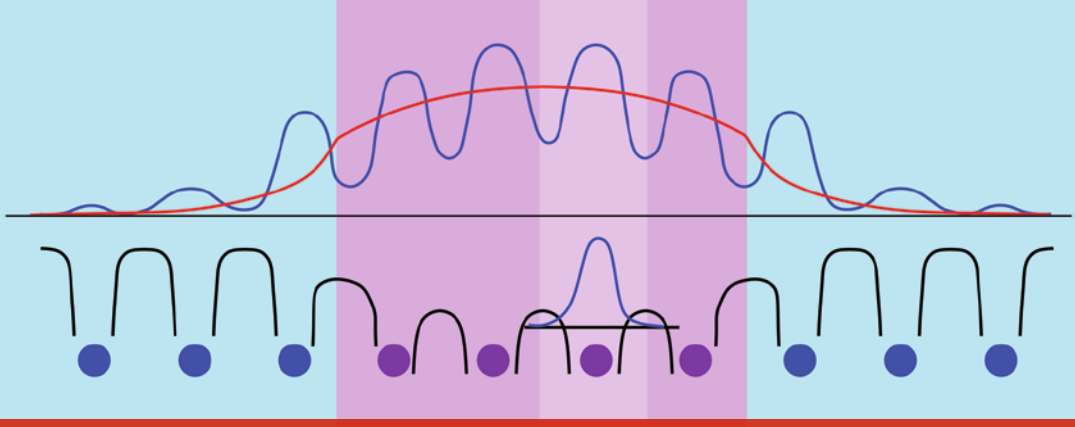


LECTURE NOTES IN COMPUTATIONAL SCIENCE AND ENGINEERING 94

Matthias Ehrhardt · Thomas Koprucki
Editors

Multi-Band Effective Mass Approximations

Springer

Lecture Notes
in Computational Science
and Engineering

94

Editors:

For further volumes:
http://www.springer.com/series/3527

Matthias Ehrhardt • Thomas Koprucki
Editors

Multi-Band Effective Mass Approximations

Advanced Mathematical Models and Numerical Techniques

Springer

Editors
Matthias Ehrhardt
Lehrstuhl für Angewandte
Mathematik und Numerische Analysis
Bergische Universität Wuppertal
Wuppertal, Germany

Thomas Koprucki
Forschungsgruppe "Partielle
Differentialgleichungen"
Weierstraß-Institut für Angewandte
Analysis und Stochastik
Berlin, Germany

ISSN 1439-7358 ISSN 2197-7100 (electronic)
ISBN 978-3-319-01426-5 ISBN 978-3-319-01427-2 (eBook)
DOI 10.1007/978-3-319-01427-2
Springer Cham Heidelberg New York Dordrecht London

Library of Congress Control Number: 2014943912

Mathematics Subject Classification (2010): 34L40, 35J10, 35Q41, 65L15, 65N06, 65N30, 81Q05

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Foreword

With the rapid development of nano sciences, the structural properties of semiconductors can be arranged on an atomic scale. This has led to massive down scaling of electronic integrated circuits, compact semiconductor light sources with highest power densities, and tiny sensors that monitor various physical properties in complex environments. This development will continue in the future, and with ongoing progress in the synthesis of complex nanostructures, the inclusion of a wider field of chemical elements and as result even more functionality and better performance will be feasible.

At the heart of understanding the electronic, optical, or magnetic properties in nanostructures is the dispersion relation for electrons and holes. It represents a quantum mechanical property of the electron, namely the energy versus its wave vector. The latter can also be viewed as momentum, using the de Broglie relation. The dispersion relation contains a plethora of information, namely the phase velocity (which in classical electromagnetics is related to the refractive index), the group velocity, and the effective mass, only to name few. In free space, solving Schrödinger's equation for a single electron gives the well-known parabolic dispersion relation. Now if the electron is located in a semiconductor crystal, it is surrounded by a periodic arrangement of nuclei, a large number of core and valence electrons. Inclusion of the respective potentials in the Hamiltonian leads to a large coupled many-particle quantum mechanical problem, which, for nanostructures, cannot be solved with current numerical or analytical methods.

It is due to three formidable approximations that we can study the physical properties of nanostructures with the sophisticated mathematical and numerical methods that are presented in this book. First, the core electrons of the fully occupied orbitals can be lumped together with the nuclei, which leads to potentials of ionic cores. This removes all the equations for the core electrons from the system. The next simplification is called Born-Oppenheimer approximation: the ion cores are much heavier than the remaining valence electrons. Therefore, they move much slower and are basically stationary to the electrons. As a consequence, the electronic

properties can be calculated using fixed nuclei positions, and the nuclei dynamics can be separated into an interaction Hamiltonian. At last, the so-called mean field approximation treats all valence electrons (except the electron of interest) as average background potential. That way, the electron–electron coupling can be treated by a single uncoupled effective potential, and the remaining equation resembles a single particle Schrödinger equation.

A fundamental property of this single particle equation for crystals is the periodicity of the ionic core potentials, which leads directly to Bloch wave functions as solutions, and a separation of the Hamiltonian into a part independent of the wave vector and dependent on the wave vector, containing a **k·p** term (therefore the name **k·p** is sometimes used for this equation). The presence of the periodic potential introduces band gaps in the dispersion relation and a plethora of significant deviations from the free-electron case. Time-dependent or stationary perturbation theory can be applied to solve the **k·p** Schrödinger equation in an elegant fashion, where the terms containing the wave vector are treated again as perturbation in the Hamiltonian. Hence, the solution is more exact for small wave vector magnitudes, depending on the order of perturbation included. In order to study realistic nanostructures, many more perturbations need to be added to the single particle Hamiltonian. These perturbations and the mathematics and numerics to solve the resulting Schrödinger equation is the subject of this book. They include the band-to-band coupling, spin–orbit interaction, the presence of hetero interfaces, mechanical strain, and surfaces or carrier scattering and their statistics. This way, the electronic dispersion relation (band structure) or even the carrier dynamics of complex semiconductor nanostructures can be calculated with high numerical efficiency. It gives us information such as the effective masses, the strength and energies of optical transitions or the spin–orbit interaction, or the density of states for charge carriers, which are of fundamental importance to understand electronic, optical, or magnetic properties in nano devices.

The book starts with three chapters on the physical models, from a multi-band description aiming at quantum transport properties of carriers within the multi-band formalism, to a focus on state-of-the-art **k·p** models for quantum dots, emphasizing symmetry considerations. The second part is devoted to numerical methods for solving the **k·p** type equation framework, with one chapter on the finite element method, and the second one on the plane wave expansion. In the third part, applications of the **k·p** method are presented, demonstrating the capabilities of the framework for describing challenging but nonetheless realistic situations in band structure calculations. In the final chapter, advanced mathematical topics are discussed, such as a time-dependent effective mass multi-band formalism dealing with carrier dynamics, and the topic of transparent boundary conditions for termination of the simulation domain.

The reader of this book will gain a detailed insight into the status of the multi-band effective mass method for semiconductor nanostructures. Both users of the **k·p** method and advanced researchers who want to advance the **k·p** method

further will find helpful information to work with this method and use it as a tool to characterize the physical properties of semiconductor nanostructures.

Kassel, Germany
June 2014

Bernd Witzigmann

Preface

The operational principle of modern semiconductor nanostructures, such as quantum wells, quantum wires, or quantum dots, relies on quantum mechanical effects. The goal of numerical simulations using quantum mechanical models in the development of semiconductor nanostructures is threefold: First, they are needed for a deeper understanding of experimental data and of the operational principle. Second, is to predict and optimize in advance qualitative and quantitative properties of new devices in order to minimize the number of prototypes needed. Semiconductor nanostructures are embedded as an active region in semiconductor devices. Finally, the results of quantum mechanical simulations of semiconductor nanostructures can be used by upscaling methods to deliver parameters needed in semi-classical models for semiconductor devices such as quantum well lasers. This book covers in detail all these three aspects using a variety of illustrating examples.

Multi-band effective mass approximations have been increasingly attracting interest over the last decades, since it is an essential tool for effective models in semiconductor materials. This book is concerned with several mathematical models from the most relevant class of $\mathbf{k}\cdot\mathbf{p}$-Schrödinger Systems. We will present both mathematical models and state-of-the-art numerical methods to solve adequately the arising systems of differential equations. The designated audience is graduate and Ph.D. students of mathematical physics, theoretical physics and people working in quantum mechanical research or semiconductor/opto-electronic industry who are interested in new mathematical aspects.

The principal audience of this book is graduate and Ph.D. students of (mathematical) physics, research lecturer of mathematical physics, and research people working in semiconductor, opto-electronic industry for a professional reference.

Wuppertal, Germany — Matthias Ehrhardt
Berlin, Germany — Thomas Koprucki
June 2014

Contents

List of Contributors

Aleksey D. Andreev Hitachi Cambridge Laboratory, Cambridge, UK

Anton Arnold Institut für Analysis und Scientific Computing, Technische Universität Wien, Wien, Austria

Luigi Barletti Dipartimento di Matematica e Informatica "U.Dini", Università di Firenze, Firenze, Italy

Dieter Bimberg Institut für Festkörperphysik, Technische Universität Berlin, Berlin, Germany

Stefan Birner Walter Schottky Institute and Physics Department, Technische Universität München, Garching, Germany

Institute for Nanoelectronics, Technische Universität München, München, Germany

Matthias Ehrhardt Bergische Universität Wuppertal, Lehrstuhl für Angewandte Mathematik und Numerische Analysis, Wuppertal, Germany

Giovanni Frosali Dipartimento di Matematica e Informatica "U.Dini", Università di Firenze, Firenze, Italy

Gerald Hönig Institut für Festkörperphysik, Technische Universität Berlin, Berlin, Germany

Dirk Klindworth Technische Universität Berlin, DFG Research Center MATHEON, Berlin, Germany

Thomas Koprucki Forschungsgruppe "Partielle Differentialgleichungen", Weierstraß–Institut für Angewandte Analysis und Stochastik, Berlin, Germany

Oliver Marquardt Tyndall National Institute, Cork, Ireland

Paul-Drude Institute, Berlin, Germany

Omar Morandi Institut de Physique et Chimie des Matériaux de Strasbourg, Université de Strasbourg, CNRS UMR 7504, Strasbourg Cedex 2, France

Eoin P. O'Reilly Tyndall National Institute, Cork, Ireland

Andrei Schliwa Institut für Festkörperphysik, Technische Universität Berlin, Berlin, Germany

Stefan Schulz Tyndall National Institute, Cork, Ireland

Sebastian Steiger Network for Computational Nanotechnology, Purdue University, West Lafayette, IN, USA

Stanko Tomić Joule Physics Laboratory, School of Computing, Science and Engineering, University of Salford, Manchester, UK

Ratko G. Veprek Condensed Matter Theory Group, School of Physics, University of Sydney, Sydney, NSW, Australia

Nenad Vukmirović Scientific Computing Laboratory, Institute of Physics Belgrade, University of Belgrade, Belgrade, Serbia

Andrea Zisowsky Institut für Mathematik, Technische Universität Berlin, Berlin, Germany

Part I
Physical Models

Chapter 1
Kinetic and Hydrodynamic Models for Multi-Band Quantum Transport in Crystals

Luigi Barletti, Giovanni Frosali, and Omar Morandi

Abstract This chapter is devoted to the derivation of **k·p** multi-band quantum transport models, in both the pure-state and mixed-state cases. The first part of the chapter deals with pure-states. Transport models are derived from the crystal periodic Hamiltonian by assuming that the lattice constant is small, so that an effective multi-band Schrödinger equation can be written for the envelopes of the wave functions of the charge carriers. Two principal approaches are presented here: one is based on the Wannier-Slater envelope functions and the other on the Luttinger-Kohn envelope functions. The concept of Wannier functions is then generalized, in order to study the dynamics of carriers in crystals with varying composition (heterostructures). Some of the most common approximations, like the single band, mini-bands and semi-classical transport, are derived as a limit of multi-band models. In the second part of the chapter, the mixed-state (i.e. statistical) case is considered. In particular, the phase-space point of view, based on Wigner function, is adopted, which provides a quasi-classical description of the quantum dynamics. After a theoretical introduction to the Wigner-Weyl theory, a two-band phase-space transport model is developed, as an example of application of the Wigner formalism to the **k·p** framework. The third part of the chapter is devoted to quantum-fluid models, which are formulated in terms of a finite number of macroscopic moments of the Wigner function. For mixed-states, the maximum-entropy closure of the moment equations is discussed in general terms. Then, details are given on the

L. Barletti • G. Frosali (✉)
Dipartimento di Matematica e Informatica "U.Dini", Università di Firenze, Viale Morgagni 67/a, 50134 Firenze, Italy
e-mail: barletti@math.unifi.it; giovanni.frosali@unifi.it

O. Morandi
Institut de Physique et Chimie des Matériaux de Strasbourg, Université de Strasbourg, CNRS UMR 7504, 23 rue du Loess, BP 43, 67034 Strasbourg Cedex 2, France
e-mail: morandi@dipmat.univpm.it

M. Ehrhardt and T. Koprucki (eds.), *Multi-Band Effective Mass Approximations*, Lecture Notes in Computational Science and Engineering 94,
DOI 10.1007/978-3-319-01427-2_1,

multi-band case, where "multi-band" is to be understood in the wider sense of "multi-component wave function", including therefore the case of particles with spin or spin-like degrees of freedom. Three instances of such systems, namely the two-band **k**·**p** model, the Rashba spin-orbit system and the graphene sheet, are examined.

1.1 Introduction

The derivation of mathematical models for particle transport in semiconductor devices of last generation cannot discard quantum effects. In particular, quantum dynamics involving two or several energy bands of a semiconductor, forces to consider "quantum multi-band models", that are the subject of the present chapter. As an example, particle transitions between the conduction and the valence band states could in some case increase considerably the peak-to-valley ratio of the resonant current through a double-barrier. Resonant interband tunneling diodes (RITD) are examples of devices which exploit this phenomenon; they are of paramount importance in nanotechnology for their applications to high-speed miniaturized systems [101, 118].

Quantum multi-band models have largely been formulated and analyzed in the recent past (see references throughout this chapter). Similarly to other models for semiconductor devices, they can be divided in two main classes: pure-state (non statistical) and mixed-state (statistical) models. The former are based on wave functions, and, therefore, on Schrödinger equation, while the latter require a density-operator framework which can be more conveniently formulated in terms of the phase-space formalism provided by Wigner functions. Such a quasi-classical description provides some advantages in terms of simplicity in the physical interpretation and the availability of feasible method for the inclusion of irreversible processes like thermalization or phase breaking mechanisms. Another category of models is that of quantum-fluid models which, analogously to classical fluids, are formulated in terms of a finite number of macroscopic moments of the Wigner function. The equations of a quantum fluid, therefore, can be deduced from the underlying phase-space description.

In this chapter, we describe the multi-band models that have recently been formulated in both classes (statistical and non-statistical). Attention is given to the definitions of the relevant quantities which characterize each model and to the advantages and disadvantages of each model compared to others. The technical details of the derivations of the various models, as well as the rigorous proofs of consistency and existence of the solutions, are diverted directly to the papers where the models have been described and that are cited in the chapter.

The chapter is organized as follows. In Sect. 1.2 we briefly recall the **k**·**p** envelope-function theory. We first discuss the Wannier-Slater approach and, secondly, the Luttinger-Kohn approach. Since in many practical cases the dynamics of carriers in crystals with varying composition (heterostructures) is of considerable

interest, then a section is devoted to non-uniform materials, introducing the concept of generalized Wannier functions. Since the microscopic simulation of the particle transport in a fully quantum many-band context is extremely complex and usually unfeasible, the last part of Sect. 1.2 is devoted to review a few theoretical works where some of the most common approximations, like the single band, mini-bands and semi-classical transport, are derived as a limit of a multi-band model. In Sect. 1.3 we deal with the statistical kinetic models based on the Wigner-function approach. The multi-band **k·p** models can be easily adapted to the Wigner-Weyl framework, providing a quasi-classical description of the quantum motion which offers some advantages in terms of simplicity in the physical interpretation. As an example of the application of the Wigner formalism to the **k·p** framework, a two-band model is considered. Section 1.4 is devoted to quantum fluid models. After a short review of quantum hydrodynamics in the single-band/scalar case, we turn to the multi-band case, indicating by this term a general framework where discrete degrees of freedom are involved. Rather than developing a general theory (which is probably beyond to come), we shall examine three instances, namely the two-band **k·p** model, the Rashba spin-orbit system and the transport of electrons on a graphene sheet.

1.2 Envelope k·p Models

The study of periodic crystalline solids leads naturally to a representation of the electronic ground state in terms of Bloch waves. They are the Schrödinger wave functions of particles with a given momentum that interact with the periodic "frozen" ionic lattice. As an alternative, the particle motion can be represented in terms of localized orbitals or Wannier functions. They are formally defined in terms of a suitable unitary transformation applied to the Bloch orbital. In the simplest case the link between the Bloch and the Wannier waves is the Fourier transform from the reciprocal to the direct lattice space. The Wannier representation is a useful starting point for various formal developments, such as the semiclassical theory of electron dynamics or more generally, the envelope function methods [74].

Many different methods have been employed for the study of the electronic structure and the optical properties of the semiconductors. In particular, the envelope function, the tight-binding, the pseudo-potential, and the density functional methods constitute the most common approaches. For the simplicity and ease of interpretation, the envelope-function method is one of the most developed approach for modeling the electrons motion in bulk semiconductor and in heterostructures. The envelope-function approach is a quite general theory and applies both to periodic and quasi-periodic materials. Various details concerning the physical composition of the sample and the band structures can be easily included. Moreover, the envelope-function method represents the theoretical basis of the common effective-mass approximation [5]. It provides a detailed description of conduction and valence subbands near the center of the Brillouin zone and it is widely used to calculate

confined-state energies and subband structure of III–V and II–VI semiconductor systems. A simple and rather complete description of the spectrum of a bulk semiconductor around the band gap includes the conduction band and the three highest valence bands, namely the heavy-hole, light-hole, and spin-split-off bands. More complex semiconductors like strained-layer heterostructures, strained crystals and indirect-gap material are characterized by a more structured valence band. A variety of envelope models, from two to a total of eight bands are found in literature. All the electronic properties of the materials enter in the model via few macroscopic parameters (Luttinger-Kohn parameters) that are easily characterized by direct measurement on the bulk crystals [13, 36].

One of the main interest in the study of multi-band models is the possibility to reproduce the band-to-band tunneling. This phenomenon explains the transition of a particle between two disconnected parts of the band diagram. In many situations, the various branches of the energy spectrum of a semiconductor belong to spectral regions which are separated by a certain energy gap (forbidden region). Particles can travel from one band to one other (for example from the conduction to the valence band) by two distinguished mechanisms: (1) overcoming the energy gap by gaining or loosing energy; (2) tunneling the energy gap. The first mechanism is an incoherent process that involves the interaction of the particle with some other particle or dynamical field. The second mechanism takes place when a static field (and in this case this phenomenon is known as Zener tunneling) or a (usually strong) discontinuity in the ionic lattice is present. The particle transition from one band to the other can be easily characterized from a mathematical point of view. Let Ψ represent the Schrödinger wave function of a particle in the solid. The conservation of the total particle probability is expressed by the normalization of the L^2 norm $\|\Psi(t)\|_{L^2} = 1$ for each time t. The description of the particle motion, in terms of a multi-band approach, requires the projection of the wave function in a suitable orthonormal set of given functions $\phi_{n,\mathbf{R}}$

$$\Psi(\mathbf{x}, t) = \sum_{n,\mathbf{R}} f_n(\mathbf{R}, t)\phi_{n,\mathbf{R}}(\mathbf{x}) \ .$$

The index n is denoted band index and $\mathbf{R}$ are the vectors of the lattice sites. The new unknowns of the problem $f_n(\mathbf{R}, t)$ are defined envelope functions. The conservation of the total probability and the orthonormality of the $\phi_{n,\mathbf{R}}$ imply that the sum over n and $\mathbf{R}$ of the squared modulus of the envelope functions is constant in time $\sum_n \|f_n\|^2_{\ell^2_{\mathbf{R}}} = 1$, where $\|f_n\|^2_{\ell^2_{\mathbf{R}}} \equiv \sum_{\mathbf{R}} |f_n(\mathbf{R}, t)|^2$. In general $\frac{\partial}{\partial t}\|f_n\|_{\ell^2_{\mathbf{R}}} \neq 0$ holds true. Since $\|f_n\|^2_{\ell^2_{\mathbf{R}}}$ is interpreted as the probability to find the particle in the n-th band, the previous relation is the mathematical formulation of the band transition process. In particular, the situation where there exists an index n such that $\frac{\partial}{\partial t}\|f_n\|_{\ell^2_{\mathbf{R}}} \simeq 0$, is indicated by single band motion.

In the following sections we introduce some of the most common $\mathbf{k}\cdot\mathbf{p}$ models.

1.2.1 Wannier-Slater Envelope Functions Approach

The motion of electrons in a semiconductor is governed by the effective one-electron potential $U(\mathbf{x})$, generated by the crystal ions with the periodicity property

$$U(\mathbf{x}+\mathbf{R}) = U(\mathbf{x})$$

for all vectors $\mathbf{R}$ of the Bravais lattice $\mathscr{L}$ of the crystal. The total single particle Hamiltonian is $H = \frac{p^2}{2m} + U(\mathbf{x})$, where $\mathbf{p} = -i\hbar\nabla$ is the momentum operator and $p = |\mathbf{p}|$, m is the electron mass, and $\hbar$ is the Planck's constant over 2π. From the Bloch's theorem it is well-known that the one-electron Hamiltonian H has a complete system of eigenfunctions $|n,\mathbf{k}\rangle$ with eigenvalues $E_n(\mathbf{k})$ (that in the solid state framework are usually defined energy bands)

$$H\,|n,\mathbf{k}\rangle = E_n(\mathbf{k})\,|n,\mathbf{k}\rangle\ . \tag{1.1}$$

In the position representation the Bloch functions can be written in the form

$$b_n(\mathbf{x},\mathbf{k}) = \mathrm{e}^{i\mathbf{k}\cdot\mathbf{x}}u_{n,\mathbf{k}}(\mathbf{x}) \equiv \langle\mathbf{x}|n,\mathbf{k}\rangle, \tag{1.2}$$

where $\mathbf{k}$ is the quasimomentum (or crystal momentum) running over the first Brillouin zone B. The first Brillouin zone is a special primitive cell in reciprocal space $\mathscr{L}^*$, defined as all points $\mathbf{k}\in\mathbb{R}^3$ which are closer to the Γ point than to any other point of $\mathscr{L}^*$. The Bloch wave are completely characterized by their behavior in a single Brillouin zone, in fact

$$u_{n,\mathbf{k}}(\mathbf{x}+\mathbf{R}) = u_{n,\mathbf{k}}(\mathbf{x})\,.$$

The Bloch functions are normalized as

$$\int_{\mathbb{R}^3} b_n(\mathbf{x},\mathbf{k})b_n(\mathbf{x},\mathbf{k}')\,\mathrm{d}\mathbf{x} = |B|\delta_{nn'}\delta(\mathbf{k}-\mathbf{k}')\,,$$

where $|B|$ denotes the measure of the Brillouin zone. Hereafter, the delta function with continuous variable denotes the Dirac's distribution and with discrete suffix denotes the Kronecker's delta. The set of the Bloch functions provides a (generalized) basis of the L^2 space. Any function $\Psi\in L^2$ can be thus expanded as

$$\Psi(\mathbf{x}) = \sum_n \int_B \varphi_n(\mathbf{k})b_n(\mathbf{x},\mathbf{k})\,\mathrm{d}\mathbf{k}, \tag{1.3}$$

where

$$\varphi_n(\mathbf{k}) = \int_{\mathbb{R}^3} \overline{b}_n(\mathbf{x},\mathbf{k})\Psi(\mathbf{x})\,\mathrm{d}\mathbf{x}\,, \tag{1.4}$$

with the bar denoting complex conjugation.

Following [114], we define the Wannier-Slater envelope functions by the Fourier transform of $\varphi_n(\mathbf{k})$

$$f_n(\mathbf{x}) = \frac{1}{(2\pi)^{3/2}} \int_B \varphi_n(\mathbf{k})\, \mathrm{e}^{i\mathbf{x}\cdot\mathbf{k}}\, \mathrm{d}\mathbf{k}\,. \tag{1.5}$$

If Ψ represents the particle wave function in the crystal, the square modulus of $f_n(\mathbf{x})$ is proportional to the probability to find the particle in the n-th energy band. We remark that, by definition, the Fourier transform of the Wannier envelope functions f_n is compactly supported in the first Brillouin zone. Consequently, any spatial oscillation with wave length smaller than the lattice constant cannot be present in f_n. For that reason, the envelope functions can be considered a class of intrinsically smooth functions. On the contrary, the original function Ψ and the Bloch functions $b_n(\mathbf{x}, \mathbf{k})$ may display high oscillations induced by the periodic ionic potential. When projected on the Wannier-Slater basis, these high frequency oscillations are filtered to high energy bands. One of the advantages of the envelope function theory is that for many purposes, the knowledge of the full wave function is not longer necessary and only few bands have to be considered. The relationship between the Wannier envelope functions and the full wave function is given by

$$f_n(\mathbf{x}) = \int_{\mathbb{R}^3} \overline{a}_n(\mathbf{x}', \mathbf{x})\Psi(\mathbf{x}')\, \mathrm{d}\mathbf{x}'\,, \tag{1.6}$$

where the Wannier functions a_n are defined by

$$a_n(\mathbf{x}, \mathbf{x}') = \frac{1}{(2\pi)^{3/2}} \int_B b_n(\mathbf{x}, \mathbf{k})\mathrm{e}^{-i\mathbf{x}'\cdot\mathbf{k}}\, \mathrm{d}\mathbf{k}\,. \tag{1.7}$$

Conversely, the wave function in terms of the Wannier functions is given by

$$\Psi(\mathbf{x}) = \sum_n \frac{1}{|B|} \int_{\mathbb{R}^3} a_n(\mathbf{x}, \mathbf{x}')\, f_n(\mathbf{x}')\ \mathrm{d}\mathbf{x}'. \tag{1.8}$$

In this section we discuss the derivation of the Wannier-Slater model, that essentially consists in replacing as unknown of the problem the original wave function Ψ with the envelope functions defined in Eq. (1.5). According to (1.1), an ideal homogeneous material is completely described by the Bloch functions. The utility of the expansion given in (1.8) becomes evident when an external non periodic field is added to the ionic potential. For the sake of simplicity, we consider the effect of a purely electrostatic potential denoted by V. In particular, V may take in account different effects, like the device energy-band offset for the heterojunctions, the bias voltage applied across the device, the contribution from the doping impurities and from the self-consistent field produced by the mobile electronic charge.

The dynamics of the Wannier envelope functions can be deduced from the Schrödinger equation for Ψ and Eq. (1.6). Following [1] (see also [14] for the details of the derivation), we have

$$i\hbar \frac{\partial f_n}{\partial t}(\mathbf{x},t) = \tilde{E}_n(-i\nabla)\, f_n(\mathbf{x},t) + \sum_{n'} \int_{\mathbb{R}^3} V^W_{nn'}(\mathbf{x},\mathbf{x}')\, f_{n'}(\mathbf{x}',t)\ \mathrm{d}\mathbf{x}'. \tag{1.9}$$

Here,

$$V^W_{nn'}(\mathbf{x},\mathbf{x}') = \frac{1}{|B|}\int_{\mathbb{R}^3} \overline{a}_n(\mathbf{y},\mathbf{x})\, V(\mathbf{y})\, a_n(\mathbf{y},\mathbf{x}')\ \mathrm{d}\mathbf{y} \tag{1.10}$$

are matrix-elements of the external potential with respect to the Wannier functions and $\tilde{E}_n(-i\nabla)$ are pseudo-differential operators associated to the energy bands with a cut-off outside the Brillouin zone, namely

$$\tilde{E}_n(-i\nabla)\, f_n(\mathbf{x}) = \frac{1}{(2\pi)^3}\int_B E_n(\mathbf{k})\, f_n(\mathbf{x}')\, \mathrm{e}^{i\mathbf{k}\cdot(\mathbf{x}-\mathbf{x}')}\ \mathrm{d}\mathbf{x}'\ \mathrm{d}\mathbf{k}.$$

The system (1.9) defines an infinite hierarchy of coupled equations. As stated before, the envelope function f_n represents the probability to find the electron in the n-th band. This is equivalent to say that the f_n envelope functions describe particles whose energy is around the band edge E_n. For this reason it seems reasonable to ignore the contribution of all the envelope functions of the remote bands $E_{n'}$ such that $|E_n - E_{n'}| \gg 1$. We remark that, despite this argument is roughly correct, in some case the remote bands could influence significatively the particle motion around the Fermi energy [63, 115].

1.2.1.1 Approximated Dynamics

In order to introduce some simple and quite general approximations it is convenient to write the set of equations (1.9) in the Fourier transformed space. We obtain

$$i\hbar \frac{\partial \varphi_n}{\partial t}(\mathbf{k}) = E_n(\mathbf{k})\varphi_n(\mathbf{k}) + \sum_{n'}\int_B \langle n,\mathbf{k}|V|n',\mathbf{k}'\rangle \varphi_n(\mathbf{k}')\, d\mathbf{k}', \tag{1.11}$$

where we use the Dirac's bracket notation in order to compact the notation. After some algebra the previous system can be recast as (see [84] for more details)

$$i\hbar \frac{\partial \varphi_n}{\partial t}(\mathbf{k}) = E_n(\mathbf{k})\varphi_n(\mathbf{k}) + \int_B \tilde{V}(\mathbf{k}-\mathbf{k}')\varphi_n(\mathbf{k}')\, d\mathbf{k}' \tag{1.12}$$

$$-i\frac{\hbar^2}{m_0}\sum_{n'\neq n}\int_B \tilde{V}(\mathbf{k}-\mathbf{k}')\varphi_{n'}(\mathbf{k}')\frac{(2\pi)^3}{\Omega}\int_{\text{u-cell}} \overline{u_{n,\mathbf{k}}}(\mathbf{x})\frac{\mathbf{k}-\mathbf{k}'}{\triangle E_{n,n'}}\cdot \nabla u_{n'}(\mathbf{k}',\mathbf{x})\, d\mathbf{x}\, d\mathbf{k}',$$

where $\tilde{V}$ denotes the Fourier transform of the potential V, Ω is the volume of the unitary cell (u-cell) and

$$\triangle E_{n,n'}(\mathbf{k},\mathbf{k}') \equiv E_{n'}(\mathbf{k}') - E_n(\mathbf{k}) + \frac{\hbar^2}{2m_0}\left(k'^2 - k^2\right) .$$

This set of equations is still too complex for practical applications. Anyway, it is in a form that can be easily simplified. The idea is to expand with respect to $\mathbf{k}$ the last term of Eq. (1.12). In particular, at the leading order we get

$$-i\sum_{n'\neq n}\frac{\hbar^2\mathbf{P}_{n,n'}}{m_0\triangle E_{n,n'}}\cdot\int_B\left(\mathbf{k}-\mathbf{k}'\right)\tilde{V}(\mathbf{k}-\mathbf{k}')\varphi_{n'}(\mathbf{k}')\,d\mathbf{k}' , \tag{1.13}$$

where

$$\mathbf{P}_{n,n'} = \frac{(2\pi)^3}{\Omega}\int_{\text{u-cell}}\overline{u_{n,\mathbf{0}}}(\mathbf{x})\nabla u_{n',\mathbf{0}}(\mathbf{x})\,\mathrm{d}\mathbf{x} . \tag{1.14}$$

After the $\mathbf{k}$ expansion is performed, the set of equations can be restored in terms of the original variable f_n by applying the inverse Fourier transform. We give here the final result in the simplest case where only two bands ("conduction" and "valence") are taken into account and the parabolic band approximation (with positive curvature for the conduction and negative for the valence band) is assumed

$$\begin{cases} i\hbar\dfrac{\partial f_c}{\partial t} = -\dfrac{\hbar^2}{2m_c^*}\,\Delta f_c + Vf_c + E_c f_c - \dfrac{\hbar^2\mathbf{P}_{c,v}\cdot\nabla V}{m_0 E_g}\,f_v \\ i\hbar\dfrac{\partial f_v}{\partial t} = \dfrac{\hbar^2}{2m_v^*}\,\Delta f_v + Vf_v + E_v f_v - \dfrac{\hbar^2\mathbf{P}_{c,v}\cdot\nabla V}{m_0 E_g}\,f_c \end{cases} . \tag{1.15}$$

Here m_n^* is the effective mass in the n-band. These equations describe the intraband dynamics and contain an interband coupling, proportional to the momentum matrix element $\mathbf{P}$, that is responsible for tunneling between different bands induced by the applied electric field proportional to the gradient of V.

1.2.2 Luttinger-Kohn Envelope Functions

A different definition of envelope function was given by Luttinger and Kohn [70] (see also [15, 30]). The crucial observation is that the set of the periodic functions $u_{n,\mathbf{k}}(\mathbf{x})$ is complete for each fixed value of the quasi-momentum $\mathbf{k}$. For that reason, all the previous expansion formulae are still valid if we substitute formally $u_{n,\mathbf{k}}(\mathbf{x})$ with $u_{n,\mathbf{0}}(\mathbf{x})$. More precisely, Eq. (1.8) becomes

$$u_{n,\mathbf{k}} \ : \Psi(\mathbf{x}) = \frac{1}{|B|(2\pi)^{3/2}} \sum_n \sum_j f_n(\mathbf{R}_j) \int_B u_{n,\mathbf{k}}(\mathbf{x}) \mathrm{e}^{-i\left(\mathbf{x}-\mathbf{R}_j\right)\cdot\mathbf{k}} \, \mathrm{d}\mathbf{k}$$

$$u_{n,\mathbf{k}} \to u_{n,\mathbf{0}} \ : \Psi(\mathbf{x}) = \frac{1}{|B|} \sum_n \sum_j f_n(\mathbf{R}_j) u_{n,\mathbf{0}}(\mathbf{x}) \int_B \mathrm{e}^{-i\left(\mathbf{x}-\mathbf{R}_j\right)\cdot\mathbf{k}} \, \mathrm{d}\mathbf{k}$$

$$\Psi(\mathbf{R}_i) = \sum_n \sum_j \delta_{i,j} f_n(\mathbf{R}_j) u_{n,\mathbf{0}}(\mathbf{R}_i) = \sum_n f_n(\mathbf{R}_i) u_{n,\mathbf{0}}(\mathbf{R}_i)$$

where the $\mathbf{R}_i$ denote the atomic sites and we substituted the integral over the space with a sum over the atomic sites (which gives the correct normalization). Interpolation over all the space gives the expansion formula

$$\Psi(\mathbf{x}) = \sum_n f_n(\mathbf{x}) u_{n,\mathbf{0}}(\mathbf{x}) \ . \tag{1.16}$$

This formula is the starting point of the Luttinger-Kohn (L-K) procedure. The Schrödinger equation takes the form

$$i\hbar \frac{\partial \varphi_n(\mathbf{k})}{\partial t} = \sum_{n'} \int_B \mathscr{H}^{kp}_{nn'}(\mathbf{k}, \mathbf{k}') \, \varphi_{n'}(\mathbf{k}') \, \mathrm{d}\mathbf{k}' \tag{1.17}$$

where, coherently with Eq. (1.5), $\varphi_n(\mathbf{k})$ is the Fourier transform of the envelope function $f_n(\mathbf{x})$. The Hamiltonian matrix elements are

$$\mathscr{H}^{kp}_{nn'}(\mathbf{k}, \mathbf{k}') = \left[\left(E_n + \frac{\hbar^2 k^2}{2m_0} \right) \delta_{nn'} + \frac{\hbar}{m_0} \mathbf{k} \cdot \mathbf{P}_{nn'} \right] \delta(\mathbf{k} - \mathbf{k}') + \tilde{V}(\mathbf{k} - \mathbf{k}') \delta_{nn'} \ . \tag{1.18}$$

In order to proceed, it is convenient to consider the quasi-unitary transformation Θ that diagonalizes the Hamiltonian in the momentum space up to second order in $\mathbf{k}$. The new Hamiltonian reads

$$\mathscr{H}^{LK} = \Theta^{-1} \mathscr{H}^{kp} \Theta \ , \tag{1.19}$$

the matrix elements of Θ are

$$\langle n, \mathbf{k} | \Theta | n', \mathbf{k}' \rangle_{kp} = \left(\delta_{nn'} - \frac{\hbar}{m_0} \frac{\mathbf{P}_{nn'} \cdot \mathbf{k}}{\Delta E_{nn'}} \right) \delta(\mathbf{k} - \mathbf{k}') \, , \tag{1.20}$$

where $\langle \mathbf{x} | n, \mathbf{k} \rangle_{kp} = u_{n,0}(\mathbf{x}) e^{i\mathbf{k}\mathbf{x}}$. Equation (1.19) defines a basis rotation. The elements of the new basis correspond to an expansion of the $u_{n,\mathbf{k}}(\mathbf{x})$ functions up to the first order in $\mathbf{k}$

$$\langle \mathbf{x} | n, \mathbf{k} \rangle_{LK} = \mathrm{e}^{i\mathbf{k}\cdot\mathbf{x}} \left(u_{n,\mathbf{0}}(\mathbf{x}) + \mathbf{k} \frac{\partial u_{n,\mathbf{0}}(\mathbf{x})}{\partial \mathbf{k}} \bigg|_{\mathbf{0}} \right) \ . \tag{1.21}$$

Applying the Fourier transform, we obtain the evolution equation for the Luttinger-Kohn envelope function. As a final result, we report here the complete Luttinger-Kohn four band model (we ignore the spin degeneracy) for direct band gap semiconductors like GaAs. This model takes into account two light and two heavy holes branches. The equation of motion takes the form of a Schrödinger equation with effective Hamiltonian. The expansion given in (1.16) describes the projection of the solution on the four degenerate Bloch wave functions at the center of the Brillouin zone. The set of the four envelope functions can be arranged in a vector

$$F = (f_{3/2}, f_{1/2}, f_{-1/2}, f_{-3/2}) ,$$

where the $\pm 3/2$ and the $\pm 1/2$ components describe the heavy and the light hole band, respectively (see [20] for more details). The Hamiltonian of the system is

$$\mathscr{H}^{LK} = \begin{pmatrix} P+Q & -S & R & 0 \\ -S^* & P-Q & 0 & R \\ R^* & 0 & P-Q & S \\ 0 & R^* & S^* & P+Q \end{pmatrix} , \tag{1.22}$$

where the symbol $*$ denotes the formal adjoint. The matrix elements are given in terms of the Luttinger parameters $\gamma_1, \gamma_2, \gamma_3$ [35]

$$\begin{aligned} P &= \frac{\hbar^2}{2m}\gamma_1 \Delta , \\ Q &= \frac{\hbar^2}{2m}\gamma_2 \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} - 2\frac{\partial^2}{\partial z^2} \right) , \\ R &= \frac{\hbar^2}{2m}\sqrt{3}\left[-\gamma_2 \left(\frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2} \right) + 2i\gamma_3 \frac{\partial}{\partial y}\frac{\partial}{\partial x} \right] , \\ S &= \frac{\hbar^2}{m}\sqrt{3}\gamma_3 \left(\frac{\partial^2}{\partial x^2} - i\frac{\partial^2}{\partial y^2} \right) \frac{\partial}{\partial z} . \end{aligned}$$

1.2.3 Non Uniform Materials and Generalized Wannier Functions

In the previous sections, the derivation of the **k·p** models have been focused on the description of a uniform bulk material in the presence of some small perturbation field. In many practical cases, the crystal translation symmetry is broken (or considerably affected) by various factors like the presence of a strong external electric or magnetic field, impurity, or, in graded mixed semiconductor and graded heterojunctions, by the variation of chemical composition. An important

theoretical effort was devoted to understand how the simple models based on the effective mass equation should be modified in order to incorporate all the aforementioned effects. The first attempts to describe the dynamics of carriers in crystals with varying composition was based on the phenomenological assumptions of a position-dependent energy gap and of a position-dependent effective mass. In this approach, the particles move in a force field, the strength of which depends on the band they occupy [68, 112]. Although this assumption could be considered reasonable for systems that deviate slightly from the ideal case (like for example in semiconductors with highly diluted impurities), serious problems arise when abrupt junctions between two different materials are considered. In this case, a common approximation consists in discarding the difference between the band edge Bloch functions in the two different materials. The heterostructures are thus characterized in terms of their bulk properties alone. Theoretical studies proved the foundation of such an hypothesis for slowly varying perturbation and for some particular abrupt junctions [48].

The ground properties of a non-uniform material and the study the interface effects could be in principle deduced by a microscopic approach based for example on the density functional theory. Anyway, the computational cost demanded for the solution of such models is extremely high and the inclusion of dynamical effects is still unfeasible. One of the main difficulties connected with the study of the interfaces is to provide an accurate approximation of the Bloch functions between the two media. A number of envelope-function models incorporating this difference have been proposed in recent years, but most of these requires extensive microscopic calculations. A widely used approach assumes a "symmetrized" form of the contact Hamiltonian. Few examples of different solutions to the question are given in [4, 30, 80, 85]. The description of position-dependent material properties is most easily attached by using a localized basis in the direct space such as the Wannier functions presented in the previous section. The Wannier functions play a central role in qualitative as well as quantitative aspects of the one-electron theory of solids. Differing from the elements of the Bloch basis, the Wannier functions are not eigenfunctions of the Hamiltonian. They have the theoretical advantage that in many cases they are exponentially localized around a given lattice cell. This particularly useful property makes the Wannier functions an interesting object for the description of a wide class of complex materials [59]. In particular, based on the theory of ordinary differential equations, Kohn in 1959 showed a rigorous result of exponential localization of the Wannier functions for a one-dimensional lattice [64]. An important contribution toward the generalization to higher dimension was given by Cloizeaux that proved the exponential localization in arbitrary dimension for a single isolated Bloch band [37]. One of the difficulties for a possible extension of this result that consider many bands relies on the lake of regularity around the regions where the bands cross [91]. For more recent achievements for two and three dimensions insulators see [31, 94]. Despite these theoretical results, it is anyway quite difficult to construct such a maximally localized Wannier basis. Indeed, for the application of the envelope function method to complex electronic devices, other approaches are also considered in literature. As an example, in the following,

we will discuss the Foreman approach to non-uniform lattices [49]. In this method, the evolution equations for the envelope functions take the form of local, spatially varying, **k**·**p** equations, with additional terms describing the explicit effects of the interface or inhomogeneity. In order to describe a graded crystal, Foreman defined the following quasi-periodic ionic potential

$$\mathscr{W}(\mathbf{x}) = \frac{1}{\Omega(2\pi)^d} \sum_i \mathrm{W}(\mathbf{x}, \mathbf{R}_i) \int_B e^{i(\mathbf{R}_i - \mathbf{x})\mathbf{k}} \, \mathrm{d}\mathbf{k} \, . \tag{1.23}$$

Here, the potential $\mathrm{W}(\mathbf{x}, \mathbf{R}_i)$ is periodic with respect the first variable, $\mathrm{W}(\mathbf{x}+\mathbf{R}_j, \mathbf{R}_i) = \mathrm{W}(\mathbf{x}, \mathbf{R}_i)$ for all i, j and the $\mathbf{R}_i$ run on the atomic sites. The idea of this representation is that the second variable of W labels the macroscopic position on the crystal lattice while the first gives the fine resolution around each Wigner-Seitz cell. The integral on the quasi-momentum in (1.23), can be viewed as a continuous function approximating the Dirac's delta in $\mathbf{R}_j$ (more precisely, it tends to a Dirac's delta when the first Brillouin zone extends to infinity). With this remark, it becomes evident that the real potential $\mathscr{W}(\mathbf{x})$ is nearly equal to the value of the periodic potential $\mathrm{W}(\mathbf{x}, \mathbf{R}_i)$ with $\mathbf{R}_i$ equal to the coordinate of the cell containing $\mathbf{x}$. By using this definition it is possible to describe a large class of non-homogeneous media (more details are given in [48]). In particular, in this approach the existence of a common symmetry group of $\mathrm{W}(\mathbf{x}, \mathbf{R}_i)$ for all $\mathbf{R}_i$ and, consequently, a unique first Brillouin are assumed.

The essential element of the Foreman approach is to obtain a localized-in-space basis that extends the definition given by Luttinger and Kohn. The new basis functions are defined as

$$\chi_n(\mathbf{x}, \mathbf{R}_j) = u_{n,\mathbf{0}}\left(\mathbf{x}, \mathbf{R}_j\right) \frac{1}{(2\pi)^d} \int_B e^{i(\mathbf{R}_j - \mathbf{x})\mathbf{k}} \, \mathrm{d}\mathbf{k} \, , \tag{1.24}$$

where, coherently with (1.16), the periodic function $u_{n,\mathbf{0}}$ is the center-band periodic part of the Bloch wave related to a "virtual" bulk lattice obtained by the periodical rearrangement of the $\mathbf{R}_j$ Wigner-Seitz cells. More precisely, $u_{n,\mathbf{0}}(\mathbf{x}, \mathbf{R}_j)$ is obtained by the eigenvalue equation

$$\left[-\frac{\hbar^2}{2}\Delta_{\mathbf{x}} + \mathrm{W}(\mathbf{x}, \mathbf{R}_i)\right] u_{n,\mathbf{0}}\left(\mathbf{x}, \mathbf{R}_j\right) = E_n^j \, u_{n,\mathbf{0}}\left(\mathbf{x}, \mathbf{R}_j\right) \, . \tag{1.25}$$

It is possible to prove that the wave function ψ can be expanded in the extended L-K basis as

$$\psi(\mathbf{x}, t) = \sum_{n,j} f_n(\mathbf{R}_j, t)\chi_n(\mathbf{x}, \mathbf{R}_j) \, . \tag{1.26}$$

Since the aim of the use of the Foreman approach is to provide a theoretical basis for the description of the particle motion in a graded semiconductor, here we

consider only the single band transport. In this case the evolution equation takes the form of an effective mass equation with variable coefficients. After cumbersome calculations we obtain

$$i\hbar \frac{\partial f_n}{\partial t} = -\nabla \frac{\hbar^2}{2m^*(\mathbf{x})} \nabla f_n + E_n(\mathbf{x}) f_n \,, \tag{1.27}$$

where m^* denotes the effective mass of the particles and the envelope functions $f_n(\mathbf{x})$ are the $f_n(\mathbf{R}_j)$ given in (1.26), interpolated over the continuous space $\mathbb{R}^3$. As a matter of fact, the envelope functions are coupled together by various terms, that for sake of simplicity are discarded in the previous equation. They take into account the possibility that a particle passes from one band to the other. They are expressed by cumbersome expressions which involve the mismatching between the periodic Bloch functions of different virtual crystals.

1.2.4 *Application of* **k·p** *Models to Heterostructures and Resonant Tunneling*

The progress of the lithographic technology and design of new miniaturized devices lead to a dramatic reduction of the field effect transistor (FET) gate lengths. Reaching the nanometric dimensions, the transport in FETs suffers from inevitable limitations due to the emergence of various quantum effects such as the interband tunneling, particle depletion and gate-oxide tunneling. Despite these quantum effects constitute a disadvantage for the conventional single band devices, they open new possibility for designing tunneling based devices. In particular, the interband tunneling diodes (IRTD) have been considered as promising components for future circuit technologies both for high speed applications and optoelectronic devices. The physics of the tunneling phenomenon in semiconductor heterostructures has been a subject of considerable investigation since the early work of Tsu and Esaki [110]. A number of theoretical works based on the **k·p** technique has been developed for understanding and simulating the behavior of tunneling structures [35,39,57,65,92].

Intraband resonant-tunneling diodes are normally grown by molecular-beam epitaxy in III–V materials, typically using GaAs/GaAlAs materials. By using these materials, devices characterized by high-frequency response and peak-to-valley current ratios particularly large have been designed. For this reason, interband-based devices are extremely attractive for use in high-frequency oscillators, logic circuits, and a variety of other digital and analog applications. Their high performances are the result of some specifically designed transport mechanisms that involve interband tunneling between conduction and valence-band states. These devices are constituted by a succession of layers of different materials connected by abrupt junctions. The special feature of IRTD is that the resulting band diagram has the shape of a quantum well in conduction (valence) band where the energy of the bounded states is lower (higher) than the top (bottom) of the valence (conduction)

band of the contacts [39, 57]. The study of the transport properties of such a devices requires the application of multidimensional quantum transport models (see Sect. 1.2.5.2 below).

1.2.5 Some Limits: Single and Mini-Band Transport

The microscopic simulation of the particle transport in a fully quantum many-band context is extremely complex and usually unfeasible. For that reason, the study of some particular dynamical regimes where approximated model can be derived, is of special interest. In the following, we will consider few theoretical works where some of the most common approximations like the single band, mini-bands and semi-classical transport are derived as a limit of a multi-band model. We will focus on their application to low-dimensional structures.

1.2.5.1 Single Band Limit and Non-Parabolicity

Single band transport is one of the most common assumption for the particle motion in solid state physics. It states that in the presence of regular electric and magnetic external fields, the particle motion is confined in a single energy band. With this approximations the motion of electrons in a periodic potential is replaced by the motion of a fictitious particle in vacuum but with a modified mass called the effective mass of the electron.

All the equations of motion previously derived, (1.9), (1.15), (1.22) and (1.27), show that the single band motion is just the simplest approximation of a **k**·**p** model, where the interband coupling terms are neglected. This formal assumption can be made rigorous in some appropriate asymptotic limit. In a series of works, Teufel and coworkers developed a formal setting where the limit of the single band dynamics arises under the fondant hypothesis that there exists an energy gap that delimitates one or a selected group of bands [105]. If initially the single particle wave function is spanned by the Bloch functions of these isolated bands, in the limit of an infinitely high energetic separation, the same is true for all time. This limit is generally defined as adiabatic decoupling. For the applications, it can be considered a sufficiently accurate approximation when the external fields are weak compared to the fields generated by the ionic cores [31, 106]. A different approach is taken in [15] where the electrons are assumed to move in both a periodic potential and a macroscopic one. The typical period of the crystal lattice potential is assumed to be small compared with the typical length scale of the macroscopic potential (homogenization limit). They show that in this case the exact **k**·**p** dynamics is well approximated (in strong sense) by an effective mass model.

Despite the adiabatic or homogenization limit leads to the single band motion, the effects of the other bands cannot be completely discarded. The presence of the remote bands enters in the single band limit through the definition of the explicit

form of the single band parameters like for example the particle effective mass or the Kane non-parabolicity. In particular, one of the first relevant prediction of the **k·p** theory was the strong mixing of heavy and light holes away from the gamma point, that in many cases gives rise to high non-parabolicities on the holes dynamics.

1.2.5.2 Devices with Reduced Dimensionality

Concerning the electronic applications, apart from the IRDT devices which exploit resonant interband tunneling, one other interesting class of quantum devices for high frequency oscillators or fast switching commutators, is constituted by the high electron mobility transistor (HEMT). These devices use quantum confinement in order to collect the electrons in a surface with reduced dimensionality (one or two dimension). The particles travel from the source to the drain contact by the application of an electric field tangent to the low-dimensional surface. In these conditions, the mobility of the carrier increases considerably. Similar situation is found in devices like ultrashort channel doublegate MOSFETs, where electrons are extremely confined in the direction transverse to the current flow.

Concerning the mathematical modeling of such a devices, one of the strategies that takes advantage of this reduction of dimensionality is to treat differently the particle diffusion process along the direction orthogonal to interface from the transport along the current directions. Differing from the IRTD, which are intrinsically many-band devices, the transport in FET channel with reduced dimensionality is essentially single band. The quantum confinement produces a series of mini-band (typically, for practical application, six or eight are significant [75]) along which the particles move. From symmetry reasons, band transitions from a mini-band to the other are strongly suppressed and require scattering with phonons. An interesting and still open question is the mathematical characterization of the mechanism in which the three dimensional particle gas present at the contacts (source or drain) becomes one or two-dimensional and enters in the channel [33].

One of the first mathematical study of multi-band models for the simulation of such kind of devices, was presented in [21]. There, a quite general study of the coupled Schrödinger-Poisson model describing electrostatic interaction and confinement was considered. In particular, this approach puts light on the delicate mathematical problem of the possible existence of bounded resonant states embedded in the continuous spectrum generated by a countable set of semi-infinite contacts. More specifically, the problem of considering the transition from a 3D to a 2D gas from a mathematical point of view, is addressed in [76, 77]. There a lattice where the crystal symmetry is broken by the presence of a potential barrier is considered. Techniques of singular perturbation applied to the three-dimensional Schrödinger-Poisson system show the decomposition of the particle motion along two orthogonal directions.

These theoretical studies have also some practical implications. In fact, by exploiting this domain decomposition, the computational cost of the simulation of a real device could be significantly reduced. The 3D quantum system is replaced by 1D Schrödinger equations (describing the bounded states) coupled to 2D equations

(for the transport). This consideration was the starting point of various theoretical works focused on the accurate simulation of the particle motion in confined gas. In many practical cases, simple models based on drift-diffusion equations are able to capture and reproduce quantitatively the main particle transport characteristics. Anyway, in the modern integrated electron devices, the scale length of the active region approaches the distance between two successive carrier interactions with the crystal. The particles move in a medium where scattering processes are not dominant. In this situation, the basic assumptions of the hydrodynamic models are no longer valid and a more accurate description of the particle lattice interaction is necessary. In an intermediate level, where quantum mechanically coherent effects can be discarded, a good characterization of the motion of interacting particles is generally achieved by the direct application of kinetic Boltzmann models [55]. The usual approach for the study of the transport along the channel consists in the application of a series of semi-classical Boltzmann equation (one for each mini-band). multi-band models like **k·p** provide the details of the mini-bands dispersion relations (a complete review of the application of the Boltzmann mini-band dynamics in superlattices can be found in [28]). The application of Boltzmann transport benefits from the solid knowledge acquired in the semiclassical device physics, especially in the physics of scattering. Many numerical methods are developed in the last decades for the determination of approximated solutions. Among them, the deterministic and the Monte Carlo methods revealed itself to be particularly efficient.

A deterministic Boltzmann method means essentially that the Boltzmann scattering kernel is directly discretized and numerically solved [10]. These methods derive from the multigroup approaches that were originally designed for the reduction of the impracticable numerical complexity of the problem [95]. One of the advantages of this approach is that it is completely free from statistical errors and noise that characterize some less expensive methods like the Monte Carlo. Deterministic models are applied to semiconductors with reduced dimensionality like for example Si based MOSFET [11] or to some new materials which are structurally 1D like carbon nanotube [12] or 2D like graphene [69]. These models provide a very accurate and complete physical characterization of the quantum mechanical phenomena occurring in the device. However, the intensive computational cost makes it still unsuitable for commercial device simulators. In order to reduce the computational complexity, Monte Carlo methods provide a good alternative to the hydrodynamics models [87, 88]. The numerical techniques used in the Monte Carlo simulators reflect the underlying physics of the transport of classical carriers. The scattering process is view in terms of a sequence of alternating free flight and instantaneous scattering events that change abruptly the momentum of the particles. As mentioned before, the parabolic band approximation may be inadequate for the simulations when a more sophisticated description of the energy bands is warranted. In [97, 98] the non-parabolicity effects have been consistently included via the approximated the Kane's formula. Numerical Monte Carlo simulations show that other important thermal effects like the heat generation or hot particle current could play a significant role on the electronic performances of the devices [45, 89].

1.3 Wigner Approach

1.3.1 Introduction

In 1932 Eugene Wigner introduced a quantum-mechanical phase-space distribution (usually defined quasi-probability) with the purpose of studying the quantum correction of the thermodynamic equilibrium for a gas in the presence of external fields [117]. Since then, the Wigner function formalism has been successfully applied to many different domains, from solid state dynamics to quantum optics. One of the major advantage of this approach is the possibility to express the quantum dynamics in a phase-space formulation, easily comparable with the classical analogue. The quantum mechanical quasi-distribution, however, is not positive definite. From a physical point of view the lack of positiveness of the Wigner distribution should be ascribed to the presence of quantum interference effects. They induces some coherent correlation between different regions of the phase-space. Technically, the Wigner function is defined through a Fourier transform combined with a simple change of variables in the density matrix. Moyal discovered the existence of a rich theoretical structure behind this transformation [86]. Nowadays, the quantum phase-space is considered to be an alterative logically complete formulation of the quantum mechanics and it is generally referred as Wigner-Weyl quantization procedure. The theory considers only summable functions defined on the classical phase-space. All the physical quantities are computed through phase-space integrals of suitable functions weighted by the Wigner distribution (method of the momenta).

More into details, in the Wigner-Weyl theory both states and observables are represented by functions of the phase-space coordinates. One key concept is the association of any quantum mechanical operator with a phase-space counterpart, called *symbol*, which is a complex-valued function on phase-space (correspondence principle). The correspondence between the operator $\hat{\mathscr{A}}$ and the function $\mathscr{A}(\mathbf{x},\mathbf{p})$ is provided by the map $\mathrm{Op}_\hbar[\mathscr{A}] = \hat{\mathscr{A}}$

$$\left(\hat{\mathscr{A}}\psi\right)(\mathbf{x}) = \left(\mathrm{Op}_\hbar\left[\mathscr{A}\right]\psi\right)(\mathbf{x}) = \frac{1}{(2\pi\hbar)^3}\int \mathscr{A}\left(\frac{\mathbf{x}+\mathbf{y}}{2},\mathbf{p}\right)\psi(\mathbf{y})\, e^{\frac{i}{\hbar}(\mathbf{x}-\mathbf{y})\cdot\mathbf{p}}\,\mathrm{d}\mathbf{y}\,\mathrm{d}\mathbf{p}\,, \tag{1.28}$$

called *Weyl quantization* [47]. Here, ψ is a generic function. Hereafter, in order to avoid confusion, we make distinction between the quantum mechanical operator and the associated symbol by denoting the first with the hat. The inverse of $\mathrm{Op}_\hbar$ is given by the Wigner transform

$$\mathscr{A}(\mathbf{x},\mathbf{p}) = \mathrm{Op}_\hbar^{-1}\left[\hat{\mathscr{A}}\right](\mathbf{x},\mathbf{p}) = \int \mathscr{K}_{\hat{\mathscr{A}}}\left(\mathbf{x}+\frac{\boldsymbol{\eta}}{2},\mathbf{x}-\frac{\boldsymbol{\eta}}{2}\right)e^{-\frac{i}{\hbar}\mathbf{p}\cdot\boldsymbol{\eta}}\,\mathrm{d}\boldsymbol{\eta}\,, \tag{1.29}$$

where $\mathscr{K}_{\hat{\mathscr{A}}}(\mathbf{x},\mathbf{y})$ is the formal kernel of the operator $\hat{\mathscr{A}}$. With this language, the Wigner function is just the phase space counterpart of the density operator

$\hat{\rho}$, describing the state of the system (the kernel $\mathscr{K}_{\hat{\rho}}$ being the density matrix). Explicitly,

$$f(\mathbf{r},\mathbf{p}) = \frac{1}{(2\pi)^3}\int \left\langle \mathbf{x}+\frac{\boldsymbol{\eta}\hbar}{2}\,\middle|\,\hat{\rho}\,\middle|\,\mathbf{x}-\frac{\boldsymbol{\eta}\hbar}{2}\right\rangle e^{-ip\cdot\eta}\,\mathrm{d}\boldsymbol{\eta}\,. \tag{1.30}$$

According to the correspondence principle, the classical position and momentum variables are associated to the quantum Hermitian operators $\hat{\mathbf{x}}$ and $\hat{\mathbf{p}}$. They satisfy the usual quantum commutators

$$\begin{aligned}\hat{\mathbf{x}} &= \mathrm{Op}_\hbar^{-1}(\mathbf{x})\,,\\ \hat{\mathbf{p}} &= \mathrm{Op}_\hbar^{-1}(\mathbf{p})\,,\\ [\hat{\mathbf{x}},\hat{\mathbf{p}}] &= i\hbar\,.\end{aligned}$$

One important property of the Weyl map is to provide an algebra for the phase-space functions in terms of the non-commutative #-product defined as (the arrows indicate on which operator the gradients act)

$$\begin{aligned}\mathscr{A}\#\mathscr{B} &\equiv \mathscr{A}\; e^{\frac{i\hbar}{2}\left(\overleftarrow{\nabla}_{\mathbf{x}}\cdot\overrightarrow{\nabla}_{\mathbf{p}}-\overleftarrow{\nabla}_{\mathbf{p}}\cdot\overrightarrow{\nabla}_{\mathbf{x}}\right)}\;\mathscr{B}\\ &= \sum_n \left(\frac{i\hbar}{2}\right)^n \frac{1}{n!}\mathscr{A}(\mathbf{x},\mathbf{p})\left[\overleftarrow{\nabla}_{\mathbf{x}}\cdot\overrightarrow{\nabla}_{\mathbf{p}}-\overleftarrow{\nabla}_{\mathbf{p}}\cdot\overrightarrow{\nabla}_{\mathbf{x}}\right]^n\mathscr{B}(\mathbf{x},\mathbf{p})\,,\end{aligned} \tag{1.31}$$

with the property

$$\mathrm{Op}_\hbar^{-1}\left[\hat{\mathscr{A}}\;\hat{\mathscr{B}}\right] = \mathscr{A}\#\mathscr{B}. \tag{1.32}$$

Basic notions of the operator quantum mechanics are formulated in the phase space with the help of the #-product. The expectation value of the operator $\hat{\mathscr{A}}$ can be obtained by the expected value of the function $\mathscr{A}(\mathbf{r},\mathbf{p})$ under the "measure" f

$$\mathrm{Tr}\left\{\hat{\mathscr{A}}\;\hat{\rho}\right\} = \int \mathscr{A}(\mathbf{x},\mathbf{p})\,f(\mathbf{x},\mathbf{p},t)\;\mathrm{d}\mathbf{p}\,\mathrm{d}\mathbf{x}\,. \tag{1.33}$$

Concerning the application of the Wigner formalism to electronic devices, the first momenta of the Wigner quasi-distribution are particularly relevant. In fact, the particle density n is given by the first momentum of the Wigner function with respect to the $\mathbf{p}$ variable

$$n(\mathbf{x},t) = \int_{\mathbb{R}^3} f(\mathbf{x},\mathbf{p},t)\,\mathrm{d}\mathbf{p}\,, \tag{1.34}$$

and the current density $\mathbf{j}$ is given by the second momentum

$$\mathbf{j}(\mathbf{x},t) = -\frac{q}{m}\int_{\mathbb{R}^3} \mathbf{p}\, f(\mathbf{x},\mathbf{p},t)\, \mathrm{d}\mathbf{p}. \tag{1.35}$$

More details on the Wigner-Weyl correspondence principle can be found in [67,90]. One of the simplest cases of application of the Wigner formalism is the study of the ballistic transport in the effective mass approximation. In this case, the Wigner function defined in Eq. (1.30) evolves according to the following equation

$$\frac{\partial f}{\partial t} + \frac{\mathbf{p}}{m^*}\cdot\nabla_{\mathbf{x}} f + \frac{i}{\hbar}\Theta[\delta V] f = 0, \tag{1.36}$$

where m^* is the electron effective mass in the selected band and

$$(\Theta[\delta V] f)(\mathbf{x},\mathbf{p}) = \frac{1}{(2\pi\hbar)^3}\int \mathrm{e}^{-i(\mathbf{p}-\mathbf{p}')\xi/\hbar}\, \delta V(\mathbf{x},\xi)\, f(\mathbf{x},\mathbf{p}')\, \mathrm{d}\xi\, \mathrm{d}\mathbf{p}' \tag{1.37}$$

is a pseudo-differential operator with symbol

$$\delta V(\mathbf{x},\xi) = V\left(\mathbf{x}+\frac{\xi}{2}\right) - V\left(\mathbf{x}-\frac{\xi}{2}\right). \tag{1.38}$$

Equation (1.36) describes the motion of an ensemble of independent electrons moving in a semiconductor crystal in the presence of an external field and it is denoted *quantum transport equation*. The essential difference between the classical and the Wigner evolution equation is given by the presence of the nonlocal potential term $\Theta[\delta V]$, while in classical mechanics the electrostatic potential is treated as a localized force.

The similarity between classical and quantum phase-space transport has been widely exploited in order to obtain corrections to the classical evolution equations that incorporate some quantum effects. Although the full quantum treatment of scattering is usually impracticable (non-Markovian processes with memory effects could affect the ultrafast response of a quantum gas in a strongly out-of-equilibrium regime), under some reasonable approximations, the particle dynamics becomes essentially a classical process and the Wigner collision operator simplifies into the well-known Boltzmann collision operator [32, 66]. Other simpler and more phenomenological approaches based on the relaxation time approximation have been considered [60, 72]. It is worth mentioning that the Wigner formalism can be obtained also as a particular case of the more general Green many-body formalism developed by Kadanoff and Baym (see i.e. [58]). This framework can be taken as an alternative starting point for the derivation of Wigner kinetic models that include relaxation processes like particle-particle scattering [60]. The classical-quantum analogy of the Wigner function approach allows to describe in a natural way open-boundary systems where carrier and energy is exchanged with the environment [51]. Anyway, some criticisms has been raised concerning this point [102]. It is still an

open question whether the classical inflow boundary conditions which are generally assumed for the Wigner function (and are mathematically well posed [73]) are always appropriate for modeling the particle reservoirs.

The numerical solution of the Wigner dynamics encounters many difficulties. The typical approach is the application of a specifically designed finite-difference method [51] or splitting methods [8, 44]. Recently, the affinity of Quantum Monte Carlo method is successfully applied [93, 99]. An extensive description of the affinity method can be found in [90].

1.3.2 Multi-Band Wigner Models

In Sect. 1.2 we introduced various **k·p** models for the quantum motion in a crystal. All of these approaches are described in terms of an effective Hamiltonian acting on the Schrödinger multi-component envelope function. As an alternative, the multi-band **k·p** models can be easily adapted to the Wigner-Weyl framework. The use of a quasi-classical description of the quantum motion leads some advantages in terms of simplicity in the physical interpretation and the availability of feasible methods for the inclusion of irreversible processes like thermalization or phase breaking mechanisms [43, 111, 116].

As an example of the application of the Wigner formalism to the **k·p** framework, we consider the two-band model given by Eq. (1.15). For the sake of simplicity, in Sect. 1.3.1 we discussed the application of the Wigner-Weyl theory to scalar fields. Anyway, the theory is completely general and applies without modification also to spinorial or vectorial fields (for more details see [17, 83]). In our case, the unknown is the two-component vector $(f_c(\mathbf{x}), f_v(\mathbf{x}))$ constituted by the conduction and valence envelope functions. Generally speaking, an electron gas is a quantum mechanical mixture and is conveniently described by the two-by-two density matrix

$$\rho(\mathbf{x},\mathbf{y}) = \sum_j \lambda_j \begin{pmatrix} f_c^j(\mathbf{x})\overline{f_c^j}(\mathbf{y}) & f_c^j(\mathbf{x})\overline{f_v^j}(\mathbf{y}) \\ f_v^j(\mathbf{x})\overline{f_c^j}(\mathbf{y}) & f_v^j(\mathbf{x})\overline{f_v^j}(\mathbf{y}) \end{pmatrix} , \tag{1.39}$$

where f_c^j, f_v^j are given functions and specify the valence and conduction band components of the solution. The statistical coefficients λ_j are normalized to one, $\sum_j \lambda_j = 1$, and are the statistical probability that the physical state is represented by the vector (f_c^j, f_v^j). The simplest way to apply the quantum phase-space theory, is to define the Wigner function f^W of the system by using (1.30). The evolution equation of the f^W is obtained by the formal application of the map $\mathrm{Op}_\hbar^{-1}$ given in (1.29) to the von Neumann equation

$$i\hbar\,\mathrm{Op}_\hbar^{-1}\frac{\partial\hat{\rho}}{\partial t} = \mathrm{Op}_\hbar^{-1}\left[\hat{\rho}, \widehat{\mathscr{H}_{env}}\right] , \tag{1.40}$$

where $\widehat{\mathscr{H}_{env}}$ is the Hamiltonian of the two-band system

$$\widehat{\mathscr{H}_{env}} = \begin{pmatrix} -\dfrac{\hbar^2}{2m_c^*}\,\Delta + V + E_c & -\dfrac{\hbar^2 \mathbf{P}\cdot\nabla V}{m_0 E_g} \\ -\dfrac{\hbar^2 \mathbf{P}\cdot\nabla V}{m_0 E_g} & \dfrac{\hbar^2}{2m_v^*}\Delta + V + E_v \end{pmatrix}. \tag{1.41}$$

By using Eq. (1.32), Eq. (1.40) gives the evolution equation for f^W

$$i\hbar \frac{\partial f^W}{\partial t} = f^W \# \mathscr{H}_{env} - \mathscr{H}_{env} \# f^W. \tag{1.42}$$

Despite this straightforward application of the Wigner-Weyl theory could be used in order to deduce the particle dynamics (see for example [52]), it is usually more convenient to proceed differently. In the literature various approaches are presented [18, 29]. All of them share the same basic idea: search for a suitable basis set of projectors that simplify the Wigner dynamics. This makes sometime the derivation of the Wigner model a little involved but provides the advantage to describe the particle gas dynamics by more transparent equations. According to [81, 82], the quantum phase-space formulation of the two-band envelope model is given by the following set of equations

$$\frac{\partial h_c}{\partial t} = -\nabla_{\mathbf{p}}\mathscr{H}^+ \cdot \nabla_{\mathbf{x}} h_c + \nabla_{\mathbf{x}}\mathscr{H}^+ \cdot \nabla_{\mathbf{p}} h_c - 2\xi\, \Re(h_{cv}) \tag{1.43}$$

$$\frac{\partial h_v}{\partial t} = -\nabla_{\mathbf{p}}\mathscr{H}^- \cdot \nabla_{\mathbf{x}} h_v + \nabla_{\mathbf{x}}\mathscr{H}^- \cdot \nabla_{\mathbf{p}} h_v + 2\xi\, \Re(h_{cv}) \tag{1.44}$$

$$\frac{\partial h_{cv}}{\partial t} = -\frac{i}{\hbar}\left(\mathscr{H}^+ - \mathscr{H}^-\right) h_{cv} + \nabla_{\mathbf{x}} V \cdot \nabla_{\mathbf{p}} h_{cv} + \xi\,(h_c - h_v) \tag{1.45}$$

where

$$\mathscr{H}^{\pm}(\mathbf{x},\mathbf{p}) = \pm\sqrt{P_R^2 + \Omega^2} + V,$$

$$\Omega(\mathbf{p}) = \frac{E_g}{2} + \frac{|\mathbf{p}|^2}{2m^*}, \qquad P_R(\mathbf{x}) = -\hbar\frac{\mathbf{P}\cdot\nabla V}{E_g m_0}$$

$E_g = E_c - E_v$ is the band gap, $\Re$ denotes the real part and

$$\xi = \frac{P_R}{P_R^2 + \Omega^2}\,\frac{\nabla V \cdot \mathbf{p}}{m^*}.$$

Equations (1.43)–(1.45) are correct only up to the second order in $\hbar$, the exact equations of motion can be found in [83]. When $\xi = 0$ (no band coupling), the Wigner system simplifies into two decoupled Liouville semi-classical evolution equation for

a gas of particles in conduction (valence) band with classical distribution h_c (h_v). This simple consideration puts in evidence the strong similarity between the Wigner framework and the classical dynamics. The distribution h_{cv} is a quantum-mechanical object and has not classical counterpart. It describes the band-to-band interference effects and accounts for the band tunneling phenomena. A detailed description of the physical meaning of the system of equations (1.43)–(1.45) can be found in [81]. Anyway, as remarked before, the single band limit is usually not trivial. In fact, when the external field is not negligible some mulSi descrivono le band-to-band effetti di interferenza e spiegare i fenomeni della band tunnelingtliband corrections still modify the simple single band effective mass approximation. According to (1.43) and (1.44), inside each band, the particles follow the trajectories defined by the scalar Hamiltonian $\mathscr{H}^{\pm}$ (the plus sign accounts for the conduction and the minus the valence band). Explicitly,

$$\begin{cases} \dot{\mathbf{x}} &= \pm \dfrac{1}{\sqrt{1+\left(\frac{P_R}{\Omega}\right)^2}} \dfrac{\mathbf{p}}{m^*} \\ \dot{\mathbf{p}} &= \nabla V \pm \dfrac{1}{\sqrt{1+\left(\frac{\Omega}{P_R}\right)^2}} \dfrac{\hbar}{m_0 E_g} \nabla \left(\nabla V \cdot \mathbf{P}\right). \end{cases} \tag{1.46}$$

This shows that the Wigner evolution system reduces to the classical Newton equations for a particle with mass m^* when $\nabla V \to 0$ or $E_g \to \infty$.

1.4 Hydrodynamic Models

1.4.1 Introduction

The technological progress made in the last decade on the fabrication of microelectronics devices downscaled the active regions of integrated transistors to the submicrometer and ultra submicrometer dimensions. Carriers travel from the source to the drain over a distance of less than 100 nm and experience rapidly changing transport conditions. New interesting questions emerge concerning the quantum transport in such extremal conditions.

Induced by strong electric field, dramatic out-of-equilibrium regimes could appear and particles have enough kinetic energy to reach regions of the band diagram where the parabolic band approximation is not longer valid. The details of the band structure should be taken into account and a full-quantum out-of-equilibrium approach becomes necessary. Moreover, high-field transport is essentially nonlocal in both space and time and simpler particle methods like the conventional drift-diffusion do not apply. Toward the final goal of a full quantum description of the dynamics, many intermediate steps should be achieved. The first

approaches to describe an electron-hole gas in a semiconductor devices were based on some macroscopic thermodynamic considerations. During the last decades, a continuous increasing of the degrees of sophistication generates a quite variegated class of models, starting from the drift-diffusion to the high order hydrodynamic models. The main differences concern the number of equations and macroscopic quantities considered, and the assumptions made for the description of a local or global equilibrium. At the easier level, the evolution of the mean charge density and current are characterized by the presence of a local microcanonical equilibrium (drift-diffusion models). On the other limit, the most complete models maintain a high number of unknowns (despite the inclusion of higher moments beyond the particle energy and energy flux leads to extremely cumbersome expressions) and are generally defined hydrodynamics approaches. Particle motion in a solid state structure displays generally a irreversible dynamics and the system is driven toward some equilibrium state (at least locally). These energy dispersive phenomena are included in a macroscopic approach through the so called closure relationship. The basis idea is to require the fulfilling of some detailed balance equation. This can be rigorously obtained via a suitable asymptotic limit of the Boltzmann kernel (Chapman-Enskog procedure). An equation relating the particle density and energy is found. For the modern devices, the relaxation processes take place in a time scale comparable with the switching time. This is a complex regime and many phenomena cannot be captured by the original Chapman-Enskog limit. Various extensions of the Chapman-Enskog procedure have been developed for the description of realistic devices. In particular, after the (seminal) contribution given in [6, 7], different models based on the semi-classical dynamics and denoted "extended thermodynamics" have been proposed. The application to the semiconductor physics [2] and the mathematical structure based on this thermodynamic approach have been deeply investigated [3]. Further extensions are also present in literature. They are based on the idea of quantum entropy [108, 109], and include phenomena which are intrinsically of quantum mechanical nature like multi-band motion and particle statistics.

In the following sections we will give the details of the quantum fluid-dynamic approach to multi-band systems, where "multi-band" has to be intended in a wide sense and stands for "quantum systems with multi-component wave function". Systems of this kind include the envelope and $\mathbf{k}\cdot\mathbf{p}$ models treated in the previous sections but, of course, also include particles with spin or spin-like degrees of freedom.

1.4.2 Scalar Quantum Fluid Equations

Before treating the multi-band/spinorial case, let us review the single-band/scalar case.

We can define Quantum Fluid Dynamics (QFD) as the mathematical description of a quantum system in terms of macroscopic, fluid-like, variables (such as density,

current, temperature, etc.). The origin of QFD dates back to 1926 [71] when E. Madelung realized that Schrödinger equation can be put in the fluid-dynamic form:

$$\begin{cases} \dfrac{\partial n}{\partial t} + \dfrac{\partial}{\partial x_i}(n u_i) = 0, \\ \dfrac{\partial u_i}{\partial t} + u_j \dfrac{\partial u_j}{\partial x_i} + \dfrac{\partial V}{\partial x_i} - \dfrac{\hbar^2}{2} \dfrac{\partial}{\partial x_i} \dfrac{\Delta \sqrt{n}}{\sqrt{n}} = 0, \end{cases} \tag{1.47}$$

where the probability density n and current $n\mathbf{u}$ are related to the wave function ψ by $\psi = \sqrt{n}\, \mathrm{e}^{iS/\hbar}$ and $n\mathbf{u} = \nabla S$ (and the particle mass was set equal to 1). Here and thereafter we adopt the Einstein summation convention on repeated indices. Madelung equations have the form of an irrotational, compressible and isothermal Euler system with an additional term, of order $\hbar^2$, interpreted as a "quantum potential" or a "quantum pressure". This was later named *Bohm potential* after D. Bohm, who based on it his celebrated, although controversial, interpretation of quantum mechanics, the so-called *Bohmian mechanics* [25, 26, 46].

Madelung equations hold for a *pure*, i.e. non-statistical, state, and are equivalent to Schrödinger equation. The natural extension of Madelung's theory, therefore, would be a fluid equivalent of the von Neumann equation (i.e. the Schrödinger equation for statistical, or *mixed*, states), but such extension can hardly be achieved. The situation can be better understood in terms of Wigner functions (see Sect. 1.3.1). Indeed, the fluid variables correspond to moments of the Wigner function, e.g.,

$$\begin{aligned} n(\mathbf{x}, t) &= \frac{1}{(2\pi\hbar)^d} \int_{\mathbb{R}^d} w(\mathbf{x}, \mathbf{p}, t)\, d\mathbf{p} = \varrho(\mathbf{x}, \mathbf{x}, t) \\ n u_k(\mathbf{x}, t) &= \frac{1}{(2\pi\hbar)^d} \int_{\mathbb{R}^d} p_k w(\mathbf{x}, \mathbf{p}, t)\, d\mathbf{p} = \frac{\hbar}{2i} \left(\frac{\partial \varrho}{\partial x_k} - \frac{\partial \varrho}{\partial y_k} \right) (\mathbf{x}, \mathbf{x}, t), \end{aligned} \tag{1.48}$$

(where the corresponding expressions in terms of the time-dependent density matrix $\varrho(\mathbf{x}, \mathbf{y}, t)$ have also been shown). This suggests that Madelung equations can also be derived in a "kinetic way" by taking the moments of the Wigner equation. In this way, assuming a pure-state initial Wigner function, Madelung equations are indeed obtained [56]. However, for a general Wigner function, extra moments appear that cannot be expressed in terms of n and $\mathbf{u}$. More in general, taking whatever set of moments of Wigner equation will always result into equations containing moments that cannot be reduced to the initially chosen set.

Of course, such situation is not at variance with the analogous situation of classical statistical mechanics: any marginalization of the phase-space distribution reduces the information, so that Liouville or Boltzmann equations cannot be equivalent to any finite set of moment equations. A fluid description, nevertheless, can be recovered by means of suitable fluid asymptotics and closure assumptions.

The analogy with the classical case suggests a possible approach to QFD:

Quantum fluid equations can be derived from Wigner equations in exactly the same way as classical fluid equations can be derived from Boltzmann equation.

The application of this idea implies starting from a quantum equivalent of Boltzmann equation, which is represented by the Wigner equation (1.36)[1] endowed with a collisional term:

$$\frac{\partial w}{\partial t} + \mathbf{p} \cdot \nabla_x w + \Theta_\hbar[V]w = Q(w).$$

Unfortunately, it is by no means clear how to construct a collisional operator *à la Boltzmann* in the quantum case. At least, the most reasonable proposals [34] are far too complicated to be really useful (involving e.g. nonlocality-in-time effects). On the other hand, similarly to what happens in the classical case, in order to derive a fluid asymptotics it is not necessary to know all the details of collisions but only fairly general features, like the conservation properties. Then, a reasonable and effective choice for the collisional term is that of a BGK-like operator [9, 24]:

$$\frac{\partial w}{\partial t} + \mathbf{p} \cdot \nabla_x w + \Theta_\hbar[V]w = \frac{g[w] - w}{t_c}. \tag{1.49}$$

Here, t_c is a typical collisional time (assumed to be constant, for the sake of simplicity) and $g[w]$ is a Wigner function representing the local equilibrium reached by the system because of collisions. The reason for the notation "$g[w]$" is that, as we shall see in details below, the local equilibrium is assumed to share a set of moments with w.

The central question has now been shifted to $g[w]$: how should it be chosen? A general answer to this question is furnished by the (quantum version of) the Maximum Entropy Principle (QMEP) [41, 42], that, in a rather generic way, can be stated as follows:

QMEP: $g[w]$ is the most probable state compatible with the observed macroscopic moments of w,

where "most probable" means that $g[w]$ is a maximizer of a suitable entropy functional and "compatible with the observed macroscopic moments of w" implies that the maximization problem is subject to a set of constraints. To be more concrete, let us first of all introduce the shorthand

$$\langle f \rangle(\mathbf{x}, t) := \int_{\mathbb{R}^d} f(\mathbf{x}, \mathbf{p}, t)\, d\mathbf{p} \tag{1.50}$$

for the integrals with respect to $\mathbf{p}$, and fix a set of N moments

$$M_k := \langle \varphi_k\, w \rangle, \qquad k = 1, \ldots, N,$$

[1] Equation (1.36) has been rewritten here with a slightly different notation. Moreover the mass is set equal to 1

by which we want to characterize our system macroscopically. The functions $\varphi_k(\mathbf{p})$ define the moments and usually (but not necessarily) are polynomials. Moreover, let $s : \mathbb{R}^+ \to \mathbb{R}$ be a regular, convex, function and let ϱ be a density operator. The *von Neumann entropy* [113] associated to the state ϱ is

$$\mathscr{S}(\varrho) = \mathrm{Tr}\{-k_B\, s(\varrho/N_0)\}, \tag{1.51}$$

where Tr denotes the operator trace, k_B is Boltzmann constant and N_0 is a constant introduced for dimensional reason. Then, by exploiting the mapping between phase-space function and operators furnished by the Weyl quantization $\mathrm{Op}_\hbar$ (see definition (1.28)), the Quantum Maximum Entropy Principle can be precisely stated by requiring that

$$g[w] = g[\langle\varphi_1 w\rangle, \ldots, \langle\varphi_N w\rangle]$$

where $g[M_1, \ldots, M_N]$ satisfies the following constrained maximization problem.

Problem. *Fixed N functions $M_k(\mathbf{x}, t)$, find $g[M_1, \ldots, M_N]$ that maximizes the entropy functional $f \mapsto \mathscr{S}(\mathrm{Op}_\hbar(f))$, among all Wigner functions f that share the moments $M_1, \ldots M_N$. In short:*

$$g[M_1, \ldots, M_N] = \max\{\mathscr{S}(\mathrm{Op}_\hbar(f)) \mid \langle\varphi_k f\rangle = M_k,\ k = 1, \ldots, N\}\ . \tag{1.52}$$

In [41] the following necessary condition[2] for the maximizer $g[M_1, \ldots, M_N]$ is formally proven.

Theorem 1.1. *A necessary condition for $g[M_1, \ldots, M_N]$ to be a solution of the above constrained maximization problem is that N Lagrange multipliers $\lambda_k(\mathbf{x}, t)$, $k = 1, \ldots, N$ exist such that*

$$g[M_1, \ldots, M_N] = \mathrm{Op}_\hbar^{-1}\left\{(s')^{-1}\left[\mathrm{Op}_\hbar\left(\sum_{k=1}^{N}\lambda_k\varphi_k\right)\right]\right\}\ . \tag{1.53}$$

Note that $(s')^{-1}$ is the energy distribution function associated to the entropy s. For example,

$$(s')^{-1}(\eta) = \frac{1}{\mathrm{e}^\eta + \delta}$$

for an entropy function of the form

$$s(f) = f\log f + \delta^{-1}(1-\delta f)\log(1-\delta f) \tag{1.54}$$

[2] A rigorous proof of existence and uniqueness of the constrained minimization problem has been recently obtained by Méhats and Pinaud [78, 79] for the moments up to first degree (density and current).

(where $\delta = 1, 0, -1$ correspond, respectively, to Fermi-Dirac, Maxwell-Boltzmann and Bose-Einstein statistics). The "quantum Maxwellian" (1.53) is a complicated object, involving a back-and-forth Weyl quantization as well as solving for $\lambda_1, \ldots, \lambda_N$ in function of $M_1, \ldots, M_N$ from the constraint equations

$$\langle \varphi_k g[M_1, \ldots, M_N] \rangle = M_k, \qquad k = 1, \ldots, N.$$

Then, although the numeric treatment of $g[M_1, \ldots, M_N]$ is devisable [53, 54], usually a semiclassical expansion is performed in order to obtain explicit models.

Once the local equilibrium state has been obtained with the QMEP, it can be used for the closure of the moment equations. Since the procedure depends on the type of the sought fluid equations, let us illustrate it by means of two examples: isothermal Euler equations and drift-diffusion equations. The two examples are based on the standard scalar Hamiltonian; we shall discuss the multi-band theory afterwards.

First of all, let us rewrite Eq. (1.49) in dimensionless variables. Let $\mathbf{x}_0$, t_0 and E_0 be the reference length, time and energy. Reference temperature and momentum are naturally related to E_0 by

$$k_B T_0 = E_0, \qquad \frac{p_0^2}{m} = E_0.$$

Then, in Eq. (1.49) we switch to dimensionless quantities

$$\mathbf{x} \to x_0 \mathbf{x}, \qquad t \to t_0 t, \qquad \mathbf{p} \to p_0 \mathbf{p}, \qquad V \to E_0 V,$$

(for the sake of simplicity the new dimensionless variables are denoted by the same symbols as the old ones), which yields

$$\frac{1}{t_0} \frac{\partial w}{\partial t} + \frac{p_0}{m x_0} p \cdot \nabla_x + \frac{E_0}{x_0 p_0} \Theta_{\frac{\hbar}{x_0 p_0}}[V] = \frac{1}{t_c} \left(g[w] - w \right).$$

We rewrite the last equation by introducing the *semiclassical parameter*

$$\epsilon = \frac{\hbar}{x_0 p_0} \tag{1.55}$$

and the energy time scale

$$t_E = \frac{m x_0}{p_0}$$

(i.e. the order of time for a particle of kinetic energy E_0 to travel a distance x_0), obtaining:

$$\frac{1}{t_0} \frac{\partial w}{\partial t} + \frac{1}{t_E} \mathbf{p} \cdot \nabla_x w + \frac{1}{t_E} \Theta_\epsilon[V] w = \frac{1}{t_c} \left(g[w] - w \right). \tag{1.56}$$

Now, two different asymptotic regimes, leading to different types of fluid equations, can be analyzed.

1.4.2.1 Hydrodynamic Regime

In this regime the system is observed on the time-scale t_E and collisions are assumed to act on a much shorter time-scale; then we put

$$\tau := \frac{t_c}{t_E} \ll 1, \qquad t_0 = t_E. \tag{1.57}$$

The corresponding Wigner-BGK equation takes therefore the *hydrodynamic scaling* form:

$$\tau \frac{\partial w}{\partial t} + \tau\, \mathbf{p} \cdot \nabla_x w + \tau\, \Theta_\epsilon[V]w = g[w] - w. \tag{1.58}$$

We choose as macroscopic variables the density $n = \langle w \rangle$ and the velocity[3] $u_k = \langle p_k w \rangle / n$, $k = 1, \ldots, d$, yielding $d + 1$ constraints in the entropy maximization problem. Moreover, since we are assuming that the temperature is constant, we have to consider the additional constraint

$$\frac{1}{2}\langle p^2 w \rangle = T$$

where T is a dimensionless, constant temperature. Imposing such a constraint is equivalent to changing the entropy functional (1.51) into the (dimensionless) free-energy functional

$$\mathscr{S}(\varrho) = \mathrm{Tr}\{T s(\varrho) + H\varrho\}, \tag{1.59}$$

where[4] $H = \mathrm{Op}_\epsilon(\frac{1}{2}p^2)$ The constrained *minimization* of such functional, analogously to what stated by Theorem 1.1, leads to a quantum Maxwellian of the form

$$g[w] = g[\langle w \rangle, \langle pw \rangle]$$

[3] Recalling definition (1.50), a comparison with (1.48) shows that the physical density and velocity are given by $N_0 n$ and $p_0 u$, where $N_0 = (p_0/2\pi\hbar)^2$.

[4] In dimensionless variables, all the identities involving Weyl quantization are obtained from the original ones by the formal substitution $\hbar \mapsto \epsilon$.

where $g[n, n\mathbf{u}]$ is (formally) determined by

$$\begin{cases} g[n, n\mathbf{u}] = \mathrm{Op}_\epsilon^{-1} \left\{ (s')^{-1} \left[\mathrm{Op}_\epsilon \left(\dfrac{|\mathbf{p} - \mathbf{B}|^2}{2T} + \dfrac{A}{T} \right) \right] \right\}, \\ \langle g[n, n\mathbf{u}] \rangle = n, \\ \langle p_k g[n, n\mathbf{u}] \rangle = n u_k, \qquad k = 1, \ldots, d, \end{cases} \tag{1.60}$$

where A and $\mathbf{B} = (B_1, \ldots, B_d)$ are the Lagrange multipliers.

Let now w_τ be solution of Eq. (1.58) and assume that the limit $w_\tau \to w_0$ for $\tau \to 0$ exists with finite moments $n = \langle w_0 \rangle$ and $n\mathbf{u} = \langle p w_0 \rangle$. Then, from (1.58) we obtain $w_0 = g[w_0]$. Taking the moments of both sides of Eq. (1.58) and letting $\tau \to 0$ we obtain

$$\begin{aligned} &\frac{\partial}{\partial t} \langle g[w_0] \rangle + \frac{\partial}{\partial x_i} \langle p_i g[w_0] \rangle + \langle \Theta_\epsilon[V] g[w_0] \rangle = 0, \\ &\frac{\partial}{\partial t} \langle p_i g[w_0] \rangle + \frac{\partial}{\partial x_j} \langle p_i p_j g[w_0] \rangle + \langle p_i \Theta_\epsilon[V] g[w_0] \rangle = 0. \end{aligned}$$

From the semiclassical expansion of the potential operator,

$$\begin{aligned} \Theta_\epsilon[V] &= \frac{i}{\epsilon} \{V, w(t)\}_\# = \frac{i}{\epsilon} (V \# w - w \# V) \\ &= -\sum_{k=0}^{\infty} (-1)^k \left(\frac{\epsilon}{2} \right)^{2k} \sum_{|\alpha| = 2k+1} \nabla_x^\alpha V \, \nabla_p^\alpha w \end{aligned} \tag{1.61}$$

(see (1.95)), we immediately obtain

$$\langle p_i \Theta_\epsilon[V] g[w_0] \rangle = \langle g[w_0] \rangle \frac{\partial V}{\partial x_i} = n \frac{\partial V}{\partial x_i}, \tag{1.62}$$

and then, using $g[w_0] = g[n, n\mathbf{u}]$, the moment equations read as follows:

$$\begin{aligned} &\frac{\partial n}{\partial t} + \frac{\partial J_i}{\partial x_i} = 0 \\ &\frac{\partial J_i}{\partial t} + \frac{\partial}{\partial x_j} \langle p_i p_j g[n, n\mathbf{u}] \rangle + n \frac{\partial V}{\partial x_i} = 0, \end{aligned} \tag{1.63}$$

which is a formally closed, Euler-like, QFD system. Systems of this kind, and their semiclassical expansions, have been derived in [40, 61], for Maxwell-Boltzmann statistics and in [16, 107] for Fermi-Dirac and Bose-Einstein statistics.

1.4.2.2 Diffusive Regime

In this regime the collisions are still assumed to act on a time-scale much shorter than t_E, but the system is observed on a time-scale much larger than t_E; then we put

$$\tau := \frac{t_c}{t_E} \ll 1, \qquad \frac{t_E}{t_0} = \tau \tag{1.64}$$

(so that $t_0 = t_E^2/t_c$). The corresponding Wigner-BGK equation takes in this case the *diffusive scaling* form:

$$\tau^2 \frac{\partial w}{\partial t} + \tau\, \mathbf{p} \cdot \nabla_x w + \tau\, \Theta_\epsilon[V]w = g[w] - w. \tag{1.65}$$

The only macroscopic moment needed is the density $n = \langle w \rangle$ and the additional constraint of constant temperature is also to be imposed, implying the use of the free-energy functional (1.59). Then, the QMEP in this case leads to a quantum Maxwellian of the simpler form

$$g[w] = g[\langle w \rangle]$$

where $g[n]$ is given by

$$\begin{cases} g[n] = \mathrm{Op}_\epsilon^{-1} \left\{ (s')^{-1} \left[\mathrm{Op}_\epsilon \left(\dfrac{p^2}{2T} + \dfrac{A}{T} \right) \right] \right\}, \\ \langle g[n] \rangle = n, \end{cases} \tag{1.66}$$

where the single Lagrange multiplier A (the "chemical potential") is needed.

Diffusive equations can now be obtained from Eq. (1.65) by using the "Chapman-Enskog" method. Let us assume that the solution w_τ of Eq. (1.65), for $\tau \to 0$, has a limit $w_\tau \to w_0$ with finite density $n = \langle w_0 \rangle$. Letting $\tau \to 0$ in Eq. (1.65) we still obtain $w_0 = g[w_0]$ but, contrarily to the previous case, the equation for the density

$$\tau \frac{\partial}{\partial t} \langle w_\tau \rangle + \frac{\partial}{\partial x_i} \langle p_i w_\tau \rangle = 0 \tag{1.67}$$

only gives, in the limit, the condition

$$\langle \mathbf{p} g[w_0] \rangle = 0, \tag{1.68}$$

i.e. the equilibrium state carries no current. The diffusive equations must be sought at next order of the Chapman-Enskog expansion

$$w_\tau = g[w_\tau] + \tau w_1.$$

Substituting this ansatz into Eq. (1.65), and letting $\tau \to 0$, yields

$$w_1 = -\left(\mathbf{p}\cdot\nabla_x + \Theta_\epsilon[V]\right) g[w_0]$$

and, therefore, from Eqs. (1.67) and (1.62), we obtain the diffusive equation

$$\frac{\partial n}{\partial t} = \frac{\partial}{\partial x_i}\left(J_i + n\,\frac{\partial V}{\partial x_i}\right), \qquad J_i = \frac{\partial}{\partial x_j}\left\langle p_i\, p_j\, g[n]\right\rangle, \tag{1.69}$$

where also $g[w_0] = g[n]$ has been considered, showing that (1.69) is a formally closed system. Quantum diffusive systems of this kind have been derived in [42] for Boltzmann statistics and in [16, 107] for Fermi-Dirac and Bose-Einstein statistics.

1.4.3 Spinorial and Multi-Band QFD

Let us now turn to the spinorial/multi-band case, indicating by this term a general framework where discrete degrees of freedom are involved. The general idea and the methodologies are not different from the scalar case illustrated in the previous subsection. However, in the multi-band case, peculiar issues have to be addressed, namely:

- the fluid description has to include spin-related moments, whose choice is not evident a priori;
- the technical difficulties, e.g. in the semiclassical expansion of the local-equilibrium states, increase sensibly;
- Hamiltonians unbounded from below may make problematic the realization of quantum Maxwellians with given moments (this is the case e.g. of graphene, see Sect. 1.4.3.3).

Although there is no general answer to these issues, they will be illustrated by means of the particular examples that follow the general discussion.

Let us restrict our framework to spin $\frac{1}{2}$-like systems, whose state space is $L^2(\mathbb{R}^d, \mathbb{C}^2)$, i.e. to systems whose wave functions are two-components spinors.[5] In this case, the density operator ϱ is a 2×2 matrix ϱ_{ij}. Thus, we can associate ϱ with a 2×2 matrix of Wigner functions $w_{ij} = w_{ij}(\mathbf{x},\mathbf{p})$ given by

$$w_{ij} = \mathrm{Op}_\hbar^{-1}(\varrho_{ij}) \tag{1.70}$$

[5] Of course, there are also examples of multi-band QFD equations with an arbitrary number of bands, see e.g. [96].

which turns out to be hermitian

$$\overline{w_{ij}(\mathbf{x},\mathbf{p})} = w_{ji}(\mathbf{x},\mathbf{p})$$

(see also Sect. 1.3, where the Wigner functions have been introduced in the scalar case). Recalling that the Pauli matrices together with the identity matrix,

$$\sigma_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

are a basis of the real space of 2×2 hermitian matrices, (orthogonal with respect to the hermitian product $\frac{1}{2}\operatorname{Tr}(AB^*)$), it will be convenient to consider the Pauli components of the Wigner matrix w, which are four real-valued functions on phase-space,

$$w = w_0\sigma_0 + \vec{w}\cdot\vec{\sigma}, \qquad \vec{w} = (w_1, w_2, w_3),$$

denoting by $\vec{w}$ the vector spinorial part of w. In this way we have a very simple and classical-fashioned formula for the expected value of an observable A in the state ϱ, extending the scalar identity (1.33):

$$\frac{1}{2}\operatorname{Tr}(\varrho A) = \frac{1}{2}\operatorname{Tr}(A\varrho) = \sum_{k=0}^{3}\int_{\mathbb{R}^{2d}} a_k(\mathbf{x},\mathbf{p})\, w_k(\mathbf{x},\mathbf{p})\, d\mathbf{x}\, d\mathbf{p}, \tag{1.71}$$

where ϱ and A were also decomposed along the Pauli basis:

$$\varrho = \sum_{k=0}^{3} \mathrm{Op}_\hbar(w_k)\sigma_k, \qquad A = \sum_{k=0}^{3} \mathrm{Op}_\hbar(a_k)\sigma_k.$$

Such identity, allows to define the *local density* of the observable A in the state ϱ as

$$\frac{1}{2}\operatorname{Tr}(\varrho A \mid \mathbf{x}) = \sum_{k=0}^{3}\int_{\mathbb{R}^3} a_k(\mathbf{x},\mathbf{p})\, w_k(\mathbf{x},\mathbf{p})\, d\mathbf{p}. \tag{1.72}$$

1.4.3.1 The Two-Band k·p Model

The content of this section is based on the paper [17]. The "mini-band" case, where the periodicity of energy bands has to be taken into account, has been treated (with partially different techniques) in [27].

The simplest possible description of an electron in a semiconductor crystal with two energy bands (e.g. "valence" and "conduction") is represented by the 2×2 Hamiltonian

$$H = \begin{pmatrix} -\frac{\hbar^2}{2m}\Delta + \gamma & -\hbar\boldsymbol{\alpha}\cdot\nabla \\ \hbar\boldsymbol{\alpha}\cdot\nabla & -\frac{\hbar^2}{2m}\Delta - \gamma \end{pmatrix}, \tag{1.73}$$

where we put

$$\boldsymbol{\alpha} = \frac{\hbar}{m}\mathbf{K} \quad \text{and} \quad \gamma = E_g/2 \tag{1.74}$$

(this corresponds to the two-band version of the Hamiltonian (1.18) introduced in Sect. 1.2.2). The **k**·**p** model has to be completed by adding an "external" potential term qV (where $q > 0$ denotes the elementary charge), accounting for all electric fields except the crystal one. The electric potential $V(\mathbf{x})$ can be either fixed or self-consistently given by a Poisson equation.

The "free" **k**·**p** Hamiltonian H is the quantization of the classical matrix-valued symbol

$$h(\mathbf{p}) = \begin{pmatrix} \frac{p^2}{2m} + \gamma & -i\boldsymbol{\alpha}\cdot\mathbf{p} \\ i\boldsymbol{\alpha}\cdot\mathbf{p} & \frac{p^2}{2m} - \gamma \end{pmatrix} \tag{1.75}$$

or, in Pauli components,

$$h(\mathbf{p}) = \frac{p^2}{2m}\sigma_0 + \boldsymbol{\alpha}\cdot\mathbf{p}\,\sigma_2 + \gamma\,\sigma_3 = h_0(\mathbf{p})\sigma_0 + \vec{h}(\mathbf{p})\cdot\vec{\sigma}, \tag{1.76}$$

where

$$h_0(\mathbf{p}) = \frac{p^2}{2m}, \qquad \vec{h}(\mathbf{p}) = (0, \boldsymbol{\alpha}\cdot\mathbf{p}, \gamma). \tag{1.77}$$

The dispersion relation for the free Hamiltonian H is easily obtained by computing the eigencouples of the symbol $h(\mathbf{p})$, which yields the *energy bands*

$$E_\pm(\mathbf{p}) = \frac{p^2}{2m} \pm \sqrt{(\boldsymbol{\alpha}\cdot\mathbf{p})^2 + \gamma^2} = \frac{p^2}{2m} \pm |\vec{h}(\mathbf{p})| \tag{1.78}$$

and the corresponding normalized energy eigenvectors

$$\psi_\pm^p = \frac{1}{\sqrt{2(1 \pm \nu_3(\mathbf{p}))}} \begin{pmatrix} \nu_3(\mathbf{p}) \pm 1 \\ \nu_1(\mathbf{p}) + i\,\nu_2(\mathbf{p}) \end{pmatrix}, \tag{1.79}$$

where we have introduced

$$\vec{\nu}(\mathbf{p}) = (\nu_1(\mathbf{p}), \nu_2(\mathbf{p}), \nu_3(\mathbf{p})) = \frac{\vec{h}(\mathbf{p})}{|\vec{h}(\mathbf{p})|} = \frac{(0, \boldsymbol{\alpha}\cdot\mathbf{p}, \gamma)}{\sqrt{(\boldsymbol{\alpha}\cdot\mathbf{p})^2 + \gamma^2}}. \tag{1.80}$$

The two eigenprojections $P_\pm(\mathbf{p})$, that we call *band-projections*, are therefore given by

$$P_\pm(\mathbf{p}) = \psi_\pm^p \otimes \psi_\pm^p = \frac{1}{2}(\sigma_0 \pm \vec{\nu}(\mathbf{p}) \cdot \vec{\sigma}) \tag{1.81}$$

and we can clearly write

$$h(\mathbf{p}) = E_+(\mathbf{p})P_+(\mathbf{p}) + E_-(\mathbf{p})P_-(\mathbf{p}). \tag{1.82}$$

The *local band densities* are the local densities of the observables $P_\pm$ and, according to (1.72), are given by

$$n_\pm(\mathbf{x}) = \mathrm{Tr}(\varrho P_\pm \mid \mathbf{x}) = \int_{\mathbb{R}^3} \left[w_0(\mathbf{x}, \mathbf{p}) \pm \vec{\nu}(\mathbf{p}) \cdot \vec{w}(\mathbf{x}, \mathbf{p}) \right] d\mathbf{p}. \tag{1.83}$$

We want now to derive diffusive QFD equations for $n_\pm$. Following the same procedure outlined in Sect. 1.4.2 for the scalar case, the starting point is the Wigner-BGK equation in the diffusive scaling which, in this case, reads as follows:

$$\begin{cases} \left(\tau \dfrac{\partial}{\partial t} + \mathbf{p} \cdot \nabla_x + \Theta_\epsilon[V] \right) w_0 + \boldsymbol{\alpha} \cdot \nabla_x w_2 = \dfrac{g_0[w] - w_0}{\tau}, \\ \left(\tau \dfrac{\partial}{\partial t} + \mathbf{p} \cdot \nabla_x + \Theta_\epsilon[V] \right) \vec{w} + \boldsymbol{\alpha} \cdot \nabla_x w_0 \, \vec{e}_2 - \dfrac{2}{\epsilon} \vec{h}(\mathbf{p}) \times \vec{w} = \dfrac{\vec{g}[w] - \vec{w}}{\tau}, \end{cases} \tag{1.84}$$

where $\vec{e}_2 = (0, 1, 0)$, $\vec{h}(\mathbf{p}) = (0, \boldsymbol{\alpha} \cdot \mathbf{p}, \gamma)$, and $\boldsymbol{\alpha}$ and γ denote now the new, dimensionless, band parameters

$$\boldsymbol{\alpha} := \frac{\hbar}{p_0} \mathbf{K}, \qquad \gamma := \frac{mE_g}{2p_0^2}. \tag{1.85}$$

Here, $g[w] = g_0[w]\sigma_0 + \vec{g}[w] \cdot \vec{\sigma}$ is the local equilibrium Wigner matrix, which is assumed to be given by the QMEP with given moments n_+ and n_- and Maxwell-Boltzmann statistics. Therefore it has the form $g[w] = g[n_+, n_-]$, subject to the constraints

$$\langle (g[n+, n_-])_\pm \rangle = n_\pm \tag{1.86}$$

(here, $w_\pm := w_0 \pm \vec{\nu} \cdot \vec{w}$ and $\langle \cdot \rangle$ is the usual shorthand for $\int \cdot d\mathbf{p}$). Similarly to the scalar case (see Theorem 1.1) it can be proved [17] that two functions $\mu_0(\mathbf{x})$ and $\mu_s(\mathbf{x})$ exist such that

$$g[n_+, n_-] = \mathrm{Op}_\epsilon^{-1} \left(\mathrm{e}^{-H_\mu} \right), \tag{1.87}$$

where $H_\mu = \mathrm{Op}_\epsilon(h_\mu)$ is a modified **k·p** Hamiltonian with symbol

$$h_\mu(\mathbf{x},\mathbf{p}) = \left[p^2/2 + V(\mathbf{x}) + \mu_0(\mathbf{x})\right]\sigma_0 + \left[|\vec{h}(\mathbf{p})| + \mu_s(\mathbf{x})\right]\vec{\nu}(\mathbf{p})\cdot\vec{\sigma}. \tag{1.88}$$

Putting $\mu_\pm(\mathbf{x}) = V(\mathbf{x}) + \mu_0(\mathbf{x}) \pm \mu_s(\mathbf{x})$ we can write

$$h_\mu(\mathbf{x},\mathbf{p}) = [E_+(\mathbf{p}) + \mu_+(\mathbf{x})]\,P_+(\mathbf{p}) + [E_-(\mathbf{p}) + \mu_-(\mathbf{x})]\,P_-(\mathbf{p}), \tag{1.89}$$

where $E_\pm(\mathbf{p})$ are the (scaled) energy bands. From (1.89) we realize that H_μ provides each energy band with its own chemical potential. The two degrees of freedom represented by μ_+ and μ_- allows the two constraints (1.86) to be satisfied.

Let us now rewrite (1.84) in the compact form

$$\tau\,\frac{\partial w}{\partial t} + Tw = \frac{g[w] - w}{\tau}. \tag{1.90}$$

The Chapman-Enskog procedure (see Sect. 1.4.2), applied to the moments

$$n_\pm = \langle w_\pm \rangle,$$

leads to the two-band quantum diffusive equations

$$\frac{\partial n_\pm}{\partial t} = \langle (T\,Tg[n_+, n_-])_\pm \rangle \tag{1.91}$$

(which are nothing else than a compact form of Eq. (1.69) for the **k·p** system). The fully-quantum diffusive model will be therefore given by Eq. (1.91) coupled with Eqs. (1.87) and (1.86), that define $g[n_+, n_-]$. Of course, such a model is extremely involved and, even though numerical approaches are envisageable, it will be interesting to approximate it semiclassically.

In order to perform such a semiclassical approximation, we need to expand $g \equiv g[n_+, n_-]$ in a formal power series of ϵ:

$$g = g^{(0)} + \epsilon g^{(1)} + \epsilon^2 g^{(2)} + \cdots. \tag{1.92}$$

The terms of the expansion can be computed as follows. First, we notice [17] that

$$g(\beta) := \mathrm{Op}_\epsilon^{-1}\left(\mathrm{e}^{-\beta H_\mu}\right), \qquad \beta \geq 0,$$

satisfies the semigroup equation

$$\begin{cases} \partial_\beta\, g(\beta) = -h_\mu \#_\epsilon\, g(\beta), & \beta > 0, \\ g(0) = \sigma_0. \end{cases} \tag{1.93}$$

The Moyal product # has a formal semiclassical expansion

$$\#_\epsilon = \#^{(0)} + \epsilon\,\#^{(1)} + \epsilon^2\#^{(2)} + \cdots, \tag{1.94}$$

where

$$a\,\#^{(k)}b = \frac{1}{(2i)^k}\sum_{|\alpha|+|\beta|=k}\frac{(-1)^{|\alpha|}}{\alpha!\,\beta!}\left(\partial_x^\alpha\partial_p^\beta a\right)\left(\partial_p^\alpha\partial_x^\beta b\right). \tag{1.95}$$

Since $H_\mu \mathrm{e}^{-\beta H_\mu} = \mathrm{e}^{-\beta H_\mu} H_\mu$ and, therefore, $h_\mu \#_\epsilon\, g(\beta) = g(\beta)\#_\epsilon h_\mu$, then, using the fact that $\#^{(k)}$ is symmetric for even k and antisymmetric for odd k, the semigroup Eq. (1.93) can be expanded as

$$-\,\partial_\beta\, g(\beta) = \sum_{k\ \text{even}}\epsilon^k\,[h_\mu\#^{(k)}g(\beta)]_+ + \sum_{k\ \text{odd}}\epsilon^k\,[h_\mu\#^{(k)}g(\beta)]_-, \tag{1.96}$$

where $[ab]_+$ and $[ab]_-$ denote, respectively, the symmetric and antisymmetric part of the matrix product ab. Substituting (1.92) and (1.94) in (1.93) yields, at leading order,

$$\begin{cases}\partial_\beta\, g^{(0)}(\beta) = -h_\mu\, g^{(0)}(\beta), \quad \beta > 0,\\ g^{(0)}(0) = \sigma_0,\end{cases}$$

and, therefore,

$$g^{(0)}(\mathbf{x},\mathbf{p};\beta) = \mathrm{e}^{-\beta h_\mu(\mathbf{x},\mathbf{p})}.$$

After straightforward calculations we obtain the Pauli components of $g^{(0)} = g^{(0)}(\beta = 1)$:

$$g_0^{(0)} = \frac{\phi_+^\mu + \phi_-^\mu}{2}, \qquad \vec{g}^{\,(0)} = \frac{\phi_+^\mu - \phi_-^\mu}{2}\vec{\nu} \tag{1.97}$$

where

$$\phi_\pm^\mu = \phi_\pm^\mu(\mathbf{x},\mathbf{p}) = \mathrm{e}^{-\left[E_\pm(\mathbf{p})+\mu_\pm(\mathbf{x})\right]}. \tag{1.98}$$

As far as the first-order term is concerned, from Eqs. (1.96) and (1.92) we obtain

$$g^{(1)}(\beta) = -\int_0^\beta \mathrm{e}^{-(\beta-\beta')h_\mu}\,[h_\mu\#^{(1)}g^{(0)}(\beta')]_-\,d\beta'. \tag{1.99}$$

After lengthy but straightforward calculations we obtain the Pauli components of $g^{(1)} = g^{(1)}(\beta = 1)$:

$$g_0^{(1)} = 0, \qquad \vec{g}^{(1)} = \eta^\mu \vec{e}_1, \tag{1.100}$$

with $\vec{e}_1 = (1, 0, 0)$ and

$$\eta^\mu = -\frac{\gamma \alpha_i}{4|\vec{h}|^2} \left[\phi_+^\mu \partial_i \mu_+ + \phi_-^\mu \partial_i \mu_- + \frac{(\phi_+^\mu - \phi_-^\mu)\partial_i(\mu_+ + \mu_-)}{2|\vec{h}| + \mu_+ - \mu_-} \right], \tag{1.101}$$

where $\phi_\pm^\mu$ are given by (1.98) and $\vec{h}(\mathbf{p}) = (0, \boldsymbol{\alpha} \cdot \mathbf{p}, \gamma)$.

Now, $\mu_\pm$ depend on ϵ because the constraint equations (1.86) contain ϵ. But, recalling that $w_\pm = w_0 \pm \vec{v} \cdot \vec{w}$, we have $g_\pm^{(0)} = \phi_\pm^\mu$ and $g_\pm^{(1)} = 0$. Thus, the constraint equations imply

$$n_\pm = \langle g_\pm \rangle = \left\langle (g^{(0)} + \epsilon g^{(1)})_\pm \right\rangle + \mathscr{O}(\epsilon^2) = \mathrm{e}^{-\left[E_\pm(\mathbf{p}) + \mu_\pm(\mathbf{x})\right]} + \mathscr{O}(\epsilon^2),$$

which allows to invert at order 2 the relation between $\mu_\pm$ and $n_\pm$:

$$\mu_\pm = -\log \frac{n_\pm}{z_\pm} + \mathscr{O}(\epsilon^2), \tag{1.102}$$

with $z_\pm = \int_{\mathbb{R}^3} \mathrm{e}^{-E_\pm(\mathbf{p})} d\mathbf{p}$. We see, therefore, that in (1.98) and (1.101) we can substitute $\mu_\pm$ with $-\log \frac{n_\pm}{z_\pm}$ by making an error of order ϵ^2 and we obtain the result below [17]:

Proposition 1.1. *The QMEP local equilibrium state $q \equiv g[n_+, n_-]$ has the semiclassical expansion* (1.92) *where*

$$g_0^{(0)} = \frac{\phi_+ + \phi_-}{2}, \qquad \vec{g}^{(0)} = \frac{\phi_+ - \phi_-}{2} \vec{v}, \tag{1.103}$$

with

$$\phi_\pm(\mathbf{x}, \mathbf{p}) = n_\pm(\mathbf{x})\, \mathrm{e}^{-E_\pm(\mathbf{p})}/z_\pm, \qquad z_\pm := \int_{\mathbb{R}^3} \mathrm{e}^{-E_\pm(\mathbf{p})} d\mathbf{p}, \tag{1.104}$$

and

$$g_0^{(1)} = 0, \qquad \vec{g}^{(1)} = \frac{\gamma \alpha_i}{4|\vec{h}|^2} \left[\partial_i(\phi_+ + \phi_-) + \frac{(\phi_+ - \phi_-)\, \partial_i \log(n_+ n_-)}{2|\vec{h}| - \log \frac{n_+ z_-}{z_+ n_-}} \right] \vec{e}_1. \tag{1.105}$$

Note that $g^{(1)}$ is certainly well defined if $\frac{n_+}{z_+} \leq \frac{n_-}{z_-}$. We can now use this result in order to approximate at leading order in ϵ the quantum diffusive equations (1.91). The limit behavior depends on how the crystal parameters $\boldsymbol{\alpha}$ and γ scale with respect to ϵ. The following result is (formally) proved in [17].

Theorem 1.2. *Assuming $\boldsymbol{\alpha}$ and γ of order 1, the solution (n_+, n_-) of the quantum diffusive equations (1.91) satisfies, up to $\mathcal{O}(\epsilon)$-terms, the following system:*

$$\begin{aligned}\partial_t n_\pm &= \partial_i \left[D_{ij}^{\pm}\, \partial_j n_\pm + \partial_j V \left(\delta_{ij} \pm 2C_{ij}^{3\pm}\right) n_\pm \right] \\ &\quad - \partial_i \left[\partial_j V \left(C_{ij}^{3+} n_+ - C_{ij}^{3-} n_- \right) \right] \mp \partial_i V \partial_j V \left(C_{ij}^{4+} n_+ - C_{ij}^{4-} n_- \right) \\ &\quad - \int_{\mathbb{R}^3} \left(c_{ij}^2 \partial_i \pm c_{ij}^3 \partial_i V \right) \frac{(\phi_+ - \phi_-)\, \partial_j \log(n_+ n_-)}{2|\vec{h}| - \log \frac{n_+ z_-}{z_+ n_-}}\, d\mathbf{p},\end{aligned} \tag{1.106}$$

with

$$\begin{aligned} D_{ij}^{\pm} &:= \frac{1}{z_\pm} \int_{\mathbb{R}^3} (\mathbf{p} \pm \nu_2 \boldsymbol{\alpha})_i (\mathbf{p} \pm \nu_2 \boldsymbol{\alpha})_j \, \mathrm{e}^{-E_\pm(\mathbf{p})} d\mathbf{p}, \\ C_{ij}^{k\pm} &:= \frac{1}{z_\pm} \int_{\mathbb{R}^3} \frac{\gamma^2 \alpha_i \alpha_j}{2|\vec{h}(\mathbf{p})|_k} \, \mathrm{e}^{-E_\pm(\mathbf{p})} d\mathbf{p}. \end{aligned} \tag{1.107}$$

Assuming instead that $\boldsymbol{\alpha}$ and γ are of order ϵ (and thus rewriting them as $\epsilon\boldsymbol{\alpha}$ and $\epsilon\gamma$, with $\boldsymbol{\alpha}$ and γ of order 1), (n_+, n_-) satisfies, up to $\mathcal{O}(\epsilon)$-terms, the following system:

$$\frac{\partial n_\pm}{\partial t} = \frac{\partial}{\partial t} \left(\frac{\partial n_\pm}{\partial t} + \frac{\partial V}{\partial x_i} n_\pm \right) \mp C_{ij} \frac{\partial V}{\partial x_i} \frac{\partial V}{\partial x_j} (n_+ - n_-), \tag{1.108}$$

where

$$C_{ij} := \frac{\alpha_i \alpha_j \gamma^2}{2(2\pi)^{3/2}} \int_{\mathbb{R}^2} \frac{\mathrm{e}^{-p^2/2}}{\left((\boldsymbol{\alpha} \cdot \mathbf{p})^2 + \gamma^2 \right)^2} \, d\mathbf{p}. \tag{1.109}$$

The intermediate cases $\boldsymbol{\alpha} \sim 1, \gamma \sim \epsilon$ and $\boldsymbol{\alpha} \sim \epsilon, \gamma \sim 1$ are less interesting and are discussed in [17]. In Eq. (1.107) notice that

$$\mathbf{p} \pm \nu_2 \boldsymbol{\alpha} = \nabla_{\mathbf{p}} E_\pm$$

are the electron semiclassical velocities in the two bands and that the response of the electrons to the electric field is mediated by an average of the tensor

$$\delta_{ij} \pm 2c_{ij}^3(\mathbf{p}) = \delta_{ij} \pm \frac{\gamma^2 \alpha_i \alpha_j}{|\vec{h}(\mathbf{p})|^3} = \frac{\partial^2 E_\pm(\mathbf{p})}{\partial p_i \partial p_j},$$

which is clearly related to the $\mathbf{k}\cdot\mathbf{p}$ effective-mass tensor [15,115]. For certain values of the band parameters $\boldsymbol{\alpha}$ and γ such effective-mass may be negative, making electrons in the lower band behave like holes. All such features disappear in the second model (1.108) since the band structure has little influence on the band dynamics and only influences the field-driven band coupling (Zener tunneling).

1.4.3.2 The Rashba Spin-Orbit System

The content of this section is based on the paper [19].

A two-Dimensional Electron Gas (2DEG) confined into an asymmetric two-dimensional quantum well, experiences the Bychkov-Rashba (B-R) spin-orbit interaction [23]. Such 2DEG is described by a two-band Hamiltonian whose symbol is

$$h(p) = \frac{p^2}{2m}\sigma_0 + \alpha_R \, \mathbf{p} \times \vec{e}_3 \cdot \vec{\sigma}, \tag{1.110}$$

where $\vec{e}_3 = (0,0,1)$ is the direction perpendicular to the well, $\mathbf{p} = (p_1, p_2, 0)$ and α_R is a constant (the Rashba constant) that depends on the characteristics of the well. Note that the electron momentum $\mathbf{p}$ is two-dimensional but the electron spin is three-dimensional. Since the constant α_R may be tuned by a suitable gate voltage, the B-R effect is expected to lead to semiconductor-based spintronic devices [38,122], with no need of built-in magnetic fields. As in the previous case, the model has to be completed by a potential term qV representing either external (e.g. the gate voltage) or self-consistent electric fields.

At variance with the previous case, here we are interested in the *spin-up* and *spin-down densities*, which are the local densities associated to the observables

$$S_\pm = \frac{1}{2}(\sigma_0 \pm \sigma_3),$$

i.e. the projections on the "spin-up" and "spin-down" eigenstates [6] with respect to the direction $\vec{e}_3$. According to the definition (1.72), the local spin densities are therefore given by

$$n_\pm(\mathbf{x}) = \mathrm{Tr}(\varrho S_\pm \mid \mathbf{x}) = \int_{\mathbb{R}^3} [w_0(\mathbf{x},\mathbf{p}) \pm w_3(\mathbf{x},\mathbf{p})]\, d\mathbf{p}. \tag{1.111}$$

[6] For notational convenience we adopt "+" and "−" to denote "up" and "down" instead of the more common "↑" and "↓".

Note that

$$n_0 = n_+ + n_- \quad \text{and} \quad n_s = n_+ - n_- \tag{1.112}$$

have, respectively, the meaning of (local) *total density* and *spin polarization*. As in the previous case, we want now to derive equations for n_+ and n_- of diffusive type. The Wigner-BGK equations, in diffusive scaling, for the Hamiltonian (1.110) (added with external field) read as follows:

$$\begin{cases} \left(\tau\dfrac{\partial}{\partial t} + \mathbf{p}\cdot\nabla_x + \Theta_\epsilon[V]\right) w_0 + \epsilon\alpha(\nabla\times\vec{e}_3)\cdot\vec{w} = \dfrac{g_0[w]-w_0}{\tau}, \\ \left(\tau\dfrac{\partial}{\partial t} + \mathbf{p}\cdot\nabla_x + \Theta_\epsilon[V]\right)\vec{w} + \epsilon\alpha(\nabla\times\vec{e}_3)w_0 - 2\alpha(\mathbf{p}\times\vec{e}_3)\times\vec{w} = \dfrac{\vec{g}[w]-\vec{w}}{\tau}, \end{cases} \tag{1.113}$$

where

$$\alpha := \frac{mx_0\alpha_R}{\hbar}$$

is the scaled Rashba constant. Note that the spin-orbit coupling is assumed to be of order ϵ, in accordance with experimental values (see [19] and references therein).

In Eq. (1.113), $g[w] = g_0[w]\sigma_0 + \vec{g}[w]\cdot\vec{\sigma}$ is the QMEP local equilibrium Wigner matrix with given moments n_+ and n_- (and Maxwell-Boltzmann statistics) and, therefore, it has the form $g[w] = g[n_+, n_-]$ and is subject to the constraints

$$\langle (g[n+, n_-])_\pm \rangle = n_\pm. \tag{1.114}$$

Such constraints have the same form as Eq. (1.86) but here $w_\pm := w_0 \pm w_3$. Analogously to the $\mathbf{k}\cdot\mathbf{p}$ case, we can prove the following: two functions $\mu_0(\mathbf{x})$ and $\mu_3(\mathbf{x})$ exist such that

$$g[n_+, n_-] = \mathrm{Op}_\epsilon^{-1}\left(\mathrm{e}^{-H_\mu}\right), \tag{1.115}$$

where H_μ is the (quantization of the) modified Bychkov-Rashba Hamiltonian

$$h_\mu(\mathbf{x},\mathbf{p}) = \left[p^2/2 + \mu_0(\mathbf{x})\right]\sigma_0 + \epsilon\alpha\,\mathbf{p}\times\vec{e}_3\cdot\vec{\sigma} + \mu_3(\mathbf{x})\sigma_3,$$

where, without loss of generality, the potential V has been incorporated in the definition of μ_0. Note that $\mu_+ := \mu_0 + \mu_3$ and $\mu_+ := \mu_0 - \mu_3$ can be interpreted as the chemical potentials of the spin-up and spin-down electron populations. Again, by writing system (1.113) in the compact form

$$\tau\frac{\partial w}{\partial t} + Sw = \frac{g[w]-w}{\tau},$$

the Chapman-Enskog procedure leads to the quantum diffusive equations in the compact form

$$\frac{\partial n_\pm}{\partial t} = \langle (SSg[n_+, n_-])_\pm \rangle \tag{1.116}$$

(a more explicit expression of the right hand side of Eq. (1.116) can be found in [19]). System (1.116)+(1.115)+(1.114) is the fully-quantum diffusive model for electrons subject to Rashba effect.

Once again, one may look for a semiclassical expansion of the model. Since the spin-orbit coupling is small (of order ϵ), at variance with the previous case, we can now compute the semiclassical expansion up to $\mathscr{O}(\epsilon^2)$-terms. In order to do that, as in the previous section, we consider the semigroup equation for $g(\beta)$, where now the Hamiltonian h_μ has a part of order ϵ which is conveniently put in evidence:

$$\begin{cases} \partial_\beta\, g(\beta) = -(h_\mu^{(0)} + \epsilon h_\mu^{(1)}) \#_\epsilon\, g(\beta), \quad \beta > 0, \\ g(0) = \sigma_0, \end{cases} \tag{1.117}$$

with

$$h_\mu^{(0)} = \left[p^2/2 + \mu_0(\mathbf{x})\right]\sigma_0 + \mu_3(\mathbf{x})\sigma_3, \qquad h_\mu^{(1)} = \alpha\, \mathbf{p} \times \vec{e}_3 \cdot \vec{\sigma}.$$

Expanding g and # in formal powers of ϵ, we get the hierarchy

$$\begin{cases} \partial_\beta\, g^{(n)}(\beta) = -h_\mu^{(0)}\, g^{(n)}(\beta) - \displaystyle\sum_{j=1}^{n} \left[h_\mu^{(0)} \#_j\, g^{(n-j)}(\beta) + h_\mu^{(1)} \#_{j-1}\, g^{(n-j)}(\beta) \right], \\ g^{(0)}(0) = \sigma_0, \qquad g^{(n)}(0) = 0, \quad n \geq 1. \end{cases} \tag{1.118}$$

It shall be convenient to rewrite these equations for the Laplace transform

$$\hat{g}(z) = \mathscr{L}\,[w]\,(z)$$

of $g(\beta)$, which yields

$$\begin{cases} \hat{g}^{(0)}(z) = R(z), \\ \hat{g}^{(n)}(z) = -\displaystyle\sum_{j=1}^{n} R(z) \left[h_\mu^{(0)} \#_j\, \hat{g}^{(n-j)}(z) + h_\mu^{(1)} \#_{j-1}\, \hat{g}^{(n-j)}(z) \right], \quad n \geq 1, \end{cases} \tag{1.119}$$

where

$$R(z) = (z + h_\mu^{(0)})^{-1} = \begin{pmatrix} \left(z + \frac{1}{2}p^2 + \mu_+\right)^{-1} & 0 \\ 0 & \left(z + \frac{1}{2}p^2 + \mu_-\right)^{-1} \end{pmatrix}. \tag{1.120}$$

Recalling the inversion formula

$$\mathscr{L}^{-1}\left[(z+f)^{-n}\right](\beta) = \frac{\mathrm{e}^{-\beta f}\,\beta^{n-1}}{(n-1)!}, \tag{1.121}$$

the zeroth order approximation to the local equilibrium ($\beta = 1$) is

$$g^{(0)} = \mathscr{L}^{-1}[R](\beta) = \begin{pmatrix} \mathrm{e}^{-\left(\frac{1}{2}p^2+\mu_+\right)} & 0 \\ 0 & \mathrm{e}^{-\left(\frac{1}{2}p^2+\mu_-\right)} \end{pmatrix}. \tag{1.122}$$

Note that $g^{(0)}$ has the form of two independent classical Maxwellians, one for each spin population. Imposing the constraints (1.114) to $g^{(0)}$ we see that, at leading order,

$$\mu_\pm^{(0)}(n_+, n_-) = -\log\frac{n_\pm}{2\pi}, \tag{1.123}$$

so that

$$g^{(0)}[n_+, n_-] = \frac{1}{2\pi}\begin{pmatrix} n_+\mathrm{e}^{-\frac{1}{2}p^2} & 0 \\ 0 & n_-\mathrm{e}^{-\frac{1}{2}p^2} \end{pmatrix}. \tag{1.124}$$

From (1.119) we see that the computation of $\hat{g}^{(n)}$ involves multiple applications of $R(z)$ to the lower-order terms $\hat{g}^{(k)}$, with $k < n$, and their derivatives. Moreover, the inverse Laplace transform can be computed explicitly. It turns out [19] that each component of $g^{(n)}$ is a finite linear combination (depending on $\mu_\pm$ and their derivatives) of terms of the form

$$\mathrm{e}^{-p^2/2} R_{rs}(\mu_+, \mu_-),$$

where

$$R_{rs}(x, y) = \sum_{k=0}^{r-1} \frac{a_k(r,s)\,\mathrm{e}^{-\mu_1}}{(r-k-1)!\,(y-x)^{s+k}} + \sum_{k=0}^{s-1} \frac{b_k(r,s)\,\mathrm{e}^{-y}}{(s-k-1)!\,(x-y)^{r+k}}, \tag{1.125}$$

and the coefficients $a_j = a_j(r,s)$, $b_j = b_j(r,s)$ are recursively given by $a_r = 1$, $b_s = 1$ and

$$
\begin{aligned}
a_k &= -\sum_{j=\max\{0,k-s\}}^{k-1} \binom{s}{k-j} a_j\,, \qquad k = 1,\dots,r-1,\\
b_k &= -\sum_{j=\max\{0,k-r\}}^{k-1} \binom{r}{k-j} b_j\,, \qquad k = 1,\dots,s-1.
\end{aligned}
\tag{1.126}
$$

In [19], the expansion terms $g^{(1)}$, $g^{(2)}$ and (partially) $g^{(3)}$ are computed explicitly, which allows to prove the following.

Theorem 1.3. *The solution (n_+, n_-) of the quantum diffusive equations (1.116) satisfies, up to $\mathcal{O}(\epsilon^3)$-terms, the following system:*

$$
\begin{aligned}
\frac{\partial n_\pm}{\partial t} - \operatorname{div}(\nabla n_\pm + n_\pm \nabla V) + \epsilon^2 \operatorname{div}\left[n_\pm \nabla\, Q(n_\pm)\right] &= \\
\pm 4\alpha^2(n_- - n_+) + \epsilon^2\alpha^2 C_\pm(n_+, n_-),&
\end{aligned}
\tag{1.127}
$$

where

$$
Q(n) = \frac{1}{6}\,\frac{\Delta\sqrt{n}}{\sqrt{n}}
$$

is the Bohm potential and the coupling terms, $C_+(n_+, n_-)$ and $C_-(n_+, n_-) = C_+(n_-, n_+)$, are given by

$$
\begin{aligned}
C_+(n_+, n_-) = {} & -2\,\operatorname{div}\left(n_+ \nabla \frac{S_{21}}{n_+}\right) + 4\left[Q(n_+) - Q(n_-)\right] S_{11} \\
& + \frac{1}{2}\log\frac{n_-}{n_+}\,\operatorname{div}\left(S_{12}\nabla\log n_- - S_{21}\nabla\log n_+ - S_{22}\nabla\log\frac{n_-}{n_+}\right) \\
& + \nabla(V + \log n_-)\cdot\left(S_{12}\nabla\log n_- - S_{21}\nabla\log n_+ - S_{22}\nabla\log\frac{n_-}{n_+}\right) \\
& + \log\frac{n_-}{n_+}\Big[(S_{22} - 2S_{23} + S_{31} - 2S_{41})\,\Delta\log n_+ \\
& \qquad + (S_{22} - 2S_{32} + S_{13} - 2S_{14})\,\Delta\log n_- \\
& \qquad + (2S_{32} + S_{41} - 4S_{33})\,|\nabla\log n_+|^2 \\
& \qquad + (2S_{23} + S_{14} - 4S_{33})\,|\nabla\log n_-|^2 \\
& \qquad + (8S_{33} - S_{32} - S_{23})\,\nabla\log n_+ \cdot \nabla\log n_-\Big] \\
& + 8\alpha^2\left(\frac{S_{21}}{n_+} - \frac{S_{12}}{n_-}\right) S_{11} + 16\alpha^2\, S_{22}\log\frac{n_-}{n_+}
\end{aligned}
$$

where $S_{rs}(n_1, n_2) = R_{rs}(-\log n_1, -\log n_2)$, *the functions* R_{rs} *being given in Eq.* (1.125).

The semiclassical drift-diffusion equations (1.127) contain complicated nonlinear coupling terms, but simpler models can be easily obtained from them.

By neglecting the $\mathscr{O}(\epsilon^2)$-terms we obtain

$$\frac{\partial n_\pm}{\partial t} - \operatorname{div}(\nabla n_\pm + n_\pm \nabla V) = \pm 4\alpha^2 (n_- - n_+), \tag{1.128}$$

which is a system of two classical drift-diffusion equations with a coupling term corresponding to a well known spin-relaxation mechanism of D'yakonov-Perel' type (going back to physical variables, the coupling factor $4\alpha^2$ becomes $4t_c m k_B T_0 \alpha_R^2/\hbar^2$, compare with Eq. (89) of [122]). The spin relaxation has a clear interpretation: collisions change randomly the momentum $\mathbf{p}$ of electrons and, consequently, the spin precession vector $\mathbf{p} \times \vec{e}_3$. Such more and more chaotic orientations of single spins make the average spin polarization disappear in a typical time of order $1/4\alpha^2$ (in scaled variables). Note that Eq. (1.128) is formally correct to first order in ϵ since Eq. (1.127) contains no $\mathscr{O}(\epsilon)$-terms.

A nice intermediate model can be obtained from (1.127) by neglecting only the $\mathscr{O}(\epsilon^2\alpha^2)$-terms. This can be justified by the fact that the values of the scaled Rashba constant for experimental spintronic devices is of order 10^{-1} or less (see [19] and references therein). We obtain in this way the model

$$\frac{\partial n_\pm}{\partial t} - \operatorname{div}(\nabla n_\pm + n_\pm \nabla V) + \frac{\epsilon^2}{6} \operatorname{div}\left[n_\pm \nabla \frac{\Delta \sqrt{n_\pm}}{\sqrt{n_\pm}} \right] = \pm 4\alpha^2 (n_- - n_+), \tag{1.129}$$

featuring the linear D'yakonov-Perel' coupling term as well as the quantum Bohm potential term.

1.4.3.3 Graphene

It is well known that graphene is a two-dimensional crystal with extraordinary electronic properties [50]. Such properties derive from the conical shape of the energy bands around the so-called "Dirac points" in the electron pseudomomentum space. Close to one of such Dirac points, electrons are described (with good approximation for standard energies) by the Hamiltonian

$$H = -ci\hbar \nabla \cdot \vec{\sigma} + V\sigma_0, \tag{1.130}$$

corresponding to the symbol

$$h(\mathbf{x}, \mathbf{p}) = c\mathbf{p} \cdot \vec{\sigma} + V(\mathbf{x})\sigma_0. \tag{1.131}$$

Here, $\mathbf{x} = (x_1, x_2)$ and $\mathbf{p} = (p_1, p_2)$ are the electron position and pseudomomentum, $c \sim 10^6$m/s is the Fermi velocity, and $V(\mathbf{x})$ is an external/self-consistent electric potential. Note that (1.131) is a two-dimensional Dirac-like Hamiltonian [104], meaning that electrons in graphene behave like two-dimensional photons (or neutrinos) with an "effective light speed" which is about $1/300$ of the real light speed and (at variance with photons and neutrinos) subject to electric forces. Graphene, besides its great interest for applications to nanoelectronics, can also be considered as an ideal laboratory where quantum-relativistic effects can be studied at nonrelativistic energies [62].

The energy bands associated with Hamilonian (1.131) are

$$E_{\pm}(\mathbf{p}) = \pm c\,|\mathbf{p}|, \tag{1.132}$$

(showing the above mentioned conical shape), from which we see that the graphene is not bounded from below. The semiclassical velocities associated to the energy bands (1.132) are

$$v_{\pm}(\mathbf{p}) = \nabla E_{\pm}(\mathbf{p}) = \pm c\frac{\mathbf{p}}{|\mathbf{p}|},$$

showing that, from the semiclassical viewpoint, electrons move at constant speed c. Finally, the Hamiltonian (1.130) acts on the Hilbert space $L^2(\mathbb{R}^2, \mathbb{C}^2)$, i.e. on two-component spinor wavefunction. Hence, electrons possess a *pseudospin* degree of freedom (the observable $\vec{\sigma} = (\sigma_1, \sigma_2, \sigma_3)$), which is physically related to the decomposition of the graphene honeycomb lattice into two equivalent sublattices [100]. Note that while the continuous degrees of freedom ($\mathbf{x}$ and $\mathbf{p}$) are two-dimensional, the pseudospin is three-dimensional. In the following, in order to write equations in Pauli components, we use "cartesian" vectors whose third components are always set to 0. Thus, for example $\mathbf{p} = (p_1, p_2, 0)$ and we adopt the convention that $\frac{\partial}{\partial x_j} = 0$ when $j = 3$.

The unboundedness from below of the graphene Hamiltonian makes the construction of QFD equations for mixed states problematic, because the moments of the equilibrium states are clearly unbounded. This difficulty is physically interpreted as due to the availability of more and more energetically convenient states. A possible way to tackle this problem would be using the Fermi-Dirac statistics, which allows an alternative description of the negative energy electron population in terms of *holes*. Another possibility would be introducing in the Hamiltonian a quadratic correction of the form

$$h(\mathbf{x}, \mathbf{p}) = c\mathbf{p} \cdot \vec{\sigma} + \left(\frac{p^2}{2m^*} + V(\mathbf{x})\right)\sigma_0,$$

where m^* is some effective mass: this correction does not modify sensibly the dynamics in the vicinity of the Dirac point but corrects the bands for larger p's (making them resemble somehow to the $\mathbf{k}\cdot\mathbf{p}$ bands (1.78)) so that the Hamiltonian

becomes bounded from below. This is the approach followed in [119–121], to which we address the interested reader.

However, in this review we prefer not to deal with such an issue, and rather focusing on the *pure-state* QFD (Madelung-like).

The (dimensional) Wigner equations associated to the Hamiltonian (1.131) are:

$$\begin{cases} \dfrac{\partial w_0}{\partial t} + c\nabla_x \cdot \vec{w} + \Theta_\hbar[V]w_0 = 0 \\ \dfrac{\partial \vec{w}}{\partial t} + c\nabla_x w_0 + \Theta_\hbar[V]\vec{w} = \dfrac{2c}{\hbar}\mathbf{p} \times \vec{w} \end{cases} . \tag{1.133}$$

Let us consider the following moments associated to the Wigner matrix w:

$$\begin{aligned} n_0 &= \langle w_0 \rangle, && \textit{total density,} \\ n_s &= \langle w_s \rangle, && \textit{pseudospin densities,} \\ J_k &= \langle p_k w_0 \rangle, && \textit{pseudomomentum current,} \\ t_{sk} &= \langle p_k w_s \rangle, && \textit{pseudospin currents (“stress-like” tensor),} \end{aligned}$$

where $s = 1, 2, 3$ and $k = 1, 2$. Then, it is easy to deduce from Eq. (1.133) the following non-closed system of moment equations:

$$\begin{cases} \dfrac{\partial n_0}{\partial t} + c\,\dfrac{\partial n_k}{\partial x_k} = 0, \\ \dfrac{\partial n_s}{\partial t} + c\,\dfrac{\partial n_0}{\partial x_s} = \dfrac{2c}{\hbar}\epsilon_{sij}\, t_{ji}, \\ \dfrac{\partial J_k}{\partial t} + c\,\dfrac{\partial t_{sk}}{\partial x_s} + n_0\,\dfrac{\partial V}{\partial x_k} = 0, \end{cases} \tag{1.134}$$

where ϵ_{ijk} is the Levi-Civita symbol. In order to find closure relations, following [22, 103], we exploit some spinorial identities that hold for pure states.

Let us consider the density matrix associated to a factorized state:

$$\rho_{ij}(\mathbf{x}, \mathbf{y}) = \psi_i(\mathbf{x})\overline{\phi}_j(\mathbf{y})$$

(the pure state corresponds to the particular case $\psi = \phi$). Since, clearly, $\det \rho = 0$, we obtain for the Pauli components, the identity

$$\vec{\rho} \cdot \vec{\rho} = \rho_0^2,$$

which, recalling (1.48), yields

$$n_1^2 + n_2^2 + n_3^2 = n_0^2, \tag{1.135}$$

(i.e. $\vec{n} \cdot \vec{n} = n_0^2$, where, of course, $\vec{n} = (n_1, n_2, n_3)$). The identity (1.135) shows that, for a pure state, n_0 is not independent on $\vec{n}$: the modulus of the spin density vector must be equal to n_0. Next, it is readily shown that

$$(\rho \nabla_x \rho)_{ij} = \mathrm{tr}\,(\nabla_x \rho)\,\rho_{ij}, \qquad \left(\nabla_y \rho \rho\right)_{ij} = \mathrm{tr}\left(\nabla_y \rho\right)\rho_{ij}, \tag{1.136}$$

where tr denotes the (purely algebraic) matrix trace. By decomposing the identities (1.136) in Pauli components we obtain:

$$\begin{cases} \vec{\rho} \cdot \nabla_x \vec{\rho} = \rho_0 \nabla_x \rho_0 \\ \vec{\rho} \cdot \nabla_y \vec{\rho} = \rho_0 \nabla_y \rho_0 \\ i\,\vec{\rho} \times \nabla_x \vec{\rho} = \vec{\rho} \nabla_x \rho_0 - \nabla_x \vec{\rho}\,\rho_0 \\ i\,\vec{\rho} \times \nabla_y \vec{\rho} = -\vec{\rho} \nabla_y \rho_0 + \nabla_y \vec{\rho}\,\rho_0 \end{cases} \tag{1.137}$$

(where the vector operations have to be understood between the $\vec{\rho}$ and not involving the gradients). From the first two of equations (1.137) we get

$$\vec{\rho} \cdot \left(\nabla_x - \nabla_y\right) \vec{\rho} = \rho_0 \left(\nabla_x - \nabla_y\right) \rho_0,$$

which, recalling (1.48), yields

$$n_s\, t_{sk} = n_0\, J_k. \tag{1.138}$$

Finally, from the second two of equations (1.137) we get

$$\vec{\rho} \times \left(\nabla_x - \nabla_y\right) \vec{\rho} = i\rho_0 \left(\nabla_x + \nabla_y\right) \vec{\rho} - i\,\vec{\rho} \left(\nabla_x + \nabla_y\right) \rho_0,$$

which, passing to the moments, yields

$$\frac{2}{\hbar}\, \epsilon_{sij}\, n_i\, t_{jk} = n_0\, \frac{\partial n_s}{\partial x_k} - n_s\, \frac{\partial n_0}{\partial x_k}. \tag{1.139}$$

Identities (1.138) and (1.139) determine the parts of the vector $\vec{t}_k = (t_{1k}, t_{2k}, t_{3k})$ that are, respectively, parallel and orthogonal to $\vec{n}$, namely:

$$n_0\, \vec{t}_k = J_k \vec{n} - \frac{\hbar}{2}\, \vec{n} \times \frac{\partial \vec{n}}{\partial x_k} \tag{1.140}$$

(where Eq. (1.135) was also used), or

$$n_0 t_{sk} = J_k n_s - \frac{\hbar}{2}\epsilon_{sij} n_i \frac{\partial n_j}{\partial x_k}. \tag{1.141}$$

By substituting (1.141) in (1.134) we obtain a closed QFD system that reads as follows:

$$\begin{cases} \dfrac{\partial n_0}{\partial t} + c\,\mathrm{div}\,\vec{n} = 0\,, \\ \dfrac{\partial \vec{n}}{\partial t} + c\nabla n_0 = \dfrac{2c}{\hbar}\mathbf{u}\times\vec{n} - \dfrac{c}{n_0}\left(\vec{n}\,\mathrm{div}\,\vec{n} - (\vec{n}\cdot\nabla)\vec{n}\right), \\ \dfrac{\partial (n_0\mathbf{u})}{\partial t} + c\,\mathrm{div}\left(\mathbf{u}\otimes\vec{n}\right) = \dfrac{c\hbar}{2}\dfrac{\partial}{\partial x_s}\left(\dfrac{1}{n_0}\epsilon_{sij} n_i \nabla n_j\right) + n_0\,\mathbf{F}\,, \end{cases} \tag{1.142}$$

where $u_k = J_k/n_0$ is the velocity field associated to the pseudomomentum and $\mathbf{F} = -\nabla V$ is the force field. It is easily shown that, if $(n_0, \vec{n}, \mathbf{u})$ is a solution of (1.142) satisfying (1.135) at the initial time, then (1.135) remains true at all times. Hence, for pure states, the first equation of system (1.142) can be dropped and $n_0 = |\vec{n}|$. The five dependent variables $(\vec{n}, \mathbf{u})$ are still one more of the four (real) dependent variable necessary to specify the two-component wavefunction (ψ_1, ψ_2) that describe the pure state in the Schrödinger picture. However, it can be shown [22] that a further constraint can be deduced on $\mathbf{u}$, which actually reduces to four the fluid variables needed to describe a pure state, namely $\vec{n}$ and the projection of $\mathbf{u}$ on $\vec{n}$.

The dynamics described by the QFD equations (1.142) is very interesting. According to the continuity equation $\partial_t n_0 + c\,\mathrm{div}\,\vec{n} = 0$, and to Eq. (1.135), the velocity field of the density n_0 is $c\,\vec{n}/|\vec{n}|$, which means that the pseudospin $\vec{n}$ determines the direction of the flow (with constant speed c). The variation of $\vec{n}$ depends on the pseudomomentum velocity field $\mathbf{u}$, as well as to space-inhomogeneities. In particular, the force does not accelerate the particle directly but, rather, affects the pseudomomentum[7] field $\mathbf{u}$, which acts as a precession vector for $\vec{n}$. It is worth remarking that, although the QFD system (1.142) may seem at first sight more complicated than the Wigner system (1.133), the latter depends on five independent variables, (x_1, x_2, p_1, p_2, t), while the former on three, (x_1, x_2, t), with all the obvious advantages from the numerical point of view. Of course, the true advantage of the QFD approach with respect to the Schrödinger picture would be evident only in the case of mixed state. We believe that Eq. (1.142) should be

[7]This shows clearly that, indeed, p is not a momentum but a "pseudomomentum" (or "crystal momentum"): the latter has rather different properties.

considered, first of all, as an interesting, alternative, description and, secondly, as a first step towards a more desirable QFD model for mixed states.

To conclude, it is worth to introduce a simplified, one-dimensional version of (1.142). Let us assume that V depends only on the first coordinate x_1 and let us look for solution of the Schrödinger equation $i\hbar\psi = H\psi$ of the form

$$\psi(x_1, x_2, t) = \phi(x_1, t)\mathrm{e}^{ik_2x_2},$$

for some constant k_2, where ϕ is a (two-component) wave function. Then, it is easy to show that the evolution of ϕ is determined by the one-dimensional Hamiltonian

$$h(x_1, p_1) = cp_1\sigma_1 + \gamma\sigma_2 + V(x_1)\sigma_0, \tag{1.143}$$

where $\gamma = \hbar ck_2$. The energy bands associated to such Hamiltonian are

$$E_\pm(p_1) = \pm\sqrt{(cp_1)^2 + \gamma^2}, \tag{1.144}$$

corresponding to having "sliced" the cones (1.132) with the plane $p_2 = \gamma/c$. The QFD equations (1.142), in this case, simplify to a system for fluid variables that depend on the one-dimensional space variable $x \equiv x_1$ and which reads as follows:

$$\begin{cases} \dfrac{\partial n_0}{\partial t} + c\dfrac{\partial n_1}{\partial x} = 0\,, \\ \dfrac{\partial \vec{n}}{\partial t} + c\,\vec{e}_1\dfrac{\partial n_0}{\partial x} = \dfrac{2c}{\hbar}\mathbf{u}\times\vec{n} - \dfrac{c}{n_0}\left(\vec{n}\dfrac{\partial n_1}{\partial x} - n_1\dfrac{\partial \vec{n}}{\partial x}\right), \\ \dfrac{\partial (n_0u)}{\partial t} + c\dfrac{\partial (n_1u)}{\partial x} = \dfrac{c\hbar}{2}\dfrac{\partial}{\partial x}\left(\dfrac{n_3}{n_0}\dfrac{\partial n_2}{\partial x} - \dfrac{n_2}{n_0}\dfrac{\partial n_3}{\partial x}\right) + n_0\,F\,, \end{cases} \tag{1.145}$$

where $\partial_x \equiv \partial/\partial x \equiv \partial/\partial x_1$, $F = -V'$ and $\mathbf{u} := (u, \gamma/c, 0)$. As in the two-dimensional case, the first equation is actually redundant, since $n_0 = |\vec{n}|$ still holds. System (1.145) can be of interest, e.g., for numerical simulations of electron tunneling through a potential barrier perpendicular to the x axis, the parameter γ being in this case related to the incidence angle of the electron on the barrier [62]. It is well known that for incidence angle equal to zero (i.e. perpendicular incidence), corresponding to $\gamma = 0$, the barrier becomes perfectly transparent: this is the so called *Klein paradox* of relativistic quantum mechanics. Klein paradox is particularly evident in our QFD formulation (1.145), since for $\gamma = 0$, the equations for n_0 and n_1 are decoupled from the others, and n_0 does not see the force field any more (see also [120]).

Acknowledgements This work was partially supported by the Italian Ministry of University (MIUR National Project "Kinetic and hydrodynamic equations of complex collisional systems", PRIN 2009, Prot. n. 2009NAPTJF_003).

References

1. E.N. Adams, Motion of an electron in a perturbed periodic potential. Phys. Rev. **85**, 41–50 (1952)
2. G. Alì, G. Mascali, V. Romano, R.C. Torcasio, A hydrodynamic model for covalent semiconductors with applications to GaN and SiC. Acta Applicandae Mathematicae (2012)
3. G. Alì, D. Bini, S. Rionero, Global Existence and Relaxation Limit for Smooth Solutions to the Euler-Poisson Model for Semiconductors. SIAM J. Math. Anal. **32** 3, 572–587 (2000)
4. T. Ando, H. Akera, Connection of envelope functions at semiconductor heterointerfaces. II. Mixings of Γ and X valleys in GaAs/Al_xGa_{1-x}As. Phys. Rev. B **40**(17), 11619–11633 (1989)
5. L.C. Andreani, A. Pasquarello, F. Bassani, Hole subbands in strained GaAs-$Ga_{1-x}Al_x$-As quantum wells: Exact solution of the effective-mass equation, Phys. Rev. B **36**(11), 5887–5894 (1987)
6. A.M. Anile, O. Muscato, Improved hydrodynamical model for carrier transport in semiconductors. Phys. Rev. B **51**(23), 16728–16740 (1995)
7. A.M. Anile, V. Romano, Non parabolic band transport in semiconductors closure of the moment equations. Cont. Mech. Thermod. **11**, 307–325 (1999)
8. A. Arnold, C. Ringhofer An operator splitting method for the Wigner-Poisson problem. SIAM J. Numer. Anal. **33**(4) 1622–1643 (1996)
9. A. Arnold, Self-consistent relaxation-time models in quantum mechanics. Commun. Part. Diff. Eqs **21**(3–4), 473–506 (1996)
10. C. Auer, F. Schürrer, Semicontinuous kinetic theory of the relaxation of electrons in GaAs. Transport theory and statistical physics **33**, 429–447 (2004)
11. C. Auer, A. Majorana, F. Schürrer, Numerical schemes for solving the non-stationary Boltzmann-Poisson system for two-dimensional semiconductor devices. In: ESAIM Proceedings **15**, 75–86 (2005)
12. C. Auer, F. Schürrer, C. Ertler, Hot phonon effects on the high-field transport in metallic carbon nanotubes. Phys. Rev. B **74**(16), 165409 (2006)
13. T.B. Bahder, Eight-band **k·p** model of strained zinc-blende crystals, Phys. Rev. B **41**(17), 11992–12001 (1991)
14. L. Barletti, Wigner envelope functions for electron transport in semiconductor devices. Transp. Theory Stat. Phys. **32**(3/4), 253–277 (2003)
15. L. Barletti, N. Ben Abdallah, Quantum transport in crystals: effective-mass theorem and **k·p** Hamiltonians. Commun. Math. Phys. **307**, 567–607 (2011)
16. L. Barletti, C. Cintolesi, Derivation of isothermal quantum fluid equations with Fermi-Dirac and Bose-Einstein statistics. J. Stat. Phys. **148** 353–386 (2012)
17. L. Barletti, G. Frosali, Diffusive limit of the two-band **k·p** model for semiconductors. J. Stat. Phys. **139**(2), 280–306 (2010)
18. L. Barletti, G. Frosali, L. Demeio L.: multi-band quantum transport models for semiconductor devices. In: C. Cercignani, E. Gabetta (eds.), Transport Phenomena and Kinetic Theory, *Modeling and Simulation in Science, Engineering and Technology*, pp. 55–89. Birkhäuser, Boston (2007)
19. L. Barletti, F. Méhats, Quantum drift-diffusion modeling of spin transport in nanostructures. J. Math. Phys. **51**(5), 053304 (2010)
20. G. Bastard, *Wave mechanics applied to semiconductor heterostructures* (Les Editions de Physique, Halsted Press, Les Ulis Cedex, 1988)
21. N. Ben Abdallah, On a multidimensional Schrödinger-Poisson scattering model for semiconductors. J. Math. Phys. **41**(7), 4241–4261 (2000)
22. I. Bialynicki-Birula, Hydrodynamic form of the Weyl equation. Acta Physica Polonica **26**(7), 1201–1208 (1995)
23. Y. Bychkov, E.I. Rashba, Properties of a 2D electron gas with lifted spectral degeneracy. JETP Lett. **39**(2), 78–81 (1984)

24. P.L. Bhatnagar, E.P. Gross, M. Krook, A model for collision processes in gases. I. Small amplitude processes in charged and neutral one-component systems. Phys. Rev., **94**, 511–525 (1954)
25. D. Bohm, A suggested interpretation of the quantum theory in terms of "hidden variables" I. Phys. Rev. **85**, 166–179 (1952)
26. D. Bohm, A suggested interpretation of the quantum theory in terms of "hidden variables" II. Phys. Rev. **85**, 180–193 (1952)
27. L.L. Bonilla, L. Barletti, M. Alvaro, Nonlinear electron and spin transport in semiconductor superlattices, SIAM. J. Appl. Math. **69**(2) 494–513 (2008)
28. L.L. Bonilla, H.T. Grahn, Non-linear dynamics of semiconductor superlattices, Rep. Prog. Phys. **68** 577–683 (2005)
29. G. Borgioli, G. Frosali, P. Zweifel, Wigner approach to the two-band Kane model for a tunneling diode. Transp. Theory Stat. Phys. **32**(3/4), 347–366 (2003)
30. M.G. Burt, The justification for applying the effective-mass approximation to microstructure. J. Phys: Condens. Matter, **4**, 6651–6690 (1992)
31. C. Brouder, G. Panati, M. Calandra, C. Mourougane, N. Marzari, Exponential localization of Wannier functions in insulators. Phys. Rev. Lett. **98**, 046402 (2007)
32. F. Buot, K. Jensen, Lattice Weyl-Wigner formulation of exact many-body quantum-transport theory and applications to novel solid-state quantum-based devices. Phys. Rev. B **42**, 9429–9457 (1990)
33. V.D. Camiola, G. Mascali, V. Romano, Numerical simulation of a double-gate MOSFET with a subband model for semiconductors based on the maximum entropy principle. Cont. Mech. Thermodyn. **70**, 710 (2011)
34. P. Carruthers, F.C. Zachariasen, Quantum Collision Theory with Phase-Space Distributions. Rev. Mod. Phys. **55**(1), 245–285 (1983)
35. C.Y.-P. Chao, S.L. Chuang, Resonant tunneling of holes in the multi-band effective-mass approximation. Phys. Rev. B **43**(9), 7027–7039 (1991)
36. S.L. Chuang, Efficient band-structure calculations of strained quantum wells. Phys. Rev. B **43**(12), 9649–9661 (1991)
37. J. des Cloizeaux, Energy bands and projection operators in a crystal: Analytic and asymptotic properties. Phys. Rev. **135**(3A), A685–A697 (1964)
38. S. Datta, B. Das, Electronic analog of the electro-optic modulator, Appl. Phys. Lett. **56**(7), 665–667 (1990)
39. Davidovich, M.A., Anda, E.V., Tejedor, C., Platero, G.: Interband resonant tunneling and transport in InAs/AlSb/GaSb heterostructures, Phys. Rev. B, **47**(8), 4475–4484 (1993)
40. P. Degond, S. Gallego, F. Méhats, Isothermal quantum hydrodynamics: derivation, asymptotic analysis, and simulation. Multiscale Model. Simul. **6**(1), 246–272 (2007)
41. P. Degond, C. Ringhofer, Quantum moment hydrodynamics and the entropy principle. J. Stat. Phys. **112**(3–4), 587–628 (2003)
42. P. Degond, F. Méhats, C. Ringhofer, Quantum energy-transport and drift-diffusion models. J. Stat. Phys. **118**(3–4), 625–667 (2005)
43. L. Demeio, L. Barletti, A. Bertoni, P. Bordone, C. Jacoboni, Wigner-function approach to multi-band transport in semiconductors. Physica B **314**, 104–107 (2002)
44. L. Demeio, P. Bordone, C. Jacoboni, Numerical simulation of an intervalley transition by the Wigner-function approach. Semicond. Sci. Technol. **19**, 1–3 (2004)
45. Di Stefano, V.: Modeling thermal effects in submicron semiconductor devices, Commun Appl. Industr. Math. **1**(1), 110–117 (2010),
46. D. Dürr, S. Teufel, *Bohmian mechanics. The physics and mathematics of quantum theory* (Springer-Verlag, Berlin, 2009)
47. G.B. Folland, *Harmonic Analysis in Phase Space* (Princeton University Press, Princeton, 1989)
48. B.A. Foreman, Exact effective-mass theory for heterostructures. Phys. Rev. B **52**, 12241–12259 (1995)

49. B.A. Foreman, Envelope-function formalism for electrons in abrupt heterostructures with material-dependent basis functions. Phys. Rev. B **54**(3), 1909–1921 (1996)
50. M. Freitag, Graphene: Nanoelectronics goes flat out. Nature Nanotechnology **3**, 455–457 (2008)
51. W.R. Frensley, Boundary conditions for open quantum systems far from equilibrium. Rev. Mod. Phys. **62**, 745–791 (1990)
52. G. Frosali, O. Morandi, A quantum kinetic approach for modeling a two-band resonant tunneling diode. Transp. Theory Stat. Phys. **36**, 159–177 (2007)
53. S. Gallego, F. Méhats, Numerical approximation of a quantum drift-diffusion model. C.R. Math. Acad. Sci. Paris **339**(7), 519–524 (2004)
54. S. Gallego, F. Méhats, Entropic discretization of a quantum drift-diffusion model. SIAM J. Numer. Anal. **43**(5), 1828–1849 (2005)
55. M. Galler, F. Schürrer, A deterministic solution method for the coupled system of transport equations for the electrons and phonons in polar semiconductors, J. Phys. A: Math. Gen. **37** 1479–1497 (2004)
56. I. Gasser, P.A. Markowich, A. Unterreiter, Quantum hydrodynamics. In: P.A. Raviart (ed.) Modeling of Collisions, pp. 179–216. Gauthier-Villars, Paris (1997)
57. J. Genoe, K. Fobelets, C. van Hoof, G. Borghs, In-plane dispersion relations of InAs/AlSb/GaSb/AlSb/InAs interband resonant-tunneling diodes. Phys. Rev. B **52**(19), 14025–14034 (1995)
58. H. Haug, S.W. Koch, *Quantum Theory of the Optical and Electronic Properties of Semiconductors* (4th Edition, World Scientific, Singapore, 2004)
59. G.J. Iafrate, J.B. Krieger, Quantum transport for Bloch electrons in inhomogeneous electric fields. Phys. Rev. B **40**(9), 6144–6148 (1989)
60. K.L. Jensen, F.A. Buot, Numerical simulation of intrinsic bistability and high-frequency current oscillations in resonant tunneling structures. Phys. Rev. Lett. **66**, 1078–1081 (1991)
61. A. Jüngel, D. Matthes, A derivation of the isothermal quantum hydrodynamic equations using entropy minimization. Z. Angew. Math. Mech. **85**(11), 806–814 (2005)
62. M.I. Katsnelson, K.S. Novoselov, A.K. Geim, Chiral tunnelling and the Klein paradox in graphene. Nature Physics **2**(9), 620–625 (2006)
63. E.O. Kane, Energy band structure in p-type Germanium and Silicon. J. Phys. Chem. Solids **1**, 82–89 (1956)
64. W. Kohn, Analytic Properties of Bloch Waves and Wannier Functions. Phys. Rev. **115**, 809–821 (1959)
65. D.Y.K. Ko, J.C. Inkson, Matrix method for tunneling in heterostructures: Resonant tunneling in multilayer systems. Phys. Rev. B **38**, 9945–9951 (1988)
66. H. Kosina, M. Nedjalkov, Review Chapter: Wigner Function Based Device Modeling. In: Handbook of Theoretical and Computational Nanotechnology (Los Angeles), 731–763 (2006)
67. H.-W. Lee, Theory and application of the quantum phase-space distribution functions, Phys. Rep. **259**, 147–211 (1995)
68. L. Leibler, Effective-mass theory for carriers in graded mixed semiconductors. Phys. Rev. B **12** 4443–4451 (1975)
69. P. Lichtenberger, O. Morandi, F. Schürrer, High field transport and optical phonon scattering in graphene, Phys. Rev. B **84**, 045406 (2011)
70. J.M. Luttinger, W. Kohn, Motion of electrons and holes in perturbed periodic fields. Phys. Rev. **97**(4), 869–882 (1955)
71. E. Madelung, Quantentheorie in hydrodynamischer Form. Zeitschr. f. Phys. **40**, 322–326 (1926)
72. R.K. Mains, G.I. Haddad, Wigner function modeling of resonant tunneling diodes with high peak-to-valley ratios. J. Appl. Phys. **64**, 5041–5044 (1988)
73. C. Manzini, L. Barletti, An analysis of the Wigner-Poisson problem with inflow boundary conditions. Nonlinear Analysis, **60**(1), 77–100 (2005)

74. N. Marzari, D. Vanderbilt, Maximally localized generalized Wannier functions for composite energy bands, Phys. Rev. B **56**, 12847–12865 (1997)
75. G. Mascali, V. Romano, A non parabolic hydrodynamical subband model for semiconductors based on the maximum entropy principle. Math. Comput. Modeling **55**(3–4), 1003–1020 (2012)
76. F. Méhats, Adiabatic Approximation of the Schrödinger–Poisson System with a Partial Confinement. SIAM J. Math. Anal. **36**(3), 986–1013 (2005)
77. F. Méhats, Analysis of a Quantum Subband Model for the Transport of Partially Confined Charged Particles. Monatsh. Math. **147**, 43–73 (2006)
78. F. Méhats, O. Pinaud, An inverse problem in quantum statistical physics. J. Stat. Phys. **140**, 565–602 (2010)
79. F. Méhats, O. Pinaud, A problem of moment realizability in quantum statistical physics. Kinet. Relat. Models **4**(4), 1143–1158 (2011)
80. A.T. Meney, B. Gonul, E.P. O'Reilly, Evaluation of various approximations used in the envelope-function method. Phys. Rev. B **50**(15), 10893–10904 (1994)
81. O. Morandi, multi-band Wigner-function formalism applied to the Zener band transition in a semiconductor. Phys. Rev. B **80**, 024301 (2009)
82. O. Morandi, Effective classical Liouville-like evolution equation for the quantum phase space dynamics. J. Phys. A: Math. Theor. **43** 365302 (2010)
83. O. Morandi, Quantum Phase-Space Transport and Applications to the Solid State Physics. In M.R. Pahlavani (ed.), Some Applications of Quantum Mechanics, 1–26. InTech, (2012)
84. O. Morandi, M. Modugno, A multi-band envelope function model for quantum transport in a tunneling diode. Phys. Rev. B **71** 235331 (1–8) (2005)
85. R.A. Morrow, K.R. Brownstein, Model effective-mass Hamiltonians for abrupt heterojunctions and the associated wave-function-matching conditions. Phys. Rev. B **30**, 678–680 (1984)
86. J.E. Moyal, Quantum mechanics as a statistical theory. Math. Proc. Cambr. Philos. Soc. **45**, 99–124 (1949)
87. O. Muscato, Validation of an Extended Hydrodynamical model for a submicron npn Bipolar Junction Transistor. Physica A **365**(2), 409–428 (2006)
88. O. Muscato, Hot electron distribution function for the Boltzmann equation with analytic bands. J. Comput. Electron. **5**, 377–380 (2006)
89. O. Muscato, V. Di Stefano, Modeling heat generation in a sub-micrometric n+ - n - n+ silicon diode. J. Appl. Phys.**104**, 124501 (2008)
90. M. Nedjalkov, D. Querlioz, P. Dollfus, H. Kosina, Review Chapter: Wigner Function Approach. In: Nano-Electronic Devices: Semiclassical and Quantum Transport Modeling, pp. 1–76, Springer-Verlag, Berlin (2011)
91. G. Nenciu, Existence of the exponentially localised Wannier functions. Comm. Math. Phys. **91**, 81–85 (1983)
92. M. Ogawa, R. Tominaga, T. Miyoshi, Multi-Band Simulation of Interband Tunneling Devices Reflecting Realistic Band Structure. IEICE Trans. Electr. **E83-C**(8) 1235–1241 (2000)
93. D. Querlioz, P. Dollfus, The Wigner Monte Carlo Method for Nanoelectronic Devices – A particle description of quantum transport and decoherence. ISTE-Wiley, (2010)
94. G. Panati, A. Pisante, Bloch bundles, Marzari-Vanderbilt functional and maximally localized Wannier functions. arXiv:1112.6197 [math-ph] (2011)
95. C. Reitshammer, F. Schürrer, A. Rossani, Atomic Beam Slowing and Cooling: Discrete Velocity Model. Phys. Rev. E. **58** 3964–3970 (1998)
96. C. Ringhofer, Sub-band diffusion models for quantum transport in a strong force regime. SIAM J. Appl. Math. **71**(6), 1871–1895 (2011)
97. V. Romano, Non parabolic band transport in semiconductors: closure of the production terms in the moment equations. Cont. Mech. Thermodyn. **12**, 31–51 (1999)
98. V. Romano, Non-parabolic band hydrodynamical model of silicon semiconductors and simulation of electron devices. Math. Meth. Appl. Sci. **24**, 439–471 (2001)

99. L. Shifren, C. Ringhofer, D.K. Ferry, A Wigner Function-Based Quantum Ensemble Monte Carlo Study of a Resonant Tunneling Diode. IEEE Trans. Electron Devices **50**(3), 769–773 (2003)
100. J.C. Slonczewski, P.R. Weiss, Band structure of graphite. Phys. Rev. **109**(2), 272–279 (1958)
101. M. Sweeney J.M. Xu, Resonant interband tunnel diodes. Appl. Phys. Lett. **54**(6), 546–548 (1989)
102. D. Taj, L. Genovese, F. Rossi, Quantum-transport simulations with the Wigner-function formalism: failure of conventional boundary-condition schemes. Europhys. Lett. **74**(6), 1060–1066 (2006)
103. T. Takabayasi, The vector representation of spinning particle in the quantum theory. Prog. Theor. Phys. **14**(4), 283–302 (1955)
104. B. Thaller, *The Dirac Equation* (Springer-Verlag, Berlin, 1992)
105. S. Teufel, *Adiabatic Perturbation Theory in Quantum Dynamics* (Springer-Verlag, Berlin, 2003)
106. S. Teufel, G. Panati, Propagation of Wigner functions for the Schrödinger equation with a perturbed periodic potential. arXiv:math-ph/0403037 (2004)
107. M. Trovato, L. Reggiani, Quantum maximum entropy principle for a system of identical particles. Phys. Rev. E **81**, 021119 (2010)
108. M. Trovato, L. Reggiani, L.: Quantum hydrodynamic models from a maximum entropy principle. J. Phys. A: Math. Theor. **43**, 102001 (2010)
109. M. Trovato, L. Reggiani, Maximum entropy principle and hydrodynamic models in statistical mechanics. La Rivista del Nuovo Cimento **35**, 99–266 (2012)
110. R. Tsu, L. Esaki, Tunneling in a finite superlattice. Appl. Phys. Lett. **22**(11), 562–564 (1973)
111. M.B. Unlu, B. Rosen, H.-L. Cui, P. Zhao, Multi-band Wigner function formulation of quantum transport. Phys. Lett. A **327**(2–3), 230–240 (2004)
112. C.M. van Vliet, A.H. Marshak, Wannier-Slater theorem for solid with nonuniform band structure. Phys. Rev. B **26**(12), 6734–6738 (1982)
113. J. Von Neumann, *Mathematical Foundations of Quantum Mechanics* (Princeton University Press, Princeton, 1955)
114. G.H. Wannier, Dynamics of band electrons in electric and magnetic fields. Rev. Mod. Phys. **34**, 645–655 (1962)
115. W.T. Wenckebach, *Essential of Semiconductor Physics* (John Wiley & Sons, Chichester, 1999)
116. G.Y. Wu, K.-P. Wu, Electron transport in a resonant-tunneling diode under the effect of a transverse magnetic field: a quantum theory in the Wigner formalism. J. Appl. Phys. **71**(3), 1259–1264 (1992)
117. E. Wigner, On the quantum correction for thermodynamic equilibrium. Phys. Rev. **40**, 749–759 (1932)
118. R.Q. Yang, M. Sweeny, D. Day, J.M. Xu, Interband tunneling in heterostructure tunnel diodes. IEEE Trans. Electron Devices **38**(3), 442–446 (1991)
119. N. Zamponi, Some fluid-dynamic models for quantum electron transport in graphene via entropy minimization. Kinet. Relat. Mod. **5**(1), 203–221 (2012)
120. N. Zamponi, L. Barletti, Quantum electronic transport in graphene: a kinetic and fluid-dynamical approach. Math. Meth. Appl. Sci. **34**, 807–818 (2011)
121. N. Zamponi, A. Jüngel, Two spinorial drift-diffusion models for quantum electron transport in graphene. Comm. Math. Sci. (in press)
122. I. Žutić, J. Fabian, S. Das Sarma, Spintronics: fundamentals and applications. Rev. Mod. Phys. **76**(2), 323–410 (2002)

Chapter 2
Electronic Properties of III-V Quantum Dots

Andrei Schliwa, Gerald Hönig, and Dieter Bimberg

Abstract Electronic properties of quantum dots are reviewed based on eight-band **k·p** theory. We will focus on the following interrelated subjects: First the role of crystallographic symmetry is evaluated. This includes the symmetry of the lattice of the substrate [wurtzite (wz) versus zinc blende (zb)] as well as different substrate orientations [zb-(001) versus zb-(111)]. Second, we discuss two different types of band alignment, type-I versus type-II, by comparing the common-anion system zb-InAs/GaAs to the common-cation system zb-GaSb/GaAs. Finally, the impact of large built-in fields resulting from piezo- and pyroelectric charges will be exemplified for the wz-GaN/AlN QD-system.

2.1 Introduction

Semiconductor quantum dots (QD) are fascinating physical subjects exhibiting electronic properties even simpler than hydrogen but in a dielectric cage, thus merging semiconductor with atomic physics. Nano-structures based on III-V-system material combinations alone cover a huge range of very different electronic and optical properties. Their tremendous tunability will be exemplified in this contribution by focusing on three cornerstones of current research: (i) InAs/GaAs-, (ii) GaAs/GaSb-, and (iii) GaN/AlN quantum dots.

(i) InAs/GaAs QDs have been subject of intense research for the last 20 years [13, 29], which led—most prominently—to the development of quantum dot

A. Schliwa (✉) • G. Hönig • D. Bimberg
Institut für Festkörperphysik, Technische Universität Berlin, Hardenbergstr. 36, 10623 Berlin, Germany
e-mail: andrei.schliwa@gmail.com; gerald.hoenig@physik.tu-berlin.de; bimberg@physik.tu-berlin.de

M. Ehrhardt and T. Koprucki (eds.), *Multi-Band Effective Mass Approximations*, Lecture Notes in Computational Science and Engineering 94,
DOI 10.1007/978-3-319-01427-2_2,

lasers [40] and single-photon emitters [48]. As InAs is a small band-gap material, including coupling between valence- and conduction bands became essential for the predictive power of any electronic structure calculations. This demand led to the development of 3D QD models going beyond effective mass theory [35]. Most transparent are calculations based on eight-band **k·p** theory [74]. Triggered by the quest for single and entangled photon emitters, later on, the role of substrate orientation went into the focus of research [65,71] which will be one topic of this chapter.

(ii) While the InAs/GaAs system exhibits a type-I quantum confinement, where both electron *and* hole are localized in the QD, GaSb/GaAs quantum dots feature a type-II band-alignment: The confinement is attractive alone for holes but repulsive for electrons. If holes are stored in the QD, eventually the band structure is modified to allow electrons to be localized nearby at the interface by means of Coulomb interaction. The difference between the two material combinations with respect to band-alignment can be traced back to the fact that InAs/GaAs share a common-anion- and GaSb/GaAs a common cation, as will be further elaborated in Sect. 2.3.2.

(iii) The third material system of huge interest is GaN/AlN. The difference to the former two is threefold: First, state-of-the-art QDs are synthesized in the wurtzite phase, as opposed to the zinc blende lattice of the preceding systems. Second, GaN has a large band gap and, third, the system GaN/AlN exhibits very strong internal piezo- and pyroelectric fields . We will see in Sect. 2.3.3 that for GaN/AlN QDs simply by variation of the QD size a large (interband) photon emission tunability ranging from 2.6 to 4.5 eV is obtained.

All three heterostructures belong to the III-V material system and coherent growth is mediated by the sufficiently large lattice mismatch leading to QD formation in the Stranski-Krastanov growth mode [68]. Strain alone already imposes significant band-shifts and band-splittings (heavy-hole/light-hole). Moreover, the lack of inversion symmetry in all III-V systems together with the inhomogeneity of the strain give rise to additional piezoelectric built-in fields.

The appeal of eight-band **k·p** theory lies in the attractive balance between accuracy, speed of computation, and transparency of the physics and the parameters used. All relevant phenomena, such as confinement, strain, piezo- and pyroelectricity, band-coupling and -splitting can be reliably treated for any shape, size and composition as long as the involved materials feature a direct band gap. This opens the opportunity to employ the model for the purpose of inverse bandstructure modeling [64] and inverse design [51].

2.1.1 Role of Lattice Symmetries (Zinc Blende vs Wurtzite)

The majority of quantum dots either carry zinc blende or wurtzite symmetry. Sometimes—as in the case of GaN—both allotropes are proven to exist. The zinc

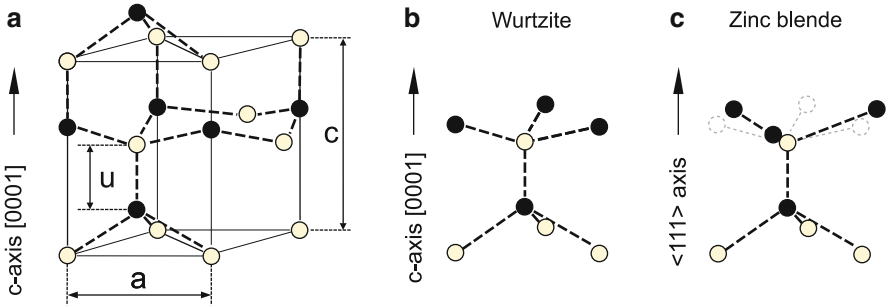


Fig. 2.1 (**a**) Wurtzite unit cell with lattice parameters a and c together with the internal parameter u. (**b** and **c**) The differences in position of second nearest neighbours between wurtzite and zinc blende lattice are shown

blende and wurtzite lattice differ only in the second-nearest neighbours [see Fig. 2.1b,c]. Analogous to the hexagonal and cubic close packing one can describe them by different layer sequences $ABABABA$ (along c-axis) or $ABCABCA$ (along <111>-axis).

These minute differences however result in very different symmetry properties: the zinc blende crystal has four threefold axes of rotation (<111>) and, thus, a higher symmetry than the wurtzite lattice with only one axis with threefold rotational symmetry (c-axis = [0001]). The latter gives rise to two important peculiarities: (1) the emergence of spontaneous polarization resulting in the pyroelectric effect (often used synonymously), and as a consequence (2) the crystal field splitting leading to a splitting of heavy and light hole (also referred to as A and B bands).

Whereas the zinc blende unit cell is sufficiently described by one lattice constant, a, for the wurtzite unit cell two lattice constants, a and c, together with an internal parameter u are required to adequately describe the lattice structure (see Fig. 2.1). The structure is composed of two interpenetrating hexagonal closed package sublattices. For the ideal wurtzite unit cell the ratio of c and a is $c/a = \sqrt{8/3} = 1.633$ and the internal parameter u has a value of $u = 3/8 = 0.375$ in fractional coordinates. The lattice is composed of tetrahedra with four atoms of one species tetrahedrally coordinated around a central atom of the other species. In case of an ideal ratio of lattice constants and ideal internal parameter the bond length and the bond angles between the nearest neighbors are equal, but the distance to the second nearest neighbor along the c-axis is about 13 % shorter than the distance to the other second nearest neighbors [2]. Hence, as was calculated by Bernardini et al. [9] already in this ideal case a spontaneous polarization arises.

GaN departs only little from the ideal situation with a $(c/a)_{\mathrm{GaN}} = 1.627$ and $u_{\mathrm{GaN}} = 0.377$, whereas AlN shows more significant deviations with $(c/a)_{\mathrm{AlN}} = 1.601$ and $u_{\mathrm{AlN}} = 0.382$, which correlates with the difference between the electronegativities of the two constituents [9].

The situation is schematically depicted in Fig. 2.2. One important consequence is the non-ideality of the tetrahedra as shown in Fig. 2.2b which amplifies the

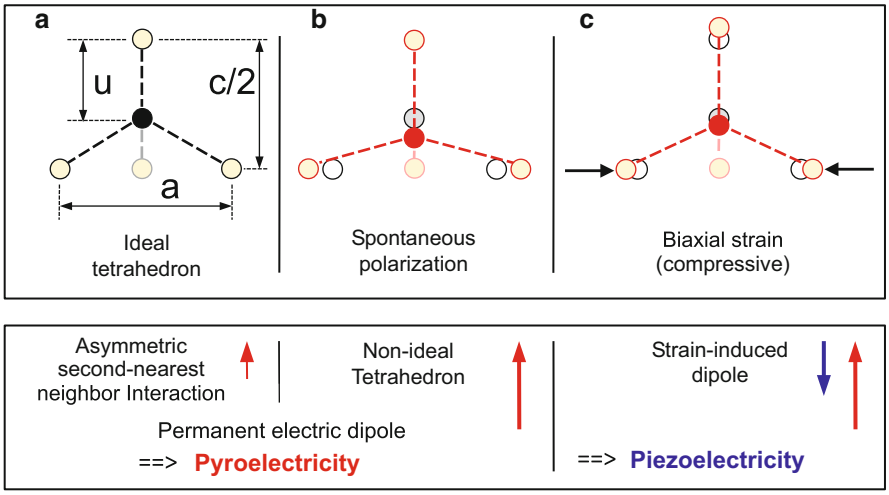


Fig. 2.2 (**a**) Ideal and (**b**,**c**) non-ideal tetrahedra. Due to asymmetric second-nearest neighbor interaction, already in the case of an ideal tetrahedron, differences in electronegativity between cation and anion lead to spontaneous polarization [3, 9]. (**b**) The pyroelectric effect is enhanced by the spontaneous polarization causing a non-ideal tetrahedron. (**c**) Strain (here biaxially compressive) leads to piezoelectricity

(pyroelectric) polarization of the lattice. For heterostructures such as wz-GaN/AlN differences between the spontaneous polarization fields ($P_{\mathrm{sp}}^{\mathrm{AlN}} = -0.09\,\mathrm{C/m^2}$ compared to $P_{\mathrm{sp}}^{\mathrm{GaN}} = -0.034\,\mathrm{C/m^2}$) result in large residual charges at the heterointerface being the root for the large built-in fields in GaN/AlN heterostructures.

Because the lattice mismatch between AlN and GaN is about 2.5 %, coherent growth leads to strain and, hence, to an additional piezoelectric field superimposed to the pyroelectric field as depicted in Fig. 2.2c.

In zinc blende crystals, in particular for the zb-III-V system, the piezoelectric effect is much smaller than for wz-III-nitride material. Nevertheless, it breaks the confinement symmetry even in circular symmetric quantum dots, and presents one important factor for the excitonic fine-structure splitting. The mechanism, how shear strain causes a piezoelectric polarization is schematically shown in Fig. 2.3.

2.2 Method of Calculation

Figure 2.4 shows schematically the modeling procedure employed in this work. It starts with an implementation of the 3D QD model structure (size, shape, chemical composition), and continues with the calculation of strain, piezoelectricity, and pyroelectricity (wurtzite only). The resulting strain and polarization fields enter the eight-band $\mathbf{k}\cdot\mathbf{p}$ Hamiltonian. Solution of the Schrödinger equation yields

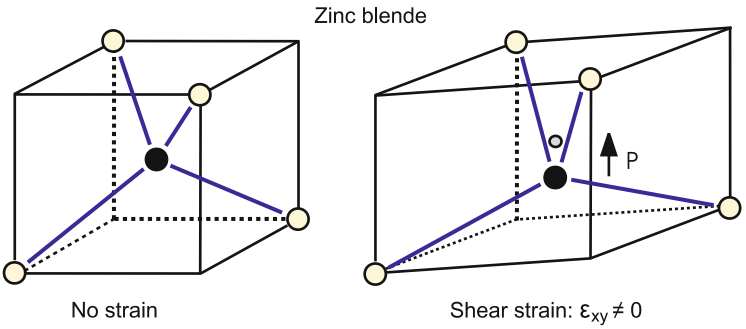


Fig. 2.3 Part of the zinc blende unit cell in absence of strain (*left*) and presence of one shear strain component, $\epsilon_{xy} \neq 0$ (*right*): Cations and anions do not share the same center of charge anymore, which results in a piezoelectric dipole

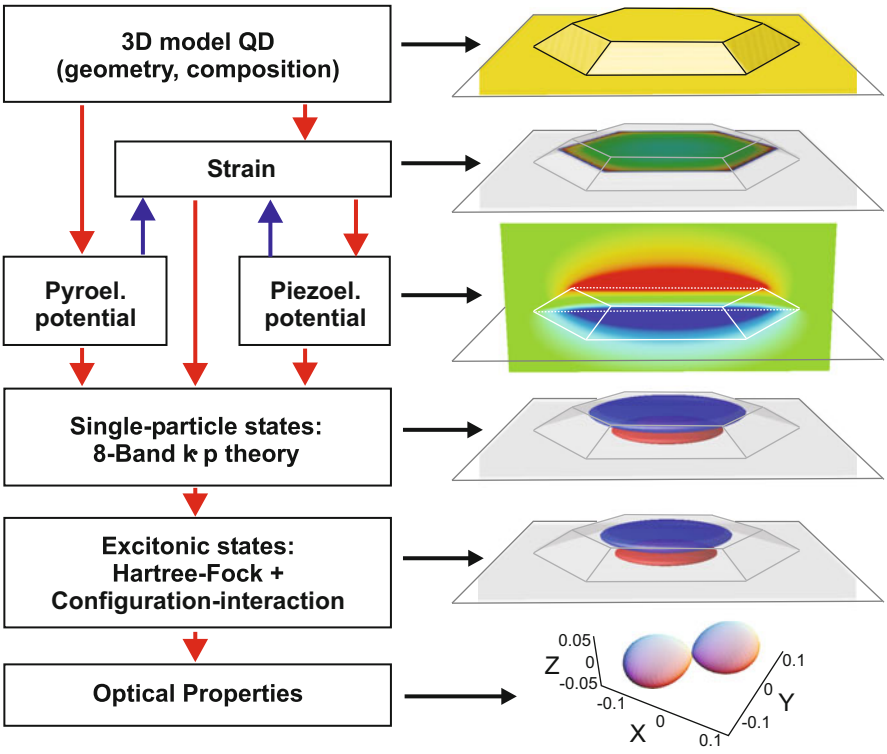


Fig. 2.4 Schematics of the modeling procedure applied in this work

electron and hole single-particle states. Coulomb interaction is accounted for by employing the configuration interaction method based either on the pristine single-particle states or on Hartree-Fock modified states. Finally, optical properties such as absorption spectra, capture cross sections or lifetimes can be calculated.

2.2.1 Calculation of Strain

Since the impact of strain on the confinement is comparable to that of the band offset resulting from the variation of the chemical composition at the heterojunctions, the wavefunctions and energies strongly depend on the underlying strain distribution. The impact of our model used for calculating the strain distribution has been analysed in a number of publications [55, 74]. Stier et al. [74] argue that the continuum elasticity model (CM) gives better results for QDs than the valence force field (VFF) model (Keating) in its linearized version (Kane). The major part of the differences in the strain distribution are attributed to the incorrect value of C_{44} in the VFF model and not to its atomistic character. Later Zunger et al. [80] introduced a generalized version of the VFF model—the G-VFF model—where C_{44} is incorporated correctly. Although the potential of the Keating model in its original version is not harmonic, it has been remarked by Kane [38], that anharmonic effects due to higher order terms are not satisfactorily treated. Therefore Klimeck and coworkers extended this model to include them correctly [42]. The same issue is addressed by Hammerschmidt et al. [30] and Migliorato et al. [49] who employed the Tersoff-potential method [76].

The choice of the most appropriate strain model depends on the choice of the model for the electronic structure calculations. Since the eight-band $\mathbf{k}\cdot\mathbf{p}$ model presents a continuum approach, an atomistic strain model cannot unfold its full potential for two reasons:

First, the mapping of the atomic positions onto a strain tensor field is associated with a loss of information. To describe the positions of four tetrahedrally coordinated In atoms around an As atom five times the three spatial dimensions $= 15$ parameters are required. The strain tensor field on the other hand is described by only six independent components at each local position.

Second, the $\mathbf{k}\cdot\mathbf{p}$ model provides only a limited number of parameters to account for the strain, the model is not sensitive to the complete information an atomistic model provides. For example for a QD having a fourfold rotational C_{4v} symmetry the strain tensor derived from the CM model has C_{4v} symmetry too. The tetrahedral configuration of the atoms in the atomistic models in contrast [55] leads to C_{2v} symmetry, i.e. the strain components are different along the [110] and [1$\bar{1}$0] directions and the p-states will split. In our approach a structural $C_{\infty v}$ or C_{4v} symmetry is reduced in the second step by the inclusion of the piezoelectric field.

2.2.2 Piezoelectricity/Pyroelectricity

2.2.2.1 Zinc Blende Crystal

Piezoelectricity is defined as the generation of electric polarization by application of stress to a crystal lacking a center of symmetry [1]. The zinc blende structure is the

simplest example of such a lattice and the strength of the resulting polarization is described by one parameter alone, e_{14}, alone for the linear case and three parameters, B_{114}, B_{124} and B_{156} for the quadratic case [10] resulting in

$$\begin{aligned}
\mathbf{P}_1 &= 2\, e_{14} \begin{pmatrix} \epsilon_{yz} \\ \epsilon_{xz} \\ \epsilon_{xy} \end{pmatrix} \quad , \\
\mathbf{P}_2 &= 2\, B_{114} \begin{pmatrix} \epsilon_{xx}\, \epsilon_{yz} \\ \epsilon_{yy}\, \epsilon_{xz} \\ \epsilon_{zz}\, \epsilon_{xy} \end{pmatrix} + \\
&+ 2\, B_{124} \begin{pmatrix} \epsilon_{yz}\, (\epsilon_{yy} + \epsilon_{zz}) \\ \epsilon_{xz}\, (\epsilon_{zz} + \epsilon_{xx}) \\ \epsilon_{xy}\, (\epsilon_{xx} + \epsilon_{yy}) \end{pmatrix} + \\
&+ 4\, B_{156} \begin{pmatrix} \epsilon_{xz}\, \epsilon_{xy} \\ \epsilon_{yz}\, \epsilon_{xy} \\ \epsilon_{yz}\, \epsilon_{xz} \end{pmatrix} \quad .
\end{aligned} \tag{2.1}$$

The resulting polarization $\mathbf{P}_{\text{zb}}$ then comprises of two components

$$\mathbf{P}_{\text{zb}} = \mathbf{P}_1 + \mathbf{P}_2 \quad .$$

2.2.2.2 Wurtzite Crystal

The total polarization $\mathbf{P}_{\text{wz}}$ in wurtzite-type semiconductors is given by

$$\mathbf{P}_{\text{wz}} = \mathbf{P}_{\text{spont}} + \mathbf{P}_{\text{piezo}} \quad ,$$

where $\mathbf{P}_{\text{piezo}}$ is the strain-induced piezoelectric polarization and $\mathbf{P}_{\text{spont}}$ the spontaneous polarization.

Due to the reduced symmetry of the wurtzite lattice three parameters, e_{15}, e_{31} and e_{33} are required to capture the piezoelectric polarization, resulting in

$$\mathbf{P}_{\text{piezo}} = \begin{pmatrix} 2\, e_{15} \epsilon_{xz} \\ 2\, e_{15} \epsilon_{yz} \\ e_{31} \epsilon_{xx} + e_{31} \epsilon_{yy} + e_{33} \epsilon_{zz} \end{pmatrix} \quad .$$

The spontaneous polarization $\mathbf{P}_{\text{spont}}$ is given by

$$\mathbf{P}_{\text{spont}} = \begin{pmatrix} 0 \\ 0 \\ P_{\text{spont}} \end{pmatrix} \quad .$$

2.2.2.3 Built-in Potential

The polarization fields $\mathbf{P}_{\mathrm{zb}}$ for the zinc blende crystal or $\mathbf{P}_{\mathrm{wz}}$ for the wurtzite crystal, respectively, induce a charge distribution ρ_{P}:

$$\rho_{\mathrm{P}}(\mathbf{r}) = -\nabla \cdot \mathbf{P}_{\mathrm{zb/wz}} \quad .$$

The resulting built-in potential is obtained by solving Poisson's equation, taking into account the material dependent static dielectric constants $\epsilon_s(\mathbf{r})$

$$\rho_{\mathrm{p}}(\mathbf{r}) = \epsilon_0 \nabla \cdot \left\{ \epsilon_s(\mathbf{r}) \nabla V_{\mathrm{p}}(\mathbf{r}) \right\} , \tag{2.2}$$

$$\Leftrightarrow$$

$$\Delta V_{\mathrm{p}}(\mathbf{r}) = \frac{\rho_{\mathrm{p}}}{\epsilon_0 \epsilon_s(\mathbf{r})} - \frac{1}{\epsilon_s(\mathbf{r})} \nabla V_{\mathrm{p}}(\mathbf{r}) \cdot \nabla \epsilon_s(\mathbf{r}) . \tag{2.3}$$

The first term on the right hand side of Eq. 2.3 refers to the true three-dimensional charge density while the second is the contribution of polarization interface charge densities due to a discontinuous $\epsilon_s(\mathbf{r})$ across heterointerfaces.

The importance of the second order term $\mathbf{P}_2$ for InGaAs/GaAs(111) quantum wells (QW) and QDs has been pointed out by Bester et al. [10]. They found that for QWs the linear and quadratic coefficients have opposite effects on the field, and for large strain the quadratic term even dominates. For InAs/GaAs QDs, however, the situation is more complex since in addition to the large strain their three-dimensional structure comes into play: The linear term generates a quadrupole-like potential which reduces a structural C_{4v}- or $C_{\infty v}$-symmetry of a QD to C_{2v} [11,29]. The effect of the quadratic terms has been evaluated recently by Bester et al. [12] for lens-shaped QDs and was found to cancel the first order potential inside the QD leading to a field free QD. The investigation was later extended to a variety of more realistic structures including truncated pyramids, and non-evenly alloyed QDs [63]. For a pyramidal model QD having a base length of 17 nm and {101} side facets the strength and distribution of the piezoelectric potential resulting from the two orders of the piezoelectric tensor are displayed in Fig. 2.5. Apart from the different orientation and sign of the two contributions, an important peculiarity of the second order potential is its restriction to the interior of the QD which is in apparent contrast to the widely extended first order field. The difference is linked to the origin of the polarization $\mathbf{P}$: $\mathbf{P}_1$ is a function of the shear-strain components alone, whereas $\mathbf{P}_2$ results mainly from the product of the diagonal and the shear-strain. However, in contrast to the shear-strain components, the diagonal elements ϵ_{ii} are large only inside the QD and its close vicinity and therefore $\mathbf{P}_2$-charges can only be created in this region.

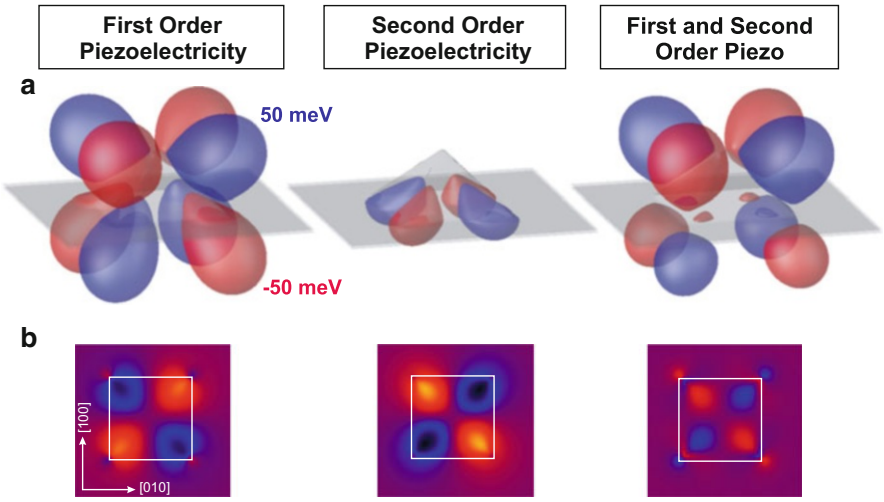


Fig. 2.5 (**a**) The piezoelectric potential isosurfaces at ±50 meV of a pyramidal InAs quantum dot with 17 nm base length shown for the linear, the quadratic part and for both. (**b**) Contour plots of the piezoelectric potential 2 nm above the wetting layer. (see [63])

2.2.3 Eight-Band **k·p** *Method: Single Particle States*

The energy levels and wavefunctions of bound electron and hole states are calculated using the eight-band **k·p** model. The theory was originally developed for the description of electronic states in bulk material [23, 37, 52]. For the use in heterostructures the envelope function version of the model has been developed and applied to quantum wells [27], quantum wires [73] and quantum dots [35,44,54,74]. Details of the principles of our implementation for zinc blende heterostructures are outlined in [73].

This model enables us to treat QDs of arbitrary shape and material composition, including the effect of strain, piezoelectricity, VB mixing and CB-VB interaction. The strain enters our model via the use of deformation potentials as outlined by Bahder [8]. Its impact on the local band edges as a function of the QD geometry will be discussed in the next sections.

The **k·p** model, when applied to small quantum structures, has in principle a few well-known drawbacks which have been examined in detail in [25, 39]. They are basically related to the fixed number of Bloch functions used for the wavefunction expansion, the restriction to the close vicinity of the Brillouin zone center Γ and to the limited ability to account for the symmetry of the underlying lattice. These problems do not arise in microscopic theories like the empirical pseudopotential [39] (EPM) or the empirical tight-binding method [43, 61] (ETB), which a priori have greater potential of accuracy. This potential, however, can only be exploited if the corresponding input parameters—the form factors in the

EPM or the tight-binding parameter and their strain dependence in the ETB—are known with sufficient accuracy. Reliable generation of these parameters, however, is highly nontrivial and yet at least controversial if not unsolved. One of the most appealing features of the **k·p** model, in contrast, is the direct availability of all the parameters entering the calculations and the corresponding transparency of the method. Additionally, the required computational expense of the method is comparatively small.

k·p models for wurtzite nitride QDs have been presented before, most prominent by Andreev and O'Reilly [4] and Fonoberov and Balandin [24]. The differences between these two models and the model used in the present work have been discussed in [81] and shall be summarized briefly here:

(i) The method introduced by Andreev and O'Reilly [4] includes all important effects except spin- orbit splitting, which has been neglected in order to reduce the dimensions of the Hamiltonian from 8×8 to 4×4. This simplification can be justified, given that the spin-orbit splitting is small in GaN (17 meV) and AlN (19 meV) and modifies the absolute value of the exciton transition energies roughly by the same amount. InN shows an even smaller spin-orbit splitting of 5 meV [77]. However, neglecting spin-orbit splitting leads to an artificial degeneracies in the hole spectra, in particular of the hole ground states [81].

(ii) Fonoberov and Balandin [24] use a 6×6 Hamiltonian for the valence bands (VBs), and the effective mass approximation for the conduction band (CB). This method neglects the coupling between VBs and CB, which is justified for large band-gap materials such as GaN and AlN. InN, in contrast, has a much smaller band gap of ≈ 0.7 meV and therefore requires the inclusion of VB/CB-coupling.

Meanwhile, also a full-fledged eight-band **k·p** implementation for wurtzite and zinc blende QDs is (freely) available within the *nextnano*[3] project. Calculations for group-III-nitride QDs using the atomistic tight-binding model [72] have been presented for InGaN/GaN-QDs [59], GaN/AlN-QDs [56], and recently for pure InN/GaN-QDs [6, 7]. An in-depth comparative discussion is beyond the scope of the work.

2.2.4 Impact of Strain on Bulk Band Structure

In the context of heterostructures it is important to use a method for electronic structure calculations that fits the following requirements:

- Sensitivity for inhomogeneous strain. This covers band shifts due to hydrostatic strain, valence-band splitting arising from biaxial strain, and a strain-dependent conduction-band curvature or in other words: a strain-dependent effective electron mass.

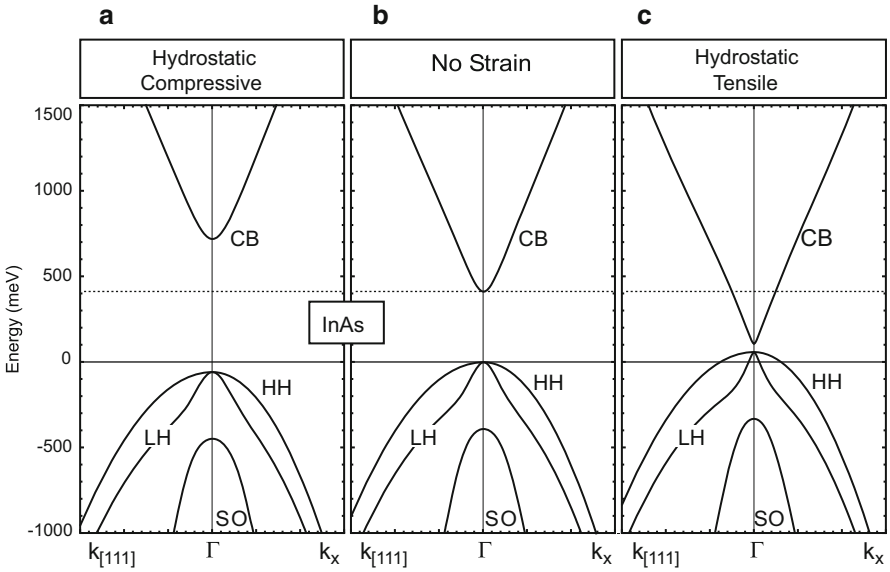


Fig. 2.6 Eight-band **k·p** bandstructure for InAs in presence of hydrostatic strain, $\epsilon_{xx} = \epsilon_{yy} = \epsilon_{zz}$: (**a**) Compressive hydrostatic strain, $\epsilon_{\mathrm{H}} < 0$, and (**c**) tensile hydrostatic strain, $\epsilon_{\mathrm{H}} > 0$. Case (**b**) refers to the absence of any strain

- Non-parabolicity effects. The Fourier-transform of a localized envelope functions spans a certain region in the vicinity of the Γ-point in the Brillouin-zone. Hence, to avoid overconfinement effects, the non-parabolic nature of the conduction band needs to be accounted for.

In the following section we will see, how eight-band **k·p** theory meets the outlined demands.

2.2.4.1 Zinc Blende Structure

Figure 2.6b shows the eight-band **k·p** bandstructure of bulk-InAs in absence of strain. Heavy hole (HH) and light hole (LH) bands are degenerate at the Γ-point, the split-off (SO) well separated due to spin-orbit interaction. The conduction band is *not parabolic* due to the conduction-band valence-band coupling mediated by the **k·p** term. This effectively decreases the electron effective mass of a localized electron state.

The impact of hydrostatic strain ($\epsilon_{\mathrm{H}} = \epsilon_{xx} + \epsilon_{yy} + \epsilon_{zz}$) is shown in Fig. 2.6a, c. Compressive strain shifts the conduction band up, thus, increasing the band gap. Tensile strain on the other hand decreases the band gap.

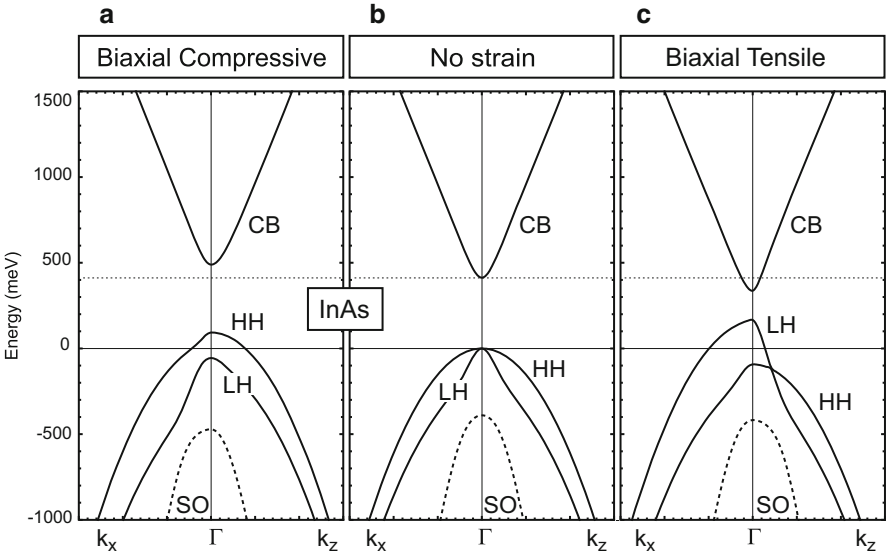


Fig. 2.7 Eight-band **k·p** bandstructure for InAs in presence of (**a**) biaxial compressive strain, with $\epsilon_{xx} = \epsilon_{yy} < 0$ and, $\epsilon_{zz} > 0$, (**c**) biaxial tensile strain, with $\epsilon_{xx} = \epsilon_{yy} > 0$ and, $\epsilon_{zz} < 0$. Case (**b**) refers to the absence of any strain. Material parameters are taken from [74]

If the main diagonal strain components are not the same, biaxial strain arises, as known from the classic case of strained quantum wells. These cases are considered in Fig. 2.7: Biaxial compressive strain (Fig. 2.7a) refers to $\epsilon_{xx} = \epsilon_{yy} < 0$ and, as a consequence of the Poisson effect, $\epsilon_{zz} > 0$, leads primarily to a splitting of the valence bands into a top lying HH- and a lower energy LH band. Note that the curvature of the HH band is 'heavier' in k_z-direction, which would be the confinement direction in case of a quantum well. Biaxial tensile strain (Fig. 2.7c) reverses the HH-LH ordering at the Brillouin-center.

Strain Dependent Electron Effective Mass

Cusack and coworker [21] were the first to point at the importance of the strain dependence of the electron effective mass in the context of strained quantum dots. The consequences for compressively strained QDs are smaller electron confinement energies as well as a decreased electron sub-level spacing.

This effect is accounted for in eight-band **k·p** theory: The hydrostatic strain strongly affects the band gap energy. As the strength of the **k·p** matrix elements critically depends on the band gap, the conduction band curvature and, hence, the electron effective mass is strongly modified by strain. This is depicted in Fig. 2.8. For tensile strain (Fig. 2.8a) the CB-curvature near the Γ-point increases, which

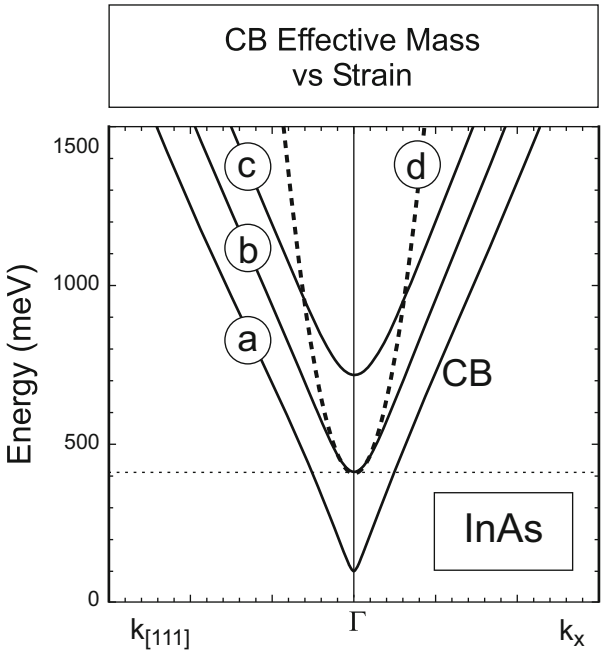


Fig. 2.8 Curvature of the conduction band as function of hydrostatic strain: (*a*) tensile strain, (*b*) no strain, (*c*) compressive strain. In case (*d*) no CB-VB coupling is present

translates into a lighter electron effective mass. For compressive strain (Fig. 2.8c), which corresponds to the case of InAs/GaAs quantum dots, on the other hand, the electron becomes heavier than in absence of strain. The impact of **k·p** coupling can be seen by comparing the curves Fig. 2.8(b and c), where in case (d) the coupling is switched off.

2.2.4.2 Wurtzite Structure

There are two major differences between the zb-InAs/GaAs system and the wz-GaN/AlN.

The first one is related to the much smaller band gap of InAs compared to GaN, which leads to a stronger CB-VB coupling for InAs. The small effect of CB-VB coupling in the GaN system is visible in Fig. 2.9a when comparing case (ii), which includes CB-VB coupling, and case (iv), where the coupling is switched off: the coupling induces no significant change of the CB-curvature near the Γ-point.

The second difference refers to the crystal-field splitting as a consequence of the spontaneous polarization present in the wz-GaN system: even if no strain is present, HH-hole and LH-hole are not degenerate at the zone-center anymore (Fig. 2.9b).

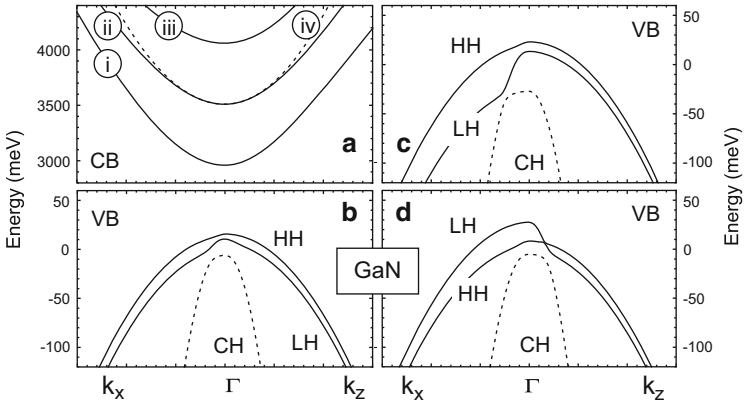


Fig. 2.9 Eight-band **k·p** bandstructure for GaN in presence of (**a**,*i*) tensile hydrostatic strain, (**a**,*ii*) no strain, (**a**,*iii*) compressive hydrostatic strain. (**a**,*iv*) No strain and no CB-VB coupling is present. (**b**) The conduction band is shown in absence of strain, whereas (**c**) refers to compressive biaxial strain, where the HH-LH splitting is enhanced, and (**d**) to tensile biaxial strain, finding a reversed HH-LH ordering, with the LH on top. Material parameters are taken from [57]

In the presence of biaxial strain the splitting is enhanced (compressive) (Fig. 2.9c) or reversed (tensile) (Fig. 2.9d).

2.2.5 Energies of Interacting Particles

Due to the resemblance of quantum dot properties to those of atoms, Coulomb interaction is best modeled using methods first developed in quantum chemistry. One of these methods which accounts for the complete spectrum of Coulomb effects, such as direct Coulomb interaction, exchange, and correlation, is the configuration interaction method [15, 16, 47, 64, 70, 79]. Its description, though, is clearly beyond the scope of this contribution. However, in the context of type-II excitons (see Sect. 2.3.2), at least selfconsistent direct Coulomb interaction is required to understand the binding nature of the spatially indirect exciton, which is shortly described in the following.

A self-consistent excitonic cycle consists of repeated calculations of one charge carrier affected not only by the pristine confinement potential but also by the Coulomb attraction of the other carrier, until convergence is reached.

The required Coulomb potential is calculated using the Poisson equation:

$$q_j |\Psi_0^j|^2 = \epsilon_0 \nabla \cdot (\epsilon_s \nabla V_0^j) \,, \tag{2.4}$$

taking into account image charge effects due to the spatial dependence of the dielectric constant, ϵ_s. The index j refers to either the electron or the hole groundstate, Ψ_0^j is the respective wavefunction and V_0^j the arising potential.

For GaN/AlN quantum dots, self-consistency effects are not as important as in the above treated case of type-II excitons. In Sect. 2.3.3.3 we report direct Coulomb energies, which are calculated using the Integral

$$J^{eh} = q_e \int \mathrm{d}\mathbf{r}\, |\Psi_0^e|^2\, V_0^h\,. \tag{2.5}$$

2.3 Discussion of Selected Topics

2.3.1 Zb(001) Versus zb(111) Substrate Orientation

The current interest in the (111)-substrate orientation results from the—compared to (001) substrates—higher surface symmetry [(001): twofold symmetry axis vs (111): threefold symmetry axis, see Fig. 2.10], which is expected to be carried over to the corresponding QD-symmetry.

A more fundamental interest in (111)-QDs stems from the fact that those QDs share properties inherited both from the zinc blende and the wurtzite lattice, since the [111] growth direction possesses many similarities to the c-axis of the wurtzite lattice. We refer to the work of Schulz and coworker [66] who published a comprehensive comparison of elastomechanic and piezoelectric properties between the zb-(001), zb-(111), and the wz-c-plane based nanostructures.

2.3.1.1 Orientation of the Piezoelectric Field

The major difference between QDs grown on either (001)- or (111)-substrate is the orientation of the piezoelectric field. As it decisively impacts the symmetry properties of the ensuing electronic states a closer inspection is carried out here.

The piezoelectric field and its orientation for (111) grown QDs is of large interest (i) first in view of its impact on the *lateral* symmetry of the confinement potential, and (ii) second with respect to the field distribution in *vertical* direction, the corresponding electron-hole alignment and the related few-particle binding energies.

(i) To compare the impact of the substrate orientation on the piezoelectric potential (Fig. 2.11) lens-shaped QDs are chosen as model system: for the (111) grown QDs, the potential shows C_{3v}-symmetry and a strong gradient along the growth direction, in contrast to the (001) grown counterpart, with only C_{2v} in-plane symmetry and no significant potential drop along the [001] axis. The field distribution of the (111) grown QD is similar to the one of *c*-plane wurtzite-type GaN/AlN or InN/GaN QDs [5, 53, 81]. The magnitude of the potential drop is much smaller as compared to nitride QDs, where the field additionally depends on pyroelectric effects, which do not occur in zinc blende crystals.

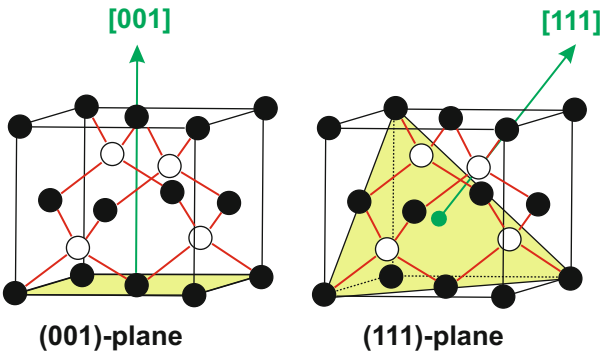


Fig. 2.10 Difference between the (001)- and the (111)-plane in the zinc blende crystal

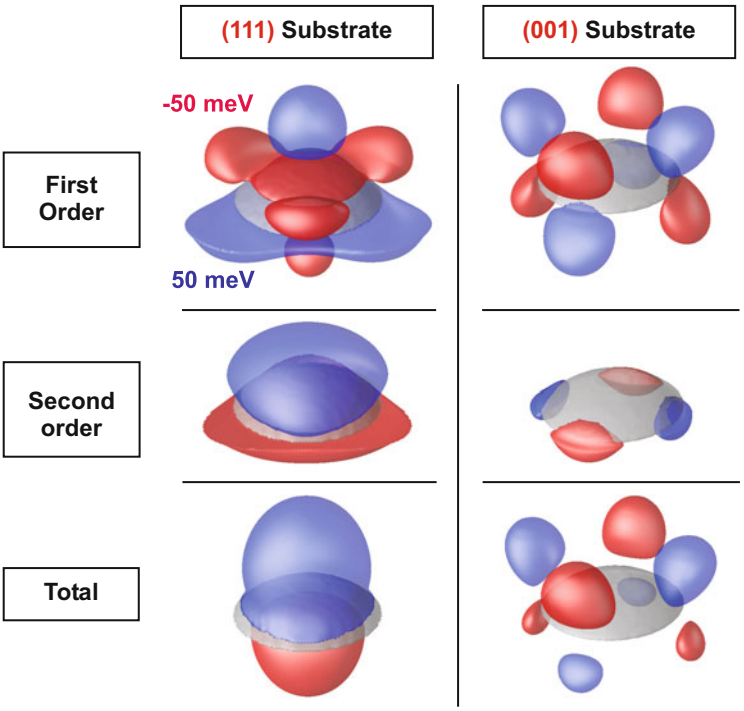


Fig. 2.11 Comparison of the piezoelectric fields (first and second order) for QDs grown on (111)B substrate to those grown on (001). Isosurfaces are shown for values of 50 meV (*blue*) and −50 meV (*red*) respectively. (see [65])

(ii) The orientation of the piezoelectric field in growth direction results from a subtle interplay between first- and second-order piezoelectric effects. The importance of second-order effects were first discovered for (111)-In(Ga)As/GaAs quantum wells (QW) [10], where the first-order piezoelectric fields alone were found to be not sufficient to explain a number of Stark-shift measurements [19, 20, 33, 60]: For increasing strain the nonlinear piezoelectricity starts to become important. It provides a small but significant contribution to the energy states in quantum wells, because the maximum In concentration in In(Ga)As QWs can hardly exceed $\approx$20 % before onset of dislocation generation.

In QDs, much larger In concentrations (up to 100 %) can be achieved, which consequently leads to much larger strain inside the QDs. As a result, non-linear piezoelectric effects cannot be neglected in QDs. As can be seen in Fig. 2.11 for (001)-grown InAs QDs (Fig. 2.11 right), first- and second-order effects compensate each other inside the QD [12], whereas in (111)-grown InAs QDs (Fig. 2.11 left and center), the second-order contributions are clearly dominant and therefore determine the orientation of the piezoelectric field.

2.3.1.2 Single Particle States

The first visible differences between the two substrate orientations regarding the electron energy states are the missing electron *p*-state splitting and a vertical charge separation. The latter further depends on the balance of first- and second order piezoelectric terms for a given composition for (111)-QDs [65]. Energy shifts occur due to different strain field distributions for varying substrate orientation [50].

The excitonic fine-structure splitting in QDs, albeit a two-particle effect, results from a distortion of electron and/or hole groundstate below C_{3v}-symmetry [67]. Such a distortion can originate from QD-elongation, from inequivalent side facets or from piezoelectric fields [67]. The role of the latter is illustrated in Fig. 2.12. For a (001)-grown lens-shaped QD (Fig. 2.12b) the hole groundstate is elongated along [110], whereas in case of a (111)B-grown QD (Fig. 2.12a), electron and hole groundstate show no deformation in any lateral direction.

An important peculiarity of (111)-QDs is the vertical distribution of the piezoelectric field inside the QD and its influence on the vertical position of electron and hole states (Fig. 2.13). In the case of zero piezoelectric field the center of mass of the electron lies above that of the hole [Fig. 2.13(left)]. This is well known from (001)-pyramidal QDs and related to a subtle interplay of the QD shape and strain [69]. Taking into account first-order piezoelectricity enhances the dipole by pulling the hole center of mass down to the QD bottom [Fig. 2.13(middle)]. The electron state moves up less pronounced than the hole state down, because its smaller effective mass makes it more 'resistant' against small potential changes. Adding second-order piezoelectric effects, however, reverses the picture completely [Fig. 2.13(right)]. The electron state moves down and the hole state up. Consequently the direction of the dipole changes. In addition, the vertical extension of the hole state increases drastically.

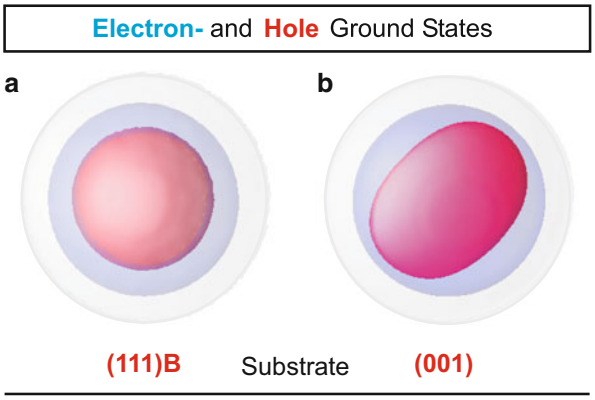


Fig. 2.12 Orientation of electron (*blue*) and hole (*red*) wave function for a lens-shaped QD on two substrate orientations. (see [65])

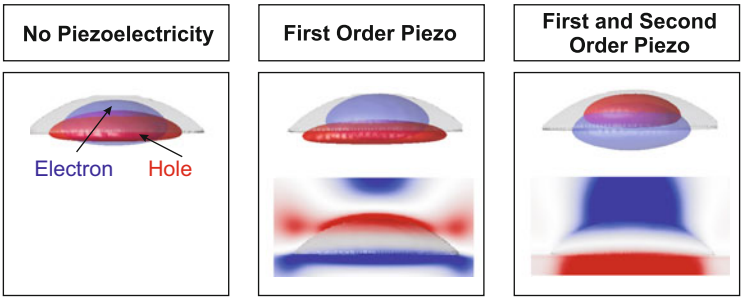


Fig. 2.13 (*upper row*) The position of electron and hole groundstate wavefunction (isosurfaces at 65 % probability density) dependent on the order of the piezoelectric field. (*lower row*) Vertical slices of the piezoelectric potential are shown. (see [65])

2.3.2 *Type-I Versus Type-II Confinement*

So far we have focused on systems where the band-alignment favors localization of both, electrons and holes. GaSb/GaAs quantum dots by contrast are very particular due to their type-II band alignment. The spatial separation of electrons and holes in type-II structures results in long exciton lifetimes [14, 28, 31, 75], enabling interesting optoelectronic applications [41]. The exclusive confinement of holes and their large localization energy makes GaSb/GaAs QDs particularly interesting for novel charge storage devices [26, 46], called nanoflash memories.

Figure 2.14 compares the local band-edge positions of dimensionally identical, but chemically different QDs: InAs/GaAs (type-I) versus GaSb/GaAs (type-II). The former represents a common anion-system with closely lying valence bands, whereas the latter is a common cation-system with the conduction bands being almost identical (see Fig. 2.14b). Hence, the band-gap difference between GaAs and

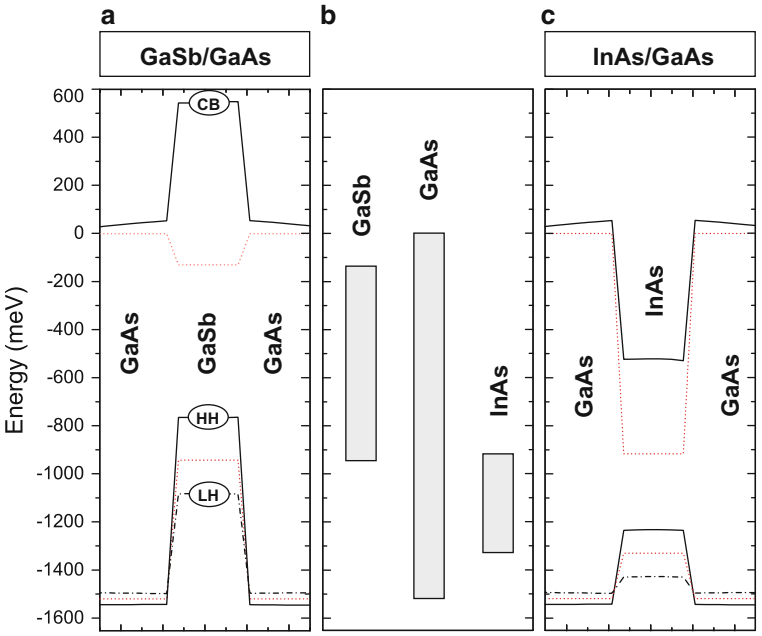


Fig. 2.14 Strained local band edge profiles are shown for quantum dots of same size but different chemical composition: (**a**) common cation GaSb/GaAs (type-II band alignment), and (**c**) common anion InAs/GaAs QD (type-I alignment). Part (**b**), as well as the *red dotted lines* in (**a**) and (**b**), show the band offsets in *absence of strain*

GaSb translates almost completely to a large valence-band offset. In the presence of strain, due to the lattice mismatch, the local band structure of a real heterostructure is strongly modified, leading to a huge increase (>700 meV for 100 % GaSb content [32]) of the GaSb conduction-band and a clear type-II band alignment (see Fig. 2.14a).

2.3.2.1 Strain Outside the Quantum Dot

If three-dimensional coherent nanostructures are put inside a crystalline matrix of different lattice constant, the resulting strain is not restricted to the inside of the nanostructure, but continues into the surrounding structure. As a result, the local bandstructure is not only altered inside but also outside the quantum dot as can be seen for both systems, InAs/GaAs and GaSb/GaAs, in Fig. 2.14. This finding stands in contrast to ideal quantum wells, where the strain is completely confined inside the layer [29], since the two-dimensional structure is allowed to relax into growth direction.

As the lattice constant of the dots investigated here is bigger than the one of the matrix, the environment is compressively strained, leading to an increase of the local

band edge in the close vicinity of the quantum dot of up to 60 meV. Hence, not only the quantum dot itself acts as repulsive force to electrons in the conduction band but even the close environment, which leads us to the problem of type-II exciton formation.

2.3.2.2 Type-II Exciton Formation

Electrons and holes localized inside the same type-I quantum dot inevitably form excitonic particles, no matter how big the Coulomb attraction is. If, however, the band edge acts repulsive towards one carrier type as in the case of type-II band alignment, exciton- and charged exciton formation is not trivial anymore, but requires additional Coulomb forces to at least form a spatially indirect composite particle.

In the case of GaSb/GaAs quantum dots, the Coulomb attraction needed is supplied by already captured holes, which create an attractive electrostatic potential around the quantum dot for binding electrons. The number of holes needed to at least compensate the strain induced increase of the band edge near the QD, depends on the composition and the vertical aspect ratio (height versus base length) [32]. In Fig. 2.15 the magnitude and the course of the additional potential caused by localized holes inside the QD and the resulting change of the band edge are shown for a flat binary GaSb/GaAs quantum dot. The flat band condition is reached for two localized hole carriers and for four carriers the potential is sufficient to bind an electron above the QD, as can be seen in Fig. 2.15c. A further increase of the number of positive charge carriers increases the Coulomb attraction and leads to a shrinkage of the electron wavefunction.

2.3.3 GaN/AlN Wurtzite Quantum Dots

Owing to a lattice mismatch of about 2.5 % GaN/AlN QDs grow strain induced in the Stranski–Krastanov growth mode. Their typical size and shape is known with great accuracy. Experimental reports on the structural properties of c-plane GaN QDs (e.g. [22, 34, 36, 62, 78]) seem to agree on the shape of the QDs, a truncated hexagonal pyramid with 30° side facets. The reported heights (h) scatter between 1.3 and 5 nm. The aspect ratios (h:d, where d denotes the lateral diameter) in most reports are in the range of 1:5–1:10.

The following section is based on the work of Winkelnkemper et al. [82]. A series of quantum dots all having the same vertical aspect ratio of 1:5 with a vertical size ranging between 0.8 and 3.6 nm is considered here. The thickness of the wetting layer is assumed to be $w = 0.2$ nm. This set of model QDs covers the major part of experimentally reported QD structures and yields excitonic transition energies and radiative lifetimes [82, Fig. 5] in good agreement with experimental values [18, 36].

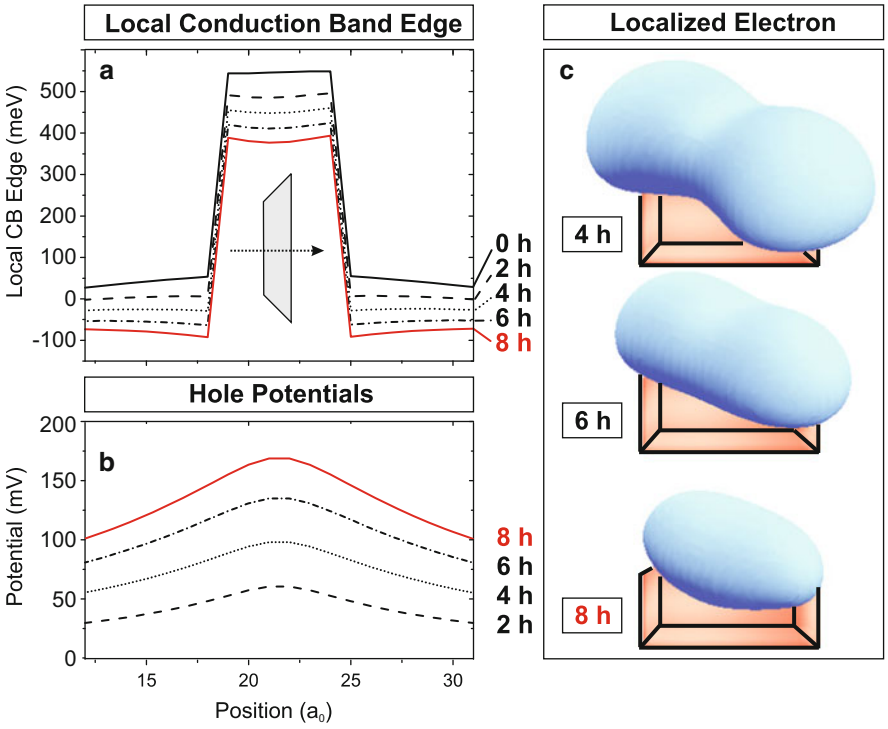


Fig. 2.15 (**a**) Course of the local conduction band edge as function of prestored hole carriers inside the GaSb/GaAs quantum dot. The QD has a height of about 3 nm and a base length of 22 nm. (**b**) Calculated electrostatic potential of the differently charged QD: 2 h refers to the occupied hole s-shell, 4 h to additionally occupation of the first p-shell and so on. (**c**) The first bound electron state is shown for different numbers of occupied hole levels

2.3.3.1 Role of Piezo- and Pyroelectric Built-in Fields

The huge built-in piezo- and pyroelectric fields within GaN/AlN QDs strongly affect the emission energies and radiative lifetimes of localized excitons within the QDs via the quantum-confined Stark effect (QCSE) [5, 17, 18, 78].

Figure 2.16 shows for two QD sizes the course of the piezoelectric and pyroelectric fields (a,b), the resulting local band edges together with the energetic positions of electron and hole groundstate (c,d). The built-in charges generate fields are as large as 8.0 MV/cm in the center of the QD leading to a strong spatial separation of electron and hole states (Fig. 2.17a,b). Depending on the size of the QDs the electron-hole overlap varies strongly, resulting in radiative lifetimes ranging from a few nanoseconds for small QDs up to as long as 100 μs for large QDs.

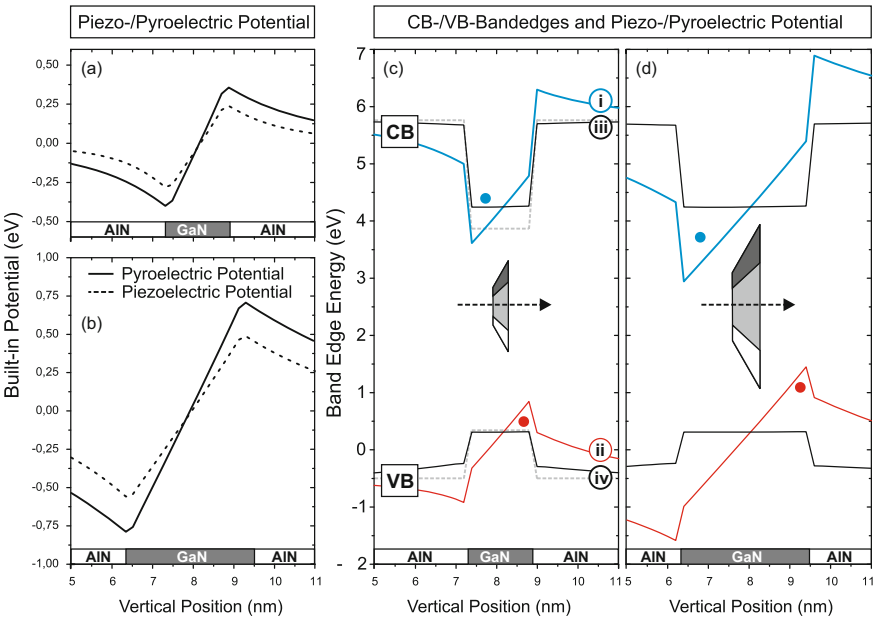


Fig. 2.16 (**a**,**b**) Piezo- and pyroelectric potentials ($e\phi$) shown for two QDs of different height: (**a**) h=1.6 nm, (**b**) h=3.2 nm. (**c**,**d**) The resulting local band edges are plotted for the cases (*i*,*ii*) strain and built-in fields, (*iii*,*iv*) strain but no built-in field, and *gray*, *dotted*: neither strain, nor piezo- and pyroelectric built-in field. The *blue* and *red dots* indicate the energetic and spatial position of the electron and hole groundstates

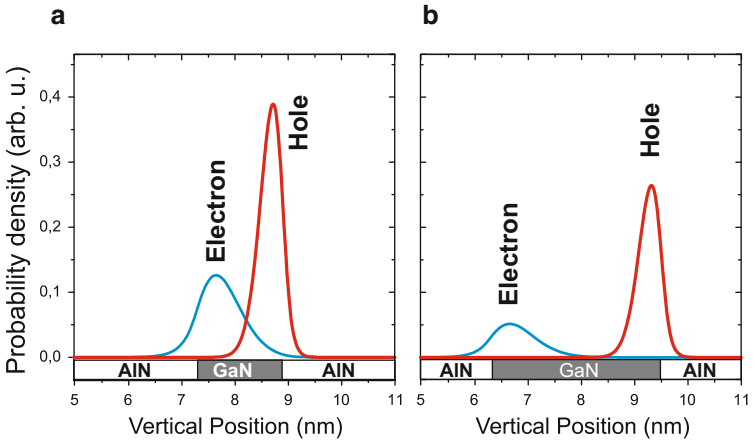


Fig. 2.17 (**a**,**b**) Probability density of electron and hole groundstate as function of QD-height

2.3.3.2 Single Particle Energy Levels

The bound hole states in GaN/AlN QDs are formed predominantly by the A- and B-band. C-band contributions are small because the biaxial strain within the QDs shifts this band to much lower energies [81, 83]. As a first approximation, for each band we expect to find a ground state with an s-shaped envelope function, which is only spin degenerate. The p-shell consists of two degenerate states and the d-shell of three. More p- and d-states with nodal planes along the growth direction are expected at much higher energies owing to the strong vertical confinement.

Due to the different parities of the bulk conduction and valence bands, the electron and hole states have a finite optical matrix element, if their envelope functions have the same parity, i.e. the allowed transition channels are s–s, p–p, s–d, etc. Each transition channel exists twice, once for the A-type holes and once for the B-type holes. Figure 2.18 shows the single-particle electron and hole energy levels of all QDs considered in this work including the five energetically lowest (highest) electron (hole) states. The electron states are all formed predominantly by the conduction band ($\approx$95 %). Therefore, s, p and d shells can be clearly distinguished. The hole spectra, however, are more complex, because hole states are not formed by either the A- or B-band, but by a mixture of both bands and even a small C-band contribution. Still each hole state can be characterized by the band that contributes most to it (see labels in Fig. 2.18). The A-band s-state ($h_0(A)$; $\approx$95 % A-band projection) and the B-band s-state ($h_1(B)$; $\approx$90 % B-band projection) are energetically well separated from the excited hole states. Both have an unambiguously s-shaped envelope function (not shown here). The splitting between both states ($\approx$9–10 meV) does not increase for smaller QDs, but is constant. It corresponds to the energy separation between A- and B-band in strained GaN. The higher excited hole states cannot be assigned to p- or d-like orbitals. Please note that, although they have been labeled according to the major band contributions, such contribution sometimes does not exceed 50 %.

2.3.3.3 Coulomb Interaction

GaN *bulk* excitons are well known for their stability even at room-temperature owing to their large exciton binding energy of 26 meV [58] which makes heterostructures based on GaN promising candidates for the realization of polariton lasers [45]. The question arises how the exciton binding energy is affected by a three-dimensional confinement as in the case of GaN/AlN QDs. Here, to display the major effect, we calculate the electron-hole binding energy, hence, leaving out selfconsistency and correlation effects. As can be seen from Fig. 2.19 height variations and the consequent change in electron-hole separation strongly affects the resulting Coulomb binding strength leading to values ranging from 160 meV for small QDs down to about 40 meV for large QDs being much larger than for GaN bulk.

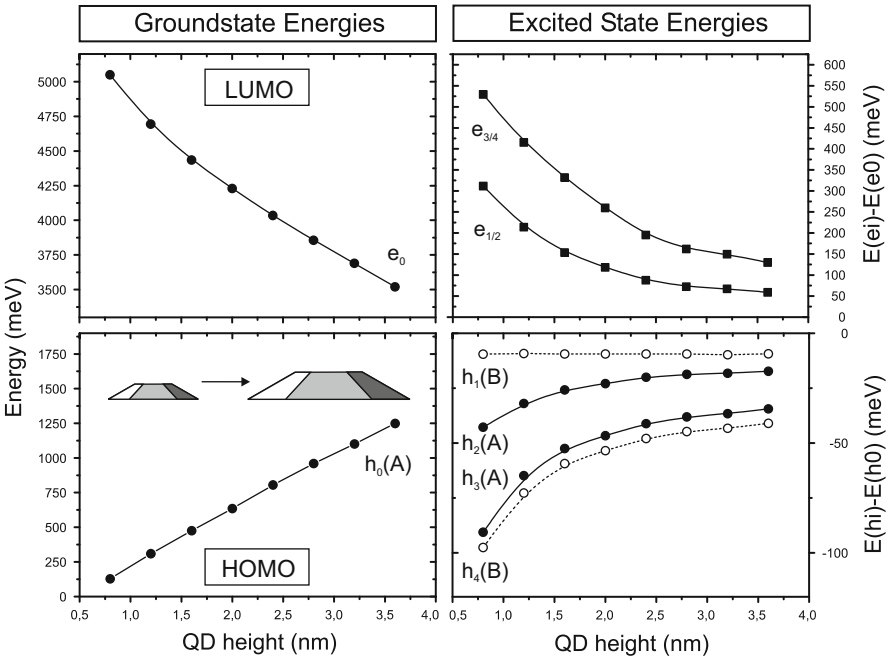


Fig. 2.18 Electron and hole single-particle energies as function of QD size. LUMO (HOMO) refers to the electron (hole) groundstate

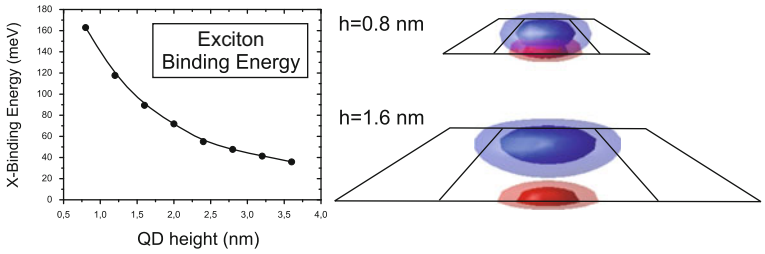


Fig. 2.19 (*left*) Electron-hole Coulomb binding energy as function of QD size. (*right*) Electron and hole probability densities are shown for two QDs of different size highlighting the different wavefunction extent and the rising electron-hole distance with increasing QD-height

2.4 Conclusion

In this chapter we gave an overview on the envelope function based eight-band **k**·**p** method for electronic structure calculations of QDs. Three types of heterostructures of particular current interest are used to exemplify the wide range of applicability of the method. It was shown that the theory is able to seamlessly account for different lattice systems, band-alignments and substrate orientations.

Acknowledgements Many colleagues contributed to the success of this work. However, three people stand out who contributed most over the past 20 years: We feel deeply indebted to Oliver Stier, Momme Winkelnkemper, and Marius Grundmann.

The work was funded by DFG in the frame of SFBs 296 and 787.

References

1. W.G. Cady *Piezoelectricity: an Introduction to the Theory and Applications of Electromechanical Phenomena in Crystals* (McGraw-Hill, 1946)
2. O.Ambacher, M. Eickhoff, A. Link, M. Hermann, M. Stutzmann, F. Bernardini, V. Fiorentini, Y. Smorchkova, J. Speck, U. Mishra, W. Schaff, V. Tilak, L.F. Eastman, Electronics and sensors based on pyroelectric AlGaN/GaN heterostructures. Phys. Stat. Sol. (c) **0**(6), 1878–1907 (2003)
3. O. Ambacher, J. Majewski, C. Miskys, A. Link, M. Hermann, M. Eickhoff, M. Stutzmann, F. Bernardini, V. Fiorentini, V. Tilak, Pyroelectric properties of Al (In) GaN/GaN hetero-and quantum well structures. J. Phys.: Condens. Matter **14**(13), 3399 (2002)
4. A.D. Andreev, E.P. O'Reilly, Theory of the electronic structure of GaN/AlN hexagonal quantum dots. Phys. Rev. B **62**(23), 851–870 (2000)
5. A.D. Andreev, E.P. O'Reilly, Optical transitions and radiative lifetime in GaN/AlN self-organized quantum dots. Appl. Phys. Lett. **79**(4), 521 (2001)
6. N. Baer, S. Schulz, P. Gartner, S. Schumacher, G. Czycholl, F. Jahnke, Influence of symmetry and Coulomb correlation effects on the optical properties of nitride quantum dots. Phys. Rev. B **76**(7) (2007)
7. N. Baer, S. Schulz, S. Schumacher, P. Gartner, G. Czycholl, F. Jahnke, Optical properties of self-organized wurtzite InN/GaN quantum dots: A combined atomistic tight-binding and full configuration interaction calculation. Appl. Phys. Lett. **87**(23), 114–231 (2005)
8. T. Bahder, Eight-band $\mathbf{k}\cdot\mathbf{p}$ model of strained zinc-blende crystals. Phys. Rev. B **41**(17), 992–1012 (1990)
9. F. Bernardini, V. Fiorentini, D. Vanderbilt, Spontaneous polarization and piezoelectric constants of III-V nitrides. Phys. Rev. B **56**(16), 10,024–10,027 (1997)
10. G. Bester, X. Wu, D. Vanderbilt, A. Zunger, Importance of second-order piezoelectric effects in zinc-blende semiconductors. Phys. Rev. Lett. **96**(18) (2006)
11. G. Bester, A. Zunger, Cylindrically shaped zinc-blende semiconductor quantum dots do not have cylindrical symmetry: Atomistic symmetry, atomic relaxation, and piezoelectric effects. Phys. Rev. B **71**(4) (2005)
12. G. Bester, A. Zunger, X. Wu, D. Vanderbilt, Effects of linear and nonlinear piezoelectricity on the electronic properties of InAs/GaAs quantum dots. Phys. Rev. B **74**(8) (2006)
13. D. Bimberg, M. Grundmann, N.N. Ledentsov, *Quantum Dot Heterostructures* (John Wiley & Sons, New York, 1999)
14. H. Born, L. Müller-Kirsch, R. Heitz, A. Hoffmann, D. Bimberg, Rapid research notes-radiative recombination in type II GaSb/GaAs quantum dots. Phys. Stat. Sol. (b) **228**(3), 4 (2001)
15. M. Braskén, M. Lindberg, D. Sundholm, J. Olsen, Full configuration interaction calculations of electron-hole correlation effects in strain-induced quantum dots. Phys. Rev. B **61**(11), 7652–7655 (2000)
16. M. Braskén, M. Lindberg, D. Sundholm, J. Olsen, Full configuration interaction calculations of electron-hole correlation effects in strain-induced quantum dots. Phys. Stat. Sol. (b) **224**(3), 775–779 (2001)
17. T. Bretagnon, S. Kalliakos, P. Lefebvre, P. Valvin, B. Gil, N. Grandjean, A. Dussaigne, B. Damilano, J. Massies, Time dependence of the photoluminescence of GaN/AlN quantum dots under high photoexcitation. Phys. Rev. B **68**(20), 205–301 (2003)

18. T. Bretagnon, P. Lefebvre, P. Valvin, R. Bardoux, T. Guillet, T. Taliercio, B. Gil, N. Grandjean, F. Semond, B. Damilano, A. Dussaigne, J. Massies, Radiative lifetime of a single electron-hole pair in GaN/AlN quantum dots. Phys. Rev. B **73**(11), 113–304 (2006)
19. C.H. Chan, M.C. Chen, H.H. Lin, Y.F. Chen, G.J. Jan, Y.H. Chen, Characterization of piezoelectric (111)B InGaAs/GaAs p-i-n quantum well structures using photoreflectance spectroscopy. Appl. Phys. Lett. **72**(10), 1208 (1998)
20. S. Cho, J. Kim, A. Sanz-Hervás, A. Majerfeld, G. Patriarche, B.W. Kim, Characterization of piezoelectric and pyroelectric properties of MOVPE-grown strained (111)A InGaAs/GaAs QW structures by modulation spectroscopy. Phys. Stat. Sol. (a) **195**(1), 260–264 (2003)
21. M. Cusack, P. Briddon, Electronic structure of InAs/GaAs self-assembled quantum dots. Phys. Rev. B (1996)
22. B. Daudin, F. Widmann, G. Feuillet, Y. Samson, M. Arlery, J. Rouvière, Stranski-Krastanov growth mode during the molecular beam epitaxy of highly strained GaN. Phys. Rev. B **56**(12), R7069–R7072 (1997)
23. P. Enders, A. Bärwolff, M. Woerner, D. Suisky, **k·p** theory of energy bands, wave functions, and optical selection rules in strained tetrahedral semiconductors. Phys. Rev. B **51**(23), 16,695–16,704 (1995)
24. V.A. Fonoberov, A.A. Balandin, Excitonic properties of strained wurtzite and zinc-blende $GaN/Al_xGa_{1-x}N$ quantum dots. J. Appl. Phys. **94**(11), 7178 (2003)
25. H. Fu, L.W. Wang, A. Zunger, Applicability of the **k·p** method to the electronic structure of quantum dots. Phys. Rev. B **57**(16), 9971–9987 (1998)
26. M. Geller, A. Marent, T. Nowozin, D. Bimberg, N. Akçay, N. Öncan, A write time of 6 ns for quantum dot–based memory structures. Appl. Phys. Lett. **92**(9), 092–108 (2008)
27. D. Gershoni, C.H. Henry, G.A. Baraff, Calculating the optical properties of multidimensional heterostructures: Application to the modeling of quaternary quantum well lasers. IEEE J. Quantum Electron. **29**(9), 2433–2450 (1993)
28. K. Gradkowski, T.J. Ochalski, D.P. Williams, J. Tatebayashi, A. Khoshakhlagh, G. Balakrishnan, E.P. O'Reilly, G. Huyet, L.R. Dawson, D.L. Huffaker, Optical transition pathways in type-II Ga(As)Sb quantum dots. J. Luminescence **129**(5), 456–460 (2009)
29. M. Grundmann, O. Stier, D. Bimberg, InAs/GaAs pyramidal quantum dots: Strain distribution, optical phonons, and electronic structure. Phys. Rev. B **52**(16), 11,969 (1995)
30. T. Hammerschmidt, Role of strain relaxation during different stages of InAs quantum dot growth. In: Physics of Semiconductors: 27th International Conference on the Physics of Semiconductors - ICPS-27, pp. 601–602. AIP (2005)
31. F. Hatami, M. Grundmann, N. Ledentsov, F. Heinrichsdorff, R. Heitz, J. Böhrer, D. Bimberg, S. Ruvimov, P. Werner, V. Ustinov, P. Kop'ev, Z. Alferov, Carrier dynamics in type-II GaSb/GaAs quantum dots. Phys. Rev. B **57**(8), 4635–4641 (1998)
32. M. Hayne, J. Maes, S. Bersier, V.V. Moshchalkov, A. Schliwa, L. Müller-Kirsch, C. Kapteyn, R. Heitz, D. Bimberg, Electron localization by self-assembled GaSb/GaAs quantum dots. Appl. Phys. Lett. **82**(24), 4355–4357 (2003)
33. R. Hogg, T. Fisher, A. Willcox, D. Whittaker, M. Skolnick, D. Mowbray, J. David, A. Pabla, G. Rees, R. Grey, J. Woodhead, J. Sanchez-Rojas, G. Hill, M. Pate, P. Robson, Piezoelectric-field effects on transition energies, oscillator strengths, and level widths in (111)B-grown (In,Ga)As/GaAs multiple quantum wells. Phys. Rev. B **48**(11), 8491–8494 (1993)
34. K. Hoshino, S. Kako, Y. Arakawa, Formation and optical properties of stacked GaN self-assembled quantum dots grown by metalorganic chemical vapor deposition. Appl. Phys. Lett. **85**(7), 1262–1264 (2004)
35. H. Jiang, J. Singh, Strain distribution and electronic spectra of InAs/GaAs self-assembled dots: An eight-band study. Phys. Rev. B **56**(8), 4696–4701 (1997)
36. S. Kako, C. Santori, K. Hoshino, S. Götzinger, Y. Yamamoto, Y. Arakawa, A gallium nitride single-photon source operating at 200K. Nat Mater **5**(11), 887–892 (2006)
37. E. Kane, *Energy Band Theory* (Handbook on Semiconductors, 1982)

38. E.Kane, Phonon spectra of diamond and zinc-blende semiconductors. Phys. Rev. B **31**(12), 7865–7876 (1985)
39. J. Kim, L.W. Wang, A. Zunger, Comparison of the electronic structure of InAs/GaAs pyramidal quantum dots with different facet orientations. Phys. Rev. B **57**(16), R9408–R9411 (1998)
40. N. Kirstaedter, N.N. Ledentsov, M. Grundmann, D. Bimberg, V.M. Ustinov, S.S. Ruvimov, M.V. Maximov, P.S. Kop'ev, Z.I. Alferov, U. Richter, P. Werner, U. Gosele, J. Heydenreich, Low threshold, large To injection laser emission from (InGa)As quantum dots. Electron. Lett. **30**(17), 1416–1417 (1994)
41. R.B. Laghumavarapu, A. Moscho, A. Khoshakhlagh, M. El-Emawy, L.F. Lester, D.L. Huffaker, GaSb/GaAs type II quantum dot solar cells for enhanced infrared spectral response. Appl. Phys. Lett. **90**(17), 173,125–173,125–3 (2007)
42. O.L. Lazarenkova, P. von Allmen, F. Oyafuso, S. Lee, G. Klimeck, Effect of anharmonicity of the strain energy on band offsets in semiconductor nanostructures. Appl. Phys. Lett. **85**(18), 4193 (2004)
43. S. Lee, F. Oyafuso, P. von Allmen, G. Klimeck, Boundary conditions for the electronic structure of finite-extent embedded semiconductor nanostructures. Phys. Rev. B **69**(4), 045,316 (2004)
44. J.A. Majewski, S. Birner, A. Trellakis, M. Sabathil, P. Vogl, Advances in the theory of electronic structure of semiconductors. Phys. Stat. Sol. (c) **1**(8), 2003–2027 (2004)
45. G. Malpuech, A. Di Carlo, A. Kavokin, J.J. Baumberg, M. Zamfirescu, P. Lugli, Room-temperature polariton lasers based on GaN microcavities. Appl. Phys. Lett. **81**(3), 412–414 (2002)
46. A. Marent, M. Geller, A. Schliwa, D. Feise, K. Pötschke, D. Bimberg, N. Akçay, N. Öncan, 10^6 years extrapolated hole storage time in GaSb/AlAs quantum dots. Appl. Phys. Lett. **91**(24), 242,109 (2007)
47. R. McWeeny, B.T. Pickup, Quantum theory of molecular electronic structure. Rep. Prog. Phys. **43**(9), 1065–1144 (2000)
48. P. Michler, A quantum dot single-photon turnstile device. Science **290**(5500), 2282–2285 (2000)
49. M. Migliorato, A. Cullis, M. Fearn, J. Jefferson, Atomistic simulation of strain relaxation in $In_xGa_{1-x}As$/GaAs quantum dots with nonuniform composition. Phys. Rev. B **65**(11), 115,316 (2002)
50. V. Mlinar, F.M. Peeters, Optical properties of (In,Ga)As capped InAs quantum dots grown on [11k] substrates. Appl. Phys. Lett. **91**(2), 021,910 (2007)
51. V. Mlinar, A. Zunger, Spectral barcoding of quantum dots: deciphering structural motifs from the excitonic spectra. Phys. Rev. B **80**(3), 035,328 (2009)
52. F.H. Pollak, Effects of homogeneous strain on the electronic and vibrational levels in semiconductors. In: T.P. Pearsall (ed.) Semiconductors and Semimetals, p. 17. Strained-Layer Superlattices: Physics (1990)
53. M. Povolotskyi, A. Di Carlo, S. Birner, Electronic and optical properties of [N11] grown nanostructures. Phys. Stat. Sol. (c) **1**(6), 1511–1521 (2004)
54. C. Pryor, Eight-band calculations of strained InAs/GaAs quantum dots compared with one-, four-, and six-band approximations. Phys. Rev. B **57**(12), 7190–7195 (1998)
55. C. Pryor, J. Kim, L.W. Wang, A.J. Williamson, A. Zunger, Comparison of two methods for describing the strain profiles in quantum dots. J. Appl. Phys. **83**(5), 2548 (1998)
56. V. Ranjan, G. Allan, C. Priester, C. Delerue, Self-consistent calculations of the optical properties of GaN quantum dots. Phys. Rev. B **68**(11) (2003)
57. P. Rinke, M. Scheffler, A. Qteish, M. Winkelnkemper, D. Bimberg, J. Neugebauer, Band gap and band parameters of InN and GaN from quasiparticle energy calculations based on exact-exchange density-functional theory. Appl. Phys. Lett. **89**(16), 161,919 (2006)
58. A. Rodina, M. Dietrich, A. Göldner, L. Eckey, A. Hoffmann, A. Efros, M. Rosen, B. Meyer, Free excitons in wurtzite GaN. Phys. Rev. B **64**(11), 115,204 (2001)
59. T. Saito, Y. Arakawa, Electronic structure of piezoelectric $In_{0.2}Ga_{0.8}N$ quantum dots in GaN calculated using a tight-binding method. Physica E **15**(3), 169–181 (2002)

60. J.L. Sanchez Rojas, A. Sacedon, F. Gonzalez Sanz, E. Calleja, E. Munoz, Dependence on the In concentration of the piezoelectric field in (111)B InGaAs/GaAs strained heterostructures. Appl. Phys. Lett. **65**(16), 2042–2044 (1994)
61. R. Santoprete, B. Koiller, R. Capaz, P. Kratzer, Q. Liu, M. Scheffler, Tight-binding study of the influence of the strain on the electronic properties of InAs/GaAs quantum dots. Phys. Rev. B **68**(23) (2003)
62. C. Santori, S. Götzinger, Y. Yamamoto, S. Kako, K. Hoshino, Y.Arakawa, Photon correlation studies of single GaN quantum dots. Appl. Phys. Lett. **87**(5), 051,916–051,916–3 (2005)
63. A. Schliwa, M. Winkelnkemper, D. Bimberg, Impact of size, shape, and composition on piezoelectric effects and electronic properties of In(Ga)As/GaAs quantum dots. Phys. Rev. B **76**(20) (2007)
64. A. Schliwa, M. Winkelnkemper, D. Bimberg, Few-particle energies versus geometry and composition of $In_xGa_{1-x}As$/GaAs self-organized quantum dots. Phys. Rev. B **79**(7) (2009)
65. A. Schliwa, M. Winkelnkemper, A. Lochmann, E. Stock, D. Bimberg, In(Ga)As/GaAs quantum dots grown on a (111) surface as ideal sources of entangled photon pairs. Phys. Rev. B **80**(16) (2009)
66. S. Schulz, M.A. Caro, E.P. O'Reilly, O. Marquardt, Symmetry-adapted calculations of strain and polarization fields in (111)-oriented zinc-blende quantum dots. Phys. Rev. B **84**(12), 125,312 (2011)
67. R. Seguin, A. Schliwa, S. Rodt, K. Pötschke, U. Pohl, D. Bimberg, Size-dependent fine-structure splitting in self-organized InAs/GaAs quantum dots. Phys. Rev. Lett. **95**(25) (2005)
68. V. Shchukin, D. Bimberg, Spontaneous ordering of nanostructures on crystal surfaces. Rev. Mod. Phys. **71**(4), 1125–1171 (1999)
69. W. Sheng, J.P. Leburton, Electron-hole alignment in InAs/GaAs self-assembled quantum dots: effects of chemical composition and dot shape. Phys. Rev. B **63**(16) (2001)
70. E. Siebert, T. Warming, A. Schliwa, E. Stock, M. Winkelnkemper, S. Rodt, D. Bimberg, Spectroscopic access to single-hole energies in InAs/GaAs quantum dots. Phys. Rev. B **79**(20) (2009)
71. R. Singh, G. Bester, Nanowire quantum dots as an ideal source of entangled photon pairs. Phys. Rev. Lett. **103**(6) (2009)
72. J.C. Slater, G.F. Koster, Simplified LCAO method for the periodic potential problem. Phys. Rev. **94**(6), 1498–1524 (1954)
73. O. Stier, D. Bimberg, Modeling of strained quantum wires using eight-band **k·p** theory. Phys. Rev. B **55**(12), 7726–7732 (1997)
74. O. Stier, M. Grundmann, D. Bimberg, Electronic and optical properties of strained quantum dots modeled by 8-band **k·p** theory. Phys. Rev. B **59**(8), 5688 (1999)
75. C.K. Sun, G. Wang, J.E. Bowers, B. Brar, H.R. Blank, H. Kroemer, M.H. Pilkuhn, Optical investigations of the dynamic behavior of GaSb/GaAs quantum dots. Appl. Phys. Lett. **68**(11), 1543–1545 (1996)
76. J. Tersoff, Modeling solid-state chemistry: interatomic potentials for multicomponent systems. Phys. Rev. B **39**(8), 5566–5568 (1989)
77. I. Vurgaftman, J.R. Meyer, Band parameters for nitrogen-containing semiconductors. J. Appl. Phys. **94**(6), 3675 (2003)
78. F. Widmann, J. Simon, B. Daudin, G. Feuillet, J. Rouvière, N. Pelekanos, G. Fishman, Blue-light emission from GaN self-assembled quantum dots due to giant piezoelectric effect. Phys. Rev. B **58**(24), R15,989–R15,992 (1998)
79. A.J. Williamson, A. Franceschetti, A. Zunger, Multi-excitons in self-assembled InAs/GaAs quantum dots: a pseudopotential, many-body approach. Europhys. Lett. **53**(1), 59–65 (2007)
80. A.J. Williamson, L. Wang, A. Zunger, Theoretical interpretation of the experimental electronic structure of lens-shaped self-assembled InAs/GaAs quantum dots. Phys. Rev. B **62**(19), 12,963–12,977 (2000)
81. M. Winkelnkemper, A. Schliwa, D. Bimberg, Interrelation of structural and electronic properties in $In_xGa_{1-x}N$/GaN quantum dots using an eight-band **k·p** model. Phys. Rev. B **74**(15) (2006)

82. M. Winkelnkemper, R. Seguin, S. Rodt, A. Hoffmann, D. Bimberg, GaN/AlN quantum dots for single qubit emitters. J. Phys.: Condens. Matter **20**(45), 454,211 (2008)
83. M. Winkelnkemper, R. Seguin, S. Rodt, A. Schliwa, L. Reissmann, A. Strittmatter, A. Hoffmann, D. Bimberg, Polarized emission lines from A- and B-type excitonic complexes in single InGaN/GaN quantum dots. J. Appl. Phys. **101**(11), – (2007)

Chapter 3
Symmetries in Multiband Hamiltonians for Semiconductor Quantum Dots

Stanko Tomić and Nenad Vukmirović

Abstract Our current understanding of the symmetries of multiband envelope function Hamiltonians for semiconductor quantum dots and their signatures in the energy level structure and wave function shapes is reviewed. We show how symmetry can be used to block-diagonalize the Hamiltonian matrix and consequently strongly reduce the computational effort. A detailed analysis of symmetries of several different model Hamiltonians reveals that the true symmetry of square-based pyramidal quantum dots is captured if either the interface effects are taken into account or additional higher energy bands are included in the multiband Hamiltonian. This indicates that multiband envelope function methods are fully capable of capturing the true atomistic symmetry of quantum dots in contrast to some widespread beliefs. In addition, we show that translational symmetry can be artificially introduced by the numerical method used, such as the plane wave method. Plane wave method introduces artificial quantum dot replica whose charges interact with charges in the real quantum dot and create an additional strain field in the real dot. This issue can be circumvented by the introduction of proper corrections in the procedure for calculation of Coulomb integrals and strain.

S. Tomić (✉)
Joule Physics Laboratory, School of Computing, Science and Engineering,
University of Salford, Manchester M5 4WT, UK
e-mail: s.tomic@salford.ac.uk

N. Vukmirović
Scientific Computing Laboratory, Institute of Physics Belgrade, University of Belgrade,
Pregrevica 118, 11080 Belgrade, Serbia
e-mail: nenad.vukmirovic@ipb.ac.rs

M. Ehrhardt and T. Koprucki (eds.), *Multi-Band Effective Mass Approximations*,
Lecture Notes in Computational Science and Engineering 94,
DOI 10.1007/978-3-319-01427-2_3,

3.1 Introduction

Quantum dots are nanostructures which provide confinement of carriers in all three spatial directions. On the fundamental side, they enable studies of interactions between electrons and photons at the single or few particle level [1,45,61,71,74,99]. Practical applications of semiconductor quantum dots include lasers [31], optical amplifiers [8], single photon sources [45,71], photodetectors [42,50,51], fluorescent biological labels [2] and solar cells [6, 11, 28, 48, 49, 67, 72].

For these reasons, there is a tremendous need to develop both accurate and computationally efficient methods for the description of electronic states in quantum dots. In other nanostructures, such as quantum wells or wires, one can exploit the translational symmetry of the structure and consequently strongly reduce the computational cost. Single quantum dots, where no translational symmetry of the structure is present, are therefore most challenging structures for numerical studies. Nevertheless, in most cases quantum dots exhibit certain symmetry which can be exploited to reduce the computational cost. The main goal of this article is to provide understanding when and how symmetry can be exploited in numerical calculations of electronic states in quantum dots.

On the other hand, we also show that the numerical method used can introduce an artificial symmetry. This is the case for the plane wave method that assumes periodic boundary conditions which introduce an artificial translational symmetry. While such method is very useful of one wants to study the quantum dot supercrystals or quantum dot arrays [3, 34, 39, 80, 82], it needs to be modified for its applications to single quantum dot structures. In such cases, one would naturally like to remove the effects of such artefacts from the results. We show how this can be done in Sect. 3.8.

The multiband **k·p** Hamiltonians [12, 14, 15, 23, 27, 33, 44, 53, 54, 59, 60, 62–64, 68, 86, 98] are capable of reproducing the bulk bandstructure more accurately than the standard 8-band Hamiltonian. Some of these, that include a large number of bands ($\gtrsim$ 15 or 30 after incorporation of the spin degeneracies), are even capable of reproducing the bulk bandstructure throughout the whole Brillouin zone. Unfortunately, these Hamiltonians have been rarely applied to nanostructures and have not been applied to QDs at all. The effect of interface band mixing [17, 21, 32, 65, 96] has also so far been analyzed only for a single interface or a quantum well structure. The goal of this work is to explore the effects of higher bands and interfaces on the electronic structure of QDs.

In this work, we focus on self-assembled quantum dots that can be produced using epitaxial techniques [9, 52]. These typically have lateral dimensions of the order of 15–30 nm and the height of the order 3–7 nm. While ab initio calculations based on density functional theory have been performed for the clusters and nanocrystals of the size up to $\sim$3 nm [18, 25, 58, 85, 97], much larger self-assembled quantum dots are still out of the range of present day computational resources. These methods also suffer from unreliability in predicting the energy gaps. Several methods that retain the atomistic details of the system, but do not involve a self-consistent calculation,

have therefore been developed and applied to self-assembled quantum dots, such as the empirical pseudopotential method [30,93–95], the tight-binding method [35,69, 70,73], and the charge patching method [36–38,90–92].

These methods directly take into account the atomistic details of the system. This leads to their high accuracy and reliability, which is however also accompanied by a significant computational cost. In the envelope function methods (better known as the **k**·**p** method), central quantities are the slowly varying envelope functions which modulate the rapidly varying atomistic wave function. The fact that the envelope functions are slowly varying implies that less memory is needed for their representation and consequently less time is needed for their computation. This makes the method very computationally efficient and therefore attractive for the applications. We will show in this chapter that this doesn't necessary causes the lost in accuracy. We will describe the procedures for improving the envelope methods in order to reach the same level of sophistication in terms of predicting the correct symmetries of states in quantum dots as in more computationally demanding atomistic methods.

3.2 Multiband Envelope Function Method

In this section, we show how one can derive the equations that envelope functions satisfy. To simplify the derivation, we do not consider the effects of strain, piezoelectricity and spin-orbit interaction. These effects have been treated on various occasions and are well documented in the literature. Instead, we focus on the effect that is less well known—the effect of interfaces.

The single-particle Hamiltonian of an electron in a semiconductor is given as

$$H = \frac{\mathbf{p}^2}{2m_0} + V_0(\mathbf{r}), \tag{3.1}$$

where $\mathbf{p}$ is the electron momentum operator, m_0 the free electron mass and $V_0(\mathbf{r})$ the crystal potential experienced by an electron. One can think of V_0 as the self-consistent potential obtained from density functional theory or as the empirical pseudopotential. The envelope representation of the electronic wave function is given as

$$\Psi(\mathbf{r}) = \sum_i \psi_i(\mathbf{r}) u_i(\mathbf{r}), \tag{3.2}$$

where the functions $u_i(\mathbf{r})$ are orthonormal and have the periodicity of the Bravais lattice, while $\psi_i(\mathbf{r})$ are slowly varying envelope functions. The most widely used choice of the functions u_i are bulk Bloch functions at the Γ point. However, there is some ambiguity in the previous statement. If we consider a quantum dot of material A embedded in material B, are u_i the Bloch functions of material A or material B?

In what follows, we will consider that u_i are Bloch functions of some effective material C which is in some sense the average of materials A and B. For example, if materials A and B are described by pseudopotentials V_A and V_B, we will assume that the pseudopotential of the average material C is $V_C = (V_A + V_B)/2$.

After the replacement of Eq. (3.2) into the eigenvalue problem of the Hamiltonian given by Eq. (3.1) and making an approximation that eliminates the non-local terms that appear in the derivation, one arrives at the system of equations for the envelope functions [13, 22]

$$-\frac{\hbar^2}{2m_0}\nabla^2\psi_m(\mathbf{r}) + \sum_n \frac{-i\hbar}{m_0}\mathbf{p}_{mn}\cdot\nabla\psi_n(\mathbf{r}) + \sum_n H_{mn}(\mathbf{r})\psi_n(\mathbf{r}) = E\psi_m(\mathbf{r}). \quad (3.3)$$

The terms in Eq. (3.3) are defined as

$$\mathbf{p}_{mn} = \frac{1}{\Omega}\int u_m(\mathbf{r}')^*\mathbf{p}u_n(\mathbf{r}')\mathrm{d}^3\mathbf{r}', \quad (3.4)$$

where the integration goes over the volume of the crystal unit cell Ω, and

$$H_{mn} = \frac{1}{\Omega}\int u_m(\mathbf{r}')^* H u_n(\mathbf{r}')\mathrm{d}^3\mathbf{r}'. \quad (3.5)$$

Eq. (3.3) can be recast into the form

$$\sum_n h_{mn}(\mathbf{R})\psi_n(\mathbf{R}) = E\psi_m(\mathbf{R}), \quad (3.6)$$

where

$$h_{mn}(\mathbf{R}) = \frac{\hbar^2k^2}{2m_0}\delta_{mn} + \frac{\hbar}{m_0}\mathbf{k}\cdot\mathbf{p}_{mn} + [u_m|H|u_n]_{\mathbf{R}} \quad (3.7)$$

and $\mathbf{k} = -i\nabla$, while the square brackets denote the averaging over a unit cell centered at $\mathbf{R}$. The term $[u_m|V|u_n]_{\mathbf{R}}$ in the Hamiltonian [Eq. (3.7)] is a constant of a given material when $\mathbf{R}$ is far away from the interface—when the averaging does not include the interface region. Since the second term in Eq. (3.7) is crucial in the Hamiltonian matrix, the envelope function method is most frequently called the $\mathbf{k}\cdot\mathbf{p}$ method. The h_{mn} operator in Eq. (3.7) is referred to as the envelope Hamiltonian or the $\mathbf{k}\cdot\mathbf{p}$ Hamiltonian.

In practical calculations, one has to restrict to a finite number of bands in the representation in Eq. (3.2). Historically, the $\mathbf{k}\cdot\mathbf{p}$ method was first applied to valence band (6-band Hamiltonian) [40, 41] and later on the conduction band was added (8-band Hamiltonian) [55]. Recently, we have applied the 14-band and 16-band Hamiltonians (that also include the effects of strain, spin-orbit interaction, crystal field splitting and remote bands) to quantum dot structures [83]. However, these

Hamiltonians consider the last term in Eq. (3.7) as a constant of a given material and do not take into account its behavior at the interface of two materials. We will show that this term is important if one wants to understand the symmetry of the envelope function Hamiltonian and therefore devote the next section to the analysis of this term.

3.3 The Effect of Interfaces

We define the "length" of the interface L_{if} as the length of the region of space that consists of all $\mathbf{R}$-vectors such that the average $[u_m|V|u_n]_{\mathbf{R}}$ encompasses the interface region. For example, in the case of the [001] interface in zincblende materials $L_{\text{if}} = a/2$, where a is the bulk lattice constant.

Since the interface region is small and the envelope functions are slowly varying, the details of the variations of the $[u_m|V|u_n]_{\mathbf{R}}$ are not of primary importance—it is only the integral of this term over the interface region that determines its role in the envelope Hamiltonian. In the flat interface model, the pseudopotentials are modeled to be equal to those of material A at one side of the interface and moving sharply to those of material B at the other side of an interface. We then obtain:

$$\begin{aligned}\int_{-L_{\text{if}}/2}^{+L_{\text{if}}/2}[u_m|H|u_n]_{z_0}\mathrm{d}z_0 = &\int_{-L_{\text{if}}/2}^{+L_{\text{if}}/2}[u_m|\frac{p^2}{2m_0}+V|u_n]_{z_0}\mathrm{d}z_0\\ &-[u_m|\frac{p^2}{2m_0}+\frac{V_{\text{A}}+V_{\text{B}}}{2}|u_n]L_{\text{if}}+\\ &+[u_m|\frac{p^2}{2m_0}+V_{\text{A}}|u_n]\frac{L_{\text{if}}}{2}+[u_m|\frac{p^2}{2m_0}+V_{\text{B}}|u_n]\frac{L_{\text{if}}}{2}.\end{aligned} \tag{3.8}$$

In Eq. (3.8), the last two terms on the right hand side represent the bulk contribution to the Hamiltonian, while the first two terms are the interface contribution. This implies that for each interface, the envelope function Hamiltonian contains an additional term of the form $\Omega_{mn}\delta(z)$ (assuming the plane of the interface is the $z = 0$ plane), with Ω_{mn} given by the expression:

$$\Omega_{mn} = \int_{-L_{\text{if}}/2}^{+L_{\text{if}}/2}[u_m|\frac{p^2}{2m_0}+V|u_n]_{z_0}\mathrm{d}z_0 - [u_m|\frac{p^2}{2m_0}+\frac{V_{\text{A}}+V_{\text{B}}}{2}|u_n]L_{\text{if}}. \tag{3.9}$$

When one chooses the bulk reference crystal as a virtual crystal being the "average" of crystals A and B, the last expression reduces to

$$\Omega_{mn} = \int_{-L_{\text{if}}/2}^{+L_{\text{if}}/2}[u_m|\frac{p^2}{2m_0}+V|u_n]_{z_0}\mathrm{d}z_0 - E_m\delta_{mn}L_{\text{if}}, \tag{3.10}$$

where E_m is the energy of the band m at the Γ point, and u_m is the corresponding Bloch functions. In the case of square-based pyramid with base width to height ratio $b/h = 2$, the total interface contribution to the Hamiltonian is

$$\begin{aligned} H_{\text{if}} = & \Omega([001])\delta\theta(z) \\ & + \Omega([101])\delta\theta(\mathbf{r}\cdot\mathbf{n}_1 - l) + \Omega([011])\delta\theta(\mathbf{r}\cdot\mathbf{n}_2 - l) \\ & + \Omega([\bar{1}01])\delta\theta(\mathbf{r}\cdot\mathbf{n}_3 - l) + \Omega([0\bar{1}1])\delta\theta(\mathbf{r}\cdot\mathbf{n}_4 - l). \end{aligned} \tag{3.11}$$

In the above equation, $\delta\theta(z)$ function represents the delta function at a given interface, with an additional constraint that the function is nonzero only at the face of the pyramid. The vectors $\mathbf{n}_i$ are the unit vectors perpendicular to the faces of the pyramid and are given as $\mathbf{n}_1 = 1/\sqrt{2}\cdot(1,0,1)$, $\mathbf{n}_2 = 1/\sqrt{2}\cdot(0,1,1)$, $\mathbf{n}_3 = 1/\sqrt{2}\cdot(-1,0,1)$, $\mathbf{n}_4 = 1/\sqrt{2}\cdot(0,-1,1)$ and $l = b/(2\sqrt{2})$. The choice of the coordinate system was made in such a way that the vertices of the pyramid are at the points $(b/2,-b/2,0)$, $(b/2,b/2,0)$, $(-b/2,b/2,0)$, $(-b/2,-b/2,0)$ and $(0,0,h)$.

The explicit form of the matrices Ω can be obtained from density functional theory or empirical pseudopotential calculations. The reader is referred to [83] for a detailed description of the extraction of the matrices Ω from empirical pseudopotentials, while the final results, in the basis $|s_a\rangle, |p_{x,b}\rangle, |p_{y,b}\rangle, |p_{z,b}\rangle$ (see Sect. 3.5 for the meaning of this basis), are given as:

$$\Omega([001]) = \begin{pmatrix} 0 & 0 & 0 & -a \\ & 0 & b & 0 \\ & & 0 & 0 \\ & & & 0 \end{pmatrix}, \tag{3.12}$$

$$\Omega([101]) = \begin{pmatrix} 0 & -c & 0 & -c \\ & 0 & d & 0 \\ & & 0 & d \\ & & & 0 \end{pmatrix}, \tag{3.13}$$

$$\Omega([011]) = \begin{pmatrix} 0 & 0 & -c & -c \\ & 0 & d & d \\ & & 0 & 0 \\ & & & 0 \end{pmatrix}, \tag{3.14}$$

$$\Omega([\bar{1}01]) = \begin{pmatrix} 0 & c & 0 & -c \\ & 0 & d & 0 \\ & & 0 & -d \\ & & & 0 \end{pmatrix}, \tag{3.15}$$

Table 3.1 Relevant material parameters of the interfaces: InAs/GaAs and GaAs/$Al_{0.35}Ga_{0.65}As$

	InAs/GaAs	GaAs/$Al_{0.35}Ga_{0.65}As$
a [eVÅ]	0.61220	0.14899
b [eVÅ]	−0.36633	0.02861
c [eVÅ]	−0.32427	−0.07039
d [eVÅ]	0.02855	0.00008

$$\Omega([0\bar{1}1]) = \begin{pmatrix} 0 & 0 & c & -c \\ & 0 & d & -d \\ & & 0 & 0 \\ & & & 0 \end{pmatrix}. \tag{3.16}$$

The relevant parameters, a, b, c and d, for two representative interfaces InAs/GaAs and GaAs/$Al_{0.35}Ga_{0.65}As$ are given in Table 3.1.

3.4 Symmetry of the Interface Hamiltonian

Proper understanding of the Hamiltonian symmetry group is of great importance for several reasons: (a) symmetry can be used to reduce the computational cost; (b) symmetry induces selection rules for certain physical processes, such as for example light absorption and emission. For this reason, we analyze the influence of the interface Hamiltonians introduced in Sect. 3.3 on the symmetry of the system. The Hamiltonian for the [001] interface in the basis $|s_a\rangle, |p_{x,b}\rangle, |p_{y,b}\rangle, |p_{z,b}\rangle$ reads:

$$H_{[001]} = \begin{pmatrix} 0 & 0 & 0 & -a \\ & 0 & b & 0 \\ & & 0 & 0 \\ & & & 0 \end{pmatrix} \delta\theta(z). \tag{3.17}$$

To represent the actions of the rotation operators on the envelope function spinors, it is more convenient to work in the basis of eigenstates of the z-component of the orbital quasi-angular momentum

$$\{u_1, \ldots, u_4\} = \{|s_a\rangle, \frac{1}{\sqrt{2}}\left(|p_{x,b}\rangle + i\,|p_{y,b}\rangle\right), \frac{1}{\sqrt{2}}\left(|p_{x,b}\rangle - i\,|p_{y,b}\rangle\right), |p_{z,b}\rangle\}. \tag{3.18}$$

In this basis, the same Hamiltonian, $H_{[001]}$, reads

$$H_{[001]} = \begin{pmatrix} 0 & 0 & 0 & -a \\ & 0 & -ib & 0 \\ & & 0 & 0 \\ & & & 0 \end{pmatrix} \delta\theta(z). \tag{3.19}$$

The action of the representation of the rotation R_φ, where $\varphi = n\pi/2$ on the spinors is given as

$$D(R_\varphi)\begin{pmatrix}\psi_1(\mathbf{r})\\ \psi_2(\mathbf{r})\\ \psi_3(\mathbf{r})\\ \psi_4(\mathbf{r})\end{pmatrix} = \begin{pmatrix}\psi_1(R_\varphi\mathbf{r})\\ e^{-i\varphi}\psi_2(R_\varphi\mathbf{r})\\ e^{i\varphi}\psi_3(R_\varphi\mathbf{r})\\ \psi_4(R_\varphi\mathbf{r})\end{pmatrix}. \tag{3.20}$$

To prove that $D(R_\varphi)$ commutes with $H_{[001]}$ it is sufficient to show that operators $H_{[001]}D(R_\varphi)$ and $D(R_\varphi)H_{[001]}$ give the same result when acting on the basis states $(e^{i\mathbf{k}\cdot\mathbf{r}},0,0,0)^\top$, $(0,e^{i\mathbf{k}\cdot\mathbf{r}},0,0)^\top$, $(0,0,e^{i\mathbf{k}\cdot\mathbf{r}},0)^\top$, $(0,0,0,e^{i\mathbf{k}\cdot\mathbf{r}})^\top$ that span the Hilbert space of spinors. By explicitly performing the calculation one obtains:

$$H_{[001]}D(R_\varphi)\begin{pmatrix}0\\ e^{i\mathbf{k}\cdot\mathbf{r}}\\ 0\\ 0\end{pmatrix} = H_{[001]}\begin{pmatrix}0\\ e^{-i\varphi}e^{i(R_\varphi^{-1}\mathbf{k})\cdot\mathbf{r}}\\ 0\\ 0\end{pmatrix} = \begin{pmatrix}0\\ 0\\ ibe^{-i\varphi}\\ 0\end{pmatrix} e^{i(R_\varphi^{-1}\mathbf{k})\cdot\mathbf{r}}\delta\theta(z) \tag{3.21}$$

and on the other hand

$$D(R_\varphi)H_{[001]}\begin{pmatrix}0\\ e^{i\mathbf{k}\cdot\mathbf{r}}\\ 0\\ 0\end{pmatrix} = \begin{pmatrix}0\\ 0\\ ib\\ 0\end{pmatrix} e^{i\mathbf{k}\cdot\mathbf{r}}\delta\theta(z) = \begin{pmatrix}0\\ 0\\ ibe^{i\varphi}\\ 0\end{pmatrix} e^{i(R_\varphi^{-1}\mathbf{k})\cdot\mathbf{r}}\delta\theta(z). \tag{3.22}$$

This implies that $D(R_\varphi)$ and $H_{[001]}$ commute only if $\varphi = n\pi$.

Next, we proceed with the proof that $H_{[001]}$ commutes with the operator $D(\sigma_v)$. For this proof, it is convenient to work in the basis

$$\{u_1,\ldots,u_4\} = \{|s_a\rangle, \frac{1}{\sqrt{2}}\left(|p_{x,b}\rangle + |p_{y,b}\rangle\right), \frac{1}{\sqrt{2}}\left(|p_{x,b}\rangle - |p_{y,b}\rangle\right), |p_{z,b}\rangle\}. \tag{3.23}$$

In this basis $H_{[001]}$ reads

$$H_{[001]} = \begin{pmatrix}0 & 0 & 0 & -a\\ b & 0 & 0 & \\ & -b & 0 & \\ & & 0 & \end{pmatrix}\delta\theta(z), \tag{3.24}$$

while the action of the operator $D(\sigma_v)$ on the spinor is given as

$$D(\sigma_v)\begin{pmatrix}\psi_1(x,y,z)\\ \psi_2(x,y,z)\\ \psi_3(x,y,z)\\ \psi_4(x,y,z)\end{pmatrix} = \begin{pmatrix}\psi_1(y,x,z)\\ \psi_2(y,x,z)\\ -\psi_3(y,x,z)\\ \psi_4(y,x,z)\end{pmatrix}. \tag{3.25}$$

It follows

$$H_{[001]}D(\sigma_v)\begin{pmatrix}0\\0\\e^{i\mathbf{k}\cdot\mathbf{r}}\\0\end{pmatrix} = H_{[001]}\begin{pmatrix}0\\0\\-e^{i(k_xy+k_yx+k_zz)}\\0\end{pmatrix} = \begin{pmatrix}0\\0\\b\\0\end{pmatrix} e^{i(k_xy+k_yx+k_zz)}\delta\theta(z). \tag{3.26}$$

On the other hand

$$D(\sigma_v)H_{[001]}\begin{pmatrix}0\\0\\e^{i\mathbf{k}\cdot\mathbf{r}}\\0\end{pmatrix} = D(\sigma_v)\begin{pmatrix}0\\0\\-b\\0\end{pmatrix}\delta\theta(z)e^{i\mathbf{k}\cdot\mathbf{r}} = \begin{pmatrix}0\\0\\b\\0\end{pmatrix} e^{i(k_xy+k_yx+k_zz)}\delta\theta(z). \tag{3.27}$$

In this way we proved that the symmetry of the interface Hamiltonian, $H_{[001]}$, is C_{2v}. Similarly, it can be shown that the symmetry of $H_{[110]} + H_{[1\bar{1}0]} + H_{[\bar{1}\bar{1}0]} + H_{[\bar{1}10]}$ term is C_{2v} too. Consequently, if the interface effects are included, the symmetry of the model will be reduced from an artificially high C_{4v} to correct C_{2v}.

3.5 The 14-Band k·p Hamiltonian

In the previous section, we have demonstrated that the inclusion of interface effects leads to the correct symmetry of the model. In this section, we will analyze how the inclusion of bands beyond the standard 8 bands affects the symmetry. For this purpose, we will investigate the 14-band Hamiltonian which includes the second conduction band Γ_{5c} (see Fig. 3.1) in addition to the standard 8 bands.

The 14-band **k·p** Hamiltonian in the basis that consists of states that originate from p bonding and antibonding (denoted as p_b and p_a) and s antibonding (denoted as s_a) states of the atoms in the bulk, see Fig. 3.1:

$$\begin{aligned}\{\ &|p_{x,a}\uparrow\rangle, |p_{y,a}\uparrow\rangle, |p_{z,a}\uparrow\rangle, |s_a\uparrow\rangle, |p_{x,b}\uparrow\rangle, |p_{y,b}\uparrow\rangle, |p_{z,b}\uparrow\rangle,\\ &|p_{x,a}\downarrow\rangle, |p_{y,a}\downarrow\rangle, |p_{z,a}\downarrow\rangle, |s_a\downarrow\rangle, |p_{x,b}\downarrow\rangle, |p_{y,b}\downarrow\rangle, |p_{z,b}\downarrow\rangle\},\end{aligned} \tag{3.28}$$

where $\uparrow$ and $\downarrow$ denote spin-up and spin-down states respectively, reads

$$H = \begin{pmatrix} |\uparrow\rangle & |\downarrow\rangle \\ G & 0 \\ 0 & G \end{pmatrix} + \begin{pmatrix} |\uparrow\rangle & |\downarrow\rangle \\ G_{\mathrm{so}} & \Gamma \\ -\Gamma^* & G^*_{\mathrm{so}} \end{pmatrix}, \tag{3.29}$$

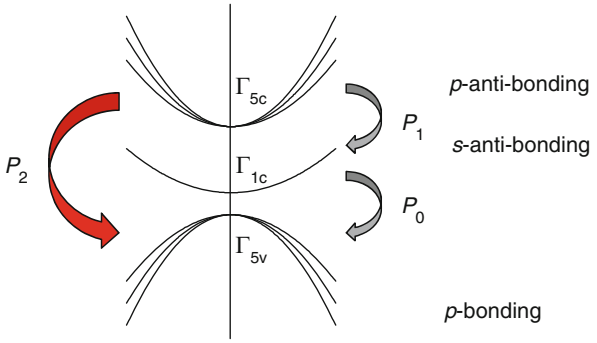


Fig. 3.1 The scheme of the band structure of the material with zincblende crystal structure around the Γ point in the first Brillouin zone, including the coupling elements between the relevant bands in the 14-band **k**·**p** Hamiltonian

where

$$G = \begin{pmatrix} |p_{x,a}\rangle & |p_{y,a}\rangle & |p_{z,a}\rangle & |s_a\rangle & |p_{x,b}\rangle & |p_{y,b}\rangle & |p_{z,b}\rangle \\ E_{p_a} & 0 & 0 & iP_1k_x & 0 & -iP_2k_z & -iP_2k_y \\ & E_{p_a} & 0 & iP_1k_y & -PiP_2k_z & 0 & -iP_2k_x \\ & & E_{p_a} & iP_1k_z & -iP_2k_y & -iP_2k_x & 0 \\ & & & E_{s_a} & iP_0k_x & iP_0k_y & iP_0k_z \\ & & & & E_{p_{x,b}} & W_1 & W_2 \\ & & & & & E_{p_{y,b}} & W_3 \\ & & & & & & E_{p_{z,b}} \end{pmatrix}, \tag{3.30}$$

$$G_{\mathrm{so}} = \begin{pmatrix} |p_{x,a}\rangle & |p_{y,a}\rangle & |p_{z,a}\rangle & |s_a\rangle & |p_{x,b}\rangle & |p_{y,b}\rangle & |p_{z,b}\rangle \\ \frac{2}{3}\Delta_{\mathrm{so}}(p_a) & -\frac{i}{3}\Delta_{\mathrm{so}}(p_a) & 0 & 0 & 0 & \frac{i}{3}\Delta_{\mathrm{cf}} & 0 \\ & \frac{2}{3}\Delta_{\mathrm{so}}(p_a) & 0 & 0 & -\frac{i}{3}\Delta_{\mathrm{cf}} & 0 & 0 \\ & & \frac{2}{3}\Delta_{\mathrm{so}}(p_a) & 0 & 0 & 0 & 0 \\ & & & 0 & 0 & 0 & 0 \\ & & & & -\frac{1}{3}\Delta_{\mathrm{so}}(p_b) & -\frac{i}{3}\Delta_{\mathrm{so}}(p_b) & 0 \\ & & & & & -\frac{1}{3}\Delta_{\mathrm{so}}(p_b) & 0 \\ & & & & & & -\frac{1}{3}\Delta_{\mathrm{so}}(p_b) \end{pmatrix}$$

and

$$\Gamma = \begin{pmatrix} |p_{x,a}\rangle & |p_{y,a}\rangle & |p_{z,a}\rangle & |s_a\rangle & |p_{x,b}\rangle & |p_{y,b}\rangle & |p_{z,b}\rangle \\ 0 & 0 & \frac{1}{3}\Delta_{\mathrm{so}}(p_a) & 0 & 0 & 0 & -\frac{1}{3}\Delta_{\mathrm{cf}} \\ 0 & 0 & -\frac{i}{3}\Delta_{\mathrm{so}}(p_a) & 0 & 0 & 0 & \frac{i}{3}\Delta_{\mathrm{cf}} \\ -\frac{1}{3}\Delta_{\mathrm{so}}(p_a) & \frac{i}{3}\Delta_{\mathrm{so}}(p_a) & 0 & 0 & \frac{1}{3}\Delta_{\mathrm{cf}} & -\frac{i}{3}\Delta_{\mathrm{cf}} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -\frac{1}{3}\Delta_{\mathrm{cf}} & 0 & 0 & 0 & \frac{1}{3}\Delta_{\mathrm{so}}(p_b) \\ 0 & 0 & \frac{i}{3}\Delta_{\mathrm{cf}} & 0 & 0 & 0 & -\frac{i}{3}\Delta_{\mathrm{so}}(p_b) \\ \frac{1}{3}\Delta_{\mathrm{cf}} & -\frac{i}{3}\Delta_{\mathrm{cf}} & 0 & 0 & -\frac{1}{3}\Delta_{\mathrm{so}}(p_b) & \frac{i}{3}\Delta_{\mathrm{so}}(p_b) & 0 \end{pmatrix}.$$

The terms in the previous equations are given as:

$$E_{p_a} = E(\Gamma_{5c}), \qquad E_{s_a} = E(\Gamma_{1c}) + \left(\frac{\hbar^2}{2m_0}\right)\gamma_c k^2$$

$$E_{p_{x,b}} = E(\Gamma_{5v}) - (P+Q) - \frac{\sqrt{3}}{2}(R^* + R),$$

$$E_{p_{y,b}} = E(\Gamma_{5v}) - (P+Q) + \frac{\sqrt{3}}{2}(R^* + R)$$

$$E_{p_{z,b}} = E(\Gamma_{5v}) - (P-2Q), \qquad E_{g0} = E(\Gamma_{1c}) - E(\Gamma_{5v}),$$

$$E_{g1} = E(\Gamma_{5c}) - E(\Gamma_{5v})$$

$$W_1 = -i\frac{\sqrt{3}}{2}(R - R^*), \qquad W_2 = -\frac{\sqrt{3}}{2}(S + S^*), \qquad W_3 = -i\frac{\sqrt{3}}{2}(S - S^*)$$

$$P = P_k + P_\epsilon, \qquad Q = Q_k + Q_\epsilon, \qquad R = R_k + R_\epsilon, \qquad S = S_k + S_\epsilon$$

$$P_k = \left(\frac{\hbar^2}{2m_0}\right)\gamma_1(k_x^2 + k_y^2 + k_z^2), \qquad Q_k = \left(\frac{\hbar^2}{2m_0}\right)\gamma_2(k_x^2 + k_y^2 - 2k_z^2)$$

$$R_k = \left(\frac{\hbar^2}{2m_0}\right)\sqrt{3}[\gamma_2(k_x^2 - k_y^2) - 2i\gamma_3 k_x k_y], \quad S_k = \left(\frac{\hbar^2}{2m_0}\right)\sqrt{6}\gamma_3(k_x - ik_y)k_z$$

$$P_\epsilon = -a_v(\epsilon_{xx} + \epsilon_{yy} + \epsilon_{zz}), \qquad Q_\epsilon = -\frac{b_{ax}}{2}(\epsilon_{xx} + \epsilon_{yy} - 2\epsilon_{zz})$$

$$R_\epsilon = -\frac{\sqrt{3}}{2}b_{ax}(\epsilon_{xx} - \epsilon_{yy}) + i\,d_{ax}\epsilon_{xy}, \qquad S_\epsilon = -\frac{d_{ax}}{\sqrt{2}}(\epsilon_{zx} - i\epsilon_{yz})$$

$$E_{P0} = 2m_0 P_0^2/\hbar^2, \qquad E_{P1} = 2m_0 P_1^2/\hbar^2, \qquad E_{P2} = 2m_0 P_2^2/\hbar^2$$

$$\begin{aligned}\gamma_c =& \frac{1}{m^*} - \frac{E_{P0}}{3}\left[\frac{2}{E_{g0}} + \frac{1}{E_{g0} + \Delta_{\rm so}(p_b)}\right] \\ &+ \frac{E_{P1}}{3}\left[\frac{1}{E_{g1} - E_{g0}} + \frac{2}{E_{g1} - E_{g0} + \Delta_{\rm so}(p_a)}\right],\end{aligned}$$

$$\gamma_1 = \gamma_1^L - \frac{1}{3}\frac{E_{P0}}{E_{g0} + \Delta_{\rm so}(p_b)/3} - \frac{2}{3}\frac{E_{P2}}{E_{g1} + \Delta_{\rm so}(p_b)/3 + 2\Delta_{\rm so}(p_a)/3},$$

$$\gamma_2 = \gamma_2^L - \frac{1}{6}\frac{E_{P0}E_{g0} + \Delta_{\rm so}(p_b)/3}{+}\frac{1}{6}\frac{E_{P2}}{E_{g1} + \Delta_{\rm so}(p_b)/3 + 2\Delta_{\rm so}(p_a)/3},$$

$$\gamma_3 = \gamma_3^L - \frac{1}{6}\frac{E_{P0}E_{g0} + \Delta_{\rm so}(p_b)/3}{-}\frac{1}{6}\frac{E_{P2}}{E_{g1} + \Delta_{\rm so}(p_b)/3 + 2\Delta_{\rm so}(p_a)/3}.$$

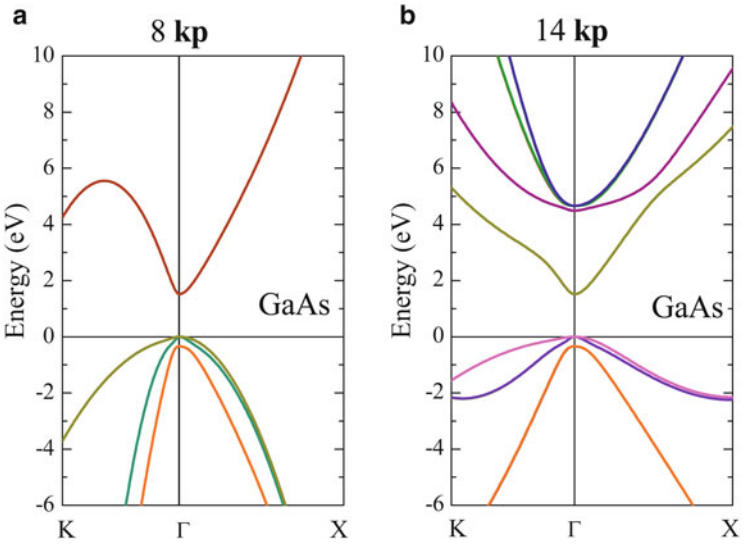


Fig. 3.2 Electronic structure of unstrained GaAs bulk material calculated using the 8-band **k·p** (**a**) and 14-band **k·p** (**b**) Hamiltonian along K–Γ–X path in the first Brillouin zone

$E(\Gamma_{5c})$, $E(\Gamma_{1c})$, $E(\Gamma_{5v})$ are the bulk band energies at the Γ point, m_0 the electron mass, $\hbar$ the reduced Planck's constant, k_x, k_y, k_z denote the components of the wave vector along the crystallographic directions [100], [010], and [001] in the vicinity of the Γ point, P_0, P_1, P_2 are the momentum matrix elements between Γ_{1c} and Γ_{5v}, Γ_{5c} and Γ_{1c}, and Γ_{5c} and Γ_{5v} states respectively, E_{P0}, E_{P1}, E_{P2} are the Kane energies related to P_0, P_1 and P_2 respectively, $\gamma_1^L, \gamma_2^L, \gamma_3^L$ are the Luttinger parameters, m^* is the effective mass in the conduction band, $\Delta_{\mathrm{SO}}(p_b)$ is the spin-orbit splitting between p-bonding states in the valence band, $\Delta_{\mathrm{SO}}(p_a)$ is the spin-orbit splitting between p-antibonding states in the conduction band, Δ_{cf} is the crystal field splitting, ϵ_{ij} are the strain tensor components, a_c and a_v are the conduction band and valence band hydrostatic deformation potentials respectively, b_{ax} and d_{ax} are the shear deformation potentials along the [001] and [111] direction respectively. The values of relevant material parameters are given in the Appendix.

In Fig. 3.2 we plot the electronic structure of unstrained GaAs calculated using the 8-band **k·p** (a) and 14-band **k·p** (b) Hamiltonian along the K–Γ–X path in the first Brillouin zone. It is visible that additional band couplings in the 14-band **k·p** Hamiltonian prevent dangerous appearance of spurious solutions that might exist in the 8-band **k·p** Hamiltonian. These spurious solutions in the 8-band **k·p** Hamiltonian are related to appearance of the artificial folding points in the lowest conduction band due to small basis size of such a Hamiltonian.

3.6 Symmetry of the 14-Band k·p Hamiltonian

3.6.1 *Symmetry of the 8-Band* k·p *Hamiltonian*

To establish the symmetry of the kinetic part of the 14-band **k·p** Hamiltonian we start with the analysis of the 8-band **k·p** Hamiltonian, which is a constituent part of the 14-band Hamiltonian. The kinetic part of the 8-band **k·p** Hamiltonian consists of two identical uncoupled 4×4 blocks. In the basis

$$\{u_1, \ldots, u_4\} = \{|s_a\rangle, |p_{x,b}\rangle, |p_{y,b}\rangle, |p_{z,b}\rangle\} \tag{3.31}$$

these blocks read:

$$H_4 = \begin{pmatrix} E_c(\mathbf{r}) & ik_x P_0 & ik_y P_0 & ik_z P_0 \\ & E_v(\mathbf{r}) & 0 & 0 \\ & & E_v(\mathbf{r}) & 0 \\ & & & E_v(\mathbf{r}) \end{pmatrix}, \tag{3.32}$$

where $E_c(\mathbf{r}) = E_{s_a}(\mathbf{r})$ and $E_v(\mathbf{r}) = E_{p_b}(\mathbf{r})$. In (3.32), the effect of remote bands was not included as its inclusion does not affect the symmetry considerations. We will show that this Hamiltonian applied to square-based pyramidal quantum dots has C_{4v} symmetry. To do this, it is sufficient to show that the blocks H_4 commute with the generators of the group—the rotation $R_{\pi/2}$ and the reflection σ_v.

To represent the actions of the rotation operators on the envelope function spinors, it is more convenient to work in the basis of eigenstates of the z-component of the orbital quasi-angular momentum (3.18) where the block H_4 reads

$$H_4 = \begin{pmatrix} E_c(\mathbf{r}) & ik_+ P_0 & ik_- P_0 & ik_z P_0 \\ & E_v(\mathbf{r}) & 0 & 0 \\ & & E_v(\mathbf{r}) & 0 \\ & & & E_v(\mathbf{r}) \end{pmatrix}, \tag{3.33}$$

where $k_\pm = \frac{1}{\sqrt{2}}\left(k_x \pm ik_y\right)$. The action of the representation of the rotation R_φ, where $\varphi = n\pi/2$ on the spinors is given by (3.20).

To prove that $D(R_\varphi)$ commutes with H_4 it is sufficient to show that operators $H_4 D(R_\varphi)$ and $D(R_\varphi) H_4$ give the same result when acting on the basis states $(e^{i\mathbf{k}\cdot\mathbf{r}}, 0, 0, 0)^\top$, $(0, e^{i\mathbf{k}\cdot\mathbf{r}}, 0, 0)^\top$, $(0, 0, e^{i\mathbf{k}\cdot\mathbf{r}}, 0)^\top$, $(0, 0, 0, e^{i\mathbf{k}\cdot\mathbf{r}})^\top$ that span the Hilbert space of spinors. By explicitly performing the calculation one gets on the one hand

$$H_4 D(R_\varphi) \begin{pmatrix} e^{i\mathbf{k}\cdot\mathbf{r}} \\ 0 \\ 0 \\ 0 \end{pmatrix} = H_4 \begin{pmatrix} e^{i(R_\varphi^{-1}\mathbf{k})\cdot\mathbf{r}} \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} E_c(\mathbf{r}) \\ -iP_0(R_\varphi^{-1}\mathbf{k})_- \\ -iP_0(R_\varphi^{-1}\mathbf{k})_+ \\ -iP_0(R_\varphi^{-1}\mathbf{k})_z \end{pmatrix} e^{i(R_\varphi^{-1}\mathbf{k})\cdot\mathbf{r}} \tag{3.34}$$

and on the other hand

$$D(R_\varphi)H_4\begin{pmatrix} e^{i\mathbf{k}\cdot\mathbf{r}} \\ 0 \\ 0 \\ 0 \end{pmatrix} = D(R_\varphi)\begin{pmatrix} E_c(\mathbf{r}) \\ -iP_0k_- \\ -iP_0k_+ \\ -iP_0k_z \end{pmatrix} e^{i\mathbf{k}\cdot\mathbf{r}} = \begin{pmatrix} E_c(R_\varphi\mathbf{r}) \\ e^{-i\varphi}(-i)P_0k_- \\ e^{i\varphi}(-i)P_0k_+ \\ -iP_0k_z \end{pmatrix} e^{i(R_\varphi^{-1}\mathbf{k})\cdot\mathbf{r}}. \tag{3.35}$$

Due to the symmetry of the dot shape it follows that $E_c(\mathbf{r}) = E_c(R_\varphi\mathbf{r})$. Furthermore, one can straightforwardly show that $(R_\varphi^{-1}\mathbf{k})_- = e^{-i\varphi}k_-$ and $(R_\varphi^{-1}\mathbf{k})_+ = e^{i\varphi}k_+$. From these identities, it follows that

$$\left[D(R_\varphi)H_4 - H_4D(R_\varphi)\right]\begin{pmatrix} e^{i\mathbf{k}\cdot\mathbf{r}} \\ 0 \\ 0 \\ 0 \end{pmatrix} = 0. \tag{3.36}$$

Using the same procedure, one can also show that $D(R_\varphi)H_4$ and $H_4D(R_\varphi)$ give the same result when acting on the other basis vectors $(0, e^{i\mathbf{k}\cdot\mathbf{r}}, 0, 0)^\top$, $(0, 0, e^{i\mathbf{k}\cdot\mathbf{r}}, 0)^\top$, $(0, 0, 0, e^{i\mathbf{k}\cdot\mathbf{r}})^\top$, which completes the proof that $D(R_\varphi)H_4$ and $H_4D(R_\varphi)$ commute.

Next, we proceed with the proof that H_4 commutes with the operator $D(\sigma_v)$. For this proof, it is convenient to work in the basis (3.23). The Hamiltonian in this basis reads

$$H_4 = \begin{pmatrix} E_c(\mathbf{r}) & i\frac{1}{\sqrt{2}}(k_x + k_y)P_0 & i\frac{1}{\sqrt{2}}(k_x - k_y)P_0 & ik_zP_0 \\ & E_v(\mathbf{r}) & 0 & 0 \\ & & E_v(\mathbf{r}) & 0 \\ & & & E_v(\mathbf{r}) \end{pmatrix}. \tag{3.37}$$

In this basis, the action of the operator $D(\sigma_v)$ on the spinor is given by (3.25). It follows

$$H_4D(\sigma_v)\begin{pmatrix} e^{i\mathbf{k}\cdot\mathbf{r}} \\ 0 \\ 0 \\ 0 \end{pmatrix} = H_4\begin{pmatrix} e^{i(k_xy+k_yx+k_zz)} \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} E_c(\mathbf{r}) \\ -iP_0\frac{1}{\sqrt{2}}(k_x+k_y) \\ -iP_0\frac{1}{\sqrt{2}}(k_y-k_x) \\ -iP_0k_z \end{pmatrix} e^{i(k_xy+k_yx+k_zz)}. \tag{3.38}$$

On the other hand

$$
\begin{aligned}
D(\sigma_v)H_4\begin{pmatrix} e^{i\mathbf{k}\cdot\mathbf{r}} \\ 0 \\ 0 \\ 0 \end{pmatrix} &= D(\sigma_v)\begin{pmatrix} E_c(\mathbf{r}) \\ -iP_0\frac{1}{\sqrt{2}}(k_x+k_y) \\ -iP_0\frac{1}{\sqrt{2}}(k_x-k_y) \\ -iP_0k_z \end{pmatrix} e^{i\mathbf{k}\cdot\mathbf{r}} \\
&= \begin{pmatrix} E_c(y,x,z) \\ -iP_0\frac{1}{\sqrt{2}}(k_x+k_y) \\ -iP_0\frac{1}{\sqrt{2}}(k_x-k_y)\cdot(-1) \\ -iP_0k_z \end{pmatrix} e^{i(k_xy+k_yx+k_zz)}
\end{aligned}
\tag{3.39}
$$

and consequently

$$
[D(\sigma_v)H_4 - H_4D(\sigma_v)]\begin{pmatrix} e^{i\mathbf{k}\cdot\mathbf{r}} \\ 0 \\ 0 \\ 0 \end{pmatrix} = 0. \tag{3.40}
$$

One can straightforwardly check this equality for other basis vectors $(0, e^{i\mathbf{k}\cdot\mathbf{r}}, 0, 0)^\top$, $(0, 0, e^{i\mathbf{k}\cdot\mathbf{r}}, 0)^\top$, $(0, 0, 0, e^{i\mathbf{k}\cdot\mathbf{r}})^\top$. That completes the proof that the kinetic part of the 8-band $\mathbf{k}\cdot\mathbf{p}$ Hamiltonian is of C_{4v} symmetry.

3.6.2 Symmetry of the Whole 14-Band Hamiltonian

We proceed with the analysis of the kinetic part of the 14-band $\mathbf{k}\cdot\mathbf{p}$ Hamiltonian. This Hamiltonian consists of two uncoupled 7×7 blocks that read (excluding the remote band effects that do not affect the symmetry considerations):

$$
H_7 = \begin{pmatrix}
|p_{x,a}\rangle & |p_{y,a}\rangle & |p_{z,a}\rangle & |s_a\rangle & |p_{x,b}\rangle & |p_{y,b}\rangle & |p_{z,b}\rangle \\
E_{c2}(\mathbf{r}) & 0 & 0 & iP_1k_x & 0 & -iP_2k_z & -iP_2k_y \\
 & E_{c2}(\mathbf{r}) & 0 & iP_1k_y & -iP_2k_z & 0 & -iP_2k_x \\
 & & E_{c2}(\mathbf{r}) & iP_1k_z & -iP_2k_y & -iP_2k_x & 0 \\
 & & & E_c(\mathbf{r}) & iP_0k_x & iP_0k_y & iP_0k_z \\
 & & & & E_v(\mathbf{r}) & 0 & 0 \\
 & & & & & E_v(\mathbf{r}) & 0 \\
 & & & & & & E_v(\mathbf{r})
\end{pmatrix}. \tag{3.41}
$$

where $E_{c2}(\mathbf{r}) = E_{p_a}(\mathbf{r})$. The 4×4 block that contains the elements P_1 is of the same form as H_4 and therefore commutes with the operators that represent the elements of the C_{4v} group. One needs therefore to understand the symmetry properties of the block that contains the P_2 elements only. This block reads:

$$H_7 = \begin{pmatrix} |p_{x,a}\rangle & |p_{y,a}\rangle & |p_{z,a}\rangle & |s_a\rangle & |p_{x,b}\rangle & |p_{y,b}\rangle & |p_{z,b}\rangle \\ 0 & 0 & 0 & 0 & 0 & -iP_2k_z & -iP_2k_y \\ & 0 & 0 & 0 & -iP_2k_z & 0 & -iP_2k_x \\ & & 0 & 0 & -iP_2k_y & -iP_2k_x & 0 \\ & & & 0 & 0 & 0 & 0 \\ & & & & 0 & 0 & 0 \\ & & & & & 0 & 0 \\ & & & & & & 0 \end{pmatrix}. \tag{3.42}$$

The most convenient basis to represent the action of the rotation operators is the basis

$$\begin{aligned} \{u_1, \ldots, u_7\} = \{ & \tfrac{1}{\sqrt{2}} \left(|p_{x,a}\rangle + i\,|p_{y,a}\rangle \right), \tfrac{1}{\sqrt{2}} \left(|p_{x,a}\rangle - i\,|p_{y,a}\rangle \right), |p_{z,a}\rangle, |s_a\rangle, \\ & \tfrac{1}{\sqrt{2}} \left(|p_{x,b}\rangle + i\,|p_{y,b}\rangle \right), \tfrac{1}{\sqrt{2}} \left(|p_{x,b}\rangle - i\,|p_{y,b}\rangle \right), |p_{z,b}\rangle \}. \end{aligned} \tag{3.43}$$

In this basis, the H_7 block reads

$$H_7 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & -P_2k_z & -P_2k_+ \\ & 0 & 0 & 0 & P_2k_z & 0 & P_2k_- \\ & & 0 & 0 & P_2k_- & -P_2k_+ & 0 \\ & & & 0 & 0 & 0 & 0 \\ & & & & 0 & 0 & 0 \\ & & & & & 0 & 0 \\ & & & & & & 0 \end{pmatrix}. \tag{3.44}$$

The action of the rotation operator on the spinor is given as

$$D(R_\varphi) \begin{pmatrix} \psi_1(\mathbf{r}) \\ \psi_2(\mathbf{r}) \\ \psi_3(\mathbf{r}) \\ \psi_4(\mathbf{r}) \\ \psi_5(\mathbf{r}) \\ \psi_6(\mathbf{r}) \\ \psi_7(\mathbf{r}) \end{pmatrix} = \begin{pmatrix} e^{-i\varphi}\psi_1(R_\varphi\mathbf{r}) \\ e^{i\varphi}\psi_2(R_\varphi\mathbf{r}) \\ \psi_3(R_\varphi\mathbf{r}) \\ \psi_4(R_\varphi\mathbf{r}) \\ e^{-i\varphi}\psi_5(R_\varphi\mathbf{r}) \\ e^{i\varphi}\psi_6(R_\varphi\mathbf{r}) \\ \psi_7(R_\varphi\mathbf{r}) \end{pmatrix}. \tag{3.45}$$

One then gets on the one hand

$$H_7 D(R_\varphi)\begin{pmatrix} e^{i\mathbf{k}\cdot\mathbf{r}} \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} = H_7 \begin{pmatrix} e^{-i\varphi} e^{i(R_\varphi^{-1}\mathbf{k})\cdot\mathbf{r}} \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ -e^{-i\varphi} P_2 (R_\varphi^{-1}\mathbf{k})_z \\ e^{-i\varphi}(-P_2)(R_\varphi^{-1}\mathbf{k})_- \end{pmatrix} e^{i(R_\varphi^{-1}\mathbf{k})\cdot\mathbf{r}} \tag{3.46}$$

and on the other hand

$$D(R_\varphi) H_7 \begin{pmatrix} e^{i\mathbf{k}\cdot\mathbf{r}} \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} = D(R_\varphi) \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ -P_2 k_z \\ -P_2 k_- \end{pmatrix} e^{i\mathbf{k}\cdot\mathbf{r}} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ -e^{i\varphi} P_2 k_z \\ -P_2 k_- \end{pmatrix} e^{i(R_\varphi^{-1}\mathbf{k})\cdot\mathbf{r}}. \tag{3.47}$$

It then follows that

$$\left[D(R_\varphi) H_7 - H_7 D(R_\varphi)\right] \begin{pmatrix} e^{i\mathbf{k}\cdot\mathbf{r}} \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} = 0 \tag{3.48}$$

only if $e^{i\varphi} = e^{-i\varphi}$, which implies $\varphi = n\pi$. One can further straightforwardly extend this result to other basis vectors.

The most convenient basis to represent the action of the $D(\sigma_v)$ operator is the basis

$$\begin{aligned} \{u_1, \ldots, u_7\} = \{ & \tfrac{1}{\sqrt{2}} \left(|p_{x,a}\rangle + |p_{y,a}\rangle\right), \tfrac{1}{\sqrt{2}} \left(|p_{x,a}\rangle - |p_{y,a}\rangle\right), |p_{z,a}\rangle, |s_a\rangle, \\ & \tfrac{1}{\sqrt{2}} \left(|p_{x,b}\rangle + |p_{y,b}\rangle\right), \tfrac{1}{\sqrt{2}} \left(|p_{x,b}\rangle - |p_{y,b}\rangle\right), |p_{z,b}\rangle\}. \end{aligned} \tag{3.49}$$

In this basis, the H_7 block reads

$$H_7 = \begin{pmatrix} 0 & 0 & 0 & 0 & -iP_2k_z & 0 & \frac{-i}{\sqrt{2}}P_2(k_x+k_y) \\ & 0 & 0 & 0 & 0 & iP_2k_z & \frac{1}{\sqrt{2}}(-i)P_2(k_y-k_x) \\ & & 0 & 0 & \frac{-i}{\sqrt{2}}P_2(k_x+k_y) & \frac{-i}{\sqrt{2}}P_2(k_y-k_x) & 0 \\ & & & 0 & 0 & 0 & 0 \\ & & & & 0 & 0 & 0 \\ & & & & & 0 & 0 \\ & & & & & & 0 \end{pmatrix}. \quad (3.50)$$

In the basis (3.49), the action of the operator $D(\sigma_v)$ on the spinor is given as

$$D(\sigma_v)\begin{pmatrix} \psi_1(x,y,z) \\ \psi_2(x,y,z) \\ \psi_3(x,y,z) \\ \psi_4(x,y,z) \\ \psi_5(x,y,z) \\ \psi_6(x,y,z) \\ \psi_7(x,y,z) \end{pmatrix} = \begin{pmatrix} \psi_1(y,x,z) \\ -\psi_2(y,x,z) \\ \psi_3(y,x,z) \\ \psi_4(y,x,z) \\ \psi_5(y,x,z) \\ -\psi_6(y,x,z) \\ \psi_7(y,x,z) \end{pmatrix}. \quad (3.51)$$

It follows

$$H_7 D(\sigma_v)\begin{pmatrix} e^{i\mathbf{k}\cdot\mathbf{r}} \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} = H_7 \begin{pmatrix} e^{i(k_x y+k_y x+k_z z)} \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

$$= \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ iP_2k_z \\ 0 \\ \frac{1}{\sqrt{2}}iP_2(k_x+k_y) \end{pmatrix} e^{i(k_x y+k_y x+k_z z)}. \quad (3.52)$$

On the other hand

$$D(\sigma_v)H_7\begin{pmatrix} e^{i\mathbf{k}\cdot\mathbf{r}} \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} = D(\sigma_v)\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ iP_2k_z \\ 0 \\ \frac{1}{\sqrt{2}}iP_2(k_x+k_y) \end{pmatrix} e^{i\mathbf{k}\cdot\mathbf{r}}$$

$$= \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ iP_2k_z \\ 0 \\ \frac{1}{\sqrt{2}}iP_2(k_x+k_y) \end{pmatrix} e^{i(k_xy+k_yx+k_zz)} \tag{3.53}$$

and consequently

$$[D(\sigma_v)H_7 - H_7D(\sigma_v)]\begin{pmatrix} e^{i\mathbf{k}\cdot\mathbf{r}} \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} = 0. \tag{3.54}$$

One can further show that this result is valid also for other basis vectors. This completes the proof that the symmetry of the kinetic part of the 14-band $\mathbf{k}\cdot\mathbf{p}$ Hamiltonian is C_{2v}.

3.7 Plane Wave Representation

Within the plane wave method [5, 7, 16, 26, 39, 79, 81], the envelope functions are expanded as a linear combination of plane waves $a_{\boldsymbol{k}}(\boldsymbol{r}) = e^{i\boldsymbol{k}\cdot\boldsymbol{r}}$

$$\psi_b(\boldsymbol{r}) = \sum_{\boldsymbol{k}} A_{b\boldsymbol{k}}a_{\boldsymbol{k}}(\boldsymbol{r}), \tag{3.55}$$

with the goal of finding the coefficients $A_{b\boldsymbol{k}}$ in the expansion. The index b takes the integer values $b \in \{1, \ldots, N_b\}$, where N_b is the number of bands in the multiband Hamiltonian. The $\boldsymbol{k}$-space is discretized by embedding the dot in a rectangular box

of dimensions $L_x^{(e)}$, $L_y^{(e)}$, and $L_z^{(e)}$ (and volume $\Omega^{(e)} = L_x^{(e)} L_y^{(e)} L_z^{(e)}$) and choosing the $\boldsymbol{k}$-vectors in the form $\boldsymbol{k} = 2\pi(n_x/L_x^{(e)}, n_y/L_y^{(e)}, n_z/L_z^{(e)})$, where n_x, n_y and n_z are integers. A wave vector cutoff is typically made by imposing the conditions $|n_x| \leq m_x^{(e)}$, $|n_y| \leq m_y^{(e)}$, and $|n_z| \leq m_z^{(e)}$ for all bands.

After making the substitution of (3.55) into the Hamiltonian eigenvalue problem one gets

$$\sum_{b\boldsymbol{k}} \mathscr{H}_{ib}(\boldsymbol{q}, \boldsymbol{k}) A_{b\boldsymbol{k}} = E A_{i\boldsymbol{q}}, \tag{3.56}$$

where

$$\mathscr{H}_{ib}(\boldsymbol{q}, \boldsymbol{k}) = \frac{1}{\Omega^{(e)}} \int_{\Omega^{(e)}} \mathrm{d}^3 \boldsymbol{r}\, a_{\boldsymbol{q}}(\boldsymbol{r})^* h_{ib} a_{\boldsymbol{k}}(\boldsymbol{r}). \tag{3.57}$$

Several characteristics have contributed to the popularity of the plane wave method: plane-wave representation of all operators in the envelope Hamiltonian is analytical, strain distribution can be calculated analytically in Fourier space [4], and a relatively small number of plane waves is sufficient for good accuracy.

The reader interested in other numerical methods for solving the eigenvalue problem of multiband Hamiltonian, such as the finite difference method [29, 47, 56, 76] and other wave function expansion methods (where the eigenfunctions of the particle in a cylinder with infinite walls [46, 77] or eigenfunctions of a harmonic oscillator [66] are used as basis set) is referred to relevant literature.

However, there is one serious shortcoming of the plane wave method when its application to single quantum dot structures is concerned: it inherently assumes periodic boundary conditions. In such a way, it artificially introduces translational symmetry of the system. This leads to artificial physical interaction of a quantum dot with its periodically replicated images through: (a) electronic coupling between states of neighboring dots; (b) propagation of strain field of neighboring dots; (c) Coulomb interaction between carrier in the dot with carriers in its artificial images.

Let the length scales where electronic coupling, strain field and Coulomb interactions become negligible respectively be $L^{(e)}$, $L^{(s)}$ and $L^{(c)}$, see Fig. 3.3 for notation. The wave function decays exponentially away from the dot, the strain field has a slower polynomial decay, while the long range Coulomb interaction has the slowest decay. Therefore, the inequality $L^{(e)} < L^{(s)} < L^{(c)}$ holds. In order to eliminate the effects of the interaction with images one would have to choose the embedding box of dimensions $L^{(c)}$ which can be quite large. This leads to the necessity of using a larger number of plane waves to accurately represent the wave function in the quantum dot region, which is undesirable since a large matrix needs to be diagonalized then.

We will show in the next section that the embedding box of the dimensions $L^{(e)}$ can still be used provided that a proper modification in the calculation procedure is performed.

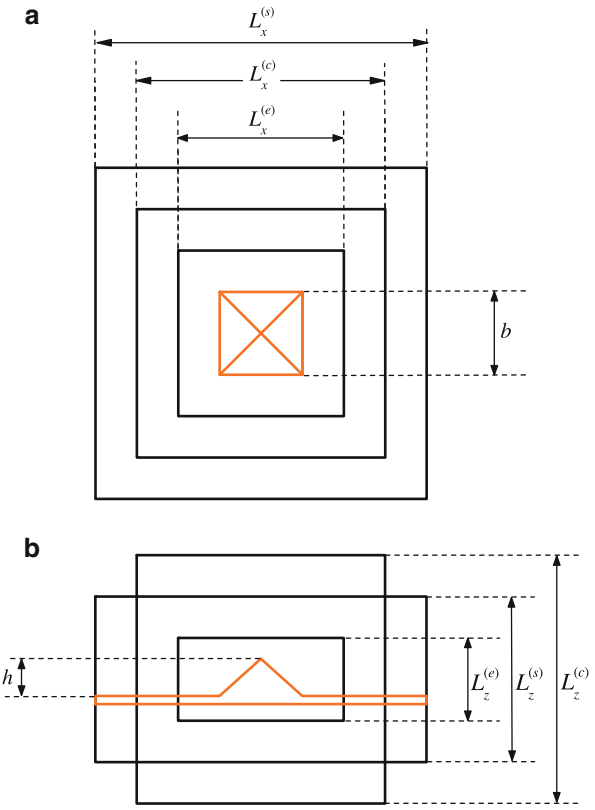


Fig. 3.3 Top view (**a**) and side view (**b**), schematic diagram of an InAs/GaAs quantum dot in the shape of a pyramid, embedded in three "Russian doll" type nested embedding boxes that are used for electronic structure [superscript (e)], strain [superscript (s)] and Coulomb integral calculations [superscript (c)], as described in the main text

3.8 Removal of Artificial Translational Symmetry Effects in Plane Wave Calculations

The essential quantities needed for the description of few particle states (such as excitons and multiexcitons) in quantum dots are the Coulomb integrals. The Coulomb integral among the states i, j, k and l is defined as

$$V_{ijkl} = \sum_{b=1}^{N_b} \sum_{b'=1}^{N_b} \int_{\Omega}^{(\mathrm{e})} \mathrm{d}^3\boldsymbol{r} \int_{\Omega}^{(\mathrm{e})} \mathrm{d}^3\boldsymbol{r}' \psi_b^{(i)}(\boldsymbol{r})^* \psi_b^{(j)}(\boldsymbol{r}) V(|\boldsymbol{r}-\boldsymbol{r}'|) \psi_{b'}^{(k)}(\boldsymbol{r}')^* \psi_{b'}^{(l)}(\boldsymbol{r}') \tag{3.58}$$

where

$$V(u) = \frac{e^2}{4\pi\varepsilon u},$$

with ε being the static dielectric constant. The integral in Eq. (3.58) can be rewritten as

$$V_{ijkl} = \int_{\Omega}^{(e)} \mathrm{d}^3\boldsymbol{r} \int_{\Omega}^{(e)} \mathrm{d}^3\boldsymbol{r}' B_{ij}(\boldsymbol{r}) V(|\boldsymbol{r} - \boldsymbol{r}'|) B_{kl}(\boldsymbol{r}'), \tag{3.59}$$

where

$$B_{ij}(\boldsymbol{r}) = \sum_{b=1}^{N_b} \psi_b^{(i)}(\boldsymbol{r})^* \psi_b^{(j)}(\boldsymbol{r}). \tag{3.60}$$

Next, we define the plane wave expansion of $B_{ij}(\boldsymbol{r})$ as

$$B_{ij}(\boldsymbol{r}) = \sum_{\boldsymbol{q} \in \mathrm{inv}\Omega^{(e)}} B_{ij}(\boldsymbol{q}) e^{i\boldsymbol{q}\cdot\boldsymbol{r}}. \tag{3.61}$$

Replacing the last expression into Eq. (3.59) one obtains

$$V_{ijkl} = \sum_{\boldsymbol{q} \in \mathrm{inv}\Omega^{(e)}} B_{ij}(\boldsymbol{q}) \sum_{\boldsymbol{q}' \in \mathrm{inv}\Omega^{(e)}} B_{kl}(\boldsymbol{q}') \int_{\Omega}^{(e)} \mathrm{d}^3\boldsymbol{r} \int_{\Omega}^{(e)} \mathrm{d}^3\boldsymbol{r}' e^{i\boldsymbol{q}\cdot\boldsymbol{r}} V(|\boldsymbol{r} - \boldsymbol{r}'|) e^{i\boldsymbol{q}'\cdot\boldsymbol{r}'}. \tag{3.62}$$

The $B_{ij}(\boldsymbol{q})$ term can be expressed in terms of the coefficients in the envelope function plane wave expansion as

$$B_{ij}(\boldsymbol{q}) = \sum_{b=1}^{N_b} \sum_{\boldsymbol{q}_1 \in \mathrm{inv}\Omega^{(e)}} A^{(i)*}_{\boldsymbol{q}_1,b} A^{(j)}_{\boldsymbol{q}_1+\boldsymbol{q},b}. \tag{3.63}$$

Next, we introduce an approximation that changes the domain of integration in one of the integrals in Eq. (3.62) from $\Omega^{(e)}$ to the whole space (which is valid when $\Omega^{(e)}$ is large enough) and make the replacement of variables from $\boldsymbol{r}$ and $\boldsymbol{r}'$, to $\boldsymbol{r}$ and $\boldsymbol{u} = \boldsymbol{r} - \boldsymbol{r}'$

$$V^{(\mathrm{a0})}_{ijkl} = \sum_{\boldsymbol{q} \in \mathrm{inv}\Omega^{(e)}} B_{ij}(\boldsymbol{q}) \sum_{\boldsymbol{q}' \in \mathrm{inv}\Omega^{(e)}} B_{kl}(\boldsymbol{q}') \left[\int_{\Omega}^{(e)} \mathrm{d}^3\boldsymbol{r} e^{i\boldsymbol{q}\cdot\boldsymbol{r}} e^{i\boldsymbol{q}'\cdot\boldsymbol{r}} \right] \left[\int \mathrm{d}^3\boldsymbol{u} V(|u|) e^{-i\boldsymbol{q}'\cdot\boldsymbol{u}} \right]. \tag{3.64}$$

Exploiting the relations

$$\int \mathrm{d}^3\boldsymbol{u} e^{-i\boldsymbol{q}'\boldsymbol{u}} V(|u|) = \frac{e^2}{\varepsilon q'^2}$$

and

$$\frac{1}{\Omega^{(e)}}\int_{\Omega}^{(e)} \mathrm{d}^3\boldsymbol{r}\, e^{i(\boldsymbol{q}+\boldsymbol{q}')\boldsymbol{r}} = \delta_{\boldsymbol{q}+\boldsymbol{q}',0}$$

one gets

$$V_{ijkl}^{(\mathrm{a0})} = \Omega^{(e)} \sum_{\substack{\boldsymbol{q}\in \mathrm{inv}\Omega^{(e)} \\ \boldsymbol{q}\neq 0}} B_{ij}(\boldsymbol{q}) B_{kl}(-\boldsymbol{q}) \frac{e^2}{\varepsilon q^2}. \tag{3.65}$$

As already pointed out, $V_{ijkl}^{(\mathrm{a0})}$ is only an approximation to V_{ijkl}. It is therefore very important to understand the nature of error introduced by using Eq. (3.65). One can interpret the initial expression given by Eq. (3.58) as the energy of the electrostatic interaction between the complex charges $B_{ij}(\boldsymbol{r})$ and $B_{kl}(\boldsymbol{r})$, both being located in volume $\Omega^{(e)}$. On the other hand, the expression given by Eq. (3.64) is the energy of the electrostatic interaction between $B_{ij}(\boldsymbol{r})$ located in volume $\Omega^{(e)}$ and $B_{kl}(\boldsymbol{r})$, located in the whole space with periodicity of the box $\Omega^{(e)}$. As a consequence, the error that is introduced by calculating V_{ijkl} using Eq. (3.65) stems from the interactions among the charge $B_{ij}(\boldsymbol{r})$ of a single quantum dot and periodically replicated charges $B_{kl}(\boldsymbol{r})$ of neighboring periodically replicated array of dots.

Now that we understand the nature of error in Eq. (3.65), we can develop a way to systematically correct it. We define the functions $B'_{ij}(\boldsymbol{r})$ equal to $B_{ij}(\boldsymbol{r})$ inside the box $\Omega^{(e)}$ and 0 in the region outside the box $\Omega^{(e)}$ and inside the box $\Omega^{(c)}$ (sides $L_x^{(c)}$, $L_y^{(c)}$, $L_z^{(c)}$) that is larger than $\Omega^{(e)}$. Fourier transform of $B'_{ij}(\boldsymbol{r})$ on the box $\Omega^{(c)}$ is then defined as

$$B'_{ij}(\boldsymbol{r}) = \sum_{\boldsymbol{Q}\in \mathrm{inv}\Omega^{(c)}} B'_{ij}(\boldsymbol{Q}) e^{i\boldsymbol{Q}\cdot\boldsymbol{r}}. \tag{3.66}$$

The relation between the Fourier transforms of B_{ij} and B'_{ij} is given as

$$\begin{aligned} B'_{ij}(\boldsymbol{Q}) = \frac{1}{\Omega^{(c)}} \sum_{\boldsymbol{q}\in \mathrm{inv}\Omega^{(e)}} & B_{ij}(\boldsymbol{q}) I_0(-L_x^{(e)}/2, L_x^{(e)}/2, Q_x - q_x) \times \\ & \times I_0(-L_y^{(e)}/2, L_y^{(e)}/2, Q_y - q_y) I_0(-L_z^{(e)}/2, L_z^{(e)}/2, Q_z - q_z). \end{aligned} \tag{3.67}$$

where

$$I_0(a,b,k) = \int_a^b \mathrm{d}x e^{ikx} = \begin{cases} \frac{e^{ikb}-e^{ika}}{ik} & k \neq 0 \\ b-a & k = 0 \end{cases} \tag{3.68}$$

Eq. (3.59) can be recast as

$$V_{ijkl} = \int_{\Omega}^{(c)} \mathrm{d}^3\boldsymbol{r} \int_{\Omega}^{(c)} \mathrm{d}^3\boldsymbol{r}' B'_{ij}(\boldsymbol{r}) V(|\boldsymbol{r}-\boldsymbol{r}'|) B'_{kl}(\boldsymbol{r}'). \tag{3.69}$$

Applying the same procedure as in the derivation of Eq. (3.65), one obtains

$$V_{ijkl}^{(\mathrm{a1})} = \Omega^{(c)} \sum_{\substack{\boldsymbol{q}\in \mathrm{inv}\Omega^{(c)} \\ \boldsymbol{q}\neq 0}} B'_{ij}(\boldsymbol{q}) B'_{kl}(-\boldsymbol{q}) \frac{e^2}{\varepsilon q^2}. \tag{3.70}$$

Eq. (3.70) differs from Eq. (3.65) since a different approximation was used in its derivation. The integral over the region $\Omega^{(c)}$ was replaced by the integral over the whole space, which is a better approximation than the replacement of the integral over the region $\Omega^{(e)}$ in the derivation of Eq. (3.65), since the region $\Omega^{(c)}$ is larger than $\Omega^{(e)}$. Consequently, the error introduced by calculating V_{ijkl} using Eq. (3.70) now originates from the interactions among the charge $B_{ij}(\boldsymbol{r})$ of a single quantum dot and periodically replicated charges $B_{kl}(\boldsymbol{r})$ with periodicity defined by the box $\Omega^{(c)}$ rather than $\Omega^{(e)}$. Therefore, Eq. (3.70) can be systematically improved by an increase in $\Omega^{(c)}$, without increasing the number of plane waves needed to represent the wave functions.

We note that a similar philosophy can be used to remove the effects of strain field introduced by neighboring boxes. The reader interested in details of this procedure is referred to [89].

Another way to correct the error introduced by Coulomb interactions is to perform a multipole expansion of the difference between the Coulomb integral and its approximation given by Eq. (3.70), i.e. to perform the Makov-Payne correction. Such a procedure has been previously applied in ab initio [43] and empirical pseudopotential[24] calculations of aperiodic systems. The calculated value of the Coulomb integral can then be corrected by adding the first few terms (monopole, dipole and quadrupole) in the multipole expansion as

$$V_{ijkl}^{\mathrm{final}} = V_{ijkl}^{(\mathrm{a1})} - \frac{e^2}{4\pi\varepsilon}\left[q_{ij}q_{kl}a_{\mathrm{mad}} + \frac{4\pi}{3\Omega^{(c)}}\boldsymbol{d}_{ij}\cdot\boldsymbol{d}_{kl} - \frac{2\pi}{3\Omega^{(c)}}(q_{ij}Q_{kl}+q_{kl}Q_{ij})\right], \tag{3.71}$$

where

$$q_{ij} = \int_{\Omega^{(c)}} B'_{ij}(\boldsymbol{r})\mathrm{d}^3\boldsymbol{r} = \delta_{ij}, \tag{3.72}$$

$$\boldsymbol{d}_{ij} = \int_{\Omega^{(c)}} B'_{ij}(\boldsymbol{r})\boldsymbol{r}\,\mathrm{d}^3\boldsymbol{r}, \tag{3.73}$$

$$Q_{ij} = \int_{\Omega^{(c)}} B'_{ij}(\boldsymbol{r})r^2\mathrm{d}^3\boldsymbol{r}, \tag{3.74}$$

are the monopole, dipole and quadrupole terms respectively. The Madelung term a_{mad} is defined in terms of the Ewald sums and the self-interaction correction term as

$$a_{\mathrm{mad}} = \sum_{\substack{\boldsymbol{R}\in \mathrm{dir}\Omega^{(c)} \\ \boldsymbol{R}\neq 0}} \frac{\mathrm{erfc}(R\eta)}{R} + \frac{4\pi}{\Omega^{(c)}} \sum_{\substack{\boldsymbol{k}\in \mathrm{inv}\Omega^{(c)} \\ \boldsymbol{k}\neq 0}} \frac{\exp(-k^2/4\eta^2)}{k^2} - 2\frac{\eta}{\sqrt{\pi}} - \frac{\pi}{\eta^2 \Omega^{(c)}}.$$

The Ewald parameter η controls the rate of convergence of the sums. A reliable value that provides fast convergence is $\eta = \pi/\sqrt{L_x^{(c)} L_y^{(c)}}$.

Next, we illustrate the described methods for the correction of the Coulomb integral calculation by analyzing the dependence of the Coulomb integrals on $\Omega^{(c)}$ (dimensions $(L_x^{(c)}, L_y^{(c)}, L_z^{(c)})$). A set of calculations was done where $(m_x^{(c)}, m_y^{(c)}, m_z^{(c)})$ (used to determine the wave vector cutoff in Eq. (3.71)) was kept at a sufficiently large value of $(35, 35, 50)$ and the box dimensions were changed. Several direct Coulomb integrals $J_{ab} = V_{aabb}$ are shown in Fig. 3.4.

As seen from Fig. 3.4, the direct integrals without corrections calculated from Eq. (3.70) converge very slowly towards the numerically exact value obtained by performing the six dimensional integration in real space. For example, the box with dimension $(L_x^{(c)}, L_y^{(c)}, L_z^{(c)}) = (100, 100, 100)$ nm is sufficient only for precision of the order of 20 % and the box $(L_x^{(c)}, L_y^{(c)}, L_z^{(c)}) = (200, 200, 200)$ nm gives a precision of the order of 10 %. Following this procedure, numerically exact value can be approached within ∼3 % by reasonable systematical enlargement of the $V^{(c)}$ box. Further improvement in the accuracy of J_{ab} appears to be very difficult.

The results obtained by adding the monopole correction only in (3.71), are sufficient for the degree of accuracy one is usually interested in. The box $(L_x^{(c)}, L_y^{(c)}, L_z^{(c)}) = (60, 60, 60)$ nm is then sufficient for the precision of 1 % or better for the direct Coulomb integrals. The results with the three corrections involved are nearly indistinguishable from the numerically exact values for the direct Coulomb integrals. The box $(L_x^{(c)}, L_y^{(c)}, L_z^{(c)}) = (60, 60, 60)$ nm then already gives the precision better than 0.1 % for the values of direct integrals.

We illustrate the use of the methods developed by performing a full configuration interaction calculation of exciton and biexciton states. The rank of the configuration interaction matrix is

$$N_r^{\mathrm{CI}} = \binom{N_e}{n_e} \cdot \binom{N_h}{n_h} \tag{3.75}$$

where lowest N_e and topmost N_h states in conduction and valence band respectively form a basis of single-particle states for configuration interaction, while n_e and

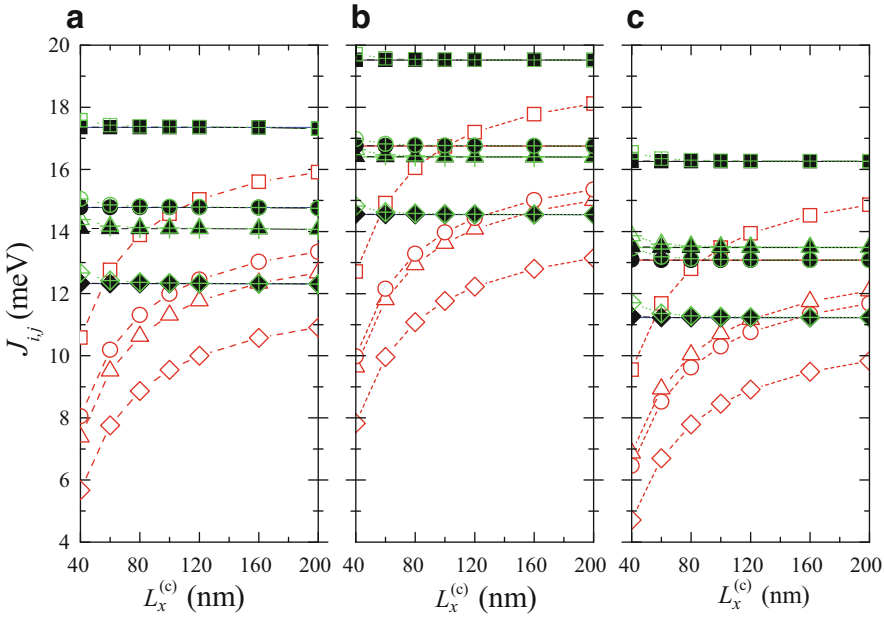


Fig. 3.4 The dependence of the values of Coulomb integrals in dot (**a**) of square-based truncated pyramidal shape with the bottom base width $b_b = 22\,\text{nm}$, the top base width $b_t = 5.5\,\text{nm}$, and the height $h = 4.425\,\text{nm}$, quantum dot (**b**) of truncated conical shape with the bottom base radius $R_b = 11\,\text{nm}$, the top base radius $R_t = 2.75\,\text{nm}$ and the height of $h = 4.425\,\text{nm}$, and quantum dot (**c**) in the shape of a lens with the radius $R = 15\,\text{nm}$ and the height $h = 4.425\,\text{nm}$ on the size of the embedding box dimension $L_x^{(c)}$. The calculation was done with $(m_x^{(c)}, m_y^{(c)}, m_z^{(c)}) = (35, 35, 50)$, $(L_x^{(c)}, L_y^{(c)}, L_z^{(c)}) = (L_x^{(c)}, L_x^{(c)}, L_x^{(c)})$. $J_{e0,h0}$ (*squares*), $J_{e1,h0}$ (*triangles*), $J_{e0,h1}$ (*circles*), and $J_{e1,h1}$ (*diamonds*). The results without corrections (*open symbols*), the results with the monopole correction only (*open symbols with cross*), and the results with monopole, dipole and quadrupole correction (*solid symbols*) are shown. The results obtained by real space integration are indistinguishable from the results obtained by including the three corrections

n_h are the number of electrons and holes that form the many-body complex. The number of Coulomb integrals needed to construct the configuration Hamiltonian is $(N_e + N_h)^4$. By exploiting the relations $V_{jilk} = V^*_{ijkl}$ and $V_{lkji} = V^*_{ijkl}$, the whole problem can be reduced to the calculation of $[(N_e + N_h)(N_e + N_h + 1)/2]^2$ integrals. Additionally, symmetry considerations imply that only Coulomb integrals V_{ijkl} whose wave functions satisfy the conservation of the total quasi-angular momentum:

$$\{m_j + m_l \equiv m_i + m_k \pmod 2\} \tag{3.76}$$

are nonzero. This additionally reduces the number of integrals that need to be calculated by a factor of 2. In our case, all 354,025 Coulomb integrals among the states from the set including first $N_e = 14$ electron and first $N_h = 20$ states were calculated. The calculation of such a big number of Coulomb integrals is performed efficiently by exploiting the following two facts: (1) $B_{ij}(\boldsymbol{q})$, (3.63), that enters the expression for Coulomb integrals, (3.70), via (3.67) and the expressions for multipole corrections (3.72), (3.73), (3.74), needs to be calculated just $(N_e + N_h)(N_e + N_h + 1)/2$ times; (2) summation in (3.67) over vectors $\boldsymbol{q}$, which should be in principle done in the domain $|n_{\boldsymbol{r}}| \leq 2m_{\boldsymbol{r}}^{(e)}$, can be done in the reduced domain $|n_{\boldsymbol{r}}| \leq m_{\boldsymbol{r}}^{(e)}$, since the relative error introduced in V_{ijkl} by this truncation is $< 10^{-5}$. One should also note that when a particular set of Coulomb integrals is calculated, it can be used for configuration interaction calculations with different values of n_e and n_h, without the need of recalculating the integrals.

In order to determine the number of single-particle states sufficient for the use in configuration interaction expansion, two sets of calculations were performed. In the first set, the number of hole states was set to $N_h = 20$ and N_e was varied. In the second set, the number of electron states was set to $N_e = 14$ and N_h was varied. The results are shown in Fig. 3.5. One can estimate from the results presented in Fig. 3.5 that $N_e = 10$ is sufficient for convergence of exciton ground state of the order of 0.2 meV and biexciton ground state of the order of 0.5 meV. For the same degree of precision, a larger number of hole states $N_h = 14$ is needed, as a consequence of smaller energy difference among hole single particle states.

3.9 Symmetries of Single Particle States in Quantum Dots

In this section, we show first how one can exploit the symmetry to block diagonalize the Hamiltonian matrix, which leads to a more efficient solution of its eigenvalue problem. Then we show how one can identify the symmetry group of the Hamiltonian and analyze the symmetries of various Hamiltonians that describe the same physical system.

Symmetry-based block diagonalization of the quantum dot **k·p** Hamiltonian matrix was performed for the first time in [88] and [87], for the cases of fourfold and sixfold symmetry respectively. The same approach can be extended to M-fold symmetry. Block diagonalization is achieved by representing the Hamiltonian in the so called symmetry adapted basis. If we denote the plane wave basis state where the envelope function of band b is equal to $a_{\boldsymbol{k}}(\boldsymbol{r})$ and the other envelope functions are zero as $|\boldsymbol{k}, b\rangle$, This is done by state where the envelope function of band b is equal to $a_{\boldsymbol{k}}(\boldsymbol{r})$ and the other envelope functions are zero) to the basis of the states characterized by a given value of the z-component of the total quasi-angular momentum m_f. In the case of M-fold symmetry, this basis is composed of the

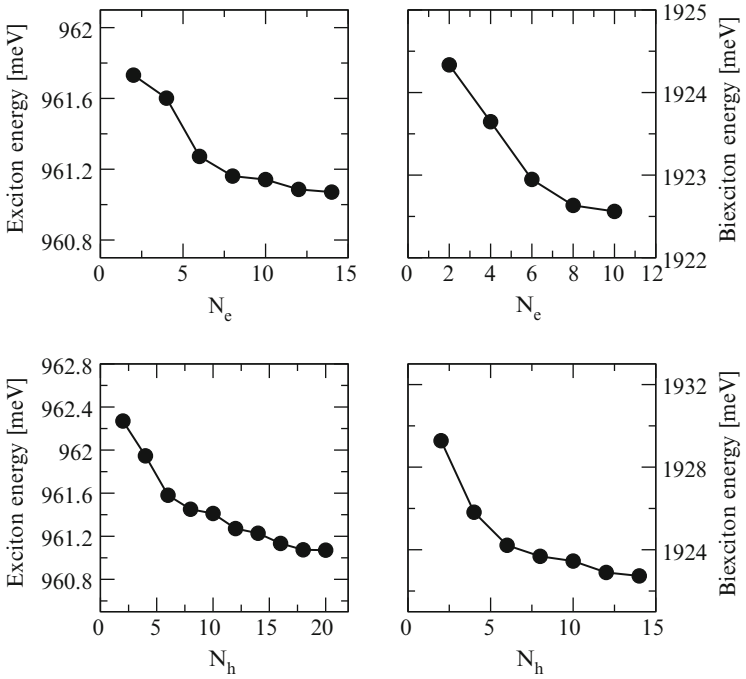


Fig. 3.5 The dependence of exciton and biexciton energy in quantum dot (a) from Fig. 3.4 on the number of electron N_e states (when $N_h = 20$) and hole N_h (when $N_e = 14$) states used for configuration interaction expansion

following elements the vectors of the symmetry adapted basis in the case of the system with M-fold symmetry are given as

$$|A_{m_f}, \boldsymbol{k}, b\rangle = \frac{1}{\sqrt{M}} \sum_{l=0}^{M-1} e^{il\phi(m_f - J_z(b))} |R_{l\phi}\boldsymbol{k}, b\rangle \tag{3.77}$$

with $\boldsymbol{k}$-vectors satisfying $k_x^2 + k_y^2 > 0$ and $0 \leq k_y < \tan(\phi)k_x$,

$$|A_{m_f}, \boldsymbol{k}, b\rangle = |\boldsymbol{k}, b\rangle \tag{3.78}$$

with $\boldsymbol{k}$-vectors satisfying $k_x = k_y = 0$ and the band b satisfying $(J_z(b) - m_f)\text{mod } M = 0$. In previous equations $\phi = 2\pi/M$, $J_z(b)$ is the z-component of the total quasi-angular momentum of the Bloch function of band b, $R_{l\phi}\boldsymbol{k} = \boldsymbol{k}'$ is the vector obtained by rotation of the vector $\boldsymbol{k}$ by an angle $l\phi$ around the z-axis

$$\begin{aligned} k_x' + ik_y' &= e^{il\phi}(k_x + ik_y), \\ k_z' &= k_z. \end{aligned} \tag{3.79}$$

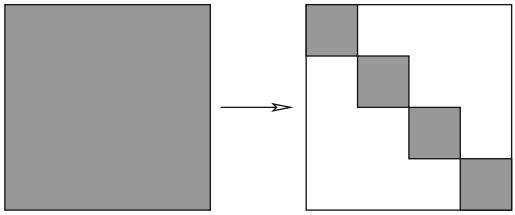


Fig. 3.6 The scheme of the Hamiltonian matrix in the plane wave basis (*left*) and in the symmetry adapted basis (*right*)

The quasi-angular momentum of the basis state m_f takes the values from the interval $[-(M-1)/2, (M-1)/2]$ with a step of 1. In the symmetry adapted basis, the Hamiltonian matrix elements are nonzero between basis states with the same m_f only. As a consequence, the Hamiltonian matrix is block diagonal (see Fig. 3.6); it is composed of M smaller matrices of approximately equal size. Since diagonalization cost of the Hamiltonian matrix is proportional to N_r^3 (where N_r is the rank of the matrix), the total cost of the diagonalization is then $\propto M \times (N_r/M)^3$, which is M^2 times faster than if symmetry were not used.

The symmetry of the Hamiltonian is not necessarily equal to the symmetry of the system. In fact, there has been a belief that $\mathbf{k} \cdot \mathbf{p}$ Hamiltonians predict a higher symmetry than the true symmetry of the system. In what follows, we analyze the square based pyramidal quantum dot with base width to height ratio $b/h = 2$ modeled with different $\mathbf{k} \cdot \mathbf{p}$ Hamiltonians. We consider the following Hamiltonians:

(a) The 8-band $\mathbf{k} \cdot \mathbf{p}$ Hamiltonian consisting of kinetic part only [without spin-orbit interaction and strain].
(b) The 8-band $\mathbf{k} \cdot \mathbf{p}$ Hamiltonian consisting of kinetic part with spin-orbit interaction taken into account [but without strain].
(c) The 8-band $\mathbf{k} \cdot \mathbf{p}$ Hamiltonian consisting of kinetic part with interface band-mixing effects taken into account [but without spin-orbit interaction and strain].
(d) The standard 8-band $\mathbf{k} \cdot \mathbf{p}$ Hamiltonian consisting of kinetic part with spin-orbit interaction and strain, as well as the strain-induced piezoelectric potential. It was assumed that piezoelectric polarization depends linearly on strain.
(e) The 8-band $\mathbf{k} \cdot \mathbf{p}$ Hamiltonian consisting of kinetic part with spin-orbit interaction and strain, as well as the strain-induced piezoelectric potential and the interface Hamiltonian.
(f) The 14-band $\mathbf{k} \cdot \mathbf{p}$ Hamiltonian consisting of the kinetic part only [without spin-orbit interaction and strain].

In what follows, we will refer to each of these Hamiltonians as models (a)-(f). We show that the inclusion of additional bands in the Hamiltonian or the inclusion of interface effects lead to the true symmetry of the system.

We start our considerations with model (a). The symmetry group of such a model applied to a pyramidal square-based quantum dot is the C_{4v} group, as demonstrated

Table 3.2 Energies (in eV) of top six hole energy levels and bottom four electron levels, for a square-based pyramidal InAs/GaAs quantum dot with base width $b = 100$ Å, and base to height ratio $b/h = 2$ calculated using different models

State	(a)	(b)	(c)	(d)	(e)	(f)
e3	1.08163	1.06674	1.08878	1.30684	1.30772	1.03852
e2	0.99336	0.96795	1.00170	1.25044	1.25570	0.97301
e1	0.99336	0.96724	1.00074	1.23439	1.23943	0.97298
e0	0.84346	0.81808	0.85115	1.12013	1.12543	0.83509
h0	−0.06722	−0.03427	−0.06475	−0.05230	−0.05052	−0.06512
h1	−0.06722	−0.03680	−0.06698	−0.06827	−0.06699	−0.06517
h2	−0.07389	−0.03765	−0.07248	−0.07840	−0.07843	−0.07263
h3	−0.07883	−0.04244	−0.07708	−0.09115	−0.09046	−0.07700
h4	−0.08518	−0.04582	−0.08119	−0.10517	−0.10565	−0.08124
h5	−0.08518	−0.04614	−0.08360	−0.10888	−0.10859	−0.08270

The letters in the first row in the table specify the model used in the calculation

in Sect. 3.5. We discuss the signatures of symmetry of the model (a) in the energy level structure [given in Table 3.2(a)] and the shape of the wavefunctions [presented in Fig. 3.7(a)]. Several pairs of energy levels [(e_1,e_2), (h_0,h_1) and (h_4,h_5)] are degenerate as a consequence of the fact that the C_{4v} group has a two dimensional representation E (the notation of [10]). The states that transform according to this representation therefore come in pairs and are degenerate. The wave function probability density isosurfaces also reflect the high symmetry of the system—they all have a perfectly C_{4v} symmetric shape.

Next, we include spin-orbit interaction in model (a) and get model (b). The eigenstates of the Hamiltonian of such a model transform according to a representation of the double C_{4v} group which is a direct product of the representation of the single C_{4v} group and the representation $\mathscr{D}_{1/2}$ according to which the spin functions transform ([10], p. 142). When the representation obtained from the direct product is reducible, the inclusion of spin-orbit interaction leads to the removal of degeneracy of energy levels. Indeed, the product $E \times \mathscr{D}_{1/2}$ is equal to $E_1' + E_2'$ (the notation of [10]). While, it is well understood that the effect of spin-orbit interaction on the states in the valence band is rather strong, we would like to point out a less known fact that the spin-orbit interaction also causes the splitting of the e_1 and e_2 states [shown in Table 3.2(b)]. The existence of this splitting was established in [88] for pyramidal quantum dots and later on analyzed again in [20] for lens-shaped quantum dots. It is important to note here that symmetry reduction is not the cause of this energy level splitting effect. Indeed, one can see in Fig. 3.7b that the probability density isosurfaces exhibit a perfect C_{4v} symmetry.

To understand the effect of interface terms in the Hamiltonian on the symmetry of the model, we analyze model (c). One can derive analytically that the [001] interface term leads to the reduction of symmetry from C_{4v} to C_{2v} (see Sect. 3.4). Similar

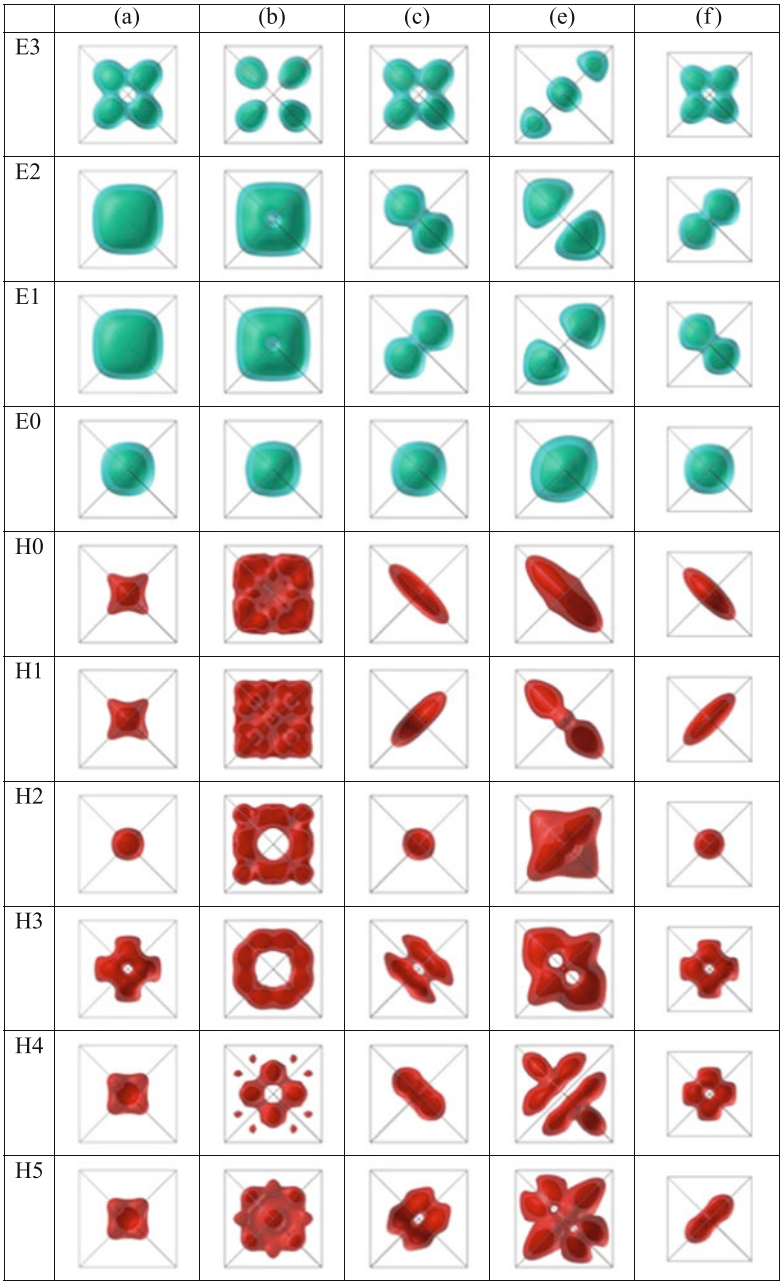


Fig. 3.7 The wavefunctions squared for top six hole states and bottom five electron states for a square-based pyramidal InAs/GaAs quantum dot with base width $b = 100$ Å, and base to height ratio $b/h = 2$ calculated using different models. The letters (a)–(c), (e)–(f) specify the model used in the calculation. The isosurfaces are plotted at 25 % (transparent) and 75 % of the maximal charge density

derivations also show that the joint effect of other four interfaces leads to the same symmetry reduction. Since the C_{2v} group has one dimensional irreducible representations only, no double degenerate eigenstates can be present. Consequently, the effect of interfaces leads to splitting of degenerate states by typically 1–3 meV, as shown in Table 3.2(c). The shapes of the wave function moduli isosurfaces also reduce their symmetry from C_{4v} [Fig. 3.7(a)] to C_{2v} [Fig. 3.7(c)].

Model (f) that also includes the second conduction band is considered next. Analytical derivations (Sect. 3.5) show that the inclusion of the additional bands also leads to reduction of symmetry to C_{2v}. The terms that contain the P_2 element which couples the top of the valence band, Γ_{5v}, [that originates from the p bonding states (denoted as p_b) of atoms in the bulk] with the second conduction band, Γ_{5c}, [that originates from the p antibonding states (denoted as p_a) of atoms in the bulk] are the only terms that prevent the C_{4v} symmetry. Therefore the P_2 element can be identified as the symmetry breaking term in model (f). This term introduces the splittings, Table 3.2(f), which are less pronounced than these of the interface terms—for example the splitting of e_1 and e_2 is less than 0.1 meV. The effect of this symmetry breaking on the wave functions is generally similar [see Fig. 3.7(f)] as in the case of interface-induced symmetry breaking.

Analytical derivations and the numerical results presented therefore indicate that the inclusion of bands beyond the standard 8 bands, as well as the inclusion of interface effects within the $\mathbf{k}\cdot\mathbf{p}$ approach both lead to a correct description of the symmetry of the system. As a consequence, a widespread belief that $\mathbf{k}\cdot\mathbf{p}$ Hamiltonians are not capable to capture the correct symmetry of the system appears not to be correct.

It has been previously well known [88] that the piezoelectric effect also reduces the symmetry from C_{4v} to C_{2v}. The results shown in Table 3.2 (d), (e) indicate that splitting of e_1 and e_2 states induced by the piezoelectric effect is stronger than the splittings induced by other effects. One should also stress that in the absence of piezoelectric effect, the strain would also give rise to symmetry reduction if it were modeled using the Valence force field model [9, 57, 75, 76, 78]

3.10 Symmetries of Exciton States in Quantum Dots

In this section, we discuss the signatures of symmetry in the exciton spectra and in the optical properties of excitons. We identify the changes in the spectrum when spin-orbit interaction is taken into account and when symmetry group of the model is lowered from C_{4v} to C_{2v}.

We start by considering model (a) of the previous section. In such a model, the single particle states transform according to single valued irreducible representations (IRs) of the C_{4v} group. We find that the h_0 state transforms according to two dimensional representation E, while the e_0 state transforms as A_1 representation

(the notation of [19]). Therefore, h_0 is fourfold degenerate (including the twofold spin degeneracy), while e_0 is twofold degenerate.

Exciton states transform according to single values IRs of the symmetry group [84]. The IRs of the exciton states in the ground exciton manifold can be obtained from the direct product of the IR of the e_0 state and complex conjugated IR of the h_0 state. In model (a) this gives $A_1 \times E = E$. Therefore, the orbital part of the exciton state transform as E. In addition, the spins of an electron and a hole that form an exciton can form either the singlet or the triplet state. As a consequence, the ground eight-dimensional manifold consists of the doubly degenerate ground state (that originates from the E symmetry of the orbital part and the singlet of the spin part) and sixfold degenerate excited state (stemming from the E symmetry of the orbital part and the triplet of the spin part). An E exciton is allowed to emit xy-polarized radiation, while it is not allowed to emit z-polarized radiation. On the other hand, due to conservation of spin in the optical transition, the singlet state is dark, while in the triplet two states are bright and one is dark. This implies that the twofold-degenerate ground exciton state is dark while the sixfold degenerate excited state consists of four bright and two dark states.

Next, we discuss the changes in the spectrum when spin-orbit interaction is included as in model (b) of the previous section. In model (b) the single particle states transform according to double valued IRs of the C_{4v} group. We find that the h_0 state transforms as $\overline{E}_2$, while the e_0 state transforms as $\overline{E}_1$. Both of them are twofold degenerate. The IRs of states in the four dimensional exciton manifold are then obtained from $\overline{E}_1 \times \overline{E}_2 = E + B_1 + B_2$. Therefore, the ground exciton manifold consists of a doubly degenerate E exciton and non-degenerate B_1 and B_2 excitons. The E exciton is allowed to emit xy-polarized radiation, while B_1 and B_2 excitons are dark. Among higher excited exciton states, the states that transforms as B_1, B_2 and A_2 are non-degenerate and dark, the states that transform as A_1 are non-degenerate and emit z-polarized radiation, while the states that transform as E are double-degenerate and emit xy-polarized radiation.

The effects that arise when the symmetry is lowered from C_{4v} to C_{2v}, as for example in models (c) and (f) (that exclude spin-orbit interaction) in the previous section, are discussed next. The subduction of the IR E of the C_{4v} group to the IRs of the C_{2v} group yields $E \rightarrow B_1 + B_2$. Therefore, the h_0 state that transformed as E in model (a), transforms now as either B_1 or B_2. The e_0 state still transforms as A_1. As a consequence, both e_0 and h_0 are two fold degenerate (including the twofold spin degeneracy). The ground exciton manifold is now four dimensional. The orbital part of the exciton states in ground exciton manifold transforms as B_1 or B_2 (depending whether h_0 transforms as B_1 or B_2). As a consequence, the ground exciton manifold consists of the non-degenerate ground state (that originates from the singlet of the spin part and is dark) and the threefold degenerate excited states (that originate from the triplet of the spin part). Two of these excited states are bright and can emit xy-polarized radiation, while one is dark.

Finally we discuss the symmetry lowering effects in models (c) or (d) (that include spin-orbit interaction). The subductions of the IRs $\overline{E}_1$ and $\overline{E}_2$ of the C_{4v}

group to the C_{2v} group yield $\overline{E}_1 \rightarrow \overline{E}$ and $\overline{E}_2 \rightarrow \overline{E}$. As a consequence, both h_0 and e_0 now transform as $\overline{E}$. The IRs of the states in the four dimensional ground exciton manifold then stem from $\overline{E} \times \overline{E} = A_1 + A_2 + B_1 + B_2$. The B_1 and B_2 states originate from the E state in the model with C_{4v} symmetry. Therefore, lowering of symmetry from C_{4v} to C_{2v} splits the E exciton into two non-degenerate B_1 and B_2 excitons. B_1 and B_2 inherit the brightness from the E exciton. The A_2 state originates from the B_2 state in the model with C_{4v} symmetry. It remains non-degenerate and dark. On the other hand, the A_1 state originates from the dark B_1 state in the model with C_{4v} symmetry. It remains non-degenerate but the A_1 state is bright and is allowed to emit z-polarized radiation. However, one should note that the bright state that originates from the dark state of the higher symmetry group is typically only weakly bright as we have verified by numerical calculations of the dipole matrix elements that correspond to such states. Among higher excited exciton states, the states that transforms as A_2 are dark, the states that transform as A_1 emit z-polarized radiation, while the states that transform as B_1 or B_2 emit xy-polarized radiation.

3.11 Conclusion

In this chapter, we have demonstrated the importance of understanding the symmetry of the $\mathbf{k} \cdot \mathbf{p}$ Hamiltonians used in electronic structure calculations. On this route, the interface term which is rarely considered was derived first, as it is essential for capturing the proper symmetry of the system. The plane wave method introduces an artificial translational symmetry accompanied by artificial Coulomb interaction between the carrier in the dot and its periodic replicas. Artifacts of this interaction can be removed by a careful modification of the procedure for calculation of Coulomb integrals. We show how one can identify the symmetry group of a certain $\mathbf{k} \cdot \mathbf{p}$ Hamiltonian. An example of a square-based pyramidal quantum dot is then used to show how the symmetry changes with the change in the level of sophistication of the model. The standard 8-band $\mathbf{k} \cdot \mathbf{p}$ Hamiltonian exhibits an artificially high C_{4v} symmetry. However, both the inclusion of the effect of interfaces and the inclusion of additional bands in the model lead to correct C_{2v} symmetry. Once the symmetry of the Hamiltonian is understood, it can be used to choose the basis in which the Hamiltonian is block diagonal and consequently largely reduce the computational effort.

Acknowledgements Stanko Tomić* wishes to thank to the STFC e-Science, UK, for providing the computational resources, and the Royal Society, London, Research Grant "High Performance Computing in Modelling of Innovative High Efficiency Photovoltaic Devices". Nenad Vukmirović was supported by European Community FP7 Marie Curie Career Integration Grant (ELECTROMAT), Serbian Ministry of Education, Science and Technological Development (project ON171017) and FP7 projects PRACE-2IP, PRACE-3IP, HP-SEE, and EGI-InSPIRE.

Appendix

The values of the relevant material parameters appearing in the kp-Hamiltonians are given in Table 3.3.

Table 3.3 Relevant material parameters of binary compound semiconductors GaAs, InAs, and AlAs

	GaAs	InAs	AlAs
a_0 [Å]	5.6503	6.0553	5.661
α [meV/K]	0.5405	0.276	0.885
β [K]	204	93	530
$E_{g0} = E(\Gamma_{1c}) - E(\Gamma_{5v})$ [eV]	1.518	0.405	3.099
$E_{g1} = E(\Gamma_{5c}) - E(\Gamma_{5v})$ [eV]	4.488	4.38	4.54
$E_{g2} = E(\Gamma_{5v}) - E(\Gamma_{1v})$ [eV]	12.50	12.64	11.95
E_{P0} [eV]	25.7	21.846	21.1
E_{P1} [eV]	0.19	0.03	0.16
E_{P2} [eV]	14.79	19.0	16.8
E_{P3} [eV]	2.3	0.6	0.1
E_{P4} [eV]	0.2	2.55	0.0 (n/a)
$E_{v,av}$ [eV]	−6.920	−6.747	−7.49
m_c^*	0.0667	0.02226	0.15
$\Delta_{\mathrm{so}}(p_a)$[eV]	0.340	0.380	0.280
$\Delta_{\mathrm{so}}(p_b)$[eV]	0.170	0.190	0.150
Δ_{cf}[eV]	0.085	0.085	0.085
c_{11} [GPa]	118.8	83.3	125.0
c_{12} [GPa]	53.8	45.3	53.4
c_{44} [GPa]	59.4	39.6	54.2
a_c [eV]	−8.013	−5.08	−5.64
a_v [eV]	0.220	1.00	2.47
b_{ax} [eV]	−1.824	−1.800	−2.3
d_{ax} [eV]	−5.062	−3.600	−3.4
e_{14} [C m^{-2}]	0.160	0.045	0.225
$\gamma_1^L, \gamma_2^L, \gamma_3^L$	7.10, 2.02, 2.91	19.67, 8.40, 9.30	3.76, 0.82, 1.42
ϵ_r	13.18	14.6	10.1

a_0 are the lattice constants, α and β are the Varshni parameters that describe the temperature dependence of the band gap (a temperature of 4K was assumed in all calculations), E_{gi} are the band gaps, E_{Pi} are the energies related to interband matrix elements of the velocity operator P_i as $E_{Pi} = 2m_0 P_i^2/\hbar^2$, $E_{v,av}$ is the average valence band edge energy at the Γ point, m_c^* is the conduction band effective mass. $\Delta_{\mathrm{so}}(p_a)$ is the spin-orbit splitting in the second conduction band, $\Delta_{\mathrm{so}}(p_b)$ is the spin-orbit splitting in the valence band and Δ_{cf} the crystal field splitting. c_{ij} are the elastic constants. a_c, a_v, b_{ax}, d_{ax} are the deformation potentials, e_{14} is the piezoelectric constant. $\gamma_1^L, \gamma_2^L, \gamma_3^L$ are the Luttinger parameters in the 6-band model. ϵ_r is the static dielectric constant

References

1. N. Akopian, N.H. Lindner, E. Poem, Y. Berlatzky, J. Avron, D. Gershoni, B.D. Gerardot, P.M. Petroff, Entangled photon pairs from semiconductor quantum dots. Phys. Rev. Lett. **96**, 130–501 (2006)
2. A.P. Alivisatos, W.W. Gu, C. Larabell, Quantum dots as cellular probes. Ann. Rev. Biomed. Eng. **7**, 55–76 (2005)
3. A.D. Andreev, Conference on In-Plane Semiconductor Lasers - From Ultraviolet to Mid-Infrared II, San Jose, CA, Jan 26–28, 1998, pp 151–161
4. A.D. Andreev, J.R. Downes, D.A. Faux, E.P. O'Reilly, Strain distributions in quantum dots of arbitrary shape. J. Appl. Phys. **86**, 297–305 (1999)
5. A.D. Andreev, E.P. O'Reilly, Theory of the electronic structure of GaN/AlN hexagonal quantum dots. Phys. Rev. B **62**, 15,851–15,870 (2000)
6. C.G. Bailey, D.V. Forbes, R.P. Raffaelle, S.M. Hubbard, Near 1 v open circuit voltage InAs/GaAs quantum dot solar cells. Appl. Phys. Lett. **98**(16), 163105 (2011)
7. G.A. Baraff, D. Gershoni, Eigenfunction-expansion method for solving the quantum-wire problem: Formulation. Phys. Rev. B **43**, 4011–4022 (1991)
8. P. Bhattacharya, S. Ghosh, A.D. Stiff-Roberts, Quantum dot opto-electronic devices. Ann. Rev. Mat. Res. **34**, 1–40 (2004)
9. D. Bimberg, M. Grundmann, N.N. Ledentsov, *Quantum dot heterostructures* (John Wiley, Chichester, 1999)
10. G.L. Bir, G.E. Pikus, *Symmetry and strain-induced effects in semiconductors* (Wiley, New York, 1974)
11. S. Blokhin, A. Sakharov, A. Nadtochy, A. Pauysov, M. Maximov, N. Ledentsov, A. Kovsh, S. Mikhrin, V. Lantratov, S. Mintairov, N. Kaluzhniy, M. Shvarts, AlGaAs/GaAs photovoltaic cells with an array of InGaAs qds. Semiconductors **43**, 514–518 (2009)
12. C. Bulutay, Pseudopotential-based full zone $\mathbf{k}\cdot\mathbf{p}$ technique for indirect bandgap semiconductors: Si, Ge, diamond and Sic. Turk. J. Phys. **30**, 287–294 (2006)
13. M.G. Burt, The justification for applying the effective-mass approximation to microstructures. J. Phys.: Condens. Matter **4**, 6651–6690 (1992)
14. M. Cardona, N.E. Christensen, G. Fasol, Relativistic band structure and spin-orbit splitting of zinc-blende-type semiconductors. Phys. Rev. B **38**(3), 1806–1827 (1988)
15. M. Cardona, F.H. Pollak, Energy-band structure of germanium and silicon: The $\mathbf{k}\cdot\mathbf{p}$ method. Phys. Rev. **142**(2), 530–543 (1966)
16. M.A. Cusack, P.R. Briddon, M. Jaros, Electronic structure of InAs/GaAs self-assembled quantum dots. Phys. Rev. B **54**, R2300–R2303 (1996)
17. J.P. Cuypers, W. van Haeringen, Coupling between γ- and x-type envelope functions at GaAs/Al(Ga)As interfaces. Phys. Rev. B **48**(15), 11,469–11,472 (1993)
18. K. Eichkorn, R. Ahlrichs, Cadmium selenide semiconductor nanocrystals: a theoretical study. Chem. Phys. Lett. **288**, 235–242 (1998)
19. J.P. Elliott, P.G. Dawber, *Symmetry in Physics* (Macmillan, London, 1979)
20. J. Even, F. Doré, C. Cornet, L. Pedesseau, Semianalytical model for simulation of electronic properties of narrow-gap strained semiconductor quantum nanostructures. Phys. Rev. B **77**, 085305 (2008)
21. B.A. Foreman, Analytical envelope-function theory of interface band mixing. Phys. Rev. Lett. **81**(2), 425–428 (1998)
22. B.A. Foreman, Analytical envelope-function theory of interface band mixing. Phys. Rev. Lett. **81**, 425–428 (1998)
23. N. Fraj, I. Saïdi, S.B. Radhia, K. Boujdaria, Band structures of AlAs, gap, and SiGe alloys: A 30 $\mathbf{k}\cdot\mathbf{p}$ model. J. Appl. Phys. **102**(5), 053703 (2007)
24. A. Franceschetti, H. Fu, L.W. Wang, A. Zunger, Many-body pseudopotential theory of excitons in inp and cdse quantum dots. Phys. Rev. B **60**, 1819–1829 (1999)

25. C.S. Garoufalis, A.D. Zdetsis, S. Grimme, High level ab initio calculations of the optical gap of small silicon quantum dots. Phys. Rev. Lett. **87**, 276,402 (2001)
26. D. Gershoni, C.H. Henry, G.A. Baraff, Calculating the optical properties of multidimensional heterostructures: Application to the modeling of quaternary quantum well lasers. IEEE J. Quantum Electron. **29**, 2433–2450 (1993)
27. C. Hermann, C. Weisbuch, $\mathbf{k}\cdot\mathbf{p}$ perturbation theory in iii-v compounds and alloys: a reexamination. Phys. Rev. B **15**(2), 823–833 (1977)
28. S.M. Hubbard, C.D. Cress, C.G. Bailey, R.P. Raffaelle, S.G. Bailey, D.M. Wilt, Effect of strain compensation on quantum dot enhanced GaAs solar cells. Appl. Phys. Lett. **92**(12), 123512 (2008)
29. H. Jiang, J. Singh, Strain distribution and electronic spectra of InAs/GaAs self-assembled dots: An eight-band study. Phys. Rev. B **56**, 4696–4701 (1997)
30. J. Kim, L.W. Wang, A. Zunger, Comparison of the electronic structure of InAs/GaAs pyramidal quantum dots with different facet orientations. Phys. Rev. B **57**(16), R9408–R9411 (1998)
31. N. Kirstaedter, N.N. Ledentsov, M. Grundmann, D. Bimberg, V.M. Ustinov, S.S. Ruvimov, M.V. Maximov, P.S. Kopev, Z.I. Alferov, U. Richter, P. Werner, U. Gosele, J. Heydenreich, Low-threshold, large T_0 injection-laser emission from (InGa)As quantum dots. Electron. Lett. **30**, 1416–1417 (1994)
32. P.C. Klipstein, Operator ordering and interface-band mixing in the Kane-like Hamiltonian of lattice-matched semiconductor superlattices with abrupt interfaces. Phys. Rev. B **81**(23), 235–314 (2010)
33. M.E. Kurdi, G. Fishman, S. Sauvage, P. Boucaud, Band structure and optical gain of tensile-strained germanium based on a 30 band $\mathbf{k}\cdot\mathbf{p}$ formalism. J. Appl. Phys. **107**(1), 013710 (2010)
34. O.L. Lazarenkova, A.A. Balandin, Miniband formation in a quantum dot crystal. J. Appl. Phys. **89**(10), 5509–5515 (2001)
35. S. Lee, J. Kim, L. Jönsson, J.W. Wilkins, G.W. Bryant, G. Klimeck, Many-body levels of optically excited and multiply charged InAs nanocrystals modeled by semiempirical tight binding. Phys. Rev. B **66**, 235,307 (2002)
36. J. Li, L.W. Wang, Energy levels of isoelectronic impurities by large scale LDA calculations. Phys. Rev. B **67**, 033,102 (2003)
37. J. Li, L.W. Wang, First principles calculations of ZnS:Te energy levels. Phys. Rev. B **67**, 205,319 (2003)
38. J. Li, L.W. Wang, Band-structure-corrected local density approximation study of semiconductor quantum dots and wires. Phys. Rev. B **72**, 125325 (2005)
39. S.S. Li, J.B. Xia, Z.L. Yuan, Z.Y. Xu, W. Ge, X.R. Wang, Y. Wang, J. Wang, L.L. Chang, Effective-mass theory for InAs/GaAs strained coupled quantum dots. Phys. Rev. B **54**, 11,575–11,581 (1996)
40. J.M. Luttinger, Quantum theory of cyclotron resonance in semiconductors: general theory. Phys. Rev. **102**, 1030–1041 (1956)
41. J.M. Luttinger, W. Kohn, Motion of electrons and holes in perturbed periodic fields. Phys. Rev. **97**, 869–883 (1955)
42. S. Maimon, E. Finkman, G. Bahir, S.E. Schacham, J.M. Garcia, P.M. Petroff, Intersublevel transitions in InAs/GaAs quantum dots infrared photodetectors. Appl. Phys. Lett. **73**, 2003–2005 (1998)
43. G. Makov, M.C. Payne, Periodic boundary conditions in *ab initio* calculations. Phys. Rev. B **51**, 4014–4022 (1995)
44. H. Mayer, U. Rössler, Spin splitting and anisotropy of cyclotron resonance in the conduction band of GaAs. Phys. Rev. B **44**(16), 9048–9051 (1991)
45. P. Michler, A. Kiraz, C. Becher, W.V. Schoenfeld, P.M. Petroff, L.D. Zhang, E. Hu, A. Imamoglu, A quantum dot single-photon turnstile device. Science **290**, 2282 (2000)
46. V. Mlinar, M. Tadić, F.M. Peeters, Hole and exciton energy levels in InP/$\mathrm{In}_x\mathrm{Ga}_{1-x}$P quantum dot molecules: Influence of geometry and magnetic field dependence. Phys. Rev. B **73**, 235,336 (2006)

47. T. Nakaoka, T. Saito, J. Tatebayashi, Y. Arakawa, Size, shape and strain dependence of the g factor in self-assembled In(Ga)As quantum dots. Phys. Rev. B **70**, 235,337 (2004)
48. A. Nozik, Quantum dot solar cells. Physica E **14**(1-2), 115–120 (2002)
49. R. Oshima, A. Takata, Y. Okada, Strain-compensated InAs/GaNAs quantum dots for use in high-efficiency solar cells. Appl. Phys. Lett. **93**(8), 083111 (2008)
50. D. Pan, E. Towe, S. Kennerly, Normal-incidence intersubband (In,Ga)As/GaAs quantum dot infrared photodetectors. Appl. Phys. Lett. **73**, 1937–1939 (1998)
51. D. Pan, Y.P. Zeng, M.Y. Kong, J. Wu, Y.Q. Zhu, C.H. Zhang, J.M. Li, C.Y. Wang, Normal incident infrared absorption from InGaAs/GaAs quantum dot superlattice. Electron. Lett. **32**, 1726–1727 (1996)
52. P.M. Petroff, S.P. DenBaars, MBE and MOCVD growth and Properties of Self-Assembling Quantum Dot Arrays in III-V Semiconductor Structures. Superlattices and Microstructures **15**, 15–21 (1994)
53. P. Pfeffer, W. Zawadzki, Conduction electrons in GaAs: Five-level $\mathbf{k}\cdot\mathbf{p}$ theory and polaron effects. Phys. Rev. B **41**(3), 1561–1576 (1990)
54. P. Pfeffer, W. Zawadzki, Five-level $\mathbf{k}\cdot\mathbf{p}$ model for the conduction and valence bands of gaas and inp. Phys. Rev. B **53**(19), 12,813–12,828 (1996)
55. C.R. Pidgeon, R.N. Brown, Interband magneto-absorption and Faraday rotation in InSb. Phys. Rev. **146**, 575–583 (1966)
56. C. Pryor, Eight-band calculations of strained InAs/GaAs quantum dots compared with one-, four-, and six-band approximations. Phys. Rev. B **57**, 7190–7195 (1998)
57. C. Pryor, J. Kim, L.W. Wang, A.J. Williamson, A. Zunger, Comparison of two methods for describing the strain profiles in quantum dots. J. Appl. Phys. **83**, 2548–2554 (1998)
58. A. Puzder, A.J. Williamson, J.C. Grossman, G. Galli, Computational studies of the optical emission of silicon nanocrystals. J. Am. Chem. Soc. **125**, 2786–2791 (2003)
59. S.B. Radhia, K. Boujdaria, S. Ridene, H. Bouchriha, G. Fishman, Band structures of GaAs, InAs, and Ge: A 24-$\mathbf{k}\cdot\mathbf{p}$ model. J. Appl. Phys. **94**(9), 5726–5731 (2003)
60. S.B. Radhia, S. Ridene, K. Boujdaria, H. Bouchriha, G. Fishman, Band structures of Ge and InAs: A 20 $\mathbf{k}\cdot\mathbf{p}$ model. J. Appl. Phys. **92**(8), 4422–4430 (2002)
61. J.P. Reithmaier, G. Sek, A. Loffler, C. Hofmann, S. Kuhn, S. Reitzenstein, L.V. Keldysh, V.D. Kulakovskii, T.L. Reinecke, A. Forchel, Strong coupling in a single quantum dot-semiconductor microcavity. Nature (London) **432**, 197–200 (2004)
62. S. Richard, F. Aniel, G. Fishman, Energy-band structure of Ge, Si, and GaAs: A thirty-band $\mathbf{k}\cdot\mathbf{p}$ method. Phys. Rev. B **70**(23), 235,204 (2004)
63. D. Rideau, M. Feraille, L. Ciampolini, M. Minondo, C. Tavernier, H. Jaouen, A. Ghetti, Strained Si, Ge, and $Si_{1-x}Ge_x$ alloys modeled with a first-principles-optimized full-zone $\mathbf{k}\cdot\mathbf{p}$ method. Phys. Rev. B **74**(19), 195–208 (2006)
64. U. Rössler, Nonparabolicity and warping in the conduction band of GaAs. Solid State Commun. **49**(10), 943–947 (1984)
65. U. Rössler, J. Kainz, Microscopic interface asymmetry and spin-splitting of electron subbands in semiconductor quantum structures. Solid State Commun. **121**(6-7), 313–316 (2002)
66. M. Roy, P.A. Maksym, Efficient method for calculating electronic states in self-assembled quantum dots. Phys. Rev. B **68**, 235,308 (2003)
67. K.A. Sablon, J.W. Little, V. Mitin, A. Sergeev, N. Vagidov, K. Reinhardt, Strong enhancement of solar cell efficiency due to quantum dots with built-in charge. Nano Letters **11**(6), 2311–2317 (2011)
68. I. Saïdi, S.B. Radhia, K. Boujdaria, Band structures of gaas, inas, and inp: A 34 $\mathbf{k}\cdot\mathbf{p}$ model. J. Appl. Phys. **104**(2), 023706 (2008)
69. Saito, T., Schulman, J.N., Arakawa, Y.: Strain-energy distribution and electronic structure of InAs pyramidal quantum dots with uncovered surfaces: Tight-binding analysis. Phys. Rev. B **57**, 13,016–13,019 (1998)
70. R. Santoprete, B. Koiller, R.B. Capaz, P. Kratzer, Q.K.K. Liu, M. Scheffler, Tight-binding study of the influence of the strain on the electronic properties of InAs/GaAs quantum dots. Phys. Rev. B **68**, 235,311 (2003)

71. C. Santori, M. Pelton, G.S. Solomon, Y. Dale, Y. Yamamoto, Triggered single photons from a quantum dot. Phys. Rev. Lett. **86**, 1502–1505 (2001)
72. R.D. Schaller, V.I. Klimov, High Efficiency Carrier Multiplication in PbSe Nanocrystals: Implications for Solar Energy Conversion. Phys. Rev. Lett. **92**, 186,601 (2004)
73. W. Sheng, S.J. Cheng, P. Hawrylak, Multiband theory of multi-exciton complexes in self-assembled quantum dots. Phys. Rev. B **71**, 035,316 (2005)
74. R.M. Stevenson, R.J. Young, P. Atkinson, K. Cooper, D.A. Ritchie, A.J. Shields, A semiconductor source of triggered entangled photon pairs. Nature (London) **439**, 179–182 (2006)
75. O. Stier, Electronic and optical properties of quantum dots and wires. Wissenschaft & Technik Verlag, Berlin (2000)
76. O. Stier, M. Grundmann, D. Bimberg, Electronic and optical properties of strained quantum dots modeled by 8-band $k \cdot p$ theory. Phys. Rev. B **59**, 5688–5701 (1999)
77. M. Tadić, F.M. Peeters, K.L. Janssens, Effect of isotropic versus anisotropic elasticity on the electronic structure of cylindrical InP/$In_{0.49}Ga_{0.51}P$ self-assembled quantum dots. Phys. Rev. B **65**, 165,333 (2002)
78. M. Tadić, F.M. Peeters, K.L. Janssens, M. Korkusiński, P. Hawrylak, Strain and band edges in single and coupled cylindrical InAs/GaAs and InP/InGaP self-assembled quantum dots. J. Appl. Phys. **92**, 5819–5829 (2002)
79. S. Tomić, Electronic structure of $In_yGa_{1-y}As_{1-x}N_x$/GaAs(N) quantum dots by ten-band $\mathbf{k} \cdot \mathbf{p}$ theory. Phys. Rev. B **73**, 125,348 (2006)
80. S. Tomić, T.S. Jones, N.M. Harrison, Absorption characteristics of a quantum dot array induced intermediate band: Implications for solar cell design. Appl. Phys. Lett. **93**(26), 263105 (2008)
81. S. Tomić, A.G. Sunderland, I.J. Bush, Parallel multi-band $\mathbf{k} \cdot \mathbf{p}$ code for electronic structure of zinc blend semiconductor quantum dots. J. Mater. Chem. **16**, 1963–1972 (2006)
82. S. Tomić, Intermediate-band solar cells: Influence of band formation on dynamical processes in InAs/GaAs quantum dot arrays. Phys. Rev. B **82**, 195,321 (2010)
83. S. Tomić, N. Vukmirović, Symmetry reduction in multiband Hamiltonians for semiconductor quantum dots: The role of interfaces and higher energy bands. J. Appl. Phys. **110**(5), 053710 (2011)
84. P. Tronc, V.P. Smirnov, K.S. Zhuravlev, Symmetry of electron states and optical transitions in GaN/AlN hexagonal quantum dots. Phys. Status Solidi B **241**, 2938–2947 (2004)
85. I. Vasiliev, S. Öğüt, J.R. Chelikowsky, First-principles density-functional calculations for optical spectra of clusters and nanocrystals. Phys. Rev. B **65**(11), 115,416 (2002)
86. P. von Allmen, Conduction subbands in a GaAs/$Al_xGa_{1-x}As$ quantum well: Comparing different $\mathbf{k} \cdot \mathbf{p}$ models. Phys. Rev. B **46**(23), 15,382–15,386 (1992)
87. N. Vukmirović, Z. Ikonić, D. Indjin, P. Harrison, Symmetry-based calculation of single-particle states and intraband absorption in hexagonal GaN/AlN quantum dot superlattices. J. Phys.: Condens. Matter **18**, 6249–6262 (2006)
88. N. Vukmirović, D. Indjin, V.D. Jovanović, Z. Ikonić, P. Harrison, Symmetry of $\mathbf{k}\cdot\mathbf{p}$ Hamiltonian in pyramidal InAs/GaAs quantum dots: Application to the calculation of electronic structure. Phys. Rev. B **72**, 075,356 (2005)
89. N. Vukmirović, S. Tomić, Plane wave methodology for single quantum dot electronic structure calculations. J. Appl. Phys. **103**, 103718 (2008)
90. L.W. Wang, Large-scale local-density-approximation band gap-corrected GaAsN calculations. Appl. Phys. Lett. **78**, 1565–1567 (2001)
91. L.W. Wang, Charge-density patching method for unconventional semiconductor binary systems. Phys. Rev. Lett. **88**(25), 256,402 (2002)
92. L.W. Wang, Generating charge densities of fullerenes. Phys. Rev. B **65**, 153,410 (2002)
93. L.W. Wang, J. Kim, A. Zunger, Electronic structures of [110]-faceted self-assembled pyramidal InAs/GaAs quantum dots. Phys. Rev. B **59**(8), 5678–5687 (1999)
94. L.W. Wang, A. Zunger, Linear combination of bulk bands method for large-scale electronic structure calculations on strained nanostructures. Phys. Rev. B **59**, 15,806–15,818 (1999)
95. A.J. Williamson, A. Zunger, InAs quantum dots: Predicted electronic structure of free-standing versus GaAs-embedded structures. Phys. Rev. B **59**(24), 15,819–15,824 (1999)

96. J.B. Xia, γ-x mixing effect in GaAs/AlAs superlattices and heterojunctions, Phys. Rev. B **41**(5), 3117–3122 (1990)
97. Z. Zhou, R.A. Friesner, L. Brus, Electronic structure of 1 to 2 nm diameter silicon core/shell nanocrystals: surface chemistry, optical spectra, charge transfer, and doping. J. Am. Chem. Soc. **125**, 15,599–15,607 (2003)
98. O. Zitouni, K. Boujdaria, H. Bouchriha, Band parameters for GaAs and Si in the 24-$\mathbf{k}\cdot\mathbf{p}$ model, Semiconductor Science and Technology **20**(9), 908 (2005).
99. A. Zrenner, E. Beham, S. Stufler, F. Findeis, M. Bichler, G. Abstreiter, Coherent properties of a two-level system based on a quantum-dot photodiode. Nature (London) **418**, 612–614 (2002)

Part II
Numerical Methods

Chapter 4
Finite Elements for k·p Multiband Envelope Equations

Ratko G. Veprek and Sebastian Steiger

Abstract This chapter applies the finite element method to the $\mathbf{k} \cdot \mathbf{p}$ equations describing electronic states in semiconductor nanostructures. It highlights advantages over other discretization methods and discusses the crucial ingredients in order to obtain accurate results. One particular issue, the appearance of unphysical or spurious solutions, is demonstrated to emerge from an inconsistency of the continuum equation system, not the discretization, and two causes are identified whose correct treatment leads to the elimination of such solutions.

4.1 Introduction

Quantum-mechanical properties of semiconductor nanostructures lie at the heart of many well-known devices from lasers to LEDs, transistors to RTDs. Usually the choice of band structure model is decisive when performing simulations of such properties. The $\mathbf{k} \cdot \mathbf{p}$ multiband envelope function approach is the method of choice for electronic structure calculations where a compromise between computational efficiency and physical accuracy is sought. The relatively straightforward interpretation of obtained wavefunctions and optical transition rates yields additional advantages over atomistic models [39]. It is natural that in order to fully leverage the power of the approach, attention should be paid not only to the theoretical model but also to the numerical methods by which the $\mathbf{k} \cdot \mathbf{p}$ equations, usually written as real-space partial differential equations, are solved.

R.G. Veprek (✉)
Condensed Matter Theory Group, School of Physics, University of Sydney, NSW 2006, Australia
e-mail: ratko.veprek@gmail.com

S. Steiger
Network for Computational Nanotechnology, Purdue University, West Lafayette, IN 47907, USA
e-mail: sebi.steiger@gmail.com

M. Ehrhardt and T. Koprucki (eds.), *Multi-Band Effective Mass Approximations*,
Lecture Notes in Computational Science and Engineering 94,
DOI 10.1007/978-3-319-01427-2_4,

There is a vast number of different methods available to numerically solve partial differential equations. One approach is the finite element method (FEM) [33]. While FEM is the prominent workhorse within aerospace and structural engineering and has widespread application in electro-magnetics and finance, it has attracted comparatively little attention in the solution of multiband Schrödinger equations. This is surprising at first glance, as the advantages of using finite elements are manifold:

- *Representation*—By using unstructured meshes and convex shaped elements, complicated geometries can be accurately represented.
- *Interfaces*—Within finite elements, boundary conditions at material interfaces do not need additional treatment. Such boundary conditions are referred to as being *natural* and are fulfilled automatically.
- *Mesh Refinement*—Meshes can be tailored to be coarse within less important regions and dense where the solution must be precisely known. Using a-priori and a-posteriori error estimates, meshes can be refined at the appropriate places using well-founded criteria.
- *Sparsity*—Contrary to plane-wave expansions, the equation systems resulting from the real-space FEM discretization are sparse. This makes them less expensive to solve and suitable for distributed computational environments.
- *Higher-Order Approximations*—Usual solutions to partial differential equations are non-linear and therefore the convergence of the approximation of the solution benefits from the usage of higher-order interpolation polynomials. Increasing the polynomial order in FEM is straightforward.

One of the reasons for the rare application of FEM is that most physics and nanoelectronics curricula do not contain a comprehensive education in numerical solution methods. Finite-difference or finite-volume discretizations on structured grids are more intuitive and straightforward to implement. Another reason is the possibility of strong spurious solutions (see Sect. 4.5). A clear understanding of their origin is a prerequisite for the successful application of FEM to the $\mathbf{k} \cdot \mathbf{p}$ envelope equations. Therefore the purpose of this chapter is twofold:

- To provide just enough FEM theory and illustrate its application to the $\mathbf{k} \cdot \mathbf{p}$ equations such that an aspiring reader can attempt an implementation.
- To convey the authors' view on the topic of spurious solutions, which anybody solving $\mathbf{k} \cdot \mathbf{p}$ equations is likely to encounter. It is our belief that there is nothing mysterious about their occurrence; yet the origin as well as the way of eliminating them are not widely known.

The chapter is organized as follows. In Sect. 4.2, the reader is familiarized with the basic concepts of finite elements. Section 4.3 gives a brief overview over common $\mathbf{k} \cdot \mathbf{p}$ models and discusses the important issue of operator ordering in nanostructures. Then a FEM discretization of the $\mathbf{k} \cdot \mathbf{p}$ equations is presented in Sect. 4.4 which is applicable to systems of any dimensionality. Section 4.5 discusses criteria for the fulfillment of a basic mathematical property of the $\mathbf{k} \cdot \mathbf{p}$ equations

that should be retained, ellipticity, following the arguments elaborated in [44, 47]. Lastly, Sect. 4.6 gives an outline of the FEM treatment of interface charges arising from spontaneous and piezo-electric polarization.

The theory covered within this chapter has been implemented by the authors in the solver *tdkp*, see http://www.tdkpaqua.net.

4.2 Basic Principles of Finite Elements

The finite-element method is based on a variational formulation of the differential equation to solve. While the theory of finite elements is rich and extensive, only rough ideas and conventions essential for the understanding of this chapter are presented. The reader interested in more profound introductions is referred to [4,51]. The *hp*-FEM theory employed here is covered extensively in [38].

The second-order partial differential operator we are concerned with has the form

$$\mathscr{L}[u] = \sum_{i \le j} \left(\partial_i c_{ij} \partial_j + \partial_j c_{ji} \partial_i + c_{ij}^R \partial_i \partial_j + \partial_i \partial_j c_{ij}^L \right) u + \sum_i \left(d_i^L \partial_i + \partial_i d_i^R \right) u + eu \tag{4.1}$$

on a domain $\Omega \subset \mathbb{R}^n$ with boundary $\partial\Omega$, operating on a function u. Here the indices i, j run over spatial dimensions and c, d and e are position-dependent coefficients. It will be shown later that the order in which the coefficients and the differential operators are placed is of crucial importance. This issue is known as *operator ordering*.

The objective is usually either to find a solution u satisfying

$$\mathscr{L}[u] = g \tag{4.2}$$

for a given function g or to find the eigensolutions (u_i, λ_i) such that

$$\mathscr{L}[u_i] = \lambda_i u_i, \tag{4.3}$$

always with respect to some boundary condition for u defined on $\partial\Omega$. Usually such boundary conditions are either of the *Dirichlet* $(u(\partial\Omega) = u_0)$ or *Neumann* $\left(\frac{\partial u}{\partial n}|_{\partial\Omega} = n_0\right)$ type. The linear equation type (4.2) is encountered in e.g. Poisson's equation or load-strain problems, and the eigenvalue problem (4.3) is found in the **k · p** equations.

Let now u be the (exact and undiscretized) solution of such a partial differential equation with $u \in H(\Omega)$, where $H(\Omega)$ is a suitable space of continuous functions. In a numerical solution procedure we seek an approximation of the solution u. Using finite elements, this is achieved by solving the equation on a distinct, finite-dimensional subset V_M of the space $H(\Omega)$, $V_M \subset H(\Omega)$, which should converge

to $H(\Omega)$ for $M \to \infty$. For this purpose, the domain Ω is partitioned into disjunct convex shaped elements $\mathscr{M}_i$, the finite elements, such that

$$\bigcup_i \mathscr{M}_i = \Omega, \quad \mathscr{M}_i \cap \mathscr{M}_j = \delta_{ij}\mathscr{M}_i + \partial\mathscr{M}_{ij}, \tag{4.4}$$

where $\partial\mathscr{M}_{ij}$ denotes the lower-dimensional boundary shared between the elements (points in 1D, edges in 2D and faces in 3D). This partition is denoted as *mesh*. Related to the mesh is the basis which spans the subset V_M of the function-space $H(\Omega)$. The basis functions are referred to as *shape functions* $N_i(x)$ and fulfil the interpolation requirement of

$$\sum_{i=1}^{M} N_i(x) = 1 \quad \forall x \in \Omega. \tag{4.5}$$

The basis functions usually have only support on a few adjacent elements, which leads to the desired property that the resulting linear matrix equation is sparse. There are many possible forms of basis functions, featuring higher-order interpolation polynomials, interpolation functions for curved boundaries or axially symmetric problems. A common approach is to choose a *nodal* basis based on Lagrange interpolation polynomials by defining M nodes x_j and demanding in addition to (4.5) that

$$N_i(x_j) = \delta_{ij}. \tag{4.6}$$

For linear interpolation polynomials, the nodes are given by the vertices of an element. For higher-order polynomials, additional nodes are required which are commonly placed on the element's edges.

For the nodal basis, let u_j denote the solution at the respective node x_j:

$$u_j = u(x_j). \tag{4.7}$$

The final equation is then solved only for the values u_j and interpolated in between by the basis functions. As a result, the approximate finite element solution $\tilde{u}(x)$ in terms of the shape functions is given as

$$\tilde{u}(x) = \sum_{j=1}^{M} u_j N_j(x). \tag{4.8}$$

How to obtain the equations for the coefficients u_j from here? This is achieved by transforming (4.2) or (4.3) into the *weak form* by multiplying each side of the equation with a *test function* $v \in H(\Omega)$ and integrating over Ω:

$$\int_\Omega \mathscr{L}[u]v\,\mathrm{d}x = \int_\Omega g v\,\mathrm{d}x \quad \forall\, v \in H(\Omega), \tag{4.9a}$$

$$\int_\Omega \mathscr{L}[u]v\,\mathrm{d}x = \lambda \int_\Omega uv\,\mathrm{d}x \quad \forall\, v \in H(\Omega). \tag{4.9b}$$

Here (4.9a) and (4.9b) refer to the problem types (4.2) and (4.3), respectively. Restricting v to the same function space as the solution u is referred to as the *Ritz-Galerkin* approach. It is clear that for the solution u, (4.9a) or (4.9b) are fulfilled for every $v \in H(\Omega)$. We now insert the approximation (4.8) and request that the resulting equation for the approximation only be fulfilled for all functions in the restricted subset $v \in V_M$ used for the approximation. This subset is given by the shape functions $\{N_i(x)\}_{i=1}^{M} = V_M$. Inserting every shape function $N_i(x)$ into (4.9a) or (4.9b) leads to M equations for the M coefficients b_j. Therefore we require $\forall\, N_i(x) \in V_M$,

$$\int_\Omega \mathscr{L}[\tilde{u}]N_i(x)\,\mathrm{d}x = \sum_{j=1}^{M} u_j \int_\Omega \mathscr{L}[N_j]N_i(x)\,\mathrm{d}x = \int_\Omega g N_i(x)\,\mathrm{d}x \tag{4.10a}$$

$$\int_\Omega \mathscr{L}[\tilde{u}]N_i(x)\,\mathrm{d}x = \lambda \sum_{j=1}^{M} u_j \int_\Omega \mathscr{L}[N_j]N_i(x)\,\mathrm{d}x = \lambda \sum_{j=1}^{M} u_j \int_\Omega N_i(x)N_j(x)\,\mathrm{d}x. \tag{4.10b}$$

The corresponding integrals can be evaluated numerically or analytically. As a result, one obtains algebraic equations for the coefficients u_j which read for (4.9a) and (4.9b)

$$(4.9a) \rightarrow \mathbf{Au} = \mathbf{g}, \tag{4.11a}$$

$$(4.9b) \rightarrow \mathbf{Au} = \lambda \mathbf{Mu}. \tag{4.11b}$$

Here $A_{ij} \equiv \int_\Omega \mathscr{L}[N_j]N_i(x)\mathrm{d}x$ is called the *stiffness matrix* and $M_{ij} \equiv \int_\Omega N_i(x)N_j(x)\mathrm{d}x$ is the *mass matrix*. One notes that discretizing the linear equation (4.2) results in the linear matrix equation (4.11a), whereas from the eigenvalue equation (4.3) one obtains the *generalized* eigenvalue problem (GEVP) in Eq. (4.11b). Eigenvectors $\mathbf{y}_i$ of the GEVP are not orthogonal but **M**-orthogonal, which corresponds to the orthogonality of the corresponding eigenfunctions within the space V_M:

$$\mathbf{y}_i^* \mathbf{M} \mathbf{y}_j = \sum_{k=1}^{M}\sum_{l=1}^{M} b_{i,k}^* b_{j,l} \int_\Omega N_k(x)N_l(x)\mathrm{d}x = \delta_{ij}. \tag{4.12}$$

The incorporation of boundary conditions will be discussed in the actual example of the $\mathbf{k}\cdot\mathbf{p}$ equations.

4.3 Nanostructure k · p Equations and Operator Ordering

4.3.1 k · p *Equations in Nanostructures*

In standard **k** · **p** theory, the Hamiltonian for a bulk crystal is written as a matrix which is second-order in the wavevector **k**:

$$\mathbf{H}(\mathbf{k}) = \sum_{i,j} \bar{\mathbf{H}}_{ij}^{(2)} k_i k_j + \sum_i \bar{\mathbf{H}}_i^{(1)} k_i + \bar{\mathbf{H}}^{(0)}. \tag{4.13}$$

In the bulk case the dimension of **H** corresponds to the number of bands which are explicitly calculated. In a nanostructure, the **k** · **p** Hamiltonian becomes a set of coupled partial differential eigenvalue equations up to second order for the vector of envelope functions $\mathbf{f}(x; k_t)$, given by

$$\begin{aligned} E\mathbf{f}(x;k_t) = &-\sum_i \sum_j \partial_i \mathbf{H}_{ij}^{(2)}(x;k_t)\partial_j \mathbf{f}(x;k_t) \\ &+ \sum_i \mathbf{H}_i^{(1L)}(x;k_t)\partial_i \mathbf{f}(x;k_t) + \sum_i \partial_i \mathbf{H}_i^{(1R)}(x;k_t)\mathbf{f}(x;k_t) + \mathbf{H}^{(0)}(x;k_t)\mathbf{f}(x;k_t). \end{aligned} \tag{4.14}$$

The indices i, j run over the quantized directions x of the system, and the equation is parametrized in terms of the transverse wavevector k_t for the directions in which translational symmetry is preserved. The particle is represented by the *envelope function vector* $\mathbf{f}(x; k_t)$ which describes at every position how the bulk Bloch states are mixed due to the symmetry breaking induced by the nanostructure.

We limit our analysis to the case of *closed* Dirichlet-type boundary conditions:

$$\mathbf{f}(x;k_t) = 0 \quad \forall x \in \partial\Omega. \tag{4.15}$$

Other boundary conditions, such as open boundary conditions arising in transport problems or by the inclusion of perfectly matched layers, can also be treated within FEM [33].

The details of the particular **k** · **p** model such as the dimensionality of the nanostructure, operator ordering, the considered number of bands, the chosen basis, the material and the crystal type determine the form of the parametric **k** · **p** matrices $\mathbf{H}_{\cdot}^{(\cdot)}(x; k_t)$. In a system of incomplete quantization, e.g. a quantum wire along the [001] crystal direction (which we take to be the z-axis), the second- and first-order terms related to the transverse direction in (4.13) are in effect added to the first- and zero-order terms in (4.14):

$$\partial_i \mathbf{H}_i^{(1L)} = \partial_i \left(\bar{\mathbf{H}}_i^{(1L)} - \bar{\mathbf{H}}_{iz}^{(2)} i k_z \right), \tag{4.16a}$$

$$\mathbf{H}_i^{(1R)}\partial_i = \left(\bar{\mathbf{H}}_i^{(1R)} - ik_z\bar{\mathbf{H}}_{zi}^{(2)}\right)\partial_i, \tag{4.16b}$$

$$\mathbf{H}^{(0)} = \bar{\mathbf{H}}^{(0)} + k_z\bar{\mathbf{H}}_{zz}^{(2)}k_z + ik_z\bar{\mathbf{H}}_z^{(1L)} + \bar{\mathbf{H}}_z^{(1R)}ik_z. \tag{4.16c}$$

Symmetry breaking (quantization) along arbitrary crystal planes can be incorporated by an appropriate rotation of the real space coordinate system into principal directions, such that Eq. (4.14) maintains its form.

4.3.2 Operator Ordering

The approach we have taken is to generate a suitable differential operator for a nanostructure from the bulk Hamiltonians by replacing the wavenumbers of the symmetry broken directions k_j by the corresponding operators $-i\,\partial_j$. However, it is unclear in which particular order a term of the type Nk_ik_j or Pk_i must be written, as there are multiple options:

$$\begin{aligned} Nk_ik_j \;&\rightarrow\; -iN\partial_i\partial_j? \quad -i\,\partial_i N\,\partial_j? \quad -i\,\partial_j N\,\partial_i? \quad -i\,\partial_i\partial_j N?\\ Pk_i \;&\rightarrow\; -iP\partial_i? \quad -i\,\partial_i P? \end{aligned}$$

This ambiguity arises from the fact that in a bulk crystal the material parameters N, P are spatially invariant and therefore the differential operators commute with them, i.e. the information of the ordering is lost due to the homogeneity of the bulk crystal. In a nanostructure, $N(=N(x))$ and $P(=P(x))$ are position dependent, so the differential operators do not commute and the ordering of the operators has a large impact. Within the $\mathbf{k}\cdot\mathbf{p}$ model, the material parameters are usually discontinuous at material interfaces, resulting in a delta function for their spatial derivatives.

A common approach is to choose an ordering which conserves the probability current and leads to a Hermitian equation system. For terms of type $-Lk_i^2$ this is given by the Ben-Daniel Duke ordering $\partial_i L\partial_i$ [1]. We note that within the single-band effective mass theory, other orderings have been suggested (see [29] and references therein). But for terms of type $-Nk_ik_j$ the distribution between the two equivalent forms $\partial_i N_+\partial_j$ and $\partial_j N_-\partial_i$ ($N = N_+ + N_-$) is again ambiguous. Choosing a symmetric distribution $N_+ = N_- = \frac{N}{2}$ is not justified by any physical argument and will in general lead to the appearance of spurious solutions, as will be shown in Sect. 4.5.

Burt demonstrated [5–8, 14] that the ambiguity could be resolved by directly deriving the envelope equations from the nanostructure Schrödinger equation such that the details of the nanostructure and the interface would not be lost by the assumption of an infinite and homogeneous crystal. This provides detailed insight into the involved approximations, such as neglecting non-local terms and certain interface-related Hamiltonian terms. Burt's theory concludes that the resulting

Table 4.1 Material parameters taken from Vurgaftman and Meyer's review [48]. The γ_i are the Luttinger parameters. The Kane parameters L', M and N are related to the Luttinger parameters by $L' = -\frac{\hbar^2}{2m_0}(\gamma_1 + 4\gamma_2)$, $M = -\frac{\hbar^2}{2m_0}(\gamma_1 - 2\gamma_2)$ and $N = -\frac{\hbar^2}{2m_0}6\gamma_3$. The Kane parameters are given in units of $\frac{\hbar^2}{2m_0}$

Material	γ_1	γ_2	γ_3	L'	M	N	$N/2$	N_+	N_-
GaAs	6.98	2.06	2.93	−15.22	−2.86	−17.58	−8.79	−13.72	−3.86
AlAs	3.76	0.82	1.42	−7.04	−2.12	−8.52	−4.26	−5.40	−3.12
InAs	20.00	8.50	9.20	−54.00	−3.00	−55.20	−27.60	−51.20	−4.00
GaP	4.05	0.49	2.93	−6.01	−3.07	−17.58	−8.79	−13.51	−4.07
AlP	3.35	0.71	1.23	−6.19	−1.93	−7.38	−3.69	−4.45	−2.93
InP	5.08	1.60	2.10	−11.48	−1.88	−12.60	−6.30	−9.72	−2.88

Hamiltonian is of the form (4.14) but not necessarily symmetric ($N_+ \neq N_-$). It should be noted that it is a general goal of $\mathbf{k}\cdot\mathbf{p}$ solvers to predict nanostructure properties by using bulk parameters taken from literature. Therefore any dependence of the parameters on the particular nanostructure is not desirable.

Foreman derived in [14, 15] the non-symmetric operator ordering for a standard eight-band zinc-blende $\mathbf{k}\cdot\mathbf{p}$ Hamiltonian from crystal symmetry considerations. This is known as *Burt-Foreman operator ordering*. Similar derivations also exist for the wurtzite Hamiltonian [28, 34, 47]. Second-order operators of the form $\partial_i\,\partial_j\,N$ do not appear in the standard $\mathbf{k}\cdot\mathbf{p}$ model (they do exist in the first-principles model of [16]). We briefly outline how the terms $-\partial_i\,N\,\partial_j$ (which are problematic for a discontinuous N) result from the perturbative inclusion of the effect of remote bands into the bands explicitly included in the $\mathbf{k}\cdot\mathbf{p}$ Hamiltonian: The second-order (i.e. effective mass) matrix element of the original single-particle crystal Hamiltonian between the lattice-periodic functions for bands s and s', explicitly including the contribution of remote bands r, reads (see Eq. (6.4) in [6])

$$-\frac{\hbar^2}{2m_0}\left(\delta_{ss'}k^2 + \frac{2}{m_0}\sum_{i,j}\sum_{r} k_i\,p_{i,sr}(E - H_{rr}(x))^{-1}\,p_{j,rs'}k_j\right). \tag{4.17}$$

The term $p_{i,sr}$ represents the momentum matrix element for direction i between the respective lattice periodic Bloch functions s and r (see Eq. (3.4) in [6]). From (4.17), terms of type $k_i\,N_+k_j$ and $k_j\,N_-k_i$ arise which cannot be assumed to be symmetric.

A particular strength of the $\mathbf{k}\cdot\mathbf{p}$ model is the ability to diagonalize the bulk Hamiltonian analytically and determine its parameters by matching the resulting dispersion relations and energy gaps to experimental values from measurements on bulk crystals. But since in a bulk crystal the split into N_+ and N_- is irrelevant, the determination of N_+ and N_- is not possible. A practical solution to this was found by Foreman [14, 15] for the zincblende valence band Hamiltonian. In bulk, the second order terms of the Hamiltonian can be parametrized in terms of the Kane parameters L', M and N [24] and it is the N term which must be split, see Table 4.1 for the relation between the Kane parameters L, M and N and the

Luttinger parameters γ_1, γ_2 and γ_3. Inspecting the definitions of these parameters, Foreman found [15] that by neglecting contributions from very remote Γ_{25} bands, formed by f-type and higher atomic orbitals, N_- can be deduced entirely from M. Using the bulk value for N, N_+ is readily determined. As a result, it was found that the splitting is highly unsymmetric with differences between N_+ and N_- of up to a factor of 10 (see Table 1 and [44]).

4.3.3 Zincblende Models

For III-V semiconductors with zincblende crystal structures, a variety of different $\mathbf{k}\cdot\mathbf{p}$ models exists. The simplest conduction-band model is the single-band effective mass model, which easily fits into the given equation frame and will not be further discussed. The simplest physical model for the valence band (vb) is given by the $\mathbf{k}\cdot\mathbf{p}$ 4×4 model describing the Γ_8 valence bands (light and heavy holes). The 6×6 model adds the Γ_7 spin-orbit split-off band. These two models do not contain a coupling between valence and conduction bands and implicitly assume a single-band description of the conduction band (cb). By including the coupling to the lowest (Γ_6) conduction band, one obtains the 8×8 model. Models taking into account 14 [36], 16 [9], 24 [2] or 30 [18] bands are not considered in this chapter.

Some implementations are based on model Hamiltonians expressed in terms of the diagonal Bloch basis of a reference crystal at the Γ point (zone-center Bloch basis). The details about the basis functions do not have to be known since any properties which rely on their form are mangled into the $\mathbf{k}\cdot\mathbf{p}$ parameters. An alternative form of the 8×8 Hamiltonian is given by the Pidgeon-Brown [31] or Enders Hamiltonian [13], which is expressed in the zone-center basis

$$S\uparrow, X\uparrow, Y\uparrow, Z\uparrow, S\downarrow, X\downarrow, Y\downarrow, Z\downarrow\,. \tag{4.18}$$

Here $\{S, X, Y, Z\}$ denote states which exhibit the same real-space symmetry as the corresponding atomic orbitals, and $\{\uparrow,\downarrow\}$ denotes the spin degree of freedom. In this basis, the spin-orbit terms of the Hamiltonian are non-diagonal at $k=0$, but the Burt-Foreman operator ordering retains a simple and clear form [14, 45]. Furthermore, some care is required for the first-order interaction terms between the valence and conduction band in the $\mathbf{k}\cdot\mathbf{p}$ 8×8 model. The self-adjoint of the first order term iPk_j is given by $-ik_jP$, not $-iPk_j$. A wrong ordering here leads to a non-Hermitian linear equation system in a nanostructure and hence imaginary eigenvalues. The detailed form of the Hamiltonian is given in [13, 14, 31, 44, 46].

The form of the 6×6 model can be derived from the 8×8 model by taking the limit of an infinite bandgap $E_g\to\infty$ and neglecting the Γ_6 conduction band. However, the valence-band parameters used in the 6×6 model differ from the ones used in the 8×8 model as the cb-vb interaction in the 6×6 model must be included perturbatively in the remaining vb coefficients.

The 4×4 model is derived from the 6×6 model by reducing it to the heavy- and light-hole bands, neglecting the split-off band. This reduction is performed by choosing a combination of basis functions which effectively diagonalizes the spin-orbit interaction and transforms the 6×6 Hamiltonian, including the operator ordering, to the new basis.

4.3.4 Wurtzite Models

Group III-nitride materials often condense in the wurtzite crystal structure. Appropriate models for the valence band of this crystal type are given by a 6×6 model including the Γ_5 bands and the Γ_1 band split by the crystal field energy [3, 12]. Due to the large band gap, it usually suffices to model the conduction band by an anisotropic single band. However, the 6×6 model can also be extended to include the conduction band[21]. The correct operator ordering was initially worked out in [28] and recently discussed in [34] and by the authors [47].

4.4 FEM Discretization and Solution of the k·p Equations

4.4.1 Weak Form

The transformation of the equations (4.14) into the weak form (4.9b) is straightforward. Similar derivations can be found in the book by Ram-Mohan [33], mostly for quantum wells, and in the articles by Park et al. [30] (valence band of quantum wires) and Johnson et al. [23] (valence band states in quantum dots). Here a general derivation applicable to systems of any dimension shall be given. For the sake of simplicity, the parametric dependence on k_t is omitted in the following.

An n-band $\mathbf{k} \cdot \mathbf{p}$ model consists of a system of n coupled equations for the envelope components $f_i(x)$. The function space of each f_i is $H(\Omega)$. First, the boundary conditions (4.15) are enforced by restricting the Hilbert space to functions compatible with the Dirichlet condition,

$$H_0(\Omega) = \{f \in H(\Omega) : f(\partial\Omega) = 0\}. \tag{4.19}$$

This condition is also propagated to the finite dimensional subset V_M used to express the approximation to the solution and the mesh of Ω. In the spirit of (4.9b), Eq. (4.14) is left-multiplied with a vectorial test function $\mathbf{w}(x)$ having $w_i(x) \in H_0(\Omega)$ and integrated over the domain Ω. Using integration by parts and Gauss' law, the second-order terms transform into

$$-\int_{\Omega} \mathbf{w}^*(x)\partial_i \mathbf{H}_{ij}^{(2)}(x)\partial_j \mathbf{f}(x)\mathrm{d}x = -\int_{\partial\Omega} \mathbf{w}^*(x)\mathbf{H}_{ij}^{(2)}(x)\partial_j \mathbf{f}(x)n_i(x)\mathrm{d}s + \int_{\Omega} (\partial_i \mathbf{w}^*(x))\,\mathbf{H}_{ij}^{(2)}(x)\,(\partial_j \mathbf{f}(x))\mathrm{d}x, \tag{4.20}$$

where n_i is the i-th component of the normal vector to the boundary element ds of $\partial\Omega$. The first integral on the right-hand side (rhs) vanishes due to the restriction on the Hilbert-space to conform to the Dirichlet boundary condition. In the second integral any possible singularities of the derivatives of $\mathbf{H}^{(2)}(x)$ caused by discontinuous parameters are now resolved. The same trick can be used to remove the singularity at material interfaces of the first-order term $\partial_i \mathbf{H}_i^{(1R)}(x)$. Note that material interfaces are sharp and do not depend on the discretization in this approach. Equation (4.14) then reads

$$\begin{aligned} E\int_{\Omega} \mathbf{w}^*(x)\mathbf{f}(x)\mathrm{d}x &= \sum_i \sum_j \int_{\Omega} (\partial_i \mathbf{w}^*(x))\mathbf{H}_{ij}^{(2)}(x)(\partial_j \mathbf{f}(x))\mathrm{d}x \\ &+ \sum_i \int_{\Omega} \mathbf{w}^*(x)\mathbf{H}_i^{(1L)}(x)(\partial_i \mathbf{f}(x))\mathrm{d}x - \sum_i \int_{\Omega} (\partial_i \mathbf{w}^*(x))\mathbf{H}_i^{(1R)}(x)\mathbf{f}(x)\mathrm{d}x \\ &+ \int_{\Omega} \mathbf{w}^*(x)\mathbf{H}^{(0)}(x)\mathbf{f}(x)\mathrm{d}x. \end{aligned} \tag{4.21}$$

There is no need to explicitly enforce special boundary conditions inside Ω at any material interface. Such boundary conditions are implicitly included in the given differential operator and are denoted in the finite element framework as *natural boundary conditions*.

4.4.2 Numerical Discretization

Clearly, Eq. (4.21) is not yet in the form of (4.10b). Based on the shape functions $N_i(x)$ used for the scalar equation, the shape function for $\mathbf{w}(x)$ and $\mathbf{f}(x)$ in the coupled equation can be defined as

$$\mathbf{w}_{mg}(x) = \mathbf{e}_m N_g(x), \tag{4.22}$$

$$\mathbf{f}(x) = \sum_h \sum_n f_{nh}\mathbf{e}_n N_h(x), \tag{4.23}$$

with $\mathbf{e}_m, \mathbf{e}_n$ being n-dimensional Cartesian unit vectors ($\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij}$). Inserting into (4.21) yields

$$
\begin{aligned}
& E \sum_h \sum_n \mathbf{e}_m^T \mathbf{e}_n f_{nh} \int_\Omega N_g(x) N_h(x) \mathrm{d}x \\
& = \sum_h \sum_n \sum_i \sum_j f_{nh} \int_\Omega \left(\mathbf{e}_m^T \mathbf{H}_{ij}^{(2)}(x) \mathbf{e}_n \right) (\partial_i N_g(x))(\partial_j N_h(x)) \mathrm{d}x \\
& + \sum_h \sum_n \sum_i f_{nh} \int_\Omega \left(\mathbf{e}_m^T \mathbf{H}_i^{(1L)}(x) \mathbf{e}_n \right) N_g(x) (\partial_i N_h(x)) \mathrm{d}x \\
& - \sum_h \sum_n \sum_i f_{nh} \int_\Omega \left(\mathbf{e}_m^T \mathbf{H}_i^{(1R)}(x) \mathbf{e}_n \right) (\partial_i N_g(x)) N_h(x) \mathrm{d}x \\
& + \sum_h \sum_n f_{nh} \int_\Omega \left(\mathbf{e}_m^T \mathbf{H}^{(0)}(x) \mathbf{e}_n \right) N_g(x) N_h(x) \mathrm{d}x .
\end{aligned}
\tag{4.24}
$$

The indices g, h run over the nodal degrees of freedom while m, n run over the bands and i, j over the spatial dimensions. $\mathbf{e}_m^T \mathbf{H} \mathbf{e}_n$ is simply the matrix element (m, n) of the matrix $\mathbf{H}$. The integrals are scalar quantities which need to be evaluated only once at the beginning of a simulation. In *tdkp* we assume that for the first- and second-order terms the coefficients $\mathbf{H}$ are element-wise constant. Using analytically evaluated element integrals leads to very fast matrix assembly times. For zero-order terms we allow non-constant coefficients and assume that a coefficient is defined by its nodal values via the shape functions,

$$
c(x) = \sum_v c_v N_v(x). \tag{4.25}
$$

The last term in (4.24) requires us to evaluate integrals of the type

$$
\int_\Omega N_v(x) N_g(x) N_h(x) \mathrm{d}x. \tag{4.26}
$$

Using this approach we can accurately include strain effects which may induce sizeable variations of the band edge within an element, whilst assuming that a steady compositional change of the material is still well represented (see also Sect. 4.6).

The boundary conditions, which are plugged into the function space via (4.19), are enforced by setting the rows and columns associated with a node located at the boundary to 0. This is done by simply discarding the respective matrix entries (rows and columns) during assembly and setting the corresponding nodal result values to 0.

The resulting generalized eigenvalue problem $\mathbf{A}'\mathbf{f} = \lambda \mathbf{M}'\mathbf{f}$ differs from the single-band case in the sense that the matrices are block sparse, each block consisting of an

$n \times n$ matrix for an n-band model. The blocks in $\mathbf{M}'$ are multiples of the unity matrix, which allows to store only the scaling factor. Furthermore, $\mathbf{M}'$ is real-valued and symmetric (this applies only to the case of closed boundary conditions). The matrix $\mathbf{A}'$ is complex-Hermitian.

4.4.3 Solving the Generalized Eigenvalue Problem

Solving the generalized eigenvalue problem can be a rather delicate task. For $\mathbf{k} \cdot \mathbf{p}$ 4×4 and 6×6 models, the spectrum of the eigenvalues reaches from $-\infty$ to the uppermost valence band edge. As energies are always given with respect to some arbitrary reference, the band edge can be adjusted such that the eigenvalues and eigenfunctions of interest correspond to smallest-magnitude values located at the upper boundary of the matrix spectrum. For the $\mathbf{k} \cdot \mathbf{p}$ 8×8 model, which also features the conduction band, the bound states of interest correspond to interior eigenvalues which are more difficult to determine. Since only a few eigenvalues need to be sought, iterative methods are favorable. For *tdkp* we currently favor the Arnoldi [40] method implemented in ARPACK [26]. To achieve fast convergence, we use the *shift-and-invert* approach in which the original problem is transformed to a problem targeting the largest eigenvalues:

$$\mathbf{A} \rightarrow (\mathbf{A} - \tau \mathbf{M})^{-1} . \tag{4.27}$$

As a result, the number of required Arnoldi iterations is reduced to typically 70–250 depending on the system, at the cost that in each Arnoldi iteration a linear equation

$$(\mathbf{A} - \tau \mathbf{M})\, \mathbf{z} = \mathbf{f} \tag{4.28}$$

must be solved for changing $\mathbf{f}$. The shift-and-invert approach has been found to be very reliable. Convergence is always achieved, which is crucial for *tdkp* when being employed for transport simulations within the simulator *AQUA* [41].

The linear problem can be solved either using iterative or direct methods. Since the equation system has to be solved several times, we commonly use a direct factorization of the matrix provided that enough memory is available.

4.5 Spurious Solutions and Equation Ellipticity

4.5.1 Examples of Spurious Solutions

Researchers doing $\mathbf{k} \cdot \mathbf{p}$ calculations are often puzzled when they obtain band structures containing inplausible results. It was noted early on that the $\mathbf{k} \cdot \mathbf{p}$ 8×8

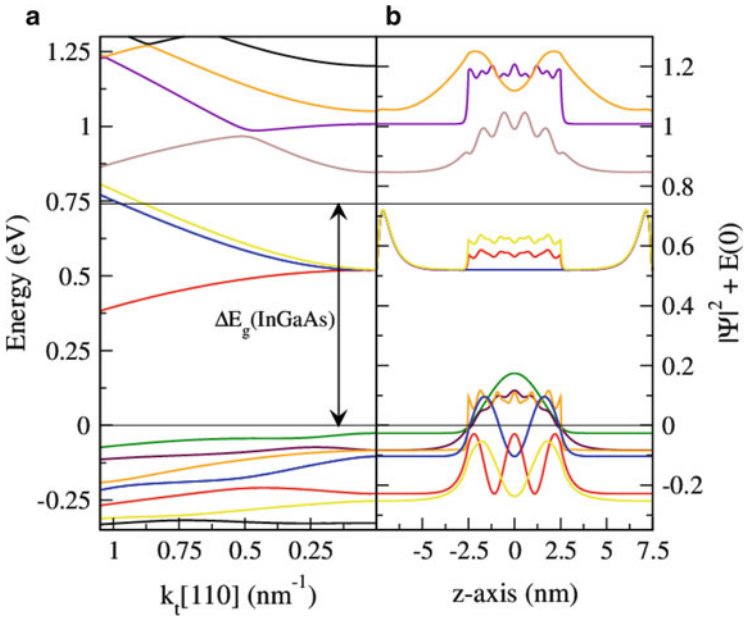


Fig. 4.1 (**a**) Band structure and (**b**) probability density at $k_t = 0$ of a 5 nm $In_{0.53}Ga_{0.47}As$ quantum well embedded in InP showing strong spurious solutions. The results were obtained using the 8×8 model, standard band structure parameters [48] at $T = 300$ K and symmetric operator ordering. The probability densities are normalized to a maximum value of 0.2 and shifted by their energy. The bulk valence band edge of the well material is located at 0.0 eV, the well conduction band edge at 0.74 eV. There are states within the bandgap, with a probability density exhibiting either an unplausible wing-band form, located at the boundary of the simulation domain, or oscillatory behavior within the well located between −2.5 and 2.5 nm. Several conduction subbands also bend downwards, i.e. into the band gap

model for zincblende could lead to unphysical, or spurious, solutions [37, 50]. An illustrative example of such solutions is given in Fig. 4.1. The figure shows the band structure and probability densities obtained for a 5 nm $In_{0.53}Ga_{0.47}As$ quantum well embedded in InP, solved by means of the 8×8 model. While certain subband shapes look reasonable, there are also bands bending into the forbidden gap. The eigenfunctions of the corresponding states can be highly oscillatory. A Fourier transform exhibits large contributions from outside the Brillouin zone, violating the initial assumption in $\mathbf{k} \cdot \mathbf{p}$ theory of slowly varying envelopes. The appearance of such solutions clearly disqualifies the method, jeopardizing its reliability and predictiveness.

While spurious solutions mostly arise in the $\mathbf{k} \cdot \mathbf{p}$ 8×8 model, they can also appear in the context of 4×4 and 6×6 models for both zincblende and wurtzite crystals, in particular when using symmetrized operator ordering. As an example we apply the 4×4 model to an (fictious) unstrained 5 nm InAs-GaAs quantum well.

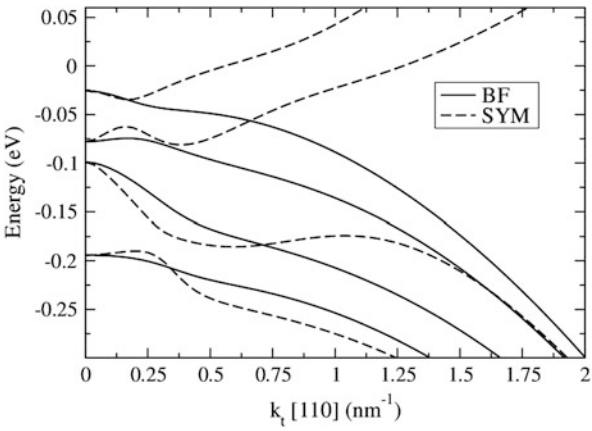


Fig. 4.2 Band structure of a 5 nm [001] InAs-GaAs quantum well in transversal [110] direction, calculated using the 4×4 model and finite elements. The *dashed lines* (labeled SYM) denote the results obtained using symmetrized operator ordering, while the *solid lines* (labeled BF) are results obtained using Burt-Foreman operator ordering. Strain effects are neglected

Our *tdkp* simulation results in the band structure depicted in Fig. 4.2. Note that the valence band edge of InAs is taken at 0 eV. The solid lines display the band structure obtained using Burt-Foreman ordering. The dashed lines result from symmetrized ordering, where the two highest resulting bands bend into the forbidden band gap, which is clearly spurious.

Extending the example to an artificial 2D square InAs-GaAs quantum wire leads to a similar picture. Again solved using the 4×4 model and neglecting strain effects, one obtains for the states closest to the uppermost valence band edge the probability densities depicted in Fig. 4.3. The figure shows 18 states organized in three rows. The uppermost row shows results calculated using symmetric operator ordering on a triangular mesh of 12,846 vertices. The second row is obtained using symmetric operator ordering on a tensorial grid of 17,161 vertices. The third row shows the results of applying Burt-Foreman operator ordering on the triangular mesh. The results for Burt-Foreman ordering on a tensorial grid are omitted as they match the results on the triangular grid up to less than 0.1 meV. The plotted states are the ones closest to the valence band edge, which is located at 0 eV. The valence band offset between InAs and GaAs is set to 0 eV. Hence there should be no particular reason for a particle to be confined within the InAs region. The symmetrized operator ordering yields solutions which are located entirely at the interface between InAs and GaAs and include highly oscillatory components. In addition, the solutions are not stable with respect to the particular meshing and triangulation used, as completely different solutions are obtained for the triangular and tensorial grids.

Further examples of spurious solutions appearing in different models for different crystals can be found in [44, 47]. All of them are connected to the symmetrized operator ordering and have in common a significant change of the effective mass parameters at a material interface.

Common approaches to reduce spurious solutions are based on adding parameters to the Hamiltonian [22], imposing special boundary conditions [35], filtering out

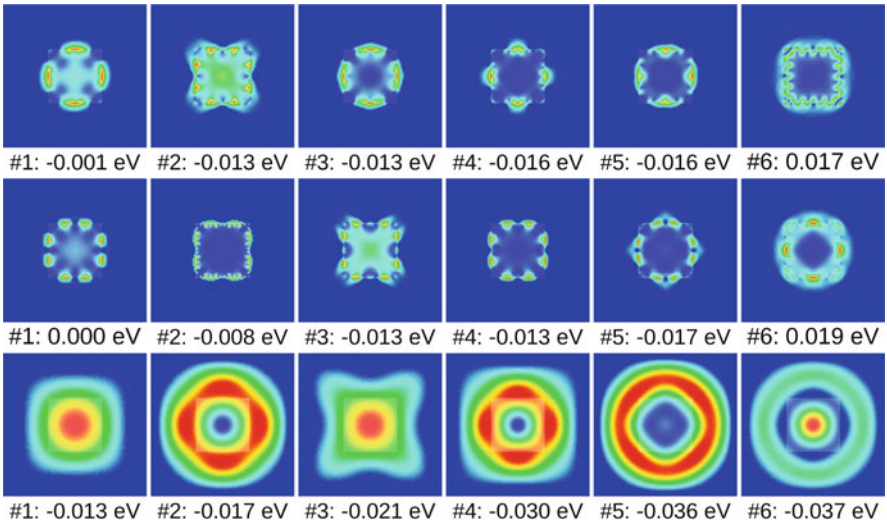


Fig. 4.3 Probability densities calculated for an artificial $(5\,\mathrm{nm})^2$ InAs nanowire embedded in GaAs, using the $\mathbf{k}\cdot\mathbf{p}$ 4×4 model and neglecting strain effects. The valence band edge and the offset between InAs and GaAs is set to 0 eV. The rows show six probabilities of the states closest to 0 eV (ignoring spin-degeneracy). The uppermost row is obtained using symmetric operator ordering on a triangular mesh. The second row is obtained using symmetric ordering on a tensorial grid. The third row is obtained using Burt Foreman operator ordering

any out-of-zone components [19] or simply removing the unwanted solutions numerically or analytically by choosing a basis set which is not able to represent highly oscillatory solutions. The fundamental weakness of these approaches is given by the fact that they treat the symptoms of the disease rather than the cause. As will be shown below, spurious solutions are a result of the posed equation system. It is the authors' view that the correct way to approach the issue is therefore to understand the reason why these solutions appear and to ask whether the equations can be formulated such that they result in a set of eigenvalue equations with conditions at the external boundaries which is free of spurious solutions. The answer to this question is yes.

Most discussions addressing the reason of spurious solutions [10, 14, 25, 27, 42, 43, 49, 50] are carried out in the framework of the transfer matrix method, except for [10] where finite differences are used. This means that the discussions are centered around a particular numerical method and focus on the validity of actual obtained solutions. Instead of discussing a particular solution method, we address the issue by a close inspection of the continuum equations themselves.

Without presenting a rigorous mathematical proof, it will be shown that the decisive *mathematical* criterion causing spurious solutions is the *ellipticity* of the coupled differential equation system given by the $\mathbf{k}\cdot\mathbf{p}$ equations. Based on this criterion, a set of conditions for the parameters in various $\mathbf{k}\cdot\mathbf{p}$ models can be derived which need to be fulfilled to retain ellipticity. Burt-Foreman operator ordering fits

into this context in that it results in (almost) elliptic equations while the symmetrized operator ordering violates this principle. Our ellipticity argument is simple but fundamental. The authors have no knowledge of spurious solutions which arise from the numerical discretization. Instead they are aware of finite-element, finite-difference and finite-volume implementations where the problem was eliminated by respecting the conditions outlaid in this section.

4.5.2 Ellipticity Criteria

4.5.2.1 Definition of Ellipticity

A scalar partial differential operator of second order

$$\mathscr{L}_2 = -\sum_{i,j=1}^{m} \partial_{i=1} c_{ij} \partial_j + \sum_{i}^{m} \left(d_i^L \partial_i + \partial_i d_i^R \right) + e \tag{4.29}$$

is termed *elliptic* if the matrix $\mathbf{C}$ composed of the coefficients c_{ij} is strictly definite:

$$\mathbf{x}^* \mathbf{C} \mathbf{x} > 0 \quad \forall \mathbf{x} \in \mathbb{R}^m. \tag{4.30}$$

The index boundary m corresponds to the dimension of the considered space. Equation (4.30) is the criterion for positive definiteness, but a similar argument can be made for a negative definite system by multiplying everything with -1.

The definiteness is equivalent to the condition that all eigenvalues of $\mathbf{C}$ are either entirely positive or entirely negative. A prominent example of a *parabolic* equation, where one eigenvalue of the second-order coefficient matrix $\mathbf{C}$ is 0, is the heat transfer equation $\partial_t u - \nabla^2 u = 0$. In case of a mix between strictly negative and positive eigenvalues, a *hyperbolic* equation is obtained, such as in the wave equation $\partial_t^2 u(x,t) - \nabla^2 u(x,t) = 0$. Elliptic equations are boundary value problems, having a solution uniquely determined by the boundary conditions, while parabolic and hyperbolic equations require a mix between boundary and initial conditions. They usually cannot be solved in a single step but have to be evolved in time.

Focusing now on the $\mathbf{k} \cdot \mathbf{p}$ equations, it is apparent from Kohn-Sham theory that the exact stationary (time-independent) Schrödinger equation of an electron many-body system, containing a nonlocal potential which is a function of the density, is elliptic. The coefficients of the second-order terms are nonzero only on the diagonal of the respective coefficient matrix and simply given by $\frac{\hbar^2}{2m_0}$. Any approximation to this equation should conserve the mathematical nature of the equation and not change the elliptic boundary value problem into a hyperbolic initial value problem. This is, as demonstrated below, exactly the case for the $\mathbf{k} \cdot \mathbf{p}$ equations with spurious solutions.

To illustrate this matter, we first extend the definition of ellipticity from the scalar case to the coupled envelope equations. In this case, the coefficient matrix $\mathbf{C}$ of the second-order terms of the envelope equations can be written as

$$\mathbf{C} = \begin{pmatrix} \mathbf{H}_{xx}^{(2)} & \mathbf{H}_{xy}^{(2)} & \mathbf{H}_{xz}^{(2)} \\ \mathbf{H}_{yx}^{(2)} & \mathbf{H}_{yy}^{(2)} & \mathbf{H}_{yz}^{(2)} \\ \mathbf{H}_{zx}^{(2)} & \mathbf{H}_{zy}^{(2)} & \mathbf{H}_{zz}^{(2)} \end{pmatrix}. \tag{4.31}$$

The scalar coefficients c_{ij} in (4.29) are hence replaced by the submatrices of the second-order components of (4.14). Calculating the eigenvalues of the matrix $\mathbf{C}$ in terms of the used parameters and checking their signs gives a simple tool to determine whether the equation is elliptic or not.

4.5.2.2 Ingredient 1: Correct Operator Ordering

To start with an actual example, let us focus on the $\mathbf{k} \cdot \mathbf{p}$ 6×6 Hamiltonian for the top valence band states in zincblende crystals [14,44], expressed in the basis of Eq. (4.18) excluding the conduction band (S). In this basis, there are no second-order coupling terms between states with different spin orientation, and all the second-order coupling is within states with equal spin. The individual coefficient matrices for spin-up and spin-down orbitals are the same, and the analysis can be restricted to one of these matrices [44]:

$$(h_{ij}^{kl}) = \frac{\hbar^2}{2m_0} \begin{pmatrix} L' & 0 & 0 & 0 & N_+ & 0 & 0 & 0 & N_+ \\ 0 & M & 0 & N_- & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & M & 0 & 0 & 0 & N_- & 0 & 0 \\ 0 & N_- & 0 & M & 0 & 0 & 0 & 0 & 0 \\ N_+ & 0 & 0 & 0 & L' & 0 & 0 & 0 & N_+ \\ 0 & 0 & 0 & 0 & 0 & M & 0 & N_- & 0 \\ 0 & 0 & N_- & 0 & 0 & 0 & M & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & N_- & 0 & M & 0 \\ N_+ & 0 & 0 & 0 & N_+ & 0 & 0 & 0 & L' \end{pmatrix}. \tag{4.32}$$

Here L', M, N are the Kane parameters and $N = N_+ + N_-$. For the symmetric operator ordering $N_{+/-} = \frac{1}{2}N$, while Foreman's estimate for N_- [17] is $N_- = M - \frac{\hbar^2}{2m_0}$. The eigenvalues of (4.32) in units of $\frac{\hbar^2}{2m_0}$ are

$$M - N_- \ (3\times), \quad M + N_- \ (3\times), \quad L - N_+ \ (2\times), \quad L + 2N_+ \ (1\times), \tag{4.33}$$

where the factor in the parenthesis denotes the eigenvalue degeneracy. The envelope equation is elliptic if all of these eigenvalues are negative (because we are inspecting

Table 4.2 Distinct eigenvalues (4.33) and the ratio (4.35) for both operator orderings, calculated using parameters from [48], listed in Table 4.1

Material	Ordering	λ_1	λ_2	λ_3	λ_4	ρ
GaAs	BF	−42.66	−6.72	−1.50	1.00	0.05
GaAs	SYM	−32.80	−11.65	−6.43	5.93	0.22
AlAs	BF	−17.84	−5.24	−1.64	1.00	0.08
AlAs	SYM	−15.56	−6.38	−2.78	2.14	0.16
InAs	BF	−156.40	−7.00	−2.80	1.00	0.02
InAs	SYM	−109.20	−30.60	−26.40	24.60	0.29
GaP	BF	−33.03	−7.14	1.00	7.50	0.33
GaP	SYM	−23.59	−11.86	2.78	5.72	0.38
AlP	BF	−15.09	−4.86	−1.74	1.00	0.09
AlP	SYM	−13.57	−5.62	−2.50	1.76	0.15
InP	BF	−30.92	−4.76	−1.76	1.00	0.06
InP	SYM	−24.08	−8.18	−5.18	4.42	0.22

the valence band). Hence one just needs to ensure that the following criteria for the valence band Kane parameters are fulfilled:

$$M - N_- < 0, \quad M + N_- < 0, \quad L - N_+ < 0, \quad L + 2N_+ < 0. \tag{4.34}$$

A simple way of visualizing a non-elliptic parametrization is obtained by summing the absolute value of the "wrong" eigenvalues and divide it by the sum of the absolute value of "correct" eigenvalues,

$$\rho = \left| \frac{\sum_{i,\lambda_i>0} \lambda_i}{\sum_{j,\lambda_j<0} \lambda_j} \right|. \tag{4.35}$$

We regard this ratio as an estimate of the degree of non-ellipticity. Ideally it should be zero. Inserting standard material parameters [11] for InAs, GaAs, AlAs, InP, GaP and AlP into (4.33) leads to the distinct eigenvalues given in Table 4.2. The table shows the values obtained for both Burt-Foreman and symmetrized operator ordering. All parameter sets yield a non-elliptic equation set, although the amount of positive eigenvalues is a lot higher for the symmetrized ordering than for the Burt-Foreman ordering. The difference is especially significant for the InAs–GaAs material system. Here, the ratio of positive to negative eigenvalues for symmetrized operator ordering is 0.29 for InAs and 0.22 for GaAs, whereas Burt-Foreman operator ordering gives small ratios of 0.02 and 0.05, respectively.

The estimate for N_- employed in the Burt-Foreman ordering [14] is only approximative and *always* leads to three constant positive eigenvalues of 1 (in units of $\frac{\hbar^2}{2m_0}$), which are small compared to the "correct" eigenvalues. Therefore Foreman's equation set is not quite elliptic. In Fig. 4.4a,c, contour plots show the ratios (4.35) for InAs, InP, AlAs, AlP, GaP and GaAs for different choices of N'

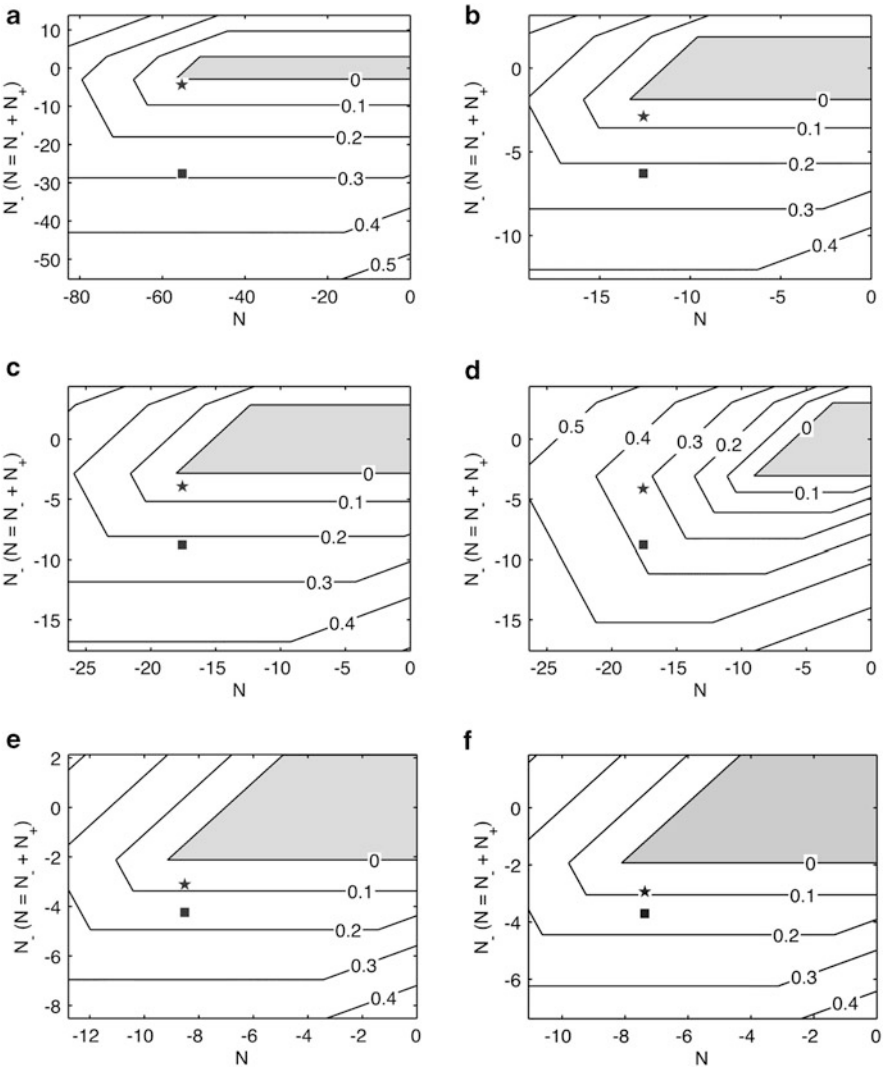


Fig. 4.4 Ratio (4.35) between positive and negative eigenvalues (4.33) of the $\mathbf{k} \cdot \mathbf{p}$ 6×6 Hamiltonian for the material parameters in Table 4.1 and different choices of N' and N_-. *filledstar* denotes Burt-Foreman operator ordering and *filledsquare* refers to the symmetric ordering. (**a**) InAs (**b**) InP (**c**) GaAs (**d**) GaP (**e**) AlAs (**f**) AlP

and N_-. The shaded regions mark the parameter space where the obtained equation is elliptic. Foreman's estimate, indicated by the symbol *filledstar*, is close to the elliptic parameter region. The symmetrized operator ($N_- = N'/2$) is indicated by the symbol *filledsquare*.

4.5.2.3 Ingredient 2 (for the 8 × 8 Model): Momentum Matrix Parameter

Some more care is required for the $\mathbf{k} \cdot \mathbf{p}$ 8×8 model which includes also the lowest conduction band. The Luttinger parameters (which are directly employed in the 4×4 and 6×6 models) contain the interaction between the valence bands and the conduction band and therefore need to be renormalized for the 8×8 model by subtracting the perturbative components. This involves the momentum matrix element P, related to the optical matrix parameter by $E_P = \frac{2m_0}{\hbar^2} P^2$, and the bandgap E_g. The parametrization therefore becomes temperature-dependent. The cb-vb coupling occurs in the standard model only via first-order terms iPk. Recall that the bulk inversion asymmetry yields a second order term for the cb-vb coupling (Kane's B parameter), which is very small and therefore usually neglected. The conduction band second-order term $\nabla \cdot A_c \nabla$ is obtained from the renormalization of the single-band effective mass m_c^* [17],

$$A_c = \frac{\hbar^2}{2m_c^*} - \frac{2}{3}\frac{P^2}{E_g} - -\frac{1}{3}\frac{P^2}{(E_g + \Delta)}, \tag{4.36}$$

where Δ is the spin orbit split-off energy. For the valence band, similar renormalizations are obtained. As shown in [44], this renormalization is delicate and can quickly lead to non-elliptic parameter sets in the valence band. The conduction band is easier to understand: the renormalized effective mass of the conduction band must simply be positive, $A_c > 0$. If the renormalization (4.36) is too strong, A_c becomes negative and the particle will resemble a hole.

Suppose now this is the case in a quantized region of a nanostructure. The band edge of this region lies below the band edge of the barrier region. A confined, artificially hole-like conduction band particle hence favors the barrier and tries to avoid the quantized region. However, the valence band components of the particle will be localized in the quantized region. Coupling these bands together, as it is done via the first-order terms, leads to high oscillations, as in every region at least one component tries to vanish. This is what can be seen in Fig. 4.1b, where the states in the bandgap are trying to avoid the quantized region.

To cope with this, both the authors and Foreman suggest to reduce the optical matrix parameter E_P (or, equivalently, P) [14, 44]. The band structure itself around the Γ-point is independent of E_P as the involved renormalizations adjust the second-order parameters of the 8×8 model such that the dispersion defined by the original effective mass parameters is retained. In our experience, a reduction of E_P by 10–20 % usually leads to numerically stable parameters and does not affect the bulk band structure around Γ. In some cases we also reduce γ_3 to obtain elliptic parameter sets in the 8×8 model. This is to our regret still a somewhat arbitrary workaround that needs further justification.

But applying this procedure to the $In_{0.53}Ga_{0.47}As/InP$ example leads to the spurious-solution-free results shown in Fig. 4.5.

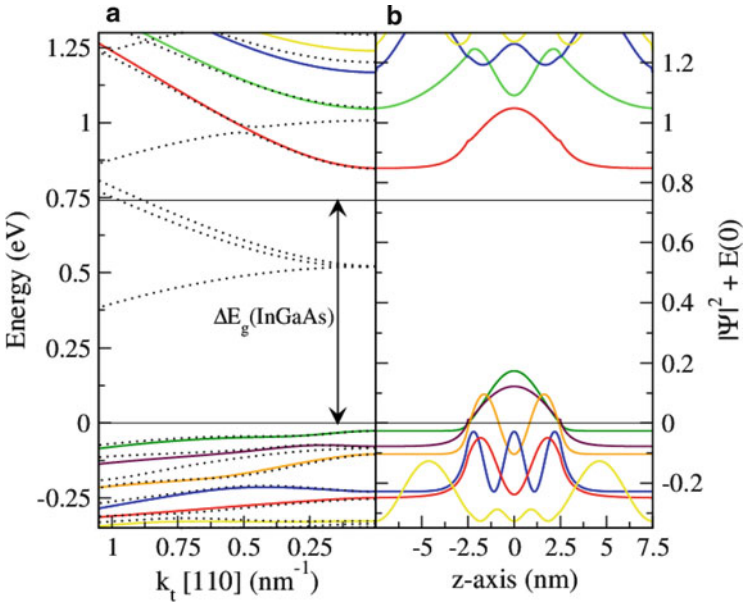


Fig. 4.5 Band structure (**a**) and probability density (**b**) of a 5 nm $In_{0.53}Ga_{0.47}As$ quantum well embedded in InP. The results were obtained using the $\mathbf{k} \cdot \mathbf{p}$ 8×8 model, Burt-Foreman operator ordering and slightly modified band structure parameters (InP: $E_P = 15.2\,\text{eV}(-26.6\,\%)$, $In_{0.53}Ga_{0.47}As$: $E_P = 19.48\,\text{eV}(-17\,\%), \gamma_3 = 6.19(-1\,\%)$). The parameters E_P and γ_3 were reduced such that the parameters meet the ellipticity criteria (4.34). The *black-dotted lines* represent the band structure obtained from symmetric ordering as plotted in Fig. 4.1

4.6 Strain and Polarization

Nanostructures are frequently composed of materials having different lattice constants. A pseudomorphic interface between these materials requires a matching of the individual crystals, which in turn leads to intrinsic strains. Strain, described by the tensor ε, in turn often induces polarizations $\mathbf{P}$ via the piezoelectric effect which to first order is given by

$$P_i = \sum_{jk} e_{ijk}\varepsilon_{jk}. \tag{4.37}$$

Here i, j, k run over the spatial dimensions and e_{ijk} is the piezoelectric stress coefficient. A change in polarization is not necessarily strain-induced but can also arise from varying spontaneous polarization across a material interface. A change of the polarization results in a charge field,

$$\rho(x) = -\nabla \cdot \mathbf{P}(x), \tag{4.38}$$

which must be included in the electronic calculation via an additional potential obtained from the solution of the respective Poisson equation. How intrinsic strain is calculated using linear elasticity theory is widely covered within literature (see e.g. [20, 32]) and is not repeated here. We also skip the theory behind polarization charges in favor of discussing how to precisely include this effect in the framework of finite elements.

The usual solution variable for calculating strain using finite elements is given by the displacement from the unstrained reference shape $\mathbf{u}(x)$. The linear small strain measure ε_{ij} is then obtained from the derivatives of the displacement,

$$\varepsilon_{ij}(x) = \frac{1}{2}\left(\partial_j u_i(x) + \partial_i u_j(x)\right). \tag{4.39}$$

The polarization charge $\rho(x)$ in turn is obtained from the derivative of the strain and therefore from the second-order derivative of the displacement. While we expect the displacement to be continuous (i.e. there are no cracks in the structure), the first-order derivative is not. This means that the polarization is expected to be discontinuous across element boundaries.

If $\mathbf{P}_L$ and $\mathbf{P}_R$ are the polarization left and right of the boundary, the total polarization can be written as

$$\mathbf{P}(x) = (\mathbf{P}_R - \mathbf{P}_L)\Theta_{AB}(x) + \mathbf{P}_L, \tag{4.40}$$

where $\Theta(x)$ is the Heaviside function ($\Theta = 0$ in the left element A and $\Theta = 1$ in the right element B). Taking the negative divergence yields the polarization charge:

$$\varrho(x) = -\mathbf{n}_{AB} \cdot (\mathbf{P}_R - \mathbf{P}_L)\delta_{AB}(x). \tag{4.41}$$

Here, $\mathbf{n}_{AB}$ denotes the normal to the boundary and $\delta_{AB}(x)$ is the delta function that vanishes everywhere except on the boundary AB.

As a result, a surface charge on the boundary is obtained. Higher-order approximation polynomials for the displacement lead to additional volume charges which are easy to handle and are not repeated here. The FEM-discretized form of Poisson's equation for the coefficients ϕ_j of the potential $\phi(x) = \sum_j \phi_j N_j(x)$ then reads

$$\sum_{j=1}^{\mathrm{M}} \phi_j \int_\Omega \varepsilon_b(x) \nabla N_i(x) \nabla N_j(x) \mathrm{d}x = \int_\Omega N_i(x) \varrho(x) \mathrm{d}x. \tag{4.42}$$

The tricky part is how to include the delta function on the right-hand side. Rewriting the rhs using $\Delta\mathbf{P}_{B_j}$ as the jump of the polarization across the element boundary B_j gives

$$\int_{\Omega} N_i(x)\varrho(x)\mathrm{d}x = -\int_{\Omega} N_i(x)\left(\sum_{B_j} \Delta\mathbf{P}_{B_j}(x)\cdot\mathbf{n}_{B_j}\delta_{B_j}(x)\right)\mathrm{d}x \quad (4.43)$$

$$= -\sum_{B_j}\int_{B_j} \Delta\mathbf{P}_{B_j}(x)\cdot\mathbf{n}_{B_j} N_i(x)\mathrm{d}s. \quad (4.44)$$

Here $\int_{B_j}\mathrm{d}s$ denotes the integration over the boundary (a face in 3D, an edge in 2D, and a point in 1D).

Note that the shape functions N_i are defined on the elements while they have to be evaluated on element boundaries. As such, for every element type also integrations over every boundary would be required. An elegant way to prevent such a tedious implementation is to use the fact that for the Lagrange polynomials these integrals are already implemented by a corresponding lower-dimensional element. As an example, within a triangular domain partition, the first-order nodal shape function restricted to an element edge containing the corresponding node is given by the linear, one-dimensional shape function. On edges not containing the given node, the used nodal shape functions N_i always vanish and the integrals are only nonzero on element boundaries B_j containing the node i. Therefore, only an appropriate mapping of the element boundary to a lower-dimensional element with the corresponding shape function is required.

This elegant way of including surface charges allows to model nanostructures subject to strain and interface charges within 2D and 3D simulations with a high accuracy. This stands in contrast to other discretization methods which in effect artificially smear out surface charges.

4.7 Conclusion

This chapter provides an overview on the essentials which need to be considered when solving the $\mathbf{k}\cdot\mathbf{p}$ equations using finite elements. Besides the FEM discretization of the Hamiltonian and the treatment of strain and interface charges, it is crucial that the continuum equations are well-posed. In the $\mathbf{k}\cdot\mathbf{p}$ case, the mathematical criterion of ellipticity translates to correct operator ordering (Burt-Foreman ordering for the zincblende models) and appropriate parameter sets. This can be achieved by limiting the magnitude of the optical matrix parameter in the $\mathbf{k}\cdot\mathbf{p}$ 8×8 model. The resulting computer simulations are robust, fast, accurate and predictable. They are excellent building blocks for the understanding of optoelectronic devices of any flavor.

Acknowledgements This work was funded by a Swiss National Science Foundation (SNSF) fellowship for prospective researchers. The chapter is based on results the authors obtained whilst working on the SNF-GAIN project, funded through SNSF grant 200021-107932.

References

1. D.J. BenDaniel, C.B. Duke, Space-charge effects on electron tunneling. Phys. Rev. **152**, 683–692 (1966)
2. S. Ben Radhia, K. Boujdaria, S. Ridene, H. Bouchriha, G. Fishman, Band structures of GaAs, InAs, and Ge: A 24-$\mathbf{k} \cdot \mathbf{p}$ model. J. Appl. Phys. **94**, 5726–5731 (2003)
3. G.L. Bir, G.E. Pikus, *Symmetry and Strain-Induced Effects in Semiconductors* (John Wiley & Sons, 1974)
4. D. Braess, *Finite Elemente* (Springer-Verlag Berlin, Heidelberg, New York, 1997)
5. M.G. Burt, A new effective-mass equation for microstructures. Semicond. Sci. Technol. **3**, 1224–1226 (1988)
6. M.G. Burt, The justification for applying the effective-mass approximation to microstructures. J. Phys.: Condens. Matter **4**, 6651–6690 (1992)
7. M.G. Burt, Direct derivation of effective-mass equations for microstructures with atomically abrupt boundaries. Phys. Rev. B **50**, 7518 (1994)
8. M.G. Burt, Fundamentals of envelope function theory for electronic states and photonic modes in nanostructures. J. Phys.: Condens. Matter **11**, R53–R83 (1999)
9. M. Cardona, N.E. Christensen, G. Fasol, Relativistic band structure and spin-orbit splitting of zinc-blende-type semiconductors. Phys. Rev. B **38**, 1806–1827 (1988)
10. X. Cartoxia, D.Z.Y. Ting, T.C. McGill, Numerical spurious solutions in the effective mass approximation. J. Appl. Phys. **93**, 3974 (2003)
11. S.L. Chuang, *Physics of Optoelectronic Devices* (John Wiley & Sons, 1995)
12. S.L. Chuang, C.S. Chang, $\mathbf{k} \cdot \mathbf{p}$ method for strained wurtzite semiconductors. Phys. Rev. B **54**, 2491 (1996)
13. P. Enders, A. Barwolff, M. Woerner, D. Suisky, **k·p** theory of energy bands, wave functions, and optical selection rules in strained tetrahedral semiconductors. Phys. Rev. B **51**, 16695 (1995)
14. B.A. Foreman, Elimination of spurious solutions from eight-band $\mathbf{k} \cdot \mathbf{p}$ theory. Phys. Rev. B **56**, R12748 (1997)
15. B.A. Foreman, Effective-mass hamiltonian and boundary conditions for the valence bands of semiconductor microstructures. Phys. Rev. B **48**, 4964 (1993)
16. B.A. Foreman, First-principles envelope-function theory for lattice-matched semiconductor heterostructures. Phys. Rev. B **72**, 165345 (2005)
17. B.A. Foreman, Choosing a basis that eliminates spurious solutions in $\mathbf{k} \cdot \mathbf{p}$ theory. Phys. Rev. B **75**, 235331 (2007)
18. N. Fraj, I. Saidi, R. Ben, K. Boujdaria, Band structures of AlAs, GaP, and SiGe alloys: A 30 $\mathbf{k} \cdot \mathbf{p}$ model. J. Appl. Phys. **21**, 8979 (2007)
19. M.J. Godfrey, A.M. Malik, Boundary conditions and spurious solutions in envelope function theory. Phys. Rev. B **53**, 16504 (1996)
20. M. Grundmann, O. Stier, D. Bimberg, InAs/GaAs pyramidal quantum dots: Strain distribution, optical phonons and electronic structure. Phys. Rev. B **52**, 11969 (1995)
21. J. Hader, J.V. Moloney, A. Thranhardt, S.W. Koch, Interband Transitions in InGaN Quantum Wells. In: Nitride Semiconductor Devices: Principles and Simulation. NUSOD Institute Newark, DE 19714-7204, USA (2007), 145–167
22. M. Holm, M.E. Pistol, C. Pryor, Calculations of the electronic structure of strained inas quantum dots in InP. J. Appl. Phys. **92**, 932 (2002)
23. H.T. Johnson, L.B. Freund, The influence of strain on confined electronic states in semiconductor quantum structures. Int. J. Sol. Struct. **38**, 1045 (2001)
24. E.O. Kane, *Energy Band Theory* In: Handbook on Semiconductors, Vol. 1, W. Paul (ed.) (1982), 194–217
25. K.I. Kolokolov, J. Li, C.Z. Ning, $\mathbf{k} \cdot \mathbf{p}$ Hamiltonian without spurious-state solutions Phys. Rev. B **68**, 161308(R) (2003)
26. R.B. Lehoucq, D.C. Sørensen, C. Yang, *Arpack users guide: Solution of large scale eigenvalue problems by implicitly restarted arnoldi methods* (SIAM, Philadelphia, 1998)

27. A.T. Meney, B. Gonul, E.P. O'Reilly, Evaluation of various approximations used in the envelope-function method. Phys. Rev. B **50**, 10893 (1994)
28. F. Mireles, S.E. Ulloa, Ordered hamiltonian and matching conditions for heterojunctions with wurtzite symmetry: GaN/$Al_xGa_{1-x}N$ quantum wells. Phys. Rev. B. **60**, 13659 (1999)
29. R.A. Morrow, K.R. Brownstein, Model effective-mass hamiltonians for abrupt heterojunctions and the associated wave-function-matching conditions Phys. Rev. B **30**, 678–680 (1984)
30. S.H. Park, D. Ahn and Y.T. Lee, Finite element analysis of valence band structures in quantum wires. J. Appl. Phys. **96** 2055 (2004)
31. C.R. Pidgeon, R.N. Brown, Interband magneto-absorption and Faraday rotation in insb. Phys. Rev. **146**, 146 (1966)
32. M. Povolotskyi, A.D. Carlo, Elasticity theory of pseudomorphic heterostructures grown on substrates of arbitrary thickness. J. Appl. Phys. **100**, 063514 (2006)
33. L.R. Ram-Mohan, *Finite Element and Boundary Element Applications in Quantum Mechanics* (Oxford University Press, 2003)
34. L.R. Ram-Mohan, K.H. Yoo, Wavefunction engineering of layered semiconductors: theoretical foundations. J. Phys.: Condens. Matter **18**, R901–R917 (2006)
35. A.V. Rodina, A.Y. Alekseev, A.L. Efros, M. Rosen, B.K. Meyer, General boundary conditions for the enevelope function in the multiband $\mathbf{k}\cdot\mathbf{p}$ model Phys. Rev. B. **65**, 125302 (2002)
36. U. Rössler, Nonparabolicity and warping in the conduction band of GaAs. Solid State Commun. **49**, 943–947 (1984)
37. M.F.H Schuurmans, G.W. t'Hooft, Simple calculations of confinemenet states in a quantum well. Phys. Rev. B. **31**, 8041 (1985)
38. C. Schwab, *p- and hp- Finite Element Methods* (Clarendon Press Oxford, 1998).
39. T. Sandu, Optical matrix elements in tight-binding models with overlap. Phys. Rev. B **72**, 125105 (2005)
40. D.C. Sørensen, Implicitly restarted Arnoldi/Lanczos methods for large scale eigenvalue calculations. Technical Report TR-96-40 (1996)
41. S. Steiger, R.G. Veprek, B. Witzigmann, Unified simulation of transport and luminescence in optoelectronic nanostructures. J. Comput. Electron. **7**, 509–520 (2008)
42. F. Szmulowicz, Envelope-function theory without spurious real solutions Europhys. Lett. **69**, 249 (2005)
43. F. Szmulowicz, Solution to spurious bands and spurious real solutions in the envelope-function approximation. Phys. Rev. B **71**, 245117 (2005)
44. R.G. Veprek, S. Steiger, B. Witzigmann, Ellipticity and the spurious solution problem of $\mathbf{k}\cdot\mathbf{p}$ envelope equations. Phys. Rev. B **76**, 165320 (2007)
45. R.G. Veprek, S. Steiger, B. Witzigmann, Reliable $\mathbf{k}\cdot\mathbf{p}$ band structure calculation for nanostructures using finite elements. J. Comput. Electron. **7**, 521–529 (2008)
46. R.G. Veprek, Computational modeling of semiconductor nanostructures for optoelectronics, Ph.D. dissertation, ETH Zürich (2009)
47. R.G. Veprek, S. Steiger, B. Witzigmann, Operator ordering, ellipticity and spurious solutions in $\mathbf{k}\cdot\mathbf{p}$ calculations of iii-nitride nanostructures. Opt. Quant. Electr. (2009)
48. I. Vurgaftman, J.R. Meyer, Band parameters for iii-v compound semiconductors and their alloys. Appl. Phys. Rev. **89**, 5815 (2001)
49. L.W. Wang, Real and spurious solutions of the 8×8 $\mathbf{k}\cdot\mathbf{p}$ model for nanostructures. Phys. Rev. B **61**, 7241 (2000)
50. S.R. White, L.J. Sham, Electronic properties of flat-band semiconductor heterostructures. Phys. Rev. Lett. **47**, 879 (1981)
51. O.C. Zienkiewicz, R.L. Taylor, *Finite Element Method Volume 1 – The Basis* (Elsevier, 5th edition, 2000)

Chapter 5
Plane-Wave Approaches to the Electronic Structure of Semiconductor Nanostructures

Eoin P. O'Reilly, Oliver Marquardt, Stefan Schulz, and Aleksey D. Andreev

Abstract This chapter is dedicated to different plane-wave based approaches to calculate the electronic structure of semiconductor nanostructures. We introduce semi-analytical and numerical methods to achieve a plane-wave based description of such systems. This includes use of plane-wave methods to calculate not just the electronic structure but also the built-in strain and the polarisation potential, with the strain and the polarisation potential each having a significant influence on the electronic properties of a semiconductor nanostructure. The advantages and disadvantages of different plane-wave based formulations in comparison to a real-space, finite element model will be discussed and we will present representative examples of semiconductor nanostructures together with their elastic and electronic properties, computed from semi-analytical and numerical approaches. We conclude that plane-wave-based methods provide an efficient and flexible approach when using **k·p** models to determine the electronic structure of semiconductor nanostructures.

5.1 Plane-Wave Approaches to Real-Space Problems

The description of real-space properties of a system can often be achieved in a very efficient manner by employing a reciprocal-space approach using plane waves. A prominent example where this approach is commonly used is in the application of density functional theory (DFT) [21, 23], to calculate the electronic properties

E.P. O'Reilly (✉) • O. Marquardt • S. Schulz
Tyndall National Institute, Cork, Ireland
e-mail: eoin.oreilly@tyndall.ie; marquardt@pdi-berlin.de; stefan.schulz@tyndall.ie

A.D. Andreev
Hitachi Cambridge Laboratory, Cambridge, UK
e-mail: andreev@phy.cam.ac.uk

M. Ehrhardt and T. Koprucki (eds.), *Multi-Band Effective Mass Approximations*,
Lecture Notes in Computational Science and Engineering 94,
DOI 10.1007/978-3-319-01427-2_5,

of many-particle systems such as atoms, molecules and crystals. Within DFT, the Schrödinger equation, $H\Psi(\mathbf{r}) = \varepsilon\Psi(\mathbf{r})$, to determine electron wavefunctions and energies needs to be solved self-consistently:

$$H\Psi(\mathbf{r}) = \left(-\frac{\hbar^2}{2m_0}\nabla^2 + V([n(\mathbf{r})];\mathbf{r})\right)\Psi(\mathbf{r}) = \varepsilon\Psi(\mathbf{r}). \tag{5.1}$$

The potential term $V([n(\mathbf{r})];\mathbf{r})$ here includes contributions such as the electron-electron Coulomb interaction, $u([n(\mathbf{r})];\mathbf{r})$, which depends on the electron density $n(\mathbf{r})$ as:

$$u([n(\mathbf{r})];\mathbf{r}) \propto \int \frac{n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}dV'. \tag{5.2}$$

The electron density $n(\mathbf{r})$ is given by the sum of the charge densities $\langle\Psi(\mathbf{r})|\Psi(\mathbf{r})\rangle$ associated with each of the filled electron states in the system. Because H depends on Ψ and Ψ on H, an iterative approach is then required to find the allowed solutions of the Schrödinger equation. Such an iterative approach requires efficient calculation techniques, and it is noteworthy that plane-wave methods are the most widely used approach when solving the Schrödinger equation within DFT.

An arbitrary electron wavefunction $\Psi(\mathbf{r})$ can be expressed in reciprocal space using plane waves via a Fourier transform taken over the volume V_{sc} of the supercell being investigated, to give:

$$\Psi(\mathbf{g}) = \frac{1}{\sqrt{V_{sc}}}\int_{V_{sc}} \Psi(\mathbf{r})e^{i\mathbf{g}\cdot\mathbf{r}}dV. \tag{5.3}$$

For a periodic system, the wavefunction $\Psi_{\mathbf{k}}(\mathbf{r})$, for a certain wave vector $\mathbf{k}$, can be expressed using the Bloch theorem as

$$\Psi_{\mathbf{k}}(\mathbf{r}) = \sum_{\mathbf{G}} c_{\mathbf{k},\mathbf{G}} \cdot e^{i(\mathbf{G}+\mathbf{k})\cdot\mathbf{r}}, \tag{5.4}$$

with the $\mathbf{G}$'s being the reciprocal lattice vectors. The accuracy with which the wavefunction is described is then determined by the number of plane waves employed in (5.4). The number of plane waves used when performing a DFT calculation is typically controlled via the choice of the cut-off energy E_c:

$$\frac{\hbar^2}{2m_0}(\mathbf{G}+\mathbf{k})^2 \leq E_c. \tag{5.5}$$

The cut-off energy can be treated as a convergence parameter, whose value is chosen as a balance between computational costs and the required energy accuracy.

Compared to a real-space description, the reciprocal space approach has a number of decisive advantages:

1. The accuracy of the calculation with respect to energies and spatial resolution is directly controlled via the single parameter E_c, and thus via the number of plane waves employed.
2. Within a numerical calculation, a plane-wave formulation is commonly highly efficient and a fast Fourier transform (FFT), or a modified FFT (depending on implementation), can be used to transform wavefunctions and charge densities from reciprocal to real space and vice versa.
3. The whole simulation cell is treated similarly, as plane waves expand throughout the whole real space.
4. The periodic nature of plane waves makes a plane-wave based formalism perfectly suited to periodic problems such as crystalline systems, where periodic boundary conditions need to be explicitly enforced in a real-space approach. However, aperiodic systems such as molecules or surfaces can also be successfully described in a plane-wave picture using a sufficiently large supercell in which the molecule is surrounded by vacuum.

Although there are several real-space methods available for DFT calculations [11, 31], the efficiencies associated with using a plane-wave framework have generally led to their being preferred, and most of the existing DFT codes rely on a plane-wave representation [9, 18, 24, 41].

For the description of the electronic properties of semiconductor nanostructures such as quantum dots, wires and wells using multi-band **k·p** models, a real-space description based on finite elements and finite differences has been the more common choice [10, 16, 32, 36, 37, 43, 48, 51]. However, a variety of different plane-wave based multi-band **k·p** models has been implemented and successfully employed in the past [2, 13, 27, 45]. These plane-wave based **k·p** methods not only retain the advantages listed above, but can also make use of one further advantage, namely that there are many nanostructure shapes and problems for which the Fourier transform can be calculated analytically. Use of these analytical Fourier transforms can then speed up the setting up of the reciprocal space calculation when using the **k·p** method to analyse semiconductor quantum dot electronic structure.

Within this chapter, we introduce both semi-analytical and numerical approaches which employ a plane-wave formalism for multi-band effective mass calculations of electronic properties of semiconductor nanostructures (Sect. 5.2). We start by introducing in Sect. 5.2.1 a semi-analytical model for a simple single-band effective mass Hamiltonian and then generalise the model to the case of an n-band **k·p** Hamiltonian. The widely generalised approach for multi-band **k·p** models employed within the S/Phi/nX package [9, 27, 29] is used in Sect. 5.2.2 as an example of how plane-wave based multi-band effective mass models can be implemented numerically. A numerical approach based on such a DFT package has the advantage that it can make use of the highly optimised minimisation routines and the mathematical framework that is commonly available in such a DFT library. We describe in Sect. 5.3 how plane-wave methods are particularly suited to calculate the elastic response and strain distribution associated with a semiconductor nanostructure embedded in a host semiconductor material. This is followed in Sect. 5.4 by a

description of how plane-wave methods also provide a very efficient approach to solve the equation and so calculate the polarisation potential associated with the nanostructure. An accurate description both of the strain distribution and of the built-in polarisation potential is essential when calculating the electronic properties of semiconductor nanostructures. The advantages and disadvantages of plane-wave frameworks in comparison with a real-space finite element or finite difference approach are discussed in Sect. 5.5 and we present examples where plane-wave based 8-band **k·p** models have been employed to investigate the electronic properties of different semiconductor nanostructures in Sect. 5.6. Finally we summarise our conclusions in Sect. 5.7.

5.2 Plane-Wave Based Formulation of Elastic and Electronic Properties

We consider a quantum dot (QD) embedded in a surrounding matrix material. The wavefunction and eigenenergy of a particle in this system can be determined within an n-band **k·p** model by solving the Schrödinger equation:

$$H^{n\times n}|\Psi\rangle = \left(H_{M0}^{n\times n} + \Delta H_{DM0}^{n\times n} + H_s^{n\times n} + \varphi + V_{ext}\right)|\Psi\rangle = \varepsilon|\Psi\rangle, \tag{5.6}$$

where we have written this Hamiltonian such that $H_{M0}^{n\times n}$ is the n-band **k·p** Hamiltonian for the bulk matrix material in equilibrium, $\Delta H_{DM0}^{n\times n} = \left(H_{D0}^{n\times n} - H_{M0}^{n\times n}\right)\chi(\mathbf{r})$, where $H_{D0}^{n\times n}$ is the the equilibrium Hamiltonian for the dot material, $\chi(\mathbf{r})$ is the dot characteristic function, equal to 1 in a dot of constant composition and equal to 0 in the surrounding matrix material; $H_s^{n\times n}$ represents the strain dependence of the electronic properties and φ and V_{ext} are a polarisation and an optional external scalar potential.

Epitaxial growth of nanostructures commonly induces strain in the system, which can significantly modify the bulk electronic properties of the Hamiltonian in (5.6) via the contribution $H_s^{n\times n}$. The strain tensor $\epsilon_{ij}(\mathbf{r})$ can be determined by minimising the elastic energy of the system [22]:

$$F = \frac{1}{2}\int_{V_{sc}} C_{ijkl}(\mathbf{r})\epsilon_{ij}(\mathbf{r})\epsilon_{kl}(\mathbf{r})dV, \tag{5.7}$$

where $C_{ijkl}(\mathbf{r})$ are the elastic constants of the system and $\epsilon_{ij}(\mathbf{r})$ is

$$\epsilon_{ij}(\mathbf{r}) = \frac{1}{2}\left(\frac{\partial u_i(\mathbf{r})}{\partial r_j} + \frac{\partial u_j(\mathbf{r})}{\partial r_i}\right) + \epsilon_{ij}^0(\mathbf{r})\chi(\mathbf{r}). \tag{5.8}$$

Here, $\mathbf{u}(\mathbf{r})$ represents the displacement of a volume element in the supercell, $\epsilon_{ij}^0(\mathbf{r})$ is a contribution arising from the bulk lattice mismatch between the materials involved, and $\chi(\mathbf{r})$ is the nanostructure characteristic function, as defined after (5.6).

Strain can also induce a piezoelectric polarisation potential in the vicinity of an epitaxially grown nanostructure, which enters the **k**·**p** Hamiltonian in (5.6) via the potential term φ. In a simple, first-order approximation, the polarisation potential in the vicinity of a nanostructure can be determined from the strain via the polarisation vector $\mathbf{P}(\mathbf{r})$, whose i-th component is given by:

$$P_i(\mathbf{r}) = \sum_{jk} e_{ijk}\epsilon_{jk}, \tag{5.9}$$

where e_{ijk} are the material piezoelectric coefficients. Although the importance of the second-order piezoelectric terms has been outlined previously [8], we refrain from discussing these contributions here for the sake of simplicity. In practice, the implementation of second-order piezoelectric contributions can be carried out in a similar manner to that used for the first order terms. Once the polarisation is known, the polarisation charge density is then given from Gausss law by $\varrho_p(\mathbf{r}) = -\nabla\cdot\mathbf{P}(\mathbf{r})$. From this, the polarisation potential φ can be determined by solving the Poisson equation:

$$\nabla(\kappa_0\kappa_r(\mathbf{r})\nabla\varphi(\mathbf{r})) = \varrho_p(\mathbf{r}), \tag{5.10}$$

where $\kappa_r(\mathbf{r})$ is the relative dielectric constant of the materials in the system; κ_0 is the permittivity of the vacuum. Strain and the piezoelectric polarisation interact with each other and thus both of them should ideally be determined within a self-consistent calculation [50]. However, as the influence of this coupling is of secondary importance in most materials [7, 12], we treat the strain and the piezoelectric polarisation as being independent of each other here.

A plane-wave approach can be used for each of the three parts of the calculation discussed above, namely calculation of the strain by minimisation of the elastic energy, the polarisation potential from the Poisson equation and finally, using these ingredients, calculation of the electronic structure of the semiconductor nanostructure. We will now demonstrate how these calculations can be performed semi-analytically as well as numerically and point out the specific features that arise from a plane-wave representation.

5.2.1 Semi-Analytical Plane-Wave Approaches

In this subsection, we introduce the plane-wave method through its application to a number of problems which can be set up analytically. The energy levels of the electrons and holes in the nanostructures are then calculated numerically by finding eigenvalues and eigenvectors of a matrix which is set up analytically. We consider first the calculation of confined states in a quantum well using a one-band effective mass Hamiltonian, showing also how the confined state energies and wavefunctions converge with the number of plane waves considered. We then extend the model to

consider a cube-shaped QD, showing again rapid convergence of the ground state energy with increasing number of plane waves. Secondly, we show how the plane-wave expansion method can be easily generalised to treat a multi-band Hamiltonian. Finally we present examples of some of the wide range of QD shapes for which the Fourier transform of the characteristic function $\tilde{\chi}(\mathbf{k})$ can be calculated analytically, thereby facilitating straightforward set up of the Hamiltonian for the nanostructure being considered.

5.2.1.1 Single-Band Plane-Wave Formulation

We consider the application of a single band effective mass model to a quantum well of width d placed at the centre of a supercell of length L, defined between $-\frac{L}{2}$ and $\frac{L}{2}$. The Schrödinger equation in the effective mass approximation (EMA) is then given by:

$$H_{\mathrm{EMA}}(x)\psi(x) = \left(-\frac{\hbar^2}{2}\frac{d}{dx}\left[\frac{1}{m_e^*(x)}\frac{d}{dx}\right] + V(x)\right)\psi(x) = \varepsilon\psi(x) \tag{5.11}$$

where $m_e^*(x)$ is the spatially dependent effective mass in the supercell and $V(x)$ is the potential, equal to $-V_0$ in the well and to 0 outside the well. The quantum well characteristic function is given by

$$\chi_1(x) = 1 \ \text{ for } \ -\frac{d}{2} \leq x \leq \frac{d}{2} \ , \tag{5.12}$$

and equals 0 otherwise. The Fourier transform of the characteristic function is then given by

$$\overline{\chi}_1(q) = \frac{1}{L}\int_{-d/2}^{d/2} e^{-iqx}dx = \frac{2}{qL}\sin\left(\frac{qd}{2}\right) \ . \tag{5.13}$$

We note that in this chapter we use two notations for the Fourier Transform of a function $f(x)$, $\overline{f}(q)$ and $\tilde{f}(q)$:

$$\overline{f}(q) = \frac{1}{L}\int_{-L/2}^{L/2} e^{-iqx}dx \ , \tag{5.14}$$

$$\tilde{f}(q) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-iqx}dx \tag{5.15}$$

with obvious generalisation for two and three dimensions; it is also clear that for periodic functions we have $\tilde{f}(q) = (L/2\pi)\overline{f}(q)$.

For the one-dimensional supercell of length L, the pth wavefunction with supercell wavenumber $q = 0$ is given from Eq. (5.4) by

$$\psi_{p0} = \frac{1}{\sqrt{L}} \sum_{n=-M}^{M} c_{pn} e^{iG_m x} \tag{5.16}$$

where $G_m = \frac{2\pi m}{L}$ and we include $2M + 1$ plane waves in the sum.

The Schrödinger equation $H_{\text{EMA}}\psi_{p0} = E_{p0}\psi_{p0}$ can then be solved by substituting the expansion from (5.16) into (5.11). We can then multiply both sides of (5.11) from the left by $\exp(-iG_i x)$ ($i = -M, \ldots M$) and integrate over the supercell length to give $2M + 1$ linear equations

$$\sum_{j=-M}^{M} H_{ij} c_{pj} = \varepsilon c_{pi} \tag{5.17}$$

where H_{ij} is given by

$$H_{ij} = \frac{\hbar^2 G_i G_j}{2m_b^*}\delta_{ij} + \overline{\chi}_1(G_i - G_j)\left[\frac{\hbar^2 G_i G_j}{2}\left(\frac{1}{m_w^*} - \frac{1}{m_b^*}\right) - V_0\right], \tag{5.18}$$

where we have assumed that the effective mass inside the well is m_w^* and outside the well (in the barrier) it is m_b^*. We can then calculate the solutions to the Schrödinger equation by finding the eigenvalues and eigenvectors of the $(2M + 1) \times (2M + 1)$ Hamiltonian matrix, H_{ij}.

Figure 5.1 shows the evolution of the calculated lowest confined conduction band state energies in a model InAs/GaAs quantum well structure for well width $d = 10$ nm, supercell period $L = 20$ nm, with band offset $V_0 = 700$ meV, and with electron relative mass $m_e = 0.023$ in the well and in the barrier. It can be seen that the three confined state energies converge very rapidly with increasing number of plane waves. Figure 5.2 shows that even the use of five plane waves ($M = 2$) can provide a good approximation to the ground state wavefunction in this case, consistent with the rapid convergence of the calculation.

The semi-analytical plane-wave method can be readily extended to treat three-dimensional nanostructures. We consider as example a cuboidal QD in a supercell of volume $L \times L \times L$ and for which the QD characteristic function is given by $\chi_3(x, y, z) = 1$ for $|x|, |y|, |z| \leq \frac{d}{2}$. The Fourier transform of the characteristic function $\tilde{\chi}_3(k_x, k_y, k_z)$ is given in this case by

$$\tilde{\chi}_3(k_x, k_y, k_z) = \tilde{\chi}_1(k_x) \times \tilde{\chi}_1(k_y) \times \tilde{\chi}_1(k_z)\,, \tag{5.19}$$

where $\tilde{\chi}_1(k_i)$ is given by (5.13). Figure 5.3 shows the calculated convergence of the lowest confined state energies in a model InAs/GaAs cuboid QD with side length $d = 10$ nm in a supercell with $L = 20$ nm as a function of increasing number of plane waves. The calculations in this case were carried out using a cubic array of reciprocal lattice vectors $\mathbf{G}$, with k_x, k_y and k_z each running from $-M$ to $+M$. It can be seen that both the ground state and lowest excited states again

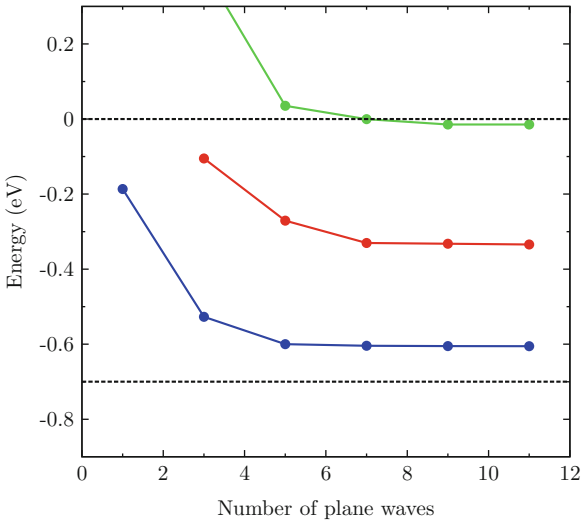


Fig. 5.1 Calculated localised energy levels in a 10 nm wide quantum well (QW) as a function of the total number of plane waves $(2M + 1)$. We used $L = 20$ nm; $V_0 = 0.7$ eV; relative mass $m_e = 0.023$ which is similar to the parameters for an InAs/GaAs QW without strain

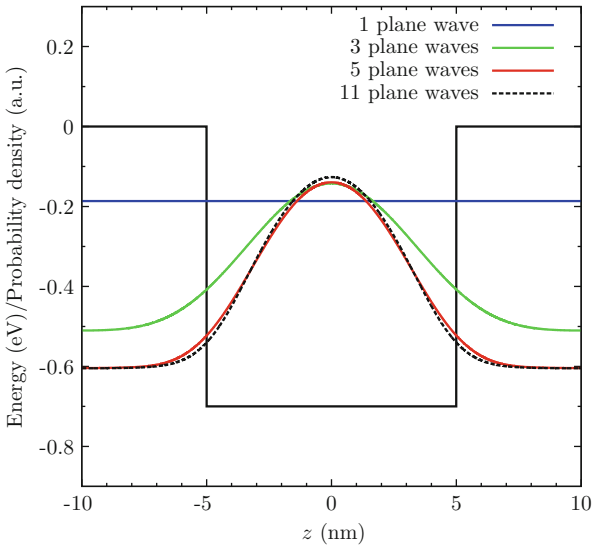


Fig. 5.2 Calculated wavefunction for the ground state level for the 10 nm QW for different values of total number of plane waves $(2M + 1)$; using the same parameters as in Fig. 5.1

converge rapidly with increasing M, confirming the value of a plane-wave approach to electronic structure calculations in semiconductor nanostructures.

5.2.1.2 Multi-Band Plane-Wave Formulation

In the envelope function method the carrier states in a nanostructure are calculated by solving a Schrödinger-like equation with an effective Hamiltonian: $H^{n\times n}\Psi = \varepsilon\Psi$. The size n and form of $H^{n\times n}$ differ depending on the particular multi-band

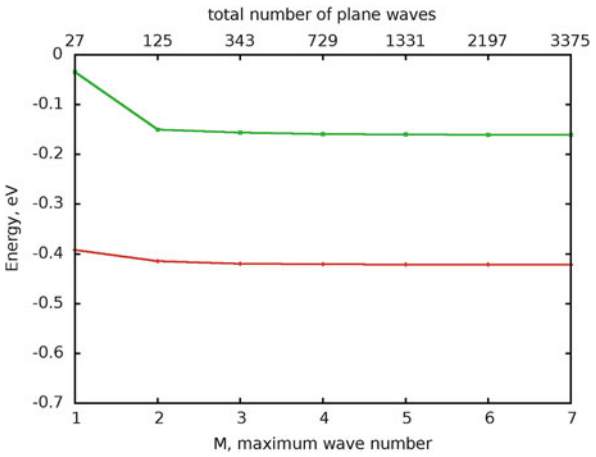


Fig. 5.3 Calculated localised energy levels in a 10 nm cubic QD versus maximum number of plane waves M; top axis shows the total number of the plane waves used, $(2M+1)^3$. We used $L = 20$ nm; $V_0 = 0.7$ eV; $m_e = 0.023$ which is similar to the parameters for an InAs/GaAs QD without strain

Hamiltonian which is chosen. From a mathematical point of view this form of the Schrödinger equation is equivalent to a system of coupled differential equations. The idea of using a plane-wave method is that it allows us to solve this system using a Fourier transform technique. From a physical point of view this corresponds to describing the carrier states in terms of a linear combination of a suitably chosen set of bulk states [3] associated with a periodic array of bulk wavevectors. The effective Hamiltonian can then be naturally represented in the form given earlier in (5.6)

$$H^{n\times n} = H_0^{n\times n} + V^{n\times n} \tag{5.20}$$

where the "perturbation" $V^{n\times n}$ describes the difference between the potential in the quantum structure considered and the potential in the bulk Hamiltonian $H_0^{n\times n}$ used for the basis states.

Two different approaches are possible to choose the plane-wave basis states. The zone-centre states in an n-band $\mathbf{k}\cdot\mathbf{p}$ Hamiltonian are given by $u_\alpha(\mathbf{r})$, where $\alpha = 1,\ldots,n$ labels the zone-centre Bloch states associated for instance in an 8-band model with conduction electron, heavy-hole, light-hole and spin-split-off bands. As we move to wavevectors away from the zone centre, there are at least two choices as how to choose the basis states associated with wavevector $\mathbf{G_m} = \mathbf{p}_i - \boldsymbol{\xi}_\mathbf{m}$, where $\mathbf{p}$ is the "quasi-momentum" label for the 3D superlattice of QDs, $(\boldsymbol{\xi}_\mathbf{m})_i = 2\pi m_i/L_i$ and $\mathbf{m}$ is the plane-wave number. In the first approach, we choose as basis states $u_\alpha(\mathbf{r})\exp(i\mathbf{G_r}\cdot\mathbf{r})$. This use of the zone-centre basis for all $\mathbf{G_m}$ is further considered in the next section, when we discuss fully numerical approaches to calculating the eigenstates and eigenenergies.

It can however also be useful to choose as basis states the eigenstates of the bulk $\mathbf{k}\cdot\mathbf{p}$ Hamiltonian at each wavevector $\mathbf{G_m}$. The basis states in this case then take the form:

$$\Psi_{\mathbf{p},\mathbf{m},S}(\mathbf{r}) = \frac{1}{\sqrt{L_1L_2L_3}}\sum_{\alpha=1}^{n} B_\alpha^S(\mathbf{p},\mathbf{k})u_\alpha(\mathbf{r})\exp(i\mathbf{G_m}\cdot\mathbf{r}) \tag{5.21}$$

where S denotes the type of state (including spin), i.e. conduction electrons, light, heavy or spin-split-off holes. This approach has the advantage that it allows if required to treat e.g. the heavy- and light-hole bands to a higher degree of accuracy than the more remote spin-split-off valence band. The operator matrix $V^{n\times n}$ in (5.20) should be obtained from the bulk-like Hamiltonian by making the substitution $k_j \rightarrow -i\partial/\partial x_j$, to take account of the spatial dependence of the band parameters. Details of the interface boundary conditions are included by an appropriate application of the differential operators at each interface. The ith solution, $\Psi_{\mathbf{p}i}(\mathbf{r})$ of the initial Schrödinger equation with quasi-momentum $\mathbf{p}$ is then found by expanding in a series with respect to the plane waves in (5.21):

$$\Psi_{\mathbf{p}i}(\mathbf{r}) = \sum_S \sum_{\mathbf{m}} C^i_{\mathbf{p},\mathbf{m},S} \Psi_{\mathbf{p},\mathbf{m},S}(\mathbf{r}) \tag{5.22}$$

where the summation over S takes into account such effects as light- and heavy-hole mixing in heterostructures. The summation over $\mathbf{m}$ is carried out over the chosen set of wave vectors in reciprocal space, which is determined by the number of plane waves taken into account for the expansion. In a simple case, the range of wave numbers $\mathbf{m}$ taken into account in the above sum for a given bulk state S is given by an extension of the 1-band approach, by setting $-M_{i,S} < m_i < M_{i,S}$, where $M_{i,S}$ is the maximum wave number for a given spatial direction i ($i = 1, 2, 3$ for x, y, z directions). It is also possible to determine the wavenumbers taken into account in the expansion by using the cut-off energy criterion discussed in the beginning of this chapter, which is equivalent to requiring $\sum_{i=1}^{3} |m_i|^2 \leq |M|^2$. Having chosen the wavevector range, the next step is to obtain the matrix $\mathbf{A}$, whose eigenvectors and eigenvalues are the coefficients $C^i_{\mathbf{p},\mathbf{m},S}$ and the energy spectrum ε_i of the QD. This matrix has the form:

$$A_{i'i} = E_S(\mathbf{p} - \boldsymbol{\xi}_{\mathbf{n}})\delta_{S'S}\delta_{\mathbf{n}',\mathbf{n}} + \sum_{\alpha'=1}^{n}\sum_{\alpha=1}^{n}[B^{i'}_{\alpha'}]^* B^i_\alpha \overline{V}^{i'i}_{\alpha',\alpha}(\mathbf{m},\mathbf{m}') \tag{5.23}$$

where the numbers i' and i denote the set of quantum numbers $(\mathbf{p}, \mathbf{m}, S)$, $E_S(\mathbf{k})$ is the energy dispersion of the bulk state of type S; and $\overline{V}_{\alpha',\alpha}(\mathbf{m},\mathbf{m}')$ is the Fourier transform of $V_{\alpha',\alpha}$:

$$\overline{V}_{\alpha',\alpha}(\mathbf{m},\mathbf{m}') = \frac{1}{L_1 L_2 L_3}\int_{\Omega_0} e^{i\boldsymbol{\xi}_{\mathbf{m}'}\mathbf{r}} V_{\alpha',\alpha}(\mathbf{r}) e^{-i\boldsymbol{\xi}_{\mathbf{m}}\mathbf{r}} dV. \tag{5.24}$$

The matrix elements $V_{\alpha',\alpha}(\mathbf{r})$ depend on real-space coordinate $\mathbf{r}$ because they depend on the applied external field, piezoelectric field, spontaneous electric field and on strain. In the multi-band Hamiltonian the matrix elements can be taken to depend *linearly* on the strain tensor components and on the built-in electric potential. Therefore the Fourier transform of $V_{\alpha',\alpha}$ is expressed through the Fourier transform of the strain tensor, built-in electric potentials and the QD characteristic function

$\chi_{QD}(\mathbf{r})$, as we discuss in more detail below. This means that using a plane-wave expansion method in conjunction with the techniques presented in this chapter, there is no need to calculate the full spatial distributions of the strain and built-in electric fields, unlike in other methods. This highly convenient and simple trick considerably reduces the time taken to set up calculations and makes the plane-wave method very efficient for the further study of QD optical properties and for modelling QD devices.

We also note that the number of bulk states (plane waves) which must be included to obtain a given level of accuracy is reduced in periodic structures with partly coupled QDs. The upper limit for the number of bulk states which can be used in the i-th direction (maximum wave number), M_i^{max}, is set, in the framework of the envelope-function approximation, by the number of the atomic layers in the period of the QD superlattice in this i-th direction: $M_i^{max} = [L_i/(a_0)_i]$ (where the square brackets indicate the integer part and $(a_0)_i$ is the average lattice constant in the i-th direction). We note that the total number of plane waves in the i-th direction is then $(2M_i^{max} + 1)$, as the summation over wave numbers m_i is carried out from $-M_i^{max}$ to $+M_i^{max}$. The envelope-function approximation (EFA) is valid only when the envelope function varies slowly over distances of order of the lattice constant. This means that in the expansion of the envelope function in the Fourier series all terms with large wave vectors, $k_i > 2\pi/(a_0)_i$, should be neglected, since in the envelope function approximation they must be negligibly small. Therefore we get the estimate presented above for the maximum number of plane waves which can be taken into account in (5.22). From the other side, this also gives a measure of the applicability of the EFA for calculation of the carrier spectrum and wavefunctions in an actual QD structure. If the number of the plane waves required in (5.22) to calculate with enough accuracy the wavefunction and carrier energy of a particular level is less than the total maximum number of plane waves $(2M_i^{max} + 1)$, then the EFA is valid for that level. This is because the terms in (5.22) with large wave vectors, $k_i > 2\pi/(a_0)_i$, which should be thrown away, do indeed then make a negligibly small contribution to the solution of the Schrödinger-like equation $H^{n\times n}\Psi = \varepsilon\Psi$. It turns out for most real QD structures in direct-gap materials that the number of plane waves required to achieve convergence is less than $(2M_i^{max} + 1)$ and therefore for these structures the EFA is valid.

5.2.1.3 Fourier Transforms for Various QD Shapes

The Fourier transform of many common dot shapes can be calculated analytically. These shapes include a sphere, a cuboid, a pyramid and a truncated pyramid, a cylinder, a hemisphere and a cone [1]. Fourier transforms can also be calculated for more complicated characteristic functions, including for instance a truncated pyramid with a linear composition gradient along the growth direction [5]. The semi-analytical approach has the major advantage that a Hamiltonian matrix such as that using (5.18) or (5.19) can be set up efficiently using analytical expressions, with

the cost however that different calculations may require input of different analytical expressions. In addition we shall see below that the analytical approach typically assumes linear variation of properties with composition and, when calculating strain and piezoelectric properties also assumes that the elastic constants and piezoelectric coefficients have the same values in the dot and surrounding matrix materials. These assumptions are usually acceptable, but it can be useful for additional flexibility to also consider approaches where the Fourier transforms are calculated numerically, as we discuss further in the next section.

We have already presented in (5.19) the form of the Fourier transform for a cuboidal QD. As discussed above, simple expressions can also be obtained for a wide range of other dot shapes. We present examples of the Fourier transform for two other dot shapes here. For a hemisphere with the centre of its base located at the origin, it can be shown that

$$\tilde{\chi}_{QD}(\boldsymbol{\xi}) = \frac{1}{2}\tilde{\chi}_{QD}^{sphere}(\boldsymbol{\xi}) + \frac{2\pi i}{\xi_3}\left\{\frac{R}{\xi_{||}}J_1(R\xi_{||}) - R^2 I_{j0}(R\xi_3, R\xi_{||})\right\} \tag{5.25}$$

where $\tilde{\chi}_{QD}^{sphere}$ is the Fourier transform for the sphere, R is the hemisphere radius, $\xi_{||} = \sqrt{\xi_1^2 + \xi_2^2}$ and I_{j0} denotes the integral

$$I_{j0}(\alpha, \beta) = \int_0^1 x \cos\left(\alpha\sqrt{1-x^2}\right) J_0(\beta x)\, dx, \tag{5.26}$$

where J_0 is a Bessel function. The integral in (5.26) is best calculated numerically.

For a cone with the centre of its base at the origin,

$$\tilde{\chi}_{QD}(\boldsymbol{\xi}) = \frac{2\pi i}{\xi_z}\left\{e^{-i\xi_z h} I_2(\xi_{||}R, \xi_z h) - \frac{R}{\xi_{||}}J_1(\xi_{||}R)\right\} \tag{5.27}$$

where R is the radius of the cone base, h is the cone height, J_1 is a Bessel function and I_2 denotes the integral

$$I_2(\alpha, \beta) = \int_0^1 x J_0(\alpha x) e^{i\beta x}\, dx. \tag{5.28}$$

This integral can be expressed as a power series or calculated numerically.

5.2.2 *Numerical Plane-Wave Approaches*

In this section, we will demonstrate an efficient numerical approach to solve the Schrödinger equation on the basis of a plane-wave framework. The approach that will be presented makes use of the computational efficiency of a plane-wave based

formulation, where the plane-wave approach retains all the real-space properties of the system under consideration, without the need for any simplifications or symmetry considerations regarding the shape or the material composition of the system being investigated.

5.2.2.1 Multi-Band k·p Approaches in a Plane-Wave Formulation

When we extend the Hamiltonian from the 1-band model of (5.11) to an n-band **k·p** model, the n-band wavefunction $\Psi^j(\mathbf{r})$ with j being the state (level) number can be formulated, as discussed in Sect. 5.2.1.2 as:

$$\Psi^j(\mathbf{r}) = \sum_{\alpha,\mathbf{m}} c^j_{\alpha,\mathbf{m}} e^{i\mathbf{G_m}\cdot\mathbf{r}} u_\alpha(\mathbf{r}) \tag{5.29}$$

where α denotes the component of the basis and $\mathbf{m}$ labels the plane waves included in the calculation. The matrix elements linking any two elements in this basis are then given by (5.24). One advantage which can be gained in a numerical application of the plane-wave approach is that any gradient operator in the Hamiltonian can be carried out analytically in Fourier space, with direct numerical calculation then only required for a subset of the terms in the Hamiltonian.

More details regarding the implementation of the numerical approach are provided for the example of an 8-band **k·p** Hamiltonian for the wurtzite lattice in [27]. A vastly generalised implementation of this technique, where the Hamiltonian is not hard-wired in the source code, but is a user-generated input file, is available within the multi-band **k·p** module of the S/Phi/nX software library [9, 27].

5.2.2.2 Electronic Minimisation Schemes for Multi-Band k·p Models

The algorithms available to numerically solve the Schrödinger equation in a plane-wave based multi-band **k·p** model and to determine electronic states and energy levels are a complex matter and thus deserve to be discussed within this section. The multi-band **k·p** model introduced in [27] has been implemented into the existing plane-wave library S/Phi/nX. This allows the highly optimised electronic minimisation schemes that are commonly available in such libraries to be employed with only minor modifications, and therefore highlights another advantage of a plane-wave based implementation. Reference [27] outlines in detail how the well-established and highly efficient conjugate-gradient (CG) minimisation scheme [33] can be modified to suit the specific requirements of multi-band **k·p** models. However, as we aim here just to outline the principles of such methods, we will in the following limit ourselves to outline the simpler steepest-descent scheme for energy minimisation.

Steepest Descent Scheme for Energy Minimisation

Within a minimisation problem that aims to finding a minimum energy of a system, a steepest descent scheme works as follows:

1. The energy to be minimised with respect to the wavefunction $|\Psi_i\rangle$ of the i-th step of the iterative minimisation is calculated via:

$$\varepsilon_i = \langle \Psi_i(\mathbf{G}) | H | \Psi_i(\mathbf{G}) \rangle.$$

2. In the next step, the search direction $|\xi_i(\mathbf{G})\rangle$ is determined with Δ_t being the step size:

$$|\xi_i(\mathbf{G})\rangle = \Delta_t \cdot (H|\Psi_i(\mathbf{G})\rangle - \varepsilon_i |\Psi_i(\mathbf{G})\rangle).$$

3. Finally, the updated wavefunction $|\Psi_{i+1}(\mathbf{G})\rangle$ is computed:

$$|\Psi_{i+1}(\mathbf{G})\rangle = |\Psi_i(\mathbf{G})\rangle - |\xi_i(\mathbf{G})\rangle.$$

This scheme is repeated until a convergence is achieved, i.e. $\varepsilon_{i+1} - \varepsilon_i$ drops below a given energy difference, or until a maximum number of iteration steps has been performed. In the latter case, the minimisation ends without having obtained a converged energy ε_i.

Further eigenvalues and wavefunctions can then be calculated by following a similar approach, while including the additional constraint that the eigenvector currently being calculated must be orthogonal to all previously calculated eigenvectors.

Steepest Descent Scheme for a Multi-band **k·p** Model

The above minimisation scheme needs to be modified to suit the specific requirements of a multi-band **k·p** model. More complex and efficient minimisation schemes such as the CG algorithm can be modified in a similar manner (see [27] for more details). First of all, the property to be minimised in a multi-band **k·p** approach is not an absolute eigenenergy of an electron or hole state. Instead, the energy with respect to the band gap is to be minimised. To illustrate this behaviour, we choose the example of an InAs quantum dot buried in GaAs, whose electronic properties, i.e. electron and hole states and eigenenergies, are to be determined using an 8-band **k·p** model. This system represents a type I nanostructure, i.e. the conduction band has its minimum inside the InAs dot, where the valence band also has its maximum (See Fig. 5.4). This means that while the minimisation of electron eigenenergies requires to search for a minimum above the conduction band edge energy, the search for hole eigenenergies requires to find a state with energy just below the valence band edge energy. What needs to be minimised is thus not the energy of the state itself, but rather the energy difference between the given state and the conduction or valence

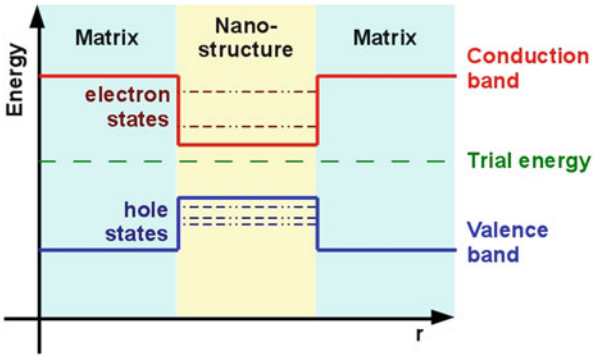


Fig. 5.4 Conduction and valence band offsets of a type I nanostructure. The trial energy is indicated by the *green dashed line* and electron and hole states are shown as *red and blue dash-dotted lines*

band edge energy of the strained dot material. Consequently, we make two changes to the steepest-descent scheme described above:

1. Instead of minimising ε, we minimise $(\varepsilon - \varepsilon_t)^2$, where ε_t is a trial energy, chosen such that the states to be calculated (i.e. electrons or holes) are the ones closest to ε_t. The value of ε_t can be treated as an arbitrary parameter in order to calculate states in a certain energy region (i.e. higher excited states) without calculating all states below.
2. The Hamiltonian H is replaced by $(H - \varepsilon_t)^2$, to allow for the calculation of a minimum energy difference with respect to the trial energy [44].

For a multi-band **k·p** model, the modified minimisation scheme is now:

1. First of all, $(H - \varepsilon_t)^2 |\Psi_i(\mathbf{G})\rangle$ is calculated as:

$$(H - \varepsilon_t)^2 |\Psi_i(\mathbf{G})\rangle = H^2 |\Psi_i(\mathbf{G})\rangle - 2\varepsilon_t \cdot H |\Psi_i(\mathbf{G})\rangle + \varepsilon_t^2 \cdot |\Psi_i(\mathbf{G})\rangle.$$

2. Correspondingly, the search direction $|\xi_i(\mathbf{G})\rangle$ is now determined as:

$$|\xi_i(\mathbf{G})\rangle = \Delta_t \cdot \left((H - \varepsilon_t)^2 - (\varepsilon - \varepsilon_t)^2\right) |\Psi_i(\mathbf{G})\rangle$$

$$= \Delta_t \cdot \left[H^2 |\Psi_i(\mathbf{G})\rangle - |\Psi_i(\mathbf{G}\rangle\langle\Psi_i(\mathbf{G})| H^2 |\Psi_i(\mathbf{G})\rangle\right.$$

$$\left. -2\varepsilon_t \left(H |\Psi_i(\mathbf{G})\rangle + |\Psi_i(\mathbf{G})\rangle\langle\Psi_i(\mathbf{G})| H |\Psi_i(\mathbf{G})\rangle\right)\right].$$

3. The final step to compute the updated wavefunction $|\Psi_{i+1}(\mathbf{G})\rangle$ remains as in the above scheme:

$$|\Psi_{i+1}(\mathbf{G})\rangle = |\Psi_i(\mathbf{G})\rangle - |\xi_i(\mathbf{G})\rangle.$$

With these modifications, the steepest-descent minimisation scheme can be employed for multi-band **k·p** calculations and more optimised schemes can be adapted to this formalism in a similar manner.

5.3 Strain Distribution in a Plane-Wave Formulation

The elastic properties of an epitaxially grown semiconductor nanostructure can be calculated both analytically and numerically, in a similar manner to the electronic properties. We first consider here the analytical approach, following which we describe one of the numerical methods which can be used to determine the elastic properties.

5.3.1 Analytical Approach

We have presented a method based on the Green's function technique to calculate the strain in QD nanostructures [1,2]. An analytical formula in the form of a Fourier series was obtained for the strain tensor for arrays of QDs of arbitrary shape taking into account the anisotropy of the elastic properties. The model assumed equal elastic constants in the dot and in the surrounding matrix material. The method gave tractable expressions for the strain distribution in materials with cubic [1] and wurtzite [2] crystal structures, with the expression for the strain distribution taking a particularly simple form in the case of materials with isotropic elastic properties [15].

The Green's tensor $G_{ln}(\mathbf{r})$ gives the displacement at $\mathbf{r}$ in the direction l due to a unit point force along direction n placed at the origin. The Green's tensor for an infinite anisotropic elastic medium [26] is the solution of the equation:

$$C_{iklm}\frac{\partial G_{ln}}{\partial x_k \partial x_m} = -\delta(\mathbf{r})\delta_{in} \ , \tag{5.30}$$

with the boundary condition $G_{ln} \rightarrow 0$ as $|\mathbf{r}| \rightarrow \infty$. In (5.30) C_{iklm} is the tensor of the elastic moduli.

To find $G_{ln}(\mathbf{r})$, we first take the Fourier transform of (5.30), which gives

$$C_{iklm}\xi_k \xi_m \tilde{G}_{ln}(\boldsymbol{\xi}) = \frac{\delta_{in}}{(2\pi)^3} \ . \tag{5.31}$$

The method of inclusions as proposed by Eshelby [17] is used to find the strain distribution in the QD structure. The displacement in a medium containing a single QD, $\mathbf{u}^s(\mathbf{r})$, can be expressed as the convolution of the Green's tensor and the forces spread over the QD surface,

$$u_i^s(\mathbf{r}) = u_i^0 \chi_{QD}(\mathbf{r}) + \int G_{in}(\mathbf{r}-\mathbf{r}')\sigma_{nk}^0(\mathbf{r}') \ dS_k' \ , \tag{5.32}$$

where $\chi_{QD}(\mathbf{r})$ is the characteristic function of the QD, equal to unity within the QD and zero outside; $\sigma_{nk}^0 = C_{nkpr}\epsilon_{pr}^0$ and σ_{nk}^0, ϵ_{pr}^0 and u_i^0 are the components of the

stress and strain tensors and the displacement caused by the "initial" strain due to the lattice mismatch. The superscript "s" indicates that this expression refers to a single QD.

The integration in (5.32) is carried out over the QD surface. Using Gauss's theorem to covert the surface integral into an integral over the QD volume, it can be shown that the strain tensor in a single QD structure is given by

$$\epsilon_{ij}^{s}(\mathbf{r}) = \epsilon_{ij}^{0}\chi_{QD}(\mathbf{r}) + \frac{1}{2}\int_{QD}\left[\frac{\partial G_{in}(\mathbf{r}-\mathbf{r}')}{\partial x_j\,\partial x_k} + \frac{\partial G_{jn}(\mathbf{r}-\mathbf{r}')}{\partial x_i\,\partial x_k}\right] C_{nkpr}\epsilon_{pr}^{0}dV', \tag{5.33}$$

where the integration is carried out over the QD volume. Using the convolution theorem and then taking the Fourier transform gives

$$\tilde{\epsilon}_{ij}^{s} = \epsilon_{ij}^{0}\tilde{\chi}_{QD}(\boldsymbol{\xi}) - \frac{(2\pi)^3}{2}\left\{\xi_i\tilde{G}_{jn}(\boldsymbol{\xi}) + \xi_j\tilde{G}_{in}(\boldsymbol{\xi})\right\} C_{nkpr}\xi_k\epsilon_{pr}^{0}\tilde{\chi}_{QD}(\boldsymbol{\xi})\ , \tag{5.34}$$

where $\tilde{\chi}_{QD}(\boldsymbol{\xi})$ is the Fourier transform of the QD characteristic function. Equation (5.34) gives the general expression for the Fourier transform of the strain tensor in a structure containing a single QD of arbitrary shape. This is a general formula valid for crystals of cubic or any other symmetry. Note that the QD shape enters only as the Fourier transform of the QD characteristic function.

The expression for an isolated QD in (5.34) can readily be extended to the 3D array of QDs implied by a supercell calculation. We have shown how to find explicit expressions for the Fourier transform of the Green's function $\tilde{G}_{jn}(\boldsymbol{\xi})$ in terms of the elastic constants C_{ij} both for cubic [1] and for wurtzite [2] structures, as well as for isotropic materials. In the isotropic case, the elastic tensor components are related by $C_{11} = C_{12} + C_{14}$, with all other tensor components equal to zero. In this case, the expressions for the Fourier transforms of the strain components, $\tilde{\epsilon}_{ij}^{\text{iso}}$, take a particularly simple form, with

$$\tilde{\epsilon}_{ij}^{\text{iso}}(\boldsymbol{\xi}) = \epsilon_0\tilde{\chi}_{QD}(\boldsymbol{\xi})\left(\delta_{ij} - \frac{3\lambda+2\mu}{\lambda+2\mu}\frac{\xi_i\xi_j}{\xi^2}\right), \tag{5.35}$$

where $\lambda = C_{12}$ and $\mu = C_{44}$ are the Lamé constants for an isotropic elastic medium. We see from (5.35) firstly that the Fourier transform can be easily set up analytically for an isotropic material, once the Fourier transform of the characteristic function is known. It can also be shown from (5.35) that the hydrostatic component of the strain tensor, $\epsilon_h \equiv \sum_{i=1}^{3}\epsilon_{ii}^{\text{iso}}$, is constant inside the QD and zero outside in the isotropic approximation and given by

$$\epsilon_h^{\text{iso}} = \epsilon_0\frac{4\mu}{\lambda+2\mu}\chi_{QD}(\mathbf{r}). \tag{5.36}$$

Thus, the deviation of the hydrostatic strain from this constant value is characteristic of the influence of elastic anisotropy on the strain distribution in QD structures.

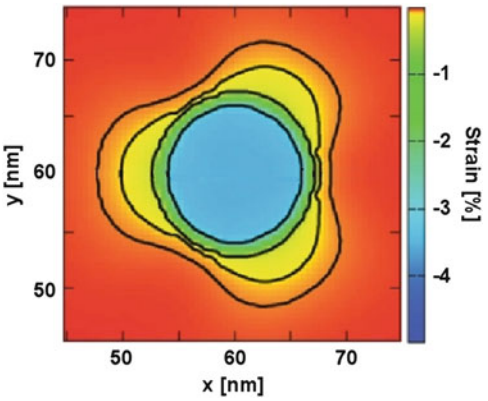


Fig. 5.5 Trace of the strain tensor in an InGaAs QD in a plane perpendicular to the (111) direction through the bottom of the QD. The hydrostatic strain is approximately constant inside the dot outside it shows a small deviation from zero, displaying a C_{3v} symmetry reflecting underlying anisotropy in the (111)-rotated elastic tensor

Calculations for a range of QD structures show that the hydrostatic strain tends to be large and approximately constant within a dot, and close to zero outside the dot. This is true both for (001)-oriented zincblende dots [1] and for wurtzite dots [2]. It is also true for (111)-oriented zincblende dots, as illustrated in Fig. 5.5 for an InGaAs model QD with a base length of 13 nm and a height of 3 nm. It can be seen that the hydrostatic strain is approximately constant inside the dot, and that outside it shows a small deviation from zero, displaying a C_{3v} symmetry that reflects the underlying anisotropy in the material elastic properties.

5.3.2 Numerical Approach

The elastic properties can be calculated numerically following a similar approach to that used for the electronic properties. We calculated the electronic states within an n-band $\mathbf{k}\cdot\mathbf{p}$ model, by requiring that square of the difference between the eigenenergy ε of a particle and a trial energy ε_t, $(\varepsilon - \varepsilon_t)^2$, is minimised with respect to the Bloch components of the wavefunction Ψ_α with $\alpha = 1, 2, \ldots, n$. In a similar way, the elastic problem requires to minimise the elastic energy F with respect to the displacements u_i with $i = 1, 2, 3$. This can be done using a calculation scheme analogous to that used for calculation of the electronic structure. For this purpose, the variational derivative of the elastic energy F in (5.7) can be expressed by a set of differential equations [35]:

$$\frac{\delta F[u_1(\mathbf{r}), u_2(\mathbf{r}), u_3(\mathbf{r})]}{\delta u_j(\mathbf{r})} = \frac{\partial}{\partial r_i}\left(C_{ijkl}(\mathbf{r})\left[\frac{\partial u_k(\mathbf{r})}{\partial r_l} + \epsilon^0_{kl}(\mathbf{r})\chi(\mathbf{r})\right]\right) \tag{5.37}$$

One can now again make use of the computational efficiency of a plane-wave based formulation, by following the same scheme applied to the gradient $H|\Psi(\mathbf{G})\rangle$ in

the calculation of electronic properties within a multi-band $\mathbf{k} \cdot \mathbf{p}$ model. Within a plane-wave framework, the gradient in (5.37) reads in reciprocal space as:

$$\frac{\delta F[u_1(\mathbf{G}), u_2(\mathbf{G}), u_3(\mathbf{G})]}{\delta u_j(\mathbf{G})} = \hat{\mathscr{G}}\left[\hat{\mathscr{R}}\left(\hat{\mathscr{G}}\left\{C_{ijkl}(\mathbf{r})\left[\hat{\mathscr{R}}\left(i\,\mathbf{G}_l u_k(\mathbf{G})\right) + \epsilon^0_{kl}(\mathbf{r})\right]\right\} i\,\mathbf{G}_i\right)\right]. \tag{5.38}$$

where we the operator $\hat{\mathscr{R}}$ denotes the discrete Fourier transformation of a function $f(\mathbf{G})$ from reciprocal space to real space. The operator $\hat{\mathscr{G}}$ transforms a wavefunction from real space to reciprocal space, correspondingly [27].

We find again for (5.38) that the real-space properties are kept without simplifications and computationally cheap gradient operations can be performed in reciprocal space with ∇ being replaced by the factor $-i\,\mathbf{G}$. The formal similarity of the gradients in the n-band Hamiltonian problem and in the case of minimisation of the elastic energy allows to employ iterative minimisation schemes such as the steepest-descent scheme for the numerical solution of the elastic problem, in a similar manner as for the solution of the electronic structure in a multi-band $\mathbf{k}\cdot\mathbf{p}$ model. For practical implementation, a number of differences between the two calculations need however to be taken into account:

1. While the electronic wavefunctions $|\Psi_i\rangle$ need to be normalised such that $\langle\Psi_i|\Psi_j\rangle = \delta_{ij}$, a similar normalisation of the displacements $\mathbf{u}$ within the solution of the elastic problem must not be enforced.
2. The electronic wavefunction $|\Psi\rangle$ is a complex vector. The displacements $\mathbf{u}$ on the other hand are physical observables that do not have any imaginary components.
3. Sharp material interfaces, e.g. between a GaN QD and an AlN matrix, and the associated discontinuities in displacement vectors, can cause difficulties in the solution of the elastic problem in a plane-wave framework, as the displacements and the resulting strains are described using a finite number of plane waves. In the vicinity of sharp interfaces, this will result in artificial oscillations of these properties that leads to artificial band offsets when incorporated in a $\mathbf{k}\cdot\mathbf{p}$ model and thus finally in erroneous electronic states. However, this problem can be overcome by a slight softening of the interfaces, e.g. using a diffusion equation. For the bulk electronic properties, however, sharp interfaces do not cause such problems, as the wavefunction of a particle will nevertheless exhibit a smooth behaviour without abrupt changes at the interfaces.
4. The elastic energy $F[\mathbf{u}]$ calculation involves minimising the absolute elastic energy, and therefore does not require the trial energy modifications introduced for the multi-band $\mathbf{k}\cdot\mathbf{p}$ model calculations.

The strain tensor $\epsilon_{ij}(\mathbf{r})$ is then calculated from the real-space displacements $\mathbf{u}(\mathbf{r})$ and modifies an n-band $\mathbf{k}\cdot\mathbf{p}$ model via the contribution $H_s^{n\times n}$ in (5.6), in which the electronic properties of the system are modified due to strain via deformation potentials. Additionally, strain and in some cases spontaneous polarisation result in an additional polarisation potential φ in (5.6) that can have a decisive influence on the electronic properties of semiconductor nanostructures.

5.4 The Polarisation Potential in a Plane-Wave Framework

The solution of the Poisson equation (5.10) to determine the polarisation potential is achieved in a plane-wave model without the need for a Poisson solver, as typically required within a real-space representation. We can re-write (5.10) as:

$$\nabla\varphi(\mathbf{r}) = -\frac{\mathbf{P}_{\text{tot}}(\mathbf{r})}{\kappa_0\kappa_r(\mathbf{r})} = -\mathbf{P}^{\kappa}(\mathbf{r}). \tag{5.39}$$

We consider as a concrete example the calculation of the polarisation potential in a wurtzite semiconductor nanostructure. There are two contributions to the total polarisation vector, $\mathbf{P}_{\text{tot}}$ in such a nanostructure, one associated with the spontaneous polarisation $\mathbf{P}_{\text{spont}}$ and the other associated with the strain-induced polarisation, $\mathbf{P}_{\text{strain}}$:

$$\mathbf{P}_{\text{tot}}(\mathbf{r}) = \mathbf{P}_{\text{spont}}(\mathbf{r}) + \mathbf{P}_{\text{strain}}(\mathbf{r}) \ . \tag{5.40}$$

The strain-induced polarisation, $\mathbf{P}_{\text{strain}}$ depends on the strain tensor ϵ_{kl} as

$$P_i^{\text{strain}}(\mathbf{r}) = e_{ikl}\epsilon_{kl} = e_{ikl}^{\text{M}}\epsilon_{kl} + \delta e_{ikl}\chi_{\text{QD}}\epsilon_{kl}, \tag{5.41}$$

where e_{ikl}^{M} and e_{ikl}^{QD} are the piezoelectric constants for the matrix and dot materials respectively, and $\delta e_{ikl} \equiv e_{ikl}^{\text{QD}} - e_{ikl}^{\text{M}}$. The spontaneous polarisation contribution, $\mathbf{P}_{\text{spont}}$, is directed along the wurtzite c-axis, and can be defined in terms of the characteristic function as

$$\mathbf{P}_{\text{spont}}(\mathbf{r}) = \mathbf{P}_{\text{spont}}^{\text{M}} + \left[\mathbf{P}_{\text{spont}}^{\text{QD}} - \mathbf{P}_{\text{spont}}^{\text{M}}\right]\chi_{\text{QD}}(\mathbf{r}) \ . \tag{5.42}$$

If we assume for now that the dielectric constant $\kappa_r(\mathbf{r})$ has the same value κ_r in the dot and in the surrounding matrix material, then it can be shown that the Fourier transform of the i-th component of the total induced electric field, $\mathbf{E}$, is given by

$$\tilde{E}_i = -\frac{\xi_i\xi_l}{\kappa_0\kappa_r\xi^2}\left[\tilde{P}_l^{\text{spont}} + \tilde{P}_l^{\text{strain}}\right], \tag{5.43}$$

where tilde denotes the Fourier transform. When using the Fourier transform technique to find the built-in electric field, we must include the additional condition that the electric field averaged over the unit cell of the QD superlattice is zero: $\overline{E}_i = 0$ (this follows from the requirement that the electric field energy is minimised). This is achieved by requiring that $\widetilde{\mathbf{E}}$ is zero at $\xi \equiv 0$. The Fourier transform of the electrostatic potential, φ, is related to the built-in electric field by $\tilde{\varphi} = -i\tilde{E}_i/\xi_i$. The Fourier transform of the polarisation potential is then made up of several contributions:

$$\tilde{\varphi} = \tilde{\varphi}^{\text{spont}} + \tilde{\varphi}^{\text{strain}} + \tilde{\varphi}^{\delta e} \tag{5.44}$$

$$\tilde{\varphi}^{\text{spont}} = -i \frac{\xi_3}{\kappa_0 \kappa_r \xi^2} \left(P^{\text{QD}}_{\text{spont}} - P^{\text{M}}_{\text{spont}} \right) \tilde{\chi}_{\text{QD}}(\boldsymbol{\xi}) \tag{5.45}$$

$$\tilde{\varphi}^{\text{strain}} = -i \frac{1}{\kappa_0 \kappa_r \xi^2} \left[2e_{15}(\xi_1 \tilde{\epsilon}_{13} + \xi_2 \tilde{\epsilon}_{23}) + \xi_3 e_{31}(\tilde{\epsilon}_{11} + \tilde{\epsilon}_{22}) + \xi_3 e_{33} \tilde{\epsilon}_{33} \right] \tag{5.46}$$

$$\tilde{\varphi}^{\delta e} = -i \frac{1}{\kappa_0 \kappa_r \xi^2} \left[2(\xi_1 \tilde{\epsilon}^{\chi}_{13} + \xi_2 \tilde{\epsilon}^{\chi}_{23})(\delta e)_{15} + \xi_3 (\delta e)_{31} (\tilde{\epsilon}^{\chi}_{11} + \tilde{\epsilon}^{\chi}_{22}) + \xi_3 (\delta e)_{33} \tilde{\epsilon}^{\chi}_{33} \right], \tag{5.47}$$

where $\tilde{\epsilon}^{\chi}_{ij}$ in (5.47) denotes the Fourier transform of the product $\chi_{\text{QD}} \epsilon_{ij}$ of the QD characteristic function and the elastic strain in the structure. The Fourier transform of the product is the convolution of the Fourier transforms of the individual terms, with $\tilde{\epsilon}^{\chi}_{ij}$ therefore given by

$$\tilde{\epsilon}^{\chi}_{ij}(\boldsymbol{\xi}) = \sum_{\widetilde{\xi}} \chi_{\text{QD}}(\boldsymbol{\xi} - \widetilde{\boldsymbol{\xi}}) \tilde{\epsilon}_{ij}(\widetilde{\boldsymbol{\xi}}) \ . \tag{5.48}$$

In deriving Eqs. (5.46) and (5.47) we use that for hexagonal III-V crystals only the following components of the piezoelectric tensor are non-zero: $e_{113} = e_{223} = e_{15}$; $e_{311} = e_{322} = e_{31}$; and $e_{333} = e_{33}$.

Equations (5.45)–(5.47) give analytical formulae for the Fourier transform of the built-in electrostatic potential. These can be used directly to calculate QD carrier spectra and wavefunctions, on the assumption that κ_r is constant.

The first contribution to the electrostatic potential in (5.44) is due to the difference in the spontaneous polarisations of the QD and matrix materials. In deriving (5.45), we explicitly used that the spontaneous polarisation is directed along (0001), i.e. $|\mathbf{P}_{\text{spont}}| \equiv P^{\text{spont}}_3$ [6, 19, 25]. The constant part $\mathbf{P}^{\text{M}}_{\text{spont}}$ of $\mathbf{P}_{\text{spont}}$ in (5.42) does not contribute to the potential because of the condition that $\varphi(\boldsymbol{\xi} = 0) = 0$. The second and third contributions to φ are due to the piezoelectric field associated with the strain distribution in the QD structure. The second term, (5.46) describes the piezoelectric field when the piezoelectric constants of the QD and matrix are equal, while the third term (5.47) arises from any difference in these constants between the two materials.

If we now extend to the case where the value of κ_r is allowed to vary between the dot and matrix material, we can then calculate directly the Fourier transform of (5.39) numerically as:

$$-i\mathbf{G}\varphi(\mathbf{G}) = -\mathbf{P}^{\kappa}(\mathbf{G}) \quad \Rightarrow \quad \varphi(\mathbf{G}) = -i \frac{\mathbf{G} \cdot \mathbf{P}^{\kappa}(\mathbf{G})}{\mathbf{G}^2}. \tag{5.49}$$

The plane-wave representation then still allows one to solve the Poisson equation via a single multiplication of the Fourier-transformed polarisation vector $\mathbf{P}^{\kappa}(\mathbf{r})$ with the factor $-i\frac{\mathbf{G}}{\mathbf{G}^2}$. The real-space polarisation potential can then be simply obtained by another Fourier transformation: $\varphi(\mathbf{r}) = \hat{\mathscr{R}}\varphi(\mathbf{G})$.

The approach which we have presented to calculate the strain, the polarisation potential and finally the electronic properties in separate steps allows one in principle to combine real-space and plane-wave based methods. The numerical plane-wave models presented above contain all the relevant real-space properties of the system, and real-space properties such as strain and polarisation potentials can thus be taken likewise from the output of other simulation methods. The results of a multi-band **k·p** calculation in the above plane-wave based formulation can furthermore be directly transformed to real space and they can also be employed in further calculations, using e.g. either a real or reciprocal space formalism to determine dipole strengths and optical spectra.

5.5 Advantages and Disadvantages of a Plane-Wave Representation

A plane-wave based approach to multi-band **k·p** models has a number of advantages, of which many have been outlined already above. To summarise, these were:

1. Gradient operations can be performed in reciprocal space via a simple multiplication of a vector with $-i\,\mathbf{G}$, which is computationally much cheaper than a finite-difference based formulation in real space. This is particularly important for the calculation of polarisation potentials, where a Poisson solver can be replaced by a simple vector multiplication.
2. Accuracy and computational cost can be controlled via the total number of plane waves employed. Furthermore, so-called spurious solutions [47, 52] that occur due to the inaccuracy of common **k·p** models in the region of higher wave vectors, can be prevented by careful choice of the plane-wave cut-off used (see, for example, discussion in [3]).
3. Real-space properties of the system such as shape and material composition can be fully contained as efficient FFT routines allow a quick transfer from real to reciprocal space and vice versa.
4. The characteristic function of most of the common dot shapes considered has an analytical Fourier transform, thereby allowing one to efficiently generate the Hamiltonian matrix for a wide range of problems.
5. A numerical, plane-wave based multi-band **k·p** model can be readily incorporated into an existing plane-wave simulation package and can thus benefit from the existing, highly optimised minimisation schemes available, that require only minor modifications for this purpose. Optimised minimisation schemes can likewise be used to find selected energy levels and wavefunctions of analytically generated Hamiltonian matrices.

In this chapter we have presented two plane-wave representation methods: semi-analytical and numerical. In the first approach all required energy levels and wavefunctions are calculated as the eigenvalues and eigenvectors of a matrix, while for the second approach a separate iteration technique is used in order to obtain the solution for each state. The two methods are complementary to each other: the semi-analytical approach can obtain many levels at once, but may have some drawbacks when self-consistent solutions are required. On the other hand the numerical approach can be much more powerful for many problems, such as when second-order piezoelectric effects need to be included or when self-consistent solutions of **k·p**-based Kohn-Sham equations are required.

On the other hand, there are also a number of potential drawbacks associated with a plane-wave based implementation of a multi-band **k·p** model, as we now describe below.

Periodic Boundary Conditions

The description of real-space properties employing a set of plane waves implicitly assumes that the system is periodic. For the example of a QD, this means that the actual system simulated is not a single, isolated dot but an array of QDs where periodic images of the system are included in all three dimensions. A sufficiently large supercell around the dot can in such cases decouple the dot from its neighbours and prevent erroneous energy levels. However, periodic boundary conditions can also modify the symmetry compared to that of an isolated dot. Even in supercells that are significantly larger than the QD they contain, symmetries can be artificially broken if the symmetry of the supercell does not match the one of the QD. For the example of a wurtzite QD, one would commonly expect to find two degenerate p-like electron states when SO coupling is neglected. However, when calculating the electronic properties of such a C_{3v}-symmetric QD in a cubic cell with periodic boundary conditions, one finds these p-like states split, as the symmetry of the supercell spoils the dot's symmetry [2, 4].

This problem can become even more significant when long-range effects, such as polarisation potentials, are present and where the long-range effects from neighbouring cells then reflect the symmetry of the supercell rather than that of the nanostructure. This type of problem can be eliminated by requiring the supercell symmetry to be consistent with the isolated QD symmetry. For the example of a wurtzite QD, a hexagonal cell should be employed, to prevent artificial symmetry effects arising from periodic images of the nanostructure.

However, periodic boundary conditions can also be of advantage, in particular for systems containing some degree of periodicity, as may for instance occur for a set of vertically stacked QDs. In such a case, a small supercell period can be chosen along the vertical direction in order to analyse the effect of interactions between neighbouring QDs on the electronic properties.

Cutoff Wave Vectors, Spatial Resolution and Spurious Solutions

The maximum number of plane waves employed in a plane-wave based formulation of multi-band **k·p** models can be fixed either by setting a maximum magnitude for the reciprocal lattice vectors used (equivalent to introducing a cut-off energy) or, more generally, by independently varying the maximum number of plane waves included along different directions in the supercell. With either set up, the plane-wave cut-off can be tuned to reduce the computational effort of a calculation and to prevent spurious solutions that occur for larger wave vectors. However, a reduction of the cut-off value also reduces the spatial resolution in real space, as well as the accuracy of the eigenenergies and can thus lead to erroneous results. Likewise the resolution for a given number of plane waves is reduced as the supercell size is increased. Consequently, the cut-off wave vector values and the supercell size are two parameters that need to be carefully chosen to ensure optimum convergence of energy levels and wavefunctions.

Non-local Behaviour of Plane Waves

Within a real-space representation using finite elements, it is possible to apply different spatial resolutions in different regions of the supercell. This means, that interface regions might be sampled on a finer, more accurate grid, whereas bulk material areas in the outer region of the supercell are sampled on a coarser grid. In a plane-wave based approach, the whole supercell is described with the same spatial resolution. Therefore the minimum (cut-off) wavelength required to provide an accurate description of the nanostructure is also employed to describe regions that show mainly bulk properties and that do not contribute to the electronic wave-functions and eigenenergies. This "waste" of plane waves in rather uninteresting regions of the supercell cannot be prevented within a plane-wave model. However, the high computational efficiency of a plane-wave based formulation generally more than compensates for this shortcoming.

Depending on the nanostructure's geometry, size and material composition, the above drawbacks can have a significant influence on the elastic, piezoelectric and electronic properties of the system and it is therefore important to keep these points in mind when a plane-wave based implementation of multi-band **k·p** models is used for the description of such a structure. In general, strain and piezoelectric effects have a slow, power law decay outside an isolated QD, whereas the confined state wavefunctions decay exponentially outside the dot. Improved convergence of electronic structure calculations can therefore be achieved if, for instance, a large supercell is used for the strain and piezoelectric calculations and the results from the centre of the large supercell are then embedded in a smaller supercell for the electronic structure calculations. Such an approach has been presented by Vukmirović and Tomić [49].

5.6 Plane-Wave Approach for (111)-Oriented Zincblende Dots

The plane-wave approach has been widely used to investigate the electronic structure of semiconductor QDs. This has included investigation both of wurtzite-based structures and of conventional (001)-oriented zincblende QD structures [2, 49]. The method is inherently flexible, and can indeed be applied across a much wider range of applications. We illustrate the application of the plane-wave method here by focusing on a specific application, namely site-controlled InGaAs QDs grown along the (111)-direction in GaAs. Such dots have been proposed as possible sources of entangled photons [34, 39, 53]. In the ideal case, these systems exhibit a threefold, C_{3v}, symmetry with respect to their geometry and underlying crystal structure. This C_{3v} symmetry is then high enough to allow for the generation of entangled photons [39, 42]. In fact, extremely small fine structure splittings have been observed using site-controlled InGaAs QDs with a (111)-orientation [14, 30]. There are however a number of interesting theoretical challenges when modelling the electronic properties of (111)-oriented site-controlled InGaAs QDs. This is due to the fact that such dots can exhibit a large base length of 50–80 nm, together with an extremely small aspect ratio with heights of only 1–2 nm along the (111) direction [20]. Due to the large dimensions of these systems, a continuum approach such as the 8-band **k**·**p** formalism appears well suited to investigate their electronic properties. On the other hand, the orientation of these QDs along the (111) direction together with the small aspect ratio require a huge supercell with a very fine discretisation mesh, when employing the conventional **k**·**p** Hamiltonian designed for the description of (001)-oriented systems. This numerical problem can be overcome by analytically rotating an 8-band **k**·**p** model such that one of the coordinate axes is parallel to the (111)-growth direction.

5.6.1 Rotated 8-Band **k**·**p** *Formalism*

The 8-band **k**·**p** Hamiltonian for zincblende QDs in a (111)-oriented simulation cell can be derived by applying the rotation matrix U to the wave vectors **k**:

$$U = \begin{pmatrix} \cos\theta\cos\phi & \cos\theta\sin\phi & -\sin\theta \\ -\sin\phi & \cos\phi & 0 \\ \sin\theta\cos\phi & \sin\theta\sin\phi & \cos\theta \end{pmatrix}, \tag{5.50}$$

where the Euler angles θ $(= \cos^{-1}(1/\sqrt{3}))$ and ϕ $(= \pi/4)$ are the azimuthal angles of the growth direction z' with respect to a conventional cell with the coordinates (x, y, z). We can compute the rotated wave vectors $\mathbf{k}'$ and strain tensor components ϵ'_{ij} as:

$$k'_i = \sum_\alpha U_{i\alpha} k_\alpha \qquad \text{and} \qquad \epsilon'_{ij} = \sum_{\alpha,\beta} U_{i\alpha} U_{j\beta} \epsilon_{\alpha\beta} \ . \tag{5.51}$$

Using a basis set similar to that used in [38],

$$\left(|S'\uparrow\rangle, |X'\uparrow\rangle, |Y'\uparrow\rangle, |Z'\uparrow\rangle, |S'\downarrow\rangle, |X'\downarrow\rangle, |Y'\downarrow\rangle, |Z'\downarrow\rangle \right)^T \tag{5.52}$$

the 8-band Hamiltonian for (111)-oriented zincblende systems, H' can now be written as [28]:

$$H'_{\rm kp} = \begin{pmatrix} H'(\mathbf{k}') & \Gamma'_{\rm so} \\ -\Gamma'^*_{\rm so} & H'^*(\mathbf{k}') \end{pmatrix}, \tag{5.53}$$

with $H'(\mathbf{k}')$ and $\Gamma'_{\rm so}$ being 4×4 matrices. $H'(\mathbf{k}')$ is then composed of matrices describing the potential energy part $H'_{\rm pe}$, the kinetic energy part $H'_{\rm ke}$, the spin-orbit (SO) interaction contribution $H'_{\rm so}$ and a strain dependent part $H'_{\rm str}$:

$$H'(\mathbf{k}') = H'_{\rm pe} + H'_{\rm ke} + H'_{\rm str} + H'_{\rm so} \ . \tag{5.54}$$

The potential energy part $H'_{\rm pe}$ of $H'_{\rm kp}$ contains terms independent of $\mathbf{k}$ and linear in $\mathbf{k}$. These terms are of the same form as in the equivalent part of the Hamiltonian $H_{\rm pe}$ with the usual (001) basis states and are given by:

$$H'_{\rm pe} = \begin{pmatrix} E_{cb} & iPk'_x & iPk'_y & iPk'_z \\ -iPk'_x & \tilde{E}_{\rm vb} & 0 & 0 \\ -iPk'_y & 0 & \tilde{E}_{\rm vb} & 0 \\ -iPk'_z & 0 & 0 & \tilde{E}_{\rm vb} \end{pmatrix}. \tag{5.55}$$

The conduction band edge is denoted by $E_{\rm cb}$ while $\tilde{E}_{\rm vb}$ denotes the average unstrained valence band edge. $E_{\rm cb}$ and $\tilde{E}_{\rm vb}$ are defined as [46]:

$$E_{\rm cb} = E_{\rm vb} + V_{\rm ext} + E_g \ , \ \tilde{E}_{\rm vb} = E_{\rm vb} + V_{\rm ext} - \frac{\Delta_{so}}{3} \ , \tag{5.56}$$

where Δ_{so} denotes the SO coupling energy, E_g the fundamental band gap, $E_{\rm vb}$ is the averaged valence band edge on an absolute scale and $V_{\rm ext}$ an optional scalar potential describing an electric field, e.g. a piezoelectric built-in field. The Kane parameter P is defined as:

$$P = \sqrt{\frac{\hbar^2 E_p}{2m_0}} \ , \tag{5.57}$$

where E_p denotes the interband optical matrix element parameter.

The kinetic energy part H'_{ke} of the Hamiltonian for (111)-oriented systems contains the rotated valence band part $h'(\mathbf{k}')$ plus the conduction band contribution, which is given by $h'^{cb} = A'\mathbf{k}'^2$, where the parameter A' is defined as:

$$A' = \frac{\hbar^2}{2m_0}\left(\frac{1}{m_e} - \frac{E_p}{E_g}\frac{E_g + \frac{2\Delta_{so}}{3}}{E_g + \Delta_{so}}\right) , \tag{5.58}$$

where m_e denotes the Γ-point conduction band effective mass. The kinetic part H'_{ke} of H'_{kp} reads:

$$H'_{\mathrm{ke}} = \begin{pmatrix} A'\mathbf{k}'^2 & 0 & 0 & 0 \\ 0 & h'_{11}(\mathbf{k}') & h'_{12}(\mathbf{k}') & h'_{13}(\mathbf{k}') \\ 0 & h'_{12}(\mathbf{k}') & h'_{22}(\mathbf{k}') & h'_{23}(\mathbf{k}') \\ 0 & h'_{13}(\mathbf{k}') & h'_{23}(\mathbf{k}') & h'_{33}(\mathbf{k}') \end{pmatrix}, \tag{5.59}$$

with

$$\begin{aligned} h'_{11} &= -\frac{1}{2}(\gamma_1 + \gamma_2 + 3\gamma_3)\,k_x'^2 - \frac{1}{2}(\gamma_1 - \gamma_2 - \gamma_3)\,k_y'^2 \\ &\quad -\frac{1}{2}(\gamma_1 - 2\gamma_3)\,k_z'^2 + \sqrt{2}(\gamma_2 - \gamma_3)\,k'_x k'_z + \frac{P^2}{E_g}k_x'^2, \\ h'_{22} &= -\frac{1}{2}(\gamma_1 - \gamma_2 - \gamma_3)\,k_x'^2 - \frac{1}{2}(\gamma_1 + \gamma_2 + 3\gamma_3)\,k_y'^2 \\ &\quad -\frac{1}{2}(\gamma_1 - 2\gamma_3)\,k_z'^2 - \sqrt{2}(\gamma_2 - \gamma_3)\,k'_x k'_z + \frac{P^2}{E_g}k_y'^2, \\ h'_{33} &= -\frac{1}{2}(\gamma_1 - 2\gamma_3)\,(k_x'^2 + k_y'^2) - \frac{1}{2}(\gamma_1 + 4\gamma_3)\,k_z'^2 + \frac{P^2}{E_g}k_z'^2 , \end{aligned}$$

$$\begin{aligned} h'_{12} &= -\sqrt{2}(\gamma_2 - \gamma_3)\,k'_y k'_z - (\gamma_2 + 2\gamma_3)\,k'_y k'_x + \frac{P^2}{E_g}k'_x k'_y, \\ h'_{13} &= -\frac{1}{\sqrt{2}}(\gamma_3 - \gamma_2)\,(k_x'^2 - k_y'^2) - (2\gamma_2 + \gamma_3)\,k'_x k'_z + \frac{P^2}{E_g}k'_x k'_z, \\ h'_{23} &= -\sqrt{2}(\gamma_2 - \gamma_3)\,k'_y k'_x - (2\gamma_2 + \gamma_3)\,k'_y k'_z + \frac{P^2}{E_g}k'_y k'_z , \end{aligned}$$

where γ_i are the Luttinger parameters for the 6-band valence band $\mathbf{k}\cdot\mathbf{p}$ Hamiltonian, defined here in units of $\hbar^2/m_0$, with m_0 being the mass of an electron. The strain dependent part H'_{str} of the 8-band $\mathbf{k}\cdot\mathbf{p}$ Hamiltonian H'_{kp} is given by:

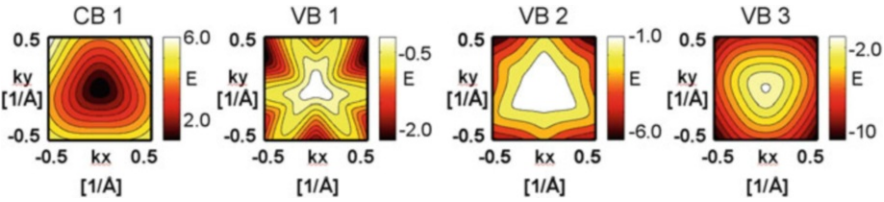


Fig. 5.6 Contour plots of the conduction band (CB 1) and the top three valence bands (VB 1—3) of InAs in the k_x–k_y plane for $k_z = 0$, obtained using the (111)-rotated 8-band **k·p** Hamiltonian of (5.53). Energies E are given in eV

$$H'_{\rm str} = \begin{pmatrix} a_c Tr(\epsilon') & -iP\epsilon'_{1j}k'_j & -iP\epsilon'_{2j}k'_j & -iP\epsilon'_{3j}k'_j \\ iP\epsilon'_{1j}k'_j & h'^{str}_{11} & h'^{str}_{12} & h'^{str}_{13} \\ iP\epsilon'_{2j}k'_j & h'^{str}_{12} & h'^{str}_{22} & h'^{str}_{23} \\ iP\epsilon'_{3j}k'_j & h'^{str}_{13} & h'^{str}_{23} & h'^{str}_{33} \end{pmatrix} . \tag{5.60}$$

The matrix elements $h'^{\rm str}_{ij}$ of the strain dependent part of the Hamiltonian can be obtained from the matrix elements $h'_{ij}(\mathbf{k}')$ by simply using the substitutions:

$$\gamma_1 k'_i k'_j \rightarrow -2a_v \epsilon'_{ij} , \tag{5.61}$$

$$\gamma_2 k'_i k'_j \rightarrow -b\epsilon'_{ij} , \tag{5.62}$$

$$\gamma_3 k'_i k'_j \rightarrow -\frac{d}{\sqrt{3}}\epsilon'_{ij} . \tag{5.63}$$

where the hydrostatic valence band deformation potential is denoted by a_v, while b and d denote the uniaxial deformation potentials. More details on the derivation of an elastic tensor and piezoelectric coefficients suited to a description of (111)-oriented nanostructures can be found in [40].

The SO related contributions $H'_{\rm so}$ and $\Gamma'_{\rm so}$ are given by:

$$H'_{\rm so} = \frac{\Delta_{so}}{3}\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -i & 0 \\ 0 & i & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \Gamma'_{\rm so} = \frac{\Delta_{so}}{3}\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & -i \\ 0 & -1 & i & 0 \end{pmatrix}, \tag{5.64}$$

and are again identical to the equivalent contributions in the (001)-oriented system, due to the isotropy of the SO interaction in zincblende systems.

As a first step to evaluate the derived 8-band Hamiltonian for (111)-oriented nanostructures, contour plots of the conduction band and the three highest valence bands are shown in Fig. 5.6 for $k'_z = 0$ in the k'_x-k'_y-plane. A threefold (C_{3v}) symmetry is clearly visible here [28].

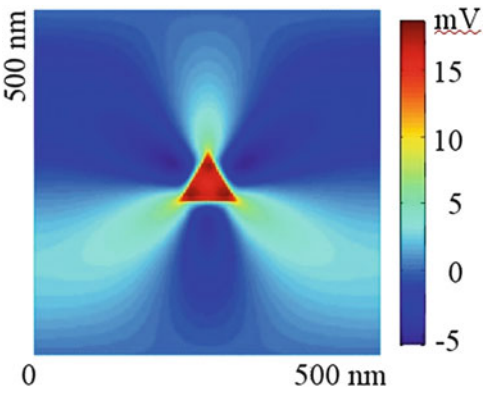


Fig. 5.7 Polarisation potential in a plane on top of a (111)-oriented InGaAs QD with a base length of 80 nm. The supercell size for this calculation was $500 \times 500 \times 30$ nm^3. The C_{3v} symmetry of the potential due to a single dot is clearly visible, but it can also be seen that the boundary conditions imposed by using a square supercell base disturb this symmetry

5.6.2 *Electronic Structure of (111)-Oriented Site-Controlled Zincblende Quantum Dots*

The formalism outlined above is well suited for application to realistic, (111)-oriented InGaAs QD structures. As an example, we consider here the calculation of strain, polarisation potential and single particle electron and hole states for an InGaAs/GaAs quantum dot with a base length of 80 nm and a height of 1.5 nm, of similar dimensions to dots which have recently been investigated experimentally [30]. The QD is taken to have a triangular, truncated pyramid shape, that is oriented top-down along the (111) direction. The piezoelectric polarisation of the QD is shown in a contour plot on top of the dot perpendicular to the growth direction in Fig. 5.7. The C_{3v} symmetry of this potential is clearly visible, though it is slightly disturbed by the cubic supercell that was employed, as was discussed in Sect. 5.5. There is in addition a potential drop across the QD along the growth direction, which tends to induce a separation of electrons and holes in the cell, due to the quantum confined Stark effect.

Figure 5.8 shows the charge densities of the first electron state and the first four hole states inside the quantum dot. Our calculations on this typical dot structure found only one localised electron state, whereas approximately 20 localised hole states were observed. The weak localisation of electrons in comparison to the large number of localised hole states is initially surprising, as the effective masses of electrons and holes parallel to the (111)-plane are of comparable magnitude. Correspondingly, the electron and hole ground state charge densities look very similar from a top view perspective (see Fig. 5.8). However, when the charge density is shown as a line scan along the (111)-axis throughout QD's center (Fig. 5.9), it can be seen that the electron state is much more weakly localised than the hole state. This is due to the fact that the effective mass of the hole along the (111) direction is much larger than the electron effective mass in this direction. As a result, the electron will be more weakly confined along the growth direction in such flat (111)-oriented InGaAs QDs. Furthermore, a spatial separation of electrons and holes can

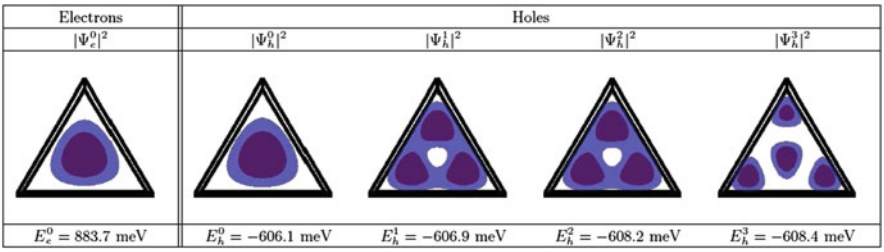


Fig. 5.8 Top-down view of the first and only localised electron and the first four localised hole states of an InGaAs QD with a base length of 80 nm and a height of 1.5 nm. The *blue and purple isosurfaces* represent 20 % and 50 % of the maximum charge density. The calculated energy of each state is given below

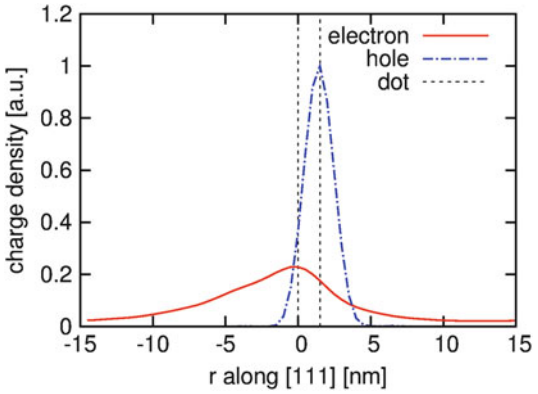


Fig. 5.9 Charge density of the electron and hole ground states, $|\Psi_e^0|^2$ and $|\Psi_h^0|^2$, in a line scan along the (111) direction through the QD's centre. It can be clearly seen that the electron is more weakly localised than the hole. Moreover, a spatial separation of electrons and holes is visible, resulting from the polarisation potential

be seen in Fig. 5.9, as a result of the polarisation potential which has its extrema at the top and bottom of the QD.

5.6.3 *Discussion of Boundary Conditions and Plane-Wave Resolution*

The calculations presented in [40] and above were performed within the plane-wave framework of the S/Phi/nX package. We now use the results of these calculations to revisit our discussion of some of the advantages and disadvantages of a plane-wave based formulation of **k·p** models and of elasticity theory.

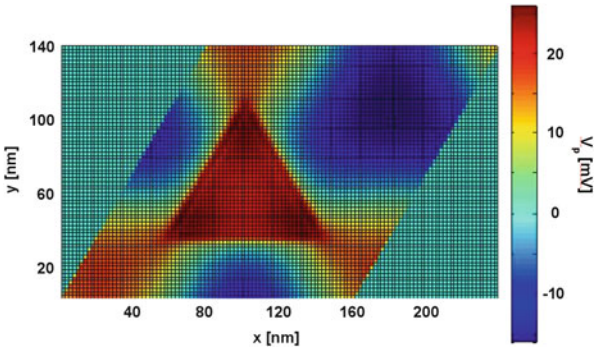


Fig. 5.10 Polarisation potential in a plane on top of a (111)-oriented InGaAs QD with a base length of 80 nm using hexagonal boundary conditions. The unit cell is smaller than the one for Fig. 5.7, because it does not disturb the symmetry, as can be clearly seen. Here, a unit cell of $160 \times 160 \times 30$ nm^3 was employed

Boundary Conditions

A square-based cuboidal supercell was employed in all three steps of the calculations described above, namely in the calculation of the elastic, piezoelectric and electronic properties of the (111)-oriented QD structure. This leads to unwanted changes in the potential due to the periodic images, as can be seen in Fig. 5.7, where the polarisation potential in the vicinity of a triangular InGaAs QD clearly exhibits asymmetries due to the supercell shape. This effect is more pronounced for long-range effects such as strain and in particular piezoelectric potentials, but it also occurs in the calculation of electronic states even when polarisation potentials are not taken into account. In such cases, artificial splittings of p-states occur, that are not expected if an ideal C_{3v}-symmetric geometry is simulated.

However, the implementation of elasticity theory and the **k·p** formalism within a plane-wave package such as S/Phi/nX allows one to remove these symmetry-disturbing effects, because the supercell geometry can be directly chosen to have a sufficiently high symmetry, e.g. a hexagonal symmetry. For this purpose, the cell can for instance be discretised with the basis vectors:

$$L_1 = a\,(1,0,0)\,,\quad L_2 = a\left(\frac{1}{2},\frac{\sqrt{3}}{2},0\right),\quad L_3 = c\,(0,0,1)\,. \tag{5.65}$$

Likewise, such a supercell geometry could also be chosen when implementing the semi-analytical approach to calculate QD properties.

For the above example of a triangular InGaAs QD grown along the (111) direction, such a hexagonal symmetry on top of the C_{3v} symmetry of the system's shape and the elastic and piezoelectric tensors can then yield the correct, C_{3v}-symmetric piezoelectric potential (see Fig. 5.10). This underlines the advantage of an implementation of the above formalisms in an existing plane-wave code where there is flexibility to choose the symmetry of the supercell used.

Discretisation and Number of Plane Waves

The real-space discretisation employed in a plane-wave description depends both on the supercell size and on the number of plane waves used in the calculation. If, for example, we use $2M + 1$ plane waves in a 1-D quantum well simulation with supercell period L (as in Figs. 5.1 and 5.2), then the minimum length scale resolvable, S_{min}, is approximately half of the shortest wavelength considered, i.e. $S_{min} \approx L/2M$. In calculations such as those presented here for (111) oriented supercells, it is clearly desirable to have higher resolution along the growth direction (where the dot height was taken as 1.5 nm) compared to the in-plane directions, where the dot base length of 80 nm is then more than 20 times larger than the height. Direction-dependent (anisotropic) resolution is straightforward to implement in a code where different numbers of plane waves can be specified along different directions. It is more difficult to implement in a code such as the S/Phi/nX package, where the use of a constant cut-off wave vector would lead to the same real-space resolution along all directions.

It is however possible to modify the elastic constants, piezoelectric coefficients and Hamiltonian used as input to such a package in order to achieve different resolution along different directions. We illustrate this principle by considering an infinite square well potential in two dimensions for which the Hamiltonian is given by

$$H = -\frac{\hbar^2}{2m_e^*}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right). \tag{5.66}$$

We suppose that the infinite square well is of width αL along the x direction, and of width L/α along the y direction, therefore requiring different plane-wave resolution along the two directions. We can however introduce a change of coordinates, so that $x' = \alpha x$ and $y' = y/\alpha$, thereby giving a well which has a square base in the new coordinate system. If we also rescale the electron mass, so that $m^*_{x'} = \alpha^2 m_e^*$ and $m^*_{y'} = m_e^*/\alpha^2$ along the x' and y' directions respectively this then gives a modified Hamiltonian with the same eigenvalues and rescaled eigenvectors compared to the original Hamiltonian of (5.66). We have generalised this re-scaling approach to undertake the calculations of the elastic, piezoelectric and electronic properties presented above.

5.7 Conclusion

In summary, we have presented an overview in this chapter that both semi-analytical and numerical plane-wave methods can be used to calculate the electronic structure of semiconductor nanostructures using a multi-band $\mathbf{k}\cdot\mathbf{p}$ Hamiltonian. We showed that plane-wave methods can be used to calculate efficiently both the strain distribution and the polarisation potential associated with a semiconductor QD

embedded in a surrounding matrix material. Knowing the strain and the polarisation potential, it is then possible to calculate efficiently and accurately the electronic structure of the QD conduction and valence states. We reviewed some of the advantages and disadvantages of plane-wave based methods compared to real-space approaches, illustrating some of the different benefits and drawbacks through representative calculations on semiconductor nanostructures. Overall we conclude that plane-wave methods provide an efficient and flexible approach when using $\mathbf{k}\cdot\mathbf{p}$ models to determine the electronic structure of semiconductor nanostructures.

Acknowledgements We thank Christopher A. Broderick for a careful reading and very useful suggestions regarding earlier versions of this chapter. We also thank the many other colleagues with whom we have had the pleasure to work on using plane-wave based methods, including J.A. Barker, S.B. Healy, S. Tomić and D.P. Williams. We acknowledge financial support for the work on (111)-oriented QDs from Science Foundation Ireland (10/IN.1/I299).

References

1. A.D. Andreev, J.R. Downes, D.A. Faux, et al., Strain distributions in quantum dots of arbitrary shape. J. Appl. Phys. **86**, 297–305 (1999)
2. A.D. Andreev, E.P. O'Reilly, Theory of the electronic structure of GaN/AlN hexagonal quantum dots. Phys. Rev. B **62**, 15851–15870 (2000)
3. A.D. Andreev, R.A. Suris, Semiconductors **30** 285–292 (1996)
4. N. Baer, S. Schulz, P. Gartner, et al., Influence of symmetry and Coulomb correlation effects on the optical properties of nitride quantum dots. Phys. Rev. B **76**, 075310, 1–14 (2007)
5. J.A. Barker, E.P. O'Reilly, Theoretical analysis of electron-hole alignment in InAs-GaAs quantum dots. Phys. Rev. B **61**, 13840–13851 (2000)
6. F. Bernardini, V. Fiorentini, D. Vanderbilt, Accurate calculation of polarization-related quantities in semiconductors. Phys. Rev. B **63**, 193201, 1–4 (2001)
7. G. Bester, X. Wu, D. Vanderbilt, et al., Importance of second-order piezoelectric effects in zinc-blende semiconductors. Phys. Rev. Lett. **96**, 187602–187605 (2006)
8. G. Bester, A. Zunger, X. Wu, et al., Effects of linear and nonlinear piezoelectricity on the electronic properties of InAs/GaAs quantum dots. Phys. Rev. B **74**, 081305(R), 1–4 (2006)
9. S. Boeck, C. Freysoldt, A. Dick, et al., The object-oriented DFT program library S/PHI/nX. Comput. Phys. Commun. **182**, 543–554 (2011)
10. P. Boucaud, S. Sauvage, Infrared photodetection with semiconductor self-assembled quantum dots. C. R. Physique **4**, 1133–1154 (2003)
11. J.R. Chelikowsky, N. Troullier, Y. Saad, Finite-difference-pseudopotential method: Electronic structure calculations without a basis. Phys. Rev. Lett. **72**, 1240–1243 (1994)
12. U.M.E. Christmas, A.D. Andreev, D.A. Faux, Calculation of electric field and optical transitions in InGaN / GaN quantum wells. J. Appl. Phys. **98**, 073522, 1–12 (2005)
13. M.A. Cusack, P.R. Briddon, M. Jaros, Electronic structure of InAs/GaAs self-assembled quantum dots. Phys. Rev. B **54**, R2300–R2303 (1996)
14. V. Dimastrodonato, L.O. Mereni, G. Juska, et al., Impact of nitrogen incorporation on pseudomorphic site-controlled quantum dots grown by metalorganic vapor phase epitaxy. Appl. Phys. Lett. **97**, 072115, 1–3 (2010)
15. J.R. Downes, D.A. Faux, E.P. O'Reilly, A simple method for calculating strain dispersions in quantum dot structures. J. Appl. Phys. **81**, 6700–6702 (1997)
16. T. Eissfeller, P. Vogl, Real-space multi-band envelope-function approach without spurious solutions. Phys. Rev. B **84**, 195122, 1–9 (2011)

17. J.D. Eshelby, The elastic field outside an ellipsoidal inclusion. Proc. R. Soc. London, Ser. A **252**, 561–569 (1959)
18. X. Gonze, B. Amadon, P.M. Anglade, et al., ABINIT: First-principles approach to material and nanosystem properties. Comput. Phys. Commun. **180**, 2582–2615 (2009)
19. N. Grandjean, M. Leroux, J. Massies, Appl. Phys. Lett, **74**, 2361 (1999)
20. S.B. Healy, R.J. Young, L.O. Mereni, et al., Physics of novel site controlled InGaAs quantum dots on (111) oriented substrates. Physica E (Amsterdam) **42**, 2761–2764 (2010)
21. P. Hohenberg, W. Kohn, Inhomogeneous Electron Gas. Phys. Rev. **136**, B864–B871 (1964)
22. B. Jogai, Three-dimensional strain field calculations in coupled InAs/GaAs quantum dots. J. Appl. Phys. **88**, 5050–5055 (2000)
23. R.O. Jones, O. Gunnarsson, The density funtional formalism, its applications and prospects. Rev. Mod. Phys. **61**, 689–746 (1989)
24. G. Kresse, J. Hafner, Ab initio molecular dynamics for liquid metals. Phys. Rev. B **47**, 558–561 (1993)
25. M. Leroux, N. Grandjean, M. Laügt, et al., Phys. Rev. B. **58**, R13371–R13374 (1998)
26. I.M. Lifshits, L.N. Rosentsverg, Zhurnal Exper. Teor. Phiziki **17**, 9, (1947) (in russian)
27. O. Marquardt, S. Boeck, C. Freysoldt, et al., Plane-wave implementation of the real-space **k·p** formalism and continuum elasticity theory. Comput. Phys. Commun. **181**, 765–771 (2010)
28. O. Marquardt, E.P. O'Reilly, S. Schulz, Asymmetric electronic properties of site-controlled (111)-oriented zinc-blende quantum dots calculated using a symmetry adapted **k·p** Hamiltonian. (submitted)
29. O. Marquardt, S. Schulz, C. Freysoldt, et al., A flexible, plane-wave based multi-band **k·p** model. Opt. Quant. Elec. **44**, 183–188 (2012)
30. L.O. Mereni, O. Marquardt, G. Juska, et al., Fine-structure splitting in large-pitch pyramidal quantum dots. Phys. Rev. B **85**, 155453, 1–13 (2012)
31. J.J. Mortensen, L.B. Hansen, K.W. Jacobsen, Real-space grid implementation of the projector augmented wave method. Phys. Rev. B **71**, 035109, 1–11 (2005)
32. A.S. Moskalenko, J. Berakdar, J. Prokofiev, et al., Single-particle states in spherical Si/SiO_2 quantum dots. Phys. Rev. B **76**, 085427, 1–9 (2007)
33. M.C. Payne, M.P. Teter, D.C. Allan, et al., Iterative minimization techniques for ab initio total-energy calculations: molecular dynamics and conjugate gradients. Rev. Mod. Phys. **64**, 1045–1097 (1992)
34. E. Pelucchi, S. Watanabe, K. Leifer, et al., Mechanisms of quantum dot energy engineering by metalorganic vapor phase epitaxy on patterned nonplanar substrates. Nano Lett. **7**, 1282–1285 (2007)
35. M. Povolotskyi, M. Auf der Maur, A. Di Carlo, Strain effects in freestanding three-dimensional nitride nanostructures. Phys. Stat. Sol. (c) **2**, 3891–3894 (2005)
36. C. Pryor, M.E. Pistol, L. Samuelson, Electronic structure of strained $InP/Ga_{0.51}In_{0.49}P$ quantum dots. Phys. Rev. B **56** 10404–10411 (1997)
37. C. Pryor, Eight-band calculations of strained InAs/GaAs quantum dots compared with one-, four-, and six-band approximations. Phys. Rev. B **57**, 7190–7195 (1998)
38. A. Schliwa, Electronic properties of self-organized quantum dots, Ph.D. dissertation, TU Berlin, Berlin (2007)
39. A. Schliwa, M. Winkelnkemper, A. Lochmann, et al., In(Ga)As/GaAs quantum dots grown on a (111) surface as ideal sources of entangled photon pair. Phys. Rev. B **80**, 161307(R), 1–4 (2009)
40. S. Schulz, M.A. Caro, E.P. O'Reilly, et al., Symmetry-adapted calculations of strain and polarization fields in (111)-oriented zinc-blende quantum dots. Phys. Rev. B **84**, 125312, 1–14 (2011)
41. M.D. Segall, P.L.D. Lindan, M.J. Probert, et al., First-principles simulation: ideas, illustrations and the CASTEP code. J. Phys: Cond. Matt. **14** 2717–2743 (2002)
42. R. Singh, G. Bester, Nanowire quantum dots as an ideal source of entangled photon pairs. Phys. Rev. Lett. **103**, 063601, 1–4 (2009)

43. O. Stier, D. Bimberg, Modelling of strained quantum wires using eight-band **k·p** theory. Phys. Rev. B **55**, 7726–7732 (1997)
44. O. Stier, *Electronic and optical properties of quantum dots and wires* (Berlin, 2000)
45. S. Tomić, Electronic structure of $In_yGa_{1-y}As_{1-x}N_x$/GaAs(N) quantum dots by ten-band **k·p** theory. Phys. Rev. B **73**, 125348, 1–10 (2006)
46. C.G. Van de Walle, Band lineups and deformation potentials in the model-solid theory. Phys. Rev. B **39**, 1871–1883 (1989)
47. R.G. Veprek, S. Steiger, B. Witzigmann, Ellipticity and the spurious solution problem of **k·p** envelope equations. Phys. Rev. B **76**, 165320, 1–9 (2007)
48. R.G. Veprek, S. Steiger, B. Witzigmann, Reliable **k·p** band structure calculation for nanostructures using finite elements. J. Comput. Electron. **7**, 521–529 (2008)
49. N. Vukmirović, S. Tomić, Plane wave methodology for single quantum dot electronic structure calculations. J. Appl. Phys. **103**, 103718–103729 (2008)
50. M. Willatzen, B. Lassen, L.L.Y. Voon, et al., Dynamic coupling of piezoelectric effects, spontaneous polarization, and strain in lattice-mismatched semiconductor quantum-well heterostructures. J. Appl. Phys. **100**, 024302, 1–6 (2006)
51. M. Winkelnkemper, S. Schliwa, D. Bimberg, Interrelation of structural and electronic properties in $In_xGa_{1-x}N$/GaN quantum dots using an eight-band **k·p** model. Phys. Rev. B **74**, 155322, 1–12 (2006)
52. W. Yang, K. Chang, Origin and elimination of spurious solutions of the eight-band **k·p** theory. Phys. Rev. B **72**, 233309, 1–4 (2005)
53. Q. Zhu, K.F. Karlsson, E. Pelucchi, et al., Transition from two-dimensional to three-dimensional quantum confinement in semiconductor quantum wires/quantum dots. Nano Lett. **7**, 2227–2233 (2007)

Part III
Applications

Chapter 6
The Multi-Band k·p Hamiltonian for Heterostructures: Parameters and Applications

Stefan Birner

Abstract In this contribution all the various definitions of the **k·p** parameters available in the literature are summarized, and equations that relate them to each other are provided. We believe that such a summary for both zinc blende and wurtzite crystals on a few pages is very useful, not only for beginners but also for experienced researchers that quickly want to look up conversion formulas. Results of **k·p** calculations for bulk semiconductors are shown for diamond, and for unstrained and strained InAs. Several examples of **k·p** calculations for heterostructures are presented. They cover spurious solutions, a spherical quantum dot and heterostructures showing the untypical type-II and type-III band alignments. Finally, self-consistent **k·p** calculations of a two-dimensional hole gas in diamond for different substrate orientations are analyzed. Wherever possible, the **k·p** results are compared to tight-binding calculations. All these calculations have been performed using the next**nano** software (next**nano**: The next**nano** software can be obtained from http://www.nextnano.com; Birner et al., IEEE Trans. Electron Dev. 54:2137–2142, 2007). Therefore, this contribution provides some specific details that are relevant for a numerical implementation of the **k·p** method.

S. Birner (✉)
Walter Schottky Institute and Physics Department, Technische Universität München, Am Coulombwall 4, 85748 Garching, Germany

Institute for Nanoelectronics, Technische Universität München, Arcisstr. 21, 80333 München, Germany
e-mail: stefan.birner@nextnano.com

M. Ehrhardt and T. Koprucki (eds.), *Multi-Band Effective Mass Approximations*, Lecture Notes in Computational Science and Engineering 94,
DOI 10.1007/978-3-319-01427-2_6,

6.1 The 8-Band k·p Hamiltonian for Bulk Materials

Compared to the single-band effective-mass model, a more accurate description of the band structure can be obtained by using the multi-band **k·p** method that has been used since the 1950s. There are a number of different **k·p** models discussed in the literature, e.g. [5, 9, 14, 16, 45, 50]. They differ mainly in the number of bands considered and their treatment of strain and spin-orbit interaction. The model that considers up to 14 bands predicts almost perfectly the bulk band structure but the computational effort turns out to be nearly as large as for the empirical tight-binding approach. In our numerical implementation, i.e. the next**nano** software, we make use of the 8-band model (e.g. [5]). It is a compromise between the accuracy and the computational cost, as well as the number of required (and known) material parameters. This model includes the lowest conduction band and the three highest valence bands. All other remote bands are treated as perturbations. Spin-orbit interaction and strain are taken into account as small perturbations.

There are basically two ways in obtaining the bulk **k·p** Hamiltonian matrix, the first one is based on a perturbative approach pioneered by Dresselhaus, Kip and Kittel [19], the second one is based on symmetry analysis (method of invariants) introduced by Luttinger [40]. A few years ago, Foreman derived six- [26] and eight-band Hamiltonians [27] from Burt's exact envelope function theory [12, 13] for heterostructures. He showed that his nonsymmetrized Hamiltonian for a homogeneous infinite sample is consistent to the bulk **k·p** Hamiltonian, and that deriving the heterostructure Hamiltonian from the bulk one using a symmetrization procedure is incorrect (symmetrized Hamiltonian). These works solved the problem of operator ordering. The reason lies in the noncommutativity of the differential operator and the (position dependent) material parameters. We note that the potential energy term of Burt's exact envelope function equation contains an extra nonlocal term ($V_{nm}(\mathbf{x}, \mathbf{x}')$) that has been neglected. In fact, using some approximations, it can be shown that the nonlocal part does not contribute for slowly varying envelope functions. At distances far away from a heterointerface the potential tends to a constant, the local periodic potential, and the nonlocal contribution is small. Close to heterointerfaces, Burt's theory leads to two correction terms to the potential function. Another view is that perturbative effects of material inhomogeneities lead to so-called interface Hamiltonians. A derivation and discussion of the Burt–Foreman theory is given in the recently published book by Lew Yan Voon and Willatzen [36] which additionally offers a detailed description and comprehensive overview on all the different **k·p** Hamiltonians used in the literature for both bulk and nanostructured semiconductors.

The key feature of the **k·p** method is the envelope function ansatz based on Bloch's theorem, according to which the electron wave function in a crystal with translational symmetries can be separated into an oscillating Bloch part which is periodic over atomic distance and a smooth envelope function which varies on a mesoscopic scale. Using Löwdin perturbation theory the rapidly oscillating Bloch functions can be eliminated from the electron Hamiltonian. Thus the resulting

electron Hamiltonian only contains the envelope functions. A detailed summary of the **k·p** method has been presented in [2, 58].

In the following we describe the **k·p** Hamiltonian for a bulk semiconductor but use the notation suited for heterostructures. The one-electron Schrödinger equation reads

$$\left(\mathbf{H}_0 + \mathbf{H}_{\mathrm{so}}\right) \Psi_{n\mathbf{k}}\left(\mathbf{x}\right) = E_n\left(\mathbf{k}\right) \Psi_{n\mathbf{k}}\left(\mathbf{x}\right) \tag{6.1}$$

$$\left(\frac{\mathbf{p}^2}{2m_0} + V\left(\mathbf{x}\right) + \frac{\hbar^2}{4m_0^2c^2}\left(\boldsymbol{\sigma} \times \nabla V\right) \cdot \mathbf{p}\right) \Psi_{n\mathbf{k}}\left(\mathbf{x}\right) = E_n\left(\mathbf{k}\right) \Psi_{n\mathbf{k}}\left(\mathbf{x}\right) \ , \tag{6.2}$$

where $\Psi_{n\mathbf{k}}$ is the Bloch function

$$\Psi_{n\mathbf{k}}\left(\mathbf{x}\right) = e^{\mathrm{i}\mathbf{k}\cdot\mathbf{x}} u_{n\mathbf{k}}\left(\mathbf{x}\right) \ , \tag{6.3}$$

composed of the product of a plane wave $e^{\mathrm{i}\mathbf{k}\cdot\mathbf{x}}$ and the periodic Bloch factors $u_{n\mathbf{k}}\left(\mathbf{x}\right)$, n is the band index and $\mathbf{k}$ is a wave vector in the first Brillouin zone which corresponds to the periodicity of the potential energy $V\left(\mathbf{x}\right)$. The operator $\mathbf{H}_{\mathrm{so}}$ approximately takes into account the relativistic effect of spin

$$\mathbf{H}_{\mathrm{so}} = \frac{\hbar^2}{4m_0^2c^2}\left(\nabla V \times \mathbf{p}\right) \cdot \boldsymbol{\sigma} \ , \tag{6.4}$$

where V is the potential energy term, $\mathbf{p}$ is the momentum operator and $\boldsymbol{\sigma}$ is the vector of the Pauli matrices $\boldsymbol{\sigma} = \left(\sigma_1, \sigma_2, \sigma_3\right)^{\mathrm{T}}$. If we insert the Bloch function $\Psi_{n\mathbf{k}}\left(\mathbf{x}\right)$ into Eq. (6.2), we obtain after canceling the plane wave $e^{\mathrm{i}\mathbf{k}\cdot\mathbf{x}}$

$$\left(\mathbf{H}_0 + \mathbf{H}_{\mathbf{k}} + \mathbf{H}_{\mathbf{k}\cdot\mathbf{p}} + \mathbf{H}_{\mathrm{so}}\right) u_{n\mathbf{k}}\left(\mathbf{x}\right) = E_n\left(\mathbf{k}\right) u_{n\mathbf{k}}\left(\mathbf{x}\right) \tag{6.5}$$

$$\left(\frac{\mathbf{p}^2}{2m_0} + V\left(\mathbf{x}\right) + \frac{\hbar^2k^2}{2m_0} + \frac{\hbar}{m_0}\mathbf{k}\cdot\mathbf{p} + \frac{\hbar^2}{4m_0^2c^2}\left(\nabla V \times \mathbf{p}\right) \cdot \boldsymbol{\sigma}\right) u_{n\mathbf{k}}\left(\mathbf{x}\right) = E_n\left(\mathbf{k}\right) u_{n\mathbf{k}}\left(\mathbf{x}\right) \ , \tag{6.6}$$

which is now written for the periodic Bloch spinor $u_{n\mathbf{k}}\left(\mathbf{x}\right)$ only. Within our approximation we consider only the part of the spin-orbit interaction Hamiltonian that is independent of $\mathbf{k}$ because the contribution of the $\mathbf{k}$ dependent part is much smaller. Solving this equation for $\mathbf{k} = 0$ (Γ point) yields the Bloch factors u_{j0} which form a complete and orthonormal basis. The Bloch factor $u_{n\mathbf{k}}$ is expanded for any value of $\mathbf{k}$ using the known Bloch factors u_{j0} at the Γ point

$$u_{n\mathbf{k}}\left(\mathbf{x}\right) = \sum_{j=1}^{8} a_j\left(\mathbf{k}\right) u_{j0}\left(\mathbf{x}\right) \ . \tag{6.7}$$

For our **k**·**p** model, the index j goes from 1 to 8 for 8×8 **k**·**p** (one conduction and three valence bands, including spin), and from 1 to 6 for 6×6 **k**·**p** (three valence bands, including spin). However, in our algorithmic implementation for 6×6**k**·**p** it goes from 3 to 8 because in this case we use the same Hamiltonian matrix (Eq. (6.10)) and omit the indices 1 and 2 related to the conduction band.

The band structure near the Γ point is described by perturbation theory around $\mathbf{k} = 0$ using a number of perturbationally defined parameters. The Γ point electron wave function is expanded into s and p orbital functions. A perturbation model that includes the spin-orbit interaction which is responsible for the splitting Δ_{so} between the Γ_7 and Γ_8 valence bands, requires a basis of eight so-called Bloch functions

$$\{|S\uparrow\rangle, |S\downarrow\rangle, |X\uparrow\rangle, |Y\uparrow\rangle, |Z\uparrow\rangle, |X\downarrow\rangle, |Y\downarrow\rangle, |Z\downarrow\rangle\}, \tag{6.8}$$

where X, Y, Z are the p-type Bloch functions referring to the three principal directions in the crystal and the arrows denote the spin. The designations S, X, Y, Z refer to the corresponding symmetry properties under operations of the tetrahedral group. For heterostructures, the envelope functions ψ that correspond to the Bloch functions in Eq. (6.8) are given by

$$\{\psi_{S\uparrow}, \psi_{S\downarrow}, \psi_{X\uparrow}, \psi_{Y\uparrow}, \psi_{Z\uparrow}, \psi_{X\downarrow}, \psi_{Y\downarrow}, \psi_{Z\downarrow}\}. \tag{6.9}$$

The structure of the bulk 8×8 **k**·**p** Hamiltonian operator $\mathbf{H}_0$ without strain and without spin-orbit coupling in the basis of Eq. (6.8) is given by

	$\lvert S\uparrow\rangle$	$\lvert S\downarrow\rangle$	$\lvert X\uparrow\rangle$ $\lvert Y\uparrow\rangle$ $\lvert Z\uparrow\rangle$	$\lvert X\downarrow\rangle$ $\lvert Y\downarrow\rangle$ $\lvert Z\downarrow\rangle$
$\lvert S\uparrow\rangle$	H_{cc}	0	$\mathbf{H}_{\mathrm{cv}}$	$\mathbf{0}$
$\lvert S\downarrow\rangle$	0	H_{cc}	$\mathbf{0}$	$\mathbf{H}_{\mathrm{cv}}$
$\lvert X\uparrow\rangle$ $\lvert Y\uparrow\rangle$ $\lvert Z\uparrow\rangle$	$\mathbf{H}_{\mathrm{vc}}$	$\mathbf{0}$	$\mathbf{H}_{\mathrm{vv}}$	$\mathbf{0}$
$\lvert X\downarrow\rangle$ $\lvert Y\downarrow\rangle$ $\lvert Z\downarrow\rangle$	$\mathbf{0}$	$\mathbf{H}_{\mathrm{vc}}$	$\mathbf{0}$	$\mathbf{H}_{\mathrm{vv}}$

(6.10)

It describes the electrons in the Γ_6 conduction band, or the Γ_7 or Γ_8 valence bands. Our choice of ordering is due to the fact that we are using the same routines within our algorithm for the setup of the 8×8 and the 6×6 Hamiltonian. In the latter case, the first two rows and the first two columns are ignored. For zinc blende, $\mathbf{H}_{\mathrm{vv}}$ is given by

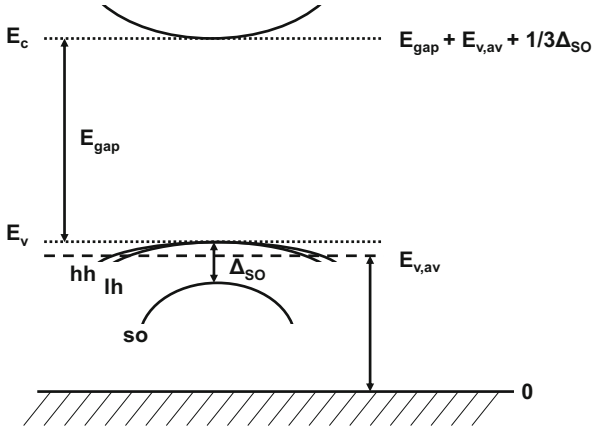


Fig. 6.1 Conduction (E_c) and valence band (E_v) alignment in a zinc blende semiconductor. The heavy hole (hh), light hole (lh) and split-off hole (so) band edges are obtained by adding the spin-orbit Hamiltonian to the average valence band edge energy $E_{v,av}$

$$\mathbf{H}_{vv} = \begin{pmatrix} E_{v,av} + \frac{\hbar^2}{2m_0}\mathbf{k}^2 & & \\ & E_{v,av} + \frac{\hbar^2}{2m_0}\mathbf{k}^2 & \\ & & E_{v,av} + \frac{\hbar^2}{2m_0}\mathbf{k}^2 \end{pmatrix} + \tag{6.11}$$

$$\begin{pmatrix} k_x L k_x + k_y M k_y + k_z M k_z & k_x N^+ k_y + k_y N^- k_x & k_x N^+ k_z + k_z N^- k_x \\ k_y N^+ k_x + k_x N^- k_y & k_x M k_x + k_y L k_y + k_z M k_z & k_y N^+ k_z + k_z N^- k_y \\ k_z N^+ k_x + k_x N^- k_z & k_z N^+ k_y + k_y N^- k_z & k_x M k_x + k_y M k_y + k_z L k_z \end{pmatrix},$$

where $E_{v,av}$ is the energy of the average of the three valence band edges, without strain shifts and without taking spin-orbit splitting into account (see Fig. 6.1). The contribution of the free electron term

$$\frac{\hbar^2}{2m_0}\mathbf{k}^2 = \frac{\hbar^2}{2m_0}\left(k_x^2 + k_y^2 + k_z^2\right) \tag{6.12}$$

could in principle be incorporated into the L and M parameters. This will be discussed further below. For wurtzite the second term has to be replaced by

$$\begin{pmatrix} k_x L_1 k_x + k_y M_1 k_y + k_z M_2 k_z & k_x N_1^+ k_y + k_y N_1^- k_x & k_x N_2^+ k_z + k_z N_2^- k_x \\ k_y N_1^+ k_x + k_x N_1^- k_y & k_x M_1 k_x + k_y L_1 k_y + k_z M_2 k_z & k_y N_2^+ k_z + k_z N_2^- k_y \\ k_z N_2^+ k_x + k_x N_2^- k_z & k_z N_2^+ k_y + k_y N_2^- k_z & k_x M_3 k_x + k_y M_3 k_y + k_z L_2 k_z \end{pmatrix}. \tag{6.13}$$

The DKK (Dresselhaus–Kip–Kittel) parameters [19] L, M, N^+, N^- can either be calculated from the appropriate matrix elements (see [30, 36]) or from the Luttinger parameters γ_1, γ_2, γ_3, κ [2, 5]

$$L = F + 2G = (-\gamma_1 - 4\gamma_2 - 1)\,\frac{\hbar^2}{2m_0} \tag{6.14}$$

$$M = H_1 + H_2 = (2\gamma_2 - \gamma_1 - 1)\,\frac{\hbar^2}{2m_0} \tag{6.15}$$

$$N^+ = F - G = (-3\gamma_3 - (3\kappa + 1))\frac{\hbar^2}{2m_0} = \frac{N}{2} - (3\kappa + 1)\frac{\hbar^2}{2m_0} \tag{6.16}$$

$$N^- = H_1 - H_2 = (-3\gamma_3 + (3\kappa + 1))\frac{\hbar^2}{2m_0} = \frac{N}{2} + (3\kappa + 1)\frac{\hbar^2}{2m_0}\,. \tag{6.17}$$

As k_i and k_j commute in bulk, it holds

$$N = N^+ + N^- = F - G + H_1 - H_2 = -6\gamma_3\frac{\hbar^2}{2m_0}\,. \tag{6.18}$$

The inverse relations for F, G, H_1 and H_2 are

$$F = \frac{1}{3}(L + 2N^+) = \left(-\frac{1}{3}\gamma_1 - \frac{4}{3}\gamma_2 - 2\gamma_3 - 2\kappa - 1\right)\frac{\hbar^2}{2m_0} = -6\sigma\frac{\hbar^2}{2m_0} \tag{6.19}$$

$$G = \frac{1}{3}(L - N^+) = \left(-\frac{1}{3}\gamma_1 - \frac{4}{3}\gamma_2 + \gamma_3 + \kappa\right)\frac{\hbar^2}{2m_0} = -6\delta\frac{\hbar^2}{2m_0} \tag{6.20}$$

$$H_1 = \frac{1}{2}(M + N^-) = \left(-\frac{1}{2}\gamma_1 + \gamma_2 - \frac{3}{2}\gamma_3 + \frac{3}{2}\kappa\right)\frac{\hbar^2}{2m_0} = -6\pi\frac{\hbar^2}{2m_0} \tag{6.21}$$

$$H_2 = \frac{1}{2}(M - N^-) = \left(-\frac{1}{2}\gamma_1 + \gamma_2 + \frac{3}{2}\gamma_3 - \frac{3}{2}\kappa - 1\right)\frac{\hbar^2}{2m_0}\,, \tag{6.22}$$

where the Foreman parameters σ, π and δ will be introduced further below. Rather than specifying the four parameters L, M, N^+, N^-, occasionally another set of parameters L, M, N, K is specified, where $K = -\frac{\hbar^2}{2m_0}2(3\kappa + 1)$ [40]. The parameters F, G, H_1 and H_2 are defined in [34]. There, also an additional fifth Luttinger parameter q [40] related to spin-orbit splitting is given which is typically neglected, and also neglected in our work. The inverse relations for the Luttinger parameters are

$$\begin{aligned}
\gamma_1 &= -\frac{1}{3}\,(L + 2M)\,\frac{2m_0}{\hbar^2} - 1 \\
\gamma_2 &= -\frac{1}{6}\,(L - M)\,\frac{2m_0}{\hbar^2} \\
\gamma_3 &= -\frac{1}{6}(N^+ + N^-)\frac{2m_0}{\hbar^2} = -\frac{1}{6}N\frac{2m_0}{\hbar^2} \\
\kappa &= -\frac{1}{6}(N^+ - N^-)\frac{2m_0}{\hbar^2} - \frac{1}{3}\,,
\end{aligned} \tag{6.23}$$

and

$$\begin{aligned}
\gamma_1 &= -\frac{1}{3}\left(F + 2G + 2H_1 + 2H_2\right)\frac{2m_0}{\hbar^2} - 1 \\
\gamma_2 &= -\frac{1}{6}\left(F + 2G - H_1 - H_2\right)\frac{2m_0}{\hbar^2} \\
\gamma_3 &= -\frac{1}{6}\left(F - G + H_1 - H_2\right)\frac{2m_0}{\hbar^2} \\
\kappa &= -\frac{1}{6}\left(F - G - H_1 + H_2\right)\frac{2m_0}{\hbar^2} - \frac{1}{3}\,.
\end{aligned} \tag{6.24}$$

The parameter H_2 is small and thus it is often neglected, e.g. in [27]. This leads to $N^- \approx M$ and $N^+ = N - N^- \approx N - M$. This is exactly equivalent to the case where κ is not known, and therefore approximated by

$$\kappa \approx -\frac{1}{6}\left(N - 2M\right)\frac{2m_0}{\hbar^2} - \frac{1}{3} = -\frac{1}{3}\left(\gamma_1 - 2\gamma_2 - 3\gamma_3 + 2\right)\,, \tag{6.25}$$

where N is defined in Eq. (6.18). Using this definition for κ, i.e. assuming $H_2 = 0$, Eqs. (6.16) and (6.17) can be expressed using the Luttinger parameters

$$N^- \approx H_1 = M = (2\gamma_2 - \gamma_1 - 1)\frac{\hbar^2}{2m_0} \tag{6.26}$$

$$N^+ = F - G = N - N^- \approx N - M = (-6\gamma_3 - (2\gamma_2 - \gamma_1 - 1))\frac{\hbar^2}{2m_0}\,. \tag{6.27}$$

For the **k·p** dispersion of bulk semiconductors without magnetic field, the contributions of the term $3\kappa + 1$ to N^+ (Eq. (6.16)) and N^- (Eq. (6.17)) effectively cancel each other. Therefore, κ is not needed and can be ignored in this case, suggesting to use only the parameter N for the bulk Hamiltonian. This misleads to effectively using $N^+ = N^- = N/2$ (Eq. (6.18)), a practice that was adopted by the whole **k·p** community until the last decade. However, Foreman identified this symmetrized **k·p** Hamiltonian to be incorrect for heterostructures [26], pointing out the noncommutativity of the momentum and position operators in heterostructures. It is thus crucial to use the correct form of the nonsymmetrized **k·p** Hamiltonian which includes the correct definitions of N^+ and N^-. In both cases the entire matrix is Hermitian whereas in the symmetrized approach, additionally, each matrix element is Hermitian. The symmetrized Hamiltonian has been derived from the bulk **k·p** Hamiltonian, whereas the nonsymmetrized version is based on Burt's exact envelope function theory for heterostructures [12], which has been extended by Foreman to multi-band **k·p**. For that reason it is usually called the Burt–Foreman Hamiltonian [36]. If the **k·p** material parameters do not depend on position, e.g. in the case of a quantum well with infinite barriers, i.e. no material interfaces, both symmetrizations lead to the same results in the case of zero magnetic field. If one is

only interested in the bulk **k·p** dispersion, there is no need to explicitly use N^+ and N^-, and thus N can be used instead (Eq. (6.18)). Unfortunately, in a lot of articles in the last decades N^+ and N^- (and also κ) have been ignored and only N has been used, which we now know is definitely not correct for heterostructures. The noncommutativity of the off-diagonal matrix elements has already been pointed out in the early work of Luttinger [40], where he derived the most general form of the 6×6 **k·p** Hamiltonian in the presence of an external homogeneous magnetic field.

Foreman introduced another set of dimensionless parameters σ, π and δ [26]

$$\begin{aligned}
\sigma &= -\frac{1}{6}F\frac{2m_0}{\hbar^2} = -\frac{1}{18}(L+2N^+)\frac{2m_0}{\hbar^2} \approx -\frac{1}{2}\delta + \overline{\gamma} \\
\pi &= -\frac{1}{6}H_1\frac{2m_0}{\hbar^2} = -\frac{1}{12}(M+N^-)\frac{2m_0}{\hbar^2} \approx \frac{3}{2}\delta + \mu \\
\delta &= -\frac{1}{6}G\frac{2m_0}{\hbar^2} = -\frac{1}{18}(L-N^+)\frac{2m_0}{\hbar^2} \approx \frac{1}{9}(\gamma_1+\gamma_2-3\gamma_3+1)\,,
\end{aligned} \tag{6.28}$$

where $\overline{\gamma}$ and μ are defined as

$$\overline{\gamma} = \frac{1}{2}(\gamma_3+\gamma_2) \tag{6.29}$$

$$\mu = \frac{1}{2}(\gamma_3-\gamma_2)\,. \tag{6.30}$$

The notation in Greek letters σ, π and δ is derived from the s, p, d (and f) orbitals of the constituent atoms. Here, the contribution of the f orbitals is neglected which is equivalent to setting $H_2 = 0$, i.e. approximating κ (Eq. (6.25)). The inverse relations show how the Luttinger parameters can be expanded to reflect the symmetry of the interaction of the bands [36]

$$\begin{aligned}
\gamma_1 &\approx -\frac{1}{3}\left(F+2G+2H_1\right)\frac{2m_0}{\hbar^2} - 1 = 2\sigma + 4\pi + 4\delta - 1 \\
\gamma_2 &\approx -\frac{1}{6}\left(F+2G-H_1\right)\frac{2m_0}{\hbar^2} = \sigma - \pi + 2\delta \\
\gamma_3 &\approx -\frac{1}{6}\left(F-G+H_1\right)\frac{2m_0}{\hbar^2} = \sigma + \pi - \delta \\
\kappa &\approx -\frac{1}{6}\left(F-G-H_1\right)\frac{2m_0}{\hbar^2} - \frac{1}{3} = \sigma - \pi - \delta - \frac{1}{3}\,.
\end{aligned} \tag{6.31}$$

They are similar to Eq. (6.24), with the exception that the term H_2 has been neglected. Consequently, the fourth parameter κ is not an independent parameter here. It depends on the choice of σ, π and δ, or γ_1, γ_2 and γ_3, respectively. The corresponding relation for κ in terms of the Luttinger parameters is given in Eq. (6.25). Finally, we list the related equations for the DKK parameters

$$L = F + 2G = (-6\sigma - 12\delta)\,\frac{\hbar^2}{2m_0} \tag{6.32}$$

$$N^+ = F - G = (-6\sigma + 6\delta)\,\frac{\hbar^2}{2m_0}$$

$$M \approx H_1 = -6\pi\,\frac{\hbar^2}{2m_0}$$

$$N^- \approx H_1 = -6\pi\,\frac{\hbar^2}{2m_0}\,.$$

The Luttinger parameters are based on the method of invariants whereas the DKK parameters are based on the perturbation method for deriving the **k·p** Hamiltonian matrix. This is the reason for the various definitions of **k·p** parameters.

For the DKK parameters L, M, N, there is another frequently used definition in the literature (e.g. used by Bir and Pikus [7] and in [53]), also called L, M, N, which often causes confusion in the numerical values of the parameters and thus even in some cases leading to incorrect use of parameters (see Sect. 6.2.2). They originally were used by Luttinger and Kohn [41]. There, they were termed A, B, C, and should not be confused with A, B, C of Eqs. (6.137), (6.138) and (6.139). These alternative LK (Luttinger–Kohn) parameters, labeled with superscript 'LK', read

$$L^{\mathrm{LK}} = A^{\mathrm{LK}} = L + \frac{\hbar^2}{2m_0} = (-\gamma_1 - 4\gamma_2)\,\frac{\hbar^2}{2m_0} \tag{6.33}$$

$$M^{\mathrm{LK}} = B^{\mathrm{LK}} = M + \frac{\hbar^2}{2m_0} = (2\gamma_2 - \gamma_1)\,\frac{\hbar^2}{2m_0} \tag{6.34}$$

$$N^{\mathrm{LK}} = C^{\mathrm{LK}} = N^{\mathrm{LK}+} + N^{\mathrm{LK}-} = N = -6\gamma_3\frac{\hbar^2}{2m_0} \tag{6.35}$$

$$N^{\mathrm{LK}+} = N^+ \approx N^{\mathrm{LK}} - \left(M^{\mathrm{LK}} - \frac{\hbar^2}{2m_0}\right) = N - M \tag{6.36}$$

$$N^{\mathrm{LK}-} = N^- \approx M^{\mathrm{LK}} - \frac{\hbar^2}{2m_0} = M\,. \tag{6.37}$$

Here, L^{LK} and M^{LK} are defined including the free electron term $\frac{\hbar^2}{2m_0}$ (see Eq. (6.12)). Expressing them using the Luttinger parameters now differs because the term '-1' that is present in Eqs. (6.14) and (6.15) has disappeared. Consequently, the diagonal term $\frac{\hbar^2}{2m_0}\mathbf{k}^2$ must be omitted in $\mathbf{H}_{\mathrm{vv}}$, i.e. the term $E_{\mathrm{v,av}} + \frac{\hbar^2}{2m_0}\mathbf{k}^2$ in Eq. (6.11) has to be replaced by $E_{\mathrm{v,av}}$. The inverse relations for the Luttinger parameters read

$$\gamma_1 = -\frac{1}{3}\left(L^{\mathrm{LK}} + 2M^{\mathrm{LK}}\right)\frac{2m_0}{\hbar^2} = -\frac{1}{3}\left(L + 2M\right)\frac{2m_0}{\hbar^2} - 1 \tag{6.38}$$

$$\gamma_2 = -\frac{1}{6}\left(L^{\mathrm{LK}} - M^{\mathrm{LK}}\right)\frac{2m_0}{\hbar^2} = -\frac{1}{6}\left(L - M\right)\frac{2m_0}{\hbar^2} \tag{6.39}$$

$$\gamma_3 = -\frac{1}{6}N^{\mathrm{LK}}\frac{2m_0}{\hbar^2} = -\frac{1}{6}N\frac{2m_0}{\hbar^2} \tag{6.40}$$

$$\kappa = -\frac{1}{6}(N^{\mathrm{LK}+} - N^{\mathrm{LK}-})\frac{2m_0}{\hbar^2} - \frac{1}{3} = -\frac{1}{6}(N^{+} - N^{-})\frac{2m_0}{\hbar^2} - \frac{1}{3}\,, \tag{6.41}$$

where the formula for γ_1 differs for the two possible definitions of the DKK or LK parameters. The term '-1' in Eq. (6.38) shows how they are related to each other. The relations for γ_2, γ_3, and κ correspond to the ones given in Eq. (6.23). In order to avoid confusion, we recommend to provide values for the Luttinger parameters when publishing or comparing material parameters, rather than the ambiguous L, M, N parameters.

For wurtzite the Rashba-Sheka-Pikus (RSP) parameters of the valence band A_1, A_2, A_3, A_4, A_5, A_6 are similar to the Luttinger parameters in zinc blende. The A_7 parameter that describes the $\mathbf{k}$ dependent spin-orbit splitting is usually neglected. From those six parameters one can obtain the nine parameters L_1, L_2, M_1, M_2, M_3, N_1^+, N_1^-, N_2^+, N_2^- as follows [25]

$$\begin{aligned}
L_1 &= \frac{\hbar^2}{2m_0}\left(A_5 + \tilde{A}\right) = -6(\sigma + \delta) \\
L_2 &= \frac{\hbar^2}{2m_0}\left(A_1 - 1\right) = -6\sigma_z \\
M_1 &= \frac{\hbar^2}{2m_0}\left(-A_5 + \tilde{A}\right) = -6\delta \\
M_2 &= \frac{\hbar^2}{2m_0}\hat{A} = -6\pi \\
M_3 &= \frac{\hbar^2}{2m_0}\left(A_2 - 1\right) = -6\pi_z \\
N_1^+ &= \frac{\hbar^2}{2m_0}\left(3A_5 - \tilde{A}\right) = N_1 - M_1 = -6(\sigma - \delta) \\
N_1^- &= \frac{\hbar^2}{2m_0}\left(-A_5 + \tilde{A}\right) = M_1 = -6\delta \\
N_2^+ &= \frac{\hbar^2}{2m_0}\left(\sqrt{2}A_6 - \hat{A}\right) = N_2 - M_2 = -6\sigma_{xz} \\
N_2^- &= \frac{\hbar^2}{2m_0}\hat{A} = M_2 = -6\pi\,,
\end{aligned} \tag{6.42}$$

where we used

$$\tilde{A} = A_2 + A_4 - 1 \tag{6.43}$$

$$\hat{A} = A_1 + A_3 - 1 \,. \tag{6.44}$$

It also holds

$$N_1 = N_1^+ + N_1^- = L_1 - M_1 = \frac{\hbar^2}{2m_0} 2A_5 = -6\sigma \tag{6.45}$$

$$N_2 = N_2^+ + N_2^- = \frac{\hbar^2}{2m_0} \sqrt{2} A_6 = -6\sigma_{xz} - 6\pi \,. \tag{6.46}$$

The 'Foreman' parameters σ, σ_z, σ_{xz}, π, π_z and δ have actually been introduced by Mireles and Ulloa [42]. The relation $N_1 = L_1 - M_1$ is due to the sixfold rotational symmetry of the Hamiltonian [17]. The related matrix elements are given in [36,42]. The contribution of the term '-1' (free electron term) for L_1, L_2, M_1, M_2 and M_3 has the same origin as in the zinc blende case. Thus one should be careful when comparing material parameters because for wurtzite there are also two definitions possible. The inverse relations to Eq. (6.42) read [17,42]

$$A_1 = \frac{2m_0}{\hbar^2} L_2 + 1 = 1 - 6\sigma_z \tag{6.47}$$

$$A_2 = \frac{2m_0}{\hbar^2} M_3 + 1 = 1 - 6\pi_z \tag{6.48}$$

$$A_3 = \frac{2m_0}{\hbar^2} (M_2 - L_2) = -6(\pi - \sigma_z) \tag{6.49}$$

$$A_4 = \frac{2m_0}{\hbar^2} \frac{1}{2} (L_1 + M_1 - 2M_3) = -3\sigma - 6(\delta - \pi_z) \tag{6.50}$$

$$A_5 = \frac{2m_0}{\hbar^2} \frac{1}{2} (L_1 - M_1) = \frac{2m_0}{\hbar^2} \frac{1}{2} N_1 = -3\sigma \tag{6.51}$$

$$A_6 = \frac{2m_0}{\hbar^2} \frac{\sqrt{2}}{2} N_2 = -\frac{6}{\sqrt{2}} (\pi + \delta) \,. \tag{6.52}$$

Finally, the 'Foreman' parameters can be expressed as

$$\sigma = -\frac{1}{6}(L_1 - M_1) \frac{2m_0}{\hbar^2} = -\frac{1}{6} N_1 \frac{2m_0}{\hbar^2} = -\frac{1}{3} A_5 \tag{6.53}$$

$$\sigma_z = -\frac{1}{6} L_2 \frac{2m_0}{\hbar^2} = -\frac{1}{6}(A_1 - 1) \tag{6.54}$$

$$\sigma_{xz} = -\frac{1}{6}(N_2 - M_2)\frac{2m_0}{\hbar^2} = -\frac{1}{6}N_2^+\frac{2m_0}{\hbar^2} = -\frac{1}{6}\left(-A_1 - A_3 + \sqrt{2}A_6 + 1\right) \tag{6.55}$$

$$= -\frac{1}{6}\left(\sqrt{2}A_6 - \hat{A}\right)$$

$$\pi = -\frac{1}{6}M_2\frac{2m_0}{\hbar^2} = -\frac{1}{6}N_2^-\frac{2m_0}{\hbar^2} = -\frac{1}{6}(A_1 + A_3 - 1) = -\frac{1}{6}\hat{A} \tag{6.56}$$

$$\pi_z = -\frac{1}{6}M_3\frac{2m_0}{\hbar^2} = -\frac{1}{6}(A_2 - 1) \tag{6.57}$$

$$\delta = -\frac{1}{6}M_1\frac{2m_0}{\hbar^2} = -\frac{1}{6}(A_2 + A_4 - A_5 - 1) = -\frac{1}{6}(\tilde{A} - A_5)\ . \tag{6.58}$$

H_{cc} in Eq. (6.10) is defined as

$$H_{\mathrm{cc}} = E_{\mathrm{c}} + k_x A_{\mathrm{c}} k_x + k_y A_{\mathrm{c}} k_y + k_z A_{\mathrm{c}} k_z \tag{6.59}$$

for zinc blende, and

$$H_{\mathrm{cc}} = E_{\mathrm{c}} + k_x A_{\mathrm{c}2} k_x + k_y A_{\mathrm{c}2} k_y + k_z A_{\mathrm{c}1} k_z \tag{6.60}$$

for wurtzite, where E_{c} is the conduction band edge (without strain shifts). The parameter A_{c} is defined as $A_{\mathrm{c}} = \frac{\hbar^2}{2m_0}S$ and S is a dimensionless parameter defined for zinc blende as [30]

$$S = 1 + 2F = \left(1 + 2\frac{1}{m_0}\sum_{n\in B}\frac{|\langle S\,|p_x|\,n\rangle|^2}{E_{\mathrm{c}} - E_n}\right)\ , \tag{6.61}$$

where the summation is over all bands apart from the valence bands. Here, F should not be confused with the parameter F of Eq. (6.19). In contrast to our definition of the valence band parameters L, M, N^+ and N^-, the conduction band parameter A_{c} includes the free electron term $\frac{\hbar^2}{2m_0}$ which corresponds to the '1' in Eq. (6.61). If the free electron term were not included, then it must be included explicitly into the H_{cc} term, analogous to the first line of Eq. (6.11) for the valence band part $\mathbf{H}_{\mathrm{vv}}$. S can also be evaluated through the experimentally determined conduction band mass m_{e} at the Γ point using the relation [38]

$$S = \frac{m_0}{m_{\mathrm{e}}} - \frac{2E_{\mathrm{P}}}{3E_{\mathrm{gap}}} - \frac{E_{\mathrm{P}}}{3\left(E_{\mathrm{gap}} + \Delta_{\mathrm{so}}\right)} = \frac{m_0}{m_{\mathrm{e}}} - E_{\mathrm{P}}\frac{E_{\mathrm{gap}} + \frac{2}{3}\Delta_{\mathrm{so}}}{E_{\mathrm{gap}}\left(E_{\mathrm{gap}} + \Delta_{\mathrm{so}}\right)}\ , \tag{6.62}$$

where E_{gap} is the (unstrained) band gap energy between the lowest conduction band edge and the highest valence band edge energy. If one wants to switch off the coupling between electrons and holes, simply setting $E_{\mathrm{P}} = 0$ eV leads to an isotropic and parabolic energy dispersion for the electrons. In this case, the S

parameter is the inverse of the effective electron mass, $S = \frac{m_0}{m_e}$, giving an intuitive meaning to this **k·p** parameter. In [54] the dimensionless F parameter (Eq. (6.61)) is given for all zinc blende materials, where $F = (S-1)/2$. However, as the band gap E_{gap} is temperature dependent, the next**nano** software by default calculates S directly from the actual band gap and the effective electron mass m_e, rather than using the S parameter of the database. We notice that different definitions of the S or A_c parameter occur in the literature depending on whether the free electron term is included or not. Thus one has to be careful when comparing different sets of material parameters. For wurtzite, the parameter A_{ci} is defined as $A_{ci} = \frac{\hbar^2}{2m_0} S_i$ with index $i = \{1, 2\}$. The index $i = 1$ refers to the direction parallel to the hexagonal c axis and $i = 2$ to the directions perpendicular to it. It holds $S_i = 1 + 2F_i$ and [17]

$$S_1 = \frac{m_0}{m_{e,\parallel}} - E_{P1} \frac{E_{\text{gap}} + 2\Delta_2}{(E_{\text{gap}} + \Delta_1 + \Delta_2)(E_{\text{gap}} + 2\Delta_2) - 2\Delta_3^2} \tag{6.63}$$

$$S_2 = \frac{m_0}{m_{e,\perp}} - E_{P2} \frac{(E_{\text{gap}} + \Delta_1 + \Delta_2)(E_{\text{gap}} + \Delta_2) - \Delta_3^2}{E_{\text{gap}}\left[(E_{\text{gap}} + \Delta_1 + \Delta_2)(E_{\text{gap}} + 2\Delta_2) - 2\Delta_3^2\right]} \,. \tag{6.64}$$

In the limit $\Delta_1 = 0$, $\Delta_2 = \Delta_3 = \frac{1}{3}\Delta_{so}$, the zinc blende result (Eq. (6.62)) is obtained.

The coupling between conduction and valence bands in Eq. (6.10) is defined as

$$\mathbf{H}_{cv} = \begin{pmatrix} H_{cv}^{SX} & H_{cv}^{SY} & H_{cv}^{SZ} \end{pmatrix} \,, \tag{6.65}$$

and

$$\mathbf{H}_{vc} = \begin{pmatrix} H_{vc}^{SX} \\ H_{vc}^{SY} \\ H_{vc}^{SZ} \end{pmatrix} \,. \tag{6.66}$$

For zinc blende, the components are given by

$$H_{cv}^{SX} = k_y B k_z + \mathrm{i} P k_x \tag{6.67}$$

$$H_{cv}^{SY} = k_z B k_x + \mathrm{i} P k_y \tag{6.68}$$

$$H_{cv}^{SZ} = k_x B k_y + \mathrm{i} P k_z \,, \tag{6.69}$$

and

$$H_{vc}^{SX} = k_z B k_y - \mathrm{i} k_x P \tag{6.70}$$

$$H_{vc}^{SY} = k_x B k_z - \mathrm{i} k_y P \tag{6.71}$$

$$H_{vc}^{SZ} = k_y B k_x - \mathrm{i} k_z P \,. \tag{6.72}$$

For wurtzite they are

$$H_{\mathrm{cv}}^{SX} = k_y B_1 k_z + \mathrm{i} P_2 k_x \tag{6.73}$$

$$H_{\mathrm{cv}}^{SY} = k_z B_2 k_x + \mathrm{i} P_2 k_y \tag{6.74}$$

$$H_{\mathrm{cv}}^{SZ} = k_x B_3 k_y + \mathrm{i} P_1 k_z \,, \tag{6.75}$$

and

$$H_{\mathrm{vc}}^{SX} = k_z B_1 k_y - \mathrm{i} k_x P_2 \tag{6.76}$$

$$H_{\mathrm{vc}}^{SY} = k_x B_2 k_z - \mathrm{i} k_y P_2 \tag{6.77}$$

$$H_{\mathrm{vc}}^{SZ} = k_y B_3 k_x - \mathrm{i} k_z P_1 \,. \tag{6.78}$$

We notice that $\mathbf{H}_{\mathrm{vc}}$ is not the Hermitian conjugate of $\mathbf{H}_{\mathrm{cv}}$ because the operator ordering is different. This ordering is the one suggested by Foreman [27]. The correct form of the inversion asymmetry parameter B has been derived by Loehr [37]. B is zero for materials that possess inversion symmetry as is the case for diamond-type crystals. Zinc blende crystals do not have inversion symmetry, thus $B \neq 0$. This also applies to wurtzite crystals, thus $B_1 \neq 0$, $B_2 \neq 0$ and $B_3 \neq 0$. For one-dimensional simulations, B only contributes for nonzero $\mathbf{k}_\parallel$ vectors. It is common practice in the **k**·**p** literature to neglect the contribution of the B parameter. However, this inversion asymmetry parameter is responsible for a spin-splitting of the bulk **k**·**p** dispersion for the bands along directions other than [001] and [111], where the splitting is suppressed by symmetry elements of the group of **k** [24] (see Fig. 6.9). For heterostructures the B parameter leads to a splitting of the states for any direction of $\mathbf{k}_\parallel$, even if the structure itself has an inversion center, i.e. no structural inversion asymmetry (SIA). In Sect. 6.2.4, we present an example (Fig. 6.15). The optical matrix parameter P (Kane momentum matrix element) that mixes the conduction and valence band states is given by

$$P = -\frac{\mathrm{i}\hbar}{m_0} \langle S | p_x | X \rangle = \sqrt{\frac{\hbar^2}{2m_0} E_{\mathrm{P}}} \tag{6.79}$$

$$E_{\mathrm{P}} = \frac{2m_0}{\hbar^2} P^2 \,, \tag{6.80}$$

where the Kane parameter E_{P} is the energy equivalent to P, whose values are usually around 22 eV for almost all semiconductors. They are similar for wurtzite

$$P_1 = -\frac{\mathrm{i}\hbar}{m} \langle S | p_z | Z \rangle = \sqrt{\frac{\hbar^2}{2m_0} E_{\mathrm{P1}}} \tag{6.81}$$

$$P_2 = -\frac{\mathrm{i}\hbar}{m} \langle S | p_x | X \rangle = -\frac{\mathrm{i}\hbar}{m} \langle S | p_y | Y \rangle = \sqrt{\frac{\hbar^2}{2m_0} E_{\mathrm{P2}}} \,. \tag{6.82}$$

In bulk, P only contributes for nonzero $\mathbf{k}$ vectors, whereas for heterostructures it is also relevant for $\mathbf{k}_\parallel = 0$.

Any algorithmic implementation of the **k·p** Hamiltonian should be the one of the wurtzite Hamiltonian. It implicitly contains the zinc blende and diamond-type cases by setting $L_1 = L_2 = L$, $M_1 = M_2 = M_3 = M$, $N_1^+ = N_2^+ = N^+$, $N_1^- = N_2^- = N^-$, $P_1 = P_2 = P$, $B_1 = B_2 = B_3 = B$ and $S_1 = S_2 = S$.

Spin-Orbit Coupling. The relativistic effect of spin is approximately taken into account by including an additional term $\mathbf{H}_{\text{so}}$ in the Schrödinger equation (Eq. (6.4)). Using the definition

$$\Delta_{\text{so}} = -3\mathrm{i}\left(\frac{\hbar^2}{4m_0^2c^2}\right)\left\langle X\left|(\nabla V\times\mathbf{p})_y\right|Z\right\rangle , \tag{6.83}$$

we can write down the spin-orbit interaction Hamiltonian for zinc blende [32]. In the basis of Eq. (6.8) it reads

$$\mathbf{H}_{\text{so}} = \frac{1}{3}\Delta_{\text{so}}\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & -\mathrm{i} & 0 & 0 & 0 & 1\\ 0 & 0 & \mathrm{i} & 0 & 0 & 0 & 0 & -\mathrm{i}\\ 0 & 0 & 0 & 0 & 0 & -1 & \mathrm{i} & 0\\ 0 & 0 & 0 & 0 & -1 & 0 & \mathrm{i} & 0\\ 0 & 0 & 0 & 0 & -\mathrm{i} & -\mathrm{i} & 0 & 0\\ 0 & 0 & 1 & \mathrm{i} & 0 & 0 & 0 & 0 \end{pmatrix} , \tag{6.84}$$

where Δ_{so} is the spin-orbit split-off energy. The spin-orbit Hamiltonian matrix can be diagonalized if one chooses the Bloch basis [38]

$$\begin{aligned}
|u_{\text{e}}\uparrow\rangle &= \left|\frac{1}{2},\frac{1}{2}\right\rangle_{\text{e}} = |S\uparrow\rangle \\
|u_{\text{e}}\downarrow\rangle &= \left|\frac{1}{2},-\frac{1}{2}\right\rangle_{\text{e}} = |\mathrm{i}S\downarrow\rangle \\
|u_{\text{hh}}\uparrow\rangle &= \left|\frac{3}{2},\frac{3}{2}\right\rangle = \frac{1}{\sqrt{2}}|(X+\mathrm{i}Y)\uparrow\rangle \\
|u_{\text{hh}}\downarrow\rangle &= \left|\frac{3}{2},-\frac{3}{2}\right\rangle = \frac{\mathrm{i}}{\sqrt{2}}|(X-\mathrm{i}Y)\downarrow\rangle \\
|u_{\text{lh1}}\rangle &= \left|\frac{3}{2},\frac{1}{2}\right\rangle = \frac{\mathrm{i}}{\sqrt{6}}[|(X+\mathrm{i}Y)\downarrow\rangle - 2|Z\uparrow\rangle] \\
|u_{\text{lh2}}\rangle &= \left|\frac{3}{2},-\frac{1}{2}\right\rangle = \frac{1}{\sqrt{6}}[|(X-\mathrm{i}Y)\uparrow\rangle + 2|Z\downarrow\rangle]
\end{aligned} \tag{6.85}$$

$$|u_{\mathrm{so1}}\rangle = \left|\frac{1}{2},\frac{1}{2}\right\rangle = \frac{1}{\sqrt{3}}\left[|(X+\mathrm{i}Y)\downarrow\rangle + |Z\uparrow\rangle\right]$$
$$|u_{\mathrm{so2}}\rangle = \left|\frac{1}{2},-\frac{1}{2}\right\rangle = \frac{\mathrm{i}}{\sqrt{3}}\left[|-(X-\mathrm{i}Y)\uparrow\rangle + |Z\downarrow\rangle\right] ,$$

leading to the eigenstates known as heavy hole (hh), light hole (lh) and spin-orbit split-off hole (so) with positive and negative angular momentum projection. The prefactors are normalization constants and these linear combinations are known as the *angular momentum representation*. The basis states for the electrons are included for completeness. They are not affected by the spin-orbit interaction. The heavy and light holes are degenerate and their eigenvalues are $\frac{1}{3}\Delta_{\mathrm{so}}$, whereas the eigenvalue of the split-off hole is $-\frac{2}{3}\Delta_{\mathrm{so}}$). Here we classified the six valence states in terms of $|\mathbf{J}, J_z\rangle$ states with the use of the Clebsch–Gordan coefficients for the angular momentum $\mathbf{J} = \mathbf{L}_{\mathrm{B}} + \mathbf{S}$ and its z component [21]. $\mathbf{L}_{\mathrm{B}}$ is the angular momentum of the Bloch orbit and $\mathbf{S}$ is the spin. The terms 'heavy' and 'light' originate from the distinct curvatures of the energy dispersions of these bands. For nonzero $\mathbf{k}$ values, the states cannot be labeled any more as pure heavy, light and split-off holes. Projecting the calculated solution in the basis of Eq. (6.8) onto the basis given in Eq. (6.85), determines if the corresponding spinors are dominated by the character of a heavy, light or split-off hole. Further details about the bulk band dispersion can be found for instance in [16]. As an example we provide the band structure of InAs in Fig. 6.9.

For wurtzite, the spin-orbit interaction Hamiltonian in the basis of Eq. (6.8) is given by

$$\mathbf{H}_{\mathrm{so}} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\mathrm{i}\Delta_2 & 0 & 0 & 0 & \Delta_3 \\ 0 & 0 & \mathrm{i}\Delta_2 & 0 & 0 & 0 & 0 & -\mathrm{i}\Delta_3 \\ 0 & 0 & 0 & 0 & 0 & -\Delta_3 & \mathrm{i}\Delta_3 & 0 \\ 0 & 0 & 0 & 0 & -\Delta_3 & 0 & \mathrm{i}\Delta_2 & 0 \\ 0 & 0 & 0 & 0 & -\mathrm{i}\Delta_3 & -\mathrm{i}\Delta_2 & 0 & 0 \\ 0 & 0 & \Delta_3 & \mathrm{i}\Delta_3 & 0 & 0 & 0 & 0 \end{pmatrix} , \tag{6.86}$$

with Δ_2 corresponding to the direction parallel to the hexagonal c axis, and Δ_3 to the plane perpendicular to it. It is usually assumed that $\Delta_2 = \Delta_3 = \frac{1}{3}\Delta_{\mathrm{so}}$, which actually yields the zinc blende spin-orbit Hamiltonian (Eq. (6.84)).

Crystal Field Splitting. In wurtzite we additionally have to consider the crystal field splitting, i.e. the energy splitting produced by the anisotropy of the hexagonal symmetry. In the basis of Eq. (6.8) it is given by

$$\mathbf{H}_{\mathrm{cr}} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \Delta_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \Delta_1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \Delta_1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \Delta_1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \tag{6.87}$$

where $\Delta_1 = \Delta_{\mathrm{cr}}$ is the crystal field splitting energy which is zero in both diamond and zinc blende materials.

Within the algorithm of the next**nano** program, it is sufficient to only implement the wurtzite spin-orbit Hamiltonian. It includes both crystal field splitting and spin-orbit coupling and reads

$$\mathbf{H}_{\mathrm{so,cr}} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \Delta_1 & -\mathrm{i}\Delta_2 & 0 & 0 & 0 & \Delta_3 \\ 0 & 0 & \mathrm{i}\Delta_2 & \Delta_1 & 0 & 0 & 0 & -\mathrm{i}\Delta_3 \\ 0 & 0 & 0 & 0 & 0 & -\Delta_3 & \mathrm{i}\Delta_3 & 0 \\ 0 & 0 & 0 & 0 & -\Delta_3 & \Delta_1 & \mathrm{i}\Delta_2 & 0 \\ 0 & 0 & 0 & 0 & -\mathrm{i}\Delta_3 & -\mathrm{i}\Delta_2 & \Delta_1 & 0 \\ 0 & 0 & \Delta_3 & \mathrm{i}\Delta_3 & 0 & 0 & 0 & 0 \end{pmatrix}. \tag{6.88}$$

It implicitly includes the zinc blende case. The basis states that diagonalize this wurtzite spin-orbit Hamiltonian are not the same as the ones we used before in the zinc blende case (Eq. (6.85)). They are [17]

$$\begin{aligned} &|\mathrm{i}S\uparrow\rangle \\ &|\mathrm{i}S\downarrow\rangle \\ &|u_1\uparrow\rangle = -\frac{1}{\sqrt{2}}|(X+\mathrm{i}Y)\uparrow\rangle \\ &|u_2\uparrow\rangle = \frac{1}{\sqrt{2}}|(X-\mathrm{i}Y)\uparrow\rangle \\ &|u_3\uparrow\rangle = |Z\uparrow\rangle \\ &|u_4\downarrow\rangle = \frac{1}{\sqrt{2}}|(X-\mathrm{i}Y)\downarrow\rangle \\ &|u_5\downarrow\rangle = -\frac{1}{\sqrt{2}}|(X+\mathrm{i}Y)\downarrow\rangle \\ &|u_6\downarrow\rangle = |Z\downarrow\rangle\ . \end{aligned} \tag{6.89}$$

After diagonalization, one obtains the following eigenvalues [17]

$$E_1 = E_\mathrm{A} = \Delta_1 + \Delta_2 \tag{6.90}$$

$$E_2 = E_\mathrm{B} = \frac{\Delta_1 - \Delta_2}{2} + \sqrt{\left(\frac{\Delta_1 - \Delta_2}{2}\right)^2 + 2\Delta_3^2} \tag{6.91}$$

$$E_3 = E_\mathrm{C} = \frac{\Delta_1 - \Delta_2}{2} - \sqrt{\left(\frac{\Delta_1 - \Delta_2}{2}\right)^2 + 2\Delta_3^2}\,, \tag{6.92}$$

where A, B and C refer to the heavy hole, light hole and crystal field split-off hole energies in wurtzite. For zinc blende one has to set $\Delta_1 = 0$ and $\Delta_2 = \Delta_3 = \frac{1}{3}\Delta_\mathrm{so}$. The zinc blende eigenvalues can further be simplified to

$$E_1 = E_2 = E_\mathrm{hh} = E_\mathrm{lh} = \frac{1}{3}\Delta_\mathrm{so} \tag{6.93}$$

$$E_3 = E_\mathrm{so} = -\frac{2}{3}\Delta_\mathrm{so}\,, \tag{6.94}$$

where heavy hole (hh) and light hole (lh) are degenerate. They are separated from the split-off hole (so) by the spin-orbit splitting energy Δ_so. Thus the valence band edges at Γ and the conduction band edges at Γ, L and X are determined as follows

$$E_\mathrm{c}^\Gamma = E_\mathrm{v,av} + \max(E_1, E_2) + E_\mathrm{gap}^\Gamma \tag{6.95}$$

$$E_\mathrm{c}^\mathrm{L} = E_\mathrm{v,av} + \max(E_1, E_2) + E_\mathrm{gap}^\mathrm{L} \tag{6.96}$$

$$E_\mathrm{c}^\mathrm{X} = E_\mathrm{v,av} + \max(E_1, E_2) + E_\mathrm{gap}^\mathrm{X} \tag{6.97}$$

$$E_\mathrm{hh} = E_\mathrm{v,av} + E_1 \tag{6.98}$$

$$E_\mathrm{lh} = E_\mathrm{v,av} + E_2 \tag{6.99}$$

$$E_\mathrm{so} = E_\mathrm{v,av} + E_3\,, \tag{6.100}$$

where we use the average of the three valence bands $E_\mathrm{v,av}$ as our reference point (Fig. 6.1). In fact, this reference is the average valence band edge energy in the absence of spin-orbit (and crystal field) splitting. This definition is valid for both zinc blende and wurtzite and is used to specify the valence band offset between different materials on a global scale [52]. Very often, however, the valence band offset is instead defined as the difference in energy with respect to the highest hole band edges between two materials.

Modified k·p Parameters. In a 6×6 **k·p** Hamiltonian all conduction bands are considered as a perturbation. In 8×8 **k·p** theory, the lowest conduction band is now included in the **k·p** Hamiltonian and not treated as a perturbation any more. Thus the related material parameters are different in 6×6 **k·p** and 8×8 **k·p** Hamiltonians.

Therefore, the L, M, N^+, N^- parameters in Eq. (6.11) must be replaced by the modified DKK parameters L', M', $N^{+\prime}$, $N^{-\prime}$ and N' because the latter include the now required correction term

$$L' = L + \frac{\hbar^2}{2m_0}\frac{E_\mathrm{P}}{E_\mathrm{gap}} = L + \frac{P^2}{E_\mathrm{gap}} \tag{6.101}$$

$$M' = M \tag{6.102}$$

$$N^{+\prime} = N^+ + \frac{\hbar^2}{2m_0}\frac{E_\mathrm{P}}{E_\mathrm{gap}} = N^+ + \frac{P^2}{E_\mathrm{gap}} \tag{6.103}$$

$$N^{-\prime} = N^- \tag{6.104}$$

$$N' = N + \frac{\hbar^2}{2m_0}\frac{E_\mathrm{P}}{E_\mathrm{gap}} = N + \frac{P^2}{E_\mathrm{gap}}\,. \tag{6.105}$$

The correction term is temperature dependent because obviously the band gap depends on temperature.

The modifications of the L_1, L_2, N_1^+, N_2^+, N_1 and N_2 parameters for wurtzite have to be done in a similar manner [4], whereas the M_1, M_2, M_3, N_1^-, N_2^- parameters remain unchanged as in the case of zinc blende

$$L_1' = L_1 + \frac{P_1^2}{E_\mathrm{gap}} \tag{6.106}$$

$$L_2' = L_2 + \frac{P_2^2}{E_\mathrm{gap}} \tag{6.107}$$

$$N_1^{+\prime} = N_1^+ + \frac{P_1^2}{E_\mathrm{gap}} \tag{6.108}$$

$$N_2^{+\prime} = N_2^+ + \frac{P_1 P_2}{E_\mathrm{gap}} \tag{6.109}$$

$$N_1' = N_1 + \frac{P_1^2}{E_\mathrm{gap}} \tag{6.110}$$

$$N_2' = N_2 + \frac{P_1 P_2}{E_\mathrm{gap}}\,. \tag{6.111}$$

For the RSP parameters the modifications read [4]

$$A_1' = A_1 + \frac{E_{\mathrm{P}2}}{E_\mathrm{gap}} \tag{6.112}$$

$$A_2' = A_2 \tag{6.113}$$

$$A_3' = A_3 - \frac{E_{\mathrm{P2}}}{E_{\mathrm{gap}}} \tag{6.114}$$

$$A_4' = A_4 + \frac{1}{2}\frac{E_{\mathrm{P1}}}{E_{\mathrm{gap}}} \tag{6.115}$$

$$A_5' = A_5 + \frac{1}{2}\frac{E_{\mathrm{P1}}}{E_{\mathrm{gap}}} \tag{6.116}$$

$$A_6' = A_6 + \frac{\sqrt{2}}{2}\frac{\sqrt{E_{\mathrm{P1}}E_{\mathrm{P2}}}}{E_{\mathrm{gap}}}\ , \tag{6.117}$$

where E_{P1} refers to the orientation parallel and E_{P2} perpendicular to the hexagonal c axis.

For zinc blende, it holds for the modified Luttinger parameters [46]

$$\gamma_1' = \gamma_1 - \frac{1}{3}\frac{E_{\mathrm{P}}}{E_{\mathrm{gap}}} \tag{6.118}$$

$$\gamma_2' = \gamma_2 - \frac{1}{6}\frac{E_{\mathrm{P}}}{E_{\mathrm{gap}}} \tag{6.119}$$

$$\gamma_3' = \gamma_3 - \frac{1}{6}\frac{E_{\mathrm{P}}}{E_{\mathrm{gap}}} \tag{6.120}$$

$$\kappa' = \kappa - \frac{1}{6}\frac{E_{\mathrm{P}}}{E_{\mathrm{gap}}}\ . \tag{6.121}$$

Finally, we list the modifications of the F, G, H_1, H_2 parameters

$$F' = F + \frac{\hbar^2}{2m_0}\frac{E_{\mathrm{P}}}{E_{\mathrm{gap}}} = F + \frac{P^2}{E_{\mathrm{gap}}} \tag{6.122}$$

$$G' = G \tag{6.123}$$

$$H_1' = H_1 \tag{6.124}$$

$$H_2' = H_2\ , \tag{6.125}$$

and Foreman's σ, π and δ parameters

$$\sigma' = \sigma - \frac{1}{6}\frac{E_{\mathrm{P}}}{E_{\mathrm{gap}}} \tag{6.126}$$

$$\pi' = \pi \tag{6.127}$$

$$\delta' = \delta\ . \tag{6.128}$$

For wurtzite, the latter are modified to

$$\sigma' = \sigma - \frac{1}{6}\frac{E_{\mathrm{P1}}}{E_{\mathrm{gap}}} \tag{6.129}$$

$$\sigma_z' = \sigma_z - \frac{1}{6}\frac{E_{\mathrm{P2}}}{E_{\mathrm{gap}}} \tag{6.130}$$

$$\sigma_{xz}' = \sigma_{xz} - \frac{1}{6}\frac{\sqrt{E_{\mathrm{P1}}E_{\mathrm{P2}}}}{E_{\mathrm{gap}}} \tag{6.131}$$

$$\pi' = \pi \tag{6.132}$$

$$\pi_z' = \pi_z \tag{6.133}$$

$$\delta' = \delta\,. \tag{6.134}$$

The next**nano** software provides full flexibility for the user with respect to choice of **k·p** parameters. The user can either specify the 6×6 **k·p** L, M, N parameters or the Luttinger parameters γ_1, γ_2, γ_3. The user can decide whether he wants to include κ or if he wants to approximate κ. The L', M', N' parameters for 8×8 **k·p** can be specified directly, or calculated automatically from the 6×6 **k·p** DKK or Luttinger parameters taking into account the temperature dependent band gap. Additionally, the modified Luttinger parameters for 8×8 **k·p** (with or without κ') can be entered instead. The user can specify the S parameter, or the program calculates S from the temperature dependent band gap and the effective electron mass. By default, a rescaling of the **k·p** parameters is not performed. However, the user can choose to rescale the parameters automatically to $S = 0$ or $S = 1$ (Eq. (6.158) or Eq. (6.159), respectively). This is sometimes necessary in order to avoid spurious solutions (see Sect. 6.2.1). In any case, a consistent set of all **k·p** parameters (DKK, Luttinger, Foreman, ...) are written out, also the ones not specified, including the 6×6 parameters if 8×8 parameters are specified. This gives as much transparency as possible to the user, especially when comparing own results with calculations and **k·p** parameters of published work by others. Additionally, the bulk **k·p** dispersion along the [001], [110] and [111] directions is part of the output, so that the user can check whether the energy dispersion for a particular choice of **k·p** parameters is meaningful. This is important to check for e.g. alloys, if strain is present or for rescaled **k·p** parameters, or if spurious solutions are present. Finally, we also output the calculated A, B, C parameters that have been used by Dresselhaus, Kip and Kittel [19]. They are of no practical relevance for the next**nano** software because they only apply along certain symmetry directions in the Brillouin zone where the energies are twofold degenerate, e.g. along the line from L to Γ to X. The DKK (Dresselhaus–Kip–Kittel) equation [19] for the energy dispersion of the holes is valid only for energies small compared to the spin-orbit splitting energy. Nevertheless, they provide insight into the choice of **k·p** parameters. From these parameters, the energy dispersion for the heavy and light holes are obtained

$$E_{\mathrm{hh}}(\mathbf{k}) = Ak^2 + \sqrt{B^2k^4 + C^2\left(k_x^2k_y^2 + k_y^2k_z^2 + k_z^2k_x^2\right)} \tag{6.135}$$

$$E_{\mathrm{lh}}(\mathbf{k}) = Ak^2 - \sqrt{B^2k^4 + C^2\left(k_x^2k_y^2 + k_y^2k_z^2 + k_z^2k_x^2\right)}\,. \tag{6.136}$$

The anisotropy is caused by C. Therefore the constant energy surfaces in reciprocal **k** space are 'warped spheres' and show a cubic symmetry. This deviation from the spherical symmetry is a direct consequence of the cubic crystal system. If C is zero, then the energy dispersion is isotropic (spherical approximation). If both, B and C are nonzero, the dispersion is nonparabolic. Both, A and B are negative with diamond being an exception, according to Yu and Cardona [56], although there is still uncertainty in the choice of parameters for diamond (see Sect. 6.2.5). A, B and C are related to the Luttinger and DKK parameters as follows

$$A = -\gamma_1\frac{\hbar^2}{2m_0} = \frac{L+2M}{3} + \frac{\hbar^2}{2m_0} \tag{6.137}$$

$$B = -2\gamma_2\frac{\hbar^2}{2m_0} = \frac{L-M}{3} \tag{6.138}$$

$$C^2 = 12\left(\gamma_3^2 - \gamma_2^2\right)\left(\frac{\hbar^2}{2m_0}\right)^2 = \frac{1}{3}(N^2 - (L-M)^2)\,. \tag{6.139}$$

Usually, the A, B, C and L, M, N parameters are given in $\frac{\hbar^2}{2m_0}$ units, and the Luttinger parameters in dimensionless units although some authors use different conventions, e.g. atomic units where $\frac{\hbar^2}{2m_0}$ leads to a factor of $1/2$ in the equations. Also opposite sign conventions for A, B, C are used, thus very often only $|A|$, $|B|$, $|C|$ are listed, and in fact only the sign of A, which is obvious, is relevant. If A and B are defined as in Eqs. (6.135) and (6.136), it is clear that A must be negative, but the signs for B and C are not determined from B^2 or C^2, respectively. Therefore, for the inverse relations it is not so obvious to assign the correct sign to γ_2 and γ_3, as the signs for B and C are not well defined

$$\gamma_1 = -A\frac{2m_0}{\hbar^2} \tag{6.140}$$

$$\gamma_2 = -\frac{B}{2}\frac{2m_0}{\hbar^2} \tag{6.141}$$

$$\gamma_3^2 = \left(\frac{B^2}{4} + \frac{C^2}{12}\right)\left(\frac{2m_0}{\hbar^2}\right)^2\,. \tag{6.142}$$

Usually for all cubic group IV, III–V and II–VI materials that we included in the next**nano** database, all three 6×6 **k·p** Luttinger parameters are positive. The only exceptions from this rule, that we are aware of, are the group IV materials

C (diamond), Sn and the II–VI materials HgS, HgSe and HgTe. For all these materials apart from diamond all three Luttinger parameters are negative. Apart from diamond, all have an unusual band structure having either a zero band gap or a negative band gap. For diamond several sets of Luttinger parameters exist (see Sect. 6.2.5). The signs of the Luttinger parameters vary among these sets. A possible reason for this is, to our believe, that some of the Luttinger parameters have been calculated from A, B and C.

It is often convenient to know estimates of effective masses in particular directions or averaged over all directions. They can be obtained by projection of the angular momentum operators onto a coordinate system that contains the desired direction. Within the next**nano** software, we write out the effective heavy and light hole masses along the [001], [110] and [111] directions. They can be extracted from the Luttinger parameters [31, 54]

$$\frac{m_0}{m_{\mathrm{hh}}^{[001]}} = \gamma_1 - 2\gamma_2 = \frac{2m_0}{\hbar^2}(-A + B) \tag{6.143}$$

$$\frac{m_0}{m_{\mathrm{lh}}^{[001]}} = \gamma_1 + 2\gamma_2 = \frac{2m_0}{\hbar^2}(-A - B) \tag{6.144}$$

$$\frac{m_0}{m_{\mathrm{hh}}^{[111]}} = \gamma_1 - 2\gamma_3 \tag{6.145}$$

$$\frac{m_0}{m_{\mathrm{lh}}^{[111]}} = \gamma_1 + 2\gamma_3 \tag{6.146}$$

$$\frac{m_0}{m_{\mathrm{hh}}^{[110]}} = \gamma_1 - \frac{1}{2}(\gamma_2 + 3\gamma_3) \tag{6.147}$$

$$\frac{m_0}{m_{\mathrm{lh}}^{[110]}} = \gamma_1 + \frac{1}{2}(\gamma_2 + 3\gamma_3)\ . \tag{6.148}$$

Also the isotropic, averaged heavy, light and split-off hole masses can be derived. For the latter, two approximations are commonly employed [54, 56]. The first one is

$$\frac{m_0}{m_{\mathrm{hh,av}}} = \frac{2m_0}{\hbar^2}\left(-A + B\left(1 + \frac{2C^2}{15B^2}\right)\right) \tag{6.149}$$

$$\frac{m_0}{m_{\mathrm{lh,av}}} = \frac{2m_0}{\hbar^2}\left(-A - B\left(1 + \frac{2C^2}{15B^2}\right)\right) \tag{6.150}$$

$$\frac{m_0}{m_{\mathrm{so,av}}} = \gamma_1 - \frac{E_{\mathrm{P}}\Delta_{\mathrm{so}}}{3E_{\mathrm{gap}}(E_{\mathrm{gap}} + \Delta_{\mathrm{so}})}\ . \tag{6.151}$$

The second one uses instead of Eqs. (6.149) and (6.150)

$$\frac{m_0}{m_{\mathrm{hh,av}}} = \frac{2m_0}{\hbar^2}\left(-A + \frac{2}{5}B\left(1 + \frac{3}{2}\sqrt{1 + \frac{4C^2}{9B^2}}\right)\right) \tag{6.152}$$

$$\frac{m_0}{m_{\mathrm{lh,av}}} = \frac{2m_0}{\hbar^2}\left(-A - \frac{2}{5}B\left(1 + \frac{3}{2}\sqrt{1 + \frac{4C^2}{9B^2}}\right)\right) . \tag{6.153}$$

Both approximations become identical if $C = 0$, corresponding to negligible warping [56].

Finally, we remark that it is sometimes useful to compare the numerical implementation of the **k·p** Hamiltonian to analytical or numerical single-band ('effective-mass') results at $\mathbf{k} = 0$, where the energy dispersion is isotropic and parabolic, and described by an effective mass m. By setting $E_{\mathrm{P}} = 0$ eV, one decouples the electrons from the holes. Then the 8×8 **k·p** Hamiltonian effectively becomes a 6×6 **k·p** Hamiltonian for the holes and a single-band Hamiltonian for the electrons, the latter being twofold degenerate due to spin. To be consistent, one then has to use the 6×6 **k·p** parameters for the holes (L, M, N^+, N^- rather than L', M', $N^{+\prime}$, $N^{-\prime}$), and the parabolic single-band effective mass m_{e} of the electron. This is achieved by setting $S = \frac{m_0}{m_{\mathrm{e}}}$. To obtain a dispersion for the holes that is both isotropic and parabolic, it requires us to set $L = M$, $N^+ = 0$ and $N^- = 0$ which implies $N = 0$. This is equivalent to setting $\gamma_2 = \gamma_3 = 0$, and $\kappa = -\frac{1}{3}$, or setting $F = G$ and $H_1 = H_2 = \frac{3}{2}F$. For instance, if we want to achieve a dispersion corresponding to an effective mass of $m_{\mathrm{h}} = 0.5m_0$ for each of the three hole bands, where the split-off band is separated from the degenerate heavy and light hole band energies by the spin-orbit splitting energy Δ_{so}, our **k·p** parameters must be given by

- $\gamma_1 = \frac{m_0}{m_{\mathrm{h}}} = 2$, $\gamma_2 = \gamma_3 = 0$, $\kappa = -\frac{1}{3}$ or
- $L = M = (-\gamma_1 - 1)\frac{\hbar^2}{2m_0} = -3\frac{\hbar^2}{2m_0}$, $N^+ = 0$, $N^- = 0$ or
- $F = G = \frac{L}{3} = -\frac{\hbar^2}{2m_0}$, $H_1 = H_2 = \frac{L}{2} = -\frac{3}{2}\frac{\hbar^2}{2m_0}$.
- (For the Foreman parameters it follows $\sigma = \delta = 1/6$ and $\pi = 1/4$. However, they implicitly assume $H_2 = 0$. This is definitely not the case here because now $H_2 \neq 0$ and its contribution is even larger than the one of F or G.)

These conditions are valid for bulk. They are also valid for heterostructures because we specified four parameters, except for the Foreman parameters. For bulk (without magnetic field), it is sufficient to specify only three parameters, and thus we can also use the Foreman parameters. To obtain an isotropic dispersion for the holes, it requires us to set $N = L - M$, i.e. $N^+ = L - 2M$ and $N^- = M$ (spherical approximation). This is equivalent to setting $\gamma_2 = \gamma_3$, leading to

$$\kappa = -\frac{1}{6}(L - 3M)\frac{2m_0}{\hbar^2} - \frac{1}{3} = -\frac{1}{3}(\gamma_1 - 5\gamma_2 + 2) . \tag{6.154}$$

For instance, setting $M = 0$, $N = L$, i.e. $N^+ = L$ and $N^- = 0$ yields an isotropic dispersion. This is equivalent to $\gamma_2 = \gamma_3 = \frac{1}{2}(\gamma_1 + 1) = -\frac{1}{6}L\frac{2m_0}{\hbar^2}$. The spherical approximation is also obtained by replacing γ_2 and γ_3 by

$$\hat{\gamma} = \frac{1}{5}(2\gamma_2 + 3\gamma_3) \ . \tag{6.155}$$

Replacing γ_2 and γ_3 by

$$\tilde{\gamma} = \frac{1}{2}(\gamma_2 + \gamma_3) \ , \tag{6.156}$$

yields the axial approximation, which is characterized by a cylindrical symmetry of the Hamiltonian, i.e. axial symmetry in the (x, y) plane.

If it holds $L = M$ (or $N = 0$), i.e. $N^+ = -M = -L$ and $N^- = M = L$, the hole dispersion is both isotropic and parabolic. This is equivalent to setting $\gamma_2 = \gamma_3 = 0$, leading to

$$\kappa = \frac{1}{3}L\frac{2m_0}{\hbar^2} - \frac{1}{3} = -\frac{1}{3}(\gamma_1 + 2) = -F\frac{2m_0}{\hbar^2} - \frac{1}{3} = 6\sigma - \frac{1}{3} \ . \tag{6.157}$$

If γ_1 is given, the four sets of parameters are related through $L = (-\gamma_1 - 1)\frac{\hbar^2}{2m_0}$, $F = -\frac{1}{3}L$ and $\sigma = \frac{L}{18}\frac{2m_0}{\hbar^2}$. In this case it holds $G = -2F = \frac{2}{3}L$, $H_1 = -3F = L$ and $H_2 = 0$. The consistent Foreman parameters are $\delta = -2\sigma = -\frac{1}{9}L\frac{2m_0}{\hbar^2}$ and $\pi = -3\sigma = -\frac{1}{6}L\frac{2m_0}{\hbar^2}$, implicitly assuming $H_2 = 0$. For these conditions it further holds $A = -\gamma_1\frac{\hbar^2}{2m_0}$, and $B = C^2 = 0$.

Temperature Dependent k·p Parameters. The following 8×8 **k·p** parameters are temperature dependent because they depend on the band gap E_{gap} at the Γ point.

- Zinc blende
  - S (Eq. (6.62))
  - the modified DKK parameters L', $N^{+\prime}$, N' (Eqs. (6.101), (6.103), (6.105))
  - the modified Luttinger parameters γ_1', γ_2', γ_3', κ' (Eqs. (6.118), (6.119), (6.120), (6.121))
  - the modified Foreman parameter σ' (Eq. (6.126))
  - F' (Eq. (6.122))
  - If the **k·p** parameters are rescaled, also the Kane parameter E_{P} (Eq. (6.158) or Eq. (6.159), respectively) is affected, and consequently also the Kane momentum matrix element P (Eq. (6.79)).
- Wurtzite
  - S_1, S_2 (Eqs. (6.63), (6.64))
  - the modified DKK parameters L_1', L_2', $N_1^{+\prime}$, $N_2^{+\prime}$, N_1', N_2' (Eqs. (6.106), (6.107), (6.108), (6.109), (6.110), (6.111))
  - the modified RSP parameters A_1', A_3', A_4', A_5', A_6' (Eqs. (6.112), (6.114), (6.115), (6.116), (6.117))
  - the modified Foreman parameters σ', σ_z', σ_{xz}' (Eq. (6.42))

– If the **k·p** parameters are rescaled, also the Kane parameters E_{P1} and E_{P2} (Eqs. (6.63), (6.64)) are affected, and consequently also the Kane momentum matrix elements P_1 and P_2 (Eqs. (6.81), (6.82)).

Summary. The zinc blende 8×8 **k·p** Hamiltonian can be parameterized by the following 10 quantities,

- the band gap energy E_{gap} at the Γ point,
- the spin-orbit split-off energy Δ_{so} (Δ_2, Δ_3, crystal field splitting Δ_1),
- the optical momentum matrix element E_{P} (E_{P1}, E_{P2}),
- the energy of the average of the three valence band edges $E_{\mathrm{v,av}}$,
- the conduction band mass at the Γ point m_{e} ($m_{\mathrm{e},\parallel}$, $m_{\mathrm{e},\perp}$),
- the Luttinger parameters γ_1, γ_2, γ_3, κ (A_1, A_2, A_3, A_4, A_5, A_6, A_7),
- the inversion asymmetry parameter B (B_1, B_2, B_3),

where the 19 wurtzite parameters are given in parenthesis if different from the zinc blende ones. To account for strain effects 4 additional parameters are required for zinc blende, and 7 for wurtzite (see [30] for details),

- the hydrostatic conduction band deformation potential a_{c}^{Γ} ($a_{\mathrm{c},\parallel}^{\Gamma}$, $a_{\mathrm{c},\perp}^{\Gamma}$),
- the valence band deformation potentials $a_{\mathrm{v,av}}$, b, d (d_1, d_2, d_3, d_4, d_5, d_6).

6.2 Applications

In this section results of **k·p** calculations for bulk semiconductors are shown for diamond, and for unstrained and strained InAs. Several examples of **k·p** calculations for heterostructures are presented. They cover spurious solutions, a spherical quantum dot and heterostructures showing the untypical type-II and type-III band alignments. Finally, self-consistent **k·p** calculations of a two-dimensional hole gas in diamond for different substrate orientations are analyzed. Wherever possible, the **k·p** results are compared to tight-binding calculations. All these calculations have been performed using the next**nano** software [8, 43].

6.2.1 Spurious Solutions

The numerical discretization of the single-band and **k·p** Schrödinger equations have been discussed in detail in [1, 30]. Everybody that numerically implements a **k·p** Hamiltonian will sooner or later encounter 'spurious solutions'. These are unphysical wave functions that look strange, e.g. if they oscillate very strongly or if they have spikes at material interfaces. Occasionally, their energies are even lying

within the forbidden band gap. These solutions might arise from incorrect operator ordering, i.e. incorrect discretization, or from the used **k·p** parameters. A detailed discussion on spurious solutions can be found in [1]. Veprek et al. [53] related the spurious solution problem to the loss of ellipticity of the differential operator. They derived a criteria that must be fulfilled by the **k·p** parameters to establish ellipticity. For all materials used in a calculation, the next**nano** software automatically outputs this information so that one can check if this criteria is fulfilled. Foreman [27] suggested to get rid of spurious solutions by setting $S = 0$ (Eq. (6.62)). This requires a rescaling of the E_P parameter

$$E_\mathrm{P} = \left(\frac{m_0}{m_\mathrm{e}}\right) \frac{E_\mathrm{gap}\left(E_\mathrm{gap} + \Delta_\mathrm{so}\right)}{E_\mathrm{gap} + \frac{2}{3}\Delta_\mathrm{so}} \,, \tag{6.158}$$

in order to still get the correct conduction band dispersion. Now E_P is fitted to the electron mass m_e, rather than having S as the fitting parameter. Essentially this implies that remote-band contributions cancel the free-electron term. An alternative is to set $S = 1$ which corresponds to entirely neglecting remote bands. The appropriate equation for E_P then reads

$$E_\mathrm{P} = \left(\frac{m_0}{m_\mathrm{e}} - 1\right) \frac{E_\mathrm{gap}\left(E_\mathrm{gap} + \Delta_\mathrm{so}\right)}{E_\mathrm{gap} + \frac{2}{3}\Delta_\mathrm{so}} \,. \tag{6.159}$$

To be consistent, the modified DKK (Eqs. (6.101), (6.103), (6.105)) or the modified Luttinger parameters (Eqs. (6.118), (6.119), (6.120), (6.121)) have to be recalculated using the new value of E_P. Analogous equations for the wurtzite case can be obtained from Eqs. (6.63) and (6.64).

According to Andlauer [1], for $(L' + 1)\, S < 0$ no oscillatory spurious solutions exist. He discussed spurious solutions for one-dimensional structures in [2]. Here we give the details of a simple two-dimensional example so that anyone interested in spurious solutions can reproduce these results. A quadratic InAs quantum wire of dimensions 5×5 nm is surrounded by a GaAs barrier of 10 nm thickness. The **k·p** parameters (without rescaling) for InAs are $\gamma_1 = 20.0$, $\gamma_2 = 8.5$, $\gamma_3 = 9.2$, $m_\mathrm{e} = 0.026 m_0$, $\Delta_\mathrm{so} = 0.39$ eV, $E_\mathrm{gap} = 0.417$ eV, $E_\mathrm{P} = 28.8$ eV and the ones for GaAs are $\gamma_1 = 6.98$, $\gamma_2 = 2.06$, $\gamma_3 = 2.93$, $m_\mathrm{e} = 0.067 m_0$, $\Delta_\mathrm{so} = 0.341$ eV, $E_\mathrm{gap} = 1.519$ eV, $E_\mathrm{P} = 21.5$ eV. The actual parameters that were used have been rescaled so that $S = 1$ (Eq. (6.159)) following the suggestions described above. For the valence band offset between InAs and GaAs we took $E_\mathrm{v,av}^\mathrm{VBO} = 0.044$ eV. This is the offset with respect to the average energy of all three hole band edges. The dimension N of the 8×8 **k·p** Hamiltonian matrix is $N = 6{,}728$ corresponding to a total of 29×29 grid points with a grid spacing of 1.0 nm in each direction. Figure 6.2 shows the probability density $\psi_1^2(x, y)$ of the ground state of this square InAs nanowire. While the 2D plot looks kind of okay, the horizontal and vertical slices through the center reveal some kinks in the density which should not be there. The energies of the ground state at $E_1 = 0.138$ eV (twofold degenerate) and of

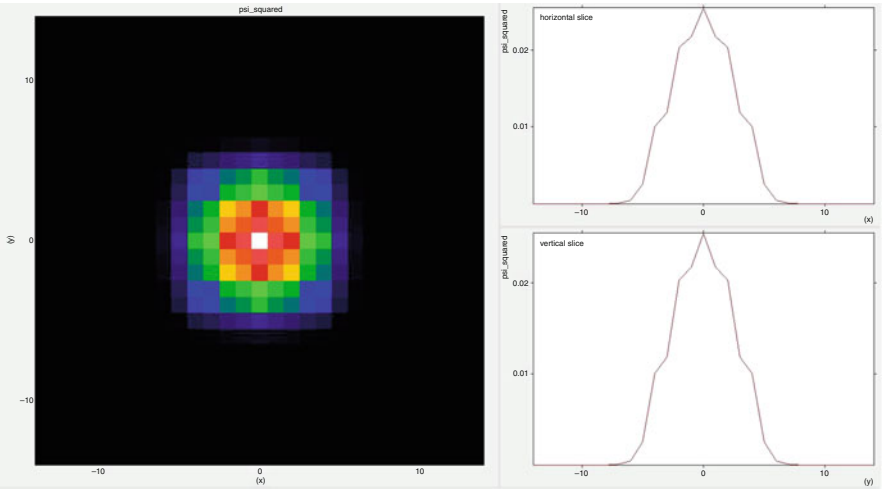


Fig. 6.2 Probability density $\psi_1^2(x, y)$ of the ground state of a square InAs nanowire. Also shown are horizontal and vertical slices through the center

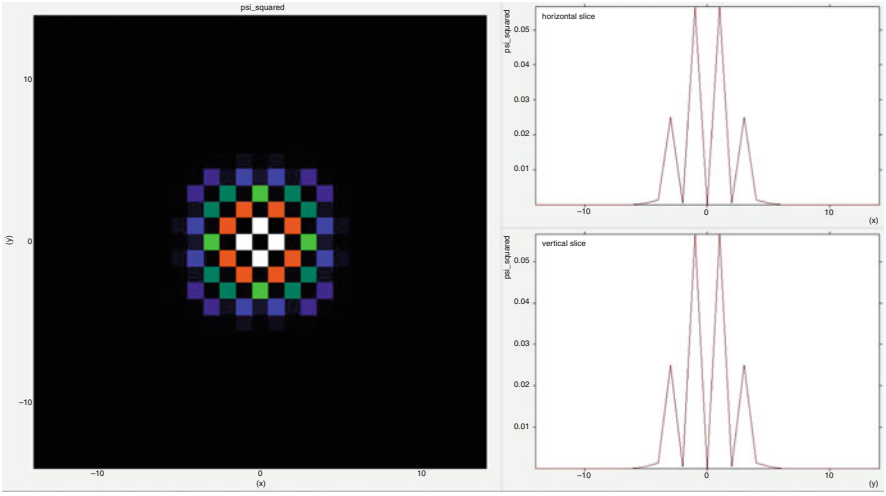


Fig. 6.3 Probability density $\psi_2^2(x, y)$ of the first excited state of a square InAs nanowire. This spurious solution shows a very pronounced oscillatory behavior. The *black squares* indicate zero probability density

the excited state at $E_2 = 0.220$ eV (fourfold degenerate) are within a reasonable range. Figure 6.3 shows the probability density $\psi_2^2(x, y)$ of the first excited state of this square InAs nanowire. This spurious solution is oscillating very strongly which can easily be seen. The black squares in the 2D plot indicate zero probability density.

It is very challenging to develop a self-consistent **k·p** Schrödinger–Poisson–current solver that will automatically detect spurious solutions if the eigenenergies are within a reasonable range (i.e. not within the band gap) and if relevant states that contribute to the density (e.g. the ground state in our example) show a 'reasonable' probability density, especially if the band edge profile is complicated due to alloy profiles, doping profiles, strain or piezoelectric fields. Rather than having a convenient 'black-box tool', the user must always check if the wave functions are reasonable. There are also situations possible where the probability density looks perfectly okay but the wave function itself is spurious [22]. In this case one is probably not able to recognize the spurious solution by looking at the density, band edge profile or energy levels. However, quantities that depend on the correctness of the wave functions, like matrix elements or calculated optical absorption spectra will likely show strange results. It is challenging to detect this. Recent progress on the topic of spurious solutions has been made by Eissfeller and Vogl. They developed a spurious-solution-free real-space multi-band envelope function approach that they termed 'symmetry adapted finite element method' (SAFE) [23].

6.2.2 Spin-Orbit Coupling in Silicon Quantum Dots

It is difficult to compare numerical results of a three-dimensional **k·p** calculation to analytical results due to its complexity. Thus one needs simple model systems where one can compare the results of different numerical implementations against each other. In this example we demonstrate the effect of spin-orbit coupling on the degeneracy of the eigenstates of a spherical silicon quantum dot (QD) with a diameter of 5 nm. We solve the 6×6 **k·p** Schrödinger equation for the hole eigenstates where we set the spin-orbit coupling energy either to $\Delta_{\mathrm{so}} = 44$ meV or to zero. For simplicity we assume infinite barriers at the QD boundaries which is a reasonable approximation as silicon nanocrystals are typically surrounded by a SiO_2 shell with a large valence band offset around 5 eV. The valence band edge energy inside the QD has been set to be 0 eV. The DKK parameters that have been used are

$$L = -6.8\frac{\hbar^2}{2m_0}\ ,\ M = -4.43\frac{\hbar^2}{2m_0}\ ,\ N = -8.61\frac{\hbar^2}{2m_0}\ , \tag{6.160}$$

that correspond to the Luttinger parameters $\gamma_1 = 4.22$, $\gamma_2 = 0.39$ and $\gamma_3 = 1.44$. They represent the anisotropy of the hole dispersion in silicon which is rather strong.

Figure 6.4 shows the energy spectrum of the spherical Si QD with (*red squares*) and without (*blue squares*) spin-orbit splitting. Without splitting there is a sixfold degeneracy of the ground state. Spin-orbit splitting reduces this degeneracy to fourfold. In general, each state is twofold degenerate due to spin. Additional geometric degeneracies arise due to the spherical symmetry of the QD. According to Burdov [11] one can calculate the ground state energy for this particular system from the L and M parameters with a high degree of accuracy using

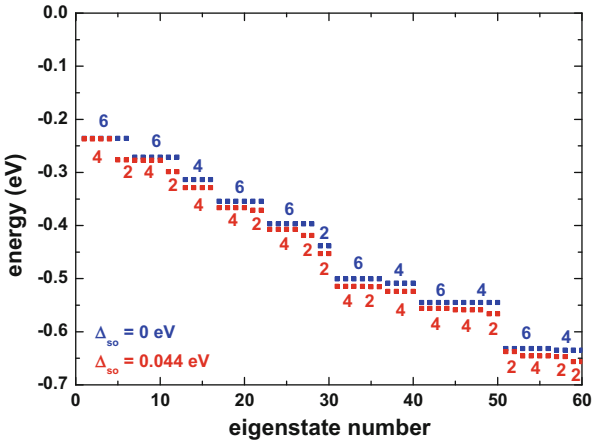


Fig. 6.4 Hole energy levels of a spherical silicon quantum dot of diameter 5 nm with (*blue squares*) and without (*red squares*) spin-orbit coupling calculated with the **k·p** method. The degeneracies of the levels are indicated by *numbers*

$$E_1 = -\frac{\hbar^2}{2m_\mathrm{h}}\frac{\pi^2}{R^2} \tag{6.161}$$

$$m_\mathrm{h} = \frac{3m_0}{(\frac{2m_0}{\hbar^2}L+1)+2(\frac{2m_0}{\hbar^2}M+1)} = \frac{3m_0}{\frac{2m_0}{\hbar^2}(L+2M)+3}\,, \tag{6.162}$$

where R is the radius of the sphere and $m_\mathrm{h} = -0.237m_0$ is the isotropic hole mass for our choice of parameters with the minus sign for the hole mass indicating that the dispersion in the bulk material is bent downwards. The calculated value for the ground state energy $E_1 = -0.254$ eV is close to our numerical value of -0.237 eV using the 6×6 **k·p** method. Burdov writes down Eq. (6.162) without the free electron term '+1' which is overall consistent within their article but their parameters [11] are unfortunately incorrect. The parameters given in Eq. (6.160) which are derived from the Luttinger parameters of Lawaetz [34] are the correct ones whereas the parameters termed L and M in Burdov's article are actually the L^LK and M^LK parameters (Eqs. (6.33), (6.34)). Thus they have to be corrected by '+1', i.e. $L^\mathrm{LK} = -5.8\frac{\hbar^2}{2m_0}$ and $M^\mathrm{LK} = -4.43\frac{\hbar^2}{2m_0}$ would have been the correct values to be consistent within their article (see also the discussion on the different and thus confusing definitions of the L and M parameters in Sect. 6.1). Consequently, their calculated hole mass of $m_\mathrm{h} = -0.19m_0$ has to be corrected by our value of $m_\mathrm{h} = -0.237m_0$.

For comparison, the results of the single-band calculations with isotropic heavy, light and split-off hole masses are shown in Fig. 6.5 for the same quantum dot. Here, each state is twofold degenerate due to spin but only one of these two energy levels is shown. The numbering of the horizontal axis, however, is taking

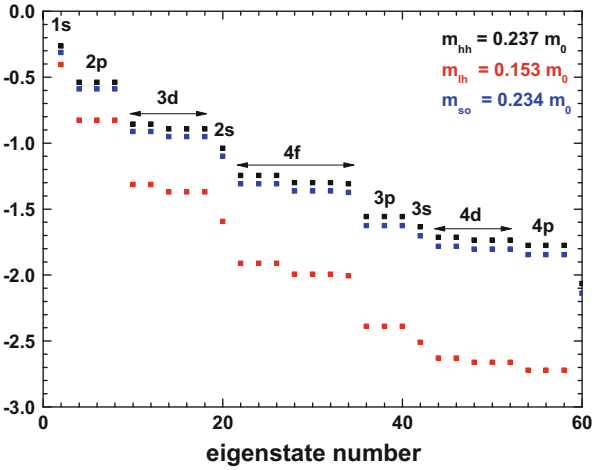


Fig. 6.5 Hole energy levels of a spherical silicon quantum dot ('artificial atom') of diameter 5 nm for heavy (*black squares*), light (*red squares*) and split-off holes (*blue squares*) calculated with the single-band Schrödinger equation. The degeneracies of the levels are indicated by the standard *atomic orbital notation* known from the electron configuration of atoms

spin into account. Figure 6.4 shows the lowest 60 eigenvalues. This numbering corresponds roughly to the first 20 eigenvalues for each hole species in Fig. 6.5. The degeneracies of the levels are indicated by the standard atomic orbital notation known from the electron configuration of atoms ($1s$, $2p$, $3d$, $2s$, $4f$, ...). From that labeling it becomes clear why quantum dots are termed 'artificial atoms'. In contrast to real atoms, their energy levels can be tuned to match technologically relevant energy regimes, while showing qualitatively similar energy spectra as atoms, at least for spherical dots. Self-organized quantum dots, however, are mainly showing an energy spectrum similar to a two-dimensional harmonic oscillator as their confinement potential in the plane perpendicular to the growth direction can often be approximated by a parabolic confinement. The fivefold degeneracy of the d levels and the sevenfold degeneracy of the f level is not reproduced well because the ideal shape of the spherical QD is approximated by a QD discretized on a rectangular grid having cubic symmetry. Obviously, the numerical single-band results are very poor in comparison to the **k·p** results with the exception of the ground state energy $E_1 = -0.265$ eV.

For this particular spherical geometry, the eigenvalues are highly degenerate, not only due to spin but also due to geometry. This sometimes causes problems for certain eigenvalue solvers as they might miss some of these degenerate eigenvalues. For instance, our implementation of the Arnoldi method that uses Chebyshev polynomials as preconditioner [51] missed some degenerate eigenvalues. For this reason it is of great advantage if any numerical software has redundancy in terms of several eigensolvers, where one can choose from, in order to check results

for consistency and accuracy, as well as performance. The ARPACK eigenvalue solver [35] took around 6 min for 60 eigenvectors where the dimension of the matrix was $N = 55{,}566$ corresponding to $21 \times 21 \times 21 = 9{,}261$ grid points with a grid resolution of 0.25 nm in each direction.

6.2.3 *Type-III Broken-Gap Band Alignment: HgTe–CdTe Quantum Well*

HgTe is an interesting material for studies of the intrinsic spin Hall effect [10] and the quantum spin hall effect [6], or spin splitting effects in general due to its large Rashba-type spin-orbit splitting. HgTe is a zero-gap semiconductor that can be embedded between CdTe layers to form a HgTe–CdTe quantum well (QW) heterostructure which shows an interesting type-III band alignment where the valence band edge in the HgTe QW lies above its conduction band edge. Due to this band alignment it is not possible to apply a single-band Hamiltonian. Thus a **k·p** or tight-binding approach is required. Large HgTe quantum wells have an inverted band structure where the highest hole state (h_1) lies above the lowest electron state (e_1). For smaller quantum well widths, the quantum confinement increases and below a critical well width, the band structure becomes normal again with the electron state above the hole state. Figure 6.6 shows the square of the calculated **k·p** wave functions of e_1 and h_1 at the crossover well width at 6.5 nm. Increasing the well width shifts the e_1 state below the h_1 state. This is shown in Fig. 6.7 where the probability density of the relevant states have been calculated with the empirical tight-binding method for a 7.8 nm HgTe quantum well. One can nicely see that in the tight-binding method the envelope of the probability density corresponds to **k·p** envelope functions. For the $sp^3d^5s^*$ tight-binding [33] calculations, we used a valence band offset of 0.4 eV. For the **k·p** calculations, we took exactly the same material parameters as in [44], including their valence band offset of 0.570 eV. In both cases, we neglected strain effects for simplicity.

Figure 6.8 shows the energies of the electron and hole states in a HgTe–CdTe quantum well as a function of HgTe QW width calculated with the 8×8 **k·p** method. The crossover of normal to inverted band structure occurs around 6.5 nm and corresponds to the situation in Fig. 6.6. The *dashed lines* indicate the energetic positions of the conduction and valence band edges of the HgTe QW. Our results for the crossover width are in good agreement to the calculations of Novik et al. [44], and also close to tight-binding calculations (not shown).

The next**nano** software package allows to perform systematic comparisons between the **k·p** and the tight-binding method for quantum wells.

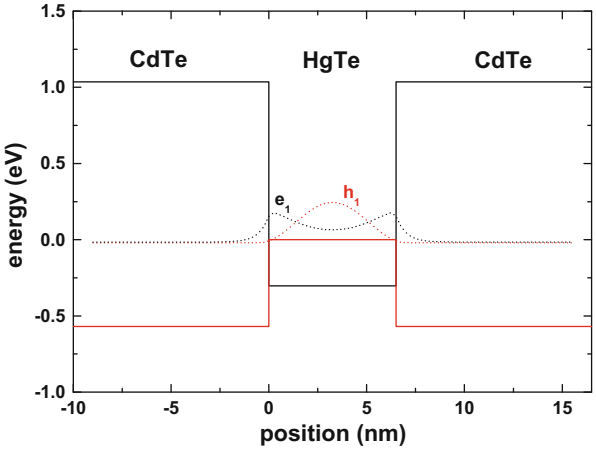


Fig. 6.6 Probability density of the lowest electron (e_1) and highest hole (h_1) eigenstates of a 6.5 nm HgTe quantum well calculated with the **k·p** method. In the **k·p** method, the eigenstates correspond to envelope functions. The conduction (*black solid line*) and valence band edges (*red solid line*) form a type-III band alignment

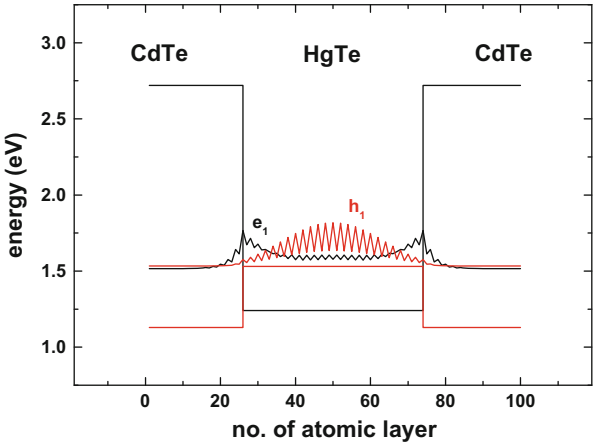


Fig. 6.7 Probability density of the lowest electron (e_1) and highest hole (h_1) eigenstates of a 7.8 nm HgTe quantum well calculated with the empirical tight-binding method

6.2.4 Type-II Broken-Gap Band Alignment: InAs–GaSb Superlattice

Type-II broken gap heterostructures have the unusual property that in one material the lowest conduction band edge energy lies below the highest valence band edge energy of the material next to it. Consequently, they cannot be modeled by a

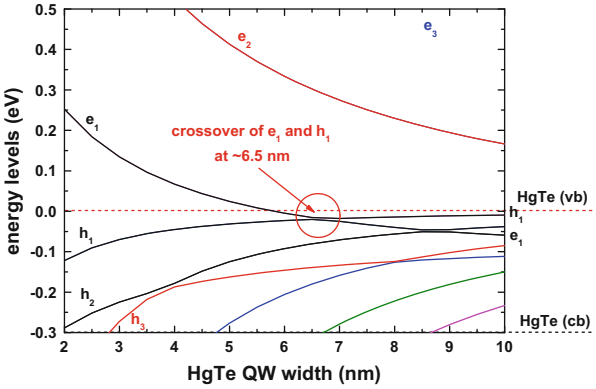


Fig. 6.8 Calculated energies of the electron and hole states in a HgTe–CdTe quantum well as a function of HgTe QW width (8×8 **k·p**). The crossover of normal to inverted band structure occurs around 6.5 nm and corresponds to the situation in Fig. 6.6. The *dashed lines* indicate the conduction and valence band edges of the HgTe QW

single-band model, similar as the type-III structures of the previous section. There will be a strong coupling of the electron and hole wave functions, in some cases even a very strong wave function hybridization, making it eventually impossible to distinguish between electron and hole states if the lowest 'conduction' band state lies below the highest 'valence' band state [3, 29, 57]. In this section, we calculate the energy dispersion of an InAs–GaSb superlattice along the in-plane directions $\mathbf{k}_\parallel$ and along the superlattice direction $\mathbf{k}_{\mathrm{SL}}$ by two different methods, the 8×8 **k·p** and the $sp^3d^5s^*$ [33] tight-binding method. Such structures are relevant for infrared detectors or to study semimetal–semiconductor phase transitions. Remarkably, for certain layer widths the in-plane subband dispersion becomes linear, indicating vanishing effective masses [3]. In such a situation, the band structure of the InAs–GaSb superlattice is similar to the band structure around the Dirac point in graphene, where a two-dimensional gas of massless Dirac fermions is formed. It is interesting to notice that such properties can also be achieved by growing InAs–GaSb superlattices. However, such topics are not part of our study here.

First, we examine the bulk band structure of InAs along the [110] and [100] directions in **k** space at zero temperature. We compare our 8×8 **k·p** model to the single-band effective mass dispersions and to the more sophisticated $sp^3d^5s^*$ tight-binding model which allows one to calculate the bulk band structure in the whole Brillouin zone accurately also for higher-lying conduction bands, e.g. the ones with minima at the X points in the Brillouin zone. The latter is also possible for a **k·p** model where more than eight bands are included [50]. Figure 6.9 shows the results. The parabolic model (*dotted lines*) is reasonable only for small **k** vectors around the Γ point. We used effective masses of $m_{\mathrm{e}} = 0.023m_0$, $m_{\mathrm{hh}} = 0.41m_0$, $m_{\mathrm{lh}} = 0.026m_0$ and $m_{\mathrm{hh}} = 0.14m_0$ for the electron, heavy hole, light hole and split-off hole, respectively. The energies for the band gap $E_{\mathrm{gap}} = 0.417$ eV and for the

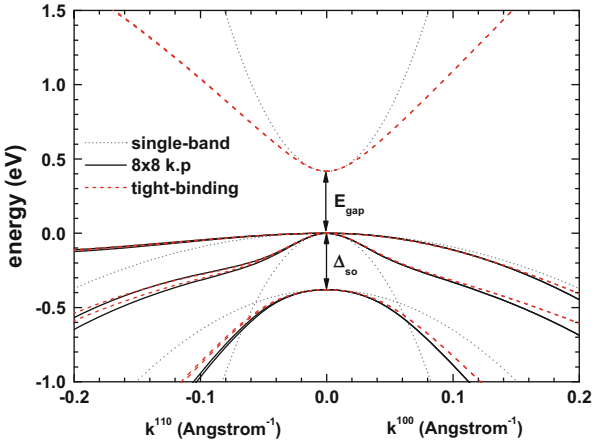


Fig. 6.9 Energy dispersion of bulk InAs along the [110] and [100] directions in **k** space calculated with the parabolic model (*dotted lines*), the 8×8 **k·p** model (*black solid lines*) and with the $sp^3d^5s^*$ tight-binding parameterization (*red dashed lines*) at $T = 0$ K. Along the [100] direction the energies are twofold degenerate while this spin degeneracy is lifted along the [110] direction for both the **k·p** and the tight-binding model

split-off energy $\Delta_{\mathrm{so}} = 0.38$ eV, which are obtained from experiment, are the same in all models and are indicated by the arrows. For the parabolic and the **k·p** model, these values are input parameters while the tight-binding parameters are fitted to yield these values. InAs is a material where the band gap and the split-off energy are of similar energy. This is typically not the case for the group IV or most other III–V materials. Along the [100] direction the energies are twofold degenerate in all models. Due to symmetry arguments, also along the [111] directions the energies are twofold spin degenerate (not shown). However, this spin degeneracy is lifted along the [110] direction for both the **k·p** (*black solid lines*) and the tight-binding model (*red dashed lines*). For both, intersubband transitions in InAs quantum wells and density calculations, where the carriers are located mainly around the Gamma point, only small **k** values are relevant. For small **k** values, the **k·p** energy dispersion is reasonably close to the tight-binding dispersion for both directions, also in terms of spin-splitting. In this calculation, the **k·p** inversion asymmetry parameter B (Eq. (6.67)) was taken to be nonzero ($B = 3.60\frac{\hbar^2}{2m_0}$ [15]). If it were zero, which is the typical assumption in practically all **k·p** calculations, a spin-splitting along the [110] direction would not be obtained. We conclude that the **k·p** model is sufficiently accurate as a description of the realistic band structure. Of course, the tight-binding approach is also a model but it has been fitted to energies and masses through the entire Brillouin zone, so one can assume that it is a very reasonable model.

We now examine the bulk band structure of biaxially, tensilely strained InAs with respect to a GaSb substrate. The biaxial strain $\varepsilon_{\parallel} = 0.0062$ is with respect to the (x, y) plane (which we call in-plane) and the strain $\varepsilon_{\perp} = -0.0067$ is with respect to the z direction (which we call out-of-plane direction). Figure 6.10

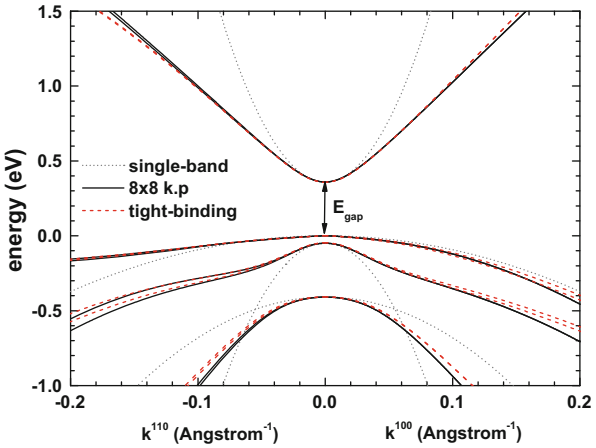


Fig. 6.10 Energy dispersion of biaxially, tensilely strained InAs along the [110] and [100] directions (in-plane directions) calculated with the parabolic model (*dotted lines*), the 8×8 **k·p** model (*black solid lines*) and with the $sp^3d^5s^*$ tight-binding parameterization (*red dashed lines*) at $T = 0$ K. Along the [100] direction the energies are still twofold degenerate for the **k·p** model while this spin degeneracy is lifted for the tight-binding model

shows the energy dispersion of biaxially, tensilely strained InAs along the [110] and [100] directions (in-plane directions) in **k** space calculated with the parabolic model (*dotted lines*), the 8×8 **k·p** model (*black solid lines*) and with the $sp^3d^5s^*$ tight-binding parameterization (*red dashed lines*) at $T = 0$ K. Along the [100] direction the energies are still twofold degenerate for the **k·p** model, while this spin degeneracy is now lifted for the tight-binding model. The single-band results have the same masses as for unstrained InAs but the band edges are shifted in the same way as for the **k·p** model. The energies of these band edges can easily be obtained by diagonalizing the bulk **k·p** Hamiltonian, that includes the deformation potentials and strain [30], at $\mathbf{k} = 0$. Alternatively, for growth along any direction, analytical equations can be used instead to obtain the shifted and split band edges due to strain for the single-band model [52]. The common zero point of energy for all three models has been set to the highest hole energy level. The band gap has decreased to $E_{\mathrm{gap}} = 0.359$ eV because the unit cell has increased, corresponding to a positive hydrostatic strain of $\varepsilon_{\mathrm{hydro}} = 0.0057$. (If the unit cell increases due to increasing temperature, the band gap also gets smaller.) The degeneracy of the heavy and light hole band edges at the Γ point is now lifted and they are separated by 0.049 eV.

Figure 6.11 shows the same as Fig. 6.10 but along the [001] (out-of-plane) and [100] (in-plane) directions. Along both directions the energies are twofold degenerate for the **k·p** model, while this spin degeneracy is lifted for the tight-binding model. Along the [001] (out-of-plane) direction the dispersion is now very different. The heavy and light hole dispersions cross along the out-of-plane direction. Now the highest hole band is 'light' along the out-of-plane direction, and 'heavy' along the in-plane directions. The first excited hole state is however 'heavy'

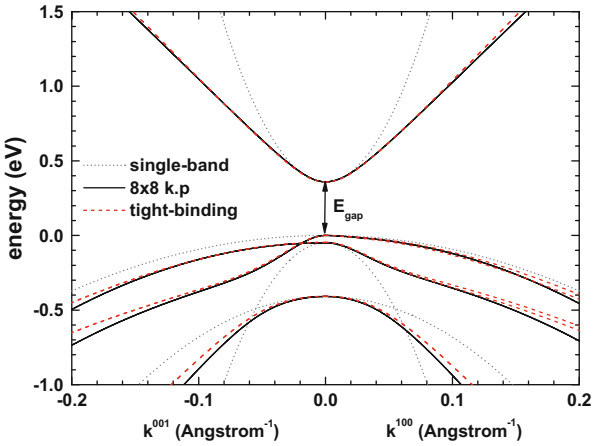


Fig. 6.11 Energy dispersion along the [001] (out-of-plane) and [100] (in-plane) directions for the same situation as in Fig. 6.10. Along the [001] (out-of-plane) direction the dispersion is now very different. Both, the **k·p** (*black solid lines*) and the tight-binding model (*red dashed lines*) show very similar results although the dispersion is now rather complicated

along the out-of-plane direction, and 'light' along the in-plane directions. Both, the **k·p** and the tight-binding model show very similar results although the dispersion is now rather complicated. This gives further confidence into our **k·p** model.

The next**nano** software package allows to calculate the **k·p** and the tight-binding bulk band structure for unstrained or biaxially strained zinc blende materials, and also for ternary alloys, making it very convenient to perform systematic comparisons between the bulk **k·p** and the tight-binding method.

Now we turn to the InAs–GaSb heterostructure. Our structure consists of a 3.7 nm (24 atomic layers) GaSb and a 6.7 nm (44 atomic layers) InAs region. Similar structures have been investigated theoretically with the **k·p** method by e.g. Grein et al. [29] or Zakharova et al. [57]. Periodic boundary conditions are used to mimic a superlattice with a periodic length of $L = 10.4$ nm corresponding to $k_{\mathrm{SL,max}} = \pi/L = 0.30\ \mathrm{nm}^{-1}$. Strain has been included assuming that the GaSb layer is unstrained and that the InAs layer is biaxially strained with respect to a GaSb substrate with $\varepsilon_{\parallel} = 0.0062$ and $\varepsilon_{\perp} = -0.0067$, i.e. InAs is tensilely strained. In order to avoid spurious solutions, the **k·p** material parameters were rescaled according to Eq. (6.159) so that $S = 1$ (see Sect. 6.2.1). These rescaled parameters were also used in the previous figures for the band structure of bulk InAs. The grid spacing resolution was 0.1 nm in the **k·p** calculation. The tight-binding calculations have been performed with the same method as in the previous section. For both the bulk **k·p** and the tight-binding calculation, the conduction and the three valence band edge energies at $\mathbf{k} = 0$ are identical. This holds for strained InAs (Figs. 6.10 and 6.11) and for GaSb. However, it cannot be avoided that slight deviations in the bulk dispersion occur for nonzero **k** vectors due to the different methods employed,

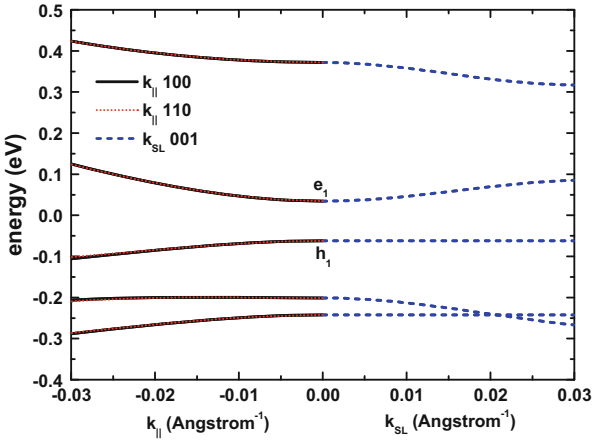


Fig. 6.12 Energy dispersion of a strained InAs–GaSb superlattice along the in-plane directions in $\mathbf{k}$ space, $\mathbf{k}_{\parallel}^{100}$ and $\mathbf{k}_{\parallel}^{110}$, and along the superlattice growth direction $\mathbf{k}_{\mathrm{SL}}$ calculated with the $\mathbf{k}\cdot\mathbf{p}$ method. For small $\mathbf{k}_{\parallel}$ vectors the dispersion is practically isotropic

in particular if strain is present. Consequently, it is expected that for heterostructure calculations, the $\mathbf{k}\cdot\mathbf{p}$ and tight-binding calculations deviate even more.

Figure 6.12 shows the energy dispersion of the strained InAs–GaSb superlattice along the in-plane directions in $\mathbf{k}$ space, $\mathbf{k}_{\parallel}^{100}$ and $\mathbf{k}_{\parallel}^{110}$, and along the superlattice growth direction $\mathbf{k}_{\mathrm{SL}}$ calculated with the 8×8 $\mathbf{k}\cdot\mathbf{p}$ method. For small $\mathbf{k}_{\parallel}$ vectors the dispersion is practically isotropic. The highest hole state h_1 does not have any curvature along the superlattice direction $\mathbf{k}_{\mathrm{SL}}$. The reason is that the ground state hole wave function is strongly confined in the individual GaSb layers (not shown) and does not couple to neighboring GaSb layers. All other states couple to neighboring layers and thus form minibands. The corresponding tight-binding results are shown for comparison in Fig. 6.13. Now the dispersion along $\mathbf{k}_{\parallel}$ is split by strain and shows a slight anisotropy. The spin-splitting is due to the fact that zinc blende materials do not have inversion symmetry. This is called bulk inversion asymmetry (BIA). Furthermore, the interfaces do not have a common atom, such as the As anion in AlAs–GaAs heterostructures. The latter can have a mirror plane if the number of arsenic layers is odd. Thus for our tight-binding calculations, no mirror plane is present. This is called structural inversion asymmetry (SIA). Consequently, both BIA and SIA were involved. The B parameter that is related to the missing inversion symmetry in zinc blende materials was assumed to be zero in the $\mathbf{k}\cdot\mathbf{p}$ calculations of Fig. 6.12. Structural asymmetry due to the atomic arrangement of the interfaces cannot be taken into account by the $\mathbf{k}\cdot\mathbf{p}$ method. However, spin-splitting due to structural asymmetry introduced by e.g. an electric field or by a heterostructure that does not have an inversion center, is automatically taken into account by the $\mathbf{k}\cdot\mathbf{p}$ method. Within the $\mathbf{k}\cdot\mathbf{p}$ model, our superlattice structure has a mirror plane, i.e. our structure is symmetric, therefore spin-splitting

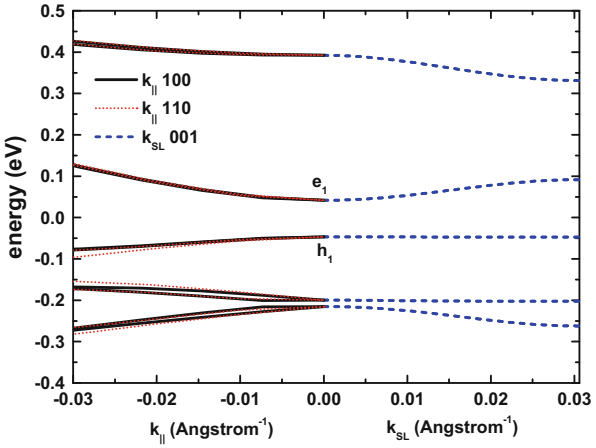


Fig. 6.13 Same as Fig. 6.12 but now showing the results of the tight-binding calculations. The dispersion along $\mathbf{k}_\parallel$ is split by strain and shows a slight anisotropy

cannot be obtained here because SIA is absent, and BIA was ignored. Both the tight-binding and the **k·p** calculations show qualitatively very similar results, apart from the crossing of the second and third holes states along the superlattice direction. Also the energy levels look very similar. There are rare cases in the literature where the B parameter was actually set to a nonzero value. Typically it is said that it can be neglected or that the parameter is not known. Cartoixà [15] lists values for a few III–V materials. In Fig. 6.14 we used his values of $B = 13.1\frac{\hbar^2}{2m_0}$ (GaSb) and $B = 3.60\frac{\hbar^2}{2m_0}$ (InAs). The isotropic dispersion for small $\mathbf{k}_\parallel$ vectors, that we had for $B = 0$, is now slightly lifted. Also the twofold spin-degeneracy along the $\mathbf{k}_\parallel$ directions is no longer present. The order of magnitude of the spin-splitting is comparable to the tight-binding results. Finally, we show the energy dispersion of a strained InAs–GaSb superlattice along the in-plane directions $\mathbf{k}_\parallel^{100}$ and $\mathbf{k}_\parallel^{110}$ calculated with the **k·p** method, where the inversion asymmetry parameter B is zero (*black solid lines*), and where it is nonzero (*red dotted lines*) to highlight the differences (Fig. 6.15). In the latter case the twofold spin degeneracy of the energy levels for nonzero $\mathbf{k}_\parallel$ is lifted due to the bulk inversion asymmetry parameter B as discussed above. Therefore the results are closer to the tight-binding calculations. We conclude that it is important to use a nonzero value for the inversion asymmetry parameter B in order to get more realistic results.

6.2.5 Valence Band Structure of Diamond

Intrinsic diamond is an insulator with an indirect band gap of 5.5 eV at room temperature. In this section we discuss the valence band structure of diamond

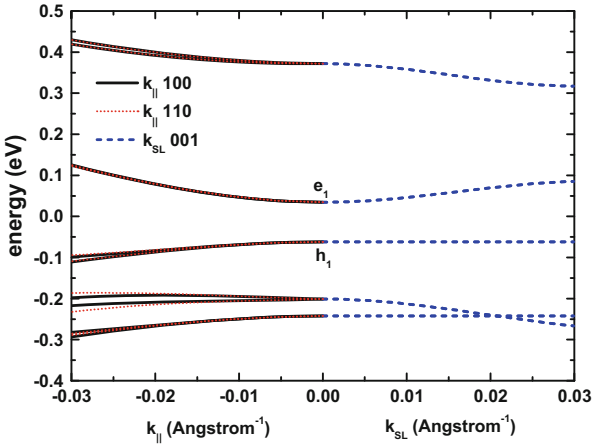


Fig. 6.14 Energy dispersion of a strained InAs–GaSb superlattice along the in-plane directions in $\mathbf{k}$ space, $\mathbf{k}_{\parallel}^{100}$ and $\mathbf{k}_{\parallel}^{110}$, and along the superlattice growth direction $\mathbf{k}_{\mathrm{SL}}$ calculated with the $\mathbf{k}\cdot\mathbf{p}$ method, where the inversion asymmetry parameter B has a nonzero value. The isotropic dispersion for small $\mathbf{k}_{\parallel}$ vectors is now slightly lifted. Also the twofold spin-degeneracy along $\mathbf{k}_{\parallel}$ is no longer present

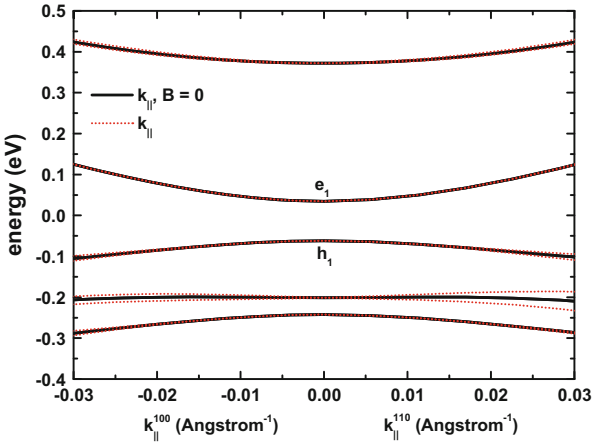


Fig. 6.15 Energy dispersion of a strained InAs–GaSb superlattice along the in-plane directions $\mathbf{k}_{\parallel}^{100}$ and $\mathbf{k}_{\parallel}^{110}$ calculated with the $\mathbf{k}\cdot\mathbf{p}$ method, where the inversion asymmetry parameter B is zero (*black solid lines*) and where it is nonzero (*red dotted lines*). In the latter case, the twofold spin degeneracy of the energy levels for nonzero $\mathbf{k}_{\parallel}$ is lifted due to the bulk inversion asymmetry parameter B

and compare the energy dispersion along several directions in $\mathbf{k}$ space. For the $\mathbf{k}\cdot\mathbf{p}$ energy dispersions a number of different sets of Luttinger parameters that can be found in the literature will be tested against each other. Surprisingly, there is still a substantial lack of information about the details of the band structure of

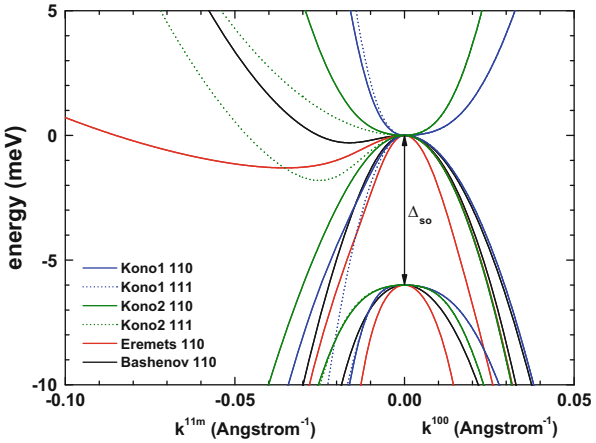


Fig. 6.16 Comparison of different sets of Luttinger parameters where the 6×6 **k·p** valence band energy dispersion leads to incorrect curvature of the hole bands. Negative x axis: energy dispersion along the [110] ($m = 0$, *solid lines*) and [111] ($m = 1$, *dotted lines*) directions; positive x axis: energy dispersion along the (100) direction (*solid lines*)

diamond. It seems that there is not much known on the precise values of the valence band masses in diamond. Willatzen et al. [55] collected eight different sets of Luttinger parameters from different authors. Further comparisons have been made by Gheeraert et al. [28] and Reggiani et al. [49]. So we have at least 13 different sets of Luttinger parameters. Only two of them are similar (the ones of Saslow et al. and van Vetchen et al.), all others deviate more or less substantially. Previous work did not compare the resulting valence band structures for these parameters. We found that for actually four of these sets of Luttinger parameters the hole dispersion bends into the opposite direction (negative mass). Figure 6.16 shows our results. We only plot the energy dispersion along the directions [110] (*solid lines*) and [111] (*dotted lines*) (left part of the figure) where the sets of parameters lead to incorrect curvature of the hole bands. The dispersion along [100] is plotted in all cases (right part of the figure, *solid lines*). In diamond the split-off energy is very small ($\Delta_{\mathrm{so}} = 6$ meV) and indicated by the arrow. As we discard these sets of Luttinger parameters in the following, we do not list the actual Luttinger parameters and their references. Instead we refer to Table IV in [55] and references therein. The parameters by Eremets et al. (*black solid lines*) and Bashenov et al. (*red solid lines*) lead to a dispersion along the [110] direction where the curvature is almost flat and has negative slope at $|k| > 0.035$ Å^{-1} (Eremets) or at $|k| > 0.016$ Å^{-1} (Bashenov). This fact was already discussed by Bashenov. Kono et al. have published two sets of parameters. For both the first set (*blue lines*) and the second set (*green lines*), the dispersion of the uppermost hole state has negative slope along all three directions [100], [110] and [111]. Furthermore, the second set (*green dotted line*) leads to negative slope for values of $|k| > 0.025$ Å^{-1} along the [111] direction. This makes us believe that not all authors actually calculated the band structure for their set

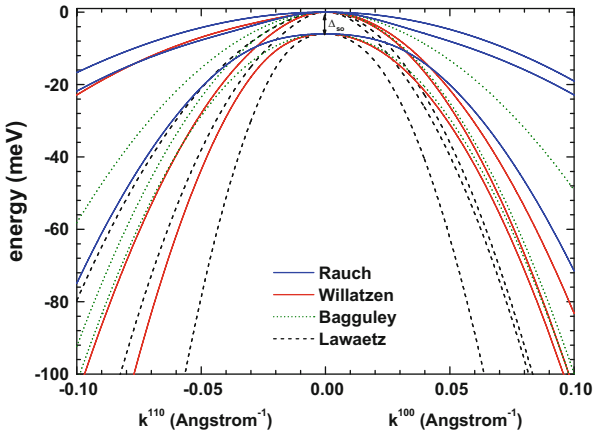


Fig. 6.17 6×6 **k·p** valence band energy dispersion along the [110] and [100] directions for different sets of Luttinger parameters

of Luttinger parameters. Luttinger parameters are usually determined by cyclotron resonance experiments. This was the method employed by Rauch [47] who derived an effective light hole mass of $m_{\mathrm{lh}} = 0.70m_0$ that showed little anisotropy and a split-off hole mass of $m_{\mathrm{so}} = 1.06m_0$ with no anisotropy. He also derived a value for the heavy hole mass of $m_{\mathrm{hh}} = 2.18m_0$ [48]. We used these values in our previous work [18] where we employed a single-band model. We note that the A, B and C parameters of the early work of Rauch are significantly different to all other sets of Luttinger parameters, and there seems to be even doubt on the correct interpretation of his experiment [28].

Figures 6.17 and 6.18 show the dispersion along [110] and [111], respectively, as well as along [100] for other sets of Luttinger parameters. We did not include the sets published by van Haeringen et al., Reggiani et al., Saslow et al., van Vetchen et al. and Hall. The purpose of the figures is to demonstrate that there is significant variance among those parameters, and that the energy dispersion of the heavy hole for the Rauch parameters corresponds to a much larger, i.e. heavier hole mass than for other sets of Luttinger parameters.

As there is currently no consensus about any experimentally derived set of Luttinger parameters, we also investigated the band structure obtained from the $sp^3d^5s^*$ tight-binding (TB) parameterization of Jancu et al. [33]. They published empirical TB parameters for diamond. However, this set does not lead to a spin-orbit splitting at $\mathbf{k} = 0$ and thus all three hole bands are degenerate at the Γ point. The reason for this is that they set their $\Delta/3$ parameter to zero. We adjusted this parameter to $\Delta/3 = 0.0024$ eV so that we now obtain a split-off energy of $\Delta_{\mathrm{so}} = 0.006$ meV. This change is negligible with respect to the overall band structure so that we can still use the original Jancu parameterization for all other parameters. Only the valence band structure at the Γ point is affected by our choice of $\Delta/3$. We compared the **k·p** dispersions for all sets of Luttinger parameters to

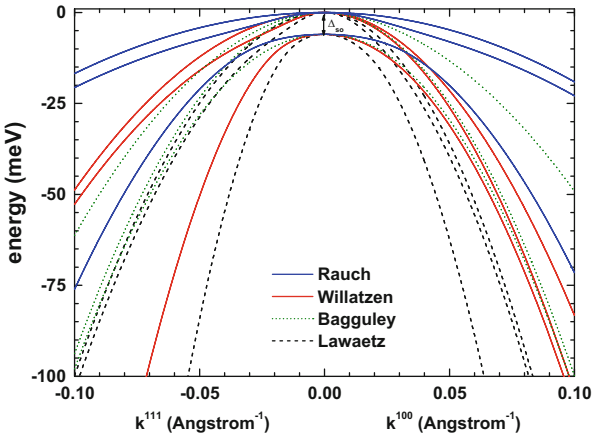


Fig. 6.18 Same as Fig. 6.17 but along the [111] and [100] directions

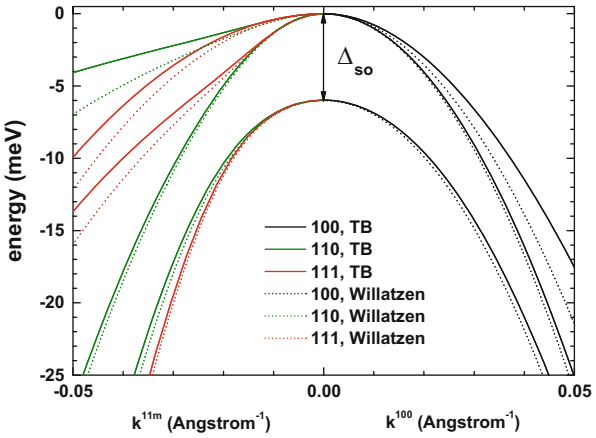


Fig. 6.19 Energy dispersion along the high symmetry directions [100] (*black*), [110] (*green*), [111] (*red*) in **k** space calculated by the tight-binding (TB) method (*solid lines*). For comparison the results obtained by diagonalizing the bulk 6×6 **k·p** Hamiltonian for each **k** vector is shown for the Luttinger parameters of Willatzen et al. (*dotted lines*)

our calculated tight-binding band structure (not shown). The Luttinger parameters of Willatzen et al. were the ones that were closest to the TB valence band structure along all three high symmetry directions as shown in Fig. 6.19. This is the reason why we use Willatzen's parameters in the following although we do not claim that this is the 'best' choice. Their parameters have been derived within the framework of density-functional theory (local density approximation, self-consistent scalar-relativistic linear muffin-tin-orbital method). Having established our set of Luttinger parameters we are now prepared to analyze diamond surfaces as a function of substrate orientation in the next section.

6.2.6 *Self-Consistent Calculations: Influence of Substrate Orientations on the Density of a Two-Dimensional Hole Gas in Diamond*

In this section we analyze the properties of the two-dimensional hole gas in surface conducting diamond as a function of substrate orientation. We are using a 6×6 **k·p** Hamiltonian to describe the charge accumulation at the diamond surface. This self-consistent **k·p** formalism [30] includes nonparabolicity effects and warping. Self-consistent solutions of the Schrödinger–Poisson equation for hydrogen-terminated diamond (surface conducting diamond) have previously been performed by Edmonds et al. [20]. They considered a single-band model with effective masses derived from the Luttinger parameters of Willatzen et al. [55] for diamond. In contrast to the single-band model with parabolic and isotropic masses, the **k·p** model allows us to compare different diamond substrate orientations, namely (100), (110) and (111) with respect to their sensitivity. However, an important ingredient for **k·p** calculations are the Luttinger parameters that describe the hole masses. For diamond we discussed these parameters in the preceding section and are now prepared to investigate the density of a two-dimensional hole gas (2DHG) in diamond for different substrate orientations at room temperature. Our model system consists of diamond where the 2DHG is induced by a constant surface charge density. In real samples the 2DHG is induced by the hydrogen termination of the diamond surface. This results in a negative electron affinity of about $\chi = -1$ eV which causes p-type surface conductivity (surface conducting diamond). We solve the 6×6 **k·p** Schrödinger–Poisson equation self-consistently for (100), (110) and (111) substrate orientations. As boundary conditions we used a negative interface charge density of $\sigma = -5 \times 10^{13}$ cm^{-2} at the diamond surface in order to induce a 2DHG, and a flat band boundary condition in the bulk diamond, i.e. zero electric field far away from the interface. The doping concentration was assumed to be n-type in the whole diamond (0.5×10^{18} cm^{-3}, nitrogen with ionization energy $E_{\mathrm{D}}^{\mathrm{ion}} = 1.7$ eV).

Figure 6.20 shows our results. It can be clearly seen that the (111) substrate orientation (*red dashed lines*) has a higher density than the (100) orientation (*black solid lines*). Even a higher density is obtained for the (110) orientation (*blue dash-dotted lines*). The corresponding 2DHG sheet densities are $\sigma_{110} = 5.8 \times 10^{12}$ cm^{-2}, $\sigma_{111} = 4.5 \times 10^{12}$ cm^{-2} and $\sigma_{100} = 3.0 \times 10^{12}$ cm^{-2}, respectively. The different results for each orientation are due to the anisotropy of the valence band structure. It can also be seen from the sheet densities that the interface charge σ is not completely canceled (screened) by the 2DHG sheet charge density. Thus there is additional band bending further away from the 2DHG region and not a flat band. This also reveals the difference between the capacitance of an ideal parallel plate capacitor and the 'quantum capacitance' [39] of a two-dimensional electron or hole gas. We used the Luttinger parameters of Willatzen et al., motivated by the discussion of the previous section. The parameters by Reggiani et al., Saslow et al., van Haeringen et al. and van Vetchen et al. show a higher anisotropy (warping) in terms of [111] vs.

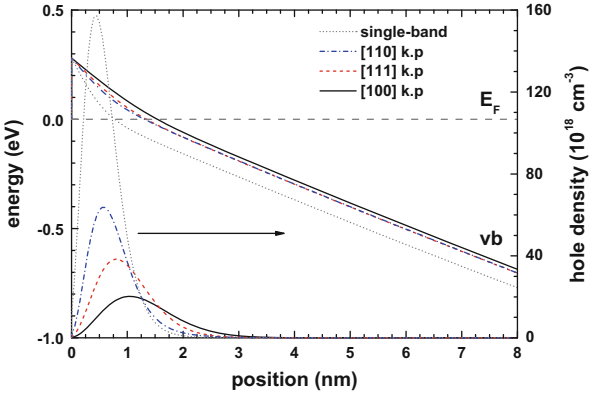


Fig. 6.20 Hole densities and valence band edges (vb) of a two-dimensional hole gas in diamond for various substrate orientations ((110)—*blue dash-dotted lines*, (111)—*red dashed lines*, (100)—*black solid lines*). The single-band results are independent of substrate orientation (*gray dotted lines*) because isotropic masses are used. The Fermi level E_F is indicated by the *gray dashed line*

[100] directions leading to a slightly larger difference with respect to the respective densities in this example (not shown). For comparison we also show the single-band results (*gray dotted lines*) obtained with the parabolic and isotropic effective masses by Rauch ($m_{hh} = 2.18m_0$, $m_{lh} = 0.70m_0$, $m_{so} = 1.06m_0$) where for each of the three valence band edges the single-band Schrödinger equation was solved. In this case the results are independent of substrate orientation. As these masses are much heavier than the masses by Willatzen (see also Figs. 6.17 and 6.18), the density is larger than for the **k·p** formalism. In fact, the single-band density is almost entirely due to the occupation of the highest heavy hole ground state. Also for **k·p** along [110] only the ground state contributes to the density. This is consistent as the mass along [110] is much heavier than along the other directions (see Fig. 6.19). For [111] the ground state and the first excited state contribute to the density because they have very similar energies, whereas for [100] the ground state and the first two excited states contribute to the density. The latter two excited states also have very similar energies. This can be seen in Fig. 6.21 which shows the three uppermost eigenstates (square of the probability amplitude shifted by its eigenenergy) for each substrate orientation at $\mathbf{k}_\parallel = 0$. In fact because of spin the six highest eigenstates are shown but the two spin states have the same energy and the same probability density at $\mathbf{k}_\parallel = 0$. Only for [110] (*blue dash-dotted lines*) the energy of the ground state lies above the Fermi energy (*gray dashed line*), compare also with Fig. 6.25. The probability densities of the uppermost three states have only one maximum for [100] (*black solid lines*) and [111] (*red dashed lines*) directions, i.e. they are derived from the ground states of the heavy, light and split-off hole but these are in fact mixed states. Only for [100] at $\mathbf{k}_\parallel = 0$, the second state is a pure heavy hole, and for [111] the ground state is 50 % heavy and 50 % light with no contribution from split-off hole at $\mathbf{k}_\parallel = 0$. All other states shown in this figure are mixed states. For

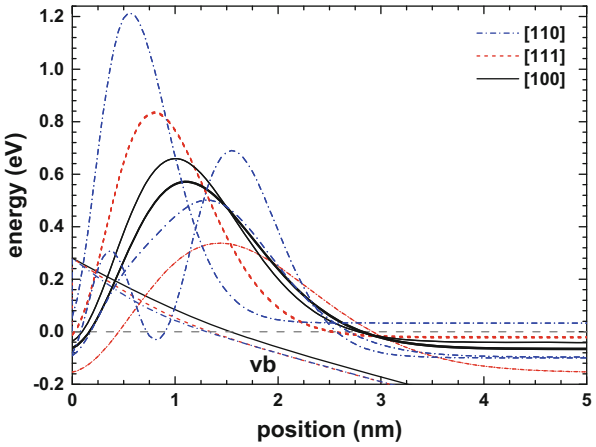


Fig. 6.21 Three uppermost eigenstates (square of the probability amplitude shifted by its eigenenergy) for each orientation at $\mathbf{k}_\parallel = 0$, valence band edge energies (vb) and Fermi level

[110] the second excited state has one node. For all orientations our calculations indicate that states with one node are not occupied (not shown). For confinement along [110] direction, the maximum of the probability density is much closer to the surface. This will have influence on the capacitance of diamond biosensors as the 2DHG density is then more sensitive to potential changes at the surface. In the figure it looks as if only the first two states are plotted for [100] and [111] but in fact three states are plotted for each. The reason is that for [100] the two excited states have almost the same energy (separated by 4 meV) and the same shape, whereas for [111] the ground state and the first excited state have almost the same energy (also separated by 4 meV) and the same shape (see also Fig. 6.25 where the energies are shown in more detail). As the triangular-like confinement potential is very strong, the 2DHG is located within the first few nanometers. This justifies to use a small quantum region of 10 nm (nonuniform grid, 151 quantum grid points) with Dirichlet boundary conditions. The $\mathbf{k}_\parallel = (k_x, k_y)$ space has been discretized on $41 \times 41 = 1{,}681$ $\mathbf{k}$ points with a maximum value of $k_{\max} = 0.18$ Å^{-1} along the k_x and k_y directions.

Figures 6.22, 6.23 and 6.24 show the $\mathbf{k}_\parallel$-resolved hole density distribution $p(k_x, k_y)$ for (110), (111) and (100) oriented diamond, respectively. In these plots one can see the amount that each $\mathbf{k}_\parallel$ point contributes to the density. The $\mathbf{k}_\parallel = (k_x, k_y)$ space refers to the rotated coordinate system and thus the k_x and k_y direction are different for each orientation (see figure captions). In all cases the growth direction is meant to be along the z direction. The $\mathbf{k}_\parallel$-resolved hole density distribution for (100) oriented diamond has to be symmetric with respect to k_x and k_y (Fig. 6.24). The $\mathbf{k}_\parallel$-resolved hole density distribution for (111) looks symmetric with respect to k_x and k_y (Fig. 6.23). A detailed analysis of the energy dispersion (see Fig. 6.25) reveals that the dispersion along k_x is in fact very similar to the

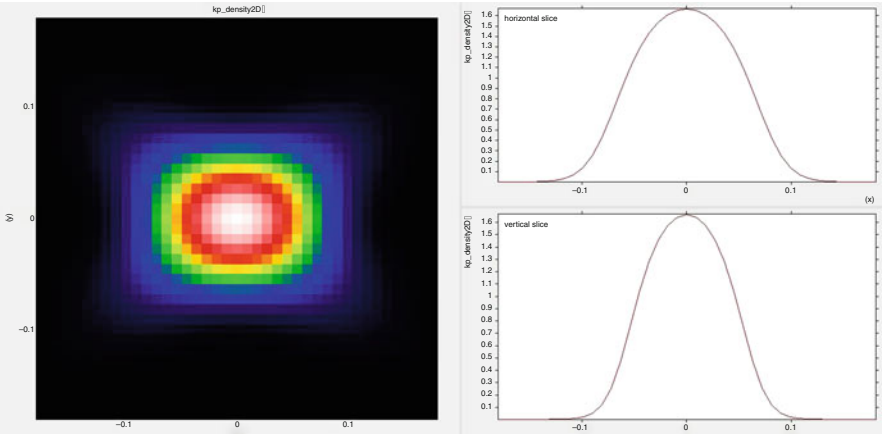


Fig. 6.22 $\mathbf{k}_\parallel$-resolved hole density distribution $p(k_x, k_y)$ for (110) oriented diamond. Here, k_x is related to [100] and k_y to $[01\bar{1}]$ direction. The confinement direction is parallel to [011]. The right part shows the horizontal and vertical slice through the center

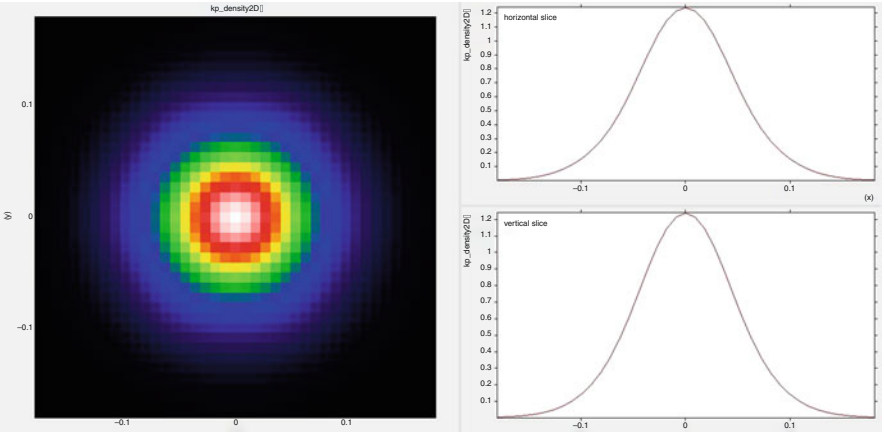


Fig. 6.23 $\mathbf{k}_\parallel$-resolved hole density distribution $p(k_x, k_y)$ for (111) oriented diamond. Here, k_x is related to $[11\bar{2}]$ and k_y to $[\bar{1}10]$ direction. The confinement direction is parallel to [111]

one along k_y (*red dotted lines*), thus justifying to expect an almost symmetric $\mathbf{k}_\parallel$-resolved density. This is not the case for (110) oriented diamond (Fig. 6.22). In this case the $\mathbf{k}_\parallel$-resolved hole density distribution is elongated along the k_x direction indicating that it is more favorable to occupy states along k_x rather than along k_y. To understand this preference it is necessary to analyze the energy dispersion of the occupied states. Figure 6.25 reveals that the energy dispersion of the ground state h_1 (*blue dash-dotted line*) has less curvature along the k_x direction rather than the k_y direction. Consequently, this favors the occupation of the states along the k_x direction for increasing energy.

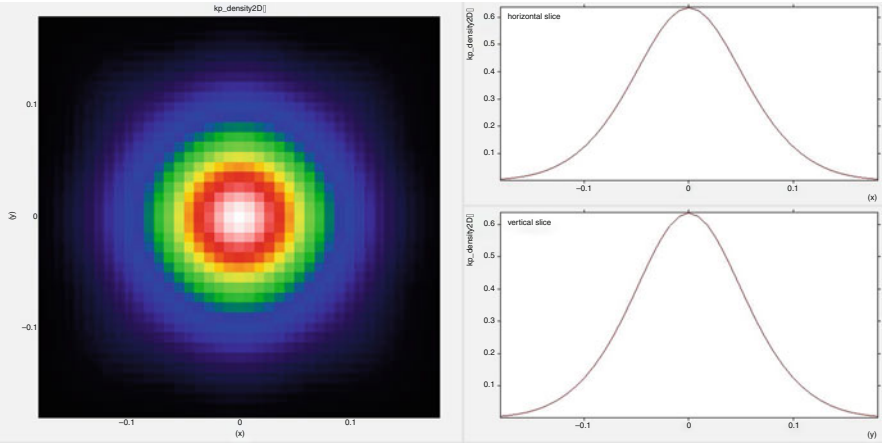


Fig. 6.24 $\mathbf{k}_\parallel$-resolved hole density distribution $p(k_x, k_y)$ for (100) oriented diamond Here, k_x is related to [100] and k_y to [010] direction. The confinement direction is parallel to [001]

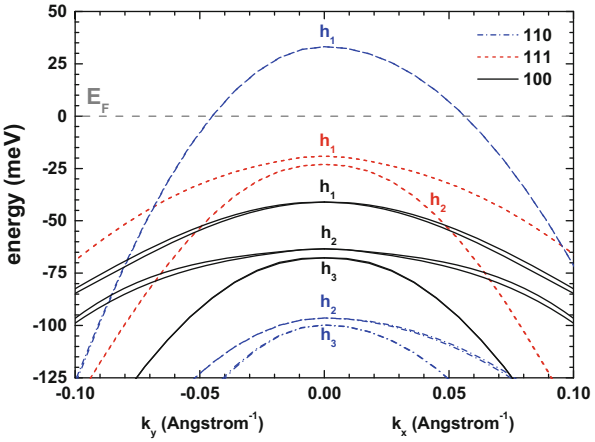


Fig. 6.25 Dispersion of the highest hole energy levels for (100), (110) and (111) oriented diamond. Only for (110) oriented diamond (*blue dash-dotted lines*) the ground state lies above the Fermi level. The k_x and k_y directions are with respect to the rotated coordinate system and thus are different for each orientation. (100) oriented diamond (*black solid lines*) is symmetric with respect to k_x and k_y whereas (111) diamond (*red dashed lines*) shows a slight asymmetry. (110) diamond shows a very pronounced asymmetry

Finally, we show the calculated energy dispersions of the highest hole eigenstates for each substrate orientation in Fig. 6.25. The energy scale is the same as for Fig. 6.21 which shows the eigenstates at $\mathbf{k}_\parallel = (k_x, k_y) = 0$. Only for (110) oriented diamond the ground state h_1 lies above the Fermi level E_F. The hole energy dispersion for (100) oriented diamond is symmetric with respect to k_x and k_y whereas for (111) diamond a slight asymmetry is present. (110) diamond shows

a very pronounced asymmetry as already discussed. Also the energies at $\mathbf{k}_\parallel = 0$ and their contributions to the density have already been discussed when analyzing the square of the probability amplitudes (Fig. 6.21). At $\mathbf{k}_\parallel = 0$ each state is twofold spin degenerate. Thus we labeled the uppermost 'three' states with h_1, h_2 and h_3 although there are actually six states. The spin-splitting at nonzero $\mathbf{k}_\parallel$ is due to structural inversion asymmetry (SIA). The triangular confinement potential induces SIA in contrast to bulk diamond which has a center of inversion, i.e. no bulk inversion asymmetry (BIA) and no SIA. Only the spin-splitting of the states h_1 and h_2 for (100) diamond is noticeable in this figure, in all other cases it can hardly be recognized.

Our calculations reveal a lot of information in terms of energies, energy dispersions, spatial extension and character of wave functions for each $\mathbf{k}_\parallel$ vector, as well as information on subband densities or even $\mathbf{k}_\parallel$-resolved densities for each substrate orientation. This information is useful in understanding and optimizing device designs, although we emphasize that reliable band structure parameters are needed as input. Our analysis is extremely useful when transitions between these hole subbands are involved, e.g. in optical absorption experiments, as different orientations show a significantly different energy spectrum, e.g. due to selection rules. However, for experiments where mainly the density is involved, e.g. in capacitance–voltage measurements, only the total density is relevant. Thus the contribution of the individual subband densities to the total density is somehow irrelevant, i.e. the precise spectrum of the energy levels is not as important as for optical absorption experiments. Depending on the orientation either one or several subbands are occupied but this is not vital as here merely the total density is relevant.

Neglecting any additional surface effects like surface relaxation or surface reconstruction, our calculations demonstrate that a (110) sample has a higher 2DHG sheet density than a (111) sample, and that the latter has a higher sheet density than a (100) sample with respect to the same boundary condition (negative interface charge). This is equivalent to saying that the change in 2DHG sheet density of a (110) sample is larger with respect to potential changes at the surface than for a (111) or for a (100) sample, i.e. a (110) sample is more sensitive than (111) and (100) samples, and consequently better suited for sensor devices.

References

1. T. Andlauer, Discretization of multiband-**k·p**-Schrödinger equations for multidimensional semiconductor nanostructures, Master thesis, Walter Schottky Institut and Physics Department, Technische Universität München (2004)
2. T. Andlauer, Optoelectronic and spin-related properties of semiconductor nanostructures in magnetic fields. Selected Topics of Semiconductor Physics and Technology **105**, Verein zur Förderung des Walter Schottky Instituts der Technischen Universität München e.V., München (2009)
3. T. Andlauer, P. Vogl, Full-band envelope-function approach for type-II broken-gap superlattices. Phys. Rev. B **80**, 035304 (2009)

4. A.D. Andreev, E.P. O'Reilly, Theory of the electronic structure of GaN/AlN hexagonal quantum dots. Phys. Rev. B **62**, 15851–15870 (2000)
5. T.P. Bahder, Eight-band **k·p** model of strained zinc-blende crystals. Phys. Rev. B **41**, 11992–12001 (1990)
6. B.A. Bernevig, T.L. Hughes, S.-C. Zhang, Quantum Spin Hall Effect and Topological Phase Transitions in HgTe Quantum Wells. Science **314**, 1757–1761 (2006)
7. G.L. Bir, G.E. Pikus, *Symmetry and strain-induced effects in semiconductors* (John Wiley & Sons, New York, 1974)
8. S. Birner, T. Zibold, T. Andlauer, T. Kubis, M. Sabathil, A. Trellakis, P. Vogl, nextnano: General Purpose 3-D Simulations. IEEE Trans. Electron Devices **54**, 2137–2142 (2007)
9. S. Boyer-Richard, F. Raouafi, A. Bondi, L. Pédesseau, C. Katan, J.-M. Jancu, J. Even, 30-band **k·p** method for quantum semiconductor heterostructures. Appl. Phys. Lett. **98**, 251913 (2011)
10. C. Brüne, A. Roth, E.G. Novik, M. König, H. Buhmann, E.M. Hankiewicz, W. Hanke, J. Sinova, L.W. Molenkamp, Evidence for the ballistic intrinsic spin Hall effect in HgTe nanostructures. Nature Physics **6**, 448–454 (2010)
11. V.A. Burdov, Electron and Hole Spectra of Silicon Quantum Dots. J. Exp. Theor. Phys. **94**, 411–418 (2002)
12. M.G. Burt, The justification for applying the effective-mass approximation to microstructures. J. Phys.: Condens. Matter **4**, 6651–6690 (1992)
13. M.G. Burt, Fundamentals of envelope function theory for electronic states and photonic modes in nanostructures. J. Phys.: Condens. Matter **11**, R53–R83 (1999)
14. M. Cardona, F.H. Pollak, Energy-Band Structure of Germanium and Silicon: The **k·p** Method. Phys. Rev. **142**, 530–543 (1966)
15. X. Cartoixà, Theoretical Methods for Spintronics in Semiconductors with Applications, Ph.D. dissertation, California Institute of Technology, Pasadena, California (2003)
16. C.Y.-P. Chao, S.L. Chuang, Spin-orbit-coupling effects on the valence-band structure of strained semiconductor quantum wells. Phys. Rev. B **46**, 4110–4122 (1992)
17. S.L. Chuang, C.S. Chang, **k·p** method for strained wurtzite semiconductors. Phys. Rev. B **54**, 2491–2504 (1996)
18. M. Dankerl, A. Lippert, S. Birner, E.U. Stützel, M. Stutzmann, J.A. Garrido, Hydrophobic interaction and charge accumulation at the diamond/electrolyte interface. Phys. Rev. Lett. **106**, 196103 (2011)
19. G. Dresselhaus, A.F. Kip, C. Kittel, Cyclotron Resonance of Electrons and Holes in Silicon and Germanium Crystals. Phys. Rev. **98**, 368–384 (1955)
20. E.T. Edmonds, C.I. Pakes, L. Ley, Self-consistent solution of the Schrödinger–Poisson equations for hydrogen-terminated diamond. Phys. Rev. B **81**, 085314 (2010)
21. H. Ehrenreich, A.W. Overhauser, Scattering of Holes by Phonons in Germanium. Phys. Rev. **104**, 331–342 (1956)
22. T. Eissfeller, private communication (2011)
23. T. Eissfeller, P. Vogl, Real-space multiband envelope-function approach without spurious solutions. Phys. Rev. B **84**, 195122 (2011)
24. R. Eppenga, M.F.H. Schuurmans, S. Colak, New **k·p** theory for $GaAs/Ga_{1-x}Al_xAs$-type quantum wells. Phys. Rev. B **36**, 1554–1564 (1987)
25. V.A. Fonoberov, A.A. Balandin, Excitonic properties of strained wurtzite and zinc-blende $GaN/Al_xGa_{1-x}N$ quantum dots. J. Appl. Phys. **94**, 7178–7186 (2003)
26. B.A. Foreman, Effective-mass Hamiltonian and boundary conditions for the valence bands of semiconductor microstructures. Phys. Rev. B **48**, 4964–4967 (1993)
27. B.A. Foreman, Elimination of spurious solutions from eight-band **k·p** theory. Phys. Rev. B **56**, R12748–R12751 (1997)
28. E. Gheeraert, S. Koizumi, T. Teraji, H. Kanda, Electronic States of Boron and Phosphorus in Diamond. Phys. Status Solidi A **174**, 39–51 (1999)
29. C.H. Grein, P.M. Young, M.E. Flatté, H. Ehrenreich, Long wavelength InAs/InGaSb infrared detectors: Optimization of carrier lifetimes. J. Appl. Phys. **78**, 7143–7152 (1995)

30. S. Hackenbuchner, Elektronische Struktur von Halbleiter-Nanobauelementen im thermodynamischen Nichtgleichgewicht. Selected Topics of Semiconductor Physics and Technology **48**, Verein zur Förderung des Walter Schottky Instituts der Technischen Universität München e.V., München (2002)
31. J.C. Hensel, G. Feher, Cyclotron Resonance Experiments in Uniaxially Stressed Silicon: Valence Band Inverse Mass Parameters and Deformation Potentials. Phys. Rev. **129**, 1041–1062 (1963)
32. J.M. Hinckley, J. Singh, Hole transport theory in pseudomorphic $Si_{1-x}Ge_x$ alloys grown on Si(001) substrates. Phys. Rev. B **41**, 2912–2926 (1990)
33. J.-M. Jancu, R. Scholz, F. Beltram, F. Bassani, Empirical $spds^*$ tight-binding calculation for cubic semiconductors: General method and material parameters. Phys. Rev. B **57**, 6493–6507 (1998)
34. P. Lawaetz, Valence-Band Parameters in Cubic Semiconductors. Phys. Rev. B **4**, 3460–3467 (1971)
35. R.B. Lehoucq, D.C. Sorensen, C. Yang, *ARPACK Users' Guide: Solution of Large-Scale Eigenvalue Problems with Implicitly Restarted Arnoldi Methods* (SIAM Publications, Philadelphia, 1998)
36. L.C. Lew Yan Voon, M. Willatzen, *The* **k·p** *method – Electronic Properties of Semiconductors* (Springer, Berlin, 2009)
37. J.P. Loehr, Parameter consistency in multienergetic **k·p** models. Phys. Rev. B **52**, 2374–2380 (1995)
38. J. Los, A. Fasolino, A. Catellani, Generalization of the **k·p** approach for strained layered semiconductor structures grown on high-index-planes. Phys. Rev. B **53**, 4630–4648 (1996)
39. S. Luryi, Quantum capacitance devices: General Theory. Appl. Phys. Lett. **52**, 501–503 (1988)
40. J.M. Luttinger, Quantum Theory of Cyclotron Resonance in Semiconductors: General Theory. Phys. Rev. **102**, 1030–1039 (1956)
41. J.M. Luttinger, W. Kohn, Motion of Electrons and Holes in Perturbed Periodic Fields. Phys. Rev. **97**, 869–883 (1955)
42. F. Mireles, S.E. Ulloa, Ordered Hamiltonian and matching conditions for heterojunctions with wurtzite symmetry: $GaN/Al_xGa_{1-x}N$ quantum wells. Phys. Rev. B **60**, 13659–13667 (1999)
43. next**nano**: The next**nano** software can be obtained from http://www.nextnano.com (2011)
44. E.G. Novik, A. Pfeuffer-Jeschke, T. Jungwirth, V. Latussek, C.R. Becker, G. Landwehr, H. Buhmann, L.W. Molenkamp, Band structure of semimagnetic $Hg_{1-y}Mn_yTe$ quantum wells. Phys. Rev. B **72**, 035321 (2005)
45. P. Pfeffer, W. Zawadzki, Five-level **k·p** model for the conduction and valence bands of GaAs and InP. Phys. Rev. B **53**, 12813–12828 (1996)
46. C.R. Pidgeon, R.N. Brown, Interband Magneto-Absorption and Faraday Rotation in InSb. Phys. Rev. **146**, 575–583 (1966)
47. C.J. Rauch, Millimeter Cyclotron Resonance Experiments in Diamond. Phys. Rev. Lett. **7**, 83–84 (1961)
48. C.J. Rauch, Millimetre cyclotron resonance in diamond. Proceedings of the International Conference on Semiconductor Physics, Exeter (The Institute of Physics and the Physical Society, London), 276–280 (1962)
49. L. Reggiani, D. Waechter, S. Zukotynski, Hall-coefficient factor and inverse valence-band parameters of holes in natural diamond. Phys. Rev. B **128**, 3550–3555 (1983)
50. S. Richard, F. Aniel, G. Fishman, Energy-band structure of Ge, Si, and GaAs: A thirty-band **k·p** method. Phys. Rev. B **70**, 235204 (2004)
51. A. Trellakis, T. Zibold, S. Andlauer, S. Birner, R.K. Smith, R. Morschl, P. Vogl, The 3D nanometer device project *nextnano*: Concepts, methods, results. J. Comput. Electron. **5**, 285–289 (2006)
52. C.G. Van de Walle, Band lineups and deformation potentials in the model-solid theory. Phys. Rev. B **39**, 1871–1883 (1989)
53. R.G. Veprek, S. Steiger, B. Witzigmann, Ellipticity and the spurious solution problem of **k·p** envelope functions. Phys. Rev. B **76**, 165320 (2007)

54. I. Vurgaftman, J.R. Meyer, L.R. Ram-Mohan, Band parameters for III-V compound semiconductors and their alloys. J. Appl. Phys. **89**, 5815–5875 (2001)
55. M. Willatzen, M. Cardona, N.E. Christensen, Linear muffin-tin-orbital and **k·p** calculations of effective masses and band structure of semiconducting diamond. Phys. Rev. B **50**, 18054–18059 (1994)
56. P.Y. Yu, M. Cardona, *Fundamentals of Semiconductors: Physics and Materials Properties* (Springer, Berlin, 1999)
57. A. Zakharova, S.T. Yen, K.A. Chao, Hybridization of electron, light-hole, and heavy-hole states in InAs/GaSb quantum wells. Phys. Rev. B **64**, 235332 (2001)
58. T. Zibold, Semiconductor based quantum information devices: Theory and simulations. Selected Topics of Semiconductor Physics and Technology **87**, Verein zur Förderung des Walter Schottky Instituts der Technischen Universität München e.V., München (2007)

Part IV
Advanced Mathematical Topics

Chapter 7
Transient Simulation of k·p-Schrödinger Systems Using Discrete Transparent Boundary Conditions

Andrea Zisowsky, Anton Arnold, Matthias Ehrhardt, and Thomas Koprucki

Abstract This chapter deals with the derivation and analysis of discrete transparent boundary conditions (TBCs) for transient systems of *Schrödinger-type equations* in one space dimension. These systems occur i.e. in the physics of *layered semiconductor devices* as the so called **k·p**-Schrödinger equations, which are a well established tool for *band structure calculations*.

The new TBCs are constructed directly for the chosen finite difference scheme, in order to ensure the stability of the underlying scheme and to completely avoid any numerical reflections. The discrete TBCs are constructed using the solution of the exterior problem with Laplace and $\mathscr{Z}$-transformation, respectively.

These discrete TBCs can easily be obtained by an inverse $\mathscr{Z}$-transformation based on FFT, but these exact discrete TBCs are non-local in time and thus very costly. Hence, as a remedy, we present approximate discrete TBCs, that allow a fast calculation of the boundary terms using a sum-of-exponentials approach.

A. Zisowsky (✉)
Institut für Mathematik, Technische Universität Berlin, Strasse des 17. Juni 136, 10623 Berlin, Germany
e-mail: Andrea.Zisowsky@web.de

A. Arnold
Institut für Analysis und Scientific Computing, Technische Universität Wien, Wiedner Hauptstr. 8, 1040 Wien, Austria
e-mail: anton.arnold@tuwien.ac.at

M. Ehrhardt
Lehrstuhl für Angewandte Mathematik und Numerische Analysis, Fachbereich C Mathematik und Naturwissenschaften, Bergische Universität Wuppertal, Gaußstr. 20, 42119 Wuppertal, Germany
e-mail: ehrhardt@math.uni-wuppertal.de

T. Koprucki
Forschungsgruppe "Partielle Differentialgleichungen", Weierstrass-Institut für Angewandte Analysis und Stochastik, Mohrenstrasse 39, 10117 Berlin, Germany
e-mail: koprucki@wias-berlin.de

M. Ehrhardt and T. Koprucki (eds.), *Multi-Band Effective Mass Approximations*, Lecture Notes in Computational Science and Engineering 94,
DOI 10.1007/978-3-319-01427-2_7,

7.1 Introduction

Many modern quantum-electronic semiconductor devices such as *resonant tunneling diodes* (RTD) [23, Chapter 14] or opto-electronic devices such as quantum-cascade lasers [24] and multi-quantum-well electro-absorption modulators [15] are based on the tunneling process of carriers through barrier structures. Typically these kind of barrier structures are *layered semiconductor heterostructures* [15, 23, 24] with a barrier thickness of a few nanometer. The transient simulation of wave packets tunneling through such nano-scale semiconductor heterostructures plays a key role in the understanding of such transport processes [22, 25, 28].

In this respect, transient simulations can be used to estimate charging and escape times [25, 28], tunneling times [22], or carrier life times [15, 27]. For the time-dependent simulation of a tunneling process usually a scalar Schrödinger equation defined by BenDaniel-Duke-type Hamiltonians [4, Chapter 3] is used [22, 27, 28]. Here the electronic band structure is approximated by a single parabolic band. These *parabolic single-band approximations* are in good agreement with the real band structure in the vicinity of the minima of the conduction bands, which is the part of the band structure that is usually occupied by the electrons. For the treatment of the holes, occupying the maxima of the valence bands, the accuracy of parabolic single-band models is often not sufficient since the valence bands possess a much more complex band structure [4, 6, 11, 12, 14, 15, 22].

However, the electronic states of the holes can be approximated appropriately by multi-band states which satisfy a so-called **k·p**-Schrödinger equation. The time-dependent **k·p**-Schrödinger equation describes the time evolution of the multi-band electronic state and can be regarded as a *linear coupled system of scalar Schrödinger equations*. The evolution is governed by the **k·p**-Schrödinger operator which is an extension to the single-band models and describes a system of bands of the band structure, e.g. the four topmost valence bands [4, 6, 12, 14].

There exists a couple of such multi-band **k·p**-models [21] including also combined models for conduction and valence bands. The later also allow for a nonparabolic approximation of the conduction bands. Such **k·p**-models can be used for devices where the parabolic conduction band approximation is not sufficient. For unipolar devices where by crossing a barrier a conduction-band to valence-band transition is possible such as *resonant interband tunneling diodes* (RITD) or for bipolar devices where additionally the hole tunneling processes are important such as for multi-quantum well electro-absorption modulators multi-band modeling is necessary. In this cases the numerical solution of the transient **k·p**-Schrödinger equation can be used to understand and to determine the tunneling properties of the corresponding semiconductor heterostructures by studying the time evolution of the multi-band electronic state.

In this chapter we discuss the appropriate numerical treatment of a transient system of **k·p**-Schrödinger-type. We notice that such type of Schrödinger systems also arise as "parabolic systems" in electromagnetic wave propagation. Artificial boundary conditions (BCs) have to be imposed to restrict the unbounded domain,

on which the partial differential equation (PDE) is originally defined, to a finite computational domain. Such BCs are called *transparent boundary conditions* (TBCs), if the solution on the whole space (restricted to the computational domain) is equal to the solution with the artificial BCs. The artificial boundary splits the problem into three parts: the interesting interior problem and a left and right exterior problem. For constant coefficients the exterior problems can be solved explicitly by the Laplace method in the continuous and $\mathscr{Z}$-transformation in the discrete case.

Claiming (spatial) C^1-continuity of the solution at the artificial boundaries yields the TBC as a Dirichlet-to-Neumann (DtN) map [2, 16]. An ad-hoc discretisation of these continuous TBC can destroy the stability of the employed numerical scheme for the PDE and induce numerical reflections [1]. To avoid this, we derive *discrete TBCs* for the fully discretised PDE. The procedure is analogous to the continuous case and uses the $\mathscr{Z}$-transformation. The inverse Laplace/$\mathscr{Z}$-transformation yields a convolution in time. Hence, the perfectly *exact* BC is non-local in time and therefore very costly for long-time simulations. While we focus her on a finite difference discretization, a similar approach based on a finite element discretization for scalar Schrödinger equation in 1D was proposed recently by A. Zlotnik and I. Zlotnik [32,33]. To reduce the numerical effort, we introduce approximate discrete TBCs. Since the inverse $\mathscr{Z}$-transformation must be accomplished numerically for Schrödinger–type systems, an additional small numerical error is induced.

This chapter is organised as follows: In Sect. 7.2 we introduce the system of **k·p**-Schrödinger equations and present as an example a quantum well structure with a double barrier that will be considered throughout this work. Next we derive in Sect. 7.3 the analytic TBC and afterwards its discrete variant in Sect. 7.4. Here, we also scrutinise the coefficients of the discrete convolution and explain our strategy to compute the coefficients by a numerical inverse $\mathscr{Z}$-transformation. In Sect. 7.5 we approximate the coefficients by a sum-of-exponentials ansatz and propose a fast evaluation of the approximate discrete TBC. Finally, in Sect. 7.6 we present the results of numerical simulations for a quantum heterostructure.

7.2 Transient k·p-Schrödinger Systems

In this section we consider **k·p**-Schrödinger systems that are well established models for *band structure calculations* [13] for one-dimensional semiconductor nanostructures. They are layered *heterostructures* consisting of layers of different semiconductor materials with abrupt, planar heterojunction interfaces between the layers [23]. Typical examples are semiconductor quantum wells and double-barrier structures or *resonant interband tunneling diodes* (RITDs) [4, 14, 23, 26].

Figure 7.1 depicts the typical valence and conduction band profile of a quantum well structure. Usually one is interested in the computation of the bound eigenstates (with eigenvalue smaller (large) than E_c (E_v) in the barrier material) in the quantum

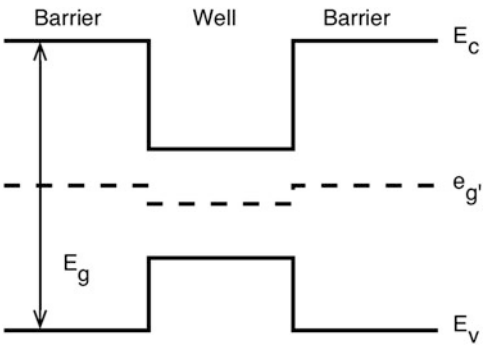


Fig. 7.1 Band edge profile for a quantum well structure. The electrons and holes are confined between the barriers in the well region

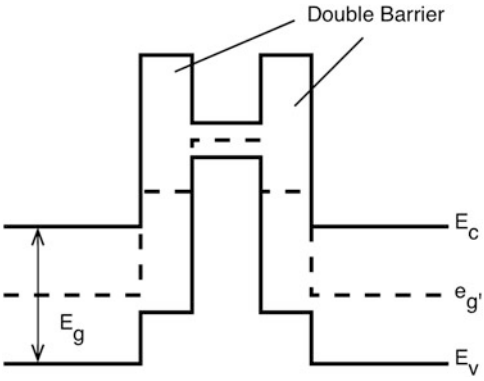


Fig. 7.2 Band edge profile for a RITD structure. By crossing from barrier to well region the electrons âĂŹbecomeâĂŹ a hole which is attracted by its potential well given by the valence band edge profile. This increases the density between the barriers and improves the performance of the tunneling device

well. Here E_g denotes the bandgap and $e_{g'}$ is the middle of the bandgap that can be assumed to be constant on each material and thus as piecewise constant functions.

The typical valence and conduction band profile for a RITD is shown in Fig. 7.2. For the RITD one is usually interested in the computation of the transmission and reflection amplitudes in dependence on the wave vector of the incoming wave together with the corresponding wave functions.

The **k**·**p**-method [19] in combination with the envelope function approximation [4, 4, 9–11, 14, 23] is a frequently used approach for the modeling of the near-band-edge electronic states in semiconductor nanostructures. Within this approach the electronic state $\Psi(\mathbf{r})$ is approximated in terms of d bands

$$\Psi_{\mathbf{k}_\parallel}(\mathbf{r}) = \exp(i\,\mathbf{k}_\parallel \cdot \mathbf{r}_\parallel) \sum_{\nu=1}^{d} \varphi_\nu(x;\mathbf{k}_\parallel) u_{\nu,\mathbf{k}=\mathbf{0}}(\mathbf{r}) \qquad \text{with} \quad \mathbf{r} = (\mathbf{r}_\parallel, x) \in \mathbb{R}^3 .$$

The index $\parallel$ indicates in-plane vectors and x denotes the growth direction of the semiconductor layers. $\mathbf{k}_\parallel = (k_1, k_2) \in \mathbb{R}^2$ is the reduced wave vector, which will be fixed for each simulation model. $u_{\nu,\mathbf{k}=\mathbf{0}}(\mathbf{r})$ are lattice periodic, zone-center Bloch functions varying on the atomic scale and $\varphi_\nu(x;\mathbf{k}_\parallel)$ are the corresponding envelope functions describing the variation of the wave function on the (larger) nanoscale.

The vector of the envelope functions $\boldsymbol{\varphi} = (\varphi_1, \cdots, \varphi_d)$ with $\boldsymbol{\varphi}(x,t) \in \mathbb{C}^d$ fulfill the one-dimensional **k·p**-*Schrödinger equation*

$$i\,\frac{\partial}{\partial t}\boldsymbol{\varphi} = \mathbf{H}\Big(\mathbf{k}_{\|}, -i\,\frac{\partial}{\partial x}\Big)\boldsymbol{\varphi}\,.$$

There is a hierarchy of **k·p**-models [21] including 4-band, 6-band and 8-band Hamiltonians. Depending on the model Hamiltonian, effects such as quantum confinement, band-mixing, spin-orbit interaction and mechanical strain can be treated consistently. The basic stage in this model hierarchy is the 4×4 Luttinger-Kohn-Hamiltonian [20] which describes the band-mixing between the heavy holes and the light holes [4, 12–14].

In our notation we will follow Bandelow et al. [5] who performed a rigorous analysis of spectral properties for the spatially one dimensional **k·p**-Schrödinger operators. The considered system reads as follows

$$\begin{aligned}
\mathrm{i}\frac{\partial}{\partial t}\boldsymbol{\varphi} &= -\frac{\partial}{\partial x}\Big(\mathbf{N}(x)\frac{\partial}{\partial x}\boldsymbol{\varphi}\Big) + \mathbf{M}_0(x)\frac{\partial}{\partial x}\boldsymbol{\varphi} - \frac{\partial}{\partial x}\big(\mathbf{M}_0^{\mathrm{H}}(x)\boldsymbol{\varphi}\big) \\
&\quad + k_1\Big(\mathbf{M}_1(x)\frac{\partial}{\partial x}\boldsymbol{\varphi} - \frac{\partial}{\partial x}(\mathbf{M}_1^{\mathrm{H}}(x)\boldsymbol{\varphi})\Big) + k_2\Big(\mathbf{M}_2(x)\frac{\partial}{\partial x}\boldsymbol{\varphi} - \frac{\partial}{\partial x}(\mathbf{M}_2^{\mathrm{H}}(x)\boldsymbol{\varphi})\Big) \\
&\quad + k_1\mathbf{U}_1(x)\boldsymbol{\varphi} + k_2\mathbf{U}_2(x)\boldsymbol{\varphi} + k_1^2\mathbf{U}_{11}(x)\boldsymbol{\varphi} + k_1k_2(\mathbf{U}_{12}(x) + \mathbf{U}_{21}(x))\boldsymbol{\varphi} \\
&\quad + k_2^2\mathbf{U}_{22}(x)\boldsymbol{\varphi} + \mathbf{v}(x)\boldsymbol{\varphi} + \mathbf{e}(x)\boldsymbol{\varphi}, \qquad x \in \mathbb{R}, t > 0, \quad k_1, k_2 \in \mathbb{R},
\end{aligned} \tag{7.1}$$

where $\boldsymbol{\varphi}(x,t) \in \mathbb{C}^d$, the mass matrix $\mathbf{N}$ and $\mathbf{e}$ are real diagonal $d \times d$-matrices. $\mathbf{U}_i$, $\mathbf{U}_{ij}$ and $\mathbf{v}$ are Hermitian $d \times d$-matrices. The $d \times d$-matrices $\mathbf{M}_0(x)$, $\mathbf{M}_1(x)$ and $\mathbf{M}_2(x)$ are skew-Hermitian. In the sequel we abbreviate

$$\mathbf{M}_S(x) := \mathbf{M}_0(x) + k_1\mathbf{M}_1(x) + k_2\mathbf{M}_2(x), \tag{7.2a}$$

$$\begin{aligned}
\mathbf{V}(x) := k_1\mathbf{U}_1(x) + k_2\mathbf{U}_2(x) + k_1^2\mathbf{U}_{11}(x) + k_2^2\mathbf{U}_{22}(x) \\
+ k_1k_2(\mathbf{U}_{12}(x) + \mathbf{U}_{21}(x)) + \mathbf{v}(x) + \mathbf{e}(x).
\end{aligned} \tag{7.2b}$$

Then $\mathbf{M}_S(x)$ is skew-Hermitian, $\mathbf{V}(x)$ is Hermitian and (7.1) reads

$$\mathrm{i}\frac{\partial}{\partial t}\boldsymbol{\varphi} = -\frac{\partial}{\partial x}\Big(\mathbf{N}(x)\frac{\partial}{\partial x}\boldsymbol{\varphi}\Big) + \mathbf{M}_S(x)\frac{\partial}{\partial x}\boldsymbol{\varphi} - \frac{\partial}{\partial x}\big(\mathbf{M}_S^{\mathrm{H}}(x)\boldsymbol{\varphi}\big) + \mathbf{V}(x)\boldsymbol{\varphi}, \tag{7.3}$$

$x \in \mathbb{R}$, $t > 0$. The real diagonal matrix $\mathbf{e}(x)$ describes the variation of the band-edges. The band-mixing due to the **k·p**-interaction of the first and second order are described by the terms containing the matrices $\mathbf{M}_\alpha$, $\mathbf{U}_\alpha$, $\alpha = 0, 1, 2$, and $\mathbf{U}_{\alpha\beta}$,

$\alpha, \beta = 1, 2$, respectively. The potential $\mathbf{v}$ can cover couplings induced by the spin-orbit interaction or by mechanical strain. When neglecting all non-diagonal coupling terms, the system would reduce to an uncoupled system of scalar Schrödinger equations corresponding to the case of uncoupled parabolic bands. In this sense the couplings can be interpreted as correction terms to the parabolic band structure approximation.

An important property of the system (7.3) is the *conservation of mass*, i.e. $||\boldsymbol{\varphi}||_{L^2}^2$ is constant in time. This property can be easily verified: we multiply (7.3) with $\boldsymbol{\varphi}^{\mathrm{H}}$ from the left and integrate by parts:

$$\begin{aligned}
\frac{\partial}{\partial t}||\boldsymbol{\varphi}||_{L^2}^2 &= \frac{\partial}{\partial t}\int_{\mathbb{R}} \boldsymbol{\varphi}^{\mathrm{H}}\boldsymbol{\varphi}\, dx = 2\ \mathrm{Im}\int_{\mathbb{R}} \mathrm{i}\boldsymbol{\varphi}^{\mathrm{H}}\boldsymbol{\varphi}_t \\
&= 2\ \mathrm{Im}\left(-\int_{\mathbb{R}} \boldsymbol{\varphi}^{\mathrm{H}}\frac{\partial}{\partial x}(\mathbf{m}\frac{\partial}{\partial x}\boldsymbol{\varphi})\, dx + \int_{\mathbb{R}} \boldsymbol{\varphi}^{\mathrm{H}}\mathbf{V}\boldsymbol{\varphi}\, dx\right. \\
&\quad \left. + \int_{\mathbb{R}} \boldsymbol{\varphi}^{\mathrm{H}}\mathbf{M}_S\frac{\partial}{\partial x}\boldsymbol{\varphi}\, dx - \int_{\mathbb{R}} \boldsymbol{\varphi}^{\mathrm{H}}\frac{\partial}{\partial x}(\mathbf{M}_S^{\mathrm{H}}\boldsymbol{\varphi})dx\right) \\
&= 2\ \mathrm{Im}\Big(\int_{\mathbb{R}} \boldsymbol{\varphi}_x^{\mathrm{H}}\mathbf{m}\boldsymbol{\varphi}_x\, dx + \int_{\mathbb{R}} \boldsymbol{\varphi}^{\mathrm{H}}\mathbf{V}\boldsymbol{\varphi}\, dx + \int_{\mathbb{R}} \underbrace{\boldsymbol{\varphi}^{\mathrm{H}}\mathbf{M}\boldsymbol{\varphi}_x + \boldsymbol{\varphi}_x^{\mathrm{H}}\mathbf{M}^{\mathrm{H}}\boldsymbol{\varphi}}_{\in\mathbb{R}}\, dx\Big) \\
&= 0.
\end{aligned}$$

The last equality follows from the fact that $\mathbf{V}$ and $\mathbf{N}$ are Hermitian and thus the imaginary part of the quadratic forms vanishes. The other term is of the form $y + y^{\mathrm{H}}$ and thus obviously real.

We now briefly review an illustrative example from [29, 30] that we will use throughout this chapter for the numerical results.

Example 7.1 (Double-barrier stepped quantum-well structure [29, 30]). We consider the GaAs/AlGaAs *double-barrier stepped quantum-well structure* (DBSQW) introduced in [28]. The variation of the band-edges $e(x)$ is depicted in Fig. 7.3. For this kind of structure an analysis of the time evolution of wave packets tunneling through the structure has been performed using a scalar Schrödinger equation [28]. We consider the more accurate four-band Luttinger-Kohn-Hamiltonian [4, 12–14] modeling the band-mixing of heavy and light holes. In atomic units adapted to the light holes ($\hbar = 1$, $m_0/(\gamma_1 + 2\gamma_2) = 1$) the coefficient matrices for the corresponding 4×4 system of Schrödinger equations are given by $\mathbf{N} = 0.5 \cdot \mathrm{diag}(\gamma, 1, 1, \gamma)$, $\mathbf{M}_0 = \mathbf{0}$,

$$\mathbf{M}_1 = \frac{1}{2}\frac{\gamma_3}{\gamma_1 + 2\gamma_2}\sqrt{3}\mathrm{i}\begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \end{pmatrix}, \qquad \mathbf{M}_2 = \frac{1}{2}\frac{\gamma_3}{\gamma_1 + 2\gamma_2}\sqrt{3}\begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix},$$

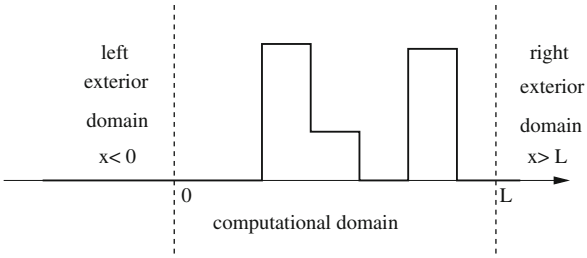


Fig. 7.3 Variation of the band-edge $e(x)$ for the GaAs/AlGaAs DBSQW structure

$$\mathbf{U}_{11,22} = \frac{1}{2}\frac{1}{\gamma_1 + 2\gamma_2}\begin{pmatrix} \gamma_1+\gamma_2 & 0 & \mp\sqrt{3}\gamma_2 & 0 \\ 0 & \gamma_1-\gamma_2 & 0 & \mp\sqrt{3}\gamma_2 \\ \mp\sqrt{3}\gamma_2 & 0 & \gamma_1-\gamma_2 & 0 \\ 0 & \mp\sqrt{3}\gamma_2 & 0 & \gamma_1+\gamma_2 \end{pmatrix},$$

$$\mathbf{U}_{12} + \mathbf{U}_{21} = \frac{\sqrt{3}\gamma_3 \mathrm{i}}{\gamma_1 + 2\gamma_2}\begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix},$$

and $\mathbf{U}_1 = \mathbf{U}_2 = \mathbf{0}$ with

$$\gamma = \frac{\gamma_1 - 2\gamma_2}{\gamma_1 + 2\gamma_2}.$$

The values of the band structure parameters for GaAs are given by

$$\gamma_1 = 6.85, \qquad \gamma_2 = 2.1, \qquad \gamma_3 = 2.9.$$

For the in-plane wave-vector $\mathbf{k}_\parallel$ we choose $k_1 = 2.3$, $k_2 = 0$. As initial condition we use a Gaussian wave packet

$$\boldsymbol{\varphi}(x,0) = (2\pi\sigma^2)^{\frac{1}{4}} \exp\left(\mathrm{i}k_r x - \frac{(x-x_0)^2}{\sigma^2}\right) \cdot \boldsymbol{\zeta}\,, \tag{7.4}$$

where $\boldsymbol{\zeta} \in \mathbb{C}^d$ is a linear combination of eigenmodes calculated via the dispersion relation of (7.3) and $\sigma = 3$, $x_0 = -2\sigma$ and $k_r = \sqrt{6.99}$. The band-edge profile $\mathbf{e}(x) = e(x)\mathbf{I}$ (with the identity matrix $\mathbf{I}$) of the DBSQW is taken from [28] and defined by

$$e(x) = \begin{cases} 0, & x \le 0 \\ \frac{15}{2}, & 0 < x \le 0.5 \\ \frac{3}{2}, & 0.5 < x \le 1 \\ 0, & 1 < x \le 1.5 \\ \frac{15}{2}, & 1.5 < x \le 2 \\ 0, & 2 < x \end{cases}. \tag{7.5}$$

The computational domain is now defined such that it contains the *significant* part of the initial data and the x-dependent part of the band-edge potential (cf. Fig. 7.3). For a strategy to soften this restriction on the initial data we refer the reader to [18]. Next we introduce in Sect. 7.3 TBCs at the left and right boundary $x = 0$ and $x = L$.

7.3 The Transparent Boundary Conditions

In this section we derive the TBCs for the **k**·**p**-Schrödinger equation (7.1). In the scalar case (classical Schrödinger equation of quantum mechanics), the Laplace-transformed equation in the exterior domain can be solved explicitly. Afterwards the solution is inverse transformed, thus yielding the analytic boundary condition, cf. [2]. Here, for a system the inverse transform can not be calculated explicitly. Nevertheless, we will present the derivation of the Laplace transformed TBC.

For the derivation we proceed as follows: we consider the Schrödinger equation in the exterior domain. A Laplace transformation with respect to time yields a system of ordinary differential equations (ODEs), that can be reduced to first order. Then the solution of this system can be given in terms of its eigenvalues and eigenvectors. Next, we prove, that half of the eigenvalues have positive real parts and thus yield solutions increasing for $x \to \infty$; the other half has negative real parts, yielding decreasing solutions. Requiring that the part of the increasing solutions in the right (and the decreasing solutions in the left) exterior domain vanish, yields the TBC.

We consider Eq. (7.1) in the bounded domain $[0, L]$ supplied with TBCs at $x = 0$ and $x = L$. Since the derivation for the left and right TBC is analogous, we focus on the right boundary at $x = L$. The TBC at $x = L$ is constructed by considering (7.1) with constant coefficients for $x > L$, the so called *right exterior problem*

$$\mathrm{i}\boldsymbol{\varphi}_t = -\mathbf{N}\boldsymbol{\varphi}_{xx} + i\mathbf{M}\boldsymbol{\varphi}_x + \mathbf{V}\boldsymbol{\varphi}, \qquad x > L,\ t > 0, \tag{7.6}$$

where $\mathbf{M} = \mathbf{M}^{\mathrm{H}}$, $\mathbf{V} = \mathbf{V}^{\mathrm{H}}$. $\mathbf{N}$ is diagonal, real and regular and given by

$$\begin{aligned} \mathbf{M} &= -\mathrm{i}\big(\mathbf{M}_0 - \mathbf{M}_0^{\mathrm{H}} + k_1(\mathbf{M}_1 - \mathbf{M}_1^{\mathrm{H}}) + k_2(\mathbf{M}_2 - \mathbf{M}_2^{\mathrm{H}})\big) \\ &= -\mathrm{i}(\mathbf{M}_S - \mathbf{M}_S^{\mathrm{H}}), \end{aligned} \tag{7.7a}$$

$$\mathbf{V} = k_1\mathbf{U}_1 + k_2\mathbf{U}_2 + k_1^2\mathbf{U}_{11} + k_2^2\mathbf{U}_{22} + k_1k_2(\mathbf{U}_{12} + \mathbf{U}_{21}) + \mathbf{v} + \mathbf{e}. \tag{7.7b}$$

If $\mathbf{M}_S$ is skew-Hermitian, then $\mathbf{M}_S - \mathbf{M}_S^{\mathrm{H}} = 2\mathbf{M}_S$ is also skew-Hermitian, thus $\mathbf{M} = -2\mathrm{i}\mathbf{M}_S$ is Hermitian. Analogously we define the *left exterior problem* for $x < 0$ and denote the occurring matrices with "$\tilde{}$".

To solve the right exterior problem for $x > L$ we apply the Laplace-transformation

$$\hat{\boldsymbol{\varphi}}(x,s) = \int_0^\infty e^{-st}\boldsymbol{\varphi}(x,t)\,dt, \quad s = \alpha + \mathrm{i}\xi, \quad \alpha > 0, \quad \xi \in \mathbb{R}, \tag{7.8}$$

to (7.6) and obtain the system of ODEs

$$\mathbf{N}\hat{\boldsymbol{\varphi}}_{xx} - \mathrm{i}\mathbf{M}\hat{\boldsymbol{\varphi}}_x = (\mathbf{V} - \mathrm{i}s\mathbf{I})\hat{\boldsymbol{\varphi}}, \qquad x > L. \tag{7.9}$$

This transformed right exterior problem (7.9) posses a *unique classical solution*:

Lemma 7.1 ([30, Lemma 3.1]). *Let* $\mathrm{Re}(s) > 0$. *Then, the boundary value problem* (7.9) *with the boundary data*

$$\hat{\boldsymbol{\varphi}}(x = L) = \hat{\boldsymbol{\Phi}} \in \mathbb{C}^d, \qquad \hat{\boldsymbol{\varphi}}(x = \infty) = 0 \tag{7.10}$$

has a unique (classical) solution.

It is exactly this linear mapping from the boundary data to the (derivative of the) solution $\hat{\boldsymbol{\varphi}}_L \mapsto \hat{\boldsymbol{\varphi}}_x(L)$, the so-called Dirichlet-to-Neumann (DtN) map, that represents the Laplace transformed TBC we are seeking for.

Next, to derive an explicit form of this TBC being accessible for numerical implementations, we define $\boldsymbol{\nu} = \hat{\boldsymbol{\varphi}}$ and $\boldsymbol{\eta} = \hat{\boldsymbol{\varphi}}_x$. Doing so, we reduce the order of the ODE and obtain a system of first order ODEs

$$\underbrace{\begin{pmatrix} \mathbf{M} & \mathrm{i}\mathbf{N} \\ -\mathrm{i}\mathbf{N} & \mathbf{0} \end{pmatrix}}_{\mathbf{A}} \begin{pmatrix} \boldsymbol{\nu}_x \\ \boldsymbol{\eta}_x \end{pmatrix} = \underbrace{\begin{pmatrix} \mathrm{i}\mathbf{V} + s\mathbf{I} & \mathbf{0} \\ \mathbf{0} & -\mathrm{i}\mathbf{N} \end{pmatrix}}_{\mathbf{B}} \begin{pmatrix} \boldsymbol{\nu} \\ \boldsymbol{\eta} \end{pmatrix}, \qquad x > L. \tag{7.11}$$

In [30, Lemma 3.4] we proved the regularity of the matrix $\mathbf{A}^{-1}\mathbf{B}$ for $\mathrm{Re}(s) > 0$ and thus we can rewrite (7.11) as

$$\begin{pmatrix} \boldsymbol{\nu}_x \\ \boldsymbol{\eta}_x \end{pmatrix} = \underbrace{\begin{pmatrix} \mathbf{0} & I \\ \mathbf{N}^{-1}(\mathbf{V} - \mathrm{i}s\mathbf{I}) & \mathrm{i}\mathbf{N}^{-1}\mathbf{M} \end{pmatrix}}_{\mathbf{A}^{-1}\mathbf{B}} \begin{pmatrix} \boldsymbol{\nu} \\ \boldsymbol{\eta} \end{pmatrix}, \qquad x > L. \tag{7.12}$$

In order to understand the structure of the solution we have to distinguish between increasing and decaying solutions of the system (7.12). For this purpose we proved using a continuity argument in [30] the *Splitting Theorem*:

Theorem 7.1 ([30, Theorem 3.2]). *The regular matrix* $\mathbf{A}^{-1}\mathbf{B}$ *has* d *eigenvalues with positive real part and* d *with negative real part.*

The next step is to transform $\mathbf{A}^{-1}\mathbf{B}$ in (7.12) to a Jordan form with

$$\mathbf{A}^{-1}\mathbf{B} = \mathbf{P}\mathbf{J}\mathbf{P}^{-1},$$

where $\mathbf{P}$ contains the left eigenvectors in columns. We sort the Jordan blocks in $\mathbf{J}$ with respect to an increasing real part of the corresponding eigenvalue. Thus $\mathbf{J}$ can be written as a block diagonal matrix $\mathbf{J} = \operatorname{diag}(\mathbf{J}_1, \mathbf{J}_2)$, where $\mathbf{J}_1$ holds all Jordan blocks to eigenvalues with negative real parts and $\mathbf{J}_2$ those with positive real parts. Due to the Splitting Theorem 7.1 $\mathbf{J}_1$ and $\mathbf{J}_2$ are $d \times d$-matrices and Eq. (7.12) reads

$$\underbrace{\begin{pmatrix} \mathbf{P}_1 & \mathbf{P}_2 \\ \mathbf{P}_3 & \mathbf{P}_4 \end{pmatrix}}_{\mathbf{P}^{-1}} \begin{pmatrix} \boldsymbol{\nu}_x \\ \boldsymbol{\eta}_x \end{pmatrix} = \begin{pmatrix} \mathbf{J}_1 & 0 \\ 0 & \mathbf{J}_2 \end{pmatrix} \begin{pmatrix} \mathbf{P}_1\boldsymbol{\nu} + \mathbf{P}_2\boldsymbol{\eta} \\ \mathbf{P}_3\boldsymbol{\nu} + \mathbf{P}_4\boldsymbol{\eta} \end{pmatrix}. \tag{7.13}$$

Obviously, the upper part of this system yields parts of the solution, which decay for $x \to \infty$ and grow for $x \to -\infty$. The opposite is true for the lower part. An analogous equation holds for the left exterior problem. Thus the *transformed transparent boundary conditions* for the left (a) and right (b) boundary is obtained by extinguishing the respectively increasing parts of the exterior solutions:

$$\tilde{\mathbf{P}}_2\hat{\boldsymbol{\varphi}}_x(0,s) = -\tilde{\mathbf{P}}_1\hat{\boldsymbol{\varphi}}(0,s), \tag{7.14a}$$

$$\mathbf{P}_4\hat{\boldsymbol{\varphi}}_x(L,s) = -\mathbf{P}_3\hat{\boldsymbol{\varphi}}(L,s). \tag{7.14b}$$

Here, $\tilde{\mathbf{P}}_1$ and $\tilde{\mathbf{P}}_2$ are the corresponding matrices for the left exterior problem $\mathbf{P}_3$ and $\mathbf{P}_4$ (resp. $\tilde{\mathbf{P}}_1$,$\tilde{\mathbf{P}}_2$) are sub-matrices of $\mathbf{P}^{-1}$ which holds the eigenvectors of $\mathbf{A}^{-1}\mathbf{B}$ and is thus regular. Therefore $(\mathbf{P}_3 \quad \mathbf{P}_4)$ has rank d and thus at least one of the two matrices $\mathbf{P}_3$ and $\mathbf{P}_4$ (resp. $\tilde{\mathbf{P}}_1$ and $\tilde{\mathbf{P}}_2$) is regular. If the matrices $\tilde{\mathbf{P}}_2$ and $\mathbf{P}_4$ are regular, then the Laplace-transformed TBC can be written in *Dirichlet-to-Neumann form*. It is not clear, if these matrices are regular in general, but for our example this hold for all tested values of s.

7.4 The Discrete Transparent Boundary Conditions

Our proposed strategy is a purely discrete one, i.e. we do not discretise the Eq. (7.14) (by a numerical inverse Laplace transformation), but instead we derive discrete TBCs for a discretisation of (7.1). For the discretisation we choose a uniform grid with the step sizes h in space and k in time: $x_j = jh, t_n = nk$ with $j \in \mathbb{Z}, n \in \mathbb{N}_0$. We discretise (7.3) using the classical unitary Crank-Nicholson scheme in time and the central differences for the first and second spatial derivatives. The discrete $\mathbf{k}\cdot\mathbf{p}$-Schrödinger equation then reads

$$\mathrm{i}\frac{h^2}{k}\left(\boldsymbol{\varphi}_j^{n+1}-\boldsymbol{\varphi}_j^{n}\right) = -\Delta^0_{\frac{h}{2}}\left(\mathbf{N}_j\,\Delta^0_{\frac{h}{2}}\boldsymbol{\varphi}_j^{n+1/2}\right)+\mathbf{M}_{Sj}\Delta^0\boldsymbol{\varphi}_j^{n+1/2}-\Delta^0\left(\mathbf{M}_{Sj}^{\mathrm{H}}\boldsymbol{\varphi}_j^{n+1/2}\right)+V_j\,\boldsymbol{\varphi}_j^{n+1/2}, \tag{7.15}$$

with the centered difference operators

$$\Delta^0_{\frac{h}{2}}\boldsymbol{\varphi}_j = \boldsymbol{\varphi}_{j+1/2}-\boldsymbol{\varphi}_{j-1/2}, \tag{7.16a}$$

$$\Delta^0\boldsymbol{\varphi}_j = \frac{1}{2}(\Delta^+ + \Delta^-)\boldsymbol{\varphi}_j = \boldsymbol{\varphi}_{j+1}-\boldsymbol{\varphi}_{j-1}, \tag{7.16b}$$

where $\Delta^+ = \boldsymbol{\varphi}_{j+1}-\boldsymbol{\varphi}_j$, $\Delta^- = \boldsymbol{\varphi}_j-\boldsymbol{\varphi}_{j-1}$ are the standard forward and backward differences and the arithmetic time average is denoted by $\boldsymbol{\varphi}_j^{n+1/2} = (\boldsymbol{\varphi}_j^{n+1}+\boldsymbol{\varphi}_j^{n})/2$.

An appropriate discretisation scheme should carry over typical properties of the continuous equation to the difference equation. This is the case for the Crank-Nicolson scheme: it conserves the discrete ℓ^2-norm on $j \in \mathbb{Z}$ and thus it is unconditionally stable for the whole space problem. The procedure to show this is analogue to the continuous case: we multiply the discrete equation with $(\varphi_j^n)^{\mathrm{H}}$ from the left and add the Hermitian of the discrete equation multiplied by $\boldsymbol{\varphi}_j^{n+1}$ from the right. This shows, that the discrete ℓ^2-norm is conserved in time for the whole space scheme.

For the case of a scalar Schrödinger equation Arnold [2] derived 1998 a discrete TBC. This discrete TBC is reflection-free and conserves the stability properties of the whole-space Crank-Nicolson scheme. The discrete TBC has the form of a discrete convolution. The convolution coefficients are a function of Legendre polynomials but can be obtained more easily by a three-term recurrence formula. Ehrhardt and Arnold [16] showed that the imaginary parts of the convolution coefficients are not decaying but oscillating and therefore introduced summed coefficients.

To derive the discrete TBC for a system of Schrödinger equations we solve the $\mathscr{Z}$-transformed system of ordinary difference equations (OΔEs) in the exterior domain. Then all its solutions are determined by eigenvalues and eigenvectors, which can distinguish between decaying and increasing solutions as indicated by the absolute value of the involved eigenvalue. We obtain the transformed discrete TBC by claiming, that no influence of increasing solutions exists.

In the exterior domain $j \geq J$ ($x_J = L$) the Crank-Nicholson scheme (7.15) simplifies to

$$\mathrm{i}\frac{h^2}{k}\left(\boldsymbol{\varphi}_j^{n+1}-\boldsymbol{\varphi}_j^{n}\right) = -\mathbf{N}\Delta^+\Delta^-\boldsymbol{\varphi}_j^{n+1/2}+\mathrm{i}h\mathbf{M}\frac{1}{2}(\Delta^+ + \Delta^-)\boldsymbol{\varphi}_j^{n+1/2}+h^2\mathbf{V}\boldsymbol{\varphi}_j^{n+1/2}. \tag{7.17}$$

A $\mathscr{Z}$-transformation given by

$$\mathscr{Z}\{\boldsymbol{\varphi}_j^n\} = \hat{\boldsymbol{\varphi}}_j(z) := \sum_{n=0}^{\infty} z^{-n}\boldsymbol{\varphi}_j^n, \qquad z \in \mathbb{C}, \qquad |z| > 1, \tag{7.18}$$

applied to (7.17) yields for $|z| > 1$ the system of second order OΔEs

$$2\mathrm{i}\frac{h^2}{k}\frac{z-1}{z+1}\hat{\boldsymbol{\varphi}}_j = -\mathbf{N}\Delta^+\Delta^-\hat{\boldsymbol{\varphi}}_j + \mathrm{i}h\mathbf{M}\frac{1}{2}(\Delta^+ + \Delta^-)\hat{\boldsymbol{\varphi}}_j + h^2\mathbf{V}\hat{\boldsymbol{\varphi}}_j, \quad j \geq J. \tag{7.19}$$

As in the continuous case we proved in [30] that this transformed right exterior problem (7.19) posses a *unique solution*:

Lemma 7.2 ([30, Lemma 4.1]). *For each $z \in \mathbb{C}$ with $|z| > 1$ the $\mathscr{Z}$-transformed exterior problem* (7.19) *with the boundary data*

$$\hat{\boldsymbol{\varphi}}_{J-1} = \hat{\boldsymbol{\Phi}}, \qquad \hat{\boldsymbol{\varphi}}_\infty = 0 \tag{7.20}$$

has a unique solution.

We proceed with the construction of the discrete TBC and define $\hat{\xi}_j = \Delta^-\hat{\boldsymbol{\varphi}}_j$ in order to write (7.19) as a system of first order OΔEs

$$\underbrace{\begin{pmatrix} \mathrm{i}\frac{h}{2}\mathbf{M} & -\mathbf{N} \\ -\mathbf{I} & \mathbf{I} \end{pmatrix}}_{\mathbf{A}} \begin{pmatrix} \Delta^+\hat{\boldsymbol{\varphi}}_j \\ \Delta^+\hat{\xi}_j \end{pmatrix} = \underbrace{\begin{pmatrix} h^2 2\frac{z-1}{z+1}\frac{1}{k}\mathrm{i}\mathbf{I} - h^2\mathbf{V} & -\mathrm{i}\frac{h}{2}\mathbf{M} \\ \mathbf{0} & -\mathbf{I} \end{pmatrix}}_{\mathbf{B}} \begin{pmatrix} \hat{\boldsymbol{\varphi}}_j \\ \hat{\xi}_j \end{pmatrix}, \quad j \geq J \tag{7.21}$$

In [30, Lemma 4.5] we proved the regularity of the matrix $\mathbf{A}^{-1}\mathbf{B} + \mathbf{I}$ for $|z| \neq 1$ and thus we can rewrite (7.21) as

$$\begin{pmatrix} \hat{\boldsymbol{\varphi}}_{j+1} \\ \hat{\xi}_{j+1} \end{pmatrix} = (\mathbf{A}^{-1}\mathbf{B} + \mathbf{I}) \begin{pmatrix} \hat{\boldsymbol{\varphi}}_j \\ \hat{\xi}_j \end{pmatrix}, \qquad j \geq J. \tag{7.22}$$

As in the continuous case there exist a *discrete version of the Splitting Theorem*:

Theorem 7.2 ([30, Theorem 4.2]). *For $|z| \neq 1$ d of the $2d$ eigenvalues of $\mathbf{A}^{-1}\mathbf{B}+\mathbf{I}$ have an absolute value strictly larger than 1 and d have an absolute value strictly smaller than 1.*

In other words, the eigenvalues $\lambda_1, \ldots, \lambda_{2n}$ of $\mathbf{A}^{-1}\mathbf{B}$ split into two commensurate groups, i.e. the solutions involving those with $|\lambda_k + 1| < 1$ for $k = 1, \ldots, n$ decay for $j \to \infty$ and those with $|\lambda_k + 1| > 1$ for $k = n, \ldots, 2n$ decay for $j \to -\infty$.

Thus, as in the continuous case, we can split the Jordan form $\mathbf{J} = \mathrm{diag}(\mathbf{J}_1, \mathbf{J}_2)$ of $\mathbf{A}^{-1}\mathbf{B}$, $\mathbf{J}_1$ containing the Jordan blocks corresponding to solutions decaying for $j \to$

∞ and $\mathbf{J}_2$ those which increase for $j \to \infty$. With the matrix of left eigenvectors $\mathbf{P}^{-1}$ the equation

$$\underbrace{\begin{pmatrix}\mathbf{P}_1 & \mathbf{P}_2\\ \mathbf{P}_3 & \mathbf{P}_4\end{pmatrix}}_{\mathbf{P}^{-1}}\begin{pmatrix}\Delta^+\hat{\boldsymbol{\varphi}}_j\\ \Delta^+\hat{\boldsymbol{\xi}}_j\end{pmatrix} = \mathbf{P}^{-1}\mathbf{A}^{-1}\mathbf{B}\begin{pmatrix}\hat{\boldsymbol{\varphi}}_j\\ \hat{\boldsymbol{\xi}}_j\end{pmatrix} = \begin{pmatrix}\mathbf{J}_1 & \mathbf{0}\\ \mathbf{0} & \mathbf{J}_2\end{pmatrix}\begin{pmatrix}\mathbf{P}_1\hat{\boldsymbol{\varphi}}_j + \mathbf{P}_2\hat{\boldsymbol{\xi}}_j\\ \mathbf{P}_3\hat{\boldsymbol{\varphi}}_j + \mathbf{P}_4\hat{\boldsymbol{\xi}}_j\end{pmatrix} \tag{7.23}$$

holds and the *transformed discrete TBCs* read

$$\tilde{\mathbf{P}}_1\hat{\boldsymbol{\varphi}}_1 + \tilde{\mathbf{P}}_2\hat{\boldsymbol{\xi}}_1 = 0, \tag{7.24a}$$

$$\mathbf{P}_3\hat{\boldsymbol{\varphi}}_J + \mathbf{P}_4\hat{\boldsymbol{\xi}}_J = 0, \tag{7.24b}$$

for the left (a) and right (b) boundary respectively. For nonsingular $\mathbf{P}_4$ and $\tilde{\mathbf{P}}_2$ (either $\mathbf{P}_4$ or $\mathbf{P}_3$ is regular, since their composition is a linear independent set of eigenvectors—the same holds for $\tilde{\mathbf{P}}_2$ and $\tilde{\mathbf{P}}_1$) and for $\hat{\mathbf{D}} = -\mathbf{P}_4^{-1}\mathbf{P}_3$ and $\hat{\tilde{\mathbf{D}}} = -\tilde{\mathbf{P}}_2^{-1}\tilde{\mathbf{P}}_1$ we write

$$\Delta^-\hat{\boldsymbol{\varphi}}_1 = \hat{\tilde{\mathbf{D}}}\hat{\boldsymbol{\varphi}}_1, \tag{7.25a}$$

$$\Delta^-\hat{\boldsymbol{\varphi}}_J = \hat{\mathbf{D}}\hat{\boldsymbol{\varphi}}_J. \tag{7.25b}$$

After an inverse $\mathscr{Z}$-transformation the *discrete TBCs* read

$$\boldsymbol{\varphi}_1^{n+1} - \boldsymbol{\varphi}_0^{n+1} - \tilde{\mathbf{D}}^0\boldsymbol{\varphi}_1^{n+1} = \sum_{k=1}^{n}\tilde{\mathbf{D}}^{n+1-k}\boldsymbol{\varphi}_1^k, \tag{7.26a}$$

$$\boldsymbol{\varphi}_J^{n+1} - \boldsymbol{\varphi}_{J-1}^{n+1} - \mathbf{D}^0\boldsymbol{\varphi}_J^{n+1} = \sum_{k=1}^{n}\mathbf{D}^{n+1-k}\boldsymbol{\varphi}_J^k. \tag{7.26b}$$

We remark that in Eqs. (7.25a) and (7.26a) the left boundary convolution is given at the interior grid point $j = 1$. Of course, the boundary condition can also be formulated at $j = 0$ using $\hat{\boldsymbol{\xi}}_j = \Delta^+\hat{\boldsymbol{\varphi}}_j$. This changes the lower row in $\mathbf{A}$ and $\mathbf{B}$ and thus the matrix $\tilde{\mathbf{D}}$ differs from $\mathbf{D}$. Posing the boundary condition at $j = 1$ as we do, has the advantage, that these matrices coincide, if the coefficients for $x < 0$ and $x > L$ are equal, what occurs often in the application. In that case we can reduce the numerical effort to calculate the convolution coefficients by half.

For a scalar Schrödinger equation Ehrhardt and Arnold [16] showed that the imaginary parts of the coefficients were not decaying but oscillating. As a remedy they introduced *summed coefficients* that decay rapidly like $O(n^{-3/2})$. Since the scalar equation is as a special case included in our system, it suggests itself to use the summed coefficients, although we cannot give any proof of the asymptotic behaviour of the systems' coefficients.

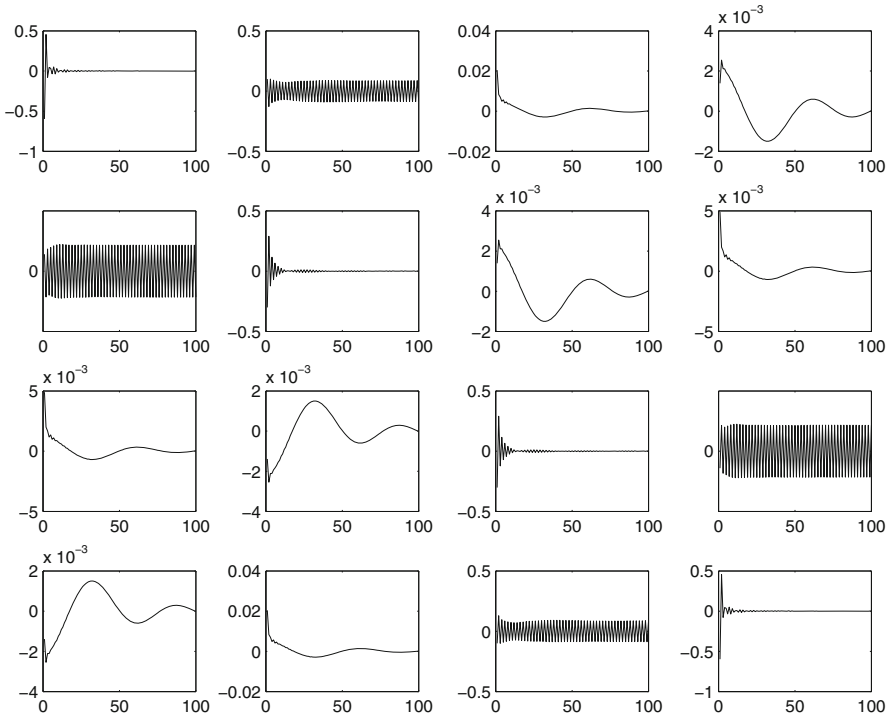


Fig. 7.4 Real parts of the convolution coefficients $\mathbf{D}_{k,l}^n$

In the Figs. 7.4, 7.5, 7.6, 7.7 we plot typical examples of the numerically calculated coefficients for a 4×4 system. Especially in Fig. 7.5 one can observe in the diagonal elements the typical oscillating behaviour known from the scalar case. The diagonal elements show the same properties as those for the scalar case. For the summed coefficients

$$\widehat{\widetilde{\mathbf{S}}}_{s,l} = \frac{z+1}{z}\widehat{\widetilde{\mathbf{D}}}_{s,l} \qquad \text{and} \qquad \hat{\mathbf{S}}_{s,l} = \frac{z+1}{z}\hat{\mathbf{D}}_{s,l} \tag{7.27}$$

the boundary conditions read

$$\boldsymbol{\varphi}_1^{n+1} - \boldsymbol{\varphi}_0^{n+1} - \tilde{\mathbf{S}}^0\boldsymbol{\varphi}_1^{n+1} = \sum_{k=1}^{n} \tilde{\mathbf{S}}^{n+1-k}\boldsymbol{\varphi}_1^k - \boldsymbol{\varphi}_1^n + \boldsymbol{\varphi}_0^n, \tag{7.28a}$$

$$\boldsymbol{\varphi}_J^{n+1} - \boldsymbol{\varphi}_{J-1}^{n+1} - \mathbf{S}^0\boldsymbol{\varphi}_J^{n+1} = \sum_{k=1}^{n} \mathbf{S}^{n+1-k}\boldsymbol{\varphi}_J^k - \boldsymbol{\varphi}_J^n + \boldsymbol{\varphi}_{J-1}^n. \tag{7.28b}$$

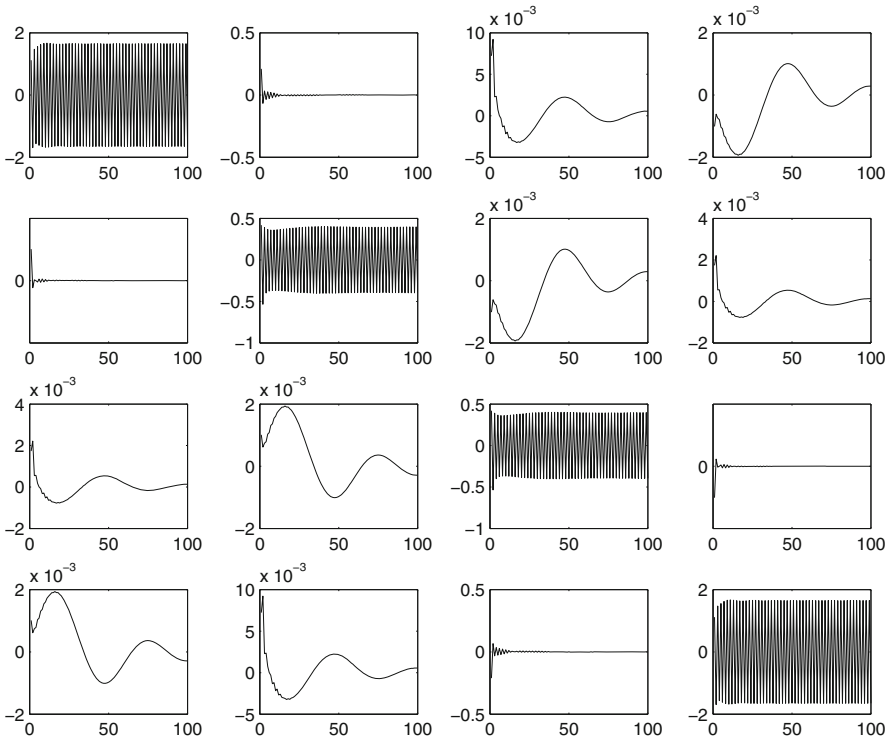


Fig. 7.5 Imaginary parts of the convolution coefficients $\mathbf{D}^n_{k,l}$

In order to compute the convolution coefficients we need to *inverse $\mathscr{Z}$-transform the transformed kernels*. We assume that the $\mathscr{Z}$-transform (7.18) is analytic for $|z| > R \geq 0$. The coefficients are then recovered by

$$\ell_n = \frac{1}{2\pi i} \oint_{S_\rho} \hat{\ell}(z) z^{n-1} dz,$$

where S_ρ denotes the circle with radius $\rho > R$. With the substitution $z = \rho e^{i\varphi}$ we obtain

$$\ell_n = \frac{\rho^n}{2\pi} \int_0^{2\pi} \hat{\ell}\left(\rho\, e^{i\varphi}\right) e^{in\varphi} d\varphi\,. \tag{7.29}$$

For $\rho > 1$, the *amplification factors* ρ^n in (7.29) will be the reason for the numerical instabilities. On the other hand, $\rho = 1$ cannot be chosen either for the application to DTBCs, due to the poor regularity of $\hat{\mathbf{D}}(z)$ on the unit circle. For the scalar Schrödinger equation, e.g., $\hat{d}\,(z)$ has two branch-points of type $\sqrt{z^2-1}$

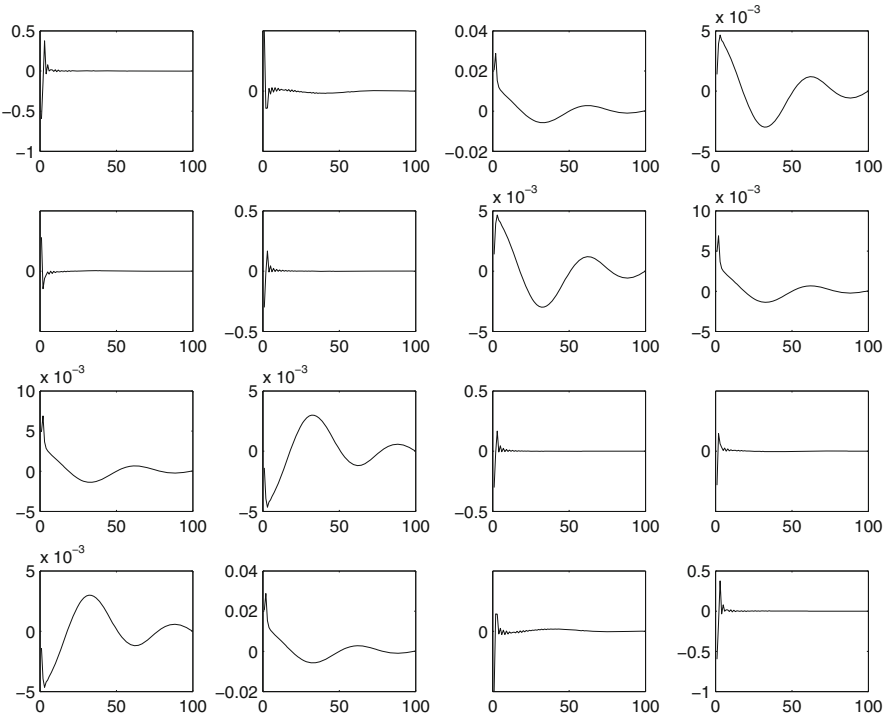


Fig. 7.6 Real parts of the summed convolution coefficients $\mathbf{S}_{k,l}^n$

(cf. [2, 16]), and hence too many quadrature points would be necessary for the numerical evaluation of (7.29). But $\hat{d}(z)$ is analytic for $|z| > 1$. So, one has to choose ρ as a compromise between more smoothness of $\hat{\ell}\big|_{|z|=\rho}$ (which allows for an efficient discretisation of (7.29)), and growing instabilities for large ρ. This situation depending on the inversion radius is depicted in Fig. 7.8.

For the numerical inverse transformation we choose a radius r and N equidistant sampling points $z_k = r\, e^{-ik2\pi/N}$. The approximate inverse transform,

$$\ell_n^N = \frac{1}{N} r^n \sum_{k=0}^{N-1} \hat{\ell}(z_k)\; e^{ink\frac{2\pi}{N}}, \qquad n = 0, \ldots N-1 \tag{7.30}$$

can then be calculated efficiently by an FFT. The numerical error of ℓ_n^N can be separated into ε_{approx}, the approximation error due to the finite number of sampling points, and the roundoff error ε_{round}, which is amplified by ρ^n. We refer the reader to [29, 30] for more details and error estimates for this numerical inversion problem.

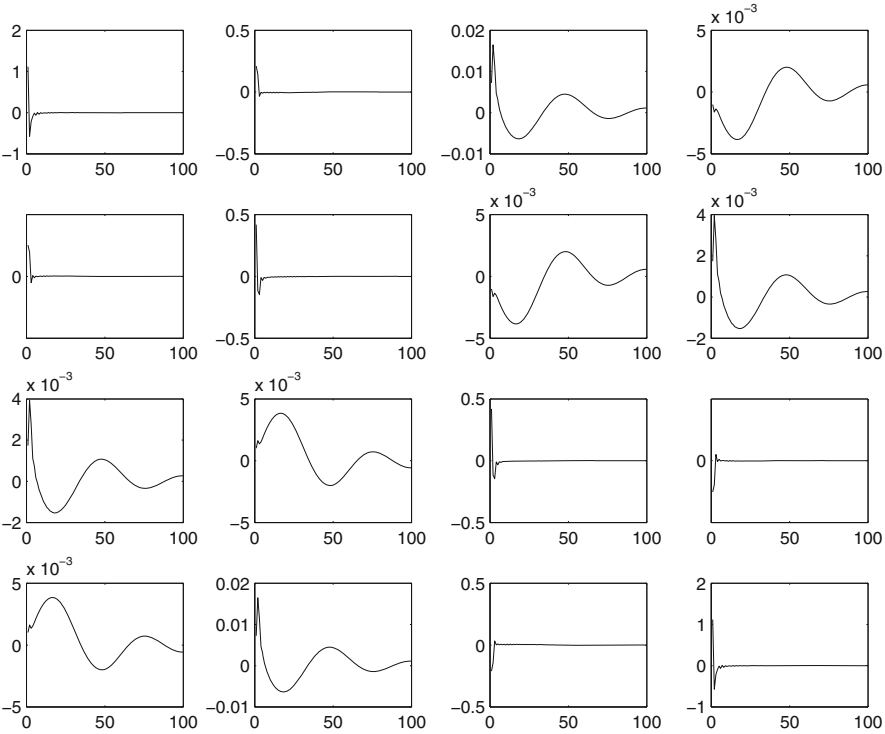


Fig. 7.7 Imaginary parts of the summed convolution coefficients $\mathbf{S}^n_{k,l}$

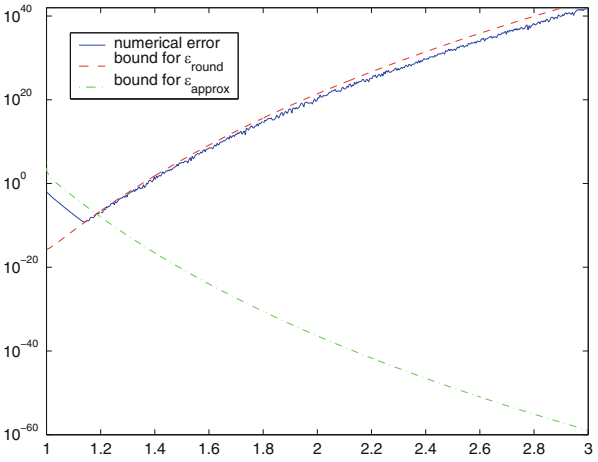


Fig. 7.8 Rounding and discretisation errors depending on the inversion radius

7.5 The Sum-of-Exponentials Approach and the Fast Evaluation of the Convolution

In order to reduce the numerical effort of the boundary convolutions (7.26), below that effort of the interior scheme, it is necessary to make some approximation. We will use the approach of Arnold, Ehrhardt and Sofronov [3] to approximate the coefficients $\tilde{s}^n_{s,l}$ by the *sum-of-exponentials* approach and show a method to evaluate the discrete convolution with the approximated convolution coefficients $\tilde{a}^n_{s,l}$ very efficiently. Afterwards we explain how these approximated convolution coefficients $\tilde{a}^n_{s,l}$ enable us to evaluate the discrete convolution efficiently.

The sum-of-exponentials approximation has to be done for each element in $\mathbf{S}$ separately. We use for each $s, \tau = 1, \ldots, d$ the following ansatz

$$\tilde{s}^n_{s,\tau} \approx \tilde{a}^n_{s,\tau} := \begin{cases} \tilde{s}^n_{s,\tau}, & n = 0, \ldots, \nu - 1 \\ \sum\limits_{l=1}^{L(s,\tau)} b_{s,\tau,l}\, q^{-n}_{s,\tau,l}, & n = \nu, \nu+1, \ldots \end{cases}, \tag{7.31}$$

where the number of summands in the approximation $L(s,\tau) \in \mathbb{N}$ and the starting index $\nu \geq 0$ (to disregard outliers) are tuneable parameters. The approximation quality of this sum-of-exponentials ansatz depends on $L(s,\tau)$, ν and the sets $\{b_{s,\tau,l}\}$ and $\{q_{s,\tau,l}\}$ for all $s, \tau = 1, \ldots, d$.

Next we present the method to calculate these sets for given $L(s,\tau)$ and ν. We consider the formal power series

$$f_{s,\tau}(x) := \tilde{s}^{\nu}_{s,\tau} + \tilde{s}^{\nu+1}_{s,\tau} x + \tilde{s}^{\nu+2}_{s,\tau} x^2 + \ldots, \qquad \text{for } |x| \leq 1\,. \tag{7.32}$$

If the Padé approximation of (7.32)

$$\tilde{f}_{s,\tau}(x) := \frac{n^{(L(s,\tau)-1)}_{s,\tau}(x)}{d^{(L(s,\tau))}_{s,\tau}(x)}$$

exists (where the numerator and the denominator are polynomials of degree $L(s,\tau) - 1$ and $L(s,\tau)$ respectively), then its Taylor series

$$\tilde{f}_{s,\tau}(x) = \tilde{a}^{\nu}_{s,\tau} + \tilde{a}^{\nu+1}_{s,\tau} x + \tilde{a}^{\nu+2}_{s,\tau} x^2 + \ldots$$

satisfies the conditions

$$\tilde{a}^n_{s,\tau} = \tilde{s}^n_{s,\tau} \quad \text{for} \quad n = \nu, \nu+1, \ldots, 2L(s,\tau) + \nu - 1$$

according to the definition of the Padé approximation rule. In [3] Arnold, Ehrhardt and Sofronov showed how to compute the coefficient sets $\{b_{s,\tau,l}\}$ and $\{q_{s,\tau,l}\}$:

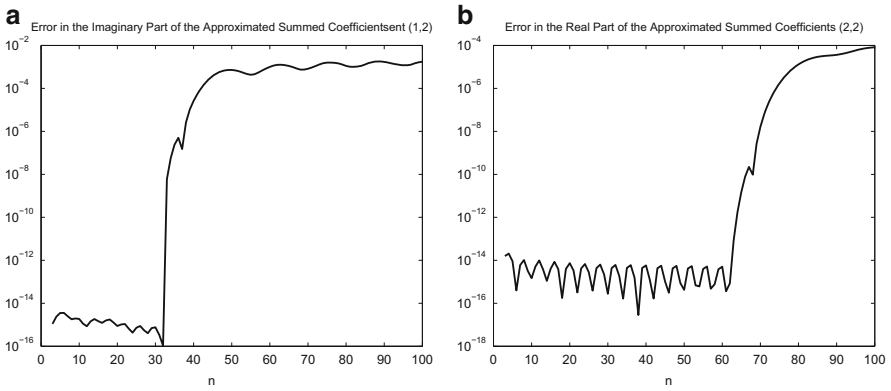


Fig. 7.9 Error $|\tilde{s}^n_{s,\tau} - \tilde{a}^n_{s,\tau}|$ versus n: imaginary part (**a**) for $s = 1$, $\tau = 2$ and (**b**) real part for $s = \tau = 2$

Theorem 7.3 ([3], Theorem 3.1). *Let $d^{L(s,\tau)}_{s,\tau}$ have $L(s,\tau)$ simple roots $q_{s,\tau,l}$ with $|q_{s,\tau,l}| > 1$, $l = 1, \ldots, L(s,\tau)$. Then*

$$\tilde{a}^n_{s,\tau} = \sum_{l=1}^{L(s,\tau)} b_{s,\tau,l}\, q^{-n}_{s,\tau,l}\,, \qquad n = \nu, \nu+1, \ldots,$$

where

$$b_{s,\tau,l} := -\frac{n^{(L(s,\tau)-1)}_{s,\tau}(q_{s,\tau,l})}{\left(d^{(L(s,\tau))}_{s,\tau}\right)'(q_{s,\tau,l})}\, q^{\nu-1}_{s,\tau,l} \neq 0\,, \qquad l = 1, \ldots, L(s,\tau)\,.$$

The asymptotic decay of the $\tilde{a}^n_{s,\tau}$ is exponential. This is due to the sum-of-exponentials ansatz (7.31) and the assumption $|q_{s,\tau,l}| > 1$, $l = 1, \ldots, L(s,\tau)$.

If we use a $[L(s,\tau) - 1|L(s,\tau)]$ Padé approximant to (7.32) then the first $2L(s,\tau) + \nu - 1$ coefficients are reproduced exactly; however, the asymptotic behaviour of $\tilde{s}^n_{s,\tau}$ and $\tilde{a}^n_{s,\tau}$ (as $n \to \infty$) differs strongly (algebraic versus exponential decay).

We note that the Padé approximation must be performed with high precision ($2L(s,\tau)-1$ digits mantissa length) to avoid a 'nearly breakdown' by ill conditioned steps in the Lanczos algorithm (cf. [8]). If such problems still occur or if one root of the denominator is smaller than 1 in absolute value, the orders of the numerator and denominator polynomials are successively reduced. For example in our numerical test case we started with $L(s,\tau) \equiv 30$ and except from two outlier values the finally reached values of $L(s,\tau)$ were between 25 and 30. Figure 7.9 shows the error $|\tilde{s}^n_{s,\tau} - \tilde{a}^n_{s,\tau}|$ versus n for the outlier with $L(1,2) = 15$ for the imaginary part of $\tilde{s}^n_{1,2}$ (a) and with $L(2,2) = 30$ for the real part of $\tilde{s}^n_{2,2}$ (b). The error increases significantly for $n > 2L(s,\tau) + 1$.

Now we describe the fast evaluation of the discrete approximate convolution

$$C_{s,\tau}^{(n+1)}(u) := \sum_{k=1}^{n+1-\nu} \tilde{a}_{s,\tau}^{n+1-k}\, u_{\tau,J}^{k}\,, \quad \text{with } \tilde{a}_{s,\tau}^{n} := \sum_{l=1}^{L(s,\tau)} b_{s,\tau,l}\, q_{s,\tau,l}^{-n}, \quad n = \nu, \nu+1, \ldots$$

that can be calculated efficiently by a simple *recurrence formula*:

Theorem 7.4 ([3], Theorem 4.1.).

$$C_{s,\tau}^{(n+1)}(u) = \sum_{l=1}^{L(s,\tau)} C_{s,\tau,l}^{(n+1)}(u) \tag{7.33}$$

with

$$C_{s,\tau,l}^{(n+1)}(u) = q_{s,\tau,l}^{-1} C_{s,\tau,l}^{(n)} + b_{s,\tau,l} q_{s,\tau,l}^{-\nu} u_{\tau,J}^{n+1-\nu}\,, \qquad n = \nu, \nu+1, \ldots \tag{7.34}$$

$$C_{s,\tau,l}^{(\nu)}(u) \equiv 0\,.$$

This is an efficient and local-in-time approximation using data only from ν levels earlier (typically $\nu = 2$). Also, there is no need to store the boundary data, which becomes computationally expensive especially for higher dimensional problems.

Note that similar recursive convolution algorithms are successfully used in other applications as well, see [7] and references therein.

Let us summarize the proposed method to evaluate approximate discrete TBCs

1. For each s, τ choose $L(s,\tau)$ and ν and calculate the exact convolution coefficients $\tilde{s}_{s,\tau}^{n}$ for $n = 0, \ldots, 2L(s,\tau) + \nu - 1$.
2. For each s, τ use the Padé approximation for the Taylor series with $\tilde{a}_{s,\tau}^{n} = \tilde{s}_{s,\tau}^{n}$, for $n = \nu, \nu+1, \ldots, 2L(s,\tau) + \nu - 1$ to calculate the sets $\{b_{s,\tau,l}\}$ and $\{q_{s,\tau,l}\}$ for all $s, \tau = 1, \ldots, d$ according to Theorem 7.3.
3. Implement the recurrence formulas (7.33) and (7.34) to calculate the approximate convolutions.

7.6 Numerical Results

Finally we present briefly the numerical results for simulating the time-dependent behaviour of the quantum well with the data of Sect. 7.1. We choose the time step $k = 0.015$ and the space step $h = 1/20$ for the computational domain $[-12, 5]$ and compute the convolution coefficients of the discrete TBC as described in Sect. 7.5.

First, to study the behaviour of the discrete TBC, we consider a system of Schrödinger equations with potential set to zero ('free Schrödinger system'). As initial condition we use the Gaussian wave packet of heavy and light holes (7.4) stimulating a slow and a fast eigenmode. Figure 7.10 shows the time-dependent behaviour of the first two components φ_1 (solid) and φ_2 (dashed). We focus on the

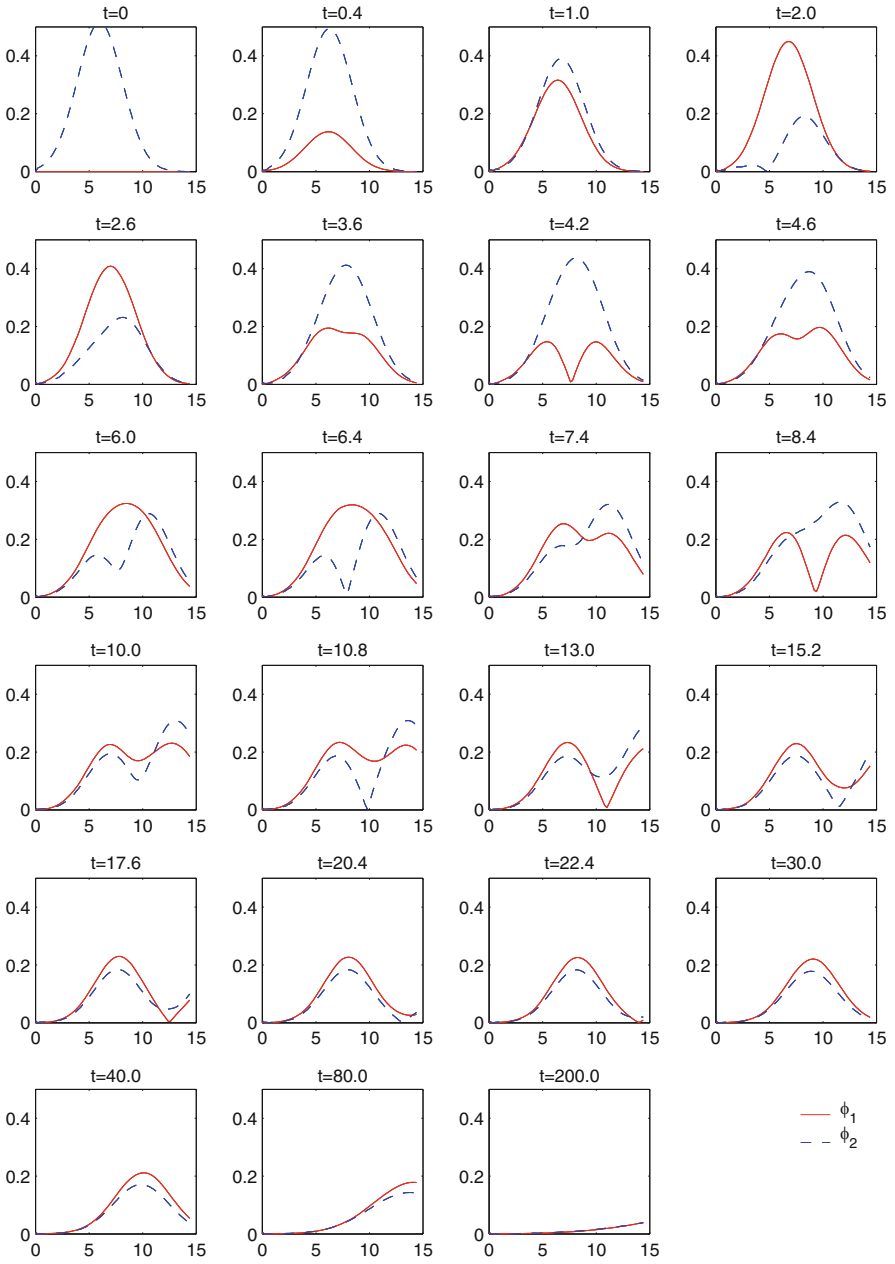


Fig. 7.10 Time dependent behaviour of φ_{1j} (*solid*) and φ_{2j} (*dashed*) for a free Schrödinger system

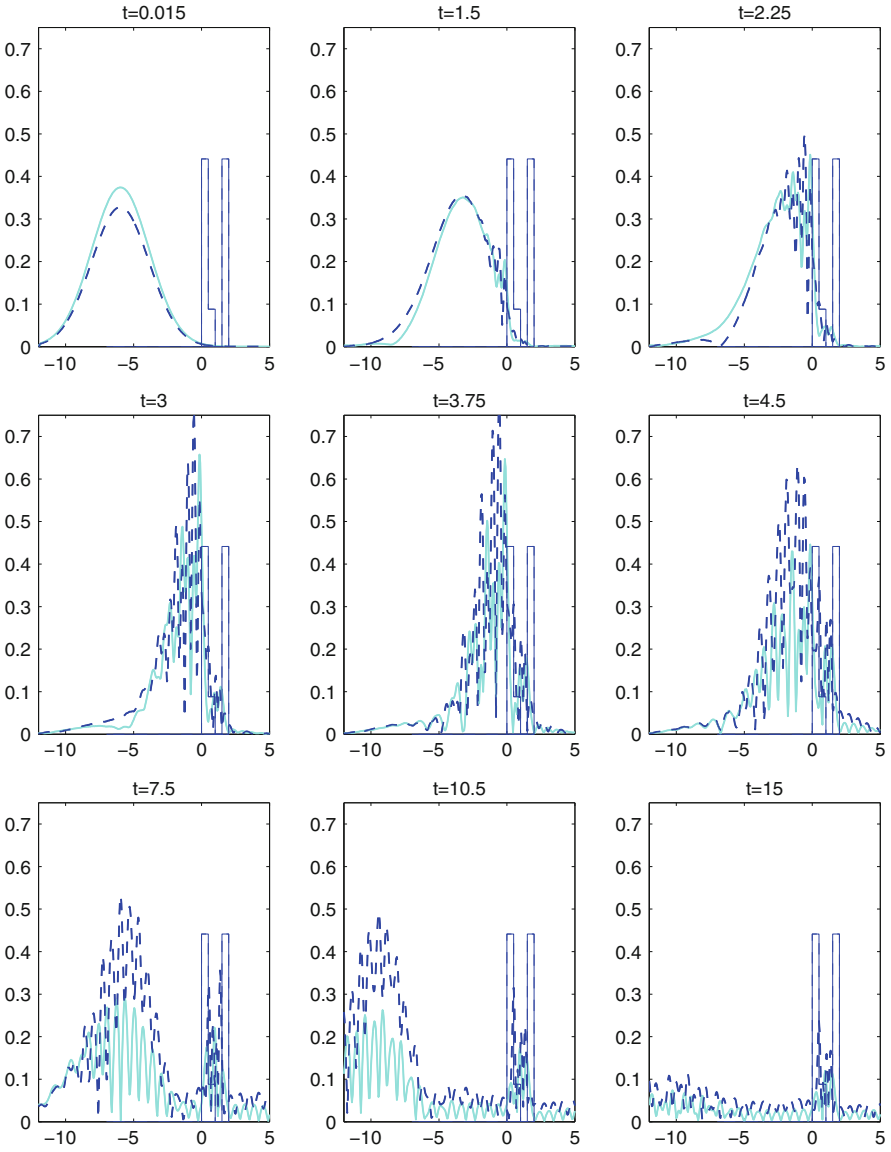


Fig. 7.11 Time dependent behaviour of φ_{1j} (*solid*) and φ_{2j} (*dashed*) for a system with DBSQW structure

first two components, since there is less mass in component three and four. The density oscillates between the components, moves to the right, fragments in two and the faster wave packet leaves the domain of computation without any visible reflections.

Secondly we consider only the faster mode and add the DBSQW structure (7.5), cf. Fig. 7.11. When the wave packet reaches the first barrier, it is partly reflected

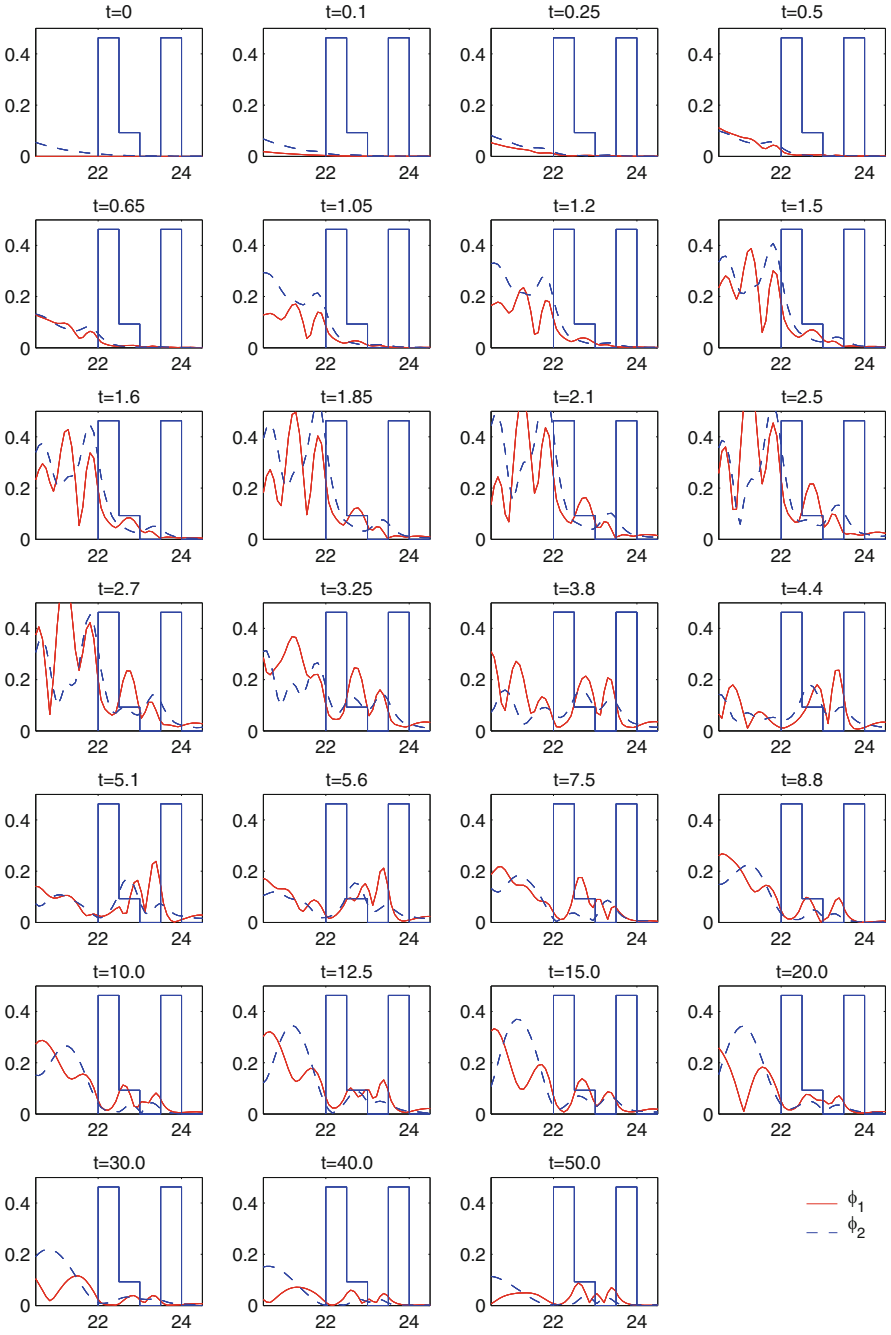


Fig. 7.12 Time dependent behaviour of φ_{1j} (*solid*) and φ_{2j} (*dashed*) for a different DBSQW system

and partly transmitted. With advancing time some part of the density accumulates between the barriers and is slowly transmitted through the second barrier, then leaving the domain of computation. The part of the density, which is reflected at the first barrier moves on to the left and after some time most part of the solution leaves the computational domain in a packet. The wave packet does not recompose smoothly. A simulation for a slightly different DBSQW structure is shown in Fig. 7.12.

The *relative* ℓ^2*-error* is defined as $e_L(t) = \|\boldsymbol{\varphi} - \boldsymbol{\varphi}_a\|_2 / \|\boldsymbol{\varphi}(.,0)\|_2$, where $\boldsymbol{\varphi}_a$ denotes the approximate solution obtained with the approximated discrete TBCs and $\boldsymbol{\varphi}$ is the solution calculated with exact discrete TBCs. When using the parameter $L(s,\tau) = 30$ initially in the sum-of-exponential approach (7.31) this relative ℓ^2-error increases moderately with respect to time (due to the interaction with the potential) but remains bounded after 1,000 time steps below $6 \cdot 10^{-3}$ which is acceptable for this application.

7.7 Conclusion

In this chapter we derived and analyzed discrete transparent boundary conditions (TBCs) for instationary one-dimensional **k·p**-Schrödinger equations modelling layered semiconductor devices. As the **k·p** method in combination with an envelope function approximation is frequently used to calculate the near bandedge electronic band structure of semiconductor heterostructures, this novel approach is a highly relevant step forward to enable stable and highly accurate computations of this whole space problem.

Our new TBCs are constructed on a purely discrete level thereby ensuring the conservation of the stability of the underlying scheme and the prevention of any numerical reflections. Since these novel exact discrete TBCs are non-local in time and thus very costly we proposed a sum-of-exponentials approach allowing for a fast calculation of the boundary convolution terms.

Future work will be concerned with the extension of our approach to nonlinear systems, cf. [31], to systems self-consistently coupled the Poisson equation, cf. [17] and to higher-dimensional problems that are of practical relevance, e.g. for the N-particle electronic Schrödinger equation or in mathematical finance.

References

1. A. Antoine, A. Arnold, C. Besse, M. Ehrhardt, A. Schädle, A review of transparent and artificial boundary conditions techniques for linear and nonlinear Schrödinger equations. Commun. Comput. Phys. **4**, 729–796 (2008)
2. A. Arnold, Numerically absorbing boundary conditions for quantum evolution equations. VLSI Design **6**, 313–319 (1998)

3. A. Arnold, M. Ehrhardt, I. Sofronov, Discrete transparent boundary conditions for the Schrödinger equation: fast calculation, approximation, and stability. Commun. Math. Sci. **1**, 501–556 (2003)
4. G. Bastard, *Wave Mechanics Applied to Semiconductor Heterostructures* (Hasted Press, 1988)
5. U. Bandelow, H.-C. Kaiser, Th. Koprucki, J. Rehberg, Spectral properties of **k·p** Schrödinger operators in one space dimension. Numer. Funct. Anal. Optim. **21**, 379–409 (2000)
6. A.C.E. Bittencourt, A.M. Cohen, G.E. Marques, Strain-induced enhancement of resonant current of holes in multilayered heterostructures. Phys. Rev. B **57**, 4525–4543 (1998)
7. G. Blakiewicz, W. Janke, Recursive convolution algorithms for time–domain simulation of electronic circuits. Comp. Meth. Sci. Techn. **7**, 91–109 (2001)
8. A. Bultheel, M. van Barel, *Linear algebra, rational approximation and orthogonal polynomials* (Studies in Computational Mathematics 6, North–Holland, 1997)
9. M.G. Burt, The justification for applying the effective-mass approximation to microstructures. J. Phys. Condens. Matter **4**, 6651–6690 (1992)
10. M.G. Burt, Direct derivation of effective-mass equations for microstructures with atomically abrupt boundaries. J. Phys. Condens. Matter **11**, R53–R83 (1998)
11. M. Cardona, *Fundamentals of Semiconductors* (Springer, Berlin, 1996)
12. C.Y.-P. Chao, S.L. Chuang, Spin-orbit-coupling effects on the valence-band structure of strained semiconductor quantum wells. Phys. Rev. B **46**, 4110–4122 (1992)
13. S.L. Chuang, Efficient band-structure calculations of strained quantum wells. Phys. Rev. B **43**, 9649–9661 (1991)
14. S.L. Chuang, *Physics of Optoelectronic Devices* (Wiley & Sons, New York, 1995)
15. P. Debernardi, P. Fasano, Quantum confined Stark effect in semiconductor quantum wells including valence band mixing and Coulomb effects. IEEE J. Quant. Electron. **29**, 2741–2755 (1993)
16. M. Ehrhardt, A. Arnold, Discrete transparent boundary conditions for the Schrödinger equation. Riv. Matem. Univ. di Parma **6**, 57–108 (2001)
17. M. Ehrhardt, A. Zisowsky, Fast Calculation of Energy and Mass preserving solutions of Schrödinger–Poisson systems on unbounded domains. J. Comput. Appl. Math. **187**, 1–28 (2006)
18. M. Ehrhardt, Discrete transparent boundary conditions for Schrödinger-type equations for non-compactly supported initial data. Appl. Numer. Math. **58**, 660–673 (2008)
19. E.O. Kane, Energy Band Theory. in Paul, W. (ed.): Handbook on Semiconductors (North-Holland, Amsterdam, New York, Oxford, 1982)
20. J.M. Luttinger, W. Kohn, Motion of electrons and holes in perturbed periodic fields. Phys. Rev. **94**, 869–883 (1955)
21. A.T. Meney, B. Gonul, E.P. O'Reilly, Evaluation of various approximations used in the envelope-function method. Phys. Rev. B **50**, 10893–10904 (1994)
22. V. Sankaran, J. Singh, Formalism for tunneling of mixed-symmetry electronic states: application to electron and hole tunneling in direct- and indirect-band-gap GaAs/$Al_xGa_{1-x}As$ structures. Phys. Rev. B **44**, 3175–3186 (1991)
23. J. Singh, *Physics of semiconductors and their heterostructures* (McGraw-Hill, New York, 1993)
24. C. Sirtori, P. Kruck, S. Barbieri, Ph. Collot, J. Nagle, M. Beck, J. Faist, U. Oesterle, GaAs/$Al_xGa_{1-x}As$ quantum cascade lasers. Appl. Phys. Lett. **73**, 3486–3488 (1998)
25. J.A. Stovneng, E.H. Hauge, Time-dependent resonant tunneling of wave packets in the tight-binding model. Phys. Rev. B **44**, 13582–13594 (1991)
26. M. Sweeny, J. Xu, Resonant interband tunnel diodes. Appl. Phys. Lett. **54**, 546–548 (1989)
27. M. Wagner, H. Mizuta, Complex-energy analysis of intrinsic lifetimes of resonances in biased multiple quantum wells. Phys. Rev. B **48**, 14393–14406 (1993)
28. J. Zhang, B. Gu, Temporal characteristics of electron tunneling in double-barrier stepped quantum-well structures. Phys. Rev. B **43**, 5028–5034 (1991)
29. A. Zisowsky, Discrete transparent boundary conditions for systems of evolution equations, Ph.D. dissertation, Technische Universität Berlin (2003)

30. A. Zisowsky, A. Arnold, M. Ehrhardt, Th. Koprucki, Discrete transparent boundary conditions for transient **k·p**-Schrödinger equations with application to quantum heterostructures. Z. Angew. Math. Mech. **85**, 793–805 (2005)
31. A. Zisowsky, M. Ehrhardt, Discrete artificial boundary conditions for nonlinear Schrödinger equations. Math. Comput. Modell. **47**, 1264–1283 (2008)
32. A. Zlotnik, I. Zlotnik, Finite element method with discrete transparent boundary conditions for the time-dependent 1D Schrödinger equation. to appear in: Kinetic and Related Models (2013)
33. A. Zlotnik, I. Zlotnik, Finite element method with discrete transparent boundary conditions for the one-dimensional non-stationary Schrödinger equation. Doklady Mathematics **86**, 750–755 (2012)

Chapter 8
Discrete Transparent Boundary Conditions for Multi-Band Effective Mass Approximations

Dirk Klindworth, Matthias Ehrhardt, and Thomas Koprucki

Abstract This chapter is concerned with the derivation and numerical testing of *discrete transparent boundary conditions* (DTBCs) for stationary *multi-band effective mass approximations* (MEMAs). We analyze the continuous problem and introduce *transparent boundary conditions* (TBCs). The discretization of the differential equations is done with the help of finite difference schemes.

A fully discrete approach is used in order to develop DTBCs that are completely reflection-free. The analytical and discrete dispersion relations are analyzed in depth and the limitations of the numerical computations are shown. We extend the results of earlier works on DTBCs for the scalar Schrödinger equation by considering alternative finite difference schemes.

The introduced schemes and their corresponding DTBCs are tested numerically on an example with a single barrier potential. The d-band **k**·**p**-model is introduced as most general MEMA. We derive DTBCs for the d-band **k**·**p**-model and test our results on a quantum well nanostructure.

D. Klindworth (✉)
Technische Universität Berlin, DFG Research Center MATHEON, Strasse des 17. Juni 136, 10623 Berlin, Germany
e-mail: klindworth@math.tu-berlin.de

M. Ehrhardt
Bergische Universität Wuppertal, Fachbereich C – Mathematik und Naturwissenschaften, Lehrstuhl für Angewandte Mathematik und Numerische Analysis, Gaußstrasse 20, 42119 Wuppertal, Germany
e-mail: ehrhardt@math.uni-wuppertal.de

T. Koprucki
Forschungsgruppe "Partielle Differentialgleichungen", Weierstraß–Institut für Angewandte Analysis und Stochastik, Mohrenstrasse 39, 10117 Berlin, Germany
e-mail: koprucki@wias-berlin.de

M. Ehrhardt and T. Koprucki (eds.), *Multi-Band Effective Mass Approximations*, Lecture Notes in Computational Science and Engineering 94,
DOI 10.1007/978-3-319-01427-2_8,

8.1 Introduction

One of the main difficulties when solving the Schrödinger partial differential equation (PDE), particularly from a numerical point of view, is to supply physically admissible *boundary conditions* to solve numerically a bounded domain equation modelling an equation originally posed on an unbounded domain. That is in order to compute a numerical solution to the Schrödinger equation one has to require a finite computational domain. Usually, this is done by introducing *artificial boundary conditions*.

If the solution of the unbounded domain restricted to the computational domain equals the approximate solution when using the artificial boundary conditions, then these boundary conditions are called *transparent boundary conditions* (TBCs). Let us note that a new approach for designing TBCs for stationary Schrödinger equations with general (linear and nonlinear) exterior potentials $V(x)$ was introduced recently by Klein et al. [2, 14]. This work is a generalization of the well-known *quantum transmitting boundary condition* of Lent and Kirkner [17] to the case of a space-dependent potential.

TBCs of time-dependent Schrödinger equations have been discussed extensively, see for example the concise review by Antoine et al. [1]. It has been shown that a fully discrete approach in deriving these TBCs, yielding so-called *discrete transparent boundary conditions* (DTBCs), implies significant numerical advantages.

On the other hand, an ad-hoc discretization of the continuous TBCs can result in unphysical reflections at the artificial boundaries and may also destroy the stability of the underlying numerical scheme [4]. Moreover, this discrete approach was successfully applied to general Schrödinger-type equations [3, 11, 12].

DTBCs (based on finite difference discretizations) for systems of time-dependent Schrödinger equations were developed by Zisowsky et al. [28, 29] (cf. also the previous Chap. 7). For stationary Schrödinger equations, however, DTBCs have been developed by Arnold [5] only for the scalar case. Recently, novel DTBCs based on a finite element discretization for the time-dependent scalar Schrödinger equation in 1D were proposed by A. Zlotnik and I. Zlotnik [30, 31].

In this chapter the numerical discretization of the differential equations in the interior computational domain is done using finite differences schemes. For an alternative approach using the finite element method the reader is referred e.g. to the work of Negulescu [23] and the references therein.

In this chapter we will directly consider the general d-band $\mathbf{k}\cdot\mathbf{p}$-model. An analysis of particular multi-band effective mass approximations (MEMAs) such as the two-band Kane-model and two-band $\mathbf{k}\cdot\mathbf{p}$-model can be found in [8] and [15].

This chapter is organized as follows: In Sect. 8.2 we will summarize the results of the stationary scalar case and extend it to some alternative discretizations. We will compare the dispersion relations of these schemes and their numerical results when applied to a single barrier potential. After that, we will derive DTBCs for the general d-band $\mathbf{k}\cdot\mathbf{p}$-model in Sect. 8.3. The DTBCs are tested using a quantum well structure. Finally, we will summarize our work in Sect. 8.4 and discuss future research directions.

8.2 Single-Band Effective Mass Approximations: The Scalar Schrödinger Equation

We start considering the stationary linear Schrödinger equation for the wave function $\psi(x) \in \mathbb{C}$

$$H\psi = E\psi, \qquad x \in \mathbb{R}, \tag{8.1}$$

where E denotes the energy of the electron and H is the Hamiltonian operator

$$H = -\frac{\hbar^2}{2m^*}\frac{\mathrm{d}^2}{\mathrm{d}x^2} + V(x), \tag{8.2}$$

with the reduced Planck constant $\hbar$, the effective mass m^* of the electron and the real-valued potential energy profile $V(x)$ of the electron at the position $x \in \mathbb{R}$.

A solution $\psi_E(x)$ of the stationary linear Schrödinger equation (8.1) is called an *energy eigenstate* with associated energy E.

Next, we consider a semiconductor of length L connected to reservoirs at $x = 0$ and $x = L$. Let us assume that the potential $V(x)$ is constant in the reservoirs, i.e. we set $V(x) = 0$ if $x \leq 0$ and $V(x) = V_L$ if $x \geq L$. Note that the assumption $V(x) = 0$ for $x \leq 0$ means no loss of generality since we are free to set the energetic zero point. Similarly, the assumption that the left boundary is located at $x = 0$ is no loss of generality.

8.2.1 The Exterior Problem and the Quantum Mechanical Dispersion Relation

The *exterior problem* is concerned with the solution of the Schrödinger equation in the exterior domains. By assumption, the potential V is constant in these domains and hence, (8.1) becomes a linear second order ordinary differential equation (ODE) with constant coefficients:

$$-\frac{\hbar^2}{2m^*}\frac{\mathrm{d}^2}{\mathrm{d}x^2}\psi = (E - V)\psi, \qquad x \in \mathbb{R}. \tag{8.3}$$

The solution of (8.3) takes the form

$$\psi(x) = \hat{\psi}\mathrm{e}^{\mathrm{i}kx}, \tag{8.4}$$

where $\hat{\psi} \in \mathbb{C}$ is an arbitrary constant and the complex *wave vector* $k = \hat{k} + \mathrm{i}\check{k}$ is the root of the characteristic polynomial

$$\frac{\hbar^2}{2m^*}k^2 = E - V.$$

If we assume that the energy satisfies $E > V$, the wave vector k is real and reads

$$k = \pm\hat{k} = \pm\sqrt{\frac{2m^*}{\hbar^2}(E - V)}. \tag{8.5}$$

Note that the resulting waves of the form (8.4) are traveling. In classical physics the energy condition $E > V$ is always fulfilled since there can only exist particles with an energy greater than the potential at that point. However, in quantum physics this is not the case and therefore, we shall, in general, not require $E > V$ inside the computational domain.

As can be seen from the *quantum mechanical momentum operator* in one dimension, $p = -\mathrm{i}\hbar\frac{\partial}{\partial x}$, the expectation value of the momentum p of the wave ψ of amplitude 1 is $\langle p\psi, \psi\rangle = \hbar k\langle\psi, \psi\rangle = \hbar k$. Hence, the expectation value of the momentum p is proportional to the wave vector k. This means for this stationary equation (8.3) that a positive wave vector corresponds to a positive momentum, i.e. a right-traveling wave, while a negative wave vector corresponds to a negative momentum, i.e. a left-traveling wave.

If the energy does not satisfy the condition $E > V$, the wave vector k is purely imaginary and takes the form

$$k_{1,2} = \pm\mathrm{i}\check{k} = \pm\sqrt{\frac{2m^*}{\hbar^2}(V - E)}, \tag{8.6}$$

i.e. this wave vector yields evanescent waves (8.4).

If we insert the solution (8.4) to (8.3) we obtain

$$\frac{\hbar^2 k^2}{2m^*}\hat{\psi} = (E - V)\hat{\psi}.$$

Since we can neglect the trivial solution $\psi \equiv 0$, the energy E satisfies the *quantum mechanical dispersion relation*

$$E = E(k) = V + \frac{\hbar^2 k^2}{2m^*}. \tag{8.7}$$

8.2.2 Transparent Boundary Conditions

In order to transform the Schrödinger equation (8.1) on the real line $x \in \mathbb{R}$ into an equivalent system posed on the bounded domain $(0, L)$ we introduce artificial boundary conditions at $x = 0$ and $x = L$. Artificial boundary conditions that form a system whose solution equals the solution of the unbounded problem on the domain $(0, L)$ are called *transparent boundary conditions* (TBCs).

In order to derive these TBCs we consider a plain wave of amplitude 1 with positive momentum coming from $-\infty$ and entering the computational domain from the left at $x = 0$

$$\psi^{\mathrm{in}} = \mathrm{e}^{\mathrm{i}\hat{k}_0 x}, \qquad x < 0, \tag{8.8}$$

where $\hat{k}_0 > 0$ denotes the propagation coefficient, cf. (8.5) of the wave vector k in the left exterior domain $x \leq 0$, i.e. with $V \equiv 0$. The incoming wave (8.8) results in a reflected, left-traveling wave

$$\psi^{\mathrm{r}} = r\,\mathrm{e}^{-\mathrm{i}\hat{k}_0 x}, \qquad x < 0, \tag{8.9a}$$

with the *reflection coefficient* r, and a transmitted, right-traveling wave

$$\psi^{\mathrm{t}} = t\,\mathrm{e}^{\mathrm{i}\hat{k}_L x}, \qquad x > L, \tag{8.9b}$$

with the *transmission coefficient* t and the propagation coefficient $\check{k}_L > 0$ of the wave vector k in the right exterior domain $x \geq L$, i.e. with $V \equiv V_L$. The propagation coefficient $\hat{k}_L$ satisfies

$$\hat{k}_L = \sqrt{\hat{k}_0^2 - \frac{2m^* V_L}{\hbar^2}}.$$

Thus, the solution in the left exterior domain has the form

$$\psi = \psi^{\mathrm{in}} + \psi^{\mathrm{r}}, \qquad x < 0, \tag{8.10a}$$

and the solution in the right exterior domain is

$$\psi = \psi^{\mathrm{t}}, \qquad x > L. \tag{8.10b}$$

We know that the wave and its first derivative are continuous at the two boundaries, cf. [16]. Hence, we can eliminate the reflection and transmission coefficients by comparing (8.10a) and its first derivative at $x = 0$ as well as (8.10b) and its first derivative at $x = L$.

The resulting *boundary value problem* (BVP) reads

$$-\frac{\hbar^2}{2m^*}\psi_{xx} + V(x)\psi = E\psi, \qquad 0 < x < L, \tag{8.11a}$$

$$\psi_x(0) + \mathrm{i}k\psi(0) = 2\mathrm{i}\hat{k}_0, \tag{8.11b}$$

$$\psi_x(L) - \mathrm{i}\sqrt{\hat{k}_0^2 - \frac{2m^* V_L}{\hbar^2}}\psi(L) = 0. \tag{8.11c}$$

Theorem 8.1 (Proposition 2.3 in [7]). *Let V be in $L^\infty(0, L)$ and real valued. Then the BVP (8.11) has a unique solution $\psi \in W^{2,\infty}(0, L)$.*

8.2.3 The Standard Discretization

After stating the BVP (8.11) and recalling that it has a unique solution, we want to propose techniques to solve it numerically and compute eigenstates for corresponding energies and potentials. First let us set $\hbar = m^* = 1$ for the remainder of this chapter. We will introduce *finite difference schemes* (FDS) to solve the BVP, using for simplicity the uniform discretization $x_j = jh$, $j = 0, \ldots, J$ with $L = jh$, of the computational domain $(0, L)$ and the approximation $f(x_j) \approx f_j$ of some function f defined in $(0, L)$. In the sequel, we will use the following finite difference quotient operators: the first order forward operator $\mathrm{D}_h^{\mathrm{fwd}} f_j := (f_{j+1} - f_j)/h$, the first order backward operator $\mathrm{D}_h^{\mathrm{bwd}} f_j := (f_j - f_{j-1})/h$, the second order centered operator $\mathrm{D}_h^{\mathrm{cen}} f_j := (f_{j+1} - f_{j-1})/2h$ and the standard second order operator $\mathrm{D}_h^{\mathrm{std}} f_j := (f_{j-1} - 2f_j + f_{j+1})/h^2$.

By applying the standard second order finite difference quotient operator we get the *second order standard FDS*

$$-\frac{1}{2}\mathrm{D}_h^{\mathrm{std}}\psi_j + V_j\psi_j = E\psi_j, \qquad j = 1, \ldots, J-1, \tag{8.12}$$

for the Schrödinger equation (8.11) with $V_j = V(x_j)$ and the approximation $\psi_j \approx \psi(x_j)$, $j = 0, \ldots, J$. We can rewrite (8.12) in the form

$$-\psi_{j+1} + 2\big(1 - (E - V_j)h^2\big)\psi_j - \psi_{j-1} = 0, \quad j = 1, \ldots, J-1. \tag{8.13}$$

This is a linear second order homogeneous ordinary difference equation (OΔE) with a spatially varying coefficient V_j.

Now let us analyze the *discrete exterior problem* of the standard FDS. We will continue with some constant potential V. The results of the exterior domains $x \le 0$ and $x \ge L$ can later be derived by inserting the respective value of the potential in the results stated.

If V is constant, (8.13) is a linear second order OΔE with constant coefficients and as shown in [19], (8.13) has a solution of the form

$$\psi_j = \hat{\psi}_h \alpha^j = \hat{\psi}_h \mathrm{e}^{\ln(\alpha) j} = \hat{\psi}_h \mathrm{e}^{\left(\ln|\alpha| + \mathrm{i}\arg(\alpha)\right) j} = \hat{\psi}_h \mathrm{e}^{\mathrm{i}k_h jh}, \tag{8.14}$$

with $\alpha \in \mathbb{C}$, cf. (8.4). We will call $\hat{\psi}_h$ the *discrete amplitude* of the discrete wave ψ_j and

$$k_h = -\mathrm{i}\frac{1}{h}\ln(\alpha) = \frac{1}{h}\big(\arg(\alpha) - \mathrm{i}\ln|\alpha|\big) \tag{8.15}$$

the *discrete wave vector*. By inserting (8.14) into (8.13) we get

$$\alpha^{j-1}\big(\alpha^2 - 2\big(1-(E-V)h^2\big)\alpha + 1\big) = 0.$$

Since we neglect the trivial solution $\alpha = 0$, we have

$$\alpha^2 - 2\big(1-(E-V)h^2\big)\alpha + 1 = 0,$$

which implies

$$\big(\alpha - \big(1-(E-V)h^2\big)\big)^2 = (E-V)h^2\big((E-V)h^2 - 2\big). \tag{8.16}$$

If the energy E satisfies $E > V$, the right hand side of (8.16) is negative if the step size h satisfies

$$h < \sqrt{\frac{2}{E-V}}, \tag{8.17}$$

and hence, the roots of (8.16) are complex and read

$$\alpha_{1,2} = 1-(E-V)h^2 \pm \mathrm{i}\sqrt{(E-V)h^2\big(2-(E-V)h^2\big)}. \tag{8.18}$$

We find that (since α_1, α_2 are complex conjugate by Vieta)

$$|\alpha_{1,2}|^2 = \big(1-(E-V)h^2\big)^2 + (E-V)h^2\big(2-(E-V)h^2\big) = 1,$$

and thus, the discrete wave vector k_h is real and takes the form

$$k_h = \pm\hat{k}_h = \pm\frac{1}{h}\arg(\alpha) = \pm\frac{1}{h}\arccos\frac{\mathrm{Re}\alpha}{|\alpha|} = \pm\frac{1}{h}\arccos\big(1-(E-V)h^2\big), \tag{8.19}$$

cf. (8.5). Note that we can neglect to add the term $n\frac{2\pi}{h}$, $n \in \mathbb{Z}$, to this formula since for any $n \neq 0$ this term diverges for $h \to 0$. On the other hand, we will see later that the discrete wave vector k_h as given in (8.19), i.e. with $n = 0$, tends to the analytical wave vector k for $h \to 0$.

The wave vector $k_h = \hat{k}_h$ corresponds to

$$\alpha_1 = 1-(E-V)h^2 + \mathrm{i}\sqrt{(E-V)h^2\big(2-(E-V)h^2\big)},$$

while $k_h = -\hat{k}_h$ is associated with

$$\alpha_2 = 1-(E-V)h^2 - \mathrm{i}\sqrt{(E-V)h^2\big(2-(E-V)h^2\big)}.$$

Hence, we get two traveling waves, $\psi_j = \alpha_1^j$ being right-traveling and $\psi_j = \alpha_2^j$ being left-traveling.

The case

$$h \geq \sqrt{\frac{2}{E-V}},$$

results in a non-negative right hand side of (8.16) and thus, a complex conjugate pair of purely imaginary wave vectors that give evanescent waves. However, this case is numerically not applicable since it defines a lower bound for the step size h.

On the other hand, if $E \leq V$, the right hand side of (8.16) is also non-negative, yielding a complex conjugate pair of purely imaginary wave vectors that give evanescent waves. Let us recall the TBCs we derived in Sect. 8.2.2. We considered an incoming wave, i.e. a traveling wave to enter the semiconductor at $x = 0$. Hence, the case $E \leq V$ is not applicable either for the exterior domains since it leads to evanescent waves only.

Now let us analyze the behavior of the discrete wave vector k_h of the standard FDS for $h \to 0$. To this end, we apply l'Hôpital's rule to get

$$\begin{aligned}\lim_{h\to 0} k_h &= \pm \lim_{h\to 0} \hat{k}_h = \pm \lim_{h\to 0} \frac{1}{h} \arccos\bigl(1-(E-V)h^2\bigr) \\ &= \pm \lim_{h\to 0} \frac{-2(E-V)h}{\sqrt{1-\bigl(1-(E-V)h^2\bigr)^2}} = \pm\sqrt{2(E-V)},\end{aligned}$$

which equals the analytical wave vector (8.5) for $\hbar = m^* = 1$.

Note, that (8.19) defines the *discrete dispersion relation*. Finally, we will state this relation in the reciprocal form. Recall the wave representation of the form $\psi_j = \hat{\psi} \mathrm{e}^{\mathrm{i}k_h jh}$ that implies $\psi_{j+1}\mathrm{e}^{-\mathrm{i}kh} = \psi_j = \psi_{j-1}\mathrm{e}^{\mathrm{i}kh}$. Applied to the O$\Delta$E (8.13), this gives

$$-\mathrm{e}^{\mathrm{i}kh} + 2(1-(E-V)h^2) - \mathrm{e}^{-\mathrm{i}kh} = 0,$$

which leads to

$$E = E_h^{\mathrm{std}}(k_h) = V + \frac{1-\cos k_h h}{h^2} = V + \frac{2}{h^2}\sin^2\frac{k_h h}{2} = V + \frac{k_h^2 h}{2} + \mathcal{O}(h^2), \tag{8.20}$$

compared to the continuous dispersion relation (8.7).

8.2.4 Discretization of the Transparent Boundary Conditions

Let us now introduce a finite difference discretization of the two Robin-type TBCs (8.11b) and (8.11c). We apply the second order centered difference operator $\mathrm{D}_h^{\mathrm{cen}}$ to

ψ_0 at the left boundary and ψ_J at the right boundary. Let $\hat{k}_0$ denote the analytical propagation coefficient of a right-traveling wave in the left exterior domain $x \leq 0$. Then the analytical propagation coefficient of a right-traveling wave in the right exterior domain is

$$\hat{k}_L = \sqrt{\hat{k}_0^2 - 2V_L}.$$

At the left boundary we have

$$\frac{\psi_1 - \psi_{-1}}{2h} + \mathrm{i}\hat{k}_0\psi_0 = 2\mathrm{i}\hat{k}_0,$$

which implies

$$-\psi_{-1} + 2\mathrm{i}\hat{k}_0 h\psi_0 + \psi_1 = 4\mathrm{i}\hat{k}_0 h. \tag{8.21a}$$

On the other hand, discretizing the right TBC gives

$$\frac{\psi_{J+1} - \psi_{J-1}}{2h} = \mathrm{i}\sqrt{\hat{k}_0^2 - 2V_L}\psi_J,$$

that can be expressed in the form

$$\psi_{J-1} + 2\mathrm{i}\sqrt{\hat{k}_0^2 - 2V_L}h\psi_J - \psi_{J+1} = 0. \tag{8.21b}$$

The two ghost points ψ_{-1} and ψ_{J+1} in (8.21) can be eliminated by subtracting the FDS (8.13) of the Schrödinger equation (8.11a) at $j = 0$ and $j = J$. By using the identities $E = \hat{k}_0^2/2$ and $E - V_L = \hat{k}_L^2/2$ we get the two second order discretized TBCs

$$\Big(\frac{1}{2}\hat{k}_0^2 h^2 - 1 + \mathrm{i}\hat{k}_0 h\Big)\psi_0 + \psi_1 = 2\mathrm{i}\hat{k}_0 h \tag{8.22a}$$

and

$$\psi_{J-1} + \Big(\frac{1}{2}(\hat{k}_0^2 - 2V_L)h^2 - 1 + \mathrm{i}\sqrt{\hat{k}_0^2 - 2V_L}h\Big)\psi_J = 0. \tag{8.22b}$$

Existence and uniqueness of the solution of the numerical scheme together with the discretized TBCs is shown in

Theorem 8.2 ([5, Theorem 2.1]). *Let* $\{V_j\}$, $j = 0, \ldots, J$, *and* $E > \max\{0, V_L\}$ *be given, and assume* $h < \min\{\sqrt{\frac{2}{E}}, \sqrt{\frac{2}{E-V_L}}\}$. *Then the discrete BVP* (8.13) *with the discretized TBCs* (8.22) *has a unique solution* $\{\psi_j\}$, $j = 0, \ldots, J$.

Let us recall the discrete waves $\psi_j = \mathrm{e}^{\mathrm{i}\hat{k}_{h,0}jh}$ in the left exterior domain $j \leq 0$ and $\psi_j = \mathrm{e}^{\mathrm{i}\hat{k}_{h,L}jh}$ in the right exterior domain $j \geq J$. Suppose that they are solutions to the difference scheme in a small vicinity of the two boundaries, i.e. $j = 0, 1$ and $j = J-1, J$. Then they should also satisfy the discretized TBCs (8.22). However, at the left boundary we get

$$\begin{aligned} 2\mathrm{i}\hat{k}_0 h &= \left(Eh^2 - 1 + \mathrm{i}\hat{k}_0 h\right)\psi_0 + \psi_1 = \left(Eh^2 - 1 + \mathrm{i}\hat{k}_0 h\right) + \mathrm{e}^{\mathrm{i}\hat{k}_{h,0}h} \\ &= \left(Eh^2 - 1 + \mathrm{i}\hat{k}_0 h\right) + \left(1 - Eh^2 + \mathrm{i}\sqrt{2Eh^2 - \mathrm{e}^2 h^4}\right) \\ &= \mathrm{i}\hat{k}_0 h + \mathrm{i}\hat{k}_0 h\sqrt{1 - \hat{k}_0^2 h^2/4} \neq 2\mathrm{i}\hat{k}_0 h, \end{aligned}$$

which is a contradiction. An analogue contradiction can be found at the right boundary.

The reason is that the TBCs are based on the analytical solution as derived in Sect. 8.2.1. The wave vector of the analytical solution, see (8.5), however, is different from the discrete wave vector (8.19). Hence, the discretized TBCs model exterior domains whose physical properties (i.e. wave vector and dispersion relation) are discretization of the analytical properties. Inside the computational domain, however, we use the FDS (8.13) that implies a discrete wave vector and a discrete dispersion relation. In other words, a wave coming from $-\infty$ and entering the semiconductor at $x = 0$ is *refracted* at the boundary $x = 0$ as it comes from a media with the analytical dispersion relation and enters a media with the discrete dispersion relation. This leads to spurious oscillations in the numerical solution.

8.2.5 Discrete Transparent Boundary Conditions

In this section we will derive the *discrete transparent boundary conditions* (DTBCs) of the single-band model. DTBCs are derived on a fully discrete level, i.e. they are deduced with the help the discrete exterior solution (8.14). We assume that the discrete exterior solution holds in a small vicinity of the two boundaries. Consequently, the refraction at the boundaries, resulting in spurious oscillations, vanishes completely.

Let us recall the right-traveling discrete wave $\psi_j = \mathrm{e}^{\mathrm{i}\hat{k}_h jh}$ and the left-traveling discrete wave $\psi_j = \mathrm{e}^{-\mathrm{i}\hat{k}_h jh}$ with discrete amplitude $\hat{\psi}_h = 1$. Let $\hat{k}_{h,0}$ denote the discrete wave vector in the left exterior domain $x \leq 0$, i.e. with $V \equiv 0$, and $\hat{k}_{h,L}$ the discrete wave vector in the right exterior domain $x \geq L$, i.e. with $V \equiv V_L$. We apply these discrete waves to the reflection and transmission conditions (8.10) and consider that they hold in a small vicinity of the two boundaries, i.e. $j = 0, 1$ and $j = J-1, J$ respectively. It yields

$$\psi_j = \psi_j^{\text{in}} + \psi_j^{\text{r}} = \mathrm{e}^{\mathrm{i}\hat{k}_{h,0}x_j} + r\,\mathrm{e}^{-\mathrm{i}\hat{k}_{h,0}x_j}, \qquad j = 0, 1,$$

and

$$\psi_j = \psi_j^{\text{t}} = t\,\mathrm{e}^{\mathrm{i}\hat{k}_{h,L}x_j}, \qquad j = J-1, J.$$

By eliminating the reflection and transmission coefficients r, t we obtain the DTBCs

$$-\psi_0 \mathrm{e}^{-\mathrm{i}\hat{k}_{h,0}h} + \psi_1 = 2\mathrm{i}\sin\hat{k}_{h,0}h, \tag{8.23a}$$

and

$$\psi_{J-1}\mathrm{e}^{\mathrm{i}\hat{k}_{h,L}h} - \psi_J = 0, \tag{8.23b}$$

cf. the TBCs (8.11b) and (8.11c).

Let us recall the discretized TBCs (8.22). We expand the exponential function and the sine function in the left DTBC (8.23a). Keeping terms up to second order gives

$$-\big(1 - \mathrm{i}\hat{k}_{h,0}h - \frac{1}{2}\hat{k}_{h,0}^2h^2\big)\psi_0 + \psi_1 = 2\mathrm{i}\hat{k}_{h,0}h.$$

If we replace the discrete wave vector $\hat{k}_{h,0}$ by the analytical wave vector $\hat{k}_0$ the above equation becomes

$$\big(\frac{1}{2}\hat{k}_0^2h^2 - 1 + \mathrm{i}\hat{k}_0h\big)\psi_0 + \psi_1 = 2\mathrm{i}\hat{k}_0h,$$

which equals the left discretized TBC. Similarly, we can deduce the right discretized TBC from the right DTBC.

Now we can reformulate Theorem 8.2 for the DTBCs.

Theorem 8.3 ([5, Theorem 2.1]). *Let* $\{V_j\}$, $j = 0, \ldots, J$, *and* $E > \max\{0, V_L\}$ *be given, and assume* $h < \min\big\{\sqrt{\frac{2}{E}}, \sqrt{\frac{2}{E-V_L}}, \frac{\pi}{\hat{k}_{h,0}}, \frac{\pi}{\hat{k}_{h,L}}\big\}$. *Then the discrete BVP* (8.13) *with the DTBCs* (8.23) *has a unique solution* $\{\psi_j\}$, $j = 0, \ldots, J$.

8.2.6 *Alternative Finite Difference Schemes*

In this section we derive and compare alternative FDSs. Our aim is to improve the convergence of the scheme or to develop schemes that solve the problem exactly if certain conditions are fulfilled.

8.2.6.1 The Numerov Discretization

We start with the so-called *Numerov discretization* [22] that is of higher order than the standard discretization.

Let us consider the Schrödinger equation of the BVP (8.11), and let us rewrite it in the form

$$\psi_{xx} = -2\big(E - V(x)\big)\psi, \qquad 0 < x < L. \tag{8.24}$$

As before, we use the uniform grid $x_j = jh,\ j = 0, \ldots, J$ with $L = jh$. From (8.24) together with the standard second order finite difference operator $\mathrm{D}_h^{\mathrm{std}}$ we find that

$$\begin{aligned} h^2\psi_j^{(iv)} &= h^2 \frac{\mathrm{d}^2}{\mathrm{d}x^2}\psi_{xx}(x)\big|_{x=x_j} \\ &= h^2\Big(-2\frac{\mathrm{d}^2}{\mathrm{d}x^2}\big((E - V(x))\psi\big)\big|_{x=x_j}\Big) \\ &= h^2\big(-2\mathrm{D}_h^{\mathrm{std}}\big((E - V_j)\psi_j\big) + \mathscr{O}(h^2)\big) \\ &= -2(E - V_{j+1})\psi_{j+1} + 4(E - V_j)\psi_j - 2(E - V_{j-1})\psi_{j-1} + \mathscr{O}(h^4). \end{aligned} \tag{8.25}$$

On the other hand, the Taylor series

$$\begin{aligned} &\psi(x \pm h) \\ &= \psi(x) \pm h\psi_x(x) + \frac{h^2}{2}\psi_{xx}(x) \pm \frac{h^3}{6}\psi_{xxx}(x) + \frac{h^4}{24}\psi^{(iv)}(x) \pm \frac{h^5}{96}\psi^{(v)}(x) + \mathscr{O}(h^6) \end{aligned}$$

gives

$$\psi(x + h) + \psi(x - h) = 2\psi(x) + h^2\psi_{xx}(x) + \frac{h^4}{12}\psi^{(iv)}(x) + \mathscr{O}(h^6),$$

which implies

$$h^2\psi_j^{(iv)} = \frac{12}{h^2}\left(\psi_{j+1} - 2\psi_j + \psi_{j-1}\right) - 12\ \psi_{xx}(x)|_{x=x_j} + \mathscr{O}(h^4).$$

If we apply (8.24) to the above equation we get

$$h^2\psi_j^{(iv)} = \frac{12}{h^2}\big(\psi_{j+1} - 2\big(1 - (E - V_j)h^2\big)\psi_j + \psi_{j-1}\big) + \mathscr{O}(h^4). \tag{8.26}$$

A comparison of (8.25) and (8.26) gives the Numerov FDS

$$\Big(1+\frac{h^2}{6}(E-V_{j+1})\Big)\psi_{j+1}-2\Big(1-\frac{5h^2}{6}(E-V_j)\Big)\psi_j + \Big(1+\frac{h^2}{6}(E-V_{j-1})\Big)\psi_{j-1}=0, \quad j=1,\dots,J-1. \tag{8.27}$$

The Numerov FDS is of fourth order if $\psi \in C^6(0,L)$ compared to second order accuracy of the standard FDS if $\psi \in C^4(0,L)$.

Now we will examine the discrete exterior problem of the Numerov FDS with V constant. We want to determine a solution of the discrete exterior problem in order to apply the DTBCs (8.23).

If V is constant (8.27) is a linear second order OΔE with constant coefficients whose solution takes the form

$$\psi_j = \hat{\psi}_h\,\alpha^j = \hat{\psi}_h\,\mathrm{e}^{\mathrm{i}k_h jh}, \tag{8.28}$$

with $\alpha \in \mathbb{C}$. Again we will refer to $\hat{\psi}_h$ as the *discrete amplitude* of the discrete wave ψ_j and

$$k_h = -\mathrm{i}\frac{1}{h}\ln(\alpha) = \frac{1}{h}(\arg(\alpha)-\mathrm{i}\ln|\alpha|) \tag{8.29}$$

as the *discrete wave vector*. Analogously to the standard FDS we get the discrete solution α by applying (8.28) to (8.27). Under the assumption that $E > V$ and the step size h satisfies

$$h < \frac{3}{\sqrt{E-V}}, \tag{8.30}$$

α is complex and reads

$$\alpha_{1,2} = 1-\frac{6(E-V)h^2}{6+(E-V)h^2} \pm \mathrm{i}\frac{\sqrt{24(E-V)h^2(3-(E-V)h^2)}}{6+(E-V)h^2}. \tag{8.31}$$

The modulus of α is

$$|\alpha_{1,2}|^2 = \Big(1-\frac{6(E-V)h^2}{6+(E-V)h^2}\Big)^2 + \frac{24(E-V)h^2\big(3-(E-V)h^2\big)}{\big(6+(E-V)h^2\big)^2} = 1.$$

Thus, the discrete wave vector k_h is real and takes the form

$$k_h = \pm\hat{k}_h = \pm\frac{1}{h}\arg(\alpha) = \pm\frac{1}{h}\arccos\frac{\mathrm{Re}\alpha}{|\alpha|} = \pm\frac{1}{h}\arccos\Big(1 - \frac{6(E-V)h^2}{6+(E-V)h^2}\Big), \tag{8.32}$$

Again we can neglect to add the term $n\frac{2\pi}{h}$, $n \in \mathbb{Z}$, to this formula since for any $n \neq 0$ this term diverges for $h \to 0$. On the other hand, it is easy to show that the discrete wave vector k_h as given in (8.32), i.e. with $n = 0$, tends to the analytical wave vector k as given in (8.5) for $h \to 0$ and $\hbar = m^* = 1$.

Hence, we get two traveling waves, the right-traveling wave $\psi_j = \alpha_1^j = \mathrm{e}^{\mathrm{i}\hat{k}_h jh}$ and the left-traveling wave $\psi_j = \alpha_2^j = \mathrm{e}^{-\mathrm{i}\hat{k}_h jh}$.

On the other hand, if the step size h does not satisfy the step size restriction (8.30) or the energy E does not satisfy the energy condition $E > V$, α is real and yields evanescent waves.

Finally, we derive the *discrete dispersion relation* of the Numerov FDS in the same way as for the standard FDS. The wave $\psi_j = \hat{\psi}\mathrm{e}^{\mathrm{i}k_h jh}$ implies $\psi_{j+1}\mathrm{e}^{-\mathrm{i}k_h h} = \psi_j = \psi_{j-1}\mathrm{e}^{\mathrm{i}k_h h}$, and hence, applied to the Numerov difference equation (8.27) we get

$$E = E_h^{\mathrm{Num}}(k) = \frac{6}{5+\cos k_h h}\frac{2}{h^2}\sin^2\frac{k_h h}{2}. \tag{8.33}$$

Now we will prove an analogon to Theorem 8.3.

Theorem 8.4. *Let* $\{V_j\}$, $j = 0, \ldots, J$, *and* $E > \max\{0, V_L\}$ *be given, and suppose*

$$h < \min\Big\{\sqrt{\frac{3}{E}}, \sqrt{\frac{3}{E-V_L}}, \sqrt{\frac{6}{|E-V_1|}}, \sqrt{\frac{6}{|E-V_{J-1}|}}, \frac{\pi}{\hat{k}_{h,0}}, \frac{\pi}{\hat{k}_{h,L}}\Big\}.$$

Then the discrete BVP (8.27) *of the Numerov FDS with the DTBCs* (8.23) *has a unique solution* $\{\psi_j\}$, $j = 0, \ldots, J$.

Proof. We show that for homogeneous DTBCs the discrete solution is zero at every grid point. Therefore, let us introduce $\varphi_j = \sigma_j\psi_j$ with $\sigma_j = 1 + \frac{h^2}{6}(E - V_j) \in \mathbb{R}$. Note that $\sigma_j > 0$ for $j = 0, 1, J-1, J$.

Now we can rewrite the Numerov FDS in the form

$$\varphi_{j+1} - 2\varphi_j + \varphi_{j-1} = -2h^2(E-V_j)\psi_j = -2h^2(E-V_j)\sigma_j^{-1}\varphi_j,$$

or

$$\mathrm{D}_h^{\mathrm{std}}\varphi_j = -2(E-V_j)\sigma_j^{-1}\varphi_j.$$

The homogeneous left DTBC can be written in the form

$$-\gamma_1\psi_0 + \psi_1 = -\gamma_1\sigma_0^{-1}\varphi_0 + \sigma_1^{-1}\varphi_1,$$

that reduces to

$$\mathrm{D}_h^{\mathrm{bwd}}\varphi_1 = \left(\gamma_1\sigma_1\sigma_0^{-1} - 1\right)\varphi_0,$$

with $\gamma_1 = \mathrm{e}^{-\mathrm{i}\hat{k}_{h,0}h}$. On the other hand, the right DTBC is

$$\psi_{J-1} - \gamma_2\psi_J = \sigma_{J-1}^{-1}\varphi_{J-1} - \gamma_2\sigma_J^{-1}\varphi_J,$$

which becomes

$$\mathrm{D}_h^{\mathrm{bwd}}\varphi_J = \left(\gamma_2^{-1}\sigma_J\sigma_{J-1}^{-1} - 1\right)\varphi_{J-1},$$

with $\gamma_2 = \mathrm{e}^{-\mathrm{i}\hat{k}_{h,L}h}$.

We multiply the sum of the Numerov FDS for $j = 1, \ldots, J-1$ by $\bar{\varphi}_j$, apply the discrete analogon of the integration by parts rule and take the imaginary part to get

$$0 = -\sigma_1\sigma_0^{-1}\mathrm{Im}\gamma_1\,|\varphi_0|^2 + \sigma_J\sigma_{J-1}^{-1}\mathrm{Im}\gamma_2^{-1}\,|\varphi_{J-1}|^2\,,$$

with

$$\mathrm{Im}\gamma_1 = -\sin\hat{k}_{h,0}h < 0, \qquad h < \frac{\pi}{\hat{k}_{h,0}},$$

and

$$\mathrm{Im}\gamma_2^{-1} = -\frac{1}{|\gamma_2|^2}\mathrm{Im}\gamma_2 = \sin\hat{k}_{h,0}h > 0, \qquad h < \frac{\pi}{\hat{k}_{h,L}}.$$

Hence, we end up with the equation

$$0 = \sigma_1\sigma_0^{-1}\sin\hat{k}_{h,0}h|\varphi_0|^2 + \sigma_J\sigma_{J-1}\sin\hat{k}_{h,L}h|\varphi_{J-1}|^2.$$

Since $\sigma_j > 0$ for $j = 0, 1, J-1, J$, the above equation implies $|\varphi_0|^2 = |\varphi_{J-1}|^2 = 0$. Thus, $\psi_0 = \psi_{J-1} = 0$ and by using the homogeneous DTBCs we get $\psi_1 = 0$ and $\psi_J = 0$. Successively applying the Numerov FDS gives $\psi_j = 0$ for $j = 0, \ldots, J$ and hence, the discrete solution vanishes at every grid point if the DTBCs are homogeneous. Thus, the coefficient matrix of the system of linear equations formed by the discrete BVP and the two DTBCs is regular. This implies that the discrete solution is unique. □

8.2.6.2 The Mickens Discretization

We recall the Schrödinger equation (8.1)

$$\psi_{xx} + 2\left(E - V\left(x\right)\right)\psi = 0, \qquad 0 < x < L. \tag{8.34}$$

Let us assume that the potential $V\left(x\right) \equiv V$ is constant. As shown in Sect. 8.2.1 the Schrödinger equation (8.34) yields a traveling wave if $E > V$ and an evanescent wave otherwise. Hence, in the case $2(E - V) \equiv c_1 > 0$, (8.34) has the solution

$$\psi\left(x\right) = a_1 \cos\left(\sqrt{c_1}x\right) + b_1 \sin\left(\sqrt{c_1}x\right), \tag{8.35a}$$

and if $2(E - V) \equiv -c_2 < 0$, (8.34) has the solution

$$\psi\left(x\right) = a_2 \cosh\left(\sqrt{c_2}x\right) + b_2 \sinh\left(\sqrt{c_2}x\right), \tag{8.35b}$$

where $a_1, a_2, b_1, b_2 \in \mathbb{C}$ are arbitrary constants. The *Mickens nonstandard finite difference discretization*

$$\psi_{j+1} - 2\cos\left(h\sqrt{c_1}\right)\psi_j + \psi_{j-1} = 0, \qquad j = 1, \ldots, J-1, \tag{8.36a}$$

if $E > V$, and

$$\psi_{j+1} - 2\cosh\left(h\sqrt{c_2}\right)\psi_j + \psi_{j-1} = 0, \qquad j = 1, \ldots, J-1, \tag{8.36b}$$

if $E \leq V$, is a so-called *exact FDS* of (8.34) with the solutions (8.35), cf. [20]. An FDS is said to be *exact* if the numerical solution equals the analytical solution at the grid points, see [18]. It is easy to see that if V is constant the OΔE (8.36a) has the solution (8.35a) whereas the OΔE (8.36b) has the solution (8.35b). Therefore, the solution of the Mickens FDS equals the analytical solution at the grid points, in other words, the Mickens FDS is exact.

Now we allow the potential $V(x)$ to vary inside the computational domain $(0, L)$. The Mickens FDS (8.36) becomes

$$\psi_{j+1} - 2D_j\psi_j + \psi_{j-1} = 0, \qquad j = 1, \ldots, J-1, \tag{8.37}$$

with

$$D_j = \begin{cases} \cos\left(h\sqrt{2\left(E - V_j\right)}\right), & E > V_j, \\ \cosh\left(h\sqrt{2\left(V_j - E\right)}\right), & E \leq V_j. \end{cases} \tag{8.38}$$

While the Mickens FDS (8.37) is exact for a constant potential V it is formally of order $\mathcal{O}(h^2)$ if the potential V is not constant, cf. [10].

In order to use the DTBCs for the Mickens FDS (8.37) we have to determine a discrete solution of the Mickens FDS in the exterior domains. The Mickens FDS is exact in the exterior domains since the potential V is assumed to be constant in these domains. Hence, the discrete solution is given by the analytical solution as derived in Sect. 8.2.1. The discrete wave vector k_h of the Mickens FDS is equal to the analytical wave vector k as given in (8.5), and the discrete dispersion relation $E_h^{\mathrm{Mic}}(\hat{k})$ is equal to the analytical dispersion relation (8.7).

Theorem 8.5. *Let* $\{V_j\}$, $j = 0, \ldots, J$, *and* $E > \max\{0, V_L\}$ *be given, and assume* $h < \min\left\{\frac{\pi}{\hat{k}_{h,0}}, \frac{\pi}{\hat{k}_{h,L}}\right\}$. *Then the discrete BVP* (8.36) *of the Mickens FDS with the DTBCs* (8.23) *has a unique solution* $\{\psi_j\}$, $j = 0, \ldots, J$.

Proof. Let us rewrite the Mickens FDS in the form

$$\mathrm{D}_h^{\mathrm{std}}\psi_j = -\frac{2}{h^2}\left(1 - D_j\right)\psi_j.$$

Since $-\frac{2}{h^2}(1 - D_j) \in \mathbb{R}$ for all $j = 1, \ldots, J-1$, this theorem is obviously a direct corollary of Theorem 8.3 and hence, the solution is unique. □

8.2.6.3 The Numerov-Mickens Discretization

Chen et al. [10] combined the Numerov discretization with the Mickens discretization and proposed the so-called *combined Numerov-Mickens finite-difference scheme*

$$\left(1 + \frac{h^2}{6}(E - V_{j+1})\right)\psi_{j+1} - 2D_j\psi_j + \left(1 + \frac{h^2}{6}(E - V_{j-1})\right)\psi_{j-1} = 0, \tag{8.39}$$

for $j = 1, \ldots, J-1$, with D_j given in (8.38).

It can be shown that the Numerov-Mickens FDS (8.39) is of order $\mathcal{O}(h^4)$, just as the Numerov FDS, and that it is an exact FDS if the potential V is constant, just as the Mickens FDS, cf. [10].

Let us now study the discrete exterior problem of the Numerov-Mickens FDS with a constant potential V. If V is constant (8.39) is a linear second order OΔE with constant coefficients whose solution takes the form

$$\psi_j = \hat{\psi}_h \alpha^j = \hat{\psi}_h \mathrm{e}^{\mathrm{i}k_h jh}, \tag{8.40}$$

with $\alpha \in \mathbb{C}$, the *discrete amplitude* ψ_j and the *discrete wave vector*

$$k_h = -\mathrm{i}\frac{1}{h}\ln(\alpha) = \frac{1}{h}\big(\arg(\alpha) - \mathrm{i}\ln|\alpha|\big). \tag{8.41}$$

By applying (8.40) to (8.39) and under the assumption that the energy E satisfies $E > V$, α is complex and reads

$$\alpha_{1,2} = \frac{\cos\bigl(h\sqrt{2(E-V)}\bigr)}{1+h^2(E-V)/6} \pm \mathrm{i}\sqrt{1-\frac{\cos^2\bigl(h\sqrt{2(E-V)}\bigr)}{\bigl(1+h^2(E-V)/6\bigr)^2}}. \tag{8.42}$$

Note that the step size h has to satisfy the condition

$$\cos^2\bigl(h\sqrt{2(E-V)}\bigr) < \bigl(1+h^2(E-V)/6\bigr)^2, \tag{8.43}$$

which is fulfilled for any step size $h > 0$ since the left hand side of (8.43) is in $[0, 1]$, while the right hand side is always greater than 1. It is easy to show that $|\alpha| = 1$ and hence, the discrete wave vector k_h of the Numerov-Mickens FDS is real and reads

$$\begin{aligned} k_h = \pm\hat{k}_h &= \pm\frac{1}{h}\arg(\alpha) = \pm\frac{1}{h}\arccos\frac{\mathrm{Re}\alpha}{|\alpha|} \\ &= \pm\frac{1}{h}\arccos\frac{\cos\bigl(h\sqrt{2(E-V)}\bigr)}{1+h^2(E-V)/6}. \end{aligned} \tag{8.44}$$

Again we shall neglect to add the term $n\frac{2\pi}{h}$, $n \in \mathbb{Z}$, to this formula since for any $n \neq 0$ this term diverges for $h \to 0$. However, the limit for $h \to 0$ of the discrete wave vector k_h as given in (8.44), i.e. with $n = 0$, is undefined. In fact, the discrete wave vector k_h does not converge to the analytical wave vector k and hence, the Numerov-Mickens FDS does not converge for $h \to 0$. Nevertheless, we will continue to analyze this FDS and formulate DTBCs with the right-traveling wave $\psi_j = \alpha_1^j = \mathrm{e}^{\mathrm{i}\hat{k}jh}$ and the left-traveling wave $\psi_j = \alpha_2^j = \mathrm{e}^{-\mathrm{i}\hat{k}jh}$.

An explicit formula of the *discrete dispersion relation* of the Numerov-Mickens FDS cannot be derived. However, the discrete wave $\psi_j = \hat{\psi}\,\mathrm{e}^{\mathrm{i}k_hjh}$ implies $\psi_{j+1}\mathrm{e}^{-\mathrm{i}kh} = \psi_j = \psi_{j-1}\,\mathrm{e}^{\mathrm{i}kh}$. Applied to the O$\Delta$E (8.39) we get

$$\cos\Bigl(h\sqrt{2(E_h^{\mathrm{NumMic}}-v)}\Bigr) - \Bigl(\frac{h^2}{6}\cos kh\Bigr)\bigl(E_h^{\mathrm{NumMic}}-V\bigr) - \cos kh = 0. \tag{8.45}$$

By numerically evaluating (8.45) for $h = 1/100$ and $V = 0$ with the MATLAB procedure `fsolve` using the tolerance 10^{-9}, we obtain the discrete dispersion relation E_h^{NumMic} as shown in Fig. 8.1.

Although the bad numerical behavior has been illustrated we shall prove

Theorem 8.6. *Let* $\{V_j\}$, $j = 0, \ldots, J$, *and* $E > \max\{0, V_L\}$ *be given, and assume*

$$h < \min\Bigl\{\sqrt{\frac{6}{|E-V_1|}}, \sqrt{\frac{6}{|E-V_{J-1}|}}, \frac{\pi}{\hat{k}_{h,0}}, \frac{\pi}{\hat{k}_{h,L}}\Bigr\}.$$

Then the discrete BVP (8.39) *of the combined Numerov-Mickens FDS with the DTBCs* (8.23) *has a unique solution* $\{\psi_j\}$, $j = 0, \ldots, J$.

Proof. Let us introduce $\varphi_j = \sigma_j \psi_j$ with $\sigma_j = 1 + \frac{h^2}{6}\left(E - V_j\right) \in \mathbb{R}$. Note that $\sigma_j > 0$ for $j = 0, 1, J-1, J$.

We rewrite the Numerov-Mickens FDS in the form

$$\mathrm{D}_h^{\mathrm{std}} \varphi_j = -\frac{2}{h^2}\left(1 - \sigma_j^{-1} D_j\right)\varphi_j .$$

Since $-\frac{2}{h^2}\left(1 - \sigma_j^{-1} D_j\right) \in \mathbb{R}$ for all $j = 1, \ldots, J-1$, this theorem is obviously a direct corollary of Theorem 8.4 and hence, the solution is unique. □

8.2.6.4 Comparison of the Discrete Dispersion Relations

In the previous sections we introduced FDSs for the BVP (8.11) and derived the corresponding discrete dispersion relations. Now we want to compare these discrete dispersion relations with the analytical quantum mechanical dispersion relation (8.7). Figure 8.1 shows the analytical and discrete dispersion relations for a step size $h = 1/100$ and a potential $V = 0$.

All discrete dispersion relations except the dispersion relation of the Mickens FDS are periodic in the wave vector k_h with the period $2\pi/h \approx 628$. We can see that for small values of the wave vector, i.e. $k_h < 100$, the dispersion relation of the Numerov FDS coincides with the analytical dispersion relation for the used level of detail in Fig. 8.1, while the dispersion relation of the combined Numerov-Mickens FDS differs significantly from the analytical dispersion relation. Particularly, for $E = 500$, the value of the energy we used in our examples, the error of the dispersion relation of the Numerov-Mickens FDS is greater than of the other FDSs. This explains the significantly greater phase error of the Numerov-Mickens FDS we observed in Sect. 8.2.6.3.

8.2.7 Numerical Example: The Single Barrier Potential

In this section we analyze the results of the four introduced FDSs in the case of a single barrier potential. We consider a semiconductor of length L composed of two different materials, e.g. GaAs (gallium arsenide) and AlGaAs (aluminium gallium arsenide), where the latter is built between two parts of the first material. Let $0 < x_1 < x_2 < L$ and let the domain $[x_1, x_2)$ be composed of AlGaAs, while the two outer domains $[0, x_1)$ and $[x_2, L]$ are composed of GaAs. $V(x) = E_c(x)$ describes the band edge profile or the variation of the conduction band edge of the semiconductor materials. We call $V_0 = \Delta E_c = E_c|_{\mathrm{AlGaAs}} - E_c|_{\mathrm{GaAs}}$ *band edge offset* between the semiconductor materials or band edge discontinuity of

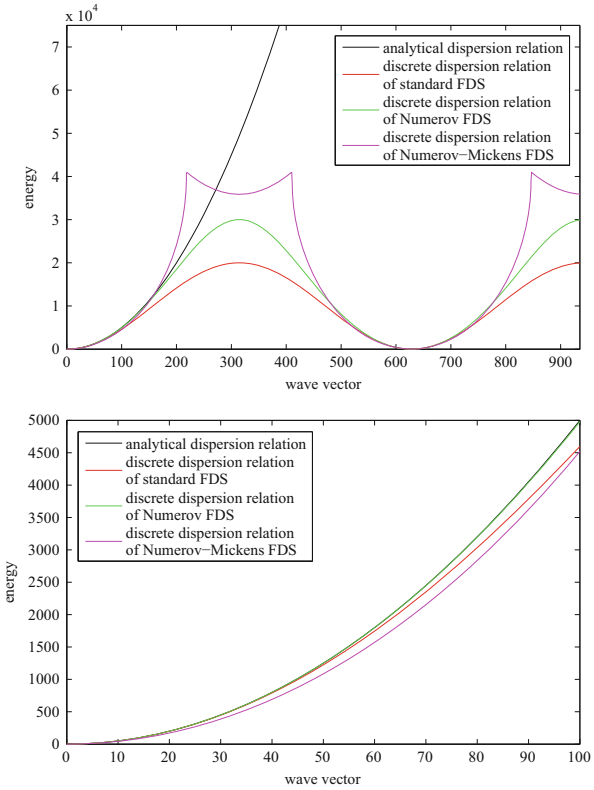


Fig. 8.1 Analytical quantum mechanical dispersion relation $E(k)$ (*black*), the discrete dispersion relation of the standard discretization $E_h^{\mathrm{std}}(k_h)$ (*red*), the discrete dispersion relation of the Numerov discretization $E_h^{\mathrm{Num}}(k_h)$ (*green*) and the discrete dispersion relation of the combined Numerov-Mickens discretization $E_h^{\mathrm{NumMic}}(k_h)$ (*magenta*) for the step size $h = 1/100$. Note that the discrete dispersion relation of the Mickens discretization $E_h^{\mathrm{Mic}}(k_h)$ coincides with the analytical quantum mechanical dispersion relation $E(k)$

the material interface. The inner domain $[x_1, x_2)$ is called *quantum barrier* if its potential $V(x) = E_{\mathrm{c}}|_{\mathrm{AlGaAs}}$ is greater than the potential $V(x) = E_{\mathrm{c}}|_{\mathrm{GaAs}}$ of the outer domains and *quantum well* if it is smaller.

For simplicity we set $E_{\mathrm{c}}|_{\mathrm{GaAs}} = 0$ and we assume that $E_{\mathrm{c}}|_{\mathrm{AlGaAs}} = 500$, i.e. the band edge offset is $V_0 = \Delta E_{\mathrm{c}} = 500$ and we have a quantum barrier at $x_1 \leq x < x_2$. Furthermore, we set $L = 1$, $x_1 = 1/3$ and $x_2 = 2/3$.

8.2.7.1 Analytical Solution

For a single barrier potential the BVP (8.11) can be solved analytically. Assuming that a right-traveling wave of amplitude $\hat{\psi} = 1$ enters the semiconductor at $x = 0$, the wave function reads

$$\psi(x) = \begin{cases} \mathrm{e}^{\mathrm{i}\sqrt{2E}x} + r\,\mathrm{e}^{-\mathrm{i}\sqrt{2E}x}, & \text{if } x \in [0, x_1), \\ a\,\mathrm{e}^{\mathrm{i}\sqrt{2(E-V_0)}x} + b\,\mathrm{e}^{-\mathrm{i}\sqrt{2(E-V_0)}x}, & \text{if } x \in [x_1, x_2), \\ t\,\mathrm{e}^{\mathrm{i}\sqrt{2E}x}, & \text{if } x \in [x_2, L]. \end{cases} \tag{8.46}$$

The coefficients r, a, b and t can be determined by using the continuity argument of the wave function and its first derivative at $x = x_1$ and $x = x_2$. For more details on the computation of the analytical solution including formulas of the coefficients r, a, b and t the interested reader is referred to [15].

Note that the solution of the so-called *transfer-matrix method* coincides with the analytical solution in the case of a single barrier potential, cf. [24].

Since we have three domains of the same length the step size h has to be of the form $h = \frac{1}{3n}$ with $n = 1, 2, \ldots$, so that the discretized domains also have the same length. Otherwise the FDSs would not solve the problem as stated above and the results would differ significantly from the analytical solution.

8.2.7.2 Numerical Results: The L^2-Error

In this section we present the numerical results of the introduced FDSs and show their discrete L^2-error. The evaluation of the discrete L^2-error, however, is not straightforward for a complex function. Since the numerical results of a stationary problem such as the BVP (8.11) has an arbitrary phase, we have to optimize the L^2-error with respect to a phase offset $\varphi \in [-\pi, \pi]$. In other words we have to solve the nonlinear problem

$$\Delta\psi_h^{\min} = \min_{\varphi\in[-\pi,\pi]} \Delta\psi_h = \min_{\varphi\in[-\pi,\pi]} \frac{1}{J+1} \sqrt{\sum_{j=0}^{J} |\psi(x_j) - \psi_h(x_j)\mathrm{e}^{\mathrm{i}\varphi}|^2}, \tag{8.47}$$

where ψ denotes the analytical solution and ψ_h is the numerical solution using the step size $h = 1/J$. In order to evaluate the minimal L^2-error $\Delta\psi_h^{\min}$ we discretize the domain $[-\pi, \pi]$ with a step size $h_\varphi = 2\pi/1000$ and analyze the L^2-error $\Delta\psi_h$ at every grid point.

In Fig. 8.2(top) the L^2-errors $\Delta\psi_h^{\min}$ of the standard FDS and the combined Numerov-Mickens FDS are plotted against the number of grid points $J = 1/h$ for the resonance energy $E = E_{\text{resonance}} \approx 544$. The error of the Numerov FDS and the error of the Mickens FDS coincide with the error of the standard discretization for the level of detail in Fig. 8.2(top). For these three FDSs the L^2-error is in $\mathcal{O}(h^2)$. The L^2-error of the combined Numerov-Mickens FDS, however, is only in $\mathcal{O}(h)$.

Figure 8.2(bottom) shows that the phase shift adjusted L^2-errors of the Numerov FDS almost coincides with the standard FDS and the Mickens FDS, i.e. the Numerov FDS turns out to be not of higher order than the standard FDS and the Mickens FDS. Although the higher order of the Numerov FDS is considered to be

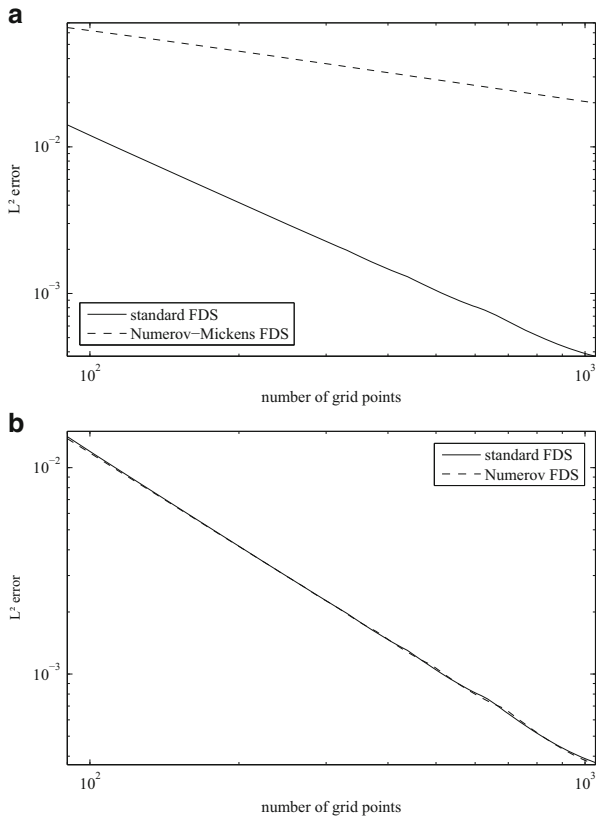


Fig. 8.2 Comparison of the L^2-errors of the numerical schemes for the resonance energy $E = E_{\text{resonance}} \approx 544$. Note that the L^2-errors of the Numerov FDS and the Mickens FDS (*dashed lines*) coincide with the L^2-error of the standard FDS (*solid line*) for the used level of detail in (*top*), while the L^2-error of the Mickens FDS coincides with the L^2-error of the Numerov FDS for the used level of detail in (*bottom*)

an advantage compared to the standard FDS and the Mickens FDS, it is this property that leads to the observed error of the scheme. By applying the identity

$$\psi_{xx} = \left(\left(\frac{\psi_x}{\psi} \right)_x + \left(\frac{\psi_x}{\psi} \right)^2 \right) \psi$$

to the Schrödinger equation (8.1) we get

$$\left(\frac{\psi_x}{\psi} \right)_x + \left(\frac{\psi_x}{\psi} \right)^2 = V(x) - E. \tag{8.48}$$

Under the assumption of the Numerov FDS, i.e. that the discretization is of fourth order, the left hand side of (8.48) is second order differentiable. The right hand side, however, is not second order differentiable as the potential V comprises two jump discontinuities at the barrier's ends. For the standard discretization, the left hand side of (8.48) is continuous but not necessarily differentiable. Hence, the jump discontinuities of the potential also lead to an error but we expect this error to be smaller than for the Numerov scheme.

An approach to improve the behavior of the Numerov FDS for a discontinuous potential V is to use the standard FDS of the Schrödinger equation at the point of discontinuity of the potential V and the Numerov FDS elsewhere. Although spurious oscillations due to possible incompatibility of the two schemes cannot be observed, numerical testing shows that the results cannot be improved significantly.

Before continuing with the multi-band case, let us note that apart from the list of FDSs we introduced in this section there are a plenty of FDSs that can be used for the stationary linear Schrödinger equation (8.1). However, the chosen FDSs demonstrate the principle of using FDSs for the Schrödinger equation clearly. The reader is referred to Simos and Williams [26] for a concise review on FDSs for the Schrödinger equation.

8.3 The General k·p-Model

In this section we introduce the general **k·p**-model. Let $d \in \mathbb{N}$ denote the number of considered bands of the semiconductor and $\mathbf{F}(x) \in \mathbb{C}^d$ the vector of the envelope functions $F_1, \ldots, F_d \in \mathbb{C}$. Let $\mathbf{m}(x), \mathbf{e}(x) \in \mathbb{R}^{d\times d}$ be diagonal, real and regular $d \times d$-matrices, $\mathbf{U}_p(x), \mathbf{U}_{pq}(x), \mathbf{v}(x) \in \mathbb{C}^{d\times d}$ Hermitian $d \times d$-matrices and $\mathbf{M}_0(x), \mathbf{M}_1(x), \mathbf{M}_2(x) \in \mathbb{C}^{d\times d}$ skew-Hermitian $d \times d$-matrices. Then we will refer to

$$\begin{aligned} E\mathbf{F}(x) = &- \frac{\partial}{\partial x}\Big(\mathbf{m}(x)\frac{\partial}{\partial x}\mathbf{F}(x)\Big) + \mathbf{M}_0(x)\frac{\partial}{\partial x}\mathbf{F}(x) - \frac{\partial}{\partial x}\left(\mathbf{M}_0^{\mathrm{H}}(x)\mathbf{F}(x)\right) \\ &+ k_1\Big(\mathbf{M}_1(x)\frac{\partial}{\partial x}\mathbf{F}(x) - \frac{\partial}{\partial x}\left(\mathbf{M}_1^{\mathrm{H}}(x)\mathbf{F}(x)\right)\Big) \\ &+ k_2\Big(\mathbf{M}_2(x)\frac{\partial}{\partial x}\mathbf{F}(x) - \frac{\partial}{\partial x}\left(\mathbf{M}_2^{\mathrm{H}}(x)\mathbf{F}(x)\right)\Big) \\ &+ k_1\mathbf{U}_1(x)\mathbf{F}(x) + k_2\mathbf{U}_2(x)\mathbf{F}(x) \\ &+ k_1^2\mathbf{U}_{11}(x)\mathbf{F}(x) + k_2^2\mathbf{U}_{22}(x)\mathbf{F}(x) + k_1k_2\left(\mathbf{U}_{12}(x) + \mathbf{U}_{21}(x)\right)\mathbf{F}(x) \\ &+ \mathbf{v}(x)\mathbf{F}(x) + \mathbf{e}(x)\mathbf{F}(x), \end{aligned} \tag{8.49}$$

with $x \in \mathbb{R}$ and $k_1, k_2 \in \mathbb{R}$, as *d-band* **k·p**-*Schrödinger equation*, cf. [6]. In order to abbreviate this physical formulation we introduce the skew-Hermitian $d \times d$-matrix

$$\mathbf{M}_{\mathrm{S}}(x) = \mathbf{M}_0(x) + k_1\mathbf{M}_1(x) + k_2\mathbf{M}_2(x), \tag{8.50a}$$

and the Hermitian $d \times d$-matrix

$$\begin{aligned}\mathbf{V}(x) &= k_1\mathbf{U}_1(x) + k_2\mathbf{U}_2(x)\\ &\quad + k_1^2\mathbf{U}_{11}(x) + k_2^2\mathbf{U}_{22}(x) + k_1k_2\big(\mathbf{U}_{12}(x) + \mathbf{U}_{21}(x)\big) + \mathbf{v}(x) + \mathbf{e}(x).\end{aligned} \tag{8.50b}$$

Then (8.49) reads

$$E\mathbf{F}(x) = -\frac{\partial}{\partial x}\Big(\mathbf{m}(x)\frac{\partial}{\partial x}\mathbf{F}(x)\Big) + \mathbf{M}_{\mathrm{S}}(x)\frac{\partial}{\partial x}\mathbf{F}(x) - \frac{\partial}{\partial x}\big(\mathbf{M}_{\mathrm{S}}^{\mathrm{H}}(x)\mathbf{F}(x)\big) + \mathbf{V}(x)\mathbf{F}(x),$$

with $x \in \mathbb{R}$.

We consider a semiconductor of length L connected to reservoirs at $x = 0$ and $x = L$. Let us assume that the matrices $\mathbf{m}$, $\mathbf{M}_{\mathrm{S}}$ and $\mathbf{V}$ are constant in the reservoirs with

$$\mathbf{m}(x) \equiv \mathbf{m}_0, \quad \mathbf{M}_{\mathrm{S}}(x) \equiv \mathbf{M}_{\mathrm{S},0}, \quad \mathbf{V}(x) \equiv \mathbf{V}_0, \qquad x \le 0,$$

and

$$\mathbf{m}(x) \equiv \mathbf{m}_L, \quad \mathbf{M}_{\mathrm{S}}(x) \equiv \mathbf{M}_{\mathrm{S},L}, \quad \mathbf{V}(x) \equiv \mathbf{V}_L, \qquad x \ge L.$$

8.3.1 The Exterior Problem and the Dispersion Relation

Let us study the exterior problem of the general **k·p**-model. In the exterior domains the matrices $\mathbf{m}$, $\mathbf{M}_{\mathrm{S}}$ and $\mathbf{V}$ are constant. Without loss of generality, we focus on the left exterior domain $x \le 0$ with $\mathbf{m}(x) = \mathbf{m}_0$, $\mathbf{M}_{\mathrm{S}}(x) = \mathbf{M}_{\mathrm{S},0}$ and $\mathbf{V}(x) = \mathbf{V}_0$. Note that the results for the right exterior domain $x \ge L$ can be derived analogously. For simplicity let us omit the subscript 0 in $\mathbf{m}_0$, $\mathbf{M}_{\mathrm{S},0}$ and $\mathbf{V}_0$. With these simplifying assumptions, (8.49) regarded on the half line $x \le 0$ is a second order system of ODEs with constant coefficients that can be written in the form

$$-\mathbf{N}\frac{\mathrm{d}^2}{\mathrm{d}x^2}\mathbf{F} + \mathrm{i}\mathbf{M}\frac{\partial}{\partial x}\mathbf{F} + (\mathbf{V} - E\mathbf{1})\mathbf{F} = \mathbf{0}, \qquad x \le 0, \tag{8.51}$$

with $\mathbf{N} = \mathbf{m}$ and $\mathbf{M} = -\mathrm{i}\mathbf{M}_{\mathrm{S}} + \mathrm{i}\mathbf{M}_{\mathrm{S}}^{\mathrm{H}} = -2\mathrm{i}\mathbf{M}_{\mathrm{S}}$. Note that $\mathbf{M}$ is Hermitian since $\mathbf{M}_{\mathrm{S}}$ is skew-Hermitian.

By introducing the standard substitution $\Phi = (\mathbf{F}, \frac{\partial}{\partial x}\mathbf{F})^\top$ we can reduce (8.51) to a first order system of ODEs with constant coefficients

$$\mathbf{A}\frac{\partial}{\partial x}\Phi = \mathbf{B}\Phi, \qquad x \le 0, \tag{8.52}$$

with

$$\mathbf{A} = \begin{pmatrix} \mathbf{M} & \mathrm{i}\mathbf{N} \\ -\mathrm{i}\mathbf{N} & \mathbf{0} \end{pmatrix} \in \mathbb{C}^{(2d)\times(2d)}, \qquad \mathbf{B} = \begin{pmatrix} \mathrm{i}\mathbf{V} - \mathrm{i}E\mathbf{1} & \mathbf{0} \\ \mathbf{0} & -\mathrm{i}\mathbf{N} \end{pmatrix} \in \mathbb{C}^{(2d)\times(2d)}.$$

Zisowsky [28] showed for the transient general **k**·**p**-model that the matrices

$$\mathbf{A}_\mathrm{t} = \mathbf{A}, \qquad \mathbf{B}_\mathrm{t} = \begin{pmatrix} \mathrm{i}\mathbf{V} + s\mathbf{1} & \mathbf{0} \\ \mathbf{0} & -\mathrm{i}\mathbf{N} \end{pmatrix},$$

with the Laplace parameter s of the Laplace-transformed exterior problem, are regular for $\mathrm{Re}(s) > 0$. Moreover she proved a *splitting theorem* saying that the matrix $\mathbf{A}_\mathrm{t}^{-1}\mathbf{B}_\mathrm{t}$ has exactly d eigenvalues with positive real part and d eigenvalues with negative real part.

Our aim is to show a similar result for the stationary general **k**·**p**-model (8.52). First let us show that **A** and **B** are regular and hence, $\mathbf{A}^{-1}\mathbf{B}$ exists and is also regular. Since $\mathbf{A}_\mathrm{t} = \mathbf{A}$, the matrix **A** is regular and $\mathbf{A}^{-1}\mathbf{B}$ exists. Since **V** is Hermitian, it is diagonalizable with the real eigenvalues $v_1, \ldots, v_d$. Let us suppose that the energy E satisfies $E \neq v_p$, for $p = 1, \ldots, d$, then the matrix $\mathrm{i}\mathbf{V} - \mathrm{i}E\mathbf{1}$ is similar to $\mathrm{diag}\,(\mathrm{i}\,(v_1 - E), \ldots, \mathrm{i}\,(v_d - E))$ which is regular. Considering that **N** is regular, the matrix **B** is regular and hence, the matrix $\mathbf{A}^{-1}\mathbf{B}$ is regular.

Thus, we can write (8.52) in the form

$$\frac{\partial}{\partial x}\Phi = \mathbf{A}^{-1}\mathbf{B}\Phi, \qquad x \le 0, \tag{8.53}$$

with

$$\mathbf{A}^{-1}\mathbf{B} = \begin{pmatrix} \mathbf{0} & \mathbf{1} \\ \mathbf{N}^{-1}\,(\mathbf{V} - E\mathbf{1}) & \mathrm{i}\mathbf{N}^{-1}\mathbf{M} \end{pmatrix} \in \mathbb{C}^{(2d)\times(2d)}. \tag{8.54}$$

The solution of (8.52) takes the form

$$\Phi(x) = \mathbf{a}\mathrm{e}^{\kappa x}, \qquad x \le 0, \tag{8.55}$$

where $\kappa = \kappa_1, \ldots, \kappa_{2d} \in \mathbb{C}$ denotes an eigenvalue and $\mathbf{a} = \mathbf{a}(\kappa) \in \mathbb{C}^{2d}$ the corresponding eigenvector of the matrix $\mathbf{A}^{-1}\mathbf{B}$. Since the vector of the envelope functions **F** is represented by the first d components of Φ, we introduce the

amplitude $\hat{\mathbf{F}} \in \mathbb{C}^d$ of $\mathbf{F}$ that contains the first d components of $\mathbf{a}(\kappa) \in \mathbb{C}^{2d}$. Then the vector of the envelope functions $\mathbf{F}$ takes the form

$$\mathbf{F}(x) = \hat{\mathbf{F}} \mathrm{e}^{\mathrm{i}kx}, \tag{8.56}$$

where $k = \hat{k} + \mathrm{i}\check{k} = -\mathrm{i}\kappa$ is called *wave vector* of $\mathbf{F}$ with the *propagation coefficient* $\hat{k}$ and the *attenuation coefficient* $\check{k}$. If the attenuation coefficient $\check{k}$ is zero, we say that $\mathbf{F}$ is traveling, while $\mathbf{F}$ is called evanescent otherwise. Again we shall refer to the vector $\mathbf{F}$ of the envelope functions as *envelope wave* since it can be written in the form of a plane wave.

By applying the solution (8.55) to the general $\mathbf{k}\cdot\mathbf{p}$-model (8.49) we get

$$\hat{\mathbf{H}}\hat{\mathbf{F}} = E\hat{\mathbf{F}}, \tag{8.57}$$

with

$$\hat{\mathbf{H}} = \hat{\mathbf{H}}(k) = k^2\mathbf{N} - k\mathbf{M} + \mathbf{V}. \tag{8.58}$$

Note that $\hat{\mathbf{H}}$ is Hermitian if k is real.

Now we propose the main theorem of this section.

Theorem 8.7 (Splitting Theorem). *Let n denote the number of positive eigenvalues of $\mathbf{N}$. Then there exists an energy $E_0^{\mathrm{e}} \in \mathbb{R}$ such that for all energies $E > E_0^{\mathrm{e}}$*

(i) there are exactly n positive and n negative wave vectors (i.e. n right and n left-traveling envelope waves),

(ii) there are exactly $2(d-n)$ complex wave vectors, $d-n$ with positive imaginary part (i.e. $d-n$ evanescent envelope waves decaying for $x \to \infty$) and $d-n$ with negative imaginary part (i.e. $d-n$ evanescent envelope waves growing for $x \to \infty$).

Moreover, there exists an energy $E_0^{\mathrm{h}} < E_0^{\mathrm{e}}$ such that for all energies $E < E_0^{\mathrm{h}}$

(iii) there are exactly $d-n$ positive and $d-n$ negative wave vectors (i.e. $d-n$ right and $d-n$ left-traveling envelope waves) and

(iv) there are exactly $2n$ complex wave vectors, n with positive imaginary part (i.e. n evanescent envelope waves decaying for $x \to \infty$) and n with negative imaginary part (i.e. n evanescent envelope waves growing for $x \to \infty$).

Numerical evidence strongly supports the validity of this theorem. However, we were unable to prove it analytically. For the simple case $\mathbf{M} = 0$ and $\mathbf{V} = \mathrm{diag}(v_1, \ldots, v_d)$ a proof can be found in [15].

Remark 8.1. In all examples considered, the d amplitudes $\hat{\mathbf{F}}(k)$ that correspond to the n positive wave vectors and the $d-n$ complex wave vectors with positive imaginary part are linearly independent. Moreover, the d amplitudes $\hat{\mathbf{F}}(k)$ associated

with the n negative wave vectors and the $d-n$ complex wave vectors with negative imaginary part are linearly independent.

8.3.2 Transparent Boundary Conditions

Let us recall the strategy we used for the single-band model in order to derive the TBCs. We considered a traveling envelope wave $\mathbf{F}^{\text{in}}$ with amplitude of norm 1 that enters the computational domain at $x = 0$. This means that depending on the energy E we require the matrix $\mathbf{N}$ to have at least one positive or negative eigenvalue in order to get at least one pair of traveling envelope waves.

Let us from now on assume that the energy E is greater than some lower bound E_0^{e} and hence, the number n of positive eigenvalues of $\mathbf{N}$ is equal to the number of pairs of traveling envelope waves.

If there are two or more pairs of traveling envelope waves, the incoming envelope wave is not unique. In this case we shall consider a unitary superposition of all right-traveling envelope waves to enter the semiconductor at $x = 0$. This means that we have to specify a priori the values of the n coefficients of the superposition of incoming envelope waves.

Let $\hat{k}_{0,l}^+$, $l = 1, \ldots, n$, denote the n positive wave vectors and $\hat{k}_{0,l}^-$, $l = 1, \ldots, n$, the n negative wave vectors in the left exterior domain. Moreover, let $\check{k}_{0,l}^+$, $l = 1, \ldots, d-n$, denote the $d-n$ complex wave vectors with positive imaginary part and $\check{k}_{0,l}^-$, $l = 1, \ldots, d-n$, the $d-n$ complex wave vectors with negative imaginary part in the left exterior domain. The wave vectors in the right exterior domain are defined analogously with subscript L instead of 0.

Note that in all considered examples, $\hat{k}_l^+ = -\hat{k}_l^-$, for $l = 1, \ldots, n$, and $\check{k}_l^+ = -\check{k}_l^-$, for $l = 1, \ldots, d-n$. Let $\hat{\mathbf{F}}_0(k)$ denote the amplitude of norm 1 in the left exterior domain that corresponds to the wave vector k, i.e. the eigenvector of norm 1 of $\hat{\mathbf{H}}(k)$ to the energy eigenvalue E. On the other hand, let $\hat{\mathbf{F}}_L(k)$ denote the corresponding amplitude in the right exterior domain.

Let us consider the superposition of all right-traveling envelope functions

$$\mathbf{F}^{\text{in}} = \sum_{l=1}^{n} \omega_l \hat{\mathbf{F}}_0(\hat{k}_{0,l}^+) \mathrm{e}^{\mathrm{i}\hat{k}_{0,l}^+ x}, \qquad x < 0, \tag{8.59}$$

with the coefficients $\omega_1, \ldots, \omega_n \in \mathbb{C}$ that satisfy the normalization condition

$$\sum_{l=1}^{n} |\omega_l|^2 = 1.$$

This incoming superposition of envelope functions is partly reflected at the left boundary at $x = 0$, yielding a superposition of left-traveling and evanescent envelope functions

$$\mathbf{F}^{\mathrm{r}} = \sum_{l=1}^{n} \hat{r}_l \hat{\mathbf{F}}_0(\hat{k}_{0,l}^{-}) \mathrm{e}^{\mathrm{i}\hat{k}_{0,l}^{-} x} + \sum_{l=1}^{d-n} \check{r}_l \hat{\mathbf{F}}_0(\check{k}_{0,l}^{-}) \mathrm{e}^{\mathrm{i}\check{k}_{0,l}^{-} x}, \qquad x < 0, \tag{8.60}$$

with the reflection coefficients $\hat{r}_l$ and $\check{r}_l$. Furthermore, the incoming waves are partly transmitted at the right boundary at $x = L$, which results in a superposition of right-traveling and evanescent envelope functions that takes the form

$$\mathbf{F}^{\mathrm{t}} = \sum_{l=1}^{n} \hat{t}_l \hat{\mathbf{F}}_0(\hat{k}_{0,l}^{+}) \mathrm{e}^{\mathrm{i}\hat{k}_{0,l}^{+} x} + \sum_{l=1}^{d-n} \check{t}_l \hat{\mathbf{F}}_0(\check{k}_{0,l}^{+}) \mathrm{e}^{\mathrm{i}\check{k}_{0,l}^{+} x}, \qquad x > L, \tag{8.61}$$

with the transmission coefficients $\hat{t}_l$ and $\check{t}_l$. Thus, the left exterior solution reads

$$\mathbf{F} = \mathbf{F}^{\mathrm{in}} + \mathbf{F}^{\mathrm{r}}, \qquad x < 0, \tag{8.62}$$

while the solution in the right exterior domain is given by

$$\mathbf{F} = \mathbf{F}^{\mathrm{t}}, \qquad x > L. \tag{8.63}$$

In order to determine the TBC at the left boundary we evaluate the envelope function $\mathbf{F}$ and its first derivative $\frac{\partial}{\partial x}\mathbf{F}$ at $x = 0$. We get

$$\mathbf{F}(0) = \sum_{l=1}^{n} \omega_l \hat{\mathbf{F}}_0(\hat{k}_{0,l}^{+}) + \sum_{l=1}^{n} \hat{r}_l \hat{\mathbf{F}}_0(\hat{k}_{0,l}^{-}) + \sum_{l=1}^{d-n} \check{r}_l \hat{\mathbf{F}}_0(\check{k}_{0,l}^{-}) \tag{8.64a}$$

and

$$\frac{\partial}{\partial x}\mathbf{F}(0) = \sum_{l=1}^{n} \mathrm{i}\hat{k}_{0,l}^{+} \omega_l \hat{\mathbf{F}}_0(\hat{k}_{0,l}^{+}) + \sum_{l=1}^{n} \mathrm{i}\hat{k}_{0,l}^{-} \hat{r}_l \hat{\mathbf{F}}_0(\hat{k}_{0,l}^{-}) + \sum_{l=1}^{d-n} \mathrm{i}\check{k}_{0,l}^{-} \check{r}_l \hat{\mathbf{F}}_0(\check{k}_{0,l}^{-}). \tag{8.64b}$$

Let us introduce

$$\mathbf{P}_0 = \left(\hat{\mathbf{F}}_0(\hat{k}_{0,1}^{-}) \cdots \hat{\mathbf{F}}_0(\hat{k}_{0,n}^{-}) \; \hat{\mathbf{F}}_0(\check{k}_{0,1}^{-}) \cdots \hat{\mathbf{F}}_0(\check{k}_{0,d-n}^{-}) \right) \in \mathbb{C}^{d \times d},$$

and

$$\mathbf{K}_0 = \mathrm{diag}\left(\mathrm{i}\hat{k}_{0,1}^{-}, \ldots, \mathrm{i}\hat{k}_{0,n}^{-}, \mathrm{i}\check{k}_{0,1}^{-}, \ldots, \mathrm{i}\check{k}_{0,d-n}^{-} \right) \in \mathbb{C}^{d \times d},$$

as well as

$$\mathbf{r} = (\hat{r}_1, \dots, \hat{r}_n, \check{r}_1, \dots, \check{r}_{d-n})^\top \in \mathbb{C}^d.$$

Then we can rewrite the envelope wave $\mathbf{F}$ and its first derivative $\frac{\partial}{\partial x}\mathbf{F}$ at $x = 0$ as

$$\mathbf{P}_0\mathbf{r} = \mathbf{F}(0) - \sum_{l=1}^{n} \omega_l \hat{\mathbf{F}}_0(\hat{k}_{0,l}^+), \tag{8.65a}$$

and

$$\mathbf{P}_0\mathbf{K}_0\mathbf{r} = \frac{\partial}{\partial x}\mathbf{F}(0) - \sum_{l=1}^{n} \mathrm{i}\hat{k}_{0,l}^+ \omega_l \hat{\mathbf{F}}_0(\hat{k}_{0,l}^+). \tag{8.65b}$$

Since the amplitudes $\hat{\mathbf{F}}_0(\hat{k}_{0,1}^-), \dots, \hat{\mathbf{F}}_0(\hat{k}_{0,n}^-), \hat{\mathbf{F}}_0(\check{k}_{0,1}^-), \dots, \hat{\mathbf{F}}_0(\check{k}_{0,d-n}^-)$ are linearly independent, cf. Remark 8.1, the matrix $\mathbf{P}_0$ is regular and hence, its inverse $\mathbf{P}_0^{-1}$ exists. Then the reflection coefficient vector $\mathbf{r}$ reads

$$\mathbf{r} = \mathbf{P}_0^{-1}\Big(\mathbf{F}(0) - \sum_{l=1}^{n} \omega_l \hat{\mathbf{F}}_0(\hat{k}_{0,l}^+)\Big),$$

cf. (8.65a). Applied to (8.65b) we get the left TBC

$$\mathbf{F}_x(0) - \mathbf{P}_0\mathbf{K}_0\mathbf{P}_0^{-1}\mathbf{F}(0) = \sum_{l=1}^{n}\Big(\mathrm{i}\hat{k}_{0,l}^+\mathbf{1} - \mathbf{P}_0\mathbf{K}_0\mathbf{P}_0^{-1}\Big)\omega_l\hat{\mathbf{F}}_0(\hat{k}_{0,l}^+). \tag{8.66}$$

At the right boundary we proceed analogously. The envelope wave $\mathbf{F}$ and its first derivative $\frac{\partial}{\partial x}\mathbf{F}$ at $x = L$ read

$$\mathbf{F}(L) = \sum_{l=1}^{n} \hat{t}_l \hat{\mathbf{F}}_L(\hat{k}_{L,l}^+)\mathrm{e}^{\mathrm{i}\hat{k}_{L,l}^+ L} + \sum_{l=1}^{d-n} \check{t}_l \hat{\mathbf{F}}_L(\check{k}_{L,l}^+)\mathrm{e}^{\mathrm{i}\check{k}_{L,l}^+ L}, \tag{8.67a}$$

and

$$\frac{\partial}{\partial x}\mathbf{F}(L) = \sum_{l=1}^{n} \mathrm{i}\hat{k}_{L,l}^+ \hat{t}_l \hat{\mathbf{F}}_L(\hat{k}_{L,l}^+)\mathrm{e}^{\mathrm{i}\hat{k}_{L,l}^+ L} + \sum_{l=1}^{d-n} \mathrm{i}\check{k}_{L,l}^+ \check{t}_l \hat{\mathbf{F}}_L(\check{k}_{L,l}^+)\mathrm{e}^{\mathrm{i}\check{k}_{L,l}^+ L}. \tag{8.67b}$$

Let us introduce

$$\mathbf{P}_L = \Big(\hat{\mathbf{F}}_L(\hat{k}_{L,1}^+)\mathrm{e}^{\mathrm{i}\hat{k}_{L,1}^+ L} \cdots \hat{\mathbf{F}}_L(\hat{k}_{L,n}^+)\mathrm{e}^{\mathrm{i}\hat{k}_{L,n}^+ L} \\ \hat{\mathbf{F}}_L(\check{k}_{L,1}^+)\mathrm{e}^{\mathrm{i}\check{k}_{L,1}^+ L} \cdots \hat{\mathbf{F}}_L(\check{k}_{L,d-n}^+)\mathrm{e}^{\mathrm{i}\check{k}_{L,d-n}^+ L}\Big) \in \mathbb{C}^{d\times d},$$

and

$$\mathbf{K}_L = \operatorname{diag}\big(\mathrm{i}\hat{k}^+_{L,1}, \ldots, \mathrm{i}\hat{k}^+_{L,n}, \mathrm{i}\check{k}^+_{L,1}, \ldots, \mathrm{i}\check{k}^+_{L,d-n}\big) \in \mathbb{C}^{d\times d},$$

as well as

$$\mathbf{t} = (\hat{t}_1, \ldots, \hat{t}_n, \check{t}_1, \ldots, \check{t}_{d-n})^\top \in \mathbb{C}^d.$$

Then we can rewrite the envelope wave $\mathbf{F}$ and its first derivative $\frac{\partial}{\partial x}\mathbf{F}$ at $x = L$ as

$$\mathbf{P}_L\mathbf{t} = \mathbf{F}(L), \tag{8.68a}$$

and

$$\mathbf{P}_L\mathbf{K}_L\mathbf{t} = \frac{\partial}{\partial x}\mathbf{F}(L). \tag{8.68b}$$

Since the amplitudes $\hat{\mathbf{F}}_L(\hat{k}^+_{L,1}), \ldots, \hat{\mathbf{F}}_L(\hat{k}^+_{L,n}), \hat{\mathbf{F}}_L(\check{k}^+_{L,1}), \ldots, \hat{\mathbf{F}}_L(\check{k}^+_{L,d-n})$ are linearly independent, cf. Remark 8.1, so are

$$\hat{\mathbf{F}}_L(\hat{k}^+_{L,1})\mathrm{e}^{\mathrm{i}\hat{k}^+_{L,1}L}, \ldots, \hat{\mathbf{F}}_L(\hat{k}^+_{L,n})\mathrm{e}^{\mathrm{i}\hat{k}^+_{L,n}L}, \hat{\mathbf{F}}_L(\check{k}^+_{L,1})\mathrm{e}^{\mathrm{i}\check{k}^+_{L,1}L}, \ldots, \hat{\mathbf{F}}_L(\check{k}^+_{L,d-n})\mathrm{e}^{\mathrm{i}\check{k}^+_{L,d-n}L}$$

and hence, the matrix $\mathbf{P}_L$ is regular and its inverse $\mathbf{P}_L^{-1}$ exists. Then the transmission coefficient vector $\mathbf{t}$ becomes

$$\mathbf{t} = \mathbf{P}_L^{-1}\mathbf{F}(L),$$

cf. (8.68a). Applied to (8.68b) we get the right TBC

$$\mathbf{F}_x(L) - \mathbf{P}_L\mathbf{K}_L\mathbf{P}_L^{-1}\mathbf{F}(L) = \mathbf{0}. \tag{8.69}$$

Let us remark that the coefficients $\omega_1, \ldots, \omega_n$ restrict the solution at the left boundary. However, in physical applications, for example the unstrained eight-band $\mathbf{k}\cdot\mathbf{p}$-model of the lowest conduction band and the three top-most valence bands, all doubly degenerate, with $\mathbf{k}_{||} = \mathbf{0}$ we will consider in our numerical examples in Sect. 8.3.5, one typically considers only a particular incoming wave. In the example mentioned above, we have $n = 2$, for $E > E_0^{\mathrm{e}}$. As indicated, the bands are doubly degenerate. Mathematically, this means that the eigenvalues of $\mathbf{A}^{-1}\mathbf{B}$ are twofold degenerate, i.e. there exist d distinct wave vectors k and for every wave vector k there exist two corresponding amplitudes, if the geometric multiplicity equals the algebraic multiplicity which is the case in our example. In this particular example we have so-called *spin-up solutions* and *spin-down solutions*. If we only consider spin-up envelope functions for example, we set the coefficient ω_1 of the incoming spin-up envelope wave to one and the coefficient ω_2 of the incoming spin-down

envelope wave to zero. Depending on the band edge profile the resulting transmitted envelope waves may also consist of spin-down solutions.

8.3.3 The Discretization

We recall the uniform grid $x_j = jh$, $j = 0, \ldots, J$ with $L = jh$, of the computational interval $(0, L)$ with $\mathbf{N}_j = \mathbf{N}(x_j)$, $\mathbf{M}_{\mathrm{S}\,j} = \mathbf{M}_{\mathrm{S}}(x_j)$, $\mathbf{V}_j = \mathbf{V}(x_j)$ and the approximation $\mathbf{F}_j \approx \mathbf{F}(x_j)$, $j = 0, \ldots, J$. In order to discretize the general $\mathbf{k}\cdot\mathbf{p}$-model (8.49) we apply the second order centered difference operator $\mathrm{D}_h^{\mathrm{cen}}$ as well as the standard second order difference operator D_h^{std} to the abbreviated continuous formulation of the general $\mathbf{k}\cdot\mathbf{p}$-model (8.49)

$$E\mathbf{F} = -\mathbf{N}\mathbf{F}_{xx} + (-\mathbf{N}_x + 2\mathbf{M}_{\mathrm{S}})\mathbf{F}_x + (\mathbf{V} - \mathbf{M}_{\mathrm{S}\,x}^{\mathrm{H}})\mathbf{F}, \tag{8.70}$$

for $x \in (0, L)$. Thus the discretization of the general $\mathbf{k}\cdot\mathbf{p}$-model leads to

$$E\mathbf{F}_j = -\mathbf{N}_j \mathrm{D}_h^{\mathrm{std}}\mathbf{F}_j + \big(-\mathrm{D}_h^{\mathrm{cen}}\mathbf{N}_j + 2\mathbf{M}_{\mathrm{S}\,j}\big)\mathrm{D}_h^{\mathrm{cen}}\mathbf{F}_j + \big(\mathbf{V}_j - \mathrm{D}_h^{\mathrm{cen}}\mathbf{M}_{\mathrm{S}\,j}^{\mathrm{H}} - E\mathbf{1}\big)\mathbf{F}_j,$$

with $j = 1, \ldots, J-1$, which implies

$$\begin{aligned} E\mathbf{F}_j =& \Big(-\frac{1}{h^2}\mathbf{N}_j + \frac{1}{2h}\Big(-\frac{1}{2h}\big(\mathbf{N}_{j+1} - \mathbf{N}_{j-1}\big) + 2\mathbf{M}_{\mathrm{S}\,j}\Big)\Big)\mathbf{F}_{j+1} \\ &+ \Big(\frac{2}{h^2}\mathbf{N}_j + \mathbf{V}_j - \frac{1}{2h}\big(\mathbf{M}_{\mathrm{S}\,j+1}^{\mathrm{H}} - \mathbf{M}_{\mathrm{S}\,j-1}^{\mathrm{H}}\big)\Big)\mathbf{F}_j \\ &+ \Big(-\frac{1}{h^2}\mathbf{N}_j - \frac{1}{2h}\Big(-\frac{1}{2h}\big(\mathbf{N}_{j+1} - \mathbf{N}_{j-1}\big) + 2\mathbf{M}_{\mathrm{S}\,j}\Big)\Big)\mathbf{F}_{j-1}, \end{aligned} \tag{8.71}$$

with $j = 1, \ldots, J-1$.

In the exterior domains $x \le 0$ and $x \ge L$, the matrices $\mathbf{N}$, $\mathbf{M}_{\mathrm{S}}$ and $\mathbf{V}$ are constant. Without loss of generality we focus on the left exterior domain $x \le 0$ with $\mathbf{N}(x) = \mathbf{N}_0$, $\mathbf{M}_{\mathrm{S}}(x) = \mathbf{M}_{\mathrm{S},0}$ and $\mathbf{V}(x) = \mathbf{V}_0$. Note that the results for the right exterior domain $x \ge L$ can be derived analogously. For simplicity let us omit the subscript 0 in $\mathbf{N}_0$, $\mathbf{M}_{\mathrm{S},0}$ and $\mathbf{V}_0$. Hence, (8.71) is a second order OΔE with constant coefficients of the form

$$E\mathbf{F}_j = \mathbf{M}^{+}\mathbf{F}_{j+1} + \mathbf{M}^{0}\mathbf{F}_j + \mathbf{M}^{-}\mathbf{F}_{j-1}, \qquad j < 0, \tag{8.72}$$

with

$$\mathbf{M}^{+} = -\frac{1}{h^2}\mathbf{N} + \frac{1}{h}\mathbf{M}_{\mathrm{S}}, \quad \mathbf{M}^{0} = \frac{2}{h^2}\mathbf{N} + \mathbf{V}, \quad \mathbf{M}^{-} = -\frac{1}{h^2}\mathbf{N} - \frac{1}{h}\mathbf{M}_{\mathrm{S}}. \tag{8.73}$$

By introducing the substitution $\Phi_j = \left(\mathbf{F}_j, \mathbf{F}_{j+1}\right)^\top$, (8.72) can be transformed into a first order OΔE with constant coefficients

$$\mathbf{A}_h \Phi_j = \mathbf{B}_h \Phi_{j-1}, \qquad j < 0, \tag{8.74}$$

with

$$\mathbf{A}_h = \begin{pmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & -\mathbf{M}^+ \end{pmatrix}, \qquad \mathbf{B}_h = \begin{pmatrix} \mathbf{0} & \mathbf{1} \\ \mathbf{M}^- & \left(\mathbf{M}^0 - E\mathbf{1}\right) \end{pmatrix}.$$

Note that $\mathbf{M}^+$ and $\mathbf{M}^-$ are not necessarily regular. However, in all examples we examined $\mathbf{M}^+$ and $\mathbf{M}^-$ are regular. In this case, $\mathbf{A}_h$ and $\mathbf{B}_h$ are regular and hence, we can write (8.74) in the form

$$\Phi_j = \mathbf{A}_h^{-1} \mathbf{B}_h \Phi_{j-1}, \qquad j < 0,$$

with the regular matrix

$$\mathbf{A}_h^{-1} \mathbf{B}_h = \begin{pmatrix} \mathbf{0} & \mathbf{1} \\ (-\mathbf{M}^+)^{-1}\mathbf{M}^- & (-\mathbf{M}^+)^{-1}\left(\mathbf{M}^0 - E\mathbf{1}\right) \end{pmatrix}.$$

Remark 8.2. In all examples we examined the geometric multiplicity of the eigenvalues of $\mathbf{A}_h^{-1}\mathbf{B}_h$ is equal to their algebraic multiplicity. Hence, the eigenvectors of $\mathbf{A}_h^{-1}\mathbf{B}_h$ are linearly independent and form a basis of $\mathbb{C}^{4d}$.

This remark is important in order to use the fact that the first order OΔE (8.74) with the initial value $\Phi_0 = \mathbf{a}$ has a solution of the form $\Phi_j = \mathbf{a}\,\alpha^j$, where $\alpha \in \mathbb{C}$ is an eigenvalue of $\mathbf{A}_h^{-1}\mathbf{B}_h$ with corresponding eigenvector $\mathbf{a} \in \mathbb{C}^{2d}$. Thus, we can set the discrete solution at the left boundary to some eigenvector $\mathbf{a}$ of $\mathbf{A}_h^{-1}\mathbf{B}_h$.

The first d components of $\Phi_j \in \mathbb{C}^{2d}$ represent the discrete solution $\mathbf{F}_j \in \mathbb{C}^d$. Therefore, we introduce the *discrete amplitude* $\hat{\mathbf{F}}_h \in \mathbb{C}^d$ that contains the first d components of $\mathbf{a}$. The discrete solution $\mathbf{F}_j$ becomes

$$\mathbf{F}_j = \hat{\mathbf{F}}_h\, \alpha^j = \hat{\mathbf{F}}_h\, \mathrm{e}^{\mathrm{i}k_h jh}, \tag{8.75}$$

with the *discrete wave vector* $k_h = (\arg(\alpha) - \mathrm{i}\ln|\alpha|)/h$.

The discrete solution as given in (8.75) implies $\mathbf{F}_{j+1}\mathrm{e}^{-\mathrm{i}k_h h} = \mathbf{F}_j = \mathbf{F}_{j-1}\,\mathrm{e}^{\mathrm{i}k_h h}$ and thus, applied to the OΔE (8.72) we obtain

$$\hat{\mathbf{H}}_h \hat{\mathbf{F}}_h = E\hat{\mathbf{F}}_h, \tag{8.76}$$

with $\hat{\mathbf{H}}_h = \hat{\mathbf{H}}_h(k_h) = \mathbf{M}^+\mathrm{e}^{\mathrm{i}k_h h} + \mathbf{M}^0 + \mathbf{M}^-\mathrm{e}^{-\mathrm{i}k_h h}$.

Now we shall state the discrete analogon of Theorem 8.7.

Theorem 8.8 (Discrete Splitting Theorem). *Let n denote the number of positive eigenvalues of* $\mathbf{N}$*. Then there exists an energy $E^{\mathrm{e}}_{0,h} \in \mathbb{R}$ such that for all energies $E > E^{\mathrm{e}}_{h,0}$*

(i) there are exactly n positive and n negative discrete wave vectors (i.e. n right and n left-traveling discrete envelope waves),
(ii) there are exactly $2\,(d-n)$ complex discrete wave vectors, $d-n$ with positive imaginary part (i.e. $d-n$ evanescent discrete envelope waves decaying for $x \to \infty$) and $d-n$ with negative imaginary part (i.e. $d-n$ evanescent discrete envelope waves growing for $x \to \infty$).

Moreover, there exists an energy $E^{\mathrm{h}}_{h,0} < E^{\mathrm{e}}_{h,0}$ such that for all energies $E < E^{\mathrm{h}}_{h,0}$

(iii) there are exactly $d-n$ positive and $d-n$ negative discrete wave vectors (i.e. $d-n$ right and $d-n$ left-traveling discrete envelope waves) and
(iv) there are exactly $2n$ complex discrete wave vectors, n with positive imaginary part (i.e. n evanescent discrete envelope waves decaying for $x \to \infty$) and n with negative imaginary part (i.e. n evanescent discrete envelope waves growing for $x \to \infty$).

Analogously to Theorem 8.7, numerical evidence strongly supports the validity of this theorem. However, an analytical proof has not been found yet.

We already pointed out that in all considered examples the geometric multiplicity of the eigenvalues of $\mathbf{A}_h^{-1}\mathbf{B}_h$ is equal to their algebraic multiplicity and hence, the eigenvectors are linearly independent. In addition, let us note:

Remark 8.3. In all considered examples the d discrete amplitudes $\hat{\mathbf{F}}_h(k_h)$ that correspond to the n positive discrete wave vectors and the $d-n$ complex discrete wave vectors with positive imaginary part are linearly independent. Moreover, the d discrete amplitudes $\hat{\mathbf{F}}_h(k_h)$ that are associated with the n negative discrete wave vectors and the $d-n$ complex discrete wave vectors with negative imaginary part are linearly independent.

8.3.4 Discrete Transparent Boundary Conditions

In order to derive the DTBCs for the general $\mathbf{k}\cdot\mathbf{p}$-model we apply the discrete solution derived in the previous section to the reflection and transmission conditions (8.62), (8.63) and assume that they hold in a small vicinity of the two boundaries, i.e. $j = 0, 1$ and $j = J-1, J$ respectively.

Let us from now on assume that the energy is greater than some lower bound $E^{\mathrm{e}}_{h,0}$ and hence, the number n of positive eigenvalues of $\mathbf{N}$ is equal to the number of purely imaginary, complex conjugate pairs of discrete wave vectors.

Suppose that there is at least one pair of discrete traveling envelope functions, in other words $n \geq 1$. If $n \geq 2$, then we have two or more pairs of traveling envelope functions and hence, the incoming envelope function is not unique. In this

case we shall proceed accordingly to the derivation of the TBCs and consider a unitary superposition of all discrete right-traveling envelope functions weighted by the coefficients $\omega_1, \ldots, \omega_n \in \mathbb{C}$.

Let $\hat{k}^+_{h,0,l}$, $l = 1, \ldots, n$, denote the n positive discrete wave vectors and $\hat{k}^-_{h,0,l}$, $l = 1, \ldots, n$, the n negative discrete wave vectors in the left exterior domain. Moreover, the $d - n$ complex discrete wave vectors with positive imaginary part in the left exterior domain are called $\check{k}^+_{h,0,l}$, $l = 1, \ldots, d - n$, and the $d - n$ complex discrete wave vectors with negative imaginary part are denoted by $\check{k}^-_{h,0,l}$, $l = 1, \ldots, d - n$. The discrete wave vectors in the right exterior domain are defined analogously with subscript L instead of 0.

In all considered examples we have $\hat{k}^+_{h,l} = -\hat{k}^-_{h,l}$, for $l = 1, \ldots, n$, and $\check{k}^+_{h,l} = -\check{k}^-_{h,l}$, for $l = 1, \ldots, d - n$.

Let $\hat{\mathbf{F}}_{h,0}(k_h)$ denote the amplitude of norm 1 in the left exterior domain that corresponds to the discrete wave vector k_h, i.e. the eigenvector of norm 1 of $\hat{\mathbf{H}}_h(k_h)$ to the energy eigenvalue E. On the other hand, let $\check{\mathbf{F}}_{h,L}(k_h)$ be the corresponding amplitude in the right exterior domain.

Then we have

$$\mathbf{F}_j = \mathbf{F}^{\mathrm{in}}_j + \mathbf{F}^{\mathrm{r}}_j = \sum_{l=1}^{n} \omega_l \hat{\mathbf{F}}_{h,0}(\hat{k}^+_{h,0,l}) \mathrm{e}^{\mathrm{i}\hat{k}^+_{h,0,l} jh} + \sum_{l=1}^{n} \hat{r}_{h,l} \hat{\mathbf{F}}_{h,0}(\hat{k}^-_{h,0,l}) \mathrm{e}^{\mathrm{i}\hat{k}^-_{h,0,l} jh} + \sum_{l=1}^{d-n} \check{r}_{h,l} \hat{\mathbf{F}}_{h,0}(\check{k}^-_{h,0,l}) \mathrm{e}^{\mathrm{i}\check{k}^-_{h,0,l} jh},$$

at the left boundary, i.e. for $j = 0, 1$, and

$$\mathbf{F}_j = \mathbf{F}^{\mathrm{t}}_j = \sum_{l=1}^{n} \hat{t}_h \hat{\mathbf{F}}_{h,L}(\hat{k}^+_{h,L,l}) \mathrm{e}^{\mathrm{i}\hat{k}^+_{h,L,l} jh} + \sum_{l=1}^{d-n} \check{t}_h \hat{\mathbf{F}}_{h,L}(\check{k}^+_{h,L,l}) \mathrm{e}^{\mathrm{i}\check{k}^+_{h,L,l} jh},$$

at the right boundary, i.e. for $j = J - 1, J$.

Let us introduce

$$\mathbf{P}_{h,0} = \Big(\hat{\mathbf{F}}_{h,0}(\hat{k}^-_{h,0,1}) \cdots \hat{\mathbf{F}}_{h,0}(\hat{k}^-_{h,0,n}) \; \hat{\mathbf{F}}_{h,0}(\check{k}^-_{h,0,1}) \cdots \hat{\mathbf{F}}_{h,0}(\check{k}^-_{h,0,d-n}) \Big) \in \mathbb{C}^{d \times d},$$

and

$$\mathbf{K}_{h,0} = \mathrm{diag}\big(\mathrm{e}^{\mathrm{i}\hat{k}^-_{h,0,1} h}, \ldots, \mathrm{e}^{\mathrm{i}\hat{k}^-_{h,0,n} h}, \mathrm{e}^{\mathrm{i}\check{k}^-_{h,0,1} h}, \ldots, \mathrm{e}^{\mathrm{i}\check{k}^-_{h,0,d-n} h} \big) \in \mathbb{C}^{d \times d},$$

as well as

$$\mathbf{r}_h = (\hat{r}_{h,1}, \ldots, \hat{r}_{h,n}, \check{r}_{h,1}, \ldots, \check{r}_{h,d-n})^\top \in \mathbb{C}^d.$$

Then we can rewrite the discrete envelope function $\mathbf{F}_j$ for $j = 0, 1$ in the form

$$\mathbf{P}_{h,0}\mathbf{r}_h = \mathbf{F}_0 - \sum_{l=1}^{n} \omega_l \hat{\mathbf{F}}_{h,0}(\hat{k}^+_{h,0,l}), \tag{8.77a}$$

and

$$\mathbf{P}_{h,0}\mathbf{K}_{h,0}\mathbf{r}_h = \mathbf{F}_1 - \sum_{l=1}^{n} \omega_l \hat{\mathbf{F}}_{h,0}(\hat{k}^+_{h,0,l}) \mathrm{e}^{\mathrm{i}\hat{k}^+_{h,0,l}h}. \tag{8.77b}$$

Since $\hat{\mathbf{F}}_{h,0}(\hat{k}^-_{h,0,1}), \ldots, \hat{\mathbf{F}}_{h,0}(\hat{k}^-_{h,0,n})$ and $\hat{\mathbf{F}}_{h,0}(\check{k}^-_{h,0,1}), \ldots, \hat{\mathbf{F}}_{h,0}(\check{k}^-_{h,0,d-n})$ are linearly independent (cf. Remark 8.3), the matrix $\mathbf{P}_{h,0}$ is regular and hence, its inverse $\mathbf{P}^{-1}_{h,0}$ exists. Then the reflection coefficient vector $\mathbf{r}_h$ is given by

$$\mathbf{r}_h = \mathbf{P}^{-1}_{h,0}\Big(\mathbf{F}_0 - \sum_{l=1}^{n} \omega_l \hat{\mathbf{F}}_{h,0}(\hat{k}^+_{h,0,l})\Big),$$

cf. (8.77a). Applied to (8.77b) we get the left DTBC

$$\mathbf{F}_1 - \mathbf{P}_{h,0}\mathbf{K}_{h,0}\mathbf{P}^{-1}_{h,0}\mathbf{F}_0 = \sum_{l=1}^{n} \big(\mathrm{e}^{\mathrm{i}\hat{k}^+_{h,0,l}}\mathbf{1} - \mathbf{P}_{h,0}\mathbf{K}_{h,0}\mathbf{P}^{-1}_{h,0}\big)\omega_l \hat{\mathbf{F}}_{h,0}(\hat{k}^+_{h,0,l}), \tag{8.78}$$

compared to the left TBC (8.66).

At the right boundary we proceed analogously. Let us introduce

$$\mathbf{P}_{h,L} = \big(\hat{\mathbf{p}}_1 \cdots \hat{\mathbf{p}}_n \ \check{\mathbf{p}}_1 \cdots \check{\mathbf{p}}_{d-n}\big) \in \mathbb{C}^{d\times d},$$

with the columns $\hat{\mathbf{p}}_l = \hat{\mathbf{F}}_{h,L}(\hat{k}^+_{h,L,l})\mathrm{e}^{\mathrm{i}\hat{k}^+_{h,L,l}hj}$ and $\check{\mathbf{p}}_l = \hat{\mathbf{F}}_{h,L}(\check{k}^+_{h,L,l})\mathrm{e}^{\mathrm{i}\check{k}^+_{h,L,l}hj}$. Moreover, we introduce

$$\mathbf{K}_{h,L} = \mathrm{diag}\big(\mathrm{e}^{-\mathrm{i}\hat{k}^+_{h,L,1}h}, \ldots, \mathrm{e}^{-\mathrm{i}\hat{k}^+_{h,L,n}h}, \mathrm{e}^{-\mathrm{i}\check{k}^+_{h,L,1}h}, \ldots, \mathrm{e}^{-\mathrm{i}\check{k}^+_{h,L,d-n}h}\big) \in \mathbb{C}^{d\times d},$$

as well as

$$\mathbf{t}_h = (\hat{t}_{h,1}, \ldots, \hat{t}_{h,n}, \check{t}_{h,1}, \ldots, \check{t}_{h,d-n})^\top \in \mathbb{C}^d.$$

Then we can rewrite the discrete envelope function $\mathbf{F}_j$ for $j = J-1, J$ in the form

$$\mathbf{P}_{h,L}\mathbf{t}_h = \mathbf{F}_J, \tag{8.79a}$$

and

$$\mathbf{P}_{h,L}\mathbf{K}_{h,L}\mathbf{t}_h = \mathbf{F}_{J-1}. \tag{8.79b}$$

Since the discrete amplitudes $\hat{\mathbf{F}}_{h,L}(\hat{k}^+_{h,L,1}), \ldots, \hat{\mathbf{F}}_{h,L}(\hat{k}^+_{h,L,n})$ and $\hat{\mathbf{F}}_{h,L}(\check{k}^+_{h,L,1}), \ldots, \hat{\mathbf{F}}_{h,L}(\check{k}^+_{h,L,d-n})$ are linearly independent (Remark 8.3), so are $\hat{\mathbf{F}}_{h,L}(\hat{k}^+_{h,L,1})\mathrm{e}^{\mathrm{i}\hat{k}^+_{h,L,1}hj}, \ldots, \hat{\mathbf{F}}_{h,L}(\hat{k}^+_{h,L,n})\mathrm{e}^{\mathrm{i}\hat{k}^+_{h,L,n}hj}$ and $\hat{\mathbf{F}}_{h,L}(\check{k}^+_{h,L,1})\mathrm{e}^{\mathrm{i}\check{k}^+_{h,L,1}hj}, \ldots, \hat{\mathbf{F}}_{h,L}(\check{k}^+_{h,L,d-n})\mathrm{e}^{\mathrm{i}\check{k}^+_{h,L,d-n}hJ}$. Thus, the matrix $\mathbf{P}_{h,L}$ is regular and its inverse $\mathbf{P}^{-1}_{h,L}$ exists. Then the transmission coefficient vector $\mathbf{t}_h$ reads

$$\mathbf{t}_h = \mathbf{P}^{-1}_{h,L}\mathbf{F}_J,$$

cf. (8.79a). Applied to (8.79b) we get the right DTBC

$$\mathbf{F}_{J-1} - \mathbf{P}_{h,L}\mathbf{K}_{h,L}\mathbf{P}^{-1}_{h,L}\mathbf{F}_J = \mathbf{0}, \tag{8.80}$$

in contrast to the right TBC (8.69).

8.3.5 *Numerical Examples*

8.3.5.1 The Free Scattering State

In our first example we want to examine the numerical result of an unstrained eight-band $\mathbf{k}\cdot\mathbf{p}$-model of the lowest conduction band and the three top-most valence bands, all doubly degenerate, with $\mathbf{k}_{||} = \mathbf{0}$, in the case of the free scattering state and compare it with the analytical solution, that can be derived from the results in Sect. 8.3.1.

In this case the 8×8-$\mathbf{k}\cdot\mathbf{p}$-Hamiltonian reduces to

$$\mathbf{H} = \mathbf{H}_0 + \mathbf{H}_\Delta + \mathbf{H}_1\frac{\partial}{\partial x} + \mathbf{H}_2\frac{\mathrm{d}^2}{\mathrm{d}x^2}, \tag{8.81}$$

where $\mathbf{H}_0$ describes the *band edge profile*, $\mathbf{H}_\Delta$ denotes the *spin orbit coupling*, $\mathbf{H}_1$ contains all *first order couplings*, i.e. the inter-band couplings, and $\mathbf{H}_2$ contains all *second order couplings*, i.e. the intra-band couplings, see [13].

The band edge profile is given by

$$\mathbf{H}_0 = \mathrm{diag}\,(E_\mathrm{c}, E_\mathrm{v}, E_\mathrm{v}, E_\mathrm{v}, E_\mathrm{c}, E_\mathrm{v}, E_\mathrm{v}, E_\mathrm{v}) \in \mathbb{R}^{8\times 8}, \tag{8.82}$$

where E_c is the conduction band edge and E_v is the valence band edge with the $E_\mathrm{g} = E_\mathrm{c} - E_\mathrm{v}$.

The spin orbit coupling matrix $\mathbf{H}_\Delta \in \mathbb{C}^{8\times 8}$ takes the form

$$\mathbf{H}_\Delta = \frac{\Delta_\mathrm{so}}{3}\begin{pmatrix} \mathbf{G} + \mathrm{i}\mathbf{G}_z & \mathbf{G}_y + \mathrm{i}\mathbf{G}_x \\ -\mathbf{G}_y + \mathrm{i}\mathbf{G}_x & \mathbf{G}_\mathrm{so} - \mathrm{i}\mathbf{G}_z \end{pmatrix}, \tag{8.83}$$

with $\mathbf{G}_{\mathrm{so}} = \mathrm{diag}\,(0,-1,-1,-1) \in \mathbb{R}^{4\times4}$ and $\mathbf{G}_x, \mathbf{G}_y, \mathbf{G}_z \in \mathbb{R}^{4\times4}$ defined by $(\mathbf{G}_x)_{ij} = \delta^{i,4}\delta^{j,3} - \delta^{i,3}\delta^{j,4}$, $(\mathbf{G}_y)_{ij} = \delta^{i,2}\delta^{j,4} - \delta^{i,4}\delta^{j,2}$, $(\mathbf{G}_z)_{ij} = \delta^{i,3}\delta^{j,2} - \delta^{i,2}\delta^{j,3}$, where $\delta^{i,j}$ is the usual Kronecker symbol with $\delta^{i,j} = 1$ if $i = j$ and $\delta^{i,j} = 0$ otherwise. The parameter Δ_{so} denotes the so-called *spin orbit splitting*.

The matrix $\mathbf{H}_1 \in \mathbb{R}^{8\times8}$ of first order couplings has the components

$$(\mathbf{H}_1)_{ij} = P_0\left(\delta^{i,1}\delta^{j,4} - \delta^{i,4}\delta^{j,1} + \delta^{i,8}\delta^{j,5} - \delta^{i,5}\delta^{j,8}\right) \tag{8.84}$$

while the matrix $\mathbf{H}_2 \in \mathbb{C}^{8\times8}$ of second order couplings takes the form

$$\mathbf{H}_2 = -\,\mathrm{diag}\,(\alpha,\mu,\mu,\lambda,\alpha,\mu,\mu,\lambda)\,, \tag{8.85}$$

where the coefficients $\alpha, \lambda, \mu \in \mathbb{C}$ are given by

$$\alpha = \frac{\hbar^2}{2m_{\mathrm{c}}} - \frac{P_0^2}{E_{\mathrm{g}}}\frac{E_{\mathrm{g}} + 2\frac{\Delta_{\mathrm{so}}}{3}}{E_{\mathrm{g}} + \Delta_{\mathrm{so}}}, \quad \lambda = \frac{P_0^2}{E_{\mathrm{g}}} - \frac{\hbar^2}{2m_0}(\gamma_1 + 4\gamma_2)\,, \quad \mu = -\frac{\hbar^2}{2m_0}(\gamma_1 - 2\gamma_2)\,,$$

with the effective mass m_{c} of the conduction band and the *Luttinger parameters* γ_1 and γ_2.

Written in the usual notation

$$-\mathbf{N}\frac{\mathrm{d}^2}{\mathrm{d}x^2}\mathbf{F} + \mathrm{i}\mathbf{M}\frac{\partial}{\partial x}\mathbf{F} + (\mathbf{V} - E\mathbf{1})\,\mathbf{F} = \mathbf{0}, \tag{8.86}$$

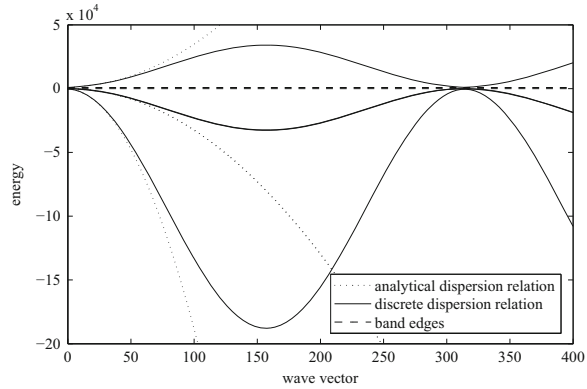
cf. (8.51), we have $\mathbf{N} = -\mathbf{H}_2$, $\mathbf{M} = -\mathrm{i}\mathbf{H}_1$ and $\mathbf{V} = \mathbf{H}_0 + \mathbf{H}_\Delta$.

For simplicity, we set $\hbar = m_0 = 1$ as well as $L = 1$. We use the dimensionless version of the parameters as given in [16] that are given by $\alpha = 3.32$, $\lambda = -18.77$, $\mu = -3.24$, $P_0 = 132.744$ and $\Delta_{\mathrm{so}} = 419.07$. According to [16], we set the band edges to $E_{\mathrm{c}} = 905.96$ and $E_{\mathrm{v}} = 0$.

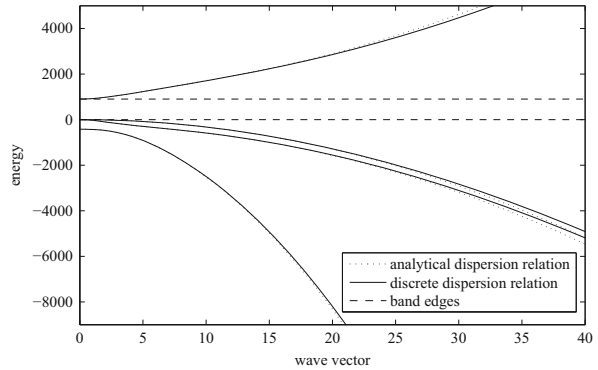
For these settings and a step size $h = 1/50$ Fig. 8.3 shows the analytical and discrete dispersion relations. The discrete dispersion relation is $2\frac{\pi}{h}$-periodic and the positive trunk of the discrete dispersion relation for these particular settings is injective in $[0, \pi/h]$. Thus, there does not exist an energy window such that we expect spurious oscillations for all admissible energies outside this window due to the wrong choice of the discrete wave vectors.

In the comparison of the numerical and analytical solution we observe a small phase error in the numerical solution. This error decreases for smaller step sizes which can be seen from the discrete L^2-error. Recall that the discrete L^2-error is the solution of the nonlinear optimization problem

$$\Delta\mathbf{F}_h^{\min} = \min_{\varphi\in[-\pi,\pi]} \Delta\mathbf{F}_h = \min_{\varphi\in[-\pi,\pi]} \frac{1}{J+1}\sqrt{\sum_{j=0}^{J}\left\|\mathbf{F}(x_j) - \mathbf{F}_j\mathrm{e}^{\mathrm{i}\varphi}\right\|^2}, \tag{8.87}$$



(a) Analytical and discrete dispersion relations.



(b) Detail view of (a).

Fig. 8.3 Analytical (*dotted line*) and discrete (*solid line*) dispersion relations of the unstrained eight-band **k**·**p**-model with $\mathbf{k}_{||} = \mathbf{0}$. The discrete dispersion relation is plotted for a step size $h = 1/50$, the bottom figure shows a detailed view

where $\mathbf{F}(x_j)$ denotes the analytical solution at $x = x_j$ and $\mathbf{F}_j$ the numerical solution using the step size $h = 1/J$. The discrete L^2-error is in $\mathcal{O}(h^2)$ which coincides with the formal order of the standard and centered difference operator we used in order to discretize the general **k**·**p**-model.

8.3.5.2 The Single Barrier Potential

In our second example we want to analyze the numerical results of the unstrained eight-band **k**·**p**-model with $\mathbf{k}_{||} = \mathbf{0}$ in the case of a single barrier potential. We consider a semiconductor of length L that is split into three parts. Let $0 < x_1 < x_2 < L$, then the three subdomains of the semiconductor are defined by $[0, x_1)$, $[x_1, x_2)$ and $[x_2, L]$. The two outer subdomains have the same physical properties and are denoted by $A = [0, x_1) \cup [x_2, L]$, while the inner subdomain is called $B = [x_1, x_2)$.

We use the same problem as in the previous example, but in the domain B we shall set the band edges to $E_{\mathrm{c}}^B = 1169.33$ and $E_{\mathrm{v}}^B = -167.60$, see [16]. Due to physical conventions, we shall refer to this band edge profile as *quantum well structure*.

Analogously to the single band case, we can compute the analytical solution since the matrices $\mathbf{N}, \mathbf{M}$ are constant the matrix $\mathbf{V}$ is piecewise constant. Let us denote the matrix $\mathbf{V}$ in the domain A by $\mathbf{V}_A$ and in the domain B by $\mathbf{V}_B$. Suppose that the energy E is greater than some lower bound E_0^{e}, cf. Theorem 8.7. Then the number of positive wave vectors is equal to the number n of positive eigenvalues of N. Thus, in each domain the envelope function takes the form

$$\mathbf{F}(x) = \sum_{l=1}^{n} \hat{a}_l \hat{\mathbf{F}}(\hat{k}_l^+) \mathrm{e}^{\mathrm{i}\hat{k}_l^+ x} + \sum_{l=1}^{d-n} \check{a}_l \hat{\mathbf{F}}(\check{k}_l^+) \mathrm{e}^{\mathrm{i}\check{k}_l^+ x} + \sum_{l=1}^{n} \hat{b}_l \hat{\mathbf{F}}(\hat{k}_l^-) \mathrm{e}^{\mathrm{i}\hat{k}_l^- x} + \sum_{l=1}^{d-n} \check{b}_l \hat{\mathbf{F}}(\check{k}_l^-) \mathrm{e}^{\mathrm{i}\check{k}_l^- x},$$

with the coefficients $\hat{a}_p, \check{a}_q, \hat{b}_p, \check{b}_q \in \mathbb{C}$, with $p = 1, \ldots, n$ and $q = 1, \ldots, d-n$. Here we used the notation of wave vectors and amplitudes from Sect. 8.3.2. Note that the amplitudes are of norm 1. In the sequel we will add a subscript A or B to the amplitudes and wave vectors in order to indicate to which domain they belong.

We consider a unitary superposition of all right-traveling envelope functions in A that enters the semiconductor at $x = 0$. Again we shall multiply these n incoming envelope waves with the coefficients $\omega_1, \ldots, \omega_n$. At $x = x_1$ this superposition of envelope functions is partly reflected. On the other hand, we expect a superposition of transmitted envelope functions in the domain $[x_2, L]$ that leaves the semiconductor at $x = L$. Thus, the envelope function reads

$$\mathbf{F}(x) = \begin{cases} \mathbf{F}_{A_1}(x) & \text{if } x \in [0, x_1), \\ \mathbf{F}_B(x) & \text{if } x \in [x_1, x_2), \\ \mathbf{F}_{A_2}(x) & \text{if } x \in [x_2, L], \end{cases} \tag{8.88}$$

with

$$\mathbf{F}_{A_1}(x) = \sum_{l=1}^{n} \omega_l \hat{\mathbf{F}}_A(\hat{k}_{A,l}^+) \mathrm{e}^{\mathrm{i}\hat{k}_{A,l}^+ x} + \sum_{l=1}^{n} \hat{r}_l \hat{\mathbf{F}}_A(\hat{k}_{A,l}^-) \mathrm{e}^{\mathrm{i}\hat{k}_{A,l}^- x} + \sum_{l=1}^{d-n} \check{r}_l \hat{\mathbf{F}}_A(\check{k}_{A,l}^-) \mathrm{e}^{\mathrm{i}\check{k}_{A,l}^- x},$$

$$\mathbf{F}_B(x) = \sum_{l=1}^{n} \hat{a}_l \hat{\mathbf{F}}_B(\hat{k}_{B,l}^+) \mathrm{e}^{\mathrm{i}\hat{k}_{B,l}^+ x} + \sum_{l=1}^{d-n} \check{a}_l \hat{\mathbf{F}}_B(\check{k}_{B,l}^+) \mathrm{e}^{\mathrm{i}\check{k}_{B,l}^+ x}$$
$$+ \sum_{l=1}^{n} \hat{b}_l \hat{\mathbf{F}}_B(\hat{k}_{B,l}^-) \mathrm{e}^{\mathrm{i}\hat{k}_{B,l}^- x} + \sum_{l=1}^{d-n} \check{b}_l \hat{\mathbf{F}}_B(\check{k}_{B,l}^-) \mathrm{e}^{\mathrm{i}\check{k}_{B,l}^- x},$$

$$\mathbf{F}_{A_2}(x) = \sum_{l=1}^{n} \hat{t}_l \hat{\mathbf{F}}_A(\hat{k}_{A,l}^+) \mathrm{e}^{\mathrm{i}\hat{k}_{A,l}^+ x} + \sum_{l=1}^{d-n} \check{t}_l \hat{\mathbf{F}}_A(\check{k}_{A,l}^+) \mathrm{e}^{\mathrm{i}\check{k}_{A,l}^+ x}.$$

We know that the solution (8.88) and its derivative are continuous, cf. [16]. In particular they are continuous at $x = x_1$ and $x = x_2$. Hence, we get a system of linear equations that can be written in the form $\mathbf{Qc} = \mathbf{s}$ with the vector

$$\mathbf{c} = (\hat{r}_1, \ldots, \hat{r}_n, \check{r}_1, \ldots, \check{r}_{d-n}, \hat{a}_1, \ldots, \hat{a}_n, \check{a}_1, \ldots, \check{a}_{d-n}, \hat{b}_1, \ldots, \hat{b}_n, \check{b}_1, \ldots, \check{b}_{d-n}, \hat{t}_1, \ldots, \hat{t}_n, \check{t}_1, \ldots, \check{t}_{d-n})^\top \in \mathbb{C}^{4d}$$

of the unknowns, the vector

$$\mathbf{s} = \begin{pmatrix} -\sum_{l=1}^{n} \omega_l \hat{\mathbf{F}}_A(\hat{k}^+_{A,l}) \mathrm{e}^{\mathrm{i}\hat{k}^+_{A,l} x_1} \\ \mathbf{0} \\ -\sum_{l=1}^{n} \mathrm{i}\hat{k}^+_{A,l} \omega_l \hat{\mathbf{F}}_A(\hat{k}^+_{A,l}) \mathrm{e}^{\mathrm{i}\hat{k}^+_{A,l} x_1} \\ \mathbf{0} \end{pmatrix} \in \mathbb{C}^{4d},$$

corresponding to the incoming waves and the coefficient matrix

$$\mathbf{Q} = \begin{pmatrix} \mathbf{Q}_{\hat{r}} & \mathbf{Q}_{\check{r}} & \mathbf{Q}_{\hat{a}} & \mathbf{Q}_{\check{a}} & \mathbf{Q}_{\hat{b}} & \mathbf{Q}_{\check{b}} & \mathbf{Q}_{\hat{t}} & \mathbf{Q}_{\check{t}} \end{pmatrix} \in \mathbb{C}^{4d \times 4d}.$$

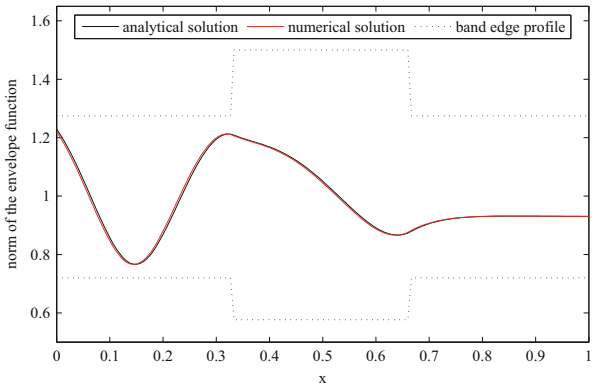
The matrices $\mathbf{Q}_{\hat{r}}, \mathbf{Q}_{\hat{a}}, \mathbf{Q}_{\hat{b}}, \mathbf{Q}_{\hat{t}} \in \mathbb{C}^{4d \times n}$ and $\mathbf{Q}_{\check{r}}, \mathbf{Q}_{\check{a}}, \mathbf{Q}_{\check{b}}, \mathbf{Q}_{\check{t}} \in \mathbb{C}^{4d \times (d-n)}$ are given in the appendix.

We shall not prove mathematically that the matrix $\mathbf{Q}$ is regular. Instead we point out that a singular matrix $\mathbf{Q}$ implies that the homogeneous case of the system of linear equations has a nonzero solution. Thus, there can exist envelope waves inside the computational domain without the existence of an incoming envelope wave which is a physical contradiction. We note that in our particular example the matrix $\mathbf{Q}$ is in fact regular and hence, the unknown coefficients $\hat{r}_p, \check{r}_q, \hat{a}_p, \check{a}_q, \hat{b}_p, \check{b}_q, \hat{t}_p, \check{t}_q$, with $p = 1, \ldots, n$ and $q = 1, \ldots, d - n$ are defined uniquely.

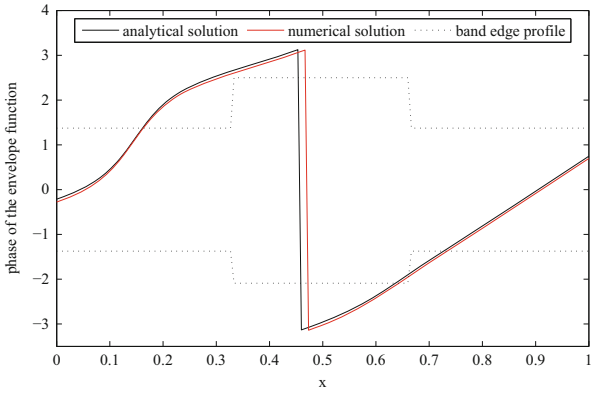
Now let us compare the analytical and numerical solutions of the quantum well structure. Figure 8.4 shows the norms and phases of the analytical and numerical solutions as well as a schematic view of the band edge profile.

As expected, we do not observe any spurious oscillations. Figure 8.3(bottom) illustrates that for the chosen energy $E = 1{,}500$ there exists a unique positive discrete wave vector in $[0, \pi/h]$. However, we observe a small phase error. This error decreases for smaller step sizes.

In Fig. 8.5 the analytical and numerical transmission coefficients are plotted against the energy E. As before, the step size $h = 1/150$ is used. Since the curve of the numerical transmission coefficient coincides with the curve of the analytical transmission coefficient for the used level of detail in Fig. 8.5 (top), only the analytical transmission coefficient is plotted. We observe that the qualitative



(a) Norm of the analytical and numerical solutions.

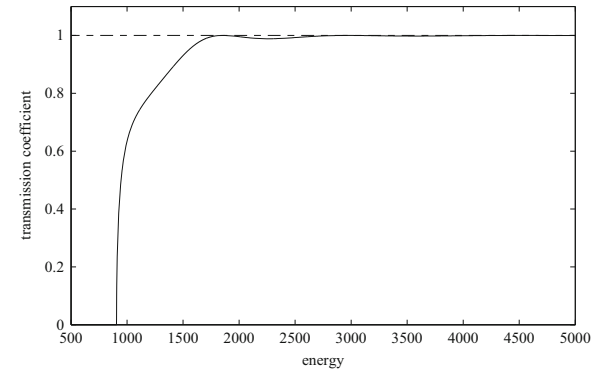


(b) Phases of the analytical and numerical solutions.

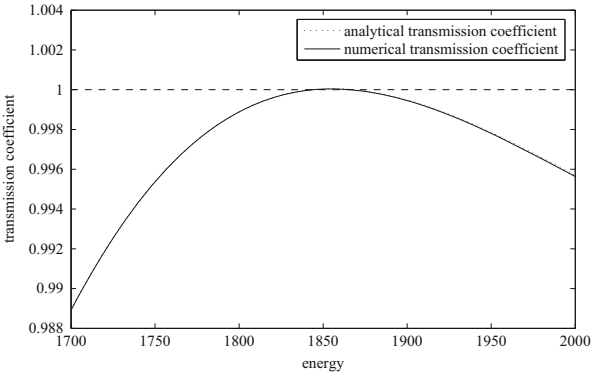
Fig. 8.4 Comparison of the analytical solution (*black*) and the numerical solution (*red*) of the quantum well structure for a step size $h = 1/150$, an energy $E = 1{,}500$. The *dotted line* indicates schematically the band end profile. The top figure shows the norm and the bottom figure the phases of the analytical and numerical solutions

behavior of the transmission coefficient of this particular quantum well structure is similar to the behavior of the transmission coefficient of the single barrier examples in the previous chapters. Note that the first resonance is located at $E \approx 1{,}856$.

Finally, we want to investigate the discrete L^2-error of the numerical scheme. Recall that we have to solve the optimization problem (8.87). Recall that the L^2-error decayed like $\mathcal{O}(h^2)$ in the free scattering state example, which confirmed the formal order of the numerical scheme. In the quantum well example, however, we observe that the numerical scheme is of order one only.



(a) Analytical transmission coefficient.



(b) Analytical (dotted line) and numerical (solid line) transmission coefficients near the first resonance at $E \approx 1856$.

Fig. 8.5 Analytical (*dotted line*) and numerical (*solid line*) transmission coefficients of the quantum well structure for a step size $h = 1/150$. The bottom figure shows the results near the first resonance at $E \approx 1{,}856$

8.4 Conclusion

In this chapter we derived DTBCs for stationary MEMAs. We first solved the continuous exterior problem and derived elementary solutions in the exterior domains and defined the TBCs. After discretizing the underlying BVP and solving the discrete exterior problem, we used the discrete elementary solutions in the exterior domains to derive DTBCs. This fully discrete approach results in reflection-free boundary conditions, while an ad-hoc discretization of the TBCs leads to spurious oscillations of the numerical solution. We tested the numerical schemes and the DTBCs in examples for which an analytical solution can be derived, i.e. for semiconductor nanostructures with piecewise constant band edges.

We reviewed DTBCs for the scalar Schrödinger equation, i.e. the single-band effective mass approximation, and analyzed alternative finite difference schemes.

Considering the numerical results we point out that the Mickens FDS is the most promising FDS among the FDSs we introduced for the scalar Schrödinger equation. If the potential is constant the Mickens FDS is an exact FDS and if the potential is not constant it is of order $\mathcal{O}(h^2)$. Which is the best possible order of convergence of the FDSs we introduced since the Numerov FDS which is formally of order $\mathcal{O}(h^4)$ is in fact also only $\mathcal{O}(h^2)$ since it requires the potential to be in $C^2(0, L)$. While the standard FDS and the Numerov FDS are also applicable, the combined Numerov-Mickens FDS, however, leads to significant errors.

We introduced the general d-band $\mathbf{k}\cdot\mathbf{p}$-model, developed the corresponding TBCs and DTBCs, tested them numerically and pointed out that these BCs depend on the choice of the elementary solutions in the exterior domain. It turned out that the numerical scheme when applied to an example with discontinuous band edges is at most of order one.

A topic of future research is the comparison of the introduced methods to solve MEMAs numerically with other methods, such as the transfer matrix method [24, 25] as well as the R-matrix method [27]. In particular, a comparison of these methods is of interest when the analytical solution cannot be derived, such as a quantum barrier structure with added bias.

Simulations of quantum cascade lasers are currently an extensively discussed topic [9]. The current density and the optical gain of quantum cascade lasers can be computed when the envelope functions are known. Based on the DTBCs for MEMAs, developed in this chapter, we plan to perform a fully discrete analysis of these simulations compared to the approach in [21], where an ad-hoc discretization of the TBCs was used.

Appendix

Definition of the Coefficient Matrix $\mathbf{Q}$

The matrices $\mathbf{Q}_{\hat{r}}, \mathbf{Q}_{\hat{a}}, \mathbf{Q}_{\hat{b}}, \mathbf{Q}_{\hat{t}} \in \mathbb{C}^{4d\times n}$ and $\mathbf{Q}_{\check{r}}, \mathbf{Q}_{\check{a}}, \mathbf{Q}_{\check{b}}, \mathbf{Q}_{\check{t}} \in \mathbb{C}^{4d\times(d-n)}$ building the coefficient matrix

$$\mathbf{Q} = \begin{pmatrix} \mathbf{Q}_{\hat{r}} & \mathbf{Q}_{\check{r}} & \mathbf{Q}_{\hat{a}} & \mathbf{Q}_{\check{a}} & \mathbf{Q}_{\hat{b}} & \mathbf{Q}_{\check{b}} & \mathbf{Q}_{\hat{t}} & \mathbf{Q}_{\check{t}} \end{pmatrix} \in \mathbb{C}^{4d\times 4d}$$

of the single potential barrier problem in Sect. 8.3 are defined by

$$\mathbf{Q}_{\hat{r}} = \begin{pmatrix} \hat{\mathbf{F}}_A(\hat{k}^-_{A,1})\mathrm{e}^{\mathrm{i}\hat{k}^-_{A,1}x_1} & \cdots & \hat{\mathbf{F}}_A(\hat{k}^-_{A,n})\mathrm{e}^{\mathrm{i}\hat{k}^-_{A,n}x_1} \\ \mathbf{0} & \cdots & \mathbf{0} \\ \mathrm{i}\hat{k}^-_{A,1}\hat{\mathbf{F}}_A(\hat{k}^-_{A,1})\mathrm{e}^{\mathrm{i}\hat{k}^-_{A,1}x_1} & \cdots & \mathrm{i}\hat{k}^-_{A,n}\hat{\mathbf{F}}_A(\hat{k}^-_{A,n})\mathrm{e}^{\mathrm{i}\hat{k}^-_{A,n}x_1} \\ \mathbf{0} & \cdots & \mathbf{0} \end{pmatrix}$$

$$\mathbf{Q}_{\check{r}} = \begin{pmatrix} \hat{\mathbf{F}}_A(\check{k}^-_{A,1})\mathrm{e}^{\mathrm{i}\check{k}^-_{A,1}x_1} & \cdots & \hat{\mathbf{F}}_A(\check{k}^-_{A,d-n})\mathrm{e}^{\mathrm{i}\check{k}^-_{A,d-n}x_1} \\ \mathbf{0} & \cdots & \mathbf{0} \\ \mathrm{i}\check{k}^-_{A,1}\hat{\mathbf{F}}_A(\check{k}^-_{A,1})\mathrm{e}^{\mathrm{i}\check{k}^-_{A,1}x_1} & \cdots & \mathrm{i}\check{k}^-_{A,d-n}\hat{\mathbf{F}}_A(\check{k}^-_{A,d-n})\mathrm{e}^{\mathrm{i}\check{k}^-_{A,d-n}x_1} \\ \mathbf{0} & \cdots & \mathbf{0} \end{pmatrix}$$

$$\mathbf{Q}_{\hat{a}} = \begin{pmatrix} -\hat{\mathbf{F}}_B(\hat{k}^+_{B,1})\mathrm{e}^{\mathrm{i}\hat{k}^+_{B,1}x_1} & \cdots & -\hat{\mathbf{F}}_B(\hat{k}^+_{B,n})\mathrm{e}^{\mathrm{i}\hat{k}^+_{B,n}x_1} \\ \hat{\mathbf{F}}_B(\hat{k}^+_{B,1})\mathrm{e}^{\mathrm{i}\hat{k}^+_{B,1}x_2} & \cdots & \hat{\mathbf{F}}_B(\hat{k}^+_{B,n})\mathrm{e}^{\mathrm{i}\hat{k}^+_{B,n}x_2} \\ -\mathrm{i}\hat{k}^+_{B,1}\hat{\mathbf{F}}_B(\hat{k}^+_{B,1})\mathrm{e}^{\mathrm{i}\hat{k}^+_{B,1}x_1} & \cdots & -\mathrm{i}\hat{k}^+_{B,n}\hat{\mathbf{F}}_B(\hat{k}^+_{B,n})\mathrm{e}^{\mathrm{i}\hat{k}^+_{B,n}x_1} \\ \mathrm{i}\hat{k}^+_{B,1}\hat{\mathbf{F}}_B(\hat{k}^+_{B,1})\mathrm{e}^{\mathrm{i}\hat{k}^+_{B,1}x_2} & \cdots & \mathrm{i}\hat{k}^+_{B,n}\hat{\mathbf{F}}_B(\hat{k}^+_{B,n})\mathrm{e}^{\mathrm{i}\hat{k}^+_{B,n}x_2} \end{pmatrix},$$

$$\mathbf{Q}_{\check{a}} = \begin{pmatrix} -\hat{\mathbf{F}}_B(\check{k}^+_{B,1})\mathrm{e}^{\mathrm{i}\check{k}^+_{B,1}x_1} & \cdots & -\hat{\mathbf{F}}_B(\check{k}^+_{B,d-n})\mathrm{e}^{\mathrm{i}\check{k}^+_{B,d-n}x_1} \\ \hat{\mathbf{F}}_B(\check{k}^+_{B,1})\mathrm{e}^{\mathrm{i}\check{k}^+_{B,1}x_2} & \cdots & \hat{\mathbf{F}}_B(\check{k}^+_{B,d-n})\mathrm{e}^{\mathrm{i}\check{k}^+_{B,d-n}x_2} \\ -\mathrm{i}\check{k}^+_{B,1}\hat{\mathbf{F}}_B(\check{k}^+_{B,1})\mathrm{e}^{\mathrm{i}\check{k}^+_{B,1}x_1} & \cdots & -\mathrm{i}\check{k}^+_{B,d-n}\hat{\mathbf{F}}_B(\check{k}^+_{B,d-n})\mathrm{e}^{\mathrm{i}\check{k}^+_{B,d-n}x_1} \\ \mathrm{i}\check{k}^+_{B,1}\hat{\mathbf{F}}_B(\check{k}^+_{B,1})\mathrm{e}^{\mathrm{i}\check{k}^+_{B,1}x_2} & \cdots & \mathrm{i}\check{k}^+_{B,d-n}\hat{\mathbf{F}}_B(\check{k}^+_{B,d-n})\mathrm{e}^{\mathrm{i}\check{k}^+_{B,d-n}x_2} \end{pmatrix}$$

$$\mathbf{Q}_{\hat{b}} = \begin{pmatrix} -\hat{\mathbf{F}}_B(\hat{k}^-_{B,1})\mathrm{e}^{\mathrm{i}\hat{k}^-_{B,1}x_1} & \cdots & -\hat{\mathbf{F}}_B(\hat{k}^-_{B,n})\mathrm{e}^{\mathrm{i}\hat{k}^-_{B,n}x_1} \\ \hat{\mathbf{F}}_B(\hat{k}^-_{B,1})\mathrm{e}^{\mathrm{i}\hat{k}^-_{B,1}x_2} & \cdots & \hat{\mathbf{F}}_B(\hat{k}^-_{B,n})\mathrm{e}^{\mathrm{i}\hat{k}^-_{B,n}x_2} \\ -\mathrm{i}\hat{k}^-_{B,1}\hat{\mathbf{F}}_B(\hat{k}^-_{B,1})\mathrm{e}^{\mathrm{i}\hat{k}^-_{B,1}x_1} & \cdots & -\mathrm{i}\hat{k}^-_{B,n}\hat{\mathbf{F}}_B(\hat{k}^-_{B,n})\mathrm{e}^{\mathrm{i}\hat{k}^-_{B,n}x_1} \\ \mathrm{i}\hat{k}^-_{B,1}\hat{\mathbf{F}}_B(\hat{k}^-_{B,1})\mathrm{e}^{\mathrm{i}\hat{k}^-_{B,1}x_2} & \cdots & \mathrm{i}\hat{k}^-_{B,n}\hat{\mathbf{F}}_B(\hat{k}^-_{B,n})\mathrm{e}^{\mathrm{i}\hat{k}^-_{B,n}x_2} \end{pmatrix},$$

$$\mathbf{Q}_{\check{b}} = \begin{pmatrix} -\hat{\mathbf{F}}_B(\check{k}^-_{B,1})\mathrm{e}^{\mathrm{i}\check{k}^-_{B,1}x_1} & \cdots & -\hat{\mathbf{F}}_B(\check{k}^-_{B,d-n})\mathrm{e}^{\mathrm{i}\check{k}^-_{B,d-n}x_1} \\ \hat{\mathbf{F}}_B(\check{k}^-_{B,1})\mathrm{e}^{\mathrm{i}\check{k}^-_{B,1}x_2} & \cdots & \hat{\mathbf{F}}_B(\check{k}^-_{B,d-n})\mathrm{e}^{\mathrm{i}\check{k}^-_{B,d-n}x_2} \\ -\mathrm{i}\check{k}^-_{B,1}\hat{\mathbf{F}}_B(\check{k}^-_{B,1})\mathrm{e}^{\mathrm{i}\check{k}^-_{B,1}x_1} & \cdots & -\mathrm{i}\check{k}^-_{B,d-n}\hat{\mathbf{F}}_B(\check{k}^-_{B,d-n})\mathrm{e}^{\mathrm{i}\check{k}^-_{B,d-n}x_1} \\ \mathrm{i}\check{k}^-_{B,1}\hat{\mathbf{F}}_B(\check{k}^-_{B,1})\mathrm{e}^{\mathrm{i}\check{k}^-_{B,1}x_2} & \cdots & \mathrm{i}\check{k}^-_{B,d-n}\hat{\mathbf{F}}_B(\check{k}^-_{B,d-n})\mathrm{e}^{\mathrm{i}\check{k}^-_{B,d-n}x_2} \end{pmatrix},$$

$$\mathbf{Q}_{\hat{t}} = \begin{pmatrix} \mathbf{0} & \cdots & \mathbf{0} \\ \hat{\mathbf{F}}_A(\hat{k}^+_{A,1})\mathrm{e}^{\mathrm{i}\hat{k}^+_{A,1}x_2} & \cdots & \hat{\mathbf{F}}_A(\hat{k}^+_{A,n})\mathrm{e}^{\mathrm{i}\hat{k}^+_{A,n}x_2} \\ \mathbf{0} & \cdots & \mathbf{0} \\ \mathrm{i}\hat{k}^+_{A,1}\hat{\mathbf{F}}_A(\hat{k}^+_{A,1})\mathrm{e}^{\mathrm{i}\hat{k}^+_{A,1}x_2} & \cdots & \mathrm{i}\hat{k}^+_{A,n}\hat{\mathbf{F}}_A(\hat{k}^+_{A,n})\mathrm{e}^{\mathrm{i}\hat{k}^+_{A,n}x_2} \end{pmatrix},$$

$$\mathbf{Q}_{\check{t}} = \begin{pmatrix} \mathbf{0} & \cdots & \mathbf{0} \\ -\hat{\mathbf{F}}_A(\check{k}^+_{A,1})\mathrm{e}^{\mathrm{i}\check{k}^+_{A,1}x_2} & \cdots & -\hat{\mathbf{F}}_A(\check{k}^+_{A,d-n})\mathrm{e}^{\mathrm{i}\check{k}^+_{A,d-n}x_2} \\ \mathbf{0} & \cdots & \mathbf{0} \\ -\mathrm{i}\check{k}^+_{A,1}\hat{\mathbf{F}}_A(\check{k}^+_{A,1})\mathrm{e}^{\mathrm{i}\check{k}^+_{A,1}x_2} & \cdots & -\mathrm{i}\check{k}^+_{A,d-n}\hat{\mathbf{F}}_A(\check{k}^+_{A,d-n})\mathrm{e}^{\mathrm{i}\check{k}^+_{A,d-n}x_2} \end{pmatrix}.$$

References

1. X. Antoine, A. Arnold, C. Besse, M. Ehrhardt, A. Schädle, A review of transparent and artificial boundary conditions techniques for linear and nonlinear Schrödinger equations. Commun. Comput. Phys. **4**, 729–796 (2008)
2. X. Antoine, C. Besse, M. Ehrhardt, P. Klein, Modeling boundary conditions for solving stationary Schrödinger equations. Preprint 10/04, University of Wuppertal, February 2010.
3. A. Arnold, M. Ehrhardt, I. Sofronov, Discrete transparent boundary conditions for the Schrödinger equation: fast calculation, approximation, and stability. Commun. Math. Sci. **1**, 501–556 (2003)
4. A. Arnold, Numerically absorbing boundary conditions for quantum evolution equations. VLSI Design **6**, 313–319 (1998)
5. A. Arnold, Mathematical concepts of open quantum boundary conditions. Trans. Theory Stat. Phys. **30**, 561–584 (2001)
6. U. Bandelow, H.-Chr. Kaiser, Th. Koprucki, J. Rehberg, Spectral properties of **k·p** Schrödinger operators in one space dimension Numer. Funct. Anal. Optimization **21**, 379–409 (2000)
7. N. Ben Abdallah, P. Degond, P.A. Markowich, On a one-dimensional Schrödinger-Poisson scattering model. ZAMP **48**, 135–155 (1997)
8. N. Ben Abdallah, J. Kefi-Ferhane, Mathematical analysis of the two-band Schrödinger model. Math. Meth. Appl. Sci. **31**, 1131–1151 (2008)
9. S. Birner, T. Kubis, P. Vogl, Simulation of quantum cascade lasers – optimizing laser performance. Photonik International **2**, 60–63 (2008)
10. R. Chen, Z. Xu, L. Sun, Finite-difference scheme to solve Schrödinger equations. Phys. Review E **47**, 3799–3802 (1993)
11. M. Ehrhardt, Discrete artificial boundary conditions, Ph.D. dissertation, Technische Universität Berlin (2001)
12. M. Ehrhardt, A. Arnold, Discrete transparent boundary conditions for the Schrödinger equation. Riv. Matem. Univ. di Parma **6**, 57–108 (2001)
13. P. Enders, M. Woerner, Exact 4×4 block diagonalization of the eight-band **k·p** Hamiltonian matrix for the tetrahedral semiconductors and its application to strained quantum wells. Semicond. Sci. Technol. **11**, 983–988 (1996)
14. P. Klein, X. Antoine, C. Besse, M. Ehrhardt, Absorbing boundary conditions for solving N-dimensional stationary Schrödinger equations with unbounded potentials and nonlinearities. Commun. Comput. Phys. **10**, 1280–1304 (2011)
15. D. Klindworth, Discrete transparent boundary conditions for multiband effective mass approximations, Diploma Thesis, Technische Universität Berlin (2009)
16. Th. Koprucki, Zu **k·p**-Schrödingeroperatoren, Ph.D. dissertation, Freie Universität Berlin (2008)
17. C. Lent, D. Kirkner, The quantum transmitting boundary method. J. Appl. Phys. **67**, 6353–6359 (1990)
18. P. Matus, Exact difference schemes for time-dependent problems. Comput. Meth. Appl. Math. **5**, 422–448 (2005)
19. R.E. Mickens, *Difference Equations: Theory and Applications* (Van Nostrand Reinhold, New York, 1990, 2nd ed.)
20. R.E. Mickens, Novel explicit finite-difference schemes for time-dependent Schrödinger equations. Comput. Phys. Commun. **63**, 203–208 (1991)
21. G. Milovanovic, O. Baumgartner, H. Kosina, Simulation of quantum cascade lasers using Robin boundary conditions. in: 9th International Conference on Numerical Simulation of Optoelectronic Devices, Gwangju Institute of Science and Technology, 2009.
22. C.A. Moyer, Numerical solution of the stationary state Schrödinger equation using discrete transparent boundary conditions. Comput. Sci. Engin. **8**, 32–40 (2006)
23. C. Negulescu, Numerical analysis of a multiscale finite element scheme for the resolution of the stationary Schrödinger equation. Numerische Mathematik **108**, 625–652 (2008)

24. R. Pérez-Alvarez, H. Rodriguez-Coppola, Transfer matrix in 1D Schrödinger problems with constant and position-dependent mass. Phys. Stat. Sol. (b) **145**, 493–500 (1988)
25. R. Pérez-Alvarez, H. Rodriguez-Coppola, V.R. Velasco, F. Garcia-Moliner, A study of the matching problem using transfer matrices. J. Phys. C: Solid State Phys. **21**, 2197–2206 (1988)
26. T.E. Simos, P.S. Williams, On finite difference methods for the solution of the Schrödinger equation. Computers & Chemistry **23**, 513–554 (1999)
27. U. Wulf, J. Kucera, P.N. Racec, E. Sigmund, Transport through quantum systems in the R-matrix formalism. Phys. Rev. **58**, 16209–16220 (1998)
28. A. Zisowsky, Discrete transparent boundary conditions for systems of evolution equations, Ph.D. dissertation, Technische Universität Berlin (2003)
29. A. Zisowsky, A. Arnold, M. Ehrhardt, Th. Koprucki, Discrete transparent boundary conditions for transient **k**·**p**-Schrödinger equations with application to quantum-heterostructures. J. Appl. Math. Mech. (ZAMM) **85**, 793–805 (2005)
30. A. Zlotnik, I. Zlotnik, Finite element method with discrete transparent boundary conditions for the time-dependent 1D Schrödinger equation. to appear in: Kinetic and Related Models (2013)
31. A. Zlotnik, I. Zlotnik, Finite element method with discrete transparent boundary conditions for the one-dimensional non-stationary Schrödinger equation. Doklady Mathematics **86**, 750–755 (2012)

Editorial Policy

1. Volumes in the following three categories will be published in LNCSE:

i) Research monographs
ii) Tutorials
iii) Conference proceedings

Those considering a book which might be suitable for the series are strongly advised to contact the publisher or the series editors at an early stage.

2. Categories i) and ii). Tutorials are lecture notes typically arising via summer schools or similar events, which are used to teach graduate students. These categories will be emphasized by Lecture Notes in Computational Science and Engineering. **Submissions by interdisciplinary teams of authors are encouraged.** The goal is to report new developments – quickly, informally, and in a way that will make them accessible to non-specialists. In the evaluation of submissions timeliness of the work is an important criterion. Texts should be well-rounded, well-written and reasonably self-contained. In most cases the work will contain results of others as well as those of the author(s). In each case the author(s) should provide sufficient motivation, examples, and applications. In this respect, Ph.D. theses will usually be deemed unsuitable for the Lecture Notes series. Proposals for volumes in these categories should be submitted either to one of the series editors or to Springer-Verlag, Heidelberg, and will be refereed. A provisional judgement on the acceptability of a project can be based on partial information about the work: a detailed outline describing the contents of each chapter, the estimated length, a bibliography, and one or two sample chapters – or a first draft. A final decision whether to accept will rest on an evaluation of the completed work which should include

– at least 100 pages of text;
– a table of contents;
– an informative introduction perhaps with some historical remarks which should be accessible to readers unfamiliar with the topic treated;
– a subject index.

3. Category iii). Conference proceedings will be considered for publication provided that they are both of exceptional interest and devoted to a single topic. One (or more) expert participants will act as the scientific editor(s) of the volume. They select the papers which are suitable for inclusion and have them individually refereed as for a journal. Papers not closely related to the central topic are to be excluded. Organizers should contact the Editor for CSE at Springer at the planning stage, see *Addresses* below.

In exceptional cases some other multi-author-volumes may be considered in this category.

4. Only works in English will be considered. For evaluation purposes, manuscripts may be submitted in print or electronic form, in the latter case, preferably as pdf- or zipped ps-files. Authors are requested to use the LaTeX style files available from Springer at http://www.springer.com/authors/book+authors/helpdesk?SGWID=0-1723113-12-971304-0 (Click on Templates → LaTeX → monographs or contributed books).

For categories ii) and iii) we strongly recommend that all contributions in a volume be written in the same LaTeX version, preferably LaTeX2e. Electronic material can be included if appropriate. Please contact the publisher.

Careful preparation of the manuscripts will help keep production time short besides ensuring satisfactory appearance of the finished book in print and online.

5. The following terms and conditions hold. Categories i), ii) and iii):

Authors receive 50 free copies of their book. No royalty is paid.
Volume editors receive a total of 50 free copies of their volume to be shared with authors, but no royalties.

Authors and volume editors are entitled to a discount of 33.3 % on the price of Springer books purchased for their personal use, if ordering directly from Springer.

6. Springer secures the copyright for each volume.

Addresses:

Timothy J. Barth
NASA Ames Research Center
NAS Division
Moffett Field, CA 94035, USA
barth@nas.nasa.gov

Michael Griebel
Institut für Numerische Simulation
der Universität Bonn
Wegelerstr. 6
53115 Bonn, Germany
griebel@ins.uni-bonn.de

David E. Keyes
Mathematical and Computer Sciences
and Engineering
King Abdullah University of Science
and Technology
P.O. Box 55455
Jeddah 21534, Saudi Arabia
david.keyes@kaust.edu.sa

and

Department of Applied Physics
and Applied Mathematics
Columbia University
500 W. 120 th Street
New York, NY 10027, USA
kd2112@columbia.edu

Risto M. Nieminen
Department of Applied Physics
Aalto University School of Science
and Technology
00076 Aalto, Finland
risto.nieminen@aalto.fi

Dirk Roose
Department of Computer Science
Katholieke Universiteit Leuven
Celestijnenlaan 200A
3001 Leuven-Heverlee, Belgium
dirk.roose@cs.kuleuven.be

Tamar Schlick
Department of Chemistry
and Courant Institute
of Mathematical Sciences
New York University
251 Mercer Street
New York, NY 10012, USA
schlick@nyu.edu

Editor for Computational Science
and Engineering at Springer:
Martin Peters
Springer-Verlag
Mathematics Editorial IV
Tiergartenstrasse 17
69121 Heidelberg, Germany
martin.peters@springer.com

Lecture Notes in Computational Science and Engineering

1. D. Funaro, *Spectral Elements for Transport-Dominated Equations.*
2. H.P. Langtangen, *Computational Partial Differential Equations.* Numerical Methods and Diffpack Programming.
3. W. Hackbusch, G. Wittum (eds.), *Multigrid Methods V.*
4. P. Deuflhard, J. Hermans, B. Leimkuhler, A.E. Mark, S. Reich, R.D. Skeel (eds.), *Computational Molecular Dynamics: Challenges, Methods, Ideas.*
5. D. Kröner, M. Ohlberger, C. Rohde (eds.), *An Introduction to Recent Developments in Theory and Numerics for Conservation Laws.*
6. S. Turek, *Efficient Solvers for Incompressible Flow Problems.* An Algorithmic and Computational Approach.
7. R. von Schwerin, ***M**ulti **B**ody **S**ystem **SIM**ulation.* Numerical Methods, Algorithms, and Software.
8. H.-J. Bungartz, F. Durst, C. Zenger (eds.), *High Performance Scientific and Engineering Computing.*
9. T.J. Barth, H. Deconinck (eds.), *High-Order Methods for Computational Physics.*
10. H.P. Langtangen, A.M. Bruaset, E. Quak (eds.), *Advances in Software Tools for Scientific Computing.*
11. B. Cockburn, G.E. Karniadakis, C.-W. Shu (eds.), *Discontinuous Galerkin Methods.* Theory, Computation and Applications.
12. U. van Rienen, *Numerical Methods in Computational Electrodynamics.* Linear Systems in Practical Applications.
13. B. Engquist, L. Johnsson, M. Hammill, F. Short (eds.), *Simulation and Visualization on the Grid.*
14. E. Dick, K. Riemslagh, J. Vierendeels (eds.), *Multigrid Methods VI.*
15. A. Frommer, T. Lippert, B. Medeke, K. Schilling (eds.), *Numerical Challenges in Lattice Quantum Chromodynamics.*
16. J. Lang, *Adaptive Multilevel Solution of Nonlinear Parabolic PDE Systems.* Theory, Algorithm, and Applications.
17. B.I. Wohlmuth, *Discretization Methods and Iterative Solvers Based on Domain Decomposition.*
18. U. van Rienen, M. Günther, D. Hecht (eds.), *Scientific Computing in Electrical Engineering.*
19. I. Babuška, P.G. Ciarlet, T. Miyoshi (eds.), *Mathematical Modeling and Numerical Simulation in Continuum Mechanics.*
20. T.J. Barth, T. Chan, R. Haimes (eds.), *Multiscale and Multiresolution Methods.* Theory and Applications.
21. M. Breuer, F. Durst, C. Zenger (eds.), *High Performance Scientific and Engineering Computing.*
22. K. Urban, *Wavelets in Numerical Simulation.* Problem Adapted Construction and Applications.
23. L.F. Pavarino, A. Toselli (eds.), *Recent Developments in Domain Decomposition Methods.*

24. T. Schlick, H.H. Gan (eds.), *Computational Methods for Macromolecules: Challenges and Applications.*

25. T.J. Barth, H. Deconinck (eds.), *Error Estimation and Adaptive Discretization Methods in Computational Fluid Dynamics*.

26. M. Griebel, M.A. Schweitzer (eds.), *Meshfree Methods for Partial Differential Equations*.

27. S. Müller, *Adaptive Multiscale Schemes for Conservation Laws*.

28. C. Carstensen, S. Funken, W. Hackbusch, R.H.W. Hoppe, P. Monk (eds.), *Computational Electromagnetics*.

29. M.A. Schweitzer, *A Parallel Multilevel Partition of Unity Method for Elliptic Partial Differential Equations*.

30. T. Biegler, O. Ghattas, M. Heinkenschloss, B. van Bloemen Waanders (eds.), *Large-Scale PDE-Constrained Optimization*.

31. M. Ainsworth, P. Davies, D. Duncan, P. Martin, B. Rynne (eds.), *Topics in Computational Wave Propagation*. Direct and Inverse Problems.

32. H. Emmerich, B. Nestler, M. Schreckenberg (eds.), *Interface and Transport Dynamics.* Computational Modelling.

33. H.P. Langtangen, A. Tveito (eds.), *Advanced Topics in Computational Partial Differential Equations.* Numerical Methods and Diffpack Programming.

34. V. John, *Large Eddy Simulation of Turbulent Incompressible Flows.* Analytical and Numerical Results for a Class of LES Models.

35. E. Bänsch (ed.), *Challenges in Scientific Computing - CISC 2002.*

36. B.N. Khoromskij, G. Wittum, *Numerical Solution of Elliptic Differential Equations by Reduction to the Interface.*

37. A. Iske, *Multiresolution Methods in Scattered Data Modelling.*

38. S.-I. Niculescu, K. Gu (eds.), *Advances in Time-Delay Systems.*

39. S. Attinger, P. Koumoutsakos (eds.), *Multiscale Modelling and Simulation.*

40. R. Kornhuber, R. Hoppe, J. Périaux, O. Pironneau, O. Wildlund, J. Xu (eds.), *Domain Decomposition Methods in Science and Engineering.*

41. T. Plewa, T. Linde, V.G. Weirs (eds.), *Adaptive Mesh Refinement – Theory and Applications.*

42. A. Schmidt, K.G. Siebert, *Design of Adaptive Finite Element Software.* The Finite Element Toolbox ALBERTA.

43. M. Griebel, M.A. Schweitzer (eds.), *Meshfree Methods for Partial Differential Equations II.*

44. B. Engquist, P. Lötstedt, O. Runborg (eds.), *Multiscale Methods in Science and Engineering.*

45. P. Benner, V. Mehrmann, D.C. Sorensen (eds.), *Dimension Reduction of Large-Scale Systems.*

46. D. Kressner, *Numerical Methods for General and Structured Eigenvalue Problems.*

47. A. Boriçi, A. Frommer, B. Joó, A. Kennedy, B. Pendleton (eds.), *QCD and Numerical Analysis III*.

48. F. Graziani (ed.), *Computational Methods in Transport.*

49. B. Leimkuhler, C. Chipot, R. Elber, A. Laaksonen, A. Mark, T. Schlick, C. Schütte, R. Skeel (eds.), *New Algorithms for Macromolecular Simulation.*

50. M. Bücker, G. Corliss, P. Hovland, U. Naumann, B. Norris (eds.), *Automatic Differentiation: Applications, Theory, and Implementations.*

51. A.M. Bruaset, A. Tveito (eds.), *Numerical Solution of Partial Differential Equations on Parallel Computers.*

52. K.H. Hoffmann, A. Meyer (eds.), *Parallel Algorithms and Cluster Computing.*

53. H.-J. Bungartz, M. Schäfer (eds.), *Fluid-Structure Interaction.*

54. J. Behrens, *Adaptive Atmospheric Modeling.*

55. O. Widlund, D. Keyes (eds.), *Domain Decomposition Methods in Science and Engineering XVI.*

56. S. Kassinos, C. Langer, G. Iaccarino, P. Moin (eds.), *Complex Effects in Large Eddy Simulations.*

57. M. Griebel, M.A Schweitzer (eds.), *Meshfree Methods for Partial Differential Equations III.*

58. A.N. Gorban, B. Kégl, D.C. Wunsch, A. Zinovyev (eds.), *Principal Manifolds for Data Visualization and Dimension Reduction.*

59. H. Ammari (ed.), *Modeling and Computations in Electromagnetics: A Volume Dedicated to Jean-Claude Nédélec.*

60. U. Langer, M. Discacciati, D. Keyes, O. Widlund, W. Zulehner (eds.), *Domain Decomposition Methods in Science and Engineering XVII.*

61. T. Mathew, *Domain Decomposition Methods for the Numerical Solution of Partial Differential Equations.*

62. F. Graziani (ed.), *Computational Methods in Transport: Verification and Validation.*

63. M. Bebendorf, *Hierarchical Matrices.* A Means to Efficiently Solve Elliptic Boundary Value Problems.

64. C.H. Bischof, H.M. Bücker, P. Hovland, U. Naumann, J. Utke (eds.), *Advances in Automatic Differentiation.*

65. M. Griebel, M.A. Schweitzer (eds.), *Meshfree Methods for Partial Differential Equations IV*.

66. B. Engquist, P. Lötstedt, O. Runborg (eds.), *Multiscale Modeling and Simulation in Science.*

67. I.H. Tuncer, Ü. Gülcat, D.R. Emerson, K. Matsuno (eds.), *Parallel Computational Fluid Dynamics 2007.*

68. S. Yip, T. Diaz de la Rubia (eds.), *Scientific Modeling and Simulations*.

69. A. Hegarty, N. Kopteva, E. O'Riordan, M. Stynes (eds.), *BAIL* 2008 – *Boundary and Interior Layers.*

70. M. Bercovier, M.J. Gander, R. Kornhuber, O. Widlund (eds.), *Domain Decomposition Methods in Science and Engineering XVIII.*

71. B. Koren, C. Vuik (eds.), *Advanced Computational Methods in Science and Engineering.*

72. M. Peters (ed.), *Computational Fluid Dynamics for Sport Simulation.*

73. H.-J. Bungartz, M. Mehl, M. Schäfer (eds.), *Fluid Structure Interaction II - Modelling, Simulation, Optimization.*

74. D. Tromeur-Dervout, G. Brenner, D.R. Emerson, J. Erhel (eds.), *Parallel Computational Fluid Dynamics 2008.*

75. A.N. Gorban, D. Roose (eds.), *Coping with Complexity: Model Reduction and Data Analysis.*

76. J.S. Hesthaven, E.M. Rønquist (eds.), *Spectral and High Order Methods for Partial Differential Equations.*

77. M. Holtz, *Sparse Grid Quadrature in High Dimensions with Applications in Finance and Insurance*.

78. Y. Huang, R. Kornhuber, O.Widlund, J. Xu (eds.), *Domain Decomposition Methods in Science and Engineering XIX*.

79. M. Griebel, M.A. Schweitzer (eds.), *Meshfree Methods for Partial Differential Equations V*.

80. P.H. Lauritzen, C. Jablonowski, M.A. Taylor, R.D. Nair (eds.), *Numerical Techniques for Global Atmospheric Models*.

81. C. Clavero, J.L. Gracia, F.J. Lisbona (eds.), *BAIL 2010 – Boundary and Interior Layers, Computational and Asymptotic Methods*.

82. B. Engquist, O. Runborg, Y.R. Tsai (eds.), *Numerical Analysis and Multiscale Computations*.

83. I.G. Graham, T.Y. Hou, O. Lakkis, R. Scheichl (eds.), *Numerical Analysis of Multiscale Problems*.

84. A. Logg, K.-A. Mardal, G. Wells (eds.), *Automated Solution of Differential Equations by the Finite Element Method*.

85. J. Blowey, M. Jensen (eds.), *Frontiers in Numerical Analysis - Durham 2010.*

86. O. Kolditz, U.-J. Gorke, H. Shao, W. Wang (eds.), *Thermo-Hydro-Mechanical-Chemical Processes in Fractured Porous Media - Benchmarks and Examples*.

87. S. Forth, P. Hovland, E. Phipps, J. Utke, A. Walther (eds.), *Recent Advances in Algorithmic Differentiation*.

88. J. Garcke, M. Griebel (eds.), *Sparse Grids and Applications.*

89. M. Griebel, M.A. Schweitzer (eds.), *Meshfree Methods for Partial Differential Equations VI.*

90. C. Pechstein, *Finite and Boundary Element Tearing and Interconnecting Solvers for Multiscale Problems*.

91. R. Bank, M. Holst, O. Widlund, J. Xu (eds.), *Domain Decomposition Methods in Science and Engineering XX*.

92. H. Bijl, D. Lucor, S. Mishra, C. Schwab (eds.), *Uncertainty Quantification in Computational Fluid Dynamics*.

93. M. Bader, H.-J. Bungartz, T. Weinzierl (eds.), *Advanced Computing*.

94. M. Ehrhardt, T. Koprucki (eds.), *Advanced Mathematical Models and Numerical Techniques for Multi-Band Effective Mass Approximations*.

95. M. Azaïez, H.E. Fekih, J.S. Hesthaven (eds.), *Spectral and High Order Methods for Partial Differential Equations ICOSAHOM 2012*.

96. F. Graziani, M.P. Desjarlais, R. Redmer, S.B. Trickey (eds.), *Frontiers and Challenges in Warm Dense Matter*.

97. J. Garcke, D. Pflüger (eds.), *Sparse Grids and Applications - Munich 2012*.

98. J. Erhel, M. Gander, L. Halpern, G. Pichot, T. Sassi, O. Widlund (eds.), *Domain Decomposition Methods in Science and Engineering XXI.*

99. R. Abgrall, H. Beaugendre, P.M. Congedo, C. Dobrzynski, V. Perrier, M. Ricchiuto (eds.), *High Order Nonlinear Numerical Methods for Evolutionary PDEs - HONOM 2013*.

100. M. Griebel, M.A. Schweitzer (eds.), *Meshfree Methods for Partial Differential Equations VII*.

101. R. Hoppe (ed.), *Optimization with PDE Constraints - ESF Networking Program 'OPTPDE'*.

102. S. Dahlke, W. Dahmen, M. Griebel, W. Hackbusch, K. Ritter, R. Schneider, C. Schwab, H. Yserentant (eds.), *Extraction of Quantifiable Information from Complex Systems.*

For further information on these books please have a look at our mathematics catalogue at the following URL: www.springer.com/series/3527

Monographs in Computational Science and Engineering

1. J. Sundnes, G.T. Lines, X. Cai, B.F. Nielsen, K.-A. Mardal, A. Tveito, *Computing the Electrical Activity in the Heart.*

For further information on this book, please have a look at our mathematics catalogue at the following URL: www.springer.com/series/7417

Texts in Computational Science and Engineering

1. H. P. Langtangen, *Computational Partial Differential Equations.* Numerical Methods and Diffpack Programming. 2nd Edition
2. A. Quarteroni, F. Saleri, P. Gervasio, *Scientific Computing with MATLAB and Octave.* 4th Edition
3. H. P. Langtangen, *Python Scripting for Computational Science*. 3rd Edition
4. H. Gardner, G. Manduchi, *Design Patterns for e-Science*.
5. M. Griebel, S. Knapek, G. Zumbusch, *Numerical Simulation in Molecular Dynamics*.
6. H. P. Langtangen, *A Primer on Scientific Programming with Python.* 4th Edition
7. A. Tveito, H. P. Langtangen, B. F. Nielsen, X. Cai, *Elements of Scientific Computing.*
8. B. Gustafsson, *Fundamentals of Scientific Computing.*
9. M. Bader, *Space-Filling Curves.*
10. M. Larson, F. Bengzon, *The Finite Element Method: Theory, Implementation and Applications.*
11. W. Gander, M. Gander, F. Kwok, *Scientific Computing: An Introduction using Maple and MATLAB.*

For further information on these books please have a look at our mathematics catalogue at the following URL: www.springer.com/series/5151