Electronic structure calculations on nitride semiconductors and their alloys

Dugdale David

Electronic Structure Calculations on Nitride Semiconductors and their Alloys

David Dugdale

A thesis submitted for the

degree of Doctor of Philosophy

at the University of Durham,

Department of Physics

September 2000

Abstract

Calculations of the electronic properties of AlN, GaN, InN and their alloys are presented. Initial calculations are performed using the first principles pseudopotential method to obtain accurate lattice constants. Further calculations then investigate bonding in the nitrides through population analysis and density of state calculations. The empirical pseudopotential method is also used in this work. Pseudopotentials for each of the nitrides are constructed using a functional form that allows strained material and alloys to be studied. The conventional **k.p** valence band parameters for both zincblende and wurtzite are obtained from the empirical band structure using two different methods. A Monte-Carlo fitting of the **k.p** band structure to the pseudopotential data (or an effective mass method for the zincblende structure) is used to produce one set. Another set is obtained directly from the momentum matrix elements and energy eigenvalues at the centre of the Brillouin zone. Both methods of calculating **k.p** parameters produce band structure in excellent agreement with the original empirical band calculations near the centre of the Brillouin zone. The advantage of the direct method is that it produces consistent sets of parameters, and can be used in studies involving a series of alloy compositions. Further empirical pseudopotential method calculations are then performed for alloys of the nitrides. In particular, the variation of the band gap with alloy composition is investigated, and good agreement with theory and experiment is found. The direct method is used to obtain **k.p** parameters for the alloys, and is contrasted with the fitting approach. The behaviour of the nitrides under strain is also studied. In particular, valence band offsets for nitride heterojunctions are calculated, and a strong forward-backward asymmetry in the band offset is found, in good agreement with other results in the literature.

Declaration

The work presented in this thesis has been carried out by the candidate (except where otherwise acknowledged) and has not been previously submitted for any degree.

_____

Ph.D. Candidate

_____

Ph.D. Supervisor

Acknowledgements

I would like to thank my supervisors Dr. Stuart Brand and Prof. Dick Abram for their help and assistance throughout my course of study. Additionally, I would like to acknowledge EPSRC for funding this work, and I am also indebted to my parents for their continued support throughout my studies. Last but not least, I would like to thank the people who have made my time in Durham such an enjoyable experience: Rick Coles, Dan Harrison, Gavin Crow, Stewart Clark, Steve Pugh, Dave Hoare, Des Ryan, Derek Middlemiss and Ian Bolland.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Semiconductors play an integral part in modern technology. Defined as having an energy gap of less than a few eV, they are the material of choice for the electronics industry due to the properties they offer and the increasing sophistication with which they can be grown.

The best known semiconductor, and that for which technology is best developed, is the group IV element Silicon. Binary compounds, formed from elements of groups III and V (or groups II and VI), have properties broadly similar to their group IV counterparts, but with increased ionicity due to the transfer of electronic charge from the group III atom to the group V atom. This ionicity has a profound effect on their semiconducting properties, causing increases in the Coulomb interaction between ions and in the fundamental energy gap. Semiconductors with larger band gaps have uses in colour displays and lasers, while smaller band gap materials typically have applications in infrared detectors. In addition to binary materials, ternary and quaternary compounds also exist with which a range of properties are possible. Heterostructures comprised of semiconducting materials can also be fabricated, allowing for further manipulation of the semiconducting properties. It is because of this degree of control that semiconductors are essential for modern electronic devices.

Within this context, semiconductor devices based on AlN, GaN and InN have re-

ceived great interest over the last few years due to the electronic and optical properties they offer. The principal reason for this interest is that they are wide band gap materials, and together with their alloys AlGaN, GaInN and AlInN, band gaps ranging from approximately 2.0 eV (InN) to 6.2 eV (AlN) are possible. Thus an immediate use of nitride LEDs is in full colour displays, for which the three primary colours red, green and blue are required. For nitride based laser diodes, light can be focused more sharply than in conventional lasers because of their short operating wavelength, and thus optical storage capacities can be greatly increased. The capacities of DVDs, for instance, would increase from 4.7 gigabytes to over 15 gigabytes going from AlGaAs to InGaN based laser diodes.

Other wide band gap semiconductors do exist, but in comparison to these and other materials the III-V nitrides have several advantageous properties. The bond strengths in the nitrides, for example, are typically larger than those found in II-VI semiconductors, with Landwehr *et al.* [1] measuring bond energies of 2.3 eV/bond in GaN and 1.2 eV/bond in ZnSe. The low field mobilities of electrons and holes in GaN are also reasonable, with reported room temperature electron mobilities as high as 900 cm^2/V s [2]. This value is significantly improved upon in AlGaN/GaN heterostructures, where mobilities of 1700 cm^2/V s have been reported [3]. Moreover, Monte Carlo simulations of GaN find a peak velocity of $\sim 3 \times 10^7$ cm/s and a saturation velocity of 1.5×10^7 cm/s [4], values which are considerably larger than those found in Si or GaAs. The breakdown voltage of GaN of 3×10^6 V/cm also compares very favourably with that of Si (2×10^5 V/cm) or GaAs (4×10^5 V/cm), making it an ideal candidate for use in high power devices. In addition, the nitrides are well suited for high temperature applications because of their high thermal stability, which is also desirable from the point of view of device processing. Thus the III-V nitrides find a particular niche in high temperature and high power applications.

Importantly, III-V nitride quantum wells and superlattices can also be fabricated. For this, AlGaN or GaN are used as the barrier and cladding layers, with GaN or InGaN

as the active layers. In addition, the lattice mismatch between AlN, GaN and InN also allows various properties of nitride heterojunctions, such as the band offset, to be tuned to some degree. This is possible because the strain in these heterostructures, which can be altered by using a suitable alloy affects these properties via the piezoelectric effect. Such heterostructure technology is necessary for the fabrication of modern devices, and ensures that the III-V nitrides have a variety of commercial applications. Indeed, the development of the III-V nitrides over the last decade has largely been fuelled by their commercial potential. What follows is a summary of the principal points in the development of the nitride materials. For a more detailed discussion of their growth and applications the reader is referred to the reviews of Ambacher [5] and Jain *et al.* [6].

The first blue GaN LED was fabricated by Pankove *et al.* [7,8] in the early 1970s. However, for the rest of this decade the development of III-nitrides was severely hampered by the low material quality of GaN, due in part to a lack of an ideal substrate. Progress began again in 1983 with the development by Yoshida *et al.* [9] of a two-step method of growth. In this, an AlN buffer layer is initally grown on the sapphire substrate at low temperature of $\sim$ 500 ^{0}C. Significantly higher quality GaN layers could then be obtained by growing GaN on this buffer layer at a higher temperature. The two-step method was further developed by Amano and co-workers [10,11], resulting in the growth of high quality GaN films with high mobility and strong photoluminescence intensity. The second breakthrough necessary for the continued development of the nitride materials was the growth of p-type GaN in 1989. This was crucial for LEDs and laser diodes, both of which rely on a p-n junction to inject holes and electrons into the active layers of the device. Previously, only n-type GaN was available. It was Amano *et al.* [12] who first obtained p-type GaN films by using Mg-doping for the acceptor impurity and a low energy electron beam irridation (LEEBI) treatment following a metal organic chemical vapour deposition (MOCVD) growth method. Nakamura refined this method, replacing the LEEBI with a N_2-ambient thermal annealing for the Mg-doped GaN [13]. Lastly, the deveopment of high quality InGaN by Nakamura and Mukai [14]

enabled the commercial development of III-V nitride based LEDs and laser diodes for emitting red to UV light. For these, incorporation of small amounts of In in GaN was found to result in the strong band to band emission at room temperature necessary in such devices [15].

The development of devices based on nitride materials has resulted in a demand for their characterisation. At present, the III-V nitrides still suffer from low material quality, and thus experimental measurements are somewhat compromised. Theoretical methods are obviously not affected by such factors, and are therefore extremely important in reliably determining various quantities which are not accessible through experiment. These theoretical methods are well developed, having been successfully used to study established materials such as Si and GaAs, and they will remain essential to studies of the nitrides until the technology matures. To date, experimental and theoretical investigations have established reliable values for such quantities as lattice parameters [16–18] and fundamental band gaps [19–21], but many other parameters remain uncertain. For device modelling, studies include calculations of optical gain in both bulk and quantum well structures [22–24], as well as Monte Carlo simulations of electron transport [25,26]. Calculations such as these rely on accurate band structure models and transport parameters, and the work in this thesis is primarily directed towards providing them.

In this work, three theoretical methods have been used. These methods approach the problem of describing the electronic structure in different ways, and are particularly suited for different applications. The first of these methods, the first principles (or *ab initio*) pseudopotential method, is useful for obtaining accurate lattice parameters and also for identifying general trends in materials. These calculations are computationally expensive, but have become increasingly tractable with increases in computational power over the last two decades. One feature associated with these calculations is that although the valence band structure is accurately obtained, the fundamental band gap is underestimated.

In contrast, the empirical pseudopotential method (EPM) has the advantage of giving accurate band structure, since it is specifically designed to reproduce accepted band energies from experiment or more sophisticated first principles calculations. This method has commonly been used in calculations of effective masses and in studies of the dielectric response. Computationally, it is less expensive than first principles approaches.

The **k.p** method uses perturbation theory to obtain the band structure around a certain point in k-space (usually the Gamma point). Of all the models, the **k.p** method [27] using six or eight bands is one that is especially useful for device modelling. This is because its computational demands are light compared to the other methods discussed above. In particular, the **k.p** approach is known to provide a good description of the heavy-hole, light-hole and crystal field split-off bands at the centre of the Brillouin zone, which are of obvious importance in any consideration of optical properties.

1.1 Description of the work

The general structure of each chapter is to introduce the theoretical background for the particular method employed. Results obtained using the method are then presented and discussed.

Chapter 2

In this chapter, the first principles psuedopotential method is introduced, as are the zincblende and wurtzite structures in which they can crystallise. Using this method, initial calculations on bulk AlN, GaN and InN are performed to obtain accurate lattice constants. Additional studies addressing bonding in the nitrides are also presented; through population analysis the covalency of the bonding is determined, and from density of states calculations the specific orbitals involved in this bonding are obtained.

Chapter 3

This chapter covers the various aspects of the empirical pseudopotential method, with particular emphasis placed on the concept of pseudopotentials and how they have been obtained in this work. Local and nonlocal pseudopotentials are discussed, together with the mechanism whereby spin orbit coupling is included in calculations. Pseudopotentials for AlN, GaN and InN which are adaptable for the study of strained material and alloys are derived. These are then used to calculate band structures and dielectric functions.

Chapter 4

Chapter 4 introduces the **k.p** method and presents calculations performed using it. From the empirical pseuodpotential models for the nitride materials, two sets of **k.p** parameters are obtained. One set is obtained by specifically fitting the parameters to the empirical band structure. The other set is obtained directly from the momentum matrix elements and energy eigenvalues at the centre of the Brillouin zone. The **k.p** parameters obtained in these two ways are used to obtain values for effective masses, and are compared with others in the literature. The relative advantages of each method are discussed. Additionally, the band structure of alloys of the nitrides is also studied using the empirical pseudopotential method. In particular, the variation of the band gap with alloy composition is investigated. Consistent **k.p** parameters for alloys of various compositions are also calculated using the direct method, and constrasted with those obtained using the fitting technique.

Chapter 5

In this chapter, the behaviour of AlN, GaN and InN under strain is investigated. In particular, the empirical pseudopotential method is used to study band gap deformation potentials and the variation of the uppermost valence bands in general. The

valence band offsets of several nitride heterojunctions are also calculated using the first principles pseudopotential method. These calculations are done for both material A grown on material B and material B grown on material A. The resulting band offsets are discussed with respect to each other and in relation to others in the literature.

Conclusions

Finally, conclusions are made in Chapter 6. Suggestions for further work are also made.

Chapter 2

The First Principles Method

2.1 Introduction

One of the most popular tools available to study the electronic structure of matter is the first principles pseudopotential method. Techniques such as this, which require only a specification of the atoms present in the system being studied, are commonly referred to as *ab initio* methods, and their goal is to numerically solve the Schrödinger equation. This solution generally proceeds through an iterative search routine implemented on a computer, which terminates when solutions are converged to a sufficient predetermined accuracy. In practice, this approach has only been made tractable by advances in both computing power and mathematical techniques over the last 30 years.

In a search for a solution to the Schrödinger equation the most important quantity is the total energy. This is because an ability to calculate the total energy of a solid is the key to obtaining its physical parameters, since the minimisation of the total energy determines the equilibrium structure of that material. Indeed, by making appropriate use of total energies (or differences in total energies), such diverse quantities as lattice constants, bulk moduli, and piezoelectric constants have been accurately calculated. To obtain these quantities using the first principles pseudopotential method, successive improvements are made to the wavefunction describing the system until the total energy

is minimised to the required degree.

In forthcoming chapters alternative methods used to calculate the electronic structure of a material will be described, but in this chapter attention is focused on the first principles pseudopotential method. In this work, this has been used to investigate various properties of the nitrides. Initially, lattice constants are calculated using three different first principles codes (FHI96MD [28], CASTEP [29] and VASP [30–32] for both the zincblende and wurtzite structures of AlN, GaN and InN. For wurtzite the internal parameter u is also calculated, as is the crystal field splitting. The bonding in each of the structures is then investigated through population analysis. Valence band structures for each of the materials are also presented. Additional insight into these band structures is provided from density of states calculations.

2.2 Theory

Due to computational constraints, first principles pseudopotential calculations typically deal with systems containing several atoms per unit cell. The nuclei and electrons of these atoms all interact with each other, and the many body Hamiltonian describing this can be written

$$H = \sum_i \frac{-\hbar^2}{2M_i}\nabla_i^2 + \frac{1}{2}\sum_{i\neq j}\frac{Z_i Z_j e^2}{|\mathbf{R}_i - \mathbf{R}_j|} + \sum_k \frac{-\hbar^2}{2m_e}\nabla_k^2 + \frac{1}{2}\sum_{k\neq l}\frac{e^2}{|\mathbf{r}_k - \mathbf{r}_l|} - \sum_{k,i}\frac{Z_i e^2}{|\mathbf{r}_k - \mathbf{R}_i|} \tag{2.1}$$

where $\mathbf{R}_i$ and M_i denote the positions and masses of the ions, $\mathbf{r}_i$ denote the positions of the electrons and the Z_i are the atomic numbers of the atoms present. In this equation the units used are such that $1/4\pi\epsilon_0 = 1$. The terms in the many body Hamiltonian above describe, in order, the nuclear kinetic energy, the ion-ion interaction, the electronic kinetic energy, the electron-electron interaction and the electron-ion interaction.

In general, it is not possible to solve this equation exactly, and the first principles pseudopotential method is an attempt to provide an approximate solution.

The total energy pseudopotential method has been extensively described in several reviews [29,33,34]. It relies on several approximations to provide a solution to the many body problem outlined above. The first of these is the Born-Oppenheimer approximation, which exploits the large difference in mass between the electrons and the nuclei. Due to this mass difference, the response of the electrons to any motion of the nuclei will be essentially instantaneous, and thus the problem can be reduced to one of electrons moving in some set configuration of nuclei. Secondly, density functional theory as developed in [35,36] enables an approximate but nevertheless successful treatment of the electron-electron interactions. This effectively changes the problem from that of an interacting electron gas in the presence of nuclei to that of a single particle moving in an effective nonlocal potential. Lastly, the electron-ion interactions are successfully described using pseudopotential theory.

The main elements of density functional theory are outlined below. The concept of pseudopotentials will be discussed in more detail in chapter 3.

2.2.1 Density Functional Theory

The main thrust of density functional theory lies in its description of exchange and correlation in an electron gas. The exchange energy is a consequence of the requirement for antisymmetry of the electron wavefunctions with the same spin. This antisymmetry causes a spatial separation of the electrons which reduces the Coulomb energy of the system, and this reduction, computed using the Hartree-Fock approximation, is called the *exchange* energy. The Coulomb energy can be further reduced (at a cost in kinetic energy) if electrons with opposite spins are also spatially separated, and the difference between this and that calculated using the Hartreee-Fock approximation is termed the *correlation* energy.

2.2.1.1 The Kohn-Sham Total Energy Functional

Work by Hohenberg and Kohn [35] showed that the total energy of an electron gas is a unique functional of the electron density, and that the ground state energy of the system can be obtained through minimisation of the total energy functional with respect to the density. The density that yields this minimum is the exact single-particle ground state density. Within this formalism, Kohn and Sham [36] then proceeded to replace the formidable many-body problem with an exactly equivalent set of self-consistent one-electron equations.

The Kohn-Sham total energy functional can be written as follows:

$$\begin{aligned} E[\{\psi_i\}] &= 2\sum_i \int \psi_i \left[\frac{-\hbar^2}{2m}\right] \nabla^2 \psi_i d^3\mathbf{r} + \int V_{ion}(\mathbf{r}) n(\mathbf{r}) d^3\mathbf{r} \\ &+ \frac{e^2}{2} \int \frac{n(\mathbf{r}) n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3\mathbf{r} d^3\mathbf{r}' + E_{XC}\left[n(\mathbf{r})\right] + E_{ion}\left(R_I\right) \end{aligned} \tag{2.2}$$

where the terms are, respectively, the kinetic energy, the electron-ion interaction energy, the electron-electron interaction energy, the exchange correlation energy and the ion-ion interaction energy. $V_{ion}(\mathbf{r})$ is the static electron-ion potential and $n(\mathbf{r})$ is the electronic charge density given by

$$n(\mathbf{r}) = \sum_{i=1}^{N} \psi_i{}^*(\mathbf{r}) \psi_i(\mathbf{r}) \tag{2.3}$$

where N is the number of particles. The computational task of total energy calculations lies in determining the set of wavefunctions which minimise the total energy. This is achieved through solving the Kohn-Sham equations [36]:

$$\left[-\frac{\hbar^2}{2m}\nabla^2 + V_{ion}(\mathbf{r}) + V_H(\mathbf{r}) + V_{XC}(\mathbf{r})\right]\psi_i = \varepsilon_i\psi_i(r) \tag{2.4}$$

where ψ_i is the wavefunction of electronic state i, ε_i is the Kohn-Sham eigenvalue, $V_H(\mathbf{r})$ is the Hartree potential and $V_{XC}(\mathbf{r})$ is the exchange correlation potential. The Hartree and exchange correlation potentials can be written as

$$V_H(\mathbf{r}) = e^2 \int \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3\mathbf{r}' \tag{2.5}$$

$$V_{XC}(\mathbf{r}) = \frac{\delta E_{XC}[n(\mathbf{r})]}{\delta n(\mathbf{r})} \tag{2.6}$$

The exchange correlation potential is usually evaluated using the Local Density Approximation (LDA) described in more detail below.

2.2.1.2 The Local Density Approximation

The most common approach for the exchange correlation potential is the Local Density Approximation (LDA). The core assumption of this approach is that the density is slowly varying. The exchange-correlation at point $\mathbf{r}$ can therefore be simply approximated by that due to a homogeneous electron gas of density $n(\mathbf{r})$. Thus

$$E_{XC}[n(\mathbf{r})] = \int \varepsilon_{XC}(\mathbf{r}) n(\mathbf{r}) d^3r' \tag{2.7}$$

$$\frac{\delta E_{XC}[n(\mathbf{r})]}{\delta n(\mathbf{r})} = \frac{\partial n(\mathbf{r})\varepsilon_{XC}(\mathbf{r})}{\partial n(\mathbf{r})} \tag{2.8}$$

where

$$\varepsilon_{XC}(\mathbf{r}) = \varepsilon_{XC}^{hom}[\mathbf{n}(\mathbf{r})] \tag{2.9}$$

There are, of course, alternatives to the LDA. Of these, the most popular are those based around the generalised gradient approximation (GGA). In the GGA, the density is a function of both the electron density and its gradient, and for many materials, but not all, it has been shown to improve results for total energies and the general properties of solids. In the nitrides, however, the GGA has been shown to provide no significant improvement on the standard LDA [37] and therefore in all the calculations in this work the standard LDA is used.

2.2.1.3 Expansion using a plane wave basis set

The next task is to solve the problem of an infinite number of interacting electrons moving in the static field due to an infinite array of ions. For this, there are two difficulties which have to be overcome. The first is that a wavefunction has to be calculated for each of the infinite number of electrons. The second is that the basis set in which the wavefunctions will be expanded is also infinite.

At 0K, the ions in a perfect crystal are arranged in a regular periodic way. Thus the potential felt by the electrons due to the ions will also be periodic, with a period equal to that of the length of the unit cell, $\mathbf{l}$. Thus $V(\mathbf{r}) = V(\mathbf{r}+\mathbf{l})$ and Bloch's theorem is satisfied. The wavefunction of the infinite crystal can therefore be expressed in terms of the reciprocal space vectors of the Bravais lattice.

Bloch's theorem uses the periodicity of the crystal to reduce the problem of com-

puting a wavefunction for an infinite number of electrons to one of computing a wavefunction for each electron in the unit cell. The electronic wavefunctions are expanded in terms of a cell-periodic part and a wavelike part:

$$\psi_i(\mathbf{r}) = exp\left[i\mathbf{k}.\mathbf{r}\right] f_i(\mathbf{r}) \tag{2.10}$$

The cell periodic part $f_i(\mathbf{r})$ is then simply expanded using a discrete set of plane waves whose wave vectors are reciprocal lattice vectors of the crystal:

$$f_i(\mathbf{r}) = \sum_{\mathbf{G}} c_{i,\mathbf{G}} exp\left[i\mathbf{G}.\mathbf{r}\right] \tag{2.11}$$

where $\mathbf{G}$ are the reciprocal lattice vectors. These are defined by $\mathbf{G}.\mathbf{l} = 2\pi m$ for all $\mathbf{l}$, where $\mathbf{l}$ is a lattice vector of the crystal and m is an integer. Combining this with equation 2.10, each electronic wavefunction can be written as a sum of plane waves:

$$\psi_i(\mathbf{r}) = \sum_{\mathbf{G}} c_{i,\mathbf{k}+\mathbf{G}} exp\left[i((\mathbf{k}+\mathbf{G}).\mathbf{r}\right] \tag{2.12}$$

By using Bloch's theorem it is no longer necessary to consider an infinite number of electrons. Instead, the problem is mapped onto one of expressing the wavefunction for each electron in the unit cell in terms of an infinite number of reciprocal space k points within the first Brillouin zone. However, because the wavefunctions at k points which are very close together are essentially identical, it is possible to represent the electronic wavefunctions over some region of k space by the wavefunction at a single k point. Several schemes have been devised to best achieve this sampling, the most notable being those of Monkhorst and Pack [38] and Chadi and Cohen [39]. The denser the k-point set the more accurate the sampling, and so it is important to choose a set

which is sufficient for the calculation in question. Thus with k-point sampling the electronic wavefunctions can be expressed using only a finite number of k points.

The electronic wavefunctions at each such k point should in principle be expanded using an infinite basis set of reciprocal lattice vectors $\mathbf{G}$. However, in practice it is possible to use a finite basis set. This is because for the lower energy solutions the coefficients $c_{i,\mathbf{k}+\mathbf{G}}$ of the plane waves with small kinetic energy, given by $(\hbar^2/2m)\,|\mathbf{k}+\mathbf{G}|^2$, are generally larger than those with higher kinetic energy. Therefore a plane wave energy cutoff can be imposed to truncate the basis set to a finite size. This will result in an error in the total energy of the system, but this can be made arbitrarily small by increasing the cutoff energy. In practice, the cutoff needed depends upon the system being studied.

One final advantage of using a plane wave basis set is that the Kohn-Sham equation takes on a particularly simple form:

$$\sum_{\mathbf{G}} \left[\frac{\hbar^2}{2m} |\mathbf{k}+\mathbf{G}|^2 \delta_{\mathbf{G}\mathbf{G}'} + V_{ion}(\mathbf{k}+\mathbf{G}, \mathbf{k}-\mathbf{G}') + V_H(\mathbf{G}-\mathbf{G}') + V_{XC}(\mathbf{G}-\mathbf{G}') \right] \times c_{i,\mathbf{k}+\mathbf{G}'} = \varepsilon_i c_{i,\mathbf{k}+\mathbf{G}'} \quad (2.13)$$

In this reciprocal space representation the kinetic energy is diagonal and the potentials are described in terms of their Fourier components. This equation can be solved by traditional matrix diagonalisation techniques, with the Hamiltonian matrix elements $H_{\mathbf{k}+\mathbf{G},\mathbf{k}-\mathbf{G}}$ given by the term in square brackets. For larger and more complicated systems standard diagonalisation becomes intractable, and alernative techniques have been developed.

The Kohn-Sham equations are solved self consistently; that is, from a starting guess for the charge density the wavefunctions are calculated, and are then used to give a new charge density, which in turn gives new wavefunctions. This process is continued until consistency is reached. To assist the convergence it is important to mix the new

and old charge densities [40]. A flowchart illustrating the basic points in the procedure is shown in Figure 2.1.

2.2.2 The Pseudopotential Approximation

It was shown above how plane waves were used in a Fourier expansion of the wavefunction. Normally, plane waves would be ill suited to expanding the electronic wavefunctions because a very large number would be required to accurately describe the rapidly oscillating wavefunctions of electrons in the core region. The pseudopotential method relies on the fact that most of the physical properties of solids are determined primarily by the valence electrons. The pseudopotential approximation expolits this fact to replace the core electrons and strong nuclear potential with a much weaker pseudopotential, and is described in detail in Chapter 3.

2.2.3 Geometry Optimisation

In many calculations, the relaxed lattice parameters and atomic positions of a material are required. The Kohn-Sham equations allow for the minimisation of the total energy with respect to the electronic degrees of freedom. However, to obtain the true minimum in the total energy of a material, it is also necessary to include the ionic degrees of freedom. Thus in such calculations the ions must be allowed to move within a set scheme. For this to be possible, the forces acting on each ion must be evaluated, and this can be incorporated into the total energy pseudopotential method.

The force acting on an ion F_I can be obtained from the full derivative of the total energy of the system with respect to the ionic position $\mathbf{R}_I$:

$$\mathbf{F}_I = -\frac{dE}{d\mathbf{R}_I} \tag{2.14}$$

The ion is then allowed to move, and the electronic wavefunctions change to the

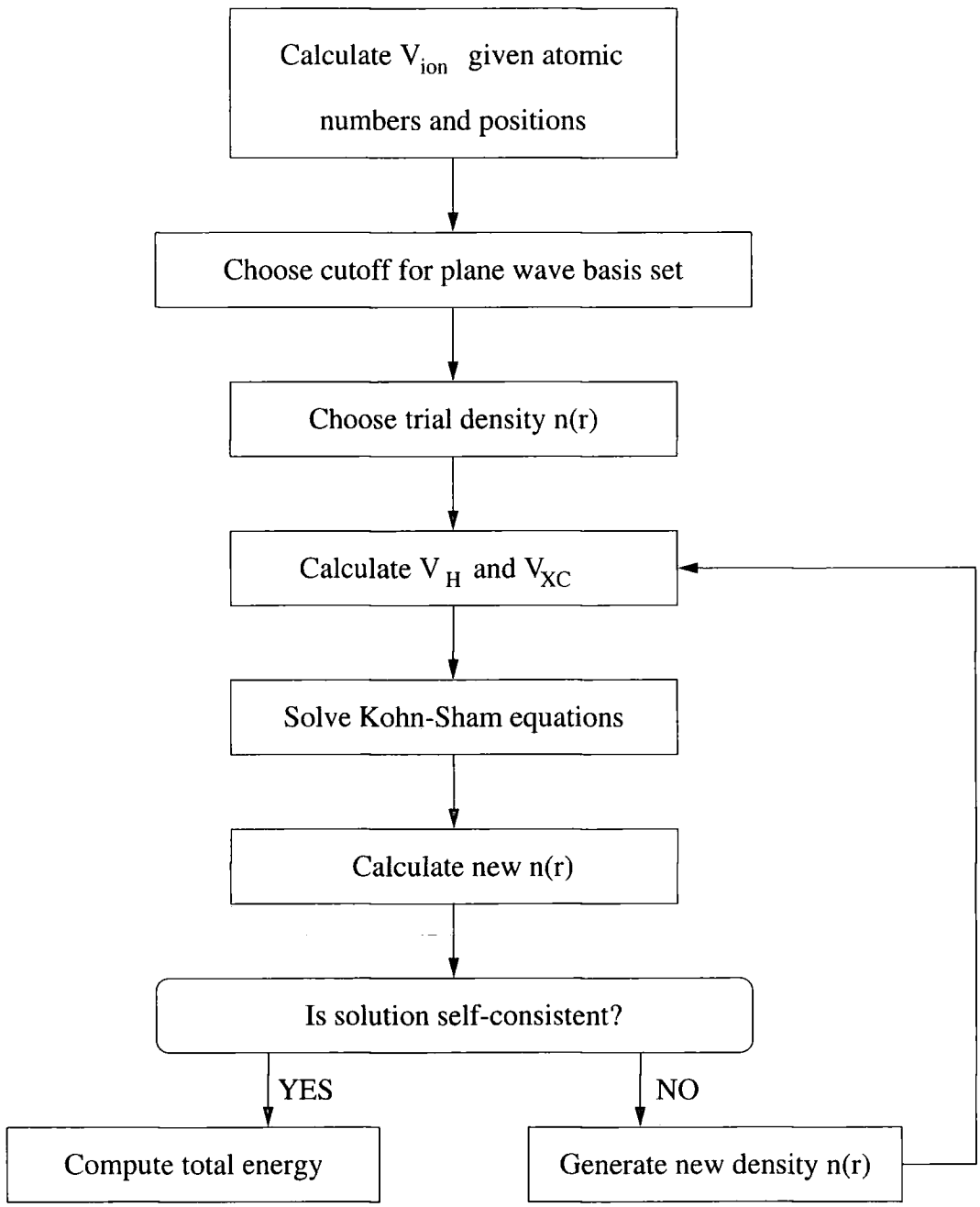

Figure 2.1: Flow chart describing the computational procedure to calculate the total energy of a solid in the first principles pseudopotential method.

self-consistent Kohn-Sham eigenstates corresponding to its new position. The force on the ion changes because of the changes in the electronic wavefunctions, as can be seen by expanding the derivative in equation 2.14:

$$\mathbf{F}_I = -\frac{\partial E}{\partial \mathbf{R}_I} - \sum_i \frac{\partial E}{\partial \psi_i}\frac{d\psi_i}{d\mathbf{R}_I} - \sum_i \frac{\partial E}{\partial \psi_i^*}\frac{d\psi_i^*}{d\mathbf{R}_I} \tag{2.15}$$

The Hellman-Feynmann theorem states that when each electronic wavefunction is an eigenstate of the Hamiltonian the last two terms in the equation above vanish, and the force felt by the ion is just the partial derivative of the total energy with respect to the ionic position. Thus in a search for the minimum total energy the ions are allowed to move and the total energy minimum corresponding to the new ionic configuration is calculated. The new forces acting on the ions are computed, and when these forces decrease to below a specified level the ionic structure is deemed to be relaxed. Similarly, the stress on a unit cell can also be evaluated. The unit cell can then be relaxed according to this stress in a similar way, and a description of this can be found in [41].

2.3 Bulk Calculations

First principles calculations have been performed on AlN, GaN and InN. Three different codes have been used for this task - FHI96MD [28], CASTEP [29] and VASP [30–32]. The results of the calculations using VASP are included for comparison, and were carried out by Steve Pugh. The three codes use different types of pseudopotentials. FHI96MD uses pseudopotentials generated according to the scheme of Troullier and Martins [42] whereas VASP and CASTEP use ultra-soft Vanderbilt pseudopotentials [43]. For GaN and InN, these pseudopotentials treated the d electrons as valence just as the s and p electrons. This was necessary because the Ga $3d$ and the In $4d$ states

have been shown to play an active part in chemical bonding [44–46]. In the rest of this chapter, all the calculations have been performed using CASTEP unless otherwise stated.

For the results in this section, it should be noted that for both the zincblende and wurtzite structures of InN a band gap of zero was obtained. This is in line with other calculations [37], and does not appear to adversely affect the calculated lattice parameters or densities of states.

2.3.1 Convergence Tests

To ensure reasonable accuracy in first principles calculations it is necessary to first perform convergence tests for each of the materials. As discussed in section 2.2.1, the two main factors which directly affect the accuracy of the calculation are the plane wave energy cutoff and the k-space sampling. The k-space sampling was done with points generated according to the Monkhorst-Pack scheme [38]. Figure 2.2 shows the convergence of the computed total energy with respect to these two parameters for wurtzite GaN. From this, it can be seen that the total energy is rapidly convergent, with a plane wave energy cutoff of 500 eV and 14 k space sampling points being sufficient to converge the total energy to better than 1meV/atom. Similar calculations were carried out for the other materials and structures, ensuring that a converged representation was used for each.

2.3.2 Equilibrium Lattice Parameters

AlN, GaN and InN can crystallise in both the zincblende and wurtzite structures. The hexagonal wurtzite structure is usually referred to as the α-phase whereas the zincblende structure is known as the β-phase. The zincblende and wurtzite structures are very similar; both structures have tetrahedral bonding and the difference between them is in the stacking order of the layers. In the zincblende form the stacking ar-

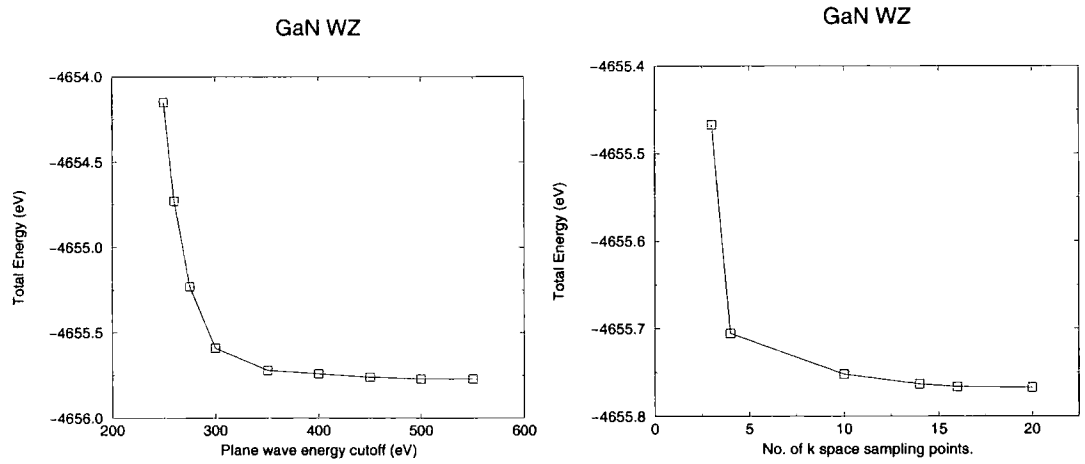

Figure 2.2: Convergence of the total electronic and ionic energy at 0K for wurtzite structure GaN with respect to the plane wave energy cutoff and the number of k-points used.

rangement is ABCABC, while in the wurtzite structure the arrangement is ABAB.

2.3.2.1 Zincblende Structure Materials

The zincblende structure is shown in Figure 2.3. It can be regarded as two interpenetrating face-centred cubic Bravais lattices displaced along the body of the diagonal by one quarter the length of the diagonal. The primitive lattice vectors are described by one lattice constant a and can be written as

$$\mathbf{a}_1 = \begin{pmatrix} 0 \\ 1/2 \\ 1/2 \end{pmatrix} a \quad ; \quad \mathbf{a}_2 = \begin{pmatrix} 1/2 \\ 0 \\ 1/2 \end{pmatrix} a \quad ; \quad \mathbf{a}_3 = \begin{pmatrix} 1/2 \\ 1/2 \\ 0 \end{pmatrix} a \tag{2.16}$$

The lattice parameters for the nitride materials in the zincblende form calculated using various first principles codes are shown in Table 2.1, together with generally accepted experimentally determined values. From this table, it can be seen that the lattice constants calculated in this work agree with these accepted values to within 2%.

The lattice constant for InN calculated using FHI96MD is unusual in that it is

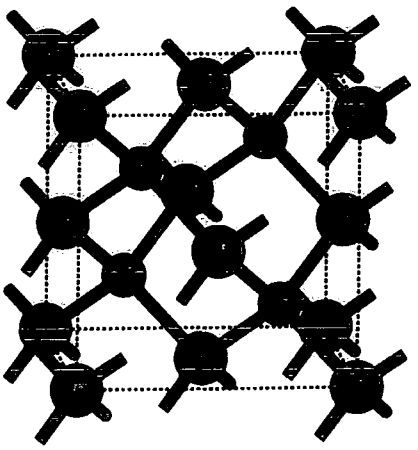

Figure 2.3: The crystal structure of zincblende GaN.

	FHI96MD	CASTEP	VASP	Other Theory	Experiment
AlN	4.30	4.31	4.34	4.35 [60]	4.38 [61]
GaN	4.42	4.46	4.45	4.46 [62]	4.50 [63]
InN	5.00	4.98	4.96	4.94 [60]	4.98 [64]

Table 2.1: Equilibrium lattice parameters in Angstroms calculated for zincblende AlN, GaN and InN.

larger than the experimentally determined value, since in general the LDA gives lattice constants which are lower than those found from experiment. As a whole, the results produced by CASTEP are in better agreement with experiment than those of the other two codes.

2.3.2.2 Wurtzite Structure Materials

The wurtzite crystal structure shown in Figure 2.4 is described by the space group $P6_3mc$ (C_{6v}^4). It can be regarded as an interpenetration of two hexagonally close-packed lattices displaced along the z-direction by uc where u is an internal parameter ideally equal to 3/8 and c is a lattice constant. For most wurtzite crystals, u deviates

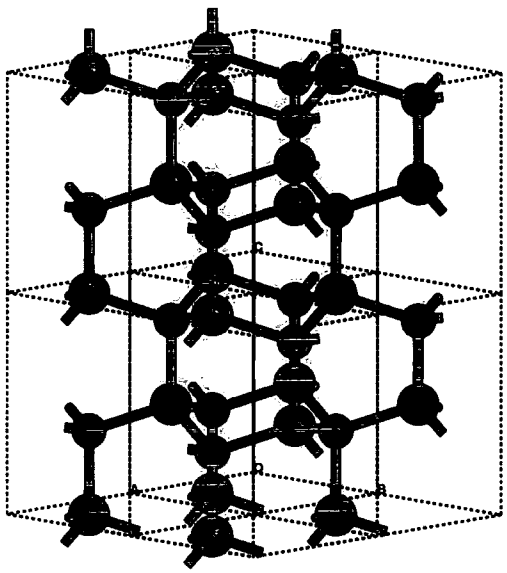

Figure 2.4: The crystal structure of wurtzite GaN.

from its ideal value, resulting in two slightly different nearest neighbour distances. The primitive lattice vectors are described by two lattice constants a and c, and can be written as

$$\mathbf{a}_1 = \begin{pmatrix} \sqrt{3}/2 \\ 1/2 \\ 0 \end{pmatrix} a \quad ; \quad \mathbf{a}_2 = \begin{pmatrix} -\sqrt{3}/2 \\ 1/2 \\ 0 \end{pmatrix} a \quad ; \quad \mathbf{a}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} c \tag{2.17}$$

The lattice parameters for the nitride materials in the wurtzite form calculated using various first principles codes are shown in Table 2.2, together with the generally accepted experimentally determined values. As for the zincblende structure, the lattice constants a and c calculated in this work agree with those from experiment to within 2%. For AlN, VASP produces the lattice parameters in the best agreement with experimental data. For GaN and InN, however, VASP produces acceptable values for a

but substantially underestimates c. A reason for this could be that the corresponding value of the internal u parameter obtained using VASP is large. For both GaN and InN the a and c values of CASTEP are in better agreement with experiment. For the internal parameter u, CASTEP produces values which are in the best agreement with those calculated by Wei and Zunger [52] for all three materials.

Values for the crystal field splitting, which splits the uppermost valence bands, are also shown in Table 2.2. Note that the value of Δ_{cf} for AlN is negative. This is discussed in more detail in chapter 4 with respect to the ordering of the valence bands. The values for the crystal field splitting calculated in this work using CASTEP show a decrease in the size of the splitting going from AlN to InN. The values obtained from VASP also follow this trend, and both sets are in approximate agreement with those from the literature. No values for the u parameter or the crystal field splitting were obtained from the calculations performed using FHI96MD, since this code does not allow an explicit relaxation of the unit cell. To calculate the lattice parameters using the FHI96MD code, total energies had to be calculated over a mesh of a and c values. The relaxed lattice parameters a and c, corresponding to the minimum of the total energy over the mesh, were then extracted.

		FHI96MD	CASTEP	VASP	Experimental values
AlN	a	3.063	3.064	3.079	3.110 [16]
	c	4.881	4.900	4.926	4.980
	u	-	0.382	0.380	0.3819 [52]
	Δ_{cf}	-	-0.246	-0.262	-0.219 [52], -0.058 [53]
GaN	a	3.133	3.155	3.149	3.190 [16]
	c	5.106	5.150	5.098	5.189
	u	-	0.376	0.377	0.3768 [52]
	Δ_{cf}	-	0.066	0.061	0.042 [52], 0.072 [53]
InN	a		3.526	3.521	3.544 [54]
	c		5.712	5.664	5.718
	u	-	0.378	0.380	0.3790 [52]
	Δ_{cf}	-	0.028	0.042	0.041 [52], 0.028 [55]

Table 2.2: Equilibrium lattice parameters in Angstroms calculated for wurtzite AlN, GaN and InN. Shown also are the crystal field splittings (in eV) calculated in this work and others in the literature.

2.3.3 Population Analysis

CASTEP can also be used to investigate the bonding in bulk crystals in more detail. In first principles calculations, the limitation of using a delocalised plane wave basis set is that it cannot be directly applied to the study of local atomic properties. However, Sanchez-Portal *et al.* [56,57] showed that it is possible to project the set of these valence electron wavefunctions onto a set composed of a linear combination of atomic orbitals (LCAO). The quality of this projection is given by the spilling parameter, which simply measures the fraction of the valence charge that is missed in the projection. Once the projection is done, bonding can be investigated through population analysis.

Population analysis is done following the formalism due to Mulliken [58], and gives such quantities as the charge on the anion and cation in the crystal. One must be careful in interpreting these results, since the absolute magnitudes of the atomic charges have little physical meaning, displaying an extreme sensitivity to the atomic basis set with which they are calculated. From the anion and cation charges, effective valence ionic valences can be obtained. These are defined as being the difference between the formal ionic charge and the Mulliken charge on the anion species in the crystal, and serve as a measure of the ionicity. Similarly, overlap populations between atoms give direct information on the covalency of the bonding [59].

Within the context of bonding, Mulliken electronegativities are traditionally used as a guide to ionicity. These electronegativities are defined as

$$X_M = \frac{A + I}{2} \tag{2.18}$$

where A is the electron affinity of an atom of the species and I is its ionization energy. The difference in the Mulliken electronegativies of two species provides an indication of the bonding between the atoms of each species; the larger the difference in the electronegativities, the greater the degree of ionicity.

Thus from calculations of this type, several indicators of the nature of bonding in crystals can be obtained.

2.3.3.1 Zincblende Structure Materials

Material	Spilling parameter	Anion charge (|e|)	Cation ++ charge (|e|)	Effective valence charge (|e|)
AlN	7×10^{-3}	-1.33	1.33	1.67
GaN	6×10^{-4}	-0.93	0.93	2.07
InN	9×10^{-4}	-0.94	0.94	2.06

Table 2.3: Spilling parameters and valence charges calculated for zincblende AlN, GaN and InN.

Table 2.3 shows the spilling parameters and atomic charges evaluated in the calculations. From this table, the spilling parameters obtained were of the order of 10^{-3} for all three materials. This indicates that the projection of the plane wave basis set onto the LCAO basis set was successful, with only 0.1% of the valence charge being missed in the projection, and that therefore good representations of the electronic bands were obtained. It can also be seen that the charge transfers for GaN and InN are approximately the same, and are significantly less than that found in AlN. Not shown in the Table 2.3, for GaN and InN the bulk of the charge transfer came from the s and p electrons, with very small change for the d electrons. However, although there is very little transfer of d electrons, it has previously been shown that they are important in bonding [45]. The basis for this is that they are close in energy to the uppermost N s states and are thus able to hybridise with them. This is discussed in more detail below in relation to density of states calculations.

Overlap populations vary from 0 to 1, with values near 1 indicating high covalency and values near 0 indicating high ionicity. For the overlap populations shown in Table 2.4, there is no clear trend for the three materials, with each having a value of approximately 0.50. This value is intermediate, indicating that the bonding can properly be described as either ionic or covalent. The effective valence charges and the differ-

ences in the Mulliken electronegativities, ΔX_M, also suggest this. However, the trend in the Mulliken electronegativity differences is not reflected in the either the overlap populations or the effective valence charges.

Material	Nearest neighbour overlap population (|e|)	Nearest neighbour bond length (Å)	ΔX_M (eV) Ref. [60]
AlN	0.51	1.865	4.08
GaN	0.56	1.932	4.15
InN	0.47	2.156	4.25

Table 2.4: Overlap populations and bond lengths calculated for zincblende AlN, GaN and InN.

2.3.3.2 Wurtzite Structure Materials

Material	Spilling parameter	Anion charge (|e|)	Cation charge (|e|)	Effective valence charge (|e|)
AlN	8×10^{-3}	-1.36	1.36	1.64
GaN	6×10^{-4}	-0.94	0.94	2.06
InN	9×10^{-4}	-0.96	0.96	2.04

Table 2.5: Spilling parameters and valence charges calculated for wurtzite AlN, GaN and InN.

Table 2.5 shows the spilling parameters and atomic charges obtained in the population analysis calculations. Again, low spilling parameters for all three materials indicate a successful projection of the PW states onto the LCAO basis states. From the table, it can be seen that the charges on the anion and cation are essentially the same as for the zincblende structure. As was found for the zincblende structure, the charge transfer for GaN and InN was almost exclusively from the s and p electrons, and given the similarity between the zincblende and wurtzite structures this is to be expected. Additionally, the effective valence charges of the three materials again indicate intermediate covalency for the bonding.

The overlap populations for the wurtzite materials are shown in Table 2.6. Note that two overlap populations and bond lengths are given. This is because the bond in

the c direction is slightly longer than the three "in-plane" bonds and thus has different values for these two quantities. For GaN and especially AlN, there is a large difference in the overlap populations for the two bonds. Some insight into this can be gained from consideration of the bond lengths, which are also shown in Table 2.6. For AlN, the difference in the bond lengths is quite large (corresponding to the large u value for this material), and this is reflected in the overlap populations. For GaN, the bond lengths do not differ by quite as much, and correspondingly the overlap populations are more similar. However, InN has overlap populations that are approximately equal, despite having quite different bond lengths.

It is also worth discussing the overlap populations for the wurtzite materials with reference to those obtained for the zincblende structure. This can be done after consideration of the number of the nearest neighbours in the two structures. In zincblende, the tetrahedral bonding is such that its 4 nearest neighbours are all the same distance away. For wurtzite however, only 3 of the bonds are of equal length, with the bond along the c direction being slightly longer. However, an *average* bond length and overlap population can be calculated for wurtzite. If this is done, the average bond length between the anion and cation in wurtzite is slightly smaller than that found in zincblende; values of 1.864 Å, 1.932 Å and 2.155 Å respectively are obtained for AlN, GaN and InN. Similarly, overlap populations of 0.50, 0.56 and 0.46 are obtained for AlN, GaN and InN respectively. These values are directly comparable to the overlaps obtained for the zincblende structure of 0.51, 0.56 and 0.47. Thus the results for wurtzite are fully consistent with those for zincblende.

The differences in the Mulliken electronegativity, ΔX_M, also suggest intermediate bonding for the nitride materials. The absolute value of ΔX_M increases from AlN to InN. This trend is not repeated in the nearest neighbour overlap populations in-plane, but does manifest itself in the nearest neighbour overlaps in the c direction.

Material	Nearest neighbour in-plane		Nearest neighbour in c direction		ΔX_M (eV)
	Overlap population (|e|)	Bond length (Å)	Overlap population (|e|)	Bond length (Å)	Ref. [60]
AlN	0.59	1.861	0.23	1.874	4.08
GaN	0.60	1.930	0.43	1.936	4.15
InN	0.46	2.153	0.46	2.162	4.25

Table 2.6: Overlap populations and bond lengths calculated for wurtzite AlN, GaN and InN. Note that in wurtzite the tetrahedral bonding is such that the bond in the c direction is longer, resulting in a lower overlap population.

2.3.4 Density of States

Much of the work contained in this thesis is concerned with the valence band structure. Insight into this, such as the s, p and d composition of the bands, can be gained from density of states (DOS) calculations. These have been performed following on from the population analysis in the previous section, and the spilling parameters obtained there give an indication of the accuracy of the DOS calculations. For all the materials considered here these were very low, indicating that the densities of states presented here are reasonably accurate.

The density of states can be evaluated as:

$$g(E) = \sum_n g_n(E) = \sum_n \int \frac{d^3\mathbf{k}}{4\pi^3} \delta(E - E_{n\mathbf{k}}) \tag{2.19}$$

2.3.4.1 Zincblende Structure Materials

The valence band structures of zincblende AlN, GaN and InN are shown in Figure 2.5. The corresponding DOS are shown in Figures 2.6, 2.7 and 2.8. In each of these figures, the upper panel shows the DOS for the binary compound. The middle and lower panels show the partial DOS for the anion and cation respectively.

For AlN, it can be seen that the top of the valence band is composed primarily from the N p orbital. The contribution to the top of the valence band from Al is relatively

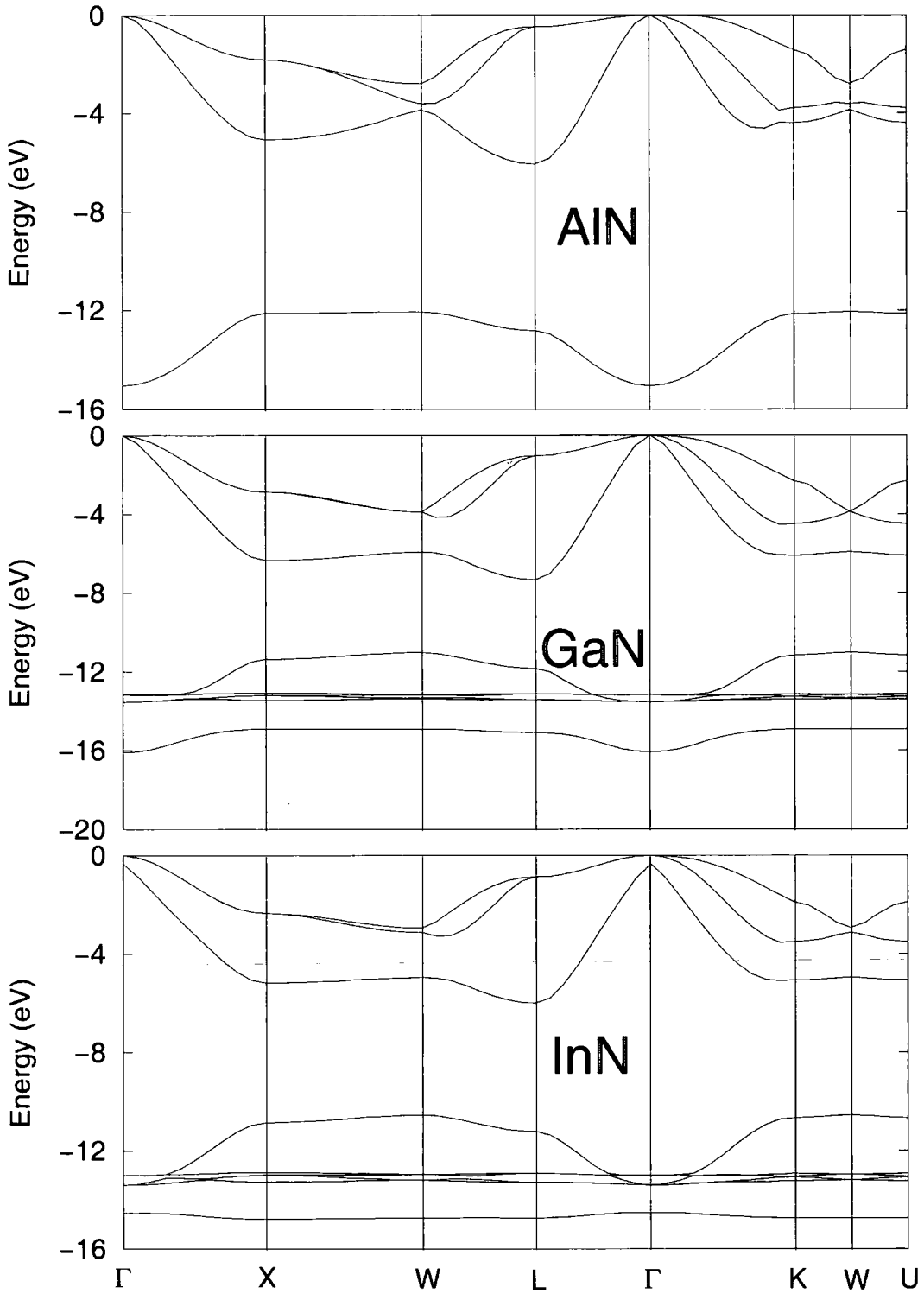

Figure 2.5: The valence band structures throughout the Brillouin zone calculated using CASTEP for zincblende structure AlN, GaN and InN.

small, being mainly due to the p orbital. For GaN and InN the $3d$ and $4d$ electrons respectively are treated as valence, and this has a significant effect on the density of states. Again, the top of the valence band is dominated by N p orbitals, though for both GaN and InN the contributions from the cation are more significant than in AlN. As can be seen from the insets in the middle panels in Figures 2.11 and 2.12, these contributions arise mainly from the cation p orbital, with the s orbitals taking part to a lesser extent. The d orbitals only make a very minor contribution to the top of the valence band.

However, the d electrons have a much larger effect on the lower bands. For both GaN and InN, the lower bands shown in Figures 2.11 and 2.12 are comprised mainly of Ga $3d$ and In $4d$ orbitals, with additional contribution from N s orbitals. It can be seen that for these bands the d orbitals are close in energy to the N s orbitals, indicating that hybridisation is possible. However, no noticeable d contributions are found from the N site, suggesting that the degree of hybridisation is small. This is backed up by the atomic populations calculated in the previous section. If the populations on each site are broken down into their constituent s, p and d parts, only about 0.01 d electrons from the Ga (and 0.03 from the In) are transferred, and none of this appears on the N site population. In the literature, one previous calculation for zincblende GaN in which the Ga $3d$ electrons were treated as valence found more evidence for hybridisation [45].

2.3.4.2 Wurtzite Structure Materials

The valence band structures of wurtzite AlN, GaN and InN are shown in Figure 2.9. The corresponding DOS are shown in Figures 2.10, 2.11 and 2.12. In each figure, the upper panel shows the DOS for the binary compound. The middle and lower panels show the partial DOS for the cation and anion respectively.

The density of states for wurtzite AlN, GaN and InN are very similar to those for the zincblende structure; at the bottom of the valence band the DOS is essentially indentical to that found in the zincblende structure, being dominated by the N s orbital

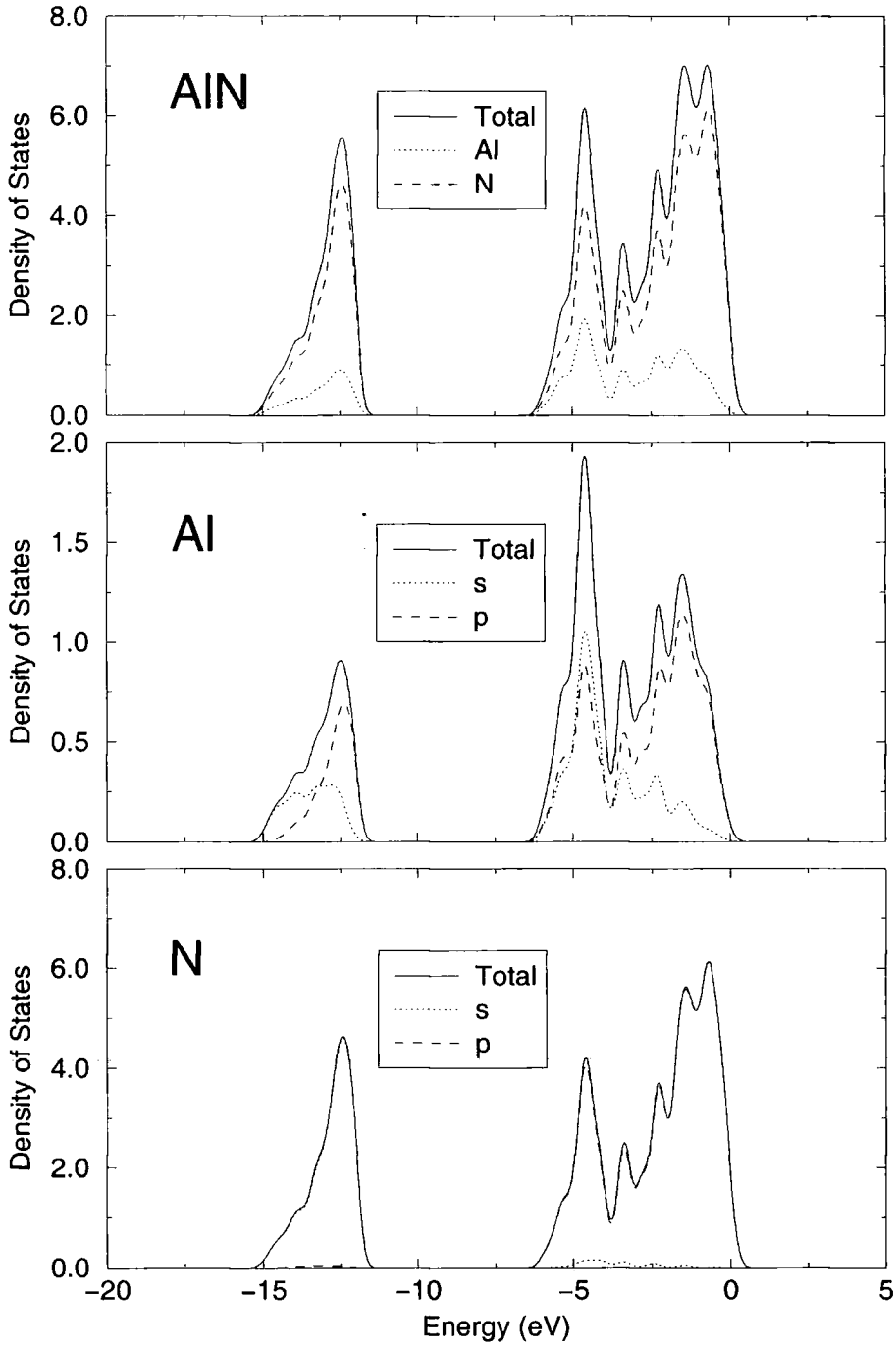

Figure 2.6: Electronic density of states (arbitrary units) for zincblende structure AlN. The upper panel shows the density of states for AlN while the middle and lower panels show the partial density of states for Al and N respectively.

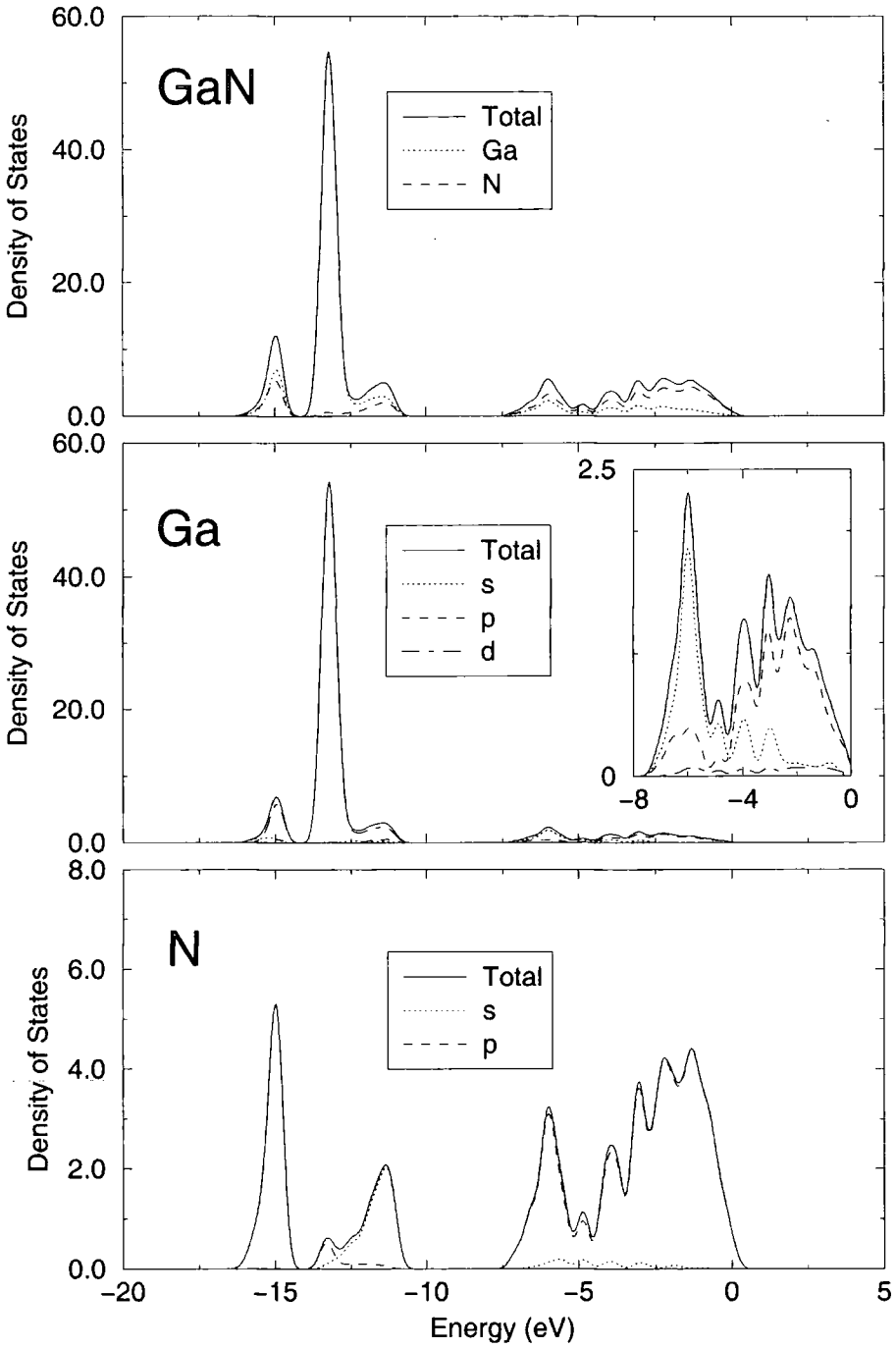

Figure 2.7: Electronic density of states (arbitrary units) for zincblende structure GaN. The upper panel shows the density of states for GaN while the middle and lower panels show the partial density of states for Ga and N respectively. Inset in the middle panel is a close up of the top of the valence band for Ga.

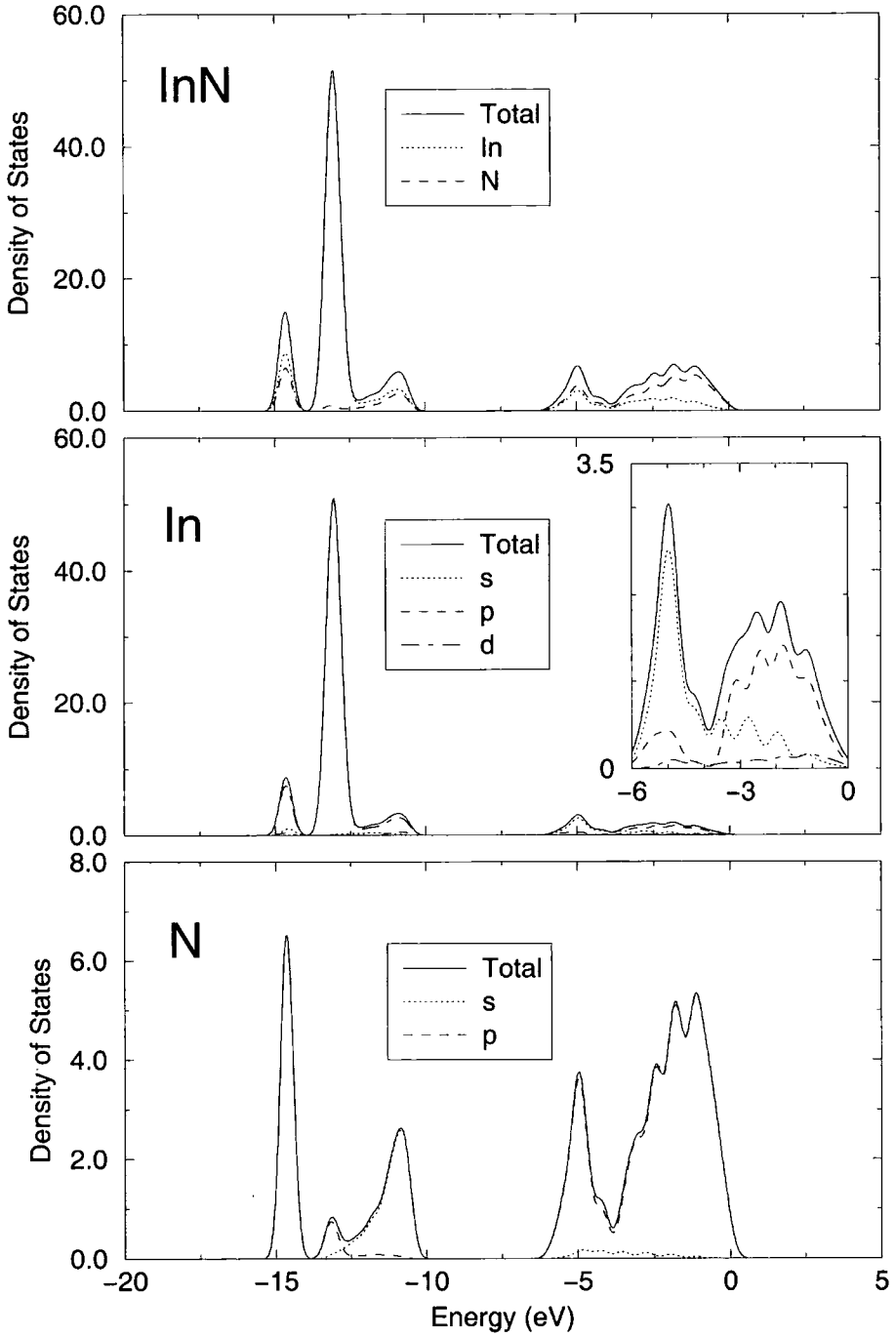

Figure 2.8: Electronic density of states (arbitrary units) for zincblende structure InN. The upper panel shows the density of states for InN while the middle and lower panels show the partial density of states for In and N respectively. Inset in the middle panel is a close up of the top of the valence band for In.

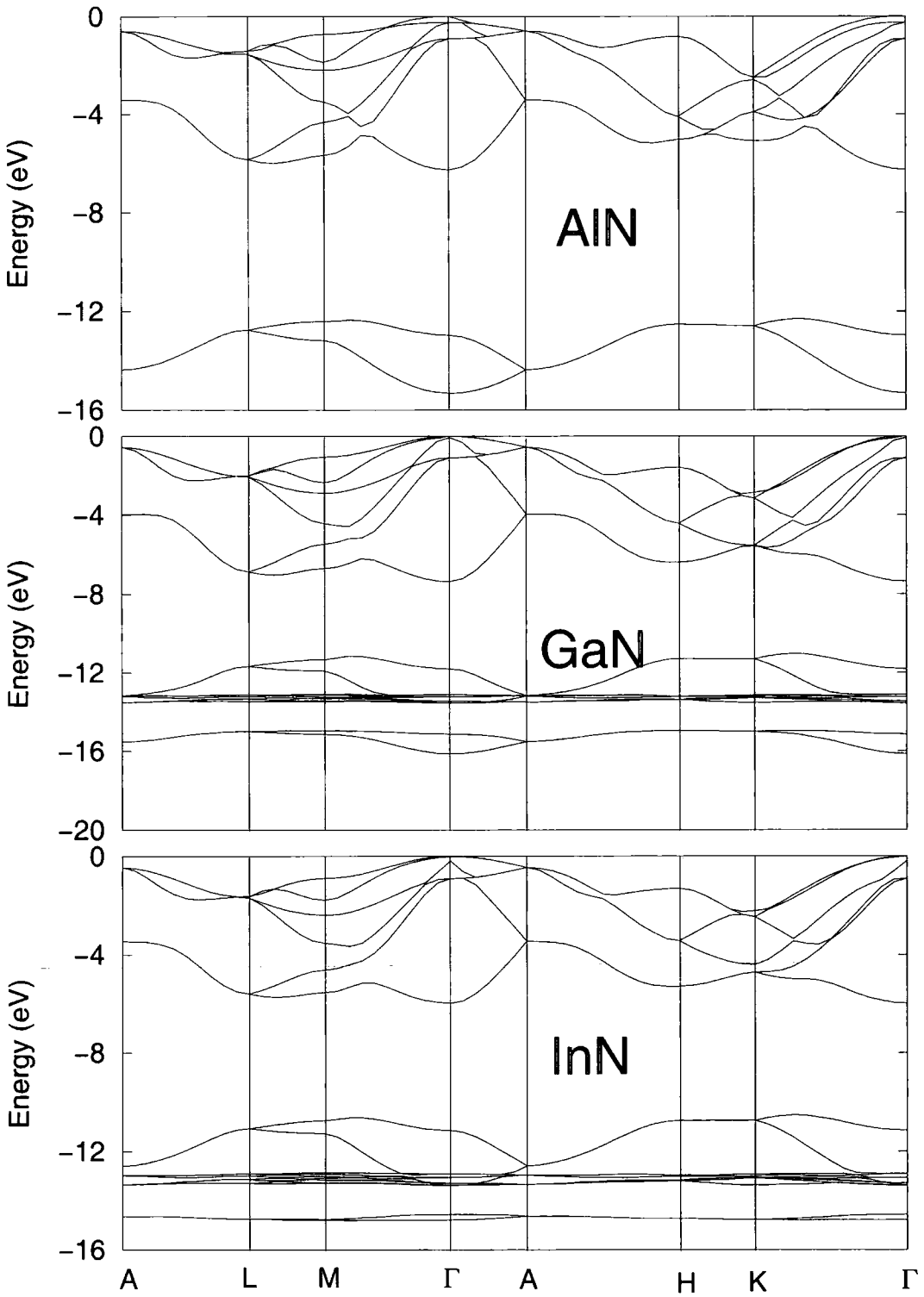

Figure 2.9: The valence band structures throughout the Brillouin zone calculated using CASTEP for wurtzite structure AlN, GaN and InN.

for AlN and by the cation d orbitals for GaN and InN. The top of the valence band is again dominated by the N p orbitals for all three materials. The most noticeable difference in the density of states between the two structures occurs at the very top of the valence band in the N p orbitals. In the zincblende structure there is a double peak whereas for wurtzite this feature becomes a large peak with a shoulder.

2.4 Summary

In this chapter, some basic properties of the nitrides have been investigated using first principles calculations. Initially, the lattice parameters describing the zincblende and wurtzite structures were determined. These were all in good agreement with experiment. Consistent crystal field splittings in the wurtzite structure were also calculated.

CASTEP was then used to further study the nitrides, starting with the bonding in each of the materials being investigated through population analysis. From this, effective valence charges and overlap populations indicated that the bonding in the nitrides is intermediate between ionic and covalent. The bulk of the charge transfer from the cation to the anion came from s and p electrons, although there was also a very small amount d transfer for GaN and InN. CASTEP was also employed to calculate the valence band structure and density of states for each of the materials. This study found the top of the valence band to be dominated by the N p orbital, whereas the bottom was dominated by either the N s orbital (in the case of AlN) or d orbitals (in the case of GaN and InN). These d electrons are close in energy to the N p states, and can thus potentially hybridise with them [61]. However, no evidence for this was found in this work.

In summary, several parameters describing the nitrides have been calculated. These were found to be consistent with both experimental results and previous calculations, establishing it as a firm basis for further calculations in the following chapters.

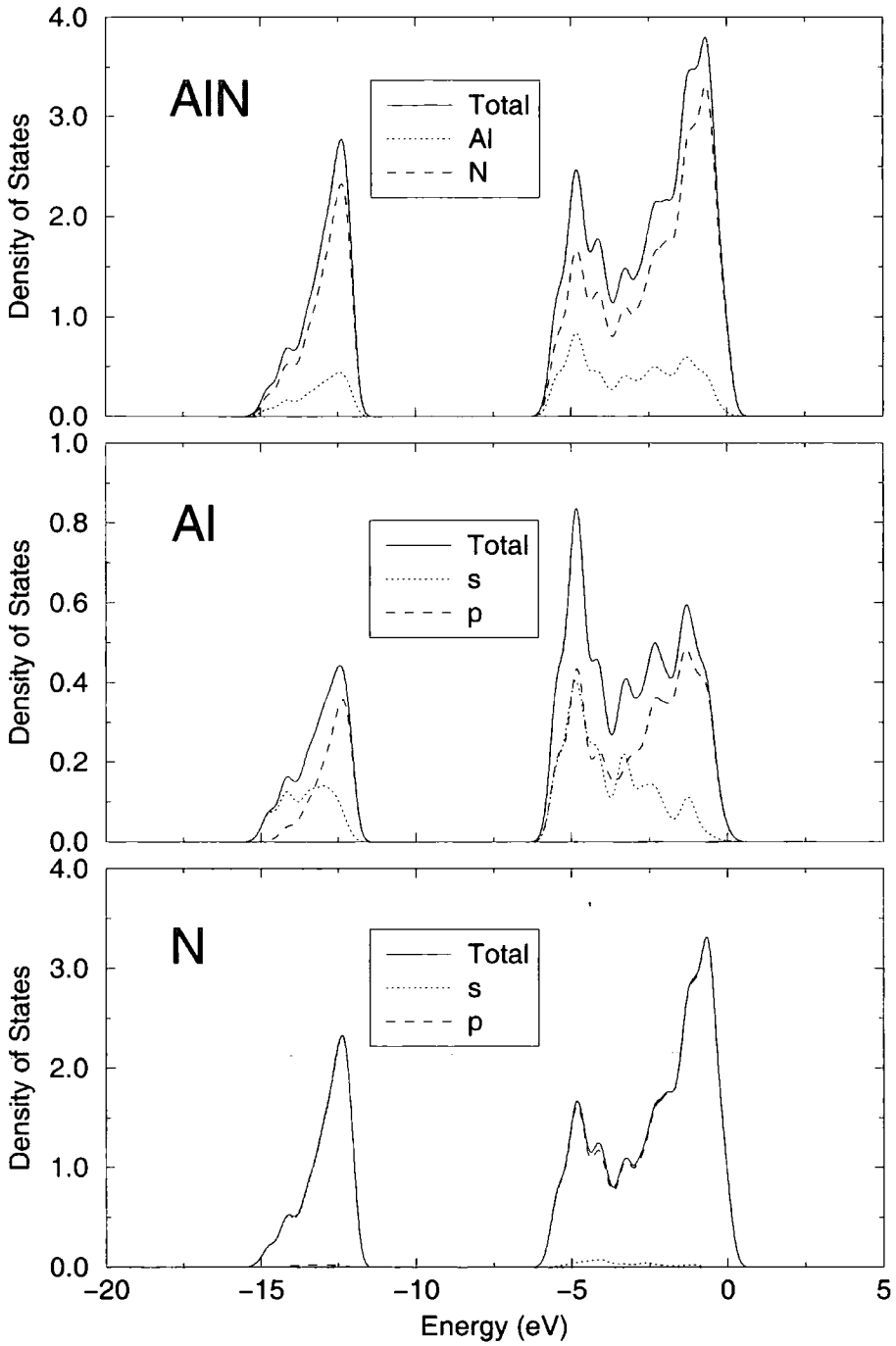

Figure 2.10: Electronic density of states for wurtzite structure AlN. The upper panel shows the density of states for AlN while the middle and lower panels show the partial density of states for Al and N respectively.

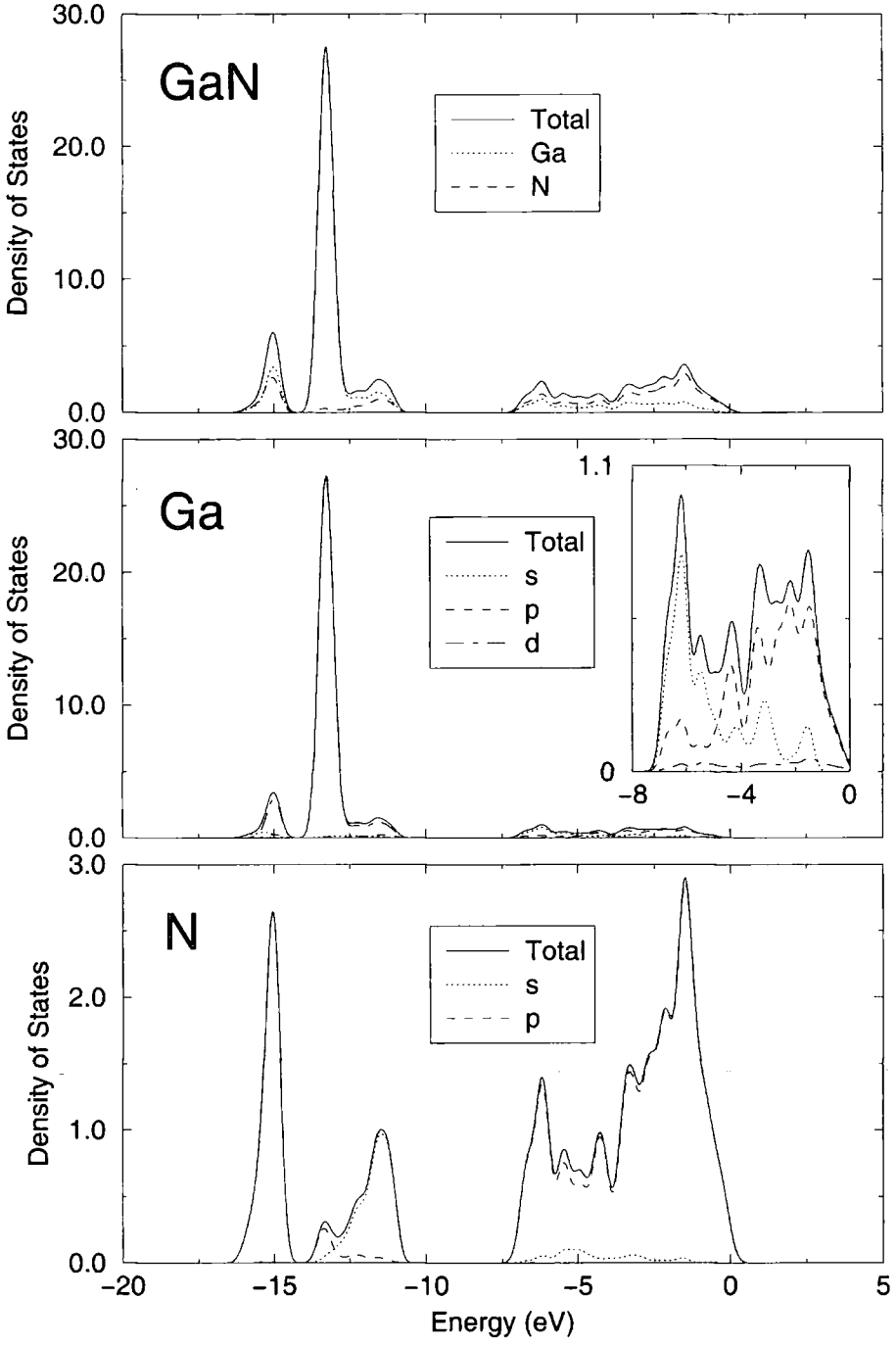

Figure 2.11: Electronic density of states for wurtzite structure GaN. The upper panel shows the density of states for AlN while the middle and lower panels show the partial density of states for Al and N respectively. Inset in the middle panel is a close up of the top of the valence band for Ga.

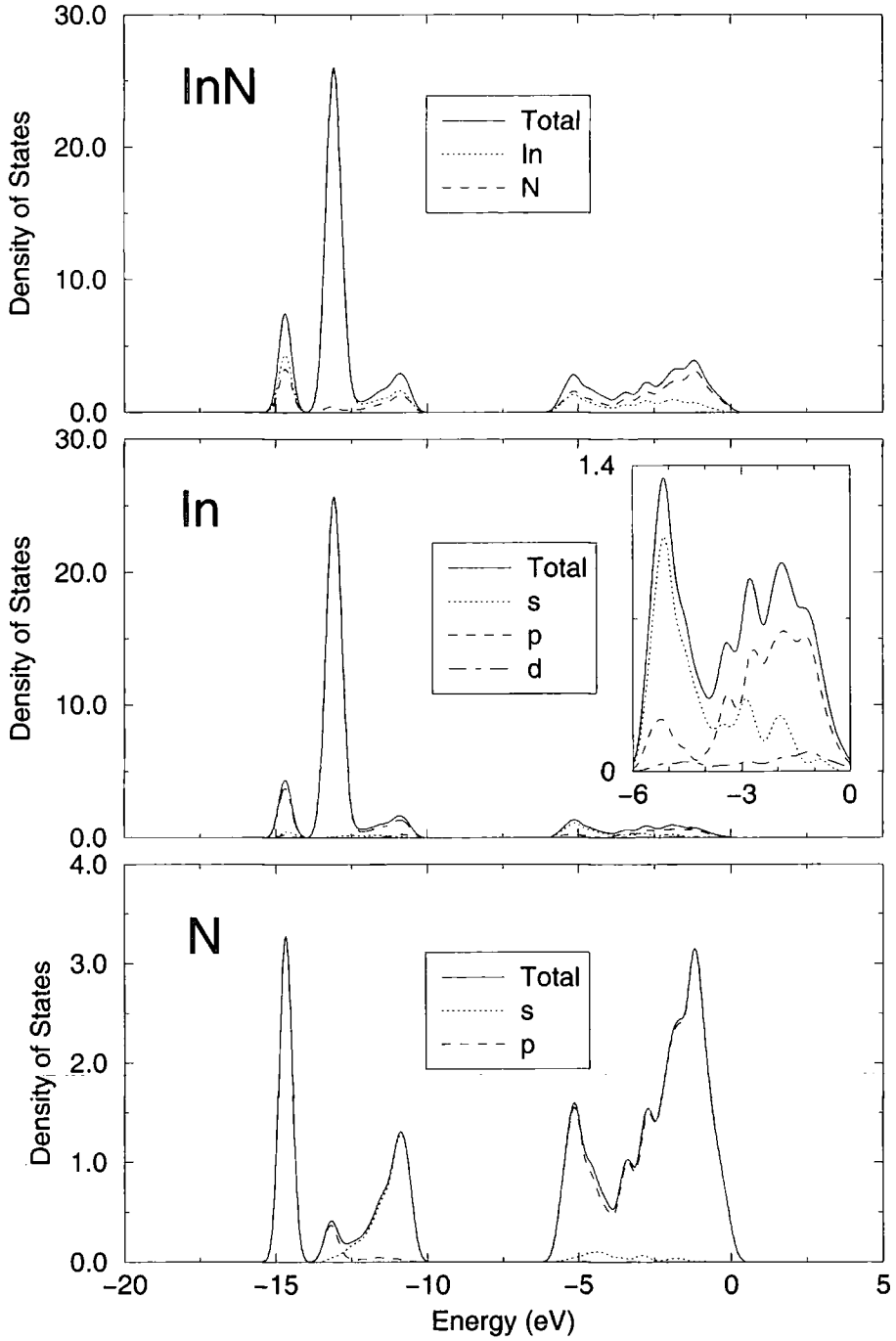

Figure 2.12: Electronic density of states for wurtzite structure InN. The upper panel shows the density of states for AlN while the middle and lower panels show the partial density of states for Al and N respectively. Inset in the middle panel is a close up of the top of the valence band for In.

Chapter 3

The Empirical Pseudopotential Method

3.1 Introduction

Studies of electronic structure rely on a knowledge of the energies and wavefunctions of single electron states throughout the Brillouin zone. There are two general approaches for the solution of this, one such being the first principles (or *ab initio*) method that was described in chapter 2. The other type of approach is *empirical.* Unlike *ab initio* methods, empirical approaches require several parameters as input. Regardless of which approach is employed, the basic requirement is for the electronic structure to be known as a function of position in **k**-space. That is, we require

$$\begin{aligned} E &= E_n(\mathbf{k}) \\ \psi &= \psi_n(\mathbf{k}) \end{aligned} \tag{3.1}$$

where n is a band index.

In practice, the choice of which method to use is largely determined by the type of

calculation to be performed. For instance, it is obviously preferable to use a method which is reasonably fast in terms of CPU time if a large number of band structure calculations throughout the Brillouin zone are required.

The standard first principles pseudopotential method for calculating band structure was described in 2. Although it can be used to obtain accurate lattice constants and to give some insight, it does not give accurate conduction band energies. This is because it is a ground state theory with correspondingly empty conduction band states, and as such cannot be expected to do so. Therefore, *ab initio* methods do not necessarily give a good representation of the actual situation. An additional drawback of first principles methods is that they are very CPU intensive.

Fortunately, empirical methods are available that are less computationally expensive than *ab initio* calculations. One such approach is the Empirical Pseudopotential Method (EPM) [62–65]. In this, the relevant parameters are adjusted so as to accurately reproduce band structure data from another source, be it experiment or more sophisticated first principles calculations. The obvious advantage of this is that accepted valence and conduction band energies throughout the Brillouin zone can be guaranteed. Several properties of materials, such as the reflectivity, are dependent upon the wavefunctions. Therefore, if these properties are successfully reproduced by the empirical pseudopotential method then the wavefunctions can be regarded as being accurate. The even less computationally intensive 6 or 8 band **k.p** method, which is described in chapter 4, can also produce accurate band structure, but as a perturbation technique its use is restricted to a limited region in **k**-space rather than the entire Brillouin zone. Although more bands can be used in this method to increase its accuracy, this results in a corresponding increase in computational effort. Thus the empirical pseudopotential method is an acceptable choice to satisfy the requirements for reasonably fast and accurate band structure.

3.2 The Pseudopotential Method

All band structure methods are required to solve the one-electron Schrödinger equation:

$$\left[\frac{-\hbar^2}{2m_0}\nabla^2 + V(\mathbf{r})\right]\psi(\mathbf{r}) = E\psi(\mathbf{r}) \tag{3.2}$$

The difficulty in solving this lies in the potential term $V(\mathbf{r})$. This is the average potential felt by each electron due to the other electrons and ion cores, and it has the periodicity of the crystal lattice. In particular, it has the property of being strong and atomic-like near the cores and weak between them. It is instructive to first look at two models which adopt each of these extremes as their view of the potential.

The tight-binding method (TBM) starts by treating the electrons as being tightly bound to the nuclei of their atoms. In a crystal where these atoms are close together, the wavefunctions will overlap, and the electronic wavefunctions are therefore approximated using a linear combination of the atomic wavefunctions. Thus this method is also known as linear combination of atomic orbitals method (LCAO). Again, the use of more orbitals results in greater accuracy, but in practice a minimal set of orbitals is usually employed. This method provides acceptable valence band structure but cannot be relied on for excited states, which are of obvious importance in any study of the optical properties.

In contrast, the nearly free electron model (NFEM) views the valence electrons as a perturbed gas of completely free electrons. That is, the potential in the one-electron Hamiltonian of equation 3.2 is considered to be much smaller than the kinetic energy term, such that the energy can be expanded using perturbation theory. As such, it is an obvious over-simplification.

It is readily apparent that neither one of these views is entirely satisfactory, and so an acceptable compromise has to be reached between the two. This can be done using the pseudopotential technique, which exploits the fact that the valence states are

of primary importance in determining the crystal properties. This approach was first employed by Fermi [66] in 1934, and subsequently by Hellman [67]. However, it was not until the late 1950's that the pseudopotential concept really began to be exploited.

3.2.1 Elements of the Pseudopotential Method

The electronic configuration of Si is $1s^2 2s^2 2p^6 3s^2 3p^2$. These electrons can be identified as either *core* or *valence*. The $1s$, $2s$ and $2p$ orbitals are fully occupied and form the core shells. Electrons in the Si $3s$ and $3p$ shells are called valence electrons, and are involved in bonding with neighbouring atoms. These electrons are nearly free, in that they are screened from seeing the full nuclear charge by the filled core shells. Thus in this pseudopotential approximation only the valence electrons are considered, with the cores frozen in an atomic-like configuration.

The wavefunctions of the valence electrons must be orthogonal to those of the core electrons. Thus it is sensible to construct the basis states in such a way as to ensure this, and this is commonly done using the Orthogonalised Plane Wave (OPW) method [68]. Importantly, constructing the basis states in this way also allows a good description of the wavefunctions in a tractable number of plane waves, and this is discussed in more detail below.

3.2.1.1 Orthogonalised Plane Waves

In this approach, the basis states are constructed from a set of plane waves which have been orthogonalised to the atomic core states. Between the cores, the plane wave component is well suited to describe the weak potential that exists there. Near the cores, however, the orthogonalisation terms force the valence electron wavefunctions to adopt the next highest core state wavefunction, effectively acting to repel the valence electrons from the core. Thus these orthogonalisation terms act like a kind of repulsive potential, and when it is combined with the attractive core potential they almost cancel, leaving behind a net, weak effective potential. This is the *pseudopotential.* This

effect is detailed in the Phillips-Kleinmann cancellation theorem [62], which explicitly demonstrates how such an orthogonality potential can be constructed. It is instructive to complete this understanding with a mathematical description, and this is done below.

The true wavefunction is expressed as the sum of a smooth wavefunction ϕ and a sum over occupied core states ϕ_t:

$$\psi = \phi + \sum_t b_t \phi_t \tag{3.3}$$

Enforcing the wavefunction to be orthogonal to the core states, i.e. $\langle \phi_t | \psi \rangle = 0$, we obtain

$$\psi = \phi - \sum_t \langle \phi_t | \phi \rangle \phi_t \tag{3.4}$$

If we then operate on ψ with the Hamiltonian, with $H = -\hbar^2/2m_0\nabla^2 + V$ where V is the core potential, we obtain

$$H\phi - \sum_t \langle \phi_t | \phi \rangle E_t \phi_t = E\phi - E \sum_t \langle \phi_t | \phi \rangle \phi_t \tag{3.5}$$

which simplifies to

$$H\phi + \sum_t (E - E_t)\phi_t \langle \phi_t | \phi \rangle = E\phi \tag{3.6}$$

This can then be rewritten in such a way as to explicitly identify the repulsive potential V_R arising from the orthogonalisation terms:

$$(H + V_R) = E\phi \tag{3.7}$$

where

$$V_R = \sum_t (E - E_t)\phi_t \langle \phi_t | \phi \rangle \tag{3.8}$$

If we then split up H into its component kinetic and core potential terms, we obtain

$$\left(\frac{-\hbar^2}{2m_0}\nabla^2 + V + V_R\right)\phi = \left(\frac{-\hbar^2}{2m_0}\nabla^2 + V_{ps}\right)\phi = E\phi \tag{3.9}$$

Hence we now have an equation for ϕ, the *pseudowavefunction.* However, although this is a pseudowavefunction it is important to note that the energy E is *not* a pseudo energy. Rather, it is the true energy corresponding to the true wavefunction. In particular, in equation 3.9 the long-range attractive core potential V and the short-ranged repulsive potential V_R have now been added together, leaving just the *pseudopotential* V_{ps}.

Figure 3.1 shows a representation of the variation of the real potential and the pseudopotential with distance from the nucleus. In this, the real potential is quite deep and therefore strong near the core. The pseudopotential, however, is weak in this region, indicating that the cancellation between the real potential and the repulsive one due to the orthogonalisation terms is quite complete. Away from the core, the pseudopotential is long ranged and attractive, and identical to that of the real unscreened Coulomb potential.

The weak form of the pseudopotential in Figure 3.1 has obvious implications for the pseudowavefunctions. Figure 3.2 is a schematic representation of the real wavefunction

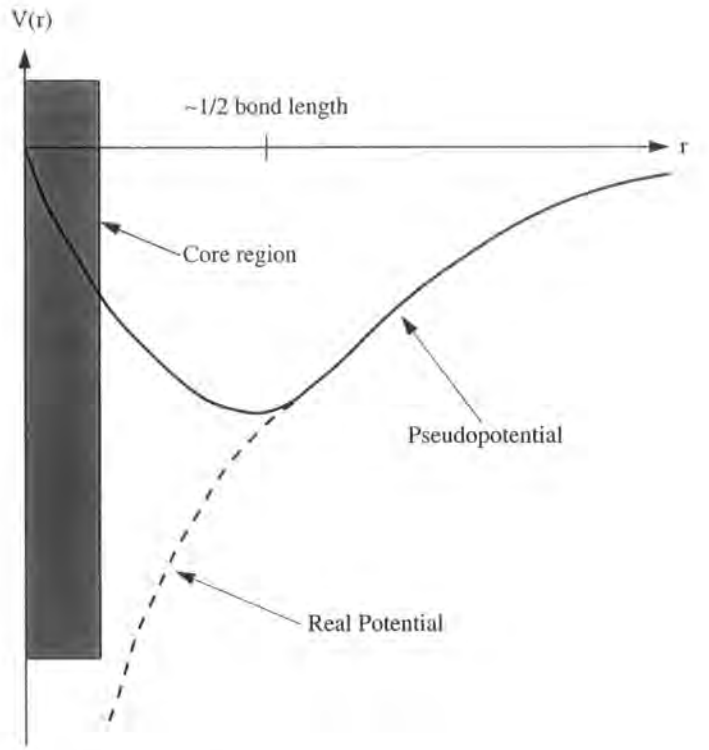

Figure 3.1: Schematic representation of the real-space pseudopotential. The solid line represents the pseudopotential and the dashed line represents the real potential.

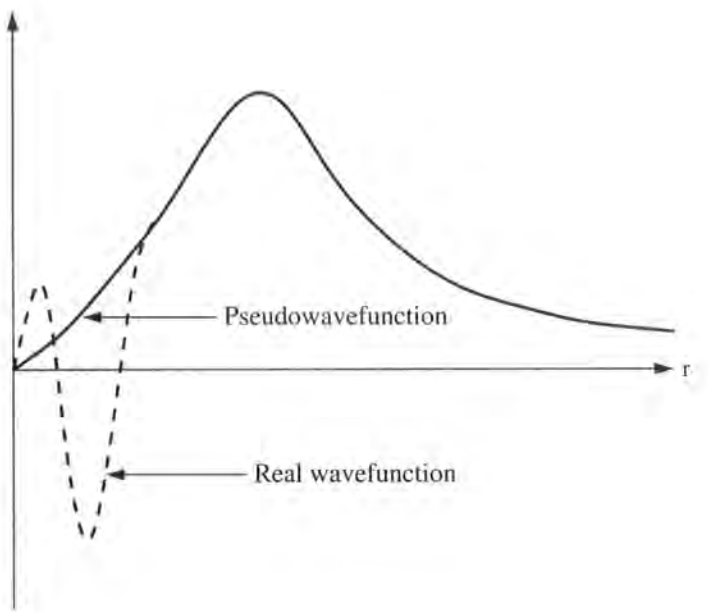

Figure 3.2: Schematic representation of the pseudowavefunction. The solid line represents the pseudowavefunction and the dashed line represents the real wavefunction.

and the pseudowavefunction, showing the variation with distance from the nucleus. From this, it can be seen that the pseudowavefunction is smooth, even in the core region. This contrasts with the real wavefunction, which has rapid spatial oscillations near the core.

These considerations serve to highlight the advantage of the pseudopotential approach. Without using pseudopotentials, these calculations would be computationally prohibitive. If a description of the true wavefunction was attempted using plane waves it would become clear that, because of the strong spatial oscillations of the wavefunction in the core region, a large number of plane waves would be required to provide an adequate representation. This number would be large enough to make such calculations intractable in terms of computing time. The advantage of using the pseudopotential lies in the resulting pseudowavefunction. This is smooth, and thus a comparatively small number of plane waves are needed in the basis state expansion to provide an acceptable description. The advantages of the pseudopotential are common not only to the empirical pseudopotential method but to *ab initio* techniques as well.

3.2.2 Solution of the Pseudo-Hamiltonian

The energies and pseudowavefunctions in the crystal are obtained by solving the Schrödinger equation:

$$\left(\frac{-\hbar^2}{2m_0} + V_{ps}\right)\psi = E\psi \tag{3.10}$$

The potential $V_{ps}(\mathbf{r})$ has to possess the periodicity of the lattice, such that

$$V(\mathbf{r}) = V(\mathbf{r} + \mathbf{R}_n) \tag{3.11}$$

where $\mathbf{R}_n$ is a lattice vector. This means that it can be written as a Fourier sum in terms of the reciprocal lattice vectors $\mathbf{G}_m$:

$$V(\mathbf{r}) = \sum_m V(\mathbf{G}_m) e^{i\mathbf{G}_m \cdot \mathbf{r}} \tag{3.12}$$

This is a local pseudopotential - there is no dependence on the angular momentum. Nonlocal pseudopotentials are discussed later. Bloch's theorem can then be employed to rewrite the wavefunction $\psi_k(\mathbf{r})$:

$$\psi_k(\mathbf{r}) = e^{\mathbf{ik}.\mathbf{r}} u_{\mathbf{k}}(\mathbf{r}) \tag{3.13}$$

where $u_{\mathbf{k}}(\mathbf{r})$ has the periodicity of the lattice. Like the potential, this can then be expanded as a Fourier sum employing N plane waves:

$$u_{\mathbf{k}}(\mathbf{r}) = \sum_{n=1}^{N} a_n(\mathbf{k}) e^{i\mathbf{G}_n \cdot \mathbf{r}} \tag{3.14}$$

When the potential in equation 3.12 and the wavefunction in equation 3.14 are put into the Schrödinger equation of 3.10, the problem is reduced to a standard matrix eigenvalue problem:

$$\begin{pmatrix} T(1) & V_{ps}(1,2) & \cdots & V_{ps}(1,N) \\ V_{ps}(2,1) & T(2) & \cdots & V_{ps}(2,N) \\ \vdots & \vdots & \ddots & \vdots \\ V_{ps}(N,1) & V_{ps}(N,2) & \cdots & T(N) \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_N \end{pmatrix} = E \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_N \end{pmatrix} \tag{3.15}$$

where the individual matrix elements are given by

$$\begin{aligned} T(i) &= \frac{\hbar^2}{2m_0}(\mathbf{k}+\mathbf{G}_i)^2 + \langle \mathbf{K}_i | V_{ps}(\mathbf{r}) | \mathbf{K}_i \rangle \\ V_{ps}(i,j) &= \langle \mathbf{K}_i | V_{ps}(\mathbf{r}) | \mathbf{K}_j \rangle \\ \mathbf{K}_n &= e^{i(\mathbf{k}+\mathbf{G}_n).\mathbf{r}} \end{aligned} \tag{3.16}$$

The solution of this proceeds using standard matrix diagonalisation routines. If N plane waves are used in the expansion at a given $\mathbf{k}$, N eigenvalues are obtained corresponding to the energies of the bands at that $\mathbf{k}$. Each of these eigenvalues also has an associated eigenvector whose coefficients $(a_1, a_2, ..., a_N)$ describe the wavefunction of that band at $\mathbf{k}$. Hence diagonalisation of the matrix results in N eigenvalues and N eigenvectors.

At this point, it is instructive to restate the advantage of the pseudopotential approach in quantitative terms. If the real potential was used in equation 3.15 then the rapid oscillations of the wavefunctions in the core would require an expansion involving of the order of 10^6 plane waves, corresponding to a $10^6 \times 10^6$ matrix. This essentially renders the problem insoluble from a computational point of view, since the number of calculations required to solve the Hamiltonian is proportional to N^3, where N is the order of the matrix. Use of the pseudopotential reduces this problem to a manageable one, with of the order of 50 plane waves per atom in the unit cell usually being sufficient to obtain a reasonably well converged representation. For the exact form of this pseudopotential, there are two main choices.

3.2.3 Choice of Pseudopotential

In the previous discussion on orthogonal plane waves, it was shown using the Phillips-Kleinmann [62] theorem in equation 3.7 how the pseudopotential could be constructed from a knowledge of the core states. However, although this theorem does serve to underline why the pseudopotential method works, in practice other techniques are generally used to obtain the pseudopotentials. For most such approaches, a parametrised

model for the potential is chosen, the final form of which is obtained by fitting to known data from experiment.

Cohen and Bergstresser [63] chose to adopt a simple local form for the pseudopotential $V_{ps}(\mathbf{r})$, where local means that there is no angular momentum dependence. This technique was successfully used for Ge and Si before being extended to other common semiconductors, and it is this form of the potential that is primarily used in this work. However, after consideration of the Phillips-Kleinmann formulation of the pseudopotential, it is obvious that it must possess some nonlocal character. This is because the pseudopotential is constructed from a summation over the core states, and can thus be split up into its constituent V_s, V_p, V_d etc. contributions. It becomes clear that if the core does not contain electrons of a certain angular momentum then there will be no repulsive potential for that component. Carbon, for example, has a core which is $1s^2$, and thus the p-valence electrons will feel no repulsive potential. It should be noted that in practice, however, the precise form of the local potential might be chosen to to give correct results for a given symmetry state. Therefore, even if there are no p or d core states it does not necessarily mean that there are no p or d components.

In addition to the local and nonlocal pseudopotentials, the effects of spin-orbit coupling must also be included, leaving a final form for the matrix elements of the pseudopotential as:

$$\langle \mathbf{K}_i | V_{ps} | \mathbf{K}_j \rangle = \langle \mathbf{K}_i | V_l + V_{nl} + V_{so} | \mathbf{K}_j \rangle \tag{3.17}$$

where V_l is the local pseudopotential, V_{nl} is the nonlocal pseudopotential and V_{so} is the contribution due to spin-orbit coupling. These are all discussed below.

3.2.4 The Local Pseudopotential

For one atom per primitive unit cell, the pseudopotential can be written as

$$V_l(\mathbf{r}) = \sum_m V(\mathbf{G}_m) e^{i\mathbf{G}_m \cdot \mathbf{r}} \tag{3.18}$$

The Fourier components $V(\mathbf{G})$ of this, which are commonly referred to as the *pseudopotential form factors*, are given by

$$V(\mathbf{G}) = \frac{1}{\Omega} \int_\Omega V(\mathbf{r}) e^{-i\mathbf{G}.\mathbf{r}} d\mathbf{r} \tag{3.19}$$

where Ω is the volume of the primitive unit cell. Different numbers of form factors are employed for the zincblende and wurtzite structures, and a description of how these are obtained can be found in Appendix A.

3.2.5 The Nonlocal Pseudopotential

It was earlier stated that in principle the pseudopotential should possess some nonlocal character, even though in actual calculations these contributions are often ignored. The inclusion of nonlocal effects in the pseudopotential method is described in the literature by various authors [63,69–72]. In this work, the nonlocal pseudopotential is incorporated by placing spherical potential wells around each ion, each of which act on a different angular momentum component of the wavefunction. The matrix element contribution from the nonlocal pseudopotential is written in the form

$$\langle \mathbf{K}_i | V_{nl} | \mathbf{K}_j \rangle = \sum_l \langle \mathbf{K}_i | A_l(E) f_l(r) \hat{P}_l | \mathbf{K}_j \rangle \tag{3.20}$$

where $A_l(E)$ is the well depth, $f_l(r)$ describes the shape of the well, and $\hat{P}_l$ is a projection operator such that the l^{th} well acts only upon the l^{th} angular component of the

wavefunction. The summation is over the angular momentum components l present in the core wavefunctions. Thus for the materials considered in this work only the $l = 0, 1, 2$ components corresponding to the s, p and d states need to be considered. In practice, the p component can effectively be included in the local part of the pseudopotential, and thus the summation above only involves the $l = 0$ and $l = 2$ components.

The well depth $A_l(E)$ is in general a function of energy. It is usually sufficient to adopt a fixed depth A_2 for the d well, but a more complicated form is used for the s well depth. A description of this, and the square well shape for $f_l(r)$, can be found in [72].

The inclusion of a nonlocal component in the pseudopotential introduces a total of up to 5 new parameters per atomic type, some of which such as the interatomic separation are fixed. In this work, the nonlocal form of the pseudopotential was only used for the zincblende structure and not for the wurtzite structure. This was because for the wurtzite structure a lot of other parameters were already involved and it was found to offer no significant improvement on the entirely local approach.

3.2.6 Spin-Orbit Coupling

For the heavier elements, the atomic spin-orbit splitting becomes important, and must therefore be included in band structure calculations. For the III-V nitrides considered in this work, however, the spin-orbit splittings are small. This is due to the top of the valence band being dominated by nitrogen, as was shown in chapter 2; since nitrogen is a light element it has a correspondingly small spin-orbit effect. Therefore, spin-orbit effects in the II-V nitrides will be much smaller than those found in, for example, GaSb, where the spin-orbit splitting is approximately 0.75 eV.

There are several approaches which incorporate spin-orbit coupling into the pseudopotential formulation [72–74]. In the Bloom and Bergstresser formulation [74], the matrix element due to the spin-orbit interaction can be written as

$$\langle \mathbf{K}_i | V_{so} | \mathbf{K}_j \rangle = (\mathbf{K}_i \times \mathbf{K}_j).\sigma_{s,s\prime} \left[-i\lambda^S . S^S(\mathbf{G}_i - \mathbf{G}_j) + \lambda^A . S^A(\mathbf{G}_i - \mathbf{G}_j) \right] \tag{3.21}$$

where

$$\begin{aligned}
\mathbf{K}_i &= \mathbf{k} + \mathbf{G}_i \\
\mathbf{K}_j &= \mathbf{k} + \mathbf{G}_j \\
\lambda^S &= (\lambda_c + \lambda_a) \\
\lambda^A &= (\lambda_c - \lambda_a) \\
\lambda_c &= \mu B^c_{nl}(K_i) B^c_{nl}(K_j) \\
\lambda_c &= \alpha\mu B^a_{nl}(K_i) B^a_{nl}(K_j) \\
B_{nl}(K) &\propto \int_0^\infty j_{nl}(Kr) R_{nl}(r) r^2 dr
\end{aligned} \tag{3.22}$$

where $j_{nl}(Kr)$ is the spherical Bessel's function of the lth angular momentum component and $R_{nl}(r)$ is the radial part of the core wavefunction.

In this work, the matrix element contributions due to the anion and cation are more simply written as (see e.g. [64]):

$$\begin{aligned}
\langle \mathbf{K}_i | V^\alpha_{so} | \mathbf{K}_j \rangle &= -\mu i \langle \nu_i | \sigma | \nu_j \rangle . (\mathbf{K}_i \times \mathbf{K}_j) \\
\langle \mathbf{K}_i | V^\beta_{so} | \mathbf{K}_j \rangle &= -\alpha\mu i \langle \nu_i | \sigma | \nu_j \rangle . (\mathbf{K}_i \times \mathbf{K}_j)
\end{aligned} \tag{3.23}$$

There is essentially no difference between this and the more sophisticated approach adopted by Bloom and Bergstresser for the materials considered here. The parameter μ is an adjustable parameter, altered so as to reproduce the required valence band splittings.

The immediate effect of including the spin-orbit coupling in the calculations is a doubling in the size of the matrices involved. Without spin, the Hamiltonian matrices were $N \times N$, where N is the number of reciprocal lattice vectors used in the plane wave expansion. Including spin, the matrices become $2N \times 2N$ with an associated increase in computational load. This doubling in matrix size can be circumvented by treating the spin-orbit interaction as a perturbation, as is done by Chelikowsky and Cohen [72], but this technique is not employed in this work.

3.3 Output of the EPM

Once parameters describing the local and nonlocal parts of the pseudopotential have been obtained, as well as the spin-orbit splitting, solution of the $2N \times 2N$ matrix can proceed by standard diagonalisation methods. Energies and their corresponding pseudowavefunctions are output as

$$E_n(\mathbf{k}), \quad n = 1, \ldots, 2N \tag{3.24}$$

$$\mathbf{c}_n(\mathbf{k}), \quad n = 1, \ldots, 2N \tag{3.25}$$

where n is the band index. The pseudowavefunctions are vector quantities, and are generally complex numbers (but not in the case of IV-IV materials with the origin chosen as the bond centre):

$$\mathbf{c}_n(\mathbf{k}) = \left[{}^{\uparrow}c_{n,1}(\mathbf{k}) \ , \ldots , \ {}^{\uparrow}c_{n,N}(\mathbf{k}), \quad {}^{\downarrow}c_{n,1}(\mathbf{k}) \ , \ldots , \ {}^{\downarrow}c_{n,N}(\mathbf{k})\right] \tag{3.26}$$

The pseudowavefunction is then given by

$$\psi_n(\mathbf{k}) = u_n(\mathbf{k})e^{i\mathbf{k}.\mathbf{r}} = \frac{1}{\Omega}\left[\sum_{j=1}^{N}\left({}^{\uparrow}c_{n,j}|\uparrow\rangle + {}^{\downarrow}c_{n,j}|\downarrow\rangle\right)e^{i\mathbf{G}_j.\mathbf{r}}\right]e^{i\mathbf{k}.\mathbf{r}} \tag{3.27}$$

where $u_n(\mathbf{k})$ is the Bloch periodic part of $\psi_n(\mathbf{k})$ and Ω is the unit cell volume. The eigenfunctions $|\uparrow\rangle$ and $|\downarrow\rangle$ denote spin-up and spin-down states respectively.

3.3.1 Overlap Integrals

In many of the calculations contained in this work, such as those on the dielectric function presented later in this chapter, it is necessary to evaluate the overlap integral between the periodic parts of the Bloch functions at specific values of the wavevector. These overlaps are given by

$$\langle u_m(\mathbf{k}_1)|u_n(\mathbf{k}_2)\rangle = \sum_{j=0}^{N}\left[{}^{\uparrow}c^*_{m,j}(\mathbf{k}_1){}^{\uparrow}c_{n,j}(\mathbf{k}_2) + {}^{\downarrow}c^*_{m,j}(\mathbf{k}_1){}^{\downarrow}c_{n,j}(\mathbf{k}_2)\right] \tag{3.28}$$

At a given wavevector $\mathbf{k}$, the bands are all orthogonal to each other, and can furthermore be normalised such that

$$\langle u_m(\mathbf{k})|u_n(\mathbf{k})\rangle = \delta_{mn} \tag{3.29}$$

where δ_{mn} is the Kronecker delta function.

The calculation of such overlap integrals should obviously be performed using the real wavefunction, which can be obtained from the pseudowavefunction through equation 3.4. However, it is actually possible to use the pseudowavefunction in these calculations with little loss of accuracy. The reason for this lies in the pseudowavefunction being a good approximation to the real wavefunction in the region between the atoms

where the valence electrons are concentrated. This is because the overlap between these valence electron states and those in the core are small. Thus the additional terms introduced in obtaining the real wavefunction should be negligible in the context of the calculations considered in this work.

3.4 Pseudopotential Fitting

As was stated at the beginning of the chapter, the empirical pseudopotential method requires input in the form of band structure information. That is, the pseudopotential is specifically tailored to reproduce band structure in good agreement with that from another source. Earlier in this chapter, it was shown how pseudopotentials for both the zincblende and wurtzite structures could be constructed. These pseudopotentials are parametrised in terms of an appropriate number of symmetric and antisymmetric form factors, $V^s(G)$ and $V^a(G)$, at the discrete set of points in reciprocal space defined by the reciprocal lattice vectors $\mathbf{G}$. The aim in any fitting procedure is to obtain a complete set of these parameters which produces sufficiently accurate band structure. One simple approach is to randomly vary the individual parameters themselves until that is achieved. However, while schemes which do this to obtain the parameters for a particular material are acceptable, the use of the resultant form factors is restricted solely to that situation; they cannot easily be adapted for use in strained material or alloys.

In this work, smooth $V(q)$ functions have instead been constructed for both the zincblende and wurtzite structures of the nitride materials, from which the form factors at the relevant points $q = G$ could be extracted. The advantage this approach offers is that it allows pseudopotentials representing strained material and alloys to be readily obtained, something which is covered in detail in chapters 4 and 5. With the midpoint between the two atoms (in the zincblende case) taken as the origin of the unit cell, the symmetric and antisymmetric functional forms of the pseudopotential used in this

work are given below. The form for the symmetric part was originally suggested by Falicov and Lin [75].

$$\begin{aligned} V^s &= \frac{a_1 q^2 + a_2}{1 + exp(a_3[a_4 - q^2])} \\ V^a &= (a_1 q^2 + a_2) exp(a_3[a_4 - q^2]) \end{aligned} \tag{3.30}$$

in which q is in units of $(2\pi/a_{zb})$, where a_{zb} is the lattice constant of the zincblende form of the relevant material. The respective form factors $V(G)$, at the particular G-values required in the band structure calculations, can then be simply obtained from these expressions.

As a summary, the parameters used in the calculations for both the zincblende and wurtzite structures are shown in Table 3.1. Note that for the wurtzite structure an entirely local pseudopotential was sufficient to provide a good description of the materials. The parameter u in Table 3.1 was chosen to be in good agreement with accepted values in the literature, and to be generally consistent with the first principles calculations of chapter 2.

Having adopted this approach, an acceptable set of the parameters a_i (i=1–4) from which the form factors are calculated must be obtained. For this, a Monte-Carlo scheme is employed, in which the empirical pseudopotential method band structure is compared with target band energies of other calculations at high symmetry points in the Brillouin zone. In this work, the target band energies used were at the Γ, L and X points. For AlN and GaN, the band energies come from the quasiparticle corrected first principles calculations of Rubio *et al.* [76]. For InN, such results were not available, and so the results of the first principles calculations of chapter 2 using VASP were used, with an appropriate "scissoring" of the band gap to correctly reproduce accepted excited state energies. It is important to note, however, that quasiparticle corrections tend to produce a k-dependent correction to the LDA results. Therefore, if the eigenvalues for

Zincblende				
Fixed	Lattice	a		
	Nonlocal	R_0^c, R_2^c		
	Nonlocal	R_0^a, R_2^a		
Fitted	Local	$V^s(q)$	$\longrightarrow$	$V^s(\sqrt{3}), V^s(\sqrt{8}), V^s(\sqrt{11})$
		$V^a(q)$	$\longrightarrow$	$V^a(\sqrt{3}), V^a(\sqrt{4}), V^a(\sqrt{11})$
	Nonlocal	α_0^c, β_0^c	$\longrightarrow$	A_0^c
		α_0^a, β_0^a	$\longrightarrow$	A_0^a
		A_2^c		
		A_2^a		
	Spin-orbit	μ		
Wurtzite				
Fixed	Lattice	a, c, u		
Fitted	Local	$V^s(q)$	$\longrightarrow$	$V^s(\sqrt{3}), \ldots, V^s\left(\sqrt{14\frac{2}{3}}\right)$
		$V^a(q)$	$\longrightarrow$	$V^a(\sqrt{3}), \ldots, V^a\left(\sqrt{14\frac{2}{3}}\right)$
	Spin-orbit	μ		

Table 3.1: The fixed and fitted parameters of the empirical pseudopotential method for the zincblende and wurtzite structures.

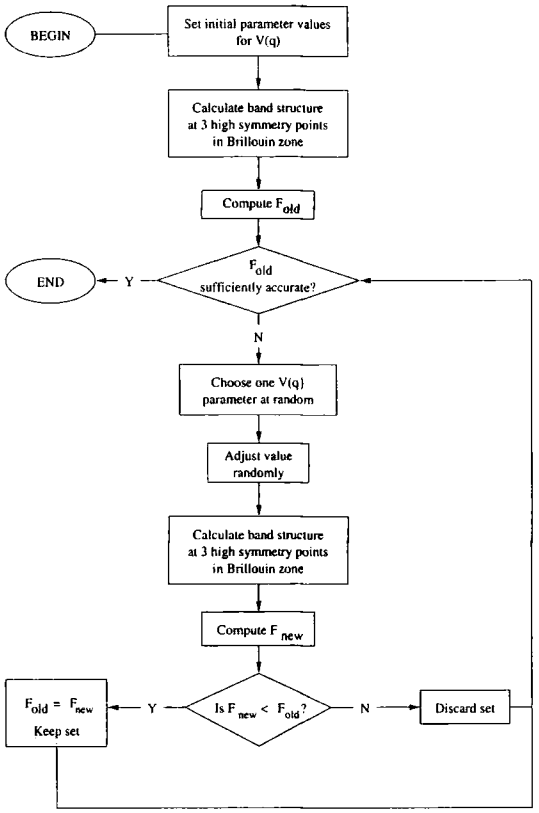

Figure 3.3: The Monte-Carlo algorithm used for the fitting of the pseudopotential parameters. See text for details.

InN across the entire Brillouin zone are rigidly shifted this will result in some error.

The a_i values in equation 3.30 were varied using a Monte-Carlo technique. The associated form factors were then calculated from these, and band structure calculations performed at the required points in the Brillouin zone. These computed band energies were then compared with the target energies, and the process systematically repeated until acceptable agreement was reached. This involved minimising the function F, given below:

$$F = \sum_i \sum_j w(i,j) \left| E_{epm}(i,j) - E_{target}(i,j) \right|^2 \tag{3.31}$$

The index i refers to the 3 high symmetry k-points used in the calculation whereas j is the band index. $w(i,j)$ is the weight attached to each particular band energy at each k-point. A schematic representation of the fitting procedure is shown in Figure 3.3

3.4.1 Requirements for the Fitted Pseudopotentials

In this work, it was important to reproduce accurate band structure for the top three valence bands and the first conduction band at the zone centre. In the wurtzite structure, the energies of the top three valence band states at the Γ point, split by the spin-orbit coupling and the crystal field, are [77]:

$$\begin{aligned} E_0 &= 0 \\ E_1 &= -\frac{1}{2}(\Delta_{so}+\Delta_{cr}) + \frac{1}{2}\sqrt{(\Delta_{so}+\Delta_{cr})^2 - \frac{8}{3}\Delta_{so}\Delta_{cr}} \\ E_2 &= -\frac{1}{2}(\Delta_{so}+\Delta_{cr}) - \frac{1}{2}\sqrt{(\Delta_{so}+\Delta_{cr})^2 - \frac{8}{3}\Delta_{so}\Delta_{cr}} \end{aligned} \tag{3.32}$$

where Δ_{so} is the spin-orbit splitting and Δ_{cr} is the crystal field splitting. Note that the zincblende structure has a higher symmetry than the wurtzite structure and has no such crystal field splitting. For the zincblende structure materials the bands are only split by the spin-orbit interaction.

The inclusion of spin results in a breaking of the degeneracy at the Γ point for the materials considered in this work. The triple degeneracy found in zincblende structure materials is split into a doublet and a singlet, whereas in the wurtzite structure the extra splitting due to the crystal field results in three separate states. A schematic diagram of this is shown in Figure 3.4.

The target band energies of Rubio *et al.* , and the first principles band structure of chapter 2 for InN, do not include spin in their calculations. Therefore, in the fitting procedure shown in Figure 3.3, spin was not included in the band structure calculations. Instead, the parameter μ, which accounts for the spin-orbit coupling Δ_{so} in the Hamiltonian, is manually altered after the fitting to give the correct splitting of the valence bands at the zone centre. However, the spin-orbit splitting for the three nitrides is not reliably known, and therefore in this work μ was altered to give a Δ_{so}

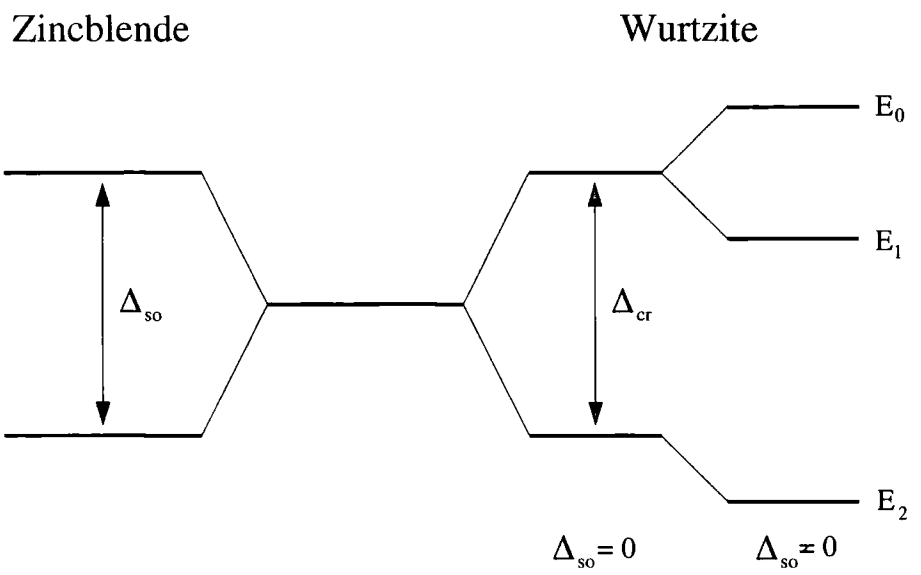

Figure 3.4: Schematic diagram of the valence band splittings due to spin-orbit coupling and crystal field effects for the zincblende and wurtzite structures at the Γ point. Note that the top band for the zincblende structure is a quadruplet. The diagram for wurtzite above is for a positive crystal field splitting as found in GaN and InN; for AlN, which has a negative crystal field splitting, the ordering of the bands is different. This can be seen in the band structures presented later in this chapter.

of 11 meV for all the materials, in accordance with Dingle *et al.* [21].

In the previous chapter, first principles calculations were used to calculate the crystal field splittings (Δ_{cr}) and other lattice parameters for the nitrides. However, as for the spin-orbit splitting there is some doubt as to what the correct values of Δ_{cr} should be. In this work, acceptable target values for the crystal field splittings were taken from various sources [21,52,53,55,77].

As well as producing acceptable band energies at the three high symmetry points in the Brillouin zone, several other requirements also had to be satisfied:

- There should be good qualitative agreement over the entire Brillouin zone between the empirical pseudopotential method band structure and that produced by the first principles calculations of chapter 2
- The $V(q)$ functions had to be smooth and physically reasonable

Moreover, for the wurtzite structure it was desirable to impose two further constraints:

	V^s				V^a			
	a_1	a_2	a_3	a_4	a_1	a_2	a_3	a_4
Zincblende								
AlN	0.0769	-0.4236	-0.3705	10.4965	0.0209	0.0940	0.1164	10.4270
GaN	0.0752	-0.5173	-0.5255	10.0740	0.0101	0.0824	0.1070	8.8359
InN	0.0732	-0.5273	-0.5732	9.5069	0.0052	0.0787	0.1016	8.2672
Wurtzite								
AlN	0.0767	-0.4478	-0.4831	10.7550	0.0018	0.1005	0.1736	10.7369
GaN	0.0837	-0.5643	-0.5485	10.5076	0.0148	0.0869	0.1265	9.3214
InN	0.0843	-0.6104	-0.5851	10.3746	0.0079	0.0857	0.0492	9.6443

Table 3.2: The coefficients of the form factor $V(q)$ functions defined in equation 3.30.

- The wurtzite $V(q)$ functions for each material had to bear some resemblance to their counterparts in the zincblende structure
- A reasonable deformation potential should be obtained for each of the three materials

The first point is desirable for consistency with the concept of transferable atomic potentials. That is, the pseudopotential is derived from the actual real space potentials of the constituent atoms, and thus there must be some correspondence in the pseudopotentials between the zincblende and wurtzite structures of the same material. The second point addresses the effect of strain on the band structure of the material, and is discussed in detail in chapter 5.

The final coefficients obtained from the fitting procedure described above are shown in Table 3.2. The resulting $V(q)$ form factor curves for each material are shown in Figure 3.5. The curves for the two structures are reasonably similar for each of the materials, with the exception of AlN. For this material, the antisymmetric $V(q)$ curve for the wurtzite structure is perhaps a bit large.

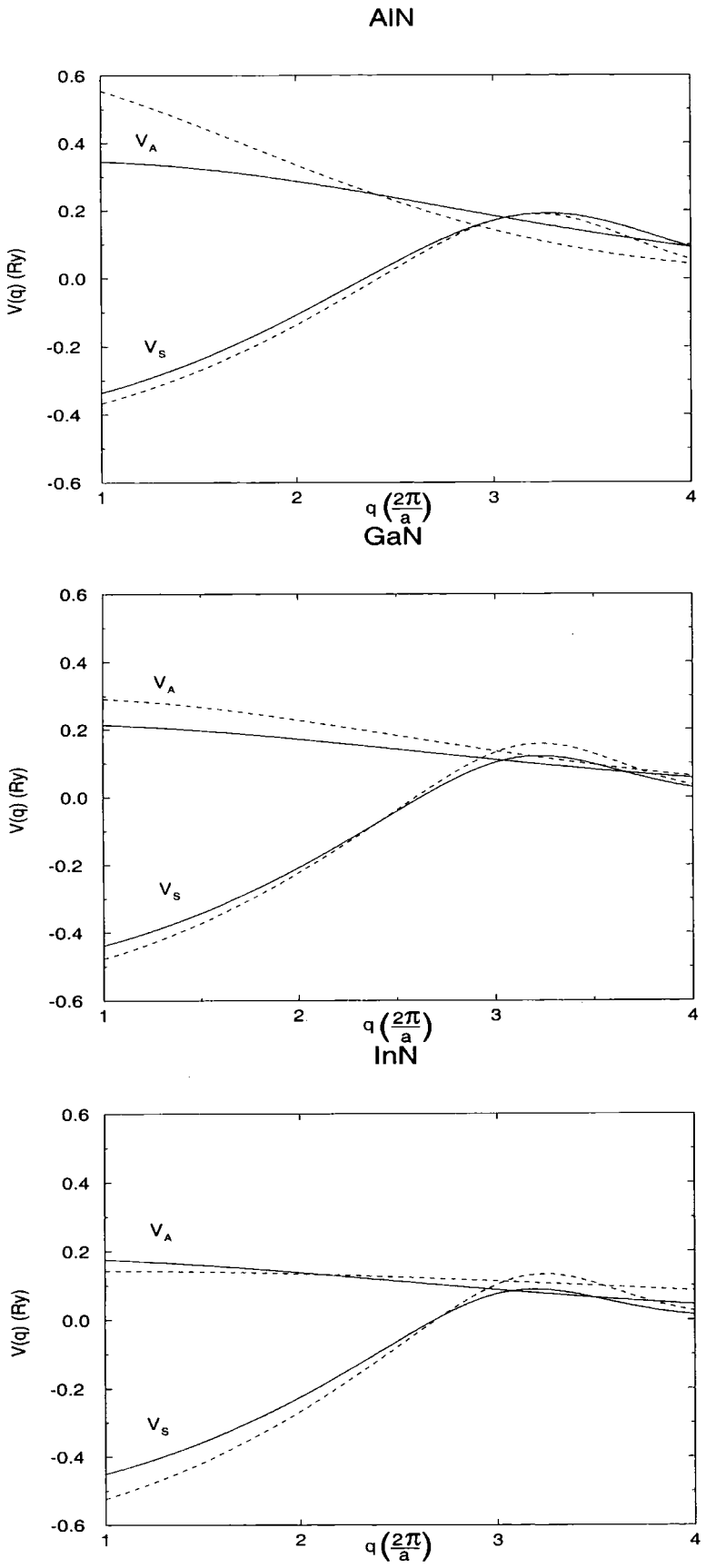

Figure 3.5: The symmetric (V_S) and antisymmetric (V_A) $V(q)$ curves for AlN, GaN and InN in the zincblende (solid line) and wurtzite (dashed line) structures.

3.5 Band Structure

The principal directions in the Brillouin zones of the zincblende and wurtzite structures are shown in Figure 3.6. Figures 3.7, 3.8 and 3.9 show the calculated empirical band structure compared with that calculated using the first principles code VASP (computed by Steve Pugh). The diamonds in the graphs are the targets used in the empirical pseudopotential fitting procedure. For GaN and InN there is good overall agreement in the band structures throughout the Brillouin zone. For AlN, particularly in the zincblende structure, the agreement is less good. The empirical band structures in Figures 3.7, 3.8 and 3.9 are consistent with the results of other groups reported in the literature [44], [55], [78].

Note that the valence band maxima all occur at the Γ-point. GaN and InN are both direct gap materials, in contrast to AlN in the zincblende structure which is indirect. Also note that for all the first principles calculations the band gap is underestimated. In particular, for the band gap of both zincblende and wurtzite InN a band gap of zero was obtained. As discussed in chapter 2, an underestimation of the band gap is typical for such calculations. Thus these band structures serve to highlight one of the advantages of the empirical pseudopotential method, in that it can reproduce accepted excited state energies from experiment.

3.6 The Dielectric Function

The optical properties of the nitrides are of obvious importance, and in particular a detailed knowledge of the dielectric function $\epsilon(\omega)$ is necessary for the development of

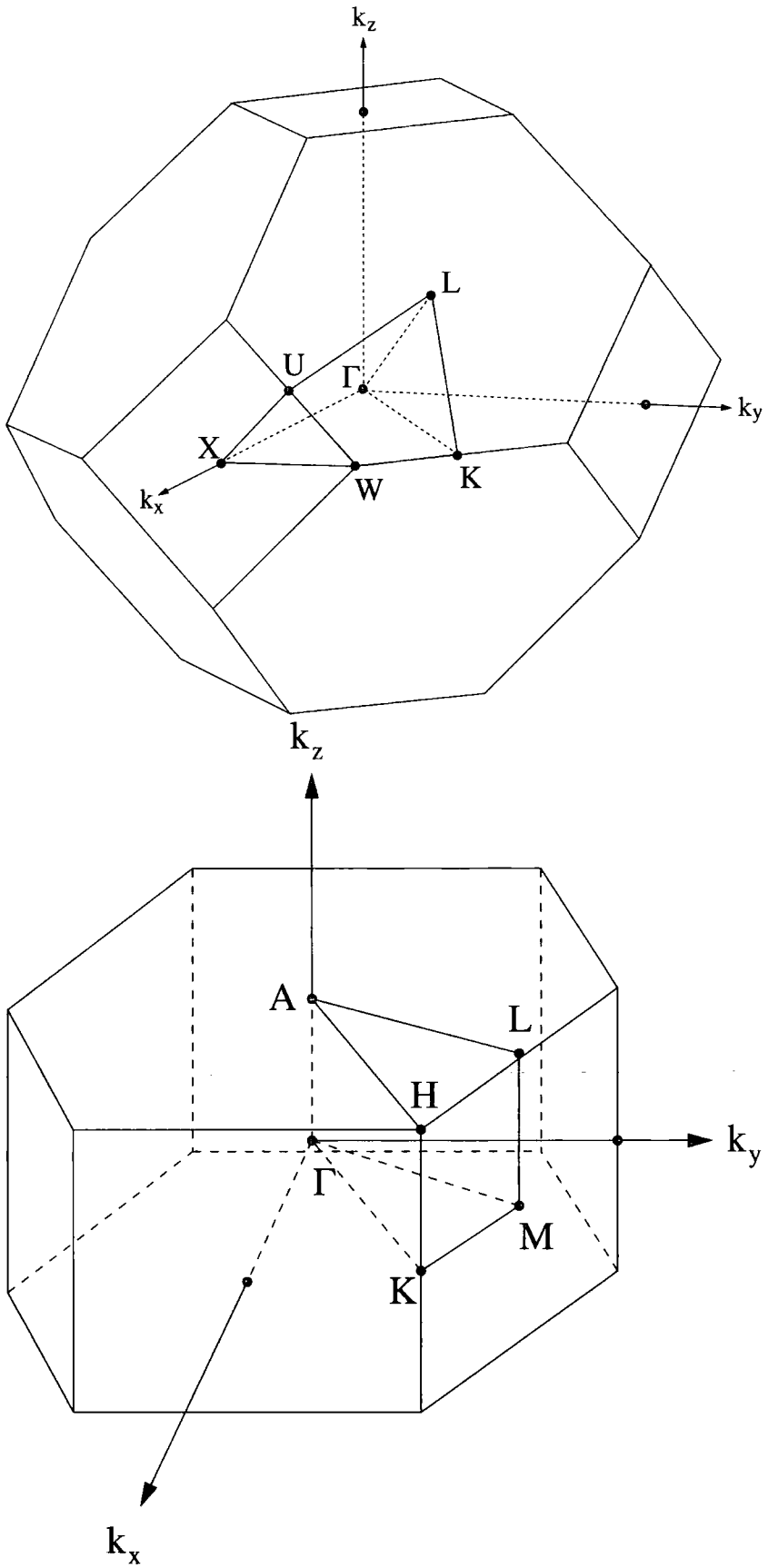

Figure 3.6: High symmetry points in the Brillouin zone for the zincblende and wurtzite structures.

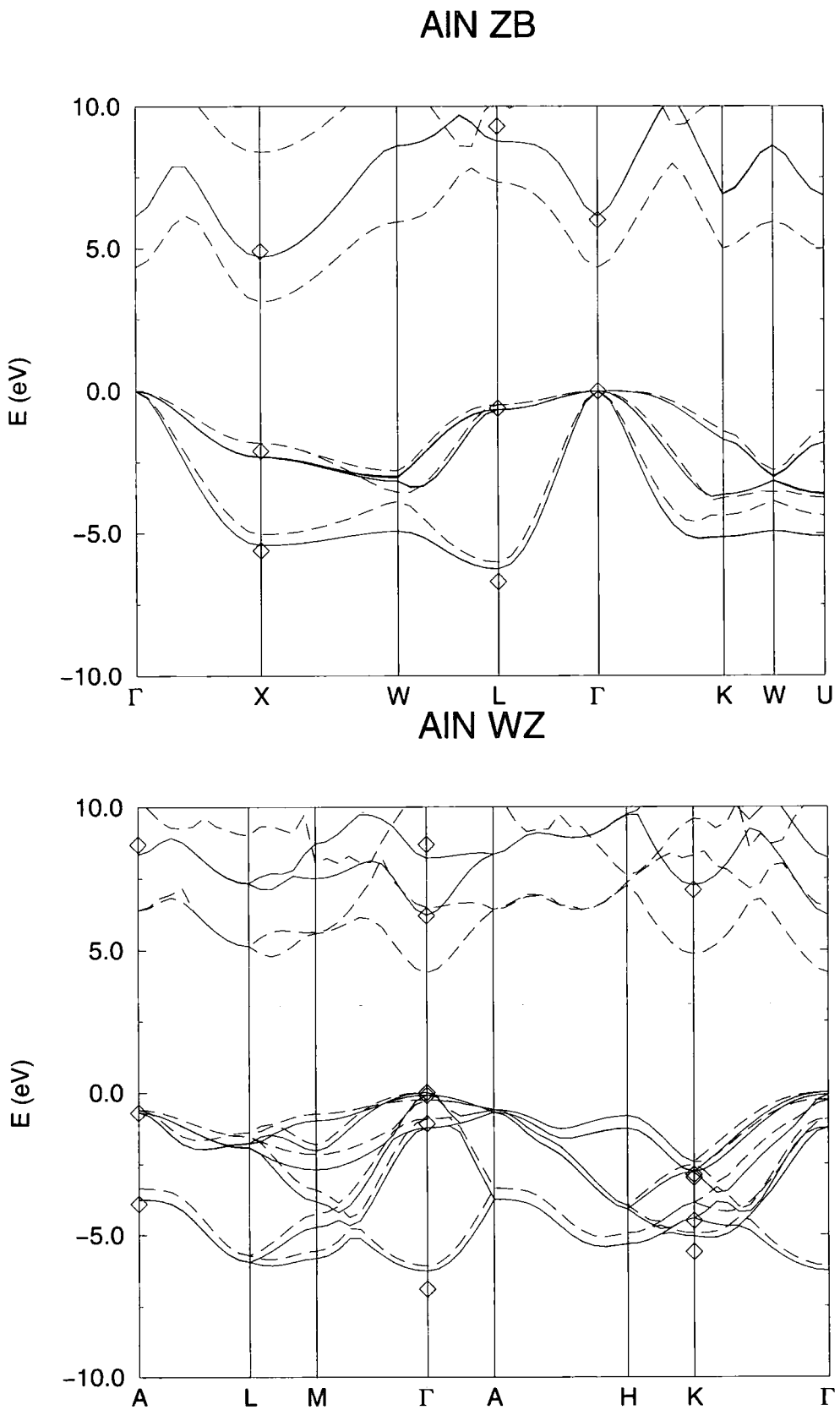

Figure 3.7: The band structures for zincblende and wurtzite structure AlN. The dashed lines are calculated using the first principles code VASP and the solid lines are calculated using the empirical pseudopotential method. The diamonds represent the targets used in the fitting procedure.

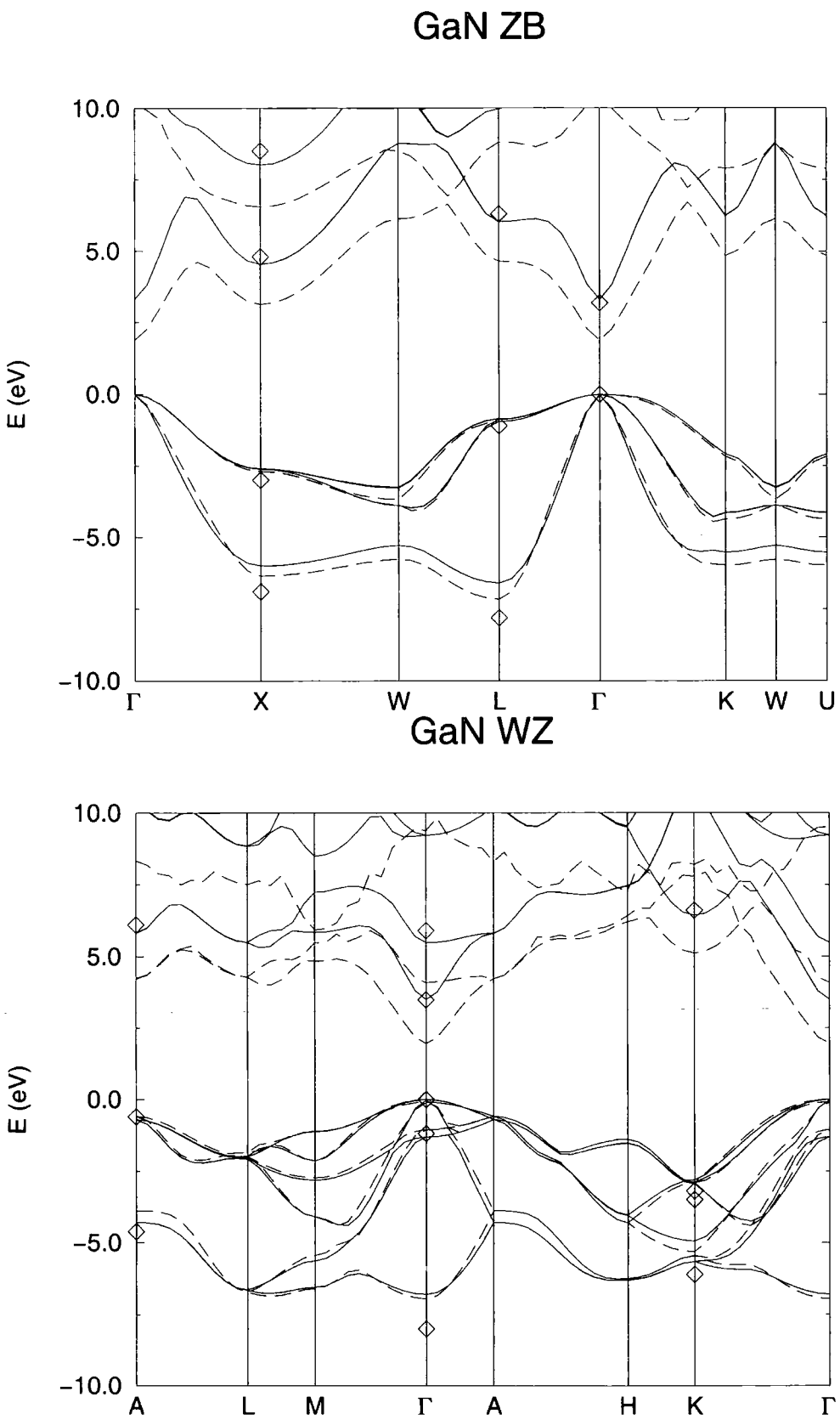

Figure 3.8: The band structures for zincblende and wurtzite structure GaN. The dashed lines are calculated using the first principles code VASP and the solid lines are calculated using the empirical pseudopotential method. The diamonds represent the targets used in the fitting procedure.

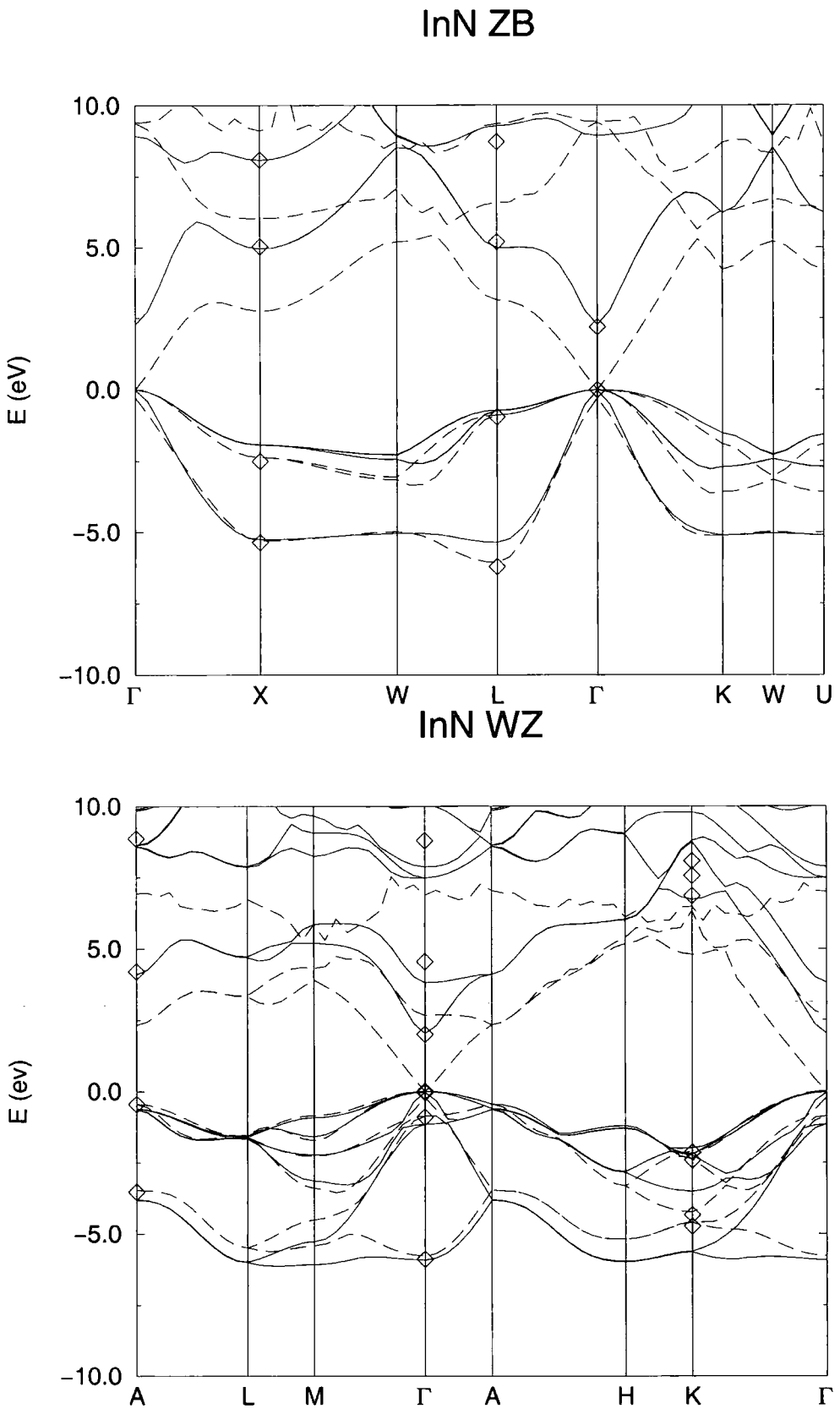

Figure 3.9: The band structures for zincblende and wurtzite structure InN. The dashed lines are calculated using the first principles code VASP and the solid lines are calculated using the empirical pseudopotential method. The diamonds represent the targets used in the fitting procedure.

optoelectronic devices. For example, Bragg reflectors can be employed as the highly reflective mirrors which are needed for laser operation, and for these it is important to know the optical response of the materials involved. Beyond the actual requirements of potential devices, calculations of the dielectric function have traditionally provided a useful contact between theory and experiment, and indeed reflectivity measurements have often been used as the data source from which the parameters for the empirical pseudopotential method have been fitted.

The optical properties can be described in terms of a complex dielectric function:

$$\epsilon(\omega) = \epsilon_1(\omega) + i\epsilon_2(\omega) \tag{3.33}$$

In general, ϵ is a function of both wavevector and frequency. The real and imaginary parts can be written as [65]

$$\begin{aligned} \epsilon_1(\mathbf{q},\omega) &= 1 + \frac{e^2}{\Omega\epsilon_0 q^2} \sum_{\mathbf{k},c,v} |\langle \mathbf{k}, c|\mathbf{k}+\mathbf{q}, v\rangle|^2 \\ &\times \left\{ [E_c(\mathbf{k}) - E_v(\mathbf{k}+\mathbf{q}) - \hbar\omega]^{-1} + [E_c(\mathbf{k}) - E_v(\mathbf{k}+\mathbf{q}) + \hbar\omega]^{-1} \right\} \end{aligned} \tag{3.34}$$

$$\epsilon_2(\mathbf{q},\omega) = \frac{\pi e^2}{\Omega\epsilon_0 q^2} \sum_{\mathbf{k},c,v} |\langle \mathbf{k}, c|\mathbf{k}+\mathbf{q}, v\rangle|^2 \delta\left(E_c(\mathbf{k}) - E_v(\mathbf{k}+\mathbf{q}) - \hbar\omega\right) \tag{3.35}$$

where c and v refer to the conduction and valence bands respectively. These equations involve computing overlaps and energies throughout the Brillouin zone, and in this work this sampling is done numerically using a Monte Carlo method. For fixed $\mathbf{q}$, a $\mathbf{k}$-point within the Brillouin zone is randomly picked and the energies and wavefunctions

in each band at $\mathbf{k}$ and $\mathbf{k}+\mathbf{q}$ are calculated. The real and imaginary parts of $\epsilon(\mathbf{q},\omega)$ corresponding to those $\mathbf{k}$ and $\mathbf{q}$ values are then evaluated at $\hbar\omega$ values ranging in small steps from 0-20 eV, and stored in histograms according to energy. This procedure is then repeated at fixed $\mathbf{q}$ for a large number of randomly chosen $\mathbf{k}$ vectors, with the real and imaginary values being added to the existing histogram values. After a certain number of such evaluations, the values in the histograms are divided by the number of $\mathbf{k}$ vectors used in the sampling. In practice, this number is chosen to ensure that the real and imaginary parts are sufficiently converged.

In evaluating the real and imaginary parts of $\epsilon(\mathbf{q},\omega)$ it is necessary to consider the the situation as $\hbar\omega \rightarrow E_c(\mathbf{k}) - E_v(\mathbf{k}+\mathbf{q})$. For the real part of the dielectric function this would lead to spikes, and thus in equation 3.34 the following approximation is made:

$$[E_c(\mathbf{k}) - E_v(\mathbf{k}+\mathbf{q}) - \hbar\omega]^{-1} + [E_c(\mathbf{k}) - E_v(\mathbf{k}+\mathbf{q}) + \hbar\omega]^{-1} \simeq$$
$$\mathcal{R}e\left\{[E_c(\mathbf{k}) - E_v(\mathbf{k}+\mathbf{q}) - \hbar\omega - i\eta]^{-1} + [E_c(\mathbf{k}) - E_v(\mathbf{k}+\mathbf{q}) + \hbar\omega + i\eta]^{-1}\right\} \quad (3.36)$$

where η is a small positive value. For $\hbar\omega \not\simeq E_c(\mathbf{k}) - E_v(\mathbf{k}+\mathbf{q})$ the right hand side of equation 3.36 is a good approximation to the left hand side. As $\hbar\omega \rightarrow E_c(\mathbf{k}) - E_v(\mathbf{k}+\mathbf{q})$ the right hand side tends to zero and thus spikes are avoided.

Similarly for $\epsilon_2(\mathbf{q},\omega)$, the Dirac delta function $\delta\left(E_c(\mathbf{k}) - E_v(\mathbf{k}+\mathbf{q}) - \hbar\omega\right)$ in equation 3.35 must also be approximated, in this case by a top hat function of finite width and unit area:

$$\delta\left(E_c(\mathbf{k}) - E_v(\mathbf{k}+\mathbf{q}) - \hbar\omega\right) \simeq \begin{cases} \frac{1}{2\delta e} & \text{for } |E_{cv} - \hbar\omega| \leq \delta e, \\ 0 & \text{otherwise.} \end{cases} \quad (3.37)$$

where δe is a small energy value.

The values of η and δe must be chosen to ensure that $\epsilon_1(\mathbf{q},\omega)$ and $\epsilon_2(\mathbf{q},\omega)$ are converged with respect to them. In practice, either the real or imaginary part of the dielectric function can be obtained from the other using the Kramers-Kronig relations shown in equation 3.38 below [79]. The numerical accuracy involved in the above approximations can be checked by comparing the directly calculated real and imaginary parts with those obtained using the Kramers-Kronig relations:

$$\begin{aligned} \epsilon_1(\omega) &= 1 + \frac{2}{\pi} P \int_0^\infty \frac{\omega' \epsilon_2(\mathbf{q},\omega')}{\omega'^2 - \omega^2} d\omega' \\ \epsilon_2(\omega) &= -\frac{2\omega}{\pi} \int_0^\infty \frac{\epsilon_1(\mathbf{q},\omega') - 1}{\omega'^2 - \omega^2} d\omega' \end{aligned} \tag{3.38}$$

The dielectric function $\epsilon(\mathbf{q},\omega)$ is obviously $\mathbf{q}$-dependent, and thus in a rigorous calculation all directions in the Brillouin zone would be sampled. The computation requirements for this, however, would be too large to make such a calculation feasible. Therefore, for the calculations in this work several appropriate directions for $\mathbf{q}$ within the Brillouin zone are chosen.

3.6.1 Zincblende Structure Materials

Equations 3.34 and 3.35 were employed to calculate the real and imaginary parts of the dielectric function for zincblende nitrides. These calculation were done for $\mathbf{q}$ in the (100), (110) and (111) directions in the Brillouin zone. The calculated dielectric responses in these three directions can be seen in Figure 3.10. From this, it can be seen that the dielectric function is approximately isotropic.

The empirical pseudopotential method was used for the calculations with an expansion involving two different numbers of plane waves. Figure 3.11 shows the real and imaginary parts of the dielectric function for InN calculated using 65 and 137 plane

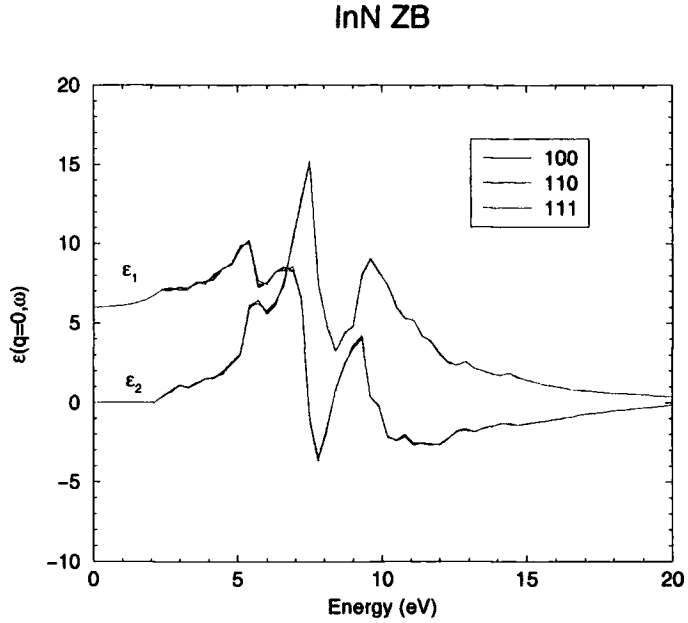

Figure 3.10: The calculated real and imaginary parts of the dielectric function for zincblende structure InN as a function of frequency at fixed wavevector (**k**=0). Calculations are performed in the 100 direction (black lines), the 110 direction (red lines) and the 111 direction (green lines) using 137 plane waves.

waves. From this, it is clear that a reasonably converged representation of the dielectric function in the zincblende structure can be obtained by using just 65 plane waves. However, 137 plane waves were used in this work because the empirical pseudopotential calculations are acceptably fast using this number.

The dielectric functions for zincblende structure AlN, GaN and InN are shown alongside those for the wurtzite structure in Figures 3.13, 3.14 and 3.15 respectively. All three graphs have the typical shape for both the real and imaginary parts of the dielectric function. There is little experimental information on the dielectric reponse for zincblende AlN and InN, but the dielectric response of zincblende GaN has been investigated experimentally [92–94]. Logothetidis *et al.* [92] observe two adjacent peaks in ϵ_2 at 7.0 and 7.6 eV. In this work, only one peak was found at 8.4 eV, though there is a small shoulder at approximately 6.75 eV. A feature of less sophisticated calculations such as the one in this work is that spectral features are shifted to higher energy, and this is discussed below in the results for wurtzite. First principles calculations of the

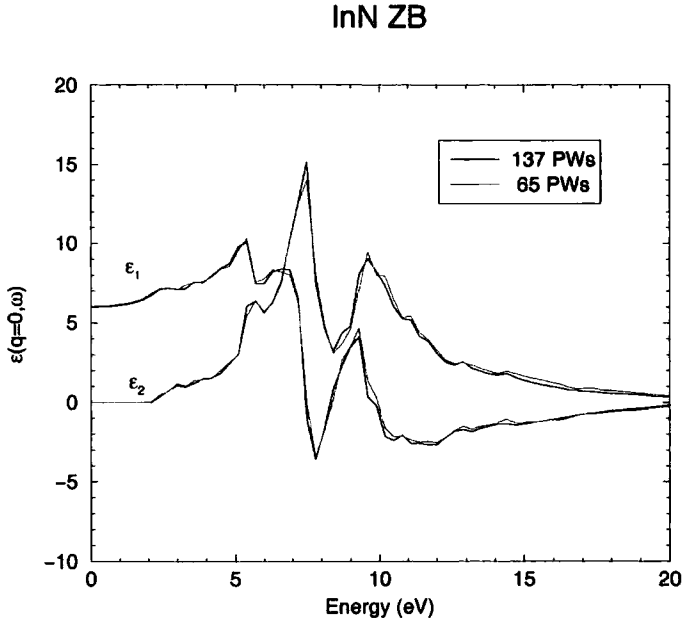

Figure 3.11: The calculated real and imaginary parts of the dielectric function for zincblende structure InN as a function of frequency at fixed wavevector (**k**=0). The calculations are performed using 137 plane waves (black lines) and 65 plane waves (red lines).

dielectric function exist for all three materials in the zincblende structure [12]. These authors find three distinct groups of peaks in ϵ_2 for both AlN and GaN instead of the two that are found in this work. More relevant to this work, Wang *et al.* [95] have also calculated the dielectric response of GaN in both zincblende and wurtzite structure using the empirical pseudopotential method. As in this work, they find two distinct peaks in ϵ_2. More generally, there is excellent agreement in the features of both the real and imaginary parts of the dielectric function. For GaN, the value for ϵ_∞ obtained in this work of 6.1 is larger than the theoretical value of 5.41 of Karch *et al.* [96], and is also larger than the experimental values for wurtzite of 5.29 [93] and 5.35 [94]. A reason for this could be due to the empirical model for GaN having a slightly smaller band gap (3.16 eV) than that found in other calculations or experiment.

3.6.2 Wurtzite Structure Materials

As for the zincblende structure , the real and imaginary parts of the dielectric function for AlN, GaN and InN in the wurtzite structure were calculated using equations 3.34 and 3.35. For this, 263 plane waves were used in the empirical pseudopotential method expansion. Unlike the zincblende structure, the dielectric function is not isotropic in the wurtzite structure because of its reduced symmetry. Instead, calculations of the dielectric response must be performed in two different directions, one perpendicular to the c-axis of the wurtzite crystal and one parallel to it. However, in the perpendicular (in-plane) direction the dielectric function is isotropic. This is shown in Figure 3.12, which plots the dielectric function in both the (100) and (210) in-plane directions. From this, it can be seen that there is essentially no difference in either ϵ_1 or ϵ_2 in the two directions. Therefore, the calculations in this work are done in the (001) c-axis direction and the (100) in-plane direction. Note that in the case of experiment, most measurements of the dielectric function are restricted to the in-plane direction.

The dielectric functions for wurtzite AlN, GaN and InN are shown alongside those for the zincblende structure in Figures 3.13, 3.14 and 3.15 respectively. Experimental and theoretical studies of the dielectric response exist for the wurtzite form of all three materials. The experimental studies have largely been based on reflectivity and ellipsometry measurements [80,85–90]. Unfortunately, these studies have been compromised by poor material quality, and in particular surface roughness and oxide layers are thought to account for some of the deficiencies in the measured values.

Theoretical calculations [46,83–85,91,92] have successfully reproduced some gross experimental features in the dielectric function such as positions of absorption peaks, but have otherwise failed to obtain the general shape of the curve. More recently, Benedict *et al.* [93,94] have included the electron-hole interaction within a first-principles scheme to obtain results in improved agreement with experiment. Djurisic *et al.* [95,96] have also incorporated excitonic effects to successfully model the dielectric response

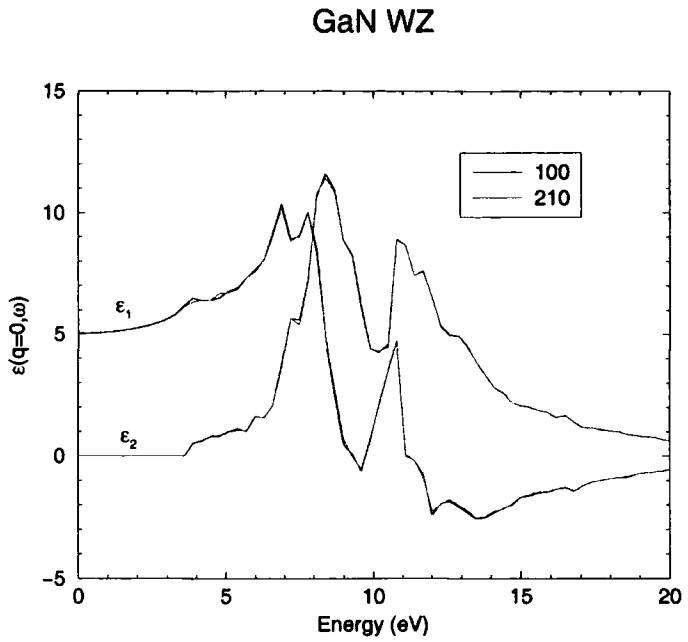

Figure 3.12: The real and imaginary parts of the dielectric function for wurtzite structure GaN as a function of frequency at fixed wavevector ($\mathbf{k}$=0) in two in-plane directions. The black lines show the calculation performed in the 100 direction and the red lines correspond to a calculation done in the 210 direction.

response for the wurtzite form of all three materials. As for the zincblende structure, Wang *et al.* [95] have also used the empirical pseudopotential method to calculate the dielectric response for GaN in the wurtzite structure. The spectral features these authors find for both ϵ_1 and ϵ_2 are in excellent agreement with those found in this work.

In comparing the wurtzite dielectric function in the in-plane and c-axis directions several features deserve comment. First, while there are differences between the two directions, the general shape of the graph is the same. For InN, the dielectric function is quite similar in the two directions, but for AlN and GaN there are more obvious differences. For these two materials, the main difference is in the positions of the peaks in the spectrum. In particular, the initial peak in ϵ_1 in the *c*-axis direction coincides with a smaller peak in the in-plane direction. This difference is more pronounced in ϵ_2, where the initial peak in the *c*-axis direction appears only as a shoulder in the in-plane direction. This kind of behaviour, in which small in-plane peaks correspond to

spectra [97,98]. More recently, Benedict *et al.* [94] have calculated the in-plane and c-axis dielectric response for AlN and GaN. They too report a significantly enhanced initial peak for ϵ_2 in the c-axis direction compared to the in-plane direction.

So far, no measurements of the dielectric response for wurtzite AlN, GaN or InN in the c-axis direction are available. However, experimental studies of the in-plane dielectric function do exist for AlN [88,90], GaN [80,85,94] and InN [87], and can be contrasted with those obtained in this work. In comparison to experiment, the principal difference is that the calculations presented here produce sharper and stronger peaks than those observed experimentally. The experimental explanation for this is thought to be the sample quality, with surface roughness and oxide layers in particular held to be responsible for decreased magnitudes in ϵ_2. Note however that the theoretical results in this work can always be artificially broadened by choosing a less dense set of energy histograms as described earlier in the chapter. The other main difference is that the spectral features calculated in this work occur at higher energies than those measured experimentally. This is consistent with the work of Benedict *et al.* [93] and Albrecht *et al.* [99], in which it was shown that inclusion of the electron-hole interaction shifts the spectral features to lower energy in better agreement with experiment. Thus, since the calculations in this work do not include this interaction, it is to be expected that the peaks are at higher energies than suggested by experiment.

More generally, InN has a larger value for ϵ_∞ than GaN, which in turn has a large value that AlN. For GaN, the value for ϵ_∞ of 5.03 eV slightly understimates the measured values of 5.29 [81] and 5.35 [82].

3.6.3 Discussion

It is also instructive to compare the dielectric functions of the zincblende and wurtzite structures shown in Figures 3.13, 3.14 and 3.15. Noticeably, the shapes of the curves

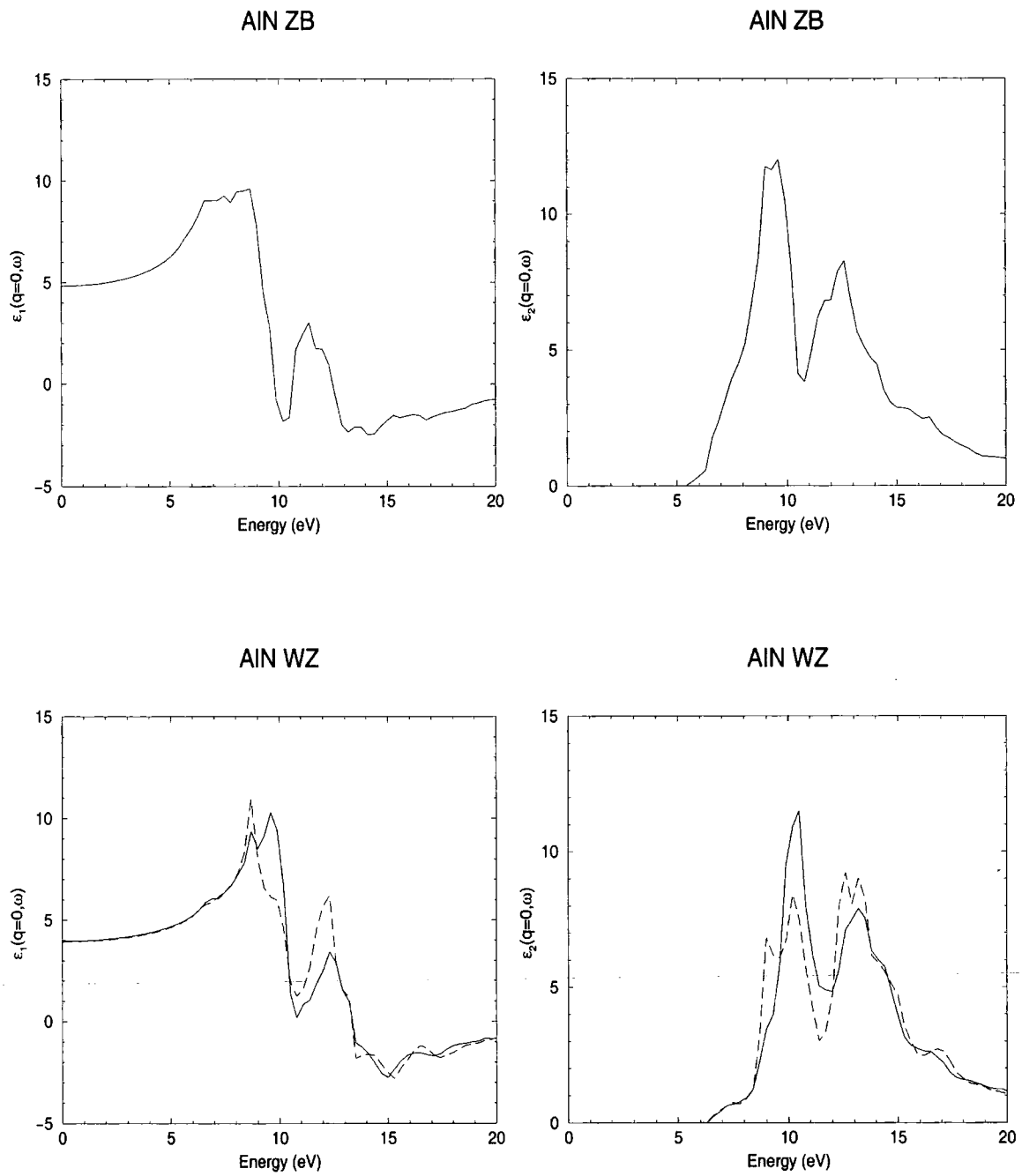

Figure 3.13: The real and imaginary parts of the dielectric function for zincblende and wurtzite structure AlN as a function of frequency at fixed wavevector ($\mathbf{k}$=0). For the wurtzite structure, calculations are done in the in-plane (solid line) and c-axis (dashed line) directions.

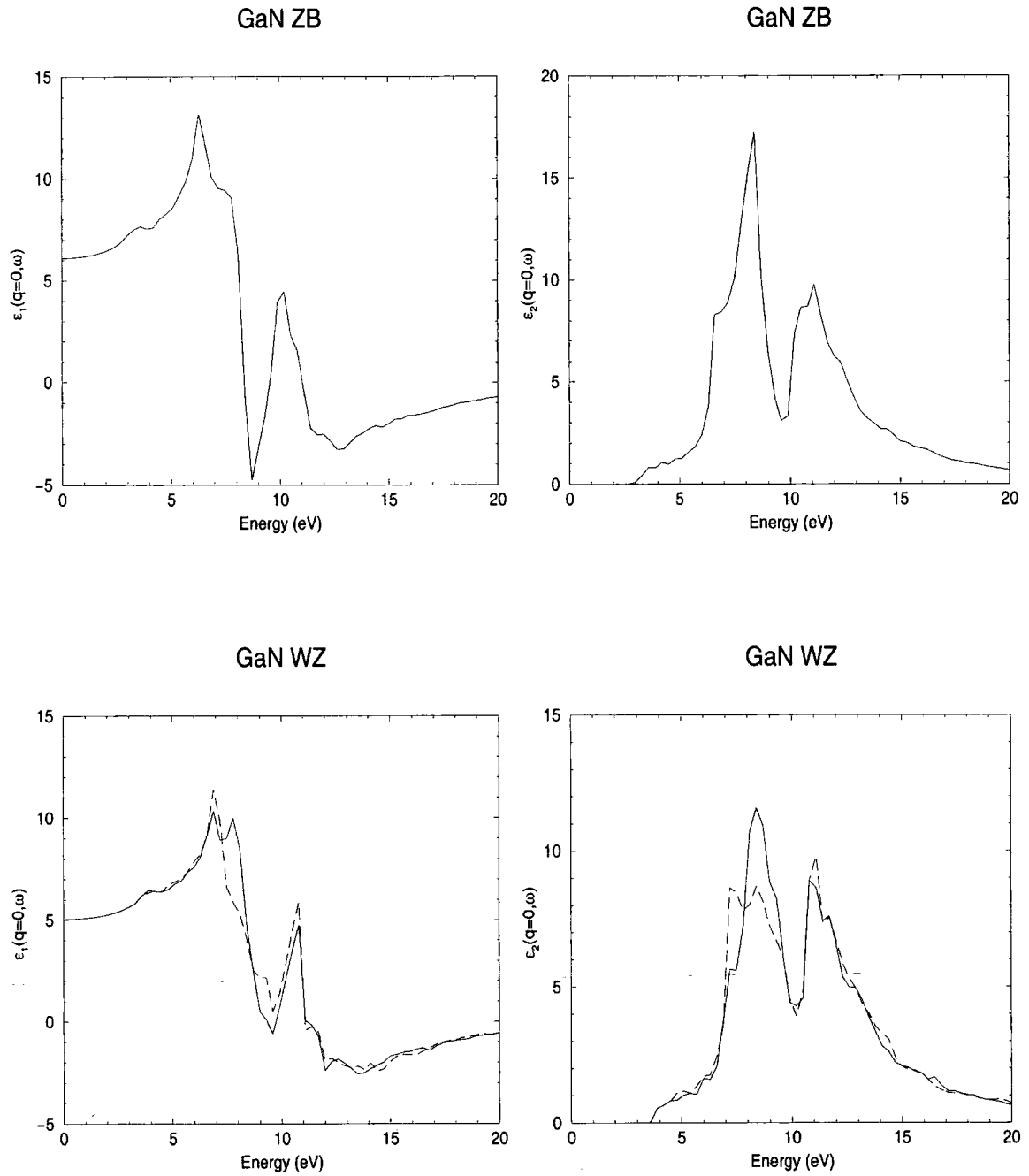

Figure 3.14: The real and imaginary parts of the dielectric function for zincblende and wurtzite structure GaN as a function of frequency at fixed wavevector ($\mathbf{k}$=0). For the wurtzite structure, calculations are done in the in-plane (solid line) and c-axis (dashed line) directions.

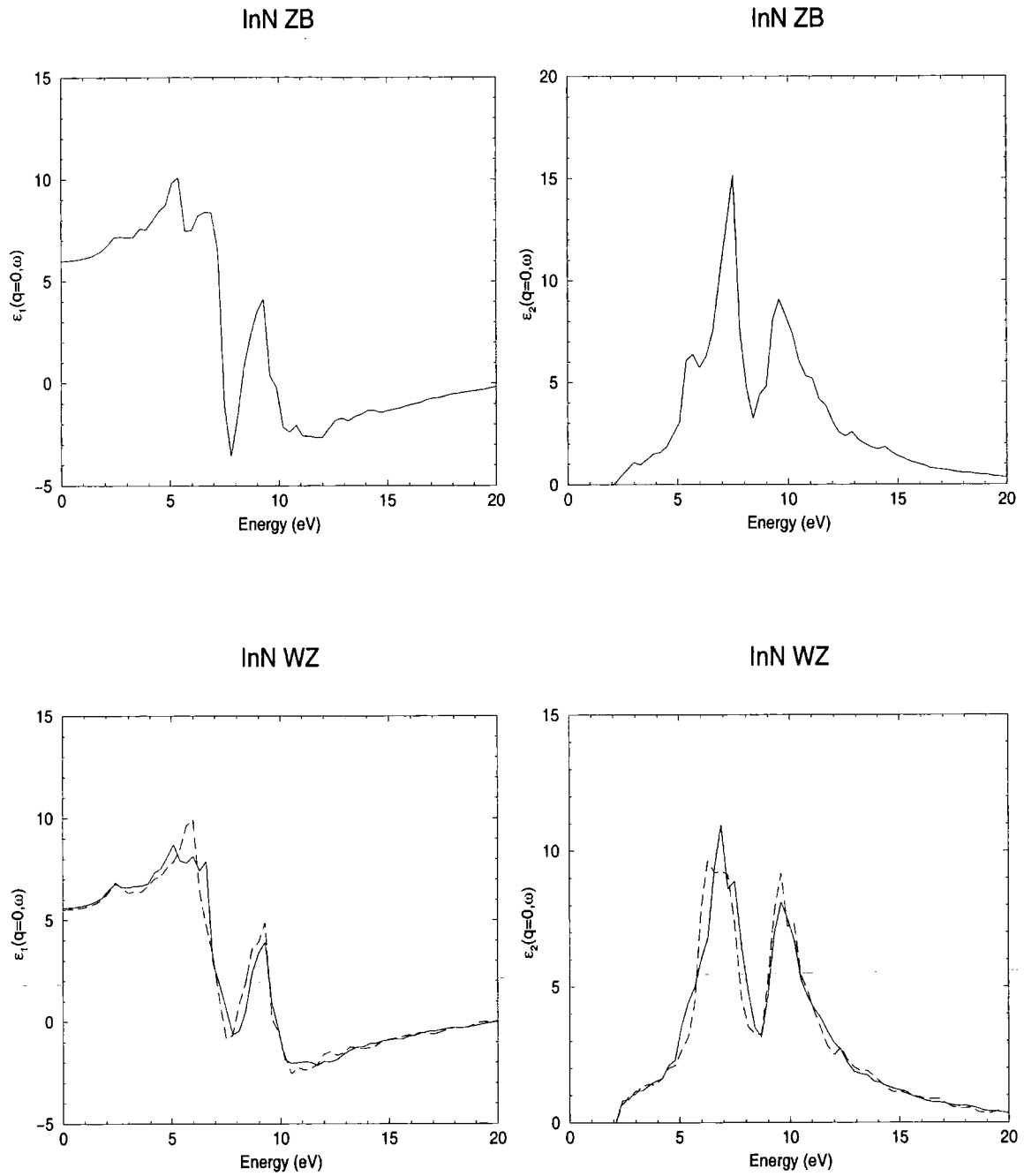

Figure 3.15: The real and imaginary parts of the dielectric function for zincblende and wurtzite structure InN as a function of frequency at fixed wavevector ($\mathbf{k}$=0). For the wurtzite structure, calculations are done in the in-plane (solid line) and c-axis (dashed line) directions.

are very similar for the different structures of each material. In particular there is marked similarity for InN, where the positions of the peaks agree very well. However, for AlN and GaN the spectral features in the wurtzite structure are at slightly higher energy. It is interesting to note that for each of the materials the zincblende dielectric function is in the best agreement with the wurtzite calculations done in-plane. This is to be expected given that the zincblende and wurtzite structures differ in their stacking arrangements in the c direction.

For each of the materials the peaks in the zincblende case are higher, particularly for the imaginary part ϵ_2. The value of ϵ_∞ also differs for the two structures. For all three materials the value of ϵ_∞ in the wurtzite structure is lower than it is in zincblende. It can also be seen that the values of ϵ_∞ increase going from AlN to InN, and both these results are consistent with the band gaps of the materials. AlN has the largest band gap and thus the energy terms in equations 3.34 and 3.34 will be large, lowering the value of the dielectric function. Similarly, zincblende AlN and GaN have a lower band gap than that they do in the wurtzite structure, which explains their larger dielectric function. However, for InN this argument does not work. Even though wurtzite InN has a smaller band gap than it does in the zincblende structure, ϵ_∞ is smaller, indicating that the wavefunction overlaps for wurtzite InN in equations 3.34 and 3.35 are perhaps smaller.

3.7 Summary

In this chapter, the main elements of the pseudopotential concept were introduced and pseudopotentials describing AlN, GaN and InN were obtained. These parameter sets were used for band structure calculations throughout the Brillouin zone and for investigating the dielectric response. Importantly, the pseudopotentials in this work were specifically constructed in such a way as to facilitate study of strained material and alloys. This involved adopting a functional form for the symmetric and antisymmetric

parts of the pseudopotential, from which the form factors for strained or unstrained material could be obtained. These empirical pseudopotential method descriptions form the basis for the further calculations on strained and unstrained material, as well as alloys, in the following chapters.

Chapter 4

The k.p Method and its Applications

4.1 Introduction

An alternative to the empirical pseudopotential method for the calculation of band structure is the **k.p** method. Primarily, it is used to describe the band structure close to a particular **k** point in great detail. It is because of this that it lends itself particularly well to studies of semiconductors, since in these materials the free carriers are normally confined to a small region of **k** space. For the nitrides, these studies have concentrated on the Γ point as the main region of interest.

The method again requires some input, but unlike the pseudopotential method it does not rely on knowledge of the crystal potential. Rather, it requires certain matrix elements which can be derived from experiment or more fundamental calculations, and consideration of the symmetry properties of the material in question. Computationally, the 6 or 8 band **k.p** method is much faster than the pseudopotential method since it involves matrices of much smaller size. Thus in cases where energies or matrix elements used in the calculation of physical properties must be evaluated many times, the **k.p** approach can be considered preferable to the empirical pseudopotential method.

4.2 Theory

The one-electron Schrödinger equation is written:

$$H\psi(\mathbf{r}) = \left[\frac{-\hbar^2}{2m_0}\nabla^2 + V(\mathbf{r})\right]\psi(\mathbf{r}) = E\psi(\mathbf{r}) \tag{4.1}$$

where m_0 is the free electron mass and $V(\mathbf{r})$ is the preiodic potential. The wavefunction of the state with wavevector $\mathbf{k}$ in band n is expressed in its Bloch form:

$$\psi_{nk}(\mathbf{r}) = u_{nk}(\mathbf{r})e^{i\mathbf{k}.\mathbf{r}} \tag{4.2}$$

Substitution of this into the Schrödinger equation leads to

$$\left[-\frac{\hbar^2}{2m_0}\nabla^2 - \frac{i\hbar^2}{m_0}\mathbf{k}.\nabla + \frac{\hbar^2k^2}{2m_0} + V(\mathbf{r})\right]u_{nk}(\mathbf{r}) = E_n(\mathbf{k})u_{nk}(\mathbf{r}) \tag{4.3}$$

If we consider a point $\mathbf{k} = \mathbf{k}_0$, then equation 4.3 has a complete set of eigenfunctions $u_{nk_0}(r)$ which completely span the space of functions periodic in the real space Bravais lattice, and can be rewritten as:

$$\left[H_{k_0} + \frac{\hbar}{m_0}(\mathbf{k} - \mathbf{k}_0).\nabla + \frac{\hbar^2}{2m_0}(k^2 - k_0^2)\right]u_{nk}(\mathbf{r}) = E_n(\mathbf{k})u_{nk}(\mathbf{r}) \tag{4.4}$$

where

$$H_{k_0} = \frac{p^2}{2m_0} + \frac{\hbar}{m_0}\mathbf{k}_0.\mathbf{p} + \frac{\hbar^2k_0^2}{2m_0} + V(\mathbf{r}) \tag{4.5}$$

It is then possible to express $u_{nk}(\mathbf{r})$ as a linear combination of the Bloch periodic parts of the infinite set of bands at this wavevector $\mathbf{k}_0$:

$$u_{nk}(\mathbf{r}) = \sum_m c_{mn}(\mathbf{k} - \mathbf{k}_0) u_{m\mathbf{k}_0}(\mathbf{r}) \tag{4.6}$$

This expansion is such that only a few terms are necessary to obtain a satisfactory representation of $u_{nk}(\mathbf{r})$ when $\mathbf{k}$ is sufficiently close to $\mathbf{k}_0$. For the work presented here, attention is focused on the conduction and valence bands at the zone centre ($\mathbf{k}_0 = 0$). Substituting this expansion into equation 4.4 and multiplying on the left by $u^*_{n0}(\mathbf{r})$ we obtain:

$$\sum_m \left(\left[E_n(\mathbf{k}) - \left(E_n(0) + \frac{\hbar^2 k^2}{2m_0} \right) \right] \delta_{nm} - \frac{\hbar}{m_0} \mathbf{k}.\mathbf{p}_{nm} \right) c_{mn} = 0 \tag{4.7}$$

where

$$\mathbf{p}_{nm} = \int u^*_{n0}(\mathbf{r}) \mathbf{p} u_{m0}(\mathbf{r}) d\mathbf{r} \tag{4.8}$$

and the integration is performed over the unit cell. Equations 4.6 to 4.8 define the **k.p** representation, and the problem is reduced to diagonalisation of the Hamiltonian matrix contained in the left hand side of equation 4.7. The terms $\mathbf{k}.\mathbf{p}_{nm}$ are the off-diagonal elements of the Hamiltonian. Note that equation 4.6 is valid for all $\mathbf{k}$ when m runs over the infinite basis, but in practice it is necessary to use a finte basis set. Use of $\simeq 30$ basis states from the bands around the fundamental gap can give a reasonable description of the band structure throughout the Brillouin zone. However, the method is commonly used with a small number (typically 8 or less) of basis states, and the results are correspondingly accurate for a small range of $\mathbf{k}$.

All the states are divided into two classes A and B, where those in A are the basis states from the bands of interest. Those in B are from the bands which are energetically remote from the bands of interest, but their interactions with the states in A are such that they have a significant effect on the resulting band structure. These interactions can be treated perturbatively following a method introduced by Löwdin [100]. In this, the **k.p** coupling between states in A are treated exactly, while the interactions of these states with those in B are treated as perturbations by defining an effective matrix element between any two valence band wavefunctions as

$$H'_{ij} = H_{ij} + \sum_k \frac{H_{ik}H_{kj}}{E_i - E_k} \tag{4.9}$$

where i and j are are in A and k is in B.

In this work, the states of interest (those in A) are the six uppermost valence bands and the first two conduction bands at the zone centre. At the zone centre the spin-orbit splitting is such that the valence band consists of a fourfold multiplet ($|3/2, \pm 3/2\rangle$, $|3/2, \pm 1/2\rangle$) and a doublet ($|1/2, \pm 1/2\rangle$). We use these to form the basis set of the Hamiltonian, and their form is addressed in the next section. Thus with this approach, the problem is reduced to one of diagonalising an 8×8 matrix, with the effects of the remote bands included perturbatively in the Hamiltonian.

Additionally, the spin-orbit interaction is considered for the materials studied in this work. This interaction increases with atomic number (or mass), leading to a lowering of the symmetry and consequent splitting of the valence bands. The spin-orbit interaction contribution to the Hamiltonian in equation 4.1 can be written as [27]:

$$H_{SO} = \frac{\hbar}{4m^2c^2} \left[\nabla V \times \mathbf{p}\right] . \sigma \tag{4.10}$$

where σ are the usual Pauli operators on the electron spin variable. This results in two

additional terms in equation 4.5:

$$H_{kSO} = \frac{\hbar}{4m^2c^2}\left[\nabla V \times \mathbf{p}\right].\sigma + \frac{\hbar^2}{4m_0^2c^2}\left[\nabla V \times \mathbf{k}\right].\sigma \tag{4.11}$$

Within the **k.p** method it is appropriate to treat this interaction by adding it to the Löwdin perturbation Hamiltonian. The justification for this is that the spin-orbit splitting is typically of the order of 0.1 eV compared to a band gap of a few eV (e.g. GaAs etc.). Moreover, H_{kSO} is usually very small compared to H_{SO}, and thus is commonly neglected.

4.3 Zincblende Structure Materials

For optical devices, the primary region of interest lies around the fundamental energy gap which separates the occupied states in the valence band from the unoccupied states in the conduction band at zero temperature. For all the materials considered in this work, the top of the valence band lies at the zone centre. If we consider GaN, then Ga has an outer shell configuration of $4s^24p^1$ while for nitrogen the outer shell configuration is $2s^22p^3$. These electronic configurations determine the conduction and valence band states around the band gap, and can be considered by reference to standard atomic physics. In this, the orbital electronic wavefunctions are classified as s, p, d, etc. according to their orbital angular momentum l. The p states corresponding to $l = 1$ are triply degenerate, and can be chosen to be eigenstates of L_z, the z-component of the orbital angular momentum operator, with eigenvalues $m_l = 1, 0, -1$. The spherical harmonics $|l, m_l\rangle$ corresponding to these states can thus be written as:

$$\begin{aligned} |1,\ 1\rangle &= -(x+iy)/\sqrt{2} \\ |1,\ 0\rangle &= z \\ |1,-1\rangle &= (x-iy)/\sqrt{2} \end{aligned} \tag{4.12}$$

However, the spin-orbit interaction must also be considered. This is usually expressed as :

$$H_{SO} = \lambda \mathbf{L}.\mathbf{s} \tag{4.13}$$

where λ is the spin-orbit coupling. The eigenfunctions of this equation are eigenstates of the total angular momentum $\mathbf{J} = \mathbf{L} + \mathbf{S}$ and its z component J_z. For the p states, with $l = 1$ and $s = 1/2$ the possible values for the eigenvalues of j are $j = l + s = 3/2$ and $j = l - s = 1/2$. For each of these there are $2j + 1$ values of m_z, the eigenvalue of J_z.

Thus in zincblende structure materials the inclusion of the spin-orbit interaction splits the p-like valence bands into an upper fourfold degenerate Γ_8 state and a lower twofold degenerate Γ_7 state. At the bottom of the conduction band is a twofold degenerate Γ_6 state derived from s orbitals. Thus it is natural to choose these eight zone centre states to form the basis set of the Hamiltonian:

$$
\begin{aligned}
u_1 &= |S\uparrow\rangle \\
u_2 &= (1/\sqrt{2})|(X+iY)\uparrow\rangle \\
u_3 &= (1/\sqrt{6})|(X+iY)\downarrow\rangle - \sqrt{\frac{2}{3}}|Z\uparrow\rangle \\
u_4 &= (1/\sqrt{3})\left[|(X+iY)\downarrow\rangle + |Z\uparrow\rangle\right] \\
u_5 &= |S\downarrow\rangle \\
u_6 &= (1/\sqrt{2})|(X-iY)\downarrow\rangle \\
u_7 &= -(1/\sqrt{6})|(X-iY)\uparrow\rangle - \sqrt{\frac{2}{3}}|Z\downarrow\rangle \\
u_8 &= -(1/\sqrt{3})\left[|(X-iY)\uparrow\rangle + |Z\downarrow\rangle\right]
\end{aligned} \tag{4.14}
$$

where $|S\rangle$ denotes an s orbital and $|X\rangle$, $|Y\rangle$ and $|Z\rangle$ indicate p_x, p_y and p_z orbitals respectively. The arrows represent the spin orientation. In this basis, the Hamiltonian follows as given in Table 4.1 [27,101–103].

In Table 4.1

$$
\begin{aligned}
k_{||}^2 &= k_x^2 + k_y^2 \\
k_+ &= k_x + ik_y \\
k_- &= k_x + ik_y
\end{aligned} \tag{4.15}
$$

where all the wavevectors are expressed in units of $\sqrt{\hbar^2/2m_0}$. The other parameters in Table 4.1 are given by:

$$
H = \begin{pmatrix}
sk^2 + E_c & \sqrt{\frac{1}{2}}iPk_+ & -\sqrt{\frac{2}{3}}iPk_z & \sqrt{\frac{1}{3}}iPk_z & 0 & 0 & -\sqrt{\frac{1}{6}}iPk_- & -\sqrt{\frac{1}{3}}iPk_- \\
-\sqrt{\frac{1}{2}}iPk_- & \begin{matrix}(2\gamma_2-\gamma_1)k_z^2+E_v\\ -(\gamma_1+\gamma_2)k_{||}^2\end{matrix} & 2\sqrt{3}k_-k_z & -\sqrt{6}\gamma_3k_-k_z & 0 & 0 & \begin{matrix}\sqrt{3}\gamma_2(k_x^2-k_y^2)\\ -2\sqrt{3}i\gamma_3k_xk_y\end{matrix} & \begin{matrix}\sqrt{6}(k_x^2-k_y^2)\\ -2\sqrt{6}i\gamma_3k_xk_y\end{matrix} \\
\sqrt{\frac{2}{3}}iPk_z & 2\sqrt{3}\gamma_3k_+k_z & \begin{matrix}-(\gamma_1+2\gamma_2)k_z^2\\ +(\gamma_2-\gamma_1)k_{||}^2+E_v\end{matrix} & \begin{matrix}2\sqrt{2}\gamma_2k_z^2\\ -\sqrt{2}\gamma_2k_{||}^2\end{matrix} & -\sqrt{\frac{1}{6}}iPk_- & \begin{matrix}-\sqrt{3}\gamma_2(k_x^2-k_y^2)\\ +2\sqrt{3}i\gamma_3k_xk_y\end{matrix} & 0 & -3\sqrt{2}\gamma_3k_-k_z \\
-\sqrt{\frac{1}{3}}iPk_z & -\sqrt{6}\gamma_3k_+k_z & \begin{matrix}2\sqrt{2}\gamma_2k_z^2\\ -\sqrt{2}\gamma_2k_{||}^2\end{matrix} & \begin{matrix}-\gamma_1k_z^2+E_v\\ -\gamma_1k_{||}^2-\Delta\end{matrix} & -\sqrt{\frac{1}{3}}iPk_- & \begin{matrix}-\sqrt{6}\gamma_2(k_x^2-k_y^2)\\ +2\sqrt{6}i\gamma_3k_xk_y\end{matrix} & 3\sqrt{2}\gamma_3k_-k_z & 0 \\
0 & 0 & \sqrt{\frac{1}{6}}iPk_+ & \sqrt{\frac{1}{3}}iPk_+ & sk^2+E_c & \sqrt{\frac{1}{2}}iPk_- & -\sqrt{\frac{2}{3}}iPk_z & \sqrt{\frac{1}{3}}iPk_z \\
0 & 0 & \begin{matrix}-\sqrt{3}\gamma_2(k_x^2-k_y^2)\\ -2\sqrt{3}i\gamma_3k_xk_y\end{matrix} & \begin{matrix}-\sqrt{6}\gamma_2(k_x^2-k_y^2)\\ -2\sqrt{6}i\gamma_3k_xk_y\end{matrix} & -\sqrt{\frac{1}{2}}iPk_+ & \begin{matrix}(2\gamma_2-\gamma_1)k_z^2+E_v\\ -(\gamma_1+\gamma_2)k_{||}^2\end{matrix} & 2\sqrt{3}\gamma_3k_+k_z & -\sqrt{6}\gamma_3k_+k_z \\
\sqrt{\frac{1}{6}}iPk_+ & \begin{matrix}\sqrt{3}\gamma_2(k_x^2-k_y^2)\\ +2\sqrt{3}i\gamma_3k_xk_y\end{matrix} & 0 & 3\sqrt{2}\gamma_3k_+k_z & \sqrt{\frac{2}{3}}iPk_z & 2\sqrt{3}\gamma_3k_-k_z & \begin{matrix}-(\gamma_1+2\gamma_2)k_z^2\\ +(\gamma_2-\gamma_1)k_{||}^2+E_v\end{matrix} & \begin{matrix}2\sqrt{2}\gamma_2k_z^2\\ -\sqrt{2}\gamma_2k_{||}^2\end{matrix} \\
\sqrt{\frac{1}{3}}iPk_+ & \begin{matrix}\sqrt{6}\gamma_2(k_x^2-k_y^2)\\ 2\sqrt{6}i\gamma_3k_xk_y\end{matrix} & -3\sqrt{2}\gamma_3k_+k_z & 0 & -\sqrt{\frac{1}{3}}iPk_z & -\sqrt{6}\gamma_3k_-k_z & \begin{matrix}2\sqrt{2}\gamma_2k_z^2\\ -\sqrt{2}\gamma_2k_{||}^2\end{matrix} & \begin{matrix}-\gamma_1k_z^2+E_v\\ -\gamma_1k_{||}^2-\Delta\end{matrix}
\end{pmatrix}
$$

Table 4.1: The zincblende **k.p** Hamiltonian matrix

$$
\begin{aligned}
P &= \sqrt{\frac{2m}{\hbar^2}} P_{Kane} \\
P_{kane} &= -i\frac{\hbar}{m}\langle S|p_x|X\rangle \\
p_x &= \frac{\hbar}{i}\frac{\partial}{\partial x}
\end{aligned}
\tag{4.16}
$$

where E_c, E_v and Δ are the conduction band, valence band and spin-orbit splitting energies at the Γ point. γ_1, γ_2 and γ_3 are the *modified* Luttinger parameters, and are related to the *unmodified* Luttinger parameters γ_1^L, γ_2^L and γ_3^L by the following equations [103]:

$$
\begin{aligned}
\gamma_1 &= \gamma_1^L - \frac{E_p}{3E_g} \\
\gamma_2 &= \gamma_2^L - \frac{E_p}{6E_g} \\
\gamma_3 &= \gamma_3^L - \frac{E_p}{6E_g} \\
s &= \frac{1}{m_c} - \frac{2E_p}{3}\left(\frac{1}{E_g} + \frac{1}{2(E_g+\Delta)}\right)
\end{aligned}
\tag{4.17}
$$

where E_g is the band gap and

$$
E_p = \frac{2m_0}{\hbar^2}P^2 \tag{4.18}
$$

The unmodified Luttinger parameters are related to the effective masses and are discussed below.

4.3.1 Calculation of the k.p parameters

The empirical pseudopotential models for zincblende structure AlN, GaN and InN that were presented in chapter 3 all used the nonlocal form for the pseudopotential. However, for the **k.p** calculations presented here the pseudopotentials used are local. This was necesssary in order to to obtain **k.p** parameters using the direct method, and is explained below. Note that new, entirely local pseudopotentials were generated for this purpose.

Two different methods were employed to obtain **k.p** parameters from the empirical pseudopotential models. The first method exploits the fact that the valence bands are sufficiently parabolic near the zone centre to enable a simple calculation of the **k.p** parameters from the curvature of the bands. In the second method, the **k.p** parameters are obtained directly from the zone centre wavefunctions and energies.

4.3.1.1 Effective Mass Method

In this method, the **k.p** parameters are obtained from the effective masses calculated at the Γ-point in the empirical band structure. For the three uppermost valence bands and the lowest conduction band, the effective masses are calculated simply from the second derivative of the band energy with respect to wavevector. The values of the γ_i are then obtained from the effective masses using equations 4.19 [102]:

$$\begin{aligned}
\gamma_1 &= \frac{1}{2}\left(\frac{1}{m_{lh}} + \frac{1}{m_{hh}}\right) - \frac{E_p}{3E_g} \\
\gamma_2 &= \frac{1}{4}\left(\frac{1}{m_{lh}} - \frac{1}{m_{hh}}\right) - \frac{E_p}{6E_g} \\
\gamma_3 &= \frac{1}{4}\left(\frac{1}{m_{lh}} + \frac{1}{m_{hh}} - \frac{2}{m_{hh}^L}\right) - \frac{E_p}{6E_g}
\end{aligned} \tag{4.19}$$

In these equations, m_{hh}, m_{lh} and m_{so} are the effective masses (in units of the free electron mass) of the heavy-hole, light-hole, and spin split-off valence bands re-

spectively. m_{hh}^{L} is the heavy hole effective mass in the L direction. Unlike GaN and InN, AlN has an indirect band gap, but in the equations above the direct gap should be used. Note that the unmodified Luttinger parameters can be identified with the effective masses by comparing equations 4.17 and 4.19.

4.3.1.2 Direct Method

k.p parameters have also been calculated using a method that requires knowledge of the wavefunctions and energies at the centre of the Brillouin zone, $\mathbf{k} = 0$. When spin-orbit splitting is ignored, the top of the valence band is six-fold degenerate (including spin) and the zone centre valence band wavefunctions have Γ_{15} symmetry, commonly denoted by the labels X, Y and Z. As described previously, it is possible to adequately describe the valence band states at nonzero $\mathbf{k}$ using an expansion in terms of X, Y and Z and employing Löwdin's perturbation theory [100] to account for the remote bands. Local pseudopotentials had to be used to generate the empirical band structure for this method. Unlike local pseudopotentials, nonlocal pseudopotentials are dependent on $\mathbf{k}$, and perturbation theory using the zone centre wavefunctions and energies cannot account for this dependence.

The contributions to the unmodified Luttinger parameters from states with symmetries Γ_1, Γ_{15} and Γ_{12} (corresponding to singlets, triplets and doublets respectively) can be written as [104,105]:

$$\begin{aligned}
\sigma &= \frac{1}{3m_0}\sum_{\nu}^{\Gamma_1} |\langle X|p_x|\nu\rangle|^2/(\varepsilon_\nu - \varepsilon) \\
\pi &= \frac{1}{3m_0}\sum_{\nu}^{\Gamma_{15}} |\langle X|p_y|\nu\rangle|^2/(\varepsilon_\nu - \varepsilon) \\
\delta &= \frac{1}{6m_0}\sum_{\nu}^{\Gamma_{12}} |\langle X|p_x|\nu\rangle|^2/(\varepsilon_\nu - \varepsilon)
\end{aligned} \tag{4.20}$$

The unmodified Luttinger parameters (γ_i^L) are obtained from σ, π and δ using the

following relations [104]:

$$\begin{aligned} {\gamma_1}^L &= -1 + 2\sigma + 4\pi + 4\delta \\ {\gamma_2}^L &= \sigma - \pi + 2\delta \\ {\gamma_3}^L &= \sigma + \pi - \delta \end{aligned} \tag{4.21}$$

Pseudowavefunctions and energies were used to evaluate the terms and sums in equations 4.20. In these equations, 100 bands were used in each summation. Figure 4.1 show the convergence of σ, π and δ with respect to the number of bands used in each summation. From these, it is clear that the sums are rapidly convergent, and although 100 bands are used the parameters are converged to within 1% of their final values after just 20 bands. For σ, the greatest contribution by far is from the first conduction band. Similarly, the principal contributions to π and δ come from the first triplet and doublet respectively.

4.3.2 Results

The calculated values of the **k.p** parameters for zincblende materials are shown in Table 4.2. From this table, it is clear that the parameters obtained from the two methods are very similar. In particular, the values of γ_1 and γ_2 are within 1% of each other for the two methods, and the only significant difference between the two sets is in the value of γ_3. This parameter is larger for the direct method in all three cases, and the agreement is only to within about 10%. Reasonable values are also obtained for E_p and the parameter s, which is required by **k.p** theory to be less than unity.

The parameters from Table 4.2 were used in **k.p** calculations to obtain valence band structures close to the zone centre for AlN, GaN and InN. Figures 4.2, 4.3 and 4.4 show the **k.p** band structures of AlN, GaN and InN in the $\Gamma - L$ and $\Gamma - X$ directions. It is clear from these figures that the bands are parabolic close to the zone centre. For all

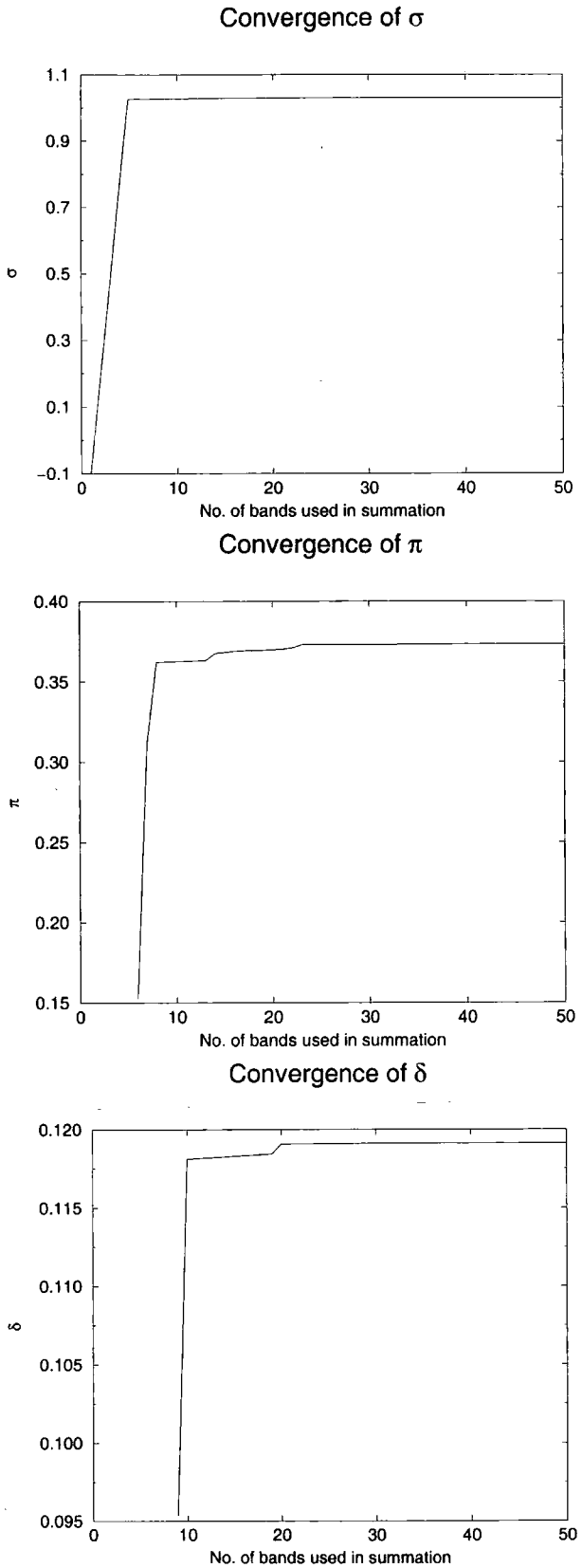

Figure 4.1: Convergence of σ, π and δ for zincblende structure GaN

	AlN		GaN		InN	
Method	(i)	(ii)	(i)	(ii)	(i)	(ii)
γ_1^L	1.842	-1.841	3.030	3.031	3.400	3.399
γ_2^L	0.431	-0.431	0.897	0.895	1.157	1.155
γ_3^L	0.755	0.772	1.282	1.285	1.468	1.480
γ_1	0.631	0.630	0.780	0.781	0.710	0.709
γ_2	-0.175	-0.175	-0.228	-0.230	-0.188	-0.190
γ_3	0.150	0.166	0.157	0.160	0.123	0.135
s	0.67	0.67	0.92	0.92	0.88	0.88
E_p	22.3	22.3	22.4	22.4	18.4	18.4

Table 4.2: **k.p** parameters for zincblende structure AlN, GaN and InN calculated from (i) the effective mass method and (ii) the direct method. The superscript ^{L} refers to the unmodified Luttinger parameters.

three of these materials, the agreement between the **k.p** band structure of both methods and that of the original empirical band structure is good. As expected, the agreement is less good away from the zone centre. In particular, there is no noticeable difference in the quality of fit that each methods produces for any of the three materials. Thus, although there is an appreciable difference in the Luttinger parameter γ_3 obtained by the two methods, this difference has a negligible effect on the resulting band structure.

The parameter sets obtained using the direct method can also be used to calculate values for the effective masses through use of equation 4.19. The effective masses calculated in this way are contrasted in Table 4.3 with those calculated from the curvature of the empirical band structure. For all three materials in the X direction, the effective masses of the light hole and heavy hole bands are in good agreement for both methods. However, due to the large value obtained for γ_3, the heavy hole mass in the L direction is slightly overestimated by the direct method.

In Table 4.3, the results of this work are also compared with the effective masses obtained by other workers. Additionally, the empirical pseudopotential parameters of Fan *et al.* [106] for AlN and GaN have been taken and used to calculate effective masses using both of the methods presented in this work. As in this work, the empirical pseudopotential parameters of Fan *et al.* were adjusted to fit the band energies obtained

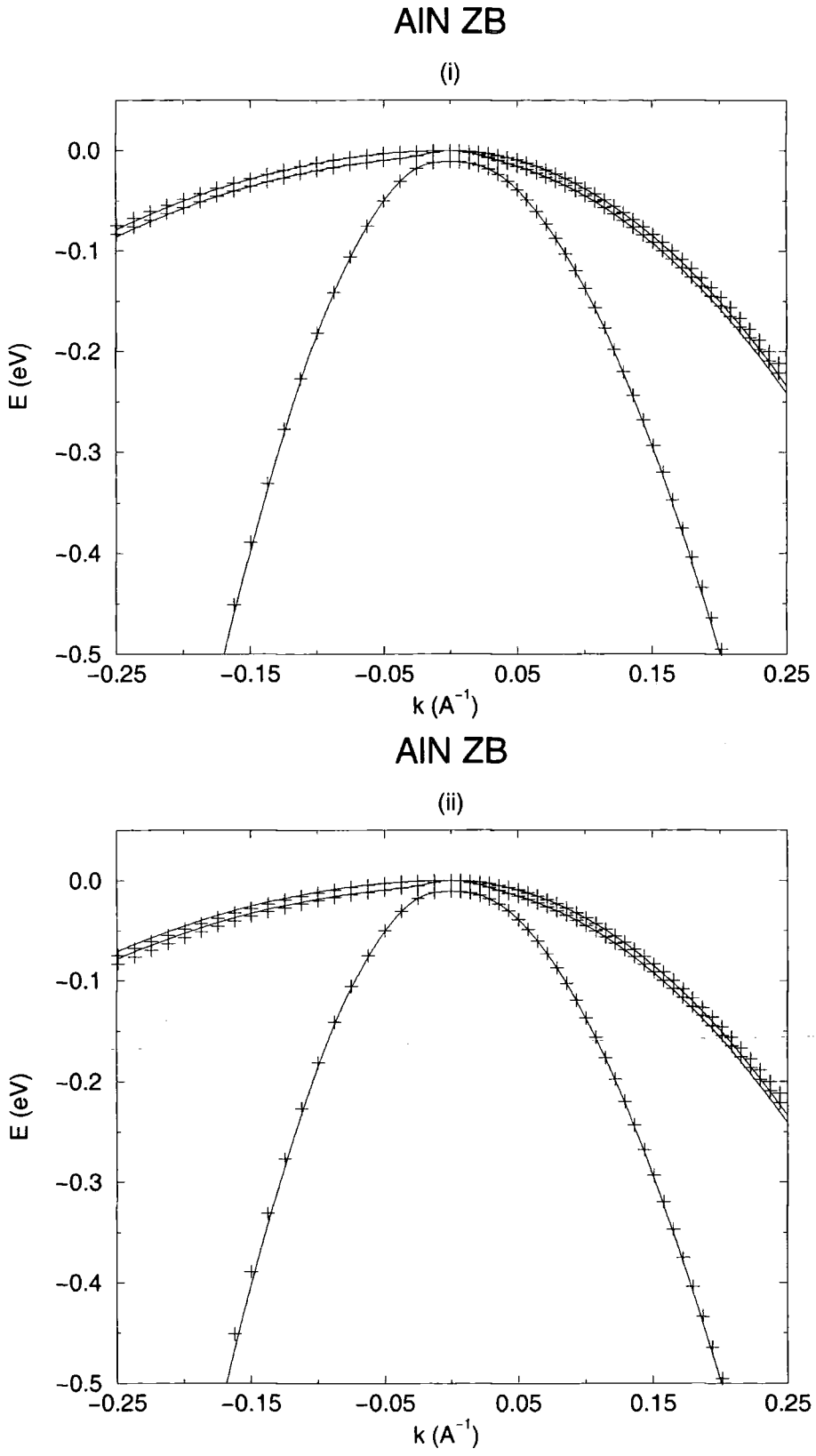

Figure 4.2: Band structure close to the zone centre (Γ-point) for zincblende structure AlN. The pluses represent the original band structure and the solid lines represent the **k.p** band structure. Negative k indicates the $\Gamma - L$ direction and positive k indicates the $\Gamma - X$ direction. (i) and (ii) show the empirical pseudopotential band structure and the **k.p** band structures obtained by the effective mass method and the direct method respectively.

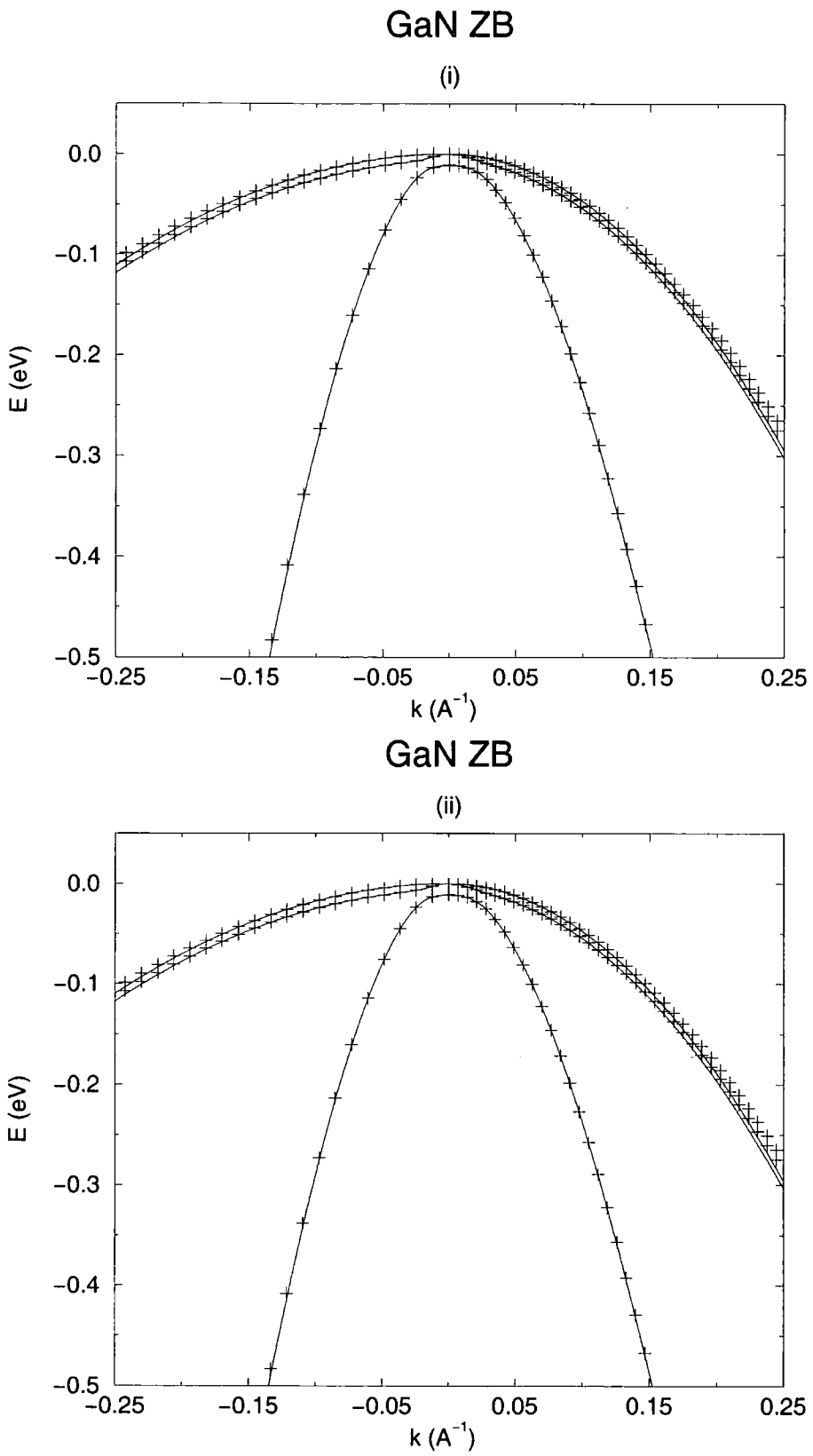

Figure 4.3: Band structure close to the zone centre for zincblende structure GaN. See Figure 4.2 for an explanation of the notation.

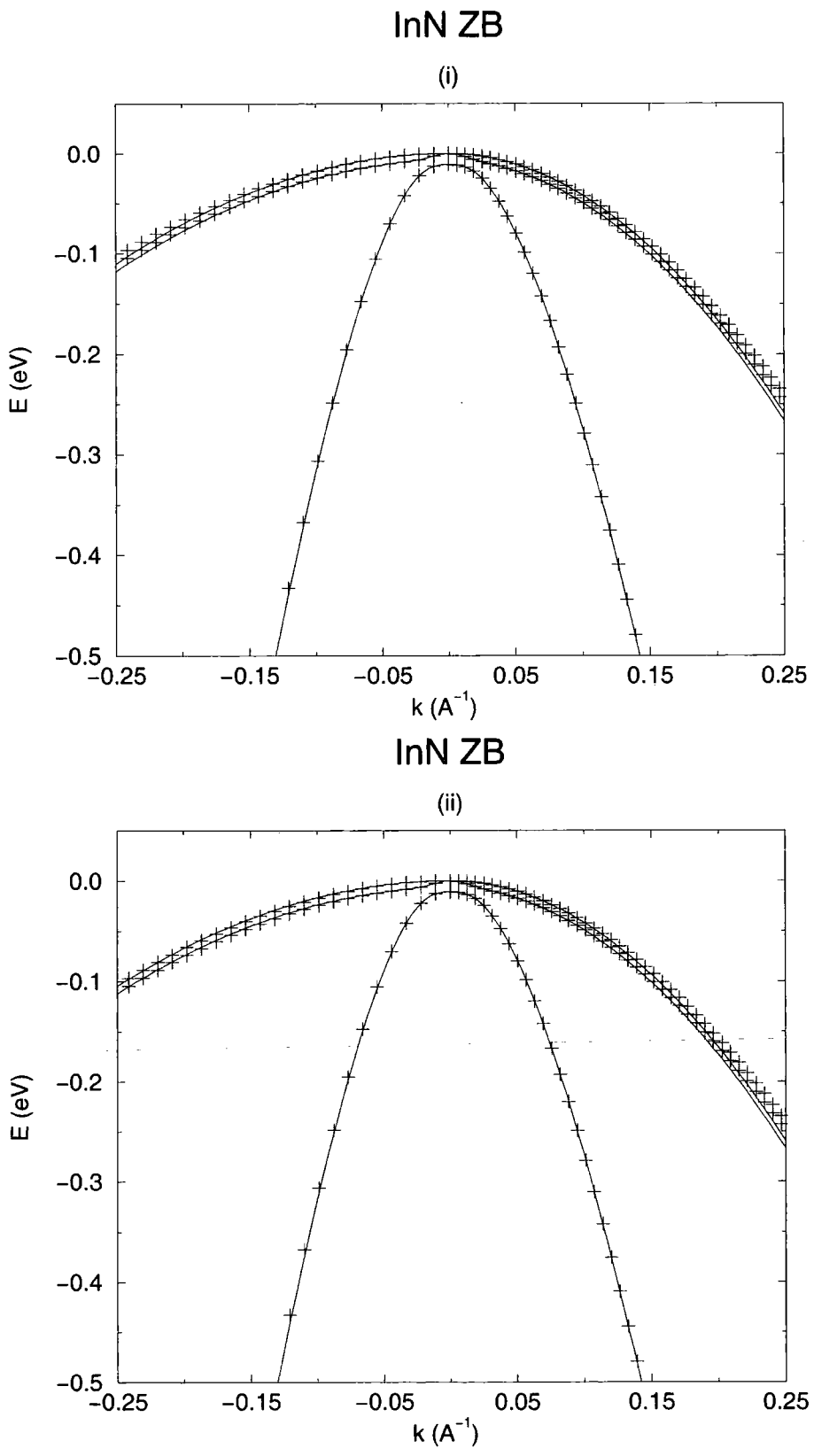

Figure 4.4: Band structure close to the zone centre for zincblende structure InN. See Figure 4.2 for an explanation of the notation.

	This work		Fan [106]		Meney [107]
	(i)	(ii)	(i)	(ii)	
AlN					
m_{so}	0.54		0.53		0.41
m_{lh}	0.37	0.37	0.35	0.35	0.28
m_{hh}	1.02	1.02	1.04	1.03	0.80*
m_{hh}^{L}	3.01	3.36	2.80	2.82	1.00*
m_c	0.23		0.21		0.21
GaN					
m_{so}	0.33		0.32		0.33
m_{lh}	0.21	0.21	0.21	0.21	0.21
m_{hh}	0.81	0.81	0.76	0.75	0.80*
m_{hh}^{L}	2.14	2.17	1.92	1.93	1.00*
m_c	0.13		0.13		0.15
InN					
m_{so}	0.30				0.31
m_{lh}	0.18	0.18			0.19
m_{hh}	0.92	0.92			0.80*
m_{hh}^{L}	2.15	2.28			1.00*
m_c	0.11				0.14

Table 4.3: Effective masses (in units of the free electron mass) for zincblende structure materials calculated (i) from the empirical band structure and (ii) from the directly obtained **k.p** parameters. The effective masses m_{so}, m_{lh} and m_{hh} are taken in the X direction. The effective mass m_{hh}^{L} is taken in the L direction. Additionally, we have taken the empirical band structure for AlN and GaN of Fan *et al.* and calculated effective masses from it using methods (i) and (ii). Values obtained by Meney *et al.* [107] are also shown (the values indicated by * are assumed rather than calculated).

from the quasiparticle calculations of Rubio *et al.* [76], and so there should be reasonable agreement in the effective masses obtained. As in the case of the results calculated using our own empirical band structure, there is excellent agreement between the effective masses obtained using both methods. In particular, the heavy hole mass in the L direction is in much better agreement. Also, there is good agreement between the effective masses obtained using Fan's bandstructure and those calculated using our own. In comparing the results of this work with those of Meney *et al.* [107], there is quite good agreement for GaN and InN, but less good agreement for AlN. As a whole, however, these results indicate that the direct method can generally be relied upon to produce accurate values for the effective masses.

4.4 Wurtzite Structure Materials

A six band **k.p** model has been used to describe the top of the valence band in a nitride material with the wurtzite structure. Following Chuang and Chang [108], the basis functions have been taken as:

$$\begin{aligned}
|u_1\rangle &= \frac{-1}{\sqrt{2}}|(X+iY)\uparrow\rangle \\
|u_2\rangle &= \frac{1}{\sqrt{2}}|(X-iY)\uparrow\rangle \\
|u_3\rangle &= |Z\uparrow\rangle \\
|u_4\rangle &= \frac{1}{\sqrt{2}}|(X-iY)\downarrow\rangle \\
|u_5\rangle &= \frac{-1}{\sqrt{2}}|(X+iY)\downarrow\rangle \\
|u_6\rangle &= |Z\downarrow\rangle
\end{aligned} \tag{4.22}$$

where $|X\rangle$, $|Y\rangle$ and $|Z\rangle$ indicate the symmetry of the valence band zone centre states, and the arrows represent the spin orientation. The z-direction corresponds to the c-axis of the wurtzite crystal.

The **k.p** Hamiltonian matrix in the basis defined by equation 4.4, where row/column i is associated with basis state $|u_i\rangle$, is

$$\begin{bmatrix} F & -K^* & -H_+^* & 0 & 0 & 0 \\ -K & G & H_- & 0 & 0 & \Delta \\ -H_+ & H_-^* & \lambda & 0 & \Delta & 0 \\ 0 & 0 & 0 & F & -K & H_- \\ 0 & 0 & \Delta & -K^* & G & -H_+^* \\ 0 & \Delta & 0 & H_-^* & -H_+ & \lambda \end{bmatrix}$$

where

$$\begin{aligned} F &= \Delta_1 + \Delta_2 + \lambda + \theta \\ G &= \Delta_1 - \Delta_2 + \lambda + \theta \\ \lambda &= \frac{\hbar^2}{2m_0}\left[A_1 k_z^2 + A_2(k_x^2 + k_y^2)\right] \\ \theta &= \frac{\hbar^2}{2m_0}\left[A_3 k_z^2 + A_4(k_x^2 + k_y^2)\right] \\ K &= \frac{\hbar^2}{2m_0} A_5 (k_x + ik_y)^2 \\ H_+ &= \frac{\hbar^2}{2m_0} A_6 (k_x + ik_y) k_z + iA_7(k_x + ik_y) \\ H_- &= \frac{\hbar^2}{2m_0} A_6 (k_x + ik_y) k_z - iA_7(k_x + ik_y) \\ \Delta &= \sqrt{2}\Delta_3 \end{aligned} \tag{4.23}$$

In the above expressions, $\Delta_1 = \Delta_{cr}$, the crystal field splitting energy, and $\Delta_2 = \Delta_3 = \Delta_{so}/3$, where Δ_{so} is the spin orbit splitting energy. The terms involving A_i (i=1–6) arise from the contribution of remote bands which are calculated using Löwdin's perturbation theory [100]. These parameters are analogous to the Luttinger parameters used in the zincblende case, and thus are commonly referred to as the Luttinger-like parameters. The Hamiltonian matrix above differs from that in reference [108] by the inclusion of terms linear in k which arise from the **k.p** term in the Hamiltonian and involve the coefficient A_7.

The **k.p** parameters can be calculated by using the pseudopotential wavefunctions and energies at the zone centre to evaluate the formulae for them. These are [108]:

$$A_1 = \frac{2m_0}{\hbar^2}L_2, \quad A_2 = \frac{2m_0}{\hbar^2}M_3, \quad A_3 = \frac{2m_0}{\hbar^2}(M_2 - L_2),$$
$$A_4 = \frac{2m_0}{\hbar^2}\left(\frac{L_1 + M_1}{2}\right), \quad A_5 = \frac{2m_0}{\hbar^2}\frac{N_1}{2}, \quad A_6 = \frac{2m_0}{\hbar^2}\frac{N_2}{\sqrt{2}} \tag{4.24}$$

where

$$
\begin{aligned}
L_1 &= \frac{\hbar^2}{2m_0}\left(1+\sum_\gamma \frac{2p^x_{X\gamma}p^x_{\gamma X}}{m_0\left(E_0-E_\gamma\right)}\right)\\
L_2 &= \frac{\hbar^2}{2m_0}\left(1+\sum_\gamma \frac{2p^y_{Y\gamma}p^y_{\gamma Y}}{m_0\left(E_0-E_\gamma\right)}\right)\\
M_1 &= \frac{\hbar^2}{2m_0}\left(1+\sum_\gamma \frac{2p^y_{X\gamma}p^y_{\gamma X}}{m_0\left(E_0-E_\gamma\right)}\right)\\
M_2 &= \frac{\hbar^2}{2m_0}\left(1+\sum_\gamma \frac{2p^z_{X\gamma}p^z_{\gamma X}}{m_0\left(E_0-E_\gamma\right)}\right)\\
M_3 &= \frac{\hbar^2}{2m_0}\left(1+\sum_\gamma \frac{2p^x_{Z\gamma}p^x_{\gamma Z}}{m_0\left(E_0-E_\gamma\right)}\right)\\
N_1 &= \frac{\hbar^2}{m_0^2}\sum_\gamma \frac{p^x_{X\gamma}p^y_{\gamma Y}+p^y_{X\gamma}p^x_{\gamma Y}}{m_0\left(E_0-E_\gamma\right)}\\
N_2 &= \frac{\hbar^2}{m_0^2}\sum_\gamma \frac{p^x_{X\gamma}p^z_{\gamma Z}+p^z_{X\gamma}p^x_{\gamma Z}}{m_0\left(E_0-E_\gamma\right)}
\end{aligned}
\tag{4.25}
$$

and $p^y_{X\gamma} = \langle X|p^y|\gamma\rangle$, $p^y = (\hbar/i)\,(\partial/\partial y)$.

The terms involving A_7 arise from matrix elements of the type $\langle X|k_x p^x|Z\rangle$ which result directly from the **k.p** term in the Hamiltonian rather than from remote band effects. Such terms vanish by symmetry in the zincblende structure but not in the wurtzite structure. In much previous work the parameter A_7 has been assumed to be zero, but recently it has been shown that its inclusion can give an improved description of the valence bands [109] and thus it is used here.

There are several matrix elements which involve A_7 in equation 4.4. To evaluate A_7 it is therefore necessary to consider one of these, such as H_+:

$$
H_+ = \frac{\hbar^2}{2m_0}A_6(k_x+ik_y)k_z + iA_7(k_x+ik_y) \tag{4.26}
$$

The term H_+ above is derived from the u_2 and u_3 states, and is defined as

$$\begin{aligned}\langle u_2|H|u_3\rangle &= \frac{1}{\sqrt{2}}\langle (X - iY)\uparrow |H|Z\uparrow\rangle \\ &= \frac{1}{\sqrt{2}}\langle X|H|Z\rangle + \frac{i}{\sqrt{2}}\langle Y|H|Z\rangle \end{aligned} \tag{4.27}$$

where H is the Hamiltonian in equation 4.7. The term corresponding to A_7 is derived from the $(\hbar/m_0)\mathbf{k}.\mathbf{p}$ term in H. For this, only the terms $\langle X|p^x|Z\rangle$ and $\langle Y|p^y|Z\rangle$ are nonzero. The term in A_7 is thus

$$\frac{\hbar}{m_0\sqrt{2}}\left[k_x\langle X|p^x|Z\rangle + ik_y\langle Y|p^y|Z\rangle\right] \tag{4.28}$$

Additionally, $\langle X|p^x|Z\rangle = \langle Y|p^y|Z\rangle$ by symmetry, such that the term corresponding to A_7 can be simplified to

$$\frac{\hbar}{m_0\sqrt{2}}\langle X|p^x|Z\rangle\left[k_x + ik_y\right] \tag{4.29}$$

Therefore, the term A_7 in equation 4.26 can be written as

$$A_7 = \frac{-i\hbar}{m_0\sqrt{2}}\langle X|p^x|Z\rangle \tag{4.30}$$

Other terms linear in k can also appear in the Hamiltonian matrix due to a linear k term in the spin-orbit interaction or through remote band effects of the **k**-independent part of the spin-orbit interaction in association with the **k.p** term. These terms either vanish or are very small in zincblende structure materials, and are expected to be so here as well, particularly in view of the very small spin-orbit interaction in the nitrides.

This view appears to be confirmed by the results presented later in this chapter.

4.4.1 Calculation of the k.p parameters

The wurtzite empirical band structures of AlN, GaN and InN of Chapter 3 have been used as the basis for a series of **k.p** calculations, and again two methods have been employed to to obtain the **k.p** parameters from them. It is not possible to calculate the parameters from the effective masses as was done in the zincblende case because the effective masses themselves are difficult to define in the wurtzite structure, since many of the bands are extremely non-parabolic. This causes problems with the calculation of convergent values for the second derivative of the band energy with respect to wave-vector at points progressively closer to the Γ point. For this reason, a fitting method was adopted instead. As for the zincblende structure, a direct method was used to obtain another set of **k.p** parameters.

4.4.1.1 Fitting Method

In this method, the values of the A_i were obtained through a simple Monte-Carlo fitting procedure to the band structure. The empirical pseudopotential method was used to calculate the band structure at several k-points over the range shown in Figure 4.7. The **k.p** method was then used to calculate the band structure at k-points in these two directions, using a set of A_i values. These A_i values were then varied using a Monte-Carlo technique, and the process systematically repeated until the **k.p** eigenvalues were in good agreement with those of the empirical pseudopotential calculations. This Monte-Carlo approach involved minimising the value F, where

$$F = \sum_i \sum_j w(i,j) \left| E_{kp}(i,j) - E_{epm}(i,j) \right|^2 \tag{4.31}$$

Here, $E_{kp}(i,j)$ and $E_{epm}(i,j)$ refer to the empirical and **k.p** eigenvalues of band j at k-

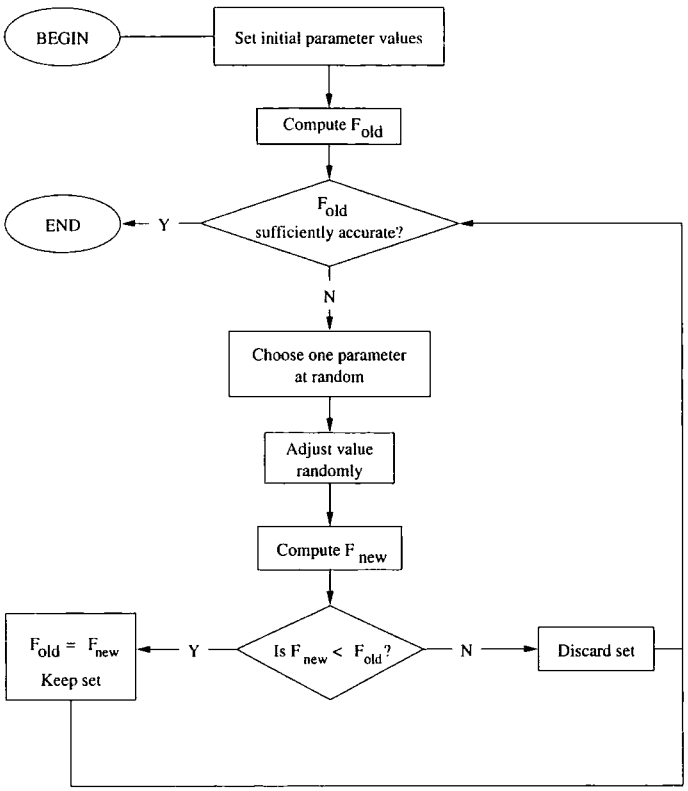

Figure 4.5: The Monte Carlo algorithm used to fit **k.p** parameters to empircial band structure.

point i respectively. The parameter $w_{i,j}$ is a weight attached to each k point and band. For this work, the empirical and **k.p** energies used in the fitting were calculated in three directions from the zone centre; one in the k_z-direction, one in an in-plane direction, and one in a direction involving k_x, k_y and k_z. These directions were necessary to obtain an accurate value for the parameter A_6, which only contributes to the Hamiltonian in 4.4 when there are nonzero values for both $(k_x + ik_y)$ and k_z. In this approach, the band structure at the zone centre was deemed to be important, and thus bands at k points close to the zone centre were typically weighted more than those further away. Also, the k-points in the in-plane directions were weighted more than those in the k_z direction, since the band structures in these directions generally have more structure, due in part to the inclusion of A_7. A schematic representation of the fitting procedure is shown in Figure 4.5.

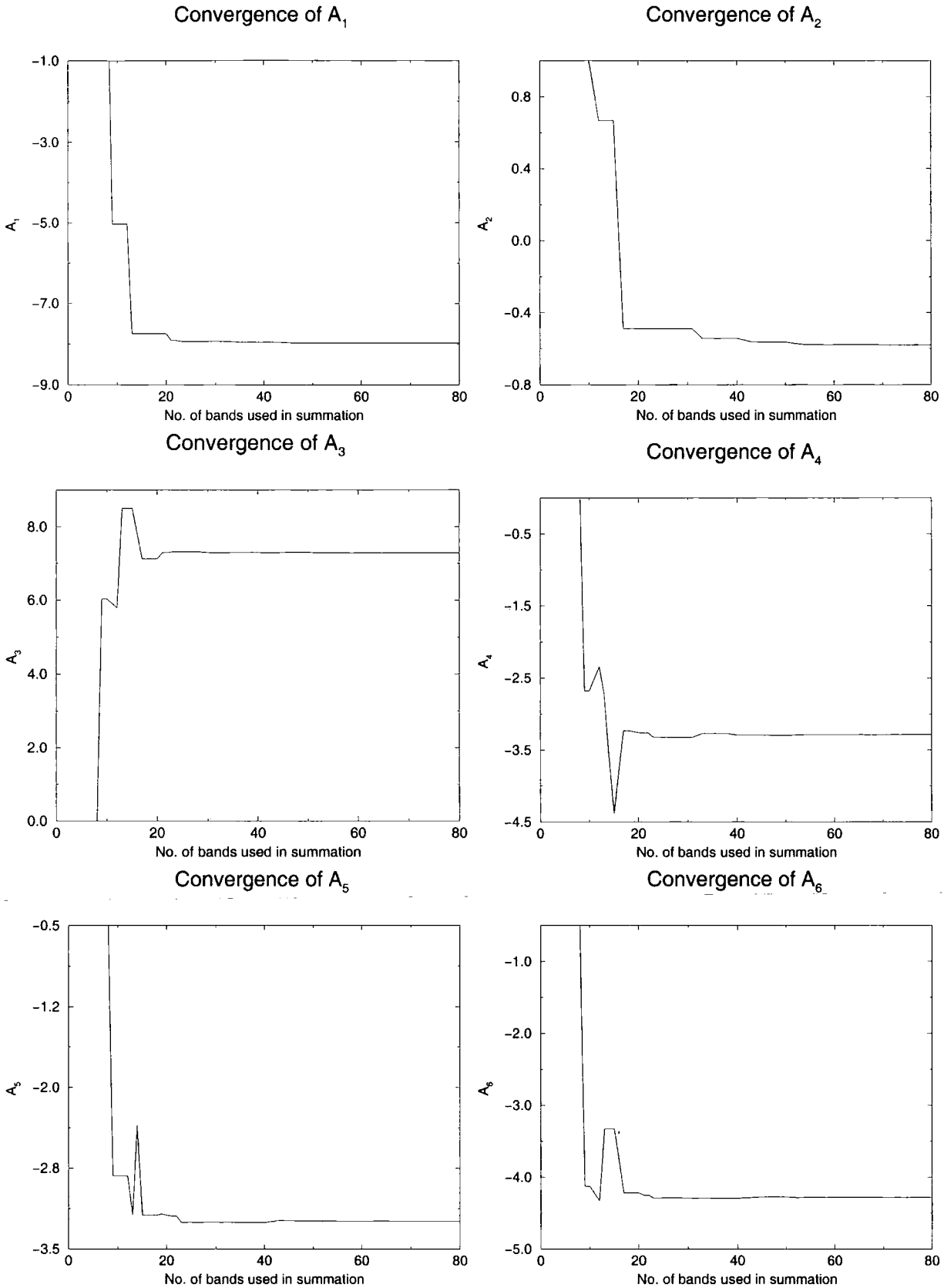

Figure 4.6: Convergence of the **k.p** parameters for wurtzite structure GaN

4.4.1.2 Direct Method

In the second method, the **k.p** parameters were evaluated directly from equations 4.24, 4.25 and 4.30. To obtain these parameters, 250 bands were included in each summation in equation 4.25, but in practice it is possible to use much fewer bands with no significant loss of accuracy. This can be seen in Figure 4.6, which shows the convergence of the **k.p** parameters with respect to the number of bands used in the summations in equation 4.25. It is important to note that this summation over bands is actually done for the L, M and N parameters of Equation 4.25, and the actual Luttinger-like parameters are then derived from these. From Figure 4.6, it is clear that the sums for all of the parameters are rapidly convergent, with all of the **k.p** parameters converged to within 1% of their final value after a summation over ≈ 60 bands. Also note that in the summation over bands, entire doublets or triplets should be included. In a triplet therefore, all three bands should be included rather than just one or two. However, since the sums are converged using 250 bands, in practice this makes no observable difference.

In comparison with the zincblende case, these parameters converge much more slowly. While for the zincblende structure materials only the first few bands were important in the summation, there are still significant contributions from more energetically remote bands in the wurtzite case. However, the expansion does still converge, and thus the parameters obtained using this method are unique for a given empirical band structure.

4.4.2 Results

The calculated values of A_i (i=1-7) for AlN, GaN and InN, together with those from other calculations, are presented in Table 4.4. The parameters from the two different methods were used in **k.p** calculations of the band structures of GaN, AlN and InN in the region of interest close to the Γ-point. The resulting band structures are shown in

Figures 4.7, 4.8 and 4.9. Note that AlN has a negative crystal field splitting, and thus the ordering of the bands is different from that of GaN and InN. For all three materials, both the Monte-Carlo fitting approach and the direct method produce band structure in very good agreement with that of the original empirical pseudopotential calculation. In particular, the inclusion of A_7 accurately models the lifting of the degeneracy near the anticrossing feature seen in GaN and InN. Also, note that these bands are extremely non-parabolic, and thus obtaining effective masses (and subsequently **k.p** parameters) from them via parabolic line fits is not preferable to the methods presented here.

From Table 4.4, it is clear that the parameters obtained by the Monte-Carlo fitting and the direct approach are somewhat different from each other for each of the materials. For AlN, the difference in the A_i values is typically about 10%. For GaN, the agreement is better, with differences of around 6%, and for InN the values generally differ by about 13%. However, despite these differences, both methods give a very similar quality of fit to the original band structure, as can be concluded from the results shown in Figures 4.7, 4.8 and 4.9. This demonstrates the potential inconsistency of the fitting method, in that two different sets of parameters appear to result in equally good agreement with the initial band structure. Indeed, for the fitting procedure a range of equally valid parameter sets exist, and because of this it can not be relied upon to give consistent **k.p** parameters when independent fits are made for a series of alloy compositions. In contrast to the fitting approach, the method of obtaining **k.p** parameters directly from the zone centre wavefunctions and energies has a firm physical basis and presents an unambiguous route to calculating these quantities from a given original band structure. Consequently, this method can be applied to a series of alloy compositions with the expectation that consistent **k.p** parameters will be obtained.

In addition to the results calculated in this work, Table 4.4 also shows the **k.p** parameters that have been reported in the literature by other workers. The most information exists for GaN. For this, the results of this work are most similar to those of Yeo *et al.* [77], who also employ an empirical pseudopotential method. Additionally, a

	Present Work		Ref. [77]		Ref. [110]		Ref. [53]	Ref. [111]
	Fit	Direct	Fit	Direct*	Fit	Quasicubic		
AlN								
A_1	-4.367	-4.711			-3.86	-3.74	-4.06	-3.82
A_2	-0.518	-0.476			-0.25	-0.23	-0.26	-0.22
A_3	3.854	4.176			3.58	3.51	3.78	3.54
A_4	-1.549	-1.816			-1.32	-1.76	-1.86	-1.16
A_5	-1.680	-1.879			-1.47	-1.52	-2.02	-1.33
A_6	-2.103	-2.355			-1.64	-1.83	-	-1.25
A_7	0.204	0.096			-	0.00	-	0.00
Δ_1 (meV)	-93.2	-93.2			-215.0	-	-58.5	-219.0
Δ_2 (meV)	3.7	3.7			6.8	6.3	6.8	6.6
Δ_3 (meV)	3.7	3.7			5.7	6.3	6.8	6.7
GaN								
A_1	-7.706	-7.979	-7.24	-7.17	-6.40	-6.36	-6.27	-6.40
A_2	-0.597	-0.581	-0.51	-0.44	-0.50	-0.51	-0.96	-0.80
A_3	7.030	7.291	6.73	6.64	5.90	5.85	5.70	5.93
A_4	-3.076	-3.289	-3.36	-3.62	-2.55	-2.92	-2.84	-1.96
A_5	-3.045	-3.243	-3.35	-3.57	-2.56	-2.60	-3.18	-2.32
A_6	-4.000	-4.281	-4.72	-4.04	-3.06	-3.21	-4.96	-3.02
A_7	0.194	0.179	-	-	0.20	0.00	< 0.27	0.35
Δ_1 (meV)	22.3	22.3	21.0	21.0	36.0	-	72.0	24.0
Δ_2 (meV)	3.7	3.7	3.7	3.7	5.0	6.3	5.4	5.4
Δ_3 (meV)	3.7	3.7	3.7	3.7	5.9	6.3	6.8	6.8
InN								
A_1	-9.470	-10.841	-9.28					
A_2	-0.641	-0.651	-0.60					
A_3	8.771	10.100	8.68					
A_4	-4.332	-4.864	-4.34					
A_5	-4.264	-4.825	-4.32					
A_6	-5.546	-6.556	-6.08					
A_7	0.278	0.283	0.00					
Δ_1 (meV)	37.3	37.3	17.0					
Δ_2 (meV)	3.7	3.7	3.7					
Δ_3 (meV)	3.7	3.7	3.7					

Table 4.4: **k.p** parameters for wurtzite structure AlN, GaN and InN. The A_i are in units of $\hbar^2/2m_0$, except A_7 where the units are eVÅ. * These are values calculated in this work using the direct method from the empirical band structure parameters of Ref. [77].

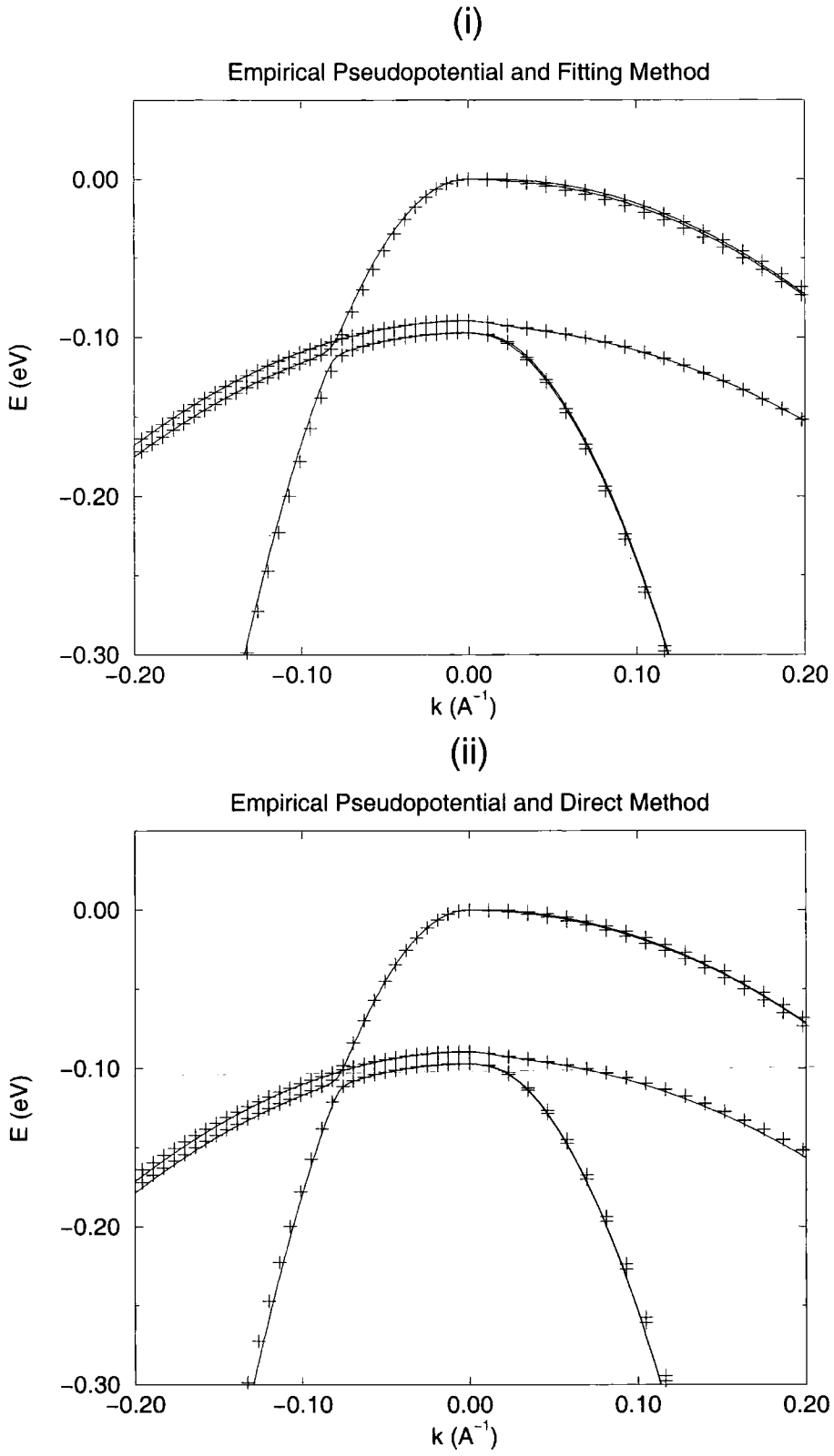

Figure 4.7: Band structure close to the zone centre (Γ-point) for wurtzite structure AlN. The pluses represent the original band structure and the solid lines represent the **k.p** band structure. Negative k indicates a c-axis ($\Gamma - A$) direction and positive k indicates an in-plane ($\Gamma - M$) direction. (i) and (ii) show the empirical pseudopotential band structure and the **k.p** band structures obtained via the fitting method and the direct method respectively.

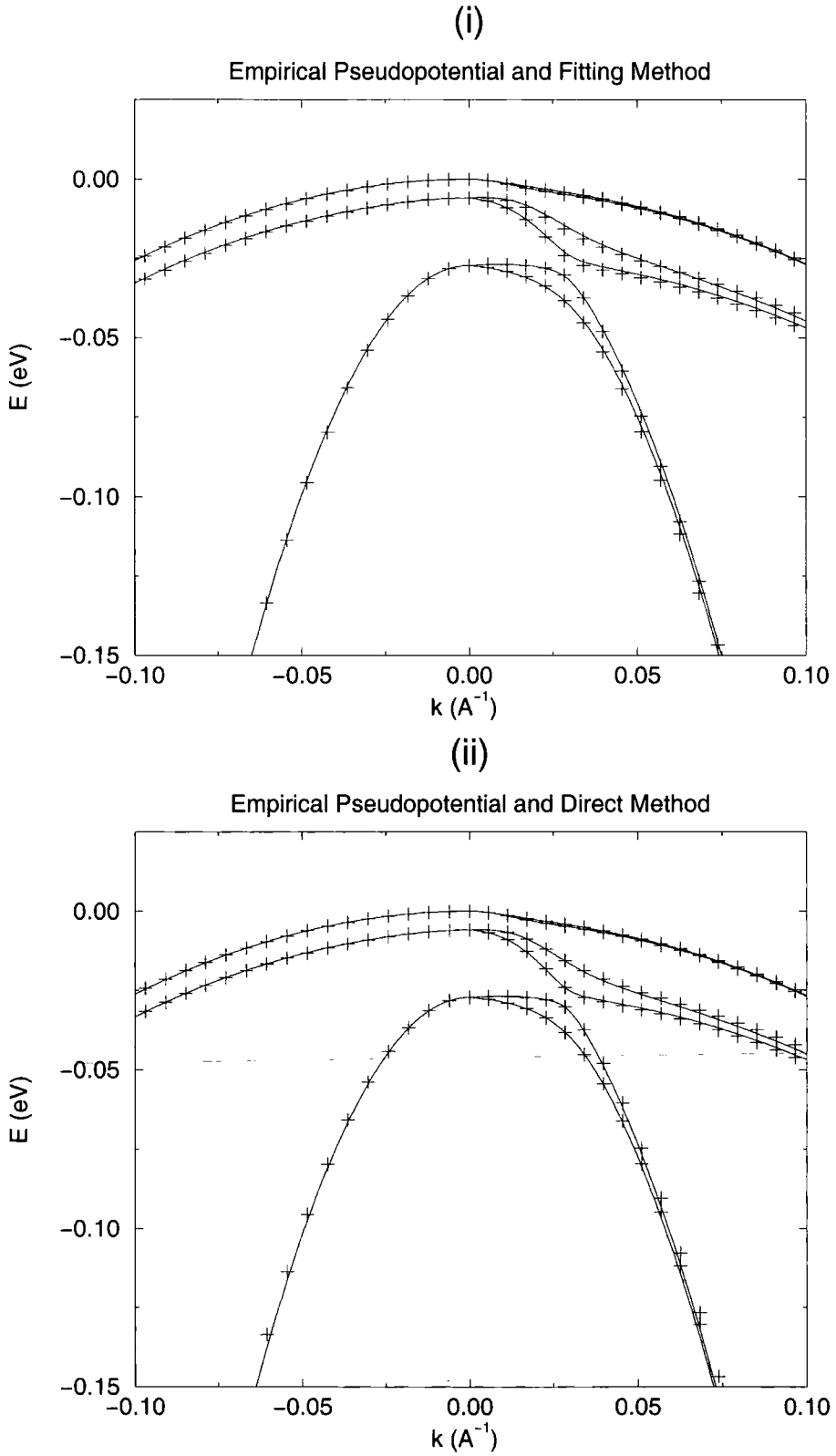

Figure 4.8: Band structure close to the zone centre for wurtzite structure GaN. See Figure 4.7 for an explanation of the notation.

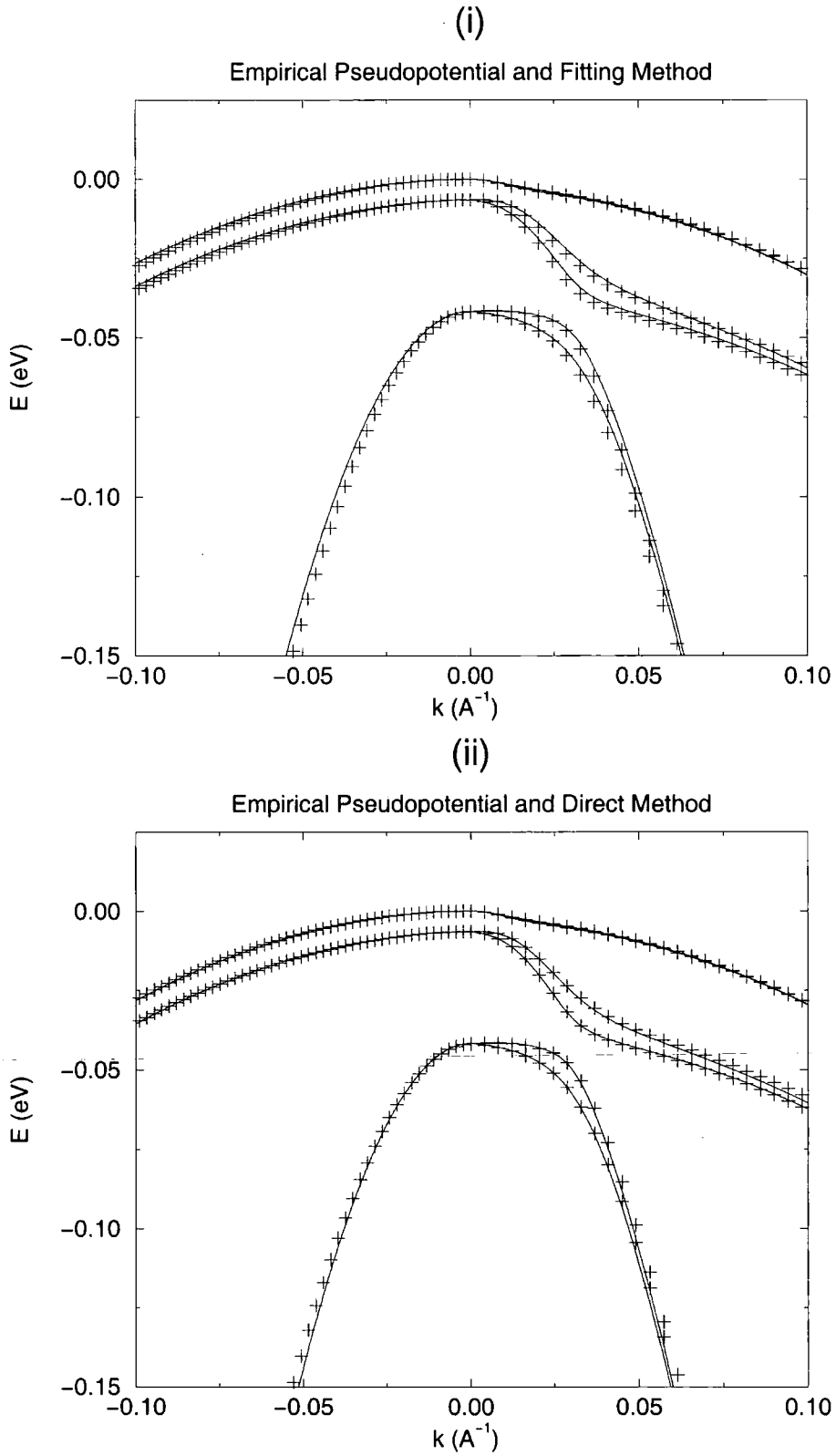

Figure 4.9: Band structure close to the zone centre for wurtzite structure InN. See Figure 4.7 for an explanation of the notation.

new set of A_i (i=1-6) have been calculated directly using Yeo's empirical band structure. There is good agreement between this new set and that originally obtained by Yeo *et al.* , who calculated their values using a 3-dimensional fitting procedure within the Brillouin zone. The other values quoted in Table 4.4 are based on first principles calculations, and while qualitatively similar to the results of this work, the individual A_i parameters can differ significantly. This is to be expected, given that the empirical band structure used in this work is different from the first principles band structures used in the fittings for those results.

The results of this work in Table 4.4 can also be considered with respect to the cubic approximation to the wurtzite structure [53,112]. This approximation is based on the similarity between the zincblende and wurtzite structures, in that they are both tetrahedrally bonded but with different stacking arrangements. For the cubic approximation, the following relations can be established between the Luttinger-like parameters:

$$\begin{aligned} A_1 &= A_2 + 2A_4 \\ A_3 &= -2A_4 \\ A_3 + 4A_5 &= \sqrt{2}A_6 \end{aligned} \tag{4.32}$$

Although there is no requirement for these relationships to be satisfied in our calculations, we note that substitution of the values of Table 4.4 in equation 4.32 show that the cubic approximation is satisfied approximately. For AlN, comparing the actual value on the left hand side of equation 4.32 with the value determined by the right hand side, the relations are satisfied to within 20% for the fitting method and to within 13% for the direct method. For GaN, the values are within 13% for the fitting method and 10% for the direct method. For InN, the approximation holds somewhat better,

with values for both methods within about 5% of those predicted.

The Luttinger-like parameters presented in Table 4.4 can also be used to calculate the valence band effective masses in the k_z direction ($m^{\|}$) and in the in-plane direction ($m^{\perp}$) using the following equations [53]:

$$\begin{aligned}
\frac{1}{m_{hh}^{\|}} &= -(A_1 + A_3) \\
\frac{1}{m_{lh}^{\|}} &= -(A_1 + A_3) \\
\frac{1}{m_{cr}^{\|}} &= -A_1 \\
\frac{1}{m_{hh}^{\perp}} &= -(A_2 + A_4 - A_5) \\
\frac{1}{m_{lh}^{\perp}} &= -(A_2 + A_4 + A_5) - 2\frac{A_7^2}{|\Delta_{cr}|} \\
\frac{1}{m_{cr}^{\perp}} &= -A_2 + 2\frac{A_7^2}{|\Delta_{cr}|}
\end{aligned} \tag{4.33}$$

Here, m_{hh}, m_{lh} and m_{cr} are the effective masses of the heavy-hole, light-hole and crystal field split-off bands respectively.

The resulting valence band effective masses for the wurtzite materials calculated using both the fitted and directly obtained **k.p** parameters are shown in Table 4.5. For both methods, the effective masses are quite similar, as is to be expected given that they are produced from the same empirical band structure. However, there are some differences in the masses, most notably in the in-plane direction. In this direction, the Luttinger-like parameter A_7 plays an important role, and thus large discrepancies in this parameter between the two methods, such as in AlN, can cause quite different effective masses. This again demonstrates the potential inconsistency of the fitting method, in that the various parameter sets obtained could produce quite different effective masses. For the direct method, only one set of effective masses is obtained. Moreover, it can be argued that the masses are more accurate than those calculated from the fitted

k.p parameters. This is because in the zincblende case the direct method produced effective masses in very good agreement with those obtained from the curvature of the empirical band structure near the zone centre, and by extension it is reasonable to believe that this is also true for the wurtzite case.

Table 4.5 also compares our calculated effective masses with others in the literature. These other calculations obtain their parameters by fitting to first principles or empirical band structure, and are qualitatively similar to our own results. In particular, the effective masses of this work are in the best agreement with those of Yeo *et al.* [77]. For AlN, however, the effective masses of this work are significantly lower than those obtained by other workers. From chapter 3, the antisymmetric $V(q)$ curve for AlN was quite large, and this is probably the origin of any discrepancies.

4.5 Wurtzite Alloys

Alloys of the nitride materials are very important for modern devices. For example, InGaN quantum wells are very important for light emitting devices. This is because incorporation of small amounts of In in GaN considerably increases the luminescence efficiency [15,113–115]. In addition, alloys also afford a degree of control over various properties that would be unavailable with bulk materials.

4.5.1 The Virtual Crystal Approximation

In this work, the empirical pseudopotential method has been used to study the band structure of wurtzite alloys involving AlN, GaN and InN. For this, a simple virtual vrystal approximation (VCA) model was adopted. In the VCA model, the alloy is described as a uniform crystal comprised of virtual atoms which are a concentration-weighted average of the real atoms. Thus for the alloy $A_xB_{1-x}N$, virtual atoms replace those of A and B, and the VCA implicitly assumes that the A-N and B-N bond lengths are the same. However, this is not the case for a real alloy, where the A-N and B-

	Present Work		Ref. [77]		Ref. [110]		Ref. [53]	Ref. [111]
	Fit	Direct	Fit	Direct*	Fit	Quasicubic		
AlN								
$m_{cr}^{\parallel}$	0.23	0.21			0.26	0.27	0.25	0.26
$m_{lh}^{\parallel}$	1.95	1.87			3.53	4.41	3.53	3.57
$m_{hh}^{\parallel}$	1.95	1.87			3.53	4.41	3.53	3.57
$m_{c}^{\parallel}$	0.24	0.24			0.35	0.33	0.33	0.32
$m_{cr}^{\perp}$	1.33	1.89			4.05	4.41	3.81	4.54
$m_{lh}^{\perp}$	0.28	0.24			0.33	0.29	0.24	0.37
$m_{hh}^{\perp}$	2.58	2.42			11.14	2.18	10.42	20.00
$m_{c}^{\perp}$	0.25	0.25			0.35	0.33	0.25	0.33
GaN								
$m_{cr}^{\parallel}$	0.13	0.13	0.14	0.14	0.16	0.16	0.15	0.16
$m_{lh}^{\parallel}$	1.48	1.47	1.96	1.90	2.00	1.96	1.10	2.12
$m_{hh}^{\parallel}$	1.48	1.47	1.96	1.90	2.00	1.96	1.10	2.12
$m_{c}^{\parallel}$	0.14	0.14	0.19	0.19	0.23	0.19	0.20	0.18
$m_{cr}^{\perp}$	0.67	0.74	1.96	1.48	1.49	1.96	1.10	0.29
$m_{lh}^{\perp}$	0.17	0.16	0.14	0.14	0.18	0.16	0.15	0.42
$m_{hh}^{\perp}$	1.59	1.59	1.87	2.03	2.04	1.20	1.65	2.27
$m_{c}^{\perp}$	0.15	0.15	0.17	0.17	0.19	0.19	0.18	0.20
InN								
$m_{cr}^{\parallel}$	0.11	0.09	0.10					
$m_{lh}^{\parallel}$	1.43	1.35	1.67					
$m_{hh}^{\parallel}$	1.43	1.35	1.67					
$m_{c}^{\parallel}$	0.10	0.10	0.11					
$m_{cr}^{\perp}$	0.58	0.56	1.67					
$m_{lh}^{\perp}$	0.12	0.11	0.11					
$m_{hh}^{\perp}$	1.41	1.45	1.61					
$m_{c}^{\perp}$	0.10	0.10	0.10					

Table 4.5: Effective masses (in units of the free electron mass) for wurtzite structure AlN, GaN and InN. Also shown are the results of other calculations which obtain the parameters by fitting to first-principles or empirical band structure. *These are our directly calculated values using the empirical band structure parameters of Ref. [77]

N bond lengths tend to preserve their respective binary crystal values and the local atomic environments are not identical. Despite the rather simplistic model of the crystal structure, the VCA has been successfully applied to other III-V systems, and is used in this work with a recognition of its limitations. In particular, it is interesting to compare the VCA results with those of more sophisticated calculations. We might expect rather better agreement for GaAlN alloys than for those containing In, since the AlN and GaN bond lengths are rather similar but much different for InN.

For the alloy $A_xB_{1-x}N$, the lattice constants of the virtual crystal are given by Vegard's law which predicts a simple linear interpolation between the lattice constants of AN and BN:

$$\begin{aligned} a_{ABN} &= x a_{AN} + (1-x) a_{BN} \\ c_{ABN} &= x c_{AN} + (1-x) c_{BN} \end{aligned} \tag{4.34}$$

A form must also be adopted for the variation of the the internal parameter u, which has a strong affect on the crystal field splitting. In the absence of any information on this quantity for the ternary nitride alloys, this is also assumed to vary according to Vegard's Law:

$$u_{ABN} = x u_{AN} + (1-x) u_{BN} \tag{4.35}$$

Once the structural parameters of the alloy have been obtained using equations 4.34 and 4.35 above, the parameters which describe the pseudopotential can also be derived.

4.5.2 Calculation of Form Factors for Alloys

It was shown in chapter 3 that the form factor at a particular G-value is given by:

$$V(\mathbf{G}) = \frac{1}{\Omega}\int_{\Omega} V(\mathbf{r})e^{-i\mathbf{G}.\mathbf{r}}d\mathbf{r} \tag{4.36}$$

As discussed further in chapter 5, the concept of transferable atomic pseudopotentials allows the pseudopotential obtained for a bulk material also to be used for its associated alloys. As for the lattice constants, in our approximate model the form factors for alloys were obtained using Vegard's Law. The symmetric and antisymmetric form factors for $A_{1-x}B_xN$ at a given G are taken to be:

$$V_{alloy}(\mathbf{G}) = V_A(\mathbf{G})(1-x) + V_B(\mathbf{G})x \tag{4.37}$$

Here, the individual $V_A(G)$, $V_B(G)$ are obtained using equation 4.36 at the G-values relevant to the alloy and using the new alloy volume. Figure 4.10 shows the symmetric and antisymmetric form factors for the wurtzite alloy $Ga_{0.5}In_{0.5}N$, together with the $V_s(q)$ and $V_a(q)$ curves for the two binary compounds.

4.5.3 Results

A band structure property of particular interest is the variation of the band gap with alloy composition. This has been studied for the wurtzite structure materials using two different calculations for unstrained and strained alloys respectively. For the unstrained case, both the a and c lattice parameters for the alloy are obtained using Vegard's Law. This corresponds to the ideal situation where the alloy is grown on a lattice-matched substrate. For the strained case, the alloy is under biaxial strain such that the in-plane lattice parameter a is held at the value of one of the binary compounds for all

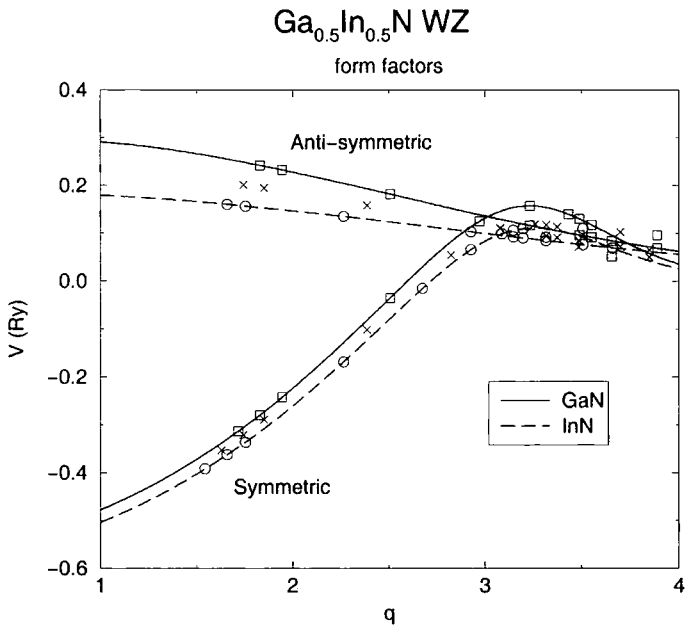

Figure 4.10: The symmetric and antisymmetric form factors for wurtzite structure $Ga_{0.5}In_{0.5}N$ alloy calculated from the $V(q)$ curves of bulk GaN and InN. The form factors for GaN and InN are represented by squares and circles respectively. The alloy form factors are indicated by crosses.

alloy compositions. The practical relevance of this is the common situation where the alloy GaInN, say, is grown on GaN. In the strained case, the value of c for the other binary compound at this value of a is obtained from first principles calculations using CASTEP.

The structure of the top of the valence band has also been studied for several alloy compositions of GaAlN, GaInN and AlInN. In addition, **k.p** parameters have been obtained for these materials using the direct method described in Chapter 3.

4.5.3.1 GaAlN

The variation of the band gap with alloy composition for $Ga_{1-x}Al_xN$ is shown in Figure 4.11. From this, it can be seen that in this work the band gap varies linearly with composition for both the strained and unstrained cases. There is little difference between the variation of the band gap for the strained and unstrained cases. For both of these, the results of this work are in good agreement with the experimental results

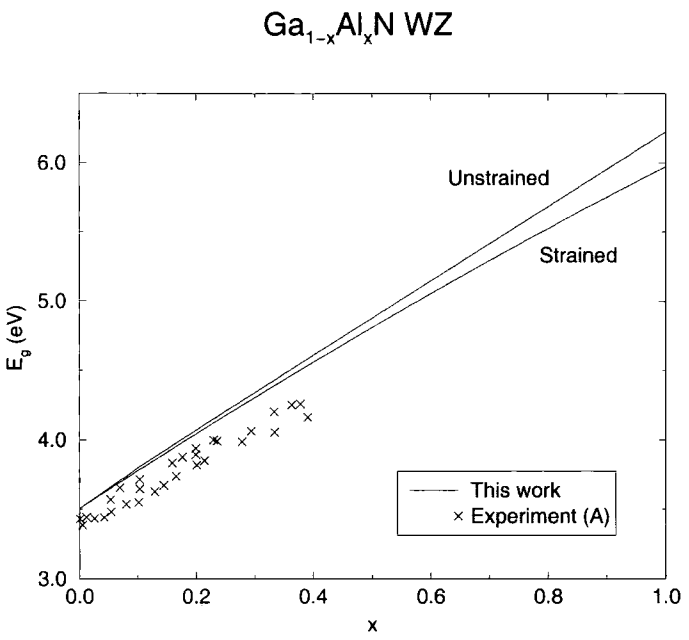

Figure 4.11: The band-gap E_g as a function of alloy composition x in the wurtzite alloy $Ga_{1-x}Al_xN$. Shown is a comparison to experimental data (A) compiled in ref. [116]

of Bergmann *et al.* [116]. Note that there is a slight offset in the theoretical and experimental values of E_g at $x = 0$. This is due to the choice of 3.50 eV for the band gap of GaN in the original fit, and discounting this the actual gradients are comparable.

Band structure diagrams close to the zone centre for GaAlN for the compositions $x = 0.05$, 0.10, 0.15 and 0.20 are shown in Figure 4.12. This figure also shows **k.p** band structures calculated using the direct method, and as for the bulk materials there is excellent agreement between the two band structures. The **k.p** parameters calculated using the direct method are shown in Table 4.6. In this table, the Luttinger-like parameters are seen to vary consistently with composition.

From Figure 4.12, it can be seen that with increasing AlN composition the crystal field splitting decreases and becomes negative, resulting in a flip-over in the normal ordering of the bands at approximately $x = 0.1$. The reasons for this behaviour are investigated in chapter 5. Experimentally, Hangleiter *et al.* [117] observe TE mode gain in GaAlN quantum wells, consistent with the heavy hole being band above the

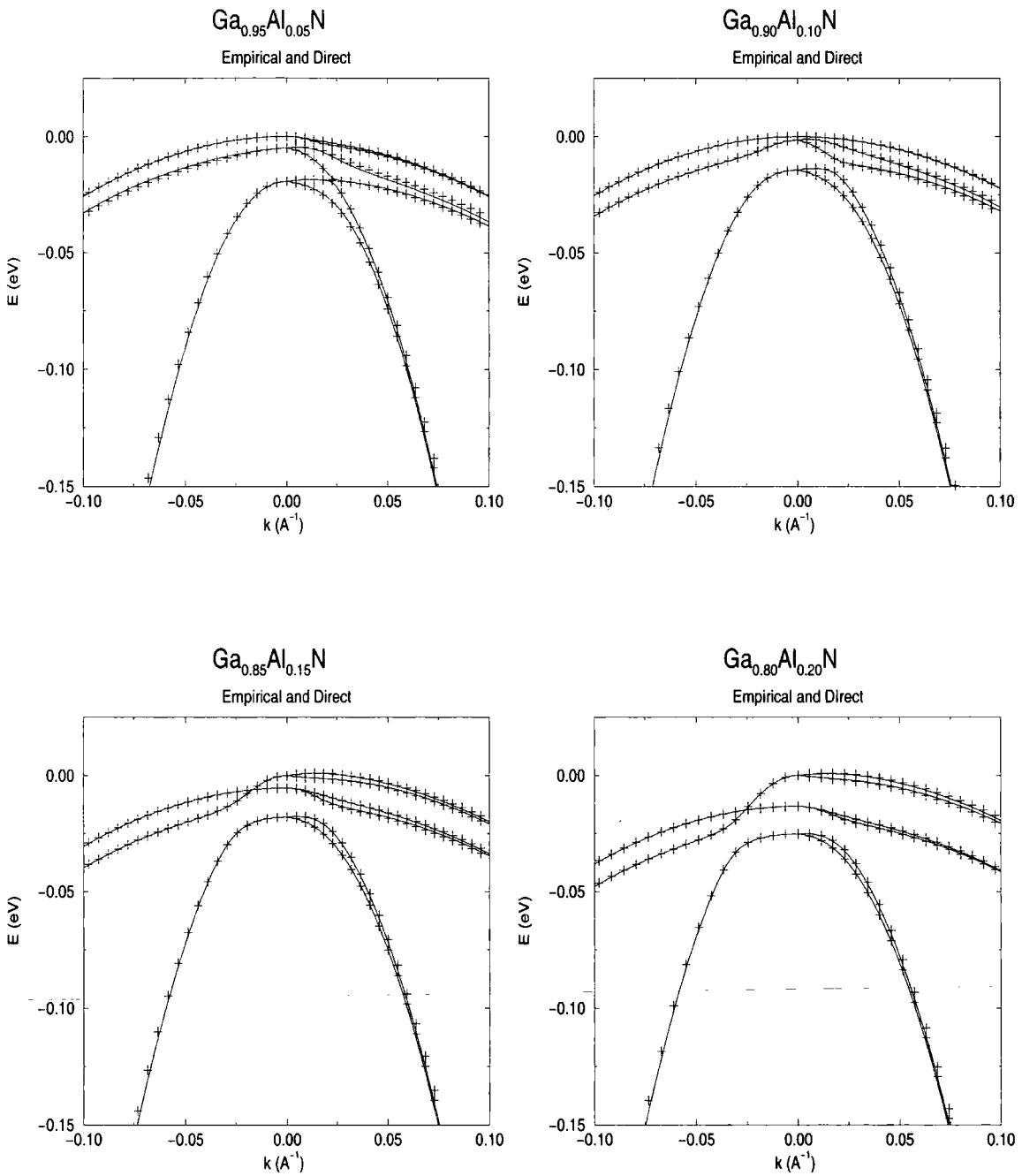

Figure 4.12: Band structure close to the zone centre (Γ-point) for the wurtzite alloy $Ga_{1-x}Al_xN$. The pluses represent the empirical band structure and the solid lines represent the **k.p** band structure calculated using directly obtained **k.p** parameters. Negative k indicates a c-axis ($\Gamma - A$) direction and positive k indicates an in-plane ($\Gamma - M$) direction.

	$Ga_{1-x}Al_xN$						
Composition	A_1	A_2	A_3	A_4	A_5	A_6	A_6
0.05	-7.68	-0.57	7.00	-3.15	-3.12	-4.10	0.17
0.10	-7.41	-0.57	6.74	-3.03	-3.00	-3.93	0.16
0.15	-7.17	-0.56	6.50	-2.92	-2.89	-3.78	0.16
0.20	-6.94	-0.55	6.29	-2.83	-2.80	-3.65	0.15

Table 4.6: **k.p** parameters for the wurtzite alloy $Ga_{1-x}Al_xN$. The A_i are in units of $\hbar^2/2m_0$, except A_7 where the units are eVÅ.

light hole band. Furthermore, they assign this gain to localised exciton states. Both these topics are discussed in more detail for GaInN.

4.5.3.2 GaInN

The variation of the band gap with alloy composition for $Ga_{1-x}In_xN$ is shown in Figure 4.13. From this, it can be seen that the band gap dependence on alloy composition exhibits significant bowing, and is very similar to that of experiment [116,118]. Note that there is again a slight offset in the value of E_g at $x = 0$ due to the original GaN fit, but that the rates of variation of E_g are comparable. The difference in the variation of the band gap between the strained and unstrained cases is much larger than for GaAlN. However, the lattice mismatch of GaN and InN is large compared to GaN and AlN, and thus this should be expected.

The band gap bowing of GaInN is accounted for in [119] by the large lattice mismatch between GaN and InN. However, $In_xGa_{1-x}As$ also has a large lattice mismatch and has been shown to exhibit bowing which is much smaller and relatively composition independent compared to $In_xGa_{1-x}N$ [120]. More particularly, recent theoretical work by Bellaiche *et al* [121] attributes the large band gap bowing instead to In-localized hole states in the upper valence band. The virtual crystal approximation adopted in this work cannot include this effect, but nevertheless appears to give a good description of the variation of the alloy band gap.

Band structure diagrams close to the zone centre for GaInN for the compositions

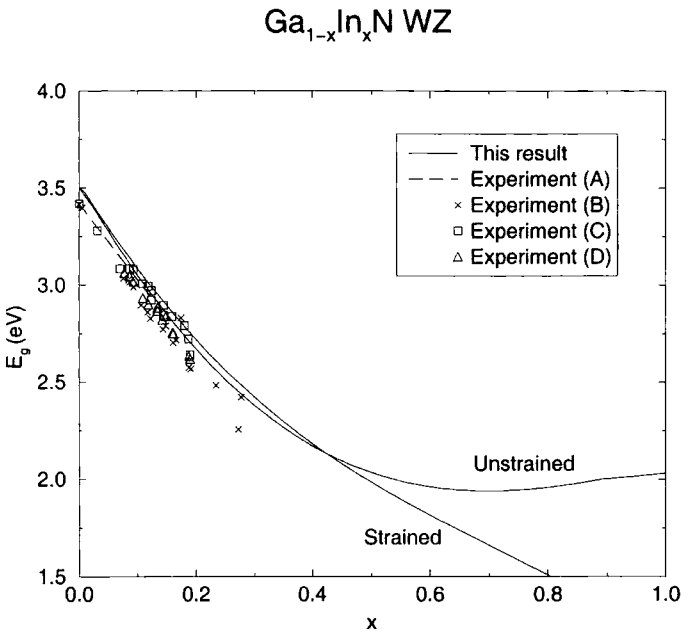

Figure 4.13: The band-gap E_g as a function of alloy composition x in the wurtzite alloy $Ga_{1-x}In_xN$. Shown is a comparison to a linear fit to experimental data (A) in [119], experimental data points (B) compiled in ref. [116] together with PR data (C) and PL data (D) from ref. [118].

$x = 0.05$, 0.10, 0.15 and 0.20 are shown in Figure 4.14. This figure also shows **k.p** band structures calculated using the direct method, and as for the bulk materials there is excellent agreement between the two band structures. The **k.p** parameters calculated using the direct method are shown in Table 4.7. Again, the **k.p** parameters vary smoothly with composition, and this feature of the direct method is discussed in more detail below.

From Figure 4.14, it can be seen that with increasing In composition the crystal field splitting decreases rapidly and becomes negative, resulting in a flip-over in the normal ordering of the bands. To try to understand this phenomenon, the behaviour of both GaN and InN under strain is investigated in chapter 5. The flip-over also deserves further comment with reference to information from experiment. Laser diodes based on GaInN quantum wells (with In composition typically up to 20%) exhibit TE mode lasing, which is usually associated with the heavy hole band. In turn, TM mode

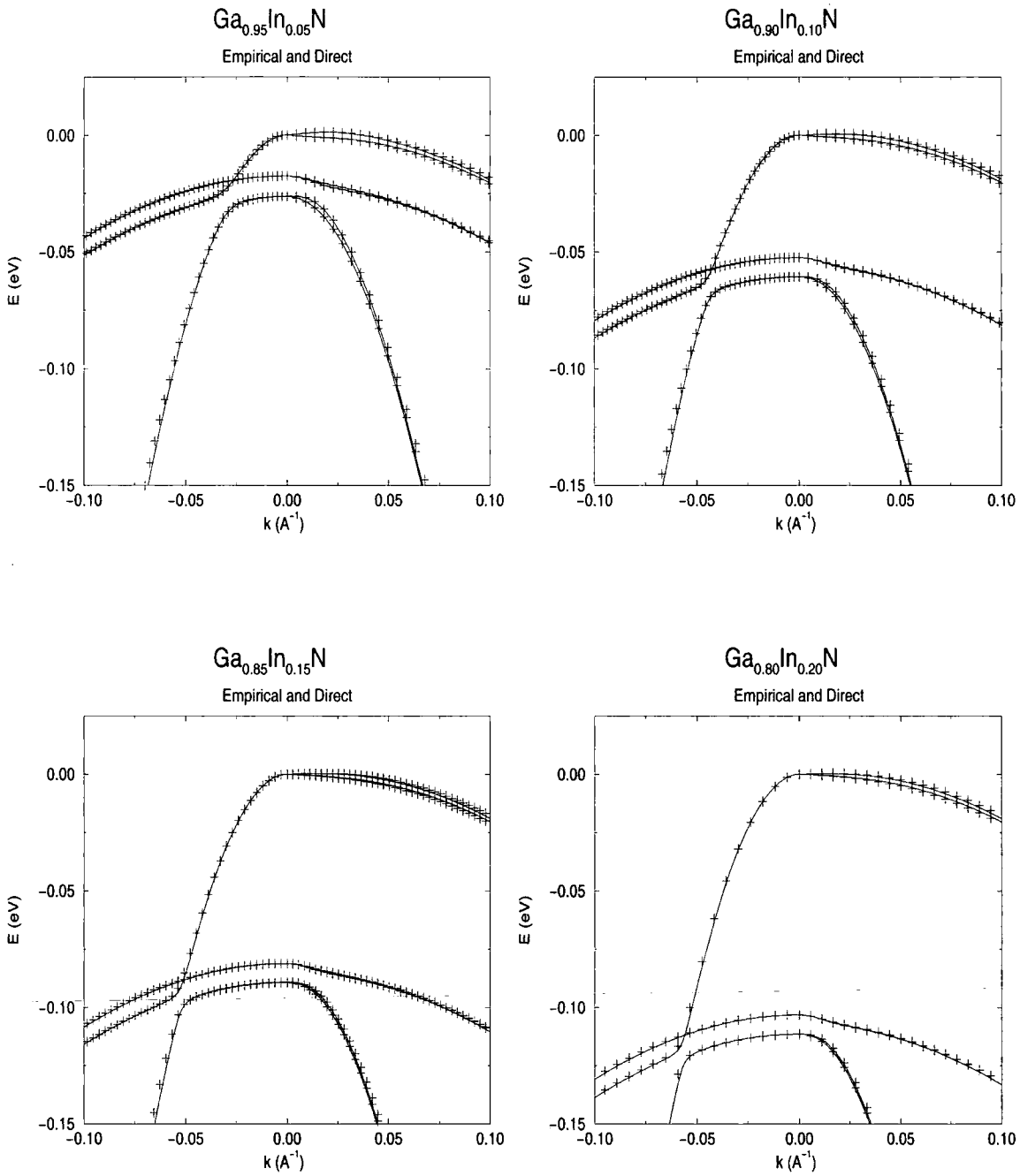

Figure 4.14: Band structure close to the zone centre (Γ-point) for the wurtzite alloy $Ga_{1-x}In_xN$. The pluses represent the empirical band structure and the solid lines represent the **k.p** band structure calculated using directly obtained **k.p** parameters. Negative k indicates a c-axis ($\Gamma - A$) direction and positive k indicates an in-plane ($\Gamma - M$) direction.

	$Ga_{1-x}In_xN$						
Composition	A_1	A_2	A_3	A_4	A_5	A_6	A_6
0.05	-8.36	-0.59	7.66	-3.46	-3.41	-4.54	0.21
0.10	-8.80	-0.60	8.09	-3.67	-3.61	-4.84	0.24
0.15	-9.25	-0.61	8.53	-3.87	-3.81	-5.15	0.27
0.20	-9.69	-0.62	8.96	-4.08	-4.02	-5.45	0.29

Table 4.7: **k.p** parameters for the wurtzite alloy $Ga_{1-x}In_xN$. The A_i are in units of $\hbar^2/2m_0$, except A_7 where the units are eVÅ.

lasing is associated with the light hole band. Experimentally however, TE mode gain is reported to be much greater than the TM mode gain; Frankowsky *et al.* [122] measured the optical gain spectra for $Ga_{0.9}In_{0.1}N$ and found a strong polarization dependence, with the TE mode gain greater than the TM mode gain. Since the heavy hole band couples almost exclusively to TE polarized light (with the electric field vector in the plane of the layer), this would imply that the heavy hole band is above the light hole band, in conflict with the results of this work.

Reasons for this disagreement probably lie in the limitations of the model used for the alloy in this work. In particular, it is most likely due to the failure of the virtual crystal approximation to describe the local atomic environment. Also, Chichibu *et al.* [15] showed experimentally that the concentration of In in GaInN alloys is not uniformly distributed. They measured a weak temperature dependence of the emission spectra in GaInN compared to that of GaN, and attributed the emission from GaInN samples to localised excitons. These localised states were explained in terms of In-rich regions due to fluctuations in the alloy composition, and the virtual crystal approximation or any bulk band structure calculation obviously cannot model this. Also, because there is a large degree of strain in the system, there is probably some error in using Vegard's Law to obtain the alloy lattice parameters.

It was stated earlier in this chapter that one of the advantages of using the direct method to calculate **k.p** parameters is that it can be expected to produce consistent sets of parameters for a series of alloy compositions. To investigate this, directly

obtained and fitted **k.p** parameters were calculated for the whole composition range of $Ga_{1-x}In_xN$. Figure 4.15 shows the variation of both the directly obtained and fitted Luttinger-like parameters with composition in GaInN. From this, it can be seen that the directly obtained **k.p** parameters are smoothly varying, and show strong bowing. The fitted **k.p** parameters also exhibit this behaviour, with the possible exception of A_2 and A_4. It should be noted, however, that similar weightings were used for the fitted **k.p** parameters. From identical empirical band structure, other groups may obtain significantly different **k.p** parameters because of different fitting procedures or different weightings. With the direct method, different groups would obtain identical **k.p** parameters.

4.5.3.3 AlInN

The variation of the band gap with composition for $Al_{1-x}In_xN$ is shown in Figure 4.16. As for GaInN, the results of this work indicate that there is significant bowing of the band gap with composition. As for GaInN, the variation of the band gap is different for the strained and unstrained cases, especially at large x . This should be expected given the large lattice mismatch between AlN and InN.

The theoretical prediction of Wright and Nelson [123], which predicts less bowing than found in this work, is also shown in Figure 4.16. However, the theoretical result of reference [123] was obtained by extension of their zincblende calculations to the wurtzite structure, which was done on the basis that the band gaps of the two structures were only slightly different. Experimental results are also shown in Figure 4.16. The results of this work are in good agreement with those of Kim *et al.* [124] and Guo *et al.* [125]. However, the agreement with the early work of Kubota *et al.* [126], who measure a very large band gap bowing, is less good.

Band structure diagrams close to the zone centre for AlInN for the compositions $x = 0.05$, 0.10, 0.15 and 0.20 are shown in Figure 4.17. This figure also shows **k.p** band structures calculated using the direct method, and it can be seen that there is excellent

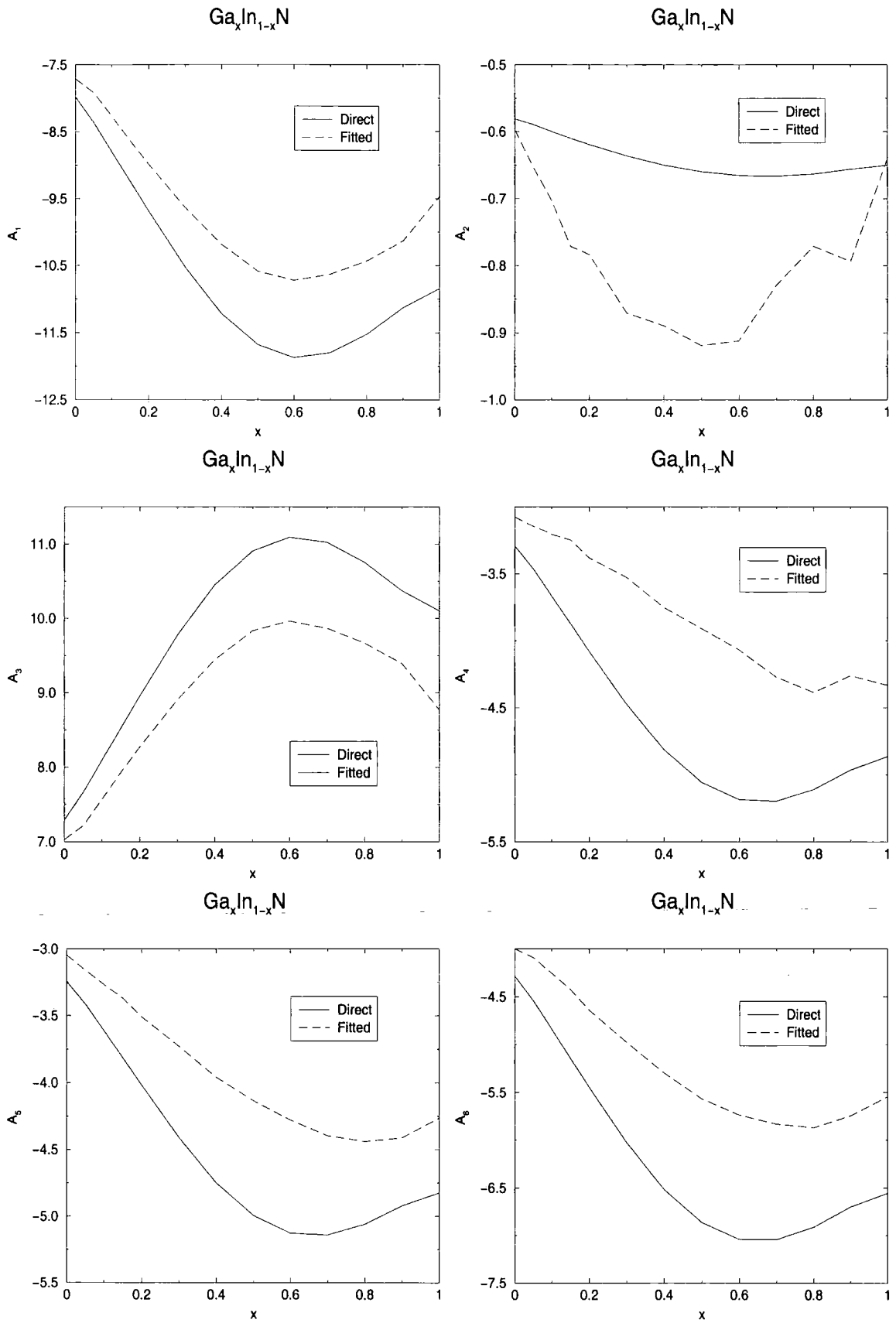

Figure 4.15: The variation of direct and fitted Luttinger-like parameters as a function of alloy composition in wurtzite structure $Ga_{1-x}In_xN$

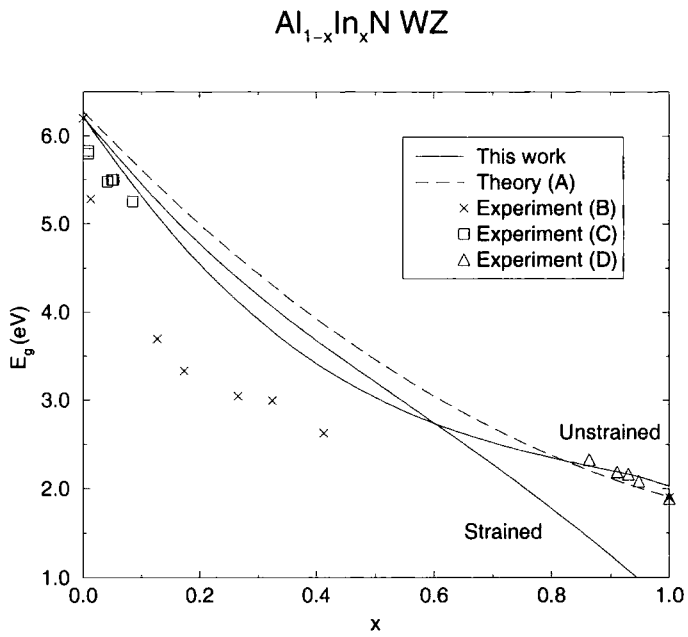

Figure 4.16: The band-gap E_g as a function of alloy composition x in the wurtzite alloy $Al_{1-x}In_xN$. Shown is a comarison to the theoretical results of Wright and Nelson (A) [123] and the experimental results of Kubota *et al.* (B) [126], Kim *et al.* (C) [124] and Guo *et al.* (D) [125].

agreement between the two band structures. The **k.p** parameters calculated using the direct method are shown in Table 4.8. In this table, the Luttinger-like parameters are again seen to vary consistently with composition. From Figure 4.17, it can also be seen that with increasing In composition the crystal field splitting decreases, becoming more negative. Again, this behaviour is discussed with respect to strain on bulk AlN and InN in chapter 5.

4.5.4 Discussion

For the variation of the band gap with composition for the wurtzite alloys GaAlN, GaInN and AlInN, the results of this work appear to be in good agreement with experiment. The band gap of GaAlN was shown to vary linearly with composition, in contrast to GaInN and AlInN where strong bowing was found. This agreement is in spite of the relatively simple model employed in this work.

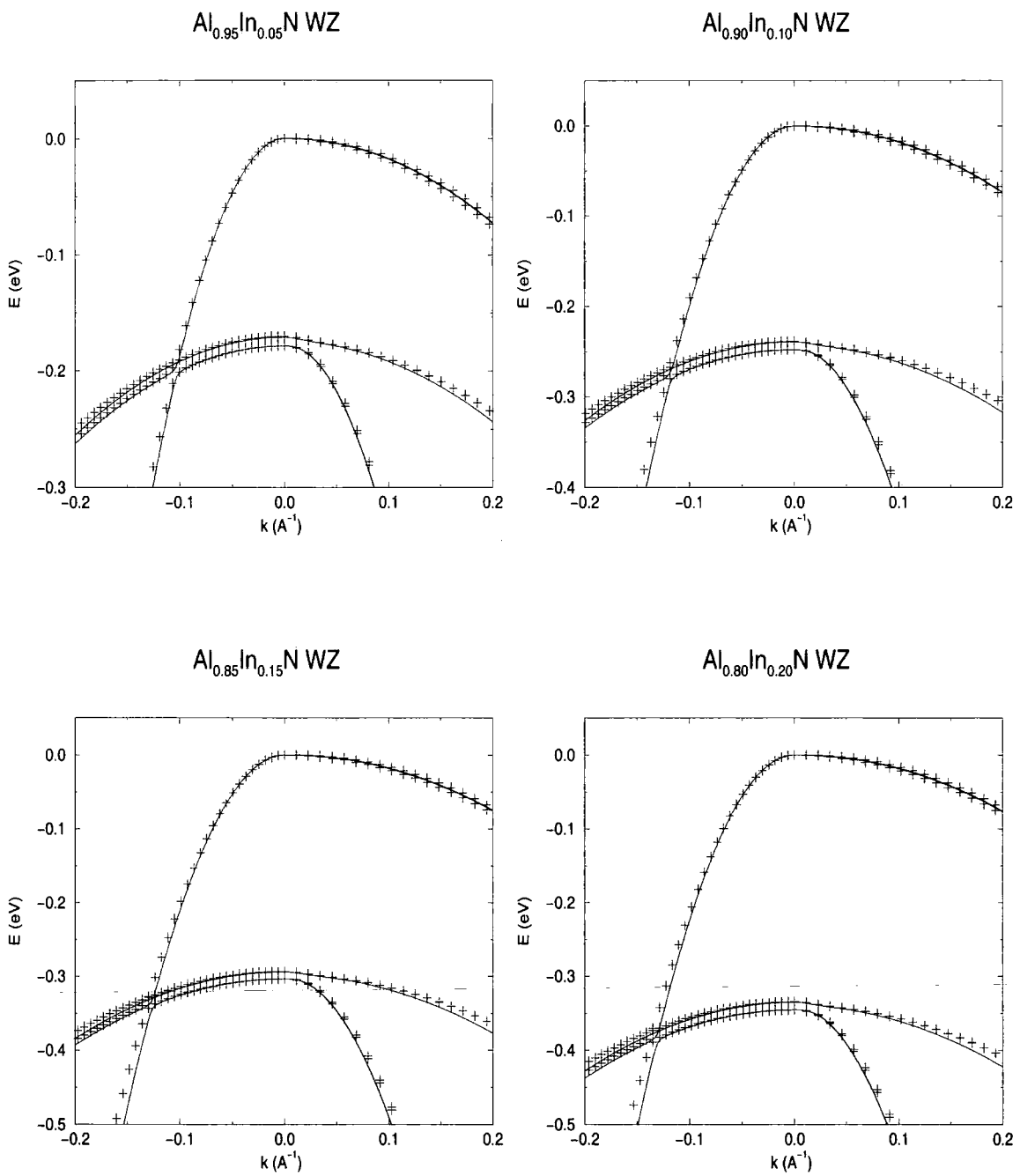

Figure 4.17: Band structure close to the zone centre (Γ-point) for the wurtzite alloy $Al_{1-x}In_xN$. The pluses represent the original band structure and the solid lines represent the **k.p** band structure calculated using directly obtained **k.p** parameters. Negative k indicates a c-axis ($\Gamma - A$) direction and positive k indicates an in-plane ($\Gamma - M$) direction.

	$Al_{1-x}In_xN$						
Composition	A_1	A_2	A_3	A_4	A_5	A_6	A_6
0.05	-4.98	-0.49	4.42	-1.95	-1.99	-2.52	0.14
0.10	-5.27	-0.50	4.69	-2.10	-2.11	-2.70	0.17
0.15	-5.58	-0.51	4.98	-2.24	-2.24	-2.90	0.19
0.20	-5.90	-0.52	5.29	-2.39	-2.38	-3.10	0.22

Table 4.8: **k.p** parameters for the wurtzite alloy $Al_{1-x}In_xN$. The A_i are in units of $\hbar^2/2m_0$, except A_7 where the units are eVÅ.

For the ordering of the uppermost valence bands, however, the agreement with experiment is less good. Optical gain experiments for both GaAlN and GaInN show a strong anisotropy, with the gain for the TE mode much greater than that of the TM mode. In both cases, the origin of the optical gain is attributed to localised excitons. The implication of having TE mode gain is that the heavy hole band is above the light hole band in these alloys, contrary to the situation found in this work.

4.6 Summary

In this chapter, **k.p** parameters have been derived for both the zincblende and wurtzite forms of AlN, GaN and InN. In the zincblende case, effective masses and Luttinger parameters were calculated from the curvature of the bands at the zone centre. The Luttinger parameters were also obtained directly from the wavefunctions and energies at the zone centre. Both methods produced essentially identical **k.p** band structure, in good agreement with the empirical band structure of Chapter 3 near the zone centre. The effective masses of the two methods were also in very good agreement.

In wurtzite, it was not possible to calculate the effective masses from the band structure because the valence bands were non-parabolic. Instead, a Monte-Carlo scheme was adopted, in which the Luttinger-like parameters were systematically varied so as to give **k.p** bandstructure in good agreement with that of the empirical pseudopotential method. As in the zincblende case, the **k.p** parameters were also obtained directly

using the zone centre wavefunctions and energies, and again the two methods produced **k.p** band structure of comparable quality. The effective masses calculated from the Luttinger-like parameters were also in reasonable agreement for the two methods.

In particular, for the wurtzite case the direct method can be considered favourable to alternative methods which employ some kind of fitting approach. This is because it produces a unique set of **k.p** parameters for a given empirical band structure, as opposed to a fitting technique for which an infinite number of near optimum sets exist. The existence of all these possible sets has implications for the effective masses calculated from them, in that two parameter sets that seem to produce equally good band structure will generally produce different effective masses. The arbitrary nature of a fitting technique means that there is no guarantee that the effective masses obtained will exactly correspond to the actual physical values. The direct technique has a sound theoretical basis and has also been shown to produce accurate effective masses for zincblende structure materials. It is thus reasonable to expect that the effective masses obtained for wurtzite structure materials using the direct method are also accurate. Also note that, in principle, a better fitted result could be obtained simply by increasing the number of points used in the fitting procedure, with an associated increase in computational effort. However, since the direct method requires only one calculation and is faster, it should be regarded as preferable.

Lastly, nitride alloys were investigated. Despite the simple nature of the virtual crystal approximation, good agreement for the variation of the band gap with alloy composition was found between the results of this work and those in the literature. For GaInN, there was found to be a flip-over in the ordering of the uppermost valence bands which is not seen in existing experimental measurements. As for bulk materials, **k.p** calculations were also performed for a series of alloy compositions using the direct method. It was found that this method has the particular advantage of producing consistent sets of parameters. Different groups using different fitting schemes would obtain different **k.p** parameters, but the direct method produces just one parameter

set.

Chapter 5

Strain

5.1 Introduction

Consideration of strain is important in any study of the nitrides. This is because nitride epilayers are usually strained, since there is typically a large mismatch in the lattice constants between the nitride materials and the substrates on which they are grown. Thus it is essential to understand how various important properties of the materials vary with strain. From the point of view of device design, such quantities include the band gap and the band structure in general.

Additionally, nitride heterojunctions are an important class of device, being used for such applications as LEDs and MOSFETs. An understanding of strain, and the lineup of the potentials across the interface, is necessary to optimise the performance of such devices. As part of this optimisation, alloys of the nitride materials afford a degree of control that would be unavailable with bulk materials. This is because alloys of the nitrides have lattice constants which differ from the bulk materials, and therefore the band offsets of heterojunctions involving these will be affected. Thus by constructing a heterojunction from nitride alloys of appropriate composition, some degree of control of the band offset, and thus of the transport across the interface, can be achieved. This is indeed the case with devices; AlGaN or GaN are often used as barrier and cladding

layers, and InGaN or GaN as active layers. Thus the offset calculations which are presented later in this chapter complement the alloy calculations that were performed in chapter 4.

In chapter 3 pseudopotentials for AlN, GaN and InN were obtained. These were specifically constructed so as to be easily adaptable for use in strained material, and this property is exploited in this chapter. It is demonstrated how the pseudopotential form factors for the nitride materials under strain are obtained. These descriptions are then used to study deformation potentials, band gaps and the valence band structure in general.

5.2 Theory

A full discussion of several aspects of strain pertinent to this work can be found in, for example, [127]. Presented here is a short summary of the major issues. Consider the two materials A and B shown in Figure 5.2 (a). These two materials have an obvious lattice mismatch, and this is the case to be found in all the nitride materials considered in this work. Due to this lattice mismatch, strain effects will be present in any heterostructures made from these materials. This is obvious if one considers growing an epilayer of material B on a substrate of material A, as is shown in Figure 5.2 (b). The atoms of B are forced to align themselves with those of A with the in-plane lattice constant of B matching that of A, and thus the epilayer is under biaxial compression. As a result of this, the epilayer of B expands along the growth direction such that instead of being described by one lattice constant a_l, B is now described by a_s and a_r, such that $a_r > a_l > a_s$.

For a sufficiently thin epilayer it is reasonable to assume that all the strain is incorporated in the layer. The net strain in the layer plane, $\epsilon_{||}$, is given by

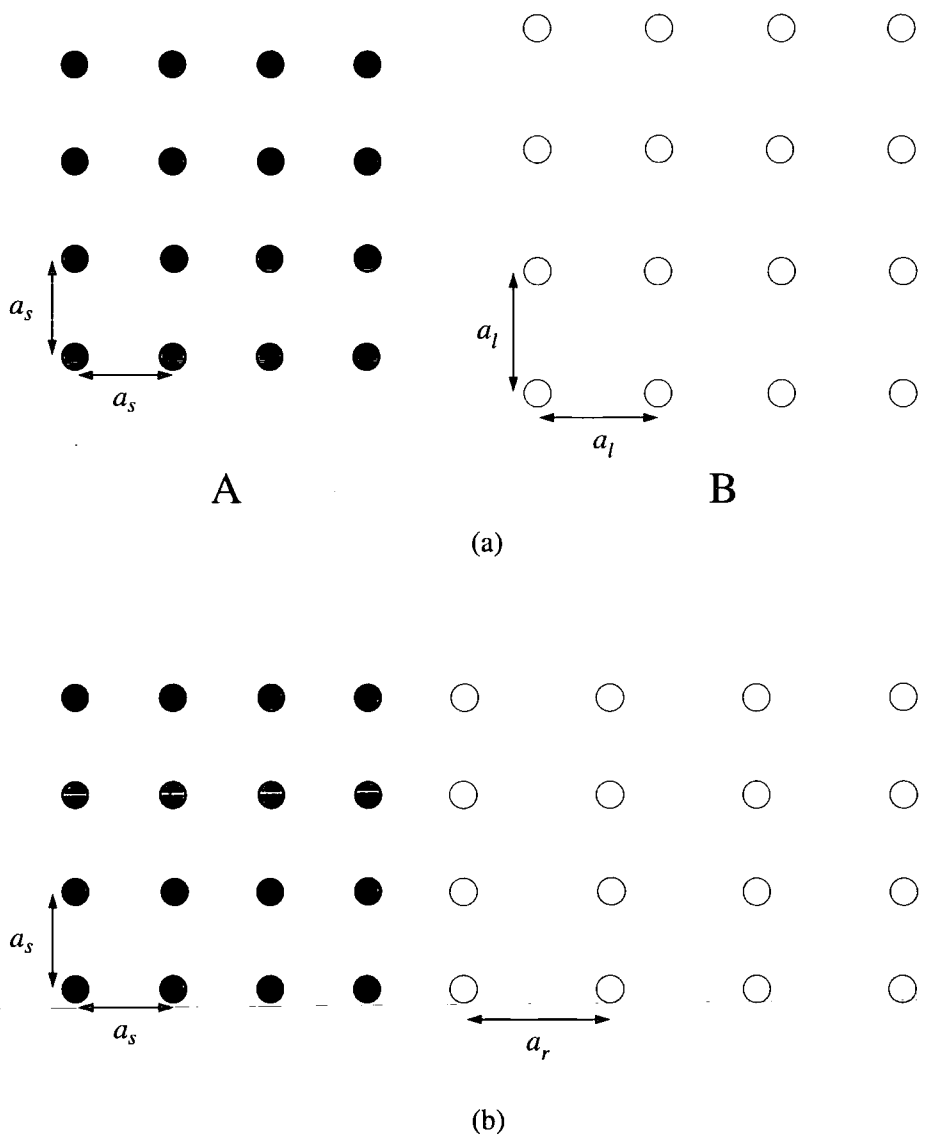

Figure 5.1: Schematic diagram showing exaggerated lattice mismatch between substrate and epilayer, resulting in strained layer growth. In (a), it is clear that the bulk lattice parameters of materials A and B are not equal, such that $a_l > a_s$. In (b) is shown the result of growing an epilayer of material B on a substrate of material A. The result of this is that the epilayer is forced to relax in the growth direction, such that $a_r > a_l > as$.

$$\epsilon_{||} = \epsilon_{xx} = \epsilon_{yy} = (a_s - a_l)/a_l \tag{5.1}$$

The strain in the epilayer along the growth direction, $\epsilon_{\perp}$, is given by

$$\epsilon_{\perp} = -\frac{2\sigma}{1-\sigma}\epsilon_{||} \tag{5.2}$$

where σ is Poissons's ratio. For the tetrahedral semiconductors considered in this work σ is approximately $\frac{1}{3}$, and thus $\epsilon_{||} \sim \epsilon_{\perp}$. It is then useful to resolve the total strain into an axial component ϵ_{ax} and a hydrostatic component ϵ_{vol}:

$$\begin{aligned} \epsilon_{ax} &= \epsilon\perp - \epsilon_{||} = -2\epsilon_{||} \\ \epsilon_{vol} &= \epsilon_{xx} + \epsilon_{yy} + \epsilon_{zz} \sim \epsilon_{||} \end{aligned} \tag{5.3}$$

Hydrostatic strain is such that the crystal symmetry is not affected; the only change is in the lattice parameters of the material. For biaxial strain, such as in the lattice mismatch case considered above, the symmetry of the crystal will change.

5.2.1 Critical Layer Thickness

Consideration of the critical layer thickness is important in any study of strained epilayers. The discussion of this issue here will be brief; for a fuller discussion with particular emphasis on the nitride materials see [6].

When an epilayer is grown on a substrate with some degree of lattice mismatch there exists a critical thickness h_c; below this thickness the epilayer is pseudomorphically strained with an in-plane lattice constant equal to that of the substrate, but beyond h_c it becomes energetically more favourable for the strain to be relieved via dislocations.

The critical layer thickness depends on such factors as the in-plane strain and material properties. Two main approaches have been proposed to model this dependence; one based on energy minimisation and one known as the force balance approach. An extensive discussion of both of these models can be found in [128]. From the force balance model, the critical layer thickness can be calculated to be

$$h_c = \frac{1-\sigma/4}{4\pi(1+\sigma)} b\epsilon_{\|}^{-1} \left(ln(\rho_c h_c/q)\right) \tag{5.4}$$

where σ is Poisson's ratio and b is the dislocation Burgers vector. In general, critical layer thicknesses calculated using this equation are smaller than those measured in experiment.

Experimental measurements of h_c for AlN grown on GaN obtain values varying from $\sim$ 4 Å [129] to 30 Å [130]. Estimates of the critical thickness also vary, with reported values lying between 10 Å [131] and 30 Å [132]. Additionally, Grandjean *et al.* [129] find no difference in the critical thickness for AlN grown on GaN and GaN grown on AlN. Akasaki and Amano [133] have studied $Al_xGa_{1-x}N$ and $In_xGa_{1-x}N$ grown on GaN (with x in the range $0.05 \leq x \leq 0.2$). They measure h_c to be 300 – 700nm for $Al_xGa_{1-x}N$/GaN and $\sim$ 400nm for $In_xGa_{1-x}N$/GaN, and observe no significant change in h_c with composition x. For GaN/InN and AlN/InN, the critical thickness is estimated to be 6 Å [134]. The implications of this for calculations of band offsets are discussed later in this chapter.

5.2.2 Strain in the Empirical Pseudopotential Method

Strain enters the empirical pseudopotential method primarily through the pseudopotential term V_{ps}. It is this that holds information about the crystal structure, and thus any changes induced by applying strain to a material will be strongly reflected in the pseudopotential form factors that describe this term. The kinetic and spin-orbit terms

also depend upon the lattice parameters, and so strain will also have an effect on these terms, though to a lesser degree.

5.2.2.1 Calculation of Form Factors for Strained Material

In chapter 3 symmetric and antisymmetric form factor curves were obtained for AlN, GaN and InN in the wurtzite structure, and these curves can be extended for studies of strained material. This approach is justified when the origin of the form factor curves is considered. In chapter 3, the form factors describing the material were explicitly formulated from the *atomic* pseudopotentials of the anion and cation. Hence the form factor curves can be applied to strained material in an appropriate way described below.

From chapter 3, the form factor at a particular G-value is written as:

$$V(\mathbf{G}) = \frac{1}{\Omega}\int_{\Omega} V(\mathbf{r})e^{-i\mathbf{G}.\mathbf{r}}d\mathbf{r} \tag{5.5}$$

From this equation it is clear that the pseudopotential form factors depend upon the volume Ω as well as the individual G-values. In strained material both these quantities will change, resulting in different form factors. Thus to obtain these form factors this equation is simply evaluated at the new volume and G-values of the strained material.

As an alternative to this method, strain can also be applied in the EPM by assuming an approximately linear variation in the form factors in the region about each reciprocal lattice vector G [135]. However, for large changes this is perhaps less appropriate.

5.3 Results

Using the empirical pseudopotentials obtained in Chapter 3 for AlN, GaN and InN, various properties of the nitrides under strain have been studied. In this work, particular attention has been focused on the behaviour of the uppermost valence bands and the conduction band close to the Brillouin zone centre.

5.3.1 Band Gap Deformation Potential

The energy shift under hydrostatic strain for the conduction band is characterised by the band gap deformation potential a_g. This is defined as

$$a_g = \frac{dE_g}{d\log\Omega} \tag{5.6}$$

where E_g is the band gap, and Ω the unit cell volume.

In this work, band gap deformation potentials have been calculated using the empirical pseudopotential method for the zincblende and wurtzite structures of AlN, GaN and InN. For the wurtzite materials, calculation of the deformation potential was incorporated into the actual pseudopotential fitting procedure in order to ensure values that were consistent with those obtained for the zincblende structure. Deformation potentials were also calculated within the group using the first principles code VASP. The results obtained using these two theoretical approaches, together with those from the literature, are shown in Table 5.1.

From Table 5.1 it can be seen that for the zincblende materials the empirical approach tends to produce larger values than the first principles calculations. This is repeated for the wurtzite materials, where the empirical values are again larger than the first principles results. For both zincblende and wurtzite structures there is a definite trend; the deformation potential of AlN is greater than that of GaN which is greater than that of InN. Compared with other results in the literature, the empirical values in this work tend to overestimate the deformation potential for all three materials. The implication of this is that the empirical pseudopotentials used in this work overestimates the effect of strain.

Material		First principles	Empirical	Ref. [106]	Ref. [136]	Ref. [46]
ZB	AlN	−9.0	−14.6	−9.8		−9.0
	GaN	−7.0	−11.2	−7.7	−6.4	−7.4
	InN		−7.7		−3.0	−2.2
WZ	AlN	−9.0	−12.4		−9.0	−8.8
	GaN	−6.8	−11.0		−6.9	−7.8
	InN		−9.1		−2.8	−4.1

Table 5.1: The band gap deformation potential (eV) for zincblende and wurtzite structure AlN, GaN and InN obtained in this work using both first principles and empirical approaches. Other empirical (Ref. [106]) and first principles (Ref. [46,136]) results in the literature are shown for comparison.

5.3.2 Variation of the Valence Bands with Strain

The variation of the uppermost valence bands with strain is of obvious importance in semiconductor devices, and this has been studied for AlN, GaN and InN in the wurtzite structure. Figures 5.2, 5.3 and 5.4 show the band structures of the heavy hole, light hole and spin split-off bands close to the centre of the Brillouin zone for AlN, GaN and InN under hydrostatic strain using the empirical pseudopotential method.

From Figure 5.2, it can be seen that for AlN the crystal field splitting increases for both positive and negative strain. For GaN shown in 5.3, the crystal field splitting also increases for negative strain. However, for positive strain the crystal field splitting decreases and becomes negative, such that there is a flip-over in the normal ordering of the bands at about 2%. This behaviour is repeated for InN, as shown in Figure 5.4, where there is a flip-over at just over 6%.

First principles calculations using CASTEP have also been used to study the variation of the uppermost valence bands with strain. These calculations were performed starting at the theoretical lattice parameters obtained in chapter 2. For AlN, the crystal field splitting was found to increase for negative strain but decrease for positive strain. For GaN, the calculations confirm the trend found in the empirical calculations. For positive strain the crystal field splitting decreases, and there is a flip-over in the normal ordering of bands. However, this flip-over occurs at a hydrostatic strain of just over

20%, much larger than in the empirical case. However, it is important to note that the first principles crystal field splittings calculated for GaN using CASTEP is much larger than that obtained by using the empirical pseudopotential method. From chapter 2, a crystal field splitting of 66 meV is obtained for GaN using CASTEP, which is much larger than the 22 meV adopted in the empirical model. Therefore, it is reasonable for the flip-over to occur at a higher strain. For negative strain, the first principles calculations find that the crystal field splitting increases. For InN, the crystal field splitting again decreases with positive strain, resulting in a flip-over at approximately 30%, and increases with negative strain. However, in the first principles calculations the atoms within the unit cell were allowed to relax, resulting in changes to the u parameter. In contrast, the u parameter does not change in the empirical calculations. Therefore, since the band structure of the emprical pseudopotential method is dependent on the value of u, a direct comparison between results for the two types of calculations is not possible.

The variation of the valence bands with strain should also be discussed with reference to the results for alloys in chapter 4. For GaAlN, there was found to be a flip-over in the uppermost valence bands with increasing AlN composition, and this can be identified as being due to two competing effects. Results for AlN under positive strain presented earlier in this chapter showed that the crystal field splitting became more negative with increasing strain. In contrast, the crystal field splitting of GaN became more positive on application of negative strain. Thus for GaAlN, the flip-over in the ordering of the bands occurs slowly because of these two effects. For GaInN, a flip-over occurred at very low InN compositions. For GaN, positive strain resulted in a decrease in the crystal field splitting. By contrast, the crystal field splitting of InN increased for negative strain up to 4% but decreased thereafter. Thus from this consideration the behaviour of the uppermost valence bands for GaInN is not unexpected. AlInN also showed a flip-over in the valence bands at small InN compositions, and this can again be explained; the crystal field splitting of AlN decreases (becoming more negative) with

positive strain, while the crystal field splitting of InN decreases under large negative strains.

5.4 Valence Band Offsets

The nitrides have many potential uses in heterojunction devices. In these, transport can either be across the interface (such as in lasers, photodiodes, LEDs) or along the interface (such as in MOSFETs, MESFETs etc.) These devices rely on band engineering to provide the required degree of optical and/or carrier confinement. The valence and conduction band offsets of the heterojunction play an important role in determining the level of confinement, and thus for this type of engineering some knowledge of the band offsets is essential. A detailed discussion of the various aspects involved in band offset engineering can be found in [137].

First principles calculations present one way in which knowledge of the valence and conduction band offsets can be obtained. The first step in these calculations is to build the supercell describing the heterojunction. One of these supercells for an InN-GaN heterojunction is shown in Figure 5.5. (Throughout the rest of this chapter the notation A-B is adopted, whereby A-B denotes a heterojunction where an overlayer of semiconductor A is grown on substrate B. In the context of the calculations presented in this work, this means that the in-plane lattice parameter of an A-B heterojunction is held at that of material B).

The supercell in Figure 5.5 contains 6 wurtzite unit cells along the c-axis. On one side of the interface there are 3 cells of GaN and on the other side there are 3 cells of InN. The size of the supercell required in these calculations is determined primarily by one factor. This is that the heterojunction supercell must be thick enough such that charge densities and potentials are approximately bulk-like in the central regions on

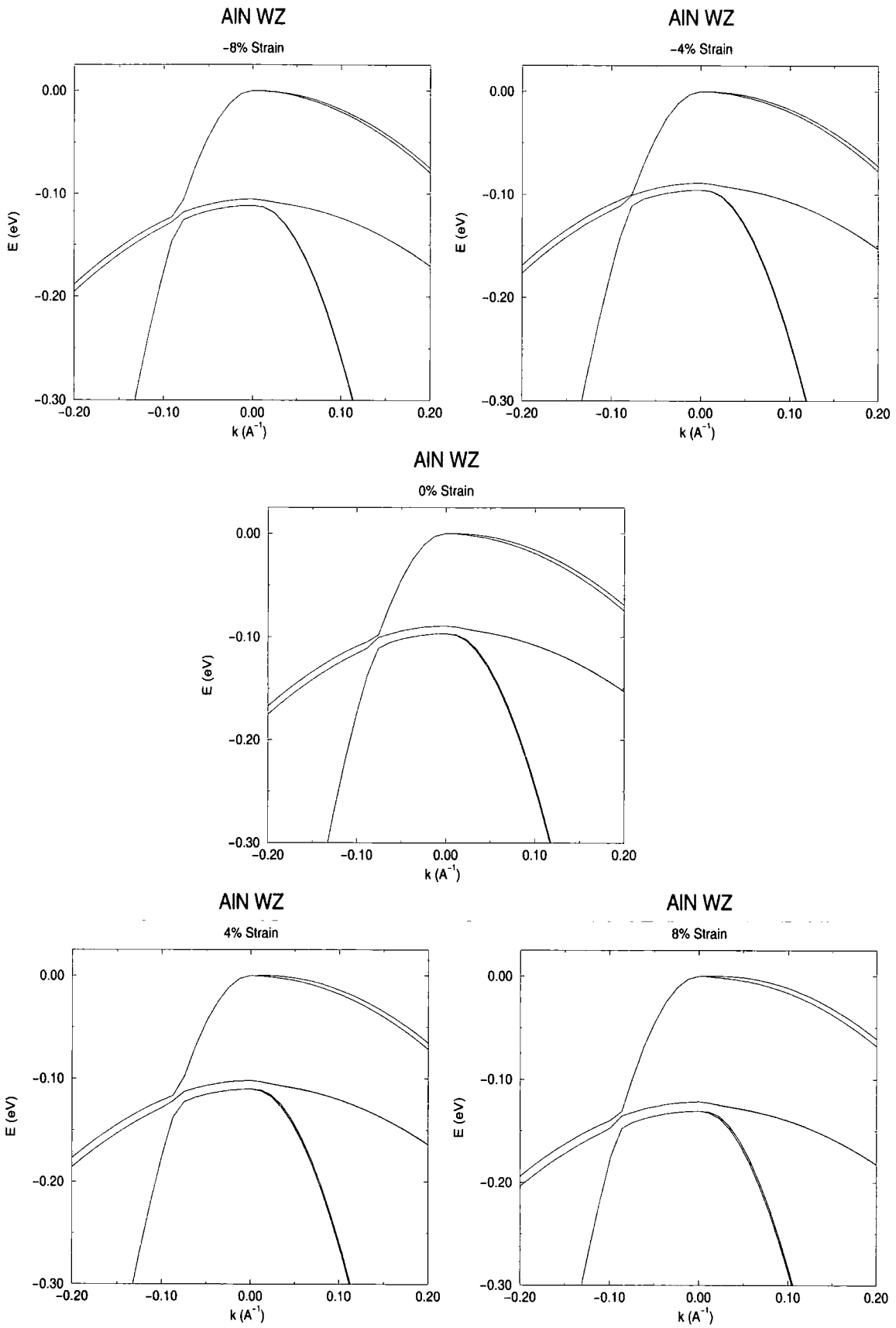

Figure 5.2: Band structure close to the zone centre (Γ-point) for wurtzite structure AlN under hydrostatic strain. Negative k indicates a c-axis ($\Gamma - A$) direction and positive k indicates an in-plane ($\Gamma - M$) direction.

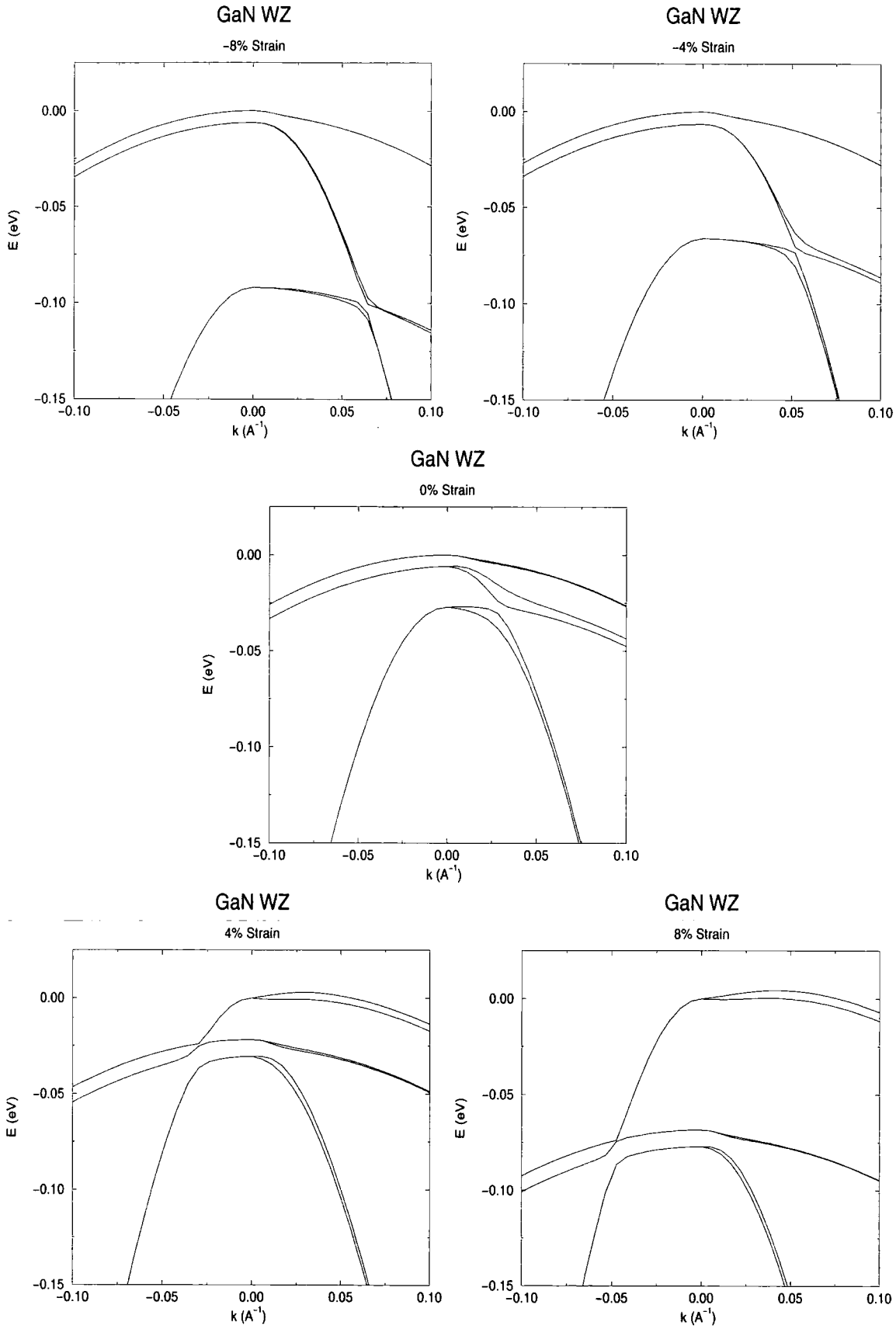

Figure 5.3: Band structure close to the zone centre (Γ-point) for wurtzite structure GaN under hydrostatic strain. Negative k indicates a c-axis ($\Gamma - A$) direction and positive k indicates an in-plane ($\Gamma - M$) direction.

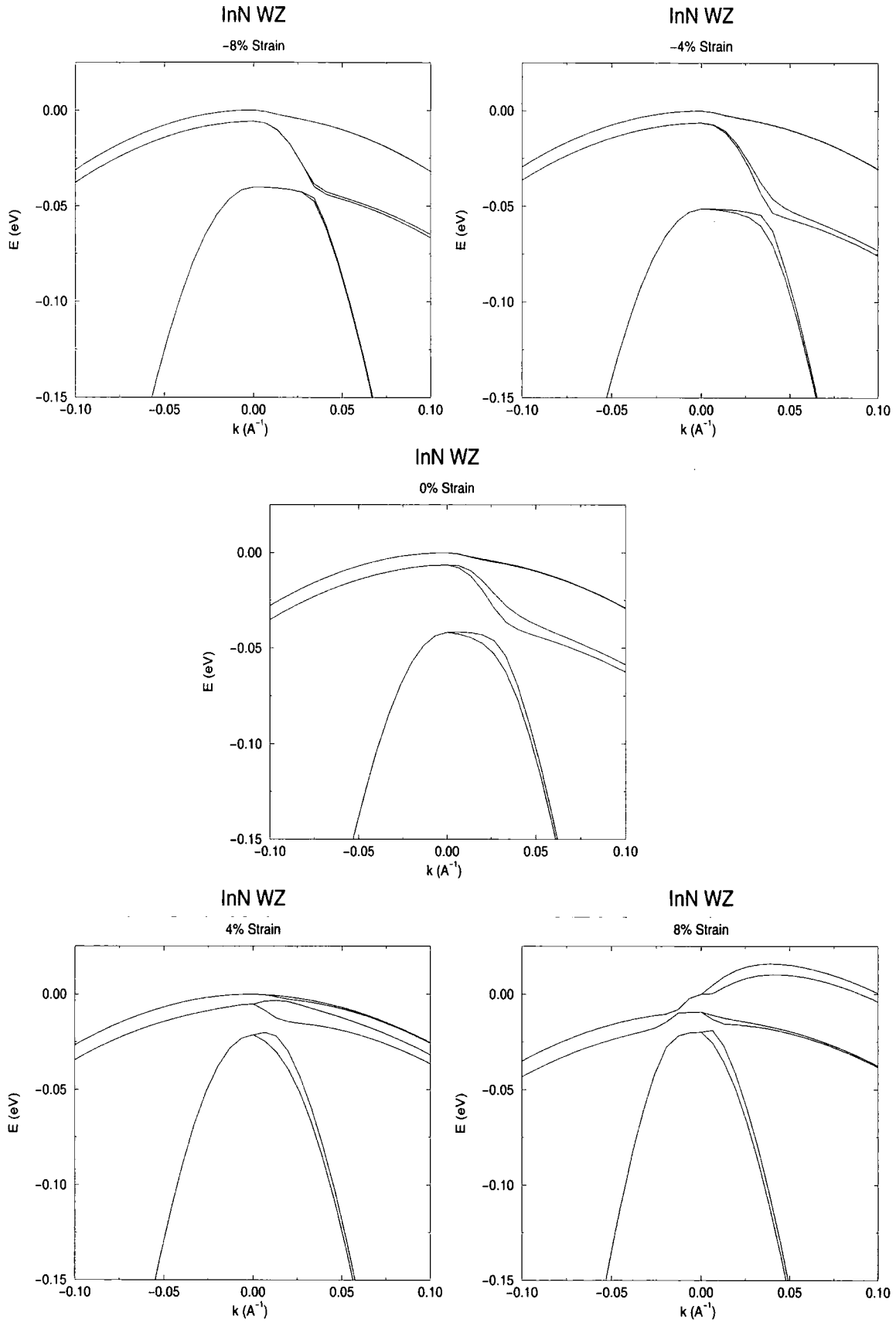

Figure 5.4: Band structure close to the zone centre (Γ-point) for wurtzite structure InN under hydrostatic strain. Negative k indicates a c-axis ($\Gamma - A$) direction and positive k indicates an in-plane ($\Gamma - M$) direction.

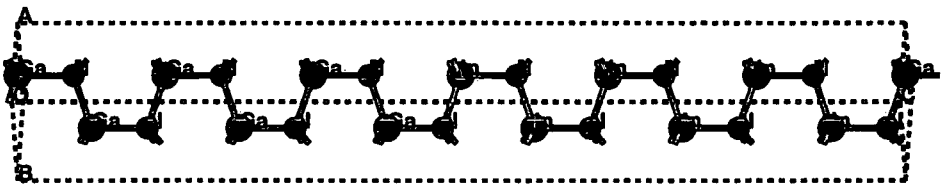

Figure 5.5: The supercell for wurtzite structure InN-GaN

each side of the interface. This indicates that the supercell gives a sufficiently accurate representation of the essentially infinite interface found in a real heterojunction.

In this work, valence band offsets have been calculated for several wurtzite structure nitride heterojunctions. In these calculations, the c lattice parameter of the supercell is allowed to relax while the in-plane lattice constant is held at the relaxed theoretical value of bulk AlN, GaN or InN. In addition, all the atomic positions within the supercell are also allowed to relax.

In order to evaluate the valence band offset of a heterojunction, it is necessary to have a knowledge both of the bulk band structure of the two materials on either side of the heterojunction and of the lineup of the electrostatic potential across the interface. In first principles calculations band energies are determined with respect to the average electrostatic potential in that particular material, which is arbitrary. Therefore, to evaluate the relative positions of the valence band maxima, it is necessary to perform a calculation in which the two average electrostatic potentials can be directly compared, and this is done in a supercell calculation of the interface.

The valence band offset of an A-B heterojunction is written

$$\Delta E_{VBO}(A-B) = (E_v(A) - E_v(B)) + \Delta V(A-B) \tag{5.7}$$

Here, $E_v(A)$ and $E_v(B)$ are the positions of the top of the valence bands with respect to the average electrostatic potential in bulk material of A and B. These terms are evalu-

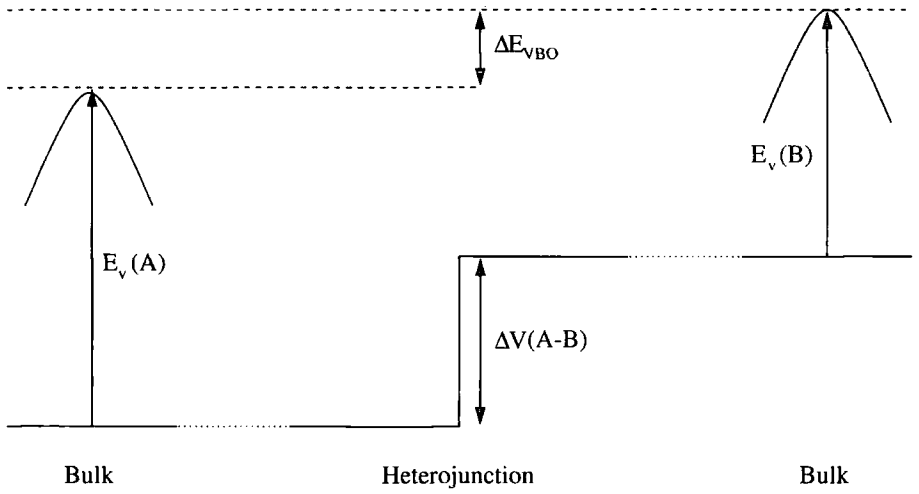

Figure 5.6: Schematic diagram showing how the valence band offset at an A-B heterojunction is evaluated.

ated from separate calculations for the two materials in uniformly strained geometries; that is, in order to evaluate $E_v(A) - E_v(B)$ for an AlN-GaN heterojunction, the bulk calculation for AlN is performed at the relaxed in-plane lattice constant of GaN. The second term $\Delta V(A-B)$ in the equation above is the lineup of the electrostatic potential across the interface, generated by the electronic and ionic charge distributions. This term is evaluated from the supercell calculation using a macroscopic averaging technique described below.

The quantities involved in evaluating the valence band offset are illustrated in Figure 5.6. Here, the lineup of the electrostatic potentials ΔV across the heterojunction is shown in the centre of the diagram. The position of the valence band maxima with respect to these potentials is known from calculations on each of the bulk materials (shown on the left and right of the diagram), and thus the band offset between them ΔE_{VBO} can be determined.

5.4.1 The Macroscopic Average

The electrostatic potential and the electronic charge density are periodic along the c-axis of the supercell due to pseudomorphism. Therefore, the first obvious step is to

compute a microscopic planar average of both the electrostatic potential (comprised of the Hartree, exchange-correlation and ionic terms) and the electronic charge density. These planar averages will be referred to as $\overline{V}(z)$ and $\overline{n}(z)$ respectively.

The planar average of the electrostatic potential can be seen in Figure 5.7 (a). From this, it can be seen that the potentials on either side of the interface are periodic functions that join smoothly across it. Due to lattice mismatch, the period of the potential on each side is not equal.

For lattice matched heterojunctions, where the periods of the potential on either side of the heterojunction are equal, Baldereschi *et al* [138] proposed a macrosopic averaging procedure to subtract bulk effects from $\overline{V}(z)$. This entailed integrating along the supercell over a period λ centred at z:

$$\overline{\overline{V}}(z) = \frac{1}{\lambda}\int_{z-\lambda/2}^{z+\lambda/2} \overline{V}(z')dz' \tag{5.8}$$

However, in systems with a lattice mismatch this approach is not sufficient to subtract the bulk effects on each side of the interface. This is because the potential and the electronic charge density have different periods on either side of the interface. However, the technique can be extended to accommodate this after Colombo *et al.* [139]. This involves performing a double integration over the supercell, once with the period on the left side of the interface and once with the period on the right:

$$\overline{\overline{V}}(z) = \frac{1}{\lambda\lambda'}\int_{z-\lambda/2}^{z+\lambda/2}\int_{z'-\lambda'/2}^{z'+\lambda'/2} \overline{V}(z'')dz''dz' \tag{5.9}$$

Here, λ and λ' are the periods of the left and right sides of the interface respectively. Figure 5.7 shows how the macroscopic averaging technique is applied to the electrostatic potential for an InN-GaN supercell. Figure 5.7 (a) shows the simple planar average potential $\overline{V}(z)$, Figure 5.7 (b) shows the potential after integrating once in equation 5.9

and Figure 5.7 (c) shows the final macroscopically averaged potential $\overline{\overline{V}}(z)$. A similar procedure is used to obtain the macroscopically average electronic charge density.

Note that the macroscopically averaged potential does not exhibit a simple step-like jump at the interface. This is because of the polar nature of nitride heterojunctions in the wurtzite structure that are joined along the c-axis. Also note that there is a strong and uniform electric field in the bulk regions on either side of the interface. This is due to charge transfer across the heterojunction resulting in different charge distributions on either side. This electronic charge transfer is from the In side of the heterojunction to the Ga side.

The macroscopically averaged electronic charge density is shown in Figure 5.8. This shows flat plateaus in the GaN and InN regions between interfaces, demonstrating that the materials have relaxed to become bulk-like in those regions, which of course is required for accuracy in these calculations. Note that the plateau values of the charge density for the bulk-like GaN and InN regions are not equal due to lattice mismatch.

As mentioned previously, the macroscopically averaged electrostatic potential is sloped on either side of the interface instead of being flat because of the strong electric fields. Therefore, there is some degree of ambiguity in determining the offset at the heterojunction, since the offset will depend where on each side of the interface the potential is evaluated. However, from the macroscopically averaged electronic charge density, the centre of the bulk-like regions on either side of the interface are well defined. Therefore, to obtain the band offsets in this work, the potential is evaluated at the two centres of the bulk-like regions. By choosing the offset in this way the error is estimated to be within 0.1 eV.

5.4.2 Details of the calculations

For the geometry optimisations of the supercell, the valence electron wavefunctions were expanded in a plane wave basis set with a cut-off of 500eV. The relaxations were performed using CASTEP, with a conjugate gradients routine chosen to minimise the

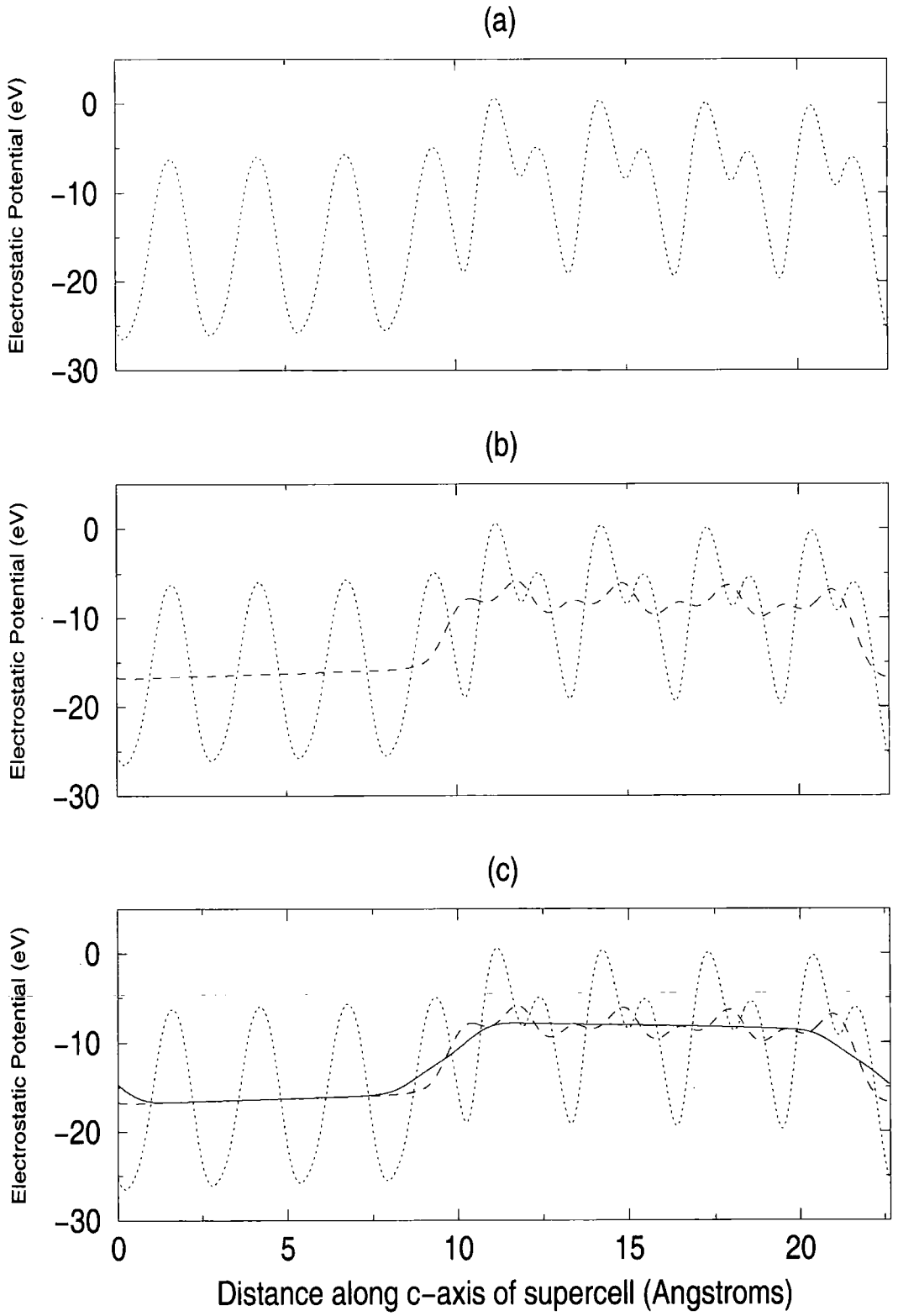

Figure 5.7: The stages involved in calculating the macroscopically averaged electrostatic potential. See text for details.

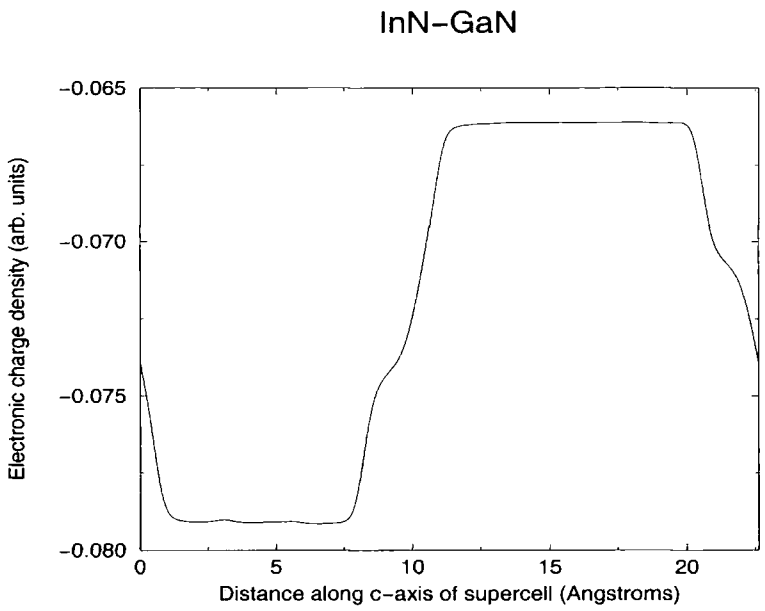

Figure 5.8: The macroscopically averaged electronic charge distribution.

total energy. Integrations over the Brillouin zone were done using 10 symmetrised k-points in the Brillouin zone generated according to the Monkhorst-Pack scheme [38]. In the calculations the electron-ion interactions were described by the ultrasoft pseudopotentials of Vanderbilt [43]. It has been shown that the Ga $3d$ states and the In $4d$ states play an active part in chemical bonding [45,46], and thus for both Ga and In the d electrons were treated as valence in addition to the s and p. The forces on the ions were calculated using the Hellmann-Feynmann theorem, and again minimised using a conjugate gradients technique. A finite basis set correction was employed to account for changes in the supercell geometry as a result of the relaxation.

As discussed previously in this chapter, there is a large lattice mismatch between InN and GaN or AlN. Consequently, heterojunctions between InN and either GaN or AlN will be highly strained systems with significant relaxations in the atomic positions at the interface. There will also be large relaxations in the c lattice parameter for these systems. Therefore, for some of the heterojunction supercells considered in this work a series of single point energy calculations were initially performed to obtain an estimate of the relaxed supercell c value. For these, the total energy was calculated for

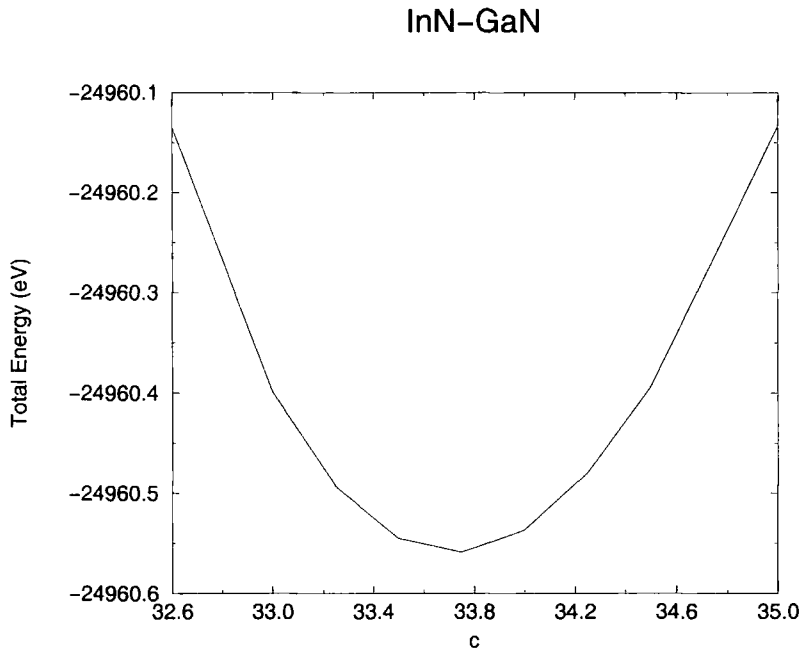

Figure 5.9: The total energy of the InN-GaN supercell for various values of the c lattice parameter. The lattice parameter a was held at the relaxed value of bulk GaN.

successive values of c, with a held fixed at the respective in-plane lattice constant of the heterojunction. A graph showing the results of such a calculation for an InN-GaN supercell is shown in Figure 5.4.2.

For InN-GaN and AlN-GaN, geometry optimisations were performed for supercells containing up to 6 wurtzite unit cells (3 on either side of the interface), corresponding to 6 atoms of Ga, 6 atoms of In or Al, and 12 atoms of N. However, it was found that there was little difference in the band offsets calculated using 4 and 6 unit cells respectively, with the discrepancy being less than 0.03 eV. Additionally, for the supercell calculations using 4 unit cells, the macroscopically averaged electronic charge densities exhibit well-defined plateau regions, indicating that 4 unit cells are sufficient for these calculations. This is consistent with previous calculations, which find it is necessary to employ layers 4 to 6 atoms thick on either side of the interface to extract accurate band offsets [140,141].

Additionally, the method used in this work was also used to obtain the valence

band offset for the well-studied AlAs-GaAs heterojunction. A value of 0.44 eV was calculated, which is in good agreement with the theoretical and experimentally observed values of other groups, which obtain values of between 0.40 and 0.55 eV [142–144].

5.4.3 Results

Heterojunction	a_{sub}	c_{sub}	$c_{superlattice}$
AlN-GaN	3.155	5.150	19.976
GaN-AlN	3.064	4.900	20.199
InN-GaN	3.155	5.150	22.611
GaN-InN	3.526	5.712	21.234
AlN-InN	3.526	5.712	19.334
InN-AlN	3.064	4.900	22.476

Table 5.2: Substrate a and c lattice parameters together with relaxed superlattice c values (all in Angstroms) for the heterojunctions considered in this work. 2 unit cells of each material are on either side of the heterojunction.

Table 5.2 shows the "substrate" a and c lattice parameters. Also shown are the relaxed superlattice c values. Note that the substrate lattice parameters used here are from the theoretical calculations of chapter 2. These parameters differ from the experimentally determined values by up to $\simeq 2\%$, and as discussed later this will have an effect on the calculated band offset. The values of $c_{supercell}$ for AlN-GaN and GaN-AlN differ only by approximately 0.20 Å, as would be expected since the lattice mismatch between AlN and GaN is only 3 %. For the heterojunctions involving InN however, the lattice mismatch is much larger, at about 12 % for GaN and 15 % for AlN. This is reflected in the two relaxed $c_{superlattice}$ values for each heterojunction, which differ by up to 3 Å.

5.4.3.1 AlN-GaN Heterojunctions

Table 5.3 shows the valence band offsets calculated in this work with others in the literature. For all of the offsets in this table, the offset is such that the top of the valence band of GaN is above that of AlN. The most significant feature in the results from this

	Theory					
	This work	Ref. [145]	Ref. [146]	Ref. [147]	Ref. [52]	Ref. [148]
AlN-GaN	0.15	0.20				0.44
				0.7 ± 0.1	0.81	
GaN-AlN	0.77	0.85	0.80			0.73
	Expt.					
			Ref. [134]	Ref. [149]		
AlN-GaN			0.57 ± 0.22	1.36 ± 0.07		
GaN-AlN			0.60 ± 0.24			

Table 5.3: Valence band offsets (in eV) for AlN-GaN and GaN-AlN heterojunctions. For all the offsets presented above the top of the valence band of GaN is above that of AlN. Note that for references [148] and [147] the offset is calculated for the zincblende (001) and (110) interfaces respectively.

work is the large forward-backward asymmetry of the offset, which has previously been reported in the theoretical calculations of Bernardini *et al.* [145] and Nardelli *et al.* [148]. The origin of this feature can be traced to either a band edge shift in the bulk band structure (the $E_v(A) - E_v(B)$ term in equation 5.7) or to a change in the lineup of the potentials in the supercell (the $\Delta V(A - B)$ term in equation 5.7). In this work, this asymmetry is almost exclusively due to band edge shifts (0.57 eV) rather than the potential lineup in the actual supercell (0.02 eV). The asymmetry has been explained by Bernardini *et al.* [145] in terms of the different natures of the valence band maxima in AlN and GaN. In AlN, the singlet at the top of the valence band (formed from hybridised N $2s$ orbitals and Al p_z states) is pushed downwards by biaxial compression; in GaN the doublet at the top of the valence band (formed from hybridised N $2s$ and Ga p_{xy} states) is pushed upwards by biaxial compression. Nardelli *et al.* obtained their results from studying a zincblende (001) interface. This result can be compared with ones for wurtzite heterojunctions because of the similarity of the wurtzite and zincblende structures; these primarily differ in their stacking arrangements, and hence their band lineups should be similar. Van de Walle *et al.* [147] have also studied the valence band offset for the nonpolar (110) zincblende interface. They obtain a value

for the valence band offset between *unstrained* materials of 0.7 ± 0.1 eV by subtracting out the contribution due to strain. These authors also report a difference of only 0.1 eV between the band offsets of (110) and (001) interfaces.

From Table 5.3, it can be seen that the results in this work are in excellent agreement with other theoretical calculations, particularly those of Bernardini *et al.* [145] and Majewski *et al.* [146]. For both of these, the offsets calculated in this work are within 0.1 eV of their reported values. The offset obtained by Van de Walle *et al.* , which is corrected for strain, also falls between the two values in this work. For the calculations of Wei and Zunger [52], the in-plane lattice constant of the wurtzite heterojunction is taken to be the average of the two nitride materials on either side, and hence it would be reasonable to expect their offset to lie within the two values presented in this work. Their reported value of 0.81 eV does not satisfy this expectation. However, it should be noted that for GaN-InN heterojunctions these authors report no appreciable change in the offset with substrate lattice constant, and thus any discrepancies are quite likely due to this. Finally, the offsets obtained by Nardelli *et al.* exhibit forward-backward asymmetry and are comparable in magnitude to the ones presented in this work.

From Table 5.3, it can be seen that there is some disagreement between the theoretically calculated band offsets and those measured from experiment. In particular, the results from x-ray photoemission spectroscopy of Martin *et al.* [134] do not show any significant forward/backward asymmetry, although they do obtain a similar magnitude for the offset. Lastly, the anomalously large offset obtained by Waldrop and Grant [149] is not in agreement with any of the other results shown in Table 5.3.

5.4.3.2 GaN-InN Heterojunctions

The offsets calculated in this work are compared with others in the literature in Table 5.4. For all of the results in this table, the offset is such that the top of the valence band of InN is above that of GaN. As for the AlN-GaN and GaN-AlN heterojunctions, the offsets obtained in this work exhibit considerable asymmetry and are of similar

	Theory				Expt.
	This work	Ref. [147]	Ref. [52]	Ref. [148]	Ref. [134]
InN-GaN	0.85				0.93 ± 0.25
		0.3 ± 0.1	0.48	0.70	
GaN-InN	0.02				0.59 ± 0.24

Table 5.4: Valence band offsets (in eV) for InN-GaN and GaN-InN heterojunctions. For all the offsets presented above the top of the valence band of InN is above that of GaN. Note that for references [148] and [147] the offset is calculated for the zincblende (001) and (110) interfaces respectively.

magnitude. The offset of Van de Walle *et al.* , of 0.3 ± 0.1 eV [147], which again is corrected for strain, lies between the two values of this work. Wei and Zunger [52] calculate a value for the offset of 0.48 eV using an in-plane lattice constant equal to the average of the GaN and InN a values. Since this offset lies between 0.02 eV and 0.82 eV, this is in general accord with the results presented in this work. However, these authors report a weak dependence of the band offset on the lattice constant used in the calculation. They find that using the substrate lattice constant for GaN or InN respectively instead of their average only accounts for changes in the band offset of the order of 0.1 eV. Nardelli *et al.* [148] have also evaluated the valence band offset of the GaN-InN based heterojunction, and obtain a value of 0.70 eV. However, this value was calculated for a heterojunction with the in-plane lattice parameter fixed at the lattice constant of AlN. As such, it should be in general agreement with that obtained here for InN-GaN, and this is true to within 0.15 eV.

Experimental investigation of the band offset in InN-GaN or GaN-InN is hindered in that achieving a pseudomorphic interface between the two compounds is problematic. Indeed, the critical layer thickness is estimated to be 6 Å [134]. Therefore, real heterojunctions are likely to be comprised of a messy interface region with fully relaxed GaN and InN on either side. However, reported values do exist for offsets in these heterojunctions. An experimental investigation from Martin *et al.* [134] finds valence band offsets for the InN-GaN and Gan-InN cases of 0.93 eV and 0.59 eV respectively. These

are in qualitative agreement with the offsets calculated in this work since they also exhibit forward-backward asymmetry, though to a lesser degree. However, the value they obtain for the offset for GaN-InN of 0.59 $\pm$ 0.24 eV is not in good agreement with the offset calculated in this work of 0.02eV. This discrepancy is most likely due to the state of the interface region mentioned above; strain in their sample is likely to be relieved by dislocations, in contrast to the "perfect" heterojunction modelled in this work.

In contrast to AlN-GaN and GaN-AlN heterojunctions, the forward-backward asymmetry in the valence band offset calculated in this work is largely caused by a change in the lineup of the potential in the supercell. The change in the offset of 0.80 eV is comprised of a band edge shift of -0.59 eV and a large change in the supercell potential lineup of 1.39 eV. Martin *et al.* suggest that the forward-backward asymmetry is explained by strain-induced piezoelectric fields. The lattice constants a and c of GaN are about 12 % lower than those of InN, and thus heterojunctions comprised of these two materials will be highly strained. In contrast, heterojunctions involving AlN and GaN, whose lattice constants agree to within $\sim$ 3 %, will not be so strongly strained. The nitrides have negative piezoelectric constants such that increases in the lattice constant result in static electric fields which decrease the band gap. Hence in the nitrides strain acts to decrease the valence band offsets, and thus nitride heterojunctions involving InN will exhibit a degree of forward-backward asymmetry.

The dependence of the asymmetry on strain also has some relevance to the origin of the discrepancy between the results of this work and those of Martin *et al.* . Specifically, if some of the strain in their sample is relieved by dislocations then it is reasonable to expect that the forward-backward asymmetry they observe would be smaller than that calculated in this work, and this is indeed the case.

	Theory			Expt.
	This work	Ref. [52]	Ref. [148]	Ref. [134]
AlN-InN	-0.22			1.32 ± 0.14
		1.25 ± 0.1		
InN-AlN	1.58		1.37	1.71 ± 0.20

Table 5.5: Valence band offsets (in eV) for InN-AlN and AlN-InN heterojunctions. A negative value for the band offset indicates that the valence band of AlN is above that of InN. Note that for references [148] the offset is calculated for the zincblende (001) interface.

5.4.3.3 InN-AlN Heterojunctions

The offsets calculated in this work are compared with others in the literature in Table 5.5. Again there is a strong forward-backward asymmetry in the offsets calculated in this work. This effect is stronger than was found in GaN-InN and InN-GaN, as would be expected given that the lattice mismatch of $\sim$ 15 % is somewhat larger than in those heterojunctions. The resulting strain is such that the valence band of AlN is actually above that of InN for an AlN-InN heterojunction. For InN-AlN, a large valence band offset is found. These results are reasonably consistent with the calculations of Wei and Zunger [52], who obtain a value of 1.25 ± 0.1 eV. This value lies within the two values obtained in this work, as would be expected for a calculation using an average of the AlN and InN in-plane lattice parameters. The result of Nardelli *et al.* [148] is also in reasonable agreement with the values found in this work.

As for GaN-InN and InN-GaN, the extreme lattice mismatch between AlN and InN makes experimental investigation of heterojunctions between them problematic, with a critical layer thickness, like GaN-InN, estimated to be 6 Å(Martin *et al.* [134]). Therefore, as for GaN-InN, it is reasonable to assume that in practice a heterojunction formed from these materials will consist of an interface region containing many defects and dislocations, with fully relaxed AlN and InN away from it on either side. It is thus important to take care when comparing the results in this work with those of experiment, such as those of Martin *et al.* . These authors report forward-backward

asymmetry that is not as pronounced as in this work. In particular, although their result of 1.71 ± 0.20 eV for InN-AlN is in good agreement with that found in this work, there is considerable discrepancy in the value for AlN-InN. Again, this is likely to be due to there being many dislocations at their sample interface instead of the uniformly strained heterojunction modelled in this work. If some of the strain is relieved by dislocations, and the forward-backward asymmetry is partly due to strain induced electric fields, then this does serve as an explanation for why they do not obtain the extreme asymmetry found in this work.

As for GaN-InN and InN-GaN, it is instructive to decompose the forward-backward asymmetry into its band edge and potential lineup components. For this, the asymmetry of 1.8 eV is comprised of a band edge shift of 2.20 eV and a change in the potential lineup of -0.4 eV. This asymmetry can again be explained by large strain-induced piezoelectric fields.

5.4.4 Discussion

The valence band offsets calculated for the nitride heterojunctions considered above are in good general agreement with the results of other groups. In particular, the forward-backward asymmetry in the band offset that is observed in experimental measurements and in other theoretical calculations is reproduced in this work. This dependence on the substrate lattice constant has obvious implications for device design which are discussed in more detail below.

In the results for the band offsets there are several features which deserve further comment. Firstly, the positions of the AlN, GaN and InN valence bands with respect to each other is worth discussing. For the AlN-GaN and GaN-AlN heterojunctions, GaN is above AlN. For the InN-GaN and GaN-InN heterojunctions, InN is above GaN. For InN-AlN, as expected InN is above AlN, but for the AlN-InN heterojunction AlN is actually higher than InN. A qualitative explanation for this behaviour can be found by examining the effect of the d electrons in each of the materials [52], and is

reproduced here. In AlN the unoccupied Al d orbital lies above the N p level, and thus p-d repulsion pushes the valence band maximum down in energy. For both GaN and InN the situation is different. In GaN there is an occupied Ga $3d$ level and in InN there is an occupied $4d$ level, both of which are below the N p energy. Thus for these two materials p-d repulsion acts to push the valence band maximum up in energy. The degree of this shift can be approximated as $\sim V_{pd}^2/(E_p - E_d)$ [52]. Therefore, the closer the anion p and cation d levels are in energy, the greater the shift of the valence band maximum. Since the matrix element V_{pd} increases as the nearest-neighbour bond length decreases, then compared to other III-V systems this shift is also greater due to the shorter anion-cation bond lengths. Thus from this consideration the valence band maximum of InN should be above that of GaN, which itself should be above that of AlN. This is reflected in the results of this work, with the only discrepancy being for AlN-InN, which has the valence band maximum of AlN above that of InN. This is explained by the large piezoelectric fields induced by the high degree of strain in this heterojunction.

For devices based upon nitride heterojunctions there are several practical limitations to consider. Foremost of these is the matter of the critical layer thickness. For GaN-AlN systems, the relatively small lattice mismatch results in a critical layer thickness of approximately 30 Å. In contrast, the critical layer thickness for nitride heterojunctions involving InN are estimated to be 6 Å. Moreover, one of the principal features of nitride epitaxial layers grown by MBE and CVD is a network of threading defects which relieve the strain. Thus for the heterojunctions modelled in this work, in which the layers are pseudomorphically strained, the calculated offsets must be used with caution when comparing them with those measured from real heterojunctions, particularly those involving InN. In real systems, strain will be present to an unknown extent. Martin *et al.* attributed the forward-backward asymmetry of the band offsets to strain-induced piezoelectric fields, and expressed a preference for measurements with the smallest piezoelectric effect. On this basis, they rejected the cases with the smallest

valence band discontinuities and applied rough corrections for the piezoelectric fields. The corrected values they obtained were

GaN-AlN 0.70 ±0.24 eV

InN-GaN 1.05 ±0.25 eV

InN-AlN 1.81 ±0.20 eV

which are in reasonable agreement with the values for those heterojunctions obtained in this work of 0.77 eV, 0.85 eV and 1.58 eV.

Perhaps the most interesting result in the calculations above is the forward-backward asymmetry in the band offsets. This effect is strongest for the AlN-InN/InN-AlN heterojunctions, as should perhaps be expected with the large lattice mismatch present in these systems, but it is also significant for the more closely lattice matched AlN-GaN/GaN-AlN heterojunctions. Therefore, the results in this work would seem to suggest that the forward-backward asymmetry found in these systems is caused by strain-induced electric fields. From the point of view of band engineering, the important result here is the dependence of the band offset on the lattice constant of the "substrate" nitride. This implies that the band offset can be tuned to some degree by using a nitride alloy substrate of appropriate composition on one side of the interface.

5.5 Summary

In this chapter, the behaviour of AlN, GaN and InN under strain was investigated. This study focused on conduction band deformation potentials and the behaviour of the uppermost valence bands. For all the nitrides, the empirical deformation potentials calculated in this work were larger than those found by experiment. For GaN, positive strain was found to result in a negative crystal field splitting and a flip-over in the normal ordering of the bands at the zone centre. This behaviour was repeated for InN.

Valence band offset calculations were also performed for various nitride heterojunctions. The results obtained for these were in line with other theoretical results and with experiment. A strong forward-backward asymmetry was displayed by all the heterojunctions, and particularly those involving InN. An explanation for this effect was found in piezoelectric fields generated as a result of the strain. In heterojunction devices, this strain induced polarization provides an additional degree of freedom by which the carrier concentration can be significantly altered, and thus device behaviour can effectively be engineered.

Chapter 6

Conclusions

In this thesis, work investigating the electronic structure of III-V nitride materials and their alloys has been presented. For this, three different but complementary methods were employed. Being theoretical, these methods avoid any problems due to poor quality of the nitride materials, and thus provide a useful tool for such studies. Moreover, these methods are all well established, having been successfully used in similar investigations on materials such as Si and GaAs.

In Chapter 2, the first principles pseudopotential method was introduced. In addition to a discussion of its physical basis, the important approximations necessary for it were discussed. Using this method, lattice parameters describing the zincblende and wurtzite structures for the nitride materials were obtained.

Bonding in the zincblende and wurtzite structures was also investigated via population analysis. Through a projection of the plane wave basis states onto LCAO states, both atomic populations and overlap populations gave indications of the relative covalent/ionic nature of the bonding. GaN and InN had similar anion and cation valence charges suggesting intermediate covalency, and this conclusion was also backed up by the results for overlap populations. The lower atomic populations for AlN implied slightly less ionic character in the bonding than GaN and InN, but bond overlaps indicated similar covalency. In summary, bonding in the three nitride materials was found

to be similar, though no definite trends were found.

From partial density of states calculations, the valence bands were resolved into the respective anion and cation s, p and d angular momentum components. In this way, the top of the valence band in the nitrides was identified as being dominated by the N p orbitals. In AlN, the lower bands were largely composed of N s orbitals, in contrast to GaN and InN where the most significant contributions came from $3d$ and $4d$ orbitals respectively. For these, the d electrons were explicitly treated as valence. The d orbitals and the N s orbitals are close in energy, but no real evidence of s-d hybridisation was found.

In chapter 3, the empirical pseudopotential method (EPM) was described, with particular emphasis on the pseudopotential concept. Pseudopotentials describing the zincblende and wurtzite materials were obtained using a functional form that could be easily extended for use in strained material and alloys. Using these pseudopotentials, full zone band structures were calculated that agreed well with the first principles calculations of chapter 2. In addition, the dielectric response of the three nitride materials in both structures was investigated. For all three materials, these reproduced general features in the dielectric function. On this basis, our EPM models for the III-V nitride materials were satisfactory and could be extended for further work on strained material and alloys.

Chapter 4 introduced the **k.p** method and explained its particular suitability for device modelling. Using the EPM band structure of chapter 3 as a basis, **k.p** parameters for both zincblende and wurtzite structures were obtained using different methods. For the zincblende structure, one set of **k.p** parameters were obtained indirectly from the effective masses (which themselves were calculated from the curvature of the valence bands). Another set was calculated directly from the EPM momentum matrix elements and energy eigenvalues at the zone centre. For the wurtzite structure, one set of **k.p** parameters was obtained by using a simple fitting procedure in which the parameters were varied in a Monte Carlo scheme until satisfactory agreement with the

EPM band structure was reached. As for the zincblende structure, a second set was calculated directly from the EPM momentum matrix elements and energy eigenvalues at the zone centre. For both the zincblende and wurtzite structures, the fitted and direct **k.p** parameters produced equally good band structure. The direct method for obtaining **k.p** parameters was advocated over an arbitrary fitting approach on the basis that it provided an unambiguous route to obtaining **k.p** parameters while producing band structure of the same quality. Effective masses for both structures were also calculated. For the zincblende structure, effective masses obtained using the direct method were in good agreement with those calculated simply from the curvature of the bands, and with others in the literature. For the wurtzite structure, effective masses were calculated using both the fitting approach and the direct method. These two methods gave effective masses in reasonable agreement with each other, and were similar to others in the literature.

Also in this chapter, various aspects of nitride alloys were studied using the empirical pseudopotential method within the virtual crystal approximation. In particular, the variation of the band gap with alloy composition was investigated. In this work, both GaInN and AlInN were found to exhibit strong band gap bowing, in contrast to AlGaN where the band gap was found to vary almost linearly with composition. Both of these results were in good agreement with experiment, despite the simple nature of the model used to describe them. Contrary to experiment, for GaInN a flip-over in the normal ordering of the bands was found for small In compositions. This discrepancy was explained by the simple nature of the model used in this work.

The direct method of obtaining **k.p** parameters was also used to produce consistent parameter sets for a series of alloy compositions in the wurtzite structure. To contrast this with the fitting method, fitted parameters were also calculated for GaN. The direct method gave parameters which were all smoothly varying, but this was not the case for all of the parameters obtained from the fitting method. This highlighted the arbitrary nature of a fitting approach, in that the **k.p** parameters obtained in such a way are

dependent on the particular weightings used in the procedure. On this basis, the direct method can be viewed as preferable to those in which parameters are fitted.

Chapter 5 began by addressing the properties of the nitride materials under strain using the empirical pseudopotential method. For this, deformation potentials describing the variation of the band gap with strain were obtained. Strain was also found to have a profound effect on the ordering of the uppermost valence bands, with small positive strain resulting in a flip-over in the normal ordering of the bands for both GaN and InN.

Band offsets for several nitride heterojunctions were also calculated in this work, and were generally found to be in good agreement with other results reported in the literature. The salient feature of the calculated band offsets was the strong forward/backward asymmetry all the heterojunctions exhibited. This effect was largest for heterojunctions involving InN where the lattice mismatch was quite large, but was still significant for the two GaN-AlN systems. This demonstrated the flexibility afforded by nitride heterojunctions, in that offsets can potentially be engineered towards a desired value through use of alloys of appropriate composition. Previously, the origin of the asymmetry has been attributed to strain-induced piezolectric fields. This view is reinforced by the results of this work, in which it was found that the largest lattice mismatch (and correspondingly the greatest strain) resulted in the greatest asymmetry.

In conclusion, several methods have been employed to study the electronic properties of the III-V nitrides and their alloys. Using these, a consistent picture of their electronic structure has been constructed, from which band structure details and transport parameters important in device design and optimisation have been obtained.

6.1 Suggestions for further work

Following on from the work contained in this thesis, there are several areas in which further work could be done. In chapter 2, bonding in the III-V nitrides was investigated

using population analysis. To help establish this method for use in more complicated calculations it would be interesting to extend this study to investigate the relationship between the calculated overlap populations and bond lengths in strained material.

There is also great scope for further investigation of valence band offsets in the nitride materials. In chapter 5, the band offsets of the nitride heterojunctions were found to depend upon the nitride material was used for the substrate. This forward/backward asymmetry could be investigated more fully by studying the exact relationship between the band offset and the substrate lattice constant used in the calculation. This would help confirm the origin of this effect and how it could be exploited in device design. Additional insight could also be gained by using population analysis to study the bonding at the heterojunction interface.

The empirical pseudopotentials obtained for the nitride materials could also be used to study other nitride alloys. In particular, GaAsN is an alloy which exhibits extremely large band gap bowing, with the substitution of 1% nitrogen for arsenic in GaAs resulting in a decrease in the band gap from 1.42 to 1.25 eV [150]. This bowing is such that in contrast to the nitride alloys considered in this work, the parametrisation of the band gap reduction requires a composition dependent bowing parameter. This suggests that the origin of the band gap variation differs from that in conventional nitride alloys, and this could be investigated using the pseudopotentials developed in this work.

Appendix A

Form Factors

A.1 Zincblende

There are two atoms in the zincblende unit cell. Denoting the position of these atoms as $\mathbf{r}_\alpha$ and $\mathbf{r}_\beta$, the *crystal pseudopotential* form factors are obtained by placing the respective *atomic pseudopotentials* at these positions:

$$V(\mathbf{G}) = \frac{1}{\Omega}\int \left[V_\alpha(\mathbf{r})e^{-i\mathbf{G}.\mathbf{r}_\alpha} + V_\beta(\mathbf{r})e^{-i\mathbf{G}.\mathbf{r}_\beta}\right] e^{-i\mathbf{G}.\mathbf{r}}d\mathbf{r} \tag{A.1}$$

It is convenient to take the midpoint between the two atoms as the origin of the unit cell. Thus, the positions $\mathbf{r}_\alpha$ and $\mathbf{r}_\beta$ of the two atoms are given by

$$\begin{aligned} \mathbf{r}_\alpha &= \tfrac{\mathbf{a}_0}{8}(1,1,1) = \tau \\ \mathbf{r}_\beta &= -\tfrac{\mathbf{a}_0}{8}(1,1,1) = -\tau \end{aligned}$$

and our expression for $V(\mathbf{G})$ now becomes

$$V(\mathbf{G}) = \frac{1}{\Omega}\int \left[V_\alpha(\mathbf{r})e^{-i\mathbf{G}.\tau} + V_\beta(\mathbf{r})e^{i\mathbf{G}.\tau}\right] e^{-i\mathbf{G}.\mathbf{r}}d\mathbf{r} \tag{A.2}$$

But the atomic pseudopotential form factors as given by equation 3.12 are

$$\begin{aligned} V_\alpha(\mathbf{G}) &= \frac{1}{\Omega_\alpha}\int V_\alpha(\mathbf{r})e^{-i\mathbf{G}.\mathbf{r}} \\ V_\beta(\mathbf{G}) &= \frac{1}{\Omega_\beta}\int V_\beta(\mathbf{r})e^{-i\mathbf{G}.\mathbf{r}} \end{aligned} \tag{A.3}$$

where the atomic volumes Ω_α and Ω_β are usually taken to be equal to $\Omega/2$. Thus we can now write

$$V(\mathbf{G}) = \frac{1}{2}\left[e^{i\mathbf{G}.\tau}V_\alpha(\mathbf{G}) + e^{-i\mathbf{G}.\tau}V_\beta(\mathbf{G})\right] \tag{A.4}$$

It is then convenient to define symmetric and antisymmetric form factors, V^s and V^a:

$$\begin{aligned} V^s(\mathbf{G}) &= \frac{1}{2}(V_\alpha(\mathbf{G}) + V_\beta(\mathbf{G})) \\ V^a(\mathbf{G}) &= \frac{1}{2}(V_\alpha(\mathbf{G}) - V_\beta(\mathbf{G})) \end{aligned} \tag{A.5}$$

Substituting these into the expression for $V(\mathbf{G})$ we obtain

$$\begin{aligned} V(\mathbf{G}) &= V^s(\mathbf{G})\cos(\mathbf{G}.\tau) + iV^a(\mathbf{G})\sin(\mathbf{G}.\tau) \\ &= V^s(\mathbf{G})S^s(\mathbf{G}) + iV^a(\mathbf{G})S^a(\mathbf{G}) \end{aligned} \tag{A.6}$$

where we have introduced the symmetric and antisymmetric structure factors, $S^s(\mathbf{G}) = \cos(\mathbf{G}.\tau)$ and $S^a(\mathbf{G}) = \sin(\mathbf{G}.\tau)$. With the form factors written this way, several observations can be made. Foremost, it is clear that the antisymmetric form factors vanish entirely for those materials possessing the diamond structure, since there is only one atomic type and thus $V_\alpha = V_\beta$. In many materials the difference in the potentials of the two atom types is small, and thus the antisymmetric part of the pseudopotential

is also be expected to be small. However, N is a small atom and thus its potential should be considered to be rather different from that of Al and particularly Ga and In. Therefore the antisymmetric part of the pseudopotential will be sizeable for the nitrides.

It is also common practice to make the assumption that the pseudopotentials of the two atoms are spherically symmetric. Thus the symmetric and antisymmetric form factors can now simply be described by $V^s(G)$ and $V^a(G)$, since they depend on just the magnitude of $\mathbf{G}$. The form factors with $G^2 > 11(\frac{2\pi}{a_0})^2$ are commonly neglected. This is because $V(\mathbf{G})$ decreases as $\frac{1}{G^2}$ for large G, and so imposing this cutoff is an acceptable compromise between accuracy and speed.

The resulting zincblende $\mathbf{G}$ vectors and their associated structure factors are shown in Table A.1.Note that the contribution to the pseudopotential from $\mathbf{G}_0$ is constant, and $\neq 0$. However, this simply shifts all the energies up or down and effectively sets the origin of the band structure. As such, it is usually set to zero. Also note that the symmetric structure factor for $\mathbf{G}_4$ is zero, as is the antisymmetric structure factor for $\mathbf{G}_8$. Thus the local pseudopotential in the zincblende structure can be described by 3 symmetric and 3 antisymmetric form factors: $V^s(\sqrt{3}), V^s(\sqrt{8}), V^s(\sqrt{11}), V^a(\sqrt{3}), V^a(\sqrt{4})$ and $V^a(\sqrt{11})$.

A.2 Wurtzite

In the wurtzite case, the pseudopotential form factors can be expressed in a similar way to those in the zincblende structure:

$$V(\mathbf{G}) = V^s(\mathbf{G})S^s(\mathbf{G}) + iV^a(\mathbf{G})S^a(\mathbf{G}) \tag{A.7}$$

An ideal wurtzite unit cell (with $u = a/c = 3/8$), with the associated $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$

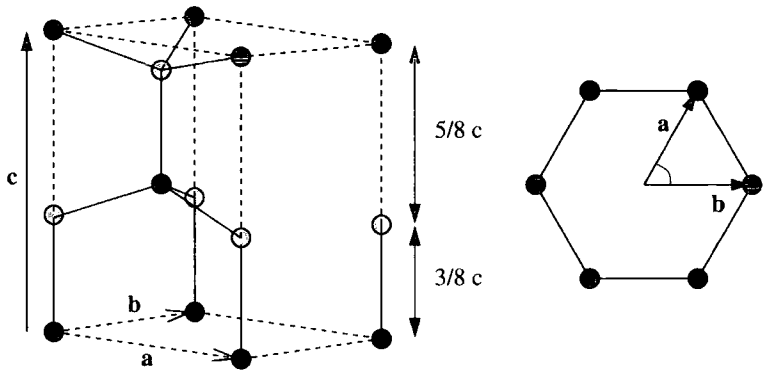

Figure A.1: The crystal structure of wurtzite GaN.

primitive vectors, is shown in Figure A.1. If the following definitions are made

$$\begin{aligned} \mathbf{r}_1 &= \left[\frac{1}{6}, \frac{1}{6}, \frac{1}{2}\left(\frac{1}{2}+u\right)\right] \\ \mathbf{r}_2 &= \left[\frac{1}{6}, \frac{1}{6}, \frac{1}{2}\left(\frac{1}{2}-u\right)\right] \end{aligned} \tag{A.8}$$

then the atoms of one type are at $+\mathbf{r}_1$ and $-\mathbf{r}_2$, and atoms of the other type are at $-\mathbf{r}_1$ and $+\mathbf{r}_2$. Again, like in the zincblende case, spherical symmetry is assumed, and thus $V^s(\mathbf{G})$ and $V^a(\mathbf{G})$ only depend upon the magnitude of $\mathbf{G}$. The symmetric and antisymmetric structure factors are this time given by [65]:

$$\begin{aligned} S^s(\mathbf{G}) &= \cos\left[2\pi\left(\frac{G_x}{6}+\frac{G_y}{6}+\frac{G_z}{4}\right)\right]\cos\left[2\pi\left(\frac{G_z u}{2}\right)\right] \\ S^a(\mathbf{G}) &= \cos\left[2\pi\left(\frac{G_x}{6}+\frac{G_y}{6}+\frac{G_z}{4}\right)\right]\sin\left[2\pi\left(\frac{G_z u}{2}\right)\right] \end{aligned} \tag{A.9}$$

For the wurtzite structure it is necessary to use more form factors than in the zincblende case. The wurtzite $\mathbf{G}$-vectors and their associated structure factors are shown in Table A.1. It is important to note that the values for G^2 and the structure factors in this table are calculated assuming an ideal wurtzite structure with $u = 3/8$. In the non-ideal case used in this work, not only will the values for G^2 be slightly

	G	G^2	$\lvert S^s(\mathbf{G})\rvert$	$\lvert S^a(\mathbf{G})\rvert$
Zincblende	(1,1,1)	3	$\frac{\sqrt{2}}{2}$	$\frac{\sqrt{2}}{2}$
	(2,0,0)	4	0	1
	(2,2,0)	8	1	0
	(3,1,1)	11	$\frac{\sqrt{2}}{2}$	$\frac{\sqrt{2}}{2}$
Wurtzite	(1,0,0)	$2\frac{2}{3}$	$\frac{1}{2}$	0
	(0,0,2)	3	0.71	0.71
	(1,0,1)	$3\frac{5}{12}$	0.33	0.80
	(1,0,2)	$5\frac{2}{3}$	0.35	0.35
	(2,1,0)	8	1	0
	(1,0,3)	$9\frac{5}{12}$	0.80	0.33
	(2,0,0)	$10\frac{2}{3}$	$\frac{1}{2}$	0
	(2,1,2)	11	0.71	0.71
	(2,0,1)	$11\frac{5}{12}$	0.33	0.80
	(0,0,4)	12	0.00*	1.00
	(2,0,2)	$13\frac{2}{3}$	0.35	0.35
	(1,0,4)	$14\frac{2}{3}$	0.00*	0.50

Table A.1: The reciprocal lattice **G** vectors and the associated structure factors for zincblende and wurtzite materials. Note that for the wurtzite case the values of G^2 have been calculated assuming the ideal case of $u = a/c = 3/8$. Those used in the calculations presented in this work are evaluated exactly. Also note that the structure factors marked with an asterisk will be nonzero for the non-ideal case considered in this work.

different, but in particular the structure factors for the $(2,0,0)$ and $(1,0,4)$ **G**-vectors will be nonzero. Thus in this work 12 symmetric and 9 antisymmetric form factors are employed for the wurtzite structure materials.

Appendix B

Publications

Much of the work contained in this thesis has been previously published in journals. The references for these are:

- S. K. Pugh, D. J. Dugdale, S. Brand and R. A. Abram, "Electronic structure calculations on nitride semiconductors", Semicond. Sci. Tech. **14**, 23, (1999)
- S. K. Pugh, D. J. Dugdale, S. Brand and R. A. Abram, "Band-gap and k.p parameters for GaAlN and GaInN alloys", J. Appl. Phys. **86**, 3768, (1999)
- D. J. Dugdale, S. Brand and R. A. Abram, "Determination of band structure parameters in nitride alloys for use in quantum well calculations", Physica Status Solidi B **216**, 351 (1999)
- D. J. Dugdale, S. Brand and R. A. Abram, "Direct calculation of **k.p** parameters for wurtzite AlN, GaN, and InN", Phys. Rev. B **61** 12933, (2000)

Bibliography

[1] G. Landwehr, A. Waag, F. Fischer, H.-J. Lugauer and K. Schüll, Physica E **3**, 158 (1998)

[2] S. J. Pearton, ed., *GaN and Related Materials*, 509–534, Gordon and Breach, New York (1997)

[3] M. A. Khan, Q. Chen, J. W. Wang and C. J. Sun, Inst. Phys. Conf. Ser. **142**, 985 (1995)

[4] B. Gelmont, K. Kim and M. Shur, J. Appl. Phys. **74**, 1818 (1993)

[5] O. Ambacher, J. Phys. D **31**, 2653 (1998)

[6] S. C. Jain, M. Willander, J. Narayan and R. V. Overstraeten, J. Appl. Phys. **87**, 965 (2000)

[7] S. J. Pearton, ed., *GaN and Related Materials*, 1–9, Gordon and Breach, New York (1997)

[8] J. I. Pankove, A. Miller and J. E. Berkeyheiser, J. Lumin. **5**, 84 (1972)

[9] S. Yoshida, S. Misawa and S. Gonda, Appl. Phys. Lett. **42**, 427 (1983)

[10] H. Amano, N. Sawaki, I. Akasaki and T. Toyoda, Appl. Phys. Lett. **48**, 353 (1986)

[11] H. Amano, I. Akasaki, T. Kozawa, K. Hiramatsu, N. Sawak, K. Ikeda and Y. Ishi, J. Lumin. **40–41**, 121 (1988)

[12] H. Amano, H. Kito, K. Hiramatsu and I. Akasaki, Japn. J. Appl. Phys. Part 2 **28**, L2112 (1989)

[13] S. Nakamura, S. Mukai, M. Senoh and N. Iwasa, Japn. J. Appl. Phys. Part 2 **31**, L139 (1992)

[14] S. Nakamura and S. Mukai, Japn. J. Appl. Phys. Part 2 **31**, L1457 (1992)

[15] S. Chichibu, T. Azuhata, T. Sota and S. Nakamura, Appl. Phys. Lett. **70**, 2822 (1997)

[16] H. Schulz and K. H. Thiemann, Solid State Commun. **23**, 815 (1997)

[17] A. F. Wright and J. S. Nelson, Phys. Rev. B **51**, 7866 (1995)

[18] I. Gorczyca and N. E. Christensen, Physica B **185**, 410 (1993)

[19] T. Lei, M. Fanciulli, R. J. Molnar, T. D. Moustakas, R. Graham and J. Scanlon, Appl. Phys. Lett. **95**, 944 (1991)

[20] O. Madelung, M. Schulz and H. Weiss, eds., *Landolt-Börnstein, Numerical Data and Functional Relationships in Science and Technology*, volume III/17a, Springer-Verlag (1982)

[21] R. Dingle, D. D. Sell, S. E. Stokowski and M. Ilegems, Phys. Rev. B **4**, 1211 (1971)

[22] M. Suzuki and T. Uenoyama, Appl. Phys. Lett. **6**, 3378 (1996)

[23] J. B. Jeon, B. C. Lee, Y. M. Sirenko, K. W. Kim and M. A. Littlejohn, J. Appl. Phys. **82**, 386 (1997)

[24] Y. C. Yeo, T. C. Chong, M. F. Li and W. J. Fan, J. Appl. Phys. **84**, 1813 (1998)

[25] J. Kolnik, I. H. Oguzman, K. F. Brennan, R. P. Wang, P. P. Ruden and Y. Wang, J. Apply. Phys. **78**, 1033 (1995)

[26] B. E. Foutz, S. K. O'Leary, M. S. Shur and L. F. Eastman, J. Appl. Phys. **85**, 7727 (1999)

[27] E. O. Kane, *Semiconductors and Semimetals*, volume 1, chapter 3, Academic Press (1966)

[28] M. Bockstedte, A. Kley, J. Neugebauer and M. Scheffler, Comput. Phys. Commun. **107**, 187 (1997)

[29] M. C. Payne, M. P. Teter, D. C. Allan, T. A. Arias and J. D. Joannopoulos, Rev. Mod. Phys. **64**, 1045 (1992)

[30] G. Kresse and J. Hafner, Phys. Rev B **47**, 558 (1993)

[31] G. Kresse and J. Furthmüller, Comput. Mat. Sci. **6**, 15 (1996)

[32] G. Kresse and J. Furthmüller, Phys. Rev. B **54**, 11169 (1996)

[33] J. A. Ihm, A. Zunger and M. L. Cohen, J. Phys. C **12**, 4409 (1979)

[34] P. Denteneer and W. van Haeringen, J. Phys. C **18**, 4127 (1985)

[35] P. Hohenberg and W. Kohn, Phys. Rev. **136**, 864 (1964)

[36] W. Kohn and L. J. Sham, Phys. Rev. **140**, 1133 (1965)

[37] C. Stampfl and C. G. V. de Walle, Phys. Rev. B **59**, 5521 (1999)

[38] H. J. Monkhorst and J. D. Pack, Phys. Rev. B **13**, 5188 (1976)

[39] D. J. Chadi and M. L. Cohen, Phys. Rev. B **8**, 5747 (1973)

[40] G. P. Srivastava, J. Phys. A **17**, L317 (1984)

[41] S. P. Francis and M. C. Payne, J. Phys. Condens. Matter **2**, 4395 (1990)

[42] N. Troullier and J. L. Martins, Phys. Rev. B **43**, 1993 (1991)

[43] D. Vanderbilt, Phys. Rev. B **32**, 8412 (1985)

[44] J. H. Edgar, ed., *Properties of Group III Nitrides*, 135–156, Inspec. London (1994)

[45] V. Fiorentini, M. Methfessel and M. Scheffler, Phys. Rev. B **47**, 13353 (1993)

[46] N. E. Christensen and I. Gorczyca, Phys. Rev. B **50**, 4397 (1994)

[47] R. Miotto, A. C. Ferraz and G. P. Srivastava, Solid State Commun. **115**, 67 (2000)

[48] I. Petrov, E. Mojab, R. C. Powell and J. E. Greene, Appl. Phys. Lett. **60**, 2491 (1992)

[49] U. Grossner, J. Furthmüller and F. Bechstedt, Phys. Rev. B **58**, R1722 (1998)

[50] S. Strite, J. Ruan, Z. Li, A. Salvador, H. Chen, D. J. Smith, W. J. Choyke and H. Morkoc, J. Vac. Sci. Technol. B **9**, 1924 (1991)

[51] S. Strite, D. Chandrasekhar, D. J. Smith, J. Sariel, H. Chen, N. Teraguchi and H. Morkoc, J. Cryst. Growth **127**, 204 (1993)

[52] S. H. Wei and A. Zunger, Appl. Phys. Lett. **69**, 2719 (1996)

[53] M. Suzuki, T. Uenoyama and A. Yanase, Phys. Rev. B **52**, 8132 (1995)

[54] K. Osamura, S. Naka and Y. Murakami, J. Appl. Phys. **46**, 3432 (1975)

[55] J. A. Majewski, M. Städele and P. Vogl, Internet. J. Nitride Semicond. Res. **1**, 30 (1996)

[56] D. Sanchez-Portal, E. Artacho and J. M. Soler, Solid State Commun. **95**, 685 (1995)

[57] D. Sanchez-Portal, E. Artacho and J. M. Soler, J. Phys. Condens. Matter **8**, 3859 (1996)

[58] R. S. Mulliken, J. Chem. Phys. **23**, 1833 (1955)

[59] M. D. Segall, R. Shah, C. J. Pickard and M. C. Payne, Phys. Rev. B **54**, 16317 (1996)

[60] R. G. Pearson, Inorg. Chem. **27**, 734 (1988)

[61] R. Miotto, G. P. Srivastava and A. C. Ferraz, Physica B **292**, 97 (2000)

[62] J. C. Phillips and L. Kleinmann, Phys. Rev. **116**, 287 (1959)

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

[64] M. L. Cohen and V. Heine, Solid State Physics **24**, 73 (1970)

[65] M. L. Cohen and J. R. Chelikowsky, *Electronic Structure and Optical Properties of Semiconductors*, Springer-Verlag (1989)

[66] E. Fermi, Nuovo Cimento **11**, 157 (1934)

[67] H. J. Hellmann, J. Chem. Phys. **3**, 61 (1935)

[68] C. Herring, Phys. Rev. **57**, 1169 (1940)

[69] J. C. Philips and K. C. Pandey, Phys. Rev. Lett. **30**, 787 (1973)

[70] J. R. Chelikowsky and M. L. Cohen, Phys. Rev. Lett. **32**, 674 (1974)

[71] D. Brust, Phys. Rev. B **4**, 3497 (1971)

[72] J. R. Chelikowsky and M. L. Cohen, Phys. Rev. B **14**, 556 (1976)

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

[74] S. Bloom and T. K. Bergstresser, Solid State Commun. **6**, 465 (1970)

[75] L. M. Falicov and P. J. Lin, Phys. Rev. **141**, 562 (1966)

[76] A. Rubio, J. L. Corkhill, M. L. Cohen, E. L. Shirley and S. G. Louie, Phys. Rev. B **48**, 11810 (1993)

[77] Y. Yeo, T. C. Chong and M. F. Li, J. Appl. Phys. **83**, 1429 (1998)

[78] S. J. Jenkins, G. P. Srivastava and J. C. Inkson, J. Phys. Condens. Matter **6**, 8781 (1994)

[79] F. Stern, *Elementary Theory of Optical Properties of Solids*, volume 15 of *Solid State Physics: Advances in Research and Applications*, chapter 4, Academic (1963)

[80] S. Logothetidis and J. Petalas, Phys. Rev. B **50**, 18017 (1994)

[81] T. Azuhata, T. Sota, K. Suzuki and S. Nakamura, J. Phys. Condens. Matter **7**, L129 (1995)

[82] A. S. Barker and M. Ilegems, Phys. Rev. B **7**, 743 (1973)

[83] R. Wang, P. B. Ruden, J. Kolnik, I. Oguzman and K. F. Brennan, J. Phys. Chem. Solids **58**, 913 (1997)

[84] K. Karch and F. Bechstedt, Phys. Rev. B **56**, 3560 (1997)

[85] W. R. L. Lambrecht, B. Segall, J. Rife, W. R. Hunter and D. K. Wickenden, Phys. Rev. B **51**, 13516 (1995)

[86] A. J. Fischer, W. Shan, J. J. Song, Y. C. Chang, R. Horning and B. Goldenberg, Appl. Phys. Lett. **71**, 1981 (1997)

[87] Q. Guo, O. Kato, M. Fujisawa and A. Yoshida, Solid State Commun. **83**, 721 (1992)

[88] Q. Guo, M. Nishio and H. Ogawa, Phys. Rev. B **55**, 15987 (1997)

[89] T. Wethkamp, K. Wilmers, N. Esser, W. Richter, O. Ambacher, H. Angerer, G. Jungk, R. L. Johnson and M. Cardona, Thin Solid Films **313–314**, 745 (1998)

[90] T. Wethkamp, K. Wilmers, C. Cobet, N. Esser, W. Richter, O. Ambacher, M. Stutzmann and M. Cardona, Phys. Rev. B **59**, 1845 (1999)

[91] Z. Yang and Z. Xu, J. Phys. Condens. Matter **8**, 8303 (1996)

[92] J. L. P. Hughes, Y. Wang and J. E. Sipe, Phys. Rev. B **55**, 13630 (1997)

[93] L. X. Benedict and E. L. Shirley, Phys. Rev. B **59**, 5441 (1999)

[94] L. X. Benedict, T. Wethkamp, K. Wilmers, C. Cobet, N. Esser, E. L. Shirley, W. Richter and M. Cardona, Solid State Commun. **112**, 129 (1999)

[95] A. B. Djurisic and E. H. Li, Appl. Phys. Lett. **73**, 868 (1998)

[96] A. B. Djurisic and E. H. Li, J. Appl. Phys. **85**, 2848 (1999)

[97] M. Cardona, Solid State Commun. **1**, 109 (1963)

[98] M. Cardona, J. Appl. Phys. **36**, 2181 (1965)

[99] S. Albrecht, L. Reining, R. D. Sole and G. Onida, Phys. Rev. Lett. **80**, 4510 (1998)

[100] P. O. Löwdin, J. Chem. Phys. **19**, 1396 (1951)

[101] E. O. Kane, ed., *Handbook on Semiconductors*, volume 1, chapter 4, North-Holland (1982)

[102] R. Eppenga, M. F. H. Schuurmans and S. Colak, Phys. Rev. B **36**, 1554 (1987)

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

[104] B. A. Foreman, Phys. Rev. B **48**, 4964 (1993)

[105] P. N. Stavrinou and R. van Dalen, Phys. Rev. B **55**, 15456 (1997)

[106] W. J. Fan, M. H. Li, T. C. Chong and J. B. Xia, J. Appl. Phys. **79**, 188 (1996)

[107] A. T. Meney, E. P. O'Reilly and A. R. Adams, Semicond. Sci. Technol. **11**, 897 (1996)

[108] S. L. Chuang and C. S. Chang, Phys. Rev. B **54**, 2491 (1996)

[109] G. B. Ren, Y. M. Liu and P. Blood, Appl. Phys. Lett. **74**, 1117 (1999)

[110] K. Kim, W. R. L. Lambrecht, B. Segall and M. Schilfgaarde, Phys. Rev. B **56**, 7363 (1997)

[111] J. A. Majewski, M. Städele and P. Vogl, Mat. Res. Soc. Symp. Proc. **449**, 887 (1997)

[112] G. L. Bir and G. E. Pikus, *Symmetry and Strain-Induced Effects in Semiconductors*, Wiley (1974)

[113] S. Chichibu, T. Azuhata, T. Sota and S. Nakamura, Appl. Phys. Lett. **69**, 4188 (1996)

[114] S. Chichibu, S. Wada and S. Nakamura, Appl. Phys. Lett. **71**, 2346 (1997)

[115] Y. Narukawa, Y. Kawakami, S. Fujita, S. Fuita and S. Nakamura, Phys. Rev. B **55**, 1938R (1997)

[116] M. J. Bergmann and H. C. Casey, J. Appl. Phys. **84**, 1196 (1998)

[117] A. Hangleiter, G. Frankowsky, V. Härle and F. Scholz, Mat. Sci. Eng. B **43**, 201 (1997)

[118] C. Wetzel, T. Takeuchi, S. Yamaguchi, H. Katoh, H. Amano and I. Akasaki, Appl. Phys. Lett. **73**, 1994 (1998)

[119] M. D. McCluskey, C. G. V. de Walle, C. P. Master, L. T. Romano and N. M. Johnson, Appl. Phys. Lett. **72**, 2725 (1998)

[120] R. Magri, S. Froyen and A. Zunger, Phys. Rev. B **44**, 7947 (1991)

[121] L. Bellaiche, T. Mattila, L. W. Wang, S. H. Wei and A. Zunger, Appl. Phys. Lett. **74**, 1842 (1999)

[122] G. Frankowsky, F. Steuber, V. Härle, F. Scholz and A. Hangleiter, Appl. Phys. Lett. **68**, 3746 (1996)

[123] A. F. Wright and J. S. Nelson, Appl. Phys. Lett. **66**, 3465 (1995)

[124] K. S. Kim, A. Saxler, P. Kung, M. Razeghi and K. Y. Lim, Appl. Phys. Lett. **71**, 800 (1997)

[125] Q. Guo, H. Ogawa and A. Yoshida, J. Cryst. Growth **146**, 462 (1995)

[126] K. Kubota, Y. Kobayashi and K. Fujimoto, J. Appl. Phys. **66**, 2984 (1989)

[127] E. P. O'Reilly, Semicond. Sci. Technol. **4**, 121 (1989)

[128] S. C. Jain, A. H. Harker and R. A. Cowley, Philos. Mag. A **75**, 1461 (1997)

[129] N. Grandjean and J. Massies, Appl. Phys. Lett. **71**, 1816 (1997)

[130] Z. Sitar, M. J. Paisley, B. Yan, J. Ruan, W. J. Choyke and R. F. Davis, J. Vac. Sci. Technol. B **8**, 316 (1990)

[131] S. Krishnankutty, R. M. Kolbas, M. A. Khan, J. N. Kuznia, J. M. V. Hove and D. T. Olson, J. Electron. Mater. **21**, 437 (1992)

[132] A. D. Bykhovski, B. L. Gelmont and M. S. Shur, J. Appl. Phys. **78**, 3691 (1995)

[133] J. I. Pankove and T. D. Moustakas, eds., *GaN*, 459–472, Academic, New York (1998)

[134] G. Martin, A. Botchkarev, A. Rockett and H. Morkoc, Appl. Phys. Lett. **68**, 2541 (1996)

[135] L. R. Saravia and D. Brust, Phys. Rev. B **178**, 1240 (1969)

[136] K. Kim, W. R. L. Lambrecht and B. Segall, Phys. Rev. B **53**, 16310 (1996)

[137] A. Franciosi and C. G. V. de Walle, Surface Science Reports **25**, 1 (1996)

[138] A. Baldereschi, S. Baroni and R. Resta, Phys. Rev. Lett. **61**, 734 (1988)

[139] L. Colombo, R. Resta and S. Baroni, Phys. Rev. B **44**, 5572 (1991)

[140] N. E. Christensen, Phys. Rev. B **37**, 4528 (1988)

[141] C. G. V. de Walle and R. M. Martin, Phys. Rev. B **35**, 8154 (1987)

[142] W. I. Wang and F. Stern, J. Vac. Sci. Tech. B **3**, 1280 (1985)

[143] A. D. Katnani and R. S. Bauer, Phys. Rev. B **33**, 1106 (1986)

[144] J. S. Batey and S. L. Wright, Appl. Phys. **59**, 1200 (1986)

[145] F. Bernardini and V. Fiorentini, Phys. Rev. B **57**, 9427 (1998)

[146] J. A. Majewski, M. Städele and P. Vogl, Mat. Res. Soc. Symp. Proc. **482**, 1 (1998)

[147] C. G. V. de Walle and J. Neugebauer, Appl. Phys. Lett. **70**, 2577 (1997)

[148] M. B. Nardelli, K. Rapcewicz and J. Bernholc, Phys. Rev. B **55**, 7323 (1997)

[149] J. R. Waldrop and R. W. Grant, Appl. Phys. Lett. **68**, 2879 (1996)

[150] M. Weyers, M. Sato and H. Ando, Japn. J. Appl. Phys. Part 2 **31**, L853 (1992)