Highly Selective Gas Sensors using GaN Nano Wires

Dissertation for the acquisition of the academic degree Doktor der Ingenieurwissenschaften (Dr.-Ing.)

Submitted to the Faculty of Electrical Engineering/Computer Science of the University of Kassel

By Kristian Frank

Disputation: 21.04.2023

Erklärung

Hiermit versichere ich, dass ich die vorliegende Dissertation selbständig, ohne unerlaubte Hilfe Dritter angefertigt und andere als die in der Dissertation angegebenen Hilfsmittel nicht benutzt habe. Alle Stellen, die wörtlich oder sinngemäß aus veröffentlichten oder unveröffentlichten Schriften entnommen sind, habe ich als solche kenntlich gemacht. Dritte waren an der inhaltlichen Erstellung der Dissertation nicht beteiligt; insbesondere habe ich nicht die Hilfe eines kommerziellen Promotionsberaters in Anspruch genommen. Kein Teil dieser Arbeit ist in einem anderen Promotions- oder Habilitationsverfahren durch mich verwendet worden.

.....
Unterschrift

Kassel, 24.09.2022

Acknowledgements

This doctoral thesis was made possible by the support of many colleagues, friends and family members.

First, I would like to thank Prof. Bernd Witzigmann for giving me the opportunity to be a part of his research group and for having an open door whenever support or consultation was needed. Furthermore I want to thank Prof. Hartmut Hillmer for overtaking the supervision of my doctoral thesis after Prof. Witzigmann moved to the University of Erlangen and for the inspiring lectures that made me shift the focus of my studies to the field of semiconductor devices in the first place.

I'd like to thank my colleagues from the CEP group Klaus Mayer, Friedhard Römer, Ugur Akcakoca, Maximilian Bettenhausen, Marko Bjelica, Akshay Shedbalkar, Abdul-Aleem Jamali, Hamed Kamrani, Lukas Wetterau, Soenke Grüssing and Paul Mertin for creating an outstanding working environment and for plenty valuable discussions in the field of physics and beyond. Special thanks to Friedhard Römer and Klaus Mayer for their support whenever there were questions regarding physics or programming.

Next I want to thank Daniel Bischof for helping me set up a reasonable development environment for this thesis and for having a quick solution to any IT issue I encountered. Furthermore I would like to thank Regina Brylla and Alina Bernhardt for helping out in all organizational matters.

Moreover, I'd like to thank Katharina Liebisch for last minute proof reading and for the moral backup on many occasions.

Last but not least I want to thank our project partners from TU Braunschweig and the University of Freiburg for the collegial cooperation throughout the whole project and the Federal Ministry of Education and Research (BMBF) whose financial support of the project "WireControl" enabled this thesis work.

Highly Selective Gas Sensors using GaN Nano Wires

A doctoral thesis by Kristian Frank

Computational Electronics and Photonics group
Department of Electrical Engineering and Computer Sciences
University of Kassel

Wilhelmshöher Allee 71, 34121 Kassel
July 21, 2024

Abstract

Over recent years, a variety of nanostructure gas sensor designs have been published, utilizing structures such as planar high electron mobility transistors (HEMTs) and one-dimensional structures like nanorods, nanowires, and nanotubes. Even though these sensors have demonstrated responses to certain target gases, generally all designs also showed cross-sensitivites to other gases in the same order of magnitude. However, a concept for developing highly selective nanowire sensors that are sensitive to only one specific gas species and meet the requirements for industrial mass production is still missing.

Sensors based on one-dimensional structures are more promising for the detection of low gas concentrations compared to planar designs due to their higher surface-to-volume ratios. Gallium nitride (GaN) is a wide band gap (≈ 3.4 eV) semiconductor material which recently has been focused on as a candidate for high performance gas sensors due to its high carrier saturation velocity, fast electron mobility, and thermal, mechanical, and chemical stability [1].

Prior to the beginning of this project, our project partners from TU Braunschweig and the University of Freiburg presented a highly selective NO_2 nanowire sensor based on tin oxide (SnO_2) [2], showing sensor responses for NO_2 which were several orders of magnitudes higher relative to the responses for six investigated comparative gases. This outstandingly high selectivity was gained by a functionalization of the semiconductor surface with organic self-assembled monolayers (SAMs) and could be explained with the aid of density functional theory (DFT-) simulations. The DFT-calculations revealed that the energetic position of the SAM-gas frontier orbitals in respect to Fermi level of the

nanowire is the key factor to determine if a charge transfer between semiconductor and gas molecule can occur or not. This insight leads to the promising strategy of fabricating sensor structures tailored to be sensitive towards single gas species by choosing suitable organic functionalities and matching of the nanowire-SAM Fermi level [2]. However, sensors based on irregularly grown and bent SnO_2 nanowires don't provide reproducible results and hence prevent a reliable prediction of sensor responses. GaN nanowires are offering the potential to overcome these impediments. Furthermore the choice of GaN as active sensing material is promising for the fabrication of sensors that are long term stable and which can operate over a wide range of temperatures.

This thesis investigates macroscopic device simulations of GaN nanowires in the context of gas sensing applications. To this end, straight nanowires with bottom and top contacts are modeled as the basic device geometry. Various sensor structures, including resistors, pn-diodes, and transistor-like npn-junctions, are investigated to identify optimal sensor designs in terms of structure, doping concentrations, and morphology. Numerical modeling is performed using the proprietary semiconductor drift-diffusion solver Quatra (QUAntized TRAnsport). As Quatra did not support high-field dependencies of carrier mobilities the source code has been extended by a Canali and a Transfered-Electron model as part of this thesis work. In order to account for the surface states introduced by docking gas molecules Quatra supports the incorporation of a Shockley-Read-Hall based surface trap model. The data required for the trap densities and energy trap levels were supposed to be gained through atomistic DFT-simulations carried out at the University of Freiburg. Besides the knowledge of trap charge densities and energy levels, literature research identified an intrinsic Fermi level pinning and the oxidation condition of the nanowire surface as additional key parameters that need to be precisely defined for a reliable prediction of the sensor behavior.

Hochselektive Gas-Sensoren auf der Basis von GaN Nanodrähten

Eine Doktorarbeit von Kristian Frank

Fachgebiet Theorie der Elektrotechnik und Photonik
Fachbereich Elektrotechnik/Informatik
Universität Kassel

Wilhelmshöher Allee 71, 34121 Kassel
July 21, 2024

Zusammenfassung

In den vergangenen Jahren wurden diverse Konzepte für auf Nanostrukturen basierende Gassensoren publiziert. Dabei wurden planare elektronische Bauteile wie High Electron Mobility Transistoren (HEMT) oder eindimensionale Strukturen wie Nanosäulen, Nanodrähte oder Nanoröhrchen zur Detektion bestimmter Gase verwendet. Während zwar immer die Sensitivität der Bauteile bezüglich des bevorzugten Gases demonstriert werden konnte, zeigten sich jedoch in allen Studien Sensitivitäten bezüglich analysierter Vergleichsgase die zwar generell kleiner waren, aber in der gleichen Größenordnung lagen. Es fehlt bisher ein Nanodraht-Sensordesign das dazu geeignet ist gezielt selektiv, reproduzierbar und langzeitstabil eine spezifische Gassorte zu detektieren und gleichzeitig die Anforderung an einen industriellen Produktionsablauf zu erfüllen.

Aufgrund ihres größeren Verhältnisses von aktiver Oberfläche zu Volumen sind Sensorentwürfe die auf eindimensionalen Strukturen basieren vielversprechender für die hochempfindliche Detektion geringer Gaskonzentrationen als planare Konzepte. Als aussichtsreicher Kandidat für die Hochleistungsdetektion bestimmter Gassorten ist Galliumnitrid in letzter Zeit vermehrt in den Fokus gerückt [1]. Galliumnitrid ist ein Halbleiter mit großer Bandlücke (≈ 3.4 eV), der sich durch hohe Elektronen-Sättigungsgeschwindigkeiten, einen hohe Elektronen-Mobilität, sowie durch seine thermische wie mechanische und chemische Stabilität auszeichnet.

Vorab dieses Projekts haben unsere Projektpartner von der TU Braunschweig und der Universität Freiburg eine Studie zu einem hochselektiven NO_2 -Nanodrahtsensor auf der Basis von Zinnoxid (SnO_2) veröffentlicht [2]. Die hier ausgewertete Sensorempfindlichkeit bezüglich NO_2 war um mehrere Größenordnungen höher als für alle sechs untersuchten

Vergleichsgase. Diese herausstechend hohe Selektivität konnte erreicht werden durch eine Funktionalisierung der Oberfläche mit organischen self-assembled monolayers (SAMs). Dieses Verhalten konnte mit Hilfe von Dichtefunktional Theorie (DFT-) Simulationen erklärt werden. Die DFT-Berechnungen zeigten, dass ein Ladungsaustausch zwischen angedocktem Gasmolekül und dem Nanodraht nur möglich ist, wenn die energetische Lage der SAM-Gas Grenzorbital in der Nähe des Fermi-Levels des SAM-Nanodraht-Systems liegt [2]. Mit dieser Erkenntnis sollte es möglich sein auf einzelne Gassorten maßgeschneiderte Nanodrahtsensoren zu entwerfen, die hochselektive und hochempfindliche Sensorempfindlichkeiten aufweisen. Der Nachteil des in [2] präsentierten Sensorentwurfs sind die unregelmäßig und gekrümmt gewachsenen SnO_2 Nanodrähte. Hierdurch ist keine exakte Reproduzierbarkeit der Ergebnisse zwischen verschiedenen Sensoren des gleichen Typs zu erwarten. Hier stellt die Verwendung von GaN anstelle von SnO_2 eine vielversprechende Alternative dar. Darüber hinaus sollte es mit GaN als aktivem Sensormaterial aufgrund seiner Robustheit möglich sein langzeistabile Sensoren zu fertigen, die über eine große Spanne an Temperaturen operieren können.

Der Fokus dieser Arbeit liegt auf makroskopischen Bauteil-Simulationen für GaN Nanodrähte im Anwendungsfeld der Gassensorik. Dafür werden gerade gewachsenen Nanodrähte, die am Boden und der Oberseite elektrisch kontaktiert sind, als grundlegende Bauteilgeometrie angenommen. Hierauf basierend werden verschiedener Sensorstrukturen wie Widerstände, PN-Dioden oder transistorartige NPN-Übergänge modelliert, mit dem Ziel optimale Sensorkonzepte bezüglich Struktur, Dotierkonzentrationen und Morphologie zu identifizieren. Die numerische Modellierung des Ladungstransports wird mit Hilfe des hauseigenen Drift-Diffusions-Lösers Quatra (QUAntized TRAnsport) durchgeführt. Da Quatra zu Beginn der Arbeit über kein Modell für die Abhängigkeit der Ladungsträgermobilitäten bezüglich hoher Feldstärken verfügte, wurde der Quellcode von Quatra um ein Canali-Modell und ein Transfered-Electron-Modell im Rahmen dieser Arbeit erweitert. Um die Wechselwirkungen zwischen Halbleiter und andockenden Gasmölekülen zu berücksichtigen bietet Quatra ein Modell für Oberflächenzustände, welches letztendlich einem Term für Shockley-Read-Hall-Rekombinationen entspricht. Die hierfür benötigten Informationen über die Zustandsdichten und deren zugehörigen Energieniveaus sollten durch atomistische DFT-Simulationen an der Universität Freiburg gewonnen werden. Neben der Kenntnis der Zustandsdichten und Energieniveaus, zeigt die Literaturrecherche, dass auch die Kenntnis über ein mögliches intrinsisches Fermi-Level-Pinning der Nanodrahtoberfläche und deren eventuelle Oxidierung unerlässlich ist, um das Nanodrahtverhalten korrekt zu beschreiben und vorherzusagen.

Contents

1 Introduction

In recent times the influence of harmful gases emitted by our modern industrial society regarding the creation of environmental issues encountered a surging awareness. Reducing the human foot print for climate change, driven by the emission of greenhouse gases and the prevention of diseases caused by pollutants have become key factors to consider for industrial production. This holds for products, like combustion engines in cars as well as for production processes. The automotive industry and vehicles belong to the largest sources of nitrogen oxides (NO_x), such as nitric oxide (NO) or nitric dioxide (NO_2), which are among the most harmful gases affecting human health and they can produce acid rain and ozone [2, 3]. In contrast to the negative repercussions of NO_x emissions, hydrogen (H_2) is one of the most promising candidates to substitute carbon-based fuels, providing a clean and renewable energy source [4]. As hydrogen is higly volatile and inflammable, safety measures like reliable leakage detection become crucial during production, storage and transport [4]. Furthermore, miniature gas sensors could be used in health care for the respiratory monitoring [5]. These examples shall illustrate how the selective detection of specific gases is a critical requirement for several industrial branches [2].
Nanoscaled gas sensors promise to provide several advantages over the currently available macroscaled sensors as they offer the prospective of having sensors with outstanding sensitivities which would enable the detection of tiny amounts of certain target gases. Such sensors could be incorporated in small devices of everydays live like smartphones or smartwatches with the purpose of monitoring health data. Furthermore would this omnipresence of sensors allow large scaled environmental evaluations for a variety of parameters by interpreting every individual measurement as part of a grid information. Further advantages could be cheapness, low power consumption and long term stability. As promising these benefits are, just as high are the scientific hurdles to overcome for the realization of marketable devices. Practical sensors have to work long term stable apart from lab conditions which means outside of a well defined environment. Electrical gas sensors have to chemically react with their surrounding in order to generate electrical signals. This means that a sensor could generally be exposed to any arbitrary gas species, which must not alter the sensor behavior or encourage device degradation. Besides that, not only is it essential that sensor responses are reproducible without regular recalibration and that the sensors must be able to be reactivated within a brief time, the gas species to detect will usually be present as a small part in a mix of a large variety of gas species. Therefore, it is required that these sensors not only provide a large sensor response, but also that the response is highly selective for the target gas. However modern chemo resistive semiconductor gas sensors offer only a high sensitivity for a broad spectrum of gases and hence a rather weak selectivity. Novel organic-inorganic hybrid gas sensors showed to be a promising approach to overcome this issue. In this setting organic functional molecule groups are assembled on the semiconductor surface in order to adjust the gas/functional group highest molecular orbital (HOMO) or lowest unoccupied molecular

orbital (LUMO) energy levels with respect to the semiconductor Fermi level [2]. With this adjustment only the desired gas offers a HOMO or LUMO energy position close to the semiconductor Fermi level so that a charge transfer can occur, for other gases a charge transfer becomes impeded leaving the sensor resistance unchanged. Additionally leads the usage of nanowires instead of thin films to an improvement in sensitivity since nanowires offer a very large surface-to-volume ratio as desired in surface effect dependent sensing.

In this thesis work Gallium Nitride (GaN) nanowires, with surfaces functionalized in the aforementioned manner, acting as gas sensor devices, shall be examined by numerical simulations. GaN is a promising material for nanowire sensing application as it's direct and large band gap of around 3.4 eV enables operations in harsh conditions by offering a wide range thermal and radiative stability. [6] Moreover, its chemical and mechanical robustness provides the ability to produce sensors with reliable and predictable sensor responses. In this way different sensor structures such as pure resistors, resistive core/shell-structures, diodes or transistors are supposed to be analyzed. The simulations shall give insight into the nanowire gas sensors electrical behavior, in order to make the sensor response understandable and furthermore predictable. Eventually these theoretical results shall be used to optimize the gas sensor devices according to parameters like structure, morphology and doping profiles. As a metric for the sensor performance, all structures will be analyzed regarding linearity and responsiveness. Here, the sensor response will mean the sensitivity of the sensor towards a change in target gas concentration and is given by the percentage change of the nanowire current relative to its reference current in air environment[1].

This thesis work was part of the BMBF project "WireControl" and has been carried out in close cooperation with our project partners from TU Braunschweig and the University of Freiburg. Within this project TU Braunschweig was responsible for the production of the GaN nanowire samples including the surface functionalization and for the experimental evaluation. At the University of Freiburg numerical simulations on an atomistic level have been conducted using Density Functional Theory (DFT) calculations. Hereby a deeper understanding of the interactions of certain gas species with the functionalized semiconductor surfaces were meant to be gained. In addition, the DFT calculations should provide estimation approaches for parameters such as trap energy levels and densities to be included in the surface model used in the device simulations presented in this work.

[1] A formula representation for sensor response and linearity is given in section 3.2.

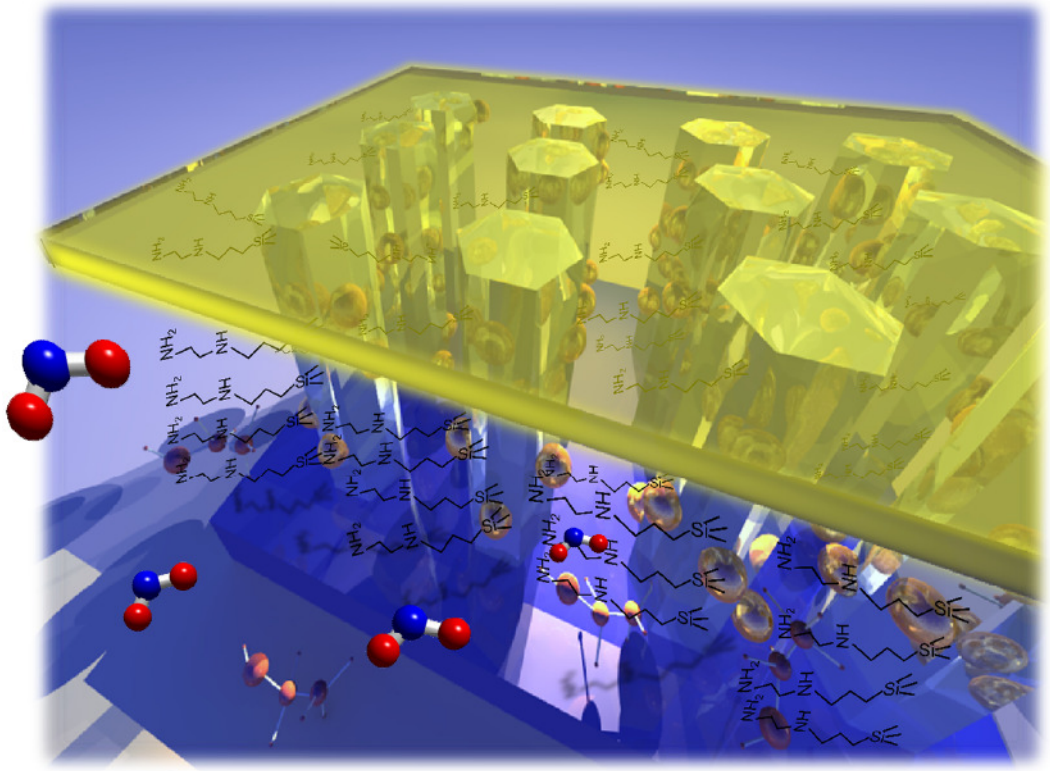

Figure 1.1: Gas sensor using an array of functionalized nanowires. [7]

Fig. 1.1 shows a schematic, which illustrates the general sensing principle of electrical nanowire gas sensors as they are analyzed in this thesis work. Such sensors consists of regularly grown nanowire columns, which are electrically contacted on the top and the bottom, enabling comparative resistance measurements. The nanowire surfaces are functionalized with amine terminated self-assembled monolayers[2] (SAMs). Due to the amines electron donating character the SAMs serve as adsorption sites for the electron affine NO_2 molecules [2]. The energetic conditions of the frontier molecules adsorbed on the surface, relative to the Fermi-Level of the semiconductor-functional group system, determine if the electrical resistance of the nanowires will be affected by the adsorption of the gas molecules or not [2].

[2]N -[3-(Trimethoxysilyl)propyl]ethylenediamine (en-APTAS 1) [2]

1.1 State of the Art

The development of nanostructures serving as high sensitive and selective gas sensors has been an active field of research in recent years. Several different basic approaches for sensor structures such as high-electron mobility transistors (HEMT) [8, 9] or one dimensional nanostructures like nanowires(NWs) [6, 10, 11], nanorods(NRs) [3, 12] and nanotubes(NTs) [13] have been employed. Moreover, gas sensors have been reported based on various semiconductor materials like ZnO, SnO_2 or GaN. Different strategies have been reported to improve sensor sensitivity including the employment of hybrid structures [3], surface treatments [12] or surface functionalization [2, 10]. One dimensional structures have proven to be more suitable for the detection of low gas concentration due to their high surface-to-volume ratio [12]. A common issue of the so far reported nanostructure gas sensors preventing an industrial production is having practical limitations for at least one of the key parameters. A high complexity of the device or the manufacturing process and the use of expensive functionalization materials makes the sensor design cost-intensive. The necessity of high operating temperatures or a weak selectivity towards a single gas species might limit the practical use. Furthermore the sensor responses need to be reproducible and long term stable. Sim et al. [10] reported a NO_2 sensor based on suspended GaN nanowires with a surface functionalized by a Pt-Pd alloy deposition. Here, the Schottky contact forms a depletion region at the metal-semiconductor interface which altered when the sensor was exhibited to certain gases. Thereby the resulting conducting channel in the nanowire could be shrinked or expanded [10]. The Sensor showed a responsivity between 4.35 % at 150 °C and 25.57 % at 350 °C with an almost linear response [10]. The selectivity found to be given by a two- to four times higher response for NO_2 against to two other comparative gases. A NO sensor based on InGaN/GaN multi-quantum well-embedded p-i-n GaN nanorods and a much less complex design based on hydrogen peroxide treated GaN nanorods without intentional doping have been published by Redeppa et al. [3, 12]. The diode based sensor showed responses regarding NO concentrations of 100 ppm of around 30 % and 115 % at 35 °C without and with UV illumination respectively. The response to NO was about three times higher compared to the response regarding NO_2 . The sensor based on the surface treated nanorods showed a response of 30.21 % at an operating temperature of 50 °C, but no data according selectivity has been reported. A sensor based on high crystalline GaN nanowires top contacted via a graphene channel has been presented by Shin et al. [1]. The sensor showed a distinct response of 23 %. A comparison with four other gases resulted in responses between 1.7 % and 5 %. No information was given on how accurate the sensor distinguishes between NO_2 and NO. However, the graphene channel improved the sensitivity regarding NO_2 compared to a reference GaN nanowire sensor from 3.1 % to 23 %. Hoffmann et al. [2] reported a NO_2 sensor based on SnO_2 nanowires with surfaces functionalized by organic self-assembled monolayers (SAMs) offering an outstanding selectivity and sensitivity at room temperature. The sensor showed a response of 2100 % for NO_2 concentrations as little as 0.4 ppm, while the responses to NO and several other comparative gases were two to three orders of magnitudes lower. DFT-calculations revealed that charge

transfer was only possible in case of a suitable alignment of the SAM frontier molecular orbitals with respect to the Fermi level of SAM-nanowire system [2]. The large variety of organic functionalities makes this approach seem promising for the tailored production of nanowire sensors with highly selective responses to certain gases. [2]. However irregularly grown arrays of SnO_2 nanowires with arbitrary nanowire intersections will be an obstacle for reproducibility and hence for predictable sensor responses. Therefore the results presented in [2] have been a motivation for this project as GaN enables the fabrication of arrays with regular and straight grown nanowires suitable for device simulation and promising for long term stability. Arrays of GaN nanorods[3] had been realized by our partners at TU Braunschweig [14] prior to the beginnging of the project. The combination of GaN nanostructures and a surface functionalization by organic SAMs has the potential to overcome the impediments for an industrial manufacturing of highly selective gas sensors. An overview[4] of reported sensor structures in recent years can be found in the following table.

Table 1.1: Characteristics of previously reported nanostructure gas sensors

Author	Year	Structure	Target-Gas (concentration)	Operating Temperature	Light Assisted	Sensor Response
Sim et al. [10]	2013	GaN NW resistors Pt-Pd functionalized	NO_2 (100 ppm)	350 °C	No	25.57 %
Hoffmann et al. [2]	2014	SnO_2 functionalized with organic SAMs	NO_2 (0.4 ppm)	Room Temp.	Yes	2100 %
Abdullah et al. [4]	2014	GaN NWs	H_2 (100 ppm)	Room Temp.	No	127 %
Bishop et al. [15]	2015	BGaN/GaN Superlattice	NO_2 (450 ppm)	250 °C	No	30 %
Reddeppa et al. [3]	2019	InGaN/GaN qw-embedded p-i-n GaN NRs	NO (100 ppm)	35 °C	Yes	115 %
Khan et al. [6]	2020	GaN NWs ZnO functionalized	SO_2 (10 ppm)	Room Temp.	No	12.1 %
Reddeppa et al. [12]	2021	GaN NRs H_2O_2 treated	NO (100 ppm)	50 °C	Yes	30.21 %
Shin et al. [1]	2021	GaN NWs with graphene channel	NO_2 (100 ppm)	Room Temp.	Yes	23 %

[3] with diameters down to 400 nm
[4] non-complete

The sensor response values in table 1.1 represent the peak sensitivity that can be found in the individual publications. In general the time dependence of sensor responses don't show on/off-characteristics, instead peak levels need to be built up by time as chemical reactions between sensor and gas species take place. Furthermore measurements at different operation temperatures show deviating responses.

1.2 Methodology

Throughout this thesis work the electrical behavior of organic-inorganic hybrid gas sensors based on GaN shall be numerically analyzed with the aim of making the sensor behavior predictable and in order to identify optimum sensor concepts. Starting from purely resistive GaN nanowire sensors, different approaches such as diodes, core/shell-structures or transistors shall be investigated. For each case the influence of geometrical dimensions as well as of different doping profiles might be critical parameters on which the studies will focus. Two different sensor approaches are shown exemplary in the figures 1.9 and 1.10.

1.2.1 Sensing Principle

The sensing principles of the devices to be analyzed are based on a resistance change caused by docking gas molecules on the functionalized semiconductor surface introducing either a charge transfer between the two systems or via electrostatic effects. Apart from pure resistors, a PN-junction in reverse bias or a transistor approach could be used, where adsorbed gas molecules lead to a carrier inversion, building up a small conductive channel close to the surface. While diode or transistor based devices should offer a sharper sensor response combined with a lower energy consumption, for sensors based on pure nanowire resistors a higher linearity should be expected.

A visual idea of the sensing principle can be gained from figure 1.1, assuming the depicted nanowires to form an array of resistive sensors. These resistors are electrically contacted on the bottom and the top and carry functional molecule groups on the surface where surrounding gas molecules can dock. The functional groups are used for a Fermi level adjustment in order to provide selectivity regarding certain gas species. Figure 1.2 further elucidates the sensing principle. As can be seen here, gas molecules docking on the functional group increase the nanowire resistance by a charge transfer from the nanowire to the functional group / gas system and the resulting carrier depletion in the semiconductor. The sensor can be reactivated by breaking up the bond using UV-illumination. For other gas species or a different semiconductor - functional group constellation also a charge transfer in the opposite direction would be possible which would result in a reduction of the nanowire resistance. The Fermi level position of the

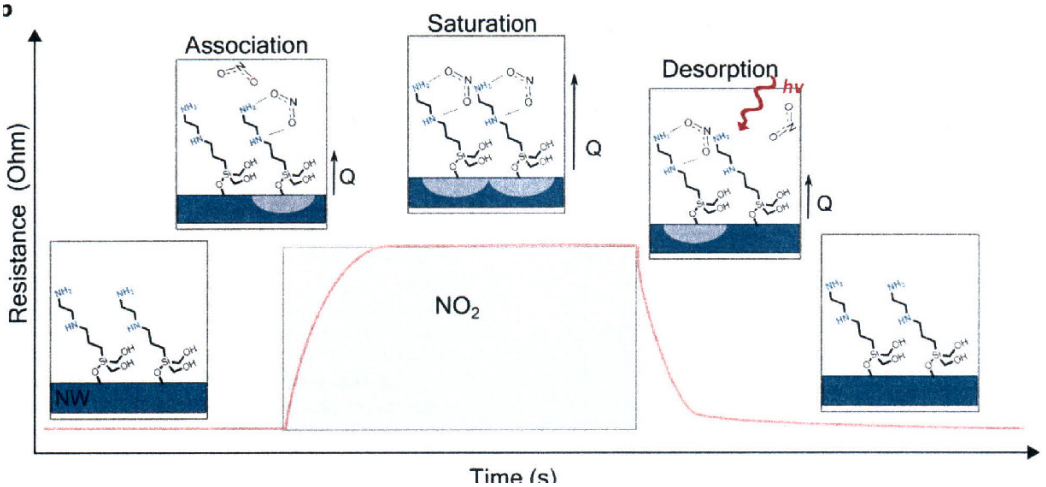

Figure 1.2: Sensing principle [2]

semiconductor - functional group system is the determining parameter for an occurrence and the direction of the charge transfer, as shall be shown by figure 1.3. In the example

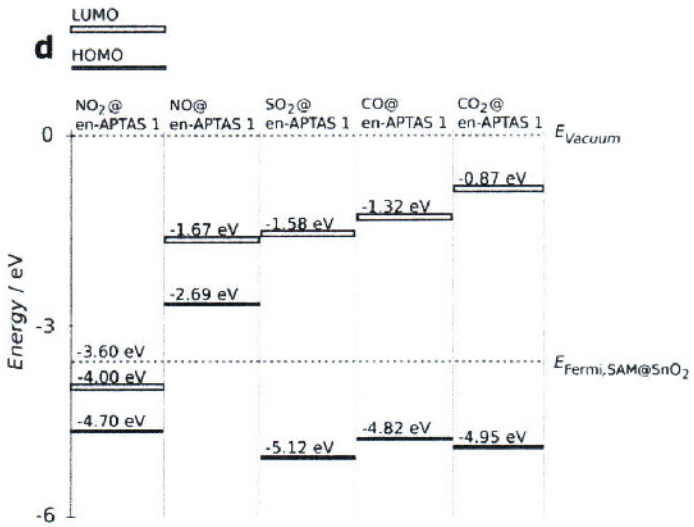

Figure 1.3: Relative position of gas LUMO and HOMO states [2]

shown in figure 1.3 only NO and NO_2 would have triggered as sensor response (with opposite sign), since only for these two gases the Fermi level of the semiconductor - functional group system lies close to the gases HOMO or LUMO respectively. This scenario shows how the Fermi level adjustment is critical for selective sensing.

1.2.2 Modeling

The part of numerical modeling will be undertaken by the use of the proprietary semiconductor drift-diffusion solver Quatra (QUAntized TRAnsport) which implements the Scharfetter-Gummel box method for the discretization of the electron/hole continuity equations and the *Poisson equation*. Together those three equations form a coupled system of non-linear differential equations commonly called the *drift-diffusion model*. When combined with additional models for recombination processes (especially for recombination at surface) and mobility dependencies (e.g. according to temperature, doping and electric field strength), this model exhaustively describes carrier transport in semiconductors without incorporated quantum structures like quantum wells, quantum wires or quantum dots. The drift-diffusion model is given by the continuity equations for electrons and holes and the *Poisson-equation* as follows:

$$\begin{aligned} q\frac{\partial}{\partial t}n &= \nabla\cdot(-q\mu_n n\nabla\phi + qD_n\nabla n) - R \\ -q\frac{\partial}{\partial t}p &= \nabla\cdot(-q\mu_p p\nabla\phi - qD_p\nabla p) + R \\ \nabla\cdot(\varepsilon\nabla\phi) &= -q(p - n + N_\mathrm{D} - N_\mathrm{A} + c_\mathrm{s}). \end{aligned} \tag{1.1}$$

An explicit derivation of the drift-diffusion model will be given in the carrier transport section 2.2 in chapter 2.
For the solution of quantized populations[5] Quatra enables a FEM $\mathbf{k}\cdot\mathbf{p}$- Schrödinger solver. The coupling of bulk and quantized populations is given by the Poisson equation and capture [16].

High-Field Mobility Model

At the beginning of this thesis work carrier mobilities could either be chosen region wise constant or having a doping and temperature dependence according to the Arora model. A new set of Arora parameters has been added to the material file in order to provide an optimum representation for wurtzite GaN low field mobilities. The parameter set originates from Mnatsakanov et al. [17] where a fitting of the Arora model for wurtzite n/p-GaN according to experimental data is given. In order to represent carrier mobilities accurately in the presence of high electric fields, two field dependent mobility models have been implemented in the C++ source code of Quatra. Those new models can be connected with the Arora model via its resulting low field mobility.

[5] Commonly needed for the simulation of optoelectronic devices like LEDs or semiconductor lasers.

The first high field representation of the carrier mobility is a model derived by Canali et al. [18] in 1975. The Canali-model has been developed by fitting a formula for the drift velocity according to experimental data taken from measurements in silicon. Therefore this model is commonly used to calculate the high field mobility in silicon or semiconductors with band structures alike [19]. For the field dependent mobility according to the Canali-model the low field mobility μ_{low} is multiplied with a factor which decays for an increasing electric field strength E

$$\mu(E) = \mu_{\text{low}} \frac{1}{\left(1 + \left(\frac{\mu_{\text{low}} E}{v_{\text{sat}}}\right)^{\beta}\right)^{\frac{1}{\beta}}}. \tag{1.2}$$

Therefore the drift velocity of the carriers v_d, which is given by the product of the mobility and the electric field strength

$$v_d = \mu(E) E \tag{1.3}$$

can be separated into two regimes.

In the low field regime the term $\left(\frac{\mu_{\text{low}} E}{v_{\text{sat}}}\right)^{\beta}$ is negligible, hence the mobility is field independent and simply given by the low field mobility.

$$v_d \approx \mu_{\text{low}} E \tag{1.4}$$

The drift velocity is directly proportional to the electric field strength, resistors driven in that regime behave according to the *Ohmic law*.

In the high field regime the previously neglected term dominates the denominator of eq. (1.2). In this case the drift velocity approaches the saturation velocity v_{sat}

$$v_d \approx v_{\text{sat}}. \tag{1.5}$$

In figure 1.4 the dependence of the drift velocity on the electric field strength according to the canali model is illustrated. The interval in the double logarithmic plot where the slope of the curve is one corresponds to the linear low field regime. After the linear regime the drift velocity saturates against the value of v_{sat}

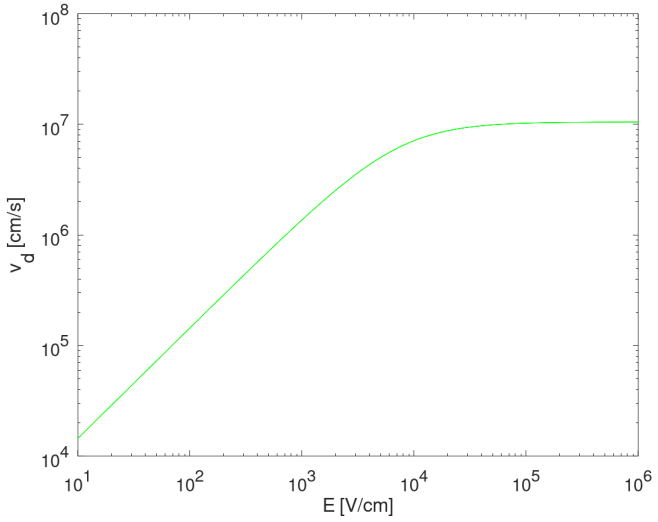

Figure 1.4: The drift velocity depicted over the electric field strength calculated with the Canali-Model. Example values for silicon at 300 K taken from [18]. $\beta = 1.3$, $\mu_{\text{low}} = 1450\ \frac{\text{cm}^2}{\text{Vs}}$, $v_{\text{sat}} = 1.05 \cdot 10^7\ \frac{\text{cm}}{\text{s}}$

The simplicity of the Canali formulation doesn't allow more complicated drift velocity-field characteristics as needed for an accurate fitting of the mobilities in GaN. Drift velocity data obtained by Monte Carlo simulations [20] show a pronounced peak well above the saturation velocity followed by a negative differential velocity in III/V-compound semiconductors [21]. Therefore the so-called Transfered-Electron model [19, 21–23] is commonly used for those kind of semiconductors. The here given extended form of the Transfered-Electron model has been suggested by several studies for an optimum fitting for wurtzite GaN [21, 23, 24]

$$\mu(E) = \frac{\mu_{\text{low}} + \mu_1 \left(\frac{E}{E_0}\right)^{\alpha} + \frac{v_{\text{sat}}}{E_1} \left(\frac{E}{E_1}\right)^{\beta-1}}{1 + \gamma \left(\frac{E}{E_0}\right)^{\alpha} + \left(\frac{E}{E_1}\right)^{\beta}}. \tag{1.6}$$

In figs. 1.5 and 1.6 the Transfered Electron model is compared to the Canali model evaluated for wurtzite GaN. The double logarithmic plot in 1.5 gives a good estimate of the linear regime of the mobility, while the linear plot in 1.6 elucidates the characteristics of the peak regime. The comparison reveals how the strictly increasing Canali model fails to represent the velocity peak while predicting a more extended linear regime with the parameter set given. The Canali model also shows a much earlier saturation (below 200 $\frac{\text{kV}}{\text{cm}}$ compared to 600 $\frac{\text{kV}}{\text{cm}}$ of the TE-Model). Both models are in a good agreement for field strengths below 40 $\frac{\text{kV}}{\text{cm}}$. All parameter values for figs. 1.5 and 1.6 have been taken

from [20] except for the the β-parameter of the Canali model (β_{canali}) and the low field mobility[6] which originates from [21].

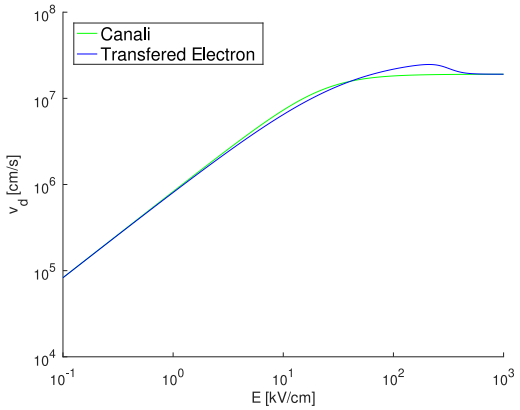

Figure 1.5: Double logarithmic plot of the drift velocity-field characteristic. Comparison of the Canali and the Tranfered Electron model for wurtzite GaN. Parameter set-up: $\mu_{\text{low}} = 830\ \frac{\text{cm}^2}{\text{Vs}}$, $\mu_1 = 0$, $\alpha = 1$, $\beta = 7.2$, $\beta_{canali} = 1.7$, $\gamma = 6.2$, $v_{\text{sat}} = 1.9 \cdot 10^7\ \frac{\text{cm}}{\text{s}}$

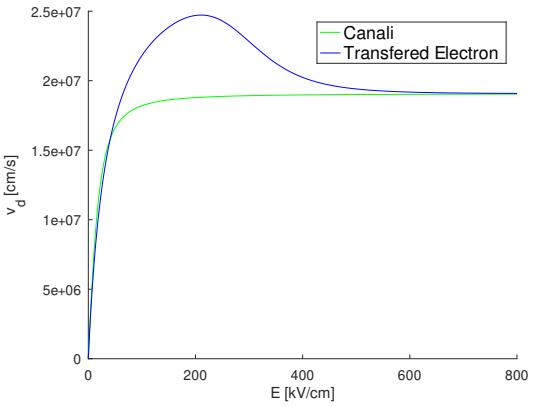

Figure 1.6: Linear plot of the drift velocity-field characteristic. Like in fig. 1.5 again a comparison of the Canali and the Tranfered Electron model for wurtzite GaN using the same parameter set-up. Here the focus is put on the hundred $\frac{\text{kV}}{\text{cm}}$ regime.

Surface Model

In order to use Quatra to simulate the behavior of a gas sensor, the changes on the electronic structure introduced by docking gas molecules at the nanowire surface need to be represented in the drift-diffusion model eq. (1.1). For this purpose Quatra provides a surface trap model that takes care of charge conservation and the fact that charges act as sources of electric fields equally. On the one hand this is realized by a recombination term based on a Shockley-Read-Hall model [25, 26] entering the continuity equations in the drift-diffusion model [27–29]

$$\begin{aligned} R_{\text{s}} &= \frac{v_{\text{sn}} v_{\text{sp}} (np - n_{\text{T0}} p_{\text{T0}})}{v_{\text{sn}}(n + n_{\text{T0}}) + v_{\text{sp}}(p + p_{\text{T0}})} \\ &= N_{\text{s}} \frac{C_{\text{n}} C_{\text{p}} (np - n_{\text{i}}^2)}{C_{\text{n}}(n + n_{\text{T0}}) + C_{\text{p}}(p + p_{\text{T0}})}. \end{aligned} \tag{1.7}$$

Here n, p are the local carrier densities and n_{i} is the intrinsic carrier density, which according to the mass-action-law [30] is given by the square root of the product of

[6] Assuming a doping concentration of $7 \cdot 10^{16}$ cm^3 at 300 K.

the equilibrium carrier densities. N_s represents the density of the surface traps. The expressions

$$\begin{aligned} v_\mathrm{sn} &= N_\mathrm{s} C_\mathrm{n} \\ v_\mathrm{sp} &= N_\mathrm{s} C_\mathrm{p} \end{aligned} \tag{1.8}$$

are the surface recombination velocities of electrons and holes respectively, with C_n and C_p being capture parameters given by the product of the thermal velocity and the respective capture cross sections

$$\begin{aligned} C_\mathrm{n} &= v_\mathrm{th} \sigma_\mathrm{n} \\ C_\mathrm{p} &= v_\mathrm{th} \sigma_\mathrm{p}. \end{aligned} \tag{1.9}$$

Here C_n represents the probability that an electron from the conduction band is captured by a trap while C_p stands for the probability of a hole captured from valence band. The densities n_T0 and p_T0 can be expressed by

$$\begin{aligned} n_\mathrm{T0} &= n_\mathrm{i} \exp(\frac{E_\mathrm{T} - E_\mathrm{i}}{k_\mathrm{B} T}) \\ p_\mathrm{T0} &= n_\mathrm{i} \exp(\frac{E_\mathrm{i} - E_\mathrm{T}}{k_\mathrm{B} T}) \end{aligned} \tag{1.10}$$

where E_T and E_i are the regarding trap and the intrinsic energy levels [28]. Those densities can be seen as carrier densities when the Fermi level would fall together with the trap energy level [25].

On the other hand one has to consider that an occupied acceptor trap level generates a negative charge while an unoccupied donor trap level generates a positive charge[7]

$$\begin{aligned} c_\mathrm{sA} &= -N_\mathrm{sA} f_\mathrm{T} \\ c_\mathrm{sD} &= N_\mathrm{sD} (1 - f_\mathrm{T}), \end{aligned} \tag{1.11}$$

where $f_\mathrm{T} \in [0, 1]$ denotes the trap level occupancy factor. In thermal equilibrium the neutrality condition demands that the surface charges are compensated by counter charges in the semiconductor bulk. Due to this charge accumulation a gradient of the potential is introduced in the surfaces region, leading to a potential difference between the surface and the inner bulk material. This effect is taken care of by Quatra's trap charge model by offering an occupancy factor in correspondence to eq. (1.7) yielding trap charge densities of

$$\begin{aligned} c_\mathrm{sA} &= -N_\mathrm{sA} \frac{C_\mathrm{n} n + C_\mathrm{p} p_\mathrm{T0}}{C_\mathrm{n}(n + n_\mathrm{T0}) + C_\mathrm{p}(p + p_\mathrm{T0})} \\ c_\mathrm{sD} &= N_\mathrm{sD} \frac{C_\mathrm{p} p + C_\mathrm{n} n_\mathrm{T0}}{C_\mathrm{n}(n + n_\mathrm{T0}) + C_\mathrm{p}(p + p_\mathrm{T0})} \end{aligned} \tag{1.12}$$

[7] The surface charges densities are given normalized to the elementary charge here, since the elementary charge is already considered as a factor in 1.1.

and entering them via the surface charge term c_s into *Gauss's law*[8].

In addition to this trap level induced surface charge (or also used singularly) an arbitrary constant charge can be impressed at every surface. Quatra understands a surface to be every intersection of two different materials [27]. Such a constant surface charge can for example be used to model ions that are docking at the nanowires, and hence introducing a charge to the surface, but do not participate in any kind of charge transfer. This would be a common scenario for gas molecules docking at oxidized surfaces, where the oxide layer prevents any charge transfer. Altogether the surface charge term to be considered in *Gauss's law* is then given by [29]

$$c_\mathrm{s} = c_\mathrm{sA} + c_\mathrm{sD} + c_\mathrm{const}. \tag{1.13}$$

As can be seen from eqs. (1.7)–(1.12) the necessary parameters to describe the model completely are the trap densities $N_{A,s/D,s}$, their corresponding trap levels E_T and the electron and hole capture parameters $C_{n/p}$. These trap charge expressions go along very well with the charge transfer mechanisms according to an alignment of HOMO/LUMO states with the semiconductor Fermi level as postulated by Hoffmann et al. [2]. Therefore the given surface trap model gives rise to be a promising approach for the representation of the surface effects to be analyzed.

However the needed information for total trap density and energy trap levels in the semiconductor-amine-gas system is still missing.

To this end atomistic Density Functional Theory (DFT) calculations might be needed in order to determine the energetic positions of the gas molecules HOMO and LUMO when docked to the functional group at the semiconductor surface. Subsequently these results could be incorporated in the above trap model to simulate the behavior of the sensor device. Figure 1.7 shows an exemplary result for the density of states for gained by DFT calculations carried out at the Fraunhofer IWM Freiburg for ZnO nanowire sensors. The gaps in the valence band arise from OH-groups on top and on the bottom of the semiconductor as can be seen on the right side of fig. 1.8. The states introduced by docking gas molecules can be modeled via a fixed energy level (Dirac shaped distribution) see 1.8. Moreover DFT calculations enable the determination of charge transfer inbetween the semiconductor/functional group/gas molecule system.

In fig. 1.8 the negative region on the z-axis represents the semiconductor material. Hence integration over this region yields the net charge transfer from the semiconductor material to the SAM/gas system for a single molecule docked to the system. Nonetheless this information is insufficient to describe the trap model as the probability for the capture of gas molecules by the SAMs as well as the feasible SAM density remain uncertain. Therefor the total trap density has to be fitted against experimental results.

[8] Which is given by the third equation of 1.1

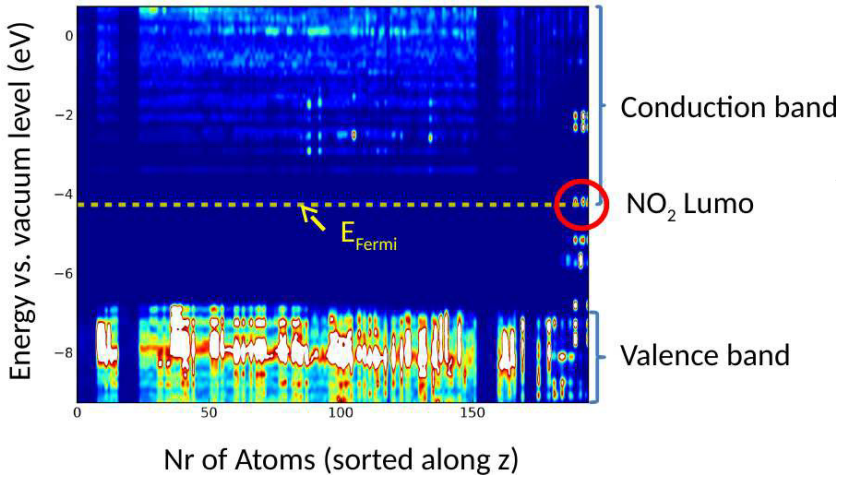

Figure 1.7: Result of DFT calculations for the density of states. Here exemplarily shown for a system of ZnO, functional group and NO_2. [31]

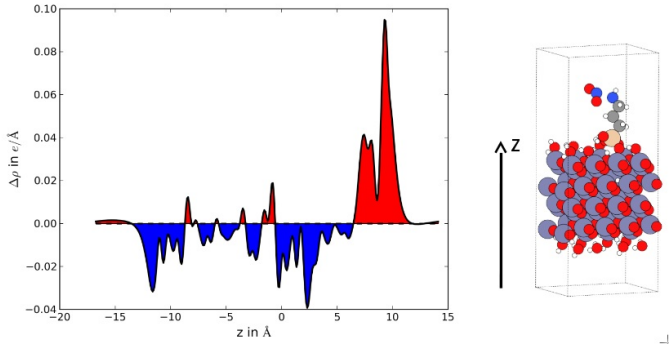

Figure 1.8: Charge transfer from ZnO to gas molecule. Red fillings indicate charge accumulation, blue fillings indicate charge depletion [31]

Geometry generation

Geometry development including doping definition and meshing is realized by using Sentaurus Structure Editor which expects script files written in the programming language Scheme as input from which it can create output files in the DF-ISE format that is necessary for usage in Quatra.
In the figures 1.9 and 1.10 an example of a cylindrical nanowire resistor and a NPN transistor shall be shown respectively. In the left picture the surface charge is positioned right at the interface of the semiconductor and the (light blue) air volume, while in the right picture it is placed on top of thin oxide layer (air volume not shown here).

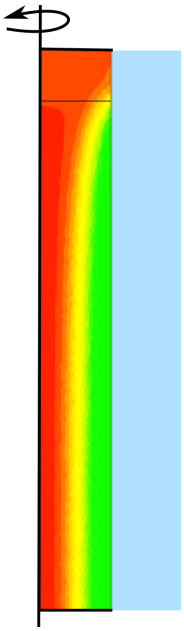

Figure 1.9: Current Distribution in a nanowire sensor

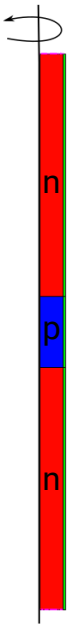

Figure 1.10: NPN transistor with depicted doping profiles

1.3 Challenges

Fermi level pinning at the nanowire surface

Nanowires offer an extraordinarily large surface-to-volume ratio compared to bulk devices. While on one hand this property makes nanowires a promising choice for sensing purposes, on the other hand it strongly demands to put a focus on undesired surface effects throughout the design of such nanowire sensors.

At semiconductor surfaces the electronic structure is varied from that within bulk. This results from the abrupt ending of the crystal periodicity and is due to interactions of surface atoms with their surrounding environment. Dangling bonds originating from unsatisfied surface atoms as well as adsorbed molecules can introduce surface states. Depending on the nature and the densities of these surface states the Fermi level can be pinned at the surface leading to a space charge region in the vicinity of the surface [29, 32, 33].

It is fairly obvious that in common semiconductor manufacturing, crystals in which the Fermi level is determined by surface defects instead of by doping concentrations are highly undesirable, as they would not show the expected semiconducting behavior. For gas sensing purposes on the other hand, devices in which the position of the Fermi level is mainly dependent on surface conditions would be promising to offer extraordinarily high sensing responses. Simple nanowire resistors (or better insulators) for example, could show a transistor-like behavior when additional surface states introduced by docking gas molecules shift the Fermi level towards the conduction band edge, and hence opening up a conducting channel in the nanowire. On the other hand, to sensors that shall offer a high linearity, an inherent Fermi level pinning of the nanowire crystal could be a major obstacle.

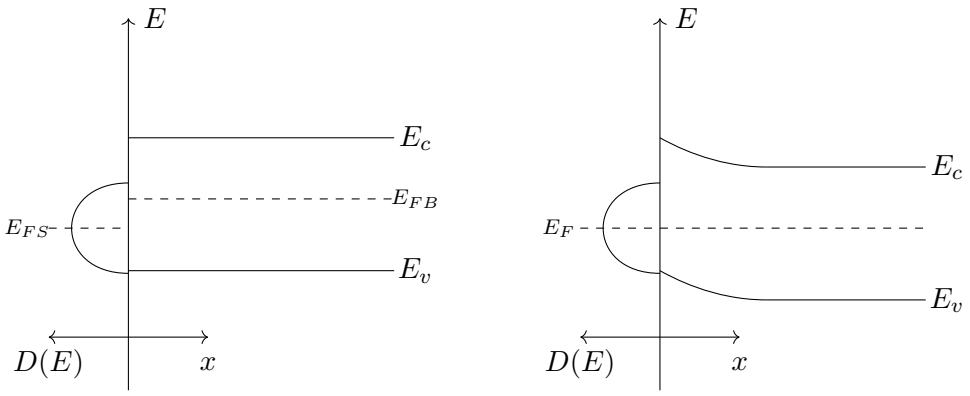

Figure 1.11: When the Fermi level at the surface is pinned due to a high density of surface states the bands in the bulk need to bent in order to equalize the two different Fermi levels. This figure which describes the process qualitatively was taken from [32].

Several experimental and theoretical studies [33, 34] found the Fermi level to be pinned at the GaN m-plane surface while some others found the opposite [35–37]. Thus the literature study gives rise that the electronic structure at the m-plane surface strongly depends on the growth conditions. The following table presents an overview of the reviewed articles regarding Fermi level pinning including their key messages.

Table 1.2: Fermi level pinning for non-polar GaN: Literature overview

Source	**Fermi Level Pinning**	**Method**	**Key Message**
[34, 38]	Yes	DFT	Fermi level gets pinned at 0.5 - 0.7 eV below conduction band minimum by unoccupied gallium dangling bonds
[33]	Yes	IV/ SPC/ PL	Fermi level is pinned at nanowire sidewall / Deoxidation unpins the Fermi level
[35]	No	DFT	Occupied N-derived surface state lies just below valence band maximum and the empty Ga-derived state lies above conduction band minimum
[36]	No	STM/ STS/ DFT	No surface states in the band gap due to dangling bonds. No Fermi level pinning for defect concentrations $< 2 \cdot 10^{12}$ cm^{-2}
[37]	Yes	STM/ STS	N and Ga derived dangling bond surface states are outside the band gap. Fermi level pinning found due to high density of defect states, but no pinning because of surface states

For a deeper understanding of the Fermi level pinning in the context of nanowire sensing section 2.4 provides a theoretical example for n-doped cylindrical nanowires.

With the insights from section 2.4 several consequences of a Fermi level pinning at the surface in respect of sensing purposes can be concluded, given in the following enumeration[9]:

1. Fermi level pinning might limit the minimum diameter for the nanowires when the sensor design is based on permanent current flow. On the other hand nanowires around the critical diameter could be promising to be used as transistor sensors, offering a very large sensor response. Therefore the functionalization of the surface need to work in a way that adsorbed gas molecules influence the critical diameter and hence opening up or closing the inner channel. Sensors based on closed channels[10] are furthermore desirable in respect of energy consumption.

2. The minimum diameter depends on the doping concentration. For example using eq. (2.182) in order to estimate the critical diameter, GaN wires at a pinned surface Fermi level of 0.55 eV below the conduction band minimum[11] show critical diameters of ≈ 80 nm and ≈ 49 nm for an n-doping of $6.25 \cdot 10^{17}$ cm^{-3} and $1.8 \cdot 10^{18}$ cm^{-3} respectively [33].

3. The electric field accompanied by the potential barrier of the space charge region confines electrons in the center of the wire while generated holes would be driven towards the surface. Hence the recombination of electrons with surface traps or holes accumulated at the surface will be impeded.

4. The wire conductivity is not simply proportional to the total diameter but to $(d - d_{critical})^2$.

5. Space-charge-limited currents might occur, as described below in the following subsection.

6. Only for nanowires having a diameter equal or larger than the critical diameter the pinned Fermi level at the surface designates the height of the potential barrier. In wires below the critical diameter the scarcity of volume carriers limits the built up barrier. Instead the Fermi level drops, meaning a reduction of available carriers in the conduction band.

7. Heat dissipation limited to a small channel might have a negative influence regarding device degradation.

This list elucidates the importance of a proper knowledge of the condition of the functionalized in respect of device simulation.

[9] without claims of completeness
[10] to be opened up by docking gas molecules
[11] A value numerically predicted by [34, 38] as well as experimentally found by [33].

Space-charge-limited currents

Another effect found in various studies [33, 39, 41] for thin and low doped nanowires, is that the voltage-characteristics is dominated by space-charge-limited (SCL) currents. In [33, 39] the occurrence of SCL currents was shown to be a consequence of the Fermi level pinning at the surface.

In all studies mentioned above, nanowires having a vertical length of around 1 μm have been used at voltages from 0 to a few volts. Therefore the occurring electric field strengths must have lain well below 10 $\frac{\mathrm{MV}}{\mathrm{m}}$, hence the mobility can be seen independent of the electric field strength. In contrast to an expected ohmic behavior, SCL currents show a quadratic dependency of the applied voltage in this field strength regime [30].

This gives rise to the assumption that sensor designs which desire a linear sensor response[12] demand a sufficiently large combination of wire diameter and doping concentration. Furthermore Fermi level pinning as a source of carrier depletion needs to be considered.

Oxidation of the nanowire surface

During the project phase our partners at TU Braunschweig indicated that the growth of GaN nanowires with pristine surfaces would probably not be feasible. Instead the formation of an oxide layer at the GaN surface was observed for the fabricated samples. If so, the original sensor concept based on charge transfer between gas molecules and semiconductor would not be possible anymore. Therefore a different strategy would have to be found and additional simulations considering the oxide layer had to be carried out accordingly.

Uncertainty factors due to surface functionalization

The trap charge density that is introduced to the semiconductor for certain amounts of gas concentrations, which is needed for the prediction of the sensor response, is a factor of uncertainty. The density of SAMs that can be placed on the nanowire surface and their trapping cross sections regarding certain gas molecules would need to be known. Therefore an iterative fitting procedure in order to match the simulation results with the practical experiments might be necessary. Furthermore the strategy of incorporating energy trap levels taken from results of DFT-simulations does not provide mutual interactions between the nanowire-SAM/SAM-gas systems in a self-consistent manner.

[12] for example for sensors that shall give information on gas concentrations instead of purely detecting the occurrence of certain gases

1.4 Outline

This thesis is organized as follows. In the first part of Chapter 2 basics of crystalline structures are introduced while the second part of Chapter 2 is devoted to the fundamentals of semiconductor physics. Governing equations of carrier transport are explicitly derived from basic axioms leading to the the drift-diffusion-model and recombination processes needed for the surface model. Chapter 3 introduces the simulation set-up and presents the results for different exemplarily chosen experiments. A conclusion of the results is given in chapter 4. The Appendix contains a highlighted overview of the parameters for the new high field mobility models, as well as the used material file with the complete parameter configuration. Furthermore example geometries for the different device structures and a simulation command file can be found here.

2 Fundamentals of Semiconductor Physics

2.1 Semiconductor/GaN Fundamentals

2.1.1 Crystal Structures

Ideal crystals are characterized by a periodically arrangement of identical atomic groups. The periodicity can be described mathematically by an invariance regarding a spatial translation vector

$$\mathbf{T} = \sum_{i=1}^{3} u_i \mathbf{a}_i, \quad u_i \in \mathbb{Z}. \tag{2.1}$$

Here u_i are integers and $\mathbf{a}_i$ represent translation vectors for the i-th primitive base vector, meaning that a particle that is getting dislocated by the translation vector $\mathbf{T}$ finds the exactly same conditions as before the translation. So the physical properties of the crystal can be described by the properties of a primitive cell respecting periodic boundaries. A primitive cell can be defined as the parallelepiped formed by the three base vectors a_i. Figure 2.1 shows an example with the base originated from the center of a lattice point [42].

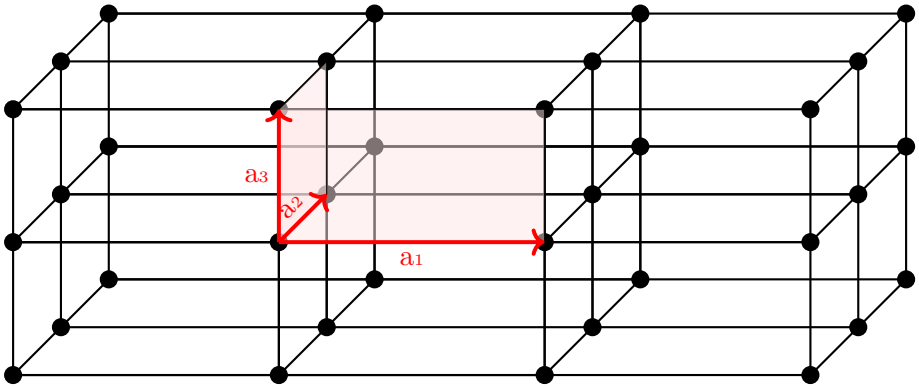

Figure 2.1: Visualization of a primitive cell [42]

The volume of such a primitive cell is then given by the parallelepiped product

$$V_\mathrm{c} = \left|\mathbf{a}_1 \cdot (\mathbf{a}_2 \times \mathbf{a}_3)\right|. \tag{2.2}$$

Wigner-Seitz Cell

Another way to define a primitive cell is the so-called *Wigner-Seitz cell*. Here the volume of the primitive cell is given by the set of points having the least distance to a chosen lattice point compared to any other lattice point. In order to construct such a cell, in a first step, for an arbitrarily chosen atom, one has to draw lines connecting all directly neighboring atoms. Then the Wigner-Seitz cell is found for the smallest volume confined by the intersection of bisector planes drawn perpendicular to each line [30]. The concept is shown in 2.2 in 2D, where the bisector planes decay to bisector lines and the enclosed volume to an area.

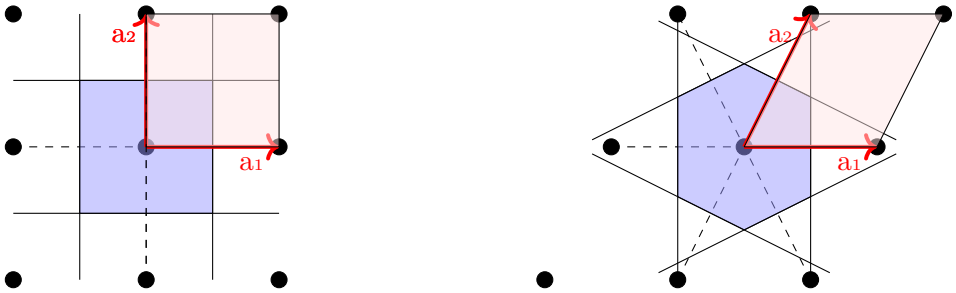

Figure 2.2: Two examples comparing the generation a primitive cell by primitive base vectors (red area) and the corresponding Wigner-Seitz cell (blue area) in 2D.

Just as for the primitive cells derived using primitive base vectors, whole crystal space can be covered by connecting the primitive Wigner-Seitz-Cells.

Reciprocal Lattice

Since any physical property keeps the periodicity of the crystal lattice

$$f(\mathbf{R}) = f(\mathbf{R} + \mathbf{T}), \tag{2.3}$$

it is always possible to be expressed in terms of a *Fourier series expansion*

$$f(\mathbf{R}) = \sum_{G} c_{\mathbf{G}} \exp(\jmath \mathbf{G} \cdot \mathbf{R}). \tag{2.4}$$

Here the amplitudes $c_{\mathbf{G}}$ have to fulfill $c^*_{\mathbf{G}} = c_{-\mathbf{G}}$ for $f(\mathbf{R})$ to be a real function. Besides the identification of the amplitudes, vectors $\mathbf{G}$ have to be found in the way that the

invariance regarding a translation $\mathbf{T}$ is kept. Therefore the primitive base vectors of a reciprocal lattice are defined as

$$\begin{aligned}
\mathbf{b}_1 &= 2\pi \frac{\mathbf{a}_2 \times \mathbf{a}_3}{\mathbf{a}_1 \cdot (\mathbf{a}_2 \times \mathbf{a}_3)} = \frac{2\pi}{V_\mathrm{c}} \mathbf{a}_2 \times \mathbf{a}_3 \\
\mathbf{b}_2 &= 2\pi \frac{\mathbf{a}_3 \times \mathbf{a}_1}{\mathbf{a}_1 \cdot (\mathbf{a}_2 \times \mathbf{a}_3)} = \frac{2\pi}{V_\mathrm{c}} \mathbf{a}_3 \times \mathbf{a}_1 \\
\mathbf{b}_3 &= 2\pi \frac{\mathbf{a}_1 \times \mathbf{a}_2}{\mathbf{a}_1 \cdot (\mathbf{a}_2 \times \mathbf{a}_3)} = \frac{2\pi}{V_\mathrm{c}} \mathbf{a}_1 \times \mathbf{a}_2.
\end{aligned} \tag{2.5}$$

Obviously each reciprocal lattice vector is orthogonal to two of the primitive vectors of the direct lattice

$$\mathbf{b}_i \cdot \mathbf{a}_j = 2\pi \delta_{ij}. \tag{2.6}$$

Now every lattice point in the reciprocal lattice can be expressed by a reciprocal vector

$$\mathbf{G} = \sum_{i=1}^{3} v_i \mathbf{b}_i \tag{2.7}$$

with v_i being integer numbers. Using such a reciprocal lattice vector in the Fourier expansion for a translation in the direct lattice, we get

$$f(\mathbf{R} + \mathbf{T}) = \sum_G c_\mathbf{G} \exp\left(\jmath \mathbf{G} \cdot (\mathbf{R} + \mathbf{T})\right) = \sum_G c_\mathbf{G} \exp(\jmath \mathbf{G} \cdot \mathbf{R}) \exp(\jmath \mathbf{G} \cdot \mathbf{T}) \tag{2.8}$$

where the argument of second exponential function yields a multiple integer of 2π

$$\exp(\jmath \mathbf{G} \cdot \mathbf{T}) = \exp\left(\jmath (2\pi \sum_{i=1}^{3} u_i v_i)\right) = 1 \tag{2.9}$$

since u_i and v_i are given as integers. Hence, we get $f(\mathbf{R} + \mathbf{T}) = f(\mathbf{R})$ as demanded by the lattice periodicity.

For the sake of visualization a simple cubic direct lattice shall be taken as an example.

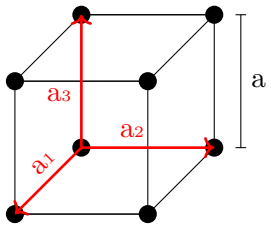

Figure 2.3: Simple cubic lattice [42]

Here, all primitive vectors of the direct lattice have the same length of the lattice constant a

$$\begin{aligned} \mathbf{a}_1 &= a\hat{\mathbf{x}} \\ \mathbf{a}_2 &= a\hat{\mathbf{y}} \\ \mathbf{a}_3 &= a\hat{\mathbf{z}}. \end{aligned} \tag{2.10}$$

The volume of the primitive cell is then given by

$$V_\mathrm{c} = \mathbf{a}_1 \cdot (\mathbf{a}_2 \times \mathbf{a}_3) = a^3 \tag{2.11}$$

and the primitive translation vectors of the reciprocal lattice are given via eq. (2.5)

$$\begin{aligned} \mathbf{b}_1 &= \frac{2\pi}{a}\hat{\mathbf{x}} \\ \mathbf{b}_2 &= \frac{2\pi}{a}\hat{\mathbf{y}} \\ \mathbf{b}_3 &= \frac{2\pi}{a}\hat{\mathbf{z}}, \end{aligned} \tag{2.12}$$

enclosing a likewise cubic volume of the reciprocal cell

$$V_\mathrm{rc} = \frac{(2\pi)^3}{V_\mathrm{c}} = \left(\frac{2\pi}{a}\right)^3. \tag{2.13}$$

The primitive Wigner-Seitz cell of the reciprocal lattice is called ***first Brillouin zone*** [42]. It's borders cut the reciprocal lattice vectors at

$$\begin{aligned} \pm\frac{1}{2}\mathbf{b}_1 &= \pm\frac{\pi}{a}\hat{\mathbf{x}} \\ \pm\frac{1}{2}\mathbf{b}_2 &= \pm\frac{\pi}{a}\hat{\mathbf{y}} \\ \pm\frac{1}{2}\mathbf{b}_3 &= \pm\frac{\pi}{a}\hat{\mathbf{z}}. \end{aligned} \tag{2.14}$$

The Crystal Structure of GaN

Realistic semiconductor crystals do not crystallize in a simple cubic lattice. For example Silicon and Germanium crystallize in a diamond lattice, which can be seen as the interpenetration of two face-centered cubic lattices [30]. In contrast to those single-element semiconductors or most III-IV compounds which crystallize in the zincblende structure, GaN has got a wurtzite[1] crystal structure. In the wurtzite structure every atoms four nearest neighbors are equidistantly distributed forming tetrahedrons [30] as

[1] Zincblende structure is also possible

can be seen in 2.4. The corresponding Brillouin zone for a wurtzite crystal is also shown in this figure.

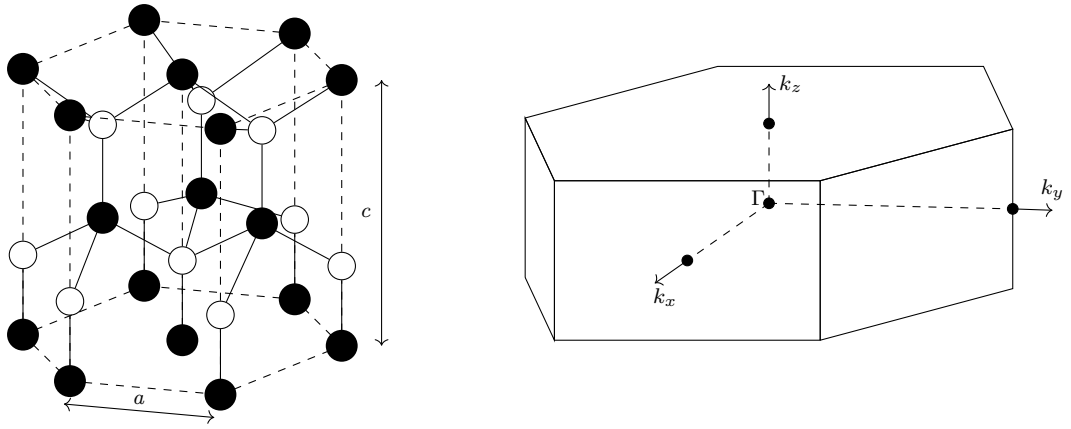

Figure 2.4: Left: GaN wurtzite crystal with lattice constants 'a' and 'c'. Right: Brillouin zone of the GaN wurtzite structure [30].

2.1.2 Band Structure / Density of States / Fermi Distribution / Carrier Densities

The energetic conditions of electrons bound to a nucleus by Coulomb force can be described by the Schrödinger equation, which has been postulated by Erwin Schrödinger in 1926 [43]. Here the state of the particles is not given in a deterministic manner but as wavefunctions holding the information of the probability for the occupation of a certain position[2] and momentum. In a crystal, the valence electrons are affected by a periodic potential caused by the Coulomb forces of a periodic arrangement of atoms. The proximity of the neighbouring atoms causes overlapping wavefunctions (or orbitals), so that the probability to find one particular electron is not limited to a certain atom but spans over the whole crystal. This enables the crystal to be conductive. In such an arrangement the energy states get separated into bands of allowed quasi-continuous energies interrupted by gaps of forbidden states. Approximate solutions for these band structures can be found by solving the Schrödinger equation for an one electron problem applying the later introduced Bloch theorem [30, 44, 45]. Since the Schrödinger Equation is a postulate, it cannot be derived from any other fundamental physical law. Nonetheless there are ways to get an intuitive access to it [44, 46]. Therefore the total energy of

[2]The square of the wave functions can be interpreted as the probability of presence. Those regions which have a none vanishing wavefunction, meaning that an electron could be found with a certain probability, are also called orbitals.

an electron is considered to be consisting of the kinetic energy of the particle and the potential energy arising from external sources.

$$E = E_{\text{kin}} + E_{\text{pot}} \tag{2.15}$$

For a non-relativistic particle the mass can assumed to be constant and the kinetic energy reads

$$E_{kin} = \frac{mv^2}{2} = \frac{p^2}{2m}. \tag{2.16}$$

The potential energy depends on the position in the (in general time dependent) force field and shall be denoted as $E_{pot} = V(\mathbf{R}, t)$. Using this notation and multiplying the total energy with the desired wave function yields

$$E\Psi = \frac{p^2}{2m}\Psi + V(\mathbf{R}, t)\Psi. \tag{2.17}$$

Assuming the solution of the wave function for a free electron which are in the form of plane waves

$$\Psi(\mathbf{R}, t) = \Psi_0 e^{-\jmath(\omega t - \mathbf{k}\cdot\mathbf{R})} = \Psi_0 e^{-\frac{\jmath}{\hbar}(Et - \mathbf{p}\cdot\mathbf{R})}, \tag{2.18}$$

using $E = \hbar\omega$ and the crystal momentum $\mathbf{p} = \hbar\mathbf{k}$, where $\hbar = \frac{h}{2\pi}$ is the reduced Planck constant, and analyzing the argument of the exponential function it gives rise that the energy and momentum terms in 2.17 can be substituted by time and spatial derivatives of the wave function. Evaluation of the derivative gets

$$\frac{\partial}{\partial t}\Psi(\mathbf{R}, t) = -\jmath\frac{E}{\hbar}\Psi(\mathbf{R}, t) \qquad => \qquad E\Psi = \jmath\hbar\frac{\partial}{\partial t}\Psi \tag{2.19}$$

and

$$\nabla \cdot \nabla\Psi(\mathbf{R}, t) = \Delta\Psi(\mathbf{R}, t) = -\frac{p^2}{\hbar^2}\Psi(\mathbf{R}, t) \qquad => \qquad p^2\Psi = -\hbar^2\Delta\Psi. \tag{2.20}$$

Substituting 2.19 and 2.20 into 2.17 one yields the space and time dependent Schrödinger Equation

$$-\frac{\hbar^2}{2m}\Delta\Psi(\mathbf{R}, t) + V(\mathbf{R}, t)\Psi(\mathbf{R}, t) = \jmath\hbar\frac{\partial}{\partial t}\Psi(\mathbf{R}, t). \tag{2.21}$$

Since the potential in semiconductor crystals is caused by the atomic cores and hence it can be assumed to be time invariant, here the potential shall be considered only having a spatial dependence. Assuming that the frequencies which with the excited particles decay from one quantum state to another are slow compared to the eigenfrequencies of the quantum states, which is generally the case [44], the Schrödinger equation can be solved stationary to give insight to the energetic distribution of carriers. In this case the wavefunctions are having sharp energies and hence are given by harmonic oscillators with

frequencies of $w_n = \frac{E_n}{\hbar}$ and amplitudes of only spatial dependence. Here the Schrödinger equation can be simplified to[3]

$$\begin{aligned} -\frac{\hbar^2}{2m}\Delta\psi(\mathbf{R},t) + V(\mathbf{R})\psi(\mathbf{R},t) &= \jmath\hbar\frac{\partial}{\partial t}\psi(\mathbf{R},t) \\ \left(-\frac{\hbar^2}{2m}\Delta\Psi(\mathbf{R}) + V(\mathbf{R})\Psi(\mathbf{R})\right)e^{-\frac{\jmath}{\hbar}Et} &= E\underbrace{\Psi_0 e^{-\jmath\mathbf{k}\cdot\mathbf{R}}}_{\Psi(\mathbf{R})}e^{-\frac{\jmath}{\hbar}Et} \\ \underbrace{\left(-\frac{\hbar^2}{2m}\Delta + V(\mathbf{R})\right)}_{\widehat{H}(\mathbf{R})}\Psi(\mathbf{R}) &= E\Psi(\mathbf{R}). \end{aligned} \tag{2.22}$$

So, in the stationary case, the wave functions Ψ are described as eigenfunctions of the so-called Hamiltonian operator $\widehat{H}$ with corresponding eigenvalues E.

$$\widehat{H}\Psi(\mathbf{R},\mathbf{k}) = E\Psi(\mathbf{R},\mathbf{k}) \tag{2.23}$$

As can be seen from eq. (2.22) the Hamiltonian operator is the sum of the operators for kinetic and potential energy and hence it serves as the operator of total energy.

The solution of this eigenvalue equation for a free particle (constant potential energy)

$$\Delta\Psi(\mathbf{R}) = -\frac{2m}{\hbar^2}(E - V_0)\Psi(\mathbf{R}) \tag{2.24}$$

is obviously described by a homogeneous Helmholtz equation

$$\left(\Delta + k^2\right)\Psi(\mathbf{R}) = \mathbf{0} \tag{2.25}$$

with the crystal momentum[4]

$$k = \frac{\sqrt{2m(E - V_0)}}{\hbar}. \tag{2.26}$$

Solutions of this equation can be found as plane waves of the form eq. (2.18), having a total energy of

$$E = \frac{\hbar^2 k^2}{2m} + V_0. \tag{2.27}$$

For the 1D issue of a single particle being confined in-between infinitely high potential walls, a so-called potential well, the wave functions must vanish outside of the well.

[3] This result can also be found by a separation of variables, so that a general solution of the time dependent Schrödinger equation could be generated as a superposition of weighted stationary solutions multiplied with their corresponding oscillators.

[4] Since the total energy must be greater than the potential energy $E - V_0 > 0$, the solutions must be non-attenuated, meaning, infinitely spatial distributed plane waves, having a sharply defined momentum or energy according to the Heisenberg uncertainty principle.

Assuming again a constant potential within the well, it becomes obvious that the solutions of the wave functions here are also given in the shape of plane waves. In order to be continuous those waves must also be zero at the well's borders, therefore solutions are only possible as standing waves with half wave lengths being multiple integers of the well thickness $l = n\frac{\lambda}{2}$. Using $\lambda = \frac{2\pi}{k}$ and eq. (2.26) to express this finding in terms of the energy

$$l = n\frac{\lambda}{2} = n\frac{\pi}{k} = n\frac{\hbar\pi}{\sqrt{2m(E - V_0)}}$$

we find the allowed energies to be quantized with energy eigenvalues of

$$E_n = \frac{1}{2m}\left(\underbrace{\hbar\frac{n\pi}{l}}_{p=\hbar k}\right)^2 + V_0. \tag{2.28}$$

This equation gives some interesting insights into the conditions of the energy states in a quantum well. First of all, the energy level of the ground state will never fall together with the bottom of the well, but will always be shifted up by an energy of $\frac{1}{2m}\left(\frac{\hbar\pi}{l}\right)^2$. The kinetic energy (or the crystal momentum) of subsequent eigenstates leap by energy quantums proportional to the square of their quantum numbers. Furthermore the crystal momentum is reciprocally dependent of the square of the quantum well's width $k_n \propto \frac{1}{l^2}$, which means that narrow quantum wells will lead to eigenstates of higher kinetic energy. The broader the well gets the more wavelengths fit into it and hence the eigenstates move closer together, moreover the ground states moves closer to the well's bottom. Even though this gives rise that for an infinitely large well the quantization of the energy states vanishes, the solution does not transit to the solution of a free particle, since still the boundary condition of a vanishing field would have to be fulfilled in infinity. Figure 2.5 illustrates these solutions of the stationary Schrödinger equation.

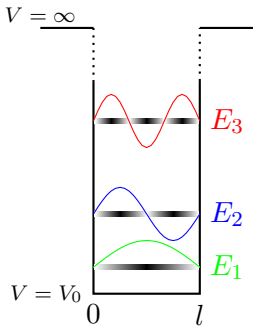

Figure 2.5: First three allowed energy states within a infinitely high quantum well with corresponding wave functions and the probability of presence depicted as shaded areas. Here dark areas represent spots of high probability, completely white areas stand for a probability of zero.

Bloch Theorem

The calculation of energy states in a semiconductor crystal using realistic potentials is difficult and requires the use of expensive computational methods. Nonetheless several realistic physical properties of particles in a semiconductor potential can be found by solving the stationary Schrödinger equation for a one electron problem. Here the Coulomb forces among the valence electrons are neglected, hence the potential energy of the electron solely arises from the periodically arranged atomic cores. In this case the potential energy shows to have the periodicity of the unit cell [45]

$$V(\mathbf{R} + \mathbf{R}_l) = V(\mathbf{R}), \tag{2.29}$$

where $\mathbf{R}_l$ represents a lattice vector. The Bloch theorem states that the wave functions for particles in such a periodic potential are given by the so-called Bloch functions [30]

$$\Psi(\mathbf{R}, \mathbf{k}) = u_n(\mathbf{R}, \mathbf{k}) e^{(\jmath \mathbf{k} \cdot \mathbf{R})} \tag{2.30}$$

where u_n and Ψ have the periodicity of the direct lattice

$$u_n(\mathbf{R} + \mathbf{R}_l) = u_n(\mathbf{R}). \tag{2.31}$$

The index n denotes the different eigenfunctions belonging to a fixed wave vector $\mathbf{k}$ [45].

Putting the Bloch functions into the stationary Schrödinger equation (2.22) and using the vector theorem [47]

$$\Delta(\Phi \mathbf{A}) = \Phi \Delta \mathbf{A} + 2 \nabla \Phi \cdot \nabla \mathbf{A} + \mathbf{A} \Delta \Phi$$

one gets

$$\left(-\frac{h^2}{2m}\Delta - \frac{\jmath h^2}{m}\mathbf{k}\cdot\nabla + \frac{h^2k^2}{2m} + V(\mathbf{R})\right) u_n = E_n(\mathbf{k})u_n(\mathbf{R},k). \tag{2.32}$$

Here the **k** dependence of the kinetic energy in the Hamiltonian identifies the possible energy states as energy bands, as they are functions of the wave vector. In general an eigenvalue equation like this allows multiple solutions for a fixed wave vector, resulting in several bands of allowed energy states. This degeneracy is counted by the index n. For energy states within the band gap eq. (2.32) demands complex wave vectors for which the Bloch theorem is not satisfied. Those states are forbidden [45]. The bandstructure of GaN can be seen in figure 2.6 .

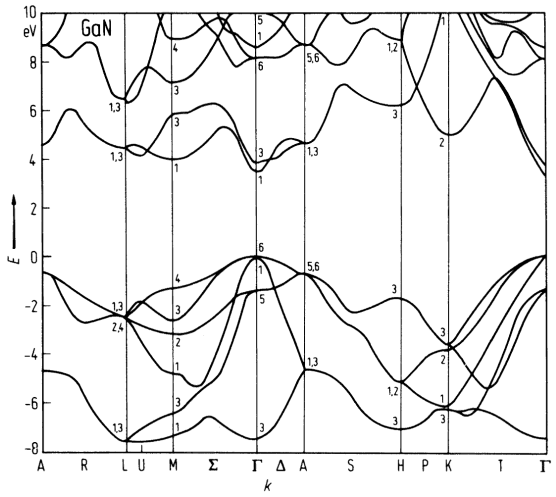

Figure 2.6: The bandstructure of wurtzite GaN. [Source Ioffe Institute [48].]

2.1.3 Density of states and Fermi distribution in thermal equilibrium

The density of states $D(E)$ represents the amount of energy states within the energy interval $E + dE$ allowed to be occupied by electrons. In 2.6 it can be seen that for a sufficiently small momentum k the energy relation $E(k)$ (also called *dispersion relation*) at the Γ-Point still have the parabolic shape of the free electron solution[5] eq. (2.27). In this regime the particles in a semiconductor crystal still behave like free particles but the influence of the periodic crystal potential changes the dispersion relation in some ways. First by defining the energy to be zero at the maximum of the valence band, the potential energy V_0 becomes the energy of the conduction band minimum (representing the band gap). Furthermore the curvature will be different compared to that of the free electron solution. As can be seen by eq. (2.27) this curvature is determined by the

[5]For the highest valence bands (light holes and heavy holes) and the lowest conduction band.

particle mass which can given by deriving E two times regarding k. Hence, with this quadratic approximation the influence of the potential is already considered by a so-called effective mass

$$m^* = \hbar^2 \left(\frac{d^2 E}{dk^2} \right). \tag{2.33}$$

So the $E(k)$-relation for an electron close to conduction band minimum can be approximated by the parabolic equation

$$E = \frac{\hbar^2 k^2}{2m^*} + E_\mathrm{c}. \tag{2.34}$$

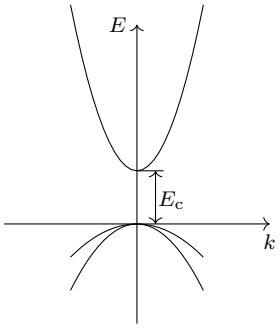

Figure 2.7: Illustration of the dispersion relation for valence and conduction band.

Because of the parabolic shape of the dispersion relation, valid **k**-vectors have to fullfill

$$k^2 - \frac{2m^*}{\hbar^2}(E - E_\mathrm{c}) = 0 \tag{2.35}$$

$$\Rightarrow k_x^2 + k_y^2 + k_z^2 - \rho^2 = 0 \tag{2.36}$$

hence for any energy level E they describe a spherical surface[6] with radius $\rho = k = \frac{\sqrt{2m^*(E-E_\mathrm{c})}}{\hbar}$.

[6] On the so-called Ewald sphere

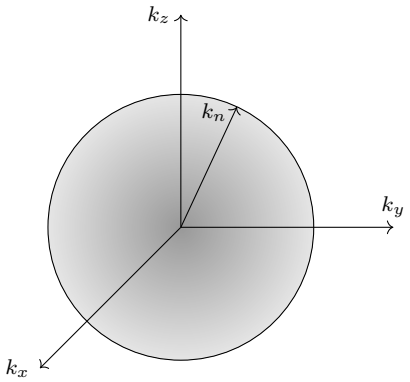

Figure 2.8: For energies lying in the conduction band ($E > E_\mathrm{c}$), the states occupy a spherical volume in **k**-space with radius $k_n = \frac{\sqrt{2m^*(E-E_\mathrm{c})}}{\hbar}$ [42].

The differential volume to accommodate additional states for an infinitesimal deviation of the wave vector thus is given by spherical shells

$$dV = 4\pi k^2 dk. \tag{2.37}$$

The differential increment of the number of states in such a shell is given by the shell volume divided by the volume of a singular state

$$dN = \frac{dV}{V_s}. \tag{2.38}$$

Since due to the crystal periodicity the components of the wave vector have to fulfill the condition $k_i = n_\mathrm{i}\frac{2\pi}{L_i}$ each state occupies a volume of [42]

$$V_s = \frac{(2\pi)^3}{L^3} = \frac{(2\pi)^3}{V_\mathrm{c}}. \tag{2.39}$$

Integration over the whole spherical volume delivers the total number of states in k-space

$$\begin{aligned} N(k) &= \frac{V_\mathrm{c}}{(2\pi)^3}\int_V dV = \frac{V_\mathrm{c}}{(2\pi)^3}\int_0^k 4\pi k'^2 dk' \\ &= \frac{V_\mathrm{c}}{3(2\pi^2)}k^3. \end{aligned} \tag{2.40}$$

In order to get the number of states in energy space eq. (2.34) can be used to express the crystal momentum in terms of the energy

$$N(E) = \frac{V_\mathrm{c}}{3(2\pi^2)}\left(\frac{m^*}{\hbar^2}\right)^{\frac{3}{2}}(E-E_\mathrm{c})^{\frac{3}{2}}.$$

Deriving the number of states regarding the energy gives the rate

$$\frac{dN}{dE} = \frac{V_\mathrm{c}}{4\pi^2}\left(\frac{m^*}{\hbar^2}\right)^{\frac{3}{2}}\sqrt{E - E_\mathrm{c}}. \tag{2.41}$$

As the density of states $D(E)$ is defined as the volume related differential rate of the number of states and the energy $D(E) = \frac{1}{V_\mathrm{c}}\frac{dN}{dE}$, we get

$$D(E) = \frac{1}{4\pi^2}\left(\frac{m^*}{\hbar^2}\right)^{\frac{3}{2}}\sqrt{E - E_\mathrm{c}}. \tag{2.42}$$

Carrier densities

The density of states can be used to determine the carrier concentration within a semiconductor. Therefore each state needs to multiplied with it's occupation probability and summed up over the corresponding energy band[7]. The electron density is then given by

$$n = \int_{E_\mathrm{c}}^{\infty} 2D(E)f(E)dE \tag{2.43}$$

where f(E) is the occupation probability for a state in the conduction band. Since holes in the valence band occur if a state is not occupied, here the complementary function $1 - f(E)$ can be used, resulting in

$$p = \int_{-\infty}^{E_v} 2D(E)[1 - f(E)]dE. \tag{2.44}$$

Fermi Distribution

As fermions, electrons are liable to the Pauli exclusion principle. The energetic distribution of those kind of particles in an electron gas is governed by *Fermi-Dirac-Statistics*

$$f(E) = \frac{1}{1 + \exp\left(\frac{E - E_\mathrm{F}}{k_\mathrm{B}T}\right)}. \tag{2.45}$$

Here, E_F is the so-called *Fermi energy*, which is defined as the energy having an occupation probability of $\frac{1}{2}$.

[7] An additional factor of 2 has to be added, in order to consider the two different electron spins allowed per each state.

Using Fermi-Dirac statistics, the electron concentration in the conduction band becomes

$$n = \int_{E_c}^{\infty} \frac{1}{2\pi^2} \left(\frac{m^*}{\hbar^2}\right)^{\frac{3}{2}} \frac{\sqrt{E - E_c}}{1 + \exp\left(\frac{E - E_F}{k_B T}\right)} dE \tag{2.46}$$

Unfortunately there is no closed analytic expression for this integral, hence n has to be evaluated numerical considering Fermi-Dirac statistics. For non-degenerated n-doped (or intrinsic) semiconductors the Fermi energy lies below the conduction band energy. If the distance from the conduction band is several $k_B T$ the exponential function becomes large and approximately determines the denominator. In this case $f(E)$ can be approximated by *Boltzmann statistics*

$$f(E) = \exp\left(-\frac{E - E_F}{k_B T}\right). \tag{2.47}$$

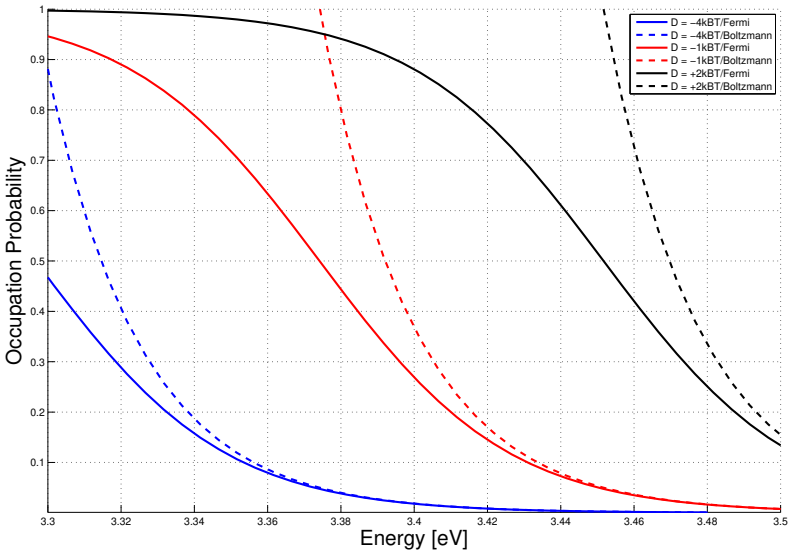

Figure 2.9: Comparison of Fermi- and Boltzmann statistics for given distances 'D' of the Fermi-level from the conduction band edge (arbitrarily chosen at 3.4 eV). For a Fermi-level lying at distance of 4 $k_B T$ below the band edge, Boltzmann statics approximates the Fermi probability function at the band edge and above to a very good level. While at a distance of -1 $k_B T$ already a distinct deviation can be seen, in the degenerated case Boltzmann statistic fails completely, having a probability much greater than 1 at the band edge

With Boltzmann statistics both carrier concentrations can be evaluated analytically, as the solutions of eqs. (2.43) and (2.44)

$$n = N_\mathrm{c} \exp\left(-\frac{E_\mathrm{c} - E_\mathrm{F}}{k_\mathrm{B}T}\right) \tag{2.48}$$

$$p = N_\mathrm{v} \exp\left(-\frac{E_\mathrm{F} - E_\mathrm{v}}{k_\mathrm{B}T}\right). \tag{2.49}$$

Intrinsic Carrier Concentration

In intrinsic semiconductors every electron in the conduction band originates from a hole in the valence band. So it's obvious that

$$n = p = n_\mathrm{i} \tag{2.50}$$

and

$$np = n_\mathrm{i}^2 = N_\mathrm{c} N_\mathrm{v} \exp\left(-\frac{E_\mathrm{g}}{k_\mathrm{B}T}\right). \tag{2.51}$$

Equation (2.51) is the *mass action law* [30] with the intrinsic carrier concentration n_i and the band gap energy $E_\mathrm{g} = E_\mathrm{c} - Ev$.

The equality of carrier concentrations enables the calculation of the Fermi energy for an intrinsic semiconductor E_i using eqs. (2.48) and (2.49)

$$N_\mathrm{c} \exp\left(-\frac{E_\mathrm{c} - E_\mathrm{F}}{k_\mathrm{B}T}\right) = N_\mathrm{v} \exp\left(-\frac{E_\mathrm{F} - E_\mathrm{v}}{k_\mathrm{B}T}\right) \tag{2.52}$$

$$\Rightarrow E_\mathrm{i} = E_\mathrm{F} = \frac{E_\mathrm{c} + E_\mathrm{v}}{2} + \frac{k_\mathrm{B}T}{2} \ln \frac{N_\mathrm{v}}{N_\mathrm{c}}. \tag{2.53}$$

The intrinsic Fermi level lies roughly in the middle of the band gap [30].

Non-Degenerated Semiconductor

For moderately n-doped semiconductors the Fermi energy lies several $k_\mathrm{B}T$ below the conduction band edge. So equations eqs. (2.48) and (2.49) can still be used to determine the carrier concentrations and hence also the mass action law(2.51) stays valid. But now the majority of electrons does not originate from the valence band anymore but from ionized impurities, so that

$$n \neq p. \tag{2.54}$$

Unlike the carrier concentrations the total amount of negative charges have to be compensated by positive charges. This leads to the charge neutrality condition

$$n + N_{\mathrm{A}}^{-} = p + N_{\mathrm{D}}^{+}. \tag{2.55}$$

With N_{A}^{-} and N_{D}^{+} being the ionized acceptors and donors respectively. So considering a purely n-doped crystal the charge neutrality demands

$$n = p + N_{\mathrm{D}}^{+}. \tag{2.56}$$

Substituting p via eq. (2.56) in the mass action law n can be determined by a quadratic equation

$$n^2 - N_{\mathrm{D}}^{+} n - n_{\mathrm{i}}^2 = 0$$

$$n = \frac{N_{\mathrm{D}}^{+}}{2} + \frac{1}{2}\sqrt{(N_{\mathrm{D}}^{+})^2 + (2n_{\mathrm{i}})^2} \tag{2.57}$$

showing that $n \approx N_{\mathrm{D}}^{+}$ is a good approximation of the electron density if $N_{\mathrm{D}}^{+} >> 2n_{\mathrm{i}}$. Assuming furthermore a complete ionization of the impurities $N_{\mathrm{D}}^{+} = N_{\mathrm{D}}$, the carrier densities can be approximated by

$$n \approx N_{\mathrm{D}} \tag{2.58}$$

$$p \approx \frac{n_{\mathrm{i}}^2}{N_{\mathrm{D}}} \tag{2.59}$$

With this approximation the Fermi energy would then result from Equation (2.48)

$$N_{\mathrm{D}} \approx N_{\mathrm{c}} \exp\left(-\frac{E_{\mathrm{c}} - E_{\mathrm{F}}}{k_{\mathrm{B}}T}\right) \tag{2.60}$$

$$\Rightarrow E_{\mathrm{F}} \approx E_{\mathrm{c}} + k_{\mathrm{B}}T \ln\frac{N_{\mathrm{D}}}{N_{\mathrm{c}}} \tag{2.61}$$

In the case that incomplete ionization has to be considered, the amount of ionized impurities can be determined by [30]

$$N_{\mathrm{D}}^{+} = \frac{1}{1 + g_{\mathrm{D}} \exp(\frac{E_{\mathrm{F}} - E_{\mathrm{D}}}{k_{\mathrm{B}}T})} \tag{2.62}$$

and

$$N_{\mathrm{A}}^{-} = \frac{1}{1 + g_{\mathrm{A}} \exp(\frac{E_{\mathrm{A}} - E_{\mathrm{F}}}{k_{\mathrm{B}}T})} \tag{2.63}$$

respectively. Here E_{D} and E_{A} are the energy levels of the impurities, g_{D} and g_{A} consider the band degeneracy[8] and the fact that each state can be occupied by two electrons of opposite spin. Assuming again that $n \approx N_{\mathrm{D}}^{+}$ the Fermi level can be calculated from

$$N_{\mathrm{c}} \exp\left(-\frac{E_{\mathrm{c}} - E_{\mathrm{F}}}{k_{\mathrm{B}}T}\right) \approx \frac{1}{1 + 2\exp(\frac{E_{\mathrm{F}} - E_{\mathrm{D}}}{k_{\mathrm{B}}T})}. \tag{2.64}$$

[8] In general $g_{\mathrm{D}} = 2$ and $g_{\mathrm{A}} = 4$ covering for a non-degenerated conduction band and two degenerated valence bands (light holes and heavy holes) at $\mathbf{k} = \mathbf{0}$, like it is true for most semiconductors [30]

Non-Equilibrium Fermi Distribution

If the semiconductor is exhibited to an outer stimulus[9], carrier concentrations leave the condition of thermal equilibrium. Here, the probabilities of carrier generation and recombination differ and the mass action law is not valid anymore. Since relaxation processes are much faster than recombination processes each band can be seen as being in equilibrium even though the two carrier species among each other are not. Therefore two separate quasi Fermi levels[10] can be defined. Using these auxiliary functions the carrier concentrations can be determined in the previous manner, just substituting the Fermi level by the corresponding quasi Fermi level.

$$n = N_\mathrm{c} \exp\left(-\frac{E_\mathrm{c} - E_\mathrm{Fn}}{k_\mathrm{B}T}\right) = n_\mathrm{i} \exp\left(-\frac{E_\mathrm{i} - E_\mathrm{Fn}}{k_\mathrm{B}T}\right) \tag{2.65}$$

$$p = N_\mathrm{v} \exp\left(-\frac{E_\mathrm{Fp} - E_\mathrm{v}}{k_\mathrm{B}T}\right) = n_\mathrm{i} \exp\left(-\frac{E_\mathrm{Fp} - E_\mathrm{i}}{k_\mathrm{B}T}\right). \tag{2.66}$$

This turns the mass action law changes for carrier concentrations beyond thermal equilibrium to

$$np = n_\mathrm{i}^2 \exp\left(\frac{E_\mathrm{Fn} - E_\mathrm{Fp}}{k_\mathrm{B}T}\right). \tag{2.67}$$

From eqs. (2.65) and (2.66) the quasi Fermi energies can be calculated for a known carrier concentration [30]

$$E_\mathrm{Fn} = E_\mathrm{c} + k_\mathrm{B}T \ln\frac{n}{N_\mathrm{c}} = E_\mathrm{i} + k_\mathrm{B}T \ln\frac{n}{n_\mathrm{i}} \tag{2.68}$$

$$E_\mathrm{Fp} = E_\mathrm{v} - k_\mathrm{B}T \ln\frac{p}{N_\mathrm{v}} = E_\mathrm{i} - k_\mathrm{B}T \ln\frac{p}{n_\mathrm{i}}. \tag{2.69}$$

2.2 Carrier Transport

In order to model carrier transport within semiconductors, here the *Drift-Diffusion-Model* will be used, where two equations for the temporal derivatives of the respective carriers and a third equation for the spatial derivative of the potential form a coupled non-linear system of equations. With this approach the macroscopic behavior of bulk semiconductor devices can be described. Before deriving the Drift-Diffusion Model from *Maxwell's Equations*, the singular parts the model is build on shall be closer introduced.

[9] Like a light illumination or a bias voltage

[10] Also called imref (Fermi spelled backwards) functions.

2.2.1 Drift Current

In the presence of an electric field, carriers in partially filled bands encounter an acceleration, where for moderate field strengths the resulting mean carrier drift velocity is proportional to the acting field

$$\mathbf{v}_n = -\mu_n \mathbf{E} \tag{2.70}$$

$$\mathbf{v}_p = \mu_p \mathbf{E}. \tag{2.71}$$

Here, the constants of proportionality μ_n and μ_p are the so-called carrier mobilities for electrons and holes. The reason for the negative sign in the electron drift velocity is that physically electrons move in the direction opposite of the acting direction of the electric field. Different scattering processes within the material, like phonon scattering, scattering at impurities or carrier-carrier scattering impede the acceleration of electrons and holes. This indicates that the mobilities show a distinct temperature and doping dependency. Both mobilities together completely represent the material influence on the conduction current. Considering eq. (2.70) in the drift current density, which is the sum of the products of the electron and hole space charge densities and their corresponding drift velocities

$$\mathbf{J} = nq\mathbf{v}_n + pq\mathbf{v}_p = q(\mu_n n + \mu_p p)\mathbf{E} \tag{2.72}$$

and comparing it to the *Ohmic law*

$$\mathbf{J} = \sigma \mathbf{E}, \tag{2.73}$$

one can see that the conductivity σ depends on the carrier densities and the corresponding mobilities, holding the material properties.

$$\sigma = q(\mu_n n + \mu_p p) \tag{2.74}$$

In the case of sufficiently high electric fields, where the influence of the field on scattering processes cannot be neglected, the linear relation for the mean drift velocity does not hold any more and an additional field dependence of the mobilities must be considered. Here, the *Ohmic law* is not valid. A detailed insight in the mobility field dependence and different modeling strategies are given in section 1.2.2.

2.2.2 Diffusion Current

Unlike in metals, carrier densities in semiconductors devices can vary significantly in space. For example in the case that carriers are locally excited due to an outer stimulus, like a local light illumination, or if the semiconductor simply consists of differently doped regions. A non-uniform carrier concentration causes equalization processes, where carriers

diffuse from regions of higher concentration to those of lower concentration. Diffusion processes introduce a current flow apart from the aforementioned drift current, not driven by an electric field but purely by a concentration gradient

$$\mathbf{J}_{n(\text{diff})} = qD_n \nabla n \tag{2.75}$$
$$\mathbf{J}_{p(\text{diff})} = -qD_p \nabla p. \tag{2.76}$$

For the total current densities those diffusion currents superimpose with the drift currents. The constants of proportionality D_n and D_p are the so-called diffusion coefficients. Since they are describing a capability of the carriers to move, one can already assume that they are depending on the carrier mobilities. In order to derive their mobility dependence[11], thermal equilibrium shall be considered. As it is shown in the next section, in thermal equilibrium there is no net current flow, hence the drift and diffusion currents cancel each other out

$$q\mu_n n\mathbf{E} = -qD_n \nabla n. \tag{2.77}$$

Where for the spatial derivative the integral from eq 2.43

$$\nabla n(\mathbf{R}) = \nabla \int_{E_c(\mathbf{R})}^{\infty} 2D(E) f_n(E) dE \tag{2.78}$$

has to be considered again. Equation (2.78) identifies the spatial dependence of the carrier density as a function of the energetic position of the conduction band edge. This is because here the energetic distance from Fermi level is inherently covered. By this energetic difference the occupation probability of the states in the conduction band is designated, as shall be illustrated by the example of a simple diode formed by an intrinsic and a N^+-semiconductor region shown in fig.(2.10).

[11] For non-degenerated semiconductors given by the so-called *Einstein relation* [49]

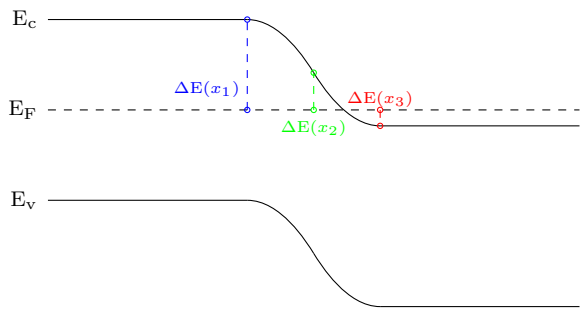

Figure 2.10: In this example three different regions for the carrier density can be identified - the intrinsic ($x < x_1$) and the $N^+ - region$ region ($x > x_3$), where the carrier density will be constant, though with different values each, and a depletion region ($x_1 <= x <= x_3$) in which the energetic position of the conduction band[12] varies regarding the Fermi level.

Using the chain rule, equation 2.78 can be transformed to [13]

$$\nabla n(\mathbf{R}) = \nabla n(E_\mathrm{c}(\mathbf{R})) = \frac{\partial n}{\partial E_\mathrm{c}} \nabla E_\mathrm{c}. \tag{2.79}$$

Now, the latter derivative in eq. (2.79) can be rewritten in terms of the electric field [30]

$$\nabla E_\mathrm{c} = q\mathbf{E}. \tag{2.80}$$

Substituting these results into eq. (2.77), the equation of the equilibrium current reads

$$q\mu_n n\mathbf{E} = -q^2 D_n \frac{\partial n}{\partial E_\mathrm{c}} \mathbf{E}, \tag{2.81}$$

where by comparing the coefficients of $\mathbf{E}$ and rearranging the terms, the diffusion coefficient finally results in a general expression giving the relation between diffusion coefficient and carrier mobility.

$$D_n = -\frac{\mu_n n}{q} \left(\frac{\partial n}{\partial E_\mathrm{c}} \right)^{-1} = -\frac{\mu_n}{q} \frac{n}{n'(E_\mathrm{c})}, \tag{2.82}$$

As mentioned in section 2.1.3, assuming Fermi-Dirac statistics there is no closed analytic expression for the carrier densities represented by $\mathrm{F}(E_\mathrm{c}(\mathrm{x}))$. But again for carrier

[12] And hence the carrier density

[13] The continuous nature of $D(E)f_n(E)$ (for a 3d density of states) is sufficient to ensure a continuous differentiable primitive integral in 2.78 according to the fundamental theorem of calculus. On the other hand, this shows that the validity of the diffusion coefficients is limited to the continuum case (a volume of (quasi-)continuous states), in comparison to quantum structures like wells or dots, where the population is determined by capture rates.

concentrations where the Fermi-level is substantially below the conduction band edge, substituting Fermi-Dirac statistics with Maxwell-Boltzmann statistics delivers an appropriate approximation for the carrier densities. In this case a closed expression for diffusion coefficient can be found by evaluating

$$\begin{aligned}\left(\frac{\partial n}{\partial E_\mathrm{c}}\right)^{-1} &= \left(\frac{\partial}{\partial E_\mathrm{c}} N_\mathrm{c} \exp\left(-\frac{E_\mathrm{c}-E_\mathrm{F}}{k_\mathrm{B}T}\right)\right)^{-1} \\ &= \left(-\frac{1}{k_\mathrm{B}T} N_\mathrm{c} \exp\left(-\frac{E_\mathrm{c}-E_\mathrm{F}}{k_\mathrm{B}T}\right)\right)^{-1} \\ &= \left(-\frac{n}{k_\mathrm{B}T}\right)^{-1} \\ &= -\frac{k_\mathrm{B}T}{n}.\end{aligned} \tag{2.83}$$

With this result eq. (2.82) simplifies and the electron diffusion coefficient can be calculated directly from the electron mobility. Analogous treatment of the hole current yields a similar result for the hole diffusion current, leading to the *Einstein relation*

$$D_n = \frac{k_\mathrm{B}T}{q}\mu_n \tag{2.84}$$

$$D_p = \frac{k_\mathrm{B}T}{q}\mu_p. \tag{2.85}$$

It has to be kept in mind that this simple relation from eqs. (2.84) and (2.85) is only valid for non-degenerated semiconductors. If the *Fermi-level* is close to the conduction band edge, or above, the diffusion coefficients must be evaluated numerically.

It might be worthwhile to notice the $\frac{n}{n(E_\mathrm{c})'}$ structure in eq. (2.82), meaning that by taking the inverse and integration regarding E_c, a relation between conduction band bending and carrier density can be found

$$\begin{aligned}n(E_c(\mathbf{R})) &= n(E_\mathrm{c}(\mathbf{R_0}))\ \exp(-\frac{\mu_n}{qD_n}[E_\mathrm{c}(\mathbf{R}) - E_\mathrm{c}(\mathbf{R_0})]) \\ &= n(E_\mathrm{c}(\mathbf{R_0}))\ \exp(\frac{1}{k_\mathrm{B}T}[E_\mathrm{c}(\mathbf{R_0}) - E_\mathrm{c}(\mathbf{R})]).\end{aligned} \tag{2.86}$$

This equation, stating an exponential decay of the electron density for an upward bending conduction band, can be expressed by means of the electrostatic potential using $E_\mathrm{c} = -q\Phi + const$ [50]

$$n(\Phi(\mathbf{R})) = n(\Phi(\mathbf{R_0}))\ \exp(\frac{q}{k_\mathrm{B}T}[\Phi(\mathbf{R}) - \Phi(\mathbf{R_0})]). \tag{2.87}$$

The hole density can be derived in an analogous manner. From simply taking the logarithm of (2.87) an expression for the potential can be found as

$$\begin{aligned}\Phi(\mathbf{R}) &= \Phi(\mathbf{R_0}) + \frac{k_\mathrm{B}T}{q} \ln\left(\frac{n(\mathbf{R})}{n(\mathbf{R_0})}\right) \\ &= \Phi(\mathbf{R_0}) + \frac{k_\mathrm{B}T}{q} \left[\ln n(\mathbf{R}) - \ln n(\mathbf{R_0})\right]. \end{aligned} \tag{2.88}$$

According to eq. (2.88) the diffusion of carriers shapes a non-constant potential profile throughout the semiconductor. With this varying potential an electric field goes along which causes a drift current counteracting the diffusion current. The field strength that is build up when thermal equilibrium is reached[14], can be found by using $\mathbf{E}(\mathbf{R}, t) = -\nabla\Phi$[15] with eq. (2.88).

$$\begin{aligned}\mathbf{E}(\mathbf{R}) &= \underbrace{-\nabla\Phi(\mathbf{R_0})}_{\mathbf{0}} - \frac{k_\mathrm{B}T}{q} \left[\nabla \ln n(\mathbf{R}) - \underbrace{\nabla \ln n(\mathbf{R_0})}_{\mathbf{0}}\right] \\ &= -\frac{k_\mathrm{B}T}{q} \nabla \ln n(\mathbf{R}). \end{aligned} \tag{2.89}$$

Applying the insights from above to the example of a pn-junction in thermal equilibrium (see fig. 2.11), one can draw the following conclusions:

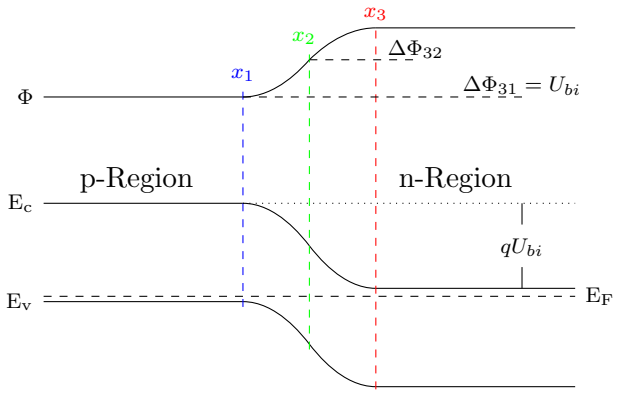

Figure 2.11: Example built-in voltage

When the pn-diode is in thermal equilibrium, for electrons in the n-Region a potential barrier $\Delta\Phi_{31} = \Phi(x_3) - \Phi(x_1) = \frac{k_\mathrm{B}T}{q} \ln\frac{n(x_3)}{n(x_1)} = U_{bi}$ has to be overcome in order to move

[14] Since thermal equilibrium was the key assumption for the derivation of the equations above
[15] Which is derived in eq. (2.129) in the next chapter.

into the p-Region. This potential difference which is commonly called built-in voltage[16], is determined by the ratio of the carrier densities in the n- and p-Region. On the other hand, in terms of the electron density eq. (2.87) states a exponential decay regarding the potential difference[17]. For example comparing the electron densities of n- and p-Region $n(x_1) = n(x_3)\exp(\frac{-q}{k_\mathrm{B}T}[\Phi(x_3) - \Phi(x1)]) = n(x3)\exp(\frac{-qU_{bi}}{k_\mathrm{B}T})$.

Interestingly, the fact that we derived the above conclusions only from eq. (2.82) shows that in a semiconductor consistent of only one material[18], the condition of carrier densities and potential in thermal equilibrium is purely governed by the ratio of carrier mobilities and diffusion coefficients.

2.2.3 Total Current Density

The superposition of the drift and diffusion parts yield the total current density

$$\mathbf{J} = \mathbf{J}_\mathrm{drift} + \mathbf{J}_\mathrm{diffusion} = q(\mu_n n + \mu_p p)\mathbf{E} + q(D_n \nabla n - D_p \nabla p). \tag{2.90}$$

For spatially constant diffusion coefficients this formulation becomes especially handy, since the current density now consists of the superposition of one drift vector and one diffusion vector

$$\mathbf{J} = q(\mu_n n + \mu_p p)\mathbf{E} + q\nabla(D_n n - D_p p). \tag{2.91}$$

Current dependency on quasi Fermi level

Another formulation for electron and hole currents, which is especially useful for the interpretation of the results in chapter 3, connects the current densities with the gradient of their corresponding quasi Fermi level. As a restriction for the derivation Boltzmann approximation shall be assumed. With this assumption the electron current density can be determined from eq. (2.90) to be

$$\mathbf{J}_n = q\mu_n n\mathbf{E} + \mu_n k_\mathrm{B} T \nabla n. \tag{2.92}$$

Using eq. (2.80) and Boltzmann approximation for the carrier densities outside of thermal equilibrium eq. (2.65) this can be expressed as

$$\mathbf{J}_n = \mu_n n \nabla E_\mathrm{c} + \mu_n k_\mathrm{B} T \, \nabla \left(N_\mathrm{c} \exp\left(\frac{E_\mathrm{Fn} - E_\mathrm{c}}{k_\mathrm{B} T} \right) \right). \tag{2.93}$$

[16] or diffusion voltage
[17] when moving into a region of lower potential
[18] homojunction

Equating the second gradient to

$$\begin{aligned}\nabla\left(N_{\mathrm{c}}\exp\left(\frac{E_{\mathrm{Fn}}-E_{\mathrm{c}}}{k_{\mathrm{B}}T}\right)\right) &= N_{\mathrm{c}}\exp\left(\frac{E_{\mathrm{Fn}}-E_{\mathrm{c}}}{k_{\mathrm{B}}T}\right)\frac{1}{k_{\mathrm{B}}T}\nabla\left(E_{\mathrm{Fn}}-E_{\mathrm{c}}\right)\\ &= \frac{n}{k_{\mathrm{B}}T}\left(\nabla E_{\mathrm{Fn}}-\nabla E_{\mathrm{c}}\right)\end{aligned}$$

the electron current density equation transforms to

$$\begin{aligned}\mathbf{J}_n &= \mu_n n\nabla E_{\mathrm{c}} + \mu_n n\nabla E_{\mathrm{Fn}} - \mu_n n\nabla E_{\mathrm{c}}\\ \mathbf{J}_n &= \mu_n n\nabla E_{\mathrm{Fn}}\end{aligned} \tag{2.94}$$

showing the current density to be directly proportional to the product of the carrier density and the gradient of the quasi Fermi level. Similar treatment of the hole current density equation delivers the analogous result

$$\mathbf{J}_p = \mu_p p\nabla E_{\mathrm{Fp}}. \tag{2.95}$$

This formulation of the current densities give further insight into the Fermi level. Since in thermal equilibrium there are no net current densities $\mathbf{J}_n = \mathbf{J}_p = \mathbf{0}$, one can conclude that

$$\nabla E_{\mathrm{F}} = \mathbf{0}, \tag{2.96}$$

demanding a constant Fermi level throughout the whole semiconductor.

2.2.4 Drift-Diffusion Model

In order to find a dependency between the semiconductor current and charge density (described by the so-called continuity equation), Maxwell's equation shall be considered where both quantities occur as source terms. The set of Maxwell's equations defines a coupled system of partial differential equations governing all macroscopic electromagnetic phenomena [47].

$$\nabla\times\mathbf{E} \;=\; -\frac{\partial}{\partial t}\mathbf{B} \tag{2.97}$$

$$\nabla\times\mathbf{H} \;=\; \mathbf{J} \;+\; \frac{\partial}{\partial t}\mathbf{D} \tag{2.98}$$

$$\nabla\cdot\mathbf{D} \;=\; \rho \tag{2.99}$$

$$\nabla\cdot\mathbf{B} \;=\; 0 \tag{2.100}$$

With the following vectorial and scalar field quantities:

- $\mathbf{E}$:= electric field strength V/m
- $\mathbf{D}$:= electric flux density As/m^2
- $\mathbf{H}$:= magnetic field strength A/m
- $\mathbf{B}$:= magnetic flux density Vs/m^2
- $\mathbf{J}$:= current density A/m^2
- ρ := space charge density As/m^3

The first equation is *Faraday's law of induction*, stating that a temporal change in the magnetic flux density causes a rotation of the electric field. The second equation is *Ampere's Circuital law* describing that an existing current density or a temporal change in the electric flux density[19] leads to a rotation of the magnetic field. Third and fourth equation are the *Gauss' law* for electric and magnetic charges respectively. Here the electrical space charge density is qualified as source of the electrical flux density and the magnetic field is determined to be source free (non-existence of magnetic monopoles).

In the condition of eqs. (2.97) - (2.100) Maxwell's equations are under-determined. Additional equations can be found in the so-called constitutive equations, that are holding material information, connecting the electric and magnetic field strengths with their according flux densities. Here, the case of linear, time invariant, isotropic and region wise homogeneous materials shall be considered. The electric and magnetic flux densities are determined by the permittivity ε_0 and permeability μ_0 of free space

- $\varepsilon_0 := 8.854 \cdot 10^{-12}\ As/Vm$
- $\mu_0 := 4\pi \cdot 10^{-7}\ Vs/Am$

and the electric material polarization $\mathbf{P}$ arising from the dislocation of atoms and electrons (also of ions and molecules) by an acting electric field within the matter

$$\mathbf{D} = \varepsilon_0 \mathbf{E} + \mathbf{P}. \tag{2.101}$$

With the former assumptions the material polarization depends linearly on the electric field strength [51]

$$\mathbf{P} = \varepsilon_0(\varepsilon_\mathrm{r} - 1)\mathbf{E} \tag{2.102}$$

[19] The term $\frac{\partial}{\partial t}\mathbf{D}$, called Displacement current, was added by Maxwell to Ampere's Circuital law resulting in waves as solutions of the electric and magnetic field in the case of non-static fields.

and hence also the flux density is directly proportional to the electric field strength

$$\mathbf{D} = \varepsilon_0 \varepsilon_\mathrm{r} \mathbf{E} = \varepsilon \mathbf{E}. \tag{2.103}$$

Here ε_r is the so-called relative permittivity covering the influence of material on the flux density. For anisotropic materials the permittivity becomes a second-rank tensor (dyad).

A relative magnetic permeability μ_r can be derived analogously, but, since only non-magnetic materials ($\mu_\mathrm{r} \approx 1$) are considered in this thesis work it shall be neglected here. This leaves the relation between magnetic flux density and field strength to be

$$\mathbf{B} = \mu_0 \mathbf{H} = \mu \mathbf{H}. \tag{2.104}$$

At the interface between two materials homogeneity is not valid anymore resulting in non-continuous flux densities. This issue can be addressed by applying appropriate boundary conditions. Those boundary conditions are a direct consequence of Maxwell's equations and can be derived by integrating them on an infinitesimal junction over the interface [52]. Assuming a surface normal vector $\hat{\mathbf{n}}$ pointing from material 2 to material 1 we get

$$\hat{\mathbf{n}} \times \mathbf{E}_1 = \hat{\mathbf{n}} \times \mathbf{E}_2 \tag{2.105}$$
$$\hat{\mathbf{n}} \times \mathbf{H}_1 = \hat{\mathbf{n}} \times \mathbf{H}_2 + \mathbf{J}_\mathrm{s} \tag{2.106}$$
$$\hat{\mathbf{n}} \cdot \mathbf{D}_1 = \hat{\mathbf{n}} \cdot \mathbf{D}_2 + \rho_s \tag{2.107}$$
$$\hat{\mathbf{n}} \cdot \mathbf{B}_1 = \hat{\mathbf{n}} \cdot \mathbf{B}_2. \tag{2.108}$$

The tangential component of the electric field and the normal component of the magnetic flux density have to be continuous at the junction of an interface. The tangential component of the magnetic field and the normal component electric flux density can vary by a surface current density $\mathbf{J}_s[A/m]$ and a surface charge density $\rho_s[As/m^2]$ respectively.

Special cases of the boundary conditions can be defined for interfaces between two dielectric materials, where no surface charges or currents can be applied and hence all vectorial field components have to be continuous at the junction

$$\hat{\mathbf{n}} \times \mathbf{E}_1 = \hat{\mathbf{n}} \times \mathbf{E}_2 \tag{2.109}$$
$$\hat{\mathbf{n}} \times \mathbf{H}_1 = \hat{\mathbf{n}} \times \mathbf{H}_2 \tag{2.110}$$
$$\hat{\mathbf{n}} \cdot \mathbf{D}_1 = \hat{\mathbf{n}} \cdot \mathbf{D}_2 \tag{2.111}$$
$$\hat{\mathbf{n}} \cdot \mathbf{B}_1 = \hat{\mathbf{n}} \cdot \mathbf{B}_2 \tag{2.112}$$

and at the interface between a dielectric material and a perfect electric conductor (material 2)

$$\hat{\mathbf{n}} \times \mathbf{E}_1 = 0 \tag{2.113}$$
$$\hat{\mathbf{n}} \times \mathbf{H}_1 = \mathbf{J}_\mathrm{s} \tag{2.114}$$
$$\hat{\mathbf{n}} \cdot \mathbf{D}_1 = \rho_s \tag{2.115}$$
$$\hat{\mathbf{n}} \cdot \mathbf{B}_1 = 0 \tag{2.116}$$

where all field values inside the perfect conductor have to vanish [52].

For all calculations in the following section a given sufficiently continuous differentiability of the field quantities is assumed. With this assumption the continuity equation can be derived by using Schwarz's theorem. Taking the divergence of Ampere's law eq. (2.98) and considering the fact that the curl of a vector field is always divergence free, one gets

$$\underbrace{\nabla \cdot (\nabla \times \mathbf{H})}_{0} = \nabla \cdot (\mathbf{J} + \frac{\partial}{\partial t}\mathbf{D})$$

$$0 = \nabla \cdot \mathbf{J} + \nabla \cdot \frac{\partial}{\partial t}\mathbf{D}$$
$$0 = \nabla \cdot \mathbf{J} + \frac{\partial}{\partial t}\nabla \cdot \mathbf{D}$$

$$-\frac{\partial}{\partial t}\rho \overset{(2.99)}{=} \nabla \cdot \mathbf{J}. \tag{2.117}$$

With the definition of the divergence of a vectorial field F [53] as a closed integral over surface enclosing an infinitesimal volume

$$\nabla \cdot F = \lim_{V \to 0} \frac{1}{V} \oiint F \cdot d\mathbf{A}, \tag{2.118}$$

eq. (2.117) can be identified as a formulation of charge conservation.

$$-\frac{\partial}{\partial t}\rho = \nabla \cdot \mathbf{J} = \lim_{V \to 0} \frac{1}{V} \oiint \mathbf{J} \cdot d\mathbf{A} \tag{2.119}$$

Charges can not simply vanish but a change in charge density needs to be caused by a net inflow or outflow of current over the surface of the infinitesimal volume.

Now the current density $\mathbf{J}$ in the continuity equation shall be substituted by the semiconductor total current eq. (2.90). For reasons of clarity the single drift and diffusion

parts will be reordered according to the electron current density and hole current density respectively.

$$-\frac{\partial}{\partial t}\rho = \nabla \cdot (\underbrace{q\mu_n n\mathbf{E} + qD_n\nabla n}_{\mathbf{J}_n} + \underbrace{q\mu_p p\mathbf{E} - qD_p\nabla p}_{\mathbf{J}_p}) \tag{2.120}$$

The charge density in a semiconductor can be expressed by [52]

$$\rho = q(p - n + N_\mathrm{D} - N_\mathrm{A}) \tag{2.121}$$

where N_D and N_A are the amount of ionized donors and acceptors respectively. Assuming them to be constant in time eq. (2.120) becomes

$$-q\frac{\partial}{\partial t}(p - n) = \nabla \cdot (\underbrace{q\mu_n n\mathbf{E} + qD_n\nabla n}_{\mathbf{J}_n} + \underbrace{q\mu_p p\mathbf{E} - qD_p\nabla p}_{\mathbf{J}_p}). \tag{2.122}$$

Since charge conservation must be valid for each kind of carrier species separately, the continuity equation can be decomposed into two equations, one for each kind of carrier. Here one has to consider that in a semiconductor electron hole pairs can be generated or electrons and holes can recombine. In eq. (2.122) this is considered by the difference of p and n, since for each generated electron also a hole with equal amount of charge (but opposite sign) is generated, leaving the total charge density unchanged. When separating the equation this needs to be handled by adding a coupled net recombination rate R for electron-hole pairs, with opposite sign in each equation

$R =$ Recombination Rate $-$ Generation Rate.

After the separation for the two different carrier species and adding of the net recombination rate we get the so far decoupled equation system

$$q\frac{\partial}{\partial t}n = \nabla \cdot (q\mu_n n\mathbf{E} + qD_n\nabla n) - R \tag{2.123}$$

$$-q\frac{\partial}{\partial t}p = \nabla \cdot (q\mu_p p\mathbf{E} - qD_p\nabla p) + R. \tag{2.124}$$

It can easily be seen that the summation of eqs. (2.123) and (2.124) results in the original continuity equation (2.122).

Now in this formulation each equation depends on the unknown and not independent vectorial electric field distribution. With the electric fields vector components the equation system has five unknown variables for only two equations. In order to derive an equation to determine the electric field, Maxwell's equations shall be revisited. Because of the divergence free character of a rotational field, from the Gauss law for magnetic charges

eq. (2.100) one can see that the magnetic flux density can be expressed by a vector potential

$$\nabla \cdot \mathbf{B} = \nabla \cdot (\nabla \times \mathbf{A}) = 0. \tag{2.125}$$

The potential $\mathbf{A}$ is commonly referred to as the magnetic vector potential. With

$$\mathbf{B} = \nabla \times \mathbf{A}. \tag{2.126}$$

Faraday's law can be transformed by using Schwarz's theorem again

$$\nabla \times \mathbf{E} = -\frac{\partial}{\partial t} \nabla \times \mathbf{A}$$

$$\nabla \times (\mathbf{E} + \frac{\partial}{\partial t}\mathbf{A}) = 0. \tag{2.127}$$

Since

$$\nabla \times (\nabla \phi) = 0$$

eq. (2.127) identifies $\mathbf{E} + \frac{\partial}{\partial t}\mathbf{A}$ as a gradient field and an expression for $\mathbf{E}$ as a sum of a scalar and vector potential is found

$$\mathbf{E} = -\nabla \phi - \frac{\partial}{\partial t}\mathbf{A}. \tag{2.128}$$

The negative sign for the scalar potential is only a historical definition in order to carry the conventional sign of the voltage [47]. Using the formulation of the potentials eqs. (2.126) and (2.129) Maxwell's equations change their appearance[20]. As it seems reasonable to describe carrier transport processes in semiconductors, for the further analysis the simplification of electro-quasistatics shall be used, where the temporal changes of the flux densities are seen as slow enough to neglect induction phenomena but the effect of displacement currents are still considered. It can easily be seen that from a vanishing temporal derivative magnetic flux density, a vanishing time derivative of the magnetic vector potential follows, and hence the expression of the electric field simplifies to

$$\mathbf{E} = -\nabla \phi. \tag{2.129}$$

With these assumptions Gauss law for electric charges 2.99 now becomes

$$\begin{aligned} \nabla \cdot (\varepsilon \mathbf{E}) &= \rho \\ \nabla \cdot (\varepsilon \nabla \phi) &= -\rho. \end{aligned} \tag{2.130}$$

[20] Interestingly Maxwell postulated the equation in the sense of potentials, while the nowadays commonly known form originates from O. Heaviside [47].

In the case of a spatially independent permittivity the scalar potential ϕ satisfies the *Poisson equation*

$$\Delta\phi = -\frac{\rho}{\varepsilon}. \tag{2.131}$$

With substituting eq. (2.129) into the two continuity equations (2.123) and (2.124), now only a third unknown (given by the potential) remains which can be expressed by eq. (2.130). Those three equations together form a coupled non-linear system of equations.

$$q\frac{\partial}{\partial t}n = \nabla\cdot(-q\mu_n n\nabla\phi + qD_n\nabla n) - R \tag{2.132}$$

$$-q\frac{\partial}{\partial t}p = \nabla\cdot(-q\mu_p p\nabla\phi - qD_p\nabla p) + R \tag{2.133}$$

$$\nabla\cdot(\varepsilon\nabla\phi) = -\rho \overset{(2.121)}{=} -q(p - n + N_\mathrm{D} - N_\mathrm{A}). \tag{2.134}$$

This equation system is called the drift-diffusion model and will be used in this thesis work to describe the carrier transport in semiconductors.

Since the assumptions of electro-quasistatics were used, also the effects of this simplification on the carrier transport in semiconductors shall be analyzed. Substituting the magnetic vector potential into Ampere's law (2.98) and taking the time derivative of the equation yields

$$\begin{aligned}\nabla\times\nabla\times\frac{\partial}{\partial t}\mathbf{A} &= \mu\frac{\partial}{\partial t}(\,\mathbf{J} + \varepsilon\frac{\partial}{\partial t}\mathbf{E})\\ 0 &= \frac{\partial}{\partial t}(\,\mathbf{J}_\sigma + \mathbf{J}_\mathrm{D} + \frac{\varepsilon}{\sigma}\frac{\partial}{\partial t}\mathbf{J}_\sigma)\end{aligned} \tag{2.135}$$

Here the current density is considered to consist of a drift current (denoted as $\mathbf{J}_\sigma$) and a diffusion current (denoted as $\mathbf{J}_\mathrm{D}$) furthermore the displacement current is expressed in terms of the drift current. Since the time derivative of the total current (including displacement current) vanishes, the electro-quasistatic assumptions only allow a constant total current density ($\mathbf{J}_\mathrm{c}$), but to show transient behavior among the single components

$$\mathbf{J}_\mathrm{c} = \mathbf{J}_\sigma + \mathbf{J}_\mathrm{D} + \frac{\varepsilon}{\sigma}\frac{\partial}{\partial t}\mathbf{J}_\sigma. \tag{2.136}$$

Equation (2.136) defines an inhomogeneous ordinary differential equation for the drift current

$$\frac{\partial}{\partial t}\mathbf{J}_\sigma = -\frac{\sigma}{\varepsilon}\mathbf{J}_\sigma - \frac{\sigma}{\varepsilon}(\,\mathbf{J}_\mathrm{c} - \mathbf{J}_\mathrm{D}\,), \tag{2.137}$$

with the solution

$$\mathbf{J}_\sigma = \mathbf{J}_\sigma(0)e^{-\frac{\sigma}{\varepsilon}t} + (\mathbf{J}_\mathrm{c} - \mathbf{J}_\mathrm{D})[1 - e^{-\frac{\sigma}{\varepsilon}t}]. \tag{2.138}$$

The derivative of eq. (2.138) gives information on the displacement current

$$\frac{\varepsilon}{\sigma}\frac{\partial}{\partial t}\mathbf{J}_\sigma = [(\mathbf{J}_\mathrm{c} - \mathbf{J}_\mathrm{D}) - \mathbf{J}_\sigma(0)]e^{-\frac{\sigma}{\varepsilon}t}. \tag{2.139}$$

From eq. (2.139) it can be seen that the displacement current has to vanish from sufficiently large time t and hence also the time derivative of the electric field has to show the same evanescent behavior.

In the non-transient case ($t \to \infty$) the drift current density becomes

$$\mathbf{J}_\sigma \overset{t\to\infty}{=} \mathbf{J}_\mathrm{c} - \mathbf{J}_\mathrm{D} \tag{2.140}$$

and hence without any external bias

$$\mathbf{J}_\sigma = -\mathbf{J}_\mathrm{D}. \tag{2.141}$$

For the charge density we get from Gauss law

$$\begin{aligned} \nabla \cdot \frac{\varepsilon}{\sigma}\mathbf{J}_\sigma &= \rho \\ \nabla \cdot \mathbf{J}_\sigma &= \frac{\sigma}{\varepsilon}\rho \end{aligned} \tag{2.142}$$

and from continuity equation

$$-\frac{\partial}{\partial t}\rho = \nabla \cdot (\mathbf{J}_\sigma + \mathbf{J}_\mathrm{D}). \tag{2.143}$$

Comparing both equations one gets an ordinary differential equation describing the temporal dependence of the charge density

$$\frac{\partial}{\partial t}\rho = -\frac{\sigma}{\varepsilon}\rho - \nabla \cdot \mathbf{J}_\mathrm{D} \tag{2.144}$$

with the solution

$$\rho = \rho(0)e^{-\frac{\sigma}{\varepsilon}t} - \frac{\varepsilon}{\sigma}\nabla \cdot \mathbf{J}_\mathrm{D}[1 - e^{-\frac{\sigma}{\varepsilon}t}]. \tag{2.145}$$

In case of steady state, ρ becomes

$$\rho \overset{t\to\infty}{=} -\frac{\varepsilon}{\sigma}\nabla \cdot \mathbf{J}_\mathrm{D} \tag{2.146}$$

showing that the resulting charge density at any point $\mathbf{R}$ is caused by a non-vanishing source strength of the diffusion current.

2.3 Recombination processes

In this section the net recombination rate R, appearing as a source term in eqs. (2.132) and (2.133) shall be quantified. In general this recombination rate is given by a superposition of different effects causing radiative and non-radiative extinction of carriers. Here the most importent effects considered in the semiconductor models used in this thesis work shall be introduced.

2.3.1 Radiative recombination

In direct semiconductors like Gallium Nitride radiative recombination processes are likely to occur, since here the maxima of the valence and conduction band fall together in momentum-space (***k**-space*) only separated by the energy of the bandgap. This means that an electron and a hole can directly recombine without having other particles (like phonons) involved for the sake of momentum conservation. In this way an electron-hole pair get extinct by the recombination of an electron from the conduction band and a hole from the valence band under the emission of a photon with the band energy E_g, taking care of the conservation of energy. Fig 2.12 exemplifies this process.

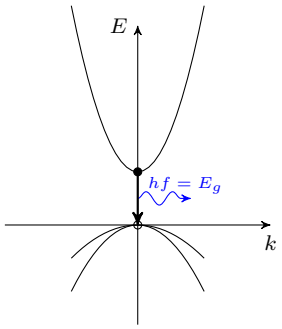

Figure 2.12: Radiative recombination: Extinction of one electron-hole-pair by emission of one photon carrying the bandgap energy

With this considerations it becomes evident that electrons and holes are equally needed in order for such an interband transition to take place. Therefore the radiative recombination rate should be proportional to the product of electron and hole densities

$$R_n = R_p = Bnp, \tag{2.147}$$

where B represents the average capture coefficient [30, 52].

In order to find the radiative net recombination rate R_{rad} also generation processes with the rate G_{th}, existing due to the constantly occurring thermal excitation of carriers, need to be considered

$$R_{rad} = R_n - G_n = R_p - G_p = Bnp - G_{th}. \tag{2.148}$$

At thermal equilibrium the net recombination rate vanishes since recombination and generation processes equal each other out and we find the generation rate to be [30, 52]

$$\begin{aligned} Bn_0p_0 - G_{th} &= 0 \\ \Rightarrow G_{th} &= Bn_0p_0. \end{aligned} \tag{2.149}$$

It makes sense to assume that this relation is also valid outside of thermal equilibrium since thermal generation merely depends on temperature and the band gap energy and hence is independent of the actual carrier densities (as long the valence band isn't heavily depleted of electrons to be thermally excited), so that the radiative net recombination rate can be given as

$$R_{rad} = B(np - n_0p_0) = B(np - n_\mathrm{i}^2). \tag{2.150}$$

Since n, p are the equilibrium carrier densities increased by the carrier densities excited by the outer stimulus $\Delta n = \Delta p$

$$n = n_0 + \Delta n \tag{2.151}$$

$$p = p_0 + \Delta p \tag{2.152}$$

the expression for $\mathbf{R}_{rad}$ can be rearranged for further insight:

$$\begin{aligned} R_{rad} &= B[(n_0 + \Delta n)(p_0 + \Delta p) - n_0p_0] \\ &= B[\Delta np_0 + \Delta pn_0 + \Delta n\Delta p] \\ &= B[\Delta n(p_0 + \Delta p) + \Delta pn_0] \end{aligned}$$

and with $\Delta n = \Delta p$

$$R_{rad} = B\Delta n(n_0 + p_0 + \Delta n). \tag{2.153}$$

This shows that the radiative net recombination rate is proportional to the product of the excess carrier density and the sum of the equilibrium densities and the excess carrier density. With this, some simplifications for border cases can be derived.

Low-Injection-Regime

In the low injection regime where the density of the majority carriers exceeds the excess carrier density by orders of magnitude, the expression within the braces in eq. (2.153) is strongly dominated by the majority carrier density. Therefore the densities of minority and excess carriers can be neglected.

For the example of a p-semiconductor one would could make the following assumptions, see fig. fig. 2.13:

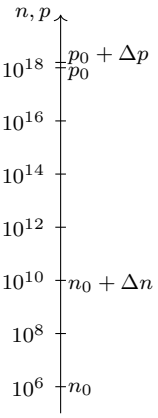

Figure 2.13: Example of carrier densities in a p-semiconductor right after an outer stimulus led to the generation of excess electron-hole-pairs.

$$p_0 \approx N_A >> n_0 + \Delta n. \tag{2.154}$$

Using this assumption, the radiative net recombination rate simplifies to

$$R_{rad} \approx BN_A\Delta n. \tag{2.155}$$

Since R was defined as the difference of the rates of recombination and thermal generation, the temporal change of the excess carrier density[21] is given by the negative net recombination rate

$$\frac{d\Delta n(t)}{dt} = -R = -BN_A\Delta n(t). \tag{2.156}$$

The solution of this ordinary differential equation states an exponential decay of the excess carrier density conveying the system back to thermal equilibrium

$$\Delta n(t) = \Delta n(0)e^{-BN_At} = \Delta n(0)e^{-\frac{t}{\tau_n}} \tag{2.157}$$

and is identifying the inverse of the product of BN_A as a time constant

$$\tau_n = \frac{1}{BN_A} \tag{2.158}$$

commonly referred to as carrier lifetime.

An analogous approach using the example of a n-semiconductor would yield the carrier lifetime for the excess hole densities as

$$\tau_p = \frac{1}{BN_D}. \tag{2.159}$$

[21] introduced by an outer stimulus

High-Injection-Regime

The second border case is the high injection scenario where the excess carriers density even surpasses the majority carrier density by a large amount $\Delta n >> n_0 + p_0$. Here the bracketed term of eq. (2.153) is now dominated by Δn, furthermore the non-equilibrium densities can be approximated by the density of the excess carriers

$$p \approx n \approx \Delta n. \tag{2.160}$$

With these assumptions the radiative net recombination rate becomes approximately

$$R_{rad} \approx B\Delta n^2 \approx Bn^2. \tag{2.161}$$

2.3.2 Non-Radiative Recombination

Besides the emission of light, recombination can also take place under the emission of phonons[22] or due to impacts among the carriers, where the energy is transmitted from one carrier to another[23]. The most common sources for phonon emissions are defects within the crystal like impurities or dislocations [30, 52]. These defects, commonly referred to as traps, introduce additional states that can be occupied by the carriers. While in most applications recombination due to trap states are undesired processes in this thesis work the introduction of surface traps by a suitable treatment of the semiconductor surface shall be used to control the sensor behavior. Therefore, the model now introduced, represents the trapping mechanism used to model the corresponding surface recombination processes as shown in section 1.2.2.

Recombination through trap states (Shockley-Read-Hall-Recombination)

In the early 1950s, based on insights given by Robert N. Hall, William Shockley and William Thornton Read first developed a model representing the recombination of electrons and hole through traps, energetically suited within the band gap [25, 26]. For the derivation of the model four basic processes have been considered, which are electron capture, electron emission, hole capture and hole emission (see fig. 2.14).

[22] The quasi-particles assigned to vibrations of crystal lattice (heat)
[23] Auger Processes

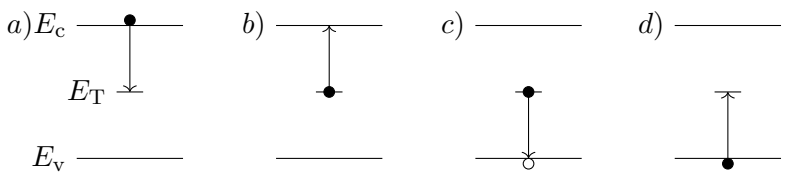

Figure 2.14: a) Electron capture, b) Electron emission, c) Hole capture, d) Hole emission Here the lines in the center of the band gap represent trap levels, opaque circles depicture electrons, transparent circles holes [25]

Electron capture: In the case of electron capture, an electron from the conduction occupies a trap level that lies at an energetic position within the band gap. In this process the energy difference of the conduction band and the trap level is most commonly conserved by emission of heat[24]. Obviously the rate of electron capture depends on the electron density in the conduction band, the trap density N_t and their regarding probability of occupation f_t and average capture cross section c_n [25, 52]

$$R_e = nN_t(1 - f_t)c_n. \tag{2.162}$$

Here the expression $1 - f_t$ represents the probability that a trap level is free and hence can be occupied by an electron.

Electron emission: For electron emission electrons leap from the trap level into the conduction band due to thermal excitation. Therefore the rate of emission needs to be proportional to the amount of occupied trap levels and the average emission rate. In general the emission rate would also depend on the availability of free states in the conduction band, for non-degenerated semiconductors this factor can be ignored though [25, 52].

$$G_e = N_t f_t e_n. \tag{2.163}$$

Hole capture: Hole capture means a recombination of an electron suited at a trap level with an hole in the valence band. Hence the corresponding recombination rate is proportional to the hole density, the amount of occupied trap levels and an averaged hole capture rate c_p [25, 52].

$$R_h = pN_t f_t c_p. \tag{2.164}$$

Hole emission: Hole generation through a trap state occurs when electrons from the valence band get thermally excited to occupy a trap level. For this process an excitation energy

[24] Energy conservation by emission of a photon is also possible

lower than the band gap energy is sufficient. Assuming non-degenerated semiconductors the valence band can be expected to be well-filled, whereby the dependence of the electron availability can be neglected. This leaves the process to be proportional to the amount of free trap states $N_t(1 - f_t)$ and the average emission rate [25, 52].

$$G_h = N_t(1 - f_t)e_p. \tag{2.165}$$

Assuming further that both carriers equalize among each other[25] the net recombination rate yields[26] [25, 52]

$$R_{\mathrm{SRH}} = N_t \frac{c_n c_p (np - n_{\mathrm{i}}^2)}{c_n(n + n_1) + c_p(p + p_1)}. \tag{2.166}$$

Where the respective auxiliary carrier densities n_1 and p_1 are

$$n_1 = n \exp(\frac{E_{\mathrm{T}} - E_{\mathrm{Fn}}}{k_{\mathrm{B}}T}) \overset{2.65}{=} N_{\mathrm{c}} \exp(\frac{E_{\mathrm{T}} - E_{\mathrm{c}}}{k_{\mathrm{B}}T}) \overset{2.65}{=} n_{\mathrm{i}} \exp(\frac{E_{\mathrm{T}} - E_{\mathrm{i}}}{k_{\mathrm{B}}T}) \tag{2.167}$$

and

$$p_1 = p \exp(\frac{E_{\mathrm{Fp}} - E_{\mathrm{T}}}{k_{\mathrm{B}}T}) \overset{2.66}{=} N_{\mathrm{v}} \exp(\frac{E_{\mathrm{v}} - E_{\mathrm{T}}}{k_{\mathrm{B}}T}) \overset{2.66}{=} n_{\mathrm{i}} \exp(\frac{E_{\mathrm{i}} - E_{\mathrm{T}}}{k_{\mathrm{B}}T}) \tag{2.168}$$

the carrier densities when the Fermi level would fall together with the regarding trap level [25]. Considering eqs. (2.167) and (2.168) it becomes obvious that only those defects having an energy trap level suited close to the middle of the band gap[27] can act as effective recombination centers [30]. Deviations of the trap level from the intrinsic Fermi level would lead to an exponential growth of one of the two terms of the denominator, and hence leading to a strong decay of recombination rate. This enables to simplify eq. (2.166) to [30]

$$R_{\mathrm{SRH}} = N_{\mathrm{T}} \frac{c_n c_p (np - n_{\mathrm{i}}^2)}{c_n(n + n_{\mathrm{i}}) + c_p(p + n_{\mathrm{i}})}. \tag{2.169}$$

Auger-Recombination

Auger recombination processes occur due to the collision of carriers. For example this could happen when two conduction band electrons collide in the vicinity of a hole. Through the impact, energy and momentum are transferred from one electron to another while the energy emitting electron recombines with a hole [54]. Analogous to this, the recombination energy can also be transferred to second hole instead. On the other hand,

[25] Meaning $R_{ec} - G_{ee} = 0$ and $R_{hc} - G_{he} = 0$

[26] A thorough and comprehensive derivation of the following formula is given in [25]

[27] More precise at energies around the intrinsic Fermi level.

electron-hole-pairs can be generated due to impact ionization. This gives rise to four different processes to be analysed in order to find the Auger net recombination rate.

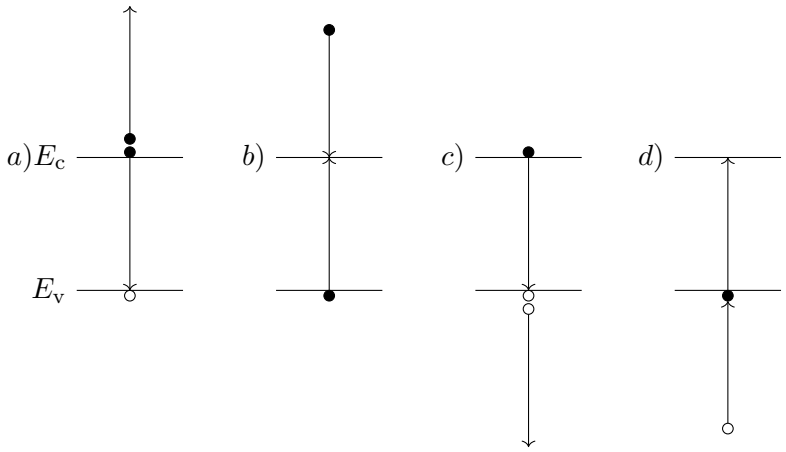

Figure 2.15: a) Electron capture, b) Electron emission, c) Hole capture, d) Hole emission The four processes contributing to Auger recombination [52]

For electron capture an electron from the conduction band recombines with a hole in the valence band while giving it's excitation energy to another electron in the conduction band. After this the high energetic electron will go through a relaxation process, most likely to conserve the energy by emission of lattice vibrations (heat). This capture process requires three particles, two electrons in the conduction band and a hole from the valence band to occur, therefore the recombination rate depends quadratic on electron density and a linearly on the hole density [52]:

$$R_e = c_n n^2 p. \tag{2.170}$$

Electron emission in the sense of a Auger process means the generation of electron-hole-pairs due to impact ionization. Here the impact is caused by an electron having at least twice the band gap energy. In this way the high energetic electron remains in the conduction band after the abrupt relaxation that causes the valence electron to leap into the conduction band. The emission rate is proportional to the electron density [52]:

$$G_e = e_n n. \tag{2.171}$$

Analogous to electron capture the recombination energy can be taken by a hole in the valence band. For this process only one electron in the conduction band but two holes in

the valence band are needed. The corresponding rate is than given by [52]:

$$R_h = c_p n p^2. \tag{2.172}$$

The fourth process is hole emission. Here, impact ionization caused by a high energetic hole in the valence band leads to a valence electron leaping to the conduction band. The rate for this process can be given by [52]:

$$G_h = e_p p. \tag{2.173}$$

Again the principle of detail balancing can be used to reduce the number of unknown variables for the rate equation. We get

$$\begin{aligned} R_e - G_e = 0 &= c_n n_0^2 p_0 - e_n n_0 \\ \Rightarrow e_n &= c_n n_\mathrm{i}^2 \end{aligned} \tag{2.174}$$

and

$$\begin{aligned} R_h - G_h = 0 &= c_p n_0 p_0^2 - e_p p_0 \\ \Rightarrow e_p &= c_p n_\mathrm{i}^2. \end{aligned} \tag{2.175}$$

With this the Auger net recombination rate is given as the sum of both individual net rates, which yield [52]

$$\begin{aligned} R_{Aug} &= c_n n^2 p - c_n n n_\mathrm{i}^2 + c_p n p^2 - c_p p n_\mathrm{i}^2 \\ &= (c_n n + c_p p)(np - n_\mathrm{i}^2). \end{aligned} \tag{2.176}$$

2.4 Fermi-Level-Pinning

In the following the consequences of a pinned Fermi level at the nanowire surface shall be discussed in the context of the sensing behavior.

When the Fermi level is pinned at the surface, the neutrality condition demands a bending of the bands in the bulk material in order to match the surface level [32]. Hence a space charge region is built up in the vicinity of the surface. Given the example of a n-semiconductor

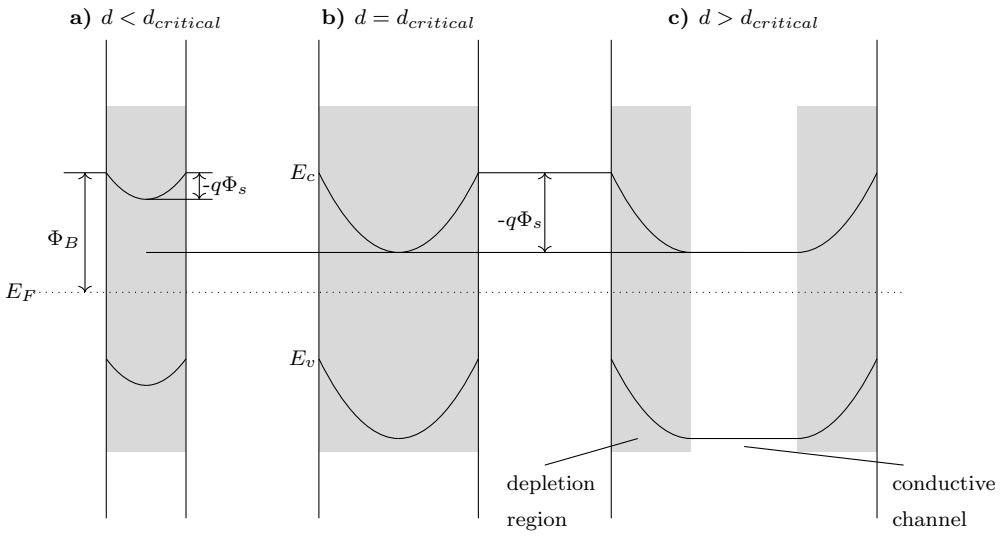

Figure 2.16: Shaded areas represent a space charge region. a) The diameter of the nanowire is below the critical diameter. Here, the carrier concentration in the bulk volume does not suffice to compensate for the Fermi level difference. Hence the a smaller potential barrier compared to wires with larger diameters builts up. The nanowire becomes completely depleted, the Fermi level drops as the scarcity of carriers becomes the dominating factor. b) The space charge region is still spanned throughout the whole wire. Therefore nanowires around the critical diameter are still completely insulating, but the volume is just big enough to built up a potential barrier to equalize the bulk Fermi level with the Fermi level of the surface. c) In nanowires bigger than the critical diameter the space charge region only spans over an outer part of the wire. In the inner a conducting channel remains that defines the resistivity of the wire [39].

the bands in the bulk are bent upwards, creating a potential barrier for electrons to overcome in order to reach the surface. Hereby sensor principles based on surface recombination are hindered as electrons are confined to an inner cylindrical region around the nanowire center. This inner region forms a conducting channel while the depleted outer parts are insulating. Hence, the resistivity of the nanowire does not depend on the cross section of the nanowire but only on the cross section of the inner region [39]. Furthermore, in thin wires below a critical diameter, the volume charge in the depleted area is too small to build up a sufficiently large potential barrier. This case results in vastly depleted and hence insulating nanowires. Here the potential barrier is lower compared to wires above the critical diameter [39], as can be seen in fig.fig. 2.16. In order to predict the critical diameter of a nanowire the Schottky-model shall be used to determine the

surface potential. In this model an abrupt transition from the neutral region to a space charge region with a constant charge distribution is assumed. Furthermore the potential is assumed to be pinned at the surface. In this case the surface potential is simply given by difference of the different Fermi levels [32]

$$-q\Phi_s = E_{FB} - E_{FS}. \tag{2.177}$$

On the other hand the potential can be found by solving the poisson equation for cylindrical coordinates. The Laplace operator in cylindrical coordinates is given by

$$\Delta\Phi = \frac{1}{\rho}\frac{\delta}{\delta\rho}\left(\rho\frac{\delta\Phi}{\delta\rho}\right) + \frac{1}{\rho^2}\frac{\delta^2\Phi}{\delta^2\phi} + \frac{\delta^2\Phi}{\delta^2\rho}.$$

Because of the cylindrical symmetry of the potential in nanowires the derivatives according to the z-coordinate and the angle ϕ become zero. Hereby the poisson equation to solve[28]

$$\Delta\Phi_s = \frac{qN_D}{\epsilon} \tag{2.178}$$

simplifies to

$$\frac{1}{\rho}\frac{\delta}{\delta\rho}\left(\rho\frac{\delta\Phi}{\delta\rho}\right) = \frac{qN_D}{\epsilon} \tag{2.179}$$

with the solution [39, 40]

$$\Phi_s = \frac{qN_D}{4\epsilon}\left(r_{ch}^2 - \frac{d^2}{4} - 2r_{ch}^2\ln\frac{2r_{ch}}{d}\right). \tag{2.180}$$

Here r_{ch} is the radius of the inner conducting channel and d the diameter of the penetrating space charge region.

[28] Assuming the charge density only arises from the completely ionized impurities.

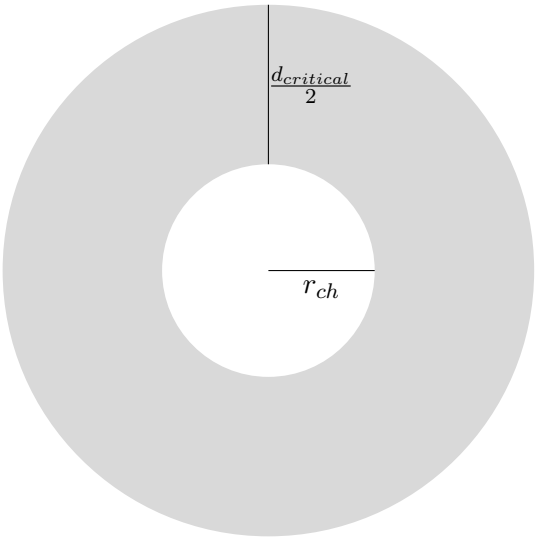

Figure 2.17: Nanowire cross-section. The shaded area represents the depleted space charge region, the inner circle with radius r_{ch} shows the conducting channel.

The critical diameter is found where the radius of the conducting channel vanishes ($r_{ch} = 0$). Here the surface potential is given by

$$\Phi_s = \frac{-qN_D}{16\epsilon} d_{critical}^2. \tag{2.181}$$

Substituting this solution in eq. (2.177) and solving for the critical diameter yields

$$d_{critical} = \sqrt{\frac{16\epsilon\,(E_{FB} - E_{FS})}{q^2 N_d}}. \tag{2.182}$$

Interestingly this result for the potential differs only by a factor $2\sqrt{2}$ from that of the same problem in 1D[29], for example to be found in [32].

[29] Where in contrast to the cylindrical geometry two of the coordinates are assumed to have infinite extension.

3 Simulation Results for Electronic Properties of GaN Nanowires

3.1 Introduction of Analyzed Sensor Structures

The fundamental evaluation set-up is given by the following four nanowire base structures:

- Resistors
- PN-Diodes
- N^+N^--Diodes
- Transistors.

For each group of sensor devices, analyses have been carried out according to variations of morphology and doping concentrations. Furthermore two different basic sensor concepts have been analyzed. The first concept is assuming sensors offering a GaN/functional group/Air interface where the sensing behavior is based on charge transfer between GaN and the functional groups due to introduced surface states. Additionally to this concept simulations have been carried out having a 10 nm thick oxide layer incorporated at nanowire/Air-interfaces, as experimental results by our partners at TU Braunschweig indicated that structures without oxidised surfaces won't be feasible in practice. Therefore the functional groups have to be realized on the oxide layer rather than on the GaN surface directly in order to provide sensor selectivity by designed molecule capture. The thusly introduced dipole moment at the nanowire surface is the source of the electrostatic effects, providing the sensor behavior. In the simulations this effect is considered by adding a fixed surface charge on top of the oxide layers. For both concepts the quantities of trap levels and energies as well as the fixed surface charges are simply given by parameter variations as data by atomistic DFT-calculations was still lacking by the end of this thesis work. All simulations have been solved as 2D-problems making use of the wire's cylindrical symmetry. A graphical overview of the simulated sensor structures is given in fig. 3.1.

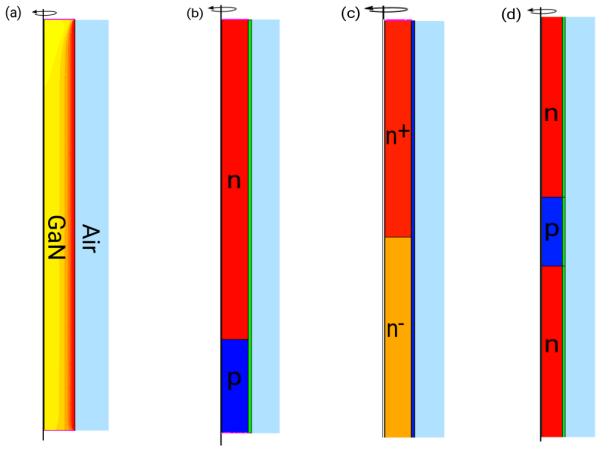

Figure 3.1: Overview of analyzed sensor structures: a) pure nanowire resistor, b) pn-diode, c) n^+n^--junction, d) npn-transistor

3.2 Discussion and Results

Throughout the simulations linearity (or more precisely the linearity deviation "ΔL") and responsitivity have been determined to serve as quality metrics for the individual sensor designs. Those two metrics are evaluated according to

$$Response = \frac{\Delta I}{I_{\text{base}}} \cdot 100\%$$

with $\Delta I = I_{\text{gas}} - I_{\text{air}}$ and $I_{base} = I_{air}$ if the target gas is reductive[1] or $\Delta I = I_{\text{air}} - I_{\text{gas}}$ and $I_{base} = I_{gas}$ when the target gas is oxidative[2], and

$$\Delta L = max\left(\left|\frac{Response(C_{\text{s}}) - g(C_{\text{s}})}{g(C_{\text{s}})}\right|\right) \cdot 100\%,$$

where C_{s} is the surface charge density and $g(C_{\text{s}})$ represents the response of a linear sensor.

[1] Correspondent to sensors aiming for an increase in conductance
[2] Aiming for an increasing resistance

3.2.1 Resistors

The resistors have been modeled as single nanowires electrically contacted at bottom and top. For all resistors a wire length of 1 μm was assumed. As mentioned before, simulations have been carried out with and without oxide layer at the nanowire surface. The difference between those simulation results shall also be discussed in this section. Both structures are depicted in figs. 3.2 and 3.3. In the left part of this figure the radial 2D-cut of the actual cylindrical nanowire shall emphasize the general simulation area, the right part focuses on meshing. As can be seen the mesh is refined at the nanowire/oxide-interface where the surface charge is seated, as changes of analyzed values are expected to be most prominent in the vicinity of this region.

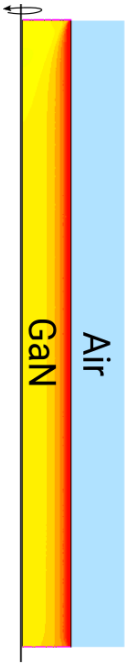

Figure 3.2: 2D cut of the cylindrical nanowire resistor. Here no oxide layer is added to the surface, offering a direct GaN/Air-interface. In this set-up carrier exchange between GaN and gas molecule affects the wire resistance.

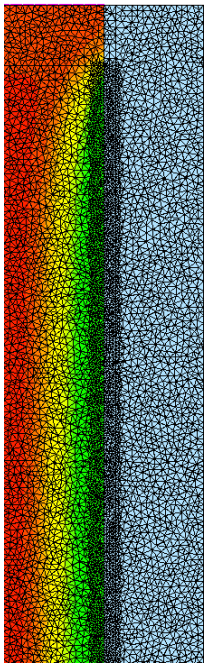

Figure 3.3: Nanowire with illustrated current density and meshing. The mesh is refined at a GaN/Oxide-interface region. Here surface charges oriented at Oxide/Air-interface are the source of changes in the nanowire resistance.

3.2.2 Results using surface trap charge model

All surface traps have been modeled with a trap charge energy 0.5 eV below the conduction band energy. The trap charge density has been varied between 0 cm^{-2} and $1 \cdot 10^{12}$ cm^{-2} representing the amount of docked gas molecules at the nanowire surface. In order to explain the sensors working principle the results for nanowires having diameters of 80 nm, 100 nm and 150 nm at doping concentrations of $1 \cdot 10^{17}$ cm^{-3} and $6.25 \cdot 10^{17}$ cm^{-3} are shown exemplarily and will be discussed compared to each other at the end of the section.

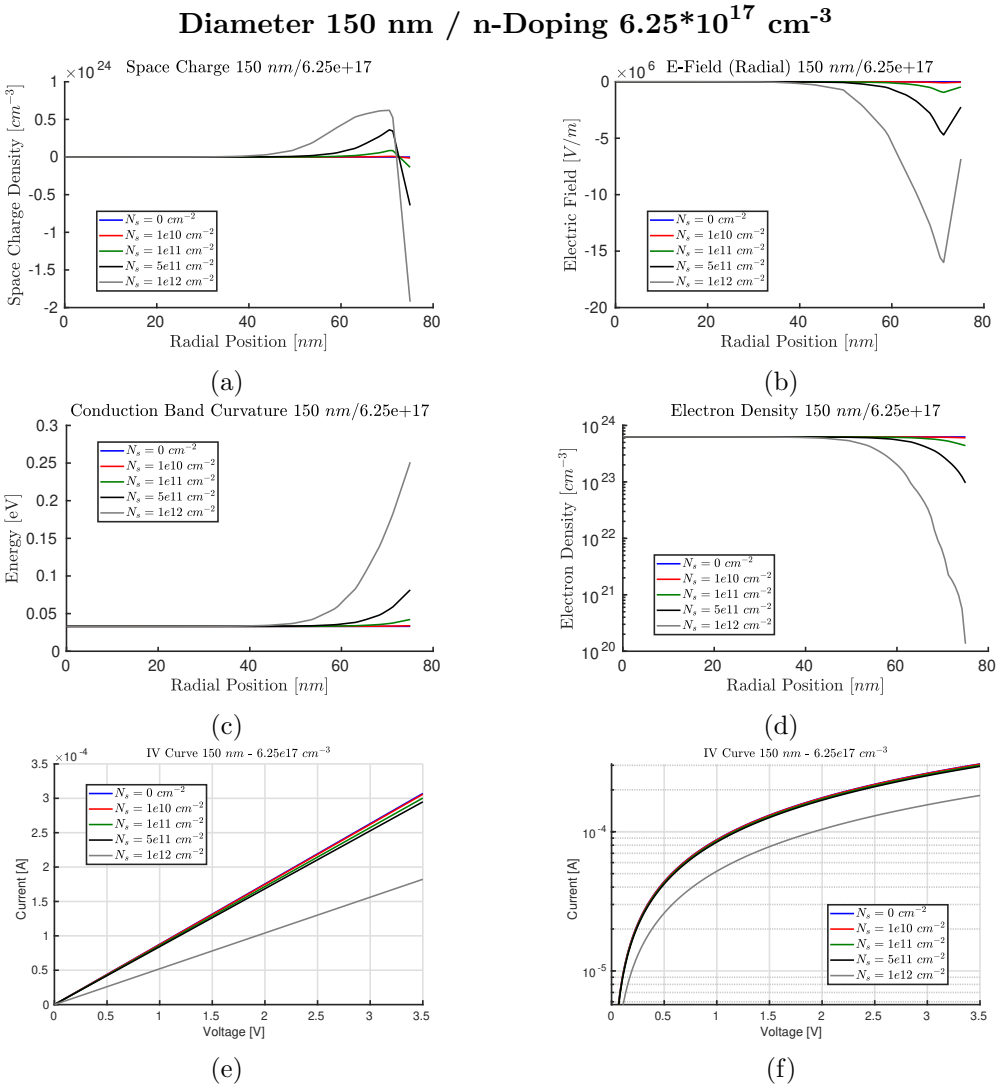

Figure 3.4: a) A positive space charge is built up in the vicinity of the surface as a consequence of captured electrons by the negative surface trap charges. b) Due to the charges an electric field has established hindering electrons from moving further towards the surface. c) The conduction band rises close to the surface as the electron density drops in this region, as can be seen in figure d). The results indicate an inner conducting channel of around 40 nm radius still unaffected at the highest surface charge of 10^{12} cm^{-2}. Figures e) and f) show the sensors IV curve in linear and semi-logarithmic scale.

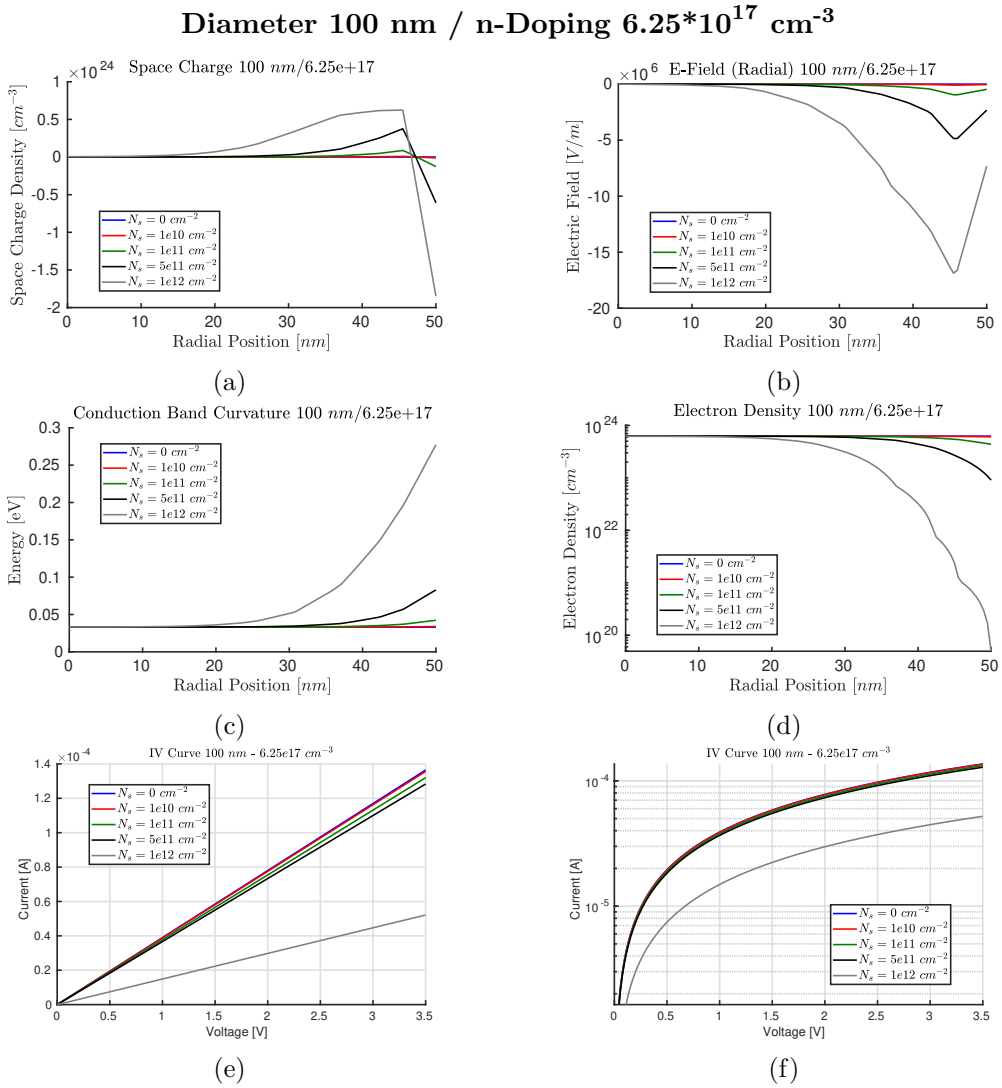

Figure 3.5: Compared to the results of the 150 nm diameter, the linear plot of the space charge now shows a distinct plateau close to the surface, which lies in the order of the doping concentration. This is due to the fact that in this region the electron density is orders of magnitudes lower than the doping concentration. The overall affected area also spans wider, leaving an unaffected channel having a radius of only around 15 nm at the highest surface charge density.

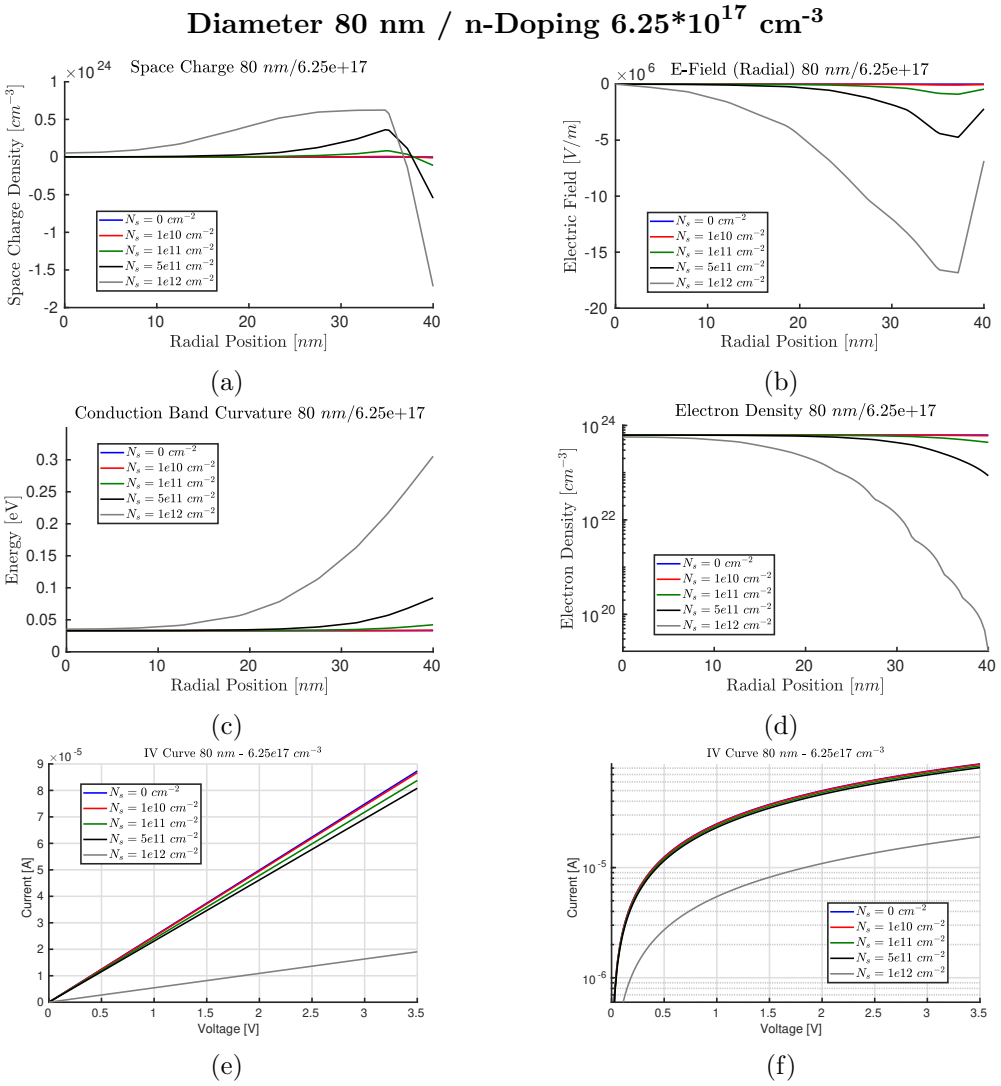

Figure 3.6: When decreasing the nanowire diameter further to 80 nm, the carrier depletion reaches deeper into the nanowire, now affecting the whole nanowire at the highest surface charge density. An outer annulus with a width around 15 nm is largely depleted in this case.

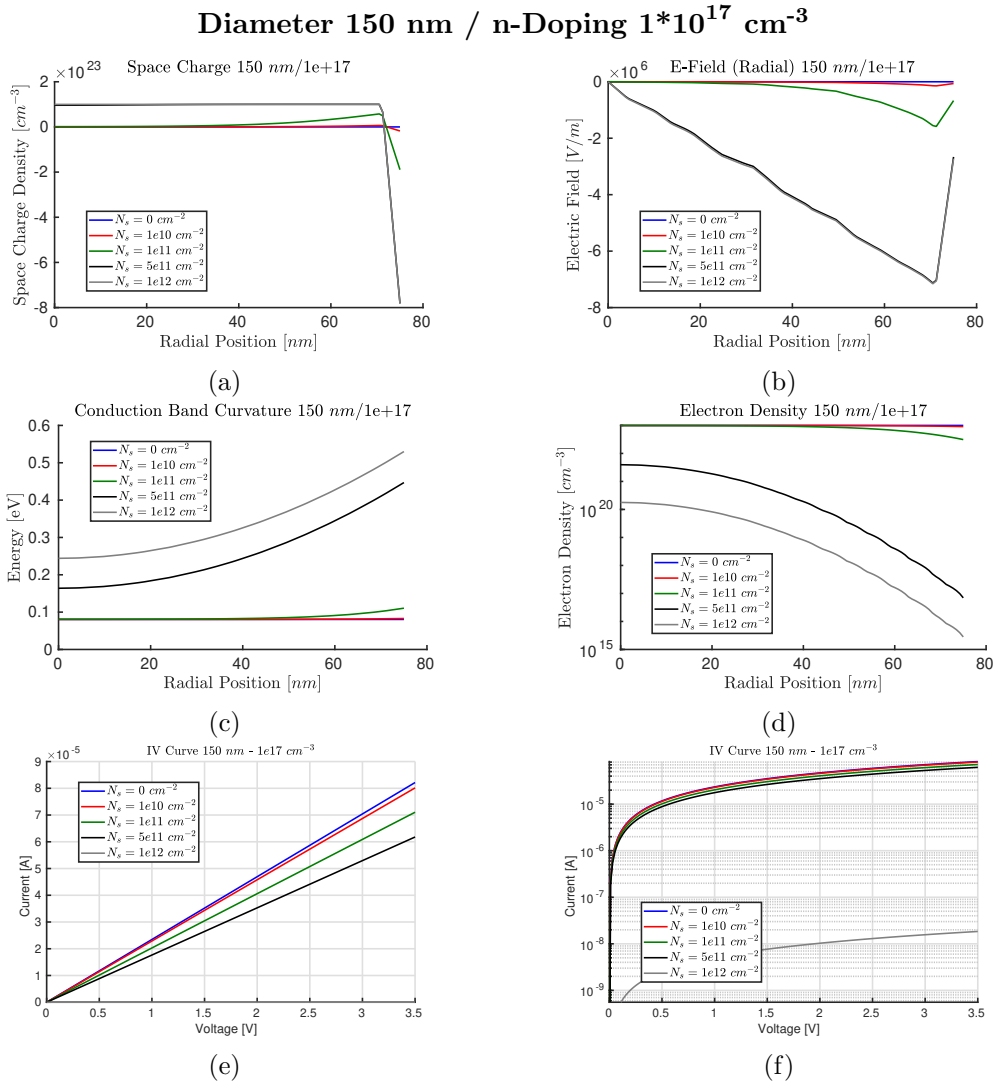

Figure 3.7: In lower doping case, a space charge region spans over the whole nanowire already for the scenario with second highest doping of $5 \cdot 10^{-3}$ cm^{-3}. The Fermi level drops as the electron density is reduced over the complete wire for the two scenarios with highest doping. However, the electron density is still more than an order of magnitude higher for the second highest surface charge compared to the highest.

The results above show that with the given assumption of the energy trap levels, the acceptor traps will become partially populated by electrons in the proximity, leading to a negatively charged surface. The electric field created by this negative layer of charges drives electrons from the vicinity of the surface towards the center of the nanowire as a potential barrier is built up for electrons to overcome in order to reach the surface.

Due to this, the outer part of the nanowire becomes depleted and hence insulating, leaving an inner conductive channel. Therefore the conductivity of the nanowire is no longer determined by the nanowire diameter, but by the effective diameter of this inner conducting channel. As desired for the sensor purpose, the conductivity decreases with increasing trap charge density as shown by the lowering currents in the IV-curve.

When reducing the nanowire diameter this effect becomes more prominent, as the ratio of surface to volume increases. The outer depletion area becomes wider as there are less volume carriers compared to the surface traps. Furthermore the ratio of the radii of the inner conductive and the outer insulating perimeters decreases. These effects result in a higher sensor responsitivity, which can be seen by comparing the IV-Characteristics.

In the case of higher doping[3], at 80 nm the nanowire shows a distinct deviation in the IV-Characteristics even for small amounts of surface charge indicating a very high responsitivity. On the other hand the space charge region spans over the whole nanowire for high surface charge concentrations. There is no unaffected inner channel anymore as the amount of volume carriers was not sufficiently high enough to shield the nanowire center. This result is in a good agreement with Calarco et al. [33, 39], where a critical diameter of around 80 nm was found for GaN nanowires with an intrinsic[4] doping of $6.25 * 10^{17}$ cm^{-3}. In devices like that, the IV-Characteristics is determined by space-charge-limited-currents[5] leading to a non-linear relation of voltage and current [33]. Therefore nanowires having a thickness around the critical diameter or below are not applicable for linear sensing devices.

In the lower doping scenario[6] the aforementioned surface effects become much more prominent compared to the case of the nanowire with higer doping. For the same amount of surface charge the space charge region penetrates deeper into the nanowire as there are less volume carriers within the inner bulk perimeter to satisfy the surface charge. Hence the conducting channel becomes smaller. Even for the resistor with the largest diameter of 150 nm a surface charge of $1 \cdot 10^{12}$ cm^{-2} causing a carrier depletion affecting the whole wire. The carrier density reduces even in the center and thus the Fermi level in the semicondutor drops. This shows that at a doping concentration of $1 \cdot 10^{17}$ cm^{-3} sensors having a diameter of 150 nm would already operate around a critical diameter. The critical diameter can be increased by a reduction of the doping or decreased by an increase of the doping concentration.

Therefore the doping concentration can be used as a key parameter for the sensing behavior. Moreover, as both measures, reducing the intrinsic n-type doping or reducing the nanowire diameter, show similar consequences they can be considered to compensate for each other if one of the options is practically not feasible.

[3] n = $6.25 * 10^{17}$ cm^{-3}
[4] or unintentional
[5] not considered here
[6] n = $1 * 10^{17}$ cm^{-3}

Current over Surface Charge

Figure 3.8 visualizes the sensitivity of the nanowire resistors regarding certain amounts of surface charge. Therefore the current is evaluated at a fixed voltage for a varying surface charge. Again the curves emphasize the major impact of doping concentrations and nanowire diameter on the sensor response identifying low doped and preferably thin nanowires as optimal choice in order to gain high sensor responses.

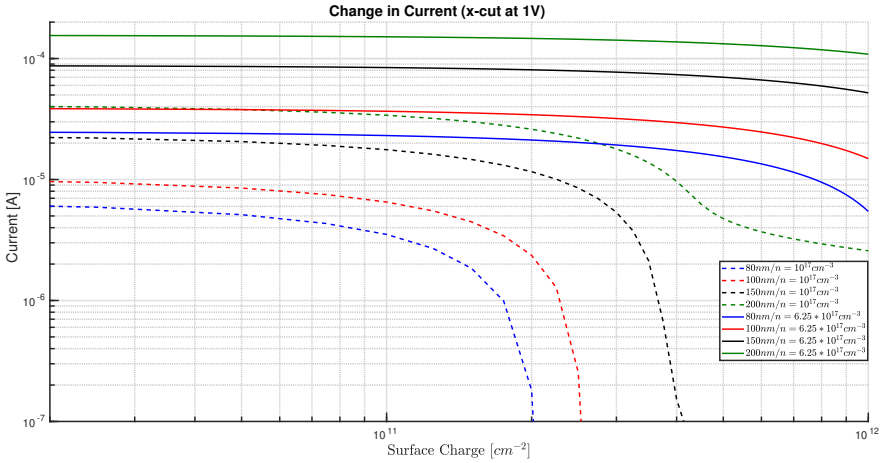

Figure 3.8: Current vs. surface charge

3.2.3 Results Including Oxide Layer at the Nanowire Surface

All following analyses will be carried out under the assumption of an oxidized GaN surface. Hence, in all sensor structure models an additional oxide layer with a thickness of 10 nm has been placed on the surface, separating the nanowire from the air environment. As for those devices a charge transfer between nanowire and gas molecules is prevented by the oxide layer, here the sensing principle will be based on electrostatic effects. Therefore it is assumed that the functional molecule groups are seated on top of the oxide layer and a dipole moment is created by docking gas molecules. An existing dipole moment on the surface reduces or increases the carrier concentrations in the outer region of the nanowire by electrostatic attraction or repulsion. This variation in the carrier density will directly affect the nanowire conductance and hence the current flow. In order to simulate this effect a fixed density of positive surface charges seated at the oxide-gas interface is added to the model.

Resistors

The resistors analyzed here had a length of 1 μm and variation of diameters between 150 nm and 200 nm and doping concentrations between $N = 10^{15}$ cm^{-3} and $N = 10^{17}$ cm^{-3}.

The figures 3.9 - 3.12 show how due the positive surface charge the conduction band shows a down bending along the radial axis. This effect increases for an increasing surface charge, hence the electron density becomes elevated in the outer part of the nanowire. Therefore the conductivity increases resulting in a higher current flow. Without surface charge no band bending occurs. Also for devices with oxide layer this effect becomes more prominent for smaller diameters as well as for lower doping concentrations, since the relative increase is higher in those set-ups. The influence of diameter and doping shall be illustrated by the comparisons shown in fig. 3.14. Comparable to the results presented in the previous section simulations with negative surface charge reveal the opposite behavior, meaning an up-bending of the conduction band close to the surface (carrier depletion).

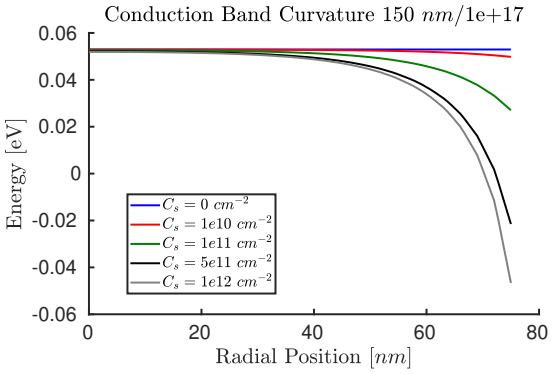

Figure 3.9: Conduction Band Resistor 150 nm and n-doping of $1 \cdot 1017$

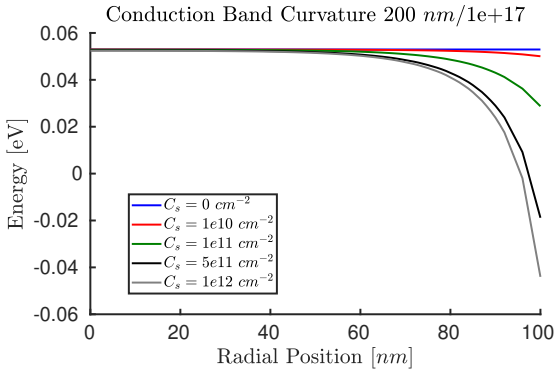

Figure 3.10: Conduction Band Resistor 200 nm and n-doping of $1 \cdot 1017$

Fig. 3.13 shows exemplarily how the presence of a surface charge affects the IV-characteristics.

In fig. 3.14 the sensor response of resistor sensors for different diameters and doping concentrations is shown in a double logarithmic scale. Curves with a slope of 1 indicate a linear response. Furthermore, it can be seen that the response increases for lower doping and smaller diameters. Results also show that the linearity of the sensor suffers for very low doping concentrations.

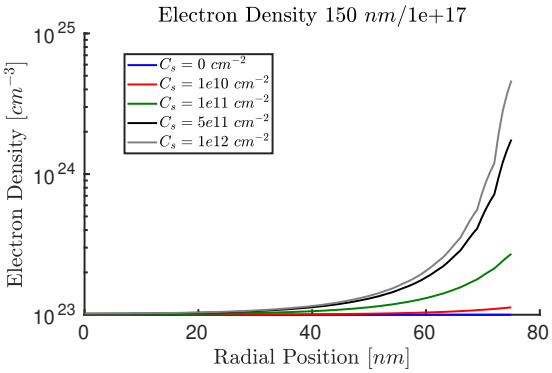

Figure 3.11: Electron Density Resistor 150 nm and n-doping of $1 \cdot 1017$

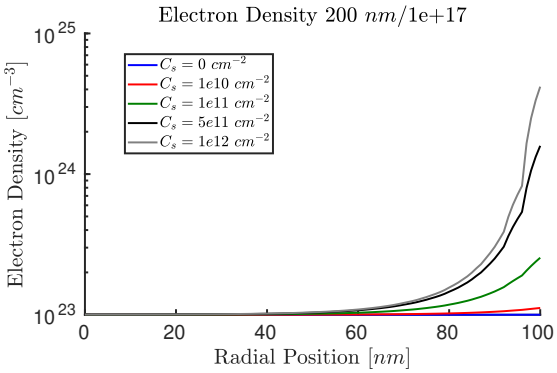

Figure 3.12: Electron Density Resistor 200 nm and n-doping of $1 \cdot 1017$

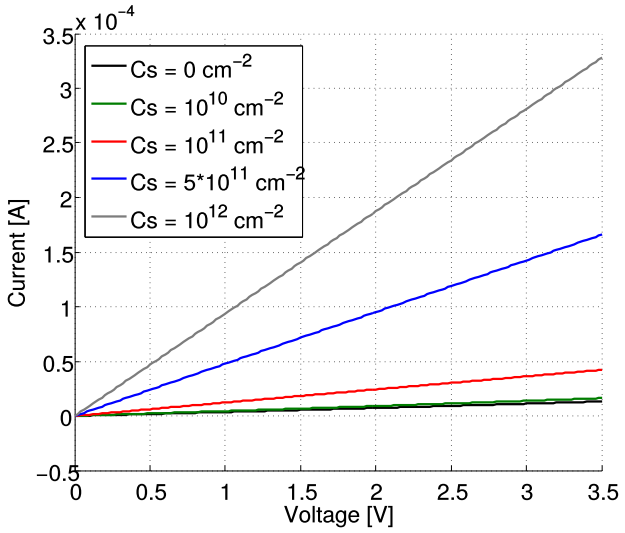

Figure 3.13: IV-Curve behaviour for increasing surface charge

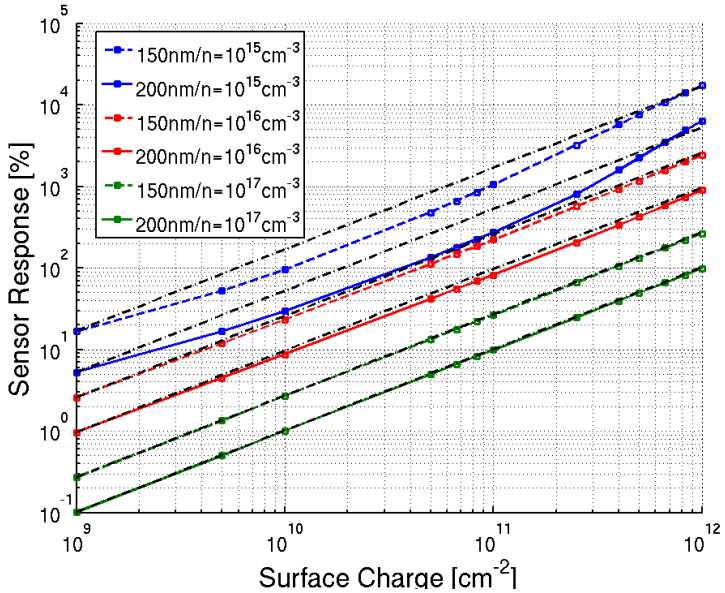

Figure 3.14: Resistor Sensor Response: Comparison of different diameters and doping concentrations. Dashed lines represent linear sensor response for comparison.

3.2.4 PN-Diodes

The analyzed diodes have a diameter of 150 nm, they consist of a 1 μm long n-region and a 300 nm long p-region. A 10 nm thick oxide layer was added to the nanowire surface, therefore the surface charge is placed at the oxide/air-interface. Constant doping profiles have been chosen with n = 10^{17} cm^{-3} and p = $5 \cdot 10^{16}$cm^{-3}.

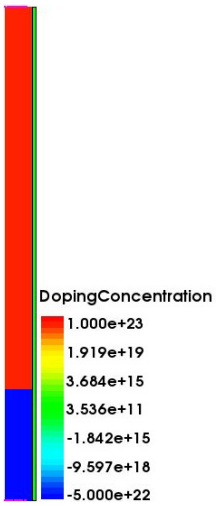

Figure 3.15: 2D cut of the cylindrical nanowire PN-Diode including a 10 nm thick oxide layer. The doping concentrations shown in this figure are not given in cm^{-3} but in m^{-3}.

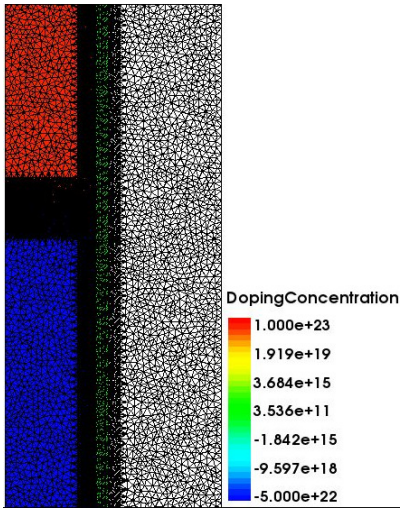

Figure 3.16: Nanowire with illustrated meshing. The mesh is refined at a GaN/Oxide-interface region and over the PN-junction (with an additional 25 nm long interface in each region.)

The results shown in this section demonstrate that a positive surface charge shifts the threshold voltage to lower values while the IV-behavior stays unaffected in reverse bias. Unlike the resistor sensors, here, the generation of a conducting channel near the surface is prevented by the potential difference at the contacts. While the surface charge leads to a carrier inversion across the p-region, the pn-junction shifts close to the p-contact. Here, as expressed by a steep slope in the band structure (dashed line graphs in fig. 3.17), a strong electric field prevents the electrons to further recombine with the holes.

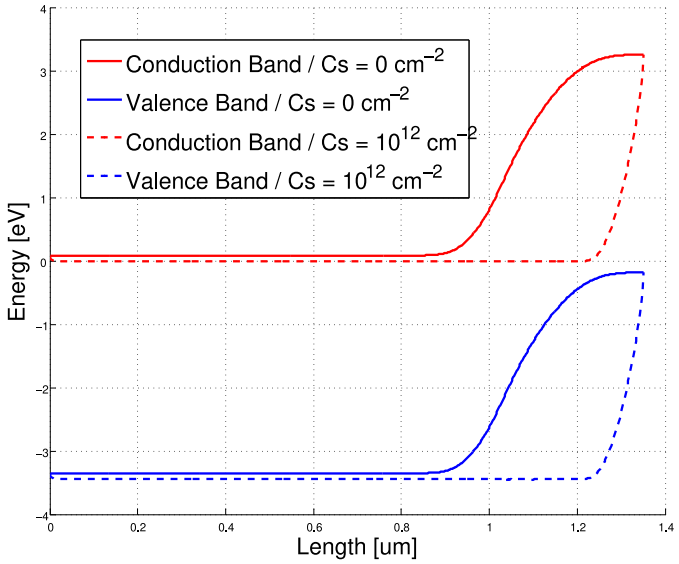

Figure 3.17: Comparison of the band structure along the y-axis with and without surface charges

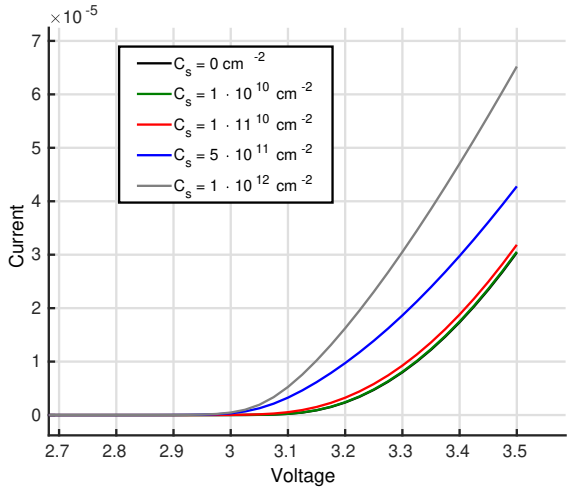

Figure 3.18: Influence of surface charge for diode IV-curve

The sensor response for the investigated diode is depicted in fig 3.19. The results show no considerable improvement in the maximum response compared to the resistor sensors while loosing their linear characteristics. Combined with the higher degree of complexity, diode sensors appear as a less desirable sensor design.

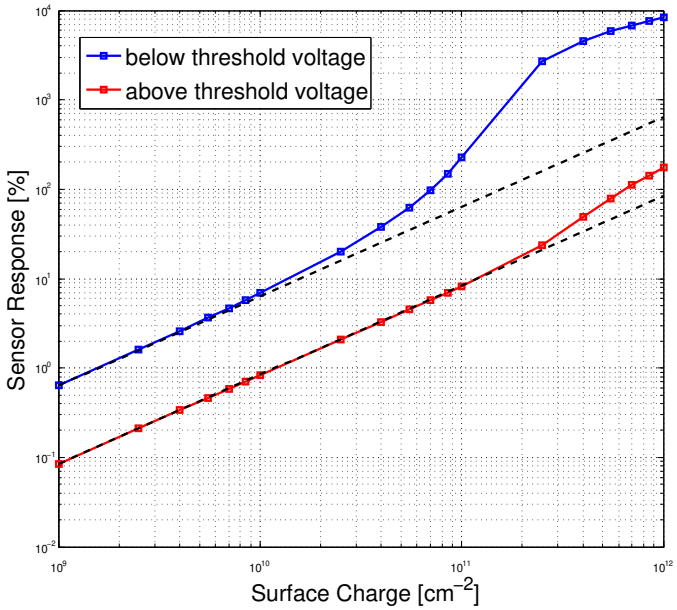

Figure 3.19: Diode Sensor Response: Blue and red curves show the response of the diode when operated slightly below and above threshold voltage. The two dashed black curves serve as comparison with a linear sensor.

3.2.5 N^+N^-- Junctions

Junctions formed by highly n-doped (n^+) and low n-doped (n^-) regions build up a much smaller potential barrier compared to pn-diodes. As a consequence, a linearity compared to those of resistors can be expected for voltages exceeding the rather small threshold voltage. Two different sensor set-ups are compared here. Here, each differently doped n-region has a length of 500 nm. The doping concentration of the highly doped region is fixed at 10^{18} cm^{-3} while for the lower doped region concentrations of 10^{16} cm^{-3} and $5 \cdot 10^{16}$ cm^{-3} are distinguished.
Figures 3.20 and 3.21 illustrate the rather small electron barrier formed by the junction already leading to weakly non-linear IV-Curves (fig. 3.22) when no surface charge is applied. Considering a surface charge density high enough to remove the energy barrier the conduction band reveals the behavior of two series resistors, as can be seen by the two straight lines. The different slopes represent the non-identical resistances.

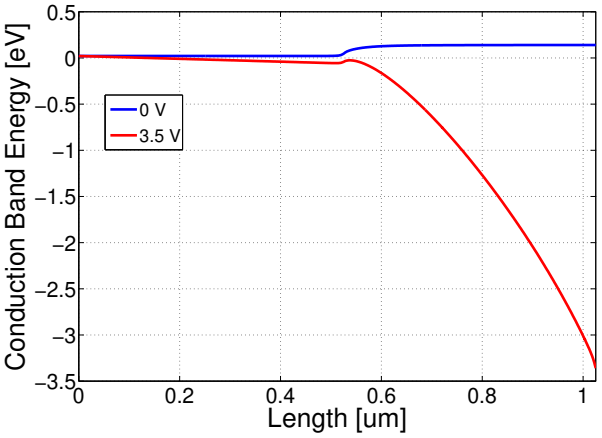

Figure 3.20: Conduction band: No surface charge

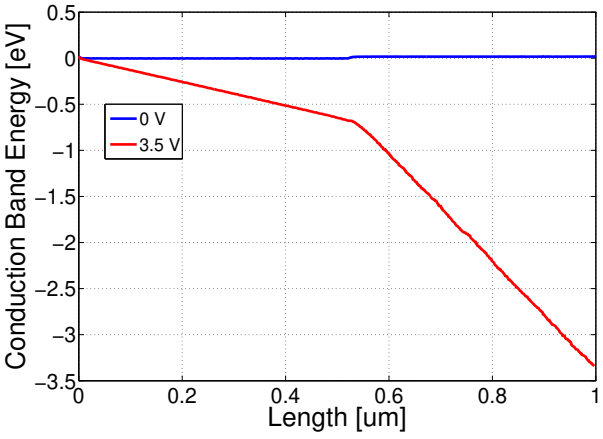

Figure 3.21: Conduction band: Including surface charge

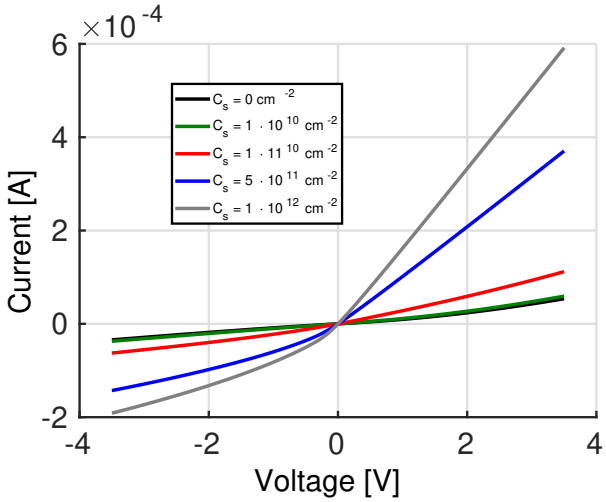

Figure 3.22: IV Curve of a sensor based on a n^+n^--junction where $n^+ = 10^{18}$ and $n^- = 10^{16}$.

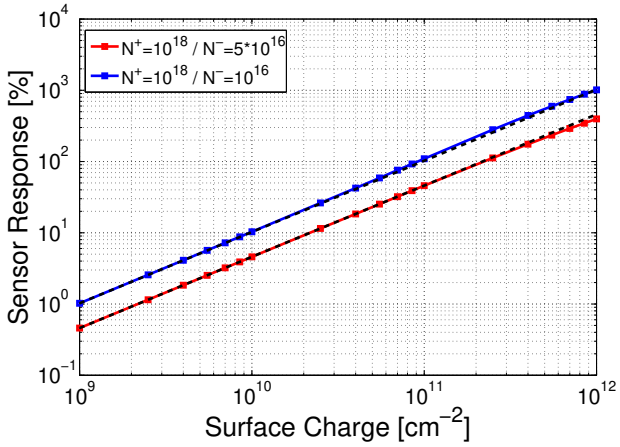

Figure 3.23: Sensor response of the n^+n^--junction.

Fig. 3.22 shows the corresponding IV-Curve. In forward bias the IV-curves of the device are weakly non-linear for low surface charge densities. With increasing surface charge the non-linearity vanishes and the graphs look like resistor curves. In reverse bias the contact potential hinders electrons to populate the vicinity of the contact in the low doped region. In this way the resistance of that region becomes voltage dependent, resulting in a damped slope for increasing negative voltages. In fig. 3.23 the sensor response is depicted. A higher response was shown by the set-up with the higher doping contrast. Both curves show a good linearity while lying in the range of the resistor responses.

3.2.6 Transistors

In this section the simulation results for two transistor set-ups are compared to each other. Both structures had a diameter of 200 nm and 1 μm long n-regions. The two different p-regions had length of 300 nm and 700 nm respectively. Again, a 10 nm thick oxide layer was added to the surface. In contrast to diodes, npn-transistors with equally doped n-regions don't have a potential difference at the contacts. The doping concentrations were chosen in both cases as $n = 10^{17}\text{cm}^{-3}$ and $p = 5 \cdot 10^{16}\text{cm}^{-3}$.

Results using a 300 nm long p-Region

Since there is no gate electrode the transistor is generally high-ohmic without the presence of a surface charge. Nonetheless, the band structure depicted in fig. 3.24 shows that due to the short p-region and the doping contrast, the potential barrier created by the pn-junctions is relatively low which is reducing the blocking behavior of the device. As there is no plateau build up for the potential, the depletion region spans over the whole p-semiconductor. Due to this, for a certain "threshold voltage" the potential barrier is low enough for the electrons to overcome, meaning the transistor is loosing its insulating character. For a sufficiently large surface charge a carrier inversion is reached in the p-region, opening a conducting channel near the surface and hence increasing the conductance. In this case the device is behaving like a resistor, or more precisely like three resistors in series connection. This effect is illustrated by the red line in fig.3.25 and the transistors IV-curve shown in fig. 3.27.

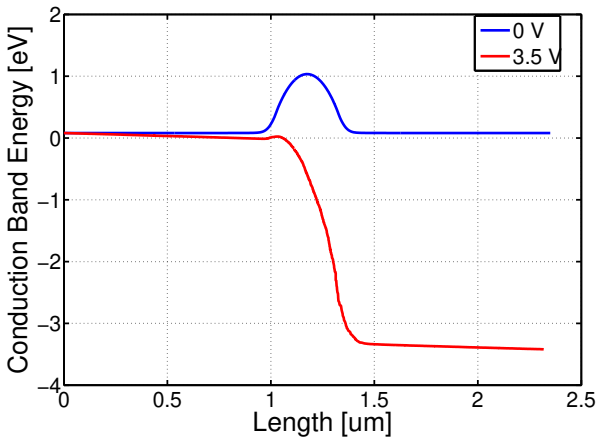

Figure 3.24: Conduction band near the surface along the y-axis: No surface charge

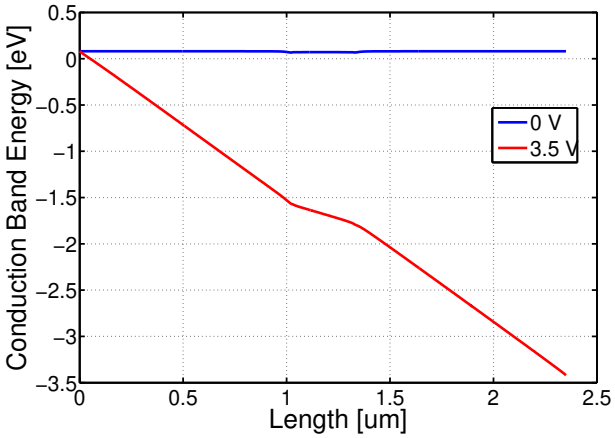

Figure 3.25: Conduction band near the surface along the y-axis: Including surface charge

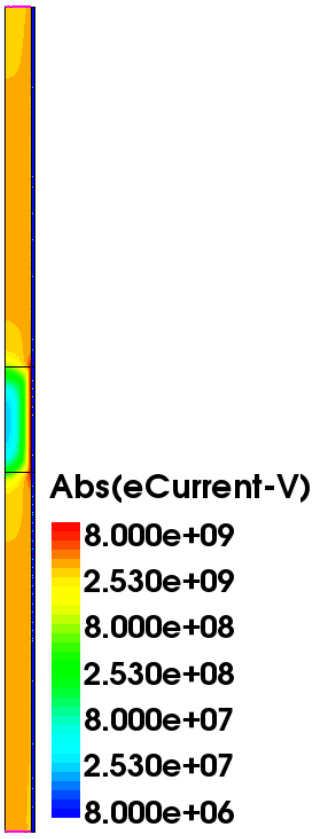

Figure 3.26: The current density in the device illustrates how the surface charge creates a conductive channel at the surface, acting like a parallel resistor to the junctions bridging the p-region

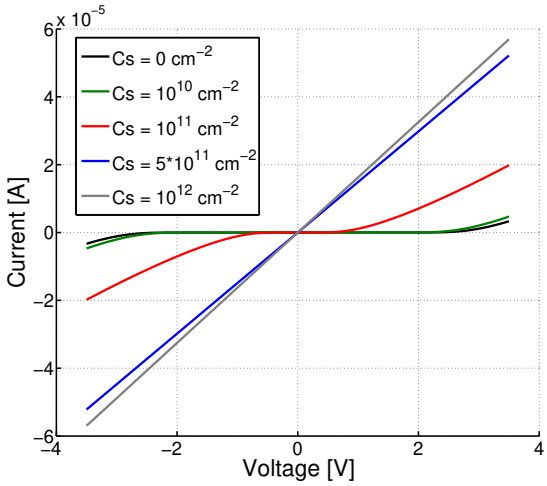

Figure 3.27: Transistor IV Curve: An increasing surface charge reduces the potential barrier and hence the threshold voltage. When completely flattened the resulting parallel resistor dominates the IV-behavior.

Like for the diode, the sensor response of the transistor (fig.3.28) was again evaluated slightly above and below the threshold voltage of the uncharged case. Large responses can be reached when driven below threshold voltage but the response would be highly non-linear. Driven above the threshold voltage the response is approximately linear over a large range of surface charge density but also lower compared to the response gained below threshold. This results from a much higher initial current serving as reference.

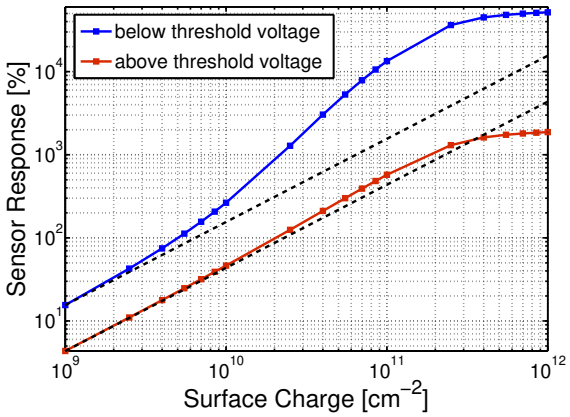

Figure 3.28: Transistor Sensor Response: Blue and red curves show the response of the transistor when operated slightly below and above threshold voltage. The two dashed black curves serve as comparison with a linear sensor.

Results using a 700 nm long p-Region

With the incorporation of a 700 nm long p-region, a more distinct potential barrier is built up. As can be seen by the red curve fig. 3.29 this barrier does not vanish for the maximum voltage of 3.5 V, leaving the device insulating. In fig. 3.30 the influence of a surface charge with a density of $1 \cdot 10^{12}$ cm^{-2} on the band structure is illustrated. Again a carrier inversion could be reached, making the device conductive. Due to the much better blocking behaviour in the case charge free surface, higher responses compared to the transistor with shorter p-region can be yield for surfaces charges sufficiently large to eliminate the potential barrier. However, since the device is insulating in the charge free case, the linearity is weak and low gas concentrations might not be detectable as the resulting surface charge could be insufficient for an effective reduction of the potential barrier.

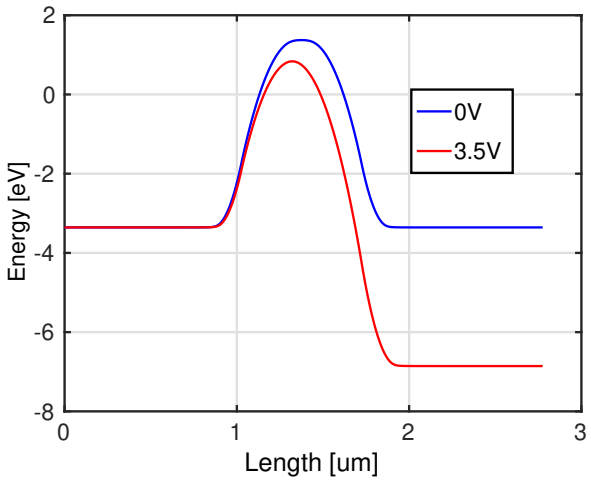

Figure 3.29: Conduction band along the y-axis: No surface charge

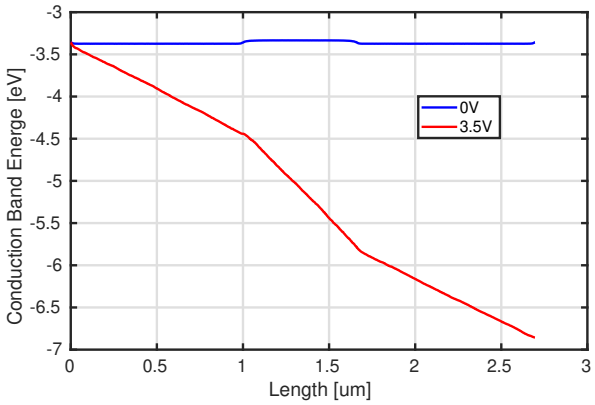

Figure 3.30: Conduction band along the y-axis: Including surface charge

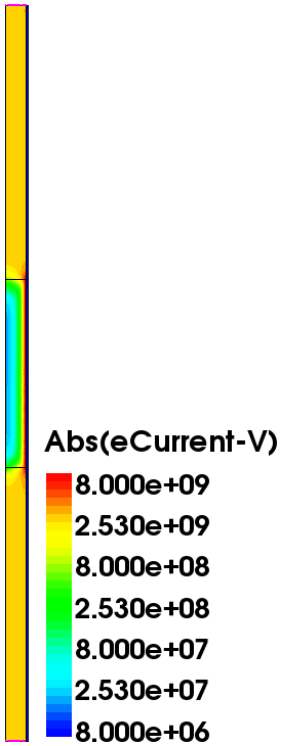

Figure 3.31: The current density in the device illustrates how the surface charge creates a conductive channel at the surface, acting like a parallel resistor to the junctions bridging the p-region

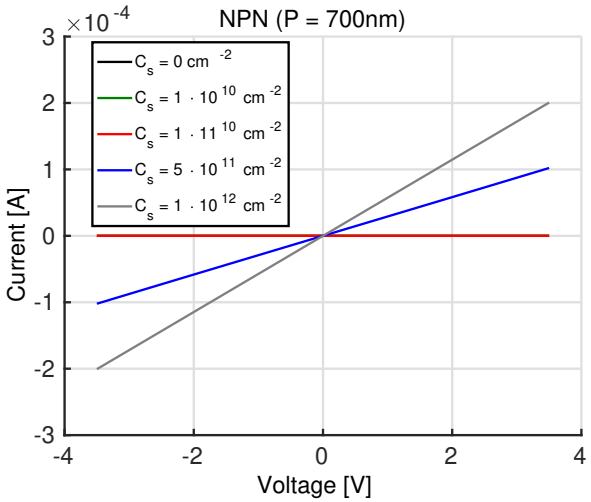

Figure 3.32: Transistor IV Curve: An increasing surface charge reduces the potential barrier and hence the threshold voltage. When completely flattened the resulting parallel resistor dominates the IV-behavior.

3.2.7 Comparison

In this section all analyzed sensor structures shall be compared according to their sensor responses and linearity deviations. Results are illustrated in figures 3.33 and 3.34. Sensor responses are evaluated at a surface charge density of $C_s = 10^{11}$ cm^{-2} where most sensors still behaved linear to a good degree. Even though the resistor sensors with the lowest doping concentrations showed the highest sensor responses they were not considered here because of their weak linearity. Due to that, two resistor sensors are chosen for the comparison, both having a diameter of 150 nm with doping concentrations of n = 10^{17} cm^{-3} and n = 10^{16} cm^{-3}. For the n$^+$n$^-$-junction the higher contrast design is compared since it showed the better overall performance. Even though the transistor with 700 nm p-region showed the highest responsivity of all devices (3043 %), the transistor with the shorter p-region is chosen for the comparison below, as it offers a better linearity and the possibility to detect lower gas concentrations.
Fig. 3.33 and 3.34 illustrate the performance of the different sensor devices. It can be seen that for resistors the lower doped sensor offers a significantly higher sensor response (218.3 % compared to 26.5 % in case of higher doping) while the resistor with the highest doping concentrations offers the best linearity (2.6 %) of all analyzed sensors. The diode driven above threshold showed the lowest sensor response with a value of 8.1 % while already having a maximum deviation of 105.8 % from a linear sensor device. This large deviation only occurs for high surface charge densities, for a surface charge up to 10^{11} cm^{-2} the maximum deviation stays below 3.5 %. Driven below the threshold voltage the diode's response reaches comparable values to that of lowly doped resistor

but suffering from a very high linearity deviation (1121.5 %). The n^+n^--junction lies in between the results of the two resistors, having a sensor response of 109.4 % and linearity deviation of 9.6 %. For the transistor driven above threshold a high response of 574.4 % can be observed, at a linearity deviation of 57 %. Similarly like the diode the linearity deviation rises for high surface charge densities while it stays below 24 % for densities up to 10^{11} cm^{-2}. The highest response (2221.1 %) was gained for the transistor driven slightly below threshold voltage but this sensor would also be highly non-linear.

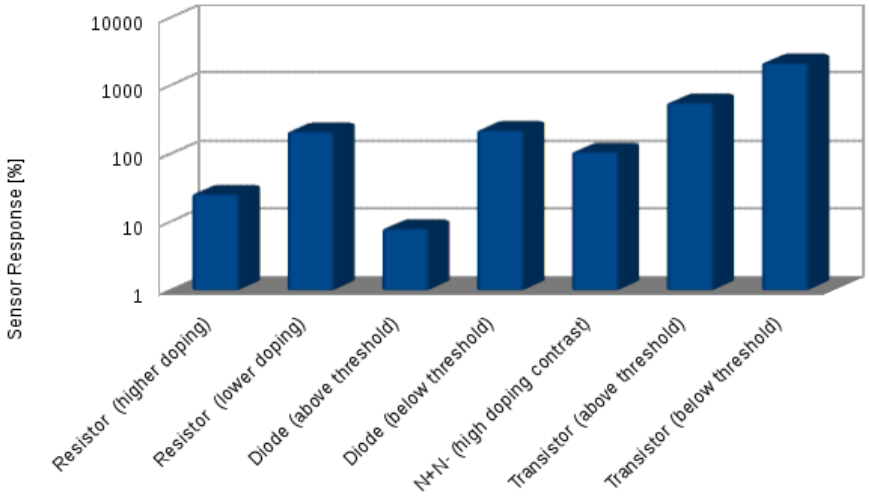

Figure 3.33: Comparison of sensor responses for different exemplarily chosen sensor structures.

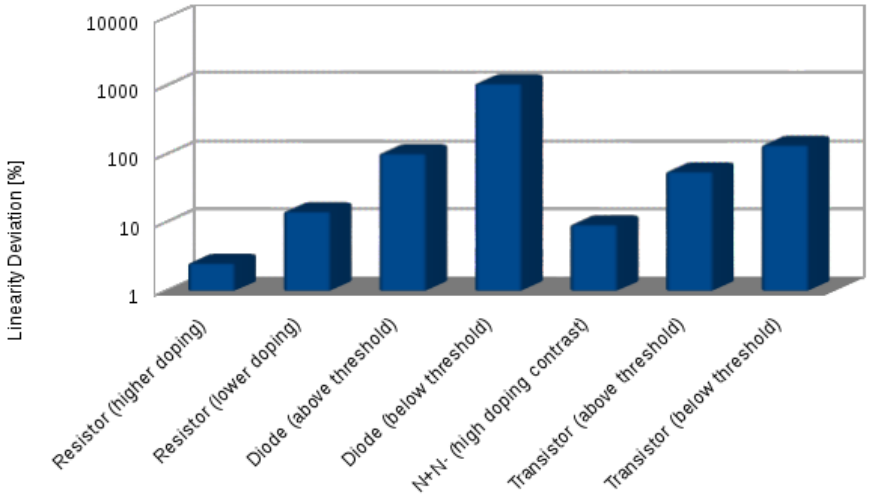

Figure 3.34: Deviation from linear sensor response

4 Conclusions and Outlook

4.1 Conclusions

The usage of GaN nanowires enables promising sensor structures for the detection of specific gases. The results presented in the previous chapter show that an approach of combining drift-diffusion-simulations with a surface model introduced in chapter 1.2.2 can be applied in order to understand the behavior of nanowire gas sensors. Moreover, the results show that GaN nanowires with oxidized surfaces are in general still applicable for gas sensing purposes. If in this case high selectivities compared to the approach of Hoffmann et al. [2] can be reached by a suitable surface functionalization remains unclear and has to be shown by experimental results. An intrinsic pinning of the nanowire surface would on the one hand superimpose the sensing mechanism and on the other hand introduce constraints regarding the device dimensions due to the critical diameter. In order to determine the sensor behavior, this bias in the energetic position of the surface needs to be incorporated into the device simulations. Therefore a precise knowledge of the nanowire surface conditions regarding the state of oxidation and Fermi level pinning has shown to be crucial for a reliable predictability of sensor responses. Even though the original project goal of having a working sample of a nanowire gas sensor with sensor responses predictable by device simulations[1] was not achieved, the results gained are still capable of giving insight to general phenomena of nanowire electronics and to serve as a guidance for the design of nanowire gas sensors. Critical diameters including their doping dependence predicted by the simulation results are in a good agreement with the results observed and theoretically explained by Calarco et al. [33, 39]. A sufficiently large trap charge density will determine the Fermi level at the surface causing a bending of the bands and hence an accumulation or depletion of carriers in the volume close to the surface. This variation in the carrier density will directly affect the nanowire conductance and hence the current flow. An increase of the surface-to-volume ratio goes along with an increase of the surface trap densities compared to the volume carrier concentrations. Therefore there are less carriers within the adjacent volume to recombine with the surface traps and the depletion region grows deeper towards the nanowire center, reducing the conductive part of the nanowire. This is comparable to a diodes PN-junction width, where the depletion region expands deeper into the bulk material when lower doping concentration are incorporated for one of the carrier species.
In order to facilitate the realization of low doping concentrations, simulations regarding a GaN/AlGaN/GaN core-shell sensor had been carried out. The approach proved to be impractical for all analyzed material compositions, as leakage currents over the AlGaN-barrier could not be prevented, resulting in non-linear IV-characteristics.

[1] realized by simulation results of atomistic Density Functional Theory for energy trap levels and densities subsequently incorporated in a macroscopic device simulation

For the comparison of the sensing performance of the individual sensor structures, devices with incorporated oxide layer on the GaN surface have been chosen, as devices without oxide layer could currently not be realized according to our partners at TU Braunschweig. The results can be summarized as follows.
Pure nanowire resistors showed higher responses for lower diameters and doping concentrations, but the linearity suffers when the combined set-up of both parameters is chosen too low. Diodes showed low responses when driven above threshold voltage and a highly non-linear behavior when driven below, making the concept less preferable compared to resistors, especially considering the higher device complexity. Reverse biasing proofed not to be practical for sensing applications for all analyzed diode sensor structures. On the other hand a diode driven slightly below threshold voltage could work in a similar way. With this approach a low power sensor with considerably high sensor response could in general be realized. However those kind of sensors would not be useful in applications requiring a high sensor linearity, but could be used for low power pure "on/off-sensors". Sensors based on n^+n^--junctions showed comparable results to those of resistors. Transistors showed very high responsivities combined with a comparatively moderate non-linearity for low surface charge concentrations. By varying the size of the p-region or the doping concentrations and driven at a suitable operation voltage they can also be tailored to act as "on/off"-sensors, triggering at a specific amount of surface charge (gas concentration). But they are also coming with a higher degree of complexity compared to pristine nanowire resistors.

4.2 Outlook

Till the end of this thesis work nor the realization of physical GaN sensor samples could be realized, neither results from the atomistic DFT-calculations for the energy levels of the HOMO/LUMO states in the system of semiconductor and functional group could be gained. Due to the lack of this data, trap energy levels and densities had to be guessed and do not represent a realistic gas-SAM-nanowire interaction. In general, the results presented above still have to be validated by practical experiments. If necessary, corrections have to be made for the simulation set up with a subsequent reevaluation of optimal sensor designs in an iterative process of simulation and experimental validation.

5 Appendix

5.1 New parameters for doping and field dependent carrier mobilities

```
# Field dependent mobility model
# Canali model
field_dependence_vsat_C_e    = 1.9064e5  #m*s^-1 ioffe.ru
field_dependence_vsat_C_h    = 1.9064e5  #m*s^-1 ioffe.ru
field_dependence_beta_C_e    = 0.5       #sdevice manual
field_dependence_beta_C_h    = 1.2       #sdevice manual

# Transfered Electron model
field_dependence_vsat_TE_e  = 1.9064e5 #m*s^-1 ioffe.ru
field_dependence_mu1_TE_e   = 0        #m^2*V^-1*s^1 sdevice manual
field_dependence_F0_TE_e    = 22089360 #V*m^-1
field_dependence_F1_TE_e    = 22089360 #V*m^-1
field_dependence_alpha_TE_e = 1        #sdevice manual
field_dependence_beta_TE_e  = 7.2044   #sdevice manual
field_dependence_gamma_TE_e = 6.1973   #sdevice manual

field_dependence_vsat_TE_h  = 1.9064e5  #m*s^-1 ioffe.ru
field_dependence_mu1_TE_h   = 0.0       #m^2*V^-1*s^1  sdevice manual
field_dependence_F0_TE_h    = 4e5       #V*m^-1
field_dependence_F1_TE_h    = 4e5       #V*m^-1
field_dependence_alpha_TE_h = 0.0       #sdevice manual
field_dependence_beta_TE_h  = 4.0       #sdevice manual
field_dependence_gamma_TE_h = 0.0       #sdevice manual

# MNatsakanov parameters for Arora Model 300K (all values taken from [4])
mobility_arora_Amin_e      = 0.0055 # [m2/Vs]
mobility_arora_alphamin_e  = 1
mobility_arora_Ad_e        = 0.0945 # [m2/Vs]
mobility_arora_alphad_e    = 1
mobility_arora_AN_e        = 2e23   # [m-3]
mobility_arora_alphaN_e    = 1
mobility_arora_Aa_e        = 1
mobility_arora_alphaa_e    = 1
mobility_arora_Amin_h      = 0.0003 # [m2/Vs]
mobility_arora_alphamin_h  = 0
mobility_arora_Ad_h        = 0.0167 # [m2/Vs]
mobility_arora_alphad_h    = 1
mobility_arora_AN_h        = 3e23   # [m-3]
mobility_arora_alphaN_h    = 1
mobility_arora_Aa_h        = 2
mobility_arora_alphaa_h    = 1
```

5.2 Examplary Sensor Geometries

5.2.1 Resistor

```
(define geoname  "n@node@_geo")
(define gridname "n@node@_mesh")

(define wire_mat "GaN")
(define ambient "Gas")

(define ref_wire 0.0001)
(define ref_air 0.01)

(define contact_hgth 0.05)

(define wire_lgth 1.0)
(define wire_top (+ contact_hgth wire_lgth))
(define top_cont_hgth (+ wire_top contact_hgth))
(define btm_cont_hgth (+ 0 contact_hgth))
(define wire_rad @radius@)
(define air_top top_cont_hgth)
(define air_btm 0)
(define air_wdth (+ wire_rad 0.2))
(define interface_wdth 0.01)

(define ndop @n_doping@)
(define ndoping (* ndop 1e6))

(sde:delay-graphics-update #t)

(sdegeo:create-rectangle (position 0 air_btm 0)
(position air_wdth top_cont_hgth 0) ambient "Gas_Surrounding")
(sdegeo:create-rectangle (position wire_rad air_btm 0)
(position (+ wire_rad interface_wdth) top_cont_hgth 0) ambient "Gas_Interface")
(sdegeo:create-rectangle (position 0 0 0)
(position wire_rad btm_cont_hgth 0) wire_mat "GaN_btm_cont")
(sdegeo:create-rectangle (position 0 btm_cont_hgth 0)
(position wire_rad wire_top 0) wire_mat "GaN_wire1")
(sdegeo:create-rectangle (position 0 wire_top 0)
(position wire_rad top_cont_hgth 0) wire_mat "GaN_top_cont")

(sdegeo:define-contact-set "ncontact" 0.05 (color:rgb 1 0 0) "//")
(sdegeo:define-contact-set "pcontact" 0.05 (color:rgb 0 0 1) "++")
(sdegeo:define-2d-contact (find-edge-id (position (/ wire_rad 2) 0 0)) "ncontact")
(sdegeo:define-2d-contact (find-edge-id  (position (/ wire_rad 2) top_cont_hgth 0))
"pcontact")

(sdedelaunizer:set-parameters "type" "boxmethod")
(sdedr:set-meshing-engine "noffset")
(sdenoffset:create-global "usebox" 1 "maxnumpoints" 800000 "recoverholes" 1)
(sdedelaunizer:set-parameters "maxPoints" 800000)
```

```
(sdedr:define-constant-profile "ndoping_def" "ArsenicConcentration" ndoping)

(sdedr:define-constant-profile-region "nregion_plac" "ndoping_def" "GaN_btm_cont")
(sdedr:define-constant-profile-region "nregion_plac2" "ndoping_def" "GaN_wire1")
(sdedr:define-constant-profile-region "nregion_plac3" "ndoping_def" "GaN_top_cont")

;(sdedr:define-refinement-size     "Gas_size" (* ref_air 2) (* ref_air 2) (* ref_air 2)
                                             (/ ref_air 10) (/ ref_air 10) (/ ref_air 10))
;(sdedr:define-refinement-region   "Gas_ref"  "Gas_size" "Gas_Surrounding")
;(sdedr:define-refinement-size     "Wire_size" (* ref_wire 2) (* ref_wire 2) (* ref_wire 2)
                                             (/ ref_wire 10) (/ ref_wire 10) (/ ref_wire 10))
;(sdedr:define-refinement-region   "Wire_ref" "Wire_size" "GaN_wire1")
;(sdedr:define-refinement-region   "Wire_ref2" "Sub_Size" "GaN_conetip")
;(sdenoffset:create-noffset-interface "region" "Gas_Surrounding" "GaN_wire1"
                                               "hlocal" ref_wire "factor" 1.1)
;(sdedr:define-refinement-size     "Oxide_size" (* ref_Ox 2) (* ref_Ox 2) (* ref_Ox 2)
                                               (/ ref_Ox 10) (/ ref_Ox 10) (/ ref_Ox 10))
;(sdedr:define-refinement-region   "Oxide_ref"     "Oxide_size" "Oxide_Substrate")

(define gas_fact 0.05)
(define wire_fact 0.05)
(define interface_fact wire_fact)

(sdenoffset:create-noffset-block "region" "GaN_btm_cont"     "maxlevel" 50
"maxedgelength"  (* wire_fact  0.15))
(sdenoffset:create-noffset-block "region" "GaN_wire1"        "maxlevel" 50
"maxedgelength"          (* wire_fact  0.15))
(sdenoffset:create-noffset-block "region" "GaN_top_cont"     "maxlevel" 50
"maxedgelength"  (* wire_fact  0.15))
(sdenoffset:create-noffset-block "region" "Gas_Surrounding" "maxlevel" 50  "maxedgelength"
(* gas_fact  0.15))
(sdenoffset:create-noffset-block "region" "Gas_Interface"    "maxlevel" 50
"maxedgelength" (* interface_fact  0.15))

(sde:delay-graphics-update #f)
(sde:save-model geoname)
;(sde:build-mesh "noffset" (string-append " -F dfise -x -o" geoname) geoname)
(sde:build-mesh "mesh" "-discontinuousData -noffset -d -F dfise" geoname)
;(ise:build-mesh "mesh" "-discontinuousData -P " geoname)
;(sde:build-mesh "snmesh" "-rect -w 4 -u -d -dfise" geoname)
```

5.2.2 PN-Diode

```
(define geoname  "n@node@_geo")
(define gridname "n@node@_mesh")

(define wire_mat "GaN")
(define ambient "Gas")

(define ref_wire 0.0001)
(define ref_air 0.01)

(define contact_hgth 0.025)

(define wire_lgth_n 1)
(define wire_lgth_p 0.3)
(define nwire_top (+ contact_hgth wire_lgth_n))
(define pwire_top (+ nwire_top wire_lgth_p))
(define top_cont_hgth (+ pwire_top  contact_hgth))
(define btm_cont_hgth (+ 0 contact_hgth))
(define wire_rad @radius@)
(define air_top top_cont_hgth)
(define air_wdth (* wire_rad 1.25))
(define interface_wdth 0.01)
(define oxide_thickness 0.01)
(define ox_GaN_interf_width 0.005)
(define junction_dpth 0.025)
(define wire_interf_penetration_factor 3)

(define ndop @n_doping@)
(define pdop @p_doping@)
(define ndoping (* ndop 1e6))
(define pdoping (* pdop 1e6))

(sde:delay-graphics-update #t)

(sdegeo:create-rectangle (position (+ wire_rad oxide_thickness) 0 0)
(position (+ (+ wire_rad oxide_thickness) air_wdth) pwire_top  0)
ambient "Gas_Surrounding")
(sdegeo:create-rectangle (position (+ wire_rad oxide_thickness) 0 0)
(position (+ (+ wire_rad oxide_thickness) interface_wdth) pwire_top 0)
ambient "Gas_Interface")
(sdegeo:create-rectangle (position wire_rad 0 0)
(position (+ wire_rad oxide_thickness) pwire_top 0)
"Oxide" "Oxide_layer")


(sdegeo:create-rectangle (position 0 0 0)
(position (- wire_rad (* ox_GaN_interf_width wire_interf_penetration_factor))
nwire_top 0) wire_mat "GaN_wire_n")
(sdegeo:create-rectangle (position
(- wire_rad (* ox_GaN_interf_width wire_interf_penetration_factor)) 0 0)
(position wire_rad nwire_top 0)
wire_mat "GaN_Ox_int_n")
(sdegeo:create-rectangle (position 0 nwire_top 0)
```

```
(position (- wire_rad (* ox_GaN_interf_width wire_interf_penetration_factor))
pwire_top 0) wire_mat "GaN_wire_p")
(sdegeo:create-rectangle (position
(- wire_rad (* ox_GaN_interf_width wire_interf_penetration_factor))
nwire_top 0) (position wire_rad pwire_top 0)
wire_mat "GaN_Ox_int_p")

(sdegeo:create-rectangle (position 0 (- nwire_top junction_dpth) 0)
(position (- wire_rad ox_GaN_interf_width) nwire_top 0)  wire_mat "Junction_n")
(sdegeo:create-rectangle (position 0 nwire_top 0)
(position (- wire_rad ox_GaN_interf_width) (+ nwire_top junction_dpth) 0)
wire_mat "Junction_p")

(sdegeo:define-contact-set "ncontact" 0.05 (color:rgb 1 0 0) "//")
(sdegeo:define-contact-set "pcontact" 0.05 (color:rgb 0 0 1) "++")
(sdegeo:define-2d-contact (find-edge-id (position (/ wire_rad 2) 0 0)) "ncontact")
(sdegeo:define-2d-contact (find-edge-id  (position (/ wire_rad 2) pwire_top 0)) "pcontact")

(sdedelaunizer:set-parameters "type" "boxmethod")
(sdedr:set-meshing-engine "noffset")
(sdenoffset:create-global "usebox" 1 "maxnumpoints" 800000 "recoverholes" 1)
(sdedelaunizer:set-parameters "maxPoints" 800000)

(sdedr:define-constant-profile "ndoping_def" "ArsenicConcentration" ndoping)
(sdedr:define-constant-profile "pdoping_def" "BoronConcentration" pdoping)

(sdedr:define-constant-profile-region "nregion_plac2" "ndoping_def" "GaN_wire_n")
(sdedr:define-constant-profile-region "pregion_plac1" "pdoping_def" "GaN_wire_p")
(sdedr:define-constant-profile-region "nregion_plac4" "ndoping_def" "GaN_Ox_int_n")
(sdedr:define-constant-profile-region "pregion_plac3" "pdoping_def" "GaN_Ox_int_p")
(sdedr:define-constant-profile-region "nregion_plac6" "ndoping_def" "Junction_n")
(sdedr:define-constant-profile-region "pregion_plac5" "pdoping_def" "Junction_p")

(define gas_fact 0.05)
(define wire_fact 0.05)
(define interface_fact (/ wire_fact 2.5))

(sdenoffset:create-noffset-block "region" "GaN_wire_n"
"maxlevel" 50 "maxedgelength" (* wire_fact  0.15))
(sdenoffset:create-noffset-block "region" "GaN_wire_p"
"maxlevel" 50 "maxedgelength" (* wire_fact  0.15))
(sdenoffset:create-noffset-block "region" "GaN_Ox_int_n"
"maxlevel" 50 "maxedgelength" (* interface_fact  0.1))
(sdenoffset:create-noffset-block "region" "GaN_Ox_int_p"
"maxlevel" 50 "maxedgelength" (* interface_fact  0.1))
(sdenoffset:create-noffset-block "region" "Junction_n"
"maxlevel" 50 "maxedgelength" (* interface_fact  0.1))
(sdenoffset:create-noffset-block "region" "Junction_p"
"maxlevel" 50 "maxedgelength" (* interface_fact  0.1))
(sdenoffset:create-noffset-block "region" "Gas_Surrounding"
"maxlevel" 50 "maxedgelength" (* gas_fact  0.15))
(sdenoffset:create-noffset-block "region" "Oxide_layer"
"maxlevel" 50 "maxedgelength" (* interface_fact  0.15))
(sdenoffset:create-noffset-block "region" "Gas_Interface"
```

```
"maxlevel" 50 "maxedgelength" (* interface_fact  0.15))

(sde:delay-graphics-update #f)
(sde:save-model geoname)
(sde:build-mesh "mesh" "-discontinuousData -noffset -d -F dfise" geoname)
```

5.2.3 Transistor

```
(define geoname  "n@node@_geo")
(define gridname "n@node@_mesh")

(define wire_mat "GaN")
(define ambient "Gas")

(define ref_wire 0.0001)
(define ref_air 0.01)

(define contact_hgth 0.025)

(define wire_lgth_n 1)
(define wire_lgth_p 0.3)
(define nwire_top (+ contact_hgth wire_lgth_n))
(define pwire_top (+ nwire_top wire_lgth_p))
(define top_cont_hgth (+ (+ nwire_top pwire_top)  contact_hgth))
(define btm_cont_hgth (+ 0 contact_hgth))
(define wire_rad @radius@)
(define air_top top_cont_hgth)
(define air_btm 0)
(define air_wdth (* wire_rad 2))
(define interface_wdth 0.01)
(define oxide_thickness 0.01)

(define ndop @n_doping@)
(define pdop @p_doping@)
(define ndoping (* ndop 1e6))
(define pdoping (* pdop 1e6))

(sde:delay-graphics-update #t)

(sdegeo:create-rectangle (position (+ wire_rad oxide_thickness) 0 0)
(position (+ (+ wire_rad oxide_thickness) air_wdth)(+ nwire_top pwire_top)
0)
ambient "Gas_Surrounding")
(sdegeo:create-rectangle (position wire_rad 0 0)
(position (+ wire_rad oxide_thickness) nwire_top 0)
"Oxide" "Oxide_layer1")
(sdegeo:create-rectangle (position wire_rad nwire_top 0)
(position (+ wire_rad oxide_thickness) pwire_top 0)
"Oxide" "Oxide_layer2")
(sdegeo:create-rectangle (position wire_rad pwire_top 0)
(position (+ wire_rad oxide_thickness) (+ nwire_top pwire_top) 0)
"Oxide" "Oxide_layer3")
```

```
(sdegeo:create-rectangle (position (+ wire_rad oxide_thickness) air_btm 0)
(position (+ (+ wire_rad oxide_thickness) interface_wdth) (+ nwire_top pwire_top) 0)
ambient "Gas_Interface")

(sdegeo:create-rectangle (position 0 0 0)
(position wire_rad nwire_top 0)
wire_mat "GaN_wire_n")
(sdegeo:create-rectangle (position 0 nwire_top 0)
(position wire_rad pwire_top 0)
wire_mat "GaN_wire_p")
(sdegeo:create-rectangle (position 0 pwire_top 0)
(position wire_rad (+ nwire_top pwire_top) 0)
wire_mat "GaN_wire_n2")

(sdegeo:define-contact-set "ncontact" 0.05 (color:rgb 1 0 0) "//")
(sdegeo:define-contact-set "pcontact" 0.05 (color:rgb 0 0 1) "++")
(sdegeo:define-2d-contact (find-edge-id
(position (/ wire_rad 2) 0 0)) "ncontact")
(sdegeo:define-2d-contact (find-edge-id
(position (/ wire_rad 2) (+ nwire_top pwire_top) 0)) "pcontact")

(sdedelaunizer:set-parameters "type" "boxmethod")
(sdedr:set-meshing-engine "noffset")
(sdenoffset:create-global "usebox" 1 "maxnumpoints" 800000 "recoverholes" 1)
(sdedelaunizer:set-parameters "maxPoints" 800000)

(sdedr:define-constant-profile "ndoping_def" "ArsenicConcentration" ndoping)
(sdedr:define-constant-profile "pdoping_def" "BoronConcentration" pdoping)

(sdedr:define-constant-profile-region "nregion_plac2" "ndoping_def" "GaN_wire_n")
(sdedr:define-constant-profile-region "nregion_plac3" "ndoping_def" "GaN_wire_n2")
(sdedr:define-constant-profile-region "pregion_plac1" "pdoping_def" "GaN_wire_p")

(define gas_fact 0.05)
(define wire_fact 0.05)
(define interface_fact (/ wire_fact 2))

(sdenoffset:create-noffset-block "region" "GaN_wire_n"
"maxlevel" 50   "maxedgelength" (* wire_fact  0.15))
(sdenoffset:create-noffset-block "region" "GaN_wire_n2"
"maxlevel" 50   "maxedgelength" (* wire_fact  0.15))
(sdenoffset:create-noffset-block "region" "GaN_wire_p"
"maxlevel" 50   "maxedgelength" (* interface_fact  0.15))
(sdenoffset:create-noffset-block "region" "Gas_Surrounding"
"maxlevel" 50  "maxedgelength"  (* gas_fact  0.15))
(sdenoffset:create-noffset-block "region" "Oxide_layer1"
"maxlevel" 50   "maxedgelength" (* interface_fact  0.15))
(sdenoffset:create-noffset-block "region" "Oxide_layer2"
"maxlevel" 50   "maxedgelength" (* interface_fact  0.15))
(sdenoffset:create-noffset-block "region" "Oxide_layer3"
"maxlevel" 50   "maxedgelength" (* interface_fact  0.15))
(sdenoffset:create-noffset-block "region" "Gas_Interface"
"maxlevel" 50   "maxedgelength" (* interface_fact  0.15))
```

```
(sde:delay-graphics-update #f)
(sde:save-model geoname)
(sde:build-mesh "mesh" "-discontinuousData -noffset -d -F dfise" geoname)
```

5.3 Examplary Simulation Command File

```
 files {

        resultdir = Results_Node@node@
        input     = n@previous@_geo_msh
        mdb       = quatra_matl.par
        SingleFileMDB
}

output {
        #AreaFactor=1e12
        #Luminescence { }
        #SRHPerEntity
        #InitialSolution
}


surface {
        Region   = [ GaN_wire1 ]
        Adjacent = [ Gas_Interface ]
        #TrapChargeOff
        AcceptorTraps {
          EnergyLevel          = 2.89  # eV with respect to valence band edge
          Density              = @Trap_Dens@  # cm^-2
          ElectronCaptureCoeff = @< @Trap_Dens@/@v_sr@ >@  # scm^-3
          HoleCaptureCoeff     = @< @Trap_Dens@/@v_sr@ >@  # scm^-3The file
        }


        #DonorTraps {
        #  EnergyLevel          = -0.5  # eV with respect to conduction band edge
        #Density                = @Trap_Dens@  # cm^-2
        # ElectronCaptureCoeff = @< @Trap_Dens@/2*@v_sr@ >@  # scm^-3
        # HoleCaptureCoeff     = @< @Trap_Dens@/2*@v_sr@ >@  # scm^-3
        #}

        #Charge { Constant = @Cs@ }
}


parameters{

        temperature            = @temperature@
        Mobility               = Arora
        FieldDependentMobility = TransferedElectron
```

```
        BodyOfRevolution
        SRH
        Auger {}
        Radiative
        #Lumirecomb

        Doping{
                Units = SI
        }

        #PolarizationCharges{
        #       InterfacesOnly
        #}
        #SpreadFunction = PredictorFermi
}


numerical {
        bulk_mindens_3d = 1e4

        #wellcont_mindens     = 1e-10
        #capture_max_exponent = 22.0 # up to now: 5
        pot_guess_const_neutral_zone
        dev_solver {
#               linear_solver          = pardiso
                linear_solver          = umfpack
                iterations             = 900
                osc_observation_cycles = 10

                bank_rose
                verbose
            }
#       dev_presolver {
#                 linear_solver   = pardiso
#                 linear_solver   = umfpack
#                 iterations      = 800
#                 bank_rose
#                 verbose
#           }

         ramper_charge {
     min_iterations = 2
     relative_error = 5e-5
   }
}


stationary{

        ncontact {
                voltage = 0.0
        }

        #boundaryCharge {
```

```
        #       min = @Min@
        #       max = @Max@
        #       minstep = 0.005
        #       maxstep = 0.01
        #       initstep = 0.01
        #       write_state = chrg_dens
        #}

        acceptorSurfTrapDens {
                min         = @Min@
                max         = @Max@
                minstep     = 0.0025
                maxstep     = 0.01
                initstep    = 0.005
                write_state = chrg_dens
        }


        #polChargeDensity {
          #min         = 0.0
          #max         = 0.5
          #minstep     = 0.025
          #maxstep     = 0.05
         # initstep    = 0.05
         # extrapolation
        #  write_state = sqw_piezo
        #}

        pcontact {
                #min       = 2.4
                #max       = 3.5
                #minstep   = 0.002
                #maxstep   = 0.02
                #initstep  = 0.02
                min        = 3.5
                max        = 0.0
                minstep    = -0.01
                maxstep    = -0.02
                initstep   = -0.02
                #extrapolation

                output {
                  step = -0.05
                 # luminescence
                }
        }

}
```

6 List of publications

1. Hamed Kamrani, Feng Yu, Kristian Frank, Klaas Strempel, Muhammad Fahlesa Fatahilah, Hutomo Suryo Wasisto, Friedhard Römer, Andreas Waag, Bernd Witzigmann, Thermal performance analysis of GaN nanowire and fin-shaped power transistors based on self-consistent electrothermal simulations, Microelectronics Reliability, Volume 91, Part 2, 2018, Pages 227-231, doi.org/10.1016/j.microrel.2018.10.007.

2. Bernd Witzigmann, Feng Yu, Kristian Frank, Klaas Strempel, Muhammad Fahlesa Fatahilah, Hans Werner Schumacher, Hutomo Suryo Wasisto, Friedhard Römer, Andreas Waag, Performance analysis and simulation of vertical gallium nitride nanowire transistors, Solid-State Electronics, Volume 144, 2018, Pages 73-77.

3. Muhammad Fahlesa Fatahilah, Feng Yu, Klaas Strempel, Kristian Frank, Friedhard Römer, Andrey Bakin, Bernd Witzigmann, Hutomo Suryo Wasisto, and Andreas Waag, 3D GaN nanowire arrays for E-mode vertical field-effect transistors, IWN conference, Kanazawa, Japan, 2018.

4. Klaas Strempel, Feng Yu, Jana Hartmann, Hendrik Spende, Kristian Frank, Friedhard Römer, Bernd Witzigmann, Andrey Bakin, Hergo-Heinrich Wehmann, Hutomo Suryo Wasisto, and Andreas Waag, Vertical field-effect transistors based on 3D GaN nanofin arrays, IWN conference, Kanazawa, Japan, 2018.

5. Hutomo Suryo Wasisto, Feng Yu, Klaas Strempel, Muhammad Fahlesa Fatahilah, Kristian Frank, Hamed Kamrani, Friedhard Römer, Hans Werner Schumacher, Bernd Witzigmann, Andreas Waag, Vertical 3D GaN nanowire field-effect transistors, Material Research Society MRS-id conference, Bali, Indonesia, 31 July - 2 August 2018.

6. Muhammad Fahlesa Fatahilah, Feng Yu, Klaas Strempel, Kristian Frank, Friedhard Römer, Bernd Witzigmann, Andrey Bakin, Hutomo Suryo Wasisto, Andreas Waag, 3D GaN nanowire arrays for vertical field-effect transistors, ICPS conference, Montpellier, France, 2018.

7. Klaas Strempel, Feng Yu, Jana Hartmann, Hendrik Spende, Kristian Frank, Friedhard Römer, Bernd Witzigmann, Sönke Fündling, Hergo-Heinrich Wehmann, Hutomo Suryo Wasisto, Andreas Waag, Vertical field-effect transistors based on 3D GaN fin arrays, ICPS conference, Montpellier, France, 2018.

8. Feng Yu, Klaas Strempel, Muhammad Fahlesa Fatahilah, Kristian Frank, Friedhard Romer, Andrey Bakin, Bernd Witzigmann, Hutomo Suryo Wasisto, Andreas Waag,

Vertical GaN nanowire field-effect transistors, GaN Marathon 2.0, Padova, Italy, April 18-19, 2018.

9. Klaas Strempel, Jana Hartmann, Feng Yu, Hendrik Spende, Muhammad Fatahilah, Friedhard Römer, Kristian Frank, Bernd Witzigmann, Sönke Fündling, Hutomo Suryo Wasisto, Andreas Waag, 3D GaN fin arrays for vertical field-effect transistors, ICNS Conference, Strasbourg, France, July 2017.

10. Kristian Frank, Friedhard Römer and Benrd Witzigmann, "GaN Nanowires for Highly Selective Gas Sensing", in *Center for Interdisciplinary Nanostructure Science and Technology (CINSaT)*, Poster presentation, February 2017.

11. Kristian Frank, Friedhard Römer and Benrd Witzigmann, "GaN Nanowires for Highly Selective Gas Sensing", in *Center for Interdisciplinary Nanostructure Science and Technology (CINSaT)*, Poster presentation, Octobre 2016.

12. Kristian Frank, Friedhard Römer and Benrd Witzigmann, "GaN Nanowires for Sensing", in *Center for Interdisciplinary Nanostructure Science and Technology (CINSaT)*, Poster presentation, February 2015.

Bibliography

[1] J. Shin, S. Han, S. Noh, Y.-T. Yu, and J. S. Kim, "Room-temperature operation of light-assisted NO_2 gas sensor based on GaN nanowires and graphene," *Nanotechnology*, vol. 32, p. 505201, sep 2021.

[2] M. Hoffmann, J. Prades, L. Mayrhofer, F. Hernandez-Ramirez, T. Järvi, M. Moseler, A. Waag, and H. Shen, "Sensors: Highly selective sam-nanowire hybrid no 2 sensor: Insight into charge transfer dynamics and alignment of frontier molecular orbitals (adv. funct. mater. 5/2014)," *Advanced Functional Materials*, vol. 24, 02 2014.

[3] M. Reddeppa, B.-G. Park, N. D. Chinh, D. Kim, J.-E. Oh, T. G. Kim, and M.-D. Kim, "A novel low-temperature resistive no gas sensor based on ingan/gan multi-quantum well-embedded p–i–n gan nanorods," *Dalton Trans.*, vol. 48, pp. 1367–1375, 2019.

[4] Q. Abdullah, F. Yam, Z. Hassan, and M. Bououdina, "Hydrogen gas sensing performance of gan nanowires-based sensor at low operating temperature," *Sensors and Actuators B: Chemical*, vol. 204, pp. 497–506, 2014.

[5] M. Folke, L. Cernerud, M. Ekström, and B. Hök, "Critical review of non-invasive respiratory monitoring in medical care," *Medical & biological engineering & computing*, vol. 41, pp. 377–83, 08 2003.

[6] M. A. H. Khan, B. Thomson, J. Yu, R. Debnath, A. Motayed, and M. V. Rao, "Scalable metal oxide functionalized gan nanowire for precise so2 detection," *Sensors and Actuators B: Chemical*, vol. 318, p. 128223, 2020.

[7] A. C. Gad and A. Waag, "Prozesstechnologien für die halbleiter-nanodraht-basierte sensorik mit hohem durchsatz," *Presentation given by Institut für Halbleitertechnik TU Braunschweig at WireControl Annual Meeting in Munich*, 2015.

[8] S. A. Eliza and A. K. Dutta, "Ultra-high sensitivity gas sensors based on gan hemt structures," in *International Conference on Electrical & Computer Engineering (ICECE 2010)*, pp. 431–433, 2010.

[9] G. Steinhoff, M. Hermann, W. J. Schaff, L. F. Eastman, M. Stutzmann, and M. Eickhoff, "ph response of gan surfaces and its application for ph-sensitive field-effect transistors," *Applied Physics Letters*, vol. 83, no. 1, pp. 177–179, 2003.

[10] J. Sim, K. Kim, S. Song, and J. Kim, "Suspended gan nanowires as no2 sensor for high temperature applications," *Analyst*, vol. 138, pp. 2432–2437, 2013.

[11] O. Lupan, V. Ursaki, G. Chai, L. Chow, G. Emelchenko, I. Tiginyanu, A. Gruzintsev, and A. Redkin, "Selective hydrogen gas nanosensor using individual zno nanowire with fast response at room temperature," *Sensors and Actuators B: Chemical*, vol. 144, no. 1, pp. 56–66, 2010.

[12] M. Reddeppa, T. K. Phung Nguyen, B.-G. Park, S.-G. Kim, and M.-D. Kim, "Low operating temperature no gas sensors based hydrogen peroxide treated gan nanorods," *Physica E: Low-dimensional Systems and Nanostructures*, vol. 116, p. 113725, 2020.

[13] T. Han, A. Nag, S. Chandra Mukhopadhyay, and Y. Xu, "Carbon nanotubes and its gas-sensing applications: A review," *Sensors and Actuators A: Physical*, vol. 291, pp. 107–143, 2019.

[14] S. Li and A. Waag, "Gan based nanorods for solid state lighting," *Journal of Applied Physics 111, 071101*, 2012.

[15] C. Bishop, J. P. Salvestrini, Y. Halfaya, S. Sundaram, Y. El Gmili, L. Pradere, J. Y. Marteau, M. B. Assouar, P. L. Voss, and A. Ougazzaden, "Highly sensitive detection of no2 gas using bgan/gan superlattice-based double schottky junction sensors," *Applied Physics Letters*, vol. 106, no. 24, p. 243504, 2015.

[16] F. Römer, "Quatra reference manual," *Computational Electronics and Photonics Group University of Kassel*, 2016.

[17] T. T. Mnatsakanov, M. E. Levinshtein, L. I. Pomortseva, S. N. Yurkov, G. S. Simin, and M. Asif Khan, "Carrier mobility model for gan," *Solid-State Electronics*, vol. 47, no. 1, pp. 111–115, 2003.

[18] C. Canali, G. Majni, R. Minder, and G. Ottaviani, "Electron and hole drift velocity measurements in silicon and their empirical relation to electric field and temperature," *IEEE Transactions on Electron Devices*, vol. 22, no. 11, pp. 1045–1047, 1975.

[19] V. O. Turin, "A modified transferred-electron high-field mobility model for gan devices simulation," *Solid-State Electronics*, vol. 49, no. 10, pp. 1678–1682, 2005.

[20] M. Farahmand, C. Garetto, E. Bellotti, K. F. Brennan, M. Goano, E. Ghillino, G. Ghione, J. D. Albrecht, and P. P. Ruden, "Monte carlo simulation of electron transport in the iii-nitride wurtzite phase materials system: binaries and ternaries," *IEEE Transactions on Electron Devices*, vol. 48, pp. 535–542, March 2001.

[21] F. Schwierz, "An electron mobility model for wurtzite gan," *Solid-State Electronics*, vol. 49, no. 6, pp. 889–895, 2005.

[22] J. Piprek, *Semiconductor Optoelectronic Devices*. Academic Press, 2003.

[23] V. M. Polyakov and F. Schwierz, "Influence of electron mobility modeling on dc i-v characteristics of wz-gan mesfet," *IEEE Transactions on Electron Devices*, vol. 48, pp. 512–516, March 2001.

[24] S. K. O'Leary, P. Siddiqua, W. A. Hadi, B. E. Foutz, M. S. Shur, and L. F. Eastman, *Electron Transport Within III-V Nitride Semiconductors*. Cham: Springer International Publishing, 2017.

[25] W. Shockley and W. T. Read, "Statistics of the recombinations of holes and electrons," *Phys. Rev.*, vol. 87, pp. 835–842, Sep 1952.

[26] R. N. Hall, "Electron-hole recombination in germanium," *Phys. Rev.*, vol. 87, pp. 387–387, Jul 1952.

[27] F. Römer, M. Deppner, and B. Witzigmann, "Computational modelling of surface effects in ingan/gan quantum disk nano wire leds," *SPIE Photonics West, San Francisco*, 2013.

[28] F. Römer and B. Witzigmann, "Modelling surface effects in nano wire optoelectronic devices," *Journal of Computational Electronics*, vol. 11, pp. 431–439, Dec 2012.

[29] R. B. Darling, "Defect-state occupation, fermi-level pinning, and illumination effects on free semiconductor surfaces," *Phys. Rev. B*, vol. 43, pp. 4071–4083, Feb 1991.

[30] S. M. Sze and K. K. NG, *Physics of Semiconductor Devices*. Wiley, 2007.

[31] G. Coco, L. Mayrhofer, and M. Moseler, "Wirecontrol annual meeting," *Presentation given by Project Group of the University of Freiburg*, 2015.

[32] H. Ibach, *Physics of Surfaces and Interfaces*. Springer-Verlag, 2006.

[33] R. Calarco, T. Stoica, O. Brandt, and L. Geelhaar, "Surface-induced effects in gan nanowires," *Journal of Materials Research*, vol. 26, pp. 2157 – 2168, 09 2011.

[34] C. G. Van de Walle and D. Segev, "Microscopic origins of surface states on nitride surfaces," *Journal of Applied Physics*, vol. 101, no. 8, p. 081704, 2007.

[35] J. E. Northrup and J. Neugebauer, "Theory of GaN($10\bar{1}0$) and ($11\bar{2}0$) surfaces," *Phys. Rev. B*, vol. 53, pp. R10477–R10480, Apr 1996.

[36] M. Bertelli, P. Löptien, M. Wenderoth, A. Rizzi, R. Ulbrich, M. C. Righi, A. Ferretti, L. Martin-Samos, C. Bertoni, and A. Catellani, "Atomic and electronic structure of the nonpolar gan (1 1 ¯ 00) surface," *Physical Review B*, vol. 80, 09 2009.

[37] L. Ivanova, S. Borisova, H. Eisele, M. Dähne, A. Laubsch, and P. Ebert, "Surface states and origin of the fermi level pinning on nonpolar gan(11¯00) surfaces," *Applied Physics Letters*, vol. 93, no. 19, p. 192110, 2008.

[38] D. Segev and C. G. Van de Walle, "Electronic structure of nitride surfaces," *Journal of Crystal Growth*, vol. 300, no. 1, pp. 199–203, 2007. First International Symposium on Growth of Nitrides.

[39] T. Richter, H. Lüth, R. Meijers, R. Calarco, and M. Marso, "Doping concentration of gan nanowires determined by opto-electrical measurements," *Nano Letters*, vol. 8, no. 9, pp. 3056–3059, 2008. PMID: 18687013.

[40] H. A. R. Wegener, "The cylindrical field-effect transistor," *IRE Transactions on Electron Devices*, vol. 6, no. 442, 1959.

[41] A. A. Talin, F. Léonard, B. S. Swartzentruber, X. Wang, and S. D. Hersee, "Unusually strong space-charge-limited current in thin wires," *Phys. Rev. Lett.*, vol. 101, p. 076802, Aug 2008.

[42] C. Kittel, *Einführung in die Festkörperphysik*. Wiley, 2005.

[43] E. Schrödinger, "Quantisierung als eigenwertproblem," *Annalen der Physik*, vol. 385, no. 13, pp. 437–490, 1926.

[44] D. Meschede, *Gerthsen Physik*. Springer-Verlag, 24 ed., 2010.

[45] Bergmann and Schaefer, *Band 6: Festkörper*. de Gruyter, 2005.

[46] A. Wallraff, "Lecture slides physik iv," *ETH Zurich*, 2009.

[47] K. J. Langenberg, *Theorie elektromagnetischer Wellen*. Department Theoretische Elektrotechnik at the faculty of Electrical Enginieering/Computer Science, University of Kassel, 2003.

[48] "Ioffe institute." NSM Archive, http://www.ioffe.ru/SVA/NSM/Semicond/GaN/bandstr.html, Dec 2020.

[49] S. Wang, *Fundamentals of Semiconductor Theory and Device Physics*. Prentice-Hall Inc., 1989.

[50] Synopsys, "Sentaurus tm device user guide," *Version I-2013.12*, Dec 2013.

[51] P. Leuchtmann, *Einführung in die elektromagnetische Feldtheorie*. Pearson Studium, 2005.

[52] S. L. Chuang, *Physics of Photonic Devices*. Wiley, 2009.

[53] I. Bronstein, K. Semendjajew, G. Musiol, and H. Mühlig, *Taschenbuch der Mathematik*. Harri Deutsch Verlag, 5 ed., 2001.

[54] J.-J. Huang, H.-C. Kuo, and S.-C. Shen, *Nitride Semiconductor Light-Emitting Diodes (LEDs): Materials, Technologies and Applications*. Woodhead Publishing, 2018.

List of Figures