

The Title

The Author

The Date

Inhaltsverzeichnis

Einleitung	v
1 Simulation in der Halbleiterphysik	1
1.1 Aktueller Stand in der Halbleitersimulation	1
1.2 Neuer Ansatz für Nichtgleichgewichts-Situationen	3
1.3 Eingabesystem	5
2 Bandstruktur	9
2.1 Grundlagen	10
2.2 Darstellung des kp-Hamiltonoperators	14
2.3 Diskretisierung	22
2.4 Einband-Schrödingergleichung	25
2.5 Magnetfeld	26
2.6 Zusammenfassung	28
3 Elektronische Struktur im Nichtgleichgewicht	29
3.1 Pauli-Master-Gleichung	30
3.2 Neues Transportmodell für dissipativen Transport	32
3.3 Vergleich der Methoden	39
3.4 Zusammenfassung und Ausblick	44
4 Halbleitergleichungen	45
4.1 Poissongleichung	45
4.2 Stromgleichungen	47
4.3 Spannungsgleichung	49
4.4 Numerik partieller Differenzialgleichungen	54
4.5 Zusammenfassung	59
5 Numerische Lösung des Gesamtproblems	61
5.1 Mathematische Formulierung des Gleichungssystems	61
5.2 Selbstkonsistente Lösung der Schrödinger- und Poissongleichung	64
5.2.1 Nichtlineare Poissongleichung	64
5.2.2 Newton-Verfahren	65

5.2.3 Schrödinger-Poisson	66
5.3 Lösung des Nichtgleichgewichtsproblems	68
5.3.1 Blockiteratives Verfahren	68
5.3.2 Gekoppelte Lösung	69
5.4 Zusammenfassung	71
6 Simulation einzelner Quantenpunkte	73
6.1 Experimenteller Aufbau	74
6.2 Grundgeometrie der Quantenpunkte	75
6.3 Inhomogenes Legierungsprofil	88
6.4 Optische Eigenschaften	90
6.5 Exzitonische Korrektur	94
6.6 Simulation von Quantenpunkten mit Magnetfeld	98
7 Weitere Anwendungen	103
7.1 Nano-MOSFET	103
7.1.1 Double-Gate-MOSFET	103
7.1.2 Vergleich mit Single-Gate-MOSFET	115
7.1.3 Zusammenfassung	118
7.2 Si/SiGe-Quantenkaskaden-Strukturen	119
7.3 Zweidimensionale Löchgase in GaN/AlGaIn-Heterostrukturen	124
A Diskretisierung der Poisson- und Stromgleichung	129
B Diskretisierung der Spannungsgleichung	141
C Kp-Hamiltonoperator für Wurtzit	151
D Modelle für die Dichteberechnung in 1D, 2D und 3D	157
E Diskretisierung des Hamiltonoperators	165
F Fermi-Funktionen	173
G Modellierung von Metall-Oxid-Silizium- Kontakten	177
H Ladungsträgerbeweglichkeit und Rekombination	181
I Bandkantenverlauf in Heterostrukturen	185
Danksagung	195

Einleitung

Die Entwicklung in der Halbleiterindustrie ist seit Jahrzehnten durch eine kontinuierliche Miniaturisierung geprägt. Moderne Technologien ermöglichen es mittlerweile, Bauelemente im Nanometerbereich herzustellen. Dies hat den Begriff Nanoelektronik geprägt.

Der Erfolg der Elektronik beruht dabei größtenteils auf der Entwicklung des Silizium-MOSFETs (Metal-Oxide-Semiconductor-Field-Effect-Transistor) und der auf ihm basierenden CMOS-Technologie zur Herstellung integrierter Schaltkreise. Daneben spielen (3)-(5)-Verbindungshalbleiter aufgrund ihrer direkten Bandlücke in optoelektronischen Bauelementen eine herausragende Rolle.

Schließlich sind Nitrid-Halbleiter für Anwendungen in der Mikrowellentechnik in den Mittelpunkt des Interesses gerückt. Dies liegt an ihrer hohen Bandlücke, die den Einsatz im Bereich der Leistungselektronik ermöglicht.

Das Vordringen in den Bereich der Nano-Physik erfordert eine Weiterentwicklung und Neubestimmung der bestehenden Konzepte zur Modellierung und theoretischen Beschreibung der Bauelemente. Die Verkleinerung der Strukturgrößen rückt insbesondere die Gesetze der Quantenmechanik in den Vordergrund. Bisherige theoretische Arbeiten kann man grob in zwei Klassen unterteilen. Die eine Gruppe beschränkt sich auf Situationen im thermodynamischen Gleichgewicht und legt den Schwerpunkt auf die Berechnung der elektronischen Struktur. Die andere Gruppe stellt Transportrechnungen in den Mittelpunkt, z.B. bei der Simulation oben genannter MOSFETs, macht aber starke Näherungen bei der elektronischen Struktur. Es ist das Hauptanliegen dieser Arbeit, ein neues Modell vorzustellen, das beide Ansätze verbindet. Konkret bedeutet dies, dass Bauelemente in Nichtgleichgewichtssituationen unter möglichst genauer Berücksichtigung der Bandstruktur untersucht werden sollen. Dies erfordert zwingend die Berücksichtigung der mechanischen Verspannungen in Heterostrukturen sowie die daraus resultierenden piezoelektrischen Effekte. In Wurtzit-Nitriden muss zudem die spontane Polarisierung miteinbezogen werden. Gleichzeitig sollten die Probleme mit vertretbarem Rechenaufwand lösbar sein. Mit dieser Arbeit wird somit ein Rahmen geschaffen, der ein detailliertes Verständnis von experimentellen Ergebnissen ermöglicht und zugleich Hilfsmittel ist für den Entwurf neuartiger Bauelemente.

In den ersten fünf Kapiteln werden die verwendeten Methoden erläutert und

die zugrunde liegenden Annahmen diskutiert. Kapitel 1 gibt einen Überblick und stellt kurz dar, in welcher Weise die berücksichtigten Gleichungen - das sind die Verspannungsgleichung, die Mehrband-kp-Schrödingergleichung, die Poissongleichung und die Stromgleichung - voneinander abhängen. In Kapitel 2 wird die Berechnung der elektronischen Struktur im Rahmen der 8×8 -kp-Theorie erläutert. Dabei gehen wir insbesondere auch auf den Einfluss der Verspannungen ein. Kapitel 3 behandelt die elektronische Struktur für Nichtgleichgewichtssituationen. Im Wesentlichen wird hier ein Modell vorgeschlagen, das stationäre quantenmechanische Zustände mit einer semiklassischen Transportgleichung verbindet. Kapitel 4 stellt schließlich die Halbleitergleichungen zusammen und erläutert die dazugehörigen Randbedingungen. Mechanische Verspannungen werden im Rahmen der Elastizitätstheorie berechnet. Das Kapitel endet mit der Darstellung des von uns verwendeten Diskretisierungsverfahrens, nämlich der Methode der finiten Volumen. Die selbstkonsistente Lösung der genannten Gleichungen kann mitunter recht kompliziert sein. Auf diesen Problemkreis gehen wir in Kapitel 5 ein.

Kernstück dieser Arbeit ist die Simulation von einzelnen Quantenpunkten. Hier sind Simulationsrechnungen unerlässlich zur Interpretation von experimentellen Ergebnissen. Wir zeigen anhand eines gegebenen Experiments, wie durch systematische Rechnungen Rückschlüsse auf die Geometrie und Größe der Quantenpunkte gezogen werden können. Geht es bei Quantenpunkten vor allem um optische Eigenschaften und die elektronische Struktur im Nichtgleichgewicht, d.h. unter angelegter Spannung, so stehen bei MOS-Transistoren Strom-Spannungskennlinien im Vordergrund. Der technologische Fortschritt ermöglicht die Herstellung von Transistoren mit Gate-Längen im Nanometerbereich. Unser Ziel ist es, die Auswirkungen von quantenmechanischen Effekten auf das elektronische Verhalten eines Double-Gate-MOSFETs darzustellen. Dabei spielen insbesondere Kurzkanaleffekte eine Rolle. Als weitere Anwendung untersuchen wir SiGe-Quantenkaskadenstrukturen, die für die Entwicklung von Lasern auf Silizium-Basis von Bedeutung sind. Der Schwerpunkt liegt hier bei der Berechnung der Bandstruktur und den daraus resultierenden Streuzeiten für strahlende und nicht-strahlende Übergänge. Schließlich untersuchen wir zweidimensionale Löchergase in GaN/AlGaIn-Heterostrukturen. Es zeigt sich, dass trotz der hohen Ionisierungsenergien der Akzeptoren hohe Löcherdichten erreichbar sind durch gezieltes Ausnutzen von piezo- und pyroelektrischen Polarisationsladungen.

Kapitel 1

Simulation in der Halbleiterphysik

1.1 Aktueller Stand in der Halbleitersimulation

Der effiziente Entwurf von Halbleiterbauelementen erfordert zwingend leistungsfähige Simulationswerkzeuge zur Vorhersage des elektrischen Verhaltens der entworfenen Halbleiterstrukturen. Vor diesem Hintergrund hat es in den letzten zwei Jahrzehnten eine umfangreiche Forschungstätigkeit gegeben, die einerseits neue Methoden in der Modellierung von Bauelementen hervorbrachte und andererseits umfangreiche Programmpakete zur Verfügung stellte. Durch die zunehmende Miniaturisierung in der Halbleitertechnologie hat dabei in den letzten Jahren eine Verschiebung des Hauptaugenmerks auf die Beschreibung von quantenmechanischen Effekten stattgefunden.

Die bisher vorhandenen Simulatoren können grob in zwei Klassen unterteilt werden. Die erste Kategorie beschränkt sich auf Situationen im thermodynamischen Gleichgewicht und legt den Schwerpunkt auf eine möglichst genaue Beschreibung der elektronischen Struktur. In der Literatur sind dazu zahlreiche Arbeiten erschienen, die sich mit Heterostrukturen wie Quantenpunkten, Quantendrähten und Quantengraben befassen. Die Bandstruktur wird dabei mit Hilfe der Einband-Schrödingergleichung [1], der Mehrband-kp-Schrödingergleichung [2][3][4] oder im Rahmen der Tight-Binding-Theorie beschrieben [5]. Diese Arbeiten haben jedoch gemeinsam, dass sie sich auf spezielle Probleme beziehen und für ein Simulationsprogramm nicht allgemein genug gehalten sind. Unseren Bedürfnissen am nächsten kommt der Simulator SCRAPS von S. Laux, A. Kumar und F. Stern. Dieser verfügt über ein flexibles Eingabesystem, hat aber die wesentliche Einschränkung, dass nur Elektronen bei parabolischer Einband-Näherung quantenmechanisch, die Löcher dagegen klassisch beschrieben werden. Eine akkurate Modellierung verspannter elektronischer Strukturen in Heterobauelementen macht jedoch die Lösung einer Mehrband-kp-Schrödingergleichung unumgäng-

lich.

Die zweite Gruppe der Bauelementsimulatoren stellt die Berechnung von Strom-Spannungskennlinien in den Vordergrund und nimmt im Gegenzug erhebliche Vereinfachungen bei der Behandlung der elektronischen Struktur in Kauf. Vor allem im Umfeld der Elektrotechnik sind hier mehrere, zum Teil kommerzielle Programmpakete entwickelt worden [6][7][8][9]. Diese sind jedoch in der Regel auf spezielle klassische Bauelemente wie Feldeffekttransistoren zugeschnitten und oft auf Silizium beschränkt. Entsprechend ihres Einsatzgebietes in der industriellen Forschung beinhalten sie oft zusätzliche Module zur Prozess-Simulation [7][8]. Der Berechnung der Ströme liegt die Annahme zugrunde, dass die Ladungsträgerverteilung nicht allzu weit vom Gleichgewichtszustand abweicht. Dann kann man von der Boltzmann-Gleichung mit Relaxationsansatz ausgehen, deren erstes Moment auf die Drift-Diffusionsgleichung führt. Die Berücksichtigung des zweiten Moments ergibt zusätzlich die hydrodynamische Gleichung. Auch hier hat es in den letzten Jahren Bemühungen gegeben, quantenmechanische Effekte wenigstens ansatzweise zu berücksichtigen. Um die Rechenzeiten gering zu halten, wird aber immer noch von einer Lösung der Schrödingergleichung abgesehen und stattdessen die Ladungsträgerdichte im Rahmen der Dichtegradientenmethode störungstheoretisch korrigiert [10]. Eindimensionale Tunnelströme können mithilfe eines WKB-Ansatzes beschrieben werden [11]. Alternativ zu den Drift-Diffusionsgleichungen gibt es bewährte Ansätze zur direkten Lösung der Transportgleichung mittels Monte-Carlo-Methoden [12][13][14], die jedoch meist die klassische Boltzmann-Gleichung zugrunde legen.

Erst seit wenigen Jahren wird versucht, diese Methoden in Richtung Quantentransport zu erweitern. Einerseits geschieht dies durch Übergang von der Fermi-Verteilung zur Wignerfunktion [15], welcher zu der bereits erwähnten Dichtegradientenmethode führt. Dieser Ansatz ist zumindest ausreichend zur Beschreibung der Strom-Spannungskennlinien in Kanälen mit Breiten im Nanometerbereich, er liefert jedoch keinerlei Aussagen über die quantisierten Zustände. Fischetti beschränkt sich auf Bauelemente mit Abmessungen, die klein gegen die Wellenlänge des Elektrons sind, und beschreibt den Transport im Rahmen der Pauli-Master-Gleichung als Streuproblem [16][17]. Am weitesten gehen Lake und Mitarbeiter [18], die in ihrem Simulator NEMO das Transportproblem mit Nichtgleichgewichts-Greenfunktionen beschreiben. Diese haben jedoch einerseits das Problem, dass sehr viele Näherungen eingehen, die das physikalische Verhalten des simulierten Bauelements beinahe unüberschaubar machen, andererseits ist der Rechenaufwand selbst in einer Dimension derart hoch, dass eine Erweiterung dieses Verfahrens auf zwei- und dreidimensionale Probleme wenig sinnvoll erscheint.

Vor diesem Hintergrund verfolgen wir das Ziel, die Ansätze aus den beiden beschriebenen Gruppen von Simulatoren zu kombinieren. Zum einen soll die elektronische Struktur möglichst genau beschrieben werden, zum anderen sollen auch Nichtgleichgewichtssituationen erfasst werden. Dies wird im nächsten Abschnitt

genauer erläutert.

1.2 Neuer Ansatz für Nichtgleichgewichts-Situationen

Für die realistische Simulation von Halbleiter-Nanostrukturen ist es unerlässlich, ihr Verhalten auch bei angelegter äußerer Spannung beschreiben zu können. Als Kern dieser Arbeit wird ein Verfahren vorgestellt, das es erlaubt, derartige Situation in vernünftiger Zeit sinnvoll zu behandeln. Als Nebenprodukt ist ein Programmpaket namens NEXTNANO³ entstanden, welches darauf ausgelegt ist, für Halbleiter-Nanostrukturen die elektronische Struktur sowie die daraus resultierende Ladungsdichte für Nichtgleichgewichts-Situationen selbstkonsistent zu berechnen. Dabei nehmen wir an, dass sich die Ladungsträger in einem lokalen thermodynamischen Gleichgewicht befinden. Dies erlaubt uns die Beschreibung der Nichtgleichgewichts-Situation über zwei Quasi-Ferminiveaus, eines für Elektronen und eines für Löcher. Es zeigt sich, dass für die Berechnung dieser Quasi-Ferminiveaus die semiklassische Drift-Diffusionsgleichung gut geeignet ist. Die Kombination von stationären quantenmechanischen Zuständen mit einer Transportgleichung, die in vernünftiger Zeit lösbar ist, stellt eine wesentliche Neuerung im Vergleich zu den bereits existierenden Simulationsmethoden dar. Gleichzeitig lassen wir für die Berechnung der Verspannungen und der Bandstruktur eine beliebige Kristallorientierung zu ebenso wie Magnetfelder. Die Unterteilung des Bauelements in Quantenregionen, wo die Dichte quantenmechanisch berechnet wird, und klassischen Regionen mit einer Dichteberechnung mithilfe der Thomas-Fermi-Zustandsdichte ermöglicht schließlich die effiziente Simulation von Strukturen vom Mikro- bis zum Nanometerbereich.

In Abb. 1.1 sind die wesentlichen Zusammenhänge zwischen den Eingabedaten und dem Simulationsergebnis, sowie die Abhängigkeiten der physikalischen Modelle untereinander grafisch dargestellt.

Die wichtigsten Eingabedaten sind die Materialgeometrie, das Dotierprofil sowie diverse Randbedingungen, wie elektrische Kontakte, mechanischer Druck auf die Probe und angelegtes Magnetfeld. Die selbstkonsistente Lösung wird im Prinzip durch das elektrostatische Potenzial und die Quasi-Ferminiveaus für Elektronen und Löcher vollständig beschrieben.

Zunächst werden die inneren Verspannungen, die durch Gitterfehlpassungen an Heterogrenzflächen verursacht werden, im Rahmen der Elastizitätstheorie berechnet. Als Randbedingung geht hier ein eventuell vorhandener äußerer hydrostatischer Druck ein. Die Verspannungen führen zu einem piezoelektrischen Polarisationsfeld, welches zusammen mit pyroelektrischer Polarisation zu festen Raumladungen führt. Ebenso kommt es über die Deformationspotenziale zur Verschiebung der Leitungs- und Valenzbandenergien und zu einer Aufspaltung der Ener-

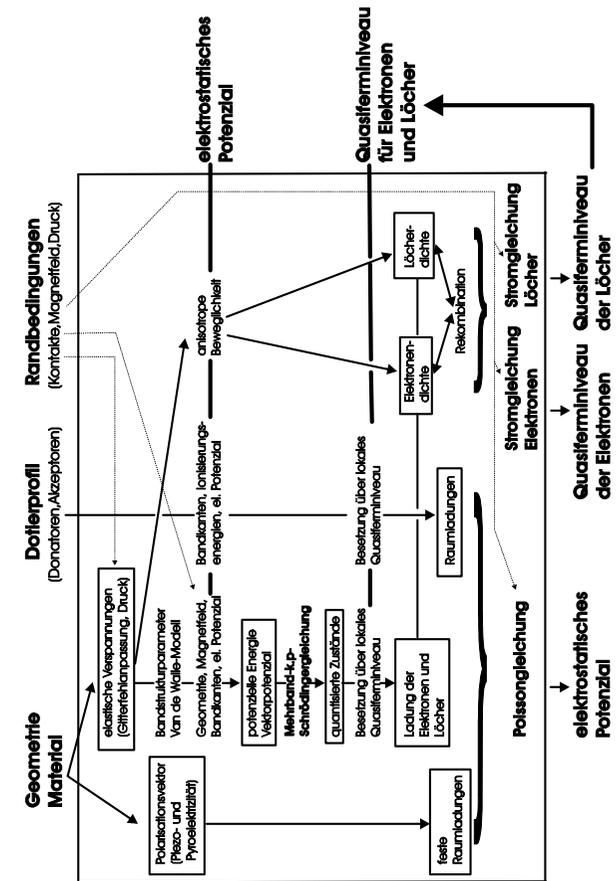


Abbildung 1.1: Überblick über die Zusammenhänge zwischen den physikalischen Modellen. Von der Eingabeseite ist das Problem durch die Materialgeometrie, das Dotierprofil und die Randbedingungen definiert. Die selbstkonsistente Lösung wird vollständig durch das elektrische Potenzial und die Quasi-Ferminiveaus bestimmt.

gjeniveau. Zusammen mit dem elektrostatischen Potenzial ergibt sich daraus der räumliche Verlauf der potenziellen Energie, das Vektorpotenzial wird durch das angelegte Magnetfeld bestimmt. Nun können durch Lösen der Mehrband-kp- oder der Einband-Schrödingergleichung die quantisierten Zustände berechnet werden. Die Ladungsdichte erhalten wir durch Besetzen dieser Zustände über das lokale Quasi-Ferminiveau. Dies ist eine der zentralen Annahmen, auf die in Kap. 3 näher eingegangen wird. Einen weiteren Beitrag zur Dichte liefern Dotierungen, d.h. Donatoren oder Akzeptoren, deren energetische Lage durch das elektrostatische Potenzial und ihre Ionisierungsenergien bestimmt wird. Aufgrund der Annahme eines lokalen thermodynamischen Gleichgewichts werden die Störstellen ebenfalls über das lokale Quasi-Ferminiveau besetzt. Bei gegebener Dichte ergibt sich nun das elektrostatische Potenzial aus der Poissongleichung. Dabei werden als Randbedingungen ohmsche Kontakte, Schottky-Kontakte und Oberflächenladungen berücksichtigt.

Zur Bestimmung der Quasi-Ferminiveaus muss eine Stromgleichung für Elektronen und Löcher gelöst werden. Dazu nehmen wir an, dass der Strom proportional der Ladungsträgerdichte und dem Gradienten des Quasi-Ferminiveaus ist. Der Proportionalitätsfaktor ist abhängig von der Relaxationszeit, dem elektrischen Feld in Bewegungsrichtung und der Verspannung, und stellt im Allgemeinen einen Tensor zweiter Stufe dar. Die Stromgleichungen für die Elektronen und Löcher sind zudem über Rekombinationsterme gekoppelt. Als Randbedingungen werden wiederum ohmsche Kontakte und Schottky-Kontakte in Betracht gezogen. Diese Zusammenhänge werden in Kap. 4 genauer erläutert.

Eine selbstkonsistente Lösung des Problems ist gegeben, wenn das elektrostatische Potenzial und die Quasi-Ferminiveaus Werte erreicht haben, für die alle diese Beziehungen gleichzeitig erfüllt sind. Numerisch stellt dies ein gewaltiges Problem dar, das in Kap. 5 dargestellt wird.

1.3 Eingabesystem

Geometrieverarbeitung

Um komplexe physikalische Systeme wie InGaAs-Quantenpunkte und Si-Transistoren effizient handhaben zu können, war es nötig, ein flexibles Eingabesystem für beliebige (3)-(5)- und (4)-(4)-Halbleiter vorzusehen. Dieses wurde im Wesentlichen von Herrn Dr. Günther Zandler entwickelt und ist so ausgelegt, dass es beliebige Geometriestrukturen verarbeiten kann. Dazu wird das Bauelement auf ein dreidimensionales, inhomogenes Gitter abgebildet, wobei jedem Gitterpunkt eine Materialnummer zugeordnet wird. In der Eingabedatei müssen hierzu Materialregionen spezifiziert werden, für die standardmäßig Geometrieobjekte wie Kugeln, Ellipsoide, Obelisk, Quader usw. zur Verfügung stehen, welche zudem beliebig kombiniert werden können. In einer separaten Datenbasis stehen für jedes

Material alle benötigten Parameter zur Verfügung, wie Bandenergien, effektive Massen, dielektrische, piezo- und pyroelektrische Konstanten, Deformationspotenziale, kp-Parameter, usw. Neben Silizium und Germanium und den typischen binären Halbleiterverbindungen wie GaAs, InAs, AlAs, GaN, AlN, InN sind außerdem Legierungen wie $\text{Al}_x\text{Ga}_{1-x}\text{As}$, $\text{Al}_x\text{Ga}_{1-x}\text{N}$ oder $\text{Si}_x\text{Ge}_{1-x}$ zugelassen. Das Legierungsprofil wird hierbei durch die ortsabhängige Legierungskonzentration x bestimmt. Im Eingabesystem sind dafür Standardfunktionen wie lineare oder gaußförmige Verteilungen, die beliebig im Raum gedreht werden können, vorgesehen. Materialparameter von ternären Legierungshalbleitern werden mittels linearer oder auch nichtlinearer Interpolation berechnet. Daraus kann man ersehen, dass im Prinzip für jeden Gitterpunkt ein eigener vollständiger Datensatz ermittelt und abgespeichert werden muss.

Simulationsgitter

Die Gitterpunkte des Simulationsgitters befinden sich zwischen den Punkten des Materialgitters. Daraus ergibt sich, dass sich die Punkte des Simulationsgitters auf den Grenzflächen befinden. Da viele physikalische Größen, wie z.B. die Ladungsträgerdichte, an Heteroübergängen unstetig sind, werden an Grenzflächen ortsgleiche Gitterpunkte eingefügt, welche die entsprechenden links- und rechtsseitigen Grenzwerte repräsentieren. Den so definierten Hyperflächen wird zusätzliche physikalische Information zugewiesen wie Oberflächenladungen oder Oberflächenzustände. Dies kann zum einen über das Eingabesystem erfolgen, zum anderen geschieht dies automatisch, z.B. bei Grenzflächenladungen aufgrund der pyro- und piezoelektrischen Polarisationsfelder.

Quantenregionen

Quantenregionen sind Gebiete, in denen die Dichte quantenmechanisch berechnet wird. Diese können mit oben genannten Geometrieobjekten eingegeben werden und sind völlig unabhängig von den Materialregionen definiert. Dabei sind beliebig viele Quantenregionen zugelassen, die sich jedoch nicht überlappen dürfen. Grob gesagt wird das Bauelement in eine klassische Region und eine bestimmte Zahl disjunkter Quantenregionen partitioniert. Quantenmechanische Größen wie Wellenfunktionen sind damit nur in der jeweiligen Quantenregion definiert. Infolgedessen müssen intern Funktionen bereitgestellt werden, welche einen raschen Austausch von Daten zwischen dem globalen Simulationsgitter und den Teilgittern in den Quantenregionen erlauben.

Kontakte

Die Randbedingungen werden über Kontaktregionen definiert, welche ebenfalls beliebige Form haben dürfen. Gegenwärtig sind spannungskontrollierte ohmsche

Kontakte und spannungskontrollierte Schottky-Kontakte implementiert. Daneben lassen sich auch mathematische Randbedingungen vom Typ Dirichlet und von Neumann direkt angeben. Nichtverschwindende elektrische Felder werden über äquivalente Oberflächenladungen spezifiziert.

Dotierung

Als Dotierung sind grundsätzlich Donatoren und Akzeptoren möglich. Dazu müssen in der Eingabedatei die Dotierkonzentration, das Dotierprofil sowie die Ionisierungsenergien angegeben werden. Das Dotierprofil kann linear oder gaußförmig sein. Es sind beliebig viele Dotierregionen erlaubt, die sich auch überlappen dürfen.

Modelle für die elektronische Struktur

Grundsätzlich können mehrere Minima des Leitungs- und Valenzbandes berücksichtigt werden. Standardmäßig sind dies das Γ -, X- und L-Tal im Leitungsband sowie das schwere, leichte und abgespaltene Lochband im Valenzband. In klassischen Regionen wird die Dichte mithilfe der Thomas-Fermi-Näherung für alle Bänder unter Berücksichtigung eines Nichtparabolizitätsfaktors berechnet. In Quantenregionen kann der Anwender für jedes Bandminimum ein eigenes Modell spezifizieren. Zur Auswahl stehen die Thomas-Fermi-Näherung, die Einband-Schrödingergleichung, die Sechsband-kp-Schrödingergleichung für das schwere, leichte und abgespaltene Lochband sowie die Achtband-kp-Schrödingergleichung für das schwere, leichte und abgespaltene Lochband und das Γ -Tal. Dabei führen Verpannungs- und Anisotropieeffekte zu einer erheblichen Komplizierung. Während diese Effekte in der Achtband-kp-Schrödingergleichung direkt enthalten sind, müssen sie bei der Behandlung der X- und L-Täler in Einband-Näherung noch berücksichtigt werden. Zunächst einmal resultiert aus der mechanischen Verpannung ein energetisches Aufspalten der X- und L-Täler, welches durch die Deformationspotenziale beschrieben wird. Dies macht es notwendig, für die verschiedenen Energieniveaus eigene Schrödingergleichungen zu lösen. Gleichzeitig können zwei X- und L-Täler mit gleicher energetischer Lage unterschiedliche Orientierung der Hauptachsen des effektiven-Massentensors haben, was nochmals eine getrennte Behandlung erfordert.

Weiterhin wird zugelassen, nur die Energie-Eigenwerte bis zu einer bestimmten Energie zu berücksichtigen und die darüber liegenden Zustände in Thomas-Fermi-Näherung zu behandeln. Ebenso ist ein Modell implementiert, welches die gebundenen Zustände herausfiltert und nur diese quantenmechanisch berücksichtigt.

Transportmodell

Wie schon erwähnt, baut unser Simulationsverfahren auf der semiklassischen Transportgleichung auf. Zusätzlich kann die ballistische Transmissionswahrscheinlich-

keit aus den stationären Zuständen berechnet werden. Die Stromgleichungen für die Elektronen und die Löcher sind über Rekombinationsterme gekoppelt. Die wichtigsten Mechanismen sind Auger-Rekombination, Shockley-Read-Hall-Rekombination und direkte Rekombination, welche in der Eingabedatei spezifiziert werden müssen.

Kapitel 2

Bandstruktur

In nanostrukturierten Bauelementen wie Quantenpunkten und Nano-Transistoren können Quantisierungseffekte nicht mehr vernachlässigt werden. Um diese korrekt zu beschreiben, ist die Lösung der Schrödingergleichung notwendig. Die elektronische Struktur berücksichtigen wir im Rahmen der Mehrband-kp-Theorie. In Heterostrukturen kann es zudem aufgrund unterschiedlicher Gitterkonstanten zu mechanischen Verspannungen kommen. Die dadurch bedingten Potenzialverschiebungen müssen in den Strukturrechnungen berücksichtigt werden. Da wir auch Situationen untersuchen wollen, in denen die Koordinatenachsen, entlang derer diskretisiert wird, nicht mit den Kristallachsen übereinstimmen, musste eine Rotation des kp-Hamiltonoperators auf beliebige Kristallrichtungen durchgeführt werden. Dies hat eine erhebliche Komplizierung des Problems zur Folge. In der Tat ist die Darstellung derart komplex, dass man keine analytische Darstellung des Hamiltonoperators mehr hinschreiben kann, sondern beim Aufstellen der Matrix in mehreren Schritten vorgehen muss. Die Diskretisierung erfolgt auf einem inhomogenen Gitter. Dabei ist zu beachten, dass die resultierende Hamiltonmatrix hermitesch sein muss, um Stromerhaltung zu gewährleisten. Darüber hinaus wollen wir in Einzelfällen die elektronische Struktur in Halbleiter-Nanostrukturen unter angelegten Magnetfeldern berechnen. Da dies im Rahmen der kp-Theorie zu erheblichen zusätzlichen Problemen führt, wurden zunächst Magnetfelder in Einband-Näherung berücksichtigt. Die Hauptschwierigkeit war dabei, eine Diskretisierung für dreidimensionale Geometrien auf inhomogenem Gitter mit ortsabhängigen Materialkonstanten zu finden, die den Hamiltonoperator invariant unter Eichtransformationen lässt. Die zu diagonalisierenden Matrizen sind dünn besiedelt und haben in drei Dimensionen typischerweise eine Dimension von $10^6 \times 10^6$. Zu ihrer Lösung verwenden wir iterative Methoden. Dabei hat sich das Jacobi-Davidson-Verfahren am besten bewährt.

Das Kapitel gliedert sich folgendermaßen. Zunächst werden die Grundlagen der kp-Theorie einschließlich mechanischer Verspannungen erläutert. Anschließend geben wir die konkrete Form des von uns verwendeten Hamiltonoperators an und gehen daraufhin auf die Diskretisierung ein. Am Ende zeigen wir, wie Magnet-

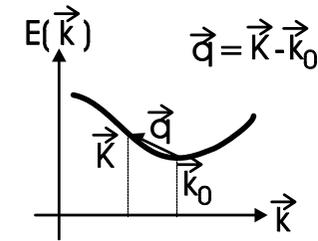


Abbildung 2.1: Energiedispersion mit einem Minimum im Punkt \vec{k}_0 .

felder in Einband-Näherung behandelt werden können.

2.1 Grundlagen

Im Folgenden sollen die Grundaussagen der kp-Theorie kurz zusammengefasst werden. Zunächst geben wir eine Darstellung des nichtrelativistischen Hamiltonoperators für die Einhüllenden-Wellenfunktionen in einem äußeren Feld an. Anschließend wird der Einfluss von relativistischen Korrekturen sowie elastischen Verspannungen erläutert.

Blochfunktionen

Die Blochfunktionen $\Psi_{n\vec{k}}(\vec{x})$ sind die exakten Lösungen der Gleichung

$$\mathbf{H}_0 \Psi = E \Psi \quad (2.1)$$

mit

$$\mathbf{H}_0 = -\frac{\hbar^2}{2m} \nabla^2 + V_0(\vec{x}) \quad (2.2)$$

mit dem gitterperiodischen Potenzial $V_0(\vec{x})$, den Eigenenergien $E_n(\vec{k})$ und der freien Elektronenmasse m . Aufgrund der Periodizität können die Blochfunktionen geschrieben werden als

$$\Psi_{n\vec{k}}(\vec{x}) = e^{i\vec{k}\vec{x}} u_{n\vec{k}}(\vec{x}),$$

wo $u_{n\vec{k}}(\vec{x})$ die periodischen Blochfaktoren sind [21].

Für die Dispersion in Abb. 2.1 mit dem Energieminimum im Punkt \vec{k}_0 kann

man folgende Luttinger-Kohn-Basis definieren ($\frac{1}{\sqrt{V}}$ als Normierungsfaktor):

$$\varphi_{n\vec{q}} = \frac{1}{\sqrt{V}} \Psi_{n\vec{k}_0} e^{i\vec{q}\vec{x}} \quad (2.3)$$

$$\langle \varphi_{n\vec{q}}, \varphi_{n'\vec{q}} \rangle = \delta_{nn'} \delta_{\vec{q}\vec{q}} \quad (2.4)$$

$$\langle \varphi_{n\vec{q}}, \vec{p} \varphi_{n'\vec{q}} \rangle = \vec{p}_{nn'} \delta_{\vec{q}\vec{q}} \quad (2.5)$$

$$\vec{p}_{nn'} \equiv \langle \Psi_{n\vec{k}_0}, \vec{p} \Psi_{n'\vec{k}_0} \rangle \quad (2.6)$$

Externes Feld

In einem äußeren Feld ist der Hamiltonoperator folgendermaßen definiert:

$$\mathbf{H} = \mathbf{H}_0 \left(\vec{p} + \frac{e}{c} \vec{A}(\vec{x}) \right) + U(\vec{x}), \quad (2.7)$$

wobei

$$\mathbf{H}_0(\vec{p}) = \frac{\vec{p}^2}{2m} + V_0(\vec{x})$$

$$\vec{E} = -\vec{\nabla}U(\vec{x}) + \frac{1}{c} \frac{\partial \vec{A}(\vec{x})}{\partial t}$$

Dabei bezeichnet $U(\vec{x})$ das äußere elektrische Potenzial und $\vec{A}(\vec{x})$ das Vektorpotenzial. V_0 ist das gitterperiodische Potenzial aus Gl. (2.2). Die Schrödingergleichung lautet:

$$\mathbf{H}\Psi(\vec{x}) = i\hbar \frac{\partial \Psi(\vec{x})}{\partial t} \quad (2.8)$$

Die Lösungen $\Psi(\vec{x})$ werden nach den Blochfunktionen $\Psi_{n\vec{k}_0}$ des ungestörten Hamiltonoperators \mathbf{H}_0 im Punkt \vec{k}_0 entwickelt:

$$\Psi(\vec{x}) = \sum_n F_n(\vec{x}) \Psi_{n\vec{k}_0}, \quad (2.9)$$

mit den im Vergleich zur Gitterkonstante langsam variierenden Einhüllenden-Wellenfunktionen $F_n(\vec{x})$. Dabei bezeichnet n den Bandindex.

Die Einhüllenden-Wellenfunktionen werden als Fourier-Reihe geschrieben mit den Fourier-Koeffizienten $c_{n\vec{k}} = \int F_n(\vec{x}) e^{-i\vec{k}\vec{x}} d\vec{x}$:

$$F_n(\vec{x}) = \sum_{\vec{k}} c_{n\vec{k}} \cdot e^{i\vec{k}\vec{x}}. \quad (2.10)$$

Diese Formel wird in Gl. (2.8) eingesetzt und man erhält nun folgendes Gleichungssystem:

$$\sum_n \sum_k \left[(H_1 + H_2)_{n'\vec{k}', n\vec{k}} + E_n \delta_{nn'} \delta_{\vec{k}\vec{k}'} \right] c_{n\vec{k}} = i\hbar \frac{\partial c_{n'\vec{k}'}}{\partial t} \quad (2.11)$$

Dabei gilt:

$$(H_0)_{n'\vec{k}', n\vec{k}} = E_n \delta_{nn'} \delta_{\vec{k}\vec{k}'}, \quad E_n = E_n(\vec{k}_0)$$

$$(H_1)_{n'\vec{k}', n\vec{k}} = \frac{1}{2m} \left[\hbar^2 \vec{k}^2 \delta_{\vec{k}\vec{k}'} + \frac{e\hbar}{c} (\vec{k}' + \vec{k}) \cdot \vec{A}_{\vec{k}-\vec{k}'} + \left(\frac{e}{c} \right)^2 A_{\vec{k}-\vec{k}'}^{(2)} \right] \delta_{nn'} + U_{\vec{k}'-\vec{k}} \delta_{nn'}$$

$$(H_2)_{n'\vec{k}', n\vec{k}} = \frac{1}{m} \left(\hbar \vec{k} \delta_{\vec{k}\vec{k}'} + \frac{e}{c} \vec{A}_{\vec{k}'-\vec{k}} \right) p_{n'n}$$

Die Potenziale lauten in Fourier-Darstellung:

$$\vec{A}(\vec{x}) = \sum_{\vec{q}} \vec{A}_{\vec{q}} e^{i\vec{q}\vec{x}}, \quad \vec{A}^2(\vec{x}) = \sum_{\vec{q}} A_{\vec{q}}^{(2)} e^{i\vec{q}\vec{x}}, \quad U(\vec{x}) = \sum_{\vec{q}} U_{\vec{q}} e^{i\vec{q}\vec{x}} \quad (2.12)$$

Bisher wurde noch keine störungstheoretische Näherung gemacht, sondern nur der Hartree-Hamiltonoperator in die Luttinger-Kohn-Basis umgeschrieben. Man fasst nun die Bänder in zwei Gruppen zusammen. Die energetisch benachbarten Bänder in Gruppe A werden weiterhin exakt behandelt, während die restlichen Bänder der Gruppe B störungstheoretisch betrachtet werden. Dann ergibt sich mit $m, m' \in A$ und $s \in B$:

$$\sum_{m\vec{k}} H_{m'\vec{k}', m\vec{k}} \cdot \bar{c}_{m\vec{k}} = i\hbar \frac{\partial}{\partial t} \bar{c}_{m'\vec{k}'}, \quad \bar{c}_{m\vec{k}} = c_{m\vec{k}} \cdot e^{-iE_m t/\hbar} \quad (2.13)$$

$$H_{m'\vec{k}', m\vec{k}} = E_m \delta_{mm'} \delta_{\vec{k}\vec{k}'} + (H_1)_{m'\vec{k}', m\vec{k}} + (H_2)_{m'\vec{k}', m\vec{k}} + \sum_{s\vec{k}'} \frac{(H_2)_{m'\vec{k}', s\vec{k}'} \cdot (H_2)_{s\vec{k}', m\vec{k}}}{E_m - E_s}$$

mit

$$(H_0)_{n'\vec{k}', n\vec{k}} = E_n \delta_{nn'} \delta_{\vec{k}\vec{k}'}, \quad E_n = E_n(\vec{k}_0)$$

$$(H_1)_{n'\vec{k}', n\vec{k}} = \frac{1}{2m} \left[\hbar^2 \vec{k}^2 \delta_{\vec{k}\vec{k}'} + \frac{e\hbar}{c} (\vec{k}' + \vec{k}) \cdot \vec{A}_{\vec{k}-\vec{k}'} + \left(\frac{e}{c} \right)^2 A_{\vec{k}-\vec{k}'}^{(2)} \right] \delta_{nn'} + U_{\vec{k}'-\vec{k}} \delta_{nn'}$$

$$(H_2)_{n'\vec{k}', n\vec{k}} = \frac{1}{m} \left(\hbar \vec{k} \delta_{\vec{k}\vec{k}'} + \frac{e}{c} \vec{A}_{\vec{k}'-\vec{k}} \right) p_{n'n}$$

Geht man nun in die Ortsraumdarstellung zurück, erhält man schließlich:

$$\sum_m H_{m'm} F_m(\vec{x}, t) = i\hbar \frac{\partial}{\partial t} F_{m'}(\vec{x}, t) \quad (2.14)$$

mit

$$\begin{aligned}
H_{m'm} &= E_m \delta_{mm'} + \frac{\hbar^2}{2m} \vec{k}^2 \delta_{mm'} + \frac{\hbar}{m} \vec{K} \vec{p}_{m'm} + \\
&+ \frac{\hbar^2}{m^2} \sum_{s \in \text{Klasse B}} \frac{\left(\vec{K} \vec{p}_{m's} \right) \left(\vec{K} \vec{p}_{sm} \right)}{E_m - E_s} + U(\vec{x}) \delta_{mm'} \quad , \quad m, m' \in \text{Klasse A} \\
\vec{K} &= -i \vec{\nabla} + \frac{e}{\hbar c} \vec{A}(\vec{x})
\end{aligned} \tag{2.15}$$

Dabei bezeichnet A die Klasse der Bänder, welche exakt behandelt werden, und B die Klasse der Fernbänder, über die summiert wird.

Man beachte, dass die Lage des Bandminimums \vec{k}_0 nicht in die Darstellung des Hamiltonoperators für die Einhüllenden-Funktionen eingeht. Lediglich die Basiszustände (Blochzustände) hängen von \vec{k}_0 ab.

Übergang zum Volumenhalbleiter (engl. bulk)

Die entsprechende Darstellung im Volumenhalbleiter (engl. bulk), d.h. für verschwindende äußere Felder ($U(\vec{x}) = 0, \vec{A}(\vec{x}) = 0$), ergibt sich, indem man für die Einhüllenden-Funktionen ebene Wellen ansetzt mit dem Wellenvektor \vec{k} :

$$F_m(\vec{x}) = \chi_m e^{i\vec{k}\vec{x}}, \quad m \in \text{Klasse A}$$

Damit folgt sofort aus Gl. (2.15):

$$\begin{aligned}
H_{m'm} &= E_m \delta_{mm'} + \frac{\hbar^2}{2m} \vec{k}^2 \delta_{mm'} + \frac{\hbar}{m} \vec{k} \vec{p}_{m'm} + \\
&+ \frac{\hbar^2}{m^2} \sum_{s \in \text{Klasse B}} \frac{\left(\vec{k} \vec{p}_{m's} \right) \left(\vec{k} \vec{p}_{sm} \right)}{E_m - E_s} + U(\vec{x}) \delta_{mm'} \quad , \quad m, m' \in \text{Klasse A}
\end{aligned} \tag{2.16}$$

Oft hat man Heterostrukturen, die in (x, y) -Richtung homogen sind. Dann handelt es sich faktisch um ein eindimensionales Problem im Ortsraum. Dementsprechend besitzt die Einhüllenden-Wellenfunktion nur eine z -Abhängigkeit, die Dispersion in (x, y) -Richtung wird wieder durch ebene Wellen beschrieben:

$$F_m(\vec{x}) = \chi_m(z) e^{i(k_x x + k_y y)}, \quad m \in \text{Klasse A}$$

Den kp -Hamiltonoperator erhält man nun aus Gl. (2.15) durch die Substitution $\vec{K} = -i \vec{\nabla} \rightarrow (k_x, k_y, -i \partial_z)^T$.

Spin-Bahn-Wechselwirkung

In der bisherigen Darstellung wurden relativistische Effekte nicht berücksichtigt. Ausgangspunkt dafür ist die Schrödingergleichung in zweikomponentiger Form,

welche man als nichtrelativistischen Grenzfall aus der Dirac-Gleichung erhält [22]:

$$\left[\frac{p^2}{2m} + V_0(\mathbf{r}) + \frac{\hbar}{4m^2 c^2} \left(\vec{\nabla} V_0 \times \vec{p} \right) \cdot \vec{\sigma} + \frac{\hbar^2}{8m^2 c^2} \vec{\nabla}^2 V_0 - \frac{p^4}{8m^3 c^2} \right] \Psi(\mathbf{r}) = \epsilon \Psi(\mathbf{r}) \tag{2.17}$$

Dabei bezeichnet m die freie Elektronenmasse und $V_0(\mathbf{r})$ das Kristallpotenzial. Der dritte, vierte, und fünfte Term in Gl. (2.17) stellen relativistische Korrekturen dar. Im Einzelnen ist $\frac{\hbar}{4m^2 c^2} \left(\vec{\nabla} V_0 \times \vec{p} \right) \cdot \vec{\sigma}$ die Spin-Bahnkopplung, welche die charakteristischen Aufspaltungen der Valenzbandenergien verursacht. Die beiden letzten Terme stellen skalar-relativistische Korrekturen dar und sind nur für schwere Atome für die innersten Schalen von Bedeutung.

In den Basisfunktionen (2.3) muss nun der Spin mitberücksichtigt werden:

$$\varphi_{n\vec{q}\sigma} = \frac{1}{\sqrt{V}} \Psi_{n\vec{k}_0} e^{i\vec{q}\vec{x}} s_\sigma \tag{2.18}$$

mit

$$s_\sigma = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Im Weiteren werden die Spineigenzustände mit $|\uparrow\rangle$ und $|\downarrow\rangle$ („Spin up“ bzw. „Spin down“) bezeichnet. Die Spin-Bahn-Wechselwirkung führt zu einer Aufspaltung der Valenzbänder am Γ -Punkt (Abb. 2.2). Zusätzlich beeinflussen noch Verspannungseffekte die Bandstruktur. Darauf wird weiter unten eingegangen.

2.2 Darstellung des kp -Hamiltonoperators

Im Folgenden wird kurz die von uns implementierte Form der Hamiltonmatrix im erläutert. Zunächst soll dabei ein Volumenhalbleiter zugrunde gelegt werden, der Übergang zu Heterostrukturen wird im nächsten Abschnitt beschrieben. Dabei beschränken wir uns auf Zinkblende, die Darstellung für Wurtzit findet sich in Anhang C. Als Basiszustände wählen wir $|s \uparrow\rangle, |s \downarrow\rangle$ für das Leitungsband und $|x \uparrow\rangle, |y \uparrow\rangle, |z \uparrow\rangle, |x \downarrow\rangle, |y \downarrow\rangle, |z \downarrow\rangle$ für das Valenzband. Diese definieren somit die Klasse A in Gl. (2.15):

$$\text{Klasse A} = \{|s \uparrow\rangle, |s \downarrow\rangle, |x \uparrow\rangle, |y \uparrow\rangle, |z \uparrow\rangle, |x \downarrow\rangle, |y \downarrow\rangle, |z \downarrow\rangle\}$$

Klasse B umfasst alle restlichen Bänder. Der Hamiltonoperator (ohne Verspannungen und Spin-Bahn-Wechselwirkung) lautet dann wie folgt, wobei die x -Achse

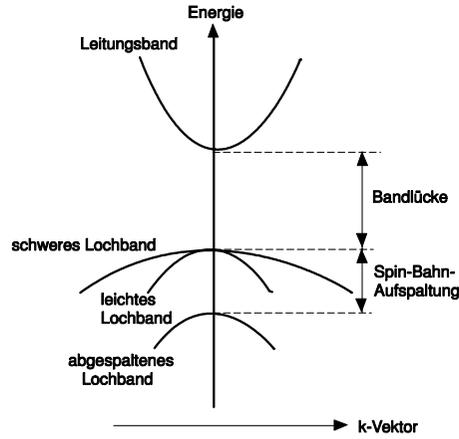


Abbildung 2.2: Bandstruktur von (3)-(5)-Halbleitern. Die Spin-Bahn-Wechselwirkung führt zur Spin-Bahn-Aufspaltung im Valenzband.

der [100]-, die y -Achse der [010]- und die z -Achse der [001]-Richtung entspricht.

	$s \uparrow$	$s \downarrow$	$x \uparrow$	$y \uparrow$	$z \uparrow$	$x \downarrow$	$y \downarrow$	$z \downarrow$
$s \uparrow$	H_{cc}		H_{cv}					
$s \downarrow$								
$x \uparrow$	T	H_{cv}	H_{vv}			0		
$y \uparrow$								
$z \uparrow$			0			H_{vv}		
$x \downarrow$								
$y \downarrow$	0			H_{vv}				
$z \downarrow$								

$$H_{vv} = \begin{pmatrix} E_v + \frac{\hbar^2}{2m} k^2 & & & & & & & & \\ & E_v + \frac{\hbar^2}{2m} k^2 & & & & & & & \\ & & E_v + \frac{\hbar^2}{2m} k^2 & & & & & & \\ Lk_x^2 + M(k_y^2 + k_z^2) & Nk_x k_y & & Nk_x k_z & & & & & \\ Nk_x k_y & Lk_y^2 + M(k_x^2 + k_z^2) & & Nk_y k_z & & & & & \\ Nk_x k_z & Nk_y k_z & & M(k_x^2 + k_y^2) + Lk_z^2 & & & & & \end{pmatrix} + \quad (2.19)$$

Dabei sind L, M, N die Dresselhausparameter [19]:

$$L = \frac{\hbar^2}{m^2} \sum_{n \in B} \frac{|\langle x | p_x | n \rangle|^2}{E_0 - E_n} \quad (2.20)$$

$$M = \frac{\hbar^2}{m^2} \sum_{n \in B} \frac{|\langle x | p_y | n \rangle|^2}{E_0 - E_n}$$

$$N = \frac{\hbar^2}{m^2} \sum_{n \in B} \frac{\langle x | p_x | n \rangle \langle n | p_y | y \rangle + \langle x | p_y | n \rangle \langle n | p_x | y \rangle}{E_0 - E_n}$$

E_c und E_v sind die Bandkanten des Leitungs- bzw. des Valenzbandes im Volumenhalbleiter. Weiter gilt:

$$H_{cc} = \begin{pmatrix} E_c & 0 \\ 0 & E_c \end{pmatrix} + \begin{pmatrix} H_{cc\uparrow} & 0 \\ 0 & H_{cc\downarrow} \end{pmatrix} \quad (2.21)$$

mit

$$H_{cc\uparrow} = H_{cc\downarrow} = \frac{\hbar^2}{2m} s(k_x^2 + k_y^2 + k_z^2) \quad (2.22)$$

Der Parameter s ist definiert als

$$s = \frac{1}{m} \sum_{n \in B} \frac{|\langle s | p_x | n \rangle|^2}{E_c - E_n} + \frac{1}{m},$$

wobei über alle Bänder außer die Valenzbänder summiert wird.

Die Kopplung zwischen Leitungs- und Valenzbändern lautet wie folgt:

$$H_{cv} = \begin{pmatrix} H_{cv}^{sx} & H_{cv}^{sy} & H_{cv}^{sz} & 0 & 0 & 0 \\ 0 & 0 & 0 & H_{cv}^{sx} & H_{cv}^{sy} & H_{cv}^{sz} \end{pmatrix} \quad (2.23)$$

$$H_{cv}^{sx} = Bk_y k_z + iPk_x$$

$$H_{cv}^{sy} = Bk_x k_z + iPk_y$$

$$H_{cv}^{sz} = Bk_x k_y + iPk_z \quad (2.24)$$

Der Parameter B tritt aufgrund fehlender Inversionssymmetrie auf:

$$B = \frac{2\hbar^2}{m^2} \sum_{n=1}^{\Gamma_{15}} \frac{\langle s | p_x | u_n \rangle \langle u_n | p_y | z \rangle}{\frac{E_c + E_v}{2} - E_n} \quad (2.25)$$

Die Summation läuft hier über alle Zustände $|u_n\rangle$, die gemäß der Darstellung Γ_{15} transformieren.

Für P gilt:

$$P = -\frac{i\hbar}{m} \langle s | p_x | x \rangle \quad (2.26)$$

Spin-Bahn-Wechselwirkung

Der Hamilton-Operator für die Spin-Bahn-Wechselwirkung lautet allgemein (vgl. Gl. (2.17))

$$H_{SO} = \frac{\hbar}{4m^2c^2} \left(\vec{\nabla}V \times \vec{p} \right) \cdot \vec{\sigma}, \quad (2.27a)$$

mit dem elektrischen Potenzial V , dem Impulsoperator \vec{p} und dem Vektor der Pauli-Matrizen $\vec{\sigma} = (\sigma_1, \sigma_2, \sigma_3)^T$. Er wirkt auf die zweikomponentigen Blochfunktionen, die den Spin beinhalten. Schließt man diese Wechselwirkung in die kp -Theorie ein, so erhält man folgende zwei Beiträge:

$$H_1 = \frac{\hbar}{4m^2c^2} \left(\vec{\nabla}V \times \vec{p} \right) \cdot \vec{\sigma} \quad (2.28)$$

$$H_2 = \frac{\hbar}{4m^2c^2} \left(\vec{\nabla}V \times \vec{k} \right) \cdot \vec{\sigma} \quad (2.29)$$

Üblicherweise wird angenommen:

$$\|H_2\| \ll \|H_1\| \quad (2.30)$$

Dann ist die Spin-Bahn-Wechselwirkung unabhängig vom \vec{k} -Vektor. Mit der Definition der Konstante

$$\Delta_0 = -\frac{3i\hbar}{4m_0^2c^2} \langle x | \left(\vec{\nabla}V \times \vec{p} \right)_y | z \rangle$$

lautet die Spin-Bahn-Wechselwirkung in der Basis $\{|x \uparrow\rangle, |y \uparrow\rangle, |z \uparrow\rangle, |x \downarrow\rangle, |y \downarrow\rangle, |z \downarrow\rangle\}$ [20]:

$$H_{SO} = \frac{\Delta_0}{3} \begin{pmatrix} 0 & -i & 0 & 0 & 0 & 1 \\ i & 0 & 0 & 0 & 0 & -i \\ 0 & 0 & 0 & -1 & i & 0 \\ 0 & 0 & -1 & 0 & i & 0 \\ 0 & 0 & -i & -i & 0 & 0 \\ 1 & i & 0 & 0 & 0 & 0 \end{pmatrix}$$

Oft wählt man von vornherein eine Basis, in der der Spin-Bahn-Wechselwirkungsoperator diagonal ist. Die entsprechenden Eigenzustände werden schweres Loch

(hh), leichtes Loch (lh) und abgespaltenes Loch (sh) genannt:

$$|u_{hh\uparrow}\rangle = \frac{1}{\sqrt{2}} |(x+iy) \uparrow\rangle \quad (2.31)$$

$$|u_{hh\downarrow}\rangle = \frac{i}{\sqrt{2}} |(x-iy) \downarrow\rangle \quad (2.32)$$

$$|u_{lh1}\rangle = \frac{i}{\sqrt{6}} [| (x+iy) \downarrow\rangle - |2z \uparrow\rangle] \quad (2.33)$$

$$|u_{lh2}\rangle = \frac{1}{\sqrt{6}} [| (x-iy) \uparrow\rangle + |2z \downarrow\rangle] \quad (2.34)$$

$$|u_{sh1}\rangle = \frac{1}{\sqrt{3}} [| (x+iy) \downarrow\rangle + |z \uparrow\rangle] \quad (2.35)$$

$$|u_{sh2}\rangle = \frac{i}{\sqrt{3}} [| -(x-iy) \uparrow\rangle + |z \downarrow\rangle] \quad (2.36)$$

Durch Projektion der gefundenen Lösung in der Basis $\{|x \uparrow\rangle, |y \uparrow\rangle, |z \uparrow\rangle, |x \downarrow\rangle, |y \downarrow\rangle, |z \downarrow\rangle\}$ auf obige Eigenzustände kann man feststellen, ob der jeweilige Spinor den Charakter eines schweren, leichten oder abgespaltenen Loches hat.

Verspannungen

Verspannungen werden durch den Verzerrungstensor beschrieben:

$$\epsilon_{ij} = \frac{1}{2} \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) \quad (2.37)$$

Dabei bezeichnet der ortsabhängige Vektor \vec{u} die Auslenkung eines Atoms aus der Gleichgewichtslage im unverspannten Zustand. Die Verspannungen werden im Rahmen der Elastizitätstheorie (vgl. Kapitel 4.3) berechnet. Im verspannten Kristall lautet der Hamiltonoperator nun:

$$\mathbf{H}(\epsilon) = \frac{p^2}{2m} + V_\epsilon(\vec{x}) + \frac{\hbar}{4m_0^2c^2} \left(\vec{\nabla}V_\epsilon \times \vec{p} \right) \cdot \vec{\sigma}, \quad (2.38)$$

wobei V_ϵ das Potenzial im verspannten Kristall bedeutet. Gl. (2.38) unterscheidet sich von der Gleichung für den unverspannten Zustand nur durch Substitution von V_0 durch V_ϵ :

$$\mathbf{H}_0 = \frac{p^2}{2m} + V_0(\vec{x}) + \frac{\hbar}{4m_0^2c^2} (\nabla V_0 \times \vec{p}) \cdot \vec{\sigma} \quad (2.39)$$

Wir betrachten den Einfluss der Verspannungen störungstheoretisch und beschränken uns außerdem auf lineare Terme. Die Basisvektoren \vec{a}_i' des Bravais-Gitters im verspannten Kristall ergeben sich durch Deformation der Basisvektoren \vec{a}_i im unverspannten Zustand:

$$\vec{a}_i' = (1 + \epsilon) \vec{a}_i \quad (2.40)$$

Damit kann man nun für die Potenziale schreiben

$$V_\epsilon [(1 + \epsilon) \vec{x}] - V_0(\vec{x}) = \sum_{ij} V_{ij}(\vec{x}) \epsilon_{ij} \quad (2.41)$$

mit

$$V_{ij}(\vec{x}) = \frac{1}{2 - \delta_{ij}} \lim_{\epsilon \rightarrow 0} \frac{V_\epsilon [(1 + \epsilon) \vec{x}] - V_0(\vec{x})}{\epsilon_{ij}} \quad (2.42)$$

Durch die Transformation (2.40) wird der Impulsoperator \vec{p}' zu

$$\vec{p}' = (1 + \epsilon) \vec{p}.$$

Der Hamiltonoperator $H'(\epsilon)$ im verspannten Zustand kann nun geschrieben werden als

$$H(\epsilon) = H_0 + H_\epsilon + H_{\epsilon \vec{k}} + H_{\epsilon so} \quad (2.43)$$

mit

$$\begin{aligned} H_\epsilon &= -\frac{(\vec{p}' \cdot \epsilon \vec{p}')}{m} + \mathbf{V} \epsilon \\ H_{\epsilon \vec{k}} &= -2 \frac{(\vec{p}' \cdot \epsilon \vec{k})}{m} \\ H_{\epsilon so} &= \frac{\hbar}{4m^2 c^2} \left\{ \left((\epsilon \vec{\nabla} V_0) \cdot (\vec{\sigma} \times \vec{p}') \right) - \left(\vec{\nabla}(\epsilon V_0) \cdot (\vec{\sigma} \times \vec{p}') \right) - \left((\vec{\sigma} \times \vec{\nabla} V_0) \cdot (\epsilon \vec{p}') \right) \right\} \end{aligned}$$

In Komponenten erhält man schließlich

$$H_\epsilon^{n'n} = \sum_{ij} \left(-\frac{(p_i p_j)_{n'n}}{m} + V_{n'n}^{ij} \right) \epsilon_{ij} = \sum_{ij} D_{n'n}^{ij} \epsilon_{ij} \quad (2.44)$$

$$H_{\epsilon \vec{k}}^{n'n} = -2 \frac{(\vec{p}'_{m'm} \epsilon \vec{k})}{m} + \frac{\hbar}{m} \sum_{s \in B} \frac{(\vec{k} \vec{p}'_{m's}) H_\epsilon^{sm} + H_\epsilon^{m's} (\vec{k} \vec{p}'_{sm})}{E_m - E_s}, \quad (2.45)$$

wo

$$D_{n'n}^{ij} = -\frac{(p_i p_j)_{n'n}}{m} + V_{n'n}^{ij} \quad (2.46)$$

den Deformationspotenzial-Operator darstellt. Im Weiteren berücksichtigen wir nur noch den Anteil $H_\epsilon^{n'n}$ in der kp-Matrix.

In unserer Basis $\{|x\rangle, |y\rangle, |z\rangle\}$ hat der Deformationspotenzial-Tensor drei nichtverschwindende Elemente:

$$l = D_{xx}^{XX}, \quad m = D_{yy}^{XX}, \quad n = D_{xy}^{XY} \quad (2.47)$$

Die Matrix \mathbf{H}_ϵ hat nun die Form:

$$\mathbf{H}_\epsilon = \begin{pmatrix} l\epsilon_{xx} + m(\epsilon_{yy} + \epsilon_{zz}) & n\epsilon_{xy} & n\epsilon_{xz} \\ n\epsilon_{xy} & m(\epsilon_{xx} + \epsilon_{zz}) + l\epsilon_{yy} & n\epsilon_{yz} \\ n\epsilon_{xz} & n\epsilon_{yz} & m(\epsilon_{xx} + \epsilon_{yy}) + l\epsilon_{zz} \end{pmatrix}$$

Die Parameter l, m, n hängen vom absoluten Deformationspotenzial a_v und den Scher-Deformationspotenzialen b und d ab via:

$$\begin{aligned} l &= a_v + 2b \\ m &= a_v - b \end{aligned} \quad (2.48)$$

$$n = \sqrt{3}d \quad (2.49)$$

Die Auswirkung des Verspannungstensors ϵ auf das Leitungsband wird durch das absolute Deformationspotenzial für das Leitungsband a_c beschrieben

$$E_c \rightarrow E_c + a_c Sp(\epsilon)$$

Eine Volumenänderung, die der Spur des Spannungstensors entspricht, führt über das absolute Deformationspotenzial zu einer Verschiebung des Schwerpunktes der Bänder. Zusätzlich haben die Außerdiagonalelemente des Spannungstensors über die Scher-Deformationspotenziale die Aufspaltung der Valenzband-Energien zur Folge.

Drehung des Koordinatensystems

Da die Kristallachsen nicht notwendig mit den Koordinatenachsen, d.h. der x -, y - und z -Achse des Simulationssystems, zusammenfallen, muss eine Drehung des Hamiltonoperators durchgeführt werden. Die Drehmatrix lautet allgemein

$$\mathbf{R} = \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{pmatrix}$$

wobei $\mathbf{R}^{-1} = \mathbf{R}^T$ gilt. Die Komponenten ergeben sich aus der Definition:

$$\vec{k}_{\text{Simulations-System}} = \mathbf{R} \vec{k}_{\text{Kristall-System}} \quad (2.50)$$

Zur Verdeutlichung ist in Abb. 2.3 ein Beispiel dargestellt. Material B ist auf Material A aufgewachsen mit Wachstumsrichtung [311]. Diese ist zugleich die z -Achse im Simulationssystem. Dann gilt konkret für die Drehmatrix:

$$\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}_{\text{Simulations-System}} = \mathbf{R} \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix}_{\text{Kristall-System}} \quad (2.51)$$

Man kann dies als passive Drehung interpretieren, d.h. der Vektor wird festgehalten und die Basis $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$ transformiert gemäß

$$\vec{e}_i^{\text{Simulations-System}} = \mathbf{R}^{-1} \vec{e}_i^{\text{Kristall-System}} \quad (2.52)$$

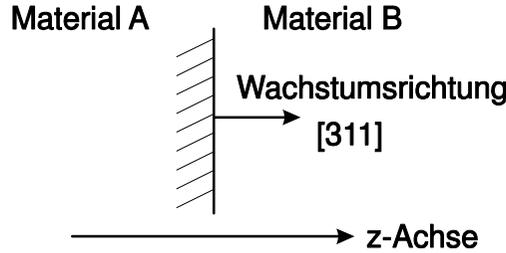


Abbildung 2.3: Einfaches Beispiel für eine Geometrie, in der die Kristallachsen nicht mit den Simulationsachsen zusammenfallen. Material B wird auf Material A aufgewachsen. Die Wachstumsrichtung ist die [311]-Richtung. Im Simulationssystem ist dies zugleich die z-Achse.

Transformation der Basisfunktionen

Für die Valenzbänder haben wir die Basiszustände $\{|x\rangle, |y\rangle, |z\rangle\}$ gewählt. Man kann sie schreiben als

$$f_i(\vec{r}) \equiv \vec{r} \cdot \vec{e}_i, \quad i = 1, 2, 3.$$

Ihr Transformationsverhalten unter der Drehung ergibt sich bei einer passiven Transformation, d.h. bei festgehaltenem Vektor und einer Basistransformation gemäß Gl. (2.52), aus

$$\mathcal{P}_{\mathbf{R}} f_i(\vec{r}) = f_i'(\vec{r}) = \vec{r} \cdot \mathbf{R}^{-1} \vec{e}_i = \vec{r} \cdot \sum_j R_{ij} \vec{e}_j = \sum_j R_{ij} \vec{r} \cdot \vec{e}_j = \sum_j R_{ij} f_j(\vec{r}). \quad (2.53)$$

Vektoriell bedeutet dies

$$\begin{pmatrix} |x\rangle' \\ |y\rangle' \\ |z\rangle' \end{pmatrix} = \mathbf{R} \begin{pmatrix} |x\rangle \\ |y\rangle \\ |z\rangle \end{pmatrix}. \quad (2.54)$$

Damit transformieren die Matrixelemente des Hamiltonoperators in der Basis $\{|x\rangle, |y\rangle, |z\rangle\}$ wie

$$H'(\vec{k}) = \mathbf{R} H(\vec{k}) \mathbf{R}^{-1} \quad (2.55)$$

$$H'_{ij}(\vec{k}) \equiv \langle f_i' | H(\vec{k}) | f_j' \rangle. \quad (2.56)$$

Zusätzlich gilt $\vec{k} = \mathbf{R}^{-1} \vec{k}'$, so dass sich letztlich ergibt:

$$H'(\vec{k}') = \mathbf{R} H(\mathbf{R}^{-1} \vec{k}') \mathbf{R}^{-1}. \quad (2.57)$$

Die s-Zustände sind isotrop und daher unter Drehungen invariant. Die Rotation aus Gl. (2.57) ist technisch sehr aufwändig und kann nicht mehr mit analytischen Formeln angegeben werden. Die Hamiltonmatrix wird in mehreren Schritten numerisch aufgesetzt.

2.3 Diskretisierung

Bisher wurde die kp-Matrix im Volumenhalbleiter diskutiert. Beim Übergang zu Heterostrukturen erhält man nun ortsabhängige Parameter im Hamilton-Operator. Dasselbe gilt für die Bandkanten E_c und E_v , welche zusätzlich das äußere elektrische Potenzial enthalten ($e = |e|$ Elementarladung):

$$\begin{aligned} E_c(\vec{r}) &= E_{c0}(\vec{r}) - e\Phi(\vec{r}) \\ E_v(\vec{r}) &= E_{v0}(\vec{r}) - e\Phi(\vec{r}) \end{aligned} \quad (2.58)$$

Dabei ist zu beachten, dass E_{c0} und E_{v0} die „natürlichen“ Bandenergien im van de Walle-Modell darstellen (vgl. Abschnitt 4.3, Anhang I).

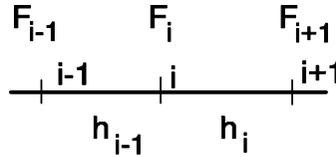
Bei der Diskretisierung des Hamiltonoperators müssen mehrere Aspekte beachtet werden. Zum einen muss an Heteroübergängen Stromerhaltung gelten. Dies erreicht man durch Symmetrisieren der Differenzialoperatoren. Zum anderen ist für uns ein inhomogenes Gitter unverzichtbar. Dadurch ist die diskretisierte Hamiltonmatrix nicht mehr hermitesch. Wir werden jedoch weiter unten zeigen, dass sich das Eigenwertproblem in ein hermitesches überführen lässt.

Ein weiteres Problem wirft die Drehung des Hamiltonoperators in beliebige Kristallrichtungen auf. Der rotierte Hamiltonoperator wird symbolisch, d.h. in folgender Form aufgesetzt:

$$\begin{aligned} H &= M_{11}(\vec{r}) k_x^2 + M_{22}(\vec{r}) k_y^2 + M_{33}(\vec{r}) k_z^2 + \\ &M_{12}(\vec{r}) k_x k_y + M_{13}(\vec{r}) k_x k_z + M_{23}(\vec{r}) k_y k_z + \\ &N_1(\vec{r}) k_x + N_2(\vec{r}) k_y + N_3(\vec{r}) k_z + C(\vec{r}) \end{aligned} \quad (2.59)$$

Dabei stellen $M_{11}, M_{22}, M_{33}, M_{12}, M_{13}, M_{23}, N_1, N_2, N_3$ und C jeweils komplexe 8×8 -Matrizen dar.

Bei der Diskretisierung der Differenziale macht man den Kohn-Luttinger-Ansatz. Dieser besteht in der Substitution $ik_\alpha \rightarrow \frac{\partial}{\partial x_\alpha}$. Der Operator $\frac{\partial}{\partial x_\alpha}$ wirkt auf die Einhüllenden-Funktionen. Um stromerhaltende Grenzflächenbedingungen für die Einhüllenden-Funktionen an Heteroübergängen zu gewährleisten, werden die Differenzialoperatoren symmetrisiert[26][27][28][29]. Mit ortsabhängigen Material-

Abbildung 2.4: Diskretisierung der Ableitungen $\frac{\partial}{\partial x_\mu}$ bzw. $\frac{\partial^2}{\partial x_\mu^2}$ (vgl. Text)

parametern ergibt sich:

$$ik_\mu \rightarrow \frac{\partial}{\partial x_\mu} \quad (2.60)$$

$$P \frac{\partial}{\partial x_\mu} \rightarrow \frac{1}{2} \left[P \frac{\partial}{\partial x_\mu} + \frac{\partial}{\partial x_\mu} P \right] \quad (2.61)$$

$$Q \frac{\partial}{\partial x_\mu} \frac{\partial}{\partial x_\nu} \rightarrow \frac{1}{2} \left[\frac{\partial}{\partial x_\mu} Q \frac{\partial}{\partial x_\nu} + \frac{\partial}{\partial x_\nu} Q \frac{\partial}{\partial x_\mu} \right] \quad (2.62)$$

Das Differenzial erster Ordnung k_μ

Zunächst sollen die Differenziale erster Ordnung k_μ behandelt werden. Um wieder eine hermitesche diskretisierte Matrix zu erhalten, müssen Differenziale erster Ordnung, welche in der oberen Dreiecksmatrix des Hamiltonoperators stehen, mit Vorwärtsdifferenzen approximiert werden. Die entsprechenden Differenziale in der unteren Dreiecksmatrix, welche zu denen in der oberen Dreiecksmatrix selbstadjungiert sind, werden mit Rückwärtsdifferenzen angenähert. Betrachten wir konkret die Situation in einer Dimension, wie sie in Abb. 2.4 für den Gitterpunkt i dargestellt ist. Dann gilt für die Vorwärtsdifferenz zunächst:

$$\begin{aligned} \left(P \frac{\partial}{\partial x_\mu} \right) F &= P_i \frac{F_{i+1} - F_i}{h_i} \\ \left(\frac{\partial}{\partial x_\mu} P \right) F &= \frac{P_{i+1} F_{i+1} - P_i F_i}{h_i} \end{aligned}$$

F bezeichnet dabei die Einhüllenden-Wellenfunktion. Für die ortsabhängigen Materialparameter P_i führen wir die Abkürzung ein:

$$P_{i+1/2} \equiv \frac{P_{i+1} + P_i}{2}. \quad (2.63)$$

Damit folgt aus Gl. (2.61)

$$P \frac{\partial}{\partial x} F \Big|_i \rightarrow \frac{P_{i+1/2} F_{i+1} - P_i F_i}{h_i} \quad (2.64)$$

Analog gilt für die Rückwärtsdifferenz:

$$P \frac{\partial}{\partial x} F \Big|_i \rightarrow \frac{P_i F_i - P_{i-1/2} F_{i-1}}{h_{i-1}} \quad (2.65)$$

Würde man die Differenziale aus den Gln. (2.64) und (2.65) direkt einsetzen, so wäre die Matrix für $h_i \neq h_{i-1}$ nicht hermitesch. Wir ersetzen also in Gl. (2.64) h_i und in Gl. (2.65) h_{i-1} mit $0.5 \cdot (h_{i-1} + h_i)$. Damit lautet die Diskretisierung für Differenziale erster Ordnung:

$$P \frac{\partial}{\partial x} F \Big|_i^{\text{vorwärts}} \rightarrow 2 \frac{P_{i+1/2} F_{i+1} - P_i F_i}{h_{i-1} + h_i} \quad (2.66)$$

$$P \frac{\partial}{\partial x} F \Big|_i^{\text{rückwärts}} \rightarrow 2 \frac{P_i F_i - P_{i-1/2} F_{i-1}}{h_{i-1} + h_i} \quad (2.67)$$

Diese Formeln gelten völlig analog auch in zwei und drei Dimensionen. Ein konkretes zweidimensionales Beispiel wird in Anhang B diskutiert.

Das Differenzial zweiter Ordnung k_μ^2

Für Ableitungen zweiter Ordnung gilt mit den Bezeichnungen in Abb. 2.4:

$$Q \frac{\partial}{\partial x} \frac{\partial}{\partial x} F \Big|_i = \frac{1}{\frac{1}{2}(h_{i-1} + h_i)} \left[Q_{i+1/2} \frac{F_{i+1} - F_i}{h_i} - Q_{i-1/2} \frac{F_i - F_{i-1}}{h_{i-1}} \right]$$

und weiter:

$$\begin{aligned} Q \frac{\partial^2}{\partial x^2} F \Big|_i &= \frac{2Q_{i-1/2}}{h_{i-1}(h_{i-1} + h_i)} F_{i-1} + \frac{2Q_{i+1/2}}{h_i(h_{i-1} + h_i)} F_{i+1} + \\ &\left[-\frac{2Q_{i-1/2}}{h_{i-1}(h_{i-1} + h_i)} - \frac{2Q_{i+1/2}}{h_i(h_{i-1} + h_i)} \right] F_i \end{aligned} \quad (2.68)$$

Gemischte Differenziale zweiter Ordnung $k_\mu k_\nu$

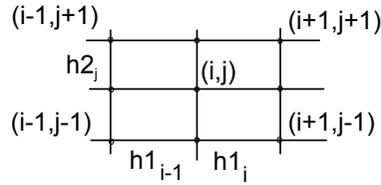
Diese treten nur in zwei und drei Dimensionen auf:

$$Q \partial_x \partial_y \rightarrow \frac{1}{2} (\partial_x Q \partial_y + \partial_y Q \partial_x)$$

Die Diskretisierung auf dem Gitter in Abb. 2.5 lautet wie folgt:

$$\frac{1}{2} (\partial_x Q \partial_y + \partial_y Q \partial_x) F = \quad (2.69)$$

$$\begin{aligned} &\frac{1}{2} \left[Q_{i+1,j} \frac{F_{i+1,j+1} - F_{i+1,j-1}}{h_{2j-1} + h_{2j}} - Q_{i-1,j} \frac{F_{i-1,j+1} - F_{i-1,j-1}}{h_{2j-1} + h_{2j}} \right] \frac{1}{h_{1i-1} + h_{1i}} + \\ &\frac{1}{2} \left[Q_{i,j+1} \frac{F_{i+1,j+1} - F_{i-1,j+1}}{h_{1i-1} + h_{1i}} - Q_{i,j-1} \frac{F_{i+1,j-1} - F_{i-1,j-1}}{h_{1i-1} + h_{1i}} \right] \frac{1}{h_{2j-1} + h_{2j}} = \end{aligned}$$

Abbildung 2.5: Diskretisierung gemischter Ableitungen $\frac{\partial}{\partial x_\mu} \frac{\partial}{\partial x_\nu}$ (vgl. Text)

$$\frac{1}{(h1_{i-1} + h1_i)(h2_{j-1} + h2_j)} \left(\frac{Q_{i-1,j} + Q_{i,j-1}}{2} F_{i-1,j-1} + \frac{-Q_{i+1,j} - Q_{i,j-1}}{2} F_{i+1,j-1} + \frac{-Q_{i-1,j} - Q_{i,j+1}}{2} F_{i-1,j+1} + \frac{Q_{i,j+1} + Q_{i+1,j}}{2} F_{i+1,j+1} \right)$$

Hermitizität

Die diskretisierte Matrix kann in ein Produkt einer hermiteschen Matrix $\tilde{\mathbf{H}}$ und einer positiv definiten Diagonalmatrix zerlegt werden: $\mathbf{H} = \mathbf{D}\tilde{\mathbf{H}}$

Damit lässt sich das Eigenwertproblem in ein hermitesches überführen mittels

$$\mathbf{D}^{-1/2} \mathbf{D} \tilde{\mathbf{H}} \mathbf{D}^{1/2} \mathbf{D}^{-1/2} F = \mathbf{D}^{-1/2} E F \longleftrightarrow \quad (2.70)$$

$$\mathbf{D}^{1/2} \tilde{\mathbf{H}} \mathbf{D}^{1/2} \Phi = E \Phi$$

mit

$$\Phi = \mathbf{D}^{-1/2} F$$

Man kann das Problem auch als verallgemeinerte Eigenwertgleichung schreiben:

$$\tilde{\mathbf{H}} F = \mathbf{D}^{-1} E F \quad (2.71)$$

In Anhang E wird konkret das Aufstellen der kp-Matrix anhand eines zweidimensionalen Beispiels dargestellt.

2.4 Einband-Schrödingergleichung

Die Einband-Schrödingergleichung ergibt sich aus Gl. (2.14), wenn man nur ein Band betrachtet. Gl. (2.14) gilt auch für Bandextrema, die sich nicht am Γ -Punkt befinden. Für das Leitungsband hat man:

$$-\frac{\hbar^2}{2} \vec{\nabla} \left[\frac{1}{m_{eff}(\vec{r})} \vec{\nabla} \right] \Psi(\vec{r}) + E_c(\vec{r}) \Psi(\vec{r}) = E \Psi(\vec{r}) \quad (2.72)$$

Für Löcher gilt:

$$+\frac{\hbar^2}{2} \vec{\nabla} \left[\frac{1}{m_{eff}(\vec{r})} \vec{\nabla} \right] \Psi(\vec{r}) + E_v(\vec{r}) \Psi(\vec{r}) = E \Psi(\vec{r}) \quad (2.73)$$

Die effektive Masse ist dabei richtungsabhängig. Man kann sie aus den Luttingerparametern bestimmen. Die Bandkanten $E_c(\vec{r})$ und $E_v(\vec{r})$ enthalten das elektrische Potenzial ($e = |e|$ Elementarladung):

$$\begin{aligned} E_c(\vec{r}) &= E_{c0}(\vec{r}) - e\Phi(\vec{r}) \\ E_v(\vec{r}) &= E_{v0}(\vec{r}) - e\Phi(\vec{r}) \end{aligned} \quad (2.74)$$

Es ist zu beachten, dass hier - anders als beim kp-Hamiltonoperator - die verspannungsinduzierte Verschiebung der Bandenergien in $E_{c0}(\vec{r})$ und $E_{v0}(\vec{r})$ enthalten ist (siehe auch Abschnitt 4.3). Diese wurde für die Valenzbänder und den Γ -Punkt im Leitungsband bereits oben dargestellt. Als weitere Leitungsbandminima berücksichtigen wir das X- und das L-Tal, welche im k -Raum des kubischen Kristalls acht- bzw. sechsfach entartet sind. Diese Entartung wird durch eine Scherspannung aufgehoben. Der entsprechende Korrekturterm zum Hamiltonoperator lautet [23]:

$$H_\epsilon = \Xi_d S p(\epsilon) + \Xi_u \left(\vec{k}_\epsilon \cdot \vec{\epsilon} k_\epsilon \right).$$

Dabei bezeichnet der Einheitsvektor \vec{k}_ϵ die Richtung des jeweiligen Minimums im k -Raum. Die Deformationspotenziale sind jeweils für das X- und das L-Tal definiert: $\Xi_d^{X/L}$ und $\Xi_u^{X/L}$.

2.5 Magnetfeld

In einem äußeren Magnetfeld wirft die Diskretisierung zusätzliche Probleme auf. Insbesondere führt eine direkte Diskretisierung des Differenzialoperators $k_\mu \rightarrow k_\mu - g A_\mu$ mit $g = \frac{q}{\hbar c}$ zur Brechung der Eichinvarianz des Hamiltonoperators. K. G. Wilson [30] hat diese Problemstellung im Rahmen der Gitter-Quantenfeldtheorie behandelt. Seine Ergebnisse wurden von M. Governale [31] aufgegriffen und auf einen zweidimensionalen Hamiltonoperator auf einem homogenen Rechtecksgitter mit homogenem Material im Ortsraum angewendet. Darauf aufbauend haben wir dieses Konzept auf drei Dimensionen und für inhomogene Gitter mit ortsabhängigen Materialkonstanten erweitert.

Die Einband-Schrödingergleichung für Elektronen lautet nun für beliebige Bandminima (vgl. Gl. (2.72) und Gl. (2.14)):

$$\mathbf{H} \Psi(\vec{r}) + E_c(\vec{r}) \Psi(\vec{r}) = E \Psi(\vec{r})$$

mit

$$\mathbf{H} = -\frac{\hbar^2}{2} \left(\vec{\nabla} + ig\vec{A} \right) \frac{1}{m_{eff}(\vec{r})} \left(\vec{\nabla} + ig\vec{A} \right) \quad (2.75)$$

Das Vektorpotenzial ergibt sich bei symmetrischer Eichung aus dem Magnetfeld durch:

$$\vec{A}(\vec{r}) = -\frac{1}{2} \vec{r} \times \vec{B} \quad (2.76)$$

Aus der Quantenmechanik [32] ist bekannt, dass unter einer Eichtransformation

$$\vec{A}(\vec{r}) \rightarrow \vec{A}(\vec{r}) + \vec{\nabla}\Lambda(\vec{r}) \quad (2.77)$$

die Wellenfunktionen wie

$$\Psi(\vec{r}) \rightarrow e^{ig\Lambda(\vec{r})} \Psi(\vec{r}) \quad (2.78)$$

transformieren. Für den Hamiltonoperator gilt entsprechend:

$$\mathbf{H} \rightarrow e^{ig\Lambda(\vec{r})} \mathbf{H} e^{-ig\Lambda(\vec{r})} \quad (2.79)$$

Man kann nun zeigen, dass folgende Diskretisierungsvorschriften bis auf einen Abbruchfehler der Ordnung $\frac{1}{6} (ghA^\mu)^3$ bei einem Gitterabstand h mit dem Ausdruck für die direkte Diskretisierung übereinstimmen:

$$Qk_x^2 \Psi|_{i,j} \rightarrow \frac{2}{h1_{i-1} + h1_i} \left[-\frac{Q_{i-1/2,j}}{h1_{i-1}} e^{igh1_{i-1}A_{i-1/2,j}^x} \Psi_{i-1,j} + \left(\frac{Q_{i-1/2,j}}{h1_{i-1}} + \frac{Q_{i+1/2,j}}{h1_i} \right) \Psi_{i,j} - \frac{Q_{i+1/2,j}}{h1_i} e^{-igh1_i A_{i+1/2,j}^x} \Psi_{i+1,j} \right] \quad (2.80)$$

$$Qk_y^2 \Psi|_{i,j} \rightarrow \frac{2}{h2_{j-1} + h2_j} \left[-\frac{Q_{i,j-1/2}}{h2_{j-1}} e^{igh2_{j-1}A_{i,j-1/2}^y} \Psi_{i,j-1} + \left(\frac{Q_{i,j-1/2}}{h2_{j-1}} + \frac{Q_{i,j+1/2}}{h2_j} \right) \Psi_{i,j} - \frac{Q_{i,j+1/2}}{h2_j} e^{-igh2_j A_{i,j+1/2}^y} \Psi_{i,j+1} \right]$$

Dabei wurde das Gitter in Abb. 2.6 zugrunde gelegt. Hinsichtlich der Indices gilt z.B.: $A_{i-1/2,j} = \frac{1}{2} (A_{i-1,j} + A_{i,j})$. Eine Eichtransformation

$$A_{i-1/2,j}^x \rightarrow A_{i-1/2,j}^x + \frac{1}{h1_{i-1}} (\Lambda_{i,j} - \Lambda_{i-1,j}) \quad (2.81)$$

$$A_{i,j-1/2}^y \rightarrow A_{i,j-1/2}^y + \frac{1}{h2_{j-1}} (\Lambda_{i,j} - \Lambda_{i,j-1})$$

führt z.B. für das Differenzial k_x^2 am Gitterpunkt (i,j) in der oberen Nebendiagonale zu einem zusätzlichen Faktor $e^{-igh1_i A_{i+1/2,j}^x} \rightarrow e^{ig\Lambda_{i,j}} e^{-ig\Lambda_{i+1,j}}$ und in der unteren Nebendiagonale zu einem Faktor $e^{igh1_{i-1} A_{i-1/2,j}^x} \rightarrow e^{ig\Lambda_{i,j}} e^{-ig\Lambda_{i-1,j}}$. Dies entspricht der Transformation in Gl. (2.79). Der Hamiltonoperator ist damit hermitesch und eichinvariant.

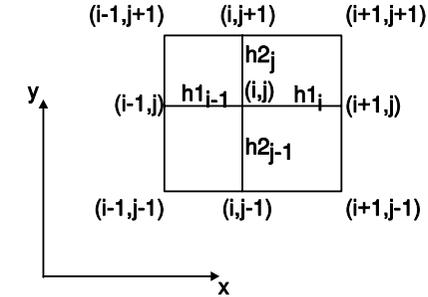


Abbildung 2.6: Zweidimensionales inhomogenes Gitter, auf dem der Hamiltonoperator mit Magnetfeld diskretisiert wird (vgl. Text). Das Vorgehen in drei Dimensionen ist analog.

2.6 Zusammenfassung

Die elektronische Struktur berechnen wir in effektiver-Massennäherung. Standardmäßig werden das schwere Loch, das leichte Loch und das abgespaltene Loch im Valenzband und das Γ -, X - und L -Tal im Leitungsband berücksichtigt. Diese Bänder kann man unabhängig voneinander in Einband-Näherung beschreiben. Im Rahmen der Mehrband-kp-Theorie werden die drei Valenzbänder an das Leitungsband mit Minimum im Γ -Punkt gekoppelt. Ein wesentlicher Aspekt ist hierbei, dass beliebige Kristallorientierungen zugelassen sind. Magnetfelder können in Einband-Näherung inkludiert werden. Die Diskretisierung des Hamiltonoperators haben wir exemplarisch für zweidimensionale inhomogene Gitter angegeben. Technische Details zum Aufstellen der kp-Matrix finden sich in Anhang E.

Kapitel 3

Elektronische Struktur im Nichtgleichgewicht

Im vorhergehenden Kapitel ging es um die Berechnung der elektronischen Struktur in nanostrukturierten Halbleiterbauelementen. Dabei wurde die Mehrbandk_p-Theorie zugrunde gelegt. Hauptanliegen dieser Arbeit ist es, eine Methode zu finden, die elektronische Struktur für Nichtgleichgewichts-Situationen zu beschreiben. Dazu muss die Lösung der Schrödingergleichung mit der Behandlung des Transportproblems verknüpft werden. In den letzten Jahren hat es dazu Arbeiten gegeben, die Nichtgleichgewichts-Greenfunktionen verwenden. Diese Methode konnte erfolgreich auf eindimensionale Quantenbauelemente angewendet werden [18]. Das Hauptproblem ist, dass dieses Verfahren selbst für eindimensionale Probleme einen horrenden Rechenaufwand erfordert, so dass an eine Erweiterung auf zwei oder drei Dimensionen vorerst nicht zu denken ist. Konzeptionell schwierig ist dabei vor allem die Berücksichtigung der diversen Streumechanismen, so dass hier viele Näherungen eingehen. Dies macht aber das physikalische Verhalten des Gesamtsystems schwer überschaubar. Angesichts dieser Schwierigkeiten haben wir uns an den - bisher gut verstandenen - semiklassischen Drift-Diffusionsgleichungen orientiert und ein Verfahren entwickelt, das es erlaubt, die quantenmechanische Dichteverteilung voll zu berücksichtigen. Die Nichtgleichgewichts-Situation beschreiben wir dabei durch das bewährte Konzept der nichtlokalen Quasi-Ferminiveaus. Die Beispiele am Ende dieser Arbeit, insbesondere die Berechnung von Quantenpunkten und Double-Gate-MOSFETs (Metal-Oxide-Semiconductor-Field-Effect-Transistor), zeigen, dass diese Methode auch in zwei und drei Dimensionen erfolgreich angewendet werden kann. Unser Ansatz verwendet die gleichen Basisfunktionen wie die Pauli-Master-Gleichung, die von Fischetti [17] anhand eindimensionaler Bauelemente untersucht wurde. In diesem Kapitel wird daher zunächst die Pauli-Master-Gleichung kurz erläutert und anschließend unser Verfahren vorgestellt. Schließlich sollen die Resultate von Fischetti mit unserer Methode verglichen werden. Dabei diskutieren wir auch die Grenzen des semiklassischen Ansatzes.

3.1 Pauli-Master-Gleichung

Die Anwendbarkeit der Pauli-Master-Gleichung wurde ausführlich von M. V. Fischetti [17][16] untersucht. Im Wesentlichen löst er die Schrödinger- und Poissongleichung und verwendet die so erhaltenen Eigenzustände als Basisfunktionen. Dadurch erhält er aus der Liouville-Gleichung die Pauli-Master-Gleichung, aus der die Besetzung der Zustände folgt. Die Übergangsraten werden mit Fermis Goldener Regel berechnet.

Die Rechtfertigung der semiklassischen Boltzmann-Transportgleichung beruht im Prinzip auch auf der Master-Gleichung, wobei hier als Basis ebene Wellen, d.h. die Eigenfunktionen des Hamiltonoperators eines freien Teilchens, benutzt werden und das äußere Potenzial als Störung betrachtet wird. Außerdem nimmt man an, dass durch die Streuungen die Phaseninformation herausgemittelt wird. Das elektrische Feld variiert dabei nur wenig im Vergleich zur mittleren freien Weglänge des Elektrons, so dass die Ebene-Wellen-Darstellung lokal gültig bleibt. Dies entspricht der Vorstellung eines punktförmigen Teilchens mit wohldefiniertem Impuls und Ort. Im Gegensatz dazu wird bei der Pauli-Master-Gleichung durch die Wahl der Wellennatur der Elektronen Rechnung getragen. Eine wesentliche Einschränkung bleibt jedoch, dass die Dichtematrix als diagonal angenommen wird.

Grundlagen der Pauli-Master-Gleichung

Seien $\Psi_n(\vec{r})$ eine Basis des Einteilchen-Hilbertraums, der das Bauelement beschreibt, und seien die Anfangszustände des N -Elektronensystems zur Zeit $t = 0$ gegeben durch $\Psi^n(\vec{r}) = \sum_{\mu} a_{\mu}^n(t) \Psi_{\mu}(\vec{r})$. Dann lautet die Dichtematrix

$$\rho_{\mu\nu} = \sum_{n=1}^N a_{\mu}^n(t) a_{\nu}^{n*}(t). \quad (3.1)$$

Die Bewegungsgleichung ist die Liouville-Gleichung:

$$\frac{\partial \rho}{\partial t} = \frac{i}{\hbar} [\rho, H] + \left(\frac{\partial \rho}{\partial t} \right)_{res} - \frac{\rho - \rho^{(eq)}}{\tau_s}. \quad (3.2)$$

Dabei beschreibt der Term $\left(\frac{\partial \rho}{\partial t} \right)_{res}$ die Wechselwirkung mit den Reservoirs (Kontakten), der Ausdruck $\frac{\rho - \rho^{(eq)}}{\tau_s}$ mit der Streuzzeit τ_s und der Gleichgewichtsverteilung $\rho^{(eq)}$ modelliert die Wechselwirkung aufgrund der Streuung im Bauelement [33]. Der Hamiltonoperator $H = H_0 + H_{int}$ besteht aus den Komponenten $H_0 = -\frac{\hbar^2}{2m} \nabla^2 - eV(\vec{r})$ ($V(\vec{r})$ Potential) und H_{int} , welcher die Streuung beschreibt. Die Herleitung der Pauli-Master-Gleichung erfolgt nun wie bei Kohn und Luttinger [34] mit dem Unterschied, dass als Basiszustände die Eigenfunktionen von

H_0 statt der ebenen Wellen verwendet werden. Es ergibt sich schließlich [17] für die Diagonalelemente des Dichteoperators $\rho_{\mu\vec{k}\sigma} \equiv \rho_{\mu\vec{k}\sigma, \mu\vec{k}\sigma}$:

$$\frac{\partial \rho_{\mu\vec{k}\sigma}}{\partial t} = \sum_{\nu\vec{k}'\sigma' \neq \mu\vec{k}\sigma} \left[W_{\mu\vec{k}\sigma, \nu\vec{k}'\sigma'} \rho_{\nu\vec{k}'\sigma'} - W_{\nu\vec{k}'\sigma', \mu\vec{k}\sigma} \rho_{\mu\vec{k}\sigma} \right] + \left(\frac{\partial \rho_{\mu\vec{k}\sigma}}{\partial t} \right)_{res} \quad (3.3)$$

$W_{\nu\vec{k}'\sigma', \mu\vec{k}\sigma}$ bezeichnet hier die Übergangsrate vom Zustand $|\mu\vec{k}\sigma\rangle$ in den Zustand $|\nu\vec{k}'\sigma'\rangle$. Diese lautet mit Fermis goldener Regel

$$W_{\nu\vec{k}'\sigma', \mu\vec{k}\sigma} = \frac{2\pi}{\hbar} \left| \langle \mu\vec{k}\sigma | H_{int} | \nu\vec{k}'\sigma' \rangle \right|^2 \delta \left[E_{\mu}(k) - E_{\nu}(k') \right] \quad (3.4)$$

Phänomenologisches Modell für Kontakte

Ein wichtiger Aspekt ist die Frage, inwieweit die Außerdiagonalelemente der Dichtematrix vernachlässigt werden dürfen. Van Hove [35] hat gezeigt, dass Gl. (3.3) den Übergang eines geschlossenen Systems (d.h. ohne den Term $\left(\frac{\partial \rho_{\mu\vec{k}\sigma}}{\partial t} \right)_{res}$) in den Gleichgewichtszustand korrekt beschreibt, vorausgesetzt der Anfangszustand ($t = 0$) ist „quasidiagonal“. Das bedeutet, dass zum Zeitpunkt $t = 0$ die Elemente $\rho_{\mu\vec{k}\sigma, \nu\vec{k}'\sigma'}$ verschwinden für $E_{\mu}(\vec{k}) - E_{\nu}(\vec{k}') > \delta$, wo $\frac{\delta}{\hbar}$ von der Größenordnung der Relaxationsrate $\frac{1}{\tau}$ ist. Fischetti argumentiert nun, dass Gl. (3.3) in gleicher Weise gilt, solange die Kontakte keine Außerdiagonalelemente injizieren, d.h. $\left(\frac{\partial \rho_{\mu\vec{k}\sigma, \nu\vec{k}'\sigma'}}{\partial t} \right)_{res} = 0$ für $E_{\mu}(\vec{k}) - E_{\nu}(\vec{k}') > \delta$. Dies trifft zu für delokalisierte Zustände, welche im Idealfall ebene Wellen sind. Daraus ergibt sich für die Kontakte folgendes phänomenologisches Modell: Man nimmt an, dass das Reservoir versucht, Ladungsneutralität in der Kontaktregion herzustellen, indem es Elektronen als ebene Wellen ins Bauelement injiziert. Dazu führt Fischetti für den Kontakt ein Ferminiveau ein, das die Besetzung der injizierten Zustände widerspiegelt. Zusätzlich verwendet er einen Korrekturparameter, mit dem die Fermi-Verteilung justiert wird, um insgesamt Ladungsneutralität und Stromerhaltung zu gewährleisten. Das Verfahren ist insgesamt recht kompliziert, insbesondere sind die Auswirkungen der Korrekturparameter auf den Strom schwer überschaubar. Interessant ist, dass er die Wellenfunktionen stationär mit gemischten Randbedingungen (vgl. nächster Abschnitt) berechnet, und diese stationären Lösungen schließlich in links- und rechtspropagierenden Zustände zerlegt. In Abb. 3.1 ist das Vorgehen Fischettis schematisch dargestellt.

Algorithmus

Zunächst wird für das Startpotenzial die Schrödingergleichung mit gemischten Randbedingungen gelöst. Daraus kann man gebundene sowie links- und rechtspropagierenden Zustände extrahieren. Anschließend werden die Übergangsraten be-

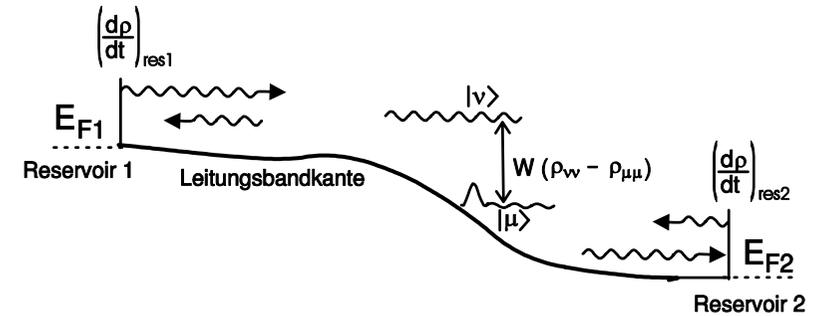


Abbildung 3.1: Schematische Darstellung des Vorgehens von Fischetti am Beispiel eines nin-Resistors. Die Dichtematrix wird als diagonal angenommen. Die Pauli-Master-Gleichung liefert die Besetzung der Zustände. An den Reservoiren wird ein phänomenologisches Kontaktmodell verwendet. Die Kontakte versuchen Ladungsneutralität herzustellen, indem sie delokalisierte Elektronen injizieren.

rechnet und die Besetzung der Zustände über die Pauli-Master-Gleichung bestimmt. Für die Wechselwirkung mit Kontakten wird ein phänomenologischer Ansatz für den Term $\left(\frac{\partial \rho_{\mu\vec{k}\sigma, \nu\vec{k}'\sigma'}}{\partial t} \right)_{res}$ gemacht. Dabei bekommen die Reservoire jeweils ein Quasi-Fermi-niveau zugewiesen, welches während des Selbstkonsistenzzyklus so angepasst werden muss, dass Ladungsneutralität in den Kontaktregionen und Stromerhaltung gewährleistet ist. Dann kann die Poisson-Gleichung neu gelöst werden usw. Die Besetzung der Zustände wird dabei durch eine diagonale Dichtematrix beschrieben.

3.2 Neues Transportmodell für dissipativen Transport

Die wichtigste Transportgleichung auf dem Gebiet der Bauelementphysik ist die Boltzmann-Gleichung. Sie ist Grundlage in allen Simulationsprogrammen, die realistische Bauelementstrukturen rechnen können. In den letzten Jahren hat es zunehmend Bemühungen gegeben, quantenmechanische Effekte störungstheoretisch zu integrieren. Zunächst sollen kurz die Grundideen und daraufhin die Erweiterung im Rahmen der Dichte-Gradienten-Methode beschrieben werden. Anschließend wird ein neuer Ansatz für Transport in Quantenbauelementen vorgestellt.

Boltzmann-Gleichung

Die Boltzmann-Gleichung ist eine Bilanzgleichung für die Zahl der Teilchen $f(\vec{r}, \vec{k}, t)$ $d^3\vec{r} d^3\vec{k}$, die sich zur Zeit t im Volumenelement $d^3\vec{r} d^3\vec{k}$ des Phasenraumes befinden [36]. Sie werden in das Volumenelement hinein- und herausgestreut und durch äußere Felder beschleunigt:

$$\begin{aligned} & \left(\frac{\partial}{\partial t} + \vec{v}_g \cdot \vec{\nabla}_r + \frac{q\vec{F}}{\hbar} \right) f(\vec{r}, \vec{k}, t) \\ &= \sum_{k'} \left\{ W_{kk'} f(\vec{r}, \vec{k}', t) [1 - f(\vec{r}, \vec{k}, t)] \right. \\ & \quad \left. - W_{k'k} [1 - f(\vec{r}, \vec{k}', t)] f(\vec{r}, \vec{k}, t) \right\} \end{aligned} \quad (3.5)$$

Dabei bezeichnen \vec{F} die elektrische Feldstärke, $\vec{v}_g = \frac{1}{\hbar} \vec{\nabla}_k E(\vec{k})$ die Gruppengeschwindigkeit und $W_{kk'}$ die Übergangsrate vom Zustand mit Impulsvektor \vec{k}' in den Zustand \vec{k} . Im Allgemeinen können nur im thermodynamischen Gleichgewicht exakte analytische Lösungen für die Boltzmann-Gleichung angegeben werden. Oft macht man einen Relaxationszeitansatz, durch den der Stoßterm vereinfacht dargestellt wird. Im Rahmen des Drift-Diffusionsmodells wird für den inhomogenen Transport in Halbleiterbauelementen eine Momentenentwicklung der Boltzmann-Gleichung vorgenommen und lediglich das nullte und erste Moment berücksichtigt. Das nullte Moment liefert die Kontinuitätsgleichung, das erste Moment einen Ausdruck für den Strom:

$$\vec{j} = \mu n \vec{\nabla} E_F. \quad (3.6)$$

Hier bedeutet μ die Beweglichkeit, n die Teilchendichte und E_F das Quasi-Ferminiveau. Der Strom ist also proportional zur Teilchendichte und zum Gradienten des Quasi-Ferminiveaus. Dabei macht man die zusätzliche Annahme, dass sich die Ladungsträger nahe des thermodynamischen Gleichgewichts befinden, so dass man für die Verteilungsfunktion eine Fermi-Verteilung ansetzen kann. Dieses lokale thermodynamische Gleichgewicht wird durch das lokale Quasi-Ferminiveau $E_F(\vec{r})$ beschrieben:

$$n(\vec{r}) = N_C(\vec{r}) \int_{E_C}^{\infty} f(E - E_F(\vec{r})) dE,$$

mit der Fermifunktion f , der Bandkante E_C und der Zustandsdichte N_C .

Wigner-Verteilung und Dichte-Gradienten-Methode

In quantenmechanischen Systemen ist es nicht möglich, eine Verteilungsfunktion im Phasenraum zu definieren, die als Wahrscheinlichkeitsdichte interpretiert werden kann. Wigner [37] hat als Erster eine Funktion angegeben, die formal analog

zur klassischen Wahrscheinlichkeitsdichte ist und diese als klassischen Grenzfall umfasst. Seien Ψ_n die Eigenzustände und E_n die Eigenenergien der Schrödinger-Gleichung des Systems [40]

$$-\frac{\hbar^2}{2m} \vec{\nabla}^2 \Psi_n(\vec{x}) + U(\vec{x}) \Psi_n(\vec{x}) = E_n \Psi_n(\vec{x}) \quad (3.7)$$

mit dem selbstkonsistenten Potenzial $U(\vec{x})$. Dann gilt für die Dichtematrix im thermischen Gleichgewicht

$$\rho(\vec{x}, \vec{x}') = \sum_n e^{-\beta E_n} \Psi_n(\vec{x}) \Psi_n(\vec{x}'). \quad (3.8)$$

Die Wigner-Funktion lautet nun

$$f_W(\vec{x}, \vec{p}) = \frac{1}{(\pi\hbar)^3} \int e^{\frac{2i}{\hbar} \vec{p} \cdot \vec{y}} \rho(\vec{x} + \vec{y}, \vec{x} - \vec{y}) d^3\vec{y} \quad (3.9)$$

Sie lässt sich als Fourier-Transformierte der Dichtematrix in Schwerpunktskoordinaten interpretieren. Wigner konnte zeigen [37][38], dass man Informationen über das Quantensystem erlangen kann, ohne die Schrödinger-Gleichung (3.7) explizit zu lösen. M. G. Ancona konnte daraus störungstheoretisch die Dichte-Gradienten-Methode herleiten [39][40]. Im Prinzip handelt es sich hierbei um verallgemeinerte Drift-Diffusionsgleichungen, d.h. um die bekannte Drift-Diffusionsgleichung ergänzt um einen Korrekturterm, der Quantisierungseffekten in niedrigster Ordnung Rechnung trägt. Die Transportgleichung lautet jetzt:

$$\vec{j} = \mu n \vec{\nabla} E_F + 2\mu b n \vec{\nabla} \left(\frac{\vec{\nabla}^2 s}{s} \right) \quad (3.10)$$

Dabei gilt $s \equiv \sqrt{|n|}$ und $b = \frac{\hbar^2}{12mq}$. Der Strom hängt nun auch vom Gradienten der Dichte ab. Dadurch werden auch in gewissem Maße nichtlokale Effekte mitberücksichtigt. In der Tat können damit in niedrigster Näherung Quantisierungseffekte in Inversionskanälen beschrieben werden [39]. Die Dichte-Gradienten-Methode ist die auf dem Gebiet der Bauelementsimulation am häufigsten verwendete Methode, um quantenmechanische Effekte in Inversionskanälen zu berücksichtigen.

Fassen wir kurz zusammen: Alle Simulationsprogramme, die realistische zwei- und dreidimensionale Strukturen rechnen können, basieren bisher auf der Boltzmann-Gleichung. Im Rahmen der Drift-Diffusionsmethode nimmt man an, dass sich die Ladungsträger lokal im thermodynamischen Gleichgewicht befinden, welches durch eine lokale Quasi-Ferminienergie charakterisiert wird. Quantisierungseffekte werden durch Hinzunahme eines Korrekturterms, der vom Gradienten der Ladungsträgerdichte abhängt, aufgenommen. Von einer Lösung der Schrödinger-Gleichung wird jedoch in der Regel abgesehen.

Neue Transportmethode

Wir behalten das bisher sehr erfolgreiche Konzept der lokalen Quasi-Fermienergien bei, gehen aber insofern einen Schritt weiter, als wir die Mehrband-kp-Schrödingergleichung vollständig lösen und dadurch eine sehr genaue Berücksichtigung der Bandstruktur erreichen. Die Ladungsträgerdichte wird in unserem Modell vollständig quantenmechanisch berücksichtigt. Wir legen damit dieselben Basisfunktionen zugrunde wie Fischetti in der Pauli-Master-Gleichung und fordern gleichfalls eine diagonale Dichtematrix. Ebenso erhalten wir adäquate Randbedingungen an ohmschen Kontakten, indem wir Ladungsneutralität fordern. Anders als Fischetti nehmen wir an, dass die Ladungsträgerverteilung nicht allzuweit vom Gleichgewicht entfernt ist. Dies erlaubt uns, die Dichte mit einer Fermiverteilung zu charakterisieren. Die Ladungsträgerdichte erhalten wir nun, indem wir die globalen quantenmechanischen Zustände mit dem lokalen Quasi-Fermienergielevel besetzen:

$$n(\vec{r}) = \sum_i |\Psi_i^n(\vec{r})|^2 f\left(\frac{E_{Fn}(\vec{r}) - E_i^n}{k_B T}\right)$$

$$p(\vec{r}) = \sum_i |\Psi_i^p(\vec{r})|^2 f\left(\frac{E_{Fp}(\vec{r}) - E_i^p}{k_B T}\right)$$

n ist die Elektronendichte, p die Löcherdichte. Die lokalen Quasi-Fermienergielevel erhält man aus der globalen Stromerhaltung $\text{div}(\vec{j}_n) = 0$ bzw. $\text{div}(\vec{j}_p) = 0$. Für den Strom nehmen wir an, dass er proportional zur Ladungsträgerdichte und zum Gradienten des Quasi-Fermienergielevels ist, was genau dem klassischen Grenzfall entspricht (vgl. Gl. 3.6). Dieses Konzept hat den Vorteil, dass es problemlos auf zwei und drei Dimensionen erweiterbar ist. Es beschreibt die Umverteilung der Ladungsträger unter angelegter Spannung in vernünftiger Weise und liefert ohne weiteren Aufwand die quantenmechanischen Zustände des Systems. Daraus können sofort die optischen Eigenschaften bestimmt werden.

Grundmodell

Zur Erläuterung und zum Vergleich mit den anderen Methoden beschränken wir uns hier auf eindimensionale Bauelemente, die nur Elektronen als freie Ladungsträger enthalten. Die Schrödingergleichung lösen wir in Einband-Näherung. Das entsprechende Gleichungssystem lautet nun:

$$-\text{div}[\varepsilon(z) \text{grad } \Phi(z)] = 4\pi e [-n + N_D^+] \quad (3.11)$$

$$-\frac{\hbar^2}{2} \nabla \left[\frac{1}{m^*(z)} \nabla \right] \Psi_n(z) + V(z) \Psi_n(z) = E_n \Psi_n(z) \quad (3.12)$$

$$\text{div}(\vec{j}) = 0 \quad (3.13)$$

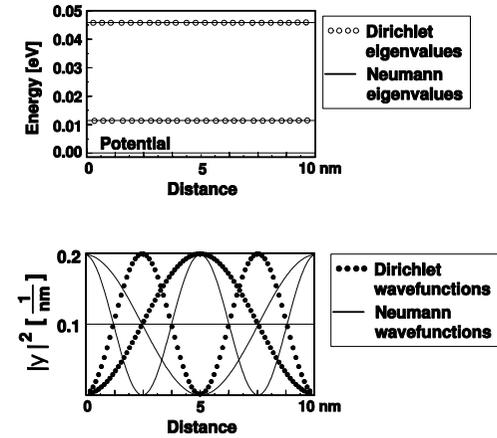


Abbildung 3.2: Eigenwerte bei flachem Potential von 0 eV mit Dirichlet- und von Neumann-Randbedingungen. Der unterste von Neumann Eigenwert hat den Wert Null und liegt damit genau auf dem Potenzialboden.

Dabei gilt für die Dichte der ionisierten Donatoren sowie für den Strom:

$$N_D^+(z) = \frac{N_D(z)}{1 + g_D e^{(E_{Fn}(z) - E_D)/k_B T}} \quad (3.14)$$

$$j(z) = \mu_n(z) n(z) \nabla_z E_{Fn}(z) \quad (3.15)$$

Hier ist $N_D(z)$ die Donatordichte, E_D die Ionisierungsenergie, g_D der Entartungsfaktor der Energieniveaus und $\mu(z)$ die Elektronenbeweglichkeit. Gleichung (3.11) stellt die Poissongleichung dar. Dabei bezeichnet n die Elektronendichte und N_D^+ die Dichte der ionisierten Donatoren. Die Einband-Schrödingergleichung (3.12) liefert die Zustandsdichte der quantisierten Zustände (mit den Wellenfunktionen $\Psi_n(z)$ und den zugehörigen Eigenwerten E_n), wobei $V(z) = E_{c0}(z) - e\Phi(z)$ ($e = |e|$ Elementarladung) der Bandkantenverlauf ist. Die Elektronendichte $n(z)$, welche in die Poissongleichung eingeht, erhält man durch Besetzen der quantisierten Zustände $\Psi_n(z)$ über das lokale Quasi-Fermienergielevel $E_{Fn}(z)$. Die Stromgleichung (3.13) stellt schließlich den Zusammenhang zwischen Elektronendichte und Quasi-Fermienergie her. Zunächst wollen wir die Randbedingungen für obige Gleichungen genauer erläutern.

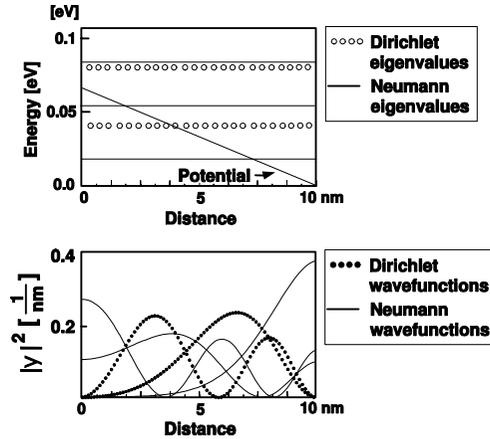


Abbildung 3.3: Eigenwerte bei schrägem Potenzial mit Dirichlet- und von Neumann-Randbedingungen.

Randbedingungen für Schrödingergleichung

Für die Schrödingergleichung kommen Dirichletsche Randbedingungen, d.h. $\Psi(z)|_{Rand} = 0$, bzw. von Neumannsche Randbedingungen, d.h. $\frac{\partial \Psi(z)}{\partial z}|_{Rand} = 0$, in Betracht. Die Wirkung dieser Randbedingungen soll zunächst anhand eines flachen eindimensionalen Potenzialverlaufs auf einem Gebiet von 10 nm Länge untersucht werden (Abb. 3.2).

Bei Dirichlet-Randbedingungen haben die Eigenfunktionen Kosinusform, wie man sie bei einem Potenzialtopf mit unendlich hohen Wänden erhält. Setzt man nun von Neumann Randbedingungen, so erhält man dieselben Eigenwerte, die dazugehörigen Eigenfunktionen sind dagegen um eine Phase von $\pi/2$ verschoben. Zusätzlich tritt der Eigenwert Null auf mit einer konstanten Eigenfunktion.

Ist das Potential nicht flach, sondern verläuft es etwas schräg, so ergibt sich die Situation wie in Abb. (3.3). Löst man die Schrödingergleichung nur in einer inneren Region des Bauelements, so kann man entweder Dirichlet- oder von Neumann-Randbedingungen verwenden. Die Zustände werden schließlich auf eins normiert, d.h. $\int |\Psi(z)|^2 dz = 1$. Dies bedeutet, dass ein vollständig besetzter Zustand genau einem Elektron entspricht. An Kontakten machen wir die Annahme, dass Ladungsneutralität vorliegt und die Zustandsdichte dieselbe wie im Volumenhalbleiter ist. Mit reinen Dirichlet-Randbedingungen würde die Dichte an den Kontakten jedoch Null werden, wohingegen sie bei von Neumann-

Randbedingungen zu den Kontakten hin ansteigen würde. Daher ist es sinnvoll, einen Mischzustand anzunehmen. Konkret heißt dies, dass wir die Schrödingergleichung einmal mit Dirichlet- und einmal mit von Neumann-Randbedingungen lösen und die Zustände jeweils auf 1/2 normieren: $\int |\Psi(z)|^2 dz = 1/2$. Man kann sich dies anhand der Abb.(3.2) plausibel machen. Für denselben Eigenwert erhält man dort einmal eine Kosinus- und einmal eine Sinusfunktion. Die Summe über beide ist konstant. Da diese Eigenfunktionen jeweils auf 1/2 normiert sind, entspricht die vollständige Besetzung dieses Mischzustandes wieder genau einem Elektron.

Randbedingungen für Poisson- und Stromgleichung

Die Lösung der Schrödingergleichung liefert uns die Zustandsdichte im Bauelement. Diese Zustandsdichte wird über das Quasi-Ferminiveau besetzt. Um Ladungsneutralität zu gewährleisten, lösen wir zunächst für den Fall des thermodynamischen Gleichgewichts mit konstantem Quasi-Ferminiveau E_{F0} die Poisson- und Schrödingergleichung selbstkonsistent, wobei wir als Randbedingungen an den Kontakten $\frac{\partial \Phi(z)}{\partial z} = 0$ fordern. Dies entspricht der Annahme von Ladungsneutralität an ohmschen Kontakten. Nach dem Gaußschen Satz ist dann die Gesamtladung gleich Null, d.h. die Ladung der freien Ladungsträger ist gleich der Ladung der ionisierten Donatoren. Man erhält daraus das eingebaute Potential (engl. built-in potential). Die Werte dieses Potentials an den Kontakten Φ_l (linker Kontakt) und Φ_r (rechter Kontakt) werden nun für die Rechnungen im Nichtgleichgewicht als Dirichlet Randbedingungen verwendet.

Konkret heißt dies bei Anlegen einer Spannung U :

$$E_{Fn}|_{linker\ Kontakt} = E_{F0} \quad (3.16)$$

$$E_{Fn}|_{rechter\ Kontakt} = E_{F0} - eU \quad (3.17)$$

$$\Phi|_{linker\ Kontakt} = \Phi_l \quad (3.18)$$

$$\Phi|_{rechter\ Kontakt} = \Phi_r + U \quad (3.19)$$

Daraus ergibt sich, dass die Ladungsträgerdichte an einem ohmschen Kontakt bei angelegter Spannung gleich der Ladungsträgerdichte im Gleichgewicht sein soll.

Numerische Lösung

Man erkennt, dass die selbstkonsistente Lösung dieses Gleichungssystems eindeutig durch Angabe des Potentials $\Phi(z)$ und des Quasi-Ferminiveaus $E_{Fn}(z)$ bestimmt ist

Methodisch gliedert sich unser Vorgehen dementsprechend in zwei Blöcke. Im ersten Schritt wird bei festgehaltenem, ortsabhängigen Quasi-Ferminiveau die Schrödinger- und Poissongleichung selbstkonsistent gelöst. Daraus ergeben sich das neue Potential und die neuen quantisierten Zustände. Im nächsten Schritt

hält man das Potenzial und die Zustände fest und ermittelt durch Lösen der Stromgleichung das neue Quasi-Ferminiveau. Dieser Zyklus wird solange wiederholt, bis Selbstkonsistenz erreicht ist.

3.3 Vergleich der Methoden

In diesem Abschnitt sollen die vorgestellten Methoden verglichen und diskutiert werden. Wir betrachten dazu einen 300 nm breiten nin-Resistor aus Silizium, wie er in Abb. 3.4 dargestellt ist.

Er besitzt auf beiden Seiten n-dotierte Bereiche von 100 nm Länge mit einer Dotierkonzentration von $N_D = 10^{17} \text{ cm}^{-3}$. Der mittlere Bereich ist intrinsisch mit einer Hintergrundkonzentration von $N_D = 10^{13} \text{ cm}^{-3}$. An beiden Enden befinden sich ohmsche Kontakte. Zur Untersuchung der Quanteneigenschaften haben wir einen Quantentopf von 10 nm Breite eingefügt, der 175 nm vom linken Kontakt entfernt ist. Für ihn sollen dieselben Materialeigenschaften wie für das Silizium gelten, mit Ausnahme eines Sprungs in der Leitungsbandkante von 0.1 eV. Die effektive Elektronenmasse beträgt $m^* = 0.32m_0$, die dielektrische Konstante hat den Wert $\varepsilon = 11.7$. Für die Donatoren wurde eine Ionisierungsenergie von $E_D = 15 \text{ meV}$ sowie ein Entartungsfaktor von $g_D = 2$ angenommen. Als Modell für die Beweglichkeit gilt folgender Zusammenhang:

$$\mu_{n,E0}(N_D) = \mu_{n,min} + \frac{\mu_{n,D}}{1 + (N_D/N_{n,ref})^{\alpha_n}}$$

$$\mu_n(\vec{E}) = \frac{\mu_{n,E0}}{\left[1 + \left(\mu_{n,E0} \frac{|\vec{E}|}{v_{n,s}}\right)^{\kappa_n}\right]^{1/\kappa_n}}$$

Dabei bedeuten: $\mu_{n,min} = 55.2 \text{ Vs/cm}^2$, $N_{n,ref} = 1.072 \times 10^{17} \text{ cm}^{-3}$, $\mu_{n,D} = 1374 \text{ Vs/cm}^2$, $\alpha_n = 0.73$, $v_{n,s} = 1.03 \times 10^7 \text{ cm/s}$ und $\kappa_n = 2.0$. \vec{E} ist das elektrische Feld. Die Temperatur beträgt $T = 300 \text{ K}$.

nin-Resistor ohne Quantentopf

Zunächst betrachten wir den nin-Resistor ohne Quantentopf. In Abb. 3.5 sind die Leitungsbandenergien und die Elektronendichten eingezeichnet. Man sieht, dass die Lösung der Pauli-Master-Gleichung [17] und die Lösung der neu entwickelten Methode sehr gut übereinstimmen. Als Stromdichte erhält Fischetti $6.8 \times 10^4 \text{ A/cm}^2$ für die Pauli-Master-Gleichung, unsere Methode liefert eine Stromdichte von $3.6 \times 10^4 \text{ A/cm}^2$. Man könnte zunächst glauben, dass die gute Übereinstimmung trivial ist, da bei dieser Struktur quantenmechanische Effekte keine so große Rolle spielen. Wir möchten jedoch darauf hinweisen, dass dies keineswegs der Fall ist, da wir die Dichte vollständig quantenmechanisch mit selbstkonsistenten nichtlokalen Quasi-Ferminiveaus berechnet haben.

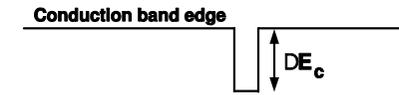
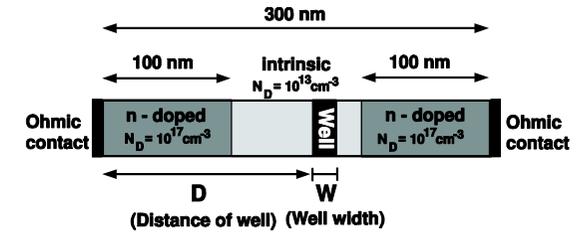


Abbildung 3.4: Das obere Bild zeigt die Geometrie des nin-Resistors. Der Quantentopf ist $D = 175 \text{ nm}$ vom linken Kontakt entfernt und $W = 10 \text{ nm}$ breit. Der Verlauf der Leitungsbandkante ist unten schematisch dargestellt. Die Topftiefe beträgt $\Delta E_C = 0.1 \text{ eV}$.

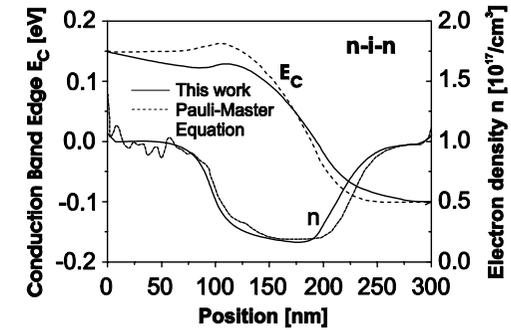


Abbildung 3.5: Leitungsbandkanten E_C und Elektronendichten n für nin-Struktur. Es werden die Pauli-Master-Gleichung (gestrichelt) und das hier vorgestellte Verfahren (durchgezogen) verglichen.

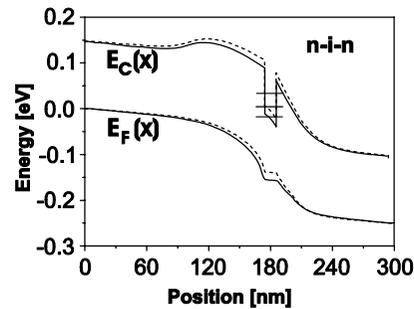


Abbildung 3.6: Vergleich zwischen klassischer (gestrichelt) und quantenmechanischer (durchgezogen) Rechnung für einen nin-Resistor mit Quantentopf. Es sind jeweils die Leitungsbandkante E_C und das Quasi-Ferminiveau E_F eingezeichnet. Man sieht, dass im Bereich der gebundenen Zustände das Quasi-Ferminiveau praktisch konstant ist.

nin-Resistor mit Quantentopf

Nun untersuchen wir den nin-Resistor mit dem Quantentopf. Das Leitungsbandprofil und das Quasi-Ferminiveau werden in Abb. 3.6 für klassische und quantenmechanische Rechnung gegenübergestellt. Im Quantentopf befinden sich drei stark gebundene Zustände. Ihre Wellenfunktionen und die Elektronendichte sind in Abb. 3.7. Quantenmechanisch ergibt sich eine kleinere Elektronendichte als bei klassischer Rechnung mit Thomas-Fermi-Näherung. Dies führt zu einem etwas stärkeren Potenzialabfall im Bereich des Quantentopfes, wodurch sich ein etwas größerer Strom ergibt. Mit klassischer Drift-Diffusion erhalten wir eine Stromdichte von $J = 2.2 \times 10^4$ A/cm², während die neue Methode $J = 2.5 \times 10^4$ A/cm² liefert. Die lokalisierten Zustände tragen faktisch nicht zum Stromfluss bei. Folglich ist das lokale Quasi-Ferminiveau praktisch konstant im Bereich der gebundenen Zustände.

Daraus kann man nun das Konzept der nichtlokalen Quasi-Ferminiveaus rechtfertigen. In der Tat entsteht kein Widerspruch für Situationen nahe des thermodynamischen Gleichgewichts, da das lokale Variieren der Besetzung eines Zustands aufgrund des lokalen Fermi-niveaus vernachlässigbar ist. Dafür gibt es drei Gründe:

- Gebundene Zustände tragen nicht zum Stromfluss bei und führen daher nicht zu einem Gradienten des Quasi-Ferminiveaus.
- Das Quasi-Ferminiveau variiert am stärksten in Regionen, wo die Dichte

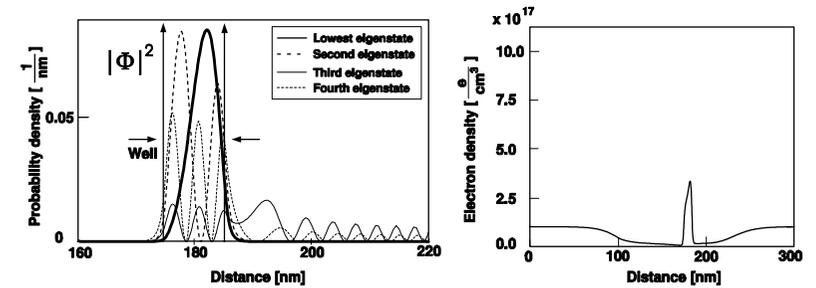


Abbildung 3.7: Eigenzustände im Quantentopf für den nin-Resistor aus Abb. 3.6 (links). Rechts ist die daraus resultierende Elektronendichte dargestellt.

sehr gering ist, z.B. in Barrieren.

- Ausgedehnte Zustände, die in unserem Verfahren formal als gebundene, stationäre Zustände behandelt werden sind entweder aufgrund ihrer hohen Energie überhaupt nicht besetzt, oder sie treten in Regionen hoher Dichte (z.B. nahe am Kontakt) auf, wo das Quasi-Ferminiveau nahezu konstant ist.

Tunnelbarriere

In Abb. 3.8 zeigen wir Resultate für eine Si-SiO₂-Si-Tunnelodiode mit einer 3 nm dicken und 3 eV hohen Barriere bei einer angelegten Spannung von 1 V. Man erhält für den Verlauf der Bandkante und der quantenmechanischen Dichte wieder eine sehr gute Übereinstimmung. Beim Strom treten hier jedoch Abweichungen auf. Während die Pauli-Master-Gleichung eine Stromdichte von 10^{-8} A/cm² liefert, ergibt sich bei unserem Zugang eine deutlich höhere Stromdichte von 10^{-3} A/cm². Dies hat einen interessanten Grund, der anhand von Abb. 3.9 erläutert werden soll. Die Dichteverteilung in der Barriere ist bei uns symmetrisch, da sie von Wellenfunktionen gebildet wird, die von links und von rechts in die Barriere eindringen. Im Gegensatz dazu verwendet Fischetti zur Berechnung des Tunnelstroms Streuzustände, die von links nach rechts laufen. Dadurch, dass bei uns die Wellenfunktionen aus beiden Richtungen eingehen, wird die Barrierendicke effektiv halbiert. Man könnte sich in unserem Modell durch einen Trick behelfen, indem man die rechte Grenze der Quantenregion an das rechte Ende der Barriere setzt. Dann erhält man auch hier nur Wellenfunktionen, die von links in die Barriere eindringen. Damit ergibt sich für die halbe Barrierendicke derselbe Tunnelstrom wie vorher, nämlich 10^{-3} A/cm². Dieses Artefakt tritt bei

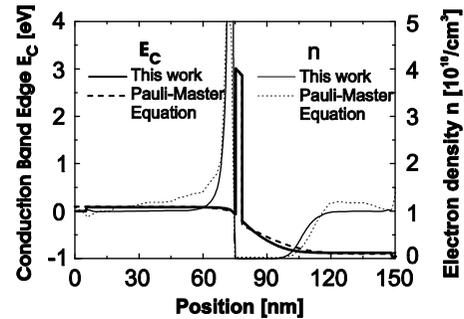


Abbildung 3.8: Vergleich der Pauli-Master-Gleichung mit dem vorgestellten Verfahren anhand einer Si-SiO₂-Si-Tunnelbarriere. Die Leitungsbandkante E_C und die Elektronendichte n stimmen gut überein.

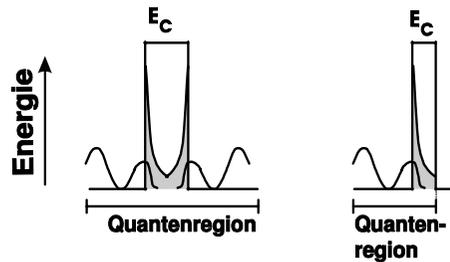


Abbildung 3.9: Erklärung für den unterschiedlichen Tunnelstrom bei ballistischem Transport und dem vorgestellten semiklassischen Verfahren. Erstreckt sich die Quantenregion über das gesamte Bauelement (links), so erhält man eine symmetrische Dichteverteilung innerhalb der Barriere, da Wellenfunktionen von links und von rechts in die Barriere eindringen. Dadurch wird die effektive Barriendicke halbiert. Man erhält denselben Tunnelstrom wie in der rechten Grafik, wo die Quantenregion am rechten Barrierenrand abgeschnitten wird, so dass von rechts keine Wellenfunktionen eindringen können.

allen semiklassischen Modellierungen des Stroms auf, insbesondere auch bei der Dichte-Gradienten-Methode [36].

3.4 Zusammenfassung und Ausblick

In diesem Abschnitt haben wir ein Verfahren entwickelt, das es erlaubt, die elektronische Struktur im Nichtgleichgewicht zu berechnen. Im Wesentlichen handelt es sich um eine Kombination der stationären Lösungen der Schrödingergleichung mit einem semiklassischen Drift-Diffusionsmodell. Die Methode liefert für das elektrische Potenzial und die Dichteverteilung sehr gute Übereinstimmung mit der Pauli-Master-Gleichung. Auch Ströme können gut beschrieben werden, wie das Beispiel des Double-Gate-MOSFET (Metal-Oxide-Semiconductor-Field-Effect-Transistor) in Kapitel 7 zeigt. Schwächen hat das Modell jedoch bei der Berechnung von Tunnelströmen. Es bietet sich daher an, das beschriebene Verfahren zur selbstkonsistenten Lösung der Poisson-, Strom- und Schrödingergleichung beizubehalten, um so den Verlauf des Potenzials, der Quasi-Ferminiveaus und der quantenmechanischen Zustände zu erhalten. Aus den so erhaltenen stationären Zuständen kann dann die ballistische Transmissionswahrscheinlichkeit durch Tunnelbarrieren über eine Greenfunktions-Methode ermittelt werden.

Kapitel 4

Halbleitergleichungen

Die effiziente und zuverlässige Lösung der Poissongleichung ist Grundlage jeder Bauelementsimulation. Für unsere Problemstellungen ist es insbesondere notwendig, Heterobarrieren in sinnvoller Weise behandeln zu können. In diesem Kapitel stellen wir kurz die Poissongleichung mit den dazugehörigen Randbedingungen dar. Technische Details zum Aufstellen des Gleichungssystems finden sich in Anhang A. Unter angelegten Spannungen hat man kein konstantes Fermi-niveau mehr. Die Nichtgleichgewichts-Situation wird nun über örtlich variierende Quasi-Fermi-niveaus beschrieben (vgl. Kapitel 3). Diese erhält man durch Lösen der Stromgleichung, welche hier kurz zusammen mit der Behandlung der Randbedingungen an Kontakten behandelt wird. Details zum Aufstellen der Matrix bzw. zu den Rekombinationsmechanismen können in Anhang A bzw. H nachgelesen werden.

In Heterobauelementen kommt es aufgrund unterschiedlicher Gitterkonstanten zu mechanischen Verspannungen. Diese werden im Rahmen der Kontinuumsmechanik berechnet. Zunächst stellen wir die Grundlagen der Elastizitätstheorie dar und diskutieren schließlich die Spannungsgleichung. Technische Details zum Aufstellen der Matrix wurden in Anhang B ausgelagert.

Um oben genannte Gleichungen einer numerischen Behandlung zugänglich zu machen, müssen sie auf ein diskretes Gitter abgebildet werden. Zur Diskretisierung verwenden wir die Finite-Volumen-Methode. Diese soll kurz erläutert und gegenüber dem Verfahren der finiten Elemente abgegrenzt werden. Abschließend diskutieren wir anhand eines Beispiels das Diskretisierungsschema.

4.1 Poissongleichung

Die Poissongleichung wird auf dem gesamten Bauelement, d.h. in Halbleiter- und Isolatormaterialien, gelöst. Sie ergibt sich aus der Maxwellgleichung für die dielektrische Verschiebung $\text{div} \vec{D} = \rho$ mit $\vec{D} = \varepsilon \vec{E} = -\varepsilon \text{grad} \Phi$ und lautet:

$$-\text{div}(\varepsilon(\vec{r}) \text{grad} \Phi(\vec{r})) = \rho(\vec{r})$$

Dabei bezeichnen $\varepsilon(\vec{r})$ die ortsabhängige dielektrische Konstante und $\Phi(\vec{r})$ das elektrische Potenzial. Die elektrische Ladungsdichte $\rho(\vec{r})$ setzt sich zusammen aus Elektronen, Löchern, ionisierten Störstellen und piezo- und pyroelektrischen Polarisationsladungen. Einzelheiten über die Dichteberechnung sind in Kap. 3 und im Anhang D genauer beschrieben, auf die Herkunft der piezoelektrischen Polarisationsladungen wird in Abschnitt 4.3 eingegangen.

Mit dem Potenzial ergeben sich die Bandkanten zu $E_c(\vec{r}) = E_{c0}(\vec{r}) - e\Phi(\vec{r})$ für das Leitungsband und $E_v(\vec{r}) = E_{v0}(\vec{r}) - e\Phi(\vec{r})$ für das Valenzband ($e = |e|$ Elementarladung), wobei E_{c0} und E_{v0} die Bandenergien im betreffenden Volumenhalbleiter (engl. bulk) darstellen.

Ist die Dichte $\rho(\vec{r})$ konstant, d.h. unabhängig vom Potenzial, so spricht man von der linearen Poissongleichung, welche in diskretisierter Form durch ein lineares Gleichungssystem beschrieben werden kann. In der Regel hängt die Dichte $\rho(\vec{r})$ jedoch vom Potenzial ab, so dass man es mit einem nichtlinearen Problem zu tun hat, dessen numerische Behandlung in Kap. 5 erläutert wird.

Kontaktregionen sind von der Simulation ausgenommen, an ihren Grenzflächen werden entsprechende Randbedingungen gesetzt. Mathematisch unterscheidet man im Wesentlichen zwei Arten von Randbedingungen (Rand ∂G):

1. Dirichlet Randbedingungen

$$\Phi|_{\partial G} = \text{const}$$

2. von Neumann Randbedingungen

$$\frac{\partial \Phi}{\partial \vec{n}}|_{\partial G} = \text{const}$$

Eingebautes Potenzial

Zur Bestimmung der Randwerte an Kontakten ist es wichtig, das eingebaute Potenzial (engl. built-in potential) zu kennen. Dazu wird für den Fall des thermodynamischen Gleichgewichts mit konstantem Fermi-niveau die Poisson- und Schrödingergleichung selbstkonsistent gelöst, wobei wir Ladungsneutralität für das gesamte Bauelement fordern, d.h. die Ladung der freien Ladungsträger ist gleich der Ladung der ionisierten Störstellen und der Polarisationsladungen. Nach dem Gaußschen Satz folgt daraus, dass das Flächenintegral des elektrischen Feldes über die gesamte Randfläche gleich der Oberflächenladung sein muss.

Ohmscher Kontakt

Ein ohmscher Kontakt liegt vor, wenn die Dicke der Raumladungszone an einem Metall-Halbleiter-Übergang so gering ist, dass die Ladungsträger die entsprechende Energiebarriere leicht durchtunneln können [41]. Das Potenzial in der

Nähe des Kontaktes stellt sich so ein, dass lokal Ladungsneutralität vorliegt. Bei der Simulation werden standardmäßig alle Grenzflächenphänomene direkt an der Grenzschicht außer Acht gelassen [42] und als Randbedingung für die Berechnung des eingebauten Potentials das elektrische Feld auf Null gesetzt.

Schottky-Kontakt

Bei einem Schottky-Kontakt führen Oberflächenzustände dazu, dass die Leitungsbandkante einen festen Wert (Schottky-Barriere), der von der Ionisierungsenergie der Oberflächenzustände abhängt, oberhalb des Fermi-niveaus fixiert ist (engl. pinning). Diese Randbedingung vom Typ Dirichlet führt zu einem nichtverschwindenden elektrischen Feld, welches durch die aus den besetzten Oberflächenzuständen resultierende Ladung hervorgerufen wird. Nähere technische Details und die Realisierung von MOS-Kontakten (metal oxide semiconductor) sind in Anhang G zu finden.

Nichtgleichgewicht

Wird nun eine Spannung U an den ohmschen oder Schottky-Kontakt angelegt (Nichtgleichgewichts-Situation), so lautet die Randbedingung in beiden Fällen

$$\Phi|_{\text{Kontakt}} = \Phi|_{\text{built-in}} - U.$$

Für die Fermienergie gilt mit dem Gleichgewichtsferminiveau E_{F0} und dem Bandindex c : $E_{Fc}|_{\text{Kontakt}} = E_{F0} - eU$. Daraus ergibt sich, dass die Ladungsträgerdichte an einem Kontakt bei angelegter Spannung gleich der Ladungsträgerdichte im thermodynamischen Gleichgewicht ist.

Heteroübergang

Die Bandkantensprünge an Heteroübergängen werden mikroskopisch durch elektrische Dipolschichten erzeugt. Hier ist es üblich, Stetigkeit des elektrostatischen Potentials zu postulieren, und den Potentialsprung direkt in die Bandkanten E_{c0} und E_{v0} der Volumenhalbleiter zu inkludieren. Infolgedessen müssen E_{c0} und E_{v0} in der Datenbasis für alle Materialien aufeinander abgestimmt sein, was im Rahmen des van-de-Walle Modells erfolgt [43] (vgl. Abschnitt 4.3 und Anhang I). Details über die Diskretisierung der Poissongleichung finden sich in Anhang A.

4.2 Stromgleichungen

Wie in Kap. 3 ausführlich erläutert, lautet der Teilchenstrom für die Elektronen \vec{j}_n und Löcher \vec{j}_p in unserem Modell:

$$\begin{aligned}\vec{j}_n(\vec{r}) &= -\hat{\mu}_n n(\vec{r}) \nabla E_{Fn}(\vec{r}) \\ \vec{j}_p(\vec{r}) &= \hat{\mu}_p p(\vec{r}) \nabla E_{Fp}(\vec{r})\end{aligned}\quad (4.1)$$

Dabei bezeichnen der Tensor zweiter Stufe $\hat{\mu}_{n/p}$ die Beweglichkeit, $n(\vec{r})$ und $p(\vec{r})$ die Elektronen- bzw. Löcherdichte und $E_{Fn/p}(\vec{r})$ das Quasi-Ferminiveau der Elektronen bzw. der Löcher. Man beachte, dass $n(\vec{r})$ und $p(\vec{r})$ Teilchendichten darstellen und damit stets positiv sind.

Die Kontinuitätsgleichungen lauten dann:

$$\begin{aligned}-\text{div } \vec{j}_n(\vec{r}) &= R(\vec{r}) \\ -\text{div } \vec{j}_p(\vec{r}) &= R(\vec{r})\end{aligned}\quad (4.2)$$

Die Stromgleichungen für die Elektronen und Löcher sind über die Generations- bzw. Rekombinationsrate gekoppelt: $R(\vec{r})$ beinhaltet beide Prozesse und hat positives Vorzeichen, falls die Rekombinationsprozesse überwiegen. Die wichtigsten Mechanismen sind [47][51][52]:

1. Shockley-Read-Hall-Rekombination, bei der die Rekombination über eine tiefe Störstelle in der Bandmitte erfolgt.
2. Auger-Rekombination, bei der Impuls und Energie des rekombinierenden Ladungsträgers von einem Ladungsträger der gleichen Sorte aufgenommen wird. Es handelt sich um einen indirekten Übergang, an dem ein drittes Teilchen beteiligt ist. Infolgedessen findet diese Art der Rekombination hauptsächlich in Bereichen mit hoher Ladungsträgerdichte statt.
3. Direkte Rekombination von Elektron und Loch über einen strahlenden Übergang, tritt nur in Halbleitern mit direkter Bandlücke auf.

In die Formeln dieser Rekombinationsraten gehen in erster Linie die Elektronen- und Löcherdichte ein, welche wiederum von den Quasi-Ferminiveaus abhängen. Formeln für die Rekombinationsraten sowie Modelle für die Beweglichkeiten sind in Anhang H aufgeführt.

Randbedingungen

Eine angelegte Spannung U an ohmschen wie an Schottky-Kontakten wird als Dirichlet-Randbedingung implementiert. Mit dem Gleichgewichtsferminiveau E_{F0} gilt ($e = |e|$ Elementarladung):

$$\begin{aligned}E_{Fn}|_{\text{Kontakt}} &= E_{F0} - eU \\ E_{Fp}|_{\text{Kontakt}} &= E_{F0} - eU.\end{aligned}$$

An Heteroübergängen nehmen wir Stetigkeit der Quasi-Ferminiveaus an.

4.3 Verspannungsgleichung

Infolge äußerer Kräfte oder Gitterfehlanspassungen kommt es zu inneren Verzerrungsfeldern im Bauelement, die in jedem Punkt durch die dort herrschende Spannung charakterisiert werden können. Diese werden im Rahmen der Elastizitätstheorie berechnet.

Zum besseren Verständnis sollen zunächst kurz die Grundaussagen der Kontinuumsmechanik erläutert werden. Ausgangspunkt ist das Spannungsprinzip von Euler-Cauchy ([48], S.90):

Als Folge äußerer Kräfte existiert auf jeder Fläche des Körpers (Schnittfläche zwischen Teilkörpern oder äußere Begrenzungsfläche) mit einem Flächennormaleinheitsvektor $\vec{n}(\vec{x}, t)$ ein Vektorfeld von Spannungsvektoren $\vec{t}(\vec{x}, \vec{n}, t)$. Fällt die Fläche mit der Oberfläche des Körpers zusammen, sind die Spannungsvektoren $\vec{t}(\vec{x}, \vec{n}, t)$ gleich den aus den Oberflächenkräften folgenden Spannungsvektoren

Spannungen innerhalb des Bauelements werden mit Schnittbetrachtungen beschrieben. Is $d\vec{A}$ das Flächenelement der Schnittfläche im betrachteten Punkt und $d\vec{f}$ der resultierende Kraftvektor auf dieses Flächenelement, so gilt für den Spannungsvektor:

$$\vec{t}(\vec{x}, \vec{n}, t) = \frac{d\vec{f}}{d\vec{A}}$$

Der Spannungsvektor ist ein Maß für die innere Kraft im betrachteten Punkt des Körpers. Er ist abhängig vom Ort, von der Zeit und von der Orientierung der Schnittfläche. Es gilt das Cauchysche Spannungsprinzip:

Jedes Schnittflächenelement in einem Punkt P mit der gleichen Tangentialebene hat den gleichen Vektor \vec{n} und führt damit zum gleichen Spannungsvektor \vec{t} , d.h. unterschiedliche Oberflächenkrümmungen im Punkt P haben keinen Einfluss auf \vec{t} , solange \vec{n} sich nicht verändert.

Daraus kann man folgern, dass die Gesamtheit aller denkbaren Spannungsvektoren für einen materiellen Punkt P den Spannungszustand in diesem Punkt definiert.

Es gilt das Cauchysche Fundamentaltheorem:

Der Spannungsvektor $\vec{t}_n = \vec{t}$ im Punkt \vec{x} einer gegebenen Schnittfläche mit dem Normalenvektor \vec{n} ist vollständig durch drei Spannungsvektoren $\vec{t}_{e_i} \equiv \vec{t}_i$ bestimmt, die auf den drei Koordinatenflächen wirken, die sich gegenseitig in \vec{x} durchdringen. Der Spannungsvektor \vec{t}_n ist eine lineare Funktion von \vec{n} . Die Gleichung

$$\vec{t}(\vec{x}, \vec{n}, t) = \vec{n} \cdot \hat{\sigma}(\vec{x}, t)$$

in Komponenten:

$$t_i = \sigma_{ji} n_j$$

beschreibt den Zusammenhang des von \vec{n} abhängigen Spannungsvektors \vec{t} mit dem

von \vec{n} unabhängigen Spannungstensor $\hat{\sigma}$. Der Spannungszustand in \vec{x} ist somit entweder durch drei Spannungsvektoren $\vec{t}_{e_i} \equiv \vec{t}_i$ oder durch 9 Tensorkomponenten $\sigma_{ij} \vec{e}_i \vec{e}_j$ eindeutig bestimmt.

Befindet sich der Gesamtkörper im Gleichgewicht, so gilt:

$$\nabla_{\vec{x}} \cdot \hat{\sigma} + \rho \vec{k} = \vec{0} \quad (4.3)$$

bzw. in integraler Form

$$\int_V (\nabla_{\vec{x}} \cdot \hat{\sigma} + \rho \vec{k}) dV = \vec{0} \quad (4.4)$$

Dabei sind $\rho \vec{k} dV$ Volumenkräfte. $\nabla_{\vec{x}}$ bezeichnet die Vektordivergenz und ist folgendermaßen definiert:

$$(\nabla_{\vec{x}} \cdot \hat{\sigma})_i = \sum_j \frac{\partial}{\partial x_j} \sigma_{ij} \quad (i, j = 1..3) \quad (4.5)$$

Bei der Integration ergibt sich mit dem Gaußschen Satz

$$\int_V \nabla_{\vec{x}} \cdot \hat{\sigma} dV = \int_A \vec{t} dA = \int_A \vec{n} \cdot \hat{\sigma} dA \quad (4.6)$$

Unter Annahme kleiner Verzerrungen kann man das Hooksche Gesetz schreiben als

$$\hat{\sigma} = \mathbf{E}^{(4)} \hat{\epsilon}, \quad (4.7)$$

wo die Verzerrungen durch den zweistufigen Tensor $\hat{\epsilon}$ und der Elastizitätstensor durch den 2-stufigen Tensor $\mathbf{E}^{(4)}$ beschrieben werden. Für die Komponenten des Verzerrungstensors gilt:

$$\epsilon_{ij} = \frac{1}{2} \left(\frac{\partial u^i}{\partial x_j} + \frac{\partial u^j}{\partial x_i} \right), \quad (4.8)$$

wobei $\vec{u} = (u^1, u^2, u^3)$ die Auslenkung eines Teilchens aus der Ruhelage bezeichnet. Der Verzerrungstensor ist symmetrisch: $\epsilon_{ij} = \epsilon_{ji}$. Beim Übergang zur Voigt-Notation (11 \rightarrow 1, 22 \rightarrow 2, 33 \rightarrow 3, 23 \rightarrow 4, 13 \rightarrow 5, 12 \rightarrow 6) macht man folgende Transformation:

$$\begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \\ \epsilon_4 \\ \epsilon_5 \\ \epsilon_6 \end{bmatrix} = \begin{bmatrix} \epsilon_{11} \\ \epsilon_{22} \\ \epsilon_{33} \\ 2\epsilon_{23} \\ 2\epsilon_{13} \\ 2\epsilon_{12} \end{bmatrix}, \quad \begin{bmatrix} \sigma_1 \\ \sigma_2 \\ \sigma_3 \\ \sigma_4 \\ \sigma_5 \\ \sigma_6 \end{bmatrix} = \begin{bmatrix} \sigma_{11} \\ \sigma_{22} \\ \sigma_{33} \\ \sigma_{23} \\ \sigma_{13} \\ \sigma_{12} \end{bmatrix} \quad (4.9)$$

Dadurch lässt sich die Relation (4.7) auch in Matrix-Vektor-Schreibweise darstellen:

$$\sigma_i = E_{ij}\epsilon_j \quad (i, j = 1..6) \quad (4.10)$$

Im Kristallsystem gilt für die Elastizitätsmatrix \hat{E} für Zinkblende bzw. Wurtzitstruktur:

$$\hat{E}_{Zinkblende} = \begin{pmatrix} c_{11} & c_{12} & c_{12} & & & \\ c_{12} & c_{11} & c_{12} & & & \\ c_{12} & c_{12} & c_{11} & & & \\ & & & c_{44} & & \\ & & & & c_{44} & \\ & & & & & c_{44} \end{pmatrix}$$

$$\hat{E}_{Wurtzit} = \begin{pmatrix} c_{11} & c_{12} & c_{13} & & & \\ c_{12} & c_{11} & c_{13} & & & \\ c_{13} & c_{13} & c_{33} & & & \\ & & & c_{44} & & \\ & & & & c_{44} & \\ & & & & & c_{66} \end{pmatrix}$$

Daraus lässt sich gemäß der Voigt-Notation der Elastizitätstensor $\mathbf{E}^{(4)}$ gewinnen:

$$E_{mn} = E_{ijkl} \quad , \quad m, n = 1 \dots 6, \quad i, j, k, l = 1 \dots 3$$

Stimmen die Kristallachsen nicht mit den Simulationsachsen (x, y, z) überein, muss der Elastizitätstensor $\mathbf{E}^{(4)}$ in das Simulationssystem transformiert werden mit der Rotationsmatrix R_{ij} (vgl. Kapitel 2):

$$E_{ijkl}^{(x,y,z)} = R_{im}R_{jn}R_{ko}R_{lp}E_{mnop}^{Kristall}$$

Zur Berechnung der Verspannungen im Halbleiterbauelement wird Gleichung (4.3) diskretisiert. Wir verwenden dazu Boxintegration in Verbindung mit dem Gaußschen Satz.

Berücksichtigung von Gitterfehlpassungen

Wir nehmen an, dass sich die Auslenkungen u^i auf ein Referenzgitter eines Referenzmaterials mit einer Gitterkonstante a_B beziehen. Die physikalisch gültige Verspannung ist dann

$$\epsilon_{ij}^{phys} = \frac{a_B}{a}\epsilon_{ij} - \delta_{ij}\frac{(a - a_B)}{a}$$

Dabei bezeichnet a die am betrachteten Gitterpunkt gültige Gitterkonstante.

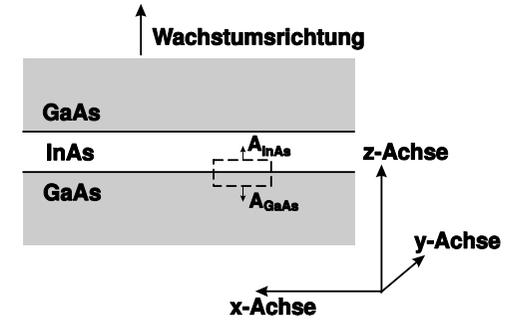


Abbildung 4.1: Schichtstruktur, bestehend aus einem dünnen InAs-Film, der pseudomorph auf GaAs-Substrat aufgewachsen wurde. Das GaAs ist unverspannt, die Verspannung in der InAs-Schicht ist homogen. Zur Berechnung des Verzerrungstensors im InAs führen wir Boxintegration über das gestrichelt eingezeichnete finite Volumen aus. Aufgrund der Homogenität der Verspannung braucht nur der Kraftfluss durch die beiden Begrenzungsflächen A_{GaAs} und A_{InAs} betrachtet werden.

Direkte Berechnung der Verspannung für Schichtstrukturen

Für eine ideale Schichtstruktur, wie sie in Abb. 4.1 dargestellt ist, kann die Verspannung unter gewissen Annahmen direkt berechnet werden. Im angegebenen Beispiel wurde eine Schicht InAs pseudomorph auf ein GaAs-Substrat aufgewachsen. Ist die Schichtbreite klein im Vergleich zur Ausdehnung in x - und y -Richtung, so ist das Substrat unverspannt und der Verspannungstensor in der InAs-Schicht homogen. Als Zwangsbedingung gilt, dass das InAs lateral die Gitterkonstante von GaAs annimmt. Daraus folgt sofort für den Verspannungstensor (in Voigt-Notation) im (x, y, z) -System:

$$\begin{aligned} \epsilon_1^{(x,y,z)} &= \epsilon_2^{(x,y,z)} = \epsilon_{||} \\ \epsilon_6^{(x,y,z)} &= 0 \end{aligned} \quad (4.11)$$

Dabei gilt $\epsilon_{||} = \frac{a_{GaAs} - a_{InAs}}{a_{InAs}}$, wo a_{GaAs} bzw. a_{InAs} die Gitterkonstante in GaAs bzw. InAs bezeichnet. Aus der Gleichgewichtsbedingung (4.3) $\nabla_{\vec{x}} \cdot \hat{\sigma} = \vec{0}$ ergibt sich durch Integration über die Box in Abb. 4.1 gemäß Gl. (4.6) unter Berücksichtigung der Homogenität des Verzerrungstensors und $A = A_{GaAs} = A_{InAs}$ (Begrenzungsfläche des finiten Volumens in z -Richtung, vgl. Abb. 4.1)

$$0 = \int_V \nabla_{\vec{x}} \cdot \hat{\sigma} dV = \int_A \vec{t} dA = \int_A \vec{n} \cdot \hat{\sigma} dA = A \begin{pmatrix} \sigma_{31}^{InAs} & \sigma_{31}^{GaAs} \\ \sigma_{32}^{InAs} & \sigma_{32}^{GaAs} \\ \sigma_{33}^{InAs} & \sigma_{33}^{GaAs} \end{pmatrix}. \quad (4.12)$$

Da das GaAs unverspannt ist, folgt aus Gl. (4.12) sofort die Bedingung (in Voigt-Notation) für die InAs-Schicht:

$$\sigma_3^{(x,y,z)} = \sigma_4^{(x,y,z)} = \sigma_5^{(x,y,z)} = 0. \quad (4.13)$$

Wegen $\sigma_i = E_{ij}\epsilon_j$ stellen die Gleichungen (4.11) und (4.13) insgesamt sechs Bedingungen für die sechs Komponenten des Verzerrungstensors ϵ_j . Dadurch kann der Verzerrungstensor im (x, y, z) -System $\epsilon^{(x,y,z)}$ für die InAs-Schicht sofort bestimmt werden.

Piezoelektrische Ladungen

Durch die Verspannung des Kristallgitters kommt es bei (3)-(5)-Halbleitern zur Ausbildung eines makroskopischen elektrischen Dipols, dessen Größe proportional zum Verzerrungstensor ist [50]. Die piezoelektrische Polarisation $\vec{P}_{piezo}(\vec{r})$ ist ein Vektor, der Verspannungstensor ϵ ein zweistufiger Tensor. Demzufolge beschreiben die piezoelektrischen Konstanten $e_{ijk}(\vec{r})$ einen Tensor dritten Grades. Es gilt:

$$P_i^{piezo} = e_{ijk} \epsilon_{jk} \quad , \quad i, j, k = 1, 2, 3 \quad (4.14)$$

Beim Übergang zur Voigt-Notation (11 \rightarrow 1, 22 \rightarrow 2, 33 \rightarrow 3, 23 \rightarrow 4, 13 \rightarrow 5, 12 \rightarrow 6) werden der erste Index beibehalten und die letzten beiden kontrahiert gemäß folgender Vorschrift:

$$e_{ijk} = \begin{cases} e_{im} & , \quad i = 1, 2, 3; m = 1, 2, 3 \\ \frac{1}{2}e_{im} & , \quad i = 1, 2, 3; m = 4, 5, 6 \end{cases} \quad (4.15)$$

Dann kann Gl. (4.14) geschrieben werden als:

$$P_i^{piezo} = e_{i1}\epsilon_1 + e_{i2}\epsilon_2 + e_{i3}\epsilon_3 + e_{i4}\epsilon_4 + e_{i5}\epsilon_5 + e_{i6}\epsilon_6 \quad (4.16)$$

Im **Kristallsystem** kann die Zahl der Konstanten erheblich reduziert werden. Für Zinkblende gilt $e_{14} = e_{25} = e_{36} \neq 0$, alle anderen Komponenten verschwinden. Damit ist die piezoelektrische Polarisation:

$$\vec{P}_{piezo}^{Zinkblende}(\vec{r}) = e_{14}(\vec{r}) \begin{pmatrix} 2\epsilon_{23} \\ 2\epsilon_{13} \\ 2\epsilon_{12} \end{pmatrix} \quad (4.17)$$

In Wurtzit ist entsprechend $e_{31} = e_{32} \neq 0$, $e_{15} = e_{24} \neq 0$, $e_{33} \neq 0$. Daraus ergibt sich für die piezoelektrische Polarisation:

$$\vec{P}_{piezo}^{Wurtzit}(\vec{r}) = \begin{pmatrix} 2e_{15}\epsilon_{13} \\ 2e_{15}\epsilon_{12} \\ e_{31}\epsilon_{11} + e_{31}\epsilon_{22} + e_{33}\epsilon_{33} \end{pmatrix} \quad (4.18)$$

Will man entsprechende Relationen im Simulationssystem (x, y, z) haben, muss der piezoelektrische Tensor e_{ijk} aus Gl. (4.14) in das Simulationssystem transformiert werden mit der Rotationsmatrix R_{ij} (vgl. Kapitel 2):

$$e_{ijk}^{(x,y,z)} = R_{im}R_{jn}R_{ko}e_{mno}^{Kristall}$$

Die piezoelektrische Polarisation führt zu Polarisationsladungen. Die dazugehörige Ladungsdichte errechnet sich aus folgender Beziehung:

$$\rho_{piezo}(\vec{r}) = -\text{div } \vec{P}_{piezo}(\vec{r}) \quad (4.19)$$

Verspannungsinduzierte Verschiebung der Bandkanten

Die relative Lage der Bandkanten verschiedener Materialien zueinander beschreiben wir im Modell der „natürlichen“ Diskontinuitäten [43][44]. Man nimmt an, dass die Energien der Bandkanten auf einer absoluten Skala angegeben werden können. Der Sprung in der Bandenergie am Übergang zwischen unverspannten („natürlichen“) Materialien ergibt sich dann aus der Differenz der absoluten Energien. Der Einfluss der Verspannungen wird in diesem Modell unabhängig von den „natürlichen“ Diskontinuitäten mittels der Deformationspotenziale behandelt. Zur Bestimmung der absoluten Energien verwendet man für jeden Halbleiter als Referenz-Niveau ein gemittelttes Potenzial. Entscheidend ist nun die Festlegung dieser Referenz-Energien auf einer absoluten Skala. Dabei ist insbesondere auf Transitivität der Banddiskontinuitäten zu achten. Van de Walle benutzt im Rahmen der „Model Solid Theory“ einen Modell-Festkörper, der aus neutralen Atomen besteht [43][45]. Da bei neutralen Atomen das Potenzial schnell genug abfällt, erhält man einen absoluten Nullpunkt, der als Bezugspunkt für die mittleren Potenziale dient [51].

Eine Volumenänderung, die durch die Diagonalelemente des Spannungstensors bestimmt wird, führt über das absolute Deformationspotenzial zu einer Verschiebung des Schwerpunktes der Bänder. Zusätzlich wird durch die spurfreien Scherungskomponenten des Spannungstensors eine Aufspaltung der Bandenergien verursacht. Diese Bandverschiebungen können für das Valenzband und das Γ -Tal im Leitungsband durch Diagonalisieren des kp-Hamiltonoperators im Volumenhalbleiter berechnet werden. Die Bedeutung der entsprechenden Deformationspotenziale kann den Angaben in Abschnitt 2.2 entnommen werden. Im Gegensatz zum Γ -Tal spalten das X - und das L -Tal auf, wenn der Spannungstensor nicht diagonal ist. Diese Aufspaltung wird durch das uniaxiale Deformationspotenzial beschrieben [46].

4.4 Numerik partieller Differentialgleichungen

In der mathematischen Literatur gibt es zahlreiche Methoden zur Lösung partieller Differentialgleichungen, die unter dem Oberbegriff der „Methode der gewichte-

ten Residuen" zusammengefasst werden können. Auch die häufigsten Verfahren in den Ingenieursanwendungen, die Finite-Elemente-Methode und die von uns verwendete Finite-Volumen-Methode, fallen darunter. Zum besseren Verständnis soll im Folgenden kurz das Grundprinzip der Methode der gewichteten Residuen dargestellt werden [53].

Grundprinzip

Gegeben sei das Randwertproblem

$$N(u) = 0, \quad u \in G \quad (4.20)$$

mit der Randbedingung $R(u) = 0$ auf dem Rand ∂G des Gebietes G . Zunächst wird eine Näherungslösung $u_N(x)$ der Differenzialgleichung als *kontinuierliche* Funktion angesetzt und als endliche Reihe geschrieben:

$$u_N(x) = u_R(x) + \sum_{k=0}^N a_k \cdot \Phi_k(x) \quad (4.21)$$

Dabei sind $u_N(x)$ die Näherungslösung, $u_R(x)$ eine Funktion, die die inhomogenen Randbedingungen erfüllt, $\Phi_k(x)$ vorgegebene Ansatzfunktionen (engl. trial functions) und a_k die gesuchten Koeffizienten.

Für die Ableitung der Funktionen u_N gilt dann (für $u_R = 0$ geschrieben):

$$\frac{d^P}{dx^P} u_N = \sum_{k=0}^N a_k \cdot \frac{d^P}{dx^P} \Phi_k(x) \quad (4.22)$$

Entwickelt man die Ableitungen wieder nach den Ansatzfunktionen, so benutzt man die Schreibweise

$$\frac{d^P}{dx^P} u_N = \sum_{k=0}^{N'} a_k^{(P)} \cdot \Phi_k(x) \quad (4.23)$$

Der Fehler, der sich beim Einsetzen des Ansatzes (4.21) in die Differenzialgleichung (4.20) ergibt, wird als Residuum $R(x)$ bezeichnet:

$$R(x) = N(u_N(x)) \quad (4.24)$$

Zur Bestimmung der noch unbekanntenen Koeffizienten a_k wird verlangt, dass das mit den $N + 1$ verschiedenen Gewichtsfunktionen (engl. test functions) $\omega_j(x)$ gewichtete Residuum bei Integration über G verschwindet, d.h. dass das Residuum zu allen Testfunktionen orthogonal ist:

$$\int_G \omega_j(x) \cdot R(x) dx = 0, \quad j = 0, \dots, N \quad (4.25)$$

Bei den Finite-Elemente-Methoden werden als Ansatzfunktionen $\Phi_k(x)$ lokal definierte, stückweise von Null verschiedene Funktionen niedriger Ordnung gewählt (engl. shape functions). Gewichtsfunktionen $\omega_j(x)$ können z.B. die Ansatzfunktionen selbst sein (Galerkin-FE-Methode).

Finite-Volumen-Methode

Die Finite-Volumen-Methode kann mit der Methode der gewichteten Residuen hergeleitet werden (Teilbereichsmethode). Als Ansatzfunktionen $\Phi_k(x)$ wählt man stückweise konstante Funktionen. Für das Gebiet G wird eine disjunkte Zerlegung in Teilgebiete G_j vorgenommen. Die Gewichtsfunktionen $\omega_j(x)$ sind definiert als

$$\omega_j(x) = \begin{cases} 1 & \text{für } x \in G \\ 0 & \text{sonst.} \end{cases} \quad (4.26)$$

Das Residuum soll in jedem der Teilgebiete verschwinden,

$$\int_{g_j} R(x) dx = 0.$$

Die Methode arbeitet also mit der integralen Form der Erhaltungsgleichungen. Die Volumenintegrale werden in Oberflächenintegrale umgewandelt. Die Finite-Volumen-Methode ist die im Ingenieurbereich am häufigsten verwendete Methode. Sie bietet den Vorteil, auf komplexe Geometrien anwendbar zu sein. Außerdem ist die Konservativität der Flüsse in der Finite-Volumen-Methode am einfachsten zu gewährleisten.

Zunächst wird das Simulationsgebiet in sog. Kontrollvolumina eingeteilt (siehe Abb. 4.2). In der Mitte eines Kontrollvolumens befindet sich der zugehörige Gitterpunkt. In Abb. 4.2 wird dies für ein zweidimensionales Gitter illustriert. Jedes Kontrollvolumen setzt sich aus vier Quadranten Q1-Q4 zusammen (in drei Dimensionen entsprechend acht Oktanten). Da sich die Materialpunkte jeweils zwischen vier Simulationspunkten befinden (in drei Dimension zwischen acht; vgl. auch Abschnitt 1.3), kann jedem Quadranten ein eigenes Materialsystem zugeordnet werden.

Bilanzgleichung in integraler Form

Die oben dargestellten partiellen Differenzialgleichungen haben die Form

$$\nabla_{\vec{x}} \cdot \hat{\sigma} + \rho \vec{k} = \vec{0} \quad (4.27)$$

Im Fall der elastomechanischen Verspannungen ist $\hat{\sigma}$ ein Tensor zweiter Stufe (Matrix) und $\rho \vec{k}$ ein Tensor erster Stufe (Vektor). Es handelt sich hier demnach um eine Gleichung mit drei Komponenten. Entsprechend ist $\nabla_{\vec{x}}$ als Vektordivergenz aufzufassen mit der Definition

$$(\nabla_{\vec{x}} \cdot \hat{\sigma})_i = \sum_j \frac{\partial}{\partial x_j} \sigma_{ij} \quad (i, j = 1..3) \quad (4.28)$$

Für die Poisson- und Stromgleichung sind die elektrische Flussdichte bzw. die Stromdichte Tensoren erster Stufe (Vektoren), die Quellterme, d.h. die Ladung

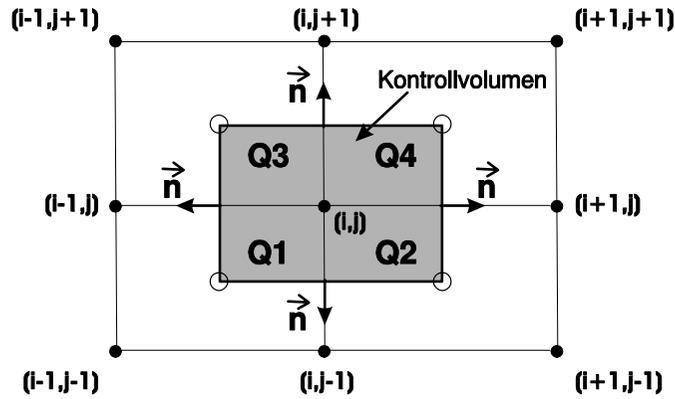


Abbildung 4.2: Kontrollvolumen, unterteilt in vier Quadranten, für den Gitterpunkt (i,j) in zwei Dimensionen. Die Normalenvektoren längs der Oberfläche sind eingezeichnet. Die Kreise bezeichnen jeweils einen Materialgitterpunkt, welcher die Materialparameter für den entsprechenden Quadranten festlegt.

bzw. die Rekombinationsrate, sind Skalare. Dann entspricht die Divergenz der ersten Komponente der Vektordivergenz aus Gl. (4.28), entsprechend reduziert sich Gl. (4.27) auf eine Komponente.

Mit Hilfe des Gaußschen Satzes lässt sich Gl. (4.27) in integraler Form schreiben mit dem Normalenvektor \vec{n} :

$$\int_V \nabla_{\vec{x}} \cdot \hat{\sigma} dV = \int_A \vec{n} \cdot \hat{\sigma} dA = - \int_V \rho \vec{k} dV \quad (4.29)$$

Ausführen der Integration

Das Integral in Gl. (4.29) wird stückweise über die einzelnen vier Quadranten (bzw. acht Oktanten in drei Dimensionen) ausgeführt. In Abb. 4.3 wird dies anhand des ersten Quadranten dargestellt. Im Folgenden soll die entsprechende Gleichung für die erste der drei Komponenten exemplarisch hergeleitet werden. Es gilt:

$$(\vec{n} \cdot \hat{\sigma})_1 = \sigma_{j1} n_j = \sigma_{11} n_1 + \sigma_{21} n_2 + \sigma_{31} n_3$$

Die Normalenvektoren in Abb. 4.3 sind $\vec{n}_1 = (-1, 0, 0)^T$ und $\vec{n}_2 = (0, -1, 0)^T$. Mit $\rho \vec{k} = 0$ folgt dann für die erste Komponente aus Gl. (4.29):

$$\int (\sigma_{11} n_1 + \sigma_{21} n_2 + \sigma_{31} n_3) dA = -\sigma_{11} \frac{h_{2j-1}}{2} - \sigma_{21} \frac{h_{1i-1}}{2} = 0 \quad (4.30)$$

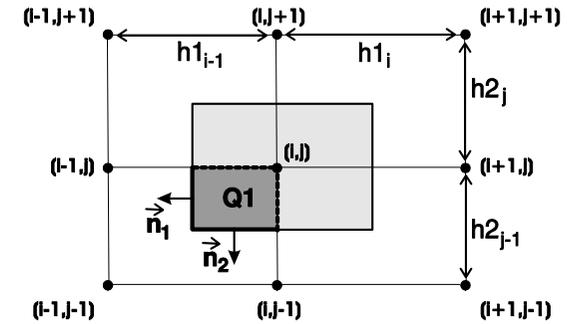


Abbildung 4.3: Integration über Oberfläche des ersten Quadranten $Q1$ am Gitterpunkt (i,j) .

Aufgrund des Hookschen Gesetzes (4.7) gilt nun $\sigma_{ij} = E_{ijkl} \epsilon_{kl}$, so dass in Gl. (4.30) schließlich alle Variationen von ϵ_{kl} auftreten können:

$$\begin{aligned} & - \frac{h_{2j-1}}{2} (E_{1111} \epsilon_{11} + E_{1112} \epsilon_{12} + E_{1113} \epsilon_{13} + E_{1121} \epsilon_{21} \\ & + E_{1122} \epsilon_{22} + E_{1123} \epsilon_{23} + E_{1131} \epsilon_{31} + E_{1132} \epsilon_{32} + E_{1133} \epsilon_{33}) \\ & - \frac{h_{1i-1}}{2} (E_{2111} \epsilon_{11} + E_{2112} \epsilon_{12} + E_{2113} \epsilon_{13} + E_{2121} \epsilon_{21} \\ & + E_{2122} \epsilon_{22} + E_{2123} \epsilon_{23} + E_{2131} \epsilon_{31} + E_{2132} \epsilon_{32} + E_{2133} \epsilon_{33}) = 0 \end{aligned}$$

Flusserhaltende Diskretisierung

Da die Spannungsgleichung drei Komponenten hat, muss der Lösungsvektor pro Gitterpunkt drei Koordinaten aufweisen. Es sind dies die Auslenkungen aus dem unverspannten Zustand (u^1, u^2, u^3) . Die Diskretisierung gemäß Gl. (4.8) sollte flusserhaltend sein, sie lautet in unserem Beispiel bei Integration über die

Teilfläche $A_{\vec{n}_1}$:

$$\begin{aligned}\epsilon_{11} &= \frac{\partial u^1}{\partial x} = \frac{u_{i,j}^1 - u_{i-1,j}^1}{h1_{i-1}} \\ \epsilon_{12} &= \frac{1}{2} \left(\frac{\partial u^1}{\partial y} + \frac{\partial u^2}{\partial x} \right) = \frac{1}{2} \left[\frac{u_{i,j}^1 - u_{i,j-1}^1}{h2_{j-1}} \right. \\ &\quad \left. + \frac{1}{2} \left(\frac{u_{i,j}^2 - u_{i-1,j}^2}{h1_{i-1}} + \frac{u_{i,j-1}^2 - u_{i-1,j-1}^2}{h1_{i-1}} \right) \right]\end{aligned}$$

Bei Integration über die Teilfläche $A_{\vec{n}_2}$ ergibt sich entsprechend:

$$\begin{aligned}\epsilon_{11} &= \frac{\partial u^1}{\partial x} = \frac{1}{2} \left(\frac{u_{i,j}^1 - u_{i-1,j}^1}{h1_{i-1}} + \frac{u_{i,j-1}^1 - u_{i-1,j-1}^1}{h1_{i-1}} \right) \\ \epsilon_{12} &= \frac{1}{2} \left(\frac{\partial u^1}{\partial y} + \frac{\partial u^2}{\partial x} \right) = \frac{1}{2} \left[\frac{1}{2} \left(\frac{u_{i,j}^1 - u_{i,j-1}^1}{h2_{j-1}} + \frac{u_{i-1,j}^1 - u_{i-1,j-1}^1}{h2_{j-1}} \right) \right. \\ &\quad \left. + \frac{u_{i,j}^2 - u_{i-1,j}^2}{h1_{i-1}} \right]\end{aligned}$$

Dadurch erhält man für das resultierende lineare Gleichungssystem eine symmetrische Matrix. Dieses Vorgehen gilt im Prinzip genauso für die Poisson- und Stromgleichung, ist dort allerdings erheblich einfacher. Dabei ist bei der Poissongleichung der Lösungsvektor das Potenzial, bei der Stromgleichung das Quasi-Ferminiveau.

Hier konnten nur die Grundgedanken bei der Diskretisierung der Verspannungsgleichung angegeben werden. In Anhang B sind technische Details beim Aufstellen der Matrix für die Verspannungsgleichung für ein dreidimensionales Problem zu finden.

4.5 Zusammenfassung

In diesem Kapitel wurden kurz die Poisson-, die Strom- und die Verspannungsgleichungen zusammengestellt und mögliche Randbedingungen diskutiert. Numerisch lösen wir diese Gleichungen mit der Finite-Volumen-Methode. Das Diskretisierungsschema wurde beispielhaft anhand der Verspannungsgleichung in zwei Dimensionen erläutert. Im Anhang finden sich detaillierte Informationen über das Aufstellen der entsprechenden Gleichungssysteme. Die Verspannungen führen zu einer Verschiebung und Aufspaltung der Bandenergien. Dieser Effekt wird für die Valenzbänder und für das Γ -, X -, und L -Tal im Leitungsband im Rahmen des van de Walle-Modells berücksichtigt.

Kapitel 5

Numerische Lösung des Gesamtproblems

Das elektrische Verhalten eines Bauelements wird in unserem Modell im Wesentlichen durch vier partielle Differenzialgleichungen beschrieben: die Spannungsgleichung, die Poissongleichung, die Mehrband-kp-Schrödingergleichung und die Stromgleichung. Die Grundlagen der kp-Theorie und das Aufstellen der Hamiltonmatrix wurden in Kapitel 2 und Anhang E erörtert. Die Lösung dieses Eigenwertproblems liefert die Eigenzustände und die Eigenenergien. Für die Berechnung der Ladungsdichte und des Stroms wird das Konzept lokaler Quasi-Fermienergien verwendet (vgl. Kapitel 3). Die daraus resultierende Stromgleichung sowie die Poissongleichung waren Gegenstand von Kapitel 4. Dort wurde auch die Spannungsgleichung erläutert. Wie bereits in Kapitel 1 angedeutet, hängen diese Gleichungen in komplizierter Weise voneinander ab, was eine selbstkonsistente Lösung des Gesamtproblems sehr schwierig macht. Dieser Themenkomplex wird in diesem Kapitel behandelt. Wir gehen dazu in mehreren Schritten vor: Zunächst widmen wir uns der Lösung der klassischen nichtlinearen Poissongleichung und anschließend der selbstkonsistenten Lösung der Schrödinger- und Poissongleichung. Bei Lösung des Nichtgleichgewichtsproblems wird abwechselnd bei festgehaltenem Potenzial das Quasi-Ferminiveau bzw. bei festgehaltenem Quasi-Ferminiveau das Potenzial neu bestimmt. In bestimmten Fällen hat es sich als zweckmäßig erwiesen, die Poisson- und Stromgleichung zu koppeln.

5.1 Mathematische Formulierung des Gleichungssystems

Die zugrunde liegenden Differenzialgleichungen wurden bereits in Kapitel 4 erläutert und sollen hier zur besseren Übersicht noch einmal kurz zusammengestellt werden:

$$\operatorname{div}(\mathbf{E}^{(4)}(\vec{r})\hat{\epsilon}(\vec{r})) = \vec{0} \quad (5.1)$$

$$-\operatorname{div}\left[\varepsilon(\vec{r})\vec{\nabla}\Phi(\vec{r})\right] = -n(\Phi, F_i^\nu, E_{Fn}, \vec{r}) + p(\Phi, F_i^\nu, E_{Fp}, \vec{r}) + N_D^+(\Phi, E_{Fn}, \vec{r}) - N_A^-(\Phi, E_{Fp}, \vec{r}) - \operatorname{div}\vec{P}_{\text{piezo}}(\hat{\epsilon}, \vec{r}) - \operatorname{div}\vec{P}_{\text{pyro}}(\vec{r}) \quad (5.2)$$

$$\sum_{n=1}^8 H_{nn}(\Phi, \vec{r}) F_n^\nu(\vec{r}) = E_\nu F_n^\nu(\vec{r}) \quad (5.3)$$

$$-\operatorname{div}\vec{j}_n(\vec{r}) = \operatorname{div}\left[\hat{\mu}_n(\Phi, \vec{r})n(\Phi, E_{Fn}, \vec{r})\vec{\nabla}E_{Fn}(\vec{r})\right] = R(\Phi, E_{Fn}, E_{Fp}, \vec{r}) \quad (5.4)$$

$$-\operatorname{div}\vec{j}_p(\vec{r}) = -\operatorname{div}\left[\hat{\mu}_p(\Phi, \vec{r})p(\Phi, E_{Fp}, \vec{r})\vec{\nabla}E_{Fp}(\vec{r})\right] = R(\Phi, E_{Fn}, E_{Fp}, \vec{r}) \quad (5.5)$$

In der Spannungsgleichung (5.1) gilt für die Komponenten des Verzerrungstensors

$$\epsilon_{ij} = \frac{1}{2}\left(\frac{\partial u^i}{\partial x_j} + \frac{\partial u^j}{\partial x_i}\right), \quad (5.6)$$

wobei $\vec{u} = (u^1, u^2, u^3)^T$ die Auslenkung eines Atoms aus der Ruhelage bezeichnet. Der 4-stufige Tensor $\mathbf{E}^{(4)}$ ist der Elastizitätstensor. Der piezoelektrische Polarisationsvektor hängt von der Verspannung und den ortsabhängigen piezoelektrischen Konstanten $e_{14}(\vec{r})$ bzw. $e_{15}(\vec{r})$, $e_{31}(\vec{r})$ und $e_{33}(\vec{r})$ ab. Im Kristallsystem gilt:

$$\vec{P}_{\text{piezo}}^{\text{Zinkblende}}(\vec{r}) = e_{14}(\vec{r})\begin{pmatrix} 2\epsilon_{23}(\vec{r}) \\ 2\epsilon_{13}(\vec{r}) \\ 2\epsilon_{12}(\vec{r}) \end{pmatrix} \quad (5.7)$$

$$\vec{P}_{\text{piezo}}^{\text{Wurtzit}}(\vec{r}) = \begin{pmatrix} 2e_{15}(\vec{r})\epsilon_{13}(\vec{r}) \\ 2e_{15}(\vec{r})\epsilon_{12}(\vec{r}) \\ e_{31}(\vec{r})\epsilon_{11}(\vec{r}) + e_{31}(\vec{r})\epsilon_{22}(\vec{r}) + e_{33}(\vec{r})\epsilon_{33}(\vec{r}) \end{pmatrix} \quad (5.8)$$

Pyroelektrische Polarisation $\vec{P}_{\text{pyro}}(\vec{r})$ tritt nur in Wurtzit-Kristallen auf und ist verspannungsunabhängig.

In der Poissongleichung (5.2) bezeichnet $\varepsilon(\vec{r})$ die dielektrische Konstante und $\Phi(\vec{r})$ das elektrische Potenzial. Die Dichte der ionisierten Donatoren bzw. Akzeptoren lautet:

$$N_D^+(\vec{r}) = \frac{N_D(\vec{r})}{1 + g_D e^{[E_{Fn}(\vec{r}) - E_D(\vec{r})]/k_B T}} \quad (5.9)$$

$$N_A^-(\vec{r}) = \frac{N_A(\vec{r})}{1 + g_A e^{-[E_{Fp}(\vec{r}) - E_A(\vec{r})]/k_B T}} \quad (5.10)$$

Hier sind $N_{D/A}(\vec{r})$ die Dotierkonzentrationen für Donatoren bzw. Akzeptoren, $E_{Fn/p}(\vec{r})$ das ortsabhängige Quasi-Ferminiveau der Elektronen bzw. der Löcher,

k_B die Boltzmannkonstante, T die Temperatur und $E_{D/A}(\vec{r})$ die Energieniveaus der Donator- bzw. Akzeptorzustände. Letztere sind bestimmt durch die Ionisierungsenergien $E_{D/A}^{ion}$ der Donatoren bzw. Akzeptoren, die Leitungs- bzw. Valenzbandkante im Volumenhalbleiter E_{C0} bzw. E_{V0} und das elektrische Potenzial Φ . E_{C0} und E_{V0} beinhalten bereits die Verschiebung der Bandkanten durch die Verspannungen:

$$\begin{aligned} E_D(\vec{r}) &= E_{C0}(\vec{r}) - e\Phi(\vec{r}) - E_D^{ion} \\ E_A(\vec{r}) &= E_{V0}(\vec{r}) - e\Phi(\vec{r}) + E_A^{ion} \end{aligned} \quad (5.11)$$

Die g -Faktoren sind $g_D = 2$ für Donatoren sowie $g_A = 4$ für Akzeptoren [54]. In den Hamiltonoperator $H_{mn}(\Phi, \vec{r})$ in der Achtband-kp-Schrödingergleichung (5.3) geht das elektrische Potenzial $\Phi(\vec{r})$ ein, infolgedessen hängen die Eigenwerte E_ν und die dazugehörigen Spinoren der Einhüllenden-Wellenfunktion F_n^ν ($n = 1, \dots, 8$) vom Potenzial ab.

Für die Beweglichkeit $\hat{\mu}_{n/p}(\Phi, \vec{r})$ in der Stromgleichung der Elektronen (Gl. 5.4) bzw. Löcher (Gl. 5.5) existieren mehrere Modelle, die eine Abhängigkeit vom lokalen elektrischen Feld und der Dotierkonzentration beinhalten. Die Rekombinationsraten $R(\Phi, E_{Fn}, E_{Fp}, \vec{r})$ werden durch die lokalen Elektronen- und Löcherdichten, letztendlich also durch das Potenzial und die Ferminiveaus bestimmt. In Anhang H sind Modelle für die Beweglichkeiten und Rekombinationsraten angegeben.

Dichteberechnung

Die Ladungsträgerdichte ergibt sich durch Besetzen der elektronischen Zustände mit dem lokalen Quasi-Fermi-niveau. Wir geben hier nur den Ausdruck für die Achtband-kp-Zustände in drei Dimensionen an. Dann hängt die Dichte nur vom lokalen Quasi-Fermi-niveau sowie den Spinoren und ihren Eigenwerten ab. Da allgemein für jedes Band aber auch die klassische Thomas-Fermi-Dichte spezifiziert werden kann, besteht allgemein zusätzlich eine Abhängigkeit vom Potenzial.

$$\begin{aligned} n(\Phi, E_{Fn}, \vec{r}) &= \sum_{\nu \in \text{Elektronen}} \left[\sum_{n=1}^8 |F_n^\nu(\vec{r})|^2 f(E_{Fn}(\vec{r}) - E_\nu) \right] \\ p(\Phi, E_{Fp}, \vec{r}) &= \sum_{\nu \in \text{Löcher}} \left[\sum_{n=1}^8 |F_n^\nu(\vec{r})|^2 f(E_\nu - E_{Fp}(\vec{r})) \right] \end{aligned} \quad (5.12)$$

In Anhang D sind nähere Details zur Dichteberechnung zu finden.

Lösungsvektor

Das Gleichungssystem lässt sich in drei Blöcke gliedern. Der erste beinhaltet die Spannungsgleichung Gl. (5.1), deren Lösungsvektor für jeden Gitterpunkt

(i, j, k) die Auslenkung der Atome aus der Ruhelage in x -, y - und z -Richtung enthält: $\vec{u}_{ijk} = (u_{ijk}^1, u_{ijk}^2, u_{ijk}^3)$. Der zweite Block umfasst die Poisson- und die Mehrband-kp-Schrödingergleichung (Gl. 5.2 und Gl. 5.3). Nimmt man die Quasi-Fermi-niveaus als konstant an, so ist die selbstkonsistente Lösung dieser beiden Gleichungen durch das elektrische Potenzial eindeutig bestimmt. Aus den Stromgleichungen (Gl. 5.4 und Gl. 5.5) ergeben sich schließlich für festgehaltenes Potenzial die Quasi-Fermi-niveaus für die Elektronen und die Löcher.

Da die Verspannung in unserem Modell weder von Potenzial noch von den Quasi-Fermi-niveaus abhängig ist, wird als Erstes die Spannungsgleichung gelöst. Im Folgenden kann man sich dann auf die selbstkonsistente Lösung des zweiten und dritten Blocks beschränken.

Eingebautes Potenzial (engl. built-in potential)

Der Bandkantenverlauf im thermodynamischen Gleichgewicht ist durch das eingebaute Potenzial bestimmt. Dieses wird zunächst berechnet, bevor angelegte Spannungen berücksichtigt werden. Man erhält es durch selbstkonsistentes Lösen der Poisson- und Schrödingergleichung bei einheitlichem Fermi-niveau für Elektronen und Löcher. Als Randbedingung wird dabei für ohmsche Kontakte das elektrische Feld auf Null gesetzt. Bei Schottky-Kontakten nehmen wir an, dass die Leitungsbandkante relativ zum Fermi-niveau fixiert ist (vgl. Kapitel 4).

5.2 Selbstkonsistente Lösung der Schrödinger- und Poissongleichung

5.2.1 Nichtlineare Poissongleichung

In diesem Abschnitt wird kurz die nichtlineare Poissongleichung mit klassischer Thomas-Fermi-Dichte diskutiert, da deren Lösung immer als Startwert für die weitere Rechnung verwendet wird. Der Einfachheit halber wird hier nur ein Leitungs- und ein Valenzband berücksichtigt. Wir nehmen an, dass sich die Ladungsträger im globalen thermischen Gleichgewicht mit dem Fermi-niveau E_F befinden, dann erhält man durch Diskretisierung der Poissongleichung folgendes Problem:

$$\mathbf{A}\vec{\Phi} - \vec{b}(\vec{\Phi}) = 0, \quad (5.13)$$

wo \mathbf{A} die $N \times N$ Poissonmatrix und N die Zahl der Gitterpunkte ist. Für ein dreidimensionales $50 \times 50 \times 50$ -Gitter ist also N gleich $50^3 = 125.000$. $\vec{\Phi} = (\Phi_1, \dots, \Phi_N)^T$ bezeichnet einen N -dimensionaler Vektor, dessen Komponente i den Wert des Potenzials am Gitterpunkt i enthält. Für die integrierte Ladung gilt:

$$b_i = e^2(p_i - n_i + N_{D,i}^+ - N_{A,i}^- - C)V_i \quad (5.14)$$

Dabei stellt V_i das Volumen der Box mit Nummer i dar (vg. Kapitel 4.1 und Anhang A). Die Polarisationsladungen sind in der Konstante C enthalten, die uns aber nicht weiter interessiert, da sie unabhängig vom Potenzial ist. Im Einzelnen gilt (vgl. auch Gl. 5.2):

$$n_i = 2 \left(\frac{2\pi k_B T m_{n,i}^*}{h^2} \right)^{3/2} F_{1/2} \left[\frac{-E_{c0,i} + e\Phi_i + E_F}{k_B T} \right] \quad (5.15)$$

$$p_i = 2 \left(\frac{2\pi k_B T m_{p,i}^*}{h^2} \right)^{3/2} F_{1/2} \left[\frac{E_{v0,i} - e\Phi_i - E_F}{k_B T} \right] \quad (5.16)$$

$$N_{D,i}^+ = \frac{N_{D,i}}{1 + g_D e^{(E_F - E_{D,i})/k_B T}} \quad (5.17)$$

$$N_{A,i}^- = \frac{N_{A,i}}{1 + g_A e^{-(E_F - E_{A,i})/k_B T}} \quad (5.18)$$

$F_{1/2}$ bedeutet hier das Fermi-Integral der Ordnung $1/2$ [55][56][57]. Diese werden in Anhang F erläutert. Man beachte, dass alle vier Terme in den Gl. (5.15)-(5.18) vom Potenzial abhängen (vgl. auch Gl. 5.11). Die Lösung von Gl. (5.13) entspricht daher der Nullstellensuche des nichtlinearen Funktionalen $\vec{F}[\vec{\Phi}] = \mathbf{A}\vec{\Phi} - \vec{b}(\vec{\Phi})$.

5.2.2 Newton-Verfahren

Zur Lösung nichtlinearer Gleichungen wird ein verallgemeinertes Newtonverfahren verwendet. Dieses dient allgemein zur Nullstellenbestimmung einer Funktion $\vec{F}(\vec{x}) = 0$ [58]. Zunächst führt man eine Taylorentwicklung durch:

$$0 = \vec{F}(\vec{x}_0 + d\vec{x}) = \vec{F}(\vec{x}_0) + \frac{\partial \vec{F}}{\partial \vec{x}} d\vec{x}$$

wobei $\mathbf{DF} = \frac{\partial \vec{F}}{\partial \vec{x}}$ die Jacobimatrix bezeichnet. Daraus ergibt sich die Newtonkorrektur:

$$d\vec{x} = -(\mathbf{DF})^{-1} \times \vec{F} \quad (5.19)$$

Da hier Konvergenz nur für sehr gute Startwerte erreichbar ist, wird das Verfahren um eine zusätzliche Komponente erweitert [58]. Bei jedem Newtonschritt minimiert man die skalare Funktion $f = \frac{1}{2} \vec{F}^T \vec{F}$ entlang des Korrekturvektors $d\vec{x}$ (Linienminimierung). An diesem Minimum wird ein neuer Korrekturvektor ermittelt. Es handelt sich hier also um eine Mischstrategie. Solange man sich weit weg von der Lösung befindet, gibt die Newtonrichtung die Richtung des steilsten Abstiegs vor. In dieser Richtung wird Minimierung des Funktionalen durchgeführt. Dadurch ist gewährleistet, dass man überhaupt erst in die Nähe der Lösung gelangt. Befindet man sich schließlich genügend nahe der Nullstelle, greift das Newtonverfahren, welches quadratische Konvergenz aufweist. Die Jacobimatrix ist hier eine $N \times N$ -Matrix. Die Lösung der Gleichung (5.19) wird mit Hilfe von konjugierten Gradientenverfahren [59] berechnet.

5.2.3 Schrödinger-Poisson

Im vorigen Abschnitt haben wir uns auf die klassische Thomas-Fermi-Dichte beschränkt. In Wirklichkeit ergibt sich die Dichte jedoch aus den quantenmechanischen Zuständen, die über das lokale Quasi-Ferminiveau besetzt werden. Wenn das Quasi-Ferminiveau als fest angenommen wird - im Gleichgewichtsfall ist es außerdem räumlich konstant - besteht unsere Aufgabe nun darin, die Poissongleichung selbstkonsistent mit der Schrödingergleichung zu lösen. Das Problem ist hierbei, dass die quantenmechanischen Zustände über eine Eigenwertgleichung vom Potenzial abhängen. Konkret wollen wir hier folgendes Gleichungssystem lösen (vgl. Gl. 5.2, 5.3):

$$-\text{div} \left[\varepsilon(\vec{r}) \vec{\nabla} \Phi(\vec{r}) \right] = -n(\Phi, F_i^\nu, \vec{r}) + p(\Phi, F_i^\nu, \vec{r}) \quad (5.20)$$

$$+ N_D^+(\Phi, \vec{r}) - N_A^-(\Phi, \vec{r})$$

$$- \text{div} \vec{P}_{\text{piezo}}(\vec{r}) - \text{div} \vec{P}_{\text{pyro}}(\vec{r})$$

$$\sum_{n=1}^8 H_{mn}(\Phi, \vec{r}) F_n^\nu(\vec{r}) = E_\nu F_n^\nu(\vec{r}) \quad (5.21)$$

Da das Quasi-Ferminiveau festgehalten wird, tritt es nicht als Argument in den Variablen auf. Die Dichten n und p errechnen sich gemäß Gl. (5.12), die Besetzung der Störstellen ergibt sich aus Gl. (5.9).

Zur Lösung derartiger Gleichungssysteme haben sich zwei Verfahren bewährt:

1. Relaxationsverfahren [60]
2. Prädiktor-Korrektor-Verfahren [61]

Prädiktor-Korrektor-Verfahren

Diese Methode hat sich als am besten erwiesen und wird daher standardmäßig verwendet. Die Idee besteht darin, eine möglichst genaue Näherung für die Quantendichte als Funktion des Potenzials zu finden und damit die nichtlineare Poissongleichung zu lösen. Aus dem resultierenden Potenzial werden dann die Energie-Eigenwerte neu berechnet und daraus eine neue Darstellung der Quantendichte ermittelt usw. Man nimmt dabei an, dass bei geringfügigen Potenzialänderungen die Wellenfunktionen konstant bleiben und die Eigenwerte entsprechend der Potenzialverschiebung variieren. Es zeigt sich, dass folgender Ausdruck den Anforderungen an die Prädiktor-Quantendichte genügt (mit festen Eigenfunktio-

nen F_n^ν und Eigenwerten E_ν):

$$\begin{aligned}\tilde{n}_q[\Phi + \delta\Phi] &= \sum_{\nu \in \text{Elektronen}} \left[\sum_{n=1}^8 |F_n^\nu(\vec{r})|^2 f(E_{F_n}(\vec{r}) - E_\nu + e\delta\Phi(\vec{r})) \right] \\ \tilde{p}_q[\Phi + \delta\Phi] &= \sum_{\nu \in \text{Löcher}} \left[\sum_{n=1}^8 |F_n^\nu(\vec{r})|^2 f(-E_{F_p}(\vec{r}) + E_\nu - e\delta\Phi(\vec{r})) \right]\end{aligned}$$

Bezeichnet $\Phi^{(k)}$ das Potenzial aus dem vorhergehenden Iterationsschritt, für welches das Eigenwertproblem gelöst wurde, so folgt:

$$\begin{aligned}\tilde{n}_q[\Phi] &= \sum_{\nu \in \text{Elektronen}} \left[\sum_{n=1}^8 |F_n^{\nu(k)}(\vec{r})|^2 f(E_{F_n}(\vec{r}) - E_\nu^{(k)} + e(\Phi(\vec{r}) - \Phi^{(k)}(\vec{r}))) \right] \\ \tilde{p}_q[\Phi] &= \sum_{\nu \in \text{Löcher}} \left[\sum_{n=1}^8 |F_n^{\nu(k)}(\vec{r})|^2 f(-E_{F_p}(\vec{r}) + E_\nu^{(k)} - e(\Phi(\vec{r}) - \Phi^{(k)}(\vec{r}))) \right]\end{aligned}\tag{5.22}$$

Mit diesem Ausdruck für die Dichte wird nun die nichtlineare Poissongleichung gelöst. Dabei benützt man am besten ein Newtonverfahren mit Zeilenminimierung.

Insgesamt ergibt sich folgendes Vorgehen:

- Löse nichtlineare Poissongleichung mit dem Ausdruck (5.22) für die Dichte:

$$-\text{div}(\varepsilon \text{grad } \Phi) = \rho[\tilde{n}_q[\Phi; \Phi^{(k)}, F_n^{\nu(k)}, E_\nu^{(k)}], \tilde{p}_q[\Phi; \Phi^{(k)}, F_n^{\nu(k)}, E_\nu^{(k)}], \Phi]$$

Das so berechnete Potenzial Φ entspricht dem neuen Potenzial $\Phi^{(k+1)}$.

- Löse Mehrband-kp-Schrödingergleichung für das neue Potenzial:

$$\sum_{n=1}^8 H_{mn}(\Phi^{(k+1)}, \vec{r}) F_n^\nu(\vec{r}) = E_\nu F_n^\nu(\vec{r})$$

Daraus erhält man die Eigenzustände und Energien für den nächsten Iterationsschritt: $E_\nu^{(k+1)} = E_\nu$, $F_n^{\nu(k+1)}(\vec{r}) = F_n^\nu(\vec{r})$.

- Abbruchkriterium:

$$\|\Phi^{(k+1)} - \Phi^{(k)}\|_2 \leq \delta$$

5.3 Lösung des Nichtgleichgewichtsproblems

Wie bereits dargelegt, wird die selbstkonsistente Lösung des Systems (5.1)-(5.5) durch die Auslenkung der Teilchen aus der Ruhelage für jeden der N Gitterpunkte ($\vec{u}_{ijk} = (u_{ijk}^1, u_{ijk}^2, u_{ijk}^3)^T$, $ijk = 1, \dots, N$), das elektrische Potenzial $\vec{\Phi} = (\Phi_1, \dots, \Phi_N)^T$ und das Quasi-Ferminiveau der Elektronen $\vec{E}_{F_n} = (E_{F_n,1}, \dots, E_{F_n,N})^T$ bzw. der Löcher $\vec{E}_{F_p} = (E_{F_p,1}, \dots, E_{F_p,N})^T$ eindeutig bestimmt. Da die Verspannung in unserem Modell unabhängig ist vom Potenzial und den Quasi-Ferminiveaus, kann die Spannungsgleichung absepariert werden, und das Problem reduziert sich auf die Gln. (5.2)-(5.5). Dieses System lösen wir entweder blockiterativ oder gekoppelt.

5.3.1 Blockiteratives Verfahren

Beim blockiterativen Verfahren werden abwechselnd die Schrödinger- und Poissongleichung bei festgehaltenem Quasi-Ferminiveau und die Stromgleichungen bei festgehaltenem Potenzial und festen quantenmechanischen Zuständen gelöst. Der Lösungsalgorithmus lautet demnach wie folgt:

- Löse für gegebenes Potenzial $\vec{\Phi} = (\Phi_1, \dots, \Phi_N)^T$ die Stromgleichungen (5.4) und (5.5). Wenn man die Dichten als Parameter ansieht, so hat man lineare Gleichungen für die Quasi-Ferminiveaus. Daraus können direkt $\vec{E}_{F_n} = (E_{F_n,1}, \dots, E_{F_n,N})$ und $\vec{E}_{F_p} = (E_{F_p,1}, \dots, E_{F_p,N})$ bestimmt werden. Dann werden die Dichten und die Rekombinationsraten neu berechnet. Dieses Vorgehen wird solange wiederholt, bis Konvergenz der Quasi-Ferminiveaus erreicht ist. Das entsprechende Relaxationsverfahren lautet (k bezeichnet den Iterationsschritt):

1. Berechne für gegebene $\vec{E}_{F_n}^{(k)}$ und $\vec{E}_{F_p}^{(k)}$ die Elektronen- und Löcherdichten $n^{(k)}$ und $p^{(k)}$ sowie die Rekombinationsrate $R^{(k)}$ und löse damit die linearen Stromgleichungen:

$$\begin{aligned}-\text{div} \left[\hat{\mu}_n n^{(k)} \vec{\nabla} E_{F_n}(\vec{r}) \right] &= -R^{(k)} \\ -\text{div} \left[\hat{\mu}_p p^{(k)} \vec{\nabla} E_{F_p}(\vec{r}) \right] &= R^{(k)}\end{aligned}\tag{5.23}$$

2. Relaxiere das Quasi-Ferminiveau mit dem Relaxationsparameter ω :

$$\begin{aligned}\vec{E}_{F_n}^{(k+1)} &= \omega \vec{E}_{F_n} + (1 - \omega) \vec{E}_{F_n}^{(k)} \\ \vec{E}_{F_p}^{(k+1)} &= \omega \vec{E}_{F_p} + (1 - \omega) \vec{E}_{F_p}^{(k)}\end{aligned}\tag{5.24}$$

3. Wiederhole ab Schritt 1, bis

$$\begin{aligned}\|\vec{E}_{F_n}^{(k+1)} - \vec{E}_{F_n}^{(k)}\|_2 &\leq \epsilon \\ \|\vec{E}_{F_p}^{(k+1)} - \vec{E}_{F_p}^{(k)}\|_2 &\leq \epsilon.\end{aligned}\tag{5.25}$$

- Löse die Poisson- und Schrödingergleichung (Gln. 5.2 und 5.3) selbstkonsistent für gegebene Quasi-Ferminiveaus $\vec{E}_{Fn} = (E_{Fn,1}, \dots, E_{Fn,N})$ und $\vec{E}_{Fp} = (E_{Fp,1}, \dots, E_{Fp,N})$, wie im vorigen Abschnitt beschrieben.
- Abbruchkriterium:

$$\begin{aligned} \|\vec{\Phi}^{(neu)} - \vec{\Phi}^{(alt)}\|_2 &\leq \delta \\ \|\vec{E}_{Fn}^{(neu)} - \vec{E}_{Fn}^{(alt)}\|_2 &\leq \delta \\ \|\vec{E}_{Fp}^{(neu)} - \vec{E}_{Fp}^{(alt)}\|_2 &\leq \delta \end{aligned}$$

5.3.2 Gekoppelte Lösung

Das blockiterative Verfahren kann unter Umständen zu Konvergenzproblemen führen. Dann ist es zweckmäßig, die Poisson- und die Stromgleichungen zu koppeln, d.h. man fasst die Gleichungen (5.2), (5.4) und (5.5) zu einer zusammen. Daneben muss aber noch die Schrödingergleichung (5.3) erfüllt sein. Dabei gehen wir vor wie beim Prädiktor-Korrektor-Verfahren. Pro Iterationsschritt wird das Eigenwertproblem gelöst und die quantenmechanischen Zustände werden festgehalten. Anschließend erfolgt die Berechnung des gekoppelten Systems mit der störungstheoretischen Prädiktor-Dichte (5.22). Dies entspricht der Nullstellensuche eines nun $3N$ -dimensionalen Funktionals \vec{F} :

$$\vec{F} \left[\vec{\Phi}, \vec{E}_{Fn}, \vec{E}_{Fp} \right] = \begin{pmatrix} -\operatorname{div} \left[\varepsilon(\vec{r}) \vec{\nabla} \Phi(\vec{r}) \right] - \rho \left(\vec{\Phi}, \vec{E}_{Fn}, \vec{E}_{Fp} \right) \\ -\operatorname{div} \left[\hat{\mu}_n \tilde{n}_q \left(\vec{\Phi}, \vec{E}_{Fn}, \vec{E}_{Fp} \right) \vec{\nabla} E_{Fn}(\vec{r}) \right] + R \left(\vec{\Phi}, \vec{E}_{Fn}, \vec{E}_{Fp} \right) \\ -\operatorname{div} \left[\hat{\mu}_n \tilde{p}_q \left(\vec{\Phi}, \vec{E}_{Fn}, \vec{E}_{Fp} \right) \vec{\nabla} E_{Fp}(\vec{r}) \right] - R \left(\vec{\Phi}, \vec{E}_{Fn}, \vec{E}_{Fp} \right) \end{pmatrix} \quad (5.26)$$

In dieser Darstellung steht jede Zeile für einen N -dimensionalen Vektor. Die Diskretisierung der Ausdrücke $\operatorname{div}[\varepsilon(\vec{r}) \operatorname{grad} \Phi(\vec{r})]$ usw. wurde in Kapitel 4.1 und Anhang A erläutert. Für die erste Zeile erhält man z.B. einen Ausdruck der Form $\mathbf{A}\vec{\Phi} - \vec{b}(\vec{\Phi})$, wo \mathbf{A} die Poissonmatrix bedeutet und $b_i = \rho_i V_i$ (V_i das Volumen der Box i). Die Matrix für die Stromgleichung $\mathbf{M}_{n/p} \left(\vec{\Phi}, \vec{E}_{Fn}, \vec{E}_{Fp} \right)$ hängt selbst von der Dichte und damit vom Potenzial und den Quasi-Ferminiveaus ab. Für die Dichten gilt, wenn $\Phi^{(k)}$ das Potenzial mit den dazugehörigen quantenmechanischen Zuständen $F_n^{\nu(k)}(\vec{r})$ aus dem vorhergehenden Iterationsschritt ist:

$$\begin{aligned} \tilde{n}_q[\Phi, E_{Fn}] &= \sum_{\nu \in \text{Elektronen}} \left[\sum_{n=1}^8 |F_n^{\nu(k)}(\vec{r})|^2 f(E_{Fn}(\vec{r}) - E_\nu^{(k)} + e(\Phi(\vec{r}) - \Phi^{(k)}(\vec{r}))) \right] \\ \tilde{p}_q[\Phi, E_{Fp}] &= \sum_{\nu \in \text{Löcher}} \left[\sum_{n=1}^8 |F_n^{\nu(k)}(\vec{r})|^2 f(-E_{Fp}(\vec{r}) + E_\nu^{(k)} - e(\Phi(\vec{r}) - \Phi^{(k)}(\vec{r}))) \right] \end{aligned}$$

$$\begin{aligned} \rho(\Phi, E_{Fn}, E_{Fp}) &= -\tilde{n}_q[\Phi, E_{Fn}] + \tilde{p}_q[\Phi, E_{Fp}] + N_D^+[\Phi, E_{Fn}] - N_A^-[\Phi, E_{Fp}] \\ &\quad - \operatorname{div} \vec{P}_{piezo}(\vec{r}) - \operatorname{div} \vec{P}_{pyro}(\vec{r}) \end{aligned}$$

Dann kann das Funktional geschrieben werden als:

$$\vec{F} \left[\vec{\Phi}, \vec{E}_{Fn}, \vec{E}_{Fp} \right] = \begin{pmatrix} \mathbf{A}\vec{\Phi} - \vec{b} \left(\vec{\Phi}, \vec{E}_{Fn}, \vec{E}_{Fp} \right) \\ \mathbf{M}_n \left(\vec{\Phi}, \vec{E}_{Fn} \right) \vec{E}_{Fn} + \vec{R} \left(\vec{\Phi}, \vec{E}_{Fn}, \vec{E}_{Fp} \right) \\ \mathbf{M}_p \left(\vec{\Phi}, \vec{E}_{Fp} \right) \vec{E}_{Fp} - \vec{R} \left(\vec{\Phi}, \vec{E}_{Fn}, \vec{E}_{Fp} \right) \end{pmatrix} \quad (5.27)$$

mit $\vec{\rho} = (\rho_1, \dots, \rho_N)$ und $\vec{R} = (R_1, \dots, R_N)$.

Für die Nullstellensuche verwenden wir wieder das Newtonverfahren mit Zeilenminimierung. Dazu ist die Berechnung der Jacobi-Matrix nötig (vgl. Abschnitt 5.2). Diese lautet: $\mathbf{Df} = \frac{\partial \vec{F}}{\partial \vec{x}}$ mit $\vec{x} = \left(\vec{\Phi}, \vec{E}_{Fn}, \vec{E}_{Fp} \right)^T$.

Bei der Berechnung der Ableitungen ist zu beachten, dass die Strommatrix selbst eine Funktion des Quasi-Ferminiveaus ist. Betrachten wir dazu das Funktional $\vec{f}[\vec{E}_F] = \mathbf{M}(\vec{E}_F) \vec{E}_F$. Dann ist die Jakobimatrix $\mathbf{Df} = \frac{\partial \vec{f}}{\partial \vec{E}_{Fn}}$ definiert als:

$$\begin{aligned} (\mathbf{Df})_{ij} &= \frac{\partial f_i}{\partial E_{F,j}} = \frac{\partial}{\partial E_{F,j}} \sum_k \mathbf{M}_{ik} E_{F,k} = \\ &= \sum_k \mathbf{M}_{ik} \left(\frac{\partial}{\partial E_{F,j}} E_{F,k} \right) + \sum_k \left(\frac{\partial}{\partial E_{F,j}} \mathbf{M}_{ik} \right) E_{F,k} = (\mathbf{M})_{ij} + (\mathbf{D}_{E_F} \mathbf{M})_{ij}, \end{aligned} \quad (5.28)$$

wo die Definition

$$(\mathbf{D}_{E_F} \mathbf{M})_{ij} = \sum_k \left(\frac{\partial}{\partial E_{F,j}} \mathbf{M}_{ik} \right) E_{F,k}$$

eingeführt wurde. Mit den Bezeichnungen in Gl. (5.27) gilt nun:

$$\mathbf{DF} = \begin{pmatrix} \mathbf{A} - \frac{\partial \vec{b}}{\partial \vec{\Phi}} & -\frac{\partial \vec{b}}{\partial \vec{E}_{Fn}} & -\frac{\partial \vec{b}}{\partial \vec{E}_{Fp}} \\ \mathbf{D}_\Phi \mathbf{M}_n + \frac{\partial \vec{R}}{\partial \vec{\Phi}} & \mathbf{M}_n + \mathbf{D}_{E_{Fn}} \mathbf{M}_n + \frac{\partial \vec{R}}{\partial E_{Fn}} & \frac{\partial \vec{R}}{\partial \vec{E}_{Fp}} \\ \mathbf{D}_\Phi \mathbf{M}_p - \frac{\partial \vec{R}}{\partial \vec{\Phi}} & -\frac{\partial \vec{R}}{\partial \vec{E}_{Fn}} & \mathbf{M}_p + \mathbf{D}_{E_{Fp}} \mathbf{M}_p - \frac{\partial \vec{R}}{\partial \vec{E}_{Fp}} \end{pmatrix} \quad (5.29)$$

Gesamtalgorithmus

Der Gesamtalgorithmus lautet nun:

- Löse das Funktional (5.27) für die quantenmechanischen Zustände aus dem vorhergehenden Iterationsschritt $F_n^{\nu(k)}(\vec{r})$, daraus ergibt sich der neue Vektor $\vec{x}^{(k+1)} = \left(\vec{\Phi}^{(k+1)}, \vec{E}_{Fn}^{(k+1)}, \vec{E}_{Fp}^{(k+1)} \right)^T$.

- Löse Mehrband-kp-Schrödingergleichung für das neue Potenzial:

$$\sum_{n=1}^8 H_{mn} (\Phi^{(k+1)}, \vec{r}) F_n^\nu(\vec{r}) = E_\nu F_n^\nu(\vec{r})$$

Daraus erhält man die Eigenzustände und Energien für den nächsten Iterationsschritt: $E_\nu^{(k+1)} = E_\nu$, $F_n^{\nu(k+1)}(\vec{r}) = F_n^\nu(\vec{r})$.

- Abbruchkriterium:

$$\|\vec{\Phi}^{(k+1)} - \vec{\Phi}^{(k)}\|_2 \leq \epsilon$$

Klassische Lösung

Wird die Dichte rein klassisch, d.h. mit Thomas-Fermi-Näherung berechnet, so reduziert sich das Gesamtproblem auf die Lösung des Funktionals (5.26). Da in den Ausdruck für die Dichte keine quantenmechanischen Zustände eingehen, ist dieser identisch mit der Prädiktor-Dichte in Gl. (5.22).

Unterprobleme

Ebenso ist es möglich, nur Teile gekoppelt zu lösen, z.B. das Potenzial und das Quasi-Ferminiveau der Elektronen. Dann lautet der $2N$ -dimensionale Lösungsvektor $\vec{x} = (\vec{\Phi}, \vec{E}_{Fn})^T$. Entsprechend reduziert sich die Jacobi-Matrix auf eine $2N \times 2N$ -Matrix. Für den Fall $\vec{x} = \vec{\Phi}$ ist der Algorithmus äquivalent zur Lösung der Schrödinger- und Poissongleichung mit dem Prädiktor-Korrektor-Verfahren.

5.4 Zusammenfassung

In diesem Kapitel wurden Verfahren vorgestellt, mit denen sich das gekoppelte System der Halbleitergleichungen lösen lässt. Da die Verspannungen nicht vom Potenzial und den Quasi-Ferminiveaus abhängen, kann man die Spannungsgleichung abseparieren und ganz zu Beginn als lineares Gleichungssystem lösen. Daraufhin wird das eingebaute Potenzial für das thermodynamische Gleichgewicht berechnet durch selbstkonsistente Lösung der Schrödinger- und Poissongleichung mit entsprechenden Randbedingungen. Nach diesen Vorarbeiten muss nun das allgemeine Problem behandelt werden. Die Lösung für den Nichtgleichgewichtsfall ist durch das Potenzial und die Quasi-Ferminiveaus eindeutig bestimmt. Die quantenmechanischen Zustände hängen über die Schrödingergleichung vom Potenzial ab. Als numerische Verfahren wurden zwei Methoden vorgestellt. Zunächst kann das Schrödinger-Poisson-System und das Stromproblem blockiterativ gelöst werden. Dies führt mitunter aber zu Konvergenzproblemen, so dass die Lösung eines gekoppelten Systems nötig wird.

Kapitel 6

Simulation einzelner Quantenpunkte

Quantenpunkte sind nulldimensionale Systeme und daher sowohl für die Grundlagenforschung als auch im Hinblick auf Anwendungen wie Fotodioden und Quantenpunktlaser von großem Interesse. Es gibt bisher bereits eine ganze Reihe von theoretischen Untersuchungen der quantenmechanischen Zustände in einem Quantenpunkt, welche meist eine Pyramidenform zugrunde legen. Neuere experimentelle Ergebnisse zeigen jedoch, dass diese Annahme nicht aufrecht erhalten werden kann. Im Folgenden richten wir unser Interesse auf überwachsene InAs-GaAs-Quantenpunkte, welche durch selbstorganisiertes Wachstum im so genannten Stranski-Krastanov-Modus hergestellt wurden. Dabei werden mit Molekularstrahl-Epitaxie einige Monolagen InAs auf das GaAs-Substrat aufgebracht. Bei einer bestimmten Schichtdicke kommt es zur spontanen Ausbildung von Quantenpunkten aufgrund der unterschiedlichen Gitterkonstanten von GaAs und InAs (ca. 7%). Form und Größe dieser Quantenpunkte hängen sehr stark von den Wachstumsbedingungen ab. Die Quantenpunktschicht wird daraufhin mit GaAs überwachsen. Dabei verändern sie durch Diffusionsprozesse ihre Form und Indium-Konzentration.

In diesem Abschnitt wollen wir systematisch die elektronischen Eigenschaften von Quantenpunkten unter angelegter Spannung, d.h. im Nichtgleichgewicht, betrachten. Dazu untersuchen wir die optische Übergangsenergie in Abhängigkeit von der Form der Quantenpunkte und der räumlichen Verteilung der Indium-Konzentration. Daraus können Rückschlüsse auf ein inneres elektrisches Feld gezogen werden. Weiterhin erlaubt die kp-Theorie sofort die Bestimmung der optischen Übergangswahrscheinlichkeiten in Abhängigkeit von der Polarisierung des eingestrahelten Lichts. Da die Geometrie der Quantenpunkte in der Regel nicht bekannt ist, stellt es unser Hauptanliegen dar, gemessene Daten in Einklang zu bringen mit Annahmen über die Form der betrachteten Quantenpunkte.

Zunächst legen wir experimentelle Ergebnisse, die von der Gruppe von M. S. Skolnick erzielt wurden, zugrunde und untersuchen ihre elektronische Struktur.

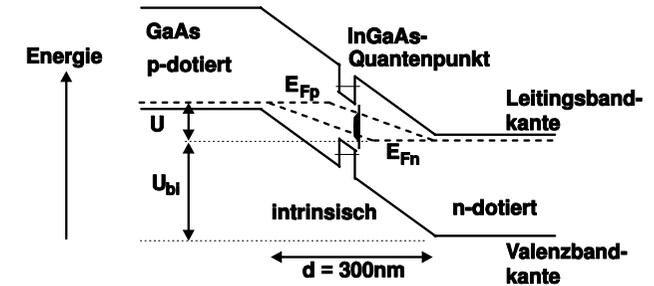


Abbildung 6.1: Banddiagramm der Probe, die von Fry et al. gemessen wurde. Es handelt sich um InGaAs-Quantenpunkte, die in (001)-Richtung auf GaAs gewachsen wurde. Die Quantenpunktschicht ist eingebettet in die intrinsische Region (Weite d) einer pin-Diode. Wird eine Spannung U angelegt, so liegt in der Quantenpunktschicht ein elektrisches Feld von $E = \frac{U+U_{bi}}{d}$ an. U_{bi} ist das eingebaute Potenzial und beträgt ca. 1.5V. Die Quasi-Fermienergie für die Elektronen E_{Fn} und die Löcher E_{Fp} sind gestrichelt eingezeichnet.

Dazu werden verschiedene Geometrien für Quantenpunkte und ihre Auswirkungen auf die elektronische Struktur diskutiert. Dann ziehen wir aus der gemessenen Stark-Verschiebung Rückschlüsse auf die Geometrie und Indium-Konzentration der betrachteten Quantenpunkte. Zusätzlich sollen ihre optischen Eigenschaften untersucht werden. Als Nächstes untersuchen wir die diamagnetische Verschiebung aufgrund eines angelegten Magnetfeldes. Dazu legen wir experimentelle Resultate von I. E. Itskevich et al. zugrunde [77]. Dadurch, dass man das magnetische Feld in verschiedenen Richtungen anlegen kann, sind recht genaue Schlüsse über die Größe des Quantenpunktes möglich.

6.1 Experimenteller Aufbau

Die Zusammensetzung der untersuchten Proben aus [72] ist in Abb. 6.1 dargestellt. Die InAs-Quantenpunktschicht wurde mit Molekularstrahl-Epitaxie auf GaAs in (001)-Richtung gewachsen. Die Schicht befindet sich im intrinsischen Bereich einer pin-Diode. Durch Anlegen einer Spannung U kann das elektrische Feld am Quantenpunkt variiert werden. Dabei ist U_{bi} die eingebaute Spannung (engl. built-in potential, vgl. Abschnitt 4.1) und d die Weite der intrinsischen Region. Dann errechnet sich das elektrische Feld aus der Formel $E = \frac{U+U_{bi}}{d}$. Durch die angelegte Spannung U befindet sich die Probe im Nichtgleichgewicht, die La-

dungsträgerverteilung wird demzufolge durch das Quasi-Ferminiveau der Elektronen E_{Fn} und der Löcher E_{Fp} beschrieben. Die Diode ist in Sperr-Richtung gepolt, so dass sich der in Abb. 6.1 eingezeichnete Verlauf für die Quasi-Ferminiveaus ergibt. Die quantisierten Zustände im Quantenpunkt sind nicht besetzt. Will man das elektrische Feld in die andere Richtung messen, muss man eine eigene Probe wachsen, die in eine nip-Diode eingebettet ist.

Es wird nun monochromatisches Licht eingestrahlt. Entspricht die Wellenlänge dem energetischen Abstand zwischen dem quantisierten Elektron- und Lochzustand, so wird ein Exziton angeregt. Das Elektron und das Loch verlassen den Quantenpunkt, fließen über die Kontakte ab und bilden so den Fotostrom. Dieser wird experimentell bestimmt [72].

6.2 Grundgeometrie der Quantenpunkte

Die genaue Geometrie der Quantenpunkte ist in der Regel nicht bekannt, da beim Überwachsen mit GaAs Diffusionsprozesse stattfinden und dadurch die Quantenpunkte ihre Form verändern. Ziel dieses Kapitels ist es, aus experimentellen Daten wie der Stark-Verschiebung Rückschlüsse auf die genaue Gestalt der Quantenpunkte zu machen. In unseren Rechnungen legen wir die kp-Parameter, Deformationspotenziale, elastischen Konstanten und piezoelektrischen Konstanten aus [2] zugrunde.

Zunächst sollen aber grundsätzliche Überlegungen über mögliche Geometrien und die daraus resultierenden physikalischen Effekte angestellt werden. Wir beschränken uns dabei auf zwei Fälle: pyramidenförmige und linsenförmige Quantenpunkte. Wie wir später sehen werden, ist die Annahme einer Pyramidenform zwar wenig realistisch, in den meisten bisherigen theoretischen Untersuchungen wurde sie aber zugrunde gelegt. Die Verspannungen sind hier besonders groß und demzufolge können die damit verbundenen physikalischen Effekte gut erläutert werden.

Pyramidenförmige Quantenpunkte

Es soll der Quantenpunkt in Abb. 6.2 näher untersucht werden. Die Struktur wurde auf [001]-GaAs-Substrat gewachsen und besteht aus einer InAs-Schicht (wetting layer) und einem pyramidenförmigen Quantenpunkt mit einer Basislänge von 20 nm und einer Höhe von 10 nm. Die Wachstumsrichtung entspricht der z -Achse in Abb. 6.2. Aufgrund der unterschiedlichen Gitterkonstante von InAs und GaAs - die Gitterkonstante von InAs ist um ca. 8% größer - kommt es zu mechanischen Verspannungen in der Struktur. In derartigen Situationen ist die Annahme pseudomorphen Wachstums gerechtfertigt, d.h. an Materialgrenzflächen nehmen beide Materialien dieselbe Gitterkonstante an und die Verspannung ist rein elastisch ohne Versetzungen. Dann können die Verzerrungsfelder

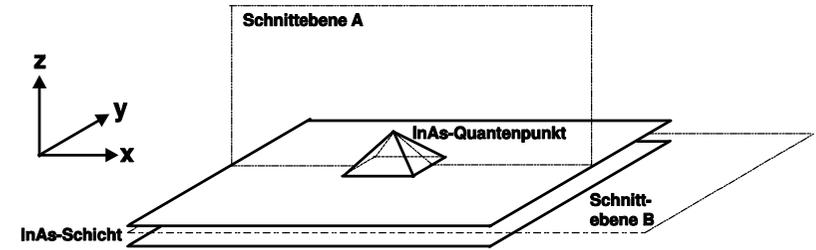


Abbildung 6.2: Pyramidenförmiger InAs-Quantenpunkt und InAs-Schicht (wetting layer) auf GaAs-Substrat. Der Quantenpunkt wurde mit GaAs überwachsen. Der Indium-Gehalt beträgt 100%. Die Wachstumsrichtung (z -Achse) entspricht der Kristallrichtung [001]. Eingezeichnet sind die Schnittebenen A und B, auf die in den anderen Abbildungen Bezug genommen wird.

im Rahmen der Elastizitätstheorie berechnet werden. Betrachten wir zunächst die Verhältnisse in einer InAs-Schicht, die pseudomorph auf ein GaAs-Substrat in [001]-Richtung (entspricht der z -Achse) aufgewachsen wurde (Abb. 6.3). Aufgrund der größeren Gitterkonstante von InAs werden die Gitterabstände in der InAs-Schicht in x - und y -Richtung zusammengestaucht und infolgedessen in z -Richtung gedehnt, d.h. $\epsilon_{xx} = \epsilon_{yy} < 0$ und $\epsilon_{zz} > 0$. Die Verspannungen in der InAs-Schicht sind homogen, es treten keine Scherspannungen auf. Das GaAs-Substrat ist unverspannt. In einem Quantenpunkt sind die Verspannungen dagegen stark inhomogen. Abb. 6.4 zeigt die Komponenten des Deformationstensors auf der Schnittebene A aus Abb. 6.2. In der InAs-Schicht (wetting layer) ergeben sich dieselben Verhältnisse wie sie oben diskutiert wurden, d.h. homogene Verspannung und keine Scherspannungen. Im Quantenpunkt selber sieht man deutlich, dass die Komponente ϵ_{zz} an der Basis positiv ist, zur Spitze hin jedoch stark abnimmt und negativ wird. Ursache dafür sind die Begrenzungsflächen des Quantenpunktes, die zur Spitze hin zusammenlaufen und aufgrund der kleineren Gitterkonstante im GaAs die Atome im Quantenpunkt in z -Richtung zusammenstauchen. Umgekehrt folgt daraus, dass die Komponenten ϵ_{xx} und ϵ_{yy} zur Spitze hin stark zunehmen und positiv werden. In Abb. 6.5 sind die Verhältnisse entlang einer Schnittlinie durch die Pyramidenspitze senkrecht zur Basisebene dargestellt. Man sieht deutlich, dass die Diagonalelemente des Deformationstensors an der Pyramidenspitze ihr Vorzeichen umkehren. Außerdem ist festzuhalten, dass anders als in einer InAs-Schichtstruktur hier auch das GaAs-Substrat verspannt ist. Diese Verspannungen der GaAs-Umgebung werden aber über eine Länge von 10-20 nm wieder abgebaut. Die Scherspannungen sind vor allem an den Kanten der Pyramidenoberfläche ausgeprägt, wie in Abb. 6.4 zu erkennen ist.

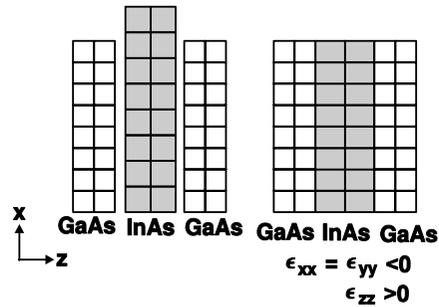


Abbildung 6.3: Verspannungen in einer InAs-Schicht, die pseudomorph auf GaAs-Substrat in $[001]$ -Richtung aufgewachsen wurde (rechts). Aufgrund der größeren Gitterkonstante im InAs (links) werden die Atome in x - und y -Richtung zusammengestaucht und infolgedessen in z -Richtung gedehnt. Dadurch erhält man $\epsilon_{xx} = \epsilon_{yy} < 0$ und $\epsilon_{zz} > 0$ in der InAs-Schicht.

Die Verspannungen führen über die Deformationspotenziale zu einer Verschiebung der Bandkanten. In Abb. 6.6 ist die Valenzbandkante entlang Schnittebene B aus Abb. 6.2 dargestellt. Das inhomogene Spannungsfeld führt dazu, dass die Valenzbandkante in der Basisebene von der Mitte zu den Rändern des Quantenpunktes um ca. 200 meV ansteigt. Dadurch wird die Grundzustandswellenfunktion des Loches nach außen gedrängt wird (vgl. Abb. 6.6). Das Elektron bleibt in der Mitte lokalisiert, da es aufgrund der deutlich kleineren effektiven Masse weniger stark auf Änderungen der Leitungsbandkante reagiert. Man beachte, dass bisher piezoelektrische Polarisationsladungen nicht berücksichtigt wurden. Diese werden durch die Scherspannungen verursacht (vgl. Abschnitt 4). Entsprechend unserer Spannungsberechnungen erhält man an den Kanten jeweils Dipole. Diese sind in Abb. 6.7 dargestellt. Insgesamt handelt es sich demnach um einen Oktopol, dessen Beitrag zum elektrischen Feld die Symmetrie des Problems weiter reduziert. Dies beeinflusst die Wellenfunktionen des Elektrons und des Loches, welche in Abb. 6.8 dargestellt sind. Der Grundzustand des Lochs ist nun nicht mehr symmetrisch, sondern in zwei diagonal gegenüberliegenden Ecken lokalisiert. Dasselbe gilt auch für den ersten angeregten Lochzustand. Der Grundzustand des Elektrons befindet sich weiterhin in der Mitte des Quantenpunktes. Der erste angeregte Elektronenzustand ist aufgrund der negativen Ladung in den beiden anderen Ecken lokalisiert als das Loch. Auf die Übergangsenergie von Elektron und Loch haben die piezoelektrischen Polarisationsladungen dagegen praktisch keinen Einfluss. Sie beträgt in unserem Fall 1.043 eV ohne bzw. 1.042 eV mit piezoelektrischem Feld, d.h. der Unterschied

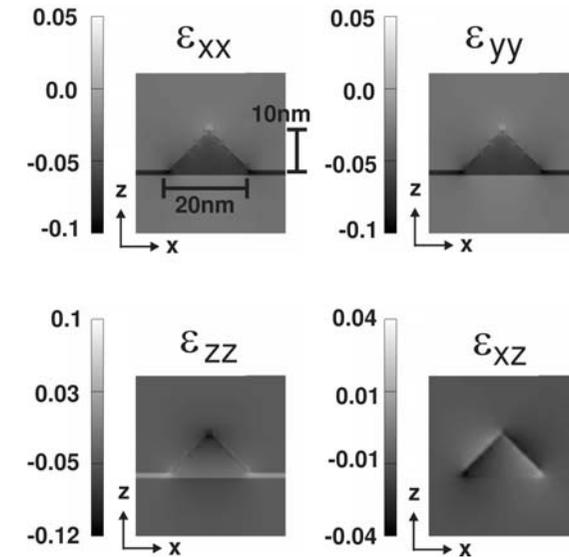


Abbildung 6.4: Komponenten des Spannungstensors in der Schnittebene A aus Abb. 6.2.

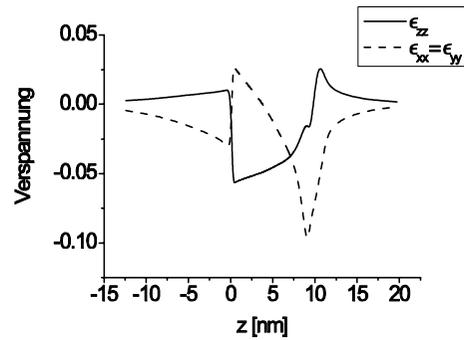


Abbildung 6.5: Diagonalelemente des Deformationstensors entlang einer Schnittlinie durch die Pyramidenspitze senkrecht zur Basisebene.

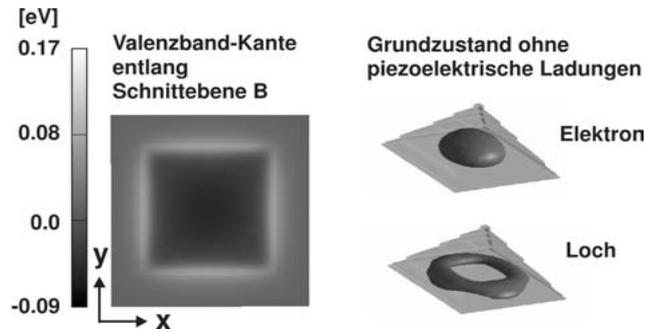


Abbildung 6.6: Valenzbandkante an der Basis des Quantenpunktes entlang der Schnitt Ebene B aus Abb. 6.2. Das inhomogene Verspannungsfeld führt zu einem Anstieg der Valenzbandkante zu den Rändern des Quantenpunktes hin. Dadurch wird die Grundzustandswellenfunktion des Loches nach außen gedrängt. Das Elektron mit der deutlich kleineren effektiven Masse bleibt in der Mitte lokalisiert. Dargestellt ist jeweils eine Isofläche für die Aufenthaltswahrscheinlichkeit von 60%.

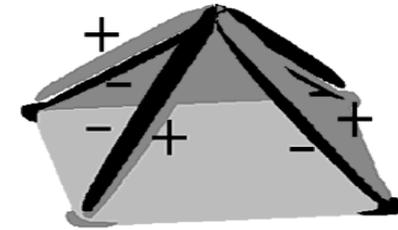


Abbildung 6.7: Piezoelektrische Polarisationsladungen für den pyramidenförmigen Quantenpunkt. Dargestellt sind die Isoflächen für eine Ladungsdichte von $3 \cdot 10^{19} \frac{e}{cm^3}$. Die Polarisationsladungen treten entlang der Kanten auf, da dort die Scherspannung groß ist. Die positiven Ladungen sind hell, die negativen dunkel gezeichnet.

ist im Bereich von 0.1%.

Linsenförmige Quantenpunkte

Wir wollen nun den pyramidenförmigen Quantenpunkt mit einem linsenförmigen vergleichen. Dieser hat eine Basisweite von 20 nm und eine Höhe von 5 nm und besteht wieder aus reinem InAs. Die Geometrie ist in Abb. 6.9 dargestellt. Für die Koordinatenachsen und die Schnittebenen gilt die gleiche Konvention wie in Abb. 6.2. Der wesentliche Unterschied zum pyramidenförmigen Quantenpunkt ist, dass die Begrenzungsflächen in wesentlich flacherem Winkel zur InAs-Schicht verlaufen. Dies führt zu einem homogeneren Verspannungsfeld im Quantenpunkt, wie in Abb. 6.10 zu erkennen ist. Die Scherspannungen verschwinden in der Mitte und sind maximal an den Begrenzungsflächen, insgesamt aber kleiner als bei der Pyramidenform. In Abb. 6.11 sind wieder die Diagonal-Komponenten des Verzerrungstensors entlang einer Schnittlinie durch die Spitze orthogonal zur InAs-Schicht dargestellt. Man sieht, dass die Verspannungen hier wesentlich homogener sind als in Abb. 6.5. Insgesamt kommen sie den Verhältnissen in einer InAs-Schicht deutlich näher. Aus der Scherspannung resultieren wieder piezoelektrische Polarisationsladungen, die nun aber kleiner und nicht so scharf lokalisiert sind, wie in Abb. 6.12 zu sehen ist. Der Hauptunterschied zum pyramidenförmigen Quantenpunkt besteht darin, dass durch die homogenere Verspannung das Valenzband innerhalb des Quantenpunktes nicht so stark variiert und infolgedessen der Grundzustand des Loches in der Mitte des Quantenpunktes lokalisiert ist (vgl. Abb. 6.13). Anders als bei der Pyramidenform hat

Wellenfunktionen unter Berücksichtigung der piezoelektrischen Polarisationsladungen

Elektronen



Löcher



Grundzustand

1. angeregter Zustand

2. angeregter Zustand

Abbildung 6.8: Wellenfunktionen der Elektronen und der Löcher für unseren pyramidenförmigen Quantenpunkt. Die piezoelektrischen Polarisationsladungen bilden ein Oktopol, welcher die Symmetrie reduziert. Der Grundzustand des Loches ist nunmehr in zwei diagonal gegenüber liegenden Ecken lokalisiert.

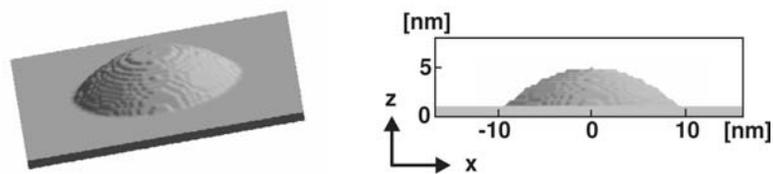


Abbildung 6.9: Geometrie des betrachteten linsenförmigen InAs-Quantenpunktes. Seine Höhe beträgt 5 nm, seine Basisweite 20 nm.

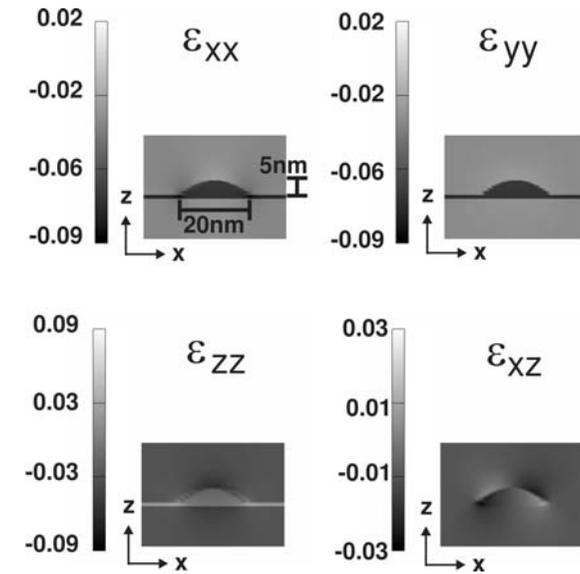


Abbildung 6.10: Komponenten des Verzerrungstensors in der Schnittebene A, wie sie in Abb. 6.2 dargestellt ist, jedoch für den linsenförmigen Quantenpunkt.

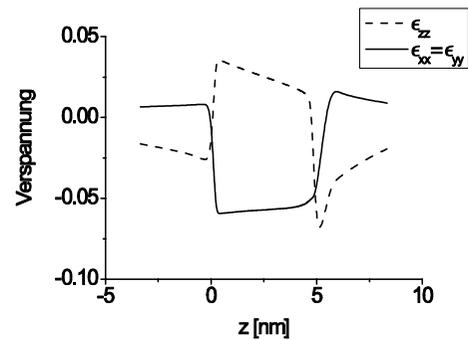


Abbildung 6.11: Diagonalelemente des Deformationstensors entlang der Schnittlinie durch die Spitze senkrecht zur Basisebene für den linsenförmigen Quantenpunkt.

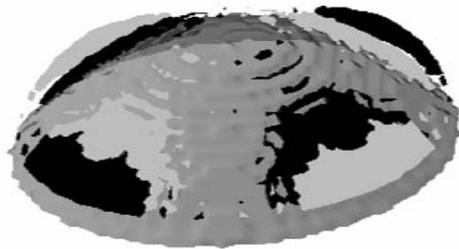


Abbildung 6.12: Piezoelektrische Polarisationsladungen für den linsenförmigen Quantenpunkt. Dargestellt sind die Isoflächen für eine Ladungsdichte von $10^{19} \frac{e}{cm^3}$. Die positiven Ladungen sind hell, die negativen dunkel gezeichnet. Im Vergleich zur Pyramide sind sie kleiner und nicht so scharf lokalisiert.

Wellenfunktionen für linsenförmigen Quantenpunkt

Elektronen



Löcher



Grundzustand

1. angeregter Zustand

2. angeregter Zustand

Abbildung 6.13: Wellenfunktionen der Elektronen und der Löcher für den linsenförmigen Quantenpunkt. Hauptunterschied zum pyramidenförmigen Quantenpunkt ist, dass der Loch-Grundzustand in der Mitte lokalisiert ist. Der Überlapp zwischen Elektron- und Lochwellenfunktion ist aber Voraussetzung für die optische Erzeugung von Exzitonen.

man hier also einen deutlichen Überlapp zwischen der Grundzustanswellenfunktion des Elektrons und des Loches. Dies ist aber Voraussetzung für die optische Generation eines Elektron-Loch-Paares, wie sie im Experiment vorkommt. Allein diese Tatsache legt den Schluss nahe, dass die gemessenen Quantenpunkte Linsenform haben müssen. Diese Vermutung wird gestützt durch AFM-Messungen (atomic force microscope) [72], die jedoch nur an unbedeckten Quantenpunkten durchgeführt werden können. Auch hier weisen die Quantenpunkte Linsenform auf. Ein weiterer interessanter Aspekt für Anwendungen ist, inwiefern ein Dipolmoment zwischen der positiven und negativen Ladung vorliegt. In den bisher betrachteten Fällen, d.h. mit konstanter Indium-Konzentration innerhalb des Quantenpunktes, ist die Lochwellenfunktion näher an der Basis und die Elektronwellenfunktion näher an der Spitze lokalisiert. Dies führt zu einem Widerspruch zu experimentellen Messungen, auf die im nächsten Abschnitt eingegangen wird.

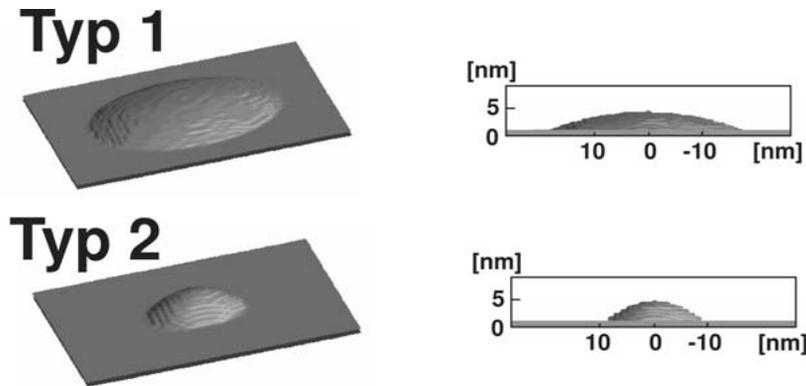


Abbildung 6.14: Die beiden betrachteten Typen von linsenförmigen Quantenpunkten: Typ 1 ist sehr flach mit einem Basisdurchmesser von 30 nm, Typ 2 ist eher schmal mit einem Basisdurchmesser von 15 nm.

Dimensionierung des gemessenen Quantenpunktes

Nachdem wir uns auf linsenförmige Quantenpunkte festgelegt haben, sollen im Weiteren die konkreten geometrischen Abmessungen bestimmt werden. Wir beschränken uns dabei auf die zwei Grenzfälle, die in Abb. 6.14 dargestellt sind: einen sehr flachen Quantenpunkt mit einem Basisdurchmesser von über 30 nm und einen eher schmalen mit einem Basisdurchmesser von gut 15 nm. Diese werden im Folgenden mit Typ 1 bzw. Typ 2 bezeichnet.

Konstante Indium-Konzentration

Die Quantenpunkte bestehen nicht aus reinem InAs, sondern aus einer InAs-GaAs-Legierung, $\text{In}_x\text{Ga}_{1-x}\text{As}$. Zunächst wird konstante Indium-Konzentration x innerhalb des Quantenpunktes vorausgesetzt. Die Indium-Konzentration bestimmt die Topftiefe und hat damit wesentlichen Einfluss auf die Übergangsenergie zwischen dem Grundzustand des Elektrons und des Loches. Da die Quantenpunkte recht flach sind, wird die Quantisierungsenergie des Grundzustandes überwiegend durch die Höhe des Quantenpunktes bestimmt. Dies gilt besonders für den Typ 1, während bei Typ 2 laterale Quantisierungseffekte eine merkliche Rolle spielen. Experimentell beträgt die Übergangsenergie 1.08 eV bei einer Temperatur von 200 K. Je größer die Höhe des Quantenpunktes, umso kleiner ist die Quantisierungsenergie und umso geringer die Indium-Konzentration, um den vorgegebenen energetischen Abstand zu erreichen. Es wurden eine Reihe von

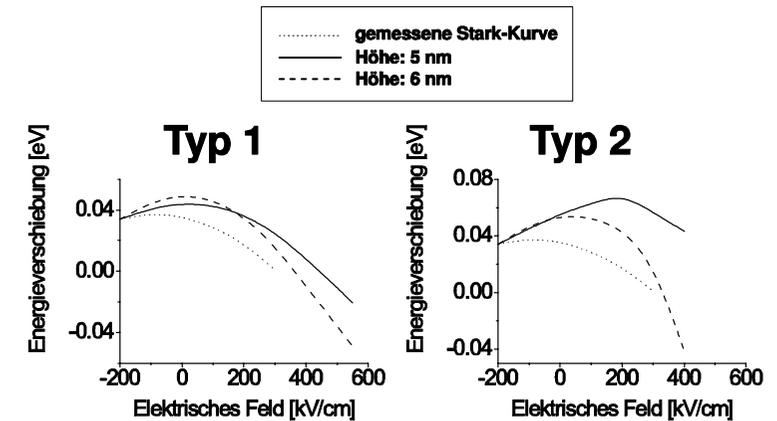


Abbildung 6.15: Berechnete Stark-Verschiebung der Übergangsenergie zwischen dem Grundzustand des Elektrons und des Loches in Abhängigkeit vom angelegten elektrischen Feld. Die experimentell bestimmten Werte aus [72] sind gepunktet eingezeichnet. Für den flachen linsenförmigen Quantenpunkt (Typ 1, links) wurde eine Höhe von 5 nm bzw. 6 nm bei einer Indium-Konzentration von 70% bzw. 60% zugrunde gelegt, für den schmaleren Quantenpunkt (Typ 2, rechts) eine Höhe von 5 nm bzw. 6 nm bei einer Indium-Konzentration von 80% bzw. 70%.

Rechnungen durchgeführt, die bei konstantem Legierungsprofil die Abhängigkeit der Übergangsenergie vom angelegten elektrischen Feld zeigen. Dabei wurde für den Quantenpunkt vom Typ 1 eine Höhe von 5 nm bzw. 6 nm mit einer Indium-Konzentration von 70% bzw. 60% zugrunde gelegt. Für den Quantenpunkt vom Typ 2 war entsprechend die Höhe 5 nm bzw. 6 nm bei einer Indium-Konzentration von 80% bzw. 70%. Die Ergebnisse sind in Abb. 6.15 dargestellt, die experimentell bestimmten Werte von Fry et al. [72] sind gepunktet eingezeichnet. Bei Anlegen eines elektrischen Feldes erhält man eine parabolische Abhängigkeit, die so genannte Stark-Parabel. Ihre Krümmung ist umso stärker, je höher der Quantenpunkt ist. Dann können nämlich die Schwerpunkte der Elektron- und Lochwellenfunktion durch ein elektrisches Feld stärker gegeneinander verschoben werden und dementsprechend größer ist die Verschiebung der energetischen Lage. Insgesamt kann aus der Krümmung der gemessenen Stark-Kurve und dem Absolutwert der Übergangsenergie auf die mittlere Indium-Konzentration und die ungefähre Höhe des Quantenpunktes geschlossen werden. In unserem Fall beträgt die Höhe zwischen 5 nm und 6 nm und die durchschnittliche Legierungskonzentration des Indiums liegt zwischen 60% und 70% für Typ 1 bzw. 70% und 80% für Typ 2.

Inneres Dipolmoment

Der bedeutendste Unterschied zwischen den berechneten Stark-Kurven bei konstanter Indium-Konzentration und den experimentellen Werten liegt darin, dass letztere ihren Scheitelpunkt bei einem negativen elektrischen Feld haben. Das positive elektrische Feld am Scheitelpunkt bei den gerechneten Stark-Parabeln erklärt sich daraus, dass die Elektron-Wellenfunktion in Richtung Spitze und die Lochwellenfunktion in Richtung Basis des Quantenpunktes verschoben ist. Daraus resultiert ein inneres Dipolmoment, das durch ein positives elektrisches Feld kompensiert werden muss, um den Scheitelpunkt der Stark-Parabel zu erreichen. Sämtliche Rechnungen mit pyramidenförmigen und linsenförmigen Quantenpunkten mit konstanter Indium-Konzentration führen zu einem inneren Dipolmoment, das im Vergleich zu den experimentellen Werten das falsche Vorzeichen hat. Die einzige mögliche Erklärung ist, dass im Experiment das Loch näher an der Spitze und das Elektron näher an der Basis lokalisiert ist [72], im Unterschied zu allen Rechnungen, die wir bisher vorgestellt haben. Man kann dies erreichen, wenn man ein inhomogenes Legierungsprofil annimmt, bei dem die Indium-Konzentration zur Spitze hin zunimmt. Dies wurde bereits von Fry et al. [72] vorgeschlagen. Wir werden darauf im nächsten Abschnitt genauer eingehen.

Beim Vergleich der gerechneten Kurven für Typ 1 und Typ 2 in Abb. 6.15 fällt auf, dass das innere Dipolmoment größer ist für den schmalen Quantenpunkt (Typ 2) als für den sehr flachen (Typ 1), d.h. ohne äußeres Feld sind die Ladungsschwerpunkte der Elektron- und Lochwellenfunktion stärker separiert. Der Grund liegt darin, dass bei dem schmalen Quantenpunkt (Typ 2) das Verspan-

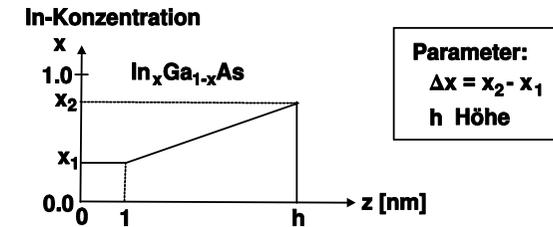


Abbildung 6.16: Betrachtetes Legierungsprofil für die Indium-Konzentration. Es wird bestimmt durch die beiden Grenzwerte für die Indium-Konzentration x_1 und x_2 und die Höhe h des Quantenpunktes. Da die durchschnittliche Indium-Konzentration durch die Übergangsenergie festgelegt ist, bleiben als Parameter noch $\Delta x = x_2 - x_1$ und die Höhe h .

nungsfeld inhomogener ist als bei dem sehr flachen Quantenpunkt (vergleiche die beiden extremen Fälle aus vorigem Abschnitt). Durch die Verspannungen wird aber die Lochwellenfunktion in Richtung Basis gedrängt, wie anhand der Pyramide deutlich gezeigt wurde.

Zusammenfassend bleibt festzuhalten, dass unser Quantenpunkt eine Höhe zwischen 5 nm und 6 nm und eine durchschnittliche Indium-Konzentration zwischen 60% und 70% für Typ 1 bzw. 70% und 80% für Typ 2 hat. Das Legierungsprofil muss inhomogen sein.

6.3 Inhomogenes Legierungsprofil

Einfaches lineares Legierungsprofil

Wir versuchen nun, den Scheitelpunkt der Stark-Parabel in Richtung negatives elektrisches Feld zu verschieben, indem wir ein inhomogenes Legierungsprofil annehmen, wobei die Indium-Konzentration zur Spitze hin zunehmen soll. Das Loch mit der höheren effektiven Masse reagiert sensibler auf diese Änderung als das Elektron. Damit kann man also erreichen, dass das Loch näher zur Spitze rückt und das Elektron seine Position fast nicht verändert. Zunächst nehmen wir an, dass die Indium-Konzentration im Bereich des Wetting-Layers, der mit 1 nm Dicke angesetzt wird, konstant ist und schließlich linear bis zur Spitze hin zunimmt. In Abb. 6.16 ist das entsprechende Legierungsprofil dargestellt. Als Parameter haben wir die beiden Grenzwerte für die Indium-Konzentration x_1 und x_2 und die Höhe h des Quantenpunktes. Da die durchschnittliche Indium-Konzentration durch die Übergangsenergie festgelegt ist, bleiben noch zwei Freiheitsgrade, die zu variieren sind, nämlich die Änderung der Legierungskonzentration $\Delta x = x_2 - x_1$

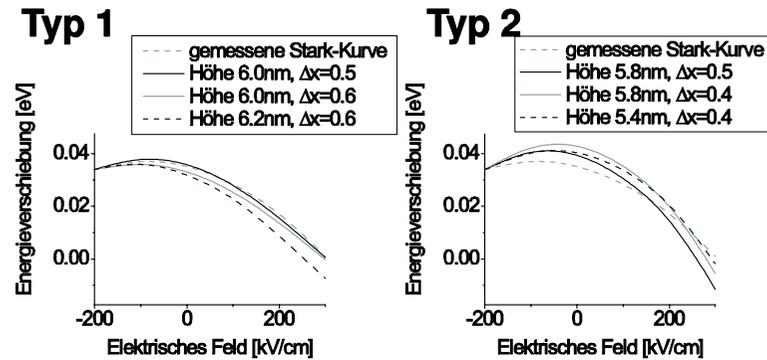


Abbildung 6.17: Stark-Kurven für beide Quantenpunkt-Typen (Typ 1 links, Typ 2 rechts). Es wurde einmal die Höhe h des Quantenpunktes und einmal der Abfall der Legierungskonzentration Δx variiert.

und die Höhe h . Wir wollen damit im Wesentlichen die Lage des Scheitelpunktes der Stark-Parabel und die Krümmung einstellen. Dazu wurden eine Reihe von Rechnungen mit unterschiedlichen Parameter für unsere beiden Quantenpunkte durchgeführt. In Abb. 6.17 sind einige charakteristische Beispiele dargestellt. Betrachten wir zunächst den Quantenpunkt vom Typ 1. Durch leichtes Vergrößern der Höhe wird die Krümmung der Parabel etwas größer. Ein steileres Legierungsprofil schiebt den Scheitelpunkt etwas nach rechts. Ähnliches Verhalten kann man beim schmalen Quantenpunkt (Typ 2) feststellen. Durch Verringern des Abfalls der Legierungskonzentration wird der Scheitel der Stark-Parabel in Richtung positiven elektrischen Feldes verschoben, die Krümmung der Parabel dagegen wenig beeinflusst. Für den Grenzfall einer konstanten Indium-Konzentration liegt der Scheitelpunkt im Bereich eines positiven elektrischen Feldes. Verringert man nun die Höhe des Quantenpunktes, so wird die Krümmung etwas kleiner, ein Verhalten, das wir schon vorher festgestellt haben. Insgesamt kann man feststellen, dass wir durch Annahme eines einfachen linearen Legierungsprofils das richtige Vorzeichen für das innere Dipolmoment erhalten. Insgesamt sind wir der experimentell gemessenen Kurve schon recht nahe gekommen, insbesondere für den Quantenpunkt vom Typ 1. Um eine noch bessere Übereinstimmung zu erhalten müssen jedoch noch weitere Modifikationen am Legierungsprofil vorgenommen werden.

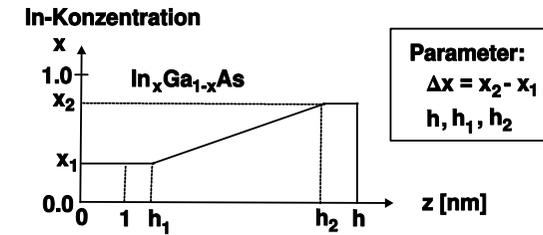


Abbildung 6.18: Abgeschnittenes lineares Legierungsprofil für die Indium-Konzentration. Es wird bestimmt durch die beiden Grenzwerte für die Indium-Konzentration x_1 bzw. x_2 beim Abstand h_1 bzw. h_2 von der Basis und durch die Höhe h des Quantenpunktes. Da die durchschnittliche Indium-Konzentration durch die Übergangsenergie festgelegt ist, bleiben als Parameter noch $\Delta x = x_2 - x_1$, h_1 , h_2 und die Höhe h .

Abgeschnittenes lineares Legierungsprofil

Wir betrachten nun ein abgeschnittenes lineares Legierungsprofil, das in Abb. 6.18 dargestellt ist. Man hat hier insgesamt fünf Freiheitsgrade, nämlich x_1 , x_2 , h_1 , h_2 und h . Nach umfangreichen Rechnungen, auf die wir hier nicht näher eingehen wollen, wurden folgende Werte für die zwei Quantenpunkte gefunden: Für den Quantenpunkt vom Typ 1 variiert die Indium-Konzentration im Bereich von $h_1 = 1$ nm bis $h_2 = 4.7$ nm linear zwischen $x_1 = 0.42$ und $x_2 = 0.84$, die Höhe beträgt $h = 6.1$ nm. Im Fall des Quantenpunktes vom Typ 2 liegt die Indium-Konzentration zwischen $x_1 = 0.6$ und $x_2 = 1.0$ im Bereich von $h_1 = 1$ nm bis $h_2 = 3.5$ nm, die Höhe beträgt $h = 5.1$ nm. Die berechneten Stark-Parabeln für den Exziton-Grundzustand sind in Abb. 6.19 dargestellt, sie stimmen mit den gemessenen Werten gut überein.

6.4 Optische Eigenschaften

Bisher konnten wir zeigen, dass sich die Stark-Verschiebung des Exziton-Grundzustands für beide Quantenpunkte durch entsprechende Wahl des Legierungsprofils theoretisch beschreiben lässt. Im Folgenden sollen nun die optischen Eigenschaften des Quantenpunktes genauer untersucht werden. Da der Dunkelstrom der Diode klein ist, wird der gemessene Strom fast vollständig durch die optisch erzeugten Exzitonen verursacht. Der Abfall des Fotostroms bei zunehmender Feldstärke kann mit der Verringerung der optischen Generationsrate erklärt werden, welche wiederum auf den abnehmenden Überlapp der Wellenfunktionen zurückzuführen ist. Daraus kann man experimentell auf den Verlauf der optischen

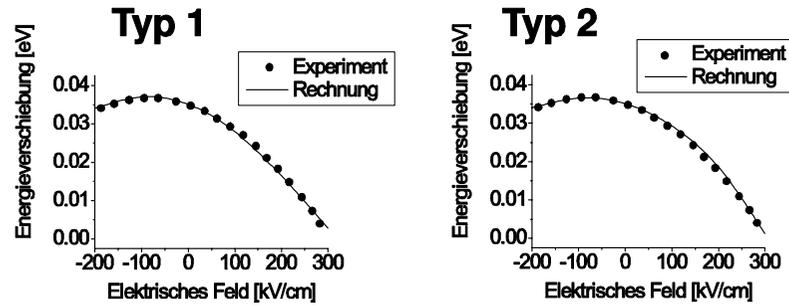


Abbildung 6.19: Vergleich zwischen berechneten und experimentellen Stark-Kurven. Es wurde ein Legierungsprofil gemäß Abb. 6.18 zugrunde gelegt. Für den Quantenpunkt vom Typ 1 (links) gilt: $h_1 = 1 \text{ nm}$, $h_2 = 4.7 \text{ nm}$, $x_1 = 0.42$, $x_2 = 0.84$, $h = 6.1 \text{ nm}$, für den Quantenpunkt vom Typ 2 ist: $x_1 = 0.6$, $x_2 = 1.0$, $h_1 = 1 \text{ nm}$, $h_2 = 3.5 \text{ nm}$, $h = 5.1 \text{ nm}$.

Matrixelemente schließen.

Berechnung der Interband-Matrixelemente

Die Übergangswahrscheinlichkeit berechnen wir mit Fermis Goldener Regel. Hier geht im Wesentlichen das Dipolmatrixelement $|\langle f | \vec{\epsilon} \cdot \vec{p} | i \rangle|^2$ ein mit dem Anfangszustand $|i\rangle$ und dem Endzustand $|f\rangle$. $\vec{\epsilon}$ ist der Polarisationsvektor des einfallenden Lichts, \vec{p} der Dipoloperator. In Achtband-kp-Theorie sind die Spinoren gegeben durch:

$$\Psi(\vec{r}) = \sum_{l=1}^8 F_l(\vec{r}) u_l(\vec{r}).$$

Dabei sind die $F_l(\vec{r})$ die Einhüllenden-Funktionen und $u_l(\vec{r})$ die Blochfunktionen. Unter der Annahme, dass die gitterperiodischen Blochfunktionen schneller variieren als die Einhüllenden-Funktionen, ergibt sich für das Matrixelement [22]:

$$\langle f | \vec{\epsilon} \cdot \vec{p} | i \rangle = \sum_{l,m=1}^8 \vec{\epsilon} \cdot \langle u_l | \vec{p} | u_m \rangle \int F_{f_l}^*(\vec{r}) F_{i_m}(\vec{r}) d^3r + \sum_{l,m=1}^8 \delta_{lm} \vec{\epsilon} \cdot \int F_{f_l}^*(\vec{r}) \vec{p} F_{i_m}(\vec{r}) d^3r \quad (6.1)$$

In unserem Experiment finden Übergänge vom Leitungs- ins Valenzband, also Interband-Übergänge statt. Für diesen Fall verschwindet der zweite Term in

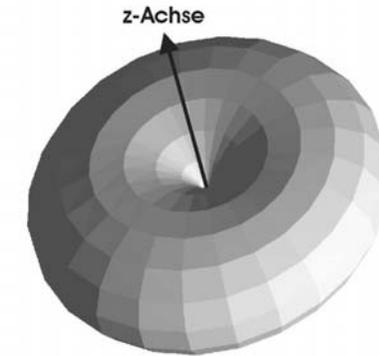


Abbildung 6.20: Optisches Matrixelement für den Übergang vom schweren Loch zum Elektron als Funktion der Polarisation des eingestrahelten Lichts. Die z-Achse entspricht der Wachstumsrichtung.

Gl. (6.1). Das Matrixelement hängt nun direkt vom Überlapp der Einhüllenden-Funktionen und den Impulsmatrixelementen der Basiszustände ab. Von den Impulsmatrixelementen sind nur folgende sechs ungleich Null:

$$\begin{aligned} \langle s \uparrow | p_x | x \uparrow \rangle &= \langle s \uparrow | p_y | y \uparrow \rangle = \langle s \uparrow | p_z | z \uparrow \rangle = \\ \langle s \downarrow | p_x | x \downarrow \rangle &= \langle s \downarrow | p_y | y \downarrow \rangle = \langle s \downarrow | p_z | z \downarrow \rangle. \end{aligned} \quad (6.2)$$

Experimentell ergibt sich, dass optische Anregung nur für Licht mit einem Polarisationsvektor senkrecht zur Wachstumsrichtung (z-Achse) stattfindet. Dies kann rechnerisch exakt reproduziert werden. In Abb. 6.20 ist das optische Matrixelement $|\langle f | \vec{\epsilon} \cdot \vec{p} | i \rangle|^2$ als Funktion der Polarisationsrichtung dargestellt. Es verschwindet für einen Polarisationsvektor in z-Richtung und ist isotrop für einen Polarisationsvektor in der (x,y)-Ebene. Dieses Resultat kann man damit erklären, dass der Grundzustand den Charakter eines schweren Loches hat. Die Eigenzustände des schweren Loches lauten (vgl. Kapitel 2):

$$|HH \uparrow \rangle = \frac{1}{\sqrt{2}} |(x + iy) \uparrow \rangle \quad (6.3)$$

$$|HH \downarrow \rangle = \frac{i}{\sqrt{2}} |(x - iy) \downarrow \rangle \quad (6.4)$$

Zusammen mit den Impulsmatrixelementen aus Gl. (6.2) folgt, dass das optische Matrixelement für z-polarisiertes Licht verschwinden muss.

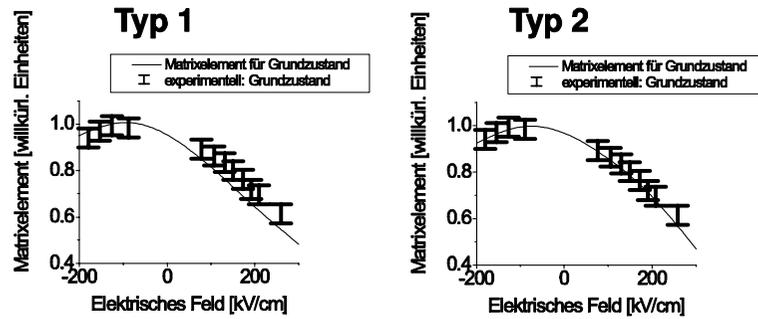


Abbildung 6.21: Optische Matrixelemente $|\langle f | \vec{\epsilon} \cdot \vec{p} | i \rangle|^2$ für den flachen Quantenpunkt (Typ 1, links) und den schmalen (Typ 2, rechts). Der berechnete Verlauf stimmt gut mit dem Verlauf der experimentellen Übergangswahrscheinlichkeit überein.

In Abb. 6.21 sind die berechneten Matrixelemente für beide Quantenpunkte als Funktion des angelegten elektrischen Feldes dargestellt. Sie stimmen gut mit den experimentellen Werten überein. Man beachte, dass hier nur die relative Änderung, nicht der Absolutwert betrachtet wird. Der geometrische Überlapp der Loch- und Elektronwellenfunktion ist recht groß, wie in Abb. 6.22 zu sehen ist. Dort sind die Grundzustandswellenfunktionen ohne angelegtes Feld dargestellt. Das Loch ist gegenüber dem Elektron zur Spitze hin verschoben. Dies erklärt das innere Dipolmoment, welches weiter oben diskutiert wurde.

Angeregter Exzitonzustand

Nun betrachten wir den nächst höheren energetischen Übergang, der experimentell gemessen wurde und circa 70 meV über der Energie des Grundzustandsexzitons liegt. Zunächst kommen dafür der Übergang vom Grundzustand des Loches zum ersten angeregten Elektronenzustand oder der Übergang vom ersten angeregten Lochzustand zum ersten angeregten Elektronenzustand infrage. Um dies zu entscheiden, betrachten wir wieder die optischen Matrixelemente. Es zeigt sich, dass die Übergangswahrscheinlichkeit zwischen dem Grundzustand des Loches und dem ersten angeregten Elektronenzustand praktisch verschwindet, dieser Übergang ist also verboten. Man kann dies einsehen, wenn man die zugehörigen Wellenfunktion betrachtet (Abb. 6.23, links). Der Elektronenzustand hat negative Parität, der Lochzustand positive. Dadurch ist ein optischer Übergang zwischen beiden nicht möglich. Abgesehen davon ist der geometrische Überlapp

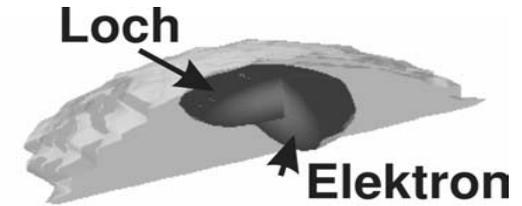


Abbildung 6.22: Elektron- und Lochzustand für verschwindendes angelegtes Feld beim schmalen Quantenpunkt (Typ 2). Das Loch ist näher an der Spitze des Quantenpunktes, das Elektron in der Mitte lokalisiert. Dies erklärt das innere Dipolmoment. Es sind die Isoflächen für eine Aufenthaltswahrscheinlichkeit von 50% der maximalen Ladungsdichte gezeichnet.

recht klein. Wir können also folgern, dass der nächste, energetisch höhere Übergang zwischen dem angeregten Loch und dem angeregten Elektron stattfindet. Die entsprechenden Wellenfunktionen sind in Abb. 6.23 rechts dargestellt. Die energetische Lage der angeregten Zustände wird überwiegend durch die laterale Quantisierung bestimmt, also die Breite des Quantenpunktes. In Abb. 6.24 ist die Stark-Verschiebung der Übergangenergie des angeregten Exzitonenzustands dargestellt. Für den sehr flachen Quantenpunkt (Typ 1) liegt die Kurve energetisch unter der gemessenen, für den schmalen Quantenpunkt liegt sie etwas zu hoch. Daraus folgt, dass der gemessene Quantenpunkt etwas breiter sein muss als unser Quantenpunkt vom Typ 2. Aus der energetischen Lage des ersten angeregten Exzitonenzustands kann man daher auf die Breite eines Quantenpunktes schließen.

6.5 Exzitonische Korrektur

Bisher haben wir für die Übergangenergien die Einteilchen-Zustände des Quantenpunktes betrachtet. Die entsprechenden Operatoren lauten:

$$\begin{aligned} \mathbf{H}_{el}^0 &= \frac{\vec{p}^2}{2m_{el}} + V_{el}^{ext} \\ \mathbf{H}_{hl}^0 &= -\frac{\vec{p}^2}{2m_{hl}} + V_{hl}^{ext} \end{aligned} \quad (6.5)$$

Für die externen Potenziale gilt:

$$\begin{aligned} V_{el}^{ext}(\vec{r}_{el}) &= E_c(\vec{r}_{el}) - e\Phi(\vec{r}_{el}) \\ V_{hl}^{ext}(\vec{r}_{hl}) &= E_v(\vec{r}_{hl}) - e\Phi(\vec{r}_{hl}), \end{aligned}$$

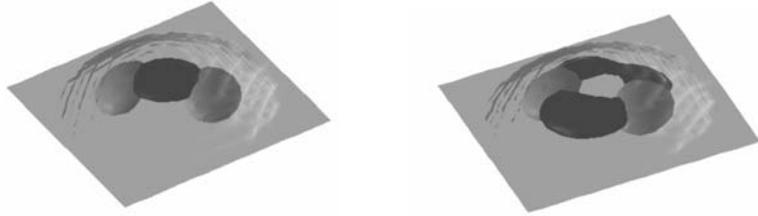


Abbildung 6.23: Die Darstellung zeigt den ersten angeregten Zustand des Elektrons (hell, in beiden Bildern gleich) und den Grundzustand des Loches (dunkel, links) bzw. den ersten angeregten Zustand des Loches (dunkel, rechts) für den Quantenpunkt vom Typ 2. Es sind jeweils die Isoflächen für eine Aufenthaltswahrscheinlichkeit von 50% der maximalen Ladungsdichte gezeichnet.

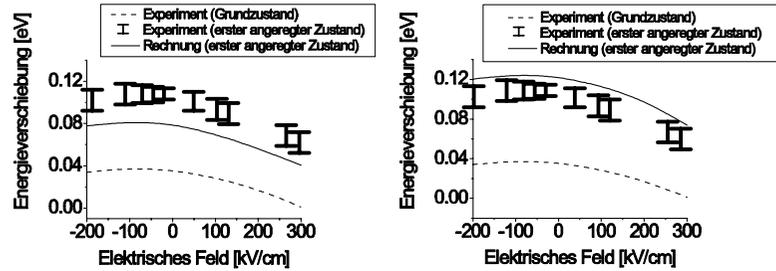


Abbildung 6.24: Stark-Verschiebung des ersten angeregten Exzitonenzustandes f für den flachen Quantenpunkt (Typ 1, links) und den schmalen (Typ 2, rechts). Je breiter der Quantenpunkt ist, desto größer ist die laterale Quantisierung und damit die Übergangsenergie für den ersten angeregten Exzitonenzustand.

mit dem elektrischen Potenzial Φ und den Bandkanten E_c bzw. E_v ($e = |e|$ Elementarladung). Die entsprechenden Einteilchen-Wellenfunktionen sind $\Psi_{el}(\vec{r}_{el})$ und $\Psi_{hl}(\vec{r}_{hl})$:

$$\begin{aligned} \mathbf{H}_{el}^0 \Psi_{el}(\vec{r}_{el}) &= E_{el}^0 \Psi_{el}(\vec{r}_{el}) \\ \mathbf{H}_{hl}^0 \Psi_{hl}(\vec{r}_{hl}) &= E_{hl}^0 \Psi_{hl}(\vec{r}_{hl}) \end{aligned} \quad (6.6)$$

Für das Exziton wollen wir nun im Rahmen der Hartree-Näherung zumindest die Coulomb-Wechselwirkung mitberücksichtigen. Die Wellenfunktion wird als Produkt von Elektron- und Lochwellenfunktion angesetzt:

$$\Omega(\vec{r}_{el}, \vec{r}_{hl}) = \Psi_{el}(\vec{r}_{el}) \cdot \Psi_{hl}(\vec{r}_{hl}) \quad (6.7)$$

Der Hamilton-Operator lautet nun:

$$\begin{aligned} \mathbf{H}_{EX} &= \mathbf{H}_{el}^0 + \mathbf{H}_{hl}^0 - V_{Coulomb} \\ V_{Coulomb} &= \frac{e^2}{4\pi\epsilon_0\epsilon} \int d\vec{r}_{el} d\vec{r}_{hl} \frac{|\Psi_{el}(\vec{r}_{el})|^2 |\Psi_{hl}(\vec{r}_{hl})|^2}{|\vec{r}_{el} - \vec{r}_{hl}|} \end{aligned} \quad (6.8)$$

Die Minimierung des Energiefunktional $(\mathbf{H}_{EX})_{\Omega}$ mit der Normierung der Einteilchen-Wellenfunktionen als Randbedingung liefert das Gleichungssystem [51]:

$$\begin{aligned} [\mathbf{H}_{el}^0 - V^{hl}(\vec{r}_{el})] \Psi_{el}(\vec{r}_{el}) &= \lambda_{el} \Psi_{el}(\vec{r}_{el}) \\ [\mathbf{H}_{hl}^0 - V^{el}(\vec{r}_{hl})] \Psi_{hl}(\vec{r}_{hl}) &= \lambda_{hl} \Psi_{hl}(\vec{r}_{hl}) \end{aligned} \quad (6.10)$$

mit $V^{hl}(\vec{r}_{el}) = \frac{e}{4\pi\epsilon_0\epsilon} \int d\vec{r}_{hl} \frac{|\Psi_{hl}(\vec{r}_{hl})|^2}{|\vec{r}_{el} - \vec{r}_{hl}|}$ und $V^{el}(\vec{r}_{hl}) = \frac{-e}{4\pi\epsilon_0\epsilon} \int d\vec{r}_{el} \frac{|\Psi_{el}(\vec{r}_{el})|^2}{|\vec{r}_{el} - \vec{r}_{hl}|}$. Diese Wechselwirkungspotenziale ergeben sich aus Lösung der Poissongleichung:

$$\begin{aligned} \vec{\nabla} \cdot (\epsilon \vec{\nabla} V^{el}) &= -e |\Psi_{el}|^2 \\ \vec{\nabla} \cdot (\epsilon \vec{\nabla} V^{hl}) &= e |\Psi_{hl}|^2 \end{aligned} \quad (6.11)$$

Die Exzitonenergie erhält man schließlich aus der selbstkonsistenten Lösung der Gleichungen (6.10) und (6.11):

$$E_{EX} = \lambda_{el} - \lambda_{hl} + \frac{1}{2} (\langle \Psi_{hl} | V^{el} | \Psi_{hl} \rangle - \langle \Psi_{el} | V^{hl} | \Psi_{el} \rangle) \quad (6.12)$$

Dabei ist noch ein Korrekturterm nötig, da in den beiden Eigenwerten λ_{el} und λ_{hl} die Coulombwechselwirkung doppelt gezählt wird und nun durch die beiden Matrixelemente wieder abgezogen werden muss [51]. Die exzitonische Korrektur ist also:

$$\Delta E_{EX} = E_{el}^0 - E_{hl}^0 - E_{EX}.$$

Die Bindung von Elektron und Loch wird durch das äußere Potenzial bestimmt, die Coulomb-Wechselwirkung stellt nur eine Korrektur zu den Einteilchen-Energien dar. Die exzitonische Korrektur hängt stark von der Dimension des

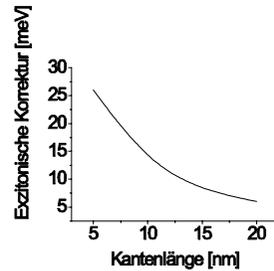


Abbildung 6.25: Exzitonische Korrektur für einen idealisierten würfelförmigen Quantenpunkt mit unendlich hohen Potenzialwänden. Es ergibt sich eine $\frac{1}{r}$ -Abhängigkeit, wie sie aufgrund der Coulomb-Weschselwirkung zu erwarten ist.

Quantenpunktes und der Tiefe des Potenzialtopfes ab. In Abb. 6.25 ist die exzitonische Korrektur für einen idealisierten Würfel mit mit unendlich hohen Potenzialwänden als Funktion der Kantenlänge dargestellt. Es ergibt sich ein $\frac{1}{r}$ -Abfall, der von der Coulomb-Wechselwirkung herrührt. Für unsere beiden Quantenpunkte erhält man einen Wert von 12 meV. Sie liegt demzufolge bei einer Größenordnung von circa einem Prozent der Grundzustandsenergie des Exzitons.

Zusammenfassung

In diesem Abschnitt ging es darum, aus experimentellen Messungen an überwachsenden InGaAs-Quantenpunkten auf die Größe und Form zu schließen. Als Grundgeometrie legen wir dabei linsenförmige Quantenpunkte zugrunde. Diese Annahme wird durch experimentelle Befunde an freien Quantenpunkten gestützt. Die gemessene Stark-Verschiebung ergibt, dass ohne äußeres Feld der Ladungsschwerpunkt des Loches näher an der Spitze als der des Elektrons lokalisiert ist. Dies lässt sich theoretisch nur erklären, wenn man ein inhomogenes Legierungsprofil für die Indium-Konzentration annimmt, wobei der Indium-Gehalt zur Spitze hin auf 100% zunimmt. Zur weiteren Eingrenzung der Geometrie muss man die Stark-Verschiebung des ersten angeregten Exzitonzustands heranziehen. Der energetische Abstand zur Grundzustandsenergie wird im Wesentlichen durch die laterale Quantisierung, also die Breite des Quantenpunktes, bestimmt. In unseren Rechnungen haben wir Einteilchen-Zustände betrachtet. Die exzitonische Korrektur liegt im Bereich von einem Prozent der Übergangsenergie.

6.6 Simulation von Quantenpunkten mit Magnetfeld

Bisher wurden die optischen Eigenschaften von Quantenpunkten in Abhängigkeit von äußeren elektrischen Feldern ausführlich untersucht. Nun wollen wir das Verhalten in Magnetfeldern näher betrachten. Dabei legen wir experimentelle Ergebnisse von I. E. Itskevich et al. [77] zugrunde. In unserem Modell haben wir bisher Magnetfelder in Einband-Näherung implementiert, der Hamilton-Operator lautet (vgl. 2.75) mit dem Vektorpotenzial \vec{A} , dem Magnetfeld \vec{B} , der effektiven Masse m_{eff} , dem Bohr-Magneton μ_B und dem effektiven g-Faktor g_{eff} :

$$\mathbf{H} = -\frac{\hbar^2}{2} \left(\vec{\nabla} + i\frac{q}{\hbar c} \vec{A}(\vec{r}) \right) \frac{1}{m_{eff}(\vec{r})} \left(\vec{\nabla} + i\frac{q}{\hbar c} \vec{A}(\vec{r}) \right) + \frac{1}{2} g_{eff} \mu_B \vec{B}$$

Während sich die diamagnetische Verschiebung direkt aus der Simulation ergibt, muss der g-Faktor g_{eff} als Parameter vorgegeben werden.

Experiment

Die Messungen in Ref. [77] wurden an InAs-Quantenpunkten durchgeführt, die im Stranski-Krastanov-Modus auf GaAs-Substrat in [001]-Richtung gewachsen wurden. Da über ihre Größe und Form nichts bekannt ist, nehmen wir in den folgenden Rechnungen die Form eines Semi-Ellipsoids mit einem Basisdurchmesser L und einer Höhe H an. Experimentell wurde die optische Übergangsenergie zwischen dem Loch- und Elektrongrundzustand in Abhängigkeit vom äußeren Magnetfeld im Bereich von 0 T bis 23 T bestimmt. Das Feld war dabei einmal in [001]- und einmal in [100]-Richtung orientiert.

Die experimentellen Messdaten zeigen, dass die Änderung der Exzitonenergie proportional zum Quadrat der Magnetfeldstärke ist. Diese Abhängigkeit ist typisch für die diamagnetische Verschiebung [32]. Man kann daher annehmen, dass in unserem Fall die Zeemann-Aufspaltung gegenüber der diamagnetischen Verschiebung vernachlässigbar ist. In der Tat zeigen andere Messungen, dass in InAs-Quantenpunkten der g-Faktor deutlich kleiner ist als im Volumenhalbleiter. Bei den in Ref. [78] untersuchten Quantenpunkten beträgt er z.B. -0.8 für Elektron und Loch zusammen im Vergleich zu -15.6 im InAs-Volumenhalbleiter.

Experimentell ist nur die gesamte Aufspaltung messbar. Gerade deshalb ist es wichtig, Aussagen über die diamagnetische Verschiebung in Abhängigkeit von der Größe, Form und Indium-Konzentration zu treffen. Dann können aus den Rechnungen und den experimentellen Ergebnissen Rückschlüsse auf den betreffenden g-Faktor gezogen werden.

Diamagnetische Verschiebung in Quantenpunkten

Wir wollen im Folgenden Simulationsergebnisse für die diamagnetische Verschiebung an InAs-Quantenpunkten mit Wachstumsrichtung [001] vorstellen, wobei das Magnetfeld in [001]- und in [100]-Richtung angelegt wurde. Als Grundgeometrie legen wir ein Semi-Ellipsoid mit einem Basisdurchmesser L und einer Höhe H zugrunde. Zunächst nehmen wir außerdem eine konstante Indium-Konzentration an, welche sich bei gegebener Höhe H und Länge L aus der Übergangsenergie ergibt.

Variation der Höhe H

Die diamagnetische Verschiebung des Elektron- bzw. Lochgrundzustands kann man schreiben als [77]:

$$\Delta E = \frac{\gamma e^2 B^2 \langle x^2 \rangle}{2m^*}.$$

Hier ist m^* die effektive Masse, B das angelegte Magnetfeld, γ ein geometrischer Formfaktor und $\sqrt{\langle x^2 \rangle}$ ist ein Maß für die Ausdehnung der Wellenfunktion in der Ebene senkrecht zum magnetischen Feld. Deshalb erwartet man folgendes Verhalten: Die Energieverschiebung wird umso größer sein, je größer der Quantenpunkt in der Ebene senkrecht zum magnetischen Feld ist. Außerdem ist die Energieverschiebung proportional zum Quadrat des magnetischen Feldes. In Abb. 6.26 sind die Simulationsergebnisse für einen Quantenpunkt mit einem Basisdurchmesser $L = 20$ nm für verschiedene Höhen dargestellt, wobei das Magnetfeld in [100]-Richtung angelegt wurde. Da die Exzitonenergie für verschwindendes Magnetfeld 1.255 eV beträgt, erhält man für die Höhe $H = 4$ nm eine Indium-Konzentration von $x_{In} = 80\%$, für $H = 5$ nm $x_{In} = 70\%$ und für $H = 6$ nm $x_{In} = 65\%$. Wie erwartet nimmt die Energieverschiebung mit der Höhe zu und hat eine quadratische Abhängigkeit von der Magnetfeldstärke. Hat das Magnetfeld dagegen [001]-Richtung, so ist die Energieverschiebung praktisch unabhängig von der Höhe des Quantenpunktes, sie variiert für Höhen von 4 nm bis 6 nm nur um circa 5%. Dies liegt daran, dass die Ausdehnung der Wellenfunktion in lateraler Richtung nur wenig von der Höhe des Quantenpunktes beeinflusst wird.

Variation des Basisdurchmessers L

Variiert man nun die Basislänge L , so erhält man genau umgekehrte Verhältnisse. Die Energieverschiebung hängt für ein Magnetfeld in [100]-Richtung praktisch nicht von der Basislänge ab. Die Resultate mit Magnetfeld in [001]-Richtung sind in Abb. 6.27 für verschiedene Basisdurchmesser L und eine Höhe von 5 nm dargestellt. Wie erwartet nimmt die Energieverschiebung mit zunehmender Basislänge L zu. Die Indium-Konzentration x_{In} beträgt jeweils 55%, 70% und 85% für eine Basislänge von 30 nm, 20 nm bzw. 10 nm.

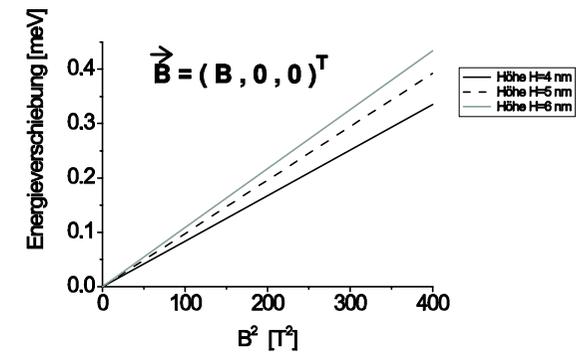


Abbildung 6.26: Diamagnetische Energieverschiebung für einen Quantenpunkt mit Basisdurchmesser 20 nm und verschiedenen Höhen H . Die Indium-Konzentration x_{In} beträgt 80% für $H = 4$ nm, 70% für $H = 5$ nm und 65% für $H = 6$ nm.

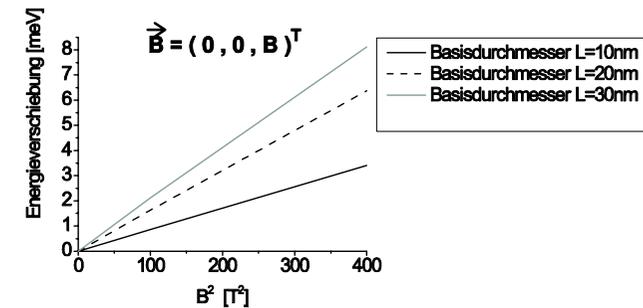


Abbildung 6.27: Diamagnetische Energieverschiebung für einen Quantenpunkt mit Höhe 5 nm und verschiedenen Basislängen L . Die Indium-Konzentration x_{In} beträgt jeweils 55%, 70% und 85% für eine Basislänge von 30 nm, 20 nm bzw. 10 nm.

Abhängigkeit vom Legierungsprofil

Schließlich kann man noch die Abhängigkeit der diamagnetischen Verschiebung von der Form des Legierungsprofils untersuchen. Dazu wählen wir eine Quantenpunkt mit Höhe $H = 6$ nm und Basisdurchmesser $L = 10$ nm. Dabei war die Indium-Konzentration einmal konstant ($x_{In} = 70\%$, Fall A), einmal variierte es linear von der Basis zur Spitze von 70% bis 90% (Fall B) und schließlich genau umgekehrt von 90% bis 70% (Fall C). Für ein Magnetfeld in [001]-Richtung ergibt sich bei einer Feldstärke von $B = 20$ T eine Energieverschiebung von 3.61 meV (Fall A), 3.16 meV (Fall B) und 4.18 meV (Fall C). Man kann dies folgendermaßen plausibel machen. Für zur Spitze hin zunehmende Indium-Konzentration sind Elektron und Loch näher an der Spitze lokalisiert. Da die laterale Breite des Quantenpunktes zur Spitze hin abnimmt, 'sehen' sie daher eine geringere laterale Breite. Infolgedessen nimmt die Energieverschiebung durch das Magnetfeld ab.

Zusammenfassung

In Abb. 6.28 wird die simulierte diamagnetische Verschiebung für ein Magnetfeld in [001]- bzw. [100]-Richtung mit den experimentellen Daten verglichen. Dabei wurde die Höhe als $H=6$ nm und der Basisdurchmesser als $L=11$ nm angenommen. Unter der Annahme, dass die Zeemann-Verschiebung eine untergeordnete Rolle spielt, kann man somit aus den Magnetfeldmessungen auf die ungefähre Größe der Quantenpunkte schließen. Weitergehende Aussagen, z.B. über das Legierungsprofil, sind dagegen nicht möglich. In Ref. [78] wurde für InAs-Quantenpunkte der effektive g-Faktor experimentell bestimmt. Für das Magnetfeld in [001]-Richtung beträgt er -0.8 . Daraus erhält man für eine Feldstärke von $B = 20$ T eine paramagnetische Verschiebung von circa 0.5 meV. Diese ist tatsächlich klein gegenüber der diamagnetischen Verschiebung von knapp 4 meV. Anders verhält es sich bei Magnetfeldrichtung in [100]-Richtung. Hier ist der g-Faktor abhängig von der Feldstärke und kann für $B=10$ T Werte bis zu -0.9 erreichen [78]. Dann ist aber die paramagnetische Verschiebung von derselben Größenordnung wie die berechnete diamagnetische Verschiebung. Deshalb ist im Rahmen unseres Modells für eine Orientierung des Magnetfeldes in (100)-Richtung keine konkrete Aussage möglich.

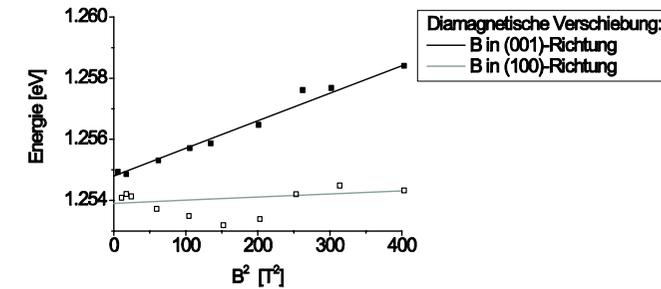


Abbildung 6.28: Vergleich der Messergebnisse aus Ref. [78] mit der berechneten diamagnetischen Verschiebung. Die experimentellen Werte sind durch die Quadrate bezeichnet (offene Quadrate für Magnetfeld in (001)-Richtung und ausgefüllte Quadrate in (100)-Richtung). Für die Rechnung wurde ein Quantenpunkt mit Höhe $H=6$ nm und Basisdurchmesser $L=11$ nm mit einer Indium-Konzentration von 75% zugrunde gelegt.

Kapitel 7

Weitere Anwendungen

7.1 Nano-MOSFET

Die fortschreitende Miniaturisierung in der Halbleitertechnologie stellt die Entwicklung von Silizium-Transistoren vor zunehmende Probleme. Insbesondere führen Kurzkanaleffekte in gewöhnlichen Si-MOSFETs (Metal-Oxide-Semiconductor-Field-Effect-Transistor) mit Gatelängen unter 100 nm zu erheblichen Beeinträchtigungen im elektronischen Verhalten. Vor diesem Hintergrund stellt der Double-Gate-MOSFET ein vielversprechendes Konzept für einen MOS-Transistor mit Gatelängen bis zu 20 nm dar [74][75]. Die Grundidee besteht darin, den Kanal möglichst gut zu kontrollieren, indem man die Kanalbreite möglichst klein wählt und an beiden Seiten einen Gatekontakt vorsieht. Dadurch können Kurzkanaleffekte unterdrückt werden, zudem ergibt sich ein höherer Strom als bei nur einem Gate [74]. Aufgrund der geringen Kanalbreite von weniger als 10 nm dürfen quantenmechanische Effekte dagegen nicht mehr vernachlässigt werden. In diesem Kapitel wird untersucht, inwiefern quantenmechanische Effekte das Schaltverhalten eines Double-Gate-MOSFETs beeinträchtigen und dabei insbesondere auf Kurzkanaleffekte eingegangen. Anschließend vergleichen wir die Eigenschaften mit denen eines Single-Gate-MOSFETs. Die Simulationsergebnisse sind im Rahmen einer Kooperation mit der Forschungsabteilung der Firma Infineon Technologies entstanden.

7.1.1 Double-Gate-MOSFET

Struktur

In Abb. 7.1 ist der Aufbau des betrachteten Transistors skizziert. Er besteht aus einem Siliziumstreifen von 120 nm Länge und 9 nm Breite. An beiden Enden befinden sich ohmsche Kontakte mit stark n-dotierten Kontaktregionen von ca. 40 nm Ausdehnung. Das zugehörige Dotierprofil ist ebenfalls in Abb. 7.1 dargestellt. Da das elektrische Verhalten im Wesentlichen durch die Kanalregion

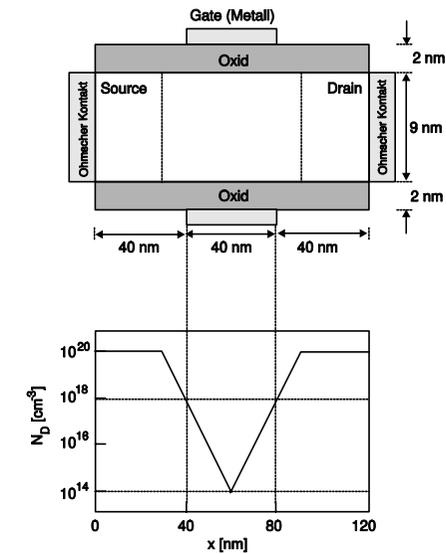


Abbildung 7.1: Geometrie des betrachteten Double-Gate-MOSFETs. Der Kanal besteht aus Silizium mit stark n-dotierten Kontaktregionen. Die Gatekontakte bestehen aus Aluminium. Das Dotierprofil ist logarithmisch gezeichnet und fällt im Gatebereich stark ab.

unter dem Gatekontakt bestimmt wird, möchte man dort eine möglichst geringe Dotierung haben, da sich z.B. bei einem Kanalquerschnitt von $9\text{ nm}\times 9\text{ nm}$, einer Kanallänge von 40 nm und einer Dotierkonzentration von $N_D = 10^{18}\text{ e/cm}^3$ nur etwa vier Dotieratome in der Kanalregion befinden, welche statistisch verteilt sind und daher von Bauelement zu Bauelement zu fluktuierenden Schwellenspannungen führen. Da ein scharfes Dotierprofil nur schwer erzielbar ist, wird nur bis einen Abstand von 10 nm vor die Gate-Region dotiert, woraus sich das angegebene Profil ergibt. Dies hat jedoch den Nachteil, dass in den Gebieten zwischen Source und Gate bzw. Gate und Drain die Dotierkonzentration abnimmt, was zu einem seriellen Widerstand führt. Insgesamt ist der erreichbare Strom dann nur etwa halb so groß wie bei einer konstanten Dotierung bis zum Gatebereich.

Elektronendichte im Kanal

Anhand der Abbildungen 7.2-7.4, welche jeweils für unterschiedliche Gate-Spannungen einen Querschnitt durch den Kanal darstellen, kann die Wirkung des Transistors studiert werden. Bei einer Gate-Spannung von -0.3 V (Abb. 7.2) ist der Kanal geschlossen, die Leitungsbandkante liegt 150 meV über der Fermienergie und die Elektronendichte ist sehr gering. Bei quantenmechanischer Rechnung ist die Dichte kleiner als bei klassischer Rechnung und sie fällt am Heterotübergang zum Oxid auf Null ab, da die Wellenfunktionen nur sehr wenig in die Barriere eindringen können. Dies macht plausibel, dass der quantenmechanisch berechnete Strom kleiner ist als der klassisch berechnete. Die Schwellenspannung liegt bei etwa -0.1 V , dann beginnt sich der Kanal zu öffnen. Bei einer Gate-Spannung von 0.1 V (Abb. 7.3) liegt das Leitungsband in den Grenzregionen zum Oxid bereits etwas unterhalb der Fermienergie, die Dichten erreichen Werte der Größenordnung 10^{19} e/cm^3 . Bei klassischer Rechnung ist dabei die Elektronendichte in der Kanalmitte klein und nimmt nach außen hin stark zu, während sie bei quantenmechanischer Rechnung mehr in der Mitte lokalisiert ist. Dabei ist zu beachten, dass das Dichteprofil im Wesentlichen durch das unterste Subband geprägt ist. Bei diesem ist die effektive Elektronenmasse mit $0.9m_0$ relativ hoch, so dass die Wellenfunktionen zum Teil in den Potenzialmulden zum Oxid hin lokalisiert sind. Man erhält damit zwei Dichtemaxima. Diese Effekte sind bei vollständig offenem Kanal noch deutlicher ausgeprägt (Abb. 7.4). Insgesamt ist die integrierte quantenmechanische Dichte aufgrund der zusätzlichen Quantisierungsenergie in allen Fällen kleiner als die klassische, so dass der quantenmechanische Strom ebenfalls kleiner ist.

Kennlinien

Die Ausgangskennlinie (d.h. Gate-Spannung konstant, Drain-Spannung wird variiert) ist in Abb. 7.5 für klassische und quantenmechanische Rechnung bei unterschiedlichen Gate-Spannungen dargestellt. Insgesamt kann man festhalten, dass

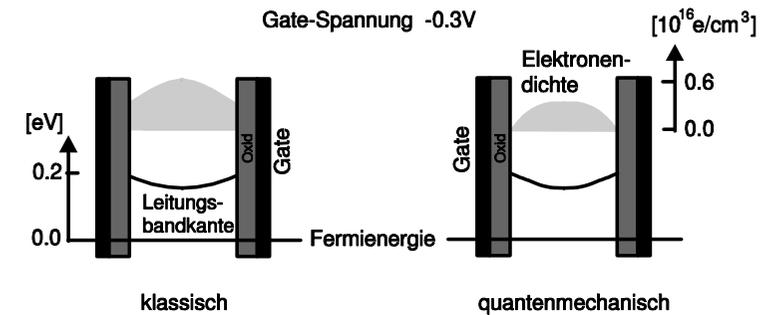


Abbildung 7.2: Klassische und quantenmechanische Dichte bei geschlossenem Kanal. Die Gate-Spannung beträgt -0.3 V .

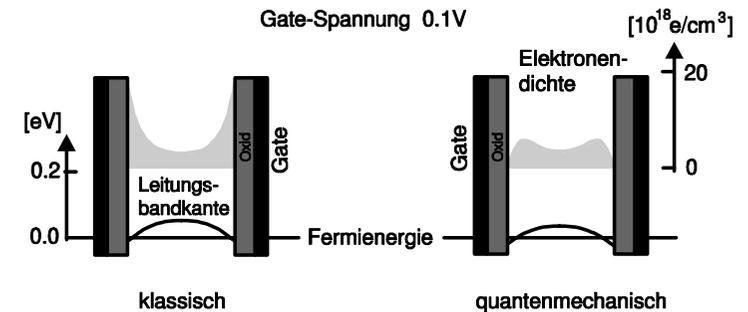


Abbildung 7.3: Klassische und quantenmechanische Dichte bei leicht geöffnetem Kanal. Die Gate-Spannung beträgt 0.1 V .

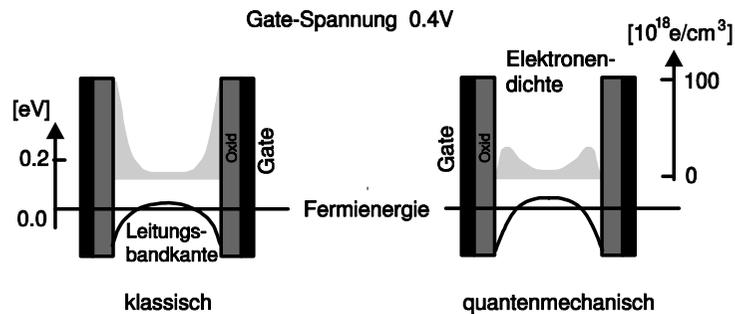


Abbildung 7.4: Klassische und quantenmechanische Dichte bei offenem Kanal. Die Gate-Spannung beträgt 0.4V.

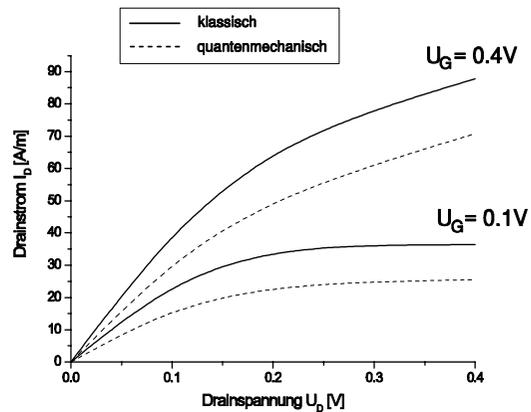


Abbildung 7.5: Ausgangskennlinie (Gate-Spannung konstant, Drain-Spannung wird variiert) für den Double-Gate-MOSFET mit Kanalbreite von 9 nm und Gate-Länge von 40 nm. Der quantenmechanische Strom (gestrichelt) ist circa 30% kleiner als der klassische (durchgezogen).

der quantenmechanisch berechnete Strom etwa 30% kleiner ist als der klassisch berechnete. Als Modell für die Beweglichkeit wurde folgender Zusammenhang verwendet:

$$\mu_{n,E0}(N_D) = \mu_{n,min} + \frac{\mu_{n,D}}{1 + (N_D/N_{n,ref})^{\alpha_n}}$$

$$\mu_n(\vec{E}) = \frac{\mu_{n,E0}}{\left[1 + \left(\mu_{n,E0} \frac{|\vec{E}|}{v_{n,s}}\right)^{\kappa_n}\right]^{1/\kappa_n}}$$

Dabei gilt: $\mu_{n,min} = 55.2 \text{ Vs/cm}^2$, $N_{n,ref} = 1.072 \times 10^{17} \text{ cm}^{-3}$, $\mu_{n,D} = 1374 \text{ Vs/cm}^2$, $\alpha_n = 0.73$, $v_{n,s} = 1.03 \times 10^7 \text{ cm/s}$ und $\kappa_n = 2.0$. \vec{E} ist das elektrische Feld.

Zweidimensionaler Verlauf des Potentials und des Fermi-niveaus

In den Abbildungen 7.6 und 7.7 sind der Bandkantenverlauf und das Quasi-Fermi-niveau der Elektronen für eine Gate-Spannung von 0.4V und eine Drain-Spannung von 0.2V dargestellt. Da, wie oben erwähnt, die Dotierung nicht ganz bis zur Gate-Region reicht, bildet sich in den Bereichen zwischen Gate und Drain bzw. Gate und Source eine leichte Barriere aus, an der die Dichte kleiner ist als im Kanalbereich am Gate. Dort fällt daher das Quasi-Fermi-niveau stark ab. Wie in Kapitel 3 beschrieben, lösen wir die Schrödingergleichung und die Transportgleichung im gesamten Bauelement, also auch im Oxid. Da die quantenmechanische Dichte im Oxid sehr klein ist, hat das Quasi-Fermi-niveau dort einen starken Gradienten.

Eine Auswahl von Wellenfunktionen ist in Abb. 7.8 dargestellt. Man sieht, dass diese Wellenfunktionen den sog. Airy-Funktionen sehr ähneln. Diese erhält man durch Lösen der Schrödingergleichung in einem schrägen Potenzial. Da alle Eigenfunktionen orthogonal zueinander sein müssen, hat der linke Wellenberg immer den größten Wert.

Transferkennlinien

Entscheidend für das An- und Abschalten des Transistors ist die Transferkennlinie, welche für den Double-Gate-MOSFET in Abb. 7.9 für klassische und quantenmechanische Rechnung dargestellt ist. Die quantenmechanische Kurve hat die gleiche Steilheit wie die klassische, sie ist nur um circa 25 mV nach rechts, d.h. zu höheren Gate-Spannungen hin, verschoben. Konkret bedeutet dies, dass das An- und Abschaltverhalten des Transistors sich durch quantenmechanische Effekte nicht verändert, aber die Einsatzspannung etwas zunimmt. Die Transferkennlinie weist für den klassischen und den quantenmechanischen Fall eine Steilheit von 60 mV/Dekade auf. In Abb. 7.10 sind die entsprechenden Kennlinien für einen Double-Gate-MOSFET mit einer Kanalbreite von 3 nm statt 9 nm, dargestellt.

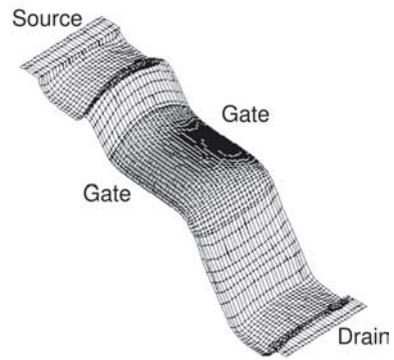


Abbildung 7.6: Leitungsbandkante bei einer Drain-Spannung von 0.2 V und einer Gate-Spannung von 0.4 V.

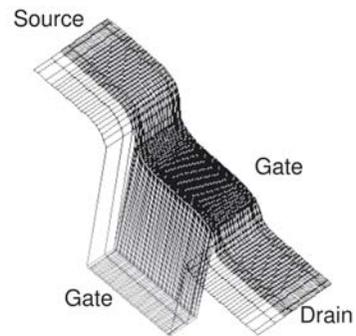


Abbildung 7.7: Verlauf des Quasi-Ferminiveaus bei quantenmechanischer Rechnung für eine Drain-Spannung von 0.2 V und eine Gate-Spannung von 0.4 V.

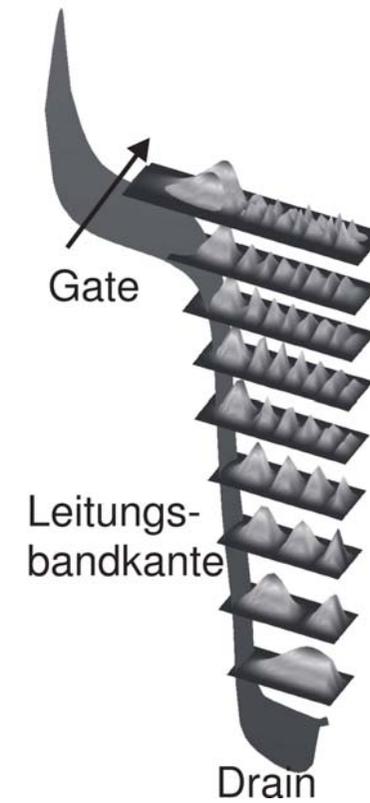


Abbildung 7.8: Die Grafik stellt den Bandkantenverlauf im Gate- bis Drainbereich dar, sowie eine Auswahl der Wellenfunktionen in diesem Bereich.

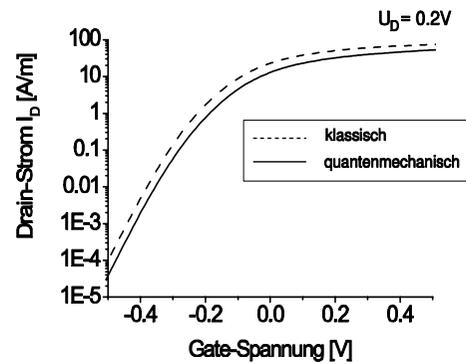


Abbildung 7.9: Transferkennlinie für einen Double-Gate-MOSFET mit einer Kanalbreite von 9 nm bei klassischer (gestrichelt) und quantenmechanischer (durchgezogen) Rechnung. Die quantenmechanische Kennlinie hat dieselbe Steigung, ihre Einsatzspannung ist 20 mV größer.

Jetzt ist die Einsatzspannung schon um 80 mV nach oben verschoben, d.h. die Änderung ist umso größer, je kleiner die Kanalbreite ist. Man kann sich dieses Verhalten folgendermaßen vorstellen: Je schmaler der Silizium-Kanal, desto größer ist die Quantisierungsenergie für die Elektron-Wellenfunktionen in lateraler Richtung. Dies führt zu einer höheren energetischen Barriere, die die Elektronen überwinden müssen. Schaltet man den Transistor in den Durchlassbereich, so muss man eine um die Quantisierungsenergie höhere Gate-Spannung anlegen, um die Barriere abzubauen, die Transfer-Kennlinie verschiebt sich entsprechend nach rechts. Entsprechend der Quantisierungsenergien in einem Quantentopf würde man eine $1/w^2$ -Abhängigkeit der Verschiebung der Einsatzspannung erwarten, wobei w die Kanalbreite darstellt. In Abb. 7.11 ist die Grundzustandsenergie in einem Quantentopf als Funktion der Topfbreite dargestellt (gestrichelt) bei einer effektiven Masse von $m = 0.56m_0$. Wie in Abb. 7.11 zu erkennen ist, ändert sich die Verschiebung der Einsatzspannung mit der Kanalbreite aber nicht so stark wie obiges Modell vorhersagt. Der Grund liegt darin, dass beim Schließen des Kanals die Leitungsbandkanten zum Oxid hin ansteigen und in der Kanalmitte eine Mulde bilden. Dies wurde vorher in Zusammenhang mit Abb. 7.2 diskutiert. Verbreitert man nun den Kanal, so nimmt die Topfbreite nicht in gleichem Maße zu, die Barrierenhöhe durch die Quantisierungsenergie folgt demnach nicht einem $1/w^2$ -Gesetz (w Kanalbreite).

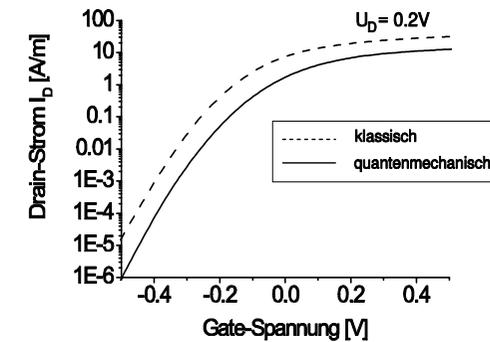


Abbildung 7.10: Transferkennlinie für einen Double-Gate-MOSFET mit einer Kanalbreite von 3 nm bei klassischer (gestrichelt) und quantenmechanischer (durchgezogen) Rechnung. Die quantenmechanische Kennlinie hat dieselbe Steigung, ihre Einsatzspannung ist 80 mV größer.

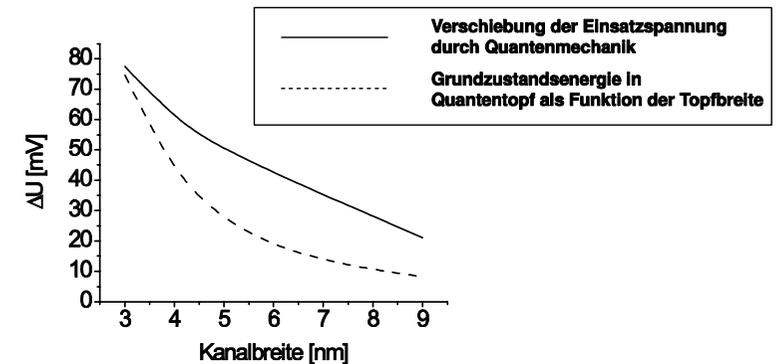


Abbildung 7.11: Verschiebung der Einsatzspannung des Double-Gate-MOSFETs bei Vergleich zwischen klassischer und quantenmechanischer Rechnung für verschiedene Kanalbreiten. Im Vergleich dazu (gestrichelt) die Quantisierungsenergie des Grundzustands in einem Quantentopf als Funktion der Topfbreite (effektive Masse $m = 0.56m_0$).

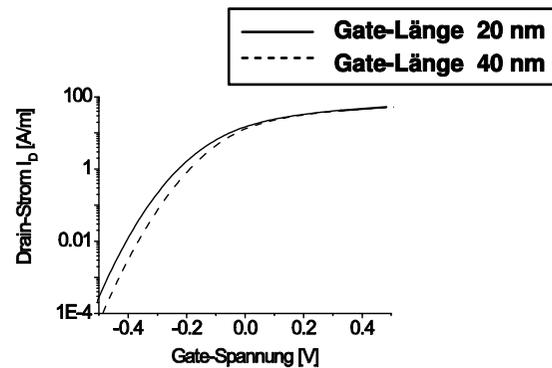


Abbildung 7.12: Transferkennlinie (quantenmechanisch) für einen Double-Gate-MOSFET mit 9 nm Kanalbreite und einer Gate-Länge von 40 nm (gestrichelt) bzw. 20 nm (durchgezogen). Für die kürzere Gate-Länge ergibt sich ein schlechteres Abschaltverhalten.

Kurzkanaleffekte

Im Folgenden soll die Abhängigkeit der Transferkennlinie von der Gate-Länge untersucht werden. Wie in Abb. 7.12 zu sehen ist, führt eine kürzere Gate-Länge von 20 nm statt 40 nm zu einem schlechteren Abschaltverhalten. Im Bereich unterhalb der Schwellenspannung ist der Strom bei kürzerem Gate höher. Man kann sich diesen so genannten Kurzkanaleffekt folgendermaßen vorstellen. Je kürzer der Kanal ist, umso stärker beeinflusst die Drain-Spannung den Potenzialverlauf in der Kanalregion. Im abgeschalteten Zustand bedeutet dies, dass die Barriere im Gate-Bereich stärker abgesenkt wird. Dieser Effekt wird auch mit dem Kürzel DIBL (Drain Induced Barrier Lowering) bezeichnet. In Abb. 7.13 ist die Schwellenspannung in Abhängigkeit von der Gate-Länge dargestellt. Die quantenmechanisch berechnete Schwellenspannung zeigt dabei dasselbe Verhalten wie die klassisch berechnete, d.h. sie ist für alle Gate-Längen um circa 25 meV größer aufgrund der lateralen Quantisierung, wie oben bereits diskutiert wurde. Für eine Kanalbreite von 3 nm ist die Änderung der Schwellenspannung mit der Gate-Länge geringer, der Unterschied zwischen quantenmechanisch und klassisch berechneter Schwellenspannung beträgt nun 80 meV. Insgesamt kann man folgern, dass die Kurzkanaleffekte bei quantenmechanischer und klassischer Rechnung in etwa vergleichbare Auswirkungen haben.

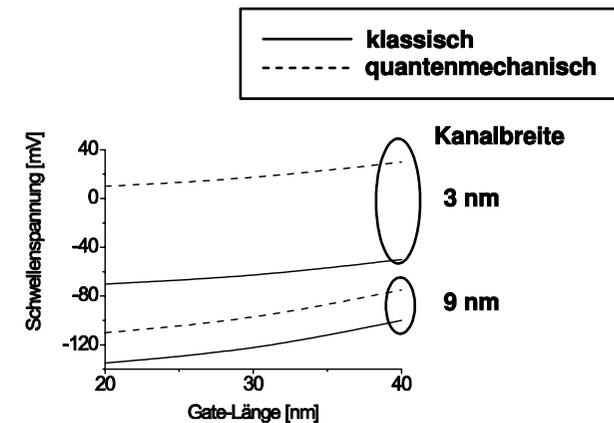


Abbildung 7.13: Vergleich der Schwellenspannung des Double-Gate-MOSFETs bei klassischer (durchgezogen) und quantenmechanischer (gestrichelt) Rechnung als Funktion der Gate-Länge. Für eine Kanalbreite von 3 nm ist die Abhängigkeit von der Gate-Länge geringer als für eine Kanalbreite von 9 nm.

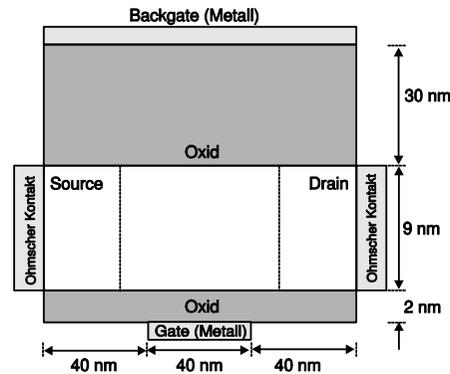


Abbildung 7.14: Der betrachtete Single-Gate-MOSFET hat im Prinzip die gleiche Geometrie wie der Double-Gate-MOSFET mit dem Unterschied, dass der Rückkontakt durch eine 30 nm dicke Oxidschicht vom Silizium-Kanal getrennt ist. Der Rückkontakt liegt auf Erde. Gesteuert wird der Kanal damit nur durch das Gate.

7.1.2 Vergleich mit Single-Gate-MOSFET

Im Folgenden sollen die Resultate für den Double-Gate-MOSFET kurz mit Ergebnissen für einen Single-Gate-MOSFET verglichen werden. Wir nehmen dazu eine identische Geometrie an, das zweite Gate ist jedoch durch eine 30 nm dicke Oxidschicht vom Silizium-Kanal getrennt, so dass der Kanal nur durch das Gate vorne gesteuert wird. Der Rückkontakt ist geerdet (Abb. 7.14). Diese Struktur entspricht einem Single-Gate-MOSFET auf SOI-Basis (Silicon On Insulator). Die Transferkennlinie für eine Drain-Spannung von 0.2 V ist in Abb. 7.15 dargestellt. Einen wesentlichen Unterschied zum Double-Gate-MOSFET stellt das deutlich schlechtere Schaltverhalten dar. Während der Double-Gate-MOSFET eine Steilheit der Ausgangskennlinie von 60 mV/Dekade aufweist, beträgt sie beim SOI-Transistor nur 100 mV/Dekade. Beim Vergleich der klassischen und der quantenmechanischen Transferkennlinie sieht man außerdem, dass die Steilheit der klassischen Kurve (90 mV/Dekade) besser ist als die Steilheit der quantenmechanischen (100 mV/Dekade). Daraus folgt, dass der Single-Gate-MOSFET aufgrund quantenmechanischer Effekte schlechter steuerbar wird.

Betrachtet man die Transferkennlinie für eine Kanalbreite von 3 nm (Abb. 7.16), so erhält man hier ähnliche Verhältnisse wie beim Double-Gate-MOSFET. Die quantenmechanische Kennlinie ist wieder aufgrund der lateralen Quantisierung um circa 80 mV gegenüber der klassischen verschoben, die Steilheit beträgt in beiden Fällen 60 mV/Dekade im Bereich unterhalb der Schwellenspannung.

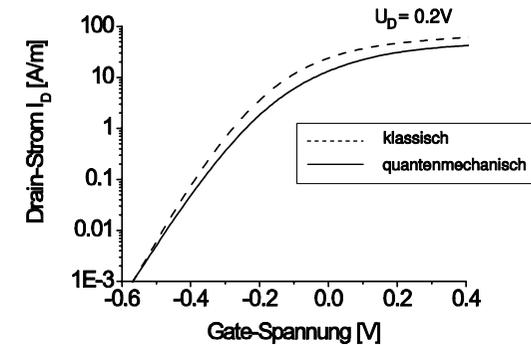


Abbildung 7.15: Single-Gate-MOSFET mit einer Kanalbreite von 9 nm bei klassischer (gestrichelt) und quantenmechanischer (durchgezogen) Rechnung. Die quantenmechanische Kennlinie hat eine geringere Steigung als die klassische.

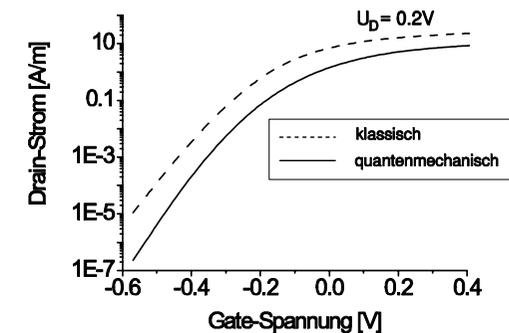


Abbildung 7.16: Single-Gate-MOSFET mit einer Kanalbreite von 3 nm bei klassischer (gestrichelt) und quantenmechanischer (durchgezogen) Rechnung. Die quantenmechanische Kennlinie hat dieselbe Steigung, ihre Einsatzspannung ist 80 mV größer.

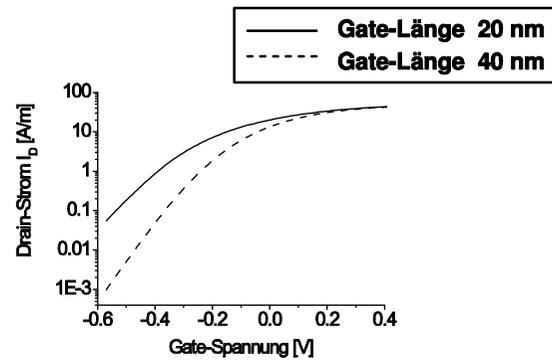


Abbildung 7.17: Transferkennlinie (quantenmechanisch) für einen Single-Gate-MOSFET mit 9 nm Kanalbreite und einer Gate-Länge von 40 nm (gestrichelt) bzw. 20 nm (durchgezogen). Für die kürzere Gate-Länge ergibt sich ein deutlich schlechteres Abschaltverhalten.

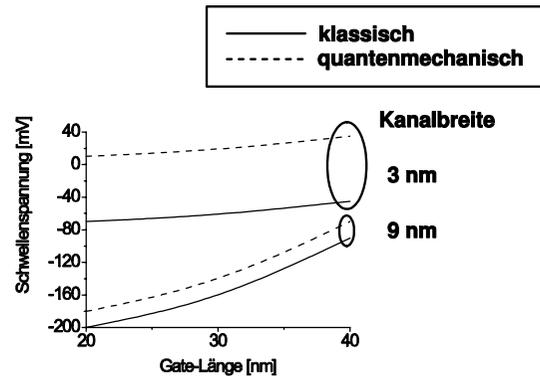


Abbildung 7.18: Vergleich der Schwellenspannung des Single-Gate-MOSFETs bei klassischer (durchgezogen) und quantenmechanischer (gestrichelt) Rechnung als Funktion der Gate-Länge. Für eine Kanalbreite von 3 nm ist die Abhängigkeit von der Gate-Länge geringer als für eine Kanalbreite von 9 nm.

Der Grund liegt darin, dass das einzelne Gate den Kanal umso besser beeinflussen kann, je größer die Gate-Länge im Vergleich zur Kanalbreite ist. Dementsprechend erwarten wir, dass Kurzkanaleffekte für den Single-Gate-MOSFET mit Kanalbreite 9 nm besonders stark ausgeprägt sind. In der Tat verschlechtert die Reduzierung der Gate-Länge auf 20 nm das Abschaltverhalten drastisch, wie in Abb. 7.17 deutlich zu sehen ist. Für kleinere Kanalbreiten sind die Kurzkanaleffekte dagegen wieder schwächer ausgeprägt. In Abb. 7.18 sind wieder die Schwellenspannungen als Funktion der Gate-Längen dargestellt.

7.1.3 Zusammenfassung

In diesem Kapitel wurden Simulationsergebnisse von zweidimensionalen Transistorstrukturen vorgestellt. Dabei ging es uns vor allem um die Unterschiede zwischen klassischer und quantenmechanischer Simulation. Insgesamt bleibt festzuhalten, dass bei quantenmechanischer Rechnung sowohl bei Double-Gate-MOSFETs als auch bei Single-Gate-MOSFETs der quantenmechanische Strom kleiner ist als der klassische. Während sich bei Double-Gate-MOSFETs die Ein-Ausschaltcharakteristik durch quantenmechanische Effekte nicht verändert, ist bei Single-Gate-MOSFETs eine Verschlechterung festzustellen. Kurzkanaleffekte liegen bei klassischer und quantenmechanischer Rechnung in derselben Größenordnung. Sie sind umso stärker ausgeprägt, je kürzer die Gate-Länge im Vergleich zur Kanalbreite wird. Dann kann der Kanal immer schlechter durch das Gate kontrolliert werden.

7.2 Si/SiGe-Quantenkaskaden-Strukturen

Die Mikroelektronik wird heute durch Silizium klar dominiert. Deshalb gibt es intensive Bemühungen, Lichtemitter auf Siliziumbasis zu finden, um optische und elektronische Komponenten auf einem Chip integrieren zu können. Das größte Hindernis für auf Silizium basierende Lichtemitter ist die indirekte Bandlücke in Silizium. Dies spielt jedoch in unipolaren Lasern wie z.B. Quantenkaskaden-Lasern keine Rolle, da die strahlenden Übergänge zwischen Subbändern desselben Bandes stattfinden. Dadurch eröffnet sich die Möglichkeit, Quantenkaskaden-Strukturen in Silizium/Germanium-Technologie zu realisieren. Bei pseudomorphem Wachstum von Si/SiGe-Strukturen auf einem Silizium-Substrat tritt der Bandkantensprung in erster Linie im Valenzband auf. Dann muss - anders als bei Quantenkaskaden-Lasern mit (3)-(5)-Halbleitern - die Kaskaden-Struktur im Valenzband auf Löcherbasis vorgesehen werden. Für derartige Strukturen wurde kürzlich Elektrolumineszenz gemessen [62].

Um Lasertätigkeit zu realisieren, ist es notwendig, Besetzungsinversion zwischen den beteiligten Zuständen zu erreichen. Dazu muss man die Streuzeiten und Streukanäle, mit denen die Zustände entvölkert werden, kennen. Als Streumechanismen kommen strahlende und nichtstrahlende Übergänge in Betracht. Als nichtstrahlender Mechanismus dominiert in den von uns betrachteten Strukturen die Emission optischer Phononen aufgrund der Deformationspotenzial-Streuung. In Quantenkaskaden-Strukturen muss sich zwischen den aktiven Regionen ein Miniband ausbilden, über welches die Löcher die aktive Zone verlassen können.

Intersubband-Elektrolumineszenz in Si/SiGe-Quantenkaskaden-Strukturen

Als ersten Schritt in Richtung Si/SiGe-Laser haben wir Si/SiGe-Quantenkaskaden-Strukturen entworfen, an denen vertikale radiative Übergänge vom zweiten schweren Loch-Subband (HH1, engl. heavy hole) zum ersten schweren Loch-Subband (HH2) gemessen wurden. Die Bandstruktur wurde mit 6×6 -kp-Theorie berechnet und ist in Abb. 7.19 dargestellt. Die Schichtdicken und die jeweilige Ge-Konzentration sind der Bildunterschrift zu entnehmen. Die Struktur wurde mit einer MBE-Anlage (Molecular Beam Epitaxy) gewachsen. Dabei ist zu beachten, dass bei Wachstumstemperaturen um 350°C keine scharfen Grenzflächen wachsen. Diese Segregationseffekte wurden im Rahmen eines kinetischen Modells berücksichtigt. Die Löcher werden vom untersten HH Subband des Quantengrabens (QW, engl. Quantum Well) in das HH2-Niveau des aktiven Quantengrabens QW1 injiziert, wo der strahlende Übergang in den HH1-Zustand erfolgt. Die Schichten QW2-5 sind so ausgelegt, dass die schweren Löcher ein Miniband bilden, durch das die Löcher im HH1-Zustand die aktive Zone verlassen können.

Abb. 7.20 zeigt das Elektrolumineszenz-Spektrum der betrachteten Probe. Experimentell ergibt sich eine Übergangsenergie von 146 meV , die sehr gut mit dem

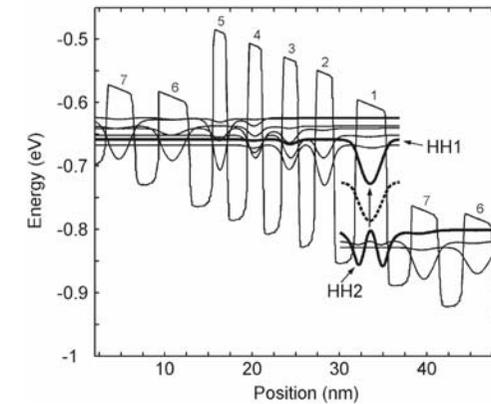


Abbildung 7.19: Berechnete Bandstruktur bei einem angelegten Feld von 55 kV/cm . Eingezeichnet sind die Bandkante des schweren Lochbandes (HH) sowie die HH1- und HH2-Zustände für $k_{\parallel} = 0$. Das leichte Loch (gestrichelt eingezeichnet) mischt sich stark mit dem abgespaltenen Loch. Die Schichtdicke (\AA) und der Ge-Gehalt der Quantengraben und der Si-Barrieren sind $38/0.4$, $23/0.43$, $18/0$, $21/0.43$, $20/0$, $19/0.43$, $22/0$, $17/0.43$, $24/0$, $38/0.24$, $24/0$, $35/0.21$ und $26/0$.

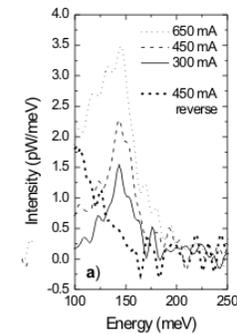


Abbildung 7.20: Elektrolumineszenz-Spektren in Abhängigkeit vom Strom durch die Probe. Auch für Sperr-Richtung ist ein Spektrum eingezeichnet

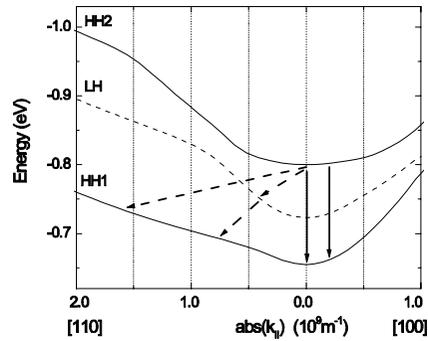


Abbildung 7.21: Energiedispersion der Subbänder HH2, LH1 und HH1 im aktiven Quantengraben QW1. Die durchgezogenen Pfeile stellen die strahlenden Übergänge dar, die experimentell gemessen wurden.

berechneten Wert von 143 meV übereinstimmt. Weiterhin haben wir die Abhängigkeit der Emissionsspektren vom Strom durch die Probe untersucht. Wie man in Abb. 7.20 erkennen kann, ist die Übergangsenergie unabhängig von der Stromstärke. Dagegen nimmt mit zunehmendem Strom die Linienbreite zu. Die Verbreiterung zu kleineren Linien hin wird durch strahlende Übergänge bei höheren Werten von k_{\parallel} verursacht. In Abb. 7.21 ist klar zu erkennen, dass mit zunehmendem k_{\parallel} die Übergangsenergie abnimmt. Bei Polung in Sperr-Richtung wird keine Elektrolumineszenz beobachtet.

Nichtstrahlende Übergänge

Wir haben zusätzlich die Zerfallsrate des Zustands HH2 aufgrund optischer Deformationspotenzial-Streuung berechnet. Dazu wurde das Loch-Phonon-Matrixelement $\langle i | \mathbf{D} \cdot \mathbf{u} | f \rangle$ ausgewertet, wobei $\langle i |$ den Anfangs- und $| f \rangle$ den Endzustand bezeichnen. Der Deformationspotenzial-Tensor hat folgende Gestalt (in der Basis $|x \uparrow\rangle, |y \uparrow\rangle, |z \uparrow\rangle, |x \downarrow\rangle, |y \downarrow\rangle, |z \downarrow\rangle$):

$$\mathbf{D} = \frac{d_0}{a} \left[\begin{pmatrix} d_x & 0 \\ 0 & d_x \end{pmatrix}, \begin{pmatrix} d_y & 0 \\ 0 & d_y \end{pmatrix}, \begin{pmatrix} d_z & 0 \\ 0 & d_z \end{pmatrix} \right] \quad (7.1)$$

mit

$$d_x = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad d_y = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad d_z = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Die optische Deformationspotenzial-Konstante hat einen Wert von $d_0 = 36$ meV, a ist die Gitterkonstante. Die Auslenkung eines Atoms ist gegeben durch $\mathbf{u}(\vec{q}) \sim e^{i\vec{q}\cdot\vec{r}}$. Dabei ist \vec{q} der Impuls, den das Phonon aufnimmt. Für die Phonondispersion nehmen wir konstante Energie an. Das Phonon-Spektrum besteht aus den Si-Si (64.3 meV), Si-Ge (50.8 meV) und den Ge-Ge (37.4 meV) Moden.

Die Streuzeit ergibt sich nun durch Anwenden von Fermis Goldener Regel und Integration über alle Endzustände. Die Gesamtdepopulationszeit für den HH2-Zustand beträgt 400 fs. Sie ist beträchtlich kleiner als die typischen Werte von (3)-(5)-Quantenkaskadenlasern. In Abb. 7.21 sind die beiden betrachteten Kanäle eingezeichnet. Bei Streuung in den HH1-Zustand beträgt die Streuzeit 614 fs, bei Streuung in das leichte Lochband (LH) 1170 fs. Obwohl der Deformationspotenzial-Tensor die HH-LH-Streuung gegenüber der HH-HH-Streuung begünstigt, ist hier letzterer Mechanismus schneller, da die Dispersion in [110]-Richtung sehr flach und damit die Zustandsdichte sehr groß ist.

Ausblick: Konzept für Laser

Bisher haben wir eine Quantenkaskadenstruktur entworfen, die Bandstruktur und die nichtstrahlenden Übergänge berechnet. Experimentell konnte Elektrolumineszenz gemessen werden. Ziel ist es letztlich, eine Laser zu bauen. Aufgrund der obigen Rechnungen kommt dabei nur die Konstellation in Betracht, die in Abb. 7.22 dargestellt ist. Der Laserübergang findet statt vom schweren Loch im Quantengraben QW7 (HH2) zum schweren Loch (HH1) bzw. leichten Loch (LH1) im Quantengraben QW1 (HH1). Konkret heißt dies, dass das schwere Loch HH2 über nichtstrahlende Streuung entvölkert, dabei in den Quantengraben QW1 gestreut wird und diesen schließlich über Tunnelprozesse verlässt. Die Tunnelzeit liegt dabei in der Größenordnung von 2-3 ps. Um Besetzungsinversion zu erreichen, muss die Streuzeit für den nichtstrahlenden Übergang größer werden als die Tunnelzeit. Die Streuzeit kann man variieren, indem man die Barrierendicke zwischen Quantengraben QW7 und Quantengraben QW1 vergrößert. In Abb. 7.22 ist rechts die berechnete Gesamtstreuzeit der nichtstrahlenden Übergänge als Funktion der Barrierendicke dargestellt. Für wachsende Barrierendicke steigt sie an, da dann der Überlapp der Wellenfunktionen kleiner wird, und erreicht bei einer Dicke von 3 nm einen Wert von 15 ps erreicht. Gleichzeitig muss aber die Kopplung an das Lichtfeld hinreichend groß sein, um Lasertätigkeit zu erreichen. Mit zunehmender Barrierendicke nimmt aber das optische Matrixelement zwischen dem schweren Loch im Quantengraben QW7 (HH2) und dem schweren Loch (HH1) im Quantengraben QW1 ab. Als nächster Schritt ist geplant, konkret diese Strukturen herzustellen und die Machbarkeit eines Lasers auf dieser Grundlage experimentell zu untersuchen.

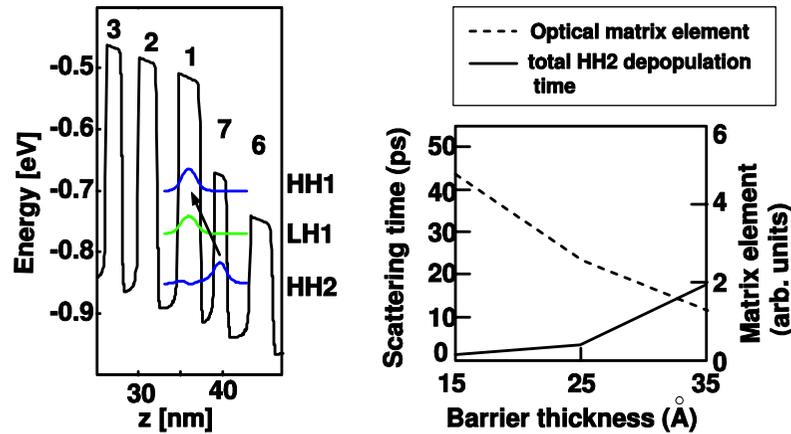


Abbildung 7.22: Konzept für einen Laser auf Si/SiGe-Basis. Links ist der Verlauf des Valenzbandes dargestellt. Der Laserübergang findet zwischen dem schweren Loch im Quantengraben QW7 (HH2) und dem schweren Loch im Quantengraben QW1 (HH1) statt. HH2 wird über nichtstrahlende Phononen-Streuung entvölkert. Das Teilchen verlässt schließlich durch Tunneln den Quantengraben QW1. Um Besetzungsinversion zu erreichen, muss die Streuzeit des nichtstrahlenden Übergangs größer sein als die Tunnelzeit. Rechts sind die Streuzeiten als Funktion der Barrierendicke zwischen Quantengraben QW7 und Quantengraben QW1 dargestellt ebenso wie das optische Matrixelement für den Übergang HH2-HH1.

7.3 Zweidimensionale Löchergase in GaN/AlGaIn-Heterostrukturen

Effiziente elektronische und optoelektronische Bauelemente auf Nitrid-Basis erfordern eine ähnlich hohe Konzentration von freien Elektronen und Löchern. Dies ist sehr schwer in Nitriden zu erreichen, da die Ionisierungsenergien der für p-Dotierung verwendeten Mg-Akzeptoren sehr groß sind. In der Tat stellte dies lange Zeit das Haupthindernis für die Entwicklung von Hetero-Bipolartransistoren oder Laserdioden im UV-Bereich dar. Erst vor kurzem wurde vorgeschlagen, dass der geringe Ionisierungsgrad der tiefen Mg-Akzeptoren in Heterostrukturen durch Ausnutzen von piezo- und pyroelektrischen Polarisationsfeldern deutlich verbessert werden kann [79]-[83]. Unverspannte Nitrid-Schichten zeigen spontane Polarisation, während in verspannten Heterostrukturen zusätzlich piezoelektrische Polarisation auftritt. Der Unterschied in der Polarisation induziert Flächenladungen an Heterogrenzflächen. Diese wiederum erzeugen starke elektrische Felder, welche die Energieniveaus der Akzeptoratome unter das Fermi-niveau drücken können. In diesem Abschnitt stellen wir ausführliche theoretische Rechnungen vor über die elektronische Struktur von GaN/AlGaIn-Quantengraben und Übergittern. Es zeigt sich, dass die Dichte des zweidimensionalen Elektronengases im Wesentlichen durch die Polarisationsladung an GaN/AlGaIn-Grenzflächen bestimmt wird. Andere Faktoren, wie die Barrierendicke, die Temperatur und das Dotierprofil spielen eine untergeordnete Rolle.

Die elektronische Struktur im Valenzband berechnen wir mit 6×6 -kp-Theorie, wobei die Abhängigkeit der Eigenzustände vom lateralen Wellenvektor vollständig berücksichtigt wird. Die Löcherdichte ergibt sich durch Summation über die diskreten Eigenzustände, gewichtet mit der Fermi-Verteilungsfunktion, und durch Integration über die laterale Brillouin-Zone (vgl. Kapitel 2 und Anhang D). Die kp-Parameter einschließlich der Deformationspotenziale stammen aus ab-initio-Rechnungen [84] und wurden für AlGaIn-Legierungen linear interpoliert. Die Berechnung der Verspannungen erfolgte im Rahmen der Elastizitätstheorie (vgl. Kapitel 4). Dabei wurden die elastischen Konstanten in den entsprechenden Volumenhalbleitern zugrunde gelegt [85].

Zunächst wollen wir Resultate für eine GaN/AlGaIn-Heterostruktur vorstellen. Diese besteht aus einer GaN-Substrat-Schicht und einer in [0001]-Richtung pseudomorph aufgewachsenen AlGaIn-Barriere, die durch einen Schottky-Kontakt abgeschlossen wird. Um die Polarisationsladung an der Grenzfläche zu erhöhen, muss die Polarität der Probe vom N-Typ sein, d.h. der spontane Polarisationsvektor muss aus dem Substrat herauszeigen. Für die GaN-Schicht nehmen wir eine homogene Mg-Konzentration von $1.6 \times 10^{17} \text{ cm}^{-3}$ an. Die angrenzende AlGaIn-Barriere besteht aus einer 5 nm dicken undotierten Spacer-Schicht und einer homogen Mg-dotierten Schicht ($[\text{Mg}] = 5 \times 10^{19} \text{ cm}^{-3}$). In unseren Rechnungen wurde die Barrierendicke von 15 nm bis 35 nm und der Al-Gehalt von 0% bis

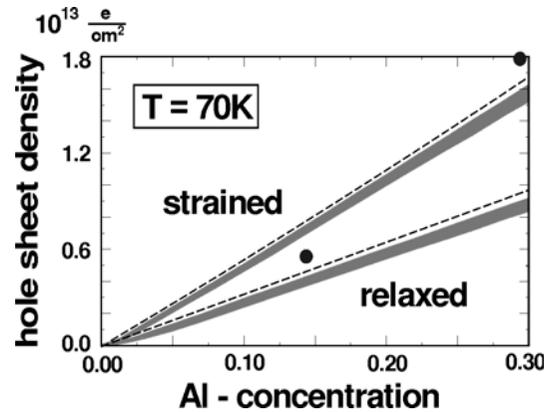


Abbildung 7.23: Berechnete Flächenladungsdichte des Löchergases für die verspannte und relaxierte GaN/AlGaN-Heterostruktur mit Schottky-Kontakt als Funktion der Al-Konzentration. Die grau unterlegten Linien bezeichnen die berechneten Dichten für Barrieren-Dicken von 15 nm (untere Kante der grau unterlegten Balken) bis 35 nm (obere Kante). Die gestrichelte Linie ist die Ladungsdichte der Polarisationsladung an der GaN/AlGaN-Grenzfläche. Die beiden Kreise stellen experimentelle Ergebnisse dar [86].

30% variiert. Für die Aktivierungsenergie der Mg-Akzeptoren E_{Mg} und die Barrierenhöhe des Ni-Schottky-Kontakts $e\Phi_b$ gelten die Beziehungen $E_{MG}(x) = (0.7 \text{ eV})x + 0.17 \text{ eV}$ [87] und $e\Phi_b(x) = (1.3 \text{ eV})x + 0.55 \text{ eV}$ [88], wobei x die Al-Konzentration bezeichnet.

Die berechnete Flächenladungsdichte des Löchergases ist in Abb. 7.23 als Funktion der Al-Konzentration zusammen mit der Polarisationsladung dargestellt. Es ist klar zu erkennen, dass das zweidimensionale Löchergas fast so groß ist wie die Ladungsdichte der Polarisationsladung. Diese wird aber nie vollständig abgeschirmt. Die untere Kurve in Abb. 7.23 legt eine vollständig relaxierte AlGaN-Schicht zugrunde, d.h. die Grenzflächenladung wird nur durch spontane Polarisation erzeugt. Dies entspricht der Situation, wo die Barrierendicke die kritische Dicke überschreitet. Einige experimentelle Ergebnisse deuten darauf hin, dass ab einer bestimmten Al-Konzentration ein Übergang von verspannter zu relaxierter Barriere stattfindet. Die beiden Kurven in Abb. 7.23 sollten daher als Grenzfälle betrachtet werden. Dennoch zeigen unsere Rechnungen, dass die zweidimensionalen Löchergase Ladungsdichten der Größenordnung 10^{13} cm^{-2} erreichen, sogar für den Fall vollständig relaxierter Strukturen. Wir möchten darauf hinweisen,

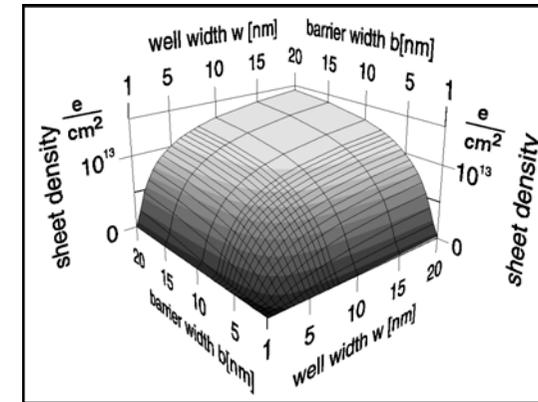


Abbildung 7.24: Berechnete Flächenladungsdichte des Löchergases in einem GaN/AlGaN-Übergitter mit 30% Al-Gehalt in den Barrieren als Funktion der Grabenweite (well width w) und Barrierendicke (barrier width b).

dass die Löcherdichte weitgehend unabhängig von Änderungen anderer Parameter, wie der Temperatur T , der Barrierenhöhe des Schottky-Kontakts $e\Phi_b$ oder der Valenzbanddiskontinuität ΔE_V an der GaN/AlGaN-Grenzfläche ist. Dazu haben wir systematisch T (von 70 K bis 300 K), $e\Phi_b$ (um 50%), ΔE_V (von 0.25 eV bis 0.6 eV) sowie die Dicke der Spacer-Schicht (von 0 bis 5 nm) variiert. Jedesmal änderte sich die Ladungsdichte des Löchergases um weniger als 5%.

Als weiteres Beispiel für ein mögliches Bauelement mit hohen Löcherdichten untersuchen wir periodische GaN/AlGaN-Übergitter bei einer Temperatur von $T=70 \text{ K}$ mit einer Aluminium-Konzentration von 30% und verschiedenen Dicken der GaN-Schicht (Graben) und AlGaN-Schicht (Barriere). In unseren Rechnungen haben wir in den Barrieren als Dotierung eine homogene Mg-Konzentration von $5 \times 10^{19} \text{ cm}^{-3}$ angenommen, in der GaN-Schicht betrug sie $1.6 \times 10^{17} \text{ cm}^{-3}$. Die resultierende Ladungsdichte des Löchergases ist in Abb. 7.24 dargestellt. Die Dichte ist für kurze Graben- und Barrierendicken gering, sie nimmt aber für größere Dicken zu und sättigt schließlich bei einer Übergitterperiode von 40 nm. Dies kann man einsehen, wenn man den Verlauf der Akzeptor-Energieniveaus im Vergleich zum Fermi-niveau betrachtet. In den Abbildungen 7.25 und 7.26 sind die Verhältnisse für ein Übergitter mit einer Periodenlänge von 4 nm bzw. 16 nm dargestellt. In Übergittern erhält man durch die Polarisationsladungen an den Grenzflächen innere elektrische Felder mit wechselnden Vorzeichen in den GaN-Gräben und AlGaN-Barrieren. Daraus ergibt sich ein sägezahnför-

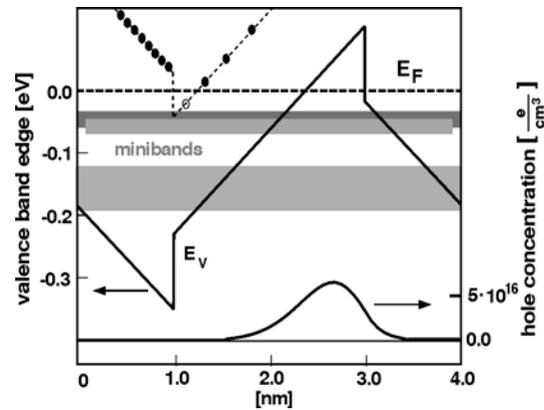


Abbildung 7.25: Verlauf der Valenzbandkante E_V im GaN/AlGaIn-Übergitter (30% Al-Gehalt). Die GaN-Grabenweite und AlGaIn-Barrierendicke betragen jeweils 2 nm. Die gestrichelte Linie markiert das Fermi-Niveau. Die Energien der neutralen und ionisierten Mg-Akzeptor-Niveaus werden durch schwarze bzw. weiße Kreise dargestellt und durch die gepunktete Linie verbunden. Die Minibänder sind grau unterlegt dargestellt. Die Dichte des Löchergases ist unten eingezeichnet.

miger Verlauf der Valenzbandkante. Bei kurzen Übergitterperioden liegen die Akzeptor-Niveaus oberhalb der Fermienergie (Abb. 7.25). Entsprechend ist die Löcherdichte, die von ionisierten Akzeptoren geliefert wird, praktisch Null. In Übergittern mit größeren Schichtdicken sind dagegen viele Akzeptoren ionisiert (Abb. 7.26). Die ionisierten Akzeptoren und die generierten Löcher schirmen die Polarisationsladungen ab und reduzieren damit das innere elektrische Feld. Folglich sättigen die Löcherdichten mit zunehmender Übergitterperiode. Unsere Rechnungen sagen einen Sättigungswert der Flächenladungsdichte des Löchergases von $1.5 \times 10^{13} \text{ cm}^{-2}$ für GaN/AlGaIn-Übergitter mit 30% Aluminium-Gehalt voraus. Dieser Wert liegt in derselben Größenordnung wie die Löcherdichte in der GaN/AlGaIn-Heterostruktur, die weiter oben untersucht wurde.

In den Abbildungen 7.25 und 7.26 sind außerdem jeweils die untersten drei Minibänder eingezeichnet. Bei dem Übergitter mit Periode 4 nm liegt das erste Miniband 35.8 meV unterhalb des Fermi-niveaus. Aufgrund der kurzen Barrierendicke ist der Überlapp zwischen den Lochwellenfunktionen aus benachbarten GaN-Gräben beträchtlich. Infolgedessen sind die Minibänder 23.56 meV, 23.84 meV und 71.4 meV breit. Bei einer Übergitterperiode von 16 nm liegen dagegen die obersten zwei Minibänder über dem Fermi-niveau, und alle drei Bänder haben

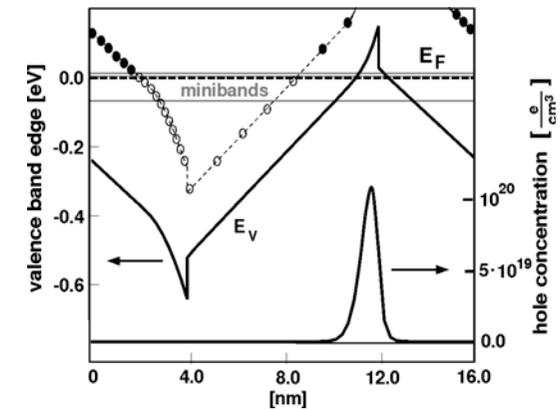


Abbildung 7.26: Siehe Abb. 7.25, aber mit einer GaN-Grabenweite und AlGaIn-Barrierendicke von jeweils 8 nm.

eine sehr kleine Breite von jeweils 1.5 meV

Wir haben weitere Möglichkeiten ausgelotet, um noch höhere Lochdichten in GaN/AlGaIn-Übergittern durch Optimieren des Dotierprofils zu erhalten. Dabei sind wir einem Vorschlag aus Ref. [79] gefolgt und haben Übergitter mit einer Modulationsdotierung untersucht, wo die Hälfte des Grabens und die Hälfte der Barriere in der Umgebung der positiven Grenzflächenladung homogen mit $[\text{Mg}] = 5 \times 10^{19} \text{ cm}^{-3}$ dotiert ist. Es zeigt sich, dass dieses Dotierprofil die Löcherdichte speziell für kurze Übergitterperioden erhöht. Konkret erhöht sie sich von $3.9 \times 10^9 \text{ cm}^{-2}$ auf $5.2 \times 10^{11} \text{ cm}^{-2}$, von $7.7 \times 10^{12} \text{ cm}^{-2}$ auf $1.0 \times 10^{13} \text{ cm}^{-2}$ bzw. von $1.4 \times 10^{13} \text{ cm}^{-2}$ auf $1.5 \times 10^{13} \text{ cm}^{-2}$ bei Übergitterperioden von 4, 12 bzw. 40 nm (jeweils gleich dicke GaN- und AlGaIn-Schichten). Diese starke Änderung für kurze Übergitterlängen kann man anhand der Abb. 7.25 verstehen. Da die Region, in der die Akzeptoren ionisiert werden, sehr klein ist, variiert die Löcherdichte linear mit der Dotierkonzentration. Für große Übergitterlängen sättigt sie dagegen wieder (Abb. 7.24 und 7.26).

Zusammenfassend zeigen unsere selbstkonsistenten Rechnungen, dass in GaN/AlGaIn-Heterostrukturen und Übergittern zweidimensionale Löcherdichten von bis zu $1.5 \times 10^{13} \text{ cm}^{-2}$ erreichbar sind. Diese hohen Ladungsdichten sind Folge der starken piezo- und pyroelektrischen Eigenschaften der Wurtzit-Nitride.

Anhang A

Diskretisierung der Poisson- und Stromgleichung

In Kapitel 4.1 wurden die Poisson- und die Stromgleichung dargestellt und die Randbedingungen kurz erläutert. Im Folgenden wollen wir näher auf die Diskretisierung dieser Gleichungen und ihre numerische Lösung eingehen. Dabei stehen insbesondere Fragestellungen der technischen Umsetzung im Vordergrund. Als Beispiel dient ein konkretes zweidimensionales Bauelement.

Allgemeines

Bei der Poisson- und Stromgleichung handelt es sich um Flussgleichungen folgender Art:

$$-\operatorname{div} \vec{F}(\vec{x}) = k(\vec{x}). \quad (\text{A.1})$$

Durch Integration über ein finites Volumen kann sie unter Verwendung des Gaußschen Satzes umgeschrieben werden zu:

$$-\iint_{V_B} \operatorname{div} \vec{F} dV = -\iint_{A_B} \vec{F} d\vec{A} = \iiint_{V_B} k(\vec{x}) dV \quad (\text{A.2})$$

Der Gaußsche Satz gilt in dieser Form für ein, zwei und drei Dimensionen gleichermaßen. Diese Gleichung wird für jeden Gitterpunkt aufgestellt. Das finite Volumen V_B ist dabei eine Box, deren Randfläche A_B durch die Mittelsenkrechten der Verbindungslinie zweier benachbarter Gitterpunkte definiert ist. Die Oberflächenintegration kann man leicht durchführen, indem man jeweils den Inhalt einer Teilfläche mit der Normalkomponente von \vec{F} multipliziert. Dabei ist $\vec{F} = \varepsilon(\vec{r}) \operatorname{grad}\Phi$ für die Poissongleichung bzw. $\vec{F} = \vec{j}_n$ für die Stromgleichung der Elektronen. Die Normalkomponenten zeigen jeweils in eine Koordinatenrichtung. Die Größe k wird als stückweise konstant angenommen, d.h. für das Integral auf der rechten Seite von Gl. (A.2) ergibt sich $\iiint_{V_B} k(\vec{x}) dV = k_i V_B$.

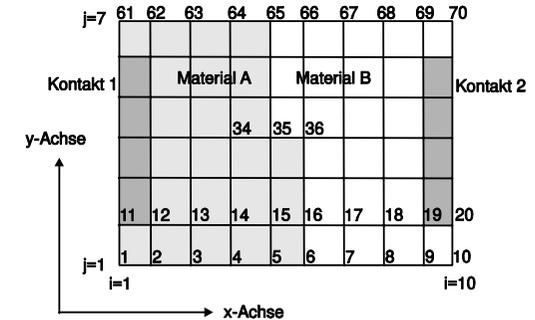


Abbildung A.1: Zweidimensionales Bauelement, bestehend aus Material A und Material B und den beiden Kontakten 1 und 2. Eingezeichnet sind die Gitterlinien des Rechtecksgitters. Man beachte, dass an einem Materialübergang der Gitterpunkt auf der Grenzfläche liegt.

Im Falle der Poissongleichung ist k die Dichte, im Falle der Stromgleichung die Rekombinationsrate. Auf diese Weise erhält man für jeden inneren Punkt eine lineare Gleichung, in die alle direkt benachbarten Punkte eingehen. Daraus ergibt sich bei spaltenweiser Durchnummerierung aller Punkte im Prinzip eine Matrixstruktur mit einer Haupt- und 2 bzw. 4 bzw. 6 (in 1D bzw. 2D bzw. 3D) Nebendiagonalen. Diese klare Struktur wird aber durch die spezielle Behandlung der Heteroübergängen zerstört. Dies wird weiter unten dargestellt.

Gitter

Das Bauelement wird auf ein inhomogenes Rechtecksgitter abgebildet. Ein Beispiel in zwei Dimensionen hierfür mit 10×7 -Gitterpunkten ist in Abb. A.1 gegeben. Die Gitterpunkte werden in der dort dargestellten Weise von 1 bis 70 durchnummeriert. Jeweils vier benachbarte Gitterpunkte bilden ein Rechteck (in 3D jeweils acht einen Quader), dem eine Materialart zugewiesen werden kann.

Differenzenstern

Es soll nun anhand der Poissongleichung das Aufstellen der Matrix erläutert werden. Dazu wird der Reihe nach für jeden Gitterpunkt der Differenzenstern gebildet und in die Matrix übernommen. Betrachten wir konkret den Gitterpunkt mit der Nummer 35, d.h. $i = 5$ und $j = 4$. Das Kontrollvolumen (vgl. Kapitel 4.4) kann in vier Quadranten (in 3D acht Oktanten) unterteilt werden, die in Abb. A.2 mit Q1-Q4 bezeichnet sind. Für die Quadranten Q1-Q3 sind die

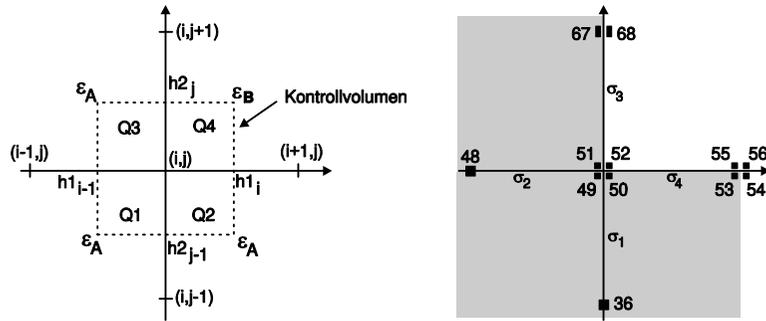


Abbildung A.2: Kontrollvolumen für den Gitterpunkt $(i = 5, j = 4)$, über das die Integration durchgeführt wird (links). Es besteht aus den Quadranten Q1-Q4. Die Integration wird stückweise über Q1-Q4 durchgeführt. Im Bild rechts sind die Nummern der Punkte, die in den Differenzstern eingehen, eingezeichnet (siehe Text). Innerhalb des Kontrollvolumens gibt es vier Grenzflächen zwischen je zwei Quadranten, denen eine Flächenladung $\sigma_1 - \sigma_4$ zugewiesen werden kann.

Materialparameter des Materials A, für den Quadranten Q4 die Parameter des Materials B gültig. Die Integration aus Gl. (A.2) $\iint_{A_B} \vec{F} d\vec{A}$ wird stückweise für die Quadranten ausgeführt. Im Einzelnen gilt mit dem Potenzial Φ , der dielektrischen Konstante ε_A bzw. ε_B für das Material A bzw. B und der Beziehung $\vec{F} = \varepsilon(\vec{r}) \text{grad}\Phi(\vec{r})$:

$$-\iint_{Q_1} \vec{F} d\vec{A} = +\varepsilon_A \frac{\Phi_{i,j} - \Phi_{i-1,j} h2_{j-1}}{h1_{i-1}} + \varepsilon_A \frac{\Phi_{i,j} - \Phi_{i,j-1} h1_{i-1}}{h2_{j-1}} \quad (\text{A.3})$$

$$-\iint_{Q_2} \vec{F} d\vec{A} = -\varepsilon_A \frac{\Phi_{i+1,j} - \Phi_{i,j} h2_{j-1}}{h1_i} + \varepsilon_A \frac{\Phi_{i,j} - \Phi_{i,j-1} h1_i}{h2_{j-1}} \quad (\text{A.4})$$

$$-\iint_{Q_3} \vec{F} d\vec{A} = +\varepsilon_A \frac{\Phi_{i,j} - \Phi_{i-1,j} h2_j}{h1_{i-1}} - \varepsilon_A \frac{\Phi_{i,j+1} - \Phi_{i,j} h1_{i-1}}{h2_j} \quad (\text{A.5})$$

$$-\iint_{Q_4} \vec{F} d\vec{A} = -\varepsilon_B \frac{\Phi_{i+1,j} - \Phi_{i,j} h2_j}{h1_i} - \varepsilon_B \frac{\Phi_{i,j+1} - \Phi_{i,j} h1_i}{h2_j} \quad (\text{A.6})$$

zwei ortsgleiche Punkte vier ortsgleiche Punkte

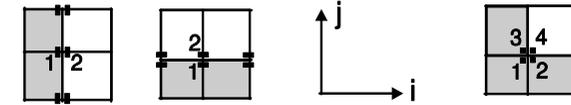


Abbildung A.3: Da die Gitterpunkte auf der Grenzfläche liegen, werden ortsgleiche Punkte eingeführt, um unstetige Größen wie Bandkanten und Ladungsdichten beschreiben zu können. Verläuft die Heterogrenzfläche entlang einer Koordinatenlinie, werden genau zwei ortsgleiche Punkte gesetzt, ansonsten vier. Sie werden in der angegebenen Weise durchnummeriert.

Ist Q die integrierte Ladung im Kontrollvolumen, so lautet Gl. (A.2) schließlich in diskretisierter Form:

$$\begin{aligned} & \left(-\varepsilon_A \frac{h1_{i-1} + h1_i}{2 \cdot h2_{j-1}} \right) \Phi_{i,j-1} + \left(-\varepsilon_A \frac{h2_{j-1} + h2_j}{2 \cdot h1_{i-1}} \right) \Phi_{i-1,j} + \\ & + \left(-\frac{\varepsilon_A h2_{j-1} + \varepsilon_B h2_j}{2 \cdot h1_i} \right) \Phi_{i+1,j} + \left(-\frac{\varepsilon_A h1_{i-1} + \varepsilon_B h1_i}{2 \cdot h2_j} \right) \Phi_{i,j+1} + \\ & + \left(\varepsilon_A \frac{h1_{i-1} + h1_i}{2 \cdot h2_{j-1}} + \varepsilon_A \frac{h2_{j-1} + h2_j}{2 \cdot h1_{i-1}} + \frac{\varepsilon_A h2_{j-1} + \varepsilon_B h2_j}{2 \cdot h1_i} + \frac{\varepsilon_A h1_{i-1} + \varepsilon_B h1_i}{2 \cdot h2_j} \right) \Phi_{i,j} \\ & = Q \end{aligned} \quad (\text{A.8})$$

Heterogrenzflächen

Wir haben das Gitter so gewählt, dass die Gitterpunkte auf den Grenzflächen an Materialübergängen liegen. Dies wirft die Frage auf, welcher Wert der Variable an diesen Punkten zuzuordnen ist, da wichtige physikalische Größen wie die Bandkanten und die klassische Teilchendichte dort einen Sprung machen. Wir lösen dieses Problem, indem wir ortsgleiche Punkte einführen, die den Wert im jeweiligen Quadranten (in 3D Oktanten) widerspiegeln. Um die Zahl der Punkte nicht unnötig aufzublähen, setzen wir an „geraden“ Grenzflächen Doppelpunkte und ansonsten vier (in 3D acht) ortsgleiche Punkte. Abb. A.3 illustriert die Verhältnisse je nach Lage des Heteroüberganges. In dieser Konvention sind auf dem Gitter in Abb. A.1 die Menge der Punkte {6, 16, 26, 45, 55, 65, 22, 32, 42, 11, 51, 20, 60, 29, 39, 49} Doppelpunkte und die Menge der Punkte {36, 35, 12, 52, 19, 59} Vierfachpunkte. Nummeriert man nun das Gitter einschließlich der ortsgleichen Punkte erneut durch, so erhält man 104 Gitterpunkte. Physikalische Größen werden nunmehr auf einem Vektor der Dimension 104 abgebildet. Für den oben betrachteten Differenzstern des Gitterpunktes mit $i = 5$ und $j = 4$ ergeben sich daraus die Punktnummern, wie sie in der rechten Grafik in Abb.

A.2 dargestellt sind. Für jeden Mehrfachpunkt werden außerdem zusätzliche Informationen wie Grenzflächenladungen gespeichert. In zwei Dimensionen gibt es dafür innerhalb eines Kontrollvolumens vier, in drei Dimensionen zwölf mögliche Grenzflächen. Die dazugehörigen Flächenladungen sind in Abb. A.2 mit σ_1 bis σ_4 bezeichnet. In unserem Fall sind σ_1 und σ_2 Null, da hier keine interne Grenzfläche vorliegt.

Nun kann man die integrierte Ladung Q in Gl. (A.7) schreiben, wenn ρ die Ladungsträgerdichte ist:

$$Q = \rho_{49} \frac{h_{1i-1} h_{2j-1}}{4} + \rho_{50} \frac{h_{1i} h_{2j-1}}{4} + \rho_{51} \frac{h_{1i-1} h_{2j}}{4} + \rho_{52} \frac{h_{1i} h_{2j}}{4} + \sigma_1 \frac{h_{2j-1}}{2} + \sigma_2 \frac{h_{1i-1}}{2} + \sigma_3 \frac{h_{2j}}{2} + \sigma_4 \frac{h_{1i}}{2}.$$

Die Bandkante E_C hat an Heteroübergängen eine Diskontinuität. Dabei gilt die Beziehung ($e = |e|$ Elementarladung):

$$E_C(\vec{r}) = E_{C0}(\vec{r}) - e\Phi(\vec{r}). \quad (\text{A.9})$$

Mikroskopisch wird der Sprung in der potenziellen Energie durch Dipole verursacht. Wir machen die Annahme, dass diese Diskontinuität voll in der Bandkante des Volumenhalbleiters $E_{C0}(\vec{r})$ enthalten ist. Dann ist das Potenzial $\Phi(\vec{r})$ an der Grenzfläche stetig. Aufgrund der Relation in Gl. (A.9) wird das Potenzial $\Phi(\vec{r})$ aber ebenso wie die Bandkante $E_C(\vec{r})$ auf das Gitter mit den Mehrfachpunkten abgebildet: $\Phi(\vec{r}) \rightarrow \Phi_1, \dots, \Phi_{104}$.

Bei der Stromgleichung lösen wir eine Gleichung für das Quasi-Ferminiveau $E_F(\vec{r})$. Auch hier nehmen wir zunächst Stetigkeit des Quasi-Ferminiveaus an Grenzflächen an. Die Abbildung auf das Gitter inklusive der Mehrfachpunkte, d.h. $E_F(\vec{r}) \rightarrow E_F^1, \dots, E_F^{104}$ erlaubt es uns jedoch, zusätzlich Modelle zu implementieren, die zu unstetigem Quasi-Ferminiveau an Heteroübergängen führen.

Insgesamt erhalten wir durch die Diskretisierung der Poisson- und Stromgleichung nun eine 104×104 -Matrix \mathbf{A} . Die Poissongleichung liefert ein lineares Gleichungssystem:

$$\mathbf{A}\vec{\Phi} = \vec{b}$$

mit $\vec{\Phi} = (\Phi_1, \dots, \Phi_{104})^T$.

Aufstellen der Poissonmatrix

Beim Aufstellen der Poissonmatrix ist darauf zu achten, dass die Matrix symmetrisch und positiv definit ist. Zunächst soll das Vorgehen anhand des Gitterpunktes ($i = 6, j = 2$) dargestellt werden. Dabei handelt es sich um einen Doppelpunkt. Die angrenzenden Gitterpunkte haben im 104-dimensionalen Gesamtvektor folgende Positionen, wobei beim Aufstellen des Differenzensterns bei ortsgleichen

Mehrfachpunkten immer der erste genommen wird:

Gitterpunkt	Position in Gesamtvektor (1..104)
$(i = 6, j = 2)$	21
$(i = 5, j = 2)$	20
$(i = 7, j = 2)$	23
$(i = 6, j = 1)$	6
$(i = 6, j = 3)$	37

Aus Gl. (A.7) ergibt sich ($i = 6, j = 2$):

$$\begin{aligned} & \left(-\varepsilon_A \frac{h_{1i-1} + h_{1i}}{2 \cdot h_{2j-1}} \right) \Phi_6 + \left(-\varepsilon_A \frac{h_{2j-1} + h_{2j}}{2 \cdot h_{1i-1}} \right) \Phi_{20} + \\ & + \left(-\frac{\varepsilon_A h_{2j-1} + \varepsilon_B h_{2j}}{2 \cdot h_{1i}} \right) \Phi_{23} + \left(-\frac{\varepsilon_A h_{1i-1} + \varepsilon_B h_{1i}}{2 \cdot h_{2j}} \right) \Phi_{37} + \\ & + \left(\varepsilon_A \frac{h_{1i-1} + h_{1i}}{2 \cdot h_{2j-1}} + \varepsilon_A \frac{h_{2j-1} + h_{2j}}{2 \cdot h_{1i-1}} + \frac{\varepsilon_A h_{2j-1} + \varepsilon_B h_{2j}}{2 \cdot h_{1i}} + \frac{\varepsilon_A h_{1i-1} + \varepsilon_B h_{1i}}{2 \cdot h_{2j}} \right) \Phi_{21} \\ & = Q \end{aligned} \quad (\text{A.10})$$

Zusätzlich haben wir die Bedingung:

$$\Phi_{22} - \Phi_{21} = 0 \quad (\text{A.11})$$

Um eine symmetrische Matrix zu erhalten, wird die Gleichung (A.11), multipliziert mit einem Faktor α , von Gl. (A.10) abgezogen und die so erhaltene Bedingung in Zeile 21 der Poissonmatrix geschrieben. Die Gleichung. (A.11) steht dann in der Zeile 22. Es ergibt sich in der Poissonmatrix folgende Struktur:

$$\left(\begin{array}{c|cccccccc} & 6 & \dots & 20 & 21 & 22 & 23 & \dots & 37 \\ \hline 21 & M_6 & \dots & M_{20} & M_{21} + \alpha & -\alpha & M_{23} & \dots & M_{37} \\ 22 & & & & -\alpha & \alpha & & & \\ \dots & & & & \dots & 0 & \dots & \dots & \dots \end{array} \right) = \begin{pmatrix} Q \\ 0 \\ 0 \\ 0 \\ \dots \end{pmatrix}$$

Dabei ist $M_6 = -\varepsilon_A \frac{h_{1i-1} + h_{1i}}{2 \cdot h_{2j-1}}$ der Koeffizient zu Φ_6 in Gl. (A.10) usw., der Faktor α sollte dieselbe Größenordnung wie M_{21} haben.

An einem Vierfachpunkt stellt sich die Situation folgendermaßen dar. Wir wollen hier kurz das Vorgehen anhand des schon mehrfach erwähnten Gitterpunktes ($i = 5, j = 4$) beschreiben. Die Nummerierung der Gitterpunkte kann aus der rechten Grafik in Abb. (A.2) entnommen werden. Da das Potenzial als stetig angenommen wird, gelten die Relationen:

$$\Phi_{50} - \Phi_{49} = 0 \quad (\text{A.12})$$

$$\Phi_{51} - \Phi_{49} = 0 \quad (\text{A.13})$$

$$\Phi_{52} - \Phi_{49} = 0 \quad (\text{A.14})$$

Weiterhin verwenden wir beim Aufstellen des Differenzensterns immer das Potenzial des ersten Punktes bei ortsgleichen Mehrfachpunkten. Dann ergibt sich aus Gl. (A.7):

$$\left(-\varepsilon_A \frac{h1_{i-1} + h1_i}{2 \cdot h2_{j-1}}\right) \Phi_{36} + \left(-\varepsilon_A \frac{h2_{j-1} + h2_j}{2 \cdot h1_{i-1}}\right) \Phi_{48} + \quad (\text{A.15})$$

$$+ \left(-\frac{\varepsilon_A h2_{j-1} + \varepsilon_B h2_j}{2 \cdot h1_i}\right) \Phi_{53} + \left(-\frac{\varepsilon_A h1_{i-1} + \varepsilon_B h1_i}{2 \cdot h2_j}\right) \Phi_{67} +$$

$$+ \left(\varepsilon_A \frac{h1_{i-1} + h1_i}{2 \cdot h2_{j-1}} + \varepsilon_A \frac{h2_{j-1} + h2_j}{2 \cdot h1_{i-1}} + \frac{\varepsilon_A h2_{j-1} + \varepsilon_B h2_j}{2 \cdot h1_i} + \frac{\varepsilon_A h1_{i-1} + \varepsilon_B h1_i}{2 \cdot h2_j}\right) \Phi_{49}$$

$$= Q \quad (\text{A.16})$$

Der Gitterpunkt ($i = 5, j = 4$) besteht aus vier ortsgleichen Punkten, so dass zum Aufstellen der Poissonmatrix vier Bedingungen, nämlich Gl. (A.12) und Gl. (A.15) benötigt werden. Dabei ist zu beachten, dass die Matrix symmetrisch und positiv definit sein sollte. Man erreicht dies, indem man die Gleichung (A.12), multipliziert mit einem Faktor α , von Gl. (A.15) abzieht und die so erhaltene Bedingung in Zeile 49 schreibt. Die Gleichungen (A.13) und (A.14) werden zunächst umgeschrieben zu:

$$\Phi_{51} - \Phi_{50} = 0 \quad (\text{A.17})$$

$$\Phi_{52} - \Phi_{50} = 0 \quad (\text{A.18})$$

Gleichung (A.17) wird, mit dem Faktor α multipliziert, direkt in Zeile 51 geschrieben, Gleichung (A.18) in Zeile 52. Um Symmetrie zu gewährleisten, werden die Gleichungen (A.17) und (A.18) von der Gleichung (A.12) abgezogen und die so erhaltene Relation in Zeile 50 geschrieben. Man erhält dann für die Poissonmatrix folgende Struktur:

$$\left(\begin{array}{c|cccccccc} & 36 & \dots & 48 & 49 & 50 & 51 & 52 & 53 & \dots & 67 \\ \hline 49 & M_{36} & \dots & M_{48} & M_{49} + \alpha & -\alpha & 0 & 0 & M_{53} & \dots & M_{67} \\ 50 & & & & -\alpha & 3\alpha & -\alpha & -\alpha & & & \\ 51 & & & & 0 & -\alpha & \alpha & 0 & & & \\ 52 & & & & 0 & -\alpha & 0 & \alpha & & & \\ \dots & & & & \dots & 0 & 0 & 0 & \dots & \dots & \dots \end{array} \right) = \begin{pmatrix} Q \\ 0 \\ 0 \\ 0 \\ \dots \end{pmatrix}$$

Dabei ist $M_{36} = -\varepsilon_A \frac{h1_{i-1} + h1_i}{2 \cdot h2_{j-1}}$ der Koeffizient zu Φ_{36} in Gl. (A.15) usw., der Faktor α sollte dieselbe Größenordnung wie M_{49} haben.

Randbedingungen

Randbedingungen müssen am Rand des Simulationsgebietes berücksichtigt werden. Zusätzlich gibt man durch Kontaktregionen Randbedingungen vor. Zunächst soll die Behandlung von Gitterpunkten am Simulationsrand betrachtet

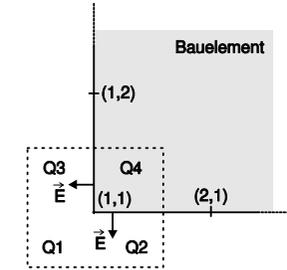


Abbildung A.4: Kontrollvolumen für den Randpunkt ($i = 1, j = 1$). Nur der Quadrant Q_4 liegt innerhalb der Bauelementregion.

werden. Dort wird typischerweise verschwindendes elektrisches Feld gefordert. Es handelt sich demnach um eine Randbedingung vom Typ von Neumann mit Feld $\vec{E} = 0$. Wählen wir dazu den Punkt ($i = 1, j = 1$) als Beispiel. In Abb. A.4 ist die Situation dargestellt. Man sieht, dass sich nur der Quadrant Q_4 innerhalb der Bauelementregion befindet. Das Volumenintegral über das Kontrollvolumen reduziert sich demnach auf das Integral über den Quadranten Q_4 . Dieses Integral wird wieder analog zu oben in ein Flächenintegral verwandelt. Von den oben angegebenen Integralen ist jetzt nur das Integral Gl. (A.6) relevant:

$$- \iint_{Q_4} \vec{F} d\vec{A} = -\varepsilon_B \frac{\Phi_{i+1,j} - \Phi_{i,j}}{h1_i} \frac{h2_j}{2} - \varepsilon_B \frac{\Phi_{i,j+1} - \Phi_{i,j}}{h2_j} \frac{h1_i}{2} \quad (\text{A.19})$$

Zusätzlich muss das Oberflächenintegral durch die Begrenzungsflächen berücksichtigt werden:

$$\iint_{Rand} \vec{F} d\vec{A} = E \frac{h1_i}{2} + E \frac{h2_j}{2} \quad (\text{A.20})$$

Daraus erhält man nun durch Addition von Gl. (A.19) und Gl. (A.20) den Differenzenstern für den Punkt ($i = 1, j = 1$):

$$\left(\varepsilon_B \frac{h2_j}{2 \cdot h1_i} + \varepsilon_B \frac{h1_i}{2 \cdot h2_j}\right) \Phi_{i,j} + \left(-\varepsilon_B \frac{h2_j}{2 \cdot h1_i}\right) \Phi_{i+1,j} + \left(-\varepsilon_B \frac{h1_i}{2 \cdot h2_j}\right) \Phi_{i,j+1}$$

$$= E \frac{h1_i}{2} + E \frac{h2_j}{2} + Q$$

mit der integrierten Ladung Q im Quadrant Q_4 .

Kontakte

Randbedingungen können weiterhin durch Kontakte vorgegeben werden. Dazu muss man so genannte Kontaktregionen definieren. In Abb. A.1 sind diese dunkel unterlegt eingezeichnet. An den Grenzflächen zwischen Kontaktmaterial und

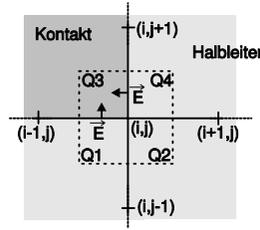


Abbildung A.5: Kontrollvolumen für den Gitterpunkt $(i = 2, j = 2)$. Quadrant $Q3$ liegt in der Kontaktregion und wird bei der Integration über das finite Volumen nicht berücksichtigt. Stattdessen wird das elektrische Feld als Randbedingung gesetzt.

dem übrigen Bauelement werden nun Randbedingungen gesetzt. In Abb. A.5 wird dies anhand des Gitterpunkts $(i = 2, j = 2)$ aus Abb. A.1 illustriert. Dabei legen wir von Neumann Randbedingungen zugrunde, d.h. wir geben ein elektrisches Feld \vec{E} vor. Das Kontrollvolumen umfasst nun die Quadranten Q1, Q2 und Q4, d.h. die Integration erstreckt sich nur über diese drei, Quadrant Q3 wird nicht berücksichtigt. Bei der Umwandlung des Volumenintegrals in ein Oberflächenintegral muss aber der Fluss durch die Grenzflächen miteinbezogen werden. Konkret erhält man daher folgende vier Beiträge:

$$-\iint_{Q1} \vec{F} d\vec{A} = +\varepsilon_A \frac{\Phi_{i,j} - \Phi_{i-1,j}}{h1_{i-1}} \frac{h2_{j-1}}{2} + \varepsilon_A \frac{\Phi_{i,j} - \Phi_{i,j-1}}{h2_{j-1}} \frac{h1_{i-1}}{2} \quad (\text{A.21})$$

$$-\iint_{Q2} \vec{F} d\vec{A} = -\varepsilon_A \frac{\Phi_{i+1,j} - \Phi_{i,j}}{h1_i} h2_{j-1} + \varepsilon_A \frac{\Phi_{i,j} - \Phi_{i,j-1}}{h2_{j-1}} \frac{h1_i}{2} \quad (\text{A.22})$$

$$-\iint_{Q4} \vec{F} d\vec{A} = -\varepsilon_B \frac{\Phi_{i+1,j} - \Phi_{i,j}}{h1_i} \frac{h2_j}{2} - \varepsilon_B \frac{\Phi_{i,j+1} - \Phi_{i,j}}{h2_j} \frac{h1_i}{2} \quad (\text{A.23})$$

$$-\iint_{Rand} \vec{F} d\vec{A} = -E \frac{h1_{i-1}}{2} - E \frac{h2_j}{2} \quad (\text{A.24})$$

Durch Summation der vier Beiträge (A.21)-(A.24) kann wieder analog zu oben der Differenzenstern assembliert werden. Die Punkte innerhalb der Kontaktregion - das sind in Abb. (A.1) die Punkte $\{i = 1, j = 3, 4, 5\}$ für den Kontakt 1 - werden nicht mitsimuliert. Wir weisen ihnen in der Poissonmatrix über eine Dirichlet-Bedingung einen festen Wert zu, der aber nicht von Bedeutung ist. Im Prinzip könnte man diese Punkte aus dem Poisson-Gleichungssystem eliminieren.

Dirichlet-Randbedingungen

Zur Berechnung der eingebauten Spannung wird bei ohmschen Kontakten als Randbedingung verschwindendes elektrisches Feld angenommen. Bei angelegter Spannung hat man dann eine Randbedingung vom Typ Dirichlet (vgl. Kapitel 4.1), d.h.

$$\Phi = const. \quad (\text{A.25})$$

Technisch wird dies realisiert, indem man für alle Gitterpunkte in der Kontaktregion und für alle Punkte auf der Kontaktgrenzfläche - das sind in Abb. (A.1) die Punkte $\{i = 1, 2, j = 2, 3, 4, 5, 6\}$ für den Kontakt 1 - die Bedingung in Gl. (A.25) direkt in die Matrix schreibt.

Stromgleichung

Bei der Stromgleichung gehen wir analog vor. Hier gilt für den Stromterm (für Elektronenstrom) aus Gl. (A.1) (vgl. Kapitel 4.2):

$$\vec{F} = \vec{j}_n = \mu n \vec{\nabla} E_{Fn}$$

Bei der Auswertung der Fluss-Integrale benötigen wir dabei die Dichte n auf den Begrenzungsflächen des Kontrollvolumens, also zwischen zwei Gitterpunkten. Wir bestimmen sie durch Bildung des Mittelwertes. Die zu Gl. (A.3) analogen Gleichungen lauten nun für den Gitterpunkt $(i = 5, j = 4)$ (siehe auch Abb. A.2):

$$-\iint_{Q1} \vec{F} d\vec{A} = +\mu \left(\frac{n_{48} + n_{49}}{2} \right) \frac{E_{i,j}^{Fn} - E_{i-1,j}^{Fn}}{h1_{i-1}} \frac{h2_{j-1}}{2} + \quad (\text{A.26})$$

$$\mu \left(\frac{n_{36} + n_{49}}{2} \right) \frac{E_{i,j}^{Fn} - E_{i,j-1}^{Fn}}{h2_{j-1}} \frac{h1_{i-1}}{2}$$

$$-\iint_{Q2} \vec{F} d\vec{A} = -\mu \left(\frac{n_{50} + n_{53}}{2} \right) \frac{E_{i+1,j}^{Fn} - E_{i,j}^{Fn}}{h1_i} h2_{j-1} + \quad (\text{A.27})$$

$$\mu \left(\frac{n_{36} + n_{50}}{2} \right) \frac{E_{i,j}^{Fn} - E_{i,j-1}^{Fn}}{h2_{j-1}} \frac{h1_i}{2}$$

$$-\iint_{Q3} \vec{F} d\vec{A} = +\mu \left(\frac{n_{48} + n_{51}}{2} \right) \frac{E_{i,j}^{Fn} - E_{i-1,j}^{Fn}}{h1_{i-1}} \frac{h2_j}{2} - \quad (\text{A.28})$$

$$\mu \left(\frac{n_{67} + n_{51}}{2} \right) \frac{E_{i,j+1}^{Fn} - E_{i,j}^{Fn}}{h2_j} \frac{h1_{i-1}}{2}$$

$$-\iint_{Q4} \vec{F} d\vec{A} = -\mu \left(\frac{n_{52} + n_{55}}{2} \right) \frac{E_{i+1,j}^{Fn} - E_{i,j}^{Fn}}{h1_i} \frac{h2_j}{2} - \quad (\text{A.29})$$

$$\mu \left(\frac{n_{52} + n_{68}}{2} \right) \frac{E_{i,j+1}^{Fn} - E_{i,j}^{Fn}}{h2_j} \frac{h1_i}{2}$$

Die Beweglichkeit wird entsprechend Anhang H ermittelt. Die Dichten n werden mit dem gegebenen Quasi-Ferminiveau E_{F_n} berechnet und als Parameter betrachtet. Insgesamt erhält man ein lineares Gleichungssystem:

$$\mathbf{A}\vec{E}_{F_n} = \vec{b}$$

mit $\vec{E}_{F_n} = (E_{F_n,1}, \dots, E_{F_n,104})^T$ und der 104×104 -Matrix \mathbf{A} . Als Lösung ergibt sich ein neues Quasi-Ferminiveau E_{F_n} . Dadurch ändert sich aber wieder Dichte und damit die Strommatrix. Zur vollständigen Lösung des Stromproblems muss daher diese Rechnung mehrfach iteriert werden, bis Selbstkonsistenz erreicht ist.

Anhang B

Diskretisierung der Verspannungsgleichung

Die grundlegenden Aspekte bei der Berechnung der elastischen Verspannungen wurden in Kapitel 4.3 dargestellt. In diesem Abschnitt soll näher auf die Diskretisierung und die technischen Gesichtspunkte beim Aufsetzen der Matrix eingegangen werden. Dabei wird wir für einen bestimmten Gitterpunkt Schritt für Schritt der Differenzenstern assembliert.

Allgemeines

Die elastomechanische Verspannungsgleichung lautet ohne äußere Kräfte (vgl. Kapitel 4.3):

$$\nabla_{\vec{x}} \cdot \hat{\sigma} = \vec{0} \quad (\text{B.1})$$

Dabei stellt $\hat{\sigma}$ einen Tensor zweiter Stufe (Matrix) dar. Es handelt sich hier demnach um eine Gleichung mit drei Komponenten. Die Definition der Vektordivergenz $\nabla_{\vec{x}}$ soll kurz in Erinnerung gerufen werden:

$$(\nabla_{\vec{x}} \cdot \hat{\sigma})_i = \sum_j \frac{\partial}{\partial x_j} \sigma_{ij} \quad (i, j = 1..3) \quad (\text{B.2})$$

Mit Hilfe des Gaußschen Satzes lässt sich Gl. (B.1) in integraler Form schreiben mit dem Normalenvektor \vec{n} :

$$\int_V \nabla_{\vec{x}} \cdot \hat{\sigma} dV = \int_A \vec{n} \cdot \hat{\sigma} dA = 0 \quad (\text{B.3})$$

Im Folgenden wollen wir uns mit der Diskretisierung der Gl. (B.3) im Rahmen der Methode der finiten Volumen genauer befassen.

Differenzenstern

Die Gleichung (B.1) umfasst drei Komponenten. Führt man für jeden Gitterpunkt die Integration über das jeweilige Kontrollvolumen (vgl. Kapitel 4.4) durch, so erhält man für N Gitterpunkte ein lineares Gleichungssystem der Dimension $3N \times 3N$. Die Lösung des Systems wird durch die Auslenkung $\vec{u}_{i,j,k} = (u_{i,j,k}^1, u_{i,j,k}^2, u_{i,j,k}^3)$ eines Teilchens aus der Ruhelage für jeden Gitterpunkt (i, j, k) vollständig beschrieben, wobei der obere Index jeweils die Auslenkung in x_1 -, x_2 - und x_3 -Richtung bedeutet. Die Gitterpunkte (i, j, k) werden von 1 bis N durchnummeriert und jedem Punkt die drei Komponenten der Auslenkung zugewiesen. Der $3N$ -dimensionalen Lösungsvektor hat dann folgende Gestalt:

$$\vec{u} = (u_1^1, u_1^2, u_1^3, u_2^1, u_2^2, u_2^3, \dots, u_N^1, u_N^2, u_N^3) \quad (\text{B.4})$$

Für die Komponenten des Verzerrungstensors gilt:

$$\epsilon_{ij} = \frac{1}{2} \left(\frac{\partial u^i}{\partial x_j} + \frac{\partial u^j}{\partial x_i} \right) \quad (\text{B.5})$$

Wie wir weiter unten sehen werden, gehen bei der Bildung der finiten Differenzen am Gitterpunkt (i, j, k) alle benachbarten Gitterpunkte entlang der Koordinatenachsen und der Diagonalen ein. Zählt man den Punkt (i, j, k) mit, so umfasst der Differenzenstern U (1..57) insgesamt 27 Punkte mit jeweils drei Komponenten bei folgender Zuordnung:

$u_{i,j,k-1}^1$	$u_{i,j,k-1}^2$	$u_{i,j,k-1}^3$	$u_{i,j-1,k}^1$	$u_{i,j-1,k}^2$	$u_{i,j-1,k}^3$	$u_{i-1,j,k}^1$	$u_{i-1,j,k}^2$	$u_{i-1,j,k}^3$		
1	2	3	4	5	6	7	8	9		
$u_{i,j,k}^1$	$u_{i,j,k}^2$	$u_{i,j,k}^3$	$u_{i+1,j,k}^1$	$u_{i+1,j,k}^2$	$u_{i+1,j,k}^3$	$u_{i,j+1,k}^1$	$u_{i,j+1,k}^2$	$u_{i,j+1,k}^3$		
10	11	12	13	14	15	16	17	18		
$u_{i,j,k+1}^1$	$u_{i,j,k+1}^2$	$u_{i,j,k+1}^3$	$u_{i-1,j-1,k}^1$	$u_{i-1,j-1,k}^2$	$u_{i-1,j-1,k}^3$					
19	20	21	22	23	24					
$u_{i+1,j-1,k}^1$	$u_{i+1,j-1,k}^2$	$u_{i+1,j-1,k}^3$	$u_{i-1,j+1,k}^1$	$u_{i-1,j+1,k}^2$	$u_{i-1,j+1,k}^3$					
25	26	27	28	29	30					
$u_{i+1,j+1,k}^1$	$u_{i+1,j+1,k}^2$	$u_{i+1,j+1,k}^3$	$u_{i-1,j,k-1}^1$	$u_{i-1,j,k-1}^2$	$u_{i-1,j,k-1}^3$					
31	32	33	34	35	36					
$u_{i+1,j,k-1}^1$	$u_{i+1,j,k-1}^2$	$u_{i+1,j,k-1}^3$	$u_{i-1,j,k+1}^1$	$u_{i-1,j,k+1}^2$	$u_{i-1,j,k+1}^3$					
37	38	39	40	41	42					
$u_{i+1,j,k+1}^1$	$u_{i+1,j,k+1}^2$	$u_{i+1,j,k+1}^3$	$u_{i,j-1,k-1}^1$	$u_{i,j-1,k-1}^2$	$u_{i,j-1,k-1}^3$					
43	44	45	46	47	48					
$u_{i,j+1,k-1}^1$	$u_{i,j+1,k-1}^2$	$u_{i,j+1,k-1}^3$	$u_{i,j-1,k+1}^1$	$u_{i,j-1,k+1}^2$	$u_{i,j-1,k+1}^3$					
49	50	51	52	53	54					
$u_{i,j+1,k+1}^1$	$u_{i,j+1,k+1}^2$	$u_{i,j+1,k+1}^3$								
55	56	57								

(B.6)

In Abb. B.1 ist das Kontrollvolumen (vgl. Kapitel 4.3) für den Gitterpunkt (i, j, k) dargestellt. Dieses kann in acht Oktanten unterteilt werden. Für die Nummerierung gilt folgende Konvention:

- Oktant 1 $(i-, j-, k-)$
- Oktant 2 $(i+, j-, k-)$
- Oktant 3 $(i-, j+, k-)$
- Oktant 4 $(i+, j+, k-)$
- Oktant 5 $(i-, j-, k+)$
- Oktant 6 $(i+, j-, k+)$
- Oktant 7 $(i-, j+, k+)$
- Oktant 8 $(i+, j+, k+)$

Der grau unterlegte Oktant in Abb. B.1 ist demnach Oktant 6. Mit diesem wollen wir uns im Folgenden näher befassen.

Integration

Die Integration wird gemäß Gl. (B.3) über die Begrenzungsfläche des Kontrollvolumens durchgeführt. Dieses Integral setzt sich aus den acht Integralen über die Oberflächen der entsprechenden acht Oktanten zusammen:

$$\int_A \vec{n} \cdot \hat{\sigma} dA = \sum_{m=1}^8 \int_{\text{Oktant } m} \vec{n} \cdot \hat{\sigma} dA = 0 \quad (\text{B.7})$$

Die Begrenzungsflächen für den Oktanten 6 sind in Abb. B.1 mit A_{n1} , A_{n2} und A_{n3} bezeichnet. Die zugehörigen Normalenvektoren sind $\vec{n}_1 = (1, 0, 0)^T$, $\vec{n}_2 = -(0, 1, 0)^T$ und $\vec{n}_3 = (0, 0, 1)^T$. Für die Flächeninhalte gilt:

$$A_{n1} = \frac{h2_{j-1}h3_k}{4}, \quad A_{n2} = \frac{h1_i h3_k}{4}, \quad A_{n3} = \frac{h1_i h2_j}{4}$$

Dabei bedeutet $h1_i$ den Abstand zwischen den Punkten i und $i+1$ auf der x -Achse, $h2_j$ den Abstand zwischen den Punkten j und $j+1$ auf der y -Achse und $h3_k$ den Abstand zwischen den Punkten k und $k+1$ auf der z -Achse. Mit diesen Bezeichnungen und der Relation

$$(\vec{n} \cdot \hat{\sigma})_k = \sigma_{jk} n_j$$

gilt nun:

$$\int_{\text{Oktant 6}} \vec{n} \cdot \hat{\sigma} dA = \begin{pmatrix} \sigma_{11} \\ \sigma_{12} \\ \sigma_{13} \end{pmatrix} A_{n1} - \begin{pmatrix} \sigma_{21} \\ \sigma_{22} \\ \sigma_{23} \end{pmatrix} A_{n2} + \begin{pmatrix} \sigma_{31} \\ \sigma_{32} \\ \sigma_{33} \end{pmatrix} A_{n3} \quad (\text{B.8})$$

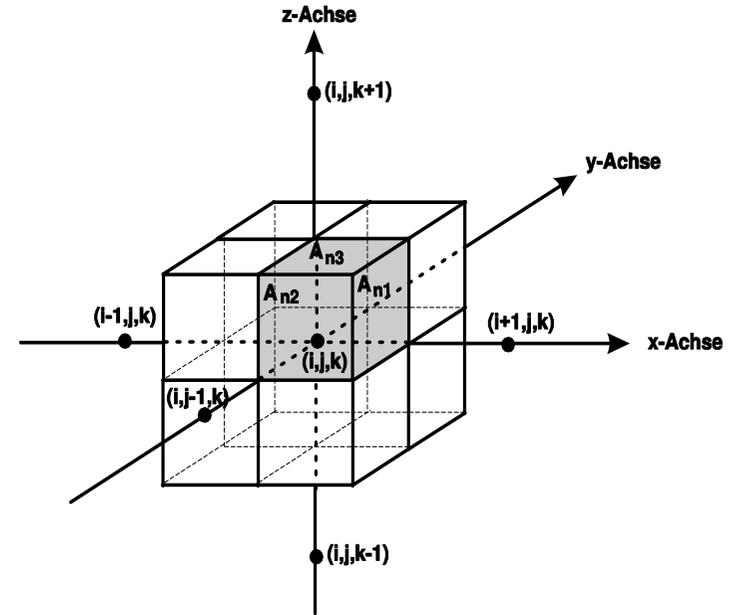


Abbildung B.1: Kontrollvolumen des Gitterpunktes (i, j, k) . Die Begrenzungsflächen sind die Mittelsenkrechten zu den Verbindungslinien zwischen zwei benachbarten Gitterpunkten. Das Kontrollvolumen kann in acht Teilvolumina (Oktanten) unterteilt werden. Jedem Oktanten ist eindeutig eine Materialsorte zuzuordnen. Oktant 6 ist grau unterlegt dargestellt. Bei der Diskretisierung muss der Fluss durch die Flächen A_{n1} , A_{n2} und A_{n3} berechnet werden.

Für eine kompakte Darstellung des Spannungs- und Verzerrungstensors gehen wir über zur Voigt-Notation (11 \rightarrow 1, 22 \rightarrow 2, 33 \rightarrow 3, 23 \rightarrow 4, 13 \rightarrow 5, 12 \rightarrow 6). Dabei macht man folgende Transformation:

$$\begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \\ \epsilon_4 \\ \epsilon_5 \\ \epsilon_6 \end{bmatrix} = \begin{bmatrix} \epsilon_{11} \\ \epsilon_{22} \\ \epsilon_{33} \\ 2\epsilon_{23} \\ 2\epsilon_{13} \\ 2\epsilon_{12} \end{bmatrix}, \quad \begin{bmatrix} \sigma_1 \\ \sigma_2 \\ \sigma_3 \\ \sigma_4 \\ \sigma_5 \\ \sigma_6 \end{bmatrix} = \begin{bmatrix} \sigma_{11} \\ \sigma_{22} \\ \sigma_{33} \\ \sigma_{23} \\ \sigma_{13} \\ \sigma_{12} \end{bmatrix} \quad (\text{B.9})$$

Das Hooksche Gesetz lässt sich nun in Matrix-Vektor-Schreibweise darstellen:

$$\sigma_i = E_{ij}\epsilon_j \quad (i, j = 1 \dots 6) \quad (\text{B.10})$$

Jedem Oktanten kann eindeutig eine Materialsorte zugewiesen werden. Für das gegebene Material ist zunächst die Elastizitätsmatrix \hat{E} im Kristallsystem gegeben. Für den Oktant 6 lautet sie (hier für Zinkblende-Struktur):

$$\hat{E}_{\text{Zinkblende}}^{\text{Oktant 6}} = \begin{pmatrix} c_{11} & c_{12} & c_{12} & & & \\ c_{12} & c_{11} & c_{12} & & & \\ c_{12} & c_{12} & c_{11} & & & \\ & & & c_{44} & & \\ & & & & c_{44} & \\ & & & & & c_{44} \end{pmatrix}$$

Stimmen die Kristallachsen nicht mit den Simulationsachsen (x, y, z) überein, muss die Elastizitätsmatrix in das Simulationssystem rücktransformiert werden. Dazu muss sie zunächst auf die Form des vierstufigen Tensors ${}^{(4)}\mathbf{E}$ gebracht und mit der Rotationsmatrix R_{ij} (vgl. Kapitel 1.2) transformiert werden gemäß:

$$E_{ijkl}^{(x,y,z)} = R_{im}R_{jn}R_{ko}R_{lp}E_{mnop}^{\text{Kristall}}$$

Anschließend wird sie wieder in die Voigt-Darstellung zurückgeführt. Gleichung (B.8) lautet nun:

$$\begin{aligned} \int_{\text{Oktant 6}} \vec{n} \cdot \hat{\sigma} dA &= \begin{pmatrix} \sigma_1 \\ \sigma_6 \\ \sigma_5 \end{pmatrix} A_{n1} - \begin{pmatrix} \sigma_6 \\ \sigma_2 \\ \sigma_4 \end{pmatrix} A_{n2} + \begin{pmatrix} \sigma_5 \\ \sigma_4 \\ \sigma_3 \end{pmatrix} A_{n3} = \quad (\text{B.11}) \\ &= \begin{pmatrix} E_{11}\epsilon_1 + E_{12}\epsilon_2 + E_{13}\epsilon_3 + E_{14}\epsilon_4 + E_{15}\epsilon_5 + E_{16}\epsilon_6 \\ E_{61}\epsilon_1 + E_{62}\epsilon_2 + E_{63}\epsilon_3 + E_{64}\epsilon_4 + E_{65}\epsilon_5 + E_{66}\epsilon_6 \\ E_{51}\epsilon_1 + E_{52}\epsilon_2 + E_{53}\epsilon_3 + E_{54}\epsilon_4 + E_{55}\epsilon_5 + E_{56}\epsilon_6 \end{pmatrix} A_{n1} - \\ &\quad - \begin{pmatrix} E_{21}\epsilon_1 + E_{22}\epsilon_2 + E_{23}\epsilon_3 + E_{24}\epsilon_4 + E_{25}\epsilon_5 + E_{26}\epsilon_6 \\ E_{41}\epsilon_1 + E_{42}\epsilon_2 + E_{43}\epsilon_3 + E_{44}\epsilon_4 + E_{45}\epsilon_5 + E_{46}\epsilon_6 \end{pmatrix} A_{n2} + \\ &\quad + \begin{pmatrix} E_{31}\epsilon_1 + E_{32}\epsilon_2 + E_{33}\epsilon_3 + E_{34}\epsilon_4 + E_{35}\epsilon_5 + E_{36}\epsilon_6 \\ E_{41}\epsilon_1 + E_{42}\epsilon_2 + E_{43}\epsilon_3 + E_{44}\epsilon_4 + E_{45}\epsilon_5 + E_{46}\epsilon_6 \\ E_{51}\epsilon_1 + E_{52}\epsilon_2 + E_{53}\epsilon_3 + E_{54}\epsilon_4 + E_{55}\epsilon_5 + E_{56}\epsilon_6 \end{pmatrix} A_{n3} \end{aligned}$$

Diskretisierung

Die Diskretisierung erfolgt flusserhaltend, dann erhält man eine symmetrische Matrix. Im Einzelnen gilt für die Differenziale für die Teilfläche A_{n1} :

$$\begin{aligned} \epsilon_1 &= \epsilon_{11} = \frac{\partial u^1}{\partial x} = \frac{u_{i+1,j,k}^1 - u_{i,j,k}^1}{h1_i} \\ \epsilon_2 &= \epsilon_{22} = \frac{\partial u^2}{\partial y} = \frac{1}{2} \left(\frac{u_{i,j,k}^2 - u_{i,j-1,k}^2}{h2_{j-1}} + \frac{u_{i+1,j,k}^2 - u_{i+1,j-1,k}^2}{h2_{j-1}} \right) \\ \epsilon_3 &= \epsilon_{33} = \frac{\partial u^3}{\partial z} = \frac{1}{2} \left(\frac{u_{i,j,k+1}^3 - u_{i,j,k}^3}{h3_k} + \frac{u_{i+1,j,k+1}^3 - u_{i+1,j,k}^3}{h3_k} \right) \\ \epsilon_4 &= 2\epsilon_{23} = 2 \cdot \frac{1}{2} \left(\frac{\partial u^2}{\partial z} + \frac{\partial u^3}{\partial y} \right) = 2 \cdot \frac{1}{2} \left[\frac{1}{2} \left(\frac{u_{i,j,k+1}^2 - u_{i,j,k}^2}{h3_k} + \frac{u_{i+1,j,k+1}^2 - u_{i+1,j,k}^2}{h3_k} \right) + \right. \\ &\quad \left. \frac{1}{2} \left(\frac{u_{i,j,k}^3 - u_{i,j-1,k}^3}{h2_{j-1}} + \frac{u_{i+1,j,k}^3 - u_{i+1,j-1,k}^3}{h2_{j-1}} \right) \right] \\ \epsilon_5 &= 2\epsilon_{13} = 2 \cdot \frac{1}{2} \left(\frac{\partial u^1}{\partial z} + \frac{\partial u^3}{\partial x} \right) = 2 \cdot \frac{1}{2} \left[\frac{1}{2} \left(\frac{u_{i,j,k+1}^1 - u_{i,j,k}^1}{h3_k} + \frac{u_{i+1,j,k+1}^1 - u_{i+1,j,k}^1}{h3_k} \right) + \right. \\ &\quad \left. \frac{u_{i+1,j,k}^3 - u_{i,j,k}^3}{h1_i} \right] \\ \epsilon_6 &= 2\epsilon_{12} = 2 \cdot \frac{1}{2} \left(\frac{\partial u^1}{\partial y} + \frac{\partial u^2}{\partial x} \right) = 2 \cdot \frac{1}{2} \left[\frac{1}{2} \left(\frac{u_{i,j,k}^1 - u_{i,j-1,k}^1}{h2_{j-1}} + \frac{u_{i+1,j,k}^1 - u_{i+1,j-1,k}^1}{h2_{j-1}} \right) + \right. \\ &\quad \left. \frac{u_{i+1,j,k}^2 - u_{i,j,k}^2}{h1_i} \right] \end{aligned}$$

Für die obigen Differenziale wird der jeweilige Differenzenstern U (1...57) gebildet:

$$\begin{aligned}
\epsilon_1 &\rightarrow U_{\epsilon_1}(10) = \frac{-1}{h1_i}, U_{\epsilon_1}(13) = \frac{1}{h1_i} \\
\epsilon_2 &\rightarrow U_{\epsilon_2}(5) = \frac{-1}{2 \cdot h2_{j-1}}, U_{\epsilon_2}(11) = \frac{1}{2 \cdot h2_{j-1}}, U_{\epsilon_2}(14) = \frac{1}{2 \cdot h2_{j-1}}, U_{\epsilon_2}(26) = \frac{-1}{2 \cdot h2_{j-1}} \\
\epsilon_3 &\rightarrow U_{\epsilon_3}(12) = \frac{-1}{2 \cdot h3_k}, U_{\epsilon_3}(15) = \frac{-1}{2 \cdot h3_k}, U_{\epsilon_3}(21) = \frac{1}{2 \cdot h3_k}, U_{\epsilon_3}(45) = \frac{1}{2 \cdot h3_k} \\
\epsilon_4 &\rightarrow U_{\epsilon_4}(6) = \frac{-1}{2 \cdot h2_{j-1}}, U_{\epsilon_4}(11) = \frac{-1}{2 \cdot h3_k}, U_{\epsilon_4}(12) = \frac{1}{2 \cdot h2_{j-1}}, U_{\epsilon_4}(14) = \frac{-1}{2 \cdot h3_k} \\
&\quad U_{\epsilon_4}(15) = \frac{1}{2 \cdot h2_{j-1}}, U_{\epsilon_4}(19) = \frac{1}{2 \cdot h3_k}, U_{\epsilon_4}(27) = \frac{1}{2 \cdot h2_{j-1}}, U_{\epsilon_4}(44) = \frac{1}{2 \cdot h3_k} \\
\epsilon_5 &\rightarrow U_{\epsilon_5}(10) = \frac{-1}{2 \cdot h3_k}, U_{\epsilon_5}(12) = \frac{-1}{h1_i}, U_{\epsilon_5}(13) = \frac{-1}{2 \cdot h3_k}, U_{\epsilon_5}(15) = \frac{1}{h1_i} \\
&\quad U_{\epsilon_5}(19) = \frac{1}{2 \cdot h3_k}, U_{\epsilon_5}(43) = \frac{1}{2 \cdot h3_k} \\
\epsilon_6 &\rightarrow U_{\epsilon_6}(4) = \frac{-1}{2 \cdot h2_{j-1}}, U_{\epsilon_6}(10) = \frac{1}{2 \cdot h2_{j-1}}, U_{\epsilon_6}(11) = \frac{-1}{h1_i}, U_{\epsilon_6}(13) = \frac{1}{2 \cdot h2_{j-1}} \\
&\quad U_{\epsilon_6}(14) = \frac{1}{h1_i}, U_{\epsilon_6}(25) = \frac{-1}{2 \cdot h2_{j-1}}
\end{aligned}$$

Diese Differenzensterne werden nun in Gl. (B.11) eingesetzt. Für den Fluss durch Grenzfläche A_{n1} ergibt sich:

$$\begin{aligned}
&\begin{pmatrix} E_{11}\epsilon_1 + E_{12}\epsilon_2 + E_{13}\epsilon_3 + 2E_{14}\epsilon_4 + 2E_{15}\epsilon_5 + 2E_{16}\epsilon_6 \\ E_{61}\epsilon_1 + E_{62}\epsilon_2 + E_{63}\epsilon_3 + 2E_{64}\epsilon_4 + 2E_{65}\epsilon_5 + 2E_{66}\epsilon_6 \\ E_{51}\epsilon_1 + E_{52}\epsilon_2 + E_{53}\epsilon_3 + 2E_{54}\epsilon_4 + 2E_{55}\epsilon_5 + 2E_{56}\epsilon_6 \end{pmatrix} A_{n1} \\
&\rightarrow \begin{pmatrix} E_{11}U_{\epsilon_1} + E_{12}U_{\epsilon_2} + E_{13}U_{\epsilon_3} + 2E_{14}U_{\epsilon_4} + 2E_{15}U_{\epsilon_5} + 2E_{16}U_{\epsilon_6} \\ E_{61}U_{\epsilon_1} + E_{62}U_{\epsilon_2} + E_{63}U_{\epsilon_3} + 2E_{64}U_{\epsilon_4} + 2E_{65}U_{\epsilon_5} + 2E_{66}U_{\epsilon_6} \\ E_{51}U_{\epsilon_1} + E_{52}U_{\epsilon_2} + E_{53}U_{\epsilon_3} + 2E_{54}U_{\epsilon_4} + 2E_{55}U_{\epsilon_5} + 2E_{56}U_{\epsilon_6} \end{pmatrix} A_{n1}
\end{aligned}$$

Die Differenziale durch die Teilfläche A_{n2} lauten:

$$\begin{aligned}
\epsilon_1 &= \epsilon_{11} = \frac{\partial u^1}{\partial x} = \frac{1}{2} \left(\frac{u_{i+1,j,k}^1 - u_{i,j,k}^1}{h1_i} + \frac{u_{i+1,j-1,k}^1 - u_{i,j-1,k}^1}{h1_i} \right) \\
\epsilon_2 &= \epsilon_{22} = \frac{\partial u^2}{\partial y} = \frac{u_{i,j,k}^2 - u_{i,j-1,k}^2}{h2_{j-1}} \\
\epsilon_3 &= \epsilon_{33} = \frac{\partial u^3}{\partial z} = \frac{1}{2} \left(\frac{u_{i,j,k+1}^3 - u_{i,j,k}^3}{h3_k} + \frac{u_{i,j-1,k+1}^3 - u_{i,j-1,k}^3}{h3_k} \right)
\end{aligned}$$

$$\begin{aligned}
\epsilon_4 &= 2\epsilon_{23} = 2 \cdot \frac{1}{2} \left(\frac{\partial u^2}{\partial z} + \frac{\partial u^3}{\partial y} \right) = 2 \cdot \frac{1}{2} \left[\frac{1}{2} \left(\frac{u_{i,j,k+1}^2 - u_{i,j,k}^2}{h3_k} + \frac{u_{i,j-1,k+1}^2 - u_{i,j-1,k}^2}{h3_k} \right) + \frac{u_{i,j,k}^3 - u_{i,j-1,k}^3}{h2_{j-1}} \right] \\
\epsilon_5 &= 2\epsilon_{13} = 2 \cdot \frac{1}{2} \left(\frac{\partial u^1}{\partial z} + \frac{\partial u^3}{\partial x} \right) = 2 \cdot \frac{1}{2} \left[\frac{1}{2} \left(\frac{u_{i,j,k+1}^1 - u_{i,j,k}^1}{h3_k} + \frac{u_{i,j-1,k+1}^1 - u_{i,j-1,k}^1}{h3_k} \right) + \frac{1}{2} \left(\frac{u_{i+1,j,k}^3 - u_{i,j,k}^3}{h1_i} + \frac{u_{i+1,j-1,k}^3 - u_{i,j-1,k}^3}{h1_i} \right) \right] \\
\epsilon_6 &= 2\epsilon_{12} = 2 \cdot \frac{1}{2} \left(\frac{\partial u^1}{\partial y} + \frac{\partial u^2}{\partial x} \right) = 2 \cdot \frac{1}{2} \left[\frac{u_{i,j,k}^1 - u_{i,j-1,k}^1}{h2_{j-1}} + \frac{1}{2} \left(\frac{u_{i+1,j,k}^2 - u_{i,j,k}^2}{h1_i} + \frac{u_{i+1,j-1,k}^2 - u_{i,j-1,k}^2}{h1_i} \right) \right]
\end{aligned}$$

Daraus kann analog zu oben der Differenzenstern abgelesen werden. Für die Differenziale durch die Teilfläche A_{n3} ergibt sich schließlich:

$$\begin{aligned}
\epsilon_1 &= \epsilon_{11} = \frac{\partial u^1}{\partial x} = \frac{1}{2} \left(\frac{u_{i+1,j,k}^1 - u_{i,j,k}^1}{h1_i} + \frac{u_{i+1,j,k+1}^1 - u_{i,j,k+1}^1}{h1_i} \right) \\
\epsilon_2 &= \epsilon_{22} = \frac{\partial u^2}{\partial y} = \frac{1}{2} \left(\frac{u_{i,j,k}^2 - u_{i,j-1,k}^2}{h2_{j-1}} + \frac{u_{i,j,k+1}^2 - u_{i,j-1,k+1}^2}{h2_{j-1}} \right) \\
\epsilon_3 &= \epsilon_{33} = \frac{\partial u^3}{\partial z} = \frac{u_{i,j,k+1}^3 - u_{i,j,k}^3}{h3_k} \\
\epsilon_4 &= 2\epsilon_{23} = 2 \cdot \frac{1}{2} \left(\frac{\partial u^2}{\partial z} + \frac{\partial u^3}{\partial y} \right) = 2 \cdot \frac{1}{2} \left[\frac{u_{i,j,k+1}^2 - u_{i,j,k}^2}{h3_k} + \frac{1}{2} \left(\frac{u_{i,j,k}^3 - u_{i,j-1,k}^3}{h2_{j-1}} + \frac{u_{i,j,k+1}^3 - u_{i,j-1,k+1}^3}{h2_{j-1}} \right) \right] \\
\epsilon_5 &= 2\epsilon_{13} = 2 \cdot \frac{1}{2} \left(\frac{\partial u^1}{\partial z} + \frac{\partial u^3}{\partial x} \right) = 2 \cdot \frac{1}{2} \left[\frac{u_{i,j,k+1}^1 - u_{i,j,k}^1}{h3_k} + \frac{1}{2} \left(\frac{u_{i+1,j,k}^3 - u_{i,j,k}^3}{h1_i} + \frac{u_{i+1,j,k+1}^3 - u_{i,j,k+1}^3}{h1_i} \right) \right] \\
\epsilon_6 &= 2\epsilon_{12} = 2 \cdot \frac{1}{2} \left(\frac{\partial u^1}{\partial y} + \frac{\partial u^2}{\partial x} \right) = 2 \cdot \frac{1}{2} \left[\frac{1}{2} \left(\frac{u_{i,j,k}^1 - u_{i,j-1,k}^1}{h2_{j-1}} + \frac{u_{i,j,k+1}^1 - u_{i,j-1,k+1}^1}{h2_{j-1}} \right) + \frac{1}{2} \left(\frac{u_{i+1,j,k}^2 - u_{i,j,k}^2}{h1_i} + \frac{u_{i+1,j,k+1}^2 - u_{i,j,k+1}^2}{h1_i} \right) \right]
\end{aligned}$$

Auf diese Weise werden der Reihe nach für alle Oktanten die Differenzensterne assembliert und aufaddiert. Da die Gleichung drei Komponenten aufweist, erhält man für jede Gleichungskomponente einen Differenzenstern U_1 (1...57), U_2 (1...57) und U_3 (1...57). Die nichtverschwindenden Elemente werden nun in die Matrix übernommen.

Matrix Aufstellen

Das Rechtecksgitter habe die Dimension $n1 \times n2 \times n3$, d.h. für die Indices gilt: $i = 1, \dots, n1$, $j = 1, \dots, n2$ und $k = 1, \dots, n3$. Dann besteht das Gitter aus insgesamt $N = n1 \cdot n2 \cdot n3$ Gitterpunkten. Da wir pro Punkt drei Komponenten haben, ergibt sich folgende Basis:

$$\vec{u} = (u_1^1, u_1^2, u_1^3, \dots, u_m^1, u_m^2, u_m^3, \dots, u_N^1, u_N^2, u_N^3) \quad (\text{B.12})$$

Dabei gilt für den Punkt (i, j, k) die Zuordnung $m = (k-1) \cdot n2 \cdot n1 + (j-1) \cdot n1 + i$. Es soll nun beispielhaft das Element $U_2(24)$ für den Gitterpunkt (i, j, k) in die Matrix übernommen werden. Die Zeile des entsprechenden Matrixelements ist:

$$\text{Zeile} = [(k-1) \cdot n2 \cdot n1 + (j-1) \cdot n1 + i - 1] \cdot 3 + 2.$$

Laut der Definition des Differenzensterns (B.6) entspricht der Eintrag der Nummer 24 dem Gitterpunkt $u_{i-1, j-1, k}^3$. Daraus folgt für die Spalte des Matrixelements:

$$\text{Spalte} = [(k-1) \cdot n2 \cdot n1 + ((j-1) - 1) \cdot n1 + (i-1) - 1] \cdot 3 + 3.$$

Auf diese Weise werden der Reihe nach für jeden Gitterpunkt alle nichtverschwindenden Einträge des Differenzensterns in die Matrix übernommen.

und

$$H_{cv} = \begin{pmatrix} H_{cv}^{sx} & H_{cv}^{sy} & H_{cv}^{sz} & 0 & 0 & 0 \\ 0 & 0 & 0 & H_{cv}^{sx} & H_{cv}^{sy} & H_{cv}^{sz} \end{pmatrix}$$

mit

$$\begin{aligned} H_{cv}^{sx} &= B_1 k_y k_z + iP_2 k_x \\ H_{cv}^{sy} &= B_2 k_x k_z + iP_2 k_y \\ H_{cv}^{sz} &= B_3 k_x k_y + iP_1 k_z \end{aligned} \quad (C.8)$$

Für die Dipolmatrixelemente gilt:

$$\begin{aligned} P_1 &= -\frac{i\hbar}{m} \langle s|p_z|z \rangle \\ P_2 &= -\frac{i\hbar}{m} \langle s|p_x|x \rangle = -\frac{i\hbar}{m} \langle s|p_y|y \rangle \end{aligned}$$

Spin-Bahn-Wechselwirkung

Die Spin-Bahn-Wechselwirkung lautet in der Basis $\{|x \uparrow\rangle, |y \uparrow\rangle, |z \uparrow\rangle, |x \downarrow\rangle, |y \downarrow\rangle, |z \downarrow\rangle\}$:

$$H_{SO} = \begin{pmatrix} 0 & -i\Delta_2 & 0 & 0 & 0 & \Delta_3 \\ i\Delta_2 & 0 & 0 & 0 & 0 & -i\Delta_3 \\ 0 & 0 & 0 & -\Delta_3 & i\Delta_3 & 0 \\ 0 & 0 & -\Delta_3 & 0 & i\Delta_2 & 0 \\ 0 & 0 & -i\Delta_3 & -i\Delta_2 & 0 & 0 \\ \Delta_3 & i\Delta_3 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Die Basiszustände, die diese Matrix diagonalisieren, sind [25]:

$$|u_{hh\uparrow}\rangle = -\frac{1}{\sqrt{2}}|(x+iy)\uparrow\rangle \quad (C.9)$$

$$|u_{ih\uparrow}\rangle = \frac{1}{\sqrt{2}}|(x-iy)\uparrow\rangle \quad (C.10)$$

$$|u_{sh\uparrow}\rangle = |z\uparrow\rangle \quad (C.11)$$

$$|u_{hh\downarrow}\rangle = \frac{1}{\sqrt{2}}|(x-iy)\downarrow\rangle \quad (C.12)$$

$$|u_{ih\downarrow}\rangle = -\frac{1}{\sqrt{2}}|(x+iy)\downarrow\rangle \quad (C.13)$$

$$|u_{sh\downarrow}\rangle = |z\downarrow\rangle \quad (C.14)$$

Ist Δ_{SO} die Spin-Bahnaufspaltung, so gilt:

$$\Delta_1 = \Delta_2 = \frac{1}{3}\Delta_{SO}.$$

Kristallfeld-Aufspaltung

In Wurtzit tritt zusätzlich die sog. Kristallfeld-Aufspaltung auf. Sie wird durch die (3×3) -Matrix

$$\begin{pmatrix} \Delta_1 & 0 & 0 \\ 0 & \Delta_1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

in den Unterräumen $\{|x \uparrow\rangle, |y \uparrow\rangle, |z \uparrow\rangle\}$ und $\{|x \downarrow\rangle, |y \downarrow\rangle, |z \downarrow\rangle\}$ berücksichtigt.

Verspannungen

Die Verspannungseffekte werden durch den Hamiltonoperator

$$H_\epsilon^{ij} = \sum_{(\alpha,\beta)=x,y,z} D_{ij}^{\alpha\beta} \epsilon_{\alpha\beta}$$

mit den Deformationspotenzial-Matrizen $\mathbf{D}^{\alpha\beta}$ beschrieben.

In der Basis $\{|x \uparrow\rangle, |y \uparrow\rangle, |z \uparrow\rangle\}$ und $\{|x \downarrow\rangle, |y \downarrow\rangle, |z \downarrow\rangle\}$ ergibt sich:

$$H_\epsilon = \begin{pmatrix} l_1\epsilon_{xx} + m_1\epsilon_{yy} + m_2\epsilon_{zz} & n_1\epsilon_{xy} & n_2\epsilon_{xz} \\ n_1\epsilon_{xy} & m_1\epsilon_{xx} + l_1\epsilon_{yy} + m_2\epsilon_{zz} & n_2\epsilon_{yz} \\ n_2\epsilon_{xz} & n_2\epsilon_{yz} & m_3(\epsilon_{xx} + \epsilon_{yy}) + l_2\epsilon_{zz} \end{pmatrix}$$

Die Deformationspotenzial-Matrizen $\mathbf{D}^{\alpha\beta}$ haben die Form:

$$\begin{aligned} D^{xx} &= \begin{pmatrix} l_1 & 0 & 0 \\ 0 & m_1 & 0 \\ 0 & 0 & m_3 \end{pmatrix}, & D^{xy} &= \begin{pmatrix} 0 & n_1/2 & 0 \\ n_1/2 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\ D^{zz} &= \begin{pmatrix} m_2 & 0 & 0 \\ 0 & m_2 & 0 \\ 0 & 0 & l_2 \end{pmatrix}, & D^{xz} &= \begin{pmatrix} 0 & 0 & n_2/2 \\ 0 & 0 & 0 \\ n_2/2 & 0 & 0 \end{pmatrix} \\ D^{yy} &= \begin{pmatrix} m_1 & 0 & 0 \\ 0 & l_1 & 0 \\ 0 & 0 & m_3 \end{pmatrix}, & D^{yz} &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & n_2/2 \\ 0 & n_2/2 & 0 \end{pmatrix} \end{aligned} \quad (C.15)$$

Die Auswirkung des Verspannungstensors ϵ auf das Leitungsband wird durch das absolute Deformationspotenzial für das Leitungsband a_c beschrieben

$$E_c \rightarrow E_c + a_c Sp(\epsilon).$$

Spezielle Wurtzit Parameter

Für Wurtzit werden gewöhnlich die Parameter A_1, \dots, A_6 anstelle von L_1, \dots, N_2 verwendet. Es gelten folgende Beziehungen:

$$L_1 = N_1 + M_1 = \frac{\hbar^2}{2m_0}(A_5 + A_4 + A_2 - 1) \quad (\text{C.16})$$

$$L_2 = \frac{\hbar^2}{2m_0}(A_1 - 1) \quad (\text{C.17})$$

$$M_1 = \frac{\hbar^2}{2m_0}(A_4 + A_2 - A_5 - 1) \quad (\text{C.18})$$

$$M_2 = \frac{\hbar^2}{2m_0}(A_1 + A_3 - 1) \quad (\text{C.19})$$

$$M_3 = \frac{\hbar^2}{2m_0}(A_2 - 1) \quad (\text{C.20})$$

$$N_1 = \frac{\hbar^2}{2m_0}2A_5 \quad (\text{C.21})$$

$$N_2 = \frac{\hbar^2}{2m_0}\sqrt{2}A_6 \quad (\text{C.22})$$

Für die Deformationspotenziale gelten ähnliche Relationen:

$$l_1 = n_1 + m_1 = d_5 + d_4 + d_2 \quad (\text{C.23})$$

$$l_2 = d_1 \quad (\text{C.24})$$

$$m_1 = d_4 + d_2 - d_5 \quad (\text{C.25})$$

$$m_2 = d_1 + d_3 \quad (\text{C.26})$$

$$m_3 = d_2 \quad (\text{C.27})$$

$$n_1 = 2d_5 \quad (\text{C.28})$$

$$n_2 = \sqrt{2}d_6 \quad (\text{C.29})$$

Anhang D

Modelle für die Dichteberechnung in 1D, 2D und 3D

Die Dichte wird entweder klassisch in Thomas-Fermi-Näherung oder quantenmechanisch durch Lösen der Schrödingergleichung berechnet. In einer Quantenregion kann es sinnvoll sein, Eigenzustände nur innerhalb eines bestimmten Energiebereichs bei der Dichteberechnung zu berücksichtigen und für darüber liegende Zustände die Thomas-Fermi-Zustandsdichte zugrunde zu legen. Diese Energie bezeichnen wir fortan als Trennenergie E_{Trenn} . Will man die Dichte voll quantenmechanisch berücksichtigen, so muss man genügend viele Eigenzustände berechnen und die Trennenergie gleich dem höchsten Eigenwert setzen. Dann wird der Beitrag der Thomas-Fermi-Zustandsdichte oberhalb der Trennenergie durch das Ferminiveau unterdrückt. In Einzelfällen kann es sinnvoll sein, nur vollständig gebundene Zustände miteinzubeziehen und im Übrigen die Dichte klassisch zu berechnen. Ein Beispiel dafür ist in Abb. D.1 gegeben. In Folgenden sollen die Ausdrücke für die Dichten in einer, zwei und drei Dimensionen kurz hergeleitet und zusammengestellt werden.

Klassische Dichten

Die Thomas-Fermi-Dichte kann folgendermaßen abgeleitet werden: Ausgangspunkt ist die Zustandsdichte im Volumenhalbleiter:

$$g(E) = \sum_{\gamma} \delta(E - E_{\gamma}) |\Psi_{\gamma}(x, y, z)|^2 \quad (D.1)$$

Dabei gilt für die Wellenfunktionen der freien Zustände:

$$\Psi_{\vec{k}}(\vec{r}) = \frac{1}{\sqrt{V}} e^{i\vec{k}\vec{r}} \quad (D.2)$$

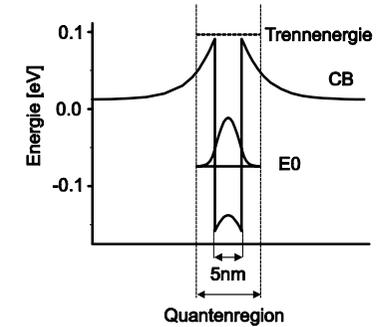


Abbildung D.1: Schematische Darstellung eines AlGaAs/GaAs-Quantentopfes mit einem gebundenen Zustand. Nur dieser wird quantenmechanisch berücksichtigt, oberhalb der Trennenergie wird die Dichte klassisch berechnet.

Mit der Formel $n = \int g(E)f(E)dE$ erhält man für die Elektronendichte:

$$n(x, y, z) = \sum_{\vec{k}} |\Psi_{\vec{k}}(x, y, z)|^2 f\left(\frac{E_{\vec{k}} - E_F}{kT}\right), \quad (D.3)$$

Die Zustände werden über die Fermifunktion f und das Ferminiveau E_F besetzt. Im Folgenden wird eine parabolische Dispersion zugrunde gelegt (mit der effektiven Masse m^* und der Bandkante E_C):

$$E_{\vec{k}}(x, y, z) = E_C(x, y, z) + \frac{\hbar^2 k^2}{2m^*} \quad (D.4)$$

Die Summation in Gl. (D.3) kann durch eine Integration ersetzt werden:

$$\sum_{\vec{k}} = \frac{V}{(2\pi)^3} \int_0^{\infty} d\vec{k} \quad g_{Tal/Spin} \quad (D.5)$$

Hier bedeutet V das Normierungsvolumen. Der Faktor $g_{Tal/Spin}$ enthält die Spin- und Talentartung. Dann ergibt sich für die Elektronendichte:

$$n(x, y, z) = g_{Tal/Spin} \frac{1}{8\pi^{3/2}} \left(\frac{2m^*}{\hbar^2}\right)^{3/2} (kT)^{3/2} F_{1/2} \left(\frac{E_F - E_C}{kT}\right) \quad (D.6)$$

Dabei bezeichnet $F_{1/2}$ das Fermi-Integral der Ordnung 1/2, E_F ist die Fermienergie. Mit der Abkürzung $N_c = g_{Tal/Spin} \frac{1}{8\pi^{3/2}} \left(\frac{2m^*}{\hbar^2}\right)^{3/2} (kT)^{3/2}$ erhält man schließlich:

$$n(x, y, z) = N_c F_{1/2} \left(\frac{E_F - E_C}{kT}\right) \quad (D.7)$$

Nichtparabolizitäten werden durch den sog. Nichtparabolizitäts-Faktor α_{eV} berücksichtigt. Dieser ist folgendermaßen definiert (E kinetische Energie):

$$\frac{\hbar^2 k^2}{2m} = E(1 + \alpha_{eV} E)$$

Definitionsgemäß hat er die Einheit $[\alpha_{eV}] = \frac{1}{eV}$. Mit der Transformation

$$\alpha = \alpha_{eV} kT$$

wird α dimensionslos und es gilt nun für die Dichte:

$$n(x, y, z) = N_C F_{1/2} \left(\frac{E_F - E_C}{kT} \right) + \frac{5}{2} \alpha N_C F_{3/2} \left(\frac{E_F - E_C}{kT} \right) \quad (\text{D.8})$$

In Quantenregionen wird ab einer gewissen Trennenergie - wie oben beschrieben - die Dichte klassisch berechnet. Dieser Anteil ergibt sich durch Integration über die Thomas-Fermi-Zustandsdichte von E_{Trenn} ins Unendliche (Gl. D.5). Dadurch erhält man die sog. unvollständigen Fermi-Integrale:

$$n(x, y, z) = N_C F_{1/2} \left(\frac{E_F - E_C}{kT}, \max\left[\frac{E_{edge} - E_C}{kT}, 0\right] \right) + \frac{5}{2} \alpha N_C F_{3/2} \left(\frac{E_F - E_C}{kT}, \max\left[\frac{E_{edge} - E_C}{kT}, 0\right] \right) \quad (\text{D.9})$$

Die unvollständigen Fermi-Integrale werden in Anhang F näher erläutert. Für die Löcherdichten muss man nur $\frac{E_F - E_C}{kT}$ durch $-\frac{E_F - E_V}{kT}$ ersetzen.

Quantenmechanische Dichten in Einband-Näherung

Quantenmechanische Zustände werden entweder in Einband-Näherung oder im Rahmen der kp-Theorie berechnet. Im Folgenden sollen kurz die Formeln für die Dichteberechnung mit den Zuständen aus der Einband-Schrödingergleichung angegeben werden.

3D-Zustände Die Dichte für die dreidimensionalen Zustände der Einband-Schrödingergleichung ergibt sich einfach durch Besetzen über das Fermineiveau:

$$n(x, y, z) = g_{Tal/Spin} \sum_{\gamma} |\Psi_{\gamma}(x, y, z)|^2 f\left(\frac{E_{\gamma} - E_F}{kT}\right)$$

2D-Zustände (homogen in z -Richtung) Die quantenmechanischen Zustände sind die Lösungen der zweidimensionalen Schrödingergleichung. Wir wollen die Dichte herleiten für eine Quantenregion, die homogen in z -Richtung ist. Dann kann man die Wellenfunktion ansetzen als:

$$\Psi_{\gamma}(x, y, z) = \frac{1}{\sqrt{L_z}} e^{ik_z z} \frac{1}{\sqrt{A}} \Phi_n(x, y)$$

Hier ist $\Phi_n(x, y)$ die Lösung der zweidimensionalen Schrödingergleichung für $k_z = 0$ mit Eigenwert E_n . Für die kinetische Energie gilt nun:

$$E_{\gamma}(n, k_z; x, y) = E_n + \frac{\hbar^2 k_z^2}{2m_z(x, y)}$$

Allgemein ist $n = \int g(E) f(E) dE$ (Gl. D.1), so dass sich für die Elektronendichte ergibt:

$$n(x, y, z) = \sum_{\gamma} |\Psi_{\gamma}(x, y, z)|^2 f\left(\frac{E_{\bar{k}} - E_F}{kT}\right),$$

wo $f\left(\frac{E_{\bar{k}} - E_F}{kT}\right)$ die Fermifunktion bedeutet. Die Summe kann aufgespalten werden in eine Summation über die quantisierten Zustände und eine Integration über k_z :

$$\sum_{\gamma} = \sum_n \sum_{k_z} = \sum_n \frac{L_z}{2\pi} \int_0^{\infty} dk_z g_{Tal/Spin}$$

Dadurch erhält man:

$$n(x, y, z) = g_{Tal/Spin} \frac{1}{2\pi} \sum_n \frac{1}{A} |\Phi_n(x, y)|^2 \int dk_z f\left(\frac{E_{\bar{k}} - E_F}{kT}\right)$$

Daraus folgt für die Dichte (bei Integration bis ∞):

$$n(x, y, z) = \sum_n \frac{|\Phi_n(x, y)|^2}{A} \sqrt{\frac{m_z(x, y) kT}{\hbar^2}} \frac{1}{2\pi} F_{-1/2} \left(\frac{E_F - E_n}{kT} \right) \quad (\text{D.10})$$

Im Allgemeinen wird oberhalb der Trennenergie die Thomas-Fermi-Zustandsdichte zugrunde gelegt. Dann müssen wir für die gebundenen Zustände nur bis zu dieser Energie integrieren:

$$n(x, y, z) = g_{Tal/Spin} \frac{1}{2\pi} \sum_n \frac{1}{A} |\Phi_n(x, y)|^2 \int_0^{\frac{E_{Trenn} - E_n}{kT}} dk_z f\left(\frac{E_{\bar{k}} - E_F}{kT}\right)$$

Dies lässt sich umformen zu:

$$n(x, y, z) = \sum_n |\Phi_n(x, y)|^2 \sqrt{\frac{m_z(x, y) kT}{\hbar^2}} \frac{1}{2\pi} \frac{1}{2} g_{\text{Tal/spin}} \left[F_{-1/2} \left(\frac{E_F - E_n}{kT} \right) - F_{-1/2} \left(\frac{E_F - E_n}{kT}, \max \left[\frac{E_{\text{edge}} - E_n}{kT}, 0 \right] \right) \right]$$

(D.10)

1D-Zustände (homogen in (x, y) -Richtung) Mit den eindimensionalen Wellenfunktionen erhält man analog zu oben folgenden Ausdruck für die Dichte:

$$n_{1D}(x, y, z) = \sum_{\gamma} g_{\text{valley/spin}} \frac{m^* kT}{2\pi \hbar^2} |\Phi_n(z)|^2 \ln \left(\frac{1 + e^{-\epsilon_{\gamma} + \epsilon_{F\gamma}}}{1 + e^{-\epsilon_{\text{edge}} + \epsilon_{F\gamma}}} \right) \quad (\text{D.11})$$

Ableitungen der Dichten

Zur Berechnung der Newton-Korrektur für die nichtlineare Poissonsgleichung muss die Ableitung der Dichte nach dem Potenzial berechnet werden (vgl. Kapitel 5). Es werden im Folgenden die wichtigsten Relationen angegeben, die in diesem Zusammenhang verwendet werden:

Für die Fermi-Integrale gilt:

$$\frac{d}{dx} F_j(x) = F_{j-1}(x)$$

Es ist zu beachten, dass für die unvollständigen Fermi-Integralen die Integrationsgrenzen ebenfalls vom Potenzial abhängen:

$$F_j[\epsilon(\vec{r}), b(\vec{r}), c(\vec{r})] \equiv \frac{1}{\Gamma(j+1)} \int_{b(\vec{r})}^{c(\vec{r})} du \frac{u^j}{1 + e^{u-\epsilon(\vec{r})}} \quad (\text{D.12})$$

Betrachten wir dazu folgendes Funktional:

$$G(y) = \int_{\Phi(y)}^{\Psi(y)} g(z, y) dz$$

Dann gilt für die Ableitung:

$$\frac{d}{dy} G(y) = \int_{\Phi(y)}^{\Psi(y)} \frac{\partial g(z, y)}{\partial y} dz + g(\Psi(y), y) \frac{d\Psi(y)}{dy} - g(\Phi(y), y) \frac{d\Phi(y)}{dy}$$

Für das Fermi-Integral (D.12) folgt daraus mit partieller Integration:

$$\begin{aligned} \nabla_{\vec{r}} F_j[\epsilon(\vec{r}), b(\vec{r}), c(\vec{r})] &= F_{j-1}[\epsilon(\vec{r}), b(\vec{r}), c(\vec{r})] \nabla_{\vec{r}} \epsilon(\vec{r}) \\ &- \frac{1}{\Gamma(j+1)} \frac{b(\vec{r})^j}{1 + e^{b(\vec{r})-\epsilon(\vec{r})}} \nabla_{\vec{r}} [b(\vec{r}) - \epsilon(\vec{r})] \\ &+ \frac{1}{\Gamma(j+1)} \frac{c(\vec{r})^j}{1 + e^{c(\vec{r})-\epsilon(\vec{r})}} \nabla_{\vec{r}} [c(\vec{r}) - \epsilon(\vec{r})] \quad (\text{D.13}) \end{aligned}$$

Die Γ -Funktion erfüllt einige nützliche Relationen:

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi} \quad (\text{D.14})$$

$$\Gamma\left(-\frac{1}{2}\right) = -2\sqrt{\pi} \quad (\text{D.15})$$

$$\Gamma\left(\frac{3}{2}\right) = \frac{\sqrt{\pi}}{2} \quad (\text{D.16})$$

$$\Gamma(x+1) = x\Gamma(x) \quad (\text{D.17})$$

KP-Zustände in 1D

Die kp-Zustände bei 8×8 -kp-Theorie lauten bei Quantisierung in z -Richtung und Homogenität in (x, y) -Richtung

$$\Psi_{\nu}(x, y, z) = \sum_{i=1}^8 \chi_{i\nu}(z) e^{i(k_x x + k_y y)} u_i \quad (\text{D.18})$$

mit den Blochfunktionen u_i :

$$\begin{aligned} u_1 &= |s \uparrow\rangle \\ u_2 &= |s \downarrow\rangle \\ u_3 &= |x \uparrow\rangle \\ u_4 &= |x \downarrow\rangle \\ u_5 &= |y \uparrow\rangle \\ u_6 &= |y \downarrow\rangle \\ u_7 &= |z \uparrow\rangle \\ u_8 &= |z \downarrow\rangle \quad (\text{D.19}) \end{aligned}$$

Für die Zustandsdichte gilt allgemein:

$$g(E) = \sum_{\gamma} \delta(E - E_{\gamma}) |\Psi_{\gamma}(x, y, z)|^2 \quad (\text{D.20})$$

Daraus ergibt sich für die Dichte der Ausdruck:

$$\begin{aligned} n &= \int g(E) f(E) dE = \int \sum_{\gamma} \delta(E - E_{\gamma}) |\Psi_{\gamma}(x, y, z)|^2 f(E) dE = \\ &= \sum_{k_z} \sum_{k_x, k_y} |\Psi_{\gamma}(x, y, z)|^2 f(E_{\gamma}) \end{aligned} \quad (\text{D.21})$$

Weiterhin ist:

$$|\Psi_{\gamma}(x, y, z)|^2 = \sum_{i,j=1}^8 \chi_{i\gamma} \chi_{j\gamma}^* u_i u_j^*$$

Nun muss man noch die Blochanteile $u_i u_j^*$ loswerden. Dabei argumentiert man folgendermaßen: Wenn wir die Dichte an der Stelle (x, y, z) berechnen, meinen wir die durchschnittliche Dichte in dem finiten Volumen am Gitterpunkt (x, y, z) . Man nimmt nun an, dass die Funktionen $u_i u_j^*$ stark variieren, so dass sich die Integration separieren lässt zu:

$$\int_{\text{Box}} \sum_{i,j=1}^8 \chi_{i\gamma} \chi_{j\gamma}^* u_i u_j^* d^3 r = \int_{\text{Box}} \sum_{i,j=1}^8 \chi_{i\gamma} \chi_{j\gamma}^* d^3 r \int_{\text{Box}} u_i u_j^* d^3 r$$

Mit

$$\int u_i u_j^* d^3 r = \delta_{ij}$$

und

$$\sum_{k_x, k_y} \rightarrow \frac{1}{(2\pi)^2} \int dk_x dk_y$$

erhält man schließlich:

$$n = \sum_{k_z} \frac{1}{(2\pi)^2} \int dk_x dk_y \sum_i |\chi_{i\gamma}|^2 f(E_{\gamma}) \quad (\text{D.22})$$

1D-Integration

Wenn die Energiedispersion in der 2D-Brillouinzone isotrop ist, d.h. $E(k_x, k_y) = E(k_{\parallel})$ mit $k_{\parallel}^2 = k_x^2 + k_y^2$, so kann die Integration in Gl. (D.22) auf ein eindimensionales Integral zurückgeführt werden mit $\int dk_x dk_y \rightarrow \int 2\pi k_{\parallel} dk_{\parallel}$:

$$n = \sum_{k_z} \frac{1}{2\pi} \int dk_{\parallel} k_{\parallel} \sum_i |\chi_{i\gamma}|^2 f(E(k_{\parallel})) \quad (\text{D.23})$$

2D-Integration, Diskretisierung der 2D-Brillouinzone

Die zweidimensionale Brillouinzone wird homogen diskretisiert, wobei a_k die Fläche ist, die ein diskreter k-Punkt im k-Raum belegt. Dann folgt für die Dichte:

$$\begin{aligned} n &= \sum_{k_z} \frac{1}{(2\pi)^2} \int dk_x dk_y \sum_i |\chi_{i\gamma}|^2 f(E_{\gamma}) = \\ &= \sum_{k_z} \sum_{k_x, k_y} f(E_{\gamma}) \sum_i |\chi_{i\gamma}|^2 w(k_x, k_y), \end{aligned} \quad (\text{D.24})$$

wobei $w(k_x, k_y) = \frac{1}{(2\pi)^2} a_k$.

Zustandsdichte-Integration

Eine weitere Möglichkeit besteht darin, zuerst die Zustandsdichte zu berechnen und dann daraus die Dichte zu ermitteln:

$$\begin{aligned} n &= \sum_{k_z} \frac{1}{(2\pi)^2} \int dk_x dk_y \sum_i |\chi_{i\gamma}|^2 f(E_{\gamma}) = \\ &= \sum_{k_z} \frac{1}{(2\pi)^2} \int dE \int dk_x dk_y \delta(E - E_{\gamma}) f(E) \sum_i |\chi_{i\gamma}|^2 = \\ &= \sum_{k_z} \frac{1}{(2\pi)^2} \int g(E) dE \end{aligned} \quad (\text{D.25})$$

mit

$$g(E) = \int dk_x dk_y \delta(E - E_{\gamma}) f(E) \sum_i |\chi_{i\gamma}|^2 \quad (\text{D.26})$$

Zur Berechnung der Zustandsdichte $g(E)$ existiert ein eigenes Programm.

Anhang E

Diskretisierung des Hamiltonoperators

In diesem Kapitel wird näher auf die numerische Behandlung der Schrödinger-Gleichung eingegangen. Dazu muss zunächst einmal der Hamilton-Operator auf ein diskretes Gitter abgebildet werden. Die Eigenwerte und Eigenzustände erhält man schließlich durch Diagonalisieren der resultierenden Matrix. Die Diskretisierung ist technisch recht aufwändig, da wir für Quantenregionen eine beliebige Geometrie zulassen. Im Folgenden soll unser Vorgehen anhand eines zweidimensionalen Beispiels illustriert werden.

Basis der Hamiltonmatrix

Das gesamte Bauelement wird - wie bereits erwähnt - auf einem Rechtecksgitter dargestellt. In unserem Beispiel handelt es sich dabei um ein Gitter mit 10×11 Gitterpunkten (Abb. E.1). Die Quantenregion ist durch eine gewisse Zahl von zusammenhängenden Punkten definiert. Diese werden in unserem Beispiel von 1 bis 35 durchnummeriert. Die Basis der Hamiltonmatrix umfasst demzufolge 35 Gitterpunkte. Im Weiteren richten wir unser Augenmerk auf die 8×8 -kp-Theorie. Dann müssen je Gitterpunkt 8 Spinorkomponenten berücksichtigt werden. Daraus ergibt sich schließlich folgende $8 \cdot 35$ -komponentige Basis für die Eigenzustände:

$$\text{Basis} = \left[\left(|s \uparrow\rangle |s \downarrow\rangle |x \uparrow\rangle |y \uparrow\rangle |z \uparrow\rangle |x \downarrow\rangle |y \downarrow\rangle |z \downarrow\rangle \right) \Big|_1, \dots, \dots, \left(|s \uparrow\rangle |s \downarrow\rangle |x \uparrow\rangle |y \uparrow\rangle |z \uparrow\rangle |x \downarrow\rangle |y \downarrow\rangle |z \downarrow\rangle \right) \Big|_{35} \right]^T$$

Der Index bezeichnet dabei den Gitterpunkt.

Diskretisierung

Beim Aufstellen der Hamiltonmatrix geht man alle Punkte von 1 bis 35 durch und bildet jeweils den so genannten Differenzstern. Dieser ergibt sich aus

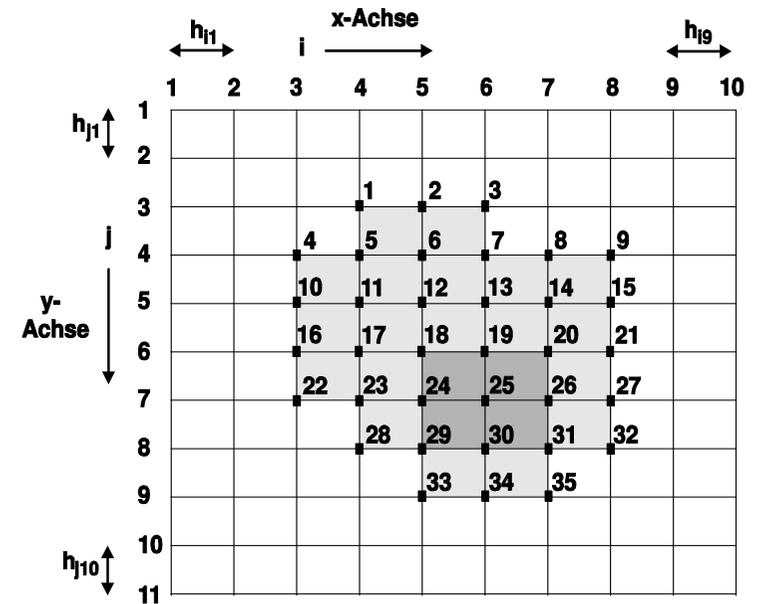


Abbildung E.1: Das Bauelement wird auf ein 10×11 -Rechtecksgitter abgebildet. Die Quantenregion ist grau unterlegt dargestellt. Die Gitterpunkte innerhalb der Quantenregion sind von 1 bis 35 durchnummeriert.

Die Diagonalmatrix \mathbf{D} hat schließlich die Gestalt

$$D = \begin{pmatrix} d_1, d_1, d_1, d_1, d_1, d_1, d_1, d_1, \dots, d_{25}, d_{25}, d_{25}, d_{25}, d_{25}, d_{25}, d_{25}, d_{25}, \dots, \\ \dots, d_{35}, d_{35}, d_{35}, d_{35}, d_{35}, d_{35}, d_{35}, d_{35} \end{pmatrix}^T$$

mit

$$d_{25} = \frac{1}{(h_{i5} + h_{i6})(h_{j6} + h_{j7})}$$

Matrix beschreiben

Jetzt muss nur noch der Differenzenstern in die Hamiltonmatrix geschrieben werden. Da wir uns auf dem Gitterpunkt 25 befinden, gehen die neun 8×8 -Blöcke in die Zeilen $(25 - 1) \cdot 8 + 1$ bis $25 \cdot 8$ der Systemmatrix ein. Die jeweilige Spalte ergibt sich aus der Nummer des entsprechenden Nachbarpunktes gemäß folgender Zuordnung:

Position relativ zu (i,j)	Spalte in Systemmatrix
$(i - 1, j - 1)$	18
$(i, j - 1)$	19
$(i + 1, j - 1)$	20
$(i - 1, j)$	24
(i, j)	25
$(i + 1, j)$	26
$(i - 1, j + 1)$	29
$(i, j + 1)$	30
$(i + 1, j + 1)$	31

Konkret heisst dies, dass die Matrix $\tilde{H}_{i-1,j-1}^{25}$ in die Zeilen $(25 - 1) \cdot 8 + 1$ bis $25 \cdot 8$ und die Spalten $(18 - 1) \cdot 8 + 1$ bis $18 \cdot 8$ der Systemmatrix geschrieben wird.

Anhang F

Fermi-Funktionen

Wie in Anhang D deutlich wird, treten in den Halbleitergleichungen sehr häufig sog. Fermi-Integrale auf. Diese sind definiert als:

$$F_j(x) = \frac{1}{\Gamma(j+1)} \int_0^\infty \frac{t^j}{e^{t-x} + 1} dt \quad (\text{F.1})$$

Deshalb ist eine effiziente Auswertung dieser Funktionen sehr wichtig. In der Regel kommt man dabei mit analytischen Näherungsformeln aus. Wir konnten uns hier meist auf bereits bekannte Näherungen stützen, wobei aber viele Parameter durch lineare Ausgleichsrechnung neu bestimmt werden mussten. Diese recht technischen Dinge sind Gegenstand des Kapitels.

Allgemeines

Über die Berechnung der Fermifunktionen gibt es viele Veröffentlichungen. Zunächst werden wir eine allgemeine Methode vorstellen zur Auswertung der Fermi-Integrale beliebiger Ordnung j für beliebige x . In den darauf folgenden Kapiteln werden dann Näherungsformeln behandelt für die Fälle $j = 1/2$, $j = -1/2$ sowie die Umkehrung des Fermi-Integrals der Ordnung $j = 1/2$.

Die Methoden, die für die Berechnung dieser Integrale vorgeschlagen wurden, können in vier Kategorien unterteilt werden [55]:

- Reihenentwicklungen
- spezielle numerische Quadratschemata
- Tabellierungen
- Approximation durch Polynome oder rationale Funktionen

Michele Goano [55] gibt eine Reihenentwicklung an, mit der sich Fermifunktionen beliebiger Ordnung auswerten lassen. Die Grundidee besteht darin, den

Integrationsbereich aufzuteilen in die beiden Intervalle $[0, x]$ und (x, ∞) und für beide Teilbereiche unterschiedliche geometrische Reihendarstellungen zu verwenden:

$$\frac{1}{e^{t-x} + 1} = \begin{cases} \sum_{n=1}^{\infty} (-1)^{n-1} (e^{x-t})^n & \text{für } t > x \\ 1 - \sum_{n=1}^{\infty} (-1)^{n-1} (e^{t-x})^n & \text{für } t < x \end{cases} \quad (\text{F.2})$$

Die Integration wird gliedweise durchgeführt. Die Techniken zur effizienten Auswertung dieser Ausdrücke sind in [55] dargestellt. Da oft eine begrenzte Genauigkeit ausreicht, sind analytische Näherungsformeln sehr attraktiv. Diese werden in den folgenden Abschnitten behandelt.

Analytische Näherung für das Fermi-Dirac-Integral $F_{1/2}(x)$

X. Aymerich-Humet et al. [68] geben folgende Formel an:

$$F_{1/2}(x) \approx \left[\frac{3\sqrt{2}}{\left[b + x + (|x - b|^c + a)^{1/c} \right]^{3/2}} + \frac{e^{-x}}{\Gamma(3/2)} \right]^{-1} \quad (\text{F.3})$$

mit den Parametern

$$\begin{aligned} a &= 9.60 \\ b &= 2.13 \\ c &= 2.40 \end{aligned}$$

Der relative Fehler ist hierbei kleiner als 0.53%.

Genauer, dafür aber auch aufwändiger ist die Darstellung von P. Van Halen und D. L. Pulfrey [69]. Sie nehmen dabei eine Fallunterscheidung bezüglich x vor:

$$F_{1/2}(x) = \begin{cases} \sum_{r=1}^7 a_r e^{rx} & \text{für } x \leq 0 \\ \sum_{r=1}^7 b_r x^{r-1} & \text{für } 0 \leq x \leq 2 \\ \sum_{r=1}^7 c_r x^{r-1} & \text{für } 2 \leq x \leq 5 \\ x^{3/2} \sum_{r=1}^7 \frac{d_r}{x^{2(r-1)}} & \text{für } x \geq 5 \end{cases} \quad (\text{F.4})$$

Die Koeffizienten von Stephen A. Wong [70] wurden etwas modifiziert und sind am Ende des Kapitels angegeben. Man erreicht damit eine relative Genauigkeit von $4.9 \cdot 10^{-6}$.

Analytische Näherung für das Fermi-Dirac-Integral $F_{-1/2}(x)$

Hier haben wir die Methode von P. Van Halen und D. L. Pulfrey [69] etwas abgewandelt. Die Reihenentwicklung lautet nun:

$$F_{-1/2}(x) = \begin{cases} \sum_{r=1}^7 a_r e^{rx} & \text{für } x \leq -2 \\ \sum_{r=1}^7 a'_r e^{rx} & \text{für } -2 \leq x \leq 0 \\ \sum_{r=1}^9 b_r x^{r-1} & \text{für } 0 \leq x \leq 2.5 \\ \sum_{r=1}^9 c_r x^{r-1} & \text{für } 2.5 \leq x \leq 5 \\ x^{1/2} \sum_{r=1}^7 \frac{d_r}{x^{2(r-1)}} & \text{für } x \geq 5 \end{cases} \quad (\text{F.5})$$

Die Parameter, die in [69] angegeben sind, liefern ungenaue Resultate. Wir haben sie daher mittels linearer Ausgleichsrechnung neu bestimmt und am Ende des Kapitels aufgelistet. Die erzielte relative Genauigkeit liegt nun bei $6 \cdot 10^{-6}$.

Analytische Näherung für das Fermi-Dirac-Integral $F_{1/2}^{-1}(\eta)$

Eine einfache und recht genaue Formel wurde von N. G. Nilsson gefunden [71]:

$$x = F_{1/2}^{-1}(\eta) = \ln \eta + \frac{\eta}{[64 + 0.05524\eta(64 + \sqrt{\eta})]^{1/4}} \quad (\text{F.6})$$

Man erreicht damit eine Genauigkeit von $5 \cdot 10^{-5}$ für Werte bis zu $\eta = 34.3$. Als Alternative kann man die Nullstelle der Gleichung

$$F_{1/2}(x) - \eta = 0 \quad (\text{F.7})$$

mit der Newton-Raphson-Methode suchen, wobei man die Beziehung

$$\frac{d}{dx} F_{1/2}(x) = F_{-1/2}(x) \quad (\text{F.8})$$

ausnutzt. Die Funktionen $F_{1/2}(x)$ und $F_{-1/2}(x)$ können dabei mit dem Algorithmus von Goano [55] mit der gewünschten Genauigkeit berechnet werden.

Der Quotient $F_{\frac{1}{2}}/F_{-\frac{1}{2}}$

Der Funktionsverlauf des Quotienten $F_{\frac{1}{2}}/F_{-\frac{1}{2}}$ in Abhängigkeit vom Argument ist in Abb.(F.1) dargestellt [52]. Insbesondere gilt:

$$F_{\frac{1}{2}}(x)/F_{-\frac{1}{2}}(x) \rightarrow 1 \quad \text{für } x \ll -1 \quad (\text{F.9})$$

$$F_{\frac{1}{2}}(x)/F_{-\frac{1}{2}}(x) \rightarrow \frac{2}{3}x \quad \text{für } x \gg 1 \quad (\text{F.10})$$

Dieses Verhalten muss man bei der Implementierung für große und kleine Argumente berücksichtigen.

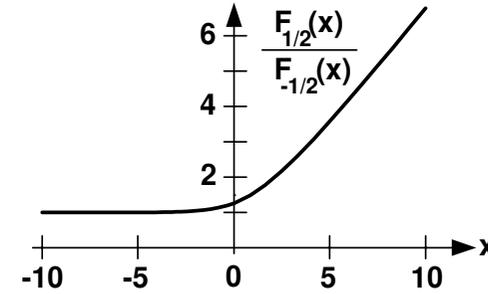


Abbildung F.1: Quotient der Fermi-Integrale der Ordnung 1/2 und -1/2.

Zusammenfassung

Zusammenfassend haben wir, ausgehend von bekannten Näherungen für Fermi-Integrale, diese modifiziert und die Parameter neu bestimmt. Man erreicht damit eine für uns in der Regel ausreichende Genauigkeit. Als Ergebnis liegt nun eine Programmbibliothek vor, die es gestattet, die Fermifunktionen in allen Varianten effizient auszuwerten. Die Koeffizienten sind für die Ordnung $j = 1/2$ und $j = -1/2$ (vgl.Gl. [F.4],[F.5]) in folgender Tabelle zusammengestellt:

		a	b	c	d	a'
$j = 1/2$	1	1.000000	0.765148	0.780957	0.752253	-
	2	-0.353568	0.604887	0.572544	0.928195	-
	3	0.192439	0.190034	0.214193	0.680839	-
	4	-0.122973	$2.00194 \cdot 10^{-2}$	$1.384327 \cdot 10^{-2}$	25.7829	-
	5	0.077134	$-4.126438 \cdot 10^{-3}$	$-5.549494 \cdot 10^{-3}$	-553.636	-
	6	-0.036228	$-4.70959 \cdot 10^{-4}$	$6.488149 \cdot 10^{-4}$	3531.43	-
	7	0.008346	$1.5007147 \cdot 10^{-4}$	$-2.840505 \cdot 10^{-5}$	-3254.65	-
$j = -1/2$	1	1.000000	0.6048983	0.4162655	1.1283792	0.9999729
	2	-0.707107	0.380116	0.7898693	-0.4644314	-0.7061996
	3	0.577346	$5.924798 \cdot 10^{-2}$	-0.3246935	-1.5891441	0.5674493
	4	-0.499848	$-1.432051 \cdot 10^{-2}$	0.1893185	-121.5185	-0.4484073
	5	0.4445297	$-4.55487 \cdot 10^{-3}$	$-7.263595 \cdot 10^{-2}$	3262.095	0.2957803
	6	-0.381935	$1.636784 \cdot 10^{-3}$	$1.707887 \cdot 10^{-2}$	-5256.636	-0.1309140
	7	0.238611	$1.377651 \cdot 10^{-4}$	$-2.418336 \cdot 10^{-3}$	-329073.72	$2.721918 \cdot 10^{-2}$
	8	-	$-1.183038 \cdot 10^{-4}$	$1.906250 \cdot 10^{-4}$	-	-
	9	-	$1.391864 \cdot 10^{-5}$	$-6.451653 \cdot 10^{-6}$	-	-

Anhang G

Modellierung von Metall-Oxid-Silizium- Kontakten

In diesem Kapitel soll kurz auf die Modellierung von Metall-Oxid-Halbleiter Kontakten (MOS-Struktur) eingegangen werden. Formal kann dieser MOS-Kontakt wie ein Schottky-Kontakt behandelt werden. Bei einem Schottky-Kontakt (vgl. Abb. G.1) wird neben der angelegten Spannung nur die Barrierenhöhe Φ_B angegeben. Dahinter steckt die Annahme, dass die Bandkante aufgrund von Oberflächenzuständen gegenüber dem Fermineveau fixiert ist (engl. pinning).

Für MOS-Strukturen werden in der elektrotechnischen Literatur stets zwei Werte angegeben [76]: Die Austrittsarbeit im Metall Φ_M und die Austrittsarbeit im Silizium Φ_S (siehe Abb. G.2). Die Austrittsarbeit Φ_M ist die Energie, die benötigt wird, um ein Elektron von der Fermienergie aus dem Metall herauszulösen, also auf Vakuumniveau zu bringen: $q\Phi_M = E_{vac} - E_{F,M}$. Entsprechend ist für die Austrittsarbeit im Silizium $\Phi_S = E_{vac} - E_{F,S}$. Die Austrittsarbeit ist daher abhängig von der Dotierung. Daneben ist noch die Elektronenaffinität definiert. Sie beschreibt den energetischen Abstand von der Leitungsbandkante zur Vakuumenergie. Es ist χ_i die Elektronenaffinität des Oxids und χ_S die Elektronenaffinität des Siliziums. Diese Größen sind in Abb. G.2 für ein MOS-System, das



Abbildung G.1: Modell für einen Schottky-Kontakt an einer Halbleiter-Metall-Grenzfläche. Im einfachsten Fall ist er durch die Barrierenhöhe Φ_B vollständig bestimmt. Dahinter steckt die Annahme, dass die Bandkante durch Oberflächenzustände gegenüber dem Fermineveau fixiert ist.

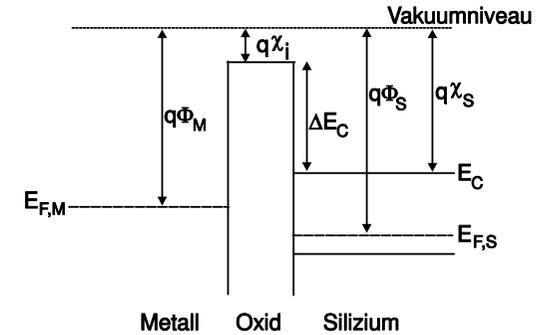


Abbildung G.2: Banddiagramm einer MOS-Struktur, die sich nicht im thermodynamischen Gleichgewicht befindet. Es bezeichnen Φ_M die Austrittsarbeit des Metalls, Φ_S die Austrittsarbeit im Silizium, χ_i die Elektronenaffinität des Oxids, χ_S die Elektronenaffinität des Siliziums, $E_{F,M}$ das Fermineveau des Metalls, $E_{F,S}$ das Fermineveau des Siliziums, E_C die Leitungsbandkante des Siliziums und ΔE_C die Leitungsbandkanten-Diskontinuität an der Grenzfläche vom Silizium zum Oxid.

sich nicht im thermodynamischen Gleichgewicht befindet, eingezeichnet. Befindet sich die MOS-Struktur nun im thermodynamischen Gleichgewicht, so bilden sich Raumladungen aus, aus denen sich das eingebaute Potenzial ergibt (Abb. G.3). Dadurch verschiebt sich die Vakuumenergie entsprechend, das Fermineveau ist nun im gesamten Bauelement konstant. Formal kann diese Randbedingung wie ein Schottky-Kontakt behandelt werden. Es gilt für die Barrierenhöhe Φ_B :

$$\Phi_B = \Phi_M - \chi_i.$$

Weiterhin ist

$$\Delta E_C + q\chi_i = q\chi_S.$$

Daraus folgt:

$$q\Phi_B = q\Phi_M - q\chi_S + \Delta E_C \quad (G.1)$$

Man kann dies auch als Funktion der Differenz der Austrittsarbeiten schreiben:

$$\Phi_B = (\Phi_M - \Phi_S) + (E_C - E_{F,S}) + \Delta E_C \quad (G.2)$$

Für Gl. (G.2) ist zu beachten, dass die Größe Φ_S und der Ausdruck $E_C - E_{F,S}$ von der Dotierung im Silizium abhängig sind, während in Gl.(G.1) nur dotierungunabhängige Materialgrößen eingehen.

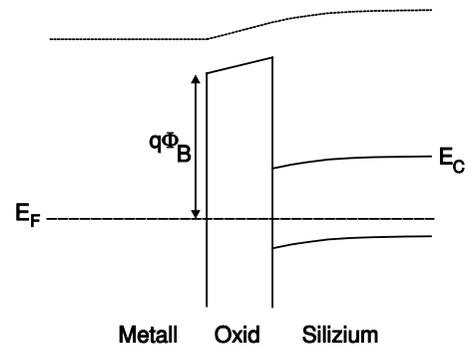


Abbildung G.3: Banddiagramm für die MOS-Struktur im thermodynamischen Gleichgewicht. Durch das eingebaute Potenzial verbiegt sich die Vakuumenergie entsprechend. Als Randbedingung kann man formal einen Schottky-Kontakt mit Barrierenhöhe $q\Phi_B = q\Phi_M - q\chi_S + \Delta E_C$ ansetzen.

Anhang H

Ladungsträgerbeweglichkeit und Rekombination

Der Stromfluss wird im Rahmen der Drift-Diffusions-Näherung beschrieben, die daraus resultierende Stromgleichung wurde in Kapitel 4.2 dargestellt. Darin tritt als Parameter die Ladungsträgerbeweglichkeit auf, welche sämtliche Streuprozesse modelliert. Dieser Abschnitt stellt einige empirische Modelle hierfür vor. Die Rekombinationsraten koppeln die Kontinuitätsgleichung für das Leitungs- und Valenzband. In der Praxis sind drei Mechanismen von Bedeutung, welche hier kurz diskutiert werden.

Ladungsträgerbeweglichkeit

In die Beweglichkeit gehen im Wesentlichen die effektive Masse m^* und die mittlere Streuzeit τ ein:

$$\mu(\vec{r}) = \frac{e\tau(\vec{r})}{m^*(\vec{r})}. \quad (\text{H.1})$$

Die mittlere Streuzeit τ wird durch eine Vielzahl von Streumechanismen bestimmt: akustische und optische Phononen, ionisierte und neutrale Störstellen sowie Streuung der Teilchen untereinander. Dafür gibt es eine Reihe von Modellbildungen, die Abhängigkeiten von Temperatur, Dotierung und Feldstärke beinhalten [9]:

- Dotierungsabhängigkeit:

Die Streuung der Elektronen und Löcher an den ionisierten Donatoren und Akzeptoren kann im Rahmen des Chaughey/Thomas-Modells behandelt werden. Mit zunehmender Dotierkonzentration $N_D + N_A$ (für Donatoren und Akzeptoren) nimmt die Beweglichkeit ab und nähert sich asymptotisch einem Minimalwert an:

$$\mu_{p,n}(N_D + N_A) = \mu_{p,n \min} + \frac{\mu_{p,nD}}{1 + \left(\frac{N_D + N_A}{N_{p,n,ref}}\right)^{\alpha_{p,n}}} \quad (\text{H.2})$$

Die Größen $\mu_{p,n \min}$, $\mu_{p,nD}$, $N_{p,n,ref}$ und $\alpha_{p,n}$ sind Parameter, der Index p/n bedeutet Loch oder Elektron.

- Temperaturabhängigkeit:

Die Streuung an akustischen Phononen bewirkt eine starke Temperaturabhängigkeit der Beweglichkeit, da die Besetzung der Phononen eine Funktion der Temperatur ist:

$$\mu_{p,n}(T) = \mu_{p,n} \left(\frac{T}{T_0}\right)^{-\gamma_{p,n}} \quad (\text{H.3})$$

Die Potenz $\gamma_{p,n}$ ist ein Modellparameter.

- Feldstärkeabhängigkeit:

Bei hohen Feldstärken kommt es durch Emission von optischen Phononen zur Driftsättigung. Dies wird durch folgendes Modell beschrieben:

$$\mu_{p,n}(\vec{E}) = \frac{\mu_{p,n}}{\left[1 + \left(\frac{|\vec{E}|}{E_{p,n}}\right)^{a_{p,n}}\right]^{\beta_{p,n}}} \quad (\text{H.4})$$

Hier sind wieder $\mu_{p,n}$, $E_{p,n}$, $a_{p,n}$ und $\beta_{p,n}$ Modellparameter, \vec{E} ist das elektrische Feld.

Generations- und Rekombinationsraten

Wie in Kapitel 4.2. dargestellt, koppeln die Rekombinationsterme die Transportgleichung für die Elektronen und die Löcher. $R(\vec{r})$ beinhaltet Generations- bzw. Rekombinationsrate und hat positives Vorzeichen, falls die Rekombinationsprozesse überwiegen. Im Folgenden werden die drei wichtigsten Mechanismen vorgestellt und phänomenologische Modelle angegeben [9].

- Shockley-Read-Hall-Rekombination (R_{SRH})

Die Rekombination über eine tiefe Störstelle in der Bandmitte wird durch die SRH-Formel beschrieben:

$$R_{SRH} = \frac{p \cdot n - n_i^2}{\tau_p(n + n_i) + \tau_n(p + n_i)} \quad (\text{H.5})$$

Die Ladungsträgerdichte im thermodynamischen Gleichgewicht ist mit $n_i = n_0 = p_0$ bezeichnet. Im Gleichgewicht verschwindet demnach die Rekombination. Die Streuzeiten $\tau_{n,p}$ hängen von der Störstellen ab. Sie werden mithilfe der halbempirischen Kendall-Beziehung beschrieben:

$$\tau_{p,n}(N_A + N_D) = \frac{\tau_{p,n0}}{1 + \frac{N_D + N_A}{N_{n,p,ref}}} \quad (\text{H.6})$$

- Auger-Rekombination (R_{Aug})

Bei der Auger-Rekombination werden Impuls und Energie des rekombinierenden Ladungsträgers von einem Ladungsträger der gleichen Sorte aufgenommen. Da hier ein drittes Teilchen beteiligt ist, tritt diese Art der Rekombination hauptsächlich in Bereichen mit extrem hoher Ladungsträgerdichte auf. Die empirische Formel für die Auger-Rekombination lautet:

$$R_{Aug} = C_n n(p^2 - n_i^2) + C_p p(n^2 - n_i^2) \quad (\text{H.7})$$

Die Rate skaliert mit der dritten Potenz der Ladungsträgerdichte, da es sich um einen Dreiteilchen-Prozess handelt. C_n und C_p sind als Parameter gegeben. Auch hier gilt, dass die Rate im Gleichgewicht $np = n_i^2$ gegen Null geht.

- Direkte Rekombination (R_{dir})

Die direkte Rekombination von Elektron und Loch erfolgt hauptsächlich über einen strahlenden Übergang, welcher nur bei Halbleitern mit direkter Bandlücke, wie zum Beispiel III-V Materialien, möglich ist. Allerdings gibt es auch bei indirekten Halbleitern die Möglichkeit der Rekombination über ein Phonon. Die Wahrscheinlichkeit hierfür ist jedoch sehr gering:

$$R_{dir} = C(np - n_i^2) \quad (\text{H.8})$$

Der Koeffizient C ist ein Modellparameter.

Anhang I

Bandkantenverlauf in Heterostrukturen

In diesem Abschnitt wird das Modell der 'natürlichen Banddiskontinuitäten' kurz vorgestellt. Die wesentlichen Resultate wurden dabei von Matthias Sabathil in seiner Diplomarbeit beschrieben [51]. Sie sollen hier aber der Vollständigkeit halber und zum besseren Verständnis aufgeführt werden.

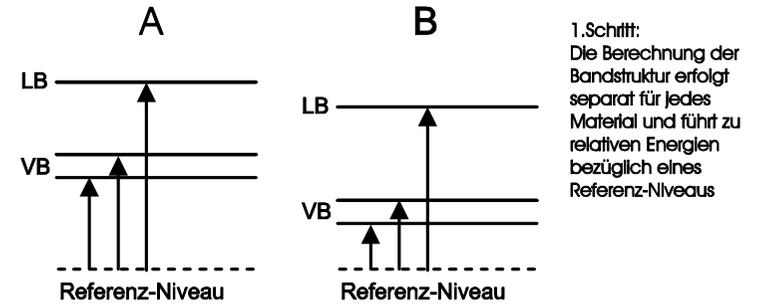
Bei der Behandlung von Heteroübergängen sind zwei Effekte zu beachten:

- Diskontinuitäten der Bandkanten am Grenzübergang zwischen zwei Materialien aufgrund der unterschiedlichen energetischen Lage der Bandkanten relativ zu einem gemeinsamen Referenzniveau.
- Veränderungen in der Bandstruktur durch mechanische Verspannungen.

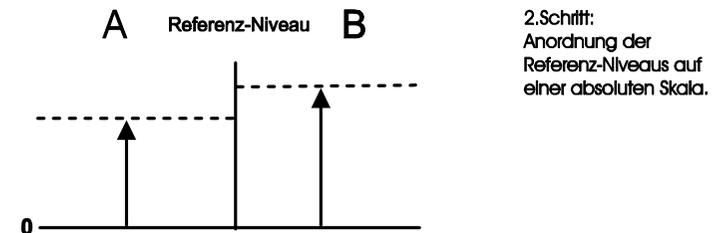
Im Rahmen des Modells der 'natürlichen Banddiskontinuitäten' nimmt man an, dass die Bandenergien verschiedener Materialien auf einer absoluten Energieskala angegeben werden können. Der Energiesprung am Übergang zwischen zwei unverspannten ('natürlichen') Materialien kann daher aus der Differenz der entsprechenden absoluten Energien berechnet werden. Die Verschiebung der Bänder durch Verspannungen wird von den 'natürlichen' Diskontinuitäten als vollkommen abgekoppelt betrachtet und muss separat mittels linearer Deformationspotenziale bestimmt werden.

Das Hauptproblem besteht darin, die Bandenergien auf einem absolutem Niveau anzugeben. Im Wesentlichen geht man dabei folgendermaßen vor:

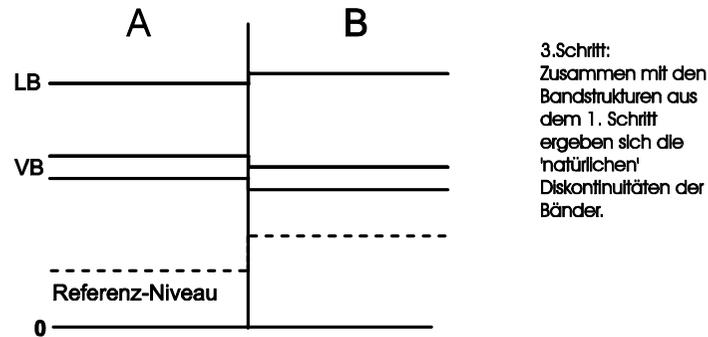
- Zunächst wird die Bandstruktur für die einzelnen Materialien mit herkömmlichen Dichte-Funktional-Rechnungen (DFT) bezüglich eines Referenz-Niveaus bestimmt. Bei diesem Referenz-Niveau handelt es sich entweder um ein gemittelttes Potenzial oder um tief liegende atomare Niveaus.



- Der entscheidende Schritt ist nun die Festlegung dieser Referenz-Niveaus auf einer absoluten Skala. Dies geschieht im Rahmen der 'Model Solid Theory' von van de Walle [43]. Dabei wählt man das mittlere Potenzial als Referenz-Energie. Dazu wird ein Modell-Festkörper konstruiert, der aus neutralen Atomen besteht. Da bei einem neutralen Atom das Potenzial schnell genug auf Null abfällt, erhält man einen absoluten Nullpunkt, der als Bezugspunkt für die errechneten mittleren Potenziale dient. Um die atomaren Potenziale zu erhalten, muss man eine Konfiguration für die Besetzung der s- und p-Orbitale finden, die derjenigen im Kristall ähnlich ist. Dazu verwendet man Daten aus 'tight-binding'-Rechnungen, welche die Bandstruktur aus der Kopplung atomarer Orbitale bestimmen [51].

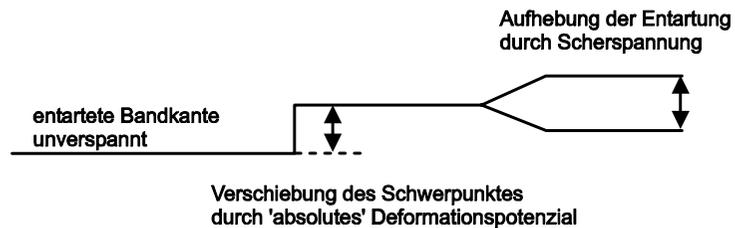


- Eine wichtige Bedingung für die Einführung einer absoluten Energieskala ist die Transitivität der Bandkantenkontinuitäten. Dies bedeutet, dass der Bandkantensprung zwischen zwei bestimmten Materialien konstant ist, unabhängig davon, wieviel andere Materialübergänge dazwischen liegen.



Lineare Deformationspotenziale

Die Verschiebung der Bandkanten unter Verspannung wird durch die linearen Deformationspotenziale beschrieben. Eine Volumendeformation des Kristalls führt über die absoluten Deformationspotenziale zu einer Verschiebung des Schwerpunktes entarteter Bänder. Scherspannungen haben dagegen über die Scher-Deformationspotenziale eine Aufhebung der Entartungen bei gleichzeitigem Festhalten des Schwerpunktes zur Folge.



Der Unterschied zwischen den zwei Arten von Deformationspotenzialen liegt darin, dass die relative Aufspaltung von Bändern sowohl experimentell als auch theoretisch sehr gut zugänglich ist und somit die zugehörigen Deformationspotenziale wohl definiert sind. Die absolute Verschiebung des Schwerpunktes macht sich hingegen nur an den Grenzen zu anderen Materialien bemerkbar und ist deshalb genau wie die 'natürlichen' Banddiskontinuitäten eine schwer zu bestimmende Größe. Die Berechnung der absoluten Deformationspotenziale verläuft daher analog zu der der Banddiskontinuitäten und beruht im Wesentlichen auf vergleichbaren Annahmen. Die Definition der Deformationspotenziale findet sich in Kapitel 2.

Literaturverzeichnis

- [1] A. Kumar, S.E. Laux, und F. Stern, Phys. Rev. B, 42, 5166 (1990); S.E. Laux, in Proc. 5th Int. Conf. on Numerical Analysis of Semiconductor Devices and Integrated Circuits (NASECODE V), edited by J.J.H. Miller (Boole, Dun Laoghaire, Ireland, 2987), p. 270-275.
- [2] M. Grundmann, O. Stier, und D. Bimberg, Phys. Rev. B **59**, 5688 (1999).
- [3] C. Pryor, Phys. Rev. B **57**, 7190 (1998); *ibid.*, 60, 2869 (1999)
- [4] M.A. Cusack, P.R. Briddon, and M. Jaros, Phys. Rev. B **54**, R2300 (1996).
- [5] A. Di Carlo, Proc. 12th Int. Conf. on Nonequilibrium Carrier Dynamics in Semiconductors (Santa Fe, 2001), Physica B (2002) , im Druck
- [6] H. Kosina and S. Selberherr, IEEE Trans. Computer-Aided Design **13**, 201 (1994)
- [7] Seong-Don Kim, et al., IEEE Trans. Electron Devices **49**, 457 (2002)
- [8] W. Fichtner, K. Esmark, W. Stadler, Proceedings of IEDM 2001, Washington, DC, USA
- [9] R. Stenzel, A. Lerch, W. Klix, K.-H. Roach, Proc. of the 11th Conference on Microwaves, Radio Communication and Electromagnetic Compatibility MIOP01, Stuttgart 2001, pp. 268-272.
- [10] G. Paasch and H. Ubensee, Phys. Stat. Sol. (b) **113**, 165 (1982)
- [11] C. M. Wu, E. S. Yang, Solid-State Electronics **22**, 241 (1979)
- [12] K. Kometer, G. Zandler, and P. Vogl. Phys. Rev. B **46**, 1382 (1992)
- [13] F. Chirico, A. Di Carlo, P. Lugli, Phys. Rev. B **64**, 45314 (2001)
- [14] F. Gamiz, M. V. Fischetti, J. Appl. Phys. **89**, 5478 (2001)
- [15] D. K. Ferry, R. Akis, D. Vasileska, Electron Device Meeting, IEDM 2000 Technical Digest Papers, pp. 287-290

- [16] M. V. Fischetti, Phys. Rev. B **59**, 4901 (1999)
- [17] M. V. Fischetti, J. Appl. Phys. **83**, 270 (1998)
- [18] R. Lake, G. Klimeck, R. C. Bowen, D. Jovanovic, J. Appl. Phys. **81**, 7845 (1997)
- [19] G. Dresselhaus, A. F. Kip, and C. Kittel, Phys. Rev. **98**, 368 (1955)
- [20] J. M. Hinckley and J. Singh, Phys. Rev. B **41**, 2912 (1990)
- [21] G. L. Bir and G. E. Pikus, *Symmetry and strain-induced effects in semiconductors*, (John Wiley & sons, New York, 1974)
- [22] G. Bastard, *Wave Mechanics Applied To Semiconductor Heterostructures*, Les Editions de Physique
- [23] C. Herring, E. Vogt, Phys. Rev. **101**, 944 (1956)
- [24] S. L. Chuang and C. S. Chang, Phys. Rev. B **54**, 2491 (1996)
- [25] Francisco Mireles and Sergio E. Ulloa, Phys. Rev. B **62**, 2562 (2000)
- [26] Y. X. Liu, D. Z. Y. Ting, and T. C. McGill, Phys. Rev. B **54**, 5675 (1996)
- [27] Q. G. Zhu and H. Kroemer, Phys. Rev. B **27**, 3519 (1983)
- [28] C. Aversa and J. E. Sipe, Phys. Rev. B **49**, 14542 (1994)
- [29] R. Balian, D. Bessis, and G. A. Mezincescu, Phys. Rev. B **51**, 17624 (1995)
- [30] K. G. Wilson, Phys. Rev. D **10**, 2445 (1974)
- [31] M. Governale and C. Ungarelli, Phys. Rev. B **58**, 7816 (1998)
- [32] F. Schwabl, *Quantenmechanik* (Springer, Berlin, 1998)
- [33] W. Jones and N. H. March, *Theoretical Solid State Physics* (Wiley, 1973), Vol. 2
- [34] W. Kohn and J. M. Luttinger, Phys. Rev. **108**, 590 (1957)
- [35] L. Van Hove, Physica XXI, 517 (1955)
- [36] A. Schenk, *Advanced Physical Models for Silicon Device Simulation* (Springer, Wien, 1998)
- [37] E. Wigner, Phys. Rev. **40**, 749 (1932)
- [38] I. Oppenheim and J. Ross, Phys. Rev. **107**, 28 (1957)

- [39] M. G. Ancona and H. F. Tiersten, Phys. Rev. B **35**, 7959 (1987)
- [40] M. G. Ancona and G. J. Iafrate, Phys. Rev. B **39**, 9536 (1989)
- [41] T. Simlinger, Dissertation, Technische Universität Wien (1996)
- [42] C. Fischer, Dissertation, Technische Universität Wien (1994)
- [43] C. G. van de Walle and R. M. Martin, Phys. Rev. B **35**, 8154 (1987)
- [44] C. G. van de Walle and R. M. Martin, Phys. Rev. Lett. **62**, 2028 (1989)
- [45] A. Qteish and R. J. Needs, Phys. Rev. **45**, 1317 (1992)
- [46] P. Y. Yu and M. Cardona, *Fundamentals of Semiconductors*, (Springer, 1996)
- [47] S. Selberherr, *Analysis and Simulation of Semiconductor Devices* (Springer, Wien, 1984)
- [48] J. Altenbach, H. Altenbach, *Einführung in die Kontinuumsmechanik* (B. G. Teubner Stuttgart 1994)
- [49] C. Pryor, J. Kim, L. W. Wang, A. J. Williamson, A. Zunger, J. Appl. Phys. **83**, 2548 (1998)
- [50] I. Vurgaftman, J. R. Meyer, L. R. Ram-Mohan, J. Appl. Phys. **89**, 5815 (2001)
- [51] M. Sabathil, Diplomarbeit, Technische Universität München (2001)
- [52] R. Wetzler, Diplomarbeit, Technische Universität München (1998)
- [53] M. Manhart, *Numerische Fluidmechanik I*, Vorlesungsskript
- [54] K. W. Boer, *Survey of Semiconductor Physics*, Volume 2 (Van Nostrand Reinhold, 1992)
- [55] M. Goano, Solid-State Electronics **36**, 217 (1993)
- [56] X. Aymerich-Humet, F. Serra-Mestres, J. Millan, Solid-State Electronics **24**, 981 (1981)
- [57] P. Van Halen, D. L. Pulvrey, J. Appl. Phys. **57**, 5271 (1985)
- [58] W. Press, S. Teukolsky, W. Vetterling, B. Flannery, *Numerical Recipes* (Cambridge University Press, 1992)
- [59] *Templates for the Solution of Algebraic Eigenvalue Problems*, edited by Z. Bai, J. Demmel, J. Dongarra, A. Ruhe, H. van der Horst, (SIAM, Philadelphia, 2000)

- [60] T. Kerkhoven, A. T. Galick, U. Ravaoli, J. H. Arends, Y. Saad, J. Appl. Phys. **68**, 3461 (1990)
- [61] A. Trellakis, A. T. Galick, A. Pacelli, U. Ravaoli, J. Appl. Phys. **81**, 7880 (1997)
- [62] G. Dehlinger, L. Diehl, U. Genser, H. Sigg, J. Faist, K. Ensslin, D. Grützmacher, and E. Müller, Science **290**, 2277 (2000)
- [63] R. A. Kaindl, M. Wurm, K. Reimann, M. Woerner, T. Elsaesser, C. Miesner, K. Brunner, and G. Abstreiter, Phys. Rev. Lett. **86**, 1122 (2001)
- [64] K. Reimann, R. A. Kaindl, M. Woerner, Phys. Rev. B **65**, 45302 (2001)
- [65] J. Faist, F. Capasso, C. Sirtori, D. L. Sivco, A. L. Hutchinson, M. S. Hybertsen, and A. Y. Cho, Phys. Rev. Lett. **76**, 411 (1996)
- [66] M. Woerner and T. Elsaesser, Phys. Rev. B **51**, 17490 (1995)
- [67] S. Fukatsu, K. Fujita, H. Yaguchi, Y. Shiraki and R. Ito, Appl. Phys. Lett. **59**, 2103 (1991)
- [68] X. Aymerich-Humet, F. Serra-Mestres, J. Millan, Solid-State Electronics **24**, 981 (1981)
- [69] P. Van Halen, D. L. Pulvrey, J. Appl. Phys. **57**, 5271 (1985)
- [70] Stephen A. Wong, Sean P. McAlister, Solid-State Electronics **37**, 61 (1994)
- [71] N. G. Nilsson, Appl. Phys. Lett. **33**, 653 (1978)
- [72] P. W. Fry, I. E. Itskevich, D. J. Mowbray, et al., Phys. Rev. Lett. **84**, 733 (2000)
- [73] P. N. Brunkov, A. Patane, A. Levin, L. Eaves, und P. C. Main, Phys. Rev. B **65**, 85326 (2002)
- [74] T. Schulz, W. Rösner, L. Risch, T. Aeugle und U. Langman, Proceedings of the Third Workshop on Innovative Circuits and Systems for Nano Electronics (Nano-EI 98), B1/1 (1998)
- [75] K. Kim und J. G. Fossum, IEEE Transactions on Electron Devices **48** (2), 294 (2001)
- [76] S. M. Sze, *Physics of Semiconductor Devices* (John Wiley & Sons, New York, 1981)
- [77] I. E. Itskevich, M. Henini, H. A. Carmona, L. Eaves, and P. C. Main, D. K. Maude and J. C. Portal, Appl. Phys. Lett. **70**, 505 (1997)

- [78] G. Medeiros-Ribeiro, M. V. B. Pinheiro, V. L. Pimentel, E. Marega, Appl. Phys. Lett. **80**, 4229 (2002)
- [79] P. Kozodoy, Y. P. Smorchkova, M. Hansen, H. Xing, S. P. DenBaars, U. K. Mishra, A. W. Saxler, R. Perrin, W. C. Mitchel, Appl. Phys. Lett. **75**, 2444 (1999)
- [80] P. Kozodoy, M. Hansen, S. P. DenBaars, and U. K. Mishra, Appl. Phys. Lett. **74**, 3681 (1999).
- [81] Kumakura and N. Kobayashi, Jpn. J. Appl. Phys. **38**, L1012 (1999).
- [82] M. S. Shur, A. D. Bykhovski, and R. Gaska, Solid State Electronics **44**, 205 (2000)
- [83] L. Hsu and W. Wahukiewicz, Appl. Phys. Lett. **74**, 2405 (1999)
- [84] J. A. Majewski, M. Stadele, and P. Vogl, Mater. Res. Soc. Symp. Proc. **449**, 887 (1997)
- [85] O. Ambacher, J. Appl. Phys. **31**, 2653 (1998)
- [86] S. Hackenbuchner, J. A. Majewski, G. Zandler, O. Ambacher, and P. Vogl, MRS Symp. Proc. **639**, 1023 (2001)
- [87] M. Stutzmann, O. Ambacher, A. Cros, M. S. Brandt, and H. Angerer, Mater. Sci. Eng. B **50**, 212 (1997)
- [88] O. Ambacher, A. Link, S. Hackenbuchner, M. Stutzmann, R. Dimitrov, M. Murphy, J. Smart, J. R. Shealy, B. Green, W. J. Schaff, and L. F. Eastman, Mater. Res. Soc. Symp. Proc. **622** (2000)

Teile dieser Arbeit wurden bereits veröffentlicht:

- *2DEGs and 2DHGs Induced by Spontaneous and Piezoelectric Polarization in AlGaN/GaN Heterostructures*, O. Ambacher, A. Link, S. Hackenbuchner, M. Stutzmann, R. Dimitrov, M. Murphy, J. Smart, J. R. Shealy, B. Green, W. J. Schaff, and L. F. Eastman, MRS Symp. Proc. **622** (2000)
- *Polarization induced 2D hole gas in GaN/AlGaN heterostructures*, S. Hackenbuchner, J. A. Majewski, G. Zandler, O. Ambacher, and P. Vogl, MRS Symp. Proc. **639**, 1023 (2001)
- *Polarization induced 2D hole gas in GaN/AlGaN heterostructures*, S. Hackenbuchner, J. A. Majewski, G. Zandler, O. Ambacher, and P. Vogl, J. Cryst. Growth **230**, 607 (2001)
- *Towards Fully Quantum Mechanical 3D Device Simulations*, M. Sabathil, S. Hackenbuchner, J. A. Majewski, G. Zandler, and P. Vogl, J. Comp. Electronics **1**, 1 (2002)
- *Nonequilibrium band structure of nano-devices*, S. Hackenbuchner, M. Sabathil, J. A. Majewski, G. Zandler, P. Vogl, E. Beham, A. Zrenner, P. Lugli, Physica B **314**, 145 (2002)
- *Midinfrared intersubband electroluminescence of Si/SiGe quantum cascade structures*, I. Bormann, K. Brunner, S. Hackenbuchner, G. Zandler, G. Abstreiter, S. Schmult, and W. Wegscheider, Appl. Phys. Lett. **80**, 2260 (2002)

Danksagung

Zunächst gilt mein besonderer Dank Prof. Dr. P. Vogl für die Vergabe des Themas, die Betreuung der Arbeit und die Anleitung und Förderung, die er mir in den zurückliegenden Jahren zukommen ließ.

Weiterhin danke ich Dr. G. Zandler für seine tatkräftige Unterstützung. Insbesondere hat er durch Programmieren des Eingabesystems den Grundstein für unsere Simulationssoftware gelegt. Herzlicher Dank geht auch an Dr. J. A. Majewski für seine große Hilfsbereitschaft und die zahlreichen konstruktiven Diskussionen. Selbstverständlich danke ich meinen beiden Kollegen Stefan Birner und Matthias Sabathil für die Mitarbeit an unserem Simulationsprogramm. Mein Dank gilt auch den anderen Kollegen am Lehrstuhl, Denis Mamaluy, Christian Strahberger, Uwe Hansen und Thomas Hammerschmidt, für die gute Arbeitsatmosphäre.

Daneben danke ich allen, die sich um die Weiterentwicklung unserer Simulationssoftware verdient gemacht haben. Hier sind vor allem Ralph Oberhuber, Michael Povoltzkyi und Martin Städele zu erwähnen, die bei der Fehlersuche behilflich waren.

Schließlich danke ich den Kollegen von der Experimentalphysik für die gute Zusammenarbeit und die vielen Diskussionen, insbesondere Dr. Ambacher, Ingo Bormann, Dr. Brunner, Evelin Beham, Frank Ertl, Angela Link.