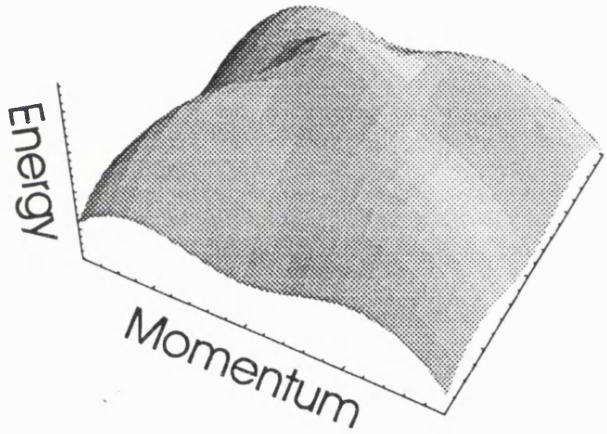
Optimisation of *p*-type two-dimensional semiconductor structures for FET applications

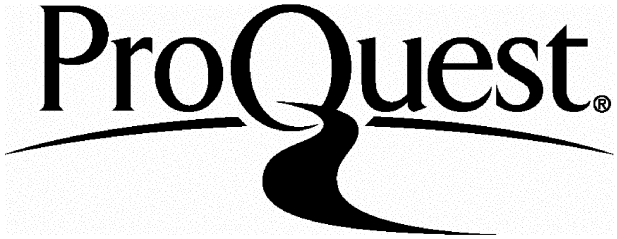
Rob van Dalen



A thesis submitted for the degree of Doctor of Philosophy

University College London

Submitted July 1998



Abstract

The phenomenally high electron mobility that can be achieved in GaAs-based heterostructures has led to the development of *n*-type Heterostructure Field Effect Transistors (HFETs) with enhanced power handling capabilities and high-frequency performance. Performance of *p*-channel HFETs is substantially lower, essentially because of the lower hole mobility as compared to that of electrons. The latter has almost completely prevented their use, despite the intrinsic advantages of complementary circuits.

An investigation to enhance the performance of *p*-type FETs for power applications by means of selective doping and wavefunction engineering was performed, both theoretically as well as experimentally. A self-consistent multiband 'k-dot-p' effective mass model was set up to provide detailed bandstructure information, this model was then used in conjunction with an RPA scattering model to provide estimates for the hole mobility. The experimental mobilities obtained for a set of 60Å channel-delta-doped InGaAs-AlGaAs QWs show an improvement in the mobility by a factor 2.5 when moving the impurity plane from the centre of the QW toward its interface, in qualitative agreement with the theoretical predictions. This verifies the capabilities of such a model to optimise *p*-type performance, as well as illustrates the possible gain in device performance by careful design of the structure.

A related investigation in this thesis concerns the elementary question of how to connect a wavefunction across an interface between two different materials, which has long been a major cause of controversy in effective mass theory. Insight in this problem was not provided until the derivation of an 'exact' envelope function theory, and it was shown that the resulting boundary conditions for [001] growth differ considerably from the conventional ones. In this thesis, the extension to arbitrary growth directions is provided in the form of a set of boundary condition rules that is to replace the traditional symmetrisation procedure.

Contents

List of figures

Chapter 4

Chapter 5

Chapter 6

List of tables

Chapter 2

Chapter 5

Chapter 6

Chapter 1 : Introduction and overview

The motives and incentives behind the work performed during this Ph.D. are outlined in this chapter. As the title suggests, the ultimate goal was to investigate the possibilities to enhance the performance of *p*-type Field Effect Transistors (FETs), as *n*-channel FETs currently outperform their *p*-channel counterparts by a significant margin. A considerable amount of work was performed in setting up a theoretical model to analyse heterostructure designs for their suitability in FET devices.

This chapter starts with the bandstructure concept in bulk and compositionally varying semiconductors. It shown that the quantum confinement associated with a heterostructure causes the momentum along the growth direction to take discrete values, resulting in the formation of multiple bands. For the electron case, these bands exhibit a parabolic dependence on the in-plane momentum, whereas it is shown that the valence bands exhibit a strong non-parabolic and anisotropic behaviour because of the interaction between the heavy and light hole bands. The latter means that the valence bands are susceptible to tailoring, for example by choosing a particular growth direction or the application of strain.

Next, the relevance of bandstructure information for device optimisation is highlighted. A short introduction to the link between an experimentally verifiable quantity such as the carrier mobility and the bandstructure is presented. Here, an effective mass is introduced which is directly related to the curvature of the energy bands, and it is shown that a high curvature (or low mass) enhances the performance of electro- and electro-optical devices such as field-effect transistors or lasers.

Section 1.2 contains a brief introduction to the chosen bandstructure model, along with a short discussion on the boundary condition problem intrinsic to this particular model. As the subject of research concerns the optimisation of device layers which contain carriers as well as doping, the embedding of this bandstructure model in a Poisson solver is discussed. Such a combination, often referred to as a self-consistent Poisson-Schrödinger solver, can provide information about subband populations, wavefunctions and corresponding bandstructure. Using a fixed library of material parameters such as effective masses and dielectric constants, these calculations only require user information about the composition of the various layers to perform a simulation, without the need of additional fitting parameters.

The bottleneck for *p*-type FET operation is discussed and is shown to be the low hole mobility in GaAs. The need for improvements of *p*-channel FETs is highlighted by discussing relevant applications which specifically require the presence of *p*-channel FETs. The latter is substantiated by showing a relevant implementation in the form of a (simplified) circuit diagram. Current state-of-the-art hole mobilities in GaAs-based heterostructures are presented which were obtained via a literature survey. This chapter then ends with a breakdown of the topics that are dealt with in the subsequent in-depth chapters.

1.1 - The bandstructure concept in semiconductors and its relevance for devices

Electrons (and holes) in a crystal lattice exhibit a discrete energy spectrum. This is a direct result of the symmetry of the electron's surrounding charge distribution. As the electron and hole energies depend on the velocity of the particle, or to be more exact the momentum $\hbar\underline{k}$, the energy at which an electron can exist takes the form of a set of energy bands $E(\underline{k})$. Knowledge of these bands, or the bandstructure, is of key importance when it comes to explaining electrical and optical properties of solids. An important bandstructure feature is the occurrence of energy gaps, which are energy zones without any allowed states. The origin of these energy gaps can easily be understood by looking at a one-dimensional Kronig-Penney model [Kit86]. Such a model shows that the two-standing wave solutions around a gap pile up the electrons at different regions, which explains their difference in energy.

The presence of an energy gap is characteristic of any semiconductor material and plays an important part in explaining the fundamental differences in electrical properties between conductors, semiconductors and metals. Similarly, the poor optical performance of a material like silicon is directly related to the fact that it has an indirect band gap, i.e. the energy extrema in the energy gap are not at the same *k*-point. This is in contrast to a material as for example GaAs which has a direct gap, and is routinely used for optical devices. On a more quantitative level one finds that various semiconductor material or device properties (for example transport mobility, cyclotron resonance or laser gain) are intimately related to the curvature of the bands around the band gap.

In the next paragraphs the general form of the band structure in bulk and compositionally varying structures is discussed. Furthermore, the need for bandstructure information to perform theoretical device calculations, either to make predictions about their electrical or optical properties, is

outlined. As an illustration, some qualitative relations are given between bandstructure features and particular electrical or optical properties that are desirable from a device point of view.

1.1.1 - The bandstructure concept, application to bulk crystals

Although the majority of this thesis is concerned with compositionally varying structures, it is instructive to gain some insight in the bandstructure of bulk semiconductors first. An important feature in such a periodic array of atoms, or lattice, is the existence of a certain number of symmetry operations. Such a lattice can be uniquely described by defining a basis (a primitive cell) along with a set of lattice translation vectors that transform the lattice into itself. Fig.1.1 shows a Zinc-Blende lattice with its primitive (Wigner-Seitz) cell. Unlike simple cubic structures, the primitive cell in a zinc-blende lattice contains two atoms, one of each species. When viewing fig.1.1 one has to remember that the atoms that are lying on the circumference of the primitive cell have to be 'shared' with adjacent cells. Alternatively, the primitive cell can be displaced from the chosen point of origin and this confirms that its volume contains two atoms.

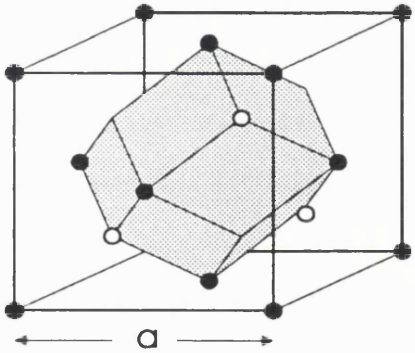


Fig.1.1 : A Zinc-Blende lattice with its primitive cell. The lattice can be seen as a double fcc-lattice where, in the case of GaAs, both Ga and As sit on their own separate fcc-lattice. Monoatomic Si has the same crystal structure but this is usually referred to as the diamond structure because of the additional (inversion) symmetry.

A so-called reciprocal lattice is introduced which is the set of all the vectors K for which a plane wave e^{iKr} has the same periodicity as the primitive cell. Bloch's theorem states that one only needs to solve the wavefunction within a primitive cell and that the area in k-space that needs to be evaluated can be restricted to the so-called (first) Brillouin zone. This can easily be understood by realising that moving from k to a position out of the Brillouin zone by adding a reciprocal lattice vector K is equivalent to imposing a modulation with the same periodicity as the Wigner-Seitz cell on the primitive cell solution. Such a modulation is obviously equivalent to moving to a new solution at the original value of the momentum. At this point, it is worthwhile to point out that the Bloch theorem is a direct result of the presence of translational symmetry in a periodic structure only. Translation symmetry is however not the only symmetry present. A more detailed discussion on symmetry, especially on the use of the additional symmetries to further simplify the bandstructure problem, follows in chapter 2.

Fig.1.2 shows the calculated bandstructure of GaAs. The various labels Γ,*L*,*X* correspond to certain translational symmetry points in the Brillouin zone. The indices denote the rotational symmetries for the particular band at that point in *k*-space. The number of outer shell electrons in GaAs is exactly that required to fill all bands up to the lowest bands labelled Γ_{15}. The latter triple degenerate combination of bands will be designated as the valence band. Also shown in fig.1.2 is that the next band (Γ_1), the conduction band, is separated from the valence band by an energy gap. The latter is characteristic of any semiconductor.

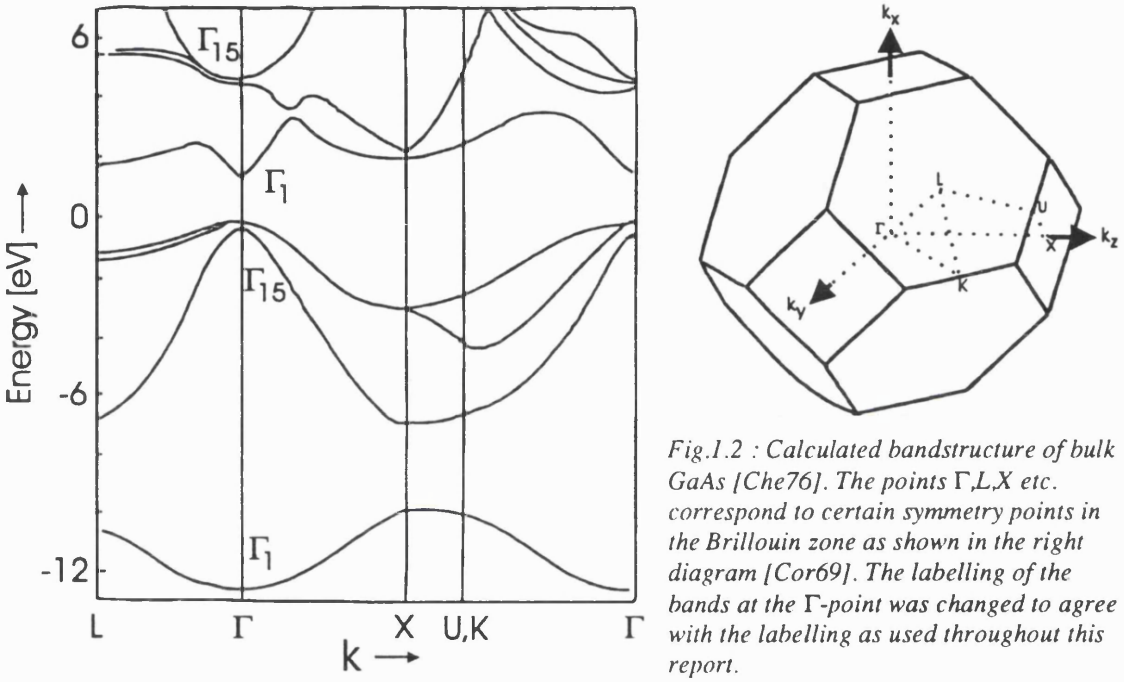


Fig.1.2 : Calculated bandstructure of bulk GaAs [Che76]. The points Γ,*L*,*X etc. correspond to certain symmetry points in the Brillouin zone as shown in the right diagram [Cor69]. The labelling of the bands at the* Γ*-point was changed to agree with the labelling as used throughout this report.*

To see why a full knowledge of the bandstructure is generally not required, one has to realise that all electrical and optical activity is ultimately driven by the presence of electrons and holes. At finite temperatures and subject to any impurities present, a certain number of electrons (holes) are able to occupy the conduction (valence) band, given by Fermi-Dirac statistics. From a device point of view, the electron and hole population can be controlled by intentionally doping the semiconductor with donors or acceptors and produce devices like diodes, transistors (BJTs), field effect transistors (FETs) or semiconductor lasers. Two observations can now be made:

- The zone edge momentum vector corresponding to the *X*-point (fig.1.2) corresponds to $k=2\pi/a$ [Sze81] where *a* is the lattice constant of the specific material. For typical III-V materials one thus obtains that the zone edge momentum is of the magnitude of $1.1\mathring{A}^{-1}$. In contrast, even a very high carrier density of $4\cdot10^{18}$ cm^{-3} in bulk (well beyond the onset of tunnelling in abrupt p-n diodes [Sze66]) corresponds to a Fermi-vector of only $0.05\mathring{A}^{-1}$ (a similar argument applies for two-dimensional structures).

- The spread in energy due to thermal activation, i.e. the transition region which separates the occupied and empty states, is of the order k_bT. The latter is only 26 meV at room temperature.

As a result, the only states expected to be occupied are those near the band gap (i.e. the Γ-point for GaAs). The conduction and valence bands in bulk can as a first approximation be described by parabolas near the band edge, in which the curvature is determined by a material constant m^*. This so-called effective mass will prove to be an essential factor in explaining the electrical properties, the lower electron mass of GaAs as compared to Si being the reason for the superior electrical properties of the former material. This subject will be treated in more detail in the next section and chapter 4.

1.1.2 - Application to heterostructures

Nowadays MBE and related technologies allow the growth of compositionally varying structures with monolayer accuracy. Relevant material systems include the binary $Si_{1-x}Ge_x$, and the quaternaries $(Al,In)_xGa_{1-x}As$, $In_x(Ga,Al)_{1-x}As$ or AlInAsP. This report is mainly limited to the quaternary $(Al,In)_xGa_{1-x}As$ and, to a lesser extend, $In_xGa_{1-x}As$ on InP. The model that is going to be introduced in subsequent chapters can however equally well be applied to other material systems.

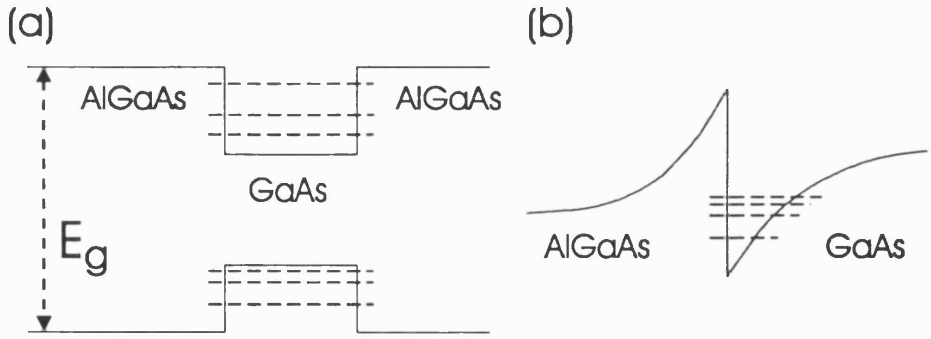


Fig.1.3 : Band diagrams of (a) a quantum well, (b) a heterojunction.

The fact that the energy gap is a material dependent quantity enables the fabrication of quantum wells or heterojunctions in which the electrons (holes) experience a confining potential (fig.1.3) (see for example [Din74,Bas88,Wei91,Sin93]). This confinement forces the momentum in the growth direction to take discrete values causing the formation of multiple confined states. One way of looking at this is by imagining these multiple bands to be cross-sections of the original bulk band at the various quantised momenta (fig.1.4).

This simplified picture of the origin and the form of these multiple bands in a heterostructure works really well for the conduction band. As mentioned in the previous section, the shape of the energy bands in bulk can be approximated by a parabola. From this it follows that the cross sections, corresponding to the multiple subbands in a heterostructure, essentially have the same parabolic dependence. For the valence band the situation is a bit more complicated due to the fact that this band consists of three separate bands, giving rise to quantised states associated with the heavy, light and split-off hole band. The interaction between these states causes these bands to be highly non-parabolic and often anisotropic (fig.1.5).

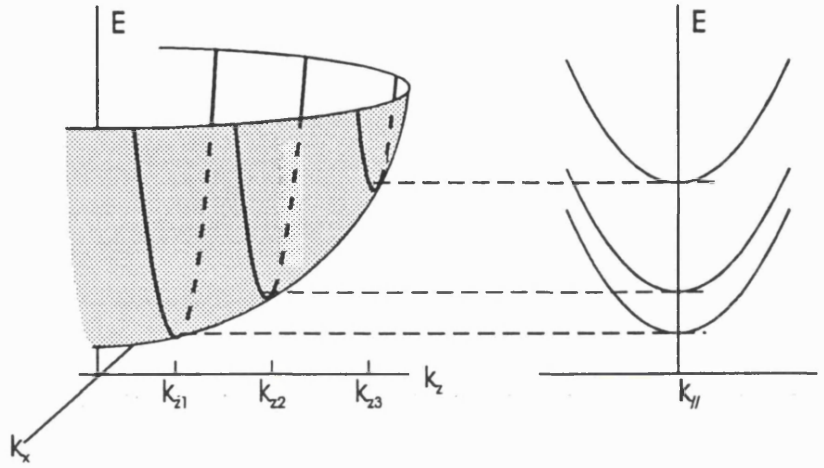


Fig.1.4 : Confinement causes the momentum in the growth direction to take discrete values (k_{z1}, k_{z2},...) which results in the formation of multiple bands.

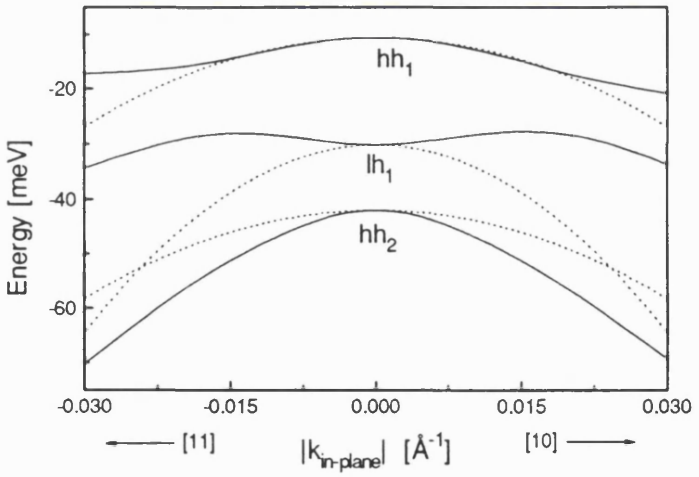


Fig.1.5 :Bandstructure diagram for the first three confined states in a 80Å $Al_{0.33}Ga_{0.67}As$-GaAs QW grown along [001]. Full line: including heavy/light hole interaction; dotted: diagonal approximation (no interaction). The interaction between the heavy and light hole bands cause the bands to be highly non-isotropic as well as non-parabolic.

Although these interactions generally have a negative effect on device performance as they tend to enhance the in-plane effective mass (see next section), it also means that the valence bands are much more susceptible to tailoring. In *n*-type structures the effective mass is essentially set by the particular choice of material whereas in *p*-type the effective mass can be influenced considerably by heterostructure design, choice of growth direction or the introduction of strain.

1.1.3 - The relation between electrical/optical properties and the bandstructure

In the previous section, it was illustrated that the energy versus momentum relation for holes is susceptible to tailoring. At this point it is worthwhile to point out which particular bandstructure and wavefunction characteristics affect the electrical or optical properties of semiconductor heterostructures. As this work is mainly concerned with electrical properties, specific attention is paid to scattering. Here, a qualitative picture of the relation between the above mentioned bandstructural features and ionised impurity scattering is given. The assessment of the importance of scattering for *p*-type field-effect-transistor operation is contained in section 1.3.

Fig.1.6 shows the classical and quasi-classical picture of an electron being scattered from a charged impurity. Obviously, the magnitude of the deflection depends on the distance (or better: the impact parameter *b*) at which the electron passes the impurity. So, knowledge of the form and position of the wavefunction with respect to the position of the impurities can directly provide qualitative information about the scattering rate and the average drift velocity of the carriers under the influence of an electric field. It explains the high performance of Field Effect Transistors utilising modulation-doping to separate the carriers from the impurities [Dae91]

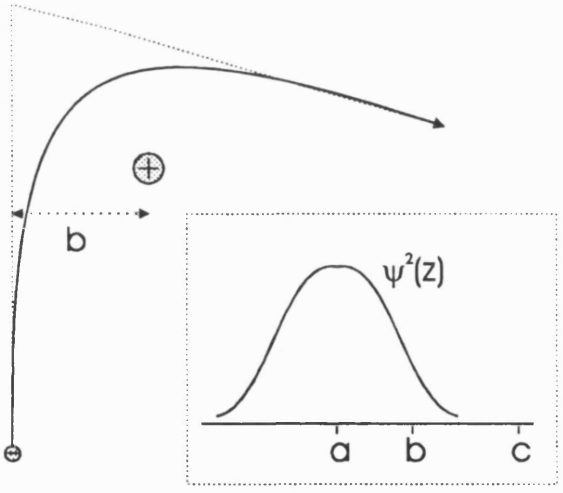


Fig.1.6 : Classical picture of (Rutherford) scattering of an electron of a positively charged impurity. The scattering angle is determined by the square of the electron's momentum and the impact parameter b. The latter is the closest distance at which the electron would pass if there was no electrostatic interaction [Rid93]. Inset : extension to the quantum-mechanical case. Obviously, the scattering time increases when moving impurities from a position with maximum overlap (a) to a position (b) with less overlap. A further improvement in mobility is expected in a modulation doped structure (c).

When evaluating scattering rates, an important quantity is the number of available states that can take part in the scattering process. The more available states, the higher the scattering rate. Scattering in a degenerate system at low temperatures can only take place at the Fermi energy, and the number of states that can participate is essentially governed by the curvature of the particular

band(s) at that (Fermi) energy. This connection between curvature and the number of states that can participate in the scattering process is graphically illustrated in fig.1.7.

The most common parameter to quantify this curvature is through the effective mass m^*. The energy dispersion of the conduction and bulk valence bands is in good approximation given by $E(k) = \hbar^2 k^2 / 2m^*$, where the effective mass m^* depends on the material system as well as the band under consideration. In this case, the number of states that can participate is independent of energy. The above parabolic approximation breaks down in the case of strong band mixing such as present in the valence bands in heterostructures. Here the effective mass is a function of energy and has to be numerically evaluated at the Fermi energy by using the definition [Zaw82]

$$\frac{1}{m^*(E)} = \frac{1}{\hbar^2 k} \frac{\partial E}{\partial k}\bigg|_{E_f} \tag{1.1}$$

Obviously, (1.1) reduces to $m^*(E)=m^*$ for parabolic bands.

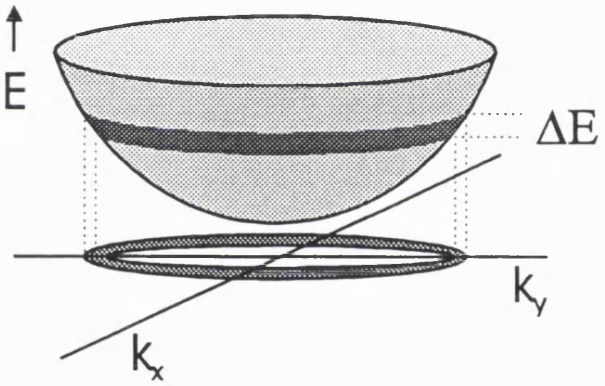


Fig.1.7 : Scattering processes in a degenerate system at low temperature only take place around the Fermi level. Allowing a certain thermal broadening (ΔE) around E_f, one obtains that the number of available states (proportional to the shaded projection in k-space on the x/y-plane) that can take part in a transition is given by the slope (mass) of the energy band around E_f.

From fig.1.7 it follows that a higher curvature, or lower mass, leads to a smaller number of states that can participate in a scattering event. A low mass is therefore desirable for effective charge transport. However, the effective mass does not just affect the speed of the carriers via the scattering rate. Fig.1.8 shows a schematic of the distribution of carriers in *k*-space under the influence of an external electric field. Such a displacement of the Fermi-circle with respect to the Γ-point yields a net momentum and net current. The displaced Fermi-contour may be maintained in a steady state, such that the acceleration of the carriers in the presence of an electric field is compensated for by scattering losses. As the mean increment in velocity in an electric field between two scattering events is given by $qE\tau/m^*$, one finds that the velocity of the carriers has an additional factor $\propto 1/m^*$ which originates from the fact that a light particle accelerates faster than a heavy particle under the application of an identical force.

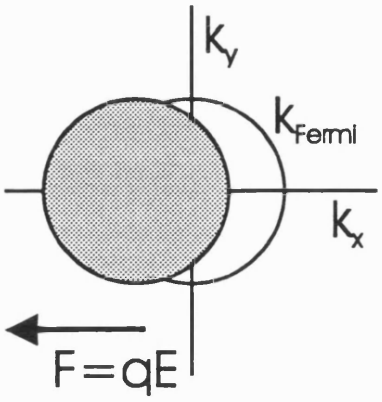


Fig.1.8 : Collective motion of carriers in an electric field. At zero field the Fermi level has its centre at the Γ-point (k=0). In this case, there are as many carriers travelling in the +x direction as in the opposite -x direction so that there is no net current. Under the influence of an external force the contour is displaced in a steady state where the magnitude of the displacement is governed by the scattering time τ [Kit86].

Summarising, the transport velocity of carriers in an electric field is expected to be proportional to $(m^*)^{-2}$. In reality the effects of screening have to be included in the evaluation. As screening reduces the scattering and becomes more effective at higher mass, one finds a dependency of $(m^*)^{-\gamma}$ with γ=1..2 for the steady state carrier velocity.

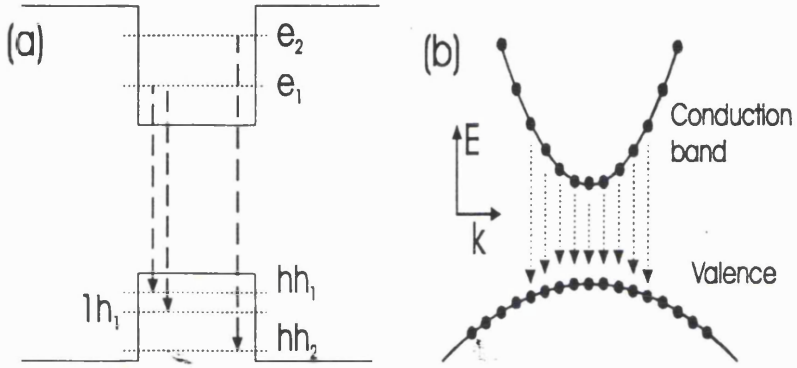


Fig.1.9 :(a) Allowed optical transitions in a fictitious direct band gap semiconductor QW. The $e_1 \rightarrow hh_2$, $e_2 \rightarrow hh_1$ and $e_2 \rightarrow lh_1$ transitions are prohibited due to parity considerations [Wei91] (b) Illustration of 'vertical' transitions in k-space from a single conduction to a single valence band. One has sum over all vertical k-transitions to obtain the emission [Zor93].

The influence of the bandstructure and wavefunctions on the optical properties of a semiconductor heterostructure will now briefly be discussed. Emission (or absorption) of a photon occurs when an electron makes a transition from a certain band to a lower (higher) lying band. The bands considered are generally the conduction and highest heavy hole band. Fig.1.9a shows the allowed optical transitions in a flat band QW structure, which highlights the fact that these transitions are only allowed between states with the same parity. Such an optical transition in a semiconductor is vertical, i.e. occurs between states with the same k vector (fig.1.9b). The intensity of emission at a certain wavelength is then given by the combination of the number of states that can participate in a transition at that wavelength and their population. Obviously, if there are no electrons present in the conduction band then there will be no emission possible. This population is determined by the position of the (quasi-)Fermi level(s) which is a function of the current-injection.

Laser emission principally occurs across the Γ-point. Any emission at a higher energy can therefore essentially be considered as the loss of an electron which now will not contribute to the emission process at the wavelength of interest. One would therefore like to minimise the number of states at energies that do not participate in the emission at the particular wavelength. This again involves tailoring the density of states and effective mass. Bandstructure engineering can therefore offer a reduction in threshold current. A lower density of states can be achieved by using for example strained structures [Sin93] or choosing non-conventional growth directions [Bat89,Men92]. The quest for a density of states in which losses via emission at non-relevant wavelengths are minimised also explains the move from 3D laser structures towards 2D, with a push to go to ever lower dimensions [Wei91,Zor93]. At this point, the connection between wavelength, threshold-current and bandstructure has been established. Furthermore, it was shown that the parity of the wavefunctions determines whether an optical transition is allowed or forbidden. However, the wavefunctions influence optical transitions in more ways than just through parity. Symmetry on atomic scale of the wavefunctions involved in a transition governs the polarisation of the emitted photons. Wavefunction engineering can be used here to tailor the polarisation to the specific needs. One example is the use of [311]-substrates to produce Vertical Cavity SEmiconductor Lasers (VCSELs) that exhibit a clamped polarisation direction [Tat97], in contrast to [001]-grown equivalents which exhibit random and even bi-stable polarisation properties.

1.2 - Modelling semiconductor heterostructures

The use of bandstructure calculations to optimise or analyse semiconductor heterostructures has been pointed out. The theoretical model chosen in this work to perform these bandstructure calculations is described in this section, along with a summary of some alternative approaches and a discussion on why this particular model was favoured. Further, a brief introduction to the boundary condition problem intrinsic to this type of bandstructure model is given.

The previous paragraph already qualitatively described the relation between the bandstructure and electrical as well as optical properties. Some of the techniques that can be used to quantitatively test bandstructure calculations are presented in this section and examples from the literature are given. This section ends with an outline of how bandstructural data can be used to obtain quantitative results for carrier densities and mobilities by combining a bandstructure model with a scattering model. These experimentally verifiable quantities are of high importance when designing transport devices, and are the subject of experimental study in this work.

1.2.1 - The effective mass model

There are various models to calculate wavefunctions and eigenvalues in semiconductor structures. Obviously, any model is intimately related to the Schrödinger equation, which in its original many-body form is too complicated to solve. Rewriting the original Schrödinger equation as a one-electron problem by making several approximations, one finds that the various bulk crystal models vary essentially only in the choice of trial functions. Examples of trial functions [Ash76] are expansions in plane waves (PW), orthogonalised plane waves (OPW) or linear combinations of atomic orbitals (Tight binding or LCAO). Such models are however computationally demanding and introducing compositionally varying structures is only thought to make the calculations more complex. One is therefore inclined to abandon the 'ab initio' route and choose a method that relies on a certain number of parameters which can be obtained by experiments. At this point it is worthwhile to re-iterate that one is essentially only interested in a small region in k-space as generally only a fraction of the Brillouin zone is filled with electrons or holes, and this allows an efficient parametrisation.

Effective mass theory has proved to be a very convenient and successful method to explain the electrical and optical properties of heterostructures. Conduction band eigenvalues and wavefunctions are simply obtained by solving a single differential equation. The situation for the valence band is slightly more complicated due to the fact that this band consists of three separate bands, being the heavy, light and split-off bands. As a result, one finds that valence bands modelling in the effective mass approximation consists of solving a system of coupled differential equations of order four or higher. It proves that this still provides an enormous simplification over exact models, yet giving results that are in good agreement with more elaborate approaches [Sch85a,Sch85b]. Furthermore, the method relies on just a limited number of parameters which are reasonably well known, its results are easy to interpret, it allows for a straightforward extension to strained layer structures and can, using a straightforward algebraic transformation, be applied to different growth directions.

1.2.2 - The boundary condition problem

Effective mass theory, although proven to work extremely well, has long been the subject of debate concerning its application to compositionally varying structures (a review on this topic is contained in [Bur92]). The question of how to connect a wavefunction across an interface between two different materials has always been a major cause of controversy as the original derivation

[Lut55] is strictly not valid near an interface. As a result, this derivation can not supply the applicable 'boundary condition' which therefore has to be obtained using quantum mechanical or other heuristic arguments.

As discussed in the previous section, the conduction band effective mass equation is a single differential equation. The boundary condition problem manifests itself in the kinetic energy contribution to the effective mass equation, containing a combination of a double differential operator and a (position dependent) effective mass. The question is which 'operator ordering' to use, or rather where to put the operators with respect to this mass. Several choices can be excluded by basic quantum mechanical arguments but a unique solution can not be obtained just using simple reasoning. Yet, a sort of general consensus was reached for the conduction band model involving symmetrising that particular Hamiltonian element [Mor84].

The valence band boundary condition problem is however much more complicated, one now has to solve a set of coupled differential equations with multiple occurrences of combinations of operators (single or double differential) and effective mass parameters. Despite the controversy over the 'ad hoc' solution to the conduction band case, the same symmetrisation procedure was commonly used for the valence band Hamiltonian [Men92].

Insight in the boundary condition problem was not rigorously provided until the development of an 'exact' envelope function theory [Bur87,Bur92]. Applied to the valence band case, the boundary conditions obtained for [001]-growth using this model showed considerable differences to those obtained using the symmetrisation procedure [For93].

The investigation in the boundary condition as contained in this work was driven by interest in using non-[001] growth direction for both electrical as well as optical devices (see chapter 6 as well as the conclusions and recommendations). A substantial part of this thesis concerns the derivation of the 'new' boundary conditions for non-[001] growth directions and this is one of the major achievements contained in this work. It resulted in the publication of two theoretical papers, the first outlining a general approach to deriving such boundary conditions for any growth direction [Sta97] and the second providing a set of analytic rules, again applicable to any growth direction, that replace the conventional symmetrisation rule [Dal98]. It was shown that the use of these new boundary conditions has a substantial effect on the calculated bandstructure.

1.2.3 - Self-consistent calculations

In the literature, the majority of valence band calculations concern calculations on flat band quantum wells without charge, i.e. without ionised acceptors/donors or occupied states in the 2DHG. The underlying reason is that inclusion of charge complicates matters considerably in that one now has to *simultaneously* satisfy the Hamiltonian and Poisson's equation [Ste84,Har89]. This is straightforward to see as the Hamiltonian's eigenstates are determined by the confining potential which in turn is affected by the eigenvalues (or rather, their position with respect to the Fermi energy) through Poisson's equation. The solution to the problem is to use a so-called self-consistent procedure in which an iterative approach is used to obtain such a simultaneous solution of both equations. Starting with a trial confining potential function, one makes subsequent changes and improvements in the potential (typically requiring 50 iterations or more) such that the charge distribution as given by Poisson's equation matches that as given by the Hamiltonian. The exact implementation will be discussed in chapter 5.

1.2.4 - Testing the bandstructure model and employing it to improve semiconductor devices

Using the above theoretical model, one can obtain information about the subband populations, the wavefunctions and corresponding bandstructure for a given heterostructure. The next step is to test its predictions and to use the model to improve electro- or electro-optical devices. Some of the techniques that can be used to put bandstructure theory to the test will now be discussed. In addition, some devices of which the development and operation rely on bandstructure and wavefunction engineering will be highlighted.

Optical techniques provide a powerful means to test bandstructure calculations. Standard Photo-Luminescence (PL), or its more sensitive variant PL-E, can give information about the quantisation energies of the various confined states. Here, electron to heavy hole transitions can be distinguished from electron to light hole transitions via a polarisation experiment [Rei90]. Such experiments also show that QW-structures grown on [110] directions exhibit anisotropic optical properties [Ger91]. The latter is caused by heavy, light and split-off hole mixing at the Γ-point and well understood theoretically [Noj92,Noj93,Kaj95]. This effect is present in all but the high-symmetry directions [001] and [111] and provides a means to test theoretical calculations of the symmetry features of the hole eigenstates for various growth directions [Hen95,Tsi95,Win96].

Information about the density of states of the various bands, including occurrences of van Hove singularities [Kem96], can be obtained from the intensity of the peaks in the luminescence data [She95a]. The ultimate optical experiment to test bandstructure calculations was performed by Kash et al. [Kas94]. It was shown that the valence band dispersion of GaAs-AlGaAs QWs, including warping, can be measured with milli-electronvolt accuracy by using a rather complex Photo-Luminescence experiment and analysis technique. Its results were shown to be in excellent agreement with effective mass theory. The application of bandstructure calculations to enhance optical devices has to date led to a push to investigate non-conventional growth directions. For example, theoretical predictions show that lower-threshold currents can be achieved in [111]-structures as compared to conventional [001]-grown lasers [Bat89,Men92]. Vertical cavity semiconductor laser structures (VCSELs) grown on [311] utilise optical anisotropy to control and stabilise the polarisation of the laser beam [Tat97]. Finally, the quantum cascade laser as developed by AT&T-Bell is probably the best illustration of the use of wavefunction engineering in an optical device [Fai94].

Magneto-electric measurements can also give information about the bandstructure. Cyclotron resonance can be used to measure m^* at the Fermi energy so that, by measuring at various carrier densities, one can obtain data about the band curvature at different energies [Liu88]. Measurements on *p*-type structures have qualitatively confirmed the theoretical predictions of the variation of the effective mass as a function of growth direction [Col95,Col96]. Magneto-tunnelling experiments on double barrier structures can be used to map out the bandstructure [Hay94].

Standard magneto-transport measurement techniques such as Hall, van der Pauw or Shubnikov-de Haas measurements do not measure variables that can directly be related to the bandstructure. The outcome of such experiments are quantities like carrier density and quantum- or transport-mobility that can only be obtained by combining the bandstructure model with a scattering model. Work on *n*-type heterostructures has revealed that theoretical quantum mobilities (as measured in a Shubnikov-de Haas experiment) are in good quantitative agreement with experimental results [Hai96] whereas theoretical transport mobilities (Hall, van der Pauw) of the lowest subband generally differ from experimental results by a factor two [Mas91,Gon94,Hai97].

This work will mainly be restricted to the measurement, analysis and simulation of van der Pauw experiments. The main reasons for this is that the mobility in the *p*-type structures of interest is generally too low to give Shubnikov-de Haas oscillations at reasonable magnetic fields.

Furthermore, it is shown in the next section that the relatively poor performance of *p*-type FETs is strongly linked to the low hole mobility in GaAs-related compounds. As this thesis is mainly concerned with the optimisation of high density FETs, it follows that an investigation into methods to enhance the transport mobility at room temperature is required. Despite the expected discrepancy in absolute values between the theoretical and experimental values for the transport mobility, such a combination of a bandstructure and scattering model can provide relatively accurate predictions of the transport mobility when comparing between structures. It can therefore give predictions of the distance between carriers and impurities and so provide a qualitative view of the shape of the wavefunction. Furthermore, it can give vital insight in the scattering processes present (intra- or interband) and, by virtue of the latter, hint at multiple subband population.

The model adopted in this work is a fully self-consistent effective mass model which implicitly includes the interactions between the heavy, light and split-off hole bands when modelling [001] structures. Other growth directions such as [110] and [111] are dealt with using a four band model. Its results have proved to agree with literature results, both for dispersion relations [And85,Iko92,Men92] as well as for wavefunction composition [And87] and corresponding optical properties [Noj93]. The combination of such an effective mass model in conjunction with an RPA scattering model (see chapter 4) has so far, to the author's knowledge, only been applied to the electron case in two-dimensional structures. The results obtained here are also in good agreement with results published in the literature [Hai95,Hai96,Hai97].

1.3 - Field effect transistors

To be able to evaluate the use of the previously described theoretical model to optimise FET-performance, and *p*-type FETs in particular, one has to establish what mechanism limits FET-operation. In this section, basic FET operation is described and some of the parameters that are used in the literature to depict the effectiveness of a FET are presented. Via a simple FET-model and a literature survey, the bottle-neck in *p*-channel FETs is evaluated and the current state-of-the-art in high frequency operation is presented. This section concludes with an explanation of why there is a need for the intrinsically slower (as compared to *n*-type) *p*-type FETs.

1.3.1 - Introduction

Field effect transistors (FETs) come in a variety of types and have a possibly even larger set of acronyms to classify them. All FETs are essentially similar in their operation that a (large) current running across two terminals (so-called source and drain terminals) can be controlled by applying a voltage to a third terminal (the gate) which is electronically insulated from the conducting channel. As a result, they are characterised by a high input impedance as compared to bipolar transistors which are controlled by a gate current rather than a gate voltage.

The difference between the various FET structures is essentially the way the gate is isolated from the current-carrying channel. To name a few [Str90],

- JFET : Junction FET, gate is isolated via a reverse-biased p-n-junction
- MESFET : Metal-Semiconductor FET, via a Schottky barrier
- MISFET : Metal Insulator FET, via an insulating layer.
- MOSFET : Metal-Oxide-FET. This is a subtype of MISFET implemented on Silicon, in which the gate is separated from the active channel by the material's native oxide.

The advent of MBE technology and demand for FETs with high-frequency operation and increased power handling capabilities brought the development of the HFET (Heterostructure FET). Other acronyms that are being used for this type of device include HEMT (High Electron Mobility Transistor), MODFET (MOdulation Doped FET) and HIGFET (Heterostructure Insulated Gate FET) to name but a few. An HFET is essentially a subtype of a MESFET in which the phenomenally high mobility (combined with a high density) that can be achieved in a heterostructure, employing two-dimensional electron (hole) transport, is utilised to improve performance. Fig.1.10 shows a schematic of such a HFET.

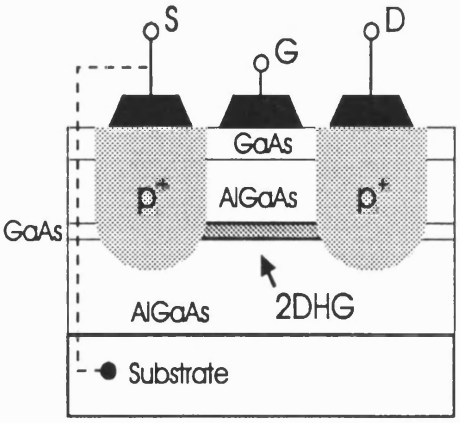


Fig.1.10 : Schematic of a AlGaAs p-type HFET

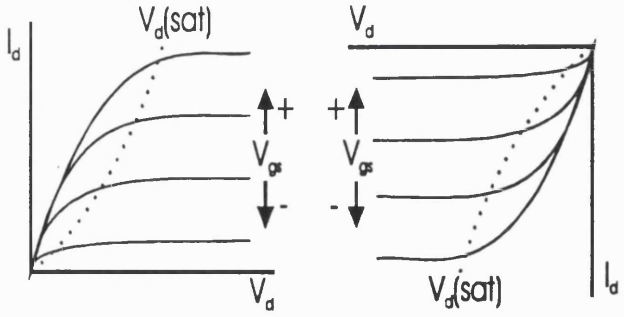


Fig.1.11: Typical output characteristics for (left) an n- and (right) an p-type FET [Sze81].

An important characteristic is the threshold voltage V_t, which is the minimum gate voltage required to induce carriers in the channel. A distinction is made between devices that conduct even at zero gate bias ('normally-on' or 'depletion mode') and devices that require a bias to induce carriers and consequently become conductive ('normally-off' or 'enhancement-mode'). A final point to be made is that this thesis is mainly concerned with heterostructures to be used in depletion mode devices. However, the bandstructure model, the scattering model and the many ideas and principles described in this thesis can equally well be used for enhancement mode devices.

1.3.2 - Basic FET operation, bottleneck for high-speed *p*-FETs

The gate in a FET structure can essentially be pictured as a parallel capacitor which, by varying the voltage, can either accumulate or deplete carriers in the underlying semiconductor and consequently raise or lower the conductivity. For example; setting V_{gs} (the gate voltage wrt. the source) below the threshold voltage V_t in a *p*-type structure, so making the gate voltage more negative than the source by more than the threshold voltage, causes the accumulation of holes in the channel. This results in a rise in conductivity, as reflected in the typical output characteristic shown in fig.1.11

In the simplest capacitor model, the carrier density (either holes or electrons) per unit area under the gate is given by [Wei91]

$$n_{2D} = \frac{C_g}{eL_g W_g}\left(V_{gs} - V_t\right) \tag{1.2}$$

where C_g is the capacitance of the gate contact, L_g is the gate length and W_g is the gate width. As a result, the current flowing through the conducting channel is given by

$$I_{ds} = n_{2D} e v W_g = C_g\left(\frac{v}{L_g}\right)\left(V_{gs} - V_t\right) \tag{1.3}$$

where v is the transport velocity of the current-carrying particle. This velocity is of course dependent on the applied electric field and is related to the transport mobility μ by

$$v = \mu E \tag{1.4}$$

However, the velocity reaches a maximum at the so-called critical field E_c above which the velocity is given by the saturation drift velocity v_s and no longer increases with increasing field via (1.4). The reason for this is that at high electric fields the carriers are accelerated to such energies that that loss mechanisms like phonon emission (*n*-type Si), intervalley scattering (*n*-type GaAs) or impact ionisation (*p*-type Si/GaAs) counters any attempt to increase the velocity further by

increasing the electric field [Sze81]. For electrons in GaAs, $v_s \approx 10^7$ cm/s and $E_c \approx 2.10^3$ V/cm, for holes $v_s \approx 6.10^6$ cm/s and $E_c \approx 6.10^4$ V/cm [Fre96]. From these values it follows that sub-micron electron devices in GaAs are essentially limited by a material constant (the energy separation between the Γ and L valley $E_{L\text{-}\Gamma}$) whereas hole devices are limited by the low-field mobility μ which, via various scattering mechanisms, depends on layer design. Although the latter is generally true, in some QW structures a small correlation has been seen between the low-field mobility and the saturation velocity [Rob96] which has been attributed to the influence of the doping profile on the well shape [Wat98].

To evaluate the efficiency of FETs in terms of material properties, the following characteristics are being used in the literature:

- Intrinsic transconductance g_m

  This quantity reflects the effectiveness by which the gate modulates the current in the conducting channel. It is defined and given by [Wei91]

$$g_m = \frac{\partial I_{ds}^{sat.}}{\partial V_{gs}} = C_g \left(\frac{v}{L_g} \right) \tag{1.5}$$

  In addition, there is also the so-called extrinsic transconductance which includes the parasitic resistance effects between the source-gate contacts. Although the latter is a better reflection of the effectiveness of the device, the intrinsic transconductance is a better measure of the quality of the conducting channel. Both are however closely related.

- Unity gain cut-off frequency (f_T)

  This is defined as the frequency at which the short circuit current gain decreases to unity. It is given by [Fen96],

$$f_T = \frac{g_m}{2\pi C_g} = \frac{1}{2\pi} \left(\frac{v}{L_g} \right) \tag{1.6}$$

Both the transconductance and cut-off frequency depend linearly on the transit time v/L_g of the electrons or holes under the gate, which explains industry's efforts to push gate lengths to smaller dimensions. Alternatively, one also expects to be able to enhance performance by increasing the carrier's drift velocity. The high electron mobility combined with a high density that can be achieved in two-dimensional structures have resulted in high-frequency *n*-type GaAs-based FETs with unity gain frequencies above 100 GHz or power FETs capable of handling over 5W with

more than 10dB gain at frequencies of 12 GHz (see [Suz96] for an overview of state of the art GaAs-based heterostructures).

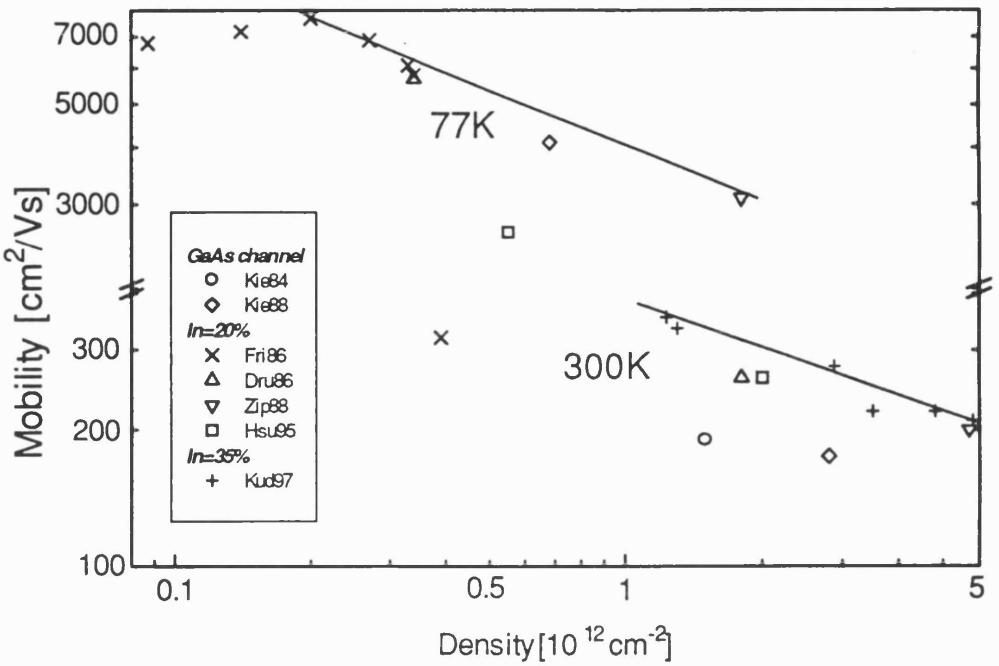


Fig.1.12 : Literature results for the hole mobility vs. carrier density in 2DHGs at both 77 and 300K (the solid lines are just guides to the eye). All structures consist of a modulation doped AlGaAs-InGaAs QWs on GaAs except [Zip88] which is a double QW structure.

The performance of *p*-channel FETs is substantially lower than their *n*-channel counterparts of equal gate lengths. *N*-channel devices are mainly limited by unwanted parasitic capacitance effects and the fundamental limit set by the saturation velocity. In contrast, *p*-type devices operate in the mobility limited regime [Abr97] which, combined with the fact that the valence bands are much more susceptible to band structure tailoring, provides the scope to significantly improve their performance by wavefunction and bandstructure engineering. It has been shown that the hole properties of GaAs-based FETs can be enhanced by using a strained InGaAs channel due to strain-induced splitting of the valence bands [Jaf87]. Recently, it was shown that the indium concentration in pseudomorphic InGaAs on GaAs could be raised above 30%, achieving state of the art room temperature transconductances and mobilities [Kud96,Kud97] (see fig.1.12). The need for *p*-type FETs will now be outlined in the following discussion on complementary circuits.

1.3.3 - Complementary circuits

The poor performance of *p*-type GaAs-based FETs has almost completely prevented their use in complementary circuits, as such a combination of *n*- and *p*-type FETs would be limited by the poorer *p*-type device. At this point, one could ask why one would want to use the intrinsically

slower *p*-type FETs altogether. To answer this, one only has to return to fig.1.11 where it can be seen that there are some fundamental differences in output characteristics between *n*- and *p*-type FETs with respect to the direction of the current flow and the effect of the gate on the magnitude. Combining both type structures can provide unique circuit features which make the integration of *p*-type devices essential in several key areas. Two of these areas are discussed below.

- **VLSI (Very Large Scale Integration)**

Power consumption becomes a key issue when designing electronic circuits that contain several hundred thousand logic gates. Various low-power logic families have been developed that only use *n*-channel devices (for example DCFL) but, in addition to various other intrinsic limitations or problems, are inferior in terms of power consumption to complementary logic [Hai89].

The inverter as pictured in fig.1.13 is a good illustration of why complementary logic is so power efficient. When the input voltage (V_{in}) is low, the bottom (enhancement) *n*-type FET is non-conducting so that there is no current flow through the FETs. In this situation, the output voltage (V_{out}) is high as it is pulled up to the supply voltage since the top *p*-type FET is a in a low-resistive state. A similar situation holds for when V_{in} is high, in which case the top *p*-type FET is non-conducting and V_{out} is pulled down to the ground level as the *n*-type FET is now in a low resistive state. The important conclusion to be drawn is that the circuit does not dissipate *any* power in a static state, the only power dissipation occurs during a switching cycle.

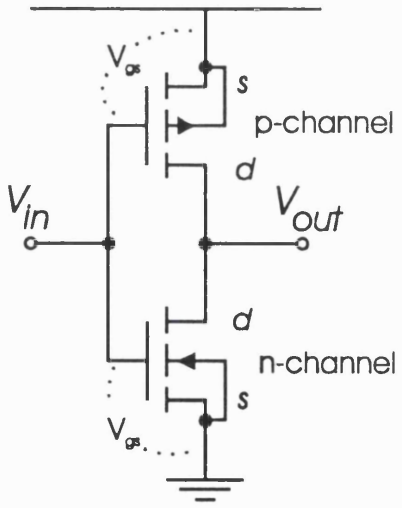


Fig.1.13 : Basic inverter circuit implemented in complementary logic. Both devices are enhancement-mode (normally-off) FETs [Str90].

To put all of this in context, Motorola has recently started fabricating GaAs-based HFET integrated circuits which employ a range of circuit technologies including complementary logic [Abr93,Ber95,Abr97]. Using a 150Å $In_{0.2}As_{0.8}As$ channel, their *p*-type devices exhibit f_t =10GHz

for a 0.5μm gate length. Typically, the complementary logic is used for the lower-frequency high-integration part of the device, other logic families taking care of the high-frequency processes.

- **High power FETs**

A consequence of the poor mobility in *p*-type devices is that such a device can not deliver large amounts of power, as the current through the device is given by the carrier density multiplied by the mobility. Applications that would benefit from an improved *p*-FET performance include power amplifiers with a circuit diagram similar to fig.1.13, delivering relatively large currents rather than just a logic level. Design of such a power HFET involves maximising the carrier density without serious degradation of the mobility.

The latter device is the main motivation for the research described in this thesis. As electrostatics impose a maximum carrier density that can be obtained using modulation doping (see [Hir84]), doping of the active heterostructure layer becomes essential to achieve very high densities. Such doped-channel devices are expected to exhibit a high drain current capability, a large breakdown voltage and a high transconductance [Hon96]. Wavefunction engineering techniques are then used in such channel-doped structures to minimise the degradation of the mobility. Although only ionised impurity scattering and depletion-mode devices with doped channels are considered in this thesis, the techniques presented here are generally applicable. Similar techniques to control the shape and position of the wavefunction can be used to enhance the transport characteristics when other scattering mechanisms are dominant, such as interface roughness scattering in for example SiGe heterostructures.

1.4 - Overview of this thesis

Following this introduction, chapter 2 starts by introducing the theory behind bandstructure calculations in semiconductors. The bandstructure is of course given by Schrödinger's equation which is however too complicated to be solved directly. A substantial amount of space is devoted to the derivation of the chosen bandstructure model, a so-called effective mass or $k \cdot p$ ('k-dot-p') model, as this model will be very important in the analysis of the experimental data. Consecutive simplifications are introduced in the general Schrödinger equation to lower its complexity by reducing it to a one-electron Hamiltonian. A brief introduction to the use of group theory in Zinc-Blende structures is provided. Group theory is then employed to obtain the symmetries of the various eigenfunctions at the band edge, which subsequently allows the construction of the final

effective mass model. The Pikus-Bir formalism to account for strain in such an effective mass model is also presented here, and the effects of strain on the bandstructure are discussed.

Chapter 3 discusses the elementary question of how to connect wavefunctions across an interface in a compositionally varying structure. Insight in this problem was not rigorously provided until recently, and subsequent investigations on [001] revealed marked differences between calculations using the conventional boundary conditions and those using the latter framework. Despite this, it was found that the traditional boundary conditions were still being used in the literature as there was no concise framework to extend the boundary condition evaluation to non-[001] growth directions. The latter will be dealt with in the results chapter, as significant new insight in this problem was obtained during this Ph.D.

The last theoretical chapter (chapter 4) concerns carrier transport in heavily doped structures. An introduction to ionised impurity scattering is given using the easy to interpret Thomas-Fermi approximation, after which a more elaborate Random Phase Approximation (RPA) model is introduced. The latter scattering model, which connects wavefunction and bandstructure to the experimentally verifiable carrier mobility, can provide theoretical predictions for the device performance of ionised impurity scattering limited structures. Its evaluation is slightly more complicated in the valence band than in the conduction band case, due to the presence of strong non-parabolicity and band mixing effects. The former is highlighted in a discussion on the definition of the effective mass parameter, the latter leads to an extra angular dependent factor in the equation for the scattering rate. The latter factor simply reduces to unity in the electron case. Although electron mobility calculations using this framework have been presented in the literature, its application to *p*-type structures is thought to be the first of its kind.

Chapter 5 concerns the numerical implementation of Poisson's equation and the effective mass Hamiltonian. Section 5.1 starts with a discussion on self-consistent methods to find a simultaneous solution to both Poisson's equation and the Hamiltonian. Numerical implementation of Poisson's equation is shown to be relatively straightforward, in sharp contrast to the valence band effective mass model. Solving for valence band eigenvalues and wavefunctions consists of finding the solutions of a set of coupled differential equations, which proves to be time consuming when dealt with using standard numerical routines. A different approach was therefore chosen to solve the Hamiltonian, which utilises an expansion method. Here, the solutions are expected to converge to the exact solutions with increasing number of expansion states. Particular attention is paid to convergence speed and advantages and drawbacks of such a method. Section 5.2 contains a

summary of all the experimental parameters that are required in the Hamiltonian, Poisson's equation and the scattering model, such as effective masses and deformation potentials. A parameter of specific importance is the band line-up of the conduction and valence bands in a compositionally varying structure, for which analytic expressions are given in terms of the band gap and offset-ratios.

Next is the results chapter (chapter 6). The derivation of non-[001] Hamiltonians in this thesis is unique in the sense that it is not based on the customary fully group-theoretical approach which, due to the nature of its derivation, is incapable of supplying information about the boundary conditions. As a result, new theoretical insight into the boundary condition problem for non-[001] growth was obtained. Analytic rules valid for all growth directions that are to replace the conventional symmetrisation rules for the operator ordering are given, which are regarded as one of the key findings in this work.

The second part of this chapter concerns the application of the theoretical model to calculate the carrier mobility in semiconductor heterostructures. First, the model is applied to *n*-channel structures, where the theoretical values for the carrier mobility are compared to both theoretical results published in the literature as well as experimental values. It is shown that a scaling factor of order 2 has to be introduced to bring theory and experiment into agreement, independent on the design of the structure or even the material system used. Thus, it proves that such a model can even provide reasonably accurate absolute predictions, simply by accepting the presence of such a scaling factor.

The attention is then moved to *p*-type structures. A set of δ-doped 60Å $In_{0.15}Ga_{0.85}As$-$Al_{0.33}Ga_{0.67}As$ hole structures was designed to test the scope of wavefunction engineering to optimise the hole mobility in high density heterostructures. The effects of diffusion and segregation on the theoretical hole mobility is evaluated, after which a comparison between the ionised impurity scattering limited theoretical mobilities and the experimental hole mobilities is presented. After all numerical effort involved in such a calculation, it is disappointing to find that a scaling factor of approximately 8 is required to bring theory and experiment into agreement. Although the presence of such a scaling factor was expected from the *n*-type evaluation, it would have been more satisfactory to find a similar scaling factor to that obtained for electron mobility calculations.

The results chapter then ends with a discussion on the origin of this scaling factor. A literature survey provided evidence that the scaling factor in the electron mobility calculations is caused by

electron-electron interaction. It is then postulated that the presence of the large scaling factor in the hole mobility evaluation is most likely not solely caused by hole-hole interactions, but by a combination of the latter and the occurrence of quantum confinement effects in the screening charge distribution. A first attempt to include these quantum confinement effects reveal an improvement in the overall agreement between theory and experiment and a new scaling factor of 4.5 that is in better agreement with the electron result.

Finally, the main achievements and conclusions of this work are discussed in chapter 7, as well as suggestions for further work for a broad range of subjects and fields.

Chapter 2 : Theory

The use of bandstructure calculations to predict or even enhance device performance was outlined in the previous chapter. It was shown that the low hole mobility in GaAs-related compounds is the limiting factor in *p*-type FETs. The mobility was proved to be strongly dependent on the effective mass which in turn is directly linked to the *E(k)* dispersion of the appropriate band. The position of the wavefunctions with respect to doping impurities was also found to be an important factor when evaluating the mobility in heavily doped structures. All this proves that bandstructure and wavefunction engineering can be a powerful means to enhance performance.

This chapter starts with a section on bandstructure calculations in solids. The general approach to perform these calculations in bulk materials is introduced, without any a priori constraint on whether one is dealing with a semiconductor, conductor or an insulator. Several approximations are introduced that render the original Schrödinger equation into a more manageable equivalent, which is unfortunately still found to be extremely difficult to solve numerically.

The next section explains how group theory can be used to obtain information about the degeneracy and symmetry properties of the bulk zone-centre (i.e. at the Γ-point) wavefunctions. A short introduction to group theory is included, but the majority of the section is devoted to the direct application of this theory to Zinc-Blende structures. Explicit forms of the valence band zone-centre solutions are given where the effects of spin-orbit coupling have been taken into account. These zone-centre wavefunctions will prove to be important in the derivation of the valence band effective mass model that is going to be introduced in section 2.3.

The actual derivation of the effective mass model is performed in several stages. Starting with the simplest valence band effective mass model that is applicable to bulk only, consecutive enhancements are made. The generalisation to compositionally varying structures is performed, the effects of choosing a non-conventional growth direction and the inclusion of strain are discussed. Although this chapter contains an extensive review of effective mass theory, the intrinsic problem of how to connect a wavefunction across an interface in a compositionally varying structure is, because of its complexity, dealt with separately in the next chapter. The subsequent use of this bandstructure model alongside a scattering model to calculate transport mobilities will be the subject of chapter 4.

2.1 - Bandstructure calculations in solids

Any bandstructure model has to be linked to the Schrödinger equation. The original Schrödinger equation is unfortunately far too complicated to solve directly, so that various simplifications and approximations need to be introduced to obtain a more manageable system. This section can be seen as a preparation for the sections 2.2 and 2.3 where respectively the general form of the valence band orbitals is derived and a bandstructure model is set up that is going to be used throughout this thesis. This section presents the customary approximations that are required to obtain a one-electron Schrödinger equation, introduces spin-orbit coupling and discusses the implications of the Bloch theorem.

2.1.1 - The one-electron Schrödinger equation

The energy levels of the electronic states of atoms, molecules and solids (involving n electrons) are given by the time-independent Schrödinger equation,

$$H\psi(\underline{r}_1,\underline{r}_2,\ldots,\underline{r}_n) = E\psi(\underline{r}_1,\underline{r}_2,\ldots,\underline{r}_n) \tag{2.1}$$

where $\underline{r}_i$ is the three dimensional coordinate of the i-th electron. The electrons in the completely filled inner shells are expected to be tightly bound to the nucleus and these electrons will therefore be little affected by outside forces. As a result, ions consisting of a nucleus and its surrounding inner shells can be introduced, which limits the many-body problem to only the outer-shell electrons. Making the additional assumption of having a perfect lattice and neglecting any lattice oscillations, the Hamiltonian takes the form [Bas75,Hae91]

$$H = \sum_{i=1}^{n}\left[-\frac{\hbar^2}{2m_0}\nabla_i^{\,2} - \frac{Ze^2}{4\pi\varepsilon_0}\sum_{\underline{R}}\frac{1}{|\underline{r}_i - \underline{R}|} + \frac{e^2}{8\pi\varepsilon_0}\sum_{j=1,j\neq i}^{n}\frac{1}{|\underline{r}_i - \underline{r}_j|}\right] \tag{2.2}$$

kinetic energy — *ion-potential* — *electron-electron interaction*

where i and j sum over all outer-shell electrons, $\underline{R}$ runs over all lattice sites and Z is the electronic charge of the ion.

The Hamiltonian (2.2) has an explicit dependency on the coordinates of *all* outer-shell electrons involved which renders it completely unsolvable. Instead, one would like to write the electron-electron interaction as a simple potential function, i.e. a potential that only depends on the coordinate of the particular electron under consideration, so that one would obtain a one-electron Schrödinger equation of the form

$$\left[-\frac{\hbar^2}{2m_0}\nabla_i^2 + U(\underline{r}_i)\right]\psi_i(\underline{r}_i) = E\psi_i(\underline{r}_i) \tag{2.3}$$

The easiest way to obtain a Hamiltonian of this form is to adopt the Hartree method [Ash76] in which the many electron wavefunction is written as a product of single electron wavefunctions

$$\psi(\underline{r}_1, \underline{r}_2, \ldots, \underline{r}_n) = \psi_1(\underline{r}_1)\psi_2(\underline{r}_2)\cdot\ldots\cdot\psi_n(\underline{r}_n) \tag{2.4}$$

One then obtains for the potential $U(\underline{r})$ (see, e.g. fig.2.1)

$$U(\underline{r}) = -\frac{Ze^2}{4\pi\varepsilon_0}\left(\sum_{\underline{R}}\frac{1}{|\underline{r}_i - \underline{R}|}\right) + \frac{e^2}{4\pi\varepsilon_0}\left(\int d^3r' \frac{n(\underline{r}')}{|\underline{r} - \underline{r}'|}\right) \tag{2.5}$$

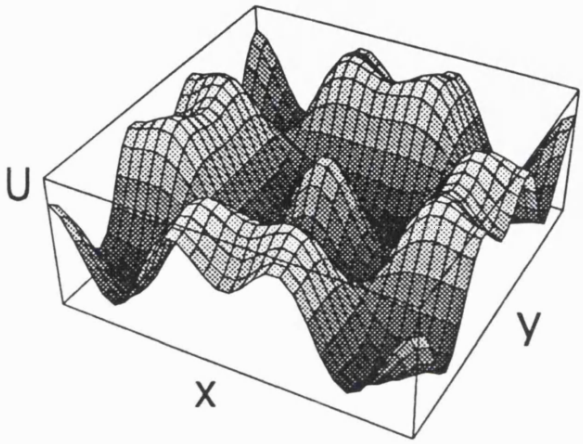


Fig 2.1 : Example of a potential function $U(\underline{r})$. Shown is a so-called pseudopotential for Si, i.e. an empirical potential in which the three adjustable parameters [Bru64,Hae86] were chosen such that some features of the corresponding energy bands agree (to some extent) with the experimental result. Note that the four corners as well as the central peak correspond to a lattice point.

Using this approximation, each electron is effectively put in a static coulomb field associated with a given ion arrangement and electron density $n(r)$. There are however two main objections against such an approach :

- The wavefunction as given by (2.4) does not satisfy the Pauli-principle, it does not implicitly prohibit two electrons from being in the same state. This can be solved by employing the Hartree-Fock method in which the electron wavefunction is written as a Slater determinant.
- The electrons are completely uncorrelated in this approach. As Coulomb repulsion will try to keep two electrons apart, the present approach is expected to overestimate the electron energy.

Including these effects in an exact way would introduce non-local potential terms and would render the approach again into a many-body problem. However, both effects can be included by adding an extra potential term, the so-called exchange and correlation potential $U_{xc}(r)$ for which a parametrised form is available for both the conduction [Hed71] and valence band [Bob96]. The equation as given in (2.3) is therefore the starting point of the bandstructure model that is going to be introduced in a later section.

2.1.2 - Spin-orbit interaction

The fact that an electron carries spin has completely been ignored in the previous discussion. The electron spin interacts with its own angular momentum and this gives rise to a correction, or rather an additional potential term in the Hamiltonian, that has to be taken into account to obtain the correct electron energies. The interaction can be qualitatively understood by looking at a classical system in which a spinning electron moves around a charged nucleus [Gas74]. As the nucleus is moving in the electron rest frame, it follows that the electron is enclosed by a current flow and therefore sees an external magnetic field (fig.2.2).

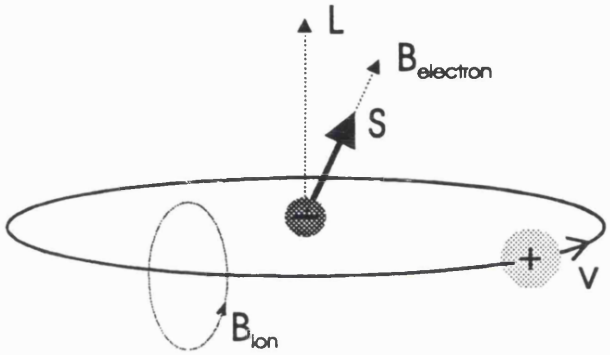


Fig. 2.2 : Classical picture of spin-orbit interaction as seen from the electron's rest frame. Here, the ion moves around the electron with velocity v, thereby generating a magnetic field B_{ion} that interacts with the electron spin.

This magnetic field interacts with the spin of the electron (or more precisely its magnetic moment), giving rise to an energy perturbation proportional to $B_{ion} \cdot S \propto (v \times E) \cdot S \propto (p \times \nabla U) \cdot S$. The correct form of the spin-orbit interaction is given by [Ell54]

$$H_{so} = \frac{1}{2m_0^2 c^2}\left[\nabla U(\underline{r}) \times \underline{p}\right] \cdot S \tag{2.6}$$

where S is the spin operator. It will prove advantageous to write the spin-orbit interaction H_{so} in the following form, assuming that the potential U is spherically symmetric $U(\underline{r})=U(|\underline{r}|)$,

$$H_{so} = \frac{1}{2m_0^2 c^2}\left(\frac{1}{r}\frac{\partial U}{\partial r}\right) L \cdot S \tag{2.7}$$

in which L is the angular momentum operator. As the spin-orbit coupling is considered a perturbation, the used assumption can be thought of as the zero-th order correction [Lut55] or tight-binding limit [Kan56].

2.1.3 - Application to bulk: the Bloch theorem

Although this thesis is mainly concerned with bandstructure calculations in compositionally varying structures, it will prove useful to have a look at the bulk case first. One can then apply the Bloch theorem that states that the eigenfunctions of (2.3) obey [Kit86]

$$\psi(r+R) = e^{ikr}\psi(r) \tag{2.8}$$

As a logical consequence of (2.8) one finds that the probability density given by $|\psi(r)|^2$ follows the underlying Brillouin zones, something which one intuitively would have expected. An alternative way of formulating Bloch's theorem is to say that any eigenfunction can be written as

$$\psi(r) = u_{nk}(r)e^{ikr} \tag{2.9}$$

Rewriting the one-electron Schrödinger equation including spin-orbit interaction (2.3,2.6) in terms of the cell-periodic functions $u_{nk}(r)$, $u_{nk}(r+R) = u_{nk}(r)$, one obtains

$$\begin{aligned} H_k u_{nk}(\underline{r}) &= \left(\underbrace{\frac{p^2}{2m_0}}_{H_0} + U(\underline{r}) + \underbrace{\frac{\hbar}{4m_0^2c^2}\left[\nabla U \times \left(\underline{p} + \hbar\underline{k}\right)\right]\cdot\underline{\sigma}}_{H_{so}} + \underbrace{\frac{\hbar}{m_0}\underline{k}\cdot\underline{p}}_{H_{k\cdot p}} \right) u_{nk}(\underline{r}) \\ &= \left(E_n - \frac{\hbar^2k^2}{2m_0} \right) u_{nk}(\underline{r}) \end{aligned} \tag{2.10}$$

The k-dependent factor in the spin-orbit term will be neglected in this work as the crystal momentum $\hbar k$ is much smaller then the atomic momentum p.

The advantage of (2.10) over (2.3) is that $u_{nk}(r)$ only has to be solved within just one Wigner-Seitz primitive cell, in contrast to $\psi(r)$ which has to be solved for the whole crystal. The application of the Bloch theorem has added two more elements to the Hamiltonian, the so-called ***k·p*** ('k-dot-p') term as well as a plain energy term with the usual k^2 dependence. Even without actually solving for the eigenvalues, one can make some interesting observations. Ignoring the k-dependence in H_k, one finds parabolic energy bands with energy

$$E_n(k) = \frac{\hbar^2k^2}{2m_0} + E_n^0 \tag{2.11}$$

i.e. one obtains the familiar empty lattice results. A closer look reveals that the above result also applies when one treats the k-dependent part as a perturbation and uses for $u_{nk}(r)$ the solutions of the unperturbed problem $(H_0 + H_{so})u_{nk} = (E - \hbar^2k^2/2m_0)u_{nk}$. From these observations it follows that

- The occurrence of electrons with effective mass $m^* \ll m_0$ as well as the occurrence of negative hole mass is solely governed by the $\boldsymbol{k}\cdot\boldsymbol{p}$ interaction.
- This $\boldsymbol{k}\cdot\boldsymbol{p}$ interaction causes the wavefunctions at finite k to deviate from the band edge solutions $u_{n0}(r)$.

At this point it is discouraging to find that, even after having made a considerable number of approximations, the bulk one-electron Schrödinger equation is still extremely difficult to solve. This is basically due to the complex form of the potential term. Several methods have been suggested to solve (2.10) [Ash76] which differ essentially in the choice of trial functions (the basis functions). There are now however two possible ways to continue, either solve (2.10) from first principles (*'ab initio'*) or choose a method that relies on experimental parameters. As the method eventually will be employed to solve for compositionally varying structures in which the computational problem is expected to be even more complex than for the bulk case, one is inclined to abandon the ab-initio route and decide for a method that relies on material constants. For both cases, group theory can substantially simplify the problem, in the first case by reducing the numerical work necessary and the second case by reducing the number of constants. Symmetry and group theory will be discussed in the next section.

2.2 - Symmetry considerations

The presence of symmetry in a physical problem is expected to reduce the complexity. In the case where one is interested in solving for the cell-periodic functions $u_{nk}(r)$ as described in the previous section, one expects that symmetry will put constraints on the general form of these cell-periodic functions. Identification of symmetry operations and the analysis of the effects on the particular physical problem is generally done within a very abstract and mathematical framework that is denoted by group theory. Although the underlying mathematics of group theory are very extensive, actual application of this theory to a physical problem is rather straightforward. In this section, a brief introduction to group theory is given, with the emphasis on the application to systems containing the symmetry of a Zinc-Blende lattice. Explicit forms for the zone-centre wavefunctions are obtained, first without and later including spin-orbit coupling.

2.2.1 - Group theory and its implications

Group theory is an extremely powerful method to obtain information like degeneracy of eigenlevels and the general form of eigenfunctions by looking at the symmetry of the problem involved. That such a link exist is not too surprising as the wavefunctions in a perfectly ordered material are not expected to be of arbitrary shape. Group theory is set up in such a way that this sort of information can be retrieved by just using character tables and symmetry groups (which will both be treated later on) without actually needing to solve the Hamiltonian.

This section is mainly concerned with outlining the fundamentals of group theory and linking it to the physical problem of solving the Hamiltonian for eigenvalues and wavefunctions. It is emphasised that this section does not contain a complete review of group theory but is rather more focused on its application. More information about the specific use of group theory in electron band calculations can for example be found in [Cor69,Bas75,Alt94].

Starting with the definition of a group, a group $\mathbf{G}$ is defined as a set of elements $A,B,C,...$ where

- any product is itself an element of $\mathbf{G}$
- $(AB)C=A(BC)$, i.e. the elements are associative
- contains the identity element E, $AE=EA$
- for any element there is an inverse which is also an element of $\mathbf{G}$.

An example of a group is the set of transformations which leave the crystal lattice invariant. Following [Cor69], a transformation is denoted by T and its effect on a vector $\underline{r}$ is given by

$$\underline{r}' = \{R(T)|t(T)\}\underline{r} \quad or\ shorthand \quad \underline{r}' = T\underline{r} \tag{2.12}$$

where $\underline{r}$ and $\underline{r}'$ describe the same coordinate but in a different reference frame. Here, the new coordinate is obtained by a rotation $R(T)$ followed by a translation $t(T)$. Some examples of a rotation, or point operation, $R(T)$ are :

- C_{ni} : a proper rotation through $2\pi/n$ in the right hand sense about the i-axis
- E : identity operation
- σ : a reflection
- I : an inversion

Every rotation $R(T)$ is partnered by an operator $P(T)$. When followed by a scalar function $f(\underline{r})$, $P(T)$ performs the coordinate change given by $R(T)^{-1}$. An example of such an operator is : 'replace (x,y,z) by $(-x,y,z)$'. In a rather more mathematical notation,

$$P(T)f(r') = f(r)$$

or (2.13)

$$P(T)f(Tr) = f(r).$$

The question now is to establish the set of symmetry operations that leave the eigenvalue equation (2.2) or its one-electron equivalent (2.3) invariant. Obviously, the potential $U(\underline{r})$ in the Schrödinger equation obeys $U(\{R(T)/t(T)\}\underline{r}) = U(\underline{r})$ if T is chosen such that it leaves the crystal invariant, and so do the other Hamiltonian elements. The group of the Schrödinger equation is therefore the space group of the crystal, i.e. the symmetry transformations that leave the lattice invariant also leave the Hamiltonian invariant,

$$H(T\underline{r}) = H(\underline{r}) \tag{2.14}$$

The Bloch theorem (2.9) that was introduced in sub-section 2.1.3 which states that every bulk wavefunction in a periodic structure can be written as a cell-periodic function u_{nk} multiplied by a plane wave is a direct consequence of the presence of translation symmetry. Translations symmetry is however not the only symmetry present in common crystalline structures but merely forms a subgroup of the total symmetry group $\mathfrak{G}$. In addition, the point group $\mathfrak{G}_0$ which contains all the remaining transformations $\{R(T)/\underline{0}\}$ is also a subgroup of $\mathfrak{G}$.

The consequences of the additional symmetry operations contained in $\mathfrak{G}_0$ are now investigated. Obviously, the presence of rotational or inversion symmetry puts constraints on the general form of the cell-periodic structure. To obtain these constraints, a link between abstract group theory and the physical problem of solving the Hamiltonian for eigenfunctions and eigenvalues has to be made. First, this involves a more detailed look on the group theory without explicitly dealing with a particular problem (being an eigenvalue problem or otherwise). Secondly, an examination into the effects of the presence of a symmetry transformation on the physical problem has to be made, and finally both need to be connected.

The group concept was introduced in the previous paragraphs, and the elements of the (sub)groups were visualised in terms of rotations and translations. It now proves convenient to work in a more abstract way and represent these group elements by matrices which, naturally, have to satisfy all group postulates. It is straightforward to see that there exists an infinite number of representations as there is no constraint on the order of the matrices that one can choose. However, all these representations can be constructed from a number of basis (so-called irreducible) representations. These irreducible representations consist of a set of matrices $\Gamma(T)$, one for each transformation, of the lowest possible order. As $\Gamma(T)$ is only uniquely determined within a similarity transformation,

one would like to avoid working with an explicit form of the matrices. Instead, an invariant quantity called the character of T, which is the trace of the matrix $\Gamma(T)$ and defined by $\chi(T) = \sum_i \Gamma(T)_{ii}$, is used throughout group theory.

Returning to the physical problem of solving the Hamiltonian, it can be shown that the Hamiltonian H and transformation operator $P(T)$ commute when T corresponds to a symmetry transformation of the lattice. From this it follows that

$$\begin{aligned} & H(r)\ \ \phi_n(r) && = E\ \ \phi_n(r) \\ \Rightarrow\ & H(r)\{P(T)\phi_n(r)\} && = E\{P(T)\phi_n(r)\} \end{aligned} \tag{2.15}$$

which says that *when ϕ_n is an eigenfunction at energy E, then $P(T)\ \phi_n$ is also an eigenfunction with the same eigenvalue*, a very important observation. Now consider an l-fold degenerate level of the Schrödinger equation, $H(\underline{r})\phi_n(\underline{r})=E\phi_n(\underline{r})$, $n=1..l$, for which the set ϕ_n forms a set of linearly independent eigenfunctions corresponding to this eigenvalue. It then follows that $P(T)\ \phi_n$ must be a linear combination of this set so that one can write

$$P(T)\phi_n(\underline{r}) = \sum_{m=1}^{l} \Gamma(T)_{mn}\, \phi_m(\underline{r}) \tag{2.16}$$

In words, ϕ_n transforms under an operation $P(T)$ as the n^{th} row of $\Gamma(T)$. This suggests that *a set of basis functions for a l-degenerate level forms a set of basis functions for an l-degenerate representation of this subgroup of the Schrödinger equation.* The proof consists in showing that all group postulates are satisfied. But this means that a subgroup Γ of order l corresponds to an l-degenerate energy level. The connection between group theory and the eigenvalue problem has now been made.

The latter observation has huge implications. Given a system, being a molecule or a crystal, it is not difficult to obtain its symmetry group. Group theory in turn, by various routines with which this work does not concern itself, can supply the appropriate character table. Using only this table and the symmetry operations, one can extract elementary information about the degeneracy of the various eigenlevels as well as the symmetry properties of the eigenfunctions. The latter can be obtained using the projection operator $\wp^\Gamma$ which projects any function onto the group Γ

$$\wp^\Gamma = \frac{l_\Gamma}{g} \sum_T \chi^*_\Gamma(T) P(T) \tag{2.17}$$

where T runs over all transformations, l_Γ is the number of basisfunctions in Γ and g is the total number of symmetry operations.

2.2.2 -Application to Zinc-blende

The previous discussion showed the connection between abstract group theory and the eigenvalue problem in periodic structures. The theory will now be applied to obtain information about the symmetry of the eigenstates of the band edge (i.e. at the Γ-point) solutions for the bulk-problem (2.3) in which the potential $U(r)$ has the symmetry of the Zinc-Blende lattice. There are 24 symmetry operations divided into 5 classes (Γ_1, Γ_2, Γ_{12}, Γ_{15}, Γ_{25}), which respectively contain the identity transformation (E), rotations over $\pi/2$ followed by inversion (C_4^2), rotations over π followed by inversion (IC_4), rotations over π followed again by inversion (IC_2) and rotations over $2\pi/3$ (C_3). The symmetry operations corresponding to the classes Γ_1, Γ_{15} and Γ_{12} are given in table 2.1 [Bas66].

Class	*Elements*	*P(T)*
E	E	(x,y,z)

Class	*Elements*	*P(T)*
IC_2	IC_{2xy}	(-y,-x,z)
	IC_{2xz}	(-z,y,-x)
	IC_{2yz}	(x,-z,-y)
	$IC_{2x,-y}$	(y,x,z)
	$IC_{2x,-z}$	(z,y,x)
	$IC_{2y,-z}$	(x,z,y)

Class	*Elements*	*P(T)*
IC_4	IC_{4z}^{-1}	(y,-x,-z)
	IC_{4z}	(-y,x,-z)
	IC_{4x}^{-1}	(-x,z,-y)
	IC_{4x}	(-x,-z,y)
	IC_{4y}^{-1}	(-z,-y,x)
	IC_{4y}	(z,-y,-x)

Table 2.1 a,b,c : Symmetry operations corresponding to Γ_1, Γ_{15} and Γ_{12}. The column P(T) gives the coordinate transformation of a point (x,y,z) where the origin is set to a lattice point (i.e. Ga or As site).

The character table corresponding to the zinc-blende lattice is given in table 2.2. Re-iterating that the characters are simply the trace of the transformation matrix that describes the effect of a lattice transformation on the eigenfunctions, it follows directly that the degeneracy of a level corresponding to a particular symmetry group can directly be obtained from $\chi(E)$ in the character table (printed in italic). For example, the eigenlevel corresponding to Γ_{15} is threefold degenerate.

	E	*$3C_4^2$*	*$6IC_2$*	*$6IC_4$*	*$8C_3$*
Γ_1	*1*	1	1	1	1
Γ_2	*1*	1	-1	-1	1
Γ_{12}	2	2	0	0	-1
Γ_{15}	*3*	-1	1	-1	0
Γ_{25}	*3*	-i	-i	i	0

Table 2.2 : Character table corresponding to a Zinc-blende lattice [Bas66].

The customary method to establish the symmetry of the eigenfunctions of a particular level is by projecting a set of trial functions onto the appropriate group using (2.17), the lattice symmetry operations, and the character table. Alternatively, one can use the explicit form of the irreducible matrix representations to generate these symmetry properties via (2.16). These explicit matrix representations are however generally not provided, but the forms for Zinc-Blende can for example be found in [Bas75]. The latter method will now be used to obtain the symmetries and to explain and clarify the underlying physics.

The irreducible representations of Γ_{15} for the transformations IC_{2xy} and IC_{2xz} are given by

$$\Gamma_{15}(IC_{2xy}) = \begin{pmatrix} 0 & -1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad , \quad \Gamma_{15}(IC_{2xz}) = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{pmatrix} \quad etc \tag{2.18}$$

It is easy to see that three functions that transform as x,y and z satisfy (2.16) and (2.18). Further investigation also reveals that the atomic orbitals d_{zx}, d_{xy} and d_{yz}, also possess the Γ_{15} symmetry. Similarly, as the irreducible representations of Γ_{12} are given by

$$\Gamma_{12}(IC_{4z}^{-1}) = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \quad , \quad \Gamma_{12}(IC_{4x}^{-1}) = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2}\sqrt{3} \\ -\frac{1}{2}\sqrt{3} & -\frac{1}{2} \end{pmatrix} \quad etc \tag{2.19}$$

it follows that the lowest orbitals that transform as Γ_{12} are of the form $\left(x^2 - y^2\right)$ and $\left(2z^2 - x^2 - y^2\right)/\sqrt{3}$, corresponding to the d-orbitals $d_{x^2-y^2}$ and $d_{3z^2-r^2}$.

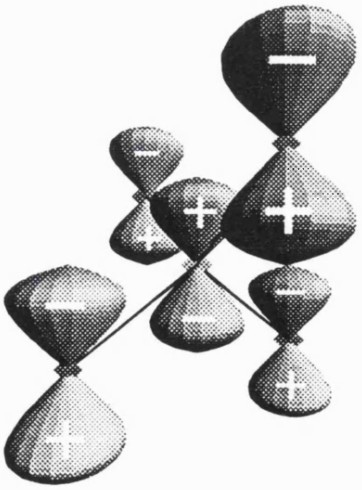


Fig.2.3 : Bonding p_z function in Zinc-Blende. A primitive cell contains the centre (Ga) atom and 1/4-th of each surrounding (As) atom. In an anti-bonding p_z function the sign of the surrounding As wavefunctions is inverted.

Although the previously derived atomic orbitals form a basis of the various point groups of a zinc-blende lattice, one might ask how the previous atomic orbital picture ties in with such a complicated lattice. Since a primitive cell contains not one but two atoms (i.e. one Ga and one As), a single atomic orbital is not a proper description of the zone-centre wavefunction. One can however form 'bonding' (+) and 'anti-bonding' (-) orbitals [Ell54,Kan66] (see fig.2.3) which have the same transformation properties.

At this point it is useful to put the previous findings in context by looking at the implications when setting up a bandstructure model. Following a tight-binding approach, one can choose to develop the cell-periodic functions $u_{nk}(r)$ in a set of atomic orbitals. Simply limiting the expansion to the *s*-, *p*- and *d*-orbitals without using group theory would mean that the model would have to minimise the energy of a linear set of nine orbitals. Group theory however proves that, without losing any accuracy, the trial functions can be limited to the following orbitals when evaluating the eigenstates at the Γ-point:

Γ_1 : s^+, s^-

Γ_{15} : p^+, p^-, d'^+, d'^-

Γ_{12} : d^+, d^-

where d holds the $d_{x^2-y^2}$ and $d_{3z^2-r^2}$ orbitals and d' holds d_{zx} d_{yz} d_{xy}. Fig.2.4 gives a graphical picture of the symmetry properties of these orbitals.

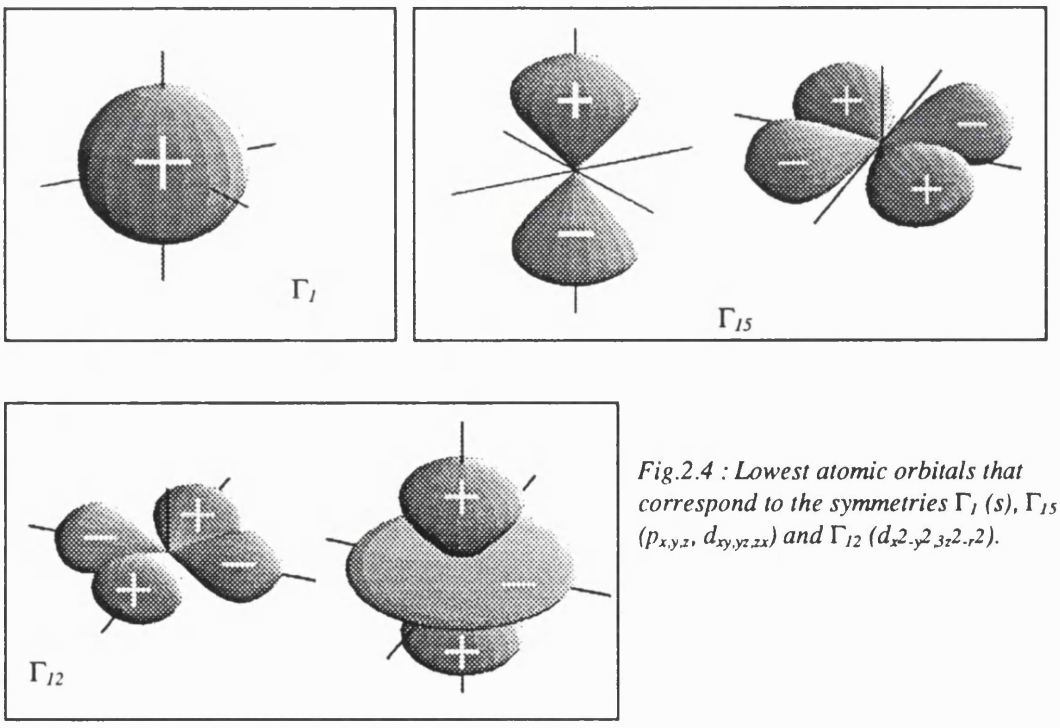


Fig.2.4 : Lowest atomic orbitals that correspond to the symmetries Γ_1 (s), Γ_{15} ($p_{x,y,z}$, $d_{xy,yz,zx}$) and Γ_{12} ($d_{x2-y2,3z2-r2}$).

As was shown in fig.1.2 in chapter 1, the valence band states of GaAs (as well as all other III-V materials) correspond to the Γ_{15} symmetry. Since the tight-binding approximation is expected to be valid for such a closely bound state, the valence band states will be approximated by pure p^+ states. From now, the three valence band eigenfunctions will therefore be denoted by $|X\rangle$, $|Y\rangle$ and $|Z\rangle$.

2.2.3 - Inclusion of spin-orbit coupling

The transformation properties of the zone-centre eigenfunctions without spin-orbit interaction were derived in the previous section. It turns out that the inclusion of this spin-orbit interaction is rather straightforward and does not have to be done in the abstract group-theoretical approach by using so-called 'double-groups' (including an additional spin group) such as done in [Dre55b]. Instead, operator methods are going to be used to set up linear combinations of the previously derived $|X\rangle$, $|Y\rangle$, $|Z\rangle$ states to form eigenfunctions of the Schrödinger Hamiltonian with spin-orbit coupling. One way of doing this is by diagonalising H_{so} in matrix form [Kan56]. There is however a much faster procedure [Sin93] which utilises the fact that spin-orbit interaction in the case of a spherically symmetric potential $U(r)$ is proportional to $L \cdot S$ (2.7). Since the $|X\rangle$,$|Y\rangle$, $|Z\rangle$ functions are already eigenfunctions of L^2 and S^2 and since the total angular momentum operator J satisfies

$$\begin{aligned} J^2 &= (L+S)^2 \\ &= L^2 + 2L \cdot S + S^2 \end{aligned} \tag{2.20}$$

it follows that for the solutions to be eigenfunctions of H_{so}, it is sufficient for them to be eigenfunctions of J^2. This observation will turn out to be very important later on when it proves beneficial to orientate the spin along different crystallographic axis. This renders a diagonalisation procedure into a time-consuming and laborious problem. As J by itself can not classify the states uniquely, one can introduce a new commuting observable which is chosen to be J_α , the angular momentum operator for a particular direction α.

The linear combination of p-orbitals that forms an eigenfunction of the band edge Hamiltonian including spin-orbit interaction will from now on be denoted by

$$\left|J,m_j\right\rangle = \left|\chi_1 X + \chi_2 Y + \chi_3 Z\right\rangle \uparrow + \left|\chi_4 X + \chi_5 Y + \chi_6 Z\right\rangle \downarrow \tag{2.21}$$

where m_j is the quantum number corresponding to J_α and $\chi_{1..6}$ are to be determined constants. The freedom to choose a particular quantisation direction will be explored next, and analytical expressions for the $|J,m_j\rangle$ states will be given in sub-section 2.2.5.

2.2.4 - Choosing a preferred quantisation direction for the angular momentum

A $|J,m_j\rangle$ state such as given in equation (2.21) corresponds to a valence band eigenfunction at the Γ-point. Restricting these states to the heavy (J=3/2, m_j=±3/2), light (J=3/2, m_j=±1/2) and split-off (J=1/2, m_j=±1/2) holes gives a total of six states. Clearly, choosing a particular basis set $|J,m_j\rangle$ that corresponds to a particular choice of quantisation axis would not, when using this set in a bandstructure model, affect characteristics like accuracy or completeness. Re-iterating that a set $|J,m_j\rangle$ at present is only known to correspond to the band edge solutions (k=0), one can ask the question if the freedom in the quantisation direction α can be utilised to make these states resemble the exact wavefunctions as closely as possible in the case where $k\neq 0$. The latter would obviously be of great benefit as it would reduce the numerical problem. It turns out the set can be chosen such that it forms the exact wavefunctions even at finite momentum. This can be done quite simply by choosing the quantisation axis for the angular momentum along the same direction as that of the direction of the momentum one is interested in. Or stated differently, when mapping out $E(\underline{k})$ with $\underline{k}$ parallel to the direction α, it is easiest to work with a basis set in which the angular momentum (and spin) are orientated along the same direction α. This can easily be proved for the case where the momentum is chosen along the z-direction, $\underline{k}=(0,0,k_z)$; other directions can be checked in the same way.

For this case, it is straightforward to prove that the $k\cdot p$ interaction term $H_{k\cdot p}$ commutes with J_{001}

$$\begin{aligned}\left[H_{k\cdot p}, J_{001}\right] &\propto \left[k\cdot p\ ,\ L_z + S_z\right] \\ &\propto \left[p_z\ ,\ xp_y - yp_x + S_z\right] \\ &= 0\end{aligned} \tag{2.22}$$

As a result, it follows that the basis states obtained by orientating the angular momentum along α are not just eigenfunctions of the band edge Hamiltonian but are also eigenfunctions for the finite-k case as long as the momentum is restricted to the direction α.

2.2.5 - Explicit form of the zone centre wavefunctions

In the tight binding approximation the valence band states are of the form $|X\rangle = R(r)x, |Y\rangle = R(r)y, |Z\rangle = R(r)z$. Rewriting these states as spherical harmonics Y_{lm}, it is easy to prove using the operators L_z, L_+, L_- [Gas74] that operating the angular momentum operators $L_{x,y,z}$ on the valence band states has the following effect

$$L_x\begin{cases} X &= 0 \\ Y &= iZ \\ Z &= -iY \end{cases}, \quad L_y\begin{cases} X &= -iZ \\ Y &= 0 \\ Z &= iX \end{cases}, \quad L_z\begin{cases} X &= iY \\ Y &= -iX \\ Z &= 0 \end{cases} \tag{2.23}$$

Similarly, it can be shown that

$$S_x\begin{cases} \uparrow &= \frac{1}{2}\downarrow \\ \downarrow &= \frac{1}{2}\uparrow \end{cases}, \quad S_y\begin{cases} \uparrow &= \frac{1}{2}i\downarrow \\ \downarrow &= -\frac{1}{2}i\uparrow \end{cases}, \quad S_z\begin{cases} \uparrow &= \frac{1}{2}\uparrow \\ \downarrow &= -\frac{1}{2}\downarrow \end{cases} \tag{2.24}$$

Using (2.23) and (2.24) and orientating the angular momentum J along the z-direction, the following set $|J,m_j\rangle$ of band edge eigenfunctions is obtained for the Hamiltonian including spin-orbit interaction,

$$\begin{aligned} \left|\frac{3}{2},+\frac{3}{2}\right\rangle &= \frac{1}{\sqrt{2}}(X+iY)\uparrow \\ \left|\frac{3}{2},+\frac{1}{2}\right\rangle &= -\frac{1}{\sqrt{6}}\left[(X+iY)\downarrow -2Z\uparrow\right] \qquad \left|\frac{1}{2},+\frac{1}{2}\right\rangle = -\frac{1}{\sqrt{3}}\left[(X+iY)\downarrow +Z\uparrow\right] \\ \left|\frac{3}{2},-\frac{1}{2}\right\rangle &= -\frac{1}{\sqrt{6}}\left[(X-iY)\uparrow +2Z\downarrow\right] \qquad \left|\frac{1}{2},-\frac{1}{2}\right\rangle = -\frac{1}{\sqrt{3}}\left[(X-iY)\uparrow -Z\downarrow\right] \\ \left|\frac{3}{2},-\frac{3}{2}\right\rangle &= -\frac{1}{\sqrt{2}}(X-iY)\downarrow \end{aligned} \tag{2.25}$$

As mentioned earlier, the $|3/2,\pm 3/2\rangle$ states correspond to the heavy, $|3/2,\pm 1/2\rangle$ to the light and $|1/2,\pm 1/2\rangle$ to the split off holes. It is important to note that the above set is not uniquely defined as each element can be multiplied with a phase factor $\alpha+i\beta$ of unit length without changing any of the physics. One therefore finds different sets in the literature; compare the above set to, for example, [Lut55]. This issue will be discussed in a subsequent section.

At this point, the choice of orientating the angular momentum J along the z-direction is rather arbitrary as a different 'preferred' direction could just as well have been chosen. It will prove advantageous to choose J along the growth direction when setting up a bandstructure model for compositionally varying structures as will be done in the following section. The above set will therefore be of importance when looking at [100]-growth only. Adopting a Cartesian reference frame with unit vectors $\underline{a}=(a_1,a_2,a_3)$, $\underline{b}=(b_1,b_2,b_3)$, $\underline{c}=(c_1,c_2,c_3)$, the general form of the set $|J,m_j\rangle$ in which the angular momentum is orientated along the vector $\underline{c}$ was presented in [Dal98] and has the form

$$\left|\frac{3}{2},+\frac{3}{2}\right\rangle=\frac{1}{\sqrt{2}}\left|(a_1+ib_1)X+(a_2+ib_2)Y+(a_3+ib_3)Z\right\rangle\uparrow$$
$$\left|\frac{3}{2},+\frac{1}{2}\right\rangle=\frac{2}{\sqrt{6}}\left|c_1X+c_2Y+c_3Z\right\rangle\uparrow-\frac{1}{\sqrt{6}}\left|(a_1+ib_1)X+(a_2+ib_2)Y+(a_3+ib_3)Z\right\rangle\downarrow \quad (2.26)$$
$$\left|\frac{1}{2},+\frac{1}{2}\right\rangle=-\frac{1}{\sqrt{3}}\left|c_1X+c_2Y+c_3Z\right\rangle\uparrow-\frac{1}{\sqrt{3}}\left|(a_1+ib_1)X+(a_2+ib_2)Y+(a_3+ib_3)Z\right\rangle\downarrow$$

The corresponding $|J,-m_j\rangle$ states can be constructed using the Kramers operator [Kan56] on their positive counterparts. Using the general form defined in (2.21), one obtains that the negative counterparts are given by[1]

$$\left|J,-m_j\right\rangle=\left|\chi_4^*X+\chi_5^*Y+\chi_6^*Z\right\rangle\uparrow-\left|\chi_1^*X+\chi_2^*Y+\chi_3^*Z\right\rangle\downarrow \quad (2.27)$$

For the standard reference frame, i.e. inserting the vectors $\underline{a}$=(100), $\underline{b}$=(010) and $\underline{c}$=(001), these equations reduce to the set (2.25). Explicit forms for the case where the quantisation axis is orientated along (110) or (111) can for example be found in [Lau71,Sta97].

At this point, its worthwhile to summarise what has been achieved in this chapter. So far :

- a one-electron Schrödinger equation operating on cell-periodic functions $u_{nk}(r)$ has been obtained that will be the starting point of the derivation of the bandstructure model in the next section.
- knowledge of the degeneracies and symmetries of the various band edge levels has been acquired when excluding spin-orbit coupling.
- inclusion of spin-orbit coupling has been shown to only consist of forming linear combinations of the eigenfunctions of each level.
- an explicit representation of the valence band states at the band edge ($k=0$) has been derived *including* spin-orbit splitting.
- the freedom to choose a preferred direction for the angular momentum in this representation has been shown to be a powerful means to optimise the set when looking at finite momentum.

Using this information, the next section deals with the derivation and underlying motivations of setting up a bandstructure model within the so-called effective mass (or $k\cdot p$) framework.

[1] It was chosen to include a sign change to comply with most of the current literature.

2.3 - The k·p approach

The Schrödinger equation and its one-electron equivalent that are the starting point of any bandstructure model were presented in the first chapter. As outlined in the previous section, bandstructure calculations are numerically very demanding even in the simplest case when dealing with bulk material. As this thesis is concerned with compositionally varying structures, which is expected to make the physical problem even more difficult, a method that partially relies on material constants rather than a first principles method was chosen. Effective mass theory, also often referred to as $k \cdot p$ ('k-dot-p') theory, provides such a method. The latter approach significantly simplifies the physical problem and provides an easy interpretation, of course at the cost of requiring some input in the form of physical parameters.

The method was pioneered by Kane who used it to calculate the valence bands in bulk germanium and silicon [Kan56] as well as the conduction and valence bands in Indium Antimonide [Kan57]. Although this model is only applicable to bulk materials, it proves to be a useful tool to understand the physics behind the processes that shape the bandstructure. As it is much simpler than its more elaborate counterpart the Luttinger-Kohn [Lut55] model (which is also applicable to compositionally varying structures) and since the latter model reduces to the Kane model in the bulk case, this review of $k \cdot p$ theory will start with the Kane model.

Starting with the simplest bulk model, various improvements are introduced and the transition to compositionally varying structures is discussed. Special attention is paid to the use of this model to non-conventional growth directions such as [110] or [111]. Finally, section 2.4 deals with the inclusion of strain effects which is essential when modelling non-lattice matched compounds like InGaAs on GaAs.

Although this section covers $k \cdot p$ theory in extensive detail, the intrinsic problem of how to connect wavefunctions across an interface in a compositionally varying structure will not be discussed here. Because of its complexity, it was chosen to address the latter subject separately in chapter 3.

2.3.1 - The Kane model

The Kane model [Kan56,Kan57] that is introduced in this section is in many aspects the basis of the effective mass model that is introduced later. Although it is only applicable to bulk structures, it is based on the same ideas and principles as its more elaborate counterpart that is capable of modelling compositionally varying structures. Because of the explicit restriction to bulk materials, the mathematics involved in deriving the Kane model are rather straightforward. As a result, the model proves a very instructive tool to obtain a physical picture of the origin of the effective mass parameters as well as the limitations of the model itself.

As seen in the discussion of the one-electron Schrödinger equation, the shape of the conduction and valence bands near the Γ-point is governed by the $k \cdot p$ term in the Hamiltonian. This term also causes the cell-periodic functions $u_{nk}(r)$ to exhibit a k-dependency. The following observations can now be made:

I. The combination of all zone-centre solutions u_{n0} form a complete set, i.e. every possible solution u_{nk} corresponding to finite momentum k can be expanded in this set [Lut55].

II. As we are mostly interested in the region near the band edge, typically just a fraction of the Brillouin zone being filled with electrons or holes, $H_{k.p}$ can be treated as a perturbation of the zone centre Hamiltonian ($H_{k.p} \ll H_0$) [Dre55a].

III. The bands of interest, being the lowest conduction band (Γ_1-symmetry) and the topmost valence bands (Γ_{15}-symmetry) are typically well separated in energy from all other bands (fig.2.5a)

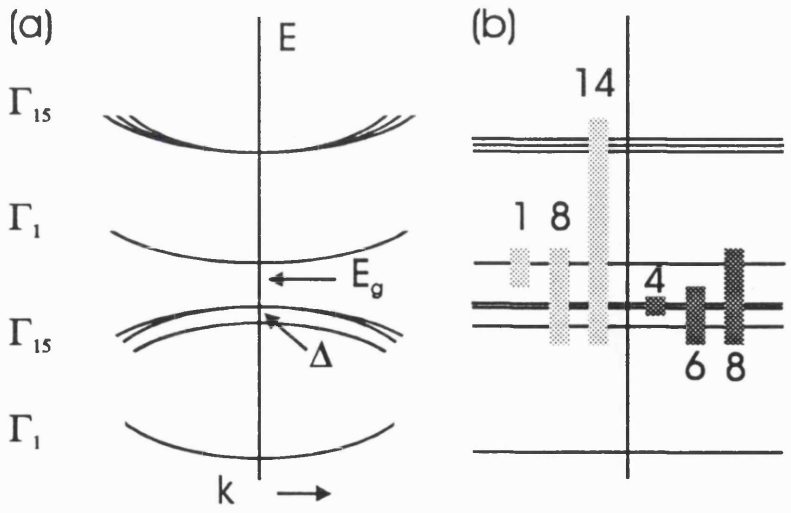


Fig. 2.5 : (a) Schematic picture depicting the relatively large separation between the conduction/valence bands and all other bands at the Γ-point. (b) The bands explicitly included in the Kane expansion to model the conduction band (light) or valence band (dark) for the various schemes present in the literature. The numeral denotes the number of bands included (including spin degeneracy).

The Kane model now proceeds by expanding the cell-periodic functions in just a limited set of zone-centre wavefunctions close in energy to the band under investigation and then uses perturbation theory to account for the effects of the not included, so-called remote, bands. The latter is expected to be a good approximation as perturbation theory states that the energy shift induced by a remote level scales as $1/\Delta E$ where ΔE is the energy separation between the interacting levels. The influence of the remote levels is therefore expected to be relatively small. In contrast, any levels in the vicinity of the band under investigation need to be explicitly included in the expansion as perturbation methods are not expected to accurately account for them.

The question is now what should be regarded as a remote level and what not. Clearly, this would depend on whether one is interested in the conduction or valence band, the choice of semiconductor, the physical property under investigation and the accuracy required. For a system with a relatively large band gap like GaAs one can, when solving for the conduction band, to a good approximation limit the expansion to just this band (a 1-band model, fig.2.5b) and treat all other bands as a perturbation. The situation is slightly more complicated when modelling the valence bands as it consists of more than one band. Here, the expansion needs to explicitly include the heavy and light hole bands (a 4-band model) to obtain results with a reasonable accuracy for GaAs-related compounds. The spin-orbit valence bands do not necessarily have to be included in the expansion due to the relatively large spin-orbit splitting in this compound. Some examples of the number of bands required to model the various material systems with reasonable accuracy are given in fig.2.6b.

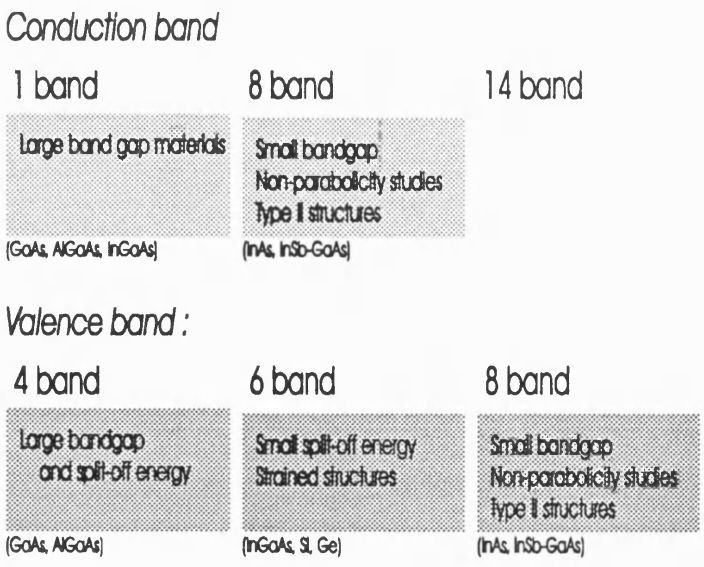


Fig 2.6 : Examples of the number of bands required for modelling various material systems.

Note that these examples are just guidelines and that models of higher order than that suggested for a particular material system can be found in the literature. Furthermore, some applications specifically require a high order model, such as optical anisotropy studies in GaAs QWs which require at least a 6-band model [Kaj95].

To clarify the method, especially with respect to the use of symmetry, the approach as taken in [Bas88] is followed. Here, a model is set up that as a first approximation neglects the presence of the remote bands. It is now assumed that the following limited expansion in zone-centre ($\underline{k}=\underline{0}$) solutions $u_{n0}(r)$ is sufficient to 'describe' $u_{nk}(r)$ in the case of finite momentum k,

$$\begin{aligned} u_{nk} = \; & c_1\left|iS\uparrow\right\rangle + c_2\left|\frac{3}{2},+\frac{1}{2}\right\rangle + c_3\left|\frac{3}{2},+\frac{3}{2}\right\rangle + c_4\left|\frac{1}{2},+\frac{1}{2}\right\rangle \\ & c_5\left|iS\downarrow\right\rangle + c_6\left|\frac{3}{2},-\frac{1}{2}\right\rangle + c_7\left|\frac{3}{2},-\frac{3}{2}\right\rangle + c_8\left|\frac{1}{2},-\frac{1}{2}\right\rangle \end{aligned} \tag{2.28}$$

The physical problem to calculate the bandstructure has thus been reduced to finding the appropriate values of the eight constants $c_{1..8}$. The above expansion has to be an eigenfunction of the original Hamiltonian H_k (2.10) so that the constants are obtained by solving $H_k u_{nk}=E u_{nk}$ for a particular value of the momentum k. Using that each element u_{n0} is an eigenstate of H_0+H_{so} and taking the inner products with respectively $\langle iS\uparrow|$, $\langle 3/2,+1/2|$, ..., the following Hamiltonian in matrix form can be obtained

$$H = \begin{pmatrix} E_g+\frac{\hbar^2k^2}{2m_0^2} & \sqrt{\frac{2}{3}}P\hbar k_z & P\hbar k_+ & \frac{-1}{\sqrt{3}}P\hbar k_z & 0 & -\frac{1}{\sqrt{3}}P\hbar k_- & 0 & -\sqrt{\frac{2}{3}}P\hbar k_- \\ \sqrt{\frac{2}{3}}P\hbar k_z & \frac{\hbar^2k^2}{2m_0^2} & 0 & 0 & \frac{1}{\sqrt{3}}P\hbar k_- & 0 & 0 & 0 \\ P\hbar k_- & 0 & \frac{\hbar^2k^2}{2m_0^2} & 0 & 0 & 0 & 0 & 0 \\ \frac{-1}{\sqrt{3}}P\hbar k_z & 0 & 0 & \Delta+\frac{\hbar^2k^2}{2m_0^2} & \sqrt{\frac{2}{3}}P\hbar k_- & 0 & 0 & 0 \\ 0 & \frac{1}{\sqrt{3}}P\hbar k_+ & 0 & \sqrt{\frac{2}{3}}P\hbar k_+ & E_g+\frac{\hbar^2k^2}{2m_0^2} & \sqrt{\frac{2}{3}}P\hbar k_z & P\hbar k_- & \frac{-1}{\sqrt{3}}P\hbar k_z \\ -\frac{1}{\sqrt{3}}P\hbar k_+ & 0 & 0 & 0 & \sqrt{\frac{2}{3}}P\hbar k_z & \frac{\hbar^2k^2}{2m_0^2} & 0 & 0 \\ 0 & 0 & 0 & 0 & P\hbar k_- & 0 & \frac{\hbar^2k^2}{2m_0^2} & 0 \\ -\sqrt{\frac{2}{3}}P\hbar k_+ & 0 & 0 & 0 & \frac{-1}{\sqrt{3}}P\hbar k_z & 0 & 0 & \Delta+\frac{\hbar^2k^2}{2m_0^2} \end{pmatrix} \tag{2.29}$$

Here, the zero-point of energy is chosen at the top of the valence bands, Δ is the spin-orbit splitting and $k_\pm$ is defined as $k_\pm = \left(k_x \pm k_y\right)/\sqrt{2}$. The Hamiltonian contains only one adjustable parameter which is the momentum matrix element P defined by

$$P = \frac{-i}{m_0}\langle S|p_x|X\rangle = \frac{-i}{m_0}\langle S|p_y|Y\rangle = \frac{-i}{m_0}\langle S|p_z|Z\rangle \tag{2.30}$$

In words: P is the volume integral over a primitive cell of the product of a conduction band state $|S\rangle$ and a momentum operator working on a valence band orbital. From symmetry arguments it follows that these are the only non-zero momentum matrix elements (fig.2.7).

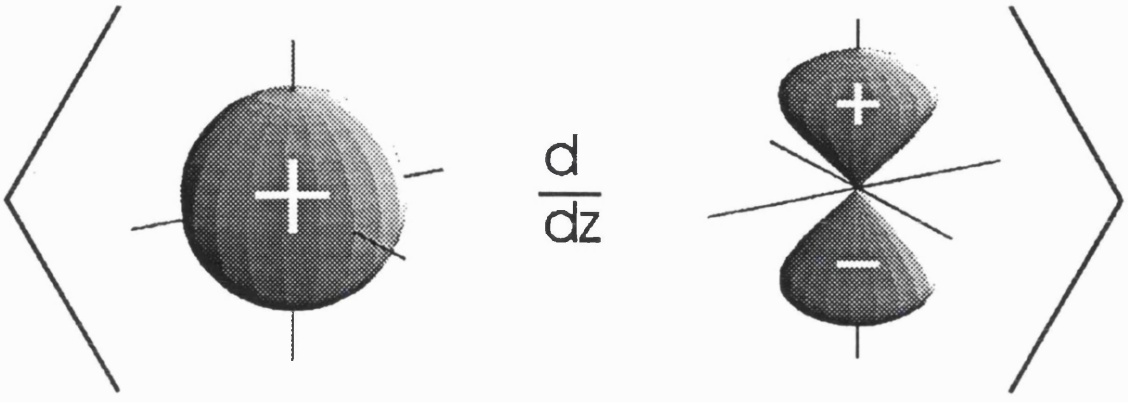


Fig.2.7: The inner product $\langle S|p_z|Z\rangle$ is the integral over space of an s-orbital (left) multiplied by $\partial/\partial z$ of a p_z-orbital (right). This integral (as well as $\langle S|p_x|X\rangle$ and $\langle S|p_y|Y\rangle$) is non-zero since it is even in x, y and z. In contrast to this, all inner products involving two p-orbitals and a differential are zero as the total function is odd in at least one coordinate.

The original eigenfunction problem has now been reduced to finding the eigenvalues of an n by n matrix, n here being eight. It follows directly from the fact that (2.29) is not a diagonal matrix at non-zero momentum k that the k·p interaction causes the wavefunctions away from the band edge to be a mixture of the original band edge s- and p-type zone-centre solutions. Returning to the eigenvalues, one obtains for the dispersion at small k [Kan57],

$$\begin{aligned}
E_c &= E_g + \frac{\hbar^2 k^2}{2m_0} + \frac{P^2 k^2}{3}\left(\frac{2}{E_g} + \frac{1}{E_g + \Delta}\right) \\
E_{hh} &= \frac{\hbar^2 k^2}{2m_0} \\
E_{lh} &= \frac{\hbar^2 k^2}{2m_0} - \frac{2P^2 k^2}{3E_g} \\
E_{so} &= -\Delta + \frac{\hbar^2 k^2}{2m_0} - \frac{P^2 k^2}{3\left(E_g + \Delta\right)}
\end{aligned} \tag{2.31}$$

Isotropic solutions are found, i.e. the energies are only a function of the length of $\underline{k}$ and not of its direction, as well as a dispersionless heavy hole band. Both features do not agree with the experimental results for III-V materials. This indicates that the remote bands can not be neglected.

The previous results show why it was chosen to introduce the Kane model by looking at an 8-band instead of a simpler 4- or 6-band model. Obviously, the only interaction between the bands present is the interaction between the conduction and the valence bands so that both lower order models would have given trivial empty lattice results. Or stated differently; the $k \cdot p$ term does not cause the p-orbitals to interact with themselves, so that any deviations from empty-lattice results are solely due to interactions with other bands. However, it turns out that 8-band and higher models have intrinsic problems in the form of so-called spurious solutions [Sch85b]. These are unphysical solutions that arise due to the fact that the Hamiltonian matrix (2.29) gives four (spin degenerate) solutions, whereas there are only *three* physical ones as the conduction and light hole band are joined in the complex plane. It was therefore chosen to treat the conduction and valence bands separately in this work. As a result, the conduction band will be considered a remote state when modelling the valence bands, and vice versa.

2.3.2 - The inclusion of remote bands

The previous discussion showed that an expansion in just the heavy, light and split-off hole states plus the conduction band, thereby totally neglecting the remote bands, is insufficient to accurately model the valence bands in III-V materials. As the expansion will be limited even further to either just the conduction band or just the valence bands, it is obvious that the remote states will further increase in importance and have to be accounted for.

Because of their large energy separation, it is however expected that these remote bands can be accounted for by using a perturbation method, denoted as Löwdin renormalisation [Löw51]. Following this paper, the resulting additional perturbation terms in the Hamiltonian are now derived up to order two for this specific case.

The band edge solutions u_{n0} are divided into two classes, a set (a) close in energy to the energy region of interest and a remote set (b) which is well separated in energy. The first set (a) will be explicitly included in the expansion, whereas the second set (b) will only be included via a perturbation method. The combined set (a) + (b) form a complete orthonormal set so that an arbitrary wavefunction can be expanded in this set as

$$|u_{nk}\rangle = \sum_{l=1}^{r} c_l |l\rangle + \sum_{\mu=1}^{s} c_\mu |\mu\rangle \tag{2.32}$$

the summation over l being the summation over r closely spaced levels and the summation over μ being the summation over the remote bands. This expansion is now used to solve the perturbed problem including the $k\cdot p$ interaction energy $H_{k\cdot p} = \frac{\hbar}{m_0} k \cdot p$. Inserting (2.32) in the eigenvalue equation

$$\left[H_0 + H_{so} + H_{k\cdot p}\right]\left\{\sum_{l=1}^{r} c_l |l\rangle + \sum_{\mu=1}^{s} c_\mu |\mu\rangle\right\} = E\left\{\sum_{l=1}^{r} c_l |l\rangle + \sum_{\mu=1}^{s} c_\mu |\mu\rangle\right\} \tag{2.33}$$

and taking the inner product with $\langle m|$ and $\langle \nu|$ where m corresponds to an element from the near set (a) and ν corresponds to the remote set (b), one obtains

$$\begin{aligned} \sum_l c_l\left[(E_l - E)\delta_{lm} + \langle m|H_{k.p}|l\rangle\right] \; + \; \sum_\mu c_\mu \langle m|H_{k.p}|\mu\rangle \; &= 0 \\ \sum_\mu c_\mu\left[(E_\mu - E)\delta_{\mu\nu} + \langle \nu|H_{k.p}|\mu\rangle\right] \; + \; \sum_l c_l \langle \nu|H_{k.p}|l\rangle \; &= 0 \end{aligned} \tag{2.34}$$

The above relations were derived by using the orthonormality relations as well as the fact that the expansion states are band edge solutions at energies E_l and E_μ. Near the band edge, at small k, the second equation of (2.34) can be approximated by

$$c_\nu \approx \frac{1}{E - E_\nu} \sum_l c_l \langle \nu|H_{k.p}|l\rangle \tag{2.35}$$

which inserted into the first equation of (2.34) results in the following relation between the expansion constants of the near set only,

$$\sum_l c_l \left((E_l - E)\delta_{lm} + \langle m|H_{k.p}|l\rangle + \sum_\nu \frac{\langle m|H_{k.p}|\nu\rangle\langle \nu|H_{k.p}|l\rangle}{E - E_\nu} \right) = 0 \tag{2.36}$$

Such a transition from an expansion in just $|l\rangle$, totally neglecting the remote bands, to an expansion in the same limited set, but where the effects of the remote bands have been included by perturbation theory, modifies the Hamiltonian operating on the basis $|l\rangle$ by adding an extra term $H_{ml}^{Löwdin}$ given by

$$
\begin{aligned}
H_{ml}^{\text{löwdin}} &= \sum_{\nu} \frac{\langle m|H_{k.p}|\nu\rangle\langle\nu|H_{k.p}|l\rangle}{E-E_{\nu}} \\
&= \frac{\hbar^2}{m_0^2}\sum_{\nu} \frac{\langle m|\underline{k}\cdot\underline{p}|\nu\rangle\langle\nu|\underline{k}\cdot\underline{p}|l\rangle}{E-E_{\nu}} \\
&= \frac{\hbar^2}{m_0^2}k_{\alpha}k_{\beta}\sum_{\nu} \frac{\langle m|p^{\alpha}|\nu\rangle\langle\nu|p^{\beta}|l\rangle}{E-E_{\nu}} \\
&= \frac{\hbar^2}{m_0^2}k_{\alpha}k_{\beta}\sum_{\nu} \frac{p_{m\nu}^{\alpha}p_{\nu l}^{\beta}}{E-E_{\nu}}
\end{aligned}
\qquad (2.37)
$$

where the summation over ν runs over all remote bands. The Einstein convention has been used that any index occurring both as a lower as well as a higher index should be interpreted as the sum over all allowed values of that index. Furthermore, the general momentum matrix element $p_{m\nu}^{\alpha}=\langle m|p_{\alpha}|\nu\rangle$ has been defined which replaces the earlier momentum matrix element P (2.30) that was restricted to just the interaction between a valence band p-orbital and a conduction band s-orbital. Whence, the Hamiltonian in matrix form for an expansion in the limited n-dimensional set $|l\rangle$ *limited to just one level* (i.e. working with an explicit expansion in just Γ_1 or Γ_{15}), including all other bands by perturbation techniques, takes the form

$$
H_{jj'}\cdot\begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} = E\begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} \quad ,where \quad \begin{aligned} H_{jj'} &= \hbar k_{\alpha}\hbar k_{\beta}D_{jj'}^{\alpha\beta} \\ D_{jj'}^{\alpha\beta} &= \frac{1}{2m_0}\delta_{jj'}\delta_{\alpha\beta} + \frac{1}{m_0^2}\sum_{\nu}\frac{p_{j\nu}^{\alpha}p_{\nu j'}^{\beta}}{E-E_{\nu}} \end{aligned} \qquad (2.38)
$$

where the indices m,l have respectively been replaced by the more common j,j'. Whereas there was only one adjustable parameter P in the previous Hamiltonian (2.29), the presence of the momentum matrix elements $p_{j\nu}^{\alpha}$ and $p_{\nu j'}^{\beta}$ in (2.38) suggests that the inclusion of the remote states will induce additional parameters. The same approach as used in section 2.1 can be followed where simple symmetry arguments along with an explicit knowledge of the various orbitals were used to obtain the non-zero matrix elements and adjustable parameters. This approach will be used, with some severe limitations on the number of participating remote states ν, in the results chapter (chapter 6) when discussing the boundary condition problem. Other parametrisation approaches however exist which prove to be more convenient in the multi-band case. The question of how to parametrise this Hamiltonian is postponed until sub-section 2.3.4. Instead, the previous model is now going to be extended to the case of compositionally varying structures.

2.3.3 - The Luttinger-Kohn envelope model

Luttinger and Kohn developed a *k·p* model [Lut55] which resembles the Kane model but does not have the restriction to be only applicable to uniform structures. At the time of its derivation, it was meant for theoretical calculations of impurity states and analysis of cyclotron resonance in bulk materials, but the general way it treats the potential term allows it to be used for compositionally varying samples as well. Here, the original derivation, which is completely performed in *k*-space, is not going to be repeated as it is very abstract and tends to hide the physics. Such a mathematically rigorous derivation is not required in this work, yet it is worth to point out that the approach that is going to be followed here does not provide any information about the presence of some of the restrictions to the use of this model in structures with abrupt potentials. One therefore has to resort to such a *k*-space derivation when looking at for example the boundary conditions (see chapter 3).

It was proved in [Lut55] that any wavefunction in a compositionally varying structure can be written as

$$\psi(\underline{r}) = \sum_j F_j(\underline{r})\, u_{j0}(\underline{r}) \tag{2.39}$$

where F_j are so-called envelope functions which vary slowly on atomic scale. The zone centre solutions u_{j0} guarantee completeness on atomic scale whereas the envelope functions guarantee completeness on a larger scale. This expansion looks very similar to the expansion used in the Kane model, the only difference being that the combination of the constants c_j and the Bloch-exponential $e^{i\underline{k}\cdot\underline{r}}$ have been replaced by the more general $F_j(\underline{r})$. This expansion is now going to be used to describe a wavefunction in a compositionally varying structure. Returning to the general Hamiltonian (2.3), and using the expansion (2.39), one now has to solve

$$\left[-\frac{\hbar^2}{2m_0}\nabla^2 + U(\underline{r})\right]\sum_j F(\underline{r})u_{jo}(\underline{r}) = E\sum_j F(\underline{r})u_{jo}(\underline{r}) \tag{2.40}$$

One has to realise that the potential $U(r)$ as well as the zone-centre expansion states $u_{jo}(r)$ now not only have a microscopic but also a macroscopic dependence as one allows for compositional changes. Assuming that the system under investigation consists of a number of stacked layers with different composition, this has the effect that the potential $U(r)$ exhibits abrupt changes near each interface.

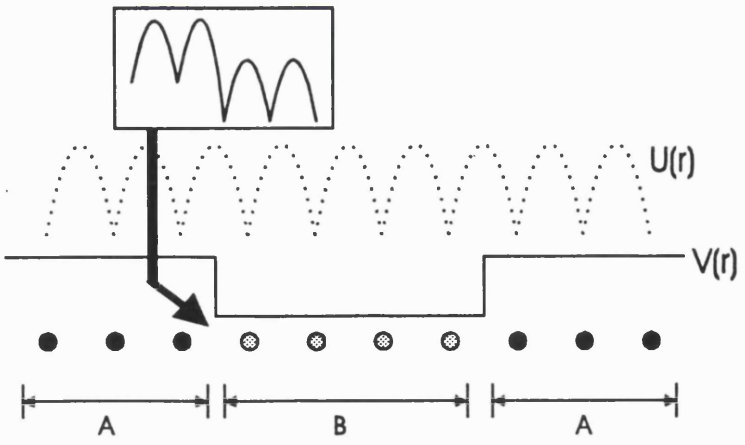


Fig.2.8 : Illustration of the periodic potential U(r) in a compositionally varying structure with regions of composition 'A' and 'B'. The inset shows the true (although fictitious) form of the potential U(r) near an interface. Assuming that the microscopic form of U(r) changes little when changing composition, the potential can be approximated by U(r)+V(r) where U(r) is now the potential of bulk 'A' and V(r) is a piecewise constant modulation.

Following the original paper, an electrostatic potential $V(r)$ is introduced to describe the variation in potential as one moves from one region of a certain composition to one with a different composition. A simplified picture is given in fig.2.8 where it is assumed that the change in composition only changes the base level of the potential and not the microscopic form. As a result, one effectively assumes to be working with the *same* material composition throughout the sample as the invariance in the microscopic form of the potential U implies that the zone-centre cell-periodic functions u_{j0} are identical throughout the whole structure. This approximation is dealt with in more detail in chapter 3.

Assuming that a compositionally varying structure can be described solely by using such a macroscopic potential $V(r)$ that is acting on the electrons or holes, equation (2.40) transforms into

$$\left[-\frac{\hbar^2}{2m_0}\nabla^2 + U(\underline{r}) + V(\underline{r})\right]\sum_j F(\underline{r})u_{jo}(\underline{r}) = E\sum_j F(\underline{r})u_{jo}(\underline{r}) \qquad (2.41)$$

where the zone centre cell-periodic wavefunctions u_{n0} form the solutions when $V(r)=0$. The potential can therefore be thought of as a position dependent perturbation. The resulting effective mass equations for the compositionally varying case can now be obtained in a similar way as in the derivation of the Kane model. The summation over the zone-centre cell-periodic functions u_{j0} is again divided into two classes, a near set that will be explicitly included in the expansion and a remote set that will be included via perturbation techniques (section 2.3.2). Using that the envelope functions vary slowly on atomic scale, taking the inner product with the various cell-periodic functions and eliminating the smaller contributions in favour of dominant ones, one obtains [Lut55]

$$\sum_{j'}\left[\left(\frac{\hbar}{i}\frac{\partial}{\partial x_\alpha}\right)\left(\frac{\hbar}{i}\frac{\partial}{\partial x_\beta}\right)D_{jj'}^{\alpha\beta}+V(\underline{r})\delta_{jj'}\right]F_{j'}(\underline{r})=EF_{j'}(\underline{r})$$

$$D_{jj'}^{\alpha\beta}=\frac{1}{2m_0}\delta_{jj'}\delta_{\alpha\beta}+\frac{1}{m_0^2}\sum_{\nu}\frac{p_{j\nu}^{\alpha}p_{\nu j'}^{\beta}}{E-E_\nu} \tag{2.42}$$

Strong similarities are observed between the Kane (2.38) and the above Luttinger-Kohn model (2.42). The latter differential system for modelling compositionally varying structures can simply be obtained from the bulk Kane equations by making the following replacements :

- changing the momenta $\hbar k_i$ to the quantum-mechanical operators p_i.
- replacing the expansion coefficients c_j by position dependent envelope functions F_j.
- adding a position dependent potential $V(\underline{r})$ on the diagonal.

Although these two models look very similar, these transitions make the problem of finding the eigenvalues considerably more complicated for the Luttinger-Kohn model as compared to the Kane model. In the latter case, finding the eigenvalues merely consists of finding the roots of a determinant. In the former case however, one is dealing with a system of coupled differential equations, i.e. a set of differentials which have to be satisfied simultaneously.

2.3.4 - Parametrised form of the Luttinger Hamiltonian

The Hamiltonian that has been derived in the previous discussion is not of much use due to the fact that neither the inner products $p_{j\nu}$ nor the energies of the remote bands E_ν are all known with sufficient accuracy. However, lattice and orbital symmetry can, similarly to the approach taken in the Kane model in the case where the remote states were completely ignored, be used to reduce the equations into a parametrised form with just a limited number of constants. These constants can then be fitted by experiment.

For the conduction band, an expansion in just Γ_1 is thought to be sufficient as this band is well separated in energy from all other bands,

$$u_{nk}(r)=F_s(r)|S\rangle \tag{2.43}$$

so that one finds that one has to solve a single differential equation with the relevant Hamiltonian,

$$
\begin{aligned}
H_s &= \frac{\hbar^2 k^2}{2m_0} + \frac{\hbar k_\alpha \hbar k_\beta}{m_0^2} \sum_v \frac{p_{sv}^\alpha p_{vs}^\beta}{E - E_v} \\
&= \sum_\alpha \frac{\hbar^2 k_\alpha^2}{2m_0} \left(1 + \frac{2}{m_0} \sum_v \frac{|p_{sv}^\alpha|^2}{E - E_v} \right) \\
&= \frac{\hbar^2}{2m^*} \left(k_x^2 + k_y^2 + k_z^2 \right)
\end{aligned}
\tag{2.44}
$$

where the directions x,y and z have been taken as equal. An effective mass m^* has been introduced which is given by

$$
m^* = m_0 \left(1 + \frac{2}{m_0} \sum_v \frac{|\langle S|p_x|v\rangle|^2}{E - E_v} \right)^{-1} \tag{2.45}
$$

The fact that no cross-terms like $k_x k_y$ appear in (2.44) follows from the fact that an s-state is spherically symmetric. Therefore, any product of the form $\langle S|\partial/\partial x|v\rangle\langle v|\partial/\partial y|S\rangle$ is zero since $\langle S|\partial/\partial x|$ is odd in x and even in y, whereas $|\partial/\partial y|S\rangle$ is even in x and odd in y. So, for a compositionally varying structure in which the composition has just a z-dependency, one finds the well-known effective mass equation

$$
\left[-\frac{\hbar^2}{2m^*} \frac{\partial^2}{\partial z^2} + V(z) + \frac{\hbar^2}{2m^*} \left(k_x^2 + k_y^2 \right) \right] F_s(z) = E\, F_s(z) \tag{2.46}
$$

The valence band problem is more complicated due to the fact that it consists of a set of closely spaced bands which have to be explicitly included in the expansion. One therefore ends up with a set of coupled differential equations in which one intuitively expects to have more adjustable parameters as compared to the conduction band case. Again, the valence band Hamiltonian $H_{jj'}$, where j,j' run over the heavy, light and split-off bands, can be simplified considerably by forcing it to be invariant under all lattice symmetry operations and subsequently replacing the remaining unknowns by constants which can be fitted by experiment.

Turning to the valence band case, a parametrised form of the matrix $H_{jj'}$ for an explicit expansion in the set $|X\rangle$, $|Y\rangle$ and $|Z\rangle$ is evaluated. According to (2.42), this matrix contains elements of the form $k_\alpha k_\beta$, where α and β run over x,y,z. At this point, one would like to impose restrictions on the possible values of α and β to reduce the complexity of the matrix and the number of constants involved. From symmetry[†] considerations it then follows that the most general Hamiltonian for this expansion is of the form [Dre55a]

[†] The general parametrised form (2.47) was actually derived for the diamond symmetry rather than for Zinc-Blende. The loss of some of the symmetry operations when looking at Zinc-Blende would introduce an extra parameter in the Hamiltonian but since its effect is very small, it is normally neglected [Bro85].

$$D=\begin{pmatrix} Lk_x^2+M(k_y^2+k_z^2) & Nk_xk_y & Nk_xk_z \\ Nk_yk_x & Lk_y^2+M(k_x^2+k_z^2) & Nk_yk_z \\ Nk_zk_x & Nk_zk_y & Lk_z^2+M(k_x^2+k_y^2) \end{pmatrix} \tag{2.47}$$

in which

$$\begin{aligned} L &= \frac{\hbar^2}{2m_0} + \frac{\hbar^2}{m_0^2}\sum_{\nu}\frac{\left|\langle X|p_x|\nu\rangle\right|^2}{E-E_\nu} \\ M &= \frac{\hbar^2}{2m_0} + \frac{\hbar^2}{m_0^2}\sum_{\nu}\frac{\left|\langle X|p_y|\nu\rangle\right|^2}{E-E_\nu} \\ N &= \frac{\hbar^2}{m_0^2}\sum_{\nu}\frac{\langle X|p_x|\nu\rangle\langle\nu|p_y|Y\rangle}{E-E_\nu} \end{aligned} \tag{2.48}$$

The Hamiltonian given in (2.42) has now been rewritten in a parametrised form which only contains three adjustable parameters; L, M and N. It is however customary to replace these parameters by the Luttinger parameters, which are defined as [Lut55]

$$\begin{aligned} \frac{\hbar^2}{2m_0}\gamma_1 &= -\frac{1}{3}(L+2M) \\ \frac{\hbar^2}{2m_0}\gamma_2 &= -\frac{1}{6}(L-M) \\ \frac{\hbar^2}{2m_0}\gamma_3 &= -\frac{1}{6}N \end{aligned} \tag{2.49}$$

At this point one has to realise that the desired form of the Hamiltonian has not yet been obtained, as it was shown in sub-section 2.2.3 that the states $|X\rangle$, $|Y\rangle$ and $|Z\rangle$ are not eigenstates of the spin-orbit term and are thus not the most convenient set for an expansion. The above Hamiltonian has therefore to be rewritten for an expansion in $|J,m_j\rangle$ such as given by for example (2.25) in which the spin and angular momentum have been orientated along [001]. Such a basis transformation is a trivial exercise so that, using an expansion

$$u_{nk} = F_1\left|\tfrac{3}{2},+\tfrac{3}{2}\right\rangle + F_2\left|\tfrac{3}{2},+\tfrac{1}{2}\right\rangle + F_3\left|\tfrac{3}{2},-\tfrac{1}{2}\right\rangle + F_4\left|\tfrac{3}{2},-\tfrac{3}{2}\right\rangle + F_5\left|\tfrac{1}{2},+\tfrac{1}{2}\right\rangle + F_6\left|\tfrac{1}{2},-\tfrac{1}{2}\right\rangle \tag{2.50}$$

one ends up with the following Hamiltonian

$$H=\begin{pmatrix} P+Q & -S & R & 0 & S/\sqrt{2} & \sqrt{2}R \\ & P-Q & 0 & -R & \sqrt{2}Q & \sqrt{\tfrac{3}{2}}S \\ & & P-Q & -S & -\sqrt{\tfrac{3}{2}}S^\dagger & \sqrt{2}Q \\ & & & P+Q & -\sqrt{2}R^\dagger & S^\dagger/\sqrt{2} \\ & & & & P+\Delta & 0 \\ & & & & & P+\Delta \end{pmatrix} \tag{2.51}$$

with elements

$$[001]\quad \begin{cases} P=\gamma_1\left(k_x^2+k_y^2+k_z^2\right) \\ Q=\gamma_2\left(k_x^2+k_y^2-2k_z^2\right) \\ S=2\sqrt{3}\left(k_x-ik_y\right)k_z \\ R=\sqrt{3}\left[\gamma_2\left(k_x^2+k_y^2\right)+2i\gamma_3 k_x k_y\right] \end{cases} \tag{2.52}$$

Only the upper diagonal in the Hamiltonian has been indicated as the remaining elements follow from the requirement that the Hamiltonian must be Hermitian. Furthermore, the macroscopic potential $V(r)$ associated with the band line-up in compositionally varying structures has been omitted from the diagonal elements and a factor $-\hbar^2/2m_0$ has been absorbed in each term. The boxed area contains the Hamiltonian elements when limiting the interactions to just the heavy and light hole states, i.e. using a 4-band model.

It is important to note that the Hamiltonian (2.51) is really only valid for the specific choice of $|J,m_j\rangle$ basis states (2.25) and resulting elements (2.52). To illustrate, when comparing the above Hamiltonian to the one presented in [Edw94] that is based on a different set, one can see various sign differences in the Hamiltonian elements. Or stated differently, there is no unitary transformation that can account for phase differences in the basis set without altering the Hamiltonian slightly. Various forms of P,Q,R,S with different phase conventions can be found in the literature, but the 6-band Hamiltonian as given above is only valid when using (2.52). At this point, it suffices to note that the 4-band Hamiltonian does not require a specific knowledge of the phase of the $|J,m_j\rangle$-set and is valid for all possible phase choices. The influence of the phase will be discussed in detail in chapter 6.

The above Hamiltonian and specific elements will be used to model [001]-grown structures. For the compositionally varying case, the momentum elements $\hbar k_z$ should then be interpreted as the operator $\hbar/i\ \partial/\partial z$, whereas k_x and k_y remain normal vector components or better: quantum numbers which originate from the plane wave form $exp[i(k_x x+k_y y)]$ of the wavefunction perpendicular to the growth direction.

2.3.5 - Non-[001] growth directions

The Hamiltonian that will be employed to model structures grown on [001] has been presented in the previous sub-section. As we have made no particular steps or approximations that restrict this Hamiltonian to just one particular growth direction, one might be surprised that it has been

specified to be used for this growth direction only. In effect, there is no restriction of this kind so that the Hamiltonian (2.51,2.52) could be used for, say, growth on (311) substrates. Performing a coordinate transformation in which the k-vectors (k_x,k_y,k_z) are replaced by $(k_{x'},k_{y'},k_{z'})$ with z' parallel to the growth direction would only change the framework that $\hbar k_{z'}$ is now to be interpreted as the operator $\hbar i\partial/\partial z$ and $k_{x'}$ and $k_{y'}$ as vector components. This way, a similar system of coupled differential equations as for the [001] case can be obtained, yet now applicable for the case where the composition varies when moving along z'. The problem associated with such an approach is however that the resulting equation is rather messy, or rather ill-conditioned, and needs to be 'cleaned up' by a unitary transformation to render it into a form in which it is usable. This comes as no surprise as the set $|J,m_j\rangle$ was designed to be a set of eigenfunctions of H_0, H_{so} and $H_{k.p}$ for the case where the momentum k was limited to just the direction [001]. One therefore can not expect this Hamiltonian in matrix form to take a diagonal form when deviating from this preferred direction. The solution to this problem is however straightforward as a general form for the set $|J,m_j\rangle$ was given in 2.2.5 where it was also shown that the most convenient quantisation axis for the angular momentum was that of the same direction in k-space one is interested in. Despite the convenient implementation, one can still find articles in the literature that start from a [001] Hamiltonian and then make a number of unitary transformations to obtain a Hamiltonian that can be employed to model a completely different growth direction [Gol95]. Such a method generally does not provide the easiest Hamiltonian and is considered to only obscure matters. The only type of unitary transformation that is considered worthwhile is that described in sub-section (2.3.6) where it is shown that a 6-band Hamiltonian can be written into two 3x3 Hamiltonians, thereby significantly reducing the numerical problem. As the latter is best performed using a starting Hamiltonian that utilises an expansion in the appropriate $|J,m_j\rangle$-set for that particular growth direction, the approach as taken by Goldoni [Gol95] is not recommended.

Using the general form of the basis set $|J,m_j\rangle$ (2.26,2.27), it can be proved that the Hamiltonian (2.51) is in fact valid for all growth directions [Dal98]. Using this basis, the elements of this Hamiltonian for growth on for example (110) and (111) are now given by

$$[110]\begin{cases} P &= \gamma_1\left(k_x^2+k_y^2+k_z^2\right) \\ Q &= \gamma_2 k_x^2-\left(\frac{1}{2}\gamma_2-\frac{3}{2}\gamma_3\right)k_y^2-\left(\frac{1}{2}\gamma_2+\frac{3}{2}\gamma_3\right)k_z^2 \\ S &= 2\sqrt{3}\left(\gamma_3 k_x-\gamma_2 i k_y\right)k_z \\ R &= \sqrt{3}\left[\gamma_2 k_x^2-\frac{1}{2}(\gamma_2+\gamma_3)k_y^2-\frac{1}{2}(\gamma_2-\gamma_3)k_z^2-2i\gamma_3 k_x k_y\right] \end{cases} \tag{2.53}$$

$$[111]\begin{cases} P &= \gamma_1\left(k_x^2+k_y^2+k_z^2\right) \\ Q &= -\gamma_3\left(k_x^2+k_y^2\right)+2\gamma_3 k_z^2 \\ S &= -\frac{\sqrt{6}}{3}(\gamma_2-\gamma_3)(k_x+ik_y)^2+\frac{2}{\sqrt{3}}(2\gamma_2+\gamma_3)(k_x-ik_y)k_z \\ R &= -\frac{2\sqrt{6}}{3}(\gamma_2-\gamma_3)(k_x+ik_y)k_z+\frac{1}{\sqrt{3}}(\gamma_2+2\gamma_3)(k_x-ik_y)^2 \end{cases} \tag{2.54}$$

Comparing these Hamiltonian elements to those presented in for example [Men91,Iko92] reveals small phase differences in the various elements. As discussed in 2.2.5, these differences are due to phase differences in the $|J,m_j\rangle$ expansion states. Again the same restriction as in the [001] case applies that the 6-band Hamiltonian (2.51), unlike the 4-band Hamiltonian, is only valid for the above sets P,Q,R,S that were specifically derived using the same basis as that used in the derivation of the Hamiltonian. As stated before, this will be dealt with in more detail in chapter 6.

2.3.6 - A unitary transformation to render the Hamiltonian in block-diagonal form

The Hamiltonian as presented in (2.51), although perfectly adequate to model any semiconductor structure, is not the most convenient choice as it can be reduced to a much simpler block diagonal form by means of a basis rotation. This will render the 6x6 Hamiltonian into two 3x3 Hamiltonians, thereby significantly reducing the numerical problem. One possible choice of this transformation is given by [Bro85,For93]

$$\begin{pmatrix} |u_1\rangle \\ |u_2\rangle \\ |u_3\rangle \\ |u_4\rangle \\ |u_5\rangle \\ |u_6\rangle \end{pmatrix} = \begin{pmatrix} -\alpha^* & & & \alpha & & \\ & \beta & \beta^* & & & \\ & & & & \beta & \beta^* \\ \alpha^* & & & \alpha & & \\ & \beta & -\beta^* & & & \\ & & & & \beta & -\beta^* \end{pmatrix} \begin{pmatrix} \left|\frac{3}{2},+\frac{3}{2}\right\rangle \\ \left|\frac{3}{2},+\frac{1}{2}\right\rangle \\ \left|\frac{3}{2},-\frac{1}{2}\right\rangle \\ \left|\frac{3}{2},-\frac{3}{2}\right\rangle \\ \left|\frac{1}{2},+\frac{1}{2}\right\rangle \\ \left|\frac{1}{2},-\frac{1}{2}\right\rangle \end{pmatrix} \quad or \quad u' = Uu \tag{2.55}$$

where the basis states are rotated within the heavy hole, light hole or split-off bands only. Using

$$\alpha = \beta = \frac{1+i}{2} \tag{2.56}$$

one then obtains that the Hamiltonian for any growth direction in the case of zero in-plane momentum can be written as [For93]

$$H=\begin{pmatrix} H_+^{3x3} & 0 \\ 0 & H_-^{3x3} \end{pmatrix} \qquad H_\pm^{3x3}=\begin{pmatrix} P+Q & R\mp iS & \sqrt{2}R\pm\frac{i}{\sqrt{2}}S \\ & P-Q\mp iC & \sqrt{2}Q\mp i\sqrt{\frac{3}{2}}S \\ & & P+\Delta\pm iC \end{pmatrix} \tag{2.57}$$

The Hamiltonian matrix element C has been introduced because of its relevance later on when discussing the boundary conditions. At this point, it suffices to set this element to zero.

If one makes the (gentle) approximation that $\gamma_2 \approx \gamma_3$ in the light hole $\Leftrightarrow$ spin-orbit interaction term, α and β can be chosen such that (2.57) is even valid for non-zero in-plane momentum in the case of [001]-growth. This unitary transformation is actually *exact* when limiting the expansion to four bands for this particular growth direction, i.e. when reducing a 4x4 to two 2x2 Hamiltonians.

2.3.7-Luttinger's theory of invariants

Using a group theoretical approach, Luttinger found an analytical expression for a four band Hamiltonian in matrix form in terms of k-vectors and angular momentum matrices. Without magnetic field, this Hamiltonian takes the form,

$$\begin{aligned} H=-\frac{\hbar^2}{2m_0}\Big[&(\gamma_1+5/2\gamma_2)k^2-2\gamma_2\left(k_x^2J_x^2+k_y^2J_y^2+k_z^2J_z^2\right) \\ &-4\gamma_3\left(k_xk_y\{J_xJ_y\}+k_yk_z\{J_yJ_z\}+k_zk_x\{J_zJ_x\}\right)\Big] \end{aligned} \tag{2.58}$$

where $\{xy\}=xy+yx$ and $J_{x,y,z}$ are the angular momentum matrices corresponding to spin 3/2. This expression for the Hamiltonian is of particular value when setting up an effective mass model for non-[001] structures, as the latter only involves performing the appropriate coordinate transformation on momentum vectors and angular momentum matrices.

The reason why it was decided to follow a different approach in this work is simply that the 'theory of invariants' has the disadvantage of not giving a lot of insight into the physical problem. Furthermore, in its present form it is restricted to deriving the Hamiltonian matrix elements of 4-band models only. In principle, this does not prevent the use of (2.58) to set up a higher order model, as (2.51) reveals that the extra matrix elements present in a 6-band model are directly related to the elements of the 4-band model. However, this link could obviously not have been made had one not performed the $|J,m_j\rangle$ derivation. In addition, there is the restriction that the

Hamiltonian (2.51) was derived using an explicit form for the set $|J,m_j\rangle$ and that this 6-band Hamiltonian is strictly only valid for elements P,Q,R,S that correspond to the same phase choice.

It turns out that 'The theory of invariants' can still be of great use, even when looking at effective mass models of order higher than four (see chapter 6). However, one has to acknowledge that the group theoretical result lacks some of the flexibilities of the approach used in this thesis.

2.4 - Extension to strained materials

Most of the earlier work on III-V materials concerned structures grown using the AlGaAs/GaAs material system. As GaAs and AlAs are (almost) lattice matched, it proves to be relatively easy to grow good quality, dislocation free, heterostructures using this ternary. However, advances in growth procedures now also allow pseudomorphic growth of a variety of materials with different lattice parameters such as InGaAs on GaAs. In this type of growth, all layers adopt the in-plane lattice parameter of the underlying substrate while keeping the formation of defects to a minimum. Although more difficult to manufacture from a growth point of view, it turns out that strain introduces a number of interesting effects which can be used to enhance electrical as well as optical device performance. Strain can for example be used to lower the hole effective mass [Jaf87] to improve transport properties [Kud97], or to lower the laser threshold current [Ahn94]. Furthermore, strained Zinc-Blende structures can exhibit a piezo-electric field due to the lack of inversion symmetry [Smi86]. The presence of such an internal electric field can significantly affect the wavefunctions and carrier distribution. As a result, it can be seen as a means to tailor the electron or hole distribution [Lu94,Hit95].

Obviously, one would like to include strain effects in the previously presented effective mass models. It turns out that the extension of the earlier presented effective mass equations to strained structures is relatively straightforward and that the effects of strain can be accommodated by simply adding some new strain terms in the Hamiltonian. The derivation of these additional terms is in many respects identical to the derivation of the original unstrained effective mass model. To be able to derive these terms, one first needs to establish how strain affects the periodic lattice, and this will be the subject of the introductory paragraphs. The actual derivation and explicit form of the strain terms will be presented in sub-section 2.4.2.

2.4.1 - Introduction

This discussion about strain starts with an evaluation of the effects of strain on the periodic array of atoms. This chapter is only concerned with pseudomorphic growth of strained layers. In this case, it is assumed that there are negligible dislocations and that the lattice constant of the strained layer parallel to the interface takes that of the underlying substrate. InGaAs grown on a GaAs substrate will be compressively strained as InGaAs has a larger lattice constant than GaAs. In contrast, $In_xGa_{1-x}As$ grown on InP can either be compressively or tensile strained depending on the indium concentration of the InGaAs layer. Such in-plane distortion of the atomic lattice is however not the only distortion that will occur (fig.2.9).

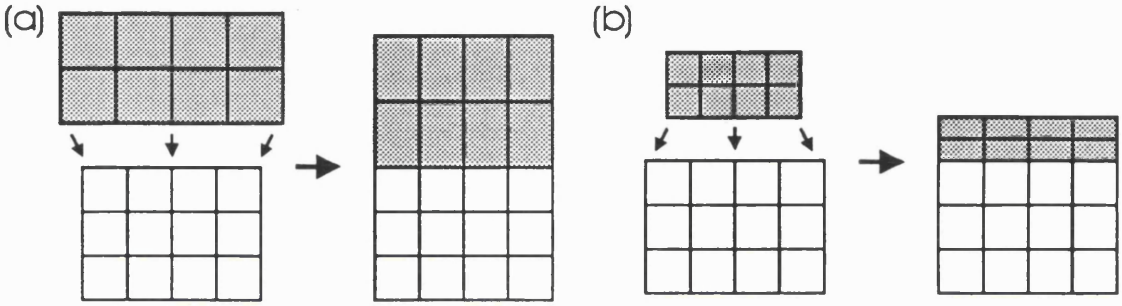


Fig.2.9 : Growth of a layer with a bulk lattice constant different from the substrate induces compressive (a) or tensile (b) strain. For pseudomorphic layers, the in-plane lattice constant matches the lattice constant of the substrate. The resulting change in volume of a primitive cell is partially compensated by an opposite change of the lattice constant perpendicular to the substrate.

Some of the basic notations and approaches to the strained layer problem are now introduced. First, the strain tensor $\varepsilon_{\alpha\beta}$ is defined as :

$$\underline{a}_i' = (1+\underline{\underline{\varepsilon}})\underline{a}_i \quad or \quad \underline{a}_i^{\alpha\,\prime} = a_i^{\alpha} + \sum_{\beta} \varepsilon_{\alpha\beta} a_i^{\beta} \tag{2.63}$$

where a_i is the lattice vector in the direction i in the unstrained and a_i' in the strained case. This strain tensor represents the three-dimensional lattice deformation. It is important to note that the symbol e is frequently used in the literature to describe strain instead of ε which has a slightly different definition,

$$e_{\mu\nu} = \begin{cases} \varepsilon_{\mu\nu} & when\ \mu = \nu \\ \varepsilon_{\mu\nu} + \varepsilon_{\nu\mu} & when\ \mu \neq \nu \end{cases} \tag{2.64}$$

In addition to the strain tensor, there is another tensor called the stress tensor $T_{\alpha\beta}$ which denotes the force in the direction α acting on a plane with the normal in the direction β (fig.2.10). Obviously, the introduction of strain, i.e. displacing the atoms from their preferred positions, will introduce stress. Hooke's law states that the stress is proportional to the strain so that, using symmetry arguments, one finds the following relation for zinc-blende materials [Sin93]

$$\begin{pmatrix} T_{xx} \\ T_{yy} \\ T_{zz} \\ T_{yz} \\ T_{zx} \\ T_{xy} \end{pmatrix} = \begin{pmatrix} C_{11} & C_{12} & C_{12} & & & \\ C_{12} & C_{11} & C_{12} & & & \\ C_{12} & C_{12} & C_{11} & & & \\ & & & C_{44} & & \\ & & & & C_{44} & \\ & & & & & C_{44} \end{pmatrix} \begin{pmatrix} e_{xx} \\ e_{yy} \\ e_{zz} \\ e_{yz} \\ e_{zx} \\ e_{xy} \end{pmatrix} \tag{2.65}$$

which shows that strain and stress are related via three elastic constants (C_{11},C_{12},C_{44}). All other elements in the matrix are zero.

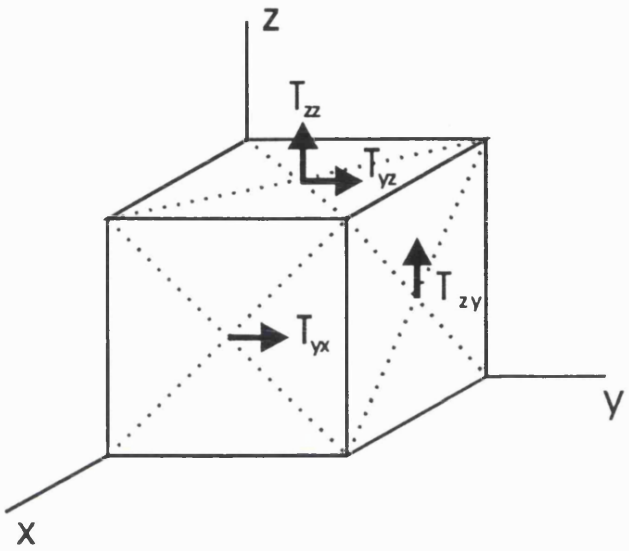


Fig.2.10: Examples of some stress elements. The vectors denote the direction of the force and the starting point shows the plane on which this force is acting.

As mentioned earlier, the lattice parameter parallel to the interface of the strained layer $a_{//}$ takes the lattice spacing of the underlying substrate a_s in the case of pseudomorphic growth, so that

$$\varepsilon_{//} = \frac{a_s}{a_l} - 1 = \varepsilon \tag{2.66}$$

Assuming no shear, one only has to work out the effect on the lattice constant perpendicular to the interface. As a plane with the normal parallel to the substrate is free to move, the position of the next plane of atoms will be such that there is no stress in any direction on this plane, i.e. $T_{xz} = T_{yz} = T_{zz} = 0$. Using (2.65) it is then straightforward to show that the strain tensor takes the following elements for [001] growth,

$$[001]\begin{cases} \varepsilon_{xx} = \varepsilon_{yy} = \varepsilon \\ \varepsilon_{zz} = -\dfrac{1}{\sigma^{001}}\varepsilon \\ \varepsilon_{xy} = \varepsilon_{yz} = \varepsilon_{xz} = 0 \\ \sigma^{001} = \dfrac{C_{11}}{2C_{12}} \end{cases} \tag{2.67}$$

where σ is the Poisson ratio which describes the ratio of the lattice deformation in the direction parallel and perpendicular to the substrate. The procedure for different growth directions is essentially the same but, as it involves a change in coordinate frame, is rather laborious. Results for other growth directions, as well as a well-documented outline of the procedure itself can for example be found in [Hin90].

An explicit form of the three-dimensional displacement of atoms under strain has been presented in terms of the elastic constants for the case of pseudomorphic growth. The corresponding effects on the bandstructure will be the next subject of discussion.

2.4.2 - The Pikus-Bir strain Hamiltonian

Previously, it was shown that the equilibrium positions of the atoms change under strain and as a result one expects these lattice changes to affect the bandstructure. As typical strain values (or dimensional changes) are only of the order of 1%, the changes in energy are expected to be relatively small. Furthermore, the microscopic form of the wavefunctions and the electron distribution in the strained case are anticipated to resemble the unstrained results. As a result, the effects of strain on the bandstructure are expected to be sufficiently small that they can be accounted for using perturbation theory.

The explicit form for this strain Hamiltonian was derived by Pikus and Bir [Pik74]. In contrast to the *k·p* derivation, it is less straightforward to write the strain effects in terms of a perturbation potential as even an infinitesimal amount of strain will cause the position of an ion to have a large shift when looking at relatively widely separated ions (fig.2.11). Or stated differently, unlike the *k·p* case there is no obvious *small* potential term that can account for the differences between the strained (U_ε) and unstrained potential (U_0).

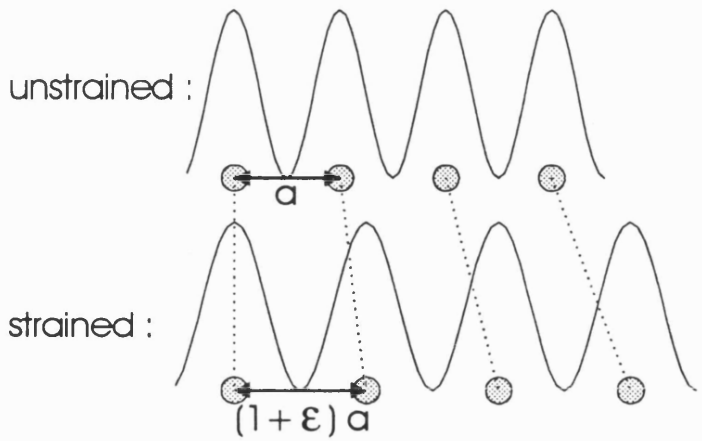


Fig.2.11 : Schematic picture of the effects of strain on the periodic potential.

The approach suggested by Pikus and Bir starts by introducing a coordinate transformation $(x,y,z)\rightarrow(x',y',z')$ to give the wavefunctions and periodic potentials in the strained case the same periodicity as those in the unstrained case

$$\underline{r}' = \left(1+\underline{\underline{\varepsilon}}\right)^{-1}\underline{r} \tag{2.68}$$

This transformation is directly related to the lattice distortion in terms of basis vectors (2.63). Obviously, the unprimed coordinate is the only real physical coordinate, but such a procedure provides a means to obtain an approximate description of the periodic potential and cell-periodic wavefunctions in the strained case. To lowest order,

$$\begin{aligned} U_\varepsilon\left(\left[1+\underline{\underline{\varepsilon}}\right]\underline{r}\right) &\approx U_0(\underline{r}) \\ u^\varepsilon_{n0}\left(\left[1+\underline{\underline{\varepsilon}}\right]\underline{r}\right) &\approx u_{n0}(\underline{r}) \end{aligned} \quad or \quad \begin{aligned} U_\varepsilon(\underline{r}) &\approx U_0(\underline{r}') \\ u^\varepsilon_{n0}(\underline{r}) &\approx u_{n0}(\underline{r}') \end{aligned} \tag{2.69}$$

In addition to this primed coordinate, primed operators are introduced which are equivalent to their unprimed partners apart from the fact that they are now working on r' rather than r. Retaining only the terms to first order in ε from now onwards, the primed momentum operator is related to its unprimed counterpart by

$$\underline{p} = \frac{\hbar}{i}\nabla = \left(1-\underline{\underline{\varepsilon}}\right)\underline{p}' \tag{2.70}$$

The underlying reason for the introduction of these operators that act on r' is that one ultimately wants to write the original Hamiltonian in terms of the coordinate r' after which its potential and zone-centre solutions can be approximated by (2.69). The relevance of these operators can be stressed by remarking that these primed operators are the *only* operators of relevance after having performed a coordinate transformation. As $p = \hbar k$, it follows that a coordinate change will affect the momentum k similarly to (2.70). Such was to be expected as the coordinate change can be regarded as a purely mathematical operation and that a physical quantity like a plane wave e^{ikr} should therefore retain exactly the same form. In this work, it was chosen not to introduce the primed momentum k' as one is ultimately interested in calculating $E(k)$ in the strained system and not $E(k')$ (note that [Bir74] not only chooses to introduce k' but also defines k and k' differently from (2.70)).

Using that the wavefunction takes the following form in the primed reference frame

$$u^\varepsilon_{nk}(\underline{r})e^{i\underline{k}\underline{r}} = u_{nk}\left(\left[1+\underline{\underline{\varepsilon}}\right]\underline{r}'\right)e^{i\underline{k}\left[1+\underline{\underline{\varepsilon}}\right]\underline{r}'} \tag{2.71}$$

one obtains that the original bulk strained Hamiltonian (top line) takes the following expression to lowest order in ε in terms of primed coordinates (bottom line) :

$$\begin{array}{ll} \left[\frac{p^2}{2m_0}+U_\varepsilon(\underline{r})\right] & u^\varepsilon_{nk}(\underline{r})e^{i\underline{k}\underline{r}} = E u^\varepsilon_{nk}(\underline{r})e^{i\underline{k}\underline{r}} \\ \left[\frac{p'^2}{2m_0}-\frac{\underline{p}'\underline{\varepsilon}\underline{p}'}{m_0}+U_\varepsilon\left(\left[1+\underline{\underline{\varepsilon}}\right]\underline{r}'\right)\right] & u^\varepsilon_{nk}\left(\left[1+\underline{\underline{\varepsilon}}\right]\underline{r}'\right)e^{i\underline{k}\left[1+\underline{\underline{\varepsilon}}\right]\underline{r}'} = E u^\varepsilon_{nk}\left(\left[1+\underline{\underline{\varepsilon}}\right]\underline{r}'\right)e^{i\underline{k}\left[1+\underline{\underline{\varepsilon}}\right]\underline{r}'} \end{array} \tag{2.72}$$

Rewriting this Hamiltonian, similar to (2.10), in terms of just the cell-periodic functions gives

$$\left[\frac{p'^2}{2m_0}+\frac{\hbar\underline{k}\underline{p}'}{m_0}+\frac{\hbar^2k^2}{2m_0}-\frac{\underline{p}'\underline{\varepsilon}\underline{p}'}{m_0}-\frac{\hbar\underline{\varepsilon}\underline{k}\underline{p}'}{m_0}+U_\varepsilon\left(\left[1+\underline{\underline{\varepsilon}}\right]\underline{r}'\right)\right]u^\varepsilon_{nk}\left(\left[1+\underline{\underline{\varepsilon}}\right]\underline{r}'\right)= E u^\varepsilon_{nk}\left(\left[1+\underline{\underline{\varepsilon}}\right]\underline{r}'\right) \tag{2.73}$$

Re-iterating that the potential and the cell-periodic functions in (2.73) have the same periodicity as in the unstrained case, the same expansion in zone-centre $|J,m_j\rangle$ states as outlined in section 2.3 can be performed which is the basis of the effective mass approach. A comparison of the various Hamiltonian elements reveals that first three elements in the Hamiltonian (2.73) are identical to those present in the unstrained Hamiltonian (2.10), the first being the kinetic energy and the last two being due to the $k\cdot p$ interaction. Thus, the important observation can be made that the transition from an unstrained to a strained system simply involves adding the fourth and fifth term to the Hamiltonian as well as any first order changes in the potential as compared to U_0.

From here onwards, the procedure is exactly the same as performed in the discussion of the $k\cdot p$ model. Again, the effective mass approach is adopted and the cell-periodic functions u^ε_{nk} are expanded in $|X\rangle$, $|Y\rangle$ and $|Z\rangle$. The following parametrised strain Hamiltonian for an expansion in $|X\rangle$, $|Y\rangle$ and $|Z\rangle$ is obtained

$$D=\begin{pmatrix} l\varepsilon_{xx}+m(\varepsilon_{yy}+\varepsilon_{zz}) & n\varepsilon_{xy} & n\varepsilon_{xz} \\ n\varepsilon_{yx} & l\varepsilon_{yy}+m(\varepsilon_{xx}+\varepsilon_{zz}) & n\varepsilon_{yz} \\ n\varepsilon_{zx} & n\varepsilon_{zy} & l\varepsilon_{zz}+m(\varepsilon_{xx}+\varepsilon_{yy}) \end{pmatrix} \tag{2.74}$$

However, rather than using l,m and n to parametrise this strain Hamiltonian, it is customary to use the deformation potentials a,b and d which render the [001] strain Hamiltonian in an easy form. These deformation potentials are defined as

$$\begin{aligned} a &= \frac{l+2m}{3} \\ b &= \frac{l-m}{3} \\ d &= \frac{n}{\sqrt{3}} \end{aligned} \tag{2.75}$$

Rewriting (2.74) into a $|J,m_j\rangle$ basis plus inserting the various strain components such as given in (2.67) for [001], the strain Hamiltonian takes the following form for the growth directions [001] [Hin90], [110] and [111] [Hen95]

$$H_\varepsilon = -\begin{pmatrix} \Delta E_g - \frac{1}{2}\delta & 0 & \chi & 0 \\ 0 & \Delta E_g + \frac{1}{2}\delta & 0 & \chi \\ \chi & 0 & \Delta E_g + \frac{1}{2}\delta & 0 \\ 0 & \chi & 0 & \Delta E_g - \frac{1}{2}\delta \end{pmatrix} \tag{2.76}$$

where

$$[100]\begin{cases} \Delta E_g &= 2a\dfrac{C_{11}-C_{12}}{C_{11}}\varepsilon \\ \delta &= 2b\dfrac{C_{11}+2C_{12}}{C_{11}}\varepsilon \\ \chi &= 0 \end{cases} \tag{2.77}$$

$$[110]\begin{cases} \Delta E_g &= a\dfrac{C_{11}-C_{12}+6C_{44}}{C_{11}+C_{12}+2C_{44}}\varepsilon \\ \delta &= \dfrac{1}{2}(b+\sqrt{3}d)\dfrac{C_{11}+2C_{12}}{C_{11}+C_{12}+2C_{44}}\varepsilon \\ \chi &= -\dfrac{1}{2}(\sqrt{3}b-d)\dfrac{C_{11}+2C_{12}}{C_{11}+C_{12}+2C_{44}}\varepsilon \end{cases} \tag{2.78}$$

$$[111]\begin{cases} \Delta E_g &= 4a\dfrac{4C_{44}}{C_{11}+2C_{12}+4C_{44}}\varepsilon \\ \delta &= 2\sqrt{3}d\dfrac{C_{11}+2C_{12}}{C_{11}+2C_{12}+4C_{44}}\varepsilon \\ \chi &= 0 \end{cases} \tag{2.79}$$

From this, three different strain effects can be seen:

- The energy gap changes under strain, generally widening for compressive strain and narrowing for tensile strain.
- The heavy hole and light hole band edges are separated under strain. In a compressively strained QW, the light hole bands are shifted away in energy from the lowest heavy hole level. For tensile strain the opposite is the case. Here, the light hole bands move towards the first

heavy hole level and can even cause them to cross, thereby forcing the first confined level to be a light hole state.

- Strain induces additional mixing between the heavy and light hole states in [110] structures, even at zero in-plane momentum.

At this point, the link between strain and the Hamiltonian is complete. Although the above discussion is restricted to just the interaction between the heavy and light hole bands, the presented Hamiltonian (2.76) can straightforwardly be extended to a 6-band model. As the parametrised Hamiltonians for the $k \cdot p$ (2.51) and strain interaction (2.74) show the same symmetry properties, the additional strain Hamiltonian elements required for such an extension are again directly related to the matrix elements present in a 4-band model.

Furthermore, the Hamiltonian (2.76) and its elements for the growth directions [001] and [111] now give a mathematical justification for the use of strain to enhance transport or optical properties. As shown in chapter 1, the interaction between heavy and light hole states causes the valence bands to be highly non-parabolic and generally increases the hole mass. Such interaction is obviously determined by the presence of non-zero off-diagonal elements R and S in the Hamiltonian (2.55). However as such a Hamiltonian represents a set of coupled equations, it is straightforward to see that the strength of this interaction is not just determined by the absolute values of these off-diagonal elements. Rather, the quantities of interest are $R/\Delta E$ or $S/\Delta E$ where ΔE represents the energy difference between the heavy and light hole solutions. The interactions between the heavy and light hole bands are therefore expected to reduce when applying compressive strain for [001] and [111] growth. In contrast, strain causes additional interactions in [110] structures which are expected to reduce performance.

2.4.3 - Piezo-electric fields

The strain contribution to the effective mass equations was presented in the previous sub-section 2.4.2. It was shown that changes in the periodicity of the lattice affect the bandstructure as the corresponding change in the periodicity of the wavefunctions affect the associated momentum. This obviously affects the outcome of the kinetic energy term in the Schrödinger equation, thereby modifying the dispersion. Although the change in dispersion will also affect the corresponding wavefunctions, the primarily affected quantity is the energy. To illustrate, zone-centre wavefunctions in a QW are expected to change very little, as the energy shifts due to strain ΔE_g, δ and χ (2.76) are generally much smaller than the well potential. Any change in wavefunctions

under strain at relatively large values of the momentum k, especially with respect to heavy/light hole composition, is entirely caused by the energy shift δ between the heavy and light hole levels and its effect on the $k \cdot p$ interaction. However, in addition to the interaction of strain through the momentum, there is another mechanism by which strain directly affects the wavefunctions through changes in the periodic potential U_ε (2.73) as compared to the unstrained potential U_0. This is the piezo-electric effect.

As seen in chapter 1, the Zinc-Blende lattice consists of two elements which each occupy their own fcc-lattice. Such a lattice does not have inversion symmetry so that the transformation $(x,y,z) \rightarrow -(x,y,z)$ does not transform the lattice into itself. As a result, one can not simply assume the electronic properties of two structures which are grown on a different face (see fig.2.12a) to be identical. And indeed, strain combined with a lack of inversion symmetry can cause the formation of macroscopic piezo-electric fields, which are caused by non-symmetric charge accumulation around the individual lattice elements (see fig.2.12b).

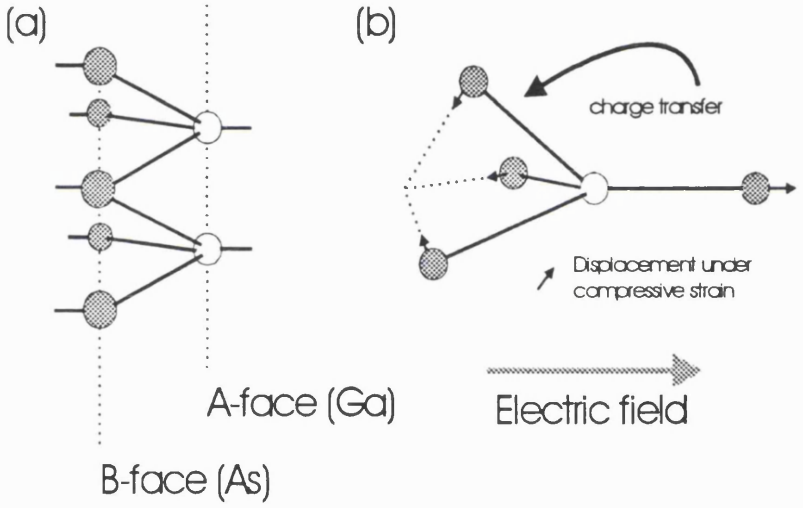


Fig.2.12 : Origin and substrate dependence of the piezo-electric effect. (a) Definition of the A- and B-face in GaAs. (b) Displacement of the various atoms under compressive strain and the accompanying charge transfer in the case of growth along (111).

The piezo-electric field in a Zinc-Blende structure is given by [Smi86]

$$E_i = -\frac{2e_{14}\varepsilon_{jk}}{\varepsilon_0 \varepsilon_r} \quad ; \quad i \neq j \neq k \tag{2.80}$$

Here, the subscript i denotes the direction of the field (i=x,y,z), e_{14} is the piezo-electric constant [Sin93] and ε_{jk} is (due to the restriction on j and k) an off-diagonal strain component. The polarisation parallel to the growth direction is strongest in a strained layer for growth along the [111] direction and zero for the other high symmetry directions [001] and [110] [Sun94,Car95]. The latter can easily be checked by using the explicit forms for the strain elements ε_{jk} as presented in [Hin90].

The piezo-electric field is of the order of 2.5 10^5 V/cm for $In_{0.2}Ga_{0.8}As$ grown on GaAs for the growth direction [111] [Sun94,Car95]. As the latter corresponds to a 250 meV shift in potential energy across a 100Å QW, it is obvious that such a field can have a serious effect on the wavefunctions and carrier distribution. As a result, it can be seen as a means to tailor the electron and/or hole distribution [Lu94,Hit95].

The latter ends this section on strain in semiconductors. It is worthwhile to re-iterate what has been achieved here :

- The necessary definitions to describe the effects of strain on the lattice have been presented as well as the physical interpretation.
- It was shown that the effects of strain can be included in the previously derived effective mass model by introducing strain elements in the one-electron Hamiltonian.
- A parametrised method similar to that employed in the derivation of the $k \cdot p$ model of section 2.2 was presented, and analytic forms of the strain Hamiltonian matrix elements were given for growth on [001], [110] and [111].

The aim of this chapter was to set up a bandstructure model capable of modelling compositionally varying structures. Starting with the simplest (Kane) approach, successive improvements were introduced and the extension of this model to compositionally varying structures was discussed. However, its full aim has not been achieved as the elementary question of how to connect a wavefunction across an interface in a compositionally varying structures has not yet been introduced. Rather than to introduce an ad-hoc approach, it was decided to treat this problem in detail in the next chapter, both because of its importance in certain material systems as well as to establish a clear insight into effective mass theory in terms of a tight-binding (orbital) approach.

Chapter 3 : The boundary condition problem

The previous chapter contains a broad overview of effective mass theory in terms of its development, derivation and interpretation. There, effective mass equations for the conduction and valence band were presented for the bulk case as well as for the case in which the electrons or holes experience a macroscopic varying potential. Although this potential was specifically introduced to account for spatial changes in the composition, the presented differential equations prove to have an intrinsic problem when applied to compositionally varying structures in which the mass parameters are position dependent. This is the so-called boundary condition problem.

So far in this thesis, as well as in the majority of the literature, the bulk description has conveniently been used to establish the effective mass equations, where necessary swapping vectors for operators to extend the model to compositionally varying structures. The ordering of the vectors in the bulk case is clearly unimportant as one is effectively working with scalars. However, the ordering becomes of crucial importance in compositionally varying structures where one is solving a set of differential equations with position dependent parameters. To illustrate, the ordering of the differentials as presented in the previous chapter can give complex energies which is clearly unphysical. At this point, one could have the impression that this is just a mathematical problem rather than a physical one, but it is not. One is essentially interested in solving the *fundamental* problem of how to connect an envelope function across an interface between two materials with different mass parameters.

This chapter concentrates solely on the boundary condition problem. The first section establishes the connection between operator ordering and envelope functions. The traditional approach to overcome non-real eigenvalues is illustrated using the one-band conduction band Hamiltonian. Various objections to this approach are discussed, highlighting the need for a more fundamental approach. The second section discusses Burt's 'exact envelope function theory' [Bur87,Bur92] and presents the results obtained by Foreman [For93] for the boundary conditions for growth along [001].

This chapter can be seen as an introduction to the the results chapter (chapter 6) in which the discussion on the boundary conditions will continue. There, a formalism to extend the Burt-Foreman framework to non-[001] growth directions is presented, as well as a set of analytic boundary condition rules that are to replace the conventional symmetrisation rule. The relevance of

adopting these 'new' boundary conditions is also highlighted in chapter 6, where it is shown that the choice of boundary conditions can have a strong effect on the dispersion or on the outcome of mobility calculations in highly doped *p*-type materials such as InGaAs-InP.

3.1 - Introduction

The question of how to connect the envelope function across an interface between two materials with different mass parameters, or the boundary condition problem, has been a major cause of controversy in effective mass theory . Its existence originates from the fact that the effective mass equation for compositionally varying structures carries a combination of differential operators as well as positionally dependent mass parameters. Clearly, the position of these parameters with respect to the operators (the operator ordering) is expected to influence the theoretical dispersion relations. Worse, particular choices of ordering can even lead to unphysical results.

To illustrate the latter, as well as to connect the operator ordering to the continuity of the envelope functions, the effective mass model for the conduction band within the 1-band approximation is examined first. In this approximation, the conduction band effective mass equation takes the form (2.50)

$$\left[-\frac{\hbar^2}{2m^*}\frac{\partial^2}{\partial z^2}+V(z)+\frac{\hbar^2}{2m^*}\left(k_x^2+k_y^2\right)\right]F_s(z)=E\,F_s(z) \tag{3.1}$$

In the present form, a position dependent mass in the kinetic energy term causes the Hamiltonian to become non-Hermitian at the interfaces. This is clearly a violation since the solutions are required to have real energies. In addition, it violates the requirement of current conservation (see [Gas74] for the definition of the current flux). Obviously, the order of the differential operators with respect to m^* should be replaced by a more physical combination.

Investigations into what operator ordering to use is however not just a mathematical problem, but corresponds (as mentioned earlier) to the fundamental question of how to connect the envelope functions across an interface. These boundary conditions are directly linked to the Hamiltonian and can be obtained by integrating it over an infinitesimal interval across an interface. To clarify this, the effects on the continuity of the envelope function *F(z)* are evaluated for three (at this point arbitrary) choices of operator ordering denoted by (a),(b) and (c),

$$(a)\quad \frac{1}{m^*}\frac{\partial^2}{\partial z^2} \qquad\qquad \Rightarrow\ F(z)\ and\ \frac{\partial F(z)}{\partial z}\ continuous$$

$$(b)\quad \frac{\partial}{\partial z}\frac{1}{m^*}\frac{\partial}{\partial z} \Rightarrow \frac{1}{m^*}\left(\frac{\partial}{\partial z}F(z)\right) continuous \Rightarrow F(z)\,continuous, \frac{\partial F(z)}{\partial z}\ discontinuous$$

$$(c)\quad \frac{\partial^2}{\partial z^2}\frac{1}{m^*} \Rightarrow \frac{\partial}{\partial z}\left(\frac{1}{m^*}F(z)\right) continuous \Rightarrow F(z)\ and \frac{\partial F(z)}{\partial z}\ discontinuous$$

Fig. 3.1 is a graphical illustration of the effects of these particular choices for the operator ordering on the continuity of the envelope function.

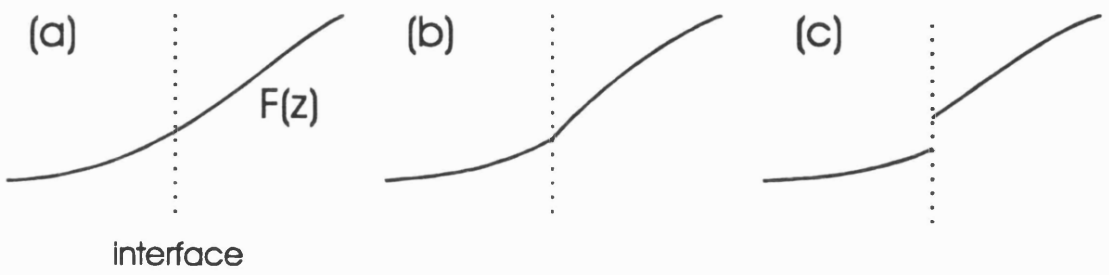


Fig.3.1 : Various schemes to connect an envelope function F(z) across an interface. (a) Both F and ∂F/∂z continuous; (b) F continuous, ∂F/∂z discontinuous; (c) both F and ∂F/∂z discontinuous.

At this point, one has to realise that there are an infinite number of choices possible for the ordering as one could for example split up the mass term and write the kinetic energy in the form $(m^*)^{\alpha}\cdot\partial^2/\partial z^2\cdot(m^*)^{\beta}$, $\alpha+\beta=-1$. Clearly, one is looking for a choice which renders the Hamiltonian in a Hermitian form and satisfies the requirement for current conservation. However, the question is if such an heuristic approach can give sufficient restrictions to obtain unambiguous boundary conditions. Furthermore, as each operator ordering corresponds to a specific boundary condition, one has to investigate the physical interpretation of the corresponding continuity requirement of the envelope functions.

3.2 - The conventional approach

As seen in the previous section, the conduction band effective mass equation for compositionally varying structures is not Hermitian in its presence form, nor conforms to the requirement of current conservation. Both physical objections can be overcome by replacing the kinetic energy operator by its symmetrised form

$$-\frac{\hbar^2}{2m^*}\frac{\partial^2}{\partial z^2} \rightarrow -\frac{\partial}{\partial z}\left(\frac{\hbar^2}{2m^*}\right)\frac{\partial}{\partial z} \qquad (3.2)$$

which corresponds to the example denoted by (b) in the previous section. Integrating this symmetrised form of the kinetic energy across an interface, the familiar boundary condition of continuity of $(1/m^*)F_s(z)$ is obtained.

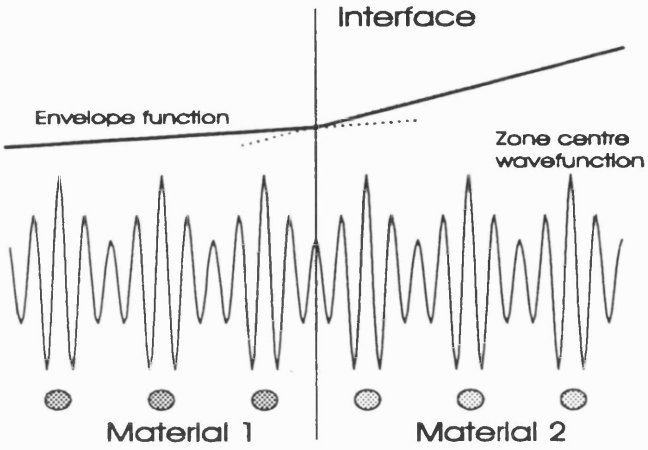


Fig.3.2 : Plot of a fictitious periodic $|S\rangle$ zone centre wavefunction and an envelope function near an interface.

Although at first sight this might seem a logical approach to circumvent the problems at the boundary, one can question the validity of the procedure. There are at least three possible objections to it :

- At this point there is absolutely no theoretical evidence that the symmetrisation procedure of (3.2) is the correct one as there are still an infinite number of choices possible [Mor84].
- The integration over the interface used to obtain the boundary conditions is not expected to be valid, since the k-space derivation of the effective mass equations [Lut55] assumed the potential V to be 'smooth' and a discrete jump at an interface can hardly be regarded as smooth.
- The kink in the envelope functions at the interface introduced by this boundary condition can not literally be present. Since the wavefunction is given by $\psi = F_s(r)|S\rangle$ this would give an infinite second derivative and therefore infinite energy at the interface unless $|S\rangle$ is zero at this interface (fig.3.2). The latter is however completely unrealistic.

The last point is disastrous from a conceptual and physical point of view. Despite this controversy and lack of knowledge, the same symmetrisation procedure was commonly used to make the valence band Hamiltonian Hermitian. The procedure is however even more questionable in the latter case, as such a valence band Hamiltonian features a number of matrix elements that contain a combination of first and second order differentials [see for example (2.51)]. Adding to the controversy, the requirement of real eigenvalues does not necessarily mean that each element by itself has to be Hermitian.

Despite this, effective mass theory seemed to work extremely well so that the boundary condition problem was considered more a fundamental than a practical problem. As the initial material systems of interest such as AlGaAs-GaAs feature a fairly constant set of mass parameters, a particular choice of operator ordering is not expected to have a large effect on the theoretical dispersion relations. However, this is less obvious for InGaAs-InP or emerging material systems such as CdTe-compounds, which feature considerable changes in mass parameters. As a result, an in depth investigation of the origin of the boundary conditions was long overdue. Although various attempts to solve this problem were made, insight in the problem of how to connect envelope functions across an interface was not provided until the development of an 'exact' envelope function theory by Burt [Bur87,Bur92]. The next section presents (some) of the results obtained in that work.

3.3 - Exact envelope function theory

The effective mass equations in chapter 2 were essentially derived for bulk structures only. As a result, no specific arrangements were made to allow for changes in effective mass parameters. It is therefore no surprise that only general arguments (such as the requirement that the Hamiltonian must be Hermitian) have so far been employed to obtain the correct operator ordering, as the original derivation of the effective mass equations by itself can not supply such information. However, as it was shown that these arguments cannot provide unambiguous boundary conditions, not even in the relatively simple conduction band case, it is clear that a more fundamental approach is needed.

Obviously, a derivation of the effective mass equations which specifically allows for changes in the composition and mass parameters is required. Such an approach was first provided by the development of an exact envelope function theory by Burt [Bur87,Bur92]. The thorough treatment of the problem in both papers is rather mathematical and too extensive to be repeated, but as suggested by himself [Bur96,unpublished] there is an alternative and less general approach to show the principles involved.

Starting again from the expansion (2.39)

$$\psi(\underline{r}) = \sum_j F_j(\underline{r})\, u_{j0}(\underline{r}) \tag{3.2}$$

one would now like to explicitly allow for changes in the mass parameters as well as link these to the microscopic structure of the semiconductor materials. One now makes an assumption involving

the origin of the effective mass parameters. From the bulk Löwdin-perturbation term (2.37) it follows that any change in mass parameters , either being m^* or γ_1, γ_2, γ_3, when crossing an interface that separates material 1 from material 2 can either be caused by

- differences in the *form* of the band edge orbitals $u_{j0}(r)$ in both materials, i.e. differences in the momentum matrix elements.
- differences in the *energies* E_j at which the various bands are located.

Following Burt's suggestion, the restriction is made that only the *energies* of the band edge states change when going from material 1 to material 2 and not the orbitals itself. A compositionally varying structure will from now on be described by position dependent band energies E_j but invariant cell-periodic functions u_{j0}. The important thing to note is that, in contrast to the approach taken in the previous chapter, one does not now implicitly assume all bands to be equally influenced by a change in composition. Obviously, the final equations will still feature a compositionally varying term such as $V(r)$ that describes the energy as a function of position of the band under consideration (either being the conduction or valence bands). However, changes in the mass parameters will now explicitly be allowed through variations in energies of the remote bands (fig.3.3).

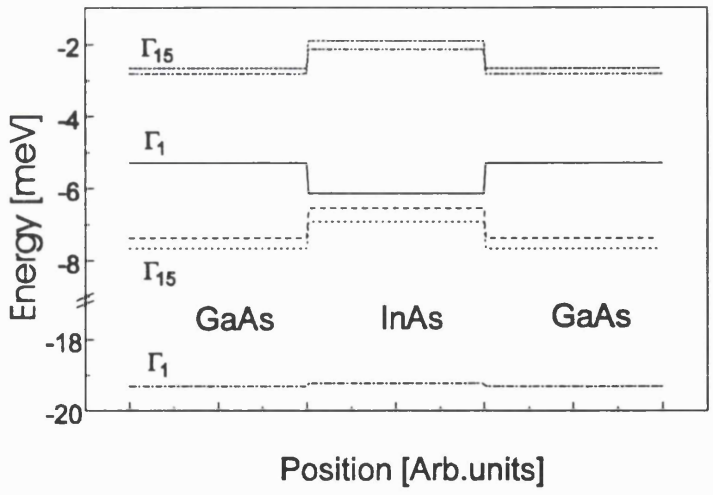


Fig.3.3 : Band line-up of the lowest Γ-point states in a GaAs-InAs QW (excluding strain, data from [Wal89,Lan82], arbitrary zero-point of origin). The solid line corresponds to the conduction band and would have modelled for the piecewise constant electrostatic potential V(r) that was introduced in the previous chapter.

Using the above approximation, the derivation of the effective mass equations (2.33-2.38) is now repeated where particular attention is paid to the operator ordering. Inserting the expansion (3.2) in the general Hamiltonian and using that the cell-periodic wavefunctions u_{j0} correspond to the bulk eigenfunctions at energies E_j, one obtains

$$\sum_j \left[E_j(r)-E\right]F_j(r)u_{j0}(r)-\frac{\hbar^2}{2m_0}\left(\nabla^2 F_j(r)\right)u_{j0}(r)-\frac{\hbar^2}{m_0}\nabla F_j(r)\cdot\nabla u_{j0}(r)=0 \qquad (3.3)$$

were $E_j(r)$ is position dependent as one now specifically allows for compositional changes. Again the assumption is made that the expansion can be divided into two classes, a dominant set $|l\rangle$ consisting of Γ_1 or Γ_{15} orbitals, l=1..r, which requires explicit treatment and a remote set $|\mu\rangle$, μ=1..s, which can be included using perturbation theory. Taking the inner product with respectively a dominant and a remote state gives,

$$\begin{aligned}\left(E_l(r)-E-\frac{\hbar^2\nabla^2}{2m_0}\right)F_l-\frac{i\hbar}{m_0}\sum_{\mu'}p_{l\mu'}\nabla F_{\mu'}=0\\ \left(E_\mu(r)-E-\frac{\hbar^2\nabla^2}{2m_0}\right)F_\mu-\frac{i\hbar}{m_0}\sum_{\mu'}p_{\mu\mu'}\nabla F_{\mu'}-\frac{i\hbar}{m_0}\sum_{l}p_{\mu l}\nabla F_l=0\end{aligned} \qquad (3.4)$$

where, as a result of the approximation that the orbitals are identical throughout the structure (but not their energies), the momentum elements $p_{l\mu'}\left(=\langle u_{\mu'0}|p|u_{l0}\rangle\right)$, $p_{\mu\mu'}$ and $p_{\mu l}$ do not have a position dependence. Note that the inner product $p_{ll'}$ is zero because of symmetry considerations as the explicit expansion in $|l\rangle$ (the dominant states) is restricted to just Γ_1 or Γ_{15} in this work (1-band conduction band, 4- or 6-band valence band model). Now, using that the remote states are relatively far separated in energy from the energy region of interest, it follows that E_μ-E is the dominant term in the second equation of (3.4). Neglecting the kinetic energy term as well as the term proportional to $p_{\mu\mu'}$ as $F_\mu \ll F_l$ results in the approximation

$$F_\mu(r)\approx\frac{i\hbar}{m_0}\sum_l\frac{p_{\mu l}}{E_\mu(r)-E}\nabla F_l \qquad (3.5)$$

Inserting this in the expression for the dominant states, the following system of coupled differential equations is obtained

$$\left(E_l(r)-\frac{\hbar^2\nabla^2}{2m_0}-E\right)F_l+\frac{\hbar^2}{m_0}\left(\sum_{l,\mu}\nabla_\alpha\frac{p^\alpha_{l\mu}p^\beta_{\mu l}}{E_\mu(r)-E}\nabla_\beta\right)F_l=0 \qquad (3.6)$$

The second part of the above equation can be seen as the Löwdin interaction term and is very similar to (2.37) apart from the fact that it is now carrying operators rather than vectors. To be able to compare the above equation to the original effective mass, it is rewritten as

$$\sum_{j'}\left[\left(\frac{\hbar}{i}\frac{\partial}{\partial x_\alpha}\right)D_{jj'}^{\alpha\beta}\left(\frac{\hbar}{i}\frac{\partial}{\partial x_\beta}\right)+E_j(\underline{r})\delta_{jj'}\right]F_{j'}(\underline{r})=EF_j(\underline{r})$$

$$D_{jj'}^{\alpha\beta}=\frac{1}{2m_0}\delta_{jj'}\delta_{\alpha\beta}+\frac{1}{m_0^2}\sum_{\nu}\frac{p_{j\nu}^{\alpha}p_{\nu j'}^{\beta}}{E-E_\nu} \qquad (3.7)$$

The important difference to (2.42) is that an ***exact***[1] and unambiguous operator ordering has been obtained. Clearly, one has to keep track of the operator ordering during the parametrisation of the effective mass model to be able to derive the correct boundary conditions. Simply lumping the contributions of all remote bands on the mass parameters together (to obtain m^* or $\gamma_{1,3}$) and neglecting the operator ordering as done in (2.44,2.47) cannot provide the correct boundary conditions as there are just not sufficient basic quantum mechanical constraints to subsequently provide unambiguous operator ordering.

For an expansion in a single band, i.e. for the conduction band case, using the ordered form (3.7) results in the same ordering as one would have obtained by using the conventional symmetrisation rule (3.2). The derivation of the boundary conditions for the valence band are however more difficult to obtain, and will be discussed in the next section.

This chapter then ends with some concluding remarks about the effects of the employed approximations in the derivation (3.2-3.7) and comments about more recent approaches towards the boundary condition problem.

3.4 - Explicit form of the boundary conditions on [001]

The correct boundary conditions for growth along the [001] direction were derived by Foreman [For93]. When working with the *ordered* form of the Löwdin interaction (3.7), one finds that it is convenient to replace the previously used Luttinger parameters, which were chosen to relate directly to the effective masses involved, by a new set of three parameters which explicitly identify the contribution corresponding to remote bands of a specific symmetry Γ_1, Γ_{15} or Γ_{12} (neglecting the small contribution of Γ_{25}) [For93]. Using the tight-binding picture as done in subsection 2.2.2 where the atomic orbitals upto *d* corresponding to the various symmetry groups were presented, it is straightforward to show that the only non-zero contributions to the band parameters *L,M,N* in the parametrised Hamiltonian (2.47) are of the form

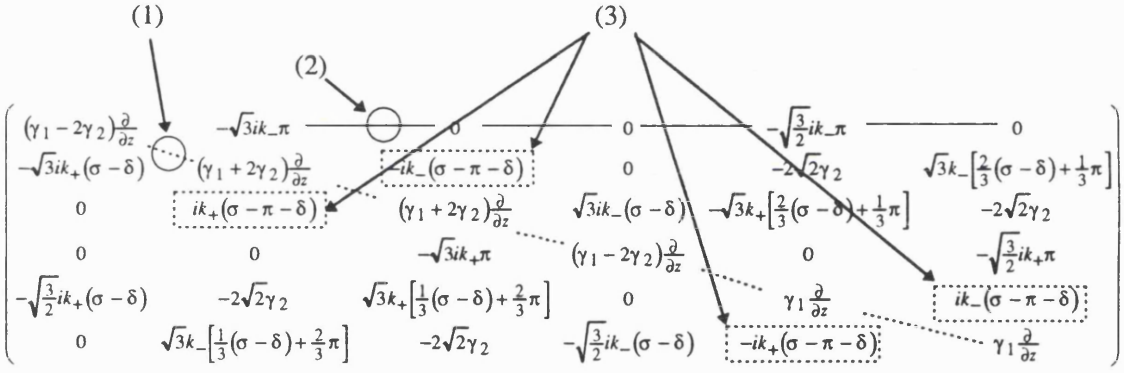
[1] Of course, in our derivation there is also the assumption that any wavefunction can be written as a product of a zone-centre solution and an envelope function without any constraints.

$$\sigma = (1/3m_0)\sum_{v}^{\Gamma_1} \left|\langle X|\hat{p}_x|v\rangle\right|^2 / (E_v - E)$$
$$\pi = (1/3m_0)\sum_{v}^{\Gamma_{15}} \left|\langle X|\hat{p}_y|v\rangle\right|^2 / (E_v - E) \qquad (3.8)$$
$$\delta = (1/6m_0)\sum_{v}^{\Gamma_{12}} \left|\langle X|\hat{p}_x|v\rangle\right|^2 / (E_v - E)$$

Each of these interactions will generate terms in the Hamiltonian with a specific ordering of the differential parameters with respect to the band parameters. Identifying the contribution to the Luttinger parameters γ_1, γ_2, γ_3, one finds

$$\gamma_1 = -1 + 2\sigma + 4\pi + 4\delta$$
$$\gamma_2 = \sigma - \pi + 2\delta \qquad (3.9)$$
$$\gamma_3 = \sigma + \pi - \delta$$

One can now proceed as done in (2.47-2.52) and derive the *ordered* Hamiltonian elements for a specific growth direction using the appropriate angular momentum basis for growth on [001]. Integrating this ordered Hamiltonian across a boundary, one then obtains the boundary condition of continuity of F and DF where D is given by [For93]



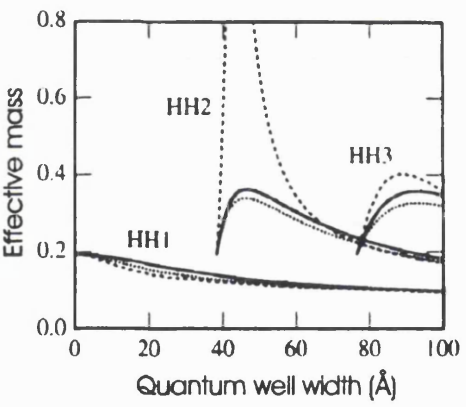
(3.10)

The above (correct) boundary condition matrix differs considerably from what one would have obtained had one used the conventional symmetrisation procedure. This boundary condition matrix exhibits

1) a considerable asymmetry with respect to the diagonal, something one would not have found using a symmetrisation procedure.
2) reduced coupling between the heavy hole band and the light/split-off hole bands as these interactions do not feature the *dominant* contribution from the remote states with s-type symmetry (σ)
3) a new coupling term, coupling both light hole states as well as the split-off-hole states.

As a result, significant changes in the dispersion relations or related quantities like subband effective masses are expected when adopting the new boundary conditions. For example, it was shown in [For93] that, using the conventional symmetrised boundary conditions, the calculated zone-centre effective mass for the second subband in an $In_{0.2}Ga_{0.8}As/GaAs$ QW can have triple the value of that obtained using the new boundary conditions (fig.3.4). Furthermore, it was shown in [Men94] that 4- and 6-bands calculations which do not implicitly include the conduction band improve significantly when adopting the new boundary conditions when compared to an 8-band model (where the higher *s*-orbitals are explicitly included) (fig.3.5). The latter results clearly show that the new boundary conditions are more physical than the old ones, as one does not expect :

- the effective mass to be so sensitive with respect to well width
- the valence band dispersion to display large variations when comparing the various (4-, 6-, 8-band) models. Obviously, one expects the accuracy to increase when including more bands in the expansion but with a gentle convergence.



▲ *Fig.3.4 : Zone centre heavy hole effective mass for an $In_{0.2}Ga_{0.8}As/GaAs$ QW. Shown are the calculated results using three different boundary conditions, Burt-Foreman (solid), symmetrised (dashed) and uncoupled (dotted) (after [For93]).*

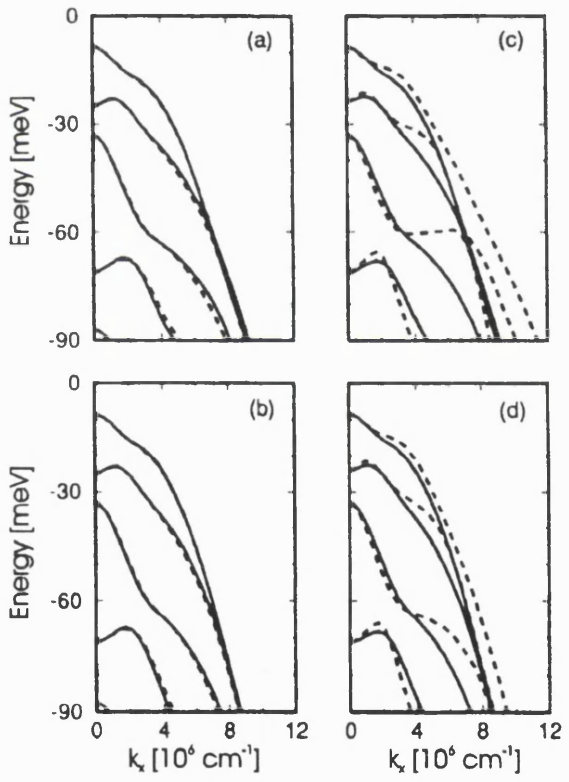


Fig.3.5 : Valence band dispersion for a 100Å $In_{0.16}Ga_{0.84}As_{0.14}Sb_{0.86}/Al_{0.4}Ga_{0.6}As$ QW (solid:Burt-Foreman boundary conditions, dashed: symmetrised). The following bands were explicitly included in the expansion: (a) CB,HH,LH,SO; (b) CB,HH,LH; (c) HH,LH,SO; (d) HH,LH. (after [Men94])

At this point, the need for implementing the new boundary conditions has been established as it was shown that valence band effective models that do not explicitly include the interaction with the conduction band states (such as the 4- and 6-band models that are used throughout this thesis) are very sensitive to the choice of boundary conditions. Despite the numerical evidence and new theoretical insight in the origin of the boundary conditions, it was found that the symmetric boundary conditions were still being used in the literature when looking at non-[001] growth as the exact envelope function theory had so far only been applied to [001]. This was addressed in two papers [Sta97,Dal98], in which the first one provides the boundary conditions for [110] growth and the second one extends the formalism to arbitrary growth directions. These theoretical results will be discussed in detail in the next chapter. The relevance of using the correct operator ordering will be highlighted again in chapter 6 where dispersion relations are presented for non-[001] growth directions for the various boundary conditions, along with its effects on theoretical hole mobility calculations.

3.5 - Concluding remarks

One may ask if the approach taken in section 3.3 is truly the final word on the boundary condition problem. Here, a rather severe approximation of invariance of the cell-periodic wavefunctions was used to obtain the final effective mass equations. Furthermore, the new boundary conditions again result in a kink in the wavefunction at the interface which, as explained in section 3.2, was exactly one of the reasons to question the traditional approach.

However, section 3.3 only contains a simplified derivation of Burt's exact envelope theory, whose strength lies in the complete circumvention of the boundary condition problem by using a set of envelope functions and derivatives that are constant even at an abrupt interface. The (mathematically rigorous) approach takes the change in each zone-centre wavefunction in consideration although the basis functions remain the same. So in this approach the momentum matrix elements are truly invariant, a feature that was exploited in the derivation presented here. The actual derivation by Burt indicated that the effects of a change in zone-centre wavefunctions are only of small importance. Furthermore, the kink in the envelope functions was shown to result from the effective mass approach of eliminating small envelope functions in favour of the dominant ones (3.4,3.5) and, more importantly, was shown to approximate a rapid change in the derivative in the region of the interface as obtained using the exact theory.

Despite the satisfactory results of the latter model, more sophisticated approaches have followed in the literature [For96] where the effects of position dependent material properties are taken into account in even more detail. The latter approach generates new interface effects and even the loss of continuity of the envelope functions at an interface. At this point, one has to ask if such a new model is actually called for. There are various arguments that can be made against this :

- Effective mass theory is generally considered a engineering method, i.e. a relatively straightforward model to gain insight. One could argue that the latter extension, simply due to the loss of continuity of the envelope functions at an interface, turns the whole model into a cumbersome mathematical exercise.
- Foreman's derivation already had to neglect the interaction with the Γ_{25} bands as there are simply not enough parameter constants available $(\gamma_1,\gamma_2,\gamma_3)$ to uniquely determine the individual contributions.
- Those parameters, as well as band-offsets and other parameters, are only known with limited accuracy. Realising that the old effective mass theory worked extremely well or that the 85:15 offset rule in AlGaAs managed to survive for a decade [Wei91], it is questionable if any improvement beyond Burt's approach is worthwhile at this point.

In conclusion, the exact envelope function theory by Burt's is considered in this thesis to be the basic framework to describe electrons and holes in compositionally varying structures. As mentioned in section 3.4, the effects of choosing particular boundary conditions on the theoretical dispersion relations will be shown in the Results chapter (chapter 6).

Chapter 4 : Carrier transport in heavily doped structures

The previous two theory chapters only concerned themselves with bandstructure theory. A general effective mass bandstructure model was set up in chapter 2 whereas chapter 3 dealt specifically with the extension to compositionally varying structures. Although several techniques were mentioned in the introductory chapter to measure the bandstructure directly, the ultimate aim in this thesis is to use the derived bandstructure model to optimise *p*-type semiconductor structures for use in power FET-applications.

Although some qualitative conclusions about the expected (electrical or optical) performance can be drawn by comparing different bandstructure diagrams (see chapter 1), one would ultimately want to be able to make quantitative predictions about the carrier mobility. To realise this, the derived bandstructure model has to be extended with a scattering model. The latter is the main subject of this chapter.

Various scattering mechanisms exist. To name a few [Rid93] :

- Phonon scattering, i.e. scattering due to thermal vibrations of the crystal lattice. Divided into
  - optical and
  - acoustic phonon scattering.
- Alloy scattering, due to fluctuations in the composition in a ternary compound
- Interface roughness scattering
- Impurity scattering
  - of neutral impurities
  - of charged impurities

each which requires a different scattering model. This thesis is concerned with highly channel-doped structures such that ionised impurity scattering is expected to be the only relevant scattering mechanism at low temperatures (77 Kelvin). At higher temperatures, optical phonon scattering is expected to contribute to the scattering processes. However, ionised impurity scattering is the only scattering mechanism that is going to treated in detail in this work, as it is expected to be the dominant mechanism even at room temperature.

This chapter is set up as follows. The following section contains an introduction to scattering. Here, scattering is described in terms of momentum vectors, scattering angles, and momentum loss. The effect of scattering on the mobility is qualitatively explained, as well as the effect of screening. Section 4.2 describes the Thomas-Fermi scattering model which is extended to the more sophisticated RPA (Random Phase Approximation) in section 4.3. The role of band mixing is examined (4.4) and the explicit forms of the relevant equations of the scattering model are presented for structures with one (4.5) or multiple (4.6) occupied subbands. This chapter then ends with a discussion on the definition of the effective mass parameter in the scattering model and its relation to the bandstructure. The latter is important as the two main existing definitions give different results when applied to non-parabolic bands of which the valence bands are an extreme example. A simple model to account for non-parabolicity effects in the conduction band is introduced and its relevance is highlighted for highly doped *n*-channel InGaAs-InP structures.

4.1 - Introduction

Quantum mechanics states that electrons in a perfect periodic lattice can move without any electrical resistance. Such behaviour, known as ballistic transport, is in contrast to what one intuitively would expect as the moving electron has to pass ions and other electrons on its passage. A perfect periodic potential does not cause electrons to scatter, all resistivity is solely due to deviations of the potential from being ideal, for example due to lattice vibrations or the presence of ionised impurities (more examples were given in the previous section).

Elastic scattering processes (such as ionised impurity scattering) can only involve states at the Fermi-energy in degenerate systems at low temperatures. Fig.4.1a schematically shows the effect on the momentum when an electron moving in the x-direction under the influence of an electric field E_x is scattered from its initial state with momentum k in the first subband to a final state with momentum k' in the second subband. In the process, some of its momentum (i.e. kinetic energy) is converted to potential energy. In addition, the momentum of interest (i.e. parallel to the electric field) is further reduced due to the angular deflection. The momentum loss q associated with such an event is given by

$$q = \left[k^2 + k'^2 - 2kk'\cos(\alpha)\right]^{1/2} \tag{4.1}$$

where α is the angle over which the electron is scattered (see fig.4.1b). So far this section has only mentioned interband scattering, but scattering generally mainly occurs within the band itself (intra-band scattering). A schematic picture of this is given in fig.4.1c.

Scattering obviously affects the collective motion of carriers under the influence of an external force such as an electric field. As illustrated in chapter 1, an electric field will cause the Fermi-contour to shift parallel to the direction of the field, resulting in a net momentum and current. The displaced Fermi-contour may be maintained in a steady state, governed by the scattering time τ as the mean increment in velocity between two scattering events is given by $qE\tau/m^*$. By calculating this scattering time, one can thus determine the response of a system. More specifically, such an approach would provide theoretical insight in the conductivity and carrier mobility in a semiconductor structure. The carrier mobility is the main parameter of interest in this work, as it was shown in chapter 1 that the relatively low hole mobility in GaAs is the reason for the poor performance of *p*-type FETs.

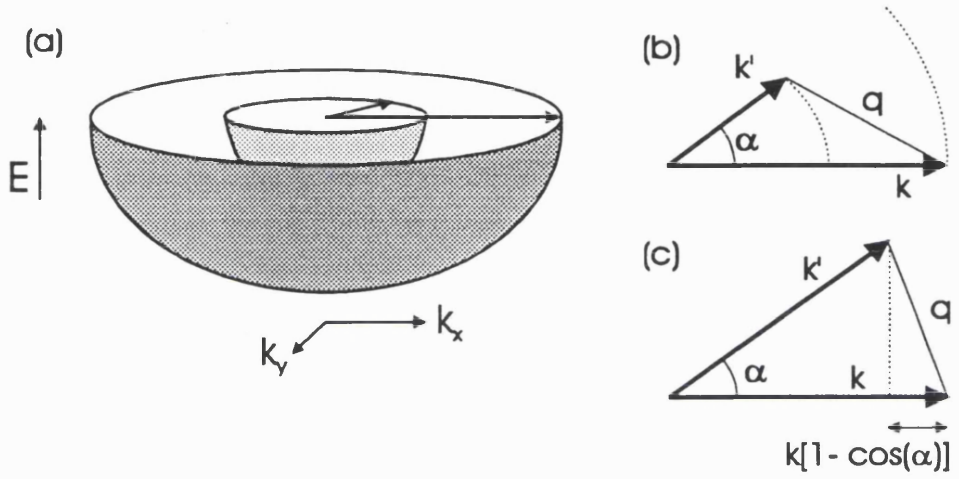


Fig.4.1 : (a) A 3-dimensional representation of an interband scattering process. (b) Schematic view of the momentum loss q with respect to the initial (k) and final (k') momentum and angle of deflectance. (c) As (b) but now for an intra-band scattering process. Highlighted is the momentum loss along k due to the transition.

Although ionised impurity scattering can conveniently be pictured as a Rutherford scattering process (fig.1.6), considerable complications arise from the fact that the electron is moving in surrounding 'sea' of particles. As a result, the scattering process is complicated as one is really looking at a many-body process. Obviously, the impurity does not only interact with the particle under observation, but also with its surroundings. The latter gives rise to screening (as will be discussed in the next section) which needs to be specifically included in the evaluation, as neglecting it would result in a significant overestimation of the scattering. Screening clearly plays an essential role in ionised impurity scattering and will be included using the Random-Phase-Approximation (RPA) which has proved to be very successful in the literature [Fle90,Hai95,Hai96,Hai97]. There are however various ways to account for screening, the simplest way being so-called Thomas-Fermi approximation (TFA) which, although not very accurate, gives a good insight in the screening mechanism. The next section will now introduce the

physics behind screening by looking at the relatively simple TFA. Section 4.3 will then extend this model to the more accurate RPA.

4.2 - Ionised impurity scattering in the Thomas-Fermi approximation

As discussed in the previous section, the presence of an ionised impurity in an otherwise perfect crystal introduces a potential fluctuation which can cause the electrons to scatter. This electrostatic potential fluctuation is equal to the Coulomb potential of the impurity itself (ϕ^{ext}) and the accompanying screening potential (ϕ^{s}) which can be seen as a response of the system to the introduction of the charged impurity. The latter potential literally screens the impurity from the passing electrons (hence its name) thereby dramatically reducing the scattering rate.

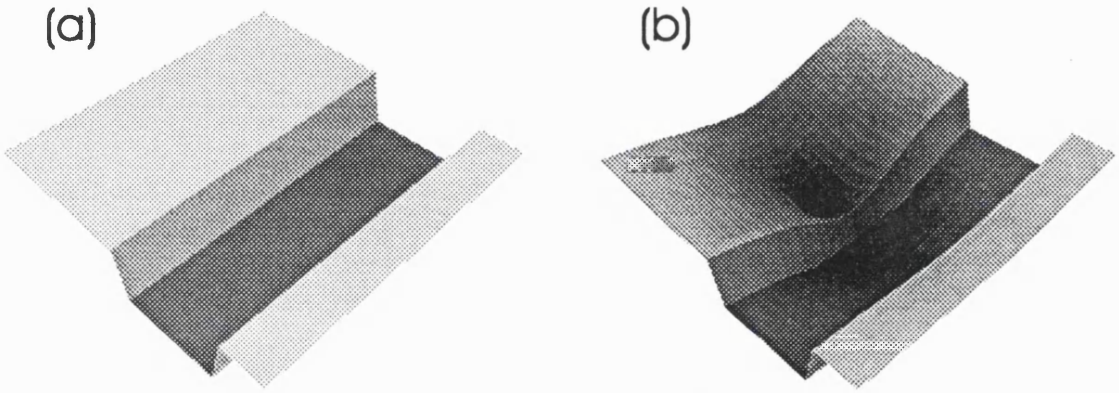


Fig.4.2 : Illustration of the effect of a remote impurity on the potential felt by a passing electron.

Fig 4.2 schematically shows the effect of a remote impurity on the potential felt by a passing electron. Obviously, the electrons in the channel will be Coulomb attracted so that one expects a higher electron density near the impurity as compared to elsewhere in the channel. The latter is the screening charge. To first order, the wavefunctions in the QW do not change but only the eigenvalues of the quantised levels, $i=1..n$, so that the induced charge is given by [Ste67]

$$\begin{aligned} \rho_{ind}(r_{//}, z) &= \sum_i \Delta E_i(r_{//}) \frac{\partial n_i}{\partial E} |f_i(z)|^2 \\ &= \sum_i \frac{m^* e}{\pi \hbar^2} \bar{\phi}_{ii}(r_{//}) \cdot |f_i(z)|^2 \end{aligned} \tag{4.2}$$

where the summation over i runs over all occupied subbands with corresponding envelope function $f_i(z)$. Here, the coordinate $r_{//}$ represents the in-plane vector $(x,y,0)$ whereas the coordinate z is orientated perpendicular to the QW interface. The expectation value $\bar{\phi}_{ij}$ of the total potential fluctuation ϕ as experienced by an electron has been defined by

$$\bar{\phi}_{ij}(r_{//}) = \int dz\, f_i(z) f_j(z) \phi(r_{//}, z) \tag{4.3}$$

For $i=j$, (4.3) represents the energy shift in the quantised level i at the in-plane coordinate $r_{//}$ due to the presence of the impurity. The problem is now that the total potential fluctuation $\phi(r_{//},z)$ is not just given by the impurity potential but also includes the contribution of the induced charge that one is trying to calculate. As a result, one one has to solve a rather complex Poisson equation in Fourier space [Yok86] to obtain the true (total) potential fluctuation,

$$\phi(q,z) = \phi^{ext}(q,z) - \sum_i \frac{m^* e^2}{2\varepsilon_0 \varepsilon_r \pi \hbar^2 q} \bar{\phi}_{ii}(q) \cdot \int |f_i(z)|^2 e^{-q|z-z'|} dz' \tag{4.4}$$

Various solution methods and approximations have been proposed to solve (4.4), ranging from full self-consistent solutions [Yok86] to severe approximations with respect to the wavefunctions involved [Lee83a]. However, a much more elegant solution to the problem exists. Fermi's Golden rule (see for example [Sin93]) states that the scattering rate for a transition $|n,k\rangle \rightarrow |n',k'\rangle$, where n,n' denote the band number and k,k' the momentum, is determined by the inner product $\left|\langle \psi_{nk} | \phi | \psi_{n'k'} \rangle\right|^2$. As a result, there is no need to know the exact form of this potential as only its projection on the initial and final eigenstates is required. Concentrating on the projections on the envelope functions (4.3), the original (Fourier transformed) Poisson's equation (4.4) transforms into a simple linear relation between the true (total) potential and the potential due to just the ionised impurities. This is known as linear response theory [Mah90]. The Fourier transformed projections of the potential can then be expressed as

$$\bar{\phi}_{nn'}(q) = \sum_{m,m'} \varepsilon^{-1}(q)\Big|_{nn',mm'} \bar{\phi}^{ext}_{mm'}(q) \tag{4.5}$$

where the indices n,n',m,m' run over the various quantised states. The screening is thus accounted for by a dielectric response matrix $\varepsilon_{nn',mm'}$. At this point, one has to appreciate the strength of the solution method. The total potential is expected to be fairly complicated due to the screening, yet the relevant physical quantities can be obtained by solving a simple matrix equation. Limiting the discussion to the TFA, i.e. a long wavelength limit $q=0$ in the screening elements, it is straightforward to show that the static dielectric response matrix takes the form

$$\varepsilon^{TFA}_{nn',mm'} = \delta_{nm}\delta_{n'm'} + \frac{m^* e^2}{2\varepsilon_0 \varepsilon_r \pi \hbar^2 q} \delta_{nn'} \delta_{mm'} \Theta\left(E_{fm}\right) \tag{4.6}$$

where Θ is the Heaviside step function and E_{fm} is the Fermi energy of the m-th subband. As noted earlier, this approach is not particularly accurate [Hai96]. The next section will therefore discuss a more advanced approach.

4.3 - The Random Phase Approximation

A more accurate dielectric response matrix than in the TFA can be obtained in the Random Phase Approximation (RPA) [Mah90]. Its derivation is rather complex, so that this thesis will limit itself to simply accepting the resulting dielectric response matrix. Fortunately, this dielectric matrix has many similarities to that obtained in the TFA and even reduces to the latter in some cases. The RPA dielectric response matrix is given by [Fle90,Gon94]

$$\varepsilon^{RPA}_{nn',mm'}(q) = \delta_{nm}\delta_{n'm'} + \frac{m^* e^2}{2\varepsilon_0 \varepsilon_r \pi \hbar^2 q} F_{nn',mm'}(q)\,\Pi_{mm'}(q) \tag{4.7}$$

where, defining $k_{\pm} = \left|k_{fi} \pm k_{fj}\right|$,

$$\begin{aligned} F_{nn',mm'}(q) &= \int_{-\infty}^{\infty} dz\, f_n(z) f_{n'}(z) \int_{-\infty}^{\infty} dz'\, f_m(z') f_{m'}(z') e^{-q|z-z'|} \\ \Pi_{mm'}(q) &= \begin{cases} 1 - \mathrm{Re}\left\{\left[1-\left(2k_{fm}/q\right)^2\right]^{1/2}\right\} & m = m' \\ 1 - \dfrac{1}{q^2}\Theta(q-k_+)\left[\left(q^2-k_+^2\right)\left(q^2-k_-^2\right)\right]^{1/2} & \begin{cases} m \neq m' \\ m,m' \text{ occupied} \end{cases} \end{cases} \end{aligned} \tag{4.8}$$

The only differences between the results for the two frameworks are that the two delta-functions in the TFA dielectric response matrix have been replaced by a new function $F_{nn',mm'}(q)$, as well as that the Heaviside stepfunction has been replaced by the electron density correlation function Π. The similarities are even more striking when one realises that $F_{nn'mm'}$ reduces to a diagonal matrix in the case of a symmetrical structure which elements would have been retained in TFA framework had one not made the long-wavelength approximation. Furthermore, the density correlation function Π reduces to 1 in the case where the momentum transfer q is smaller than both $2k_{fm}$ and k_+, something which is always satisfied for a system containing just one occupied band.

Note that the electron density correlation function $\Pi_{mm'}$ has only been presented for the case where both bands m,m' are occupied. In practice, one has to truncate the dielectric response matrix at some point to obtain a workable system. In this work, it was chosen to limit the sum over the indices n, n',m and m' to just the occupied bands. Although scattering at low temperatures will only occur between occupied bands, one has to keep in mind that the empty bands will still affect the mobility via their influence on the dielectric response matrix. Their influence is however expected to be relatively small [Hai95].

A full description of the interaction potential, given by the impurity and the screening, has been obtained. One can therefore now set up a scattering model, and this will be the subject of section

4.5. However, before such a model can be set up, some comments about the application the dielectric response matrix to Fermi's Golden Rule have to be made.

4.4 - Ehrenreich's overlap function

Section 4.2 explained the framework that was chosen to include screening in the scattering potential. A dielectric response matrix was set up whose derivation utilised the fact that the exact form of the total scattering potential ϕ is not required to calculate scattering rates, but only its projection on the initial and final wavefunctions. The latter follows, as mentioned, directly from Fermi's Golden Rule. The consequent discussion only involved the projections of the total scattering potential on the envelope functions, which were directly related to the projection of the impurity potential via (4.5).

At this point, one has completely ignored the fact that the initial and final wavefunctions are not simply given by the envelope functions f_n. The latter only describe the slowly varying part, on an atomic scale, of the wavefunction. The true wavefunction in the effective mass approximation is given by $\psi(r) = \sum_i f_i(r) u_{i0}(r)$, where u_{i0} are zone-centre solution of the original Schrödinger Hamiltonian. Here, the summation over i runs over all bands explicitly used in the effective mass expansion. The question is now if (and if yes, how) the presence of the zone-centre solutions affect the scattering rates. Or alternatively, how is the projection of the potential on the envelope functions related to the quantity of interest which is the projection on the initial and final wavefunctions ?

The situation is straightforward for the conduction band case. Using a 1-band model to describe the conduction band in a compositionally varying structure with growth direction along z, the wavefunction is given by $\psi^c_{nk}(r) = f^c_n(z) e^{i(k_x x + k_y y)} |S\rangle$, where $|S\rangle$ is a rapidly varying function on atomic scale with the same transformation properties as an atomic s-orbital. One thus obtains,

$$\left|\langle \psi^c_{nk}(r) | \phi(r) | \psi^c_{nk'}(r) \rangle\right|^2 = \left|\bar{\phi}_{nn'}(q)\right|^2 \cdot \left|\langle S|S\rangle\right|^2 = \left|\bar{\phi}_{nn'}(q)\right|^2 \quad , \quad q = k' - k \tag{4.9}$$

The requirement $q=k'-k$ arises from the plane wave form of the envelope functions perpendicular to the growth direction in combination with the requirement of momentum conservation. The overlap function $|\langle S|S\rangle|^2$ reduces to unity as the initial and final cell-periodic function are identical

(both given by $|S\rangle$), so that the initial and final cell-periodic functions do not appear in the final expression.

However, the situation is more complicated in the valence band case as the hole eigenfunctions away from the band edge are no longer pure heavy, light or split-off holes, but form a mixture of the three. As a result, one does not expect the cell-periodic solutions corresponding to two different momentum vectors to be identical. Extending the approach as followed in (4.9), one introduces

$$\begin{aligned} \left|\langle \psi_{nk}|\phi|\psi_{nk'}\rangle\right|^2 &= \left|\bar{\phi}_{nn'}\right|^2 \cdot G_{nk,n'k'} \\ G_{nk,n'k'} &= \left|\langle u_{nk}|u_{n'k'}\rangle\right|^2 \end{aligned} \tag{4.10}$$

where G is a multiplicative function that was first introduced by Ehrenreich [Ehr57]. This overlap function contains all the effects imposed by the cell-periodic part of the wavefunction. The implementation and interpretation is straightforward. As expected, G reduces to unity for scattering between identical cell-periodic states. Scattering between states with orthogonal cell-periodic states is prohibited, as reflected in $G=0$.

Analytic expressions for $G_{nk,n'k'}$ in the case of a bulk semiconductor can be found in [Wil71] where it was found that the overlap function takes the form $[1+3\cos^2(\alpha)]/4$ for a heavy hole intraband transition, where α is the angle between the initial and final momentum vector. A significant deviation from unity is seen, which highlights the need for the explicit inclusion of the overlap function. Its effect on scattering rates is re-examined in chapter 6 where numerical results of G will be presented for a variety of QW structures.

4.5 - Calculating scattering rates

At this point, the total framework has been set out to evaluate the relevant scattering rates using Fermi's Golden rule. From a mathematical point of view it is most convenient to restrict the approach to delta-doped structures, i.e. where the impurities are essentially contained within one atomic layer, as it leads to the simplest form of equation. Since it also offers a relatively easy interpretation compared to bulk doping, all experimental structures that are discussed in chapter 6 employ δ-doping. Using equation (4.5) in which the total potential fluctuation is written in terms of a dielectric response matrix and the definition of Ehrenreich's overlap function (4.10), one obtains directly from Fermi's Golden rule that the scattering rate per unit angle for a transition $|n,k\rangle \rightarrow |n',k'\rangle$ in a delta-doped structure is given by (see [Hai96] although the overlap factor is set to one in that article)

$$\Gamma_{nn'}(\theta) = \frac{m^* e^4 n_D}{8\pi\hbar^3(\varepsilon_0\varepsilon_r q)^2}\sum_{m,m'}\left|\varepsilon^{-1}_{nn',mm'}(q)\int dz\, f_n(z) f_{n'}(z) e^{-q|z-z_i|}\right|^2 \cdot G_{nn'}(k,k') \quad (4.11)$$

where n_D is the number of ionised donors in the δ-plane located at z_i. The total scattering rate for a given transition is obtained by integrating (4.11) over all angles. One can then define the quantum and transport lifetime for a transition n to n' as

$$\frac{1}{\tau^q_{nn'}} = \int_{-\pi}^{\pi}\Gamma_{nn'}(\alpha)d\alpha \quad , \quad \frac{1}{\tau^t_{nn'}} = \int_{-\pi}^{\pi}\Gamma_{nn'}(\alpha)[1-\cos(\alpha)]d\alpha \quad (4.12)$$

These definitions only differ in the sense that the transport lifetime contains a scaling factor *[1-cos(α)]* within the integral. The interpretation of the above scattering times is straightforward. The quantum lifetime reflects the true mean time between scattering events of all possible angles. This lifetime is of relevance in for example a Shubnikov-de Haas experiment where any scattering event will disturb the cyclotron motion of the carrier. In contrast, one does not expect all scattering events to carry the same weight when evaluating the electrical resistivity (or mobility) of a structure such as done when performing a Hall- or van der Pauw measurement. Clearly, an intra-band scattering event over a small angle will have negligible effect on the current flow as both the initial as the final state will have their momentum essentially parallel to each other. The latter is the reason for the presence of the scaling factor, as the momentum loss *along the direction of original motion* is given by k_F*[1-cos(α)]* (see fig.4.1c).

Ionised impurity scattering in *n*-type δ-doped two-dimensional systems has previously been considered by Masselink and Thobel [Mas91,Tho93]. There, the evaluation was restricted to just a single occupied subband. It can be shown that the equations presented in this thesis reduce to those of Masselink and Thobel when making the same restriction. In the situation where there are no inter-band scattering processes, the transport scattering time $\tau=\tau_{11}{}^T$ then directly determines the carrier transport mobility via

$$\mu = \frac{e\tau}{m^*} \quad (4.13)$$

(and similarly for the quantum mobility). This one-band scattering model with its simple conversion to carrier mobility is essentially the description that will be used to evaluate the experimental results for the hole mobility in a set of single occupied subband InGaAs-AlGaAs QWs in chapter 6 in this thesis. However, the actual model that has been implemented is the multi-band model as described in the previous sections. It was used to gain insight in the mechanisms limiting the mobility in a set of channel doped *n*-type devices that were previously studied by J.Roberts [Rob96,Rob98] at the IRC. This evaluation led to useful insight in inter-subband scattering which was consequently used in the optimisation of the *p*-type structures grown

specifically for this study. The next section deals with the question of how to combine several inter- and intra-scattering times between various occupied subbands to a mobility that is measured experimentally in a magneto-transport measurement.

4.6 - Scattering between multiple occupied bands

When the system under investigation contains more than one occupied subband, one has to solve the Boltzmann transport equation to correctly account for all the inter- and intraband transitions to obtain the transport lifetime τ_n^t of the individual levels. Siggia and Kwok showed that the Boltzmann transport equation can be written as a set of coupled linear equations [Sig70]

$$\sum_{n'=1}^{N} K_{nn'}\tau_{n'}^{t} = 1 \tag{4.14}$$

where

$$K_{nn'} = \begin{cases} \dfrac{1}{\tau_{nn}^{t}} + \sum_{m\neq n}\dfrac{1}{\tau_{nm}^{q}} & n = n' \\ \dfrac{k_{fn'}}{k_{fn}}\left(\dfrac{1}{\tau_{nn'}^{t}} - \dfrac{1}{\tau_{nn'}^{q}}\right) & n \neq n' \end{cases} \tag{4.15}$$

For a system with just two occupied subbands, one obtains the following transport lifetimes,

$$\begin{aligned} \tau_1^t &= \frac{K_{22} - K_{12}}{K_{11}K_{22} - K_{12}K_{21}} \\ \tau_2^t &= \frac{K_{11} - K_{21}}{K_{11}K_{22} - K_{12}K_{21}} \end{aligned} \tag{4.16}$$

At this point, it is useful to point out two interesting limits,

- As expected, one obtains for a system with two *identical* bands that the transport lifetime of each band is half of what one would have obtained when one had neglected interband scattering
- For the case where the second subband is just about to get populated, $k_{f2}=0$, the transport lifetime of the first subband is given by

$$\frac{1}{\tau_1^t} = \frac{1}{\tau_{11}^t} + \frac{1}{\tau_{12}^q} \tag{4.17}$$

  as scattering to the second subband involves a loss of all momentum.

After solving (4.14), the corresponding mobilities for each subband can be obtained from the subband carrier lifetimes using the generalisation of (4.13),

$$\mu_i = \frac{e\tau_i}{m^*} \tag{4.18}$$

These individual mobilities and their population can then be combined to obtain the Hall-density and Hall-mobility, i.e. the values one would measure in a Hall-measurement. For a system with n-occupied levels ($i=1..n$) with densities n_i and mobilities μ_i, one obtains [Pet58]

$$n^{hall} = \frac{\left(\sum_{i=1}^{n} n_i \mu_i\right)^2}{\sum_{i=1}^{n} n_i \mu_i^2} \qquad \mu^{hall} = \frac{\sum_{i=1}^{n} n_i \mu_i^2}{\sum_{i=1}^{n} n_i \mu_i} \tag{4.19}$$

assuming that the Hall-factor (r_H) equals unity. The latter is a good approximation for degenerate structures at low temperatures.

4.7 - Comments on the definition of the effective mass

The effective mass m^* was defined in equation (1.1) in chapter 1. This is however not the only way to define an effective mass, and as a result there are various definitions present in the literature which are referred to with a variety of names. For example, there is the 'carrier concentration mass' [Pet90], 'optical mass' [Bla87], effective density of states mass [Bla87], as well as those more appropriate to describe carrier transport which are the momentum (m^*) and acceleration mass (M^*) [Zaw82,Rot92] to which this work will limit itself. There is considerable ambiguity in the literature about which mass one should use when modelling carrier transport. This is probably due to the fact that both definitions give the same mass when applied to an isotropic parabolic band. However, as pointed out by for example [Bla87], considerable differences are found when applied to non-parabolic bands of which the hole bands form an extreme example. Despite this, various authors have tried to extend the transport equations beyond the parabolic approximation to the case of warped bands, yet without agreeing on the fundamental choice of mass to use in such an approach (compare [Tak85] to [Ye87]). An important point to make is that most of the authors that choose to use the acceleration mass completely gloss over the intrinsic problem that this definition of mass results in infinite or even negative mass points when applied to the valence bands or when applied to the conduction band when looking at relatively large momentum [Tya91]. Obviously, this casts serious doubt on whether the acceleration mass is the proper definition to use in the scattering model of section 4.5. The physical interpretation of both definitions is now briefly discussed.

The equation of motion of an electron with momentum k under the influence of an external force F is given by [Gas74,Rot92]

$$F = \hbar \frac{dk}{dt} \tag{4.20}$$

Turning to the classical picture, the motion of a localised wavepacket made up of wavefunctions near this particular wavevector (or momentum) k is now considered. Such a wavepacket travels with a velocity given by the group velocity, where ω is the frequency associated with a wavefunction of energy E [Gas74,Zaw82,Rot92],

$$\begin{aligned} \underline{v} = \frac{d\omega}{d\underline{k}} &= \frac{1}{\hbar}\nabla_k E \\ &= \left(\frac{1}{\hbar^2 k}\nabla_k E\right)\cdot\hbar\underline{k} \\ &= \frac{1}{m^*}\hbar\underline{k} \end{aligned} \tag{4.21}$$

where the $E(k)$ dispersion is assumed to be spherical. A 'momentum mass' which relates momentum to velocity was introduced

$$\frac{1}{m^*} = \frac{1}{\hbar^2 k}\frac{\partial E}{\partial k} \tag{4.22}$$

which corresponds to the definition used in (1.1). This is one possible definition of the effective mass, but one could alternatively have defined a mass by relating force to acceleration,

$$\underline{a} = \frac{d\underline{v}}{dt} = \frac{d}{dt}\left(\frac{1}{\hbar}\nabla_k E\right) = \frac{dk}{dt}\frac{d}{dk}\left(\frac{1}{\hbar}\nabla_k E\right) = \frac{\nabla_k^2 E}{\hbar^2}\cdot\underline{F} = \frac{\underline{F}}{M^*} \tag{4.23}$$

where an acceleration mass has been introduced which is now given by

$$M^{*-1} = \frac{1}{\hbar^2}\frac{\partial^2 E}{\partial k^2} \tag{4.24}$$

At this point it is important to note that both definitions are physically correct. Therefore, the question is not which of the definitions is wrong, but which one is applicable in the previously derived scattering model.

As illustrated in fig.1.7, the mass in equation (4.11) reflects the total number of states that can particupate in the scattering process and is thus related to the density of states at the Fermi energy. The same holds for the dielectric response matrices (4.6,4.7) as outlined in the TFA approach. Both occurrences of m^* clearly correspond to the momentum mass (1.1,4.22) as the density of states is inverse proportional to the slope of the $E(k)$ dispersion. The problem of choosing the correct definition of the effective mass really only applies to equation (4.13) and its generalisation to the multiband case (4.18). This section now concentrates at establishing the appropriate definition of mass in the equations governing the conversion of scattering time to carrier mobility.

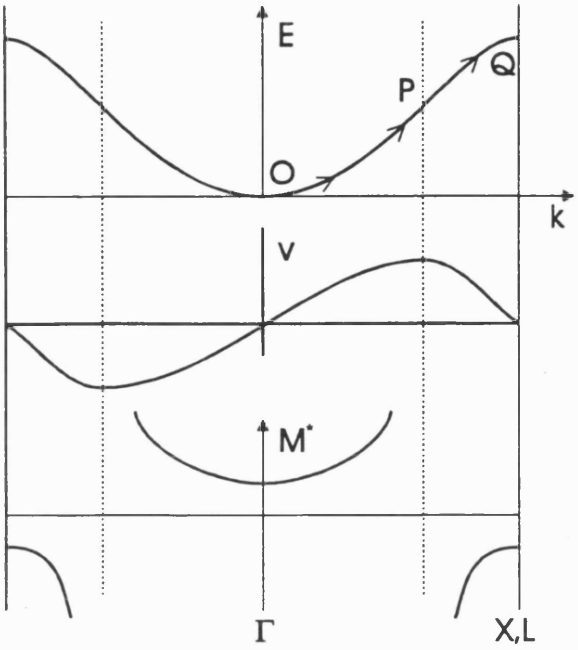


Fig.4.3 : Energy vs. momentum plot of a simple band. Also shown are the corresponding velocity as well as the acceleration mass at each point in k-space (after [Tya91]).

Fig.4.3 illustrates the motion of an electron in a simple band. Consider a single electron at momentum $k=0$ in an otherwise empty band. On application of an external force F, the electron is accelerated towards the right according to (4.20), as indicated by the arrows. Near the Γ-point, the velocity increases linearly with k and the electron behaves as a free particle with mass M^*. At higher momentum (i.e. approaching the point indicated by P) the lattice begins to interfere with its motion, slowing down the acceleration and consequently increasing the acceleration mass. Beyond point P the electron actually experiences a strong retarding force such that an increase in momentum actually decreases the velocity, as described by a negative acceleration mass. So far there are no physical problems, the description in terms of an acceleration mass M^* is correct.

However, problems arise when applying this mass to the previously introduced scattering model. Here, an ensemble of electrons confined within a Fermi-circle is modelled. As scattering can only take place at the Fermi-level, the whole group of carriers was described by the dynamics at the Fermi-energy only. The latter is exactly the cause of the problem. Consider a Fermi-contour that coincides with the inflection point P in fig.4.3. As these Fermi-electrons have an infinite acceleration mass, the application of field will not increase their velocity. By simply looking at the Fermi-contour only, one could now make the incorrect conclusion that the system would not exhibit a current flow under the application of an external electric field. The situation is even worse in the case where the Fermi-circle lies in the region between P and Q, i.e. beyond the inflection point. An application of an external field would actually reduce the speed of the Fermi-

contour electrons, yet the conclusion that the application of an external field would lead to an anti-current is clearly unphysical. Yet, as mentioned earlier, the description itself is correct.

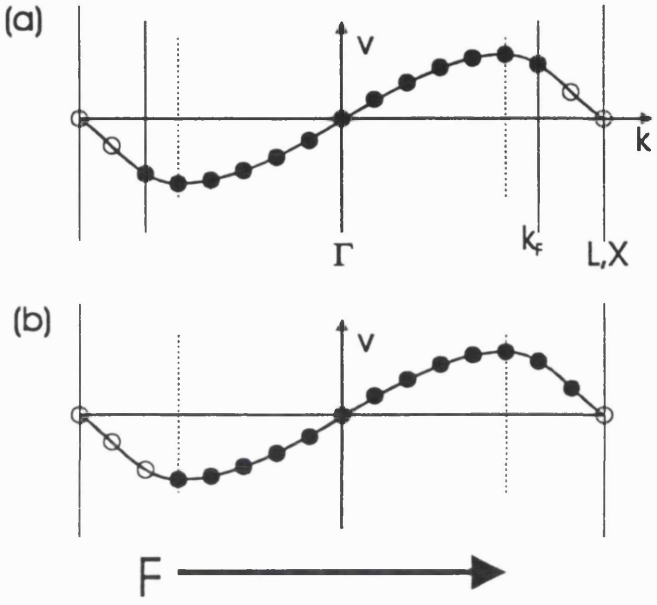


Fig.4.4 : Illustration of the collective motion of carriers under the influence of an external force F. The system displayed in (a) exhibits no net current flow, as there are as many carriers moving parallel to the +k axis as there are moving in the opposite direction. The application of an electric field will however change the electron distribution as displayed in (b), resulting in a positive flux parallel to the force.

The solution lies in the fact that one is in fact modelling an ensemble of carriers, as illustrated in fig.4.4. Point to note is that an external force not only changes the velocity of the carriers but also changes the distribution over *k*-space such as shown previously in fig.1.8. So, the velocity of the carrier at Fermi-contour with a velocity parallel to *F* might reduce, yet there will still be increase in current. The latter is illustrated in fig.4.4b which shows that the application of a field over a certain period of time effectively moves an electron from $-k_F$ (so moving in a direction *opposite* to the field) to a position near $+k_F$ (i.e. moving *along* the field direction), thereby clearly inducing a flux in the direction of the force. Re-iterating, the relevant quantity is therefore not the acceleration but the velocity at the Fermi-energy as the current will be driven by an excess amount of carriers near the section of the Fermi-contour parallel to the direction of the field. Because of this, it is obvious that the momentum mass at the Fermi-contour (relating momentum to velocity) is the quantity of interest.

So, at this point the relevant choice of mass to model the collective motion of carriers in a semiconductor has been established. As implied in the first paragraph of this section, there is considerable controversy in the literature about this subject, a subject so elementary that one would have thought that this would have been clearly settled by now. This section now ends with some of the very few comments that can be found in the literature about this subject. The momentum mass (4.22) is described by [Wet96] as the 'relevant definition of effective mass for most optical and transport related experiments'. This definition of mass is described by [Rot92] as

'useful for mobility calculations', quoting [Zaw82]. The last author seems to be the only one to make a (very) strong statement about this subject: '... it is the momentum effective mass that should be regarded as the basic quantity ... common use of M^* ... in semiconductors with parabolic band structures should be regarded as accidental.' Other papers that correctly identify m^* include [Lee83b,Tak85,Che97]. Finally, an important comment to make is that one could in principle use M^* to model an ensemble of electrons, yet that one then has to sum over the acceleration of all carriers and not just those at the Fermi-contour (see for example Lei's force-balance equation as presented in [Wen94], this approach however exceeds the scope of this thesis).

4.8 - Non-parabolicity effects in the conduction band

The previous section confirmed the validity of the effective mass definition (1.1,4.22). As shown in chapter 1, the valence bands feature considerable non-parabolic effects as well as warping, so that the mass at the Fermi-energy is expected to deviate considerably from the band edge mass m_0^*. By numerically evaluating the effective mass using a six-band model to describe the valence bands, non-parabolicity effects as well as warping are implicitly included in the analysis. However, as mentioned in sub-section 2.3.4, a simple 1-band model will be used for the conduction band so that non-parabolicity effects are not implicitly included here. Although the dispersion relation is expected to closely resemble a parabola, the electron mass (or curvature) at the Fermi-energy will deviate significantly from that at the band edge for relatively highly doped samples. This will directly affect both the population of the various levels as well as the scattering times.

Various descriptions and results have been presented in the literature to describe non-parabolicity effects in the conduction band [Ros84,Pfe96]. All methods involve solving a multiband $k \cdot p$ model and consequently parametrising the obtained $E(k)$ curves and/or mass. It was chosen to follow the (bulk) approach of Pfeffer [Pfe96] in which an effective two-level formula is introduced to describe the energy dependence of the in-plane effective mass

$$m^*(E) = m_0^* \left(1 + \frac{2E}{E_{fnd}} \right) \tag{4.25}$$

containing a single parameter E_{fnd} which is the so-called fundamental band gap. Here, the energy E is measured from the band edge, i.e. m_0^* is the band edge effective mass. For use in heterostructures, this work will neglect any quantisation effects on the effective mass [Eke87] and simply measure E from the band edge of the first quantised level. As a result of this

approximation, one obtains that the effective masses at the Fermi energy of all subbands are identical in a system with more than one degenerate band.

In the case where the interactions between the conduction and remote bands are limited to just the heavy and light hole bands, one obtains that the fundamental band gap is simply given by the energy gap [Ros84]. This is a reasonable approximation for small band gap materials because of the relative large separation in energy between the conduction band and the excluded bands as compared to the energy gap, i.e. the interaction with the heavy and light hole bands is expected to be dominant. However, for larger-gap semiconductors one has to implicitly include the interactions with the higher lying conduction bands to obtain an accurate description of the non-parabolicity. The latter is demonstrated by Pfeffer who, using a 14-band k·p model, obtained E_{fnd}=0.98 eV for GaAs which is considerably lower than its band gap. For InGaAs on InP, E_{fnd} is set to 0.7 eV which corresponds to the result obtained by [Wet96] for energies around 160 meV up the band.

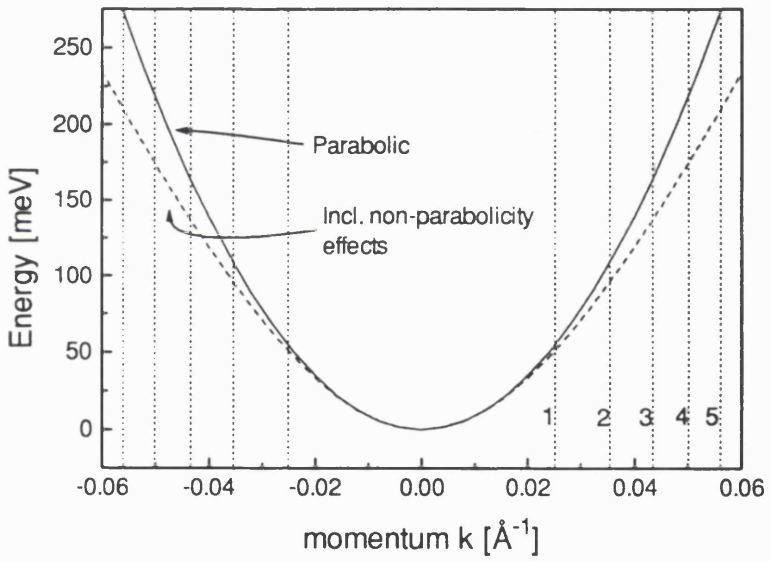


Fig.4.5 : Comparison between E(k) curves with and without non-parabolicity effects using Pfeffer's model. The fundamental band gap and the effective mass are respectively set to 0.7 eV and 0.044m_0, corresponding to unstrained $In_{0.52}Ga_{0.48}As$ on InP. The vertical dotted lines denoted with 1..5 correspond to the Fermi momentum k_F at carrier densities of 1..5·10^{12} cm^{-2}.

The ionised scattering model is now complete. A full description of scattering and screening was presented and linked to the dispersion relation via the effective mass. The next chapter will discuss the numerical interpretation of both the *k·p* as well as the scattering model.

Chapter 5 : Implementation of the theoretical models

The previous chapters outlined the models that will be employed in this work to analyse the transport properties of two dimensional electron and hole gases. Chapters 2 and 3 introduced the bandstructure model that will be used to calculate the eigenvalues of the various subbands, the corresponding wavefunctions, carrier densities and the effective masses at the Fermi-energies for a given heterostructure. This information is then subsequently used in the scattering model of chapter 4 to calculate the theoretical carrier mobility.

This chapter now deals with the implementation of those theoretical models. It is divided into two sections. Section 5.1 deals with the numerical implementation of the equations, describing and evaluating the chosen solutions methods. Section 5.2 then presents the necessary material parameters that are required in the above models. Examples include mass parameters, dielectric constants and deformation potentials. A specifically important parameter is the line-up of the conduction and valence band in a compositionally varying structure. Numerical values of the relevant material parameters and analytic expressions for the band line-up are presented for the material systems $Al_xGa_{1-x}As$, $Al_xIn_{1-x}As$, $In_xGa_{1-x}As$ and InP. The combination of the implementation contained in section 5.1 and the material parameters finally results in a workable model that will be utilised in chapter 6 to calculate dispersion relations and evaluate scattering times in two-dimensional electron and hole gasses.

5.1 - Numerical implementation

As mentioned in the introductory chapter, the fact that one is solving for a structure that contains charge leads to the additional complexity that both the Hamiltonian and Poisson's equation have to be solved simultaneously. Such a solution can be obtained by using a self-consistent procedure where, starting with an initial guess for the confining potential, subsequent changes in this potential are made until it has converged to the form in which it simultaneously satisfies both equations. A routine that performs the latter is often referred to as a Self-Consistent Poisson-Schrödinger (SCPS) solver.

The self-consistent approach is the starting point of this section, and is outlined in the general overview (§5.1.1). The internal representation of compositionally varying structures is also discussed here because of its large impact. It was chosen to describe continuous functions such as potentials, envelope functions and doping profiles at a finite number of discrete mesh points. Such a discretisation renders the differential equations such as found in Poisson's equation and the Hamiltonian into finite difference form. As a result, both take the form of a matrix equation.

Subsequently, the exact implementation of Poisson's equation (§5.1.2) and the effective mass equations (§5.1.3) are presented. Explicit forms of the matrix equations are given. As speed proves to be of crucial importance, it was chosen to solve the Hamiltonian in a time-saving expansion method whose advantages and implementation will be discussed in detail. Sub-section 5.1.4 then briefly discusses the implementation of Fermi-Dirac statistics, required to convert Fermi-energy and dispersion relations to a carrier density. Section 5.1 ends with a short discussion on the implementation of the scattering model, which is far less complicated than the Poison-Schrödinger solver.

5.1.1 - General overview

The effective mass equations that can supply information like dispersion relations, eigenvalues and wavefunctions in a heterostructure were respectively given in (2.46) for the electron case and (2.51,2.57) for the valence band case. Both Hamiltonians contain the confinement potential $V(r)$ (see subsection 2.3.3) that describes the variation of the potential energy as a function of position of the band under investigation (being either the conduction or valence band). As this is a function of material composition, it enables the fabrication of QWs and heterojunctions such as outlined in chapter 1.

The first observation is that chargeless structures are relatively easy to solve. Without charge, the confinement potential basically reduces to a simple material dependent function that describes the variation of the potential energy between the layers of different composition due to the differences in crystal structure (see fig.5.1a, or fig.3.3). Expressions for the band line-up will be presented in §5.2.2. As a result, the confinement potential is locally flat (ergo the name 'flat band' structures) and the eigenlevels and dispersion are obtained simply by solving the effective mass equations for the confinement as described by those analytic expressions for the band line-up.

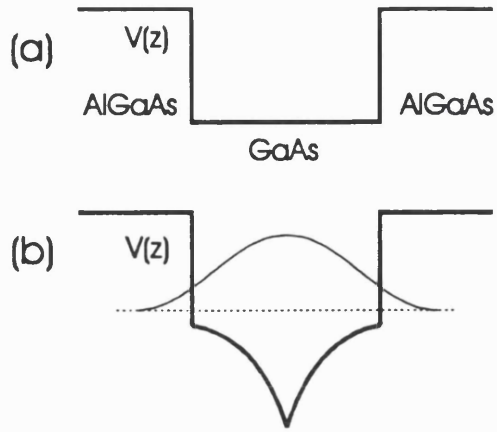


Fig.5.1: (a) Schematic picture of the confinement potential V(z) in a chargeless QW structure consisting of a GaAs well sandwiched between AlGaAs barriers. (b) Now for the case where the centre of the QW contains a δ-doped layer (i.e. the impurities are contained in a single monolayer). The potential V now contains an electrostatic part due to the presence of the impurities as well as the presence of (Coulomb attracted) carriers.

The situation is however more complicated in structures that contain charge, as pictured in fig.5.1b. The confinement potential is now no longer given by just the band line-up but contains an electrostatic component due to the presence of the impurities and carriers, among other contributions that were absent in the case where there was no charge. The total confinement potential is given by a number of terms [Ste84,Har89] (neglecting the image charge contribution as it is expected to be small [Ste84])

$$V(z) = -e\phi(z) + V_h(z) + V_{xc}(z) \tag{5.1}$$

where

- $\phi(z)$ is the electrostatic potential due to the presence of charge, given by Poisson's equation.
- $V_h(z)$ describes the variation of the potential energy as a function of material composition. It describes the effects of the changes in the surrounding lattice and electron configuration when moving from one layer to another with a different material composition.
- $V_{xc}(z)$ is the local exchange and correlation potential which is a many-body correction to the one-electron Schrödinger equation for high-density systems (see sub-section 2.2.1).

In a chargeless structure, $V_h(z)$ was the only contribution to the quantisation potential. In contrast, the confinement is now given by three different contributions. Allowing the structure to have charge doesn't just complicate the problem by the fact that one now also has to solve for the electrostatic and exchange/correlation components, but it has a rather larger impact. As mentioned above, the Schrödinger equation, or its simplification being the effective mass equations, requires knowledge of the confining potential before one can start solving for eigenvalues, electron/hole distributions and densities. As these quantities in turn affect the confining potential through $\phi(z)$ and $V_{xc}(z)$, it follows that the particular equations cannot simply be solved individually but that *one has to find a solution that simultaneously satisfies the Poisson as well as the Schrödinger equation!*

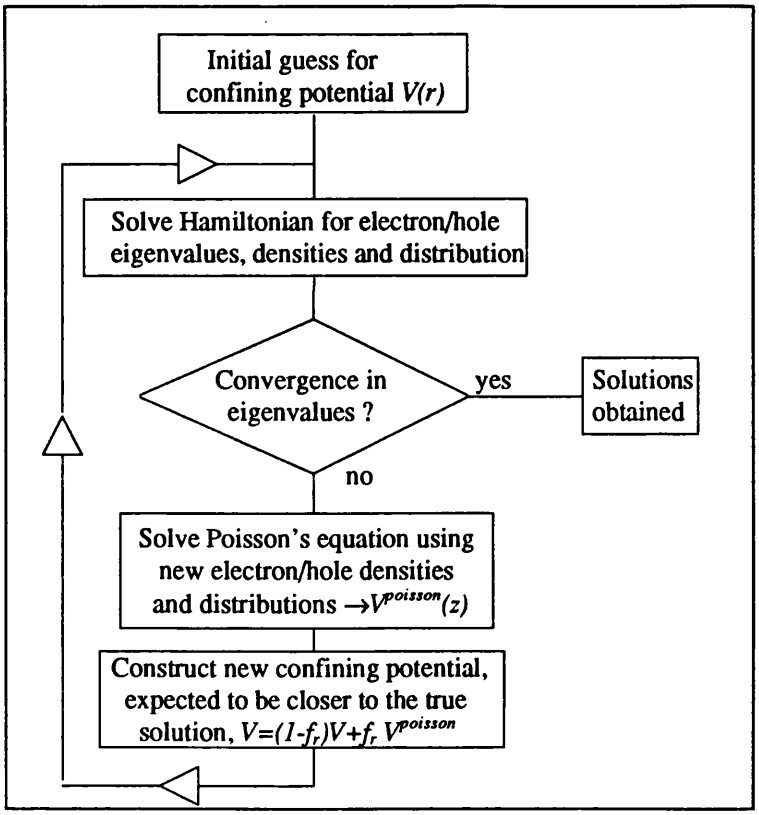


Fig.5.2: Flow chart of the self-consistent procedure that is used to obtain a simultaneously solution to both the Poisson and Schrödinger equation. Here, f_r is a so-called relaxation factor ($0<f_r<1$) which limits the feedback in the iterative process, thereby improving the stability of the routine.

Such a simultaneous solution can be obtained by using a self-consistent approach in which successively better solutions of both equations are obtained in an iterative approach. Fig.5.2 shows the corresponding flowchart of such a process. Starting with an initial guess for the confinement potential $V(z)$ including its electrostatic and exchange-correlation component (5.1), the Hamiltonian is solved for the electron/hole eigenfunctions, distributions and densities. The next step involves solving Poisson's equation to obtain the true electrostatic contribution $\phi(z)$ corresponding to these densities and distributions, which defines a new confinement potential $V^{Poisson}(z)$. If both confinement potentials were equal, one would have obtained a simultaneous solutions of both equations, but generally this is not the case. One then proceeds by constructing a confinement potential which is expected to be closer to the exact solution,

$$V(z) \rightarrow (1-f_r)V(z) + f_r V^{poisson}(z) \tag{5.2}$$

where f_r is a so-called relaxation factor ($0<f_r<1$, typically a few percent). The above procedure is then repeated using this new guess for the confinement potential. After a number of iterations, it is expected that the solution will have converged to a point where it satisfies both the Poisson and the Schrödinger equation (within a very small margin of error, depending on the number of iterations, the feedback as determined by the relaxation factor, and speed of the convergence).

Although the flowchart given in fig.5.2 correctly depicts the principles behind the self-consistent procedure, the actual process is slightly more complicated. As discussed above, the feedback is limited quite severely by means of a relaxation factor when constructing the new confinement potential of the next iteration. This is a typical feature of any self-consistent routine as such a process is intrinsically unstable in practice, especially in the early iterations as the potential at that stage normally deviates quite severely from the true solution. Rather than use a fixed value for this relaxation factor, it is allowed to vary from iteration to iteration to speed up convergence. As a consequence, the progress of the iterative procedure has to be closely monitored to protect the process from becoming unstable, something that will not be expanded on here.

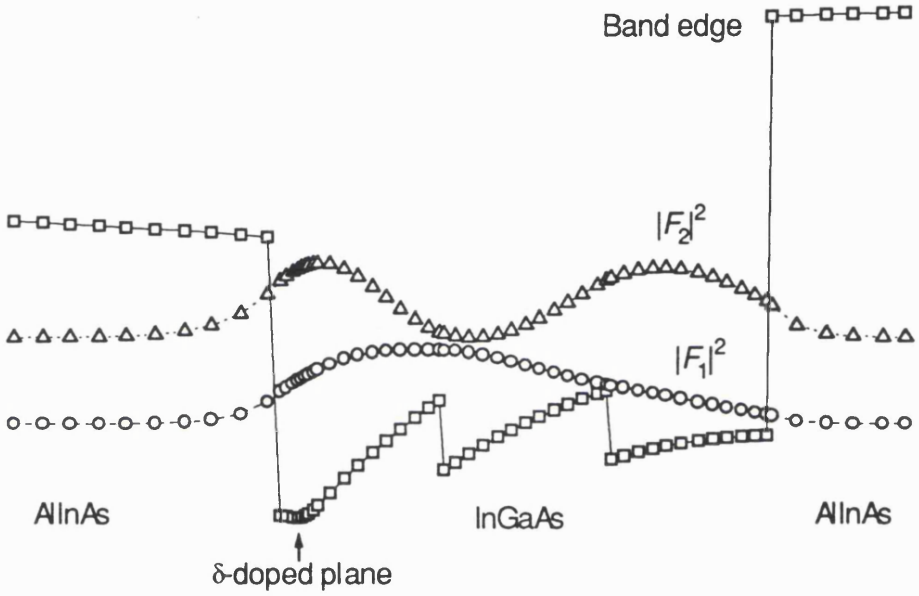


Fig.5.3 : The confining potential and envelope functions in a doped and graded InGaAs QW sandwiched by AlInAs barriers. The relevant quantities are described by their values at a number of discrete mesh points. δ-Doped layers are described by a layer of small but finite width, typically 5Å wide. (For illustrative purposes, the number of mesh points has been reduced substantially)

An important aspect of the total implementation is the framework chosen to represent a compositionally varying structure, especially with respect to quantities like the confinement potential, envelope functions and doping distribution. As will be shown, the particular representation chosen has a large effect on how Poisson's and the effective mass equations are solved. This issue completes this general section on the implementation of the Poisson-Schrödinger solver.

The most flexible way to represent continuous functions is by introducing a mesh and discretising the particular quantities at the various mesh points (fig.5.3). By automatically adding mesh-points

in regions with high band curvature (i.e. high density), a typical structure is typically described by a non-uniform mesh of the order of one thousand points. In such a discrete model, any derivative is replaced by its finite difference form,

$$\left.\frac{\partial y}{\partial x}\right|_{x(i)} \rightarrow \frac{y[i+1]-y[i]}{h_i}$$
$$\left.\frac{\partial^2 y}{\partial x^2}\right|_{x(i)} \rightarrow \frac{h_2 y[i-1]-(h_1+h_2)y[i]+h_1 y[i+1]}{h_{i-1}h_i(h_{i-1}+h_i)} \qquad (5.3)$$

where $y[i]$ is the discrete value of the quantity y at the i-th mesh point and h_i is the difference between adjacent mesh points; $h_i=x[i+1]-x[i]$. The adoption of a discrete mesh transforms any differential equation (such as Poisson's and the effective mass equations) into a matrix equation. The next sections will deal with their specific implementation.

At this point, it has to be mentioned that the basic framework of the program, consisting of the Poisson solver combined with the implementation of a single band effective mass equation, was designed by M.C.Hellon at Philips Redhill. It was then subsequently subjected to improvements and changes by the author's supervisor. However, severe changes in the convergence routine were made by the author as comparison of its results with those obtained by Stern and Das Sarma [Ste84] revealed differences of the order of several meV. Exchange and correlation was included, as well as strain, and the model was extended to include new material systems. Stringent monitoring was introduced in the Poisson equation to avoid the (occasional) occurrence of iterative steps that would result in a divergent iterative series, as well as to speed up the iterative process. Finally, the program was extended to render it capable of solving the valence band Hamiltonian. The latter resulted in a considerable increase in program size, approximately doubling the size of the original program.

5.1.2 - Solving Poisson's equation

As mentioned in the previous section, the presence of electronic charge in a semiconductor heterostructure affects the confinement potential through the associated electrostatic potential which is given by Poisson's equation (assuming that the dielectric constant is identical throughout the structure)

$$\varepsilon_0\varepsilon_r\frac{d^2\phi(z)}{dz^2} = \rho_e(z)-\rho_h(z)-\rho_I(z) = \rho_{tot}(z) \qquad (5.4)$$

where $\rho_e(z)$, $\rho_h(z)$, $\rho_I(z)$ respectively are the electron, hole and ionised impurity charge density. The assumption is made that the regions sufficiently far away from the heterostructure can be treated as if they were bulk, so that the electron density distribution is split up into a quantum and bulk region :

$$\rho_e(z) = \begin{cases} e\sum_i n_i |F_i(z)|^2 & z_0 \le z \le z_1 \\ \rho_e^{3D}(z) & z < z_0 , z > z_1 \end{cases} \tag{5.5}$$

where the summation over i runs over all populated electron subbands with density n_i and envelope function $F_i(z)$. The coordinates z_0 and z_1 define the 'quantum region' and are set manually. A similar split-up such as (5.5) is made for the hole density. All densities, either electron, hole or ionised doping densities, follow from the position of the quantised states or doping levels with respect to the Fermi level as given by Fermi-Dirac statistics [Har89].

Using the finite difference form of a second order differential as presented in (5.3), it is easy to see that Poisson's equation (5.4) takes the following matrix form,

$$\begin{pmatrix} \cdot & & \\ & \cdot & \\ \cdot\ h_{i-1} & (-h_{i-1}-h_i)\ h_i & \cdot \\ & \cdot & \\ & \cdot & \end{pmatrix} \begin{pmatrix} \phi_{i-1} \\ \phi_i \\ \phi_{i+1} \end{pmatrix} = \begin{pmatrix} \cdot & & \\ & \cdot & \\ \cdot\ \cdot & h_i h_{i+1}(h_i + h_{i+1})/2 & \cdot\ \cdot \\ & \cdot & \\ & \cdot & \end{pmatrix} \begin{pmatrix} \\ \rho_i \\ \\ \end{pmatrix} \tag{5.6}$$

or $\boldsymbol{A}\ \phi = \boldsymbol{B}\ \rho$, in which $\boldsymbol{A}$ is a tridiagonial matrix and $\boldsymbol{B}$ is diagonal, ϕ_i and ρ_i respectively denote the electrostatic potential and charge density at the i-th mesh point. However, rather than solve (5.6) directly, a slightly different approach is employed (for stability reasons) which uses two Newtonian-like interpolation schemes, one for the initial and a faster but more critical one for the final iterations. The electrostatic potential ϕ is thus solved in an iterative approach using standard Numerical Algorithm Group (NAG) routines.

It has to be pointed out that Poisson's equation (5.4) by itself does not uniquely determine the potential, but only when supplied with two boundary conditions. A simple boundary condition follows from the assumption that the substrate remains charge neutral. This effectively pins the Fermi level at mid gap at a semiconductor-substrate interface when using a semi-insulating (SI) substrate. The other boundary condition follows from an experimental observation of Fermi-level pinning at the air-semiconductor surface. This boundary conditions turn out to be dependent on the type of material under consideration, and will therefore be discussed and evaluated later on in section 5.2.4 when explicitly dealing with the various parameters and material constants.

5.1.3 - Solving the effective mass Hamiltonians

The effective mass equations for the conduction band (2.46) and valence band (2.51,2.57) were presented in chapter 2. It was shown that the conduction band effective mass model takes the form of the single differential equation, in contrast to that of the valence band which consists of a set of coupled differential equations. The latter significantly complicates the problem of obtaining the eigenvalues and wavefunctions as compared to the relatively simple conduction band problem. Despite this additional complication, the solution method chosen in this work to solve the hole-equations are based on the conduction band approach as originally implemented by M.C.Hellon. This section now starts with a brief explanation of the implementation of the conduction band model. Various alternative approaches are described, along with their drawbacks. The attention is then shifted towards the valence band case, where its numerical complexity is highlighted.

The conduction band effective mass equation is straightforward to solve for a flat band structure. In this case the solutions in each individual layer can be written down immediately (either consisting of exponentials or cosine/sines functions). The problem then merely consist of matching these solutions at each interface. For a flat band quantum well, the problem reduces to finding the solutions of a simple trigonometric equation [Gas74]. The situation is more complicated in the case of non-flat band structure, for which several solution methods have been proposed in the literature. By splitting up the structure into small sections and treating each section as being flat, the solution in each section can again be written down directly. The solutions in adjacent sections can then be linked via a 2x2 transfer matrix [Wei91]. The combined transfer matrices describe the evolution of the envelope functions over the heterostructure and can be used to obtain the various quantised states, replacing the trigonometric equation corresponding to a simple QW. However, as generally several sections are required to realistically represent a continuously varying potential, the eigenvalues and eigenfunctions can no longer be obtained analytically. An alternative approach is to use a variational method to obtain an approximate form of the envelope functions such as suggested by Fang and Howard [Fan66]. This method however has the drawback that the choice of trial functions depends on the particular problem, so that a different set is needed for each type of heterostructure. As mentioned earlier, it was chosen to use a discrete mesh to describe the (compositionally varying) structures and its corresponding envelope functions. Obviously, such a discrete model imposes no restrictions on the form of the confinement potential or envelope functions. In finite difference form, the kinetic energy operator in the conduction band effective mass equation takes the form (see fig.5.4)

$$\frac{\partial}{\partial z}\frac{1}{m^*}\frac{\partial F(z)}{\partial z}\bigg|_{z(i)} \rightarrow \frac{h_{i+1}m_+^{*-1}F_{i-1} - \left[h_i m_-^{*-1} + h_{i-1}m_+^{*-1}\right]F_i + h_{i-1}m_+^{*-1}F_{i+1}}{h_{i-1}h_i\left(h_{i-1}+h_i\right)} \tag{5.7}$$

where the subscripts +/- denote the mass directly left or right of the i-th mesh point. Away from any interface, the above equation reduces to the form given in (5.3). In a similar approach to that described in section 5.1.2, the conduction mass equation is then written in matrix form. It is easy to see from (5.7) that the Hamiltonian takes the form of a simple tridiagonal matrix which is (again) solved using NAG-routines.

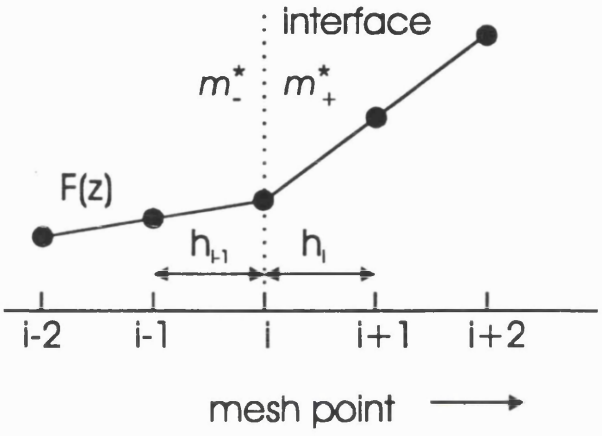


Fig. 5.4 : Definitions used in the finite difference formulae. Specific care has to be taken when writing derivatives in finite difference form around an interface, assumed to be at the i-th mesh point.

So far, there are few numerical problems. Both Poisson's as well as the conduction band effective mass equation transform into relatively simple tridiagonal matrix form which do not require extensive computer power to be solved. As will be shown, this is in sharp contrast to the valence band effective mass equations which will now be discussed.

Again, various solution methods have been proposed in the literature to solve the valence band effective mass Hamiltonian to obtain the hole eigenvalues, envelope functions and dispersion. As for the conduction band case, the problem is relatively easy to solve when applied to flat band structures as the general solution in each layer can again be written down directly. Matching this general solution, now consisting of heavy- and light- (and for 6-band models the split-off) hole contributions, the problem for a finite quantum well reduces to finding the roots of an 8x8 (4 band model, [And87]) or 12x12 determinant (6-band model, [Edw94]). Variational methods, with their drawbacks, are probably the easiest solution method for non-flat band structures and have been used in the literature to solve the heterojunction problem [Bro85,Gol95]. Brute force numerical approaches, which can handle arbitrary confining potentials, represent the other side of the scale in terms of required computational power, in which the set of N coupled differential equations of order 2 is either solved directly, or written as a set of $2N$ coupled differential equations of order 1.

As for the conduction band problem, a discrete mesh is employed to describe the compositionally varying structure and its confinement potential. As this thesis is concerned with modelling device structures, i.e. structures with charge, one has to remember that the Hamiltonian is not the only equation that has to be solved but that Poisson's equation needs to be satisfied simultaneously. As a result, speed is of crucial importance.

The solution method chosen in this work has successfully been used by Ando [And85] and is a combination of the brute force and expansion method. In this approach, the valence band Hamiltonian is solved in the so-called diagonal approximation in which all off-diagonal elements are set to zero[1]. As a result, the set of differential equations decouples and reduces to three single differential equations (one each for the heavy, light and split-off bands) which closely resemble the electron problem. As this is a Sturm-Liouville problem [Kre88], it follows that all three sets form a complete set so that *any* envelope function can be expanded in such a set,

$$F^{4band} = \begin{pmatrix} \sum_n a_n F_n^{hh} \\ \sum_m b_m F_m^{lh} \\ \sum_m b'_m F_m^{lh} \\ \sum_n a'_n F_n^{hh} \end{pmatrix} \quad or \quad F^{6band} = \begin{pmatrix} \sum_n a_n F_n^{hh} \\ \sum_m b_m F_m^{lh} \\ \sum_l c_l F_l^{so} \end{pmatrix} \quad \begin{matrix} 1 \le n \le N_{hh} \\ , \quad 1 \le m \le N_{lh} \\ 1 \le l \le N_{so} \end{matrix} \tag{5.8}$$

where F_n^{hh}, F_m^{lh} and F_l^{so} respectively are the heavy, light and split-off solutions for the envelope functions in the diagonal approximation. Using such an expansion, one then has to solve for the expansion coefficients rather than directly for the envelope functions (see for example the appendix in [Sin93]). Inserting the 6 band expansion (5.8) into the Hamiltonian $H_{\pm}^{3x3}$ (2.57), one obtains

$$\begin{matrix} H_{hh}\sum_n a_n F_n^{hh} & (R \mp iS)\sum_m b_m F_m^{lh} & \left(\sqrt{2}R \pm \frac{i}{\sqrt{2}}S\right)\sum_l c_l F_l^{so} & = 0 \\ (R \mp iS)^{\dagger}\sum_n a_n F_n^{hh} & (H_{lh} \mp iC)\sum_m b_m F_m^{lh} & \left(\sqrt{2}Q \mp i\sqrt{\frac{3}{2}}S\right)\sum_l c_l F_l^{so} & = 0 \\ \left(\sqrt{2}R \pm \frac{i}{\sqrt{2}}S\right)^{\dagger}\sum_n a_n F_n^{hh} & \left(\sqrt{2}Q \mp i\sqrt{\frac{3}{2}}S\right)^{\dagger}\sum_m b_m F_m^{lh} & (H_{so} \pm iC)\sum_l c_l F_l^{so} & = 0 \end{matrix} \tag{5.9}$$

Re-iterating, F_n^{hh} are the solutions of $H_{hh}F_n^{hh} = E\,F_n^{hh}$, and similarly for F_m^{lh} and F_l^{so}. By using the standard technique of taking the inner product with $F_{n'}^{hh}$ in the first, $F_{m'}^{lh}$ in the second and $F_{l'}^{so}$ in the third differential equation, (5.9) can be transformed in a matrix equation of order $N_{hh}+N_{lh}+N_{so}$. Rather than give the full matrix , this section will limit itself to just the matrix

[1] Note that the results corresponds to the band edge solutions, i.e. zero in-plane momentum (k_x=k_y=0), in the case of [001] and [111] growth.

elements caused by the first row/second column element of H_+^{3x3} (boxed in the above equation, or see fig.5.5). All other Hamiltonian matrix elements can, with a bit of effort, be obtained in a similar manner. The particular matrix elements $H_{n'm}$ are given by

$$\begin{aligned}
&\left\langle F_n^{hh}|R-iS|F_m^{lh}\right\rangle \\
&=\left\langle F_n^{hh}(z)\middle|\sqrt{3}\left\{\gamma_2(z)\left(k_x^2+k_y^2\right)-2i\gamma_3(z)k_xk_y-2\left(k_x-ik_y\right)k_z\right\}\middle|F_m^{lh}(z)\right\rangle \qquad (5.10)\\
&=\sum_{layer=1}^{N_{layer}}\sqrt{3}\left[\gamma_2|_{layer}\left(k_x^2+k_y^2\right)-2i\gamma_3|_{layer}k_xk_y\right]\times\left\langle F_n^{hh}(z)\middle|F_m^{lh}(z)\right\rangle_{layer}\\
&\qquad -2\sqrt{3}\gamma_2|_{layer}\left(k_x-ik_y\right)\times\frac{1}{i}\left\langle F_n^{hh}(z)\middle|\frac{\partial}{\partial z}\middle|F_m^{lh}(z)\right\rangle_{layer}
\end{aligned}$$

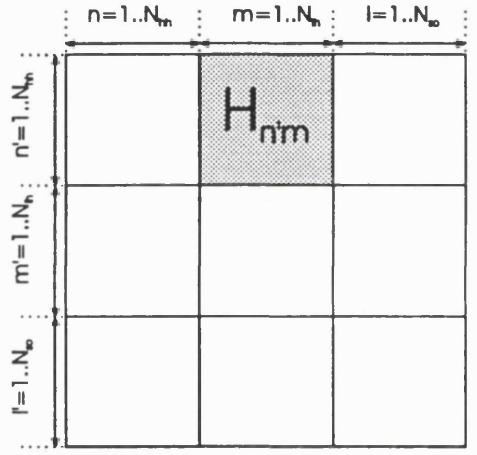


Fig.5.5: Schematic of the Hamiltonian $H_\pm^{3x3}$ in matrix form. Highlighted are the elements described in the text.

The actual implementation in FORTRAN of the whole routine is rather straightforward. Having obtained the envelope functions in the diagonal approximation, one numerically evaluates the various inner products of which two examples are as shown in the above equation. One finds that 11 of such inner products need to be considered to be able to model [001], [110] and [111] growth, corresponding to various bracketed operators and ordering. Having obtained these, the band energies $E(k_\perp)$ can be obtained at any in-plane momentum $k_\perp$ by solving the matrix equation depicted in fig.5.5 of order $N_{hh}+N_{lh}+N_{so}$.

At this point, one might ask if the above method really offers any speed advantage over the much simpler brute force methods. The answer is that it does, and quite considerably. Simply writing the 6-band Hamiltonian in its block-diagonal form (2.57) as a matrix equation, similar to the approach taken when discussing Poisson's equation, would have rendered the problem into finding the eigenvalues of two $3N$x$3N$ matrices, where N is the number of mesh points (typically 500 or more)

The order of the matrix follows directly from the fact that one is solving for the discrete values of the heavy, light and split-off envelope functions at all N mesh points. As the double differentials of again all three functions (heavy, light and split off) are involved, it follows from the finite difference form of the derivatives (5.3) that as much as 9 non-zero elements can be found in each row. In contrast, the problem reduces to finding the eigenvalues of three tridiagonal NxN matrices for the diagonal approximation to obtain the expansion states. After having obtained these expansion states and the various inner products, $E(k_\perp)$ can be evaluated at any in-plane momentum $k_\perp$ by solving a matrix equation of the order of just 30..45 (using 10 or 15 expansion states) whereas the brute force method would have to go through the same computationally demanding procedure for each value of the in-plane momentum. Generally a significant number of points of the dispersion $E(k_\perp)$ is required, for example to perform Fermi-Dirac statistics (§5.1.5), so that the effectiveness of the approach to speed up the calculation is obvious.

5.1.4 - Guide and evaluation of the Hamiltonian expansion method

The expansion method used to solve the valence band effective mass equation is evaluated in this section. It proves that one can optimise the convergence by introducing infinite potential barriers, thereby actively controlling the form of the higher lying expansion states. The fastest convergence is obviously achieved when the expansion states resemble the real physical solution as much as possible. The introduction of these potential barriers is similar to the split-up of the structure into a 'bulk' and 'quantum' region such as introduced in §5.1.2 when discussing the implementation of Poisson's equation. It will now be shown that the expansion method is a very powerful method, generally requiring only a limited number of expansion states, when following some simple guidelines about the placement of the artificial infinite barriers.

In theory, the expansion method as specified in the previous section is complete in the sense that it could directly be applied to any problem. There is however the practical problem that a typical quantum well does not support more than just a couple of bound states, which are typically insufficient to reach convergence. One could of course simply proceed by adding some of the unconfined states above the band edge in the expansion. But, as they are expected to have little in common with the bound solutions, it is unlikely that they will behave as ideal expansion states. As suggested by Ando [And85], the solution is to extend the number of bound states by introducing infinite barriers at a distance L from the heterostructure (fig.5.6), respectably at coordinates z_0 and z_1.

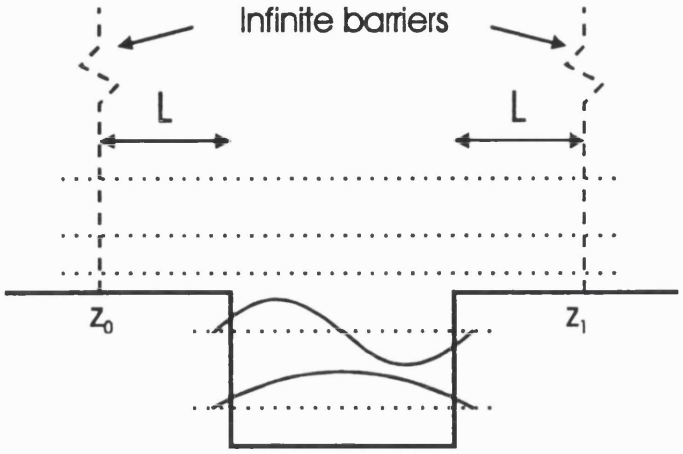


Fig.5.6 : Infinite barriers are introduced at a distance L from the interface(s), respectively at the coordinates z_0 and z_1, to optimise the expansion method. The presence of such barriers transforms the higher lying free states into bound states which can then be (efficiently) used in the expansion procedure.

Obviously, these barriers are only introduced to speed up convergence and should not affect the final $E(k)$ energies. One then finds that there is a trade-off between speed of converge and influence of these barriers on the final results :

- A relatively high separation length guarantees that the initial problem will not be affected by the introduction of the infinite barriers, but convergence will be slow. As mentioned earlier, the higher lying expansion states are simply too different from the confined solution that one is trying to calculate.
- A short separation length will guarantee fast convergence but will at the same time affect the resulting eigenvalues.

As a result, these barriers should be placed rather carefully. Fig.5.7 shows the effect of adding such barriers on the eigenvalues as a function of separation length L. Several interesting observations can be made. The quantisation energy (i.e. at $k_\perp=0$) of the light hole states is clearly much more strongly affected than that of the heavy hole states at the same separation length. This was to be expected as the light hole, due to its lower mass, penetrates the barriers further than the heavy holes and thus interacts more strongly with the infinite barriers. Increasing the momentum to $k_x=0.03\text{Å}$ adds to the interaction for both heavy hole states, in contrast to the light hole level which displays a reduction in interaction. Both effects are attributed to band mixing. The eigenstates are clearly no longer pure heavy or light hole states at this distance from the band edge. As a result, the heavy hole states acquire some light hole characteristics with a corresponding increase in interaction, and similarly for the light hole bands which acquire some heavy hole characteristics. Summarising, the influence of the infinite barriers can be neglected when put at least 40 Å from an interface.

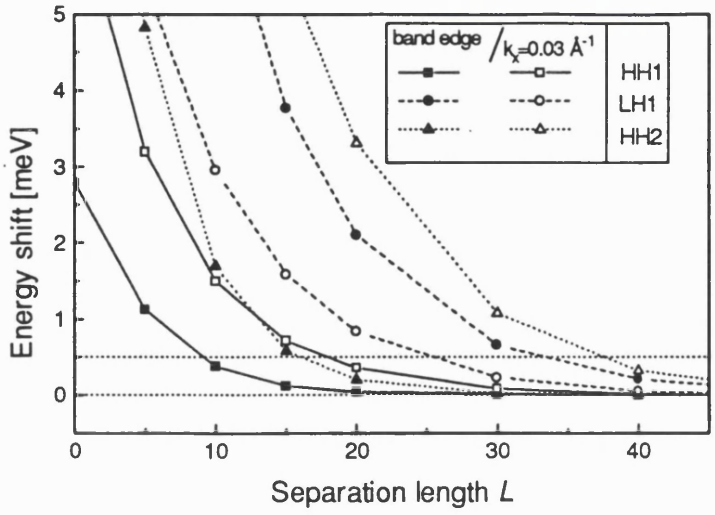


Fig.5.7 : Shift in eigenvalues in the first three quantised states in a 100Å $Al_{0.3}Ga_{0.7}As$-GaAs QW as induced by adding infinite barriers at a distance L from the interfaces (4 band model).

Fig.5.8 shows the influence of the separation length L on the convergence speed of the calculated dispersion corresponding to a 100Å $Al_{0.3}Ga_{0.7}As$-GaAs QW. Comparing (a) in which $L=100Å$ to (b) where $L=50Å$ reveals that a smaller separation length speeds up the convergence considerably, especially in the higher levels. A quantitative idea of the number of expansion states needed to accurately model p-type quantum well structures can be obtained from fig.5.9 where the deviation in Fermi energy, as compared to the fully converged value, is plotted versus the number of expansion states. The structure under investigation is the same QW as described above. Results are plotted for two hole densities, 5.10^{11} and 10^{12} cm^{-2}, and two temperatures As expected, an increase in density or temperature requires an implicit inclusion of more expansion states to achieve convergence, caused by the resulting occupation of regions with relatively high momentum k. However, the general result is that typically 10 expansion states are sufficient to achieve full convergence in QW structures.

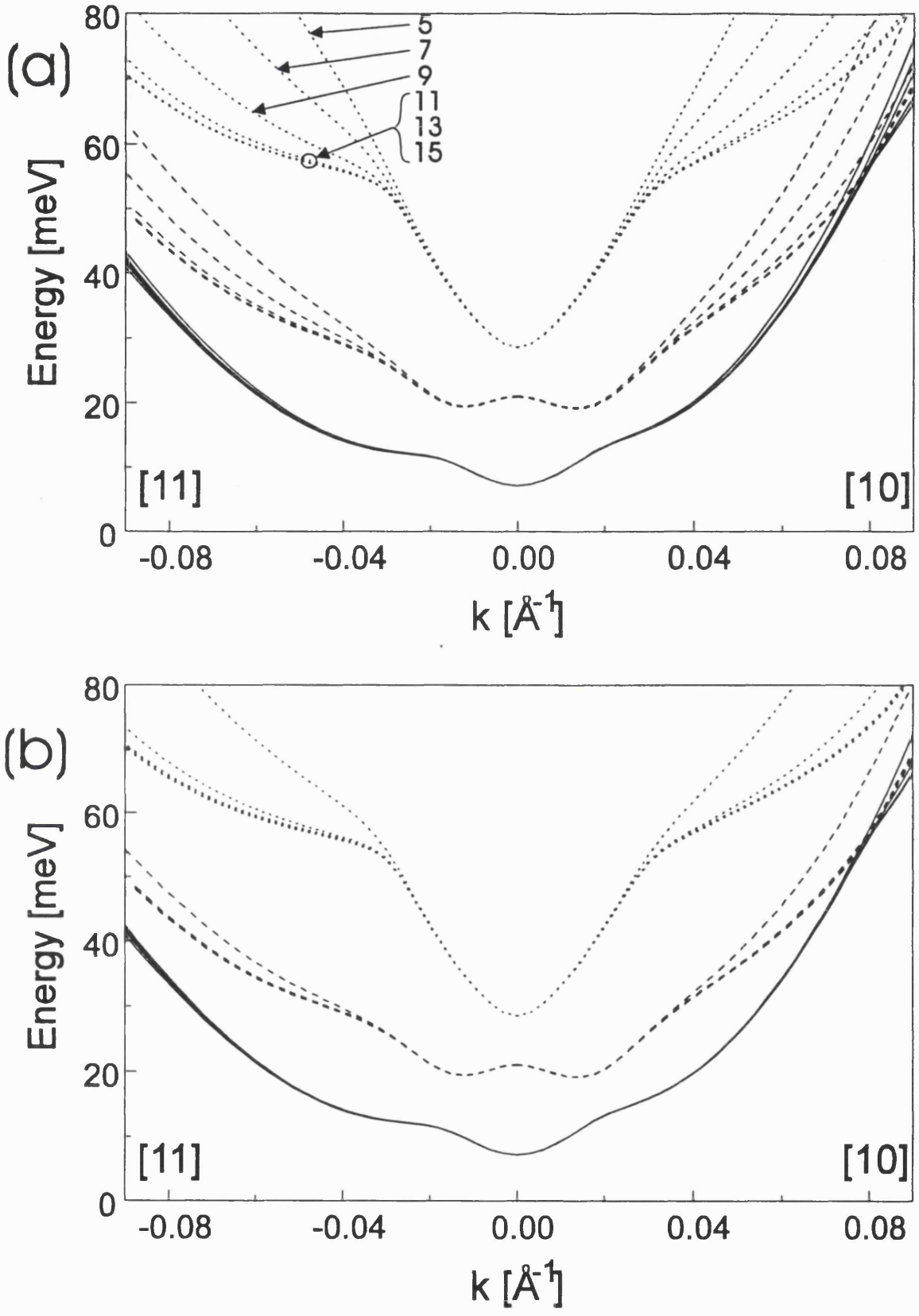


Fig 5.8: E(k) diagrams for the three lowest quantised states in a 100Å $Al_{0.3}Ga_{0.7}As$-GaAs QW calculated using a 6-band model, respectively obtained with a separation length of 100 Å (top) and 50Å (bottom). The number of expansion states was increased from 5 per band (heavy, light and split-off) to 15 in steps of two.

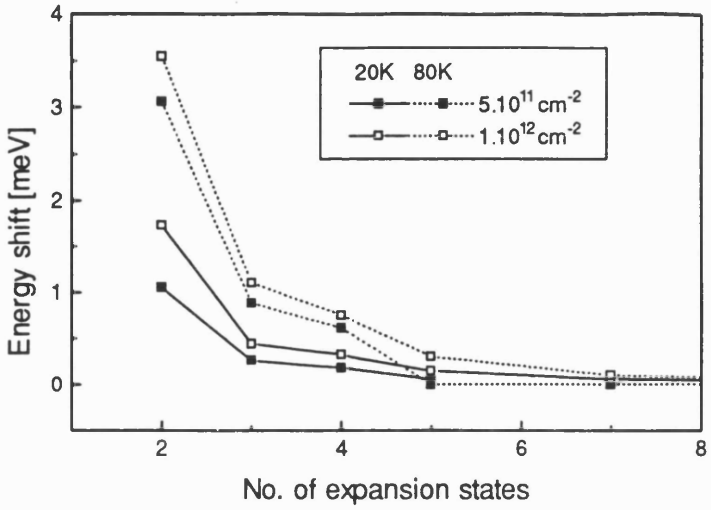
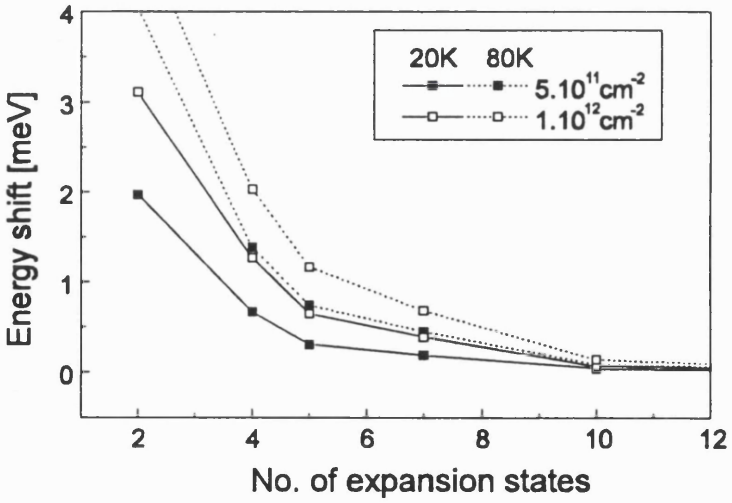


Fig.5.9 : Shift in Fermi energy when lowering the number of expansion states (hh, lh and so) from the from their value required to achieve full convergence. The results are plotted for two carrier densities as well as two temperatures. The structure used is the quantum well as described in fig.5.8, with the infinite barriers placed 50Å away from the QW.

For heterojunctions, the optimisation of the convergence by introducing infinite barriers is slightly more complicated because of two reasons. First, one cannot so easily identify where to put the infinite barrier in the lower band gap material. Secondly, the envelope functions corresponding to finite k in such a triangular shaped potential deviate more from the band edge solutions which form the expansion states. Fig.5.10 shows the deviation in Fermi energy, as compared to the fully converged value, versus the number of expansion states for an $Al_{0.3}Ga_{0.7}As$-GaAs heterojunction. One observes that more expansion states are required, when compared to the quantum well case as presented in fig.5.9, to obtain results with the same accuracy. Again, an increase in density or temperature has a negative influence on the convergence. Although the expansion method exhibits a slower convergence when applied to a heterojunction than compared to a QW, the number of expansion states can still be kept within reasonable limits by correctly choosing the position of the barriers.

Fig.5.10: As fig.5.9 but now for an $Al_{0.3}Ga_{0.7}As$-GaAs heterojunction. Infinite barriers are placed at 50Å and 250Å away from the interface, respectively in the AlGaAs barrier and the GaAs.



Summarising the above results, it follows that fast convergence can be obtained by introducing infinite barriers around the region of interest. Convergence is typically achieved within 10 expansion states for each band when applied to a QW structure, by placing the barriers at 50Å from the interfaces. At such a spacing, the introduction of the barriers will not adversely affect the dispersion. For a heterojunction, convergence typically requires 15 expansion states, provided the barriers are placed at 50Å from the interface in the barrier material and at 250Å from the interfaces in the lower band gap material.

The fast convergence is in contrast to that observed by Ando who used the same expansion method [And85] to calculate the cyclotron effective mass in an AlGaAs-GaAs heterojunction. His observation that at least one hundred expansion states are necessary to achieve convergence can be explained by his choice for the position of the infinite barriers, which are set at a relatively large distance from the heterojunction (respectively 200Å into the barrier material and 1500Å into the GaAs). As discussed, this results in a set of expansion states that differ considerably from the true eigenfunctions and these states have therefore only a limited effect in such an expansion. Furthermore, the cyclotron effective mass is probably more sensitive to the number of expansion states as it is determined by the band structure curvature at the Fermi energy whereas the hole density is proportional to the area under the Fermi-energy contour.

5.1.5 - Fermi-Dirac statistics on hole bands

So far, Fermi-Dirac (FD) statistics have not entered the discussion about the implementation of the self-consistent model presented in this thesis. Although the previous section already presented some results that require the application of Fermi-Dirac statistics to hole bands (see fig.5.9 and fig.5.10), some comments about the exact implementation of FD-statistics will now be made.

For a two-dimensional electron gas in which the dispersion relation takes a simple parabolic form, the electron density at a temperature T is given by an analytic formula,

$$n = \sum_m \frac{m^* kT}{\pi\hbar^2} \ln\left[1 + \exp\left(\frac{E_f - E_m}{kT}\right)\right] \qquad (5.11)$$

where the summation over m runs over the various quantised levels at energy E_m and E_f is the Fermi energy. Unfortunately, due to the valence band warping there is no equivalent analytic formula for the hole density. As a result, the Fermi-integral has to be performed numerically, i.e. one has to evaluate

$$p = \int g_2^{valence}(E) \frac{dE}{1+\exp\left[\frac{E_F - E}{kT}\right]} \tag{5.12}$$

where $g_2^{valence}(E)$ is the two dimensional hole density of states (DOS). As outlined in section 5.1, valence band calculations are relatively demanding, which was exactly the reason for the implementation of the expansion method of sub-section 5.1.3. Although a considerable reduction over exact models was achieved, a full evaluation over *k*-space of the FD-integral (5.12) in each iteration would render the self-consistent procedure into a time-consuming and impractical routine. A simple method to increase the speed is to evaluate the DOS only every couple of iterations. Justification for this follows from the expectation that the form of the dispersion relations (thus the DOS) changes little per iteration, as the change in quantisation energies of various subbands are typically of the order of 0.1 meV after the initial iterations. Yet, even when evaluating the FD-integral every 5 iterations such as done in this work, some approximations to (5.12) have to be made, as follows.

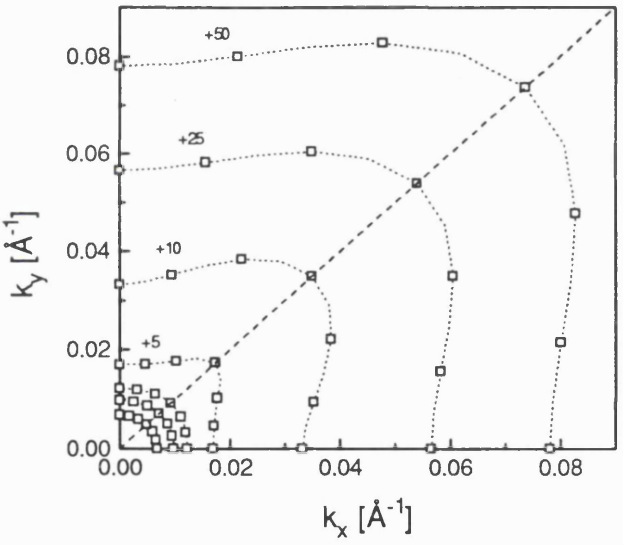


Fig .5.11: Energy contours for the lowest quantised (heavy) hole state in a flat band 100Å $Al_{0.3}Ga_{0.7}As$-GaAs QW. The respective energies are +1, +2, +3, +5, +10, +25 and +50 meV above the band edge minimum which is at 7.19 meV above the GaAs band edge.

First, a limited number of energy contours for the various hole bands are calculated (fig.5.11), exploiting any symmetries present in the *E(k)* diagrams. From this, a discrete density of states function is derived (fig.5.12) which is then integrated by using that the contribution to the integral of each segment is given by an analytical formula similar to (5.11). At this point, one is able to evaluate the hole density as a function of Fermi-energy and temperature in warped bands (fig.5.13).

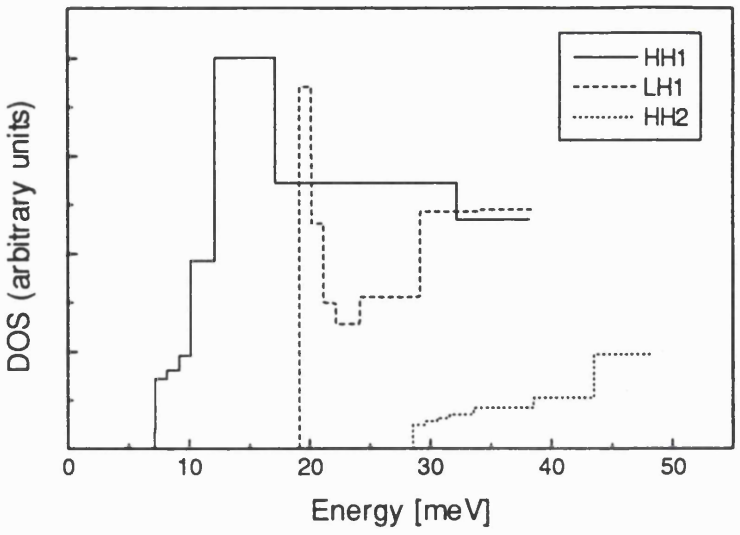


Fig.5.12 : Discrete DOS (Density of states) for the first three quantised states in the QW as described in the previous figure.

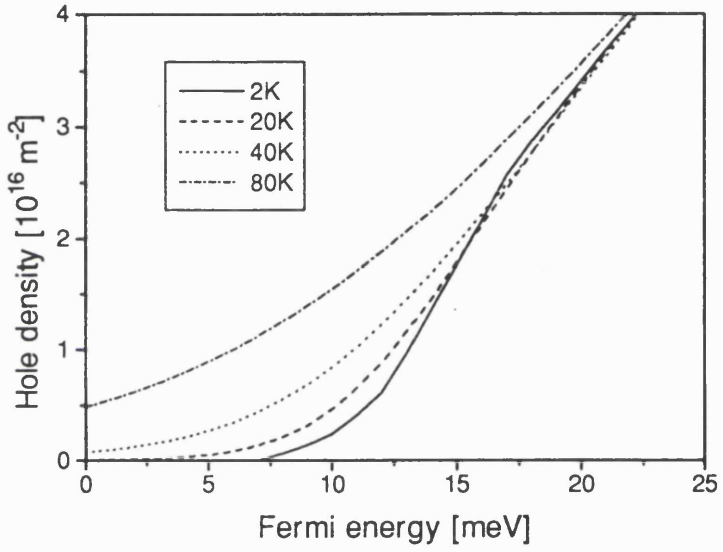


Fig.5.13 : Hole density as a function of Fermi energy for the flat band QW as described in fig.5.11.

The scattering model that will be used to convert the obtained carrier densities and envelope functions into experimentally verifiable quantities like carrier mobility and Hall density will now be discussed.

5.1.5 - The scattering model

The implementation of the scattering model of chapter 4 is much more straightforward than the self-consistent model of the previous sections. The implementation is outlined in the flow chart of fig.5.14 which shows that the process simply consists of evaluating the scattering times per unit angle of the various allowed transitions at a number of scattering angles, after which the total (transport or quantum) scattering rate can be evaluated by integrating over all angles. Note that

only scattering over an angle within the range $[0,\pi]$ has to be evaluated, as the scattering rate corresponding to a deflection $-\alpha$ is equal to that over an angle α.

To minimise the numerical effort involved, (again) several time-saving techniques were implemented, such as making full use of the symmetry of the dielectric response matrix, evaluation of allowed transitions only and a dynamic allocation of points in the curve containing the scattering rate as a function of angle (see fig.5.15).

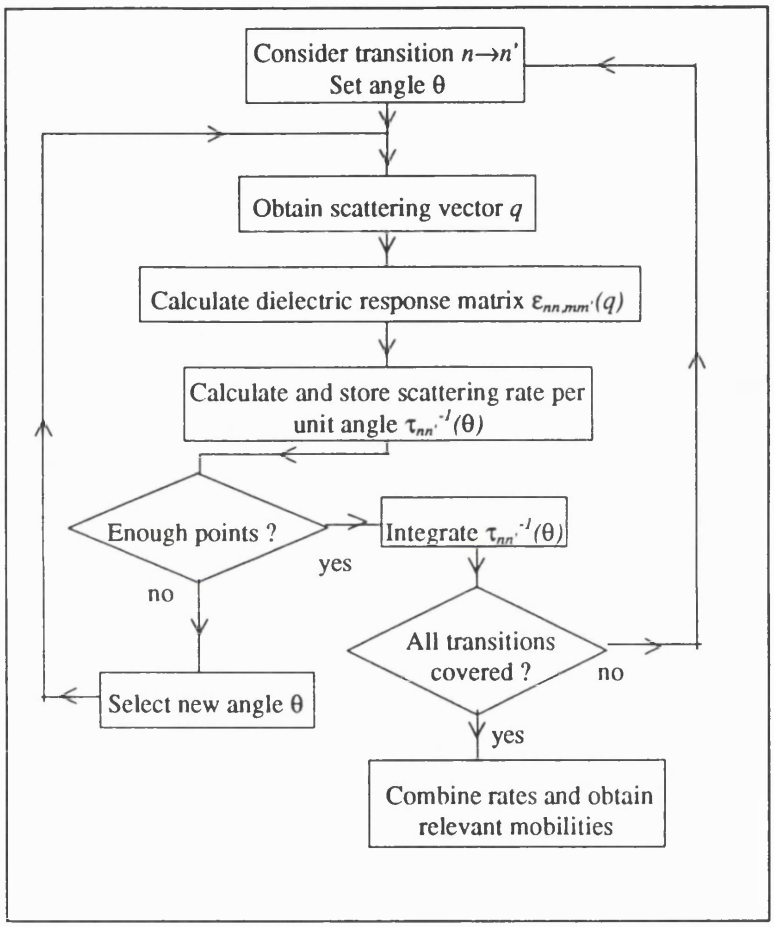
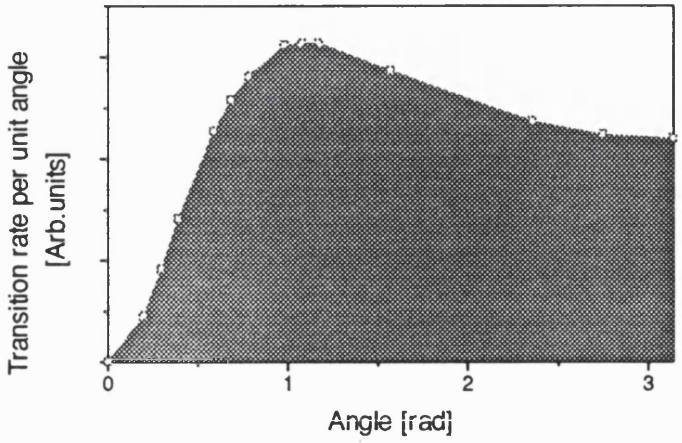


Fig.5.14 : Flow chart corresponding to the chosen implementation of the scattering model.

Fig.5.15: Example of the calculated transition rates per unit angle for an arbitrary n-type structure and transition. Notice the dynamic allocation of data points. The total transition rate is proportional to the shaded area under the curve.



At this point, a fairly extensive insight in the numerical problems involved with performing self-consistent valence band calculations has been presented, along with details of the implementation of Poisson's equation, the effective mass equations and the Fermi-Dirac integral. In addition, the implementation of the ionised impurity scattering model has been presented here. Section 5.2 will now present the required material parameters to actually use the outlined models.

5.2 - Experimental parameters

This section presents all the necessary material constants that were introduced in the theoretical models of chapter 2 to 4. The section is split up in three sub-divisions which respectively deal with the various material constants such as band gaps, mass parameters and deformation potentials (§5.2.1), the band line-up between adjacent heterostructure layers with different material composition (§5.2.2) and parametrised forms of many-body corrections such as exchange and correlation (§5.2.3). This chapter then ends with sub-section 5.2.4 which discusses the interaction of the semiconductor with the outside world via its air-semiconductor interface. Here, surface pinning effects are discussed which force the Fermi-energy to take a fixed value at such an interface. The latter is of vital importance for Poisson's equation, as it provides the remaining boundary condition as discussed in §5.1.2.

5.2.1 - Material constants

A summary of the material parameters that are used in the self-consistent Poisson-Schrödinger solver and the ionised impurity scattering model are presented here. Table 5.1 shows the specific material parameters for the binaries GaAs, AlAs, InAs and InP. A ternary like $In_xGa_{1-x}As$ can be seen as a mixture between the two binaries GaAs and InAs and, as a result, the material constants for such a ternary are obtained by linear interpolation between the two values. An exception is the unstrained band gap E_g^0 for which a linear interpolation is a rather crude approximation. Expressions for the band gap of $Al_xGa_{1-x}As$, $In_xGa_{1-x}As$ and $In_xAl_{1-x}As$ as function of the composition x are therefore presented in (5.13). It should be noted that considerable differences can be found between some of the experimental values quoted here and those used by other authors. This is mainly due to the fact that it is (especially) difficult to obtain accurate results for the higher Luttinger parameters (γ_2 and γ_3) as well as the deformation potentials a,b and d, and that various techniques (experimental or theoretical) can give widely different results (see for example [Zol96] for a list of experimental values for the deformation potential a).

	GaAs	*AlAs*	*InAs*	*InP*
a [Å]	5.6533	5.6611	6.0584	5.8687
E_g^0 *(T=0K) [eV]*	1.51914	3.13	0.4180	1.4236
$m^*(\Gamma)$ *[m_0]*	0.0667	0.15	0.0239	0.0765
γ_1	6.85	3.45	20.4	4.95
γ_2	2.10	0.68	8.3	1.65
γ_3	2.90	1.29	9.1	2.35
C_{11} *[10^{11} Dyn/cm]*	11.879$^{\$}$	12.02	8.329	10.11
C_{12} *[10^{11} Dyn/cm]*	5.376$^{\$}$	5.70	4.526	5.61
C_{44} *[10^{11} Dyn/cm]*	5.95	5.89	3.959	4.56
e_{14} *[C/m^2]*	-0.16$^{\#}$	-0.23$^{\#}$	-0.05$^{\#}$	-0.04$^{\#}$
a [eV/Å]	-9.77$^{\$}$		-6.0	-6.4
b [eV/Å]	-1.7$^{\$}$		-1.8	-2.0
d [eV/Å]	-5.4		-3.6	-5.0
Δ *[eV]*	0.341	0.281	0.38	0.11
$\varepsilon_r(0)$	13.18	10.06	15.15	12.5

Table 5.1 : Material parameters for GaAs, AlAs, InAs and InP. All parameters are from [Lan82] or [Lan87] apart from the entries marked # which were taken from [Sin93] or the ones marked $ which are from [Cha92].

As mentioned above, linear interpolation between the results for the various binaries does not provide accurate results for the band gap in a ternary. Thus, the following values for the unstrained band gap at 300K are used [Dat91],

$$\begin{aligned} E_g^0(Al_xGa_{1-x}As) &= \begin{cases} 1.420 + 1.087x + 0.438x^2 & x < 0.43 \\ 1.905 + 0.10x + 0.16x^2 & x > 0.43 \end{cases} \\ E_g^0(Al_xIn_{1-x}As) &= \begin{cases} 0.37 + 1.91x + 0.74x^2 & x < 0.68 \\ 1.8 + 0.4x & x > 0.68 \end{cases} \\ E_g^0(In_xGa_{1-x}As) &= 1.424 - 0.7x + 0.4x^2 \end{aligned} \tag{5.13}$$

The temperature dependence of the band gap of all materials is set to the temperature dependence of GaAs, such that the unstrained band gap at a temperature T is given by

$$E_g^0(T) = E_g^0(T = 0K) - \frac{5.408 \cdot 10^{-4} T^2 \, eV}{T(T + 204K)} \tag{5.14}$$

5.2.2 - Band line-ups

When modelling compositionally varying structures, it is important to know how the conduction and valence bands line up with each other. The line-up of the band under consideration determines the effective potential energy $V_h(z)$ which via (5.1) provides the confinement potential $V(z)$ required in the effective mass equations.

The band line-up is usually expressed in terms of the band offsets ΔE_c and ΔE_v (or their ratio $Q=\Delta E_c/\Delta E_v$) which may be related, as a first approximation, to the differences in work functions (similar to the approach taken when evaluating an ideal Schottky barrier [Sze81]) and the band gaps of the materials involved. However, more detailed investigations reveal that the offsets are considerably influenced by the occurrence of charge redistribution across the interfaces as a result of the differences in the adjacent electron configurations [Wal86]. The latter, combined with the relatively insensitivity of the lower quantised levels with respect to the offset, explains the significant differences between the presently accepted and early values for the offsets in for example the $Al_xGa_{1-x}As/GaAs$ system [Cha74,Din74].

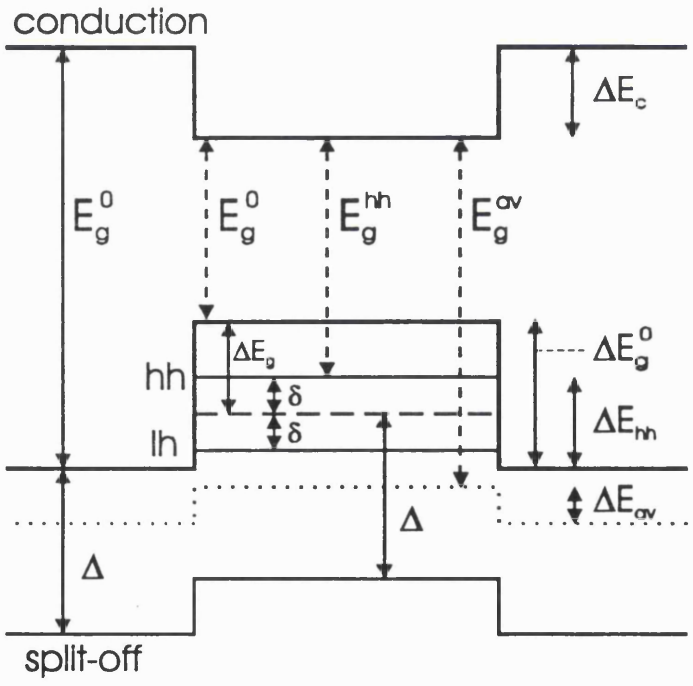


Fig.5.16: Definition of the various symbols used in the band offset equations. The band diagram depicts a compressively strained QW structure in which the heavy and light hole bands have been split up via ΔE_g andδ as defined in the Pikus-Bir strain Hamiltonian of chapter 2. Shown are three definitions for the band gap: E_g^0 is the unstrained band gap, E_g^{hh} the true band gap and E_g^{av} the difference between the conduction band and the average energy of the heavy, light and split-off band (represented by a dotted line). All three definitions are used in the literature to describe the valence band offset (thus specifying either ΔE_v, ΔE_{hh} or ΔE_{av}).

As outlined in section 2.5.2, strain causes the heavy and light hole bands to split. As a result, a variety of band gaps can be defined in a strained structure such as (see fig.5.16)

- The unstrained band gap E_g^0
- The actual band gap that is given by the energy difference between the conduction and heavy hole band E_g^{hh} in a compressively strained structure (E_g^{lh} in a tensile strained structure)
- The average band gap E_g^{av} given by the energy difference between the conduction band and the average energy of the heavy, light and split-off band.

This ambiguity results in many different ways to describe the offsets. In this work will employ the following results (again, see fig.5.16 for the various definitions),

$$\text{AlGaAs-GaAs: } \begin{cases} \Delta E_c &= 0.62\,\Delta E_g^0 \\ \Delta E_v &= 0.38\,\Delta E_g^0 \end{cases} \quad \text{[Bry93]}$$

$$\text{InGaAs-GaAs: } \begin{cases} \Delta E_c &= 0.66\,\Delta E_g^{hh} \\ \Delta E_{hh} &= 0.34\,\Delta E_g^{hh} \end{cases} \quad \text{[Am92]} \qquad (5.15)$$

$$\text{InGaAs-InP: } \begin{cases} \Delta E_c &= 0.36\,\Delta E_g^{av} \\ \Delta E_{av} &= 0.64\,\Delta E_g^{av} \end{cases} \quad \text{[Lee90]}$$

Using the energy gaps, spin-orbit splitting and deformation potentials given in table 5.1, the band line-up of all relevant bands (conduction, heavy, light and split-off hole) have been uniquely defined via (5.15) and the Pikus-Bir Hamiltonian (2.76).

5.2.3 - Many body effects: the exchange and correlation potential

As discussed in §2.1.1, the one-electron Schrodinger equation that is the starting point of the derivation of the effective mass model of section 2.3 contains a many-body correction which is the exchange and correlation potential. It originates from the fact that both the Pauli-principle and the Coulomb repulsion tend to keep all electrons as far apart as possible, therby lowering the potential energy. For the conduction band, a simple analytic parametrisation was given by Hedin and Lundqvist [Hed71,Ste84]

$$V_{xc}(z) = -\left[1 + 0.7734 \cdot \frac{r_s}{21} \ln\left(1 + \frac{21}{r_s}\right)\right] \frac{2}{\pi\alpha r_s} Ry^* \qquad (5.16)$$

where

$$\alpha = (4/9\pi)^{1/3}$$
$$r_s = \left[\tfrac{4}{3}\pi (a^*)^3 n(z)\right]^{-1/3}$$
$$a^* = \frac{4\pi\varepsilon_0\varepsilon_r\hbar^2}{m^* e^2} \qquad (5.17)$$
$$Ry^* = \frac{m^* e^4}{2(4\pi\varepsilon_0\varepsilon_r\hbar)^2}$$

Here, a^* and Ry^* are the effective Bohr radius and effective Rydberg. For GaAs, they respectively take the values 104Å and 5.3 meV. Inclusion of exchange and correlation results in a potential correction of typically 40 meV at an electron density of 10^{18} m^{-3}. As the effect of the exchange and correlation correction is much smaller for holes [Bob96], it is neglected in this work.

5.2.4 - Surface pinning

As Poisson's equation (5.4) only operates on the second derivative of the potential, this only supplies a unique potential function within two boundary conditions. Such a boundary condition consists of specifying either the absolute value or the derivative of the potential at a given point. One boundary condition was already presented in 5.1.2 where it was postulated that the substrate remains in a charge neutral state. Effectively, it is assumed that the surface impurities on the growth side of the substrate completely shield this substrate from the rest of the heterostructure. Thus, the potential at this semiconductor-substrate interface is pinned at charge equilibrium. The remaining boundary condition will now be discussed.

It is well known that the surface-air interface exhibits surface states that effectively pin the potential at a fixed level. The position of such a pinning level could in principle be determined by comparing self-consistent electrostatic calculations for the carrier density with the corresponding experimental results, but there is an alternative (and more sensitive) technique in which information about the pinning potential is obtained from a photoreflectance experiment. In such an experiment, as suggested by [Hoo89], the electric field present in the (intrinsic) top layer in a [s-i-n^+] or [s-i-p^+] structure (s denotes the surface) is measured from which the pinning potential can be deduced. A good review of the technique as well as a summary of the various experimental results can be found in for example [She95b]. The surface pinning potentials that will be used in this work are now presented.

For *n*-type structures, the pinning potential at an [001] GaAs-Air interface is assumed to be pinned at mid gap [Yin92]. Although a similar result was quoted for *p*-type structures in this paper, more recent research showed a reduced surface state density (i.e. less effective pinning behaviour) and the presence of two pinning levels at respectively 0.25V and 0.50V above the valence band [Yan94], i.e. well below mid gap as compared to the *n*-type case. A possible explanation for the different pinning behaviour and reduced surface state density of *n*- and *p*-type GaAs comes from an STM study by Pashley [Pas93] concerning the formation of donor kink sites. Returning to the implementation of Poisson's equation and the boundary condition in particular, in this work the potential is assumed to be pinned at mid gap in an *n*-type structure and 0.3V above the valence band in a *p*-type structure. Fig.5.17 shows a graphical picture of the boundary condition implementation.

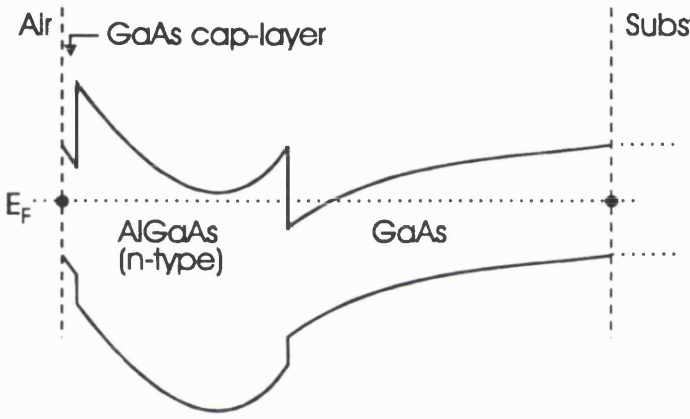


Fig.5.17 : In order to be able to solve Poisson's equation, two boundary conditions have to be supplied. The first is obtained by assuming the substrate to be charge neutral. The second follows from the experimental observation that the surface potential is pinned midgap for n-type and around 0.3 V above the GaAs valence band for p-type structures.

With this information, the electrostatic potential has been uniquely defined through Poisson's equation. Using this, the implementation of the various equations as presented in section 5.1, material constants, band-line ups and parametrised form of the many-body correction, the total framework of finding the wavefunctions and dispersion in a compositionally varying structure including charge has been presented, along with that to obtain the theoretical carrier mobility. The next chapter will now present band structure calculations and theoretical mobilities for a selection of *n*- and *p*-type structures.

Chapter 6 : Results

A significant proportion of this thesis has so far been devoted to the outline, derivation, implementation and discussion of a valence band effective mass model. Chapter 2 contained the essentials of the derivation of this model, whereas chapter 3 concentrated on the intrinsic boundary condition problem. The reason for the extensive coverage is the fact that substantial advances were made during this Ph.D. in the boundary condition problem for growth on non-conventional growth directions. This led to two publications on this matter [Sta97,Dal98] of which the results will now be presented in the first section of this chapter. Here, a formalism is introduced that provides the correct boundary conditions for arbitrary growth directions via relatively simple analytic expressions for the operator ordering. Furthermore, the relevance of using the correct boundary conditions is highlighted, especially when calculating a theoretical momentum effective mass for use in a scattering model.

Section 6.2 focuses on the use of the valence band effective mass model to calculate the eigenstates in heterostructures. Numerical results for the dispersion $E(k)$ as a function of growth direction are presented, along with the effects of the spin-orbit interaction, strain, and spin-splitting such as observed in non-symmetric structures. Qualitative results about their respective influence on the transport properties are given.

Theoretical electron mobility calculations for a range of GaAs-AlGaAs and InGaAs-InAlAs-InP QWs are presented in section 6.3. These results are compared with others presented in the literature, both to confirm the validity of the model and to gain insight into the accuracy of the RPA scattering model. It is shown that a scaling factor of approximately 2 is needed to bring the theoretical and experimental electron transport mobilities into agreement, confirming similar findings in the literature. As it is found that this factor is fairly independent of sample structure or even material system, it is thus established that the RPA approach can give reasonably accurate absolute values for the mobility by simply accepting the presence of such a scaling factor, whose origin will be discussed later on. Also in this section, special attention is paid to the effects of interband scattering. It is shown that the over-estimation of the electron transport mobility in graded QW structures in a previous evaluation by Roberts [Rob96,Rob98] is due to the failure to include this scattering process. The improvement in mobility obtained in his work is shown to be severely limited by the latter process.

The final section in this chapter contains an evaluation of hole transport in two-dimensional semiconductor structures. The previously obtained insight into the effects of strain and spin-splitting on the valence band dispersion, combined with the relevance of interband scattering, was utilised to design a set of channel doped *p*-type $In_{0.15}Ga_{0.85}As$-$Al_{0.33}Ga_{0.67}As$ QW structures to evaluate the possibilities and scope of wavefunction engineering to enhance the hole transport mobility. It is shown that the RPA approach can provide reasonably accurate relative predictions for *p*-type structures when comparing between various doping schemes, but unfortunately provides absolute values of the hole mobilities that are generally a factor 8 too high in comparison with experimental results. As for the electron case, a scaling factor is introduced to bring the experimental and theoretical mobilities in agreement, but it is disappointing to find a scaling factor that is so much larger than that introduced for *n*-type calculations. This chapter then finishes with a discussion on the origin of this scaling factor. Whereas the scaling factor in *n*-type materials is attributed to electron-electron interaction, it is postulated that the larger fraction of the hole scaling factor is most likely caused by the over-estimation of the screening in the RPA model due to quantum confinement effects in the screening states. A first approach to account for this in the scattering model leads to a new scaling factor of 4.5 which is in better agreement with the electron factor, as well as an overal improvement between experimental and theoretical results.

6.1 - Boundary conditions for arbitrary growth directions

It was shown in chapter 3 that the 'new' (Burt-Foreman) boundary conditions improve the consistency and behaviour of the $k \cdot p$ model, and that adopting these boundary conditions can have a significant influence on the resulting bandstructure. Nevertheless, it was found that the (old) symmetric boundary conditions were still being used when looking at non-[001] growth as only these boundary conditions had been published in the literature. The underlying reason why little effort was made to extend the formalism to non-[001] growth was in the author's opinion probably due to (at least) three reasons,

- The mathematics involved in Burt's review paper tends to swamp any physical insight.
- Neither Foreman [For93] nor Meney [Men94] present any information about how the boundary conditions for [001] were obtained.
- A general under-estimation of the need to adopt the new boundary conditions when evaluating experimental quantities of interest such as electrical or optical properties in technologically relevant material systems. Foreman chose InGaAs-GaAs to illustrate the effects but limited the discussion to the band edge masses which only exhibit significant changes in the second heavy

hole level. Meney provided full *E(k)* dispersions but these unfortunately only showed significant changes in the InGaAsSb-AlGaSb system for the chosen (relatively large) well width of 100Å. The influence of the boundary conditions on the dispersion however increases dramatically with decreasing well width, as will be shown later in this chapter.

Thus, there was a clear need for a comprehensive framework to derive the new boundary conditions for non-[001] growth. This subject was addressed in two papers. In the first paper [Sta97], a framework was presented to set up an effective mass Hamiltonian in a similar way as done in the bulk approach (2.47-2.50) with end result (2.51) for growth on [001], but now including information about the operator ordering. As such an approach is not limited to just [001] as shown in §2.3.5, it enabled the derivation of the boundary conditions for [110] which were subsequently also presented there. It was shown that the effects of adopting the new boundary conditions for the commonly used material system $In_{0.53}Ga_{0.47}As/InP$ were dramatic for this particular growth direction. Although this approach can be repeated for any growth direction, the procedure can be rather tedious and requires an explicit knowledge of the angular momentum basis set for the growth direction under consideration. In the second paper [Dal98], a more satisfactory approach was presented in which a set of analytic rules provides the correct operator ordering for any growth direction. The above work on the boundary condition problem was performed in collaboration with Dr.Paul N.Stavrinou at the IRC for Semiconductor Materials in Oxford, presently at the IRC for Semiconductor Materials at Imperial College.

6.1.1 - Framework to extend the Burt-Foreman formalism to non-[001]

Using the insight into effective mass theory obtained in chapter 2, a framework to derive the boundary conditions for any growth direction is presented here. At this point, it is important to point out that the orbital-picture of effective mass theory proved essential in this work. Only because of the acquired knowledge of the symmetry properties of the various zone-centre solutions and the explicit derivation of the effective mass Hamiltonians using a tight-binding view, thereby explicitly avoiding Luttinger's 'Theory of invariants' (§2.3.7), was it possible to contribute to and eventually solve the boundary condition problem within Burt's framework for non-[001] growth. With hindsight, it is therefore that the theory chapter had a strong focus on group theory, symmetry properties and zone-centre solutions for the case where spin and momentum were orientated along non-conventional directions. This obtained insight in effective mass theory in terms of orbitals and inner products is now going to be used to solve the boundary condition problem.

It was shown in chapter 2 that the valence band effective mass Hamiltonian in the compositionally varying case essentially only differs from that for bulk in the sense that one is now working with operators $\hat{k}$ rather than momentum vectors k. As a result, almost the whole of chapter 2 dealt with the bulk case only, making the substitution $k \leftrightarrow \hat{k}$ whenever necessary. Chapter 3 showed that such an approach cannot provide the correct boundary conditions, as all information about the operator ordering is lost when parametrising the Hamiltonian (2.38). However, the solution is obvious as the boundary conditions can simply be obtained by parametrising the Hamiltonian, but taking the correctly *ordered* form as its origin : [identical to (3.7)]

$$H_{jj'} = \hbar\hat{k}_\alpha D_{jj'}^{\alpha\beta}\hbar\hat{k}_\beta$$
$$D_{jj'}^{\alpha\beta} = \frac{\hbar^2}{2m_0}\delta_{jj'}\delta_{\alpha\beta} + \frac{\hbar^2}{m_0^2}\sum_{\nu}\frac{\langle j|p^\alpha|\nu\rangle\langle\nu|p^\beta|j'\rangle}{E - E_\nu} \tag{6.1}$$

As the bulk unordered Hamiltonian over an $|X\rangle$, $|Y\rangle$ and $|Z\rangle$ basis (2.47) was the starting point of all Hamiltonians for all growth direction, a similar but ordered equivalent is going to be set up. Thus, the effect of adopting the ordered Löwdin interaction term on the individual terms in each element of the Hamiltonian over this basis $|X\rangle$, $|Y\rangle$ and $|Z\rangle$ needs to be evaluated.

The element $H_{12} = \langle X|H|Y\rangle$ is taken as an example, all other elements can be obtained in a similar approach. As this particular term is proportional to k_xk_y in the bulk case, the evaluation of (6.1) can be restricted to just the contributions proportional to $\hat{k}_x\hat{k}_y$ or $\hat{k}_y\hat{k}_x$ Following Foreman [For93], the individual contributions of the remote states of symmetry Γ_1, Γ_{15} and Γ_{12}, thereby neglecting the small contribution of Γ_{25} , are evaluated. It then follows from fig.6.1 that the only non-zero contributions to this Hamiltonian matrix element are

$$H_{12} = \frac{\hbar^2}{m_0^2}\sum_{\nu=S}^{[\Gamma_1]}\hat{k}_x\frac{\langle X|\partial/\partial x|\nu\rangle\langle\nu|\partial/\partial y|Y\rangle}{E - E_\nu}\hat{k}_y + \frac{\hbar^2}{m_0^2}\left\{\sum_{\substack{\nu=X,Y,Z,\\ d_{xy},d_{xz},d_{yz}}}^{[\Gamma_{15}]} + \sum_{\nu=d_{x^2-y^2},d_{3z^2-r^2}}^{[\Gamma_{12}]}\right\}\hat{k}_y\frac{\langle X|\partial/\partial y|\nu\rangle\langle\nu|\partial/\partial x|Y\rangle}{E - E_\nu}\hat{k}_x \tag{6.2}$$

Using the parameters as defined in (3.8), the following parametrised form of this ordered Hamiltonian element is obtained,

$$H_{12} = -\frac{\hbar^2}{2m_0}\left(6\hat{k}_x\sigma\hat{k}_y + 6\hat{k}_y\pi\hat{k}_x + 6\hat{k}_y\delta\hat{k}_x\right) \tag{6.3}$$

The other elements are derived in a similar way, and the following Hamiltonian over an $|X\rangle$, $|Y\rangle$ and $|Z\rangle$ basis is found [Sta97]

$$H=\frac{\hbar^2}{2m_0}\begin{pmatrix} \hat{k}_xA\hat{k}_x+\hat{k}_yB\hat{k}_y+\hat{k}_zB\hat{k}_z & \hat{k}_xC_1\hat{k}_y+\hat{k}_yC_2\hat{k}_x & \hat{k}_xC_1\hat{k}_z+\hat{k}_zC_2\hat{k}_x \\ \hat{k}_yC_1\hat{k}_x+\hat{k}_xC_2\hat{k}_y & \hat{k}_yA\hat{k}_y+\hat{k}_xB\hat{k}_x+\hat{k}_zB\hat{k}_z & \hat{k}_yC_1\hat{k}_z+\hat{k}_zC_2\hat{k}_y \\ \hat{k}_zC_1\hat{k}_x+\hat{k}_xC_2\hat{k}_z & \hat{k}_zC_1\hat{k}_y+\hat{k}_yC_2\hat{k}_z & \hat{k}_zA\hat{k}_z+\hat{k}_xB\hat{k}_x+\hat{k}_yB\hat{k}_y \end{pmatrix} \tag{6.4}$$

where

$$\begin{aligned} A&=1-6\sigma-12\delta & C_1&=-6\sigma+6\delta \\ B&=1-6\pi & C_2&=-6\pi \end{aligned} \tag{6.5}$$

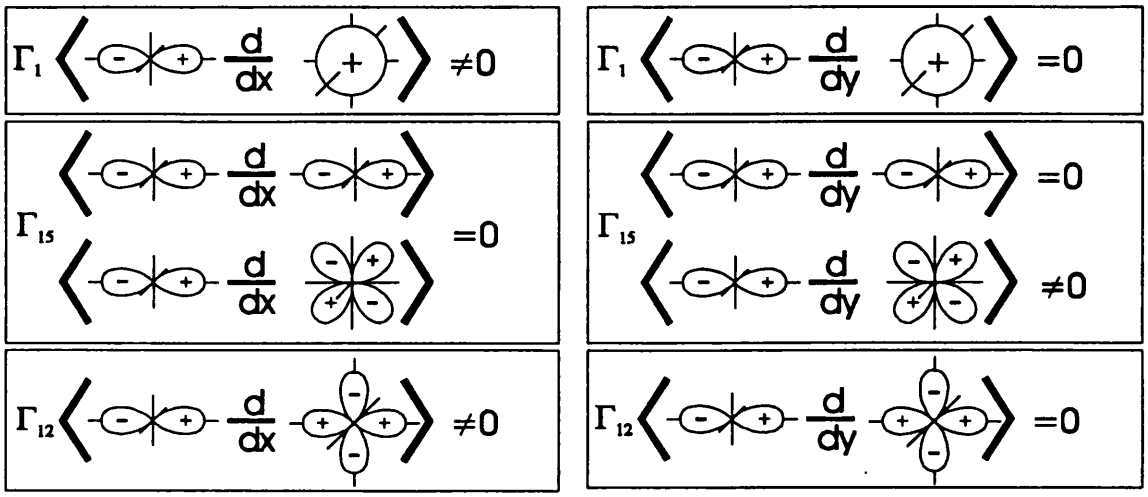


Fig.6.1 : Schematic of the various inner products $\langle X|p_x|v\rangle$ and $\langle X|p_y|v\rangle$. The z-coordinate corresponding to the quantisation direction is taken perpendicular to the paper, the in-plane coordinates x and y are respectively taken as the horizontal and vertical direction.

At this point, a concise formalism has been provided to set up an ordered Hamiltonian for *any* growth direction. By replacing the bulk parametrised Hamiltonian over an $|X\rangle$, $|Y\rangle$ and $|Z\rangle$ basis (2.38) by the properly ordered version as given above, the $k\cdot p$ formalism has been extended such that it is now valid not just in bulk but also at an interface. Using the ordered Hamiltonian (6.4), one can now, identically to the approach shown in chapter 2, set up an effective mass Hamiltonian for an angular momentum basis $|J,m_j\rangle$ (2.26) for any growth direction. Using this approach, the boundary conditions for a 4-band [110] effective mass model were derived and presented in [Sta97]. Similar results to those found by Foreman when looking at [001]-growth were obtained such as reduced interface coupling between the heavy and light hole states as well as a new interface term coupling both light-hole states. Explicit forms for the boundary conditions will be presented in §6.1.4. Numerical results for the dispersion as a function of the choice of boundary conditions for a variety of material systems will be given in section 6.2.

6.1.2 - Analytic rules to obtain the Burt-Foreman boundary conditions

A general approach to obtain the boundary conditions for any growth direction was presented in the previous discussion. It was shown that the ordered Hamiltonian over an $|X\rangle$, $|Y\rangle$ and $|Z\rangle$ basis (6.4) is a natural starting point and that, along with an appropriate coordinate transformation, the problem merely consists of recasting the Hamiltonian into the more usual angular momentum basis $|J,m_j\rangle$ for the desired growth orientation. However, this procedure can rapidly become laborious, needs to be performed for every growth direction individually and therefore prevents general observations with respect to the interface coupling. A more satisfactory approach would be to work from *un*ordered bulk-like Hamiltonians, which are readily available in the literature [Iko92,Men92] or easily obtained from the theory of invariants [Lut56] and, where necessary, to replace terms by the ordered form prescribed from the exact theory. Essentially, one would like to have a set of rules to set up a properly ordered Hamiltonian, much like the (now shown to be incorrect) symmetrisation rule that used to be employed before the development of the exact envelope function theory.

It was shown in [Dal98] that, with the aid of the general $|J,m_j\rangle$-basis, a set of analytic rules can be found that provide the correct operator ordering within Burt's framework for each possible growth direction. One obtains, again regardless of growth direction, that the overall form of a 6-band effective mass Hamiltonian with operator ordering takes the form (boxed are the elements corresponding to a 4-band model)

$$H=\begin{pmatrix} \boxed{\begin{matrix} P+Q & -S_1 & R & 0 \\ & P-Q & \Theta C & -R \\ & & P-Q & -S_2 \\ & & & P+Q \end{matrix}} & \begin{matrix} S_1/\sqrt{2} & \sqrt{2}R \\ \sqrt{2}Q & \Pi\sqrt{\frac{3}{2}}S_3 \\ -\Pi^{\dagger}\sqrt{\frac{3}{2}}S_4^{\dagger} & \sqrt{2}Q \\ -\sqrt{2}R^{\dagger} & S_2^{\dagger}/\sqrt{2} \end{matrix} \\ \begin{matrix} P+\Delta & -\Theta C \\ & P+\Delta \end{matrix} & \end{pmatrix} \tag{6.6}$$

where the subscripts on S terms are introduced to distinguish elements that have a specific operator ordering. Note that if the ordering is neglected, e.g. as in the bulk descriptions often used in the literature, then S_1, S_2, S_3 and S_4 will all reduce to a single term S.

An important new feature of the above Hamiltonian is a new element C that couples both light hole bands as well as both split-off hole bands. This element was already introduced in the block diagonal Hamiltonian (2.57). Its presence is solely due to the inclusion of the operator ordering and is therefore expected to be only non-zero at an interface. Further, phase terms of unit length, Π

and Θ, have been introduced in order to accommodate any differences in phase of existing $|J,m_j\rangle$ sets used in the literature compared to the set given in (2.26,2.27), an issue that will be returned to in section 6.1.3; for the present set both Π and Θ would be equal to 1. Complying with the latter and ignoring any ordering, the above Hamiltonian of course reduces to the previously given bulk Hamiltonian (2.51).

The relevance of the above Hamiltonian, and its bulk counterpart, should be noted, and was partially highlighted in §2.3.7. There, it was outlined that the convenient group theoretical 'Theory of invariants' can only provide 4x4 Hamiltonians, but as all the elements of a 6x6 model are related to the latter, in principle it does not prevent the use of this theory when setting up the higher order model.

The rules for the appropriate operator ordering in the various elements in the effective mass Hamiltonian are now presented. As mentioned earlier, the derivation of the above Hamiltonian and these rules was only possible because of the orbital approach developed by the author, accumulated into the general basis set (2.26). The necessary linear algebra involved with the basis transformation to $|J,m_j\rangle$ [Dal98] was performed with the aid of MATHEMATICA (see appendices A1,A2,A3). As the algebra is straightforward but cumbersome, this section will restrict itself to just presenting the results obtained. Identical to the approach taken in 2.3.5, a new orthonormal coordinate system (x',y',z') (with z' parallel to the growth direction) is introduced in the Hamiltonian over an $|X\rangle$, $|Y\rangle$ and $|Z\rangle$ basis (6.4), after which basis transformation to the appropriate $|J,m_j\rangle$-set is performed. Thus, the general Hamiltonian (6.6) is obtained with the following operator ordering rules:

- P and Q are symmetrised with respect to the operator $\hat{k}_{z'}$.
- All terms in S_i involving $\hat{k}_{z'}^2$ are to be symmetrised.
- Any terms in S_i containing a single operator $\hat{k}_{z'}$, i.e. with a unordered form such as $\hat{k}_{z'}(f\sigma + g\pi + h\delta)$ where f,g and h are constant coefficients, should be ordered as follows;

$$
\begin{aligned}
S_1&: \left(f\sigma - \frac{f-g}{2}\pi - \frac{f-h}{2}\delta\right)\hat{k}_{z'} + \hat{k}_{z'}\left(\frac{f+g}{2}\pi + \frac{f+h}{2}\delta\right)\\
S_2&: \left(\frac{f+g}{2}\pi + \frac{f+h}{2}\delta\right)\hat{k}_{z'} + \hat{k}_{z'}\left(f\sigma - \frac{f-g}{2}\pi - \frac{f-h}{2}\delta\right)\\
S_3&: \left(\frac{f}{3}\sigma + \frac{f+3g}{6}\pi + \frac{f+3h}{6}\delta\right)\hat{k}_{z'} + \hat{k}_{z'}\left(\frac{2f}{3}\sigma - \frac{f-3g}{6}\pi - \frac{f-3h}{6}\delta\right)\\
S_4&: \left(\frac{2f}{3}\sigma - \frac{f-3g}{6}\pi - \frac{f-3h}{6}\delta\right)\hat{k}_{z'} + \hat{k}_{z'}\left(\frac{f}{3}\sigma + \frac{f+3g}{6}\pi + \frac{f+3h}{6}\delta\right)
\end{aligned}
\qquad (6.7)
$$

- All terms in R are symmetrised with respect to the operator $\hat{k}_{z'}$ and do not contain any combination of σ and $\hat{k}_{z'}$.
- The term C, which is only non-zero at an interface, always takes the form;

$$-(\hbar/m_0)(k_{x'} - ik_{y'})\left[(\sigma - \pi - \delta)\hat{k}_{z'} - \hat{k}_{z'}(\sigma - \pi - \delta)\right] \tag{6.8}$$

Summarising, two significant results have been obtained. First is the occurrence of an element C that couples the light hole bands as well as the split-off bands. This element always takes the same form, independent of growth direction. Second, a strong asymmetry in the σ-coefficient for the S-elements is observed. Integrating across a boundary, it follows that coupling between heavy and light hole states *never* includes the *dominant* contribution from remote states with s-symmetry (Γ_1), which is in contrast to the result one would have obtained if the normal symmetrisation procedure had been used. As a result, significant changes in the calculated $E(k)$ diagrams for *all* growth directions are expected when adopting the new boundary conditions, such as shown for [110] in [Sta97].

6.1.3 : Influence of the phase of the $|J,m_j\rangle$-basis on the general form of the Hamiltonian

This section explains how, with the introduction of the two phase terms Π and Θ, existing Hamiltonians which are ***not*** derived from the set (2.26,2.27) may be recast into the general form (6.6). The appearance of different Hamiltonians in the literature basically stems from the differences in the basis set used to derive them. Although two commuting observables have been chosen and a particular projection axis has been specified, the $|J,m_j\rangle$ states of (2.26,2.27) are by no means unique in the sense that each may be multiplied by a unit length factor of $(\mu+i\nu)$ without changing any of the physics. Typically, using a $|J,m_j\rangle$ basis corresponding to a different phase convention becomes apparent in the off-diagonal elements of the Hamiltonian which then differ by some (complex) phase term from the general form of the Hamiltonian used in this work. To illustrate, there are at least two general forms of 4-band Hamiltonians in the literature which exhibit phase differences in the off-diagonal elements, compare for example [Men92,She95,Hen96] to [Iko92,For93]. The phase convention in this work corresponds to the latter notation. The following section now demonstrates that it is possible to account for the phase of the $|J,m_j\rangle$ set by introducing two phase factors Π and Θ.

Consider the unitary transformation $H'=U^{\dagger}HU$, where U is a diagonal matrix with elements [Dal98],

$$U_{ij}=\delta_{ij}u_i\ ,\ \boldsymbol{u}=\left(1,\exp[i\alpha],\exp[i\beta],\exp[i(\alpha+\beta)],\exp[i\alpha],\exp[i\beta]\right) \tag{6.9}$$

and δ_{ij} is the Kronecker delta. As a result of the transformation the following elements in the Hamiltonian are multiplied by a phase factor

$$\begin{array}{ll} S_{1,2} & \rightarrow \exp[i\alpha]S_{1,2} \\ R & \rightarrow \exp[i\beta]R \\ S_{3,4} & \rightarrow \exp[i(\beta-\alpha)]S_{3,4} \\ C & \rightarrow \exp[i(\beta-\alpha)]C \end{array} \tag{6.10}$$

To demonstrate how this unitary transformation can be used to account for the choice in $|J,m_j\rangle$ states in (6.6), a simple 4-band Hamiltonian without C-element is considered first (i.e. only the top two transformations are of relevance). It follows from (6.10) that one can multiply all the S as well as R elements in such a 4-band Hamiltonian with a phase factor without changing the eigenvalues. Effectively one is now working with new off-diagonal elements, $S_{new}=\exp(i\alpha)S$ and $R_{new}=\exp(i\beta)R$, which are based on a different $|J,m_j\rangle$ set as compared to the original matrix elements.

The latter is now extended to a full 6-band model with C-elements. Again, S_{new} and R_{new} represent 'arbitrary' off-diagonal elements which correspond to a yet unknown phase and $|J,m_j\rangle$ set (for example, they could have been obtained directly from the theory of invariance [Lut56]). As the C-element is not invariant in the transformation (6.10) one finds that a phase factor $\Theta=\exp[i(\beta-\alpha)]$ has to be introduced to account for any difference in basis states. Similarly as $S_{3,4}$ transform differently from $S_{1,2}$, although they were all defined to reduce to the same element neglecting any ordering, another phase factor Π has to be introduced which is placed in the Hamiltonian elements containing $S_{3,4}$. Using (6.10), one obtains,

$$\Pi=\frac{\exp[i(\beta-\alpha)]}{\exp[i\alpha]}=\exp[i(\beta-2\alpha)] \tag{6.11}$$

Both factors can be obtained by comparing the chosen R and S to the expressions one would have obtained when using the set (2.26,2.27). As the latter set will always give a real and positive σk_z^2

component in R and a real and positive $\sigma k_{x'}\hat{k}_{z'}$ component in S (this can be checked by evaluating $\langle 3/2,+3/2|H|3/2,\pm 1/2\rangle$), this leads to the following expressions for Π and Θ

$$\begin{aligned}\Pi &= -\left(s^{\sigma}_{x'z'}{}^{*}\right)^{2} r^{\sigma}_{x'x'}\, Sign(P)\\ \Theta &= -s^{\sigma}_{x'z'}{}^{*} r^{\sigma}_{x'x'}\, Sign(P)\end{aligned} \tag{6.12}$$

where $r^{\sigma}_{x'x'}$ is the normalised $\sigma k^2_{x'}$ coefficient in R and $s^{\sigma}_{x'z'}$ is the normalised $\sigma k_{x'}\hat{k}_{z'}$ coefficient in S. Here, the factor Sign(P) was introduced to eliminate the effects of an overall sign change in the Hamiltonian often performed to obtain positive hole energies. To illustrate, the next section will now use the presented analytic rules to derive the appropriate boundary conditions for growth on [111].

6.1.4 - Explicit form for the boundary conditions for growth along [111]

Two frameworks to derive boundary conditions for arbitrary growth directions were presented in the previous sections, the first using implicit knowledge of $|J,m_j\rangle$ set and the second using a set of analytic rules. Following [Dal98], the second (ultimately more appealing) framework will now be used to obtain the boundary conditions for growth along [111].

Using the 'Theory of invariants' [Lut56], taking the axes along (1,1,-2), (-1,1,0) and (1,1,1) and dropping the primes from now onwards, the following unordered Hamiltonian elements for growth along the [111] direction are obtained.

$$\begin{aligned}P &= -\frac{1}{2}\gamma_1(k_x^2+k_y^2+\hat{k}_z^2)\\ Q &= -\frac{1}{2}\gamma_3(k_x^2+k_y^2-2\hat{k}_z^2)\\ S &= \frac{i}{\sqrt{6}}(\gamma_2-\gamma_3)k_+^2-\frac{i}{\sqrt{3}}(2\gamma_2+\gamma_3)k_-\hat{k}_z\\ R &= -\frac{\sqrt{3}}{6}(\gamma_2+2\gamma_3)k_-^2+\frac{\sqrt{6}}{3}(\gamma_2-\gamma_3)k_+\hat{k}_z\end{aligned} \tag{6.13}$$

where $k_{\pm}=k_x\pm k_y$. Both the R and S elements feature terms linear in $\hat{k}_z$. In these elements, σ,π,δ are substituted for the Luttinger parameters after which the ordered replacements can be obtained using the rules of §6.1.2,. Explicitly,

$$
\begin{aligned}
S_1&: \ (3\sigma - 2\pi)\hat{k}_z + \hat{k}_z(\pi + 3\delta) \\
S_2&: \ (\pi + 3\delta)\hat{k}_z + \hat{k}_z(3\sigma - 2\pi) \\
S_3&: \ (\sigma + 2\delta)\hat{k}_z + \hat{k}_z(2\sigma - \pi + \delta) \\
S_4&: \ (2\sigma - \pi + \delta)\hat{k}_z + \hat{k}_z(\sigma + 2\delta) \\
R&: \ (2\pi - 3\delta)\hat{k}_z + \hat{k}_z(2\pi - 3\delta)
\end{aligned}
\tag{6.14}
$$

To obtain the phase terms, a closer look at the bulk equations (6.13) reveals that the sign of P is negative, R contains a component $-\frac{1}{2}\sqrt{3}\sigma k_x^2$ and S a component $-\sqrt{3}i\sigma k_x \hat{k}_z$, so that working from (6.12) this implies,

$$
\left.
\begin{aligned}
Sign(P) &= -1 \\
r_{xx}^{\sigma} &= -1 \\
s_{xz}^{\sigma} &= -i
\end{aligned}
\right\}
\text{ leading to } \Pi = +1 \text{ and } \Theta = -i
\tag{6.15}
$$

Including the C term completes the transformation of the Hamiltonian. The boundary conditions are now obtained by integrating across an interface, resulting in the requirement of continuity of $\boldsymbol{F}$ and $\Omega\boldsymbol{F}$ where $\boldsymbol{F}$ is a 6-component vector containing the envelopes and Ω is given by

$$
\begin{pmatrix}
(\gamma_1 - 2\gamma_3)\frac{\partial}{\partial z} & \frac{2}{\sqrt{3}}k_-(\pi + 3\delta) & \frac{\sqrt{6}}{3}ik_+(2\pi - 3\delta) & 0 & -\frac{2}{\sqrt{6}}k_-(\pi + 3\delta) & -\frac{2}{\sqrt{3}}ik_+(2\pi - 3\delta) \\
-\frac{2}{\sqrt{3}}k_+(3\sigma - 2\pi) & (\gamma_1 + 2\gamma_3)\frac{\partial}{\partial z} & 2k_-(\sigma - \pi - \delta) & \frac{\sqrt{6}}{3}ik_+(2\pi - 3\delta) & -2\sqrt{2}\gamma_3\frac{\partial}{\partial z} & \sqrt{2}k_-(2\sigma - \pi + \delta) \\
-\frac{\sqrt{6}}{3}ik_-(2\pi - 3\delta) & 2k_+(\sigma - \pi - \delta) & (\gamma_1 + 2\gamma_3)\frac{\partial}{\partial z} & \frac{2}{\sqrt{3}}k_-(3\sigma - 2\pi) & \sqrt{2}k_+(2\sigma - \pi + \delta) & -2\sqrt{2}\gamma_3\frac{\partial}{\partial z} \\
0 & \frac{\sqrt{6}}{3}ik_-(2\pi - 3\delta) & -\frac{2}{\sqrt{3}}k_+(\pi + 3\delta) & (\gamma_1 - 2\gamma_3)\frac{\partial}{\partial z} & \frac{2}{\sqrt{3}}ik_-(2\pi - 3\delta) & \frac{2}{\sqrt{6}}k_+(\pi + 3\delta) \\
\frac{2}{\sqrt{6}}k_+(3\sigma - 2\pi) & -2\sqrt{2}\gamma_3\frac{\partial}{\partial z} & -\sqrt{2}k_-(\sigma + 2\delta) & \frac{2}{\sqrt{3}}ik_+(2\pi - 3\delta) & \gamma_1\frac{\partial}{\partial z} & -2k_-(\sigma - \pi - \delta) \\
-\frac{2}{\sqrt{3}}ik_-(2\pi - 3\delta) & -\sqrt{2}k_+(\sigma + 2\delta) & -2\sqrt{2}\gamma_3\frac{\partial}{\partial z} & -\frac{2}{\sqrt{6}}k_-(3\sigma - 2\pi) & -2k_+(\sigma - \pi - \delta) & \gamma_1\frac{\partial}{\partial z}
\end{pmatrix}
\tag{6.16}
$$

As expected, the boundary condition matrix does not feature σ-terms in the first and fourth row, i.e. the heavy hole bands do not couple with the light or split-off hole bands via the remote states of symmetry Γ_1. As a result, along with the additional coupling that originates from the C-term, the boundary condition matrix shows profound differences from that which would have been obtained had one worked from a symmetrised Hamiltonian [Iko92,Men92]. At this point, the theoretical discussion on the boundary conditions ends. This results chapter will now continue with presenting numerical results for both the dispersion and mobilities.

6.2 - Bandstructure calculations

This section presents numerical results obtained with the valence band effective mass model derived in chapter 2. A selection of band dispersion diagrams *E(k)* is given to illustrate the various mechanisms that shape or alter the band structure. Starting with flat band QW structures, dispersion relations are shown for a variety of growth directions, where it is shown that growth on [111] substrates can lead to a reduction in the density of states and thereby significant enhance electrical and optical device properties as discussed in chapter 1. Furthermore, the effects of strain on the dispersion are discussed, as well as the occurrence of spin-splitting in non-symmetric systems. The boundary conditions briefly re-appear when their influence of the dispersion is highlighted.

Unless stated otherwise, all calculations are performed using a 6-band Hamiltonian for the growth direction [001] and a 4-band Hamiltonian for all other ([110] and [111]) growth directions. The new boundary conditions are employed throughout this report, unless clearly indicated. All material parameters used are those presented in section 5.2. The various external parameters that control the effective mass solver are chosen such that they have no influence on the final dispersion. Specifically, a mesh size of sufficient density is used to eliminate any influence of the discretisation procedure, the number of expansion states is set to at least 15 per band to guarantee convergence, and the infinite barriers are placed at least 50Å away from any interface to minimise interference with the confined states. Finally, all self-consistent calculations are performed using an explicit Fermi-Dirac integration.

6.2.1 - A comparision between results obtained using the new and conventional boundary conditions

This section evaluates the effect of the boundary conditions on the bandstructure. Fig.6.2 and 6.3 show the calculated in-plane dispersion for growth on [110] for the technologically relevant material system $In_{0.532}Ga_{0.468}As/InP$. Here, differences between the results obtained using the exact and conventional symmetrised boundary conditions are dramatic even at fairly low in-plane momentum. As expected, the effect is more pronounced in smaller wells as the interface obviously stronger affects the eigenvalues in the case of enhanced penetration into the barrier regions. Fig.6.4 and 6.5 show similar diagrams for growth on [001], respectively using the InGaAs-InP material system and $Al_{0.33}Ga_{0.67}As/GaAs$ system. Unsurprisingly, the boundary conditions are shown to

have a far larger influence on the dispersion in a material system with a relatively large mismatch in terms of effective mass parameters such as InGaAs/InP as compared to GaAs-based structures. Furthermore, the growth direction [110] is shown to be more sensitive to the choice of boundary conditions as compared to [001].

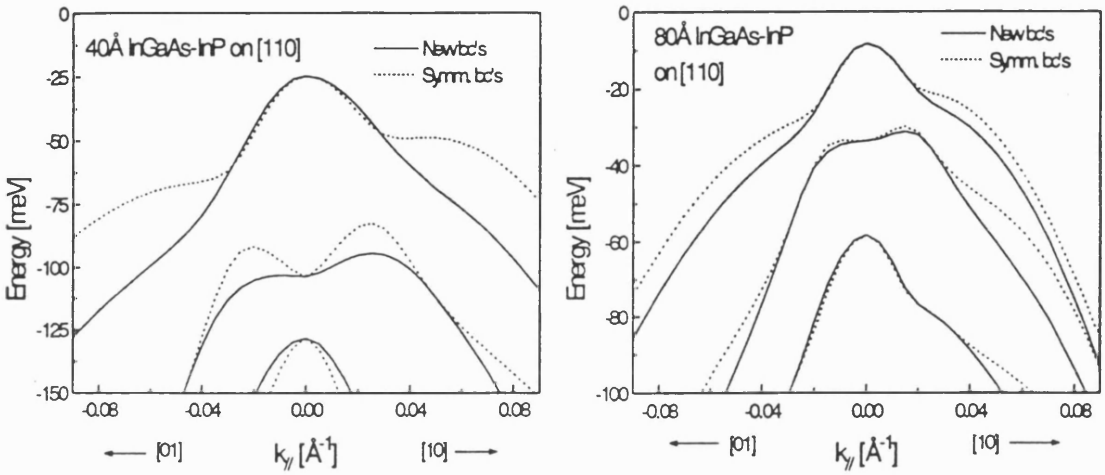


Fig.6.2 and 6.3: Calculated in-plane valence band dispersion for a 40Å and 80Å wide $In_{0.532}Ga_{0.468}As$/InP QW grown in the [110] direction. The solid line shows the result obtained using the Burt-Foreman boundary conditions, the dotted line using the conventional symmetrised boundary conditions (the result published in [Sta97] differs slightly from that presented here as it uses a different band offset).

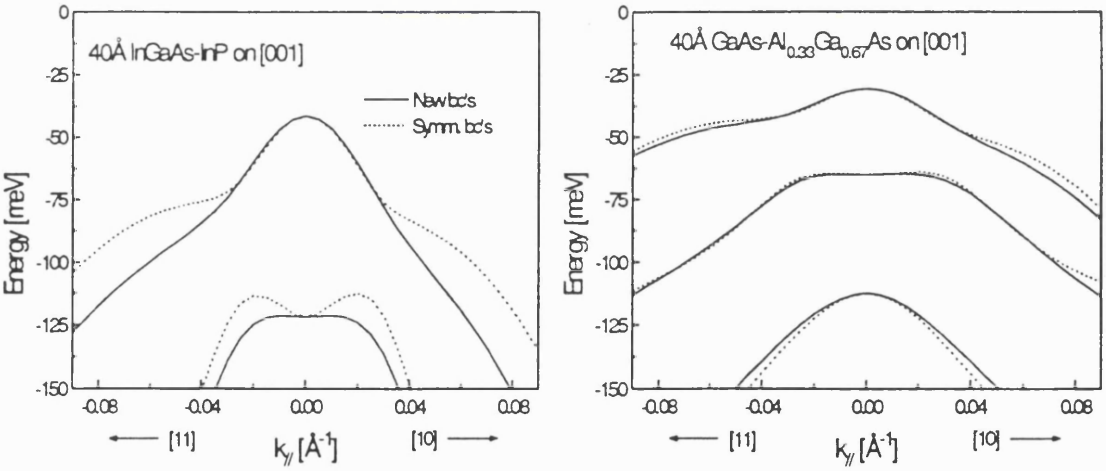


Fig.6.4 and 6.5 : As the previous figures, but now respectively for a 40Å $In_{0.532}Ga_{0.468}As$-InP and $Al_{0.33}Ga_{0.67}As$-GaAs QW grown along [001].

It follows from the above results that it is essential to use the exact boundary conditions when modelling material systems with a relatively large mismatch in material constants, especially for smaller well widths and when looking at the growth direction [110]. Fig.6.6 illustrates the effect in [001] InGaAs containing 15% Indium, as this material composition coincides with the experimental structures that are going to be evaluated later on in this chapter. The well width used in this diagram is smaller than that of the experimental structures to allow comparison between the

above material systems. Although only small deviations are observed in this dispersion diagram, the influence on the effective mass can still considerable as presented in fig.6.7.

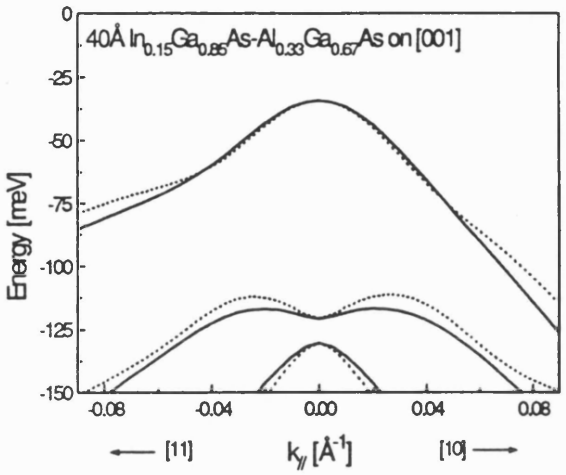


Fig.6.6: As above, now for an InGaAs QW containing 15% Indium.

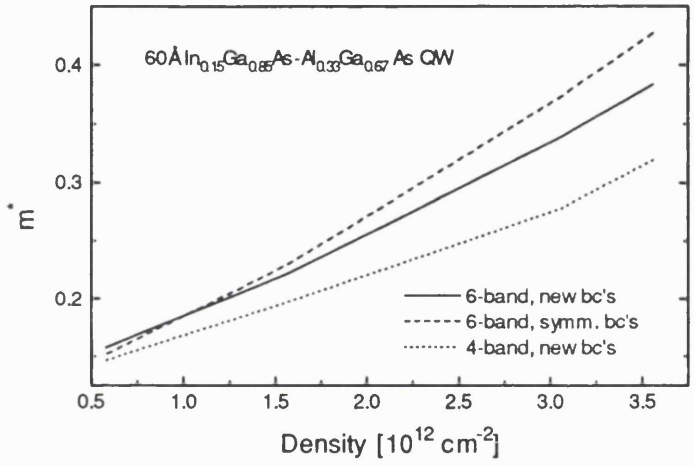


Fig.6.7 : Effective mass at the Fermi energy in a set of 60Å wide centre δ-doped $In_{0.15}Ga_{0.85}As$-$Al_{0.33}Ga_{0.67}As$ [001] QWs (described in detail in §6.4.1). Shown are the theoretical results using a 6-band model for the new and symmetrised boundary conditions, as well as using a 4-band model.

Failure to use the correct boundary conditions can lead to a 10% error in the effective mass at the Fermi level at densities above $3 \cdot 10^{12}$ cm^{-2} thus leading to errors in the mobility, assuming that $\mu \propto (m^*)^{-2}$, of about 20%. Such a deviation is half the error induced when excluding the spin-orbit bands in the effective mass model.

Summarising, the boundary conditions are not of critical importance when looking at $In_xGa_{1-x}As$ on GaAs as the difference in the dispersion between that obtained using the conventional and new boundary conditions is only of the order of a few meV at the relatively small well width of 40Å, and this difference is expected to be even smaller in wider QW structures. The boundary conditions are expected to have even less influence in the ternary $Al_xGa_{1-x}As$. However, the boundary condition problem is of vital importance for the technologically relevant InGaAs-InP compound or emerging systems such as CdTe.

6.2.2 - Influence of the growth direction on the dispersion

Currently, heterostructures grown along non-conventional growth directions have been attracting considerable interest as it has been shown that changing the orientation can significantly affect the resulting bandstructure [Iko92,Men92]. This is illustrated in fig.6.8 in which bandstructure calculations are presented for the growth directions [001], [110] and [111]. It is shown that a change in growth direction affects the dispersion dramatically. It can even result in a change in the order of the heavy and light hole bands, as illustrated by the change of the second quantised state from a heavy hole state in a [001] structure to a light hole state in [110] and [111] grown structures.

As outlined in §1.2.4 when discussing current areas of research and practical use of bandstructure engineering for devices, it follows directly from the contour plots that [111] exhibits a reduced density of states which make it attractive for use in low-threshold lasers. Similarly, the anisotropic contour plot of [110] in which clear differences between the [01] and [10] direction can be seen, suggests that [110] possesses anisotropic optical properties that can be employed to produce Vertical Cavity Semiconductor Lasers (VCSELs) that feature stable and controllable polarisation properties. When looking at transport properties, the reduced density of states in [111] and corresponding lower effective mass is again evidence of a possible superior performance when compared to [001] structures.

Although non-[001] growth directions will briefly re-appear in the next sub-section when presenting the effects of strain on the dispersion, especially when discussing the occurrence of internal piezo-electric fields, this thesis will from now on mainly concern itself with growth along [001] as all experimental work that will be discussed later on is based on [001] structures. The conclusions and recommendations section will briefly outline some suggestions for further work on non-[001].

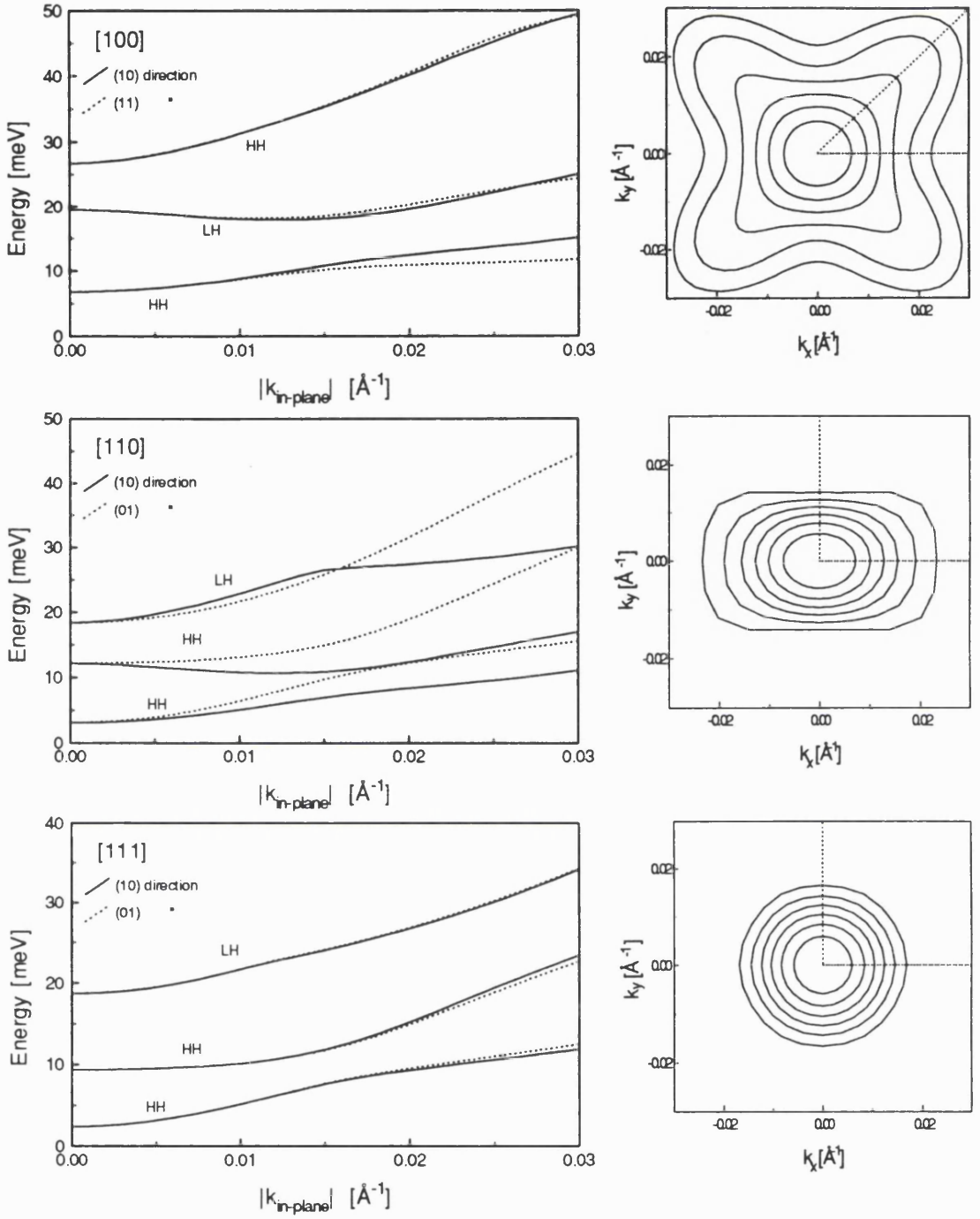


Fig.6.8 : Calculated bandstructure, using a 4 band model, for a 120 Å GaAs QW with $Al_{0.34}Ga_{0.66}As$ barriers grown on the substrate orientations [001], [110] and [111]. The right-hand side plots show the corresponding energy contours in k-space of the hh_1 level. Energy difference between adjacent lines is 1 meV. The well thickness and composition were chosen to be able to compare the results with calculations performed in [Iko92].

6.2.3 - The influence of strain

It was shown in §2.4.2 that strain affects the bandstructure via three effects: it widens (reduces) the band gap for compressively (tensile) strained structures, splits up the heavy and light hole band and introduces extra heavy⇔light hole coupling in [110] structures. In this section, the effect of strain on the dispersion is studied, especially with respect to its influence via the induced splitting of the heavy and light hole band edge. As outlined in chapter 1, the interaction between the heavy and light hole bands causes the valence bands to be highly non-parabolic and non-isotropic. As a result, moving the light holes away from the lowest heavy hole states is expected to reduce the effective mass and enhance the electrical and optical properties.

Fig.6.9 shows the quantisation energy, i.e. at zero in-plane momentum, of the lowest quantised states in the InGaAs-AlGaAs system (N.B. hole energies are taken as positive). Increasing the indium composition in the QW results in a linear shift of the light hole levels away from the lowest heavy hole level at low Indium composition. Such was expected from the Bir-Pikus Hamiltonian (2.76,2.77). The quantisation energy of the LH levels however no longer increases with higher Indium composition as the spin-orbit bands (located at 380 meV above the heavy hole band edge in InAs (see table 5.1 and fig.5.16) starts to interact and effectively starts to push down the light hole levels. However, the split up of the heavy and light hole level is fairly monotonic up to 20% Indium. Similar results are obtained for the InGaAs-InP system (fig.6.10).

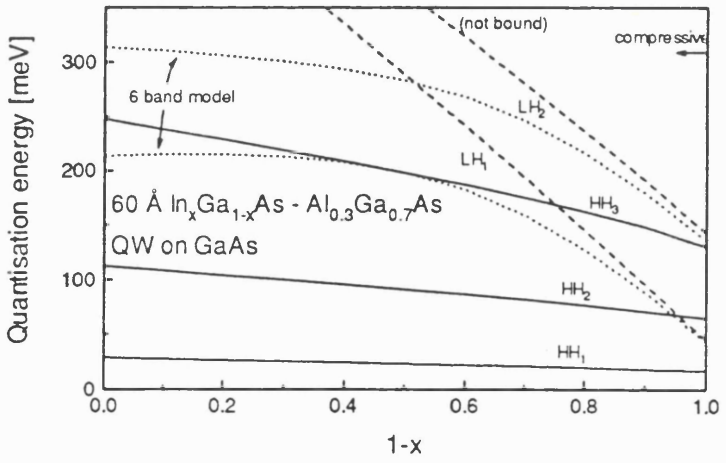


Fig.6.9 : Quantisation energy wrt. hh-band edge vs. Indium composition for a 60Å $In_xGa_{1-x}As$-$Al_{0.3}Ga_{0.7}As$ QW grown on a [001] GaAs substrate. The dashed lines show the results for the light holes when neglecting the spin-orbit interaction (4-band), the dotted lines when explicitly including it in a 6-band model. Note that the heavy hole quantisation energy is not affected by the spin-orbit bands.

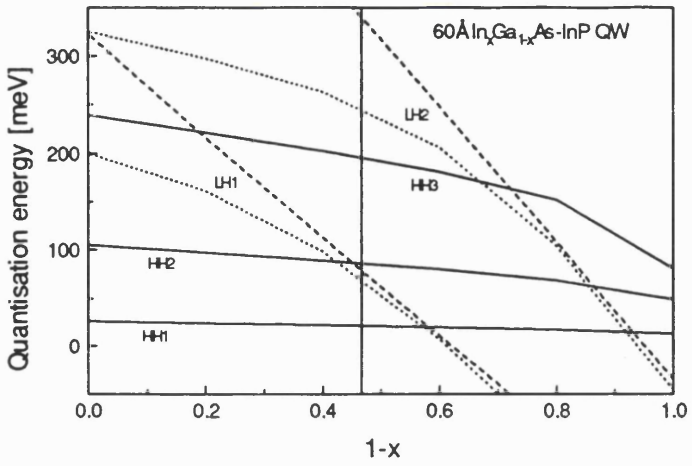


Fig.6.10 : Same as fig.6.9, now for a 60Å $In_xGa_{1-x}As$-InP QW grown along [001]. Result are identical to that obtained by [Cha92] (note that this paper does not use the hh-band edge as the zero-point of origin such as done here).

The next subject to investigate is the effect of strain on the dispersion. Fig.6.12 shows the valence band dispersion for the lowest three quantised states in a series of InGaAs-AlGaAs QWs with varying indium composition. Again, it is shown that the light hole level lh_1 is pushed away from the lowest heavy hole level (hh_1) with increasing strain, passing the second heavy hole level at an indium composition between 5 and 10%. Concentrating on the hh_1 level, it is clear that the curvature of this level increases (thus the mass decreases) with increasing indium composition. This effect on the Fermi-level effective mass is highlighted in fig.6.11. It is obvious that the strain can be utilised to lower the mass and thereby significantly enhance the device properties. It is therefore no surprise that the highest reported *p*-type mobilities in the literature were achieved in structures that contain a high indium composition (see fig.1.12).

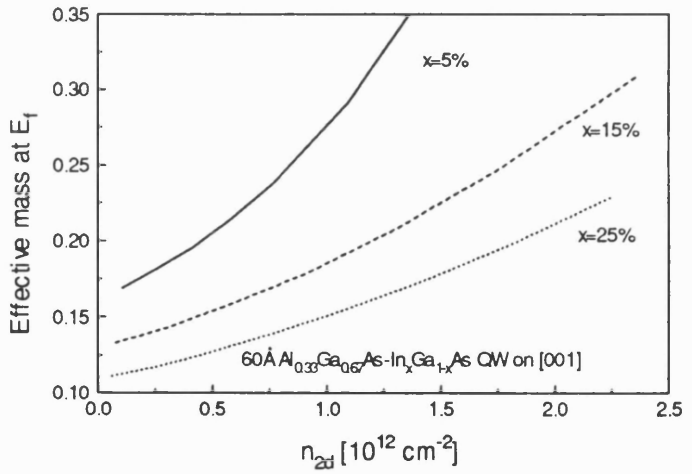


Fig.6.11 : Effective mass at the Fermi-level as a function of carrier density in a flat band 60Å $In_xGa_{1-x}As$-$Al_{0.33}Ga_{0.67}As$ QW on [001]. Results are shown for 5,15 and 25% In in the channel.

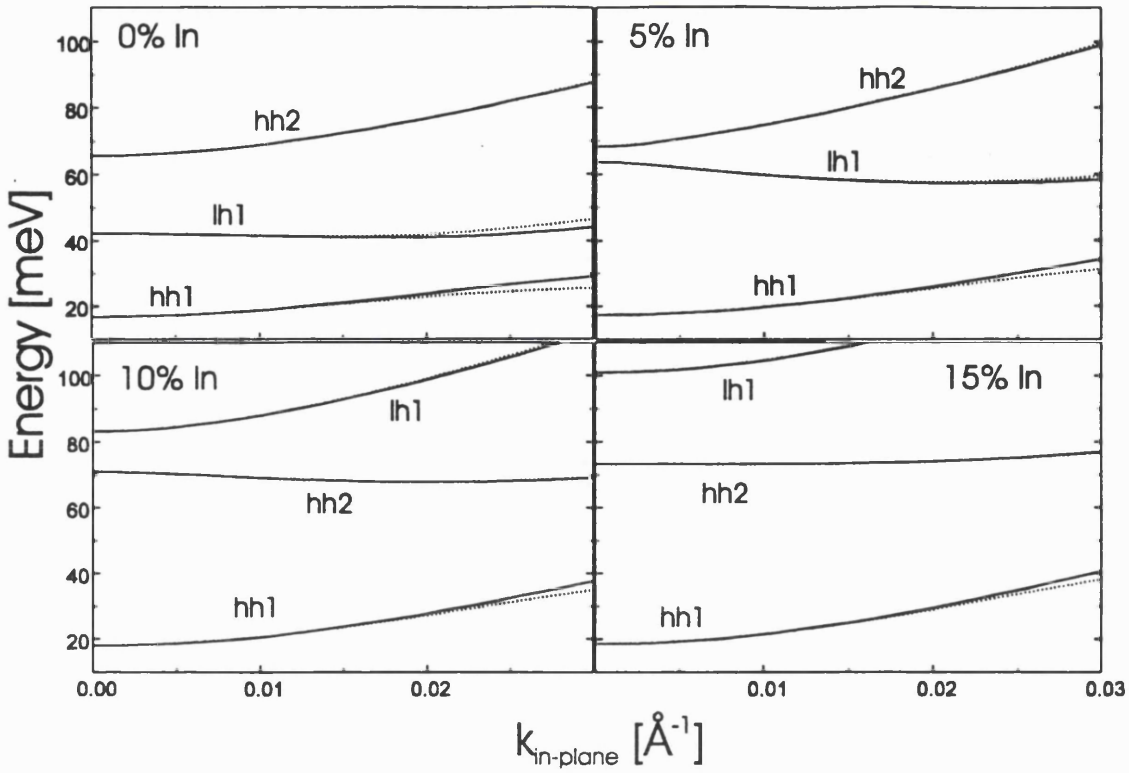


Fig.6.12: Valence band E(k) diagrams for pseudomorphic 60Å $In_xGa_{1-x}As$-$Al_{0.33}Ga_{0.67}As$ QWs on [001], respectively containing 0,5,10 and 15% Indium in the channel (solid lines: along [10], dotted:along [11]). As the In% (and therefore strain) increases, the light hole bands (of which lh_1 is only shown) are pushed away from the hh_1 band, which reduces the hh/lh coupling and results in a lower hh_1-mass.

So far, only strain effects in [001] structures have been dealt with. As outlined in §2.4.3, strain can induce internal piezo-electric fields whose effect on the confinement potential and envelope functions will now be illustrated. The inset in fig.6.13 shows the envelope functions of the lowest quantised states in a 60Å $In_{0.15}Ga_{0.85}As$-$Al_{0.33}Ga_{0.67}As$ QW grown on a [111]B substrate (located to the right of the QW). As a result of the piezo-electric field, the QW is skewed and the hh_1 state takes up an asymmetric position with respect to the centre of the QW. The use of such internal fields to optimise device properties is illustrated in the main figure, which pictures the band diagram of a similar QW but which now includes a δ-doped plane (N_{ac}=2·10^{12} cm^{-2}). As shown, the field effectively pulls the envelope functions away from the impurities, resulting in an expected increase in carrier mobility in structures that are limited by ionised impurity scattering. At this point, it has to be remarked that the use of piezo-electric fields in [Hit95] is fundamentally different from that depicted here, as piezo-electric fields were utilised in that work to reduce interface scattering rather than ionised impurity scattering. The main object was therefore not to achieve maximum separation between carriers and impurities, but to achieve a symmetric confinement potential to avoid the envelope functions from piling up at an interface.

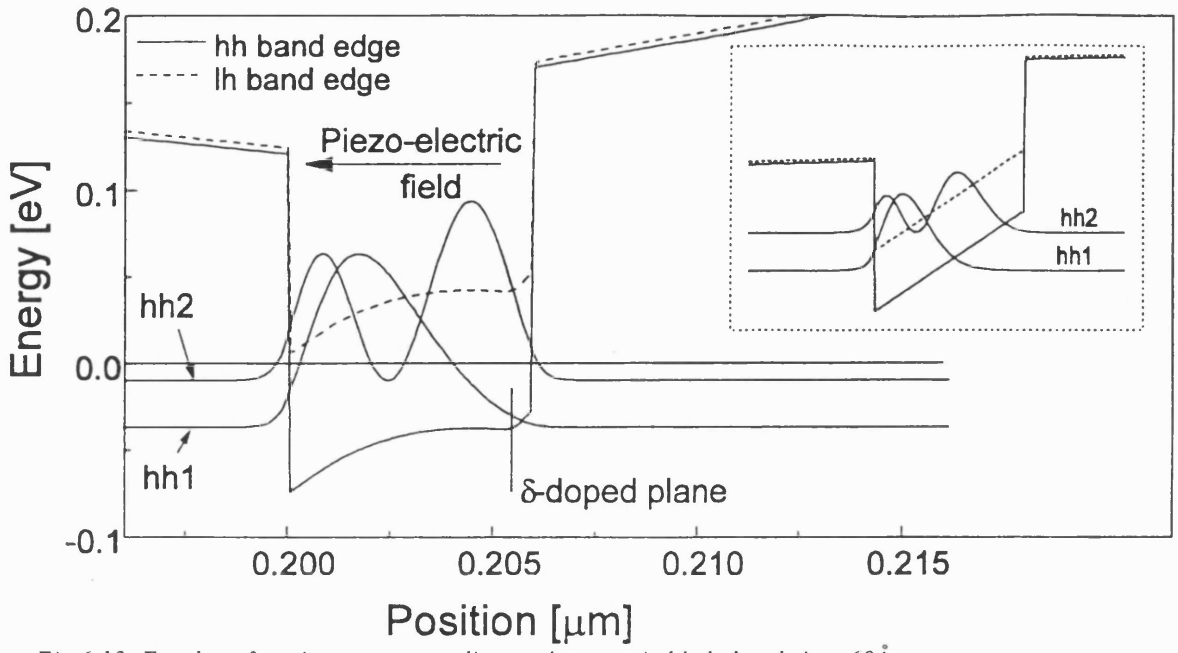


Fig.6.13: Envelope functions corresponding to the occupied hole bands in a 60Å $In_{0.15}Ga_{0.85}As$-$Al_{0.33}Ga_{0.67}As$ QW on [111]B, containing a δ-doped plane (N_{ac}=2.10^{12} cm^{-2}) in the channel located at 5Å from the bottom interface (hh_1 and hh_2 respectively contain 1.22 and 0.33·10^{12} holes cm^{-2}). The presence of the piezo-electric field results in a shift in the lowest quantised state away from the impurities. The inset shows the band diagram without the presence of any impurities.

6.2.4 - Spin splitting in non-symmetric structures

Section 6.2 now finishes with a brief discussion of the effects of spin-splitting on the dispersion relations and hole mobilities. All *E(k)* bands presented so far in this thesis were double degenerate and contained both a spin-up and -down component. As outlined in [And85,Bas88], the use of non-symmetric structures results in spin-splitting which lifts this spin-degeneracy. The former showed the presence of significant spin-splitting effects in AlGaAs-GaAs heterojunctions, both on the dispersion and the effective masses. An investigation into the magnitude of the spin-splitting in graded QW structures is required to analyse the scope of wavefunction engineering in *p*-type structures through the use of graded QW channels.

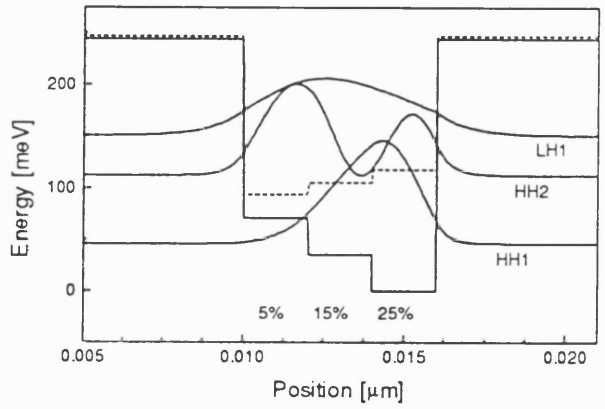


Fig.6.14 : Calculated band diagram for a flat band graded 60Å $In_xGa_{1-x}As$-$Al_{0.33}Ga_{0.67}As$ QW on [001]. The channel contains three regions of each 20Å wide, respectively containing 5,15 and 25% Indium.

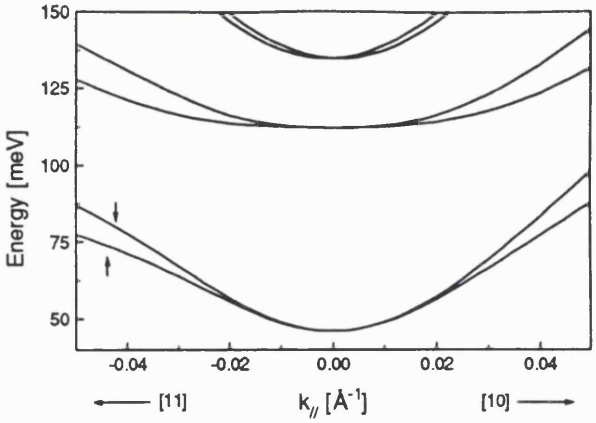


Fig.6.15: Dispersion relations for the graded structure of fig.6.14. A considerable spin-splitting is observed in the lowest subband (hh_1) at in-plane momentum above 0.03Å^{-1}, which would correspond to a hole density of about 1.4·10^{12} cm^{-2}.

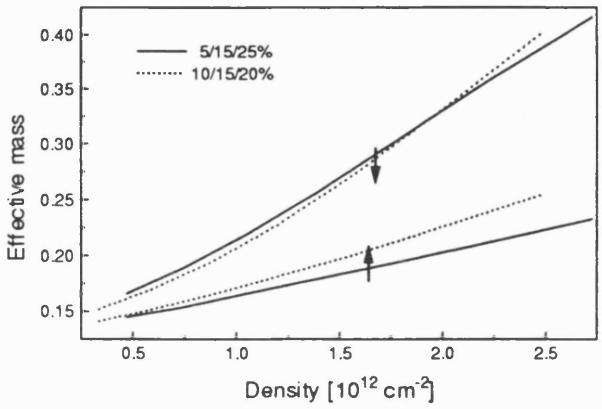


Fig.6.16 : Effective mass in the presence of spin splitting for two graded flat band structures, respectively that of fig.6.14 and a similar structure of which the well consists of 3x20Å $In_xGa_{1-x}As$ with x=10/15/20%.

Although the presence of spin-splitting results in a lower (i.e. desirable) mass for the band labelled spin-down in fig.6.15, the spin-up band increases considerably in mass. This is illustrated in fig.6.16 where results for both masses are shown for two graded structures. Spin-splitting is therefore expected to negatively affect the transport properties as the band with the heavier mass would accommodate the majority of the carriers. At this point, is has to be noted that fig.6.15 and 6.16 correspond to the flat band case and can therefore only be used as a guideline to evaluate the effect of spin-splitting on the effective masses in a doped structure. The presence of charge will obviously affect the dispersion and effective masses through the confinement potential, so that the exact effect of spin-splitting has to be evaluated for each doping density and distribution individually. However, the first indication is that the use of non-symmetric structures (either through grading or internal electric- or piezo-electric fields) has an adverse effect on the transport properties of *p*-type structures.

6.3 - Electron mobility calculations

The previous section showed some numerical results obtained with the valence band effective mass model. The ultimate aim of this thesis is to use that model to simulate and optimise the electrical properties of *p*-type heterostructures by combining the effective mass model with the RPA scattering model of chapter 4. As no results are available in the literature to test the latter combination for *p*-type structures (the calculations presented later on in this chapter are thought to be the first of their kind), this thesis will now briefly discuss the results of the model when applied to *n*-type systems. The reason for this is threefold

- Electron mobility calculations using the formalisms of this thesis have been presented in the literature, offering a possibility to check the model employed in this work.
- Earlier work performed at the IRC revealed significant differences between experimental and theoretical carrier mobilities in a set of *n*-type structures. It is hoped that the more advanced model of this work can reveal the reason for this discrepancy
- Electron mobility calculations are relatively straightforward to interpret and can be used to gain elementary insight in the intra- and inter-band scattering processes. This knowledge was subsequently used to design a set a *p*-type structures.

These three issues will be treated in the same order as presented above. Starting with a comparison between the theoretical calculations performed by Hai et al. [Hai95,Hai96,Hai97] and those obtained in this work, the emphasis then shifts towards interpretation of experimental data. Experimental results obtained by Roberts et al. [Rob96,Rob98] are compared to theoretical calculations, where the relevance of interband scattering is highlighted for wavefunction engineered structures.

6.3.1 - Comparison with results published in the literature

The papers by Hai et al. [Hai95,Hai96,Hai97] provide detailed information about the RPA formalism and give extensive numerical results for the ionised impurity limited mobility in delta doped structures. This sub-section presents a number of calculations of electron mobilities that are directly comparable to the results published by Hai. Some small differences between both sets of results can be found, which are attributed to differences in the implementation of non-parabolicity effects. This work only considers non-parabolicity effects for the in-plane direction, in contrast to the approach taken by Hai who also includes non-parabolicity effects on the quantisation energies

via [Eke87]. A Fermi-energy of respectively 25.53 and 4.11 meV was quoted in [Hai96] for the two occupied subbands in δ-doped GaAs with a background acceptor concentration of 10^{14} cm^{-3} and total electron density of 10^{12} cm^{-2}. Similar results could be obtained by manually setting the in-plane and perpendicular electron effective mass to $0.08m_0$, thereby excluding any non-parabolicity effects ($E_{fnd}\rightarrow\infty$), giving Fermi energies of 25.65 and 3.84 meV. This choice of mass parameters was then subsequently used in all electron calculations.

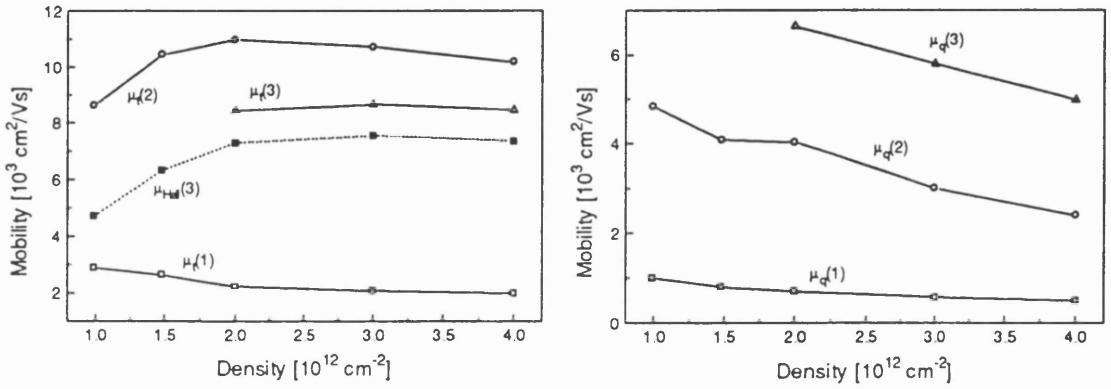


Fig.6.17 and 6.18: Theoretical transport (left) and quantum (right) subband mobilities in δ-doped GaAs as a function of electron density. A background acceptor concentration of 10^{14} cm^{-3} was assumed.

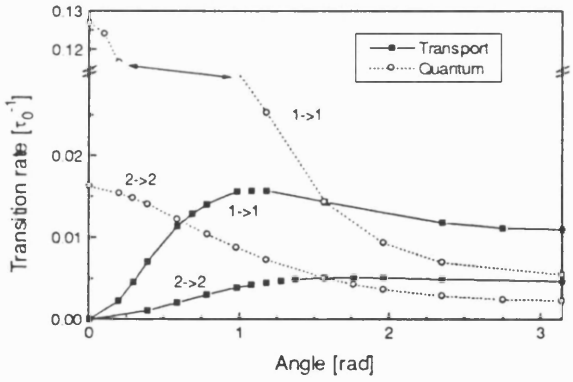


Fig.6.19: Scattering time versus angle for the interband transitions 1→1 and 2→2 for previously used δ-doped structure at an electron density of 10^{12} cm^{-2} (τ_0=0.955 10^{-14} s).

Fig.6.20: Subband transport mobility as a function of well width for a centre δ-doped $Al_{0.3}Ga_{0.7}As$-GaAs QW with N_d=5 10^{12} cm^{-2} and a background acceptor concentration of 10^{14} cm^{-3}.

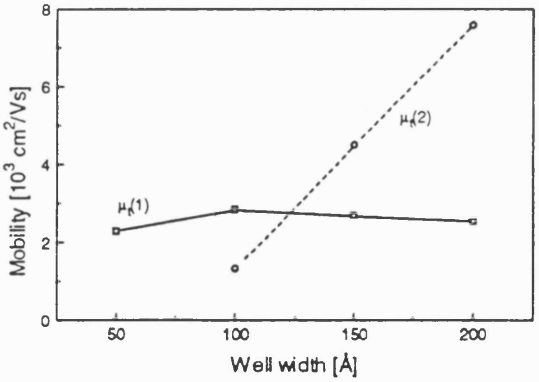


Fig.6.17 to 6.20 show some results of the application of the RPA model of this work to *n*-type systems. All results presented here agree closely (within 10%) to those obtained by Hai et al. From this, the conclusion can be made that the implementation of the scattering model is correct. The scattering model will now be employed to analyse the experimental Hall mobility of a set of *n*-type structures in which wavefunction engineering techniques were used to enhance their transport properties.

6.3.2 - Comparison with experimental data.

In his thesis, J.Roberts [Rob96,Rob98] presents experimental carrier (Hall) mobilities and densities for three types of structures. These structures consist of :

I. centre δ-doped GaAs QWs with well widths of 100, 50 and 25Å (Layer structure: 100Å GaAs cap layer, 2000Å $Al_{0.33}Ga_{0.67}As$, GaAs QW, 2000Å $Al_{0.33}Ga_{0.67}As$, 1000Å GaAs, [001] substrate)

II. 100Å doped GaAs QWs with a various doping schemes; uniformly doped (UD), centre δ-doped (Cδ) and edge δ-doped (Eδ). Layer structure similar to (I)

III. Edge δ-doped 100Å $In_xGa_{1-x}As$-$Al_{0.48}In_{0.52}As$-InP QW structures with a various channel compositions. Again three structures were investigated, having channels which were (a) uniform (x=0.53), (b) 3-step graded (3x33Å of composition x=0.25/0.53/0.80) and (c) 5-step graded (5x20Å of composition x=0.20/0.33/0.53/0.67/0.80), between 2000Å AlInAs barriers with a 100Å InAs cap layer.

Tables 6.1 to 6.3 show the experimental and theoretical results obtained for the samples described above. All parameters required in the self-consistent and mobility calculations were presented in chapter 5. The effective mass m^* given in all tables corresponds to the effective mass at the Fermi-level using the non-parabolicity expression and fundamental band gaps presented in section 4.8. Although the growth conditions were set up to keep the donor concentration N_D constant within each set, the experimental carrier concentration was generally found to vary considerably within such a set. This was attributed to differences in donor incorporation. The impurity concentration used in the theoretical models was therefore adjusted to agree with the experimental result. The dielectric constant ε_r was set to that of the channel layer.

Well width Å	Parameters		Theoretical					Experimental		μ_t / μ_H
	m^*	N_D	n_1	μ_1	n_2	μ_2	μ_t	n_H	μ_H	
100	0.084	5.12	4.12	2358	0.50	1081	2291	4.5	1220	1.88
50	0.084	4.64	4.10	1779			1779	4.1	990	1.80
25	0.078	3.12	2.56	1489			1489	2.6	400	3.72

Table 6.1: Theoretical and experimental carrier densities and mobilities for the set of centre δ-doped GaAs QWs as described in detail in (I). The columns $n_{1,2}$ and $\mu_{1,2}$ respectively contain the theoretical carrier density and mobility of the first and second subband. The column μ_t contains the theoretical value for the mobility that one would measure in a Hall experiment at carrier density n_H. The last columns states the ratio of the theoretical prediction and the experimentally observed (Hall) mobility.

Doping dist'n	Parameters		Theoretical					Experimental		μ_t / μ_H
	m^*	N_D	n_1	μ_1	n_2	μ_2	μ_t	n_H	μ_H	
UD	0.081	4.06	3.14	2528	0.32	869	2472	3.3	970	2.54
Cδ	0.080	3.60	3.02	2338			2338	3.0	1030	2.27
Eδ	0.079	3.44	2.84	3605			3605	2.8	1100	3.28

Table 6.2: Similar to above, now for the set (II) of channel doped GaAs QWs with varying doping distribution (UD=uniformly doped, Cδ=centre doped, Eδ=edge doped).

Well comp'n	Parameters		Theoretical					Experimental		μ_t / μ_H
	m^*	N_D	n_1	μ_1	n_2	μ_2	μ_t	n_H	μ_H	
uniform	0.072	7.56	5.60	2685	1.55	3600	2933	7.0	1300	2.25
3 step	0.065	5.92	4.00	3357	1.53	2910	3246	5.5	1590	2.04
5 step	0.062	4.52	3.33	4387	0.81	2198	4149	4.0	2630	1.58

Table 6.3: Similar to table 6.1, now for a set of InGaAs-AlInAs on InP QWs with varying well composition, specified under (III) (uniform=53% In, 3 step=graded QW consisting of 3 layers with increasing In%, 5 step=similar to 3 step)

The first observation is that the theoretical calculations tend to over-estimate the Hall mobility by a factor of approximately two. The slightly higher deviations observed in the 25Å centre δ-doped QW of (I) is attributed to the effect of interface roughness scattering. The deviation in the edge-doped structure of (II) will be discussed below. The slightly lower deviation in the 5 step InP structure (III) is attributed to the fact that the form of the envelope functions in this structure, and therefore the mobility, was found to be very sensitive to the deformation potential *a* of which there is considerable uncertainty (see chapter 5).

The fact that the discrepancy between theory and experiment is fairly constant, even when comparing different samples structures or even different material systems, is a very important observation. As a direct result of this, the calculations do not just provide the correct trend but can even provide fairly accurate absolute predictions by simply rescaling the theoretical results by a

factor 2. This scaling factor will be used in the theoretical electron mobility calculations below. Its origin will be addressed later in this work when comparing the theoretical and experimental mobilities of a set of *p*-type structures.

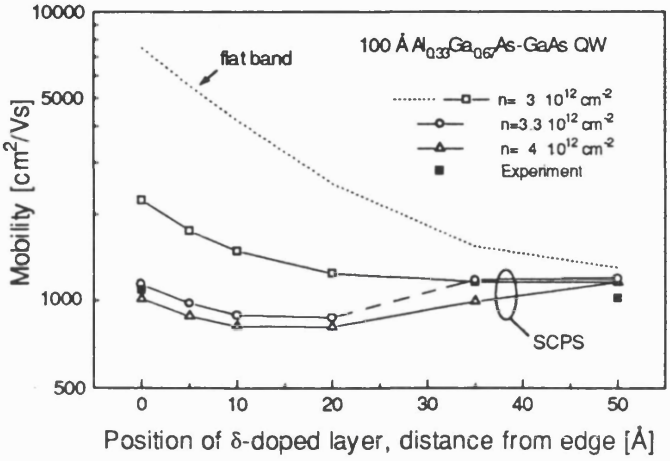


Fig.6.21: Calculated Hall mobility (scaled by a factor 2) for a 100Å δ-doped $Al_{0.33}Ga_{0.67}As$-GaAs QW at various carrier densities as a function of doping position. This structure corresponds to that of sample set (II). The dotted line depicts the results for the flat band approximation, whereas all other calculation correspond to full SCPS solutions. The dashed section in the results for n=3.3 10^{12} cm^{-2} represents a transition from a single (right) to two (left) occupied subbands.

Fig.6.21 is very illustrative when addressing the scope and objects of wavefunction engineering to enhance carrier mobility. Here, the theoretical Hall mobility is plotted as a function of doping position for the channel δ-doped QWs of set (II). One intuitively expects the mobility to increase when moving the impurity plane away from the centre of the well, where it has maximum overlap with the envelope function of the first subband, towards the interface. The latter is clearly observed for the flat band case (dotted line) for which an improvement in mobility of a factor 5.7 is predicted. However, this is clearly too simple an approximation as the electrons (or envelope functions) are Coulomb attracted to the impurities so that their distribution is expected to follow the impurities rather than remain symmetric with respect to the centre of the QW as predicted in the flat band approximation. Nevertheless, these flat band results suggest that a significant gain could be achieved by using a combination of δ-doping and wavefunction engineering techniques, where one uses a graded QW to compensate for the Coulomb interaction. The latter was exactly the underlying reason for the application of a graded channel in the InP structures of set (III).

Returning to fig.6.21, theoretical predictions for the mobility using full self-consistent calculations are also presented in this graph. A smaller but still considerable improvement in the ionised impurity limited mobility of a factor 1.9 is calculated when moving the impurity plane from the centre to the edge of the channel at an electron density of $3 \cdot 10^{12}$ cm^{-2}. Yet, this improvement is not observed in the experimental data which hardly shows any improvement in mobility when moving the impurities away from the carriers, unlike similar measurements by [Mas91]. The explanation

follows from the calculations at a slightly higher electron density of $3.3 \cdot 10^{12}$ cm^{-2}. Here, it is found that the second subband becomes populated as the impurity plane moves within 20Å of the interface, resulting in the occurrence of interband scattering. The calculated mobility at this carrier density agrees very well with the experimental data. Population of the second subband was not taken into account in table 6.2 and explains the poor agreement between the experimental and theoretical values for the edge doped structure. Although the original calculation indicated that the second subband was located at 7.6 meV above the Fermi level, the small shift necessary (about 8% of the spacing between the first and second subband) to bring this subband below the Fermi-level can easily be explained by the omission of non-parabolicity effects on the perpendicular (confinement) mass. Note that interband scattering from e_1 to e_2 (and vice versa) is prohibited in a symmetric single δ-doped system such as the centre-δ-doped structure of (II). The latter explains why there is little difference in mobility between the centre-δ-doped case when comparing between the densities $3 \cdot 10^{12}$ cm^{-2} (1 band populated) and $4 \cdot 10^{12}$ cm^{-2} (2 bands populated) when there is considerable difference between the mobility of the corresponding edge-δ-doped structures.

Fig.6.22 presents the results for the $In_xGa_{1-x}As$-$Al_{0.48}In_{0.52}As$-InP QW structures (III) and shows carrier mobility versus carrier density for the various channel compositions: uniform, 3-step and 5-step. As noted, the differences between the theoretical and experimental value of the 5-step structure is thought to be caused by the uncertainty in the deformation potential *a*. The important observation to make is that the theoretical model does not predict significant differences between the 3- and 5-step structure. The observed improvement in mobility is thus thought to be caused by the accompanying reduction in impurity concentration and occupancy of the second subband, not by any improvements in wavefunction engineering.

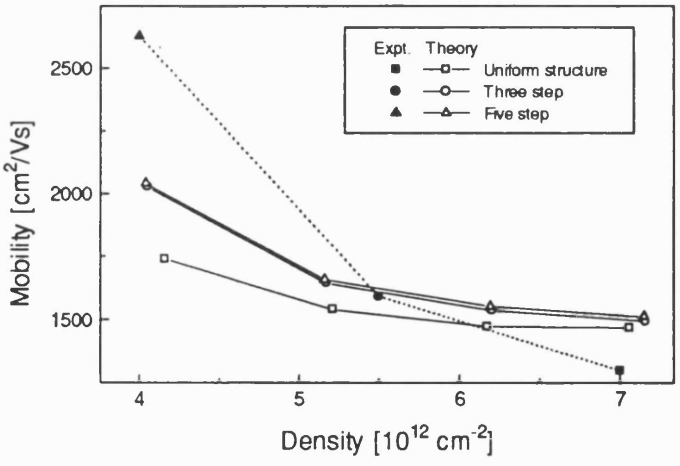


Fig.6.22: Theoretical and experimental Hall mobility vs. carrier density for the set of InP samples (III). Results are presented for the structure with a uniform channel as well for the graded structures. (theoretical results are again scaled by a factor 2)

The failure of the single band Thobel model [Tho93] for the InP-based structures is obvious, as Roberts et al. presented theoretical predictions based on this model, that were a factor of at least 15 higher than experimentally observed, in sharp contrast to the values presented here which are out by just a factor 2. The theoretical investigation of the earlier work failed to include

1) Interband scattering processes
2) Reduction in Fermi-velocity due the reduction in Fermi momentum through the occupation of a second subband
3) Non-parabolicity effects

Intraband scattering	$\tau_{1\to1}{}^{T}$	$= 3.6\ 10^{-13}$ s
	$\tau_{2\to2}{}^{T}$	$= 1.0\ 10^{-13}$ s
Interband scattering	$\tau_{1\to2}{}^{T}=\tau_{2\to1}{}^{T}$	$= 3.2\ 10^{-13}$ s
	$\tau_{1\to2}{}^{Q}=\tau_{2\to1}{}^{Q}$	$= 1.3\ 10^{-13}$ s

Table 6.4 : Individual scattering times corresponding to the 3-step InP structure of (III). Parameters used in the calculation are identical to those used to generate table 6.3.

The effect of the omission of the interband scattering processes can be illustrated by comparing the individual transition times for the various transitions in the 3-step InP structure (see table 6.4). It is shown that the interband scattering times are comparable to the intraband scattering time of the first subband, and are thus expected to strongly affect the transport mobility of this band. Neglecting the interband scattering would result in the transport mobilities $\mu_1{}^{T}$=9720 cm^2/Vs and $\mu_2{}^{T}$=2668 cm^2/Vs of the two individual subbands, giving a Hall mobility μ_H=9050 cm^2/Vs, which is a factor 2.8 higher than obtained in the proper calculation. The omission of non-parabolicity effects accounts for another factor of between 1.48 and 2.2 difference between the two calculations, assuming $\mu \propto (m^*)^{\gamma}$ with γ=1...2. Both corrections account for the majority of the differences found between the predictions of this work and Roberts' model.

6.4 - Theoretical and experimental hole mobilities

The previous section concerned scattering in *n*-type systems. There, it was shown that the RPA formalism can provide reasonably accurate absolute values of the mobility when introducing a scaling factor of 2. Despite the fact that the presence of such an ad hoc factor is rather unsatisfactory considering the large amount of computational work involved in the band structure and scattering calculations, it is clear that the combination of SCPS solver and scattering model provides a powerful means to investigate and optimise the performance of *n*-type structures. This

section will now extend the latter work to *p*-type structures, as the ultimate aim of this thesis is to investigate the possibility to use wavefunction engineering techniques to optimise the hole transport properties of GaAs-based heterostructures. Obviously, a similar model for hole transport would be an essential tool to investigate the latter.

The ionised impurity scattering model that was presented in chapter 4 will be compared to experimental data obtained for a set of *p*-type structures. The exact composition of structures will be described in §6.4.1 where all the insight obtained in this chapter, either band structure considerations or concerning scattering, is used to design a set of structures to evaluate the scope of wavefunction engineering to optimise of the electrical properties of two-dimensional hole gases. Experimental values for the hole mobility and density are given in §6.4.2 where it is shown that varying the doping distribution can significantly affect the ionised impurity limited mobility as well as change its temperature dependence.

The next sub-section starts with a numerical evaluation of Ehrenreich's overlap function. It is shown that this multiplicative factor can have a large influence on theoretical hole mobility calculations. Results for various QW structures and compositions are presented, which are shown to deviate from the bulk results that can be found in the literature.

Before actually presenting theoretical hole mobilities, the effects of diffusion and segregation on the hole mobility are presented. It is shown that segregation can significantly affect structures with non-symmetric doping distribution. The latter will be taken into account when comparing theoretical and experimental hole mobilities such as is done in §6.4.5.

As mentioned earlier, ionised impurity scattering calculations in two-dimensional *p*-type structures using the RPA framework are thought to be the first of their kind. As in the *n*-type case, it is found that reasonable agreement between theory and experiment can be obtained by introducing a scaling factor, which is unfortunately found to be much larger than for electron mobility calculations. The origin of the scaling factor is discussed in §6.4.7. Although the electron scaling factor can be explained by electron-electron interaction, it is shown that the origin of hole scaling factor is most likely not solely caused by hole-hole interaction but by a combination of the latter and quantisation effects in the screening charge distribution. Accounting for these quantisation effects can be done by using an appropriate effective mass in the dielectric response matrix of the screening model. The first indications are that this would result in a new scaling factor that is now in better agreement with the electron result.

6.4.1 - Design of a set of *p*-type heterostructures

All the techniques presented in the previous section and all the insight gained in impurity scattering can now be used to design a set of *p*-type heterostructures to explore the scope of wavefunction engineering to enhance the hole mobility, as well as to test the scattering model for the application to hole systems. As one specifically wants to investigate ionised impurity scattering, the investigation will be limited to partially or fully doped channel structures only, to ensure that ionised impurity scattering is the dominant scattering mechanism. It was chosen to use δ-doping because of the easy theoretical interpretation (section 4.5) as well as the relatively large head room to affect the scattering while limiting the allowed doping region to just the channel.

As spin-splitting in asymmetric heterostructures is expected to negatively affect the electrical performance (see §6.2.4), the use of graded structures is avoided at the present time. The magnitude of spin-splitting effects was carefully evaluated in the structures containing an asymmetric doping distribution, moving to symmetrical doping distributions once the results of the first set confirmed that the growth of *p*-type structures was under control. The latter specifically concerns the accurate growth of the δ-doped planes, as evidence for segregation would have meant that the growth of inverted structures (i.e. with the impurities located between QW and substrate) would have to be avoided; the resulting presence of impurities in the channel would have made the theoretical investigation very difficult.

It was chosen to use strained InGaAs channels as it was pointed out earlier in this chapter (§6.2.3) that the presence of strain can significantly enhance the transport properties. Since this work focuses on the electrical properties and consequently wants to avoid any growth related problems, it was chosen to use a lower In-composition than that used in current state-of-the-art hole structures such as by Kudo [Kud97] which contain 35% Indium. As high-quality $Al_{0.33}Ga_{0.67}As$ as well as $In_{0.15}Ga_{0.85}As$ was routinely obtained at the IRC, the latter two material compositions were selected to be used in the experimental structures of this work. To minimise the effect of the air-semiconductor and semiconductor-substrate interface (see §5.2.4), relatively wide AlGaAs barrier layers (2000Å) were employed.

The following comment is possibly the most important. It was shown in the previous section that interband scattering can completely compensate for any gain in mobility associated with the reduced overlap between the carrier distribution and the impurities when moving a δ-doped plane through the channel towards the interface. From this, it follows that the occupancy of a second

subband has to avoided at all costs to obtain a usable system to evaluate the effects of the doping distribution on the hole mobility. The latter implies the use of relatively high barriers as well as narrow QWs. As 60Å to 80Å wide InGaAs hole structures are routinely used in high performance *p*-type structures [Hsu95,Kud96,Kud97], it was estimated that interface scattering would not be of significant importance when using a 60Å well width, such as in the samples that will be investigated in this work.

Two sets of *p*-type 60Å $In_{0.15}Ga_{0.85}As$-$Al_{0.33}Ga_{0.67}As$ QWs, all on undoped GaAs [001] substrates, were grown by MBE. The exact composition of both sets of samples, denoted by (A) and (B), is given by :

(A) The first set consists of a 1000 Å buffer layer of GaAs grown at 580°C, followed by 2000Å $Al_{0.33}Ga_{0.67}As$ over which the temperature was gradually lowered to 510°C, a 60Å $In_{0.15}Ga_{0.85}As$ QW containing a Be-delta-doped plane, 2000Å $Al_{0.33}Ga_{0.67}As$ over which the temperature was increased to 580°C, capped by 100Å GaAs. Six samples were grown, delta-doped at 2.0 and $3.5 \cdot 10^{12}$ cm^{-2} at the centre of the QW, at 15Å from the top QW interface, and at 5Å from the top QW interface (see table 6.5). The composition of the structures was checked by high-resolution X-ray diffraction and proved to be very close to that specified, verifying the accurate calibration of the MBE machine.

Sample no.	*δ-Doping conc. [10^{12} cm^{-2}]*	*Doping position*
U11042	2.0	Centre
U11043	3.5	Centre
U11044	3.5	3/4
U11045	2.0	3/4
U11047	2.0	Edge
U11048	3.5	Edge

Table 6.5 : Doping concentration and position of the δ-doped plane for the set (A). Three different positions for the impurity plane were used, all inside the QW channel. Centre: centre of the well at 30Å from both interfaces, 3/4: 15Å from the top interface/ 45Å from the bottom, Edge: 5Å from the top interface/55Å from the bottom.

(B) The second set consists of two $In_xGa_{1-x}As$-$Al_{0.33}Ga_{0.67}As$ samples that are similar to those above apart from the fact that they have four δ-doped planes. The structure was set up to contain a symmetric confinement potential to avoid the lifting of Kramer's degeneracy, as well as to counteract the Coulomb interaction and provide a relatively large separation between carriers and impurities. Two of the δ-doped planed were placed in the well at 5Å from the top and bottom interfaces, the first sample doped at 1.2 (U11080) and the second at $1.7 \cdot 10^{12}$ cm^{-2}

(U11081). The two other δ-doped planes were placed in the barriers at 30Å from the top and bottom interfaces, respectively doped at 0.65 and $0.80 \cdot 10^{12}$ cm^{-2}. These doping densities were chosen to supply 10^{12} carriers cm^{-2} to the QW, the rest of the carriers are expected to be lost to the surface and substrate via surface and substrate pinning. By introducing the latter impurity planes, it was hoped to achieved a greater range of carrier mobilities which would offer enhanced means to verify the scattering model. An effort to slightly tailor the QW such that it would consist of a parabolic confinement potential with an Indium composition of around 12% to around 18% at its centre proved unsuccessful. Careful monitoring of the growth run, as well as X-ray diffraction, revealed little variation of the In% over the QW and an average composition of about 13%, being less than the average that was hoped for (15%).

From this point, the majority of the discussion will specifically concern the previously described structures, unless specifically stated. Experimental hole mobilities will now be presented for those structures.

6.4.2 - Experimental hole mobilities

Hall measurements were performed on the previously described *p*-type 60Å $In_{0.15}Ga_{0.85}As$-$Al_{0.33}Ga_{0.67}As$ QWs using a commercial Hall-apparatus, a Biorad HL5200. Because of the simplicity of the measurement procedure, this work will not expand on the latter. A good description of the Hall measurement technique and experimental set-up can be found in [Jas97].

All measurements were performed in the van der Pauw geometry. Because of the perpendicular cleavage planes in [001], such measurements generally do not require any complex processing to form the appropriate geometry and contacts. Square shaped structures of approximately 3x3 mm^2 were manually cleaved, after which InZn contacts were attached at the four corners using a simple soldering iron under a microscope. The samples where then subsequently annealed for 3 minutes at 400°C in a forming gas environment ($10\%H_2/90\%N_2$). Good quality contacts were obtained for the samples with reasonably high carrier density ($>2 \cdot 10^{12}$ cm^{-2}) for which two-terminal resistances R of typically less than 100kΩ, and as low as 40kΩ (U11081), were measured at room temperature. These values are however not to be interpreted as being twice the contact resistance R_c as they include the resistance of the semiconductor structure itself. The experimental sheet resistivity ρ_s was found to be comparable to the latter two-terminal resistances, and it is thus estimated that the contact resistances were of the order of a few kΩ or less. As a 4-probe measurement technique such as a Hall measurement is essentially immune to the presence of contact resistances, no measurement problems were expected. At low temperatures (T=77K), the

contacts on these high density structures showed a slightly less ideal behaviour. Resistivities of up to 230kΩ between contacts were measured for some samples (U11043) which was now more than three times the experimental sheet resistivity at that temperature, showing evidence of higher contact resistances. Limited asymmetric behaviour when reversing the polarity provided evidence that the contacts showed minor Schottky-effects, yet accurate experimental data (accurate within a few percent) for the Hall density and mobility could be obtained by performing several measurements at various current settings. Sample U11081 again proved to possess the best characteristics, showing perfect ohmic behaviour at 77K with resistances less than 21kΩ at a sheet resistivity of about 9 kΩ/□. The latter proves that InZn can provide good quality contacts for highly doped structures, even at low temperature.

In contrast, it proved virtually impossible to perform measurements with reasonable accuracy (±20%) on the low density structures U11042, U11045 and U11047 when using InZn contacts. The centre-doped structure (U11042) showed the best performance and indicated that the Hall mobility was of the order of 65±5 cm^2/Vs at a Hall density of $3.9 \cdot 10^{12}$ cm^{-2}, corresponding to a sheet resistivity around 240kΩ/□ at room temperature. Of the latter three quantities, only the sheet resistivity could accurately be obtained at lower temperatures, showing a value of 155kΩ/□ around 170K after having decreased monotonously when compared to the room temperature value. The poor electrical characteristics were attributed to poor contact quality, illustrated by the fact that two-terminal resistivities of around 1MΩ at 150K and 2MΩ at 77K were measured for U11042 which were accompanied by non-ohmic behaviour.

The low density structures were therefore processed at UCL and contacted with Au/Zn/Au. Because of the particular van der Pauw geometry with narrow contact channels, the measured resistivities were still high (1.4MΩ at ρ_s=300kΩ/□) but now showed perfect ohmic behaviour. Low temperature measurements were however still difficult. Measurements at 77K regularly showed incorrect signs of the Hall-voltage at the application of a magnetic field, as well as two-terminal resistances of up to 2MΩ while the measurements indicated a reduced internal sheet resistivity of ρ_s=200kΩ/□. It was therefore concluded that the Au/Zn/Au contact resistances were of the order of 0.5MΩ at liquid nitrogen temperature, explaining the poor data quality that was obtained for these low density structures. It should be noted that all the mentioned data that was taken while experiencing such high ohmic resistances should be regarded with care, as the Biorad manual states an up upper limit of 1MΩ between terminals to guarantee reliable measurements.

Despite these problems, it proved possible to measure the low density structure as it was found that faint illumination of the Au/Zn/Au structures for a duration of a few seconds while at 77K greatly reduced the contact resistances of subsequent 77K measurements in a dark environment. This is a common technique to improvement contact resistances, associated with the formation of electron-hole pairs at those (locally damaged) contact regions and a resulting contact-semiconductor barrier lowering due to charge transfer effects. To illustrate, U11042 showed a reduction in two-terminal resistance to just 450kΩ. The question is now whether this is just a contact effect or that the illumination has significantly affected the carrier density. The latter is thought to be not the case, as persistent photo-conductivity effects usually generate just 10^{10}-10^{11} carriers cm^{-2} [Lac88] while the carrier densities involved in these experiments are almost an order of magnitude higher. This is further backed by the observation that the measured sheet resistivity (140 kΩ/□) is in agreement with the trend and absolute value of the earlier mentioned un-illuminated measurements in the InZn-contacted structure around 170K. All 77K measurements of U11042, U11045 and U11047 were subsequently performed in the dark, typically several minutes after faint illumination for a couple of seconds at that temperature. The experimental data corresponding to the samples of both sets (A) and (B) are given in table 6.6.

	Room temperature		T=77K	
Sample no.	n_{2D} $[10^{12}\ cm^{-2}]$	μ $[cm^2/Vs]$	n_{2D} $[10^{12}\ cm^{-2}]$	μ $[cm^2/Vs]$
U11042	0.41±0.1	52±5	0.86±0.1 (†)	64±6 (†)
U11043	2.30±0.1	49±3	2.60±0.1	41±2
U11044	2.46±0.1	50±3	2.75±0.05	47±1
U11045	0.72±0.2	68±1	1.27±0.05 (†)	85±10 (†)
U11047	0.57±0.2	69±4	1.19±0.1 (†)	143±5 (†)
U11048	2.20±0.15	68±4	2.66±0.05	104±1
U11080	2.71±0.1	90±3	3.13±0.05	195±2
U11081	4.0±0.15	85±3	4.35±0.05	162±2

Table 6.6 : Experimental values for the Hall density and mobility. All measurements were performed in the dark. The data marked (†) was obtained after faint illumination of the structure for a few seconds while at 77K.

Fig.6.23 shows a graphic plot of the experimental hole mobilities at room temperatures and 77K for the centre, 3/4 and edge δ-doped structures at $N_{ac}=3.5\cdot10^{12}$ cm^{-2}. The corresponding Hall densities are within 14% of the average of the measured hole density $n_{2D}=2.5\cdot10^{12}$ cm^{-2}. A clear improvement in the mobility is observed at 77K when moving the δ-doped plane from the centre, where its overlap with the hole distribution is strongest, to the edge of the QW. The small

improvement in mobility associated with moving the impurity plane over just 25% of the well width from the centre to 15Å from the interface (denoted with 3/4) is attributed to the Coulomb interaction between carriers and impurities. Unlike the edge-doped case where the doping is positioned next to a potential barrier, the envelope functions are relatively free to follow the impurity plane in the 3/4 structure, resulting in only a marginal improvement.

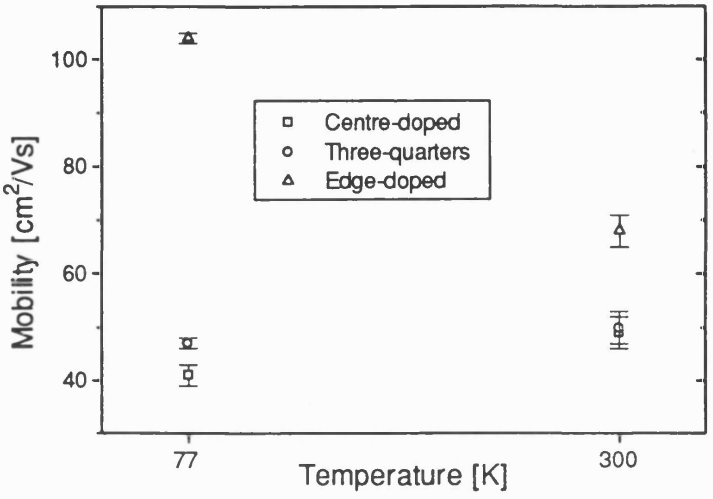


Fig.6.23 : Comparison between the experimental hole mobilities at room temperature and 77K for the high density centre, 3/4 and edge δ-doped structures U11043/44/48.

Similar observations can be made in the 300K data, but the improvement in mobility when moving the impurities towards the edge of the QW is now considerably less than at 77K. This is caused by thermal effects such as phonon scattering and the thermal occupation of higher lying subbands. The presence of these various thermal effects can clearly be demonstrated by comparing the absolute values of the mobility at 77K and 300K. The edge-doped structure features a decrease in mobility when increasing the temperature, due to the inevitable inclusion of phonon scattering and the possibility of interband scattering in such an asymmetric structure at elevated temperatures. In contrast, the centre doped structure actually experiences an increase in mobility, something one would not have expected. The explanation follows from the fact that the second subband of the centre-doped structure is thought to posses a higher mobility as compared to the first subband due to its smaller overlap with the impurities. Thus, thermal excitation actually improves the mobility due to a charge transfer to a subband with superior electrical properties, partially because of the fact that interband scattering between the first and second subband is prohibited in such a symmetric δ-doped structure. The behaviour of the intermediate 3/4-doped structure is found to be inbetween both results. Similar behaviour is found for the low density single δ-doped structures, although higher error margins, relatively large variations in density, as well as the fact that these systems are no longer degenerate at 300K make the interpretation slightly more difficult.

Finally, an upper limit for the phonon scattering times at 300K can be obtained by assuming that the mobility of the edge-doped structure is solely determined by phonon scattering. This value can then be used to estimate the upper limit for the mobility at 77K imposed by this scattering mechanism, by assuming the mobility to be inverse proportional to the phonon density. This density is proportional to $exp(-\hbar\omega_P/k_BT)$, where $\hbar\omega_P \approx 36$meV is the optical phonon energy. Using the lower limit for the phonon limited mobility at 300K of 68 cm^2/Vs, it follows that phonon scattering cannot significantly affect the mobility at 77K as a lower limit of well over 3000 cm^2/Vs is obtained. The 77K experimental data will therefore directly be compared to the theoretical ionised impurity limited hole mobility calculations.

Before presenting such hole mobility calculations, the effect of Ehrenreich's overlap function which is specific to hole transport calculations will be examined. Then, the effects of segregation and diffusion on the theoretical hole mobilities will be presented after which a comparison between the theoretical hole predictions and the above experimental results will be given (§6.4.5).

6.4.3 - Numerical results for Ehrenreich's overlap function

Theoretical values for the electron mobility were presented in section 6.3. As discussed in chapter 4, electron and hole transport calculations are essentially identical apart from the occurrence of an additional angular dependent function $G(\alpha)$ in the hole calculations which describes the effects of scattering from an initial state to a final state with a different cell-periodic composition (section 4.4). As explained, this factor reduces to unity within a one-band effective mass approximation such as for the conduction band. Here, both the initial as final cell-periodic functions are identical and given by a function with the same symmetry properties as an atomic *s*-orbital. Although analytical forms of Ehrenreich's overlap function for the valence bands have been provided for bulk [Wil71], no reference could be found for the two dimensional case. As it will prove that no obvious simple expression exists for the overlap function for the compositionally varying case, approximation schemes that simply involve a rescaling factor such as used by [Lee83b] are difficult to justify for compositionally varying structures. The overlap function will therefore be evaluated numerically in the hole mobility calculations in this work. Because of the complicating factor that the true envelope function composition following from a block-diagonalised 6-band model (§2.3.6) can only be obtained by taking the inverse of the unitary transformation of the envelope functions, it was chosen to perform the evaluation within the 4-band approximation.

Fig.6.24 shows the numerical results for bulk GaAs. To obtain this result, a slab of 1000Å GaAs was evaluated while setting the number of expansion states to 1 as there would be only one heavy and one light hole level at each value of the momentum k, in contrast to the multiple confined states in heterostructures. Good agreement is obtained with the analytical functions in [Wil71].

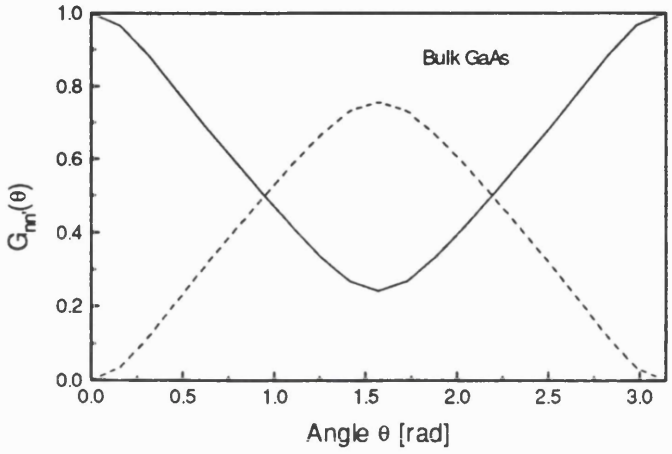


Fig.6.24: Calculated values for Ehrenreich's overlap function $G_{nn'}(\theta)$ for an interband heavy hole (solid) and intraband heavy to light hole transition (dashed) in bulk. The results agree well with the analytical expressions given in [Wil71].

Fig.6.25: Overlap function G for an intraband transition within the first heavy hole band for the two-dimensional case. Results for various QW structures are given; dotted: 100Å GaAs, dashed: 100Å $In_{0.15}Ga_{0.85}As$, solid: 60Å $In_{0.15}Ga_{0.85}As$. All structures have $Al_{0.33}Ga_{0.67}As$ barriers. Results are presented for various values of the in-plane momentum (calculated using a 4-band model)

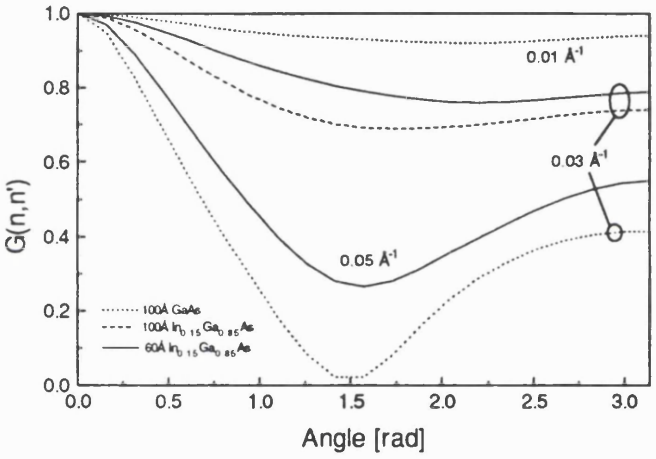


Fig.6.25 shows numerical values for the overlap function as a function of scattering angle for various QW structures. All trends can easily be explained by the fact that the eigenstates at finite momentum are no longer pure heavy hole or light hole states but form a mixture of the two. This process of band mixing is obviously stronger at larger values of the in-plane momentum. The latter explains the relatively small deviation from unity for the GaAs-QW at in-plane momentum $k_{//}$=0.01Å, as the initial and final eigenstates at this small value of the momentum are essentially still identical with an almost pure heavy hole character. At higher in-plane momentum, the large mixing of heavy and light hole properties results in large differences between the initial and final cell-periodic functions, thus leading to an overlap function that deviates significantly from unity. Illustrative is the overlap function for the GaAs well at $k_{//}$=0.03Å, where it can be seen that the initial and final cell-periodic functions that are separated by 90° in k-space are almost perpendicular. Other interesting observations are that both the inclusion of strain as well as the

reduction in well width push the overlap function towards unity. Again, this was to be expected from band mixing considerations.

Exclusion of the overlap factor from the evaluation can result in a significantly lower theoretical hole mobility than obtained in the proper evaluation. It is estimated from respectively the solid curves at k=0.03Å^{-1} and k=0.05Å^{-1} in fig.6.25 that the the overlap-factor affects the theoretical hole transport mobility in a 60Å $In_{0.15}Ga_{0.85}As$-$Al_{0.33}Ga_{0.67}As$ QW by 30% at a hole density of $1.4 \cdot 10^{12}$ cm^{-2} and as much as a factor 2 at a density of $4 \cdot 10^{12}$ cm^{-2}.

6.4.5 - Diffusion and segregation effects

Before absolute values for the theoretical transport mobility of the previously described p-type structures are given, appropriate error bars on their values have to be determined. As the numerical errors induced in the calculation are expected to be small (and very difficult to evaluate), this section will solely concentrate on the effects on diffusion and segregation, i.e. the influences of the presence of small deviations in the doping distribution from that specified for the growth run.

Fig.6.26 shows the effect of diffusion of the impurity δ-doped plane on the theoretical mobility for the centre and edge δ-plane structures of set (A) at a doping density of $3.5 \cdot 10^{12}$ cm^{-2}. Because of numerical reasons, delta-doped planes are generally given a finite width of 4Å in the SCPS calculations. Diffusion effects were simulated by using the corresponding results for the envelope functions and then calculating the mobility for various impurity distributions (truely δ-shaped or broadened), as the envelope functions are expected to be fairly insensitive to the exact form of the doping layer. All resulting mobilities were normalised to the value obtained for the case where diffusion is absent. A Gaussian broadening profile was assumed in the mobility calculations with corresponding standard deviation σ. A more frequently used parameter to describe broadening effects is the Full Width Half Maximum (FWHM) which is directly related to the standard deviation via $FWHM$=2.355σ [Sch96].

Only small effects on the mobility are observed (fig.6.26). The effect is stronger for the centre-doped than the edge-doped structure as diffusion in the former case corresponds to an effective shift away from the peak of the hole distribution of all the carriers, in contrast to the latter where broadening results in half of the impurities moving away and half actually moving closer to the peak in the hole distribution function. Values for the width of diffusion broadened delta-peaks can

be obtained from [Sch96] where a rapid reduction in standard deviation as a function of decreasing growth temperature is presented for growth of Be in GaAs, σ(660°C)=20.3Å and σ(580°C)=7.1Å. As the InGaAs regions were grown at 510°C, diffusion is not expected to significantly influence the results.

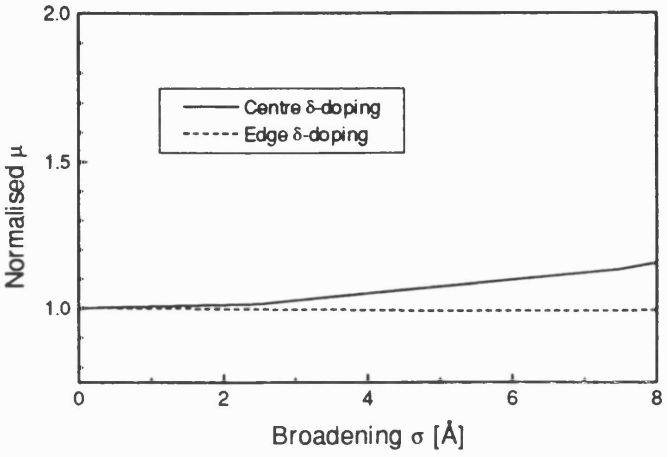


Fig.6.26 : Theoretical hole mobility in the centre and edge δ-doped structure of set (A) as a function of the broadening of the impurity δ-planes. Gaussian broadening is assumed with a standard deviation σ. All mobilities are normalised to the value corresponding to zero broadening.

Fig.6.27 : Similar to above, using the same vertical scale, now showing the effect of segregation on the theoretical (normalised) hole mobility.

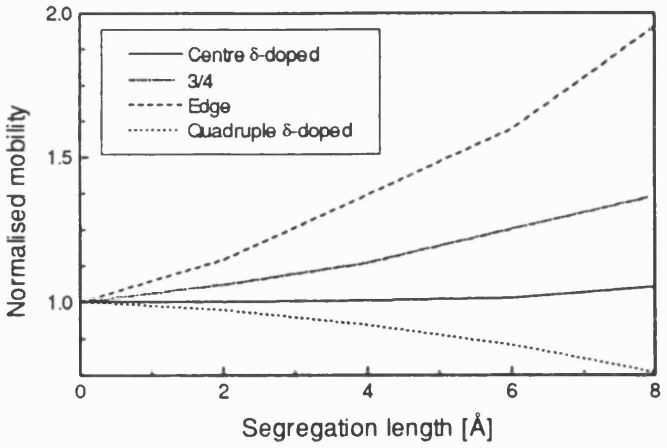


Fig. 6.27 shows the effects of segregation on the hole mobility of both the centre-, edge and quadruple δ-doped structures described in 6.4.2, again for a doping density of $3.5 \cdot 10^{12}$ cm^{-2} for the samples of set (A) and $4.4 \cdot 10^{12}$ cm^{-2} for the 4-δ-plane structure of (B). Segregation is fundamentally different from diffusion by the fact that it has a preferential direction, namely towards the surface. For simplicity, its effects were evaluated by comparing the theoretical mobilities for various slightly shifted positions of the δ-impurity planes, respectively at 0, 2, 4, 6 and 8Å from their intended positions, again simulated by 4Å wide 'delta'-distributions. This model provides an upper limit to the effect of segregation since in practice not all the impurity atoms will move away from the plane. The main difference between this evaluation and the previous one is that SCPS calculations had to be performed for each separate shift in the impurity planes, as it is incorrect to assume that such a uni-directional shift will not affect the envelope functions.

In contrast to the results found for diffusion, segregation is expected to significantly affect the mobilities as a segregation length of 4Å results in an increase in mobility by 37% in the edge-doped structure at a doping density of $3.5 \cdot 10^{12}$ cm^{-2} and a decrease in mobility of 6% in the quadruple δ-doped structure at $4.4 \cdot 10^{12}$ cm^{-2}. These consideration will have to be taken into account when comparing the experimental and theoretical hole mobilities.

6.4.6 - Comparison between theoretical and experimental hole mobilities

Theoretical calculations for the expected hole mobility in the structures described in §6.4.1 and analysed in §6.4.2 will now be presented. The rescaling factor that was shown to be of the order 2 for electron mobility calculations (§6.3.2) is an interesting subject of research, the results presented here will give the first indications about its necessity and value for hole mobility calculations.

Fig.6.28 shows a comparison between theoretical predictions for the ionised impurity scattering limited mobility as well as the experimental 77K hole mobility. The calculations were not performed at absolute zero, as required ideally in a framework where scattering processes are limited to the Fermi-contour only, but at 20K due to numerical reasons. The latter is however not expected to significantly affect the results. All calculations are self-consistent, featuring a proper Fermi-Dirac analysis over the warped hole bands. The effective mass at the Fermi energy that is required in the scattering model is numerically evaluated using (1.1). All other parameters used were supplied in chapter 5.

After performing all the numerical work described in chapter 5, it is disappointing to find that an overall scaling factor of approximately 8 is required to obtain a reasonable agreement between experiment and theory. Although the need for such a rescaling factor was to be expected, it would have been much more satisfactory if this factor had been closer (or even equal) to the scaling factor obtained from the electron mobility analysis of section 6.3. The particular factor 8 was obtained by comparing the theoretical and experimental results, concentrating on the centre and 3/4 single δ-doped structures as it was shown that those were least affected by segregation effects; the latter is expected to be the reason for the higher than expected experimental hole mobilities for the edge-doped structures. The differences between the theoretical and experimental results for set (B) can easily be explained by

- assuming slightly higher segregation effects, which would consequently result in a higher scaling factor when comparing the theoretical and experimental mobilities
- intrinsic errors in the above scaling factor concept, as no insight in the origin of this factor has been presented yet. The scaling factor could for example be carrier density dependent.
- the approximation that Ehrenreich's overlap function is calculated within a 4-band approximation which could affect the calculations at higher density.
- A deviation in the confinement potential from symmetric, either due to a residual effect of the effort to grade the QW so as to possess a dish-shape, segregation effects in the δ-doped planes or incorrect accounting for surface/substrate carrier losses in the design (i.e. incorrectly specified impurity densities of the outer two δ-doped planes that were specifically designed to 'shield' the QW from the surface and substrate). All these effects particularly affect high density structures designed to be symmetric, as such a small variation can make the system unstable and force the hole distribution to 'tip-over' at these high carrier distributions. This would result in a hole distribution much closer to one of the channel impurity planes than would be the case in a symmetric structure.

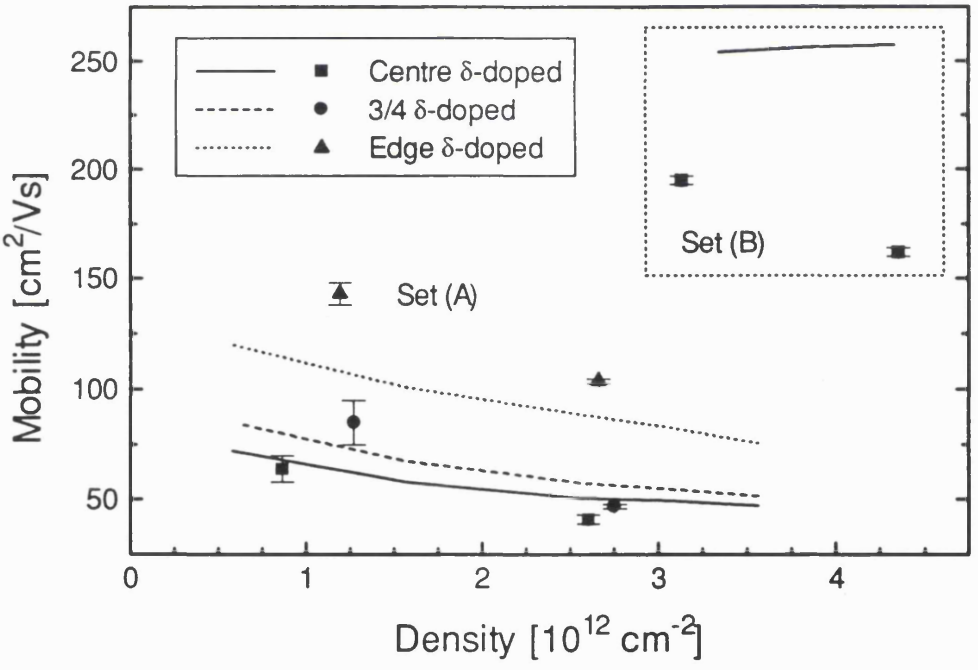


Fig.6.28 : Theoretical and experimental hole mobilities for the samples (A) and (B). The theoretical mobilities have been rescaled by a factor 8.

Summarising, reasonably accurate absolute predictions can be made for ionised impurity scattering limited *p*-type heterostructures when accepting a scaling factor of approximately 8. The next section will now discuss the origin and value of this factor.

6.4.7 - Origin and explanation of the presence of the scaling factor

At this point, one might argue that the discrepancy between the experimental and theoretical mobility could be caused by the failure to included other scattering mechanisms, so that the presence of this (large) scaling factor is simply an artefact of the restriction to ionised impurity scattering only. However, fig.6.28 shows that the RPA scattering model correctly predicts the *relative* improvement in mobility when moving the delta-doped plane from the centre to the edge of the QW. As a result, the introduction of an extra scattering time τ^* corresponding to all the un-included scattering processes can never bring the theoretical and experimental results into agreement via Matthiessen's rule, as the latter would level the theoretical prediction for the relative improvement between the centre- and edge-doped structures. This is an important observation, it states that the occurrence of this scaling factor can only be explained by the presence of too strong approximations in the implementation of the Boltzmann transport equation or in the scattering formalism.

A possible explanation for the requirement of a scaling factor 2 in electron mobility calculations follows from the work by Appel [App61] who states that the presence of electron-electron scattering in a system under the influence of an external electric field modifies the Boltzmann distribution, therby actively changing the conductivity. In terms of fig.1.8 which explained the collective motion of carriers via a shift of the Fermi-contour, the presence of electron-electron interaction results in an overall smaller equilibrium displacement of this Fermi contour than that would have been obtained when neglecting *e-e* interactions. The exact deviation is too elaborate to be evaluated here, so that this section restricts itself to just quoting the calculated ratio between the conductivity with and without *e-e* interaction, which was given as 0.58:1 for an ionised impurity limited bulk semiconductor. Adopting the latter results in a scaling factor of 1.72 which is very close to the factor 2 used in section 6.3.

Turning to the valence band case, it is very difficult to evaluate the above scaling resulting from the effects of hole-hole interaction on the Boltzmann distribution in warped valence bands. Although it is possible that the rather large deviation of a factor 8 could be caused by hole-hole interaction effects, it seems unlikely that it would result in such a large deviation.

Assuming that hole-hole interaction can only partially account for the large scaling factor, this section now proposes an alternative explanation for the difference between theoretical and experimental hole mobilities. Here, the error analysis is now limited to the ionised impurity

scattering model itself. As scattering is well understood theoretically and determined directly from Fermi's Golden rule (see chapter 4), it follows that the most likely cause for the differences between theory and experiment is an over-estimation of the screening effects as accounted for within the RPA. It is now postulated that the RPA formalism breaks down at such high (up to $0.4m_0$) effective masses at the Fermi energy.

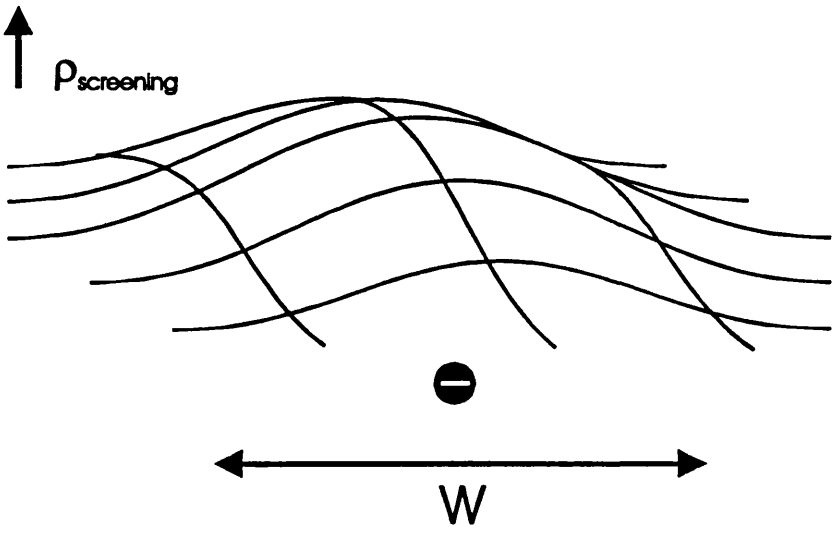


Fig.6.29 : Schematic of the screening charge density $\rho_{screening}$ surrounding a negatively charged impurity located in a p-type QW channel. The density is plotted as a function of the in-plane position. W denotes a quantitative length scale for the particular distribution.

Fig.6.29 shows a schematic picture of the screening charge around a negatively charged impurity where a parameter W describing the width of the charge distribution has been introduced. A very important observation is that the total screening charge is equal to that of the impurity (thus $+e$). This follows from the simple argument that if the charges did not cancel each other, more charge would be attracted until a state of charge neutrality was reached [Mah90]. It is now proposed that the width of the screening charge distribution in the RPA model is so narrow that quantum confinement effects start to play a role, something which is explicitly excluded from the RPA evaluation.

A qualitative value for the width can be found by using the following argument. As seen in section 4.3, the one band RPA dielectric response matrix is essentially the same as that obtained in the easy to interpret TFA. Using the TFA result, one obtains that the projection on the envelope functions of the total scattering potential, thus that due to the presence of the ionised impurity ϕ^{ext} and of the screening charge, is given in Fourier space by [see (4.6)]

$$\bar{\phi}(q) = \left(1+\frac{1}{Sq}\right)^{-1}\bar{\phi}^{ext}(q) = \frac{Sq}{1+Sq}\bar{\phi}^{ext}(q) \quad , \quad S=\frac{2\varepsilon_0\varepsilon_r\pi\hbar^2}{m^*e^2} \tag{6.17}$$

Its relation to ionised impurity scattering is given by Fermi's Golden rule from which follows that the transition rate is proportional to $\left|\bar{\phi}(q)\right|^2$ [see (4.9)]. Rather than stay with the Fourier space description, it proves useful to return to real space. It is then straightforward to see that Fermi's Golden rule does nothing more than select that particular component of the perturbation potential with the repetition period $exp(iqr_{//})$, where $r_{//}$ is the in plane coordinate and q is the momentum transfer in the scattering process $k \rightarrow k'$. The relation to the screening width W is now as follows. Equation (6.17) predicts that the screening is very efficient at low values of q. This is consistent with the picture that a wave $exp(iqr_{//})$ with a long wavelength compared with the charge distribution will simply pass this charge without noticing (see fig.6.30) as it will not be able to resolve the separation between impurity and screening charge and only see a total charge of zero. Obviously, a wave with wavelength compatible to the charge distribution will pick up the individual screening and impurity component and thus scatter.

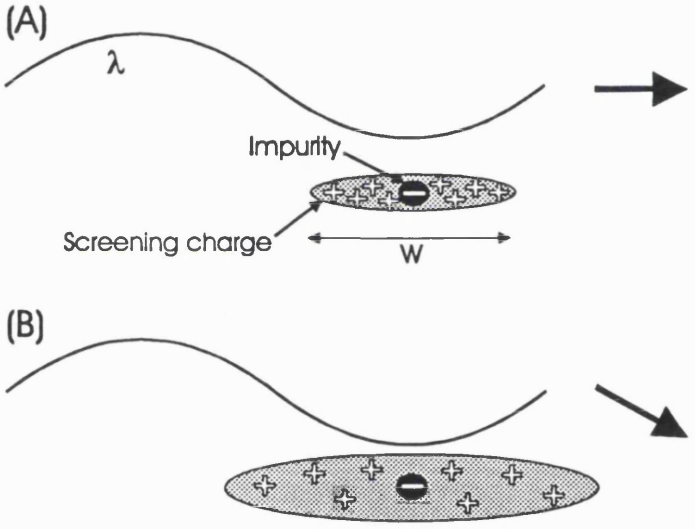


Fig.6.30 : Illustration of the effect of the screening charge distribution width W on the scattering of a plane wave with wavelength λ. (A): As the total screening charge is equal to the charge of the impurity, it follows that the plane wave with λ>W does not scatter, as it does not see the separation between impurity and screening charge and thus effectively sees a zero total charge. (B): As λ becomes comparable to W, the plane wave starts to see the charge fluctuation and scatters as a result.

Using the latter, one can now obtain a qualitative value for the distribution width W. The momentum transfer q at which the perturbation potential of the impurity is screened such that the scattering rate is only half of what one would have obtained without accounting for screening is given by $q^{-1}=(\sqrt{2}-1)\cdot S$. Using a typical effective mass for the structures of sets (A) and (B) of $0.3m_0$, one obtains a corresponding momentum transfer vector $q=0.20\text{Å}^{-1}$ and, using that the periodicity of such a plane wave is given by $W=2\pi/q$, a screening charge distribution width of 31Å. The latter corresponds to an average screening charge density of $1.3\cdot10^{13}$ cm^{-2}. Clearly, quantisation effects will play a large role at these narrow distribution widths, illustrated in fig 6.31, leading to a broader distribution and so accounting for the over-estimation of the RPA

scattering model for *p*-type structures. In contrast, the same analysis reveals a screening charge distribution width of 136Å in *n*-type GaAs which explains the absence of these quantisation effects in electron systems.

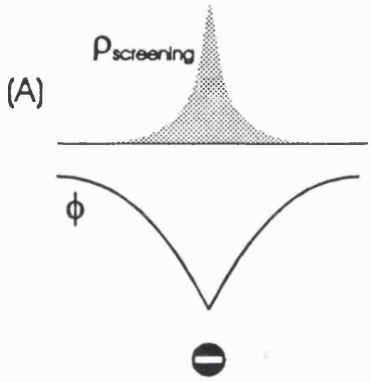


Fig.6.31 : Schematic picture of the effect of quantum confinement on the screening charge distribution. (A) Depicts a 'spiked' distribution that follows from a bulk approach, while (B) shows a distribution corresponding to $|\psi|^2$ of a quantised state.

A rather crude but convenient matter to simulate such wider distributions in hole-systems is by setting the screening mass of (4.6,4.7) to a lower value. Re-iterating, this only affects the distribution and not the amount of the screening charge. Setting the distribution width to a constant realistic value, thereby fixing the screening mass, would result in a mass dependency of the hole mobility at low temperature of $\mu \propto m^{-2}$ as outlined in chapter 1. Such a mass dependency was actually observed by Fritz et al. [Fri86] at 76K in a set of $In_{0.2}Ga_{0.8}As$-GaAs QWs for hole densities above $n_{2d}=1.5\cdot10^{11}$ cm^{-2}, and provides some confirmation that the screening mass becomes a constant at values above $m^*=0.14m_0$ [Fri86] of the Fermi level effective mass.

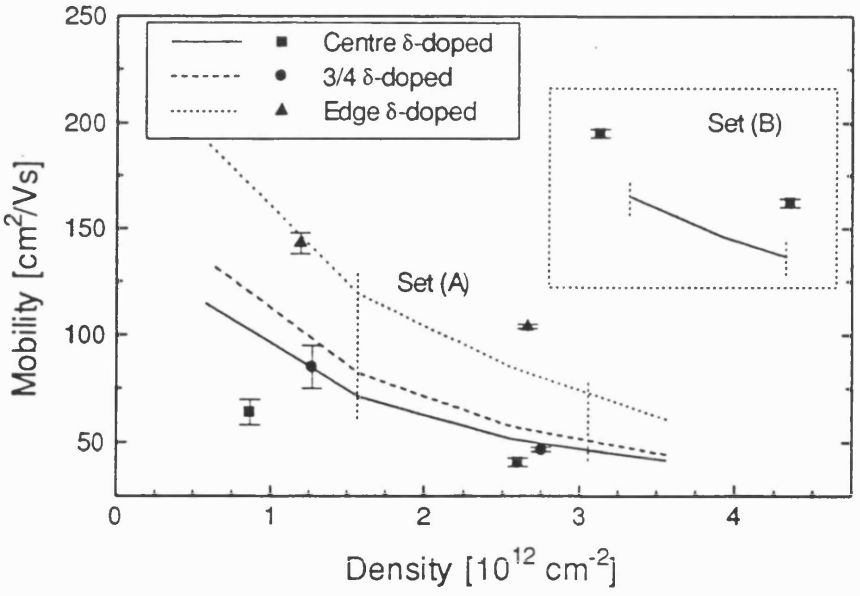


Fig.6.32 : As figure 6.28, now using a screening mass of $m^*=0.14m_0$ and a scaling factor of 4.5. The vertical lines show the predicted carrier densities, assuming that all impurities are ionised.

Finally, fig.6.32 provides a comparison between the theoretical hole mobility and experimental data similar to fig.6.28 but now using a (rather arbitrary) fixed screening mass of $0.14m_0$, rather than to set this screening mass to the effective mass at the Fermi-energy for each particular structure. Slightly better overall agreement between theory and experiment is obtained, especially with respect to the trend in the structures of set (B). Here, a scaling factor of 4.5 was used which is in better agreement with that found in electron mobility calculations and the corresponding explanation in terms of electron-electron (hole-hole) scattering.

This chapter now ends with the conclusion that a powerful framework has been set up to predict the hole mobility in ionised impurity limited structures. Good agreement between theory and experiment was obtained for a set of channel δ-doped *p*-type $In_{0.15}Ga_{0.85}As$-$Al_{0.33}Ga_{0.67}As$ QWs when introducing an overall scaling factor, whose value was explained in terms of quantum confinement effects in the screening charge density and hole-hole interaction. The following and final chapter of this thesis will now outline possible subjects of further research based on the foundations laid in this work.

chapter 7 : Conclusions and recommendations

7.1 - Conclusions

Performance of *p*-channel HFETs is substantially lower than their *n*-type counterparts, essentially because of the lower hole mobility as compared to that of electrons. The poor performance of *p*-channel devices has almost completely prevented their use, despite the intrinsic advantages of complementary circuits. An investigation to enhance the performance of *p*-type FETs for power applications by means of selective doping and wavefunction engineering was performed, both theoretically as well as experimentally.

A full-self-consistent valence band model that specifically includes valence band warping in the evaluation of the Fermi-Dirac integral was set up. It was shown that such numerically intensive calculations are within reach of middle-range computers, such as a Sun Sparc 20, when using various mild approximations. The most import one of those approximations is to use an expansion method to solve the system of coupled differential equations of the Hamiltonian, which showed fast convergence and good agreement with exact calculations.

A substantial amount of work was devoted to the fundamentals of effective mass theory. The resulting orbital (or tight-binding) picture of $k \cdot p$ theory resulted into an alternative approach to derive effective mass Hamiltonians for non-[001] growth, compared to the conventional group-theoretical 'theory of invariants'. Here, a general $|J,m_j\rangle$ basis set to set up effective mass Hamiltonians for arbitrary growth directions was derived.

As a result of the above approach, significant contributions to the elementary question of how to connect envelope functions across an interface in compositionally varying structures were presented. This work accumulated in a general form for a 6-band Hamiltonian with a corresponding set of boundary condition rules that are to replace the conventional symmetrisation rule, both valid for arbitrary growth directions. It was shown that the adoption of the new boundary conditions only has marginal affects when evaluating GaAs-based structures, but is crucial for the technologically relevant InP system, especially when looking at [110] structures.

Furthermore, an ionised impurity scattering model within the RPA formalism was set up. Literature results have shown that electron mobility calculations within this formalism require a

scaling factor of order 2 to bring the experimental and theoretical results within agreement, independent of the heterostructure design. It was shown here that this factor is also independent of material system, so that it was concluded that the RPA formalism can give reasonably accurate *absolute* values for the mobility, simply by accepting the presence of such a scaling factor.

It was shown that interband transitions in non-symmetric structures with multiple occupied bands can have a large effect on the electron mobility. From this, it was deduced that wavefunction engineering is only effective in systems with a single occupied subband, as optimisation of the first subband generally results in a reduced performance of higher lying states.

Turning to *p*-type systems, the scope of wavefunction engineering techniques to enhance the hole mobility in highly (channel) doped systems was investigated. It was shown that the use of graded QW structures has to be avoided, as the resulting spin-splitting of the valence bands would have a negative effect on the hole mobility. Selective doping however offers significant gain in mobility. Experimental measurements on a set of channel δ-doped *p*-type $In_{0.15}Ga_{0.85}As$-$Al_{0.33}Ga_{0.67}As$ QWs revealed enhancements in the hole mobility of a factor 2.5 at a carrier density around $2.5 \cdot 10^{12}$ cm^{-2} when moving the impurity plane away from the centre of the well to the QW interface.

Hole mobility calculations were presented using the same RPA formalism as used in the *n*-type calculations, believed to be the first of their kind. Additional complexity when compared to the electron case arises from the fact that the hole bands exhibit considerable band mixing effects, i.e. do not stay pure heavy of light hole like at finite momentum. This results in an additional angle dependent factor in the scattering equations, Ehrenreich's overlap function, for which numerical values were presented. Significant changes in the angular dependence were found when comparing the two-dimensional results to the bulk case. For the latter, good agreement with the results published in the literature was obtained.

As for the electron case, reasonable accurate predictions for the absolute value of the hole mobility could be obtained by introducing a scaling factor. After all numerical work involved, it is disappointing to find that this factor is now of the order 8 as compared to the earlier mentioned factor 2 in electron mobility calculations. It was proposed that the electron mobility scaling factor is caused by electron-electron scattering effects. Such an explanation had not earlier been proposed as papers published in the literature that explicitly deal with the RPA tend to accept this factor without giving any explanation of its origin. In contrast, it was postulated that the hole mobility scaling factor is not solely caused by hole-hole interaction, but by a combination of the latter and

quantum confinement effects on the screening charge distribution. A first effort to account the quantum confinement effects resulted in a slightly better overall agreement between theory and experiment, especially with respect to the trend in the mobility for the second set of samples. Further, it resulted in a new scaling factor of 4.5, which is in better agreement with that found in electron mobility calculations and the corresponding explanation in terms of electron-electron (hole-hole) scattering.

Suggestions for future work

Using the results obtained in this work and the frameworks set out, a solid basis for the theoretical analysis of electron and hole systems has been presented. The first results show that a considerable gain in the electrical properties can be obtained by selective doping or wavefunction engineering techniques. The next step is now to actually design and produce such high-performance devices. It was shown that the *n*-type InP-based graded channel structures evaluated by Roberts et al. at the IRC were severely limited by the occupation of the second subband and resulting interband scattering. It would be worthwhile to re-design these structures so that they have only one occupied subband, and then test the true scope of the use of such graded channels to enhance the mobility.

It was shown that such graded structures have to be avoided in hole systems because of the adverse effect of the resulting spin-splitting on the hole mobility. An alternative to using graded structures is to switch to a growth direction such as [111] in which strained QWs exhibits a strong internal piezo-electric field. Rather than utilise this field to minimise interface roughness scattering as in [Hit95], it would be interesting to use it to minimise ionised impurity scattering in *p*-type structures.

However, it was shown to be able to significantly enhance the mobility by selective doping only. As state-of-the-art hole mobilities require a high indium composition, only a comparison between various doping schemes at elevated indium content can truly establish the scope to enhance the transport mobility in highly-doped structures such as for power FET applications.

Wavefunction engineering could be useful for SiGe heterostructures for a slightly different reason. Here, the back interface exhibits a much rougher structure than the interface closer to the substrate. Wavefunction engineering could be used to pull away the carriers from this rough interface much like it was used to separate the carriers from the impurities in this work.

Optimisation of optical devices has only briefly been discussed in this thesis. Again, the theoretical models developed here can be of great value. Theoretical calculations of the expected optical anisotropy in [110] grown vertical cavity structures (not shown in this thesis) proved to be in excellent agreement with those performed by Nojima [Noj93]. Photo-Luminescence studies could greatly benefit from such calculations, offering enhanced means to interpretation of (polarisation dependent) PL-data.

Apart from the immediate practical application of the outcome of this thesis, there are several other subjects and fields that are would benefit from additional work.

Although it was shown in the literature that the new boundary conditions are more physical than those obtained using the conventional symmetrisation procedure, there is no solid experimental proof of their validity. Mapping out the bandstructure of an InP-based (preferable [110]) orientated) narrow well QW using the Photoluminescence technique of [Kas94] could resolve this. The boundary conditions have a significant effect on whether or not the higher lying subbands exhibit a van Hoof singularity in their density of states. As these singularities have a strong effect on PL-spectra, theoretical modelling of these singularities combined with a PL-study [Kem96] could also provide some evidence of the correctness of the new boundary conditions, as could cyclotron resonance measurements. Finally, a careful comparison between laser gain calculations similar to [Men92] for InP systems and experimental data, could possible also help to resolve the matter.

Theoretical rather than experimental evidence of the correctness of the boundary conditions could for example be obtained by comparing calculated results for the bandstructure using an effective mass model to those obtained in an alternative model that is not affected by the boundary conditions problem, such a ab-initio or pseudo-potential calculations. Some work is currently undertaken by the author and P.N.Stavrinou, in collaboration with the University of Durham.

Although the electron scaling factor could be explained in terms of electron-electron interaction, as theoretical evaluations for the bulk case reveal a scaling factor close to that observed here, there is a lack of theoretical insight in two-dimensional electron structures and hole systems in general. Additional justification for the presence of such a scaling factor, as well as theoretical estimates, would greatly enhance the appeal for RPA electron and hole mobility calculations.

Appendices

A1 - Using Mathematica to obtain explicit forms for the zone-centre wavefunctions

The development of the $k \cdot p$ model for arbitrary growth directions, especially to establish the relation between the Hamiltonian and the zone-centre wavefunctions, has relied heavily on the use of the mathematical package Mathematica. This software package proved essential for the derivation (and consequent publication) of the [110] boundary conditions [Sta97] and subsequent development of an analytic set of rules to replace the previously used symmetrisation procedure [Dal98]. This appendix will now outline how such a mathematics package can be utilised to obtain explicit forms for the zone-centre wavefunctions $|J,m_j\rangle$. The approach taken here will restrict itself to obtaining the solution for just a single growth direction, rather than derive the general solution (2.26) in terms of arbitrary orthogonal basis vectors.

In the following program, a six-element vector (a,b,c,d,e,f) has been defined which denotes the general linear combination of p-orbitals

$$\left|J,m_j\right\rangle = (aX + bY + cZ)\uparrow + (dX + eY + fZ)\downarrow \qquad \text{(A1.1)}$$

similar to that used in chapter 2. The program starts by defining the action of the momentum ($L_{x,y,z}$) and spin operators ($S_{x,y,z}$) on this general solution. After this, the coordinate transformation ($x'=[1,1,-2]/\sqrt{6}$, $y'=[-1,1,0]/\sqrt{2}$, $z'=[1,1,1]/\sqrt{3}$) is performed on the angular momentum operators, to obtain the action of $L \cdot S$ and the angular momentum operator J along [111] on this general solution, where the spin has been orientated along [111]. Finally, the appropriate constants $a...f$ corresponding to the state $|1/2,+1/2\rangle$ are calculated. Note that such a state corresponds to a simultaneous solution of $L \cdot S\ |J,m_j\rangle = -|J,m_j\rangle$ and $J_{111}\ |J,m_j\rangle = 1/2\ |J,m_j\rangle$. The former equation is equivalent to the requirement that $J=1/2$, using (2.20) and the fact that $J^2|J,m_j\rangle = J(1+J)|J,m_j\rangle$ ($\hbar=1$).

```
"##### Define action of the operators lx..z, sx..z #####";

lx[{x_,y_,z_,xx_,yy_,zz_}]:={   0,-I z, I y,    0,-I zz, I yy}
ly[{x_,y_,z_,xx_,yy_,zz_}]:={ I z,   0,-I x, I zz,    0,-I xx}
lz[{x_,y_,z_,xx_,yy_,zz_}]:={-I y, I x,   0,-I yy, I xx,    0}

sx[{x_,y_,z_,xx_,yy_,zz_}]:=  { xx, yy, zz,  x,  y,  z}/2
sy[{x_,y_,z_,xx_,yy_,zz_}]:=I {-xx,-yy,-zz,  x,  y,  z}/2
sz[{x_,y_,z_,xx_,yy_,zz_}]:=  {  x,  y,  z,-xx,-yy,-zz}/2
```

```
"########## Perform coordinate transformation ##########";

lx2[{x_,y_,z_,xx_,yy_,zz_}]:=1/Sqrt[6] (   lx[{x,y,z,xx,yy,zz}]
                                         +ly[{x,y,z,xx,yy,zz}]
                                       -2 lz[{x,y,z,xx,yy,zz}] )
ly2[{x_,y_,z_,xx_,yy_,zz_}]:=1/Sqrt[2] (  -lx[{x,y,z,xx,yy,zz}]
                                         +ly[{x,y,z,xx,yy,zz}] )
lz2[{x_,y_,z_,xx_,yy_,zz_}]:=1/Sqrt[3] (   lx[{x,y,z,xx,yy,zz}]
                                         +ly[{x,y,z,xx,yy,zz}]
                                         +lz[{x,y,z,xx,yy,zz}] )

"#################### Calculate L.S ####################";

ls[{x_,y_,z_,xx_,yy_,zz_}]:=( lx2[sx[{x,y,z,xx,yy,zz}]]
                             +ly2[sy[{x,y,z,xx,yy,zz}]]
                             +lz2[sz[{x,y,z,xx,yy,zz}]] );

"############ Calculate J along [111] ##################";

j111[{x_,y_,z_,xx_,yy_,zz_}]:=lz2[{x,y,z,xx,yy,zz}]+
                              sz[{x,y,z,xx,yy,zz}];

"### Calc. simultaneous eigenfunction of J111 and L.S ###";

Solve[j111[{a,b,c,d,e,f}]==1/2 {a,b,c,d,e,f} &&
       ls[{a,b,c,d,e,f}]==-1  {a,b,c,d,e,f}      ,{a,b,c,d,e,f}];

Simplify[%];
Cancel[%]
```

Output

```
              f                   f                   f
{{a -> -(-------), b -> -(-------), c -> -(-------),
           Sqrt[2]             Sqrt[2]             Sqrt[2]

        I                          -((1 + I Sqrt[3]) f)
   d -> - (I + Sqrt[3]) f, e -> --------------------)}
        2                                2
```

Normalising the latter by multiplying each element by $(\sqrt{2})/3$ gives as a final result

$$\left|\frac{1}{2},\frac{1}{2}\right\rangle = -\frac{1}{3}(X+Y+Z)\uparrow -\frac{1-i\sqrt{3}}{3\sqrt{2}}X\downarrow -\frac{1+i\sqrt{3}}{3\sqrt{2}}Y\downarrow +\frac{\sqrt{2}}{3}Z\downarrow \qquad \text{(A1.2)}$$

which agrees with the general solution (2.26) for the particular choice of axis and growth direction.

A2 - Obtaining effective mass Hamiltonians and boundary conditions

After having obtained a set of basis functions, the zone-centre solutions, one can derive the corresponding effective mass Hamiltonian by evaluating the inner product $\langle m|H|n\rangle$ where m,n run over all the elements of the basis set. Again, Mathematica can be of great help as the mathematics tend to become very laborious when deviating from the high-symmetry directions, especially when one wants to retain the operator ordering to subsequently derive the boundary conditions. The next example shows how this task was implemented in Mathematica. Some repetitive statements were

cut out for clarity. The particular growth direction under consideration is [110]. The constants *s,p,d* respectively denote the mass parameters σ,π and δ.

```
"Define the inner products <a|H|b>, a,b=X,Y,Z
 two momentum vectors are defined, {kx,ky,kz} and {lx,ly,lz}, where k
 is to be interpreted as the operator BEFORE and l as the operator
 AFTER the mass-parameters (s,p,d)";

XHX:=kx(-1+6s+12d)lx + ky(-1+6p)ly + kz(-1+6p)lz;
XHY:=6 kx(s-d)ly + 6 ky p lx;
XHZ:=6 kx(s-d)lz + 6 kz p lx;
YHX:=6 ky(s-d)lx + 6 kx p ly;
YHY:=ky(-1+6s+12d)ly + kx(-1+6p)lx + kz(-1+6p)lz;
YHZ:=6 ky(s-d)lz + 6 kz p ly;
ZHX:=6 kz(s-d)lx + 6 kx p lz;
ZHY:=6 kz(s-d)ly + 6 ky p lz;
ZHZ:=kz(-1+6s+12d)lz + kx(-1+6p)lx + ky(-1+6p)ly;

"Perform the coordinate tranformation, new vectors along
 k1=[0,0,-1], k2=[-1,1,0] and k3=[1,1,0]";

kx:=   -1/Sqrt[2] k2 +1/Sqrt[2] k3;
ky:=    1/Sqrt[2] k2 +1/Sqrt[2] k3;
kz:=-k1                          ;

lx:=   -1/Sqrt[2] l2 +1/Sqrt[2] l3;
ly:=    1/Sqrt[2] l2 +1/Sqrt[2] l3;
lz:=-l1                          ;

"Define heavy hole state J=3/2, mj=+3/2";
 |3/2,+3/2> = a1 X^ + b1 Y^ + c1 Z^ + d1 X. + e1 Y. + f1 Z. ";

a1:= I/2;
b1:=-I/2;
c1:= 1/2 Sqrt[2];
d1:= 0;
e1:= 0;
f1:= 0;

"Define light hole state J=3/2 mj=+1/2";

a2:= I/ Sqrt[3];
b2:= I/ Sqrt[3];
c2:= 0;
d2:=-1/(2 Sqrt[3]);
e2:=+1/(2 Sqrt[3]);
f2:= I/Sqrt[6];

{The light hole state with mj=-1/2 and the heavy hole state with
mj=-3/2 are defined in a similar way}

"############ Calculate Hamiltonian elements ##########";

"Calculate Hhh =<3/2,3/2|H|3/2,3/2>"

Hhh=Conjugate[a1] (a1 XHX + b1 XHY + c1 XHZ) +
    Conjugate[b1] (a1 YHX + b1 YHY + c1 YHZ) +
    Conjugate[c1] (a1 ZHX + b1 ZHY + c1 ZHZ);

Apart[%];
%/.k2 l1->k1 l2;
Cancel[%]
```

```
"Calculate L =<3/2,3/2|H|3/2,1/2>"

L  =Conjugate[a1] (a2 XHX + b2 XHY + c2 XHZ) +
    Conjugate[b1] (a2 YHX + b2 YHY + c2 YHZ) +
    Conjugate[c1] (a2 ZHX + b2 ZHY + c2 ZHZ) +
    Conjugate[d1] (d2 XHX + e2 XHY + f2 XHZ) +
    Conjugate[e1] (d2 YHX + e2 YHY + f2 YHZ) +
    Conjugate[f1] (d2 ZHX + e2 ZHY + f2 ZHZ);

Apart[%];
%/.k2 l1->k1 l2;
Cancel[%]

{The other matrix elements in a 4-band Hamiltonian are calculated in a
similarl way}

"### Set up Hamiltonian matrix using calc.elements ####";

ham={ {Hhh,L,M,DI1},{LL2,Hlh,DI2,MM}
      ,{MM2,DI3,Hlh,LL},{DI4,MM3,LL3,Hhh} };

"############ Calculate Boundary Conditions ############
 replace kz,lz-> d/dz and retain only terms of the form
 form kx d/dz (s,p,d), ky d/dz (s,p,d) or d/dz (s,p,d) d/dz";

Apart[%];
%/.k1 l1-> 0;
%/.k2 l2-> 0;
%/.k1 l2-> 0;
%/.k2 l1-> 0;
%/.k1 l3-> 0;
%/.k2 l3-> 0;
%/.k3 l3-> d_dz;
%/.k3 l1-> I K1;
%/.k3 l2-> I K2;
Cancel[%];
Apart[%]
```

Output

```
Hhh =<3/2,3/2|H|3/2,3/2>
 (-2 k1 l1 + 12 d k1 l1 - 2 k2 l2 + 3 d k2 l2 -
    2 k3 l3 + 9 d k3 l3 + 6 k1 l1 p + 12 k2 l2 p +
    6 k3 l3 p + 6 k1 l1 s + 6 k2 l2 s) / 2

L =<3/2,3/2|H|3/2,1/2>
Sqrt[3] (-3 d k3 l2 + 2 I d k1 l3 - d k2 l3 -
    2 I k3 l1 p + 2 k2 l3 p - 2 I k1 l3 s - 2 k2 l3 s)

{Similar output for other Hamiltonian elements}

Boundary condition, continuity of :
{(3^(1/2)*(-3*d + 2*p)*(d_dz))/2,
   2*(K1 + I*K2)*(d + p) - 2*I*(-I*K1 + K2)*s,
   ((-2 + 7*d + 10*p + 8*s)*(d_dz))/2,
   3^(1/2)*(2*d*K1 + I*d*K2 - 2*I*K2*p - 2*K1*s +
      2*I*K2*s)}, {0,
   (3^(1/2)*(-3*d + 2*p)*(d_dz))/2,
   3^(1/2)*(3*I*d*K2 + 2*K1*p),
   ((-2 + 9*d + 6*p)*(d_dz))/2}}
```

This matrix is identical of that published in [Sta97]. Obviously, the whole approach can be extended to a 6 band model.

A3 - Derivation of the boundary condition rules

The derivation of the boundary condition rules for arbitrary growth directions (§6.1.2) is considered to be much more complicated than the derivation of the boundary conditions for one specific growth direction such as shown in appendix A2. Again, the derivation relied heavily on the use of Mathematica to perform the derivation. In contrast to the latter simulation, it proved impossible to set up a single routine that would give explicit forms of the final rules directly. Instead, several calculations that were all limited to a specific element of the Hamiltonian, a single mass parameter or both, were performed. The combination of all these calculations was then used to piece together the final analytical boundary condition rules. One of such calculations is shown below, which examines the presence of combinations of the mass parameter σ and the momentum operator along the growth direction (k_3 or l_3, where the former is located BEFORE σ and the latter AFTER σ).

```
"Define the inner products <a|H|b>, a,b=X,Y,Z";

XHX:=kx(-u+6s+12d)lx+ky(-u+6p)ly+kz(-u+6p)lz;
XHY:=kx(6s-6d)ly+ky(6p)lx;
XHZ:=kx(6s-6d)lz+kz(6p)lx;
YHX:=ky(6s-6d)lx+kx(6p)ly;
YHY:=ky(-u+6s+12d)ly+kx(-u+6p)lx+kz(-u+6p)lz;
YHZ:=ky(6s-6d)lz+kz(6p)ly;
ZHX:=kz(6s-6d)lx+kx(6p)lz;
ZHY:=kz(6s-6d)ly+ky(6p)lz;
ZHZ:=kz(-u+6s+12d)lz+kx(-u+6p)lx+ky(-u+6p)ly;

"For the moment, the evaluation is restricted to the sigma (s) elements
 only"

u=0;
p=0;
d=0;

"Perform coordinate tranformation, new vectors along
 k1=[a1,a2,a3], k2=[b1,b2,b3] and k3=[c1,c2,c3]";

kx:= a1 k1 +b1 k2 + c1 k3;
ky:= a2 k1 +b2 k2 + c2 k3;
kz:= a3 k1 +b3 k2 + c3 k3;

lx:= a1 l1 +b1 l2 + c1 l3;
ly:= a2 l1 +b2 l2 + c2 l3;
lz:= a3 l1 +b3 l2 + c3 l3;

"Define heavy hole state J=3/2, mj=+1/2
 |3/2,+1/2> = a1 X^ + b1 Y^ + c1 Z^ + d1 X. + e1 Y. + f1 Z.
 s designates complex conjugate";

wa0s:=a1-I b1;
wb0s:=a2-I b2;
wc0s:=a3-I b3;
wd0s:=0;
we0s:=0;
wf0s:=0;
```

```
"Similarly, J=3/2, mj=-1/2";
wa2:=-a1+I b1;
wb2:=-a2+I b2;
wc2:=-a3+I b3;
wd2:=-2 c1;
we2:=-2 c2;
wf2:=-2 c3;

"Calculate <3/2,3/2|H|3/2,-1/2> excluding normfactor"
ip=(wa0s (wa2 XHX + wb2 XHY + wc2 XHZ) +
    wb0s (wa2 YHX + wb2 YHY + wc2 YHZ) +
    wc0s (wa2 ZHX + wb2 ZHY + wc2 ZHZ) +
    wd0s (wd2 XHX + we2 XHY + wf2 XHZ) +
    we0s (wd2 YHX + we2 YHY + wf2 YHZ) +
    wf0s (wd2 ZHX + we2 ZHY + wf2 ZHZ) );

"As one is interested in the boundary conditions, only the elements
 that contain the operator along (111) need to be retained"
Expand[%];
%/.k1^2->0;
%/.k2^2->0;
%/.k1 k2->0;
%/.k1 l1->0;
%/.k1 l2->0;
%/.k2 l1->0;
%/.k2 l2->0;
%/.l1^2->0;
%/.l2^2->0;
%/.l1 l2->0;
Simplify[%]
```

Output

```
6 (I a1 c1 + I a2 c2 + I a3 c3 + b1 c1 + b2 c2 + b3 c3)

       2              2              2
  (I a1  k3 l1 + I a2  k3 l1 + I a3  k3 l1 + a1 b1 k3 l1 +

                                                    2
    a2 b2 k3 l1 + a3 b3 k3 l1 + I a1 b1 k3 l2 + b1  k3 l2 +

                         2                          2
    I a2 b2 k3 l2 + b2  k3 l2 + I a3 b3 k3 l2 + b3  k3 l2 +

        2              2              2
    I a1  k1 l3 + I a2  k1 l3 + I a3  k1 l3 + a1 b1 k1 l3 +

                                                    2
    a2 b2 k1 l3 + a3 b3 k1 l3 + I a1 b1 k2 l3 + b1  k2 l3 +

                         2                          2
    I a2 b2 k2 l3 + b2  k2 l3 + I a3 b3 k2 l3 + b3  k2 l3 +

    I a1 c1 k3 l3 + b1 c1 k3 l3 + I a2 c2 k3 l3 + b2 c2 k3 l3 +

    I a3 c3 k3 l3 + b3 c3 k3 l3) s
```

The outcome of the above evaluation seems a rather complicated expression at first sight, but a closer look at the first part of the equation between brackets (in bold) reveals that the whole expression reduces to zero when using a orthogonal reference frame in which $\underline{a}\cdot\underline{c} = \underline{b}\cdot\underline{c} = 0$. From this, the important conclusion can be drawn that the off-diagonal Hamiltonian element denoted by R in (6.6) does not contain any combination of σ and the momentum operator acting along the growth direction. Similar calculations were performed for all other Hamiltonian elements.

References

[Abr93] J.K.Abrokwah, J.H.Huang, W.Ooms, C.Shurboff, J.A.Hallmark, R.Lucero, J.Gilbert, B.Bernhardt, G.Hansell
A Manufacturable Complementary GaAs process
IEEE GaAs IC Symp. 1993 127

[Abr97] J.K.Abrokwah, B.Bernhardt and M.Lamacchia
Complementary GaAs (CGaAs™): New enhancements
Sol.State.Elec.41 1997 1433

[Ahn94] D.A.Ahn and S.L.Chuang
The theory of strained layer quantum-well lasers with bandgap renormalisation
IEEE J.of Quant. Elec.30 1994 350

[Alt94] S.L.Altmann
Band theory of solids: An introduction from the point of view of symmetry
Clarendon Press 1994 Oxford

[And85] T.Ando
Hole subband at GaAs/AlGaAs heterojunctions and quantum wells
J.Phys.Soc.Jap.54 1985 1528

[And87] L.C.Andreani, A.Pasquarello and F.Bassani
Hole subbands in strained GaAs-$Ga_{1-x}Al_xAs$ quantum wells: Exact solution of the effective mass equation
Phys.Rev.B 36 1987 5887

[App61] J.Appel
Electron-electron scattering and transport phenomena in non-polar semiconductors
Phys.Rev.122 1961 1760

[Arn92] G.Arnaud, J.Allègre, P.Lefebre, H.Mathieu, L.K.Howard, D.J.Dunstan
Photoreflectance and piezophotoreflectance studies of strained-layer $In_xGa_{1-x}As$-GaAs quantum wells
Phys.Rev.B 46 1992 15290

[Ash76] N.W.Ashcroft and N.D.Mermin
Solid state physics
Holt, Rinehart and Winston 1976 New York

[Bas66] F.Bassani
Methods of band calculations applicable to III-V compounds
Semiconductors and semimetals, vol.1 (edited by R.K.Willardson and A.C.Beer)
Academic press 1966 Orlando

[Bas75] F.Bassani and G.Pastori Parravicini
Electronic states and optical transitions in solids (edited by R.A.Ballinger)
Pergamon press 1975 Oxford

[Bas88] G.Bastard
Wave mechanics applied to semiconductor heterostructures
Les Editions de Physique 1988 Paris

[Bat89] W.Batty, U.Ekenberg, A.Ghiti and E.P.O'Reilly
Valence subband structure and optical gain of GaAs-AlGaAs (111) quantum wells
Semicon.Sc.and Tech.4 1989 904

[Ber95] B.Bernhardt, M.LaMacchia, J.Abrokwah, J.Hallmark, R.Lucero, B.Mathes, B.Crawford, D.Foster, K.Clauss, S.Emmert, T.Lien, E.Lopez, V.Mazzotta, B.Oh
Complementary GaAs (CGaAs™): A high performance BiCMOS alternative
IEEE GaAs IC Symp. 1995 18

[Bir74] G.L.Bir and G.E.Pikus
Symmetry and strain-induced effects in semiconductors
Wiley 1974 New York

[Bla87] J.S.Blakemore
Key papers in physics, no.1 (GaAs)
American Institute of Physics 1987 New York

[Bob96] P.A.Bobbert, H.Wieldraaijer, R.van der Weide, M.Kemerink, P.M.Koenraad, J.H.Wolter
Exchange-correlation energy of a hole gas including valence band coupling
Phys.Rev.B 56 1997 3664

[Bro85] D.A.Broido and L.J.Sham
Effective masses of holes at GaAs-AlGaAs heterojunctions
Phys.Rev.B.31 1985 888

[Bru64] D.Brust
Electronic spectra of crystalline Germanium and Silicon
Phys.Rev.134A 1964 1337

[Bry93] G.W.Bryant, J.L.Bradshaw, R.P.Leavitt, M.S.Tobin, J.T.Pham
Determining band offsets with triple quantum-well structures
Appl.Phys.Lett.63 1993 1357

[Bur87] M.G.Burt
An exact formulation of the envelope function-method for the determination of electronic states in semiconductor microstructures
Semicon.Sc.and Techn. 2 1987 460 (erratum 2 1987 701)

[Bur92] M.G.Burt
The justification for applying the effective-mass approximation to microstructures
J.Phys.Condens.Matter 4 1992 6651

[Bur96] M.G.Burt
Private communications 1996

[Car95] L. de Caro and L.Tapfer
Strain and piezoelectric fields in arbitrarily orientated semiconductor heterostructures. I. Multiple quantum wells
Phys.Rev.B 51 1995 4374

[Cha74] L.L.Chang, L.Esaki, R.Tsu
Resonant tunneling in semiconductor double barriers
Appl.Phys.Lett.24 1974 593

[Cha92] C.Y.Chao and S.L.Chuang
Spin-orbit coupling effects on the valence band structure of strained semiconductor quantum wells

[Che76] J.R.Chelikowski, M.L.Cohen
Nonlocal pseudopotential calculations for the electronic structure of eleven diamond and zinc-blende semiconductors
Phys.Rev.B.14, p.556, 1976

[Che97] M.C.Cheng, L.Y.Guo, R.M.Fithen, Y.S.Luo
A study of the non-parabolic hydrodynamic modelling of a sub-micrometer n^+-n-n^+ device
J. of Phys. D - Applied Physics 30 1997 2343

[Col95] B.E.Cole, W.Batty, Y.Imanaka, Y.Shimamoto, J.Singleton, J.M.Chamberlain, N.Miura, M.Henini and T.Cheng
Effective mass anisotropy in GaAs-(Ga,Al)As two-dimensional hole systems: comparison of theory and very high-field cyclotron resonance experiments
Jour.of Physics -Cond.Matt. 7 1995 L675

[Col96] B.E.Cole, S.O.Hill, Y.Imanaka, Y.Shimamoto, W.Batty, J.Singleton, J.M.Chamberlain, N.Miura, M.Henini, T.Cheng
Effective mass anisotropy and many-body effects in 2D GaAs/(Ga,Al)As hole gases observed in very high magnetic fields: comparison of theory and experiment
Surf.Sc. 361/362 1996 464

[Cor69] J.F.Cornwell
Selected topics in solid state physics, vol. X. (edited by E.P.Wohlfarth)
Group theory and electronic energy bands in solids
North-Holland publishing 1969 Amsterdam

[Dae91] H.Daembkes
Modulation-Doped Field Effect Transistors: Principles, Design and Technology
IEEE Press 1991 New York

[Dal98] R.van Dalen and P.N.Stavrinou
General rules for constructing valence band effective mass Hamiltonians with correct operator order for heterostructures with arbitrary orientations
Semicon.Sc.and Tech.13 1998 11

[Dat91] Data in Science and Technology
Semiconductors, group IV elements and III-V compounds
Springer Verlag 1991 Berlin

[Din74] R.Dingle, W.Wiegmann, C.H.Henry
Quantum states of confined carriers in very thin $Al_xGa_{1-x}As$-GaAs- $Al_xGa_{1-x}As$ heterostructures
Phys.Rev.Lett.33 1974 827

[Dre55a] G.Dresselhaus, A.F.Kip and C.Kittel
Cyclotron resonance of electrons and holes
Phys.Rev.98 1955 368

[Dre55b] G.Dresselhaus
Spin-orbit coupling effects in zinc blende structures
Phys.Rev.100 1955 580

[Dru86] T.J.Drummond, T.E.Zipperian, I.J.Fritz, J.E.Schirber, T.A. Plut
P-channel, strained quantum well, field-effect transistor
Appl.Phys.Lett.49 1986 461

[Edw94] G.Edwards, E.C.Valdares and F.W.Sheard
Hole subband states of GaAs/$Al_xGa_{1-x}As$ qunatum wells within the 6x6 Luttinger model
Phys.Rev.B 50 1994 8493

[Ehr57] H.Ehrenreich
Electron scattering in InSb
J.Phys.Chem.Sol. 2 1957 131

[Eke87] U.Ekenberg
Enhancement of nonparabolicity effects in a quantum well
Phys.Rev.B 36 1987 6152

[Ell54] R.J.Elliot
Theory of the effect of spin-orbit coupling on magnetic resonance in some semiconductors
Phys.Rev.96 1954 266

[Fai94] J.Faist, F.Capasso, D.L.Sivco, A.L.Hutchinson, C.Sirtori, S.N.G.Chu and A.Y.Cho
Quantum cascase laser: Temperature dependence of the performance characteristics and high T_0 operation
Appl.Phys.Lett.65 1994 2901

[Fan66] F.F.Fang and W.E.Howard
Negative field-effect mobility on (100) Si-surfaces
Phys.Rev.Lett.16 1966 797

[Fen96] M.Feng, P.J.Apostolakis and W.H.Chang
GaAs MESFET: discrete, power and MMIC microwave devices
Properties of Galium Arsenide, third edition
EMIS data review series 16 (INSPEC) (edited by R.Brozel and G.E.Stillman)
Short run press 1996 Exeter

[Fle90] R.Fletcher, E.Zaremba,M.D'Iorio, C.T.Foxon, J.J.Harris
Persistant photoconductivity and two-band effects in GaAs/$Al_xGa_{1-x}As$ heterojunctions
Phys.Rev.B 41 1990 10649

[For93] B.A.Foreman
Effective-mass Hamiltonian and boundary conditions for the valence bands of semiconductor microstructures
Phys.Rev.B 48 1993 4964

[For96] B.A.Foreman
Envelope-function formalism for electrons in abrubt heterostructures with material-dependent basis functions
Phys.Rev.B 54 1996 1909

[Fre96] J.Freyer and M.Claasen
GaAs IMPATT diodes
Properties of Galium Arsenide, third edition
EMIS data review series 16 (INSPEC) (edited by R.Brozel and G.E.Stillman)
Short run press 1996 Exeter

[Fri86] I.J.Fritz, T.J.Drummond, G.C.Osbourn, J.E.Schirber and E.D.Jones
Electrical transport of holes in GaAs/InGaAs/GaAs single strained quantum wells
Appl.Phys.Lett.48 1986 1678

[Gas74] S.Gasiorowicz
Quantum Physics
Wiley 1974 New York

[Ger91] D.Gershoni, G.A.Baraff, S.N.G.Chu, L.N.Pfeiffer, K.West
Anisotropic optical properties of (110)-oriented quantum wells
Phys.Rev.B 44 1991 1930

[Gol95] G.Goldoni and F.M.Peeters
Hole subbands and effective masses in p-doped [113]-grown heterojunctions
Phys.Rev.B.51 1995 17806

[Gon94] L.R.González, J.Krupski, T.Szwacka
Low-temperature electron mobility in a δ-doped semiconductor
Phys.Rev.B 49 1994 11111

[Hae91] W. van Haeringen
Lecture notes of the course 'Quantum theory van de vaste stof deel 1'
University of Eindhoven, 1991

[Hai89] D.Haigh and J.Everard
GaAs technology and its impact on circuits and systems
Peter Peregrinus Ltd. 1989 London

[Hai95] G.Q.Hai, N.Studart, F.M.Peeters
Multiband electron transport in δ-doped semiconductor systems
Phys.Rev.B 52 1995 8363

[Hai96] G.Q.Hai, N.Studart, F.M.Peeters, P.M.Koenraad, J.H.Wolter
Intersubband-coupling and screening effects on the electron transport in a quasi-two-dimensional δ-doped semiconductor system
J.Appl.Phys.80 1996 5809

[Hai97] G.Q.Hai and N.Studart
Quantum transport in δ-doped quantum wells
Phys.Rev.B 55 1997 186

[Har89] J.J.Harris, J.A.Pals, R.Woltjer
Electronic transport in low-dimensional structures
Rep.Prog.Phys.52 1989 1217

[Hay94] R.K.Hayden, L.Eaves, M.Henini, E.C.Valdares, O.Kühn, D.K.Maude, J.C.Portal, T.Takamasu, N.Miura and U.Ekenberg
Probing the anisotropic dispersion of hole states in (100) and (311)A AlAs/GaAs/AlAs quantum wells
Semi.Sc.&Tech. 9 1994 298

[Hed71] L.Hedin and B.I.Lundqvist
Explicit local exchange-correlation potentials
J.Phys.C 4 1971 2064

[Hen95] R.H.Henderson and E.Towe
Strain and crystallographic orientation effects on interband optical matrix elements and band gaps of [11*l*]-orientated III-V epilayers
J.Appl.Phys.78 1995 2447

[Hin90] J.M.Hinckley and J.Singh
Influence of substrate composition and crystallographic orientation on the band structure of pseudomorphic Si_Ge alloy films
Phys.Rev.B 42 1990 3546

[Hir84] K.Hirakawa, H.Sakaki and J.Yoshita
Concentration of electrons in selectively doped GaAlAs/GaAs heterojunction and its dependence on spacer-layer thickness and gate electric field
Appl.Phys.Lett.45 1984 253

[Hit95] L.J.Hitchens, P.A.Houston, M.Hopkinson and G.J.Rees
Enhanced mobility piezoelectric AlInAs/InGaAs quantum well structures on (111)B InP structures
Electr.Lett.31 1995 2215

[Hon96] B.W.P.Hong
Delta-doped channel III-V field effect transistors (FETs)
Delta doping of semiconductors (edited by E.F.Schubert)
Cambridge University Press 1996 Cambridge

[Hoo89] C.van Hoof, K.Deneffe, J.DeBoeck, D.J.Arent, G.Borghs
Franz-Keldysh oscillations originating from a well-controlled electric field in the GaAs depletion region
Appl.Phys.Lett.54 1989 608

[Hsu95] R.T.Hsu, W.C.Hsu, J.S.Wang, M.J.Kao, Y.H.Wu, J.S.Su
Study of zinc-delta-doped strained-quantum-well InGaAs/GaAs p-channel heterostructure field-effect transistors
Jap.J.of.Appl.Phys. part 1 35 1996 2085

[Iko92] Z.Ikonic, V.Milanovic, D.Tjapkin
Valence subband structure of [100]-, [110]-, and [111]-grown GaAs-(Al,Ga)As quantum wells and the accuracy of the axial approximation
Phys.Rev.B 46 1992 4285

[Jaf87] M.Jaffe, Y.Sekiguchi and J.Singh
Theoretical formalism to understand the role of strain in the tailoring of hole masses in p-type $In_xGa_{1-x}As$ (on GaAs substrates) and $In_{0.53+x}Ga_{0.47-x}As$ (on InP substrates) modulation-doped field-effect transistors
Appl.Phys.Lett.51 1987 1943

[Jas97] R.Jaszek
Electrical properties of modulation doped InAs quantum well heterostructures.
Doctoral thesis 1997 Imperial College London

[Kaj95] Y.Kajikawa
Well-width dependence of the optical anisotropies in (001) and (110) semiconductor quantum wells: The effect of the spin-orbit split-off bands
Phys.Rev.B 51 1995 51

[Kan56] E.O.Kane
Energy band structure in p-type germanium and silicon.
J.Phys.Chem.Solids 1 1957 82

[Kan57] E.O.Kane
Bandstructure of Indium antimonide
J.Phys.Chem.Solids 1 1957 249

[Kan66] E.O.Kane
The k·p method
Semiconductors and semimetals, vol.1 (edited by R.K.Willardson and A.C.Beer)
Academic press 1966 Orlando

[Kas94] J.A.Kash, M.Zachau, M.A.Tischler, U.Ekenberg
Optical measurements of warped valence bands in quantum wells
Surf.Science 305 1994 251

[Kem96] M.Kemerink, P.M.Koenraad and J.H.Wolter
Effect of strain on second-order van Hove singularity in $Al_xGa_{1-x}As/In_yGa_{1-y}As$ quantum wells
Phys.Rev.B 54 1996 10644

[Kie84] R.A.Kiehl, A.C.Gossard
P-channel (Al,Ga)As/GaAs modulation-doped logic gates
IEEE Elec.Dev.Lett. 5 1984 420

[Kie88] R.A.Kiehl, S.Tiwari, S.L.Wright, M.A.Olson
P-channel quantum-well heterostructure misfet
IEEE Elect.Dev.Lett.9 1988 309

[Kit86] C.Kittel
Introduction to solid state physics, sixth edition
Wiley 1986 New York

[Kre88] E.Kreyszig
Advanced Engineering Mathematics, sixth edition
Wiley 1988 New York

[Kud96] M.Kudo, T.Mishima, H.Matsumoto, I.Ohbu and T.Tanimoto
Highly strained $In_{0.35}Ga_{0.65}As$/GaAs layers grown by molecular beam epitaxy for high hole mobility transistors
Jour.Elec.Mat.25 1996 944

[Kud97] M.Kudo, H.Matsumoto, T.Tanimoto, T.Mishima, I.Ohbu
Improved hole transport properties of highly strained $In_{0.35}Ga_{0.65}As$ channel double-modulation-doped structures grown by MBE on GaAs
Jour.Cryst.Growth 175/176 1997 910

[Lac88] D.E.Lacklison, J.J.Harris, C.T.Foxon, J.Hewett, D.Hilton and C.Roberts
A comparison of photoconduction effects in (Al,Ga)As and GaAs/(Al,Ga)As heterostructures
Semicon.Sci.Techn.3 1988 633

[Lan82] Landolt-Börnstein, vol. 17a
Numerical data and functional relationships in science and technology
Springer-Verlag 1982 Berlin

[Lan87] Landolt-Börnstein, vol. 22
Numerical data and functional relationships in science and technology
Springer-Verlag 1987 Berlin

[Lau71] L.D.Laude, F.H.Pollak and M.Cardona
Effects of uniaxial stress on the indirect exciton spectrum of silicon
Phys.Rev.B 3 1971 2623

[Lee83a] J.Lee, H.N.Spector and V.K.Aurora
Quantum transport in a single layered structure for impurity scattering
J.of Appl.Phys.54 1983 6995

[Lee83b] H.J Lee and D.C.Look
Hole transport in pure and doped GaAs
J.Appl.Phys.54 1983 4446

[Lee90] C.D.Lee and S.R.Forrest
Effects of lattice mismatch on $In_xGa_{1-x}As$/InP heterostructures
App.Phys.Lett.57 1990 468

[Liu88] C.T.Liu, S.Y.Lin and D.C.Tsui
Cyclotron resonance measurements of electron effective mass in strained AlGaAs/InGaAs/GaAs pseudomorphic structures
Appl.Phys.Lett.53 1988 2510

[Löw51] P.O.Löwdin
A note on the Quantum-Mechanical perturbation theory
Jour.Chem.Phys.19 1951 1396

[Lu94] S.S.Lu and C.L.Huang
Piezoelectric field effect transistor (PEFET) using $In_{0.2}Ga_{0.8}As$/GaAs strained layer structure on (111)B GaAs substrate
Electr.Lett.30 1994 823

[Lut55] J.M.Luttinger and W.Kohn
Motion of electrons and holes in perturbed periodic fields
Phys.Rev.97 1955 869

[Lut56] J.M.Luttinger
Quantum theory of cyclotron resonance in semiconductors: General theory
Phys.Rev.102 1956 1030

[Mah90] G.D.Mahan
Many particle physics, second edition
Plenum Press 1990 New York

[Mas91] W.T.Masselink
Ionised-impurity scattering of quasi-two-dimensional qunatum confined carriers
Phys.Rev.Lett.66 1991 1513

[Men92] A.T.Meney
Orientation dependence of subband structure and optical properties in GaAs-AlGaAs quantum wells: [001], [111], [110] and [310] growth directions
Superlat.&Microstruc.11 1992 31

[Men94] A.T.Meney and B.Gonul
Evaluation of various approximations used in the envelope function method
Phys.Rev.B 50 1994 10893

[Mor84] R.A.Morrow and K.R.Brownstein
Model effective mass Hamiltonians for abrupt heterojunctions and associated wavefunction matching conditions
Phys.Rev.B 30 1984 678

[Noj92] S.Nojima
Anisotropical optical transitions in semiconductor quantum wells
Jpn.J.Appl.Phys. 31 1992 L765

[Noj93] S.Nojima
Anisotropy of optical transistions in (110)-orientated quantum wells
Phys.Rev.B 47 1993 13535

[Pas93] M.D.Pashley, K.W.Haberern, R.M.Feenstra, P.D.Kirchner
Different Fermi-level pinning behaviour on n-type and p-type GaAs (001)
Phys.Rev.B 48 1993 4612

[Pet58] R.L.Petritz
Theory of an experiment for measuring the mobility and density of carriers in the space-charge region of a semiconductor surface
Phys.Rev.110 1958 1254

[Pet90] T.L.Peterson, F.Szmulowics and P.M.Hemenger
Electrical transport analysis of ultra-pure silicon
J.of Cryst.Growth 106 1990 16

[Pfe96] P.Pfeffer and W.Zawadzki
Five-level k·p model for the conduction and valence bands of GaAs and InP
Phys.Rev.B 53 1996 12813

[Pik74] G.L.Pikus and G.E.Bir
Symmetry and strain induced effects in semiconductors
Wiley 1974 New York

[Rei90] J.P.Reithmaier, R.Höger and H.Riechert
Confinement of light hole valence-band states in psuedomorphic InGaAs/Ga(Al)As quantum wells
Appl.Phys.Lett.57 1990 957

[Rid93] B.K.Ridley
Quantum processes in Semiconductors, third edition
Clarendon Press 1993 Oxford

[Rob96] J.M.Roberts, J.J.Harris, L.Hart, M.Hopkinson, C.Roberts
Mobility and saturation drift velocity enhancement in highly doped GaAs and $In_xGa_{1-x}As$ structures designed for use in power FET devices
Electr.Lett.32 1996 494

[Rob98] J.M.Roberts, J.J.Harris, N.J.Woods, M.Hopkinson
Investigation of delta-doped quantum wells for power FET applications
Superlat.and Microstruc. 23 1998 187

[Ros84] U.Rössler
Nonparabolicity and warping in the conduction band of GaAs
Solid State Comm.49 1984 943

[Rot92] L.M.Roth
Transport of carriers in semiconductors
Handbook on semiconductors, vol.1 (edited by T.S.Moss)
North Holland 1992 Amterdam

[Sch85a] J.N.Schulman and Y.C.Chang
Band mixing in semiconductor superlattices
Phys.Rev.B.31 1985 2056

[Sch85b] M.F.H.Schuurmans and G.W.'t Hoofd
Simple calculations of confinement states in a quantum well
Phys.Rev.31 1985 8041

[Sch96] E.F.Schubert
Delta-doping of semiconductors
Cambridge University Press 1996 Cambridge

[She95a] G.Shechter, L.D.Shvartsman, J.E.Golub
Analysis of the orientational effects on infrared absorbtion spectra in p-type semiconductor quantum wells
Superlat.&Microstr. 19 1996 383

[She95b] H.Shen and M.Dutta
Franz-Keldysh oscillations in modulation spectroscopy
J.Appl.Phys.78 1995 2151

[Sig70] E.D.Siggia and P.C.Kwok
Properties of electrons in semiconductor inversion layers with many occupied electric subbands. I. Screening and impurity scattering
Phys.Rev.B 2 1970 1024

[Sin93] J.Singh
Physics of Semiconductors and their heterostructures
McGraw-Hill 1993 New York

[Smi86] D.L.Smith
Strain generated electric fields in [111] growth axis strained layer superlattices
Solid State Comm.57 1986 919

[Sta97] P.N.Stavrinou and R. van Dalen
Operator ordering and boundary conditions for valence band modeling: Application to [110] heterostructures
Phys.Rev.B 55 1997 15456

[Ste67] F.Stern and W.E.Howard
Properties of semiconductor surface inversion layers in the electrical quantum limit
Phys.Rev.163 1967 816

[Ste84] F.Stern and S.Das Sarma
Electron energy levels in GaAs-$Ga_{1-x}Al_xAs$ heterojunctions
Phys.Rev.B 30 1984 840

[Str90] B.G.Streetman
Solid State Electronic Devices, 3th edition
Prentice-Hall 1990 New York

[Sun94] D.Sun and E.Towe
Strain-generated internal fields in pseudomorphic (In,Ga)As/GaAs qunatum well structures on {111} GaAs substrates
Jpn.J.Appl.Phys. part 1 33 1994 702

[Suz96] K.Suzue, S.N.Mohammed and H.Morkoc
Properties of III-V quantum wells and superlattices
EMIS data review series 15 (INSPEC) (edited by P.Bhattacharya)
Short run press 1996 Exeter

[Sze66] S.M.Sze and G.Gibbons
Avalanche breakdown voltages of abrupt and graded p-n junctions in Ge, Si, GaAs and GaP
Appl.Phys.Lett. 8 (1966) 111

[Sze81] S.M.Sze
Physics of semiconductor devices, 2nd edition
Wiley 1981 New York

[Tak85] K.Takeda, N.Matsumoto, A.Taguchi,H.Taki,E.Ohta and M.Sakata
Hole mobility of GaAs GaP and $GaAs_{1-x}P_x$ mixed compound semiconductors
Phys.Rev.B 32 1985 1101

[Tat97] K.Tateno, Y.Ohiso, C.Amano, A.Wakatsuki, T.Kurokawa
Growth of vertical-cavity surface-emitting laser structures on GaAs (311)B substrates by metalorganic chemical vapor deposition
Appl.Phys.Lett.70 1997 3395

[Tho93] J.L.Thobel, L.Baudry, F.Desenne, M.Charef, R.Fauquembergue
Theoretical investigation of impurity scattering limited mobility in quantum wells: The influence of wave function modeling
J.Appl.Phys. 73 1993 233

[Tsi95] E.G.Tsiisishvili
Intrinsical optical anisotropy of quantum wells in cubic crystals
Phys.Rev.B 52 1995 11172

[Tya91] M.S.Tyagi
Introduction to semiconductor materials and devices
Wiley 1991 New York

[Wal89] C.G.van der Walle
Band lineups and deformation potentials in the model-solid theory
Phys.Rev.B 39 1989 1871

[Wat98] J.R.Watling, A.B.Walker, J.J.Harris, J.M.Roberts
Monte Carlo simulation of electron transport in highly delta-doped GaAs/AlGaAs quantum wells
Semicon.Sci.and Tech.13 1998 43

[Wei91] C.Weisbuch and B.Vinter
Quantum Semiconductor Structures, fundamentals and applications
Academic Press 1991 San Diego

[Wen94] X.M.Weng, X.L.Lei
Electron transport in non-parabolic Kane bands
Journ. of Phys. Cond. Matt.6 1994 6287

[Wet96] C.Wetzel, R.Winkler, M.Drechsler, B.K.Meyer, U.Rössler, J.Scriba, J.P.Kotthaus, V.Härle, F.Scholz
Electron effective mass and nonparabolicity in $Ga_{0.47}In_{0.53}As/InP$ quantum wells
Phys.Rev.B 53 1996 1038

[Wil71] J.D.Wiley
Polar mobility in III-V compounds
Phys.Rev.B. 4 1971 2485

[Win96] R.Winkler, A.I.Nesvizhskii
Anisotropic hole subband states and interband optical absorption in [mmn]-orientated quantum wells
Phys.Rev.B 53 1996 9984

[Yan94] D.Yan, E.Look, X.Yin, F.H.Pollak, J.M.Woodall
Air stabilized (001) p-type GaAs fabricated by molecular beam epitaxy with reduced surface state density
Appl.Phys.Lett.65 1994 186

[Ye87] L.X.Ye, R.A.Stewart, J.N.Churchill
On the effective mass in modeling nonstationairy transport
J.of Appl.Phys.61 1987 685

[Yin92] X.Yin, H.M.Chen, F.H.Pollak, Y.Chan, P.A.Montano, P.D.Kirchner, G.D.Pettit, J.M.Woodall
Photoreflectance study of the surface fermi level at (001) n-type and p-type GaAs-surfaces
Jour.Vac.Sci.and Tech. 10 1992 131

[Yok86] K.Yokohama and K.Hess
Monte Carlo studyof electronic transport in $Al_{1-x}Ga_xAs/GaAs$ single well structures
Phys.Rev.B 33 1986 5595

[Zaw82] W.Zawadzki
Mechanisms of electron scattering in semiconductors
Handbook on semiconductors, vol.1 (edited by T.S.Moss)
North Holland 1982 Amsterdam

[Zip88] T.E.Zipperian, L.R.Dawson, T.J.Drummond, J.E.Schirber, I.J.Fritz
GaAs/(In,Ga)As, p-channel, multiple strained quantum well field effect transistors with high transconductance and high peak saturated drain current
Appl.Phys.Lett.52 1988 975

[Zol96] S.Zollner and M.Cardona
Intra- and intervalley deformation potentials for electrons in GaAs
Properties of Galium Arsenide, third edition
EMIS data review series 16 (INSPEC) (edited by R.Brozel and G.E.Stillman)
Short run press 1996 Exeter

[Zor93] P.S.Zory, jr.
Quantum well lasers
Academic Press 1993 San Diego

Acknowledgements

London, 21-7-'98

Well, this is it then. Before you lies the result of over 3 years work performed at the IRC for Semiconductor Materials in London. However, this has not been a solo project. Many people have contributed, either directly through *deep* scientific discussions or simply by making Imperial College and life more pleasant. I would like to extend my thanks to the following people,

First of all, my supervisor Jeff Harris for allowing me to join the transport group, his work to actually get me there, and foremost, to teach me semiconductor physics !

To the former direction of the IRC, Prof.B.A.Joyce, and the new director of the centre, Prof.G.Parry, for making my stay possible.

Paul Stavrinou, for his collaboration on the subject of the boundary conditions and the discussions on effective mass theory. We should write some more papers !

Christine Roberts, for the fabrication of my hole samples. Sorry for making you grow my rather unglamorous structural designs that were bound to have a mobility of less than 50 cm^2/Vs !

Paul Koenraad of the University of Eindhoven (The Netherlands), for helpful discussions on the RPA screening model.

Rys Jaszek and Nick Woods, for having lunch with me. Caroline Bryan, for cheering me up during the writing of the thesis. Pippa Steans, for drinking my coffee. Jing Zhang, for $2\pi/a$ and generally being right (damn it). Ray Murray, for being Scottish. Mark Whitehead, Phil Siverns, Surama Malik, Jim Neave, Eng Soon Tok, Anna Lees, Stuart Holmes, Tillman Sander, Hai Ping Yu, Gesine Grosche, Sam Lycett and many more that made life at the IRC more pleasant.

To all my friends, you know who you are ! Special thanks go out to Gosia, for being a special friend... And last but not least, to my parents...

Publications

[Sta97] P.N.Stavrinou and R. van Dalen
Operator ordering and boundary conditions for valence band modeling: Application to [110] heterostructures
Phys.Rev.B 55 1997 15456

[Dal98] R.van Dalen and P.N.Stavrinou
General rules for constructing valence band effective mass Hamiltonians with correct operator order for heterostructures with arbitrary orientations
Semicon.Sc.and Tech.13 1998 11

Operator ordering and boundary conditions for valence-band modeling: Application to [110] heterostructures

Paul N. Stavrinou
Interdisciplinary Research Centre for Semiconductor Materials, Department of Engineering Science, University of Oxford, Parks Road, Oxford OX1 3PU, United Kingdom

Rob van Dalen
Interdisciplinary Research Centre for Semiconductor Materials, Imperial College, Prince Consort Road, London SW7 2BZ, United Kingdom

(Received 5 December 1996)

A general method to derive the operator ordering and boundary conditions within the framework of the recently developed exact envelope-function theory is presented. An ordered form of the familar Luttinger-Kohn Hamiltonian is derived which can be regarded as the natural starting point for calculations involving heterostructures, such that transforming with an appropriate angular momentum basis, the derived ordering provides an unambiguous prescription for the boundary conditions across the interface. As an example, specific expressions are derived for structures with confinement along the [110] direction where the resulting boundary conditions are found to differ significantly from those currently used in the literature. The effect on the calculated in-plane dispersions is examined for two material systems, where it is found the greatest differences occur in systems with a large difference in Luttinger parameters. Typically, the results show that employing the traditional boundary conditions tends to exaggerate any negative curvature of the valence bands near the zone center. [S0163-1829(97)06724-6]

Effective-mass theory is a powerful and convenient method for obtaining valence-band descriptions of heterostructures. In such an approach, the wave function is expanded in a number of zone-center solutions of the original bulk Hamiltonian close to the energy region of interest, each multiplied by a so-called envelope function.[1–3] Rather than solving for the wave function itself, one now solves for these envelope functions. The question of how these envelopes connect across a boundary has now been settled by the development of an exact envelope-function theory.[3] Based on this work, boundary conditions were derived specifically for structures confined along [001], and found to differ from those conventionally used.[4] Moreover it was shown that application of these boundary conditions can result in significant changes to the band structure.[4,5] Following Burt's approach, we present a starting valence-band Hamiltonian applicable to heterostructures from which boundary conditions for any growth direction can be obtained. As an example, boundary conditions for [110] confinement are derived, and found to differ considerably from those currently used in the literature.

Ignoring linear k terms, the valence-band maximum in bulk GaAs and related materials occurs at the Γ-point, i.e., $\mathbf{k}=\mathbf{0}$. Neglecting spin-orbit splitting, the symmetry of the zone-center valence-band wave functions is described by bonding p functions in the tight-binding picture, and corresponds to the symmetry Γ_{15}. As usual, these valence-band zone-center wave functions are denoted by $|X\rangle$, $|Y\rangle$, and $|Z\rangle$. One then assumes that an expansion using just these states, and accounting for the excluded (remote) states by perturbation theory, is sufficient to describe the valence bands at nonzero $\mathbf{k}$. Treating the expansion in such a way modifies the Hamiltonian by adding an extra interaction term, the so-called Löwdin interaction.[6] During the development of the exact envelope-function theory, Burt showed that for heterostructures this interaction should be replaced by a version in which the differential operators $\hat{k}$ have a specific ordering

$$H_{jj'}^{\text{int}}=(\hbar/m_0)^2\sum_{\alpha,\beta=x,y,z}\hat{k}_\alpha\left(\sum_\nu\frac{\langle j|\hat{p}_\alpha|\nu\rangle\langle\nu|\hat{p}_\beta|j'\rangle}{\varepsilon-\varepsilon_\nu}\right)\hat{k}_\beta, \tag{1}$$

where the indices j and j' run over the expansion states $|X\rangle$, $|Y\rangle$, and $|Z\rangle$ while ν refers to the remote states. In the notation adopted by Foreman, the interactions between the remote states and the p-type valence-band states may be described using the following quantities which identify the individual contributions of the remote states of symmetry Γ_1, Γ_{15}, and Γ_{12} (neglecting the small contribution of Γ_{25}):

$$\begin{aligned}\sigma&=(1/3m_0)\sum_\nu^{\Gamma_1}|\langle X|\hat{p}_x|\nu\rangle|^2/(\varepsilon_\nu-\varepsilon),\\ \pi&=(1/3m_0)\sum_\nu^{\Gamma_{15}}|\langle X|\hat{p}_y|\nu\rangle|^2/(\varepsilon_\nu-\varepsilon),\\ \delta&=(1/6m_0)\sum_\nu^{\Gamma_{12}}|\langle X|\hat{p}_x|\nu\rangle|^2/(\varepsilon_\nu-\varepsilon),\end{aligned} \tag{2}$$

where ε is the valence-band energy in the absence of spin-orbit splitting. These expressions may be visualised by considering a tight-binding view. The lowest atomic orbitals

compatible with Γ_1 and Γ_{12} (respectively, s and $d_{x^2-y^2}$ and $d_{3z^2-r^2}$) are even functions with respect to x, y, and z; therefore the only nonzero inner products are of the form $\langle X|\hat{p}_x|v\rangle$. In contrast, the lowest Γ_{15} states to make a contribution are d_{xy}, d_{yz}, and d_{zx}, as a result the only non-zero inner products are of the form $\langle X|\hat{p}_y|v\rangle$.

For each layer, an effective-mass Hamiltonian over $|X\rangle$, $|Y\rangle$, and $|Z\rangle$ bases can be written as

$$H=\left(\frac{\hbar^2}{2m_0}\right)\begin{bmatrix} \hat{k}_xA\hat{k}_x+\hat{k}_yB\hat{k}_y+\hat{k}_zB\hat{k}_z & \hat{k}_xC_1\hat{k}_y-\hat{k}_yC_2\hat{k}_x & \hat{k}_xC_1\hat{k}_z-\hat{k}_zC_2\hat{k}_x \\ \hat{k}_yC_1\hat{k}_x-\hat{k}_xC_2\hat{k}_y & \hat{k}_yA\hat{k}_y+\hat{k}_zB\hat{k}_z+\hat{k}_xB\hat{k}_x & \hat{k}_yC_1\hat{k}_z-\hat{k}_zC_2\hat{k}_y \\ \hat{k}_zC_1\hat{k}_x-\hat{k}_xC_2\hat{k}_z & \hat{k}_zC_1\hat{k}_y-\hat{k}_yC_2\hat{k}_z & \hat{k}_zA\hat{k}_z+\hat{k}_xB\hat{k}_x+\hat{k}_yB\hat{k}_y \end{bmatrix}, \tag{3}$$

where

$$A=1-6\sigma-12\delta, \quad B=1-6\pi,$$

$$C_1=6\delta-6\sigma, \quad C_2=6\pi.$$

The diagonal potential-energy term, describing the band edges, has been omitted for clarity. It is important to note that the ordering of the differential operators above follows naturally from Eq. (1) and the symmetry properties of the zinc-blende lattice. For a homogenous structure, Eq. (3) operates on plane-wave envelope functions. The differential operators then reduce to components of the Bloch wave vector $\hat{k}_i \rightarrow k_i$, and the ordering of the terms becomes redundant, such that the contributions from the different symmetries are often lumped together into one constant. In this case, Eq. (3) can take the form of the familiar Luttinger-Kohn (LK) Hamiltonian[1] for the same basis. [Note: If the contribution of Γ_{25} had been included, then Eq. (3) would reduce exactly to the LK Hamiltonian.] In heterostructures, it is customary to start with the LK Hamiltonian, which correctly describes the bulk band structure of the constituent layers, and *then* introduce the operators. Clearly, with the off-diagonal terms represented by one constant, the specific ordering as observed in Eq. (3) is not immediately apparent. As we see later, this ordering is crucial when the question of boundary conditions arises and replaces the need to symmetrize the Hamiltonian.

When looking at [001] structures, it is customary to choose $\hat{k}_x$ and $\hat{k}_y$ perpendicular to the growth direction, leaving $\hat{k}_z$ to act along the confinement direction. For different confinement directions one would like to introduce a new reference frame $(\hat{k}_{x'}, \hat{k}_{y'}, \hat{k}_{z'})$ with a similar orientation as in the [001] case, i.e., the z' component being parallel to the confinement direction. Along [110], the new axes are respectively taken along $(00\bar{1})$, $(\bar{1}10)$, and (110). It is important to note that by nature of its derivation, Eq. (3) is valid for any growth condition, and one can simply replace the old operators in terms of their projections in the new frame. Therefore, Eq. (3) is the natural starting point for calculations involving heterostructures.

Including the effects of spin-orbit interaction merely consists of forming a new basis consisting of linear combinations of the $|X\rangle$, $|Y\rangle$, and $|Z\rangle$ states which are eigenfunctions of the spin-orbit operator. Such a basis can be found by recognizing that eigenfunctions of the spin-orbit operator are also eigenfunctions of $\mathbf{J}^2$, with $\mathbf{J}$ representing the total angular momentum. Since $\mathbf{J}$ does not uniquely classify the states, one can introduce a new commuting observable which we choose to be its projection $J_{z'}$ along the desired confinement direction z'. As usual, the states will be labeled $|J,m_j\rangle$, where m_j is the quantum number corresponding to $J_{z'}$. Reorienting the spin-up and -down states along the growth direction, we find the following set describing the heavy ($J_{z'}=\pm 3/2$) and light hole ($J_{z'}=\pm 1/2$) states at the Γ point for [110]-oriented heterostructures,

$$\begin{aligned} |\tfrac{3}{2},+\tfrac{3}{2}\rangle &= (i/2)|(X-Y)\uparrow\rangle+(1/\sqrt{2})|Z\uparrow\rangle, \\ |\tfrac{3}{2},+\tfrac{1}{2}\rangle &= (i/\sqrt{3})|(X+Y)\uparrow\rangle-(1/\sqrt{12})|(X-Y)\downarrow\rangle \\ &\quad +(i/\sqrt{6})|Z\downarrow\rangle, \\ |\tfrac{3}{2},-\tfrac{1}{2}\rangle &= -(1/\sqrt{3})|(X+Y)\downarrow\rangle-(i/\sqrt{12})|(X-Y)\uparrow\rangle \\ &\quad +(1/\sqrt{6})|Z\uparrow\rangle, \\ |\tfrac{3}{2},-\tfrac{3}{2}\rangle &= (1/2)|(X-Y)\downarrow\rangle+(i/\sqrt{2})|Z\downarrow\rangle. \end{aligned} \tag{4}$$

Such a set is not uniquely defined, as each state can be multiplied by a phase factor of unit length. The above basis set differs slightly from that previously reported for this orientation,[7] and was chosen to facilitate comparisons with previous work on [110] structures.[8]

The wave function is now written in terms of the basis functions (4), each multiplied by an envelope function

$$|\Psi\rangle=f_1|\tfrac{3}{2},+\tfrac{3}{2}\rangle+f_2|\tfrac{3}{2},+\tfrac{1}{2}\rangle+f_3|\tfrac{3}{2},-\tfrac{1}{2}\rangle+f_4|\tfrac{3}{2},-\tfrac{3}{2}\rangle. \tag{5}$$

Following Burt's approach, the basis sets in Eq. (4) are taken to be the same throughout the heterostructure, regardless of composition. Using the above expansion, the new reference frame k' and Eq. (3), the effective-mass Hamiltonian for [110] confinement, takes the form

$$\mathbf{H}=-\frac{\hbar^2}{2m_0}\begin{bmatrix} P+Q & L_- & M & 0 \\ L_-^\dagger & P-Q & C & M \\ M^\dagger & C^\dagger & P-Q & -L_+^\dagger \\ 0 & M^\dagger & -L_+ & P+Q \end{bmatrix}+E_v(z')\mathbf{I}, \tag{6a}$$

where

$$P=\gamma_1(k_{x'}^2+k_{y'}^2)+\hat{k}_{z'}\gamma_1\hat{k}_{z'},$$

$$Q=(\gamma_2 k_{x'}^2-\tfrac{1}{2}\gamma_2 k_{y'}^2-\tfrac{1}{2}\hat{k}_{z'}\gamma_2\hat{k}_{z'})+\tfrac{3}{2}(\gamma_3 k_{y'}^2-\hat{k}_{z'}\gamma_3\hat{k}_{z'}),$$

$$M=-(\sqrt{3}/2)[(\gamma_2+\gamma_3)k_{y'}^2+\hat{k}_{z'}(\gamma_2-\gamma_3)\hat{k}_{z'}]$$
$$+\sqrt{3}(\gamma_2 k_{x'}^2-2i\gamma_3 k_{x'}k_{y'}), \quad (6b)$$

$$L_{\pm}=-\sqrt{3}k_{y'}[\hat{k}_{z'}(3\delta)+(2\sigma-2\pi+\delta)\hat{k}_{z'}]$$
$$\pm i2\sqrt{3}k_{x'}[\hat{k}_{z'}(\pi)+(\sigma-\delta)\hat{k}_{z'}],$$

$$C=2(k_{y'}+ik_{x'})[\hat{k}_{z'}(\sigma-\pi-\delta)-(\sigma-\pi-\delta)\hat{k}_{z'}].$$

The hole energy is taken as negative, $E_v(z')$ refers to the bulk heavy- and light-hole valence-band profiles, and **I** is the identity matrix. Band parameters for each layer can be found from the more commonly known Luttinger parameters through the relationships

$$\gamma_1=-1+2\sigma+4\pi+4\delta,$$

$$\gamma_2=\sigma-\pi+2\delta, \quad (7)$$

$$\gamma_3=\sigma+\pi-\delta,$$

As the operators $\hat{k}_{x'}$ and $\hat{k}_{y'}$ are oriented parallel to the interface, they have been replaced by the in-plane wave-vector components $k_{x'}$ and $k_{y'}$. In contrast, $\hat{k}_{z'}$, acting along the confinement direction, is left as a differential operator, i.e., $\hat{k}_{z'}=-i(\partial/\partial z')$.

The boundary conditions are now obtained by integrating Eq. (6) across a boundary. Recalling that only the terms of the form $\hat{k}_{z'}f(z')$ survive, we find the requirement of continuity of F and $\mathbf{D}F$ across an interface, where F is a four component vector containing the envelopes and **D** is given by

$$\mathbf{D}=\begin{bmatrix} \frac{1}{m_{hh}}\partial/\partial z' & l_1 & n\partial/\partial z' & 0 \\ l_2 & \frac{1}{m_{lh}}\partial/\partial z' & c & n\partial/\partial z' \\ n\partial/\partial z z' & c^* & \frac{1}{m_{lh}}\partial/\partial z' & l_2^* \\ 0 & n\partial/\partial z' & l_1^* & \frac{1}{m_{hh}}\partial/\partial z' \end{bmatrix}, \quad (8a)$$

where

$$n=-(\sqrt{3}/2)(\gamma_2-\gamma_3),$$

$$c=-2(k_{x'}-ik_{y'})(\sigma-\delta-\pi),$$

$$l_1=2\sqrt{3}(\pi)k_{x'}-i3\sqrt{3}(\delta)k_{y'}, \quad (8b)$$

$$l_2=-2\sqrt{3}(\sigma-\delta)k_{x'}-i\sqrt{3}(2\sigma-2\pi+\delta)k_{y'},$$

$$\frac{1}{m_{hh}}=\left(\gamma_1-\frac{1}{2}\gamma_2-\frac{3}{2}\gamma_3\right), \quad \frac{1}{m_{lh}}=\left(\gamma_1+\frac{1}{2}\gamma_2+\frac{3}{2}\gamma_3\right),$$

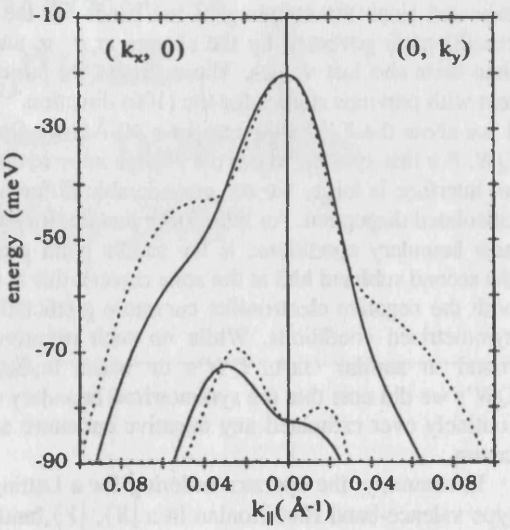


FIG. 1. In-plane dispersion of the two highest valence subbands for an $In_{0.53}Ga_{0.47}As/InP$ QW of width 40 Å with confinement along the [110] direction. Solid lines show the calculation performed using boundary conditions obtained from the ordered Hamiltonian, i.e., Eq. (8) in the present work. Also shown (dotted lines) are the results obtained using boundary conditions found from a symmetrized Hamiltonian.

Prior to the development of the exact envelope-function theory, it was necessary to symmetrize a Hamiltonian before performing the integration.[8] However, the Hamiltonian in Eq. (6), with its operator ordering found naturally within the framework of the exact envelope-function theory, shows a strong asymmetry with respect to $\hat{k}_{z'}$. It is therefore hardly surprising that, when comparing Eq. (8) to that obtained from a symmetrized Hamiltonian for this growth direction,[8] we see profound differences. The coupling of heavy- and light-hole bands of the same spin is now only through the interaction with the remote p and d states, while the traditional boundary conditions also include the larger interaction with the remote s states. Additionally we also note that the two light-hole bands are now coupled. These observations are somewhat familiar to those found in the case of [001] structures under an equivalent comparison.[4]

To evaluate the influence of the new boundary conditions, calculations were performed on two flat-band quantum-well (QW) structures. The equations were solved in an exact manner using a transfer-matrix method with the energy solutions coinciding with the complex zeros of an 8×8 determinant. In-plane dispersion diagrams $E(\mathbf{k})$ for the valence band of (110) QW structures were studied for two lattice-matched materials systems, namely, $GaAs/Al_xGa_{1-x}As$ and $In_{0.53}Ga_{0.47}As/InP$. For each structure, we applied both the new boundary conditions described here in Eq. (8) and those found from a symmetrized Hamiltonian (cf. Eq. (20) in Ref. 8). Examining only the highest two or three subbands, we found some well width dependence with the largest differences typically occurring as the well width reduced. The magnitude of the differences was certainly smaller in the $GaAs/Al_xGa_{1-x}As$ system, where the Luttinger parameters in the well and barrier materials are quite similar. This can be

expected since the coupling of the bands via the boundary conditions is governed by the change in σ, π, and δ rather than their absolute values. These results are largely consistent with previous studies for the (100) direction.[4,5] In Figure 1 we show the $E(\mathbf{k})$ diagram for a 40-Å $In_{0.53}Ga_{0.47}As/InP$ QW. For this system, where the change in σ, π, and δ across an interface is large, we see considerable differences in the calculated dispersion. An interesting feature, found using the new boundary conditions, is the saddle point predicted for the second subband hh2 at the zone center); this is in contrast with the negative electronlike curvature predicted using the symmetrized conditions. While no such occurrences were found in similar GaAs QW's or wider $In_xGa_{1-x}As/InP$ QW's we did note that the symmetrized boundary conditions routinely over estimated any negative curvature at the zone centre.

In summary, the operator ordering for a Luttinger-Kohn-type valence-band Hamiltonian in a $|X\rangle$, $|Y\rangle$, and $|Z\rangle$ basis has been presented. The nature of the ordering, which is derived within the framework of the exact envelope-function theory, has significant consequences on the boundary conditions that arise in heterostructure calculations. Starting from this Hamiltonian, specific expressions were derived for heterostructures with confinement along the (110) direction. The subsequent boundary conditions show the usual coupling between heavy-hole and light-hole (LH) bands across an interface should not involve the large interactions from remote bands of s symmetry. In addition to this we note the appearance of a term coupling LH states of opposite spin. Both of these features represent a significant departure from the previously adopted boundary conditions for this direction.

The authors gratefully acknowledge useful correspondence and discussions with Mike Burt and Bradley Foreman. We would also like to thank Gareth Parry and Jeff Harris from the ESPRC funded Semiconductor IRC for support and encouragement of this work.

[1] J. M. Luttinger and W. Kohn, Phys. Rev. **97**, 869 (1955).
[2] G. Bastard, *Wave Mechanics Applied to Semiconductor Heterostructures* (Les Editions de Physique, Paris, 1988).
[3] M. G. Burt, J. Phys. Condens. Matter **4**, 6651, (1992).
[4] B. A. Foreman, Phys. Rev. B **48**, 4964 (1993).
[5] A. T. Meney, Besire Gonul, and E. P. O'Reilly, Phys. Rev. B **50**, 10 893 (1994).
[6] P. O. Löwdin, J. Chem. Phys. **19**, 1396 (1951).
[7] L. D. Laude, F. H. Pollak, and M. Cardona, Phys. Rev. B **8**, 2623 (1971)
[8] A. T. Meney, Superlatt. Microstruct. **11**, 31 (1992)

Semicond. Sci. Technol. **13** (1998) 11–17. Printed in the UK PII: S0268-1242(98)85385-6

General rules for constructing valence band effective mass Hamiltonians with correct operator order for heterostructures with arbitrary orientations

R van Dalen† and P N Stavrinou‡§

† Interdisciplinary Research Centre for Semiconductor Materials, Imperial College of Science, Technology and Medicine, Prince Consort Road, London SW7 2BZ, UK
‡ Interdisciplinary Research Centre for Semiconductor Materials, Department of Engineering Science, University of Oxford, Parks Road, Oxford OX1 3PU, UK

Received 20 June 1997, in final form 29 September 1997, accepted for publication 23 October 1997

Abstract. Working from the recently developed exact envelope-function theory, the question of how the envelope functions connect across a boundary is addressed for arbitrary growth directions. Using angular momentum basis functions that are expressed in terms of the desired growth orientation, a general form for a six-band Hamiltonian is presented. A set of general rules has been obtained which describe how unordered bulk elements can be replaced to conform to the ordering determined from the exact theory, thereby avoiding the use of *ad hoc* symmetrization procedures. One important result of this study is that the reduced interface coupling between the heavy-hole state and light/split-off-hole state found previously for [100] and [110] orientation is a general property valid for all growth directions.

1. Introduction

Effective mass theory is a powerful and convenient method for obtaining information about the electronic bandstructure of solids, and has been particularly successful for predicting electronic and optical properties of semiconductor heterostructures [1,2]. Currently heterostructures designed for confinement along lower-symmetry directions have been attracting considerable interest with effective mass calculations demonstrating that changing the orientation can significantly affect the resulting bandstructure [3–5].

Despite the success of effective mass theory, insight into how to connect envelope functions across an interface was not rigorously provided until the development of an exact envelope-function theory by Burt [6,7] (a good overview of the use of effective mass theory in heterostructures, in particular addressing the subject of boundary conditions, is contained in the review [7]). The strength of the exact theory is that it employs the same expansion states throughout the structure thereby circumventing the boundary problem. Imposing the traditional effective mass approximations, i.e. abrupt junctions and eliminating smaller envelope functions by perturbation methods, the theory describes unambiguously the operator ordering in the Hamiltonian which leads directly to the connection rules at the boundary. More recently an alternative approach, which allows different expansion states in each layer, has been proposed by Foreman [8]. As a result new interface terms are introduced in the boundary conditions. Nevertheless the underlying operator ordering, which is the subject of the present work, is closely related to that obtained from Burt's approach.

With the operator ordering now specified, this prompted a re-examination of the widely used multi-band effective mass Hamiltoninians which without a clear derivation of the ordering had previously been handled using a rather *ad hoc* symmetrization procedure. Initial studies focused on heterostructures oriented on the technologically relevant (001) face and demonstrated significant differences in the resulting bandstructure [9,10]. More recent work, now looking at structures confined along [110],

§ Present address: Imperial College†.

demonstrated dramatic differences particularly for the higher-lying subbands near the zone centre [11]. This raises the question as to the effect the new approach will have for other orientations, which to date still use the symmetrization procedure for the operator ordering and hence boundary conditions. In a previous report [11] we presented a valence band Hamiltonian written in the X, Y and Z basis (excluding spin) in which the operator ordering was derived following Burt's approach. Along with an appropriate coordinate transformation, this may be recast into the more usual angular momentum basis $|J, m_j\rangle$ for the desired growth orientation. However, this procedure can rapidly become laborious, needs to be performed for every growth direction individually and therefore prevents general observations with respect to the interface coupling. A more satisfactory approach would be to work from *un*ordered bulklike Hamiltonians, which are readily available in the literature [5, 12] or easily obtained from the theory of invariants [13], and where necessary replace terms by the ordered form prescribed from the exact theory. This is the aim of the present paper where, upon defining a general form for a valence band Hamiltonian in the $|J, m_j\rangle$ basis applicable to any growth direction, we present rules governing the replacement of unordered terms by their ordered form.

The paper is organized as follows. In section 2 we outline the specific formulation of the envelope-function theory used in this work. Section 3 introduces the expressions for the zone-centre heavy-, light-hole and spin–orbit split-off states expressed in terms of vectors chosen for the particular growth direction. These are then used to construct the form of a general six-band Hamiltonian which suitably describes the valence band in heterostructures for any growth direction. In section 4 we summarize the operator ordering rules found which are specific to the elements of the general Hamiltonian which are essentially the key findings of the present work. Sections 5 and 6 detail the method used to derive the zone-centre expansion states and the ordering rules, while in section 7 we address how different forms of Hamiltonians, currently existing in the literature, may be accommodated within the present work. In section 8 we consolidate the results of the present work by generating, as an example, expressions relevant to the [111] growth direction. Finally a summary of the work is given in section 9.

2. Effective mass theory and operator ordering

The successful envelope-function theory formulation widely used in heterostructure calculations [1] largely follows from the early work on the effective mass theory, first described in detail by Luttinger and Kohn [2]. Generally the small linear k-terms that arise in the bulk description are neglected resulting in the valence maximum occuring at the zone-centre (Γ-point), i.e. $\boldsymbol{k} = \mathbf{0}$. With spin–orbit interaction also neglected, the symmetry of the zone-centre (Γ_{15}) valence band wavefunctions can be described by bonding p functions in the tight-binding approximation, denoted by $|X\rangle$, $|Y\rangle$ and $|Z\rangle$. In the effective mass approximation one then assumes that an expansion using just these states, accounting for the excluded (remote) states by perturbation theory [14], is sufficient to describe the valence bands at non-zero $\boldsymbol{k}$.

Adopting Foreman's notation, the coupling with the remote states is split up into the individual contributions of each symmetry type (neglecting Γ_{25}) [9],

$$\sigma = (1/3m_0)\sum_{\nu}^{\Gamma_1} |\langle X|\hat{p}_x|\nu\rangle|^2/(\varepsilon_\nu - \varepsilon)$$

$$\pi = (1/3m_0)\sum_{\nu}^{\Gamma_{15}} |\langle X|\hat{p}_y|\nu\rangle|^2/(\varepsilon_\nu - \varepsilon)$$

$$\delta = (1/6m_0)\sum_{\nu}^{\Gamma_{12}} |\langle X|\hat{p}_x|\nu\rangle|^2/(\varepsilon_\nu - \varepsilon) \qquad (1)$$

where ε is the energy of the valence band edge and the summation over ν runs over all remote states at energy ε_ν of given symmetry. These constants replace the traditional Luttinger parameters which were defined to give quantities directly related to the heavy, light and split-off masses in a bulk semiconductor. Identifying the contributions of each symmetry type in the Luttinger parameters, one finds

$$\gamma_1 = -1 + 2\sigma + 4\pi + 4\sigma$$

$$\gamma_2 = \sigma - \pi + 2\delta$$

$$\gamma_3 = \sigma + \pi - \delta. \qquad (2)$$

Working in Burt's framework [6, 7], the traditional effective mass Hamiltonian in each layer over an $|X\rangle$, $|Y\rangle$ and $|Z\rangle$ basis has to be replaced by a version with specific operator ordering [11],

$$\mathbf{D} = \frac{\hbar^2}{2m_0}\begin{pmatrix} \hat{k}_x A\hat{k}_x + \hat{k}_y B\hat{k}_y + \hat{k}_z B\hat{k}_z & \hat{k}_x C_1\hat{k}_y - \hat{k}_y C_2\hat{k}_x & \hat{k}_x C_1\hat{k}_z - \hat{k}_z C_2\hat{k}_x \\ \hat{k}_y C_1\hat{k}_x - \hat{k}_x C_2\hat{k}_y & \hat{k}_y A\hat{k}_y + \hat{k}_x B\hat{k}_x + \hat{k}_z B\hat{k}_z & \hat{k}_y C_1\hat{k}_z - \hat{k}_z C_2\hat{k}_y \\ \hat{k}_z C_1\hat{k}_x - \hat{k}_x C_2\hat{k}_z & \hat{k}_z C_1\hat{k}_y - \hat{k}_y C_2\hat{k}_z & \hat{k}_z A\hat{k}_z + \hat{k}_x B\hat{k}_x + \hat{k}_y B\hat{k}_y \end{pmatrix} \qquad (3)$$

where

$$A = 1 - 6\sigma - 12\delta \qquad C_1 = 6\delta - 6\sigma$$

$$B = 1 - 6\pi \qquad C_2 = 6\pi. \qquad (4)$$

In (3), $\hat{k}_j$ is to be interpreted as the differential operator $(1/\mathrm{i})\nabla_j$ and the ordering in each element has followed naturally from working within the framework of the exact theory. To include the effects of spin–orbit coupling, the most convenient approach is to form a new basis $|J, m_j\rangle$, comprising linear combinations of $|X\rangle$, $|Y\rangle$ and $|Z\rangle$ that are eigenfunctions of the spin–orbit operator

$$|J, m_j\rangle = |\chi_1 X + \chi_2 Y + \chi_3 Z\rangle \uparrow + |\chi_4 X + \chi_5 Y + \chi_6 Z\rangle \downarrow \qquad (5)$$

where J and m_j represent the total angular momentum and its projection along a particular and as-yet undefined axis. For each layer, the set $|J, m_j\rangle$ represents the exact heavy-hole, light-hole and split-off-hole eigenstates at the band edge. The wavefunction in a heterostructure is then

written as a linear combination of this set with each element multiplied by a so-called envelope function

$$|\psi\rangle = F_1|\tfrac{3}{2},+\tfrac{3}{2}\rangle + F_2|\tfrac{3}{2},+\tfrac{1}{2}\rangle + F_3|\tfrac{3}{2},-\tfrac{1}{2}\rangle + F_4|\tfrac{3}{2},-\tfrac{3}{2}\rangle + F_5|\tfrac{1}{2},+\tfrac{1}{2}\rangle + F_6|\tfrac{1}{2},-\tfrac{1}{2}\rangle. \quad (6)$$

The corresponding effective mass Hamiltonian in matrix form, including information about the ordering, can now simply be constructed from (3)–(6) by realizing that the individual elements of such a Hamiltonian are given by

$$H_{mn} = \sum_{i,j} \langle \{J, m_j\rangle_m | i\rangle D_{ij} \langle j | \{J, m_j\}_n\rangle \quad (7)$$

where i, j run over X, Y and Z while m, n run over the expansion states explicitly noted in (6). In the following section, expressions for $|J, m_j\rangle$ written in terms of orientation vectors are presented which enable the construction of a general form of the Hamiltonian.

3. General form for the effective mass Hamiltonian

To account for arbitrary substrate orientations, we introduce a transformation matrix $\mathbf{T}$, whose elements are related to the orientation vectors describing the principal directions of the structure. For example, a new coordinate frame r' is obtained through a proper rotation of the standard Cartesian coordinate system r, i.e. $r' = \mathbf{T}r$ or,

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}. \quad (8)$$

The row elements of $\mathbf{T}$ may be viewed as vectors a, b, c which form an orthonormal right-handed set with c pointing along the growth (confinement) direction. Thus in general, the matrix $\mathbf{T}$ describes how all relevant operators transform under a rotation of the axis. From the introduction of $\mathbf{T}$ we can express a general $|J, m_j\rangle$ basis in terms of orientation vector elements. Specifically,

$$\begin{aligned}
|\tfrac{3}{2},+\tfrac{3}{2}\rangle &= \frac{1}{\sqrt{2}}|(a_1+\mathrm{i}b_1)X+(a_2+\mathrm{i}b_2)Y+(a_3+\mathrm{i}b_3)Z\rangle\uparrow \\
|\tfrac{3}{2},+\tfrac{1}{2}\rangle &= \frac{2}{\sqrt{6}}|c_1X+c_2Y+c_3Z\rangle\uparrow \\
&\quad -\frac{1}{\sqrt{6}}|(a_1+\mathrm{i}b_1)X+(a_2+\mathrm{i}b_2)Y+(a_3+\mathrm{i}b_3)Z\rangle\downarrow \\
|\tfrac{1}{2},+\tfrac{1}{2}\rangle &= -\frac{1}{\sqrt{3}}|c_1X+c_2Y+c_3Z\rangle\uparrow \\
&\quad -\frac{1}{\sqrt{3}}|(a_1+\mathrm{i}b_1)X+(a_2+\mathrm{i}b_2)Y+(a_3+\mathrm{i}b_3)Z\rangle\downarrow .
\end{aligned} \quad (9)$$

The derivation of these expressions is explained fully in section 6. The corresponding $|J, -m_j\rangle$ states can be constructed using the Kramers operator [15] on the positive counterparts in (9) with a change in the overall sign. So written in terms of (5) we would find,

$$|J, -m_j\rangle = |\chi_4^*X+\chi_5^*Y+\chi_6^*Z\rangle\uparrow - |\chi_1^*X+\chi_2^*Y+\chi_3^*Z\rangle\downarrow . \quad (10)$$

The matrix elements for the effective mass Hamiltonian are found using (5), (7), (9) and (10). Clearly their exact form will depend on the particular growth direction chosen; nevertheless it is possible to derive a general form for this Hamiltonian. Thus regardless of growth direction, the overall form of a six-band effective mass Hamiltonian for each layer can be expressed as,

$$H = \begin{pmatrix}
P+Q & -S_1 & R & 0 & S_1/\sqrt{2} & \sqrt{2}R \\
 & P-Q & \Theta C & -R & \sqrt{2}Q & \Pi\sqrt{\tfrac{3}{2}}S_3 \\
 & & P-Q & -S_2 & -\Pi^\dagger\sqrt{\tfrac{3}{2}}S_4^\dagger & \sqrt{2}Q \\
 & & & P+Q & -\sqrt{2}R^\dagger & S_2^\dagger/\sqrt{2} \\
 & & & & P+\Delta & -\Theta C \\
 & & & & & P+\Delta
\end{pmatrix}. \quad (11)$$

Only the upper diagonal is indicated as the lower-diagonal elements follow from the requirement that the Hamiltonian must be Hermitian. As already noted, the spin–orbit operator will always be diagonal for the basis used, hence in (11) Δ represents the spin–orbit splitting energy for the layer. The subscripts on S-terms are used to distinguish elements that have a specific operator ordering which, along with the operator ordering of the P-, Q- and R-terms, will be dealt with in the following section. Note however that if the ordering is neglected, e.g. as in the bulk descriptions often used in the literature, then S_1, S_2, S_3 and S_4 will all reduce to a single term S. Finally, phase terms of unit length, Π and Θ, have been introduced in (11) in order to accommodate any differences in phase of existing $|J, m_j\rangle$ sets used in the literature compared to the set given in (9) and (10); for the present set both Π and Θ would be equal to 1. The explicit form and origin of these terms is discussed in detail later on in section 7.

4. General rules for operator ordering

We now turn to the operator ordering for each of the elements in (11). Recall that since we are considering heterostructures, with the growth (confinement) direction taken along z', it is only the placement of the operator $\hat{k}_{z'}$ with respect to the position dependent mass parameters that is important for the boundary conditions. The operators acting in the plane simply reduce to components of the Bloch wavevector, $\hat{k}_{x'(y')} \rightarrow k_{x'(y')}$. In principle the ordering for a particular growth direction can be found in a straightforward, albeit lengthy, manner through the application of (7). However in keeping with a general approach, it is also possible to obtain rules applicable to *any* direction that describe how to transform an unordered bulklike element into an element with the correct operator ordering. The attraction of this approach is that unordered elements for a particular growth direction can be found either from existing work in the literature [5, 12] or may be easily calculated using Luttinger's theory of invariants [13].

We now present the general rules governing the operator ordering of the elements of the Hamiltonian (11). The rules were obtained by exploiting the linear nature of the transformation (7) and the structure of $\mathbf{D}$ in (3): the details of the procedure are outlined in section 6. We obtained the following results.

• P and Q are symmetrized with respect to the operator $\hat{k}_{z'}$.

• All terms in S_i involving $\hat{k}_{z'}^2$ are to be symmetrized.

• Any terms in S_i containing a single operator $\hat{k}_{z'}$, i.e. with a unordered form such as $\hat{k}_{z'}(f\sigma + g\pi + h\delta)$ where f, g and h are constant coefficients, should be ordered as follows:

$$
\begin{aligned}
S_1 &: \left(f\sigma - \frac{f-g}{2}\pi - \frac{f-h}{2}\delta\right)\hat{k}_{z'} \\
&\quad + \hat{k}_{z'}\left(\frac{f+g}{2}\pi + \frac{f+h}{2}\delta\right) \\
S_2 &: \left(\frac{f+g}{2}\pi + \frac{f+h}{2}\delta\right)\hat{k}_{z'} \\
&\quad + \hat{k}_{z'}\left(f\sigma - \frac{f-g}{2}\pi - \frac{f-h}{2}\delta\right) \\
S_3 &: \left(\frac{f}{3}\sigma + \frac{f+3g}{6}\pi + \frac{f+3h}{6}\delta\right)\hat{k}_{z'} \\
&\quad + \hat{k}_{z'}\left(\frac{2f}{3}\sigma - \frac{f-3g}{6}\pi - \frac{f-3h}{6}\delta\right) \\
S_4 &: \left(\frac{2f}{3}\sigma - \frac{f-3g}{6}\pi - \frac{f-3h}{6}\delta\right)\hat{k}_{z'} \\
&\quad + \hat{k}_{z'}\left(\frac{f}{3}\sigma + \frac{f+3g}{6}\pi + \frac{f+3h}{6}\delta\right). \qquad (12)
\end{aligned}
$$

• All terms in R are symmetrized with respect to the operator $\hat{k}_{z'}$ and do not contain any combination of σ and $\hat{k}_{z'}$.

• The term C, which is only non-zero at an interface, always takes the form

$$-(\hbar/m_0)(k_{x'} - \mathrm{i}k_{y'})[(\sigma - \pi - \delta)\hat{k}_{z'} - \hat{k}_{z'}(\sigma - \pi - \delta)]. \quad (13)$$

We have frequently referred to 'symmetrized' in the above rules: by this we mean application of the traditional symmetrization procedure for the $\hat{k}_{z'}$ operator that has previously been adopted [1]. In fact as the above rules demonstrate there are two principle differences that arise from working within the exact theory, the ordering for S_i terms which are linear in $\hat{k}_{z'}$ and the introduction of the C element. A particularly significant result from (12) is in the strong asymmetry of the sigma coefficient for the S_1 and S_2 elements. Integrating across an interface, it follows that coupling between heavy- and light-hole states *never* includes the large contribution from remote states with s-type symmetry, which is in contrast to the result obtained if a symmetrization procedure is used for these elements. Previous work had found this to be the case for [100] and [110] structures [9, 11]: what the present work confirms is this is in fact a general feature for *all* growth directions.

Although we will demonstrate the application of the above rules in full in section 8 where we will derive an ordered [111] Hamiltonian working from unordered matrix elements, we will now give a brief illustration to familiarize the reader with the concept of the rules (12) and (13). At this point it is useful to recall the work by Foreman where he presented an ordered Hamiltonian for the growth direction [100] [9]. This Hamiltonian was obtained by evaluating all elements H_{mn} individually using (7) and an explicit knowledge of the basis states $|J, m_j\rangle$ for this growth direction. We will now reproduce his result for the element S_1 using the appropriate rule in (12). Neglecting any ordering, this element is written as

$$S_1 = \sqrt{3}(k_x - \mathrm{i}k_y)\gamma_3 k_z = \sqrt{3}(k_x - \mathrm{i}k_y)[\sigma + \pi - \delta]k_z. \quad (14)$$

We find that f, g and h are given by $f = g = -h = \sqrt{3}$; inserting this in the top rule of (12), we obtain

$$S_1 = \sqrt{3}(k_x - \mathrm{i}k_y)[(\sigma - \delta)\hat{k}_z + \hat{k}_z(\pi)] \quad (15)$$

which agrees with the result obtained by Foreman (see equation (5*b*) in [9]).

5. Derivation of the $|J, m_j\rangle$ states

To clarify the derivation of the $|J, m_j\rangle$ basis shown in (9) it is useful to recall the requirement that such a set should fulfil, namely that the spin–orbit operator should always be diagonal in the basis. In other words, such a set represents the eigenfunctions of the spin–orbit operator. The approach taken to derive and label the set in (9) is to recognize that eigenfunctions of the spin–orbit operator are also eigenfunctions of $\hat{J}^2$, where $\hat{J}$ represents the total angular momentum operator. Since this observable is insufficient to uniquely specify the states, one has the freedom to choose another (commuting) observable which we choose to be $\hat{J}_{z'}$, the projection of $\hat{J}$ along the growth axis. As a consequence, the states will be labelled $|J, m_j\rangle$ and we will align the spin along z' [1, 11]. The choice of operators is a natural choice since the $\boldsymbol{k}\cdot\boldsymbol{p}$ term in the effective mass formalism commutes with $\hat{J}_{z'}$ for $k_{x'} = k_{y'} = 0$, thus for this basis the Hamiltonian is well conditioned (diagonal or close to diagonal) at the band edge. Furthermore, this choice is compatible with Luttinger's theory of invariants [13] when both the $\boldsymbol{k}$-vectors and angular momentum matrices are transformed under $\mathbf{T}$.

Specifically, working from the most general linear combination of X, Y and Z states, i.e. equation (5), we have to find the χ_i coefficients which simultaneously satisfy

$$\hat{J}^2|J, m_j\rangle = \hbar^2 J(J+1)|J, m_j\rangle$$

$$\hat{J}_z|J, m_j\rangle = \hbar m_j|J, m_j\rangle \quad (16)$$

where $\hat{J}^2 = \hat{L}^2 + 2\hat{L}\cdot\hat{S} + \hat{S}^2$ and $\hat{J}_{z'} = \hat{L}_{z'} + \hat{S}_{z'}$. To find expressions for the $|3/2, +3/2\rangle$ state, the equations above would reduce to

$$2\hat{L}'\cdot\hat{S}'|\tfrac{3}{2}, +\tfrac{3}{2}\rangle = |\tfrac{3}{2}, +\tfrac{3}{2}\rangle$$

$$(\hat{L}_{z'} + \hat{S}_{z'})|\tfrac{3}{2}, +\tfrac{3}{2}\rangle = 3/2|\tfrac{3}{2}, +\tfrac{3}{2}\rangle. \quad (17)$$

Note that although the state has been written explicitly, it is understood the coefficients are still to be determined. To proceed, $\hat{L}'$ is transformed to the original reference frame, $\hat{L}' = \mathbf{T}\hat{L}$, so that the first equation in (17) transforms into

$$2[(\boldsymbol{a}\cdot\hat{L})\hat{S}_{x'} + (\boldsymbol{b}\cdot\hat{L})S_{y'} + (\boldsymbol{c}\cdot\hat{L})S_{z'}]|\tfrac{3}{2}, \tfrac{3}{2}\rangle = |\tfrac{3}{2}, \tfrac{3}{2}\rangle. \quad (18)$$

For the state in question, it is possible to reduce the number of unknown coefficients from the outset by realizing that a state with angular momentum of 3/2 cannot be built up out of $l = 1$ orbitals combined with a spin-down spinor, i.e. we set $|3/2, +3/2\rangle = \alpha_1 X\uparrow + \alpha_2 Y\uparrow + \alpha_3 X\uparrow$. Using the relations (and similarly for the other elements)

$$\hat{L}_x|X\rangle = 0 \qquad \hat{L}_x|Y\rangle = \mathrm{i}|Z\rangle \qquad \hat{L}_x|Z\rangle = -\mathrm{i}|Y\rangle$$

$$\hat{S}_{x'}\uparrow = \tfrac{1}{2}\downarrow \qquad \hat{S}_{x'}\downarrow = \tfrac{1}{2}\uparrow \tag{19}$$

we obtain six equations for the coefficients, conveniently expressed in vector form as

$$\mathrm{i}\boldsymbol{c}\times\boldsymbol{\alpha} = \boldsymbol{\alpha}$$

$$(\boldsymbol{a} + \mathrm{i}\boldsymbol{b})\times\boldsymbol{\alpha} = \boldsymbol{0} \tag{20}$$

where $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \alpha_3)$. For $|3/2, +3/2\rangle$ the second equation of (20) does not give additional restrictions as it reduces to the first equation in (20). Since $\boldsymbol{a}$, $\boldsymbol{b}$ and $\boldsymbol{c}$ form a right-handed orthonormal set it follows that the only coefficients that can satisfy (20), excluding an overall phase factor, correspond to those presented in (9). The other states listed in (9) are derived along the same manner.

6. Derivation of the ordering rules

We will now outline the procedure followed to derive the ordering rules. As mentioned earlier, the approach we take is to refer to correct ordering expressions with respect to unordered bulklike elements. It follows from the linear nature of (7) that a conversion from a Hamiltonian without operator ordering to an ordered Hamiltonian within Burt's framework will redistribute the σ-, π- and δ-terms as

$$(f\sigma + g\pi + h\delta)\hat{k}_{z'}|_{bulk} \rightarrow [(f_1\sigma + g_1\pi + h_1\delta)\hat{k}_{z'} + \hat{k}_{z'}(f_2\sigma + g_2\pi + h_2\delta)] \tag{21}$$

where $f_1 + f_2 = f$, $g_1 + g_2 = g$ and $h_1 + h_2 = h$ are the coefficients of the remote terms for a particular Hamiltonian element obtained using (7). Not shown explicitly in (21) is a reference to in-plane Bloch components $k_{x'}$ and $k_{y'}$ which are simply multiplicative constants for the total expression. A quantity is then defined, the degree of asymmetry, which describes the difference between the occurrence of an element before and after the operator $\hat{k}_{z'}$. For example, the degree of asymmetry in σ is $f_1 - f_2$. We then look to find a relationship between the degree of asymmetry for each of the three symmetries of the remote states. Referring back to (3), we note that the distribution of π and δ in the off-diagonal elements is similar to that of σ, the only difference being the reverse ordering of terms containing π and the difference in sign of the terms containing δ. These two differences have essentially the same effect regarding the comparative degree of asymmetry between the terms. The diagonal terms D_{11}, D_{22} and D_{33} only give rise to a symmetrical contribution with respect to $\hat{k}_{z'}$, therefore the degree of asymmetry is determined by the off-diagonal terms only. As a result we find that the degree of asymmetry in π and δ is identical to that of σ, only of opposite sign. Explicitly,

$$f_1 - f_2 = g_2 - g_1 = h_2 - h_1. \tag{22}$$

It is important to note that we have only transformed the momentum vector $\hat{k}$ and not the basis states $|X\rangle$, $|Y\rangle$ and $|Z\rangle$. Rotating these basis states would change the degree of asymmetry in the individual elements of **D** (3) and the above discussion. Naturally one expects to obtain the same final result for the Hamiltonian if one works with rotated basis states and consequently transforms **D**. This is straightforward to prove and confirms that the result for the degree of asymmetry as given by (21) and (22) is a general property of the Hamiltonian.

Relating the degree of asymmetry from the other terms with respect to that for the σ-term proves convenient since coefficients involving σ and $\hat{k}_{z'}$ in (11) written in the general $|J, m_j\rangle$ basis (9) are found to be *independent* of reference frame, unlike π- and δ-terms which depend on a particular growth direction [16].

Consider the spin-up heavy-hole related elements, i.e. the first row of the Hamiltonian in (11). Evaluating (7) using $|J, m_j\rangle_m = |3/2, +3/2\rangle$ as given in (9) plus the general expression (5) for $|J, m_j\rangle_n$ and concentrating only on the coupling with the remote bands with s-type symmetry (σ), we obtain (working in atomic units).

$$\begin{aligned}\langle \tfrac{3}{2}, +\tfrac{3}{2}|H|J, m_j\rangle &= -\frac{3}{\sqrt{2}}(a_1 - \mathrm{i}b_1)(\chi_1\hat{k}_x\sigma\hat{k}_x + \chi_2\hat{k}_x\sigma\hat{k}_y + \chi_3\hat{k}_x\sigma\hat{k}_z)\\ &\quad -\frac{3}{\sqrt{2}}(a_2 - \mathrm{i}b_2)(\chi_1\hat{k}_y\sigma\hat{k}_x + \chi_2\hat{k}_y\sigma\hat{k}_y + \chi_3\hat{k}_y\sigma\hat{k}_z)\\ &\quad -\frac{3}{\sqrt{2}}(a_3 - \mathrm{i}b_3)(\chi_1\hat{k}_z\sigma\hat{k}_x + \chi_2\hat{k}_z\sigma\hat{k}_y + \chi_3\hat{k}_z\sigma\hat{k}_z).\end{aligned} \tag{23}$$

Transforming the left-hand $\hat{k}_{z'}$ operator under a general rotation of the axis (5) and considering only those components associated with $\hat{k}_{z'}\sigma$ (only these will give a non-zero contribution when integrating across a boundary), we find

$$-(3/\sqrt{2})(\boldsymbol{a} - \mathrm{i}\boldsymbol{b})\cdot\boldsymbol{c}\hat{k}_{z'}\sigma(\chi_1\hat{k}_x + \chi_2\hat{k}_y + \chi_3\hat{k}_z) = 0 \tag{24}$$

which is due to the fact that $\boldsymbol{a}$, $\boldsymbol{b}$ and $\boldsymbol{c}$ form an orthonormal set. This is a very important result as this demonstrates that only the combination of $\sigma\hat{k}_{z'}$ can be present in this row of the Hamiltonian. It implies that for the boundary conditions there is no interband coupling via the remote bands of s symmetry. From this result which defines the degree of asymmetry for the σ-coefficient, combined with (21) and (22), we can set $g_2 - g_1 = h_2 - h_1 = f$ which leads to the rule in (12) governing the correct ordering of S_1 terms. The expression for S_1 terms is actually the *general* rule for all the elements involving a single $\hat{k}_{z'}$ operator in the top row of (11). For particular elements, where $|J, m_j\rangle_n$ is specified, the derivation proceeds along in the same manner. When considering the R term then in addition to (24) we also find that no combination $\sigma\hat{k}_{z'}$ occurs, i.e. $f = 0$, which shows the ordering for R is symmetric with respect to $\hat{k}_{z'}$ or the degree of asymmetry is zero. The other rules for S_2, S_3 and S_4 are evaluated in a similar manner. The procedure to obtain the expression for C follows in the same way only here the π- and δ-terms must also be considered.

7. The phase of the $|J, m_j\rangle$ basis

In this section we show how, with the introduction of two phase terms, Π and Θ, existing Hamiltonians which are not derived from the set (9) may be recast into the general form (11). The appearance of different Hamiltonians in the literature stems from the basis set used to derive them. Although we have chosen two commuting observables and set the particular projection axis involved, the $|J, m_j\rangle$ states as presented in (9) and (10) are by no means unique in that they may be multiplied by factor of unit length $(\mu + i\nu)$ without changing any of the physics. Typically, the consequence of using a $|J, m_j\rangle$ basis that differs in phase from (9) is apparent in the off-diagonal elements which differ by some (complex) phase term. For example there are at least two general forms of four-band Hamiltonians in the literature where the off-diagonal elements differ in phase (compare [3–5] and [9, 12]); for the present work we have chosen the phases of (9) and (10) to correspond to the latter notation. As we now go on to show, and demonstrate in the following section, it is possible to account for the phase of the $|J, m_j\rangle$ set.

Consider the unitary transformation $\mathbf{H}' = \mathbf{U}^\dagger \mathbf{H} \mathbf{U}$, where $\mathbf{U}$ is a diagonal matrix with elements

$$U_{ij} = \delta_{ij} u_i$$

$$u = (1, \exp[i\alpha], \exp[i\beta], \exp[i(\alpha+\beta)], \exp[i\alpha], \exp[i\beta]) \tag{25}$$

and δ_{ij} is the Kronecker delta. As a result of the transformation the following elements in the Hamiltonian are multiplied by a phase factor:

$$S_{1,2} \rightarrow \exp[i\alpha] S_{1,2}$$

$$R \rightarrow \exp[i\beta] R$$

$$S_{3,4} \rightarrow \exp[i(\beta-\alpha)] S_{3,4}$$

$$C \rightarrow \exp[i(\beta-\alpha)] C. \tag{26}$$

To demonstrate how this unitary transformation was used to account for the choice in $|J, m_j\rangle$ states in (11), we will first look at a four-band Hamiltonian without a C-element (i.e. only the top two transformations in (26) are of relevance). It follows from (26) that we can multiply all the R- and S-elements with a phase factor without changing the eigenvalues. Effectively one is now working with new off-diagonal elements, $S_{new} = \exp(i\alpha) S$ and $R_{new} = \exp(i\beta) R$, which are based on a different $|J, m_j\rangle$ set as compared to (9) and (10).

We will now extend this to a full six-band model with C-elements. Again, S_{new} and R_{new} represent 'arbitrary' off-diagonal elements which correspond to an as-yet unknown phase and $|J, m_j\rangle$ set (for example, they could have been obtained directly from the theory of invariants [13]). As the C-element is not invariant in the transformation (25) we see that we have to introduce a phase factor $\Theta = \exp[i(\beta-\alpha)]$ to account for any difference in basis states. Similarly as $S_{3,4}$ transform differently from $S_{1,2}$, although they were all defined to reduce to the same element neglecting any ordering, we see that we have to introduce another phase factor Π which is placed in the Hamiltonian elements containing $S_{3,4}$. Using (26), one obtains

$$\Pi = \frac{\exp[i(\beta-\alpha)]}{\exp[i\alpha]} = \exp[i(\beta-2\alpha)]. \tag{27}$$

Both factors can be obtained by comparing the chosen R and S to the expressions one would have obtained when using the set (9). As the latter set will always give a real and positive $\sigma k_{x'}^2$ component in R and a real and positive $\sigma k_{x'} \hat{k}_{z'}$ component in S which can be checked by evaluating $\langle 3/2, +3/2|H|3/2, \pm 1/2\rangle$, we arrive at the following expressions for Π and Θ

$$\Pi = -(s^{\sigma}_{x'z'}{}^{*})^2 r^{\sigma}_{x'x'} \mathrm{sign}(P)$$

$$\Theta = -s^{\sigma}_{x'z'}{}^{*} r^{\sigma}_{x'x'} \mathrm{sign}(P) \tag{28}$$

where $r^{\sigma}_{x'x'}$ is the normalized $\sigma k_{x'}^2$ coefficient in R and $s^{\sigma}_{x'z'}$ is the normalized $\sigma k_{x'} \hat{k}_{z'}$ coefficient in S. We have introduced the factor $\mathrm{sign}(P)$ to eliminate the effects of an overall sign change in the Hamiltonian often performed to obtain positive hole energies.

8. Application to [111]

To demonstrate the application of the rules presented in section 4, we consider the case for growth along the [111] direction. As a first step, we look to obtain the unordered bulklike elements for a four-band Hamiltonian, as the other elements can be expressed in terms of these. Luttinger's theory of invariants [13] provides us with a straightforward way to obtain the four-band Hamiltonian elements, this Hamiltonian is given by (atomic units)

$$\mathbf{H} = -\tfrac{1}{2}(\gamma_1 + \tfrac{5}{2}\gamma_2)\hat{k}^2 + \gamma_2(\hat{k}_x^2 \mathbf{J}_x^2 + \hat{k}_y^2 \mathbf{J}_y^2 + \hat{k}_z^2 \mathbf{J}_z^2)$$
$$+2\gamma_3(\hat{k}_x \hat{k}_y \{\mathbf{J}_x \mathbf{J}_y\} + \hat{k}_y \hat{k}_z \{\mathbf{J}_y \mathbf{J}_z\} + \hat{k}_z \hat{k}_x \{\mathbf{J}_z \mathbf{J}_x\}) \tag{29}$$

where $\mathbf{J}_{x,y,z}$ are the angular momentum matrices (we employ the same definitions as used in the original article) and $\{\mathbf{J}_i \mathbf{J}_j\} = \mathbf{J}_i \mathbf{J}_j + \mathbf{J}_j \mathbf{J}_i$. The new axes are taken along $(1, 1, -2)$, $(-1, 1, 0)$ and $(1,1,1)$. Performing a coordinate transformation on both the operators and the angular momentum matrices (dropping the primes from now on), we find the following unordered Hamiltonian elements,

$$P = -\tfrac{1}{2}\gamma_1(k_x^2 + k_y^2 + \hat{k}_z^2)$$

$$Q = -\tfrac{1}{2}\gamma_3(k_x^2 + k_y^2 - 2\hat{k}_z^2)$$

$$S = \frac{i}{\sqrt{6}}(\gamma_2 - \gamma_3)k_+^2 - \frac{i}{\sqrt{3}}(2\gamma_2 + \gamma_3)k_- \hat{k}_z$$

$$R = -\frac{\sqrt{3}}{6}(\gamma_2 + 2\gamma_3)k_-^2 + \frac{\sqrt{6}}{3}(\gamma_2 - \gamma_3)k_+ \hat{k}_z \tag{30}$$

where $k_\pm = k_x \pm i k_y$. In both the R- and S-elements there are terms linear in $\hat{k}_z$; for these parts we substitute σ, π

Table 1.

$$\begin{pmatrix}
(\gamma_1-2\gamma_3)\frac{\partial}{\partial z} & \frac{2}{\sqrt{3}}k_-(\pi+3\delta) & -\frac{\sqrt{6}}{3}ik_+(2\pi-3\delta) & 0 & -\frac{2}{\sqrt{6}}k_-(\pi+3\delta) & -\frac{2}{\sqrt{3}}ik_+(2\pi-3\delta) \\
-\frac{2}{\sqrt{3}}k_+(3\sigma-2\pi) & (\gamma_1+2\gamma_3)\frac{\partial}{\partial z} & 2k_-(\sigma-\pi-\delta) & \frac{\sqrt{6}}{3}ik_+(2\pi-3\delta) & -2\sqrt{2}\gamma_3\frac{\partial}{\partial z} & \sqrt{2}k_-(2\sigma-\pi+\delta) \\
-\frac{\sqrt{6}}{3}ik_-(2\pi-3\delta) & 2k_+(\sigma-\pi-\delta) & (\gamma_1+2\gamma_3)\frac{\partial}{\partial z} & \frac{2}{\sqrt{3}}k_-(3\sigma-2\pi) & \sqrt{2}k_+(2\sigma-\pi+\delta) & -2\sqrt{2}\gamma_3\frac{\partial}{\partial z} \\
0 & \frac{\sqrt{6}}{3}ik_-(2\pi-3\delta) & -\frac{2}{\sqrt{3}}k_+(\pi+3\delta) & (\gamma_1-2\gamma_3)\frac{\partial}{\partial z} & \frac{2}{\sqrt{3}}ik_-(2\pi-3\delta) & \frac{2}{\sqrt{6}}k_+(\pi+3\delta) \\
\frac{2}{\sqrt{6}}k_+(3\sigma-2\pi) & -2\sqrt{2}\gamma_3\frac{\partial}{\partial z} & -\sqrt{2}k_-(\sigma+2\delta) & \frac{2}{\sqrt{3}}ik_+(2\pi-3\delta) & \gamma_1\frac{\partial}{\partial z} & -2k_-(\sigma-\pi-\delta) \\
-\frac{2}{\sqrt{3}}ik_-(2\pi-3\delta) & -\sqrt{2}k_+(\sigma+2\delta) & -2\sqrt{2}\gamma_3\frac{\partial}{\partial z} & -\frac{2}{\sqrt{6}}k_-(3\sigma-2\pi) & -2k_+(\sigma-\pi-\delta) & \gamma_1\frac{\partial}{\partial z}
\end{pmatrix}$$

and δ for the Luttinger parameters and then using (12) we find the ordered replacements. Explicitly,

$$\begin{aligned}
S_1 &: (3\sigma-2\pi)\hat{k}_z+\hat{k}_z(\pi+3\delta) \\
S_2 &: (\pi+3\delta)\hat{k}_z+\hat{k}_z(3\sigma-2\pi) \\
S_3 &: (\sigma+2\delta)\hat{k}_z+\hat{k}_z(2\sigma-\pi+\delta) \\
S_4 &: (2\sigma-\pi+\delta)\hat{k}_z+\hat{k}_z(\sigma+2\delta) \\
R &: (2\pi-3\delta)\hat{k}_z+\hat{k}_z(2\pi-3\delta). \qquad (31)
\end{aligned}$$

To evaluate the phase terms, we return to the bulk equations (30). We find that the sign of P is negative, the component $-\frac{1}{2}\sqrt{3}\sigma k_x^2$ in R and $-\sqrt{3}i\sigma k_x\hat{k}_z$ in S, so that working from (28) this implies,

$$\left.\begin{array}{l}\mathrm{sign}(P)=-1\\ r_{xx}^{\sigma}=-1\\ s_{xz}^{\sigma}=-\mathrm{i}\end{array}\right\}\ \Pi=+1 \qquad \Theta=-\mathrm{i}. \qquad (32)$$

Including the C-term completes the transformation of the Hamiltonian. For the boundary conditions, we find the requirement of continuity of $\boldsymbol{F}$ and $\Omega\boldsymbol{F}$ where $\boldsymbol{F}$ is a six-component vector containing the envelopes and Ω is given by the matrix in table 1.

As expected, we see that the boundary condition matrix does not feature σ-terms in the first and fourth row, i.e. the heavy-hole bands do not couple with the light- or split-off-hole bands via the remote states of symmetry Γ_1. As a result, along with the additional coupling that originates from the C-term, the boundary condition matrix shows profound differences from that which would have been obtained had one worked from a symmetrized Hamiltonian [5].

9. Conclusions

In summary, we have presented analytical rules to construct six-band effective mass Hamiltonians with operator ordering for any growth direction, conforming to Burt's exact envelope-function theory framework. These rules completely circumvent the need to construct such a Hamiltonian from first principles using an appropriate $|J,m_j\rangle$ basis for the growth direction under consideration and can be directly applied to a bulk (unordered) Hamiltonian. As a result, we have proved that the reduced interface coupling for the heavy-hole bands as observed for [100] and [110] is a general property for all growth directions. Moreover, we were able to show that the additional coupling between the light-hole states as well as the split-off-hole states, which is not present in the symmetric boundary conditions (the C-term), has the same form for all growth directions.

Acknowledgments

The authors would like to thank Jeff Harris and Gareth Parry from the EPSRC funded Semiconductor IRC for support and encouragement of this work. R van Dalen gratefully acknowledges the financial support by the European Union under the TMR framework.

References

[1] Bastard G 1988 *Wave Mechanics Applied to Semiconductor Heterostructures* (Paris: Editions de Physique)
[2] Luttinger J M and Kohn W 1955 *Phys. Rev.* **97** 869
[3] Shechter G, Shvartsman L D and Golub J E 1995 *Phys. Rev.* B **51** 10857
[4] Henderson R H and Towe E 1996 *J. Appl. Phys.* **79** 2029
[5] Meney A T 1992 *Superlatt. Microstruct.* **11** 31
[6] Burt M G 1987 *Semicond. Sci. Technol.* **2** 460 (erratum **2** 701)
[7] Burt M G 1992 *J. Phys.: Condens. Matter* **4** 6651
[8] Foreman B A 1996 *Phys. Rev.* B **54** 1909
[9] Foreman B A 1993 *Phys. Rev.* B **48** 4964
[10] Meney A T, Besire Gonul and O'Reilly E P 1994 *Phys. Rev.* B **50** 10893
[11] Stavrinou P N and van Dalen R 1997 *Phys. Rev.* B **55** 15456
[12] Iconic Z, Milanovic V and Tjapkin D 1992 *Phys. Rev.* B **46** 4285
[13] Luttinger J M 1956 *Phys. Rev.* **102** 1030
[14] Löwdin P O 1951 *J. Chem. Phys.* **19** 1396
[15] Kane E O 1956 *J. Phys. Chem. Solids* **1** 82
[16] Stavrinou P N and van Dalen R 1997 unpublished calculations