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

Opto-electronic and quantum transport properties of semiconductor nanostructures

Matthias Sabathil

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: ....

Prüfer der Dissertation:

1. Univ.-Prof. Dr. P. Vogl
2. ....
3. ....

Die Dissertation wurde am 29.9.2004 bei der Technischen Universität München eingereicht und durch die Fakultät für Physik am .... angenommen.

Opto-electronic and quantum transport properties of semiconductor nanostructures

Matthias Sabathil

Contents

Danksagung vii

Abstract ix

Zusammenfassung xi

Introduction xiii

I New method for ballistic quantum transport 1

1 Introduction to ballistic quantum transport 3

2 Contact block reduction (CBR) method 7
2.1 Existing methods for quantum transport calculation 7
2.2 Green's functions and transmission 8
2.3 Basic idea of CBR 9
2.4 Contact block reduction method for the density matrix 11
2.5 Numerical details and computational costs 13
2.6 Multiband lead mode dispersion and self-energy 13
2.7 Boundary condition for decoupled device Hamiltonian 15
2.8 Mode space reduction in single-band case 16
2.9 Hole transport in T-shaped junction 18
2.9.1 Lead dispersion 19
2.9.2 Transmission function and density of states 20
2.9.3 Conductance 22
2.9.4 Carrier density 23
2.9.5 k-space distribution of the density of states 24
2.10 Resonant tunneling through a 3D quantum dot molecule 26
2.11 Conclusion 27

3 Self-consistent extension of CBR 29
3.1 Introduction 29
3.2 Description of the model 29
3.2.1 Hamiltonian 30
3.2.2 Discretization of Schrödinger and Poisson equation 31
3.2.3 Boundary conditions 34
3.3 Calculation of local density of states 35

3.4 Bound states . . . 36
3.4.1 Truly bound states . . . 37
3.4.2 Quasi bound states . . . 37
3.4.3 Resonant states . . . 39
3.5 Combined density of states . . . 41
3.5.1 Calculation of the carrier density . . . 41
3.5.2 Detection of bound states . . . 42
3.6 Self-consistent solution . . . 43
3.6.1 Main loop . . . 45
3.6.2 Predictor step . . . 46
3.6.3 Energy grid . . . 47
3.7 Calculational results . . . 47
3.7.1 Self-consistent potential and carrier density . . . 49
3.7.2 IV-Characteristics . . . 51
3.7.3 Bound states . . . 54
3.7.4 Eigenstate Convergence . . . 60
3.7.5 Computational costs . . . 62
3.7.6 Exchange and Correlation potentials . . . 63
3.7.7 Varying gate oxide thickness . . . 66
3.7.8 Local density of states . . . 66
3.8 Cone-DGFET . . . 67
3.9 Conclusion . . . 71

4 Inelastic scattering via Büttiker probes 73
4.1 Introduction . . . 73
4.2 Introduction to the method . . . 74
4.3 Additional remarks . . . 77
4.4 Numerical results . . . 79
4.5 Conclusion . . . 82

5 Summary and outlook 87

II Quantum dots 89

6 Introduction 91

7 Modeling quantum dots 93
7.1 Motivation . . . 93
7.2 Dot shape . . . 94
7.3 Alloy composition . . . 95
7.4 Strain . . . 98
7.4.1 Band-lineup . . . 99
7.4.2 Piezoelectric charges . . . 99
7.5 Poisson equation . . . 100
7.6 Hamiltonian . . . 102
7.7 Few particle states . . . 103

8 Positively charged excitons in QDs **109**
8.1 Experimental results . . . 110
8.2 Theoretical results . . . 111
8.3 Conclusion . . . 114

9 Quantum dot molecules **115**
9.1 Theory . . . 115
9.1.1 Coupled quantum systems . . . 116
9.1.2 Strain . . . 118
9.1.3 Single particle states . . . 119
9.1.4 Excitons in coupled quantum dots . . . 121
9.1.5 Stark shift in coupled quantum dots . . . 122
9.2 Comparison to experimental results . . . 123
9.3 Conclusion . . . 125

A Closed system limit for Ξ **129**

B Hamiltonian for k·p calculations **131**

C Normalization of the bound states **133**

D CBR self-consistent code **135**
D.1 User guide . . . 135
D.1.1 Simulation domain . . . 136
D.1.2 Definition of contacts . . . 136
D.1.3 Voltage sweep . . . 136
D.1.4 Setup of material grid . . . 137
D.1.5 Cutoff parameters . . . 138
D.1.6 Energy grid . . . 138
D.1.7 Bound states . . . 139
D.1.8 Self-consistent cycle . . . 139
D.2 Flow chart . . . 141

E Local spin density approximation (LSDA) **143**
E.1 Exchange . . . 143
E.2 Correlation . . . 144

F Publication list **147**
F.1 Published papers . . . 147
F.2 Accepted for publication . . . 148
F.3 Submitted papers . . . 148

Danksagung

Mein ganz besonderer Dank gilt Prof. Peter Vogl für die Ermöglichung dieser Dissertation, seine fortwährende Unterstützung, Anleitung und großzügige Förderung bei dieser Arbeit, sowie für die gewährte Freiheit bei der Interpretation des Themas. Weiterhin danke ich Dr. Denis Mamaluy für die intensive Zusammenarbeit bei der Ausarbeitung der CBR Methode und für die tagelangen, intensiven Diskussionen die wesentlich zum Spass und der Begeisterung, die ich während dieser Arbeit hatte, beigetragen haben. Bei Jacek Majewski möchte ich mich für seine Unterstützung und Geduld bedanken, die man jederzeit und scheinbar unbegrenzt in Anspruch nehmen konnte. Für die interessanten Erfahrungen im Bereich Startup-Unternehmungen, die ich im Rahmen des Munich Business Plan Wettbewerbs machen konnte, danke ich Stefan Birner der unsere Teilnahme durch seine Initiative als Teamchef von next**nano**^{3} ermöglichte.

Mein Dank gilt auch meinen anderen Kollegen Dr. Alex Trellakis, Dr. Christian Strahberger, Christian Uhl, Michael Bayer, Dr. Günther Zandler, Dr. Stefan Hackenbuchner, Tillmann Kubis, Till Andlauer, Tobias Zibold, sowie in besonderem Maße unseren beiden Sekretärinnen Veronika Enter und Liane Lindner, für die angenehme und lockere Arbeitsatmosphäre in unserer Gruppe. Für die hervorragende und intensive Zusammenarbeit mit den experimentellen Gruppen insbesondere der Quantenpunkt-Gruppe von Prof. Johnathan Finley und seinen Mitarbeitern Hubert Krenner, Michael Reimer und Emily Clark, sowie bei allen anderen Kollegen und Mitarbeitern am Walter Schottky Institut möchte ich mich recht herzlich für die vier interessanten und schönen Jahre bedanken, die ich an diesem Institut verbringen durfte.

Vor allem aber, gilt mein Dank meinen Eltern, die mich auf meinem Weg immer unterstützt haben und mir diese Ausbildung ermöglichten, sowie meiner wunderbaren Frau Brigitte und meinen Kindern Lukas und Annika die mir die Augen für die wirklich wichtigen Dinge im Leben geöffnet haben.

Abstract

In this work a novel and efficient method for the calculation of the ballistic transport properties of open semiconductor nanostructures connected to external reservoirs is presented. It is based on the Green's function formalism and reduces the effort to obtain the transmission and the carrier density to a single solution of a hermitian eigenvalue problem with dimensions proportional to the size of the decoupled device and the multiple inversion of a small matrix with dimensions proportional to the size of the contacts to the leads. Using this method, the 4-band GaAs hole transport through a 2-dimensional three-terminal T-junction device, and the resonant tunneling current through a 3-dimensional InAs quantum dot molecule embedded into an InP heterostructure have been calculated. The further extension of the method into a charge self-consistent scheme enables the efficient prediction of the IV-characteristics of highly doped nanoscale field effect transistors in the ballistic regime, including the influence of quasi bound states and the exchange-correlation interaction. Büttiker probes are used to emulate the effect of inelastic scattering on the current for simple 1D devices, systematically analyzing the dependence of the density of states and the resulting self-consistent potential on the scattering strength.

The second major topic of this work is the modeling of the optical response of quantum confined neutral and charged excitons in single and coupled self-assembled InGaAs quantum dots. For this purpose the existing device simulator next**nano**^{3} has been extended to incorporate particle-particle interactions within the means of density functional theory in local density approximation. In this way the exciton transition energies for neutral and charged excitons as a function of an externally applied electric field have been calculated, revealing a systematic reduction of the intrinsic dipole with the addition of extra holes to the exciton, a finding that is in excellent agreement with experimental measurements. The field controlled coupling of the excitons in a quantum dot molecule built of two vertically stacked quantum dots is calculated, with the predicted dependence on dot separation and size-mismatch being in excellent agreement with recent experimental results.

Zusammenfassung

In dieser Arbeit wird eine neue Methode vorgestellt, die es ermöglicht, die ballistischen Quantentransport Eigenschaften einer mit externen Ladungsträger-Reservoiren verbundenen Halbleiter-Nanostruktur, sehr effizient zu berechnen. Die Methode, die auf dem Green's Funktions Formalismus basiert, reduziert den Aufwand zur Berechnung der Transmission und der Ladungsträgerdichte auf die einmalige Lösung eines hermiteschen Eigenwertproblems mit Dimensionen proportional zu der Größe des entkoppelten Systems sowie der mehrfachen Invertierung einer kleinen Matrix mit Dimensionen proportional zu Größe der Kontaktflächen zu den Reservoiren. Mit Hilfe dieser Methode wurde der 4-Band Löchertransport durch eine zweidimensionale, drei-Terminal T-junction Struktur sowie der resonante Tunnelstrom durch ein dreidimensionales InAs Quantenpunkt-Molekül, das in eine InP-Barriere eingebettet ist, berechnet. Die Erweiterung der Methode in ein Ladungs-selbstkonsistentes Verfahren ermöglicht die effiziente Vorhersage der Strom-Spannungs-Charakteristik von hoch dotierten nano Feldeffekt-Transistoren im ballistischen Regime unter Einbeziehung der Einflüsse von quasi-gebundenen Zuständen und der Austausch-Korrelations-Wechselwirkung. Der Einfluss der inelastischen Streuung, generiert durch so genannte 'Büttiker probes', auf den Strom, die Zustandsdichte und das selbst-konsistente Potential werden anhand einfacher 1D Strukturen systematisch untersucht.

Der zweite Teil dieser Arbeit befasst sich mit der Modellierung der optischen und elektronischen Eigenschaften von neutralen und geladenen Exzitonen in InGaAs Quantenpunkten. Zu diesem Zweck wurde der Device-Simulator next**nano**^{3} erweitert um die Mehrteilchen-Wechselwirkung im Rahmen des Dichte-Funktional-Formalismus zu beschreiben. Mit dieser Methode wurden die optischen Übergangsenergien von neutralen und geladenen Exzitonen in Abhängigkeit von einem äußeren elektrischen Feld berechnet. Die theoretischen Ergebnisse zeigen eine signifikante Reduktion des intrinsischen Dipols mit zunehmender positiver Ladung der Exzitonen und wurden durch experimentelle Messungen sehr gut bestätigt. Des Weiteren zeigt der Vergleich der berechneten quantenmechanischen Kopplung der Exziton-Zustände in Quantenpunkt-Molekülen, als Funktion des elektrischen Feldes, mit neuen experimentellen Ergebnissen, eine sehr gute Übereinstimmung in Bezug auf den vorhergesagten Quantenpunkt Abstand und den erwarteten Größenunterschied.

Introduction

With the continuous advances in semiconductor technology, the dimensions of heterostructures and devices fabricated in the labs of research institutions and industrial development centers have already entered the realm of quantum mechanics. Devices, such as quantum wires, quantum dots, or nano-MosFets that exhibit a purely quantum mechanical behavior and have formerly been addressed in textbooks only, are now being fabricated and subject of sophisticated experiments that measure their electronic structure, the optical response or the quantum transport properties. However, the process of the idea creation, the design and the analysis of any novel quantum structure would be unthinkable without the tool of the theoretical and numerical modeling. Because of this, the degree of sophistication of the numerical models employed in the simulation tools for semiconductor devices has been subject to an evolution parallel to the rapid advances in the fabrication technology. While for simple heterostructures, such as one-dimensional quantum wells, analytical models are mostly sufficient, other devices as there are quantum dots, nanoscale transistors or carbon-nanotubes require models that are able to predict the electronic structure and the quantum transport properties of complex multi-dimensional devices. These models do not only demand for a properly defined Hamiltonian to predict the quantum mechanical density distribution, but also have to take into account many-particle interactions or the fully relaxed strain that leads to band deformations and polarization charges, effects that are of high importance in self-assembled quantum dots or Silicon/Germanium heterostructures. With the increased complexity of the models, the computational effort rises tremendously highlighting the need for more efficient methods to keep up with the demand of modern nanoscale physics. In this work two approaches are presented that are able to predict the ballistic quantum transport properties of large three-dimensional multi-terminal devices and the optical response of neutral and charged excitons in quantum dots and molecules using a standard 'of the shelf' PC, while leading to results that are comparable to other methods that rely on heavy computations on clusters or supercomputers.

The thesis is organized in two main parts, each of which is devoted to a different observable of a quantum nanostructure. In the first part a novel and efficient method for the calculation of the ballistic transport properties of open nanostructures is presented that has been developed within this thesis. It is based on the Green's function formalism, utilizing the fact, that the calculation of the transmission and the carrier density of an open device can be split into a single diagonalization of the hermitian Hamiltonian of the device decoupled from the leads, plus a repeated solution of a small linear system of dimensions proportional to the contact size. The method, that we have termed contact block reduction (CBR) method, provides the transmission, the local density of states and the density matrix of a multi-dimensional open nanostructure with an arbitrary number of leads in the ballistic regime. In this way, the ballistic hole transport of a GaAs/AlGaAs T-junction using a multi-band-Hamiltonian and the resonant current through an InAs quantum dot molecule embedded in an InP barrier have been

calculated for the first time.

The self-consistent implementation of the CBR method, presented in the second chapter, allows for the modeling of technologically relevant devices, such as nanoscale field effect transistors, in the regime of high temperature and high bias. In contrast to previous works, the gates are fully included into the calculation, thus enabling the realistic prediction of the gate tunneling current. We find that, in addition to the density resulting from a purely ballistic approach, it is important to take into account the bound states that occur in the gate regions at the oxide interface in order to obtain the physically correct density. Finally, it is shown, that the inclusion of the exchange and correlation interaction within the local density approximation (LDA) leads to a significant change in the IV-characteristics of such highly doped devices and therefore cannot be neglected in a self-consistent calculation. As an outlook, in the third chapter, we present an evaluation of two simple models proposed in the literature that include the effects of inelastic scattering within the device in terms of the nonequilibrium Green's function formalism. The effect of scattering on the resulting current and potential of a simple nin-resistor is analyzed in terms of the scattering strength and the model employed. We conclude that the simple models cannot substitute realistic calculations, but are well suited to get a qualitative picture of the impact of the scattering in quantum nanostructures.

The second part of this thesis is dedicated to the electronic and optical properties of self-assembled quantum dots (QDs), investigating the response of neutral and charged quantum confined excitons to external electric fields. This work has been carried out in close collaboration with the experimental group of Dr. Jonathan Finley at the Walter Schottky institute, that performs state of the art single dot spectroscopy, using photoluminescence and photocurrent techniques. In order to model these experiments, the existing device simulator next**nano**^{3} has been extended to predict the properties of few particle states, described within the means of a self-consistent density functional theory (DFT) approach.

In the first chapter of this part, a step by step introduction to the key points of quantum dot modeling is provided, as there are the choice of the dot shape and alloy profile and the definition of the entire structure including substrate, wetting- and cap-layer on a finite differences grid. An overview of the underlying physical equations is provided, emphasizing the additional parts connected to the DFT-algorithm that have been developed as a part of this work. Using this model we are able to predict the density distribution of electrons and holes in neutral and charged quantum dots as a function of the external electric field. We find that the intrinsic dipole, namely the separation of the positive and negative charges at zero external field, is reduced with the addition of excess holes to the exciton, a result which is in excellent agreement to experimental data. A QD system with possible applications in quantum information storage is the quantum dot molecule (QDM) formed of two vertically stacked quantum dots, that is investigated in the last chapter. This system features a field controlled adiabatic transition of an optically active, to a dark state, leading to an enhanced exciton lifetime. Our theoretical predictions of the quantum mechanical coupling and the Stark shift of the excitonic states as a function of the dot separation were recently confirmed by precise measurements on single a QDM, highlighting the predictive capability of the employed model.

Part I

New method for ballistic quantum transport

Chapter 1

Introduction to ballistic quantum transport

Since the 1980s the semiconductor technology has reached a level of sophistication that allows for the fabrication of heterostructures with length scales in the mesoscopic regime, meaning that the length scales of the structure are below the characteristic scattering lengths of the carriers such that coherent quantum transport effects are observed. Since these length scales generally depend very strongly on the material, the temperature, and the applied electric or magnetic fields the structural dimensions may vary between a few nanometers and a micrometer. One way of reaching the mesoscopic regime without the use of nano-fabrication techniques is to reduce the scattering rates and thus enhance the characteristic length scales up to the micrometer range. Therefore, one of the key structures that first enabled quantum transport experiments is the two dimensional electron gas (2-DEG) that forms at the interface of an undoped GaAs substrate and a modulation doped AlGaAs barrier. Due to the offset of the conduction band edges and the alignment of the Fermi-levels, a triangular quantum well is formed at the junction that confines a two-dimensional layer of electrons. Since the dopant ions are located inside the barrier, the impurity scattering inside the 2-DEG is efficiently reduced leading to high mobility samples that are a prerequisite for quantum transport experiments. Metal gates that are imprinted on the surface of the device via electron beam lithography deplete the electron gas beneath and thus offer the necessary flexibility to design tiny quantum circuits. This type of structure is still the main work-horse for the fundamental research in quantum transport phenomena, such as integer and fractional quantum Hall effect. Today, the minimization of the semiconductor heterostructures has come to a point where technologically and economically important devices such as field effect transistors, that are the building blocks of every integrated circuit, are of dimensions of few tens of nanometers, where quantum effects start to play a role even at room temperature. While for these devices it suffices to take the quantum effects into account in terms of small corrections to the semi-classical limit, the simulation of future devices with gate lengths below 10 nm, which will be in production in less then ten years according to the road map of the semiconductor industry [1] clearly demands for a more sophisticated treatment of quantum transport.

One of the most intuitive approaches to quantum transport is probably the Landauer method [2] that was introduced by Landauer in 1988 but with its roots dating back to the 1930s [3][4]. The idea is, that the current through a ballistic conductor can be expressed in terms of the probability that an electron can transmit from one contact to another. This concept, that seems to be very obvious for tunnel-junctions with a low probability of transmission has been

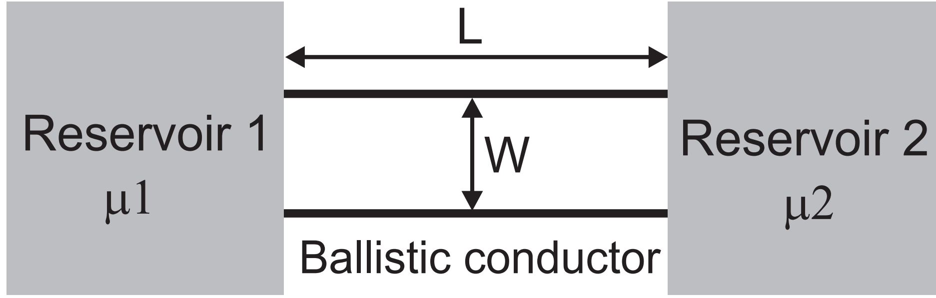


Figure 1.1: Schematic picture of a simple ballistic conductor of length L and width W, connecting to contacts 1 and 2. The contacts act as reservoirs for the carriers with well defined Fermi levels μ_1 and μ_2.

generalized to open structures with transmission close to unity by Landauer and later extended to multi-terminal structures including magnetic fields by Büttiker [5]. To explain the method in more detail we will use a simple sketch of a two-terminal structure shown in fig. 1.1 [6] that consists of a ballistic conductor of length L and width W connecting two reservoirs 1 and 2. The contacts are assumed to be much larger than the conductor such that they can be considered to be carrier reservoirs with local Fermi-levels μ_1 and μ_2. These Fermi levels can then be controlled externally to apply a bias to the structure. The ballistic conductor has a mean free path length L_m much larger then device length L, such that electrons can pass without any energy or momentum relaxation due to scattering, leading to a transmission probability equal to one for each mode. Even though each single mode has zero resistance, due to the finite width W of the conductor only a limited number of modes carry current which results in a finite current through the ballistic conductor. This effective resistance is called contact resistance since one can think of it occurring at the interface between the contacts and the conductor where the carriers have to be transmitted from the infinite number of modes in the contacts to the finite number of modes in the conductor.

The current through a ballistic conductor with M propagating modes that each have transmission probability equal to unity is determined from the product of the carrier density, the charge carried by each particle and the velocity at which the carriers propagate through the conductor. In a ballistic conductor with two terminals the net current is a balance between the influx of electrons I_1 from the first reservoir minus I_2 originating from the second reservoir. This net current I_{12} can be written as

$$I_{12} = I_1 - I_2 \tag{1.1}$$

$$I_i = \frac{2e}{L}\sum_m^M \sum_k v(E) f(\mu_i, E) dE, \quad i = 1, 2 \tag{1.2}$$

where the first sum in eq. (1.2) is over the modes M, and the second over all states with wavevector k that propagate within this mode with the direction into the device. Each state is normalized over the conductor length with $1/L$ and the spin degeneracy is taken into account with the factor of 2. With the group velocity v resulting from the dispersion $E(k)$ by

$$v(E) = \frac{1}{\hbar}\frac{dE}{dk}$$

we can rewrite eq. (1.2) into

$$I_i = \frac{2e}{L\hbar}\sum_m^M \sum_k \frac{dE}{dk} f(\mu_i, E).$$

To transform the sum over k into an integral we make use of the standard conversion

$$\sum_k \rightarrow \frac{L}{2\pi}\int dk,$$

which finally leads to

$$I_i = \frac{e}{\pi\hbar}\sum_m^M \int dk \frac{dE}{dk} f(\mu_i, E) = \frac{2e}{h}\sum_m^M \int_{\varepsilon_m}^{\infty} f(\mu_i, E) dE,$$

with ε_m being the cut-off energy of mode m. Using this expression, the net current is given by

$$I_{12} = \frac{2e}{h}\sum_m^M \int_{\varepsilon_m}^{\infty} \left[f(\mu_1, E) - f(\mu_2, E)\right] dE. \tag{1.3}$$

The constant prefactor of equation 1.3 is also called the conductance quantum

$$G_c = \frac{2e^2}{h}, \tag{1.4}$$

a universal constant with its inverse corresponding to a resistance of 12.9 kΩ. The first experimental evidence of this conductance quantum was found by van Wees *et al.* [7], who measured the conductance through a constriction imposed by a split-gate structure as depicted in the inset of figure 1.2. Two metal gates on top of a GaAs/AlGaAs heterostructure were used to deplete the two-dimensional electron gas beneath via a negative applied gate voltage. In this way the width of the channel connecting the source and drain contacts was controlled. Since the number of propagating modes M in eq. (1.3) depends strongly on the width of the channel, the measured conductance, shown in fig. 1.2 exhibits plateaus that correspond to a constant number of propagating modes. The conductance of each plateau corresponds exactly to the number of modes times the conduction quantum G_c which is the proof that ballistic quantum transport has been observed.

To obtain the ballistic transport properties of more general devices that exhibit non-unitary transmission, Landauer extended this picture of the quantized conductance by introducing a weighting factor for the transmission probability, namely the transmission function $T(E)$. With this function the ballistic current through a two-terminal device is obtained via the well known Landauer formula

$$I_{12} = \frac{2e}{h}\int T(E)\left[f(\mu_1, E) - f(\mu_2, E)\right] dE. \tag{1.5}$$

Thus, the quantum transport problem in the ballistic limit has been reduced to the calculation of the transmission function of an open device connected to the leads. In the case of quantum ballistic transport, the transmission probability T of eq. (1.5) is connected to the scattering matrix S of the system by the following relation:

$$T_{pq} = \left|S_{pq}\right|^2. \tag{1.6}$$

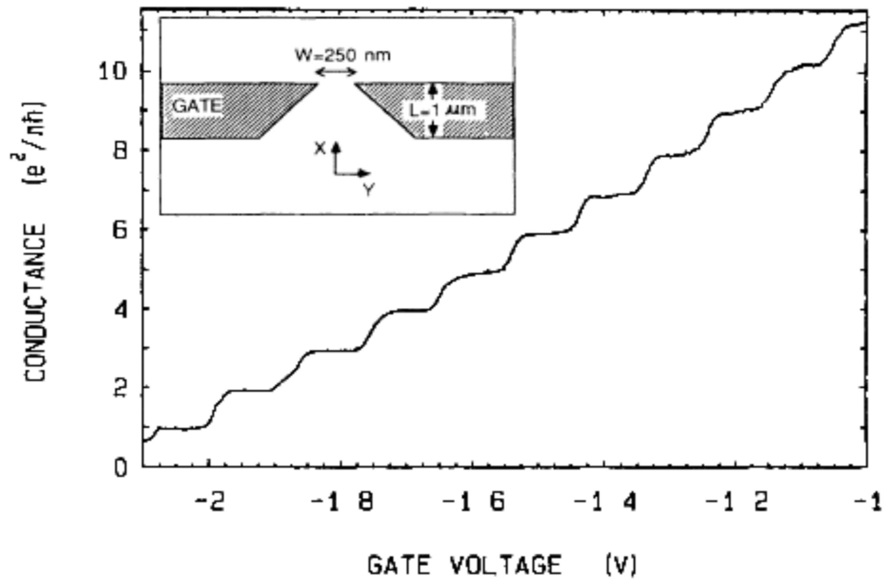


Figure 1.2: Measured conductance vs. gate voltage for the split-gate device depicted in the inset [7]. A negative voltage applied on a pair of metallic gates is used to deplete the electron gas beneath.

Therefore, by using scattering theory the relevant information needed to describe the ballistic transport of an open device can be obtained from the S-matrix. In this way the transport properties of a ballistic device under applied bias are described using equilibrium statistical mechanics to treat a special class of non-equilibrium problems for which each eigenstate remains in equilibrium, but with different reservoirs.

Chapter 2

Contact block reduction (CBR) method

2.1 Existing methods for quantum transport calculation

Several methods have been developed to calculate the ballistic transport through quantum devices that can be roughly divided into two classes. The direct approach is to calculate the S-matrix via the solution of the Schrödinger equation with scattering boundary conditions. A well-known representative of this class is the transfer matrix method [8, 9], which has been generalized to multi-dimensional systems by Frensley [10, 11], Lent *et al.* [12], and to multi-band Hamiltonians by Ting *et al.* [13, 14]. In the latter methods the scattering boundary conditions are applied via the quantum transmitting boundary method (QTBM) [12], and the transmission is obtained via the solution of a linear system with dimensions proportional to the size of the device that has to be solved repeatedly. Therefore, the published works still appear to be limited to one-dimensional tight-binding [13, 15, 16] or $\mathbf{k}{\cdot}\mathbf{p}$-based [14] multiband calculations, or two-dimensional (2D) single-band calculations [12, 17]. Three-dimensional calculations have been published [18], but are restricted to separable problems and single band Hamiltonians. Very recently, a modified version of the QTBM has been developed that expands the scattering solutions in terms of two different closed system wave functions [19]. The calculations are charge self consistent but have only been implemented for single-band situations so far. A method that is based on a similar idea as the approach presented in this work is the R-matrix method [20] that is widely used in atomic physics but has not received much attention in the semiconductor device community with the exception of the works by Wulf *et al.* [21, 22] that neglect charge self-consistency. In this scheme the S-matrix is represented in terms of the eigenfunctions of a closed system, so called Wigner-Eisenbud functions.

The second class of works is based on the Green's function method to calculate the quantum transport properties, with the coupling to the leads being introduced via the self-energy. The advantage of this approach is the well developed theory of the Green's functions that also allows considering inelastic scattering within the nonequilibrium Green's function formalism. A very efficient and widely used algorithm is the recursive Green's function method [23, 24] that has been successfully implemented for two-dimensional devices [25, 26] and for small three-dimensional structures such as nano-wires [27]. It's main advantage is that it does not only yield the retarded Green's function that is connected to the S-matrix, but also the lesser Green's function [25], which is needed if inelastic scattering is to be considered. The main drawback of the method is its restriction to devices that can be discretized into cross-sectional slices with

nearest neighbor interactions only, a condition that cannot be maintained for structures with more than two contacts, since additional contacts inevitably couple more distant slices with one another.

Thus, although a huge variety of methods has been developed in the past decades, the quantum-mechanical ballistic multi-band transport calculation of large two- and three-dimensional structures or devices with more than two Ohmic contacts still presents a significant challenge.

In this work, we present a novel and efficient Green's function method to calculate the electronic properties of open quantum systems such as the transmission, the density of states, and the carrier density in the ballistic limit. The presently developed scheme is applicable to large, multi-terminal two- and three-dimensional structures without geometric constraints and can handle single and as well as multi-band **k·p-** or tight-binding Hamiltonians.

The chapter is organized as follows. In the first sections a brief summary of the Green's function formalism is provided followed by an introduction to the CBR method. A general expression for the density matrix of an open system in terms of the eigenstates of the corresponding closed system is derived in section (2.4). The generalization of the method to multi-band Hamiltonians is presented in section (2.6). A further significant reduction of the computational effort can be achieved for the special case of single band Hamiltonians by transforming the basis into the space of propagating modes of the leads, as it is shown in section (2.8). In sections (2.9) and (2.10), respectively, we present two numerical examples. For the first time, we predict the hole conductance in a GaAs based T-junction device and the resonant current through an InAs quantum dot molecule that is embedded within InP barrier material. These examples highlight the applicability of the CBR method to multi-band **k·p** Hamiltonians and realistic three-dimensional devices.

2.2 Green's functions and transmission

A convenient way to obtain the S-matrix is via the retarded Green's function of the open device. The connection between the S-matrix and the retarded Green's function has been derived by Fisher and Lee [28]

$$S_{qp} = -\delta_{qp} + i\hbar\sqrt{v_q v_p} G^R_{qp}, \tag{2.1}$$

with the group velocities $v_{q,p}$ of the modes q and p. In this way the retarded Green's function yields the transmission between the different leads and finally the ballistic current via equation (1.5).

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

$$H_{tot} = \begin{pmatrix} H_1 & 0 & 0 & W_1 \\ 0 & \ddots & 0 & \vdots \\ 0 & 0 & H_L & W_L \\ W_1^\dagger & \cdots & W_L^\dagger & H^0 \end{pmatrix}. \tag{2.2}$$

where H_λ represents the Hamiltonian of lead λ, the Hamiltonian H^0 corresponds to the device region, and W_λ is the coupling between the device and this lead ($\lambda = 1, ...L$). The leads (acting

as reservoirs) are semi-infinite and, therefore, the total Hamiltonian H_{tot} is infinite dimensional. Since we are interested in ballistic transport of electrons through the finite device region, it is convenient to describe the device including its coupling to the leads in the following manner [6]. It is represented by a finite effective (non-hermitian) Hamiltonian $H = H^0 + \Sigma$ of the *open device*, where the influence of the external leads is included through a finite-dimensional operator Σ that can be derived from eq. (2.2) as will be explained below. The hermitian Hamiltonian H^0 represents, by contrast, a closed system, i.e. the device with no coupling to the leads or *decoupled device*. We assume throughout this work that the Hamiltonian H does not contain scattering terms. In this ballistic case, all observables of interest can, in principle, be obtained from the retarded Green's function G^R of the open device, which is defined by

$$G^R = [E - H]^{-1} = \left[E - H^0 - \Sigma\right]^{-1}, \tag{2.3}$$

where $\Sigma = \Sigma_1 + \Sigma_2 + \ldots + \Sigma_L$, and Σ_λ represents the complex self-energy due to the coupling W_λ between lead λ and the device. We can express eq. (2.3) in terms of the Green's function G^0 of the decoupled device in the following manner,

$$\begin{aligned} G^R &= [1 - G^0\Sigma]^{-1}G^0, \\ G^0 &= \left[E - H^0 + i\eta\right]^{-1}, \quad \eta \to 0+. \end{aligned} \tag{2.4}$$

The Green's function G^0 can also be expressed in terms of the eigenstates $|\alpha\rangle$ of the decoupled device Hamiltonian H^0 as follows,

$$\begin{aligned} G^0 = \sum_\alpha \frac{|\alpha\rangle \langle\alpha|}{E - \varepsilon_\alpha + i\eta}, \quad \eta \to 0+, \\ H^0 |\alpha\rangle = \varepsilon_\alpha |\alpha\rangle . \end{aligned} \tag{2.5}$$

2.3 Basic idea of CBR

The method rests on a Green's function approach that rigorously separates the open system problem into the single solution of a suitably defined closed system and the repeated solution of a small linear system of equations of size determined by the contact regions that couple the closed system to the leads. We have termed this method the contact block reduction method (CBR). We show in this work that the calculation of the charge density of the open system throughout the device can be performed with an effort comparable to a single calculation of a small percentage of the eigenstates of a closed system.

The evaluation of the retarded Green's function in eq. (2.3) requires the inversion of a large matrix that is proportional to the number of grid points or cells of the device [10, 12, 29], which can be quite demanding even in two spatial dimensions. The essence of the contact block reduction method consists in the decomposition of G^R into blocks such that the transmission function of the open device can be calculated by inverting only small matrices [30].

The first step in this procedure is to choose a real space discretization of the device region and subdivide the total grid space of N_T points into N_L boundary grid points that overlap with the leads – *contacts* – and the interior region with $N_D = N_T - N_L$ points. Typically the number N_T exceeds N_L by several orders of magnitude. The real space Hamiltonian matrix that corresponds to this discretization is assumed to couple only sites within some finite range with one another, typically only nearest neighbors. The contact grid points associated with

lead λ form the set C_λ of all points where the coupling matrix W_λ in eq. (2.2) is nonzero. The number of grid points in this set that links the device with the lead λ is denoted by N_λ. The total number N_L of boundary grid points that overlap with leads is given by

$$N_L = \sum_{\lambda=1}^{L} N_\lambda \ . \tag{2.6}$$

It is convenient to number the grid points, denoted by letters $i, j, k, l, ...$, in such a way that the total grid space is given by the following ordered set,

$$\Omega = \{C_1 \cup C_2 \cup ...C_L\} \cup \{D = \text{interior part of device}\}. \tag{2.7}$$

The self-energy in eq. (2.3) reflects the coupling of the device to the leads and is given by the expression [6]

$$\mathbf{\Sigma}_\lambda (i,j) = \begin{cases} \left[\mathbf{W}_\lambda \mathbf{G}_\lambda^R \mathbf{W}_\lambda^\dagger\right]_{ij} \ , & i,j \in C_\lambda \\ \mathbf{0} \ , & i,j \notin C_\lambda \end{cases} \tag{2.8}$$

where $\mathbf{\Sigma}_\lambda$ and $\mathbf{G}_\lambda^R$ are the self-energy matrix and the retarded Green's function matrix of lead λ, respectively. Note that we indicate all real space representations of operators, i.e. matrices, by bold letters. In the basis of the ordered set, eq. (2.7), the total self-energy matrix $\mathbf{\Sigma}$ is given by a block-diagonal matrix of the form

$$\mathbf{\Sigma} = \begin{pmatrix} \mathbf{\Sigma}_C & \mathbf{0} \\ \mathbf{0} & \mathbf{0}_D \end{pmatrix}, \tag{2.9}$$

$$\mathbf{\Sigma}_C = \mathbf{\Sigma}_1 \oplus \mathbf{\Sigma}_2 \oplus \ldots \mathbf{\Sigma}_L. \tag{2.10}$$

The submatrix $\mathbf{\Sigma}_C$ is of dimension N_L and $\mathbf{0}_D$ is a square null submatrix of dimension N_D associated with the interior part of the device. In addition, we define the Hermitian matrix

$$\mathbf{\Gamma} = i\left(\mathbf{\Sigma} - \mathbf{\Sigma}^\dagger\right), \tag{2.11}$$

which has the same structure as the self-energy matrix. Analogously, we define $\mathbf{G}_C^R$ to be the submatrix of the open device's retarded Green's function $\mathbf{G}^R$ within the contact regions. Now, let us write the Dyson equation (eq. (2.4)) for $\mathbf{G}^R$ in matrix form,

$$\mathbf{G}^R = [\mathbf{1} - \mathbf{G}^0\mathbf{\Sigma}]^{-1}\mathbf{G}^0 = \mathbf{A}^{-1}\mathbf{G}^0, \tag{2.12}$$

$$\mathbf{A} = \mathbf{1} - \mathbf{G}^0\mathbf{\Sigma} \ . \tag{2.13}$$

We now subdivide the matrix $\mathbf{G}^0$ into blocks corresponding to the contact portion $\mathbf{G}_C^0$ and the interior device portion $\mathbf{G}_D^0$, in analogy to the self energy matrix defined in equation (2.9). Then,

$$\mathbf{G}^0 = \begin{bmatrix} \mathbf{G}_C^0 & \mathbf{G}_{CD}^0 \\ \mathbf{G}_{DC}^0 & \mathbf{G}_D^0 \end{bmatrix} . \tag{2.14}$$

Note that $\mathbf{G}_C^0$ is a small matrix of size N_L, in contrast to $\mathbf{G}_D^0$ which has size $N_D \gg N_L$. Similarly, we subdivide the matrix $\mathbf{A}$ into blocks which leads to

$$\mathbf{A} = \begin{bmatrix} \mathbf{1}_C - \mathbf{G}_C^0\mathbf{\Sigma}_C & \mathbf{0} \\ -\mathbf{G}_{DC}^0\mathbf{\Sigma}_C & \mathbf{1}_D \end{bmatrix} . \tag{2.15}$$

Here, the submatrix $\mathbf{A}_C = \mathbf{1} - \mathbf{G}_C^0 \mathbf{\Sigma}_C$ is a square matrix of dimension N_L, whereas the lower diagonal block of $\mathbf{A}$ is a square unit matrix of dimension N_D. The inverse of the full matrix $\mathbf{A}$ possesses exactly the same structure as $\mathbf{A}$ itself, namely

$$\mathbf{A}^{-1} = \left[\begin{array}{cc} \mathbf{A}_C^{-1} & \mathbf{0} \\ \mathbf{G}_{DC}^0 \mathbf{\Sigma}_C \mathbf{A}_C^{-1} & \mathbf{1}_D \end{array} \right]. \tag{2.16}$$

Using this result eq. (2.16) and eq. (2.12), the full retarded Green's function of the open device can be finally written as

$$\mathbf{G}^R = \left[\begin{array}{cc} \mathbf{A}_C^{-1} \mathbf{G}_C^0 & \mathbf{A}_C^{-1} \mathbf{G}_{CD}^0 \\ \mathbf{G}_{DC}^0 \mathbf{\Sigma}_C \mathbf{A}_C^{-1} \mathbf{G}_C^0 + \mathbf{G}_{DC}^0 & \mathbf{G}_{DC}^0 \mathbf{\Sigma}_C \mathbf{A}_C^{-1} \mathbf{G}_{CD}^0 + \mathbf{G}_D^0 \end{array} \right]. \tag{2.17}$$

This important relation shows that *any* block of $\mathbf{G}^R$ can be calculated by inverting only the small submatrix $\mathbf{A}_C$. Based on this result, it is now easy to show [30] that the transmission function can be entirely expressed in terms of the small submatrix $\mathbf{G}_C^R = \mathbf{A}_C^{-1} \mathbf{G}_C^0$,

$$T_{\lambda\lambda'}(E) = \mathrm{Tr}\left[\mathbf{\Gamma}_C^{\lambda} \mathbf{G}_C^R \mathbf{\Gamma}_C^{\lambda'} \mathbf{G}_C^{R\dagger} \right], \tag{2.18}$$

$$\mathbf{\Gamma}_C^{\lambda} = i \left(\mathbf{\Sigma}_C^{\lambda} - \mathbf{\Sigma}_C^{\lambda\dagger} \right). \tag{2.19}$$

The contact elements of the decoupled Green's function $\mathbf{G}_C^0$ are to be determined through the spectral representation given in equation (2.5).

2.4 Contact block reduction method for the density matrix

In this section we show that the CBR method can be extended to observables that are functions of the carrier density in a way that maintains the advantages and the computational efficiency of this scheme. This generalization is a prerequisite for charge self-consistent calculations. The results presented in this section apply equally well to single-band and multi-band $\mathbf{k}\cdot\mathbf{p}$- or tight-binding Hamiltonians. The only assumption we make is that the basis is localized in real space so that nearest-neighbor interactions provide an adequate representation of the Hamiltonian matrix.

In order to derive a general expression for the density matrix, we first consider the spectral function $\mathbf{S}$ of the open device that is defined by [6]

$$\mathbf{S} \equiv \mathbf{G}^R \mathbf{\Gamma} \mathbf{G}^{R\dagger}. \tag{2.20}$$

Since $\mathbf{\Gamma}$ in eq. (2.11) is nonzero only within the contact region, $\mathbf{S}$ can be written as

$$\mathbf{S} = \begin{pmatrix} \mathbf{G}_C^R & \mathbf{G}_{CD}^R \\ \mathbf{G}_{DC}^R & \mathbf{G}_D^R \end{pmatrix} \begin{pmatrix} \mathbf{\Gamma}_C & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{G}_C^{R\dagger} & \mathbf{G}_{DC}^{R\dagger} \\ \mathbf{G}_{CD}^{R\dagger} & \mathbf{G}_D^{R\dagger} \end{pmatrix} \tag{2.21}$$

$$= \begin{pmatrix} \mathbf{G}_C^R \mathbf{\Gamma}_C \mathbf{G}_C^{R\dagger} & \mathbf{G}_C^R \mathbf{\Gamma}_C \mathbf{G}_{DC}^{R\dagger} \\ \mathbf{G}_{DC}^R \mathbf{\Gamma}_C \mathbf{G}_C^{R\dagger} & \mathbf{G}_{DC}^R \mathbf{\Gamma}_C \mathbf{G}_{DC}^{R\dagger} \end{pmatrix}. \tag{2.22}$$

This matrix consists of 4 blocks that may be denoted by $\mathbf{S}_{XX'}$ with elements $X, X' \in \{C, D\}$. One can easily see that the left column of the matrix in eq. (2.17) can be expressed in the form

[31]

$$\begin{aligned} \mathbf{G}^R_{XC} &= \mathbf{G}^0_{XC}\mathbf{B}_C^{-1}, \\ \mathbf{B}_C &\equiv \mathbf{1}_C - \mathbf{\Sigma}_C\mathbf{G}^0_C \ , \end{aligned} \tag{2.23}$$

where again $X \in \{C, D\}$. Inserting this into eq. (2.21) leads to the following expression for the spectral function

$$\mathbf{S}_{XX'} = \mathbf{G}^0_{XC}\mathbf{B}_C^{-1}\mathbf{\Gamma}_C\left(\mathbf{B}_C^{-1}\right)^\dagger\left(\mathbf{G}^0_{X'C}\right)^\dagger . \tag{2.24}$$

Using spectral representations for the decoupled Green's functions $\mathbf{G}^0_{XC}$, eq. (2.5), we can rewrite the matrix elements of $\mathbf{S}$ in terms of the discrete real space basis as follows

$$\begin{aligned} \mathbf{S}_{ij} &= \sum_{k,l\in C}\left[\sum_\alpha \frac{\langle i|\alpha\rangle\langle\alpha|k\rangle}{(E-\varepsilon_a+i\eta)}\right]\left[\mathbf{B}_C^{-1}\mathbf{\Gamma}_C\left(\mathbf{B}_C^{-1}\right)^\dagger\right]_{kl} \\ &\quad \times\left[\sum_\beta \frac{\langle l|\beta\rangle\langle\beta|j\rangle}{(E-\varepsilon_a-i\eta)}\right]\Bigg]_{\eta\to 0+} \\ &= 2\pi\sum_{\alpha,\beta}\langle i|\alpha\rangle\langle\beta|j\rangle\,\Xi_{\alpha\beta}(E), \end{aligned} \tag{2.25}$$

$$\Xi_{\alpha\beta}(E) = \frac{1}{2\pi}\left.\frac{\mathrm{Tr}\left\{\left[|\beta\rangle\langle\alpha|\right]_C\mathbf{B}_C^{-1}\mathbf{\Gamma}_C\mathbf{B}_C^{-1\dagger}\right\}}{(E-\varepsilon_a+i\eta)(E-\varepsilon_\beta-i\eta)}\right|_{\eta\to 0+} . \tag{2.26}$$

Equation (2.25) for the spectral function is our starting point for the derivations of the characteristics of the open system, such as the local density of states, the density matrix, and the charge density. Note that this expression only contains the inverse of the small matrix $\mathbf{B}_C$ of size N_L that is associated with the contact-device coupling. The term $\Xi_{\alpha\beta}(E)$ plays the role of a generalized density of states of the open system and is expressed in terms of the eigenstates $|\alpha\rangle, |\beta\rangle$ of the decoupled device, i.e. the closed system. In the limiting case of $\mathbf{\Gamma}_C \to 0$ (closed system), one gets $\Xi_{\alpha\beta}(E) = \delta_{\alpha\beta}\delta(E-E_\alpha)$ as shown in Appendix A. The local density of states $\rho_i(E) = (1/2\pi)S_{ii}(E)$ at grid point i follows now from

$$\rho_i(E) = \sum_{\alpha,\beta}\langle i|\alpha\rangle\langle\beta|i\rangle\,\Xi_{\alpha\beta}(E). \tag{2.27}$$

By integrating this expression over real space, the density of states is obtained as

$$N(E) = \mathrm{Tr}\ \Xi(E). \tag{2.28}$$

Finally, we can express the density in terms of $\Xi_{\alpha\beta}(E)$. The device is connected to L leads that represent particle reservoirs in local equilibrium and are characterized by lead distribution functions $f_\lambda(E)$, $\lambda = 1\ldots L$. It has been shown before that the particle density of an open system can be expressed for such a situation in the form [32, 29]

$$n_i = \frac{1}{2\pi}\int dE\sum_{\lambda=1}^{L} f_\lambda(E)\left\langle i|\mathbf{G}^R\mathbf{\Gamma}^\lambda\mathbf{G}^{R\dagger}|i\right\rangle, \tag{2.29}$$

where $\mathbf{\Gamma}^\lambda$ is the matrix defined in eq. (2.11) with self energies associated with lead λ. Note that a direct calculation of this expression would involve the solution of a very large linear system

of equations. Using the contact block reduction method and the results given by eq. (2.25), by contrast, yields a computationally efficient expression for the particle density n_i,

$$n_i = \sum_{\alpha,\beta} \langle i|\alpha\rangle \langle\beta|i\rangle \boldsymbol{\zeta}_{\alpha\beta}, \tag{2.30}$$

$$\boldsymbol{\zeta}_{\alpha\beta} = \sum_{\lambda=1}^{L} \int \boldsymbol{\Xi}^{\lambda}_{\alpha\beta}(E) f_\lambda(E)\, dE, \tag{2.31}$$

where $\boldsymbol{\zeta}_{\alpha\beta}$ is the density matrix, and the matrix $\boldsymbol{\Xi}^{\lambda}_{\alpha\beta}$ is defined analogously to eq. (2.26), with $\boldsymbol{\Gamma}_C$ being replaced by $\boldsymbol{\Gamma}^{\lambda}_C$ as defined in equation (2.19).

2.5 Numerical details and computational costs

The CBR method expresses the problem of ballistic scattering through an open nanostructure in terms of those parts of the retarded Green's function that connect the contacts to the interior device. Here we assess the numerical cost of calculating the density or local density of states in this approach. All of the matrices involved in eq. (2.26) are of small size $N_L \times N_L$. Initially, we solve the Hermitian eigenvalue problem for the decoupled device which requires $O\left(N_T N^2_{eigen}\right)$ operations for N_{eigen} eigenstates, $N_{eigen} \leq N_T$. The cost for the calculation of the transmission function [30] given by eq. (2.18) is $N_E \times \left[O\left(N_L^3\right) + N_L^2 N_{eigen}\right]$, where N_E is the number of energy grid points. The calculation of the local density of states in eq. (2.27) or the particle density n_i requires $N_{CBR} = N_E \times N_L^2 N^2_{eigen} + N_T N^2_{eigen}$ operations, since the matrix elements of $\boldsymbol{\Xi}_{\alpha\beta}(E)$ need to be calculated only once for all grid points i. As we will show below, it suffices for a mesoscopic device to calculate only a few percent of all N_T eigenvalues so that $N_{CBR} \ll N_E N_T$. This leads to an additional dramatic reduction in computer time.

2.6 Multiband lead mode dispersion and self-energy

In this section, we set up a concrete model for the self-energy eq. (2.8) that couples the device to the leads, taking into account a general multi-band Hamiltonian and a two- or three-dimensional device. Several ballistic lead models have been proposed. A recently published paper [27] treats the lead of a three-dimensional device as infinitely extended bulk in the two directions parallel to the contact surface so that the lead eigenstates are two-dimensional plane waves. In this work, we adopt the more common approach [6] to treat the lead as a semi-infinite straight quantum wire of finite circumference. For the following discussion, we pick out an individual lead of that type and calculate its self-energy. It is convenient to express it in terms of the eigenmodes $\chi(y,z)$ of an infinitely extended and homogeneous quantum wire. These modes propagate along the wire direction $\widehat{z}$ and are confined states in the perpendicular directions. All cross-sectional coordinates including band indices are lumped into the index y. With these assumptions, the wire modes are product states of one-dimensional plane waves along the lead wire and cross-sectional localized wave functions that are independent of the position z, i.e.

$$\chi(y,z) = \mathrm{e}^{ikz}\xi(y). \tag{2.32}$$

Let us discretize the quantum wire by introducing cross-sectional slices at positions $z = la$, where $l = 0, \pm1, \pm2...$ is an integer and a is the grid spacing. To keep the notation simple,

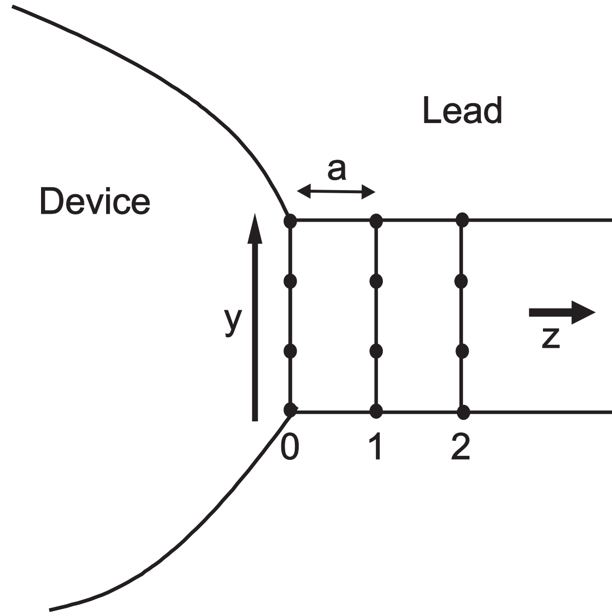


Figure 2.1: Schematic picture of local coordinate system of an individual semi-infinite lead attached to the device. The dots indicate the 2- or 3-dimensional discrete grid points. Along the lead axis z, the lattice spacing is a. The origin of the z axis lies at the contact boundary.

we only include nearest neighbor coupling in the Hamiltonian. We adopt a vector notation so that the wire eigenfunctions $\chi(y, z)$ are replaced by vectors χ_l of dimension M, where M cycles through all degrees of freedom except the position l along the wire. These wave functions obey the Schrödinger equation

$$(\mathbf{H}_{l,l} - E\mathbf{1})\,\chi_l + \mathbf{W}^\dagger\chi_{l-1} + \mathbf{W}\chi_{l+1} = 0, \tag{2.33}$$

where the matrix $\mathbf{H}_{l,l}$ is the cross-sectional part of the wire Hamiltonian, and $\mathbf{W}$ denotes the coupling between adjacent slices. To simplify notation and without loss of generality, we assume here the coupling constant between the lead slices to be the same as between the contact slice of the device and the lead. Note that $\mathbf{H}_{l,l}$ is the same for all slices l and is set up in such a way that the wave functions vanish at the circumference of the quantum wire. By employing the Bloch condition $\chi_l = \exp(ik_\mu a)\chi_{l-1}$, eq. (2.33) can be rewritten in the form of a non-Hermitian linear eigenvalue problem for the complex wire band structure $k_\mu\,(E)$ and the wire eigenfunctions χ_l^μ,

$$\begin{aligned} &\begin{pmatrix} (\mathbf{H}_{l,l} - E\mathbf{1}) & \mathbf{W} \\ \mathbf{1} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \chi_l \\ \chi_{l+1} \end{pmatrix}_\mu \\ = \;& e^{-ik_\mu a} \begin{pmatrix} -\mathbf{W}^\dagger & \mathbf{0} \\ \mathbf{0} & \mathbf{1} \end{pmatrix} \begin{pmatrix} \chi_l \\ \chi_{l+1} \end{pmatrix}_\mu . \end{aligned} \tag{2.34}$$

The solution of eq. (2.34) yields the wire mode dispersion $k_\mu\,(E)$ of the $2M$ linearly independent modes χ_l^μ which are propagating or decaying to the right and to the left, respectively. Now we consider the semi-infinite lead that extends from the contact at $l = 0$ to $l \to +\infty$ (see figure 2.1).

The self-energy couples the device with this semi-infinite lead and can be expressed in terms of those M modes χ_l^μ that propagate or decay into the leads ($l \to +\infty$). Within these M

modes, the evanescent ones have $\mathrm{Im}(k_\mu) < 0$ whereas the propagating ones are characterized by $\mathrm{Im}(k_\mu) = 0$ and group velocity $v_g^\mu > 0$. The latter can be calculated from [16]:

$$v_g^\mu (E) = -\frac{2a}{\hbar |\chi_l^\mu|^2} \mathrm{Im} \left((\chi_l^\mu)^\dagger \mathbf{W}^\dagger \chi_l^\mu e^{ik_\mu a} \right), \tag{2.35}$$

which is independent of l. Finally, the self-energy can be written in the following form, [27]

$$\mathbf{\Sigma} = -\mathbf{W}\mathbf{K}^{-1}\mathbf{\Lambda}\mathbf{K}, \tag{2.36}$$

where $\mathbf{K}$ is the $M \times M$ matrix of the column vectors χ_0^μ at the contact site ($l = 0$), and $\mathbf{\Lambda}$ is a diagonal matrix,

$$\mathbf{\Lambda} = \mathrm{diag}[\exp(-ik_1(E)a), ..., \exp(-ik_M(E)a)]. \tag{2.37}$$

We note that the wire dispersion relation can be solved analytically if the wire Hamiltonian describes either a single energy band or decoupled bands. In that case, the Hamiltonian is separable and has the form

$$H_L (y, z) = H_\perp (y) + T_\parallel (z), \tag{2.38}$$

where $H_\perp$ is the cross-sectional lead Hamiltonian that obeys a 1-D or 2-D Schrödinger equation

$$H_\perp |\xi^\mu\rangle = \varepsilon^\mu |\xi^\mu\rangle, \tag{2.39}$$

and the term $T_\parallel$ represents the kinetic energy along the propagation direction. In addition, in this case, the coupling matrix $\mathbf{W}$ in eq. (2.36) is proportional to the unit matrix, $\mathbf{W} = W\mathbf{1}$. This gives the following expression for the self-energy Σ at a contact [6]:

$$\mathbf{\Sigma}(i, j) = -W \sum_\mu^M \langle i|\xi^\mu\rangle \exp[ik^\mu a] \langle \xi^\mu|j\rangle \quad (i, j \in C_\lambda), \tag{2.40}$$

where the wire dispersion $k^\mu(E)$ is given analytically by the relation

$$E = \varepsilon^\mu + 2W (1 - \cos [k^\mu a]). \tag{2.41}$$

Thus, the cross-sectional lead modes have to be calculated only once in the single band case.

2.7 Boundary condition for decoupled device Hamiltonian

The splitting of the Hamiltonian H into a decoupled device Hamiltonian H^0 and the self energy Σ is not unique. This degree of freedom can be exploited such that the low-lying eigenstates of H^0 mimic the scattering states of the open system which greatly reduces the effort to calculate the retarded Green's function of the open device [30]. If we simply truncate the infinite system's Hamiltonian right at the contacts to define H^0, we obtain the self-energy as given by equation (2.36). This corresponds to Dirichlet (i.e. infinite barrier) boundary conditions for H^0. For single-band Hamiltonians, we have found previously that it is significantly more efficient to include the energy-independent part of the self energy into H^0 and to redefine the self-energy correspondingly. This procedure is equivalent to applying von Neumann boundary conditions to H^0 at the contacts. In this section, we generalize this procedure to multiband Hamiltonians.

By using a formal series expansion of the exponents, we decompose the diagonal matrix $\mathbf{\Lambda}$ for an individual lead into the unit matrix plus an energy dependent term,

$$\mathbf{\Lambda}(E) = \mathbf{1} + \mathbf{\Lambda}'(E), \tag{2.42}$$
$$\mathbf{\Lambda}' = \mathrm{diag}[\exp(-ik_1(E)a) - 1, ..., \exp(-ik_M(E)a) - 1].$$

We substitute this back into the self-energy to get

$$\mathbf{\Sigma} = -\mathbf{W} - \mathbf{W}\mathbf{K}^{-1}\mathbf{\Lambda}'\mathbf{K} \quad . \tag{2.43}$$

Since $\mathbf{W}$ is not Hermitian, we symmetrize it as

$$\mathbf{W}^H = \frac{1}{2}\left[\mathbf{W} + \mathbf{W}^\dagger\right] \ , \tag{2.44}$$

and redefine the decoupled device Hamiltonian and the self-energy, respectively, as

$$\mathbf{H}^N = \mathbf{H}^0 - \mathbf{W}^H \ , \tag{2.45}$$
$$\mathbf{\Sigma}^N = \mathbf{\Sigma} + \mathbf{W}^H. \tag{2.46}$$

The multi-band Hamiltonian $\mathbf{H}^N$ corresponds to a closed system with generalized von Neumann boundary conditions at the contact boundaries. Thus, the eigenfunctions of the Hamiltonian $\mathbf{H}^N$ are approximate solutions of the open-boundary problem in the low energy limit [30]. As a consequence, by solving the Dyson equation (eq. (2.4)) with $\mathbf{H}^N$ instead of $\mathbf{H}^0$, it suffices to include an incomplete set of eigenfunctions in the spectral representation of $\mathbf{H}^N$ for some limited energy range. From a numerical point of view, this is the most crucial timesaving step in the CBR method that will be illustrated in section (2.9).

2.8 Mode space reduction in single-band case

In this section, we show that the numerical effort of calculating the transmission function and the density matrix of the open system can be further reduced in the single band case by transforming the Green's functions and the self-energy into a basis of propagating mode states. Since the cross-sectional mode eigenfunctions ξ^μ of a given lead in eq. (2.40) are orthogonal and complete, we can transform the real space matrices $\mathbf{G}^0_{C,ij}$, $\mathbf{\Sigma}_{C,ij}$ and $\mathbf{\Gamma}_{C,ij}$ into mode space, where the indices i, j are elements of the contact set C . This leads to diagonal matrices $\mathbf{\Sigma}_{\mu\mu}$ (and $\mathbf{\Gamma}_{\mu\mu}$) and full matrices $\mathbf{G}^0_{\mu\nu}$ with (Greek) mode indices μ, ν,

$$\mathbf{\Sigma}_{\mu\mu} = \sum_{ij}^{M} \langle \xi^\mu | i \rangle \, \mathbf{\Sigma}_{ij} \, \langle j | \xi^\mu \rangle = -W \exp[ik^\mu a], \tag{2.47}$$

$$\mathbf{G}^0_{\mu\nu} = \sum_{ij}^{M} \langle \xi^\mu | i \rangle \, \mathbf{G}^0_{ij} \, \langle j | \xi^\nu \rangle \tag{2.48}$$

Correspondingly, one has $\mathbf{\Gamma}_{\mu\mu} = 2\mathrm{Im}\mathbf{\Sigma}_{\mu\mu}$. Therefore, the purely real elements $\mathbf{\Sigma}_{\mu\mu}$ correspond to $\mathbf{\Gamma}_{\mu\mu} = 0$ and do not contribute to the traces in the transmission function given by eq. (2.18) and the density matrix determined from eq. (2.26), because they correspond to decaying modes. In contrast to the transmission, the Green's function of the open device $\mathbf{G}^R$ is determined, in

principle, by the full self-energy $\mathbf{\Sigma}_C$ including purely real elements [33]. Here we show, however, that the error introduced by neglecting the decaying modes in the calculation $\mathbf{G}^R$ is negligible. To this end, we write the self-energy matrix in the mode basis in the following form

$$\mathbf{\Sigma} = \begin{pmatrix} \mathbf{\Sigma}_P & 0 \\ 0 & \mathbf{\Sigma}_E \end{pmatrix}, \tag{2.49}$$

where the diagonal blocks $\mathbf{\Sigma}_P$ and $\mathbf{\Sigma}_E$ correspond to the propagating and evanescent modes, respectively. We note that the dimension of $\mathbf{\Sigma}_P$ is much smaller than that of $\mathbf{\Sigma}_E$, because only a small fraction of the modes are propagating modes at low energies which is relevant for ballistic situations. Then, the Dyson equation for the retarded Green's function corresponding to the propagating modes P can be written in block-matrix form as

$$\begin{aligned} \mathbf{G}^R_{PP} &= \mathbf{G}^0_{PP} + \left[\mathbf{G}^0\mathbf{\Sigma}\mathbf{G}^0\right]_{PP} + ... \\ &= \mathbf{G}^0_{PP} + \mathbf{G}^0_{PP}\mathbf{\Sigma}_P\mathbf{G}^0_{PP} + \mathbf{G}^0_{PE}\mathbf{\Sigma}_E\mathbf{G}^0_{EP} + ... \ . \end{aligned} \tag{2.50}$$

As we will show below, the off-diagonal blocks $\mathbf{G}^0_{EP}$ are much smaller than both diagonal blocks $\mathbf{G}^0_{PP}$ and $\mathbf{G}^0_{EE}$, which gives the approximate relation

$$\mathbf{G}^R_{PP} \approx \left[\mathbf{1} - \mathbf{G}^0_{PP}\ \mathbf{\Sigma}_P\right]^{-1} \mathbf{G}^0_{PP}. \tag{2.51}$$

This equation allows one to evaluate both the transmission function and the density matrix extremely efficiently since the computational effort for solving this equation is proportional to the third power of the matrix dimension. The mixed Green's function matrix elements in eq. (2.50),

$$\mathbf{G}^0_{EP} = \sum_\alpha \frac{\langle E|\alpha\rangle\ \langle\alpha|P\rangle}{E - \varepsilon_\alpha}, \tag{2.52}$$

can be neglected because of the fact that the numerators obey

$$|\ \langle E|\alpha\rangle\ \langle\alpha|P\rangle\ | \ll |\ \langle P|\alpha\rangle\ |^2, |\ \langle E|\alpha\rangle\ |^2\ , \tag{2.53}$$

for all values of α which label the eigenstates of the decoupled device. In order to show this, we note that the cross-sectional wave functions ψ^μ that correspond to propagating modes have eigenvalues $\varepsilon^\mu < E$, while the decaying lead modes have eigenvalues $\varepsilon_\mu > E$. Since, in the single-band case, the energy separation of the eigenenergies ε_μ increases roughly quadratically with increasing μ, the energy distance between the eigenvalues of evanescent and propagating modes is much larger than the energy distance between two propagating modes. Therefore, the projection in eq. (2.52) of the device eigenstate $|\alpha\rangle$ with eigenvalue $\varepsilon_\alpha < E$ onto the propagating mode $|P\rangle$ is much larger than onto the evanescent modes $|E\rangle$ and *vice versa* for the device eigenstates with $\varepsilon_\alpha > E$. This implies the validity of the inequality eq. (2.53) as one of the scalar products on the left hand side is small for any α.

This result grossly reduces the numerical effort in calculating the inverse of the matrix $\mathbf{B}_C$ as defined in equation (2.23). This is a crucial factor that enables one to calculate the transmission through realistic 3-D devices even with large contact areas as will be illustrated in section (2.10).

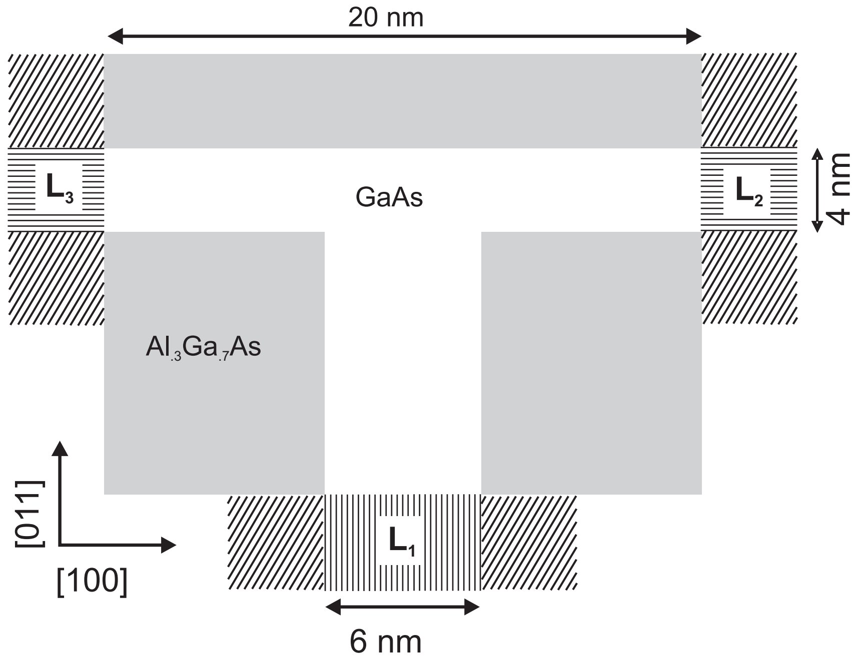


Figure 2.2: Sketch of the T-junction device used in the calculation. The leads are indicated by the patterned regions.

2.9 Hole transport in T-shaped junction

The efficiency of the CBR method for the fully quantum mechanical calculation of the transmission function, the density of states, and the carrier density, as developed in this work, rests on three key points: (i) the evaluation of all of theses observables requires only that part of the retarded Green's function $\mathbf{G}^R$ that connects the contacts to the interior device; (ii) $\mathbf{G}^R$ can be evaluated efficiently by calculating only a few percent of the eigenstates of a suitably defined Hermitian Hamiltonian $\mathbf{H}^N$, equation (2.45); (iii) for single-band Hamiltonians, the size of the contact part of the $\mathbf{G}^R$ matrix can be grossly reduced by transforming into the basis of all propagating lead modes which amount to only a few percent of all modes.

To illustrate the applicability of the method in the multi-band case, we consider a T-junction that is formed by two crossing GaAs quantum wells. Such a device can be realized by the cleaved edge overgrowth technique [34] and has been investigated experimentally extensively [35, 36, 37, 38, 39]. However, the complicated geometry has hampered realistic theoretical studies on transport properties so far [40, 41].

To be concrete, we assume the barrier material to be p-doped and analyze the ballistic quantum transport characteristics of holes in the GaAs wells. We assume the two perpendicular GaAs quantum wells to be embedded within $Al_{.3}Ga_{.7}As$ barrier material. One quantum well consists of a 6 nm thick GaAs layer that is grown along the [100] direction. After cleavage along the [011] surface, a second quantum well of 4 nm thickness is grown, followed by a cap layer of $Al_{.3}Ga_{.7}As$ (see figure 2.2).

Since this device possesses three contacts, it is not possible to reduce the problem size to one-dimensional slicing, which is a prerequisite for the recursive Green's function method [23]. We have discretized the two-dimensional structure shown in fig. 2.2, using a 20 nm by 20 nm mesh with constant grid spacing of 0.5 nm. The leads are indicated by the striped regions and

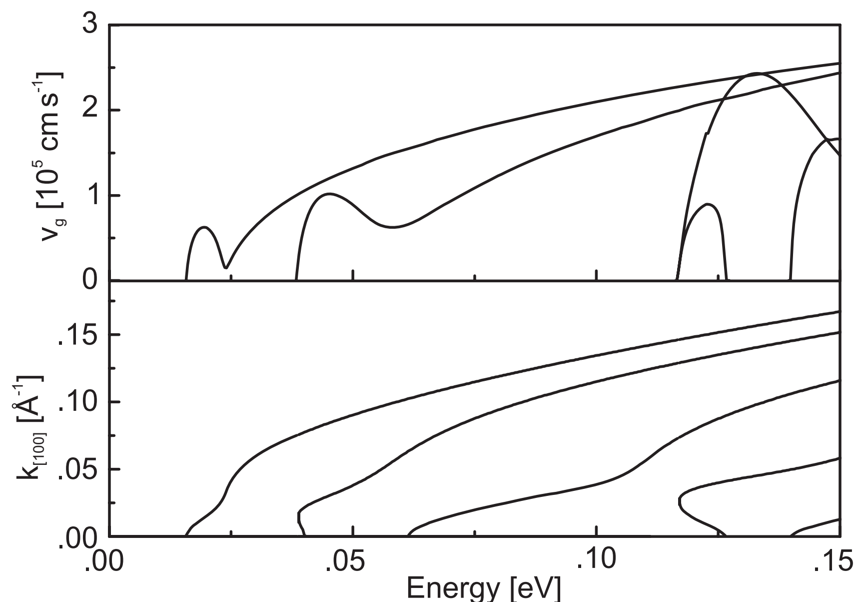


Figure 2.3: Lower box: subband structure of the lead L_1. Energy is in eV, the wave vector k lies along the [011] direction. Upper box: Corresponding group velocity in 10^5cm/s as a function of energy in each band. The perpendicular wavevector along $[01\bar{1}]$ has been set equal to zero.

labeled L_1 to L_3. Since the valence band offset (barrier height) is only 140 meV, the transverse wave-functions of the leads penetrate into the barrier regions (indicated by the diagonal stripes) in figure 2.2. Therefore, the diameter of the semi-infinite lead wires has to be considerably larger than the quantum well thickness such that the wire modes are completely decayed at the wire boundaries. We employ a four band $\mathbf{k}\cdot\mathbf{p}$-Hamiltonian (see Appendix B) to calculate the band structure for the holes. The Hamiltonian includes the $p_{3/2}$-spinor states and couples the light and heavy hole bands but neglects the $p_{1/2}$ states. With the exception of the valence band offset, the $\mathbf{k}\cdot\mathbf{p}$-parameters are taken to be the same throughout the structure.

The strong coupling between heavy and light hole bands induced by the lateral confinement results in a complicated dispersion function in the leads that has to be evaluated numerically at each energy step. We first analyze the properties of the leads — such as dispersion and group velocity — for the case of zero wave vector along the free $k_{[01\bar{1}]}$ direction and then perform the integration over k-space along the $[01\bar{1}]$ direction to obtain the conductance and the local carrier density.

2.9.1 Lead dispersion

For the first lead L_1 (see fig. 2.2), the quantization axis lies parallel to [100], whereas the propagation direction is [011]. Figure 2.3 shows the dispersion $k_{[011]}(E)$, and the corresponding group velocity $v_g(E)$ for this lead. The pronounced quantization strongly mixes the heavy and light holes. Some propagating modes only exist within a certain energy interval and then decay, leading to the maxima and semi-parabolas in the group velocity. Thus, the number of propagating modes is no longer monotonously rising with increasing energy as is the case for parabolic bands.

The second lead L_2 (see fig. 2.2) is quantized in the [011] direction, and the carriers propagate

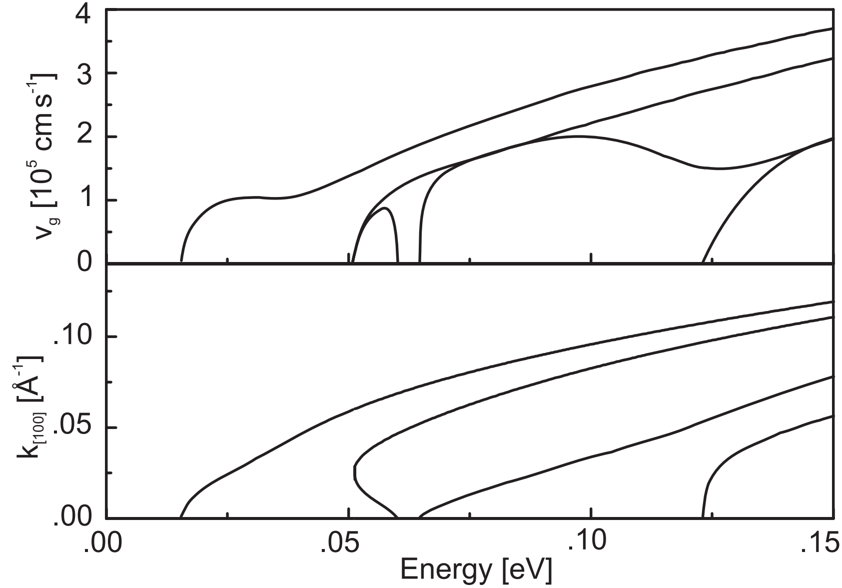


Figure 2.4: Lower box: subband structure of the lead L_2. Energy is in eV, the wave vector k lies along the [100] direction. Upper box: Corresponding group velocity in 10^5 cm/s as a function of energy in each band. The perpendicular wavevector along $[01\bar{1}]$ has been set equal to zero.

along the [100] direction. Figure 2.4 shows the dispersion $k_{[100]}(E)$, and the corresponding group velocity. As in lead L_1, the bands are strongly coupled, but exhibit a slightly higher group velocity due to the smaller effective mass in the [100] direction. The third lead L_3 has identical properties, as is obvious from figure 2.3.

It is interesting to note that the onset of the first propagating mode occurs at almost the same energy for the L_1 and L_2 leads in spite of their different width of 6 and 4 nm, respectively. This is due to the highly anisotropic mass of the heavy hole band that reaches its maximum along the [011] direction. This result would not be obtained in a single-band analysis with parabolic masses which would lead to incorrect transmission characteristics.

2.9.2 Transmission function and density of states

Using the CBR method, the transmission function between the three leads can be calculated very efficiently. The transmission function $T(E)$ shown in fig. 2.5 is a spiky function due to the reflections caused by the sharp edges in the T-junction potential. The density of states provides information regarding the total number of carriers in the device and is depicted in figure 2.6.

As we have pointed out in section 2.6, we compute the decoupled Green function by employing a transformation that corresponds to generalized von Neumann boundary conditions. We performed a series of calculations that included either all or only a fraction of the eigenstates of H^N in the spectral representation of G^0. Figure 2.5 compares the transmission functions T_{12} and T_{23} calculated exactly (solid lines) with those obtained by including only a subset of eigenstates, namely the lowest 25 % (dashed lines) or 10 % (dotted lines) of all states. This reduction corresponds to cutoff energies of 1.5 and 0.6 eV respectively. As one can deduce from figs. 2.5 and 2.6, the inclusion of 25 % of the eigenstates still guarantees basically exact results for low energies.

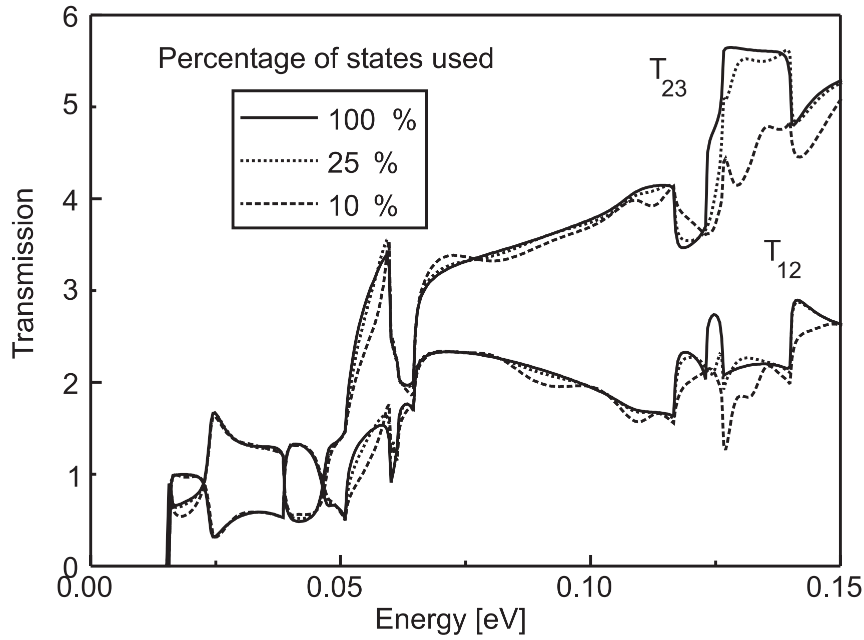


Figure 2.5: The transmission functions $T_{12}(E)$=$T_{13}(E)$ and $T_{23}(E)$ of the T-junction, plotted as a function of energy in eV. The perpendicular wavevector along $[01\bar{1}]$ has been set equal to zero. The exact calculations (solid curves) are compared to results with a reduced set of eigenstates, either using 25 % (dashed curves) or using 10 % (dotted curves) of all eigenstates.

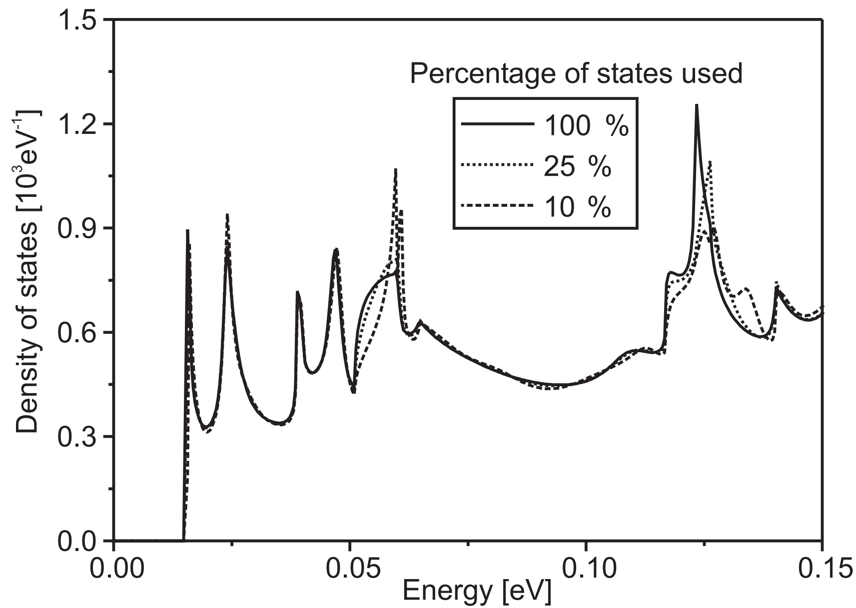


Figure 2.6: Density of states of the T-junction in units of 10^3eV^{-1} as a function of energy in eV. The perpendicular wavevector along $[01\bar{1}]$ has been set equal to zero. The exact calculations (solid curves) are compared to results with a reduced set of eigenstates, either using 25 % (dashed curves) or using 10 % (dotted curves) of all eigenstates.

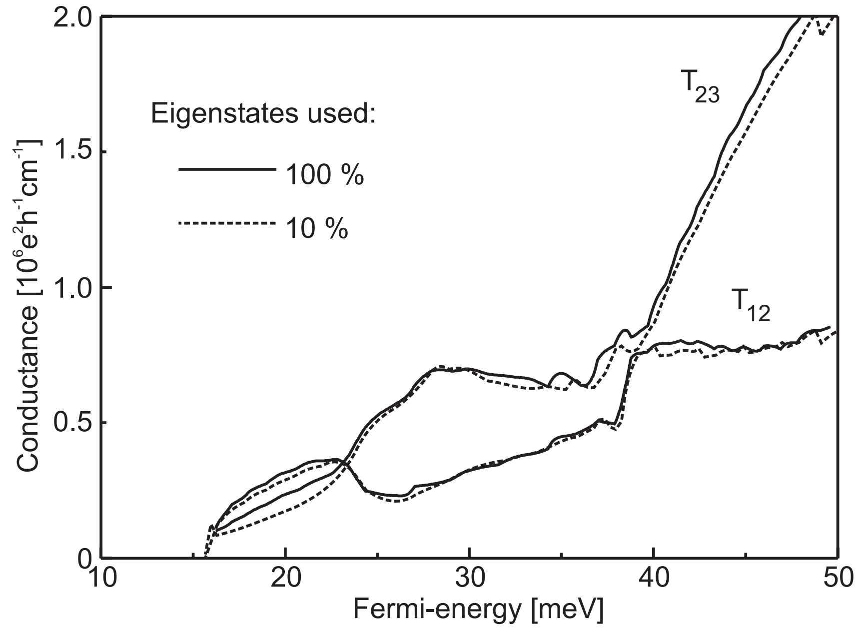


Figure 2.7: The total conductance is calculated via the integration of the transmission in k space. The exact result (solid curves) is compared to the transmission obtained using only the lowest 10 % of all eigenstates (dashed curves).

Thus, the reduction of the problem size that can be achieved by neglecting high energy eigenstates in G^0 is comparable for multi-band Hamiltonians and for single-band Hamiltonians [30], although the strong coupling between the bands requires a higher cutoff energy in the former case.

2.9.3 Conductance

So far, we only presented auxiliary quantities such as transmission and density of states. In this section, we present concrete predictions of the conductance for the T-junction shown in figure 2.2. So far, no experimental data on this structure seem to be available; however, hole transport in AlGaAs heterostructures with mean free paths of the order of μm has been reported [42, 43]. We have calculated the differential conductance at the Fermi-edge E_F between the different leads. To this end, we have integrated the transmission in k-space along the $[01\bar{1}]$ direction from $k = 0$ up to the Fermi wave vector. Numerically, this is achieved by computing the transmission function for different values of $k_{[01\bar{1}]}$ from 0 to 0.1 Å^{-1}, using a homogenous k-spacing of 2×10^{-3} Å^{-1}. Figure 2.7 displays the resulting conductance that exhibits a highly non-monotonic behavior. For Fermi energies below 40 meV, the calculations show that the preferred transport channel is T_{12} for Fermi energies below 24 meV. For higher energies, however, we see a switch towards the T_{23} channel. The sharp drop of the T_{12} conductance at the transition energy gives even rise to a negative differential resistance. For energies higher than 40 meV, the holes with higher energy and correspondingly larger momentum are less likely to be scattered at the junction than low energy holes which explains the larger conductance in the T_{23} channel.

The dashed lines in fig. 2.7 show that the resulting conductance when only 10 % of all eigenstates are taken into account. The comparison with the exact results (full lines) shows that the integrated conductance is even less sensitive to an energy cutoff than the transmission

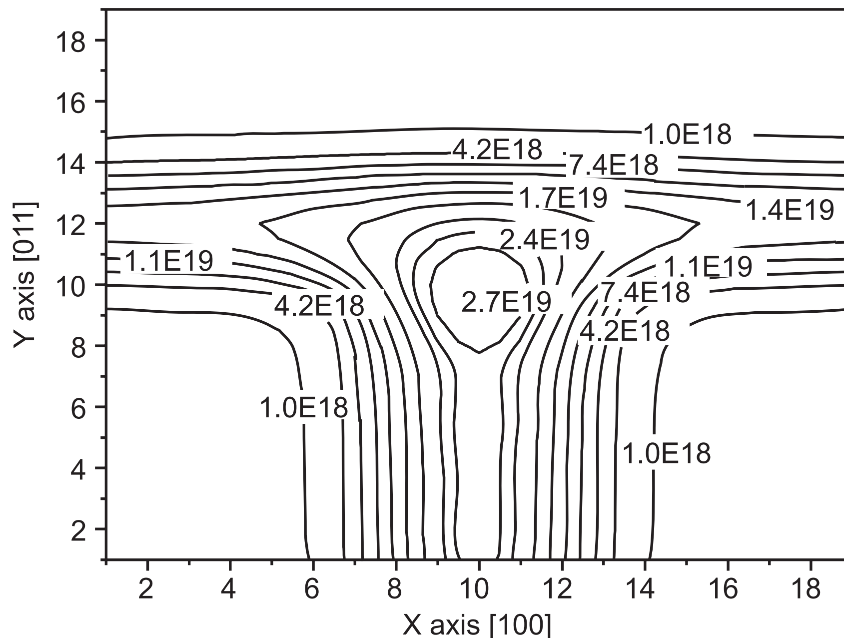


Figure 2.8: Contour plot of the hole density in units of cm^{-3} for the T-junction at zero temperature. The Fermi level is set to 0.05 eV.

function.

2.9.4 Carrier density

We have calculated the local carrier density for a Fermi level of 0.05 eV at zero temperature. Note that the potential is assumed to be constant within the entire well region within the device, since this calculation is not charge-self consistent. The integrated hole density of the device is plotted in figure 2.8. It shows a high density of holes in the lower quantum well that exhibits small oscillations along the [011] axis which is due to reflections at the upper barrier. The main feature of interest is the peak in the density right at the junction which reflects a bound state. A contour plot of this laterally localized state is shown in figure 2.9. It is bound only for small values of $k_{[01\bar{1}]}$ with a binding energy of 1 meV at $k = 0$.

In the calculation of this density, we have included only a fraction of the eigenstates in evaluating G^R. Since the contribution of the eigenstates to the density matrix scales with $1/(\Delta_\alpha \Delta_\beta)$, where $\Delta_{\alpha,\beta} = E - \varepsilon_{\alpha,\beta}$ according to eq. (2.26), it is possible to introduce a cutoff energy E_{cutoff} that obeys the condition $E^2_{\text{cutoff}} > |(E - \varepsilon_\alpha)(E - \varepsilon_\beta)|$. For given E_{cutoff}, only states α and β are taken into account in the calculation of the density that obey this condition. The magnitude of this cutoff energy has been set to 0.1 eV which is larger than the Fermi energy of 0.05 eV. This cutoff reduces the computational effort by almost a factor of 1000. Even when we reduce the cutoff energy to 0.05 eV, the resulting density differs by less than 2 % throughout the device from the one with the larger cutoff. This indicates excellent convergence of this procedure.

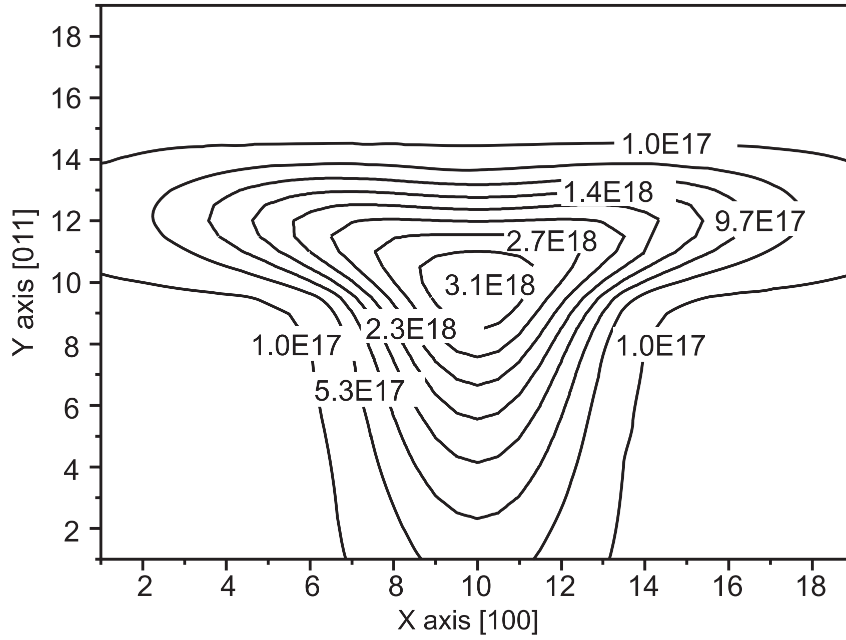


Figure 2.9: Contour plot of the lateral bound state for holes. The parameters are the same as in the previous figure.

2.9.5 k-space distribution of the density of states

Further insight into the physics of this 3-terminal device is obtained by studying the k-resolved density of states (k-DOS). For some device geometries, it has been shown that this property is accessible experimentally by resonant magneto-tunneling spectroscopy [44, 45]. In figure 2.10, the k-DOS is shown as a function of the energy and of the wave vector $k_{[01\bar{1}]}$ perpendicular to the device plane. The dark regions indicate high values of the k-DOS. Generally, this density of states possesses features that can be attributed to the individual quantum wells and others that arise from the interference between the quantum wells. The former states can be identified in fig. 2.10 by the white circles and squares, respectively, which represent the sub-band edges of the isolated 6 nm and 4 nm quantum wells in the absence of the T-junction. The position of the two bound states is marked by solid squares in this figure; they are spin degenerate at $k_{[01\bar{1}]} = 0$.

This degeneracy is lifted by the structure inversion asymmetry that splits the two states linearly proportional to k. For k-values below 0.01 Å $^{-1}$, the dispersion is approximately parabolic, and we can extract the spin-splitting to be $0.07 \pm 6 \times 10^{-4}$ eVÅ. For larger values of k, the bound states eventually merge into the continuum.

This example demonstrates that the CBR method is an efficient technique to calculate the carrier density in an arbitrary shaped open device structure accurately. The main practical limitation is given by the need to take into account all lead modes in the case of multi-band Hamiltonians, in contrast to the CBR method for the single-band case.

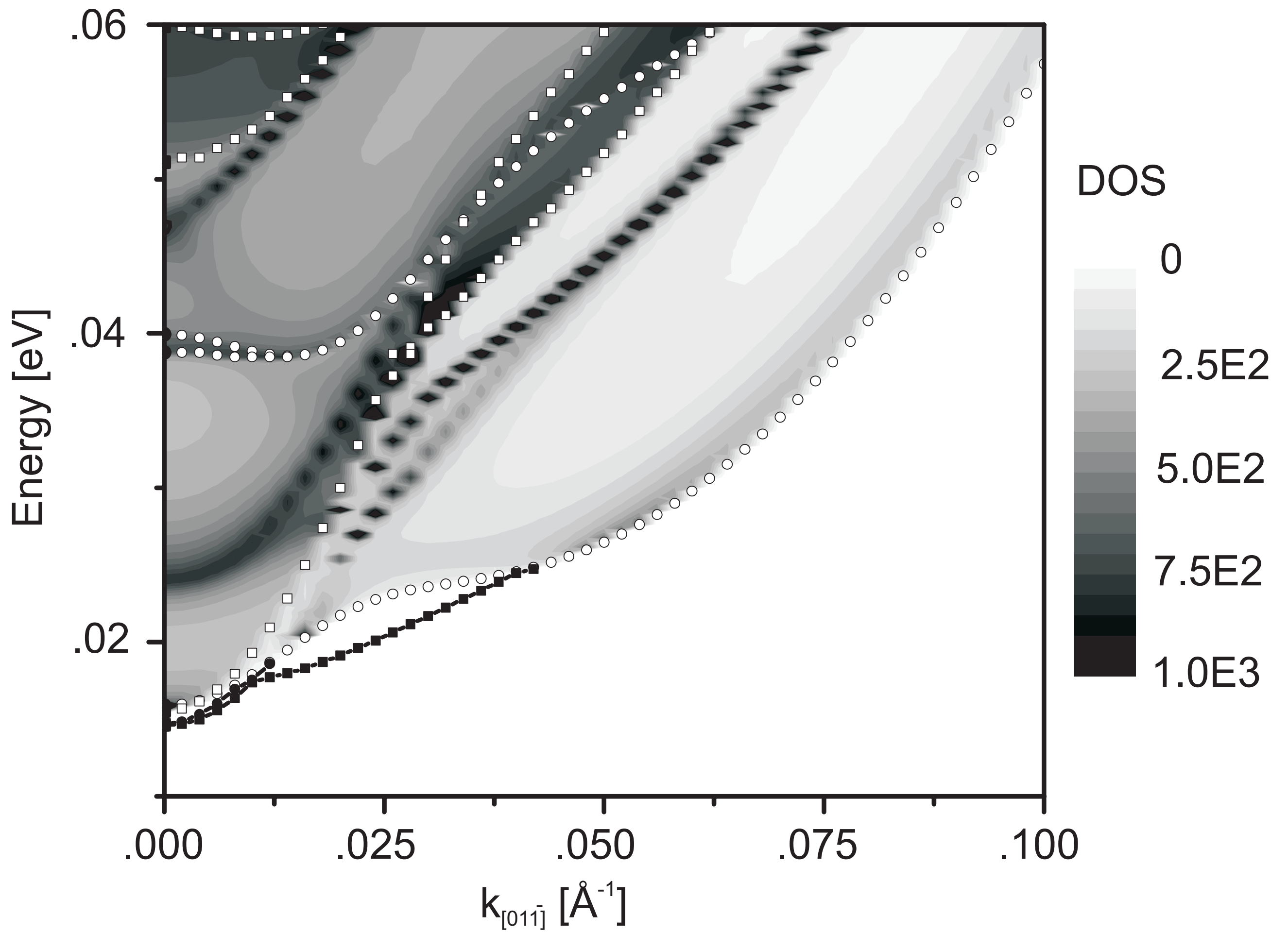


Figure 2.10: The wave-vector resolved density of states of the T-junction as a function of energy in eV and wavevector $k_{[01\bar{1}]}$ in units of Å^{-1}. Dark regions indicate a high density of states on a linear scale. The open circles and squares indicate the subband-edges of the isolated 6 nm and 4 nm quantum wells, respectively. The solid squares depict the location of the bound states.

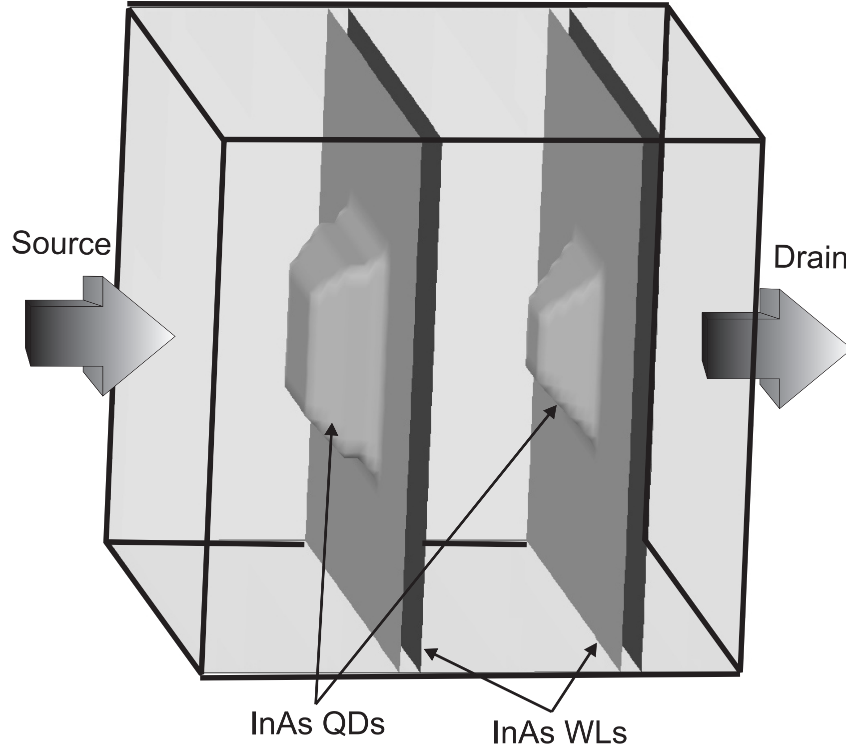


Figure 2.11: The two InAs quantum dots (QDs) on top of the InAs wetting layers (WLs) are embedded within InP barrier material. The two faces in growth direction act as source and drain contacts and consist of doped $In_{0.6}Ga_{0.4}As$.

2.10 Resonant tunneling through a 3D quantum dot molecule

To demonstrate the applicability of the CBR method to realistic 3D problems we consider a resonant tunneling diode that consists of two vertically stacked quantum dots. This structure is depicted in fig. 2.11 and is composed of a 33 nm wide InP barrier layer with two embedded InAs dots that are grown on 0.5 nm thick InAs wetting layers. On both sides of the InP layer, there are thick layers of doped $In_{0.6}Ga_{0.4}As$ that act as contacts. This structure has been investigated experimentally recently [46, 47]. Using the device-simulator next**nano**[3] [48], we have calculated the three-dimensional strain field and the corresponding piezoelectric potential. The electronic states are represented by an effective mass Hamiltonian with spatially varying masses and band offsets for the conduction band. The potential includes the strain induced deformation potential. In the present calculation, the quantum dots have been modeled by truncated pyramids of 2.5 nm height and 12 nm base width for the right dot in fig. 2.11 and of 6 nm height and 16 nm base width for the left dot, respectively.

We have chosen a grid spacing of 1 nm in the two lateral directions and 0.5 nm along the propagation direction, resulting in a total number of 128000 grid points. In spite of the large size of this device, we find the electric current to be well converged by including as few as 100 eigenstates in the computation of $\mathbf{G}^0$ in eq. (2.48) and 50 propagating lead modes out of a total of 1600 modes in the relevant bias range. The calculated resonant current is shown in fig. 2.12 and exhibits a very high peak-to-valley ratio in accord with the experiment. The two ground

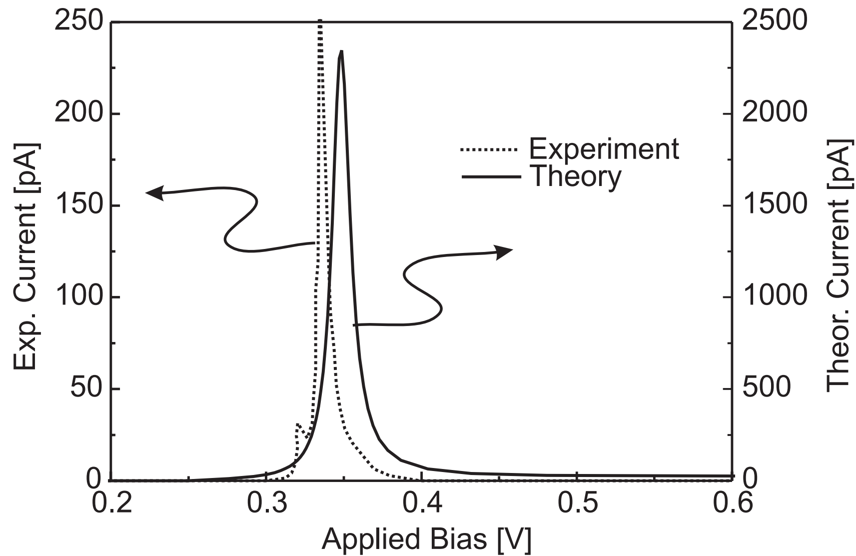


Figure 2.12: Current in units of pA for the quantum dot molecule based resonant tunneling structure as a function of the applied bias in V. The scale for the experimentally observed and the calculated current is shown on the left and right vertical axis, respectively.

states are in resonance with one another only for a small range of voltages. The position of the peak agrees very well with the experiment, whereas the maximum current and peak width exceed the measured data by one order of magnitude. Since the tunneling current depends exponentially on the barrier thickness, already small discrepancies between the assumed and the real dot shape may account for this difference, in addition to charging and scattering effects that have been ignored in these calculations.

We note that the results are well converged with respect to the lateral dimension of the structure. A more detailed description of the transport characteristics of this quantum dot molecule will be presented elsewhere.

2.11 Conclusion

In this section, we have shown that the electronic properties of a general 2- or 3-dimensional ballistic open system can be calculated efficiently in terms of a small number of eigenstates of a properly defined closed system and a few additional linear systems of small size. Any kind of open system can be arranged into an inner part of the device, contact regions, and the leads. We have shown that the Green's function formalism allows one to exploit the fact that the contact regions constitute a small fraction of the entire system very efficiently. In particular, we have found that the open device problem can be solved rigorously by breaking it down into 3 separate problems that can be tackled consecutively, namely a large Hermitian problem for the decoupled (closed) device, and two small non-Hermitian problems for the leads and the contacts. We have termed this method contact block reduction (CBR) method [30]. We have shown here that this method not only applies to the transmission function [30], but basically to any electronic observable such as the local charge density and the local density of states. As a result, this scheme avoids the repetitive solution of the entire device problem for each energy that has hampered calculations of ballistic transport for realistic devices so far.

We have shown that the CBR method is well suited to deal with multi-band Hamiltonians, either in a $\mathbf{k}\cdot\mathbf{p}$ representation or in a tight binding basis, and is applicable to devices of any shape and any number of leads in 2 or 3 spatial dimensions. The computational effort of the calculation of the transmission function and the carrier density is determined by the effort required to diagonalize the closed system's Hamiltonian. This effort can be grossly reduced since we found that it suffices for the open system observables to include only a few percent of this Hamiltonian's eigenstates. For the special case of single-band Hamiltonians, a further significant reduction of the problem size can be achieved by transforming into the basis of all propagating lead modes. We have shown that the neglect of the decaying lead modes in the Green's function introduces only negligible errors.

As one concrete application, we have predicted the hole conductance, density, and electronic bound state in a GaAs T-junction. In addition, we have studied the ballistic transport through a resonant tunneling diode consisting of two aligned InAs quantum dots and found a single sharp resonance for tunneling between the ground states of the quantum dots, in agreement with experiment.

Chapter 3

Self-consistent extension of CBR

3.1 Introduction

For the devices investigated in the previous part of this work, charge self-consistency was not of high importance, since the device was either very close to equilibrium (T-junction) or the states were unoccupied as in the case of the coupled quantum dots. For most other devices, especially those that operate at room temperature and under high bias, such as transistors, resistors or diodes, charge self-consistency cannot be neglected. Recent advances in semiconductor technology lead to a tremendous reduction of the channel lengths of metal oxide field effect transistors (MosFets). Working samples with gate lengths in the range of 10 nm have been demonstrated by several research groups [49, 50]. Design considerations to yield devices with desirable channel length have been theoretically explored in refs. [51, 52], and [53]. The smallest devices are based on the Silicon on insulator (SOI) technique that allows the fabrication of very thin Silicon layers on top of an oxide. Using this technology various kinds of new nano-MosFets have been developed, with the most prominent being the double gate MosFet (DGFET).

The combination of the quantum mechanical transport calculation needed for such nanostructures, together with charge self-consistency poses a computationally demanding problem for larger two-dimensional devices. Thus even within the purely ballistic approach only very recently a successful solution was presented by Laux *et al.* [19] but still the computational effort is quite high with several days per bias point. In this work we demonstrate that the CBR method can be applied to self-consistent calculations while keeping its efficiency, and thus reducing the computational effort by almost one order of magnitude compared to the work of Laux.

In the following we will describe the self-consistent algorithm using the example of the DGFET. Special focus is set on the treatment of bound and quasi bound states that may be present in these structures, since their contribution to the carrier density cannot be neglected. The resulting current-voltage characteristics of the DGFET are presented, analyzing the effect of the bound states and an optional inclusion of exchange-correlation effects within local density approximation (LDA).

3.2 Description of the model

To simplify the following discussion, we demonstrate the applicability of the CBR method to the self-consistent calculation of the transport properties of open nano devices on a concrete

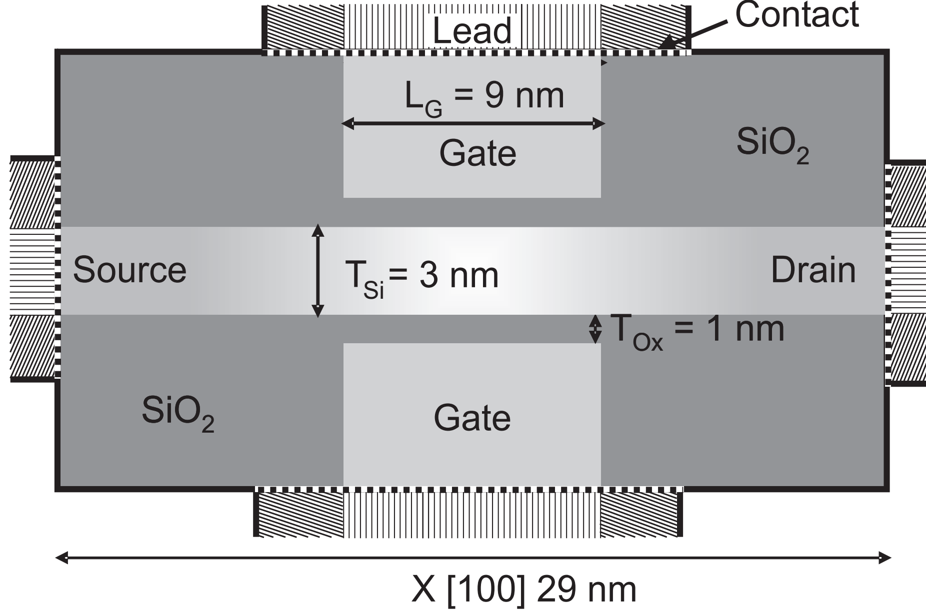


Figure 3.1: The calculated double gate MosFet structure with the attached leads indicated by the patterned regions. The device boundary is divided into the contact part (dashed line) and the outer boundary (solid line). The dimensions have been taken from [54].

example of the DGFET. The structural parameters of this device, such as an oxide thickness of 1 nm, a channel width of 3 nm and a gate length of 9 nm, are taken from the example of the Lundstroem group [54]. The simulation domain is displayed in fig. 3.1, showing the oxide (dark gray), the channel, as well as the gate regions, which are fully included into the domain to account for the gate leakage current. The position of the semi-infinite leads is indicated by the black rectangles. The gate, source and drain regions are highly doped with 2×10^{20} cm^{-3} (light gray), whereas the doping within the channel is gradually decreased to a background value of 10^{10} cm^{-3} (white region) as shown in figure 3.2. The overall dimensions of the structure are determined by the condition that all space charge regions have to be confined within the simulated device region. Since, in this work, the gates are treated in the same way as source and drain, the device region has to be considerably larger than for pure Poisson boundary gates. The benefit of this extra effort is the prediction of the charge accumulation that occurs at the interface between the gate and the oxide barrier, which has a strong influence on the resulting device characteristics. Additionally the transmission between the gates and the source and drain leads is calculated, which allows for the prediction of the gate leakage current within the ballistic model.

3.2.1 Hamiltonian

Within the calculation we employ a single band Hamiltonian with parabolic effective masses. In unstrained silicon there are six degenerate conduction band valleys along the $\Delta-$axes of the Brioullin zone, as depicted in figure 3.3. Each valley has an ellipsoidal shape with highly anisotropic effective masses. In this work we used a value of 0.19 m_0 for the transverse component and 0.92 m_0 for the longitudinal component. Due to symmetry, the six valleys can be reduced to a set of 3 distinct valleys each weighted by a degeneracy factor of 2. For the oxide we assume a spherical mass of 0.5 m_0 within the two-dimensional device plane where the

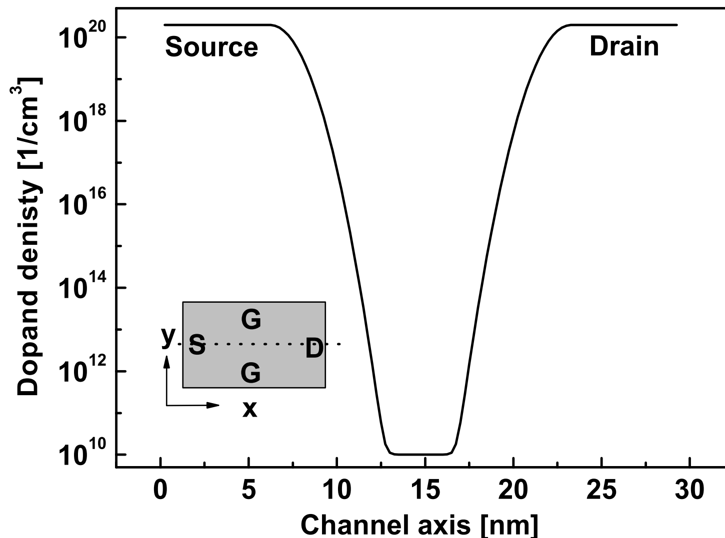


Figure 3.2: The doping profile along the channel is characterized by the two 6 nm long, highly doped source and drain regions, that are connected to the almost intrinsic center via 4 nm transition regions by a gaussian profile.

Schrödinger equation is solved. To simplify the density integration along the direction orthogonal to the device plane, the masses of the silicon are used in the entire domain including the oxide regions.

3.2.2 Discretization of Schrödinger and Poisson equation

For the numerical calculation, the structure is discretized on a rectangular tensor-product grid with a homogenous grid spacing of 2.5 Å, resulting in a total of 7100 nodes. The discretization is performed employing two complementary grids for material parameters and physical fields, as depicted in figure 3.4. The material grid (bold crosses) defines all material properties such as conduction band edge, effective masses, or dielectric constant within a finite volume surrounding each grid point. All physical quantities like wavefunctions, or the electrostatic potential are discretized on the physical grid (bold dots) which has identical grid spacing but is offset with respect to the material grid such that each grid point of the physical grid is placed at an interface between four material volumes.

To obtain a self-consistent solution for the quantum transport problem of an open device, the Schrödinger and Poisson equation have to be solved in an iterative scheme, since the quantum mechanical density resulting from the solution of the Schrödinger equation, modifies the potential and vice versa. The first step is the solution of the Schrödinger equation to calculate the eigenstates of the decoupled device. Within the effective mass model, this equation can be written as

$$-\frac{\hbar^2}{2m_0}\vec{\nabla}\left[\mathbf{M}^{-1}(\mathbf{r})\vec{\nabla}\right]\psi_\alpha(\mathbf{r})+\left[E_c(\mathbf{r})-|e|\,\phi_{\mathrm{H}}(\mathbf{r})+\phi_{\mathrm{XC}}^{\mathrm{LDA}}(\mathbf{r})\right]\psi_\alpha(\mathbf{r})=E_\alpha\psi_\alpha(\mathbf{r}) \tag{3.1}$$

where $\mathbf{M}^{-1}$ is the effective mass tensor, that is diagonal in the frame of the principal

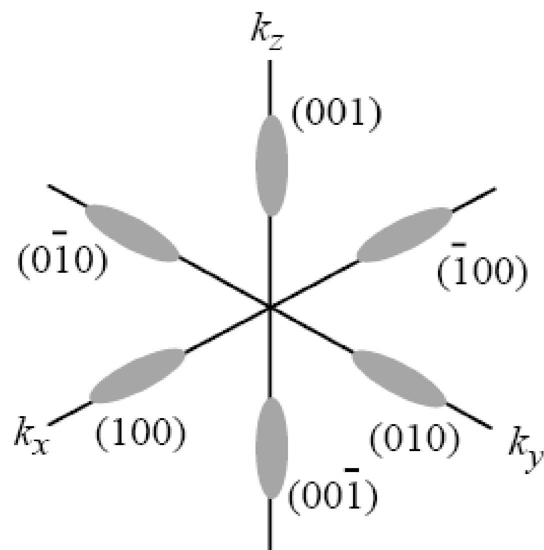


Figure 3.3: The energy iso-surfaces of the six ellipsoidal conduction band valleys of silicon.

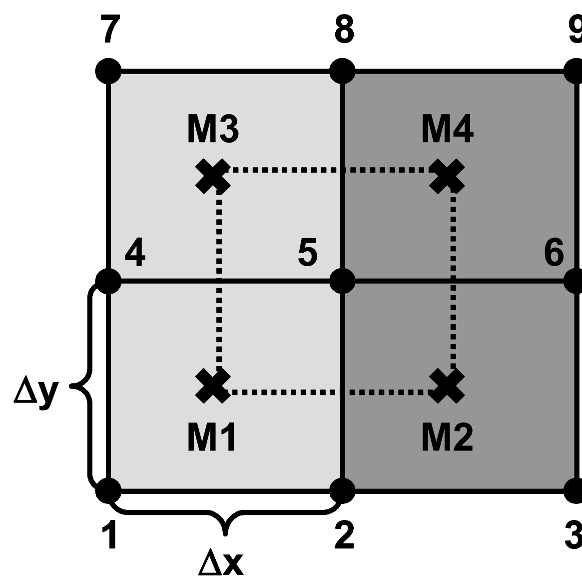


Figure 3.4: The two complementary grids used in the calculation. The numbered solid circles label the grid points on the physical grid whereas the bold crosses mark the grid points of the material grid.

axes of the band minimum, $E_c(\mathbf{r})$ the conduction band edge, $\phi_{\mathrm{H}}(\mathbf{r})$ the electrostatic potential and $\phi_{\mathrm{XC}}^{\mathrm{LDA}}(\mathbf{r})$ the (optional) exchange-correlation potential within local density approximation (LDA). Since the principal axes of the conduction band-edge in silicon coincide with the simulation coordinate system we can rewrite the term

$$\vec{\nabla}\left[\mathbf{M}^{-1}(\mathbf{r})\vec{\nabla}\right] \rightarrow \frac{\partial}{\partial x}\frac{1}{m_1(\mathbf{r})}\frac{\partial}{\partial x} + \frac{\partial}{\partial y}\frac{1}{m_2(\mathbf{r})}\frac{\partial}{\partial y} + \frac{\partial}{\partial z}\frac{1}{m_3(\mathbf{r})}\frac{\partial}{\partial z}.$$

The discretization of eq. (3.1) within the finite differences scheme is exemplified for the grid in fig. 3.4 showing the elements of the 5th row of the Hamilton matrix:

$$\begin{aligned}
H_{5,8} &= -\frac{\hbar^2}{2m_0}\frac{1}{\Delta y^2}\left(\frac{1}{m_2(M_3)} + \frac{1}{m_2(M_4)}\right), \\
H_{5,6} &= -\frac{\hbar^2}{2m_0}\frac{1}{\Delta x^2}\left(\frac{1}{m_1(M_2)} + \frac{1}{m_1(M_4)}\right), \\
H_{5,4} &= -\frac{\hbar^2}{2m_0}\frac{1}{\Delta x^2}\left(\frac{1}{m_1(M_1)} + \frac{1}{m_1(M_3)}\right), \\
H_{5,2} &= -\frac{\hbar^2}{2m_0}\frac{1}{\Delta y^2}\left(\frac{1}{m_2(M_1)} + \frac{1}{m_2(M_2)}\right), \\
H_{5,5} &= \left[E_5^c - |e|\,\phi_5\right] + H_{5,8} + H_{5,6} + H_{5,4} + H_{5,2}.
\end{aligned}$$

The boundary conditions employed for the Schrödinger equation are von Neumann conditions for the contacts (see fig. 3.1)

$$\left.\frac{\partial}{\partial \vec{n}}\psi_i\right|_{i\in\mathrm{contact}} = 0,$$

with $\vec{n}$ being the normal to the contact surface and Dirichlet conditions for the remaining boundary

$$\psi_i = 0 \qquad \forall\, i \notin \{\mathrm{device,lead}\}.$$

For a tensor product grid with constant grid spacings Δx and Δy and a rectangular simulation domain, the Hamiltonian is extremely sparse, with a diagonal and two upper and lower side-diagonals. This simple structure allows for the use of a very efficient storage and matrix-multiplication scheme leading to a fast calculation of the eigenstates.

The second major equation that has to be solved repeatedly within the self-consistent cycle, is the Poisson equation

$$-\vec{\nabla}\left(\boldsymbol{\varepsilon}(\mathbf{r})\,\vec{\nabla}\phi_{\mathrm{H}}(\mathbf{r})\right) = \left(n(\mathbf{r}) - N_D(\mathbf{r})\right), \tag{3.2}$$

where $\boldsymbol{\varepsilon}(\mathbf{r})$ is the dielectric tensor, ϕ_{H} the electrostatic potential, and $n(\mathbf{r})$ and $N_D(\mathbf{r})$ the densities of the electrons and the ionized donors, respectively. Other densities, such as interface states, piezoelectric, or pyroelectric charges would enter the Poisson as well, but in this particular example they can be neglected. The discretization of this equation is analogous to the effective mass Schrödinger equation, with the difference that von Neumann boundary conditions are applied along the entire boundary.

3.2.3 Boundary conditions

Since the applied boundary conditions are very important for the self-consistent solution, it is worth to discuss them in more detail. The Schrödinger equation 3.1 is solved for the decoupled system, using von Neumann boundary conditions at the contacts and Dirichlet for the remaining boundary. This results in a hermitian Hamiltonian and allows one to use a reduced set of eigenstates in the CBR formalism as shown in section 2.7. Therefore, iterative eigensolvers, such as Arnoldi, Jacoby-Davidson, or similar, can be efficiently applied to calculate a limited number of eigenstates. The open boundary conditions used for the calculation of the density of the open device are then introduced via the self-energy term in the Dyson equation (2.4).

In a self-consistent calculation we have to specify additional conditions for the Poisson equation and the distribution functions in the leads. These boundary conditions are very important since different ways of implementation exist that correspond to different physical pictures and lead to slightly different results. We will explain this for the example of the charge accumulation at the drain under applied bias in the ballistic regime. This effect results from high energy modes originating from the source that add to the density at the drain, leading to a increased density close to the contact compared to the equilibrium solution. This is undesirable, since the density and potential distribution at the contact boundary should converge to the one of the corresponding individual lead. Otherwise the position of the contact boundary with respect to the device would highly influence the resulting device behavior. The handles to cope with this problem are the boundary conditions for the electrostatic potential and the distribution function in the leads. A brief overview of this topic is found in the work of Laux [19], where essentially three possibilities are mentioned. The simplest way is to assume an equilibrium Fermi-Dirac distribution in each lead and to shift the quasi Fermi levels to maintain charge neutrality [55, 10]. A second approach, that is able to deal even with large currents is the introduction of a drifted distribution function [55] in the leads that accounts for a net current flow in those leads. The third method to avoid any charge accumulations at the contacts is the use of zero field boundary conditions for the electrostatic potential, while fixing the Fermi level with Dirichlet conditions [29, 56, 57, 58]. This assumes that large reservoirs are connected to the leads that ultimately determine the Fermi levels, whereas the potential in the lead is not fixed, but is actually raised in the drain and lowered in the source as a result of a charge self-consistent calculation. There is still an open debate which conditions are appropriate, but in our opinion it is also a matter of the actual device geometry. In a quasi 1D device which is infinitely extended in the two lateral directions, there is no possibility to have a lead that supports exceedingly more modes than the actual device. Therefore, in this case, the model of the lead as large reservoir no longer holds and it is reasonable to assume that the distribution in the leads is actually drifted as it is inside the device. In a small nanostructure on the other hand that is connected to rather large leads, the picture of the reservoirs makes sense and the third approach is valid as well. In our work we use this way of treating the leads, mainly because of the reasons stated in this section, but also because of simplicity, since we don't have to introduce an extra convergence parameter, such as a drift wave vector [55].

In this work von Neumann boundary conditions are used for the Poisson equation to ensure charge neutrality and Dirichlet conditions fix the chemical potential in the leads. Importantly, the lead modes are calculated self-consistently using the potential at the contacts, obtained from the solution of the Poisson equation of the entire device.

3.3 Calculation of local density of states

In section 2.4 we derived an expression for the carrier density in terms of the density matrix. In principle, within the CBR method, this is the most efficient way to calculate the density, but it turns out that, within a self-consistent scheme, the calculation of the density via the local density of states is favorable, since it allows for the use of a predictor corrector scheme to improve the convergence. In the following we will derive an expression for the local density of states within the CBR formalism. Starting with the expression for the spectral function given in eq. (2.20) we define the spectral function connected to lead λ as

$$S_{ij}^{\lambda}(E) \equiv \left\langle i|\mathbf{G}^R(E)\,\mathbf{\Gamma}^{\lambda}(E)\,\mathbf{G}^{R\dagger}(E)\,|j\right\rangle, \tag{3.3}$$

where $\mathbf{\Gamma}^{\lambda}$ corresponds to the self-energy associated with lead λ and G^R is the retarded Green's function of the open device. Since for the calculation of the density, the non-diagonal components of the spectral function, containing the information about phase coherence, are not relevant, we define the lead connected local density of states (LDOS) [6]:

$$\rho_i^{\lambda}(E) \equiv \frac{1}{2\pi} S_{ii}^{\lambda}(E) = \frac{1}{2\pi}\left\langle i|\mathbf{G}^R\mathbf{\Gamma}^{\lambda}\mathbf{G}^{R\dagger}|i\right\rangle, \tag{3.4}$$

taking into account only the diagonal elements of the spectral function. From this expression the local carrier density is obtained by integration over the energy occupying each level with the distribution function of the corresponding lead. Since $\mathbf{\Gamma}^{\lambda}$ in eq. (3.4) is nonzero only at contact λ, for the evaluation of the LDOS via eq. (3.4) only the sub-block $\mathbf{G}_{DC}^R$ is needed that couples the interior part D of the device to the contact part C. Even further we showed in eq. (2.23) that this also holds for the Green's function of the decoupled device $\mathbf{G}^0$, since $\mathbf{G}_{DC}^R$ is obtained via the Dyson equation:

$$\begin{aligned} \mathbf{G}_{DC}^R &= \mathbf{G}_{DC}^0\mathbf{B}_C^{-1}, \\ \mathbf{B}_C &\equiv \mathbf{1}_C - \mathbf{\Sigma}_C\mathbf{G}_C^0\,. \end{aligned} \tag{3.5}$$

This implies that only a rectangular sub-block of $\mathbf{G}^R$ and $\mathbf{G}^0$ is needed that is of dimension N_{tot} (total number of device grid points) times N_C (number of contact grid points). The size of the problem can be further reduced with the projection onto the subspace of the propagating modes which is of dimension $N_M \ll N_C$ as shown in section 2.8. The procedure and its numerical effort to obtain the LDOS is summarized in the following three steps, where N_E is the number of energy steps, N_{eigen} the number of eigenstates included:

1. Setup of the Green's function for the decoupled device $\mathbf{G}_{XC}^0$ from the calculated eigenstates $|\alpha\rangle$:

$$\mathbf{G}_{i\xi}^0(E) = \sum_{\alpha} \frac{\langle i|\,\alpha\rangle\,\langle\alpha|\,\xi\rangle}{E-\varepsilon_\alpha}, \tag{3.6}$$

   where $\langle i|$ is a vector of the real space basis of the device and $|\xi\rangle$ is the cross-sectional eigenfunction of a propagating mode. The numerical effort of this step is found to be

$$N_E \times N_{eigen} \times N_{tot} \times N_M$$

2. Calculation of the corresponding elements of $\mathbf{G}^R_{XC}$ via eq. (3.5).

$$N_E \times N_{tot} \times N_M^2$$

3. Evaluation of the local density of states (eq. (3.4)) which reduces to

$$\rho_i^\lambda (E) = \frac{1}{2\pi} \langle i|\mathbf{G}^R \mathbf{\Gamma}^\lambda \mathbf{G}^{R\dagger}|i\rangle = \sum_\xi \left|G^R_{i\xi}\right|^2 \mathbf{\Gamma}^\lambda_\xi \tag{3.7}$$

since $\mathbf{\Gamma}^\lambda$ is diagonal in the subspace of the propagating modes and the sum is over all propagating modes ξ of lead λ. Therefore the total costs for the calculation of $\boldsymbol{\rho}^\lambda (E)$ for all leads are

$$N_E \times N_{tot} \times N_M$$

Even though steps one and two look like a significant computational effort, they are pure matrix multiplications of dense matrices that can be very efficiently performed by system-optimized BLAS routines. Therefore in real calculations it turns out that the computation time for the calculation of the local density of states with an energy resolution of 1000 points is roughly the same as for the calculation of the eigenvectors of the decoupled Hamiltonian H^0.

3.4 Bound states

One problem that arises with the self-consistent solution of the Schrödinger-Poisson equation using the ballistic formalism is the treatment of the bound states that may exist in certain nanostructures and that are not accounted for automatically within the ballistic calculation. To explain this we will make use of the picture of the scattering states that propagate from a single lead (source-lead) into the device and are scattered and reflected into all leads connected to the device. In this picture the occupation of the scattering-state is given by the Fermi-distribution of the corresponding source-lead. States that do not propagate in any lead because they are bound within the device or have an eigenenergy below the energy of the lowest propagating mode are still accounted for in the density of states that turns into a delta function in this case (see appendix A), but no Fermi level can be associated with the state since the Fermi levels are only defined inside the leads where the state does not propagate. In this situation the purely ballistic density would not account for the part of the density that corresponds to the bound states.

Physically however, it is reasonable to include those bound states into the calculation of the density since we are only looking for steady state solutions of the open system. In this case a single inelastic scattering event within an infinitely long time would be enough to occupy the bound state that does not couple to any decay channel in the leads and therefore will remain occupied. Thus, even though inelastic scattering is not present in the ballistic approach, we propose to occupy the bound states, while calculating the transport itself within the ballistic limit.

In the following we classify the bound states into three classes that we term 'truly bound', 'quasi bound' and 'resonant' states and demonstrate the importance of their proper inclusion into the total carrier density. Additionally, we introduce a combined density of states that includes both the propagating and the bound states and present a numerical scheme for the detection of the bound states.

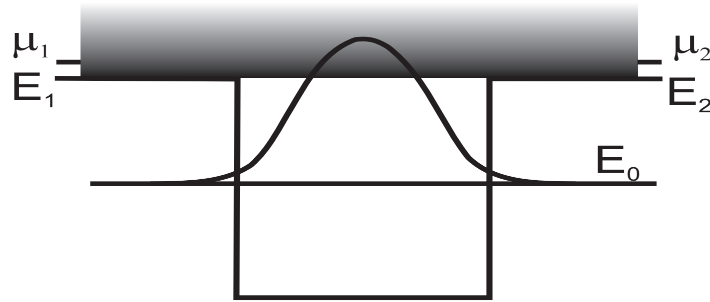


Figure 3.5: Schematic picture of a quantum well containing a truly bound state with an eigenenergy E_0 below the mode energies of the leads L_1 and L_2. The shaded grey region indicates the continous part of the electron density that is coupled to the leads.

3.4.1 Truly bound states

A 'truly bound' state is a state that does not propagate in any lead of the device and therefore is not connected to any reservoir with a well defined Fermi level. To exemplify this situation we consider the quantum well depicted in fig. 3.5 that supports a deeply bound state of eigenenergy E_0 below the energies E_1 and E_2 of the lowest propagating modes in the leads. In this case the density resulting from the ballistic formalism will not account for this state, since inelastic scattering is not allowed. However, as explained in the previous section in a steady state calculation it is reasonable to occupy this state, for example with the distribution function corresponding to the lead with the lowest Fermi-level.What seems to be physically reasonable turns out to be absolutely necessary to obtain a self-consistent solution of the Schrödinger-Poisson equation. To demonstrate this, we calculated the self-consistent potential (solid line) and density (dashed line) for a 5 nm wide and 400 meV deep, homogeneously doped quantum well using either the purely ballistic approach (see fig. 3.6a)) or assuming that the bound state is occupied (see fig. 3.6b)).

The completely unphysical results of the calculation in the case that the bound state remains unoccupied can be explained as follows. The kinetic energy of the scattering states is higher in the quantum well than in the contact regions, leading to a lower local density, since the flux has to be conserved. To compensate for this lack of density in the well, the potential is lowered which results in a even stronger confinement supporting several bound states that remain unoccupied. This behavior is catastrophic and does not lead to any converged solution. However, if the bound state is included into the density, the self-consistent calculation converges to a physically meaningful result as shown in figure 3.6. It is observed that the density in the well is higher than in the contacts as expected and the potential is actually raised in the well to account for the surplus of electrons. This result shows that in a self-consistent calculation it is crucial to take into account the truly bound states in order to obtain converged and physical results.

3.4.2 Quasi bound states

A second class that we will term 'quasi bound' states, are states that decay in most leads but propagate in at least one single lead to which they are weakly coupled via a tunneling-barrier.

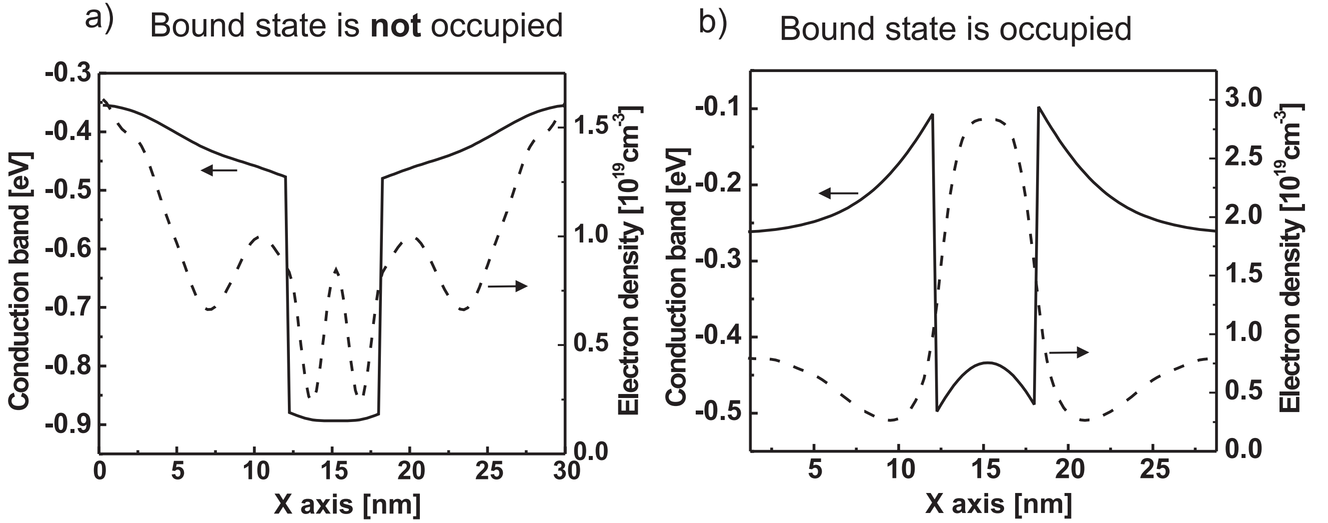


Figure 3.6: Self-consistent solution of the potential (solid line) and density (dashed line) for a quantum well. Figure a) shows the unphysical result for the case that the bound state is unoccupied, and fig. b) shows the result of a calculation that occupies the bound state explicitly.

Although these states can be associated to a contact in which they propagate and therefore be occupied with a well defined Fermi-level, the results become unphysical if the device is under applied bias. An example of this case is a single barrier structure as depicted in fig. 3.7, where a 'quasi bound' state is supported by the triangular well formed in front of the tunneling barrier. Since the eigenenergy E_0 of this state is below the cutoff-energy E_1 of the lowest propagating mode of the first lead, the quasi bound state is solely coupled to the second lead via the tunnel-coupling and will therefore be occupied by a Fermi-distribution with the Fermi-level μ_2. However, with increasing bias, the difference between E_0 and μ_2 rises and the state becomes effectively unoccupied. This result is certainly unphysical, since, in a simple capacitor picture, a charge accumulation on the left side of the barrier in fig. 3.7 is expected.

Similarly to the 'truly bound' states we argue, that in the case of very weak tunnel-coupling, the state is very likely be occupied via an eventual scattering process from the overlapping continuum of the propagating states corresponding to the left lead (grey shaded region) and will decay only very slowly via tunneling-escape. Thus in contrast to the purely ballistic formalism in which the state would be occupied with the Fermi-level μ_2 of the second lead, we propose to occupy it with the Fermi-level μ_1 of the first lead to account for the scattering that occurs in real systems. In a calculation that includes scattering [59] the occupation of the 'quasi bound' states has been found to be extremely important to yield the correct self-consistent potential in a resonant tunneling structure.

In the example of the double gate MosFet, such a quasi bound state is localized in the depletion layer that forms at the interface of the homogeneously doped gate contact and the oxide barrier. This layer arises due to the fact that the positively charged ionized donors are homogeneously distributed even close to the barrier, whereas the quantum mechanical electron density is repelled from the barrier, since the wavefunctions have to vanish inside the barrier. This charge imbalance leads to a triangular confinement potential called depletion layer. Depending on the effective mass and the depth of the potential well, bound states may or may not exist. Figure 3.8 shows a one-dimensional slice of the conduction band edge (solid

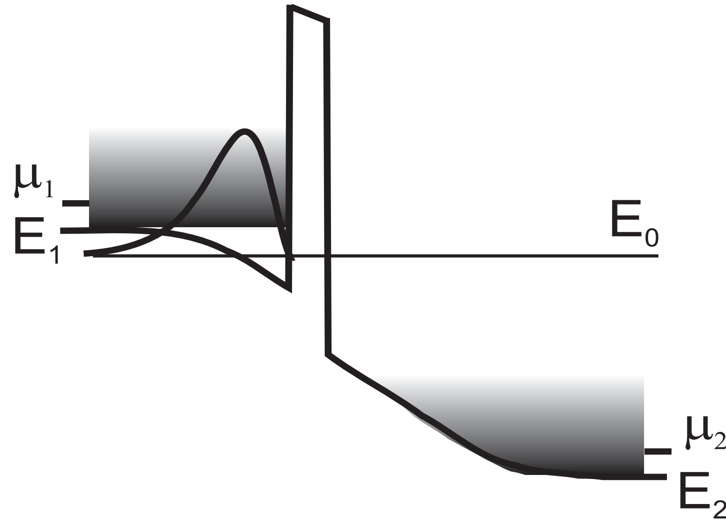


Figure 3.7: Quasi bound state that is located in front of a barrier with a bias voltage applied.

line) within one of the gates resulting from a converged self-consistent solution. At the interface to the oxide barrier a triangular potential well is formed with maximum confinement energy of about 100 meV and a width of roughly 1 nm. It turns out that only for the valleys with high effective mass (0.92 m_0) in y-direction the confinement is strong enough to support a bound state. The probability density and the eigenenergy of this bound state are indicated in fig. 3.8 by the dashed and dotted lines, respectively. Although the eigenenergy of the bound state is clearly above the conduction band edge at the interface to the gate contact, it is still several meV below the onset of the first propagating mode (indicated by the solid rectangle) of the semi-infinite lead and therefore the wavefunction has to decay exponentially into the gate. Thus the only escape channel of this state is through the 1 nm thick oxide barrier and into the source or drain contacts, depending on the applied voltages. But since the rate of this transition is rather small, we can consider this state to be a quasi bound state.

Regarding the example of the MosFet the occupation of the quasi bound state has a significant effect, since the characteristics of the device depends strongly on the charge accumulation and charge distribution in the gates as will be shown in the discussion of the numerical results.

3.4.3 Resonant states

The third class of states is the class of the resonant states which occur in double barrier structures as depicted in fig. 3.9. These states are characterized by an eigenenergy that is high enough to couple to propagating modes of all leads, but the coupling via the tunneling barriers is very weak. In principle these states are fully accounted for in the ballistic scheme since they couple to propagating modes in the leads and can therefore be occupied via the Fermi levels in the leads. Technically however, these states might pose a problem especially in a self-consistent cycle, since they appear as narrow peaks in the transmission and the density of states that have to be resolved with high accuracy. Therefore it is important to single out these resonant states within the algorithm and resolve them using a locally adjusted energy grid. Within the CBR formalism this is easily achieved since the eigenenergies of the resonance states are known approximately from the eigenenergies of the decoupled system.

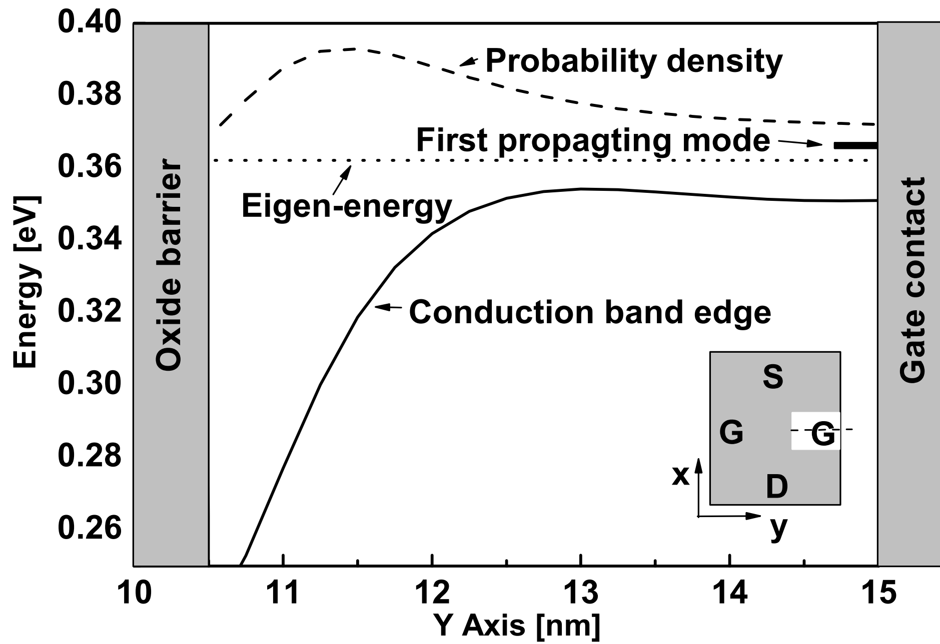


Figure 3.8: The conduction band edge (solid line) forms a triangular well in front of the oxide barrier (grey). The probability density and the eigenenergy of the bound state are indicated by the dashed and dotted lines, respectively. The onset of the first propagating mode in the gate lead is indicated by the bold dash.

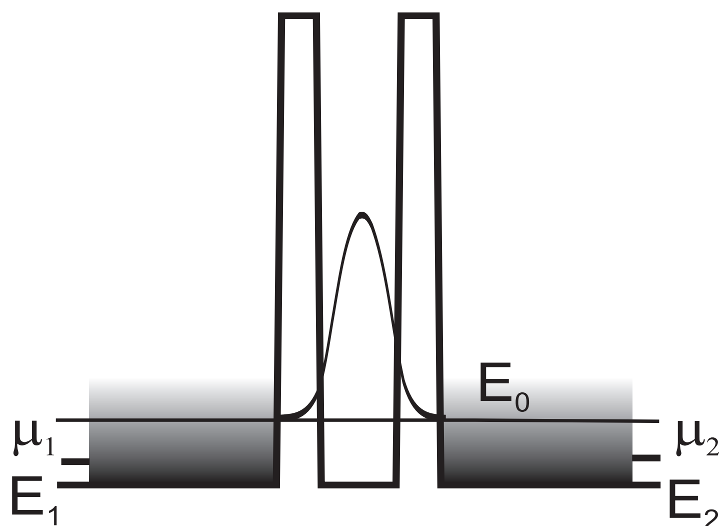


Figure 3.9: Resonant state located inside a double barrier.

3.5 Combined density of states

There are two possible solutions to obtain a physically meaningful density in ballistic devices where bound states are present that correspond to at least one of the three classes defined in the previous section. For the 'resonant' states is suffices to define an appropriate energy grid, that resolves the corresponding peaks in the density of states well enough to obtain a converging self-consistent algorithm. The 'truly' and 'quasi' bound states however require an extra treatment outside of the purely ballistic approach. By definition, the bound states are weakly coupled to the leads and therefore do not mix with the propagating states. Within this limit the corresponding eigenstates of the open device Hamiltonian H are approximately the same as the eigenstates of the decoupled device Hamiltonian H^0 that is used within the CBR method. This fact allows us to define a combined density of states $\rho_i^\lambda(E)$ associated with lead λ

$$\rho_i^\lambda(E) = \rho_i^{\lambda,\mathrm{disc}}(E) + \rho_i^{\lambda,\mathrm{cont}}(E) \tag{3.8}$$

that consists of a continuous part $\rho_i^{\lambda,\mathrm{cont}}(E)$ that accounts for the propagating states and is obtained via eq. (3.7) and a discrete part $\rho_i^{\lambda,\mathrm{disc}}(E)$ that accounts for the bound states. The latter part is defined in the following way

$$\rho_i^{\lambda,\mathrm{disc}}(E) = \sum_\alpha \left|\psi_i^{\alpha,\lambda}\right|^2 \delta(E - E_\alpha) \tag{3.9}$$

where the index α is summed over all bound and quasi bound states $\psi^{\alpha,\lambda}$ associated with lead λ. This association of the bound state to a certain lead is necessary, because the Fermi-levels are only defined inside the leads. One possible idea to find the corresponding reservoir would be to consider the overlap B^λ between the propagating states originating from lead λ and the bound state probability density

$$B^\lambda = \sum_i \left|\psi_i^\lambda\right|^2 \int dE\ \rho_i^{\lambda,\mathrm{cont}}(E).$$

In this way we obtain a rough measure for the assumed scattering from propagating states injected from contact λ to the bound state that can be used to determine the reservoir which is most strongly coupled to the bound state. One further issue has to be addressed before we can finally calculate the density of the device. The bound states found in the gate region of the DGFET device are usually bound by a few meV such they decay rather slowly into the lead. In order to account for the resulting density in a correct way, one would either have to extend the gate region inside the simulation domain which increases the computational costs significantly, or one could renormalize the bound states to account for the portion that is located outside of the simulation domain. Since the state decays exponentially into the gate this is very well possible as shown in appendix C and it turns out, that in the considered examples always more than 90 % of the state is located inside of the domain.

3.5.1 Calculation of the carrier density

The combined density of states as defined in eq. (3.8) is used to calculate the carrier density within the self-consistent cycle. To obtain the local carrier density, the local density of states

has to be integrated over the energy E and the remaining k-vector k_z

$$n(\mathbf{r}) = g_{\text{spin}} \frac{1}{2\pi} \sum_v g_{\text{valley}} \sum_{\lambda=1}^{L} \int dE \int dkz \rho^\lambda(\mathbf{r}, E) f\left(\left(E + \frac{\hbar^2 k_z^2}{2m_z} - E_F^\lambda\right) / k_B T\right)$$

which can be rewritten using the Fermi-integral $F_{-1/2}$ [60]

$$n(\mathbf{r}) = g_{\text{spin}} \frac{1}{2} \sqrt{\frac{m_z k_B T}{\hbar^2 2\pi}} \sum_v g_{\text{valley}} \sum_{\lambda=1}^{L} \int dE \rho^\lambda(\mathbf{r}, E) F_{-1/2}\left(\left(E_F^\lambda - E\right) / k_B T\right) \tag{3.10}$$

where $g_{\text{spin}} = 2$ is the spin degeneracy, m_z the effective mass in z direction perpendicular to the device plane. The first sum accounts for the different conduction band valleys v that are weighted with a degeneracy factor $g_{\text{valley}} = 2$. Additionally one has to sum over the contacts λ since each one has a different Fermi-level E_F^λ. For further details on the Fermi integrals we refer to the dissertation of Stefan Hackenbuchner [60] as well as to Alex Trellakis [61, 62] who developed and implemented a very efficient numerical algorithm for the evaluation of the Fermi-integrals. This is very important, since the density has to be calculated many times during the self-consistent cycle.

3.5.2 Detection of bound states

In the previous section we tried to classify the various sorts of bound states and introduced the combined density that accounts for bound as well as for propagating states. In the following we will now describe a simple way to detect the bound states. Thus, in order to proceed further, we need a valid criterion to assess whether a state is bound or belongs to the continuum. Considering the spectral function $\mathbf{S}_{i,j}$ written in terms of the eigenstates $|\Phi_\alpha\rangle$ and $|\Psi_\alpha\rangle$ of the open device Hamiltonian H [6]

$$\mathbf{S}(\mathbf{r}, \mathbf{r}', E) = \sum_\alpha \Phi_\alpha(\mathbf{r}) \Psi_\alpha^*(\mathbf{r}') \frac{\gamma_\alpha}{(E - \varepsilon_{\alpha 0} + \Delta_\alpha)^2 + (\gamma_\alpha/2)^2}$$

where $\varepsilon_{n\alpha 0}$ is the eigenvalue of the state in a decoupled system and Δ_α and γ_α are the shifts on the real and imaginary axis, respectively that originate from the self energy. Since, in our model, we neglect any other sources of scattering besides the coupling to the contact reservoirs, the self-energy $\Sigma_\alpha = \Delta_\alpha + i\gamma_\alpha$ for state α has to be proportional to the coupling of the eigenstate $|\psi^\alpha\rangle$ of the decoupled system to the propagating lead modes $|\xi_l^m\rangle$.

$$\Sigma_\alpha \sim \sum_{\lambda,m} |\langle \xi_\lambda^m | \psi^\alpha \rangle|^2 \equiv C_\alpha^2 \tag{3.11}$$

where the index λ denotes the lead and m is summed over all modes of the lead that are propagating at the eigenenergy $\varepsilon_{\alpha 0}$. In this way we define a measure C_α for the total coupling strength of each eigenstate $|\psi^\alpha\rangle$ to the propagating lead modes. Although this is just a crude estimate, a comparison between the values of C_α for all states reveals that this actually helps to distinguish bound states from states that belong to the continuum. Figure 3.10 shows the coupling strengths C_α resulting from eq. (3.11) as a function of the eigenenergy for all eigenstates within the desired energy range for the DGFET example. The squares labeled with numbers 1 to 5 can be clearly distinguished from the main part of the states since their coupling strength

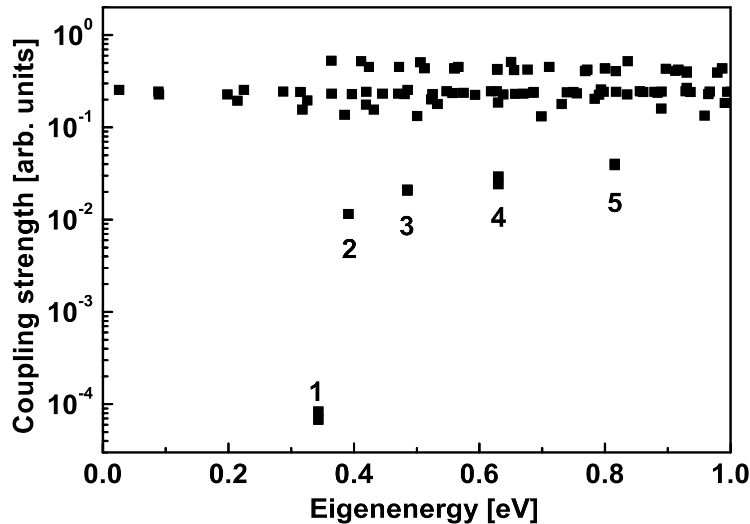


Figure 3.10: The magnitude of the coupling of each individual eigenstate to all propagating lead modes is plotted as a function of the eigenenergy. The numbered squares correspond to the twofold degenerate, weakly coupled states that can be classified as quasi bound, or resonant states.

is orders of magnitude smaller than the average. Due to this significant deviation from the mean, it is possible to find a threshold value for C_α that reliably separates these states from the continuum, which is important for the numerical implementation. The state labeled with number 1 features a very low coupling and can be classified as a quasi bound state since its eigenenergy is lower than the smallest mode energy in the corresponding gate contact. The other states labeled 2 to 5 are the so called resonant states that are weakly coupled but have eigenenergies in the range of the propagating modes of all leads.

The purely ballistic density of states corresponding to the gate contact of the DGFET, shown in fig. 3.11, exhibits the expected sharp resonance peaks labeled by the numbers 3 to 5. The quasi bound state labeled with number 1 in fig. 3.10 does not show up since the eigenenergy is below the lowest propagating mode and therefore it is not contained in the density of states corresponding to the gate. The resonance states however show up as peaks with heights varying arbitrarily from zero (peak 2) to 1000 (peak 3) since for an energy resolution of 1 meV they are very poorly resolved. This picture clearly shows why it is important to extract the quasi-bound states and increase the resolution for the resonant states since for the DOS shown in fig. 3.11 convergence will never be achieved.

3.6 Self-consistent solution

The self-consistent solution of the ballistic transport properties of an open device requires the repeated solution of the Schrödinger and Poisson equation due to their coupling via the potential and the quantum mechanical density. The algorithm that is used within the self-consistent implementation of the CBR method is presented in terms of a flow-chart in appendix D.2 and

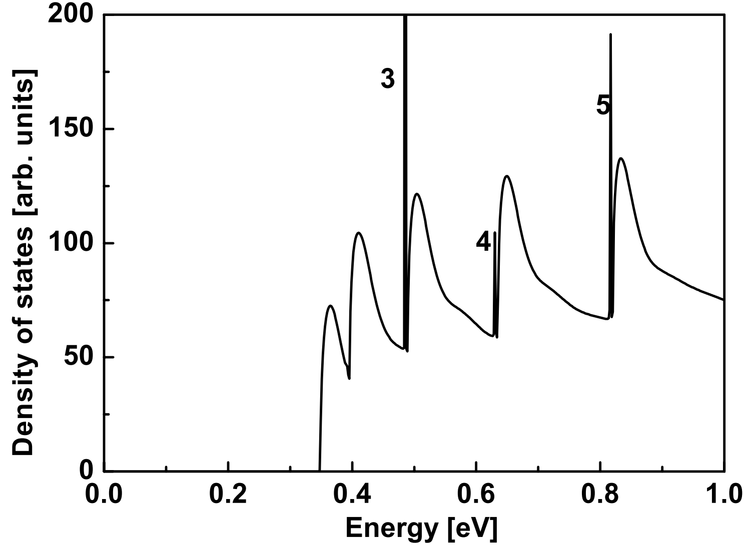


Figure 3.11: The density of states is plotted as a function of energy. The numbers label the peaks of the quasi bound states.

will be explained in more detail in the following section. In passing we note that all quantities denoted in bold letters are vectors in real space and the indices i and j are iteration counters for the outer and inner loop, respectively. We will start with the definition of the basic equations and then describe the self-consistent algorithm.

The Schrödinger equation for the decoupled device

$$\mathbf{H}^0 \left[\boldsymbol{\phi}_{\mathrm{H}}, \boldsymbol{\phi}_{\mathrm{XC}}\right] \cdot \boldsymbol{\psi}_\alpha = E_\alpha \boldsymbol{\psi}_\alpha, \tag{3.12}$$

with the Hamilton matrix $\mathbf{H}^0$, depends on the exchange-correlation potential $\boldsymbol{\phi}_{\mathrm{XC}}$ and the Hartree potential $\boldsymbol{\phi}_{\mathrm{H}}$ (see eq. 3.1). The resulting eigenfunctions $\boldsymbol{\psi}_\alpha$ form a basis that is used within the CBR formalism to obtain a local density of states $\boldsymbol{\rho}(E)$ of the open device (see eq. 3.4)

$$\mathrm{CBR}\left[\boldsymbol{\psi}_\alpha\right] \Rightarrow \boldsymbol{\rho}(E). \tag{3.13}$$

To obtain the local carrier density $\mathbf{n}$, the density of states has to be integrated over the energy and weighted with the Fermi-distribution of the leads (see eq. 3.10)

$$\mathbf{n} = g_{\mathrm{spin}} \frac{1}{2} \sqrt{\frac{m_z k_B T}{\hbar^2 2\pi}} \sum_v g_{\mathrm{valley}} \sum_{\lambda=1}^{L} \int dE\ \boldsymbol{\rho}^\lambda(E) F_{-1/2}\left(\left(E_F^\lambda - E\right)/k_B T\right). \tag{3.14}$$

Finally the potentials $\boldsymbol{\phi}_{\mathrm{H}}$ and $\boldsymbol{\phi}_{\mathrm{XC}}$ have to be calculated as a function of the density. While the exchange-correlation potential $\boldsymbol{\phi}_{\mathrm{XC}}$ is a direct function of the density within the local density approximation (LDA)

$$\boldsymbol{\phi}_{\mathrm{XC}} = \boldsymbol{\mu}_{\mathrm{X}}^{\mathrm{LDA}}\left[\mathbf{n}\right] + \boldsymbol{\mu}_{\mathrm{C}}^{\mathrm{LDA}}\left[\mathbf{n}\right], \tag{3.15}$$

the Hartree potential $\boldsymbol{\phi}_{\mathrm{H}}$ is obtained from the solution of the Poisson equation (see eq. 3.2)

$$\mathbf{A} \cdot \boldsymbol{\phi}_{\mathrm{H}} = \mathbf{n} - \mathbf{N}_D, \tag{3.16}$$

where $\mathbf{A}$ is the Poisson matrix derived from the discretized Laplacian operator, and $\mathbf{N}_D$ is the density of the positively charged donors.

These equations form a complete set that determines the ballistic transport problem of an open system within the CBR method. In principle it is possible to simply iterate the solution of the Schrödinger and Poisson equations and with enough damping this will yield a converged result. However, although the solution of the Schrödinger equation is efficient due to the use of iterative solvers, it is still costly enough that hundreds of iteration steps for each bias point do not pose a reasonable scheme. To improve the convergence of a highly non-linear set of coupled equations, such as the Schrödinger-Poisson problem, the Newton algorithm is usually the first choice. But since the exact Jacobian cannot be derived analytically and a numerical evaluation would be too costly, the simple adaptation of the method is not feasible.

For the case of a closed system this problem has been solved using a predictor corrector approach [63]. The aim of this method is to find a good approximation for the quantum density as a function of the electrostatic potential where an expression for the Jacobian is known. Within this approximation the non-linear Poisson equation can be efficiently solved using the Newton scheme resulting in a predictor update for the Hartree potential and the carrier density. If this approximation is close enough to the real quantum density, only very few updates will be necessary to yield a converged solution. For the DGFET example calculated in this work, this implies that the Schrödinger equation usually has to be solved less than 10 times until the potentials and the density are sufficiently converged.

The self-consistent cycle as implemented in the self-consistent CBR code can be divided into two loops, the main loop and the predictor step that includes the Newton algorithm.

3.6.1 Main loop

The first part of the main loop (see appendix D.2) consists of the solution of the Schrödinger equation (eq. 3.12) with the potentials $\boldsymbol{\phi}_{\mathrm{H}}^{i}$ and $\boldsymbol{\phi}_{\mathrm{XC}}^{i}$ of the ith iteration step and the subsequent calculation of the local density of states $\boldsymbol{\rho}^{i}(E)$ (eq. 3.13) and the carrier density $\mathbf{n}^{i}$ (eq. 3.14). The Hartree potential $\boldsymbol{\phi}_{\mathrm{H}}^{i}$ and the density $\mathbf{n}^{i}$ are then used to calculate the residuum $\mathbf{F}^{i}$ of the Poisson equation

$$\mathbf{F}^{i}\left[\boldsymbol{\phi}_{\mathrm{H}}^{i}\right] = \mathbf{A}\cdot\boldsymbol{\phi}_{\mathrm{H}}^{i} - \left(\mathbf{n}^{i} - \mathbf{N}_{D}\right). \tag{3.17}$$

If the magnitude of this residuum vector is below a certain threshold

$$\left|\mathbf{F}^{i}\right| < \varepsilon_{\mathrm{threshold}},$$

the calculation is assumed to be converged and the loop cycle will be terminated. If the residuum is still too large, the correction to the Hartree potential is calculated in the Predictor step which will be explained in the next subsection. The resulting correction to the Hartree potential $\Delta\boldsymbol{\phi}_{\mathrm{H}}^{i}$ and the corresponding carrier density are used to update the Hartree and exchange-correlation potentials for the next step $i+1$

$$\begin{aligned}
\boldsymbol{\phi}_{\mathrm{H}}^{i+1} &= \boldsymbol{\phi}_{\mathrm{H}}^{i} + \Delta\boldsymbol{\phi}_{\mathrm{H}}^{i},\\
\boldsymbol{\phi}_{\mathrm{XC}}^{i+1} &= \boldsymbol{\mu}_{\mathrm{X}}^{\mathrm{LDA}}\left[\mathbf{n}_{\mathrm{predictor}}^{i}\left[\Delta\boldsymbol{\phi}_{\mathrm{H}}^{i}\right]\right] + \boldsymbol{\mu}_{\mathrm{C}}^{\mathrm{LDA}}\left[\mathbf{n}_{\mathrm{predictor}}^{i}\left[\Delta\boldsymbol{\phi}_{\mathrm{H}}^{i}\right]\right].
\end{aligned}$$

This loop is repeated until convergence is obtained.

3.6.2 Predictor step

Within the predictor step (see appendix D.2) the density of states $\boldsymbol{\rho}^i$ from the ith step of the main loop is used to obtain the predictor density $\mathbf{n}^i_{\text{predictor}}$ that is defined to be a direct function of the correction to the Hartree potential $\Delta\boldsymbol{\phi}^j_{\text{H}}$ and approximately reflects the correct quantum mechanical density

$$\mathbf{n}^i_{\text{predictor}}\left[\Delta\boldsymbol{\phi}^j_{\text{H}}\right] = g_{\text{spin}}\frac{1}{2}\sqrt{\frac{m_z k_B T}{\hbar^2 2\pi}}\sum_v g_{\text{valley}}\sum_{\lambda=1}^{L}\int dE\,\boldsymbol{\rho}^i_\lambda(E)F_{-1/2}\left(\left(E_F^\lambda - E - |e|\,\Delta\boldsymbol{\phi}^j_{\text{H}}\right)/k_B T\right). \tag{3.18}$$

The idea behind this approximation is that to first order the wavefunctions and therefore the density of states, remain unchanged for small deviations in the potential and only the eigenenergies are adjusted locally to changes in the potential. This predictor density is now used to solve the non-linear Poisson equation

$$\mathbf{F}^j_{\text{predictor}}\left[\boldsymbol{\phi}^i_{\text{H}}, \Delta\boldsymbol{\phi}^j_{\text{H}}\right] = \mathbf{A}\cdot\left(\boldsymbol{\phi}^i_{\text{H}} + \Delta\boldsymbol{\phi}^j_{\text{H}}\right) - \left(\mathbf{n}^i_{\text{predictor}}\left[\Delta\boldsymbol{\phi}^j_{\text{H}}\right] - \mathbf{N}_D\right), \tag{3.19}$$

with the condition that the magnitude of the residuum $\mathbf{F}^j_{\text{predictor}}$ for the approximate density has to be smaller than the residuum $\mathbf{F}^i$ of the main loop, by some factor $\varepsilon_{\text{predictor}}$

$$\left|\mathbf{F}^j_{\text{predictor}}\right| \overset{!}{<} \varepsilon_{\text{predictor}}\cdot\left|\mathbf{F}^i\right|, \tag{3.20}$$

which is usually on the order of 10^{-4}–10^{-6} to guarantee good convergence. The minimization of the functional defined in eq. (3.19) can be performed using a Newton scheme since the derivative of the predictor density with respect to $\Delta\boldsymbol{\phi}^j_{\text{H}}$ is obtained analytically by replacing the Fermi integral $F_{-1/2}$ by $F_{-3/2}$ in equation (3.18). Within the Newton method a Taylor expansion of the functional $\mathbf{F}^j_{\text{predictor}}$ to first order

$$0 = \mathbf{F}^j_{\text{predictor}}\left[\boldsymbol{\phi}^i_{\text{H}}, \Delta\boldsymbol{\phi}^j_{\text{H}} + \delta\boldsymbol{\phi}^j_{\text{H}}\right] = \mathbf{F}^i_{\text{predictor}}\left[\boldsymbol{\phi}^i_{\text{H}}, \Delta\boldsymbol{\phi}^j_{\text{H}}\right] + \frac{\partial\mathbf{F}^i_{\text{predictor}}}{\partial\Delta\boldsymbol{\phi}_{\text{H}}}\delta\boldsymbol{\phi}^j_{\text{H}}$$

is used to determine the Newton correction

$$\delta\boldsymbol{\phi}^j_{\text{H}} = -\mathbf{J}^{-1}\mathbf{F}^i_{\text{predictor}}\left[\boldsymbol{\phi}^i_{\text{H}}, \Delta\boldsymbol{\phi}^j_{\text{H}}\right],$$

where $\mathbf{J}$ is the Jacobi matrix

$$\mathbf{J} = \left.\frac{\partial\mathbf{F}^i_{\text{predictor}}}{\partial\Delta\boldsymbol{\phi}_{\text{H}}}\right|_{\partial\Delta\boldsymbol{\phi}_{\text{H}}=\partial\Delta\boldsymbol{\phi}^j_{\text{H}}}.$$

This Jacobian is simply the Poisson matrix plus the derivative of the predictor density

$$\mathbf{J} = \mathbf{A}+\frac{\partial\mathbf{n}^i_{\text{predictor}}}{\partial\Delta\boldsymbol{\phi}_{\text{H}}}.$$

To improve the convergence for each Newton correction a line minimization procedure is employed that minimizes the functional along the direction of the Newton correction as a function of a scalar α using the bi-section method. Once this line search is converged, the predictor potential is updated for the next iteration in the Newton-cycle

$$\Delta\boldsymbol{\phi}^{j+1}_{\text{H}} = \Delta\boldsymbol{\phi}^j_{\text{H}} + \alpha\cdot\delta\boldsymbol{\phi}^j_{\text{H}}.$$

If the convergence criterion of eq. (3.20) is matched, the correction for the Hartree potential for the outer loop is updated

$$\Delta\phi_{\mathrm{H}}^{i} = \Delta\phi_{\mathrm{H}}^{j},$$

and the predictor step is finished.

It turns out, that the predictor-corrector approach works equally well for closed and open systems which is not obvious, since the density of states of an open system in the ballistic limit is strongly coupled to the leads. Therefore, the boundary conditions at the leads are expected to influence the convergence, but since in this work only von Neumann boundary conditions were used for the electrostatic potential we cannot tell whether the predictor-corrector scheme works for other boundary conditions as well. In the work of Laux *et al.* [19] a scheme similar to the predictor-corrector scheme is used for the drifted distribution contacts.

3.6.3 Energy grid

For the numerical implementation of a self consistent scheme using a continuous density of states, the energy grid is of high importance. To integrate the carrier density, we discretize the local DOS in energy space and then employ a simple numerical integration by summing up the values for each energy step weighted by the Fermi distribution and the energy grid spacing ΔE. Since the DOS is a very spiky function with peaks corresponding to highly localized states and due to the onset of the propagating modes, it is very important for the convergence to have these features properly resolved. The peaks resulting from the quasi bound states have already been eliminated by using the combined density of states (eq. 3.4), so that the main structural features in the DOS are due to the lead modes. In our DGFET example, the channel acts as a 1D wire, which leads to peaks that have a $1/\sqrt{E-E_m}$ dependence, where E_m is the mode energy. Thus using a regular grid with constant grid spacing, the integral over the peak is very poor since the relative distance between the nearest energy grid point E_i and the mode energy E_m is arbitrary. Additionally the mode energy is slightly shifted with each iteration step, leading to a varying integration error during the self-consistent cycle, which is a sure kill for any self-consistent algorithm. Thus a solution to this problem is to use the physical information we have about the system and employ an adaptive energy grid that resolves each known peak m with a local energy grid of a few tens of grid points that is fixed to the mode energy E_m. Additionally extra points are distributed in the space between the peaks to obtain a smooth enough energy grid. Figure 3.12 shows the improvement of the resolution of the peak compared to a regular grid with the same total number of energy points. Importantly, the integration error is reduced compared to the regular grid and also remains constant within the iteration, since the grid is locally fixed to the shifting mode energies. This effect can be observed in fig. 3.13, where the convergence behavior of a regular grid with a total of 7000 grid points is compared to the adaptive grid with only 1000 points. For the first iteration steps, the convergence is the same in both cases, but then the regular grid reaches a bottom at a residuum of 10^{-5} which is due to the fluctuating integration error, whereas the adaptive grid guarantees almost perfect convergence.

3.7 Calculational results

In the following section, a detailed analysis of the results of the self-consistent calculation of the double gate nano-MosFet (DGFET, see fig. 3.1) is presented. Generally, there are two classes of

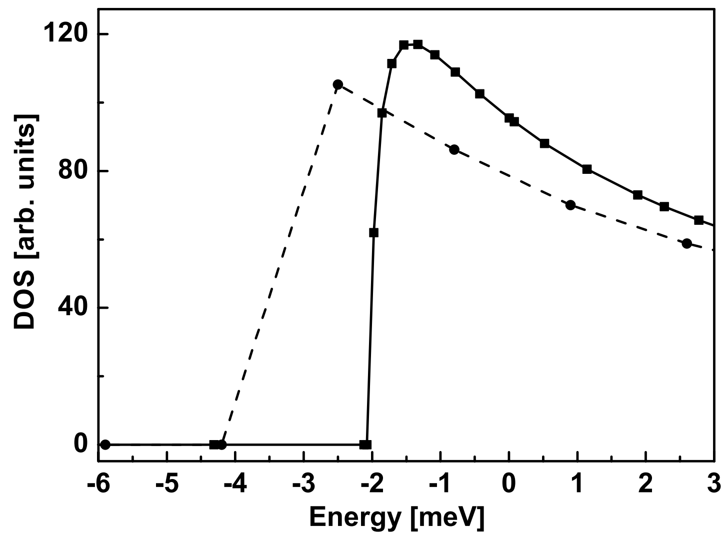


Figure 3.12: Comparison of the energy resolution in the vicinity of a peak in the density of states. Both, the regular grid (solid circles) and the adaptive grid (solid squares) have the same average grid spacing, but locally the adaptive grid may be much finer.

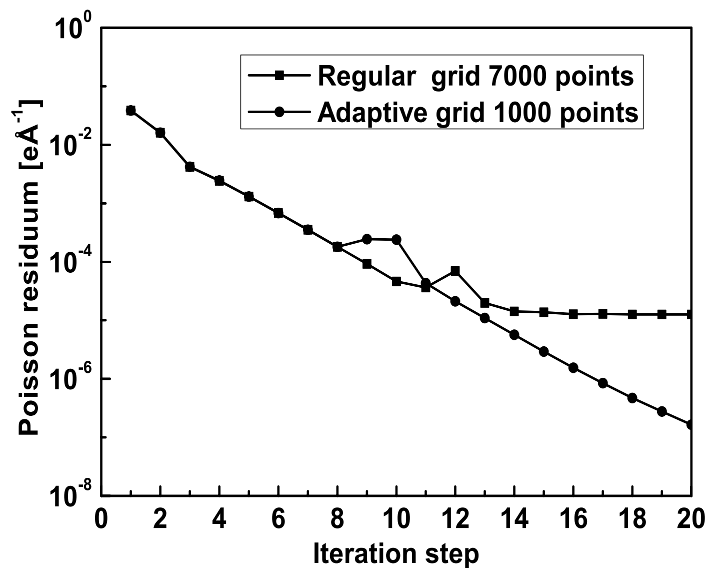


Figure 3.13: The convergence of the regular grid (solid squares) ends abruptly due to the fluctuating integration error. The adaptive grid guaranties good convergence even with less grid points.

calculations, ones that include scattering [64, 57], but are restricted to small simulation areas, thus neglecting the quantum mechanical treatment of the leads, and the ballistic approach [19], that is able to handle larger device geometries and to yield the gate leakage current, on the expense of overestimating the resulting on-current by roughly a factor of two [64]. In this work we did not include any scattering, but we will focus on the fully quantum mechanical treatment of the gates. The gate regions included in the device domain are 9 nm long and 5 nm thick, which is large enough to allow for a converged potential at the contact. For the purpose of the calculation we treat the poly-gate as single-crystal silicon with the same band-structure and effective masses and same doping as for the silicon in the source or drain contacts. Although it is not entirely clear if this approach is valid it has been used by others [19, 65], and a more realistic description is most likely restricted to the use of atomistic Hamiltonians, such as tight binding, or pseudopotentials, which is out of reach without the use of supercomputers.

The results presented in the following are obtained with the program `CBR_selfcon` that has been developed as part of this Ph.D. thesis. The program will be freely available and a short user manual is given in the appendix.

3.7.1 Self-consistent potential and carrier density

In the following we present the self-consistent solution of the potential and the electron density in the DGFET under applied bias. We will analyze the shape of the potential, shown in fig. 3.14, at the contacts with respect to the applied boundary conditions and investigate the dependence on the applied gate voltage. Since the potential is calculated demanding charge neutrality for the device, the derivative of the potential at the domain boundary has to vanish. This is especially important in the contact regions, where the potential has to be constant in the direction of the leads. The peak located in the channel close to the source belongs to the barrier that forms in the intrinsic part of the channel. Due to the large bias the barrier is already lowered considerably close to the drain, which is the so called drain induced barrier lowering (DIBL). The almost rectangular shape of the potential in the cross-section of the gate indicates that the number of occupied modes is rather high, which leads to a density distribution that is closer to the classical limit. This is in contrast to the source and drain contacts, where the dominant part of the density is contributed by the first mode only. Due to the highly doped gates that pin the potential, the bias voltage drops almost entirely over a short distance between gates and drain.

The self-consistent electron density is plotted in fig. 3.15 and exhibits the characteristic peaks at the gate-oxide interface that are due to bound states forming at the depletion layer. The peak close to the drain is slightly higher due to the applied bias. Although the device is almost open from the point of view of the transfer characteristics, the density in the channel is still very small compared to the source and drain contacts. This is a consequence of the extremely high doping in the contacts and also supports the applied boundary model since the effect of the potential lowering in the drain will be small.

To analyze the effect of the gate on the potential and the density, a slice through the device along the center axis of the channel is presented showing the change of the potential barrier (see fig. 3.16) and the corresponding density (see fig. 3.18) for different gate voltages. It is clearly seen in fig. 3.16 that within the ballistic model the entire voltage drop occurs within the intrinsic region of the channel and the electric field in the highly doped parts of the source and drain contacts is zero. The double gate is working properly by raising the barrier height almost by the same amount as the gate voltage is changed. The potential in the contacts is

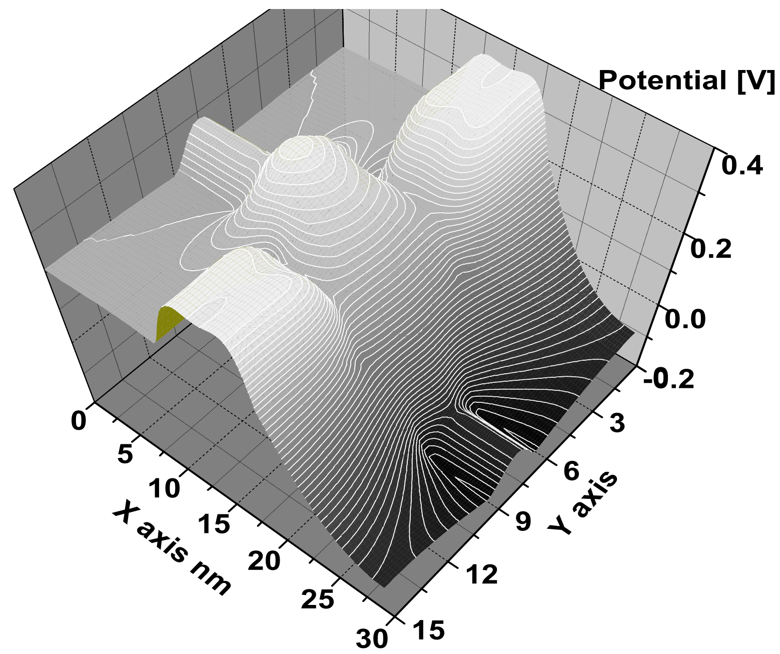


Figure 3.14: Self-consistent Hartree potential, calculated for an applied bias V_{DS} of 0.4 Volts and zero source-gate voltage.

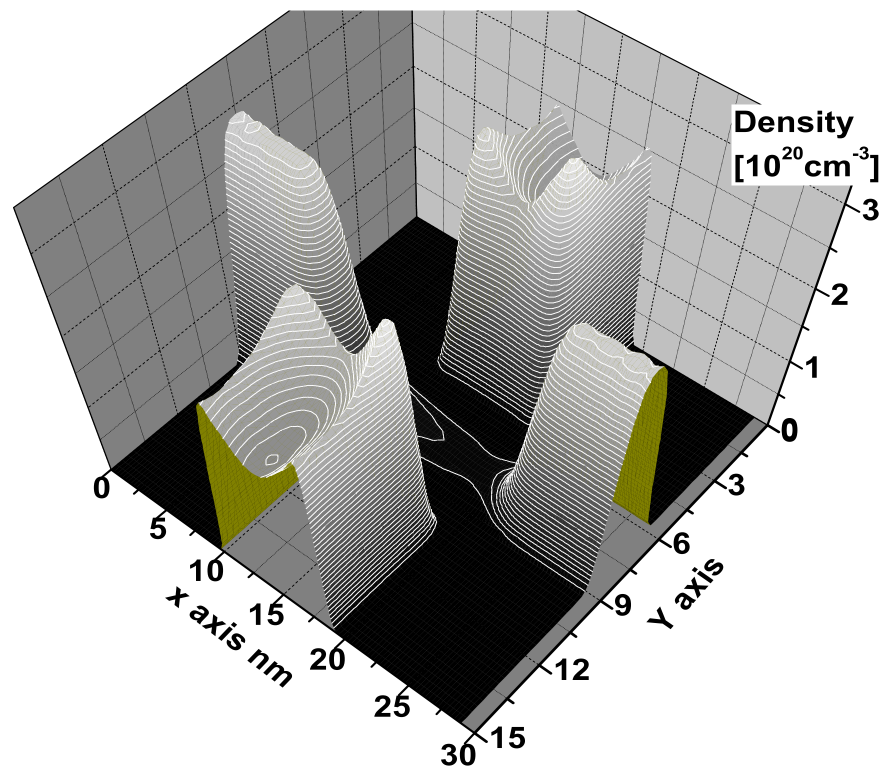


Figure 3.15: Self-consistent electron density, calculated for an applied bias V_{DS} of 0.4 Volt and zero gate-source voltage V_{GS}.

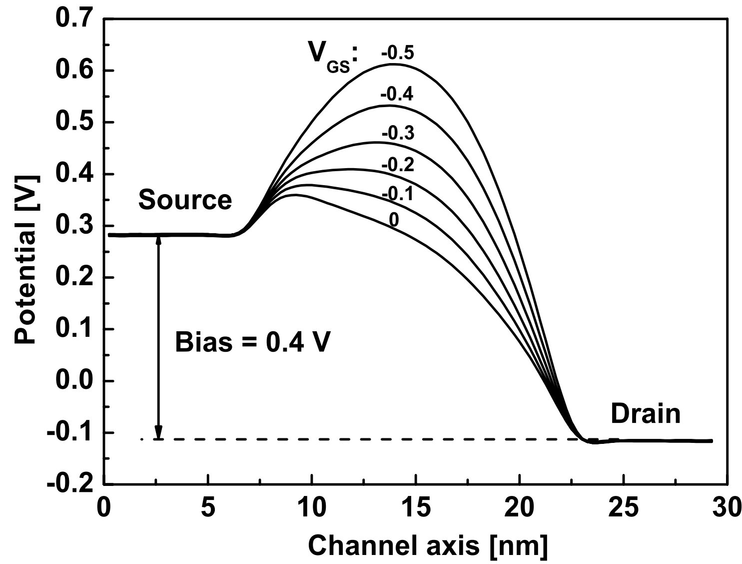


Figure 3.16: The self-consistent potential along the center axis of the channel is plotted for several source-gate voltages. The applied source-drain bias is 0.4 Volt.

almost constant for different gate voltages and thus the floating of the contacts due to the boundary conditions is very small. To analyze this important fact in more detail we plotted the difference between the applied bias, namely the difference in the local Fermi levels of source and drain and the effective bias, calculated from the difference between the electrostatic potential in source and drain. Figure 3.17 shows the reduction of the effective bias as a function of the source-gate voltage for applied bias V_{DS} of 0.4 V (full line) and 0.05 V (dashed line). For highly negative source-gate voltages the two contacts are well separated by a barrier in the channel and therefore no coupling is present that could affect the effective bias. If this barrier is lowered with increasing gate voltage V_{GS}, electrons originating from the source tunnel through the barrier and contribute to the density in the drain contact. To compensate for this additional charge, the potential in the drain is raised and simultaneously lowered in the source, resulting in a reduction of the effective bias. But as shown in fig. 3.17 even for zero gate voltage the relative bias change is always less then 5 %. Therefore we claim that for the presented device the applied von Neumann boundary conditions for the electrostatic potential are perfectly valid and the difference between the different boundary models is expected to be minor.

The density in the channel, as shown in fig. 3.18 is changed by more than a decade for a gate voltage change of 0.1 V in the off-state of the DGFET. Due to the more efficient penetration of the wavefunctions into the barrier from the source than from the drain, the density is slightly asymmetric with a minimum close to the drain contact. As the potential, the density is well converged in the source and drain contacts where it is flat.

3.7.2 IV-Characteristics

The main result of the self-consistent calculation is the IV-characteristics of the DGFET-device which will be presented in the following. The transfer characteristic of the DGFET is shown in

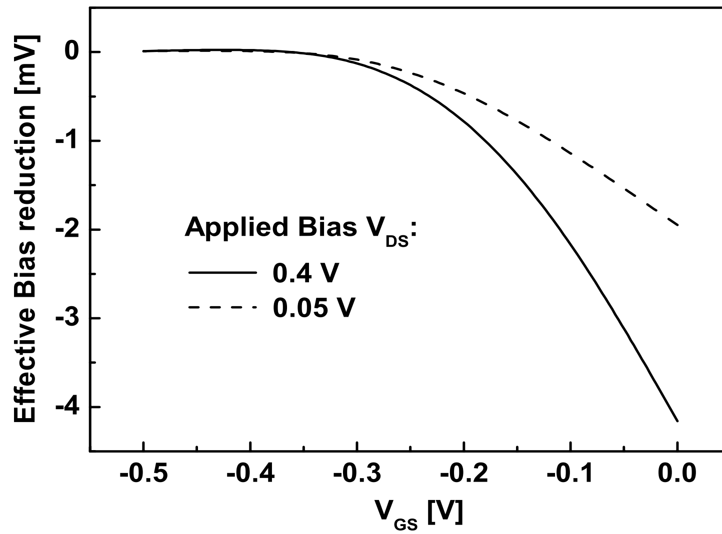


Figure 3.17: The difference between applied and effective bias is plotted as a function of the source-gate voltage V_{GS}.

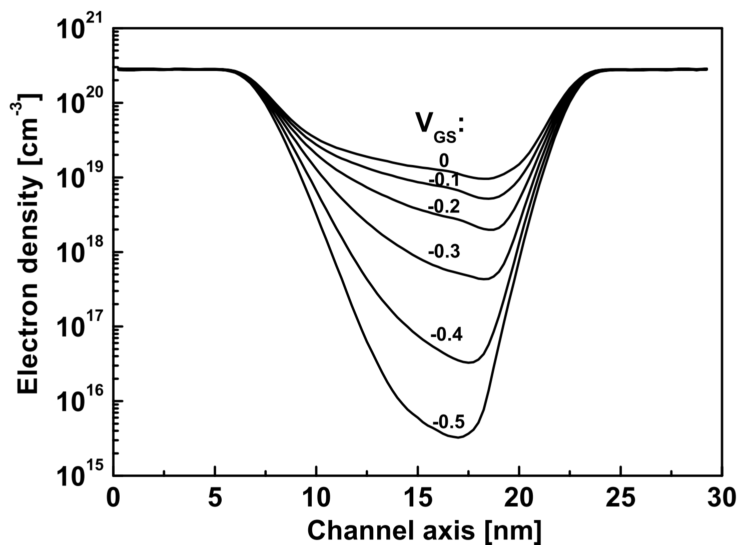


Figure 3.18: The electron density along the center axis of the channel is plotted for several source-gate voltages for an applied source-drain bias of 0.4 Volts.

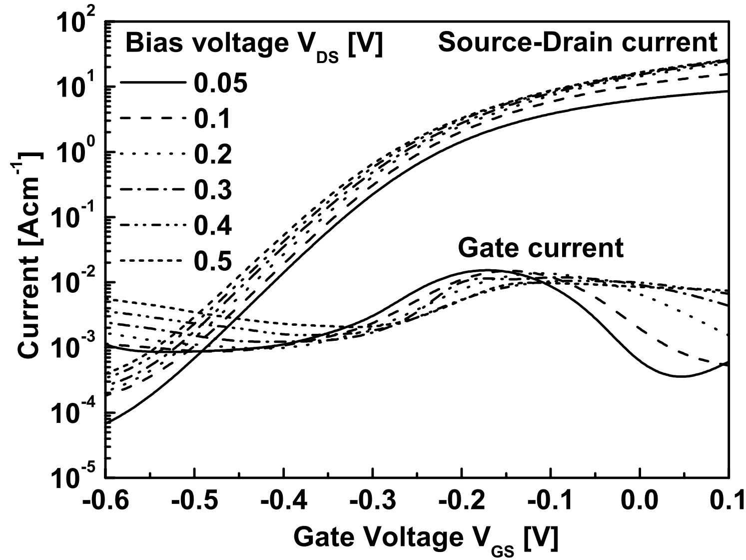


Figure 3.19: The transfer characteristics of the DGFET with the source-drain and gate currents are plotted as a function of the source-gate voltage for different applied biases.

fig. 3.19 with the source-drain as well as the gate leakage current plotted as a function of the source-gate voltage V_{GS} for different applied bias voltages. The source-drain current exhibits the characteristics expected of a transistor with an off-current that is exponentially rising with the gate voltage at a rate given by the sub-threshold slope of $S \sim 80$ [mV/dec] which is well in the expected range. The shift of the IV curve for different bias voltages is explained by the drain induced barrier lowering (DIBL) which is evaluated to 110 [mV_{GS}/V_{DS}]. The current in the on-state of the device is rather high compared to calculations that include scattering [54] and can be regarded as the upper ballistic limit. The gate-leakage current that is already summed over both gates as well as source and drain, shows a non-monotonic dependence on the gate voltage which is explained by quantum resonances in the narrow channel. We observe a crossing of the source-drain and the gate leakage curves for a gate voltage between -0.5 and -0.4 Volts which is well in the off-state of the device. The gate-leakage calculated using the single band model can only give a qualitative indication of the magnitude rather than a quantitative prediction. But a more reliable way to calculate the leakage is still out of reach since it should include a multi-band Hamiltonian that allows for inter-band tunneling as well as phonon-scattering terms, which is presently impossible to calculate for a rather large 2D structure as in the presented example.

The output characteristic of the device is shown in fig. 3.20 where the source-drain current I_{SD} is plotted as a function of the applied bias V_{DS} for different gate voltages V_{GS}. Due to the absence of scattering in the ballistic model, the electron momentum is not bound and thus the on-current does not saturate but rises continuously with increasing bias. In order to validate our results we compare the calculated ballistic current of this work with the results of a non-equilibrium Green's function calculation (NGEF) [54] that includes elastic as well as inelastic scattering. Figure 3.21 shows the source drain currents for two different bias voltages of 0.05 and 0.4 Volts. To account for the workfunction of the gates, the gate voltage of the ballistic

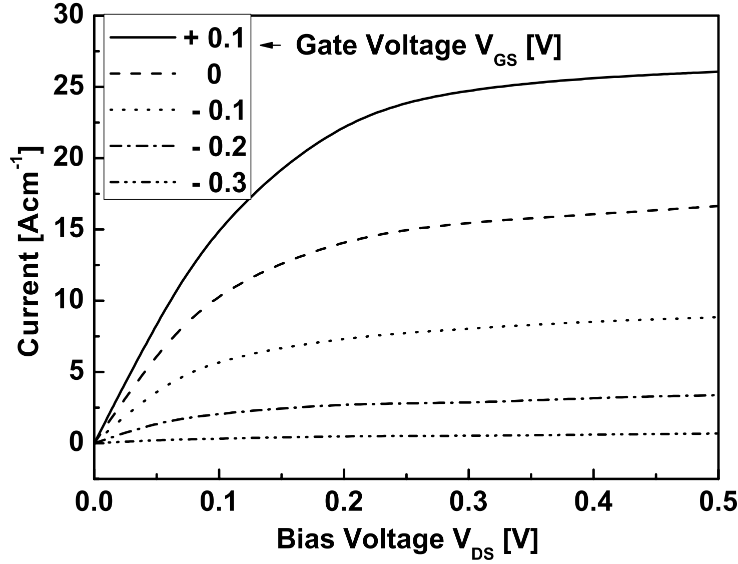


Figure 3.20: The output characteristic of the device with the source-drain current plotted as a function of applied bias for different source-gate voltages.

results was shifted by 0.4 V such that the resulting currents of both models are identical at zero source-gate voltage. Due to the steeper sub-threshold slope of the ballistic current of 76 [mV/dec], compared to 90 [mV/dec] in the scattering model, the current rises faster with increasing source-gate voltage which leads to an on-current that is higher than the current resulting from the scattering model by a factor between two and three depending on the bias voltage.

This overestimation of the current by the ballistic model is in good agreement with findings of Venugopal *et al.* [64], who compared the results of a calculations using a purely ballistic model and one that includes the effects of inelastic scattering for a comparable DGFET-device. Thus the ballistic model is very well suited to give an upper limit of the expected current and reflects the device characteristics in a qualitative way. In contrast to the calculations that include scattering, no additional simplifications are needed, such that all contacts can be included in the calculation and an estimation of the gate leakage current is obtained.

3.7.3 Bound states

In the theory part the importance of the proper treatment of the bound states was discussed and a scheme was presented to incorporate these states into the density to obtain a physically meaningful and well converged carrier density. In the following we will present an evaluation of the contribution of quasi bound states to the local carrier density in the gates. We are going to show that the bound states significantly alter the local distribution of the charge density and the potential in the vicinity of the gate compared to purely ballistic calculations which drastically affects the IV-characteristics of the device.

The self-consistent electron density in the gate region including the continuous and discrete

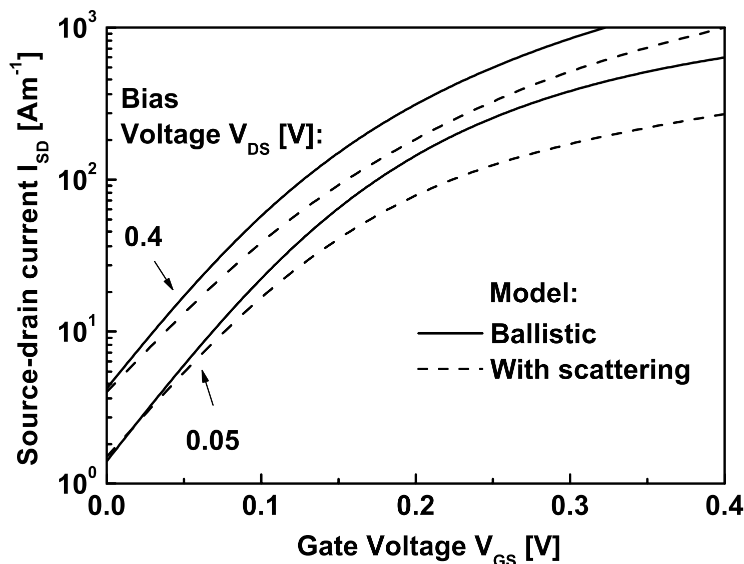


Figure 3.21: Comparing the source-drain current resulting from the ballistic model employed in this work (solid lines) with the results of a calculation including scattering (dashed lines) for two different bias voltages.

part of the density is shown in fig. 3.22 for a bias of 0.05 V and zero source-gate voltage.

The local density distribution is characterized by the depletion regions close to the oxide barriers that are located at the boundaries of the depicted gate region. The electrons can hardly penetrate the oxide barrier with a conduction band offset of 3 eV and therefore the quantum mechanical probability density vanishes in the barriers. This leads to an almost one nanometer wide depletion region at the oxide interfaces where the net charge is positive since the doping ions are assumed to be homogeneously distributed over the entire gate region. Certainly this is an approximation, since in reality the ions are randomly distributed but this cannot be accounted for in a two dimensional calculation where the structure is necessarily assumed to be homogeneous in the remaining z-direction. The pronounced peaks in the corners of the gate region in the vicinity of the oxide barrier ($x = 5$) are due to the bound states that are confined in the triangular well in front of the barrier. The discrete part of the density resulting from the quasi bound states is shown in figure 3.23. Its most prominent features are the pronounced peaks (indicated by the light regions) that account for more than one third of the total density in this region as shown in figure 3.22. It is also seen that the density is not completely zero at the contact ($x = 0$) which is explained by the small binding energy of only a few meV. The part of the bound state wavefunction that decays into the leads is calculated to properly normalize the part of the state localized in the gate region. In this example the quasi bound states are almost completely (> 90 %) localized inside the device domain. To show the importance of the inclusion of the quasi bound states into the electron density we perform a self-consistent calculation without taking into account the bound states in the density explicitly. As shown in fig. 3.24, the maxima of the electron density are shifted away from the oxide barrier by almost 1 nm which correlates with the wider depletion region. Thus the charge distribution close to the channel is changed dramatically which has a strong influence on the functionality of the

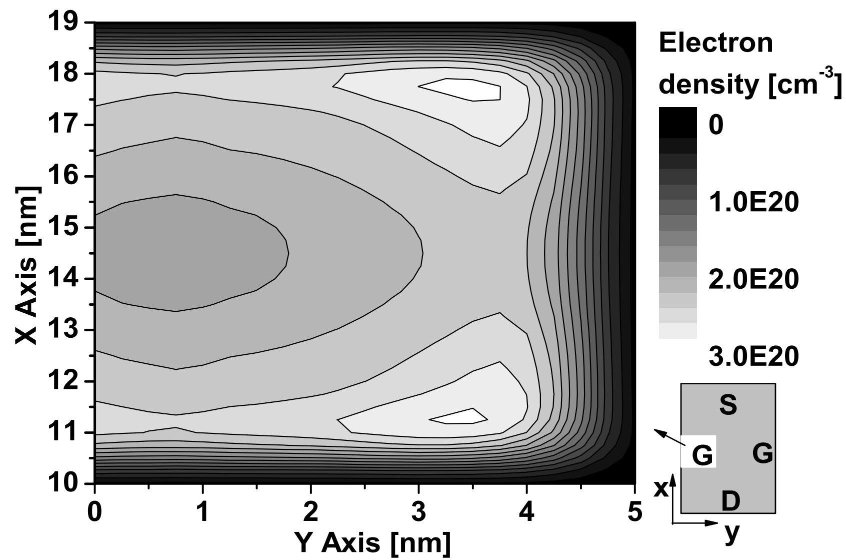


Figure 3.22: Local distribution of the electron density in the gate region including continuous part and bound states.

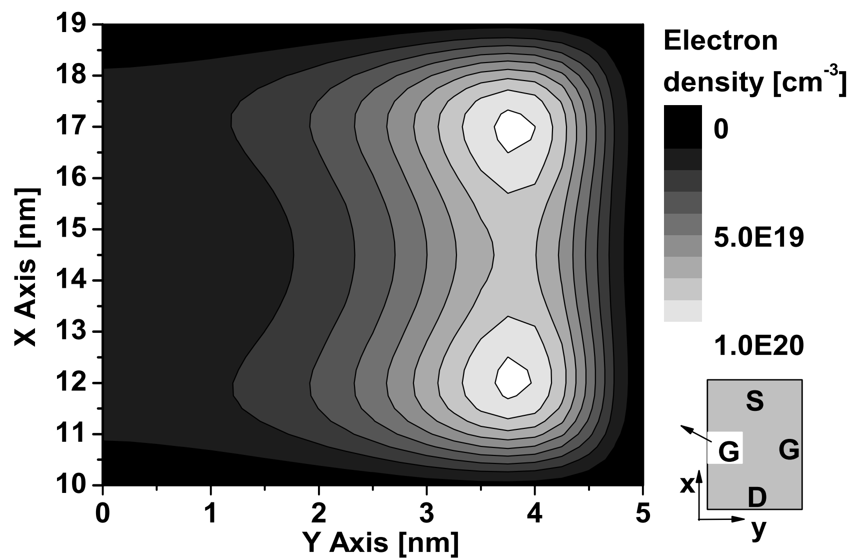


Figure 3.23: Discrete fraction of electron density resulting from the bound states localized close to the barrier.

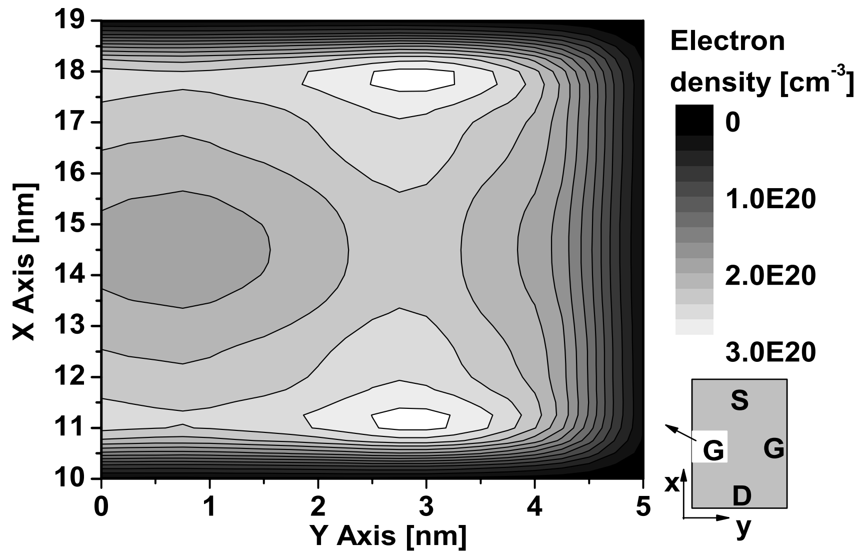


Figure 3.24: Electron density distribution in the gate region of the DGFET calculated neglecting the bound states.

gates.

The binding energy of the bound state in the gate of the DGFET is plotted in fig. 3.25 as a function of the applied gate voltage V_{GS}. It is observed that the binding energy is almost perfectly linear with the applied voltage with an average energy of more than 10 meV. This shows that the state is not an artefact of the self-consistent calculation but is physically existent inside of the DGFET. Comparing the density distribution for a slice along the x-axis through the middle of the gates as shown in fig. 3.26 the redistribution of the electron density is easily seen. Additionally there is an effect on the electron density inside the channel leading to a higher density for the calculation neglecting the bound states. This change in the channel density can be explained by the different electrostatic potentials shown in fig. 3.27 that result from the self-consistent calculation including (solid line) or neglecting (dashed line) the bound states. Since the total density of states in the gate regions is smaller for the latter case, the potential has to be significantly lower to result in the same electron density demanded by charge neutrality. The effective gate potential just at the contact to the oxide barrier that ultimately controls the functionality of the DGFET device is then changed by 40 meV. This shift in the average potential also affects the potential in the channel which causes the higher channel density observed in figure 3.26.

To evaluate the impact of the different models for the calculation of the density on the predicted device behavior we compare a measurable quantity such as the source-drain current shown in figure 3.28. Due to the higher density in the channel, the overall current is higher for the calculation neglecting the bound states (dashed line) and the transfer characteristics seems to be shifted to lower gate voltages. This leads to the conclusion that the proper inclusion of the bound states is eminent in order to obtain a realistic prediction of the device characteristics.

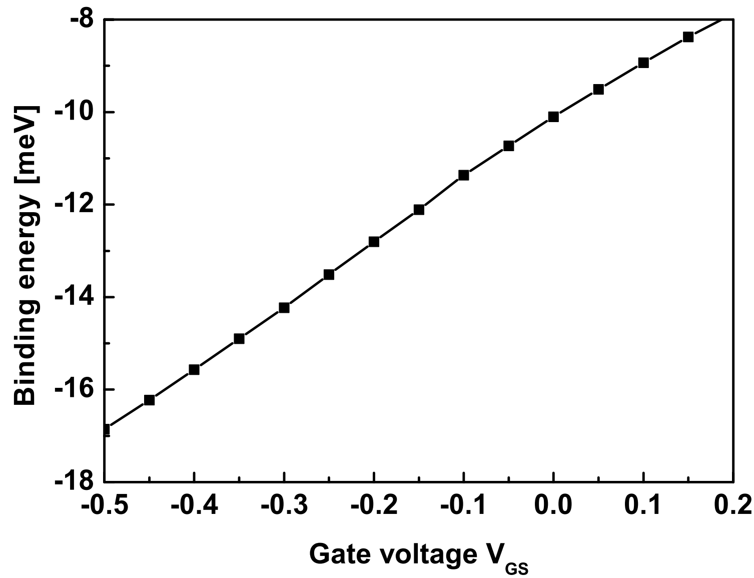


Figure 3.25: Bound state binding energy defined as the difference between the energy of the lowest propagting mode in the gate and the eigenenergy of the state.

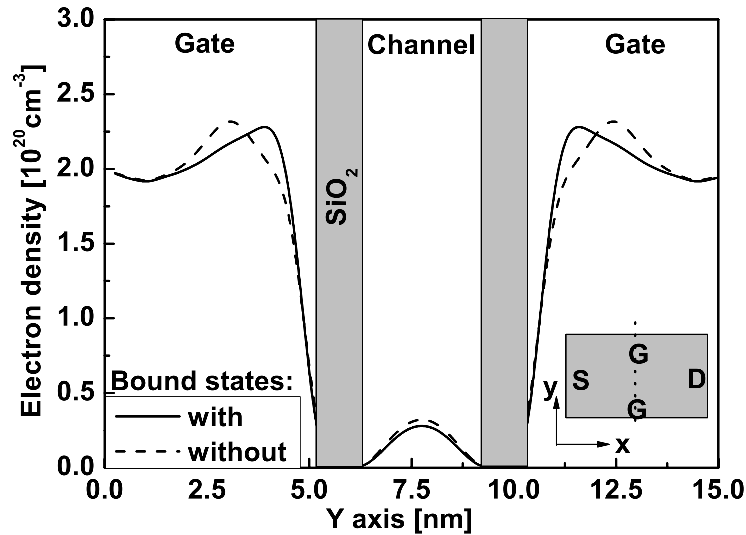


Figure 3.26: Electron density along a slice through the center of the gates. The position of the oxide barriers is indicated by the gray regions. The density is compared for self-consistent calculations including (solid line) or neglecting (dashed line) the bound states.

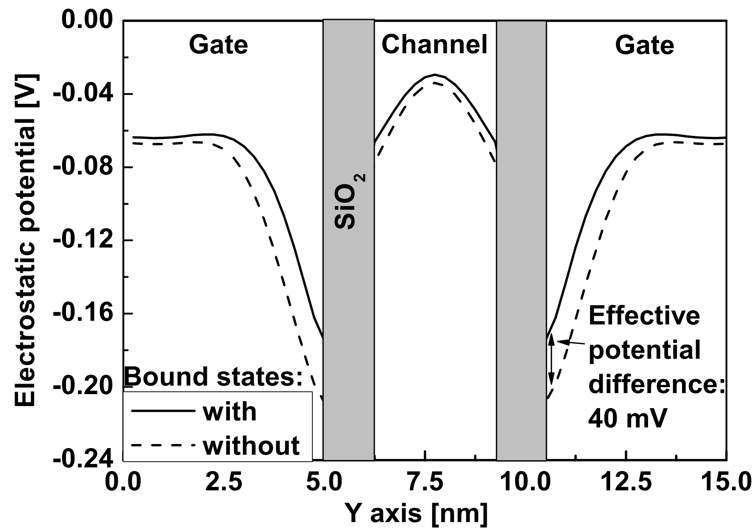


Figure 3.27: Electrostatic potential along a slice through the center of the gate regions with the oxide barriers indicated by the gray regions. The potential is compared for self-consistent calculations including (solid line) or neglecting (dashed line) the bound states.

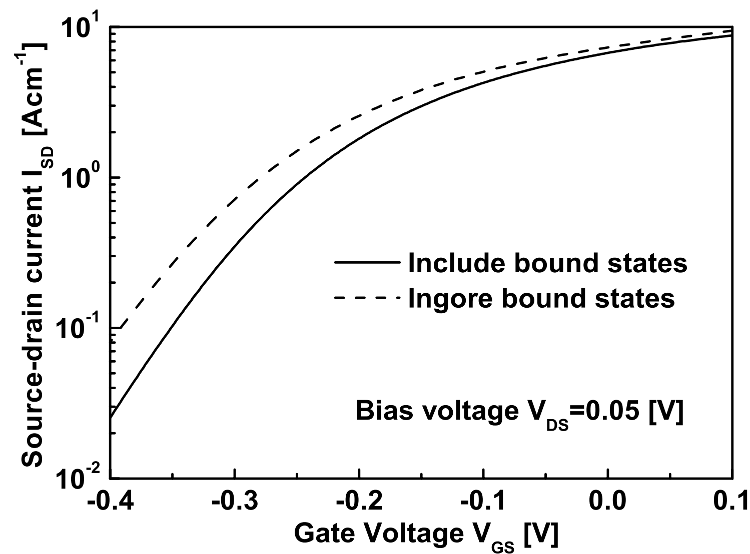


Figure 3.28: The source-drain current calculated for 0.05 V applied bias. The current resulting from the calculation including the bound states (solid line) is significantly lower than for the calculation ignoring the bound states (dashed line).

3.7.4 Eigenstate Convergence

The main advantage of the CBR method is the possibility to reduce the set of eigenstates included in the representation of G_0 to a narrow energy interval mainly determined by the applied bias and temperature of the device. This leads to a tremendous reduction in the computational effort because even in 2D calculations usually less then 10 % of all eigenvalues are needed. Since the result is not exact anymore compared to the case when all eigenvectors are included, we will now assess the error introduced with the truncation and demonstrate that it is negligible for the examples considered in this work. In the following we will analyze the convergence behavior of the self-consistent calculation as well as the resulting current for different numbers of eigenstates included in the representation of the Green's function of the decoupled device. Certainly if the energy of the highest eigenstate included is very close to the Fermi energy such that for the given temperature the Fermi distribution is not small then the total density will not be correct. But even if the highest state is far enough from the Fermi level such that it is not occupied, the local distribution of the density may still be affected since the retarded Green's function G^R coherently mixes all included states. In the first place we will look at the convergence behavior for different percentages of eigenstates included as shown in figure 3.29. Interestingly, with the exception of two cases, the convergence seems to be rather unaffected by the number of eigenstates included. The first exception is the case where only 0.7 % of all states are included which corresponds to a cutoff energy of only 250 meV above the Fermi level. For temperatures of 300 K this is clearly not enough to yield a converged density and therefore the convergence is very slow over the entire self-consistent cycle. The other exception is not yet explained since it corresponds to the second largest set of eigenstates and it is not obvious to us why the convergence stops at a residuum twice as high as for the remaining cases. But in general we conclude that the convergence behavior is very robust and almost independent of the number of included states as long as the cutoff is far enough above the Fermi level depending on the given temperature. The values of the converged source-drain current are shown in fig. 3.30 as a function of the cutoff energy relative to the highest Fermi level in the calculation. Here it is seen that a reasonable convergence is reached for a cutoff higher than 1 eV above the Fermi edge which corresponds to more than 3.6 % of all states in the example of the DGFET device. This result shows that the method is reliable and the efficiency due to the truncation is preserved even in the self-consistent calculation.

One problem that remains unsolved in this work are oscillations in the local density close to the contacts which are a relict of the truncation of the set of eigenstates used in the calculation. This phenomenon is illustrated in fig. 3.31 where the density deviation relative to the converged value is shown for a slice along the propagation axis of a narrow quantum wire with leads attached to the left and right end. In the interior of the wire roughly 20 nm apart from the contacts, the density is absolutely converged for all three lines that have been calculated using different percentages of eigenstates included. Closer to the contacts the density starts to oscillate with a maximum deviation of 2.5 % for the dotted line showing the result with only one percent of all eigenstates included. This error is reduced by more than a factor of ten when using 15 % of all states as shown by the solid line. Even though the deviation looks quite large due to the magnification of the graph it is still in the range of a few percent and therefore not a big problem. Nevertheless, one could try to artificially smoothen the density in the contact regions since the deviation is well localized in a narrow region close to the contacts. In summary, we conclude that the truncation in the set of eigenstates included in the calculation of the ballistic density works very well since the error is already very small for a few percent of

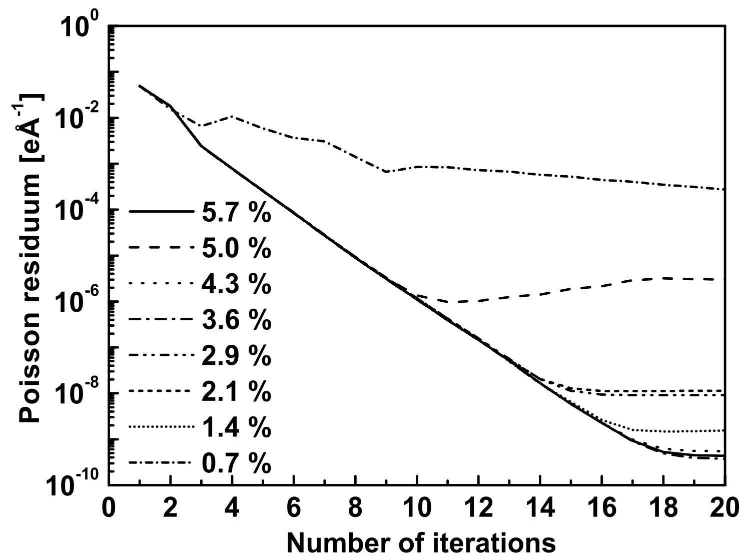


Figure 3.29: Convergence behavior of the selfconsistent calculation for different percentages of eigenstates included.

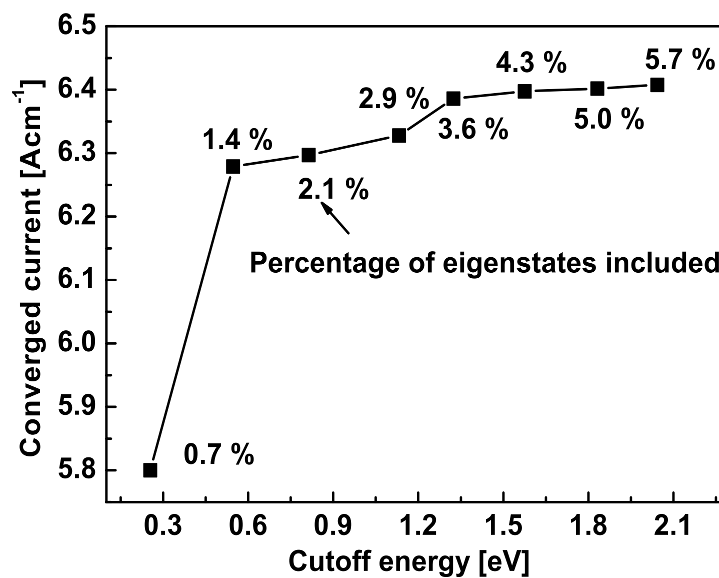


Figure 3.30: Source-drain current as a function of the cutoff energy relative to the highest Fermi level. The numbers indicate the percentage of all eigenstates included in the representation of the Green's function of the decoupled device G^0. The calculation has been performed with an applied bias of 0.05 Volts and zero gate voltage.

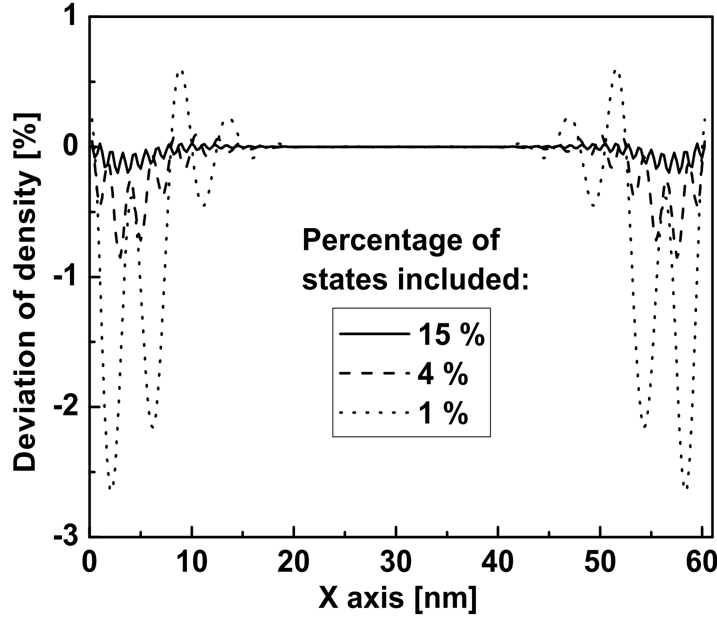


Figure 3.31: The deviation of the density from the converged value is shown for a slice along a narrow quantum well for different percentages of eigenstates included in the calculation.

all states included. Importantly, the method allows checking the convergence of the results by simply adding more states to the set which systematically reduces the error as shown in this section.

3.7.5 Computational costs

The efficiency of the self-consistent CBR method is estimated easiest for the example of the DGFET depicted in figure 3.1. The structure is discretized on a finite differences grid with 2.5 Å grid spacing, resulting in a total of 7500 grid points. Depending on the valley, up to 400 eigenstates of the decoupled Hamiltonian are taken into account, which guarantees a well converged result as shown in figure 3.30. For the desired bias range of 1 Volt only 30 propagating lead modes out of a total of 200 are needed for the computation. The energy grid is resolved on 1000 grid points including the extra points used within the adaptive energy grid to resolve the onset of the modes. Figure 3.32 shows the convergence of one single bias point with the residuum of the Poisson equation defined as the magnitude of the residuum vector resulting from equation (3.17). The convergence of the non-linear Poisson equation is monotonic until a minimum is reached below 10^{-9} eÅ^{-1}. Most importantly the convergence of the current expressed in the relative error of the current at iteration i compared to the converged current, is almost parallel to the convergence of the Poisson equation leading to a result that is correct up to the 5th digit after 10 iterations. Since the changes in the current induced by small fluctuations in the structural parameters such as the gate oxide thickness (see fig. 3.37) or the channel length are very large, a convergence of the current to a relative error smaller than 1 % is completely sufficient to describe the device characteristics within this model. This leads to

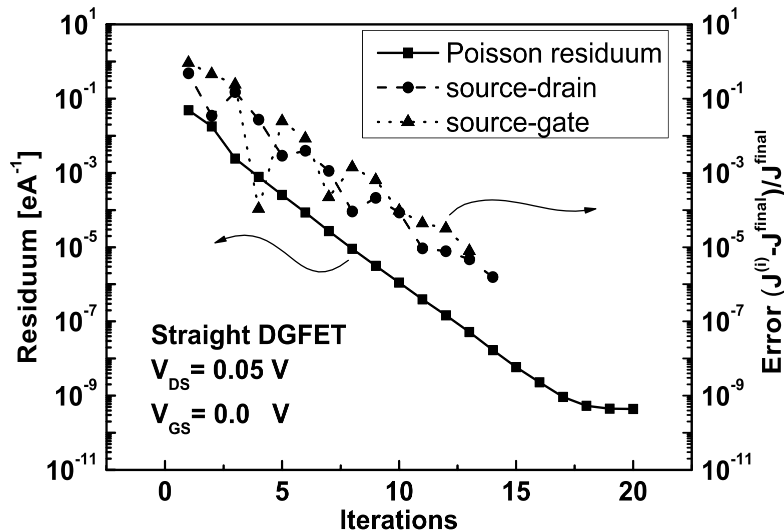


Figure 3.32: The convergence of the Poisson equation (solid squares) at a single bias point reaches a lower limit at 10^{-10} [eÅ^{-1}]. The convergence of the source-drain (solid circles) and source-gate (solid triangles) current follows roughly the trend of the Poisson residuum, leading to a current that is correct up to the 5th digit after 10 iterations.

a rather rapid convergence of only four to five steps per bias point as shown in fig. 3.33, after the potential has been calculated once for an initial bias step. The computational costs of one single step, that involves the solution of three Schrödinger equations for the different valleys, the calculation of the local density of states using the CBR method, as well a the Newton cycle to calculate a new correction for the electrostatic potential is only between 20 and 30 minutes on a ordinary PC with a Pentium IV processor with 3 GHz cpu speed. This efficiency has been achieved using 'off the shelf' eigensolver (ARPACK [66]) and BLAS routines from the Intel mathkernel library. This leads to a computation time between one and two hours per bias point, which is a bit faster than the times of a few days per bias point reported in the work of Laux [19]. Thus, we conclude that the self-consistent algorithm of the CBR method, developed within this thesis, presents an efficient and state of the art tool for the computation of the ballistic transport properties of 2D quantum nanostructures.

3.7.6 Exchange and Correlation potentials

Due to the high doping of the simulated DGFET structure and the resulting high carrier density, electron-electron interactions like exchange and correlation are not negligible anymore and have a strong influence on the device characteristics. Within density functional theory (DFT) these effects can be calculated via introducing effective potentials accounting for the exchange and the correlation within the local density approximation (LDA). These potentials depend on the density locally and can therefore be evaluated very easily. A more detailed description of the DFT and the parametrized effective potentials is given in the quantum dot section where this formalism is used to calculate charged exciton states. In this part we will analyze the differences in the total effective potential between the pure Hartree potential used in the previous sections

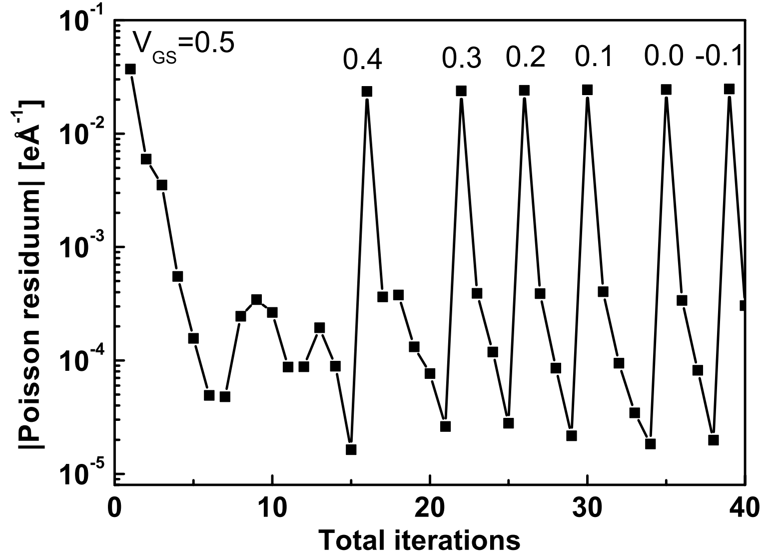


Figure 3.33: Residuum of Poisson equation for a gate voltage sweep with fixed bias voltage of $V_{DS} =$ 0.4. Each point on the graph corresponds to a solution of the Schrödinger equation of the three valleys including the calculation of the local density of states via the CBR method.

and the one including exchange and correlation (XC). Figure 3.34 shows the effective potential for the Hartree (solid line) and the LDA (dashed line) calculation along the channel axis for an applied bias of 0.05 V and zero gate voltage. It is clearly seen that the effective barrier height with XC is larger by roughly 10 meV compared to the pure Hartree calculation. This is explained by the fact that the exchange and correlation potential is effectively attractive with an average value of -70 meV in the highly doped regions. Since the XC-potential is non-linearly proportional to the density, its magnitude is much lower in the intrinsic region of the channel which leads to an effective increase in the barrier height.

The shape of the effective potential across the channel as plotted in fig. 3.35 is changed dramatically now resembling more a classical solution rather then the typical quantum well potential with the maximum in the center of the channel. Besides the different cross-sectional shape of the average potential in the channel is also higher than in the pure Hartree calculation which leads to a lower density for the same applied gate voltage. This effect is also visible in the transfer characteristics of the device shown in fig. 3.36, where the source-drain and the gate current is plotted as a function of the gate voltage for the Hartree potential (solid line) and for the XC calculation (dashed line). The source-drain current of the XC seems to be shifted about 50 mV on the gate voltage axis but has qualitatively the same shape as the Hartree curve. The more important effect is present in the gate leakage current which, in the off-region of the device below -0.2 V, is roughly one order of magnitude higher than in the Hartree calculation. This effect has strong implications on the quality of the transistor since it increases the residual current in the off-state of the device. The origin of this increased gate leakage is the different density distribution in the channel which allows the wavefunctions to get closer to the oxide barriers. The results of this section show that for these highly doped structures the exchange-correlation potential cannot be neglected anymore, since it significantly influences the device

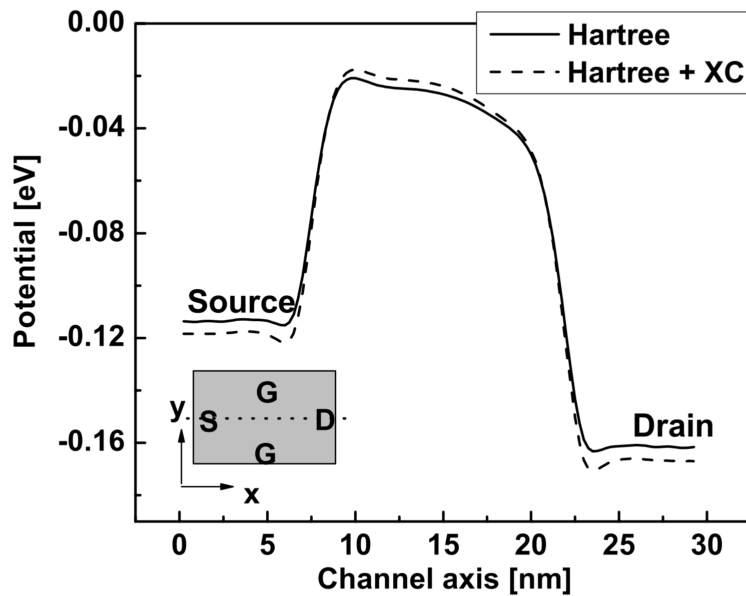


Figure 3.34: Effective potential along the channel axis for an applied bias of 0.05 V and zero gate voltage. The solid line shows the pure Hartree potential whereas the dashed line shows the sum of Hartree, exchange and correlation potential resulting from a converged self-consistent solution.

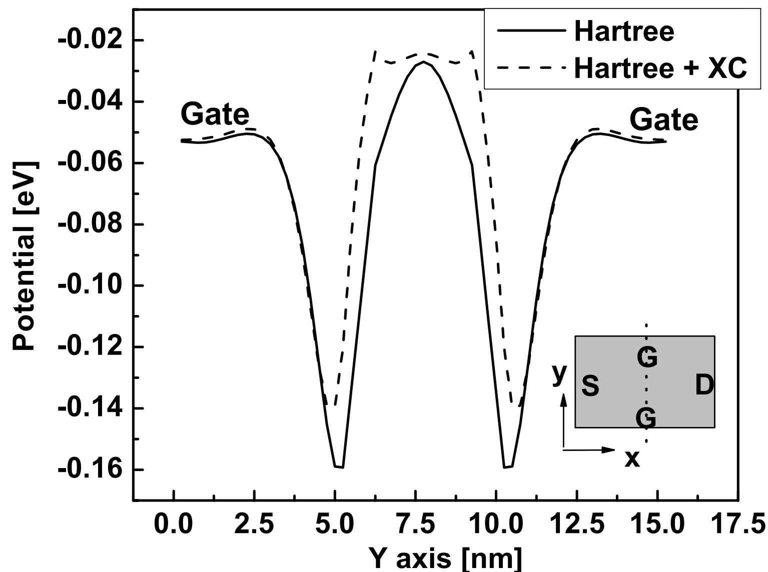


Figure 3.35: Cross-sectional slice of the effective potential for an applied bias of 0.05 V and zero gate voltage comparing the pure Hartree (solid line) and Hartree plus exchange and correlation potential (dashed line).

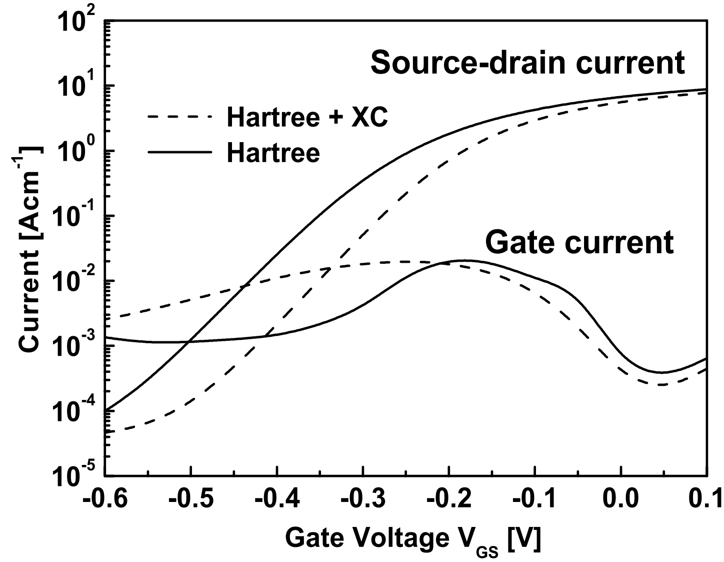


Figure 3.36: Transfer characteristics of the DGFET device comparing the results of the pure Hartree calculation (solid line) with the results from the exchange-correlation calculation (dashed line).

characteristics.

3.7.7 Varying gate oxide thickness

An important parameter that is never controlled with absolute precision in the production process of the device, is the gate oxide thickness. But since this parameter has a tremendous influence on the device behavior, we will analyze in the following, how the device characteristics is changed when the gate oxide thickness is varied. Figure 3.37 shows the change in the transfer characteristics for different gate oxide thicknesses. First looking at the source-drain current, it seems that it is not affected much. The sub-threshold slope changes from 75 mV/dec for the 12.5 Å oxide to 70 mV/dec for the thin oxide of 7.5 Å which is explained by higher electric fields perpendicular to the channel for the same applied voltage. This better gate control of the source-drain current also leads to a slightly higher on-current for the thin gate oxide. The price to be paid for this small improvement however is a gate leakage current that is orders of magnitude higher than for the thick oxide. Effectively the DGFET with the 7.5 Å barrier cannot be switched off anymore since the leakage current is almost constant at 0.1 A/cm in the off-state of the transistor. Certainly the values for the gate current are not highly accurate due to the simple effective mass model employed, but they give a qualitative idea of the strong dependence of the gate leakage on the oxide thickness.

3.7.8 Local density of states

Further insight into the physics of the DGFET in the ballistic regime is obtained by regarding the local density of states as plotted in figure 3.38. In this graph the local DOS along the center

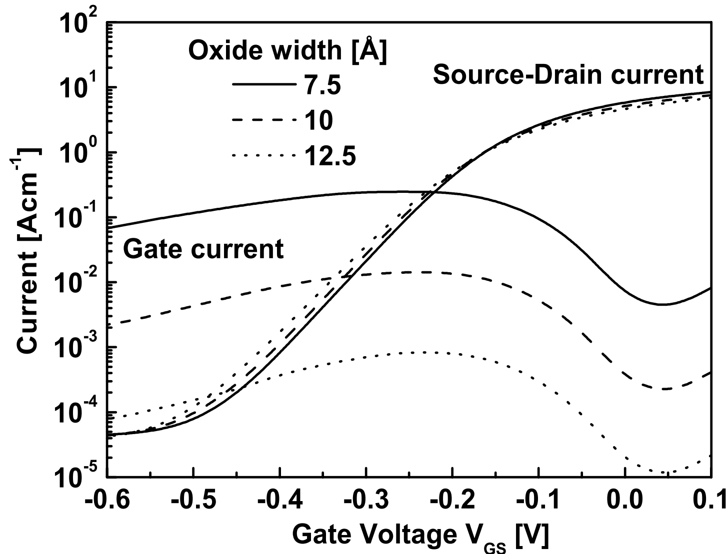


Figure 3.37: The source-drain and gate current of the DGFET device calculated for different gate oxide thicknesses of 7.5 Å (solid line), 10 Å (dashed line) and 12.5 Å (dotted line) is shown as a function of the gate voltage for an applied bias of 0.05 V. The calculation was performed taking into account the exchange and correlation potential in LDA.

axis of the channel is plotted as a function of the energy where the light areas indicate a high DOS. The conduction band edge with an applied bias of 0.3 V is indicated by the white line to guide the eye. Only a few meV above the conduction band edge in the source as well as in the drain contact, the first propagating mode appears, but with rather low density of states, which is due to the low mass in the direction of propagation. The main contribution to the DOS, as indicated by the bright regions in fig. 3.38, is due to the remaining valleys with the low mass in propagation direction that start roughly 100 meV above the first mode. However, due to the cutoff imposed by the Fermi distribution, the main contribution to the density is still resulting from the energetically lowest mode. The most interesting features are certainly the strong oscillations of the DOS in the source and dominantly in the drain contact. These are due to reflections at the barrier that are not damped by scattering within the ballistic model. In the next section we will be able to compare the ballistic DOS with the one obtained using a simple scattering model.

3.8 Cone-DGFET

The results presented in this chapter so far, have been obtained for the DFGET depicted in fig. 3.1 that has a straight Si-channel with source and drain contacts of the same shape. In realistic devices however, one could think of different geometries for the source and drain contacts that connect to leads that are wider than the channel. Such a device with cone shaped source and drain regions has been proposed and modeled by Laux *et al.* [19]. A sketch of the structure is depicted in fig. 3.39 with source and drain regions that widen from 3 nm in the

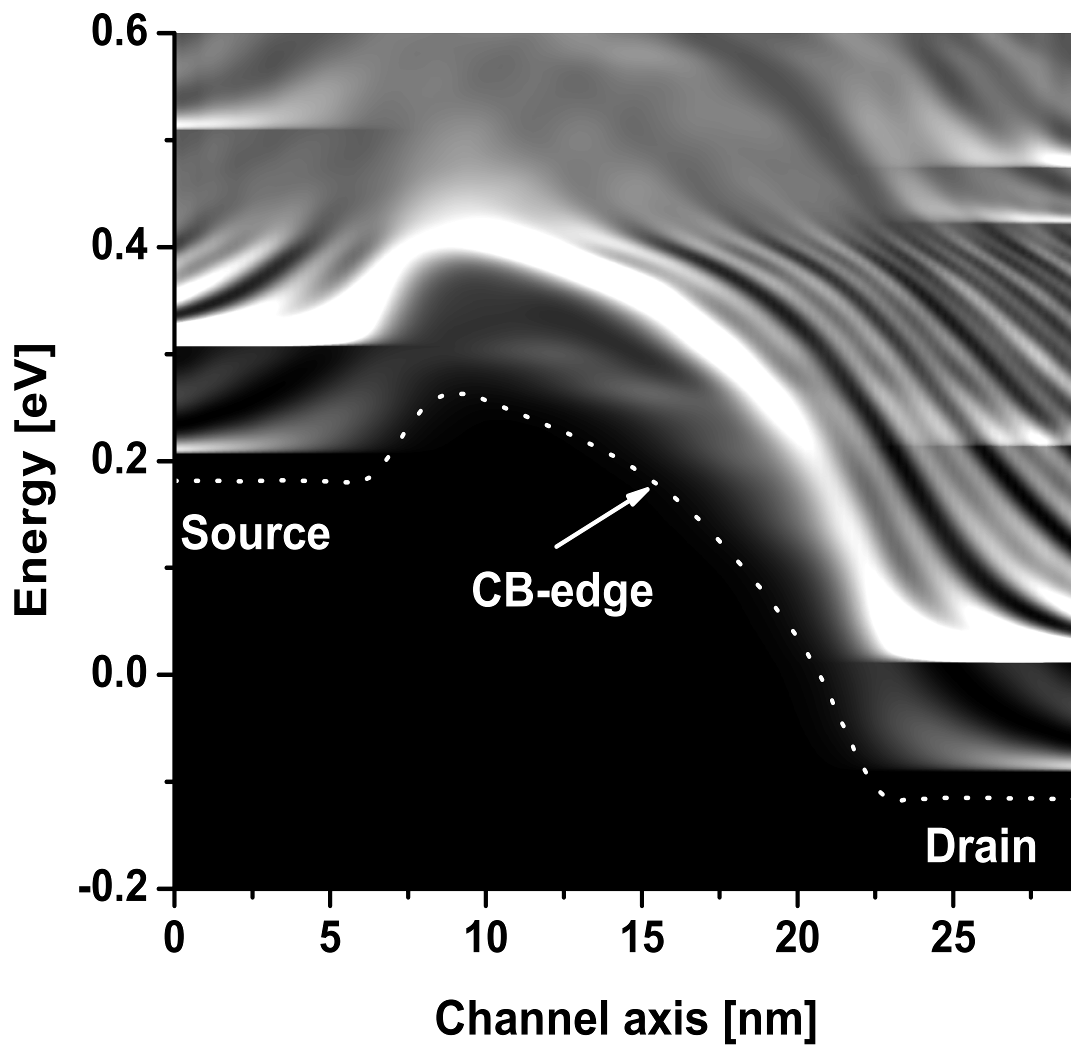


Figure 3.38: Local density of states plotted along the center axis of the channel with the light regions corresponding to high DOS.

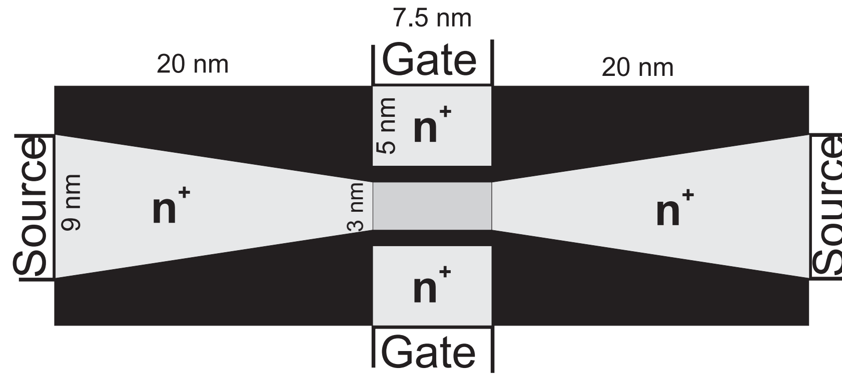


Figure 3.39: Double gate FET with cone shaped source and drain regions. The contacts are highly doped with 10^{20} cm^{-3} (light gray), while the channel is doped with 10^{16} cm^{-3} (dark gray). The oxide, with a gate-oxide thickness of 1 nm, is indicated in black.

channel to 9 nm at the contacts. The gate length is 7.5 nm, which is slightly shorter than the 9 nm of the straight DGFET, but both devices have the same gate oxide thickness of 1 nm. The contact regions are highly doped with 1×10^{20} cm^{-3} (light gray regions in fig. 3.39) and the channel is doped with 1×10^{16} cm^{-3} as indicated by the dark gray regions.Due to the geometry of the source and drain contacts in the cone-DGFET, the incoming states from the leads have a much higher probability to get reflected than in the straight DGFET, which can be observed in the density of states connected to the drain contact, depicted in figure 3.40. Both curves have been calculated for the same applied voltages and the same valley with the longitudinal mass in the direction transverse to the channel axis. The onset of the DOS is lower for the straight DGFET, since the doping is higher by a factor of two. The main interest however lies in the qualitative shape of the DOS for the two different devices. While the DOS for the straight DGFET exhibits few well developed peaks that correspond to the onset of the propagating modes in the leads, the DOS of the cone-DGFET is disturbed by strong oscillations caused by the reflections that cannot be attributed to any mode energies. This spiky DOS generally poses a problem for self-consistent calculations, where the density has to be integrated without introducing a large error, but, due to the efficiency of the CBR method, it is possible to resolve these peaks with a sufficient amount of energy points, such that convergence is achieved. In contrast to the significant qualitative difference in the DOS of both geometries, the transmission between source and drain depicted in fig. 3.41 is almost identical with the exception of minor wiggles in the case of the cone-DGFET. This finding suggests that the transmission through the device is mainly dependent on the width of the channel and that the geometry of the leads has only minor influence on the resulting IV characteristics.Finally we compare the IV characteristics calculated within our model with the results presented in [19]. Beforehand, we would like to point out the differences in the models employed, as there are the discretization, the boundary conditions and the inclusion of bound states and exchange-correlation potential. The rectangular grid used for the finite differences discretization in this work is rather inappropriate for shapes of arbitrary angle such as the cone-contacts. But regarding the fact that the real structure is build out of atoms in a cubic lattice with a lattice constant close to the one used in the discretization it seems to be still more realistic than the finite element grid used in [19]. A second important difference is also in the way the oxide is treated. While in this work the boundaries of the Silicon regions are all assumed to consist of

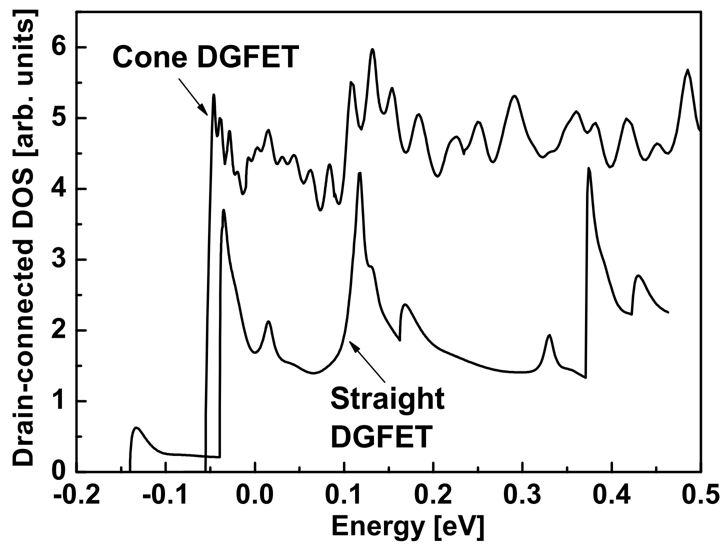


Figure 3.40: Drain connected density of states of the straight and cone-DGFET calculated for the same valley and applied voltages.

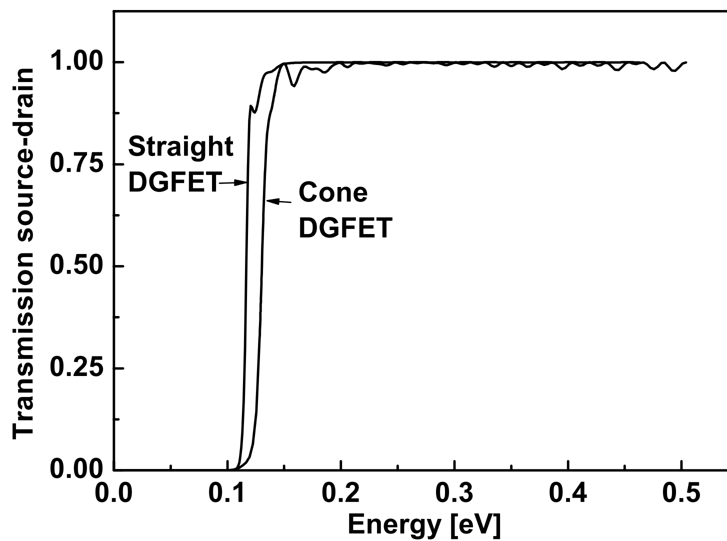


Figure 3.41: Comparing the source-drain transmission between the straight and cone-DGFET calculated for the same valley and applied voltages.

Silicon-oxide with a finite barrier height of 3 eV, the boundaries in the work of Laux are infinite barriers with the exception of the gate oxide. The comparison of the current resulting from our model with the use of the Hartree potential only (dotted line), Hartree including the bound states (dashed line) and including also exchange and correlation potential within LDA (solid line) with the data reported in the work of Laux (dash-dotted line) is presented in figure 3.42. The gate current (plotted as absolute value in the log plot) actually changes sign at a gate voltage V_{GS} of 0.2 V where the direction is inverted from current flowing into the gates for $V_{GS} > 0.2$ V to a net flow out of the gates into the device for $V_{GS} < 0.2$ V. Generally the gate current of our model is about one order of magnitude larger than Laux's which is most likely due to a difference in the effective barrier thickness within the two different discretization schemes. However, the qualitative behavior of the gate current calculated neglecting the bound states is almost identical to the result of [19]. The source drain-current in the off-state of the device is is comparable for both models neglecting the exchange-correlation potential, but in the on-state, the current of Laux is higher by roughly 10 to 20 %. In summary, the agreement is very good, keeping in mind that completely different methods have been used to calculate the ballistic quantum transport.

3.9 Conclusion

In this chapter a self-consistent formulation of the CBR method has been presented that allows for the calculation of the ballistic transport properties of realistic two-dimensional devices such as nano-MosFets under applied bias. Due to the reduction of the problem size, using only those lead modes and eigenstates of the decoupled device which contribute to the current and to the electron density, the method is extremely efficient compared to existing methods. Another important advantage of the method is the possibility to include bound and quasi bound states that would be neglected in the purely ballistic calculation, but inevitably appear in any biased quantum structure that contains barriers. In this chapter a classification of these states is provided and it has been demonstrated that such bound states exist in front of the gate oxide of a double gate MosFet and that their presence has a significant effect on the resulting device characteristics. The exchange-correlation interaction has been taken into account in local density approximation leading to a rise of the effective channel-barrier of roughly 50 meV due to the density difference between the highly doped leads and the depleted channel.

The comparison to recent results of Laux *et al.* [19] shows that our model is valid and leads to comparable results. However, we conclude that it is important to take into account the bound states and the exchange-correlation to obtain more realistic results.

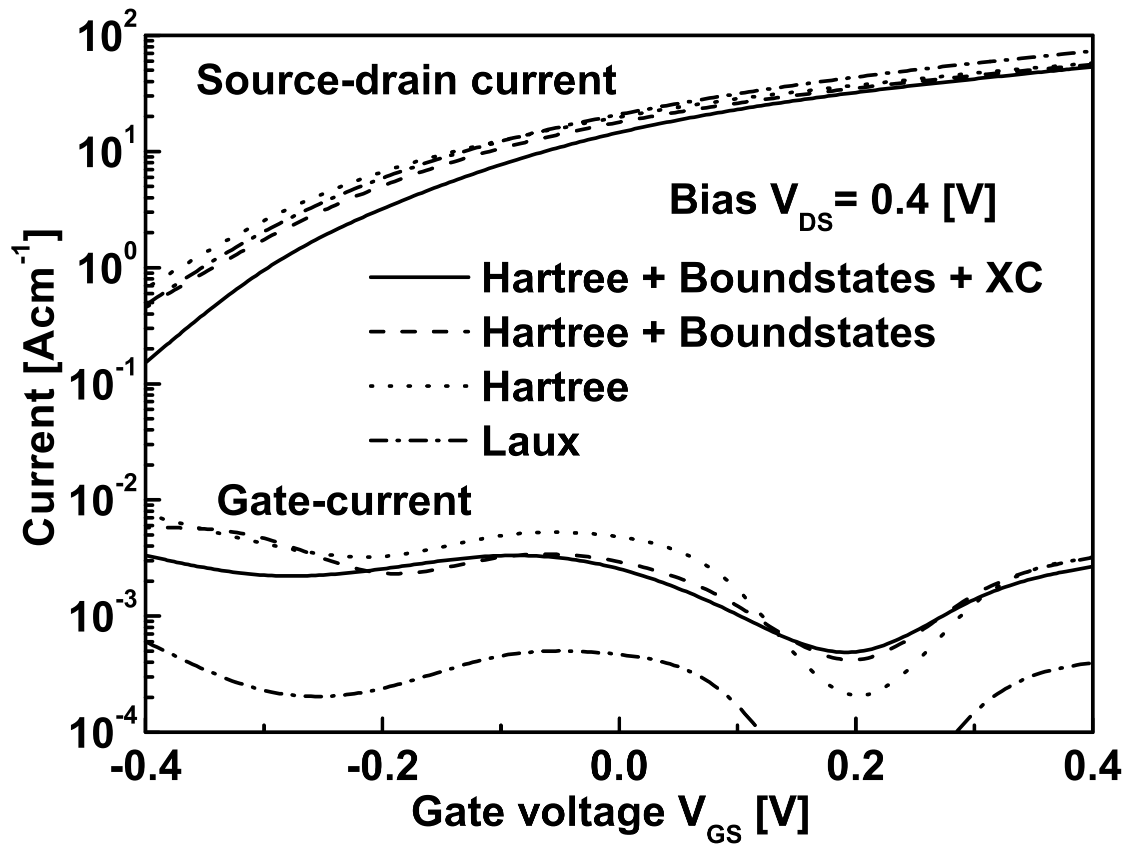


Figure 3.42: Comparing the source-drain and gate current for the three different models applied in this work with the results from Laux *et al.* [19].

Chapter 4

Inelastic scattering via Büttiker probes

4.1 Introduction

In the ballistic transport formalism as regarded in the previous part of this work, scattering, other than off the potential profile of the heterostructure, is not included. For many devices, especially for those that operate at room temperature it might be necessary to include elastic as well as inelastic scattering of the electrons with the phonon bath in order to obtain the correct device characteristics. The method of choice for this type of problems that depend on both, the inclusion of momentum and energy relaxation of the carriers, as well as the correct description of the quantum transport effects, is the non equilibrium Green's function (NEGF) method. In this formalism all types of scattering can be included in a very concise manner and the properties of the device far from equilibrium can be calculated. This method has been successfully applied to one dimensional calculations of resonant tunneling diodes [67, 24], or quantum cascade lasers [68]. The drawback of this approach is its computational effort that hampers the application to larger two or three dimensional systems. Therefore, simpler models have been proposed, that are able to capture most of the phenomenons connected to scattering, while being simple enough to minimize the computational effort. A very intuitive method is the concept of the Büttiker probes introduced by Büttiker [69] and further developed by Datta and Lundstroem [64]. The idea is, that additional contacts in the interior of the device are used to model the coupling of the carriers to a phonon bath. Depending on the coupling strength to these reservoirs, part of the carriers instead of ballistically propagating inside of the device, escape into the contacts and an equal number of thermalized carriers is reinjected to maintain current conservation. The distribution in the reservoirs can be assumed to be a Fermi distribution and the local Fermi level acts as a convergence parameter that has to be adjusted self-consistently in order to conserve the current. In this way energy relaxation is introduced into the system. It can be shown [64], that the coupling to these thermalized reservoirs, which is a free parameter of the method, is related to a carrier mobility that could be determined experimentally. Therefore the parameters are not completely arbitrary and the method retains some predictive power. Numerically, the calculation of the self-energy corresponding to the Büttiker probes is analogous to the self-energy terms accounting for the open boundaries. Importantly, since the self-energy is non-zero only for the diagonal terms, the model allows for the use of the efficient recursive Green's function method to evaluate the retarded Green's function of the device coupled to the leads and phonon reservoirs.

In the following we will briefly introduce the method, showing how potential and quasi Fermi levels are determined. The resulting currents and densities for different magnitudes

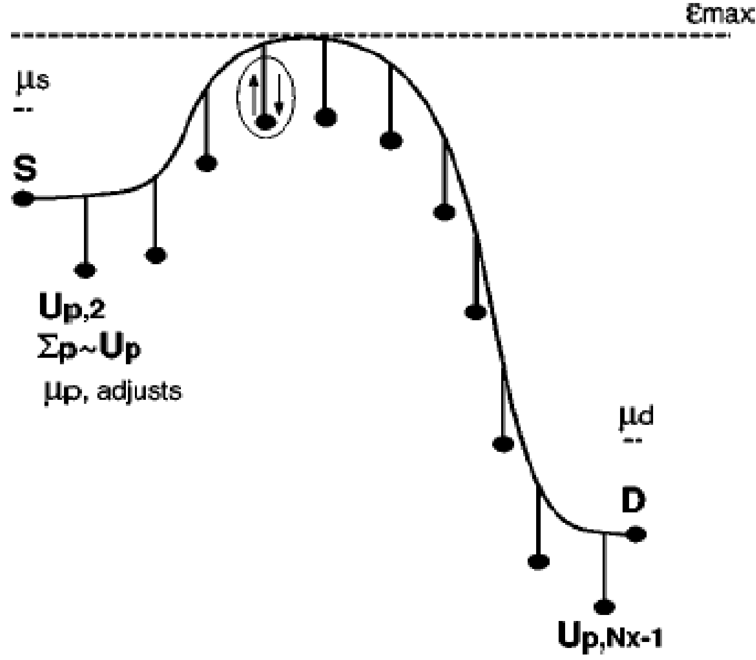


Figure 4.1: Simple sketch to illustrate the Büttiker probe method [64]. The source and drain contacts are to the left and to the right of the indicated conduction band edge and have fixed Fermi levels μ_S and μ_D. The Büttiker probes are indicated by the vertically attached black dots and have position dependent quasi Fermi levels μ_P.

of scattering are compared with an even further simplified model using constant scattering parameter η as proposed by Datta [29]. Figure 4.1 shows a sketch [64] that exemplifies the method showing the source (S) and drain (D) contacts attached to the left and right edge of the conduction band profile indicated by the solid line. The vertically attached black dots symbolize the additional contacts that act as the Büttiker probes. Electrons can scatter into the probes and a thermalized distribution of carriers corresponding to a local quasi Fermi level μ_P is reinjected into the device such that the number of particles is conserved.

4.2 Introduction to the method

To exemplify the method we use a simple device discretized on a one dimensional grid with homogeneous grid spacing Δ. Within the effective mass approximation assuming a constant effective mass m^*, the Hamiltonian results in a simple tridiagonal matrix $\mathbf{H}$:

$$\mathbf{H} = \begin{pmatrix} d_1 & -t & 0 & ... & 0 \\ -t & d_2 & -t & ... & 0 \\ 0 & -t & ... & ... & 0 \\ ... & ... & ... & ... & ... \\ 0 & 0 & ... & -t & d_N \end{pmatrix} \tag{4.1}$$

with the diagonal elements $d_i = V_i + 2t$, where V_i is the conduction band edge including the self-consistent potential at site i, and t being the constant intersite coupling element

$$t = \frac{\hbar^2}{2m^* \Delta^2}.$$

Following the approach of Datta [6] we are going to describe the method in a simple way that contains elements of the NEGF formalism but in a slightly different notation than the standard notation used in textbooks. Introducing this notation we relate the quantities like correlation functions and self-energies to the standard notation

$$\begin{aligned} \mathbf{G}^p &= -i\mathbf{G}^< \\ \mathbf{G}^p &= i\mathbf{G}^> \\ \mathbf{\Sigma}^{in} &= -i\mathbf{\Sigma}^< \\ \mathbf{\Sigma}^{out} &= i\mathbf{\Sigma}^> \end{aligned}$$

where the electron correlation function $\mathbf{G}^n$ is obtained from the retarded Green's function $\mathbf{G}^R$ and the self-energy $\mathbf{\Sigma}^{in}$ as follows

$$\mathbf{G}^n = \mathbf{G}^R \mathbf{\Sigma}^{in} \left(\mathbf{G}^R\right)^\dagger .$$

The retarded Green's function is obtained by the relation

$$\mathbf{G}^R = \left[E\mathbf{I} - \mathbf{H} - \mathbf{\Sigma}\right]^{-1} .$$

Within the Büttiker probe method the self energy matrix of a simple 1D device as sketched in fig. 4.1 has the following form:

$$\mathbf{\Sigma} = \begin{pmatrix} \Sigma_S + \Sigma_1^P & 0 & ... & 0 \\ 0 & \Sigma_2^P & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & \Sigma_D + \Sigma_N^P \end{pmatrix}.$$

with the self energies of the source and drain contacts Σ_S and Σ_D, respectively and the terms corresponding to the Büttiker probes Σ_i^P attached to grid point i ($i = 1 \dots N$). Within this approach the in-scattering matrix $\mathbf{\Sigma}^{in}$ is obtained from the imaginary part of the self-energy weighted with a local distribution function

$$\mathbf{\Sigma}^{in} = \begin{pmatrix} f_S\Gamma_S + f_1^P\Gamma_1^P & 0 & ... & 0 \\ 0 & f_2^P\Gamma_2^P & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & f_D\Gamma_D + f_N^P\Gamma_N^P \end{pmatrix} \quad (4.2)$$

with $\Gamma = i\left(\Sigma - \Sigma^\dagger\right)$ and f_i being the local Fermi distribution with quasi Fermi level μ_i. Within the self-consistent cycle the quasi Fermi level is determined such that current conservation is obeyed. The self-energy at each site is simply given by [6]

$$\begin{aligned} \Sigma_{S,D} &= -t\exp\left(ik\Delta\right) \quad (4.3) \\ \Sigma_i^P &= -ct\exp\left(ik\Delta\right) \quad (4.4) \end{aligned}$$

with the wavenumber k obtained from the dispersion of the discrete lattice

$$E = V_i + 2t\left(1 - \cos\left(k\Delta\right)\right). \tag{4.5}$$

The difference between the self-energy of the source and drain contacts $\Sigma_{S,D}$ and the self-energy of the Büttiker probes Σ_i^P is just the prefactor c with $0 \leq c \leq 1$. This parameter c is the handle to control the strength of the coupling to the probes and thus to vary the scattering. It has been shown [64] that, in the limit of weak scattering, the coupling c can be related to the mean free path λ in the following way:

$$c^2 = \frac{2\Delta}{\lambda}. \tag{4.6}$$

From the correlation function G^n properties such as the carrier density

$$n(\mathbf{r}) = \int dE \frac{1}{2\pi}\left[G^n\left(\mathbf{r},\mathbf{r}';E\right)\right]_{\mathbf{r}=\mathbf{r}'},$$

and the current density

$$J(\mathbf{r}) = \int dE \frac{1}{2\pi}\left[-\frac{ie\hbar}{2m}\left(\boldsymbol{\nabla} - \boldsymbol{\nabla}'\right)G^n\left(\mathbf{r},\mathbf{r}';E\right)\right]_{\mathbf{r}=\mathbf{r}'},$$

are obtained. It is convenient to define [6] a current operator $\mathbf{I}^{op}$ whose diagonal elements are the divergence of the current

$$I_{ii}^{op} \equiv \boldsymbol{\nabla}\cdot J_i = \frac{ie}{\hbar}\left[\mathbf{H}\mathbf{G}^n - \mathbf{G}^n\mathbf{H}\right]_{i,i}$$

Since scattering is present, current conservation is not obeyed for a single energy channel, but for the integral over the energy.

$$\int I_{ii}^{op}\left(E\right)dE = 0, \quad i \notin \text{boundary}$$

At the boundary points $i = 1$ and $i = N$, I_{ii}^{op} is not zero, but determines the total current through the device. For the self-consistent calculation we rewrite I_{ii}^{op} in terms of the retarded Green's function in the following way:

$$\begin{aligned} I_{ii}^{(op)}\left(E\right) &= \frac{e}{\hbar}\sum_{j=1}^{N} A_{ij}\left(E\right)\Sigma_j^{in}\left(E,\mu_j\right) \\ A_{ij} &= i\left[G_{ij}^R\left(t_{i-1}\left(G_{i-1,j}^R\right)^* + t_j\left(G_{i+1,j}^R\right)^*\right) - c.c.\right]. \end{aligned}$$

Now, let us define a functional

$$\begin{aligned} F_i\left(\boldsymbol{\mu}\right) &= \int I_{ii}^{(op)}\left(E,\boldsymbol{\mu}\right)dE \\ F_i\left(\boldsymbol{\mu}\right) &= \frac{e}{\hbar}\int dE \sum_{j=1}^{N} A_{ij}\left(E\right)\Sigma_j^{in}\left(E,\mu_j\right) \end{aligned}$$

which has to be zero for all $i = 2 \ldots N-1$ to satisfy the current conservation. (The first and the last elements of F_k determine the total current through the device). In order to determine roots of $F_i\left(\boldsymbol{\mu}\right)$ we can use Newton's method with the Jacobian

$$M_{ij} = \frac{\partial F_i}{\partial \mu_j} = \frac{e}{\hbar} \int dE A_{ij}(E) \beta \Sigma_j^{in}(E, \mu_j) \left(1 - \Sigma_j^{in}(E, \mu_j)\right)$$

Thus to determine the correction for the Fermi level $\triangle \boldsymbol{\mu} = \boldsymbol{\mu}^{(new)} - \boldsymbol{\mu}^{(old)}$ we solve the following linear system:

$$\sum_{j=1}^{N} M_{ij} \triangle \mu_j = -F_i, \ i = 1..N$$

with Dirichlet boundary conditions for the chemical potential $\mu_1 = \mu_S$ and $\mu_N = \mu_D$. Now using a simple iteration coupling the solution of Poisson's equation and the correction for the Fermi level a converged solution is achieved. It should be noted that since all quantities are functions of the energy, the resolution of the energy grid has to be adapted to obtain a reasonable convergence.

4.3 Additional remarks

There are two technical points that are not mentioned in the cited works, but important enough to receive additional attention. The first point is that using the self-energy defined in eq. (4.3) for the source and drain contacts together with the dispersion relation of eq. (4.5), the leads are effectively described employing the ballistic model. Since the device itself is modeled with the scattering introduced via the Büttiker probes which alters the local density of states compared to the ballistic case, the calculations result in a non-homogeneous density even for a completely homogeneous device with no bias applied as shown by the dashed line in figure 4.2. This artifact can be eliminated by treating the leads in the same way as the device to guarantee smooth interfaces. This is very easily done by adding the same self-energy term that is used inside of the device to the energy in eq. (4.5) to account for the scattering in the leads in the same way as in the contacts. The contact self-energy that includes these scattering terms now results in the expected homogenous density as shown by the solid line in figure 4.2.

The second point that comes up when using the Büttiker probe model is the fact that the probes can be regarded as channels parallel to the propagation inside of the device. Therefore, if a fraction of the carriers is propagating into the leads, only the remaining current can flow in the device, which has to be accounted for in the intersite coupling element t of the Hamiltonian defined in equation (4.1). This intersite coupling has to be rescaled in the following way

$$t' = (1 - c)t \tag{4.7}$$

to account for flux conservation. The effect of the rescaled intersite coupling is reflected in a dramatic change of the local density of states as shown in figure 4.3. The dashed line represents the DOS for the unchanged coupling element and exhibits strong oscillations that are unphysical since it is a homogeneous device where no reflections are expected. The rescaled intersite coupling t' from eq. (4.7) however results in the ordinary $1/\sqrt{E}$ behavior as expected for a homogenous one dimensional device. The local density of states plotted in fig. 4.3 is obtained from diagonal elements of the spectral function

$$\mathbf{S} = \mathbf{G}^R \boldsymbol{\Gamma} \left(\mathbf{G}^R\right)^\dagger \tag{4.8}$$

with $\boldsymbol{\Gamma}$ derived from the $\boldsymbol{\Sigma}^{in}$ of eq. (4.2) just setting all distribution functions f equal to unity.

After these two issues have been addressed we obtain the expected constant density for a homogeneous piece of conductor attached to two leads with exactly the same properties.

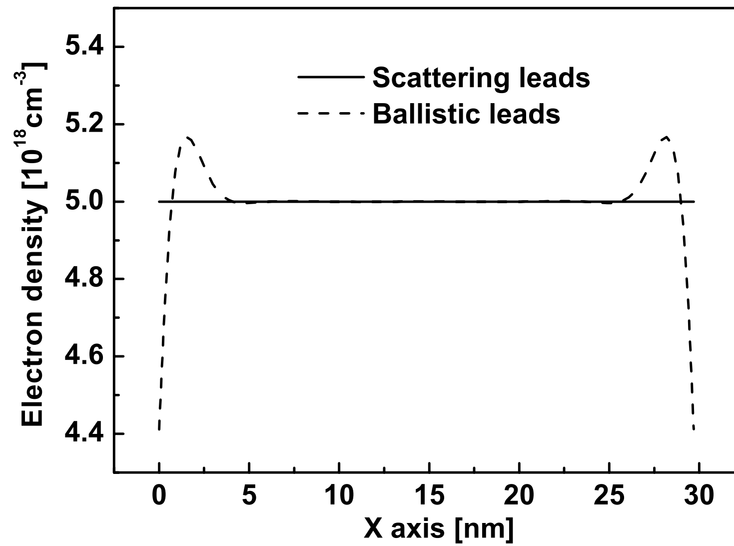


Figure 4.2: Density profile of a homogeneously doped conductor with no bias applied. Only the contacts calculated including the scattering (solid line) lead to the expected homogeneous density.

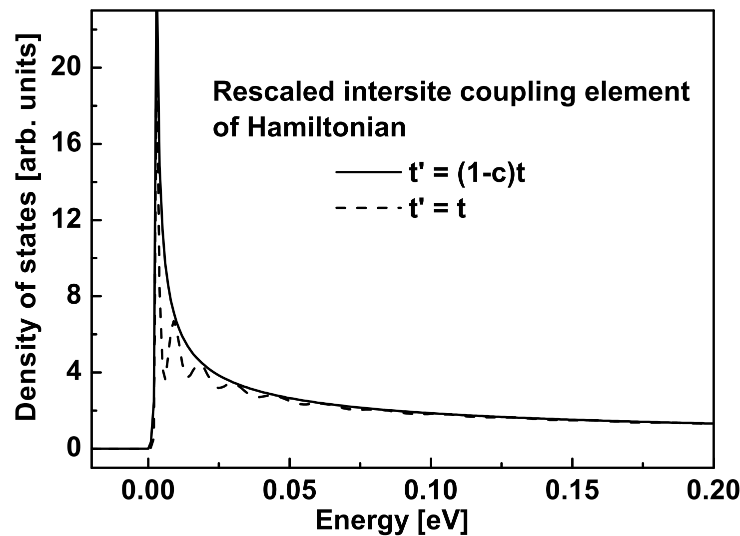


Figure 4.3: Local density of states at the contact for a homogeneous conductor with no bias applied. The dashed line shows the highly oscillating DOS resulting from the unchanged intersite coupling, while the solid line shows the expected smooth DOS resulting from the rescaled intersite coupling.

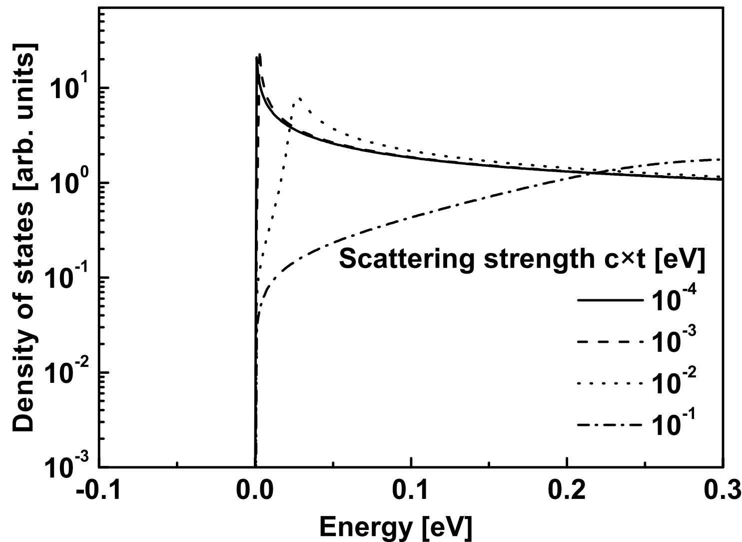


Figure 4.4: Local density of states at the source contact point for different scattering strength induced by varying the coupling term ct for the Büttiker probe self-energy.

4.4 Numerical results

In the following we are going to show the results of self-consistent calculations for two different devices, a nin resistor and a single tunneling barrier. In addition to model the Büttiker probes as real contacts we will also present results for a simplified scheme presented in [29] that replaces the probe self-energy defined in eq. (4.4) by a simple constant $\Sigma^P = -i\eta$ $(\eta > 0)$. To start with, we have a look at the local density of states obtained from the diagonal elements of the spectral function 4.8 at the source point of the device. As already mentioned in the previous section, the shape of the DOS as a function of the energy is changed with the introduction of scattering. First we look at the DOS resulting from the Büttiker probe-model shown in fig. 4.4 where the ballistic limit with very weak scattering is indicated by the solid line. As the scattering is increased by increasing the coupling c in eq. (4.4) the shape of the DOS changes and the peak shifts to higher energies. As a result of this change of the DOS the self-consistent potential and Fermi level for low density samples are strongly dependent on the scattering rates. The effect of the scattering on the DOS is even more pronounced for the simple model of a constant self-energy $-i\eta$ as shown in figure 4.5. The fact that the probe self-energy is energy independent, leads to a non-zero DOS for negative energies that slowly tends to zero following a Lorentzian lineshape. This leads to the unphysical situation that carriers with energies below the band-gap are populated since the distribution function does not vanish for negative energies. Nevertheless, for small values of η or high enough densities the error is not dramatic and meaningful results can still be obtained. To evaluate the effect of the scattering on the resulting current we choose a nin resistor of 30 nm length that is doped with 5×10^{18} cm^{-3} and has an intrinsic region of 10 nm length in the center of the device. The device is modeled on a finite differences grid with 3 Å grid spacing, the mass is taken to be 0.32 m_0 and the resolution of the energy grid is 1 meV. For an applied bias of 25 mV, fig. 4.6 shows the

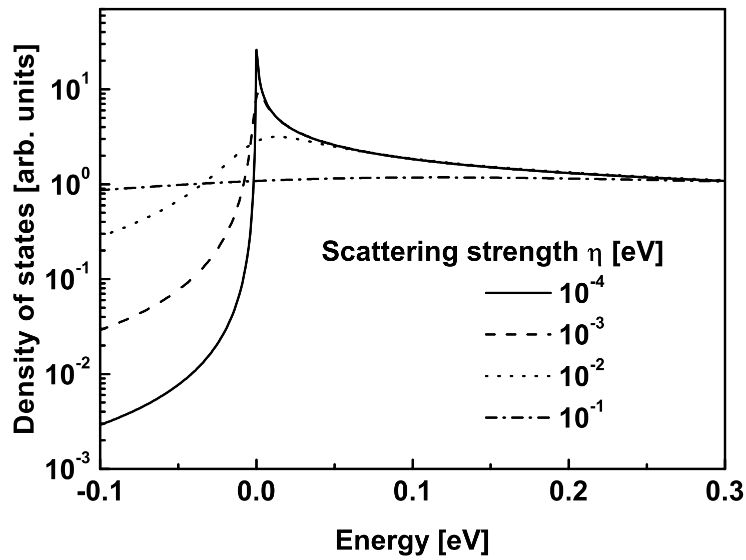


Figure 4.5: Local density of states at the source contact point for different values of the constant probe self-energy $\Sigma^P = -i\eta$.

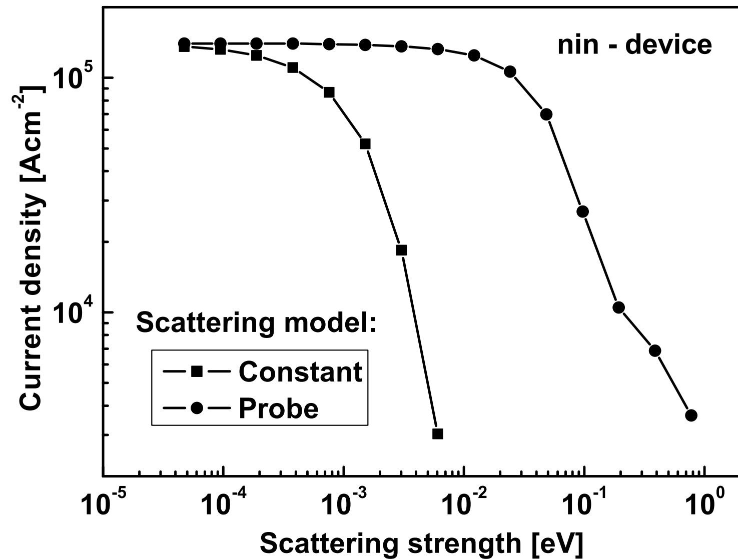


Figure 4.6: Comparing the source-drain current resulting from the two models for a *nin* resistor with an applied bias of 25 mV, as a function of the scattering strength.

resulting current density comparing the two different models, as a function of the scattering strength. It should be noted that for the model with the constant Σ^P the scattering strength is the parameter η, while for the Büttiker probe model, the scattering strength is ct. For a scattering strength below 10^{-4} eV, both models yield the ballistic limit of 1.4×10^5 Acm^{-2}. For higher scattering strengths the resulting current of the two models differ tremendously, with the simple model with constant Σ^P reducing the current much stronger than the Büttiker probe model. The last current point of the probe model curve corresponds to a coupling parameter c equal to 0.7 which means that more than two thirds of the electrons are thermalized at each grid point leading to an extremely high scattering rate. The strong effect of the constant Σ^P model is partially explained by the purely imaginary $\Sigma^P = -i\eta$ compared to the complex Σ^P of eq. (4.4), since the imaginary part of Σ is proportional to the density of states of the probe. But the difference in the currents also depends on the device which is illustrated by a plot similar to fig. 4.6 but for a homogeneously doped conductor (nnn) as shown in figure 4.7. The shapes of the two lines corresponding to the two different models are again very similar, but the difference in the scattering strength for the same resulting current is only one order of magnitude for the nnn device compared to two orders for the nin. To compare the results to a semiclassical calculation, we plotted the expected current resulting from a drift diffusion model

$$j_{\text{DD}} = \mu \cdot n \cdot \nabla E_F$$

by the dashed line in figure 4.7. Since the mean free path and therefore the mobility μ is inversely proportional to the square of the coupling parameter c (see eq. 4.6) we plotted the drift-diffusion current j_{DD} as a function of the square root of the mobility (top axis of fig. 4.7). It is observed, that for strong scattering the Büttiker probe model roughly follows the semiclassical result, while for low scattering, the ballistic limit is reached corresponding to a mobility of ~1000 [Vscm^{-2}].

Additional insight into the physics of the models is obtained by looking at the self-consistent potentials and Fermi levels. As explained in the previous part of this section, the scattering alters the density of states, which has a strong influence on the relative position of the potential and the Fermi level in a self-consistent calculation. To demonstrate this fact we plotted the potentials and the quasi Fermi levels of the nin resistor for scattering rates that correspond to a source-drain current of half of the ballistic limit. This condition results in a scattering strength η for the constant model of about 1 meV, while ct for the probe model is 50 meV. Figure 4.8 shows the potentials and Fermi levels of the constant and the probe model indicated by the dashed and dotted lines, respectively. The potential resulting from the ballistic model is shown by the solid line for comparison. While the quasi Fermi levels are almost identical for both models, the potential for the probe models is very different from both the ballistic and the constant-Σ model. To explain this we have to keep in mind, that the boundary conditions for the Fermi levels are Dirichlet conditions with fixed Fermi levels at source and drain, while the potential is calculated with von Neumann boundary conditions to guarantee the charge neutrality of the device. For a scattering strength of 1 meV as used for the constant model, the density of states is very close to the ballistic one as shown by the dashed line in figure 4.5. This leads to a potential that is very similar to the ballistic case. For the probe model on the other hand, the scattering strength has to be about 50 meV to achieve the same current reduction of 50 %, which results in a density of states that is highly distorted as shown by the dotted line in figure 4.4. Since the peak of the DOS is shifted towards higher energies, the potential has to be lowered in order to retain the same electron density. This effect will be less pronounced for higher densities, since the integrals over the DOS for different scattering strengths converge to

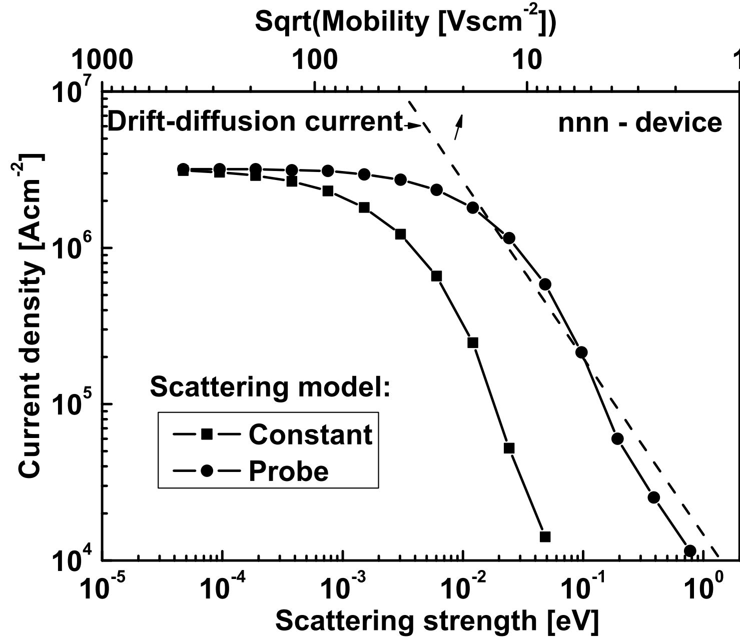


Figure 4.7: Comparing the source-drain current resulting from the two models for a *nnn* resistor with an applied bias of 25 mV, as a function of the scattering strength. The dashed line corresponding to the top scale, indicates the resulting current expected from a drift diffusion model.

the same value for higher energies.

The effect of the damping due to the scattering is nicely seen by comparing the local density of states (LDOS) along the axis of the device for different scattering rates. Figure 4.9 shows the LDOS of the nin-resistor for low scattering of 1 meV revealing the oscillations in front of the barrier that have also been observed in the DGFET (see figure 3.38). The white line indicated the band-edge resulting from the self-consistent calculation in equilibrium. These pronounced oscillations vanish due to the damping if the scattering is increased. Figure 4.10 shows the LDOS for the same structure but with a coupling of the Büttiker probes of 50 meV. There are two main differences to the low scattering limit. As expected, the oscillations are strongly damped due to the scattering introduced by the Büttiker probes. Interestingly, the energetic distance of the maximum of the DOS to the band edge is almost 50 meV compared to only a few meV for the low scattering example. This is again due to the shift of the peak of the DOS towards higher energies for strong scattering as shown in figure 4.4.

4.5 Conclusion

As a result of the investigations of the two simplified scattering models, using Büttiker probes as well as the constant scattering self-energy, we note that there are still open questions that have not been addressed in the cited publications. Besides the details of the boundary conditions and the correction to the Hamiltonian that have been shown in section 4.3, the main issue remains the strong dependence of the density of states on the scattering strength and the resulting shifts in the self-consistent potentials. This effect is present in both models although it leads to shifts of the potential in opposite directions thus lowering the potential relative to the Fermi

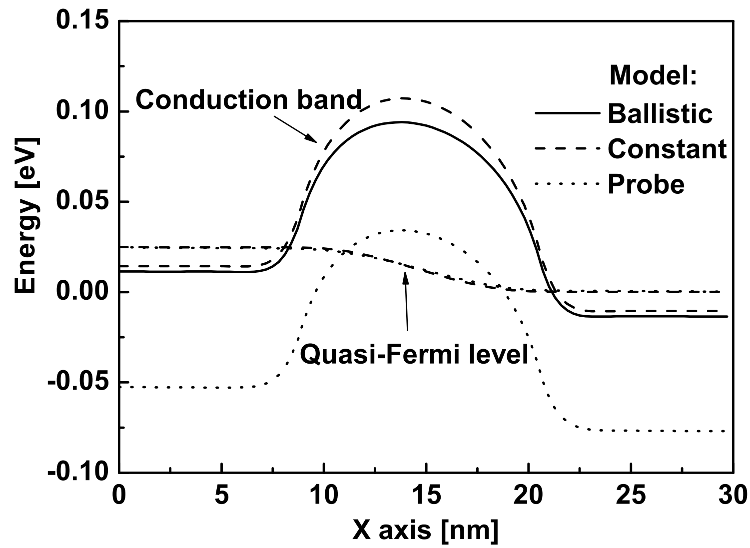


Figure 4.8: Self-consistent potentials and quasi Fermi levels of the two scattering models for a resulting source drain current of half of the ballistic limit. The potential profile for the ballistic solution is depicted by the solid line for comparison.

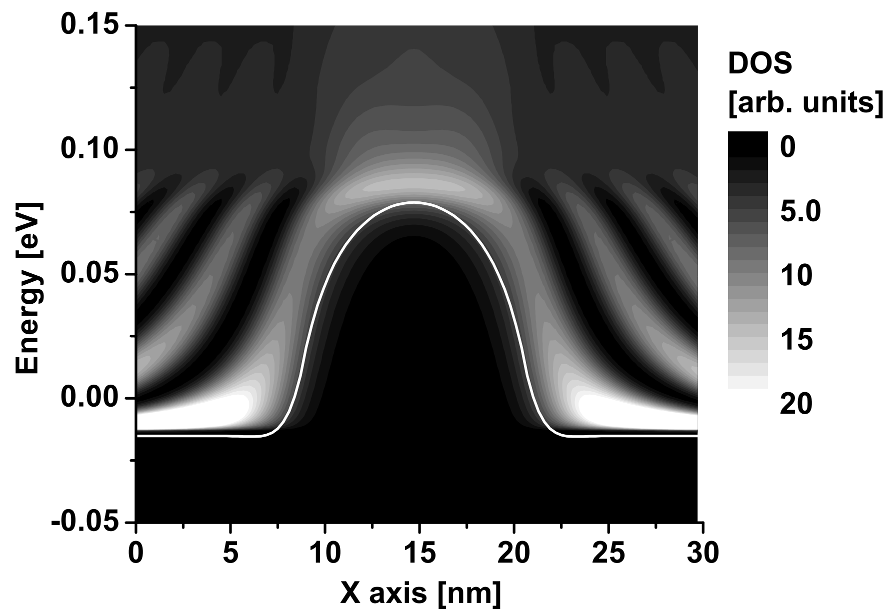


Figure 4.9: Local density of states of the nin resistor for very low scattering

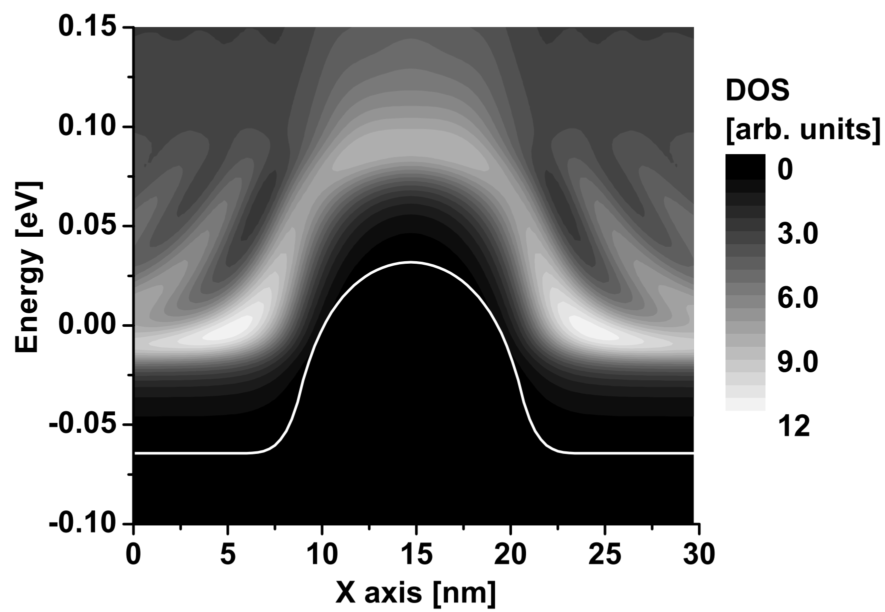


Figure 4.10: The local density of states of the nin-resistor calculated using the probe model with a coupling of 50 meV. The white line indicates the corresponding band-edge.

level for the probe model but raising it in the constant Σ model with increasing scattering strength. Since the scattering efficiency measured in the reduction of the resulting current compared to the ballistic case is much higher in the constant model, the self-energy may be smaller to yield the same mean free path length and therefore the change in the potential is less dramatic compared to the probe model as shown in figure 4.8. Thus, for the nin-resistor it seems that the constant model leads to reasonable results, while for a barrier structure this model is completely off resulting in an extremely high electron density inside the barrier depending on the scattering parameter. This arbitrariness of the result depending on the choice of the model and the parameters drastically reduces the ability to make predictions using these models. However, as a simple model to study certain effects that would be buried inside the self-consistent algorithm in a full NEGF approach the model is certainly useful.

Another issue is the possible implementation of scattering into the CBR method. Certainly the probe model is not an option since it corresponds to a contact block of the same size of the device, which renders a further reduction impossible. The only way would be the simple Datta model making use of the fact that we can put the constant and diagonal self-energy inside the Hamiltonian of the decoupled device. In this way we would obtain a non-Hermitian but still linear eigenvalue problem that might be solved yielding states with complex eigenvalues E_i that decay due to the scattering

$$\begin{aligned} \mathbf{H}^0_{NH} \left|\psi_i\right\rangle &= E_i \left|\psi_i\right\rangle, \\ \mathbf{H}^0_{NH} &= \mathbf{H}^0 - i\eta\mathbf{I}, \\ E_i &\in \mathbb{C}. \end{aligned}$$

The retarded Green's function would then be evaluated in the same way as in the ballistic case, since the dimension of the Dyson equation is still proportional to the real contacts. The main obstacle is however the need to calculate the full Green's function coupling any grid point to all other points inside of the device, leading to a computational effort to construct the Green's function of the decoupled device G^0 of

$$N_E \times N_{eigen} \times N_{tot}^2,$$

instead of

$$N_E \times N_{eigen} \times N_{tot} \times N_M,$$

as in the ballistic case. In the 2D example of the double gate MosFet presented in this work the difference would be more than a factor of 300 corresponding to the number of total grid points N_{tot} ($\sim 10^4$) divided by the number of propagating modes N_M (~ 30). This huge increase of computational effort is in no relation to the small difference in the resulting current shown in fig. 3.21, that is, in addition, strongly dependent on the parameters used. We conclude that the models are well suited to investigate certain dependencies very easily and compare to more sophisticated methods, but cannot replace them. Concerning the numerical efficiency, the method of choice would be the recursive Green's function rather than the CBR method since its efficiency seems vanish with the introduction of scattering terms inside of the device.

Chapter 5

Summary and outlook

Within this part a novel and efficient method, termed contact block reduction (CBR) method, for the calculation of the ballistic transport properties of open nanostructures, such as the transmission, the density of states and the density matrix, has been presented. Its efficiency rests upon the splitting of the problem into the calculation of a small number of eigenstates of a properly defined closed system and the solution of a few additional linear systems of small size proportional to the size of the contacts. We have shown that this method is applicable to single- and multi-band Hamiltonians and found that it suffices to include only a few percent of the eigenstates of the closed system Hamiltonian even in the multi-band case if generalized von Neumann boundary conditions are defined at the contacts. For single-band Hamiltonians, a further significant reduction of the problem size can be achieved by transforming into the basis of all propagating lead modes, which enables the calculation of the transmission through 3-dimensional structures with large contact areas such as a quantum dot resonant tunneling diode.

The method has been successfully implemented into a charge self-consistent algorithm using a predictor-corrector approach and an adaptive energy grid to achieve a reliable and fast convergence of the potential and the carrier density. In this way, the source-drain as well as the gate leakage current of nanoscale double gate MosFets have been calculated. A systematic analysis of the current-convergence with respect to the number of calculated eigenstates shows that even in the self-consistent scheme, less than 5 percent of all eigenstates are sufficient to calculate the electron density and transmission through the device up to a bias of 0.5 Volts. Another important result is the finding, that quasi-bound states are located in the gate region of the device which are not properly taken into account within the purely ballistic formalism. Therefore, in this work a method has been developed to account for the additional density resulting from the bound and quasi bound states, while calculating the current within the ballistic limit. Additionally, for highly doped devices, particle-particle interactions such as exchange and correlation are described within the means of local density formalism leading to an effective increase of the channel barrier compared to the pure Hartree potential. By comparing our results to published works, we validate our model and emphasize the efficiency of the method.

Within this work, the CBR method has been shown to be applicable to realistic devices that can be described within the ballistic limit. To exceed this limitation of the method to the ballistic regime, inelastic scattering has to be introduced into the calculation. Two simple models based on the Büttiker probes have been evaluated within this work, but so far no satisfactory integration into the CBR method has been found. However, this is not a final conclusion and following students might have brighter ideas, how to solve this problem.

A more straight forward, but nevertheless very important improvement would be the inclusion of magnetic fields, since most ballistic transport experiments are performed under the influence of strong magnetic fields. Especially spin transport phenomena could be investigated for large multi-terminal devices that are out of reach today due to the high computational effort of the existing methods.

In this thesis a first step has been taken to improve the efficiency of quantum transport calculations for semiconductor nanostructures and we are looking forward to the improvements and novel ideas that are to follow.

Part II

Quantum dots

Chapter 6

Introduction

The nanoscale devices regarded in the first part of this work, such as the double gate transistors, are artificially designed structures that have to be fabricated via lithographic techniques using either an electron beam or ultraviolet light to project the structural pattern onto the substrate. This, so called 'top down', approach has the advantage that it yields reproducible devices that can be designed with a rather high degree of precision, but so far the resolution of the techniques is not high enough for structural dimensions below a few tens of nanometers. To get into the fundamental limits of miniaturization with a single or few charged carriers fully localized within quasi-zero dimensional heterostructures, the opposite 'bottom up' approach has been proven to be the most successful. A well known example is the self-assembled growth of quantum dots by the epitaxy of semiconductors with naturally different lattice constants. In 1994 Petroff and DenBaars [70] demonstrated that the deposition of a thin pseudomorphically strained InAs-layer on a GaAs substrate (lattice-mismatch 7 %) leads to the spontaneous nucleation of small islands driven by the strain relaxation, when a critical thickness is exceeded. This, so called, Stranski-Krastanov growth mode is very efficient and results in small InAs quantum dots of lateral dimensions below 30 nm and heights below 10 nm, depending on the temperature and Indium flux during the growth. Because these quantum dots provide confinement for electrons and holes simultaneously with highly discrete energy levels similar to atoms, they are also called 'artificial atoms' that pose ideal candidates for optical spectroscopy measurements. The major drawback of this self-assembled growth is their size fluctuation that leads to a broadening of the lines in the optical spectra of QD ensembles and their random lateral distribution as shown in fig. 6.1 a). These obstacles have been partially overcome by single dot spectroscopy that uses shadow masks to extract a single dot for the optical measurement, as well as the stacking of quantum dots (see fig. 6.1 b)) that leads to much more homogenous growth for subsequent layers of QDs.

One advantage of QD's is that they provide different energy scales and physical features which can be easily varied over a wide range of values. Most important, perhaps, is that QD's also allow the control of direct quantum-mechanical electronic coupling with not only composition but externally applied voltages. These flexible systems represent therefore the ideal for theoretical and experimental investigations, where the interactions between light and matter can be studied in a fully controlled, well characterized environment, and with excellent optical and electrical probes. These features make semiconductor QD's promising candidates for applications in electro-optical devices such as QD lasers [71, 72], and in quantum information processing [73, 74, 75, 76]. In the latter case, one can exploit the optical excitation in a QD [73, 75] or its spin state [74, 76] as qubits. These high expectations are produced by experimental

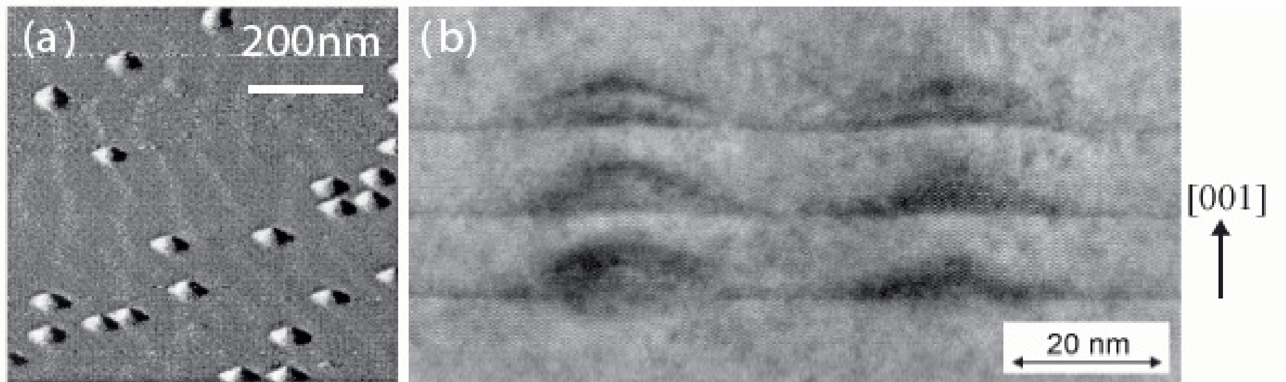


Figure 6.1: a) AFM image of a sparse array of In(Ga)As QDs deposited on GaAs. b) High resolution TEM cross-section of vertically correlated InAs–GaAs QDs via multilayer Stranski–Krastanov growth.

advances in the coherent manipulation of QD states, such as the exciton Rabi oscillations in single dots, achieved by the application of electromagnetic pulses [77, 78, 79, 80, 81].

This chapter is organized as follows. In section 7 a step by step introduction to the process of modeling a quantum dot structure, including the discussion of the underlying physical equations is given. Two investigations on quantum dots that have been performed in close collaboration with the experimental group of prof. Finley are presented. In section 8, the measured change of the intrinsic electron-hole dipole as a function of the excess charge in a single quantum dot is compared to the model revealing a charge redistribution that is highly sensitive to the dot size and alloy composition. The predictive capability of the model is demonstrated in the second example in section 9, where the quantum mechanical coupling of two vertically stacked quantum dots, a so called quantum dot molecule (QDM), has been calculated as a function of the interdot separation with the theoretical results being in excellent agreement with recent experimental measurements.

Chapter 7

Modeling quantum dots

7.1 Motivation

Quantum dots have been subject to a vast amount of experimental and theoretical works since their first evidence in the beginning of the 90's [82]. It might be due to the advent of fast enough computers, but more likely because of the complex structure of the quantum dots, that sophisticated numerical modeling was used to interpret the experimental results right from the beginning. In contrast to simpler heterostructures, such as quantum wells and resonant tunneling devices, that can roughly be described using analytical models, the complex geometry and composition of the self-assembled quantum dots demand a more detailed description. The experimentally measured photoluminescence (PL) or photocurrent (PC) is the result of a complex interplay between various effects, such as strain induced band-deformation, piezoelectric charges, quantum confinement and few-particle interactions, that have to be taken into account for a numerical model to be able to predict meaningful results. The works that lead the foundation of this field were published by Grundmann [83, 84], and Pryor [85, 86] who first used three-dimensional models that were able to calculate the strain distribution for a quantum dot and wetting layer embedded in GaAs material. With the device simulator next**nano**^{3} we extended these approaches allowing a more flexible geometry, few-particle interactions within density functional theory (DFT) and transport through quantum dots using the CBR method, introduced in the first part of this work. The quantum mechanical properties of the heterostructure are described using either single-band or multi-band Hamiltonians in $\mathbf{k}\cdot\mathbf{p}$ parametrization. Recently, works have been presented that employ Hamiltonians that include more details of the electronic structure, such as tight-binding [87] or pseudo potentials [88], but at the expense that external electric or magnetic fields, that are the key probe of any experiment, cannot be considered. Therefore it is fair to say that next**nano**^{3} is still state of the art, as the publications [89, 90, 91] and the results presented in this work demonstrate.

At the advent of the self-assembled quantum dots, the main idea of modeling could best be described as 'reverse engineering'. Since the shape and composition of the buried quantum dots was not easily detectable by experimental techniques, little was known besides their overall dimensions. With the combination of experiments that probe the optical and electronic structure as a function of externally applied electric fields, and numerical models that predict these quantities and their dependence on the dot shape and composition, it is possible to determine the parameters of the dots such as shape and composition for particular samples. Once this structure is found, the model can be verified by regarding additional properties, such as few-body states or the coupling of two quantum dots that form an artificial molecule, topics

that will be addressed throughout this part of the thesis. The beauty of a model is certainly that the influence of individual physical effects can be easily singled out by simply switching on or off the corresponding equations in the model, leading to a better understanding of the underlying physics. Many quantities that are not directly detectable by any experiment such as wavefunctions, or piezoelectric charges, can be calculated and visualized, to help the experimentalist to interpret the measurement. Finally, models are well suited to estimate certain possible applications of quantum dot systems such as quantum cryptography and quantum computation, which is of great importance in order to draw money from government sources.

In this section we describe how quantum dots are modeled using the device simulator next**nano**^{3}, giving step by step introduction starting from the choice of geometry and alloy profile, up to the calculation of few-body states. Within this work next**nano**^{3} has been extended to enable the calculation of few-particle states such as excitons or charged excitons via a self-consistent density functional approach. The involved equations, such as the definition of asymptotic boundary conditions for the Poisson equation, the calculation of the exchange and correlation potentials and energies within local spin density approximation (LSDA) and the self-consistent implementation are described. An overview of the physical equations governing the modeling of the quantum dots is given, but for more detailed information on the fundamental numerical models and their implementation in next**nano**^{3} we refer to the thesis of Stefan Hackenbuchner [92] who is the main author of the code.

7.2 Dot shape

The first models for the shape of self-assembled InAs quantum dots were steep pyramids [83] similar to those experimentally observed for surface grown quantum dots. This picture does not hold anymore for buried dots, since the interdiffusion of Gallium and Indium during the capping process changes the shapes of the dots compared to the ones on top of the surface. Additionally it is possible to grow dots with less then hundred percent Indium content which also affects the geometry and composition profile. Within the simulation we employ mainly two different types of shapes that have been proposed in the literature, namely the truncated pyramid and the lens. Even though the exact shape of each individual dot will never be captured in this way, the rough classification makes sense since properties like the piezoelectric charges or the lateral polarizability are strongly dependent on the dot shape. In fig. 7.1 the two classes of shapes are depicted with the truncated pyramid on the left and the lens shape on the right. The dimensions vary typically between 15 and 30 nm width and 3 to 5 nm height including the wetting layer. The angle of the pyramid facets may vary, but within the simulation they are restricted to discrete values due to the finite mesh size, since in a 3D calculation, the grid spacing cannot fall below 5 Å because of memory and cpu-time consumption.

As already mentioned, one of the main consequences of the choice of the dot shape is the distribution of the piezoelectric charges as shown in figure 7.2. The piezoelectric polarization is directly proportional to the off-diagonal elements of the strain tensor which are highest along the edges of the dots. Since the truncated pyramid has very pronounced edges, the resulting dipole charges align along these edges and reach far into the interior of the dot, a finding which is in good agreement with [84]. The potential caused by the strong dipoles of the truncated pyramid has a much stronger influence on the wavefunction, then the comparably weak polarization of the lens shaped dots. For very large dots this might affect the electron hole overlap and therefore the efficiency of the optical response of the quantum dot as shown in [93]. It turns

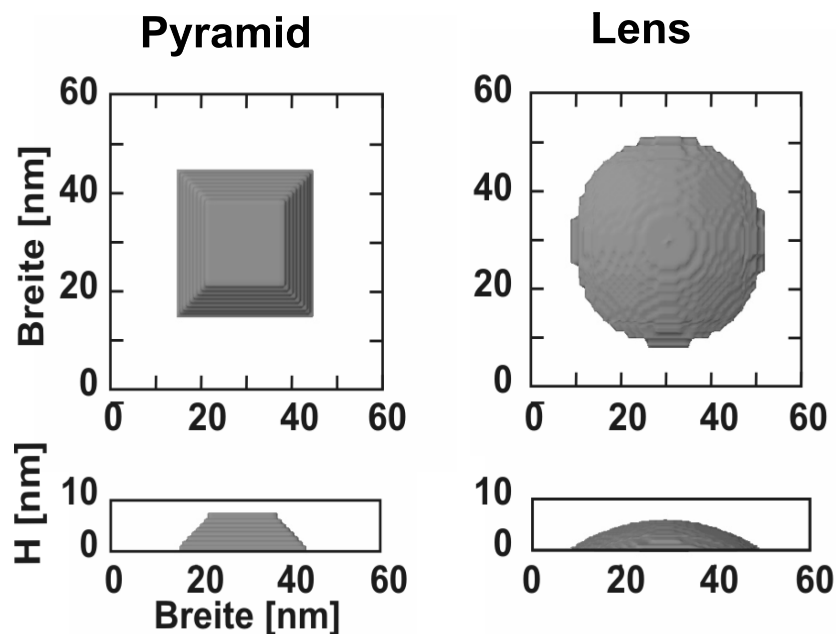


Figure 7.1: Top and side views of the two classes of quantum dot shapes used in the calculation, truncated pyramid and lens.

out that the dots may have both shapes, depending mainly on the growth conditions of the samples.

7.3 Alloy composition

In contrast to the pure InAs quantum dots, InGaAs dots contain a significant amount of Gallium, depending on the growth conditions. Since only the average amount of Gallium can be controlled during the growth, the distribution of the Indium inside of the dot during the process of self-assembly is unknown. However, the alloy profile has a strong influence on the shape of the confining potential and therefore affects the energies and the localization of the electron and hole wavefunctions. One of the first major breakthroughs in the investigation of the alloy profile of InGaAs quantum dots was the finding that the Indium content increases towards the tip of the dot [94] which leads to a localization of the hole in the tip while the electron remains delocalized over the whole dot, due to the lighter mass. This separation of the positive and negative charges results in an intrinsic dipole, that can be detected using Stark-effect spectroscopy. Within this technique, the energy of the neutral exciton is detected by means of PL or PC as a function of an external electric field applied along the growth direction. Due to the polarizability of the exciton, the field dependence of the exciton-energy is mainly quadratic, with a linear term originating from the intrinsic dipole revealing the relative position of electron and hole. Thus, comparing the experimental results to Stark shifts calculated with dots having different alloy profiles, the type of profile and the maximum and minimum Indium content can be determined. This example shows that the use of modeling is inevitable to obtain quantitative information on the composition of the quantum dot.

In the work of Fry *et al.* [94] only linear profiles have been considered where the Indium

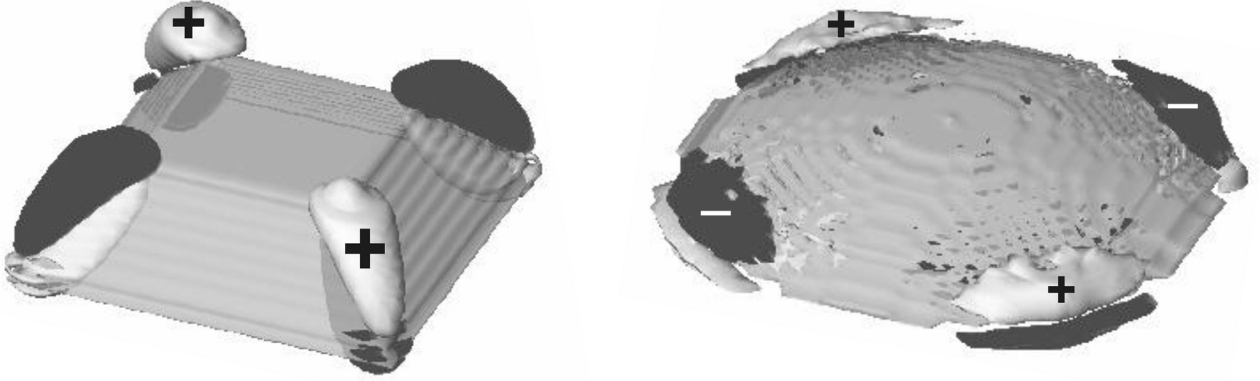


Figure 7.2: Localization of the piezo-electric charges for the two dot shapes. The positive and negative charges are indicated by the light and dark isosurfaces, respectively, that represent a charge density of 5×10^{18} cm^{-3}.

content c_{In} is varied along the growth direction

$$\text{Linear:} \qquad c_{\text{In}}(z) = c_{\min} + (z - z_{\min})(c_{\max} - c_{\min}) \quad \text{for} \quad z_{\min} \leq z \leq z_{\max},$$

with the minimum and maximum Indium content $c_{\min}$ and $c_{\max}$, and the coordinate along the center axis z with $z_{\min}$ and $z_{\max}$ pointing to the base and apex of the QD, respectively. Recently more detailed models have been proposed on the basis of experimental evidence using scanning tunneling microscopy (STM), as well as theoretical calculations of the growth process. Both consider an additional lateral variation of the Indium content. For simplicity the alloy profile is still isotropic around the center axis of the quantum dot. Figure 7.3 shows the three different types of profiles considered in this work in the upper row. In the model of Tersoff *et al.* [95] the Indium content depends solely on the angle to the center axis, with high Indium content for small angles as indicated by the light regions in figure 7.3. This leads an Indium rich zone that resembles an inverted pyramid with the tip pointing towards the bottom of the quantum dot. In a simplified way the Indium content is then described by

$$\text{Tersoff:} \qquad c_{\text{In}}(\varphi) = c_{\max} - (c_{\max} - c_{\min})\cos^2(\varphi),$$

with φ being the angle to the center axis. Finally the most refined model proposed by Migliorato *et al.* [96] confirmed by experimental evidence from Walther *et al.* [97], takes into account details of the growth dynamics that lead to a profile that resembles the horn of a trumpet and therefore will be called 'trumpet' profile throughout this work. Its Indium distribution is given by the formula

$$\text{Trumpet:} \qquad c_{\text{In}}(x,y,z) = c_{\min} + (c_{\max} - c_{\min})\exp\left\{\frac{-\sqrt{x^2+y^2}\exp(-2zC_3)}{\alpha}\right\}.$$

Note that within next**nano**^{3} the formula has been implemented in its original form from Migliorato with the parameters $C_1 = 1 - c_{\max}$ and $C_2 = 1 - c_{\min}$. The maximum concentration is again in the center axis of the dot

$$c_{\text{In}}(x=0, y=0, z) = c_{\min} + (c_{\max} - c_{\min}) \cdot 1 = c_{\max}.$$

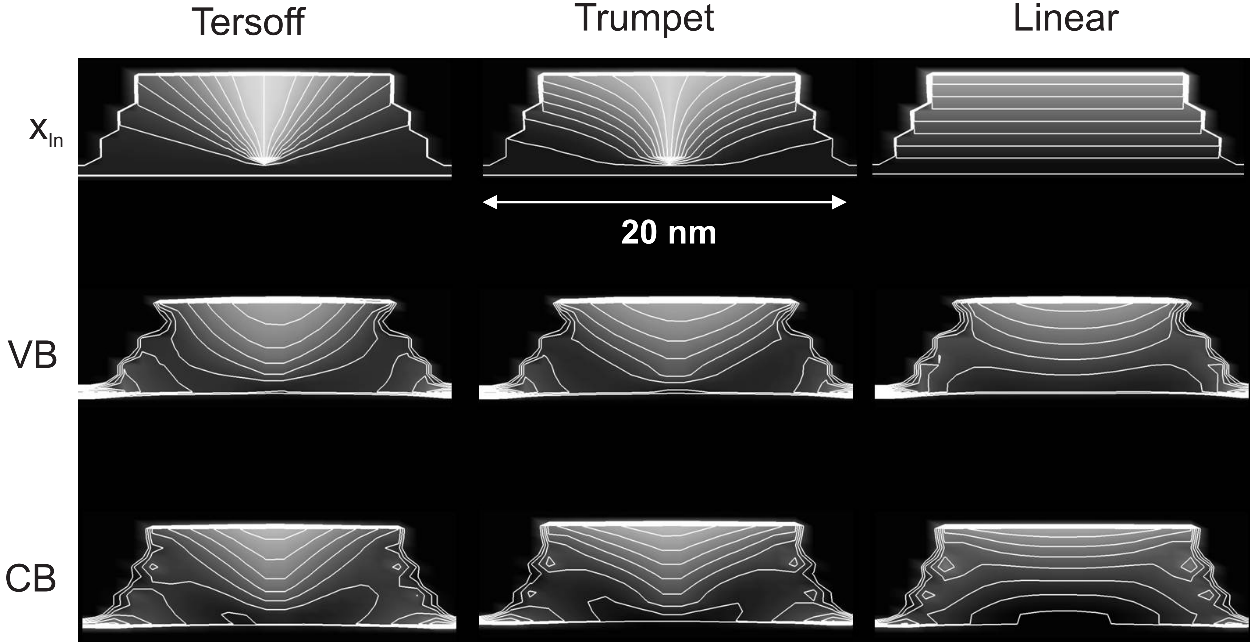


Figure 7.3: The graph shows the three different alloy profiles named Tersoff, Trumpet and Linear in the three columns. The first row shows the actual alloy profile with the light regions indicating high Indium content. The second and third row show the resulting valence band and conduction band profile, respectively. Here the light regions correspond to strong confinement.

The two remaining parameters C_3 and α can be used to vary the shape of the alloy profile while keeping the average Indium concentration fixed. For most of our calculations we used parameters of

$$\begin{aligned} 0.3 &< C_3 < 0.5, \\ 0.5 &< \alpha < 0.8, \end{aligned}$$

but they certainly depend on the size and the shape of the dot.

The resulting valence and conduction band profiles that are calculated including the deformation induced by the strain of the fully relaxed structure are shown in the two lower rows of figure 7.3. The light areas correspond to strong confinement and it is seen that the strength of the lateral confinement decreases going from the Tersoff to the linear profile. Interestingly, the linear profile with a homogeneous Indium distribution in the lateral directions, still offers a weak parabolic confinement that is due to the strain induced band deformation in the QD. A conclusive experiment to determine the lateral confinement and therefore the possible alloy profile of a QD would be the Stark-spectroscopy with laterally applied fields as shown in [91]. Unfortunately, due to technical difficulties, no satisfying results for measurements of this kind have been reported so far. However, it turns out that indications for the alloy profile can be found by modeling other experiments such as the vertical Stark shift of charged excitons, that have shown to be very sensitive to the lateral confinement [89].

Since shape and composition of the dots can be very different depending on the growth conditions there is no unique profile that fits to every dot. However, comparing the experimental

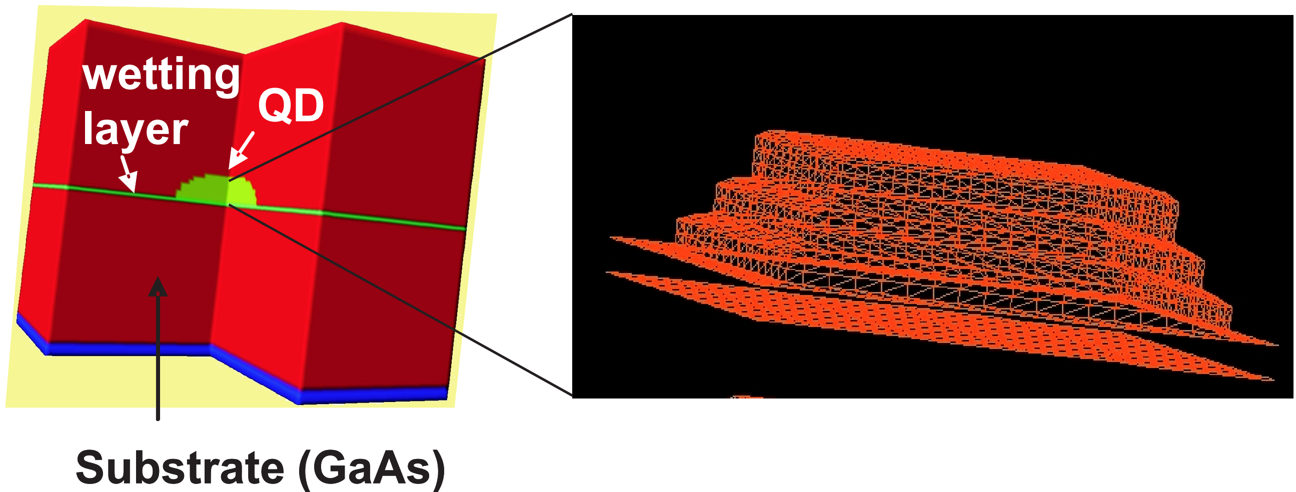


Figure 7.4: The left picture shows the simulation domain used for the calculation of the electronic properties of a single dot of roughly 20 nm width on top of a thin wetting layer embedded in GaAs material. The right picture shows the mesh that is used for the discretization of the structure.

data with calculations based on the models presented in this section it is often possible to determine the most likely profile for the particular sample measured in the particular experiment.

7.4 Strain

With the choice of shape and alloy profile of the quantum dot, the structure is not completely determined since the strain depends strongly on the environment in which the dot is embedded. As mentioned in the introduction, the dots considered in this work are overgrown such that they are buried within a GaAs cap layer with the dot placed on top of an InGaAs wetting layer of one or two monolayers thickness. Since, in the process of strain relaxation inside the QD, the substrate and cap layer will be strained as well, it is important to use a simulation domain that is large enough such that boundary effects do not alter the results. Therefore the typical dimensions of a simulation domain are three times the width of the dot as depicted in fig. 7.4 discretized on a grid with a spacing equal or below one nanometer to properly resolve the structural features of the quantum dot.

The proper calculation of the strain in this relaxed structure is of tremendous importance in order to obtain the correct deformation of the band-structure via linear deformation potentials and to calculate the piezoelectric polarization charges. To calculate the strain, there are basically two existing models that describe the strain-relaxation in semiconductor heterostructures. The first one is the valence-force-field [98] method, which is an atomistic model bound to the atomic-lattice, and secondly, the continuum model (CM) [83] that can be employed on an arbitrary grid. While the atomistic model is more precise if it comes to structures that vary on atomistic scale, it is computationally very costly, since the entire structure containing substrate, dots and cap layer has to be resolved on an atomistic grid. Therefore in our calculation we employ the CM on an inhomogeneous grid with a small lattice spacing of 5 Å in the vicinity of the dot that is gradually increased further away from the dot. The boundary conditions applied in this model are such that part of the structure, in our case the dark layer below the

GaAs substrate in fig. 7.4, is assumed to be unstrained and acts as Dirichlet boundary condition for the displacement. The remaining boundaries are best described using periodic boundary conditions, since in a real sample the dots grow in densities of roughly 10^{10} cm^{-2} which results in an average inter-dot distance below 100 nm.

The strain tensor is defined as the symmetrized change of the displacement $\mathbf{u}(\mathbf{x})$

$$\varepsilon_{ij} = \frac{1}{2}\left(\frac{\partial u^i}{\partial x_j} + \frac{\partial u^j}{\partial xi}\right)$$

and is calculated via the minimization of the elastic energy

$$E_{elastic} = \frac{1}{2}\int_V d\mathbf{r}\; C_{ijkl}(\mathbf{r})\,\varepsilon_{ij}(\mathbf{r})\,\varepsilon_{kl}(\mathbf{r})\,.$$

Numerically this amounts to the solution of a large system of linear equations proportional to the size of the entire device with the details of the implementation found in [92].

7.4.1 Band-lineup

To obtain the band-offsets in a semiconductor heterostructure, we employ the model solid theory of van de Walle [99, 100]. In this model the band-edges of different materials are defined all on the same absolute scale, such that the offset between the bands of different materials simply results from the difference in the absolute values. This scheme of the 'natural' band offsets works quite well for similar materials as the group of the III-V's, but certainly neglects any effects that are selective to the type of the interface between the two materials, or the crystal orientation. Since it is not possible to measure the energy of the band-edges on an absolute energy scale, only theoretical values exist. In the next**nano**^{3} database we use the values calculated by Zunger [101], since they result from a more sophisticated model than the old values proposed by van de Walle and Martin [100].

In a strained heterostructure the band-edges are additionally shifted relative to the position obtained from the model solid theory by an energy that is linearly proportional to the strain and calculated using linear deformation potentials. The strain tensor can be decomposed [102] into a hydrostatic component that accounts for the change of the band-gap but preserves all degeneracies due to crystal symmetry, and the shear deformations which lower the crystal symmetry and break the degeneracies in the valence band as well as in the conduction band minima that are located off the center of the Brioullin zone. An example for the resulting band-structure including the deformation due to the strain is shown in figure 7.3. For a more detailed description of this model the reader is referred to the diploma thesis of Matthias Sabathil [103], or the Ph.D. thesis of Stefan Hackenbuchner [92].

7.4.2 Piezoelectric charges

The piezoelectric polarization $\mathbf{P}$ is directly proportional to the off-diagonal elements of the strain tensor

$$\mathbf{P}_{\text{piezo}}(\mathbf{r}) = 2e_{14}(\mathbf{r})\begin{pmatrix}\varepsilon_{23}\\ \varepsilon_{13}\\ \varepsilon_{12}\end{pmatrix},$$

with the piezoelectric-constant e_{14}, that is a material property. The piezoelectric charges can be easily obtained from

$$\rho_{\text{piezo}}(\mathbf{r}) = -\text{div}\, \mathbf{P}_{\text{piezo}}(\mathbf{r}).$$

In the case of a pyramidal quantum dot the piezoelectric charges are strongly localized at the edges of the dot generating an octupole as shown in figure 7.2. The corresponding potential lowers the symmetry of the dot and leads to a splitting of the otherwise degenerate first excited electron states by up to a few meV.

7.5 Poisson equation

Within the self-consistent calculation of few-particle states in quantum dots, the repeated solution of the Poisson equation is needed to obtain the Hartree interaction between the charged particles confined in the quantum dot. Additionally the potential resulting from the constant back-ground charges has to be solved once at the beginning of the calculation. The discretization of the Poisson equation is performed on a finite volume grid using a box integration scheme [92] and the resulting system of linear equations is solved employing the conjugate gradient (cg) method.

The total potential can be split into a constant part ϕ_{bg} that does not change within the self-consistent cycle and the Hartree part ϕ_{H} that describes the Coulomb interaction between the confined particles, which is modified during the self-consistent cycle. The potential ϕ_{bg} results from the back-ground charges ρ_{bg} ,such as the piezo- and pyroelectric charges, is calculated via Poisson's equation

$$-\text{div}\left(\varepsilon(\mathbf{r})\, \text{grad}\phi_{\text{bg}}(\mathbf{r})\right) = \rho_{\text{bg}}(\mathbf{r}),$$

with von Neumann boundary conditions

$$\left.\frac{\partial\phi_{\text{bg}}(\mathbf{r})}{\partial\mathbf{n}}\right|_{\partial G} = 0,$$

with the unit-vector $\mathbf{n}$ normal to the surface of the domain ∂G. These boundary conditions are possible since the back-ground charges are polarization charges that preserve charge neutrality within the simulation domain. Since ρ_{bg} is constant, the potential ϕ_{bg} is calculated only once and then added to the total potential in each iteration. To calculate the Hartree potential created by an ensemble of electrons and holes confined in a quantum dot that does carry a net charge, different boundary conditions have to be applied, since the electric field is non-zero at the boundary. In this case the appropriate boundary conditions are vanishing potential at infinity

$$\phi_{\text{H}}(\mathbf{r}) \to 0 \quad \text{as} \quad |\mathbf{r}| \to \infty.$$

Since the electrostatic potential is long range, asymptotic boundary conditions have to be applied, in order to use a finite simulation domain [104]. To do this, the potential at the boundary is calculated using the integral expression of the potential resulting of an arbitrary charge distribution $\rho(\mathbf{r})$

$$\varphi(\mathbf{r}) = \frac{1}{4\pi\bar{\varepsilon}_r\varepsilon_0}\int d^3r' \frac{\rho(\mathbf{r})}{|\mathbf{r}-\mathbf{r}'|}, \tag{7.1}$$

with an averaged dielectric constant $\bar{\varepsilon}_r$. The potential given by eq. (7.1) is evaluated at the boundary ∂G and then applied as a Dirichlet boundary condition for the Poisson equation.

In this way the asymptotic behavior of the potential is taken into account, while keeping the numerical efficiency of the linear Poisson equation. To speed up the evaluation of eq. (7.1) at the boundary points we employ a multipole expansion of the charge distribution $\rho(\mathbf{r})$.

The electrostatic Green's function contained in eq. (7.1)

$$\frac{1}{|\mathbf{r}-\mathbf{r}'|},$$

can be expanded in terms of $r = |\mathbf{r}|$:

$$\frac{1}{|\mathbf{r}-\mathbf{r}'|} = \exp(-\mathbf{r}'\cdot\nabla)\frac{1}{r} = \frac{1}{r} + \frac{\mathbf{r}'\cdot\mathbf{r}}{r^3} + \frac{3\left(\mathbf{r}'\cdot\mathbf{r}\right)^2 - r'^2 r^2}{2r^5} + \ldots$$

We define the multipole moments of the charge with respect to a common origin that is best chosen to be the center of mass of the charge distribution:

- Monopole

$$q = \int d^3r'\rho(\mathbf{r}')$$

- Dipole-moment

$$\mathbf{p} = \int d^3r'\mathbf{r}'\rho(\mathbf{r}')$$

- Quadrupole-moment

$$Q_{ij} = \int d^3r'\rho(\mathbf{r}')\left(3x_i'x_j' - r'^2\delta_{ij}\right)$$

- ...

This leads to the following expression for the asymptotic potential φ:

$$4\pi\varepsilon_0\varphi(\mathbf{r}) = \frac{q}{r} + \frac{\mathbf{r}\cdot\mathbf{p}}{r^3} + \frac{1}{2}\sum_{i,j} Q_{ij}\frac{x_i x_j}{r^5} + \ldots \tag{7.2}$$

The potential φ resulting from eq. (7.2) is evaluated only at the boundary ∂G of the device and employed as a boundary condition for the Poisson equation of the carrier density ρ

$$\begin{aligned} -\mathrm{div}\left(\varepsilon\left(\mathbf{r}\right)\mathrm{grad}\phi_{\mathrm{H}}\left(\mathbf{r}\right)\right) &= \rho\left(\mathbf{r}\right), \\ \phi_{\mathrm{H}}\left(\mathbf{r}\right)|_{\partial G} &\equiv \varphi\left(\mathbf{r}\right)|_{\partial G}. \end{aligned} \tag{7.3}$$

In this way the potential of a charge distribution with non-zero net charge can be calculated on a finite simulation domain. Any externally applied fields can easily be introduced via the boundary conditions into equation (7.3).

7.6 Hamiltonian

The choice of the appropriate Hamiltonian is the key part for the calculation of the electronic structure of a highly quantized system such as a quantum dot. In the literature various Hamiltonians are used that differ in complexity and in the degree of physical details taken into account. The most simple approach is to use an analytic Hamiltonian with a quantum well potential in growth direction combined with a parabolic confinement in the lateral direction to emulate the electronic structure of the quantum dot [105, 106, 107, 108]. This model completely neglects all details of the dot shape and composition and is therefore not suited for a realistic comparison to experiments. However, due to its simplicity and efficiency, particle-particle interactions can be taken into account with fewer approximations. A different approach is to start with a certain shape and composition of the quantum dot and calculate the single particle wave functions using effective mass (EMA) [109, 110, 111, 104], tight binding (TB) [87, 112] or pseudopotential (PP) [88, 113, 114] Hamiltonians. This is computationally more challenging but, regarding the complex shape and composition of realistic self assembled quantum dots, more appropriate to investigate the properties of real QD's. Within this class one can also distinguish between continuum (EMA) and atomistic (TB,PP) approaches, the latter are more sensitive to the underlying crystal structure but also numerically more demanding and, to our knowledge, non of the published works includes external electric fields.

Within next**nano**^{3} we use the EMA since it is appropriate for the grid that is used throughout the simulation to solve other equations such as Poisson and strain and it is also computationally less demanding than the atomistic models which have to be solved on supercomputers or large clusters. Within EMA the quantum mechanical states of the system are described in the frame-work of the envelope function approximation, which means that the wavefunction of the quantum-heterostructure is expanded close to the band-minimum in terms of the eigenfunctions of the bulk crystal.

For a periodic bulk crystal, the electronic structure is given by the Schrödinger equation

$$\begin{aligned} \mathbf{H}_0 \Psi &= E\Psi, \\ \mathbf{H}_0 &= -\frac{\hbar^2}{2m_0}\nabla^2 + V_0(\mathbf{r}), \end{aligned}$$

with the lattice periodic potential $V_0(\mathbf{r})$ and the free electron mass m_0. Due to the periodicity of the potential the eigenfunctions Ψ can be written as Bloch functions

$$\Psi_{n\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{kr}} u_{n\mathbf{k}}(\mathbf{r}) \tag{7.4}$$

with the lattice periodic part $u_{n\mathbf{k}}(\mathbf{r})$. The bands are indexed by n and the energetic minimum of each band will be denoted by $\mathbf{k}_0$. If the periodicity of the crystal is perturbed by an external potential $U(\mathbf{r})$ that is slowly varying on the scale of the lattice constant

$$\mathbf{H} = \mathbf{H}_0 + U(\mathbf{r})$$

we can expand the solutions of the new Schrödinger equation

$$\mathbf{H}\Psi = E\Psi \tag{7.5}$$

in terms of the Bloch functions defined in eq. (7.4) around the minimum $\mathbf{k}_0$

$$\Psi(\mathbf{r}) = \sum_n F_n(\mathbf{r}) \Psi_{n\mathbf{k}_0}(\mathbf{r}). \tag{7.6}$$

Now taking into account only the energetically close band directly, while including the influence of the remote bands perturbatively we obtain the following Schrödinger equation

$$\sum_m H_{m'm} F_m(\mathbf{r}) = E F_{m'}(\mathbf{r}) \tag{7.7}$$

where the dimension of the matrix $H_{m'm}$ corresponds to the number of directly coupled bands. The elements of the Hamiltonian contain the perturbative coupling to the remote bands, the spin-orbit coupling and the band deformation due to the strain. In nextnano this Hamiltonian is included for three different sets of states. The most simple one is the single-band case with only one band included in eq.(7.7) that can be written as,

$$-\frac{\hbar^2}{2m_0}\vec{\nabla}\left[\mathbf{M}^{-1}(\mathbf{r})\vec{\nabla}\right]\psi_\alpha(\mathbf{r}) + \left[E_c(\mathbf{r}) - |e|\,\phi(\mathbf{r})\right]\psi_\alpha(\mathbf{r}) = E_\alpha\psi_\alpha(\mathbf{r}) \tag{7.8}$$

with the inverse effective mass tensor $\mathbf{M}^{-1}(\mathbf{r})$ and the position dependent conduction band edge $E_c(\mathbf{r})$. This single band Hamiltonian can be applied to any band and valley including the valence bands. However since the two uppermost p-like valence bands termed heavy- and light-hole, are degenerate in the bulk system, they become strongly coupled in heterostructures and strained systems. Additionally, the dispersion is not isotropic but warped, which is not possible to model using the single band Schrödinger equation 7.8. These effects can be described by including the hole bands explicitly into eq. (7.7) which results in a 4×4 Hamiltonian as given in the appendix. Within next**nano**[3] also six- and eigth-band Hamiltonians are implemented that include the split-off hole band and the Γ conduction band, respectively. The latter coupling of conduction and valence bands is of great importance for the calculation of the optical properties, since electron and hole states result from the same Hamiltonian and therefore form a complete and orthogonal set of states. The details of the numerical implementation with a list of eigensolvers employed can again be found in the thesis of Stefan Hackenbuchner [92]. A special topic are the matching conditions at the material interfaces, since with an inappropriate symmetrization of the operators, spurious solutions may appear. Within this work, no such solutions have been observed, but the corresponding theory has been worked out in more detail by Till Andlauer [115].

Due to the zero dimensional confinement, the electronic structure of the quantum dots is very similar to the ones of atoms, which is the reason that quantum dots are also called artificial atoms. Figure 7.5 shows the wavefunctions of a lens-shaped quantum dot calculated using an eight band $\mathbf{k}\cdot\mathbf{p}$-Hamiltonian. The states can be classified analogously to atoms by the S, P and D shell. Interestingly the hole states are not completely symmetric, although the shape of the dot has no asymmetry. This effect is due to the strain, and the strain induced piezocharges which break the symmetry of the system.

7.7 Few particle states

The states obtained from the single- or multi-band Hamiltonian of eq. (7.7) are the single particle states that build up the basis for all further calculations. To obtain the energies of the few-particle states that are investigated in optical experiments on quantum dots, particle-particle interactions have to be taken into account.

The problem of few body states in quantum dots such as excitons, biexcitons and charged excitons has been addressed in a variety of theoretical publications, which can be classified by

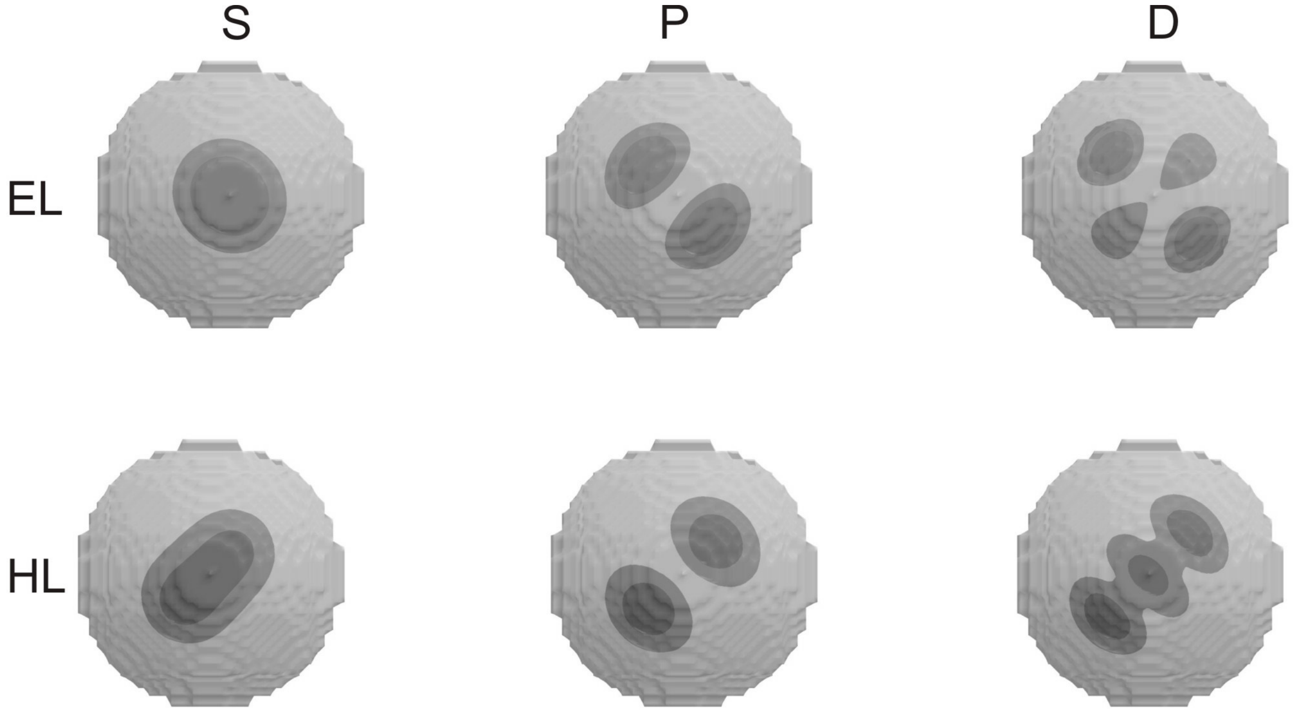


Figure 7.5: The electron (EL) and hole (HL) wavefunctions of a lens-shaped quantum dot using an eight band $\mathbf{k}\cdot\mathbf{p}$ Hamiltonian. The states can be classified in an atom-like scheme as S, P and D orbitals.

three main approaches. One popular method is the configuration interaction (CI) [109, 110, 113] where the many particle states are build of Slater determinants and the total energy is calculated by the diagonalization of the resulting Hamilton matrix. Since the convergence of the results is strongly dependent on the number of orbitals used for expansion of the states, the calculations can be very time consuming. Another approach is the quantum Monte Carlo (QMC) [113, 104] method which is considered to be the most sophisticated but also numerically most demanding.

In this work we choose the well developed method of the density functional theory (DFT) that was originally proposed by P. Hohenberg, W. Kohn and L. Sham [116, 117]. The method is employed within the local spin density approximation (LSDA) using the single particle states obtained from the effective mass Hamiltonian given by equation (7.8). A similar scheme was used previously by Martin *et al.* [111, 104] who compared the results to QMC calculations, showing that the difference in total energy is on the order of a few meV. Depending on the problem addressed in the calculation, this error can be negligible as we will show in the section of the charged quantum dots. In the following, a brief introduction to the DFT is given, highlighting the key points of the method with the parametrization of the LSDA potentials given in the appendix E.

The central assertion of the DFT method is [118], that for any interacting system a local single particle potential $V^{\mathrm{ni}}(\mathbf{r})$ exists, such that the exact ground state density $n^0(\mathbf{r})$ equals the ground state density of a non-interacting problem $n^{\mathrm{ni}}(\mathbf{r})$

$$n^{\mathrm{ni}}(\mathbf{r}) = n^0(\mathbf{r}).$$

This density

$$n^{\mathrm{ni}}(\mathbf{r}) = \sum_{i=1}^{N} |\psi_i(\mathbf{r})|^2 ,$$

is obtained from the eigenfunctions of the non-interacting Hamiltonian

$$\begin{aligned} H^{\mathrm{ni}}\psi_i &= \varepsilon_i \psi_i, \\ H^{\mathrm{ni}} &= T + V^{\mathrm{ni}}(\mathbf{r}) . \end{aligned}$$

Therefore with the exact knowledge of the auxiliary potential V^{ni}, the ground state energy and density of an interacting few-body state with N particles is obtained via the self-consistent solution of N single particle Schrödinger equations. In the case of a Fermion system V^{ni} is composed of two terms, the Hartree term

$$V_{\mathrm{Hartree}}(\mathbf{r}) = \int d\mathbf{r}' \frac{e^2}{|\mathbf{r}-\mathbf{r}'|} n(\mathbf{r}') ,$$

as well as the exchange and correlation term $V_{xc}(\mathbf{r})$ that is known exactly only for special cases. Therefore an approximation is needed in order to deal with general many particle systems, with the most prominent one being the so-called local density approximation (LDA). In this scheme the system is assumed to locally behave as a homogenous gas of charged carriers, for which exchange and correlation can be evaluated exactly as a function of the density.

Within the LDA and single-band EMA (eq. 7.8), the Kohn-Sham-equation reads

$$\left[-\frac{\hbar^2}{2m}\vec{\nabla}\left[\mathbf{M}^{-1}(\mathbf{r})\vec{\nabla}\right] + V_{\mathrm{ext}}(\mathbf{r}) + V_{\mathrm{Hartree}}(\mathbf{r}) + V_{\mathrm{xc}}^{\mathrm{LDA}}(\mathbf{r})\right]\psi_i(\mathbf{r}) = \varepsilon_i\psi_i(\mathbf{r}), \tag{7.9}$$

$$V_{\mathrm{Hartree}}(\mathbf{r}) = \int d\mathbf{r}' \frac{e^2}{|\mathbf{r}-\mathbf{r}'|} n(\mathbf{r}), \tag{7.10}$$

$$n(\mathbf{r}) = \sum_{i=1}^{N} |\psi_i(\mathbf{r})|^2 ,$$

with the LDA exchange-correlation potential

$$\begin{aligned} V_{\mathrm{xc}}^{\mathrm{LDA}}(\mathbf{r}) &= \frac{\partial}{\partial \bar{n}}\left(\bar{n}\varepsilon_{\mathrm{xc}}^{\mathrm{hom}}(\bar{n})\right)\Big|_{\bar{n}=n(\mathbf{r})} , \\ \varepsilon_{\mathrm{xc}}^{\mathrm{hom}} &= \varepsilon_{\mathrm{x}}^{\mathrm{hom}} + \varepsilon_{\mathrm{c}}^{\mathrm{hom}}, \\ \varepsilon_{\mathrm{x}}^{\mathrm{hom}} &= -\frac{e^2 3}{4}\left(\frac{3\bar{n}}{\pi}\right)^{1/3} , \\ \varepsilon_{\mathrm{c}}^{\mathrm{hom}} &= \text{parametrized.} \end{aligned}$$

In this work we use the parametrization of Perdew and Zunger [119] for the correlation, that is given in appendix E. The Kohn-Sham equations 7.9 have to be solved self-consistently for all particles until the density is converged. The total energy of the many-body state is finally obtained by

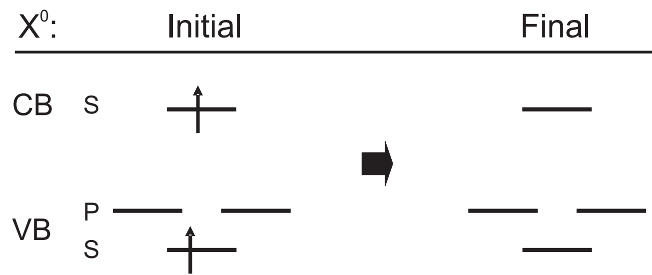
$$\begin{aligned} E_0 &= \sum_{i=1}^{N} \varepsilon_i - \frac{1}{2}\int d\mathbf{r}\ V_{\mathrm{Hartree}}(\mathbf{r})\, n(\mathbf{r}) + E_{\mathrm{xc}}(n) - \int d\mathbf{r}\ \mathbf{V}_{\mathrm{xc}}^{\mathrm{LDA}}(\mathbf{r})\, n(\mathbf{r}) , \qquad (7.11) \\ E_{\mathrm{xc}}(n) &= \int d\mathbf{r}\ \ \varepsilon_{\mathrm{xc}}^{\mathrm{hom}}(\bar{n})\big|_{\bar{n}=n(\mathbf{r})}\, n(\mathbf{r}) . \end{aligned}$$

With equations 7.9 and 7.11 everything is set up to calculate the ground state energy of few-particle states. It should be noted, that within the numerical implementation, the Hartree potential is not evaluated via eq. (7.10), but more efficiently by solving the Poisson equation given in equation (7.3). A further refinement to the LDA is obtained through the consideration of the spin-polarization

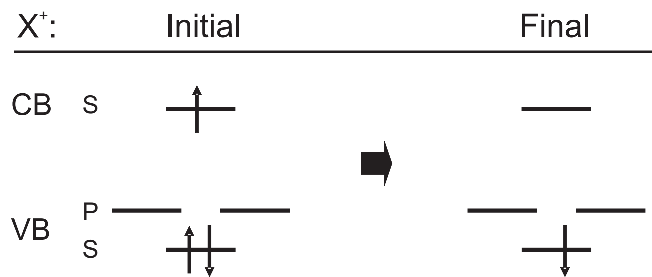
$$\zeta = \frac{n_\uparrow - n_\downarrow}{n}$$

as done in the local spin density approximation (LSDA) used by Martin *et al.* [104]. The details of this approach including the corresponding expressions for the exchange and correlation potential are given in appendix E. In this way the net spin polarization of the particles is taken into account which is important especially for few particle states and odd numbers of carriers.

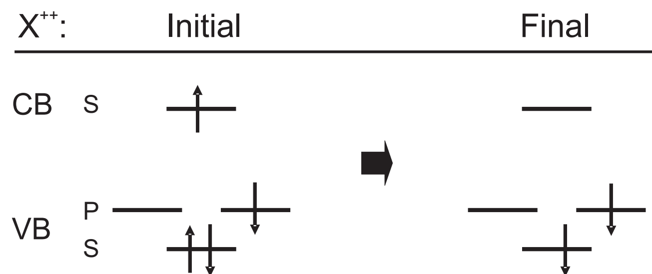
Using this method we are able to calculate the optical transition energies of neutral exciton X^0,



the positively charged X^+,



and the X^{++}



excitons by taking the difference of the total energies of the intial and final state. We note that for the X^{++} transition, there are actually two final states, a singlet with the two holes having opposite spins, and a triplet state with the spins aligned. Since the triplet state is threefold degenerate, the transition probability is enhanced and within an experiment this is the dominant transition observed. For the neutral exciton the DFT scheme is simplified taking

into account only the pure Hartree interaction, leading to the iterative solution of the electron Schrödinger equation with the potential of the hole and vice versa.

The DFT scheme presented in this section will be used in this work to calculate the Stark shift of neutral and charged excitons in self-assembled single quantum dots and quantum dot molecules. Certainly, the method is not applicable to all kinds of few-particle states. Especially biexcitons pose a problem, since their binding-energy is solely due to correlation effects, which are poorly described by LDA. The validity of the approach for the examples presented in the following is shown via the comparison to the experimental results.

Chapter 8

Positively charged excitons in QDs

The fully quantized electronic structure of self-assembled quantum dots (QDs) has recently led to many proposals for their use in next generation optoelectronic devices with true quantum mechanical functionality. Specific examples include the use of charge and spin excitations in isolated dots as quantum bits [120, 121] and the deterministic generation of single photons [122]. Each of these proposals relies fundamentally on the enhanced Coulomb interactions in such fully quantized systems that result in millielectronvolt shifts of the fundamental optical gap as the number, orbital configuration and total spin of the localized carriers changes [123]. Recently, it has been shown that the carrier occupancy of single dots can be electrically controlled using specially designed field-effect structures [124]. Fine structure in the emission spectra was identified as arising from charged exciton species from which the relative direct and exchange [124, 125] Coulomb energies for different few-body states have been extracted. Whilst these experiments enable measurement of the relative eigenenergies of different few-body states, they provide little direct information on the form of the few-body wavefunctions themselves and how they are modified as carriers are added to the dot. The nature of the exciton wavefunction determines key parameters such as the coupling strength to electromagnetic radiation and phonons and, ultimately, the decoherence rate for excitons. Therefore, a soundly based understanding of the influence of such Coulomb driven modifications of the few-body wavefunction in QDs is of central importance for their future implementation in novel quantum devices.

In this section, which has been published in [89], we investigate the form of the excitonic charge-distribution for different few-body states in individual QDs. Our approach, based on Stark effect perturbation spectroscopy, provides direct access to the excitonic permanent dipole (p_z) and polarizability (β). These quantities are intrinsic properties of the wavefunction and reflect both the nature of the zero dimensional confinement potential and the role of Coulomb interactions in the few-body state. Our results reveal a systematic reduction of p_z as excess holes are added controllably to the dot, from neutral ($X^0 = 1e + 1h$) to positively charged excitons ($X^+ = 1e+2h$ and $X^{++} = 1e+3h$). Calculations of the electro-optical response of the few-body wavefunctions provide good quantitative agreement with experiment, demonstrating that the reduction of the dipole arises from a Coulomb driven redistribution of the exciton wavefunction in the dot upon charging. Surprisingly, we find that the dominant effect of hole charging is a strong lateral expansion (contraction) of the hole (electron) component of the wavefunction. This controls the reduction of p_z due to a modification of the electron-hole overlap and, consequently, the balance between $e-h$ attractive and $h-h$ repulsive contributions to the total Coulomb energy.

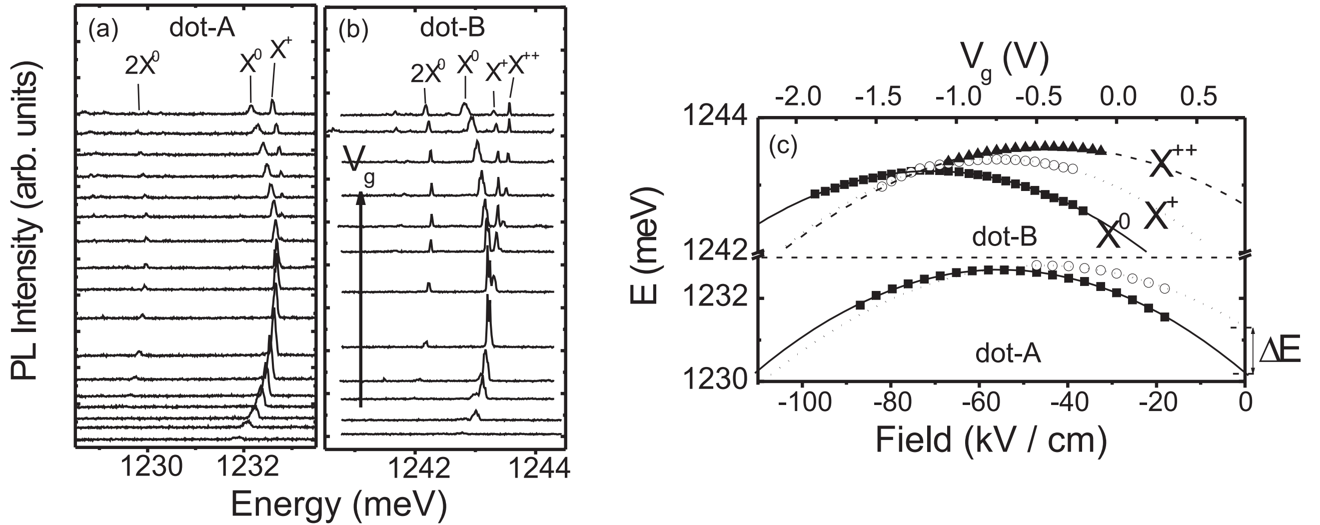


Figure 8.1: Single dot PL spectra obtained from dot A (a) and dot B (b) as a function of V_g from +1.9 V (lower trace) to −0.3 V (upper trace). Neutral (X^0) and charged (X^+, X^{++}) single excitons are observed. (c) Peak positions of the X^0, X^+ and X^{++} transitions as a function of V_g / F_z. Parabolic fits to the data are also shown from which p_z and β are extracted.

8.1 Experimental results

We begin by describing the experimental samples and the single dot spectroscopy results before presenting the calculations and discussing the form of the few-particle wavefunctions. The devices investigated consisted of a single layer of low density ($\sim 5\ \mu\text{m}^{-2}$) $\text{In}_{0.5}\text{Ga}_{0.5}\text{As}$ self-assembled QDs embedded within the intrinsic region of a p-type GaAs Schottky photodiode [126]. Such structures allow vertical electric fields up to $F_z \sim -150$ kV/cm to be applied [127] enabling the study of both Stark effect phenomena and controlled hole charging of the dot from the p-contact.

Photoluminescence (PL) spectra were recorded at T=5K under weak optical excitation such that the average QD occupancy was in the single exciton limit [128]. Typical PL results obtained from two representative dots (dot-A and dot-B) are presented in figs 1a and 1b respectively as a function of axial electric field F_z (V_g). For $|F_z| \geq 90$ kV/cm ($V_g \geq +1.75$ V) no PL-signal is observed due to field ionization of excitons, whilst PL arising from the charge neutral exciton (X^0) emerges for lower field [126]. The peak position of X^0 as a function of V_g is plotted in fig. 8.1c for both dots (filled squares). As $|F_z|$ reduces, X^0 shifts due to the quantum confined Stark effect, the shift being well described by $E(F) = E_0 - p_z F_z + \beta F_z^2$ where p_z and β are the permanent excitonic dipole moment and polarizability of the excitonic state respectively [94]. The full lines in fig. 8.1a show the result of such a parabolic fit to the X^0 data from which we extract values of $p_z = +0.72 \pm 0.07$ nm and $+0.89 \pm 0.02$ nm for dots A and B respectively. For all dots investigated, the positive sign of p_z implies that the electron component of the wavefunction is located below the hole in the dot at $F_z = 0$, signifying In-enrichment towards the dot apex [94, 129].

With decreasing reverse bias, excess holes tunnel into the dot [124] and one or more additional PL peaks (labeled X^+ and X^{++} in fig. 8.1) emerge on the *high*-energy side of X^0 arising from positively charged excitons. The bias range over which these features emerge, together

with their splitting from X^0 at $F_z = 0$, enable us to identify them as arising from single (X^+) and doubly (X^{++}) charged states. Different charge states are observed simultaneously due to the time integrated nature of the spectra and the comparatively weak tunneling coupling for holes between the dots and the p-contact [130]. By extrapolating the observed field dependencies for X^0, X^+ and X^{++} to $F_z = 0$, we measure directly the relative eigenenergies of the different charged exciton states relative to X^0 (ΔE). From a survey of more than 20 dots we measure $\Delta E = +0.7 \pm 0.3$ meV and $+2 \pm 0.1$ meV for X^+ and X^{++}. These results contrast with the case for negatively charged excitons for which large *red*shifts (up to $\Delta E \sim -5$ meV) are generally observed [124]. This difference arises due to the smaller spatial extent of the hole wavefunction and resulting positive energy contribution [131, 124].

Also plotted in fig. 8.1c are the peak positions of the various exciton states together with parabolic fits from which the resulting values of p_z and β are deduced. For X^0, figure 8.2 shows the values of p_z (fig. 8.2a) and β (fig. 8.2b) plotted versus E_0 for more than twenty individual dots. In contrast with ref. [132] β was found to be uncorrelated with E_0 and p_z, remaining approximately constant over the entire inhomogeneous linewidth with an average value of $\overline{\beta} = -34 \pm 9$ nm^2/V. For a particular QD shape and Indium composition profile, β is most sensitive to the dot height (h)[133]: this observation enables us to fix h in our calculations presented below. In contrast, p_z and E_0 (fig. 8.2a) are clearly correlated, the permanent dipole reducing from $\sim +0.8$ nm to $\sim +0.2$ nm as E_0 increases from $\sim 1230 - 1310$ meV. For a particular QD shape and height, p_z is determined by the In-composition gradient along the QD growth axis, negative dipoles (electron above hole) always being predicted for dots with a homogeneous composition [133, 94]. When taken together, these observations suggest that higher energy dots possess both a weaker composition gradient (smaller p_z) whilst their average In-content also reduces, tending to increase E_0. The absence of any $\beta - E_0$ correlation demonstrates that the effective height of the dots remains approximately constant, suggesting that the absolute In-content and composition profile are not independent variables but that a strong composition gradient also indicates a high In-content in the dot and vice versa.

Figure 8.2c shows the percentage change of the exciton dipole ($\Delta p_z^{n+} = 100 \times (p_z^{n+} - p_z^0)/p_z^0$) for the charged exciton species relative to X^0. The most important observation is that Δp_z^{n+} is negative for all dots investigated, implying that the center of gravity of the electron and hole components of the wavefunction move closer together as holes are added to the dot. From fig. 8.2c, the mean reduction of p_z upon single hole charging was measured to be $\Delta p_z^+/e = -17 \pm 5$ %, whereas for the doubly charged exciton (X^{++}) we measured $\Delta p_z^{++} \sim -40$ % relative to X^0.

8.2 Theoretical results

In order to explain our experimental observations, we performed calculations of the electronic structure and few-particle states in our dots. We use a one band effective mass Hamiltonian to calculate the single particle states including a full treatment of strain and piezoelectric effects [134]. The model QDs are square based truncated pyramids aligned along the [100] direction with a height of $h = 5$ nm needed to reproduce the experimentally observed X^0 polarizability, which depends mostly on the dot height [133], and further supported by cross-sectional TEM microscopy results (inset - Fig 8.3a) that revealed dot heights between $h \sim 4$–6 nm. The dot width (w) was chosen as a free parameter and varied in the range $w = 14$–26 nm. Three different composition profiles were investigated with linear [94], inverted pyramidal [129] and

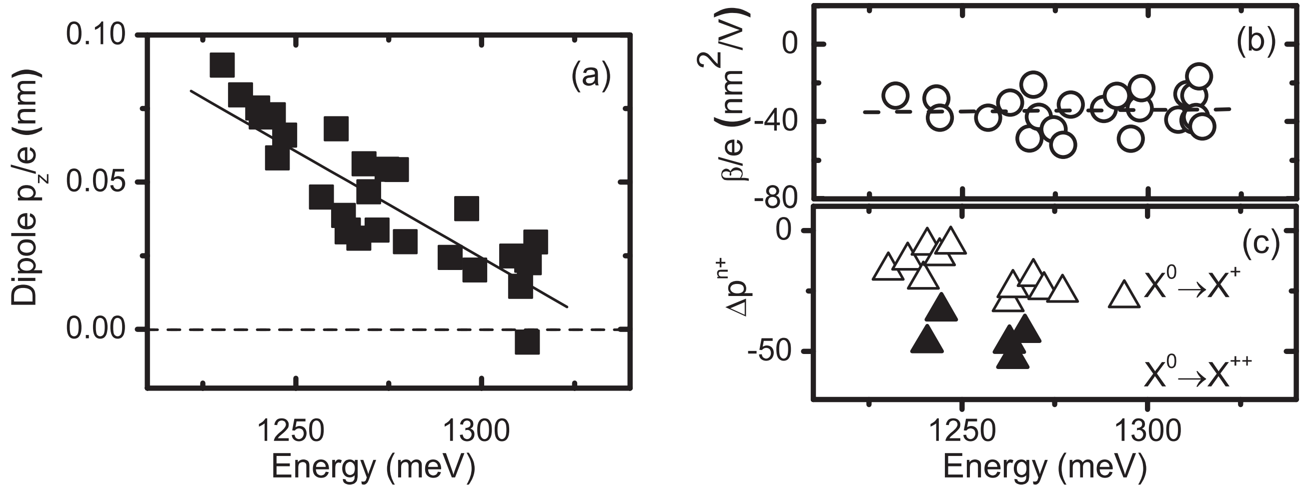


Figure 8.2: Summary of p_z (a) and β (b) values measured from many single dots for X_0 plotted as a function of E_0. (c) Percentage change of exciton dipole upon adding one ($X^0 \to X^+$ - open symbols) or two ($X^0 \to X^{++}$ filled symbols) holes.

trumpet-like [96] In-distributions. For fixed w, the lateral confinement potential in the dot becomes progressively stronger when moving from the linear to the inverted pyramidal and trumpet-like profiles.

Few body interactions were treated by self-consistent solution of the Kohn-Sham equations accounting for the exchange and correlation Coulomb potential within the local spin density approximation. For QDs, this approach tends to overestimate the correlation and exchange contributions to the total Coulomb energy by a few meV. This is comparable to the size of the Coulomb charging shifts (ΔE) and it is, therefore, unrealistic to attempt to fit the absolute transition energies of the various charge states at zero field. Instead, our approach is to investigate the response of the few body states to static electric fields, i.e. to fit E_0, p_z and β for X^0 and then to deduce the change of p_z upon adding one (Δp_z^+) or two (Δp_z^{++}) holes to the dot. The results are plotted in figs. 8.3a and 8.3b, the shaded regions showing the experimentally deduced values and their associated error for a dot with a ground state energy of $E_0 = 1250$ meV.

We first examine the change of the permanent dipole due to the addition of a single excess hole ($X^0 \to X^+$) as presented in fig. 8.3a. Similar results were obtained for the three lateral composition profiles, the dot width being the dominant parameter. Firstly, our calculations reproduce quite generally the experimentally observed reduction of p_z upon charging over the entire range of w investigated. The magnitude of $\Delta p_z^+/e$ shows two distinct regimes. For $w \geq 20$ nm $\Delta p_z^+/e$ saturates around -0.1 nm in excellent quantitative agreement with the experimentally observed value of -0.14 ± 0.03 nm (shaded area - fig. 8.3a). In contrast, for $w \leq 20$ nm the reduction of the dipole becomes progressively larger, reaching $\Delta p_z^+/e \sim -0.5$ nm for $w \sim 14$ nm, more than $3\times$ larger than the experimental result. Thus, the calculations predict that the QDs investigated here have $w \geq 20$ nm. This result is in good quantitative agreement with the lateral size of our dots deduced from TEM microscopy of ~ 23 nm (inset - fig. 8.3a) providing further support for the validity of our theoretical approach.

In an effort to understand the physical origin of the w dependence of our $\Delta p_z^+/e$ calculations we examined the form of the few-body wavefunctions for X^0 and X^+. Figure 8.4 shows

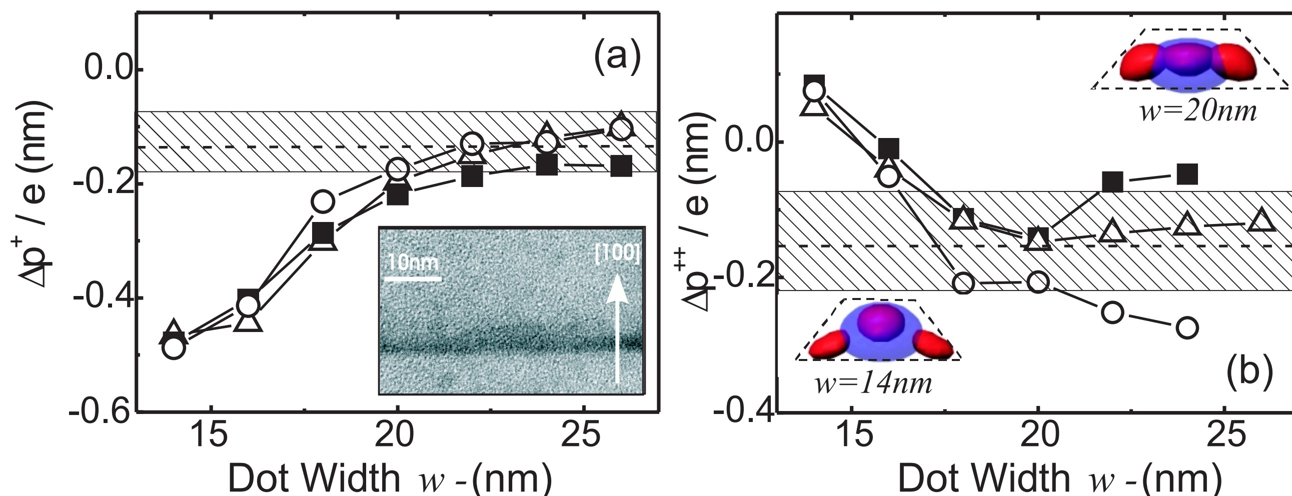


Figure 8.3: Calculated values of Δp_z^+ (a) and Δp_z^{++} (b) as a function of dot width for linear (squares), inverted pyramidal (circles) and trumpet (triangles) In-profiles. The shaded regions show the mean experimental values

calculated 50 % constant probability isosurfaces of the electron (blue) and hole (red) components of the wavefunction for two model dots with $w = 14$ and 20 nm representing the two size regimes discussed above. For the $w = 20$ nm dot, injecting an additional hole results in a marked (~ 2 nm) expansion of the 2-hole wavefunction in the *plane* of the dot and a commensurate contraction of the 1-electron wavefunction. This redistribution is driven by Coulomb interactions and reflects the tendency of the system to minimize the repulsive h–h interaction whilst maximizing the attractive e–h contribution. The potential for Coulomb limited lateral redistribution of the e and h wavefunctions upon charging in larger dots explains the origin of the saturation behavior of $\Delta p_z^+/e$ for $w \geq 20$ nm. The comparatively small reduction of the dipole (~ -0.1 nm) then arises due to enhanced electron-hole attraction in the 2h-1e system which pulls the centers of gravity of the $2h$ and $1e$ wavefunctions together. In contrast, for the $w \sim 14$ nm dot the lateral redistribution of the wavefunction is partially suppressed by the stronger lateral quantization and the hole wavefunction instead expands vertically in the QD upon charging. These effects lead to larger values of $\Delta p^+/e$ as is clearly visible in the cross sectional images of the wavefunctions presented in fig. 8.4.

The dependence of Δp^{++} on w is calculated to be opposite to that for Δp^+ as discussed above. The X^{++} state is particularly revealing since the four-particle ($3h + 1e$) wavefunction involves the first excited hole p-state. As for the case of X^+ discussed above, the calculations for the three composition profiles are similar for $w \leq 20$ nm, diverging more strongly for larger w. One of the most interesting features of the calculated curve is the prediction of an inversion of the sign of the dipole shift for small dots, in strong contrast with the behavior and discussion of X^+ presented above. A qualitative picture of the underlying physics can be obtained by examining the constant energy surfaces of the s- and p-orbital components of the heavy hole wave function as shown by the insets on fig. 8.3b. For larger dots, the three holes are coplanar and localized toward the top of the dot. This tends to concentrate the Coulomb e–h attraction and reduce the dipole as discussed for X^+ above. In contrast, for narrow dots the excited hole state shifts toward the base of the dot due to the tapering geometry. This partially balances the net e–h attraction and results in positive values of Δp_z^{++} for $w \leq 16$ nm

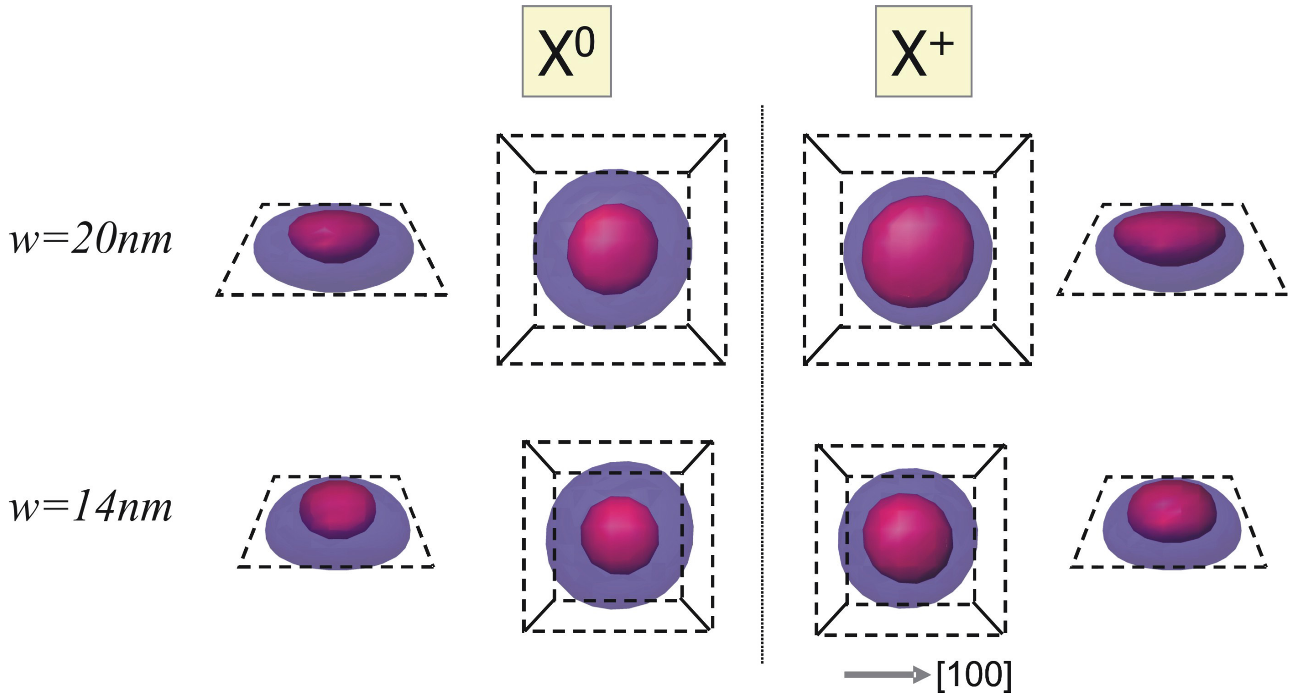


Figure 8.4: Electron and heavy hole ~ 50 % constant probability isosurfaces for X^0 and X^+ for QDs having a width of $w = 14$ and 20 nm. As described in the text, the lateral redistribution of the electron and hole components of the wavefunction can clearly be seen.

and negative values when the three holes are coplanar, such as is the case for larger dots. Again, the best agreement with experiment is obtained for $w \geq 20$ nm in excellent accord with the findings for Δp_z^+ discussed above and our TEM-microscopy results. The inverted pyramidal and trumpet like In-profiles provide the best quantitative agreement, the linear profile consistently underestimating the magnitude of Δp_z^{++}. This provides direct spectroscopic evidence for recent high-resolution structural microscopy results [129, 136] that have indicated both vertical and lateral variations of the In-content throughout the body of the dots grown by the Stranski-Krastanow method.

8.3 Conclusion

In summary, from analysis of experimental data obtained by Stark effect perturbation spectroscopy with the help of a detailed theoretical model we were able to predict the form of the charge-distribution for different few-particle states in individual InGaAs quantum dots. Our measurements reveal a very pronounced redistribution of the electron and hole components of the wavefunction as additional holes are sequentially added to the system. The magnitude of these effects, which are driven by Coulomb interactions, were found to be in good quantitative agreement with calculations of the few-body wavefunctions for several distinct few-body states (X^0, X^+ and X^{++}).

Chapter 9

Quantum dot molecules

Over the last years coherent manipulation of excitonic states in ensembles and single QDs has been demonstrated by various groups worldwide backing the proposed implementation of quantum logic using excitons as quantum bits [137, 138, 139, 140]. Since this implementation for single QDs is limited to two qubits, efforts have been made to scale this qubit system from single QD ("artificial atoms") to "artificial molecules" i.e. electronically coupled QDs [141, 142, 143]. Another application lies in the generation of polarization-entangled photons using the biexciton state in coupled QDs [144]. The most promising approach suggested provide stacked self-assembled QDs as shown in fig. 9.1, that are grown in the strain-driven Stranski-Krastanov growth mode. Here, first a seeding layer of QDs in grown. Capped by a thin buffer layer the strain profile of the seeding layer is preserved at the sample surface which leads to a preferable nucleation sites for islands in a second layer. Self-alignment can be achieved even for layer spacings where no electronic coupling effects are expected [145, 146]. Recent experiments demonstrated dephasing times of excitons in quantum dot molecules (QDMs) which are still sufficiently long to apply enough quantum logic operations needed for basic quantum algorithms [147]. First experiments on the nature of the excitonic structure of QDMs have been reported which have been focusing on single QDMs. So far no demonstration of a controlled variation and manipulation of the coupling strength in QDMs have been reported [148, 149]. In comparison to single layers of QDs [94] a strongly anomalous shift due to the Quantum Confined Stark Effect (QCSE) is predicted for QDMs [150].

In this chapter we provide a theoretical investigation of the electronic and optical properties of coupled quantum dots in external electric fields including the effects of strain and Coulomb coupling. The predicted anti-crossing of the two lowest exciton states is compared to recent experimental results, demonstrating the possibility of an externally controlled coupling of the excitonic states.

9.1 Theory

In the following section we will analyze step by step the different aspects of the physics of a quantum dot molecule, as there are quantum mechanical coupling, coupling through strain, Coulomb interaction and the effect of external fields.

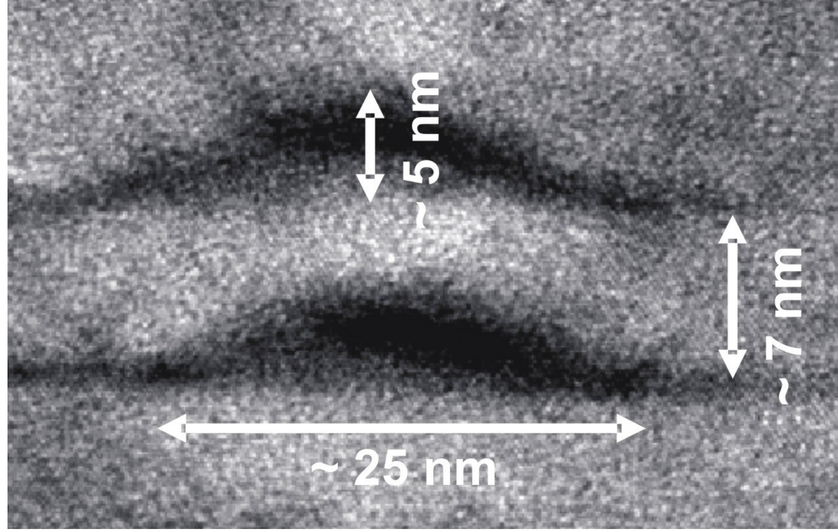


Figure 9.1: STM picture of two coupled quantum dots. The dots are almost equally sized with a height of 4–5 nm and a width of 20–25 nm. The distance between the two wetting layers is 7 nm.

9.1.1 Coupled quantum systems

A first step in the description of quantum dot molecules is to take a look back on basic quantum mechanics. A coupled quantum system such as the quantum dot molecule investigated in this section can be described in a simplified way by a coupled two state system. Consider the uncoupled eigensystem

$$H^0 \left|\psi_i^0\right\rangle = E_i^0 \left|\psi_i^0\right\rangle \quad i = 1, 2 \tag{9.1}$$

with the Hamiltonian H^0 and the corresponding eigenstates $\left|\psi_i^0\right\rangle$ and eigenenergies E_i^0. Introducing a potential V that couples the states $\left|\psi_1^0\right\rangle$ and $\left|\psi_2^0\right\rangle$, we can write the resulting coupled eigensystem in the basis of the eigenstates of the decoupled system

$$\begin{aligned} H\left|\psi_i\right\rangle &= E_i \left|\psi_i\right\rangle \\ \left|\psi_i\right\rangle &= \alpha_i \left|\psi_1^0\right\rangle + \beta_i \left|\psi_2^0\right\rangle . \end{aligned}$$

This results in a Hamiltonian given by the 2×2 matrix

$$\mathbf{H} = \begin{pmatrix} E_1^0 & \frac{1}{2}V_{12} \\ \frac{1}{2}V_{21} & E_2^0 \end{pmatrix}, \tag{9.2}$$

with the coupling elements $V_{21}^* = V_{12} = \left\langle\psi_1^0\right| V \left|\psi_2^0\right\rangle$. The new eigenvalues of this coupled Hamiltonian can be expressed in terms of the energy difference between the decoupled eigenenergies $\Delta = E_2^0 - E_1^0$

$$E_{1,2} = E_1^0 + \frac{1}{2}\Delta \pm \frac{1}{2}\sqrt{\Delta^2 + V_{12}^2}. \tag{9.3}$$

For a degenerate system ($\Delta = 0$), and attractive potential one obtains

$$E_1 - E_2 = |V_{12}|$$

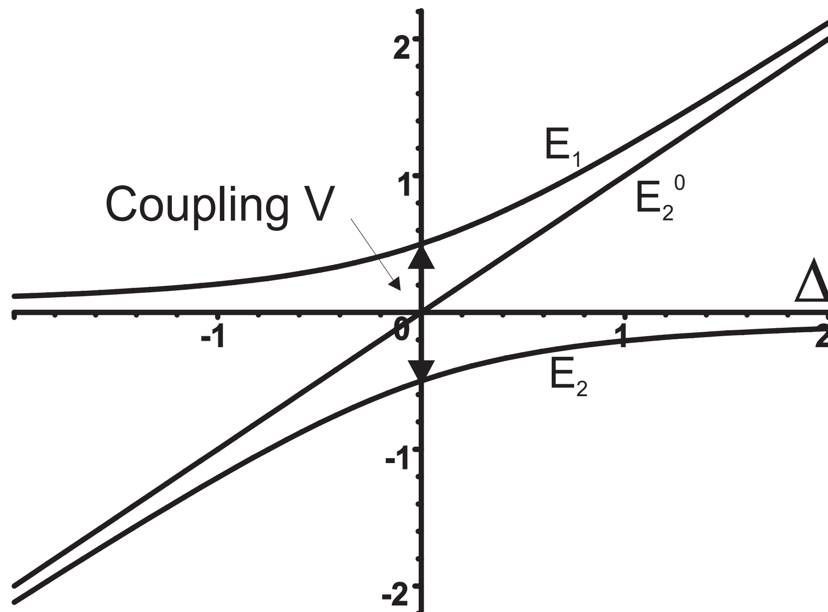


Figure 9.2: The eigenvalues of the coupled system E_1 and E_2 are plotted as a function of the energy difference of the decoupled eigenvalues $\Delta = E_2^0 - E_1^0$ where we take $E_1^0 = 0$ and the coupling $|V_{12}| = 1$.

with the corresponding eigenstates

$$|\psi_{1,2}\rangle = \frac{1}{\sqrt{2}} \left(|\psi_1^0\rangle \mp |\psi_2^0\rangle \right)$$

that are also termed bonding (+) and anti-bonding (−) state. Figure 9.2 shows the evolution of the eigenvalues of the coupled system as a function of the energy difference Δ of the decoupled eigenvalues. The graph shows the characteristic anti-crossing of coupled systems where the energy difference between the eigenvalues at zero splitting $\Delta = 0$ corresponds to the coupling energy $|V_{12}|$. Asymptotically, the upper eigenvalue of the coupled system E_1 converges to the decoupled eigenvalues E_1^0 for negative splitting $\Delta < 0$ and to E_2^0 for positive splitting. In this example the coupling was taken $V = 1$ and the eigenenergy $E_1^0 = 0$.

This simple example should illustrate qualitatively what kind of behavior is expected in a real system of coupled quantum dots, although in an optical experiment one can only observe many-particle states such as electron-hole pairs and no single particle states. However, as will be shown in the following sections, only the electrons with the lighter mass couple significantly, while the heavy holes remain localized in the individual dots, such that the simple picture remains valid.

To assess the quantum mechanical coupling without the influence of an inhomogeneous strain field, or Coulomb coupling as present in the 3D dot model, we performed 1D calculations of the eigenstates of two coupled InGaAs quantum wells with an Indium concentration of 30 % equal to the average content of the InGaAs quantum dots. The two wells are 5 nm thick each and separated by a GaAs barrier of thickness d as shown in the insets of fig. 9.3. The Indium concentration in the wells differs by 1 % in order to lift the degeneracy, leading to a splitting of the decoupled eigenstates Δ. Part a) of fig. 9.3 shows the two lowest eigenenergies of the electron states as a function of the barrier thickness d. For large barrier thicknesses above 8 nm,

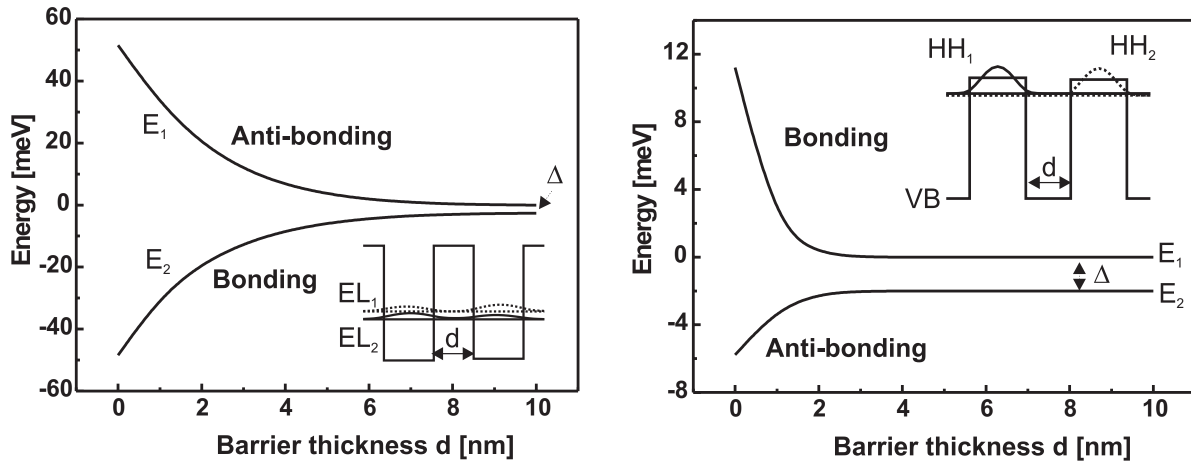


Figure 9.3: Splitting of the electron and hole states into bonding and antibonding states as a function of the barrier thickness between two coupled quantum wells.

the coupling energy is smaller than the initial splitting of the decoupled states $\Delta = 2.5$ meV and cannot be determined accurately from this graph. For smaller barrier thickness the coupling is increased and the states split into bonding and anti-bonding state with a maximum coupling energy of 100 meV for zero barrier thickness. The inset shows the two quantum wells with a barrier thickness of $d = 4$ nm including the bonding (EL_2) and anti-bonding (EL_1) states. The eigenenergies of the two highest heavy-hole states are plotted in part b) of fig. 9.3, again as a function of the barrier thickness d. Due to the heavy mass of the holes ($m^*_{HH} = 0.5$ m_0) no coupling is present until the barrier thickness drops below 2 nm. Correspondingly, the wavefunctions HH_1 and HH_2 shown in the inset are completely localized within each quantum well. This picture shows that for quantum dot molecules with an interdot barrier thicker than 2 nm, the coupling is solely due to the electrons, while the holes remain localized in the individual dots.

9.1.2 Strain

To keep the number of free parameters for the three-dimensional model of the coupled quantum dot system at a minimum, we assume that the two dots have identical shape (truncated pyramid), size ($h = 5$ nm, $w = 22$ nm) and alloy profile (trumpet profile). These parameters have been chosen according to the results of the investigation of the charged dots in the previous section, that have been grown under similar conditions. Figure 9.4 shows the simulated double dot structure including the substrate, the wetting layers (WL), a GaAs spacer of varying thickness, and the GaAs cap layer. Even though the quantum dots are identical, the relaxation of the strain in the entire structure leads to a different local strain-field in each single dot. Correspondingly the valence band-edge is shifted via the deformation potentials, thus lifting the degeneracy of the two dots. To illustrate this effect a slice of the heavy-hole valence band through the center of the quantum dots is depicted in fig. 9.5 for three quantum dot separations (distance from the tip of the lower dot to the wetting layer of the upper dot). Whereas the potential profile of the two dots is nearly identical for a large separation of 10 nm (dotted line), the influence of the strain induces a difference between the maxima of the valence band edges

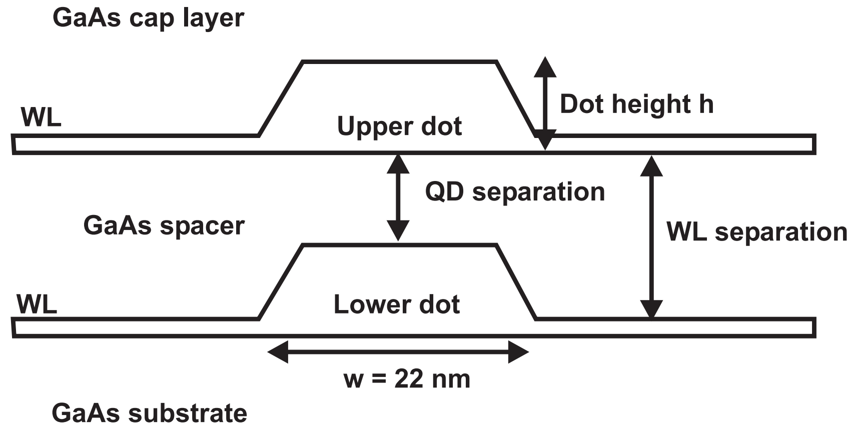


Figure 9.4: Simulated double dot structure with the two InGaAs dots of 5 nm height (h) and 22 nm base width (w).

of the two dots of 10 to 20 meV for the smallest separation of 2 nm (solid line). Additionally the potential in the barrier region is changed significantly which has a strong influence on the coupling. These findings show that it is extremely important to take into account the fully relaxed strain in the structure in order obtain physically meaningful results.

9.1.3 Single particle states

The single particle states have been calculated employing the one-band effective mass Hamiltonian of equation (7.8). Figure 9.6 shows the evolution of the single particle eigenenergies as a function of the quantum dot separation for zero applied field. The two lowest electron states, depicted in fig. 9.6a), are split by roughly 3 meV for the QD separations larger than 7 nm, which is caused by the difference of the hydrostatic strain in the two dots. With decreasing inter-dot distance, the localized states evolve into delocalized bonding (solid line) and anti-bonding (dashed line) states with a splitting of more than 20 meV for the case of 2 nm QD-separation. In contrast to the 1D calculation presented in part a) of fig. 9.3, the splitting is not symmetric with respect to the energies of the decoupled states, but exhibits an underlying trend towards higher energies, which is due to the increased overall hydrostatic strain in a system of closely coupled dots. The picture is less clear in the case of the heavy hole states depicted in part b) of fig. 9.6. For separations larger than 5 nm the eigenenergies are split by less than 1 meV and follow a trend to higher energies for larger separations that can again be attributed to the hydrostatic strain that increases the energy-gap. This trend is broken for separations below 5 nm which might be caused by the influence of the increased off-diagonal and shear-components of the strain that lead to a distortion of the confining potential as shown in fig. 9.5. Since the splitting of the heavy hole states does not increase monotonously with smaller barrier width we conclude that it is solely due to strain effects and that the quantum mechanical coupling of the heavy holes is negligible as already suggested by the 1D results shown in part b) of figure 9.3.

As an example of the corresponding wavefunctions the 30 % isosurfaces of the bonding and anti-bonding electron states are shown in fig. 9.7 for a dot separation of 2 nm.

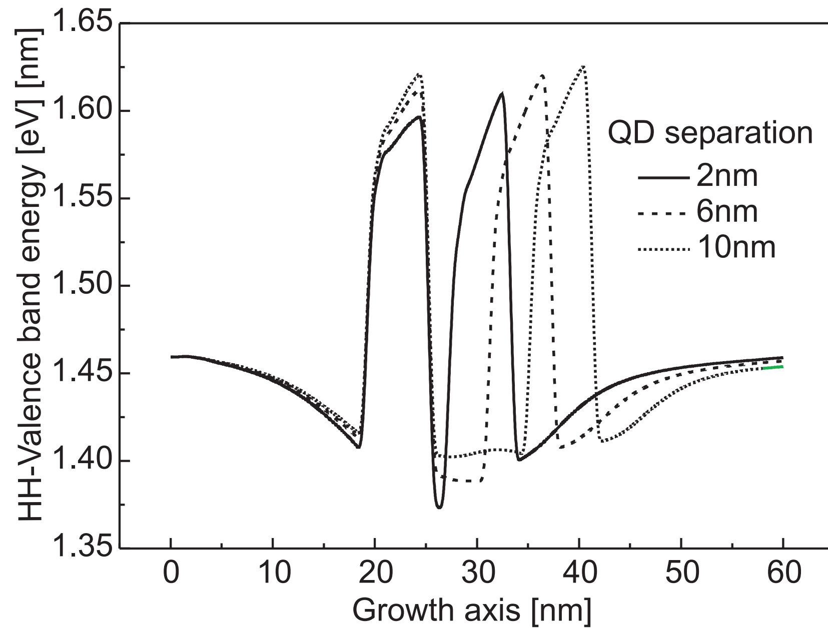


Figure 9.5: The heavy-hole valence band is shown along a slice through the center of the quantum dots for different inter-dot separations.

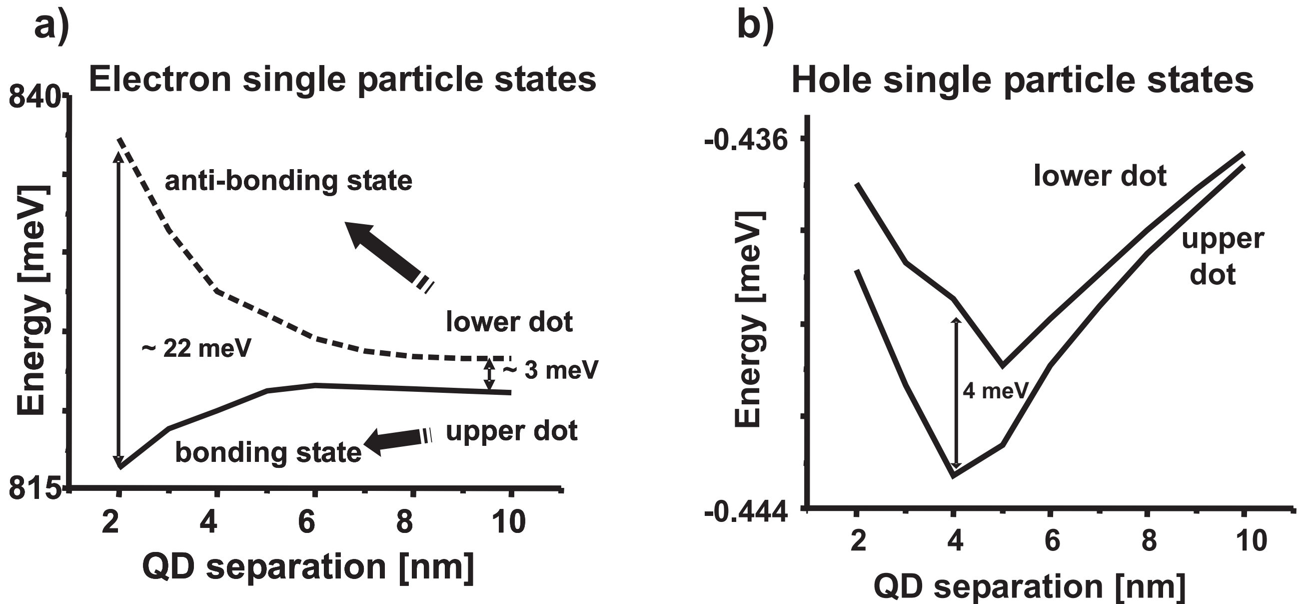


Figure 9.6: The first two single particle states of the electrons a) and holes b) are shown as a function of the QD separation.

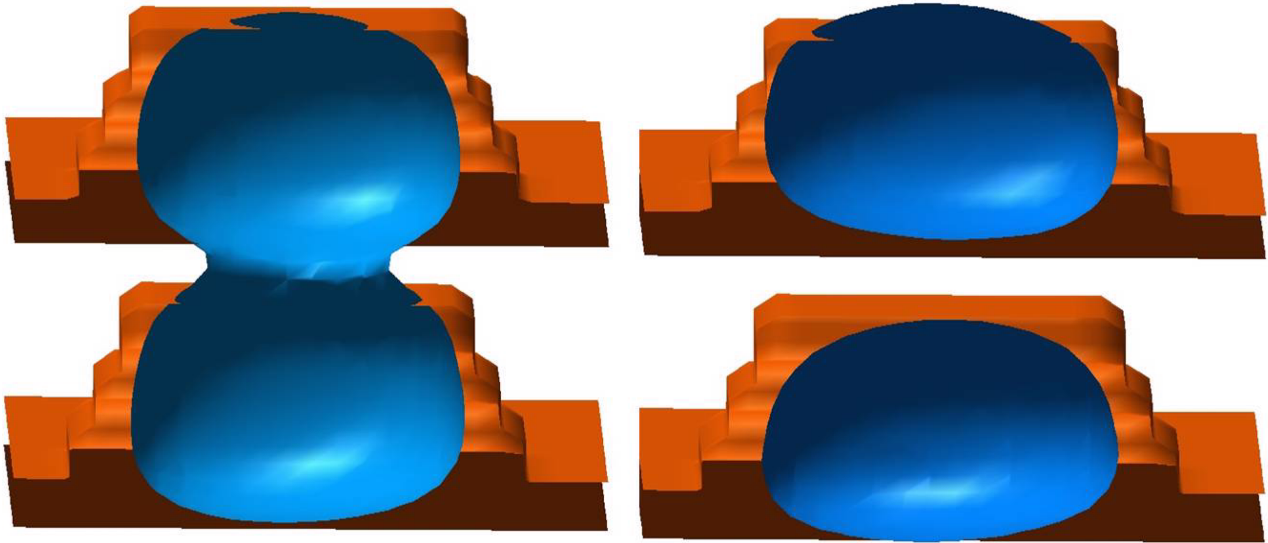


Figure 9.7: The bonding (left part) and anti-bonding (right part) electron states for a quantum dot separation of 2 nm.

9.1.4 Excitons in coupled quantum dots

So far we considered the single particle states including the effects of quantum mechanical coupling and strain. Since in an experiment only electron-hole pairs can be observed, we have to take into account additionally the Coulomb interaction between the carriers, which is done according to the mean-field scheme presented in section 7.7. With the two hole and two electron states one obtains four different configurations for the excitons. For large dot separations where no significant coupling is present we can classify the excitons into direct excitons with electron and hole in the same dot and indirect excitons where electron and hole are located in different dots as shown in figure 9.8. The energies of the two direct excitons (labeled by $|c\rangle$ and $|d\rangle$) are roughly 20 meV lower than the energy of the two indirect excitons (labeled by $|a\rangle$ and $|b\rangle$) due to the Coulomb binding energy. Each pair of lines is additionally split by a few meV because of the asymmetry in the potentials of the two dots induced by the strain. For smaller separations, our calculations predict a blue shift of the direct and indirect excitons in agreement with calculations of Bester et al. [143], a trend which cannot be obtained using less realistic models [148, 151, 152, 153] which neglect strain effects and assume identical tunneling rates for electrons and holes. In the regime of strong coupling for separations below 5 nm, the states can be characterized by the bonding and anti-bonding electron state. The sum of quantum mechanical coupling energy and the Coulomb interaction, leads to a splitting of the states by roughly 30 meV for a QD separation of 2 nm.

It is important to note that, since the Coulomb and quantum coupling energies as well as the potential shift due to the hydrostatic strain are all of the same order, they have to be taken into account properly in the calculation in order to yield quantitative predictions of the observed transition energies.

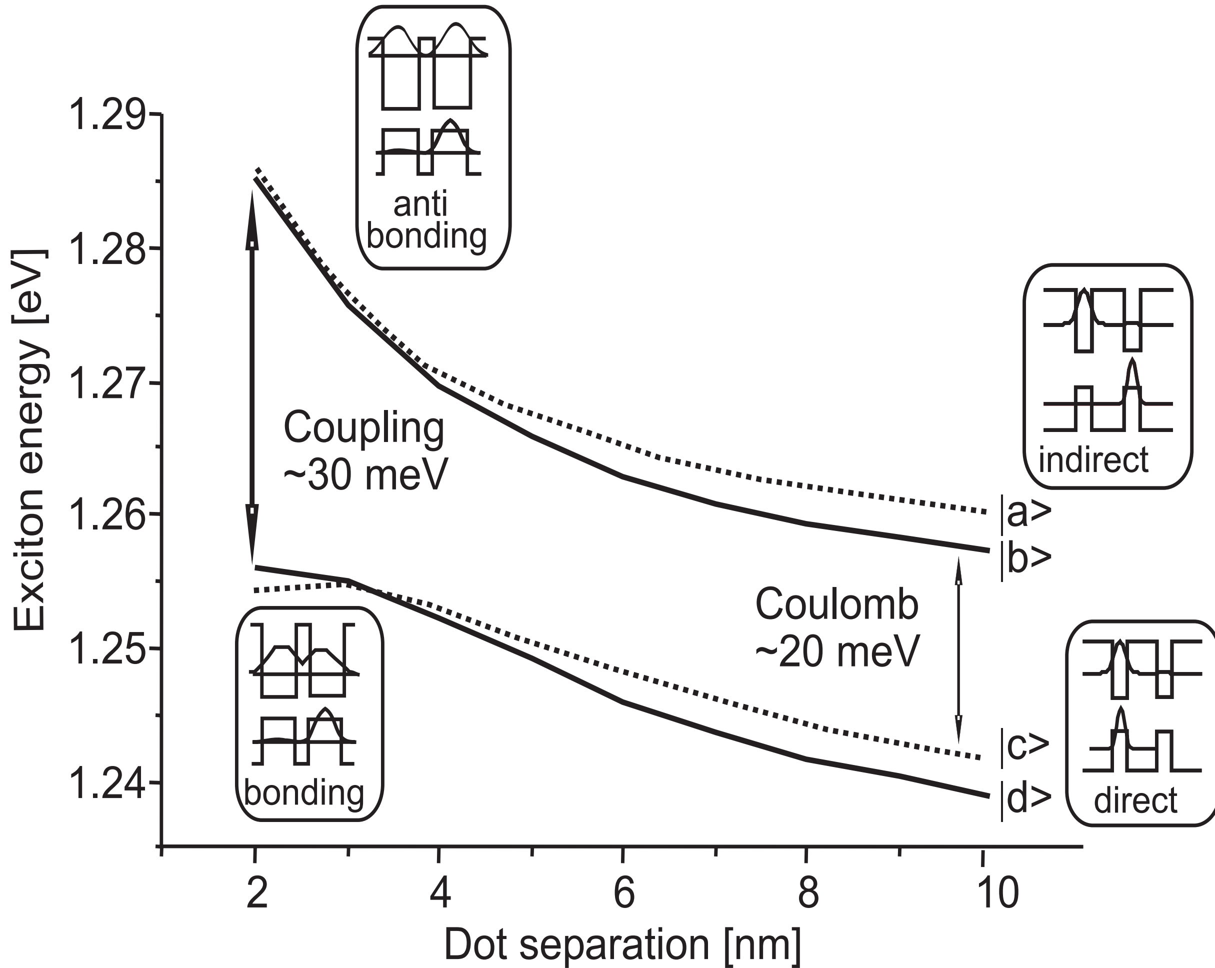


Figure 9.8: Exciton transition energies as a function of QD separation.

9.1.5 Stark shift in coupled quantum dots

In order to obtain further insight into the physics of these QDMs, external electric fields pose an ideal probe which has been demonstrated in the section of charged quantum dots. Most importantly, the nature of the excitonic states in coupled quantum dots can be changed adiabatically from direct to indirect as a function of the applied electric field, as will be demonstrated in the following section. Figure 9.9 shows the exciton energies and the electron-hole overlap for two structures with different QD-separation, namely 2 nm (right column) and 6 nm (left column).

The exciton states of the strongly coupled dots (fig. 9.9b) can be classified into bonding and anti-bonding states that exhibit clear anti-crossings with a splitting of 27 meV, which is comparable to the results of the single particle states at zero field presented in figure 9.6. As a comparison the exciton transition energy of a single dot of identical shape and composition is indicated by the orange line emphasizing the blue shift due to the increased hydrostatic strain in the coupled dots. The corresponding electron hole overlap (fig. 9.9d) shows a rather weak dependence on the electric field with all four states exhibiting a non-zero oscillator strength over the whole field range, which is a clear indication of strongly coupled states.

This picture changes significantly for the quantum dots with a large separation of 6 nm (fig. 9.9a), where the states can be classified in terms of direct and indirect excitons. The

direct excitons exhibit a small quadratic Stark shift analogously to the single dots, whereas the energy of indirect excitons is linearly shifted $E_{Stark} = -p \cdot F$ with a dipole p proportional to the distance between the two quantum dots. At electric fields of ± 15 kV/cm anti-crossings are observed that correspond to a direct state (solid/dashed line) that is optically active (bright) being changed adiabatically into an indirect exciton (dark) with increasing/decreasing electric field. This fact is additionally supported by the sharp change in the oscillator strength of the states (fig. 9.9c) in the vicinity of the anti-crossing, with identical strength right at the point of the anti-crossing. The comparison with the single dot exciton shows that the direct states of the coupled dots with 6 nm separation behave very much like single dots, while for 2 nm separation the coupling is so strong that it dominates the shape of the Stark shifts within the regarded field range. Again the anti-crossings provide information on the coupling strength which is about 4 meV for the 6 nm separation.

Comparing the results for the 6 nm separation with the analytical model of the coupled two-level Hamiltonian (eq. 9.2) we can identify the splitting Δ in equation 9.3 with the energy difference of the direct and the indirect exciton states. Since this energy splitting changes linearly with the applied electric field via the quantum confined Stark shift, the coupling of the system is perfectly controllable by the external field. This is an important finding, since the control of the coupling between two states is a key prerequisite for any application in the field of quantum computing.

9.2 Comparison to experimental results

To investigate the influence of electric field perturbation on a QDM experimentally, Hubert Krenner, Michael Reimer and Emily Clark performed photoluminescence (PL) spectroscopy on single pairs of stacked self-assembled QDs embedded in electrically active n-i Schottky junctions. The samples have been grown by molecular beam epitaxy and contained two layers of InGaAs QDs separated by a 10 nm spacer layer of GaAs. They were processed into photo diodes equipped with shadowmasks patterned with sub-μm sized apertures in order to isolate single QDMs. By applying a bias voltage across the structure the electric field (F) along the growth direction can be tuned from 0 to ~ 250 kV/cm. For low electric fields ($F < 30$ kV/cm) PL experiments were performed whereas for higher $F > 40$ kV/cm carrier the radiative lifetime of the excitonic states exceeds the tunneling time out of the QDMs.

Figure 9.10 shows the photoluminescence signal of a single QDM under the influence of an applied electric field. The upper-most curve in this waterfall plot corresponds to a field strength of ˜15 kV/cm which is linearly increasing up to ˜23 kV/cm for the lowest spectrum. For this QDM two exciton branches are clearly resolved which show a characteristic anti-crossing behavior with varying electric field F. For low electric fields $F < 17$ kV/cm only one is observed indicating that it is the excitonic ground state of the QDM. With increasing F another peak appears at higher energy which shows a strong shift with the electric field whilst the Stark shift of the first peak stays weak. A clear anti-crossing of the two line is observed for field strengths of $F \sim 18.5$ kV/cm where the minimum splitting occurs. For further increasing F the magnitudes of the Stark shifts of the two states are exchanged leading to an increasing energetic splitting. The signal of both states weakens for higher fields, due to the increased tunnel escape. The inset of fig. 9.10 shows the comparison between the extracted measured peak positions and theoretical calculation for a dot height of 4 nm and a separation of 8 nm that are in very good agreement. Please note that to improve the visibility, the theoretical curves

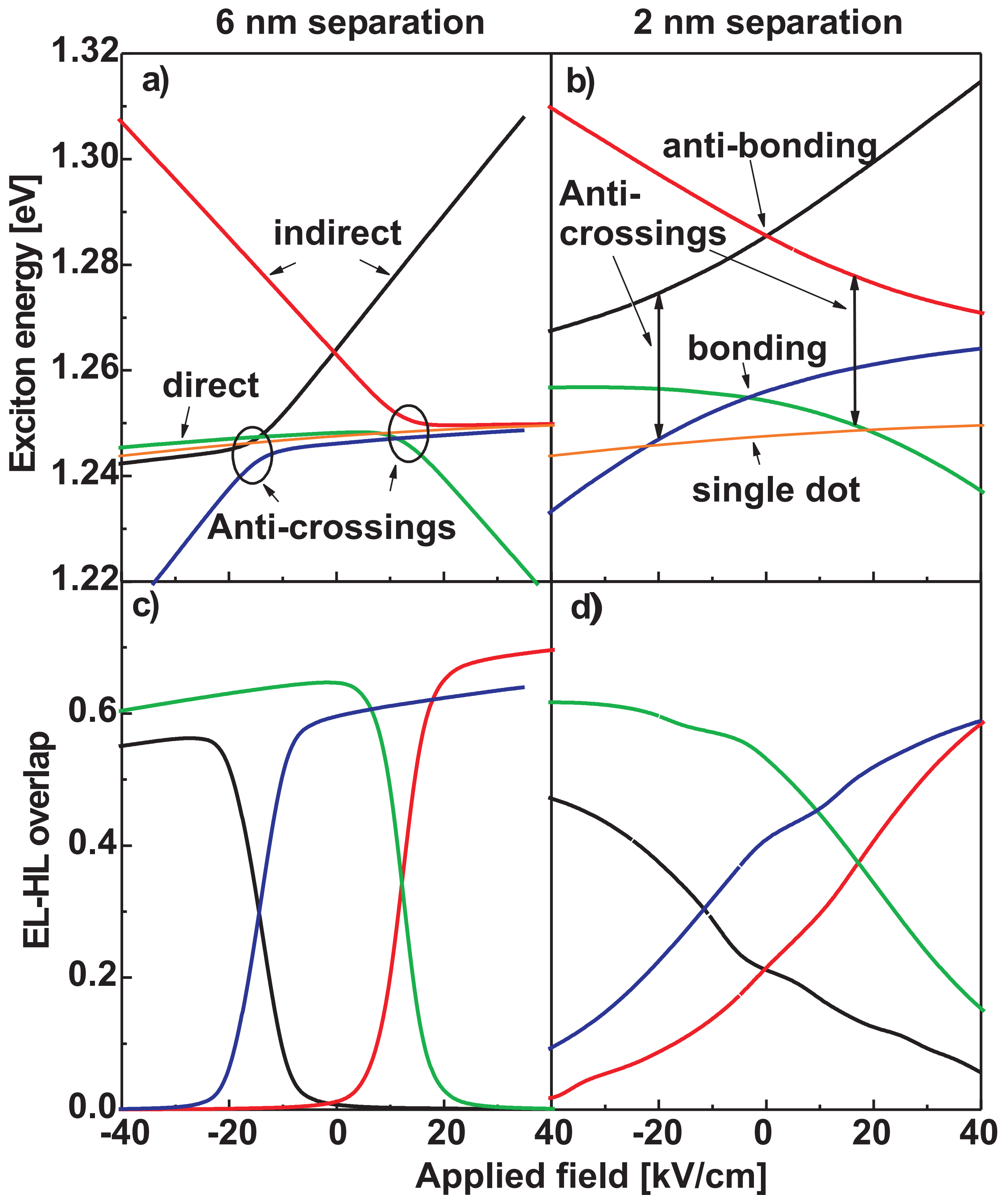


Figure 9.9: The exciton energies a),b) and the electron hole overlap c),d) is plotted as a function of the external applied electric field for QDMs with 2 nm and 6 nm dot separation. The corresponding exciton energy of the single dot is indicated by the orange line.

have been shifted by small offsets in energy (+5 meV) and field (−3.4 kV/cm), corresponding to very small changes in size and composition of the dot, that do not alter the qualitative behavior.

Comparing the measured coupling energies for various QDMs, a range between 1 and 2 meV is observed as indicated by the shaded rectangle in fig. 9.11a). This corresponds to QD separation within the theoretical model of 7–9 nm leading to a wetting layer separation of 11–13 nm for a dot height of 4 nm. This slight over-estimation of the nominal spacer thickness of 10 nm, can be explained by either a smaller dot height in the real sample or an effective increase in the interdot distance due to the bulging of the second wetting layer as seen in figure 9.1. The second characteristic quantity, the position of the anti-crossing as a function of the applied field is mainly dependent on the initial splitting between the energies of the direct and indirect exciton at zero field. Thus for two identical dots this splitting is given by the Coulomb interaction and the strain effects, but if the size or Indium concentration of one of the dots is varied, this splitting is additionally shifted on the field axis. Figure 9.11b) shows the anti-crossing field as a function of the energy difference of the direct excitons ($E^{\mathrm{EX}}_{\mathrm{lower\ dot}} - E^{\mathrm{EX}}_{\mathrm{upper\ dot}}$). The three points in the graph were obtained for a height of the lower dot of 4.0, 3.5 and 3.0 nm corresponding to energy differences of −13, 1.5 and 15 meV. Again, the experimentally measured values indicated by the shaded corridor between 14 and 20 kV/cm are well within the predicted range.

9.3 Conclusion

In this chapter we have studied the electronic structure of a system of two vertically coupled quantum dots under the influence of an external electric field. Our calculations predict that the nature of the excitonic states can be controlled via the application of an external electric field, a finding which has shown to be in excellent agreement with recent experimental results of the group of prof. John Finley. The theoretical analysis of the physics of the quantum dot molecule exhibits that the observed exciton transition energy results from a complex interplay of three dominating effects, namely the strain, the Coulomb interaction and the quantum mechanical coupling, that are all of the same order. A fact that is supported by the comparison to other works, where good agreement is found with the results of sophisticated models as used in [143] in contrast to simpler models [148, 151, 152, 153], that neglect important effects. In summary we conclude that the model applied within this work has proven to yield predictive results for complex, strongly coupled systems of quantum dots.

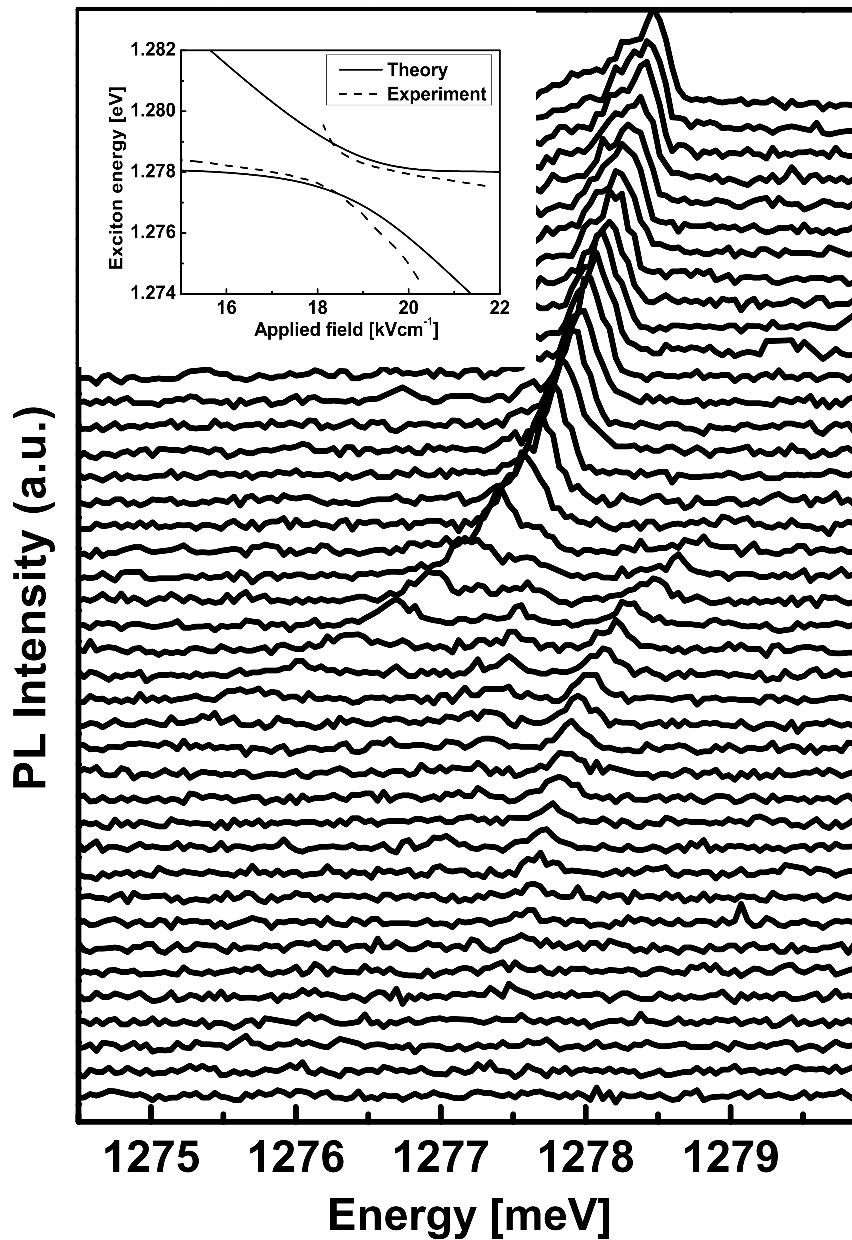


Figure 9.10: Photoluminescence of a single QDM as a function of the applied electric field ranging from 12.5 to 25 kV/cm. The two excitonic states show a clear anticrossing behavior with a minimum splitting at 18.5 kV/cm. Inset: Comparison of the extracted peak position with theoretical calculations.

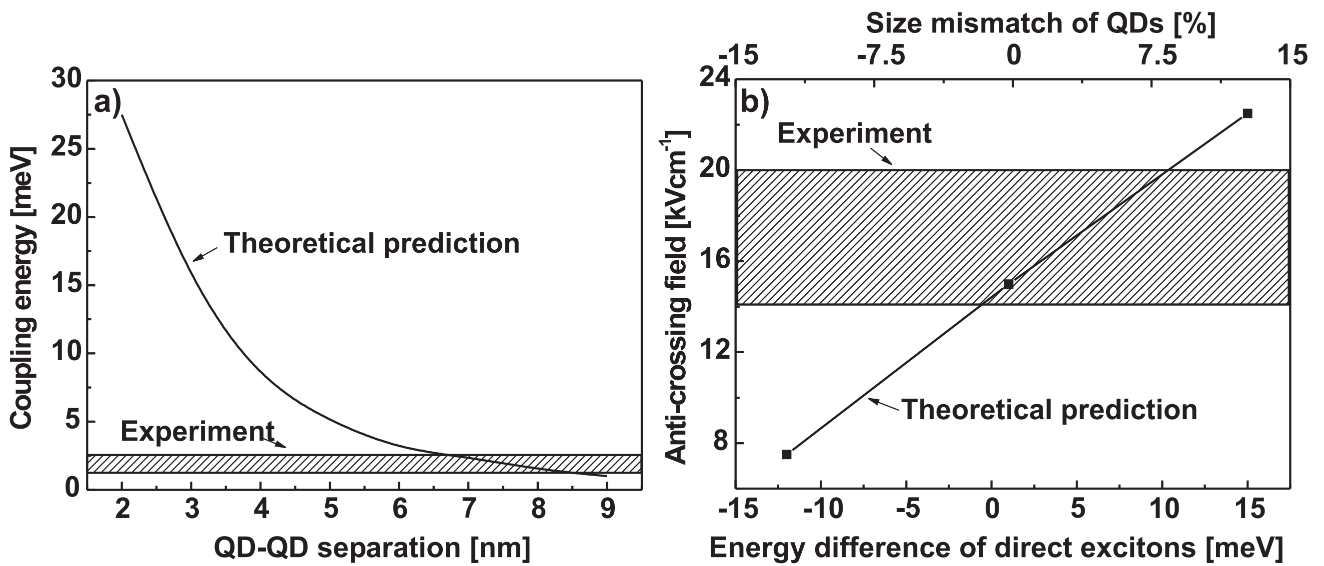


Figure 9.11: a) Calculated coupling energy as a function of QD separation (solid line). The range of the measured values is indicted by the shaded bar. b) Electric field of the anti-crossing as a function of the energy difference between the direct excitons ($E^{\mathrm{EX}}_{\mathrm{lower_dot}} - E^{\mathrm{EX}}_{\mathrm{upper_dot}}$) at zero field induced by the size mismatch between upper and lower dot. The corridor of the experimental data is indicated by the shaded rectangle.

Appendix A

Closed system limit for Ξ

In this Appendix we derive an expression for the generalized density of states in eq. (2.26) in the limit of the decoupled device. In this limit, the self-energy can be written as $\mathbf{\Sigma} = -i\eta\mathbf{1}$ where $\eta \to 0+$ is an infinitesimal positive constant. Correspondingly, the matrix in eq. (2.11) becomes $\mathbf{\Gamma} = 2\eta\mathbf{1}$, and the matrix $\mathbf{B}_C^{-1} = \left[\mathbf{1} - \mathbf{\Sigma}\mathbf{G}^0\right]^{-1}$ tends towards the unit matrix. This leads to

$$\mathbf{\Xi}_{\alpha\beta}(E) = \frac{1}{2\pi}\left[\frac{\mathrm{Tr}\left(\left[|\beta\rangle\langle\alpha|\right]2\eta\right)}{\left(E-\varepsilon_a - i\eta\right)\left(E-\varepsilon_\beta + i\eta\right)}\right] \tag{A.1}$$

$$= \mathrm{Tr}\left(\left[|\alpha\rangle\langle\beta|\right]\right)\delta_{\alpha\beta}\delta\left(E-\varepsilon_\alpha\right) = \delta_{\alpha\beta}\delta\left(E-\varepsilon_\alpha\right), \tag{A.2}$$

where the relation

$$\lim_{\eta\to 0+}\left[\frac{1}{\pi}\frac{\eta}{\left(E-\varepsilon_a - i\eta\right)\left(E-\varepsilon_\beta + i\eta\right)}\right]$$

$$= \delta_{\alpha\beta}\delta\left(E-\varepsilon_\alpha\right)\ , \tag{A.3}$$

has been used.

Appendix B

Hamiltonian for k·p calculations

In our multi-band example, we employ a 4-band Luttinger-Kohn Hamilton represented in the $p_{3/2}$ spinor basis: [14]

$$\mathbf{H}_{kp} = \begin{bmatrix} P+Q & -S & R & 0 \\ -S^* & P-Q & 0 & R \\ R^* & 0 & P-Q & S \\ 0 & R^* & S^* & P+Q \end{bmatrix},$$

For convenience, we set $\hbar^2/(2m) = 1$ in this Appendix. We denote the Cartesian wave-number components along the principal axes of the crystal by k_i. Because of the orientation of the cleaved edge structure in fig. 2.2, we introduce additionally the rotated components $k'_y = (k_y + k_z)/\sqrt{2}$ that lies along [011], and $k'_z = (k_y - k_z)/\sqrt{2}$ along the $[01\bar{1}]$ direction. With these wave vectors, the Hamiltonian matrix elements ca be written as

$$P = \gamma_1 \left(k_x^2 + k_y'^2 + k_z'^2\right) - E_v, \quad \text{(B.1)}$$

$$Q = \gamma_2 \left(k_x^2 - \frac{1}{2}k_y'^2 - \frac{1}{2}k_z'^2 + 3k'_y k'_z\right), \quad \text{(B.2)}$$

$$S = 2\sqrt{3}\gamma_3 \left(\frac{1}{\sqrt{2}}k'_y k_x - \frac{1}{\sqrt{2}}k'_z k_x - i\frac{1}{2}\left(k_y'^2 - k_z'^2\right)\right), \quad \text{(B.3)}$$

$$R = \sqrt{3}\left[-\gamma_2 \left(k_x^2 - \frac{1}{2}k_y'^2 - \frac{1}{2}k_z'^2 - k'_y k'_z\right) + \sqrt{2}i\gamma_3 \left(k_x k'_y + k_x k'_z\right)\right]. \quad \text{(B.4)}$$

Here γ_1, γ_2 and γ_3 are the Luttinger parameters. For the discretization in real space, we quantize the x and the y axis by replacing the scalars k_x and k_y by the differentials $k_x \to -i\partial/\partial x, k_y \to -i\partial/\partial y$, whereas k'_z remains a scalar parameter. We use a finite difference mesh with constant grid spacing Δx and spatially constant Luttinger parameters $\gamma_1 = 6.98, \gamma_2 = 2.06, \gamma_3 = 2.93$. The discrete first and second order derivatives of the wave function $\psi(x, y) \to \psi_{i,j}$ can be

written as

$$\frac{\partial \psi}{\partial x} = \frac{\psi_{i+1,j} - \psi_{i-1,j}}{2\Delta x}, \tag{B.5}$$

$$\frac{\partial^2 \psi}{\partial x^2} = \frac{1}{\Delta x^2}\left(\psi_{i-1,j} - 2\psi_{i,j} + \psi_{i+1,j}\right), \tag{B.6}$$

$$\frac{\partial^2 \psi}{\partial x \partial y} = \tag{B.7}$$

$$\frac{1}{4\Delta x^2}\left(\psi_{i-1,j-1} - \psi_{i+1,j-1} - \psi_{i-1,j+1} + \psi_{i+1,j+1}\right). \tag{B.8}$$

Correspondingly, the Hamiltonian can be divided into six Hermitian matrices as follows,

$$\begin{aligned} &\mathbf{H}\left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, k_z'\right) \\ = \; &\mathbf{H}_0(k_z') - i\mathbf{H}_x(k_z')\frac{\partial}{\partial x} - \mathbf{H}_{xx}\frac{\partial^2}{\partial x^2} \\ &-i\mathbf{H}_y(k_z')\frac{\partial}{\partial y} - \mathbf{H}_{yy}\frac{\partial^2}{\partial y^2} - \mathbf{H}_{xy}\frac{\partial^2}{\partial x \partial y}, \end{aligned} \tag{B.9}$$

$$\mathbf{H}_0\left(k_z'\right) = k_z'^2 \begin{bmatrix} \gamma_1 - \frac{1}{2}\gamma_2 & -i\sqrt{3}\gamma_3 & \frac{1}{2}\sqrt{3}\gamma_2 & 0 \\ & \gamma_1 + \frac{1}{2}\gamma_2 & 0 & \frac{1}{2}\sqrt{3}\gamma_2 \\ & & \gamma_1 + \frac{1}{2}\gamma_2 & i\sqrt{3}\gamma_3 \\ & & & \gamma_1 - \frac{1}{2}\gamma_2 \end{bmatrix}, \tag{B.10}$$

$$\mathbf{H}_x\left(k_z\right) = k_z' \begin{bmatrix} 0 & \sqrt{6}\gamma_3 & \sqrt{6}i\gamma_3 & 0 \\ & 0 & 0 & \sqrt{6}i\gamma_3 \\ & & 0 & -\sqrt{6}\gamma_3 \\ & & & 0 \end{bmatrix}, \tag{B.11}$$

$$\mathbf{H}_{xx} = \begin{bmatrix} \gamma_1 + \gamma_2 & 0 & -\sqrt{3}\gamma_2 & 0 \\ & \gamma_1 - \gamma_2 & 0 & -\sqrt{3}\gamma_2 \\ & & \gamma_1 - \gamma_2 & 0 \\ & & & \gamma_1 + \gamma_2 \end{bmatrix}, \tag{B.12}$$

$$\mathbf{H}_y\left(k_z\right) = k_z' \begin{bmatrix} 3\gamma_2 & 0 & \sqrt{3}\gamma_2 & 0 \\ & -3\gamma_2 & 0 & \sqrt{3}\gamma_2 \\ & & -3\gamma_2 & 0 \\ & & & 3\gamma_2 \end{bmatrix}, \tag{B.13}$$

$$\mathbf{H}_{yy} = \begin{bmatrix} \gamma_1 - \frac{1}{2}\gamma_2 & i\sqrt{3}\gamma_3 & \frac{1}{2}\sqrt{3}\gamma_2 & 0 \\ & \gamma_1 + \frac{1}{2}\gamma_2 & 0 & \frac{1}{2}\sqrt{3}\gamma_2 \\ & & \gamma_1 + \frac{1}{2}\gamma_2 & -i\sqrt{3}\gamma_3 \\ & & & \gamma_1 - \frac{1}{2}\gamma_2 \end{bmatrix}, \tag{B.14}$$

$$\mathbf{H}_{xy} = \begin{bmatrix} 0 & -\sqrt{6}\gamma_3 & i\sqrt{6}\gamma_3 & 0 \\ & 0 & 0 & i\sqrt{6}\gamma_3 \\ & & 0 & \sqrt{6}\gamma_3 \\ & & & 0 \end{bmatrix}. \tag{B.15}$$

Appendix C

Normalization of the bound states

One further issue that has not been addressed yet, is the normalization of the bound states. In principle, if the state is completely localized within the simulated device domain, there is no problem and the probability density is just normalized to one. However, in our example, the state is very weakly bound by typically just a few meV and therefore decays very slowly into the leads as shown in figure 3.8. Thus by truncating the gate region at the contact where the bound state is not completely decayed, we have to account for the portion of the state that is localized outside of the domain. This is crucial since the states are per default normalized to the domain volume where the Schrödinger equation is solved. For the scattering boundary conditions we assume that the portion of the wavefunction ψ^α that is localized in lead λ is a product state formed of the cross-sectional lead modes ξ_λ^m and an exponential function along the propagation direction z of the lead

$$\psi^\alpha(y,z) = \sum_m |\xi_\lambda^m(y)\rangle \langle \xi_\lambda^m | \psi^\alpha \rangle e^{-\kappa_\alpha^{\lambda,m} z} \qquad y,z \in L_\lambda, \tag{C.1}$$

where y denotes the cross-sectional coordinates, and L_λ denotes the volume of lead λ. The decay vector κ can be easily computed from the dispersion of the lead

$$\begin{aligned} E_m^\lambda - E_\alpha &= \frac{\hbar^2}{2m^*}\left(\kappa_\alpha^{\lambda,m}\right)^2 \qquad E_m^\lambda > E_\alpha \\ \Longrightarrow \kappa_\alpha^{\lambda,m} &\equiv \left| \sqrt{\frac{2m^*}{\hbar^2}\left(E_m^\lambda - E_\alpha\right)} \right| \end{aligned} \tag{C.2}$$

where E_m^λ is the energy of the mode m in lead λ and E_α is the eigenenergy of state α. With the expression of the decay vector we can now integrate the wavefunction in all leads to obtain the portion of the state that is localized outside of the device. Since the goal is the proper normalization of the state ψ^α we first lift its normalization, since it has already been wrongly normalized to the domain volume V_D during the computation.

$$\begin{aligned} \int_{V_D} d\mathbf{r}\, |\psi^\alpha(\mathbf{r})|^2 &= 1, \\ \int_{L_y^\lambda} dy\, |\xi_\lambda^m(y)|^2 &= 1 \end{aligned}$$

thus defining the unnormalized states

$$\begin{aligned} \psi'_\alpha(\mathbf{r}) &= \sqrt{V_D}\psi^\alpha(\mathbf{r}), \\ (\xi_\lambda^m)'(y) &= \sqrt{L_y^\lambda}\xi_\lambda^m(y), \end{aligned}$$

where L_y^λ denotes the diameter of the lead. From this we can calculate the volumes V^α occupied by the state α inside and outside of the device domain

$$\begin{aligned}
\left|V_{inside}^{\alpha}\right|^2 &= \int_{V_D} d\mathbf{r} \left|\psi'_\alpha(\mathbf{r})\right|^2 = V_D \int_{V_D} d\mathbf{r} \left|\psi^\alpha(\mathbf{r})\right|^2 = V_D, \\
\left|V_{outside}^{\alpha}\right|^2 &= \sum_{\lambda,m} \int dy \int dz \left| \left|\xi_\lambda^m(y)\right\rangle \left\langle \xi_\lambda^m\right| \left|\psi^\alpha\right\rangle e^{-\kappa_\alpha^{\lambda,m} z} \right|^2 \\
&= L_y^\lambda \sum_{\lambda,m} \int dz \left| \left\langle \xi_\lambda^m\right| \left|\psi^\alpha\right\rangle e^{-\kappa_\alpha^{\lambda,m} z} \right|^2 \\
&= L_y^\lambda \sum_{\lambda,m} \left|\left\langle \xi_\lambda^m\right| \left|\psi^\alpha\right\rangle\right|^2 \frac{1}{2\kappa}.
\end{aligned}$$

Then the normalization of the wavefunction inside the device domain is

$$\sqrt{\frac{\left|V_{inside}^{\alpha}\right|^2}{\left|V_{outside}^{\alpha}\right|^2 + \left|V_{inside}^{\alpha}\right|^2}}\psi^\alpha(\mathbf{r}).$$

With this way of normalization we can now handle even weakly bound states that leak out of the device into the leads. Typically in our example the difference between the old and the corrected norm is less than 10 %.

Appendix D

CBR self-consistent code

D.1 User guide

The CBR method is an efficient method that uses a limited set of eigenstates of the decoupled device and a few propagating lead modes to calculate the retarded Green's function of the device coupled to external contacts. From this Green's function, the density and the current is obtained in the ballistic limit using Landauers formula with fixed Fermi levels for the leads. It is important to note that since the efficiency of the calculation, but also the convergence of the results is strongly dependent on the cutoff energies for the eigenstates and modes. Thus it is important to check during the calculation if the specified number of states and modes is sufficient for the applied voltages. To summarize the code may do its job very efficiently, but its not at all a black box tool.

This text is meant as a very brief instruction how to use the program, meaning where to specify the relevant parameters and how to setup the device structure. The description of the algorithm can be found in the Ph.D. thesis of M. Sabathil and in various publications. The program has been developed within a Ph.D. thesis and therefore is not at all fool prove and has only been tested on very few devices such that it is not unexpected if it fails to converge or yields strange results.

Since the code does not feature a fancy input system like nextnano has it, almost all specifications have to be hard wired in the code. Only a few parameters are meant to be provided via the command line to allow for the efficient scan of small variations of the gate oxide for example. In the following we will describe the setup of the structure, the relevant convergence parameters and the resulting output.

D.1.1 Simulation domain

The specification of the dimensions of the simulation domain is done in the module **Device_constants** in the file 'common.f90'.

File:	common.f90	
Module:	Device_constants	
domain boundaries	`xmin,xmax,ymin,ymax`	[Å]
grid spacing	`dx,dy`	[Å]
Number of attached contacts	`Num_contacts`	
Number of materials used in device	`Num_materials`	
Lattice temperature	`Temperature`	[K]
maximum number of valleys in conduction band	`num_valleys`	
degeneracy of each valley	`valley_degeneracy`	

The total number of points is directly calculated from these values leading to a total mesh-size of Nmax=7137 points for the current example. This number is important since it determines the size of the memory usage and the computational time which both scale linearly with Nmax.

D.1.2 Definition of contacts

After the specification of the simulation domain we will determine the position of the contacts. This is done in the subroutine 'Setup_Leads' in module 'Device_Setup_routines', file 'device_setup.f90'.

The number of contacts is specified in the parameter **Num_contacts** in module 'Device_constants'. For each contact it has to be specified if it lies along the

$$\mathbf{x}_{axis} \rightarrow \texttt{low_contact_edge(contact,1) = 0}$$
$$\mathbf{y}_{axis} \rightarrow \texttt{low_contact_edge(contact,1) = 1.}$$

Then the first point of the lead is specified as follows

$$\texttt{low_contact_edge(contact,2)} = x_{\min}$$
$$\texttt{low_contact_edge(contact,3)} = y_{\min}$$

The size of the contact give by the number of grid points

$$\texttt{contact_size(num_contact)} = N_c$$

To check if the definition of the contacts is correct, you can plot the matrix '2D_Leads.dat'.

D.1.3 Voltage sweep

The main self-consistent cycle is performed in the file 'self_consistent_solution.f90', where the Fermi levels of the Contacts are set in the array `voltage`. The kind of sweep has to be hardwired in this subroutine, but can then be steered by the parameters `bias, starting_voltage, voltage_step,` and `number_of_voltage_step` The program is currently configured to read in these parameters form the command line, but this might be changed by the user.

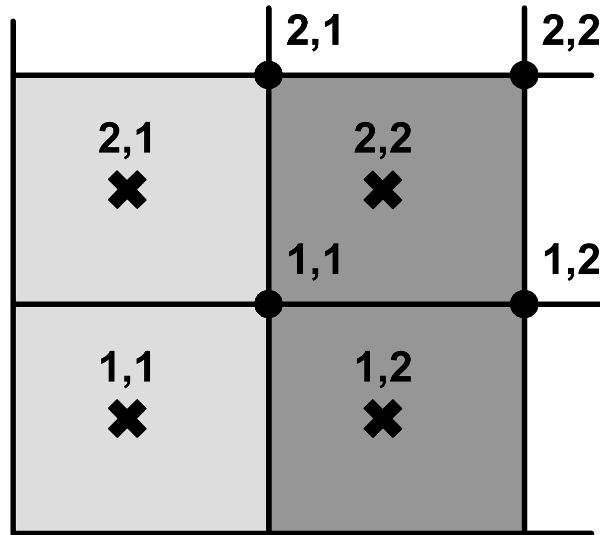


Figure D.1: The two complementary grids for the simulation. On the material grid indicated by the bold crosses, all material dependent parameters are specified. On the physical grid, indicated by the bold dots, the physical potentials and wavefunctions are defined.

D.1.4 Setup of material grid

The material grid is setup in routine 'setup_material_grid' in module 'Device_Setup_routines', file 'device_setup.f90'. The device is characterized by two complementary grids, the material grid and the physical grid. Figure D.1 depicts the two grids, where the bold crosses mark the material grid and the dots mark the physical grid. Points on the material grid correspond to some finite volume of size $dx \times dy$ that is homogeneously filled with one material. The grid points of the physical grid are placed at the interfaces between the material blocks. The material grid is shifted by half grid spacing relative to the physical grid and therefore is one grid point larger in each direction than the physical grid that has been specified by the parameters in the module 'device_constants'. On the material grid the numbers of the corresponding material at this position are stored, which then can be linked to all relevant parameters. The material grid `material_grid_xy(x,y)` is simply defined on the x-y basis together with the `dopant_grid_xy(x,y)`. The doping concentration is demanded in units [cm^{-3}] and has to have a minus sign for donors. For each material the relevant parameters are specified

File	device_setup.f90	
Routine	setup_material_grid	
number of material	`mat_num`	integer
conduction band edge	`band_edge(mat_num)`	$[eV]$
dielectric constant	`eps(mat_num)`	$[\varepsilon_0]$
inverse of transverse effective mass	`inv_effective_mass(mat_num,1)`	$[m_0^{-1}]$
inverse of transverse effective mass	`inv_effective_mass(mat_num,2)`	$[m_0^{-1}]$
inverse of longitudinal effective mass	`inv_effective_mass(mat_num,3)`	$[m_0^{-1}]$

Be sure that parameters are specified for each material, otherwise the program might crash! Your input can be immediately checked with the files '2D_material_grid.dat', '2D_cb_grid.dat',

'2D_mass_grid_mt.dat', '2D_mass_grid_ml.dat' all of which plot the parameters on the material grid. The doping density is plotted on the physical grid in the file '2D_Doping_dens.dat'.

D.1.5 Cutoff parameters

As already mentioned in the introduction the efficiency of the method relies on the truncation in energy and mode space which leads to a tremendous reduction of the computational effort. The cutoffs are specified by the number of states per valley and lead that are taken into account. During the calculation the program writes the corresponding cutoff energies to the standard output. It should be continuously checked that the lead cutoff energies are well above the Fermi levels of the corresponding contacts. The cutoff of the eigenstates of H^0 should be even higher to yield a smoother density. If too few states are taken into account density oscillations close to the contacts might appear. As a rule of thumb the cutoff should be twice as high as the relevant energy interval of the Fermi distributions. One should note however, that the computational effort roughly scales with the square of the number of eigenstates and lead modes taken into account.

		Point of specification
number of eigenstates of H^0	`Max_num_ev(1..num_valleys)`	main
number of lead-eigenmodes	`num_modes(num_contacts,num_valleys)`	setup_leads

D.1.6 Energy grid

The energy grid is adaptive to the density of states of the device. This is achieved by specially resolving the onset of the propagating modes and the resonant states that both commonly lead to very sharp peaks in the density of states. If these peaks are badly resolved no convergence will be reached. Therefore it is also very important to continuously check the DOS that is written out every iteration step in the file 'DOS_val***_ind**.dat'. The program calculates the mode energies and resolves the DOS around this energy by a local grid with very fine grid spacing. This is done for the first couple of modes in each lead while the remaining energy grid points are distributed within the interval to minimize the grid spacing. The parameters that specify the energy grid are defined in the module device_constants.

File	common.f90	
Module	device_constants	
Total number of energy grid points	`numE`	integer
maximum number of modes	`modes_per_lead_valley`	integer
Number of energy grid points in local grid	`points_per_mode`	integer
Initial energy grid spacing	`delta_E`	[eV]
Type of grid	`exponential_grid`	.true./.false.
Exponential factor	`grid_factor`	real
cutoff energy for consideration of modes	`lead_cutoff`	[eV]

The recommended grid type is an exponential grid since it provides good resolution for the $1/\sqrt{E}$ behavior of the expected peaks. The first grid point is set slightly below the onset of the

mode and then each grid point is set with increasing grid spacing

$$dE^n = \texttt{grid_factor} \times dE^{n-1},$$

starting with the initial grid spacing `delta_E`

Again no definite numbers for these parameters can be provided but the overall resolution should not be less the 1 meV, thus leading to roughly 1000 total energy grid points for low biases. Memory and computational time depend linearly on the total number of energy grid points.

D.1.7 Bound states

As explained in more detail in the section 3.4, in some structures under certain Bias conditions bound, or quasi bound states might exist. These states are defined by their very weak coupling to the propagating modes, such that their contribution to the density is not accounted for within the purely ballistic model. In the program an algorithm was developed to detect the bound states via the coupling and treat them separately to the strongly coupled states. The detection works directly by the overlap of the state with the propagating modes at its eigenenergy. If this coupling is below a certain treshold that is determined by the parameter `bound_state_treshold`, the state is considered to be a bound state. This detection system is not at all working perfectly and certainly has to be improved in upcoming versions. The overlap to the propagating modes is written to the file 'coupling_val***_ind**.dat' during the iteration. The first column shows the eigenenergies of the eigenstates, the following column the coupling to the propagating modes of each lead and the last column shows the sum of the couplings. If this sum is below the threshold the state is considered to be a bound state. Therefore it is necessary to check whether the treshold is such that the majority of states is not affected. Figure D.2 shows typical couplings that occur in the DGFET example. The numbered dots would be recognized as bound states for a threshold of 0.1. If a bound state is detected it is attributed to a certain lead, which is currently done via the coupling to all modes including the non-propagating modes. A more sophisticated way would be the overlap of the wavefunction with the continuous density resulting from each lead. Once the state is detected and attributed it is taken out of the continuous set and added to the discrete set, where it is then occupied with Fermi level of the corresponding lead.

Important: The routine responsible for the bound states is 'Extract_bound_states' in the file 'CBR_routine_new.f90'. This will be the place to change things if convergence is off, or the results are physically meaningless.

The density resulting from the discrete part is plotted in the file 'discrete_dens.dat', this should be checked regularly.

Switch off bound state detection: If this procedure is not desired or fails to produce reasonable results, it might be switched off by setting the flag **'ignore_bstates=.true'** that is currently specified in the main part.

D.1.8 Self-consistent cycle

The self-consistent calculation of the potential and electron density is performed using a predictor corrector scheme that relies on a correction step that is calculated using a quasi-classical approximation of the density within a Newton scheme. This results in two cycles per bias

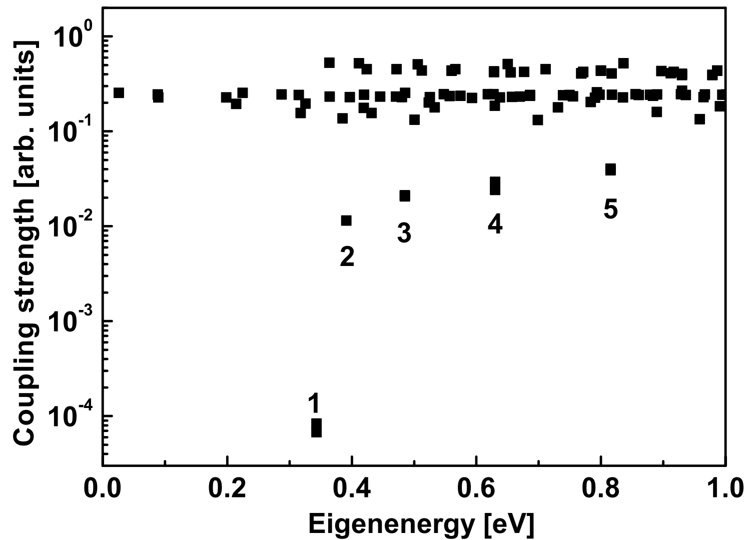


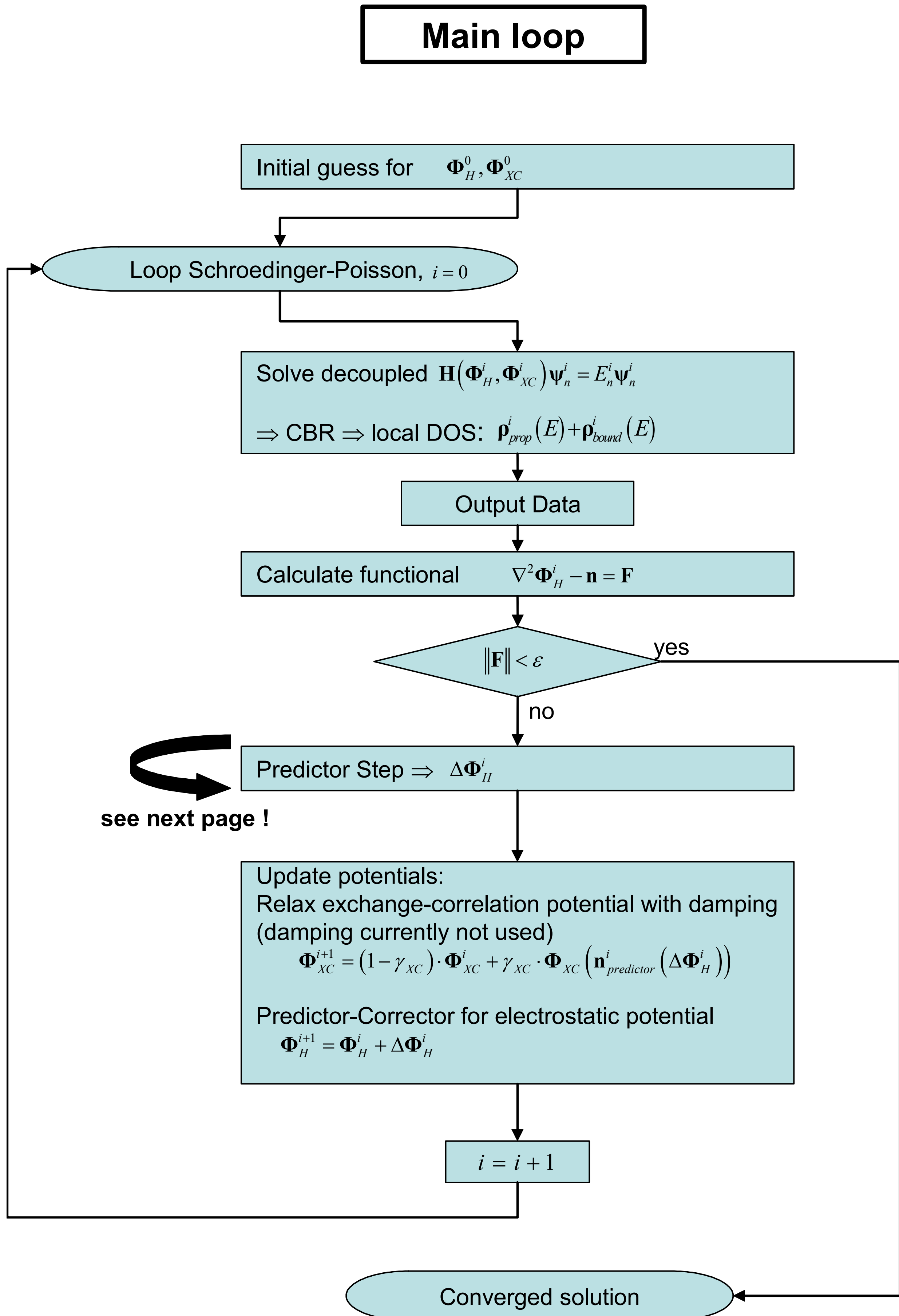
Figure D.2: Typical graph resulting from the file 'coupling_val***_ind**.dat'. The dots represent the sum of all couplings given in the last column of the file. The numbered dots would be considered as bound states, since they are well below the treshold of 0.1.

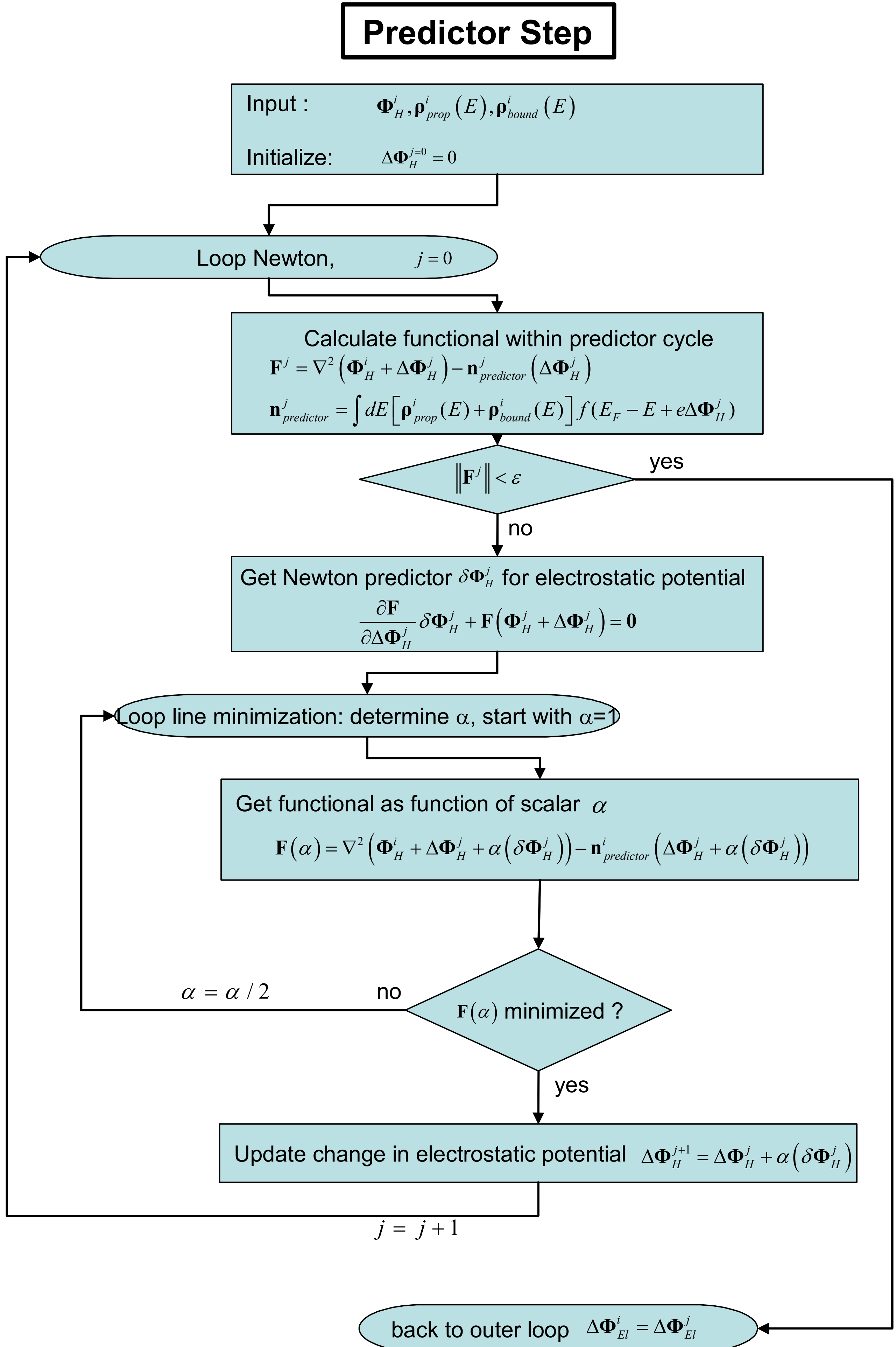
point: an outer cycle that contains the heavy calculation of the new eigenstates and the resulting density of states, and an inner cycle that contains the comparably cheap Newton scheme to generate a new correction to the potential. This method has been proven to be very robust and rather efficient. There are a few parameters that control the self-consistent cycle.

File	main.f90	
Max Number of outer iterations	`CBR_Poisson_iteration`	integer
Treshold for residuum	`abs_error`	real
Max number of iterations in Newton	`iter_newton`	integer
Max number of iterations in Linesearch	`iter_linesearch`	integer
File	common.f90	
module	device_constants	
Relative improvement of residuum	`predictor_factor`	real

The most important parameters are certainly the ones that control the outer loop, namely `CBR_Poisson_iteration` and `abs_error`. The convergence behavior can be tracked using the two files 'residuum.dat' and 'current_convergence'. The residuum file contains two columns that correspond to two different norms of the residuum. The first column is the maximum value of the residuum vector, while the second column is the squared norm of the entire vector. This last norm corresponds to the value set in `abs_error`. Another way to check the result is the convergence of the current, since usually it is fair enough if the current is converged to less than one percent error.

D.2 Flow chart




Predictor Step
Input :
Initialize:
Loop Newton, $j=0$
Calculate functional within predictor cycle
yes
no
Get Newton predictor $\delta\mathbf{\Phi}_H^j$ for electrostatic potential
Loop line minimization: determine α, start with α=1
Get functional as function of scalar α
$\alpha = \alpha / 2$
no
$\mathbf{F}(\alpha)$ minimized ?
yes
Update change in electrostatic potential
$j = j+1$
back to outer loop

Appendix E

Local spin density approximation (LSDA)

The local spin density approximation (LSDA) [111][104] accounts for a net spin polarization

$$\zeta = \frac{n_\uparrow - n_\downarrow}{n}$$

of the system in a similar way as LDA by using the expressions of exchange and correlation energy of a homogenous carrier gas that is fully polarized $\zeta = 1$ or completely unpolarized $\zeta = 0$.

$$\begin{aligned} \varepsilon_{xc}(n) &= \varepsilon_{xc}^{P}(n) \qquad \zeta = 1 \\ \varepsilon_{xc}(n) &= \varepsilon_{xc}^{U}(n) \qquad \zeta = 0 \end{aligned}$$

For a system with arbitrary polarization $0 \leq \zeta \leq 1$ an interpolation is used

$$\varepsilon_{xc}(n,\zeta) = \varepsilon_{xc}^{U}(n) + f(\zeta)\left[\varepsilon_{xc}^{P}(n) - \varepsilon_{xc}^{U}(n)\right],$$

with the interpolation formula

$$f(\zeta) = \frac{1}{2}\frac{(1+\zeta)^{4/3} + (1-\zeta)^{4/3} - 2}{2^{1/3} - 1}.$$

The single Kohn-Sham equation of the LDA is now split into two equations, one for spin-up and one for spin-down carriers

$$\left[-\frac{\hbar^2}{2m}\nabla^2 + V_{ext}(\mathbf{r}) + V_{Hartree}(\mathbf{r}) + \left(V_{xc}^{LSDA}(\mathbf{r})\right)_\sigma\right]\psi_{\alpha\sigma}^{e}(\mathbf{r}) = \varepsilon_{\alpha\sigma}^{e}\psi_{\alpha\sigma}^{e}(\mathbf{r}),$$

with a spin dependent exchange and correlation potential $\left(V_{xc}^{LSDA}\right)_\sigma$.

E.1 Exchange

The expressions for the exchange energies are

$$\begin{aligned} \varepsilon_{x}^{U}(n) &= -\frac{3}{4}\left(\frac{3}{\pi}n(\mathbf{r})\right)^{1/3} e^2, \\ \varepsilon_{x}^{P} &= 2^{1/3}\varepsilon_{x}^{U}(n), \end{aligned}$$

which results in an expression for arbitrary polarization

$$\varepsilon_x(n,\zeta) = -e^2\frac{3}{4}\left(\frac{3}{\pi}n(\mathbf{r})\right)^{1/3}\left[1+f(\zeta)\left(2^{1/3}-1\right)\right].$$

For the exchange potential we get **[119]**

$$\begin{aligned}
\mu_x(n,\zeta) &= \mu_x^U(n) + f(\zeta)\left[\mu_x^P(n)-\mu_x^U(n)\right] + \left[\varepsilon_x^P(n)-\varepsilon_x^U(n)\right]\left[sgn(\sigma)-\zeta\right]\frac{df}{d\zeta} \\
\mu_x^U(n) &= -e^2\bar{n}^{1/3}\left(\frac{3}{\pi}\right)^{1/3} \\
\mu_x^P(n) &= -2^{1/3}e^2\bar{n}^{1/3}\left(\frac{3}{\pi}\right)^{1/3} \\
\mu_x(n,\zeta) &= -e^2\bar{n}^{1/3}\left(\frac{3}{\pi}\right)^{1/3}\left[1+f(\zeta)\left(2^{1/3}-1\right)+\left[sgn(\sigma)-\zeta\right]\frac{df}{d\zeta}\left(2^{1/3}-1\right)\right] \\
\frac{df}{d\zeta} &= \frac{2}{3}\frac{(1+\zeta)^{1/3}-(1-\zeta)^{1/3}}{2^{1/3}-1}
\end{aligned}$$

E.2 Correlation

In this work we use the correlation energy of a homogeneous electron gas in the parametrization of Ceperley [154]. With the definition of a scaled length r_s

$$r_s : n = \frac{3}{4\pi r_s^3} \rightarrow r_s = \sqrt[3]{\frac{4\pi}{3}n},$$

Ceperley's parametrization for the correlation energy for $r_s \geq 1$ is given by:

$$\varepsilon_c^i = \frac{\gamma_i}{1+\beta_1^i\sqrt{r_s}+\beta_2^i r_s}$$

where $i = U$ (unpolarized, $\zeta = 0$) or P (polarized, $\zeta = 1$). The corresponding correlation potential is

$$\begin{aligned}
\mu_c^i &= \left(1-\frac{r_s}{3}\frac{d}{dr_s}\right)\varepsilon_c^i, \\
&= \varepsilon_c^i\frac{1+\frac{7}{6}\beta_1^i\sqrt{r_s}+\frac{4}{3}\beta_2^i r_s}{1+\beta_1^i\sqrt{r_s}+\beta_2^i r_s}.
\end{aligned}$$

For atomic calculations in the high density limit $r_s < 1$ and for arbitrary polarizations $0 < \zeta < 1$ we have

$$\varepsilon_c^i = A_i\ln(r_s) + B_i + C_i r_s\ln(r_s) + D_i r_s$$

with the corresponding correlation potential

$$\mu_c^i = A_i\ln(r_s) + \left(B_i - \frac{1}{3}A_i\right) + \frac{2}{3}C_i r_s\ln(r_s) + \frac{1}{3}\left(2D_i - C_i\right)r_s$$

For intermediate Spin polarizations an interpolation analogous to LSDA can be used

$$\begin{aligned}\varepsilon_c(r_s,\zeta) &= \varepsilon_c^U(r_s) + f(\zeta)\left[\varepsilon_c^P(r_s) - \varepsilon_c^U(r_s)\right] \\ \mu_c(r_s,\zeta) &= \mu_c^U(r_s) + f(\zeta)\left[\mu_c^P(r_s) - \mu_c^U(r_s)\right] + \left[\varepsilon_c^P(r_s) - \varepsilon_c^U(r_s)\right]\left[sgn(\sigma) - \zeta\right]\frac{df}{d\zeta}\end{aligned}$$

Parameters for correlation

In this work we use the parameters of Perdew and Zunger [119]

$Parameter$	U	P
γ	$-0,1423$	$-0,0843$
β_1	$1,0529$	$1,3981$
β_2	$0,3334$	$0,2611$
A	$0,0311$	$0,0155$
B	$-0,048$	$-0,0269$
C	$0,0020$	$0,0007$
D	$-0,0116$	$-0,0048$

which are given in the atomic units:

$$\begin{aligned}r_s &= \frac{r_0}{a_0} \\ r_0 &= \left(\frac{4\pi}{3}n\right)^{-1/3} \\ a_0 &= 4\pi\varepsilon_0\frac{\hbar^2}{me^2} = 0.529177 \quad \left[\mathring{A}\right] \text{ (Bohr-radius)} \\ Ha &= \frac{1}{4\pi\varepsilon_0}\frac{e^2}{a_0} = 27.212\ eV\end{aligned}$$

It is important to note that in a semiconductor with background dielectric constant and effective mass the units have to be rescaled [104]

$$\begin{aligned}a_0^* &= \frac{\varepsilon}{m^*}a_0, \\ Ha^* &= \frac{m^*}{\varepsilon^2}Ha.\end{aligned}$$

Appendix F

Publication list

F.1 Published papers

Parts of this thesis have been published in the following works:

- *Systematic reduction of the permanent exciton dipole for charged excitons in individual self-assembled InGaAs quantum dots*, J.J. Finley, M. Sabathil, R. Oulton, A.I. Tartakovskii, D.J. Mowbray, M.S. Skolnick, S. Liew, M. Migliorato, M. Hopkinson, P. Vogl, Physica E **21**, 199 (2004)

- *Prediction of a realistic quantum logic gate using the contact block reduction method*, M. Sabathil, D. Mamaluy and P. Vogl, Semiconduct. Sci. Technol. **19**, S137 (2004)

- *Contact block reduction method and its application to a 10 nm MOSFET device,* D. Mamaluy, A. Mannargudi, D. Vasiliska, M. Sabathil, and P. Vogl, Semiconduct. Sci. Technol. **19**, S118 (2004)

- *Efficient method for the calculation of ballistic quantum transport,* D. Mamaluy, M. Sabathil, and P. Vogl, J. Appl. Phys. **93**, 4628 (2003)

- *Efficient computational method for ballistic currents and application to single quantum dots*, M. Sabathil, S. Birner, D. Mamaluy, and P. Vogl, Journal of Computational Electronics **2**, 269 (2003)

- *Theory of vertical and lateral Stark shifts of excitons in quantum dots*, M. Sabathil, S. Hackenbuchner, S. Birner, J.A. Majewski, P. Vogl, J.J. Finley, phys. stat. sol. (c) 0 (4), 1181 (2003)

- *Towards fully quantum mechanical 3D device simulations*, M. Sabathil, S. Hackenbuchner, J.A. Majewski, G. Zandler, P. Vogl, Journal of Computational Electronics **1**, 81 (2002)

- *Nonequilibrium band structure of nano-devices*, S. Hackenbuchner, M. Sabathil, J.A. Majewski, G. Zandler, P. Vogl, E. Beham, A. Zrenner, P. Lugli, Physica B **314**, 145 (2002)

F.2 Accepted for publication

- *Calculation of carrier transport through quantum dot molecules*, T. Zibold, M. Sabathil, D. Mamaluy, P. Vogl, *Proceedings of the 27th International Conference on the Physics of Semiconductors* (2004)

- *Self-consistent contact block reduction method for ballistic nanodevices*, M. Sabathil, D. Mamaluy, and P. Vogl, Invited talk at the *International Workshop for Computational Electronics* (IWCE) (2004)

- *Probing Many Body Wavefunctions in Individual Self-Assembled InGaAs-GaAs Quantum Dots*, J.J. Finley and M. Sabathil, R. Oulton, D.J. Mowbray, and M.S. Skolnick, S.L. Liew and M. Hopkinson, Phys. Rev. B rapid communications.

F.3 Submitted papers

- *Direct observation of controlled quantum coupling in an individual dot molecule*, H.J. Krenner, M. Sabathil, E.C. Clark, D. Schuh, M. Bichler, P. Vogl, G. Abstreiter, and J.J. Finley, submitted to Physical Review Letters

- *Ballistic quantum transport in nano-devices*, D. Mamaluy and D. Vasileska, M. Sabathil, T. Zibold, and P. Vogl, submitted to Physical Review B

Bibliography

[1] International technology roadmap for semiconductors (ITRS), publ ished online at: http://public.itrs.net

[2] R. Landauer, Physica Scripta, **T42**, 110 (1992).

[3] J. Frenkel, Phys. Rev. B **36**, 1604 (1930)

[4] W. Ehrenberg, and H. Honl, Z. Phys. **68**, 289 (1931)

[5] M. Büttiker, IBM J. Res. Dev. **32**, 317 (1988)

[6] Supriyo Datta, *Electronic Transport in Mesoscopic Systems*, (Cambridge University Press, Cambridge, 1995).

[7] B. J. van Wees,H. van Houten, C.W.J. Beenakker, J.G. Williamson, L.P. Kouwenhoven, D. van der Marel, and C.T. Foxon, Phys. Rev. Lett. **60**, 848 (1988)

[8] E. O. Kane, in *Tunneling Phenomena in Solids*, edited by E. Burstein and S. Lundqvist (Plenum, New York, 1969), p. 1.

[9] J. N. Schulman and Y. C. Chang, Phys. Rev. B **27**, 2346 (1983)

[10] W. Frensley, Rev. Mod. Phys. **62**, 745 (1990).

[11] http://www.utdallas.edu/ ∼frensley/technical/qtrans/qtrans.html

[12] C. Lent, D. Kirkner, J. Appl. Phys. **67**, 6353 (1990).

[13] D. Z. Y. Ting, E. T. Yu, and T. C. McGill, Phys. Rev. B **45**, 3583 (1992).

[14] Y. X. Liu, D. Z. Y. Ting, and T. C. McGill, Phys. Rev. B **54**, 5675 (1996).

[15] E. S. Daniel, X. Cartoixa, W. Frensley, D. Z.-Y. TIng, and T. C. McGill, IEEE Trans. on Electron Dev., **47**, 1052 (2000).

[16] C. Strahberger and P. Vogl , Phys. Rev. B **62**, 7289 (2000).

[17] E. Polizzi, N. BenAbdallah, J. Appl. Phys. **87**, 8700 (2000).

[18] E. Polizzi, N. BenAbdallah, Phys. Rev. B **66**, 245301 (2002).

[19] S.E. Laux, A. Kumar, and M. V. Fischetti, J. Appl. Phys. **95**, 5545 (2004)

[20] L. Smrcka, Supercond. Sci. Technol. **8**, 221 (1990)

[21] E. R. Racec and Ulrich Wulf, Phys. Rev. B **64**, 115318 (2001)

[22] G. A. Nemnes, U. Wulf, P. N. Racec, J. Appl. Phys. **96**, 596 (2004)

[23] D. K. Ferry and S. M. Goodnick, *Transport in Nanostructures*, (Cambridge University Press, Cambridge, UK, New York, 1997).

[24] R. Lake, G. Klimeck, R. C. Bowen, and D. Jovanovic, J. Appl. Phys. **81**, 7845 (1997)

[25] A. Svizhenko, M.P. Anantram, T.R. Govindan, B. Biegel, and R. Venugopal , J. Appl. Phys. **91**, 2343 (2002).

[26] R. Venugopal, Z. Ren, S. Datta, and M. S. Lundstrom, and D. Jovanovic, J. Appl. Phys. **92**, 3730 (2002).

[27] C. Rivas, R. Lake, Phys. Stat. Sol. (b) **239**, 94 (2003)

[28] D. S. Fisher, and P. A. Lee, Phys. Rev. B **23**, 6851 (1981)

[29] Supriyo Datta, Superlattices and Microstructures **28**, 253 (2000).

[30] D. Mamaluy, M. Sabathil, P. Vogl, J. Appl. Phys. **93**, 4628 (2003).

[31] In the single band case, the matrix $\mathbf{B}_C$ reduces to $\mathbf{G}_C^0 \left(\mathbf{A}_C^{-1}\right)^T$.

[32] Y. Meir, N. S. Wingreen, Phys. Rev. Lett. **68**, 2512 (1992).

[33] P. S. Krstic, X.G. Zhang, W. H. Butler, Phys. Rev. B **66**, 205319 (2002).

[34] G. Schedelbeck, W. Wegscheider, M. Bichler, G. Abstreiter, Science **278**, 1792 (1997).

[35] A. Yacoby, H. L. Stormer, N. S. Wingreen, L. N. Pfeiffer, K. W. Baldwin, K. W. West, Phys. Rev. Lett. **77** 4612 (1996).

[36] H. Akiyama, T. Someya, H. Sakaki, Phys. Rev. B **53**, R4229 (1996).

[37] H. Akiyama, J. Phys. Cond. Mat. **10**, 3095 (1998).

[38] R. de Picciotto, H. L. Störmer, A. Yacobi, K. W. Baldwin, L. N. Pfeiffer, K. W. West, Physica E **6**, 514 (2000).

[39] F. Ertl ,T. Asperger, R. A. Deutschmann, W. Wegscheider, M. Bichler, G. Bohm, G. Abstreiter, Physica E **13**, 920 (2002).

[40] J. Höntschel, R. Stenzel, W. Klix, F. Ertl, T. Asperger, R. A. Deutschmann, M. Bichler, and G. Abstreiter, in: Simulation of Semiconductor Processes and Devices 2001, Springer Conference Proceedings, eds.: D. Tsoukalas, and C. Tsamis, Springer Berlin, pp. 222 (2001).

[41] D. Boese, M. Governale, A. Rosch, U. Zülicke, Physica E **12**, 730 (2002).

[42] J. P. Lu, J. B. Yau, S. P. Shukla, M. Shayegan, L. Wissinger, U. Rössler, R. Winkler, Phys. Rev. Lett. **81**, 1282 (1998)

[43] D. Sprinzak, M. Heiblum, Y. Levinson, and Hadas Shtrikman, Phys. Rev. B **55**, R10185 (1997)

[44] R. K. Hayden, D. K. Maude, L. Eaves, E. C. Valadares, M. Henini, F. W. Sheard, O. H. Hughes, J. C. Portal, and L. Cury, Phys. Rev. Lett. **66**, 1749 (1991).

[45] R. R. Marquardt, D. A. Collins, Y. X. Liu, D. Z.Y. Ting, T. C. McGill, Phys. Rev. B 53, 13624 (1996).

[46] T. Bryllert, M. Borgstrom, T. Sass, B. Gustafson, L. Landin, L. E. Wernersson, W. Seifert, and L. Samuelson, Appl. Phys. Lett. **80**, 2681 (2002).

[47] M. Borgstrom, T. Bryllert, T. Sass, B. Gustafson, L. E. Wernersson, W. Seifert, and L. Samuelson, Appl. Phys. Lett. **78**, 3232 (2001).

[48] http://www.wsi.tum.de/nextnano3/

[49] www.intel.co/labs (2001)

[50] www.amd.com/us-en/ (2002)

[51] Y. Naveh and K. Likharev, IEEE Electron Device Lett. **21**, 242 (2000)

[52] P. Solomon and S. Laux, Tech. Dig. - Int. Electron Devices Meet., 95 (2001)

[53] J. Knoch, B. Lengeler, and j. Appenzeller, IEEE Trans. Electron Devices **49**, 1212 (2002)

[54] http://falcon.ecn.purdue.edu:8080/mosfet/10nmstructure.pdf

[55] W. Pötz, J. Appl. Phys. **66**, 2458 (1989)

[56] A. Rahman, J. Guo, S. Datta, and M.S. Lundstrom, IEEE Trans. Electron Devices **50**, 1853 (2003)

[57] Z. Ren, R. Venugopal, S. Goasguen, S. Datta, and M.S. Lundstrom, IEEE Trans. Electron Devices **50**, 1914 (2003)

[58] R. Venugopal, Z. Ren, and M.S. Lundstrom, IEEE Trans. Nanotechnology **2**, 135 (2003)

[59] Gerhard Klimeck, Roger Lake, R. Chris Bowen, William R. Frensley, and Ted S. Moise, Appl. Phys. Lett. **67**, 2539 (1995)

[60] S. Hackenbuchner, *Elektronische Struktur von Halbleiter-Nanobauelementen im thermodynamischen Nichtgleichgewicht*, ISBN 3-932749-48-0

[61] A. Trellakis, A. J. Galick, and U. Ravaioli, Solid-State Electronics **41**, 771 (1997)

[62] H. M. Antia, Astrophysics Journal Supplement Series **84**, 101 (1993)

[63] A. Trellakis, A.T. Galick, U. Ravaioli, J.H. Arends, Y. Saad, J. Appl. Phys. **81**, 3461 (1997)

[64] R. Venugopal, M. Paulsson, S. Goasguen, S. Datta, and M.S. Lundstrom, J. Appl. Phys. **93**, 5613 (2003)

[65] A.S. Spinelli, R. Clerc, and G. Ghibaudo, IEEE Trans. Electron Devices **49**, 1314 (2002)

[66] http://www.caam.rice.edu/software/ARPACK/

[67] Roger Lake, and Supriyo Datta, Phys. Rev. B **45**, 6670 (1992)

[68] S.-C. Lee, and A. Wacker, Phys. REv. B **66**, 245314 (2002)

[69] C. Texier, and M. Büttiker, Phys. Rev. B **62**, 7254 (2000)

[70] P.M. Petroff, and S.P. DenBaars, *Superlattices and Microstructures* **15**, 15 (1994)

[71] O. B. Shchekin, G. Park, D. L. Huffaker, and D. G. Deppe, Appl. Phys. Lett. **77**, 466 (2000)

[72] H. Saito, K. Nishi, and S. Sugou, Appl. Phys. Lett. **78**, 267 (2001)

[73] A. Ekert and R. Jozsa, Rev. Mod. Phys. **68**, 733 (1996)

[74] D. Loss and D. P. DiVincenzo, Phys. Rev. A **57**, 120 (1998)

[75] M. S. Sherwin, A. Imamoglu, and T. Montroy, Phys. Rev. A **60**, 3508 (1999)

[76] A. Imamoglu, D. D. Awschalom, G. Burkard, D. P. DiVincenzo, D. Loss, M. Sherwin, and A. Small, Phys. Rev. Lett. **83**, 4204 (1999)

[77] 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)

[78] H. Kamada, H. Gotoh, J. Temmyo, T. Takagahara, and H. Ando, Phys. Rev. Lett. **87**, 246401 (2001)

[79] H. Htoon, T. Takagahara, D. Kulik, O. Baklenov, A. L. Holmes Jr., and C. K. Shih, Phys. Rev. Lett. **88**, 087401 (2002)

[80] A. Zrenner, E. Beham, S. Stufler, F. Findeis, M. Bichler, and G. Abstreiter, Nature London **418**, 612 (2002)

[81] X. Li, Y. Wu, D. G. Steel, D. Gammon, T. H. Stievater, D. S. Katzer, D. Park, C. Piermarocchi, and L. J. Sham, Science **301**, 809 (2003)

[82] K. Brunner, U. Bockelmann, G. Abstreiter, M. Walther, G. Böhm, G. Tränkle, and G. Weimann, Phys. Rev.Lett. 69, 3216 (1992).

[83] M. Grundmann, O. Stier, and D. Bimberg, Phys. Rev. B **52**, 11969.(1995)

[84] M. Grundmann, O. Stier, and D. Bimberg, Phys. Rev. B **59**, 5688 (1999)

[85] C. Pryor, J. Kim, L. W. Wang, A. J. Williamson, and A. Zunger, J. Appl. Phys. **83**, 2548 (1998)

[86] C. Pryor, Phys. Rev. B **60** (1999)

[87] Seungwon Lee, Lars Jönsson, John W. Wilkins, Garnett W. Bryant, and Gerhard Klimeck, Phys. Rev. B **63**, 195318 (2001)

[88] A. Francesschetti, H. Fu, L. W. Wang, and A. Zunger, Phys. Rev. B **60**, 1819 (1999)

[89] J.J. Finley, M. Sabathil, R. Oulton, A. I. Tartakovskii, D. J. Mowbray, M. S. Skolnick, S. Liew, M. Migliorato, M. Hopkinson, P. Vogl, Physica E **21**, 199 (2004)

[90] M. Povolotskyi, A. Di Carlo, P. Lugli, S. Birner, P. Vogl, IEEE Transactions on Nanotechnology **3** (1), 124 (2004)

[91] M. Sabathil, S. Hackenbuchner, S. Birner, J. A. Majewski, P. Vogl, J.J. Finley, phys. stat. sol. (c) **0**, 1181 (2003)

[92] Ph. D. thesis of Stefan Hackenbuchner

[93] S. Hackenbuchner, M. Sabathil, J. A. Majewski, G. Zandler, P. Vogl, E. Beham, A. Zrenner, P. Lugli, Physica B **314**, 145 (2002)

[94] 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. Hopkinson, M. Al-Khafaji, J.P.R. David, A.G. Cullis, G. Hill, and J.C. Clark, Phys. Rev. Lett. **84**, 733 (2000)

[95] N. Liu, J. Tersoff, O. Baklenov, A.L. Holmes, Jr., and C.K. Shih, Phys. Rev. Lett. **84**, 334 (2000)

[96] M. A. migliorato, A. G. Cullis, M. Fewrn, and J. H. Jefferson, Phys. Rev. B **65**, 115316 (2002)

[97] T. Walther, A. G. Cullis, D. J. Norris, and M. Hopkinson, Phys. Rev. Lett. **86**, 2381 (2001)

[98] M. A. Cusack, P. R. Briddon, and M. Jaros, Phys. Rev. B **56**, 4047 (1997)

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

[100] C. G. van de Walle, and R. M. Martin, Phys. Rev. Lett. **62**, 2028 (1989)

[101] A. Zunger, and S. Wei, Appl. Phys. Lett. **72**, 2011 (1998)

[102] Peter Y. Yu, Manuel Cardona, *'Fundamentals of Semiconductors'*, Spinger-Verlag Berlin, Heidelberg, New York, ISBN 3-540-61461-3, (1996)

[103] Matthias Sabathil, Diplomarbeit, Universität Konstanz (2001)

[104] J. Shumway, L.R.C. Fonseca, J.P. Leburton, Richard M. Martin, D.M. Ceperley, Physica E **8**, 260 (2000)

[105] Gustavo A. Narvaez, and Pawel Hawrylak, Phys. Rev. B **61**, 13753 (2000)

[106] Pawel Hawrylak, Physica E **11**, 53 (2001)

[107] A. Barenco, and M.A. Dupertuis, Phys. Rev. B **52**, 2766 (1995)

[108] Arno Hartmann, Yann Ducommun, Eli Kapon, Ulrich Hohenester, and Elisa Molinari, Phys. Rev. Lett. **84**, 5648 (2000)

[109] O. Stier, A. Schliwa, R. Heitz, M. Grundmann, and D. Bimberg, phys. Stat. Sol. (b) **224**, No. 1, 115 (2001)

[110] V. Türck, S. Rodt, R. Heitz, O. Stier, M.Strassburg, U. W. Pohl, and D. Bimberg, phys. Stat. Sol (b) **224**, No. 1, 217 (2001)

[111] L.R.C. Fonseca, J. L. Jimenez, J.P. Leburton, and Richard M. Martin, Phys. Rev. B **57**, 4017 (1998)

[112] Gerhard Klimeck, Fabio Oyafuso, R. Chris Bowen, *et al.* Superlattices and Microstructures, Vol. **31**, Nos 2-4, p. 171 (2002)

[113] J. Shumway, A. Franceschetti, and Alex Zunger, Phys. Rev. B **63**, 155316 (2001)

[114] Gabriel Bester, and Alex Zunger, to be published

[115] Till Andlauer, Diploma thesis, to appear in December 2004

[116] P. Hohenberg, and W. Kohn, Phys. Rev. **136**, B864 (1964)

[117] W. Kohn, and L. J. Sham, Phys. Rev. **140**, A1133 (1966)

[118] P. Vogl, lecture notes to *Quantum mechanics II*, 29 (2003)

[119] J. P. Perdew, and Alex Zunger, Phys. Rev. B **23**, 5048 (1981)

[120] A. Barenco, D. Deutsch, A. Ekert, R. Jozsa, Phys. Rev. Lett, **74**, 4083, (1995)

[121] A. Imamoglu, D.D. Awschalom, G. Burkard, D.P. DiVincenzo4, D. Loss, M. Sherwin, and A. Small, Phys. Rev. Lett. **83**, 4204, (1999)

[122] N. Gisin, G. Ribordy, W. Tittel, and H. Zbinden, Rev. Mod. Phys. **74**, 145, (2002) and references therin.

[123] P. Hawrylak. Phys. Rev. B **60**, 5597, (1999), Narvaez, P. Hawrylak, Phys. Rev. B **61**, 13753, (2000)

[124] R. J. Warburton *et al*, Nature **405**, 926, (2000), J.J. Finley *et al*, Phys. Rev. B **63**, 073307 (2001), F. Findeis *et al*, Phys. Rev. B **63**, 121309, (2001)

[125] B. Urbaszek,1 R.J. Warburton, K. Karrai, B.D. Gerardot, P.M. Petroff, and J.M. Garcia, Phys. Rev. Lett. **90**, 247403, (2003)

[126] R. Oulton, J.J. Finley, A.D. Ashmore, I.S. Gregory, D.J. Mowbray, and M.S. Skolnick, Phys. Rev. B **66**, 045313 (2002)

[127] The sign convention adopted in this work is chosen such that positive F_z is along the growth direction, i.e. orientated from base to tip of the QD.

[128] J.J. Finley, A.D. Ashmore, A. Lemaître, D.J. Mowbray, M.S. Skolnick, I.E. Itskevich, P.A. Maksym, M. Hopkinson, and T.F. Krauss, Phys. Rev. B **63**, 073307 (2001)

[129] T. Walther, A.G. Cullis, D.J. Norris, and M. Hopkinson, Phys. Rev. Lett. **86**, 2381, (2001) and D.M. Bruls *et al* Appl. Phys. Lett. **81**, 1708,(2002)

[130] M. Baier, F. Findeis, A. Zrenner, M. Bichler, and G. Abstreiter, Phys. Rev. B **64**, 195326, (2001)

[131] D.V. Regelman, E. Dekel, D. Gershoni, E. Ehrenfreund, A.J. Williamson, J. Shumway, A. Zunger, W.V. Schoenfeld and P.M. Petroff, Phys. Rev. B **64**, 165301 (2001)

[132] R.J. Warburton, C. Schulhauser, D. Haft, C. Schäflein, K. Karrai, J.M. Garcia, W. Schoenfeld, and P.M. Petroff, Phys. Rev. B **65**, 113303, (2002)

[133] J. A. Barker and E.P. O'Reilly, Phys. Rev. B **61**, 13840 (2000)

[134] S. Hackenbuchner, M. Sabathil, J.A. Majewski, G. Zandler, P. Vogl, E. Beham, A. Zrenner, P. Lugli, Physica B **314**, 145 (2002)

[135] M. Sabathil, S. Hackenbuchner, J. A. Majewski, G. Zandler, P. Vogl, Journal of Computational Electronics **1**, 81 (2002)

[136] N. Liu, J. Tersoff, O. Baklenov, A.L. Holmes, Jr., and C.K. Shih, Phys. Rev. Lett **84**, 334, (2000)

[137] T.H. Stievater, Xiaoqin Li, D.G. Steel, D. Gammon, D.S. Katzer, D. Park, C. Piermarocchi, and L.J. Sham, Phys. Rev. Lett. **87**, 133603 (2001)

[138] A. Zrenner, E. Beham, S. Stufler, F. Findeis, M. Bichler, and G. Abstreiter, Nature (London) **418**, 612 (2002)

[139] H. Htoon, T. Takagahara, D. Kulik, O. Baklenov, A.L. Holmes, Jr., and C.K. Shih, Phys. Rev. Lett. **88**, 087401 (2002)

[140] Oliver Gywat, Guido Burkard, and Daniel Loss, Phys. Rev. B **65**, 205329 (2002)

[141] Xiaoqin Li, Yanwen Wu, Duncan Steel, D. Gammon, T.H. Stievater, D.S. Katzer, D. Park, C. Piermarocchi, and L.J. Sham, Science **301**, 809 (2003)

[142] Eliana Biolatti, Rita C. Iotti, Paolo Zanardi, and Fausto Rossi, Phys. Rev. Lett. **85**, 5647 (2000)

[143] Gabriel Bester, J. Shumway, and Alex Zunger, Phys. Rev. Lett. **93**, 047401 (2004)

[144] P. Borri, W. Langbein, S. Schneider, and U. Woggon, R.L. Sellin, D. Ouyang, and D. Bimberg, Phys. Rev. B **66**, 081306(R) (2002)

[145] Qianghua Xie, Anupam Madhukar, Ping Chen, and Nobuhiko P. Kobayashi, Phys. Rev. Lett. **75**, 2542 (1995)

[146] J. Tersoff, C. Teichert, and M.G. Lagally, Phys. Rev. Lett. **76**, 1675 (1996)

[147] P. Borri, W. Langbein, U. Woggon, M. Schwab, and M. Bayer,S. Fafard, Z. Wasilewski, and P. Hawrylak , Phys. Rev. Lett. **91**, 267401 (2003)

[148] M. Bayer, P. Hawrylak, K. Hinzer, S. Fafard, M. Korkusinski, Z.R. Wasilewski, O. Stern, and A. Forchel, Science **291**, 451 (2001)

[149] G. Ortner, M. Bayer, A. Larionov, V.B. Timofeev, A. Forchel, Y.B. Lyanda-Geller, T.L. Reinecke, P. Hawrylak, S. Fafard, and Z. Wasilewski, Phys. Rev. Lett. **90**, 086404 (2003)

[150] W. Sheng and J. Leburton Phys. Rev. Lett. **88**, 167401 (2002)

[151] M. Korkusinski and P. Hawrylak Phys. Rev. B **63**, 195311 (2001)

[152] M. Korkusinski, P. Hawrylak, M. Bayer, G. Ortner, A. Forchel, S. Fafard and Z. Wasilewski, Physica E **13**, 610 (2002)

[153] M. Bayer, G. Ortner, A. Larionov, V. Timofeev, A. Forchel, P. Hawrylak, K. Hinzer, M. Korkusinski, S. Fafard and Z. Wasilewski, Physica E **12**, 900 (2002)

[154] D. M. Ceperley and B. J. Alder, Phys. Rev. Lett **45**, 566 (1980)