Physik-Department
Technische Universität München
Institut für Theoretische Physik
Lehrstuhl Univ.-Prof. Dr. P. Vogl

Semiconductor based quantum information devices:

Theory and simulations

Tobias Zibold

Vollständiger Abdruck der von der Fakultät für Physik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzender:	Univ.-Prof. Dr. Dirk Grundler
Prüfer der Dissertation:	1. Univ.-Prof. Dr. Peter Vogl 2. Univ.-Prof. Dr. Wilhelm Zwerger

Die Dissertation wurde am 03.04.2007 bei der Technischen Universität München eingereicht und durch die Fakultät für Physik am 21.05.2007 angenommen.

Information is physical.
– Rolf Landauer

Contents

Introduction **ix**

1 Simulation of semiconductor quantum information devices **1**
1.1 Introduction . . . 1
1.2 Theory . . . 2
1.2.1 Band structure . . . 2
1.2.2 Poisson equation . . . 10
1.2.3 Current equation . . . 13
1.2.4 Strain equation . . . 15
1.3 Summary . . . 18

2 Numerical techniques **19**
2.1 Introduction . . . 19
2.2 Box discretization . . . 20
2.3 Iterative solvers . . . 22
2.4 Solving linear systems of equations . . . 23
2.5 Solving large matrix eigenvalue problems . . . 25
2.6 Solving non-linear systems of equations . . . 27
2.7 Solving coupled systems of equations . . . 28
2.8 Summary . . . 32

3 Program structure **33**
3.1 Libraries . . . 33
3.2 Program flow . . . 35
3.3 Class hierarchy . . . 37
3.4 Input processing . . . 39
3.5 Geometry processing . . . 42
3.6 Summary . . . 43

4 Calculation of the charge carrier density by Brillouin zone integration **45**
4.1 Introduction . . . 45
4.2 A novel method for the two-dimensional Brillouin zone integration . . . 47
4.2.1 Integration method . . . 47
4.2.2 Irreducible wedge of the two-dimensional Brillouin zone . . . 51
4.2.3 Analysis of the convergence behavior . . . 52
4.3 A novel method for the charge density calculation in the 8-band $k \cdot p$ model . . . 56
4.4 Application to doping superlattices . . . 60

4.5 Summary and Outlook 64

5 A novel method for ballistic quantum transport in entangled two-particle systems 67

5.1 Introduction to ballistic quantum transport 67
5.2 Contact block reduction method 69
5.3 Novel method for ballistic quantum transport in entangled two-particle systems 72
5.4 Interaction matrix elements 74
5.5 Summary 75

6 Prediction of entanglement detection by I-V characteristics 77

6.1 Introduction 77
6.2 All-electric Mach-Zehnder interferometer 79
6.2.1 Introduction 79
6.2.2 Numerical Details 80
6.2.3 Results 80
6.2.4 Discussion 83
6.3 Entangled Mach-Zehnder double quantum dot device 86
6.3.1 Introduction 86
6.3.2 Numerical Details 86
6.3.3 Results 87
6.3.4 Discussion 89
6.4 Analytical model for the entangled Mach-Zehnder double quantum dot device . 90
6.4.1 Introduction 90
6.4.2 Analytical model 91
6.4.3 Relation between visibility and von Neumann entropy 93
6.4.4 Results and Discussion 95
6.5 Summary 96

7 Further Applications 99

7.1 Calculation of carrier transport through quantum dot molecules 99
7.2 Strain in AlAs quantum wells and cleaved-edge overgrown quantum wires 108

8 Summary and Outlook 113

A Grid definition 115

B Geometry processing 117

B.1 Semiellipsoids 117
B.2 Cones 120

C $k \cdot p$ and strain Hamiltonian for wurtzite 123

D Recombination and mobility models 125

E Alloys 129

E.1 Alloys of the type A_xB_{1-x} 129
E.2 Alloys of the type $A_xB_yC_{1-x-y}$ 130

E.3 Alloys of the type $A_xB_{1-x}C_yD_{1-y}$ . . . 131

F Analytical model: detailed calculations **133**
F.1 Transfer matrices . . . 133
F.2 Eigenvalues of the reduced density matrix . . . 137

Publication List **141**

Bibliography **143**

Danksagung **151**

Abstract

In this work, the theory of semiconductor single- and two-qubit quantum logic gates based on electrostatically defined quantum wires and quantum dots is studied. A concrete proposal for an all-electric Mach-Zehnder interferometer is worked out by detailed, numerical calculations of its electronic structure and ballistic quantum transport properties. It is shown that this Mach-Zehnder interferometer acts as a single-qubit gate for electrons propagating in either of two quantum wires. The proposal is based on an AlGaAs/GaAs heterostructure and includes the detailed three-dimensional geometry, material compositions, doping profiles, and bias voltages. The device is well within the reach of present-day experimental fabrication techniques. A novel quantum transport device is presented that acts as a two-qubit gate. It is based on the all-electric Mach-Zehnder interferometer and a double quantum dot. We show that this device allows for the controlled generation and detection of entanglement from DC I-V characteristics. The entanglement between the two qubits has been calculated for a fully three-dimensional model of the device. To this end, a new quantum transport method has been developed that allows one to take into account the Coulomb interaction between two electrons non-perturbatively. A dynamic, analytical model of the two-qubit device is proposed and an analytical relation is derived for quantifying the von Neumann entropy from the DC I-V characteristics.

A second topic of this work is the study of material aspects of semiconductor based quantum information processing. The ballistic tunneling current through an InAs/InP quantum dot molecule embedded in a resonant tunneling diode is calculated for a realistic, three-dimensional model of the molecule. We predict that a wealth of unique information on the size, inter-dot distance, lateral alignment, and electronic states can be obtained by measuring this tunneling current. Finally, the level ordering in AlAs cleaved-edge overgrown quantum wires and its influence on the transport properties is studied by calculating the electronic states charge self-consistently including strain, doping profiles, and piezoelectricity.

Zusammenfassung

In dieser Arbeit untersuchen wir die Theorie halbleiterbasierter 1- und 2-Qubit-Quantenlogikgatter aus Quantendrähten und Quantenpunkten. Mit Hilfe von detaillierten, numerischen Berechnungen der elektronischen Struktur sowie des ballistischen Quantentransportes haben wir einen konkreten Entwurf für ein rein elektrisches Mach-Zehnder Interferometer ausgearbeitet. Wir zeigen, dass dieses Mach-Zehnder Interferometer für Elektronen, die durch einen von zwei Quantendrähten propagieren, ein 1-Qubit-Quantenlogikgatter darstellt. Unserem Entwurf für das Bauelement liegt eine AlGaAs/GaAs Heterostruktur zugrunde. Im Einzelnen umfasst der Entwurf die detaillierte dreidimensionale Geometrie, Materialien, Dotierprofile und Gatterspannungen und sollte sich mit den heutigen experimentellen Techniken herstellen lassen. Auf der Grundlage dieses Mach-Zehnder Interferometers und eines Doppelquantenpunktes schlagen wir ein neues Quantentransport-Bauelement vor, mit dem sich ein 2-Qubit-Quantenlogikgatter realisieren lässt. Wir zeigen, dass die beiden Qubits in diesem Bauelement kontrolliert verschränkt werden können und dass diese Verschränkung über die Gleichstrom-Stromspannungskennlinie nachzuweisen ist. Zur Berechnung der Verschränkung wird ein vollständig dreidimensionales Modell des Bauelementes verwendet. Hierzu haben wir auch eine neue Quantentransportmethode entwickelt, mit der die Coulombwechselwirkung zwischen zwei Elektronen nichtperturbativ berücksichtigt werden kann. Des Weiteren wird ein dynamisches, analytisches Modell des 2-Qubit-Bauelementes erarbeitet und es wird eine analytische Beziehung hergeleitet, mit der die von Neumann Entropie aus der Gleichstrom-Stomspannungskennlinie bestimmt werden kann.

Ein zweites Thema dieser Arbeit ist die Untersuchung von Materialaspekten halbleiterbasierter Quanteninformationsverarbeitung. Hier haben wir den ballistischen Tunnelstrom durch ein in eine resonante Tunneldiode eingebettetes InAs/InP Quantenpunktmolekül berechnet, wobei das Molekül als realistisches, dreidimensionales Modell berücksichtigt wurde. Aus unseren Berechnungen folgt, dass die Messung des Tunnelstroms eine Vielzahl einzigartiger Daten über die Größe der Quantenpunkte, deren Abstand und laterale Ausrichtung sowie die elektronischen Zustände liefert. Im letzten Teil der Arbeit untersuchen wir die Anordnung der elektronischen Niveaus in CEO Quantendrähten aus Aluminiumarsenid und deren Einfluss auf die Transporteigenschaften. Hierzu berechnen wir ladungsselbstkonsistent die elektronischen Zustände unter Berücksichtigung von Verspannungen, Dotierprofilen und piezoelektrischen Ladungen.

Introduction

The power of today's computers is a consequence of the continuous miniaturization in the area of semiconductor based integrated logic and memory chips. These days, the ultimate end of this process can be anticipated by the arrival of the atomic length scale. However, the end of miniaturization is not the only limitation that information processing has to face. Rather there are computational problems that are believed to not have an efficient solution on classical computers. One of these problems is to find the prime factors of an integer. The realization that computers based on the principles of quantum mechanics are able to solve these problems efficiently has triggered an enormous interest and attempt to build such quantum computers. Up to now there is no consensus on the best approach for quantum computers. The requirements of long decoherence times and a high degree of scalability seem to be mutually exclusive.

In this thesis we focus on semiconductor based approaches for quantum computers that in general promise good scalability but suffer from short decoherence times. The topic is to study the physics of quantum logic gates build from semiconductor nanostructures by detailed, numerical calculations within realistic, three-dimensional models. The goal of these studies is to provide concrete proposals for the realization of semiconductor quantum logic gates including the detailed geometry, material compositions, doping profiles, and bias voltages. In this, our work differs fundamentally from the majority of approaches whose objectives are to identify the general possibilities for quantum information processing in semiconductors by means of simple models. These are essential in understanding the basic physical relations. For the interpretation of our results we therefore also rely on simple, analytical models. However, they are often not sufficient to support the experimentalists in a concrete experimental realization.

The physics of quantum computation poses new requirements on methods and tools for the simulation of semiconductor devices for quantum information processing. To mention the most important only, quantum mechanical many-body effects such as entanglement require including two-particle interactions non-perturbatively. Besides simulations, an important part of this thesis has therefore been the development of novel computational methods as well as the development of the new software packet next**nano**++ that has established the basis for the computationally demanding calculations.

This thesis is organized as follows: In the first chapter an overview on the theoretical background of simulating semiconductor based devices for quantum information processing is given. This includes the calculation of the electronic structure taking into account different materials, impurities, elastic strain, deformation potentials, piezo- and pyroelectric charges, as well as currents in a fully self-consistent manner. The second chapter summarizes the numerical methods for the solution of this computationally demanding problem. Here we begin with the discretization scheme for the various partial differential equations such as the Schrödinger equation, the Poisson equation, the current equation, and the strain equation. We then introduce the iterative methods for solving the resulting linear systems of equations and

large matrix eigenvalue problems. We conclude the chapter by presenting several approaches for the self-consistent solution of non-linear as well as coupled systems of equations.

The implementation of the presented physical models as well as the methods for their solution in the framework of an actual software packet is a task that easily surpasses the manpower of a single Ph.D. student. However, a modular approach as well as the use of modern object oriented programming techniques such as inheritance and generic programming in combination with the use of standard components allows for a straightforward breakdown of the overall problem. This is illustrated in Chapter 3.

In the following chapters we turn to the key elements of this work, namely the prediction and simulation of semiconductor based devices for quantum information processing. This also includes the development of novel methods for the calculation of specific properties of such devices as well as the development of simplified models for an in-depth analysis of the obtained results. It is well known that universal quantum computation requires single as well as two-qubit quantum gates. As an example of a single-qubit gate, we have analyzed a concrete quantum transport device built from two parallel quantum wires that are coupled by two potential wells (coupling windows). The two quantum wires define the qubit states for electrons that propagate through the wires. This device acts as an all-electric Mach-Zehnder interferometer and allows one to prepare the propagating electrons in an arbitrary linear combination of the two qubit states. We show theoretically that the device can be realized by state-of-the-art experimental techniques in a modulation doped AlGaAs/GaAs heterostructure. In particular, our analysis has identified the coupling windows as the most critical components of the Mach-Zehnder interferometer that require for a high degree of fabrication precision.

On the basis of the single-qubit gate we have proposed and analyzed a novel quantum transport device that acts as a two-qubit gate. Here, the idea is to couple the all-electric Mach-Zehnder interferometer via the Coulomb interaction to an electrostatically defined single electron double quantum dot. The basis of any non-trivial two-qubit quantum gate is entanglement. The calculation of the quantum properties of a two-qubit gate therefore requires methods that allow for the inclusion of many-body correlation effects such as entanglement. Thus we have developed a method for the calculation of the ballistic quantum transport in realistic, three-dimensional two-particle systems that includes the Coulomb interaction non-perturbatively. We have used this method to predict and analyze a concrete realization of the proposed two-qubit quantum transport device based on two vertically stacked two-dimensional electron gases in an AlGaAs/GaAs heterostructure. In particular, we could identify the optimum geometric arrangement of the Mach-Zehnder interferometer and the double quantum dot. To help in understanding the basic physics of the proposed device, we have developed a simplified analytical model. This model allows us to derive an explicit analytical relation between the visibility of the Mach-Zehnder interference pattern and the von Neumann entropy. Therefore, it is possible to determine the degree of entanglement from the DC I-V characteristics of the Mach-Zehnder interferometer.

While for the study of quantum logic gates the focus has been on the device physics, in Chapter 7 we turn towards the material aspects of semiconductor based quantum information processing and its characterization by current measurements. Quantum dots grown epitaxially by self-assembly are ideal candidates for the realization of spin as well as excitonic qubits. In addition, the strain induced growth of stacks of such quantum dots, also known as quantum dot molecules, offers the possibility to easily fabricate multiple coupled qubits. However, as only poor control over the growth condition is given, the individual quantum dots may vary in size, shape, alloy composition, inter-dot distance, and their lateral alignment. These properties

strongly affect the electronic structure and inter-dot coupling, but for the buried quantum dots cannot be determined by direct measurements. Optical spectroscopy is one approach for an indirect, *in-situ* measurement of the mentioned properties. We have performed detailed, numerical calculations of the ballistic tunneling current for realistic, three-dimensional models of double dot molecules. From these calculations we predict that measurements of the ballistic tunneling current provide a wealth of information on the geometry and the electronic states of quantum dot molecules and can therefore be used as an additional method for the *in-situ* characterization of quantum dot molecules. The second application in this chapter is the self-consistent calculation of the electronic states in AlAs cleaved-edge overgrown quantum wires including strain, doping profiles, and piezoelectricity. AlAs quantum wires are well suited for studying electronic correlation effects. The physics of electronic correlations, unintentional ones as well as intentional ones, is of great importance in quantum information processing. Here, our theoretical calculations have contributed considerably to the understanding of recent measurements of the transport properties of strained AlAs cleaved-edge overgrown quantum wires.

Chapter 1

Simulation of semiconductor quantum information devices

1.1 Introduction

The idea of quantum information processing dates back to Richard Feynman, who proposed in 1982 to use artificial quantum systems with specifically designed interactions to simulate and help in understanding the physics of real quantum systems [1]. The desire for quantum computers grew significantly since 1994, when Peter Shor showed that the important problem of finding the prime factors of an integer could be solved efficiently on a quantum computer [2]. The belief that this problem has no efficient solution on a classical computer is the basis of present-day cryptographic methods. Since then many different concepts for quantum computations have been brought forward and some have also been realized experimentally. Shor's algorithm could be demonstrated on a 7-qubit nuclear magnetic resonance (NMR) quantum computer by factoring the number 15 into its prime factors 3 and 5 [3].

It has turned out that the basic problem is as follows: Approaches such as NMR quantum computers or trapped ion quantum computers that are well suited in terms of long decoherence times are very hard to scale up. In contrast, approaches such as solid state based quantum computers that can be easily scaled up suffer from extremely short decoherence times. As a consequence, first successful realizations of a single- and a two-qubit system based on the electron spin in a semiconductor nanostructure could be demonstrated only recently [4, 5].

The difficulties in the experimental realization of the many existing proposals for quantum information devices based on semiconductor nanostructures show that calculational methods and tools that allow for a predictive analysis in the context of realistic, three-dimensional models are essential. Simple models may help in understanding the basic physical relations, but they are not sufficient to support the experimentalists in a concrete experimental realization. The development of an appropriate simulation tool (next**nano**++) and its application to the physics of semiconductor based devices for quantum information processes is the topic of this work.

In this chapter, we address the theoretical background for realistic, three-dimensional calculations of semiconductor nanostructures for quantum information processing. Here, the main focus is obviously the quantum mechanical determination of the electronic structure. For this, we employ the 1-band, 6-band, or 8-band $k \cdot p$ method in envelope function approximation, including the Hartree interaction as well as exchange and correlation by means of the local spin density approximation. Furthermore, we allow for arbitrarily shaped three-dimensional device geometries and for any combination of group IV elementary and III-V compound semiconduc-

tor materials and alloys. Our theory includes band offsets, deformation potentials, total elastic strain energy, the long range Coulomb potential induced by charged impurities, piezo- and pyroelectric charges, as well as surface charges in a fully self-consistent manner. The current evaluation is restricted to situations close to equilibrium where either the concept of local quasi Fermi levels or ballistic quantum transport is applicable.

1.2 Theory

1.2.1 Band structure

One of the most immediate consequences of the periodic structure of crystalline solids is the arrangement of the electronic states within bands. This so called band structure has a particular importance for semiconductors as many properties of semiconductors are determined by only a small number of these bands. In this section, we will therefore illustrate the origin of the band structure in bulk semiconductors and present the $k \cdot p$ method for band structure calculations. We will then discuss the most important $k \cdot p$ models for the valence and conduction bands in group IV elementary as well as III-V compound semiconductors.

In semiconductor hetero- and nanostructures the discrete translational symmetry of the crystal is broken. However, the envelope function approximation (EA) allows the calculation of the electronic structure of such systems on the basis of the bulk band structures. We will illustrate the EA in the context of the $k \cdot p$ method and we will briefly address the issue of discontinuities in the band energies of different semiconductor materials.

Bloch theory

A crystal is characterized by its regular, periodically repeated structure. The smallest unit of this structure is called the primitive cell. A consequence of this periodicity is that the lattice is invariant to translations by distances that are integer multiples of the lattice period. The lattice symmetry operations referred to as primitive translations can be written as

$$\mathbf{R}_n = n_1\mathbf{a}_1 + n_2\mathbf{a}_2 + n_3\mathbf{a}_3, \tag{1.1}$$

where the n_i are integers. The set of all $\mathbf{R}_n$ form the point lattice of the crystal. The number of possible types of point lattices is limited; in three dimensions there are 14 different point lattices. The crystal's real lattice is obtained from its point lattice by substituting each point of the point lattice by a basis, i.e. the arrangement of the atoms in the Wigner-Seitz cell. Real lattices whose basis does not restrict the symmetry of the Wigner-Seitz cell are called Bravais lattices. The three types of crystals that are most important for semiconductors are shown in Figure 1.1. The zinc-blende crystal has a face-centered cubic point lattice and a two atomic basis. In the case of group IV semiconductors such as silicon, germanium, and carbon, the two atoms are identical (strictly speaking, these semiconductors possess a diamond crystal), whereas in the case of III-V semiconductors such as GaAs, InAs, InP, etc. the two atoms are different. The wurtzite crystal has a hexagonal point lattice and a two atomic basis. Examples of semiconductors with wurtzite crystal structures can be found among the nitrides such as GaN, InN, AlN, etc.

The discrete translational invariance of crystalline semiconductors has an important consequence on their electronic structure as it allows associating the electronic states with discrete

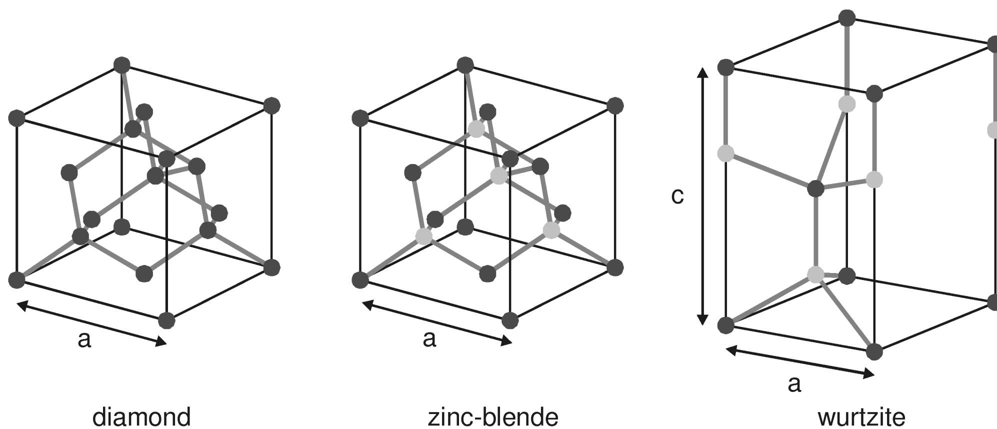

Figure 1.1: The three most common crystal structures for semiconductors: diamond (e.g. C, Si, Ge), zinc-blende (e.g. GaAs, InAs, AlAs), and wurtzite (e.g. GaN, AlN).

bands, thereby forming the so called band structure of the semiconductor. This fact is intimately linked to the Bloch theorem that we will therefore briefly summarize.

The electronic many-body Hamiltonian of a crystalline solid is given by

$$H = \sum_i \left[\frac{\mathbf{p}_i^2}{2m} + V_{\text{ion}}(\mathbf{x}_i) \right] + \frac{1}{2} \sum_{i \neq j} \frac{e^2}{4\pi |\mathbf{x}_i - \mathbf{x}_j|} \tag{1.2}$$

where $\frac{\mathbf{p}_i^2}{2m}$ is the kinetic energy of the ith electron, $V_{\text{ion}}(\mathbf{x}_i)$ is the interaction of the ith electron with the atomic cores, and the last term is the Coulomb interaction between the electrons. Here we have already made the assumption that the atomic cores can be considered as fixed at their equilibrium lattice sites. The Hamiltonian H therefore does not contain the interaction of the electrons with the quantized lattice vibrations, the phonons. For the potential of the atomic cores, translational invariance

$$V_{\text{ion}}(\mathbf{x} + \mathbf{R}) = V_{\text{ion}}(\mathbf{x}) \tag{1.3}$$

holds for all lattice vectors $\mathbf{R}$. Despite neglecting the phonons, the many-body Hamiltonian H is still too hard to be solved exactly. According to density functional theory it is however possible, at least for the many-body ground state, to replace the Hamiltonian H by an effective one-body Hamiltonian

$$H_1 = \frac{\mathbf{p}^2}{2m} + V_{\text{ion}}(\mathbf{x}) + V[n](\mathbf{x}) \tag{1.4}$$

where the electron-electron interaction is replaced by an additional one-body potential $V[n](\mathbf{x})$ that is a functional of the ground state electron density of H_1

$$n(\mathbf{x}) = \sum_{E_n < E_{\text{F}}} |\psi_n(\mathbf{x})|^2, \quad H_1 \psi_n = E_n \psi_n. \tag{1.5}$$

Here, ψ_n denotes the eigenfunctions of the one-body Hamiltonian H_1 and E_n denotes the corresponding eigenenergies. For the ground state electron density, all eigenfunction with energies smaller than the Fermi energy E_{F} are fully occupied. We take into account $V[n](\mathbf{x})$ in the context of the local spin density approximation (LSDA) as is briefly summarized in Reference [6].

It is possible to show that the additional one-body potential $V[n](\mathbf{x})$ obeys the translational symmetry of the lattice so that for all lattice vectors $\mathbf{R}$ a relation analogous to Eq. (1.3) holds

$$V[n](\mathbf{x}+\mathbf{R}) = V[n](\mathbf{x}). \tag{1.6}$$

For this reason, the eigenfunctions $\psi_n(\mathbf{x})$ of the one-body Schrödinger equation

$$H_1\psi_n(\mathbf{x}) = E_n\psi_n(\mathbf{x}) \tag{1.7}$$

with eigenvalues E_n obey Bloch's theorem: they are at the same time eigenfunctions $\psi_{n,\mathbf{k}}(\mathbf{x})$ of the translational operator $T_{\mathbf{R}}$ with eigenvalues $\exp(i\mathbf{k}\cdot\mathbf{R})$

$$T_{\mathbf{R}}\psi_{n,\mathbf{k}}(\mathbf{x}) = \exp(i\mathbf{k}\cdot\mathbf{R})\,\psi_{n,\mathbf{k}}(\mathbf{x}) \tag{1.8}$$

for all lattice vectors $\mathbf{R}$. As a consequence of Bloch's theorem, the eigenfunctions $\psi_{n,\mathbf{k}}(\mathbf{x})$ of H_1 can be written as the product of a plane wave $\exp(i\mathbf{k}\cdot\mathbf{x})$ and a Bloch factor $u_{n,\mathbf{k}}(\mathbf{x})$

$$\psi_{n,\mathbf{k}}(\mathbf{x}) = \exp(i\mathbf{k}\cdot\mathbf{x})\,u_{n,\mathbf{k}}(\mathbf{x}). \tag{1.9}$$

The $\psi_{n,\mathbf{k}}(\mathbf{x})$ are called Bloch functions. By inserting the Bloch functions into the one-body Schrödinger equation, the following relation is obtained for the Bloch factors $u_{n,\mathbf{k}}(\mathbf{x})$

$$\begin{aligned} E_n(\mathbf{k})\,u_{n,\mathbf{k}}(\mathbf{x}) &= H(\mathbf{k})\,u_{n,\mathbf{k}}(\mathbf{x}) \\ &= \left[\frac{(\mathbf{p}+\hbar\mathbf{k})^2}{2m} + V_{\text{ion}}(\mathbf{x}) + V[n](\mathbf{x})\right] u_{n,\mathbf{k}}(\mathbf{x}). \end{aligned} \tag{1.10}$$

For each integer band index n, the eigenenergies $E_n(\mathbf{k})$ depend continuously on the vector $\mathbf{k}$ in reciprocal space. This systematics of the electronic spectrum is called the band structure and a characteristic of all crystalline solids. Special to semiconductors is the fact that in the ground state each band is either fully occupied or fully empty. The highest, fully occupied band is separated from the lowest, unoccupied band by an energy gap of the order of a few eV.

A large number of methods have been developed for the calculation of the band structure of which we want to present in the following the $k\cdot p$ approximation, as this is the basis also for the quantum mechanical treatment of the electronic structure of nanostructures within next**nano**++.

The $k\cdot p$ approximation

The $k\cdot p$ approximation exploits the fact that many properties of semiconductors depend only on the position and shape of the minima and maxima $\mathbf{k}_0$ of the conduction and valence bands and a precise knowledge of the entire band structure is not needed. We therefore expand $E_n(\mathbf{k})$ around these extrema $\mathbf{k}_0$ up to second order in $|\mathbf{k}-\mathbf{k}_0|$ and rewrite the Hamiltonian $H(\mathbf{k})$ as follows:

$$H(\mathbf{k}) = H(\mathbf{k}_0) + \frac{\hbar}{m}(\mathbf{k}-\mathbf{k}_0)\cdot\mathbf{p} + \frac{\hbar^2}{2m}(\mathbf{k}-\mathbf{k}_0)^2. \tag{1.11}$$

The Bloch factors $u_{n,\mathbf{k}_0}(\mathbf{x})$ for $\mathbf{k}=\mathbf{k}_0$ obey

$$\begin{aligned} H(\mathbf{k}_0)\,u_{n,\mathbf{k}_0}(\mathbf{x}) &= \left[\frac{(\mathbf{p}+\hbar\mathbf{k}_0)^2}{2m} + V_{\text{ion}}(\mathbf{x}) + V[n](\mathbf{x})\right] u_{n,\mathbf{k}_0}(\mathbf{x}) \\ &= E_n(\mathbf{k}_0)\,u_{n,\mathbf{k}_0}(\mathbf{x}) \end{aligned} \tag{1.12}$$

and form a complete set of orthonormal functions. This allows us to expand the Bloch factors $u_{n,\mathbf{k}}(\mathbf{x})$ for arbitrary $\mathbf{k}$ in terms of the $u_{n,\mathbf{k}_0}(\mathbf{x})$

$$u_{n,\mathbf{k}}(\mathbf{x}) = \sum_{\nu} c_{n,\nu}(\mathbf{k})\, u_{\nu,\mathbf{k}_0}(\mathbf{x}) . \tag{1.13}$$

By inserting this expansion into Eq. (1.10) we obtain an eigenvalue equation for the expansion coefficients $c_{n,\nu}(\mathbf{k})$

$$E_n(\mathbf{k})\, c_{n,\nu}(\mathbf{k}) = \sum_{\mu} H_{\nu\mu}(\mathbf{k})\, c_{n,\mu}(\mathbf{k}) \tag{1.14}$$

and the Hamiltonian is represented by the following infinite dimensional matrix

$$H_{\nu\mu}(\mathbf{k}) = \left[E_\nu(\mathbf{k}_0) + \frac{\hbar^2}{2m}(\mathbf{k}-\mathbf{k}_0)^2\right]\delta_{\nu\mu} + \frac{\hbar}{m}(\mathbf{k}-\mathbf{k}_0)\cdot\mathbf{p}_{\nu\mu} \tag{1.15}$$

that couples all energy bands by means of the momentum matrix elements $\mathbf{p}_{\nu\mu}$ between the respective Bloch factors for $\mathbf{k} = \mathbf{k}_0$

$$\mathbf{p}_{\nu\mu} = \langle u_{\nu,\mathbf{k}_0} |\, \mathbf{p}\, | u_{\mu,\mathbf{k}_0} \rangle . \tag{1.16}$$

For many properties of semiconductors only a small number of energy bands are important, namely the valence and conduction bands in the immediate proximity of the Fermi energy. Therefore, the next step is to use perturbation theory to obtain a finite dimensional Hamiltonian matrix for those bands [7]. For this, we separate the bands into two groups A and B where group A contains the few bands near the Fermi energy E_F and group B contains all remaining bands. We assume that the non-diagonal coupling term

$$\frac{\hbar}{m}(\mathbf{k}-\mathbf{k}_0)\cdot\mathbf{p}_{\nu\mu} \tag{1.17}$$

is weak between bands from group A and group B and can be treated as a perturbation [8–11]. The infinite dimensional Hamiltonian matrix H then becomes the finite dimensional $k \cdot p$ Hamiltonian matrix

$$\begin{aligned} H^{\mathrm{kp}}_{\nu\mu}(\mathbf{k}) \;=\; & \left[E_\nu(\mathbf{k}_0) + \frac{\hbar^2}{2m}(\mathbf{k}-\mathbf{k}_0)^2\right]\delta_{\nu\mu} + \frac{\hbar}{m}(\mathbf{k}-\mathbf{k}_0)\cdot\mathbf{p}_{\nu\mu} \\ & + \sum_{\gamma\in B} \frac{H_{\nu\gamma}(\mathbf{k})\, H_{\gamma\mu}(\mathbf{k})}{E_\nu(\mathbf{k}_0) - E_\gamma(\mathbf{k}_0)}, \end{aligned} \tag{1.18}$$

where the indices ν and μ run only over the few bands in group A. The momentum matrix elements $\mathbf{p}_{\nu\mu}$ are customarily used as fitting parameters to match experimentally obtained results. Furthermore, many of the momentum matrix elements vanish due to symmetry.

Depending on the set of bands in group A one obtains different $k\cdot p$ models for the conduction and the valence bands. This also includes the effective mass approximation where only a single band is taken into account. Common multi-band $k \cdot p$ models are the 6-band model that includes only valence bands, and the 8-band and 14-band models that include valence as well as conduction bands. In the following, we will present the details of the effective mass approximation, the 6-band, and the 8-band model. For the 6-band and 8-band model we restrict ourselves for the moment to the zinc-blende crystal structure and refer the reader to Appendix C as concerning the case of the wurtzite crystal structure.

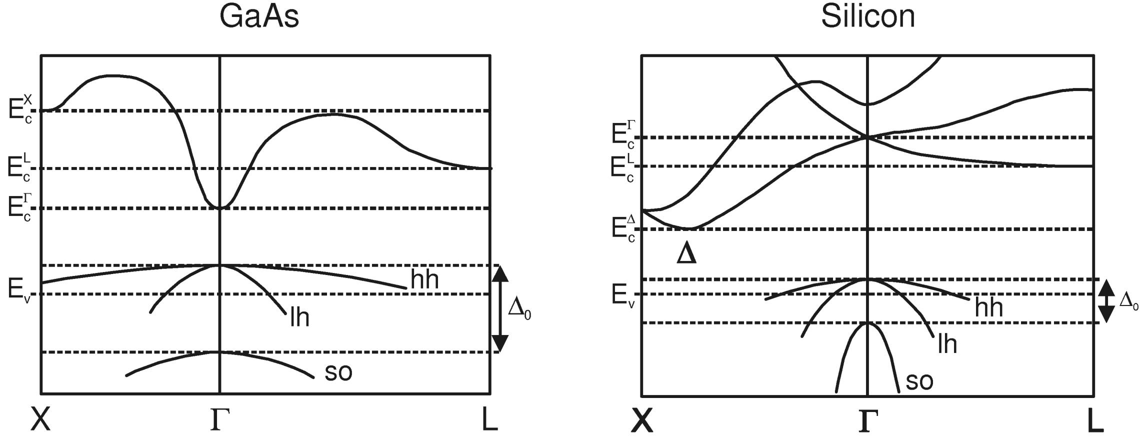

Figure 1.2: The band structures of GaAs and silicon. For GaAs, the conduction band has minima at the Γ-point, the X-point and the L-point. For silicon, there are no minima at the X-point, but along the connecting lines Δ between Γ and X. In both materials, the heavy hole (hh) and the light hole (lh) valence bands are degenerate. The degeneracy is lifted for the split-off band (so) due to the spin orbit interaction.

Conduction and valence band models

The simplest model that can be derived from the $k \cdot p$ approximation is the so called effective mass approximation (EMA). For the EMA only one band ($\nu, \mu = n$) is taken into account in group A. Thus Eq. (1.18) for the $k \cdot p$ Hamiltonian matrix is reduced to a simple scalar function of $\mathbf{k}$

$$\begin{aligned} H^{\mathrm{EMA}}(\mathbf{k}) &= E_n(\mathbf{k}_0) + \frac{\hbar^2}{2m}(\mathbf{k}-\mathbf{k}_0)^2 + \frac{\hbar^2}{m^2}\sum_{\gamma\neq n}\frac{|\mathbf{p}_{n\gamma}|^2}{E_n(\mathbf{k}_0)-E_\gamma(\mathbf{k}_0)} \\ &= E_n(\mathbf{k}_0) + \frac{\hbar^2}{2}(\mathbf{k}-\mathbf{k}_0)\cdot\frac{1}{\underline{m}^*}(\mathbf{k}-\mathbf{k}_0). \end{aligned} \tag{1.19}$$

The symmetric 3×3 matrix $\underline{m}^*$ is called the effective mass tensor. The EMA often suffices for the description of the conduction band. The conduction band has multiple local minima at special points in the Brillouin zone. For zinc-blende, these are the Γ-point ($\mathbf{k}_0 = 0$), the X-point ($\mathbf{k}_0 = \{100\}$), and the L-point ($\mathbf{k}_0 = \{111\}/\sqrt{3}$). The minima at the X-point are 3-fold, the ones at the L-point are 4-fold degenerate, as each is shared by two neighboring Brillouin zones. In silicon, there are no minima at the X-point but along Δ, which is the connecting line between Γ and X. These minima are therefore 6-fold degenerate. At the Γ-point of the bulk Brillouin zone $\underline{m}^*$ is isotropic

$$\underline{m}^* = m^*\mathbf{1}_{3\times3} \tag{1.20}$$

due to symmetry reasons. The effective mass tensors at the X-point, the L-point and along Δ are characterized by a longitudinal effective mass m_l^* and a transversal effective mass m_t^*

$$\underline{m}^* = (m_\mathrm{l}^* - m_\mathrm{t}^*)\,\hat{\mathbf{k}}_0\hat{\mathbf{k}}_0^T + m_\mathrm{t}^*\mathbf{1}_{3\times3} \tag{1.21}$$

and describe ellipsoids of revolution with the axis of symmetry pointing in the direction of $\mathbf{k}_0$.

The valence bands are created by the p orbitals in most semiconductors and are therefore threefold degenerate. For this reason, the $k \cdot p$ model for the valence bands needs to include at least three bands in the orbital basis

$$|x_1\rangle, |x_2\rangle, |x_3\rangle, \tag{1.22}$$

which for zinc-blende materials at the Γ-point results in

$$\begin{aligned} H^{\mathrm{zb},3\times 3}(\mathbf{k}) &= E_{\mathrm{v}} + \frac{\hbar^2}{2m}\mathbf{k}^2 + N\mathbf{k}\mathbf{k}^T + (L-N)\,\mathrm{diag}\left(k_1^2, k_2^2, k_3^2\right) \\ &\quad + M\,\mathrm{diag}\left(k_2^2 + k_3^2, k_1^2 + k_3^2, k_1^2 + k_2^2\right). \end{aligned} \tag{1.23}$$

The parameters L, M, N are called the Dresselhaus parameters [9] and depend on the momentum matrix elements $\mathbf{p}_{\nu\mu}$ as the effective mass tensor $\underline{m}^*$ does. So far, we have neglected spin effects. For an accurate description of the valence bands however, it is necessary to include the spin-orbit coupling. By taking into account the spin-orbit coupling we have to add the term

$$H_{\mathrm{SO}} = \frac{\hbar}{4m^2c^2}\left(\boldsymbol{\nabla} V \times \mathbf{p}\right) \cdot \boldsymbol{\sigma} \tag{1.24}$$

to the one-body Hamiltonian given by Equation (1.4). Here $\boldsymbol{\sigma}$ denotes the Pauli spin matrix vector $\boldsymbol{\sigma} = (\sigma_1, \sigma_2, \sigma_3)^T$. The solutions of Schrödinger's equation are then two-component spinors and an appropriate basis for the three valence bands including spin is given by

$$|x_1\rangle|\uparrow\rangle, |x_2\rangle|\uparrow\rangle, |x_3\rangle|\uparrow\rangle, |x_1\rangle|\downarrow\rangle, |x_2\rangle|\downarrow\rangle, |x_3\rangle|\downarrow\rangle. \tag{1.25}$$

The $k \cdot p$ Hamiltonian then becomes a 6×6 matrix that can be expressed as follows:

$$H^{\mathrm{zb},6\times 6}(\mathbf{k}) = H^{\mathrm{zb},3\times 3}(\mathbf{k}) \otimes \mathbf{1}_{2\times 2} + H_{\mathrm{SO}}^{\mathrm{zb},6\times 6}(\mathbf{k}) \tag{1.26}$$

where $H_{\mathrm{SO}}^{\mathrm{zb},6\times 6}(\mathbf{k})$ is the 6×6 matrix resulting from H_{SO} by employing the $k \cdot p$ approximation. This gives rise to the following two contributions:

$$H_{\mathrm{SO}}^1 = \frac{\hbar}{4m^2c^2}\left(\boldsymbol{\nabla} V \times \mathbf{p}\right) \cdot \boldsymbol{\sigma}, \tag{1.27}$$

$$H_{\mathrm{SO}}^2 = \frac{\hbar}{4m^2c^2}\left(\boldsymbol{\nabla} V \times \mathbf{k}\right) \cdot \boldsymbol{\sigma}. \tag{1.28}$$

Usually one makes the assumption that

$$\left\|H_{\mathrm{SO}}^2\right\| \ll \left\|H_{\mathrm{SO}}^1\right\|, \tag{1.29}$$

which renders $H_{\mathrm{SO}}^{\mathrm{zb},6\times 6}(\mathbf{k})$ to be independent of $\mathbf{k}$. The Hamiltonian matrix $H_{\mathrm{SO}}^{\mathrm{zb},6\times 6}$ is then given by

$$H_{\mathrm{SO}}^{\mathrm{zb},6\times 6} = \frac{\Delta_0}{3}\begin{pmatrix} 0 & -i & 0 & 0 & 0 & 1 \\ i & 0 & 0 & 0 & 0 & -i \\ 0 & 0 & 0 & -1 & i & 0 \\ 0 & 0 & -1 & 0 & i & 0 \\ 0 & 0 & -i & -i & 0 & 0 \\ 1 & i & 0 & 0 & 0 & 0 \end{pmatrix}, \tag{1.30}$$

where Δ_0 is the spin-orbit splitting. The basis that diagonalizes $H_{\mathrm{SO}}^{\mathrm{zb},6\times 6}$ is used to classify the valence bands into heavy hole, light hole, and split-off hole.

An even more accurate description is obtained when also the conduction band at the Γ-point is taken into account. As basis functions for the conduction band we choose

$$|s\rangle\,|\uparrow\rangle\,,|s\rangle\,|\downarrow\rangle\,. \tag{1.31}$$

The resulting 8-band $k \cdot p$ model is then described by the following 8×8 Hamiltonian matrix

$$H^{\mathrm{zb},8\times 8}(\mathbf{k}) = \begin{pmatrix} H_{\mathrm{c}}^{\mathrm{zb}}(\mathbf{k}) & H_{\mathrm{cv}}^{\mathrm{zb}}(\mathbf{k}) \\ H_{\mathrm{vc}}^{\mathrm{zb}}(\mathbf{k}) & H^{\mathrm{zb},6\times 6}(\mathbf{k}) \end{pmatrix} \tag{1.32}$$

with the 2×2 conduction band submatrix

$$H_{\mathrm{c}}^{\mathrm{zb}}(\mathbf{k}) = \left(E_{\mathrm{c}} + \frac{\hbar^2 S}{2m}\mathbf{k}^2\right) \otimes \mathbf{1}_{2\times 2}, \tag{1.33}$$

and the coupling matrix between conduction and valence bands

$$H_{\mathrm{vc}}^{\mathrm{zb}}(\mathbf{k}) = \begin{pmatrix} Bk_2k_3 + iPk_1 \\ Bk_1k_3 + iPk_2 \\ Bk_1k_2 + iPk_3 \end{pmatrix} \otimes \mathbf{1}_{2\times 2}, \qquad H_{\mathrm{cv}}^{\mathrm{zb}}(\mathbf{k}) = H_{\mathrm{cv}}^{\mathrm{zb}}(\mathbf{k})^H\,. \tag{1.34}$$

The parameter B is non-zero only for crystals without inversion symmetry as is the case for the III-V compounds. For group IV semiconductors such as silicon, $B = 0$ holds. The parameter P is related to Kane's momentum matrix element E_P according to $E_P = 2m_0\left(\frac{P}{\hbar}\right)^2$. For $B = P = 0$ the parameter S^{-1} would be equivalent to the effective mass of the conduction band.

The Hamiltonian matrices for the corresponding multi-band $k \cdot p$ models for wurtzite are given in Appendix C.

The envelope function approximation

Until now we have only considered perfectly periodic, infinite-size semiconductor crystals. However, realistic nanostructures not only have a very small finite size, but they may also contain variations in material compositions and may be subjected to non-uniform external potentials $V_{\mathrm{ext}}(\mathbf{x})$. As a consequence, the Hamiltonian describing these systems is not translation invariant and neither Bloch's theorem nor Eq. (1.9) is available.

By means of the envelope function approximation it is however possible to adopt most of the presented theory also for nanostructures. We want to explain the envelope function approximation for the example of a homogeneous semiconductor that is subject to an external potential $V_{\mathrm{ext}}(\mathbf{x})$. The electron wave functions $\psi(\mathbf{x})$ are written as superpositions of Bloch functions $\psi_{n,\mathbf{k}_0}(\mathbf{x})$

$$\psi(\mathbf{x}) = \sum_n F_n(\mathbf{x})\,\psi_{n,\mathbf{k}_0}(\mathbf{x}) \tag{1.35}$$

where the expansion coefficients $F_n(\mathbf{x})$ are denoted envelope functions and are assumed to be slowly varying as compared to the lattice constant. The next step is to transform the Hamiltonian $H(\mathbf{k}_0)$ in Eq. (1.10) to reciprocal space. For this, we expand the envelope functions as well as the external potential in terms of Fourier series:

$$F_n(\mathbf{x}) = \sum_{\mathbf{k}} c_{n\mathbf{k}} \exp(i\mathbf{k}\mathbf{x})\,, \tag{1.36}$$

$$V_{\text{ext}}(\mathbf{x}) = \sum_{\mathbf{k}} V_{\mathbf{k}}^{\text{ext}} \exp(i\mathbf{k}\mathbf{x}). \tag{1.37}$$

From the Schrödinger equation for $H(\mathbf{k}_0)$ including the external potential $V_{\text{ext}}(\mathbf{x})$

$$(H(\mathbf{k}_0) + V_{\text{ext}}(\mathbf{x}))\psi(\mathbf{x}) = E\psi(\mathbf{x}) \tag{1.38}$$

we therefore obtain the following system of equations for the envelope function expansion coefficients $c_{n\mathbf{k}}$

$$Ec_{n'\mathbf{k}'} = \sum_n \sum_{\mathbf{k}} c_{n\mathbf{k}} \left(\left[E_n(\mathbf{k}_0) + \frac{\hbar^2 \mathbf{k}^2}{2m} \right] \delta_{n'n}\delta_{\mathbf{k}'\mathbf{k}} + V_{\mathbf{k}'-\mathbf{k}}^{\text{ext}} \delta_{n'n} + \frac{\hbar \mathbf{k} \cdot \mathbf{p}_{n'n}}{m} \delta_{\mathbf{k}'\mathbf{k}} \right). \tag{1.39}$$

In the real space representation, this system of equations yields a Schrödinger like eigenvalue problem for the envelope functions

$$\begin{aligned} EF_{n'}(\mathbf{x}) &= \sum_n \left(\left[E_n(\mathbf{k}_0) - \frac{\hbar^2 \boldsymbol{\nabla}^2}{2m} \right] \delta_{n'n} + \frac{\hbar \boldsymbol{\nabla} \cdot \mathbf{p}_{n'n}}{m} + V_{\text{ext}}(\mathbf{x}) \delta_{n'n} \right) F_n(\mathbf{x}) \\ &= \sum_n H_{n'n}^{\text{EA}}(\mathbf{k}_0) F_n(\mathbf{x}). \end{aligned} \tag{1.40}$$

Thus the envelope function Hamiltonian $H_{\nu\mu}^{\text{EA}}(\mathbf{k})$ in Eq. (1.40) is identical to the infinite dimensional Hamiltonian matrix $H_{\nu\mu}(\mathbf{k})$ in Eq. (1.15) if $\mathbf{k} - \mathbf{k}_0$ is replaced by $-i\boldsymbol{\nabla}$.

For $H_{\nu\mu}^{\text{EA}}(\mathbf{k})$ it is also possible to employ the $k \cdot p$ approximation. The envelope function Hamiltonian for the conduction band in effective mass approximation is given by

$$H^{\text{EA,EMA}} = -\frac{\hbar^2}{2} \boldsymbol{\nabla} \cdot \left(\frac{1}{\underline{m}^*} \right) \boldsymbol{\nabla} + E_{\text{c}} + V_{\text{ext}}(\mathbf{x}) \tag{1.41}$$

where the conduction band edge $E_{\text{c}} = E_n(\mathbf{k}_0)$ is the energy of the conduction band at the considered minimum $\mathbf{k} = \mathbf{k}_0$. This expression may also be used for nanostructures with non-constant material composition [12, 13]. In this case, the effective mass tensor $\underline{m}^*$ as well as the conduction band edge E_{c} depend on the position, i.e. $\underline{m}^* = \underline{m}^*(\mathbf{x})$ and $E_{\text{c}} = E_{\text{c}}(\mathbf{x})$

$$H^{\text{EA,EMA}} = -\frac{\hbar^2}{2} \boldsymbol{\nabla} \cdot \left(\frac{1}{\underline{m}^*(\mathbf{x})} \right) \boldsymbol{\nabla} + E_{\text{c}}(\mathbf{x}) + V_{\text{ext}}(\mathbf{x}). \tag{1.42}$$

The same approach is also valid for the presented multi-band $k \cdot p$ Hamiltonians. In this case, all the $k \cdot p$ parameters become position dependent quantities. At first, this leads to several problems due to operator ordering issues. However, it has been shown that each of these problems can be solved [14].

The model solid

At the interface of two semiconductors, the discontinuities in band edge energies are not well defined *a priori*. This is because an absolute energy scale, on which to determine the energies of the band edges of the bulk semiconductors, is missing. Microscopically, the band edge discontinuities are caused by electric dipole layers [15]. Dipole layers correspond to a discontinuity in the electrostatic potential that in general is sought to be fully included in the band edge energies of the bulk semiconductors. Van de Walle has proposed a theory that relates all

band edge energies to the energy scale provided by a model solid consisting of neutral atoms [16, 17]. This approach also yields the basis of the separation of the strain induced band shifts between the valence and conduction bands. A different approach for setting up the reference energy scale in terms of electronic states of the atomic cores is pursued in Reference [18]. For our database, the parameters for the bulk band edge energies are mainly taken from this more recent work.

1.2.2 Poisson equation

The Poisson equation for a semiconductor nanostructure with spatially varying dielectric constant is given by

$$\nabla \cdot \left[\varepsilon \left(\mathbf{x} \right) \nabla \phi \left(\mathbf{x} \right) \right] = -\rho \left(\mathbf{x} \right) \tag{1.43}$$

where $\phi(\mathbf{x})$ denotes the electrostatic potential, $\rho(\mathbf{x})$ is the charge density, and $\varepsilon(\mathbf{x})$ in general is the position dependent dielectric tensor. The charge density is composed of holes $p(\mathbf{x})$, electrons $n(\mathbf{x})$, ionized donors $N_{\mathrm{D}}^{+}(\mathbf{x})$, ionized acceptors $N_{\mathrm{A}}^{-}(\mathbf{x})$, and fixed charges $\rho_{\mathrm{fix}}(\mathbf{x})$ as follows:

$$\rho \left(\mathbf{x} \right) = e \left[p \left(\mathbf{x} \right) - n \left(\mathbf{x} \right) + N_{\mathrm{D}}^{+} \left(\mathbf{x} \right) - N_{\mathrm{A}}^{-} \left(\mathbf{x} \right) + \rho_{\mathrm{fix}} \left(\mathbf{x} \right) \right]. \tag{1.44}$$

Contributions to the fixed charges $\rho_{\mathrm{fix}}(\mathbf{x})$ are polarization charges $\rho_{\mathrm{pol}}(\mathbf{x})$ due to piezo- and pyroelectric polarizations as well as fixed surface and volume charges $\rho_{\mathrm{surf/vol}}(\mathbf{x})$

$$\rho_{\mathrm{fix}} \left(\mathbf{x} \right) = \rho_{\mathrm{pol}} \left(\mathbf{x} \right) + \rho_{\mathrm{surf/vol}} \left(\mathbf{x} \right). \tag{1.45}$$

The determination of the polarization charges is discussed in Section 1.2.4.

The unique solution of the Poisson equation requires specifying boundary conditions. Boundary conditions have to be specified at the boundaries of the nanostructure. Furthermore, boundary conditions are used at the interfaces to contacts. Contacts are part of the nanostructure, but inside of contacts the Poisson equation is not solved. Instead, assumptions on the electrostatic potential within the contacts are made that are then incorporated in terms of appropriate boundary conditions.

From a mathematical point of view, the two most important types of boundary conditions are the Dirichlet and the von Neumann boundary conditions. For a given region Ω with boundary $\partial\Omega$, these are characterized by

$$\phi \left(\mathbf{x} \right) \big|_{\mathbf{x} \in \partial \mathbf{\Omega}} = f \left(\mathbf{x} \right) \qquad \text{(Dirichlet boundary condition)} \tag{1.46}$$

and

$$\partial_{\mathbf{n}} \phi \left(\mathbf{x} \right) \big|_{\mathbf{x} \in \partial \mathbf{\Omega}} = g \left(\mathbf{x} \right) \qquad \text{(von Neumann boundary condition),} \tag{1.47}$$

where $\mathbf{n}$ denotes the normal to $\partial\Omega$ and $f(\mathbf{x})$ and $g(\mathbf{x})$ are given functions on the boundary of Ω. Besides those, periodic boundary conditions are of interest for simulations of periodically repeated structures. The application of the boundary conditions for the electrostatic potential will be discussed in the next section in combination with the boundary conditions for the quasi Fermi levels. In the remaining part of this section, we will go into the details on how to determine the hole and electron charge densities as well as that of the ionized impurities.

For the determination of the charge carrier densities we use two different approaches. In regions, where quantum mechanical effects are dominant, the charge carrier densities are calculated on the basis of the envelope wave functions obtained from the single-band effective mass

or multi-band $k \cdot p$ equations. If quantum mechanical effects do not play a role, the charge carrier densities can be calculated classically by means of the Thomas-Fermi approximation.

To calculate the charge carrier densities from the envelope wave functions, the Fermi-Dirac statistics of either holes ($l = \mathrm{v}$) or electrons ($l = \mathrm{c}$) are used to occupy the states according to their corresponding probability densities

$$p^i_{l,j}(\mathbf{x}) = \sum_\nu \left| (F_\nu)^i_{l,j}(\mathbf{x}) \right|^2 \qquad l = \mathrm{v}, \mathrm{c}. \tag{1.48}$$

For $l = \mathrm{v}$ ($l = \mathrm{c}$) the index i runs over all included valence bands (conduction bands). The index j enumerates the states in band i and the summation runs over all components of the envelope function $F^i_{l,j}(\mathbf{x})$. For quantization in all three spatial dimensions, this immediately leads to the following expressions for the hole density $p(\mathbf{x})$ and the electron density $n(\mathbf{x})$

$$p(\mathbf{x}) = \sum_{i \in \mathrm{VB}} g^i_{\mathrm{v}} \sum_j p^i_{\mathrm{v},j}(\mathbf{x}) \, f\left(\left[E^i_{\mathrm{v},j} - E_{\mathrm{F,p}}(\mathbf{x})\right]/kT\right), \tag{1.49}$$

$$n(\mathbf{x}) = \sum_{i \in \mathrm{CB}} g^i_{\mathrm{c}} \sum_j p^i_{\mathrm{c},j}(\mathbf{x}) \, f\left(\left[-E^i_{\mathrm{c},j} + E_{\mathrm{F,n}}(\mathbf{x})\right]/kT\right), \tag{1.50}$$

where g^i_{v} and g^i_{c} account for possible spin and valley degeneracies. $f(E)$ is the Fermi distribution function. Note that strictly speaking these two expressions are correct only in equilibrium for $E_{\mathrm{F,n}}(\mathbf{x}) = E_{\mathrm{F,p}}(\mathbf{x}) = E_{\mathrm{F}}$, where E_{F} denotes the Fermi energy. The use of separate, position dependent quasi Fermi levels $E_{\mathrm{F,p}}(\mathbf{x})$ and $E_{\mathrm{F,n}}(\mathbf{x})$ for holes and electrons to describe also non-equilibrium situations is justified only if $E_{\mathrm{F,p}}(\mathbf{x})$ and $E_{\mathrm{F,n}}(\mathbf{x})$ vary slowly as compared to the envelope wave functions [19].

For structures with translational invariance in one or two of the three spatial dimensions, it is possible to separate the free motion of the charge carriers in the directions of translational invariance. The free motion of the charge carriers in these directions is characterized by the $\mathbf{k}_\parallel$ lattice wave vector. The correct density of the charge carriers is then obtained by integrating over $\mathbf{k}_\parallel$. Further details on this integration are presented in Chapter 4. Here, we want to restrict ourselves to the case of the single-band effective mass approximation with position independent effective mass tensor. In this case, the integration of the $\mathbf{k}_\parallel$ lattice wave vector can be done analytically resulting in the following expressions for the charge densities of holes and electrons

$$p(\mathbf{x}) = \sum_{i \in \mathrm{VB}} g^i_{\mathrm{v}} \left(\frac{m^i_{\mathrm{qdos,h}} kT}{2\pi\hbar^2} \right)^{(3-d)/2} \sum_j p^i_{\mathrm{v},j}(\mathbf{x}) \, \mathcal{F}_{(1-d)/2}\left(\left[E_{\mathrm{F,p}}(\mathbf{x}) - E^i_{\mathrm{v},j}\right]/kT\right), \tag{1.51}$$

$$n(\mathbf{x}) = \sum_{i \in \mathrm{CB}} g^i_{\mathrm{c}} \left(\frac{m^i_{\mathrm{qdos,e}} kT}{2\pi\hbar^2} \right)^{(3-d)/2} \sum_j p^i_{\mathrm{c},j}(\mathbf{x}) \, \mathcal{F}_{(1-d)/2}\left(\left[-E_{\mathrm{F,n}}(\mathbf{x}) + E^i_{\mathrm{c},j}\right]/kT\right), \tag{1.52}$$

where the number of quantized spatial dimensions is given by d. The function $\mathcal{F}_n(E)$ denotes the Fermi-Dirac integral of order n. Integrating out the free motion of the carriers in the direction of $\mathbf{k}_\parallel$ results in the equivalent densities of states of a two-dimensional (for $d = 1$) or one-dimensional (for $d = 2$) free electron gas that depend on the density-of-states mass

$$m^i_{\mathrm{qdos},\lambda} = \left(\det \bar{m}^{*i}_\lambda\right)^{1/d} \qquad \lambda = \mathrm{h}, \mathrm{e}. \tag{1.53}$$

Here, $\bar{m}_{\lambda}^{*i}(\mathbf{x})$ denotes the 2×2 (1×1) submatrix of the effective mass tensor $\underline{m}_{\lambda}^{*i}(\mathbf{x})$ in the direction of $\mathbf{k}_{\|}$.

In the context of the Thomas-Fermi approximation the following expressions for the charge carrier densities can be derived: For holes one finds

$$p(\mathbf{x}) = \sum_{i\in \mathrm{VB}} N_{\mathrm{v}}^{i}(T)\,\mathcal{F}_{1/2}\left(\left[E_{\mathrm{v}}^{i}(\mathbf{x}) - e\phi(\mathbf{x}) - E_{\mathrm{F,p}}(\mathbf{x})\right]/kT\right), \tag{1.54}$$

the respective formula for electrons is given by

$$n(\mathbf{x}) = \sum_{i\in \mathrm{CB}} N_{\mathrm{c}}^{i}(T)\,\mathcal{F}_{1/2}\left(\left[-E_{\mathrm{c}}^{i}(\mathbf{x}) + e\phi(\mathbf{x}) + E_{\mathrm{F,n}}(\mathbf{x})\right]/kT\right). \tag{1.55}$$

The summations run over the included valence and conduction bands, respectively. Here, $E_{\mathrm{v}}^{i}(\mathbf{x})$ and $E_{\mathrm{c}}^{i}(\mathbf{x})$ denote the position dependent band edges, $N_{\mathrm{v}}^{i}(T)$ and $N_{\mathrm{c}}^{i}(T)$ are the equivalent densities of states at the valence and conduction band edges. They are given by

$$N_{l}^{i}(T) = g_{l}^{i}\left(\frac{m_{\mathrm{dos},\lambda}^{i}(\mathbf{x})\,kT}{2\pi\hbar^{2}}\right)^{3/2} \qquad l = \mathrm{V,C} \text{ and correspondingly } \lambda = \mathrm{h,e}, \tag{1.56}$$

where $m_{\mathrm{dos,h}}^{i}(\mathbf{x})$ and $m_{\mathrm{dos,e}}^{i}(\mathbf{x})$ are the density-of-states masses for electrons and holes in the various valence and conduction bands. These density-of-states masses can be calculated from the effective mass tensors $\underline{m}_{\mathrm{h}}^{*i}$ and $\underline{m}_{\mathrm{e}}^{*i}$ according to

$$m_{\mathrm{dos},\lambda}^{i}(\mathbf{x}) = \left(\det \underline{m}_{\lambda}^{*i}(\mathbf{x})\right)^{1/3} \qquad \lambda = \mathrm{h,e}. \tag{1.57}$$

Likewise, expressions for the densities of ionized impurities can be derived in the context of the Thomas-Fermi approximation. For the ionized acceptors the resulting formula is given by

$$N_{\mathrm{A}}^{-}(\mathbf{x}) = \sum_{i\in \mathrm{Acceptors}} \frac{N_{\mathrm{A}}^{i}(\mathbf{x})}{1 + g_{\mathrm{A}}^{i}\exp\left(-\left[E_{\mathrm{F,p}}(\mathbf{x}) - E_{\mathrm{A}}^{i}(\mathbf{x})\right]/kT\right)}, \tag{1.58}$$

where

$$E_{\mathrm{A}}^{i}(\mathbf{x}) = E_{\mathrm{v}}(\mathbf{x}) - e\phi(\mathbf{x}) - E_{\mathrm{A,ion}}^{i}. \tag{1.59}$$

The summation in Eq. (1.58) runs over all different acceptor species. Each type of acceptor is characterized by its ionization energy $E_{\mathrm{A,ion}}^{i}$, degeneracy g_{A}^{i}, and in particular its spatial distribution $N_{\mathrm{A}}^{i}(\mathbf{x})$. Similarly, one finds for the ionized donors

$$N_{\mathrm{D}}^{+}(\mathbf{x}) = \sum_{i\in \mathrm{Donors}} \frac{N_{\mathrm{D}}^{i}(\mathbf{x})}{1 + g_{\mathrm{D}}^{i}\exp\left(\left[E_{\mathrm{F,n}}(\mathbf{x}) - E_{\mathrm{D}}^{i}(\mathbf{x})\right]/kT\right)}, \tag{1.60}$$

where

$$E_{\mathrm{D}}^{i}(\mathbf{x}) = E_{\mathrm{c}}(\mathbf{x}) - e\phi(\mathbf{x}) - E_{\mathrm{D,ion}}^{i}. \tag{1.61}$$

In Eq. (1.60), the summation runs over all different donor species. As for the acceptors, each donor is characterized by its ionization energy $E_{\mathrm{D,ion}}^{i}$, degeneracy g_{D}^{i}, and its spatial distribution $N_{\mathrm{D}}^{i}(\mathbf{x})$. The impurity degeneracies are normally $g_{\mathrm{A}} = 4$ for acceptors and $g_{\mathrm{D}} = 2$ for donors [20].

1.2.3 Current equation

If the contacts of a nanostructure are biased by different voltages the system is driven out of equilibrium and electric currents may flow between the contacts. This situation is characterized by the fact that the specification of a single Fermi energy for electrons and holes that is constant throughout the device is not sufficient to describe the charge distribution. Moreover, the charge distribution in general does not follow the Fermi-Dirac statistics. The non-equilibrium Green's function (NEGF) formalism [21] allows for a fully quantum mechanical treatment of non-equilibrium situations including scattering, but is computationally still to demanding for calculations in two or three spatial dimensions.

For next**nano**++ we therefore employ the following two approaches. For situations close to equilibrium and in the absence of scattering we apply a recently developed, very efficient method for the calculation of the ballistic quantum transport [22]. This method is discussed in detail in Chapter 5. The second approach has been introduced by Hackenbuchner in Ref. [19] and combines the classical drift diffusion method with a quantum mechanical approach for the determination of the charge carrier density. This method is presented briefly in the following.

The classical drift diffusion method can be derived from the Boltzmann equation whose zeroth and first moments yield the continuity equation and an expression for the current density. Taking into account holes and electrons, the continuity equations are given by

$$-\nabla \cdot \mathbf{j}_{\mathrm{p}}(\mathbf{x}) = R(\mathbf{x}), \tag{1.62}$$

$$-\nabla \cdot \mathbf{j}_{\mathrm{n}}(\mathbf{x}) = R(\mathbf{x}), \tag{1.63}$$

where $\mathbf{j}_{\mathrm{p}}(\mathbf{x})$ and $\mathbf{j}_{\mathrm{n}}(\mathbf{x})$ are the current densities for holes and electrons, respectively. The two equations are coupled in terms of generation and recombination processes that are both included in $R(\mathbf{x})$. Depending on the sign of $R(\mathbf{x})$ either generation ($R(\mathbf{x}) < 0$) or recombination ($R(\mathbf{x}) > 0$) is dominant. The current densities that result from the first moment of the Boltzmann equation are given by

$$\mathbf{j}_{\mathrm{p}}(\mathbf{x}) = \mu_{\mathrm{p}}(\mathbf{x})\, p(\mathbf{x})\, \nabla E_{\mathrm{F,p}}(\mathbf{x}), \tag{1.64}$$

$$\mathbf{j}_{\mathrm{n}}(\mathbf{x}) = -\mu_{\mathrm{n}}(\mathbf{x})\, n(\mathbf{x})\, \nabla E_{\mathrm{F,n}}(\mathbf{x}), \tag{1.65}$$

where $\mu_{\mathrm{p}}(\mathbf{x})$, $\mu_{\mathrm{n}}(\mathbf{x})$ are the mobilities and $E_{\mathrm{F,p}}(\mathbf{x})$, $E_{\mathrm{F,n}}(\mathbf{x})$ are the quasi Fermi levels for holes and electrons, respectively. The additional assumption is that the charge carriers are locally close to thermodynamic equilibrium and can be described by the Fermi-Dirac statistics. Models for the recombination $R(\mathbf{x})$ and mobility $\mu_{\mathrm{p}}(\mathbf{x})$, $\mu_{\mathrm{n}}(\mathbf{x})$ are given in Appendix D. In the classical drift diffusion approach, the charge carrier densities $p(\mathbf{x})$ and $n(\mathbf{x})$ are calculated according to Eq. (1.54) and Equation (1.55). In the quantum drift diffusion method by Hackenbuchner, Eq. (1.49) and Eq. (1.50) are used. Together with the Poisson equation and the Schrödinger equation this method allows one to determine a charge self-consistent solution for the electrostatic potential and the quasi Fermi levels. The current densities can then be calculated from the quasi Fermi levels and the charge carrier densities are calculated according to Eq. (1.64) and Equation (1.65).

To determine the quasi Fermi levels the following two linear partial differential equations have to be solved

$$-\nabla \cdot \mu_{\mathrm{p}}(\mathbf{x})\, p(\mathbf{x})\, \nabla E_{\mathrm{F,p}}(\mathbf{x}) = R(\mathbf{x}), \tag{1.66}$$

$$\nabla \cdot \mu_{\mathrm{n}}(\mathbf{x})\, n(\mathbf{x})\, \nabla E_{\mathrm{F,n}}(\mathbf{x}) = R(\mathbf{x}). \tag{1.67}$$

The numerical solution of these equations is plagued by the fact that the charge carrier densities may vary by several orders of magnitude, which renders the resulting linear system of equations to be conditioned in an extremely poor way.

As for the Poisson equation, the unique solution of Eq. (1.66) and Eq. (1.67) requires specifying boundary conditions. Boundary conditions have to be specified at the boundaries of the nanostructure as well as at the interfaces to contacts. Contacts are part of the nanostructure, but inside of contacts the current equations are not solved. Instead, assumptions on the quasi Fermi levels within the contacts are made that are then incorporated in terms of appropriate boundary conditions.

From a physical point of view two types of contacts have to be distinguished, namely Ohmic contacts and Schottky contacts:

1. An Ohmic contact is characterized by the fact that the space charge region at the metal-semiconductor interface is so small that charge carriers can easily tunnel through the corresponding energy barrier. Close to the contact the electrostatic potential is determined by the requirement of local charge neutrality

2. At a Schottky contact the conduction band edge is fixed to a given value above the Fermi level due to surface states. The energy difference between the conduction band edge and the Fermi level is denoted as Schottky barrier and depends on the ionization energy of the surface states. The charge of the occupied surface states results in a non-vanishing electric field at the contact.

In the following, we explain how these contacts are modeled in terms of boundary conditions:

1. Ohmic contact: The bias voltage V_{bias} that is applied to an Ohmic contact C fixes the quasi Fermi levels within the contact. For the quasi Fermi level at the metal-semiconductor interface ∂C we therefore choose Dirichlet boundary conditions

$$E_{\mathrm{F,p}}\left(\mathbf{x}\right)\big|_{\mathbf{x}\in\partial C} = E_{\mathrm{F,n}}\left(\mathbf{x}\right)\big|_{\mathbf{x}\in\partial C} = -eV_{\mathrm{bias}}. \tag{1.68}$$

   For the electrostatic potential we employ von Neumann boundary conditions

$$\partial_{\mathbf{n}}\phi\left(\mathbf{x}\right)\big|_{\mathbf{x}\in\partial C} = 0 \tag{1.69}$$

   that correspond to a vanishing electric field.

2. Schottky contacts: As for the Ohmic contact, the applied bias voltage V_{bias} fixes the quasi Fermi levels within the contacts and we therefore use Dirichlet boundary conditions for the quasi Fermi levels at the interface

$$E_{\mathrm{F,p}}\left(\mathbf{x}\right)\big|_{\mathbf{x}\in\partial C} = E_{\mathrm{F,n}}\left(\mathbf{x}\right)\big|_{\mathbf{x}\in\partial C} = -eV_{\mathrm{bias}}. \tag{1.70}$$

   The conduction band edge E_{c} is fixed above the Fermi energy by the Schottky barrier $-e\phi_{\mathrm{B}}$, which results in a Dirichlet boundary condition for the electrostatic potential at the interface ∂C that is given by

$$\phi\left(\mathbf{x}\right)\big|_{\mathbf{x}\in\partial C} = \frac{1}{e}E_{\mathrm{c}} + V_{\mathrm{bias}} + \phi_{B}. \tag{1.71}$$

In addition, we also allow specifying Dirichlet boundary conditions exclusively for the quasi Fermi levels by means of so called Fermi contacts.

1.2.4 Strain equation

Elasticity theory

If semiconductor materials with different lattice constants are epitaxially grown on top of each other, lattice mismatch results in a displacement of the crystal's atoms from their rest positions producing strain. We use a continuum model on the basis of elasticity theory to calculate these strain fields. This is based on the Euler-Cauchy stress principle and Cauchy's fundamental postulate [23]. In differential form, the resulting equation for the Cauchy stress tensor σ_{ij} is given by

$$\partial_j \sigma_{ij} = -f_j, \qquad i,j = 1,\ldots,3 \tag{1.72}$$

where $\mathbf{f} = (f_1, f_2, f_3)$ is an external volume force (for example gravity). We make the approximation $\mathbf{f} = 0$, i.e. we assume that all external volume forces vanish. For small strain, the stress tensor is related linearly to the strain tensor by means of Hooke's law

$$\sigma_{ij} = E_{ijkl}\varepsilon_{kl}, \tag{1.73}$$

where E_{ijkl} denotes the fourth rank elasticity tensor. The strain tensor is related to the displacement vector $\mathbf{u} = (u_1, u_2, u_3)$ according to

$$\varepsilon_{kl} = \frac{1}{2}\left(\partial_l u_k + \partial_k u_l\right). \tag{1.74}$$

The displacement vector $\mathbf{u}$ characterizes the displacement of a particle from its rest position. From Eq. (1.74) it follows immediately that the strain tensor is symmetric. Therefore, only six components of the strain tensor are independent. This allows us to introduce the following notation due to Voigt

$$\begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \\ \varepsilon_5 \\ \varepsilon_6 \end{pmatrix} = \begin{pmatrix} \varepsilon_{11} \\ \varepsilon_{22} \\ \varepsilon_{33} \\ 2\varepsilon_{23} \\ 2\varepsilon_{13} \\ 2\varepsilon_{12} \end{pmatrix}, \quad \begin{pmatrix} \sigma_1 \\ \sigma_2 \\ \sigma_3 \\ \sigma_4 \\ \sigma_5 \\ \sigma_6 \end{pmatrix} = \begin{pmatrix} \sigma_{11} \\ \sigma_{22} \\ \sigma_{33} \\ \sigma\varepsilon_{23} \\ \sigma_{13} \\ \sigma_{12} \end{pmatrix}. \tag{1.75}$$

In the Voigt notation, Hooke's law can be written as a simple matrix-vector multiplication as follows:

$$\sigma_i = E_{ij}\varepsilon_j. \tag{1.76}$$

The elasticity tensors for zinc-blende and wurtzite crystal structure in Voigt notation are given by

$$E_{\text{zb}} = \begin{pmatrix} c_{11} & c_{12} & c_{12} & & & \\ c_{12} & c_{11} & c_{12} & & & \\ c_{12} & c_{12} & c_{11} & & & \\ & & & c_{44} & & \\ & & & & c_{44} & \\ & & & & & c_{44} \end{pmatrix} \tag{1.77}$$

and

$$E_{\mathrm{wz}} = \begin{pmatrix} c_{11} & c_{12} & c_{13} & & & \\ c_{12} & c_{11} & c_{13} & & & \\ c_{13} & c_{13} & c_{33} & & & \\ & & & c_{44} & & \\ & & & & c_{44} & \\ & & & & & c_{66} \end{pmatrix}. \tag{1.78}$$

Note that the elasticity tensors are characterized by only a small number of elasticity constants due to the symmetry of the zinc-blende and wurtzite crystals.

Strain has two major effects that will be discussed in the following. First, strain is responsible for the piezoelectric polarization charges that contribute to the charge density in the Poisson equation and therefore change the electrostatic potential. Second, strain induces shifts in the band edges via the deformation potentials and therefore influences directly the electronic structure.

Polarization charges

In III-V semiconductors strain leads to a displacement of the centers of positive and negative charges within a unit cell and therefore to the formation of microscopic electric dipoles. At surfaces, interfaces, or in inhomogeneous strain fields, these microscopic dipoles are not compensated and therefore yield macroscopic, so called piezoelectric polarization charges. Semiconductors in wurtzite crystal structure show an additional permanent polarization independent of strain in the direction of the c-axis. This also leads to macroscopic polarization charges at surfaces or interfaces that are called pyroelectric charges.

The piezoelectric polarization P^{piezo} is proportional to the strain tensor

$$P_i^{\mathrm{piezo}} = e_{ijk}\varepsilon_{jk}. \tag{1.79}$$

The proportionality constants form a third rank tensor e_{ijk} that in general has 27 entries. In the crystal coordinate system the number of constants can be drastically reduced. In Voigt notation, the only non-zero component of e_{ijk} for zinc-blende crystal structure is e_{14} and the relation between the piezoelectric polarization vector and the strain tensor is given by

$$\mathbf{P}_{\mathrm{zb}}^{\mathrm{piezo}}(\mathbf{x}) = e_{14}(\mathbf{x}) \begin{pmatrix} 2\varepsilon_{23}(\mathbf{x}) \\ 2\varepsilon_{13}(\mathbf{x}) \\ 2\varepsilon_{12}(\mathbf{x}) \end{pmatrix}. \tag{1.80}$$

Similarly, for wurtzite crystal structure, there are only three non-zero components, namely e_{15}, e_{31}, and e_{33}. For the relation between the piezoelectric polarization vector and the strain tensor the following equation holds:

$$\mathbf{P}_{\mathrm{wz}}^{\mathrm{piezo}}(\mathbf{x}) = \begin{pmatrix} 2e_{15}(\mathbf{x})\varepsilon_{13}(\mathbf{x}) \\ 2e_{15}(\mathbf{x})\varepsilon_{12}(\mathbf{x}) \\ e_{31}(\mathbf{x})\varepsilon_{11}(\mathbf{x}) + e_{31}(\mathbf{x})\varepsilon_{22}(\mathbf{x}) + e_{33}(\mathbf{x})\varepsilon_{33}(\mathbf{x}) \end{pmatrix}. \tag{1.81}$$

To obtain the full polarization vector for wurtzite crystal structure the contribution

$$\mathbf{P}_{\mathrm{wz}}^{\mathrm{pyro}}(\mathbf{x}) = \begin{pmatrix} 0 \\ 0 \\ p(\mathbf{x}) \end{pmatrix} \tag{1.82}$$

of the pyroelectric polarization has to be added. Here $p(\mathbf{x})$ is the material dependent pyroelectric constant.

The polarization charges are now obtained by forming the divergence of the polarization vector

$$\rho_{\text{pol}}(\mathbf{x}) = -\nabla \cdot \mathbf{P}^{\text{pol}}(\mathbf{x}), \tag{1.83}$$

$$\mathbf{P}^{\text{pol}}(\mathbf{x}) = \begin{cases} \mathbf{P}_{\text{zb}}^{\text{piezo}}(\mathbf{x}) & \text{for zinc-blende} \\ \mathbf{P}_{\text{wz}}^{\text{piezo}}(\mathbf{x}) + \mathbf{P}_{\text{wz}}^{\text{pyro}}(\mathbf{x}) & \text{for wurtzite} \end{cases}. \tag{1.84}$$

No polarization charges are present in regions where $\mathbf{P}^{\text{pol}}(\mathbf{x})$ is constant. However at interfaces, where the piezoelectric or pyroelectric constants change abruptly from one material to another material, polarization charges may occur. The same is true if the piezoelectric constants are position independent, but the strain fields are inhomogeneous.

Strain induced shift of the band edges

Besides polarization charges, strain has a second, more direct effect on the electronic structure in semiconductor nanostructures. Strain causes a shift of the band edges of conduction bands and valence bands relative to each other by material dependent amounts. This can be described by deformation potentials. For the conduction band at the Γ-point strain results in a shift proportional to the trace of the strain tensor:

$$E_{\text{c}}^{\Gamma} \to E_{\text{c}}^{\Gamma} + a_{\text{c}} \operatorname{Tr} \varepsilon. \tag{1.85}$$

The proportionality constant is the absolute deformation potential a_{c} for the conduction band. Note that the trace of the strain tensor corresponds to the change in volume. For the conduction band valleys at the X-point, the L-point, and along Δ the effect of strain also depends on the particular position $\mathbf{k}_0$ of the valley [24]

$$E_{\text{c}}^{i}\left(\hat{\mathbf{k}}_0\right) \to E_{\text{c}}^{i}\left(\hat{\mathbf{k}}_0\right) + \Xi_{\text{d}}^{i} \operatorname{Tr} \varepsilon + \Xi_{\text{u}}^{i}\left(\hat{\mathbf{k}}_0 \cdot \varepsilon \hat{\mathbf{k}}_0\right), \quad i = X, L, \Delta. \tag{1.86}$$

One part of the shift is again due to the change in volume and therefore proportional to the trace of the strain tensor. The proportionality constant Ξ_{d}^{i} is again denoted as absolute deformation potential and may differ for $i = X, L, \Delta$. The second part also includes the off-diagonal elements of the strain tensor that describe sheer strain. The $\mathbf{k}_0$ dependency of this term may lead to a splitting of the degenerate valleys of each of the Brillouin zone points X, L, or along Δ. The proportionality constant Ξ_{u}^{i} is denoted as uniaxial deformation potential.

In the context of the $k \cdot p$ model, strain leads to an additional contribution to the $k \cdot p$ Hamiltonian matrix for the valence bands

$$H_{3\times3}(\mathbf{k}) \to H_{3\times3}(\mathbf{k}) + S(\varepsilon), \tag{1.87}$$

that for zinc-blende is given by

$$S^{\text{zb}}(\varepsilon) = n\varepsilon + (l-n)\operatorname{diag}(\varepsilon_{11}, \varepsilon_{22}, \varepsilon_{33}) + m \operatorname{diag}(\varepsilon_{22}+\varepsilon_{33}, \varepsilon_{11}+\varepsilon_{33}, \varepsilon_{11}+\varepsilon_{22}), \tag{1.88}$$

where any $\mathbf{k}$ depending terms have been neglected [19]. The parameters m, n, and l are related to the absolute deformation potential for the valence band a_{v} and to the sheer deformation potentials b and d as follows:

$$l = a_{\text{v}} + 2b, \tag{1.89}$$

$$m = a_{\mathrm{v}} - b, \tag{1.90}$$

$$n = \sqrt{3}d. \tag{1.91}$$

For wurtzite, the contribution of strain to the $k \cdot p$ Hamiltonian matrix is given in Appendix C.

To calculate the energies of the valence band edges of heavy holes, light holes and split-off holes, the following 6×6 $k \cdot p$ Hamiltonian matrix

$$H_{\varepsilon}^{\mathrm{zb},6\times 6}(\mathbf{k}) = \left[H^{\mathrm{zb},3\times 3}(\mathbf{k}) + S^{\mathrm{zb}}(\varepsilon)\right] \otimes \mathbf{1}_{2\times 2} + H_{\mathrm{SO}}^{\mathrm{zb},6\times 6} \tag{1.92}$$

has to be diagonalized for $\mathbf{k} = 0$. Besides a shift of the valence band edges, this also results in a mixture of the heavy hole, light hole, and split-off hole states that are the eigenvectors of $H_{\mathrm{SO}}^{\mathrm{zb},6\times 6}$.

1.3 Summary

In this chapter, we have given an overview on the physical background for the simulation of semiconductor nanostructures for equilibrium as well as non-equilibrium situations. We have illustrated the band structure of bulk semiconductors and we have presented the $k \cdot p$ model in combination with the envelope function approximation as a method for the calculation of the electronic structure of semiconductor nanostructures. Here we take into account holes in the heavy hole, light hole, and split-off valence bands and electrons in the conduction band valleys at the Γ-point, the X- and L-points and along Δ. A non-vanishing distribution of space charges occurs as a consequence of the ionization of impurity atoms, polarization charges, fixed charges at interfaces, or as charge carriers in terms of holes and electrons. These charges are the source of the built-in electrostatic potential that has an important influence on the electronic structure. The application of bias voltages to the contacts of a nanostructure drives the system out of thermodynamic equilibrium and causes electric currents to flow. In the context of the drift diffusion model, this is described by position dependent quasi Fermi levels for holes and electrons. In epitaxially grown nanostructures, lattice mismatch between different semiconductor materials results in non-vanishing strain fields that influence the electronic structure in terms of deformation potentials and piezoelectric charges.

Chapter 2

Numerical techniques

2.1 Introduction

The equations that need to be solved in next**nano**++ are sets of coupled partial differential equations (PDEs) in position space. This immediately necessitates their discretization on a grid. In order to achieve a highly efficient implementation, non-uniform tensor product grids are used. The mapping of the PDEs onto the grid is achieved using the box integration finite differences scheme (also termed box discretization in the following), since this discretization scheme is flux conserving in the presence of material discontinuities as they naturally occur in the simulation of heterostructures.

As a result of the discretization procedure, every differential operator in the PDEs becomes a sparse $N \times N$ matrix, where N is the number of nodes in the grid. For three spatial dimensions N is easily of the order of 10^6 (this corresponds to only 100 grid points in each direction). Here we see immediately that three-dimensional simulations require extremely large matrices that cannot be easily stored and processed, unless one explicitly exploits its sparsity structure with its many vanishing entries.

After discretization, every PDE that is a boundary value problem (such as the Poisson equation, the strain equation, and the current equation) becomes a large system of linear or non-linear equations (N equations and unknowns). The Schrödinger equation as an operator eigenvalue problem becomes a similarly large matrix eigenvalue problem with discretized wave functions as eigenvectors. Due to the sparsity of the matrices, these large matrix problems can be efficiently solved using iterative solvers such as the conjugate gradient method or Arnoldi iterations. An in-depth presentation of the subject of iterative methods for solving linear systems can be found in Reference [25]. Detailed information on a large number of methods for the iterative solution of large matrix eigenvalue problems is given in Reference [26]. Here we want to give a short overview on each of the methods employed within next**nano**++ to illustrate their application therein.

To deal with non-linearities as in the Poisson equation, standard algorithms such as the Newton method with line search are employed. Elaborate iteration schemes as the predictor-corrector approach [27] are used for the coupled Schrödinger and Poisson equations. The coupled Schrödinger, current, and Poisson equations are solved using an underrelaxation approach for the quasi Fermi levels. These methods and their application within next**nano**++ are also discussed in the following.

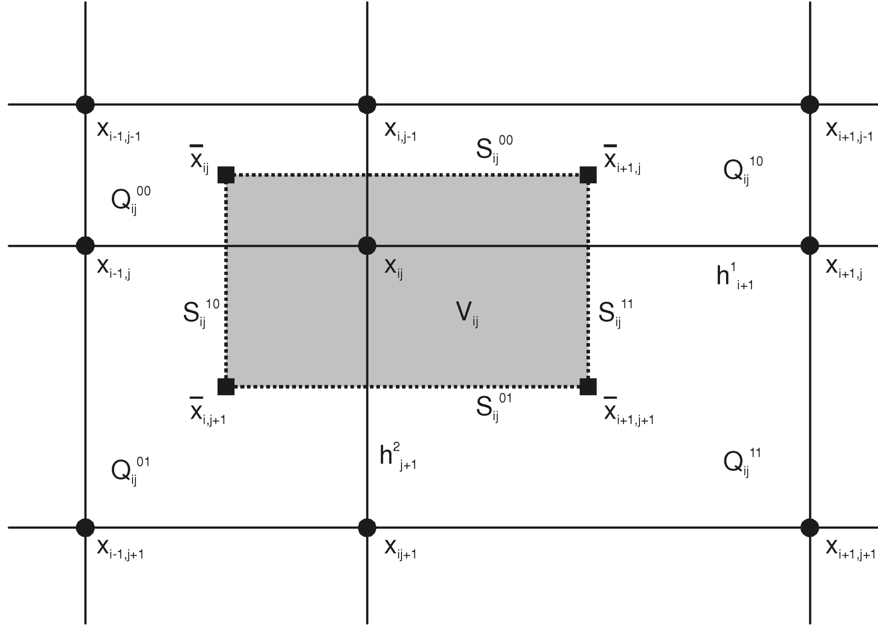

Figure 2.1: For box integration each grid node x_{ij} is surrounded by a rectangular control volume V_{ij} (gray). Material properties are assumed to be constant in the four quadrants $Q_{ij}^{\nu\mu}$ of each grid node x_{ij}.

2.2 Box discretization

Box discretization has the advantage of being flux conserving also in the presence of material discontinuities as they naturally occur in the simulation of nanostructures. For box discretization, we cover the simulation domain with N non-overlapping rectangular boxes each belonging to one of the grid nodes. We then integrate our differential equations over each box volume to obtain an integral equation that is then discretized using the finite differences scheme. We illustrate this approach here by means of the Poisson equation in two spatial dimensions with a scalar dielectric constant. All the details concerning the application of box discretization within next**nano**++ can be found in Reference [14].

The Poisson equation is given by

$$\boldsymbol{\nabla} \cdot \varepsilon(\mathbf{x})\, \boldsymbol{\nabla}\phi(\mathbf{x}) = -\rho(\mathbf{x}) \tag{2.1}$$

and we want to find an approximate solution for the electrostatic potential ϕ at the grid nodes $\mathbf{x}_{ij} = \left(x_i^1, x_j^2\right)$, where x_i^1, x_j^2 are the ith and jth elements of the ordered sets $\left\{x_1^1, x_2^1, \ldots, x_{N_1}^1\right\}$ and $\left\{x_1^2, x_2^2, \ldots, x_{N_2}^2\right\}$, respectively. As shown in Fig. 2.1, for each grid node $\mathbf{x}_{ij}$ we introduce a surrounding box

$$V_{ij} = \left[\bar{x}_i^1, \bar{x}_{i+1}^1\right] \times \left[\bar{x}_j^2, \bar{x}_{j+1}^2\right] \tag{2.2}$$

where

$$\bar{x}_i^1 = \frac{1}{2}\left(x_{i-1}^1 + x_i^1\right), \quad \bar{x}_j^2 = \frac{1}{2}\left(x_{j-1}^2 + x_j^2\right). \tag{2.3}$$

The surrounding box V_{ij} is also termed the control volume. We further define grid spacings h_i^1 and h_j^2 between neighboring nodes

$$h_i^1 = x_i^1 - x_{i-1}^1, \quad h_j^2 = x_j^2 - x_{j-1}^2. \tag{2.4}$$

Further, we introduce quadrants

$$Q_{ij}^{\mu\nu} = \left[x_{i-1+\mu}^1, x_{i+\mu}^1\right] \times \left[x_{j-1+\nu}^2, x_{j+\nu}^2\right], \quad \mu, \nu = 0, 1. \tag{2.5}$$

Each of these quadrants is assumed to be filled by a constant material and accordingly, material parameters such as the dielectric constant $\varepsilon(\mathbf{x})$ are assumed to be the same for each $\mathbf{x} \in Q_{ij}^{\mu\nu}$. To derive now a discretization for Eq. (2.1) we proceed as follows: For each grid node we integrate Eq. (2.1) over the control volume

$$\int_{V_{ij}} \mathrm{d}^2\mathbf{x}\, \boldsymbol{\nabla} \cdot \varepsilon(\mathbf{x})\, \boldsymbol{\nabla}\phi(\mathbf{x}) = -\int_{V_{ij}} \mathrm{d}^2\mathbf{x}\, \rho(\mathbf{x}). \tag{2.6}$$

The left hand side of Eq. (2.6) is cast into a surface integral by means of Gauss's theorem

$$\int_{V_{ij}} \mathrm{d}^2\mathbf{x}\, \boldsymbol{\nabla} \cdot \varepsilon(\mathbf{x})\, \boldsymbol{\nabla}\phi(\mathbf{x}) = \int_{\partial V_{ij}} \mathrm{d}s\, \mathbf{n} \cdot \varepsilon(\mathbf{x})\, \boldsymbol{\nabla}\phi(\mathbf{x}) \tag{2.7}$$

where $\mathbf{n}$ denotes the normal to the boundary ∂V_{ij} of the control volume. This boundary is given by

$$\partial V_{ij} = \bigcup_{\mu,\nu=0,1} S_{ij}^{\mu\nu}, \tag{2.8}$$

where for $\nu = 0, 1$

$$S_{ij}^{0\nu} = \left[\bar{x}_i^1, \bar{x}_{i+1}^1\right] \times \left\{\bar{x}_{j+\nu}^2\right\}, \tag{2.9}$$

$$S_{ij}^{1\nu} = \left\{\bar{x}_{i+\nu}^1\right\} \times \left[\bar{x}_j^2, \bar{x}_{j+1}^2\right]. \tag{2.10}$$

By means of the line segments $S_{ij}^{\mu\nu}$ the right hand side of Eq. (2.7) now reads

$$\int_{\partial V_{ij}} \mathrm{d}s\, \mathbf{n} \cdot \varepsilon(\mathbf{x})\, \boldsymbol{\nabla}\phi(\mathbf{x}) = \sum_{\mu,\nu=0,1} (-1)^{\nu+1} \int_{S_{ij}^{\mu\nu}} \mathrm{d}x_{\mu+1}\, \varepsilon(\mathbf{x})\, \partial_{\mu+1}^{\perp}\phi(\mathbf{x}). \tag{2.11}$$

We only consider the case $\mu = 0$, the calculations for $\mu = 1$ can be carried out in an analogous way

$$\sum_{\nu=0,1} (-1)^{\nu+1} \int_{S_{ij}^{0\nu}} \mathrm{d}x_1\, \varepsilon(\mathbf{x})\, \partial_2\phi(\mathbf{x}) = \sum_{\nu=0,1} (-1)^{\nu+1} \int_{S_{ij}^{0\nu}} \mathrm{d}x_1\, \varepsilon\left(x_1, \bar{x}_{j+\nu}^2\right) \partial_2\phi(x_1,x_2)\big|_{x_2=\bar{x}_{j+\nu}^2}. \tag{2.12}$$

We now use finite differences to approximate the derivatives at the center of the segments $S_{ij}^{\mu\nu}$

$$\partial_2\phi(x_1,x_2)\big|_{x_2=\bar{x}_{j+\nu}^2} \approx (-1)^{\nu+1} \frac{\phi\left(\bar{x}_i^1, x_{j-1+2\nu}^2\right) - \phi\left(\bar{x}_i^1, x_j^2\right)}{h_{j+\nu}^2}. \tag{2.13}$$

The remaining integral can then be easily evaluated taking into account that $\varepsilon(\mathbf{x})$ is constant within each quadrant $Q_{ij}^{\mu\nu}$

$$\int_{S_{ij}^{1\nu}} \mathrm{d}x_1\, \varepsilon\left(x_1, \bar{x}_{j+\nu}^2\right) = \frac{1}{2} \sum_{\sigma=0,1} h_{i+\sigma}^1 \varepsilon\left(\bar{x}_{i+\sigma}^1, \bar{x}_{j+\nu}^2\right), \tag{2.14}$$

Thus we obtain for the left hand side of Eq. (2.12)

$$\sum_{\nu=0,1} (-1)^{\nu+1} \int_{S_{ij}^{0\nu}} \mathrm{d}x_1 \, \varepsilon(\mathbf{x}) \, \partial_2 \phi(\mathbf{x}) \tag{2.15}$$

$$\approx \frac{1}{2} \sum_{\sigma,\nu=0,1} \varepsilon\left(\bar{x}^1_{i+\sigma}, \bar{x}^2_{j+\nu}\right) \frac{h^1_{i+\sigma}}{h^2_{j+\nu}} \left[\phi\left(\bar{x}^1_i, x^2_{j-1+2\nu}\right) - \phi\left(\bar{x}^1_i, x^2_j\right)\right]. \tag{2.16}$$

Now, forming the sum for $\mu = 0$ and $\mu = 1$ yields the following result for the discretization of the left hand side of Eq. (2.6)

$$\int_{V_{ij}} \mathrm{d}^2\mathbf{x} \, \boldsymbol{\nabla} \cdot \varepsilon(\mathbf{x}) \, \boldsymbol{\nabla}\phi(\mathbf{x}) \tag{2.17}$$

$$\approx \frac{1}{2} \sum_{\sigma,\nu=0,1} \varepsilon\left(\bar{x}^1_{i+\sigma}, \bar{x}^2_{j+\nu}\right) \left\{ \frac{h^1_{i+\sigma}}{h^2_{j+\nu}} \left[\phi\left(\bar{x}^1_i, x^2_{j-1+2\nu}\right) - \phi\left(\bar{x}^1_i, x^2_j\right)\right] \right. \tag{2.18}$$

$$\left. + \frac{h^2_{j+\nu}}{h^1_{i+\sigma}} \left[\phi\left(\bar{x}^1_{i-1+2\sigma}, x^2_j\right) - \phi\left(\bar{x}^1_i, x^2_j\right)\right] \right\}. \tag{2.19}$$

We also want to spend a few words on the right hand side of Equation (2.6). Here, the problem is that the discretized charge density $\rho(\mathbf{x})$ is composed of two components

$$\rho(\mathbf{x}) = \rho_{\text{quad}}(\mathbf{x}) + \rho_{\text{grid}}(\mathbf{x}) \tag{2.20}$$

where $\rho_{\text{quad}}(\mathbf{x})$ is defined on the quadrants $Q^{\mu\nu}_{ij}$ and has a constant value within each of the quadrants, whereas $\rho_{\text{grid}}(\mathbf{x})$ is defined on the grid nodes only. The origin of $\rho_{\text{quad}}(\mathbf{x})$ is any kind of material property that leads to fixed charges as for example doping, pyro- or piezoelectricity, or surface charges. In contrast, the origin of $\rho_{\text{grid}}(\mathbf{x})$ is the free charge carrier density due to electrons and holes. These two components have to be treated differently in evaluating the right hand side of Eq. (2.6) as follows

$$-\int_{V_{ij}} \mathrm{d}^2\mathbf{x} \, \rho(\mathbf{x}) = -\frac{1}{4} \sum_{\sigma,\nu=0,1} h^1_{i+\sigma} h^2_{j+\nu} \rho_{\text{quad}}\left(\bar{x}^1_{i+\sigma}, \bar{x}^2_{j+\nu}\right) - |V_{ij}| \, \rho_{\text{grid}}(\mathbf{x}_{ij}). \tag{2.21}$$

where $|V_{ij}|$ denotes the volume of the surrounding box V_{ij}.

2.3 Iterative solvers

The two basic numerical problems that have to be solved by next**nano**++ are linear systems of equations

$$A\mathbf{x} = \mathbf{b} \tag{2.22}$$

and large matrix eigenvalue problems

$$A\mathbf{x} = \lambda\mathbf{x}, \tag{2.23}$$

where A is a non-singular $N \times N$ matrix and $\mathbf{b}$ is a given N-vector. The matrices that arise in next**nano**++ from the discretization of the partial differential equations are sparse, having only a few non-zeros per row. In contrast to the well-known Gaussian elimination method for solving linear systems of equations, this sparseness can be fully exploited by solvers based on

matrix-vector multiplications. If a matrix has only a few non-zero elements per row, the number of operations required to compute the product of that matrix with a given vector is just a few N instead of $2N^2$ operations required for a general dense matrix-vector multiplication.

Solvers based on matrix-vector multiplications are denoted as iterative solvers. If $A\mathbf{x} = \mathbf{b}$ can be solved by a moderate number of matrix-vector multiplications and little additional work, this iterative procedure may far outperform Gaussian elimination in terms of both work and storage. Iterative solvers are known not only for the solution of linear systems of equations, but also for large matrix eigenvalue problems.

The iterative solvers used in next**nano**++ for linear systems of equations as well as for eigenvalue problems are so called Krylov subspace methods. The basic idea is to construct the Krylov subspace

$$K = \text{span}\left\{\mathbf{b}, A\mathbf{b}, \ldots, A^{k-1}\mathbf{b}\right\} \tag{2.24}$$

starting here from the only vector associated with the problem, i.e. the right hand side of Equation (2.22). One then tries to find a good approximation for $\mathbf{x}$ in K. This immediately raises the following questions: How good is the best approximate solution that is contained in K? And how can an optimal approximation from K be computed with a moderate amount of work and storage? The answers to these questions lead to a multitude of different methods of which we want to present in the following those employed within next**nano**++.

2.4 Solving linear systems of equations

In the following we will shortly discuss the three methods that are used within next**nano**++ for solving linear systems of equations. These methods are the conjugate gradient method (CG) [28], the composite step conjugate gradient method (CSCG), and the biconjugate gradient method (BiCG) [29]. The common feature of these methods is that they are all based on the Krylov subspace approximation. We do not use these methods in their original form but employ preconditioners to help in accelerating the convergence.

The conjugate gradient method (CG)

CG is the method of choice for Hermitian positive definite problems. These arise typically when discretizing the linear Poisson equation, the current equation, and the strain equation. In these cases the resulting matrix is real valued and therefore even symmetric. For the strain equation this is true only if non-periodic boundary conditions are applied. For periodic boundary conditions, discretization yields an unsymmetric matrix and CG cannot be used. In this case BiCG is our method of choice.

It turns out that CG is often also good for indefinite problems. As will be discussed in the section on eigenvalue problems, these indefinite matrices arise when discretizing the 8-band $k \cdot p$ equation. There, CG is used for the shift-invert spectral transformation that allows one to calculate inner eigenvectors using the standard Arnoldi iteration. Because CG is so widely used we want to provide a description of this algorithm here: Given an initial guess $\mathbf{x}_0$, compute $\mathbf{r}_0 = \mathbf{b} - A\mathbf{x}_0$ and set $\mathbf{p}_0 = \mathbf{r}_0$. For $k = 1, 2, \ldots$

1. Compute $A\mathbf{p}_{k-1}$.

2. Set $\mathbf{x}_k = \mathbf{x}_{k-1} + a_{k-1}\mathbf{p}_{k-1}$, where $a_{k-1} = \mathbf{r}_{k-1} \cdot \mathbf{r}_{k-1} / \left(\mathbf{p}_{k-1} \cdot A\mathbf{p}_{k-1}\right)$.

3. Compute $\mathbf{r}_k = \mathbf{r}_{k-1} - a_{k-1} A \mathbf{p}_{k-1}$.

4. Set $\mathbf{p}_k = \mathbf{r}_k + b_{k-1}\mathbf{p}_{k-1}$, where $\mathbf{b}_{k-1} = \mathbf{r}_k \cdot \mathbf{r}_k / (\mathbf{r}_{k-1} \cdot \mathbf{r}_{k-1})$.

The composite step conjugate gradient method (CSCG)

In contrast to CG, CSCG has been explicitly developed also to solve Hermitian indefinite problems. It is our method of choice if CG fails for the shift-invert spectral transformation. CSCG can also be used for the Poisson equation, the current equation, and the strain equation (only non-periodic boundary conditions), but CSCG is slower than CG by approximately a factor of 2. The composite step conjugate gradient method is similar to the standard conjugate gradient method, except that it occasionally proceeds from the iterate for step k to the iterate for step $k+2$. Such composite steps are taken to improve the stability of the recurrence relations and smooth the behavior of the residual norm. Details of composite step methods are presented in Reference [30].

The biconjugate gradient method (BiCG)

BiCG has to be used for the strain equation if periodic boundary conditions are applied. In this case, discretization yields an unsymmetric matrix. In contrast to CG and CSCG, BiCG can also deal with non-Hermitian matrices.

Preconditioners

To help in accelerating the convergence, we do not use the presented methods in their original form but employ preconditioners. Generally speaking, a preconditioner for solving a linear system of equations is an appropriately chosen matrix M. The linear system of equations $A\mathbf{x} = \mathbf{b}$ is multiplied with the inverse M^{-1} from the left and instead of $A\mathbf{x} = \mathbf{b}$ one effectively solves the modified problem

$$M^{-1}A\mathbf{x} = M^{-1}\mathbf{b} \tag{2.25}$$

with a hopefully better Krylov subspace by generating approximate solutions $\mathbf{x}_1, \mathbf{x}_2, \ldots$ satisfying

$$\mathbf{x}_k \in \operatorname{span}\left\{M^{-1}\mathbf{b}, \left(M^{-1}A\right) M^{-1}\mathbf{b}, \ldots, \left(M^{-1}A\right)^k M^{-1}\mathbf{b}\right\}. \tag{2.26}$$

At each step of the preconditioned algorithm, it is necessary to compute the product M^{-1} with a vector, or equivalently, to solve a linear system with coefficient matrix M. Therefore, M should be chosen so that such linear systems are much easier to solve than the original problem.

For next**nano**++ we use three types of preconditioners: the incomplete Cholesky [31], the Dupont-Kendall-Rachford [32], and a diagonal preconditioner. The Dupont-Kendall-Rachford preconditioner is used for the Poisson equation and for the current equation. The incomplete Cholesky decomposition is used for the strain equation with non-periodic boundary conditions. In conjunction with BiCG, we use the diagonal preconditioner for the strain equation with periodic boundary conditions. For the shift-invert spectral transformation we use either the incomplete Cholesky or the Dupont-Kendall-Rachford preconditioner.

Of the presented methods, diagonal preconditioning is the simplest approach. Here, one sets $M = \operatorname{diag}(A)$. Thus, the determination of M^{-1} is straightforward.

The incomplete Cholesky preconditioner is based on the Cholesky factorization $A = LL^H$ that can be performed for any Hermitian, positive definite matrix A, where L is a lower triangular matrix. However, if A is a sparse matrix, then the lower triangular factor L is usually much less sparse than A and the entire band "fills in" during Gaussian elimination. By restricting the lower triangular matrix L to have the sparsity pattern of the lower triangular of A, one might obtain an approximate factorization of A. The nonzeros of L are then chosen so that the product LL^H matches the non-zero entries of A, but of course, LL^H does not match A everywhere. An approximate factorization of this form is called incomplete Cholesky decomposition. The matrix $M = LL^H$ can then be used as a preconditioner. The linear system $M\mathbf{z} = \mathbf{r}$ is solved in two steps: first one solves the lower triangular system $L\mathbf{y} = \mathbf{r}$, then the upper triangular system $L^H\mathbf{z} = \mathbf{y}$. For symmetric A with specific sparsity pattern (as it is the case for the Poisson, the current, and the strain equation), the incomplete Cholesky decomposition can also be written in the form LDL^T, where D is a diagonal matrix. The determination of L and D is discussed in detail in Reference [25].

The Dupont-Kendall-Rachford preconditioner is a modification of the incomplete Cholesky preconditioner and also fill-in free. The preconditioner matrix M is also given by a product of the form LDL^T, with a lower triangular matrix L and a diagonal matrix D. For the Dupont-Kendall-Rachford method, the matrices L and D depend on a parameter η (termed DKR value). For $\eta = 0$ the Dupont-Kendall-Rachford preconditioner is equivalent to the incomplete Cholesky preconditioner.

2.5 Solving large matrix eigenvalue problems

Besides solving large systems of linear equations, the second main task in next**nano**++ is the determination of eigenvectors and eigenvalues of large matrix eigenvalue problems. With the exception of one-dimensional systems, storage and time limits the determination of eigenvectors and eigenvalues to only a few as compared to the total number N. In most cases, however, we are interested only in a few eigenvectors and eigenvalues close to the Fermi energy so that this restriction poses no actual problem.

In next**nano**++, two types of eigenvalue problems have to be solved. These are positive definite eigenvalue problems and indefinite eigenvalue problems. Positive definite eigenvalue problems arise from the single-band Schrödinger equation and from the 6-band $k \cdot p$ equation. An indefinite eigenvalue problem results from the 8-band $k \cdot p$ equation. Different methods have to be used to determine the desired eigenvectors and eigenvalues in these two cases. The reason for this is not a numerical one, as one could argue that by an appropriate shift an indefinite matrix eigenvalue problem should always become a definite matrix eigenvalue problem. Instead, the reason is a physical one: For the positive definite eigenvalue problems, we are interested in the lowest eigenvalues and corresponding eigenvectors. This is because the lowest eigenvalues are those close to the Fermi energy. However, for the indefinite eigenvalue problem, the eigenvalues close to the Fermi energy are non-extremal eigenvalues that cannot be determined in the same way as it is usually done for extremal eigenvalues.

In total, we use three different eigenvalue solvers: an eigenvalue solver based on Arnoldi iterations (from the ARPACK library [33]), the Jacobi-Davidson eigenvalue solver, and a dense matrix eigenvalue solver from the LAPACK library. The latter one is only used for one-dimensional problems. However, it can be employed for extremal as well as inner eigenvalues. The Jacobi-Davidson method as proposed by van der Voerst [34] is used as one of two methods for the

calculation of non-extremal eigenvalues. Arnoldi iterations are efficient in calculating extremal eigenvalues. This efficiency can be further increased by employing appropriate polynomial spectral transformations that amplify the lower part of the spectrum. By making use of a different type of spectral transformation known as shift-invert, the Arnoldi method can also be used to calculate non-extremal eigenvalues. For calculating non-extremal eigenvalues, the Arnoldi method with shift-invert has proven to be faster and more reliable in our applications than the Jacobi-Davidson algorithm. In the following, we will discuss in greater detail the Arnoldi iterations and spectral transformations that we use in next**nano**++ for the solution of extremal as well as non-extremal eigenvalue problems.

Extremal matrix eigenvalue problems resulting from positive definite operators

Because Arnoldi iterations are so widely used we want to provide a short description of this algorithm here. The Arnoldi method is an iterative method for the determination of a few eigenvalues and corresponding eigenvectors. For a given (complex valued) square matrix A and a given initial guess $\mathbf{q}_1$ with $\|\mathbf{q}_1\| = 1$, the Arnoldi method calculates an orthonormal basis of the Krylov subspace span $\left\{\mathbf{q}_1, A\mathbf{q}_1, A^2\mathbf{q}_1, \ldots, , A^k\mathbf{q}_1\right\}$. This is done as follows: For $j = 1, 2, \ldots$

1. $\tilde{\mathbf{q}}_{j+1} = A\mathbf{q}_j$. For $i = 1, \ldots, j$, $h_{ij} = \tilde{\mathbf{q}}_{j+1} \cdot \mathbf{q}_i$, $\tilde{\mathbf{q}}_{j+1} \leftarrow \tilde{\mathbf{q}}_{j+1} - h_{ij}\mathbf{q}_i$,
2. $h_{j+1,j} = \|\tilde{\mathbf{q}}_{j+1}\|$, $\mathbf{q}_{j+1} = \tilde{\mathbf{q}}_{j+1}/h_{j+1,j}$.

If Arnoldi iterations are used as an eigenvalue solver, one computes the so called Ritz eigenvalues of the projection of A onto the orthonormal basis of the Krylov subspace. These Ritz eigenvalues converge to the extremal eigenvalues of A.

By appropriate spectral transformations, the convergence of the Arnoldi method can be improved considerably. For extremal matrix eigenvalue problems, we use polynomial spectral transformations, either based on simple monomials x^n, Chebyshev polynomials $T_n(x)$, or Legendre polynomials $L_n(x)$. The basic idea behind this spectral transformation is that given a certain polynomial $p(x)$, the eigenvectors of the matrix A do not change by the transformation

$$A \rightarrow p(A) \tag{2.27}$$

whereas the spectrum does change. Therefore, one can determine the eigenvectors of A also by determining the eigenvectors of $p(A)$. The corresponding eigenvalues of A can be calculated afterwards: If $\mathbf{v}$ is a normalized eigenvector of A, the corresponding eigenvalue is given by $\mathbf{v} \cdot A\mathbf{v}$. By a transformation that amplifies the extremal part of the spectrum, the convergence of the Arnoldi method towards these extremal eigenvalues is significantly increased. In our applications, we have observed speed-ups by up to a factor of 10 for the Chebyshev polynomials as compared to the standard Arnoldi algorithm. By employing Chebyshev polynomials the relevant part of the spectrum can be amplified more significantly as compared to simple monomials. Therefore, the Chebyshev polynomials outrun the simple monomials by a factor of 2. However, this requires defining an energy cutoff that separates the interesting part of the spectrum from its remainder. To determine an appropriate energy cutoff in advance is difficult. In addition, if the cutoff energy is too small, the Arnoldi iterations fail to converge at all. For charge self-consistent problems, we have therefore developed the following robust strategy: In solving the eigenvalue problem for the first time, we always employ the Arnoldi method with the monomial spectral transformation that does not involve additional parameters. In solving

the eigenvalue problem for the next time, we employ the Arnoldi method with Chebyshev spectral transformation and use the highest eigenvalue found in the first iteration as a guess for the cutoff. If convergence fails nevertheless, we fall back on the monomial spectral transformation.

The advantage of using polynomials for the spectral transformation is that polynomials of a matrix A can be easily calculated. For this, one does not calculate $p(A)$ explicitly, but only matrix-vector products of $p(A)$ and a vector $\mathbf{q}$ employing the highly efficient Horner scheme.

Non-extremal matrix eigenvalue problems resulting from indefinite operators

Arnoldi iterations can also be used to determine non-extremal eigenvalues by means of the shift-invert spectral transformation. This spectral transformation is given by

$$A \rightarrow (A - \lambda)^{-1} \tag{2.28}$$

where λ denotes the shift. By the shift-invert spectral transformation, non-extremal eigenvalues around λ become extremal eigenvalues either at the lower or the upper part of the spectrum. ARPACK can then be used to specifically find those eigenvectors whose eigenvalues are either greater or less than λ.

The shift-invert spectral transformation is much more costly than polynomial spectral transformation as each Arnoldi iteration step now requires to solve the linear system of equations given by

$$(A - \lambda)\,\tilde{\mathbf{q}}_{j+1} = \mathbf{q}_j. \tag{2.29}$$

In our applications, $A - \lambda$ is Hermitian but indefinite. Nonetheless, it has turned out that in most cases CG can still be used to solve Equation (2.29).

In next**nano**++ Arnoldi iterations with a shift-invert spectral transformation are used for the calculation of eigenvectors and eigenvalues in the 8-band $k \cdot p$ model. There, we can often exploit the fact that the spectrum has a gap between the valence band and the conduction band states. If this gap exists, the energy range of the gap is known in advance at least approximately from the band structure in position space. In addition, the eigenvectors and eigenvalues that we are interested in are exactly those below and above the gap. We therefore set the shift λ to a value close to but smaller than the upper boundary of the gap and determine the eigenvectors and eigenvalues above the gap. Then we set the shift λ to a value close to but larger than the lower boundary of the gap and determine the eigenvectors and eigenvalues below the gap.

2.6 Solving non-linear systems of equations

In next**nano**++, a non-linear system of equations arises from the discretization of the non-linear Poisson equation

$$\boldsymbol{\nabla} \cdot \varepsilon(\mathbf{x})\,\boldsymbol{\nabla}\phi = -\rho(\phi)\,. \tag{2.30}$$

The non-linearity comes into play here by the non-linear dependence of the density ρ on the electrostatic potential ϕ, for example in terms of the Thomas-Fermi approximation for the charge densities. For solving Eq. (2.30) we employ the Newton-Raphson algorithm with inexact line search. First, we express Equation (2.30) in the following form

$$0 = \boldsymbol{\nabla} \cdot \varepsilon(\mathbf{x})\,\boldsymbol{\nabla}\phi + \rho(\phi) =: F(\phi) \tag{2.31}$$

in order to rewrite the problem into that of finding the zero of the function $F(\phi)$. The Newton-Raphson algorithm with inexact line search is then given as follows: For $k = 1, 2, \ldots$ and given ϕ_1

1. $\tilde{\phi}_k(\lambda) = \phi_k - \lambda \left[(\mathrm{D}_\phi F)(\phi_k)\right]^{-1} \cdot F(\phi_k)$.
2. For $\lambda = 1, \frac{1}{2}, \frac{1}{4}, \ldots,\ \phi_{k+1} = \tilde{\phi}_k(\lambda)$ with $\left\| F\left(\tilde{\phi}_k(\lambda)\right) \right\| = \min!$

In the first step of the algorithm, a new guess for the potential is made that still depends on the step length λ. This guess is based on the idea that to find the zero one should follow the direction of the steepest descent given by $\left[(\mathrm{D}_\phi F)(\phi_k)\right]^{-1} \cdot F(\phi_k)$. The second step of the algorithm describes the inexact line search: Along the direction of steepest descent one determines the step length λ for which $\|F(\phi)\|$ becomes minimal.

Except for the extra diagonal term $\mathrm{D}_\phi \rho$, the Jacobian matrix

$$(\mathrm{D}_\phi F)(\phi) = \boldsymbol{\nabla} \cdot \varepsilon(\mathbf{x}) \boldsymbol{\nabla} + (\mathrm{D}_\phi \rho)(\phi), \tag{2.32}$$

is identical to the Poisson operator $\boldsymbol{\nabla} \cdot \varepsilon(\mathbf{x}) \boldsymbol{\nabla}$. Since $\boldsymbol{\nabla} \cdot \varepsilon(\mathbf{x}) \boldsymbol{\nabla}$ is positive definite and $\mathrm{D}_\phi \rho$ is positive for all physical $\rho(\phi)$, the Jacobian $\mathrm{D}_\phi F$ must be positive as well. We therefore employ the preconditioned CG to solve the linear system in step 1 of the Newton-Raphson algorithm. As preconditioner we use either the incomplete Cholesky or the fill-in free Dupont-Kendall-Rachford method.

2.7 Solving coupled systems of equations

In next**nano**++, two sets of coupled systems of equations have to be solved. These are the coupled Schrödinger and Poisson equations for equilibrium situations and the coupled Schrödinger, current, and Poisson equations for non-equilibrium situations. We have to employ different solution strategies for the two cases. For the coupled Schrödinger and Poisson equations we use the highly efficient predictor-corrector approach developed by Trellakis *et al.* [27]. For the coupled Schrödinger, current, and Poisson equations, we have to use an approach based on underrelaxation. Accordingly, the solution of the coupled Schrödinger, current, and Poisson equations is less efficient. In the following, the two methods are described in closer detail. Both approaches can be combined with a newly developed idea that we have called subspace iterations and that helps to accelerate the calculation by saving time in the solution of the quantum mechanical eigenvalue problems [35].

Predictor-corrector approach for the coupled Schrödinger and Poisson equations

In solving the coupled Schrödinger and Poisson equations, the goal is to determine a charge self-consistent solution for the charge densities or equivalently for the electrostatic potential. The basic problem is that the Schrödinger equation

$$H(\phi)\,\psi_n = E_n \psi_n \tag{2.33}$$

depends on the electrostatic potential ϕ, whereas the Poisson equation depends on the eigenfunctions ψ_n

$$\boldsymbol{\nabla} \cdot \varepsilon(\mathbf{x}) \boldsymbol{\nabla} \phi = -\rho\left[\psi_1, \psi_2, \ldots\right]. \tag{2.34}$$

Due to the strong coupling between the equations, a simple iteration scheme does not converge. Similarly, underrelaxation stabilizes the iteration only poorly and strong charge oscillations from one iteration step to the next are observed that interfere with convergence.

This situation can be much improved by partially decoupling both partial differential equations using a predictor-corrector-type approach. To this end, the exact carrier densities $p(\mathbf{x})$ and $n(\mathbf{x})$ from Eq. (1.49) and Eq. (1.50) are replaced by predictors depending on the electrostatic potential ϕ

$$\tilde{p}[\phi](\mathbf{x}) = \sum_{i\in\mathrm{VB}} g_{\mathrm{v}}^{i} \sum_{j} p_{\mathrm{v},j}^{i}(\mathbf{x}) f\left(\left[E_{\mathrm{v},j}^{i} - e\left[\phi(\mathbf{x}) - \phi_{\mathrm{prev}}(\mathbf{x})\right] - E_{\mathrm{F,p}}(\mathbf{x})\right]/kT\right), \tag{2.35}$$

$$\tilde{n}[\phi](\mathbf{x}) = \sum_{i\in\mathrm{CB}} g_{\mathrm{c}}^{i} \sum_{j} p_{\mathrm{c},j}^{i}(\mathbf{x}) f\left(\left[-E_{\mathrm{c},j}^{i} + e\left[\phi(\mathbf{x}) - \phi_{\mathrm{prev}}(\mathbf{x})\right] + E_{\mathrm{F,n}}(\mathbf{x})\right]/kT\right), \tag{2.36}$$

where ϕ_{prev} denotes the electrostatic potential from the previous iteration step. These predictors for the quantum densities $n(\mathbf{x})$ and $p(\mathbf{x})$ are then used in the non-linear Poisson equation in order to determine the new potential $\phi(\mathbf{x})$. Using this $\phi(\mathbf{x})$ the Schrödinger equation is solved for a new set of eigenfunctions and eigenenergies. The method can be summarized as follows: For $k = 1, 2, \ldots$ and given ϕ_1

1. Solve $H(\phi_k)\psi_n^{(k)} = E_n^{(k)}\psi_n^{(k)}$ for a set of N_{ev} eigenfunctions $\psi_1^{(k)}, \psi_2^{(k)}, \ldots, \psi_{N_{\mathrm{ev}}}^{(k)}$ and corresponding eigenenergies $E_1^{(k)}, E_2^{(k)}, \ldots, E_{N_{\mathrm{ev}}}^{(k)}$.
2. Solve the non-linear Poisson equation $\boldsymbol{\nabla}\cdot\varepsilon(\mathbf{x})\boldsymbol{\nabla}\phi_{k+1} = -\tilde{\rho}\left(\phi_{k+1}\right)$, where $\tilde{\rho}$ is obtained from ρ by replacing n by $\tilde{n}$ and p by $\tilde{p}$ with $\phi_{\mathrm{prev}} = \phi_k$.

Once the iteration has converged we have $\phi = \phi_{\mathrm{prev}}$ and therefore $\tilde{p}[\phi](\mathbf{x}) = p(\mathbf{x})$, $\tilde{n}[\phi](\mathbf{x}) = n(\mathbf{x})$, i.e. the predictors yield the correct densities. This approach leads to rapid convergence and no further steps are necessary to ensure convergence.

As exit condition for the iteration we require that

$$\int \mathrm{d}^d\mathbf{x}\,\left|\tilde{p}\left[\phi_{k+1}\right](x) - \tilde{p}\left[\phi_k\right](x)\right| < \delta_{\mathrm{dens}}, \tag{2.37}$$

$$\int \mathrm{d}^d\mathbf{x}\,\left|\tilde{n}\left[\phi_{k+1}\right](x) - \tilde{n}\left[\phi_k\right](x)\right| < \delta_{\mathrm{dens}}, \tag{2.38}$$

i.e. separate convergence for the hole and electron charge densities. Here, d denotes the number of spatial dimensions of the simulated region. Typical values for δ_{dens} are 10^{-3} for 3-dimensional, $10^3\,\mathrm{cm}^{-1}$ for 2-dimensional, and $10^5\,\mathrm{cm}^{-2}$ for 1-dimensional calculations. The meaning of δ_{dens} can be best described as the residual change in the occupation of the eigenfunctions.

Underrelaxation approach for the coupled Schrödinger, current, and Poisson equations

In solving the coupled Schrödinger, current, and Poisson equations, the goal is to determine a charge self-consistent solution for the quasi Fermi levels and the charge densities. Here, the basic problem is that the Schrödinger equation

$$H(\phi)\psi_n = E_n\psi_n \tag{2.39}$$

depends on the electrostatic potential ϕ, the current equation depends on the eigenfunctions ψ_n via the charge densities

$$\nabla \cdot \mu_{\mathrm{p}} (\mathbf{x})\, p \left[\psi_1, \psi_2, \ldots\right] (\mathbf{x})\, \nabla E_{\mathrm{F,p}} (\mathbf{x}) = -R (\mathbf{x}), \tag{2.40}$$

$$\nabla \cdot \mu_{\mathrm{n}} (\mathbf{x})\, n \left[\psi_1, \psi_2, \ldots\right] (\mathbf{x})\, \nabla E_{\mathrm{F,n}} (\mathbf{x}) = R (\mathbf{x}), \tag{2.41}$$

and the Poisson equation depends on the eigenfunctions ψ_n as well as on the quasi Fermi energies $E_{\mathrm{F,n}}$ and $E_{\mathrm{F,p}}$

$$\nabla \cdot \varepsilon (\mathbf{x})\, \nabla \phi = -\rho \left[E_{\mathrm{F,n}}, E_{\mathrm{F,p}}, \psi_1, \psi_2, \ldots\right]. \tag{2.42}$$

Whereas we can still apply the predictor-corrector-type method for the Schrödinger and the Poisson equations, no such approach exists including the current equation and we therefore have to use an underrelaxation approach for the quasi Fermi energies, as a simple iteration scheme does not converge due to strong oscillations in the charge densities from one iteration step to the next. The combined predictor-corrector and underrelaxation approach can be summarized as follows: For $k = 1, 2, \ldots$ and given ϕ_1, $E_{\mathrm{F,n}}^{(1)}$, and $E_{\mathrm{F,p}}^{(1)}$

1. Solve $H(\phi_k)\, \psi_n^{(k)} = E_n^{(k)} \psi_n^{(k)}$ for a set of N_{ev} eigenfunctions $\psi_1^{(k)}, \psi_2^{(k)}, \ldots, \psi_{N_{\mathrm{ev}}}^{(k)}$ and corresponding eigenenergies $E_1^{(k)}, E_2^{(k)}, \ldots, E_{N_{\mathrm{ev}}}^{(k)}$.

2. Solve

$$\nabla \cdot \mu_{\mathrm{p}} (\mathbf{x})\, p \left[\psi_1^{(k)}, \psi_2^{(k)}, \ldots, \psi_{N_{\mathrm{ev}}}^{(k)}\right] (\mathbf{x})\, \nabla \tilde{E}_{\mathrm{F,p}}^{(k+1)} (\mathbf{x}) = -R (\mathbf{x}), \tag{2.43}$$

$$\nabla \cdot \mu_{\mathrm{n}} (\mathbf{x})\, n \left[\psi_1^{(k)}, \psi_2^{(k)}, \ldots, \psi_{N_{\mathrm{ev}}}^{(k)}\right] (\mathbf{x})\, \nabla \tilde{E}_{\mathrm{F,n}}^{(k+1)} (\mathbf{x}) = R (\mathbf{x}) \tag{2.44}$$

   for the quasi Fermi energies $\tilde{E}_{\mathrm{F,p}}^{(k+1)}$ and $\tilde{E}_{\mathrm{F,n}}^{(k+1)}$.

3. Determine new quasi Fermi energies $E_{\mathrm{F,p}}^{(k+1)}$ and $E_{\mathrm{F,p}}^{(k+1)}$ according to

$$E_{\mathrm{F,p}}^{(k+1)} = \alpha \tilde{E}_{\mathrm{F,p}}^{(k+1)} + (1 - \alpha)\, E_{\mathrm{F,p}}^{(k)}, \tag{2.45}$$

$$E_{\mathrm{F,n}}^{(k+1)} = \alpha \tilde{E}_{\mathrm{F,n}}^{(k+1)} + (1 - \alpha)\, E_{\mathrm{F,n}}^{(k)}, \tag{2.46}$$

   where α denotes the underrelaxation parameter.

4. Solve the non-linear Poisson equation $\nabla \cdot \varepsilon (\mathbf{x})\, \nabla \phi_{k+1} = -\tilde{\rho} \left[E_{\mathrm{F,p}}^{(k+1)}, E_{\mathrm{F,n}}^{(k+1)}\right] (\phi_{k+1})$, where $\tilde{\rho}$ is obtained from ρ by replacing n by $\tilde{n}$ and p by $\tilde{p}$ with $\phi_{\mathrm{prev}} = \phi_k$.

As exit condition for the iteration we require that

$$\int \mathrm{d}^d \mathbf{x}\, \left|\tilde{p}\left[\phi_{k+1}\right](x) - \tilde{p}\left[\phi_k\right](x)\right| < \delta_{\mathrm{dens}}, \tag{2.47}$$

$$\int \mathrm{d}^d \mathbf{x}\, \left|\tilde{n}\left[\phi_{k+1}\right](x) - \tilde{n}\left[\phi_k\right](x)\right| < \delta_{\mathrm{dens}}, \tag{2.48}$$

and

$$\left\| E_{\mathrm{F,p}}^{(k+1)} - E_{\mathrm{F,p}}^{(k)} \right\|_\infty < \delta_{\mathrm{Fermi}}, \tag{2.49}$$

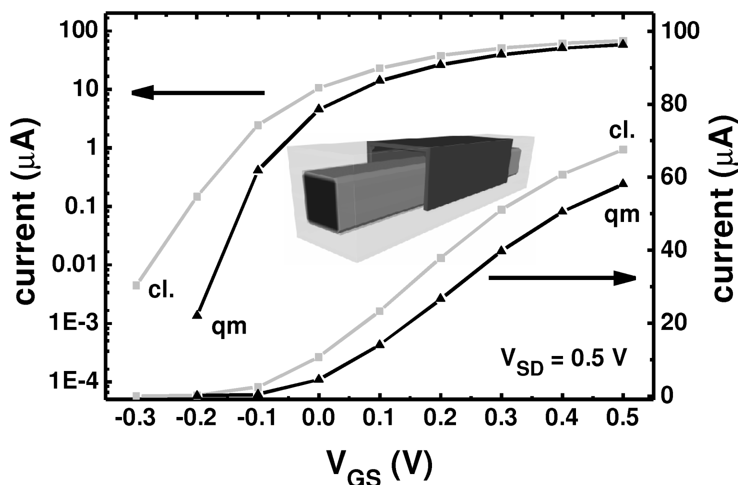

Figure 2.2: Current-voltage characteristics of a triple-gate Si MOSFET with a 5 nm × 5 nm channel and 25 nm gate length (inset). The graph shows the I-V curves for classical (gray line) and quantum mechanical (black line) calculations. In the latter case the current is smaller.

$$\left\| E_{\mathrm{F,p}}^{(k+1)} - E_{\mathrm{F,p}}^{(k)} \right\|_\infty < \delta_{\mathrm{Fermi}}. \tag{2.50}$$

For the underrelaxation parameter, values of about $\alpha = 0.2$ have turned out to be reasonable in a wide range of applications. Typical values for δ_{Fermi} should be around 10^{-5} eV or smaller.

Figure 2.2 shows the I-V characteristics of a triple-gate silicon MOSFET that has been calculated by the author using the described method. The device has a 5 nm × 5 nm channel and 25 nm gate length. The number of grid points amounts to 5×10^4. The calculation of each voltage step of the I-V curve took approximately two hours on a state-of-the-art PC and required about 200 iteration steps.

Subspace iterations

The predictor-corrector approach as well as the underrelaxation approach can be combined with a newly developed idea that helps to accelerate the calculation by saving time in the solution of the quantum mechanical eigenvalue problems [35]. We have denoted this method as subspace iterations. The idea behind this method is that the Schrödinger equation

$$H(\phi)\,\psi_n = E_n \psi_n \tag{2.51}$$

is solved exactly only in a limited number of iteration steps, whereas in the remaining steps, the Hamiltonian is diagonalized in a subspace of the full Hilbert space. This subspace is the span of only a few of the eigenvectors that are lowest in energy. In detail, this is done as follows: For given ϕ_0

1. Solve $H(\phi_0)\,\psi_n^{(0)} = E_n^{(0)} \psi_n^{(0)}$ for a set of N_{ev} eigenfunctions $\psi_1^{(0)}, \psi_2^{(0)}, \ldots, \psi_{N_{\mathrm{ev}}}^{(0)}$ and corresponding eigenenergies $E_1^{(0)}, E_2^{(0)}, \ldots, E_{N_{\mathrm{ev}}}^{(0)}$.

2. Determine ϕ_1.

Then for $k = 1, 2, \ldots$

1. If the $\psi_n^{(k-1)}$ are exact eigenvectors or if

$$\max\left(\int \mathrm{d}^d\mathbf{x}\,\left|\tilde{p}\left[\phi_{k+1}\right](x)-\tilde{p}\left[\phi_k\right](x)\right|,\int \mathrm{d}^d\mathbf{x}\,\left|\tilde{n}\left[\phi_{k+1}\right](x)-\tilde{n}\left[\phi_k\right](x)\right|\right)>f_{\mathrm{subspace}}\cdot\delta_{\mathrm{dens}}, \tag{2.52}$$

   proceed with step 4.

2. Solve $H(\phi_k)\,\psi_n^{(k)}=E_n^{(k)}\psi_n^{(k)}$ for a set of N_{ev} exact eigenfunctions $\psi_1^{(k)},\psi_2^{(k)},\ldots,\psi_{N_{\mathrm{ev}}}^{(k)}$ and corresponding eigenenergies $E_1^{(k)},E_2^{(k)},\ldots,E_{N_{\mathrm{ev}}}^{(k)}$.

3. Proceed with step 6.

4. Determine $\left(H^k\right)_{nm}=\left\langle\psi_n^{(k-1)}\right|H(\phi_k)\left|\psi_m^{(k-1)}\right\rangle$, $n,m=1,\ldots,N_{\mathrm{ev}}$.

5. Diagonalize $\left(H^{(k)}\right)_{nm}$ and obtain N_{ev} approximate eigenfunctions $\psi_1^{(k)},\psi_2^{(k)},\ldots,\psi_{N_{\mathrm{ev}}}^{(k)}$ and corresponding eigenenergies $E_1^{(k)},E_2^{(k)},\ldots,E_{N_{\mathrm{ev}}}^{(k)}$.

6. Determine ϕ_{k+1}.

The decision whether the Schrödinger equation is solved exactly or within the subspace is met in step 1 of the second block. Each exact solution is followed by at least one approximate solution. Further approximate solutions are inserted if the residuals of the electron and hole charge densities are greater than the convergence criterion δ_{dens} by at least the subspace residual factor f_{subspace}. Typical values for f_{subspace} are of the order of 10^6. Note that the exit condition is modified so that the iteration is always terminated following an exact solution of the Schrödinger equation.

2.8 Summary

In this chapter, we have introduced the techniques for the numerical solution of the semiconductor equations. The discretization of those partial differential equations results either in linear systems of equations or large matrix eigenvalue problems that can be solved by iterative algorithms. We have presented the set of algorithms that is employed by next**nano**++ and we have shown its concrete application for the solution of the semiconductor equations. The solution of the semiconductor equations aims at the determination of a charge self-consistent solution for the electrostatic potential and the quasi Fermi energies. This also requires the solution of the non-linear Poisson equation and the coupled systems of Schrödinger and Poisson as well as Schrödinger, current, and Poisson equations. For the coupled systems, two different approaches have been chosen: a predictor-corrector method for the coupled Schrödinger and Poisson equations, and an underrelaxation method for the coupled Schrödinger, current, and Poisson equations. Both approaches have been combined with the newly developed subspace iteration method.

Chapter 3

Program structure

3.1 Libraries

In the following sections we want to give an overview on the program structure of next**nano**++. Next**nano**++ is written mainly in C++ with only the core numerical routines still being implemented using Fortran. By employing a modern, object oriented language such as C++, we can benefit a lot in terms of code reuse by means of inheritance and generic programming as will be discussed exemplarily later on.

Next**nano**++ is built in a modular way from multiple libraries. These libraries are divided into the core numerical libraries and the application specific libraries. Figure 3.1 shows the core numerical libraries and their dependencies. The libraries BLAS (Basic Linear Algebra Subprograms [36]) and LAPACK (Linear Algebra PACKage [37]) are provided from external. We use the BLAS library provided by the Intel Math Kernel Library (MKL). Further, we use LAPACK 3.0 with minor changes as compared to Reference [37]. The libraries BLAS as well as LAPACK are written in Fortran. On top of BLAS and LAPACK we have placed the mixed language library `f95library`. This as well as `mathlib`, `arraylib` and `sparselib` depend on the library `systemlib` that provides a unified interface for the control of debug and range check options. The library `f95library` implements the following classes of routines: sparse matrix routines, iterative solvers for sparse matrices, Fermi-Dirac integrals, and special functions. Further, it provides the necessary Fortran bindings as well as templated C++ wrappers for accessing the library functions from within Fortran as well as C++ programs. Templated C++ wrappers are also implemented for several LAPACK and BLAS subroutines. The library `mathlib` is written completely in C++. It depends on `f95library` and implements classes for complex arithmetic (to be used as an alternative to the C++ standard library implementation for complex numbers), templated vector arithmetic, and templated second rank tensor arithmetic. The library `arraylib` is based on `mathlib` and `f95library` and is also written completely in C++. It implements a templated class for multi-dimensional arrays and relevant operations (serialization, arithmetics, diagonalization, etc.). The library `sparselib` depends on `arraylib` and `f95library`. `Sparselib` provides a templated class for sparse matrices and corresponding operations (arithmetics, linear solvers, eigenvalue solvers, etc.). `Sparselib` is also written completely in C++.

Figure 3.2 shows the application specific libraries and their dependencies. The libraries `parser` and `fileIOlib` mainly deal with input and output operations and provide classes to parse, validate, and process input files and to format the output data for use with Origin and AVS/Express. The library `discretization` provides a very general class for the discretiza-

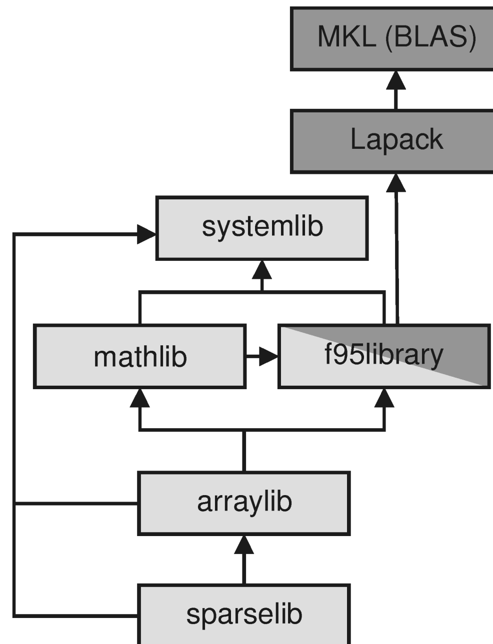

Figure 3.1: The core numerical libraries used by next**nano**++. The programming language is indicated in light gray (C++) and dark gray (Fortran). `f95library` is a mixed language library. The libraries BLAS and LAPACK are provided from extern. The arrows show the dependencies between the libraries.

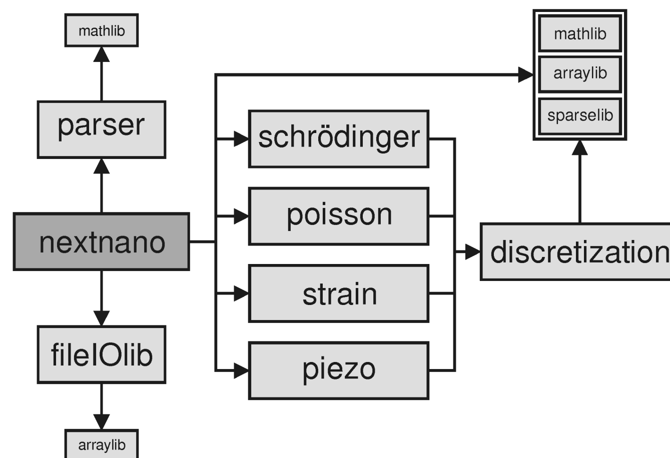

Figure 3.2: The project next**nano**++ and the application specific libraries. Next**nano**++ and all application specific libraries are completely written in C++ using object oriented programming techniques such as inheritance and generic programming. The arrows indicate the dependencies between next**nano**++, the application specific libraries as well as the core numeric libraries `mathlib`, `arraylib`, and `sparselib`.

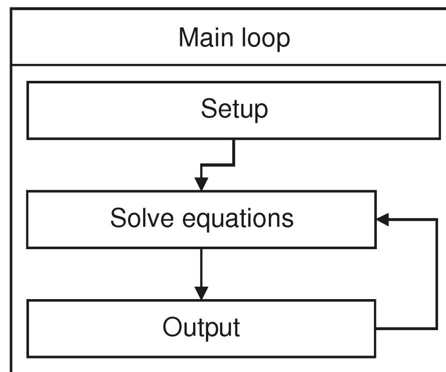

Figure 3.3: Flow chart of the main loop. The main loop iterates the bias voltages that are applied to the contacts.

tion of up to second order partial differential equations on one-, two-, and three-dimensional non-uniform tensor product grids based on the box integration finite differences scheme. This discretization library is used in the libraries `schrödinger`, `poisson`, and `strain` for the discretization of the Schrödinger equation, the Poisson equation, and the strain equation, respectively. In the library `piezo`, the discretization class is employed for the discretization of the divergence operator.

3.2 Program flow

In this section, we address the program flow of next**nano**++. For clarity, in Fig. 3.3 we show a flow chart of only the main loop and present the flow chart for the charge self-consistent solution of the Schrödinger, Poisson, and current equations separately in Figure 3.4. The main loop iterates the voltages that are applied to the contacts. Data that does not change while iterating the voltages is set up and calculated prior to the main loop in the "Setup" step. This includes (in the given order):

- Parsing, validating, and processing of the input file.
- Parsing, validating, and processing of the database file.
- Grid generation.
- Geometry processing (determination of materials, alloy compositions, doping profiles, and fixed charges).
- Calculation of the strain fields.
- Calculation of polarization charges.
- Determination of the classical bandstructure based on the model solid by van de Walle.
- Determination of the intrinsic density.
- Discretization of the Poisson equation.

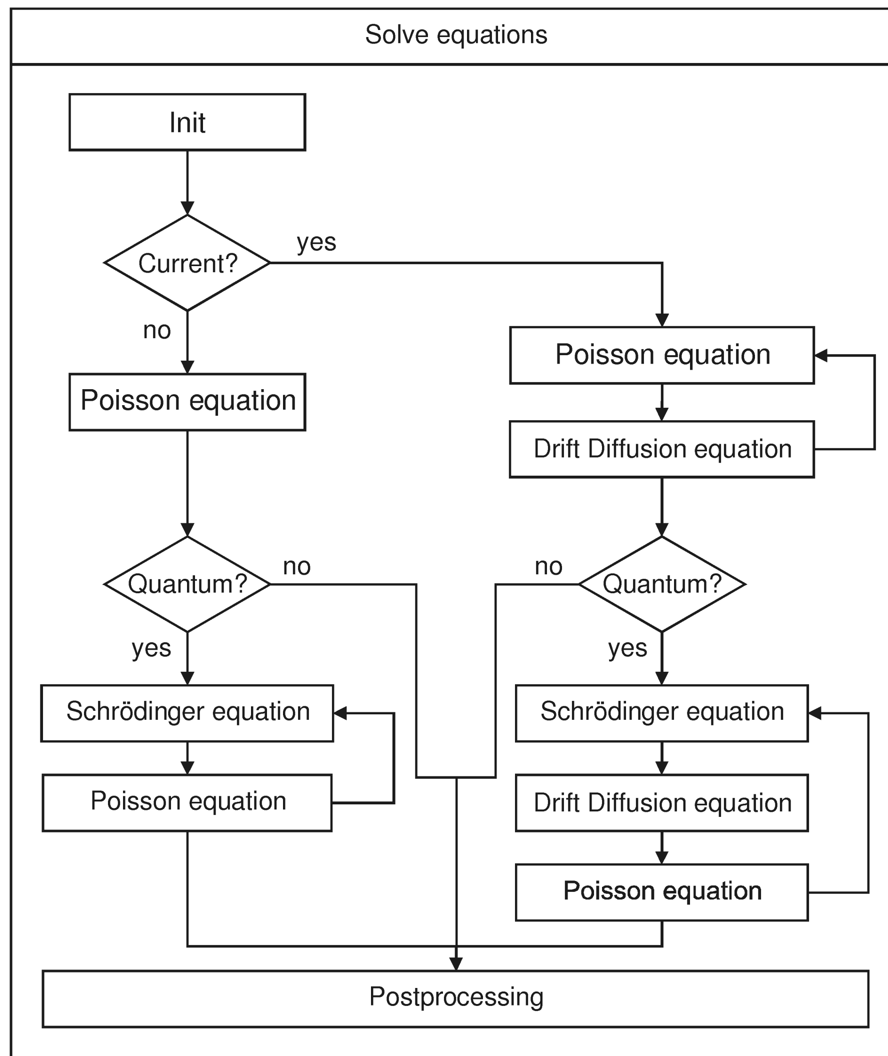

Figure 3.4: Flow chart for the charge self-consistent solution of the Schrödinger, Poisson, and current equations.

- Output of materials, alloy compositions, doping profiles, fixed charges, polarization charges, the intrinsic density, and the strain fields.

Following the "Setup" step, the charge self-consistent solution of the Schrödinger, Poisson, and current equations is calculated and the output is generated. This is repeated for the various voltages specified in the input file. The output includes the electrostatic potential, charge densities, quasi Fermi levels, current densities, wave functions, energy spectra, band structure in real space and reciprocal space, and transmission probabilities.

In Fig. 3.4 the flow chart of the charge self-consistent solution of the Schrödinger, Poisson, and current equations is shown. Data that does not change while iterating the solution of the mentioned equations is set up and calculated in the "Init" step. This includes:

- Initialization of the quasi Fermi levels employing an empirical approach.
- Update of the discretization of the Poisson equation to take into account the voltages applied to the contacts.

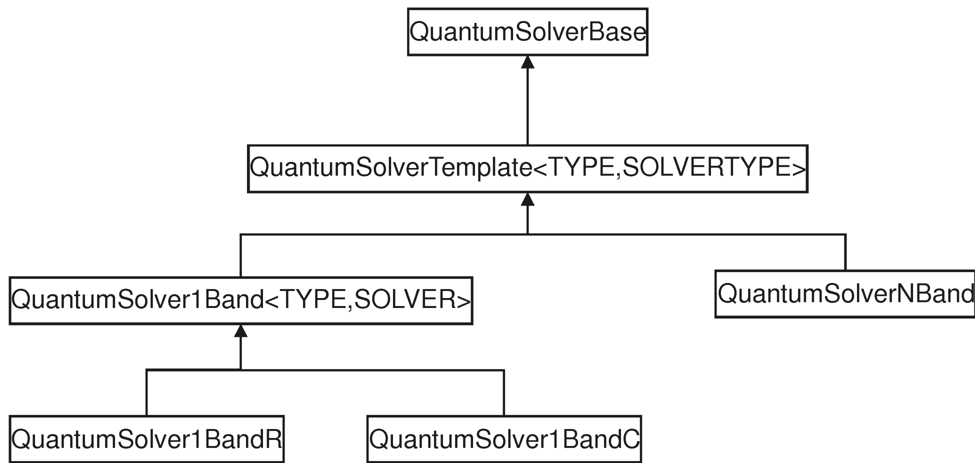

Figure 3.5: Class hierarchy of the set of classes that implement the unified interface for the solution of the single band Schrödinger equation and the multi-band $k \cdot p$ equations.

- Initialization of the electrostatic potential (either locally charge neutral or equal to the quasi Fermi level).

Following the "Init" step, the non-linear Poisson equation coupled to the current equation is solved using the Thomas-Fermi approximation for the charge carrier density. For equilibrium situations, the solution of the current equation is omitted. For the solution of the coupled system of the non-linear Poisson equation and the current equation, we employ the underrelaxation algorithm presented in Section 2.7. For the solution of the non-linear Poisson equation, we use the Newton method with inexact line search as presented in Section 2.6.

The solution for the electrostatic potential and the quasi Fermi levels is the initial guess for the solution including the quantum mechanical charge carrier densities. As described in closer detail in Sec. 2.7 we repeatedly solve the Schrödinger, the current, and the Poisson equation. Again, for equilibrium situations, the solution of the current equation is omitted. For the solution of the coupled system of the Schrödinger and the Poisson equation we use the predictor-corrector algorithm presented in Section 2.7. The iterations are repeated either until separate convergence of the electron and hole densities and the electron and hole quasi Fermi levels or until exceeding the maximum number of iterations. Following the self-consistent calculation of the Poisson, Schrödinger, and current equation, additional calculations are performed in the "Postprocessing" step. This includes calculations that use the electrostatic potential, the quasi Fermi levels, and the charge density as input parameters such as the determination of transmission probabilities using the contact block reduction method and the calculation of the current densities.

3.3 Class hierarchy

In this section we want to show exemplarily the use of inheritance within next**nano**++ for the classes that implement the quantum mechanical models for the electronic structure calculations. Further, we discuss in more detail the design of the class `simulator` that manages the data, the data flow, and the program flow.

As an example for the use of inheritance within next**nano**++ we choose the set of classes that implement the solution of the Schrödinger equation. Here, the details depend strongly on the specific model that has to be solved, whereas the basic properties are the same for all models. This suggests to implement the properties common to all models in a single base class and to derive more specialized sub classes that implement the model specific properties. Figure 3.5 shows the class hierarchy that results from this approach. The common properties that are provided by the base class `QuantumSolverBase` include interfaces to methods for:

- discretization,
- determination of eigenvectors and eigenvalues,
- calculation of the charge density and its derivative with respect to the electrostatic potential.

Depending on the presence of magnetic fields and the number of bands (either single-band or multi-band), eigenvectors are either real valued or complex valued. This is taken into account by means of the derived class `QuantumSolverTemplate<TYPE,SOLVERTYPE>` that provides a unified interface independent of the type of the eigenvectors. The calculation of the charge carrier density has to be implemented differently depending on the number of bands, the method used for determining the eigenvectors and eigenvalues, and the method used for discretization. Therefore, separate sub classes are derived from class `QuantumSolverTemplate<TYPE,SOLVERTYPE>` that either deal with the single-band Schrödinger equation or the multi-band $k \cdot p$ equation. For the single-band case, the class `Quantum1BandSolver<TYPE,SOLVER>` is the basis for the types `QuantumSolver1BandR` and `QuantumSolver1BandC` that differ in the type of the eigenvectors (either real or complex). The class `QuantumNBandSolver` is the generic class for all multi-band $k \cdot p$ models. In this case, the eigenvectors are always complex valued.

Now, we want to address the design of the main class of next**nano**++ termed `Simulator`. The class `Simulator` manages all simulation relevant data, the data flow, and the program flow. The class `Simulator` consists of objects that are instances of the following 14 classes: `Flow`, `Input`, `TensorGrid`, `Database`, `Structure`, `Cbr`, `Poisson`, `Contacts`, `Global`, `Currents`, `Bandstructure`, `Densities`, `Quantum`, and `Strain`. This "has-a" relation is depicted in Fig. 3.6 by the solid black arrows emanating from class `Simulator`. Each of these 14 classes is responsible for managing a specific facet of the simulation. The main issue concerning the data flow is that most of these 14 classes have dependencies on data from other classes. For example all classes except for class `Database`, `TensorGrid`, `Input`, and `Flow` depend on data provided by class `Global`. In Fig. 3.6 this dependencies are indicated by the dash-dotted arrows that point towards class `Global`. All objects that depend on class `Global` obtain a reference to the simulator's instance of class `Global` on construction. This minimizes the effort in passing data as arguments, while at the same time avoiding the disadvantageous use of global data. In the same way we have also treated the data dependencies between the other classes. These are depicted by the dotted arrows in Figure 3.6. For example, the class `Poisson` depends on class `Contacts`, class `Currents`, and class `Densities`. This is because in the course of solving the non-linear Poisson equation, the class `Poisson` needs to know about the voltages applied to the individual contacts (provided by class `Contacts`), the quasi Fermi levels for electrons and holes (provided by class `Currents`), and various densities (provided by class `Densities`). The dashed arrows indicate dependencies between members of a certain class and one of the other classes. For example class `Densities` possesses a reference to the list of `QuantumRegions` contained

in an object of type `Quantum` (dashed arrow). On the other hand, class `Quantum` contains a reference to an object of type `Densities` (dotted arrow). Also note that in some cases, data is also shared directly between the classes. The members `Acceptors`, `Donors`, `FixedCharges`, and `Polarization` in the simulator's instance of class `Densities` are only pointers to the actual data contained in the simulator's instances of class `Structure` and class `Strain`, respectively.

3.4 Input processing

In this section, we give an overview on the input processing implemented in next**nano**++. Input processing as implemented in next**nano**++ is a three step process. We use a parser based on the Flex scanner generator and the Bison parser generator to first parse a so called validator file and a corresponding input file [38]. The validator file as well as the input file obey the same grammar with some validator specific features disabled for input files. The grammar allows describing assignment statements as well as a hierarchical data structure in terms of nested groups. When parsed, this hierarchical data structure is translated into a tree like symbol table with each node either representing a group or an assignment. Nodes that represent groups may have sub nodes; nodes that represent assignments may not have sub nodes. An example is shown in Figure 3.7. The second step is that the validator symbol table is used to validate the input file symbol table. By this, a certain logical structure is enforced for the input file: groups and assignments are restricted to certain names and types, the number of identically named groups can be fixed, and conditions such as "group A requires the existence of group B" can be defined. This approach of validating the logical structure of the input by means of a validator file is similar to that of XML schemes.

Following the successful validation, the validator symbol table is discarded. The input file symbol table is now guaranteed to follow the logical structure defined by the validator file. All groups and assignments declared in the input file are stored in the tree like structure of the symbol table. The third step in processing the input is the automated instantiation and initialization of hierarchical linked objects right from the symbol table. To this end, a sub class of class `Group` (provided by the `parser` library) is implemented for each group that according to the validator can be present in the input file. This is illustrated in Fig. 3.8 for the group "global" shown in Figure 3.7. For each possible assignment within this group a data member of the respective type has to be provided. Similarly, for each nested group, an instance of the class corresponding to this nested group has to be provided. Further, the data members and nested group's objects have to be associated with the names that are used in the input file for the corresponding assignments and nested groups. This is done by means of the statements "define" (for data members) and "link" (for sub groups) in the default constructor of the class. The data members are now initialized automatically from the symbol table. In the same manner, objects of classes derived from class `Group` are automatically instantiated (if necessary) and initialized. Thus the data from the symbol table is transferred to a data structure of hierarchically linked, individually typed objects and can therefore be conveniently accessed from within the program in the same way as any other kind of data. In this, the procedure follows the ideas of reflective programming.

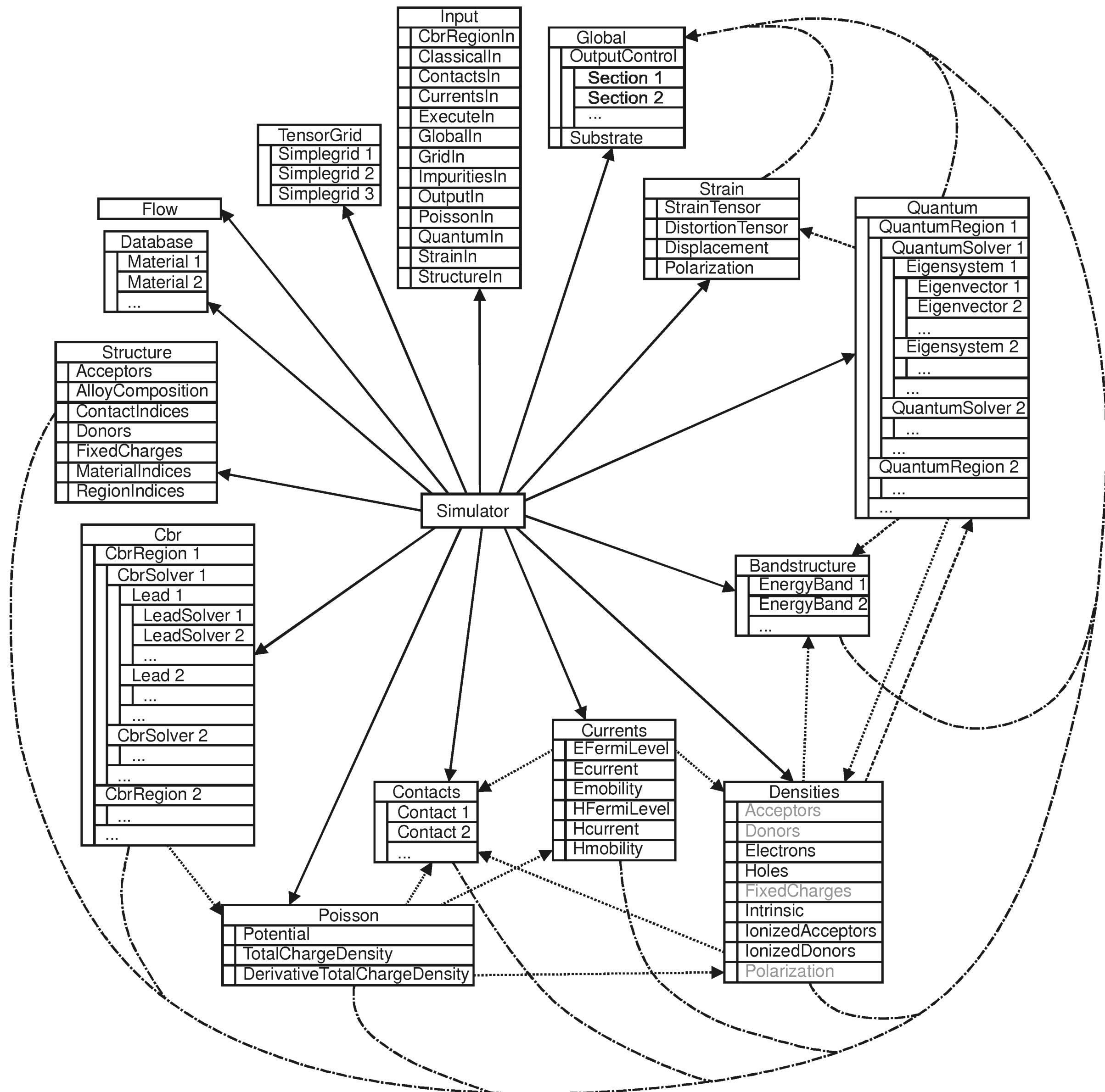

Figure 3.6: Design of the main class of next**nano**++ termed `Simulator`. The class consists of the following 14 objects: `Flow`, `Input`, `TensorGrid`, `Database`, `Structure`, `Cbr`, `Poisson`, `Contacts`, `Global`, `Currents`, `Bandstructure`, `Densities`, `Quantum`, and `Strain`. The relations between the objects and class `Simulator` are indicated by arrows. The solid black arrows indicate a "has-a" relation. The dashed, dotted, and dash-dotted arrows indicate references. Also shown are the most relevant data members of the individual objects.

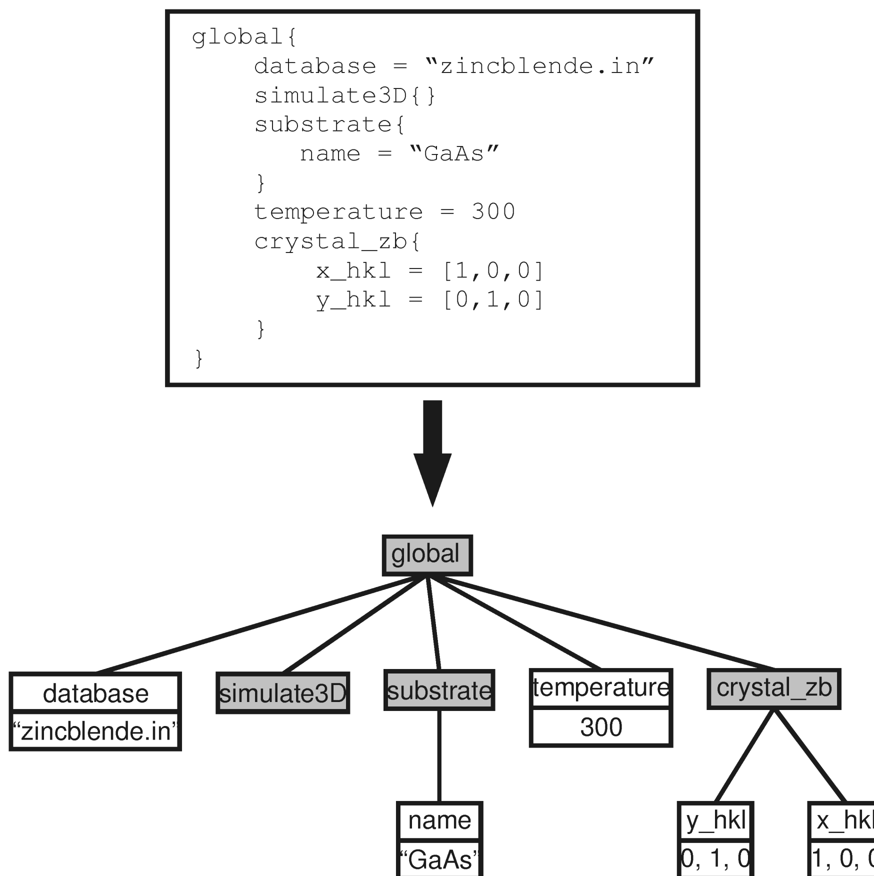

Figure 3.7: Example for the hierarchical structure of the next**nano**++ input files and the tree like symbol table that results from parsing the input file. Nodes shaded in gray represent groups and may have sub nodes. The other nodes represent assignments.

```
class Global : public Group {

    string database;
    Group simulate3D;
    ...

    Global() {
        define <string> (”database”, database);
        link <Group> (”simulate3D”, simulate3D);
        ...
    }
};
```

Figure 3.8: Example of the class `Global` derived from class `Group` for processing the input related to the group "global" (see Figure 3.7). For each assignment in group "global" a data member is provided (here only for "database") and the data member is associated with the name of the assignment by the "define" statement in the default constructor. Similarly, for each nested group of group "global" an instance of class `Group` (or a derived class) is provided (here only for "simulate3D"). The instance is associated to the name of the nested group by the "link" statement in the default constructor.

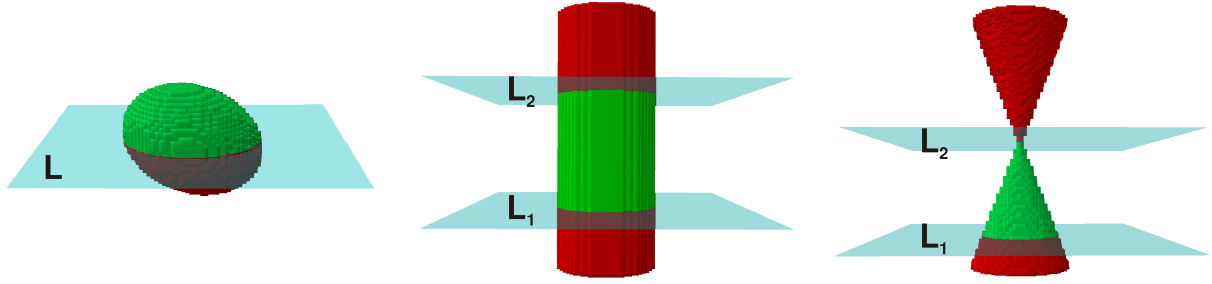

Figure 3.9: Constructive solid geometry shapes (green). From left to right: semiellipsoid, cylinder, and cone. The shapes are composed of the intersection of quadrics (green and red) and one or two planes (blue).

3.5 Geometry processing

The geometry processing as implemented in next**nano**++ is another example, where large benefit is obtained from object oriented programming techniques. In next**nano**++, the device geometry is defined by unions and differences of relatively simple geometric objects. In three spatial dimensions, these include cuboids, obelisks, cylinders, semiellipsoids, and cones. This approach for defining complex shapes from basic shapes (also called primitives) is called constructive solid geometry (CSG).

Here, the advantage is that the surface of the primitives can be easily described by relatively simple mathematical formulas. Therefore, this allows one to easily check if a specific grid point lies within or outside the primitive. The primitives used in next**nano**++ are based on even more simple geometric objects, namely planes and quadratic surfaces (also called quadrics). Planes are defined by an emission point $\mathbf{a}$ and a normal vector $\mathbf{n}$

$$\text{Plane} = \left\{ \mathbf{x} \in \mathbb{R}^3 : (\mathbf{x} - \mathbf{a}) \cdot \mathbf{n} = 0 \right\}, \tag{3.1}$$

quadrics are defined by a mid point $\mathbf{m}$, a symmetric second rank tensor T, and a scalar s

$$\text{Quadric} = \left\{ \mathbf{x} \in \mathbb{R}^3 : (\mathbf{x} - \mathbf{m}) \cdot T\,(\mathbf{x} - \mathbf{m}) = s \right\}. \tag{3.2}$$

Planes divide space into two half spaces that are given by

$$\text{Lower half space} = \left\{ \mathbf{x} \in \mathbb{R}^3 : (\mathbf{x} - \mathbf{a}) \cdot \mathbf{n} < 0 \right\}, \tag{3.3}$$

$$\text{Upper half space} = \left\{ \mathbf{x} \in \mathbb{R}^3 : (\mathbf{x} - \mathbf{a}) \cdot \mathbf{n} > 0 \right\}. \tag{3.4}$$

Quadrics divide space into inside and outside

$$\text{Inside} = \left\{ \mathbf{x} \in \mathbb{R}^3 : (\mathbf{x} - \mathbf{m}) \cdot T\,(\mathbf{x} - \mathbf{m}) < s \right\}, \tag{3.5}$$

$$\text{Outside} = \left\{ \mathbf{x} \in \mathbb{R}^3 : (\mathbf{x} - \mathbf{m}) \cdot T\,(\mathbf{x} - \mathbf{m}) > s \right\}. \tag{3.6}$$

Obelisks are defined by the section of six lower half spaces $L_1, \ldots, L_6$

$$\text{Obelisk} = \bigcap_{i=1}^{6} L_i. \tag{3.7}$$

As shown in Fig. 3.9, semiellipsoids are defined as the section of the inside of an ellipsoidal quadric I with $s = 0$ and a lower half space L

$$\text{Semiellipsoid} = I \cap L, \tag{3.8}$$

so that one half of the ellipsoid is cut off. Cylinders and cones are defined as the section of the inside of a conic quadric I with $s = 1$ and two lower half spaces L_1, L_2

$$\text{Cylinder/Cone} = I \cap L_1 \cap L_2. \tag{3.9}$$

The surface described by a conic quadric extends to infinity in both directions along the symmetry axis. The two half spaces are placed so that only a finite section of the conic quadric remains that either represents a cone, a truncated cone, or a cylinder. This is also shown in Figure 3.9. Note that a cylinder is nothing else than a cone with the cone's tip at infinity.

The cuboid is a special case of the obelisk with its faces parallel to the coordinate planes. The inside of the cuboid can therefore be easily defined in terms of six conditions as follows

$$\text{Cuboid} = \left\{ \mathbf{x} \in \mathbb{R}^3 : a_i < x_i < b_i, \quad i = 1, 2, 3 \right\}. \tag{3.10}$$

The determination of the symmetric second rank tensors T for the ellipsoidal and conic quadratics from the parameters provided in the input file is discussed in detail in Appendix B.

For two spatial dimensions the primitives are obtained from the primitives in three spatial dimensions by sections in the $x_3 = 0$ coordinate plane. They include rectangles, trapezoids, and semiellipses. In one spatial dimension, the only primitive is given by the line segment.

3.6 Summary

For next**nano**++ we have used a modular structure in terms of core numeric libraries and application specific libraries. The core numeric libraries are written in C++ and Fortran, where Fortran itself is reduced to a minimum. The application specific libraries and next**nano**++ itself are completely written in C++. This allows the heavy use of modern object oriented programming techniques such as inheritance and generic programming and therefore a high degree of code reuse all over the project. An innovative approach for the input processing based on the ideas of reflection and XML schemes has been developed. The CSG based geometry processing allows for an easy and convenient definition of even complex device geometries.

Chapter 4

Calculation of the charge carrier density by Brillouin zone integration

4.1 Introduction

The calculation of the charge density $n(x)$ along the growth axis of a heterostructure with two-dimensional translational invariance requires the evaluation of the following integral

$$n(x) = \sum_j \frac{1}{(2\pi)^2} \int_{\Omega_{\mathrm{BZ}}} \mathrm{d}^2\mathbf{k}_\parallel \, p_j\left(x, \mathbf{k}_\parallel\right) f\left(E_j\left(\mathbf{k}_\parallel\right)\right) \tag{4.1}$$

over the two-dimensional Brillouin zone Ω_{BZ} in the transversal reciprocal space. The summation runs over the energy subbands, $p_j\left(x, \mathbf{k}_\parallel\right)$ is the probability density of the eigenstates of the jth subband, $E_j\left(\mathbf{k}_\parallel\right)$ is the dispersion relation of the jth band, and $f(E)$ is the Fermi function. Instead of directly evaluating the integral in Eq. (4.1), it is possible to first calculate the local density of states (LDOS) according to

$$g(x, E) = \sum_j \int_{\Omega_{\mathrm{BZ}}} \mathrm{d}^2\mathbf{k}_\parallel \, p_j\left(x, \mathbf{k}_\parallel\right) \delta\left(E - E_j\left(\mathbf{k}_\parallel\right)\right) \tag{4.2}$$

and to determine the charge density $n(x)$ from the LDOS as follows:

$$n(x) = \frac{1}{(2\pi)^2} \int_{-\infty}^{\infty} \mathrm{d}E \, g(x, E) \, f(E) \, . \tag{4.3}$$

It has been recognized already for decades that integrals such as the one in Eq. (4.2) are exemplary for the calculation of a large number of properties of solids [39]. Correspondingly, a multitude of very sophisticated methods has been developed for the numerical evaluation of Eq. (4.2) [40–55]. These methods try to reduce the number of Brillouin zone points $\mathbf{k}_\parallel$ where $p_j\left(x, \mathbf{k}_\parallel\right)$ and $E_j\left(\mathbf{k}_\parallel\right)$ are evaluated exactly by using so called special k points and by interpolating $p_j\left(x, \mathbf{k}_\parallel\right)$ and $E_j\left(\mathbf{k}_\parallel\right)$ up to quadratic order.

However, for the calculation of the charge carrier density, no advantage is gained by the artificial separation of Eq. (4.1) into Eq. (4.2) and Equation (4.3). On the contrary, the two-dimensional integral for the LDOS $g(x, E)$ has to be performed multiple times in the evaluation of Eq. (4.3), whereas a direct evaluation of Eq. (4.1) requires only one evaluation of a two-dimensional integral. In particular this is important in our case, where Eq. (4.1) has to be

evaluated multiple times in the course of solving the Schrödinger and Poisson equations charge self-consistently. We have therefore developed a novel method that calculates the charge density by a direct integration over the Brillouin zone according to Equation (4.1). We also use interpolation schemes for $p_j\left(x, \mathbf{k}_{\|}\right)$ and $E_j\left(\mathbf{k}_{\|}\right)$ to reduce the number of Brillouin zone points $\mathbf{k}_{\|}$ where $p_j\left(x, \mathbf{k}_{\|}\right)$ and $E_j\left(\mathbf{k}_{\|}\right)$ are evaluated. For $p_j\left(x, \mathbf{k}_{\|}\right)$ a linear interpolation proves to be sufficient, whereas $E_j\left(\mathbf{k}_{\|}\right)$ requires higher order interpolation.

In many situations it is possible to exploit the symmetry of the crystal when integrating over the two-dimensional Brillouin zone. For a zinc-blende crystal with growth direction (001), the irreducible wedge of the two-dimensional Brillouin zone is only one eighth of the original two-dimensional Brillouin zone. However, in general the determination of the irreducible wedge is a complicated group theoretical problem that becomes even worse if strain has to be considered. Instead of solving this problem exactly, we have developed an approach that determines the irreducible wedge by evaluating the subband dispersion at specific points in the two-dimensional Brillouin zone.

The integral given in Eq. (4.1) can also be solved analytically for some special cases. These include the case of the single band effective mass model with parabolic dispersion if the mass is position independent. For anisotropic masses and non-diagonal mass tensors one has to calculate the correct density-of-states mass. This leads to the expressions given by Eq. (1.51) and Equation (1.52). If the mass is position dependent we allow for the following approximation: We still use Eq. (1.51) and Eq. (1.52), but we determine an average density-of-states mass $\bar{m}_{\mathrm{qdos}}$ separately for each eigenstate j. For this, we average the position dependent density-of-states mass $m_{\mathrm{qdos}}(x)$ according to the probability density $p_j(x)$ of the respective eigenstate

$$\bar{m}_{\mathrm{qdos},j} = \int \mathrm{d}x\, m_{\mathrm{qdos}}(x)\, p_j(x)\,. \tag{4.4}$$

In most other cases Eq. (4.1) can no longer be solved analytically and we employ our novel method for its numerical evaluation. In particular, this includes the case of all multi-band $k \cdot p$ models. While the calculation of the charge density according to Eq. (4.1) can be done straightforward for the 6-band $k \cdot p$ model, the following complication arises for the 8-band $k \cdot p$ model and for all other multi-band $k \cdot p$ models that deal simultaneously with conduction as well as valence bands. We will denote these kinds of multi-band $k \cdot p$ models as *mixed-band* $k \cdot p$ models. In occupying the eigenstates found within these mixed-band models, there is in general no possibility to decide if a specific eigenstate has to be occupied by an electron or by a hole. In fact it turns out that one and the same eigenstate can effectively contribute to the electron charge density and to the hole charge density.

We do not want to deny that also in the case of the mixed-band $k \cdot p$ models it is often possible to do a classification of the eigenstates into electrons and holes. In particular, this can be done if a positive band gap exists between the valence band highest in energy and the conduction band lowest in energy. While this is the case for all undoped type I heterostructures it does not include the interesting cases of type II heterostructure with broken gap. Similarly, this does also not include doping superlattices (so called nipi crystals [56, 57]) where depending on the doping concentration the band gap in position space may also become negative.

For these broken gap heterostructures we have developed a novel method for the determination of the charge carrier density for mixed-band $k \cdot p$ models that is not limited to positive spatial band gaps. By means of this method it is therefore possible to calculate the electronic structure of broken gap type II heterostructures and nipi crystals in the 8-band $k \cdot p$ model in a fully charge self-consistent way. Already at the beginning of the 1980th Altarelli has published

an article, where the electronic structure of an InAs/GaSb broken gap type II superlattice in a mixed-band $k \cdot p$ model including two valence and one conduction band is calculates in a charge self-consistent way [58]. And only recently a couple of articles have been published [59–61] where properties of AlSb/InAs/GaSb broken gap type II heterostructures are analyzed based on a charge self-consistent calculation of the electronic structure using an 8-band $k \cdot p$ approach. However, in all cases the authors do not give a description of their methods for the calculation of the charge density. Nonetheless, we believe that the method used in references [59–61] differs fundamentally from our approach, because as a matter of principle our approach always yields the total charge density (i.e. the sum of electron and hole charge density), whereas Fig. 1 in Ref. [59] indicates that the electron and the hole charge densities have been calculated separately.

This chapter is organized as follows: In Sec. 4.2 we introduce our novel method for the two-dimensional Brillouin zone integration. In this section we also analyze the convergence behavior of the method. Our novel approach for the calculation of the charge density for mixed-band $k \cdot p$ models is introduced in Section 4.3. In Sec. 4.4 we present the application of the method to doping superlattices. The chapter is concluded by a summary and outlook.

4.2 A novel method for the two-dimensional Brillouin zone integration

4.2.1 Integration method

To determine the charge density according to Eq. (4.1) we have to evaluate integrals of the form

$$I_j(x) = \int_{\Omega_{\mathrm{BZ}}} \mathrm{d}^2\mathbf{k}_\parallel \, p_j\left(x, \mathbf{k}_\parallel\right) f\left(E_j\left(\mathbf{k}_\parallel\right)\right) \tag{4.5}$$

where

$$p_j\left(x, \mathbf{k}_\parallel\right) = \sum_{\nu=1}^{N} \left|\psi_j^{(\nu)}\left(x, \mathbf{k}_\parallel\right)\right|^2 \tag{4.6}$$

denotes the probability density of eigenstate $\psi_j\left(x, \mathbf{k}_\parallel\right)$ of the jth subband with components $\psi_j^{(1)}, \ldots, \psi_j^{(N)}$ at point x along the growth axis and at point $\mathbf{k}_\parallel$ in the transversal reciprocal space. The dispersion relation of the subband is given by $E_j\left(\mathbf{k}_\parallel\right)$. The occupation is determined by the Fermi function

$$f(E) = \frac{1}{\exp\left[\left(E - E_{\mathrm{F}}\right)/k_B T\right] + 1} \tag{4.7}$$

where E_{F} denotes the quasi Fermi level and may also depend on x, i.e. $E_{\mathrm{F}} = E_{\mathrm{F}}(x)$.

A lot of work can be saved in the evaluation of Eq. (4.5) by exploiting the symmetry of the crystal in the plane orthogonal to the growth direction. Figure 4.1 shows all five different two-dimensional point lattices and the corresponding Wigner-Seitz cells. For the reciprocal lattice, the Wigner-Seitz cell is called Brillouin zone. The Brillouin zones for all five different two-dimensional reciprocal point lattices are shown in Fig. 4.2 together with the irreducible wedges that result from the full point symmetry. Instead of integrating over the full Brillouin zone Ω_{BZ}, it is sufficient to integrate over the irreducible wedge Ω_{IW} and to multiply the result by the ratio of the volume $|\Omega_{\mathrm{BZ}}|$ of the full Brillouin zone and the volume $|\Omega_{\mathrm{IW}}|$ of the irreducible

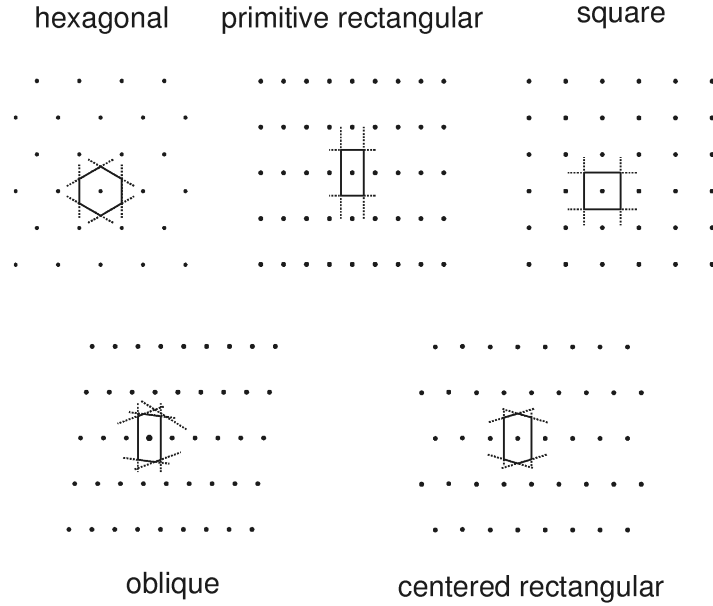

Figure 4.1: The five different two-dimensional point lattices and the corresponding Wigner-Seitz cells (solid lines). The Wigner-Seitz cells can be constructed from the perpendicular bisectors (dotted lines) of the connecting lines between the lattice point in the center of the Wigner-Seitz cell and all other lattice points.

wedge

$$I_j(x) = \frac{|\Omega_{\mathrm{BZ}}|}{|\Omega_{\mathrm{IW}}|} \int_{\Omega_{\mathrm{IW}}} d^2\mathbf{k}_\parallel p_j\left(x, \mathbf{k}_\parallel\right) f\left(E_j\left(\mathbf{k}_\parallel\right)\right). \tag{4.8}$$

In the numerical evaluation of the integral it is advantageous to perform the integration over a larger part Ω_{XIW} of the Brillouin zone than the irreducible wedge. We will call this part of the Brillouin zone the *extended* irreducible wedge (XIW). We want to emphasize that this does not require to evaluate $p_j\left(x, \mathbf{k}_\parallel\right)$ and $E_j\left(\mathbf{k}_\parallel\right)$ more often, because the values of $p_j\left(x, \mathbf{k}_\parallel\right)$ and $E_j\left(\mathbf{k}_\parallel\right)$ outside the irreducible wedge can be immediately obtained by the application of symmetry operations. Thus the integral that we will eventually evaluate is given by

$$I_j(x) = \frac{|\Omega_{\mathrm{BZ}}|}{|\Omega_{\mathrm{XIW}}|} \int_{\Omega_{\mathrm{XIW}}} \mathrm{d}^2\mathbf{k}_\parallel\, p_j\left(x, \mathbf{k}_\parallel\right) f\left(E_j\left(\mathbf{k}_\parallel\right)\right). \tag{4.9}$$

Extended irreducible wedges are only used in the case of the hexagonal and the square reciprocal point lattice. They are shown in Figure 4.3. In all three other cases, the extended irreducible wedges are identical to the irreducible wedges. However, to the boundary of the Brillouin zone the integration range is modified so that the integration can always be performed over a square or rectangle as shown in Figure 4.4. This is possible because the $k \cdot p$ dispersion does not show the translational symmetry of the lattice. For practical purposes it is even possible to restrict the integration to a small fraction of the Brillouin zone centered at $\mathbf{k}_\parallel = 0$. We specify this fraction by referring it to the radius of a sphere with the same volume as the three-dimensional bulk Brillouin zone.

The next few paragraphs will be more technical as we describe the details of our numerical

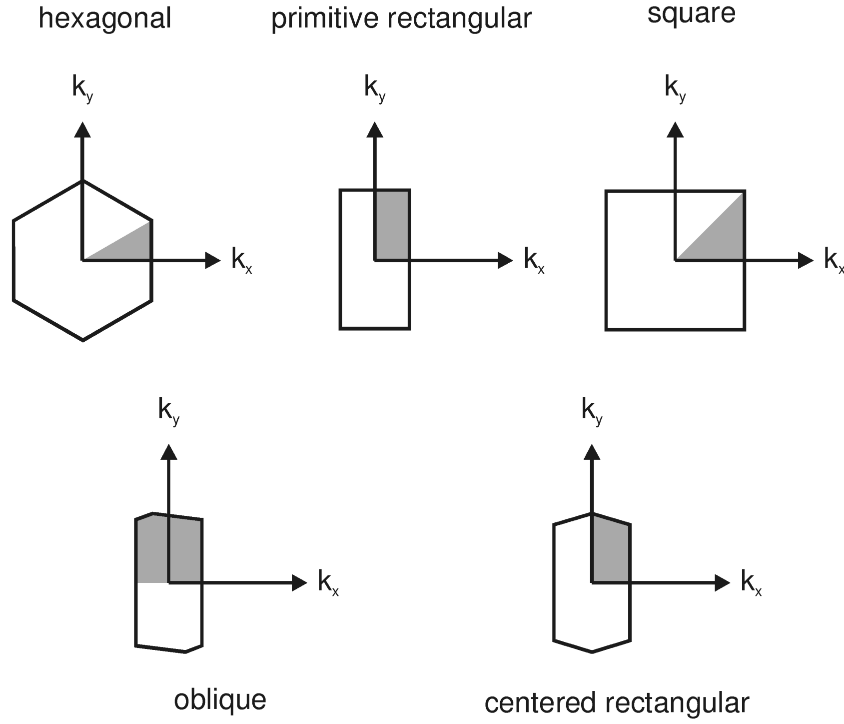

Figure 4.2: The Brillouin zones that correspond to the five different two-dimensional reciprocal point lattices. The shaded part of the Brillouin zone is the irreducible wedge resulting from the full point symmetry of the two-dimensional point lattice.

integration scheme. In essence we want to calculate the expression

$$\int_{\Omega} \mathrm{d}^2\mathbf{k}_{\parallel}\, g\left(\mathbf{k}_{\parallel}\right) \tag{4.10}$$

where Ω is either a square, a rectangle, or a rhombus with interior angles of $60°$ and $120°$. Now we subdivide Ω uniformly into finite quadrangular integration patches that are congruent to Ω. This results in a set of patches L_{p}, a set of $\mathbf{k}_{\parallel}$ points (vertices) L_{v}, and a set of quadruples L_{q} that enumerate the $\mathbf{k}_{\parallel}$ points at the four vertices of each patch. The area of each of the patches is denoted by Δ. Then, the integral (4.10) is approximated by

$$\int_{\Omega} \mathrm{d}^2\mathbf{k}_{\parallel}\, g\left(\mathbf{k}_{\parallel}\right) \approx \Delta \sum_{p \in L_{\mathrm{p}}} g_p, \tag{4.11}$$

where g_p is determined as follows: For the case of the rectangular integration ranges (square, rectangle), we apply the common two-dimensional trapezoidal rule

$$g_p = \frac{1}{4} \sum_{i=1}^{4} p_j\left(x, \mathbf{k}_{\parallel}^{(p,i)}\right) f\left(E_j\left(\mathbf{k}_{\parallel}^{(p,i)}\right)\right), \tag{4.12}$$

where $\mathbf{k}_{\parallel}^{(p,i)}$ denotes the $\mathbf{k}_{\parallel}$ point at the ith vertex of patch p. In the limit of $\Delta \to 0$, the approximation $\Delta \sum_{p \in L_{\mathrm{p}}} g_p$ converges towards $\int_{\Omega} d^2\mathbf{k}_{\parallel} g\left(\mathbf{k}_{\parallel}\right)$ quadratically in the spacing of neighboring $\mathbf{k}_{\parallel}$ points, i.e. proportional to Δ as one would expect for the trapezoidal rule. However, for the

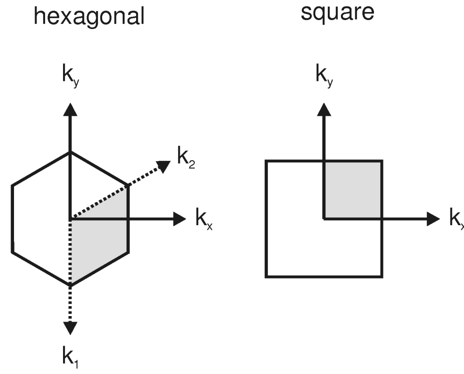

Figure 4.3: The integration ranges for the hexagonal and for the square lattice. As can be seen from Fig. 4.2 the irreducible wedges for these two lattices are actually smaller than the extended irreducible wedges.

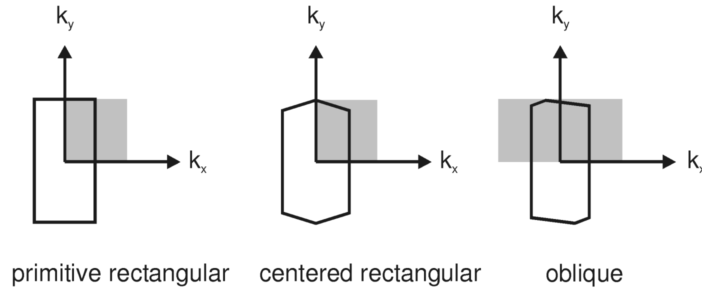

Figure 4.4: At the boundary of the two-dimensional Brillouin zone the irreducible wedge is modified to obtain integration ranges of square and rectangular shape, respectively. This procedure is justified by the fact that the $k \cdot p$ dispersion does not obey the translational invariance of the reciprocal point lattice.

case of the rhombus shaped integration range, this simple two-dimensional trapezoidal rule has to be changed. The reason for this is that to satisfy the symmetry of the hexagonal Brillouin zone (which is a multiple of a three-fold symmetry) the integration patches should be equilateral triangles instead of rhombi. This is shown in Figure 4.5. But rhombi with interior angles of 60° and 120° can be thought as being made up of two equilateral triangles, i.e.

$$\begin{aligned} g_p &= \frac{1}{6}\sum_{i=1,2,4} p_j\left(x,\mathbf{k}_\parallel^{(p,i)}\right) f\left(E_j\left(\mathbf{k}_\parallel^{(p,i)}\right)\right) + \frac{1}{6}\sum_{i=2,3,4} p_j\left(x,\mathbf{k}_\parallel^{(p,i)}\right) f\left(E_j\left(\mathbf{k}_\parallel^{(p,i)}\right)\right) \\ &= \frac{1}{6}\sum_{i=1,3} p_j\left(x,\mathbf{k}_\parallel^{(p,i)}\right) f\left(E_j\left(\mathbf{k}_\parallel^{(p,i)}\right)\right) + \frac{1}{3}\sum_{i=2,4} p_j\left(x,\mathbf{k}_\parallel^{(p,i)}\right) f\left(E_j\left(\mathbf{k}_\parallel^{(p,i)}\right)\right). \end{aligned} \tag{4.13}$$

The convergence behavior shows that we are able to calculate $\int_\Omega d^2\mathbf{k}_\parallel g\left(\mathbf{k}_\parallel\right)$ with arbitrary precision. However, this requires to know $p_j\left(x,\mathbf{k}_\parallel\right)$ and $E_j\left(\mathbf{k}_\parallel\right)$ for a correspondingly large number of $\mathbf{k}_\parallel$ points. But we can also improve the accuracy of our integration without increasing the number of $\mathbf{k}_\parallel$ points where $p_j\left(x,\mathbf{k}_\parallel\right)$ and $E_j\left(\mathbf{k}_\parallel\right)$ are evaluated exactly. The main source for integration errors has its reason in the Fermi function. The Fermi function drops from one to zero over an energy range of the order of $k_{\mathrm{B}}T$. If the difference in energies at neighboring $\mathbf{k}_\parallel$

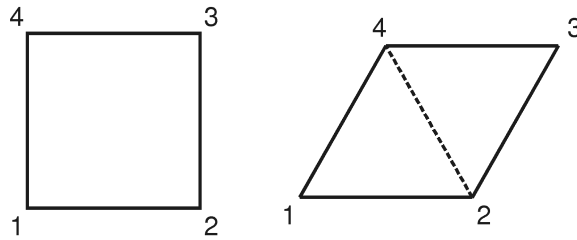

Figure 4.5: The two distinct integration patches. For the square and rectangular integration range, the integration patch is a square (left). For the rhombus shaped integration range, the patch is a rhombus with interior angles of 60° and 120° (right). This can be thought of being made up of two equilateral triangles. The vertices are numbered from 1 to 4 starting with the lower left vertex and proceeding counter-clockwise.

points is larger than $k_\mathrm{B}T$ or of the order of $k_\mathrm{B}T$, the integration error is large even if the change in $p_j\left(x, \mathbf{k}_\parallel\right)$ and in $E_j\left(\mathbf{k}_\parallel\right)$ is small. This problem can be solved by an appropriate interpolation of $p_j\left(x, \mathbf{k}_\parallel\right)$ and $E_j\left(\mathbf{k}_\parallel\right)$ between the vertices of an integration patch. Equation (4.12) and Eq. (4.13) then translate into

$$g_p = \frac{1}{4} \sum_{u,v=1}^{N} \sum_{i=1}^{4} p_j^{(u,v)} f\left(E_j^{(u,v)}\right) \tag{4.14}$$

and

$$g_p = \frac{1}{6} \sum_{u,v=1}^{N} \left[\sum_{i=1,3} p_j^{(u,v)} f\left(E_j^{(u,v)}\right) + 2 \sum_{i=2,4} p_j^{(u,v)} f\left(E_j^{(u,v)}\right) \right]. \tag{4.15}$$

The given formulas for g_p assume that we introduce $N-2$ additional points in reciprocal space along the two sides of the integration patch. These additional $\mathbf{k}_\parallel$ points as well as the $\mathbf{k}_\parallel$ points at the vertices are numbered by u, v ranging from 1 to N. The interpolated values for the probability density and for the energy dispersion at point (u, v) are then assumed to be given by $p_j^{(u,v)}$ and $E_j^{(u,v)}$, respectively. It turns out that for the probability density a bilinear interpolation is sufficient, whereas the energy dispersion has to be interpolated in a bicubic way. The reason for this is that the integration is very sensitive on the energy dispersion at minima that highly contribute to the density, e.g. at the Γ point. A linear interpolation always underestimates the contribution at minima and this underestimation is the more pronounced the larger the curvature of the minimum.

Being a local interpolation scheme, the bilinear interpolation of the probability density depends on the data of a single patch only. In contrast, we require for the bicubic interpolation continuity as well as differentiability of the energy dispersion between neighboring patches. The derivatives of the energy dispersion at the vertices of each patch are determined as for a two-dimensional spline [62]. At the boundaries of the integration range we choose the boundary conditions of a natural spline.

4.2.2 Irreducible wedge of the two-dimensional Brillouin zone

The determination of the irreducible wedge is a complicated group theoretical problem that becomes even worse if strain has to be considered. Here we follow a different approach and

determine the irreducible wedge by evaluating the subband dispersion $E_j\left(\mathbf{k}_{\|}\right)$ at specific points in the two-dimensional Brillouin zone.

Actually there is only a small set of symmetries that has to be considered for $E_j\left(\mathbf{k}_{\|}\right)$, namely C_2, C_4, C_6, D_2, D_4, and D_6. To determine the symmetry of $E_j\left(\mathbf{k}_{\|}\right)$, we evaluate $E_j\left(\mathbf{k}_{\|}\right)$ at selected $\mathbf{k}_{\|}$ points and check if $E_j\left(\mathbf{k}_{\|}\right)$ obeys the corresponding symmetry. The symmetry group C_2 is always guaranteed due to Kramer's theorem, i.e. $E_j\left(\mathbf{k}_{\|}\right) = E_j\left(-\mathbf{k}_{\|}\right)$. The remaining rotational symmetries can be easily identified. For this, we choose a reference point $\mathbf{k}_{\|}^{\mathrm{ref}}$ that is not invariant under the symmetry operations of the symmetry groups under consideration, and we calculate a reference spectrum $E_j^{\mathrm{ref}} = E_j\left(\mathbf{k}_{\|}^{\mathrm{ref}}\right)$. We apply a rotation of 60° to $\mathbf{k}_{\|}^{\mathrm{ref}}$ and calculate the spectrum $E_j\left(R_{60^\circ}\mathbf{k}_{\|}^{\mathrm{ref}}\right)$. If $E_j\left(R_{60^\circ}\mathbf{k}_{\|}^{\mathrm{ref}}\right)$ is identical to E_j^{ref} we conclude that the dispersion relation obeys C_6 symmetry. If not, we apply a rotation of 90° to $\mathbf{k}_{\|}^{\mathrm{ref}}$ and calculate the corresponding spectrum. If $E_j\left(R_{90^\circ}\mathbf{k}_{\|}^{\mathrm{ref}}\right)$ is identical to E_j^{ref}, the dispersion obeys C_4 symmetry, if not, it only obeys C_2 symmetry.

The identification of reflection axes is more involved. For this, we exploit that the Brillouin zone has the full point symmetry of the point lattice. To construct the Brillouin zone, one has to draw lines from the $\mathbf{k}_{\|} = 0$ lattice point to all other points of the reciprocal lattice. The perpendicular bisectors of these lines divide the reciprocal space into two half planes. By taking the intersection of all half planes that include the $\mathbf{k}_{\|} = 0$ lattice point we obtain the first Brillouin zone. To find the axes of reflection it is not necessary to actually construct the Brillouin zone. It is sufficient to check if one of the connecting lines between the $\mathbf{k}_{\|} = 0$ lattice point and all other points of the reciprocal lattice is an axis of reflection. This is done similarly to the identification of rotational symmetries, i.e. we apply the reflection that corresponds to a certain axis to the reference point $\mathbf{k}_{\|}^{\mathrm{ref}}$, calculate the spectrum, and compare this spectrum to the reference spectrum. However, as the number of reflection axes is identical to the number of symmetry elements of the rotation group, it is sufficient to identify a single axis of reflection. If the symmetry group of rotations is C_n and an axis of reflection is found, the overall symmetry group is D_n. Otherwise it is C_n.

4.2.3 Analysis of the convergence behavior

For the analysis of the convergence behavior of the presented method we have calculated the contribution of the Γ-band to the charge density in a small piece of homogeneously n doped GaAs for different doping concentrations ($10^{16}\,\mathrm{cm}^{-3}$, $10^{20}\,\mathrm{cm}^{-3}$) and different temperatures (4 K, 70 K, 300 K). We have chosen this example as it allows a direct comparison to the results obtained by an analytical calculation of the charge density according to Equation (1.52). The simulated structure corresponds to a GaAs slab of 1 nm in the direction of (100) and is discretized on a uniform one-dimensional grid with 0.1 nm grid spacing. This amounts to a total of 11 grid points. For all doping concentrations and for all temperatures, only the lowest

	Temperature [K]	Doping [cm^{-3}]	Occupation [cm^{-2}]
(a)	300	10^{16}	$1.657013703 \times 10^{10}$
(b)	70	10^{20}	$1.497683877 \times 10^{13}$
(c)	4	10^{20}	$1.495012554 \times 10^{13}$

Table 4.1: The results for the analytical calculation of the subband occupation.

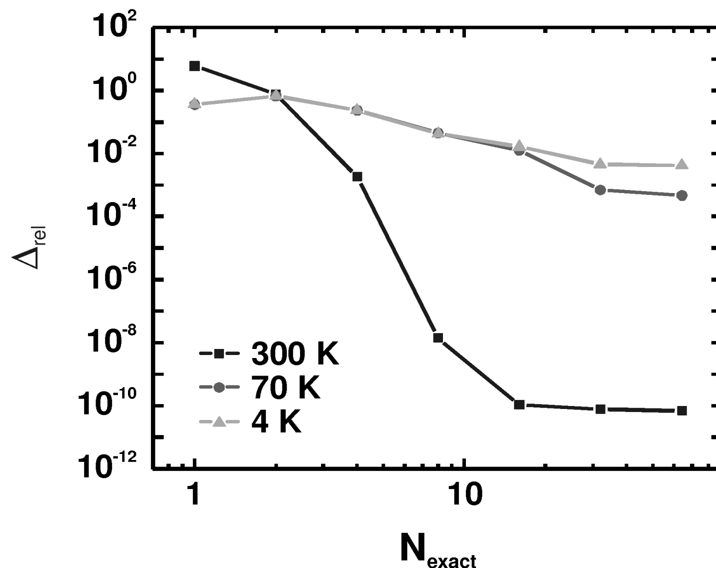

Figure 4.6: Relative error Δ_{rel} as a function of the number of exact $\mathbf{k}_{\|}$ points N_{exact} for the three temperatures $T = 300\,\text{K}$ (squares), $T = 70\,\text{K}$ (circles), and $T = 4\,\text{K}$ (triangles). The lines are guides to the eye.

subband in the Γ-band is significantly occupied. Hence we use the occupation of this subband as the criterion for the convergence of the numerical integration. Table 4.1 lists the results for the analytical integration.

To demonstrate the convergence behavior we determine the relative error Δ_{rel} of the numerical result for the occupation N_{n} from the analytical result N_{a}. This relative error Δ_{rel} is defined by

$$\Delta_{\text{rel}} = \frac{|N_{\text{n}} - N_{\text{a}}|}{N_{\text{a}}}, \tag{4.16}$$

and eventually it is limited by roundoff errors. We analyze the following aspects: (i) convergence with respect to the number N_{exact} of $\mathbf{k}_{\|}$ points where $p_j\left(x, \mathbf{k}_{\|}\right)$ and $E_j\left(\mathbf{k}_{\|}\right)$ are evaluated exactly, (ii) convergence with respect to the number N_{inter} of additional points where $p_j\left(x, \mathbf{k}_{\|}\right)$ and $E_j\left(\mathbf{k}_{\|}\right)$ are found by interpolation. We also determine the relative error δ_{rel} of the numerical result N_{n} for a certain number of additional points N_{inter} from the asymptotic value N_{n}^{∞} in the limit of large N_{inter}. This relative error is defined by

$$\delta_{\text{rel}} = \frac{|N_{\text{n}} - N_{\text{n}}^{\infty}|}{N_{\text{n}}}.$$

We begin by considering Δ_{rel} as a function of N_{exact}. The results for the three different temperatures (300 K: squares, 70 K: circles, 4 K: triangles) are shown in Figure 4.6. Both axes of Fig. 4.6 and the following figures in this section are drawn in logarithmic scale. Figure 4.6 shows that the method yields very accurate results: for a temperature of 300 K the relative error is limited to 10^{-10} due to roundoff errors. The second conclusion that can be drawn from this figure is the fact that the convergence behavior depends strongly on the temperature. This is no surprise, as we have already identified the Fermi function as the main source for integration errors. The Fermi function drops from one to zero over an energy range of the order of $k_{\text{B}}T$.

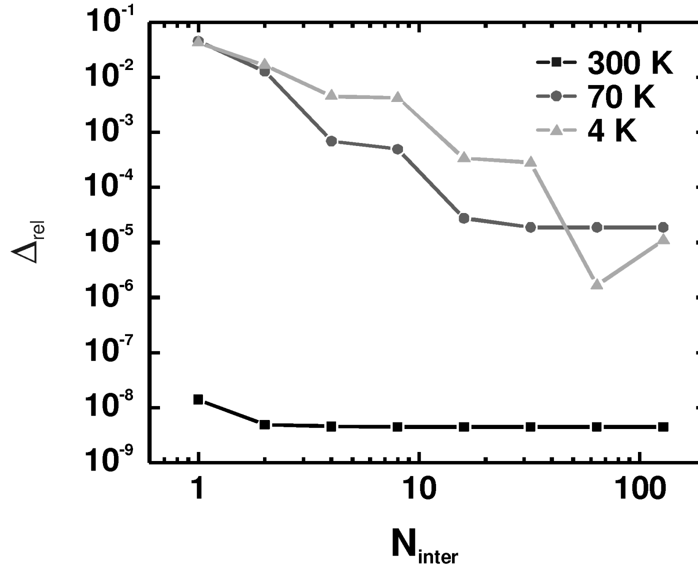

Figure 4.7: Relative error Δ_{rel} as a function of the number of interpolated $\mathbf{k}_\|$ points N_{inter} for the three temperatures $T = 300\,\text{K}$ (squares), $T = 70\,\text{K}$ (circles), and $T = 4\,\text{K}$ (triangles). The lines are guides to the eye.

If the difference in energies at neighboring $\mathbf{k}_\|$ points is larger than $k_{\text{B}}T$ or of the order of $k_{\text{B}}T$, the integration error is large even if the change in $p_j\left(x, \mathbf{k}_\|\right)$ and in $E_j\left(\mathbf{k}_\|\right)$ is small.

Figure 4.7 shows that this problem can be solved by an appropriate interpolation of $p_j\left(x, \mathbf{k}_\|\right)$ and $E_j\left(\mathbf{k}_\|\right)$ between the vertices of an integration patch. In this figure Δ_{rel} is plotted as a function N_{inter}, again for the three different temperatures (300 K: squares, 70 K: circles, 4 K: triangles). The addition of $\mathbf{k}_\|$ points where $p_j\left(x, \mathbf{k}_\|\right)$ and $E_j\left(\mathbf{k}_\|\right)$ are determined by interpolation yields basically no improvement for $T = 300\,\text{K}$ because in this case the energy spacing of the exact $\mathbf{k}_\|$ points is already smaller then $k_{\text{B}}T$. In contrast, for $T = 70\,\text{K}$ and $T = 4\,\text{K}$, the result can be significantly improved by the addition of interpolated $\mathbf{k}_\|$ points without a noticeable increase in computational effort.

For determining $E_j\left(\mathbf{k}_\|\right)$ at the interpolated $\mathbf{k}_\|$ points we use a cubic interpolation scheme. Figure 4.8 illustrates that interpolating $E_j\left(\mathbf{k}_\|\right)$ linearly is indeed not sufficient. The figure shows the relative error Δ_{rel} as a function of N_{inter} for $T = 300\,\text{K}$ for the cubic interpolation scheme as well as for a linear interpolation scheme. The expected behavior is that increasing N_{inter} should improve the accuracy of the numerical integration, i.e. Δ_{rel} should become smaller. However, this is the case only for the cubic interpolation (squares). In contrast, for the linear interpolation (circles), Δ_{rel} is increased by several orders of magnitude by the addition of a single interpolated point and the addition of further interpolated points yields absolutely no improvement.

A similar analysis for the relative error Δ_{rel} now as a function of N_{exact} is shown in Figure 4.9. Again we consider the case of $T = 300\,\text{K}$. If linear interpolation (circles) is used for $E_j\left(\mathbf{k}_\|\right)$, the convergence behavior is worse by several orders of magnitude as compared to cubic interpolation (squares). For this example, $N_{\text{inter}} = 3$ interpolated $\mathbf{k}_\|$ points have been used. Also shown in this figure is a comparison between the two patch weighting formulas given by Eq. (4.12) (triangles down) and Eq. (4.13) (triangles up) both applied assuming D_6 symmetry. Here again, a significant difference in the convergence behavior is obtained. Only using the

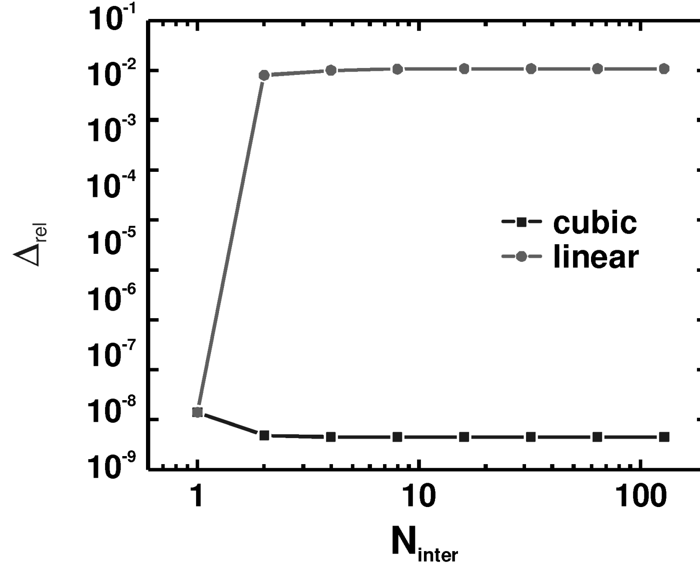

Figure 4.8: Comparison between the relative error Δ_{rel} as a function of the number of interpolated $\mathbf{k}_{\|}$ points N_{inter} for cubic (squares) and linear (circles) interpolation of $E_j\left(\mathbf{k}_{\|}\right)$ for $T = 300\,\text{K}$. The lines are guides to the eye.

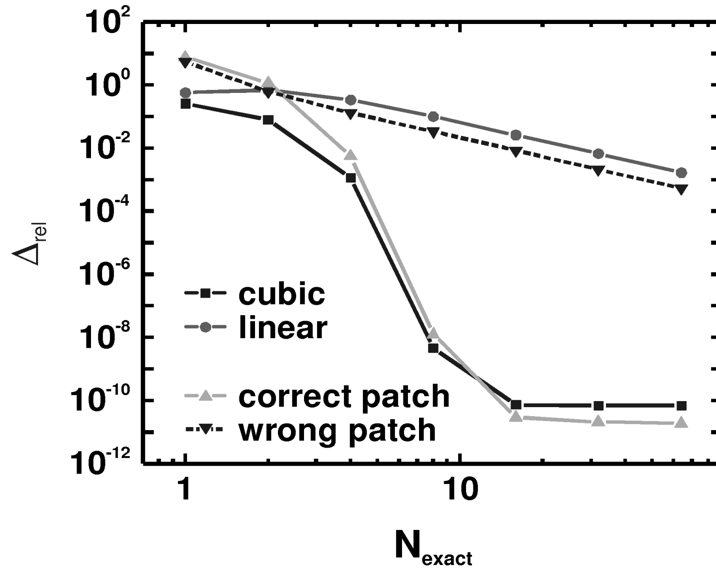

Figure 4.9: Comparison between the relative error Δ_{rel} as a function of the number of interpolated $\mathbf{k}_{\|}$ points N_{inter} for (i) cubic (squares) and linear (circles) interpolation of $E_j\left(\mathbf{k}_{\|}\right)$, (ii) patch weighting formulas given by Eq. (4.13) (triangles up) and Eq. (4.12) (triangles down). The temperature is 300 K. The lines are guides to the eye.

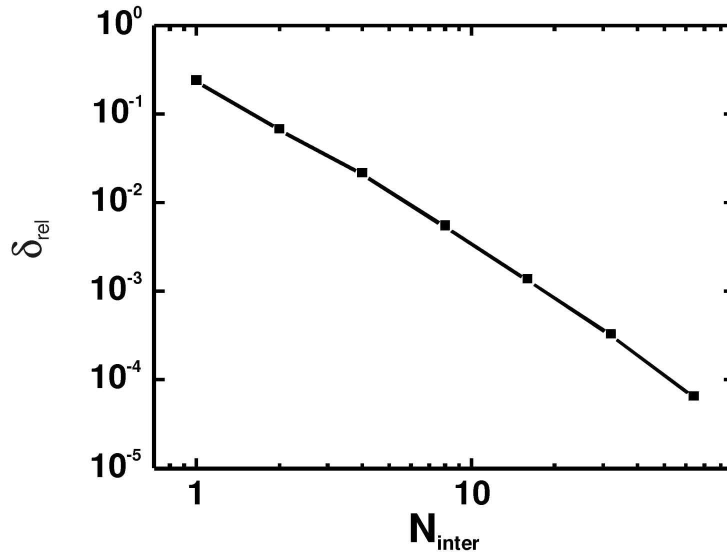

Figure 4.10: The relative error δ_{rel} as a function of the interpolated $\mathbf{k}_{\|}$ points N_{inter} for $T = 300\,\mathrm{K}$. This figure shows the quadratic convergence behavior of the integration method. The lines are guides to the eye.

patch weighting formula given by Eq. (4.13) it is possible to achieve the same level of convergence as assuming non-hexagonal symmetry. For hexagonal symmetry the patch weighting formula given by Eq. (4.12) yields results that are worse by several orders of magnitude.

Finally, in Fig. 4.10 we show the relative error δ_{rel} as a function of N_{inter}. This analysis illustrates the polynomial order that governs the convergence of the numerical integral. Again we consider the case of $T = 300\,\mathrm{K}$. The expected behavior is quadratic convergence in the spacing of neighboring $\mathbf{k}_{\|}$ points. This is confirmed by Fig. 4.10: the plotted data points can be fitted by a straight line with a slope of -2. As both axes in Fig. 4.10 are drawn in logarithmic scale, this means that δ_{rel} approaches zero quadratically.

4.3 A novel method for the charge density calculation in the 8-band $k \cdot p$ model

For the 8-band $k \cdot p$ model and for all other multi-band $k \cdot p$ models that deal with conduction as well as valence bands simultaneously (mixed-band $k \cdot p$ models) it is in general not possible to decide whether a specific eigenstate has to be occupied by an electron or by a hole. In particular, it is therefore not possible to calculate the electron and hole charge densities in the usual way simply by occupying the electronic states according to the Fermi-Dirac statistics of electrons and the hole states according to the Fermi-Dirac statistics of holes.

We have therefore developed a novel method for calculating the charge density for mixed-band $k \cdot p$ models. For this method, we do not have to classify eigenstates into electron like or hole like. Instead, we occupy all eigenstates (including the eigenstates that have energies in the energy range of the valence bands) with electrons according to the Fermi-Dirac statistic of electrons. Afterwards, we subtract an appropriately calculated positive background charge.

Neglecting doping and fixed charges, the usual procedure can be summarized as follows:

$$\rho(x) = -n(x) + p(x), \tag{4.17}$$

i.e. the total charge density is the difference of the density $p(x)$ of holes in the valence bands and the density $n(x)$ of electrons in the conduction band. In contrast, our approach is given by

$$\rho(x) = -n_{\mathrm{all}}(x) + \rho_{\mathrm{bg}}(x), \tag{4.18}$$

where $n_{\mathrm{all}}(x)$ is the electron density of electrons in all bands (conduction and valence bands) and $\rho_{\mathrm{bg}}(x)$ is the positive background charge.

This approach requires calculating a much larger number of eigenstates, at the moment restricting the application of the method to quasi-one-dimensional calculations. If in a specific mixed-band $k \cdot p$ model (as for the 8-band model) the number of valence bands is larger than the number of conduction bands, it is possible to occupy all non-contributing eigenstates with holes and subtract a negative background charge. For the 8-band $k \cdot p$ model this reduces the number of required eigenstates by a factor of 3.

In the following, we introduce our novel method and we show the equivalence between occupation with electrons and occupation with holes. For simplicity we neglect doping and fixed charges for the moment. We consider the case of a heterostructure with two-dimensional translational symmetry in the plane orthogonal to the growth direction, and we assume that the Hamiltonian has been discretized using box discretization along the growth direction in real space on a grid with box integration volumes Δ_i $(i = 1, 2, \ldots, N)$. The resulting Hamiltonian matrix $H = H_{\nu'\nu}\left(i', i; \mathbf{k}_\parallel\right)$ is not Hermitian, but can be represented by the product of a diagonal matrix D with $D_{ij} = \Delta_i^{-1}\delta_{ij}$ and a Hermitian matrix H' as follows:

$$H = D \cdot H'. \tag{4.19}$$

The Hamiltonian matrix still depends on the wave vector $\mathbf{k}_\parallel$ that describes the free motion of the crystal electrons in the plane orthogonal to the growth direction. Further, i and i' denote indices to the grid points x_i and take on the values $1, 2, \ldots, N$. The indices ν and ν' denote the components of the valence and conduction bands. In total, there are $N_{\mathrm{C}} = N_{\mathrm{VB}} + N_{\mathrm{CB}}$ such components. The N_{VB} valence band components are numbered from 1 to N_{VB}, the N_{CB} conduction band components are numbered from $N_{\mathrm{VB}} + 1$ to $N_{\mathrm{VB}} + N_{\mathrm{CB}}$. The eigenvalue equation for the envelope wave functions $\psi_j\left(x, \mathbf{k}_\parallel\right)$ of the jth subband is then given by

$$\sum_{\nu=1}^{N_{\mathrm{C}}} \sum_{i=1}^{N} H_{\nu'\nu}\left(i', i; \mathbf{k}_\parallel\right) \psi_j^{(\nu)}\left(x_i, \mathbf{k}_\parallel\right) = E_j\left(\mathbf{k}_\parallel\right) \psi_j^{(\nu)}\left(x_i, \mathbf{k}_\parallel\right), \qquad j = 1, \ldots, N \cdot N_{\mathrm{C}}, \tag{4.20}$$

where $E_j\left(\mathbf{k}_\parallel\right)$ denotes the subband dispersion relation. The probability density $p_j\left(x, \mathbf{k}_\parallel\right)$ is given by Equation (4.6). We now determine the electron density $n_{\mathrm{all}}(x)$ of the electrons in all bands (conduction and valence bands). For this, Eq. (4.1) still holds

$$n_{\mathrm{all}}(x) = \sum_{j=1}^{N \cdot N_C} \frac{1}{(2\pi)^2} \int_{\Omega_{\mathrm{BZ}}} \mathrm{d}^2\mathbf{k}_\parallel \, p_j\left(x, \mathbf{k}_\parallel\right) f\left(E_j\left(\mathbf{k}_\parallel\right)\right). \tag{4.21}$$

The summation runs from 1 to $N \cdot N_C$, i.e. we sum over subbands in the energy region of the conduction band as well as over subbands in the energy region of the valence bands. The Fermi

function $f\left(E_j\left(\mathbf{k}_\|\right)\right)$ depends on the Fermi energy E_F. Using the orthonormality condition for the $\psi_j\left(x,\mathbf{k}_\|\right)$

$$\sum_{\nu=1}^{N_\mathrm{C}}\sum_{i=1}^{N}\psi_j^{*(\nu)}\left(x_i,\mathbf{k}_\|\right)\psi_{j'}^{(\nu)}\left(x_i,\mathbf{k}_\|\right)\Delta_i=\delta_{jj'} \tag{4.22}$$

we find that the integrated electron density is given by

$$\sum_{i=1}^{N} n_\mathrm{all}\left(x_i\right)\Delta_i=\sum_{j=1}^{N\cdot N_C}\frac{1}{\left(2\pi\right)^2}\int_{\Omega_\mathrm{BZ}}\mathrm{d}^2\mathbf{k}_\| \, f\left(E_j\left(\mathbf{k}_\|\right)\right). \tag{4.23}$$

The Fermi energy is now determined so that

$$N_\mathrm{VB}\cdot N\cdot\frac{\Omega_\mathrm{BZ}}{\left(2\pi\right)^2}=\sum_{i=1}^{N} n_\mathrm{all}\left(x_i\right)\Delta_i, \tag{4.24}$$

which means that the $N_\mathrm{VB}\cdot N$ valence subbands have originally been fully occupied. To guarantee total charge neutrality

$$\sum_{i=1}^{N}\rho\left(x_i\right)\Delta_i=0 \tag{4.25}$$

we therefore find the condition

$$\sum_{i=1}^{N}\rho_\mathrm{bg}\left(x_i\right)\Delta_i=N_\mathrm{VB}\cdot N\cdot\frac{\Omega_\mathrm{BZ}}{\left(2\pi\right)^2}. \tag{4.26}$$

Even though the positive background charge $\rho_\mathrm{bg}\left(x\right)$ cannot be identified with the positively charged cores as is possible in tight-binding, the origin of the background charge are those cores. Since the cores are distributed homogeneously, we assume that $\rho_\mathrm{bg}\left(x_i\right)\Delta_i$ is independent of x_i and given by

$$\rho_\mathrm{bg}\left(x_i\right)\Delta_i=N_\mathrm{VB}\cdot N\cdot\frac{\Omega_\mathrm{BZ}}{\left(2\pi\right)^2}. \tag{4.27}$$

For the charge density $\rho\left(x\right)$ this yields the following expression

$$\rho\left(x_i\right)=-\sum_{j=1}^{N\cdot N_C}\frac{1}{\left(2\pi\right)^2}\int_{\Omega_\mathrm{BZ}}\mathrm{d}^2\mathbf{k}_\|\left(p_j\left(x_i,\mathbf{k}_\|\right)f\left(E_j\left(\mathbf{k}_\|\right)\right)-\frac{1}{N\cdot\Delta_i}\frac{N_\mathrm{VB}}{N_\mathrm{C}}\right). \tag{4.28}$$

The equivalence between occupation with electrons and occupation with holes is based on the completeness relation for the $\psi_j\left(x,\mathbf{k}_\|\right)$

$$\sum_{j=1}^{N\cdot N_\mathrm{C}}\psi_j^{*(\nu)}\left(x_i,\mathbf{k}_\|\right)\psi_j^{(\nu')}\left(x_{i'},\mathbf{k}_\|\right)=\delta_{\nu\nu'}\frac{\delta_{ii'}}{\Delta_i}. \tag{4.29}$$

By setting $\nu=\nu'$, $i=i'$ and subsequently summing over ν we find

$$\frac{N_\mathrm{C}}{\Delta_i}=\sum_{j=1}^{N\cdot N_\mathrm{C}}p_j\left(x_i,\mathbf{k}_\|\right). \tag{4.30}$$

Now we use that

$$
\begin{aligned}
&\sum_{j=1}^{N\cdot N_C} \frac{1}{(2\pi)^2} \int_{\Omega_{\mathrm{BZ}}} \mathrm{d}^2\mathbf{k}_\parallel \frac{1}{N\cdot\Delta_i} \frac{N_{\mathrm{VB}}}{N_{\mathrm{C}}} \\
= &\frac{1}{(2\pi)^2} \int_{\Omega_{\mathrm{BZ}}} \mathrm{d}^2\mathbf{k}_\parallel \frac{N_{\mathrm{C}}}{\Delta_i} - \frac{1}{(2\pi)^2} \int_{\Omega_{\mathrm{BZ}}} \mathrm{d}^2\mathbf{k}_\parallel \frac{N_{\mathrm{VB}}}{\Delta_i} \\
= &\sum_{j=1}^{N\cdot N_{\mathrm{C}}} \frac{1}{(2\pi)^2} \int_{\Omega_{\mathrm{BZ}}} \mathrm{d}^2\mathbf{k}_\parallel \, p_j\left(x_i, \mathbf{k}_\parallel\right) - \sum_{j=1}^{N\cdot N_{\mathrm{C}}} \frac{1}{(2\pi)^2} \int_{\Omega_{\mathrm{BZ}}} \mathrm{d}^2\mathbf{k}_\parallel \frac{1}{N\cdot\Delta_i} \frac{N_{\mathrm{VB}}}{N_{\mathrm{C}}} \\
= &\sum_{j=1}^{N\cdot N_{\mathrm{C}}} \frac{1}{(2\pi)^2} \int_{\Omega_{\mathrm{BZ}}} \mathrm{d}^2\mathbf{k}_\parallel \left(p_j\left(x_i, \mathbf{k}_\parallel\right) - \frac{1}{N\cdot\Delta_i} \frac{N_{\mathrm{VB}}}{N_{\mathrm{C}}} \right).
\end{aligned}
\tag{4.31}
$$

For the charge density this yields

$$
\rho\left(x_i\right) = \sum_{j=1}^{N\cdot N_C} \frac{1}{(2\pi)^2} \int_{\Omega_{\mathrm{BZ}}} \mathrm{d}^2\mathbf{k}_\parallel \left(p_j\left(x_i, \mathbf{k}_\parallel\right) \left[1 - f\left(E_j\left(\mathbf{k}_\parallel\right)\right)\right] - \frac{1}{N\cdot\Delta_i} \frac{N_{\mathrm{CB}}}{N_{\mathrm{C}}} \right). \tag{4.32}
$$

This expression for the charge density can be interpreted as follows: the overall sign is now positive, i.e. instead of occupying the subbands with electrons according to $f\left(E_j\left(\mathbf{k}_\parallel\right)\right)$ we occupy the subbands with holes according to $1 - f\left(E_j\left(\mathbf{k}_\parallel\right)\right)$ and we subtract a negative background charge $\bar{\rho}_{\mathrm{bg}}(x)$ that is given by

$$
\bar{\rho}_{\mathrm{bg}}\left(x_i\right) = -\frac{1}{\Delta_i} N_{\mathrm{CB}} \cdot N \cdot \frac{\Omega_{\mathrm{BZ}}}{(2\pi)^2}. \tag{4.33}
$$

Equation (4.28) and Eq. (4.32) are equivalent. The decision to use either the one or the other depends on the number of valence bands N_{VB} and the number of conduction bands N_{CB}. If $N_{\mathrm{CB}} < N_{\mathrm{VB}}$ as in the case of the 8-band $k \cdot p$ model one should prefer Eq. (4.32) for the following reason. From a practical point of view it is not necessary to calculate the sum over all eigenstates as Eq. (4.28) and Eq. (4.32) suggest. This is because the Fermi function drops exponentially for energies that are larger than the Fermi energy E_{F} by $k_{\mathrm{B}}T$. Therefore, eigenstates with eigenenergies larger than $E_{\mathrm{F}} + k_{\mathrm{B}}T$ hardly contribute to the sum in Equation (4.28). The same is true for the sum in Eq. (4.32) and eigenstates with eigenenergies less than $E_{\mathrm{F}} - k_{\mathrm{B}}T$. The number of eigenstates with eigenenergies smaller than E_{F} is of the order of $N \cdot N_{\mathrm{VB}}$, whereas the number of eigenstates with eigenenergies larger than E_{F} is of the order of $N \cdot N_{\mathrm{CB}}$. For Eq. (4.32) it is therefore sufficient to determine $N \cdot N_{\mathrm{CB}} + n$ eigenstates, but for Eq. (4.28) one has to determine $N \cdot N_{\mathrm{VB}} + n$ eigenstates, where n is usually small if compared to N. Thus $N \cdot N_{\mathrm{VB}} + n > N \cdot N_{\mathrm{CB}} + n$ in the 8-band case and accordingly Eq. (4.32) is computationally less demanding.

Finally we want to note that doping can be included as usual. The charge density is then given by

$$
\rho(x) = -n_{\mathrm{all}}(x) + \rho_{\mathrm{bg}}(x) - N_A^-(x) + N_D^+(x) \tag{4.34}
$$

where $N_A^-(x)$ and $N_D^+(x)$ are the densities of ionized acceptors and donors, respectively.

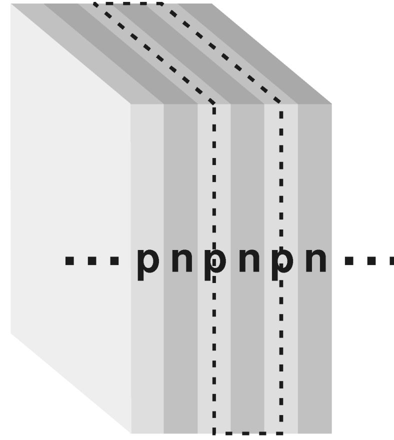

Figure 4.11: A doping superlattice is made of a single semiconducting material with periodically repeated p and n doped layers. Employing periodic boundary conditions, the calculation of such a superlattice can be limited to a single period as indicated by the dashed lines.

4.4 Application to doping superlattices

Doping superlattices, in contrast to compositional superlattices, consist of a single semiconducting material. A superstructure potential is created by periodic n and p doping. The variation of doping concentrations and doping layer thicknesses offers a wide range for tuning the properties of doping superlattices as for example the bandgap, free-carrier densities, and resulting from that, conductivity, luminescence spectra, absorption coefficients, recombination lifetimes etc.

Here we exploit that by varying the doping concentrations, a continuous transition between a situation with positive spatial band gap and a situation with negative spatial band gap is possible. This allows us to compare the results of our novel method with the results obtained with the traditional method at least in the limiting case where both methods are applicable, i.e. for positive spatial band gap. The calculations in the case of negative spatial band gap yield the intrinsic charge transfer between the n and the p doped layers. Charge transfer is expected from qualitative arguments. However, the quantitative calculation of the charge transfer is a new result because it cannot be obtained using the traditional method for calculating the charge density.

We have analyzed doping superlattices that consist of GaAs and are alternately n and p doped with doping layer thicknesses of $d_n = d_p = 10\,\text{nm}$ and doping concentrations of $N_A = 2.6 \times 10^{19}\,\text{cm}^{-3}$, $N_D = 5.2 \times 10^{19}\,\text{cm}^{-3}$ (positive spatial band gap) and $N_A = N_D = 5.2 \times 10^{19}\,\text{cm}^{-3}$ (negative spatial band gap). A sketch of the structure is shown in Figure 4.11. We have calculated the charge density and the electrostatic potential self-consistently within the 8-band $k \cdot p$ model. The temperature has been set to $T = 4\,\text{K}$.

For the calculation we have discretized the heterostructure along the (100) growth direction using a uniform grid with 0.25 nm grid spacing and periodic boundary conditions. This amounts to a total of 80 grid points for the 20 nm long structure. Using LAPACK, we have calculated the 168 8-band $k \cdot p$ states that are highest in energy. The energy subband dispersion has been evaluated at 210 exact $\mathbf{k}_{\|}$ points in the irreducible wedge (1/8th of the full two-dimensional Brillouin zone). The Brillouin zone integration has been restricted to 5% of the maximum $\mathbf{k}$ value in the bulk Brillouin zone. A total of 11 iterations were necessary to achieve a converged

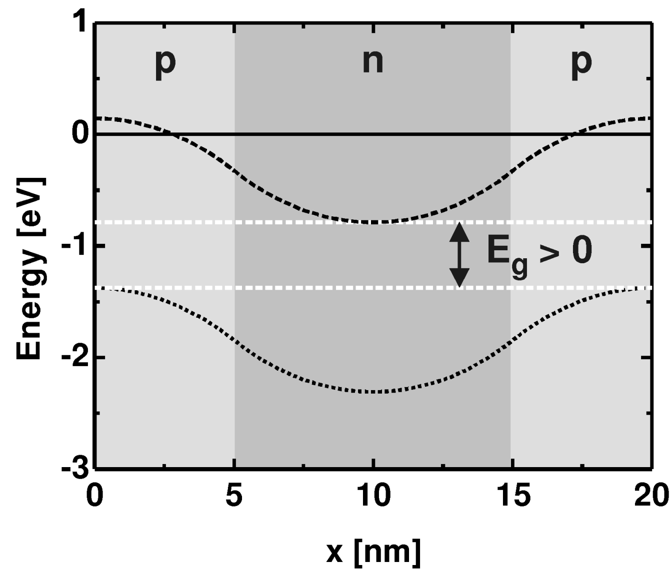

Figure 4.12: Band edges of Γ conduction band (dashed curve) and heavy hole valence band (dotted curve) along the growth axis of the doping superlattice with positive spatial band gap. The Fermi energy is $E_{\mathrm{F}} = 0\,\mathrm{eV}$.

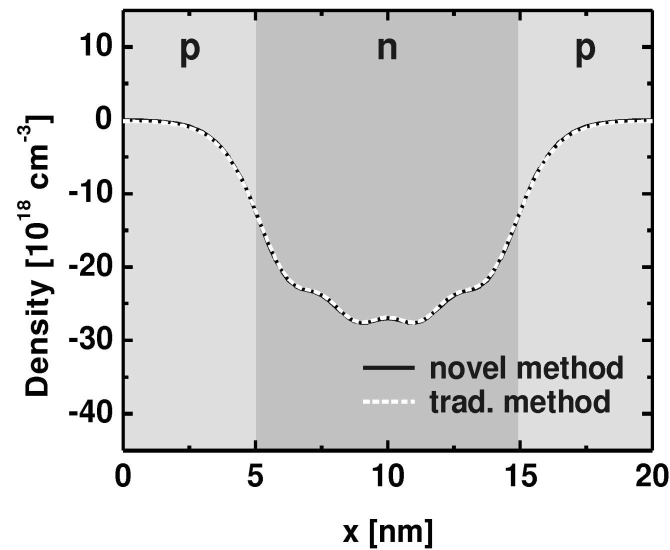

Figure 4.13: Charge carrier density along the growth direction for the doping superlattice with positive spatial band gap. The solid black curve has been calculated using the novel method, for the dashed white curve the traditional method has been used.

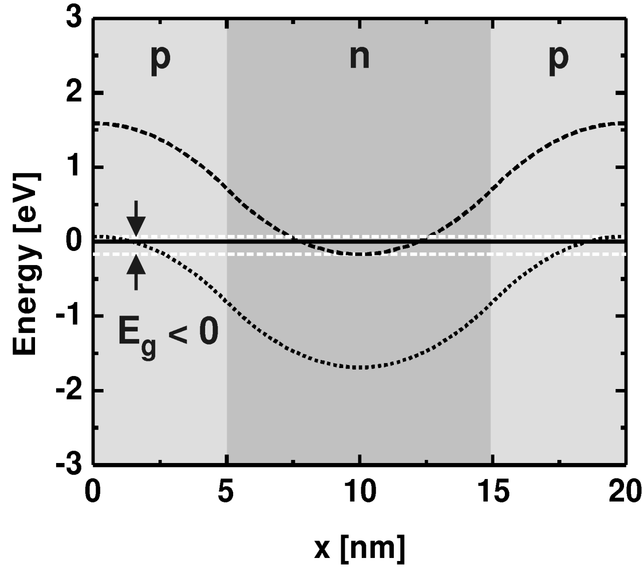

Figure 4.14: Band edges of Γ conduction band (dashed curve) and heavy hole valence band (dotted curve) along the growth axis of the doping superlattice with negative spatial band gap. The Fermi energy is $E_{\mathrm{F}} = 0\,\mathrm{eV}$.

self-consistent solution for the charge density and the electrostatic potential.

We begin with the case of positive spatial band gap. Figure 4.12 shows the band edges of the Γ conduction band (dashed curve) and the heavy hole valence band (dotted curve) including the electrostatic potential along the growth direction. The Fermi energy is $E_{\mathrm{F}} = 0\,\mathrm{eV}$. The minimum of the conduction band edge is larger than the maximum of the valence band edge, so that the spatial band gap is positive and amounts to $E_{\mathrm{g}} = 0.59\,\mathrm{eV}$. One can easily classify the eigenstates into electron like and hole like. All eigenstates that are in energy above the minimum of the conduction band edge are electron like eigenstates, all eigenstates that are in energy below the minimum of the valence band edge are hole like eigenstates. Therefore, the charge density can be calculated with our novel as well as with the traditional method. The results for the densities of charge carriers are plotted in Fig. 4.13 (novel method: solid black curve, traditional method: dashed white curve). The figure shows that the two densities coincide. This confirms that our novel method yields the same results as the traditional method in cases where both methods can be applied.

We now turn to the case of negative spatial band gap. Figure 4.14 shows the band edges of the Γ conduction band (dashed curve) and the heavy hole valence band (dotted curve) including the electrostatic potential along the growth direction. Again, the Fermi energy is $E_{\mathrm{F}} = 0\,\mathrm{eV}$. Now, the minimum of the conduction band edge is less than the maximum of the valence band edge, so that the spatial band gap is negative and amounts to $E_{\mathrm{g}} = -0.24\,\mathrm{eV}$. A classification of eigenstates into electron like or hole like is no longer possible and the traditional method fails in calculating the charge carrier density. With our novel method it is still possible to calculate the density of charge carriers as shown in Figure 4.15. The figure shows a transfer of negative charge from the p doped layers into the n doped layers. This charge transfer can be explained in detail by means of the dispersion relation of the subbands that is shown in Figure 4.16. The subband dispersion exhibits an anti-crossing of the lowest subband above the Fermi energy (denoted upper subband in the following) and the highest subband below the Fermi energy

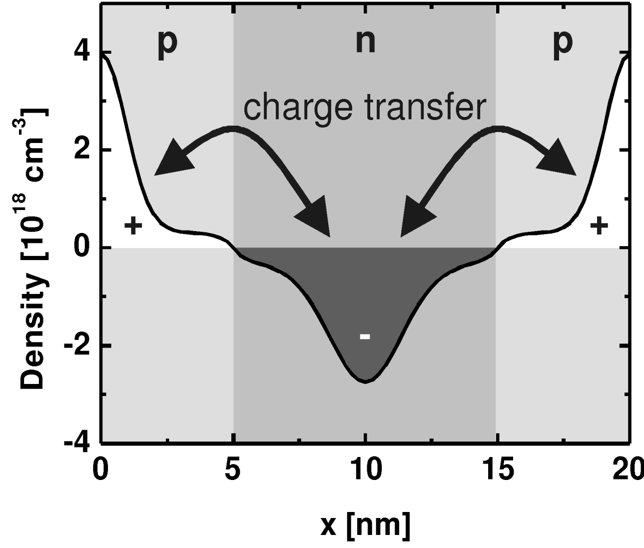

Figure 4.15: Charge carrier density along the growth direction for the doping superlattice with negative spatial band gap. This result was obtained using the novel method, the traditional method cannot be applied in this case.

(denoted lower subband in the following). This anti-crossing is shown in detail in the inset (d) of the figure. For large values of $\mathbf{k}_\parallel$ we can undoubtedly assign the property electron or hole to either of the subbands. The upper subband is electron like, the lower subband is hole like. Further, inset (c) shows the probability distribution of the eigenstate of the upper subband for $\left|\mathbf{k}_\parallel\right| = 0.05 \cdot 2\pi/a$. The probability density is localized in the n layers that form potential wells for electrons. However, at the anti-crossing the classification of the upper and the lower subband into electron and hole is no longer possible. This can be clearly seen from inset (b) that shows the probability distribution of the eigenstate of the upper subband for a $\mathbf{k}_\parallel$ point near the anti-crossing. Keeping in mind the periodic boundary conditions, the probability distribution shows two distinct maxima that correspond to localization in the n layers as well as in the p layers. This eigenstate is an example of the hybridization of conduction band and valence band states and makes the classification into electrons and holes impossible. At the zone center, the classification that holds for large values of $\mathbf{k}_\parallel$ is reversed. As shown in inset (a), the probability density of the eigenstate of the upper subband is now localized completely within the p layers that form potential wells for holes. This means that for $\mathbf{k}_\parallel = 0$ an eigenstate that has its origin in the valence bands is higher in energy than an eigenstate that has its origin in the conduction band. Because of this, the energy of the eigenstate originating from the valence band is larger than the Fermi energy, its occupation therefore lacks electrons and creates a positive charge distribution. Since the probability density of this state is mainly localized in the p layers, we thus observe a positive charge distribution in the p layers. In each case, the opposite is true for the eigenstate originating from the conduction band. We therefore observe a negative charge distribution in the n layers. Altogether this results in the observed charge transfer.

This argument shows clearly that every method based on a classification of eigenstates or even subbands must fail in this situation. First, it is not possible to assign the properties electron and hole unambiguously to a subband. Second, at certain $\mathbf{k}_\parallel$ points as the one of the anti-crossing, it is not even possible to assign the properties electron and hole to an eigenstate.

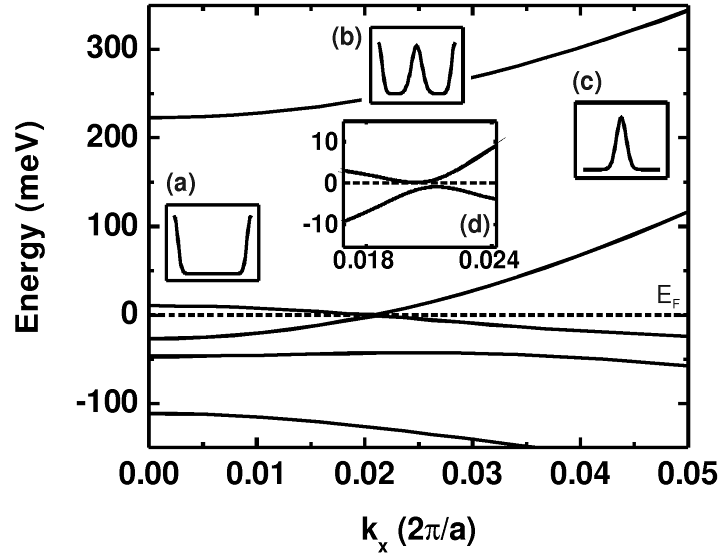

Figure 4.16: $\mathbf{k}_{\parallel}$ subband dispersion relation in the (010) direction. The insets (a), (b), and (c) show the probability density of the eigenstates of the lowest subband above the Fermi energy E_{F} for three different values of $k_{\parallel}$ $(ak_{\parallel}/2\pi = 0, 0.02, 0.05)$. The dispersion exhibits an anti-crossing of the lowest subband above the Fermi energy and the highest subband below the Fermi energy (inset (d)).

To summarize, we have shown that in the case of positive spatial band gap, the result for the density of charge carriers calculated with our novel method is identical to the result obtained by the traditional method. However, with our novel method we can also address the case of negative spatial band gap. This has allowed us to analyze the charge transfer between the p and the n doped layers quantitatively and we could explain this charge transfer in detail by the subband dispersion relation of the lowest subband above the Fermi energy and the highest subband below the Fermi energy.

4.5 Summary and Outlook

In this chapter, we have presented a novel method for the calculation of the charge carrier densities of semiconductor heterostructures by numerical integration over the two-dimensional Brillouin zone. We have compared the method with the results obtained by analytical integration for homogeneous GaAs. By this comparison we have shown that our method yields accurate results for the charge carrier densities with a moderate computational effort. Based on the method for the calculation of the charge carrier density, we have presented a second method for calculating the charge carrier density for mixed-band $k \cdot p$ models that does not require the classification of the electronic states into hole like and electron like. We have argued that such a classification fails for broken gap structures so that the traditional method for calculating the charge carrier densities is no longer applicable. We have demonstrated the method by quantitatively calculating the intrinsic charge transfer for a doping superlattice with broken gap. The method can be easily integrated into the highly efficient predictor-corrector approach for the charge self-consistent solution of the coupled Schrödinger and Poisson equations presented in Chapter 2.

In the future, the presented scheme for the fully charge self-consistent calculation of the electronic structure of broken gap heterostructures within the 8-band $k \cdot p$ approximations allows for the accurate prediction of optical properties such as transition energies and absorption coefficients of type II infrared lasers and infrared detectors.

Chapter 5

A novel method for ballistic quantum transport in entangled two-particle systems

5.1 Introduction to ballistic quantum transport

A conductor shows non-ohmic behavior if its dimensions are smaller than certain characteristic lengths: the mean free path and the phase-relaxation length [21]. If the length of a conductor becomes shorter than the mean free path the conductance approaches a limiting value. This classical ballistic limit has still nothing to do with quantum mechanics. Quantum mechanics becomes important not until the dimensions of the conductor are smaller than the phase-relaxation length and interference-related effects come into play. In present day high-mobility semiconductor heterostructures such as modulation doped GaAs/AlGaAs heterojunctions or quantum wells, mean free paths and phase-relaxation lengths of the order of 30–40 μm are relatively easy to obtain at low temperatures. Thus ballistic quantum transport plays an important role in many mesoscopic transport experiments.

The theoretical approach that has proven to be most useful in describing mesoscopic transport was introduced by Landauer [63, 64] in 1988. The roots of this approach date back even to the early 1930s [65, 66]. A generalization to multi-terminal devices including magnetic fields was proposed by Büttiker [67, 68] and is generally referred to as Landauer-Büttiker formalism. The essential idea behind the Landauer-Büttiker formalism is that the current through a ballistic conductor is determined by the probabilities of the electrons to be transmitted or reflected. The contacts of the conductor are assumed to be large electron reservoirs in equilibrium that each can be described by its own Fermi distribution with chemical potential μ. The difference in chemical potentials between the contacts is equal to the externally applied bias voltage. By the Landauer-Büttiker formula, these relations are expressed as follows:

$$I_{pq} = \frac{2e}{h} \int \mathrm{d}E \, T_{pq}(E) \left[f\left(\mu_p, E\right) - f\left(\mu_q, E\right) \right], \tag{5.1}$$

where I_{pq} is the current between contact p and contact q, $T_{pq}(E)$ is the corresponding transmission probability, μ_p and μ_q are the chemical potentials in contacts p and q, and $f(\mu, E)$ is the Fermi distribution function. Thus the Landauer-Büttiker formalism reduces the problem of calculating the ballistic current in a mesoscopic device to the determination of the transmission probabilities of an open device connected to reservoirs.

A typical problem where the Landauer-Büttiker formalism can be successfully applied is ballistic quantum transport. Here, the transmission probabilities can be determined using quantum mechanical scattering theory. For this, one first calculates the scattering matrix S of the system, and the transmission probabilities T_{pq} are then obtained according to

$$T_{pq} = |S_{pq}|^2 . \tag{5.2}$$

Several numerical methods have been developed to determine the scattering matrix for quantum devices. The transfer matrix method [69, 70] as well as its generalizations to multi-dimensional systems [71, 72] and to multi-band Hamiltonians [73, 74] follow the direct approach where the scattering matrix is determined from the solutions of the Schrödinger equation with scattering boundary conditions. An improved method (quantum transmitting boundary method) for applying the scattering boundary conditions was introduced in Reference [72]. This method still requires the repeated solution of a linear system of equations with dimensions proportional to the size of the device, but has been successfully applied to tight-binding and multi-band $k \cdot p$ calculations in one dimension [73–76] and single-band calculations in two as well as in three dimensions [72, 77–79]. In the R-matrix method [80] the scattering matrix is determined from eigenfunctions of a closed system, the so called Wigner-Eisenbud functions. This method is widely used in atomic physics but has not been applied to semiconductor devices with the exception of the works by Wulf *et al.* [81, 82]. A very efficient and widely used method is the recursive Green's function method [83, 84]. It has been successfully used for two-dimensional [85, 86] as well as small three-dimensional [87] devices. In general, Green's function methods benefit from the well developed theory of Green's function. A particular advantage of the recursive Green's function method is that it yields the retarded Green's function as well as the lesser Green's function. The latter is required if inelastic scattering is to be considered. The main disadvantage of the method is its limitation to devices that can be separated into discrete slices. This often restricts the number of contacts to only two. A recent and also very efficient method is the contact block reduction (CBR) method [88]. The CBR method can be applied to arbitrarily shaped devices with arbitrary number of contacts in up to three dimensions. Extensions to multi-band Hamiltonians as well as to charge self-consistent calculations have been demonstrated recently [6, 22].

The presented methods have in common that many-particle effects are treated only perturbatively within the Hartree approximation or the local density approximation (LDA). In particular, this means that many-particle correlation effects such as entanglement cannot be discussed. Entanglement, however, is the key point in quantum information processing, as any non-trivial two-qubit quantum gate is based on entanglement. We have developed a novel numerical method for ballistic quantum transport based on the Green's function approach that allows one to treat the interaction between two particles non-perturbatively. With this method it is therefore possible to study many-particle correlation effects such as entanglement. We show that in doing so, we can still benefit from the high computational efficiency of the CBR method.

In the following, we first provide a brief summary of the Green's function formalism by an introduction to the CBR method. Then, we present our novel method for ballistic quantum transport in entangled two-particle systems and we show that we can still benefit from the high computational efficiency of the CBR method. In Sec. 5.4 we derive interaction matrix elements for the Coulomb interaction to be used in the particular application of the method that is discussed in Chapter 6.

5.2 Contact block reduction method

Reduction to the contact block

In the following, we introduce the terminology and setup the notation that is used throughout this chapter. We denote by the term "device" a finite n-dimensional ($n = 1, 2, 3$) region that is coupled to reservoirs by an arbitrary number of leads. The device may be under applied bias and contain some spatially varying potential. The total Hamiltonian of the system, including the device and the leads, can be written in symbolic matrix form

$$H_{\text{tot}} = \begin{pmatrix} H_1 & 0 & 0 & W_1 \\ 0 & \ddots & 0 & \vdots \\ 0 & 0 & H_L & W_L \\ W_1^\dagger & \cdots & W_L^\dagger & H^0 \end{pmatrix}, \tag{5.3}$$

where H_p represent the Hamiltonian of lead p, the Hamiltonian H^0 corresponds to the device region, and W_p is the coupling between the device and this lead ($p = 1, \ldots, L$). The leads (acting as reservoirs) are semi-infinite and therefore, the total Hamiltonian H_{tot} is infinite dimensional. This infinite dimensional Hamiltonian can be reduced to an effective (non-hermitian) finite dimensional Hamiltonian $H = H^0 + \Sigma$ that describes the *open device*. In this formulation, the influence of the leads is included through a finite-dimensional operator $\Sigma = \Sigma^1 + \cdots + \Sigma^L$, the sum of the complex contact self-energies Σ^p. In the following, Σ is called the total self-energy. The Hermitian Hamiltonian H^0 represents the *decoupled device*, i.e. the device with no coupling to the leads. In the ballistic case, i.e. if the Hamiltonian H does not contain scattering terms, all observables of interest can, in principle, be calculated from the retarded Green's function G^R of the open device, which is defined by

$$G^R = (E - H)^{-1} = \left(E - H^0 - \Sigma\right)^{-1}. \tag{5.4}$$

The retarded Green's function can be obtained from the Green's function G^0 of the decoupled device by the Dyson equation

$$G^R = \left(\mathbf{1} - G^0\Sigma\right)^{-1} G^0. \tag{5.5}$$

The Green's function G^0 is defined by

$$G^0 = \left(E - H^0 + i\eta\right)^{-1}, \qquad \eta \to 0^+, \tag{5.6}$$

and can be calculated with the help of its spectral representation in terms of the eigenstates $|\alpha\rangle$ of the decoupled device Hamiltonian H^0

$$G^0 = \sum_\alpha \frac{|\alpha\rangle \langle\alpha|}{E - \varepsilon_a + i\eta}, \qquad \eta \to 0^+, \tag{5.7}$$

$$H^0 |\alpha\rangle = \varepsilon_a |\alpha\rangle. \tag{5.8}$$

The direct evaluation of G^R according to Eq. (5.5) requires the inversion of a large matrix that is proportional to the number of grid points N_T of the device. Practically, this is impossible in three spatial dimensions and can be quite demanding even in two spatial dimensions. The essence of the CBR method consists in the realization that the determination of G^R from G^0 and Σ actually requires only the inversion of a matrix that is proportional to the number of

grid points N_C that connect the device with the leads. In three spatial dimensions, N_C is of the order of $N_T^{2/3}$. The most convenient derivation of this result is obtained by dividing the device into $N_D = N_T - N_C$ interior grid points Ω_D and N_C contact grid points Ω_C that connect the device with the leads. Thus the total set of grid points is given by $\Omega = \Omega_D \cup \Omega_C$. The total self-energy Σ vanishes at all interior grid points as each contact self energy is non-zero only at grid points belonging to the respective contact. We are therefore allowed to represent Σ in block-diagonal matrix form as follows:

$$\Sigma = \begin{pmatrix} \Sigma_C & 0 \\ 0 & 0_D \end{pmatrix}. \tag{5.9}$$

The submatrix Σ_C is termed the *contact block* of Σ and is a small square matrix of dimension N_C. It is given by the sum of all contact blocks $\Sigma_C^1, \ldots, \Sigma_C^L$ of the contact self-energies $\Sigma^1, \ldots, \Sigma^L$. The large square null-matrix 0_D corresponds to the interior grid points and is of dimension N_D. Analogously to Σ we decompose the Green's function of the decoupled device into blocks

$$G^0 = \begin{pmatrix} G_C^0 & G_{CD}^0 \\ G_{DC}^0 & G_D^0 \end{pmatrix}. \tag{5.10}$$

Note again that G_C^0 is a small square matrix of dimension N_C, whereas G_D^0 is a large square matrix of dimension N_D. Now we define the matrix A by $A = \mathbf{1} - G^0\Sigma$ so that the retarded Green's function is given by $G^R = A^{-1}G^0$. For the matrix A, we again employ the introduced decomposition into blocks

$$A = \begin{pmatrix} A_C & 0 \\ -G_{DC}^0\Sigma_C & \mathbf{1}_D \end{pmatrix} \tag{5.11}$$

where $A_C = \mathbf{1}_C - G_C^0\Sigma_C$. The key point is that the inverse of A can be given explicitly in terms of the inverse of A_C

$$A^{-1} = \begin{pmatrix} A_C^{-1} & 0 \\ G_{DC}^0\Sigma_C A_C^{-1} & \mathbf{1}_D \end{pmatrix}. \tag{5.12}$$

In addition, we recognize that A and its inverse possess the same structure. Having inverted only the small square matrix A_C (the contact block of A), we can now calculate any block of the retarded Green's function according to

$$G^R = \begin{pmatrix} A_C^{-1}G_C^0 & A_C^{-1}G_{CD}^0 \\ G_{DC}^0\Sigma_C A_C^{-1}G_C^0 + G_{DC}^0 & G_{DC}^0\Sigma_C A_C^{-1}G_{CD}^0 + G_D^0 \end{pmatrix}. \tag{5.13}$$

In particular, the contact block of the retarded Green's function is given by

$$G_C^R = A_C^{-1}G_C^0. \tag{5.14}$$

Determination of the transmission function

The transmission functions follow from the retarded Green's function according to the Fisher-Lee relation [89]

$$S_{mn} = -\delta_{mn} + i\hbar\sqrt{v_m v_n}G_{mn}^R. \tag{5.15}$$

For multi-moded leads, m and n indicate the individual modes and v_m and v_n denote the corresponding phase velocities. Together with Eq. (5.2) the Fisher-Lee relation yields the transmission between lead p and lead q

$$T_{pq}(E) = \mathrm{Tr}\,\Gamma_C^p G_C^R \Gamma_C^q G_C^{R\dagger}, \qquad p \neq q, \tag{5.16}$$

$$T_{pq}(E) = \mathrm{Tr}\left(\mathbf{1}_C - i\Gamma_C^p G_C^R\right)\left(\mathbf{1}_C - i\Gamma_C^q G_C^R\right)^\dagger, \qquad p = q, \tag{5.17}$$

where $\Gamma^p = i\left(\Sigma^p - \Sigma^{p\dagger}\right)$. Note that only the contact blocks of the matrices Γ^p, Γ^q, and G^R contribute to the transmission functions as the Hermitian matrices Γ^p have the same structure as the contact self-energies Σ^p and therefore vanish at interior grid points.

Determination of the particle density

For an open system connected by leads to reservoirs with distribution functions $f_p(E)$ ($p = 1, \ldots, L$), it has been shown in Ref. [90, 91] that the particle density at grid point i can be expressed in the following form

$$n_i = \frac{1}{2\pi}\int \mathrm{d}E \sum_{p=1}^{L} f_p(E)\left[G^R\Gamma^p G^{R\dagger}\right]_{ii}. \tag{5.18}$$

From a straightforward calculation one obtains

$$n_i = \frac{1}{2\pi}\int \mathrm{d}E \sum_{p=1}^{L} f_p(E)\left[G_X^R\Gamma_C^p\left(G_X^R\right)^\dagger\right]_{ii}, \tag{5.19}$$

where $X = C$ for $i \in \Omega_C$ (contact grid points) and $X = DC$ for $i \in \Omega_D$ (interior grid points). A more efficient expression for the calculation of the particle density can be found in References [6, 22].

Approximations

The following two approximations are customarily used along with the CBR method. Firstly, the Green's function of the decoupled device G^0 is represented by a reduced number of eigenstates of the decoupled device Hamiltonian H^0 instead of the full set [88]. Thereby the proper choice of boundary conditions for the decoupled device Hamiltonian (von Neumann boundary conditions instead of Dirichlet boundary conditions) helps to improve the results of this approximation quite drastically. The second approximation concerns the number of lead modes that are taken into account in the case of multi-moded leads [22]. The matrix A_C, that has to be inverted, has the dimension of the number of contact grid points N_C, which can still be quite large at least for three spatial dimensions. However, the size of A_C can be further reduced if a change of basis is performed on the real space matrices G_C^0 and Σ_C expressing G_C^0 and Σ_C in terms of the lead modes. In doing so it proves to be sufficient for the calculation of the transmission function $T_{pq}(E)$ up to a maximum energy $E_{\max}$ to take into account only those lead modes in leads p and q that are propagating for energies below $E_{\max}$. In most cases, this approximation allows one to grossly reduce the numerical effort in calculating the inverse of A_C without a significant change in the resulting transmission function. This is indeed a crucial factor that enables one to calculate the ballistic quantum transport also for realistic three-dimensional devices with large contact areas.

5.3 Novel method for ballistic quantum transport in entangled two-particle systems

Two-particle Green's function

In the following, we extend the CBR method to the case of an open device that describes a system of two interacting, distinguishable particles. We restrict ourselves to the case that only one of the two particles contributes to the transport and the other particle is bound. We further assume that the interaction is negligible outside the device. The total Hamiltonian of the system is then identical to Eq. (5.3) except for H^0, which now becomes a two-particle Hamiltonian. H^0 is composed of the single-particle Hamiltonians H_1^0 and H_2, and the interaction term V_{12}

$$H^0 = H_1^0 + H_2 + V_{12}. \tag{5.20}$$

H_1^0 as well as H_2 are Hamiltonians of the decoupled device, but only the particle described by H_1^0 is eventually coupled to the leads. The coupling of H_1^0 to the lead Hamiltonians is treated in the same way as for Eq. (5.3) yielding an effective (non-hermitian) finite dimensional two-particle Hamiltonian

$$H = H_1^0 + \Sigma + H_2 + V_{12} = H^0 + \Sigma. \tag{5.21}$$

Here again, we first determine the two-particle Green's function G^0 of the decoupled device by means of its spectral representation Equation (5.7). Then we calculate the two-particle Green's function G^R of the open system according to the Dyson equation. In the following, we show that the same decomposition of the matrices Σ, G^0, and G^R is applicable as in the case of the CBR method for a single particle. First, we note that the self-energy Σ is given by the following tensor product

$$\Sigma = \Sigma_1 \otimes \mathbf{1}_2 \tag{5.22}$$

where Σ_1 is the total self-energy of the first particle. The reason for this representation of the self-energy Σ is that only the particle described by H_1^0 is coupled to the leads. The two particles are distinguishable, which in this case means that they are localized in separate parts of space. Thus we divide the total set of grid point Ω into two subsets

$$\Omega = \Omega_1 \cup \Omega_2 \tag{5.23}$$

where Ω_1 is the subset of grid points for particle 1, Ω_2 is the subset of grid points for particle 2. The total number of grid points is again denoted by N_T, the number of grid points in subsets Ω_1 and Ω_2 are denoted by N_1 and N_2, respectively. Further, we decompose Ω_1 into N_D interior grid points Ω_D and N_C contact grid points Ω_C as in the previous section so that we can represent Σ_1 as follows

$$\Sigma_1 = \begin{pmatrix} \Sigma_{1,C} & 0 \\ 0 & 0_{1,D} \end{pmatrix} \tag{5.24}$$

and Σ can be written according to

$$\Sigma = \begin{pmatrix} \Sigma_C & 0 \\ 0 & 0_D \end{pmatrix} \tag{5.25}$$

with $\Sigma_C = \Sigma_{1,C} \otimes \mathbf{1}_2$ and $0_D = 0_{1,D} \otimes \mathbf{1}_2$. Thus in Eq. (5.25) Σ is formally identical to the expression given in Eq. (5.9). Note, however, that Σ_C is a square matrix with dimension

$N_C \times N_2$. In close analogy to the previous section, we decompose the two-particle Green's function of the decoupled device

$$G^0 = \begin{pmatrix} G^0_C & G^0_{CD} \\ G^0_{DC} & G^0_D \end{pmatrix}, \tag{5.26}$$

define the matrix A by $A = \mathbf{1} - G^0\Sigma$ so that the retarded two-particle Green's function is given by $G^R = A^{-1}G^0$, and apply the decomposition to A

$$A = \begin{pmatrix} A_C & 0 \\ -G^0_{DC}\Sigma_C & \mathbf{1}_D \end{pmatrix}, \tag{5.27}$$

$$A_C = \mathbf{1}_C - G^0_C\Sigma_C. \tag{5.28}$$

From the previous section it is now obvious that to obtain the retarded two-particle Green's function G^R according to the Dyson equation it is sufficient to calculate the inverse of A_C. The retarded two-particle Green's function is then given by

$$G^R = \begin{pmatrix} A_C^{-1}G^0_C & A_C^{-1}G^0_{CD} \\ G^0_{DC}\Sigma_C A_C^{-1}G^0_C + G^0_{DC} & G^0_{DC}\Sigma_C A_C^{-1}G^0_{CD} + G^0_D \end{pmatrix}. \tag{5.29}$$

Thus, by means of the presented method, the numerical effort for calculating G^R is reduced from the effort of inverting a square matrix of size $N_1 \times N_2$ to the effort of inverting a square matrix of size $N_C \times N_2$.

The two-particle Green's function G^0 of the decoupled device is determined by means of its spectral representation Equation (5.7). Here we make use of the same approximation as described in the previous section and take into account only a reduced set of eigenstates. This set of eigenstates is determined as follows: First, a certain number of the lowest eigenstates of the closed system's single particle Hamiltonians H_1^0 and H_2 are calculated. Then, the interaction V_{12} is diagonalized in the product basis of these single-particle eigenstates. Having at hand single-particle eigenstates of H_1^0 and H_2, it is possible to represent the matrices Σ, G^0, and G^R in a mixed basis of position eigenstates of particle 1 and the subset of N_2^{eigen} lowest energy eigenstates $|\alpha_2\rangle$ of particle 2 ($H_2 |\alpha_2\rangle = \varepsilon_{a_2} |\alpha_2\rangle$)

$$\Sigma(i, j, \alpha_2, \alpha_2') = \langle i| \langle \alpha_2| \Sigma |_j\rangle |\alpha_2'\rangle, \tag{5.30}$$

$$G^0(i, j, \alpha_2, \alpha_2') = \langle i| \langle \alpha_2| G^0 |_j\rangle |\alpha_2'\rangle, \tag{5.31}$$

$$G^R(i, j, \alpha_2, \alpha_2') = \langle i| \langle \alpha_2| G^R |_j\rangle |\alpha_2'\rangle, \tag{5.32}$$

where $i, j = 1, \ldots, N_1$ and $\alpha_2, \alpha_2' = 1, \ldots, N_2^{\text{eigen}}$. The decomposition into the contact block and the block of interior grid points affects the indices i and j only. Thus, from the change of basis for the second particle from position eigenstates to the reduced set of energy eigenstates, it follows that the square matrix A_C is reduced in size from $N_C \times N_2$ to only $N_C \times N_2^{\text{eigen}}$. Note that a direct diagonalization of the two-particle Hamiltonian H^0 requires diagonalizing a square matrix of size $N_1 \times N_2$. For three spatial dimensions, where typically $N_1, N_2 \approx 10^6$, this is far beyond any imaginable possibilities of today's and also future computers The storage of only the diagonal of the matrix of H^0 would require 10 Tb of memory.

A further reduction in the size of the square matrix A_C is possible by means of the reduction of the number of lead modes as described in the previous section. The total number of lead modes N_{modes} can usually be chosen to be much smaller than the number of contact grid points N_C. The size of the square matrix A_C is then given by $N_{\text{modes}} \times N_2^{\text{eigen}}$.

Determination of the transmission function

In the following, we represent G_C^R in the mixed basis of position eigenstates of particle 1 and energy eigenstates of particle 2

$$G_C^R(i, j, \alpha_2, \alpha_2') = \langle i| \langle \alpha_2| G_C^R |j\rangle |\alpha_2'\rangle , \qquad i, j \in \Omega_C, \quad \alpha_2, \alpha_2' = 1, \ldots, N_2^{\text{eigen}}. \tag{5.33}$$

In this form, G_C^R represents the probability amplitude for the propagation of particle 1 from position x_j to position x_i accompanied by a transition of particle 2 from eigenstate $|\alpha_2'\rangle$ to $|\alpha_1'\rangle$. For $p \neq q$, the expression

$$T_{pq}^{\alpha_2,\alpha_2'}(E) = \text{Tr}_1\, \Gamma_C^p G_C^R \Gamma_C^q G_C^{R\dagger} \tag{5.34}$$

therefore yields the probability for the transmission of particle 1 from lead q to lead p under the condition that initially particle 2 is in state $|\alpha_2'\rangle$ and ends up in state $|\alpha_2\rangle$. For $p = q$ the corresponding expression is given by

$$T_{pq}^{\alpha_2,\alpha_2'}(E) = \text{Tr}_1 \left(\mathbf{1}_C - i\Gamma_C^p G_C^R\right) \left(\mathbf{1}_C - i\Gamma_C^q G_C^R\right)^\dagger . \tag{5.35}$$

Note that the trace in Eq. (5.34) and Eq. (5.35) is taken only with respect to the position of particle 1. If the final state of particle 2 is not determined, the corresponding transmission probability is obtained by summing over all final states of particle 2

$$T_{pq}^{\alpha_2'} = \sum_{\alpha_2} T_{pq}^{\alpha_2,\alpha_2'}. \tag{5.36}$$

5.4 Interaction matrix elements

The two lowest lying quantum states of a tunneling coupled single electron double quantum dot can be described by a very simple model Hamiltonian that takes into account only the energy splitting Δ between the ground states of the isolated quantum dots and the tunneling coupling t between the two quantum dots. This Hamiltonian is given by

$$H^D = \begin{pmatrix} -\frac{\Delta}{2} & -\frac{t}{2} \\ -\frac{t}{2} & \frac{\Delta}{2} \end{pmatrix}. \tag{5.37}$$

Typical values for Δ and t are of the order of 10 μeV [92]. In our two-particle calculations, we use this model Hamiltonian for the subsystem of the second particle instead of a realistic three-dimensional model because of the following reasons. First, the properties of the double dot system can be characterized more easily in terms of only two parameters. Second, the calculation of the interaction matrix elements between the two particles is numerically less demanding. However, we want to emphasize that our method for the calculation of the ballistic current for an interacting system of two distinguishable particles is not in principle limited to treating the second particle by a model Hamiltonian.

The eigenstates of H are linear combinations of the ground states $|0\rangle$ and $|1\rangle$ of the two isolated quantum dots

$$|B\rangle = h_{B0} |0\rangle + h_{B1} |1\rangle , \tag{5.38}$$

$$|A\rangle = h_{A0} |0\rangle + h_{A1} |1\rangle , \tag{5.39}$$

with real valued coefficients h_{Xk} ($X = A, B$, $k = 0, 1$). Here, $|B\rangle$ denotes the bonding eigenstate and $|A\rangle$ denotes the anti-bonding eigenstate. The corresponding eigenenergies are

$$E_B = -\frac{1}{2}\sqrt{t^2 + \Delta^2}, \tag{5.40}$$

$$E_A = \frac{1}{2}\sqrt{t^2 + \Delta^2}. \tag{5.41}$$

The coefficients h_{Xk} can be conveniently written in terms of E_X ($X = A, B$)

$$h_{X0} = \frac{\Delta - 2E_X}{\sqrt{t^2 + (\Delta - 2E_X)^2}}, \tag{5.42}$$

$$h_{X1} = \frac{t}{\sqrt{t^2 + (\Delta - 2E_X)^2}}. \tag{5.43}$$

For vanishing Δ, $|B\rangle$ and $|A\rangle$ become $(|0\rangle + |1\rangle)/\sqrt{2}$ and $(-|0\rangle + |1\rangle)/\sqrt{2}$, respectively. For vanishing t, $|B\rangle$ and $|A\rangle$ reduce to $|0\rangle$ and $|1\rangle$. We assume that the charge distribution of the electron in the ground states of the isolated quantum dots can be approximated by point charges at positions $\mathbf{q}_0$ and $\mathbf{q}_1$. With the Coulomb interaction

$$V_{12} = \frac{e^2}{4\pi\varepsilon_r\varepsilon_0}\frac{1}{|\mathbf{x}_1 - \mathbf{x}_2|} \tag{5.44}$$

as the interaction term between particle 1 at position $\mathbf{x}_1$ and particle 2 at position $\mathbf{x}_2$, we find the following expression for the interaction matrix elements between the single-particle eigenstates $|\alpha_1\rangle$, $|\alpha_1'\rangle$ ($H_1^0|\alpha_1\rangle = \varepsilon_{\alpha_1}|\alpha_1\rangle$) of particle 1 and single-particle eigenstates $|X\rangle$, $|Y\rangle$ ($X, Y = A, B$) of particle 2

$$\begin{aligned}(V_{12})_{\alpha_1\alpha_1' XY} &= \langle\alpha_1|\langle X| V_{12} |\alpha_1'\rangle |Y\rangle \\ &= \frac{e^2}{4\pi\varepsilon_r\varepsilon_0}\int \mathrm{d}^3x\, \psi_{\alpha_1}(\mathbf{x})\,\psi_{\alpha_1'}(\mathbf{x}) \sum_{k=0,1}\frac{h_{Xk}h_{Yk}}{|\mathbf{x} - \mathbf{q}_k|},\end{aligned} \tag{5.45}$$

where $\psi_{\alpha_1}(\mathbf{x})$ denotes the eigenstate $|\alpha_1\rangle$ in the position representation.

5.5 Summary

In this chapter, we have developed a novel method for the ballistic quantum transport in entangled two-particle systems. The method allows for the non-perturbative treatment of the Coulomb interaction between two distinguishable electrons. Therefore, many-particle correlation effects such as entanglement can be studied. The method is based on the contact block reduction (CBR) method. Approximations such as the mode space reduction and the incomplete spectral representation of the closed system Green's function that render the CBR method to be highly efficient are still valid in our novel method. Thus, our novel method is also highly efficient in terms of computational effort and it therefore allows us to study the ballistic quantum transport in entangled two-particle systems also for realistic, three-dimensional devices.

Chapter 6

Prediction of entanglement detection by I-V characteristics

6.1 Introduction

The idea to build quantum computers based on semiconductor nanostructures is at least as appealing as it is demanding. In contrast to most other proposals [93], semiconductor based concepts for quantum information processing promise a high degree of scalability because of the highly developed technology and great experience in the fabrication of semiconductor devices. However, the implementation of only a single qubit proves to be more difficult in semiconductors than it is for most other approaches. The reasons for this are the short decoherence times and strong interactions between elementary excitations in solids [94]. Already in 2001, a proof of principle of Shor's factoring algorithm was given by means of a 7-qubit NMR quantum computer [3]. But only recently the first single- and two-qubit quantum operations in semiconductors have been demonstrated with spin qubits [4, 5].

At liquid helium temperatures and below, electrons in a two-dimensional electron gas (2DEG) formed at the interface of a GaAs/AlGaAs semiconductor heterostructure show mean free paths as well as phase-relaxation lengths of the order of 30–40 μm [21]. Metal gates on the surface of such heterostructures can be used to define regions, where the motion of the electrons is confined to zero (quantum dots) or one dimension (quantum wires). Bertoni *et al.* [95] and Ionicioiu *et al.* [96] proposed a scheme for quantum computation in semiconductors that exploits the long mean free path and phase-relaxation length for electrons propagating through such quantum wires (QWRs). In their approach, a single-electron wave packet propagates through either of two parallel quantum wires. One of the quantum wires represents the state $|0\rangle$, the other quantum wire represents the state $|1\rangle$. A single-qubit rotation gate can be realized by a small opening that couples the two quantum wires. Such coupling windows between two quantum wires have been studied theoretically as well as experimentally by several authors [78, 95, 97–100]. The size of such a gate is one to two orders of magnitude smaller than the mean free path and the phase-relaxation length. This approach should therefore allow applying even multiple quantum gate operations to a qubit, despite the generally short decoherence times in solids. Except for Ref. [78] the theoretical analyses are limited to simple non-selfconsistent models. In Ref. [78] a self-consistent model is used, but the calculation is done only quasi three-dimensional.

Universal quantum computation demands not only for single-qubit gates, but also for two-qubit gates [101]. Two-qubit gates in turn are intrinsically tied to entanglement. A possible

way to entangle two quantum wire qubits is to exploit the Coulomb interaction. For this, two pairs of quantum wires are brought close to one another so that the electrons in the quantum wires can interact in a controlled way. However, the main disadvantage is that the injection of the electrons into the quantum wires has to be timed very accurately on a timescale of less than a picosecond. The possibilities to entangle two qubits as well as to perform two-qubit gates based on this idea has been studied theoretically in several articles [102–105] by means of simple models. Akguc *et al.* [98] and Reichl and Snyder [106] proposed to consider stationary scattering states of quantum wire qubits rather than wave packets and developed several analytical models to describe the entanglement of the qubits. They also studied complex networks and considered generalizations of the concept of stationary quantum wire qubits to higher dimensional spaces [107].

This stationary picture is also the basis of the approach that we employ here. Our objective however is the analysis of realistic, fully three-dimensional models of quantum wire based single- and two-qubit quantum gates that result in conceptually simple and experimentally realizable proposals of novel semiconductor devices for quantum information processing. Here we present a theoretical analysis of a ballistic GaAs/AlGaAs quantum device for the controlled generation and detection of entanglement between an all-electric Mach-Zehnder interferometer and an electrostatically defined single electron quantum dot. The Mach-Zehnder interferometer is realized by two electrostatically defined QWRs that are connected by two coupling windows. In contrast to electronic Mach-Zehnder interferometers based on quantum Hall edge channels [108], no magnetic fields are employed. We calculate the ballistic I-V characteristics of the fully three-dimensional, open device using a single band effective mass description for the electronic Hamiltonian including the Hartree potential self-consistently. The interferometer can be employed as a fully controllable single-qubit gate for quantum wire qubits.

To calculate the entanglement between the interferometer and the double quantum dot, we have developed the Green's function method presented in the preceding chapter that allows us to calculate the ballistic current of the entangled system for a realistic, three-dimensional device structure including the Coulomb interaction non-perturbatively. In addition, we have also developed a dynamic, analytical model of the device. This allows us to show that the entanglement can be determined by measuring the DC I-V characteristics of the interferometer and that the device can be employed as an externally controllable two-qubit gate.

This chapter is organized as follows: In Sec. 6.2 we first address the topic of single-qubit gates and present a concrete model of the electronic Mach-Zehnder interferometer in terms of its geometry, material composition, doping profiles, and applied bias voltages. In the process of working out this model, we have always kept in mind present-day technological possibilities. We show the operation of the Mach-Zehnder interferometer as a single-qubit gate for propagating electrons by numerical calculations of the DC I-V characteristics of the device. Further, we discuss the components of the interferometer that are most critical in view of fabrication precision. In Sec. 6.3 we turn to two-qubit gates and present our novel quantum transport device that allows for the generation and detection of entanglement between a quantum wire qubit and a double quantum dot qubit. We show the results of our numerical calculations of the DC I-V characteristics for the proposed device. We discuss that the degree of entanglement can be controlled by external gates and can be determined in a straightforward manner from the DC I-V characteristics. In Sec. 6.4 we present the analytical model of the device and derive the relation between the I-V characteristics and the degree of entanglement that we use for the interpretation of the results in Section 6.3. The chapter is concluded by a summary.

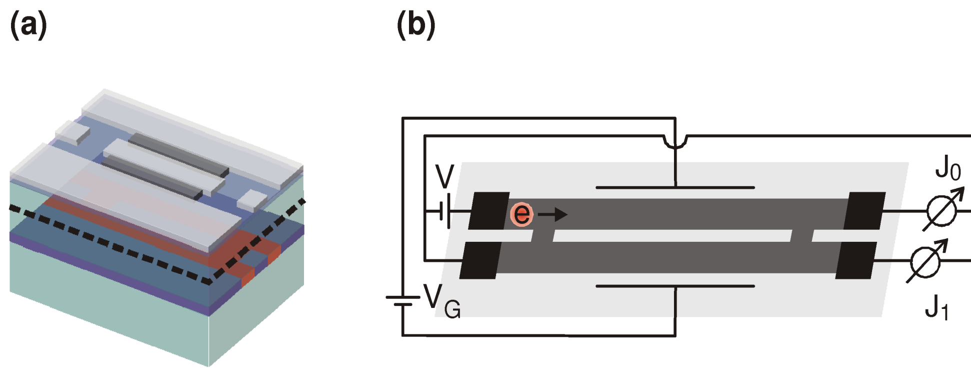

Figure 6.1: Schematic views of the proposed all-electric Mach-Zehnder interferometer. (a) The 2DEG in an AlGaAs/GaAs heterostructure (green/blue) is locally depleted by external metal gates (gray) to form two parallel quantum wires (red) that are coupled by two potential wells in the separating barrier. (b) A small bias voltage V is applied to the upper left terminal. The resulting currents J_0 and J_1 that flow to the upper and lower right terminals are measured as a function of the phase gate voltage V_G. For clarity, both figures are not drawn to scale.

6.2 All-electric Mach-Zehnder interferometer

6.2.1 Introduction

In optics, Mach-Zehnder interferometers are used to determine phase shifts as well as for the modulation of light intensities. An incoming beam of light is split by a beam-splitter. To one of the resulting beams, a phase shift is applied and at a second beam-splitter the two beams interfere. The intensities of the two beams then depend on the applied phase shift. The same interference experiment can also be done with electrons. However, because of the fermionic character of electrons in contrast to photons, an electron can interfere only with itself. A quantum optical formulation of the Mach-Zehnder interferometer in terms of photonic creation and annihilation operators that also discusses the case of particles with fermionic character is given in Reference [109].

Here we study the theory of semiconductor based electronic Mach-Zehnder interferometers and their application as single-qubit gates. A schematic view of such a Mach-Zehnder interferometer is shown in Figure 6.1 (a). The device is based on the two-dimensional electron gas (2DEG) that forms in the quantum well of a modulation doped AlGaAs/GaAs heterostructure. The 2DEG is locally depleted by the application of negative voltages to metal gates on top of the heterostructure. By this, two parallel quantum wires are formed in the 2DEG. These quantum wires are connected by two coupling windows that are realized as potential wells in the barrier separating the quantum wires. For specific energies, the coupling windows between the quantum wires act as beam-splitters and the quantum wire network behaves as an electronic Mach-Zehnder interferometer. In contrast to electronic Mach-Zehnder interferometers based on quantum Hall edge channels [108], no magnetic fields are employed. We therefore use the term *all-electric Mach-Zehnder interferometer*. In particular, this also means that the phase shift of the electron wave function between the two quantum wires is controlled electrostatically by *phase gates*. These are additional gates on top of the quantum wires, but can also be combined with the gates used to define the quantum wires.

The Mach-Zehnder interferometer is operated under ballistic transport conditions. By ap-

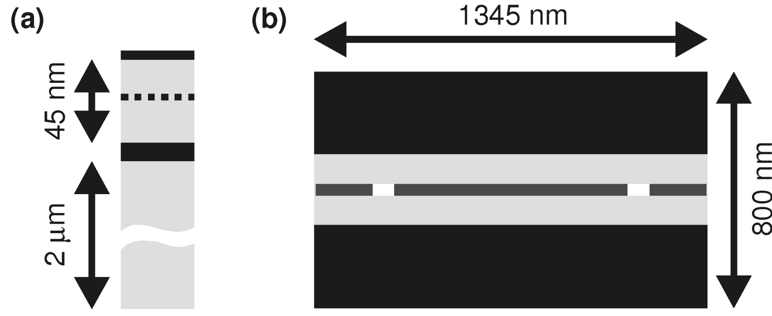

Figure 6.2: (a) GaAs/AlGaAs heterostructure with a single 10 nm wide GaAs quantum well as used in the calculation of the Mach-Zehnder interferometer. GaAs is shown in black, $Al_{0.37}Ga_{0.63}As$ is shown in gray. The GaAs cap layer is 5 nm thick. A silicon δ doping layer with a concentration of $2.5 \times 10^{12}\,\mathrm{cm}^{-2}$ is located 25 nm below the surface (dashed line). (b) Top view of the Mach-Zehnder interferometer showing the side gates (black) and the mid gates (dark gray).

plying a small DC bias voltage V as shown in Fig. 6.1 (b) between the upper left and the remaining three contacts, current flows predominantly from the upper left to the upper right contact (J_0) and from the upper left to the lower right contact (J_1). The ratio of J_0 and J_1 depends primarily on the voltage V_G that is applied to the phase gate.

6.2.2 Numerical Details

For the discretization of the all-electric Mach-Zehnder interferometer we have used a nonuniform tensor product grid with a total of 6×10^6 grid points. The Schrödinger equation has been solved on a subset of 1.5×10^6 grid points. For a converged result of the equilibrium charge density and the electrostatic potential, it was sufficient to determine the 250 eigenstates that are lowest in energy. From these 250 eigenstates, the occupation of the ones highest in energy is smaller than the occupation of those lowest in energy by eight orders of magnitude. The converged result was obtained by a total number of 12 predictor-corrector steps. For the calculation of the transmission probabilities we have taken into account the same 250 eigenstates to determine the retarded Green's function of the closed device. Further, we have taken into account the lowest 5 transversal modes in each of the four leads.

6.2.3 Results

We now turn to the results for the self-consistent calculation of the I-V characteristics of the all-electric Mach-Zehnder interferometer. The calculations in this section are based on a GaAs/AlGaAs heterostructure with a single 2DEG. This heterostructure is depicted in Figure 6.2 (a). It consists of a 5 nm thick GaAs cap layer, 45 nm $Al_{0.37}Ga_{0.63}As$, and 10 nm GaAs (2DEG layer) followed by 2 μm $Al_{0.37}Ga_{0.63}As$ (substrate). A silicon δ doping layer with a concentration of $2.5 \times 10^{12}\,\mathrm{cm}^{-2}$ is located 25 nm below the surface. The electron sheet density in the 2DEG layer has been calculated to be $2.4 \times 10^{11}\,\mathrm{cm}^{-2}$ for the unstructured sample at a temperature of 4 K.

Figure 6.2 (b) shows a top view of the Mach-Zehnder interferometer with the layout of the metal gates. The device is 800 nm wide and 1345 nm long. Here we distinguish two types

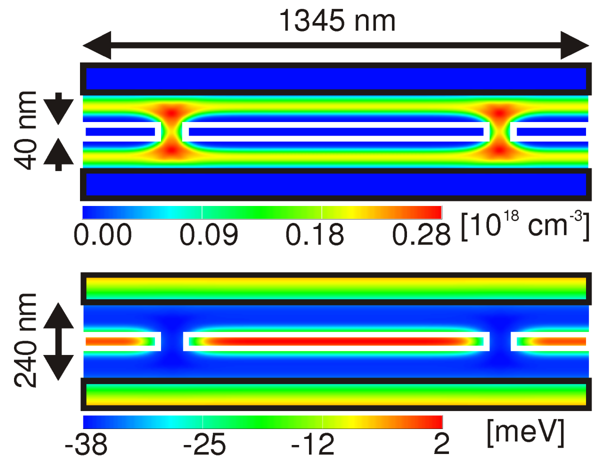

Figure 6.3: Top views of the equilibrium electron density (upper panel) and electrostatic potential (lower panel) within the 2DEG. The rectangles indicate the position of the metal gates on the surface.

of gates: side gates (black) and mid gates (dark gray). The two side gates are 280 nm wide, 1345 nm long, and biased at -0.245 V. The three mid gates are 40 nm wide and 200 nm, 800 nm, and 200 nm long, respectively. To the mid gates a voltage of -0.660 V is applied. The gates define two quantum wires (light gray) with a nominal width of 100 nm as well as two coupling windows (white) with a nominal length of 72.5 nm. The subband spacing of the two lowest subbands in each of these quantum wires is 3.1 meV. The Fermi wave length in the lowest subband amounts to $\lambda_F = 77$ nm and thus is of the same order as the nominal length of the coupling windows. To control the phase shift of the electron wave function between the two quantum wires either a small additional bias voltage V_G is applied with opposite sign to each of the two side gates, or additional gates are placed on top of the quantum wires and $\pm V_G$ is applied to these gates. The first design that combines the phase gate and the side gates requires applying larger additional bias voltages to achieve the phase shift. For the specific device that we have studied, these larger bias voltages result in a contribution of the second subband as the Fermi energy is only slightly (~ 80 μeV) below the second subband for $V_G = 0$. By the contribution of the second subband the interference is suppressed. In the following, we therefore focus on the design that employs additional gates. If however the subband spacing can be increased, the first design is to be preferred because of its greater conceptual simplicity.

Figure 6.3 shows the equilibrium charge density (upper panel) and the corresponding potential (lower panel) in the 2DEG layer of the Mach-Zehnder interferometer. The black and the white rectangles indicate the position of the metal gates on the surface of the heterostructure. The upper panel shows that the 2DEG is fully depleted underneath the gates and by this, the quantum wires and the coupling windows are formed. The quantum wires are real one-dimensional, i.e. only the first transversal quantum wire mode contributes to the density. The lower panel shows the potential barrier that separates the two quantum wires. The energy scale is chosen so that the Fermi energy coincides with zero.

Figure 6.4 (a) shows the currents J_0 (solid), J_1 (dashed), and also the detrimental current J_2 that flows from the upper left terminal to the lower left terminal (dotted). The latter is a measure of back-scattering. The currents have been calculated by assuming a DC bias voltage of 50 μV and a temperature of 30 mK. The figure shows the dependence of the currents on the

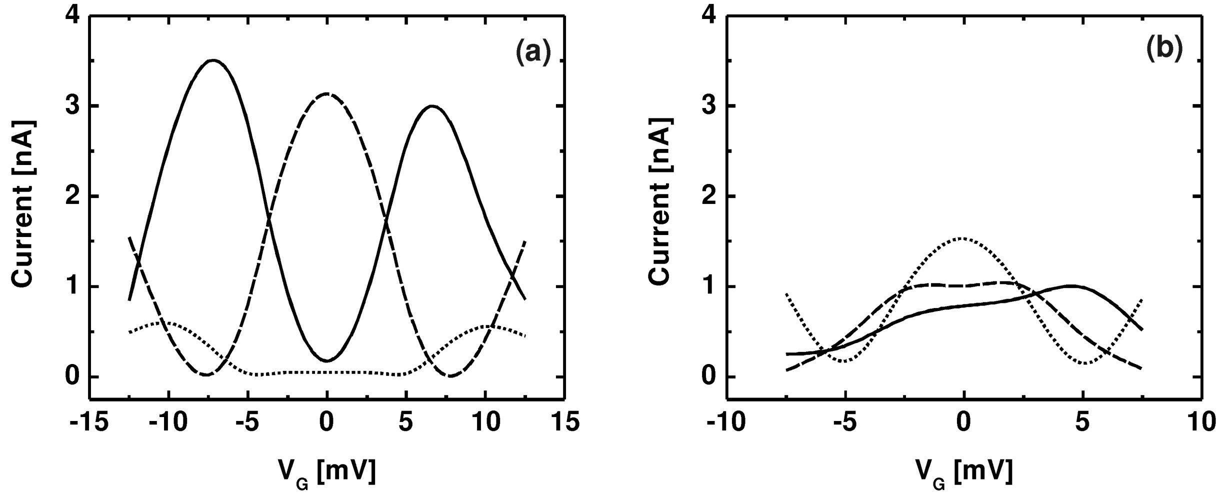

Figure 6.4: Stationary currents J_0 (solid) and J_1 (dashed) through the Mach-Zehnder interferometer as a function of the phase gate voltage V_G for two different nominal lengths of the coupling windows. (a) For a coupling window length of 72.5 nm the currents show an interference pattern with almost 100% visibility. (b) If the length of the coupling windows is reduced by only 7%, the interference pattern is already noticeably disturbed. Also shown is the current (dotted) that flows from the upper left terminal to the lower left terminal and is due to reflections.

voltage V_{G} applied to the phase gate. We first note that the detrimental current J_2 is smaller than 0.5 nA for the presented range of gate voltages. This shows that back-scattering from the upper left to the lower left terminal does not play a dominant role here. In contrast to J_2, the currents J_0 and J_1 show a strong dependence on the gate voltage. In essence, over the presented range of gate voltages, J_0 and J_1 oscillate in between 0 nA and 4 nA for more than 1.5 periods. For zero gate voltage, J_0 is minimal and J_1 is maximal, thus the two currents are phase-shifted relative to one another by π. This result confirms that the device actually behaves as a Mach-Zehnder interferometer. In the following, we will therefore denote the dependence of the currents J_0 and J_1 according to Fig. 6.4 (a) as Mach-Zehnder interference pattern. Note that the interference pattern is damped out if the absolute gate voltage is too large. The reason for this is that by increasing the gate voltage we also increase the amount of back-scattering. This can be avoided by increasing the length of the device instead of increasing the gate voltage, because the phase shift depends linearly on the length of the phase gate, whereas back-scattering is independent of the length but depends on the absolute value of the gate voltage.

For comparison, we show in Fig. 6.4 (b) the currents J_0, J_1, and J_2 for the same device but with the nominal length of the coupling windows reduced by 7%. Even this small change in the nominal length of the coupling windows disturbs the interference pattern. The strong dependence of the interference pattern on the size of the coupling windows therefore requires a very high fabrication precision.

Figure 6.5 shows the charge densities of the stationary current carrying states of the Mach-Zehnder interferometer for selected phase gate voltages. The three figures 6.5 (a), (b), and (c) correspond to gate voltages of 0 mV, −3.6 mV, and −7.5 mV, respectively. The charge densities illustrate the quantum mechanical nature of the interference pattern. The three graphs show that the charge densities change predominantly at the right terminals. This change reflects

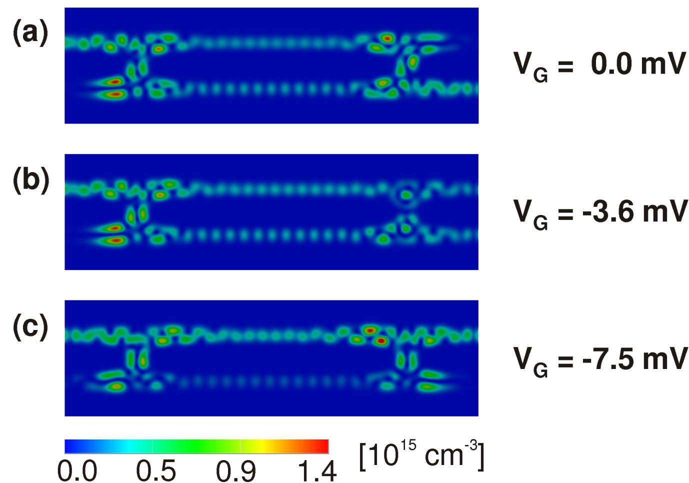

Figure 6.5: Charge densities of the current carrying scattering state for three different gate voltages. Current flows from the upper left terminal to (a) the lower right terminal, (b) both right terminals, (c) the upper right terminal.

the redistribution of the total current between J_0 and J_1. In Fig. 6.5 (a) the charge density is zero at the upper right terminal and finite at the lower right terminal. From Fig. 6.4 we see that for the corresponding gate voltage of $0\,\mathrm{mV}$ the current J_0 is minimal and the current J_1 is maximal. Similarly, in Fig. 6.5 (b) the charge density is equal at both right terminals, and also the currents J_0 and J_1 are equal for a gate voltage of $-3.6\,\mathrm{mV}$. Eventually, in Fig. 6.5 (c) the charge density is finite at the upper right terminal and vanishes at the lower right terminal. Correspondingly, the current J_0 is maximal and the current J_1 is zero for a gate voltage of $-7.5\,\mathrm{mV}$.

The change of the charge densities to the right of the second beam splitter is a consequence of the interference of the electron wave function. This interference depends on the relative phase that the wave function acquires on its way through the upper quantum wire as compared to the lower quantum wire. Depending on this relative phase, the interference at the second beam splitter leads to a (partial) extinction either at the upper right terminal or at the lower right terminal.

6.2.4 Discussion

The DC I-V characteristics shown in Fig. 6.4 (a) exhibits multiple pronounced switches. These can be attributed to rotations of a qubit on the Bloch sphere. According to reference [107] one can define the basis states $|0\rangle$ and $|1\rangle$ of a qubit by an electron that propagates through one of two quantum wires. In particular, in the stationary picture, we can relate the stationary scattering states shown in Fig. 6.5 to specific transformations of an initial state $|\phi_i\rangle$ into a final state $|\phi_f\rangle$. The initial state is characterized here by an electron that enters the device at the upper left terminal, which we denote as $|0\rangle$. On the right hand side of the first beam splitter, the electron is in an intermediate state that is a linear combination of $|0\rangle$ and $|1\rangle$ and is given approximately by $\left(|0\rangle + \exp\left(i\phi_l\right)|1\rangle\right)/\sqrt{2}$ with an unknown phase factor $\exp\left(i\phi_l\right)$ and up to an overall phase factor. On the left hand side of the second beam splitter, the electron is still in the same intermediate state $\left(|0\rangle + \exp\left(i\phi_r\left(V_\mathrm{G}\right)\right)|1\rangle\right)/\sqrt{2}$, but with a different phase factor

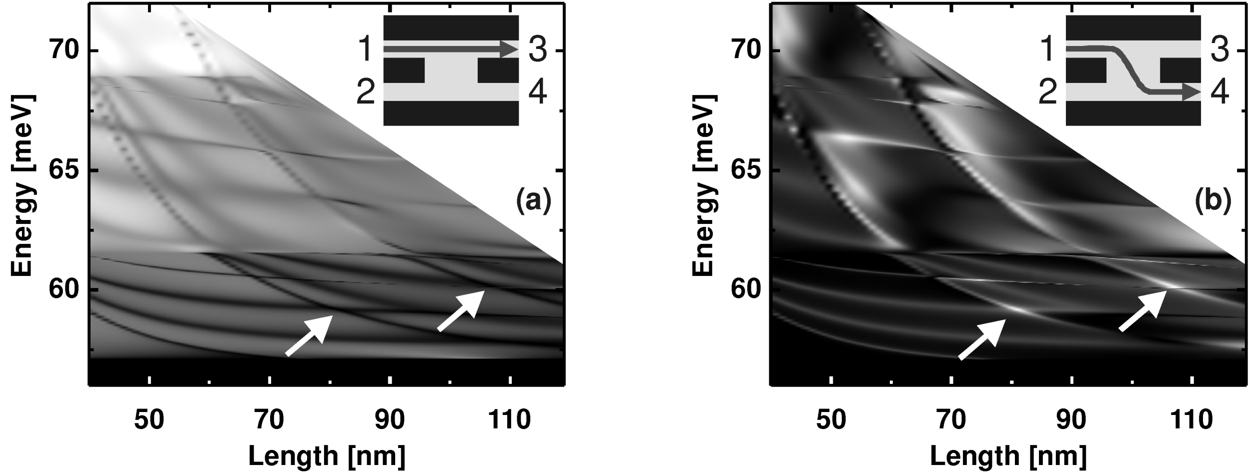

Figure 6.6: Probability for an electron entering the H shaped quantum wire device at terminal 1 to be transmitted to (a) terminal 3, and (b) terminal 4 as a function of the length l of the coupling window and the kinetic energy of the electron for fixed width w. The resonances can be related to the energies of the lowest quantum states in a box of size $l \times w$. The arrows point to crossings of resonances belonging to different transversal quantum numbers. At these crossings the transmission probability $1 \rightarrow 3$ vanishes and the transmission probability $1 \rightarrow 4$ is close to unity.

$\exp(i\phi_r(V_{\mathrm{G}}))$ that also depends on the phase gate voltage V_{G}. On the right hand side of the second beam splitter, the electron leaves the device in a state that is a linear combination of $|0\rangle$ and $|1\rangle$, where the admixture of the two states depends strongly on the phase gate voltage. For the three graphs shown in Fig. 6.5 these final states $|\phi_f\rangle$ are approximately given (up to an overall phase factor) by $|1\rangle$, $(|0\rangle + \exp(i\phi)|1\rangle)/\sqrt{2}$, and $|0\rangle$ for (a), (b), and (c), respectively, where again $\exp(i\phi)$ is an unknown phase factor.

We now want to address the question concerning the strong dependence of the interference pattern on the length of the coupling windows. This can be illustrated by means of an H shaped structure as the one depicted in the insets of Figure 6.6 (a) and (b). We have defined this structure within a 10 nm thick slab of GaAs by hard wall potentials. The structure is characterized by three parameters: the width of the quantum wires w_{QW}, the width of the coupling window w and the length of the coupling window l. For the following calculations, we have chosen $w_{\mathrm{QW}} = 60\,\mathrm{nm}$ and $w = 100\,\mathrm{nm}$. Figure 6.6 (a) shows the probability for an electron that enters the H shaped structure at terminal 1 to be transmitted to terminal 3 (we denote this transmission probability by T_{13}) as a function of the coupling window length l and the kinetic energy of the electron. Correspondingly, Fig. 6.6 (b) shows the transmission probability from terminal 1 to terminal 4, denoted as T_{14}. The energy range covers the first three modes of the quantum wires. Without the coupling window, the transmission T_{13} would therefore increase from 0 to 3 in steps of unit magnitude. However, the coupling between the two quantum wires causes T_{13} to be resonantly reduced for specific energies that depend on the length of the coupling window. At the same time T_{14} is enhanced as can be seen from Figure 6.6 (b). In the case of this hard wall potential model, the energies of these resonances are approximately equal to the energies of the bound states in a rectangular box with the

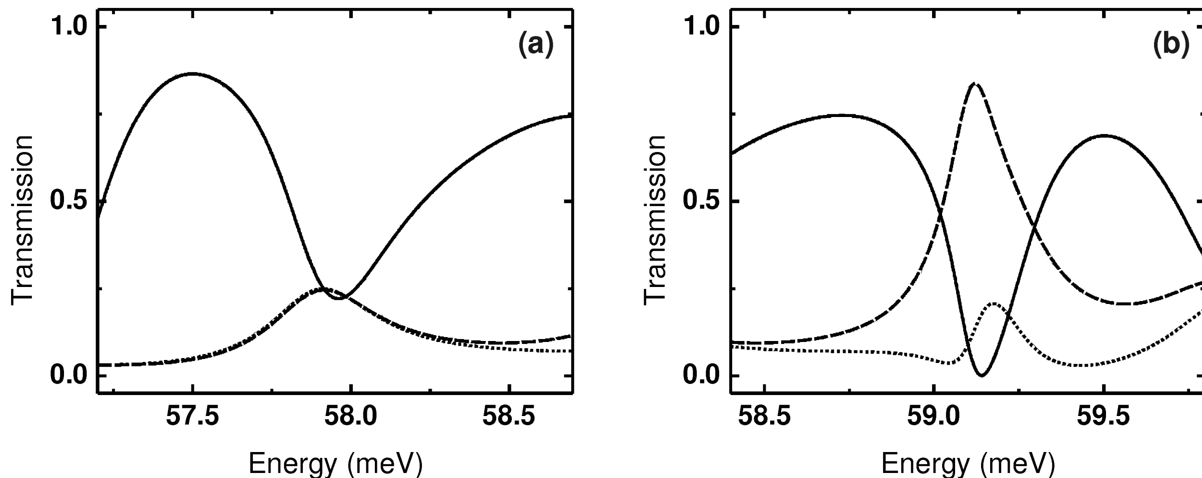

Figure 6.7: Transmission probabilities for an electron entering the H shaped quantum wire device at terminal 1 to be transmitted to terminal 3 (solid), terminal 4 (dashed), and terminal 2 (dotted) as a function of the kinetic energy of the electron for a length l of the coupling window equal to 82 nm.

dimensions of the coupling window. Thus, resonances occur if an integer multiple n of half the wave length of the electron matches the length of the coupling window. For the first mode, the transmission T_{13} drops from 1 to 0.25 at such a resonance. The transmission probabilities T_{11}, T_{12}, and T_{14} from terminal 1 to terminals 1, 2, and 4 increase from 0 to 0.25, i.e. the electron is transmitted to each of the four terminals with equal probability. In Fig. 6.7 (a) the transmission probabilities T_{12} (dotted curve), T_{13} (solid curve), and T_{14} (dashed curve) are shown for an example of such a resonance. These resonances have a typical width of only a few 100 μeV and occur for any length of the coupling window at multiple energies. However, for a beam-splitter we require that $T_{13} \approx T_{14} \approx 0.5$ and $T_{11} \approx T_{12} \approx 0$.

The dependence of the resonance energy on the length of the coupling window is determined by the number n of half wavelengths that fit through the coupling window. Therefore, resonances that belong to different n become degenerate for specific lengths. In the graphs shown in Fig. 6.6, two such crossings have been emphasized by arrows. The first one occurs at a length of 82 nm. For this coupling window length, Fig. 6.7 (b) shows the transmission probabilities T_{13} (solid curve), T_{14} (dashed curve), and T_{12} (dotted curve) for a small energy range around the energy of the crossing. At the energy of the crossing, the transmission probability T_{13} vanishes whereas the transmission probability T_{14} approaches unity. The transmission probability T_{12} is small. It follows that for a slightly smaller as well as for a slightly larger energy the transmission probabilities T_{13} and T_{14} are identical, and in addition they are indeed approximately equal to 0.5. This is exactly the condition that is required for a beam-splitter. If the length of the coupling window is changed by only 5 nm, the energy difference of the two contributing resonances is already larger than 100 μeV, which is sufficient to destroy the beam-splitting condition. The same argument also holds for the width of the coupling window. The strong dependence of the interference pattern on the size of the coupling window is therefore a consequence of the fact that the beam-splitting condition depends on the energies of two resonances.

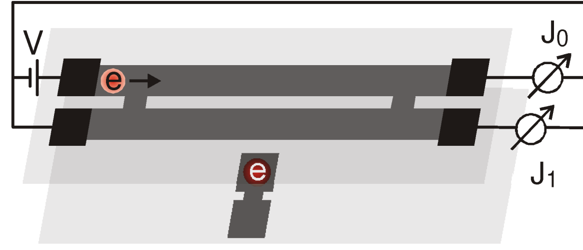

Figure 6.8: Schematic view of the proposed quantum transport device. The device is realized by two stacked GaAs/AlGaAs 2DEGs. The top 2DEG is depleted by external gates to form a Mach-Zehnder interferometer. In the bottom 2DEG two coupled quantum dots are located. For sake of clarity, the figure is not drawn to scale.

6.3 Entangled Mach-Zehnder double quantum dot device

6.3.1 Introduction

The single-qubit gate based on the all-electric Mach-Zehnder interferometer presented in the preceding section is the basis of our novel two-qubit quantum transport device, which we will discuss in the following. We have calculated the DC I-V characteristics through the device by means of our method for ballistic quantum transport in entangled two-particle systems that we have presented in Chapter 5. As we will show, the proposed device allows for the controlled generation and measurement of entanglement between a quantum wire qubit and a double quantum dot qubit. The measurement involves only the DC I-V characteristics, no higher order current correlations are required to detect entanglement.

A schematic view of the proposed device is shown in Figure 6.8. The device is based on a GaAs/AlGaAs heterostructure that consists of two vertically stacked two-dimensional electron gases (2DEGs). The 2DEGs are locally depleted by the application of negative voltages to metal gates on the surface of the heterostructure. Thus an all-electric Mach-Zehnder interferometer is formed in the upper 2DEG and a pair of tunneling coupled single electron quantum dots is formed in the lower 2DEG.

Electrons in the Mach-Zehnder interferometer and in the double quantum dot interact by means of the Coulomb interaction. In general, the interaction between two quantum mechanical systems leads to an entanglement of these systems: none of the systems can be separately represented by pure quantum states. We therefore refer to the proposed device as *Entangled Mach-Zehnder Double Quantum Dot Device*. The maximum degree of entanglement depends on the geometric arrangement of the Mach-Zehnder interferometer and the double quantum dot. By means of our studies we could determine an optimum geometric arrangement that promises a particularly high degree of entanglement.

6.3.2 Numerical Details

For the discretization of the entangled Mach-Zehnder double quantum dot device, we have used a grid with 8×10^4 grid points. For the determination of the retarded Green's function of the closed device we have taken into account the lowest 100 eigenstates and we have taken into

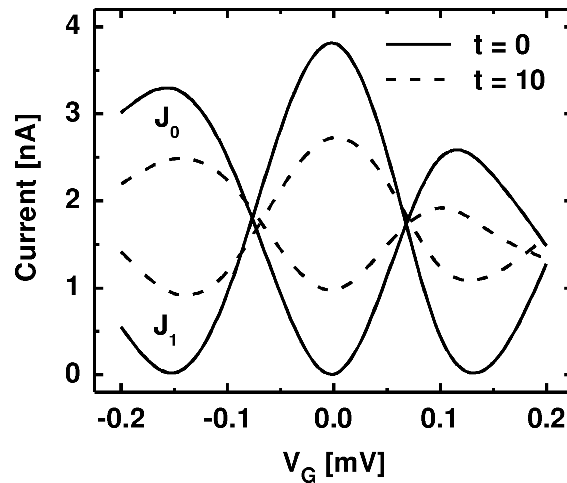

Figure 6.9: Currents J_0 and J_1 as a function of the phase gate voltage for two different tunneling couplings t in μeV. This figure shows that the visibility of the interference pattern depends on the tunneling coupling.

account the lowest 5 transversal modes in each of the four leads.

6.3.3 Results

In the following we present the results for the entangled Mach-Zehnder double quantum dot device. In these calculations the electron charge density and the corresponding potential have not been calculated self-consistently. Instead, we have defined the Mach-Zehnder interferometer by a hard-wall potential within a 10 nm thick slab of GaAs. The two quantum wires are 55 nm wide, 1000 nm long, and the lateral distance between them amounts to 20 nm. The coupling windows have a length of 85 nm. The energy spacing of the two lowest subbands is 5.3 meV, and the Fermi energy has been set to 1.6 meV in the lowest subband. For electrons with this kinetic energy, the coupling windows act as almost perfect beam-splitters. For the calculation of the interaction between the Mach-Zehnder interferometer and the double quantum dot we assume a vertical distance of 80 nm between the Mach-Zehnder interferometer and the quantum dots. One of the quantum dots is located exactly underneath the center of the Mach-Zehnder interferometer. The lateral distance between the two quantum dots is chosen to be 60 nm. The double quantum dot system is modeled according to the Hamiltonian given by Equation (5.37). The relative geometric arrangement of the Mach-Zehnder interferometer and the double quantum dot has a large influence on the operation of the device. The configuration that we have chosen here exploits the idea that the interaction due to an electron in the quantum dot underneath the center of the Mach-Zehnder interferometer is identical for both quantum wires. Thus, the phase shifts induced by an electron in this quantum dot are also identical for both quantum wires and therefore cancel. This is equivalent to the case that the electron in this quantum dot does not interact with the electrons in the quantum wires at all. Therefore, the effect of the interaction between an electron in the other quantum dot and the electrons in the quantum wires is the more pronounced. We have confirmed the assumption that this geometric arrangement is optimal for the operation of the device by approximating the phase shifts within a WKB model as well as by the presented numerical calculations.

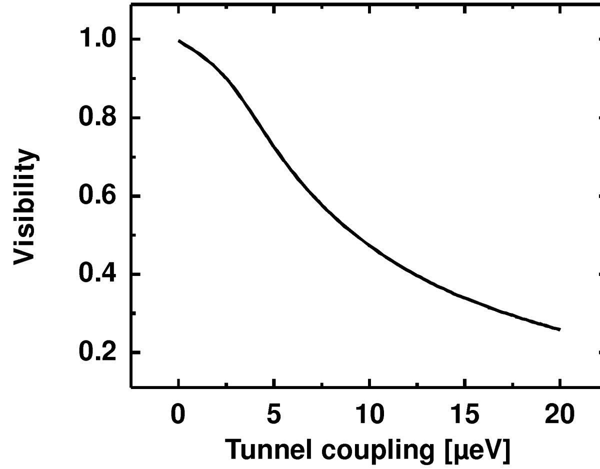

Figure 6.10: Visibility of the interference pattern as a function of the tunneling coupling. The visibility decreases monotonously with increasing tunneling coupling.

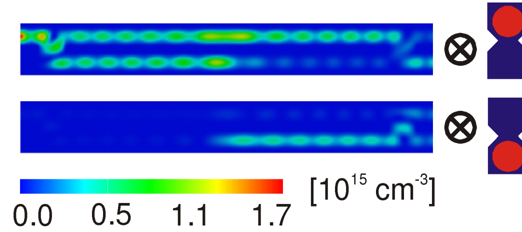

Figure 6.11: Charge densities of the current carrying states in the two basis states of the quantum dot qubit for a finite tunneling coupling of $t = 10$ μeV. The upper panel shows the charge density for the ground state, the lower panel shows the charge density for the excited state of the quantum dot qubit.

At first, we study the I-V characteristics of the device for a situation where tunneling between the quantum dots is inhibited ($t = 0$). For the ground state splitting we have chosen a value of $\Delta = 10$ μeV. The result is shown in Fig. 6.9 by the solid curves. We obtain the typical interference pattern that we have also found in the previous self-consistent calculations for the all-electric Mach-Zehnder interferometer. The pronounced asymmetry with respect to the sign of the gate voltage is caused by the asymmetric position of the phase gate along only one of the two quantum wires in combination with the additional repulsive potential due to the electron in the double quantum dot.

We now allow the electron in the double quantum dot to tunnel between the two quantum dots ($t = 10$ μeV). The currents J_0 and J_1 for this situation are shown in Fig. 6.9 by the dashed curves. We still obtain the Mach-Zehnder interference pattern, but the visibility is reduced. For simplicity we define the visibility here by $v = (J_1 - J_0) / (J_1 + J_0)|_{V_G=0}$, i.e. for zero phase gate voltage V_G. Figure 6.10 shows the visibility as a function of the tunneling coupling. The visibility is almost 1 for vanishing tunneling coupling and decreases monotonously with increasing tunneling coupling. The visibility is therefore uniquely related to the tunneling coupling.

The suppression of the visibility for non-vanishing tunneling coupling can be visualized by

means of the charge densities of the current carrying states. For a tunneling coupling $t = 10\ \mu$eV these are shown in Fig. 6.11 separately for the two basis states of the double quantum dot qubit. The upper panel corresponds to the ground state, the lower panel corresponds to the excited state. The figure shows that initially (i.e. to the left of the first beam splitter) the quantum dot qubit is in the ground state. Also immediately to the right of the first beam splitter, the quantum dot qubit is still in the ground state. However, by means of the Coulomb interaction, the probability of the quantum dot qubit to be in the ground state is lowered towards the second beam splitter, mainly for the lower quantum wire. At the same time, the probability of the quantum dot qubit to be in the excited state is increased. As a consequence of this probability redistribution, the interference at the second beam splitter is modified so that to the right of the second beam splitter the charge density is finite in the upper as well as in the lower quantum wire in both states of the quantum dot qubit. The total charge density is the sum of the charge densities for the two states of the quantum dot qubit. Therefore, the total charge density is also finite in both quantum wires to the right of the second beam splitter. This in turn reflects the fact that both currents, J_0 and J_1, are finite and thus that the visibility is suppressed.

6.3.4 Discussion

The suppression of the visibility can be understood as the result of a two-qubit quantum operation. We can think of the entangled Mach-Zehnder double quantum dot device as a two-qubit quantum gate that is based on the Coulomb interaction between the electrons in the pair of quantum wires (qubit W) and in the double quantum dot (qubit D). Following reference [107] the basis states $|0\rangle^{\mathrm{W}}$ and $|1\rangle^{\mathrm{W}}$ of qubit W are defined by an electron that propagates through either of the two quantum wires. The basis states $|0\rangle^{\mathrm{D}}$ and $|1\rangle^{\mathrm{D}}$ of qubit D are defined by the ground states of the isolated quantum dots. On the basis of this picture we have developed a simple dynamic, analytical model of the device. The details of this model are presented in the following section and in Appendix F. The model captures the basic physics of the device and reproduces Fig. 6.9 qualitatively for the transmission probabilities $T_{|0\rangle}$ and $T_{|1\rangle}$ given by Equation (6.20). The reason for the suppression of the visibility is the entanglement of the two qubits. From the point of view of each of the two qubits, the entanglement with the other qubit is nothing else than decoherence. This decoherence disturbs the interference between the electron wave function in the two quantum wires. In other words, the reduced density matrix of each of the qubits that is obtained by tracing out the degrees of freedom of the other qubit represents a mixed state if and only if the two qubits are entangled.

In the context of the analytical model, we have quantified this relation and we have derived an analytical expression for the relation between the visibility and the von Neumann entropy, which is a measure for the degree of entanglement. According to this expression, the von Neumann entropy is given by

$$S = -\sum_{i=+,-} p_i \log p_i, \tag{6.1}$$

where $p_\pm$ are the eigenvalues of the reduced density matrix of qubit W. These eigenvalues depend on the visibility as follows:

$$p_\pm = \frac{1}{2}\left(1 \pm v\right). \tag{6.2}$$

By means of these equations the von Neumann entropy depends monotonously on the visibility. A visibility of $v = 1$ corresponds to vanishing von Neumann entropy (zero degree of entangle-

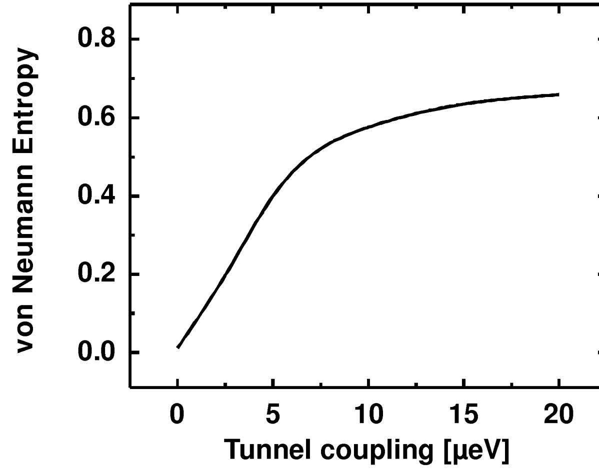

Figure 6.12: The figure shows the von Neumann entropy as a function of the tunneling coupling. The von Neumann entropy is a measure of the degree of entanglement and is uniquely related to the tunneling coupling.

ment), for a visibility of $v = 0$ we find a von Neumann entropy of $S = 1$ (maximum degree of entanglement). We have used this relation to determine the von Neumann entropy as a function of the tunneling coupling for the entangled Mach-Zehnder double quantum dot device. The result is presented in Fig. 6.12 and shows that the von Neumann entropy is uniquely related to the tunneling coupling. Thus we conclude that first, the degree of entanglement can be determined simply by measuring the DC I-V characteristics of the device. Second, the device offers a simple way for external control over the degree of entanglement as the tunneling coupling can be easily tuned by external gates.

Entanglement is essential for any kind of non-trivial two-qubit operation. The controlled generation of entanglement between the quantum wire qubit and the double quantum dot qubit therefore demonstrates that the entangled Mach-Zehnder double quantum dot device can be considered as a two-qubit quantum gate.

6.4 Analytical model for the entangled Mach-Zehnder double quantum dot device

6.4.1 Introduction

While detailed numerical calculations as the ones presented in the last two sections are extremely helpful in working out concrete proposals for novel devices, analytical models often have the advantage that the underlying physics can be illustrated more clearly. In the interpretation of the results of the previous section, we have therefore also made use of such an analytical model, where the device has been reduced to two interacting qubit systems. In the following, we present this model in detail and we demonstrate that qualitatively we can obtain the same results as in the last section. This shows on the one hand that the model indeed captures the physics of the device, and on the other hand, it validates our numerical results. On the basis of the model we further derive an analytical relation between the visibility of the interference pattern and the von Neumann entropy.

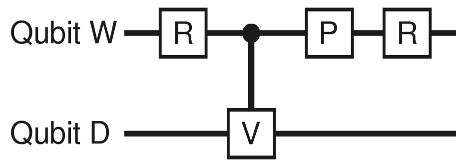

Figure 6.13: Schematic of the analytical model of the entangled Mach-Zehnder double quantum dot device. The device is represented by the two qubits W and D and a series of one- ($\hat{R}$ and $\hat{P}$) and two-qubit gates ($\hat{V}$).

6.4.2 Analytical model

The entangled Mach-Zehnder double quantum dot device can be considered as a two-qubit quantum gate based on the Coulomb interaction between a pair of quantum wires (qubit W) and a single electron double quantum dot (qubit D). The basis states $|0\rangle^{\mathrm{W}}$ and $|1\rangle^{\mathrm{W}}$ of qubit W are defined by an electron that propagates through either of the two quantum wires. For the double quantum dot qubit, we neglect excited states and define the basis states $|0\rangle^{\mathrm{D}}$ and $|1\rangle^{\mathrm{D}}$ by the ground states of the two isolated quantum dots.

Quantum gates can be described by unitary operators that map an initial state $|\phi_i\rangle$ onto a corresponding final state $|\phi_f\rangle$. The unitary operator $\hat{T}$ that represents our two-qubit quantum transport device can be decomposed into a sequence of unitary operators of less complex single- and two-qubit gates as follows:

$$\hat{T} = \hat{R} \cdot \hat{P} \cdot \hat{V} \cdot \hat{R}. \tag{6.3}$$

A schematic of this description is shown in Figure 6.13. The operators $\hat{R}$ and $\hat{P}$ represent single-qubit gates that act on qubit W. In detail, $\hat{R}$ represents a $\pi/2$ rotation gate and $\hat{P}$ represents a phase gate with variable phase angle. The sequence $\hat{R} \cdot \hat{P} \cdot \hat{R}$ corresponds to a Mach-Zehnder interferometer. The operator $\hat{V}$ represents a two-qubit gate and describes the coupling of the two qubits due to the Coulomb interaction. The advantage of this decomposition is that we do not have to develop a consistent description for the whole device. Instead, we can specify the action of each of the components individually. In particular, this allows us to specify the single-qubit gates $\hat{R}$ and $\hat{P}$ explicitly and consequently to focus on a dynamic model for the two-qubit gate $\hat{V}$.

To this end, we represent the two electrons in the pair of quantum wires and the double quantum dot by the following basis states

$$\psi_{k,i,j}(x) = \exp(ikx)\,|i\rangle^{\mathrm{W}}\,|j\rangle^{\mathrm{D}}, \qquad i,j = 0,1. \tag{6.4}$$

These states describe the two electrons as follows: The first electron propagates along the x axis in either of two states $|0\rangle^{\mathrm{W}}$ and $|1\rangle^{\mathrm{W}}$, corresponding to the two quantum wires. The second electron occupies one of the two quantum dots, represented by states $|0\rangle^{\mathrm{D}}$ and $|1\rangle^{\mathrm{D}}$. The two-particle Hamiltonian that describes the free dynamics of the system is then given by

$$H_0 = p^2/2m \otimes \mathbf{1}^{\mathrm{W}}_{2\times 2} \otimes \mathbf{1}^{\mathrm{D}}_{2\times 2} + \mathbf{1}_{L_2(\mathbf{R})} \otimes \mathbf{1}^{\mathrm{W}}_{2x2} \otimes H^{\mathrm{D}}, \tag{6.5}$$

where H^{D} is the model Hamiltonian that according to Eq. (5.37) describes a pair of tunneling coupled quantum dots in terms of the bare splitting Δ and the tunneling coupling t. We do not take into account the Coulomb interaction between the two electrons exactly, but make the following ansatz

$$V(x) = U\,|1\rangle^{\mathrm{W}}\,|1\rangle^{\mathrm{D}}\,\langle 1|^{\mathrm{W}}\,\langle 1|^{\mathrm{D}}\,\theta(x)\,\theta(L-x). \tag{6.6}$$

This describes a localized interaction with interaction strength U that is restricted to the two-particle state $|1\rangle^{\mathrm{W}}|1\rangle^{\mathrm{D}}$. The interaction only acts if the position of the quantum wire electron is in between $x=0$ and $x=L$. The resulting quantum mechanical scattering problem $H_0+V(x)$ can be solved analytically for the 8×8 transfer matrix $T_{\hat{V}}$. It is convenient to first diagonalize H_0 and use the eigenstates of H_0 as asymptotic states. Since $V(x)$ only acts on the state $|1\rangle^{\mathrm{W}}|1\rangle^{\mathrm{D}}$ the only non-trivial part of $T_{\hat{V}}$ is the 4×4 submatrix $\boldsymbol{\tau}$ that connects the asymptotic states

$$\psi_{1,2}(x)=\exp(\pm ikx)\,|1\rangle^{\mathrm{W}}|B\rangle^{\mathrm{D}}, \tag{6.7}$$

$$\psi_{3,4}(x)=\exp(\pm ikx)\,|1\rangle^{\mathrm{W}}|A\rangle^{\mathrm{D}} \tag{6.8}$$

for $x\to-\infty$ with their analogs for $x\to\infty$. The states $|B\rangle^{\mathrm{D}}$ and $|A\rangle^{\mathrm{D}}$ are defined as in Eq. (5.38) and Eq. (5.39). The explicit expression for the submatrix $\boldsymbol{\tau}$ is given by $\boldsymbol{\tau}=\boldsymbol{\mu}_{L,2}^{-1}\boldsymbol{\mu}_{L,1}\boldsymbol{\mu}_{0,2}^{-1}\boldsymbol{\mu}_{0,1}$ where for $x=0,L$

$$\boldsymbol{\mu}_{x,1}=\begin{pmatrix}\nu_x(\bar{k}_B) & 0\\ 0 & \nu_x(\bar{k}_A)\end{pmatrix}, \tag{6.9}$$

$$\boldsymbol{\mu}_{x,2}=\begin{pmatrix}\alpha_B\nu_x(\bar{q}_B) & \alpha_A\nu_x(\bar{q}_A)\\ \beta_B\nu_x(\bar{q}_B) & \beta_A\nu_x(\bar{q}_A)\end{pmatrix}. \tag{6.10}$$

The 2×2 matrix $\nu_x(k)$ is defined as follows:

$$\nu_x(k)=\begin{pmatrix}\exp(ikx) & \exp(-ikx)\\ ik\exp(ikx) & -ik\exp(-ikx)\end{pmatrix}. \tag{6.11}$$

The coefficients α_X, β_X and the wave vectors $\bar{k}_X$ and $\bar{q}_X$ $(X=A,B)$ are given by $(t,U>0)$

$$\alpha_X=\frac{tU}{\sqrt{\left(4E_X(E_X+\varepsilon)+\Delta U\right)^2+t^2U^2}}, \tag{6.12}$$

$$\beta_X=\frac{4E_X(E_X+\varepsilon)+\Delta U}{\sqrt{\left(4E_X(E_X+\varepsilon)+\Delta U\right)^2+t^2U^2}}, \tag{6.13}$$

$$\bar{k}_X=\frac{1}{\hbar}\sqrt{2m(E-E_X)}, \tag{6.14}$$

$$\bar{q}_X=\frac{1}{\hbar}\sqrt{2m\left(E-\frac{1}{2}U\mp\varepsilon\right)}, \tag{6.15}$$

where $\varepsilon=\frac{1}{2}\sqrt{t^2+(U+\Delta)^2}$. The details of the derivation of the transfer matrices are given in Appendix F.

Within this model we include the rotation and the phase gate by explicitly specifying the 8×8 transfer matrices $T_{\hat{R}}$ and $T_{\hat{P}}$ for these gates. The transfer matrix T of the complete device is then given by the product of the individual transfer matrices according to Eq. (6.3), i.e.

$$T=T_{\hat{R}}\cdot T_{\hat{P}}\cdot T_{\hat{V}}\cdot T_{\hat{R}}. \tag{6.16}$$

We obtain the transfer matrices $T_{\hat{R}}$ and $T_{\hat{P}}$ from the following representations of ideal rotation and phase gates

$$\hat{R}=\begin{pmatrix}\cos(\theta) & -\sin(\theta)\\ \sin(\theta) & \cos(\theta)\end{pmatrix}\otimes\mathbf{1}_{2\times 2}^{\mathrm{D}},\qquad \theta=\pi/2, \tag{6.17}$$

$$\hat{P} = \begin{pmatrix} e^{i\phi} & 0 \\ 0 & e^{-i\phi} \end{pmatrix} \otimes \mathbf{1}^{\mathrm{D}}_{2\times 2}. \tag{6.18}$$

Further, we assume that these gates do not introduce back-scattering, and accordingly we apply $\hat{R}$ and $\hat{P}$ separately to the asymptotic states for both directions of propagation.

The transfer matrix T maps the asymptotic states for $x \to -\infty$ onto the asymptotic states for $x \to \infty$. By rearranging and partial inversion of the transfer matrix we construct the scattering matrix S. The scattering matrix relates the initial state $|\phi_i\rangle$ to the final state $|\phi_f\rangle$

$$|\phi_f\rangle = S\,|\phi_i\rangle\,. \tag{6.19}$$

The initial state is a linear combination of the in-going asymptotic states, whereas the final state is a linear combination of the outgoing asymptotic states. We consider the case that the propagating electron enters the pair of quantum wires from the left in state $|0\rangle^{\mathrm{W}}$. We further assume that the electron in the double quantum dot is initially in the bonding state $|B\rangle^{\mathrm{D}}$. The eight final state coefficients represent the transmission amplitudes into the four states $|i\rangle^{\mathrm{W}}|X\rangle^{\mathrm{D}}$ ($i = 0, 1,\ X = A, B$) for each of the two asymptotic propagation directions (either to the right or back to the left). We denote these final state coefficients by $E^{\mathrm{r}}_{i,X}$ and $E^{\mathrm{l}}_{i,X}$, respectively. By construction, the two rotation gates and the phase gate do not introduce any back-scattering. The amount of back-scattering due to the interaction $V(x)$ depends on the ratio of the interaction strength U and the length L of the interaction region and can be made arbitrarily small. Thus, the transmission amplitudes $E^{\mathrm{l}}_{i,X}$ are negligible. We can therefore calculate the transmission probabilities $T_{|0\rangle}$ and $T_{|1\rangle}$ that the propagating electron is transmitted either into state $|0\rangle^{\mathrm{W}}$ or into state $|1\rangle^{\mathrm{W}}$, i.e. irrespective of the final state of the electron in the double quantum dot, from the transmission probabilities $E^{\mathrm{r}}_{i,X}$

$$T_{|i\rangle} = \left|E^{\mathrm{r}}_{i,A}\right|^2 + \left|E^{\mathrm{r}}_{i,B}\right|^2, \qquad i = 0, 1. \tag{6.20}$$

6.4.3 Relation between visibility and von Neumann entropy

In the following we derive the relation between the visibility of the interference pattern and the von Neumann entropy. Further details of the derivation are presented in Appendix F. For this, we assume that back-scattering is negligible. This allows us to describe the device by Eq. (6.3) instead of by the scattering matrix S. The density matrix of the final state is then given by

$$\rho = \begin{pmatrix} E_{0,0}E^*_{0,0} & E_{0,0}E^*_{0,1} & E_{0,0}E^*_{1,0} & E_{0,0}E^*_{1,1} \\ E_{0,1}E^*_{0,0} & E_{0,1}E^*_{0,1} & E_{0,1}E^*_{1,0} & E_{0,1}E^*_{1,1} \\ E_{1,0}E^*_{0,0} & E_{1,0}E^*_{0,1} & E_{1,0}E^*_{1,0} & E_{1,0}E^*_{1,1} \\ E_{1,1}E^*_{0,0} & E_{1,1}E^*_{0,1} & E_{1,1}E^*_{1,0} & E_{1,1}E^*_{1,1} \end{pmatrix}. \tag{6.21}$$

Here, the coefficients $E_{i,j}$ denote the final state probability amplitudes of the four states $|i\rangle^{\mathrm{W}}|j\rangle^{\mathrm{D}}$ ($i, j = 0, 1$). The von Neumann entropy

$$S = -\operatorname{Tr}\rho_{\mathrm{red}}\log\rho_{\mathrm{red}} \tag{6.22}$$

is a measure of the degree of entanglement of the two particles in the final state. The reduced density matrix ρ_{red} is obtained from ρ by tracing out the degrees of freedom of one of the two particles. Here we choose to trace out the electron in the double quantum dot. For the reduced density matrix we obtain

$$\rho_{\mathrm{red}} = \begin{pmatrix} T_{|0\rangle} & Z \\ Z^* & T_{|1\rangle} \end{pmatrix}, \tag{6.23}$$

where the coherence Z is given by

$$Z = E_{0,0}E_{1,0}^* + E_{0,1}E_{1,1}^*. \tag{6.24}$$

The von Neumann entropy can be easily calculated in terms of the eigenvalues

$$p_\pm = \frac{1}{2}\left(1 \pm \sqrt{4\left(|Z|^2 - T_{|0\rangle}T_{|1\rangle}\right) + 1}\right) \tag{6.25}$$

of the reduced density matrix as follows:

$$S = -\sum_{i=+,-} p_i \log p_i. \tag{6.26}$$

We now show that the eigenvalues $p_\pm$ can be expressed in terms of the visibility v of the interference pattern according to

$$p_\pm = \frac{1}{2}(1 \pm v). \tag{6.27}$$

First, we determine expressions for the transmission probabilities as a function of the angle of the phase gate. For this, we apply the $\pi/2$ rotation gate $\hat{R}$ and the two-qubit gate $\hat{V}$ to the initial state $|\phi_i\rangle$. The resulting state is completely determined by the probability amplitudes $A_{i,j}$ of the four two-particle basis states $|i\rangle^{\mathrm{W}} |j\rangle^{\mathrm{D}}$ $(i, j = 0, 1)$. Application of the phase gate $\hat{P}$ and the second $\pi/2$ rotation gate $\hat{R}$ yields final state probability amplitudes

$$E_{0,j} = \frac{1}{\sqrt{2}}\left(\exp(i\phi) A_{0,j} + A_{1,j}\right), \tag{6.28}$$

$$E_{1,j} = \frac{1}{\sqrt{2}}\left(-\exp(i\phi) A_{0,j} + A_{1,j}\right). \tag{6.29}$$

Note that these gates, being single-qubit transformations, do not change the degree of entanglement of the two electrons. For the transmission probabilities $T_{|0\rangle}$ and $T_{|1\rangle}$ we obtain

$$T_{|j\rangle} = \frac{1}{2} \pm |A_{0,0}|\,|A_{1,0}| \cos(\phi + \delta_0) \pm |A_{0,1}|\,|A_{1,1}| \cos(\phi + \delta_1), \tag{6.30}$$

where here and in the following $+$ and $-$ correspond to 0 and 1, respectively. The most general ansatz for the amplitudes $A_{i,j}$ that results from the initial state by the application of the gates $\hat{R}$ and $\hat{V}$ leads to

$$|A_{i,0}|^2 = \frac{1}{2}\cos^2\left(\frac{1}{2}(\gamma \pm \sigma)\right), \tag{6.31}$$

$$|A_{i,1}|^2 = \frac{1}{2}\sin^2\left(\frac{1}{2}(\gamma \pm \sigma)\right). \tag{6.32}$$

This ansatz takes into account that on the one hand the probabilities for the electron in the quantum wire to be in either state $|0\rangle^{\mathrm{W}}$ or state $|1\rangle^{\mathrm{W}}$ must be equal due to the first $\pi/2$ rotation gate $\hat{R}$. On the other hand the electron in the double quantum dot is allowed to be in arbitrary linear combinations of the states $|0\rangle^{\mathrm{D}}$ and $|1\rangle^{\mathrm{D}}$ that can also be different for states $|0\rangle^{\mathrm{W}}$ and $|1\rangle^{\mathrm{W}}$. These linear combinations are described by the mixing angles γ and σ. If we insert the expressions for $A_{i,j}$ into Eq. (6.30) we obtain for the transmission probabilities

$$T_{|j\rangle} = \frac{1}{2}\left(1 \pm \sqrt{\cos^2(\gamma)\sin^2\left(\bar{\delta}\right) + \cos^2(\sigma)\cos^2\left(\bar{\delta}\right)}\cos\left(\bar{\phi}\right)\right), \tag{6.33}$$

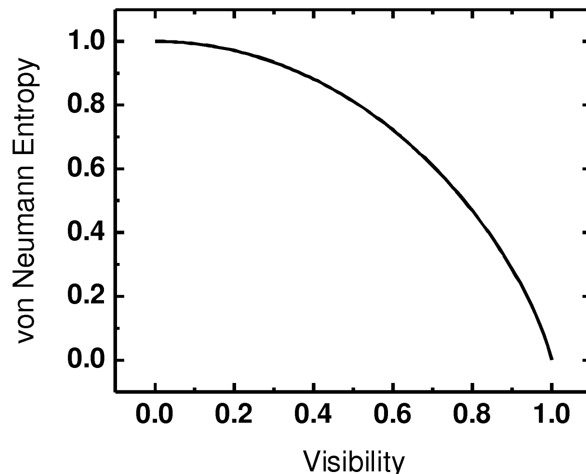

Figure 6.14: The von Neumann entropy is shown as a function of the visibility according to Eq. (6.27) and Equation (6.26). The figure shows that the von Neumann entropy decreases strictly monotonously with increasing visibility.

where $\bar{\delta} = \frac{1}{2}(\delta_0 - \delta_1)$ and $\bar{\phi} = \phi + \text{const.}$ From these expressions we read off that the transmission probabilities $T_{|0\rangle}$ and $T_{|1\rangle}$ show a sinusoidal oscillation as a function of the phase angle $\bar{\phi}$, and that the visibility of this oscillation is given by

$$v = \sqrt{\cos^2(\gamma)\sin^2\left(\bar{\delta}\right) + \cos^2(\sigma)\cos^2\left(\bar{\delta}\right)}. \tag{6.34}$$

If we now evaluate the expression $|Z|^2 - T_{|0\rangle}T_{|1\rangle}$ in an analogous way, we obtain

$$|Z|^2 - T_{|0\rangle}T_{|1\rangle} = \frac{1}{4}\left(\cos^2(\gamma)\sin^2\left(\bar{\delta}\right) + \cos^2(\sigma)\cos^2\left(\bar{\delta}\right) - 1\right) \tag{6.35}$$

$$= \frac{1}{4}\left(v^2 - 1\right). \tag{6.36}$$

Insertion of this result into Eq. (6.25) yields Equation (6.27).

In Fig. 6.14 we have plotted the dependence of the von Neumann entropy on the visibility resulting from Eq. (6.27) and Equation (6.26). The figure shows that the von Neumann entropy depends uniquely on the visibility. The von Neumann entropy decreases strictly monotonously from 1 to 0 with increasing visibility. Note that a von Neumann entropy of 1 corresponds to a maximal degree of entanglement, whereas a von Neumann entropy of 0 corresponds to zero entanglement. Thus, a maximal degree of entanglement is characterized by vanishing visibility, and correspondingly zero entanglement is characterized by full visibility.

6.4.4 Results and Discussion

In the following we show results obtained by the evaluation of Eq. (6.20). These results illustrate that the proposed analytical model captures the physics of the entangled Mach-Zehnder double quantum dot device. The parameters have been scaled with respect to the characteristic energy scale $E_{\text{char}} = \hbar^2/(2mL^2)$ as introduced in Appendix F. For the graphs in Fig. 6.15 we have chosen a total energy $E' = 1000$. The coupling strength and the ground state splitting have been set to $U' = 300$ and $\Delta' = 10$, respectively.

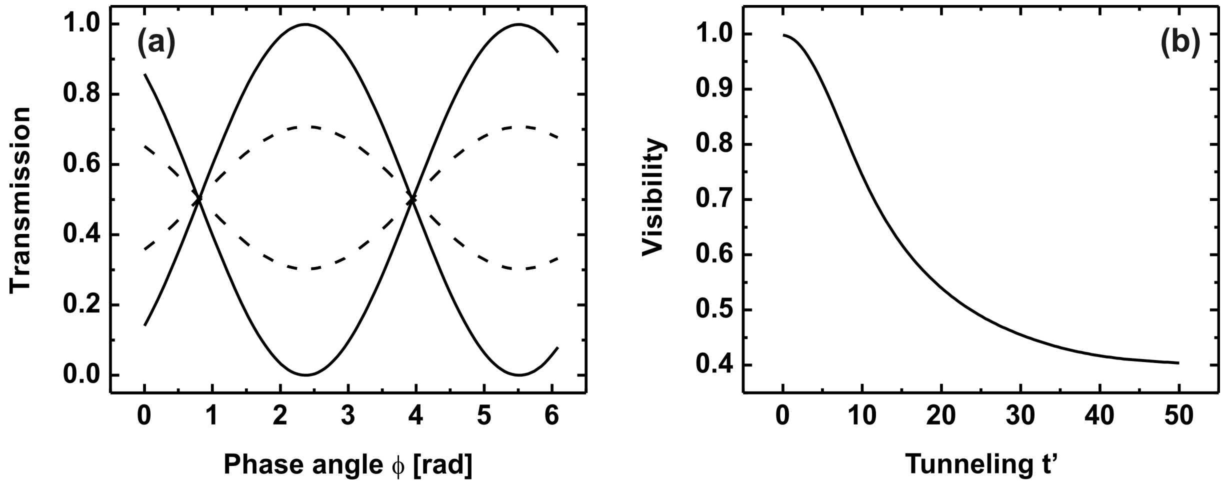

Figure 6.15: (a) Transmission probabilities as a function of the angle of the phase gate for two different tunneling couplings (solid curves: $t' = 0$, dashed curves: $t' = 50$). (b) Visibility as a function of the tunneling coupling as obtained from the analytical model.

In Fig. 6.15 (a) the transmission probabilities $T_{|0\rangle}$ and $T_{|1\rangle}$ are plotted for two different values of the tunneling coupling. The solid curves correspond to a tunneling coupling $t' = 0$, for the dashed curves the tunneling coupling has been set to $t' = 50$. The figure shows the sinusoidal dependence of the transmission probabilities on the angle of the phase gate. This corresponds directly to the interference pattern of the Mach-Zehnder interferometer. The important point is that the visibility of the interference pattern depends on the tunneling coupling in the same way as for the entangled Mach-Zehnder double quantum dot device: the solid curve shows full visibility, whereas the visibility of the dashed curve is lower than 50%.

Figure 6.15 (b) shows the visibility as a function of the tunneling coupling t' in the range between 0 and 50. The qualitative dependence of the visibility on the tunneling coupling is almost identical to that obtained for the entangled Mach-Zehnder double quantum dot device in the preceding section. We therefore conclude that our analytical model is indeed a valid description of the entangled Mach-Zehnder double quantum dot device. This in turn emphasizes that the entangled Mach-Zehnder double quantum dot device can be actually considered as a two-qubit quantum gate.

6.5 Summary

The topic of this chapter was to theoretically study semiconductor single- and two-qubit quantum gates based on electrostatically defined quantum wires and quantum dots. Our objective was the detailed, numerical analysis of realistic, fully three-dimensional models that would result in conceptually simple and experimentally realizable proposals of novel semiconductor devices.

We have worked out the detailed geometry, material compositions, doping profiles, and bias voltages for an all-electric Mach-Zehnder interferometer. Our calculations of the ballistic current through this device show that it acts as a fully controllable single-qubit gate for electrons propagating in either of two quantum wires. The fabrication of this device is within the reach of present-day technology as no lithographic feature is smaller than 40 nm. We have found

that the beam-splitters are the most critical components in terms of fabrication precision: even deviations of a few nm in the size of the beam-splitters spoil the device operation.

The all-electric Mach-Zehnder interferometer is the basis of our novel two-qubit quantum transport device. In this device, electrons that propagate through the interferometer are coupled via the Coulomb interaction to a single electron double quantum dot. The ballistic transport properties of this device have been calculated with the novel method that we have developed in Chapter 5. We have shown that the interaction between the electrons in the interferometer and the double quantum dot results in a reduced visibility of the Mach-Zehnder interference pattern. The tunneling coupling between the two quantum dots can be used to externally control this suppression. Further, our three-dimensional calculations have allowed us to identify the optimum geometric arrangement of the interferometer and the double quantum dot.

We have developed a simple analytical model of the novel device in terms of two interacting single-qubit systems. This model allows for an in-depth understanding of the device physics and the interpretation of the numerical results from a quantum information point of view: the visibility of the interference pattern is a measure for the entanglement of the two qubits as it is uniquely related to the von Neumann entropy. Thus the device realizes a non-trivial two-qubit quantum operation. In particular, it allows for the controlled generation and straightforward detection of entanglement from DC I-V characteristics.

Chapter 7

Further Applications

7.1 Calculation of carrier transport through quantum dot molecules

Introduction

Many concepts for quantum information processing in semiconductors are based on quantum dots. Two types of quantum dots have to be distinguished: electrostatically defined quantum dots and self-assembled quantum dots. In the following we will focus on the latter ones. To define a qubit based on quantum dots one can either exploit the charge degree of freedom [111, 112], excitonic degrees of freedom [113], or the spin degree of freedom [114, 115]. The possibility of coherent manipulation of excitonic qubits in single self-assembled quantum dots has been shown in a number of experiments performing Rabi oscillations by the application of electromagnetic pulses [116–120]. Another application of self-assembled quantum dots in the area of quantum information processing is the generation of entangled photons [121] and the deterministic generation of single photons, which is necessary for quantum cryptography (Ref. [122] and references therein). Beyond, self-assembled quantum dots are also of great interest for more traditional electro-optical devices such as lasers [123, 124].

Self-assembled quantum dots are grown in a self organized process driven by strain relaxation when a pseudomorphically strained layer of some material is deposited on a different substrate material and a certain critical layer thickness is exceeded (so called Stranski-Krastanov growth mode). This was first demonstrated in 1994 by Petroff and DenBaars, with InAs grown on a GaAs substrate [125].

Single quantum dots are limited to one, at maximum two qubits [126]. However, quantum computations demand for multi-qubit systems and the possibility to perform quantum gate operations between pairs of qubits. Several proposals have therefore considered the possibility to employ coupled quantum dots for the implementation of single qubits [127] as well as multiple qubits [128]. Coupled quantum dots can be realized by exploiting strain-induced correlation effects in the Stranski-Krastanov growth of multiple quantum dot layers on top of each other. The experimental realization of such vertical stacks of quantum dots [129, 130] is almost as old as that of a single layer of quantum dots by Petroff and DenBaars. Due to strain relaxation the quantum dots in latter layers grow preferentially on top of quantum dots in former layers, thereby forming self-aligned vertical stacks (Fig. 7.1).

The fabrication of semiconductor nanostructures by self-assembly has the disadvantage that the only parameters that influence the details of the structure are those that control the growth

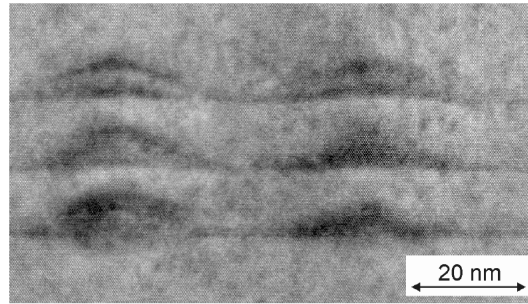

Figure 7.1: High resolution TEM cross-section of vertically correlated InAs-GaAs quantum dots via multi-layer Stranski-Krastanov growth. The picture is taken from Reference [110].

conditions. In general, this does not allow for a directed and reproducible fabrication of the desired structures. Thus it is necessary to identify suitable structures *a posteriori*, i.e. in a post-selection process. Such a process requires that quantum dots and quantum dot molecules can be characterized *in situ*. A common technique for characterizing quantum dots and quantum dot molecules is the use of optical spectroscopy [6, 19, 131]. Here, we suggest a different approach, based on ballistic transport through quantum dot molecules. Recently, Bryllert *et al.* have shown [132, 133] how to measure the ballistic tunneling current through a quantum dot molecule embedded in a resonant tunneling diode (RTD). We show that measurements of the tunneling current provide a wealth of unique information about the size, geometry, and the energy levels of quantum dot molecules and can therefore help to improve the *in situ* characterization of quantum dot molecules. In particular, we present a quantitative theoretical analysis of the ballistic current through a quantum dot molecule based on two vertically stacked quantum dots embedded within a RTD. Our calculations are based on a fully three-dimensional model of the RTD with realistically shaped quantum dot structures including the wetting layers. The local strain is calculated by minimizing the total elastic energy of the entire device. We solve the Poisson equation including the piezoelectric charges and subsequently calculate the ballistic current in terms of the contact block reduction method.

Modeling quantum dots and quantum dot molecules

As a further introduction into the topic of quantum dots and quantum dot molecules, we want to give a short overview on the subject on modeling quantum dots, respectively quantum dot molecules. Over the last years sophisticated numerical approaches have been developed for the interpretation of experimental results on quantum dots. The starting point of this development have been the articles published by Grundmann, Stier and Bimberg [134, 135] and Pryor *et al.* [136, 137] who first used three-dimensional models including the calculation of the strain distribution for a quantum dot and wetting layer embedded in GaAs. Most of the present approaches are based on a quantum mechanical description of the quantum dots in terms of single-band or multi-band Hamiltonians in $k \cdot p$ parameterization. However, more recent works also employ Hamiltonians that include more details of the electronic structure such as tight-binding [138] or pseudo potentials [139]. In the present analysis of ballistic transport properties of electrons through quantum dot molecules, our interest is focused on the spectrum and wavefunctions of the electron like states. Thus, we neglect the coupling between valence and conduction bands and choose the single-band effective mass approximation for the quantum

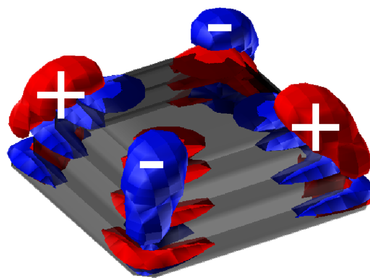

Figure 7.2: Localization of the piezoelectric charges for a quantum dot in the shape of a truncated pyramid. The positive and negative charges are indicated in red and blue, respectively.

mechanical description of our quantum dots.

The first models for the shape of self-assembled quantum dots were steep pyramids [134], similar to those observed experimentally for surface grown quantum dots. However, for buried quantum dots, interdiffusion during the capping process changes this shape. Depending on the growth conditions, a more realistic picture of buried quantum dots is that of truncated pyramids or lens shaped dots. An important role concerning the confining potential of quantum dots and therefore also concerning the energies and localization of the electron and hole wavefunctions, is played by the alloy composition. In InGaAs quantum dots for example, the Indium content is distributed inhomogeneously, increasing from the bottom to the tip of the dot. This leads to localization of the hole in the tip, while the electron remains delocalized over the whole dot. Modeling the Indium distribution accurately has proven to be one of the major achievements in the understanding of experimental data on Stark shifts [140]. More refined models of the Indium distribution were given in References [141, 142].

The growth of quantum dots is driven by strain relaxation between the substrate and the deposited material. Differences in the lattice constants of the two materials are the reason for the build up of these driving strain fields. Strain however is not completely relaxed and has a tremendous influence on the electronic structure of quantum dots. One part of this influence is by means of changes in the bandstructure in terms of deformation potentials, the second part are piezoelectric polarization charges. There are basically two models for calculating the strain relaxation in semiconductor nanostructures: The valence-force-field method [143] is an atomistic model and bound to the atomic lattice. It is therefore computationally very demanding. The second model is the continuum model as described in Section 1.2.4 and in Reference [19]. The continuum model can be employed on a grid with arbitrary resolution and is computationally less demanding. However, the relative displacement of the atoms of the crystal basis cannot be described within the continuum model.

The piezoelectric charges are readily obtained from the off-diagonal elements of the strain tensor according to Eq. (1.80) and Equation (1.83). The piezoelectric charges for a quantum dot in the shape of a truncated pyramid are shown in Figure 7.2. The piezoelectric charges are strongly localized at the edges of the dot and generate an octupole contribution to the electrostatic potential.

If multiple layers of quantum dots are grown subsequently, the quantum dots in latter layers grow preferentially on top of quantum dots in former layers, thereby building stacks of quantum dots. Again, strain is responsible for this self organized growth. If the separation between the layers is small the dots are coupled quantum mechanically and form a quantum dot molecule. The same methods that are used for modeling quantum dots can also be used to model quantum

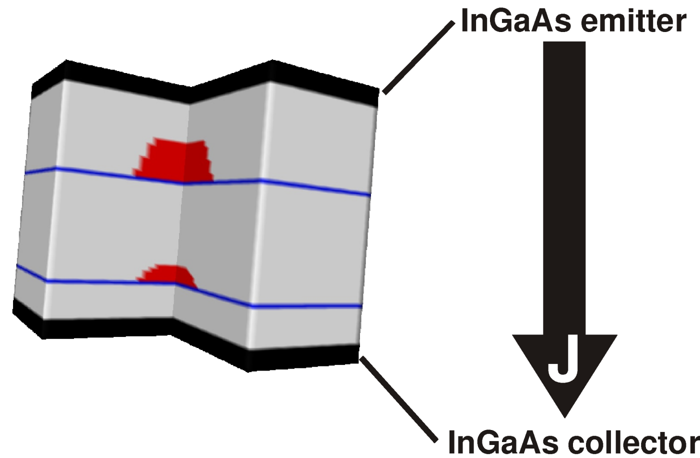

Figure 7.3: The structure of the device including the InGaAs leads (black), wetting layers (blue), and the QDs (red). The current flow from the emitter to the collector increases resonantly if two states in the upper and the lower QD are aligned in energy.

dot molecules.

Comparison to experimental results

The resonant tunneling diode (RTD) as depicted in Fig. 7.3 is composed of a 33 nm thick InP barrier with two embedded InAs quantum dots that are grown on 0.5 nm thick wetting layers. The distance between the wetting layers is 15 nm. This RTD was studied experimentally by Bryllert *et al.* [132, 133]. As depicted in the inset of Fig. 7.4, upon applying a bias, pairs of bound states in the two adjacent quantum dots can be brought into resonance with one another, thereby increasing the electron transmission probability by several orders of magnitude. Figure 7.4 shows the experimentally determined as well as the theoretically predicted current. The two quantum dots were modeled by truncated pyramids of 2.5 (5) nm height and 12 (16) nm base width, respectively. We have assumed a homogeneous alloy profile corresponding to no inter-diffusion of the dot material and the barrier material during capping. The difference ΔE_0 of the zero bias ground state energies is 87 meV. The figure shows a reasonable agreement between the calculation and experiment concerning the position of the resonance. The absolute value of the current is smaller by roughly a factor of 10 for the experimentally determined curve.

The differences between experiment and calculation may have several reasons. First of all, for our three-dimensional calculations we take into account the InP barrier region and only a small part of the InGaAs leads. The experimentally applied voltage (also denoted as *external voltage*) is therefore different from the voltage that is applied in the calculations (in the following denoted as *internal voltage drop*) by the voltage that drops in the leads. We have estimated this voltage drop and found that half of the external voltage drops in the leads, i.e. the internal voltage drop is smaller than the external voltage by a factor of 2. This factor of 2 is accounted for in Figure 7.4. However, also at the contacts to the InGaAs leads, the experimentally applied voltage might drop by an unknown amount due to poor quality of the contacts that results in Schottky barriers instead of ohmic behavior. This additional voltage drop is not accounted for in

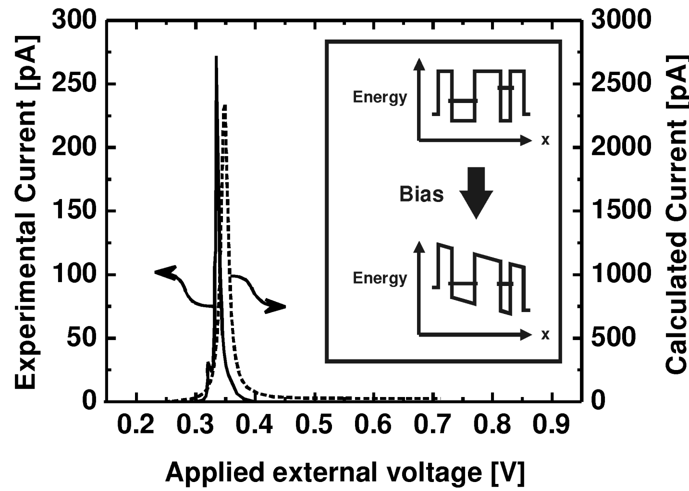

Figure 7.4: The experimental and calculated current through the InAs/InP QDM as a function of the applied bias. The inset shows the mechanism of resonant tunneling: upon applying a bias, pairs of bound state are brought into resonance and the transmission probability increases by several orders of magnitude.

Figure 7.4 and might explain the remaining discrepancy in the positions of the experimental and theoretical resonances. Parasitic resistances such as Schottky barriers at the contacts are also a possible explanation of the factor-of-10 difference in the absolute current. A further reason could be differences between the theoretically assumed and experimentally realized inter-dot distance and lateral alignment of the two quantum dots. This issue is addressed in detail in the next section. However, we want to anticipate here the following result that the linewidth of the resonance and also the peak current depend strongly on the inter-dot distance and the lateral misalignment.

To summarize, the comparison of our calculations with the experimental data from Bryllert *et al.* shows that we are able to model quantum dot molecules and calculate the ballistic transport through quantum dot molecules accurately. Despite the unaddressed issue of contact resistances, our model should therefore allow quantitative predictions on properties of the tunneling current that could be used to gain information about size, geometry, and energy levels of quantum dot molecules. This will be the topic of the following section.

Predictions

The advantage of analytical models for resonant tunneling through quantum dots consists in the understanding of experimental findings and basic physics in terms of only a few, universal quantities [144]. In Ref. [145] these are the tunneling rates Γ_{C}, Γ_{E}, and Γ_{inter} between the quantum dots and the collector and emitter, and in between the quantum dots, respectively. However, these tunneling rates are not directly related to properties of the quantum dot molecule such as the size, shape, alignment and distance of the quantum dots. Moreover, important effects that influence the energy spectrum of quantum dots, as for example strain and piezoelectric charges, are not present in such models. Thus, analytical models are good in the explanation of the underlying physics. However, the virtue to obtain quantitative predictions for concrete physical systems is missing. This part of the picture can be completed by our detailed numerical calcu-

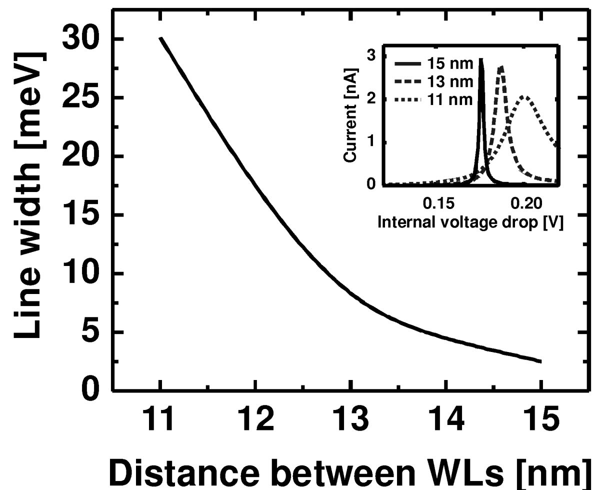

Figure 7.5: Calculated resonance line width as a function of the inter-dot distance. The inset shows the resonances for WL distances of 11, 13, and 15 nm, respectively. The peak current increases with increasing inter-dot distance, whereas the integrated current falls of exponentially.

lations. In the following, we analyze the influence of various geometric quantities of quantum dot molecules, such as size, inter-dot distance and lateral misalignment on the resonance line width and peak current. These relations can then be used for the *in situ* characterization of samples of quantum dot molecules.

Figure 7.5 shows the line width of the resonance to decrease exponentially with increasing quantum dot distance, which reflects the exponential decrease in the quantum dot coupling. The inset shows the shape of several current resonances explicitly. The peak current increases with increasing inter-dot distance, whereas the integrated current falls off exponentially. Furthermore, for decreasing inter-dot distance, the resonance is shifted to higher voltages. The reason for this is that a smaller fraction of the total internal voltage drop occurs in between the two quantum dots when decreasing the inter-dot distance. However, to bring the two quantum states into resonance, the same voltage drop in between the quantum dots is necessary, independent of the inter-dot distance. Therefore, the total internal voltage drop must be higher.

The dependence of the resonance line width on the lateral misalignment is much less pronounced as compared to the dependence on the inter-dot distance. Figure 7.6 (a) shows the line width as a function of the overlap of base areas. The overlap of base areas, as depicted in Fig. 7.6 (b), measures the lateral misalignment by the relative amount of coverage of the base area of the smaller quantum dot by the larger quantum dot and it is therefore independent of the direction of the misalignment. We have found that the line width is proportional to the overlap and changes in the order of a few meV if the overlap is decreased from 100% to 25%.

Further information on the lateral displacement of the two quantum dots can be obtained by considering additional resonances. As depicted in the inset of Fig. 7.7 (a), an additional resonance may occur if the excited state 1 in the large dot is in resonance with the ground state $0'$ of the small dot. Figure 7.7 (a) displays the ratio between the peak currents of the $1 \to 0'$ resonance and the $0 \to 0'$ resonance as a function of the relative lateral displacement of the two quantum dots. We note that the ratio $J_{1\to 0'}/J_{0\to 0'}$ is zero if the lateral displacement between the quantum dots vanishes. In fact, this means that $J_{1\to 0'}$ is zero and therefore that the excited state resonance is missing. The reason for this is that the excited state 1 in the large dot is p

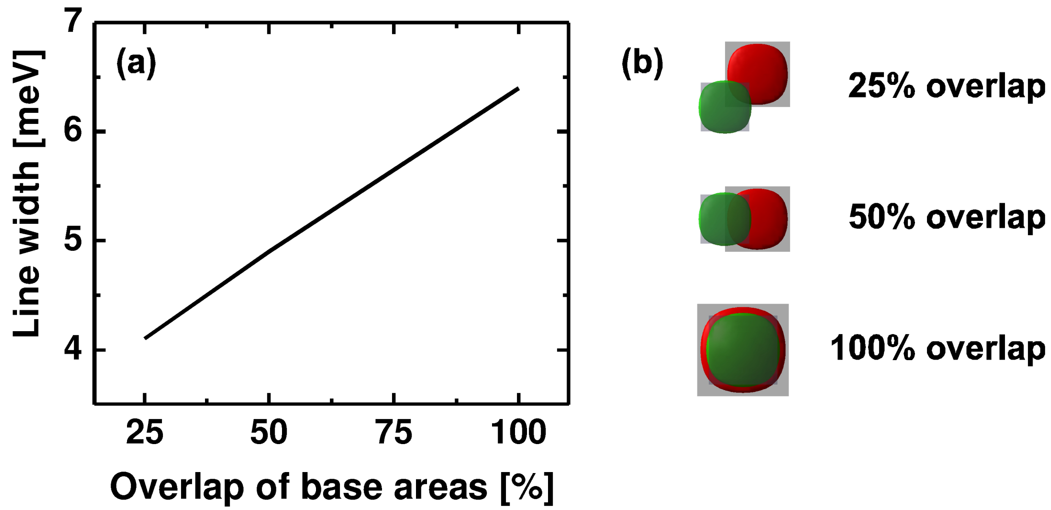

Figure 7.6: (a) Calculated resonance line width as a function of the overlap of base areas. (b) The overlap of base areas is a measure for the relative coverage of the base area of the smaller QD by the larger QD. Also shown are the according ground state wave functions.

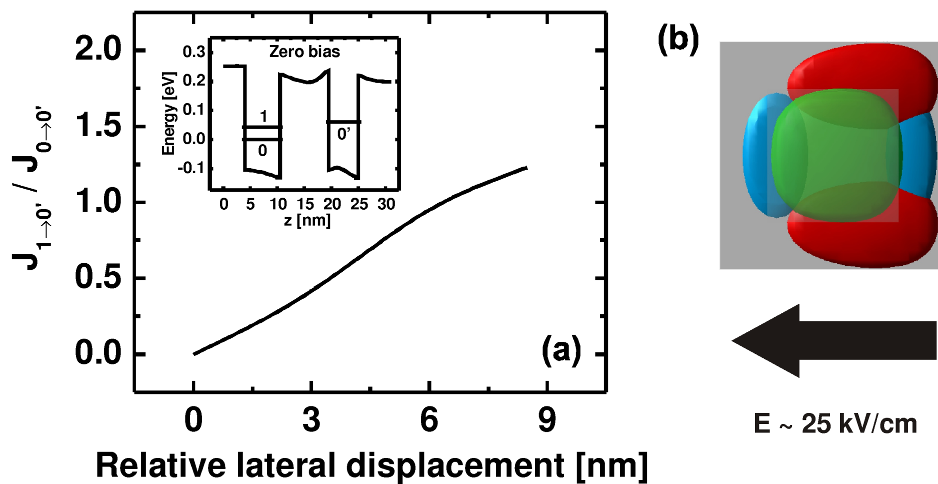

Figure 7.7: (a) Ratio between the peak currents of the $1 \to 0'$ resonance and the $0 \to 0'$ resonance as a function of the relative lateral displacement of the two QDs. The inset shows the energies of the ground state 0 and the excited state 1 in the large QD and of the ground state $0'$ in the small QD. (b) Ground state wave function of the smaller QD (green) and excited state wave functions of the larger QD (red and blue) for a lateral electric field of 25 kV/ cm.

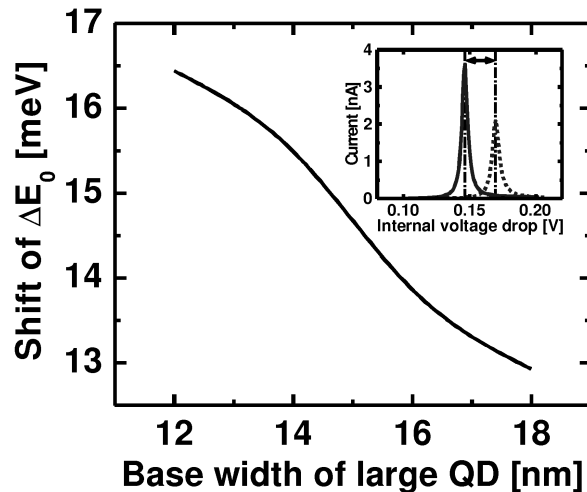

Figure 7.8: Change in the zero-bias energy difference ΔE_0 between the two QD ground states due to charging of the large QD, plotted as a function of the base width. The inset shows the shift of the current resonance for the neutral and the charged situation.

like and the ground state $0'$ in the small dot is s like. If brought into resonance they do not hybridize. If the two dots are displaced in different directions relative to the C_4 symmetry axis, the classification into p like and s like does no longer hold and the states hybridize if brought into resonance. The absence of a second resonance is therefore a very clear indication for good lateral alignment of the two quantum dots.

Another interesting technique in this context is the application of lateral electric fields. Due to a higher polarizability of the excited states in the larger quantum dot as compared to the ground state in the smaller quantum dot, the wave functions are shifted relative to each other. For an electric field of 25 kV/ cm, this is shown in Figure 7.7 (b). Therefore, the application of a lateral electric field corresponds to a lateral displacement. In particular, this means that the second resonance of an aligned quantum dot molecule can be made visible by the application of a lateral electric field. Firstly, this proves the lateral alignment, secondly, information on the energy difference between the ground state and the excited state in the large quantum dot is obtained.

Information on the size of the large quantum dot can be obtained if one allows for charging of this quantum dot. The effects of charging have been observed in Reference [133]. An emitter electron tunnels into and occupies the large quantum dot. Because of the thick InP barrier it takes some time before this electron tunnels out to the collector. In the meanwhile, the ground state of the large quantum dot is shifted to higher energies by the charging energy. The zero bias energy difference ΔE_0 between the ground state of the large and the ground state of the small quantum dot is decreased, and the resonance voltage is lowered. This is shown in the inset of Figure 7.8. The experimental DC I-V characteristics exhibits therefore a second resonance at lower voltage. The peak current of this resonance is determined by scattering and tunneling rates. In Fig. 7.8 we show our results for the calculation of the zero bias energy difference ΔE_0 as a function of the base width of the large quantum dot. ΔE_0 changes in the order of several meV for quantum dot base width in the range of 12 and 18 nm. Thus it is possible to estimate the lateral dimensions of the large quantum dot from the difference in resonance voltages between the ground state resonances with and without charging.

Summary

We have emphasized the importance of quantum dots and quantum dot molecules for optical applications as well as applications in the area of quantum computation and quantum cryptography. Next, we have given an overview about modeling of quantum dots and quantum dot molecules, pointing out the importance of strain and piezoelectric charges. We have presented results for the calculation of the ballistic carrier transport through quantum dot molecules that are embedded within a resonant tunneling diode. These results have been compared to experimental data for a similar resonant tunneling diode, and by this we have shown that we are able to model quantum dot molecules and calculate the ballistic current through quantum dot molecules accurately. Subsequently, we have presented predictions on the dependence of properties of the tunneling current on the geometry of quantum dot molecules. Further, we have discussed the application of lateral electric fields as an additional technique for the characterization of quantum dot molecules. To summarize, our calculations have shown that peak heights and positions, and the line width of the tunneling resonances through ground and excited states yield quantitative information about the energy levels, base widths, distance, and lateral alignment of the quantum dots in a quantum dot molecule. As a second source of information besides optical spectroscopy, this offers the possibility of a better *in situ* characterization of quantum dot molecule samples.

7.2 Strain in AlAs quantum wells and cleaved-edge overgrown quantum wires

Introduction

For quantum information processing in semiconductor nanostructures, a thorough understanding of electronic correlations is extremely important. On the one hand, intentional quantum mechanical correlations between electrons are the basis of multi-qubit operations. On the other hand, unintentional correlations may lead to decoherence and should therefore be minimized. One-dimensional systems of heavy electrons (effective mass m^* of order m_0, where m_0 is the free electron mass) are promising candidates for the study of electronic correlations in semiconductors. However, owing to an inherent low mobility $\mu \sim 1/m^*$ and fabrication challenges, few realizations are reported in the literature. Experiments in Si [146–148] and Si/SiGe [149–151] quantum point contacts, which begin to address this limit, have only focused on transport in (100)-plane structures with a light, isotropic in-plane mass $m^* = 0.19m_0$. Aluminum arsenide (AlAs) is an alternate heavy mass system with degenerate valleys and anisotropic mass. In AlAs the constant energy surfaces are ellipses centered at the X_i points ($i = x, y, z$) at the Brillouin zone edge, characterized by a heavy longitudinal mass $m_H^* = 1.1m_0$ and a light transverse mass $m_L^* = 0.19m_0$ [152]. The energy minima at these X points are highly sensitive to strain [153, 154].

In Ref. [154] a strain induced degeneracy crossover between two-dimensional subbands belonging to different X valleys has been demonstrated for two-dimensional AlAs/Al$_x$Ga$_{1-x}$As quantum-well systems grown along the (100) direction experimentally as well as theoretically. Here we confirm the observed crossover and predict the quantum-well width for the crossover in systems grown along the (110) direction.

The order of subbands belonging to different X valleys is also important for the analysis of transport properties of AlAs quantum wires as it affects valley degeneracies. In Ref. [155] conductance measurements of a quantum wire fabricated at the edge of an AlAs two-dimensional electron system using cleaved-edge overgrowth (CEO) were reported. Two conductance steps were observed at low electron density, with height $G_0 \sim 0.44e^2/h$ and therefore much lower than expected. Here we discuss the role of strain in relation to the observed 1D conductance.

Results and Discussion

We first address the influence of strain in AlAs quantum-wells on the level structure of the two-dimensional subbands. In two-dimensional AlAs systems, strain splits the degeneracy of the X_x, X_y, and X_z valleys from a mean value E_0 by an amount Δ [153], which is then compensated by the quantum confinement in the growth direction along the z-axis

$$\begin{aligned} E_x &= E_y = E_{xy} = E_0 + \frac{\hbar^2\pi^2}{2m_L^* W^2} - \frac{1}{3}\Delta, \\ E_z &= E_0 + \frac{\hbar^2\pi^2}{2m_H^* W^2} + \frac{2}{3}\Delta. \end{aligned} \tag{7.1}$$

These equations hold for growth in the (100) direction if we assume that the quantum well can be modeled by an infinitely deep potential well of width W. We have calculated the strain-induced splittings between E_{xy} and E_z charge self-consistently within effective mass approximation for two-dimensional AlAs/Al$_x$Ga$_{1-x}$As quantum-well systems grown along the

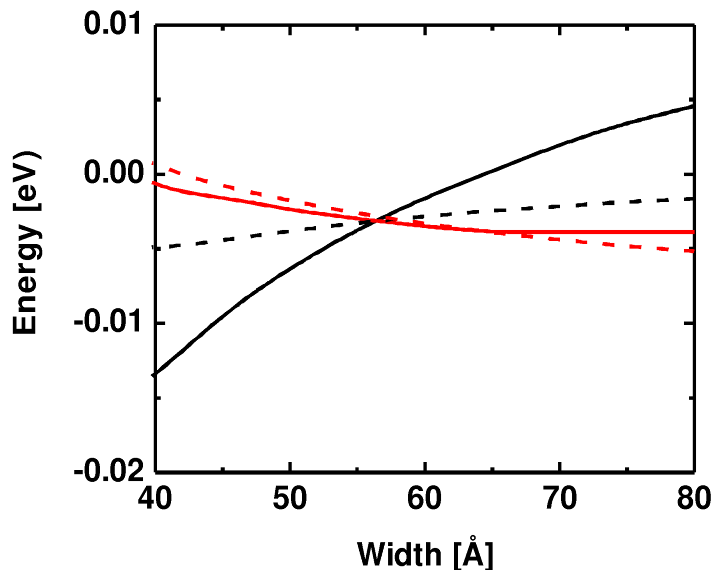

Figure 7.9: Energies E_{xy} (red curves) and E_z (black curves) of the lowest subband for two-dimensional AlAs quantum-well systems with (100) (solid lines) and (110) (dashed lines) growth direction. The crossover $E_{xy} = E_z$ depends only slightly on the growth direction.

(100) and (110) direction for various well widths. The structure consists of a single AlAs quantum-well embedded on both sides in a five periods 20 Å/10 Å GaAs/AlAs superlattice. Two buffer layers of $Al_{0.37}Ga_{0.63}As$ terminate the superlattices. Free carriers are provided by δ-doping layers ($0.75 \times 10^{-12}\,\mathrm{cm}^{-2}$) at a distance of 40 Å on both sides of the quantum well. The same structure has been analyzed experimentally as well as theoretically in Ref. [154] for the (100) growth direction only. In Fig. 7.9 we have plotted the energies E_z and E_{xy} as a function of the well width. Our calculations for the quantum-well systems grown along the (100) direction confirm the degeneracy crossover $E_{xy} = E_z$ at a well width of $W \sim 55 - 60$ Å that has been demonstrated experimentally in Ref. [154]. For wider quantum wells such as the one in Ref. [155] the strain term dominates Eq. (7.1) and $E_{xy} < E_z$, yielding dual valley degeneracy.

Only recently high-mobility two-dimensional $AlAs/Al_xGa_{1-x}As$ quantum-well systems have been grown along the (110) direction [156]. From our calculations for the quantum-well systems grown along the (110) direction we predict that the degeneracy crossover occurs for approximately the same well width. However, it is important to note that the dependence of the energy difference $|E_{xy} - E_z|$ on the well width is smaller by more than a factor of 2 for (110)-grown systems as compared to (100)-grown systems. In cases where it is important to distinguish between contributions of the E_{xy} and the E_z subband this might therefore prove to be more difficult in (110)-grown systems than it is in (100)-grown systems.

We now turn to the role of strain in AlAs quantum wires. The samples considered in Ref. [155] were fabricated with the same CEO technique also employed for ballistic GaAs wires [157]. They contain a modulation-doped AlAs quantum well [158] flanked by two $Al_xGa_{1-x}As$ dielectric barriers ($x = 0.45$). The samples are grown on a (001)-oriented GaAs substrate by molecular beam epitaxy. A 150-Å-wide quantum well resides 4000 Å below the surface, and a 1-μm-wide tantalum gate is patterned on top of the heterostructure. The samples are cleaved *in situ*, and the exposed (110) facet is overgrown with a modulation-doping sequence.

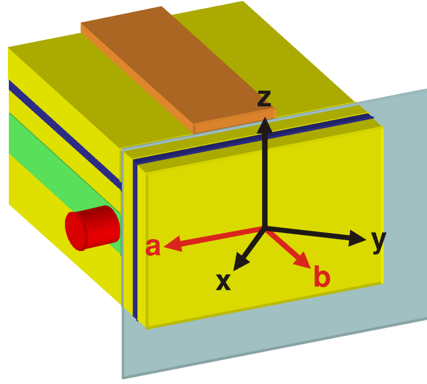

Figure 7.10: A schematic of the AlAs quantum wire sample. The protruding red cylinder represents the wire at the edge of the AlAs 2D system; yellow: AlGaAs, green: AlAs, blue: doping layers, orange: top gate, gray: cleavage plane. The figure is not drawn to scale.

A schematic of the sample is shown in Figure 7.10. For the further discussion, we define basis vectors $\hat{x}, \hat{y}, \hat{z} = [100], [010], [001]$ as unit vectors along the crystal axes and unit vectors $\hat{a} = (\hat{x} - \hat{y})/\sqrt{2} = [1\bar{1}0]$ and $\hat{b} = (\hat{x} + \hat{y})/\sqrt{2} = [110]$ The vector $\hat{a}$ denotes the direction parallel to the wire, $\hat{b}$ denotes the direction normal to the cleave plane.

We have performed calculations of the strain field in the two-dimensional $(\hat{b}, \hat{z})$ plane to determine the various spatially dependent X-band energies near the cleave plane. The subband energies have been determined charge self-consistently within effective mass approximation.

In Fig. 7.11, the spatial dependence of strain in the $(\hat{b}, \hat{z})$ plane perpendicular to the wire is depicted. Figure 7.11 (a) shows the $\varepsilon_{xx} = \varepsilon_{yy}$ components of the strain tensor. Maximum compressive strain is shown in red and can be found in the AlAs well away from the cleave plane. Figure 7.11 (b) shows the ε_{zz} component of the strain tensor. Here, maximum tensile strain is shown in red. Maximum tensile strain is found to the right of the AlAs well in the AlGaAs cleave plane. We want to note that $\varepsilon_{xx} = \varepsilon_{yy}$ is required by the symmetry of the structure. As a consequence, the X_x and X_y valleys are shifted by the same amount and the valley degeneracy $g_v = 2$ in the lowest bands of the two-dimensional well is preserved even in the wire region.

Figure 7.12 shows the effect of the strain alone on the X_z and the X_{xy} bands in the $\hat{b}$ direction in the middle of the well. Far from the cleave plane, the valley splitting of the band edges reaches 15 meV. This valley splitting is decreased down to 7 meV at the cleave plane, since both compressive $\varepsilon_{xx} = \varepsilon_{yy}$ and tensile ε_{zz} strain components decrease in magnitude towards the cleave plane. Figure 7.13 shows the results for the charge self-consistent calculation of the band edges. The two-dimensional electron gas in the AlAs quantum well has been depleted underneath the tantalum gate. Also shown in this figure are the energies of the first and second X_{xy} one-dimensional subband, the energy of the first X_z one-dimensional subband, the Fermi energy, and the probability density of the first X_{xy} one-dimensional quantum state. We note that the energies of the first and the second X_{xy} subbands are below the energy of the first X_z subband. The first X_{xy} subband is the only subband below the Fermi energy. The probability density of the corresponding quantum state is located within 20 nm from the cleave plane.

The author's present calculations have been published in Ref. [155] to help in understanding the conductance measurements that has been performed for the sample shown in Figure 7.10. In Ref. [155] two conductance steps were observed at low electron density, with height $G_0 \sim$

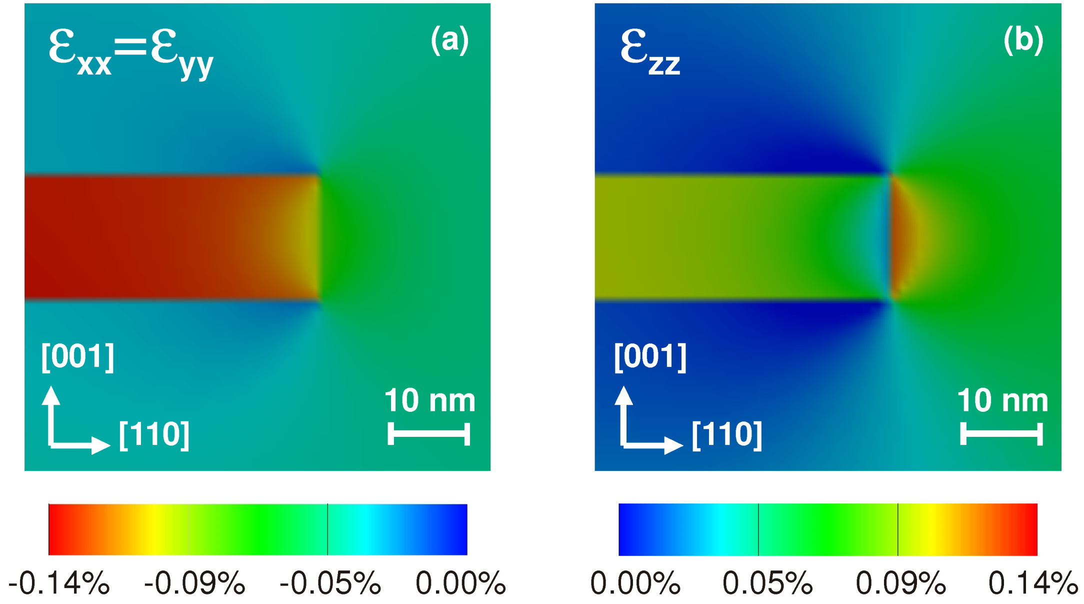

Figure 7.11: (a) Compressive $\varepsilon_{xx} = \varepsilon_{yy}$ and (b) tensile ε_{zz} strain components in the $(\hat{b}, \hat{z})$ plane.

$0.44e^2/h$. Our calculation shows that the two lowest one-dimensional subbands belong to X_{xy}. This is not *a priori* obvious, as the large splitting of 15 meV between the X_{xy} and the X_z in the well far from the cleave plane is reduced towards the cleave plane by more than a factor of 2. Thus the two conductance steps can both be attributed to the doubly degenerate valleys X_x and X_y. This is important since the valley degeneracy g_v enters the one-dimensional conductance quantum within the Landauer-Büttiker formalism according to $G_0 = g_s g_v \xi e^2/h$, where $g_s = 2$ denotes the spin degeneracy, and ξ is the transmission factor. Therefore, we can conclude that in the ballistic regime G_0 should be equal for the two lowest conductance steps and identical to $4e^2/h$ as they belong to the same doubly degenerate valleys X_x and X_y. The measurements indicate that the two lowest conduction steps are indeed equal in height, but smaller than $4e^2/h$ by almost a factor of 10. This reduction of the conductance quantum is most probably caused by a back-scattering induced suppression of the transmission coefficient.

Summary

We have presented calculations of the electronic structure of strained AlAs quantum-well and cleaved-edge overgrown quantum-wire systems. For the quantum-well systems we have confirmed the degeneracy crossover of the two lowest two-dimensional X valley subbands that has been experimentally observed for quantum wells grown along the (100) direction. For the (110) direction we have predicted that the degeneracy crossover occurs for the same quantum well width. However, the energy difference between the two lowest subbands depends less on the quantum well width by more than a factor of two for the (110) direction as compared to (100). For the quantum-wire systems, our self-consistent calculations including strain, doping profiles, and piezoelectricity have yielded the subband level ordering. This has contributed significantly to the understanding of recent measurements of the transport properties of strained AlAs cleaved-edge overgrown quantum wires.

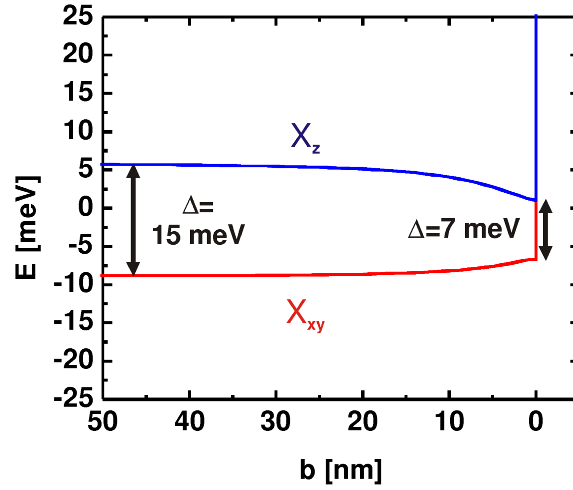

Figure 7.12: X_{xy} and X_z band structure in the center of the well along the $\hat{b}$ direction (distance is measured from the cleavage plane). The figure shows the strain-induced splitting of the valley degeneracy.

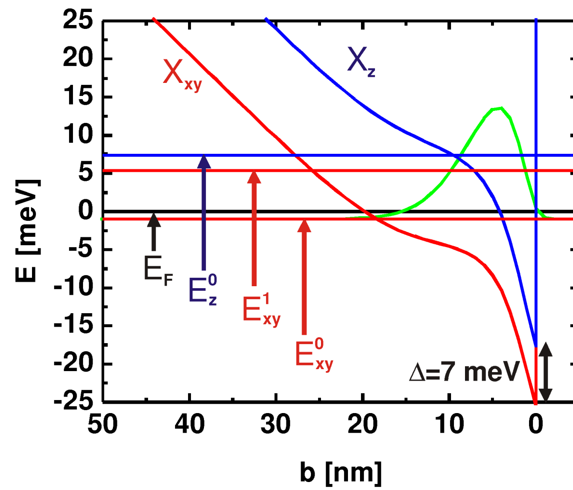

Figure 7.13: X_{xy} and X_z band structure in the center of the well along the $\hat{b}$ direction (distance is measured from the cleavage plane). The band structure has been calculated charge self-consistently within the effective mass approximation. The energies E^0_{xy}, E^1_{xy} of the first and second 1D X_{xy} subbands as well as E^0_z of the first X_z subband are indicated. $E_\mathrm{F} = 0$ is the Fermi energy. The probability density of the first X_{xy} 1D state is indicated (green curve).

Chapter 8

Summary and Outlook

The topic of this work was to study the physics of quantum logic gates built from semiconductor nanostructures by detailed, numerical calculations within realistic, three-dimensional models. The goal of these studies was to provide concrete proposals for the realization of semiconductor quantum logic gates including the detailed geometry, material compositions, doping profiles, and bias voltages.

We have succeeded to propose a concrete single-qubit gate based on two parallel quantum wires that are connected by two coupling windows. We have demonstrated by calculating the electronic structure and ballistic quantum transport properties of this device that it acts as the all-electric analog of a Mach-Zehnder interferometer. In particular, this means that the device allows for the coherent manipulation of electrons that propagate through the quantum wires. Thus, the electrons can be prepared in an arbitrary linear combination of the two qubit states defined by the wires. These studies have also aided to a better understanding of the physics of the coupling windows as coherent beam-splitters for electrons. In view of an experimental realization of the proposed all-electric Mach-Zehnder interferometer we have identified the beam-splitters as the most critical components in terms of fabrication precision. In collaboration with the group of Dr. Stefan Ludwig at the chair of Prof. Jörg Kotthaus (LMU Munich), a design for the beam-splitters as quantum mechanical current switches has been developed, experimentally tested, and will be presented in the Diploma thesis of Daniel Harbusch.

Based on the all-electric Mach-Zehnder interferometer, we have proposed a novel two-qubit quantum transport device that consists of the interferometer and an electrostatically defined single electron double quantum dot. The electrons in the interferometer and the double quantum dot interact via the Coulomb interaction. This leads to an entanglement of the two qubits. To study the proposed two-qubit gate within a realistic, three-dimensional model, we have developed a novel method for the calculation of ballistic quantum transport in entangled two-particle systems. The main result was the prediction that the entanglement causes a suppression of the visibility of the Mach-Zehnder interference pattern and can therefore be measured from its DC I-V characteristics. In particular, this does not involve any higher order current correlations. To this end, we have also developed an analytical model of the two-qubit quantum logic device, for which we have derived an analytical relation between the visibility of the interference pattern and the von Neumann entropy. Further, we have identified the optimum geometric arrangement of the Mach-Zehnder interferometer and the double quantum dot that yields maximum degree of entanglement between the two qubits.

With this thesis, we have also contributed to the area of material aspects of semiconductor

based quantum information processing: We have performed a quantitative theoretical analysis of the ballistic tunneling current through self-assembled quantum dot molecules embedded within a resonant tunneling diode. Our calculations were based on a fully three-dimensional model of the resonant tunneling diode with realistically shaped quantum dot structures. Self-assembled quantum dots and molecules are ideal candidates for the realization of either excitonic or spin qubits. With our calculations, we have shown that spectroscopic measurements of the ballistic tunneling current through quantum dot molecules provide a wealth of unique information on the size, inter-dot distance, lateral alignment, and electronic structure of these molecules. This second source of information besides optical spectroscopy is of high value, as a direct, non-destructive characterization of buried quantum dots and molecules is not possible.

Intentional quantum mechanical correlations between electrons are the basis for quantum information processing in semiconductors. Unintentional correlations on the other hand lead to decoherence and should therefore be minimized in quantum logic devices. Thus a thorough understanding of electronic correlations is extremely important. One-dimensional systems of heavy electrons such as AlAs cleaved-edge overgrown quantum wires are promising candidates for studying electronic correlations. We have performed self-consistent calculations of the electronic states in AlAs cleaved-edge overgrown quantum wires including strain, doping profiles, and piezoelectricity. This has aided considerably to the understanding of recent measurements of the transport properties of such systems.

By the development of a novel method for calculating the charge carrier density in broken gap heterostructures based on the 8-band $k \cdot p$ approximation, we have contributed methodically to the field of electronic structure calculation in semiconductor heterostructures. With our method it is now possible to calculate the electronic structure of a semiconductor heterostructure or superlattice charge self-consistently also in cases where valence band and conduction band states are hybridized. In the future, this will be applied to the calculation of inter-band transition energies and optical absorption coefficients in InAs/GaSb broken gap heterostructures. Those are of great interest for the fabrication of infrared lasers and infrared detectors.

As an additional benefit of the overall work, the simulation tool next**nano**++ has been developed and intensively used. The modular structure of next**nano**++ and the use of modern object oriented programming techniques such as inheritance and generic programming form the basis for the easy maintenance and extensibility of this software packet. The additional inclusion of magnetic fields has already led to the interesting application of g-factor engineering. Several proposals for semiconductor based quantum information processes exploit the idea of selectively manipulating the g-factor for individual electrons. The combination of magnetic fields with quantum transport will allow for realistic calculations of quantum Hall edge channels. This would allow one to study two-particle interference and entanglement in mesoscopic quantum transport devices such as Aharonov-Bohm based Mach-Zehnder interferometers. In this context, methods for the calculation of shot-noise and higher order current correlations are required. For the near future it is planned to include weak scattering into our ballistic quantum transport method in an empirical manner. This will then allow for self-consistent quantum transport calculations also far from thermodynamic equilibrium. Another project is to take into account the spin degree of freedom in quantum transport. The electron spin is of interest for quantum as well as classical information processing applications. With the inclusion of the electron spin in our quantum transport methods, the application of next**nano**++ could be extended to both, the simulation of classical spintronic devices and also spin based quantum information devices.

Appendix A

Grid definition

In Chapter 2 we have introduced the box integration finite differences scheme for the discretization of the partial differential equations on a non-uniform tensor product grid. Now, the topic of this appendix is the determination of the non-uniform tensor product grid from a set of given grid lines and spacings.

In the input file, the grid is defined by providing for each spatial dimension a list of major grid nodes $\{x_1, x_2, \ldots, x_N\}$ and a list of grid spacings $\{h_1, h_2, \ldots, h_N\}$. The goal is to determine a list of grid nodes that is composed of the major grid nodes and additional grid nodes in between the major grid nodes $\left\{x_1, x_1^1, x_1^2, \ldots, x_1^{N_1}, x_2, x_2^1, x_2^2, \ldots, x_{N-1}^{N_{N-1}}, x_N\right\}$. For each major grid node x_n, the grid spacing h_n denotes the approximate spacing to the first additional grid node on the left and on the right of x_n. If the grid spacings h_n and h_{n+1} of two neighboring major grid nodes x_n and x_{n+1} are different (we assume $h_n < h_{n+1}$ for definiteness), the grid spacings between the additional grid nodes in between x_n and x_{n+1} are chosen to increase exponentially. For $h_n = h_{n+1}$ the intermediate grid spacings are constant and the grid in between x_n and x_{n+1} is homogeneous.

Obviously, it is not always possible that the grid spacings between a major grid node x_n and the first grid node to the left and to the right are exactly given by h_n. This is because we always have to insert an integer number of additional grid nodes. If in the case of a homogeneous grid the distance between neighboring major grid nodes x_n and x_{n+1} is not an integer multiple of the grid spacings $h_n = h_{n+1}$, it is simply not possible to insert an integer number of additional grid nodes each having a spacing h_n. The same is also true for the case of an inhomogeneous, exponential grid.

We therefore use the following method to determine the grid by inserting an integer number of additional grid nodes while satisfying the conditions imposed by the given grid spacings h_n as close as possible:

For a pair of neighboring major grid nodes x_n and x_{n+1}, we calculate the distance L_n between those nodes

$$L_n = x_{n+1} - x_n. \tag{A.1}$$

For the corresponding grid spacing h_n and h_{n+1}, we distinguish the following two cases: $h_n = h_{n+1}$ and $h_n \neq h_{n+1}$. For $h_n = h_{n+1}$ we determine the number N_n of additional grid nodes in between x_n and x_{n+1} from

$$L_n = (N_n + 1)h_n = (N_n + 1)h_{n+1}. \tag{A.2}$$

This yields for N_n

$$N_n = \frac{L_n}{h_n} - 1, \tag{A.3}$$

but in general N_n is not an integer. We therefore set

$$N_n = \text{round}\left(\frac{L_n}{h_n} - 1\right) \tag{A.4}$$

and determine new grid spacings $h'_n = h'_{n+1}$ from

$$h'_{nn} = h'_{n+1} = \frac{L_n}{N_n + 1}. \tag{A.5}$$

These grid spacings now apply in between x_n and x_{n+1}. The more interesting case is that of $h_n \neq h_{n+1}$. Here, we require that

$$L_n = \sum_{i=0}^{N_n} \lambda_n^i h_n = \frac{\lambda^{N_n+1} - 1}{\lambda - 1} h_n, \tag{A.6}$$

where now we have to determine the grid factor λ_n and the number N_n of additional grid nodes. Further, we have to satisfy the condition

$$h_{n+1} = (\lambda_n)^{N_n} h_n, \tag{A.7}$$

i.e. the spacing to the left of x_{n+1} is uniquely related to h_n by means of λ_n and N_n. Also in the case $h_n \neq h_{n+1}$ it is in general not possible to find a grid factor λ_n and an integer number N_n of additional grid nodes that satisfy Eq. (A.6) and Equation (A.7). This problem is resolved as follows: First we solve Eq. (A.6) for λ_n, exploiting that $\lambda^{N_n+1} = \lambda_n h_{n+1}$

$$\begin{aligned} L_n &= \frac{\lambda_n h_{n+1} - h_n}{\lambda - 1} \\ &\Longrightarrow \lambda = \frac{L_n - h_n}{L_n - h_{n+1}}. \end{aligned} \tag{A.8}$$

Then we solve Eq. (A.7) for N_n

$$\begin{aligned} h_{n+1} &= \lambda^{N_n} h_n \\ &\Longrightarrow N_n = \log_\lambda \frac{h_{n+1}}{h_n} = \frac{\ln \frac{h_{n+1}}{h_n}}{\ln \frac{L_n - h_n}{L_n - h_{n+1}}} \end{aligned} \tag{A.9}$$

but in general this N_n is not an integer. We therefore set

$$N_n = \text{round}\left(\frac{\ln \frac{\Delta x_{n+1}}{\Delta x_n}}{\ln \frac{L_n - \Delta x_n}{L_n - \Delta x_{n+1}}}\right). \tag{A.10}$$

Now we determine λ_n from

$$f(\lambda_n) = \frac{\lambda_n^{N+1} - 1}{\lambda_n - 1} - \frac{L_n}{h_n} = 0. \tag{A.11}$$

This can be easily done using bisection as $f(\lambda_n)$ depends monotonously on λ_n. Note that by this approach, the grid spacing h_{n+1} to the left of grid node x_{n+1} is changed according to $h_{n+1} \to h'_{n+1} = (\lambda_n)^{N_n} h_n$, whereas the grid spacing h_n to the right of grid node x_n remains unchanged.

Appendix B

Geometry processing

In Sec. 3.5 we have presented the definition of quadratic surfaces (also denoted as quadrics) by means of a symmetric second rank tensor T as follows:

$$(\mathbf{x} - \mathbf{m}) \cdot T\,(\mathbf{x} - \mathbf{m}) = 0, \tag{B.1}$$

where $\mathbf{x}$ denotes the points on the surface and $\mathbf{m}$ denotes a translation vector. In this appendix, we show how this second rank tensor T can be determined from the parameters provided in the input file for the two cases of ellipsoidal and conic quadrics.

B.1 Semiellipsoids

We restrict ourselves to the case where the base plane of the semiellipsoid is parallel to the x–y plane. The semiellipsoid is then defined as shown in Fig. B.1 by specifying the base plane by $z = z_{\mathrm{base}}$, the extensions $x^{(1)}_{\mathrm{base}}$, $y^{(1)}_{\mathrm{base}}$, $x^{(2)}_{\mathrm{base}}$, $y^{(2)}_{\mathrm{base}}$ of the semiellipsoid in the base plane, and the top of the semiellipsoid x_{top}, y_{top}, z_{top} in the plane $z = z_{\mathrm{top}}$ above the base plane. From these input parameters we define vectors $\mathbf{a}'$, $\mathbf{b}'$, and $\mathbf{c}'$ as follows

$$\mathbf{a}' = \begin{pmatrix} x^{(2)}_{\mathrm{base}} \\ \frac{1}{2}\left(y^{(1)}_{\mathrm{base}} + y^{(2)}_{\mathrm{base}}\right) \\ z_{\mathrm{base}} \end{pmatrix}, \tag{B.2}$$

$$\mathbf{b}' = \begin{pmatrix} \frac{1}{2}\left(x^{(1)}_{\mathrm{base}} + x^{(2)}_{\mathrm{base}}\right) \\ y^{(2)}_{\mathrm{base}} \\ z_{\mathrm{base}} \end{pmatrix}, \tag{B.3}$$

$$\mathbf{c}' = \begin{pmatrix} x_{\mathrm{top}} \\ y_{\mathrm{top}} \\ z_{\mathrm{top}} \end{pmatrix}. \tag{B.4}$$

We want to determine the symmetric matrix

$$T = \begin{pmatrix} t_{11} & t_{12} & t_{13} \\ t_{12} & t_{22} & t_{23} \\ t_{13} & t_{23} & t_{33} \end{pmatrix}, \tag{B.5}$$

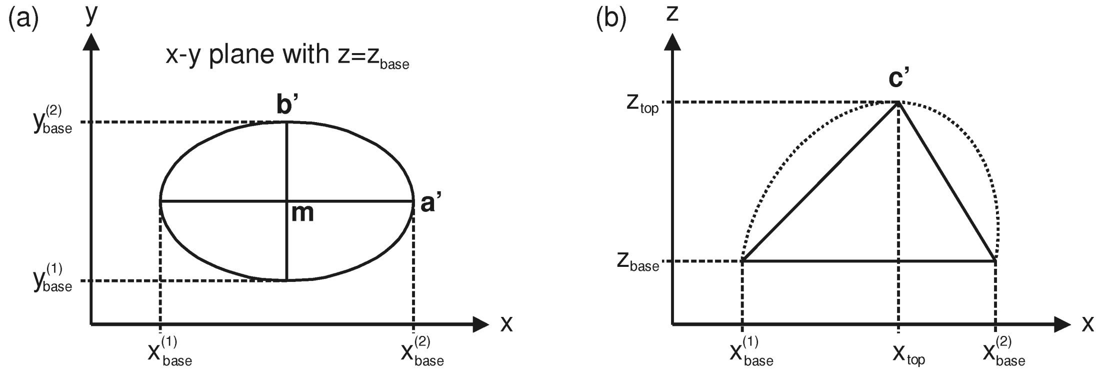

Figure B.1: Definition of semiellipsoids and cones. (a) The intersection of the quadric in the base plane. (b) For a cone, $\mathbf{c}'$ denotes the position of the cone's tip, whereas for a semiellipsoid, $\mathbf{c}'$ is the point of maximum extension in the direction normal to the base plane.

which by means of the following implicit equation

$$1 = (\mathbf{x} - \mathbf{m}) \cdot T \, (\mathbf{x} - \mathbf{m})$$

describes an ellipsoid with mid point $\mathbf{m}$

$$\mathbf{m} = \begin{pmatrix} \frac{1}{2}\left(x_{\mathrm{base}}^{(1)} + x_{\mathrm{base}}^{(2)}\right) \\ \frac{1}{2}\left(y_{\mathrm{base}}^{(1)} + y_{\mathrm{base}}^{(2)}\right) \\ z_{\mathrm{base}} \end{pmatrix} . \tag{B.6}$$

We have to determine the six coefficients t_{11}, t_{12}, t_{13}, t_{22}, t_{23}, and t_{33} from the following conditions

$$\begin{aligned} 1 &= (\mathbf{a}' - \mathbf{m}) \cdot T \, (\mathbf{a}' - \mathbf{m}) \, , \\ 1 &= (\mathbf{b}' - \mathbf{m}) \cdot T \, (\mathbf{b}' - \mathbf{m}) \, , & \text{(B.7)} \\ 1 &= (\mathbf{c}' - \mathbf{m}) \cdot T \, (\mathbf{c}' - \mathbf{m}) \, , & \text{(B.8)} \end{aligned}$$

i.e. $\mathbf{a}'$, $\mathbf{b}'$, and $\mathbf{c}'$ are points on the surface of the ellipsoid. We shift $\mathbf{a}'$, $\mathbf{b}'$, and $\mathbf{c}'$ by $\mathbf{m}$

$$\mathbf{a} = \mathbf{a}' - \mathbf{m} = \begin{pmatrix} \frac{1}{2}\left(x_{\mathrm{base}}^{(2)} - x_{\mathrm{base}}^{(1)}\right) \\ 0 \\ 0 \end{pmatrix} , \tag{B.9}$$

$$\mathbf{b} = \mathbf{b}' - \mathbf{m} = \begin{pmatrix} 0 \\ \frac{1}{2}\left(y_{\mathrm{base}}^{(2)} - y_{\mathrm{base}}^{(1)}\right) \\ 0 \end{pmatrix} , \tag{B.10}$$

$$\mathbf{c} = \mathbf{c}' - \mathbf{m} = \begin{pmatrix} x_{\mathrm{top}} - \frac{1}{2}\left(x_{\mathrm{base}}^{(1)} + x_{\mathrm{base}}^{(2)}\right) \\ y_{\mathrm{top}} - \frac{1}{2}\left(y_{\mathrm{base}}^{(1)} + y_{\mathrm{base}}^{(2)}\right) \\ z_{\mathrm{top}} - z_{\mathrm{base}} \end{pmatrix} . \tag{B.11}$$

Now, the conditions for T read

$$\begin{aligned} 1 &= \mathbf{a}\cdot T\mathbf{a} = a_1 a_1 t_{11}, && \text{(B.12)}\\ 1 &= \mathbf{b}\cdot T\mathbf{b} = b_2 b_2 t_{22}, && \text{(B.13)}\\ 1 &= \mathbf{c}\cdot T\mathbf{c} = c_1c_1t_{11} + 2c_1c_2t_{12} + 2c_1c_3t_{13} + c_2c_2t_{22} + 2c_2c_3t_{23} + c_3c_3t_{33}. && \text{(B.14)}\end{aligned}$$

These three conditions do not determine the ellipsoid completely. We have to further require that for $\mathbf{x} = \mathbf{c}$, x_z is extremal under the condition $1 = \mathbf{x}\cdot T\mathbf{x}$

$$g(\mathbf{x}) = x_z = \text{extremum!}, \tag{B.15}$$

$$f(\mathbf{x}) = \mathbf{x}\cdot T\mathbf{x} - 1 = 0. \tag{B.16}$$

We introduce a Lagrange multiplier λ. This results in

$$\begin{aligned} \mathbf{c}\cdot T\mathbf{c} &= 1, && \text{(B.17)}\\ \boldsymbol{\nabla} g(\mathbf{x})\big|_{x=c} &= \lambda \boldsymbol{\nabla} f(\mathbf{x})\big|_{\mathbf{x}=\mathbf{c}}. && \text{(B.18)}\end{aligned}$$

Only the second equation represents an additional condition and yields

$$\begin{aligned} 0 &= 2\lambda t_{1n}c_n = 2\lambda\left(t_{11}c_1 + t_{12}c_2 + t_{13}c_3\right) \Longrightarrow t_{11}c_1 + t_{12}c_2 + t_{13}c_3 = 0, && \text{(B.19)}\\ 0 &= 2\lambda t_{2n}c_n = 2\lambda\left(t_{12}c_1 + t_{22}c_2 + t_{23}c_3\right) \Longrightarrow t_{12}c_1 + t_{22}c_2 + t_{23}c_3 = 0, && \text{(B.20)}\\ 1 &= 2\lambda t_{3n}c_n = 2\lambda\left(t_{13}c_1 + t_{23}c_2 + t_{33}c_3\right) \Longrightarrow t_{13}c_1 + t_{23}c_2 + t_{33}c_3 = \frac{1}{2\lambda}. && \text{(B.21)}\end{aligned}$$

Equations (B.12), (B.13), (B.14), (B.19), (B.20), and (B.21) are still not sufficient to determine all six independent coefficients of T as only five of these equations are independent. We discard the last one involving the Lagrange multiplier. The condition that is missing can be identified by considering the case that

$$\mathbf{c} = \begin{pmatrix} 0 \\ 0 \\ c_3 \end{pmatrix}. \tag{B.22}$$

Then t_{11}, t_{22}, and t_{33} follow immediately from the first three conditions and t_{13} as well as t_{23} are found to vanish according to the remaining two conditions. The coefficient t_{12} cannot be determined from the present conditions. This coefficient fixes the orientation of the principal axes of the ellipsoid in the base plane. The information about the orientation of the principal axes of the ellipsoid in the base plane is not provided by the parameters in the input file, instead we assume that

$$t_{12} = 0. \tag{B.23}$$

The meaning of this condition is that the principal axes in the base plane are parallel to the coordinate axes. The coefficients of the tensor T are then determined by solving the following system of linear equations

$$\begin{pmatrix} a_1a_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & b_1b_1 & 0 & 0 \\ c_1c_1 & 2c_1c_2 & 2c_1c_3 & c_2c_2 & 2c_2c_3 & c_3c_3 \\ c_1 & c_2 & c_3 & 0 & 0 & 0 \\ 0 & c_1 & 0 & c_2 & c_3 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} t_{11} \\ t_{12} \\ t_{13} \\ t_{22} \\ t_{23} \\ t_{33} \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}. \tag{B.24}$$

For the coefficients we find

$$\begin{pmatrix} t_{11} \\ t_{12} \\ t_{13} \\ t_{22} \\ t_{23} \\ t_{33} \end{pmatrix} = \begin{pmatrix} \frac{1}{a_1^2} \\ 0 \\ -\frac{1}{a_1^2}\frac{c_1}{c_3} \\ \frac{1}{b_1^2} \\ -\frac{1}{b_1^2}\frac{c_2}{c_3} \\ \frac{1}{c_3^2} + \frac{1}{a_1^2}\frac{c_1^2}{c_3^2} + \frac{1}{b_1^2}\frac{c_2^2}{c_3^2} \end{pmatrix}. \tag{B.25}$$

B.2 Cones

Again, we restrict ourselves to the case where the base plane of the cone is parallel to the x–y plane. The cone is then defined by specifying the base plane by $z = z_{\text{base}}$, the extensions $x^{(1)}_{\text{base}}$, $y^{(1)}_{\text{base}}$, $x^{(2)}_{\text{base}}$, $y^{(2)}_{\text{base}}$ of the cone in the base plane, and the tip of the cone x_{tip}, y_{tip}, z_{tip} in the plane $z = z_{\text{tip}}$ above the base plane. We want to determine the symmetric matrix

$$T = \begin{pmatrix} t_{11} & t_{12} & t_{13} \\ t_{12} & t_{22} & t_{23} \\ t_{13} & t_{23} & t_{33} \end{pmatrix},$$

which by means of the following implicit equation

$$0 = (\mathbf{x} - \mathbf{m}) \cdot T\,(\mathbf{x} - \mathbf{m})$$

describes a cone with mid point $\mathbf{m}$. The mid point is given by the tip, i.e.

$$\mathbf{m} = \begin{pmatrix} x_{\text{tip}} \\ y_{\text{tip}} \\ z_{\text{tip}} \end{pmatrix}. \tag{B.26}$$

From the input parameters we define vectors **a, b**, and **c**

$$\mathbf{a} = \begin{pmatrix} x^{(2)}_{\text{base}} \\ \frac{1}{2}\left(y^{(1)}_{\text{base}} + y^{(2)}_{\text{base}}\right) \\ z_{\text{base}} \end{pmatrix} - \mathbf{m} = \begin{pmatrix} x^{(2)}_{\text{base}} - x_{\text{tip}} \\ \frac{1}{2}\left(y^{(1)}_{\text{base}} + y^{(2)}_{\text{base}}\right) - y_{\text{tip}} \\ z_{\text{base}} - z_{\text{tip}} \end{pmatrix}, \tag{B.27}$$

$$\mathbf{b} = \begin{pmatrix} \frac{1}{2}\left(x^{(1)}_{\text{base}} + x^{(2)}_{\text{base}}\right) \\ y^{(2)}_{\text{base}} \\ z_{\text{base}} \end{pmatrix} - \mathbf{m} = \begin{pmatrix} \frac{1}{2}\left(x^{(1)}_{\text{base}} + x^{(2)}_{\text{base}}\right) - x_{\text{tip}} \\ y^{(2)}_{\text{base}} - y_{\text{tip}} \\ z_{\text{base}} - z_{\text{tip}} \end{pmatrix}, \tag{B.28}$$

$$\mathbf{c} = \begin{pmatrix} x_{\text{tip}} \\ y_{\text{tip}} \\ z_{\text{tip}} \end{pmatrix} - \mathbf{m} = \mathbf{0}. \tag{B.29}$$

Obviously, the condition that **c** is a point on the surface of the cone is immediately satisfied

$$0 = \mathbf{c} \cdot T\mathbf{c}.$$

Thus, the following two conditions remain

$$0 = \mathbf{a}\cdot T\mathbf{a} = a_1a_1t_{11} + 2a_1a_2t_{12} + 2a_1a_3t_{13} + a_2a_2t_{22} + 2a_2a_3t_{23} + a_3a_3t_{33}, \tag{B.30}$$

$$0 = \mathbf{b}\cdot T\mathbf{b} = b_1b_1t_{11} + 2b_1b_2t_{12} + 2b_1b_3t_{13} + b_2b_2t_{22} + 2b_2b_3t_{23} + b_3b_3t_{33}. \tag{B.31}$$

We further assume that

$$t_{12} = 0, \tag{B.32}$$

i.e. as for the semiellipsoid, the principal axes in the base plane are parallel to the coordinate axes. Finally, we require that the direction $\mathbf{k}$ of the cone's symmetry axis is an eigenvector of T

$$T\mathbf{k} = \lambda\mathbf{k}, \tag{B.33}$$

where λ turns out to result in an overall scaling that is irrelevant and can therefore be set equal to 1. In summary, equations (B.30) – (B.33) result in the following system of linear equations

$$\begin{pmatrix} a_1a_1 & 2a_1a_2 & 2a_1a_3 & a_2a_2 & 2a_2a_3 & a_3a_3 \\ b_1b_1 & 2b_1b_2 & 2b_1b_2 & b_2b_2 & 2b_2b_3 & b_3b_3 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ k_1 & k2 & k3 & 0 & 0 & 0 \\ 0 & k1 & 0 & k2 & k3 & 0 \\ 0 & 0 & k1 & 0 & k2 & k3 \end{pmatrix} \begin{pmatrix} t_{11} \\ t_{12} \\ t_{13} \\ t_{22} \\ t_{23} \\ t_{33} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ k_1 \\ k_2 \\ k_3 \end{pmatrix}. \tag{B.34}$$

The direction of the symmetry axis can be calculated as follows

$$\mathbf{k} = \begin{pmatrix} \boldsymbol{\alpha}_x \\ \boldsymbol{\beta}_y \\ \frac{1}{2}\left(\boldsymbol{\alpha}_z + \boldsymbol{\beta}_z\right) \end{pmatrix}, \tag{B.35}$$

where

$$\boldsymbol{\alpha}' = \boldsymbol{\alpha}\frac{z_{\text{base}} - z_{\text{tip}}}{\boldsymbol{\alpha}_z}, \quad \boldsymbol{\alpha} = \frac{\mathbf{a}}{|\mathbf{a}|} + \frac{\bar{\mathbf{a}}}{|\bar{\mathbf{a}}|}, \tag{B.36}$$

$$\boldsymbol{\beta}' = \boldsymbol{\beta}\frac{z_{\text{base}} - z_{\text{tip}}}{\boldsymbol{\beta}_z}, \quad \boldsymbol{\beta} = \frac{\mathbf{b}}{|\mathbf{b}|} + \frac{\bar{\mathbf{b}}}{|\bar{\mathbf{b}}|}, \tag{B.37}$$

and

$$\bar{\mathbf{a}} = \begin{pmatrix} x^{(1)}_{\text{base}} - x_{\text{tip}} \\ \frac{1}{2}\left(y^{(1)}_{\text{base}} + y^{(2)}_{\text{base}}\right) - y_{\text{tip}} \\ z_{\text{base}} - z_{\text{tip}} \end{pmatrix}, \tag{B.38}$$

$$\bar{\mathbf{b}} = \begin{pmatrix} \frac{1}{2}\left(x^{(1)}_{\text{base}} + x^{(2)}_{\text{base}}\right) - x_{\text{tip}} \\ y^{(1)}_{\text{base}} - y_{\text{tip}} \\ z_{\text{base}} - z_{\text{tip}} \end{pmatrix}. \tag{B.39}$$

Appendix C

$k \cdot p$ and strain Hamiltonian for wurtzite

In Chapter 1 we have presented the $k \cdot p$ Hamiltonian including strain for semiconductors with zinc-blende crystal structure. Here we want to append the corresponding formulas for semiconductors with wurtzite crystal structure.

$k \cdot p$ **Hamiltonian** The valence bands are created by the p orbitals in most semiconductors and are therefore threefold degenerate. For this reason, the $k \cdot p$ model for the valence bands needs to include at least three bands in the orbital basis

$$\left|x_1\right\rangle, \left|x_2\right\rangle, \left|x_3\right\rangle. \tag{C.1}$$

For wurtzite this results in the following 3×3 Hamiltonian matrix

$$H^{\mathrm{wz},3\times3}(\mathbf{k}) = E_\mathrm{v} + \frac{\hbar^2}{2m}\mathbf{k}^2 + \begin{pmatrix} L_1k_1^2 + M_1k_2^2 + M_2k_3^2 & N_1k_1k_2 & N_2k_1k_3 \\ N_1k_2k_1 & M_1k_1^2 + L_1k_2^2 + M_2k_3^2 & N_2k_2k_3 \\ N_2k_3k_1 & N_2k_3k_2 & M_3\left(k_1^2+k_2^2\right) + L_2k_3^2 \end{pmatrix} \tag{C.2}$$

with Dresselhaus parameters L_1, L_2, M_1, M_2, N_1, and N_2 according to Reference [159]. Including spin

$$\left|x_1\right\rangle\left|\uparrow\right\rangle, \left|x_2\right\rangle\left|\uparrow\right\rangle, \left|x_3\right\rangle\left|\uparrow\right\rangle, \left|x_1\right\rangle\left|\downarrow\right\rangle, \left|x_2\right\rangle\left|\downarrow\right\rangle, \left|x_3\right\rangle\left|\downarrow\right\rangle, \tag{C.3}$$

the 6×6 Hamiltonian matrix is then given by

$$H^{\mathrm{wz},6\times6}(\mathbf{k}) = \begin{pmatrix} H^{\mathrm{wz},3\times3}(\mathbf{k}) + H_\mathrm{CF}^{\mathrm{wz},3\times3} & 0 \\ 0 & H^{\mathrm{wz},3\times3}(\mathbf{k}) + H_\mathrm{CF}^{\mathrm{wz},3\times3} \end{pmatrix} + H_\mathrm{SO}^{\mathrm{wz},6\times6} \tag{C.4}$$

where $H_\mathrm{SO}^{\mathrm{wz},6\times6}$ is due to the spin orbit coupling and $H_\mathrm{CF}^{\mathrm{wz},3\times3}$ is an additional contribution in wurtzite due to the so called crystal field splitting. The spin orbit coupling is given by

$$H_\mathrm{SO}^{\mathrm{wz},6\times6} = \begin{pmatrix} 0 & -i\Delta_2 & 0 & 0 & 0 & \Delta_3 \\ i\Delta_2 & 0 & 0 & 0 & 0 & -i\Delta_3 \\ 0 & 0 & 0 & -\Delta_3 & i\Delta_3 & 0 \\ 0 & 0 & -\Delta_3 & 0 & i\Delta_2 & 0 \\ 0 & 0 & -i\Delta_3 & -i\Delta_2 & 0 & 0 \\ \Delta_3 & i\Delta_3 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{C.5}$$

with parameters Δ_2, Δ_3. For the crystal field splitting one finds

$$H_\mathrm{CF}^{\mathrm{wz},3\times3} = \begin{pmatrix} \Delta_1 & 0 & 0 \\ 0 & \Delta_1 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{C.6}$$

The crystal field splitting is characterized by the parameter Δ_1. It lifts the degeneracy of the heavy hole band and the light hole band for $\mathbf{k} = 0$ also in the absence of strain. The inclusion of the conduction band at the Γ-point results in the following 8×8 Hamiltonian matrix

$$H^{\mathrm{wz},8\times8}(\mathbf{k}) = \begin{pmatrix} H_{\mathrm{c}}^{\mathrm{wz}}(\mathbf{k}) & H_{\mathrm{cv}}^{\mathrm{wz}}(\mathbf{k}) \\ H_{\mathrm{vc}}^{\mathrm{wz}}(\mathbf{k}) & H^{\mathrm{wz},6\times6}(\mathbf{k}) \end{pmatrix}, \tag{C.7}$$

where

$$H_{\mathrm{c}}^{\mathrm{wz}}(\mathbf{k}) = \left(E_{\mathrm{c}} + \frac{\hbar^2}{2m}\left[S_2\mathbf{k}^2 + (S_1 - S_2)\,k_3^2\right]\right) \otimes \mathbf{1}_{2\times2} \tag{C.8}$$

is the 2×2 submatrix for the conduction band with parameters S_1, S_2 and

$$H_{\mathrm{vc}}^{\mathrm{wz}}(\mathbf{k}) = \begin{pmatrix} B_1k_2k_3 + iP_2k_1 \\ B_2k_1k_3 + iP_2k_2 \\ B_3k_1k_2 + iP_1k_3 \end{pmatrix} \otimes \mathbf{1}_{2\times2}, \qquad H_{\mathrm{cv}}^{\mathrm{zb}}(\mathbf{k}) = H_{\mathrm{cv}}^{\mathrm{zb}}(\mathbf{k})^H \tag{C.9}$$

describes the coupling of the conduction and valence bands by means of the parameters B_1, B_2, B_3, P_1, and P_2.

Strain In the orbital basis $|x_1\rangle, |x_2\rangle, |x_3\rangle$ one obtains the following contribution to the $k \cdot p$ Hamiltonian due to strain

$$S^{\mathrm{wz}} = \begin{pmatrix} l_1\varepsilon_{11} + m_1\varepsilon_{22} + m_2\varepsilon_{33} & n_1\varepsilon_{12} & n_2\varepsilon_{13} \\ n_1\varepsilon_{12} & m_1\varepsilon_{11} + l_1\varepsilon_{22} + m_2\varepsilon_{33} & n_2\varepsilon_{23} \\ n_2\varepsilon_{13} & n_2\varepsilon_{23} & m_3(\varepsilon_{11} + \varepsilon_{22}) + l_2\varepsilon_{33} \end{pmatrix}. \tag{C.10}$$

This has to be added to the 3×3 Hamiltonian matrix $H^{\mathrm{wz},3\times3}(\mathbf{k})$ for spin up as well as for spin down. The parameters l_1, l_2, m_1, m_2, m_3, n_1, and n_2 depend on the deformation potentials d_1, d_2, d_3, d_4, d_5, and d_6 as follows:

$$l_1 = d_5 + d_4 + d_2, \tag{C.11}$$
$$l_2 = d_1, \tag{C.12}$$
$$m_1 = d_4 + d_2 - d_5, \tag{C.13}$$
$$m_2 = d_1 + d_3 + d_2, \tag{C.14}$$
$$m_3 = d_2, \tag{C.15}$$
$$n_1 = 2d_5, \tag{C.16}$$
$$n_2 = \sqrt{2}d_6. \tag{C.17}$$

Appendix D

Recombination and mobility models

The drift diffusion model for electronic transport introduced in Chapter 1 depends on recombination and generations rates as well as on the mobilities of the charge carriers. In this appendix, we summarize the various models that are employed within next**nano**++ for the determination of these quantities.

Recombination models Several mechanisms for the recombination of holes and electrons exist. Here we list the ones that are accounted for in next**nano**++:

1. Schockley-Read-Hall recombination: Recombination via a deep trap in the middle of the band gap
$$R_{\mathrm{SRH}} = \frac{p \cdot n - n_i^2}{\tau_p \cdot (n + n_i) + \tau_n \cdot (p + n_i)}, \tag{D.1}$$
where n_i is the intrinsic density. The scattering times τ_p and τ_n depend on the impurity concentrations N_D, N_A and are given by the semi-empiric Kendall relation
$$\tau_l = \frac{\tau_{l,0}}{1 + \frac{N_D + N_A}{N_{l,\mathrm{ref}}}}, \qquad l = p, n. \tag{D.2}$$
The intrinsic density is given by
$$n_i(\mathbf{x}) = \frac{1}{2}\left(\sum_i N_{\mathrm{C}}^i(T)\,\mathcal{F}_{1/2}\left(E_{\mathrm{F}} - E_{\mathrm{C}}^i(\mathbf{x})\right) + \sum_i N_{\mathrm{V}}^i(T)\,\mathcal{F}_{1/2}\left(E_{\mathrm{V}}^i(\mathbf{x}) - E_{\mathrm{F}}\right)\right) \tag{D.3}$$
where the Fermi energy E_{F} is determined for each $\mathbf{x}$ so that $\sum_i N_{\mathrm{C}}^i(T)\,\mathcal{F}_{1/2}(E_{\mathrm{F}} - E_{\mathrm{C}}^i(\mathbf{x})) - \sum_i N_{\mathrm{V}}^i(T)\,\mathcal{F}_{1/2}(E_{\mathrm{V}}^i(\mathbf{x}) - E_{\mathrm{F}}) = 0$.

2. Auger recombination: The empirical formula for Auger recombination is given by
$$R_{\mathrm{Auger}} = C_n n\left(p^2 - n_i^2\right) + C_p p\left(n^2 - n_i^2\right). \tag{D.4}$$
Auger recombination is a three-body effect and therefore occurs predominantly in regions where the charge carrier density is high. C_n and C_p are material dependent parameters.

3. Radiative recombination describes the direct recombination of electrons and holes by a photon emitting transition. This kind of recombination is important only in semiconductors with a direct band gap such as GaAs. For indirect band gap semiconductors, the

probability for a radiative recombination is low because this requires the additional contribution of a phonon for reasons of momentum conservation. Radiative recombination is described as follows:

$$R_{\text{radiative}} = C\left(np - n_i^2\right), \tag{D.5}$$

where C is a material dependent parameter.

Mobility models Several mobility models can be combined. The total mobility μ then results from the mobilities μ_i of the individual models according to Mathiessen's rule [39]:

$$\frac{1}{\mu} = \sum_i \frac{1}{\mu_i}. \tag{D.6}$$

In next**nano**++ we have implemented the following mobility models: constant mobility, Minimos [160], Arora [161], and Masetti [162]:

1. Constant mobility:

$$\mu_{\text{const}} = \mu_L \left(\frac{T}{300\,\text{K}}\right)^{-\zeta}. \tag{D.7}$$

The constant mobility model is independent of the impurity concentration $N_i = N_A + N_D$. It depends only on the two material dependent parameters μ_L and ζ.

2. Minimos:

$$\mu(\mathbf{x}) = \mu_{\text{min}} + \frac{\mu_L - \mu_{\text{min}}}{1 + \left(\frac{N_i}{C_{\text{ref}}}\right)^{\alpha}}, \tag{D.8}$$

where

$$\mu_{\text{min}} = \begin{cases} \mu_{\text{min},300}\left(\frac{T}{300\,\text{K}}\right)^{\gamma_1} & \text{for } T \geq T_{\text{switch}} \\ \mu_{\text{min},300}\left(\frac{2}{3}\right)^{\gamma_1}\left(\frac{T}{200\,\text{K}}\right)^{\gamma_2} & \text{for } T < T_{\text{switch}} \end{cases}, \tag{D.9}$$

$$C_{\text{ref}} = C_{\text{ref},300}\left(\frac{T}{300\,\text{K}}\right)^{\gamma_3}, \tag{D.10}$$

$$\alpha = \alpha_{300}\left(\frac{T}{300\,\text{K}}\right)^{\gamma_4}. \tag{D.11}$$

The Minimos mobility model depends on the following 9 material dependent parameters: μ_L, $\mu_{\text{min},300}$, γ_1, γ_2, γ_3, γ_4, T_{switch}, $C_{\text{ref},300}$, and α_{300}.

3. Arora:

$$\mu = \mu_{\text{min}} + \frac{\mu_d}{1 + \left(\frac{N_i}{N_0}\right)^{A^*}}, \tag{D.12}$$

where

$$\mu_{\text{min}} = A_{\text{min}}\left(\frac{T}{300\,\text{K}}\right)^{\alpha_m}, \quad \mu_d = A_d\left(\frac{T}{300\,\text{K}}\right)^{\alpha_d}, \tag{D.13}$$

$$N_0 = A_N\left(\frac{T}{300\,\text{K}}\right)^{\alpha_N}, \quad A^* = A_a\left(\frac{T}{300\,\text{K}}\right)^{\alpha_a}, \tag{D.14}$$

The Arora mobility model depends on the following 8 material dependent parameters: A_{min}, α_m, A_d, α_d, A_N, α_N, A_a, and α_a.

4. Masetti:

$$\mu = \mu_{\text{min1}} \exp\left(-\frac{P_C}{N_i}\right) + \frac{\mu_{\text{const}} - \mu_{\min 2}}{1 + \left(\frac{N_i}{C_r}\right)^{\alpha}} - \frac{\mu_1}{1 + \left(\frac{C_s}{N_i}\right)^{\beta}}. \tag{D.15}$$

The Masetti mobility model depends on the following 8 material dependent parameters: μ_{min1}, P_C, $\mu_{\min 2}$, C_r, α, μ_1, C_s, and β. In addition, the mobility μ_{const} from the constant mobility model enters the Masetti mobility model.

Appendix E

Alloys

In the database, material parameters are provided only for unary or binary materials such as Si, Ge, GaAs, AlAs, etc. Material parameters for ternary, or quaternary materials, such as $Al_xGa_{1-x}As$ or $Al_xGa_{1-x}As_ySb_{1-y}$ are determined from the material parameters of the binaries that they are composed of. For this, we use polynomial interpolation formulas. The simplest approach is linear interpolation by means of a convex combination. However, it turns out that for the calculation of some material properties of ternaries or quaternaries, a second order term should be included. This is widely known as bowing.

For the common material system $Al_xGa_{1-x}As$, the bowing parameter depends on the alloy composition x. We have developed consistent interpolation schemes for ternaries as well as quaternaries that also take such dependencies of the bowing parameter on the alloy composition into account in linear order [163]. These schemes are presented in the following.

E.1 Alloys of the type A_xB_{1-x}

For alloys of this type (and also $A_xB_{1-x}H$ where H is an inert component) we can use *linear interpolation*

$$f^{(1)}(x) = xf_{\mathrm{A}} + (1-x)\,f_{\mathrm{B}} \tag{E.1}$$

to compute the material property f as a function of the concentration x with

$$f_{\mathrm{A}} = f(\mathrm{A}), \qquad f_{\mathrm{B}} = f(\mathrm{B}). \tag{E.2}$$

A *second order interpolation* scheme $f^{(2)}(x)$ that is consistent to the components A and B can be obtained by adding a quadratic bowing parameter as

$$f^{(2)}(x) = xf_{\mathrm{A}} + (1-x)\,f_{\mathrm{B}} + x(1-x)\,f_{\mathrm{AB}}. \tag{E.3}$$

Finally, a consistent *third order interpolation* is given by

$$\begin{aligned} f^{(3)}(x) &= xf_{\mathrm{A}} + (1-x)\,f_{\mathrm{B}} + x(1-x)\,f_{\mathrm{AB}}(x) \\ &= xf_{\mathrm{A}} + (1-x)\,f_{\mathrm{B}} + x(1-x)\left[xf_{\mathrm{AB}}^{\mathrm{A}} + (1-x)\,f_{\mathrm{AB}}^{\mathrm{B}}\right] \end{aligned} \tag{E.4}$$

where $f_{\mathrm{AB}}(x)$ is a bowing parameter that depends linearly on x.

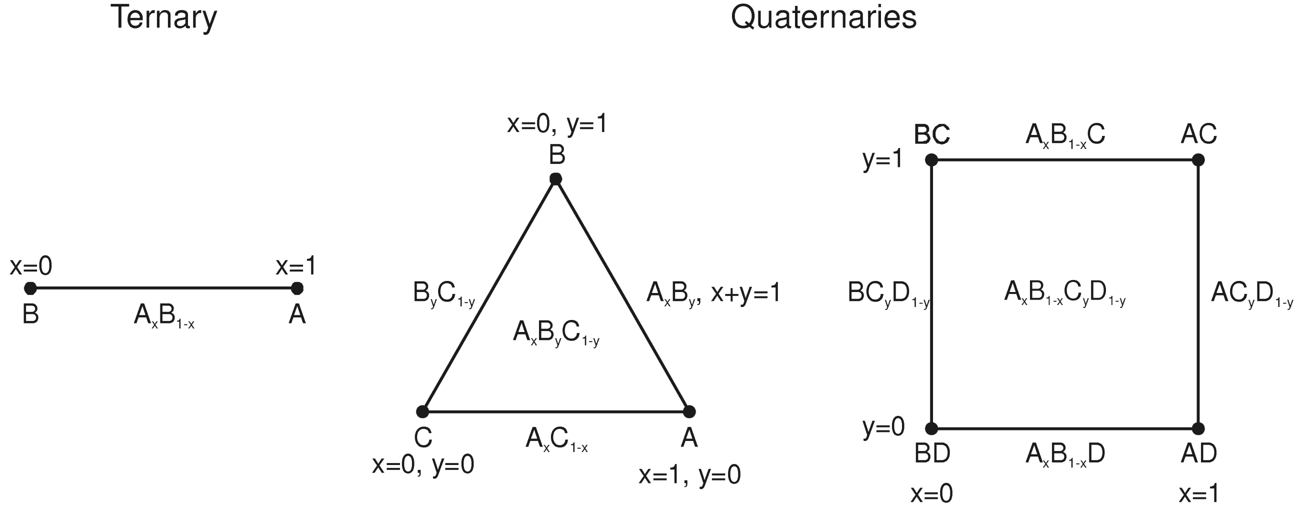

Figure E.1: Composition of ternaries A_xB_{1-x} and quaternaries $A_xB_yC_{1-x-y}$, $A_xB_{1-x}C_yD_{1-y}$ from binaries A, B, C, and D.

E.2 Alloys of the type $A_xB_yC_{1-x-y}$

For alloys of this type (and also $A_xB_yC_{1-x-y}H$ where H is an inert component) we can use *linear interpolation*

$$f^{(1)}(x,y) = xf_A + yf_B + (1-x-y)\,f_C \tag{E.5}$$

to describe the material property f as a function of the concentrations x and y with

$$f_A = f(A), \qquad f_B = f(B), \qquad f_C = f(C). \tag{E.6}$$

A *second order interpolation* scheme $f^{(2)}(x)$ that is also consistent to the three limiting one-parameter alloys A_zB_{1-z}, A_xC_{1-x}, and B_yC_{1-y} is given by

$$f^{(2)}(x,y) = xf_A + yf_B + (1-x-y)\,f_C + xyf_{AB} + x(1-x-y)\,f_{AC} + y(1-x-y)\,f_{BC} \tag{E.7}$$

where f_{AB}, f_{AC}, and f_{BC} are the bowing parameters for the limiting one-parameter alloys. Additional bowing that does not change the results for the limiting one-parameter cases can be achieved by adding a *third order term* as

$$\begin{aligned} f^{(2.3)}(x,y) = {} & xf_A + yf_B + (1-x-y)\,f_C + xyf_{AB} + x(1-x-y)\,f_{AC} + y(1-x-y)\,f_{BC} \\ & + xy(1-x-y)\,f_{ABC}. \end{aligned} \tag{E.8}$$

This extra term is also needed in order to achieve bowing on one of the three constraint surfaces

$$x = x_0, \qquad y = y_0, \qquad x + y = z_0. \tag{E.9}$$

But the most general *third order interpolation* scheme has 10 parameters. With

$$\begin{aligned} f^{(3)}(x,y) = {} & xf_A + yf_B + (1-x-y)\,f_C + xy\left[xf_{AB}^{A} + yf_{AB}^{B}\right] \\ & + x(1-x-y)\left[xf_{AC}^{A} + (1-x-y)\,f_{AC}^{C}\right] \\ & + y(1-x-y)\left[yf_{BC}^{B} + (1-x-y)\,f_{BC}^{C}\right] \\ & + xy(1-x-y)\,f_{ABC} \end{aligned} \tag{E.10}$$

these parameters can again be chosen consistent to the three limiting one-parameter alloys.

E.3 Alloys of the type $A_x B_{1-x} C_y D_{1-y}$

For alloys of this type (and also $A_x B_{1-x} C_y D_{1-y} H$ where H is an inert component) we cannot use linear interpolation to describe a material property f as a function of the concentrations x and y, since we need here at least four interpolation coefficients. For this reason, we add a *quadratic term* and use

$$f^{(1.2)}(x,y) = xyf_{\mathrm{AC}} + (1-x)\,yf_{\mathrm{BC}} + x\,(1-y)\,f_{\mathrm{AD}} + (1-x)\,(1-y)\,f_{\mathrm{BD}}$$

to interpolate f with

$$f_{\mathrm{AC}} = f\,(\mathrm{AC})\,, \qquad f_{\mathrm{BC}} = f\,(\mathrm{BC})\,, \qquad f_{\mathrm{AD}} = f\,(\mathrm{AD})\,, \qquad f_{\mathrm{BD}} = f\,(\mathrm{BD})\,. \tag{E.11}$$

To get an interpolation scheme that is consistent to the four limiting one-parameter alloys $A_x B_{1-x} C$, $A_x B_{1-x} D$, $C_y D_{1-y} A$, and $C_y D_{1-y} B$, we have to add both *quadratic and cubic terms* as

$$\begin{aligned} f^{(1.3)}(x,y) = {} & xyf_{\mathrm{AC}} + (1-x)\,yf_{\mathrm{BC}} + x\,(1-y)\,f_{\mathrm{AD}} + (1-x)\,(1-y)\,f_{\mathrm{BD}} \\ & + x\,(1-x)\,yf_{\mathrm{AB,C}} + x\,(1-x)\,(1-y)\,f_{\mathrm{AB,D}} \\ & + xy\,(1-y)\,f_{\mathrm{CD,A}} + (1-x)\,y\,(1-y)\,f_{\mathrm{CD,B}} \end{aligned} \tag{E.12}$$

in order to achieve consistency, where for instance $f_{\mathrm{AB,C}}$ is the bowing parameter for the limiting alloy $A_x B_{1-x} C$. Additional bowing that does not change the results for the limiting one-parameter cases can be achieved by adding a *fourth order term* as

$$\begin{aligned} f^{(1.4)}(x,y) = {} & xyf_{\mathrm{AC}} + (1-x)\,yf_{\mathrm{BC}} + x\,(1-y)\,f_{\mathrm{AD}} + (1-x)\,(1-y)\,f_{\mathrm{BD}} \\ & + x\,(1-x)\,yf_{\mathrm{AB,C}} + x\,(1-x)\,(1-y)\,f_{\mathrm{AB,D}} \\ & + xy\,(1-y)\,f_{\mathrm{CD,A}} + (1-x)\,y\,(1-y)\,f_{\mathrm{CD,B}} + x\,(1-x)\,y\,(1-y)\,f_{\mathrm{ABCD}}. \end{aligned} \tag{E.13}$$

This extra term is also needed in order to achieve bowing on one of the three constraint surfaces

$$x = x_0, \qquad y = y_0, \qquad y - \lambda_0 x = 0, \tag{E.14}$$

where the third type of constraint is used for strain compensation. By adding more *cubic and fourth order terms* as

$$\begin{aligned} f^{(2.4)}(x,y) = {} & xyf_{\mathrm{AC}} + (1-x)\,yf_{\mathrm{BC}} + x\,(1-y)\,f_{\mathrm{AD}} + (1-x)\,(1-y)\,f_{\mathrm{BD}} \\ & + x\,(1-x)\,y\left[xf^{\mathrm{A}}_{\mathrm{AB,C}} + (1-x)\,f^{\mathrm{B}}_{\mathrm{AB,C}}\right] + x\,(1-x)\,(1-y)\left[xf^{\mathrm{A}}_{\mathrm{AB,D}} + (1-x)\,f^{\mathrm{B}}_{\mathrm{AB,D}}\right] \\ & + xy\,(1-y)\left[yf^{\mathrm{C}}_{\mathrm{CD,A}} + (1-y)\,f^{\mathrm{D}}_{\mathrm{CD,A}}\right] + (1-x)\,y\,(1-y)\left[yf^{\mathrm{C}}_{\mathrm{CD},B} + (1-y)\,f^{\mathrm{D}}_{\mathrm{CD,B}}\right] \\ & + x\,(1-x)\,y\,(1-y)\,f_{\mathrm{ABCD}} \end{aligned} \tag{E.15}$$

we can also achieve consistency to the limiting one-parameter alloys.

Appendix F

Analytical model: detailed calculations

F.1 Transfer matrices

In this appendix we present the details of the derivation of the transfer matrices given by Eq. (6.9) and Eq. (6.10) in Sec. 6.4.2. For this, we have to solve the quantum mechanical scattering problem given by $H_0 + V(x)$, where H_0 is the Hamiltonian given by Eq. (6.5) and $V(x)$ is the interaction given by Equation (6.6). In matrix form, the total Hamilton operator $H = H_0 + V(x)$ can be expressed as follows

$$H = \frac{p^2}{2m}\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{pmatrix} + \frac{1}{2}\begin{pmatrix}-\Delta&-t&0&0\\-t&\Delta&0&0\\0&0&-\Delta&-t\\0&0&-t&\Delta\end{pmatrix} + U\begin{pmatrix}0&0&0&0\\0&0&0&0\\0&0&0&0\\0&0&0&1\end{pmatrix}\theta(x)\,\theta(L-x)\,. \quad \text{(F.1)}$$

We assume $t, \Delta > 0$, further $U > 0$. The latter corresponds to a repulsive interaction as the Coulomb interaction between two electrons. We have expressed the Hamiltonian H in terms of the basis states

$$\psi_{k,i,j}(x) = \exp(ikx)\,|i\rangle^W\,|j\rangle^D\,, \qquad i,j = 0,1. \quad \text{(F.2)}$$

These states describe the two electrons in the quantum wires and the double quantum dot as follows: the first electron propagates along the x axis in either of two states $|0\rangle^W$ and $|1\rangle^W$, corresponding to the two quantum wires. The second electron occupies one of the ground states of the two uncoupled quantum dots, represented by states $|0\rangle^D$ and $|1\rangle^D$. We determine two additional basis sets that diagonalize

$$\mathbf{1}^W_{2\times2}\otimes H^D = \frac{1}{2}\begin{pmatrix}-\Delta&-t&0&0\\-t&\Delta&0&0\\0&0&-\Delta&-t\\0&0&-t&\Delta\end{pmatrix} \quad \text{(F.3)}$$

and

$$\mathbf{1}^W_{2\times2}\otimes H^D + V = \frac{1}{2}\begin{pmatrix}-\Delta&-t&0&0\\-t&\Delta&0&0\\0&0&-\Delta&-t\\0&0&-t&\Delta\end{pmatrix} + U\begin{pmatrix}0&0&0&0\\0&0&0&0\\0&0&0&0\\0&0&0&1\end{pmatrix}. \quad \text{(F.4)}$$

The first basis set is given by

$$\psi'_{k,i,X}(x) = \exp(ikx)\,|i\rangle^W\,|X\rangle^D \quad \text{(F.5)}$$

where $X = B, A$ and

$$|B\rangle^D = \frac{\Delta - 2E_B}{\sqrt{t^2 + (\Delta - 2E_B)^2}} |0\rangle^D + \frac{t}{\sqrt{t^2 + (\Delta - 2E_B)^2}} |1\rangle^D, \tag{F.6}$$

$$|A\rangle^D = \frac{\Delta - 2E_A}{\sqrt{t^2 + (\Delta - 2E_A)^2}} |0\rangle^D + \frac{t}{\sqrt{t^2 + (\Delta - 2E_A)^2}} |1\rangle^D, \tag{F.7}$$

with energies $E_B = -\frac{1}{2}\sqrt{t^2 + \Delta^2}$ and $E_A = \frac{1}{2}\sqrt{t^2 + \Delta^2}$. The second basis set is given by

$$\psi''_{k,i,X}(x) = \exp(ikx) |i, X\rangle \tag{F.8}$$

where again, $i = 0, 1$, $X = A, B$ and

$$|0, B\rangle = |0\rangle^W |B\rangle^D, \tag{F.9}$$

$$|0, A\rangle = |0\rangle^W |A\rangle^D, \tag{F.10}$$

$$|1, B\rangle = |1\rangle^W \left(\alpha_B |B\rangle^D + \beta_B |A\rangle^D\right), \tag{F.11}$$

$$|1, A\rangle = |1\rangle^W \left(\alpha_A |B\rangle^D + \beta_A |A\rangle^D\right), \tag{F.12}$$

with energies $E_{1,B} = \frac{1}{2}(U - \varepsilon)$ and $E_{1,A} = \frac{1}{2}(U + \varepsilon)$, where $\varepsilon = \frac{1}{2}\sqrt{t^2 + (U + \Delta)^2}$ and coefficients

$$\alpha_X = \frac{tU}{\sqrt{\left(4E_X(E_X + \varepsilon) + \Delta U\right)^2 + t^2U^2}}, \tag{F.13}$$

$$\beta_X = \frac{4E_X(E_X + \varepsilon) + \Delta U}{\sqrt{\left(4E_X(E_X + \varepsilon) + \Delta U\right)^2 + t^2U^2}}. \tag{F.14}$$

For the solution of the scattering problem, we now solve the stationary two-particle Schrödinger equation $H\psi = E\psi$ separately in the three spatial regions $z < 0$, $0 < z < L$, and $z > L$. Employing the two additional basis sets, the solutions for the total energy E are given by

$$\psi^{\pm}_{i,X}(x) = \begin{cases} \exp\left(\pm i\bar{k}_X x\right) |i\rangle^W |X\rangle^D & \text{for } x < 0 \\ \exp\left(\pm i\bar{q}_{i,X} x\right) |i, X\rangle & \text{for } 0 < x < L \\ \exp\left(\pm i\bar{k}_X x\right) |i\rangle^W |X\rangle^D & \text{for } x > L \end{cases}, \tag{F.15}$$

where $\bar{k}_X = \frac{1}{\hbar}\sqrt{2m(E - E_X)}$ and $\bar{q}_{i,X} = \frac{1}{\hbar}\sqrt{2m(E - E_{i,X})}$. We determine general solutions for each of the three spatial regions by forming linear combinations of the basis functions

$$\psi_<(x) = \sum_{i=0,1} \sum_{X=B,A} \sum_{\sigma=+,-} A^{i,X}_\sigma \psi^\sigma_{i,X}(x) \quad \text{for } x < 0, \tag{F.16}$$

$$\psi_\parallel(x) = \sum_{i=0,1} \sum_{X=B,A} \sum_{\sigma=+,-} B^{i,X}_\sigma \psi^\sigma_{i,X}(x) \quad \text{for } 0 < x < L, \tag{F.17}$$

$$\psi_>(x) = \sum_{i=0,1} \sum_{X=B,A} \sum_{\sigma=+,-} C^{i,X}_\sigma \psi^\sigma_{i,X}(x) \quad \text{for } x > L. \tag{F.18}$$

Our goal is to determine the 8×8 transfer matrix T that connects the coefficient 8-vector $A = \left(A_\sigma^{i,X}\right)$ and the coefficient 8-vector $C = \left(C_\sigma^{i,X}\right)$ according to

$$C = TA. \tag{F.19}$$

This transfer matrix follows from the continuity and differentiability conditions for $\psi_<$, $\psi_\parallel$, and $\psi_>$

$$\psi_> (0) = \psi_\parallel (0) \text{ and } \partial_x \psi_> (x)\big|_{x=0} = \partial_x \psi_\parallel (x)\big|_{x=0}, \tag{F.20}$$

$$\psi_\parallel (L) = \psi_> (L) \text{ and } \partial_x \psi_\parallel (x)\big|_{x=L} = \partial_x \psi_> (x)\big|_{x=L}. \tag{F.21}$$

However, we note that these conditions are trivial for the coefficients with $i = 0$ as the interaction $V(x)$ vanishes if the electron in the quantum wires occupies the state $|0\rangle^W$. The transfer matrix T can therefore be written as follows:

$$T = \begin{pmatrix} \mathbf{1}_{4\times4} & \mathbf{0}_{4\times4} \\ \mathbf{0}_{4\times4} & \boldsymbol{\tau} \end{pmatrix} \tag{F.22}$$

where $\boldsymbol{\tau}$ is a 4×4 matrix. The submatrix $\boldsymbol{\tau}$ can be decomposed into four 4×4 matrices $\boldsymbol{\mu}_{0,1}$, $\boldsymbol{\mu}_{0,2}$, $\boldsymbol{\mu}_{L,1}$, and $\boldsymbol{\mu}_{L,2}$ according to

$$\boldsymbol{\tau} = \boldsymbol{\mu}_{L,2}^{-1} \boldsymbol{\mu}_{L,1} \boldsymbol{\mu}_{0,2}^{-1} \boldsymbol{\mu}_{0,1}, \tag{F.23}$$

where the matrices $\boldsymbol{\mu}_{x,i}$ ($x = 0, L$, $i = 1, 2$) are defined by

$$\boldsymbol{\mu}_{0,1} \begin{pmatrix} A_+^{1,B} \\ A_-^{1,B} \\ A_+^{1,A} \\ A_-^{1,A} \end{pmatrix} = \boldsymbol{\mu}_{0,2} \begin{pmatrix} B_+^{1,B} \\ B_-^{1,B} \\ B_+^{1,A} \\ B_-^{1,A} \end{pmatrix}, \tag{F.24}$$

$$\boldsymbol{\mu}_{L,1} \begin{pmatrix} B_+^{1,B} \\ B_-^{1,B} \\ B_+^{1,A} \\ B_-^{1,A} \end{pmatrix} = \boldsymbol{\mu}_{L,2} \begin{pmatrix} C_+^{1,B} \\ C_-^{1,B} \\ C_+^{1,A} \\ C_-^{1,A} \end{pmatrix}, \tag{F.25}$$

and have to be determined by means of Eq. (F.20) and Equation (F.21). Equation (F.20) yields

$$\sum_{X=B,A} \sum_{\sigma=+,-} A_\sigma^{1,X} |X\rangle^D = \sum_{X=B,A} \sum_{\sigma=+,-} B_\sigma^{1,X} \left(\alpha_X |B\rangle^D + \beta_X |A\rangle^D\right), \tag{F.26}$$

$$\sum_{X=B,A} \sum_{\sigma=+,-} \sigma i \bar{k}_X A_\sigma^{1,X} |X\rangle^D = \sum_{X=B,A} \sum_{\sigma=+,-} \sigma i \bar{q}_{1,X} B_\sigma^{1,X} \left(\alpha_X |B\rangle^D + \beta_X |A\rangle^D\right), \tag{F.27}$$

and Eq. (F.21) yields

$$\sum_{X=B,A} \sum_{\sigma=+,-} B_\sigma^{1,X} e^{\sigma i \bar{q}_{1,X} L} \left(\alpha_X |B\rangle^D + \beta_X |A\rangle^D\right) = \sum_{X=B,A} \sum_{\sigma=+,-} C_\sigma^{1,X} e^{\sigma i \bar{k}_X L} |X\rangle^D, \tag{F.28}$$

$$\sum_{X=B,A}\sum_{\sigma=+,-}\sigma i\bar{q}_{1,X}B_{\sigma}^{1,X}e^{\sigma i\bar{q}_{1,X}L}\left(\alpha_X\left|B\right\rangle^D+\beta_X\left|A\right\rangle^D\right)=\sum_{X=B,A}\sum_{\sigma=+,-}\sigma i\bar{k}_X C_{\sigma}^{1,X}e^{\sigma i\bar{k}_X L}\left|X\right\rangle^D. \tag{F.29}$$

By writing these equations in matrix form, we obtain

$$\begin{pmatrix} 1 & 1 & 0 & 0 \\ i\bar{k}_B & -i\bar{k}_B & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & i\bar{k}_A & -i\bar{k}_A \end{pmatrix}\begin{pmatrix} A_+^{1,B} \\ A_-^{1,B} \\ A_+^{1,A} \\ A_-^{1,A} \end{pmatrix}=\begin{pmatrix} \alpha_B & \alpha_B & \alpha_A & \alpha_A \\ i\bar{q}_{1,B}\alpha_B & -i\bar{q}_{1,B}\alpha_B & i\bar{q}_{1,A}\alpha_A & -i\bar{q}_{1,A}\alpha_A \\ \beta_B & \beta_B & \beta_A & \beta_A \\ i\bar{q}_{1,B}\beta_B & -i\bar{q}_{1,B}\beta_B & i\bar{q}_{1,A}\beta_A & -i\bar{q}_{1,A}\beta_A \end{pmatrix}\begin{pmatrix} B_+^{1,B} \\ B_-^{1,B} \\ B_+^{1,A} \\ B_-^{1,A} \end{pmatrix}, \tag{F.30}$$

$$\begin{aligned}&\begin{pmatrix} \alpha_B e^{i\bar{q}_{1,B}L} & \alpha_B e^{-i\bar{q}_{1,B}L} & \alpha_A e^{i\bar{q}_{1,A}L} & \alpha_A e^{-i\bar{q}_{1,A}L} \\ i\bar{q}_{1,B}\alpha_B e^{i\bar{q}_{1,B}L} & -i\bar{q}_{1,B}\alpha_B e^{-i\bar{q}_{1,B}L} & i\bar{q}_{1,A}\alpha_A e^{i\bar{q}_{1,A}L} & -i\bar{q}_{1,A}\alpha_A e^{-i\bar{q}_{1,A}L} \\ \beta_B e^{i\bar{q}_{1,B}L} & \beta_B e^{-i\bar{q}_{1,B}L} & \beta_A e^{i\bar{q}_{1,A}L} & \beta_A e^{-i\bar{q}_{1,A}L} \\ i\bar{q}_{1,B}\beta_B e^{i\bar{q}_{1,B}L} & -i\bar{q}_{1,B}\beta_B e^{-i\bar{q}_{1,B}L} & i\bar{q}_{1,A}\beta_A e^{i\bar{q}_{1,A}L} & -i\bar{q}_{1,A}\beta_A e^{-i\bar{q}_{1,A}L} \end{pmatrix}\begin{pmatrix} B_+^{1,B} \\ B_-^{1,B} \\ B_+^{1,A} \\ B_-^{1,A} \end{pmatrix} \\ &=\begin{pmatrix} e^{i\bar{k}_B L} & e^{-i\bar{k}_B L} & 0 & 0 \\ i\bar{k}_B e^{i\bar{k}_B L} & -i\bar{k}_B e^{-i\bar{k}_B L} & 0 & 0 \\ 0 & 0 & e^{i\bar{k}_A L} & e^{-i\bar{k}_A L} \\ 0 & 0 & i\bar{k}_A e^{i\bar{k}_A L} & -i\bar{k}_A e^{-i\bar{k}_A L} \end{pmatrix}\begin{pmatrix} C_+^{1,B} \\ C_-^{1,B} \\ C_+^{1,A} \\ C_-^{1,A} \end{pmatrix},\end{aligned} \tag{F.31}$$

from which we now read off the matrices $\boldsymbol{\mu}_{0,1}$, $\boldsymbol{\mu}_{0,2}$, $\boldsymbol{\mu}_{L,1}$, and $\boldsymbol{\mu}_{L,2}$ according to their definitions given by Eq. (F.24) and Equation (F.25). Introducing the 2×2 matrix

$$\nu_x\left(k\right)=\begin{pmatrix} \exp\left(ikx\right) & \exp\left(-ikx\right) \\ ik\exp\left(ikx\right) & -ik\exp\left(-ikx\right) \end{pmatrix}, \tag{F.32}$$

the matrices $\boldsymbol{\mu}_{x,i}$ can be written in a compact form as follows:

$$\boldsymbol{\mu}_{x,1}=\begin{pmatrix} \nu_x\left(\bar{k}_B\right) & 0 \\ 0 & \nu_x\left(\bar{k}_A\right) \end{pmatrix}, \tag{F.33}$$

$$\boldsymbol{\mu}_{x,2}=\begin{pmatrix} \alpha_B\nu_x\left(\bar{q}_{1,B}\right) & \alpha_A\nu_x\left(\bar{q}_{1,A}\right) \\ \beta_B\nu\left(\bar{q}_{1,B}\right) & \beta_A\nu\left(\bar{q}_{1,A}\right) \end{pmatrix}. \tag{F.34}$$

Note that in the presented form, the parameters of the model are over-determined. To get rid of the redundant parameter, one can introduce a characteristic energy scale E_{char} by the following definition

$$E_{\text{char}}=\frac{\hbar^2}{2mL^2} \tag{F.35}$$

so that the kinetic energy becomes

$$E_{\text{kin}}=E_{\text{char}}k'^2, \tag{F.36}$$

where k' is the dimensionless wave vector resulting from

$$k=\frac{k'}{L}. \tag{F.37}$$

In the same way dimensionless coordinates x' are defined by

$$x = x'L. \tag{F.38}$$

The Hamiltonian in Eq. (F.1) can now be expressed with respect to the energy scale E_{char} as follows:

$$H' = \frac{H}{E_{\text{char}}} = k'^2 \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} -\Delta' & -t' & 0 & 0 \\ -t' & \Delta' & 0 & 0 \\ 0 & 0 & -\Delta' & -t' \\ 0 & 0 & -t' & \Delta' \end{pmatrix} + U' \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \theta(x')\,\theta(1-x'), \tag{F.39}$$

where Δ', t', and U' are the splitting, tunneling, and coupling energies in units of E_{char}, respectively.

F.2 Eigenvalues of the reduced density matrix

In this appendix we present the details of the derivation of Eq. (6.27) in Sec. 6.4.3, which is the basis of the relation between the visibility and the von Neumann entropy. We begin by deriving the reduced density matrix of the first particle from the density matrix of the final state

$$\rho = \begin{pmatrix} E_{0,0}E^*_{0,0} & E_{0,0}E^*_{0,1} & E_{0,0}E^*_{1,0} & E_{0,0}E^*_{1,1} \\ E_{0,1}E^*_{0,0} & E_{0,1}E^*_{0,1} & E_{0,1}E^*_{1,0} & E_{0,1}E^*_{1,1} \\ E_{1,0}E^*_{0,0} & E_{1,0}E^*_{0,1} & E_{1,0}E^*_{1,0} & E_{1,0}E^*_{1,1} \\ E_{1,1}E^*_{0,0} & E_{1,1}E^*_{0,1} & E_{1,1}E^*_{1,0} & E_{1,1}E^*_{1,1} \end{pmatrix}. \tag{F.40}$$

To obtain the reduced density matrix ρ_{red} we have to trace out the degrees of freedom of the second particle, i.e.

$$\rho_{\text{red}} = \operatorname{Tr}_2 \rho. \tag{F.41}$$

For this purpose, we define four submatrices ρ_{ij} $(i, j = 0, 1)$ so that ρ is given by

$$\rho = \begin{pmatrix} \rho_{00} & \rho_{10} \\ \rho_{01} & \rho_{11} \end{pmatrix}. \tag{F.42}$$

These submatrices ρ_{ij} have the following form:

$$\rho_{00} = \begin{pmatrix} E_{0,0}E^*_{0,0} & E_{0,0}E^*_{0,1} \\ E_{0,1}E^*_{0,0} & E_{0,1}E^*_{0,1} \end{pmatrix}, \qquad \rho_{10} = \begin{pmatrix} E_{0,0}E^*_{1,0} & E_{0,0}E^*_{1,1} \\ E_{0,1}E^*_{1,0} & E_{0,1}E^*_{1,1} \end{pmatrix}, \tag{F.43}$$

$$\rho_{01} = \begin{pmatrix} E_{1,0}E^*_{0,0} & E_{1,0}E^*_{0,1} \\ E_{1,1}E^*_{0,0} & E_{1,1}E^*_{0,1} \end{pmatrix}, \qquad \rho_{11} = \begin{pmatrix} E_{1,0}E^*_{1,0} & E_{1,0}E^*_{1,1} \\ E_{1,1}E^*_{1,0} & E_{1,1}E^*_{1,1} \end{pmatrix}. \tag{F.44}$$

The trace over the degrees of freedom of the second particle can then be written as

$$\operatorname{Tr}_2 \rho = \begin{pmatrix} \operatorname{Tr}\rho_{00} & \operatorname{Tr}\rho_{10} \\ \operatorname{Tr}\rho_{01} & \operatorname{Tr}\rho_{11} \end{pmatrix} = \begin{pmatrix} E_{0,0}E^*_{0,0} + E_{0,1}E^*_{0,1} & E_{0,0}E^*_{1,0} + E_{0,1}E^*_{1,1} \\ E_{1,0}E^*_{0,0} + E_{1,1}E^*_{0,1} & E_{1,0}E^*_{1,0} + E_{1,1}E^*_{1,1} \end{pmatrix}. \tag{F.45}$$

We define transmission probabilities $T_{|i\rangle} = |E_{i,0}|^2 + |E_{i,1}|^2$ and the coherence $Z = E_{0,0}E^*_{1,0} + E_{0,1}E^*_{1,1}$ so that the reduced density matrix can eventually be expressed in the following form:

$$\rho_{\text{red}} = \begin{pmatrix} T_{|0\rangle} & Z \\ Z^* & T_{|1\rangle} \end{pmatrix}. \tag{F.46}$$

The eigenvalues of ρ_{red} are found to be

$$p_{\pm} = \frac{1}{2}\left(1 \pm \sqrt{4\left(|Z|^2 - T_{|0\rangle}T_{|1\rangle}\right) + 1}\right). \tag{F.47}$$

We now show that the eigenvalues $p_{\pm}$ can be expressed in terms of the visibility pattern according to Eq. (6.27). For this, we calculate in the first step explicit expressions for the transmission probabilities $T_{|i\rangle}$. In Sec. 6.4.3 we have argued that the final state probability amplitudes $E_{i,j}$ are given by

$$E_{0,j} = \frac{1}{\sqrt{2}}\left(\exp\left(i\phi\right)A_{0,j} + A_{1,j}\right), \tag{F.48}$$

$$E_{1,j} = \frac{1}{\sqrt{2}}\left(-\exp\left(i\phi\right)A_{0,j} + A_{1,j}\right) \tag{F.49}$$

in terms of probability amplitudes $A_{i,j}$ that characterize the state after the application of the gates $\hat{R}$ and $\hat{V}$ and that the most general ansatz for $A_{i,j}$ is given by

$$|A_{i,0}|^2 = \frac{1}{2}\cos^2\left(\frac{1}{2}\left(\gamma \pm \sigma\right)\right), \tag{F.50}$$

$$|A_{i,1}|^2 = \frac{1}{2}\sin^2\left(\frac{1}{2}\left(\gamma \pm \sigma\right)\right). \tag{F.51}$$

Here and in the following $+$ and $-$ correspond to 0 and 1, respectively. From Eq. (F.48) and Eq. (F.49) we determine the probabilities $|E_{i,j}|^2$

$$\begin{aligned} |E_{0,j}|^2 &= \frac{1}{\sqrt{2}}\left(\exp\left(-i\phi\right)A_{0,j}^* + A_{1,j}^*\right)\frac{1}{\sqrt{2}}\left(\exp\left(i\phi\right)A_{0,j} + A_{1,j}\right) \\ &= \frac{1}{2}\Big(A_{0,j}^*A_{0,j} + A_{1,j}^*A_{1,j} \\ &\qquad + |A_{0,j}|\,|A_{1,j}|\left[\exp\left(i\left(\phi+\delta_j\right)\right) + \exp\left(-i\left(\phi+\delta_j\right)\right)\right]\Big) \\ &= \frac{1}{2}\left(|A_{0,j}|^2 + |A_{1,j}|^2 + 2\,|A_{0,j}|\,|A_{1,j}|\cos\left(\phi+\delta_j\right)\right) \end{aligned} \tag{F.52}$$

and analogously

$$|E_{1,j}|^2 = \frac{1}{2}\left(|A_{0,j}|^2 + |A_{1,j}|^2 - 2\,|A_{0,j}|\,|A_{1,j}|\cos\left(\phi+\delta_j\right)\right). \tag{F.53}$$

The phases δ_j are defined by

$$\delta_j = \arg A_{0,j} - \arg A_{1,j}. \tag{F.54}$$

Thus

$$\begin{aligned} T_{|i\rangle} &= \frac{1}{2}\Big(|A_{0,0}|^2 + |A_{1,0}|^2 + |A_{0,1}|^2 + |A_{1,1}|^2 \\ &\qquad \pm 2\left(|A_{0,0}|\,|A_{1,0}|\cos\left(\phi+\delta_0\right) + |A_{0,1}|\,|A_{1,1}|\cos\left(\phi+\delta_1\right)\right)\Big) \\ &= \frac{1}{2} \pm |A_{0,0}|\,|A_{1,0}|\cos\left(\phi+\delta_0\right) \pm |A_{0,1}|\,|A_{1,1}|\cos\left(\phi+\delta_1\right), \end{aligned} \tag{F.55}$$

where we have used that $|A_{0,0}|^2+|A_{1,0}|^2+|A_{0,1}|^2+|A_{1,1}|^2=1$. With Eq. (F.50) and Eq. (F.51) we obtain

$$\begin{aligned} T_{|i\rangle} &= \frac{1}{2}\pm\frac{1}{2}\cos\left(\frac{1}{2}(\gamma+\sigma)\right)\cos\left(\frac{1}{2}(\gamma-\sigma)\right)\cos(\phi+\delta_0) \\ &\quad \pm\frac{1}{2}\sin\left(\frac{1}{2}(\gamma+\sigma)\right)\sin\left(\frac{1}{2}(\gamma-\sigma)\right)\cos(\phi+\delta_1). \end{aligned} \tag{F.56}$$

For convenience, we introduce the following abbreviations: $\alpha=\frac{1}{2}(\gamma+\sigma)$, $\beta=\frac{1}{2}(\gamma-\sigma)$, $\phi_j'=\phi+\delta_j$, and $\bar{\delta}=\frac{1}{2}(\delta_0-\delta_1)$ so that $T_{|i\rangle}$ is now given by

$$\begin{aligned} T_{|i\rangle} &= \frac{1}{2}\left(1\pm[\cos\alpha\cos\beta\cos\phi_0'+\sin\alpha\sin\beta\cos\phi_1']\right) \\ &= \frac{1}{2}\left(1\pm\left[\frac{1}{2}(\cos\sigma+\cos\gamma)\cos\phi_0'+\frac{1}{2}(\cos\sigma-\cos\gamma)\cos\phi_1'\right]\right) \\ &= \frac{1}{2}\left(1\pm\left[\frac{1}{2}\cos\sigma(\cos\phi_0'+\cos\phi_1')+\frac{1}{2}\cos\gamma(\cos\phi_0'-\cos\phi_1')\right]\right) \\ &= \frac{1}{2}\left(1\pm\left[\cos\sigma\cos\bar{\delta}\cos\left(\phi+\frac{\delta_0+\delta_1}{2}\right)+\cos\gamma\sin\bar{\delta}\sin\left(\phi+\frac{\delta_0+\delta_1}{2}\right)\right]\right) \\ &= \frac{1}{2}\left(1\pm\sqrt{\cos^2\sigma\cos^2\bar{\delta}+\cos^2\gamma\sin^2\bar{\delta}}\sin\bar{\phi}\right), \end{aligned} \tag{F.57}$$

where

$$\bar{\phi}=\phi+\frac{\delta_0+\delta_1}{2}+\arctan\left(-\frac{\cos\sigma\cos\bar{\delta}}{\cos\gamma\sin\bar{\delta}}\right). \tag{F.58}$$

Thus we have found that the transmission probabilities $T_{|0\rangle}$ and $T_{|1\rangle}$ show a sinusoidal oscillation as a function of the phase angle $\bar{\phi}$, and that the visibility of this oscillation is given by

$$v=\sqrt{\cos^2\sigma\cos^2\bar{\delta}+\cos^2\gamma\sin^2\bar{\delta}}. \tag{F.59}$$

The second step is to evaluate the expression $|Z|^2-T_{|0\rangle}T_{|1\rangle}$ that occurs in Eq. (6.27) for the eigenvalues of the reduced density matrix

$$\begin{aligned} |Z|^2-T_{|0\rangle}T_{|1\rangle} &= \left(E_{0,0}E_{1,0}^*+E_{0,1}E_{1,1}^*\right)\left(E_{0,0}^*E_{1,0}+E_{0,1}^*E_{1,1}\right) \\ &\quad -\left(|E_{0,0}|^2+|E_{0,1}|^2\right)\left(|E_{1,0}|^2+|E_{1,1}|^2\right) \\ &= 2\,\mathrm{Re}\,E_{0,0}E_{0,1}^*E_{1,1}E_{1,0}^*-\left(|E_{0,0}|^2|E_{1,1}|^2+|E_{0,1}|^2|E_{1,0}|^2\right). \end{aligned} \tag{F.60}$$

For the first term in Eq. (F.60) we obtain

$$\begin{aligned} 2\,\mathrm{Re}\,E_{0,0}E_{0,1}^*E_{1,1}E_{1,0}^* &= \frac{1}{2}\mathrm{Re}\left[(\exp(i\phi)A_{0,0}+A_{1,0})\left(-\exp(-i\phi)A_{0,0}^*+A_{1,0}^*\right)\right. \\ &\qquad \left.\times(-\exp(i\phi)A_{0,1}+A_{1,1})\left(\exp(-i\phi)A_{0,1}^*+A_{1,1}^*\right)\right] \\ &= \frac{1}{2}\left[\left(|A_{0,0}|^2-|A_{1,0}|^2\right)\left(|A_{0,1}|^2-|A_{1,1}|^2\right)\right. \\ &\qquad \left.+4|A_{1,0}||A_{0,0}||A_{0,1}||A_{1,1}|\sin(\phi+\delta_0)\sin(\phi+\delta_1)\right] \\ &= \frac{1}{8}\left(\cos^2\alpha-\cos^2\beta\right)\left(\sin^2\alpha-\sin^2\beta\right) \\ &\quad +\frac{1}{8}\sin2\alpha\sin2\beta\sin(\phi+\delta_0)\sin(\phi+\delta_1). \end{aligned} \tag{F.61}$$

With Eq. (F.52) and Eq. (F.53) we obtain for the second term in Eq. (F.60)

$$\begin{aligned}
-\left(|E_{0,0}|^2\,|E_{1,1}|^2+|E_{0,1}|^2\,|E_{1,0}|^2\right) &= -\frac{1}{4}\left(\frac{1}{2}\cos^2\alpha+\frac{1}{2}\cos^2\beta+\cos\beta\cos\alpha\cos(\phi+\delta_0)\right)\\
&\quad\times\left(\frac{1}{2}\sin^2\alpha+\frac{1}{2}\sin^2\beta-\sin\beta\sin\alpha\cos(\phi+\delta_1)\right)\\
&\quad-\frac{1}{4}\left(\frac{1}{2}\cos^2\alpha+\frac{1}{2}\cos^2\beta-\cos\beta\cos\alpha\cos(\phi+\delta_0)\right)\\
&\quad\times\left(\frac{1}{2}\sin^2\alpha+\frac{1}{2}\sin^2\beta+\sin\beta\sin\alpha\cos(\phi+\delta_1)\right)\\
&= -\frac{1}{8}\left(\cos^2\alpha+\cos^2\beta\right)\left(\sin^2\alpha+\sin^2\beta\right)\\
&\quad+\frac{1}{8}\sin 2\alpha\sin 2\beta\cos(\phi+\delta_0)\cos(\phi+\delta_1)\,. \qquad \text{(F.62)}
\end{aligned}$$

Obtaining $|Z|^2 - T_{|0\rangle}T_{|1\rangle}$ in the desired form as given by Eq. (6.35) requires a couple of transformations by means of relations between trigonometric functions

$$\begin{aligned}
|Z|^2-T_{|0\rangle}T_{|1\rangle} &= -\frac{1}{4}\left(\cos^2\alpha\sin^2\beta+\cos^2\beta\sin^2\alpha\right)+\frac{1}{8}\sin 2\alpha\sin 2\beta\cos 2\bar{\delta}\\
&= \frac{1}{8}\left(\sin 2\alpha\sin 2\beta\cos 2\bar{\delta}+\cos 2\alpha\cos 2\beta-1\right)\\
&= \frac{1}{8}\left(\frac{1}{2}\left(\cos 2\sigma-\cos 2\gamma\right)\cos 2\bar{\delta}+\frac{1}{2}\left(\cos 2\sigma+\cos 2\gamma\right)-1\right)\\
&= \frac{1}{8}\left(\cos 2\sigma\cos^2\bar{\delta}+\cos 2\gamma\sin^2\bar{\delta}-1\right)\\
&= \frac{1}{4}\left(\cos^2\sigma\cos^2\bar{\delta}+\cos^2\gamma\sin^2\bar{\delta}-1\right)\\
&= \frac{1}{4}\left(v^2-1\right). \qquad \text{(F.63)}
\end{aligned}$$

The eigenvalues $p_\pm$ of the reduced density matrix can now be written as follows:

$$\begin{aligned}
p_\pm &= \frac{1}{2}\left(1\pm\sqrt{4\left(|Z|^2-T_{|0\rangle}T_{|1\rangle}\right)+1}\right)\\
&= \frac{1}{2}\left(1\pm\sqrt{4\frac{1}{4}\left(v^2-1\right)+1}\right)\\
&= \frac{1}{2}\left(1\pm v\right). \qquad \text{(F.64)}
\end{aligned}$$

Publication List

- *Calculation of carrier transport through quantum dot molecules*, T. Zibold, M. Sabathil, D. Mamaluy, and P. Vogl, AIP Conf. Proc. 722, 799 (2005).
- *Aluminum arsenide cleaved-edge overgrown quantum wires*, J. Moser, T. Zibold, D. Schuh, M. Bichler, F. Ertl, G. Abstreiter, M. Grayson, S. Roddaro, and V. Pellegrini, Appl. Phys. Lett. 87, 052101 (2005).
- *Contact block reduction method for ballistic transport and carrier densities of open nanostructures*, D. Mamaluy, D. Vasileska, M. Sabathil, T. Zibold, and P. Vogl, Phys. Rev. B 71, 245321 (2005).
- *Prediction of entanglement detection by I-V characteristics*, T. Zibold, P. Vogl, and A. Bertoni, Proceedings of the 14th International Conference on Nonequilibrium Carrier Dynamics in Semiconductors, M. Saraniti and U. Ravaioli, eds., Chicago, USA, July 25-19, 2005, pp. 15-18, Springer Proceedings in Physics, vol. 110.
- *The 3D nanometer device project nextnano: Concepts, Methods, Results*, A. Trellakis, T. Zibold, T. Andlauer, S. Birner, R. K. Smith, R. Morschl, P. Vogl, J. Comput. Electron. (2006).
- *Modeling of semiconductor nanostructures with nextnano^3*, S. Birner, S. Hackenbuchner, M. Sabathil, G. Zandler, J. A. Majewski, T. Andlauer, T. Zibold, R. Morschl, A. Trellakis, P. Vogl, Acta Physica Polonica A 110 (2), 111 (2006).
- *nextnano: General Purpose 3D Simulations,* S. Birner, T. Zibold, T. Andlauer, T. Kubis, A. Trellakis, and P. Vogl, submitted to IEEE Transactions on Electron Devices.
- *Gauge invariant discretization in multiband envelope function theory and g-factors in nanowire dots*, R. Morschl, T. Andlauer, S. Birner, A. Trellakis, T. Zibold and P. Vogl, in preparation.
- *Theory of quantum computation with all-electric Mach-Zehnder interferometers and double quantum dots,* T. Zibold, P. Vogl, and A. Bertoni, in preparation.

Bibliography

[1] R. P. Feynman, Int. J. Theor. Phys. **21**, 467 (1982).

[2] P. Shor, Algorithms for quantum computation: discrete logarithms and factoring, *Proceedings, 37th Annual Symposium on Foundations of Computer Science* (IEEE Press, Los Alamitos, 1994).

[3] L. M. K. Vandersypen, M. Steffen, G. Breyta, C. S. Yannoni, M. H. Sherwood, I. L. Chuang, Nature **414**, 883 (2001).

[4] F. H. L. Koppens, C. Buizert, K. J. Tielrooij, I. T. Vink, K. C. Nowack, T. Meunier, L. P. Kouwenhoven and L. M. K. Vandersypen, Nature **442**, 766 (2006).

[5] J. R. Petta, A. C. Johnson, J. M. Taylor, E. A. Laird, A. Yacoby, M. D. Lukin, C. M. Marcus, M. P. Hanson, A. C. Gossard, Science **309,** 2180 (2005).

[6] M. Sabathil, Dissertation, Technische Universität München (2004).

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

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

[9] G. Dresselhaus, A. F. Kuip, and C. Kittel, Phys. Rev. **98**, 368 (1955).

[10] C. R. Pidgeon and R. N. Brown, Phys. Rev. **146**, 575 (1966).

[11] G. Bastard, Phys. Rev. B **24**, 5693 (1981).

[12] M. G. Burt, J. Phys.: Condens. Matter **4**, 6651 (1992).

[13] M. G. Burt, J. Phys.: Condens. Matter **11**, R53 (1999).

[14] T. Andlauer, Diploma thesis, Technische Universität München (2004).

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

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

[17] C. G. van de Walle and R. M. Martin, Phys. Rev. Lett. **62**, 2028 (1989).

[18] S.-H. Wei and A. Zunger, Appl. Phys. Lett. **72**, 2011 (1998).

[19] S. Hackenbuchner, Dissertation, Technische Universität München (2002).

[20] K. W. Boer, *Survey of Semiconductor Physics*, Volume 2 (Van Nostrand Reinhold, New York, 1992).

[21] S. Datta, *Electronic Transport in Mesoscopic Systems* (Cambridge University Press, Cambridge, England, 1995).

[22] D. Mamaluy, D. Vasileska, M. Sabathil, T. Zibold, and P. Vogl, Phys. Rev. B **71**, 245321 (2005).

[23] J. Altenbach and H. Altenbach, *Einführung in die Kontinuumsmechanik*, (B. G. Teubner, Stuttgart, 1994).

[24] C. Herring, E. Vogt, Phys. Rev. **101**, 944 (1956).

[25] A. Greenbaum, *Iterative Methods for Solving Linear System* (SIAM,Philadelphia, 1997).

[26] Z. Bai, J. Demmel, J. Dongarram A. Ruhe, and H. van der Vorst, editors. *Templates for the Solution of Algebraic Eigenvalue Problems: A Practical Guide* (SIAM, Philadelphia, 2000).

[27] A. Trellakis, A. T. Galick, A. Pacelli, and U. Ravaioli, J. Appl. Phys. **81**, 7880 (1997).

[28] M. R. Hestenes and E. Stiefel, J. Res. Nat. Bur. Stand. **49**, 409 (1952).

[29] C. Lanczos, J. Res. Nat. Bur. Stand. **49**, 33 (1952).

[30] T. F. Chan and T. Szeto, SIAM J. Sci. Comput. **17**, 1491 (1996).

[31] J. A. Meijernick and H. A. van der Vorst, Math. Comp. **31**, 148 (1977).

[32] T. Dupont, R. P. Kendall, and H. H. Rachford Jr., SIAM J. Numerical Analysis **5**, 559 (1968).

[33] See `http://www.caam.rice.edu/software/ARPACK/` for obtaining the ARPACK libraries and related publications.

[34] G. L. G. Sleijpen and H. A. van der Vorst, SIAM J. Matrix Anal. Appl. **17**, 401 (1996).

[35] R. K. Smith, *private communications.*

[36] See `http://www.netlib.org/blas/` for obtaining the BLAS libraries and related publications.

[37] See `http://www.netlib.org/lapack/` for obtaining the LAPACK libraries and related publications.

[38] J. R. Levine, T. Mason, and D. Brown, *lex & yacc* (O'Reilly & Associates, Sebastopol, 1992).

[39] N. W. Ashcroft and N. D. Mermin, *Solid State Physics* (Holt-Saunders, New York, 1976).

[40] G. Gilat and L. J. Raubenheimer, Phys. Rev. B **111**, 390 (1966).

[41] O. Jespen and O. K. Andersen, Solid State Commun. **9**, 1763 (1971).

[42] M . Taut and G. Lehmann, Phys. Status Solidi b **54**, 469 (1972).

[43] J. Rath and A. J. Freeman, Phys. Rev. B **11**, 2109 (1975).

[44] S. Kaprzyk and P.E. Minjarends, J. Phys. C: Solid State Phys. **19**, 1283 (1986).

[45] S. I. Kurganskii, O. I. Dubrovskii, and E. P. Domashevskaya, Phys. Status Solidi b **129**, 293 (1985).

[46] J. A. Ashraff and P. D. Loly, J. Phys. C: Solid State Phys. **20**, 4823 (1987).

[47] A. B. Chen, Phys. Rev. B **6**, 3291 (1977).

[48] J. F. Cooke and R. Wood, Phys. Rev. B **5**, 1276 (1972).

[49] A. H. MacDonald, S. H. Vosko, P. T. Coleridge, J. Phys. C: Solid State Phys. **12**, 2991, (1979).

[50] M. S. Methfessel, M. H. Boon, and F. M. Mueller, J. Phys. C: Solid State Phys. **16**, L949 (1983).

[51] M. H. Boon, M. S. Methfessel, and F. M. Mueller, J. Phys. C: Solid State Phys. **19**, 5337 (1986).

[52] M. S. Methfessel, M. H. Boon, and F. M. Mueller, J. Phys. C: Solid State Phys. **20**, 1069 (1987).

[53] G. Wiesenekker, G. Te Velde, and E. J. Baerends, J. Phys. C: Solid State Phys. **21**, 4263 (1988).

[54] G. Wiesenekker and E. J. Baerends, J. Phys.: Condens. Matter **3**, 6721 (1991).

[55] F. E. Harris, J. Phys.: Condens. Matter **14**, 621 (2002).

[56] G. H. Döhler, Phys. Stat. Sol. (b) **52**, 79 (1972).

[57] G. H. Döhler, Phys. Stat. Sol. (b) **52**, 533 (1972).

[58] M. Altarelli, Phys. Rev. B **28**, 842, (1983).

[59] I. Lapushkin, A. Zakharova, S. T. Yen, and K. A. Chao, J. Phys.: Condens. Matter **16**, 4677 (2004).

[60] A. Zakharova, I, Lapushkin, K. Nilsson, S. T. Yen ,and K. A. Chao, Phys. Rev. B **73**, 125337 (2006).

[61] K. Nilsson, A. Zakharova, I. Lapushkin, S. T. Yen, and K. A. Chao, Phys Rev. B **74**, 075308 (2006).

[62] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in Fortran 90* (Cambridge University Press, Cambridge, 1996).

[63] R. Landauer, IBM J. Res. Dev. **32**, 306 (1988).

[64] R. Landauer, Physica Scripta **T42**, 110 (1992).

[65] J. Frenkel, Phys. Rev. **36**, 1604 (1930).

[66] W. Ehrenberg and H. Hönl, Z. Phys. **68**, 289 (1931).

[67] M. Büttiker, Phys. Rev. Lett. **57**, 1761 (1986).

[68] M. Büttiker, IBM J. Res. Dev. **32**, 317 (1988).

[69] E. O. Kane, in *Tunneling Phenomena in Solids*, edited by E. Burstein and S. Lundqvist, (Plenum, New York, 1969).

[70] J. N. Schulman and Y. C. Chang, Phys. Rev. B **27**, 2346 (1983).

[71] W, Frensley, Rev. Mod. Phys. **62**, 745 (1990).

[72] C. Lent and D. Kirkner, J. Appl. Phys. **67**. 6353 (1990).

[73] D. Z.-Y. Ting, E. T. Yu, and T. C. McGill, Phys. Rev. B **45**, 3583 (1992).

[74] Y. X. Liu, D. Z.-Y. Ting, and T. C. McGill, Phys. Rev. B **54**, 5675 (1996).

[75] E. S. Daniel, X. Cartoixa, W. Frensley, D. Z.-Y. Ting, and T. C. McGill, IEEE Trans. Electron. Dev. **47**, 1052 (2000).

[76] C. Strahberger and P. Vogl, Phys. Rev. B **62**, 7289 (2000).

[77] E. Polizzi, N. Ben Abdallah, O. Vanbésiena, and D. Lippens, J. Appl. Phys. **87**, 8700 (2000).

[78] E. Polizzi and N. Ben Abdallah, Phys. Rev. B **66**, 245301 (2002).

[79] S. E. Laux, A. Kumar, and M. V. Fischetti, J. Appl. Phys. **95**, 5545 (2004).

[80] L. Smrčka, Superlattices and Microstructures **8**, 221 (1990).

[81] E. R. Racec and U. Wulf, Phys. Rev. B **64**, 115318 (2001).

[82] G. A. Nemnes, U. Wulf, and P. N. Racec, J. Appl. Phys. **96**, 596 (2004).

[83] D. K. Ferry and S. M. Goodnick, *Transport in Nanostructures* (Cambridge University Press, Cambridge, England, 1997).

[84] R. Lake, G. Klimeck, R. C. Bowen, and D. Jovanovic, J. Appl. Phys. **81**, 7845 (1997).

[85] A. Svizhenko, M. P. Anantram, T. R. Govindan, B. Biegel, and R. Venugopal, J. Appl. Phys. **91**, 2343 (2002).

[86] R. Venugopal, Z. Ren, S. Datta, M. S. Lundstrom, and D. Jovanovic, J. Appl. Phys. **92**, 3730 (2002).

[87] C. Rivas and R. Lake, phys. stat. sol. (b) 239, **94** (2003).

[88] D. Mamaluy, M. Sabathil, and P. Vogl, J. Appl. Phys. **93,** 4628 (2003).

[89] D. S. Fisher and P. A. Lee, Phys. Rev. B **23**, 6851 (1981).

[90] Y. Meir and N. S. Wingreen, Phys. Rev. Lett. **68**, 2512 (1992).

[91] S. Datta, Superlattices and Microstructures **28**, 253 (2000).

[92] W. G. van der Wiel, S. De Franceschi, J. M. Elzerman, R. Fujisawa, S. Tarucha, and L. P. Kouwenhoven, Rev. Mod. Phys. **75**, 1 (2003)

[93] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, England, 2000).

[94] H. J. Krenner, S. Stufler, M. Sabathil, E. C Clark, P. Ester, M. Bichler, G. Abstreiter, J. J Finley, and A. Zrenner, New J. Phys. **7**, 184 (2005).

[95] A. Bertoni, P. Bordone, R. Brunetti, C. Jacoboni , and S. Reggiani, Phys. Rev. Lett. **84**, 5912 (2000).

[96] Rado Ionicioiu, Gehand Amaratunga, and Florin Udrea, Int. J. Mod. Phys. B **15**, 125 (2001).

[97] J. Harris, R. Akis, and D. K. Ferry, Appl. Phys. Lett. **79**, 2214 (2001).

[98] G. B. Akguc, Linda. E. Reichl, A. Shaji, and Michael. G. Snyder, Phys. Rev. A **69,** 042303 (2004).

[99] A. Ramamoorthy, J. P. Bird, and J. L. Reno, APL **89**, 013118 (2006).

[100] A. Ramamoorthy, J. P. Bird, and J. L. Reno, APL **89**, 153128 (2006).

[101] David P. DiVincenzo, Phys. Rev. A **51**, 1015 (1995).

[102] R. Ionicioiu, P. Zanardi, and F. Rossi, Phys. Rev. A **63**, 050101(R) (2001).

[103] A. Bertoni, R. Ionicioiu, P. Zanardi, F. Rossi, and C. Jacoboni, Physica B **314**, 10 (2002).

[104] A. Marchi, S. Reggiani, A. Bertoni, and M. Rudan, J. Comp. Elect. **2**, 381 (2003).

[105] A. Bertoni and S. Reggiani, Semicond. Sci. Technol. **19**, S113 (2004).

[106] Linda E. Reichl and Michael G. Snyder, Phys. Rev. A **72**, 032330 (2005).

[107] Michael G. Snyder and Linda E. Reichl, Phys. Rev. A **70**, 052330 (2004).

[108] Y. Ji, Y. Chung, S. Sprinzak, M. Heiblum, D. Mahalu, and H. Shtrikman, Nature **422**, 415 (2003).

[109] R. Loudon, in *Disorder in Condensed Matter Physics*, edited by J. A. Blackman and J. Taguena (Clarendon Press, Oxford, 1991), p. 441.

[110] E. Beham, M. Betz, S. Trumm, M. Kroutvar, Y. Ducommun, H. J. Krenner, M. Bichler, A. Leitenstorfer, J. J. Finley, A. Zrenner, and G. Abstreiter, phys. stat. sol. (c) **1**, 2131 (2004).

[111] A. Barenco, D. Deutsch, A. Ekert, and R. Josza, Phys. Rev. Lett. **74**, 4083 (1995).

[112] A. Ekert and R. Jozsa, Rev. Mod. Phys. **68**, 733 (1996).

[113] E. Biolatti, R. C. Iotti, P. Zanardi, and F. Rossi, Phys. Rev. Lett. **85**, 5647 (2000).

[114] D. Loss and D. P. DiVincenzo, Phys. Rev. A **57**, 120 (1998).

[115] A. Imamoglu, D. D. Awschalom, G. Burkhard, D. P. DiVincenzo, D. Loss, M. Sherwin, and A. Small, Phys. Rev. Lett. **83**, 4204 (1999).

[116] T. H. Stievater, X. Li, D. G. Steel, D. Gammon, D. S. Katzer, D. Park, C. Piermarocchi, and L. J. Sham, Phys. Rev. Lett. **87**, 133603 (2001).

[117] H. Kamada, H. Gotoh, J. Temmyo, T. Takagahara, and H. Ando, Phys. Rev. Lett. **87**, 246401 (2001).

[118] P. Borri, W. Langbein, S. Schneider, and U. Woggon, R. L. Sellin, D. Ouyang, and D. Bimberg, Phys. Rev. B **66,** 081306(R) (2002).

[119] A. Zrenner, E. Beham, S. Stufler, F. Findeis, M. Bichler, G. Abstreiter, Nature **418**, 612 (2002).

[120] H. Htoon, T. Takagaharam, D. Kulik, O. Baklenov, A. L. Holmes, Jr., and C. K. Shih, Phys. Rev. Lett. **88**, 087401 (2002).

[121] O. Gywat, G. Burkhard, and D. Loss, Phys. Rev. B **65**, 205329 (2002).

[122] N. Gisin, G. Ribordy, W. Tittel, and H. Zbinden, Rev. Mod. Phys. **74**, 145, (2002).

[123] O. B. Shchekin, G. Park, D. L. Huffaker, and D. G. Deppe, Appl. Phys. Lett. **77**, 466 (2000).

[124] H. Saito, K. Nisho, and S. Sugou, Appl. Phys. Lett. **78**, 267 (2001).

[125] P. M. Petroff and S. P. DenBaars, Superlattices and Microstructures **15**, 15, (1994).

[126] X. Li, Y. Wu, D. Steel, D. Gammon, T. H. Stievater, D. S. Katzer, D. Park, C. Piermarocchi, and L. J. Sham, Science **301**, 809 (2003).

[127] M. S. Sherwin, A. Imamoglu, and T. Montroy, Phys. Rev. A **60**, 3508 (1999).

[128] G. Bester, J. Shumway, and A. Zunger, Phys. Rev. Lett. **93**, 047401, (2004).

[129] Q. Xie, A. Madhukar, P. Chen, and N. P. Kobayashi, Phys. Rev. Lett. **75**, 2542 (1995).

[130] J. Tersoff, C. Teichert, and M. G. Lagally, Phys. Rev. Lett. **76**, 1675 (1996).

[131] H. J. Krenner, M. Sabathil, E. C. Clark, A. Kress, D. Schuh, M. Bichler, G. Abstreiter, and J. J. Finley, Phys. Rev. Lett. **94**, 057402 (2005).

[132] M. Borgstrom, T. Bryllert, T. Sass, B. Gustafson, L.-E. Wernersson, W. Seifert, L. Samuelson, Appl. Phys. Letters **78**, 3232 (2001).

[133] T. Bryllert, M. Borgstrom, L.-E. Wernersson, W. Seifert, L. Samuelson, Appl. Phys. Letters **82**, 2655 (2003).

[134] M. Grundmann, O. Stier, and D. Bimberg, Phys. Rev. B **52**, 11969 (1995).

[135] O. Stier, M. Grundmann, and D. Bimberg, Phys. Rev. B **59**, 5688 (1999).

[136] C. Pryor, J. Kim, L. W. Wang, A. J. Williamson, and. A. Zunger, J. Appl. Phys. **83**, 2548 (1998).

[137] C. Pryor, Phys. Rev. B **60**, 2869 (1999).

[138] S. Lee, L. Jönsson, J. W. Wilkins, G. W. Bryant, G. Klimeck, Phys. Rev. B **63**, 195318 (2001).

[139] A. Francesschetti, H. Fu, L. W. Wang, and A. Zunger, Phys. Rev. B **60**, 1819 (1999).

[140] P. W. Fry, I. E. Itskevich, D, J, Mowbray, M. S. Skolnick, J. J. Finley, J. A. Barker, E. P. O'Reilly, L. R. Wilson, I. A. Larkin, P. A. Maksym, M, Hopkins, M, Al-Khafaji, J. P. R. David, A. G. Cullis, G. Hill, and J. C. Clark, Phys. Rev. Lett. **84,** 733 (2000).

[141] N. Liu, J. Tersoff, O. Baklenov, A. L. Holmes, Jr., and C. K. Shih, Phys. Rev. Lett. **84**, 334 (2000).

[142] M. A. Migliorato, A. G. Cullis, M. Fewrn, and J. H. Jefferson, Phys. Rev. B **65**, 115316 (2002).

[143] M. A. Cusack, P. R. Briddon, and M. Jaros, Phys. Rev. B **56**, 4047 (1997).

[144] G. W. Bryant, Phys. Rev. B **44**, 3782 (1991).

[145] H. Sprekeler, G. Kießlich, A. Wacker, and E. Schöll, Phys. Rev. B **69**, 125328 (2004).

[146] K. Nishiguchin and S. Oda, Appl. Phys. Lett. **76**, 2922 (2000).

[147] K. Nishiguchin and S. Oda, J. Appl. Phys. **92**, 1399 (2002).

[148] S. L. Wang, P.C. can Son, B. J. Van Wees, and T. M. Klapwijk, Phys. Rev. B **46**, 12873 (1992).

[149] D. Többen, D. A. Wharam, G. Abstreiter, J. P. Kotthaus, and F. Schaffler, Semicond. Sci. Technol. **10**, 711 (1995).

[150] D. Többen, D. A. Wharam, G. Abstreiter, J. P. Kotthaus, and F. Schaffler, Phys. Rev. B **52**, 4704 (1995).

[151] U. Wieser, U. Kunze, K. Ismail, and J. O. Chu, Appl. Phys. Lett, **81**, 1726 (2002).

[152] S. Adachi, J. Appl. Phys. **58**, R1 (1985).

[153] H. W. van Kesteren, E. C. Cosman, P. Dawson, K. J. Moore, and C. T. Foxon, Phys. Rev. B **39**, 13426 (1989).

[154] A. F. W. van de Stadt, P. M. Koenraad, J. A. A. J. Perenboom, and J. H. Wolter, Surf. Sci. **361/362**, 521 (1996).

[155] J. Moser, T. Zibold, D. Schuh, M. Bichler, F. Ertl, G. Abstreiter, M. Grayson, S. Roddaro, and V. Pellegrini, Appl. Phys. Lett. **87**, 052101 (2005).

[156] S. Dasgupta, A. Fontcuberta i Morral, M. Bichler, G. Abstreiter, and M. Grayson, *in preparation.*

[157] A. Yacoby, H. L. Stormer, K. W. Baldwin, L. N. Pfeiffer, and K. W. West, Solid State Commun. **101**, 77 (1997).

[158] E. P. De Poortere, Y. P. Shkolnikov, E. Tutuc, S. J. Papadakis, and M. Shayegan, Appl. Phys. Lett. **80**, 1583 (2002).

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

[160] S. Selberherr, A. Schütz, and H. W. Pötzl, IEEE J. Solid-State Circuits **15**, 605 (1980).

[161] N. D. Arora, J. R. Hauser, and D. J. Roulston, IEEE Trans. Electron Devices **29**, 292 (1982).

[162] G. Masetti, M. Severi, and S. Solmi, IEEE Trans. on Electron Devices **30**, 764 (1983).

[163] A. Trellakis and T. Zibold, *private communications.*

Danksagung

Ich bedanke mich bei

Prof. Peter Vogl dafür, dass er mir diese Arbeit ermöglicht hat und mir dabei stets mit viel Unterstützung, Anleitung und großzügiger Förderung zur Seite stand.

meinem Kollegen Dr. Alex Trellakis, von dem ich durch unsere gemeinsame Arbeit an next**nano**++ viel über die Simulation von Halbleiterbauelementen und insbesondere über die dabei verwendeten numerischen Methoden gelernt habe.

Dr. Andrea Bertoni für die gute Zusammenarbeit und vielen Diskussionen rund um das Thema Qubits in Quantendrähten und Verschränkung.

Dr. Stefan Ludwig und Daniel Harbusch für ihr unerschrockenes Engagement in der experimentellen Umsetzung von Teilen des vorgeschlagenen Mach-Zehnder Interferometers.

Dr. Matthew Grayson und Dr. Joel Moser für die fruchtbare Zusammenarbeit auf dem Gebiet der AlAs Quantendrähte.

meinem Kollegen Dr. Matthias Sabathil, der mir zu Beginn meiner Doktorarbeit durch seinen großen Einsatz sehr bei der Einarbeitung in mein neues Arbeitsgebiet geholfen hat.

meinem Kollegen Till Andlauer für das stets gute Arbeitsklima in unserem gemeinsamen Büro und insbesondere für die intensive Zusammenarbeit bei der Entwicklung von next**nano**++.

meinen anderen Kollegen von T33: Stefan Birner, Tillmann Kubis und Dr. Reinhard Scholz für die vielen fachlichen und persönlichen Diskussionen.

unseren Diplomanden Thomas Eißfeller und Peter Greck für das Korrekturlesen sowie diverse arbeitserleichternde Softwarelösungen.

unseren Sekretärinnen Veronika Enter und Liane Lindner für ihre große Hilfsbereitschaft bei allen verwaltungstechnischen und organisatorischen Problemen, mit denen ich mich auseinander zu setzen hatte.

allen anderen Kollegen und Mitarbeitern des Walter Schottky Instituts für das inspirierende Umfeld und all die schönen Feste und Feiern.

meiner Freundin Dr. Petra Neff für die große Bereicherung meines Lebens, die sie für mich ist. Im Speziellen danke ich ihr auch für das genaue Korrekturlesen meiner Arbeit.

meinen Eltern für ihre immerwährende Unterstützung, ohne die mir diese Ausbildung bis hin zur Promotion nicht möglich gewesen wäre.