omega_{ba}+\Omega+i\Gamma_{ba})}\Bigg]
\end{aligned} \tag{3.61}$$

where we have used the fact that $\omega_{ij} = -\omega_{ji}$ and $\Gamma_{ij} = \Gamma_{ji}$. For semiconductors in the one electron approximation, the population of the single particle state labelled a is given by the Fermi-Dirac function

$$\rho_{aa}^{(0)} = \frac{1}{e^{(E_a(\mathbf{k})-E_F)/k_BT} + 1} \tag{3.62}$$

and the sum over states labelled a becomes a sum over all bands and all wave vectors labelled $\mathbf{k}$ in the first Brillouin zone.

Within the dipole approximation, the total second harmonic susceptibility for a given applied field frequency is equal to the sum of the contributions from direct electron transitions between the states $a \rightarrow c \rightarrow b \rightarrow a$ occuring at each point in the Brillouin zone. The importance of each contribution depends on the product of the three momentum matrix elements representing the overall probability that the process $a \rightarrow c \rightarrow b \rightarrow a$ will take place [the numerators in (3.61)], and the proximity of the photon energy ($\hbar\Omega$) to energy differences between the participating states ($\hbar\omega_{ij}$) [the denominators in (3.61)].

It is immediately apparent that if all single particle states are empty then the total contribution to the second harmonic susceptibility is zero. Likewise, if all the states are occupied then we would expect, from the exclusion principle, the response to be zero. Examining (3.61) it is not obvious that this is the case. However, inspection of (3.59), which contains a sum over population differences, indicates that at least one state must be (un)occupied for a finite $\chi^{(2)}(-2\Omega;\Omega,\Omega)$. We would also expect the second-order susceptibility to be zero when the electron binding potential is symmetric. If the semiconductor in question possesses a centre of inversion, then reversing the direction of the applied field should cause the sign of the induced polarisation to change but its magnitude to remain the same. This is clearly only the case if the even power terms in the polarisation expansion (3.4) are zero, i.e., the even order susceptibilities are zero. In symmetric systems where the response is given in terms of the polarisation rather than the polarisation-current density the second-order response is zero because one of the position matrix elements between the three states involved is zero (Rosencher *et al.*, 1996). For a symmetric semiconductor system in which the response is described in terms of the current density, the second-order nonlinearities can be shown to vanish through the cancellation of contributions throughout the Brillouin zone.

Equation (3.61) was derived using a perturbational approach and so remains valid only when the interaction of light with matter is weak. We should expect the formula to reasonably describe the frequency dependence of the second harmonic susceptibility when $|H'_{nm}| << \hbar|\omega_{nm} - \omega|$, where ω is the frequency of the applied field and $\omega_{nm} = (E_n - E_m)/\hbar$ is the resonant transition frequency. This condition is well met in two regimes: high field but far from the resonant transition frequency and low field near to the resonant frequency. In both these non-resonant regimes the response is dominated by processes involving virtual transitions between states rather than processes involving energy dissipative real excitations. For a virtual transition to take place, the uncertainty principle requires that it does so in a very short space of time of the order of 10^{-12} seconds. In contrast, real transitions occur over a much longer timescale $\sim 10^{-9}$ seconds.

As the frequency of the driving field approaches zero the expression (3.61) diverges due to the $(\Omega)^{-3}$ prefactor. These divergences are not expected on physical grounds and are a result of using the velocity gauge in the derivation. The length gauge approach, in which the response is described in terms of the polarisation and the expressions for the susceptibilities involve matrix elements of the position operator, does not lead to unphysical divergences, but as discussed in Section 3.2.2, has long been regarded as unsuitable for crystal applications because the position operators are incompatible with the concept of an infinite semiconductor. However, Aversa and Sipe (1995) have recently provided a prescription for obtaining expressions for the nonlinear optical responses in semiconductors in the length gauge. They have demonstrated expressions for $\chi^{(2)}$ and $\chi^{(3)}$ that are free from the divergences that plague velocity gauge approach.

The barrier to obtaining divergence-free expressions has been the treatment of the position operator $\mathbf{r}$ in view of the extended Bloch states. The matrix elements of $\mathbf{r}$ can be resolved into two parts: an *intraband* part and an *interband* part. It is the highly singular and non-unique character of the intraband part which leads to the troublesome nature of the position operator in crystals. This problem has led many researchers to avoid the position operator in crystals by using the velocity gauge in theoretical studies of optical properties. Aversa and Sipe have developed a mathematical framework in which the obstacles to using the length gauge approach that are contained within the intraband part of the position operator can be negotiated.

In the following chapter, the expression for $\chi^{(2)}(-2\Omega;\Omega,\Omega)$ derived using the velocity gauge approach is used to investigate the second harmonic response in semiconductor quantum well systems. As already mentioned, the formula exhibits an unphysical divergence as the frequency of the applied field approaches zero. In the results that follow, these divergences will be accounted for and the conclusions reached concerning the role of material parameters such as quantum well depth and width in optimising such responses should remain valid.

Chapter 4

Second Harmonic Generation in p-type Quantum Wells

In this chapter the microscopic expression for the second-order susceptibility derived in the previous chapter and empirical pseudopotential band structure calculations described in Chapter 2 are used to determine the second harmonic response in p-type GaAs/AlAs asymmetric quantum well systems at a temperature of 0 K. The role of band structure tuning in the optimisation of such responses is investigated. Specifically, we examine the second harmonic response arising from virtual intersubband transitions at various valence band quantum well depths and widths. We also investigate the influence on the second harmonic signal of the degree of quantum well asymmetry.

4.1 $\chi^{(2)}$ Expression for p-type Quantum Wells

We are interested in optical nonlinearities arising from virtual transitions between valence subbands in GaAs/AlAs asymmetric quantum wells. The expression for the second harmonic response of a quantum well structure is given by (3.61)

$$\begin{aligned}\chi^{(2)}_{\mu\alpha\beta}(-2\omega;\omega,\omega) &= \frac{ie^3}{4\epsilon_0\hbar^2\omega^3 V m_0^3}\sum_{\mathbf{k}}\sum_{a,b,c} f_0(a,\mathbf{k}) \\ &\left[\frac{p^{\mu}_{ab}p^{\alpha}_{bc}p^{\beta}_{ca}+p^{\mu}_{ab}p^{\beta}_{bc}p^{\alpha}_{ca}}{(\omega_{ba}-2\omega-i\Gamma_{ba})(\omega_{ca}-\omega-i\Gamma_{ca})}\right. \\ &+\frac{p^{\beta}_{ab}p^{\mu}_{bc}p^{\alpha}_{ca}+p^{\alpha}_{ab}p^{\mu}_{bc}p^{\beta}_{ca}}{(\omega_{bc}+2\omega+i\Gamma_{bc})(\omega_{ca}-\omega-i\Gamma_{ca})} \\ &+\frac{p^{\beta}_{ab}p^{\mu}_{bc}p^{\alpha}_{ca}+p^{\alpha}_{ab}p^{\mu}_{bc}p^{\beta}_{ca}}{(\omega_{cb}-2\omega-i\Gamma_{cb})(\omega_{ba}+\omega+i\Gamma_{ba})} \\ &\left.+\frac{p^{\alpha}_{ab}p^{\beta}_{bc}p^{\mu}_{ca}+p^{\beta}_{ab}p^{\alpha}_{bc}p^{\mu}_{ca}}{(\omega_{ca}+2\omega+i\Gamma_{ca})(\omega_{ba}+\omega+i\Gamma_{ba})}\right] \end{aligned} \tag{4.1}$$

where p^{η}_{nm} is a matrix element of the momentum operator connecting valence subbands m and n and $f_0(a,\mathrm{k})$ is the population of band a with wave vector k. The square bracket in (4.1) now represents the contribution to the second-order susceptibility when an electron with wave vector **k** makes virtual transitions between the valence levels $a \rightarrow c \rightarrow b \rightarrow a$.

The summations are, in principle, over all bands and all wave vectors and so a detailed knowledge of the momentum matrix elements and transition frequencies throughout the Brillouin zone is required. This can be provided only by methods in electronic structure theory that offer a global description of the band structure. We therefore use the empirical pseudopotential approach described in Chapter 2 to provide momentum matrix elements and subband energies of GaAs/AlAs asymmetric quantum wells at any point in the Brillouin zone.

4.1.1 Populated States

In order to generate a finite second harmonic signal at least one of the subbands participating in the process $a \rightarrow c \rightarrow b \rightarrow a$ must be vacant. At a temperature of 0 K, the valence band is fully occupied and so we would expect intersubband transitions to be forbidden by the Pauli exclusion principle. The uppermost valence subband in unstrained GaAs/AlAs quantum well systems is the HH1 band (see Section 2.3). States in this band can be made vacant by the introduction of acceptor impurities such as beryllium atoms into the barrier layer (AlAs). Doping creates acceptor levels near to the edge of the valence band in AlAs. Electrons will choose to occupy these lower energy impurity states leaving behind vacant levels.

In this study, each quantum well was assumed to have been doped with an acceptor concentration of 2×10^{18} cm^{-3} over a 20 Å region of the barrier layer. The acceptor levels lie ≈ 30 meV above the valence band edge of AlAs. Since the HH1 state is localised in the GaAs well, approximately 500 meV above the AlAs valence band edge, we can assume that the acceptors are fully ionised. Therefore if the presence of each acceptor impurity causes the removal of a single electron from the HH1 state then the hole population per unit area in the well will be $\mathcal{N}_h = 4 \times 10^{15}$ m^{-2}. It is useful to think of the vacancies created by the removal of an electron from the GaAs layer as mobile hole carriers. Then, optical nonlinearities can be considered to be a result of virtual processes involving transitions made by hole carriers to and from the ground hole state (HH1). The hole carriers at the bottom of the valence band well constitute a two-dimensional hole gas, in the same way that electrons excited into the lowest conduction subband constitute a two-dimensional electron gas when confined by the quantum well potential.

The holes introduced into the HH1 subband occupy states at nonzero in-plane wave vectors near to the centre of the quantum well Brillouin zone. Let us assume that the energy dispersion of the HH1 subband can be described by an effective mass parabola:

$$E = -\frac{\hbar^2 k^2}{2m^*_{HH1}}. \tag{4.2}$$

Within the effective mass approximation, the Fermi level for holes is given by (Ando *et al.*, 1982)

$$E_F = -\frac{\pi\hbar^2\mathcal{N}_h}{m^*_{HH1}}. \tag{4.3}$$

Let us assume that the effective mass of the HH1 subband is equal to the inplane effective mass of heavy holes in bulk GaAs. From the empirical pseudopotential calculations (Table 2.3) we have $m^*_{HH1} = m^{[001]}_{hh} = 0.498m_0$. This gives a Fermi level for holes of 2 meV below the energy of the HH1 state at Γ. The Fermi wave vector k_F defines the area of the k_x-k_y plane containing wave vectors that supply a finite contribution to $\chi^{(2)}$. Accordingly, we restrict the summation over k points in (4.1) to those with x and y components less than $|k_x| = |k_y| = k_F$ where

$$k_F = \sqrt{\frac{2m^*_{HH1}E_F}{\hbar^2}}. \tag{4.4}$$

For the Fermi energy and effective mass given above we have for the Fermi wave vector $k_F = 0.015(2\pi/a_0)$, where a_0 is the lattice constant of GaAs.

4.1.2 Relaxation

The relaxation operator Γ in (4.1) is phenomenologically responsible for the damping from such processes as electron-phonon interactions and collisions among electrons, and ensures that the expression remains finite at resonances. The matrix element Γ_{ij} has a wave vector dependence as well as a dependence on the one electron states i and j. The determination of the values of the damping matrix for all the states participating in the generation of second harmonics is beyond the scope of this thesis. Furthermore, the lack of sufficient experimental data for intravalence band transitions in the particular structures considered here precludes the use of experimentally determined values for the Γ_{ij}. We therefore choose to use a value for each of the damping matrix elements of $\hbar\Gamma = 3$ meV representing, in the absence of a more quantitative datum, a typical linewidth in a good quality quantum well sample. The dependence of $\chi^{(2)}$ on the choice of relaxation constant has been investigated theoretically by Shaw *et al.* (1994). The study revealed a progressive broadening and flattening of the prominent features of the frequency spectra of the second harmonic signal as the damping constant increased from 1.5 meV towards 15 meV. A ten-fold change of the magnitude of $\chi^{(2)}$ was predicted over this range of relaxation rates. Although the uncertainties introduced into the magnitude of the second-order susceptibility due to the adoption of a single value for Γ may well be large, we would expect the major conclusions we arrive at to remain valid, particularly in the limit of low intensity.

4.1.3 Brillouin Zone Summation

The empirical pseudopotential approach we will employ to determine the band structure parameters in the expression for $\chi^{(2)}$ requires that the crystal should be considered infinite in size. The formula obtained for $\chi^{(2)}$ is unsatisfactory as it stands under such circumstances and must be modified slightly.

Quite generally, the number of k states in the first Brillouin zone for the quantum well system, $\mathcal{N}_{BZ}$, is equal to the number of quantum well unit cells in the whole crystal, $\mathcal{N}_C$. It follows that the volume, V, of the entire crystal is

$$V = \mathcal{N}_C V_C = \mathcal{N}_{BZ} V_C \tag{4.5}$$

where V_C is the volume of the quantum well unit cell. For the infinite crystal considered here, the volume normalisation factor $1/V$ in the expression for $\chi^{(2)}$ (4.1) is zero and the k points form a continuum. Under these circumstances, the summation over k points must be replaced by an integral over the Brillouin zone. The integral may be evaluated numerically by calculating contributions to $\chi^{(2)}$ at a random selection of $\mathcal{N}_{SZ}$ points and weighting each of these contributions by a factor of $\mathcal{N}_{BZ}/\mathcal{N}_{SZ}$. We replace the normalisation factor and the summation over k points in (4.1) by

$$\frac{1}{V}\sum_{k}^{\mathcal{N}_{BZ}} \rightarrow \frac{1}{\mathcal{N}_{BZ}V_C}\left(\frac{\mathcal{N}_{BZ}}{\mathcal{N}_{SZ}}\right)\sum_{k}^{\mathcal{N}_{SZ}} = \frac{1}{\mathcal{N}_{SZ}V_C}\sum_{k}^{\mathcal{N}_{SZ}}. \tag{4.6}$$

Further, from the discussion in Section 4.1.1 we know that it is only the volume in the Brillouin zone enclosed by the Fermi surface that supplies a finite contribution to $\chi^{(2)}$. Denoting this volume as V_{SZ} we can restrict the random sampling to this volume by including a factor of V_{SZ}/V_{BZ} in (4.6):

$$\frac{1}{V}\sum_{k}^{\mathcal{N}_{BZ}} \rightarrow \frac{1}{\mathcal{N}_{SZ}}\left(\frac{V_{SZ}}{V_{BZ}}\right)\frac{1}{V_C}\sum_{k}^{\mathcal{N}_{SZ}}. \tag{4.7}$$

By making this substitution, the expression for $\chi^{(2)}$ can be computed under the infinite crystal approach.

4.1.4 Many Body Effects

In the absence of doping, the empirical pseudopotential method provides a good description of the band structure of GaAs/AlAs quantum well systems. When hole carriers are deliberately introduced into the well by doping the barrier layer with acceptors we should anticipate changes in the subband energies through many body effects in the two-dimensional hole gas. For the doping density considered here, the dominant correction to the subband energies is expected to arise from the exchange interaction (Bandara *et al.*, 1988).

For a sufficiently localised ground state wave function, Bandara *et al.* estimated the exchange energy shift of the ground state in an n-type quantum well at the Brillouin zone centre and at the Fermi surface to be

$$E_{exch}(0) \simeq -\frac{e^2 k_F}{4\pi\epsilon}\left[1 - 0.25\left(\frac{k_F}{k_L}\right)\right] \tag{4.8}$$

and

$$E_{exch}(k_F) \simeq -\frac{e^2 k_F}{4\pi\epsilon}\left[\frac{2}{\pi} - 0.25\left(\frac{k_F}{k_L}\right)\right]. \tag{4.9}$$

Here, $k_L = \pi/L$ where L is the length of the well. The above expressions are equally applicable to p-type quantum wells. In this work, the exchange shift is taken as the average of the values at the zone centre and at the Fermi wave vector.

4.2 Geometrical Symmetry of $\chi^{(2)}$

As for any other physical properties of a medium, the second-order susceptibility must reflect the spatial symmetries of the medium. Therefore $\chi^{(2)}$ must transform into itself under all the transformations that leave the medium invariant. This has the consequence that the second-order susceptibilities vanish in centrosymmetric crystals and can lead to a reduction in the number of independent tensor components in crystals that don't possess a centre of inversion.

In this section we use symmetry operations for asymmetric quantum well structures to restrict further the summation over wave vectors in (4.1) and to generate geometric selection rules by which the nonzero, independent components of $\chi^{(2)}$ may be determined.

4.2.1 The Symmetry Operations of an Asymmetric Quantum Well System

The axes of the asymmetric quantum well structure are taken to be the same as those defined in Section 2.1.2 for symmetric wells. That is, with the z axis defined to lie parallel to the quantum well growth direction and the x and y axes defined in the plane parallel to the interfaces, lying along the edges of the conventional fcc unit cell. The operations which leave the asymmetric

Operation	x	y	z	k_x	k_y	k_z	p^x	p^y	p^z	IS
E	x	y	z	k_x	k_y	k_z	p^x	p^y	p^z	a
σ_{ad}	y	x	z	k_y	k_x	k_z	p^y	p^x	p^z	b
σ_{bd}	$-y$	$-x$	z	$-k_y$	$-k_x$	k_z	$-p^y$	$-p^x$	p^z	c
C_{2z}	$-x$	$-y$	z	$-k_x$	$-k_y$	k_z	$-p^x$	$-p^y$	p^z	d
$\mathcal{T}E$	x	y	z	$-k_x$	$-k_y$	$-k_z$	$-p^{x*}$	$-p^{y*}$	$-p^{z*}$	e
$\mathcal{T}\sigma_{ad}$	y	x	z	$-k_y$	$-k_x$	$-k_z$	$-p^{y*}$	$-p^{x*}$	$-p^{z*}$	f
$\mathcal{T}\sigma_{bd}$	$-y$	$-x$	z	k_y	k_x	$-k_z$	p^{y*}	p^{x*}	$-p^{z*}$	g
$\mathcal{T}C_{2z}$	$-x$	$-y$	z	k_x	k_y	$-k_z$	p^{x*}	p^{y*}	$-p^{z*}$	h

Table 4.1: The eight related real space positions and wave vectors under the C_{2v} point group and time reversal symmetry operations, and the Cartesian components of the momentum matrix elements at these wave vectors. The irreducible segment (IS), labelled as in Figure 4.1, to which each symmetry related wave vector belongs is also shown.

quantum well structure invariant are those belonging to the C_{2v} point group: E, σ_{ad}, σ_{bd} and C_{2z}. The crystal Hamiltonian is invariant under these four operations and their time-reversed counterparts denoted as $\mathcal{T}E$, $\mathcal{T}\sigma_{ad}$, $\mathcal{T}\sigma_{bd}$ and $\mathcal{T}C_{2z}$ where $\mathcal{T}$ denotes the time-reversal operation. The effect of these eight symmetry operations on a general point in real space (x, y, z) and in wave vector space (k_x, k_y, k_z) is shown in the second and third columns of Table 4.1. It is possible to generate the quantum well subband energies and momentum matrix elements at the eight symmetry related wave vectors from the subband energies and momentum matrix elements at any one of these wave vectors. Because the quantum well Hamiltonian is invariant under any one of the eight operations in Table 4.1, denoted as S, the energy eigenvalue at a symmetry

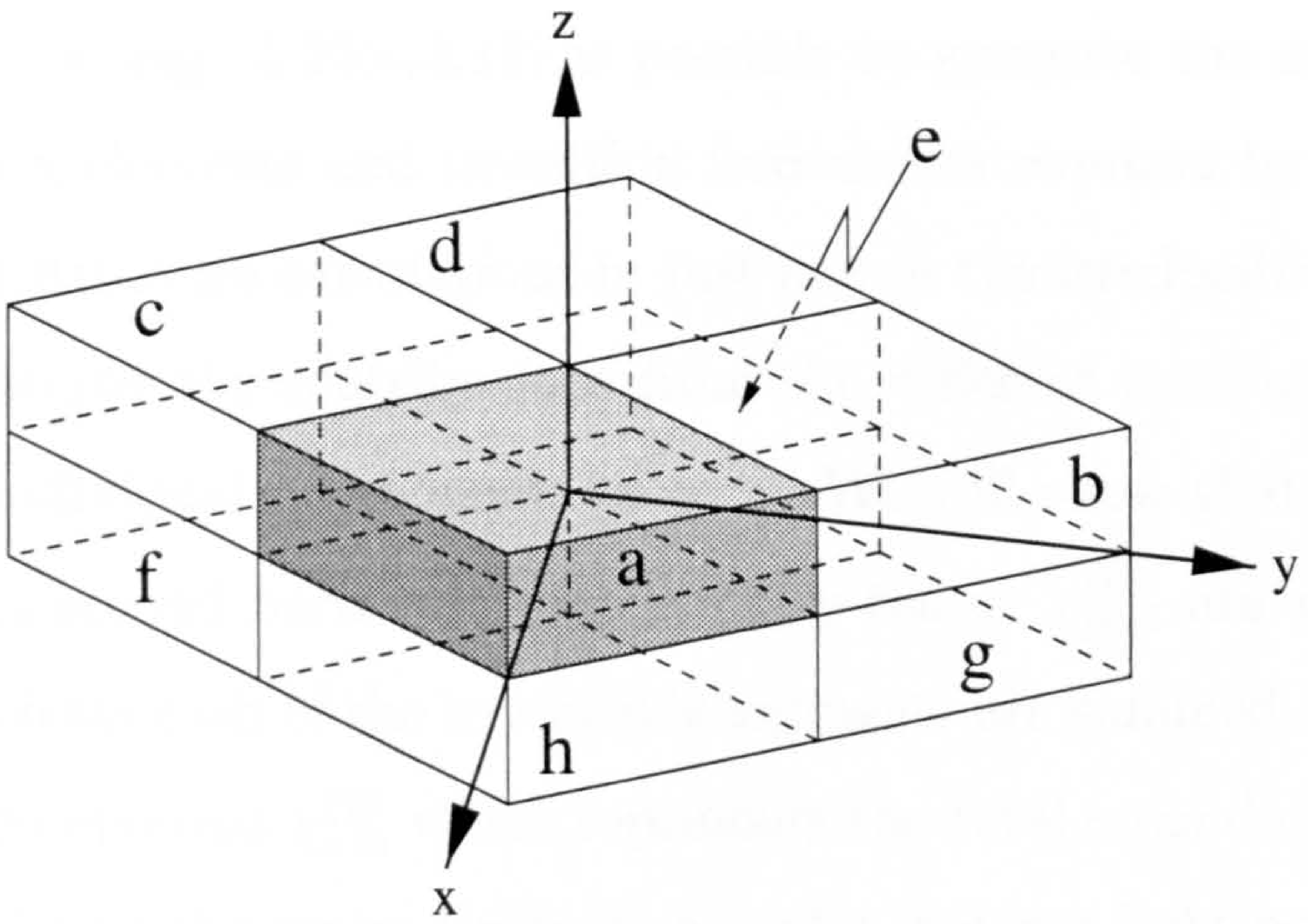

Figure 4.1: The Brillouin zone of a GaAs/AlAs asymmetric quantum well structure. The eight irreducible segments are labelled a through h.

related wave vector $\mathcal{S}\mathbf{k}$ is identical to that at the original wave vector $\mathbf{k}$:

$$E(\mathcal{S}\mathbf{k}) = E(\mathbf{k}). \tag{4.10}$$

The momentum matrix element at a wave vector $\mathcal{G}\mathbf{k}$, where $\mathcal{G}$ denotes an operator of the crystal point group, is related to the wave vector $\mathbf{k}$ by

$$p_{ab}^{\hat{\mathbf{e}}}(\mathcal{G}\mathbf{k}) = p_{ab}^{\mathcal{G}^{-1}\hat{\mathbf{e}}}(\mathbf{k}) \tag{4.11}$$

as can be established by using a straightforward group theoretical approach (Shaw, 1993b). Using the time-reversal operator we can generate the momentum matrix elements at the four remaining symmetry related wave vectors,

$$p_{ab}^{\hat{\mathbf{e}}}(\mathcal{T}\mathcal{G}\mathbf{k}) = -[p_{ab}^{\mathcal{G}^{-1}\hat{\mathbf{e}}}(\mathbf{k})]^*. \tag{4.12}$$

The Cartesian components of the momentum matrix elements at each of the symmetry related wave vectors are given in the fourth column of Table 4.1. Examination of the quantum well Brillouin zone in Figure 4.1 reveals that each symmetry related wave vector occupies one of eight irreducible segments (labelled *a*-*h*). Using (4.10)-(4.12) is possible to generate the zone-wide momentum matrix elements and transition frequencies required by (4.1) by performing band structure calculations in just one of the irreducible segments.

Let us examine the contributions from the different irreducible segments towards the individual components of $\chi^{(2)}$. We will show that certain components of the second harmonic susceptibility tensor $\chi^{(2)}_{\mu\alpha\beta}$ are zero when the contributions from each of the irreducible segments are summed. For example, consider the component $\chi^{(2)}_{xzz}$ which represents the total second-order susceptibility summed over the entire Brillouin zone for incident light polarised in the z direction and a second harmonic beam along the x axis. Each term in the expression for $\chi^{(2)}_{xzz}$ consists of the product of three momentum matrix elements: two polarised in the z direction and one in the x direction. The contributions from each of the irreducible segments can be written in terms of the p^x, p^y and p^z matrix elements in the irreducible segment a in Figure 4.1 using the symmetry transformations given in Table 4.1. For instance, the terms in expression (4.1) for $\chi^{(2)}_{xzz}$ that arise from wave vectors in segment d can be written using the product of the symmetry transformed momentum matrix elements $-p^x$, p^z and p^z. Clearly, since the denominators of (4.1) are the same in each irreducible segment, the contributions to $\chi^{(2)}_{xzz}$ from a will cancel with those from d. Likewise, terms in the expression (4.1) that represent the contributions from segment b can be written so that they contain triple products of the matrix elements p^y, p^z and p^z. These cancel with terms for segment c which consist of triple products of the matrix elements $-p^y$, p^z and p^z. Repeating the above procedure for each segment we find that the component $\chi^{(2)}_{xzz} = 0$. By this method it is possible to show that only thirteen of the twenty-seven components of $\chi^{(2)}$ are nonzero. These are $\chi^{(2)}_{xxz}$, $\chi^{(2)}_{yyz}$, $\chi^{(2)}_{xyz}$, $\chi^{(2)}_{yxz}$, $\chi^{(2)}_{xzx}$, $\chi^{(2)}_{yzy}$, $\chi^{(2)}_{zyy}$, $\chi^{(2)}_{zxx}$, $\chi^{(2)}_{xzy}$, $\chi^{(2)}_{yzx}$, $\chi^{(2)}_{zxy}$, $\chi^{(2)}_{zyx}$ and $\chi^{(2)}_{zzz}$. However, further inspection of the contributions from each of the symmetry related wave vectors shows that not all of these components are independent. Several of the components can be shown to be identical to each other. The nonzero, independent components are

$$\begin{gathered}\chi^{(2)}_{xxz} = \chi^{(2)}_{yyz} \quad \chi^{(2)}_{xyz} = \chi^{(2)}_{yxz} \quad \chi^{(2)}_{xzx} = \chi^{(2)}_{yzy} \\ \chi^{(2)}_{zyy} = \chi^{(2)}_{zxx} \quad \chi^{(2)}_{xzy} = \chi^{(2)}_{yzx} \quad \chi^{(2)}_{zxy} = \chi^{(2)}_{zyx} \\ \chi^{(2)}_{zzz}. \end{gathered} \tag{4.13}$$

For the process of second harmonic generation, the frequencies of the incident fields are indistinguishable and intrinsic permutation symmetry allows us to reduce the number of independent components even further. The nonzero independent components of the second harmonic susceptibility tensor are

$$\begin{aligned}\chi^{(2)}_{xxz}(-2\omega;\omega,\omega) = \chi^{(2)}_{yyz}(-2\omega;\omega,\omega) &= \chi^{(2)}_{xzx}(-2\omega;\omega,\omega) = \chi^{(2)}_{yzy}(-2\omega;\omega,\omega) \\ \chi^{(2)}_{xyz}(-2\omega;\omega,\omega) = \chi^{(2)}_{yxz}(-2\omega;\omega,\omega) &= \chi^{(2)}_{xzy}(-2\omega;\omega,\omega) = \chi^{(2)}_{yzx}(-2\omega;\omega,\omega) \\ \chi^{(2)}_{zxx}(-2\omega;\omega,\omega) &= \chi^{(2)}_{zyy}(-2\omega;\omega,\omega) \\ \chi^{(2)}_{zxy}(-2\omega;\omega,\omega) &= \chi^{(2)}_{zyx}(-2\omega;\omega,\omega) \\ \chi^{(2)}_{zzz}&(-2\omega;\omega,\omega). \end{aligned} \tag{4.14}$$

In the following sections, we present results of calculations of the frequency dependence of the five nonzero independent components of the second harmonic susceptibility for p-type asymmetric GaAs/AlAs quantum wells.

4.3 Computational Considerations

To ensure convergence of the series (4.1) the susceptibility contributions from approximately 10,000 wave vectors in an irreducible segment were determined.

With the symmetry techniques detailed above, the sampling set effectively consisted of 80,000 k points. For the longest period structure investigated here, the time taken to calculate the band structure parameters and the susceptibility contribution at a single wave vector was in excess of 10 minutes using a single Hewlett-Packard HP710 UNIX workstation capable of 12 MFLOPS (Million Floating Point Operations Per Second). At this rate, a calculation for the frequency dependence of a component of the second harmonic susceptibility executed on a single computer could take in excess of 70 days!

The summation of the microscopic susceptibility contributions from wave vectors within the quantum well irreducible segment represents what is sometimes referred to as an "embarrassingly parallel" computational problem. Because an evaluation of (4.1) at a single k point is independent from one at another k point, many such evaluations may take place concurrently. Using a multi-processor computer, the time taken to complete the summation could (theoretically) be reduced by a factor of $1/\mathcal{N}$, where $\mathcal{N}$ is the number of processors. (In practice, though, there is an inherent cost with parallelism, specifically the additional effort of constructing and coordinating the concurrent tasks.)

Recent developments within the high performance scientific computing community have led to the widespread acceptance of networked groups of computers as a viable parallel computing resource. The individual computers in the group resemble processing elements in a distributed memory parallel computer, with a Local Area Network (LAN) and/or a Wide Area Network (WAN) providing a means of inter-processor communication. The creation, synchronisation and coordination of processes running on a distributed parallel computer is controlled by software.

The paradigm of choice for distributed memory parallel processing is the message passing paradigm. In its simplest form the problem is divided up into independent tasks which differ from each other only by the input parameters required to perform the task. The computer code written to perform the task is usually referred to as a slave program. When running on a processor the program generates a slave process. The slave process relies on a master process for the supply of input parameters and provides the master process with the result of performing the task with the given inputs. The information which is transfered between the master and slave process travels through the network in the form of packets. These packets of data constitute the messages. When a slave process has finished and the result has been passed to the master process a new slave process is spawned with a new set of input parameters. This cycle continues until all the tasks which constitute the problem have been performed.

The *de facto* standard message passing software for LAN based parallel processing is the Parallel Virtual Machine (PVM) package (Beguelin *et al.*, 1994) which coordinates message passing between master and slave programs written in C or FORTRAN. Some of the key features of the PVM package are stated in Appendix A. In this section, we outline the development of a parallel algorithm for the calculation of the components of $\chi^{(2)}$. The performance of the algorithm is also discussed, with reference to a serial version of the computer code.

4.3.1 Parallel Implementation

The summation of independent, microscopic susceptibility contributions over a quantum well irreducible segment lends naturally to parallelisation using a master/slave strategy. The master process is responsible for the allocation of a randomly generated wave vector to a slave process. It also sums the

Number of workstations, $\mathcal{N}_w$	Observed wall clock times, τ_w^{obs} (seconds)	Expected wall clock times, $\tau_w^{exp} = \tau_1^{obs}/\mathcal{N}_w$ (seconds)
1	17160	17160
2	8711	8580
4	4333	4290
6	2890	2860
8	2191	2145
10	1758	1716
12	1468	1430
14	1247	1226
16	1100	1073

Table 4.2: Observed and predicted wall clock times τ_w^{obs} and τ_w^{exp}, for a cluster of HP710 workstations as a function of $\mathcal{N}_w$, the number of workstations.

microscopic susceptibility contributions and normalises the total to give the macroscopic second harmonic response. A slave process computes the susceptibility contribution at a given wave vector for a range of input frequencies [using the empirical pseudopotential method and (4.1)] and communicates the result to the master process. The algorithm is dynamically load balancing since the master process allocates a wave vector to a slave process only when the slave is idle. The scheme prevents workload imbalance except at the end of program execution (Schmidt and Sunderam, 1994). The serial FORTRAN program for evaluating (4.1) was parallelised by removing code within the loop over wave vectors to a slave program. Message passing constructs from the PVM library were added to enable communication between the slave program and the remainder of the serial code (the master program).

The performance of the parallel program on a homogeneous cluster of

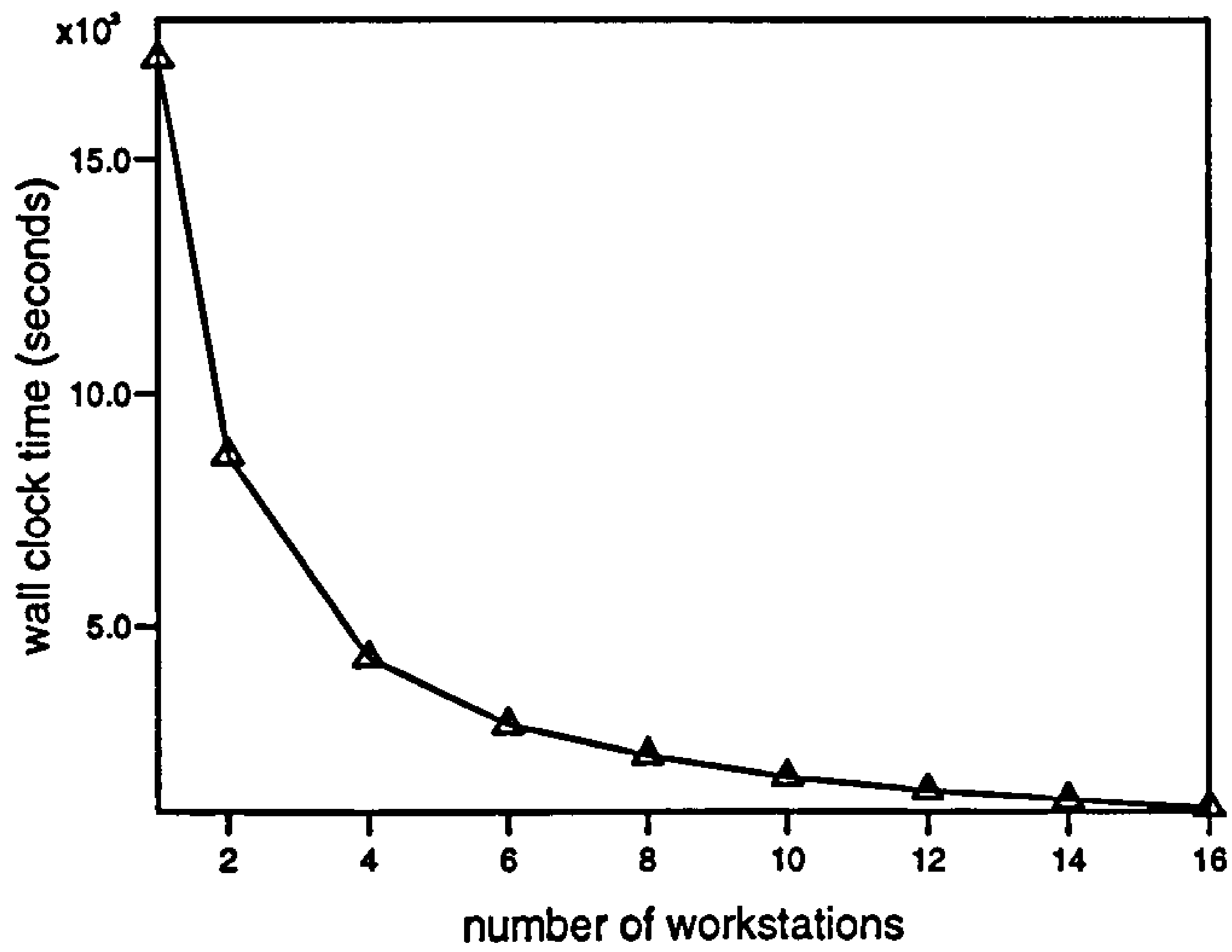

Figure 4.2: Measured wall clock times in seconds on clusters of HP710 workstations for calculations of $\chi^{(2)}_{zzz}(-2\omega;\omega,\omega)$ involving 500 wave vectors as a function of the number of workstations.

Hewlett-Packard HP710 UNIX workstations was investigated. Cluster sizes of 2, 4, 6, 8, 10, 12, 14 and 16 workstations in addition to a master workstation (HP730) were used. The results were compared to serial code execution on a single HP710 computer. The workstations communicated via an Ethernet, employing the user datagram protocol (UDP).

4.3.2 Performance Details and Discussion

Calculations of $\chi^{(2)}_{zzz}(-2\omega;\omega,\omega)$ for a GaAs/Ga$_{0.5}$Al$_{0.5}$As/AlAs quantum well system (where the alloy layer provides the step asymmetry) were performed using a test set of 500 wave vectors randomly generated in the irreducible segment by the master program. The measured wall clock times for different sized clusters of HP710 workstations are reported in Table 4.2. The table includes predicted wall clock times based on the assumption that the serial code

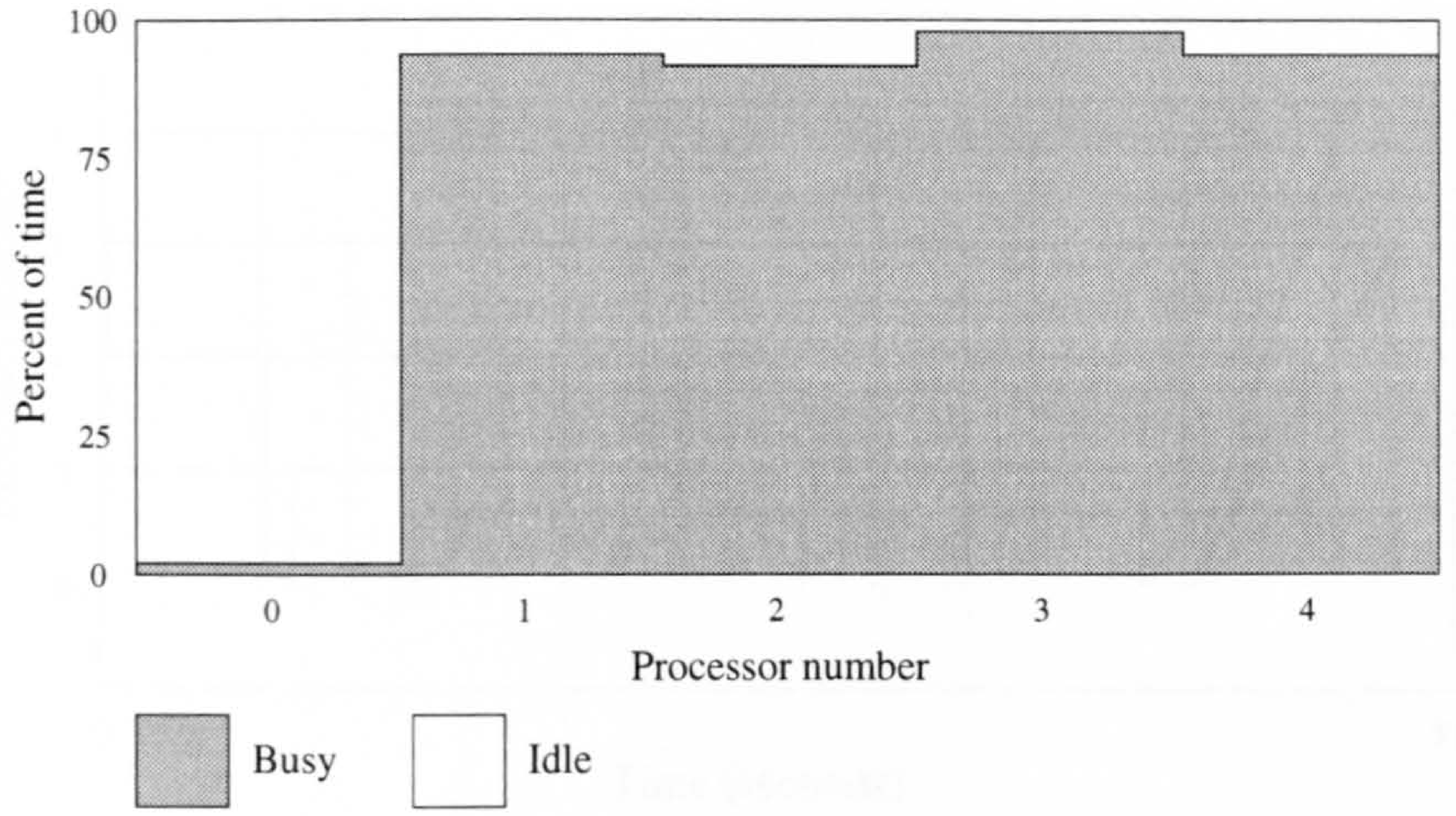

Figure 4.3: The cumulative percentage of time that each processor spent in each of the busy or idle states. The master processor is labelled 0 and the slave processors are labelled 1 to 4.

could be 100% parallelised, i.e. $\tau_w^{exp} = \tau_1^{obs}/\mathcal{N}_w$, where τ_1^{obs} is the wall clock execution time for the serial code, $\mathcal{N}_w$ is the number of workstations and τ_w^{exp} is the expected wall clock time. The table shows the observed and expected wall clock times to be in good agreement. In Figure 4.2 the observed wall clock times in seconds are plotted as a function of the number of workstations. The wall clock time for one workstation is the serial code execution time.

The load balancing and efficiency of the parallel program were examined using the ParaGraph visualisation software (Heath and Etheridge, 1994). The program was executed on a four node cluster of HP710 workstations for a test set of just 28 wave vectors. Figure 4.3 shows the cumulative percentage of time that each processor spent in each of the busy or idle states. The percentage of time is shown on the vertical axis and the processor number on the horizontal axis. Clearly, all four processors are busy for over 90% of the time with work fairly evenly distributed between them.

The "spacetime" diagram in shown in Figure 4.4 indicates processor ac-

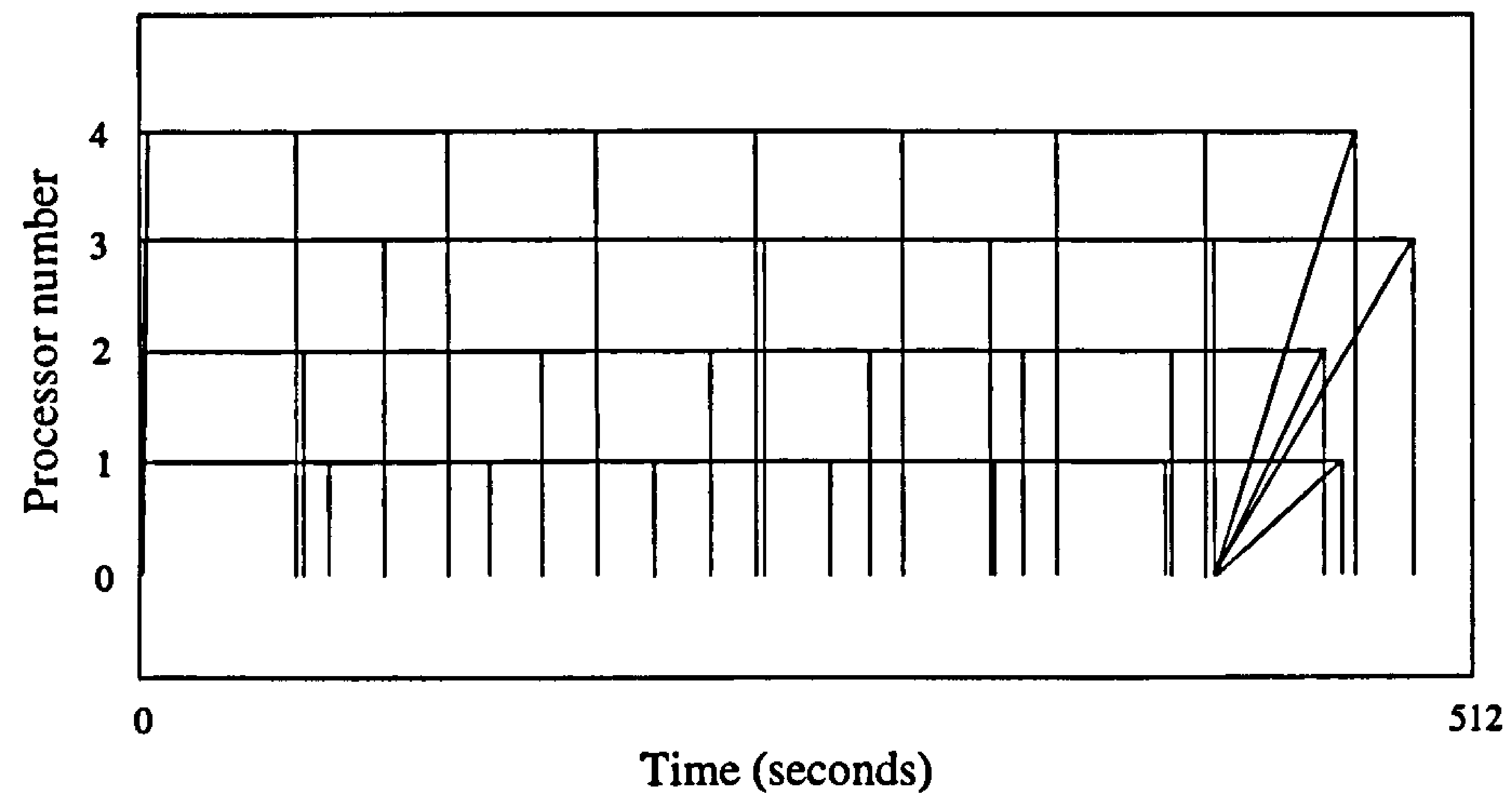

Figure 4.4: A "spacetime" diagram illustrating processor activity and inter node communication as a function of time (seconds). Solid horizontal lines represent periods of busy or overhead activity. Slanted lines depict message passing between the master and slave processors.

tivity (busy/idle) with horizontal lines, one for each processor, with the lines drawn solid if the corresponding processor is busy (or doing overhead) and blank if the processor is idle. Slanted and vertical lines represent messages passing between the master and slave processors. The slope of the lines gives an impression of how soon a given piece of data produced by one processor was needed by the receiving processor. The highly slanted lines towards the end of execution, when there are no more wave vectors in the task queue, depict the transmission of a "poison pill" message to each of the slaves, telling them to die.

The high level of granularity and low communication overhead associated with the algorithm for the calculation of the second harmonic susceptibility components make it well suited to execution on clusters of high-powered workstations coordinated by the PVM package.

4.4 Results of $\chi^{(2)}$ Calculations

The second-order susceptibility at 0 K of a 7GaAs/3Ga$_{0.5}$Al$_{0.5}$As/20AlAs quantum well system (7GaAs represents 7 monolayers, i.e. 3.5 lattice constants of GaAs) was calculated by performing an empirical pseudopotential band structure calculation at randomly selected wave vectors within the irreducible segment, followed by a full scale numerical evaluation of (4.1).

In Figure 4.5 the uppermost four valence subbands are plotted from the point P at the edge of the quantum well Brillouin zone along the [001] direction to the zone centre (Γ) and from Γ along the [100] direction, passing through the point A which marks the edge of the inplane sampling region at $k_x = 0.015(2\pi/a_0)$. The valence subbands are given a heavy hole (HH) and light hole (LH) classification. In Figure 4.6 the charge densities for the top four

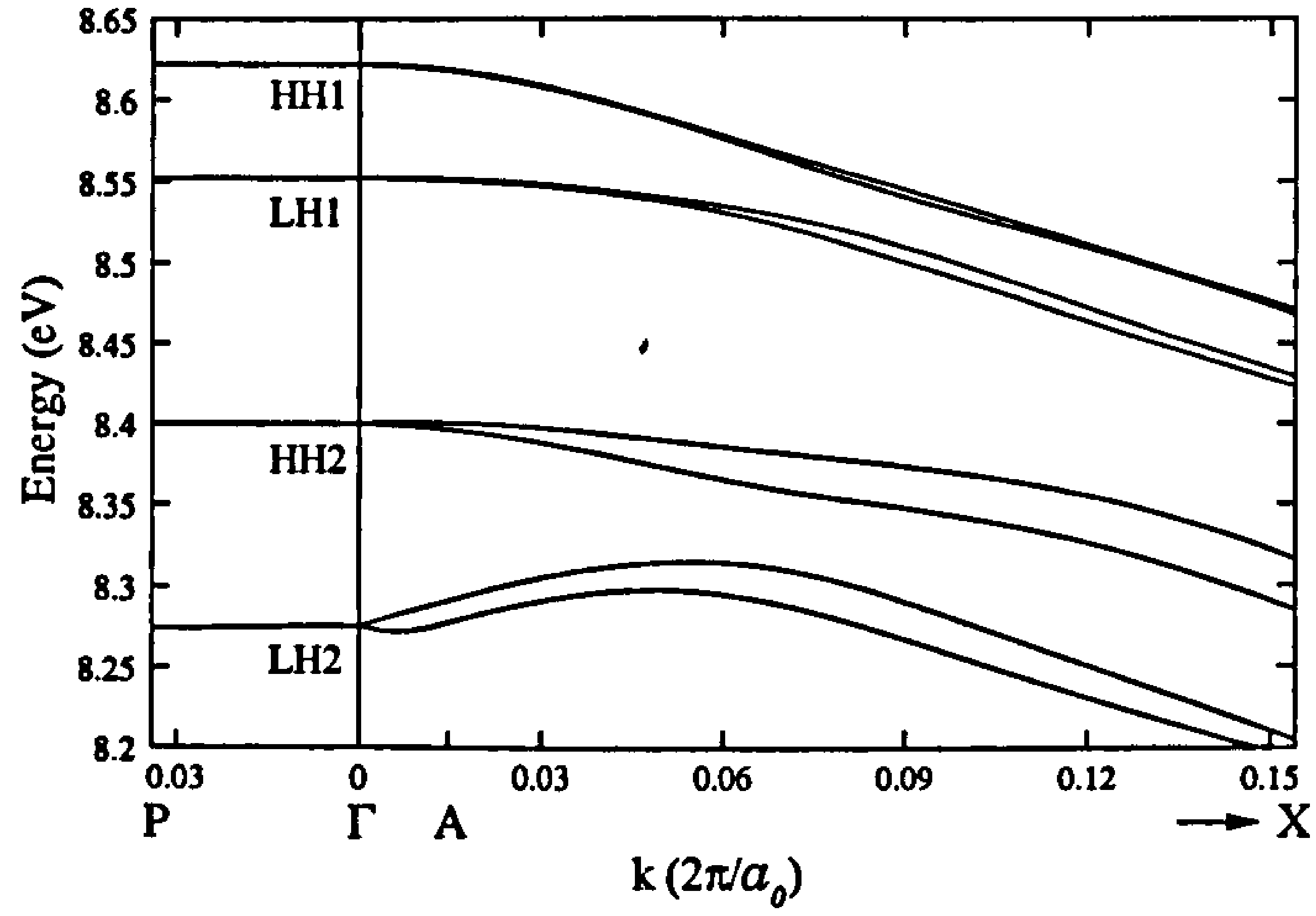

Figure 4.5: The uppermost valence subbands of the 7GaAs/3Ga$_{0.5}$Al$_{0.5}$As/20AlAs quantum well structure plotted from the point P at the edge of the Brillouin zone along z to the zone centre and through the point A along the x direction (7GaAs represents 7 monolayers, i.e. 3.5 lattice constants of GaAs). The bands are given a heavy hole (HH) and light hole (LH) classification.

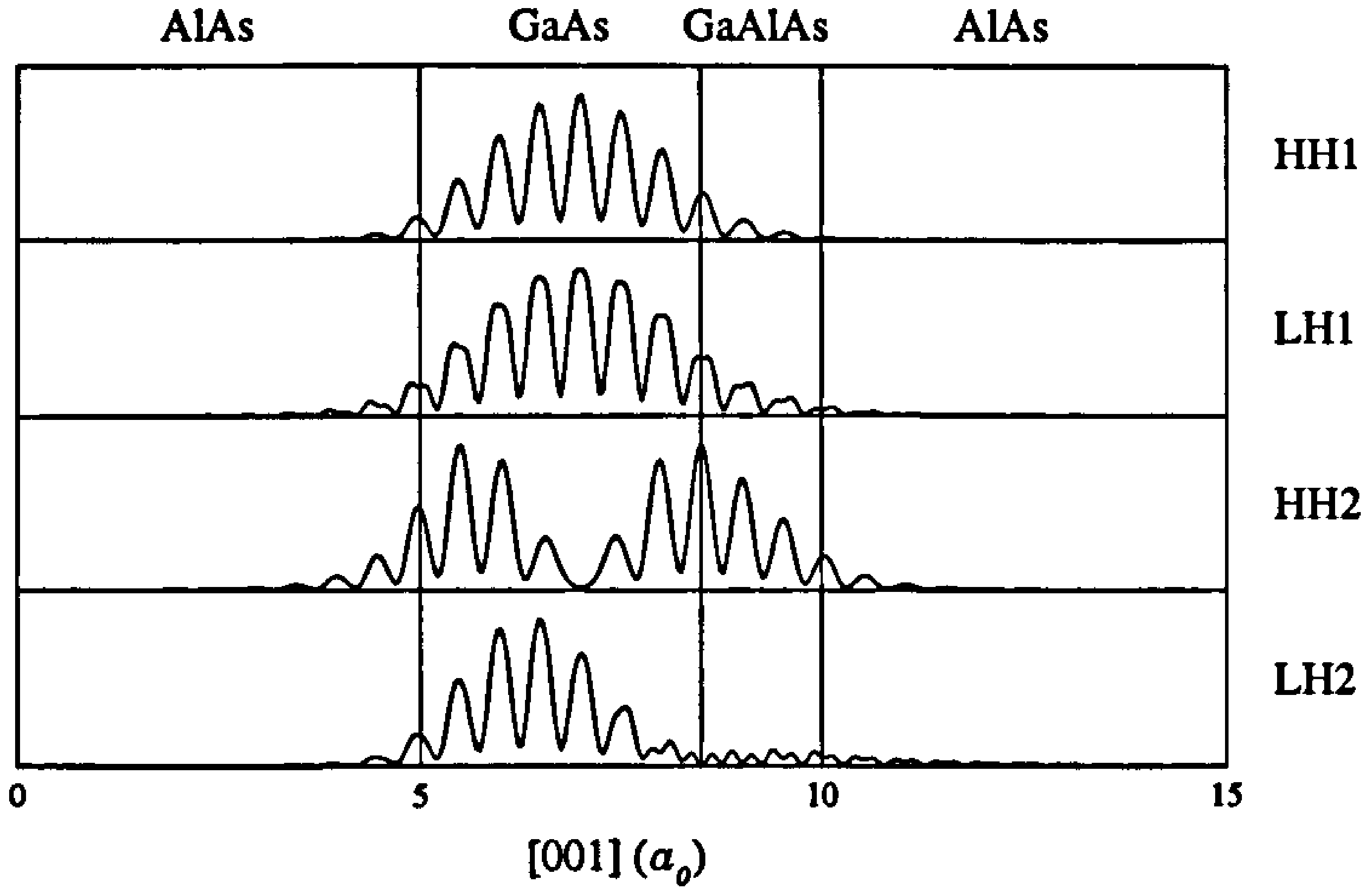

Figure 4.6: Charge density of the uppermost valence subbands at the Brillouin zone centre for the quantum well structure 7GaAs/3$Ga_{0.5}Al_{0.5}As$/20AlAs. The states are given a heavy hole (HH) and light hole (LH) classification.

valence subbands at Γ are given. A schematic level diagram for the structure is shown in Figure 4.7, indicating the zone-centre energies of the first three bound states relative to the 500 meV valence band well. The structure was assumed to have been doped with an acceptor concentration of 2×10^{18} cm^{-3} over a 20 Å region of the barrier (AlAs) layer. This results in a Fermi energy 85 meV below the bulk GaAs zone-centre valence band edge and an areal hole concentration of 4×10^{15} m^{-2} in the GaAs layer. The Fermi wave vector (k_F) defines the area of the $k_x - k_y$ plane (parallel to the interfaces) that supplies a finite contribution to $\chi^{(2)}$. Accordingly, the summation over k points in (4.1) was restricted to those wave vectors with x and y components less than $|k_x| = |k_y| = k_F = 0.015(2\pi/a_0)$, where a_0 is the lattice constant of GaAs. Exchange interactions within the two-dimensional hole gas confined in the GaAs layer were included in the calculation by applying a rigid shift to the HH1 subband energy, as discussed in Section 4.1.4. For the level of doping described above, the increase in HH1 energy was estimated at 15 meV. The

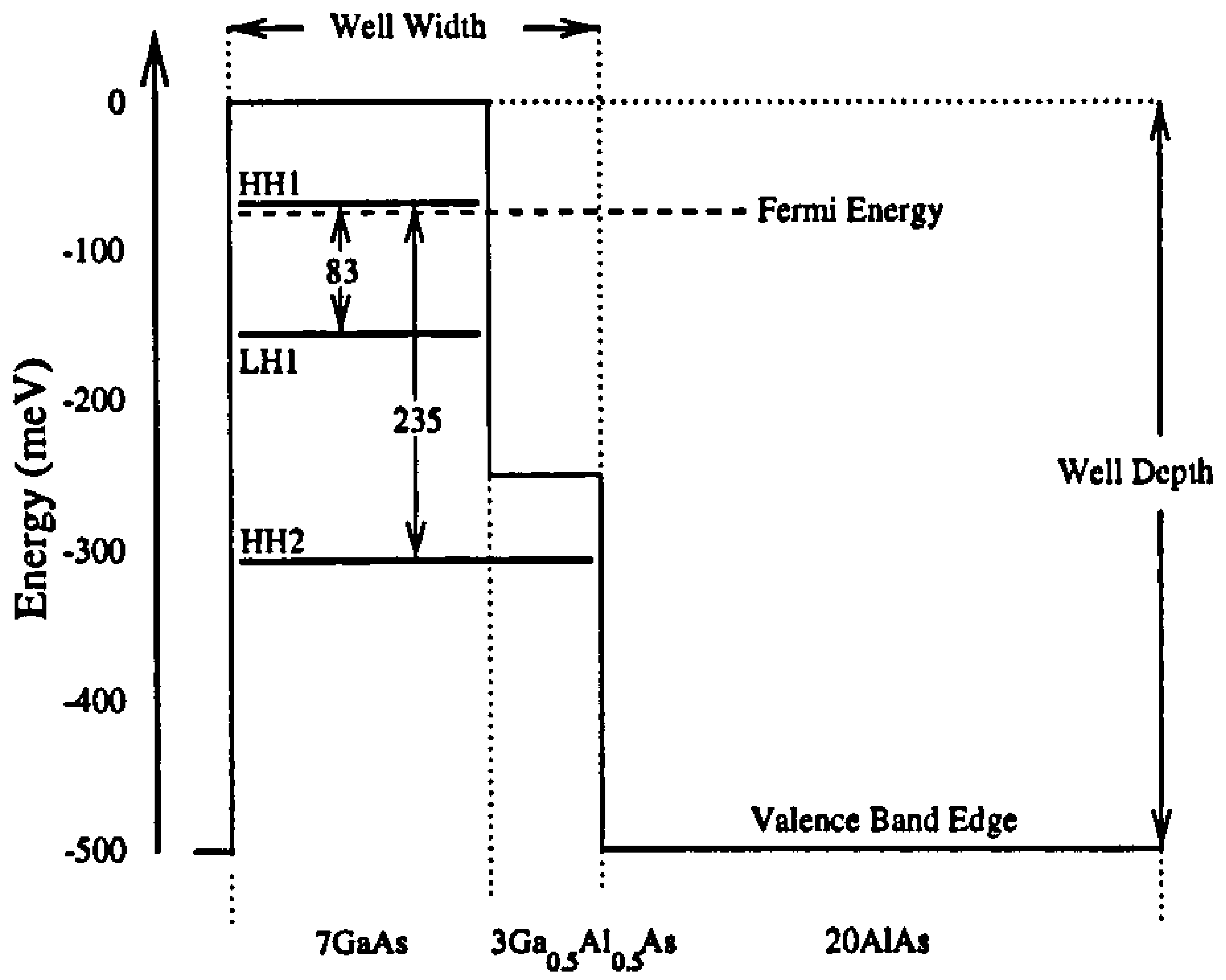

Figure 4.7: Zone-centre valence band energy level diagram of the quantum well structure $7GaAs/3Ga_{0.5}Al_{0.5}As/20AlAs$. The states are given a heavy hole (HH) and light hole (LH) classification. Transition energies between some of the key states are shown.

exchange shift to the energy of the HH1 state is included in Figure 4.7.

The frequency dependence at 0 K of the nonzero independent components of $\chi^{(2)}(-2\omega;\omega,\omega)$ is shown in Figure 4.8. It can be seen that the dominant component is $\chi^{(2)}_{zxy}$. This is convenient for practical applications since the vector of the velocity of the incident light is in the z (quantum well growth) direction (the x and y axis are chosen to lie along the edge of the conventional unit cell of the face centred cubic lattice). The $|\chi^{(2)}_{zxy}|$ spectrum contains peak values of 4.1×10^{-6} esu and 3.5×10^{-6} esu, occurring at photon energies of 82 meV and 117 meV respectively. The virtual process LH1→HH1→HH2→LH1 is responsible for the features in the spectra in Figure 4.8 at 82 meV as the HH1-LH1 transition energy coincides with the photon energy, $\hbar\omega$. This causes the second factors in the frequency denominators of the first two terms of (4.1) to be small. The peaks at 117 meV are generated by the two-photon HH1-HH2 transition, which causes the first factors in the denominators of the first and

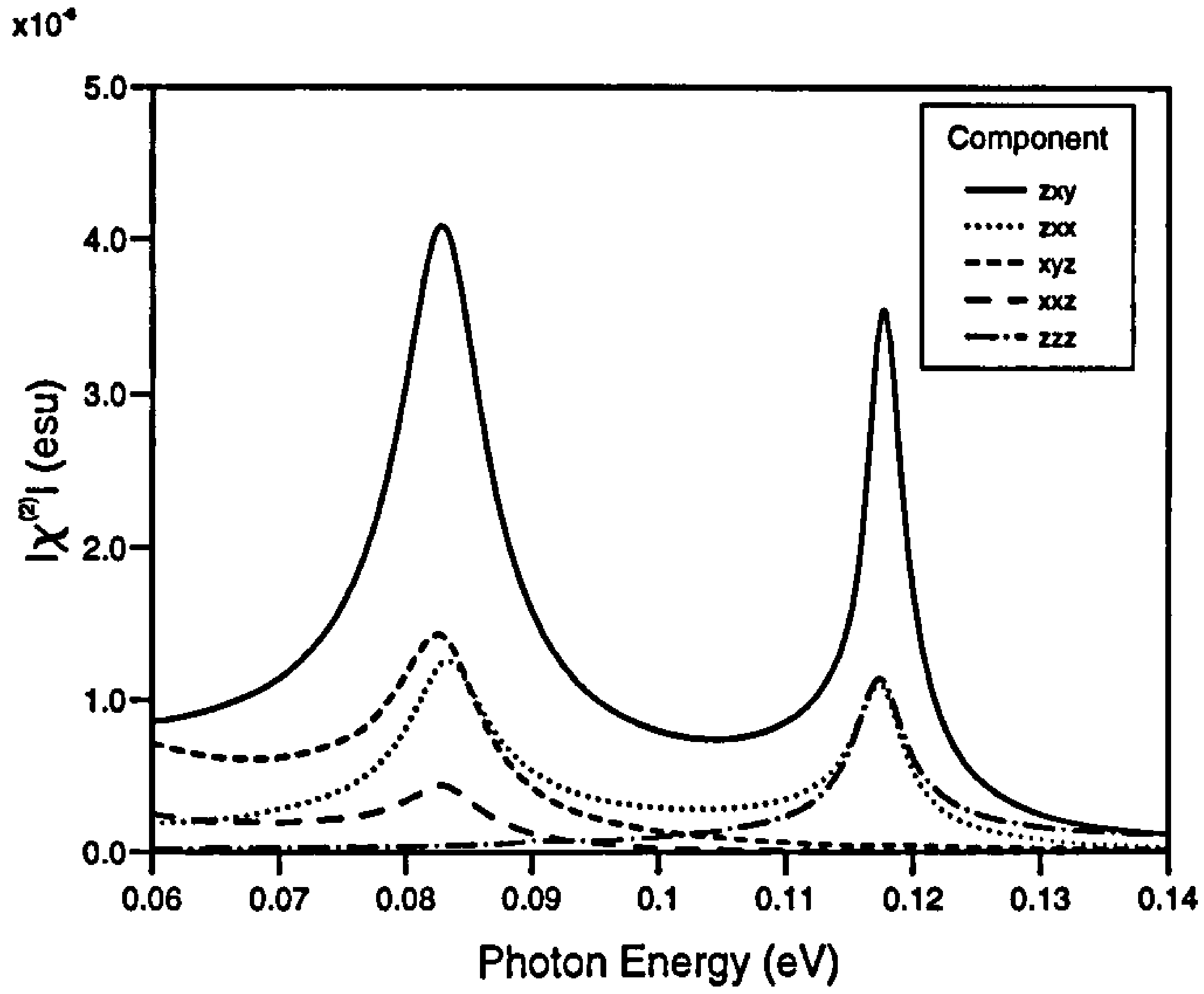

Figure 4.8: The magnitude of the nonzero independent components of $\chi^{(2)}(-2\omega;\omega,\omega)$ in esu, as a function of photon energy in eV, for the *p*-type 7GaAs/3Ga$_{0.5}$Al$_{0.5}$As/20AlAs quantum well structure at 0 K. The component labels $\mu\alpha\beta$ are those in (4.1), where z is in the growth direction. The principle features of the $|\chi^{(2)}_{zxy}|$ spectrum are labelled with the transitions responsible.

third terms of (4.1) to be small. The HH1-HH2 excitation forms a part of the virtual process HH2→LH1→HH1→HH2.

From the particle-in-a-box model (Bastard, 1988) one might have expected the $\chi^{(2)}_{zzz}$ component to be dominant. That this is not the case may be explained by examining the optical matrix elements for the transitions involved in the virtual processes discussed above i.e. HH2-LH1, LH1-HH1 and HH1-HH2. Empirical pseudopotential calculations of the matrix elements show that the z component is the largest between the HH1 and HH2 subbands. The matrix component ratios $|p^x|^2$ to $|p^z|^2$ and $|p^y|^2$ to $|p^z|^2$ for the HH2-LH1 and LH1-HH1 transitions are plotted in Figure 4.9 as a function of wave vector from the centre of the quantum well Brillouin zone to the edge of the sampling region

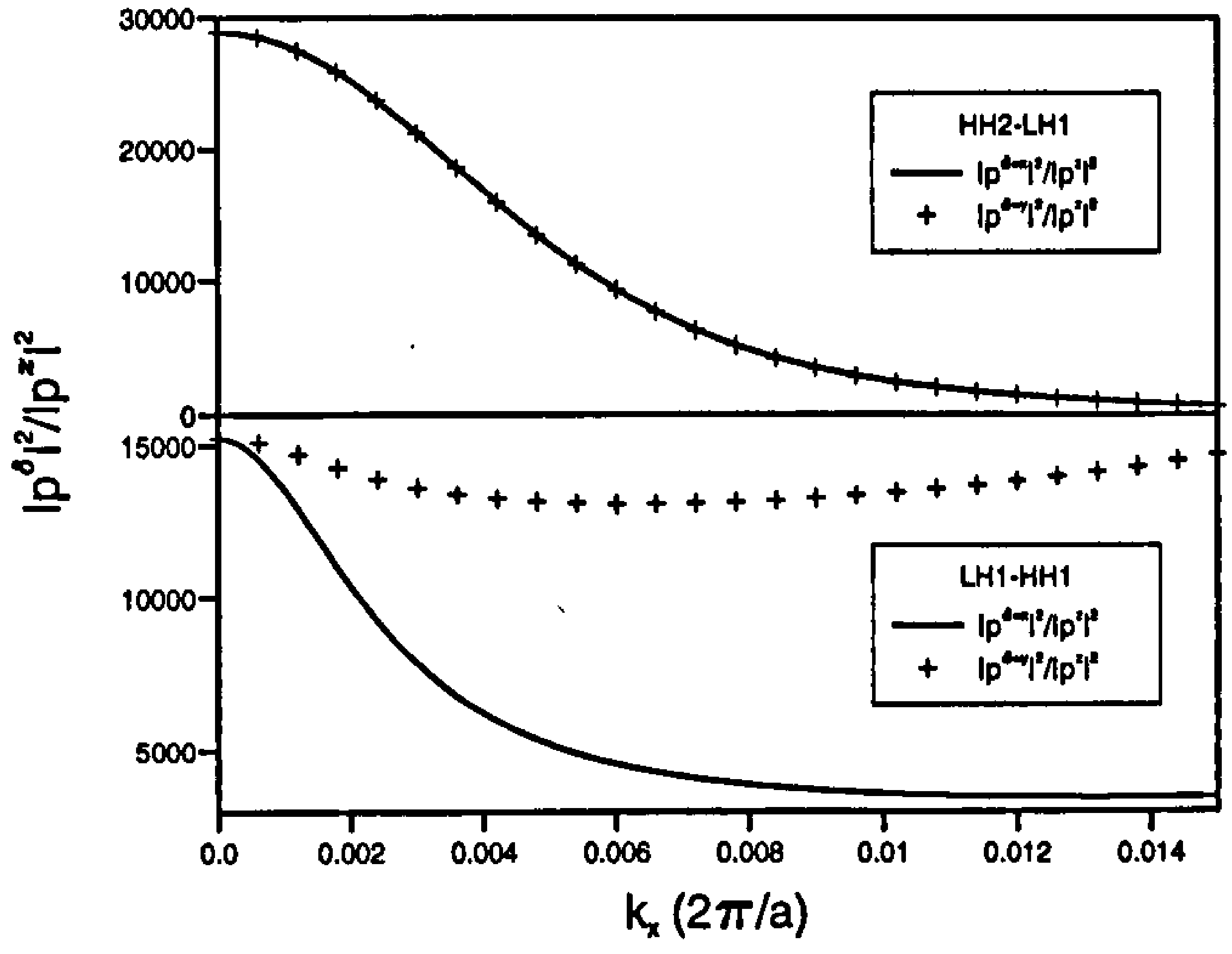

Figure 4.9: The optical matrix element component ratios $|p^x|^2$ to $|p^z|^2$ and $|p^y|^2$ to $|p^z|^2$ for the LH1-HH1 and HH2-LH1 transitions plotted as a function of wave vector from the centre of the quantum well Brillouin zone to the edge of the sampling region at $k_x = 0.015(2\pi/a_0)$, where a_0 is the lattice constant of GaAs.

at $k_x = 0.015(2\pi/a_0)$. Clearly, for both transitions along this symmetry line, the x and y components of the matrix elements are much larger than the z component. This results in a small $\chi^{(2)}_{zzz}$ when the z components of the matrix elements for all three transitions are multiplied to form the numerators of (4.1).

As a consequence of the C_{2v} symmetry of the asymmetric quantum well, certain point group operations transform the x component of the matrix element at a given point in the irreducible segment to the y component of the matrix element at the corresponding point in another segment of the quantum well Brillouin zone. Such transformations were discussed in Section 4.2.1. As a result, for example, $\chi^{(2)}_{zxx}$ is identically equal to $\chi^{(2)}_{zyy}$ when the sum over the whole quantum well Brillouin zone is performed. From inspection of the individual matrix elements we would have expected $\chi^{(2)}_{zyy}$ to differ from $\chi^{(2)}_{zxx}$ since from Figure 4.9 the magnitudes of the x and y components of the matrix elements are distinct for LH1-HH1 transitions. Consequently, the question of whether $\chi^{(2)}_{zxx}$ or $\chi^{(2)}_{zxy}$ is larger cannot be addressed by comparing single matrix elements of x and y polarisations. Rather, this is determined by the relative sizes of the matrix element products which appear in the numerators of (4.1), $p^z_{ac}p^x_{cb}p^y_{ba}$ and $p^z_{ac}p^x_{cb}p^x_{ba}$ for example, throughout the quantum well Brillouin zone. For the quantum well structure in question, it is generally found that the products involved in the dominant processes are larger for $\chi^{(2)}_{zxy}$ than for $\chi^{(2)}_{zxx}$ and as a consequence $\chi^{(2)}_{zxy}$ is approximately three times larger than $\chi^{(2)}_{zxx}$ in Figure 4.8.

The dependence of the peak value of the HH1-HH2 feature in the spectrum of $|\chi^{(2)}_{zxy}|$ on quantum well depth was analysed. This task was accomplished by performing empirical pseudopotential band structure calculations for a 7GaAs/3$Ga_{1-x}Al_xAs$/20$Ga_{1-2x}Al_{2x}As$ quantum well (where $0.1 \leq x \leq 0.5$), possessing valence band well depths spanning the 100 meV to 500 meV range of energies. For each structure studied, the peak value and photon energy of the HH1-HH2 feature was obtained from a $|\chi^{(2)}_{zxy}|$ spectrum generated by (4.1). In Figure 4.10 the HH1-HH2 transition energy is plotted for selected well depths. Clearly, as well depth decreases, the photon energy at which the excitation occurs becomes progressively lower. Because (4.1) contains an $(\omega)^{-3}$ prefactor the lower transition frequency results in an increase in the peak value of the feature. To reveal the influence of the well depth on the peak value we consider a second harmonic response described by $(\hbar\omega)^3|\chi^{(2)}_{zxy}|$. In Figure 4.11 the maximum value of the HH1-HH2 feature given by $(\hbar\omega)^3|\chi^{(2)}_{zxy}|$ is plotted as a function of well depth. The figure implies that as well depth decreases, the ability of a quantum well system to radiate second harmonics is lessened.

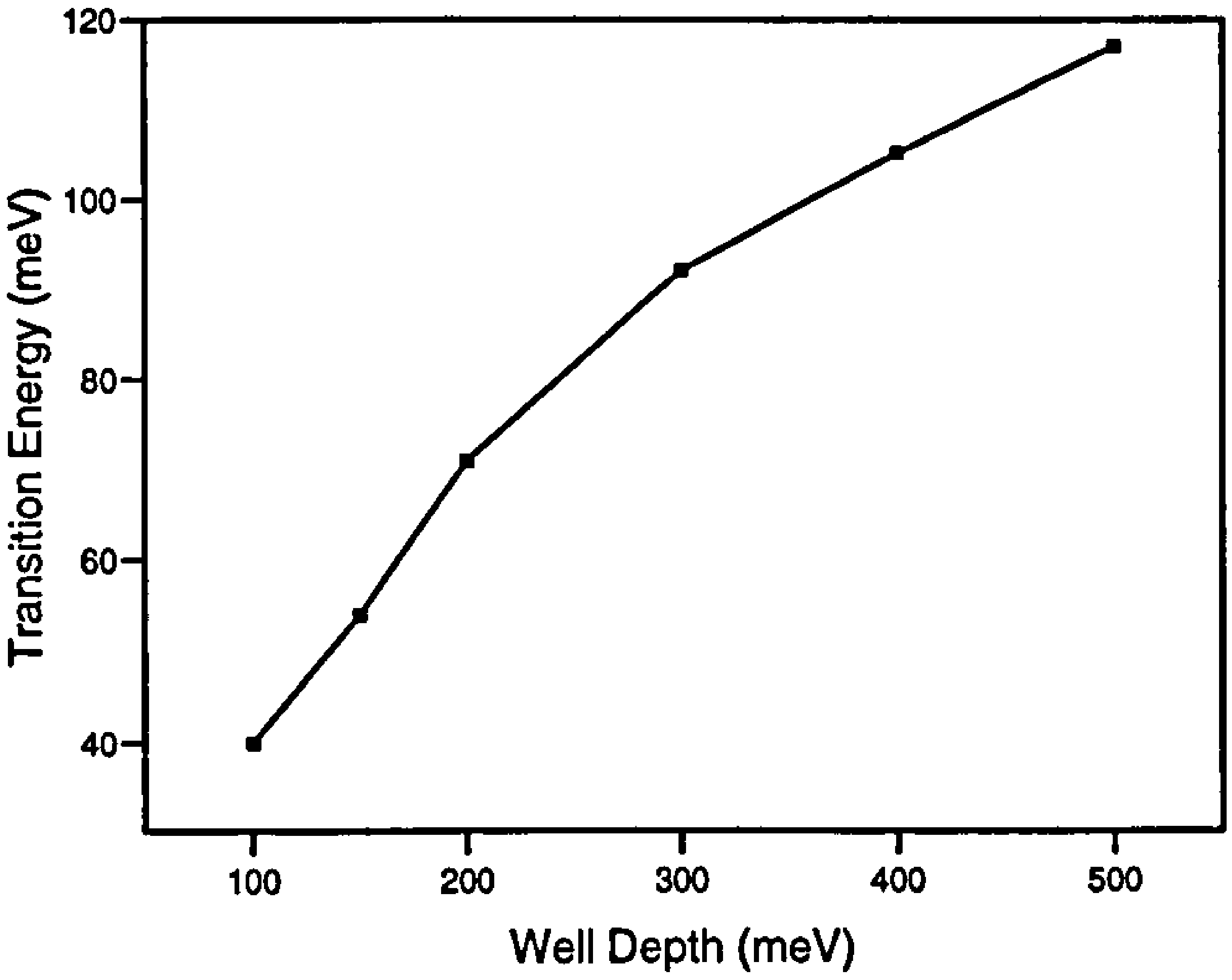

Figure 4.10: The transition energy (meV) associated with the HH1-HH2 feature in the spectrum of $|\chi^{(2)}_{zxy}|$ plotted for a selection of valence band quantum well depths (meV).

Figure 4.11 displays an interesting feature at a well depth of approximately 150 meV. The peak in $(\hbar\omega)^3|\chi^{(2)}_{zxy}|$ can be attributed to the enhancement that occurs when the HH1, LH1 and HH2 states are almost equally spaced causing both frequency denominators in the first term in (4.1) to be simultaneously small. The origin of the reduction in $(\hbar\omega)^3|\chi^{(2)}_{zxy}|$ with decreasing well depth becomes apparent when a detailed analysis of the optical matrix elements concerning transitions to and from the second heavy hole subband is undertaken. We find that the oscillator strengths of transitions involving the HH2 state in the virtual process HH2→LH1→HH1→HH2 decrease as well depth decreases, lowering the contribution to $(\hbar\omega)^3|\chi^{(2)}_{zxy}|$ via the numerators of (4.1). For example, in Figure 4.12 the magnitude of the momentum matrix element for the HH1-HH2 transition induced by light polarised in the z direction is plotted as a function of wave vector for eight different well depths. The figure

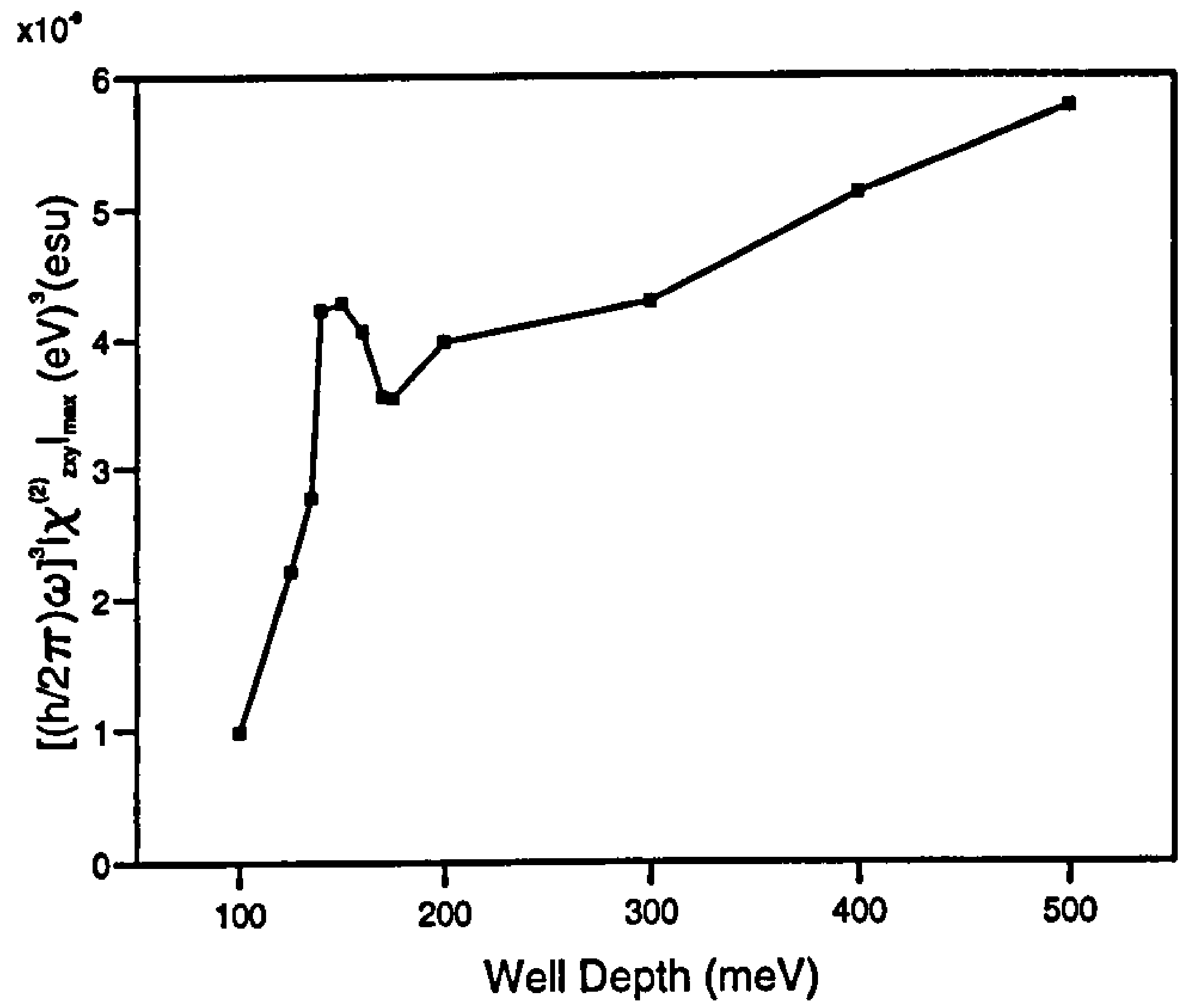

Figure 4.11: The magnitude of the peak value of the HH1-HH2 feature given by $(\hbar\omega)^3|\chi^{(2)}_{zxy}|_{max}$ in $(\mathrm{eV})^3(\mathrm{esu})$ for selected well depths (meV).

includes matrix elements at wave vectors from the point P at the edge of the Brillouin zone in the [001] direction to the zone centre (Γ) and from Γ to the point A located at the Fermi wave vector in the [100] direction. This change in the oscillator strengths is related to the confinement of the HH2 state within the quantum well. In a quantum well possessing a well depth of 500 meV, the HH1, LH1 and HH2 states are strongly bound within the well and the calculations predict large optical matrix elements for the HH1-HH2 and HH2-LH1 transitions. In the structure with a well depth of 100 meV, we find that the HH2 state is located above the semi-classical barrier and is extended in nature. In this case, the oscillator strengths for the HH1-HH2 and HH2-LH1 transitions are very much weaker. The results demonstrate that some degree of band structure optimization for the production of second harmonics is possible through a choice of layer compositions that lead to deep valence band quantum wells. Moreover, the results show that responses aris-

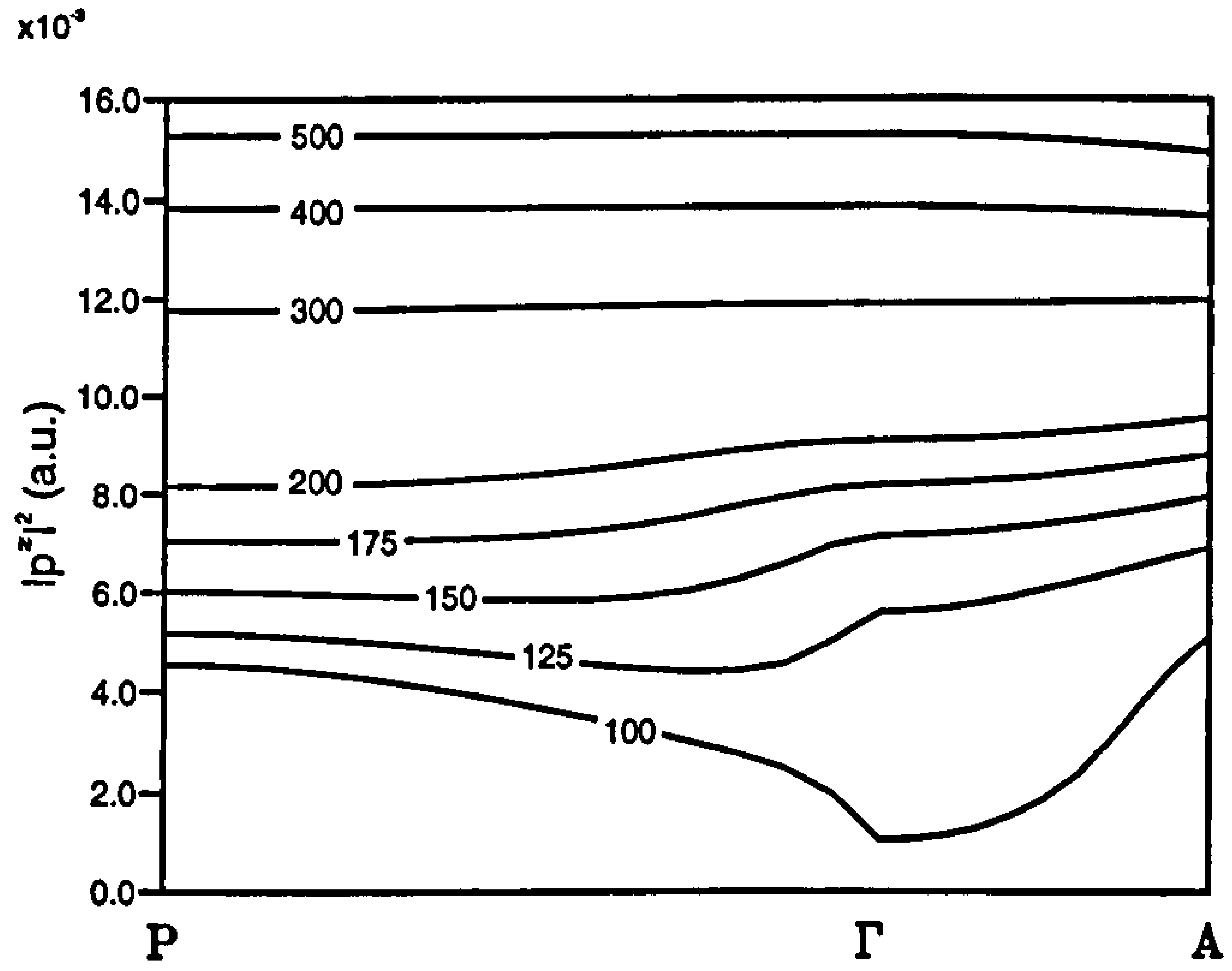

Figure 4.12: The momentum matrix elements due to z (growth direction) polarised light for the HH1-HH2 transition in eight GaAs/AlAs asymmetric quantum well structures possessing valence band well depths between 100 meV and 500 meV. The matrix elements are plotted from the zone centre (Γ) to the point P at the edge of the Brillouin zone along the [001] axis and from Γ to the Fermi wave vector located at A along the [100] axis.

ing from bound-bound-bound state transitions generate stronger second-order nonlinearities than bound-bound-extended state transitions.

The influence of quantum well width on the second harmonic response of a quantum well system was investigated by performing empirical pseudopotential calculations for the structures listed in Table 4.3. Note that all structures considered possess the same period and that increases to the well width (defined as the sum of the widths of the GaAs and alloy layers) occur at the expense of the barrier layer. Also notice that the ratio of the GaAs layer width to the alloy layer width is 2:1 for all quantum well structures, allowing us to neglect differences in the degree of quantum well asymmetry between

Structure (monolayers)	Well Width (monolayers)	Transition Energy (meV)	$(\hbar\omega)^3 \lvert\chi^{(2)}_{zxy}\rvert_{max}$ $\times 10^{-9}(\text{eV})^3(\text{esu})$
8GaAs/18AlAs/4GaAlAs	12	103	5.9
6GaAs/21AlAs/3GaAlAs	9	129	4.8
4GaAs/24AlAs/2GaAlAs	6	154	4.4
2GaAs/27AlAs/1GaAlAs	3	112	0.6

Table 4.3: Well widths and HH1-HH2 transition energies of the four quantum well structures studied. For each structure the composition of the alloy GaAlAs is $(\text{GaAs})_{0.5}(\text{AlAs})_{0.5}$. The maximum values of the transition given by $(\hbar\omega)^3|\chi^{(2)}_{zxy}|$ are in $(\text{eV})^3(\text{esu})$.

structures. Evaluation of (4.1) for each structure showed the dominant component to be $\chi^{(2)}_{zxy}$. The third column in Table 4.3 shows the HH1-HH2 transition energy for all structures. The peak values given by $(\hbar\omega)^3|\chi^{(2)}_{zxy}|$ are listed in the final column in Table 4.3. In Figure 4.13, the transition energies for the three quantum well structures possessing the widest quantum wells are plotted.

We see that the trend in these structures is one of decreasing transition energy with increasing well width. In the structure possessing the narrowest well width (3 monolayers) we notice from Table 4.3 that there is a major deviation from this trend. The reason is that there exist substantial differences in the nature of the three states in this structure in comparison to those in the other quantum well structures considered here. In the wider wells, the HH1 and LH1 state are strongly bound within the deepest section of the well (the GaAs layer) whereas the HH2 state is loosely bound, extending throughout the GaAs and alloy layer. In the narrowest well considered here, the HH1 and LH1 states are loosely bound within the GaAs and alloy layers and the HH2 subband is located above the barrier. Localisation in the wider section of

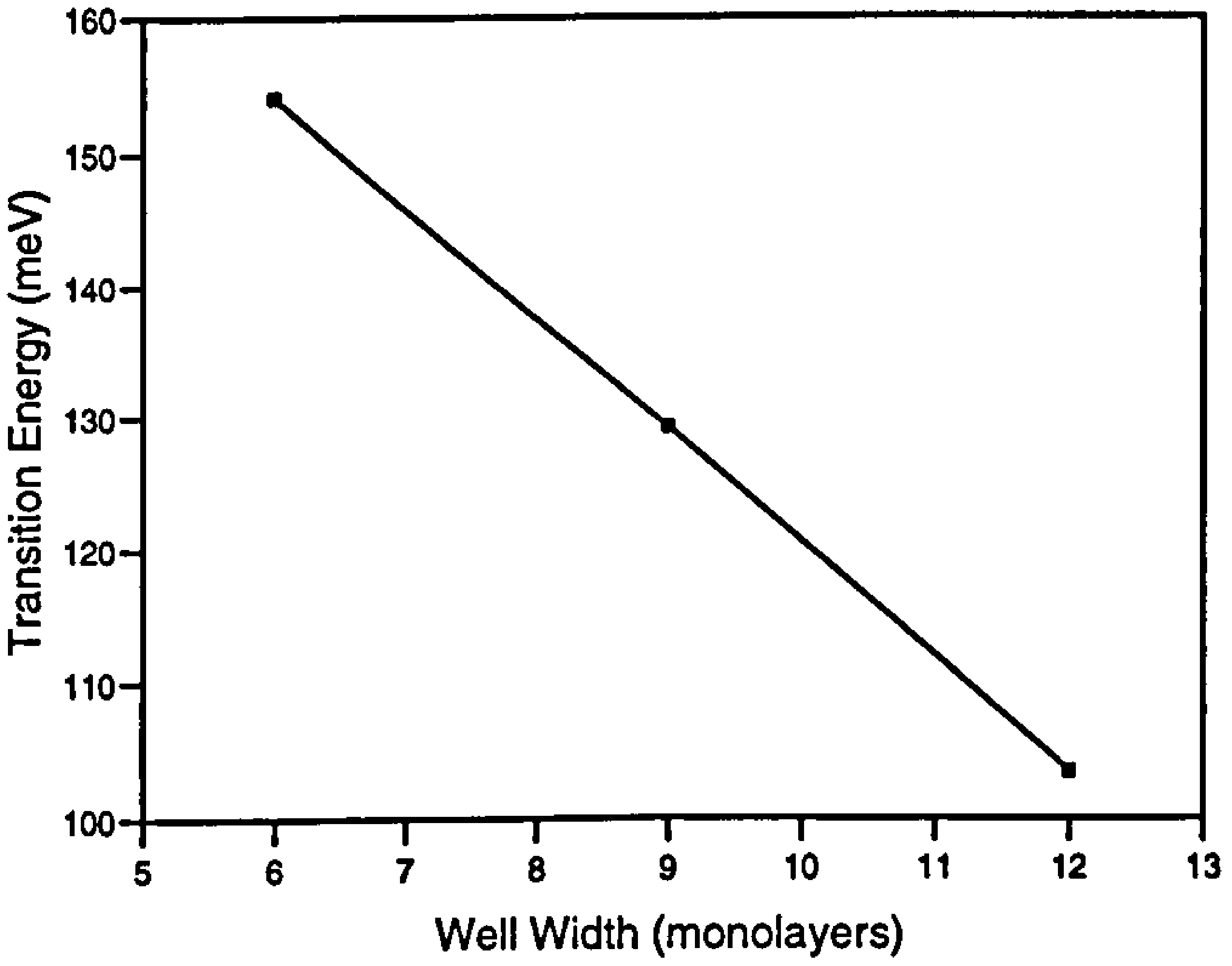

Figure 4.13: Photon energies (meV) associated with the HH1-HH2 transition in the spectrum of $|\chi^{(2)}_{zxy}|$ for three quantum well structures possessing a differing well width (monolayers).

the well produces narrower gaps between the HH1, LH1 and HH2 subbands, leading to a lower HH1-HH2 transition energy.

In Figure 4.14 the peak value of the HH1-HH2 feature is plotted for the three quantum well structures possessing the widest wells. The figure implies that as well width decreases, the second harmonic response becomes weaker. This change in well width means that the HH2 state becomes less confined. As discussed earlier, it is to be expected that the second harmonic response is weaker when one of the three states involved exhibits extended characteristics. These results show that, in common with well depth, a judicious choice of well width can enhance second-order nonlinearities in these structures. We may conclude also that change in the confinement of the participating states is the principle mechanism controlling the extent to which these systems radiate second harmonics.

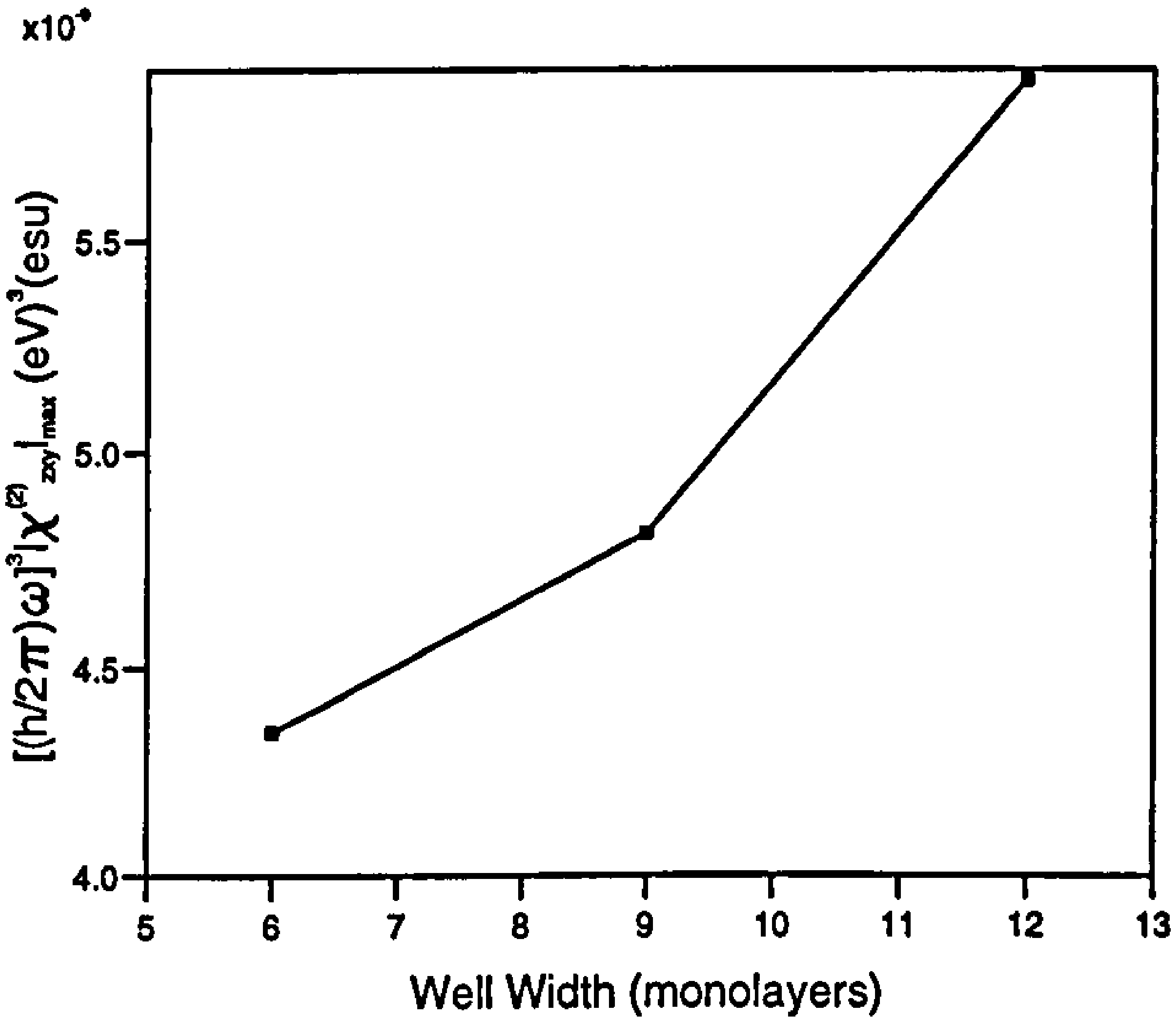

Figure 4.14: The magnitude of the HH1-HH2 peak value given by $(\hbar\omega)^3|\chi^{(2)}_{zxy}|_{max}$ in $(\mathrm{eV})^3$(esu) for selected well widths (monolayers).

The role of the asymmetry in the quantum well profile in determining the magnitude of the second harmonic response was also investigated. Calculations of the $\chi^{(2)}(-2\omega;\omega,\omega)$ spectrum, for incident wavelengths of the order 3 μm, were performed for several short period (10 lattice constant) quantum well structures with varying degrees of asymmetry. The band structures of these systems were such that, at the photon energies of interest, the second-order response was dominated by virtual excitation processes SO→HH1→C→SO and HH2→HH1→C→HH2, where SO is the spin split-off subbands and C is a continuum state lying outside the well. Due to the large number of closely spaced continuum states in the energy range of interest, the precise form and magnitude of the $\chi^{(2)}$ spectrum is determined by the detailed balance between competing processes throughout the quantum well Brillouin zone (Shaw *et al.*, 1993a).

Figure 4.15 shows the $\chi^{(2)}$ response for the 5GaAs/1Ga$_{0.5}$Al$_{0.5}$As/14AlAs

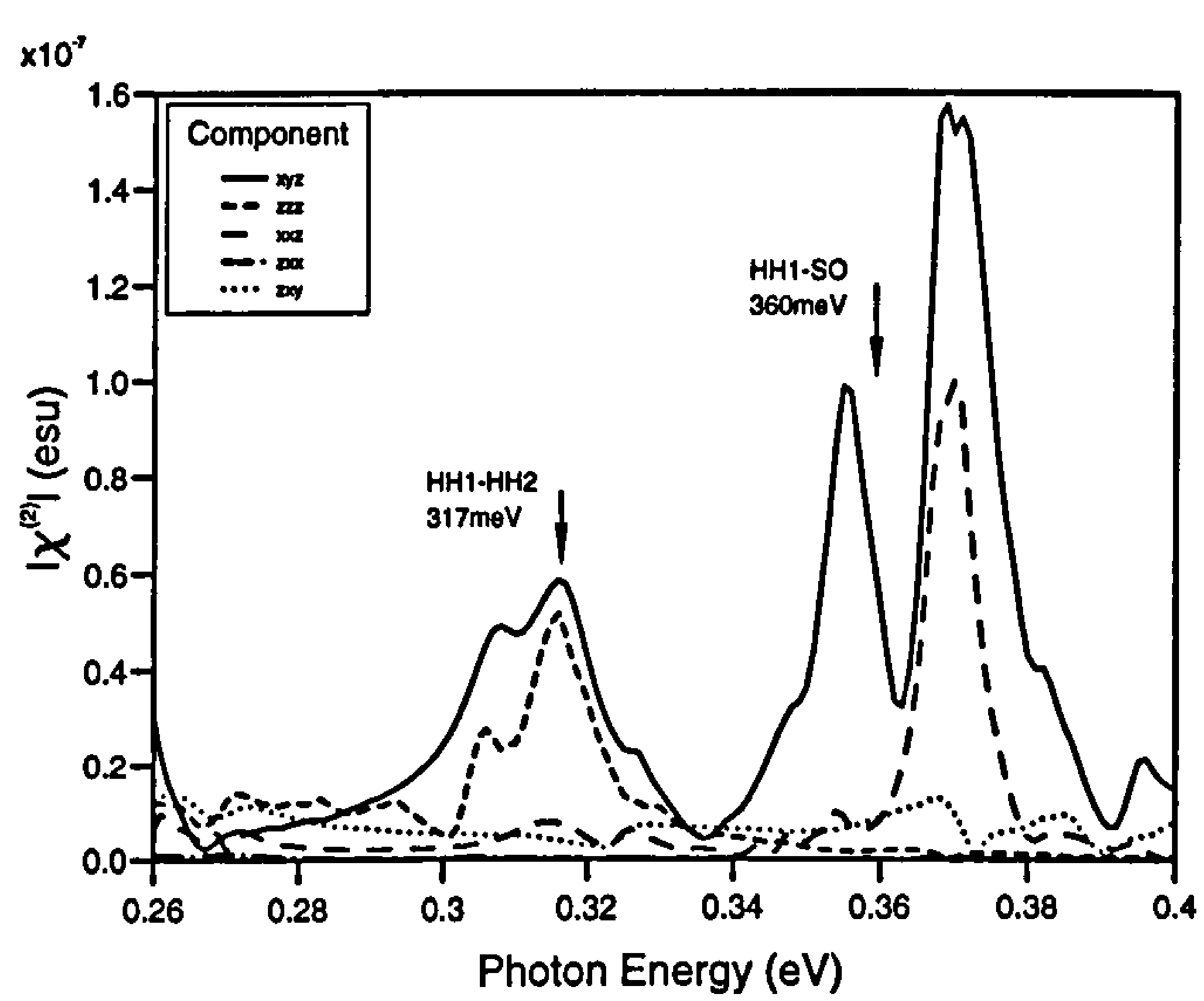

Figure 4.15: The five nonzero components of $\chi^{(2)}(-2\omega;\omega,\omega)$ (esu) plotted against incident photon energy (eV) for the 5GaAs/1$\text{Al}_{0.5}\text{Ga}_{0.5}$As/14AlAs quantum well structure at 0 K. The zone-centre subband energy separations are indicated by arrows, the shifts in the peak positions clearly illustrating the importance of contributions from other regions of the quantum well Brillouin zone.

quantum well system, doped in the barriers to a hole concentration per unit area of 4×10^{15} m^{-2}. The principal peaks in $\chi^{(2)}_{xyz}$, the dominant component for this structure, occur at 355 meV and 375 meV resulting from SO-HH1 transitions. It is clear from Figure 4.15 the energies of these peaks do not coincide with the zone-centre separation of the SO and HH1 subbands. This displacement of the maxima in $\chi^{(2)}_{xyz}$ occurs due to the summation over the quantum well Brillouin zone and the effect of two-photon excitations with continuum states away from the zone-centre, in accord with the observations by Shaw *et al.* (1993a). The peaks at 315 meV in both $\chi^{(2)}_{xyz}$ and $\chi^{(2)}_{zzz}$ occur very close to the zone-centre energy separation of the HH1 and HH2 subbands. These peaks result from the HH1-HH2 processes whose optical transition probabili-

Structure	Asymmetry	$(\hbar\omega)^3\lvert\chi^{(2)}_{xyz}\rvert_{max}$	$(\hbar\omega)^3\lvert\chi^{(2)}_{zzz}\rvert_{max}$
(monolayers)	Ξ	$\times10^{-9}(\text{eV})^3(\text{esu})$	
5GaAs/14AlAs/1GaAlAs	0.2	10.4	2.8
5GaAs/13AlAs/2GaAlAs	0.4	10.4	11.0
4GaAs/14AlAs/2GaAlAs	0.5	5.1	6.1
4GaAs/13AlAs/3GaAlAs	0.75	8.1	14.2
4GaAs/12AlAs/4GaAlAs	1.0	16.7	24.4

Table 4.4: The five quantum well structures studied and their degree of asymmetry, Ξ. For each structure the composition of the alloy GaAlAs is $(\text{GaAs})_{0.5}(\text{AlAs})_{0.5}$. The maximum values of $(\hbar\omega)^3|\chi^{(2)}_{xyz}|$ and $(\hbar\omega)^3|\chi^{(2)}_{zzz}|$, where $\hbar\omega$ is in eV and $\chi^{(2)}$ in esu.

ties become very small away from the centre of the zone.

The dependence of second-order processes on asymmetry is expected to cause an increase in $\chi^{(2)}$ as the symmetry of the quantum well is decreased. To test this, several wells of differing asymmetry were studied. Here we define the parameter Ξ, a quantitative measure of the degree of asymmetry, by

$$\Xi = \begin{cases} w_{GaAs}/w_{alloy} & \text{if } w_{GaAs} < w_{alloy} \\ w_{alloy}/w_{GaAs} & \text{if } w_{GaAs} > w_{alloy}, \end{cases} \tag{4.15}$$

where w_{GaAs} is the width of the GaAs well and w_{alloy} the width of the alloy layer. The structures investigated are listed in Table 4.4 with the corresponding degree of asymmetry. For each structure a full band structure evaluation of $\chi^{(2)}_{xyz}$ and $\chi^{(2)}_{zzz}$ was performed. The maximum values of $(\hbar\omega)^3|\chi^{(2)}_{xyz}|$ and $(\hbar\omega)^3|\chi^{(2)}_{zzz}|$ obtained for each quantum well structure are also shown in Table 4.4 [again, $(\hbar\omega)^3|\chi^{(2)}|$ has been used to account for the $(\omega)^{-3}$ factor in (4.1)] and are plotted against Ξ in Figure 4.16.

It can be seen that, in general, the magnitude of $(\hbar\omega)^3|\chi^{(2)}_{zzz}|$ increases with

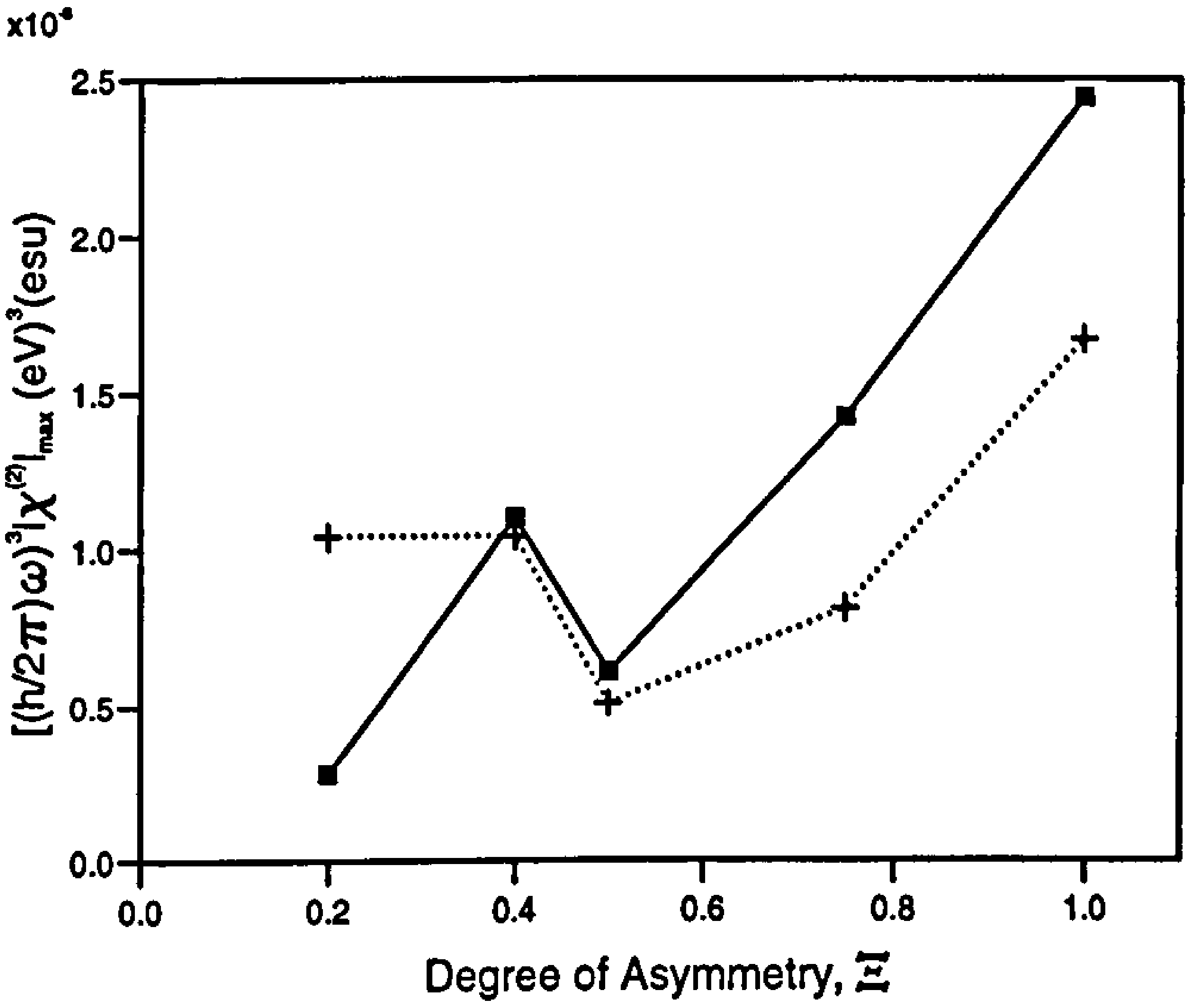

Figure 4.16: The maximum value of $(\hbar\omega)^3|\chi^{(2)}_{zzz}|$ (squares) and $(\hbar\omega)^3|\chi^{(2)}_{xyz}|$ (crosses) in $(\mathrm{eV})^3(\mathrm{esu})$ is plotted against the degree of asymmetry Ξ for each of the five structures studied.

Ξ. Now, it has also been shown (above) that the magnitude of the second-order susceptibility is sensitive to the detailed structure of the subband energies throughout the quantum well Brillouin zone. Such "band structure effects" result in fluctuations in the plot of $(\hbar\omega)^3|\chi^{(2)}_{zzz}|$ against Ξ, clearly seen in Figure 4.16. For $(\hbar\omega)^3|\chi^{(2)}_{xyz}|$ it is evident from Figure 4.16 that the magnitude of the band structure effect fluctuations is comparable to the change due to any trend with asymmetry which may exist. The xyz component of $\chi^{(2)}$ is seen to be less sensitive to asymmetry than the zzz component. That this is the case may be explained by the fact that $(\hbar\omega)^3|\chi^{(2)}_{zzz}|$ originates from processes involving three transitions polarized in the growth direction, the direction in which the symmetry has primarily been reduced, while $(\hbar\omega)^3|\chi^{(2)}_{xyz}|$ involves only one such transition. One can, therefore, generally obtain an increased second harmonic response by designing structures with a large asymmetry, though this applies mainly to the parallel incidence response, $\chi^{(2)}_{zzz}$. However, the sensitivity to the detailed band structure of the quantum well structure *can* cause the second-order response to reduce as the asymmetry is increased, negating any benefit deriving from the increased asymmetry.

It is of interest to qualitatively consider the effects of a finite temperature on the second harmonic spectra presented in this chapter. For temperatures above 0 K we would expect the Fermi energy to become increasingly more negative on the scale of Figure 4.7, and as a consequence, the hole carriers in the HH1 subband to occupy an increasingly greater area of the k_x-k_y plane. This means that contributions to $\chi^{(2)}$ would now come from virtual transition processes at points lying far away from the Brillouin zone centre where the subband energy gaps and momentum matrix elements can be very different from those in the neighbourhood of Γ. We can anticipate that a broadening of the principle features of the $\chi^{(2)}$ spectra will take place as a result. Furthermore, for finite temperatures we can expect that holes will eventually populate the ground light hole subband (LH1). This will result in contributions to the second-order susceptibility arising from processes involving LH1 as the unoccupied state (in contrast to the processes considered here where HH1 is the unoccupied state). We would expect additional features in the spectra at nonzero temperatures as a result of these processes.

It has been demonstrated that order to optimise the second harmonic response of *p*-type GaAs/AlAs quantum well systems, well parameters such as depth and width should be chosen as to maximize the confinement of the participating states. A typical set of such parameters was presented above. It was also found that a high degree of well asymmetry can result in an increased second harmonic response in these structures, although the response is highly sensitive to the detailed band structure of the quantum well structure.

Chapter 5

The Electronic Structure of Semiconductor Quantum Dots

When the concept of one-dimensional confinement of charge in a semiconductor quantum well is abstracted to three dimensions we obtain a quantum dot. While there is still a dispersion in the conduction and valence band in quantum wells, quantum dots exhibit an atomic-like spectrum and a density of states that resembles a series of delta functions. The changed dimensionality allows greater control over the optical and electronic properties of quantum dots than can be achieved by engineering the band structure of quantum wells.

In this chapter we examine the electronic structure of quantum dots fabricated from the group III-V compounds and consider the effects of material parameters such as dot shape, size and composition on energy levels, wave functions and optical transitions. We make use of the $\mathbf{k} \cdot \mathbf{p}$/EMA method described in Section 2.2.3 to determine the electronic structure of arbitrarily shaped quantum dots.

5.1 GaAs/AlAs Cubic Quantum Dots

In this section we investigate the electronic and optical properties of quantum dots consisting of a cube of GaAs embedded in an AlAs barrier. For simplicity we assume that the quantum dot conduction states are adequately described by the single-band effective mass theory, that is, the slowly varying part of the quantum dot wave function satisfies the equation

$$\left[-\frac{\hbar^2}{2m_e^*}\nabla^2 + V(\mathbf{r})\right] F(\mathbf{r}) = EF(\mathbf{r}) \tag{5.1}$$

where m_e^* is the isotropic effective mass of GaAs at the bottom of the conduction band and $V(\mathbf{r})$ is the three dimensional confining potential due to the step-like difference of the band edge potentials in GaAs and AlAs. The general solutions for the electron states are

$$\psi_n^c(\mathbf{r}) = F_n^c(\mathbf{r})u^c(\mathbf{r}) \tag{5.2}$$

where u^c is a bulk band edge Bloch function and the quantum dot states are labelled with the index n. A carrier mass of $m_e^* = 0.0804m_0$ was assumed, derived from empirical pseudopotential calculations of the band structure of bulk GaAs. The envelope functions of the valence band states in the quantum dot are solutions to the four-band effective mass equation (2.58). We assume that the Luttinger valence band parameters used in (2.58) are those of bulk GaAs given in Table 2.3.

Since periodic boundary conditions are required in the electronic structure calculation, we consider a three dimensional array of GaAs dots each of which lies inside a cubic unit cell of AlAs with sides chosen such that states in neighbouring dots do not interact (see Figure 5.1). The conduction states in the quantum dot are determined from the solution of (5.1). The confining potential is given by the conduction band offset (ΔE_c) between GaAs and

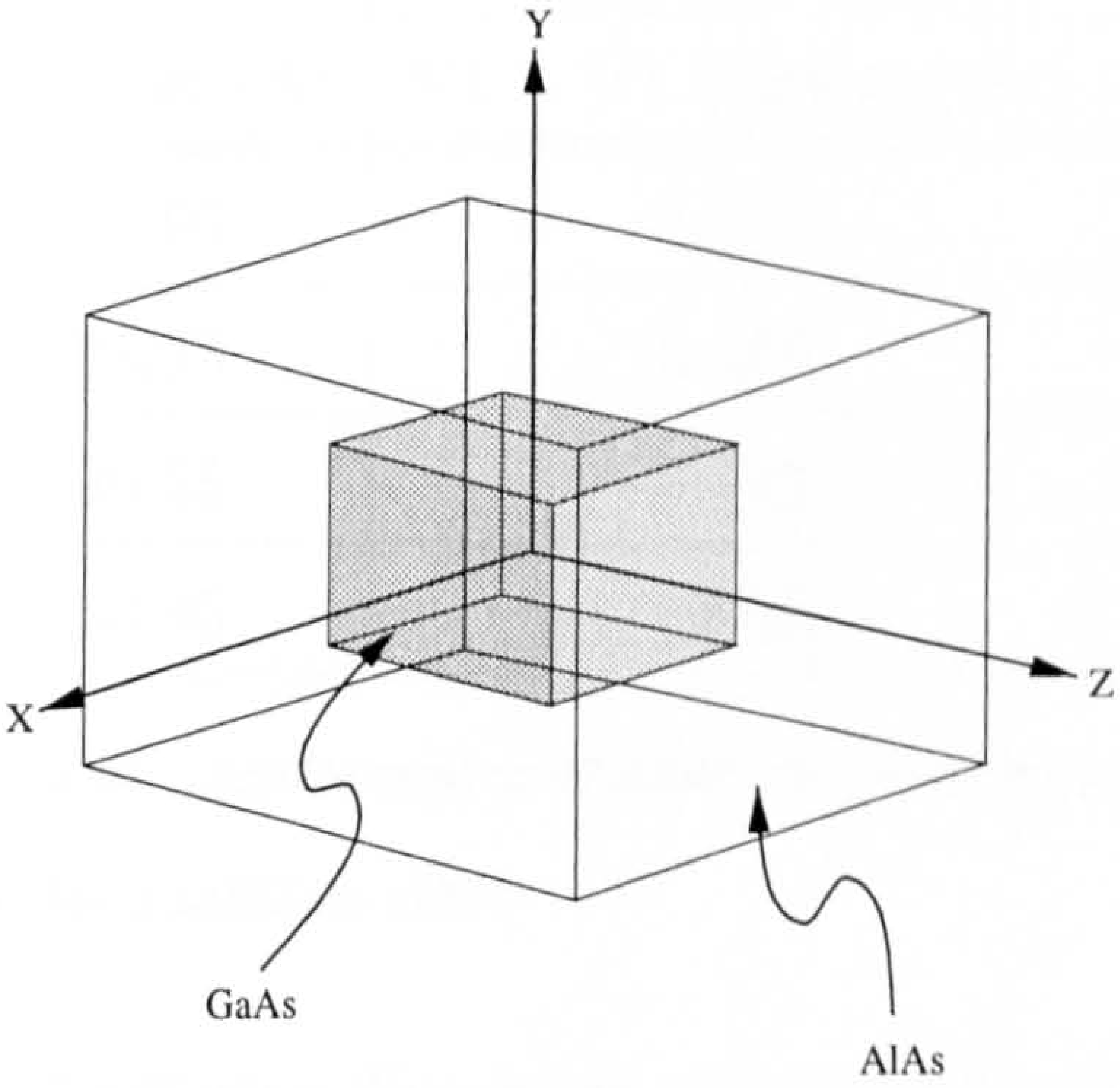

Figure 5.1: The unit cell of a cubic GaAs/AlAs quantum dot together with the axes referred to in the text.

AlAs which is taken to be 1.1 eV from the work of Van der Walle (1989). The envelope functions $F(\mathbf{r})$ are expanded in a set of normalised plane wave states

$$F(\mathbf{r}) = \sum_{\mathbf{g}} A^c(\mathbf{g})\langle \mathbf{r}|\mathbf{g}\rangle \tag{5.3}$$

where $\mathbf{g}$ is a reciprocal lattice vector of the lattice of unit cells, each containing a single cubic quantum dot. Substituting the expression for $F(\mathbf{r})$ into (5.1) and multiplying on the left by an arbitrary plane wave $\langle \mathbf{g}'|$ gives a square matrix of order n_{g}, where n_{g} is the number of plane waves appearing in the expansion (5.3). The matrix elements of the confining potential in this representation are given by (2.60)

$$\langle \mathbf{g}'|V|\mathbf{g}\rangle = \frac{\Omega_d \Delta E_c}{\Omega} \operatorname{sinc}\left(\frac{\mathrm{g}_x l}{2}\right) \operatorname{sinc}\left(\frac{\mathrm{g}_y l}{2}\right) \operatorname{sinc}\left(\frac{\mathrm{g}_z l}{2}\right) \tag{5.4}$$

where $\Omega_d = l^3$ is the volume of the dot and Ω is the volume of the GaAs/AlAs unit cell. Diagonalisation of the Hamiltonian matrix yields the conduction band energy levels and the coefficients $A^c(\mathbf{g})$ of the envelope functions.

The conduction band energies and envelope functions of four GaAs/AlAs

Structure	Dot size (Å)	V1→ C1 Energy (eV)	$\|\hat{e} \cdot p\|^2$ (a.u.)
1	16.95	2.976	0.186
2	28.25	2.495	0.408
3	39.55	2.142	0.428
4	50.85	1.935	0.434

Table 5.1: Dot dimensions, fundamental transition energies and oscillator strengths of four cubic GaAs/AlAs quantum dots.

cubic quantum dots of differing dimensions have been calculated. A basis consisting of 729 plane waves (9 in each direction x, y, z) was adequate to ensure convergence of the energy levels with respect to the confining potential. Each dot lay at the centre of a cubic barrier of AlAs of side $13a_0$. The dimensions of the four GaAs dots are given in the second column of Table 5.1. In Figure 5.2(a) the bound conduction band levels are plotted against dot size. The dashed line represents the position of the AlAs conduction band edge at Γ. Each state shown is doubly degenerate through spin. In addition to the spin degeneracy a number of other degeneracies can be seen to exist among the excited states. For example, the first excited conduction state exhibits a six-fold degeneracy including spin.

The valence band levels in GaAs/AlAs quantum dots of differing sizes were found using the multi-band effective mass procedure detailed in Section 2.2.3. The effective Hamiltonian for the valence band states was expanded in a basis of 729 plane waves per bulk band. The energy levels of the valence band states are given in Figure 5.2(b) as a function of dot size. The dashed line represents the position of the AlAs valence band edge. The confining potential for the hole carriers was assumed to be equal to the valence band offset between GaAs and AlAs (500 meV). From Figure 5.2(b) it is evident that the ground hole

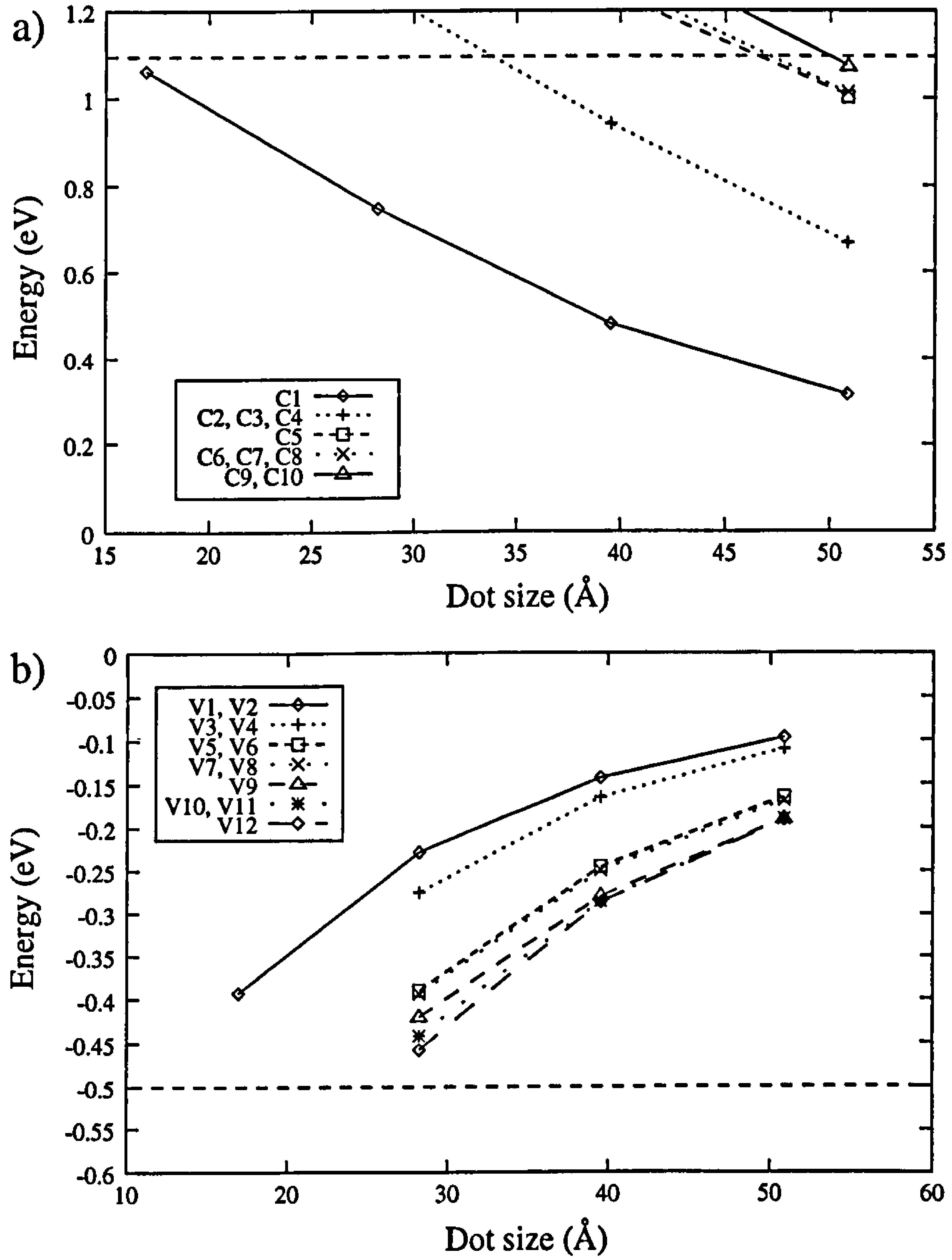

Figure 5.2: The energies of the confined states of cubic GaAs/AlAs quantum dots displayed as a function of GaAs dot size. (a) The conduction levels. (b) The valence levels. The dashed line represents the position of the AlAs conduction (valence) band edge.

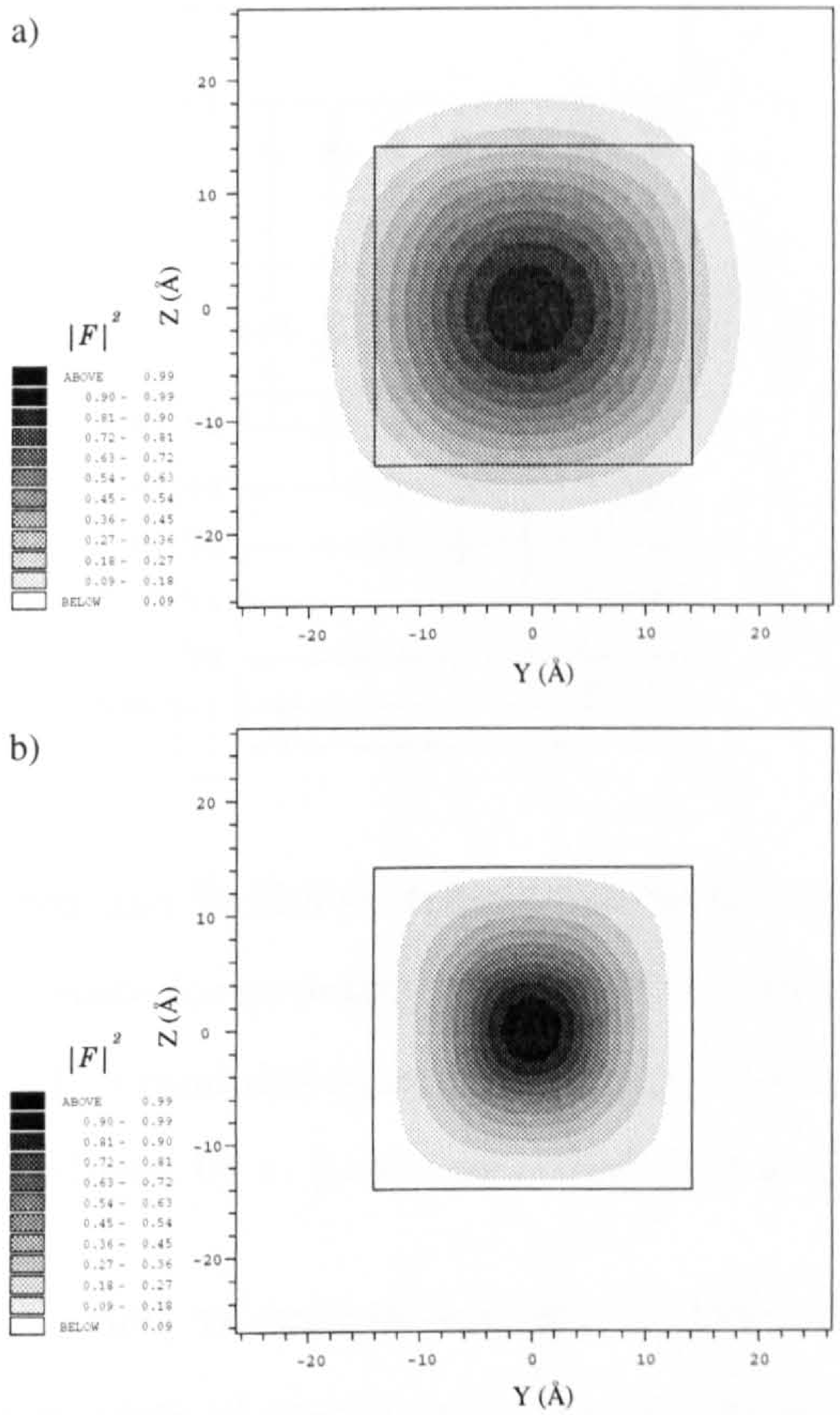

Figure 5.3: The modulus-squared envelope function of the ground conduction and valence states in structure 2 in Table 5.1. (a) State C1. (b) States V1, V2.

states in the dot are four-fold degenerate (including spin degeneracy). This is to be expected in dots of cubic or spherical symmetry due to the degeneracy of the bulk heavy hole and light hole states ($J = 3/2$) at Γ (Sercel and Vahala, 1990).

In Figure 5.3 the modulus-squared envelope function for the ground electron state and the ground hole state of structure 2 are plotted in a plane through the centre of the quantum dot. In Figure 5.3(b) the modulus-squared envelopes for each member of the quadruply degenerate set of states is summed. The form of the envelope function indicates that the distribution of charge in

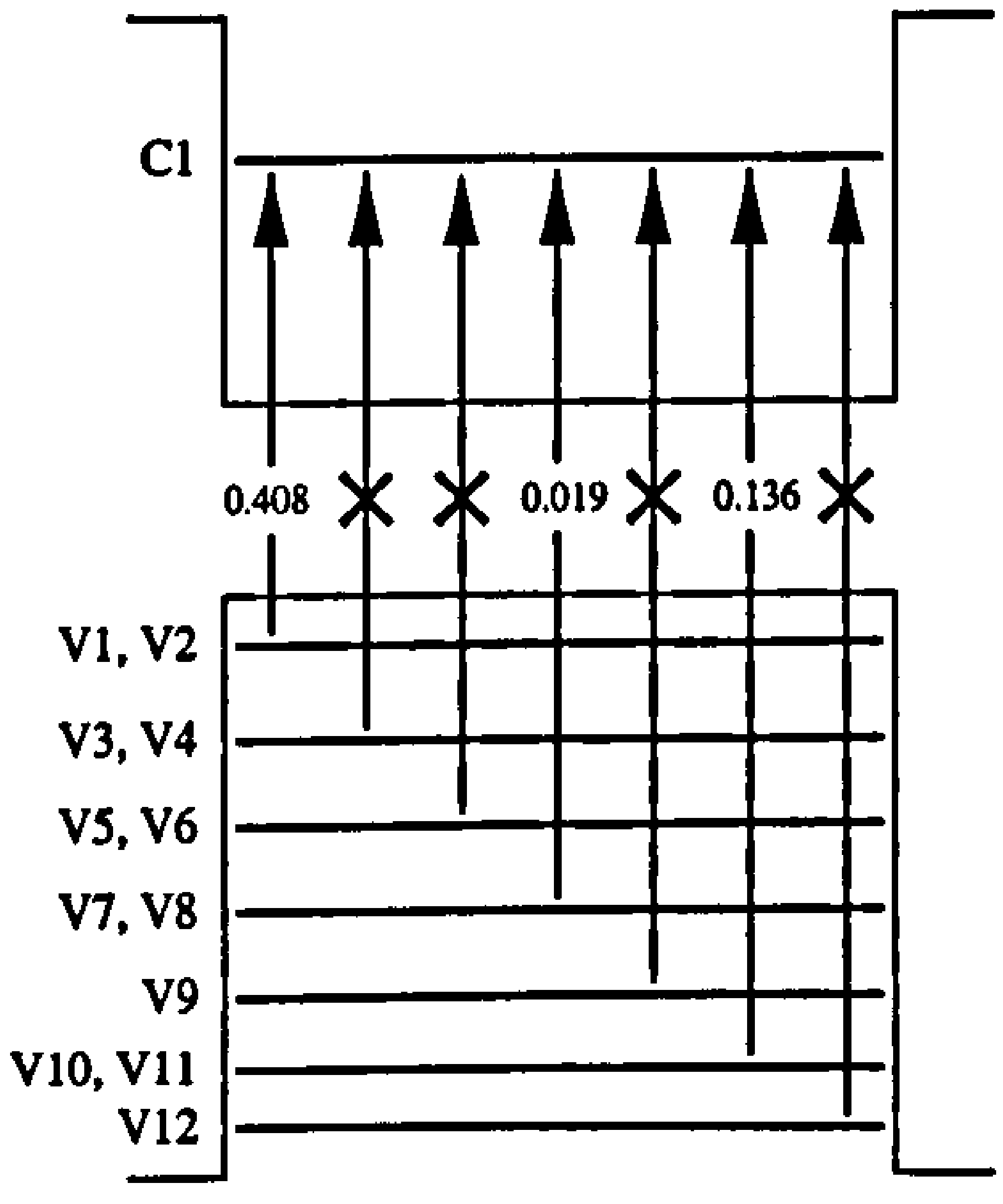

Figure 5.4: The allowed and forbidden transitions between the valence states and the ground conduction state for structure 2 in Table 5.1. Forbidden transitions are denoted with a cross. The modulus-squared of the momentum matrix element for an allowed transition induced by x, y or z polarised light is given in atomic units.

the ground conduction and valence states reflects the cubic symmetry of the dot. The low effective mass of electrons in GaAs results in a greater leakage of charge into the AlAs barrier than is predicted for hole carriers.

The modulus-squared of the momentum matrix elements for different polarisations of light between the ground conduction state (C1) and the bound valence band states (Vn) were calculated using (2.64). The results are summarised in Figure 5.4. In the figure, forbidden transitions are denoted with a cross. The allowed transitions indicate the square of the magnitude of the momentum matrix element for light polarised along the x, y or z axis. Because of the cubic symmetry, the square of the optical matrix elements induced by different polarisations of incident light are equal. In the third and fourth column of Table 5.1 the energy and the square of the momentum matrix element of the fundamental transition are given for the four structures studied. The strength of the transition can be seen to increase as the dot size increases due

State	Energy (eV)	$Vn \rightarrow C1$ (eV)	$\lvert\hat{e}\cdot p\rvert^2$ (a.u.)		
			x	y	z
HH1	-0.328	2.595	0.451	0.451	0
LH1	-0.332	2.599	0.153	0.153	0.614
HH2	-0.468	2.735	0	0	0

Table 5.2: Energies and momentum matrix elements for the $Vn \rightarrow C1$ transition in the absence of valence band mixing for structure 2 in Table 5.1.

to the greater overlap of the envelope functions for the ground conduction and valence states. The major change to the oscillator strength between structures 1 and 2 can be attributed to the fact that the C1 level is only weakly bound in the well [see Figure 5.2(a)] with the corresponding envelope function leaking considerably into the surrounding barrier material greatly reducing the overlap between the ground states.

The valence band states of structure 2 were also calculated with the R and S terms in the Luttinger-Kohn Hamiltonian matrix (2.42) set to zero (hereafter termed the diagonal approximation). In the diagonal approximation, the quantum dot states possess a heavy hole or light hole origin; they are not an admixture of both as they are in the full calculation which features nonzero R and S terms. The energy levels of the confined states for structure 2 in the absence of valence band mixing are given in the second column of Table 5.2. The levels are given a heavy hole (HH) or light hole (LH) classification. In the absence of valence band mixing, fewer bound states are predicted. Furthermore, the quadruple-degeneracies obtained in the presence of mixing are absent. The optical matrix elements between the ground valence and conduction band states without valence band mixing (in Table 5.2) also expose the inadequacies of the diagonal approximation. While the magnitude of the matrix elements for the HH1→C1 transition given in Table 5.2 is similar to the value given for the same transition for structure 2 in Table 5.1, the transition due to z polarised light is forbidden in the absence of valence band mixing.

5.2 InAs/GaAs Self-Assembled Quantum Dots

The nanometre-sized pyramids of InAs that form during the interrupted MBE growth of InAs on a GaAs substrate confine charge in three dimensions, quantising their motion in the process. The deposition of InAs on GaAs (001) proceeds first by two-dimensional (2D) growth and above a 1.7 ML coverage (Leonard *et al.*, 1994) by the formation of dislocation-free dots on a residual 2D wetting layer (Figure 5.5). It has been demonstrated (Solomon *et al.*, 1995) that the island sizes and areal densities can be controlled by varying growth parameters such as the thickness of the initial two-dimensional layer deposition, and the growth rate.

The islanding occurs in an attempt to relieve the strain in the substrate and 2D layer due to the 7% lattice mismatch between InAs and GaAs. Transmission electron microscopy measurements (Grundmann *et al.*, 1995c) have revealed pyramidal shaped dots (12 ± 1 nm base along $\langle 100 \rangle$, 5 ± 1 nm high, and side facets close to $\{011\}$) with an areal density of 10^{11} dots per square centimetre. The measurements also indicate that the quantum dots exhibit short range order, aligning along rows in the $\langle 100 \rangle$ directions. Further studies (Leonard *et al.*, 1993) have shown the 3D islands to be remarkably uniform with a size distribution $\sim$ 10%. This might make them suitable for optoelectronic applications. Islanding has also been observed during the strained heteroepitaxial growth of other materials, including SiGe on Si (Apetz *et al.*, 1995), InP on InGaP (Petroff and Denbaars, 1994), and InAs on AlAs (Leon

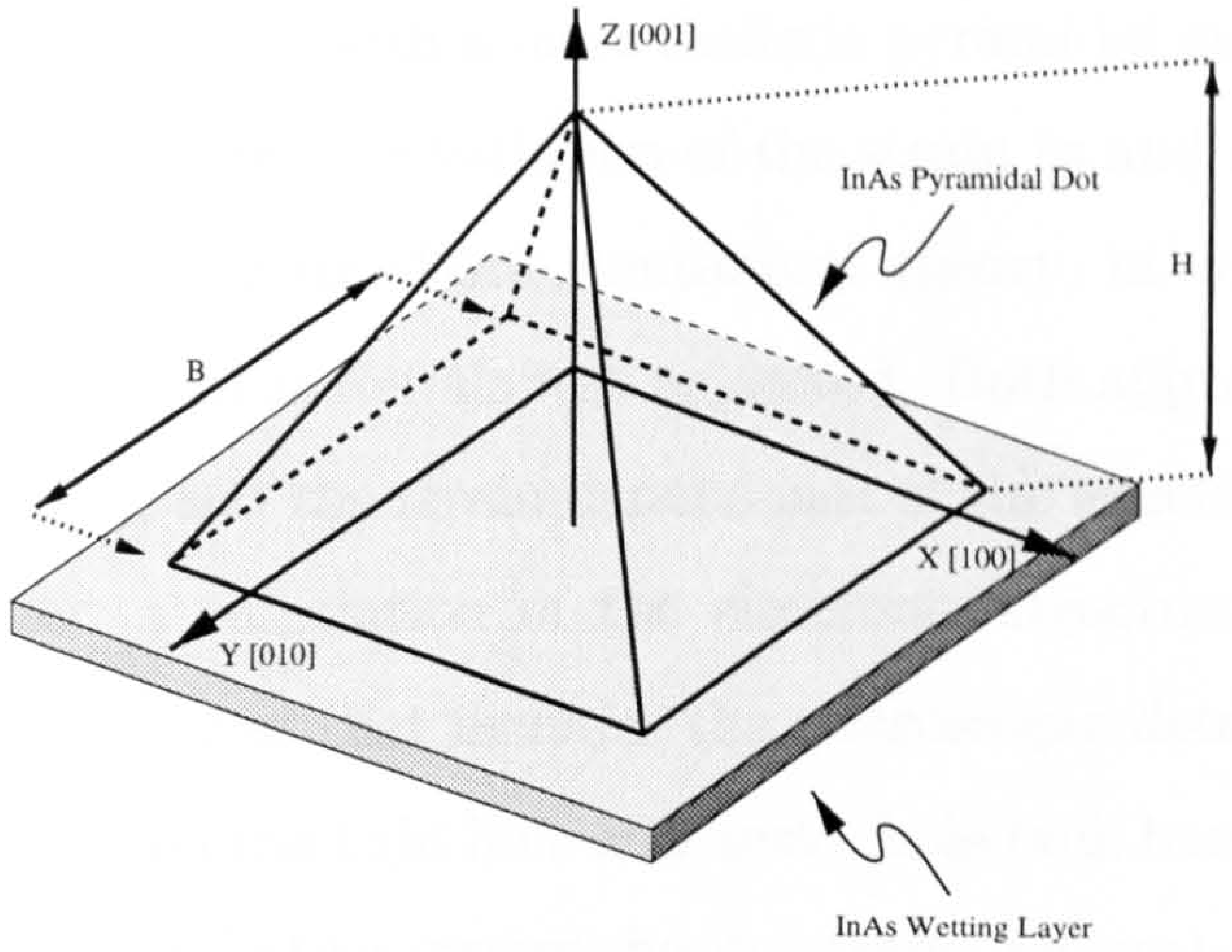

Figure 5.5: Schematic diagram of an InAs quantum dot and InAs wetting layer together with the axes referred to in the text.

et al., 1995).

The prevailing theory (Priester and Lannoo, 1995) concerning the origin of the islands states that stable 2D platelets of InAs, each containing similar numbers of atoms, form on the surface of the InAs 2D layer. These reach their equilibrium distribution and, at the critical coverage, their repulsive interaction induces a spontaneous transformation into 3D islands which keep the same size distribution and possess no misfit dislocations. These platelets can be thought of as "precursors" for the formation of 3D islands. A further increase in InAs coverage results in a subsequent growth of these islands which, at some critical size, can lead to the creation of misfit dislocations and finally to coalescence.

The first theoretical study of the electronic properties of these structures (Marzin and Bastard, 1994b) used the single-band effective mass theory to calculate the energy levels and wave functions in InAs/GaAs cone-shaped quantum dots. The strain was taken to be a constant in the InAs material and zero in the surrounding GaAs barrier. More recently, the single-band theory was applied to InAs/GaAs dots with a more realistic pyramidal geometry (Grundmann *et al.*, 1995b). There the variation of the strain in and around the InAs island was determined using elastic continuum theory, in which the atomic nature of the constituent materials was neglected. Both approaches neglected valence band mixing, and the strain-dependence of the effective masses.

In this chapter, a calculation of the electronic structure of InAs/GaAs quantum dots is presented that includes the microscopic details of the strain and the mixing between the light hole and heavy hole bulk bands and accounts for the change in the effective masses due to strain. Optical matrix elements for the dominant transitions in dots of differing sizes are obtained and a picture of the charge localisation throughout the pyramid is provided.

Firstly, however, we must investigate the physical structure (atomic positions and strain distribution) of InAs/GaAs self-assembled structures so that the strain-modified confining potential required as input to the electronic structure calculation can be determined. The procedures for achieving this objective are detailed in the following sections.

5.2.1 Atomic Positions

The positions of the atoms in the InAs islands and the surrounding GaAs layer are determined using a molecular mechanical model. In this approach the atoms are considered as point particles and the bonds as springs. The mathematics of spring deformation is used to describe the ability of the bonds to stretch and bend. The object of molecular mechanics is to predict the energy associated with a given conformation of a dot then attempt to find another conformation with a lower energy. When this process is repeated many times and an energy minima is reached, we can assume we have obtained a good approximation to the true conformation. A simple energy equation involves the sum of the total energy associated with the stretching of bonds and the total energy associated with the bending of bonds. These total energies have no meaning as absolute quantities, only differences between two or more conformations have meaning. The total energy equation together with the data (parameters) required to describe the behaviour of different kinds of atoms and bonds constitutes a force field. Many different kinds of force field have been developed over the years. Some include additional energy terms that describe other kinds of deformations. Some force fields account for coupling between bending and stretching in adjacent bonds, bond twisting terms and non-bonded interactions (e.g. Van der Waals) in order to improve the accuracy of the model.

The choice of force field depends, of course, on the type of bonding present in the material under consideration. In the III-V semiconductors considered here, which display covalent bonding with partial charge transfer, the valence force field (VFF) is most suitable. We begin by writing the total potential energy as the sum over the atoms in the dot and barrier materials

$$E_{tot} = \sum_{i}^{d} V_i^d + \sum_{i'}^{b} V_{i'}^b \qquad (5.5)$$

where V^d and V^b represent the potential of an atom of the dot and barrier material respectively. The potential assigned to the ith atom of the dot material may be written as (Briddon, 1990)

$$\begin{aligned} V_i^d = {} & \frac{1}{4}\sum_{j=1}^{n} f_r^d(\Delta r_{ij})^2 + \frac{1}{2} f_\theta^d (r_0^d)^2 \sum_{j=1}^{n-1}\sum_{k=j+1}^{n} (\theta_{ijk})^2 + f_{rr}^d \sum_{j=1}^{n-1}\sum_{k=j+1}^{n} \Delta r_{ij}\Delta r_{ik} \\ & + r_0^d f_{r\theta}^d \sum_{j=1}^{n-1}\sum_{k=j+1}^{n} (\Delta_{ij} + \Delta r_{ik})\Delta\theta_{jik} \\ & + (r_0^d)^2 f_{\theta\theta}^d \sum_{j=1}^{n-2}\sum_{k=j+1}^{n-1}\sum_{l=k+1}^{n} (\Delta\theta_{jik}\Delta\theta_{kil} + \Delta\theta_{kil}\Delta\theta_{lij} + \Delta\theta_{lij}\Delta\theta_{jik}) \qquad (5.6) \end{aligned}$$

where $\Delta r_{ij} = r_{ij} - r_{ij}^0$, $\Delta\theta_{jik} = \theta_{jik} - \theta_{jik}^0$ and r_{ij}^0 and θ_{jik}^0 are the equilibrium

	f_r (eV/Å^2)	f_θ (eV/Å^2)	f_{rr} (eV/Å^2)	$f_{r\theta}$ (eV/Å^2)	$f_{\theta\theta}$ (eV/Å^2)
GaAs	7.9689	0.3071	0.4350	0.1813	-0.0028
InAs	6.8812	0.1701	0.2222	0.1520	-0.0259

Table 5.3: Valence force field parameters for GaAs and InAs.

values of the length of the bond between atoms i and j and the angle subtended by atoms j and k at atom i. n is the coordination number. The summations have all been symmetrised and the factors of (1/2) and (1/4) inserted to avoid double counting. The force constants, f^d, characterise the strength of the different interactions that make up the valence force field for the dot material. The parameter f_r^d is the bond stretching force constant, f_{rr}^d accounts for the coupling between bonds sharing a common atom, $f_{r\theta}^d$ describes the strength of the coupling between an angle and a bond along one leg of θ and $f_{\theta\theta}^d$ describes the correlation between angles having a common leg and apex. A similar set of force constants exists for the barrier material.

In the past the VFF method has been used to determine a range of elastic and vibrational properties of covalently bonded solids. Musgrave and Pople (1962) and McMurry *et al.* (1967) carried out VFF calculations of the phonon spectrum of diamond. Martin (1970) investigated the elastic properties of ZnS using a VFF model which accounted for charge transfer through the inclusion of a Coulomb term. Here we are interested in using the VFF model to determine the positions of atoms in the dot and barrier materials, extracting from this, information concerning the distribution of strain in and around the islands. To this end, we choose to fit the five force constants for the dot material to the phonon spectrum of InAs from the experimental work of Charles *et al.* (1980). Likewise, the five force constants pertaining to GaAs were fitted to the phonon spectrum measured by Strauch and Dorner (1990).

Structure	B (Å)	H (Å)	C1 → V1 (eV)	$\lvert p^{x(y)}\rvert^2$ (a.u.)
1	57	28	1.43	0.017
2	79	40	1.33	0.316
3	102	51	1.23	0.374
4	124	62	1.11	0.400

Table 5.4: Dimensions, fundamental transition energy and C1 → V1 momentum matrix elements for $x(y)$ polarised light belonging to the four structures studied. Across-gap transitions induced by light polarised along the z direction in the quantum dot are forbidden.

The transfer of charge between the two atomic species in materials such as InAs and GaAs should be taken into account through an additional term in (5.6) representing the extra interatomic force resulting from the placement of a point charge on each of the atoms. The resulting electrostatic interaction is the cause of the TO-LO splitting observed at $\mathbf{q} = 0$ in the phonon spectra of these materials. This additional Coulomb term was neglected and instead the five force constants adjusted such that the phonon spectrum in the neighbourhood of $\mathbf{q} = 0$ was well reproduced and agreed with the values predicted from the elastic constants of the material. Although the TO-LO splitting is not recovered during the fitting procedure, the low frequency part of the spectrum was accurately obtained. By this method, the effects of the Coulomb term on the acoustic branch of the spectrum in the neighbourhood of $\mathbf{q} = 0$ were encoded into the force constants. This was important as it is primarily the longer-wavelength variation in strain that is extracted from this part of the calculation and used in the electronic structure determination described later on. The force constants used in the structure calculation are listed in Table 5.3.

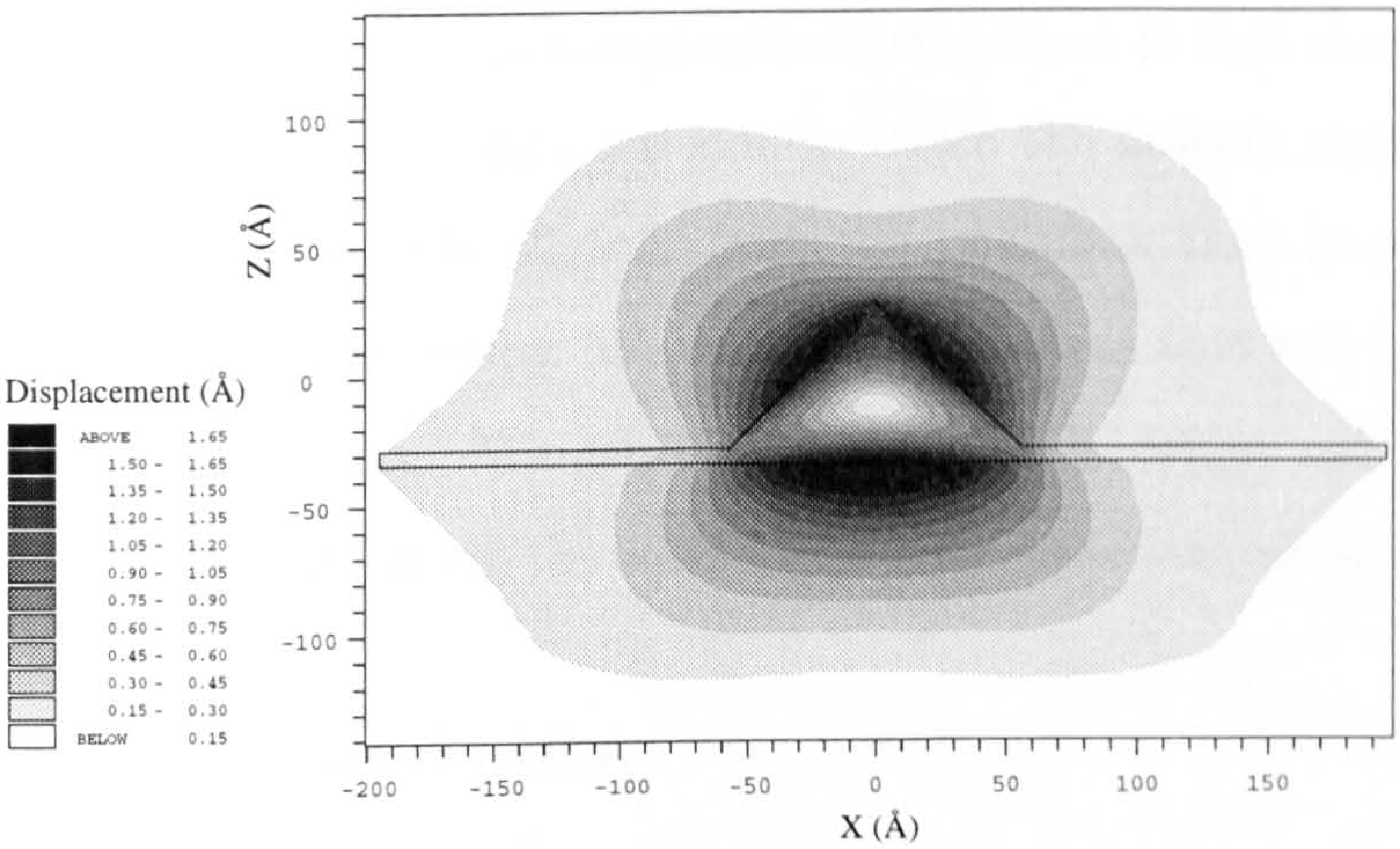

Figure 5.6: The absolute displacement of anions from their original positions in the XZ plane of the InAs/GaAs quantum dot due to the relaxation procedure. The outline of the InAs pyramid and wetting layer is also shown.

The calculation of the atomic positions in the InAs dot and GaAs barrier proceeded by placing every atom in a unit cell containing a single island and wetting later section surrounded by GaAs, at the atomic positions that would be occupied by Ga or As atoms in a similar sized unit cell of pure GaAs. The total energy of this conformation was calculated, together with the net force on each atom. Each atom was moved a distance in the direction of the net force and then the total energy was recalculated. This procedure was repeated until the total energy reached a minimum.

The atomic positions of four strained InAs/GaAs self-assembled quantum dot systems containing differing sizes of InAs islands have been determined by the above method. In each structure, the unit cell contained 2×10^6 atoms. The dimensions of each dot are given in Table 5.4. The axes used in this study are defined in Figure 5.5. Each pyramid lies on a 1.5 ML InAs wetting layer. In Figure 5.6 the magnitude of the displacement of the As atoms from their positions prior to relaxation is shown across the XZ plane of structure 4 in

Table 5.4. The outline of the pyramid and wetting layer is also shown. In the centre of the pyramid there exists a region in which the arsenic atoms bonded to indium are displaced only slightly from the positions that arsenic atoms bonded to gallium would occupy. We can anticipate that this will be a region under hydrostatic pressure. We can also expect the interface regions at the base of the pyramid and along the sides to be under the largest strain.

5.2.2 Strain Distribution

Bulk conventional cubic unit cells in the structure were assigned a strain tensor on the basis of the displacement of the eight corner arsenic atoms from their positions prior to relaxation. In Figure 5.7 a typical conventional cubic unit cell is shown before and after the relaxation process with the deformed and undeformed unit basis vectors labelled. Using (2.45) it is straightforward to generate the six nonzero independent strain tensor components listed in (2.46), for a deformed unit cell. When this procedure is repeated at every conventional

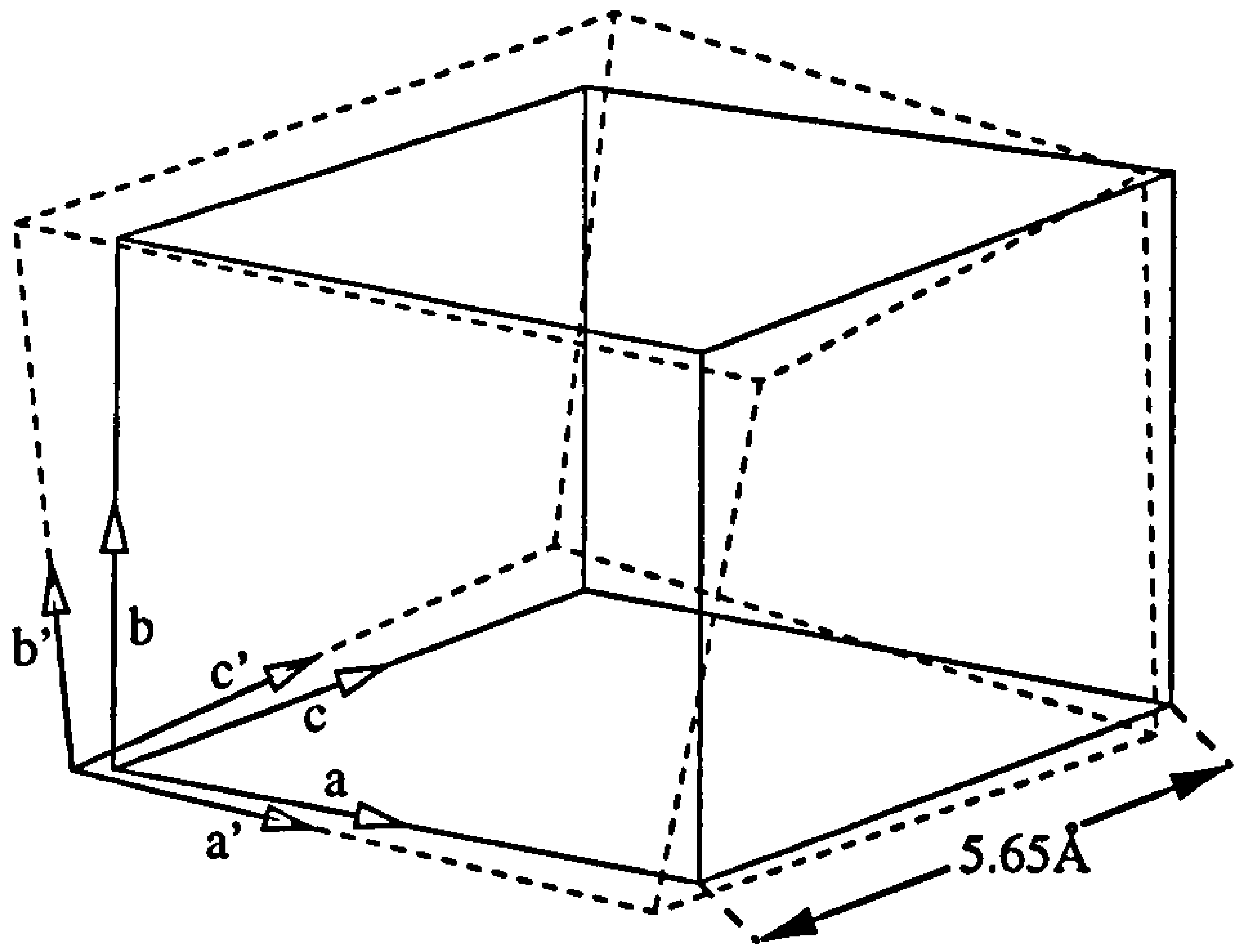

Figure 5.7: A conventional cubic unit cell inside the quantum dot structure before and after deformation.

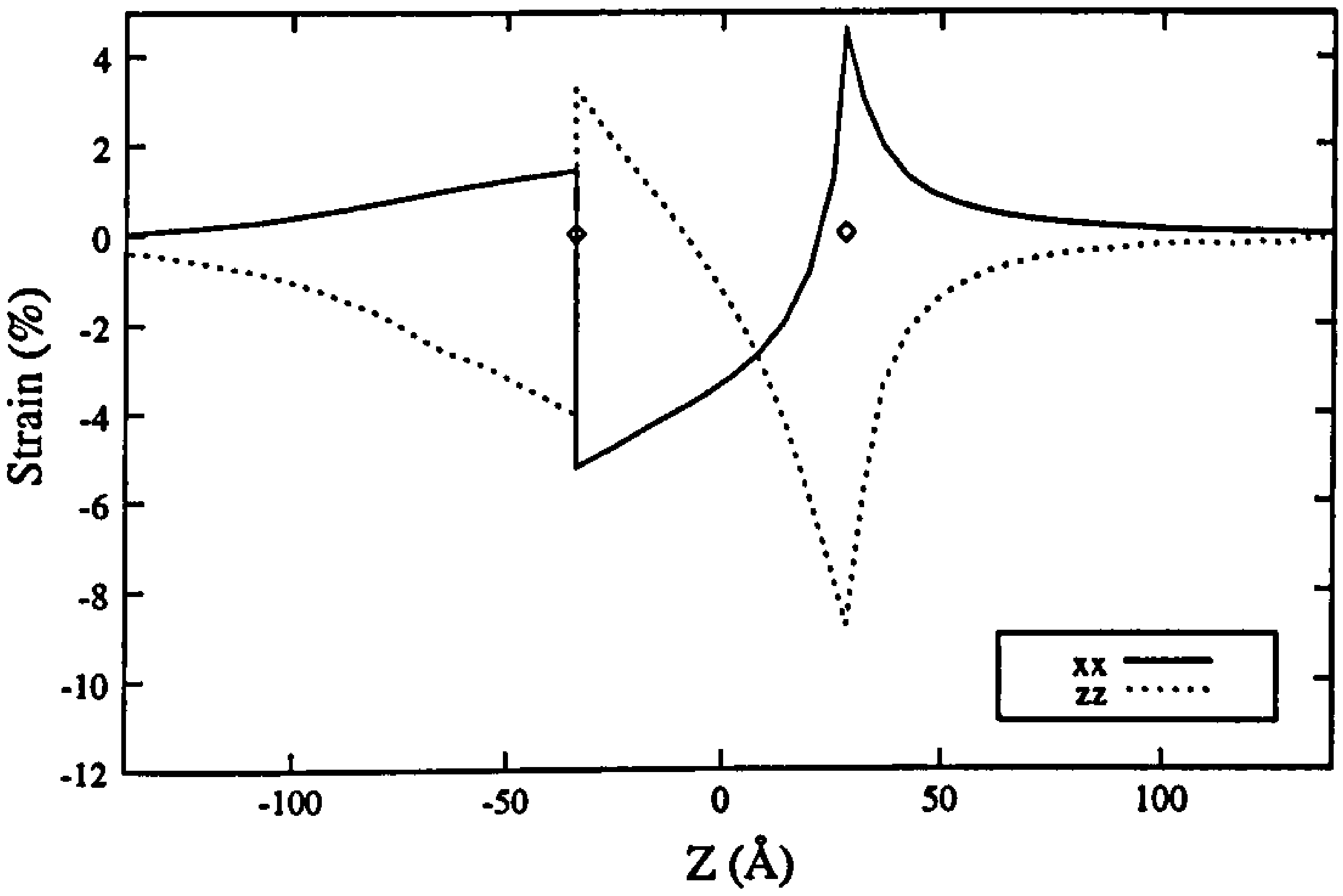

Figure 5.8: Strain tensor components ϵ_{xx} and ϵ_{zz} for structure 4 in Table 5.4 plotted along the Z axis. The diamonds indicate the spatial extent of the InAs dot.

cubic unit cell in the structure a three-dimensional pattern of the strain can be recovered. This method becomes unreliable if the conventional cubic unit cell crosses the interface between the dot and the barrier. This means that the method will be unable to reproduce the strain distribution in the interface region or in the wetting layer. Nevertheless, a good description of the strain away from these regions should be attainable. An approximation to the strain near to the interfaces is obtained using an interpolation procedure.

Figure 5.8 shows the strain tensor components ϵ_{xx} and ϵ_{zz} for structure 4 in Table 5.4 plotted with position along the Z axis in Figure 5.5. The shear strain components, ϵ_{xy}, ϵ_{xz} and ϵ_{yz} are negligible in the dot and barrier although could be appreciable at the interface (Grundmann *et al.*, 1995b). The apex and base of the pyramid are indicated by the diamond-shaped points. At these interfacial points the method for determining the strain becomes unreliable and so interpolation has to be used. By symmetry $\epsilon_{yy} = \epsilon_{xx}$. In the substrate layer below the dot we see that the GaAs lattice experiences a tensile stress in the XY plane (positive ϵ_{xx} strain) and a compressive stress in the Z direction

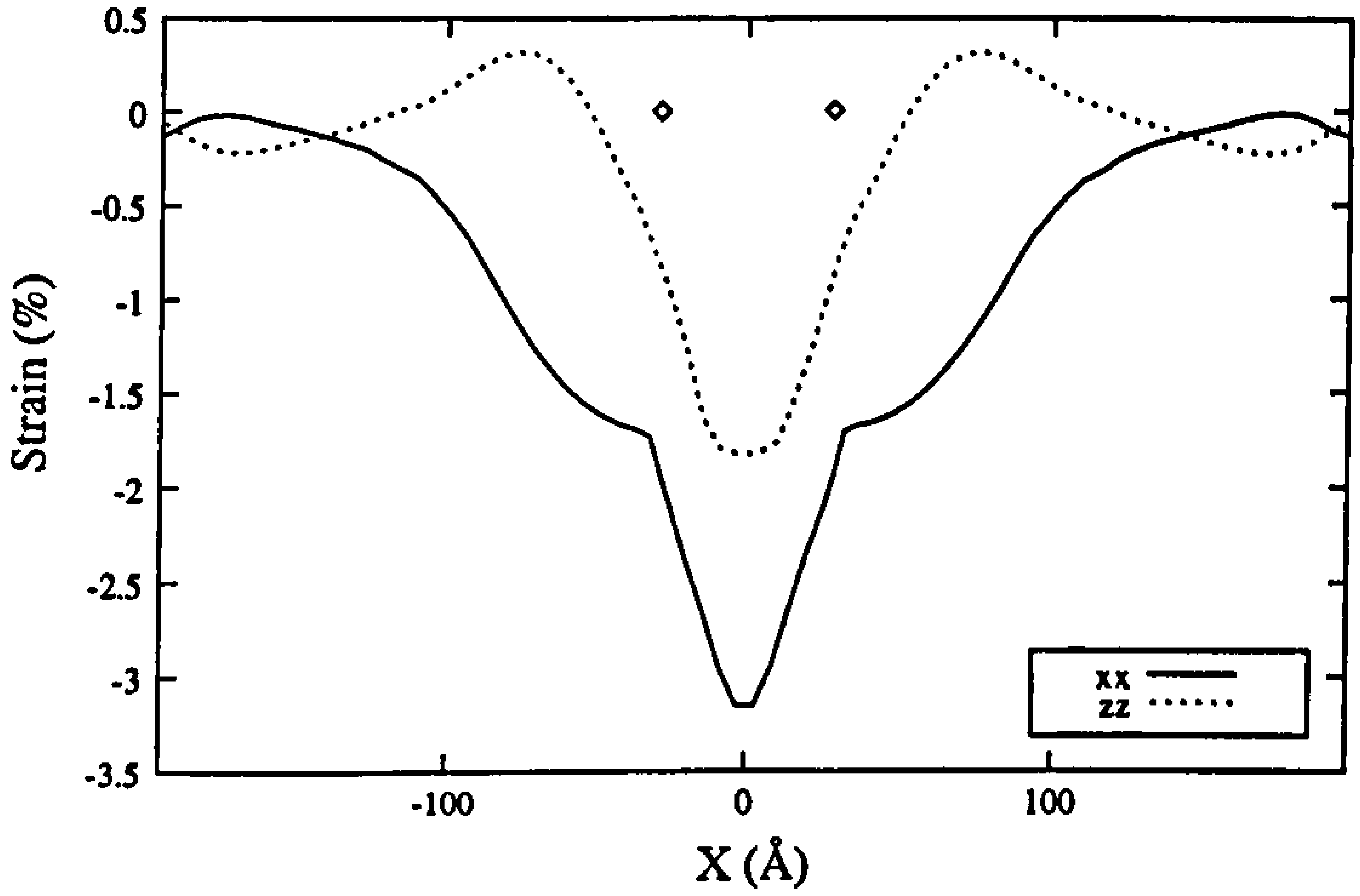

Figure 5.9: Strain tensor components ϵ_{xx} and ϵ_{zz} for structure 4 in Table 5.4 plotted along the X axis. The diamonds indicate the spatial extent of the InAs dot.

(negative ϵ_{zz}). This is because the dot is forcing the substrate lattice constant to be that of InAs (6.05 Å). In the base region of the dot, the situation is reversed. Here, ϵ_{zz} is positive and ϵ_{xx} negative because the substrate now attempts to force the dot lattice constant to be that of GaAs (5.65 Å). With increasing height within the dot, ϵ_{zz} changes its sign, becoming negative at the top of the pyramid. At the top of the pyramid the dominant forces acting on the dot originate from the GaAs at the sides, causing a compressive strain along Z (negative ϵ_{zz}) and a tensile strain (positive ϵ_{xx}) in the XY plane. In Figure 5.9 the ϵ_{xx} and ϵ_{zz} components are plotted along the X axis in structure 4. In the dot, the ϵ_{xx} and ϵ_{zz} are both negative implying the existence of a region of almost hydrostatic compression. The ϵ_{xx}, ϵ_{yy} and ϵ_{zz} components of the strain tensor are plotted in the XZ plane which includes the Z axis in Figure 5.10(a-c). The outline of the pyramid is also shown.

The hydrostatic and biaxial components of the strain, defined as

$$\epsilon_h = \epsilon_{xx} + \epsilon_{yy} + \epsilon_{zz} \text{ and } \epsilon_b = 2\epsilon_{zz} - \epsilon_{xx} - \epsilon_{yy} \tag{5.7}$$

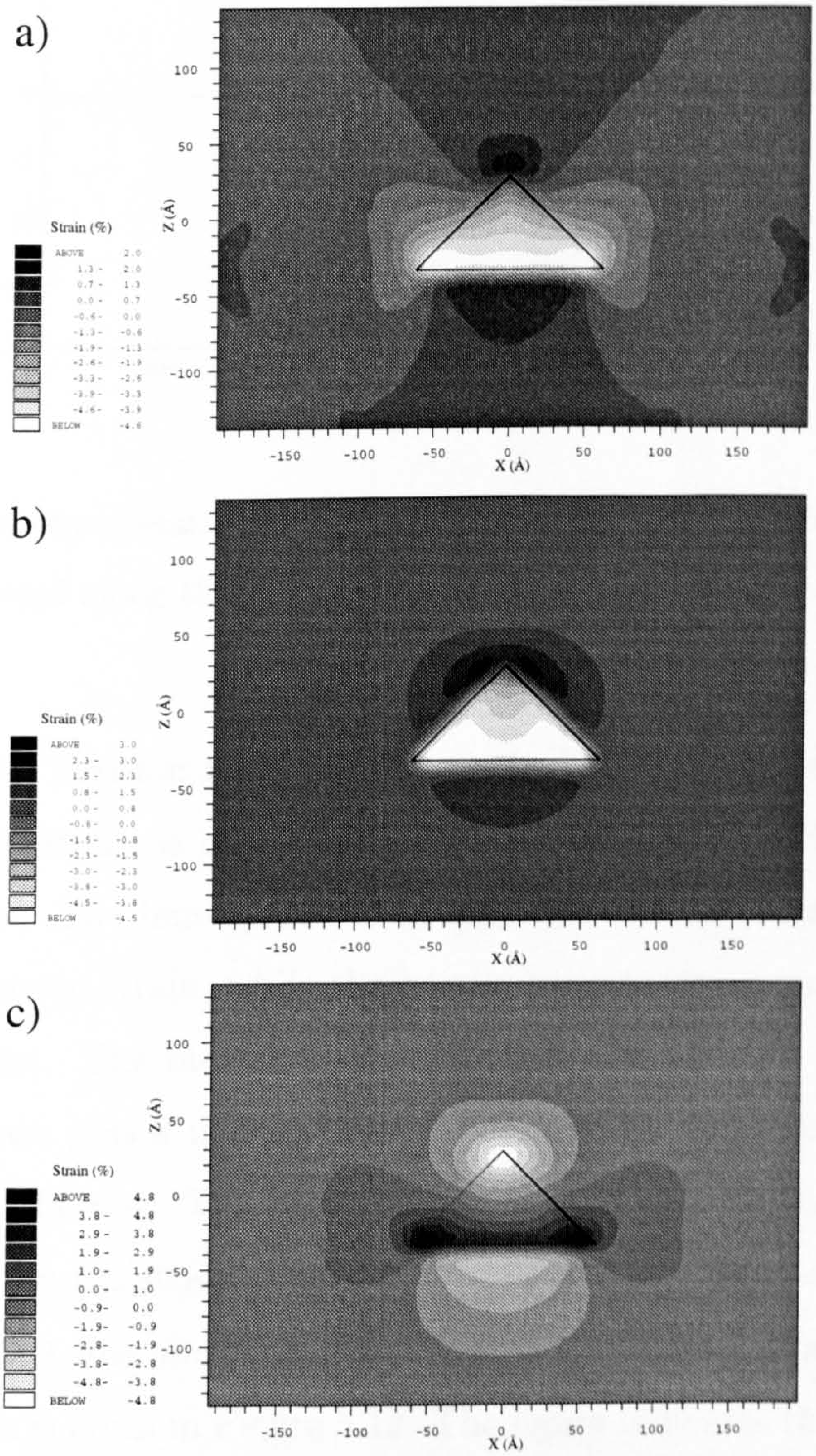

Figure 5.10: Strain tensor components plotted in the XZ plane of structure 4 in Table 5.4. The outline of the InAs pyramid is also shown. (a) ϵ_{xx}. (b) ϵ_{yy}. (c) ϵ_{zz}.

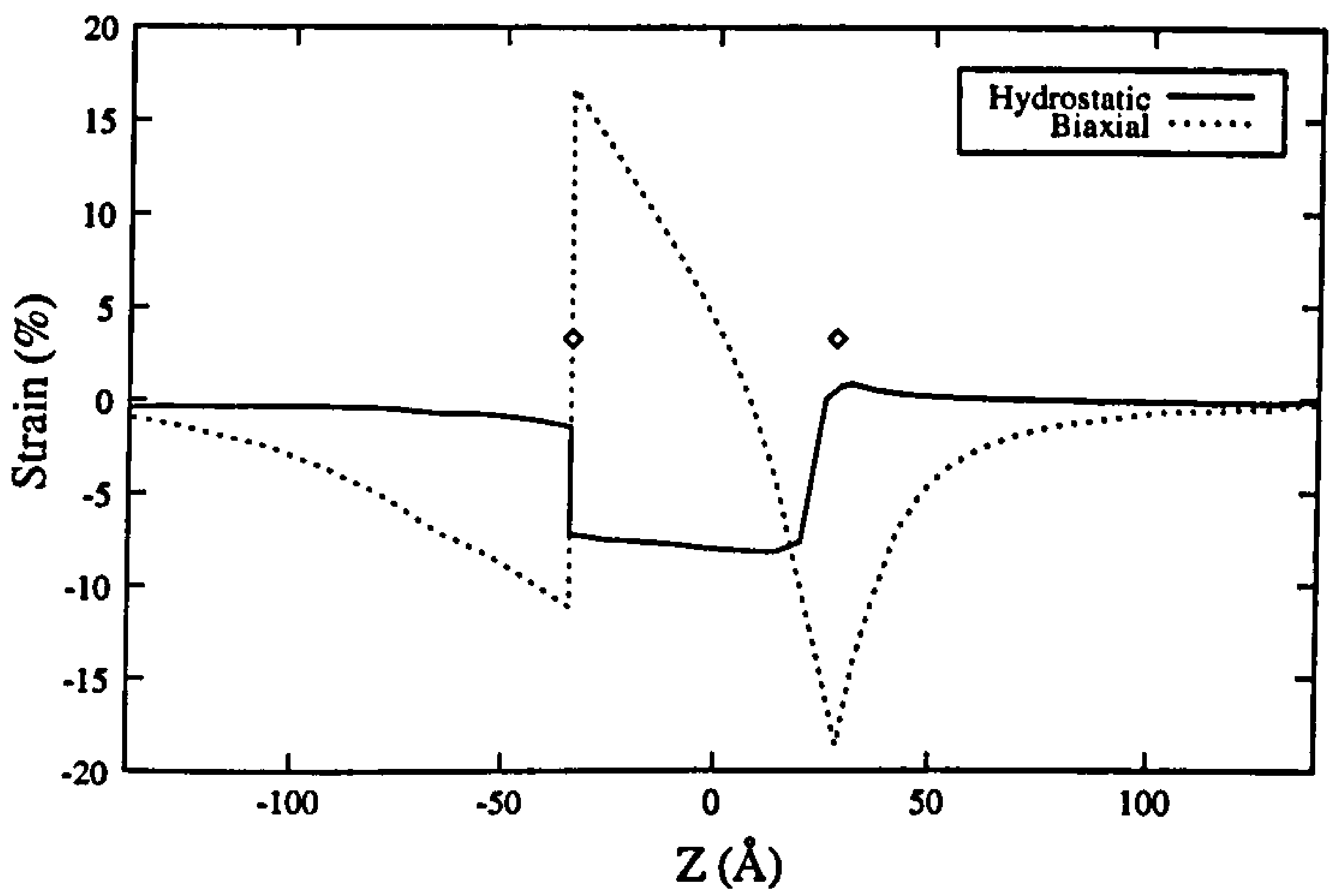

Figure 5.11: The hydrostatic and biaxial components of the strain for structure 4 in Table 5.4 plotted along the Z axis. The diamonds indicate the spatial extent of the InAs dot.

are plotted with position along the Z axis in Figure 5.11 for structure 4. The hydrostatic strain is compressive in both the dot and the surrounding barrier material. The inner part of the quantum dot exhibits almost homogeneous hydrostatic strain, while the barrier experiences a small hydrostatic stress component. The biaxial strain, which can be thought to consist of a hydrostatic strain plus a uniaxial strain, tends to be negative in the barrier and positive in the dot. The biaxial strain is zero near to the centre of the dot. In this region the strain is entirely hydrostatic in character, as predicted from Figure 5.6. A line plot for the hydrostatic and biaxial strain components along the X axis is given in Figure 5.12. The figure indicates that a significant transfer of the biaxial strain to the barrier takes place.

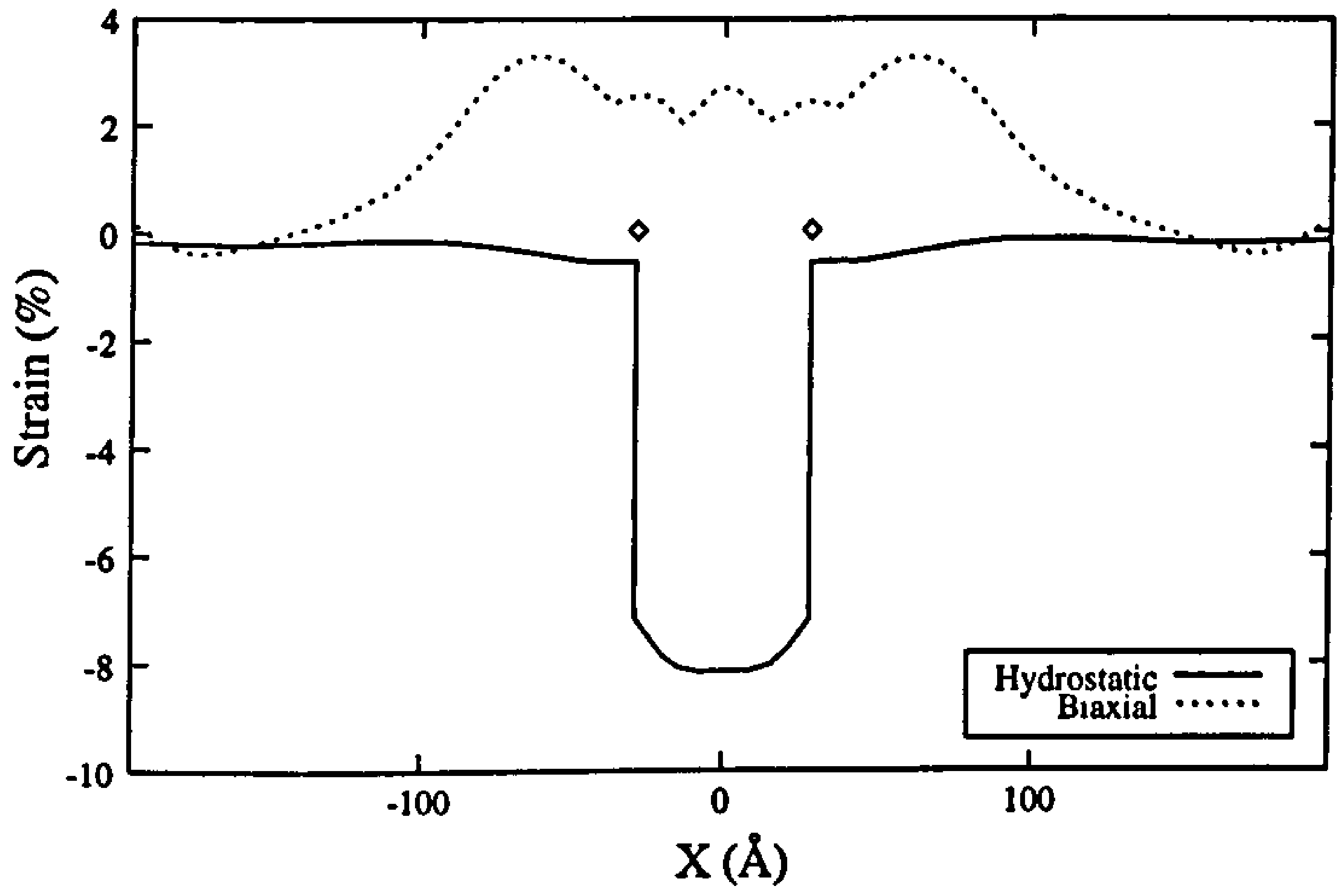

Figure 5.12: The hydrostatic and biaxial components of the strain for structure 4 in Table 5.4 plotted along the X axis. The diamonds indicate the spatial extent of the InAs dot.

5.2.3 Electronic Structure and Optical Transitions

In the absence of strain effects, the confining potential for an electron (hole) is a square well formed by the difference in the absolute energy of the conduction (valence) band edges in InAs and GaAs given by Van der Walle (1989). In the conduction band the depth of the confining potential in the absence of strain effects is 840 meV. For hole carriers the potential well is 263 meV deep. In each conventional cubic unit cell in the dot and barrier, the confining potential for each carrier type is shifted due to the strain. Since the strain varies from cell to cell, the confining potentials will also vary from cell to cell. Furthermore, degeneracies in the valence band edge will be lifted due to deviations of the unit cells from cubic symmetry. For each unit cell, the strain-induced shifts to the confining potentials are obtained by diagonalising the 8 × 8 Hamiltonian matrix (2.54). Hence, the confining potentials including the effects of strain are piecewise continuous functions of position.

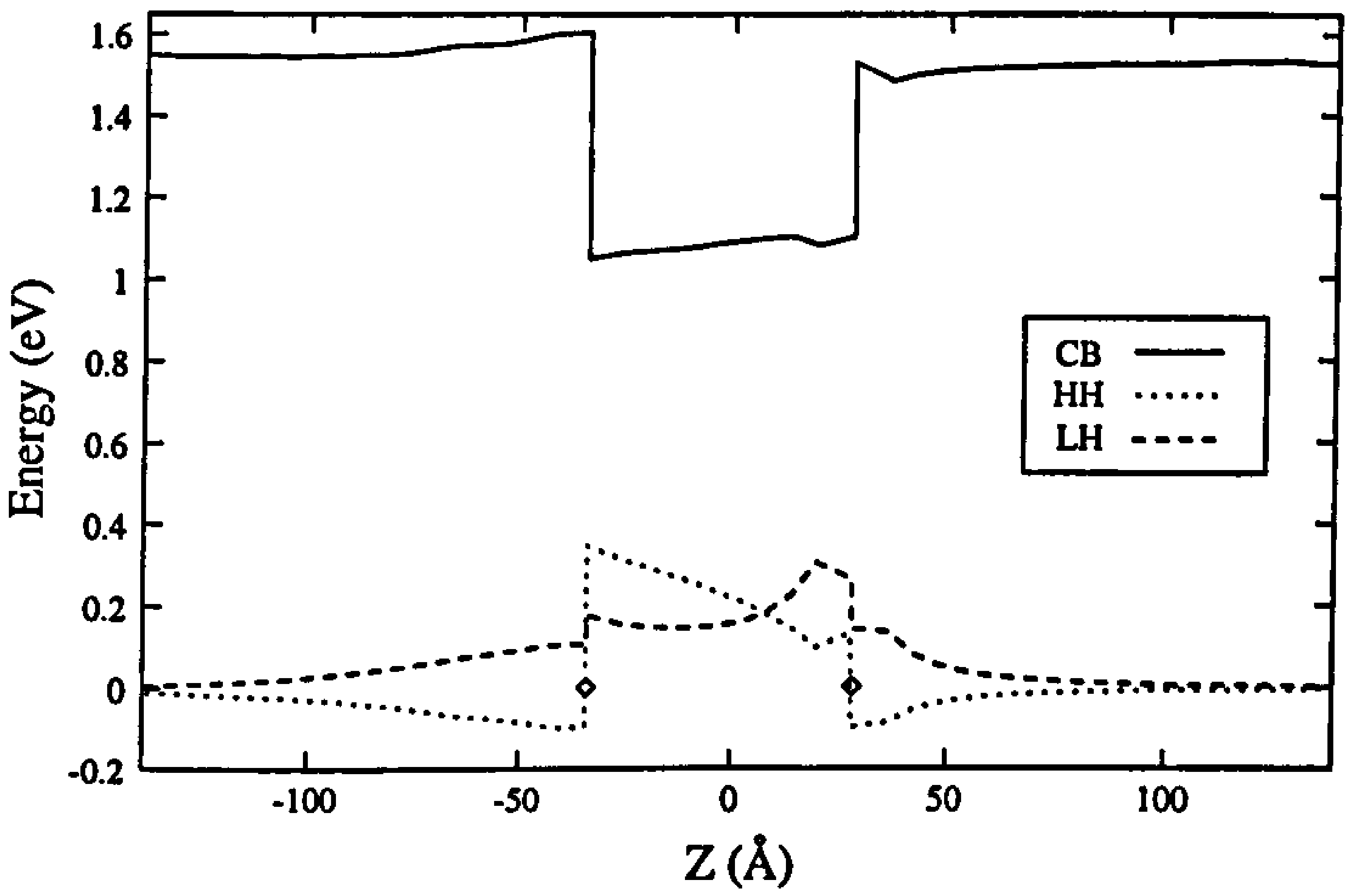

Figure 5.13: The electron, heavy hole and light hole potential profiles for structure 4 in Table 5.4 plotted along the Z axis. The diamonds indicate the spatial extent of the InAs dot.

In Figure 5.13 the confining potentials for electrons and holes are shown with the modification due to strain included. The split-off band edge is far in energy from the heavy hole and light hole band edges and plays no part in the calculation. The heavy hole and light hole potential wells are shown inverted. In the absence of strain, the heavy hole and light hole confining potentials would be the same. However, the anisotropic (biaxial) components of the strain in the dot and barrier reduce the symmetry of the conventional cubic unit cells, lifting the heavy hole-light hole degeneracy.

The compressive strain in the barrier shifts the GaAs conduction band edge slightly above the unstrained level (at 1.52 eV). From Figure 5.13 it is clear that the light hole band edge is higher in energy than the heavy hole band edge in the barrier and towards the apex of the pyramid. The heavy hole band is the uppermost band at the base of the pyramid. The direction and magnitude of the splitting of the light and heavy hole bands - in the absence of appreciable shear strain components - is dependent solely on the magnitude and sign of

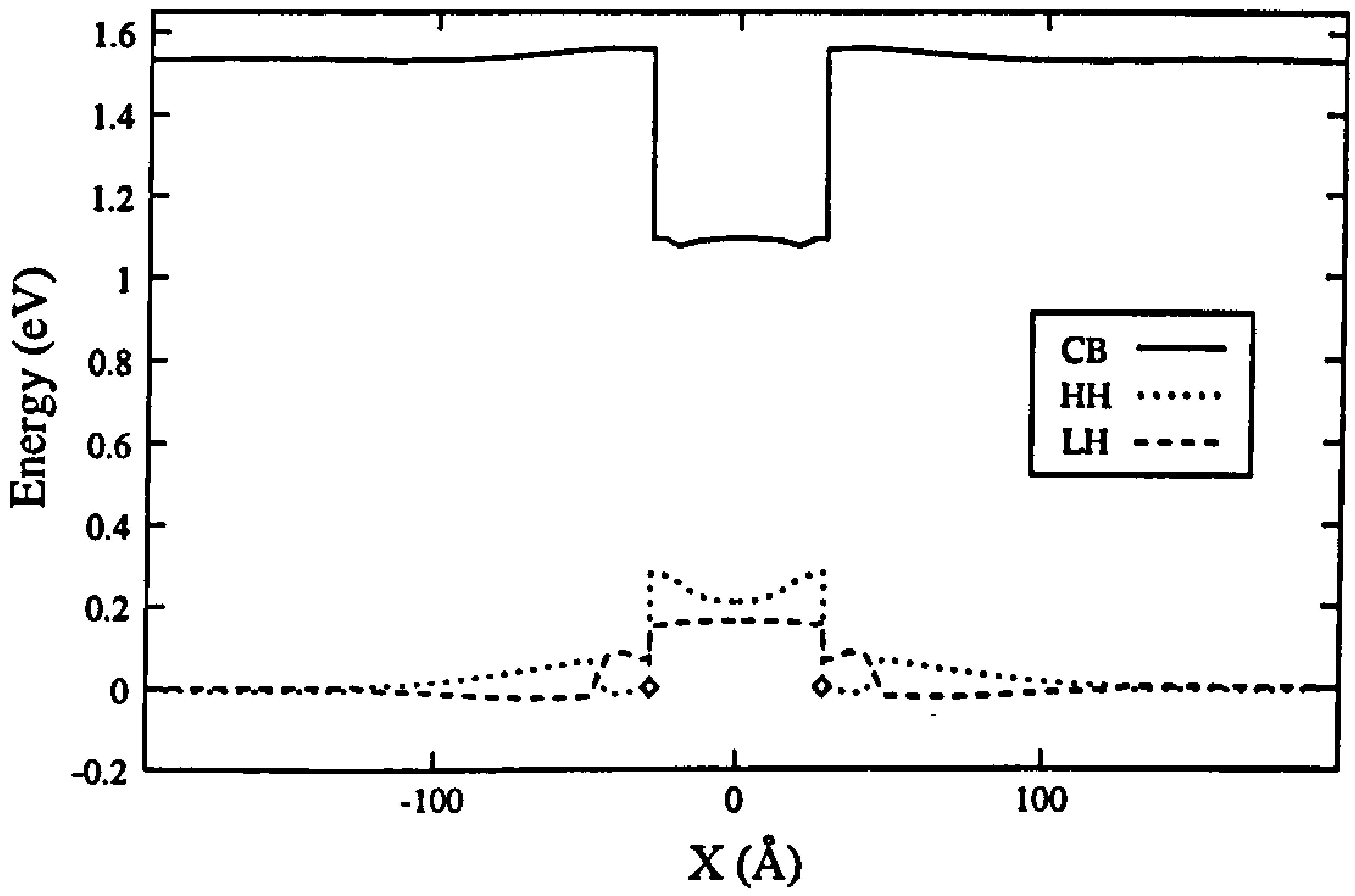

Figure 5.14: The electron, heavy hole and light hole potential profiles for structure 4 in Table 5.4 plotted along the X axis. The diamonds indicate the spatial extent of the InAs dot.

the biaxial strain, ϵ_b (see Figure 5.11). In those regions of the structure where the biaxial strain is negative the light hole band will be shifted upwards in energy and the heavy hole band downwards; in those regions where the biaxial strain is positive, the heavy hole band will be uppermost. When the biaxial strain function is zero, the light and heavy hole bands will be degenerate.

In Figure 5.14 the confining potentials are plotted along X. The homogeneous nature of the strain in the centre of the dot is highlighted by the almost constant electron confining potential. The electron confining potential is shown in the XZ plane in Figure 5.15(a). In this diagram, the energy zero is taken to be that of unstrained GaAs. The potential well for electrons possesses a depth of approximately 400 meV over much of the pyramid, deepening to 450 meV towards the base. Near to the interface, the potential is calculated using an interpolation procedure. Within the wetting layer, the potential is taken to be equal to the potential near to the base of the pyramid. The potential in the barrier is close to zero. In Figure 5.15(b) the contour plot shows

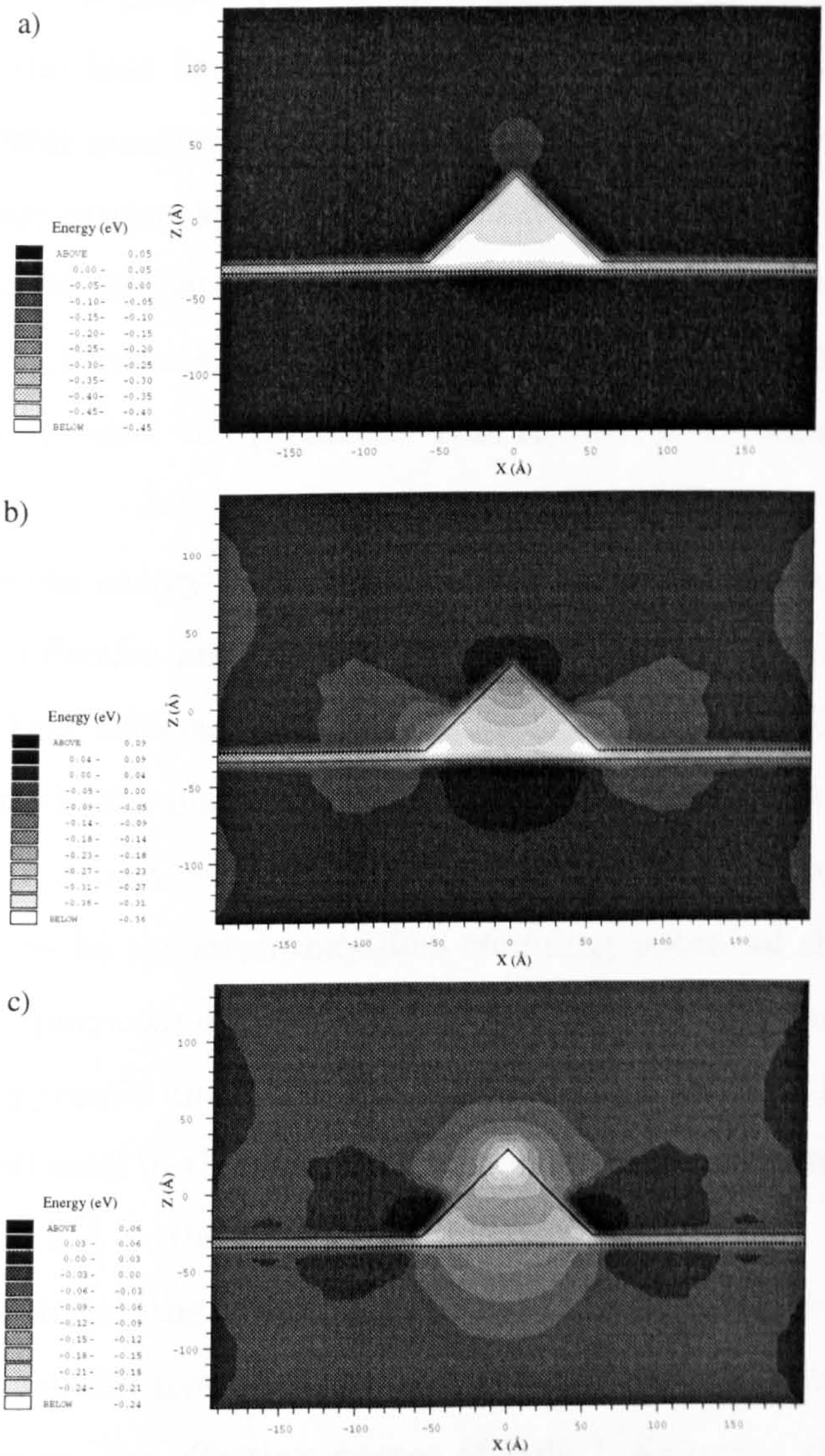

Figure 5.15: The confining potential experienced by the charge carriers in the XZ plane of structure 4 in Table 5.4. Negative values represent an attractive potential. (a) Confining potential for electrons. (b) Confining potential for heavy holes. (c) Confining potential for light holes.

the confining potential experienced by heavy hole carriers. Regions attractive to heavy holes are negative. The potential inside the well possesses a positive gradient from the base towards the apex. Over this distance the potential changes by several hundred meV. This coupled with the large effective mass of the heavy hole carriers should lead us to expect a certain amount of charge localisation near to the base. A similar plot for the light hole confining potential [Figure 5.15(c)] reveals a slowly varying attractive potential in the barrier region above and below the pyramid. This reaches a minimum towards the pyramid's apex.

To calculate the energy levels and electron (or hole) wave functions we use the multi-band effective mass theory described in Section 2.2.3. For simplicity, we assume that the conduction and valence bands are decoupled. For the conduction states in the quantum dot, the single-band effective mass equation (5.1) is solved in a basis consisting of 1331 plane waves. The $V(\mathbf{r})$ term in (5.1) was taken to be the strain-modified confining potential shown in Figure 5.15(a). For the purposes of the electronic structure calculation, the cell containing a single pyramid and portion of the wetting layer was taken to be equal in size to the cell used in the calculation of the atomic positions, i.e., 395.5 Å × 395.5 Å × 282.5 Å in the x, y and z directions respectively.

In the dot material, the compressive stress alters the curvature of the bulk bands causing the effective masses to differ from those of unstrained InAs. For this calculation, the effective masses of bulk InAs under the average hydrostatic strain present in the dot material have been used. These values were obtained by performing empirical pseudopotential band structure calculations (see Section 2.1) for the conduction→valence band momentum matrix elements (P_{cv}) of InAs under pressure. The conduction band effective mass is

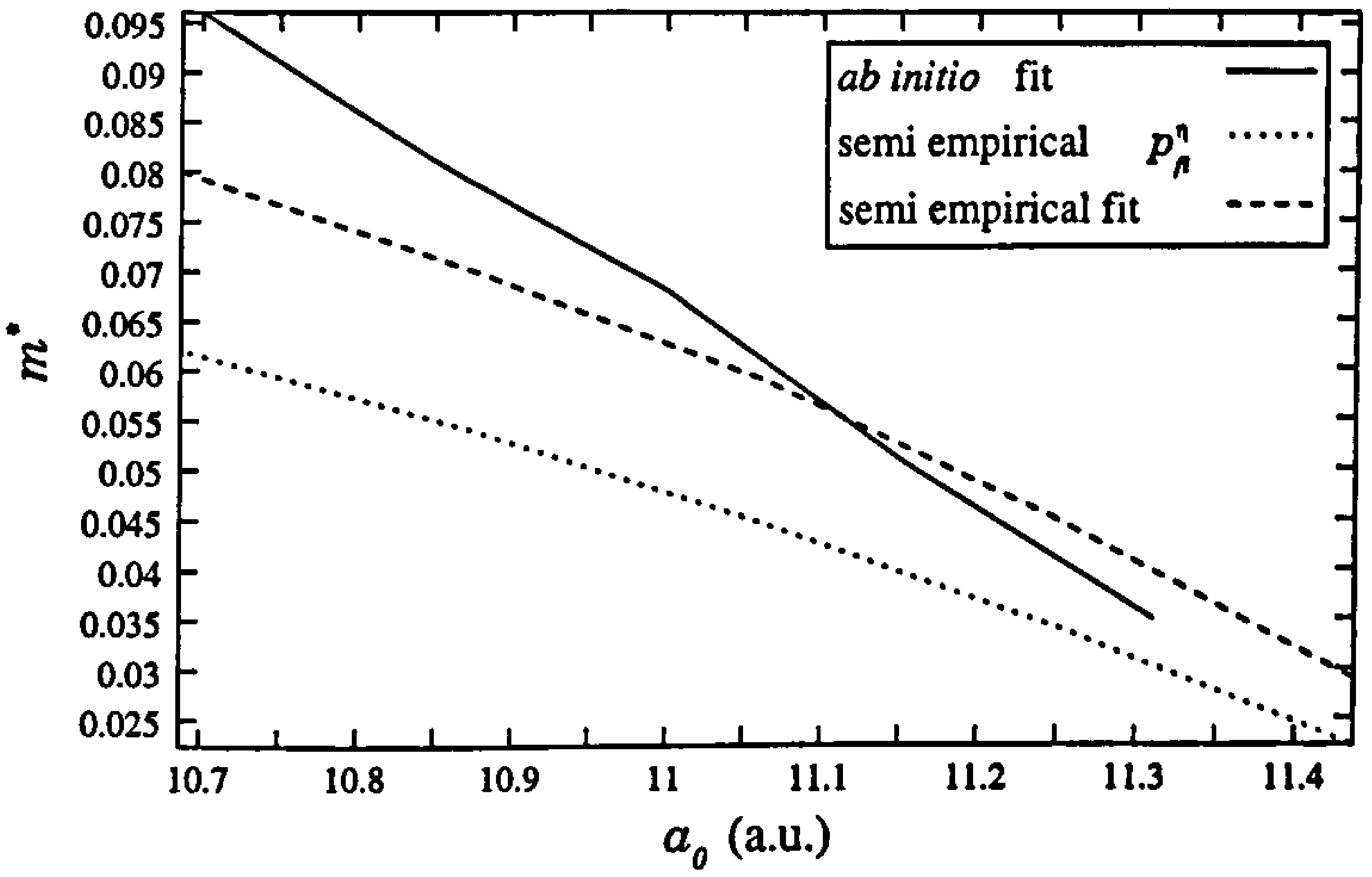

Figure 5.16: The InAs conduction band effective mass (in terms of m_0) plotted for different lattice constants (in atomic units). The results of three different calculations for the effective mass are given.

related to the P_{cv} matrix element by (Ridley, 1982)

$$\left(\frac{m_e^*}{m_0}\right)^{-1} = 1 + \frac{E_p}{E_0}\frac{(E_0 + \frac{2}{3}\Delta)}{(E_0 + \Delta)} \tag{5.8}$$

where the characteristic energy from Kane's theory E_p is related to P_{cv} by (2.66). The change in the conduction band effective mass of InAs with lattice constant is shown in Figure 5.16. The plot shows the variation of mass with lattice constant predicted in three ways: the first graph is the result of *ab initio* local density calculations for the band structure of InAs in the vicinity of Γ followed by the fitting of an effective mass parabola; the second graph was obtained from an empirical pseudopotential calculation of the P_{cv} momentum matrix element and band structure at Γ and (5.8); the last line gives the effective mass vs lattice constant trend for an empirical pseudopotential band structure calculation followed by the fitting of an effective mass parabola. It can be seen that all three techniques give the same trend but the two methods that involved the fitting of an effective mass parabola to the band

m_e^*	$m_{hh}^{[001]}$	$m_{lh}^{[001]}$	$m_{hh}^{[111]}$	$m_{lh}^{[111]}$	γ_1	γ_2	γ_3	E_p (eV)
0.04	0.341	0.027	0.917	0.026	19.67	8.37	9.29	22.2

Table 5.5: Effective masses (in terms of m_0), Luttinger-Kohn parameters and Kane's energy, E_p, for InAs. The conduction band mass is the result of an empirical pseudopotential calculation. All other values are taken from Landolt-Börnstein (1982).

structure in the neighbourhood of Γ give generally larger masses than the values obtained from the momentum matrix element calculation. This is to be expected as it is well known that the effective mass of the conduction band in these materials increases away from the zone centre due to non-parabolicity. The result using (5.8) is more representative of the conduction mass variation with lattice constant at Γ. The mean lattice constant in the InAs dot was calculated to be 11.17 Å. The second graph in Figure 5.16 yields a strained conduction band effective mass of $0.04m_0$ for this lattice constant.

The valence band states are defined by solutions of the four-band effective mass equation (2.58), with the potential function given by the three-dimensional modified heavy hole and light hole confining potentials shown in cross-section in Figures 5.15(b) and 5.15(c). The Luttinger-Kohn band structure parameters $\{\gamma_i\}$ were assumed to be those of unstrained InAs, given by Landolt-Börnstein (1982). The values used are listed in Table 5.5.

In Figure 5.17 we present the electron and hole energy levels as a function of base diameter given by the method detailed above. The electron (hole) states are plotted relative to the unstrained conduction (valence) band edge in GaAs. For dot base diameters smaller than approximately 60 Å, no bound electron states are predicted. This number increases to three for structures larger than 120 Å. In the valence band there are many confined hole states.

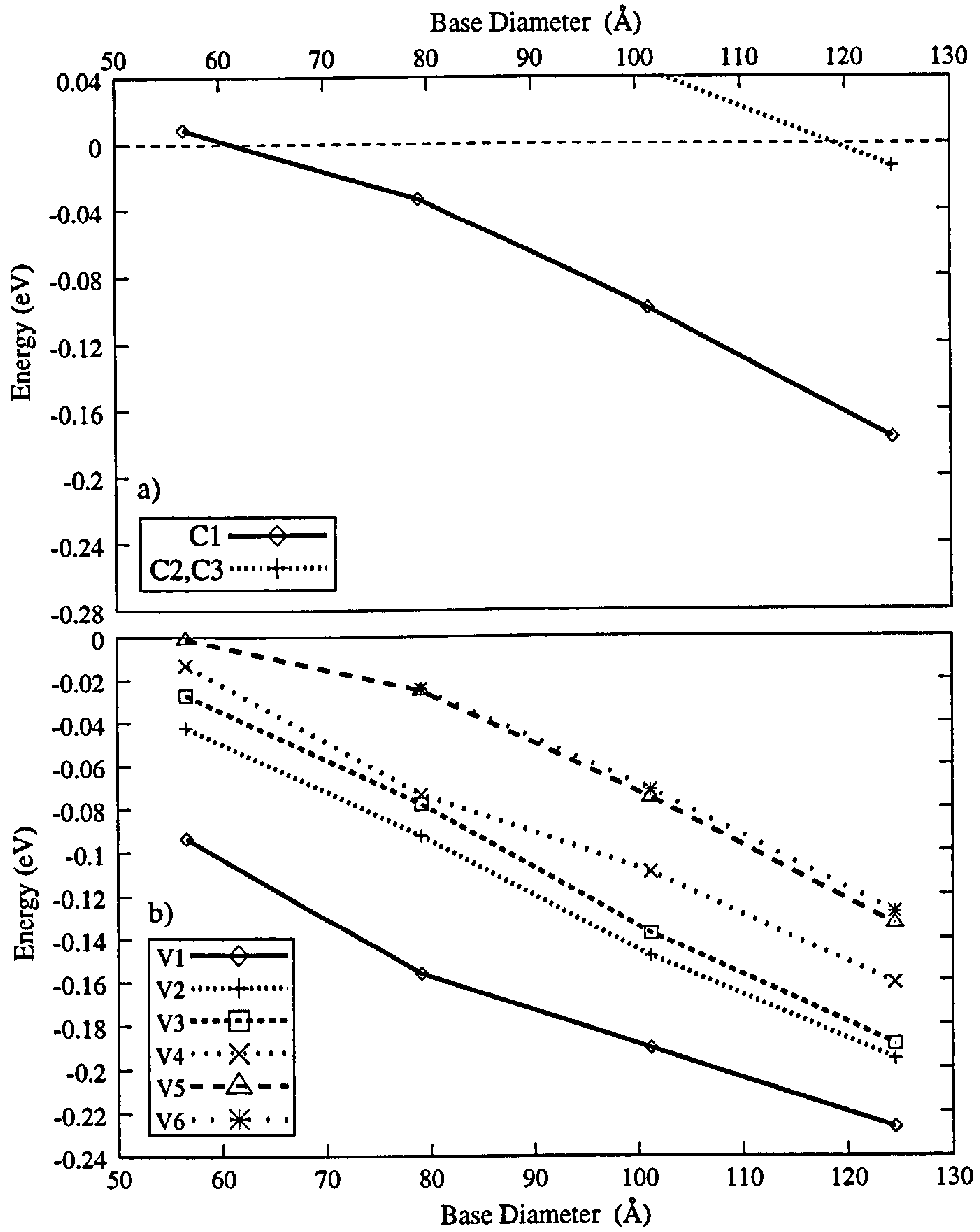

Figure 5.17: (a) Electron and (b) hole quantum dot energy levels displayed as a function of dot base size.

This is due to the larger effective mass associated with these carriers and to the nature of the light hole confining potential whose smoothly varying form leads to a quasi-continuum of tenuously bound states. Due to the C_{4v} symmetry of the pyramidal dot we would have expected the first and second excited hole states to be degenerate. This is true for the electron levels. However, in the valence band the first and second excited hole levels are split due to mixing between different bulk states which can be shown by performing calculations in the diagonal approximation.

In Figure 5.18(a) the modulus-squared envelope function of the C1 state for structure 4 is plotted in the YZ plane through the pyramid and wetting layer. The relatively isotropic character of the confining potential for electrons coupled with the small effective mass results in a state that permeates throughout the dot and penetrates the sides of the pyramid. The envelope function implies that charge does not significantly sample the apex or the base corners of the pyramid. In Figure 5.18(b) we show the envelope of the V1 state plotted in the YZ plane. Unlike the ground conduction state, the ground hole state is confined to the base of the dot due to the larger effective mass and the anisotropic nature of the heavy hole confining potential. Figure 5.18(c-e) shows the envelope functions of the V2, V3 and V4 hole states.

The calculated energies for the C1→V1 transition in each structure are given in the fourth column of Table 5.4. The fundamental transition energy of 1.11 eV for structure 4 is in excellent agreement with the photoluminescence value of 1.1 eV for a similar sized structure (Grundmann *et al.*, 1995a). The energy splitting between the ground and first excited hole state of 30 meV in Figure 5.17(b) is in good agreement with a very recent experimental study (Schmidt *et al.*, 1996) of the sub-level structure which measured a difference of $\approx$ 27 meV. However, Schmidt *et al.* also show that a larger number of

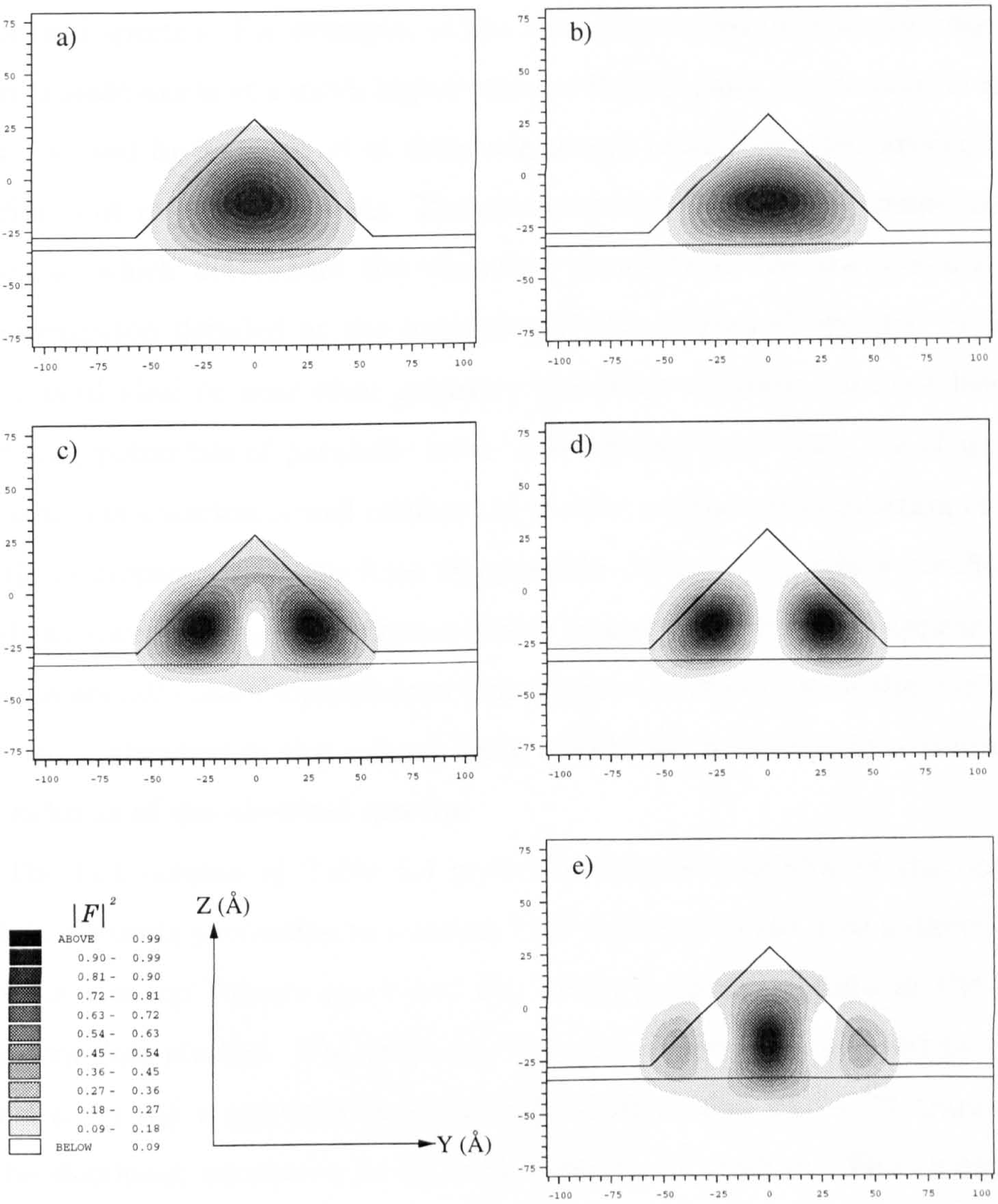

Figure 5.18: The modulus-squared envelope function for structure 4 in Table 5.4 through the YZ plane. (a) State C1. (b) State V1. (c) State V2. (d) State V3. (e) State V4.

conduction states than predicted in Figure 5.17(a) and previous calculations (Marzin and Bastard, 1994b; Grundmann *et al.*, 1995b) contribute effectively to optical spectra. For example, in the calculation presented above only one excited state exists at a much higher energy. Furthermore, the transition energies observed by Schmidt *et al.* form a sequence resembling that arising from confinement in parabolic wells. The evaluation of the confining semiclassical potential which determines the electronic structure in the particle-in-a-box approximation detailed at the beginning of this section shows that at least for dots of ideal or near ideal geometry the effect of strain does not lead to confining potentials of parabolic form. Also, from Figure 5.18, the charge of the states in question is well confined in the dot so that reexamination of any subtle discrepancies arising from the possible choice of boundary conditions (such as abrupt changes in effective mass) cannot help. It would appear that a fundamentally more sophisticated Hamiltonian consistent with the confining potentials obtained in the valence force calculation is needed to account for the richness of the observed spectra.

The last column of Table 5.4 gives the squared modulus of the optical matrix elements pertaining to incident light polarised in the x or y directions, between the top valence state and the lowest conduction state in the four quantum dots studied. The transition between these states induced by light polarised in the z (growth) direction is forbidden. The V1 → C1 transition is the dominant excitation in all of the structures studied. The change in matrix element with increasing dot size can be attributed to the increased confinement of the states in the larger dots and therefore the larger overlap between the envelope functions of the conduction and valence states. The V1 → C1 transition in structure 1 is only weakly allowed compared to that in the larger dots because the conduction state is only tenuously bound to the well (see Figure 5.17) and one would expect a poor overlap of the conduction and valence envelope functions as a result.

In summary, a calculation for the energy levels and wave functions of InAs/GaAs self-organized quantum dots of differing sizes has been performed. Account was taken of the strain modification to the confining potential, valence band mixing, and the conduction band mass in the InAs dot and the surrounding GaAs barrier. It was shown that the geometry of the system coupled with the inhomogeneous strain confines the ground electron and hole states to the base of the pyramidal dot. The calculated fundamental transition energy and the valence sub-level structure agree very well with the available experimental data. However, in the conduction band very recent experimental studies indicate that many more conduction states may be involved in the construction of optical spectra than predicted here.

Chapter 6

Concluding Remarks

The role of band structure engineering in maximising linear and nonlinear optical responses in quantum wells and dots has been investigated. To undertake this task, established theories of the energy bands and polarisation response in solids have been taken, modified and finally applied to describe electronic and optical effects in a range of semiconductor heterostructures. We summarise, in this chapter, the principle findings of this study, and provide suggestions as to how the calculations performed might be further refined and extended.

6.1 Summary of Results

The calculation of the energy bands, wave functions and optical matrix elements in III-V compound semiconductors and in semiconductor heterostructures was described. The one electron Schrödinger equation for bulk GaAs and GaAs-based quantum well structures was solved using the empirical pseudopotential method and the four-band $\mathbf{k} \cdot \mathbf{p}$/effective mass approximation. Agreement between the two band structure techniques was tested by calculating the dispersion of the valence subbands in a 50 Å GaAs/AlAs quantum well. To assure equivalent inputs, all parameters needed for the effective mass treatment were extracted from the calculated bulk GaAs and AlAs pseudopotential band structures. A good agreement for the heterostructure hole bands was achieved.

The quantum mechanical theory of virtual optical nonlinearities in semiconductors was detailed. An expression was derived for the second-order susceptibility tensor, $\chi^{(2)}(-2\omega;\omega,\omega)$, governing the extent to which a semiconductor radiates second harmonics. Using the formula, the second harmonic response in p-type GaAs/Al$_x$Ga$_{1-x}$As 85 Å period quantum well systems of differing well width and well depth was calculated. The optical nonlinearities in these structures are a result of virtual transitions between bound valence subbands. The dominant component of $\chi^{(2)}(-2\omega;\omega,\omega)$ was found to be $\chi^{(2)}_{zxy}$, in contrast to the prediction of the particle-in-a-box model that $\chi^{(2)}_{zzz}$ would be largest. The second harmonic response in several shorter (57 Å) period p-type GaAs/Al$_x$Ga$_{1-x}$As quantum well systems was studied. In these structures, $\chi^{(2)}_{xyz}$ is predicted to be the dominant component of $\chi^{(2)}(-2\omega;\omega,\omega)$. The response was found to be generally weaker and possessed a different directional dependence to that found in the longer period systems. This is because the nonlinearities in the 57 Å period systems arise from transitions involving continuum states which are generally weaker than their bound → bound counterparts. It was shown that in order to optimise second harmonic generation in these materials, the structural parameters should be chosen as to maximise the confinement of the participating states. A typical set of parameters was provided. Further, it was demonstrated that a high degree of well asymmetry can result in an increased second harmonic response in these structures, although the response is highly sensitive to the detailed band structure of the quantum well.

The four-band effective mass approximation was applied to the calculation of the valence band structure of cubic GaAs/AlAs quantum dots. The conduction band levels were calculated using a single-band approach. The calculated charge densities emphasised the high degree of confinement experienced by electrons and holes in these structures. The variation with dot size of the energies of the conduction and valence states and the interband transition strengths between them was examined. It was shown that in all the structures the maximum oscillator strength was achieved by transitions between the uppermost valence state and the lowest conduction state in the quantum dot. The transitions exhibited a directional independence on the polarisation of the incident light, as expected from the cubic symmetry. The energy levels and optical matrix elements in the absence of mixing between bulk valence bands were determined. The lifting of degeneracies and anisotropy of the optical matrix elements under these conditions showed this to be an invalid approximation, underlining the importance of contributions from different bulk bands in the construction of a quantum dot valence state.

The effective mass approximation was also used to calculate the electronic structure and the strength of the allowed optical transitions of InAs/GaAs self-assembled quantum dots of differing sizes. Account was taken of the strain modification to the three dimensional confining potential, valence band mixing and the effect of strain on the conduction band mass in the InAs dot. It was shown that the geometry of the system coupled with the inhomogeneous strain confines the ground electron and hole states to the base of the pyramidal dot. The calculated fundamental transition energy and the valence sublevel structure agree very well with the available experimental data. However, in the conduction band very recent experimental studies indicate that many more conduction states may be involved in the construction of optical spectra than predicted here. The transition dominating the optical response of these materials was shown to be that between the lowest conduction state and the highest valence state. The transition was found to be forbidden for light polarised along the principle axis of the InAs pyramid. For incident light polarised in the plane parallel to the pyramid base, large oscillator strengths were obtained for the fundamental transition.

6.2 Refinements and Extensions

The recent development of length gauge expressions for the nonlinear optical susceptibilities pertaining to infinite semiconductors (Aversa and Sipe, 1995) which are free of the unphysical (infra-red) divergences that plague the velocity gauge approach represents a significant step forward. In order to successfully apply these expressions to semiconductor heterostructures in the near-resonant, low field regime, relaxation will need to be included, probably in a phenomenological fashion. An assessment can then be undertaken of how realistic the velocity gauge expressions used in this study (and in many other theoretical studies) are. A more sophisticated treatment of the relaxation operator is required, one that accounts for frequency, wave vector and subband index. This is so that a more realistic low intensity, near-resonant evaluation of the susceptibilities may take place. However, even in its present form, this model of the optical susceptibilities has provided useful information concerning the structural parameters leading to optimal second harmonic generation in III-V quantum well systems in the non-resonant, ultra short timescale regime, where virtual optical processes dominate.

Further extensions to the work concerning the strain distribution, the electronic structure and optical transitions in self-assembled quantum dots is also possible. The inclusion of excitonic effects should be straightforward. Also, changes to the confining potential due to piezoelectric polarisations induced by shear strains could be determined, However, in order to include this effect, fresh thought will need to be given to the method by which strain tensor components are obtained from the relaxed atomic positions. This is because the current technique is unreliable at the interface, where the largest shear strains are most likely to occur. The construction of the confining potentials for different charge carriers could be further refined by accounting for the gradient of the strain in the effective Hamiltonian, as outlined by Zhang (1994). Another important step would be to include the bulk conduction and split-off bands in the basis set used to construct the quantum dot states. This would improve the accuracy of the allowed transitions and perhaps lead to better agreement between the calculated conduction states and the experimental data. Account could also be taken of the change of effective masses at the interface between the two materials. However, both of these steps represent such a significant increase in the size of the computational problem as to render its solution intractable at the present time.

Appendix A

An Overview of PVM

PVM (Parallel Virtual Machine) is a software package that allows a group of networked computers to act as a resource for parallel computing. Under PVM a user-defined grouping of serial, parallel and vector computers appears as one large distributed memory machine. The PVM system consists of two parts:

A daemon, `pvmd3`**.** This resides on the machines making up the virtual parallel computer. It provides, among other things, an inter-host point of contact.

A library of PVM interface routines. The library contains routines for passing messages, spawning processes, coordinating processes and modifying the virtual machine. Applications can be written in C or FORTRAN and are parallelised by using message passing constructs contained in the PVM library.

PVM is an on-going research project started in 1989 at Oak Ridge National Laboratory (ORNL) and now involving people from ORNL, the University of Tennessee and Carnegie Mellon University. The project produces software that is useful to other members of the scientific community and, in the interests of advancing science, this software is available freely over the Internet. The documentation and source code for more than thirty UNIX platforms can be obtained from the World Wide Web at `http://www.epm.ornl.gov/pvm/`.

References

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

Apetz, R., Vescan, L., Hartmann, A., Dieker, C., and Lüth, H. (1995), *Appl. Phys. Lett.* **67**, 445.

Aversa, C., and Sipe, J. E. (1995). *Phys. Rev. B* **52**, 14 636.

Bandara K. M. S. V., Coon, D. D., and O, Byungsung (1988). *Appl. Phys. Lett.* **53**, 1931.

Baraff, G. A., and Gershoni, D. (1991). *Phys. Rev. B* **43**, 4011.

Bass, M., Franken, P. A., Hill, A. E., Peters, C. W., and Weinreich, G. (1962a). *Phys. Rev. Lett.* **8**, 18.

Bass, M., Franken, P. A.,Ward, J. F., and Weinreich, G. (1962b). *Phys. Rev. Lett.* **9**, 446.

Bastard, G., and Brum, J. A. (1986). *IEEE J. Quantum Electron.* **22**, 1625.

Bastard, G. (1988). "Wave Mechanics Applied to Semiconductor Heterostructures", Les Editions de Physique, Les Ulis Cedex.

Beguelin, A., Dongarra, J. J., Geist G. A., Manchek, R., and Sunderam, V. S. (1994). "A Users' Guide to PVM Parallel Virtual Machine", Oak Ridge National Laboratory, ORNL/TM-12187.

Bloembergen, N. (1965). "Nonlinear Optics", Benjamin, New York.

Bloom, S., and Bergstresser, T. K. (1968). *Solid State Commun.* **6**, 465.

Boucaud, P., Julien, F. H., Yang, D. D., Lourtioz, J. M., Rosencher, E., Bois, P., and Nagle, J. (1990). *Appl. Phys. Lett.* **57**, 215.

Boyd, R. W. (1992). "Nonlinear Optics", Academic Press, Boston.

Briddon, P. R. (1990). PhD Thesis, "Defects in Solids: an *Ab Initio* Study", University of Exeter.

Butcher, P. N., and McLean, T. P. (1963). *Proc. Phys. Soc.* **81**, 219.

Butcher, P. N., and McLean, T. P. (1964). *Proc. Phys. Soc.* **83**, 579.

Butcher, P. N., and Cotter, D. (1990). "The Elements of Nonlinear Optics", Cambridge University Press, Cambridge.

Chang, L. L., Esaki, L., Howard, W. E., and Ludeke R. (1973). *J. Vac. Sci. Technol.* **10**, 11.

Chang, Y. -C., and James, R. B. (1989). *Phys. Rev. B* **39**, 12 672.

Charles, R., Saint-Cricq, N., Renucci, J. B., Renucci, M. A., and Zwick, A. (1980). *Phys. Rev. B* **22**, 4804.

Cohen, M. L., and Bergstresser, T. K. (1966). *Phys. Rev.* **141**, 789.

Coon, D. D., and Karunasiri, R. P. G. (1984). *Appl. Phys. Lett.* **45**, 649.

Esaki, L., and Tsu R. (1970). *IBM J. Res. Develop.* **14**, 61.

Fejer, M. M., Yoo, S. J. B., Byer, R. L., Harwit, A., and Harris, J. S., Jr. (1989). *Phys. Rev. Lett.* **62**, 1041.

Flytzanis, C. (1975). "Quantum Electronics" (H. Rabin and C. L. Tang eds.), Academic Press, New York.

Franken, P. A., Hill, A. E., Peters, C. W., and Weinreich, G. (1961). *Phys. Rev. Lett.* **7**, 118.

Gershoni, D., Henry, C. H., and Baraff, G. A. (1993). *IEEE J. Quantum Electron.* **29**, 2433.

Grundmann, M., Christen, J., Ledentsov, N. N., Böhrer, J., Bimberg, D., Ruvimov, S. S., Werner, P., Richter, U., Gösele, U., Heydenreich, J., Ustinov, V. M., Egorov, A. Yu., Kop'ev, P. S., and Alferov, Zh. I. (1995a). *Phys. Rev.*

Lett. **74**, 4043.

Grundmann, M., Stier, O., Bimberg, D. (1995b). *Phys. Rev. B* **52**, 11 969.

Grundmann, M., Ledentsov, N. N., Heitz, R., Eckey, L., Christen, J., Böhrer, J.,Bimberg, D.,Ruvimov, S. S., Werner, P., Richter, U., Heydenreich, J., Ustinov, V. M., Egorov, A. Yu., Zhukov, A. E., Kop'ev, P. S., and Alferov, Zh. I. (1995c). *Phys. Status Solidi B* **188**, 249.

Heath, M. T., and Etheridge, J. A. (1994). *IEEE Software* **8**, 29.

Hermann, F., and Skillman, S. (1963). "Atomic Structure Calculations", Prentice-Hall, New York.

Janz, S., Chatenoud, F., and Normandin, R. (1994). *Opt. Lett.* **19**, 622.

Jaros, M., and Wong, K. B. (1984). *J. Phys.* **C17**, L 765.

Jaros, M., Wong, K. B., and Gell, M. A. (1985). *Phys. Rev. B* **31**, 1205.

Kane, E. O. (1966). "Semiconductors and Semimetals" (R. K. Willardson and A. C. Beer, eds.) Vol. 1. Academic Press, New York.

Kirstaedter, N., Ledentsov, N. N., Grundmann, M., Bimberg, D., Ustinov, V. M., Ruvimov, S. S., Maximov, M. V., Kop'ev, P. S., Alferov, Zh. I., Richter, U., Werner, P., Gösele, U., and Heydenreich, J. (1994). *Electr. Lett.* **30**, 1416.

Kohn, W., and Luttinger, J. L. (1957). *Phys. Rev* **108**, 590.

Lambrecht, W. R. L., and Segall, B. (1988). *Phys. Rev. Lett.* **61**, 1764.

Landolt-Börnstein (1982). In "Numerical Data and Functional Relationships in Science and Technology" (O. Madelung, M. Schultz and H. Weiss, eds.), Group III vol. 17a, Springer, New York.

Ledentsov, N. N., Grundmann, M., Kirstaedter, N., Schmidt, O., Heitz, R., Böhrer, J., Bimberg, D., Ustinov, V. M., Shchukin, V. A., Egorov, A. Yu., Zhukov, A.E., Zaitsev, S., Kop'ev, P. S., Alferov, Zh. I., Ruvimov, S. S., Kosogov, P., Werner, P., Gösele, U., and Heydenreich, J. (1996). *Solid State Electronics* **40**, 785.

Leon, R., Fafard, S., Leonard, D., Merz, J. L., and Petroff, P. M. (1995). *Appl. Phys. Lett.* **67**, 521.

Leonard, D., Krishnamurthy, M., Reaves, C. M., Denbaars, S. P., and Petroff, P. M. (1993). *Appl. Phys. Lett.* **63**, 3203.

Leonard, D., Pond, K., and Petroff, P. M. (1994). *Phys. Rev. B* **50**, 11 687.

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

Luttinger, J. M., and Kohn, W. (1955). *Phys. Rev.* **97**, 8690.

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

Mäder, K. A., and Zunger, A. (1995). *Phys. Rev. B* **50**, 17 393.

Martin, R. M. (1970). *Phys. Rev. B* **1**, 4005.

Marzin, J.-Y., Gérard, J.-M., Izraël, A., Barrier, D., and Bastard G. (1994a). *Phys. Rev. Lett.* **73**, 716.

Marzin, J.-Y., and Bastard, G. (1994b). *Solid State Commun.* **92**, 437.

Medeiros-Ribeiro, G., Leonard, D., and Petroff, P. M. (1995). *Appl. Phys. Lett.* **66**, 1767.

McMurry, H. L., Solbrig, Jr., A. W., Boyter, J. K., and Noble, C. (1967). *J. Phys. Chem. Solids* **28**, 2359.

Moison, J. M., Houzay, F., Barthe, F., Leprince, L., André, E., and Vatel, O. (1994). *Appl. Phys. Lett.* **64**, 196.

Musgrave, M. J. P., and Pople, J. A. (1962). *Proc. Roy. Soc. (London)* **A268**, 474.

Patel, C. K. N. (1965). *Phys. Rev. Lett.* **15**, 1027.

Patel, C. K. N. (1966). *Phys. Rev. Lett.* **16**, 613.

Petroff, P. M., and Denbaars, S. P. (1994). *Superlattice Microst.* **15**, 15.

Pollak, F. H., and Cardona, M. (1968). *Phys. Rev.* **172**, 816.

Pollak, F. H. (1973). *Surface Science* **37**, 863.

Pollak, F. H. (1990). "Semiconductors and Semimetals" (R. K. Willardson and A. C. Beer, eds.) Vol. 32. Academic Press, New York.

Priester, C., and Lannoo, M. (1995). *Phys. Rev. Lett.* **75**, 93.

Qu, X. H., and Ruda, H. (1993). *Appl. Phys. Lett.* **62**, 1946.

Qu, X. H., and Ruda, H., Janz, S., and Springthorpe, A. J. (1994). *Appl. Phys. Lett.* **65**, 3176.

Reed, M. A. (1993). *Scientific American* **268**, 118.

Ridley, B. K. (1982). "Quantum Processes in Semiconductors", Clarendon Press, Oxford.

Rosencher, E., Bois, P., Nagle, J., and Delaître, S. (1989). *Electron. Lett.* **25**, 1063.

Rosencher, E., Fiore, A., Vinter, B., Berger, V., Bois, P., and Nagle, J. (1996). *Science* **271**, 168.

Schmidt, B. K., and Sunderam, V. S. (1994). *Concurrency: Practice and Experience* **6**, 1.

Schmidt, K. H., Medeiros-Ribeiro, G., Oestreich, M., Döhler, G. H., and Petroff, P. M. (1996). Accepted for publication in *Phys. Rev. B* 54, No. 15.

Sercel, P. C., and Vahala, K. J. (1990). *Phys. Rev. B* **42**, 3690.

Shaw, M. J., Wong, K. B., and Jaros, M. (1993a). *Phys. Rev. B* **48**, 2001.

Shaw, M. J. (1993b). PhD Thesis. "Microscopic Theory of Optical Nonlinearity in Semiconductor Superlattices", University of Newcastle upon Tyne.

Shaw, M. J., Jaros, M., Xu, Z., Fauchet, P. M., Rella, C. W., Richman, B. A., Schwettman, H. A., and Wicks, G. W. (1994). *Phys. Rev. B* **50**, 18 395.

Solomon, G. S., Trezza, J. A., and Harris, Jr., J. S. (1995). *Appl. Phys. Lett.* **66**, 3161.

Stranski, I. N., and Krastanow, Von. L. (1939). *Akad. Wiss. Lit. Mainz Math.-Natur. Kl. IIb* **146**, 797.

Strauch, D., and Dorner, B. (1990). *J. Phys. Condens. Matte.* **2**, 1457.

Szmulowicz, F., and Brown, G. J. (1995). *Appl. Phys. Lett.* **66**, 1659.

Terhune, R. W., Maker, P. D., and Savage, C. M. (1962). *Phys. Rev. Lett.* **8**, 404.

Tsang, L., and Chuang, S. L. (1992). *Appl. Phys. Lett.* **60**, 2543.

Van der Walle, C. G. (1989). *Phys. Rev B* **39**, 1871.

Weisz, G. (1966). *Phys. Rev.* **149**, 504.

West, L. C., and Eglash, S. J. (1985). *Appl. Phys. Lett.* **46**, 1156.

Zhang, Y. (1994). *Phys. Rev. B* **49**, 14 352.