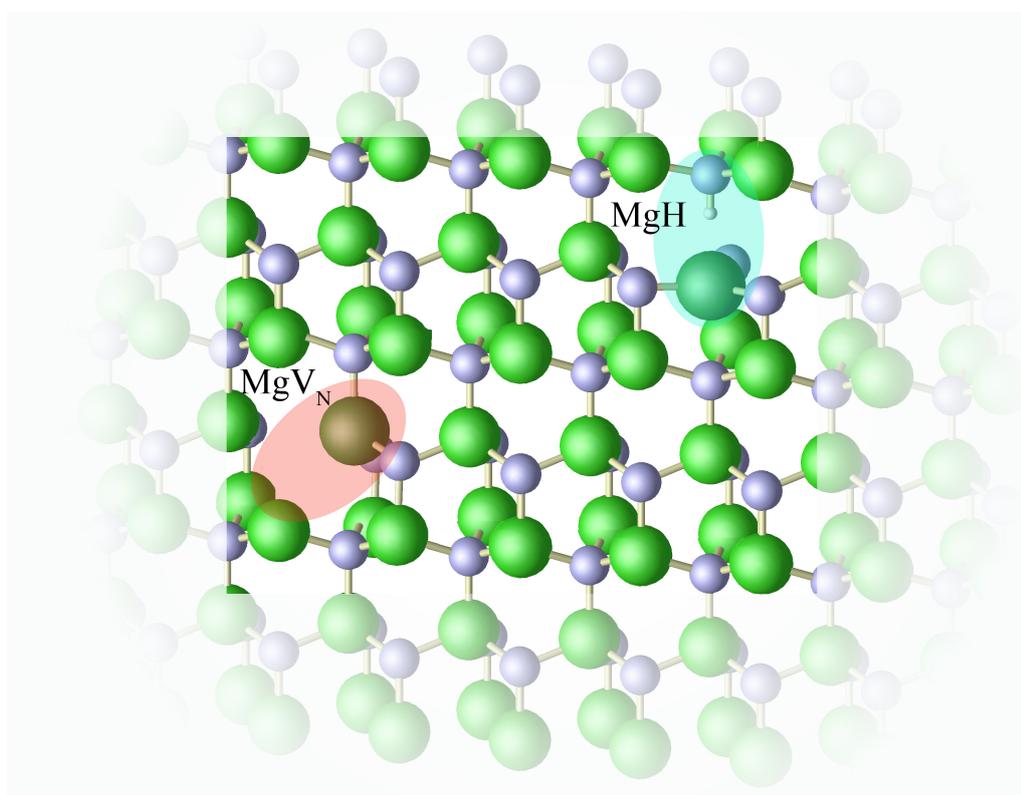


Limitierungen der p -Dotierbarkeit von Galliumnitrid

Eine Defektstudie von GaN:Mg auf Basis der Dichtefunktionaltheorie

Björn Lange, M.Sc.



Dissertation

zur Erlangung des akademischen Grades

DOCTOR RERUM NATURALIUM

vorgelegt der

Universität Gesamthochschule Paderborn

- Department Physik, Fakultät für Naturwissenschaften -

eingereicht von: Björn Lange
Matrikelnummer: 6207289
geboren am: 18.01.1982
in: Steinheim

Tag der Einreichung: 19. Juni 2012
Tag der mündlichen Prüfung: 29. August 2012

Vorsitzender: Prof. Dr. Klaus Lischka
Erstgutachter: Prof. Dr. Jörg Neugebauer
Zweitgutachter: Prof. Dr. Wolf Gero Schmidt
Beisitzer: Dr. Uwe Gerstmann

Inhaltsverzeichnis

1	Motivation	11
2	Zusammenfassende Kapitelübersicht	15
3	Mathematische Vorbetrachtungen	20
3.1	Vektorräume	20
3.2	Funktionale und Diracnotation	22
3.3	Lineare Operatoren	23
3.4	Selbstadjungierte Operatoren	25
4	Dichtefunktionaltheorie - Eine Einführung	27
4.1	Beschreibung von Vielteilchensystemen	27
4.2	Die Born-Oppenheimer Näherung	30
4.3	Dichtefunktionaltheorie	31
4.3.1	Die Hohenberg-Kohn-Theoreme	32
4.3.2	DFT in der Anwendung: Die Kohn-Sham Gleichungen	34
4.3.3	Die lokale Dichtenäherung (LDA)	37
4.3.4	Gradientenkorrekturen und das PBE-Funktional	38
4.3.5	Hartee-Fock Austausch und Hybrid-Funktionale	39
5	Dichtefunktionaltheorie für kristalline Systeme	42
5.1	Der Impulsraum-Formalismus	42
5.1.1	Diskrete Integration über die Brillouinzone	45
5.2	Die Methode der Pseudopotentiale	47
5.2.1	Die Näherung der unveränderlichen Ionen	47
5.2.2	Das Konzept der Pseudopotentiale	48
5.2.3	Die Projector Augmented Wave (PAW) Methode	50
5.3	LCAO Initialisierung	52
5.4	Bildungsenergien von Defekten	53
5.4.1	Superzellenkorrektur geladener Defekte	55
5.4.2	Bestimmung der dielektrischen Konstanten	57
5.5	Berechnung der Defektkonzentration im thermischen Gleichgewicht	59
6	Quantitativ optimierte Atomorbitale — Quamols	62
6.1	Atomzentrierte Orbitalfunktionen	62
6.2	Ein quantitatives Mass für Basisqualität: Die Spillage	64
6.3	Periodische Systeme	67
6.4	Algorithmus	68
6.4.1	Initialisierung der lokalisierten Basis	71
6.4.2	Typische Parameter	72
6.5	Ausgewählte Beispiele	73
6.5.1	Minimale versus erweiterte Basis - Stickstoffdimer	73

6.5.2	Kristalle - Silizium	78
6.5.3	Wahl der Filterfaktoren - Aluminium	84
6.5.4	Transferabilität	86
6.6	Lokalisierung und Tight-Binding - Ein Ausblick	90
6.7	Zusammenfassung und Ausblick	95
7	Ableitung theoretischer Grenzen in der p-Dotierbarkeit von GaN:Mg	97
7.1	Einleitung	97
7.2	Kristallwachstum mittels MOVPE — Die experimentellen Randbedingungen	98
7.3	Den Inversionsdomänengrenzen auf der Spur	100
7.3.1	Magnesiumnitrid in der Antibixbyitstruktur	100
7.3.2	Bestimmung der Konvergenzparameter	102
7.3.3	Intrinsische Punktdefekte	105
7.3.4	Wasserstoffdefekte und Defektkomplexe	109
7.3.5	Mg ₃ N ₂ Einschlüsse in GaN	111
7.3.6	Thermische Austreibung	116
7.3.7	Zusammenfassung	117
7.4	Hohe Mg-Akzeptorkonzentration: Gratwanderung zwischen Kompensation und Phasentrennung	118
7.4.1	Einleitung	118
7.4.2	Defekte in GaN	119
7.4.3	Vergleich der Methoden	122
7.4.4	Einfluss der Superzellenkorrektur	124
7.4.5	Das Bandlückenproblem — LDA/PBE versus HSE	126
7.4.6	Feinjustierung — Die chemischen Potentiale	129
7.4.7	Die Kompensationsmechanismen in p -GaN	131
7.4.8	Kopplung der chemischen Potentiale	133
7.4.9	Zusammenfassung	136
8	Zusammenfassung	140
A	Mg₃N₂ in der idealen Antibixbyitstruktur	144
B	Effektivmassenberechnung	145
C	Pulay Kräfte	148
D	Variation der inversen Überlappmatrix	150
E	Spline-Interpolation	150
F	Spline-Fitten	153

Björn Lange

Limitierungen der p -Dotierbarkeit von Galliumnitrid - Eine Defektstudie von GaN:Mg auf Basis der Dichtefunktionaltheorie

Dissertation, Department Physik, Fakultät für Naturwissenschaften, Universität Paderborn (2012), 166 Seiten, 50 Abbildungen, 30 Tabellen

Kurzfassung

Galliumnitrid (GaN) basierte Verbindungen haben sich zur bevorzugt verwendeten Materialklasse für die Konstruktion lichtemittierender Dioden (LEDs) entwickelt. Grund dafür ist ihre große, einstellbare Bandlücke, welche fast das komplette sichtbare Spektrum abdeckt und somit den Zugang zu Blau- und Weißlicht-LEDs ermöglicht. Eine Verbesserung der p -Leitfähigkeit in dieser Materialklasse würde die Herstellung effizienterer und hellerer LEDs ermöglichen. Die vorliegende Dissertation untersucht basierend auf *ab initio* Methoden die Limitierungen in der GaN p -Dotierbarkeit mittels Magnesium. Zur Untersuchung dieser limitierenden Faktoren werden Defektenergien unter Verwendung der Dichtefunktionaltheorie (DFT) berechnet, welche sich besonders zur Modellierung von Punktdefekten eignet. Ausgedehnte Defekte, wie zum Beispiel die in hoch p -dotiertem GaN experimentell nachgewiesenen Inversionsdomänen (IDs), lassen sich auf Grund ihrer Dimension von einigen nm³ in dieser Theorie nicht mehr effizient modellieren. Hier bilden *coarse-grained* Methoden eine alternative Möglichkeit. In ihrer Genauigkeit hängen diese jedoch von den zu Grunde liegenden atomzentrierten, atomaren Basissätzen ab. Das in dieser Dissertation vorgestellte QUAMOL Konzept generiert atomzentrierte numerische Basisfunktionen auf Grundlage von Ebenen-Wellen DFT-Beschreibungen. An einfachen halbleitenden und metallischen Systemen wird die Erzeugung solcher QUAMOLs demonstriert und ihre Genauigkeit und Transferabilität in weiteren Untersuchungen herausgestellt. Weiter werden die in GaN dominanten Punktdefekte untersucht. Basierend auf den berechneten Bildungsenergien lassen sich Stickstoffvakanzan als potentielle Kompensationszentren in Magnesium dotierten GaN identifizieren. Eine Gradwanderung zwischen einem ausreichenden Stickstoffangebot zur Vermeidung der Vakanzbildung und einem Unterangebot an Stickstoff zur Verhinderung der Phasenseparation Galliumnitrid/Magnesiumnitrid bei hohen Magnesiumkonzentrationen limitiert schlussendlich die p -Dotierbarkeit.

Björn Lange

p-doping limitations of galliumnitride - A defekt study of GaN:Mg based on density-functional theory

PhD Thesis, Department of Physics, Faculty of Science, University of Paderborn (2012)

166 pages, 50 figures, 30 tables

Abstract

Gallium nitride (GaN) based alloys have evolved into the material class of choice for blue and white light-emitting diodes (LEDs). The reason is the large, tunable band gap, which allows light emission of nearly the whole visible spectrum. Improving the *p*-conductivity of GaN would allow for more efficient and brighter LEDs based on this material class. In the present PhD Thesis limitations in the *p*-doping of GaN were investigated. The examination is carried out by calculating defect energetics by means of density-functional theory (DFT), which is the state-of-the-art *ab initio* method for modeling and describing point defects. Large defects, such as the experimentally observed inversion domains (IDs) in *p*-type GaN:Mg, are not feasible to model within the framework of DFT. An alternative are *coarse-grained* methods employing accurate, atom-centered atomic orbitals. However, the construction of accurate and, in particular, transferable atomic orbital basis sets is far away from being trivial. Within this PhD thesis the QUAMOL concept is introduced, which constructs atom-centered, numerical orbitals based on plane-wave DFT calculations. The applicability and performance of the developed approach is demonstrated for semiconducting and metallic test systems, which show that the constructed orbitals are accurate and transferable. Further, the dominant point defects in GaN are studied in detail. Based on calculated formation energies nitrogen vacancies have been identified as possible compensators aside hydrogen in GaN:Mg. A tightrope walk between providing as much nitrogen as needed to avoid vacancy formation and providing as less nitrogen as possible to hinder the phase separation GaN/Mg₃N₂ at high Mg concentrations has finally been identified as the theoretical limitation of *p*-doping in GaN.

Publikationen

- Publikationen mit direktem Bezug zur Dissertation

1. **Theoretical limits of p-doping in GaN:Mg — The malicious interplay of compensation and Mg₃N₂ formation**

Lange B; Freysoldt C.; Neugebauer J; Qimin Y.; Lyons J.L.; Janotti A.; Van de Walle C.

in Vorbereitung

2. **Reliable defect energetics in GaN:Mg**

Lange B; Freysoldt C.; Neugebauer J; Qimin Y.; Lyons J.L.; Janotti A.; Van de Walle C.

in Vorbereitung

3. **Construction and performance of fully numerical optimum atomic basis sets**

Lange B.; Freysoldt C.; Neugebauer J.

PHYSICAL REVIEW B **84**,085101 (2011)

DOI: 10.1103/PhysRevB.84.085101

4. **Native and hydrogen-containing point defects in Mg₃N₂: A density functional theory study**

Lange B; Freysoldt C.; Neugebauer J.

PHYSICAL REVIEW B **81**, 224109 (2010)

DOI: 10.1103/PhysRevB.81.224109

- sonstige Publikationen

1. **Quasiparticle bandoffsets of semiconductor heterojunctions from a generalized marker method**

Mitra C.; Lange B.; Freysoldt C.; Neugebauer J.

PHYSICAL REVIEW B **84**, 193304 (2011)

DOI: 10.1103/PhysRevB.84.193304

2. **Water adsorption on hydrogenated Si(111) surfaces**

Lange B.; Posner R.; Pohl K.; et al.

SURFACE SCIENCE **603**,60-64 (2009)

DOI: 10.1016/j.susc.2008.10.030

3. **Ammonia adsorption on G/Si(001): First-principles calculations**

Lange B.; Schmidt W. G.

SURFACE SCIENCE **602**,1207-1211 (2008)

DOI: 10.1016/j.susc.2008.01.024

Danksagung

Leider lässt sich eine wahrhafte Dankbarkeit mit Worten nicht ausdrücken.

Johann Wolfgang von Goethe, dt. Dichter, 1749-1832

... und dennoch möchte ich es versuchen.

Auf meinem Weg vom frisch gebackenem Master hin zum Doktoranden, der nun mit dieser Schrift versucht seine Doktorwürde zu erlangen, haben mich viele Menschen begleitet, denen ich an dieser Stelle ausdrücklich meinen Dank aussprechen möchte.

Zuallererst möchte ich Professor Doktor Jörg Neugebauer danken. In seinem Department fand ich einen fordernden Arbeitsplatz und konnte mich intensiv mit spannenden Themen aktueller Forschung auseinandersetzen. In vielen Diskussionen half er mir die Schwächen und Stärken meiner Ansätze zu bewerten und diese vor einem breiten wissenschaftlichen Publikum zu diskutieren, um mich so in der Wissenschaft zu etablieren. Sein unglaubliches Wissen und seine Erfahrung waren mir immer wieder ein Leuchtturm, wenn ich mich auf der hohen See der Wissenschaft in den Untiefen der Details verloren hatte.

Meine Dankbarkeit gebührt aber auch im besonderen Doktor Christoph Freysoldt. Er half mir letztendlich bei der Themenwahl, sorgte für die Finanzierung und unterstützte mich als direkter Betreuer. Mit einer unglaublichen Geduld erklärte er mir das S/PHI/nX Projekt, welches ich als methodisches Flaggschiff zu Lösung meiner Fragestellung nutzte, adaptierte und — sofern nötig — erweiterte. Es hat mir viel Spass gemacht, neben meiner Dissertation Teil des S/PHI/nX Projektes werden zu dürfen und an der Erstellung dieser Multiskalenbibliothek von Physikern für Physiker mitzuwirken.

Der eigentliche Urheber des S/PHI/nX Projektes ist Doktor Sixten Boeck, der in einer unglaublich akribisch und peniblen Art und Weise eine Entwicklungsumgebung geschaffen hat, die den Physiker von den Niederungen der Informatik abschirmt. Sein Streben nach Perfektionismus haben mich motiviert den inoffiziellen S/PHI/nX Gedanken — eine Zeile einer Formel entspricht einer Zeile Code — ebenfalls anzustreben: Mit einem unglaublichen Nutzen für die Erstellung gut lesbaren Computercodes gleich im ersten Anlauf.

Für die wunderbare Arbeitsumgebung danke ich der gesamten CM Abteilung, die mir jederzeit für Fragen und Diskussionen zur Verfügung stand und mit denen

man auch fernab der Wissenschaft beim Kanufahren ungezwungen nass werden, oder nach einer gelungen Abfahrt auf weiss-glatter Oberfläche den Alltag Revue passieren lassen konnte. Insbesondere sind hierbei Fritz Körmann, Oliver Marquardt und Ugur Aydin zu nennen, die mir kritisch ihrer Meinungen zu meinen Verstößen wider die deutsche Rechtschreibung kundgetan haben.

Ganz besonders danke ich meiner Spaßfraktion Albert Glensk, Claas Doktor Hueter und Natalie „Penny“ Tillak, die dafür gesorgt haben, dass mein Leben nicht nur aus Wissenschaft bestand, sondern auch aus Experimenten zur Rollbewegung auf nicht ganz planerer Ebene, alchemistischen Übungen zur Nahrungsmittelerzeugung und Augentraining zur visuellen Langzeitbestrahlung.

In diesem Sinnen gilt mein Dank auch Marco Paul Plate und Christine Federschmid, die mit einem unglaublichen Eifer dafür gesorgt haben, dass ich trotz Wissenschaft immer noch eines geblieben bin: Mensch.

Ein ganz besonderer Dank geht an meine Eltern. Sie haben mich immer motiviert meinen eigenen Weg zu gehen und mir die Möglichkeit gegeben, mich frei zu entfalten. Es gibt sicher leichteres als einen angehenden Physiker mit all seinen exzentrischen Ausbrüchen großzuziehen und trotzdem halten sie auch heute noch zu mir. Dafür bin ich ihnen unendlich dankbar.

Teile dieser Arbeit wurden durch das Ministerium für Bildung und Forschung im Rahmen des Projektes 03X0512G (VEKTRA) finanziert, welches ich in meinen Dank mit einschliessen möchte.

... meiner süßen Oma Mia ...
... sowie meinem lieben Freund Roland ...
... wo auch immer ihr nun seid ...

1. Motivation

Es ist absolut möglich, dass jenseits der Wahrnehmung unserer Sinne ungeahnte Welten verborgen sind.

Albert Einstein, Physiker und Nobelpreisträger, 1879-1955

In Folge der Nuklearkatastrophe im Kernkraftwerk Fukushima Daiichi im Frühjahr 2011 hat die Regierung der Bundesrepublik Deutschland den endgültigen Ausstieg aus der Kernenergie bis zum Jahre 2022 beschlossen. Dringender denn je ist die Frage zu beantworten wie Deutschland sich in der globalen Energielandschaft positioniert und wie die gesteckten Ziele des Ausstiegs erreicht werden sollen. Ein Flügel bildet der Ausbau der regenerativen Energien, die den Wegfall der deutschen Atomreaktoren kompensieren sollen. Es ist jedoch ebenso wichtig, auf energiesparende und effizienzsteigernde Maßnahmen zu setzen. In diesem Kontext ist die gestaffelte Abschaffung der traditionellen Glühbirne zu sehen. Eine Alternative zur Glühbirne stellt die lichtemittierende Diode (LED) dar. Solche Elemente besitzen zwar höhere Anschaffungskosten, haben jedoch mit ca. 30-140 lm/Watt eine signifikant höhere Lichtausbeute gegenüber handelsüblichen Glühbirnen mit ungefähr 13-20 lm/Watt. Auch die Lebensdauer von bis zu 45.000 h ist der von handelsüblichen Glühlampen von bis zu 2000 h deutlich überlegen. Die Lebensdauer einer LED hängt jedoch empfindlich von den Umgebungsparametern wie z.B. Temperatur, Feuchtigkeit und — insbesondere auch — von den durchfließenden elektrischen Strömen[1] ab. Während die Temperatur und die Feuchtigkeit durch Kühlkörper und die Kapselung der LED mit korrosionsresistenten Materialien gut in den Griff zu bekommen sind, hängen die notwendigen elektrischen Ströme von der Leitfähigkeit des gewachsenen Halbleitermaterials ab. Das Design geeigneter LED-Materialien mit geringen ohmschen Widerstand ist deswegen in den zentralen Blickpunkt der Forschung gelangt. Geringere Widerstände bedeuten eine geringere Wärmeumwandlung und geringere Betriebsströme. Insofern ist der Bau von niedrig-ohmigen LEDs gerade im Hinblick auf die Effizienz und die Lebensdauer interessant.

Für den Bau von LEDs spielen heutzutage insbesondere Halbleitermaterialien basierend auf dem III-VI Halbleiter Galliumnitrid (GaN) eine große Rolle. Grund hierfür ist die Bandlücke von 3.4 eV dieses Materials, welche durch geeignete Legierung mit den Borgruppenmetallen Aluminium (Al) und Indium (In) so beeinflusst werden kann, dass Lichtemission im gesamten sichtbaren Spektrum möglich ist. Grob gesprochen besteht eine GaN-LED aus einer lichtaktiven Schicht von $\text{Al}_x\text{In}_y\text{Ga}_{1-x-y}\text{N}$, die zwischen n -leitendem und p -leitendem GaN eingebettet ist (vgl. Abbildung 1.1). GaN, welches mittels *Metallorganischer Gasphasenepitaxie* (MOVPE)[2] hergestellt wurde, ist naturgemäss n -leitend[3]. Durch Dotierung mit Silizium (Si) oder Sauerstoff (O) kann die Anzahl der Donatoren erhöht und die Leiteigenschaft weiter verbessert werden[4]. Um p -leitendes GaN herzustellen

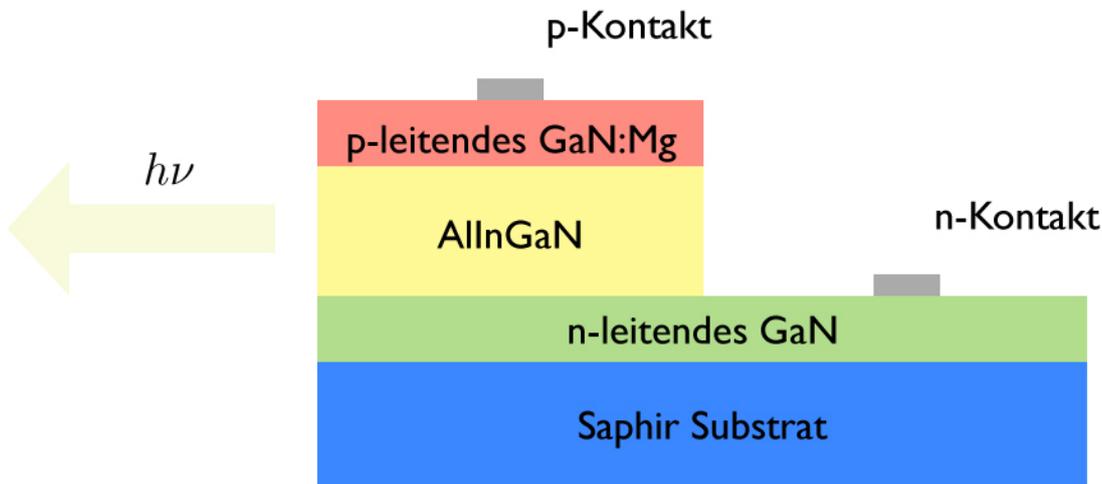


Abb. 1.1:

Schematischer vereinfachter Aufbau einer LED. Auf einem Substrat werden durch Epitaxie drei Domänen von Galliumnitrid erzeugt. Das n-leitende GaN steht mit der Elektrode im Kontakt. Diese Domäne wird dann mit der lichtaktiven Schicht überwachsen. Abgeschlossen wird das Wachstum mit einer p-leitenden GaN Schicht auf der die Kathode angebracht wird.

bedarf es einer Dotierung mittels des Erdalkalimetalls Magnesium (Mg). Dies besitzt ein Valenzelektron weniger als Gallium und fügt sich, anders als die anderen Bormetalle, ins Wirtsgitter ohne drastische Verspannungen zu bewirken. Die Löslichkeit von reinem Mg in GaN ist im Rahmen der MOVPE jedoch gering. Die Molekularstrahlepitaxie (MBE) erlaubt deutlich höhere Magnesiumkonzentrationen, setzt allerdings für den Betrieb Ultrahochvakuum voraus und ist somit für die industrielle Massenproduktion nur bedingt geeignet. Der Durchbruch gelang durch die Erkenntnis, dass in mit MOVPE gewachsenen Proben Magnesium in Form eines neutralen $[\text{Mg}_{\text{Ga}}\text{H}]$ - Komplexes eingebaut wird und in einem weiteren Wärmebehandlungsschritt durch Austreibung des Wasserstoffes aktiviert werden muss.

Dieser Effekt gilt als wohl verstanden [5, 4, 6, 7, 8]: Mg alleine baut sich als negativ geladenes, substitutionelles Mg_{Ga}^- im Kristallgitter ein. Die Ladung dieses Defektes bedingt ladungskompensierende Zentren zur Erreichung der Ladungsneutralität. Durch die Zugabe von Protonen (H^+) werden nun diese kompensierenden Zentren mitgeliefert, was im Rahmen des klassischen Modells den Einbau des Magnesiums durch die Bildung neutraler $[\text{MgH}]$ Komplexe erleichtern sollte. Allerdings können diese Defekte nicht aktiv an der Leiteigenschaft teilnehmen, da das entsprechende Defektniveau in der Bandlücke besetzt ist. Wasserstoff besitzt aber die Eigenschaft schon bei Temperaturen von 500°C - 800°C in diesem Material mobil zu sein, so dass er sich in einem anschließenden Prozessschritt thermisch austreiben lässt [6, 8].

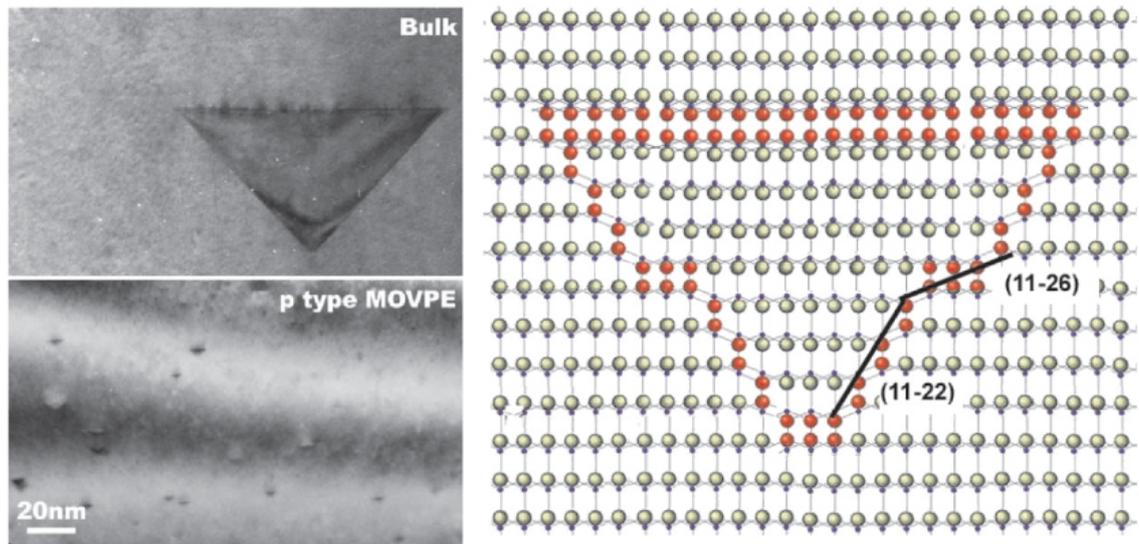


Abb. 1.2:

Die linke Abbildung zeigt die auftauchenden pyramidalen Inversionsdomänen in einer Querschnitts-TEM Aufnahme. Rechts wird ein möglicher atomarer Aufbau der Inversionsdomäne gezeigt. Die braunen Kugeln stellen hierbei Ga-Atome da, die roten Kugeln visualisieren Mg-Atome und in blau werden die Stickstoffatome dargestellt. Die Bilder sind der Referenz[11] entnommen.

Dies bedeutet dann die Aktivierung des zuvor durch den Wasserstoff passivierten Magnesiums.

Neuere Erkenntnisse rütteln jedoch an diesem Bild und zeigen Limitierungen, die sich in diesem Modell nicht erklären lassen: Experimentell beobachtet man bei Magnesiumkonzentrationen oberhalb von 10^{19} cm^{-3} einen Einbruch der relativen Wasserstoffkonzentration. Das Wasserstoff-zu-Magnesium Verhältnisses fällt von ca. 0.7-0.9 auf nur noch 0.2 ab, wie mittels Sekundärionenmassenspektroskopie (SIMS) beobachtet wird[9] (vgl. Abbildung 1.3). Eine anschließende Aktivierung ist nun nicht mehr möglich, die Magnesiumakzeptorniveaus bleiben passiviert und sind für die Erzeugung der p -Leitfähigkeit verloren. Dies deutet auf einen anderen Kompensationsmechanismus als Protonierung hin, um den das bisherige Bild erweitert werden muss. Zudem zeigen experimentell hergestellte, hoch magnesium-dotierte Proben eine deutliche rauere Oberflächenbeschaffenheit. Hochaufgelöste Transmissionselektronenmikroskopie (HR-TEM) zeigt hierbei das Auftreten von Inversionsdomänen (ID) in welcher die GaN-Wachstumsrichtung schlagartig umgekehrt wird (vgl. Abbildung 1.2). Die Grenzen dieser Inversionsdomänen (IDB) werden dabei durch eine Magnesiumnitrid-(Mg_3N_2)-artige Schicht gebildet [10, 11, 12, 13, 14].

Aus experimenteller Sicht allein ist eine Klärung dieses Phänomens schwierig, da sich die Menge des Wasserstoffes in der Probe zwar nachweisen lässt, die genaue

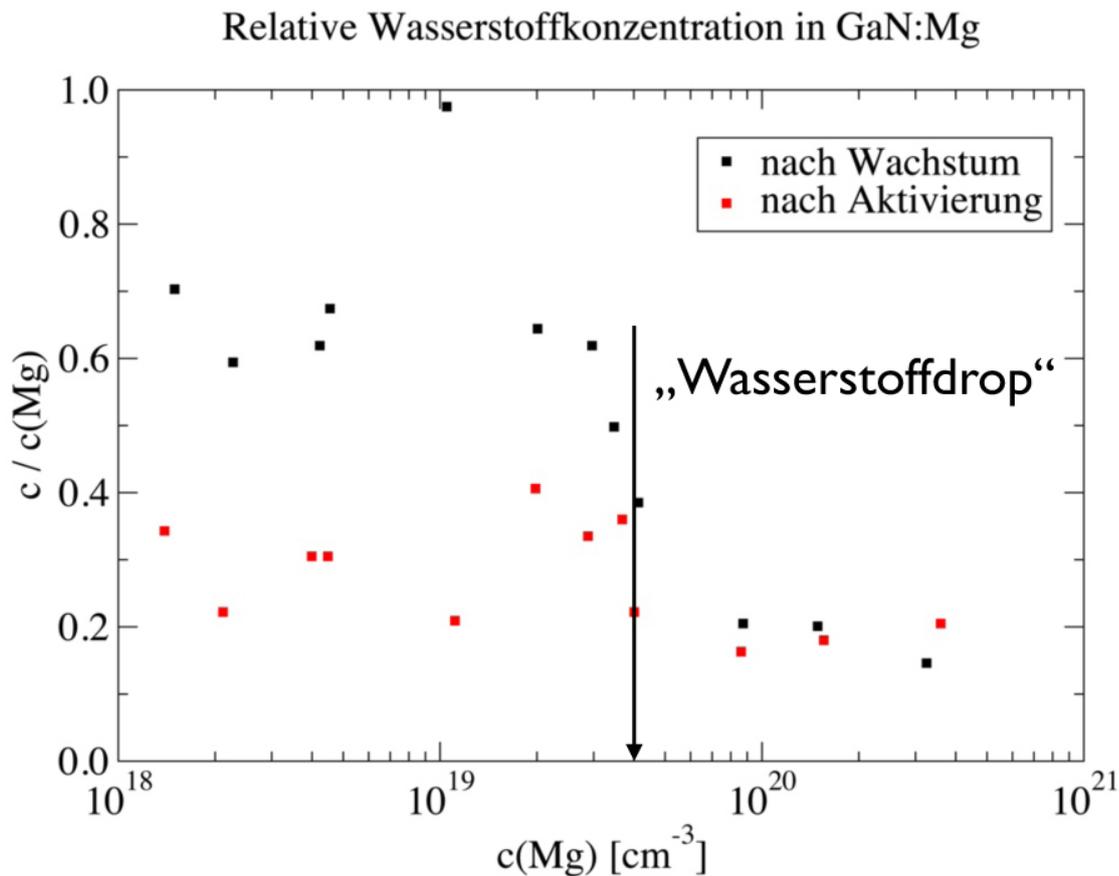


Abb. 1.3:

Relative Wasserstoffkonzentration in Abhängigkeit der Magnesiumkonzentration. Die Daten wurden mittels SIMS-Messungen erhoben[9]. Der schwarze Pfeil zeigt die ungefähre Position des Kompensationswechsels der als „Wasserstoffdrop“ bezeichnet wird.

Position oder Art des Defektes sowie seine Bildungsenergie dem Experimentator aber verschlossen bleiben. Die Klärung dieses Phänomens ist nun zentrale Aufgabe dieser Dissertation. Eine quantentheoretische Betrachtung möglicher Defekte in GaN mittels Dichtefunktionaltheorie (DFT) stellt heutzutage die geeignetste Möglichkeit dar, akkurate Defektenergetiken zu erhalten. Hinter den Mechanismen stehende wichtige Punktdefekte und Defektkomplexe können so identifiziert werden [5, 10, 11, 6, 7, 15, 16, 4]. Dabei ist die DFT als *ab initio* Theorie unabhängig von experimentellen Daten und erlaubt eine alternative Sichtweise auf das Problem, insbesondere auch auf die Facetten, die dem Experimentator verborgen bleiben. Eine Problembetrachtung mittels der DFT ist somit oftmals wegweisend für weitere Experimente und wird in dieser Dissertation verwendet, um die obige Problematik zu analysieren, die gewonnenen Erkenntnisse zu diskutieren und das anerkannte Bild des Wasserstoffkodotierungsmodells einer Revision zu unterziehen.

2. Zusammenfassende Kapitelübersicht

[...] hier stehe ich nun, zusammen mit euch, meine furchtlos lesenden Freunde! Und [...] hier fängt die Geschichte erst richtig an.

Hildegunst von Mythenmetz, Zamonischer Dichter und Buchautor
übersetzt aus dem Zamonischen von Walter Moers, Comic-Zeichner, Illustrator
und Autor, 1957 -

Die vorliegende Dissertation untersucht die Grenzen der p -Dotierbarkeit von Galliumnitrid (GaN) mittels Magnesium (Mg) basierend auf theoretischen *ab initio* Berechnungen. Ausgangspunkt sind hierbei die in Kapitel 1 vorgestellten experimentell beobachteten limitierende Phänomene. Zum einen bilden sich in hoch mit Magnesium dotierten Proben Inversionsdomänen (ID) aus, deren Grenzen aus einer magnesiumnitridartigen Schicht bestehen und die p -Leitfähigkeit durch ihre Polarisationsumkehr stören. Zum anderen zeigen Messungen mittels Sekundärionen-spektroskopie (SIMS), dass die Passivierung der Magnesiumakzeptoren nicht ausschließlich durch das Kodotiermaterial Wasserstoff (H) gegeben ist. Dies ist ein Widerspruch zu dem bisherigen anerkannten Modell, welches die Verbesserung der Magnesiumlöslichkeit in GaN allein über die Bildung von neutralen [MgH]-Defektkomplexen erklärt. Besonders drastisch fällt der Einbruch der relativen Wasserstoffkonzentration ab einer Magnesiumkonzentration von ca. $3 \cdot 10^{19} \text{ cm}^{-3}$ ins Gewicht. Eine thermische Aktivierung des Materials ist nicht mehr möglich.

Gerade der letztere Aspekt ist ein herber Rückschlag, da die Magnesiumakzeptoren die zentrale Rolle in der Verbesserung der p -Leitfähigkeit darstellen. Die Erhöhung der Löslichkeit von Mg in GaN ist nur dann sinnvoll, wenn das zusätzlich eingebrachte Magnesium seine Rolle als Akzeptor wahrnehmen kann. Da diese aber auch ohne Wasserstoff passiviert sind, muss neben dem Wasserstoff ein weiterer Kompensationsmechanismus eine Rolle spielen. Die experimentellen Daten lassen aber eine direkte Bestimmung dieses Mechanismus nicht zu. Deswegen bietet es sich an, auf eine theoretische Modellierung der Problemstellung basierend auf der Dichtefunktionaltheorie (DFT) als *ab initio* Theorie zurückzugreifen. Die aus der Theorie gewonnenen Erkenntnisse über die in p -GaN auftauchenden Defekte helfen die Details im Experiment genauer zu verstehen, um so praktische Anweisungen für die weitere experimentelle Forschung zu geben.

Die in dieser Dissertation verwendeten Modellierungen bedingen geeignete Modifikationen gängiger Algorithmen oder die Entwicklung neuer Ansätze. Dies setzt ein grundlegendes Verständnis der DFT, sowie des verwendeten mathematischen Kalküls voraus. Kapitel 3 stellt die notwendigen mathematischen Konzepte dar, die in der DFT und — besonders auch — für die in dieser Dissertation andiskutierten optimierten Basissätze eine Rolle spielen. Kapitel 4 führt dann in die Konzepte der

Dichtefunktionaltheorie ein, während das Kapitel 5 detailliert die DFT kristalliner Systeme und die Modellierung von Punktdefekten und Defektkomplexen behandelt.

Gängige Programmpakete, wie das weit verbreitete *Vienna ab initio simulation package* (VASP)[17] oder die am Max-Planck-Institut für Eisenforschung GmbH entwickelte Multiskalenbibliothek S/PHI/nX[18], benutzen in ihren Algorithmen der Dichtefunktionaltheorie eine Ebene-Wellen-Basis für die quantenmechanische Wellenfunktion. Dieser Ansatz bedingt eine Beschreibung des Problems im reziproken Raum, hat aber den Charme, dass nur ein einziger Kontrollparameter zur Einstellung der Genauigkeit des Basissatzes benötigt wird. Dies vereinfacht die Modellierung ungemein. Die quantenmechanische Beschreibung von großen Defekten mit mehreren Nanometer Ausdehnung, wie sie beispielsweise die in GaN:Mg auftretende Inversionsdomäne besitzt, ist in diesem Ansatz jedoch sehr zeitintensiv. Typische Algorithmen skalieren kubisch mit der Systemgröße. Eine volle Modellierung der Inversionsdomäne mit mehreren 10.000 bis 100.000 Atomen wäre folglich nur auf *High Performance Computing* Systemen mit entsprechender Parallelisierung denkbar und würde selbst dann Monate bis Jahre an Rechenzeit benötigen.

Wesentlich effizienter sind *coarse grained* Realraummethoden, die mit geeigneten atomzentrierten Basissätzen hantieren. Diese Methoden skalieren zu großen Teilen linear mit den Basisfunktionen und erlauben damit die effiziente Modellierung von Systemen, deren Ausmaße einige Kubiknanometer und mehr erreichen. Die Genauigkeit solcher Methoden stehen jedoch in direkter Verbindung zu dem verwendeten atomzentrierten Basissatz. Dieser bietet — im Gegensatz zu den ebenen Wellen — keine systematisch kontrollierbare Genauigkeit. Vielmehr wird durch Hinzunahme weiterer Funktionen „versucht“ den Basissatz weiter zu komplettieren. Weitere Funktionen bedeuten aber nicht unbedingt einen konsequenten Gewinn an Genauigkeit. Wäre es nun möglich die einfach zu kontrollierende Vollständigkeit einer Ebenen-Wellen-Basis auf eine atomzentrierte Orbitalbasis zu übertragen, wäre eben diese große Schwäche der atomaren Basissätze behoben.

In Kapitel 6 wird daher eine Methode vorgestellt, die aus einer Ebenen-Wellen-Rechnung geeignete atomzentrierte Basisfunktionen, sogenannte Quamols, erzeugt und so optimiert, dass die besetzten Unterräume nahezu identisch sind. Nach einer kurzen Einführung atomzentrierter Basisfunktionen in Kapitel 6.1 wird diese „Nahezu-Identität“ der Unterräume über die *Spillage* in Kapitel 6.2 quantifiziert. Diese gilt es nun zu minimieren. Dabei sollen den Orbitalfunktionen so wenig Einschränkungen wie möglich auferlegt werden. Lediglich die sphärische Symmetrie der Orbitale wird gefordert. Entgegen früherer Arbeiten, in denen atomzentrierte Orbitalfunktionen über Gaußfunktionen oder Besselfunktionen als zu Grunde liegenden Basissatz dargestellt wurden, sind die hier verwendeten Orbitalfunktionen basissatzfrei als Punkte auf einem logarithmischen Radialgitter repräsentiert.

Eine besondere Herausforderung in der Entwicklung des Algorithmus stellt dabei die Berechnung von Überlappintegralen der atomzentrierten Basisfunktionen mit den Zuständen in der Ebenen-Wellen-Repräsentation dar. Dies bedeutet die Fouriertransformation von radialen Ortsraumfunktionen auf das kartesische reziproke Raumgitter der ebenen Wellen. Diese beiden Gitter genügen jedoch völlig unterschiedlichen Ansprüchen. Die sich daraus ergebenden numerischen Schwierigkeiten wurden im Rahmen der Doktorarbeit durch den geschickten Einsatz von Spline-Funktionen kompensiert. Die notwendigen Schritte zum eigentlichen Algorithmus werden in Kapitel 6.3 und Kapitel 6.4 detailliert behandelt. An einigen ausgewählten Beispielen demonstriert Kapitel 6.5 anschließend die Erzeugung der optimierten Basisfunktionen. Diese werden auf ihrer Eigenschaften bezüglich Genauigkeit und Transferabilität untersucht. Die Ergebnisse sind vielversprechend: Es zeigt sich eine qualitativ hochwertige Übereinstimmung mit den Ebenen-Wellen-Resultaten bezüglich der Bandstruktur. Auch physikalische Größen, welche die Berechnung an modifizierten Geometrien benötigen, wie die Gitterkonstante, der Bulkmodulus oder auch Elastizitätskonstanten lassen sich mit diesen an nur einer einzigen Struktur generierten Orbitalen sehr gut beschreiben. Die für verschiedenste Modellierungsaufgaben so grundlegende Transferabilität ist also gewährleistet. Damit ist es gelungen, akkurate und transferable atomzentrierte Basissätze zu erzeugen. Durch die Minimierung der Spillage führt die Hinzunahme weiterer Funktionen zu einer systematisch kontrollierbaren Verbesserung der Genauigkeit.

Mit den Quamols ist der Grundstein zur Entwicklung oder Adaption von effizienten Realraummethoden gelegt. Kapitel 6.6 diskutiert einige Entwicklungen bezüglich der Verwendung von den Quamols in einem Tight-Binding-Modell. Die Herausforderung besteht darin die Basisfunktionen weiter zu lokalisieren, damit die Berücksichtigung von nur einigen wenigen Nachbarn quantitativ gute Ergebnisse liefert. Dies ist der nächste Meilenstein zur Entwicklung eines Algorithmus mit dem auch die Modellierung großer Systeme wie beispielsweise der Inversionsdomäne effizient vollzogen werden kann. Allerdings entfiel in dieser Dissertation die Notwendigkeit einer entsprechenden Entwicklung, da die Ergebnisse in Kapitel 7 eine explizite Behandlung der Inversionsdomänen nicht erforderlich machten. Zur Beantwortung der technologisch wichtigen Frage der p -Leitfähigkeit in GaN findet in Kapitel 7.1 eine Diskussion des Phänomens anhand der Literatur statt. Kapitel 7.2 fasst die experimentellen Randbedingungen beim Kristallwachstum mittels Metallorganischer Gasphasenepitaxie (MOVPE) zusammen. Da die Änderung des Kompensationsmechanismus mit dem Auftauchen der Inversionsdomänen einhergeht, rücken zunächst die magnesiumnitridartigen Inversionsdomänengrenzen in den Fokus der Untersuchung. Es stellt sich die Frage, inwieweit die sich schlagartig ändernde chemische Umgebung Modifikationen der Wasserstoffverteilung im Material bewirken kann. In Kapitel 7.3 erfolgt eine erste Abschätzung über die Modellierung von Magnesiumnitrid in der Antibixbytephase. Dabei werden

erstmalig Bildungsenergien von Punktdefekten und Defektkomplexen in diesem Material mittels DFT bestimmt. Basierend auf dieser Defektenergetik lassen sich Aussagen über die Defektkonzentrationen unter den GaN Wachstumsbedingungen treffen. Die dort erzielten Ergebnisse legen nahe, dass durch die Inversionsdomänen keine signifikante Modifikation der Wasserstoffkonzentration erreicht werden kann. Die geringe Wasserstoffadsorption, gepaart mit dem geringen Konzentrationsanteil der IDB in GaN führen zu Wasserstoffkonzentrationen weit unterhalb der experimentellen Befunde. Es zeigt sich aber, dass der Wasserstoff in diesem Material vornehmlich an Vakanzen bindet, was die Frage motiviert inwiefern Vakanzen in GaN eine Rolle spielen.

In Folge dessen fällt in Kapitel 7.4 der Fokus auf die Defektenergetik von GaN. Da GaN ein in der Literatur viel diskutiertes Material ist, sind schon komplette Datensätze vorhanden, mit denen verglichen werden kann. In Kapitel 7.4.1 zeigen wir diese vergleichende Gegenüberstellung der nicht modifizierten Bildungsenergien einzelner Defekte. Es zeigen sich erhebliche qualitative Unterschiede. Dies motiviert eine detaillierte Analyse der zu Grunde liegenden Berechnung, da eine akkurate Berechnung von Defektkonzentrationen mit diesen Daten fragwürdig erscheint. Die Verlässlichkeit der Aussagen basierend auf diesen Datensätzen steht deswegen auf dem Prüfstand. Nach einer Einführung in die verwendete Defektnomenklatur in Kapitel 7.4.2 findet in den Kapiteln 7.4.3 — 7.4.6 eine Entwicklung von Korrekturtermen statt, welche die Schwächen der unterschiedlichen Austausch- und Korrelationsfunktionale korrigieren und eine Validierung der Datensätze erlauben. Dazu ist es notwendig eine konsequente Korrektur der durch den Superzellenansatz auftretenden langreichweitigen Ladungswechselwirkung durchzuführen. Weiter schlägt das Bandlückenproblem der DFT voll zu Buche, wenn die mit unterschiedlichen Funktionalen berechneten Defektenergien verglichen werden. In dieser Arbeit ist es gelungen das Umladungsniveau der Stickstoffvakanz als funktionalübergreifende Referenz für das Fermi-niveau zu identifizieren, welches unabhängig von den Bandeigenschaften des Materials ist. Damit lassen sich die mit den unterschiedlichen Funktionalen berechneten Ergebnisse in Einklang bringen. Zudem ist durch die Anwendung neuester Entwicklungen der DFT wie des von Heyd, Scuseria und Ernzerhof entwickelte Hybridfunktional HSE und der Superzellenkorrektur für geladene Defekte die Genauigkeit der Defektbildungsenergien signifikant verbessert worden.

Basierend auf der validierten HSE Defektenergetik findet dann in Kapitel 7.4.7 die Berechnung der Defektkonzentrationen für hoch Mg-dotiertes GaN statt. Dabei werden alle Prozesse beleuchtet, die eine Modifikation der Wasserstoffkonzentration bewirken können, um ein komplettes Bild dieses Phänomens zu zeichnen. Es wird aus theoretischer Sicht gezeigt, dass Stickstoffvakanzen in vergleichbaren Konzentrationen zum Wasserstoff auftreten und zudem dessen Kompensationsrolle

übernehmen. Damit ist ein Kompensationsmechanismus identifiziert, der den gewünschte Aktivierungsmechanismus der Wasserstoff austreibung erheblich stört. Stickstoffvakanz sind in GaN nicht mobil und lassen sich im Gegensatz zum Wasserstoff somit nicht thermisch austreiben. Zudem wurde in Kapitel 7.4.8 erstmals die Bildung von Magnesiumnitrid als Grenze für die im Wachstumsprozess beteiligten die chemischen Potentiale diskutiert. Es gelingt in diesem Bild eine qualitative und quantitative Reproduktion des experimentellen Phänomens. Schlussendlich entpuppt sich das Wechselspiel zwischen Stickstoffvakanzbildung und der Phasenseparation GaN/Mg₃N₂ als eigentlich limitierender Faktor der *p*-Dotierbarkeit von Galliumnitrid. Damit sind die zentralen zusätzlichen Effekte aufgedeckt, um welche das ursprüngliche Modell der Wasserstoffkodotierung ergänzt werden muss, um eine Erklärung der experimentellen Befunde zu erlauben.

Abschließend findet in Kapitel 8 eine Zusammenfassung der Ergebnisse statt.

3. Mathematische Vorbetrachtungen

Physik verhält sich zu Mathematik wie Sex zu Selbstbefriedigung.

Richard P. Feynman, Physiker und Nobelpreisträger, 1918-1988

Zur Betrachtung der in der Einleitung diskutierten Fragestellungen bedient sich diese Dissertation aktueller Methoden der Festkörpertheorie, welche allesamt auf der in den 20er Jahren des vorherigen Jahrhunderts entwickelten Quantenmechanik basieren. Einen kompletten Abriss dieser Thematik zu liefern ist in diesem Rahmen nicht möglich, so dass lediglich auf einige wenige Eckpfeiler der Theorie eingegangen wird, um ein Verständnis der in dieser Arbeit getätigten Überlegungen zu ermöglichen. Für eine fundierte Diskussion empfiehlt sich die Lektüre der in den einzelnen Themenkomplexen angegebenen Literatur. Um insbesondere auf die Herleitung der Algorithmen der quantitativ optimierten Orbitalfunktionen vorzubereiten, erfolgt zunächst eine mathematische Vorbetrachtung, die an die hervorragende Einführung von Ballentine[19] angelehnt ist. Diese ist auch für das Verständnis der in Kapitel 4 dargestellten physikalischen Eckpfeiler der Dichtefunktionaltheorie hilfreich.

3.1. Vektorräume

Die Quantenmechanik benutzt in ihrer mathematischen Beschreibung einen speziellen Vektorraum: den Hilbertraum. Im allgemeinen besteht ein Vektorraum V aus einem Satz von Elementen, sogenannten Vektoren ϕ , und ist *geschlossen* bezüglich der Addition und der Multiplikation mit Skalaren c . Das bedeutet:

Definition 3.1 *Seien ϕ_1 und ϕ_2 Elemente des Vektorraumes V und c_1 und c_2 Skalare der komplexen Zahlen \mathbb{C} , so ist ϕ mit*

$$\phi = c_1\phi_1 + c_2\phi_2$$

ebenfalls ein Element des Vektorraumes V .

Aus der Vielzahl von möglichen Vektorräumen sind nun in der Physik zwei von ganz besonderem Interesse. Der Vektorraum der

1. n -dimensionalen diskreten Vektoren, in welchem die Elemente als Spalten komplexer Zahlen a_j repräsentiert werden können, z.B.

$$\phi = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \equiv (a_1, a_2, \dots, a_n)^T.$$

- ausgezeichneten Funktionen, in welchem die Elemente über Funktionen mit bestimmten Eigenschaften, wie zum Beispiel Polynome oder 2-fach differenzierbare Funktionen, dargestellt werden.

Diesen Vektoren können nun weitere Eigenschaften zugeschrieben werden.

Definition 3.2 Eine Teilmenge von Vektoren aus V bildet einen Unterraum U von V , wenn diese Teilmenge geschlossen bezüglich der Addition und der Multiplikation mit Skalaren ist.

Definition 3.3 Ein Satz von Vektoren $\{\phi_n\}$ heißt linear unabhängig, wenn die Gleichung

$$\sum_n c_n \phi_n = 0$$

nur die triviale Lösung $c_n = 0$ für alle n besitzt.

Wenn die obige Gleichung eine weitere Lösung besitzt, spricht man von linear abhängigen Vektoren. Das hat zur Folge, dass ein Vektor eines linear abhängigen Satzes in Form einer Linearkombination der verbleibenden Vektoren ausgedrückt werden kann. Dies ist der Basisgedanke.

Definition 3.4 Ein Satz linear unabhängiger Vektoren, der in der Lage ist jeden Vektor des Vektorraumes abzubilden heißt vollständige Basis des Vektorraumes.

Analog können auch in Unterräumen des Vektorraumes Basen definiert werden. Diese werden mit Bezug auf den Vektorraum als unvollständige Basen bezeichnet.

Weiter wird das *innere Produkt* definiert.

Definition 3.5 Eine Zuordnung $(\phi, \psi) \rightarrow c$ auf den Vektorraum V mit $(\phi, \psi \in V; c \in \mathbb{C})$ heißt inneres Produkt oder skalares Produkt, wenn es die folgenden Eigenschaften erfüllt:

- $(\phi, \psi) = (\psi, \phi)^*$
- $(\phi, c_1\psi_1 + c_2\psi_2) = c_1(\phi, \psi_1) + c_2(\phi, \psi_2)$
- $(\phi, \phi) \geq 0$.

Für die obigen zwei Fälle von Vektorräumen haben die inneren Produkte die folgende Gestalt

- Seien ϕ und ψ Elemente der diskreten Vektoren, dargestellt als Spalten mit den komplexen Zahlen a_1, \dots, a_n , bzw. b_1, \dots, b_n , so gilt

$$(\phi, \psi) = a_1^* b_1 + a_2^* b_2 + \dots + a_n^* b_n,$$

- Seien ϕ und ψ Elemente eines Vektorraumes der Funktionen von x und $w(x)$ eine positiv reellwertige Gewichtungsfunktion, so gilt

$$(\phi, \psi) = \int \phi^*(x) \psi(x) w(x) dx.$$

3.2. Funktionale und Diracnotation

Weiterhin existieren Zuordnungen, welche einem Vektor eines Vektorraumes eine komplexe Zahl zuordnen, die sogenannten *Funktionale*. Man spricht von linearen Funktionalen, wenn für das Funktional F

$$F(c_1\phi_1 + c_2\phi_2) = c_1F(\phi_1) + c_2F(\phi_2) \quad (3.1)$$

gilt. Gilt zudem für die Summe zweier Funktionale

$$(F_1 + F_2)(\phi) = F_1(\phi) + F_2(\phi), \quad (3.2)$$

so bilden die Funktionale ihrerseits einen linearen Raum V' , der *Dualraum* des Vektorraumes der Vektoren ϕ_n genannt wird. Über die Beziehung dieser Funktionale zu den Vektoren, welche letztendlich die berühmte Diracnotation begründet, trifft das Riesz Theorem eine Aussage.

Theorem 3.1 Riesz Theorem: *Es existiert eine eins-zu-eins Entsprechung zwischen linearen Funktionalen F in V' und Vektoren f in V , sodass alle linearen Funktionale als*

$$F(\phi) = (f, \phi) \quad (3.3)$$

dargestellt werden können. Dabei sei f ein fester Vektor und ϕ ein beliebiger Vektor des Raumes V .

Der Beweis des Riesz Theorems geschieht in zwei Schritten. Zunächst ist zu zeigen, dass jedem Vektor f ein Funktional F zuzuordnen ist. Dieses ist aber über die jeweilige Definition des inneren Produktes eines Vektorraumes explizit gegeben und somit trivial erfüllt. Es bleibt folglich noch zu zeigen, dass sich für jedes Funktional F ein entsprechender Vektor f konstruieren lässt. Dazu nehmen wir an, dass die $\{\phi_n\}$ eine orthonormale Basis des Vektorraumes V bilden, also $(\phi_n, \phi_m) = \delta_{n,m}$ gilt. Somit lässt sich jeder beliebige Vektor ψ in dieser Basis ausdrücken als

$$\psi = \sum_n c_n \phi_n.$$

Es folgt für lineare Funktionale nach Gleichung (3.1)

$$F(\psi) = \sum_n c_n F(\phi_n).$$

Wenn man nun den Vektor f als

$$f = \sum_n [F(\phi_n)]^* \phi_n$$

konstruiert, ergibt sich für das skalare Produkt

$$(f, \psi) = \sum_{n,m} F(\phi_n) c_m \delta_{n,m} = \sum_n F(\phi_n) c_n = F(\psi). \quad \text{q.e.d}$$

Die Diracnotation nennt nun Vektoren des linearen Vektorraumes V *ket*-Vektoren und bezeichnet sie mit $|\phi\rangle$. Die linearen Funktionale F des Dualraumes V' werden als *bra*-Vektoren bezeichnet und $\langle F|$ geschrieben. Das komplexwertige Ergebnis eines Funktionals wird nun

$$F(\phi) = \langle F|\phi\rangle \tag{3.4}$$

geschrieben. Gemäß dem Riesz Theorem sind nun das Funktional F und der entsprechende Vektor f gleichwertig, so dass beide mit dem gleichen Buchstaben bezeichnet werden können. Die Verwendung des bra $\langle F|$ oder des ket $|F\rangle$ ermöglicht nun die Unterscheidung, in welchem Raum das Objekt F beheimatet ist. Gemäß Gleichung (3.3) gilt entsprechend

$$\langle F|\phi\rangle = (F, \phi). \tag{3.5}$$

Deshalb wird das *braket* $\langle F|\phi\rangle$ in der Physik formal als weitere Darstellung des inneren Produktes verwendet. Es ist aber zu beachten, dass über das Riesz Theorem eine antilineare Entsprechung

$$c_1^* \langle F| + c_2^* \langle F| \leftrightarrow c_1 |F\rangle + c_2 |F\rangle \tag{3.6}$$

von bra und ket gegeben ist.

3.3. Lineare Operatoren

Neben den Funktionalen gibt es eine Klasse von Objekten, die Vektoren des Vektorraumes V wieder auf Vektoren abbilden: die sogenannten *Operatoren*.

Definition 3.6 *Seien $|\psi\rangle$ und $|\phi\rangle$ Elemente eines Vektorraumes V . Das Objekt \hat{A} mit der Eigenschaft*

$$|\psi\rangle = \hat{A}|\phi\rangle$$

heißt Operator.

Ein Operator heißt *linear* wenn er, angewendet auf eine Linearkombination von Vektoren, gemäß

$$\hat{A}(c_1|\psi_1\rangle + c_2|\psi_2\rangle) = c_1\hat{A}|\psi_1\rangle + c_2\hat{A}|\psi_2\rangle \tag{3.7}$$

zerfällt. Weiter besitzen lineare Operatoren — im folgenden nur noch als Operatoren bezeichnet — die folgenden interessanten Eigenschaften.

Die Gleichheit zweier Operatoren \hat{A} und \hat{B} ist dann gegeben, wenn ihre Wirkung auf beliebige Vektoren $|\psi\rangle$ identisch ist, d.h

$$\hat{A}|\psi\rangle = \hat{B}|\psi\rangle.$$

Dies wird durch die Kurzschreibweise $\hat{A} = \hat{B}$ ausgedrückt. Damit lässt sich das Produkt und die Summe von Operatoren definieren. Es gilt

$$(\hat{A} + \hat{B})|\psi\rangle = \hat{A}|\psi\rangle + \hat{B}|\psi\rangle \quad (3.8)$$

$$\hat{A}\hat{B}|\psi\rangle = \hat{A}(\hat{B}|\psi\rangle) \quad (3.9)$$

Aus dieser Definition folgt unmittelbar, dass die Multiplikation von Operatoren assoziativ ist, $\hat{A}(\hat{B}\hat{C}) = (\hat{A}\hat{B})\hat{C}$. Jedoch besitzt das Kommutativgesetz in der Regel keine Gültigkeit: $\hat{A}\hat{B} \neq \hat{B}\hat{A}$.

Bisher ist nur die Wirkung von Operatoren auf ket-Vektoren diskutiert worden. Ihre Wirkung auf bra-Vektoren wird in der Diracnotation über die Beziehung

$$(\langle\phi|A)|\psi\rangle = \langle\phi|(A|\psi\rangle) \quad (3.10)$$

definiert. Untersuchen wir dieses Definition aus der Sichtweise der lineare Funktionale und verwenden das Riesz Theorem

$$F_\phi(\psi) = (\phi, \psi). \quad (3.11)$$

Die Wirkung des Operators \hat{A} auf das lineare Funktional F_ϕ ist dann analog zu Gleichung (3.1) als

$$\hat{A}F_\phi(\psi) = F_\phi(\hat{A}\psi) \quad (3.12)$$

definiert. Darüber hinaus ist das neue Funktional $\hat{A}F_\phi$ definiert. Folgt man dem Riesz Theorem, muss ein ket-Vektor $|\chi\rangle$ existieren mit

$$\hat{A}F_\phi(\psi) = (\chi, \psi) = F_\chi(\psi). \quad (3.13)$$

Da nun aber $|\chi\rangle$ im selben Vektorraum wie $|\phi\rangle$ beheimatet ist, muss ein Operator \hat{A}^\dagger existieren, so dass $|\chi\rangle = \hat{A}^\dagger|\phi\rangle$ gilt. Damit folgt

$$\hat{A}F_\phi = F_{\hat{A}^\dagger\phi}$$

und weiter

$$(\hat{A}^\dagger\phi, \psi) = (\phi, \hat{A}\psi).$$

Letzteres ist aber gerade die Definition der Adjunktion eines Operators über das skalare Produkt, welche sich hier ganz natürlich aus der Diracnotation ergibt. Dies

hat zur Folge, dass der korrespondierende bra-Vektor $\langle \omega |$ des ket-Vektors $|\omega\rangle \equiv \hat{A}|\phi\rangle$ als $\langle \omega | \equiv \langle \phi | \hat{A}^\dagger$ gegeben ist. Verwendet man nun, dass $\langle \omega | \psi \rangle^* = \langle \psi | \omega \rangle$ folgt hieraus

$$\langle \phi | \hat{A}^\dagger | \psi \rangle^* = \langle \psi | \hat{A} | \phi \rangle. \quad (3.14)$$

Zusätzlich zum inneren Produkt $\langle \phi | \psi \rangle$ kann nun ein äußeres Produkt $|\psi\rangle\langle\phi|$ definiert werden, welches seinerseits einen Operator darstellt. Dies wird klar, wenn man die Wirkung dieses Objektes auf einen ket-Vektor betrachtet. Mittels Assoziativgesetz folgt

$$(|\psi\rangle\langle\phi|)|\lambda\rangle = \langle\psi|(\underbrace{\langle\phi|\lambda\rangle}_{c \in \mathbb{C}}) = c|\psi\rangle,$$

was eine Abbildung des Vektors $|\lambda\rangle$ auf $c|\psi\rangle$ bedeutet. Diese Konstrukte werden uns in Kapitel 6 als Projektoren wieder begegnen. Unter Verwendung von Gleichung (3.14) ergibt sich aus $\langle\lambda_2|\psi\rangle\langle\phi|\lambda_1\rangle = (\langle\lambda_1|\phi\rangle\langle\psi|\lambda_2\rangle)^* = \langle\lambda_2|(|\phi\rangle\langle\psi|)^\dagger|\lambda_1\rangle$ die Identifizierung

$$(|\phi\rangle\langle\psi|)^\dagger = |\psi\rangle\langle\phi|. \quad (3.15)$$

3.4. Selbstadjungierte Operatoren

Erfüllt ein Operator die Beziehung

$$\langle\psi|\hat{A}|\phi\rangle = \langle\phi|\hat{A}|\psi\rangle^* \quad (3.16)$$

so heißt er selbstadjungiert. Selbstadjungierte Operatoren nehmen in der Physik eine besondere Rolle ein. Durch diese werden messbare Observablen, wie zum Beispiel die Energie eines Systems, der Drehimpuls oder der Ort eines Teilchens repräsentiert.

Theorem 3.2 *Das Eigenwertspektrum eines selbstadjungierten Operators ist rein reell:*

$$\hat{A}|\alpha\rangle = a|\alpha\rangle \text{ mit } a \in \mathbb{R}$$

Der Beweis erfolgt direkt unter Verwendung der Definition selbstadjungierter Operatoren. Sei $|\alpha\rangle$ ein beliebiger Eigenvektor von \hat{A} zu dem Eigenwert a . Dann gilt die Folgerung:

$$a\langle\alpha|\alpha\rangle = \langle\alpha|a\alpha\rangle = \langle\alpha|\hat{A}|\alpha\rangle \quad (3.17)$$

$$= \langle\alpha|\hat{A}|\alpha\rangle^* = \langle\alpha|a\alpha\rangle^* \quad (3.18)$$

$$= a^*\langle\alpha|\alpha\rangle \quad (3.19)$$

Daraus folgt sofort die Identität $a = a^*$, was nur gültig ist für $a \in \mathbb{R}$.

Weiter gilt für selbstadjungierte Operatoren, dass Eigenvektoren zu unterschiedlichen Eigenwerten orthogonal aufeinander stehen.

Theorem 3.3 *Seien $|\alpha_1\rangle$ und $|\alpha_2\rangle$ Eigenvektoren des selbstadjungierten Operators \hat{A} mit den Eigenwerten a_1 und a_2 . Dann gilt*

$$\langle \alpha_2 | \alpha_1 \rangle = 0 \text{ f\u00fcr } a_1 \neq a_2$$

Auch hier verwendet man die Definition selbstadjungierter Operatoren. Aus

$$0 = \langle \alpha_2 | \hat{A} | \alpha_1 \rangle - \langle \alpha_1 | \hat{A} | \alpha_2 \rangle^* \tag{3.20}$$

$$= a_1 \langle \alpha_2 | \alpha_1 \rangle - a_2 \langle \alpha_1 | \alpha_2 \rangle^* \tag{3.21}$$

$$= (a_1 - a_2) \langle \alpha_2 | \alpha_1 \rangle \tag{3.22}$$

folgt unmittelbar die Aussage des Theorems.

Diese mathematischen Vorbetrachtungen bilden den Grundstock zum Verst\u00e4ndnis der mathematischen Formulierung der Quantenmechanik. Auf dieser fusst die nun im Folgenden beschriebene Dichtefunktionaltheorie.

4. Dichtefunktionaltheorie - Eine Einführung

Elektronen im Bohrschen Atommodell sind stehende Materiewellen. Um das zu verstehen, stellen Sie sich einen Elektrozaun vor, nur ohne Draht und Pfosten.

Vince Ebert, dt. Physiker und Kabarettist, 1968 -

4.1. Beschreibung von Vielteilchensystemen

Die Modellierung von Molekülen und Festkörpern — und insbesondere von Defekten — ist zentraler Bestandteil aktueller Forschung und zur Beantwortung der zentralen Fragestellung dieser Dissertation unabdingbar. Besonders die Entwicklung der *Dichtefunktionaltheorie* (DFT) in den 60er Jahren und ihre kontinuierlichen Verbesserungen bis in unsere heutige Zeit haben dafür gesorgt, dass die moderne theoretische Festkörperphysik heutzutage einen immensen Stellenwert hat. Nur sie erlaubt es effiziente und quantitativ akkurate Modelle von Punktdefekten zu entwickeln und damit für die Experimentatoren wegweisende Ergebnisse zu produzieren. Im Folgenden werden die Konzepte dieser mächtigen Theorie erörtert und dargestellt.

Atomare Ensemble können als ein Satz positiv geladener Kerne verstanden werden, welche ihrerseits von Elektronen umgeben sind. Die bindende Kraft zwischen Kernen und Elektronen ist die Coulombkraft F_{Coulomb} , hervorgerufen durch das entsprechende Coulombpotential V_{Coulomb} . Für Punktladungen q_i, q_j im Vakuum, die sich im Abstand $r_{i,j}$ voneinander befinden, ist dieses über die Beziehung

$$V_{\text{Coulomb}}(r_{i,j}) = \frac{1}{4\pi\epsilon_0} \frac{q_i q_j}{r_{i,j}} \quad (4.1)$$

gegeben. Aus der Kenntnis des Potentials und den damit verbundenen Kräften lässt sich nun klassisch die Dynamik des Systems über die Differentialgleichung beschreiben, die empirisch von Sir Isaac Newton bestimmt wurde und als zweites Newtonsches Axiom Berühmtheit erlangt hat. Im Falle zeitlich unveränderlicher Massen lässt sich diese als

$$\mathbf{F} = \frac{d\mathbf{p}}{dt} = m \frac{d^2\mathbf{r}}{dt^2} \quad (4.2)$$

schreiben. Für die Beschreibung klassischer Partikel, die sich in unserer alltäglichen Größenordnung bewegen, leistet diese Gleichung hervorragende Ergebnisse. Allerdings versagt dieser Ansatz bei der Beschreibung des Mikrokosmos. Zu Beginn des 20. Jahrhunderts gelang es Erwin Schrödinger eine Gleichung aufzustellen, welche das Verhalten von quantenmechanischen Partikeln, zu denen die Elektronen gehören, korrekt beschreibt. Dazu war es notwendig den Begriff des Teilchens fallenzulassen und quantenmechanische Objekte mittels einer komplexwertigen

Wellenfunktion Ψ zu beschreiben. Eine anschauliche Bedeutung hat hierbei nur das Betragsquadrat dieser Wellenfunktion, welches als Aufenthaltswahrscheinlichkeit des quantenmechanischen Objektes am Ort \mathbf{r} aufgefasst wird.

Die Dynamik dieser Wellenfunktion wird durch die Schrödingergleichung

$$\hat{H}(\mathbf{r}, t)\Psi(\mathbf{r}, t) = i\hbar \frac{\partial \Psi(\mathbf{r}, t)}{\partial t} \quad (4.3)$$

beschrieben. Die zeitliche Propagation der Wellenfunktion beschreibt der sogenannte *Hamiltonoperator* $\hat{H}(\mathbf{r}, t)$, in welchem nun Ausdrücke für die kinetische Energie und die Potentiale stecken. In vielen Fällen hängen die Potentiale nicht explizit von der Zeit ab, so dass der Hamiltonoperator $\hat{H}(\mathbf{r}, t) \equiv \hat{H}(\mathbf{r})$ zeitunabhängig angesetzt werden kann. In diesen Situationen wählt man die Wellenfunktion als Produktfunktion $\Psi(\mathbf{r}, t) = \psi(\mathbf{r})T(t)$. Dies führt zur sogenannten zeitunabhängigen Schrödingergleichung

$$\hat{H}\psi(\mathbf{r}) = E\psi(\mathbf{r}) \quad (4.4)$$

$$\Psi(\mathbf{r}, t) = \psi(\mathbf{r})e^{-iEt/\hbar}. \quad (4.5)$$

Das Auffinden von Lösungen der zeitunabhängigen Schrödingergleichung ist eine zentrale Aufgabe der Quantenmechanik. Für ein einziges Teilchen ist der Hamiltonoperator \hat{H} die Summe der Operatoren der kinetischen Energie \hat{T} und Energie des Teilchens im externen Potential V . In der Ortsraumdarstellung ist der Operator der kinetischen Energie als

$$\hat{T} = -\frac{\hbar^2}{2m}\Delta \quad (4.6)$$

gegeben. Dabei bezeichnet $\Delta = \nabla \cdot \nabla$ den Laplace Operator. Die potentielle Energie, welche aufgrund der obigen Separabilität der Wellenfunktion als zeitunabhängig anzunehmen ist, ist eine Funktion des Ortes $\hat{\mathbf{r}}$:

$$\hat{H} = \hat{T} + V(\mathbf{r}) \quad (4.7)$$

$$= -\frac{\hbar^2}{2m}\Delta + V(\mathbf{r}) \quad (4.8)$$

Für Fragestellungen in der Festkörperphysik gilt es primär eine Vielzahl von Teilchen zu modellieren, welche ihrerseits miteinander durch die Coulombkraft wechselwirken. Deshalb betrachten wir an dieser Stelle die Schrödingergleichung von N durch die Coulombkraft wechselwirkenden Elektronen in einem externen Potential. Die Wellenfunktion des Systems ist damit eine Vielteilchenwellenfunktion der N Elektronen. Die Wechselwirkung der Elektronen untereinander wird durch das Coulombpotential

$$V_{\text{Coulomb}} = \frac{1}{4\pi\epsilon_0} \sum_{i,j < i}^N \frac{q^2}{|\mathbf{r}_i - \mathbf{r}_j|} \quad (4.9)$$

beschrieben und ist ein Zweiteilchenpotential. Dies hat gerade im Falle von Elektronen einen interessanten Effekt zur Folge. Auf Grund des Pauli-Prinzips muss die Wellenfunktion eines Systems von N Elektronen antisymmetrisch bezüglich der Vertauschung zweier Koordinaten

$$\Psi(r_1, \dots, r_j, \dots, r_k, \dots, r_N) = -\Psi(r_1, \dots, r_k, \dots, r_j, \dots, r_N) \quad (4.10)$$

sein. Nur so ist gewährleistet, dass keine zwei Elektronen in all ihren Quantenzahlen übereinstimmen. Der Einfachhalt halber beschränkt man sich im Folgenden auf ein System von zwei Elektronen. In der Hartree-Fock Theorie wird diese antisymmetrische Gesamtwellenfunktion Ψ anhand einer Slaterdeterminante

$$\Psi(r_1, r_2) = \begin{vmatrix} \psi_1(r_1) & \psi_2(r_1) \\ \psi_1(r_2) & \psi_2(r_2) \end{vmatrix} = \psi_1(r_1)\psi_2(r_2) - \psi_1(r_2)\psi_2(r_1). \quad (4.11)$$

bestehend aus Einteilchenwellenfunktionen ψ konstruiert. Die Couloumbenergie des Systems beinhaltet dann die folgenden Energieausdrücke. Zum einem die sogenannte *Hartree-Energie* E_H , wo selbe Orbitale am selben Ort lokalisiert sind

$$E_H = \left\langle \psi_1(r_1)\psi_2(r_2) \left| \frac{1}{r_1 - r_2} \right| \psi_1(r_1)\psi_2(r_2) \right\rangle + \left\langle \psi_1(r_2)\psi_2(r_1) \left| \frac{1}{r_1 - r_2} \right| \psi_1(r_2)\psi_2(r_1) \right\rangle \quad (4.12)$$

als auch solche Beiträge, wo der Ort bei selben Orbitalen ausgetauscht ist

$$E_X = - \left\langle \psi_1(r_1)\psi_2(r_2) \left| \frac{1}{r_1 - r_2} \right| \psi_1(r_2)\psi_2(r_1) \right\rangle - \left\langle \psi_1(r_2)\psi_2(r_1) \left| \frac{1}{r_1 - r_2} \right| \psi_1(r_1)\psi_2(r_2) \right\rangle. \quad (4.13)$$

Dieser letzte Energiebeitrag wird als *Austauschenergie* E_X bezeichnet. Die Näherung der Vielteilchenwellenfunktion über eine einzige Slaterdeterminante ist in der Regel zu grob, so dass in der Hartree-Fock Theorie die Summe aus Hartree-Energie und Austausch-Energie ca. 99% der tatsächlichen Vielteilchen-Coulombenergie ausmachen[20]. Diese Energiedifferenz wird als *Korrelationsenergie* E_C bezeichnet und ist in ihrer analytischen Form nicht bekannt. Das Auffinden entsprechender Funktionale für die Austausch- und Korrelationsenergie ist auch heute noch eine zentrale Fragestellung in der DFT. An späterer Stelle werden die gängigen Näherungen für diese diskutiert.

Um nun aber zur Beschreibung von Atomen, Molekülen und Festkörpern zu kommen, müssen die Kerne in der Gleichung berücksichtigt werden. Die Kerne sind aus Protonen und Neutronen aufgebaut, wobei jedes dieser Bestandteile ungefähr um den Faktor 1800 schwerer ist, als ein einzelnes Elektron. Born und Oppenheimer haben gezeigt, dass die Gesamtenergie von Molekülen in Form einer Entwicklung nach Potenzen von $\sqrt[4]{m_{\text{Elektronen}}/M_{\text{Kerne}}}$ geschrieben werden kann[21]. Energiebeiträge basierend auf einer gekoppelten Elektronen-Kern-Bewegung tauchen erst in vierter Ordnung und höher in dieser Entwicklung auf, so dass diese gegenüber den anderen Energieausdrücken basierend auf einer reinen Kern- oder Elektronenbewegung vernachlässigt werden können. Erst im Falle von speziellen Randbedingungen, wie sie z.B. bei der Tieftemperatursupraleitung auftreten, sind die massenskalierten Energiegrößen vergleichbar und eine Vernachlässigung nicht gerechtfertigt. Für die in dieser Dissertation betrachteten Randbedingungen ist diese Näherung aber statthaft. Daraus resultiert eine Entkopplung der Kernbewegung von der Bewegung der Elektronen im Rahmen der im folgenden diskutierten Born-Oppenheimer Näherung.

4.2. Die Born-Oppenheimer Näherung

Zerlegt man den obigen Hamiltonoperator explizit in seine Kern- und seine Elektronenanteile, so besitzt er die folgende Gestalt:

$$\hat{H} = \hat{T} + \hat{V} = \hat{T}_{\text{K}} + \hat{T}_{\text{e}} + \hat{V}_{\text{KK}} + \hat{V}_{\text{eK}} + \hat{V}_{\text{ee}}. \quad (4.14)$$

Es bezeichnen hierbei \hat{T}_{K} und \hat{T}_{e} die Operatoren der kinetischen Energie für Kerne und Elektronen und \hat{V}_{KK} , \hat{V}_{eK} und \hat{V}_{ee} die Operatoren der potentielle Energien der Wechselwirkungen Kern-Kern, Elektron-Kern und Elektron-Elektron. Der elektronische Hamiltonoperator \hat{H}_0

$$\hat{H}_0 = \hat{H} - \hat{T}_{\text{K}} - \hat{V}_{\text{KK}} = \hat{T}_{\text{e}} + \hat{V}_{\text{ee}} + \hat{V}_{\text{eK}} \quad (4.15)$$

ist nun Bestandteil weiterer Betrachtungen. In diesen fließen die Kernkoordinaten nur noch als Parameter ein, so dass eine Unterscheidung zwischen elektronischen Koordinaten und Kernkoordinaten sinnig ist. Dazu sei an dieser Stelle $\mathbf{x} = \{\mathbf{r}_i\}_{i=1}^{N_{\text{e}}}$ und $\mathbf{X} = \{\mathbf{R}_k\}_{k=1}^{N_{\text{K}}}$ definiert. N_{e} bezeichnet hierbei die Anzahl der Elektronen, N_{K} die Anzahl der Kerne. Das rein elektronische Teilproblem lautet dann

$$\hat{H}_0(\mathbf{x}; \mathbf{X})\psi_i(\mathbf{x}; \mathbf{X}) = \epsilon_i(\mathbf{X})\psi_i(\mathbf{x}; \mathbf{X}) \quad (4.16)$$

mit

$$\langle \psi_i | \psi_j \rangle = \delta_{ij}. \quad (4.17)$$

Unter der Annahme, dass dieses elektronische Problem bereits gelöst ist, bilden die ψ_i eine Basis des Hilbertraumes. Die Gesamtwellenfunktion des Systems lässt sich dann gemäß

$$\psi_{\text{ges}}(\mathbf{x}, \mathbf{X}) = \sum_i \chi_i(\mathbf{X}) \psi_i(\mathbf{x}; \mathbf{X}) \quad (4.18)$$

entwickeln. Setzen wir dies in die obige Hamiltongleichung des Gesamtsystems ein und bilden das Produkt mit $\langle \psi_j |$ gewinnen wir als Schrödingergleichung

$$\langle \psi_j | \hat{H} | \psi_{\text{ges}} \rangle = E_{\text{ges}} \langle \psi_j | \psi_{\text{ges}} \rangle \quad (4.19)$$

$$\iff \sum_i \chi_i(\mathbf{X}) \langle \psi_j | \hat{H}_0 + \hat{T}_K + \hat{V}_{\text{KK}} | \psi_i \rangle = E_{\text{ges}} \sum_i \chi_i(\mathbf{X}) \langle \psi_j | \psi_i \rangle \quad (4.20)$$

$$\iff \left(\epsilon_j(\mathbf{X}) + \sum_i \langle \psi_j | \hat{T}_K | \psi_i \rangle + \hat{T}_K + \hat{V}_{\text{KK}} \right) \chi_j(\mathbf{X}) = E_{\text{ges}} \chi_j(\mathbf{X}). \quad (4.21)$$

An dieser Stelle greift nun die aus[21] motivierte Vernachlässigung der Kopplungsterme $\langle \psi_j | \hat{T}_K | \psi_i \rangle$ gegenüber den anderen Ausdrücken. Die sich daraus ergebene Näherung

$$(\epsilon_j(\mathbf{X}) + T_K + V_{\text{KK}}) \chi_j(\mathbf{X}) \approx E_{\text{ges}} \chi_j(\mathbf{X}) \quad (4.22)$$

ist die sogenannte adiabatische Näherung oder Born-Oppenheimer Näherung[21]. Hier entkoppeln Kernbewegung und Elektronenbewegung vollständig und es gilt das Problem des rein elektronischen Hamiltonoperators \hat{H}_0 zu lösen. Dieser Operator beinhaltet die kinetische Energie der Elektronen, sowie die Potentiale der Coulombinteraktion der Elektron-Elektron Wechselwirkung und der Kern-Elektron Wechselwirkung und besitzt die Gestalt:

$$\hat{H}_0 = \underbrace{\sum_i^N -\frac{\hbar^2}{2m} \Delta_i}_{T_{\text{ee}}} + \frac{1}{4\pi\epsilon_0} \left(\underbrace{\sum_{i,j < i}^n \frac{q^2}{|\mathbf{r}_i - \mathbf{r}_j|}}_{V_{\text{ee}}} - \sum_{i,k}^{N_e, N_K} \frac{Z_k q^2}{|\mathbf{r}_i - \mathbf{r}_j|} \right). \quad (4.23)$$

4.3. Dichtefunktionaltheorie

Zur Lösung des oben diskutierten elektronischen Problems werden zwei Klassen von Verfahren angewendet. Die Hartree-Fock-Ansätze und die daraus resultierenden Verfahren finden in der Quantenchemie bei der Berechnung einzelner Moleküle Verwendung. Da die typischen Verfahren jedoch vom Rechenaufwand wie N^4 und höher mit ihrer Systemgröße skalieren, sind sie zur Beschreibung von Festkörpern mit bis zu hundert Atomen für Defektzellen nur beschränkt bis gar nicht einsetzbar.

Ein anderer Ansatz zur Lösung des elektronischen Problems basiert auf der Einführung der orts aufgelösten Elektronendichteverteilung $n(\mathbf{r})$. Diese ist über die Gleichung

$$n(\mathbf{r}) := N \int \dots \int \Psi^*(\mathbf{r}, \mathbf{r}_2, \dots, \mathbf{r}_N) \Psi(\mathbf{r}, \mathbf{r}_2, \dots, \mathbf{r}_N) d^3 r_2 d^3 r_3 \dots d^3 r_N \quad (4.24)$$

definiert. Der Faktor N bezeichnet die Gesamtzahl der zu behandelnden Teilchen und dient somit der Erhaltung der Teilchenzahl. Es gilt folglich der Zusammenhang

$$N = \int n(\mathbf{r})d^3r. \quad (4.25)$$

Man bezeichnet die Teilchenanzahl N hierbei auch als Funktional der Elektronendichte n und schreibt $N = N[n]$. Ebenso ist die Elektronendichte n aufgrund ihrer obigen Definition ein Funktional der N -Teilchenwellenfunktion Ψ .

Um jedoch die Wellenfunktion durch die Elektronendichte zu ersetzen, benötigt man zwischen diesen beiden Größen eine eindeutige Abbildung. Über diese Eigenschaft und einige weitere, die zur Lösung des obigen Problems beitragen, treffen die *Hohenberg-Kohn-Theoreme* Aussagen.

4.3.1. Die Hohenberg-Kohn-Theoreme

Die Hohenberg-Kohn Theoreme[22] sind das Herzstück der Dichtefunktionaltheorie. Nach Capelle[23] lassen sich 3 Hauptaussagen treffen.

Theorem 4.1 *Hohenberg-Kohn (I)*

Im Falle eines nichtentarteten Grundzustandes kann die Grundzustandswellenfunktion als ein erweitertes Funktional¹

$$\Psi_0(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) = \Psi[n_0(\mathbf{r})] \quad (4.26)$$

der Grundzustandselektronendichte dargestellt werden.

Aus diesem ersten Theorem folgen sämtliche wichtigen physikalischen Verankerungen um die Einteilchenelektronendichte als einzige notwendige Größe zu etablieren. Aufgrund dieser eindeutigen Funktionalabbildung trägt die Elektronendichte genau den gleichen Informationsgehalt wie die N -Teilchenwellenfunktion. Weiter sind alle physikalischen Erwartungswerte O von Observablen \hat{O} , welche vorher Funktionale der Wellenfunktion waren, nun auch Funktionale der Elektronendichte. Dies folgt direkt aus der Definition des quantenmechanischen Erwartungswertes

$$O_0 = O[n_0] := \langle \Psi[n_0] | \hat{O} | \Psi[n_0] \rangle \quad (4.27)$$

einer Observablen. Der Beweis dieses aussagekräftigen Theorems ist im Folgenden skizziert.

¹Entgegen der in Abschnitt 3.2 getätigten Definition eines Funktionals findet hier nun eine Zuordnung eines Skalars aus dem Zahlenkörper \mathbb{R} auf den Raum der zwei mal differenzierbaren und quadratintegriblen N -dimensionalen Funktionen \mathbb{L}^N statt. Ein solches Funktional wird als erweitertes Funktional bezeichnet.

Angenommen es gäbe zwei Potentiale $V(\mathbf{r})$ und $V'(\mathbf{r})$, die sich um mehr als eine additive Konstante unterscheiden, aber die gleiche Einteilchenelektronendichte $n(\mathbf{r})$ als Grundzustandsdichte besitzen. Diese Potentiale definieren zwei verschiedene Hamiltonoperatoren \hat{H} und \hat{H}' mit zwei unterschiedlichen Eigenfunktionen $|\Psi\rangle$ und $|\Psi'\rangle$. Für diese gilt:

$$\langle \Psi | \hat{H} | \Psi \rangle = E_0 \quad (4.28)$$

$$\langle \Psi' | \hat{H}' | \Psi' \rangle = E'_0. \quad (4.29)$$

Wendet man nun die Wellenfunktion Ψ' auf den Hamiltonoperator \hat{H} an, ist der erhaltene Energieeigenwert echt größer als die Grundzustandsenergie E_0 . Dies folgt aus dem Variationsprinzip der Quantenmechanik.

$$\begin{aligned} E_0 < \langle \Psi' | \hat{H} | \Psi' \rangle &= \langle \Psi' | \hat{H}' | \Psi' \rangle + \langle \Psi' | \hat{H} - \hat{H}' | \Psi' \rangle \\ &= E'_0 + \int n(\mathbf{r}) [v(\mathbf{r}) - v'(\mathbf{r})] d^3r \end{aligned} \quad (4.30)$$

Zum selben Schluss kommt man auch für die Wellenfunktion Ψ und den Hamiltonoperator \hat{H}' :

$$\begin{aligned} E'_0 < \langle \Psi | \hat{H}' | \Psi \rangle &= \langle \Psi | \hat{H} | \Psi \rangle - \langle \Psi | \hat{H} - \hat{H}' | \Psi \rangle \\ &= E_0 - \int n(\mathbf{r}) [v(\mathbf{r}) - v'(\mathbf{r})] d^3r. \end{aligned} \quad (4.31)$$

Eine Addition beider Ungleichungen führt zu der Ungleichung:

$$E_0 + E'_0 < E_0 + E'_0. \quad (4.32)$$

Dies stellt einen offensichtlichen Widerspruch dar. Zu zwei Potentialen, die sich um mehr als eine additive Konstante unterscheiden, müssen folglich auch zwei verschiedene Grundzustandsdichten n gehören.

Das zweite Hohenberg-Kohn-Theorem motiviert das Variationsprinzip.

Theorem 4.2 Hohenberg-Kohn(II)

Es existiert eine Grundzustandsladungsdichte n_0 zur Grundzustandsenergie

$$E_0 = E[n_0] = \langle \Psi[n_0] | \hat{H} | \Psi[n_0] \rangle. \quad (4.33)$$

Jede andere Ladungsdichte n' liefert eine Energie, die größer oder gleich der Grundzustandsenergie ist:

$$E[n_0] \leq E[n']. \quad (4.34)$$

Diese Aussage entspricht dem Variationsprinzip der Wellenfunktion und schränkt die berechnete Energie eines Variationsprozesses nach unten ein. Beim Vollzug der Minimierung der Energie E ist jedoch stets darauf zu achten, dass die Teilchenzahl N des Systems, sprich die Normierung der Teilchendichte, erhalten bleibt. Um dies zu bewerkstelligen wird nicht $E[n]$ variiert, sondern $E[n] - \mu(\int n(\mathbf{r})d\mathbf{r} - N)$. Solche Nebenbedingungen werden in der Variationsrechnung über sogenannte Lagrangeparameter μ eingebracht, denen eine physikalische Bedeutung zugeschrieben werden kann. In diesem Fall wird μ als *chemisches Potential* der Elektronen interpretiert.

Eine weitere Aussage treffen die Hohenberg-Kohn-Theoreme über das Funktional der potentiellen Energie.

Theorem 4.3 *Hohenberg-Kohn (III)*

Wenn Ausdrücke für die Operatoren der kinetischen Energie T , der Wechselwirkungsenergie V_{ee} und der Energie des externen Potentials V_{eK} bekannt sind, existieren auch ihre Funktionale $T[n]$, $V_{ee}[n]$, sowie $V_{eK}[n]$ und es gilt

$$E[n] = T[n] + V_{ee}[n] + V_{eK}[n] = F[n] + V_{eK}[n]. \quad (4.35)$$

Zwar wird hier die Existenz der Funktionale belegt, jedoch keine Aussagen über deren explizite Gestalt getroffen. Dies ist gerade für die essentiellen Funktionale der kinetischen Energie und der Elektron-Elektron-Wechselwirkung kritisch, für welche bis heute keine exakte Form bekannt ist. Für das externe Potential $V_{eK}[n]$ hingegen lässt sich sofort ein Funktional hinschreiben:

$$V_{eK}[n] = \int n(\mathbf{r})v_{eK}(\mathbf{r})d^3r. \quad (4.36)$$

4.3.2. DFT in der Anwendung: Die Kohn-Sham Gleichungen

Die Hohenberg-Kohn Theoreme zeigen, dass ein Wechsel von der Mehrteilchenwellenfunktion zur Einteilchenladungsdichte theoretisch möglich ist und sämtliche Operatoren auch als Funktionale der Eielektronendichte existieren. Die möglichen Lösungsverfahren dieses Problems sind vielfältig. Die Minimierung des Energiefunktional nach der Elektronendichte stellt hierbei jedoch einen ineffektiven Ansatz dar. Weit verbreitet ist deswegen heute die Methode von Kohn und Sham[24, 23]. Das Vielteilchenproblem wird, wie in der Hartree-Fock-Theorie, auf ein effektives Einteilchenproblem reduziert. Dabei wird das wechselwirkende System durch ein wechselwirkungsfreies System in einem effektiven Einteilchenpotential ersetzt. Zu diesem Ansatz wird man geradezu gedrängt, da die einzelnen Energiefunktionale

$$E[n] = T[n] + V_{ee}[n] + V_{eK}[n] \quad (4.37)$$

nicht bekannt sind und man gezwungen ist diese durch geeignete Wahl zu nähern.

Eine explizite Darstellung des Funktionals der Elektron-Kern-Wechselwirkung ist über die Gleichung (4.36) bereits gegeben. Probleme bereiten die Funktionale der kinetischen Energie und der Elektron-Elektron-Wechselwirkung. Aus der Hartree-Fock Theorie ist bekannt, dass Vielteilcheneffekte bzw. Korrelationseffekte eine Rolle spielen und diese sich in diesen Funktionalen widerspiegeln müssen. Ohne Kenntnis der expliziten Form formuliert man das Problem dergestalt um, dass man auf Bekanntes aus der Hartree-Fock-Theorie zurückgreifen kann: Man zerlegt die Funktionale in ihre Einteilchenanteile und Vielteilchenanteile:

$$E[n] = T_{\text{single}}[n] + V_{\text{H}}[n] + V_{\text{eK}}[n] + E_{\text{xc}}[n]. \quad (4.38)$$

Die Vielteilcheneffekte werden nun durch ein einziges Funktional E_{xc} beschrieben, welches die Austausch- und Korrelationsenergie enthält. Das Funktional der kinetischen Einteilchenenergie ist jedoch nur ein implizites Funktional der Elektronendichte. Da schon dieses in der expliziten Form nicht bekannt ist, verwendet man die Darstellung über Einteilchenwellenfunktion

$$T_{\text{single}}[n] = T_{\text{single}}[\{\Phi_i\}] = -\frac{\hbar^2}{2m} \sum_i^N \langle \Phi_i | \Delta \Phi_i \rangle. \quad (4.39)$$

Das Potential V_{H} für die Ladungswchselwirkung der Elektronen wird aus der Hartree-Fock-Theorie übernommen:

$$V_{\text{H}} = \frac{1}{4\pi\epsilon_0} \frac{q^2}{2} \iint \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3r' d^3r. \quad (4.40)$$

Unbekannt bleibt nun lediglich das Austausch- und Korrelationsfunktionalfunktional. Gängige Näherungen für dieses werden im nächsten Abschnitt diskutiert.

Um nun das elektronische Problem zu lösen, ist die Variation der Elektronendichte

$$\left. \frac{\delta E[n]}{\delta n(\mathbf{r})} \right|_{n_0} = \left. \frac{\delta T_{\text{single}}[n]}{\delta n(\mathbf{r})} \right|_{n_0} + \underbrace{\left. \frac{\delta V_{\text{H}}[n]}{\delta n(\mathbf{r})} \right|_{n_0}}_{v_{\text{H}}[n_0]} + \underbrace{\left. \frac{\delta V_{\text{eK}}[n]}{\delta n(\mathbf{r})} \right|_{n_0}}_{v_{\text{eK}}[n_0]} + \underbrace{\left. \frac{\delta E_{\text{xc}}[n]}{\delta n(\mathbf{r})} \right|_{n_0}}_{v_{\text{xc}}[n_0]} \stackrel{!}{=} 0 \quad (4.41)$$

von Nöten. Die Herren Kohn und Sham gingen hier allerdings einen pragmatischen Weg. Sie führten ein nichtwechselwirkendes System in einem effektiven Einteilchenpotential $v_{\text{eff}}([n])$ ein. Die daraus resultierende Minimierungsvorschrift

$$\left. \frac{\delta E[n]}{\delta n(\mathbf{r})} \right|_{n_0} = \left. \frac{\delta T_{\text{single}}[n]}{\delta n(\mathbf{r})} \right|_{n_0} + \underbrace{\left. \frac{\delta V_{\text{eff}}[n]}{\delta n(\mathbf{r})} \right|_{n_0}}_{v_{\text{eff}}[n_0]} \stackrel{!}{=} 0 \quad (4.42)$$

zeigt eine Identität mit der Lösung des Vielteilchenproblems, wenn

$$v_{\text{eff}}[n_0] = v_{\text{H}}[n_0] + v_{\text{eK}}[n_0] + v_{\text{xc}}[n_0] \quad (4.43)$$

erfüllt ist. Dies kann für physikalische Systeme immer erreicht werden. Damit ist nun die Lösung des effektiven Einteilchenproblems

$$\left(-\frac{\hbar^2}{2m}\Delta + V_{\text{eff}}[n_0]\right) |\Phi_i(\mathbf{r})\rangle = \epsilon_i |\Phi_i(\mathbf{r})\rangle \quad (4.44)$$

von Interesse. Bei bekannter Grundzustandselektronendichte n_0 lässt sich dieses Problem durch geeignete Standardverfahren lösen. Doch genau dort liegt das Problem. Die Grundzustandselektronendichte ist unbekannt und zudem noch ein Funktional der Orbitalfunktionen:

$$n_0(\mathbf{r}) = \sum_i^N |\Phi_i(\mathbf{r})|^2. \quad (4.45)$$

Die Kohn-Sham-Gleichungen sind somit nichtlinear. Solche Gleichungen können aber iterativ selbstkonsistent gelöst werden, indem die Dichte in jedem Berechnungsschritt konsistent zu den Wellenfunktionen gewählt wird und das Gleichungssystem solange gelöst wird, bis die Dichte konvergiert ist. Die Gesamtenergie des Problems berechnet² sich dann zu

$$\begin{aligned} E[n] &= T_{\text{single}}[n] + V_{\text{eK}}[n] + V_{\text{H}}[n] + E_{\text{xc}}[n] \\ &= T_{\text{single}}[n] + \int v_{\text{eff}}n - v_{\text{H}}n - v_{\text{xc}}n \, d\mathbf{r} + V_{\text{H}}[n] + E_{\text{xc}}[n] \\ &= T_{\text{single}}[n] + V_{\text{eff}}[n] - 2V_{\text{H}}[n] - \int v_{\text{xc}}n \, d\mathbf{r} + V_{\text{H}}[n] + E_{\text{xc}}[n] \\ &= \sum_i \epsilon_i - V_{\text{H}}[n] - \int v_{\text{xc}}n \, d\mathbf{r} + E_{\text{xc}}[n]. \end{aligned} \quad (4.46)$$

An dieser Stelle erkennt man, dass die Gesamtenergie nicht durch einfache Summation der Kohn-Sham Orbitalenergien gegeben ist. Das bedeutet, dass die mit dem Kohn-Sham-Schema berechnete Lösung keine vollständige Lösung des wechselwirkenden Systems ist. Lediglich die Grundzustandselektronendichte und die Grundzustandsenergie kommen identisch heraus. Die Kohn-Sham-Energien stimmen jedoch zumindest qualitativ mit dem echten Energiespektrum überein[25]. Aus diesem Grund werden auch Bandstrukturen und Bandlücken mit den Kohn-Sham-Energieeigenwerten berechnet. Man muss sich jedoch im klaren darüber sein, dass diese keinen realen Bezug haben müssen. Dies fällt besonders gravierend in der Unterschätzung von Bandlücken ins Gewicht, welche bei den gängigen Näherung des Austausch- und Korrelationsfunktionals auftauchen. Diesen Näherungen widmen sich nun die nächsten Abschnitte.

²Dabei wird der Zusammenhang $V_{\text{eff}}[n] = \int v_{\text{eff}}[n]n \, d\mathbf{r}$ verwendet. Es ist allerdings zu beachten, dass die Integration über das Hartree-Potential $\int v_{\text{H}}[n(\mathbf{r})]n(\mathbf{r}) \, d\mathbf{r} = \iint n(\mathbf{r})n(\mathbf{r}')/|\mathbf{r} - \mathbf{r}'| \, d\mathbf{r}' \, d\mathbf{r} = 2V_{\text{H}}[n]$ (atomare Einheiten: $4\pi\epsilon_0 = q = 1$) ergibt. Über die Integration des Austausch- und Korrelationspotentials v_{xc} lässt sich an dieser Stelle noch keine Aussage treffen.

Kristall	Bandlücke [eV]					a_{lat} [Å]
	LDA	PBE	Art	exp ^a	Art	
Diamant	4.14	4.16	indirekt	5.4	indirekt	3.55
Si	0.48	0.58	indirekt	1.17	indirekt	5.43
Ge	0.06	0.07	direkt	0.744	indirekt	5.66
Ge ($\Gamma \rightarrow L$)	0.06	0.20	indirekt	0.744	indirekt	5.66
GaN (ZB)	1.79	1.75	direkt	3.2	direkt	4.52
GaP	1.47	1.63	indirekt	2.32	indirekt	5.45
GaAs	0.41	0.58	direkt	1.52	direkt	5.65

^a Referenz[26]

^b Referenz[27]

Tab. 4.1:

Die mit den gängigen Funktionalen LDA und PBE berechneten Bandlücken zeigen eine deutliche Unterschätzung. Die Unterschätzung kann sogar bis zu einer völligen Fehlbeschreibung des Materials gehen, wie im Falle des Germaniums, welches in den DFT Rechnungen nahezu metallisch ist. Die Berechnung erfolgte an der experimentellen Gitterkonstante.

4.3.3. Die lokale Dichtenäherung (LDA)

In Ermangelung einer geschlossenen analytischen Form für das Austausch- und Korrelationsfunktional ist man gezwungen ein solches geeignet zu konstruieren. Die lokalen Dichtenäherungen sind dabei solche, bei denen das Austausch- und Korrelationsfunktional allein von der Elektronendichte abhängt — im Gegensatz zu den später beschriebenen Korrekturverfahren, die noch den Gradienten der Dichte einbauen.

Basierend auf dem Modell eines homogenen Elektronengas von Thomas und Fermi lässt sich das Austauschfunktional im Falle geschlossener Schalen — also gleicher Elektronenanzahl mit „Spin up“ und „Spin down“ — zu

$$E_x[n] = -\frac{3}{4} \left(\frac{3}{\pi} \right)^{1/3} \int (n(\mathbf{r}))^{4/3} d^3r \quad (4.47)$$

berechnen[28]. Für offene Schalen — also bei unterschiedlicher Anzahl von „Spin up“ und „Spin down“ Elektronen — ergibt sich jenes zu

$$E_x[n] = -\frac{3}{2} \left(\frac{3}{4\pi} \right)^{1/3} \sum_{\sigma} \int (n_{\sigma}(\mathbf{r}))^{4/3} d^3r. \quad (4.48)$$

Der Laufindex σ bezeichnet hierbei die zwei möglichen Spineinstellungen[29].

Obwohl diese Ausdrücke schon 1930 bekannt waren, brauchte es weitere 50 Jahre bis Vosko, Wilk und Nusair das Korrelationsfunktional hinreichend genau bestimmen

konnten. Ceperley und Alder entwickelten später basierend auf Quantum Monte-Carlo-Rechnungen einen weiteren Ausdruck für das Korrelationsfunktional[30]. Perdew und Zunger schlugen ein Verfahren vor um die fehlerhafte Beschreibung der Selbstwechselwirkung zu korrigieren[31]. Die Addition beider Ausdrücke zu einem Austausch- und Korrelationsfunktional

$$E_{xc}^{\text{LDA}} = E_x[n] + E_c^{\text{CA}}[n] \quad (4.49)$$

ist heute der gängige Ausdruck zur Berechnung von Materialeigenschaften auf dem LDA Niveau.

Im Vergleich zu den experimentellen Befunden neigt die LDA dazu Gitterkonstanten und Bindungslängen tendenziell zu unterschätzen[29]. Bindungsenergien werden tendenziell überschätzt, weswegen LDA tendenziell zum sogenannten *Overbinding* neigt. Bandlücken werden deutlich unterschätzt, wie die Datensammlung in Tabelle 4.1 zeigt.

4.3.4. Gradientenkorrekturen und das PBE-Funktional

Um den Inhomogenitäten in realen Ladungsverteilungen Rechnung zu tragen, wurden auf Basis der Weizsäckerschen Inhomogenitätskorrektur

$$n_{\text{corr}} = \frac{|\nabla n|^2}{n}$$

Versuche unternommen, LDA zu verbessern. 1988 formulierte Becke[29], basierend auf einer Arbeit von Herman, van Dyke und Ortenberger[32], sein Austauschfunktional.

$$E_x^{\text{Becke}}[n] = E_x^{\text{LDA}}[n] - \beta \sum_{\sigma} \int n_{\sigma}^{4/3} \frac{x_{\sigma}^2}{1 + 6\beta x_{\sigma} \sinh^{-1} x_{\sigma}} \quad (4.50)$$

mit

$$x_{\sigma} = \frac{|\nabla n_{\sigma}|}{n_{\sigma}^{4/3}} \quad (4.51)$$

welches den korrekten asymptotischen Coulombverlauf im Unendlichen besitzt. Der Parameter β ist ein Fit-Parameter. Becke bestimmte ihn, indem er den exakten Hartree-Fock Austausch für die Edelgasatome bestimmte und den Parameter mittels eines *least-square*-Fits an diese Daten anpasste. Die mit diesem semi-empirischen Austauschfunktional berechneten Austauschenergien für andere Atome und 3d-Übergangselemente sind erstaunlich gut. Dies liegt daran, dass die Korrektur erst für große x_{σ} beginnt, eine Rolle zu spielen. Große x_{σ} bedeuten eine stark fluktuierende Elektronendichte n , wie sie beispielsweise in gerichteten Bindungen auftritt. Die Korrektur wirkt subtraktiv zum LDA Austausch und bewirkt somit eine Reduzierung der Elektronendichte, was zu einer Abschwächung

der Bindung führt. Dem Overbinding wird folglich entgegengewirkt.

1986 gelang es Perdew das Funktional der Korrelationsenergie weiter zu verbessern[33]. Die Kombination von Perdews Korrelationsfunktional und Beckes Austauschfunktional wird als BP86 bezeichnet und war lange Zeit das Standard Austausch- und Korrelationsfunktional. Jedoch besitzt dieses Funktional bezüglich des Korrelations- und Austauschloches das falsche Abklingverhalten in der Zweiteilchenelektronendichte

$$n_2(\mathbf{r}_1, \mathbf{r}_2) = N(N-1) \int \Psi^*(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) \Psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) d\mathbf{r}_3 \dots d\mathbf{r}_N. \quad (4.52)$$

Aufgrund des Pauliverbotes und der Coulombabstoßung dürfen sich zwei Elektronen nicht beliebig nahe kommen, so dass

$$\lim_{\mathbf{r}_1 \rightarrow \mathbf{r}_2} n_2(\mathbf{r}_1, \mathbf{r}_2) = 0 \quad (4.53)$$

gelten muss. Die Art der Nullstelle ist aber essentiell unterschiedlich. Das *Pauli-Loch* — die Elektronen besitzen den gleichen Spin — besitzt die Eigenschaft

$$\left. \frac{dn_2(\mathbf{r}_1 \otimes \mathbf{s}, \mathbf{r}_2 \otimes \mathbf{s})}{dr_{12}} \right|_{r_{12}=0} = 0, \quad (4.54)$$

während das *Coulomb-Loch* — die Elektronen besitzen einen unterschiedlichen Spin — eine Nullstelle mit der Eigenschaft

$$\left. \frac{dn_2(\mathbf{r}_1 \otimes \mathbf{s}_1, \mathbf{r}_2 \otimes \mathbf{s}_2)}{dr_{12}} \right|_{r_{12}=0} = \pm\infty \quad (4.55)$$

besitzt. Bei der Konstruktion eines neuen Austausch- und Korrelationsfunktionals haben Perdew und Wang dieses Abklingverhalten berücksichtigt und das PW91 Funktional geschaffen[34]. 1996 beseitigte Perdew einige Schwächen seine Funktionals. Insbesondere die Zahl der Parameter wurde deutlich reduziert um das Funktional verständlicher und durchschaubarer zu machen, ohne dabei etwas an Genauigkeit einzubüßen[35]. Das als GGA-PBE bekannte Funktional gilt heutzutage als Standardfunktional der Dichtefunktionaltheorie in der generalisierten Gradienten-näherung.

4.3.5. Hartee-Fock Austausch und Hybrid-Funktionale

LDA und PBE liefern sehr gute Strukturdaten. Allerdings versagen diese Funktionale bei der Vorhersage von physikalischen Observablen, die einen direkten Bezug zur Bandlücke besitzen. Aktuelle Veröffentlichungen zeigen zudem, dass auch die relative Lage des Valenzbandmaximums zum mittleren elektrostatischen Potential fehlerhaft beschrieben wird[39, 40, 41]. Der Grund für dieses Versagen ist eine zu

Kristall	Bandlücke [eV]			
	HSE	Art	exp	Art
GaN (WZ)	3.24	direkt	3.5 ^a	direkt
Mg ₃ N ₂	2.76	direkt	2.8 ^{b,c}	direkt ^b /indirekt ^c

^a Referenz[36]

^b Referenz[37]

^c Referenz[38]

Tab. 4.2:

Die mit HSE berechneten Bandlücken an der optimierten LDA Struktur für GaN und an der optimierten PBE Struktur von Mg₃N₂ zeigen eine gute Übereinstimmung zu den experimentellen Befunden. Die leichte Unterschätzung der Bandlücke von GaN ist der LDA Struktur mit der durchs *Overbinding* zu geringen Gitterkonstante geschuldet. Rechnungen an der PBE Struktur ergeben eine Bandlücke von 3.46 eV.

starke Lokalisierung des Austauschlochs in der LDA und der PBE. Hybridfunktionale, welche zu dem PBE-Austausch einen gewissen Betrag an Hartree-Fock Austausch mischen, zeigen ein deutlich delokalisiertes Austauschloch und beschreiben die Bandlücke sowie die relative Position des Valenzbandmaximums besser[42, 43, 44]. Die Konstruktion des Hybridfunktionals geschieht dabei nach der Vorschrift

$$E_{XC}^{\text{Hybrid}} = \alpha E_X^{\text{HF}} + (1 - \alpha) E_X^{\text{PBE}} + E_C^{\text{PBE}}. \quad (4.56)$$

Im Falle von $\alpha = 0.25$ erhält man das Austauschfunktional, das als *PBE0*[45, 46, 47, 48] in der Literatur bezeichnet wird. Die Verwendung des wie r^{-1} -abklingenden Austauschpotentials führt jedoch zu einem eklatanten Problem für metallische Systeme: Während Moleküle, Halbleiter und Isolatoren ein exponentielles Abklingverhalten in der Wellenfunktion besitzen, welches das r^{-1} Verhalten dämpft, schlägt in Metallen dieses Abklingverhalten voll zu Buche und führt zu divergierenden Integralen und langsamer Konvergenz. Praktisch wird versucht dies durch ein abgeschirmtes Coulombpotential zu umgehen, welches eine geringere Reichweite als das r^{-1} -Potential besitzt. Heyd, Scuseria und Ernzerhof (HSE) zerlegten dazu das r^{-1} -Potential in einen langreichweitigen und einen kurzreichweitigen Anteil:

$$\frac{1}{r} = \frac{\text{erfc}(\omega r)}{r} + \frac{\text{erf}(\omega r)}{r}. \quad (4.57)$$

Dabei bezeichnet der Ausdruck

$$\text{erf}(\omega r) = \frac{2}{\sqrt{\pi}} \int_0^{\omega r} e^{-t^2} dt \quad (4.58)$$

die Fehlerfunktion und projiziert den langreichweitigen Anteil aus dem Coulombpotential. Die komplementäre Fehlerfunktion

$$\operatorname{erfc}(\omega r) = \frac{2}{\sqrt{\pi}} \int_{\omega r}^{\infty} e^{-t^2} dt \quad (4.59)$$

projiziert den kurzreichweitigen Anteil aus dem Coulombpotential. Im Falle $\omega \rightarrow 0$ geht die Fehlerfunktion gegen Null und die komplementäre Fehlerfunktion gegen 1. Für grosse ω ist es genau umgekehrt. Der Parameter ω erlaubt folglich eine Einstellung der Reichweite des kurzreichweitigen Potentials und ist von HSE auf Basis von zahlreichen Rechnungen justiert worden[43, 44]. Für den Hartree-Fock Austausch ist $\omega_{\text{HF}} = 0.15/\sqrt{(2)}a_0^{-1}$ verwendet worden, der PBE Austausch wird mit einem $\omega_{\text{PBE}} = 0.15 \times 2^{1/3}a_0^{-1}$ abgeschirmt. Der Mischungsparameter α ist identisch zu dem in PBE0. Basierend auf diesen Rechnungen konnten HSE zeigen, dass die langreichweitige (LR) Beimischung des Hartree-Fock Austausches nur einen geringen Einfluss auf die numerischen Ergebnisse besitzt. Das eigentliche HSE-Funktional mischt folglich nur kurzreichweitig (SR) den Hartree-Fock Austausch mit dem PBE Austausch

$$E_{\text{XC}}^{\text{HSE}} = \alpha E_{\text{X}}^{\text{HF,SR}} + (1 - \alpha) E_{\text{X}}^{\text{PBE,SR}} + E_{\text{X}}^{\text{PBE,LR}} + E_{\text{C}}^{\text{PBE}} \quad (4.60)$$

Eigene Testrechnungen unter Verwendung der HSE Implementierung in VASP[17] liefern für die in dieser Dissertation betrachteten Materialien eine signifikante Verbesserung in der Beschreibung der Bandlücke (vgl. Tabelle 4.2). Insbesondere für Mg_3N_2 ist somit HSE eine echte Alternative zur Berechnung von Eigenschaften der Bandstruktur im Vergleich zur aufwendigen und rechenzeitintensiven Modellierung der primitiven Zelle mit 40 Atomen mit dem GW-Ansatz[49, 50], welcher vor HSE als akkurater Standard zur Berechnung von Bandstruktureigenschaften ohne Bandlückenproblem galt.

5. Dichtefunktionaltheorie für kristalline Systeme

Man hat den Eindruck, dass die moderne Physik auf Annahmen beruht, die irgendwie dem Lächeln einer Katze gleichen, die gar nicht da ist.

Albert Einstein, Physiker und Nobelpreisträger, 1879-1955

5.1. Der Impulsraum-Formalismus

Im Blickpunkt der Festkörpertheorie stehen die Eigenschaften von kristallinen Strukturen. Diese zeichnen sich durch ihre Symmetrien aus. Insbesondere die Ausnutzung der Translationssymmetrie erlaubt die Einführung periodischer Randbedingungen nach Born und Kármán[51, 52], durch welche die Kristalle in *primitiven Zellen* beschrieben werden können. Im Allgemeinen werden Symmetrietransformationen des dreidimensionalen Raumes mittels Operationen der Gestalt

$$\mathbf{r}' = R\mathbf{r} + \mathbf{t} =: \{R|\mathbf{t}\}\mathbf{r} \quad (5.1)$$

vollzogen. Dabei bezeichnet R eine Operation aus der Punktgruppe $O(3)$ — zum Beispiel Drehungen und Spiegelungen — und \mathbf{t} eine beliebige Translation. Im Rahmengruppentheoretischer Betrachtungen kann man zeigen, dass solche Raumoperationen eine Gruppe bilden und die Translationen $\{E|\mathbf{t}\}\mathbf{r}$ eine invariante Untergruppe der Raumtransformationen sind. E ist hierbei die identische Abbildung. Lässt man bei den Raumtransformationen nur diskrete Translationsvektoren der Gestalt

$$\mathbf{t} = \sum_{i=1}^3 n_i \mathbf{a}_i \quad n_i \in \mathbb{Z} \quad \{\mathbf{a}_i\}_{i=1}^3 \text{ ist Basis des } \mathbb{R}^3 \quad (5.2)$$

zu, so nennt man die Gruppe der Raumtransformationen *Raumgruppe*.

Um nun Darstellungen der Translationsgruppe T zu finden, untersucht man Translationen des Kristalls entlang seines ersten primitiven Gittervektors \mathbf{a}_1 . Führt man nun zwei Translationen \mathbf{t}_1 und \mathbf{t}_2 hintereinander aus, so ist anschaulich klar, dass eine solche Translation auch als direkte Translation $\mathbf{t}_3 = \mathbf{t}_1 + \mathbf{t}_2$ ausführbar ist. Um nun eine Darstellung dieser Gruppe zu finden benötigt man Abbildungen der Translationsoperatoren auf unitäre Matrizen Γ , so dass die Abbildungsvorschrift

$$\mathbf{t}_1 + \mathbf{t}_2 = \mathbf{t}_3 \longrightarrow \Gamma(\mathbf{t}_1)\Gamma(\mathbf{t}_2) = \Gamma(\mathbf{t}_3) \quad \forall \quad \mathbf{t}_1, \mathbf{t}_2, \mathbf{t}_3 \in T \quad (5.3)$$

erfüllt ist. Die Gruppe der Translation ist abelsch, da die Vektoraddition kommutativ ist. Für abelsche Gruppen, zu denen die Gruppe der Translationen gehört, sind die entsprechenden unitären Matrizen Skalare.

Betrachten wir nun im folgenden eine primitive Kristallzelle mit dem Volumen $\Omega = \mathbf{a}_1 \cdot (\mathbf{a}_2 \times \mathbf{a}_3)$. Zur Beschreibung des Kristalls verwenden wir N_1 Zellen in

\mathbf{a}_1 -Richtung, N_2 Zellen in \mathbf{a}_2 -Richtung sowie N_3 Zellen in \mathbf{a}_3 -Richtung. Das Kristallvolumen berechnet sich nun zu $V = N_1 N_2 N_3 \Omega = N \Omega$. Betrachtet man nun Translationen um den primitiven Gittervektor \mathbf{a}_1 , so ist anschaulich klar, dass eine solche Verschiebung den Kristall auf sich selbst abbildet. Auch eine N_1 -fache Anwendung dieser Translation bildet den Kristall auf sich selbst ab. Folglich gilt

$$N_1 \mathbf{a}_1 = \mathbf{0} \longrightarrow [\Gamma(\mathbf{a}_1)]^{N_1} = 1. \quad (5.4)$$

Daraus lässt sich nun die Darstellung dieser Translation zu

$$\Gamma(\mathbf{a}_1) = \sqrt[N_1]{1} \longrightarrow \Gamma(\mathbf{a}_1) = e^{-2\pi i m_1 / N_1} \quad 0 \leq m_1 \leq N_1 - 1 \quad (5.5)$$

ableiten. In der Verallgemeinerung auf dreidimensionale Translationen der Gestalt

$$\mathbf{t} = n_1 \mathbf{a}_1 + n_2 \mathbf{a}_2 + n_3 \mathbf{a}_3 \quad (5.6)$$

ergibt sich somit für deren Darstellung

$$\Gamma(\mathbf{t}) = \left(e^{-2\pi i m_1 / N_1} \right)^{n_1} \left(e^{-2\pi i m_2 / N_2} \right)^{n_2} \left(e^{-2\pi i m_3 / N_3} \right)^{n_3}. \quad (5.7)$$

Diese Darstellung hat die Form

$$\Gamma(\mathbf{t}) = e^{-i\mathbf{k}\mathbf{t}}, \quad (5.8)$$

wenn \mathbf{k} als

$$\mathbf{k} = \sum_{i=1}^3 \frac{m_i}{N_i} \mathbf{b}_i \quad (5.9)$$

angesetzt wird und die \mathbf{b}_i die primitiven Gittervektoren des reziproken Gitters bezeichnen. Die reziproken Gittervektoren werden über die Gleichung

$$\mathbf{b}_i \cdot \mathbf{a}_j = 2\pi \delta_{ij} \quad (5.10)$$

definiert.

Da nun der Hamiltonoperator mit der Translationsoperation vertauscht, besitzen beide Operatoren einen gemeinsamen Satz von Eigenfunktionen. Somit bilden die elektronischen Wellenfunktionen eine Basis der Eigenfunktionen des Translationsoperators und können mit einem entsprechenden Reziprokraumvektor \mathbf{k} indiziert werden. Wendet man nun den Translationsoperator auf eine solche Wellenfunktion an

$$\phi_{\mathbf{k}}(\mathbf{r} - \mathbf{t}) = \hat{t}\phi_{\mathbf{k}}(\mathbf{r}) = \Gamma(\mathbf{t})\phi_{\mathbf{k}}(\mathbf{r}) = e^{-i\mathbf{k}\mathbf{t}}\phi_{\mathbf{k}}(\mathbf{r}) \quad (5.11)$$

folgt auf völlig natürliche Weise das Bloch-Theorem der Festkörperphysik:

$$\phi_{\mathbf{k}}(\mathbf{r} + \mathbf{t}) = e^{i\mathbf{k}\mathbf{t}}\phi_{\mathbf{k}}(\mathbf{r}). \quad (5.12)$$

In der Festkörpertheorie wird nun der Vektor \mathbf{k} auf die erste Brillouinzone beschränkt. Dadurch wird der Bandindex n eingeführt, um für ein festes \mathbf{k} eine Zuordnung des Energiewertes treffen zu können. Für die Wellenfunktion wird der Ansatz

$$\phi_{n\mathbf{k}}(\mathbf{r}) = u_{n\mathbf{k}}(\mathbf{r})e^{i\mathbf{k}\mathbf{r}} \quad (5.13)$$

gemacht. Die Funktion $u_{n\mathbf{k}}(\mathbf{r})$ besitzt hierbei die Periodizität des Gitters und lässt sich gemäss

$$\phi_{n\mathbf{k}}(\mathbf{r}) = \frac{1}{\sqrt{\Omega}} \sum_{\mathbf{G}} c_{n\mathbf{k}}(\mathbf{G})e^{i(\mathbf{G}+\mathbf{k})\mathbf{r}} \quad (5.14)$$

nach ebenen Wellen entwickeln. Dies stellt eine Fourierentwicklung der Wellenfunktion mit den Fourierkoeffizienten $c_{n\mathbf{k}}(\mathbf{G})$ dar. Im Rahmen numerischer Methoden bedarf es eines Abbruchkriteriums für diese Fourierentwicklung. Hier findet sich die in zahlreichen Konvergenztests verwendete *Cutoffenergie* E_{cut} wieder. Man nimmt in der Fourierentwicklung nur solche \mathbf{G} -Vektoren mit, welche die Relation

$$\frac{\hbar^2}{2m}(\mathbf{G} + \mathbf{k})^2 \leq E_{\text{cut}} \quad (5.15)$$

erfüllen und prüft durch sukzessives Erhöhen dieser Cutoffenergie die Konvergenz der Fourierreihenentwicklung.

Mit $|\mathbf{G} + \mathbf{k}\rangle$ werden die Basisfunktionen der ebenen Wellen des Impulsoperators zum Eigenwert $i(\mathbf{G} + \mathbf{k})$ definiert. Diese genügen der Orthonormierung

$$\langle \mathbf{G} + \mathbf{k} | \mathbf{G}' + \mathbf{k}' \rangle = \delta_{\mathbf{G}\mathbf{G}'}\delta_{\mathbf{k}\mathbf{k}'} \quad (5.16)$$

und in der Konvergenz der Vollständigkeitsrelation

$$\sum_{\mathbf{G}+\mathbf{k}} |\mathbf{G} + \mathbf{k}\rangle \langle \mathbf{G} + \mathbf{k}| = 1. \quad (5.17)$$

Im Ortsraum lässt sich die Wellenfunktion als

$$\langle \mathbf{r} | \mathbf{G} + \mathbf{k} \rangle = \frac{1}{\sqrt{\Omega}} e^{i(\mathbf{G}+\mathbf{k})\mathbf{r}} \quad (5.18)$$

darstellen. Damit formuliert sich die Beziehung (5.14) zu

$$|\phi_{n\mathbf{k}}\rangle = \sum_{\mathbf{G}} c_{n\mathbf{k}}(\mathbf{G}) |\mathbf{G} + \mathbf{k}\rangle. \quad (5.19)$$

Die umformulierte Kohn-Sham-Gleichung

$$\begin{aligned}
 \sum_{\mathbf{G}'+\mathbf{k}'} \left(-\frac{\hbar^2}{2m} \Delta + v_H[n] + v_{xc}[n] + v_{PS}(\mathbf{r}) \right) |\mathbf{G}' + \mathbf{k}'\rangle \langle \mathbf{G}' + \mathbf{k}' | \phi_{n\mathbf{k}} \rangle \\
 = \sum_{\mathbf{G}'+\mathbf{k}'} \epsilon_{n\mathbf{k}} |\mathbf{G}' + \mathbf{k}'\rangle \langle \mathbf{G}' + \mathbf{k}' | \phi_{n\mathbf{k}} \rangle
 \end{aligned} \tag{5.20}$$

lässt sich durch Multiplikation mit $\langle \mathbf{G} + \mathbf{k} |$ in die numerisch gut zu behandelnde Matrixgleichung

$$\begin{aligned}
 \frac{\hbar^2}{2m} (\mathbf{G} + \mathbf{k})^2 c_{n\mathbf{k}}(\mathbf{G}) \\
 + \left[\sum_{\mathbf{G}'} (v_H(\mathbf{G} - \mathbf{G}') + v_{xc}(\mathbf{G} - \mathbf{G}') + v_{PS}(\mathbf{G} - \mathbf{G}')) \right] c_{n\mathbf{k}}(\mathbf{G}) \\
 = \epsilon_{n\mathbf{k}} c_{n\mathbf{k}}(\mathbf{G})
 \end{aligned} \tag{5.21}$$

überführen. Hierbei wurde als Kurzschreibweise

$$v(\mathbf{G} - \mathbf{G}') = \langle \mathbf{G} + \mathbf{k} | v | \mathbf{G}' + \mathbf{k} \rangle \tag{5.22}$$

benutzt. Aus den Entwicklungskoeffizienten $c_{n\mathbf{k}}(\mathbf{G})$ und den aus einer Fermiverteilung berechneten Besetzungszahlen $f_{n\mathbf{k}}$ kann dann die Elektronendichte im reziproken Raum zu

$$n(\mathbf{G}) = \frac{2}{\sqrt{\Omega}} \sum_{n,\mathbf{k}} f_{n\mathbf{k}} \sum_{\mathbf{G}'} c_{n\mathbf{k}}^*(\mathbf{G}') c_{n\mathbf{k}}(\mathbf{G} + \mathbf{G}') \tag{5.23}$$

berechnet werden. Aufgrund der besseren Skalierung wird die Elektronendichte oftmals direkt im Ortsraum berechnet und ergibt sich dort zu

$$n(\mathbf{r}) = \frac{2}{\Omega} \sum_{n,\mathbf{k}} f_{n\mathbf{k}} \left| \sum_{\mathbf{G}} c_{n\mathbf{k}}(\mathbf{G}) e^{i\mathbf{G}\mathbf{r}} \right|^2 \tag{5.24}$$

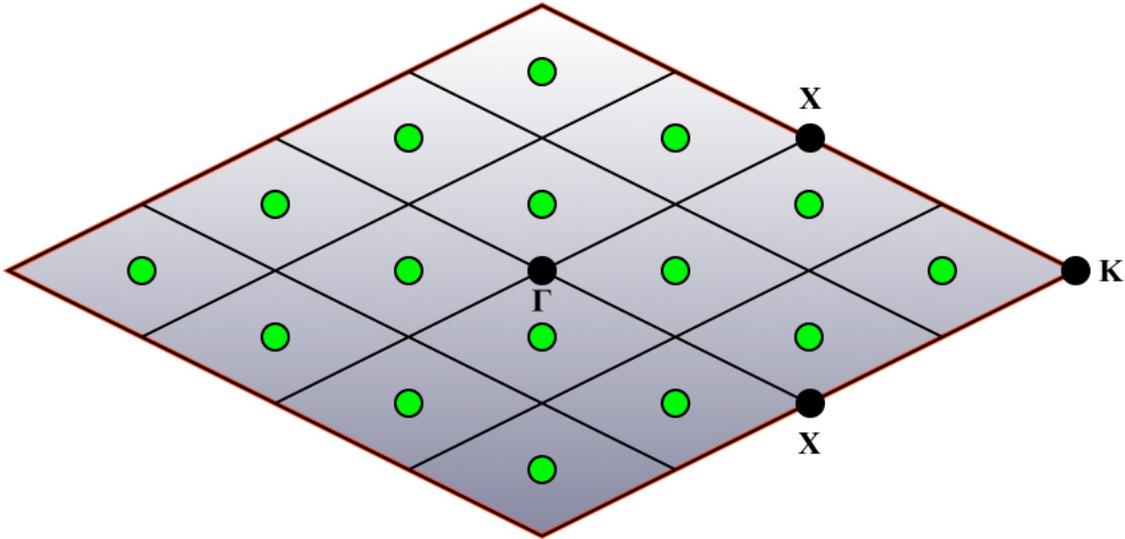
Die einzelnen Fouriertransformierten der Potentiale lassen sich mittels FFT berechnen. Vereinzelt ist es auch möglich analytische Ausdrücke direkt anzugeben.

Das Matrixeigenwertproblem (5.21) ist nun iterativ selbstkonsistent zu lösen, da die Potentiale ihrerseits von der Elektronendichte $n(\mathbf{G})$ abhängen, welche über die Lösungen $c_{n\mathbf{k}}(\mathbf{G})$ gegeben ist.

5.1.1. Diskrete Integration über die Brillouinzone

Brillouinzoneintegrale der Form

$$I = \int_{\text{BZ}} f(\mathbf{k}) d^3k = \frac{(2\pi)^3}{\Omega} \bar{f} \tag{5.25}$$

**Abb. 5.1:**

Die Methode zur \mathbf{k} -Punkt Verteilung nach Monkhorst und Pack: Die Brillouinzone (rot umrandet) wird äquidistant aufgeteilt. Gezeigt ist in dieser Skizze ein 4x4 Gitter. Mithilfe eines Satzes von erzeugenden \mathbf{k} -Punkten (hier ein einzelner Punkt mit den Koordinaten $(\frac{1}{2}\frac{1}{2})$) wird das Gitter aufgebaut. Dabei ist darauf zu achten, dass Hochsymmetriepunkte sowie die Brillouinzonengrenze ausgespart bleiben. Die problemspezifischen Kristallsymmetrien können an dieser Stelle verwendet werden um die Anzahl der \mathbf{k} -Punkte zu reduzieren.

sind ein Standardproblem im Impulsraumformalismus. Unter der begründeten Annahme, dass die Wellenfunktionen langsam mit \mathbf{k} variieren, gilt der Mittelwertsatz der Integralrechnung. Aus einer geeigneten Näherung des Mittelwertes \bar{f} kann das Integral numerisch effizient ausgewertet werden, ohne die komplette Brillouinzone abscannen zu müssen. Mögliche Verfahren zur Approximation dieses Mittelwertes sind unter anderen von Baldereschi, sowie von Chadi und Cohen vorgeschlagen worden [53, 54]. Dabei wird das Integral über einen geeigneten Satz von \mathbf{k} -Punkten gemäss

$$I = \int_{\text{BZ}} f(\mathbf{k}) d\mathbf{k} \approx \frac{(2\pi)^3}{\Omega} \sum_i w_i f(\mathbf{k}_i) \quad (5.26)$$

mit

$$\sum_i w_i = 1 \quad (5.27)$$

approximiert. Heutzutage findet das Verfahren von Monkhorst und Pack am meisten Verwendung. Hier wird der Hauptwert \bar{f} über ein äquidistantes Netz von \mathbf{k} -Punkten genähert [55] (vgl. Abbildung 5.1).

5.2. Die Methode der Pseudopotentiale

Die Behandlung von Vielteilchensystemen erfolgt in der Praxis numerisch und ist ein nichtlineares Problem. Typische Dichtefunktionalalgorithmen mit einer Ebenen-Wellen-Basis skalieren quadratisch bis kubisch mit der Anzahl der ebenen Wellen. Diese wachsen linear mit der Anzahl der Atome und müssen die Wellenfunktionen adäquat nachbilden können. Problematisch ist hierbei das starke Coulombpotential der Kerne, welches zur Lokalisierung der Wellenfunktion führt. Die Orthogonalitätsforderung der Wellenfunktionen bedingt starke Oszillationen in Kernnähe. Um diese Oszillationen darstellen zu können, bedarf es einer großen Anzahl von Ebenen-Wellen-Basisfunktionen, wodurch die Behandlung aufwendig und speicherintensiv wird. Diese Probleme werden in heutigen Implementierungen der Dichtefunktionaltheorie durch die Verwendung der *unveränderlichen Ionennäherung*, sowie der Einführung von *Pseudopotentialen*, die in der Kernregion glatte Wellenfunktionen liefern, überwunden. Diese Konzepte werden nun vorgestellt.

5.2.1. Die Näherung der unveränderlichen Ionen

In der großen Gesamtheit von molekularen Systemen und Festkörpern wird die chemische Bindung durch die Valenzelektronen dominiert. Die Rumpfelektronen spielen dabei eine eher untergeordnete Rolle. Dies motiviert die Frage, ob es nicht zulässig ist die an der Bindung nur schwach teilnehmenden Rumpfelektronen durch ihre atomaren Wellenfunktionen zu beschreiben und in der elektronischen Optimierung allein die Valenzelektronen zu variieren. Dies führt zur Näherung der unveränderlichen Ionen: Man teilt die Elektronendichte in zwei räumliche Bereiche. Der Rumpfbereich enthält die N_{core} Rumpfelektronen mit der Rumpfdichte

$$n_{\text{core}} = \sum_1^{N_{\text{core}}} |\phi_i(\mathbf{r})|^2, \quad (5.28)$$

der Valenzbereich erhält die für die Bindung interessanten Valenzelektronen und somit die Valenzdichte

$$n_{\text{valence}} = \sum_{N_{\text{core}}+1}^N |\phi_i(\mathbf{r})|^2. \quad (5.29)$$

Nun findet eine Abseparation des Kernbereichs in den Kohn-Sham-Gleichungen gemäß

$$v_{\text{eff}}[n](\mathbf{r}) = v_H[n_{\text{valence}}](\mathbf{r}) + v_{\text{xc}}[n_{\text{valence}}](\mathbf{r}) + v_{\text{core}}[n_{\text{valence}}, n_{\text{core}}](\mathbf{r}) \quad (5.30)$$

statt. Dabei sei v_{core} definiert als

$$v_{\text{core}}[n_{\text{valence}}, n_{\text{core}}](\mathbf{r}) = v_H[n_{\text{core}}](\mathbf{r}) + v_{\text{xc}}[n_{\text{valence}} + n_{\text{core}}](\mathbf{r}) - v_{\text{xc}}[n_{\text{valence}}](\mathbf{r}) + v(\mathbf{r}). \quad (5.31)$$

Hier findet nun die Näherung $n_{\text{core}} \approx n_{\text{core}}^{\text{atom}}$

$$v_{\text{core}}[n_{\text{valence}}](\mathbf{r}) = v_{\text{H}}[n_{\text{core}}^{\text{atom}}](\mathbf{r}) + v_{\text{xc}}[n_{\text{valence}} + n_{\text{core}}^{\text{atom}}](\mathbf{r}) - v_{\text{xc}}[n_{\text{valence}}](\mathbf{r}) + v(\mathbf{r}). \quad (5.32)$$

statt, in der die tatsächliche Rumpfdichte durch die Superposition atomarer Rumpfdichten

$$n_{\text{core}}^{\text{atom}}(\mathbf{r}) = \sum_1^{N_{\text{core}}} |\phi_i^{\text{atom}}(\mathbf{r})|^2 \quad (5.33)$$

genähert wird. Durch diese Näherung ist das effektive Kernpotential v_{core} lediglich ein Funktional der Valenzelektronen. Die Rumpfelektronendichte n_{core} wird durch die atomare Rumpfdichte $n_{\text{core}}^{\text{atom}}$ ersetzt und variationell „eingefroren“. Dies bezeichnet man als *frozen core approximation* [56]. Die Kohn-Sham Gleichungen

$$\left(-\frac{\hbar^2}{2m_e} + v_{\text{H}}[n_{\text{valenz}}](\mathbf{r}) + v_{\text{xc}}[n_{\text{valence}}](\mathbf{r}) + v_{\text{core}}(\mathbf{r}) \right) |\phi_i(\mathbf{r})\rangle = \epsilon_i |\phi_i(\mathbf{r})\rangle \quad (5.34)$$

enthalten nun nur noch Funktionale der Valenzelektronendichte. Bei der selbstkonsistenten Bestimmung von n_{valenz} ist allerdings nach wie vor zu beachten, dass die Orthogonalität der Valenzorbitale mit den atomaren Rumpforbitalen erhalten bleibt.

Der Vorteil dieses Verfahrens liegt in der deutlichen Reduktion der Systemgröße, da nur noch die Valenzelektronen — je nach Atomsorte in der Größenordnung 10 pro Atom — in die selbstkonsistente Behandlung des Problems einfließen. Die Abweichungen einer Rechnung mit festgehaltenem Kernbereich im Vergleich zu einem explizit behandeltem Kernbereich in den einzelnen Energiebeiträgen sind zwar eklatant, jedoch hat eine Studie gezeigt, dass sich diese Fehler im Mittel gegenseitig aufheben und Gesamtenergien in der Regel Abweichungen von nur einigen 10 meV zeigen [56].

5.2.2. Das Konzept der Pseudopotentiale

Mittels des Konzeptes der unveränderlichen Ionen ist zwar die Anzahl der zu behandelnden Elektronen deutlich verringert worden, jedoch bleiben die in Kernnähe starken Oszillationen, die einer großen Basis für ihre Beschreibung bedürfen, bestehen. Unter der Annahme, dass zur chemischen Bindung nur die Teile der Wellenfunktion beitragen, die sich am Ort der Bindung befinden, können die stark oszillierenden Anteile durch eine glatte Wellenfunktion ersetzt werden. Dabei wird das im vorherigen Abschnitt eingeführte Kernpotential durch eines ersetzt, welches hinreichend glatt ist und für $r \rightarrow 0$ nicht singular wird. Die genaue Art der Konstruktion solcher Potentiale kann mit Hilfe der gängigen Literatur [57, 58, 59, 60, 61]

nachvollzogen werden.

Insgesamt gilt es vier wohldefinierte Bedingungen zu erfüllen, um dieses Vorgehen zu legitimieren[62].

1. Die durch das Pseudopotential generierten Pseudowellenfunktionen sind in ihrem radialen Anteil R_{nl}^{PP} knotenfrei.
2. Für einen vorher gewählten Kernradius r_c , der in der Regel zwischen dem letzten Knoten und dem folgenden Maximum des Radialteils der Allelektronenwellenfunktion zu wählen ist[61], ist die Pseudowellenfunktion identisch zu der Allelektronenwellenfunktion

$$R_{nl}^{\text{PP}}(\mathbf{r}) = R_{nl}^{\text{AE}}(\mathbf{r}) \quad \text{für } r > r_{cl}. \quad (5.35)$$

3. Die im Bereich des Cutoffabstandes r_c eingeschlossene Ladung muss für beide Wellenfunktionen identisch sein

$$\int_0^{r_{cl}} |R_{nl}^{\text{PP}}(r)|^2 r^2 dr = \int_0^{r_{cl}} |R_{nl}^{\text{AE}}(r)|^2 r^2 dr. \quad (5.36)$$

4. Die Orbitalenergien ϵ_i sind identisch

$$\epsilon_{nl}^{\text{PP}} = \epsilon_{nl}^{\text{AE}}. \quad (5.37)$$

In der gängigen Praxis werden Pseudopotentiale dadurch erzeugt, dass zunächst die Energien, Wellenfunktionen und das atomare Potential der Atomsorte über die explizite Lösung der Kohn-Sham-Gleichung in Form einer Allelektronenrechnung beschrieben werden. Anhand der Allelektronenwellenfunktionen werden nun die Kernradien r_{cl} definiert. Hierbei gilt: Je kleiner der Radius, umso ähnlicher ist die Pseudowellenfunktion der Allelektronenwellenfunktion. Andererseits steigt mit kleinem Kernradius auch die Anzahl der benötigten Basisfunktionen, so dass ein entsprechender Mittelweg gefunden werden muss. Sind die r_{cl} festgesetzt, wird die Allelektronenvalenzwellenfunktion so variiert, dass sie im Kernbereich knotenfrei wird. Dabei ist die Erhaltung der Ladung im Kernbereich einzuhalten. Die Variation geschieht hierbei unter Verwendung von analytischen Fits. Aus der Kenntnis der Eigenwerte und der neuen Wellenfunktion kann nun die Schrödingergleichung invertiert und das zugehörige Potential bestimmt werden.

Neben der Methode der normerhaltenen Pseudopotentiale haben sich an bestimmte Problemstellungen angepasste Varianten entwickelt. So gibt es Pseudopotentiale, die neben den obigen Bedingungen Stetigkeit und stetige Differenzierbarkeit bis hin zur zweiten Ableitung des Potentials verlangen, sowie solche, die eine verschwindende

Krümmung am Kern fordern[57].

Von besonderer Schwierigkeit sind die Elemente der ersten Reihe des Periodensystems. Aufgrund der starken Lokalisierung der Zustände in Kernnähe ist der Kernradius sehr gering zu wählen. Eine wirkliche Reduzierung der Basisfunktionen mittels normerhaltenden Pseudopotentialen ist damit nicht möglich. Von Vanderbilt wurden sogenannte *ultrasofte Pseudopotentiale* eingeführt, welche auf die Normerhaltung im Kernbereich verzichten[60]. Der Verlauf der Wellenfunktion kann somit deutlich weicher modelliert werden und sorgt damit für eine Verringerung der notwendigen Basisfunktionen.

Pseudopotentiale sind aber dennoch mit Vorsicht zu genießen und ihre Aussagekraft immer an wohldefinierten physikalischen Referenzen zu erproben, da sie in ihrer Konstruktion sämtlichen Informationsgehalt zum physikalischen Kernbereich verlieren. Eine bessere Alternative stellen die im nächsten Kapitel diskutierten PAW-Potentiale dar.

5.2.3. Die Projector Augmented Wave (PAW) Methode

Von Blöchl[63] stammt die als *Projector Augmented Wave* (PAW) Methode bekannte Erweiterung der Pseudopotentiale, welche sich das Ziel setzt eine exakte Abbildung zwischen den pseudoisierten Wellenfunktionen (PS) und den Allelektronenwellenfunktionen (AE) zu schaffen. Im Gegensatz zu der Pseudopotentialmethode stellt also die PAW-Methode im Rahmen der Näherung der unveränderlichen Ionen eine trickreiche Allelektronenrechnung dar. Die Grundideen werden im Folgenden skizziert.

Basis der PAW-Methode ist die Existenz eines Operators \mathcal{T} , welcher die in Kernnähe glatte Pseudowellenfunktion $|\Psi^{\text{PS}}\rangle$ in die oszillierende, effektive Kohn-Sham-Einteilchenwellenfunktion $|\Psi^{\text{AE}}\rangle$ überführt

$$|\Psi^{\text{AE}}\rangle = \mathcal{T}|\Psi^{\text{PS}}\rangle \quad (5.38)$$

Diese Transformation ist dahingehend interessant, dass sie analog zum Wechsel von dem Schrödinger Bild zum Heisenberg Bild die Berechnung des Erwartungswertes $\langle A \rangle$ eines Operators A nun auf zwei Arten erlaubt: Zum einem direkt über die Kohn-Sham-Einteilchenwellenfunktionen, zum anderen über die Einführung eines Pseudooperators \tilde{A} der über

$$\begin{aligned} \langle A \rangle &= \langle \Psi^{\text{AE}} | A | \Psi^{\text{AE}} \rangle \\ &= \langle \Psi^{\text{PS}} | \tilde{A} | \Psi^{\text{PS}} \rangle \quad \text{mit } \tilde{A} = \mathcal{T}^\dagger A \mathcal{T} \end{aligned} \quad (5.39)$$

definiert ist. So kann die PAW-Methode mit minimalem Aufwand durch Neudefinition der Wirkung eines Projektors auf die Pseudowellenfunktion in bestehende

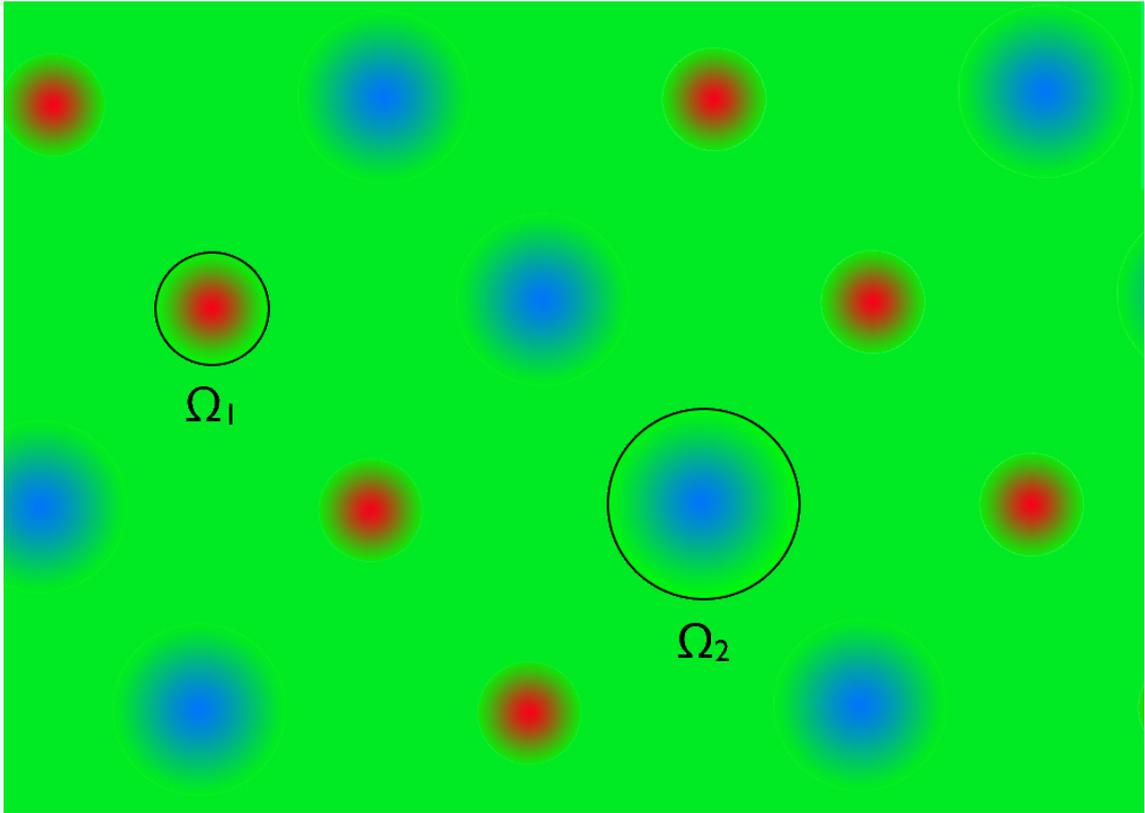


Abb. 5.2:

Separation des atomaren Umgebung in Valenzbereich und Kernbereich. Dargestellt ist ein regelmäßiges atomares Gitter zweier Atomsorten. Die Bereiche Ω_1 und Ω_2 stellen den Kernbereich der Atome dar. Ausserhalb dieses Bereiches variiert die Wellenfunktion langsam und verläuft glatt.

DFT-Codes, welche auf der Pseudopotentialmethode basieren, implementiert werden. Dazu ist lediglich der Projektionsoperator \mathcal{T} zu definieren.

Wie schon in der Pseudopotentialmethode wird der Raum in zwei Bereiche unterteilt, den Bereich in dem hauptsächlich die atomare Bindung stattfindet und die Wellenfunktionen glatt verlaufen und dem, in dem sich das oszillierende Verhalten der Wellenfunktion auf Grund der Kernnähe abspielt, der sogenannte *Augmentationsbereich* (vgl. Abbildung 5.2). Da sich die Pseudowellenfunktion und die Kohn-Sham-Einteilchenwellenfunktion nur in diesem Augmentationsbereich unterscheiden, kann der Operator \mathcal{T} als Identität dargestellt werden, der gemäß

$$\mathcal{T} = \mathbf{1} + \sum_R \mathcal{T}_R \quad (5.40)$$

additive Korrekturen im Augmentationsbereich Ω_R erfährt. Die Allelektronenlösung $|\phi_i^{\text{AE}}\rangle$ — von Blöchl als Partialwelle bezeichnet — ergibt sich dementsprechend aus einer computertechnisch gut zu behandelnden pseudoisierten Partialwelle $|\phi_i^{\text{PS}}\rangle$ als

$$|\phi_i^{\text{AE}}\rangle = (1 + \mathcal{T}_R)|\phi_i^{\text{PS}}\rangle. \quad (5.41)$$

Insbesondere kann nun jede Pseudowellenfunktion $|\Psi^{\text{PS}}\rangle$ in die pseudosierten Partialwellen entwickelt werden

$$|\Psi^{\text{PS}}\rangle = \sum_i |\phi_i^{\text{PS}}\rangle c_i \quad \text{innerhalb von } \Omega_R. \quad (5.42)$$

Aufgrund der Transformationseigenschaften \mathcal{T} gilt dies aber auch für die Wellenfunktion $|\Psi^{\text{AE}}\rangle$ und die Partialwellen

$$|\Psi^{\text{AE}}\rangle = \sum_i |\phi_i^{\text{AE}}\rangle c_i \quad \text{innerhalb von } \Omega_R \quad (5.43)$$

mit den gleichen Entwicklungskoeffizienten c_i . Blöchl führt an dieser Stelle Projektorfunktionen $|p_i\rangle$ ein, welche die Eigenschaften $\langle p_i|\Psi^{\text{PS}}\rangle = c_i$ besitzen und zudem auf Grund der Vollständigkeitsrelation innerhalb des Augmentationsbereiches die Eigenschaft $\langle p_i|\phi_j^{\text{PS}}\rangle = \delta_{ij}$ erfüllen. Mit Hilfe dieser Projektorfunktionen kann \mathcal{T} definiert werden:

$$\begin{aligned} |\Psi^{\text{AE}}\rangle &= \Psi^{\text{PS}} - \sum_i |\phi_i^{\text{PS}}\rangle c_i + \sum_i |\phi_i^{\text{AE}}\rangle c_i \\ &= \Psi^{\text{PS}} + \sum_i (|\phi_i^{\text{AE}}\rangle - |\phi_i^{\text{PS}}\rangle) \langle p_i|\Psi^{\text{PS}}\rangle \\ &= \underbrace{\left(1 + \sum_i (|\phi_i^{\text{AE}}\rangle - |\phi_i^{\text{PS}}\rangle) \langle p_i|\right)}_{\mathcal{T}} |\Psi^{\text{PS}}\rangle \end{aligned} \quad (5.44)$$

5.3. LCAO Initialisierung

Der iterative Lösungsansatz zur akkuraten Beschreibung der elektronischen Struktur bedingt im ersten Schritt schon ein effektives Potential und somit eine Initialisierung der Wellenfunktionen und der Dichte. Rechnungen aus dem Alltag zeigen, dass Algorithmen zur Minimierung der Gesamtenergie in ihrer Funktionsweise erheblich verbessert werden, wenn eine gute Initialisierung der Wellenfunktion erfolgt. Zufallszahlen liefern für metallische Systeme, in denen die Elektronendichte nahezu gleich verteilt im Kristall vorliegt sicherlich eine gute Initialisierung. Im Falle von Systemen, die durch gerichtete Bindungen dominiert werden, ist ein solcher Ansatz jedoch denkbar schlecht. Hier lokalisieren die Elektronen in Bindungorbitalen, die es zu beschreiben gilt.

Eine Möglichkeit besteht hierbei diese Bindungorbitale in eine atomzentrierte Orbitalbasis zu entwickeln[64]. Da im Rahmen der Pseudopotentialmethode die Elek-

tronenverteilung für die Pseudoatome bestimmt worden ist, bieten sich die dort berechneten Pseudoatomorbitale $|\mu_i\rangle$ als Basis an. Definiert man nun durch

$$H_{ij} = \langle \mu_i | \hat{H} | \mu_j \rangle \quad (5.45)$$

die Hamiltonmatrix des Systems und durch

$$S_{ij} = \langle \mu_i | \mu_j \rangle \quad (5.46)$$

den Überlappoperator der Pseudoatomorbitale, so ergibt sich das verallgemeinerte Eigenwertproblem

$$H|c_n\rangle = \epsilon_n S|c_n\rangle. \quad (5.47)$$

Da S eine positiv definite, hermitesche Matrix ist, existiert von dieser die Choleskyzerlegung $S = LL^\dagger$, so dass sich das verallgemeinerte Eigenwertproblem in das gewöhnliche Eigenwertproblem

$$H'|\chi_n\rangle = \epsilon_n|\chi_n\rangle \quad (5.48)$$

mit

$$H' = L^{-1}H(L^\dagger)^{-1} \text{ und } |\chi_n\rangle = L^\dagger|c_n\rangle \quad (5.49)$$

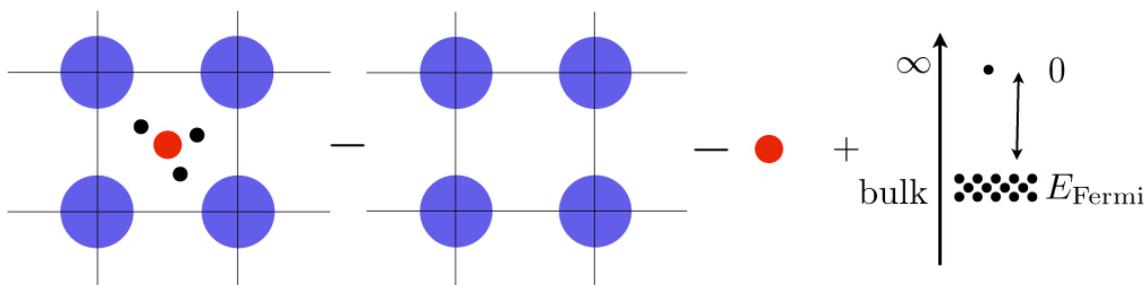
transformiert. Dessen Eigenvektoren können als gute Initialisierung der Wellenfunktion verwendet werden. Auch die in dieser Dissertation in Kapitel 6 diskutierten optimierten Basisfunktionen können mit Vorteil als Initialisierungsfunktionen verwendet werden. Da in diesen Orbitalen schon Kenntnisse über die atomaren Nachbarn verankert sind, konvergieren die mit Quamols initialisierten Rechnungen schneller und sparen erfahrungsgemäß in dem verwendeten S/PHI/nX Code bis zu 10 iterative Schritte in der elektronischen Konvergenz. Normalerweise werden für die elektronische Konvergenz der hier betrachteten Systeme 15-35 Schritte beobachtet.

5.4. Bildungsenergien von Defekten

Die Bildungsenergie eines Defektes ist die zentrale Größe bei der theoretischen Analyse von Punktdefekten. Über jene lassen sich Defekte bezüglich ihrer Stabilität vergleichen und Defektkonzentrationen abschätzen. Die Bildungsenergie E^f ist über

$$E_{X^q}^f = E_{\text{tot}}[\text{Bulk}:X^q] - E_{\text{tot}}[\text{Bulk}] - \sum_i n_i \mu_i + q(\epsilon_{\text{Fermi}} + E_{\text{VBM}}) \quad (5.50)$$

definiert[4]. Dieser Ausdruck ist vornehmlich eine Energiedifferenz des modellierten Defektes und seiner Edukte: des vorherrschenden Wirtssystems, sowie die zum Defekt notwendigen addierten oder subtrahierten Atome und Ladungen. Die einzelnen Bestandteile der Bildungsenergie sind in Abbildung 5.3 noch einmal

**Abb. 5.3:**

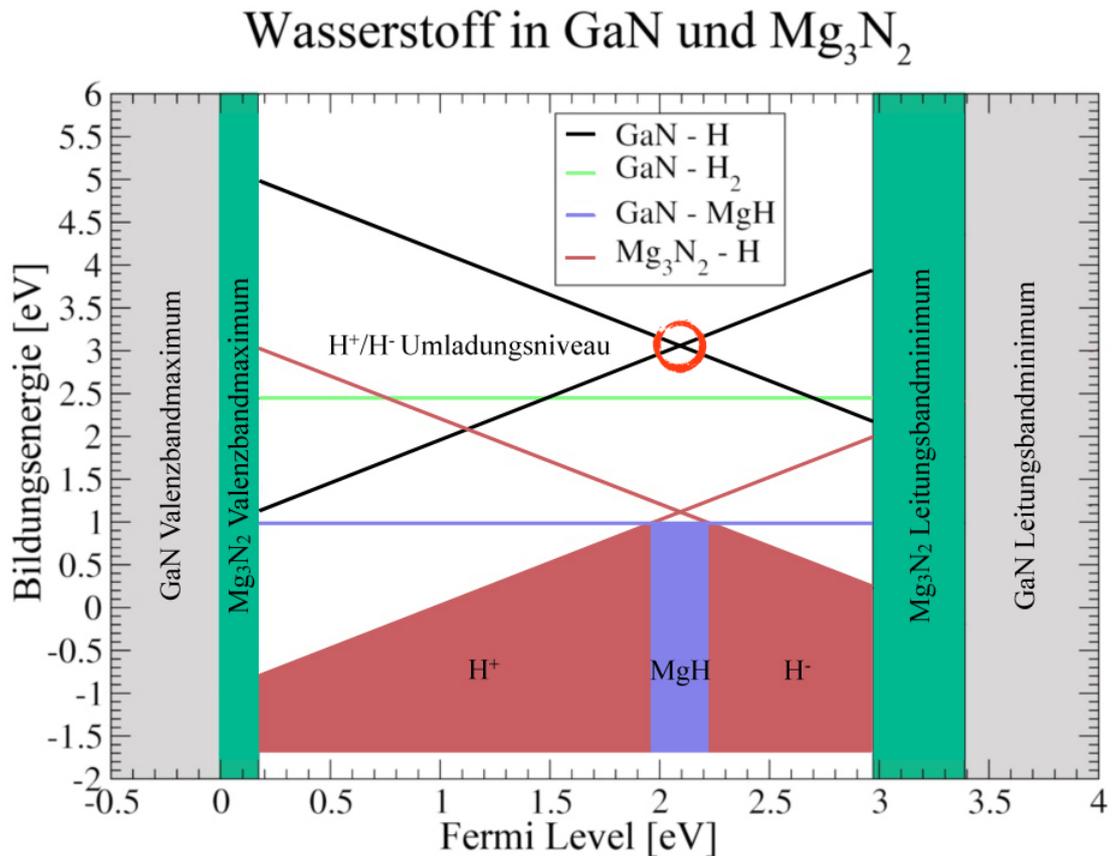
Die Bildungsenergie ist definiert als Differenzenergie des Defektes und seiner Edukte: Hostsystem (blau), Defektatom (rot) und zusätzliche Elektronen (schwarz).

zusammengefasst.

Während die Gesamtenergien einzelner Geometrien durch das zu Grunde liegenden DFT-Programm ermittelt werden, ist die Wahl des chemischen Potentials der Atomreservoirs nicht eindeutig bestimmt. Diese können über Rechnungen einzelner Atome, Gasphasenmoleküle oder über mittlere Energien einzelner Atome in stabilen Phasen abgeschätzt werden. An dieser Stelle erlaubt die Theorie eine Anknüpfung an die im Experiment realisierten Bedingungen (wie z.B. Temperatur, Partialdrücke der einzelnen Elemente, usw.).

Im Falle geladener Defekte bedarf es eines Reservoirs der Elektronen. Das chemische Potential ist über die Fermienergie im System gegeben, welche nach allgemeiner Konvention auf das Valenzbandmaximum E_{VBM} referenziert wird. Die Lage des Fermi-niveaus in der Bandlücke wird über die Gesamtheit der Defekte im Wirtssystem bestimmt, so dass die Berechnung eines einzelnen Defektes nicht ausreicht um dessen Position des Fermi-niveaus im Gesamtsystem zu erhalten. Aus diesem Grund wird die Bildungsenergie eines Defektes in Abhängigkeit zur Lage des Fermi-niveaus aufgetragen, wie es Abbildung 5.4 zeigt. In solchen Diagrammen kann in Abhängigkeit des Fermi-niveaus die Stabilität einzelner Defekte bewertet werden. Die Steigung der einzelnen Graden gibt dabei den Ladungszustand des entsprechenden Defektes an. Kreuzen sich gleiche Defektsorten zu unterschiedlichen Ladungen, so heisst dieser Punkt *Umladungsniveau*.

Da die Bandlücke in den theoretischen Rechnungen massgeblich unterschätzt wird, ist in solchen Diagrammen häufig die experimentelle Bandlücke eingetragen. Dies ist insofern mit Vorsicht zu geniessen, da ein Fermi-niveau oberhalb der theoretischen Leitungsbandkante zu Besetzungen der Leitungsbander führt und die berechneten Energien nicht vertrauenswürdig sind. In praktischen Rechnungen kann hier aber die diskrete Beschreibung des \mathbf{k} -Punktgitters aushelfen, indem \mathbf{k} -Punkte niedrig liegender Leitungsbandzustände ausgespart werden und so die Rechnung blind für die tatsächliche theoretische Bandlücke ist.

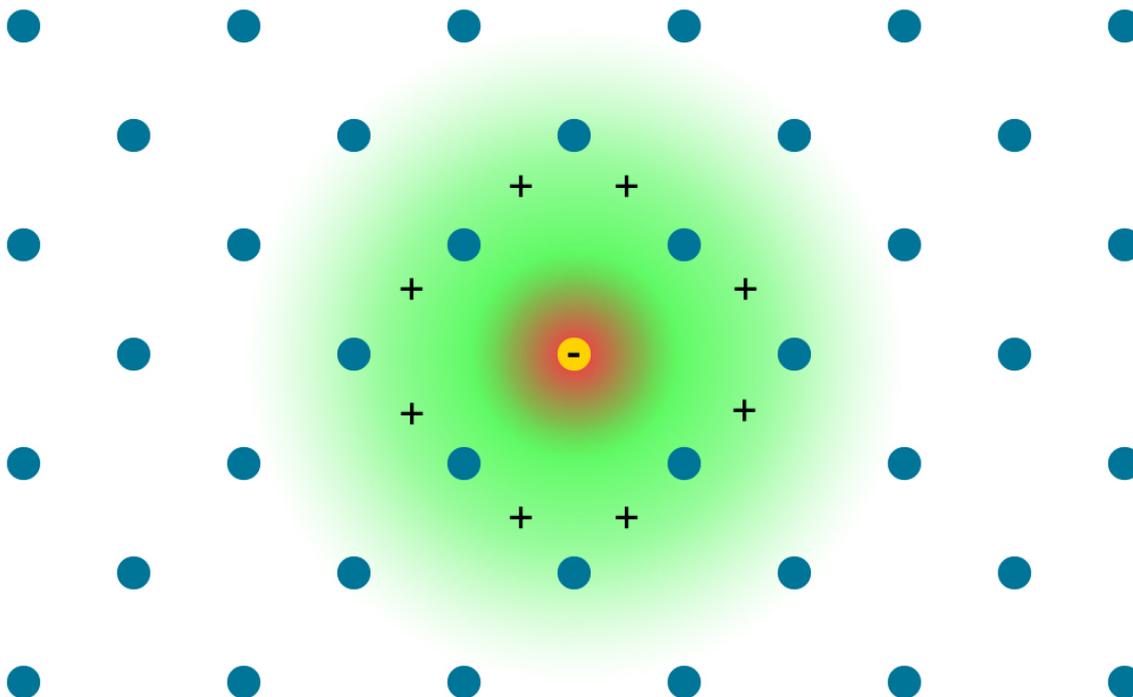
**Abb. 5.4:**

Die Bildungsenergien von Wasserstoffverunreinigungen in GaN und Mg_3N_2 als Funktion des Fermi-niveaus. Der rote Kreis markiert das Umladungsniveau H^+/H^- in GaN. Die farbigen Flächen im unteren Bereich des Graphen geben die Defekte geringster Bildungsenergie wieder.

5.4.1. Superezellenkorrektur geladener Defekte

Versucht man in einem Superezellenansatz einen geladenen Defekt zu beschreiben, so ist es notwendig die Ladung q des Defektes durch einen homogenen Ladungshintergrund zu kompensieren, da ansonsten die elektrostatische Energie divergiert. Durch die langreichweitige Coulombwechselwirkung, interagiert der Defekt mit seinen periodischen Bildern und dem homogenen Ladungshintergrund, so dass die Formationsenergie in Abhängigkeit zur Superezellenlänge L nur sehr langsam konvergiert[65].

Folglich sind in einem direkten Ansatz sehr große Superezellen von Nöten um die Formationsenergie des isolierten Defektes bestimmen zu können. Damit die Zellen berechenbar bleiben, wurden viele Versuche unternommen, um die Wechselwirkungsenergie des Defektes mit seinen periodischen Bildern und den elektrischen homogenen Hintergrund zu bestimmen und nachträglich zu korrigieren. Eine erste Abschätzung kann über die Madelungenergie

**Abb. 5.5:**

Ein geladener Defekt in einem Festkörper akkumuliert Ladung entgegengesetzten Vorzeichens, welche sein Coulombpotential abschwächt.

$$E_{\text{Madelung}} = \frac{\alpha q^2}{2\epsilon L} \quad (5.51)$$

gemacht werden. Dabei bezeichnet α die von der Superzelle abhängige Madelungkonstante.

Makov und Payne berechneten für ein Gitter von isolierten Ionen einen weiteren Korrekturterm, der wie L^{-3} skaliert. In realen Systemen ist diese Korrektur dann noch durch die dielektrische Konstante zu skalieren. Jedoch führt dies nicht immer zum gewünschten Erfolg, sodass im praktischen Vorgehen die Vorfaktoren wie freie Parameter behandelt und zudem noch weitere Korrekturterme L^{-n} hinzugefügt werden. Dabei geht jedoch ein beachtlicher Teil Physik verloren.

Interessant ist daher der Ansatz von Freysoldt[66, 67], der ein Korrekturschema auf elementaren elektrostatischen Überlegungen aufbaut. Ein geladener Defekt wird von einer kompensierenden Ladungsakkumulation in seinem Potential abgeschwächt (vgl. Bild 5.5). Sein langreichweitiges Potential ist dann über

$$v^{\text{lang}}(\mathbf{r}) = \frac{1}{\epsilon} \int \frac{n^{\text{model}}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3\mathbf{r}' \quad (5.52)$$

in atomaren Einheiten gegeben. Die in diesem Ausdruck eingeführte Ladungsdichte n^{model} kann für lokalisierte Defektladungen durch eine Gaußladung modelliert werden. Das komplette Defektpotential $v^d(\mathbf{r})$ wird dann als Summe des langreichweitigen Teils und einer kurzreichweitigen Modulation

$$v^d(\mathbf{r}) = v^{\text{lang}}(\mathbf{r}) + v^{\text{kurz}}(\mathbf{r}) \quad (5.53)$$

geschrieben. Praktisch kann das Defektpotential als Differenz der elektrostatischen Potentiale

$$v^d(\mathbf{r}) = v^{\text{el}}[X^q](\mathbf{r}) - v^{\text{el}}[X^0](\mathbf{r}) \quad (5.54)$$

berechnet werden. Das kurzreichweitige Potential ist damit

$$v^{\text{kurz}}(\mathbf{r}) = v^{\text{el}}[X^q](\mathbf{r}) - v^{\text{el}}[X^0](\mathbf{r}) - v^{\text{lang}}(\mathbf{r}) + \delta v. \quad (5.55)$$

Der Offset δv sorgt dafür, dass $v^{\text{kurz}}(\mathbf{r})$ für große Entfernungen vom Defektzentrum gegen Null abklingt. Aus dem langreichweitigen Verhalten kann nun die Madelungenergie $E^{\text{Gitter}}[n^{\text{model}}]$ bestimmt werden. Der kurzreichweitige Teil bestimmt die Wechselwirkungsenergie mit dem kompensierenden Hintergrund

$$\int_{\Omega} -\frac{q}{\Omega} v^{\text{kurz}}(\mathbf{r}) d^3\mathbf{r} = -qE^{\text{WW}}. \quad (5.56)$$

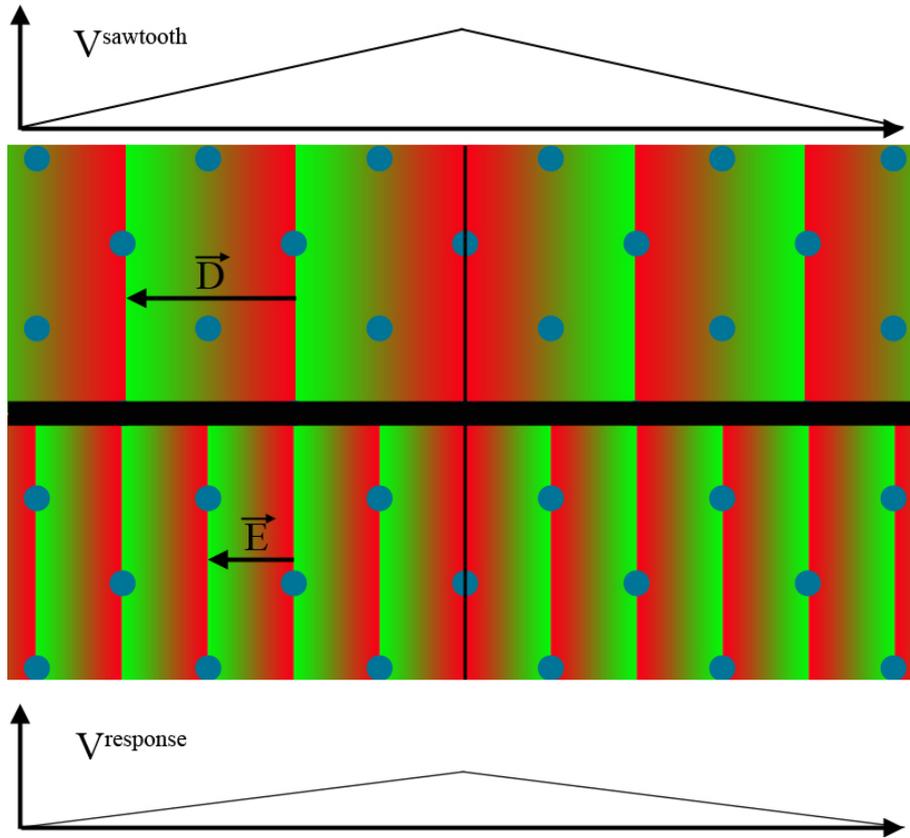
Nach Freysoldt[67] herrscht nun die Identität $E^{\text{WW}} = \delta v$, so dass sich die Bildungsenergie des isolierten Defektes entsprechend Gleichung (5.50) zu

$$\begin{aligned} E_{X^q}^f &= E_{\text{tot}}[\text{Bulk} : X^q] - E_{\text{tot}}[\text{Bulk}] - \sum_i n_i \mu_i + q(E_{\text{Fermi}} + E_{\text{VBM}}) \\ &\quad - E^{\text{Gitter}}[q^{\text{model}}] + q\delta v \end{aligned} \quad (5.57)$$

ergibt. Das Trickreiche an der Identität $E^{\text{WW}} = \delta v$ ist ihre Unabhängigkeit von dem Referenzsystem, so dass in der Praxis die Bildungsenergie des isolierten geladenen Defektes allein aus der Superzellenrechnung des Bulks und des geladenen Defektes bestimmt werden kann.

5.4.2. Bestimmung der dielektrischen Konstanten

Die Anwendung des eben diskutierten Korrekturverfahrens zur Berechnung der Bildungsenergie von isolierten geladenen Defekten setzt die Kenntnis der dielektrischen Konstante ϵ voraus. Diese kann zum Beispiel aus dem Experiment genommen werden. Konsistenter ist es jedoch die Konstante für das modellierte Material theoretisch zu bestimmen. Dazu dient das folgende Computerexperiment.


Abb. 5.6:

Entlang einer Kristallachse wird ein sägezahnförmiges Potential V^{sawtooth} angelegt, welches einem konstanten Erregungsfeld \mathbf{D} entspricht. Der Kristall antwortet auf dieses Feld über die Ausprägung eines elektrischen Feldes \mathbf{E} , welches im elektrostatischen Potential als Sägezahnpotential wahrgenommen werden kann. Die farbigen Verläufe symbolisieren hierbei die Feldvektoren mit ihrer Richtung von rot nach grün. Das Verhältnis der Feldamplituden entspricht gerade der dielektrischen Konstanten.

Setzt man einen Festkörper einer konstanten elektrischen Erregung \mathbf{D} aus, so regiert dieser durch den Aufbau des elektrischen Feldes \mathbf{E} . Diese beiden Felder sind über die Materialgleichung

$$\mathbf{D} = \epsilon_0 \mathbf{E} + \mathbf{P} \quad (5.58)$$

miteinander verknüpft. Im Falle eines isotropen und linearen Kristalls ist die Polarisierbarkeit \mathbf{P} als

$$\mathbf{P} = \epsilon_0 \chi \mathbf{E} \quad (5.59)$$

gegeben. Damit vereinfacht sich die Materialgleichung zu

$$\mathbf{D} = \epsilon_0 (1 + \chi) \mathbf{E} = \epsilon \mathbf{E}, \quad (5.60)$$

so dass sich die dielektrische Konstante als $\epsilon = |\mathbf{D}|/|\mathbf{E}|$ ergibt. Ändert man die

Materialsystem	dielektrische Konstante $[\frac{\text{As}}{Vm}]$	
	ϵ	ϵ_∞
GaN (LDA)	10.3	5.6
GaN (PBE)	11.8	6.2
GaN (exp)	10.4 ^a	5.8 ^b
Mg ₃ N ₂ (LDA)	13.6	5.1
Mg ₃ N ₂ (PBE)	9.9	5.2
Mg ₃ N ₂ (exp)	—	4.45 ^c

^a[68], ^b[69], ^c[70]

Tab. 5.1:

Dielektrische Konstanten für die Materialsysteme GaN und Mg₃N₂. Die theoretischen Werte sind mit dem im Text beschriebenen Verfahren ermittelt worden. ϵ_∞ entspricht der dielektrischen Konstanten im Falle hochfrequenter Wechselfelder, so dass die Ionen der elektrischen Anregung nicht folgen ($\omega_\infty \gg \omega_{\text{Phonon}}$).

Orientierung der elektrischen Erregung mit der Zeit, so können die Ionen diesem sich ändernden Feld ab einer Frequenz ω_∞ ($\omega_\infty \gg \omega_{\text{Phonon}}$) nicht mehr folgen und die elektrische Abschirmung wird allein von den Elektronen bestimmt. Dies entspricht der dielektrischen Konstanten ϵ_∞ .

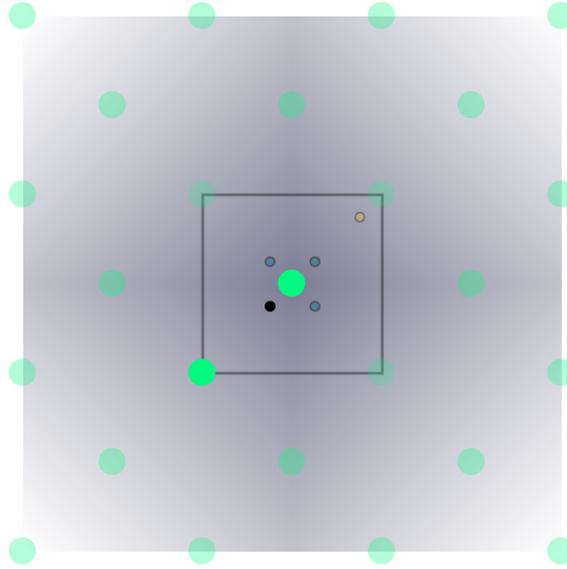
Dieses Experiment kann auch im Rahmen der Dichtefunktionaltheorie vollzogen werden. Dabei wird in die Superzelle entlang einer Hauptachse ein Sägezahnpotential induziert (siehe Abbildung 5.6). Die Elektronendichte und die Ionen reagieren entsprechend und bilden ein entsprechendes elektrostatisches Potential aus, welches auch die Sägezahncharakteristik zeigt. Aus den Steigungen der jeweiligen Flanken können die Stärken der elektrischen Erregung und des resultierenden elektrischen Feldes extrahiert werden. Zur Bestimmung der rein elektronischen Abschirmung werden die Ionen in der DFT Rechnung fixiert und dürfen nicht auf Grund des angelegtem Potentials relaxieren.

Für die in dieser Dissertation relevanten Materialsysteme GaN und Mg₃N₂ wurden nach diesem Verfahren die dielektrischen Konstanten bestimmt (Tabelle 5.1).

5.5. Berechnung der Defektkonzentration im thermischen Gleichgewicht

Die Fermienergie in Gleichung (5.50) ist ein freier Parameter, jedoch in der Gesamtheit des Systems implizit über die Ladungsneutralität

$$c(h) - c(e) + \sum_{X^q} qc(X^q) = 0 \quad (5.61)$$


Abb. 5.7:

Ein Defekt (schwarz) wird in einem 2-dimensionalen Kristall (grün) in einer 2 Atome enthaltenen Superzelle beschrieben. Der Defekt besitzt in dieser Zelle eine weiter äquivalente Positionen (gelb). Die Anzahl Zu jeder Position gibt es weitere drei äquivalenter Konfigurationen (blau). Die Gesamtzahl der Defektpositionen ist somit zwei, mit jeweils vier möglichen Konfigurationen.

definiert. Im Falle des thermodynamischen Gleichgewichtes kann die Konzentration $c(X^q)$ eines Defektes X im Ladungszustand q über die Boltzmannverteilung

$$c(X^q) = N_0 N_{\text{conf}} \exp\left(-\frac{E^f(X^q)}{k_b T}\right) \quad (5.62)$$

berechnet werden. N_0 bezeichnet hierbei die Anzahl möglicher Defektpositionen, N_{conf} die Anzahl möglicher Konfigurationen des Defektes in der betrachteten Superzelle (vgl. Abbildung 5.7). Der Ausdruck $k_b T$ quantifiziert die thermische Energie zur Temperatur T .

Die Konzentrationen der freien Ladungsträger sind unter der Annahme eines parabolischen Bandes über die Beziehungen

$$c(e) = 2 \left(\frac{m_e^* k_b T}{2\pi \hbar^2}\right)^{3/2} e^{(E_{\text{Fermi}} - E_{\text{CBM}})/k_b T} \quad (5.63)$$

und

$$c(h) = 2 \left(\frac{m_h^* k_b T}{2\pi \hbar^2}\right)^{3/2} e^{(E_{\text{VBM}} - E_{\text{Fermi}})/k_b T} \quad (5.64)$$

gegeben[26]. Damit sind für die Gesamtheit der Defekte und der freien Ladungsträger Bestimmungsgleichungen gegeben, welche eine eindeutige Lösung für die Fermi-

5.5 Berechnung der Defektkonzentration im thermischen Gleichgewicht

energie und Defektkonzentrationen liefern. Die S/PHI/nX Bibliothek[18] löst dieses numerisch schlecht konditionierte Problem über ein Bisektionsverfahren.

6. Quantitativ optimierte Atomorbitale — Quamols

The important thing in science is not so much to obtain new facts as to discover new ways of thinking about them.

Das Bedeutsame an der Wissenschaft ist nicht unbedingt die Erzeugung neuer Ergebnisse, als vielmehr die Entdeckung neuer Wege über sie nachzudenken.

Sir William Lawrence Bragg, australischer bzw. britischer Physiker, 1890 - 1971

6.1. Atomzentrierte Orbitalfunktionen

Ein atomzentriertes Orbital ist das Paradebeispiel einer lokalisierten, das bedeutet am Atom verankerten und in der Reichweite beschränkten, Funktion. Ihre mathematische Form besteht aus dem Produkt einer radialen Funktion $R_{nl}(r)$ mit einer Kugelflächenfunktion $Y_{lm}(\hat{\mathbf{r}})$

$$\mu_\alpha(\mathbf{r}) = R_{nl}(r)Y_{lm}(\hat{\mathbf{r}}) \quad \text{mit } \alpha = \{\tau, n, l, m\}. \quad (6.1)$$

Dabei beschreibt \mathbf{r} den Abstandsvektor mit Richtung $\hat{\mathbf{r}}$ zum atomaren Zentrum τ , l die Drehimpulsquantenzahl und m die magnetische Quantenzahl. Der Index n entspricht der Hauptquantenzahl, kann aber in der kommenden Verallgemeinerung als Zählindex der Radialfunktionen zur zugehörigen Drehimpulsquantenzahl interpretiert werden. Die Lösungen atomarer Schrödingergleichungen ergeben, dass die radiale Form der Orbitale sich in Kernnähe wie r^l verhält und für große Abstände exponentiell abklingt. Einen natürlichen Ansatz bilden deswegen die sogenannten Slaterorbitale[71] (STO)

$$R_{nl}^{\text{STO}}(r) \sim r^l \sum_i c_{ni} e^{-\zeta_{ni} r}. \quad (6.2)$$

Diese geben das korrekte Abklingverhalten wieder, besitzen aber den Nachteil, dass die Berechnung von Mehrzentrenintegralen nicht analytisch vollzogen werden kann, sondern numerisch ausgewertet werden muss. Dies war zu Beginn der computer-gestützten Modellierungen ein gravierendes Problem, weswegen die sogenannten Gaußorbitale[72] (GTO)

$$R_{nl}^{\text{GTO}}(r) \sim r^l \sum_i c_{ni} e^{-\beta_{ni} r^2} \quad (6.3)$$

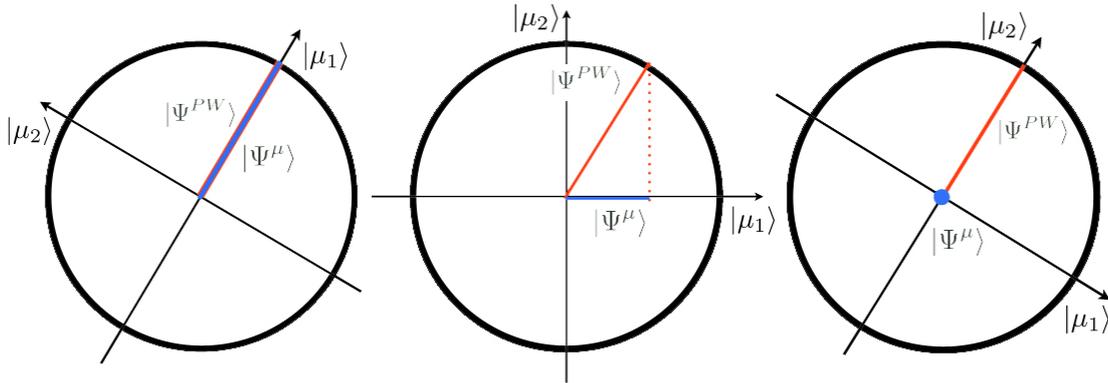
eingeführt wurden, welche eine analytische Beschreibung der Mehrzentrenintegrale erlauben. Allerdings besitzen diese GTOs im Gegensatz zu den STOs zwei wesentliche Nachteile. Zum einen erlauben die GTOs im kernnahen Bereich keine

diskontinuierliche Ableitung, die sich dort aber aus der Quantenmechanik ergibt. Zum anderen besitzen die GTOs gegenüber den STOs ein zu starkes Abklingverhalten. Zwar kann mit beiden Orbitaltypen eine komplette Basis aufgebaut werden, aber grundlegend werden mehr GTOs benötigt, als STOs. Die Koeffizienten c_{ni} , β_{ni} und ζ_{ni} werden dabei an Lösungen des atomaren Problems angepasst.

Diese atomaren Orbitale bilden einen chemisch intuitiven Basissatz, da sie — im Gegensatz zu ebenen Wellen — an Lösungen der atomaren Schrödingergleichung angelehnt sind. Allerdings ist die quantitative Verbesserung eines atomaren Basissatzes zur Beschreibung der chemischen Bindung bei Molekülen und Festkörpern nicht trivial, da es hier eine Vielzahl von möglichen Stellschrauben gibt. Die Anzahl der Basisfunktionen pro l -Kanal, die radiale Form, sowie der maximale Drehimpuls l spielen in der Optimierung solcher Basissätze eine entscheidende Rolle. Allerdings neigen atomzentrierte Basissätze zu dem Problem der Überbestimmtheit. Generell würde ein an einem einzigen Zentrum lokalisierter Basissatz genügen, um eine akkurate Beschreibung auch an anderen Zentren zu ermöglichen. Durch die Verankerung des Basissatzes an vielen Zentren läuft man in Gefahr linearer Abhängigkeiten zu erschaffen, was gerade für große Basissätze ein Problem darstellt. Reduziert man die Basissätze, neigen die atomaren Orbitale dazu auch Unzulänglichkeiten in der Beschreibung von nahe liegenden Fremdatomen auszugleichen — der sogenannte Basissatzsuperpositionsfehler (BSSE). Der BSSE macht sich dann besonders bei sich ändernden Strukturen bemerkbar, durch eine deutliche Verschlechterung der Genauigkeit bemerkbar. Die Transferabilität leidet immens.

Neuere Ansätze gehen deswegen dazu über Atomorbitale auf Basis von Ebenen-Wellen-Rechnungen zu erzeugen. So soll sichergestellt werden, dass für eine kleine Basis die einzelnen Atomorbitale optimal die Ebenen-Wellen-Zustände beschreiben. Dies kann beispielsweise mittels einer unitären Transformation geschehen, was zu den sogenannten Wannierorbitalen[73, 74, 75] führt. Diese unitäre Transformation hat jedoch die Schwäche, dass die erzeugten Orbitale alle Informationen über das Referenzsystem besitzen und die sphärische Symmetrie der Orbitale aufgegeben wird. Die Basis büßt deswegen bei strukturellen Veränderungen des Systems sofort signifikant an Beschreibungskraft ein. Bei Beibehaltung der Symmetrie stellt sich die Frage unter welchen Randbedingungen die Orbitalfunktionen optimiert werden. Die im nächsten Abschnitt vorgestellte Spillage[76] wird hierbei in einigen Arbeiten[77, 78, 79, 80, 81] als Optimierungskriterium verwendet. Allerdings beschränken sich diese Arbeiten auf eine Optimierung bezüglich einer zu Grunde liegenden Basis wie STOs, GTOs oder Besselfunktionen. Die dadurch eingebüßte Flexibilität der Orbitalfunktionen ist schwer abzuschätzen.

In dieser Dissertation soll nun erstmals die Optimierung von numerischen — das bedeutet basisfreien — Orbitalfunktionen unter Verwendung des Spillagegedanken


Abb. 6.1:

Im 2-dimensionalen lässt sich die Spillage als Normverlust eines Zustandes interpretieren. Der Kreis repräsentiert alle möglichen normierten Zustände des \mathbb{R}^2 , aus der ein Repräsentant (rot) gewählt worden ist. μ_1 und μ_2 bezeichnen eine vollständige Basis des \mathbb{R}^2 . Im Falle einer nicht vollständigen Basis ist nun nur noch ein einziger Basisvektor — hier μ_1 — zur Beschreibung des Zustandes vorhanden. Eine Projektion des Zustandes auf diesen Basisvektor resultiert nun in einem Normverlust: der Spillage.

vollzogen werden.

6.2. Ein quantitatives Mass für Basisqualität: Die Spillage

Das Ziel des Quamol Ansatzes ist es Basisfunktionen zu generieren, die in der Lage sind gegebene Zustände in der Darstellung ebener Wellen akkurat nachzubilden. Die Entwicklung eines ebenen Wellenzustandes in eine nicht-orthogonale, lokalisierte Basis $|\mu_\alpha\rangle$ lässt sich analog zur Entwicklung in orthogonale Basisfunktionen herleiten. Sei $|\Psi_n\rangle$ der zu entwickelnde Zustand und $|\mu_\alpha\rangle$ ein Element der neuen Basis mit $S_{\alpha\beta} = \langle\mu_\alpha|\mu_\beta\rangle$ als entsprechender Überlappmatrix. Gesucht sind im folgenden die Koeffizienten c_α , so dass die Basisentwicklung

$$|\Psi_n\rangle = \sum_{\alpha} c_{\alpha n} |\mu_\alpha\rangle \quad (6.4)$$

gilt. Die Multiplikation mit $\langle\mu_\beta|$ führt zu

$$\langle\mu_\beta|\Psi_n\rangle = \sum_{\alpha} c_{\alpha n} S_{\beta\alpha} \quad \text{mit} \quad S_{\beta\alpha} = \langle\mu_\beta|\mu_\alpha\rangle. \quad (6.5)$$

Um sich der Überlappmatrix auf der rechten Seite der Gleichung zu entledigen, multipliziert man mit $S_{\gamma\beta}^{-1}$ und summiert über β . Das rechtsseitige Matrixprodukt der Überlappmatrix und ihres Inversen liefert das Kroneckersymbol $\delta_{\gamma\alpha}$. Dieses führt in der Summe und anschließender Variablensubstitution $\gamma \rightarrow \alpha$ zu

$$c_{\alpha n} = \sum_{\beta} S_{\alpha\beta}^{-1} \langle\mu_\beta|\Psi_n\rangle. \quad (6.6)$$

Das Gleichheitszeichen der Entwicklung (6.4) besitzt natürlich nur dann Gültigkeit, wenn die zugrundeliegende lokalisierte Basis vollständig ist. Ist sie nicht vollständig, projiziert die Transformation (6.6) den Zustand $|\Psi_n\rangle$ in den entsprechenden Raum der lokalisierten Basis, von dem im weiteren angenommen wird, dass dieser ein Unterraum zu dem ursprünglichen Raum der ebenen Wellen ist. Um nun zwischen Projektion und ursprünglichen Zustand unterscheiden zu können, verwendet man an dieser Stelle $|\Psi_n^{\text{PW}}\rangle$ zur Kennzeichnung des Zustandes im Raum der ebenen Wellen und $|\Psi_n^\mu\rangle$ zur Kennzeichnung der Projektion. Diese sind über

$$|\Psi^\mu\rangle = \sum_{\alpha,\beta} |\mu_\alpha\rangle S_{\alpha\beta}^{-1} \langle\mu_\beta|\Psi_n^{\text{PW}}\rangle \quad (6.7)$$

$$= \hat{P}|\Psi_n^{\text{PW}}\rangle \quad (6.8)$$

miteinander verknüpft. Der so definierte Projektionsoperator \hat{P} ist *idempotent*

$$\hat{P}\hat{P} = \sum_{\alpha,\beta,\gamma,\delta} |\mu_\alpha\rangle S_{\alpha\beta}^{-1} \langle\mu_\beta|\mu_\gamma\rangle S_{\gamma\delta}^{-1} \langle\mu_\delta| \quad (6.9)$$

$$= \sum_{\alpha,\gamma,\delta} |\mu_\alpha\rangle \delta_{\alpha\gamma} S_{\gamma\delta}^{-1} \langle\mu_\delta| \quad (6.10)$$

$$= \sum_{\alpha,\delta} |\mu_\alpha\rangle S_{\alpha\delta}^{-1} \langle\mu_\delta| \quad (6.11)$$

$$= \hat{P} \quad (6.12)$$

und *selbstadjungiert*

$$\hat{P}^\dagger = \sum_{\alpha,\beta} |\mu_\beta\rangle (S^{-1})_{\alpha\beta}^\dagger \langle\mu_\alpha| \quad (6.13)$$

$$= \sum_{\alpha,\beta} |\mu_\beta\rangle S_{\beta\alpha}^{-1} \langle\mu_\alpha| \quad (6.14)$$

$$= \hat{P}. \quad (6.15)$$

Die Qualität dieser Projektion wird über die sogenannte *Spillage*

$$\mathbb{S}_n = \frac{\langle\Psi_n^{\text{PW}}|1 - \hat{P}|\Psi_n^{\text{PW}}\rangle}{\langle\Psi_n^{\text{PW}}|\Psi_n^{\text{PW}}\rangle} \quad (6.16)$$

quantifiziert. Da der Raum der Orbitalfunktionen ein Unterraum des Raumes der ebenen Wellen ist, gilt

$$0 \leq \langle\Psi_n^{\text{PW}}|\hat{P}|\Psi_n^{\text{PW}}\rangle \leq \langle\Psi_n^{\text{PW}}|\Psi_n^{\text{PW}}\rangle = 1, \quad (6.17)$$

und somit

$$0 \leq \mathbb{S}_n \leq 1. \quad (6.18)$$

Dabei bedeutet $\mathbb{S}_n = 0$ eine verlustfreie Projektion des Zustandes $|\Psi_n^{\text{PW}}\rangle$, während $\mathbb{S}_n = 1$ bedeutet, dass der Zustand $|\Psi_n^{\text{PW}}\rangle$ keine Komponente im Unterraum $\{\mu_\alpha\}$ hat. Abbildung 6.1 zeigt eine visuelle Interpretation dieser Größe.

Die für einen einzelnen Zustand $|\Psi_n^{\text{PW}}\rangle$ definierte Spillage lässt sich für eine Gesamtheit von Zuständen als

$$\mathbb{S} = \frac{1}{\mathbb{N}^{\text{PW}}} \sum_n f_n \langle \Psi_n^{\text{PW}} | 1 - \hat{P} | \Psi_n^{\text{PW}} \rangle \quad (6.19)$$

generalisieren, wobei die f_n als Filterfaktoren fungieren und

$$\mathbb{N}^{\text{PW}} = \sum_n f_n \langle \Psi_n^{\text{PW}} | \Psi_n^{\text{PW}} \rangle \quad (6.20)$$

eine Normierungsgröße darstellt: die sogenannte *Raumnorm*. Die Filterfaktoren f_n erlauben eine Selektion der zu beschreibenden Zustände. Ersetzt man jene beispielsweise durch die Besetzungszahlen gemäss einer Fermi-Dirac-Verteilung gibt die Spillage die Qualität der Projektion bezüglich des besetzten Kohn-Sham-Unterraumes an.

Mit der Spillage als quantitatives Maß der Projektionsgüte ist nun ein Vergleich verschiedener Orbitalbasen direkt möglich. Zudem erlaubt die Variationsrechnung die Bestimmung von Gradienten und somit die Möglichkeit eines gradientengestützten Optimierungsverfahrens für die Basis. Die Variation der Spillage \mathbb{S} im Hinblick auf die Orbitale ist hierbei die zentrale Grösse und wird nun im Folgenden bestimmt. Kennt man die Variation der inversen Überlappmatrix im Hinblick auf die atomaren Basisfunktionen (Anhang D) lässt sich damit die Variation des Projektors \hat{P} nach den Orbitalen $\langle \mu_\tau |$ zu

$$\frac{\hat{P}}{d\langle \mu_\tau |} = \sum_{\alpha, \beta} |\mu_\alpha\rangle \frac{dS_{\alpha\beta}^{-1}}{d\langle \mu_\tau |} \langle \mu_\beta | + \sum_{\alpha} |\mu_\alpha\rangle S_{\alpha\tau}^{-1} \quad (6.21)$$

$$= - \sum_{\alpha, \beta, \epsilon} |\mu_\alpha\rangle S_{\alpha\tau}^{-1} |\mu_\epsilon\rangle S_{\epsilon\beta}^{-1} \langle \mu_\beta | + \sum_{\alpha} |\mu_\alpha\rangle S_{\alpha\tau}^{-1} \quad (6.22)$$

$$= \sum_{\alpha} |\mu_\alpha\rangle S_{\alpha\tau}^{-1} (1 - \hat{P}) \quad (6.23)$$

berechnen. Die Variation der Spillage bezüglich des Orbitals $\langle \mu_\tau |$ liefert somit

$$\frac{d\mathcal{S}}{d\langle\mu_\tau|} = -\frac{1}{\mathbb{N}^{\text{PW}}} \sum_n f_n \left\langle \Psi_n^{\text{PW}} \left| \frac{\hat{P}}{d\langle\mu_\tau|} \right| \Psi_n^{\text{PW}} \right\rangle \quad (6.24)$$

$$= -\frac{1}{\mathbb{N}^{\text{PW}}} \sum_n f_n (1 - \hat{P}) |\Psi_n^{\text{PW}}\rangle \sum_\alpha \langle \Psi_n^{\text{PW}} | \mu_\alpha \rangle S_{\alpha\tau}^{-1} \quad (6.25)$$

$$= -\frac{1}{\mathbb{N}^{\text{PW}}} \sum_n f_n c_{\tau n}^* (1 - \hat{P}) |\Psi_n^{\text{PW}}\rangle. \quad (6.26)$$

Den zentralen Beitrag zu diesem Gradienten bildet das Residuum $(1 - \hat{P})|\Psi_n^{\text{PW}}\rangle$, welches im allgemeinen keinerlei sphärische Symmetrie mehr aufweist. Diesbezüglich ist Variation der Spillage nach den Radialfunktionen der Orbitale von Interesse. Diese ergibt sich durch Anwendung der Kettenregel der Differentiation zu

$$\frac{d\mathcal{S}}{d\langle R_{nl}|} = \frac{d\mathcal{S}}{d\langle\mu_\tau|} \cdot \frac{d\langle\mu_\tau|}{d\langle R_{nl}|}. \quad (6.27)$$

Der Operator $\frac{d\langle\mu_\tau|}{d\langle R_{nl}|}$ beinhaltet die Multiplikation mit den Kugelflächenfunktionen und die Projektion auf die radiale Basis. Die explizite Form wird an späterer Stelle formuliert werden.

6.3. Periodische Systeme

Durch die periodischen Randbedingungen, welche zur Lösung der Festkörper angenommen werden, ergibt sich für die Eigenzustände des Systems die Beschreibung durch Blochwellen $|\Psi_n^{\text{PW}}\rangle \rightarrow |\Psi_{nk}^{\text{PW}}\rangle$. Eine vergleichbare Beschreibung ist nun auch für den Projektor \hat{P} erforderlich um dessen Anwendung auf die Zustände technisch realisieren zu können. Dazu führt man an dieser Stelle atomare Blochwellen (AO-Blochwellen)

$$\mu_{\alpha\mathbf{k}}(\mathbf{r}) = \sum_{\mathbf{R}} \mu_\alpha(\mathbf{r} - \mathbf{R}) e^{i\mathbf{k}\mathbf{R}} \quad (6.28)$$

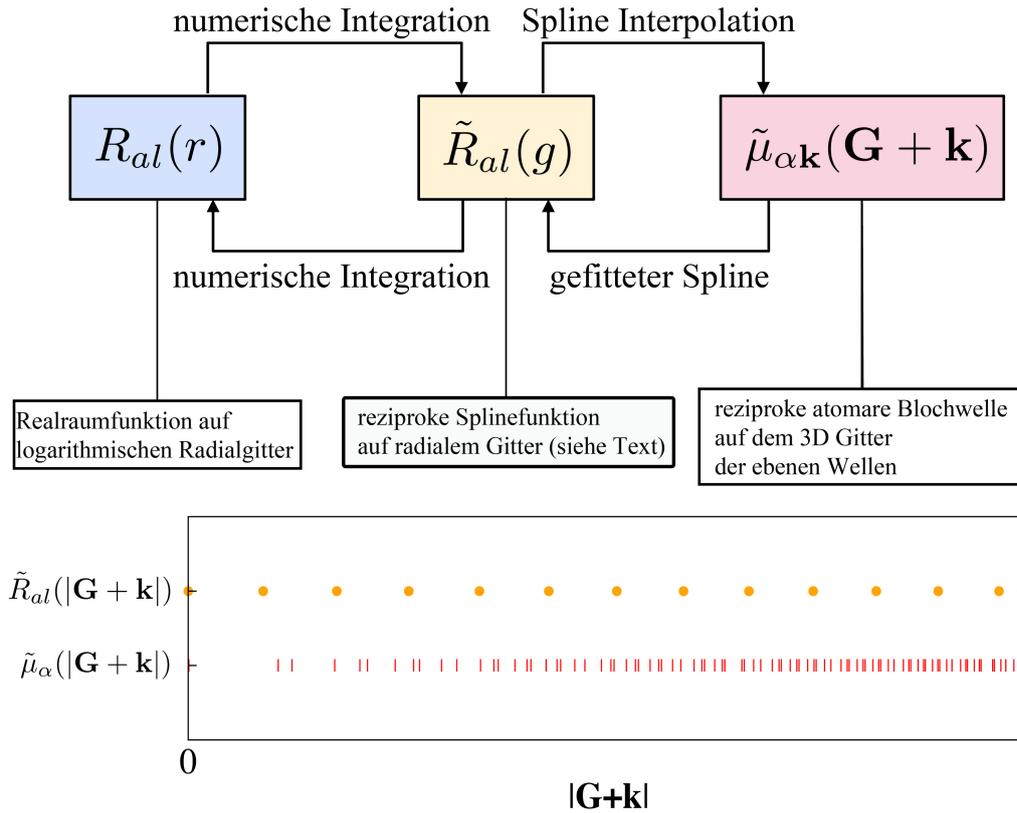
ein. \mathbf{R} ist hierbei ein Gittervektor des periodischen Systems. Die entsprechende AO-Blochwelle im reziproken Raum erreicht man durch die Fouriertransformation

$$\tilde{\mu}_{\alpha\mathbf{k}}(\mathbf{G} + \mathbf{k}) = \frac{1}{\sqrt{\Omega}} \int_{\Omega} d\Omega \mu_{\alpha\mathbf{k}}(\mathbf{r}) e^{-i(\mathbf{G}+\mathbf{k})\mathbf{r}}, \quad (6.29)$$

wobei Ω das Volumen der periodischen Zelle bezeichnet. Mit diesen Größen lässt sich nun der \mathbf{k} -Punkt abhängige Projektor

$$\hat{P}_{\mathbf{k}} = \sum_{\alpha\beta} |\tilde{\mu}_{\alpha\mathbf{k}}\rangle S_{\mathbf{k},\alpha\beta}^{-1} \langle \tilde{\mu}_{\beta\mathbf{k}}| \quad (6.30)$$

angeben, welcher in den Ausdrücken für die Spillage


Abb. 6.2:

Darstellung der Repräsentationskette der Orbitalfunktionen in dem entwickelten Algorithmus. Das jeweils genutzte Verfahren zum Repräsentationswechsel ist an den Pfeilen angegeben. Der untere Graph illustriert die Regelmäßigkeit des radialen Gitters im reziproken Raum, verglichen mit der radialen Projektion des Blochwellengitters.

$$\mathbb{S} = \frac{1}{\mathbb{N}^{\text{PW}}} \sum_{n\mathbf{k}} w_{\mathbf{k}} f_{n\mathbf{k}} \langle \Psi_{n\mathbf{k}}^{\text{PW}} | 1 - \hat{P}_{\mathbf{k}} | \Psi_{n\mathbf{k}}^{\text{PW}} \rangle \quad (6.31)$$

und dem Gradienten

$$\frac{d\mathbb{S}}{d\langle \mu_{\tau} \rangle} = -\frac{1}{\mathbb{N}^{\text{PW}}} \sum_{n\mathbf{k}} w_{\mathbf{k}} f_{n\mathbf{k}} c_{\mathbf{k},\tau n}^* (1 - \hat{P}_{\mathbf{k}}) | \Psi_{n\mathbf{k}}^{\text{PW}} \rangle \quad (6.32)$$

Verwendung findet. Mit $w_{\mathbf{k}}$ werden die \mathbf{k} -Punkt-Gewichte der diskretisierten Brillouinzone bezeichnet.

6.4. Algorithmus

Mit den in den vorherigen Kapiteln eingeführten Größen kann nun die Beschreibung des eigentlichen Algorithmus erfolgen. Ausgehend von einer atomzentrierten Orbitalfunktionen der Gestalt

$$\mu_{\alpha}(\mathbf{r}) = R_{nl}(r) Y_{lm}(\hat{\mathbf{r}}) \quad (6.33)$$

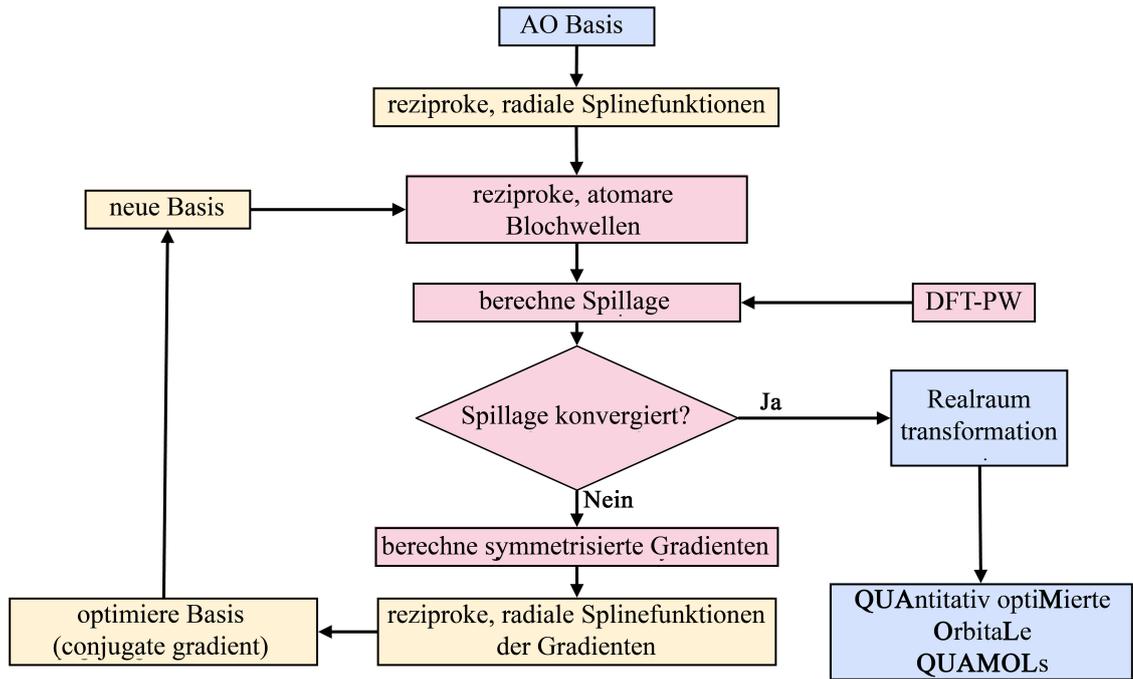


Abb. 6.3:

Flussdiagramm des Optimierungsverfahrens. Die gewählten Farben zeigen die zugrundeliegende Repräsentation der Orbitalfunktionen gemäß Abb. 6.2.

soll der auf einem Gitter numerisch repräsentierte, radiale Teil $R_{nl}(r)$ bei festgehaltenen $Y_{lm}(\hat{\mathbf{r}})$ dergestalt variiert werden, dass die Spillage bezüglich der zu Grunde liegenden Referenz der Ebenen-Wellen-Zustände minimal wird. Diese optimierten Orbitalfunktionen werden hier als **QU**Antitativ optiMIerte Ato**MO**rbita**Le**, kurz Quamols, bezeichnet.

Die Berechnung der Spillage und ihres Gradienten bedingen Auswertung von Überlappintegralen der Gestalt

$$\langle \mu_\alpha | \Psi_n^{\text{PW}} \rangle = \sum_{\mathbf{k}} w_{\mathbf{k}} \sum_{\mathbf{G}} \mu_{\alpha\mathbf{k}}^*(\mathbf{G} + \mathbf{k}) \Psi_{n\mathbf{k}}(\mathbf{G} + \mathbf{k}). \quad (6.34)$$

Dies bedarf einer Transformation der Orbitale in den reziproken Raum der ebenen Wellen. Der Gradient selber muss dann seinerseits zurück in eine radiale Form gebracht werden um die sphärische Symmetrie nicht zu brechen. Dies setzt in jedem Schritt der Optimierung eine Transformation in den radialen Raum voraus, welche auf Grund der ungleichmäßigen Stützpunktverteilung im $(\mathbf{G} + \mathbf{k})$ -Raum zu verrauschten Radialfunktionen führt. Diese Transformationen sind zudem auf Grund der Vielzahl von Gitterpunkten (≈ 40.000) rechenintensiv. Um das Rauschen zu minimieren und um die Anzahl an Stützpunkten deutlich zu reduzieren, führen wir an dieser Stelle radiale Orbitalfunktionen im reziproken Raum ein. Diese werden durch kubische Splines auf einem Hilfsgitter repräsentiert, welches eine deutlich homoge-

nere Stützstellenverteilung aufweist und in der Anzahl der Gitterpunkte signifikant geringer ist. In der Regel reichen 100 Punkte zur akkuraten Beschreibung aus. Abbildung 6.2 zeigt eine Gegenüberstellung der beiden Gitter. Die Transformationen aus dem radialen Ortsraum in den radialen reziproken Raum und *vice versa* sind hierbei durch eine ebene Wellenentwicklung in Kugelwellen gegeben. Die Transformation in den reziproken Raum geschieht über die Transformationsvorschrift

$$\tilde{R}_{nl}(g) = \sqrt{\frac{2}{\pi}} \int_0^{\infty} R_{nl}(r) j_l(gr) r^2 dr. \quad (6.35)$$

Die Rücktransformation geschieht über die Vorschrift

$$R_{nl}(r) = \sqrt{\frac{2}{\pi}} \int_0^{\infty} \tilde{R}_{nl}(g) j_l(gr) g^2 dg. \quad (6.36)$$

Die j_l bezeichnen die sphärischen Besselfunktionen erster Art. Diese Transformationen können durch eine numerische Integration erfolgen und sind in ihren Kosten durch die geringere Anzahl an Gitterpunkten (≈ 100) deutlich moderater als die direkte Transformation $\mu(\mathbf{G} + \mathbf{k}) \rightarrow R(r)$.

Die Auffaltung der radialen Orbitalfunktionen im reziproken Raum geschieht in zwei Schritten. Zunächst wird \tilde{R}_{nl} auf die $|\mathbf{G} + \mathbf{k}|$ -Stützstellen im $(\mathbf{G} + \mathbf{k})$ -Raum über kubische Splines interpoliert. Die Beschreibung dieser Methode befindet sich im Anhang E. Aus diesen ergeben sich dann die atomaren Blochwellen zu

$$\mu_{\alpha\mathbf{k}}(\mathbf{G} + \mathbf{k}) = \sqrt{\frac{(2\pi)^3}{\Omega}} \tilde{R}_{nl}(|\mathbf{G} + \mathbf{k}|) Y_{lm} \left(\frac{\mathbf{G} + \mathbf{k}}{|\mathbf{G} + \mathbf{k}|} \right) e^{i(\mathbf{G} + \mathbf{k})\mathbf{r}_\tau}. \quad (6.37)$$

Dabei bezeichnet \mathbf{r}_τ die Position des Atoms τ in der Zelle und Y_{lm} sind die reellwertigen, normierten Kugelflächenfunktionen.

Die Rücktransformation der aufgefalteten Orbitale in den radialen reziproken Raum geschieht ebenfalls in zwei Schritten. Zunächst erfolgt die Berechnung der radialen Blochwellen

$$\begin{aligned} \tilde{R}_{nl}(g_\xi) &= \sum_{\mathbf{G} + \mathbf{k}} \delta(g_\xi - |\mathbf{G} + \mathbf{k}|) \\ &\times \frac{1}{n_\tau(2l + 1)} \sum_{\tau, m=-l}^l \sqrt{\frac{2\Omega}{\pi}} \mu_{\alpha\mathbf{k}}(\mathbf{G} + \mathbf{k}) Y_{lm} \left(\frac{\mathbf{G} + \mathbf{k}}{|\mathbf{G} + \mathbf{k}|} \right) e^{-i(\mathbf{G} + \mathbf{k})\mathbf{r}_\tau} \end{aligned} \quad (6.38)$$

über eine Mittelung aller beitragenden Atomzentren τ und aller $(2l + 1)$ Orbitale zur Drehimpulsquantenzahl l . Die Projektion mit den Kugelflächenfunktionen Y_{lm} garantiert die Bewahrung der sphärischen Symmetrie und ist bei der analogen

Berechnung der radialen Gradienten im reziproken Raum essentiell. Im zweiten Schritt muss nun auf die Stützstellen des radialen Gitters rückgerechnet werden. Eine Interpolation führt nun jedoch nicht ans Ziel, da die obigen berechneten radialen Blochwellen signifikanten numerischen Schwankungen unterliegen. Deswegen wird an dieser Stelle die Methode des Splinefittens verwendet, welche in Anhang F genauer erläutert wird. Da die Anzahl der Punkte für $|\mathbf{G} + \mathbf{k}| \rightarrow 0$ abnimmt und sogar leere Intervalle möglich sind, behilft man sich an dieser Stelle damit, Intervalle adaptiv zusammenführen, bis jedes neue Intervall zumindest einen Punkt beheimatet. Zudem wird $R_{nl}(g)$ symmetrisch nach $-\infty$ fortgesetzt, was das Verfahren weiter stabilisiert. Bei ungeraden l ist die Punktsymmetrie zu wählen, bei geradem l die Achsensymmetrie.

Dieses Verfahren kann ebenso für die Transformation des Gradienten verwendet werden. Eleganter ist es den resultierenden Gradienten direkt gemäss Kettenregel der Differentiation zu berechnen. Dieser ergibt sich zu

$$\frac{dS}{d\langle \tilde{R}_{nl}(g_x) |} = \frac{d\langle \tilde{R}_{nl}(g_\xi) |}{d\langle \tilde{R}_{nl}(g_x) |} \times \frac{d\langle \mu_{\alpha\mathbf{k}}(\mathbf{G} + \mathbf{k}) |}{d\langle \tilde{R}_{nl}(g_\xi) |} \times \frac{dS}{d\langle \mu_{\alpha\mathbf{k}}(\mathbf{G} + \mathbf{k}) |}. \quad (6.39)$$

Die expliziten Formen sind mit

$$\frac{d\langle \mu_{\alpha\mathbf{k}}(\mathbf{G} + \mathbf{k}) |}{d\langle \tilde{R}_{nl}(g_\xi) |} = \sqrt{\frac{(2\pi)^3}{\Omega}} \sum_{\tau,m} |g_\xi\rangle e^{i(\mathbf{G}+\mathbf{k})\mathbf{r}_\tau} \langle Y_{lm} | \mathbf{G} + \mathbf{k} \rangle \langle \mathbf{G} + \mathbf{k} | \quad (6.40)$$

und

$$\frac{d\langle \tilde{R}_{nl}(g_\xi) |}{d\langle \tilde{R}_{nl}(g_x) |} = \sum_{i,p} |g_x\rangle \frac{d\chi_{i,p}}{dR_{nl}(g_x)} (g_\xi(i) - g_x(i))^p \langle g_\xi | \quad (6.41)$$

gegeben. Der letzte Ausdruck ergibt sich aus der Variation der Splinekoeffizienten bezüglich der Funktionswerte an den Stützstellen. Für weitergehende Lektüre sei hier auf den Anhang E und F verwiesen.

Mit diesen Transformationen ist es nun möglich den Algorithmus aufzubauen, der basierend auf der Spillage und des entsprechenden Gradienten die Orbitale anhand der Methode der konjugierten Gradienten optimiert. Das Flussdiagramm ist in Abbildung 6.3 dargestellt. Dieser Algorithmus ist in die S/PHI/nX Bibliothek[18] implementiert worden.

6.4.1. Initialisierung der lokalisierten Basis

Die Optimierung der Orbitalfunktionen ist — vergleichbar zur Optimierung der elektronischen Struktur — hochgradig nichttrivial. Eine gute Initialisierung der

Basis sowohl in der Form, als auch in der Anzahl der Basisfunktionen, ist auf Grund der komplexen Spillage-Oberfläche wichtig. Bezüglich der Anzahl der Basisfunktionen ist es gut, dem Konzept des *balancierten Basissätze*[20] zu folgen. Die wesentlichen Beiträge zur Elektronenverteilung werden durch s - und p -Funktionen gegeben. Höhere Funktionen im Sinne der Drehimpulsquantenzahl dienen, mit Ausnahme der Beschreibung von Übergangselementen, der Polarisation. So ist eine $3s2p1d$ -Basis balanciert, während eine $1s2p1d$ -Basis ein zu hohes Gewicht auf die Polarisierungsfunktionen legt und somit als *überpolarisiert* gilt. Eine $5s4p1d$ -Basis hingegen legt ein zu schwaches Gewicht auf die Polarisationsfunktionen und gilt als *unterpolarisiert*. Erfahrungsgemäß stellt eine balancierte Basis eine gute Initialisierung dar. Ein zu hohes Gewicht auf den Polarisationsfunktionen kann in der Optimierung dazu führen, dass Unzulänglichkeiten in den s - und p -Anteilen in den höheren Funktionen ausgeglichen werden, was zu Artefakten in der Darstellung führen kann. Dies werden wir am Beispiel des Stickstoffdimers in Abschnitt 6.5.1 und am Beispiel des Silizium in Abschnitt 6.5.2 diskutieren.

Gleichermaßen wichtig ist die radiale Form der Initialisierung. Atomare Lösungen des Atoms, beziehungsweise des Pseudoatoms bei der Verwendung von Pseudopotentialen, bilden hierbei erfahrungsgemäß gute Initialisierungen, die um weitere Gaußfunktionen der Gestalt

$$R_{nl}(r) = r^l \cdot e^{-\gamma r^2} \quad (6.42)$$

oder gar um Linearkombinationen

$$R_{nl}(r) = r^l \cdot \sum_i c_i e^{-\gamma_i r^2} \quad (6.43)$$

ergänzt werden können. Einen fundierten Datensatz bildet hierbei der *EMSL Basis-set Exchange*[82], welcher eine Bandbreite unterschiedlichster Orbitale bereithält, die in dieser Arbeit als Initialisierung für Polarisationsfunktionen gedient haben.

In dieser Arbeit werden die Basisfunktionen aus einer Mischung der pseudoisierten Atomorbitale und weiteren Gaußfunktionen, entnommen aus Jensens *polarization consistent* (PC) Basissatz[82], initialisiert. Eine solche Initialisierung wird dann als pseudoisierter PC-Basissatz bezeichnet.

6.4.2. Typische Parameter

Zur Erzeugung der Eigenzustände $|\Psi_n^{\text{PW}}\rangle$ in der Darstellung ebener Wellen werden Dichtefunktionalrechnungen durchgeführt. Die Art der Beschreibung des Austausch- und Korrelationsfunctionals wird zu Beginn der Diskussion des Materialsystems angegeben, wie auch der benutzte Energiecutoff der ebene Wellen und das \mathbf{k} -Punkt Gitter. Während für Dichtefunktionalrechnungen nicht-zentrierte \mathbf{k} -Punkt Gitter

Basisname	#	S_{init}	S_{opt}	E_{tot} [eV]	ΔE_{tot} [eV]
minimal ($1s1p$)	4	$4.68 \cdot 10^{-2}$	$1.02 \cdot 10^{-2}$	-540.6616	4.56
minimal + d ($1s1p1d$)	9	$3.29 \cdot 10^{-2}$	$3.29 \cdot 10^{-3}$	-543.4634	1.76
pseudoisierter PC0 ($2s2p$)	8	$6.39 \cdot 10^{-3}$	$2.29 \cdot 10^{-3}$	-543.6593	1.56
balancierter PC0 ($2s1p$)	5	$3.57 \cdot 10^{-2}$	$3.43 \cdot 10^{-3}$	-543.0280	2.20
pseudoisierter PC1 ($2s2p1d$)	13	$3.97 \cdot 10^{-3}$	$9.99 \cdot 10^{-5}$	-545.1124	0.112
balancierter PC1 ($3s2p1d$)	14	$3.16 \cdot 10^{-3}$	$6.97 \cdot 10^{-5}$	-545.1287	0.095
balancierter PC2 ($4s3p2d1f$)	30	$1.67 \cdot 10^{-3}$	$8.08 \cdot 10^{-6}$	-545.2076	0.016
unterpolarisiert ($4s3p1d$)	18	$1.98 \cdot 10^{-3}$	$5.61 \cdot 10^{-5}$	-545.1341	0.090
überpolarisiert ($1s3p2d1f$)	27	$4.69 \cdot 10^{-3}$	$2.34 \cdot 10^{-4}$	-544.8947	0.329

Tab. 6.1:

Vergleich der Spillage nach Initialisierung und Optimierung für unterschiedlich Basisgrößen. Gegeben ist zudem die Anzahl der Funktionen (#) sowie die Gesamtenergie berechnet in einem LCAO Ansatz und der Vergleich dieser mit der Ebenen-Wellen-Rechnung.

bevorzugt werden, sind für die Splineinterpolation zentrierte \mathbf{k} -Punkt Gitter vorteilhaft. Die explizite Mitnahme des Γ -Punktes erlaubt die Nutzung der physikalischen Eigenschaften der Orbitalfunktionen an diesem Punkt, wie

$$R_{nl}(0) = 0 \quad \forall \quad l > 0 \quad (6.44)$$

und

$$R'_{nl}(0) = 0 \quad \forall \quad l > 1. \quad (6.45)$$

Weiterhin liefern dichte \mathbf{k} -Punkt Gitter mehr Punkte nahe des Γ -Punktes, was die Stabilität des Splinefittens verbessert. Deswegen werden die Eigenzustände bei konstant gehaltener Elektronendichteverteilung für den Quamolansatz an dichteren, zentrierten \mathbf{k} -Punkt Gitter nachgerechnet um obige Vorteile nutzen zu können. Die typischen Gittergrößen belaufen sich auf 500...1000 Punkte für das logarithmische Gitter der Radialfunktionen im Ortsraum, und 50...200 Punkte für das lineare Gitter der Radialfunktionen im reziproken Raum.

Die Filterfaktoren f_n werden — soweit nicht anders spezifiziert — in Übereinstimmung mit den Besetzungszahlen des physikalischen Systems gewählt.

6.5. Ausgewählte Beispiele

6.5.1. Minimale versus erweiterte Basis - Stickstoffdimer

Zur Illustration des Quamolansatzes verwenden wir zu Beginn ein wohl verstandenes System und diskutieren den Einfluss der Basissatzgröße an einem einfachen Molekül

Größe	σ_s	σ_s^*	π_p	σ_p
1s1p	$4.51 \cdot 10^{-3}$	$1.43 \cdot 10^{-4}$	$9.48 \cdot 10^{-4}$	$3.60 \cdot 10^{-3}$
1s1p1d	$1.60 \cdot 10^{-3}$	$1.38 \cdot 10^{-5}$	$2.56 \cdot 10^{-4}$	$1.16 \cdot 10^{-3}$
2s2p	$1.95 \cdot 10^{-4}$	$3.62 \cdot 10^{-5}$	$9.11 \cdot 10^{-4}$	$2.37 \cdot 10^{-4}$
2s1p	$1.00 \cdot 10^{-3}$	$8.13 \cdot 10^{-5}$	$1.00 \cdot 10^{-3}$	$3.36 \cdot 10^{-4}$
2s2p1d	$2.97 \cdot 10^{-5}$	$1.42 \cdot 10^{-6}$	$2.78 \cdot 10^{-5}$	$1.33 \cdot 10^{-5}$
3s2p1d	$1.38 \cdot 10^{-5}$	$1.42 \cdot 10^{-6}$	$2.11 \cdot 10^{-5}$	$1.22 \cdot 10^{-5}$
4s3p2d1f	$1.64 \cdot 10^{-6}$	$2.93 \cdot 10^{-7}$	$2.19 \cdot 10^{-6}$	$1.77 \cdot 10^{-6}$
4s3p1d	$1.06 \cdot 10^{-5}$	$1.02 \cdot 10^{-6}$	$1.89 \cdot 10^{-5}$	$6.69 \cdot 10^{-6}$
1s3p2d1f	$1.41 \cdot 10^{-4}$	$7.25 \cdot 10^{-7}$	$8.71 \cdot 10^{-6}$	$7.55 \cdot 10^{-5}$

Tab. 6.2:

Vergleich der Zustandsspillage für verschiedene Basissätze. Die Beiträge sind auf die Gesamtspillage normiert.

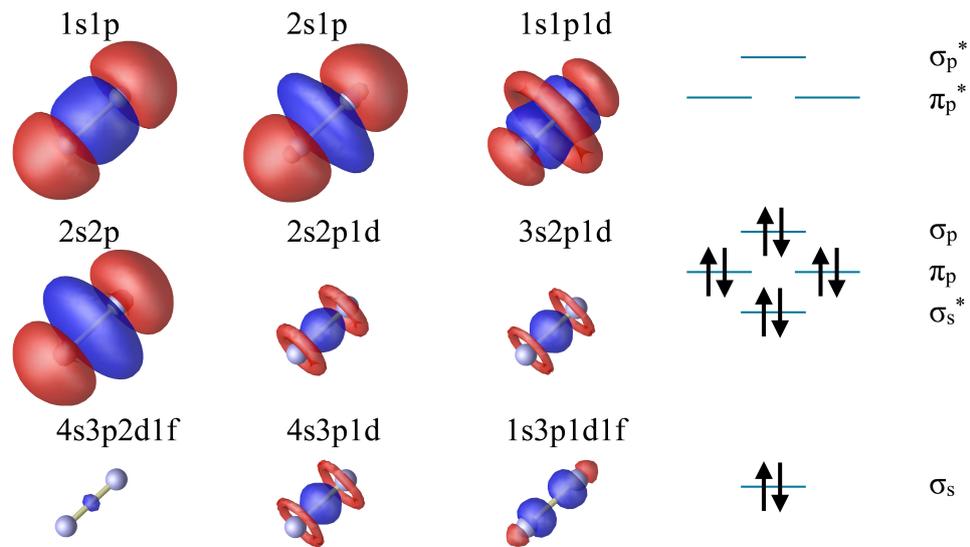
Größe	d [bohr]	Δd [%]	E_{tot} [eV]	ΔE_{tot} [eV]
ebene Wellen	2.085	—	-545.2239	—
1s1p	2.233	+7.1	-541.0198	4.20
1s1p1d	2.126	+2.0	-543.4987	1.73
2s2p	2.124	+1.8	-543.6892	1.53
2s1p	2.137	+2.5	-543.0879	2.14
2s2p1d	2.088	+0.11	-545.1124	$112 \cdot 10^{-3}$
3s2p1d	2.088	+0.12	-545.1314	$93 \cdot 10^{-3}$
4s3p2d1f	2.086	+0.03	-545.2076	$16 \cdot 10^{-3}$
4s3p1d	2.087	+0.09	-545.1341	$90 \cdot 10^{-3}$
1s3p2d1f	2.077	-0.4	-544.8974	$327 \cdot 10^{-3}$

Tab. 6.3:

Vergleich der Ergebnisse der Strukturoptimierung von N_2 unter Verwendung unterschiedlicher optimierter Basissätze.

mit σ - und π -Bindungen: das Stickstoffdimer. Die elektronische Struktur dieses Systems ist im Sinne des LCAO-Bildes gut verstanden. Die s - und p -Orbitale des Stickstoffes prägen bei der atomaren Bindung drei σ - und zwei π -Molekülorbitale aus, die mit Elektronen gefüllt werden. Ein minimaler Basissatz besteht somit aus einer s - und einer p -Funktion.

Die Modellierung des Systems mittels DFT erfolgt unter Verwendung der lokalen Dichtenäherung (LDA). Konvergierte Ergebnisse lassen sich mit einer Cutoffenergie von 35 Ry erreichen. Das Molekül wird in einer Box der Kantenlänge 15 Bohr beschrieben, weswegen eine Γ -Punkt Rechnung im Sinne der \mathbf{k} -Punkt-Konvergenz ausreichend ist. Der berechnete Bindungsabstand beläuft sich auf 2.085 Bohr in guter Übereinstimmung mit dem experimentellen Befund von 2.074 Bohr[27].

**Abb. 6.4:**

Links: Visualisierung der Residualdichte. Blaue Bereiche zeigen fehlende Dichte im Vergleich zur Ebenen-Wellen-Basis. Rote Bereiche zeigen additive Dichte an. Die Isofläche beträgt: $5 \cdot 10^{-3} \text{ e}/\text{\AA}^3$

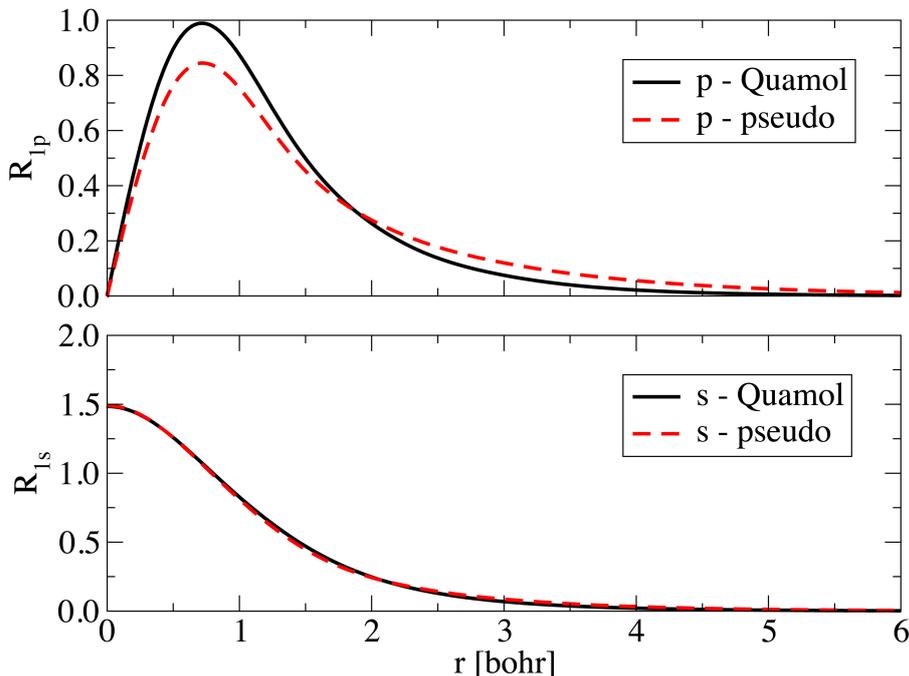
Rechts: LCAO-Termschema des Stickstoffdimers mit Besetzung.

Für unterschiedliche Basissatzgrößen führen wir nun Optimierungen der Basisfunktionen aus. Zudem sind die pathologischen Fälle einer unter- und einer überpolarierten Basis betrachtet worden, um ihre Effekte zu studieren. Die Spillage, sowie Gesamtenergie und die Abweichung zur ebenen Wellenrechnung sind in Tabelle 6.1 angegeben. Für die Diskussion ebenfalls interessant ist die zustandsaufgelöste Spillage

$$S_n = \frac{\langle \Psi_n | 1 - \hat{P} | \Psi_n \rangle}{\text{NPW}} \quad (6.46)$$

in Tabelle 6.2.

Eine Optimierung der minimalen Basis reduziert die anfängliche Spillage der aus dem PAW-Potential entnommenen Orbitalfunktionen um einen Faktor 5 auf $1 \cdot 10^{-2}$, was energetisch immer noch eine Diskrepanz von 4.6 eV bedeutet. Die dominanten Beiträge zur Spillage stammen von den bindenden σ_s und σ_p Zustand. Die Hinzunahme einer d -Funktion zur Polarisierung senkt die Spillage aller Zustände, während eine weitere s -Funktion — allein, oder in Kombination mit einer p -Funktion — sich einzig positiv auf die Beschreibung der σ -artigen Zustände auswirkt, jedoch auf den π_p Zustand keinen signifikanten Effekt hat. Eine Kombination beider Erweiterungen zu einer $2s2p1d$ Basis liefert eine Spillage von $1 \cdot 10^{-4}$ mit einer energetischen Diskrepanz von 110 meV. Das Hinzufügen weitere Funktionen zu einer balancierten $4s3p2d1f$ Basis senkt die Spillage um eine weitere Größenordnung. Die energeti-

**Abb. 6.5:**

Vergleich der Basisfunktionen aus der Initialisierung und nach der Optimierung. Verwendet wurde die minimale sp -Basis für das Stickstoffmolekül.

sche Abweichung beträgt nun nur noch 14 meV. Die verbesserte Beschreibung der Zustände lässt sich über die Residualdichte

$$\mathbb{R}(\mathbf{r}) = \sum_{n,\mathbf{k}} w_{\mathbf{k}} f_n (\Psi_{n\mathbf{k}}^{\text{PW}}(\mathbf{r}) \Psi_{n\mathbf{k}}^{\text{PW}*}(\mathbf{r}) - \Psi_{n\mathbf{k}}^{\mu}(\mathbf{r}) \Psi_{n\mathbf{k}}^{\mu*}(\mathbf{r})) \quad (6.47)$$

visuell darstellen. Abbildung 6.4 zeigt diese für die unterschiedlichen Basisgrößen. Blaue Bereiche zeigen hierbei fehlende Dichte im Vergleich zur Ebenen-Wellen-Referenz. Rote Bereiche hingegen sind zusätzliche Dichte, hervorgerufen durch den beschränkten atomaren Basissatz. Dies wird beispielsweise durch eine Mangelkompensation einer s -artigen Funktion durch eine am Nachbaratom lokalisierten p -Funktion hervorgerufen. Mit zunehmender Basisgröße reduzieren sich beide Effekte. Zu sehen ist ebenfalls, dass eine unterpolarisierte Basis ein qualitativ vergleichbares Bild liefert, während eine überpolarisierte Basis — sowohl in der $1s1p1d$ -Basis aber dominant in der $1s3p2d1f$ -Basis — die Beschreibung qualitativ verändert. Dieser visuelle Unterschied lässt sich auch an der Beschreibung der Bindung quantitativ fassen, wenn die Ergebnisse der Geometrieoptimierung in Tabelle 6.3 betrachtet werden. Während der Bindungsabstand mit steigender Basisgröße von oben gegen das Ergebnis der Ebenen-Wellen-Rechnung konvergiert, kommt die Bindung für die überpolarisierte Basis zu kurz heraus. Zwar wird der

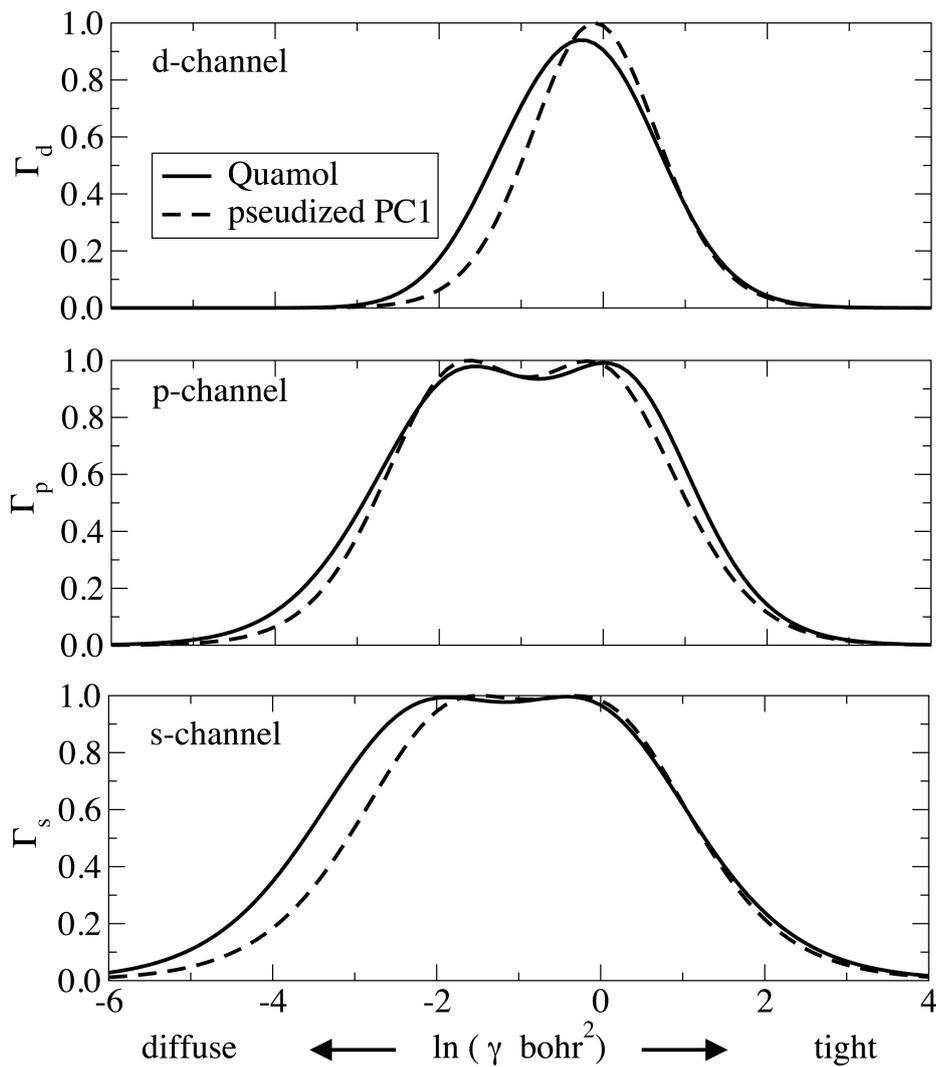


Abb. 6.6:

Vollständigkeitsprofil der Basisfunktionen nach der Initialisierung und nach der Optimierung für die $2s2p1d$ -Basis zur Beschreibung des Stickstoffmoleküls.

experimentelle Wert zufällig besser reproduziert, jedoch nicht auf Grund einer exakteren physikalischen Beschreibung. Weder energetisch noch in der Spillage lässt sich diese fehlerhafte Beschreibung direkt herauslesen. Bei der Beschreibung des Siliziums an späterer Stelle werden weitere Artefakte, induziert durch eine überpolarisierte Basis, sichtbar.

Der Effekt der Optimierung lässt sich an der radialen Form der Orbitalfunktionen diskutieren. Im Falle der minimalen sp -Basis sind die Radialfunktion in Abbildung 6.5 wiedergegeben. Man erkennt, dass hauptsächlich der Radialteil der p -Funktion von der Optimierung durch Kontraktion modifiziert worden ist. Eine signifikante Modifizierung der s -Funktion lässt sich nicht beobachten. Im Falle mehrerer

Basisfunktionen pro l -Kanal lässt sich ein direkter Vergleich der radialen Initialfunktionen und ihrer optimierten Gegenstücke nicht mehr ziehen. Die Verwendung mehrerer Funktionen und die fehlende Forderung nach Orthogonalität erlaubt das Bilden von Linearkombinationen, so dass eine optimierte Funktion nicht mehr eindeutig einer Initialisierungsfunktion zugeordnet werden kann. Abhilfe schafft hier der Übergang zu den sogenannten *completeness*-Profilen nach Chong[83]. Abbildung 6.6 zeigt solche Profile für die $2s2p1d$ -Basis. Dem Grundgedanken nach wird hier die Repräsentation von Gaußfunktionen mit unterschiedlichen Exponenten γ abgefragt. Je näher die durch $\Gamma_l(\gamma)$ evaluierte Projektion

$$\Gamma_l(\gamma) = \sum_{\alpha,\beta,m} \langle g_{lm}^\gamma | \mu_\alpha \rangle S_{\alpha,\beta}^{-1} \langle \mu_\beta | g_{lm}^\gamma \rangle \quad (6.48)$$

an eins liegt, umso besser ist diese Gaußfunktion repräsentiert. Die dargestellten Profile zeigen, dass durch die Optimierung in diesem Fall mehr Exponenten besser repräsentiert werden, als in der Initialisierung. Dies kann sogar dazu führen, dass, wie im Falle der d -Funktion, die Repräsentation eines einzelnen Exponenten abgesenkt wird, um eine größere Breite zu erreichen. In der praktischen Anwendung ist das aber nicht problematisch, da die Orbitalfunktion und nicht die Gaußfunktion bestmöglich repräsentiert werden muss.

Nach Betrachtung des Stickstoffmoleküls lassen sich die folgenden Kernaussagen treffen.

- Minimale Basissätze sind selbst nach der Optimierung nicht ausreichend um quantitativ hochwertige Ergebnisse zu erzielen.
- Balancierte Basissätze sind den nicht balancierten Basissätzen vorzuziehen, da sich in diesen Artefakte einschleichen können, die zu unphysikalischen Sachverhalten führen.
- Eine Vergrößerung des Basissatzes liefert eine systematische Konvergenz der Spillage um circa eine Größenordnung pro hinzugenommener Schale.

6.5.2. Kristalle - Silizium

Silizium kristallisiert in der Diamantstruktur und ist ein kovalent gebundenes Materialsystem. Jedes Siliziumatom prägt zu seinen nächsten Nachbarn vier identische Bindungen aus. Von diesen ist bekannt, dass die bindenden Orbitale durch sp^3 -Hybride beschrieben werden können. Ein sp -Basissatz bildet folglich in diesem System die minimale Basis. Silizium ist in der Literatur ein viel diskutiertes Materialsystem und erlaubt somit einen Vergleich mit anderen Ansätzen, die ebenfalls auf einer Optimierung der Spillage basieren.

Basisname	Größe	S_{init}	S_{opt}
minimal basis	$1s1p$	$4.1 \cdot 10^{-2}$	$6.2 \cdot 10^{-3}$
minimal basis + d	$1s1p1d$	$2.2 \cdot 10^{-2}$	$1.7 \cdot 10^{-4}$
PC1 - Pseudo	$2s2p1d$	$7.1 \cdot 10^{-4}$	$1.1 \cdot 10^{-4}$
PC2 - Pseudo	$3s3p2d1f$	$2.2 \cdot 10^{-4}$	$2.8 \cdot 10^{-5}$
PC2 - Reduced	$2s2p1d1f$	$6.3 \cdot 10^{-5}$	$3.7 \cdot 10^{-5}$
PC2 - Reduced	$1s2p1d1f$	$7.3 \cdot 10^{-5}$	$5.4 \cdot 10^{-5}$
Basisname	Größe	S_{opt}	$S_{\text{opt}} / S_{\text{opt}}^{\text{Quamol}}$
STO[77]	$1s1p$	$7.6 \cdot 10^{-3}$	1.23
PLATO[78]	$2s2p1d$	$\approx 2.7 \cdot 10^{-2}$	245.45
MIN S[77]	$1s1p$	$4.4 \cdot 10^{-3}$	0.71
MIN S[77]	$1s1p1d$	$1 \cdot 10^{-4}$	0.59

Tab. 6.4:

Gesamtspillage für unterschiedliche Basissätze am Beispiel des Silizium. MIN S ist die theoretisch berechnete untere Schranke für die Spillage[77] (vgl. Text).

Zur Modellierung des Materialsystems findet die lokale Dichtenäherung (LDA) zur Beschreibung des Austausch- und Korrelationsfunktionals Verwendung. Ein entsprechendes PAW-Potential ist der VASP-Datenbank[84] entnommen worden. Konvergierte Ergebnisse erreicht man bei Verwendung eines Energiecutoffs von 20 Ry und eines $4 \times 4 \times 4$ Monkhorst-Pack \mathbf{k} -Punktgitters. Zur Stabilisierung des Splinefittens sind die Wellenfunktionen bei festgehaltener Dichte auf einem Γ -zentrierten $10 \times 10 \times 10$ \mathbf{k} -Punktgitter berechnet worden. Für das radiale Gitter im reziproken Raum sind 100 Punkte verwendet worden.

Neben einem minimalen Basissatz sind erweiterte Basissätze erzeugt worden. Als Ausgangspunkt dienen hierbei balancierte Basissätze der pseudoisierten PC-

Method	c_{11} [GPa]	c_{12} [GPa]	c_{44}^0 [GPa]	c_{44} [GPa]	B_0 [GPa]
DFT-PW	160.6	61.6	103.2	73.5	94.6
LCAO sp	174.0	66.3	124.1	89.1	102.2
LCAO spd	171.3	68.3	110.9	77.9	102.7
LCAO $2s2pd$	168.6	60.3	107.3	76.6	96.4
LCAO $2s2pdf$	165.2	61.4	106.2	75.7	96.0
DFT-PW[78]	162.0	63.19	107.4	77.22	96.13
Plato sp [78]	186.8	75.03	139.5	91.45	112.3
Plato $2s2pd$ [78]	156.9	59.84	111.9	76.85	92.19

Tab. 6.5:

Elastische Konstanten für Silizium. c_{44}^0 bezieht sich auf die c_{44} Konstante ohne ionische Relaxationen.

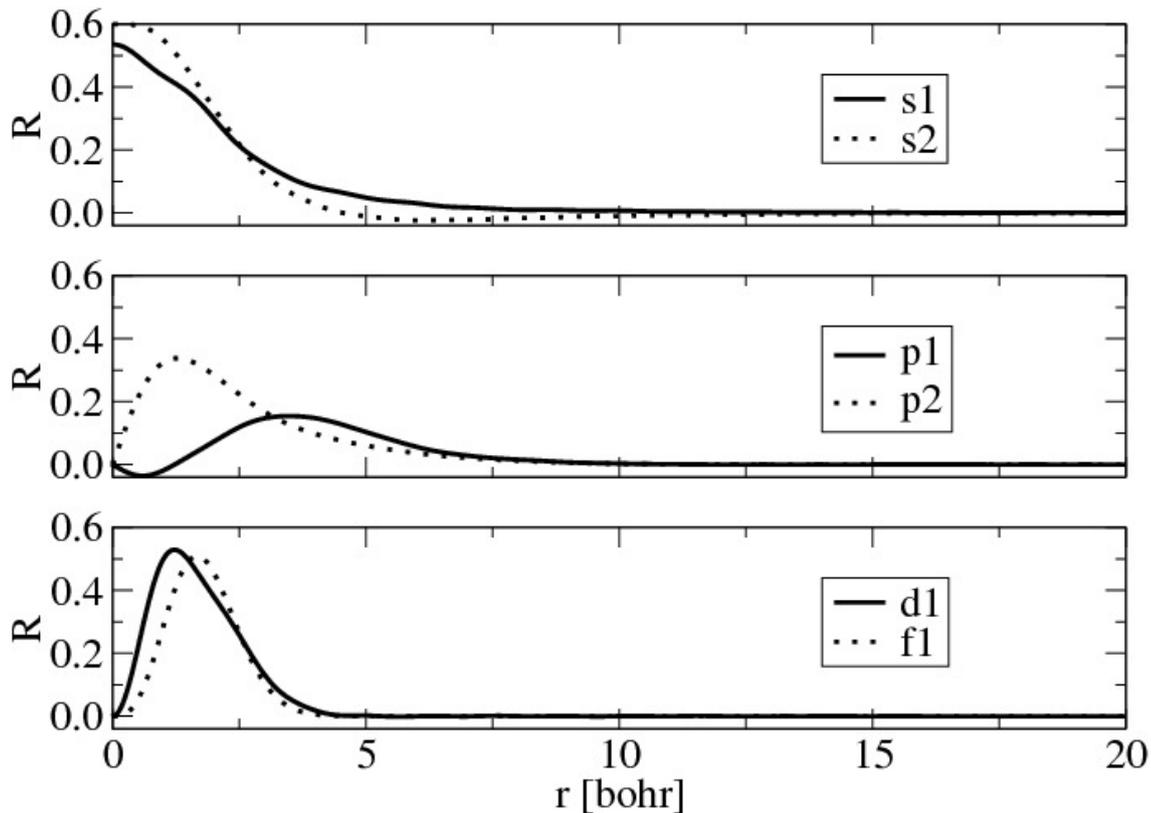


Abb. 6.7:
Darstellung des optimierten 2s2pfd Basissatzes.

Variante. Um mit anderen Basissätzen aus der Literatur vergleichbar zu sein, sind diese anschliessend in ihrer Größe reduziert worden. Da das Weglassen einzelner Funktionen die Spillage signifikant verschlechtert, ist hierbei eine Eigenwertzerlegung verwendet worden, um die Funktionen zu identifizieren, die einen minimalen Beitrag zur Spillagereduktion liefern. Das Verfahren wird an dieser Stelle grob skizziert.

Der Beitrag einer einzelnen Basisfunktion zur Zustandsentwicklung ist gemäß Gleichung (6.4) über die Entwicklungskoeffizienten $c_{\alpha n}$ gegeben. Da die zugrundeliegende Basis nicht orthonormal ist, ist nun die Koeffizientenmatrix

$$P_{\alpha\beta} = \sum_{n,\mathbf{k}} w_{\mathbf{k}} f_{n\mathbf{k}} c_{\mathbf{k},\alpha n}^* c_{\mathbf{k},\beta n} \quad (6.49)$$

aufzustellen. Eine Diagonalisierung dieser Matrix liefert Eigenwerte mit hinzugehörigen Eigenfunktionen. Die Eigenwerte geben ein quantitatives Kriterium für die Wichtigkeit der einzelnen Eigenfunktion zur Beschreibung des Zustandsraumes. Allerdings mischen dabei in die Eigenfunktionen beliebige l -Komponenten und Atomsorten, so dass die sphärische Symmetrie sowie die Atomzugehörigkeit nicht mehr zuzuordnen sind. Um diese Information zu bewahren, definiert man eine Koeffizientenmatrix, die als zusätzliche Indizes die Spezies i_s und die Drehimpulszahl l erhält

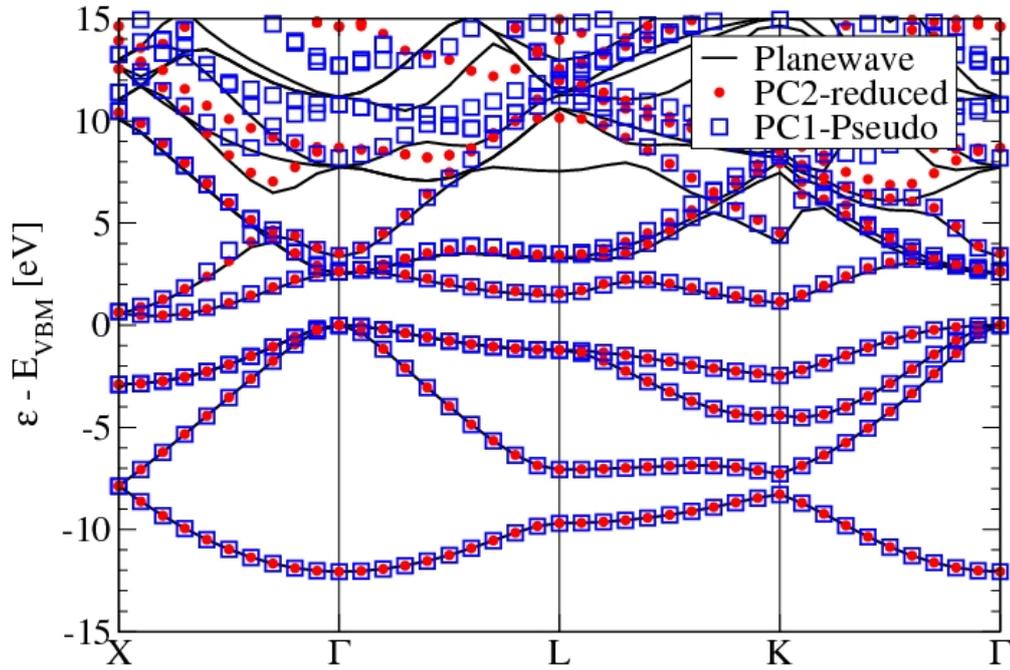


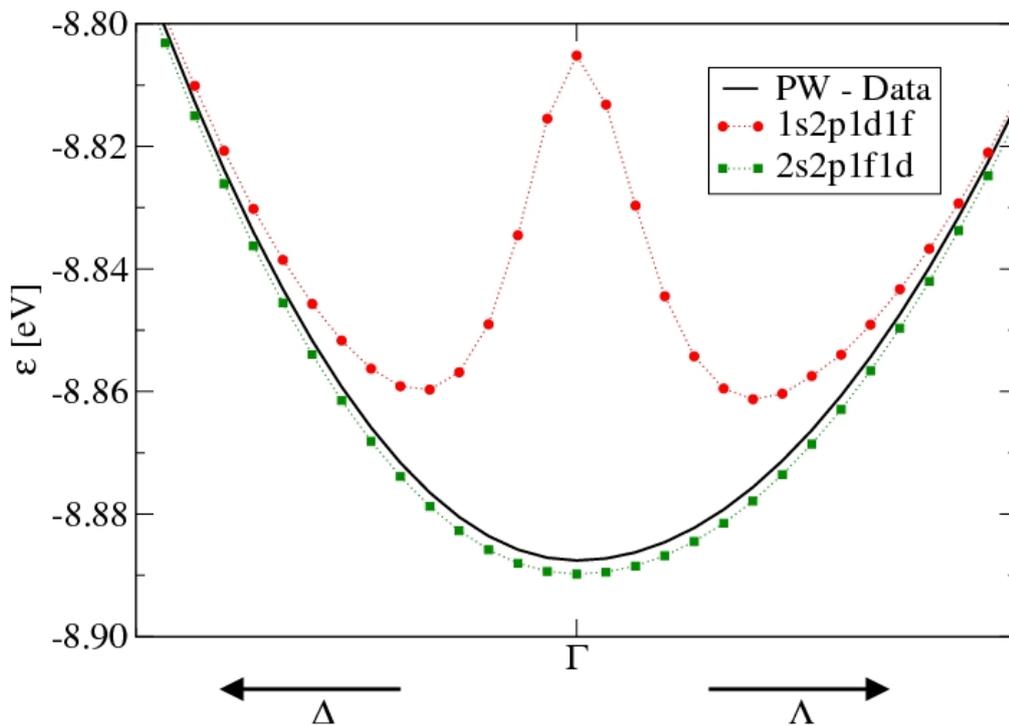
Abb. 6.8:

Silizium Bandstruktur berechnet unter Verwendung unterschiedlicher Basissätze. Das Valenzbandmaximum ist hierbei auf 0 eV referenziert.

$$P_{\alpha'\beta'}^{i_s,l} = \frac{1}{n_{\text{Atoms}}} \frac{1}{2l+1} \sum_{n,\mathbf{k}} \sum_{i_a,m} w_{\mathbf{k}} f_n c_{\mathbf{k},\alpha',n,i_s,i_a,l,m}^* c_{\mathbf{k},\beta',n,i_s,i_a,l,m}. \quad (6.50)$$

Diese Koeffizientenmatrix unterläuft hierbei einer Mittelung über alle Atome zur Atomsorte i_s , sowie der $2l+1$ Orbitalfunktionen zur Drehimpulsquantenzahl l . α' und β' sind hierbei nur noch als Nummerierung der einzelnen Orbitalfunktionen zu einer festgehaltenen Atomsorte und zu einer festgehaltenen Drehimpulszahl l zu verstehen. Diese Matrizen lassen sich diagonalisieren, ohne einer Vermischung unterschiedlicher l -Komponenten oder Atomsorten zu unterliegen. Allerdings gehen in diese Koeffizientenmatrizen nur die entsprechenden *on-site* Beiträge ein. Jene bilden damit nur eine grobe Abschätzung der tatsächlich reduzierten Basis. Durch eine nachträgliche Optimierung dieser neuen Basis kann dieser Makel aber beseitigt werden. Als Richtwert können solche Funktionen aus der Basis entfernt werden, deren Eigenwert kleiner ist, als das Produkt aus Spillage und maximalen Eigenwert.

Die so konstruierten Basissätze sind mit der resultierenden Spillage in Tabelle 6.4 wiedergegeben. Ein Vergleich mit anderen auf der Spillage basierenden Optimie-

**Abb. 6.9:**

Die Vergrößerung des untersten Bandes der Siliziumbandstruktur zeigt nahe des Γ -Punktes einen signifikanten Fehler (rote Kreise), wenn die unterpolarisierte $1s2p1d1f$ -Basis verwendet wird. Die balancierte $2s2p1f1d$ -Basis (grüne Quadrate) ist frei von diesem Fehler und reproduziert das Ergebnis der ebenen Wellen (schwarze Linie).

rungsmethoden zeigt, dass die hier optimierten Funktionen durchweg eine geringere Spillage aufweisen und stellenweise sogar die von Sánchez-Portal[77] vorhergesagte theoretische untere Schranke erreichen. Diese hängt jedoch von den Details der zu Grunde liegenden Ebenen-Wellen-Rechnung ab. Insbesondere das zugrunde liegende k -Punkt Gitter spielt hier eine Rolle, denn die in [77] verwendete punktweise Minimierung unterschätzt die untere Schranke für endliche k -Punktsätze stärker als die hier diskutierte spline-gestützte Variante. Diskrepanzen lassen sich auf diese Unterschätzung zurückführen.

Es stellt sich die Frage, ob die Orbitale auch eine ausreichende Transferabilität besitzen, oder ob durch die Optimierung die Orbitale nur für das System, für welches sie optimiert wurden, akkurat sind. Um die Transferabilität der Orbitale zu prüfen, werden in der Regel nicht symmetrische Verzerrungen des Materials berechnet. Tabelle 6.5 zeigt die mit den optimierten Basisfunktionen berechneten Elastizitätskonstanten. Diese besitzen eine vergleichbar gute Übereinstimmung mit

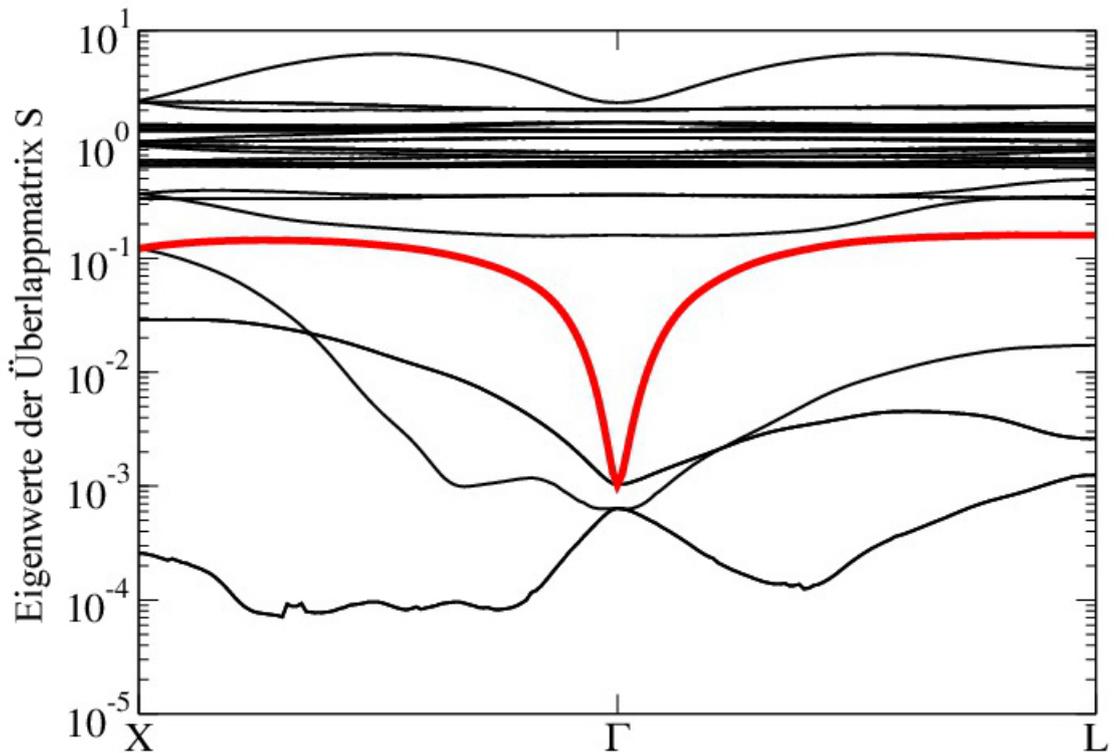


Abb. 6.10:

Entwicklung der Eigenwerte der Überlappmatrix $S(\mathbf{k})$ entlang des Pfades $X \rightarrow \Gamma \rightarrow L$ für den unterpolarisierten $1s2p1d1f$ -Basissatz. Die starke rote Linie hebt den Eigenwert hervor, welcher nahe des Γ -Punktes rapide singularär wird und das Problem um einen Freiheitsgrad beraubt.

anderen Ebenen-Wellen-Resultaten[78]. Von einer deutlichen Verminderung der Transferabilität kann also nicht gesprochen werden.

Die optimierten Orbitalfunktionen besitzen auch im Falle des unendlich ausgedehnten Kristalls eine Lokalisierung, wie Abbildung 6.7 zeigt. Ein deutliches Signal in der radialen Funktion lässt sich außerhalb von 10 Bohrradien nicht wahrnehmen. Die Reproduktion der Bandstruktur ist in Abbildung 6.8 wiedergegeben. Die besetzten Zustände werden mit einem Maximalfehler von 50 meV korrekt beschrieben. Mit steigender Größe des Basissatzes nimmt zudem die qualitative Beschreibung der unbesetzten Zustände zu.

Im Falle einer überpolarisierten Basis, die hier mit dem oben beschriebenen Verfahren der Eigenwertzerlegung erzeugt wurde, liefert die Bandstruktur ein interessantes Bild. Anhand der Spillage allein stellt sich die konstruierte $1s2p1d1f$ -Basis als gute Basis dar. Ihre Spillage ist gerade mal doppelt so groß wie die der fast balancierten $3s3p2d1f$ -Basis. Die Bandstruktur (Abbildung 6.9) zeigt jedoch eine minderwertige Beschreibung des niedrigsten, s -artigen Bandes nahe dem Γ -Punkt. Die Erklärung hierfür ist die erzwungene Beschreibung eines s -artigen Zustandes

durch am Nachbaratom lokalisierte p -Funktionen. Dies gelingt zwar über weite Teile der Bandstruktur recht gut, führt aber am Γ -Punkt zu eklatanten Fehlern, da sich hier die p -Funktionen symmetriebedingt aufheben. Der zuständige Überlapp wird singulär und verliert somit einen Freiheitsgrad wie Abbildung 6.10 zeigt. Solche Artefakte in der Bandstruktur können also auf eine fehlerhafte Beschreibung eines Zustandes durch Funktionen höherer Drehimpulsquantenzahlen zurückgeführt werden und bilden ein sicheres Kriterium dafür, dass die verwendete Basis überpolarisiert ist.

Im folgenden sind die Kernaussagen dieses Abschnittes noch einmal zusammengefasst.

- Eine Eigenwertzerlegung nach *on-site* Beiträgen erlaubt die Reduktion der Basissatzgröße ohne die Spillage wesentlich zu verschlechtern.
- Überpolarisierte Basisfunktionen neigen zu einem unphysikalischen Ausgleich von Unzulänglichkeiten von Orbitalfunktionen niedrigeren Drehimpulses durch Orbitalfunktionen höheren Drehimpulses, welche an einem anderem Zentrum lokalisiert sind (Superpositionsfehler). Diese Fehler fallen besonders bei der Betrachtung von Hochsymmetriepunkten ins Auge.
- Die optimierten Orbitalfunktionen weisen für das Materialsystem Silizium eine geringere Spillage auf als andere in der Literatur diskutierte Basisfunktionen, die unter der Minimierung der Spillage erzeugt wurden. Zudem besitzen die optimierten Orbitalfunktionen eine gute Transferabilität.

6.5.3. Wahl der Filterfaktoren - Aluminium

Metalle eignen sich in idealer Weise für das Studium der Filterfaktoren, da die Verwendung von Besetzungszahlen nun Bänder abschneidet, die das Fermi-niveau kreuzen. Aluminium bietet sich hierbei als Materialsystem an. Es kristallisiert in der kubischen flächenzentrierten Form. Die Modellierung des elektronischen Austausch- und Korrelationsfunktionals geschieht hier auf dem Niveau der generalisierten Gradientennäherung (GGA-PBE). Ein entsprechendes PAW-Potential ist der VASP-Datenbank[84] entnommen worden.

Metallische Systeme benötigen deutlich höhere \mathbf{k} -Punkt Gitter zur Konvergenz, was dem Quamolansatz zu Gute kommt. Ein $24 \times 24 \times 24$ \mathbf{k} -Punkt Gitter und ein Energiecutoff von 24.25 Ry (330 eV) liefern konvergierte Ergebnisse. Erneut wird die elektronische Dichte an einem nichtzentrierten \mathbf{k} -Punkt Gitter erzeugt um dann die Wellenfunktionen für das zentrierte Gitter zu bestimmen.

In einem ersten Schritt werden die Filterfaktoren in Übereinstimmung mit den Besetzungszahlen des Materialsystem gesetzt. Um einen relativ weichen Übergang

Basisname	Größe	S_{opt}	ΔE [meV]
PC0 - Pseudo	$2s2p$	$1.7 \cdot 10^{-3}$	
PC0 - Reduced	$1s1p$	$2.5 \cdot 10^{-3}$	396
PC1 - Pseudo	$2s2p1d$	$1.2 \cdot 10^{-4}$	
PC1 - Reduced	$1s1p1d$	$1.4 \cdot 10^{-4}$	125
PC2 - Pseudo	$3s3p2d1f$	$1.9 \cdot 10^{-5}$	
PC2 - Reduced	$2s2p1d1f$	$2.8 \cdot 10^{-5}$	112
PC2 - Reduced*	$2s2p1d1f$	$1.9 \cdot 10^{-4}$	117
Basisname	Größe	S_{opt}	$S_{\text{opt}}/S_{\text{opt}}^{\text{Quamol}}$
PAO[77]	$1s1p$	$1.57 \cdot 10^{-2}$	6.28
PAO[77]	$1s1p1d$	$7 \cdot 10^{-4}$	5.00
MIN S[77]	$1s1p$	$9 \cdot 10^{-4}$	0.36
MIN S[77]	$1s1p1d$	$2.9 \cdot 10^{-5}$	0.20

Tab. 6.6:

Spillage und Differenz der Gesamtenergie (selbst-konsistente LCAO vs. ebene-Wellen) für Aluminium. Der '*' bezeichnet einen Wechsel der Filterfaktoren in der PC2 Basis.

zu gewährleisten ist hierbei eine Fermiverteilung von 0.05 eV verwendet worden. Die optimierten Orbitalfunktionen werden — wie schon beim Silizium — generiert und anschließend über die Eigenwertzerlegung in ihrer Größe reduziert. Tabelle 6.6 zeigt die hier erreichte Spillage sowie einen Vergleich mit Literaturwerten. Es zeigt sich eine signifikante Verbesserung. Die theoretische untere Schranke wird diesmal nicht erreicht. Allerdings stellt Aluminium deutlich größere Ansprüche an das k-Punktgitter als Silizium, so dass die in [77] mit nur 10 irreduziblen k-Punkten berechnete untere Schranke wahrscheinlich deutlich unterschätzt wird. Die hier zugrunde liegenden Ebenen-Wellen-Rechnungen für Aluminium verwenden ca. 400 irreduzible k-Punkte.

Die mit diesen optimierten Orbitalfunktionen berechnete Bandstruktur (Abbildung 6.11) zeigt eine gute Übereinstimmung des besetzten Unterraumes. Abweichungen sind hier nicht größer als 48 meV. Energetisch höher liegende Zustände werden bis 20 eV über dem Fermi Level zumindest noch qualitativ gut beschrieben, auch wenn Abweichungen bis zu einigen eV verzeichnet werden können. Eine Änderung der Besetzungszahlen in Übereinstimmung mit einem künstlichen Fermi-niveau 13 eV oberhalb des tatsächlichen — ebenfalls mit einer Fermiverteilung von 0.05 eV — und entsprechender Optimierung hebt die Spillage zwar um eine Größenordnung an, verschlechtert die Energetik aber nur minimal wie Tabelle 6.6 zeigt. Die Bandstruktur zeigt im neu definierten Unterraum deutliche Verbesserungen. Die quantitative Übereinstimmung der Bänder bis zum künstlichen Fermi-niveau ist nun auch innerhalb eines Fehlers von 45 meV. Die qualitative Beschreibung hoher liegender Bänder reicht entsprechend weiter in den hohen

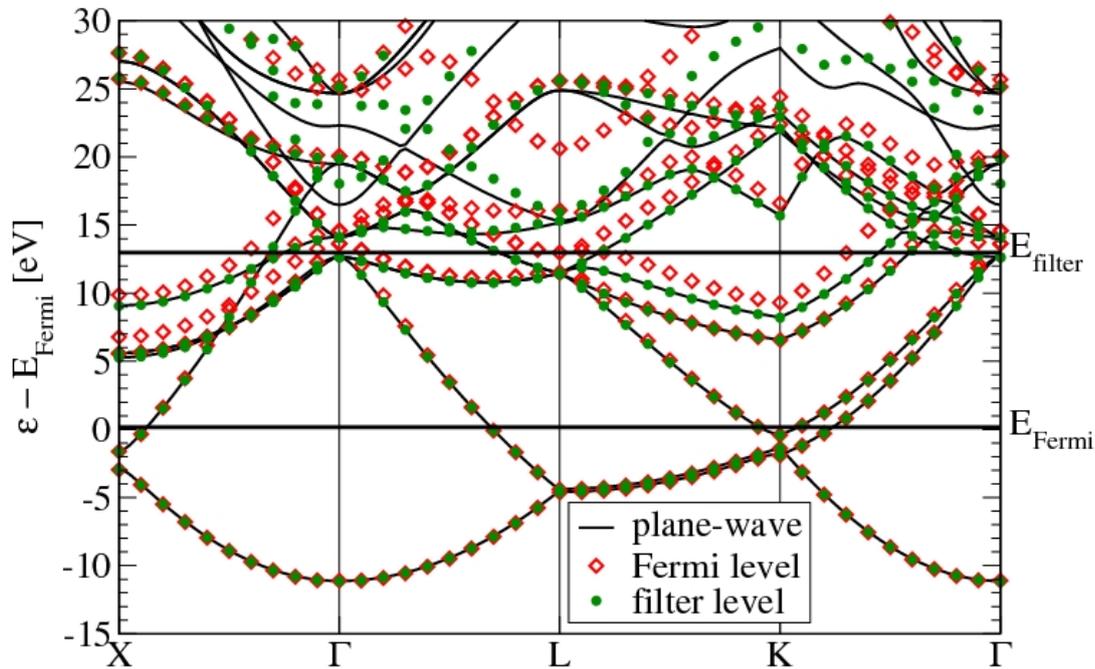


Abb. 6.11:

Bandstrukturrechnung für Aluminium mit unterschiedlichen Filterfaktoren. Die Energien sind hierbei auf das Fermi-niveau referenziert.

Energiebereich.

Die Filterfaktoren erlauben folglich eine gewichtete Fokussierung auf Teile der Bandstruktur, denen ein gesondertes Interesse gilt. So würden zur Effektivmassenberechnung von Elektronen und Löchern nur den Bandkanten der Bandlücke von Null verschiedene Filterfaktoren zugewiesen werden.

6.5.4. Transferabilität

Eine wichtige Eigenschaft eines atomaren Basissatzes ist seine Transferabilität. Die erzeugten Orbitalfunktionen sollen nicht nur eine adäquate Beschreibung des Systems ermöglichen, für welches sie generiert wurden, sondern ebenfalls Strukturoptimierungen zulassen. Dazu darf eine Verzerrung der atomaren Struktur die Beschreibung durch den Basissatz nicht eklatant verschlechtern. Bereits am Silizium ließ sich erkennen, dass mit den Orbitalfunktionen die Elastizitätskonstanten in guter Qualität bestimmt werden. Ein nicht ganz so stringenter, jedoch ebenfalls aussagekräftiger Schnelltest ist die Bestimmung der optimalen Gitterkonstante und

	Si			Al		
Basis	$1s1p1d$			$1s1p1d$		
	DFT-PW	LCAO	Δ	DFT-PW	LCAO	Δ
E_{tot} [H]	-8.0389	-8.0368	57 meV	-2.1054	-2.1047	19 meV
a_0 [bohr]	10.217	10.223	+0.06 %	7.636	7.6364	+0.01 %
B_0 [GPa]	94.6	102.7	+8.6%	79	96	+22%
Basis	$2s2p1d1f$			$2s2p1d1f$		
	DFT-PW	LCAO	Δ	DFT-PW	LCAO	Δ
E_{tot} [H]	-8.0389	-8.0382	19 meV	-2.1054	-2.1053	3 meV
a_0 [bohr]	10.217	10.218	+0.01 %	7.636	7.634	0.03 %
B_0 [GPa]	94.6	96.0	+1.5 %	79	82	+4%

Tab. 6.7:

Vergleich der strukturellen Eigenschaften von Silizium und Aluminium. Zur Berechnung sind Gesamtenergieberechnungen für verschiedene Gitterkonstanten durchgeführt worden. Die daraus erhaltene Energieoberfläche ist dann an die Murnaghansche Zustandsgleichung gefittet worden. DFT-PW betitelt die Ebenen-Wellen-Referenzdaten. LCAO betitelt die selbstkonsistente LCAO Berechnung unter Verwendung der angegebenen Basis.

des Bulkmodulus durch Fitten der Energieoberfläche an die Zustandsgleichung nach Murnaghan[85]. Tabelle 6.7 und Tabelle 6.8 geben hier die Ergebnisse für die Festkörpersysteme Si und Al, sowie für GaN und NaCl wieder. Schon ein minimaler Basissatz liefert qualitativ richtige Ergebnisse. Der Bulkmodulus zeigt jedoch noch Abweichungen von bis zu 20%. Durch Hinzunahme weiterer Funktionen kann die quantitative Übereinstimmung jedoch gesteigert werden. Bei der Verwendung einer $2s2p1d1f$ Basis erreichen wir eine Übereinstimmung bis auf wenige Prozent. Zur Bestimmung des Bulkmodulus werden in der Regel Variationen der Gitterkonstante bis hin zu 2% verwendet.

Drastischere Änderungen der atomaren Struktur ergeben sich beispielsweise bei der Modellierung der NH_3 *umbrella* Schwingungsmode bis zum molekularen Umklapp-Prozess. Als optimierter Basissatz ist hierbei eine $2s2p1d$ -Basis für den Stickstoff, sowie eine $1s1p$ -Basis für den Wasserstoff gewählt worden. Die Basissätze wurden an der molekularen Gleichgewichtsstruktur optimiert. Zur Bestimmung der Energiebarriere startet man von dem planaren Molekül und bewegt das Stickstoffatom senkrecht aus dieser Ebene heraus. Die Wasserstoffatome dürfen in ihrer Ebene frei relaxieren. Der entsprechende Energieverlauf ist in Abbildung 6.12 gezeigt. Die Lage des Energieminimums und die Höhe der Energiebarriere sind in Tabelle 6.9 zusammengefasst. Während die Gesamtenergiedifferenz mit 1.32 eV groß ist, wird die Energiebarriere bis auf 19 meV reproduziert.

	GaN			NaCl		
Basis	1s1p1d + 1s1p1d			1s1p + 1s1p1d		
	DFT-PW	LCAO	Δ	DFT-PW	LCAO	Δ
E_{tot} [H]	-24.8054	-24.7975	215 meV	-15.4976	-15.4946	82 meV
a_0 [bohr]	6.041	6.047	+0.1 %	10.6878	10.7411	+0.5 %
B_0 [GPa]	211	233	+11%	24.3	29.3	+21%
Basis	2s2p1d1f + 2s2p1d1f			2s2p1d1f + 1s1p1d1f		
	DFT-PW	LCAO	Δ	DFT-PW	LCAO	Δ
E_{tot} [H]	-24.8054	-24.8037	46 meV	-15.4976	-15.4976	6 meV
a_0 [bohr]	6.041	6.043	+0.03 %	10.688	10.688	0.00 %
B_0 [GPa]	211	215	+2 %	24.3	25.5	+5%

Tab. 6.8:

Vergleich der strukturellen Eigenschaften von Galliumnitrid und Natriumchlorid. Die Berechnung der Daten erfolgte analog zu Tabelle 6.7

	ΔE_{opt} [eV]	$d_{\text{N-H}_3}$ [Å]	E_{Barriere} [meV]
DFT-PW	0.00	0.37	157
LCAO initial guess	10.08	0.67	1630
LCAO Quamol	1.32	0.38	176

Tab. 6.9:

Vergleich unterschiedlicher Kerndaten der NH_3 *umbrella* Schwingungsmode. Als LCAO Basis ist für das Stickstoff eine $2s2pd$ und für den Wasserstoff eine sp -Basis verwendet worden. Der Abstand $d_{\text{N-H}_3}$ bezeichnet den optimalen Abstand des Stickstoffes von der H_3 -Ebene. ΔE_{opt} bezeichnet die Energiedifferenz der optimierten Struktur in der Ebenen-Wellen-Rechnung und der optimierten Struktur in der LCAO Rechnung unter Verwendung des atomaren Basissatzes. E_{Barriere} beziffert die berechnete Energiebarriere des Umklappprozesses wie in Abb. 6.12 dargestellt.

Ein weiterer Transferabilitätstest ist die Reaktion der Basis auf eine sich ändernde chemische Umgebung im Sinne von Bindungswinkeln und insbesondere Mehrfachbindungen. Für diesen Test sind Silan, Silen und Silin zur Modellierung in der *trans*-Konfiguration, vgl. Abbildung 6.13, gewählt worden. Die Si-Si Bindung ändert sich hierbei von einer Einfachbindung bis hin zu einer Dreifachbindung mit einer entsprechen Verkürzung der Bindungslänge. Der H-Si-Si Bindungswinkel variiert zwischen 110° und 125° . Als Basis ist für das Si eine $2s2p1d1f$ -Basis und für den Wasserstoff eine $2s2p1d$ -Basis gewählt worden. Diese Basis wird für jeweils eine Geometrie optimiert und dann auf die zwei verbleibenden Konfigurationen angewandt. Die Spillage, sowie die Unterschiede in der Gesamtenergie im Hinblick auf das Resultat der Ebenen-Wellen-Rechnung sind in Tabelle 6.11 zusammengefasst. Signifikante Einflüsse der Referenzstruktur lassen sich nicht feststellen, die Gesamtenergie wird mit jeder Basis innerhalb von 40 meV reproduziert.

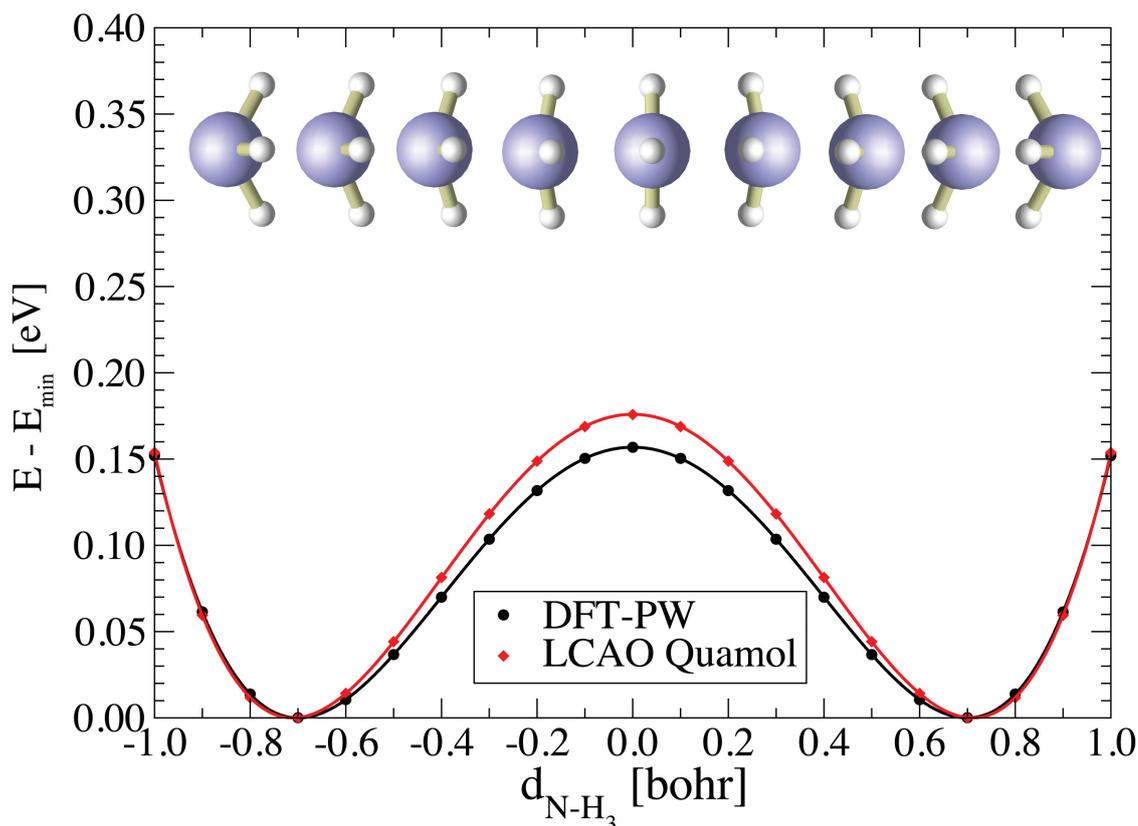


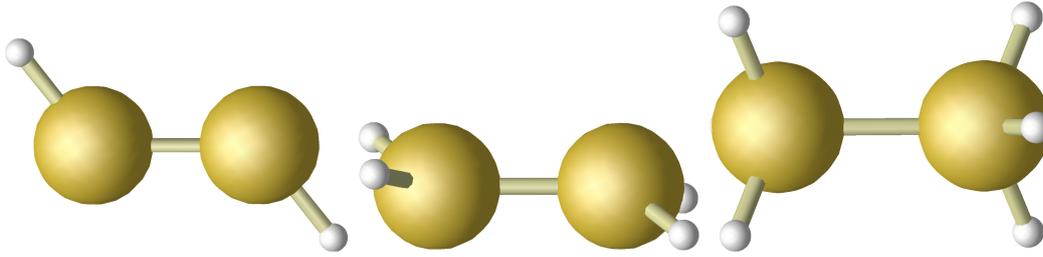
Abb. 6.12: Energieoberfläche der NH_3 *umbrella* Mode. Der Reaktionspfad ist oberhalb der Kurve visualisiert.

Zusammenfassend lässt sich also sagen:

- Moderate Verzerrungen der ionischen Struktur lassen sich mit den optimierten Orbitalfunktionen abbilden und energetisch richtig beschreiben.
- Moderate Änderungen der chemischen Umgebung (unterschiedliche Anzahl der Bindungspartner und *closed-shell* Bindungsgeometrien) werden durch die optimierten Orbitalfunktionen innerhalb geringer energetischer Fehler gut beschrieben.

	$d_{\text{Si-Si}}$ [Å]	$d_{\text{Si-H}}$ [Å]	$\alpha_{\text{H-Si-Si}}$ [°]	$\alpha_{\text{H-Si-H}}$ [°]
<i>trans</i> - Si_2H_2	2.09	1.50	125.3	—
<i>trans</i> - Si_2H_4	2.15	1.49	119.5	113.4
<i>trans</i> - Si_2H_6	2.32	1.50	110.4	108.5

Tab. 6.10: Strukturelle Eigenschaften von *trans*-Disilan, *trans*-Disilen und *trans*-Disilin. Berechnet wurden diese mittels eines ebenen-Wellen LDA-DFT Ansatzes.

**Abb. 6.13:**

Visualisierung von *trans*-Disilin (links), *trans*-Disilen (mitte) und *trans*-Disilan (rechts). Die gelben Sphären stellen die Siliziumatome dar, die weißen Sphären entsprechen den Wasserstoffatomen..

Si ₂ H ₂ Basis	Si ₂ H ₂	Si ₂ H ₄	Si ₂ H ₆
Spillage	$4.93 \cdot 10^{-5}$	$5.13 \cdot 10^{-5}$	$4.41 \cdot 10^{-5}$
ΔE [meV]	20	26	33
Si ₂ H ₄ Basis	Si ₂ H ₂	Si ₂ H ₄	Si ₂ H ₆
Spillage	$7.26 \cdot 10^{-5}$	$4.37 \cdot 10^{-5}$	$3.45 \cdot 10^{-5}$
ΔE [meV]	22	25	29
Si ₂ H ₆ Basis	Si ₂ H ₂	Si ₂ H ₄	Si ₂ H ₆
Spillage	$7.40 \cdot 10^{-5}$	$5.91 \cdot 10^{-5}$	$3.40 \cdot 10^{-5}$
ΔE [meV]	29	31	28

Tab. 6.11:

Sich ergebene Spillage und Energiedifferenzen der selbstkonsistenten LCAO Berechnung im Vergleich zur Ebenen-Wellen-Rechnung. Die verwendete atomare Basis ist in der linken Spalte angegeben.

6.6. Lokalisierung und Tight-Binding - Ein Ausblick

Die optimierten Orbitalfunktionen zeigen eine natürliche Lokalisierung. Betrachtet man jedoch die abstandsabhängige Entwicklung der Diagonalelemente der Hamiltonmatrix und der Überlappmatrix, so ist kein exponentielles Abklingverhalten erkennbar (vgl. Abbildung 6.14). Folglich liefert der Überlapp von Orbitalfunktionen an entfernten Atomen einen nicht verschwindenden Beitrag. Dies spiegelt sich in minimalen Fluktuationen im langreichweitigen Teil der radialen Orbitalfunktionen wieder. Die Auswirkungen dieser Fluktuationen für einen Tight-Binding Ansatz, der in der Regel nur Wechselwirkungen bis zum übernächsten Nachbar berücksichtigt, werden nun im Folgenden diskutiert.

Dazu werden die Hamiltonmatrizen für Silizium basierend auf einer optimierten $2s1p1d$ -Basis im reziproken Raum an verschiedenen \mathbf{k} -Punkten konstruiert. Die abstandsabhängigen Beiträge werden bis zu einem Radius von 10 Bohr berücksichtigt. Die Lösung des verallgemeinerten Eigenwertproblems der \mathbf{k} -abhängigen Hamilton-

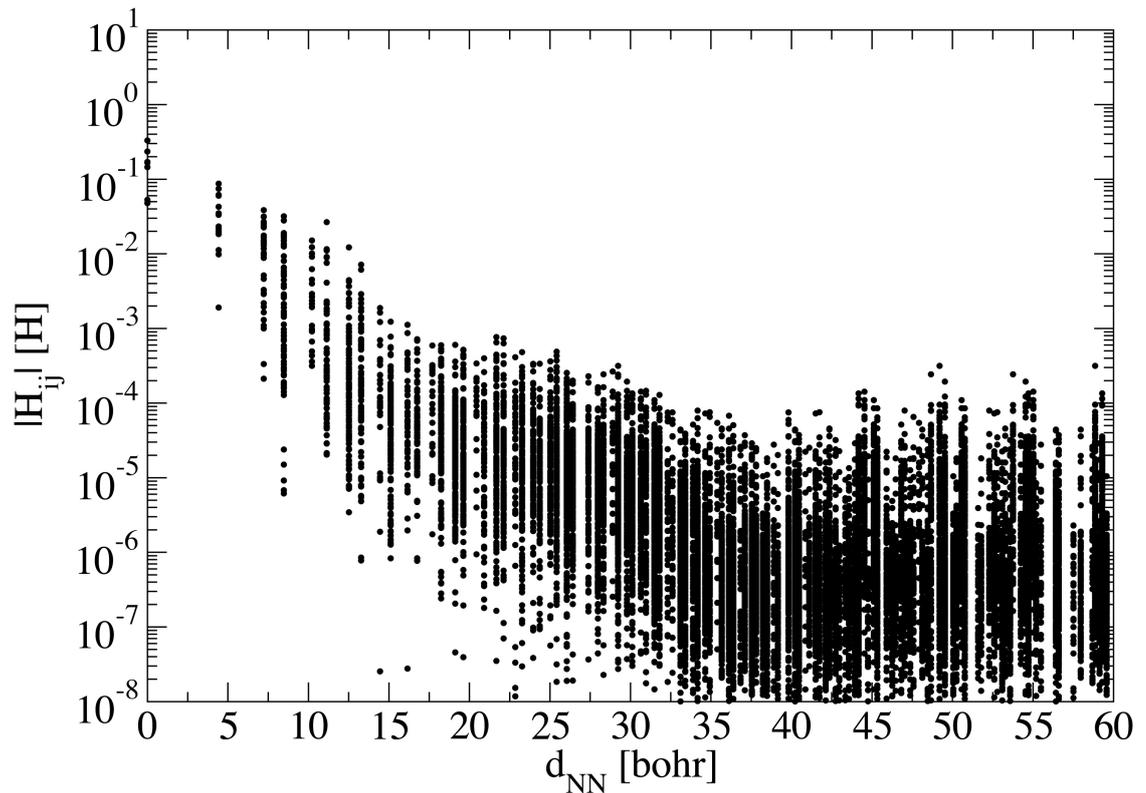


Abb. 6.14:

Entwicklung der Hauptdiagonalelemente der entfernungsabhängigen Hamiltonmatrizen für bulk Silizium. Als Basissatz ist eine $2s1p1d$ -Basis verwendet worden, die für den besetzten Unterraum optimiert wurde.

matrizen gibt die Eigenwerte an den entsprechenden \mathbf{k} -Punkten wieder und somit die Bandstruktur. Ein Vergleich dieser Bandstruktur mit der LCAO Bandstruktur zeigt eklatante, qualitative Abweichungen (vgl. Abbildung 6.15). Sollen die Quamols für Tight-Binding-Zwecke eingesetzt werden, ist es unabdingbar, diese Fluktuationen im langreichweitigen Teil der Orbitalfunktionen zu dämpfen. Tatsächlich lassen sich die Fluktuationen für Moleküle beseitigen, in dem Superzellen mit einer Kantenlänge von 40 Bohr und mehr verwendet werden, um die Wechselwirkung der Atome mit ihren periodischen Abbildern zu reduzieren. Dies ist jedoch bei der Beschreibung von Festkörpern nicht möglich. Hier bedarf es eher einer in den Algorithmus eingebetteten Lokalisierungsbedingung, um die Wechselwirkungen mit weit entfernten Atomen zu dämpfen. Ein erster Schritt wäre die Definition von Cutoffradien, hinter denen die Orbitalfunktionen auf Null gesetzt werden. Dies würde jedoch zu harten Sprüngen in den Orbitalfunktionen führen und somit zu einem unkontrollierbarem Verhalten in den Fouriertransformationen. Vielmehr bietet es sich an das HSE-Konzept für das Austauschpotential[43] auf die Quamols zu übertragen und die Orbitalfunktionen zu abschirmen. Dafür verwenden wir an dieser Stelle als Abschirmfunktion die Fermi-Dirac-Funktion

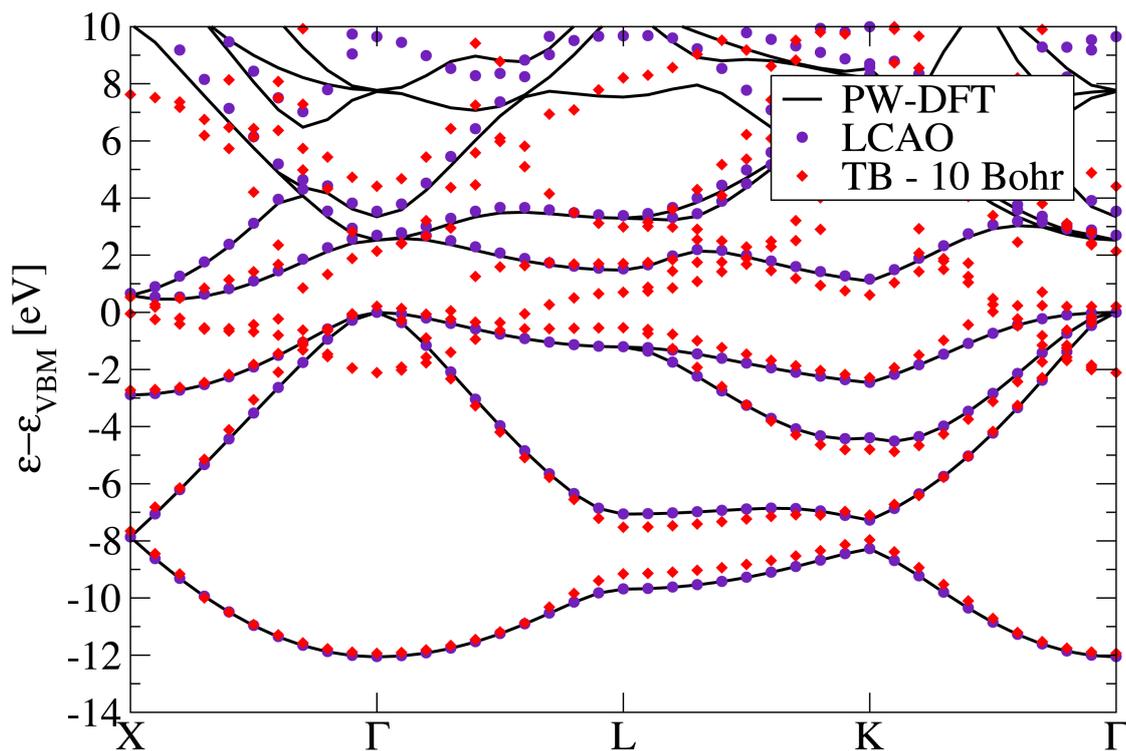


Abb. 6.15:

Vergleich der Bandstrukturen für bulk Silizium berechnet in der Ebenen-Wellen Basis, der $2s1p1d$ -Quamolbasis und dem in dieser Basis verwendeten Tight-Binding-Ansatz, welcher Wechselwirkungen innerhalb einer 10 Bohr Sphäre mitnimmt.

$$\Pi(r; r_0, \beta) = \frac{1}{1 + e^{\beta(r_0 - r)}}. \quad (6.51)$$

Diese variiert zwischen 0 und 1. Dabei gibt r_0 die Position an, an der die Funktion den Wert 0.5 hat, der Cutoffradius. β ist mit der Breite des Überganges verknüpft. Hierbei ist der Übergang umso härter, je größer β ist. Weiterhin hat die Fermi-Dirac-Funktion, im Vergleich zu einer Stufenfunktion, den Vorteil der Differenzierbarkeit in jedem Punkt und wird somit keine Kanten in die Radialfunktionen induzieren. Die Lokalisierung $\bar{R}(r)$ des Radialteils $R(r)$ einer Orbitalfunktion lässt sich nun über

$$\bar{R}(r) = (1 - \Pi(r; r_0, \beta))R(r) \quad (6.52)$$

erreichen. Das Optimierungsfunktional ist nun durch

$$\begin{aligned} \mathbb{F} = & \mathbb{S} \\ & + \kappa \sum_{\alpha} \langle \mu_{\alpha} | \Pi_{\alpha}(r) | \mu_{\alpha} \rangle \\ & - \sum_{\alpha, \beta' (l(\alpha))} \epsilon_{\alpha\beta'} (\langle \mu_{\alpha} | \mu_{\beta'} \rangle - \delta_{\alpha\beta'}) \end{aligned} \quad (6.53)$$

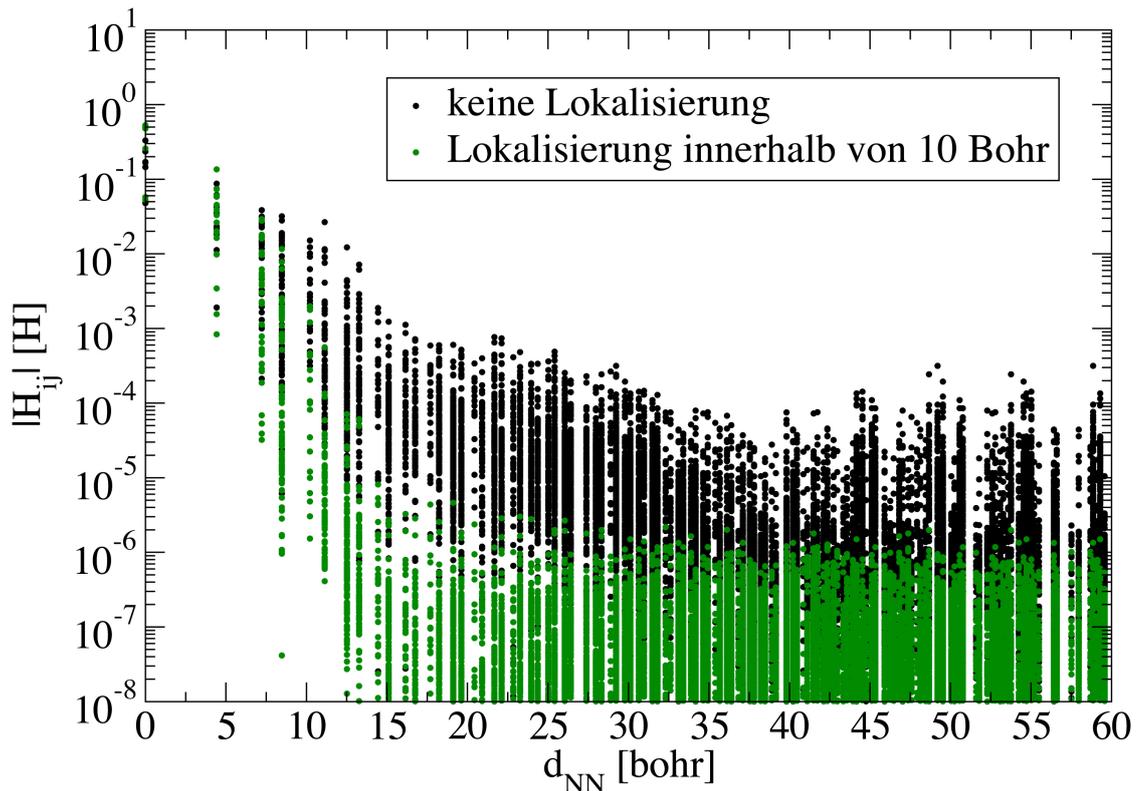


Abb. 6.16:

Entwicklung der Hauptdiagonalelemente der entfernungsabhängigen Hamiltonmatrizen für bulk Silizium. Als Basissatz ist eine $2s1p1d$ -Basis verwendet worden, die für den besetzten Unterraum ohne (schwarz) und mit (grün) Anwendung des Lokalisierungskriteriums optimiert wurden.

zu ersetzen. Die Spillage wird nun um einen additiven Term ergänzt, der die Lokalisierung bewertet. Da dies jedoch für unterschiedliche Normierungen der Orbitalfunktionen zu unterschiedlichen Gewichtungen führen würde, ist nun eine explizite Orthonormalisierung erforderlich. Dafür sorgt der dritte Term im Funktional. Mit κ kann die Lokalisierungsbedingung gewichtet werden. Unsere Rechnung haben gezeigt, dass die Optimierung nicht sensibel auf κ reagiert, wenn κ in einem Bereich von 0.1 - 1.0 gewählt wird. Deutlich höhere Werte legen ein zu starkes Gewicht auf die Lokalisierung und verhindern die Optimierung. Deutlich kleinere Werte schwächen die Lokalisierung so weit ab, dass die langreichweitigen Fluktuationen der Funktion wieder auftreten. Der Lagrangeparameter $\epsilon_{\alpha\beta}$ bestimmt sich über das Stabilitätskriterium zu

$$\epsilon_{\alpha\beta} = \left\langle \mu_\beta \left| \frac{dS}{d\langle \mu_\alpha \rangle} \right. \right\rangle + \kappa \langle \mu_\beta | \Pi_\alpha(r) | \mu_\alpha \rangle \text{ für } l(\alpha) = l(\beta). \quad (6.54)$$

Vollzieht man nun mit diesem Funktional die Optimierung der Radialfunktionen für Silizium und wählt als Cutoffradius 10 Bohr sowie eine Breite der Fermiverteilung von einem Bohr, so erhält man Orbitalfunktionen, die deutlich besser loka-

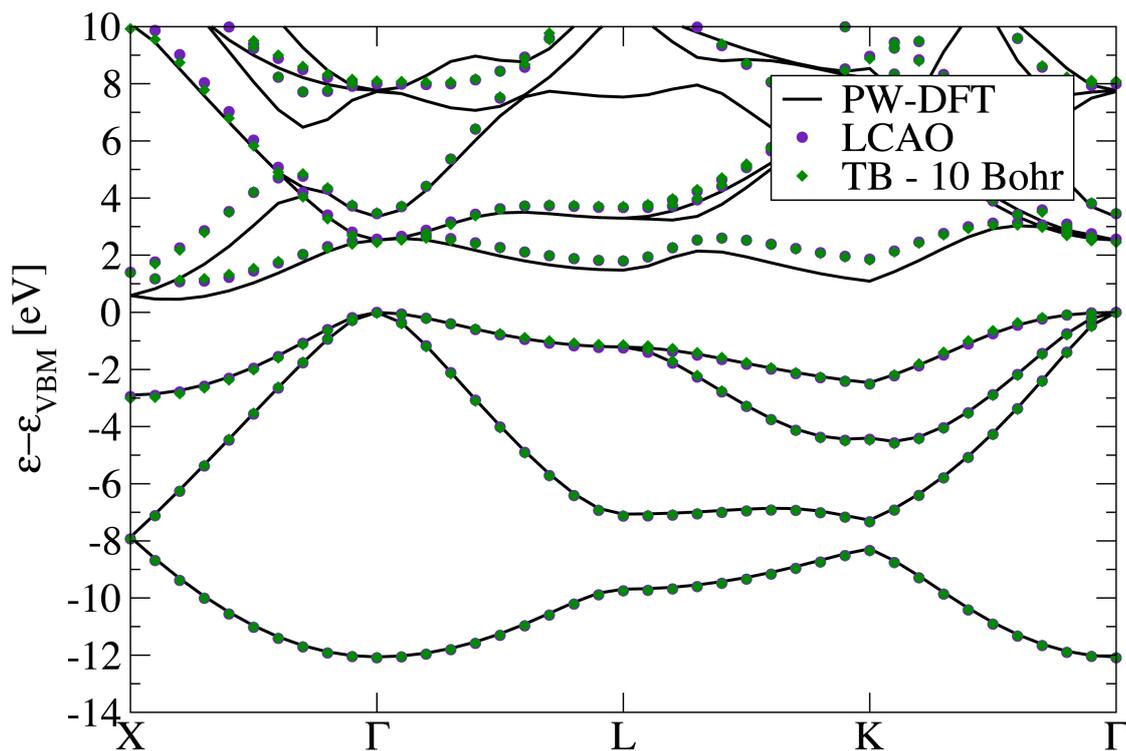


Abb. 6.17:

Vergleich der Bandstrukturen für bulk Silizium berechnet in der Ebenen-Wellen-Basis, der $2s1p1d$ -Quamolbasis und dem in dieser Basis verwendeten Tight-Binding-Ansatz, welcher Wechselwirkungen innerhalb einer 10 Bohr Sphäre mitnimmt. Die Basisoptimierung ist unter Verwendung der Lokalisierungsbedingung für einen Cutoffradius von 10 Bohr und einer Übergangsbreite von einem Bohr vollzogen worden.

lisiert sind, aber noch eine akzeptable Spillage aufweisen. Die Matrixelemente der Hamiltonmatrix klingen deutlich schneller ab, wie Abbildung 6.16 zeigt. Berechnet man nun die Bandstruktur mittels eines Tight-Binding-Ansatzes und nimmt alle Wechselwirkungen innerhalb von 10 Bohr (sechsnächster Nachbar) mit, so ergibt sich die Bandstruktur in Abbildung 6.17. Diese zeigt eine gute Übereinstimmung mit der LCAO Bandstruktur. Die vorherigen qualitativen Differenzen für den optimierten Unterraum sind mit diesen neuen Funktionen beseitigt. Jedoch bleibt eine quantitative Abweichung in der Umgebung von \mathbf{X} und \mathbf{K} in den untersten unbesetzten Bändern, welche vor Berücksichtigung der Lokalisierungsbedingung nicht vorhanden war. Eine explizite Mullikensche Populationsanalyse[86] offenbart einen signifikanten d -Charakter des Bandes ($\sim 30\%$ pro Atom) an diesen \mathbf{k} -Punkten, wo hingegen die unbesetzten Bandabschnitte, die eine gute Übereinstimmung zur Ebenen-Wellenrechnung zeigen, von s - und p -Beiträgen dominiert werden. Dies legt die Interpretation nahe, dass die Lokalisierung das d -Orbital in der Optimierung so einschränkt, dass es lediglich seine Aufgabe als Polarisationsfunktion wahrnehmen kann und die quantitative Übereinstimmung in der nicht lokalisierten Variante durch langreichweitige Wechselwirkungen zufällig erzeugt wurde. Eine explizite Hinzunah-

me der 4 untersten unbesetzten Bänder durch Auffüllen mit künstlichen Elektronen verbessert die Übereinstimmung in der lokalisierten Variante auch ohne Vergrößerung des Basissatzes signifikant.

6.7. Zusammenfassung und Ausblick

In diesem Kapitel ist eine Methode vorgestellt worden, um atomare Orbitalfunktionen basierend auf Ebenen-Wellen-Rechnungen zu generieren und optimieren. Grundlage der Optimierung bildet als quantitatives Maß die Spillage: der gewichtete Normverlust bei der Projektion eines Ebenen-Wellen-Zustandes auf die atomzentrierte Basis. Dabei verwenden die hier vorgestellten Orbitale keine repräsentierende Basis, was sie im Vergleich zu anderen spillagbasierten Verfahren unterscheidet. Der analytische Ausdruck der Spillage ermöglicht den Zugang zu gradientenbasierten Minimierungsverfahren. Eine Herausforderung stellt hier die Berechnung der Projektionen der im reziproken Raum beheimateten ebenen Wellen mit den im Ortsraum beheimateten Radialfunktionen dar. Das sich aus den abgespeicherten $\mathbf{G} + \mathbf{k}$ Vektoren der ebenen Wellen ergebene reziproke radiale Gitter ist nahe des Ursprunges nicht dicht genug um die reziproken Orbitalfunktionen repräsentieren zu können. Dafür verfügt es über eine extrem dichte Anzahl an Stützpunkten in dem Bereich, wo die Radialfunktionen schon längst abgeklungen sind. Dies macht eine direkte Fouriertransformation zwischen diesen beiden Räumen instabil und rechenintensiv. Erst der in dieser Arbeit verwendete Übergang zu reziproken Splinefunktionen ermöglicht eine effiziente und stabile Fouriertransformation zwischen diesen beiden Räumen und erlaubt einen stabilen Algorithmus zu Erzeugung der Quamols.

Die optimierten Basisfunktionen sind für unterschiedliche chemische Umgebungen bestimmt und die entsprechenden LCAO Resultate mit den Ebenen-Wellen-Resultaten verglichen worden. Es zeigt sich durchgehend eine gute Übereinstimmung. Bezüglich der Spillage unterbieten die Quamols andere Verfahren deutlich und reichen vereinzelt sogar an die theoretisch vorhergesagt untere Schranke heran. Die Diskrepanzen zu dieser lassen sich auf die Details der zu Grunde liegenden Ebenen-Wellen-Rechnung zurückführen und sind keine Unzulänglichkeiten des Quamolansatzes. Die erzeugten Funktionen sind auch für elektronischen Berechnungen von abweichenden atomaren Strukturen ohne große Einbußen an Genauigkeit verwendbar und in diesem Sinne transferabel. Elastizitätskonstanten, Bulkmoduli und Schwingungsmoden unterschiedlicher Systeme stimmen mit den Ebenen-Wellen-Resultaten bis auf wenige Prozent überein.

Um große Systeme zu betrachten ist die Anzahl der betrachteten Wechselwirkungen im System zu reduzieren. Dies ist im Falle gut lokalisierter Orbitale in einem Tight-Binding-Ansatz möglich. Aus den mit den Quamols erzeugten Hamiltonma-

trizen folgt jedoch, dass die natürliche Lokalisierung der Orbitalfunktionen nicht ausreichend ist, um akkurate Ergebnisse zu erzielen. Es zeigen sich gravierende Abweichungen in der Bandstruktur verglichen mit den Ebenen-Wellen-Resultaten. Um nun eine explizite Lokalisierung zu fordern ist das Minimierungsfunktional um eine Lokalisierungsbedingung erweitert worden. Dem Grundgedanken nach bestraft diese Bedingung Signal in der Radialfunktion ausserhalb eines festgelegten Radius. Dabei ist sicherzustellen, dass der Übergang stetig und differenzierter erfolgt. Am Beispiel des Siliziums wurde gezeigt, dass mit diesen Lokalisierungsverfahren gute Ergebnisse erzielt werden, wenn bei der Konstruktion des Tight-Binding-Hamiltonians Wechselwirkungen bis hin zum gewählten Cutoffradius berücksichtigt werden.

Zu Beginn dieser Dissertation sind die Quamols entstanden mit dem Ziel, Defekte in großen Superzellen zu berechnen. Gerade der Tight-Binding-Ansatz erlaubt hier die effektive Behandlung mehrerer tausend Atome. Die Quamols fungieren dabei als akkurater, atomzentrierter Basissatz. Allerdings haben die mit der Modellierung von Magnesiumnitrid erhaltenen Ergebnisse gezeigt, dass die Inversionsdomäne nicht für die signifikante Modifikation der Wasserstoffkonzentration in magnesiumdotiertem Galliumnitrid verantwortlich ist. Aus diesem Grund sind die nötigen Entwicklungen, um effizient atomare Strukturen mit mehreren 10.000 Atomen zu berechnen nicht weiter angegangen worden. In dieser Dissertation verbleiben die Quamols als Orbitalfunktionen zu Analyse Zwecken, sowie geeignete Initialisierungsfunktionen für die Ebenen-Wellen Rechnungen bezüglich des magnesiumdotierten Galliumnitrids und seiner Defekte.

Dennoch ist mit der Entwicklung der Quamols der Grundstein gelegt große Systeme effizient und akkurat zu behandeln. Die Quamols lassen sich aber auch hervorragend zu Analyse Zwecken einsetzen. Für die Mullikensche Populationsanalyse von Ebenen-Wellen-Resultaten bilden sie einen systemnäheren Basissatz als die sonst Verwendeten atomaren Lösungen der einzelnen beteiligten Atomsorten. Ebenso finden sie auch in der Bandstrukturanalyse zur Charakterbestimmung eines Zustandes Anwendung. Besonders interessant ist die Erzeugung von abstandsabhängigen Hamiltonmatrizen mit der Möglichkeit gezielt Wechselwirkungen einzelner Orbitale an- und abzuschalten um so beispielsweise den Einfluss der besetzten Gallium d -Orbitale in GaN zu analysieren.

7. Ableitung theoretischer Grenzen in der p -Dotierbarkeit von GaN:Mg

Give me a fruitful error any time, full of seeds, bursting with its own corrections. You can keep your sterile truth for yourself.

Vilfredo Frederico Pareto, italienischer Ingenieur und Soziologe, 1848 - 1923

7.1. Einleitung

Die Verbindung von Elementen der dritten Hauptgruppe (Al, Ga und In) mit Stickstoff — die sogenannten Gruppe-III-Nitride — sind die vornehmlich verwendeten Materialsysteme für die Entwicklung von Licht emittierenden Dioden (LEDs)[87, 5]. Grund dafür ist die große Bandlücke dieser Materialsysteme, welche durch geeignete Legierungen Licht des kompletten sichtbaren Spektrums emittieren können. Ein Durchbruch in der Entwicklung der Galliumnitrid (GaN) basierten Dioden konnte durch die Erkenntnis erzielt werden, dass eine hohe p -Leitfähigkeit in mit Magnesium (Mg) dotiertem GaN durch ein thermische Aktivierungsverfahren nach dem Wachstum erreicht werden kann. Im theoretischen Modell bindet Wasserstoff an den Magnesiumakzeptor in Form eines neutralen [MgH] Defektkomplexes und passiviert diesen[5, 4, 6, 7]. Allerdings ist die Löslichkeit des neutralen Komplexes in GaN signifikant größer als die des Akzeptors allein, womit mehr Magnesium in das Wirtssystem gebracht werden kann. In der thermischen Aktivierung dissoziiert dann der Komplex in das gewünschte Mg_{Ga}^- und ein mobiles H^+ , welches schon bei verhältnismäßig niedrigen Temperaturen durch das Material diffundieren kann und an der Oberfläche gasförmigen Wasserstoff bildet[6, 8].

Dieses theoretische Konzept bildet aktuell die Grundlage experimenteller Forschung. Die Befunde weisen aber darauf hin, dass der Aktivierungseffekt durch die Magnesiumkonzentration limitiert ist. Experimentell zeigt sich ein verminderter Wasserstoffeinbau ab Magnesiumkonzentrationen oberhalb von 10^{19} cm^{-3} [9]. Eine nachträgliche Aktivierung führt nur zu einer minimalen Steigerung der p -Leitfähigkeit. Experimentelle Resultate deuten auf drei Mechanismen für diese Beobachtung hin:

1. Ab einer Magnesiumkonzentration von 10^{19} cm^{-3} bilden sich Inversionsdomänen (vgl. Abb. 7.1), an deren Grenzen sich eine parasitäre Magnesiumnitrid (Mg_3N_2) Phase befindet[12, 10, 13, 14]. Das in dieser Phase gebundene Magnesium ist als Akzeptor verloren. Weiter ist unklar, in welcher Art und Weise diese Phase Wasserstoff binden kann und sich somit negativ auf die Bildung der [MgH] Komplexe auswirkt.
2. Die thermische Austreibung des Wasserstoffes geschieht nicht vollständig. Experimentelle Messungen mittels *Sekundärionenmassenspektroskopie* (SIMS) zeigen ein relatives Wasserstoff-zu-Magnesium-Verhältnis von 10-20 %.

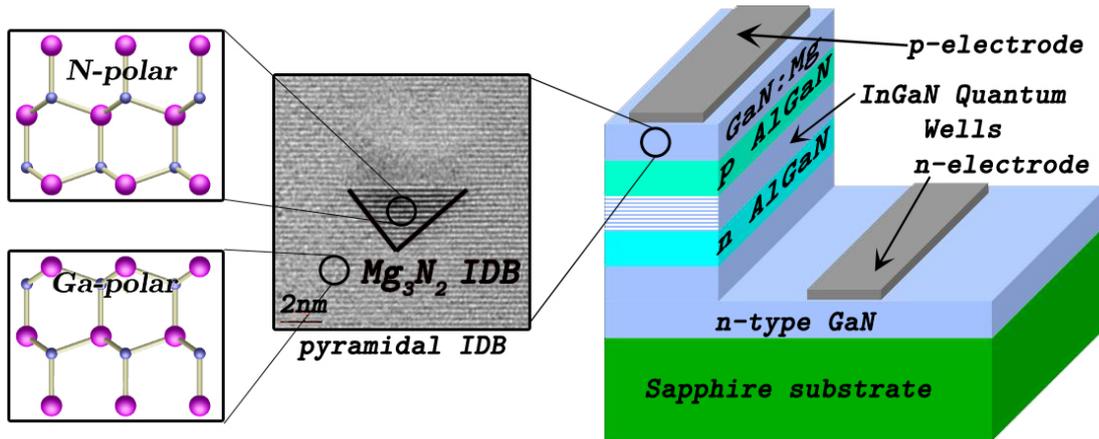


Abb. 7.1:

Aufbau einer GaN Diode. Im p -leitenden GaN:Mg kann es bei hohen Magnesiumkonzentrationen zur Ausformung von Inversionsdomänen kommen, deren Grenzen durch Magnesiumnitrid gebildet werden.

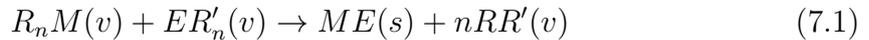
3. Neben dem Wasserstoff sind noch andere kompensierende Mechanismen denkbar, die bei hohen Magnesiumkonzentrationen zu Tage treten und die Rolle des Wasserstoffes übernehmen, jedoch keine oder eine schwierigere thermische Aktivierung erlauben.

Die Analyse und das Verständnis dieser einschränkenden Mechanismen, ebenso wie die Möglichkeit die Mechanismen zu kontrollieren sind der Schlüssel dazu, die heutigen Limitierungen der p -Dotierung zu überwinden und die Effizienz solcher Bauelemente signifikant zu steigern. Dabei muss das bisherige Bild eine Revision erfahren, da es die drei zentralen experimentellen Beobachtungen nicht erklären kann. In den folgenden Abschnitten werden nun mittels einer Dichtefunktionalstudie mögliche Mechanismen untersucht um ein erklärendes Modell zu entwickeln.

7.2. Kristallwachstum mittels MOVPE — Die experimentellen Randbedingungen

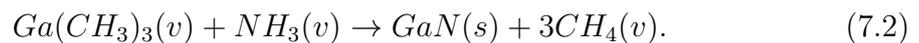
Zum Verständnis der einschränkenden Mechanismen in der p -Dotierbarkeit von GaN werden typische Defekte in GaN und — für die Inversionsdomänengrenzen — in Mg_3N_2 betrachtet. Aus deren Defektenergetiken können unter Annahme des thermodynamischen Gleichgewichtes Konzentrationsprofile erstellt werden. Dabei ist es notwendig die vorherrschenden experimentellen Parameter zu kennen, die implizit in den chemischen Potentialen, bezüglich der Temperatur aber auch explizit in der Boltzmannverteilung eine Rolle spielen. An dieser Stelle wird nun die Kristallwachstumsmethode MOVPE grob umrissen.

Die metall-organische Gasphasenepitaxie — *metal organic vapor-phase epitaxy* (MOVPE) — wurde in ihrem Ursprung von Manasevit entwickelt[2] und gilt heutzutage als Standardverfahren zum Wachstum von Bauelementen, deren Anwendung von Laserdioden und LEDs bis hin zu Solarzellen reicht[88]. Die treibende Reaktion jedes MOVPE Wachstumsprozesses lässt sich als



darstellen. Dabei repräsentieren R und R' jeweils ein organisches Radikal (wie z.B. Methyl, Ethyl oder höhere Kohlenstoffradikale), Wasserstoff oder ihre Kombinationen. M ist ein Gruppe-II- oder Gruppe-III-Metall und E ein Element der fünften oder sechsten Hauptgruppe. Der Index n gibt an, ob das entsprechende Wachstum ein II-VI Wachstum ($n=2$) oder ein III-V Wachstum ($n=3$) ist. Mit (v) und (s) wird schlussendlich die Gasphase, beziehungsweise die feste Phase bezeichnet.

Galliumnitrid wird in der Regel aus Galliumtrimethyl und Ammoniak hergestellt. Die Reaktion hat die Gestalt



Im Experiment wird Galliumtrimethyl und Ammoniak über separate Rohrleitungen in die Reaktorkammer gebracht. Der Fluss kann dabei über die Zuleitungen gesteuert werden. In der Reaktorkammer befindet sich ein Substrat als Wachstumskeim, an welchem die obige Reaktion stattfindet. Die Reaktorkammer ist dabei beheizt. Bei Galliumnitrid beträgt diese Wachstumstemperatur ca. 1275 K[9].

Die Dotierung mit Magnesium geschieht über die Zuschaltung einer Magnesiumquelle, wie zum Beispiel Dicyclopentadienylmagnesium $Mg(C_5H_5)_2$ [89]. Experimentell können dabei Dotierungen von 10^{18} - 10^{20} cm^{-3} erreicht werden[90, 9]. Eine neuere Arbeit erreicht sogar qualitativ hochwertige Proben mit einer Magnesiumkonzentration von $2 \cdot 10^{20}$ cm^{-3} unter Verwendung der Metall-Modulations-Epitaxie (MME)[91]. Den Proben ist gemein, dass sie anschliessend thermisch aktiviert werden müssen. Hierbei beobachtet man, dass die benötigte Aktivierungstemperatur mit steigender Magnesiumkonzentration sinkt. Sind für moderat dotierte Proben mit einer Magnesiumkonzentration von 10^{18} - 10^{19} cm^{-3} Temperaturen bis zu 725° C notwendig, erreicht man die thermische Aktivierung von hoch dotierten Proben schon bei 500° C[91]. Dies deutet auf einen sich ändernden Passivierungsmechanismus oder eine sich ändernde Wasserstoffkinetik hin und kann eine Ursache des im Experiment[9] beobachteten verminderten Wasserstoffeinbaus sein.

Um die theoretische Modellierung dieses Prozesses so einfach wie möglich zu halten, werden die chemischen Reservoirs der Prozessprodukte über sogenannte chemische Potentiale repräsentiert. Diese werden auf möglichst einfache Systeme referenziert.

Das chemische Potential des Gallium beispielsweise referenziert man auf die Galliumbulkphase anstatt auf das Galliumtrimethyl um die Gesamtheit der an der Reaktion beteiligten chemischen Potentiale auf das notwendige Minimum zu beschränken. Die chemischen Potentiale sind dann in der theoretischen Modellierung frei wählbare Parameter, deren Wertebereich im Zusammenspiel mit dem Experiment sinnig abzustecken ist.

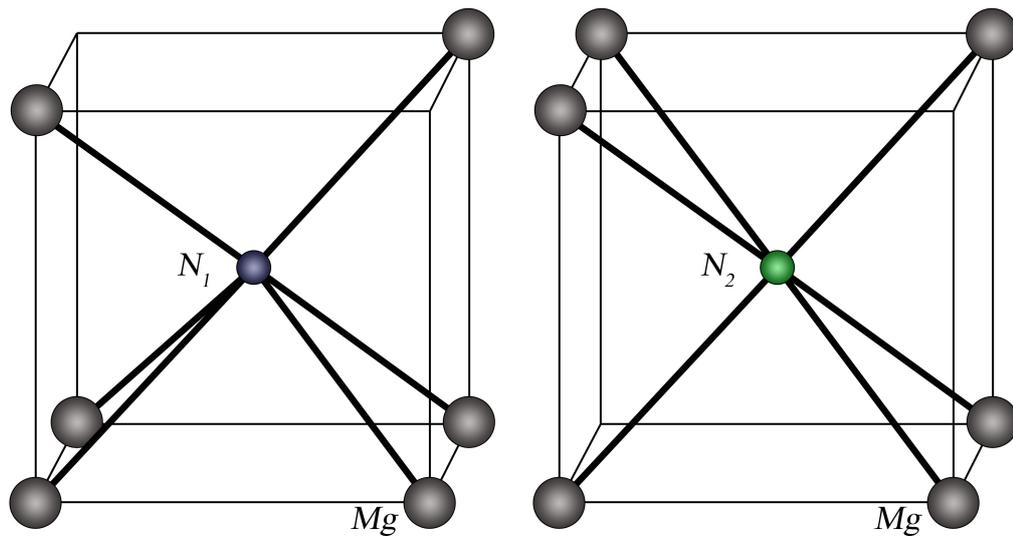
Die vorgestellte Theorie betrachtet das Kristallwachstum im thermischen Gleichgewicht. Es ist jedoch gezeigt worden, dass durch kinetische Prozesse an der Oberfläche die Defektkonzentration dort signifikant von der im thermischen Gleichgewicht abweichen kann[92]. Durch Überwachsen gelangen dann diese Defekte in den Bulkbereich, so dass die Ergebnisse hier grundlegend kritisch in diesem Zusammenhang zu diskutieren sind. Auch die Behandlung der Referenzsysteme bei 0 K induziert systematische Fehler in die berechneten Konzentrationen und die auf den chemischen Potentialen basierenden Phasendiagramme. Abschätzungen zu Folge beläuft sich die Ungenauigkeit im Vergleich zum Übergang zu endlichen Temperaturen auf einige Zehntel eV. So besitzen Magnesiumnitrid und Magnesium in der 0 K Beschreibung bei dem zur Dotierung notwendigen chemischen Magnesiumpotential schon stabile Bulkphasen, die aber experimentell nicht in Erscheinung treten. Magnesiumnitrid zersetzt sich bei den Wachstumstemperaturen und reines Magnesium liegt in einer flüssigen Phase vor. Auch hier sind folglich die Phasengrenzen kritisch zu hinterfragen und zu diskutieren.

7.3. Den Inversionsdomänengrenzen auf der Spur

Im Folgenden wird zunächst Magnesiumnitrid in den Fokus der Untersuchung gerückt. Leichtmetallnitride, zu denen auch Mg_3N_2 gehört, werden im Hinblick ihrer Eigenschaften als Wasserstoffspeichermedien in der aktuellen Forschung diskutiert[93, 94, 95]. Typischerweise bedingt die reversible Wasserstoffspeicherung in diesen Materialien eine Kristallphasentransformation, welche in den Inversionsdomänengrenzen nicht beobachtet wird. Das wirft die Frage auf, ob es andere Mechanismen gibt, welche den Einbau von Wasserstoff in die Inversionsdomänengrenzen begünstigen. Da jedoch schon das Bulkmaterial in der Literatur kaum behandelt wird, fokussieren wir uns an dieser Stelle auf Mg_3N_2 in der thermodynamisch stabilen Antixbyitstruktur. Die Diskussion der wasserstoffbeinhaltenden Defekte und nativer Punktdefekte in diesem Bulkmaterial geschieht hier nun erstmals auf dem Niveau der Dichtefunktionaltheorie und erlaubt die Vorhersage chemischer Trends für die Inversionsdomänengrenzen.

7.3.1. Magnesiumnitrid in der Antixbyitstruktur

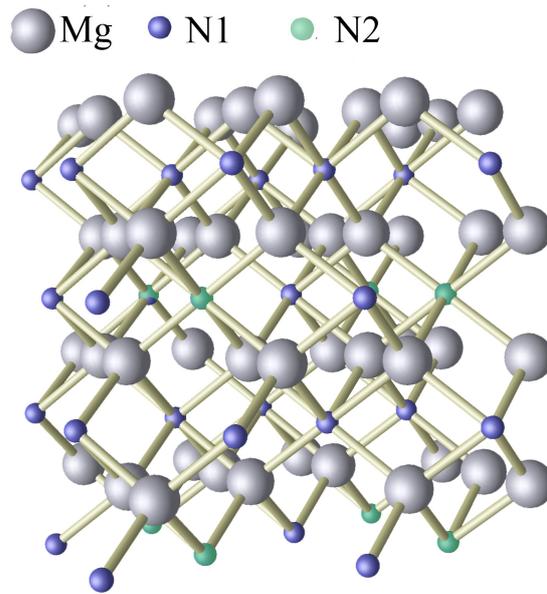
Magnesiumnitrid kristallisiert als Mg_3N_2 in der Antixbyitstruktur[96, 97], welche im Folgenden untersucht wird. Nach Mokhtari und Akbarzadeh kann die Antixbyitstruktur über die Kalziumfluoridstruktur konstruiert werden[98]. Kalziumfluorid

**Abb. 7.2:**

In der idealen Antibixbyitstruktur unterscheiden sich die beiden inäquivalenten Stickstofftypen durch die Lage der Magnesiumstrukturvakanz in dem umgebenen Magnesiumkubus. Liegen die Strukturvakanz auf einer Flächendiagonalen handelt es sich um den Stickstofftyp 1 (N_1), liegen sie auf der Raumdiagonalen handelt es sich um den Typ 2 (N_2).

(CaF_2) kristallisiert in einer kubischen Struktur. Dabei bilden die Kalziumkationen ein *fcc*-Gitter, dessen Tetraederlücken von den Fluoranionen gefüllt werden. Ein künstliches Mg_2N vertauscht nun die Rolle von Anionen und Kationen und resultiert in einer *Anti*-Kalziumfluoridstruktur. Die Stickstoffatome bilden dann ein *fcc*-Gitter dessen acht Tetraederlücken von Magnesiumatomen besetzt werden.

Um nun die richtige Stöchiometrie des Magnesiumnitrid (Mg_3N_2) zu gewährleisten, müssen zwei Magnesiumatome so entfernt werden, dass jedem Stickstoffatom sechs Nachbarn verbleiben. Diese Strukturvakanz richtig einzuführen ist nicht trivial und gelingt in der kubischen Zelle erst in einer $2 \times 2 \times 2$ Superzelle mit einer Gesamtzahl von 80 Atomen. Die Koordinaten dieser Strukturvakanz sind im Anhang A zu finden. Diese Zelle lässt sich symmetriebedingt in eine primitive *bcc* Zelle mit 40 Atomen reduzieren. Interessant hierbei ist, dass die Stickstoffatome nicht mehr ununterscheidbar sind. Je nachdem ob die von den acht Magnesiumatomen in der *Anti*- CaF_2 -Struktur eingeführten zwei Strukturvakanz auf einer Flächendiagonalen oder auf einer Raumdiagonalen liegen, spricht man nun zur Unterscheidung von den Stickstofftypen N_1 und N_2 (vgl. hierzu auch Abbildung 7.2). Dabei ist der Stickstoff vom Typ 1 dreimal häufiger anzutreffen, als der des Typs 2. Trotz dieser inäquivalenten Stickstoffatome verhalten sich die Magnesiumatome alle äquivalent zueinander. Ein jedes bindet zu drei Stickstoffatomen des Typs 1 und zu einem Stickstoffatom des Typs 2. Diese ideale Antibixbyitstruktur, dargestellt in Abbildung 7.3, stellt die Initialisierungsstruktur für die Bestimmung der Bulkeigenschaften dar. In den

**Abb. 7.3:**

Visualisierung der idealen Antibixbyitstruktur von Mg_3N_2 . Die grauen Atome zeigen die Magnesiumpositionen im Gitter an. Die kleineren kolorierten Atome zeigen die Positionen der Stickstoffatome vom Typ 1 (blau) und vom Typ 2 (grün).

folgenden Rechnungen darf diese Struktur frei relaxieren.

7.3.2. Bestimmung der Konvergenzparameter

Für die Berechnung der Bulkigenschaften ist die Multiskalenbibliothek S/PHI/nX[18] verwendet worden. Zum Zeitpunkt der Berechnungen waren mit diesem Programmpaket DFT-basierende Berechnungen lediglich mittels normerhaltenden Pseudopotentialen möglich. Zur Modellierung sind im Rahmen dieser Arbeit entsprechende Potentiale mit dem Programm FHI98PP[99] generiert³ worden. Die Modellierung des Austausch- und Korrelationsfunktional geschieht mittels des GGA-PBE Funktional. Ein Energiecutoff von 680 eV (50 Ry) und ein \mathbf{k} -Punkt Gitter von $2 \times 2 \times 2$ (offcenter) genügen, um eine Energiekonvergenz innerhalb von 1 meV zu gewährleisten. Die in der Tabelle 7.1 angegebenen Werte der theoretischen Gitterkonstante und des Bulkmodulus sind innerhalb von 0.01 Å und 1 GPa auskonvergiert. Die Gitterkonstante, der Bulkmodulus, sowie der Parameter B'_0 sind durch Fit der Energiekurve an die Murnaghansche Zustandsgleichung[85] bestimmt worden und zeigen eine sehr gute Übereinstimmung zu anderen theoretischen Rechnungen[98] und auch zu experimentellen Befunden[96, 38, 37]. Die berechnete

³ Für die Potentiale sind als cut-off Radien ($r_s^{\text{cut}}, r_p^{\text{cut}}, r_d^{\text{cut}}, r_f^{\text{cut}}$) in atomaren Einheiten verwendet worden:

N $2s^2p^3d^0$ (1.5,1.5,1.5) Troullier-Martin Type,

Mg $3s^3p^0d^0$ (1.2,1.6,1.5) Haman-Type,

H $1s^12p^03d^04f^0$ (1.276,1.276,0.350,1.276) Troullier-Martin Type.

	diese Arbeit	FP-LAPW [98]	Exp. [96, 38, 37]
a_{lat} [Å]	9.977	10.037	9.9528
B_0 [GPa]	108	110	—
B'_0	3.76	4.02	—
Bandlücke [eV]	1.59 (Γ - Γ)	1.56 (Γ - Γ)	3.15, 2.8(direkt) 2.85(indirekt)
Distanz [Å]			
Mg—N1 (2x)	2.09	2.10	2.084
Mg—N1 (2x)	2.17	2.18	2.160
Mg—N1 (2x)	2.19	2.20	2.179
Mg—N2 (6x)	2.15	2.16	2.145
Int. Koordinaten			
v	0.969	0.969	0.9784
x	0.389	0.389	0.3890
y	0.153	0.153	0.1520
z	0.382	0.383	0.3823

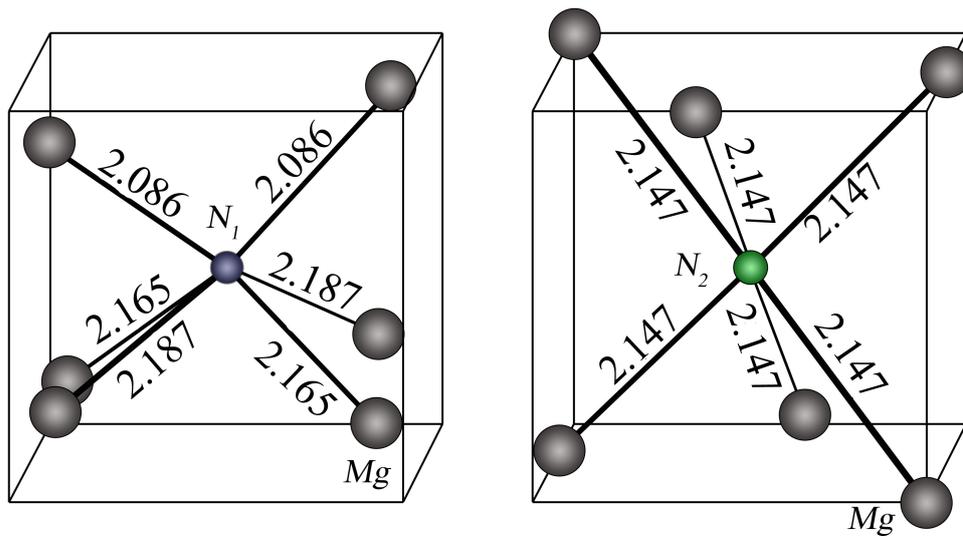
Tab. 7.1:

Vergleich der strukturellen, elastischen und elektronischen PBE-Parameter von Magnesiumnitrid mit verfügbaren theoretischen und experimentellen Daten der Literatur.

Bandlücke zeigt die typische Unterschätzung durch die Verwendung des PBE Funktionals.

Die Relaxation verzerrt die in Abbildung 7.2 dargestellten Kuben dergestalt, dass der Stickstoff des Typs 2 sechs identisch lange Bindungen zu seinen nächsten Magnesiumnachbarn ausprägt und diese Unterstruktur die hohe D_{3d} Symmetrie aufweist. Die Stickstoffe des Typs 1 finden sich in der geringeren C_2 Symmetrie wieder. Sie prägen 3 x 2 identische Bindungen aus (vgl. Abbildung 7.4). Dieses Verhalten wird durch vier interne Parameter beschrieben, die ebenfalls in sehr guter Übereinstimmung zu den experimentellen Befunden und anderen theoretischen Arbeiten sind.

Auffällig ist die Dehnung der Bindung im Vergleich zur Summe der kovalenten Radien (2.05 Å[?]). Diese lässt sich in einem Zwei-Zentren Bild erklären: Da die Koordinationszahl doppelt so groß ist, wie es die Anzahl der Valenzelektronen vermuten lässt, sind die bindenden Orbitale mit jeweils einem einzigen Elektron besetzt. Dies resultiert in einer schwächeren chemischen Bindung und hat den längeren Bindungsabstand im Vergleich zu vollbesetzten Bindungsorbitalen zur

**Abb. 7.4:**

Bindungslängen (in Å) der relaxierten Koordinationspolyeder im Magnesiumnitrid. Die Kuben geben die Positionen der Magnesiumatome in der idealisierten Antixbyitstruktur wieder. Grundlage der Daten bilden die PBE-Geometrien.

Folge.

Eine weitere Eigenart dieser Struktur sind die verzerrten Oktaederlücken, wie sie in Abbildung 7.5 dargestellt sind. Diese Lücken, die sich zwischen zwei Stickstoffatomen des Typ 2 befinden, stellen geeignete Positionen dar, um kleinere Atome (wie z.B. Wasserstoff) als Störstellen einzufangen und werden in der Bestimmung möglicher Punktdefekte eine Rolle spielen.

	Defektabstand [Å]		Defektabstand [Å]
Mg—V _{N₁} (2x)	2.15 (+2.9%)	V _{Mg} —N ₁ (2x)	2.28 (+9.1%)
Mg—V _{N₁} (2x)	2.29 (+5.8%)	V _{Mg} —N ₁ (2x)	2.35 (+8.6%)
Mg—V _{N₁} (2x)	2.32 (+6.2%)	V _{Mg} —N ₁ (2x)	2.38 (+9.0%)
Mg—V _{N₂} (6x)	2.27 (+5.9%)	V _{Mg} —N ₂ (6x)	2.33 (+8.4%)

Tab. 7.2:

Mg—V_N Distanzen der N1/N2 Vakanz im einfach positiv geladenen Ladungszustand und V_{Mg}—N Abstände der Mg Vakanz im 2-fach negativ geladenen Zustand in Mg₃N₂. Die prozentualen Angaben beziehen sich auf die Abstände im Bulk. Die geklammerten Ausdrücke (2x) bzw. (6x) geben die Anzahl der äquivalenten Bindungen an. Die Abstände basieren auf den PBE-Geometrien.

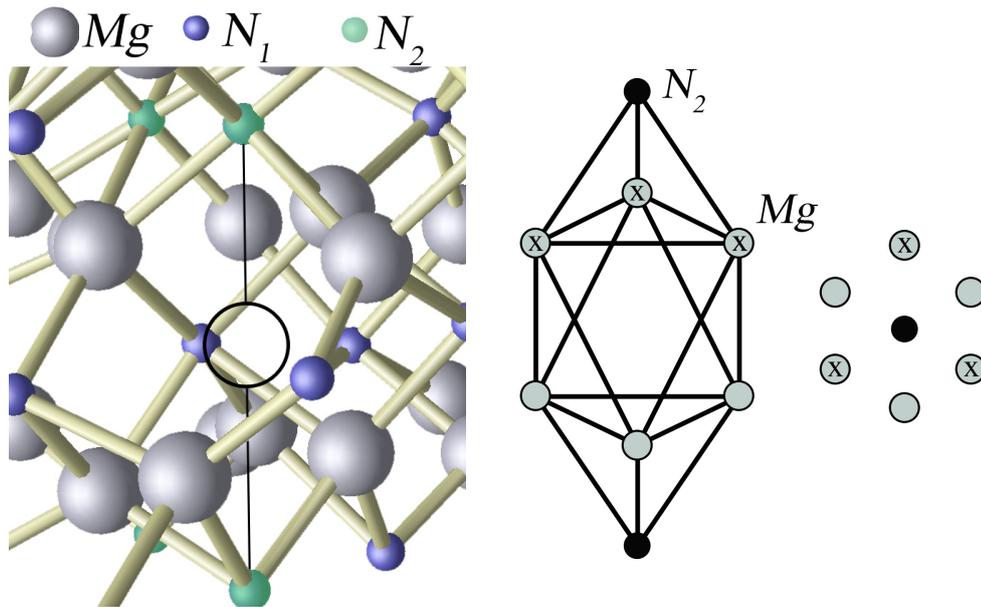


Abb. 7.5:

Die verzerrte Oktaederlücke in Mg_3N_2 . Links: Die relaxierte Struktur mit der symbolisierten Lücke in der $\text{N}_2\text{—N}_2$ Verbindungslinie. Rechts: Idealisierte Darstellung des Oktaederkäfigs.

7.3.3. Intrinsische Punktdefekte

Zu den intrinsischen Punktdefekten in Mg_3N_2 gehören Fehlstellen und Zwischen-gitterbesetzungen der Elemente Magnesium und Stickstoff. Wir diskutieren im folgenden die Magnesiumvakanz V_{Mg} , die Besetzung der Strukturvakanz mit Magnesium I_{Mg} , die Stickstoffvakanz vom Typ 1 V_{N_1} und Typ 2 V_{N_2} , sowie die sogenannten *Splitinterstitials* des Stickstoffs $I_{\text{N}_{1/2}}$. Bei letzterem handelt es sich um die Ersetzung eines Stickstoffatoms durch ein Stickstoffmolekül. Abbildung 7.6 zeigt diese intrinsischen Defekte schematisch.

Diese Defekte sind in der 40 Atome enthaltenden primitiven Zelle als auch in der 80 Atome enthaltenden kubischen Superzelle modelliert und berechnet worden. Die Bildungsenergie ergibt sich hierbei gemäß Gleichung (5.50) plus der Superzellenkorrektur[67]. Für die chemischen Potentiale sind magnesiumreiche Bedingungen angenommen worden. Die konkurrierenden Systemphasen sind Mg_{bulk} und Mg_3N_2 .

$$\mu(\text{Mg}) = E_{\text{tot}}[\text{Mg}, \text{bulk}] \quad (7.3)$$

$$2\mu(\text{N}) = E_{\text{tot}}[\text{Mg}_3\text{N}_2, \text{bulk}] - 3\mu(\text{Mg}) \quad (7.4)$$

Für die relevanten Defekte stimmen die ladungskorrigierten Bildungsenergien innerhalb von 0.1 eV in beiden Superzellen überein. Vereinzelt sind die Abweichungen

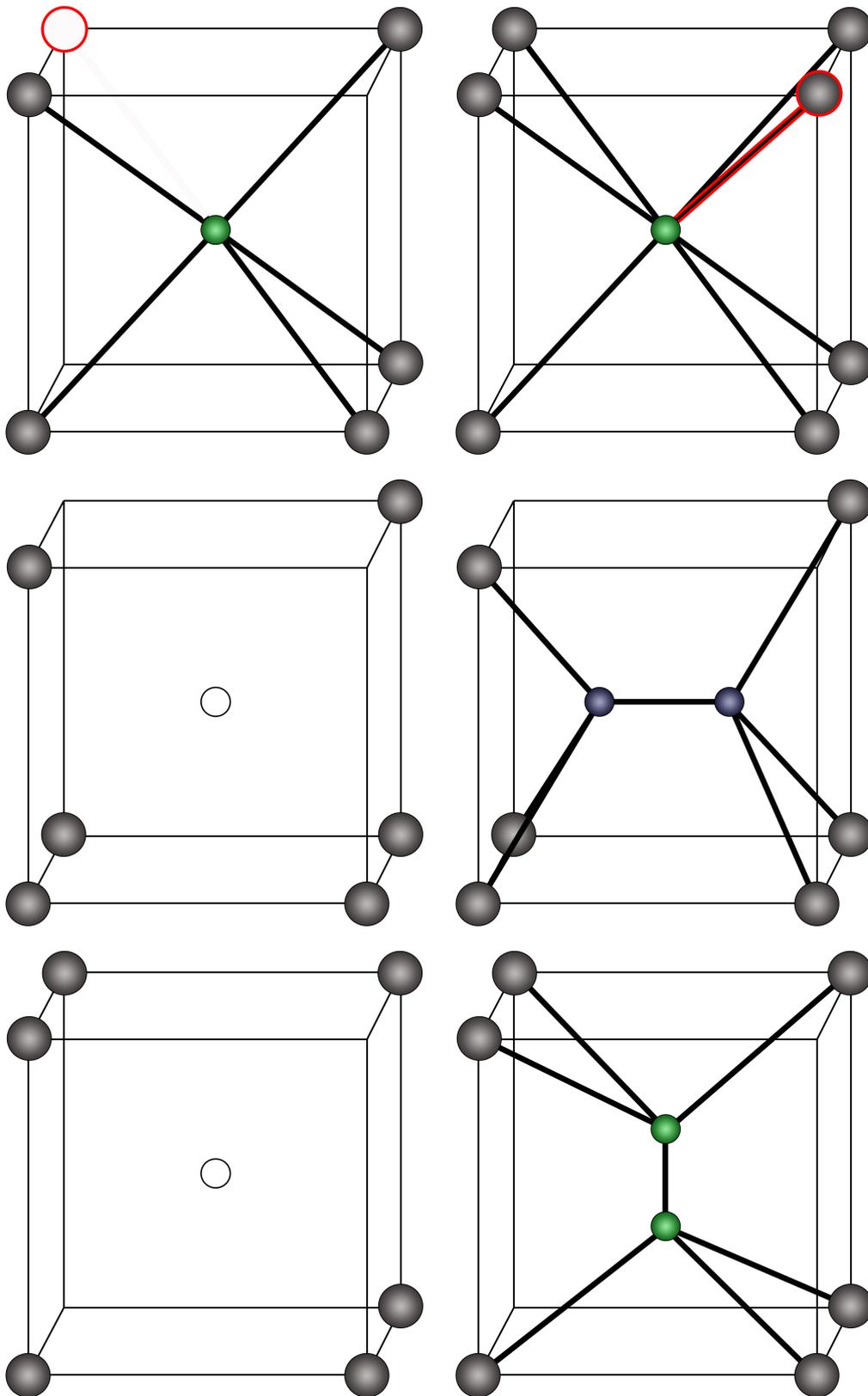


Abb. 7.6:

Visualisierung der intrinsischen Punktdefekte in Mg₃N₂. Gezeigt sind die Magnesiumvakanz und der Magnesiumstörstelle (oben), sowie die Stickstoffvakanzanzen und Stickstoffstörstellen des Typ I (mitte) und des Typ II (unten).

I _{N1/N2} Abstand			
	[Å]		
Ladungszustand	-1	0	+1
Mg—N1 (2x)	1.95	1.99	2.04
Mg—N1 (2x)	2.05	2.08	2.12
Mg—N1 (2x)	2.11	2.21	2.32
Mg—N1 (2x)	2.17	2.22	2.28
N1—N1	1.576	1.409	1.301
Mg—N2 (6x)	2.00	2.07	2.15
N2—N2	1.58	1.41	1.30

I _{Mg} Abstand			
	[Å]		
Ladungszustand	0	+1	+2
Mg—N (3x)	2.09	2.09	2.09
Mg—N	2.17	2.16	2.16

Tab. 7.3:

N—Mg Abstände des Stickstoffsplinterstitials und der Magnesiumstörstelle für unterschiedliche Ladungszustände in Mg₃N₂ berechnet in der PBE Näherung. (*nx*) gibt die Anzahl äquivalenter Bindungen an.

etwas größer. Die Diskrepanz übersteigt aber nicht 0.4 eV.

Die charakteristischen Bindungslängen des Defektes zum nächsten Nachbarn sind in Tabelle 7.2 für die Vakanzen und in Tabelle 7.3 für die Störstellen wiedergegeben. Für die Vakanzen beobachten wir eine Relaxation der umgebenden Atome weg vom Defektzentrum. Die Abstände erhöhen sich um 0.06—0.19 Å (3—9%) im Vergleich zum defektfreien System. Grund hierfür ist die Abnahme der Koordinationszahl und daher eine Zunahme der Elektronen in jeder Zwei-Zentren Bindung. Der kovalente Anteil der Bindungen zu den benachbarten Atomen wird gestärkt. Auch der ionische Bindungsanteil wird durch die kürzeren Abstände zu diesen Nachbarn verstärkt.

Die Stickstoffstörstellen ihrerseits bilden N₂-Komplexe. Unabhängig vom Stickstofftyp in der atomaren Struktur variiert die N—N Bindungslänge zwischen 1.57 Å im negativ geladenen Fall und 1.30 Å im positiv geladenen Fall. Typische vergleichbare Komplexe weisen eine deutlich kürzere Bindungslänge auf, z.B. Hydrazin (H₂N—NH₂: 1.45 Å) oder Diazin (HN=NH: 1.24 Å). Energetisch ist die Störstelle des Typs 2 jener des Typs 1 bevorzugt. Dies lässt sich mit der nahezu störungsfreien Integration der Störstelle des Typs 2 in die umgebene Struktur begründen. Die Typ 1 Störstelle hingegen findet eine spezielle Umgebung vor, die nur durch entsprechende Verzerrungen besetzt werden kann.

		Bildungsenergie [eV]			
		<i>bcc</i> -Zelle 40 Atome		<i>sc</i> -Zelle 80 Atome	
	Defekt	DFT	+ E_{align}	DFT	+ E_{align}
V_{Mg}	-2	4.61	6.11	4.97	6.33
	-1	4.63	5.16	4.85	5.27
	0	4.85	4.85	4.88	4.88
I_{Mg}	0	4.18	4.18	3.84	3.84
	+1	1.13	1.48	1.26	1.63
	+2	-1.74	-0.33	-1.35	-0.07
V_{N_1}	-1	5.17	5.60	4.99	5.24
	0	2.41	2.41	2.49	2.49
	+1	-0.08	0.08	-0.03	0.14
	+2	-0.75	0.16	-0.60	0.20
	+3	-1.49	0.70	-1.17	0.70
V_{N_2}	-1	4.88	5.26	4.96	5.24
	0	2.46	2.46	2.45	2.45
	+1	-0.09	0.09	-0.04	0.16
	+2	-0.89	0.03	-0.72	0.23
	+3	-1.89	0.33	-1.52	0.69
I_{N_1}	-1	5.76	6.09	5.82	6.01
	0	4.39	4.39	4.44	4.44
	+1	3.29	3.78	3.41	3.68
I_{N_2}	-1	5.04	5.35	5.08	5.31
	0	3.91	3.91	3.96	3.96
	+1	2.96	3.36	3.10	3.39

Tab. 7.4:

Berechnete Bildungsenergien stabiler intrinsischer Defekte in Mg_3N_2 gemäß PBE-DFT Superzellenrechnung sowie nach Anwendung der Superzellenkorrektur für geladenen Defekte (E_{align}).

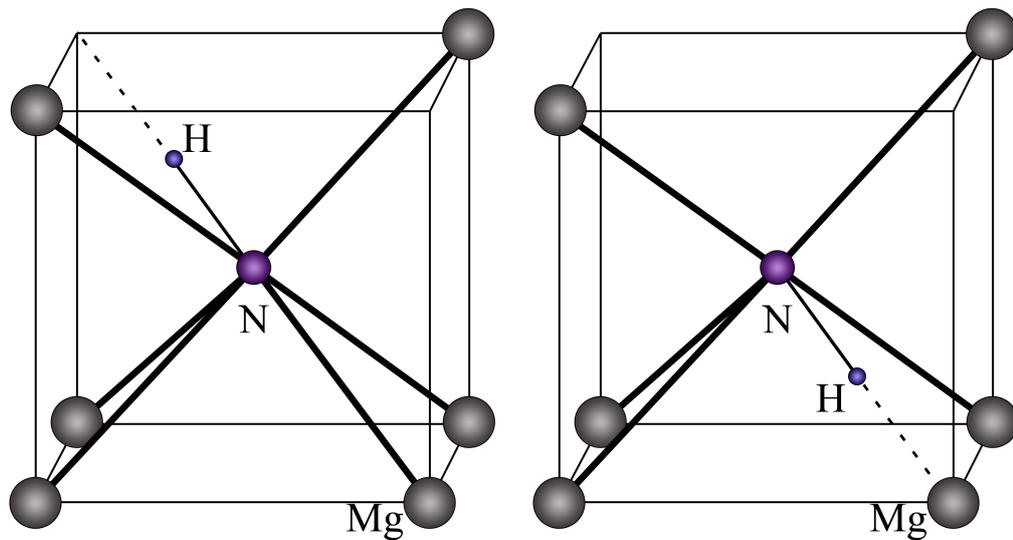


Abb. 7.7:

Darstellung der denkbaren Wasserstoffstörstellenkonfigurationen in Magnesiumnitrid. Vergleichbar zum GaN gibt es im Mg₃N₂ für den Wasserstoff die Möglichkeit an den Stickstoff zu binden. Dabei kann er eine sogenannte *anti-bindende* Position (links) besetzen, oder eine *bindungszentrierte* Position (rechts).

Die Magnesiumstörstelle bindet an der Position der Strukturvakanz zu den umgebenden vier Stickstoffatomen. Dabei entsprechen drei dieser Bindungen der typischen kurzen Mg—N₁-Bindung. Die vierte Bindung reagiert geringfügig auf den Ladungszustand und variiert zwischen 2.17 Å im ladungsfreien Zustand und 2.16 Å im zweifach positiv geladenen Zustand. In diesem Zustand liegt das Magnesiumatom vergleichbar zu den anderen Magnesiumatomen vor und fügt sich am besten in das System ein.

7.3.4. Wasserstoffdefekte und Defektkomplexe

Für den Wasserstoff gibt es verschiedene Positionen zur Adsorption. Die Strukturvakanz [V_{Mg}—*n*H] sowie die Oktaederlücke H_{okt} bieten genug freien Raum um den Wasserstoff aufzunehmen. Weiter betrachten wir die sogenannte *anti-bindende* Position H_{N,ab} und die *bindungszentrierte* Position H_{N,bc} (vgl. Abbildung 7.7) zum Stickstoff. Diese Defekte sind im folgenden für unterschiedliche Ladungszustände modelliert worden.

Das chemisches Wasserstoffpotential ist auf das halbe Wasserstoffmolekül für T = 0 K referenziert:

$$2\mu(\text{H}) = E_{\text{tot}}[\text{H}_2, T = 0 \text{ K}]. \quad (7.5)$$

Im positiven Ladungszustand (H⁺) ist die anti-bindende Position nahe des Stick-

	Bildungsenergie [eV]		
	Defekt	bcc -Zelle 40 Atome	
		DFT	+ E_{align}
$H_{N_1,ab}$	-1	4.55	4.86
	0	1.62	1.62
	+1	-1.09	-0.74
$H_{N_2,ab}$	-1	4.49	4.77
	0	1.54	1.54
	+1	-1.18	-0.81
H_{okt}	-1	2.42	2.78
	0	2.05	2.05
	+1	1.49	1.84
$H_{2,okt}$	0	1.44	1.44

Tab. 7.5:

Bildungsenergien der wasserstoffbeinhaltenen Defekte in Mg_3N_2 . Angegeben sind die Werte als direktes Resultat der PBE-DFT Superzellenrechnung und nach Addition der Superzellenkorrektur für geladenen Defekte (E_{align}).

stoffatoms am stabilsten. Der Unterschied in der Bildungsenergie bezüglich des Stickstofftyps beträgt hier 0.1 eV zugunsten des Typ 2. Der Bindungsabstand des Wasserstoffes zum Stickstoff ist charakteristisch für eine kovalente Bindung, wie sie beispielsweise im Hydrazin ($d_{N-H} = 1.03 \text{ \AA}$) auftaucht. Das Wasserstoffion in der bindungszentrierten Position relaxiert ohne Barriere in die anti-bindende Position und ist damit thermodynamisch nicht stabil. Das negativ geladene Wasserstoffion in der Oktaederlücke bildet den energetisch stabilsten negativ geladenen Defekt. Die Abstände des Wasserstoffatoms zu den benachbarten Magnesiumatomen variieren zwischen 1.96 \AA und 2.52 \AA . Für eine kovalente Mg—H Bindung sind diese Abstände zu lang, so dass hier rein ionische Aspekte den Wasserstoff auf seiner Position halten. Das Wasserstoffion in der Strukturvakanz relaxiert barrierefrei in die Oktaederlücke, so dass dieser Defekt thermodynamisch ebenfalls nicht stabil ist. Die berechneten Bildungsenergien sind in Tabelle 7.5 wiedergegeben. Die so identifizierten stabilen Wasserstoffdefekte stimmen gut mit bekannten Wasserstoffdefekten in anderen Nitriden überein[5, 6, 7, 8].

In Kombination mit den intrinsischen Defekten lassen sich Defektkomplexe bilden, die den eben diskutierten Wasserstoffdefekten recht ähnlich sind. Die Stickstoffvakanz bietet eine vergleichbare Umgebung zur Oktaederlücke: Eine Leerstelle umge-

[V _N —H]-Komplex		[V _{Mg} —H]-Komplex			
	Distanz [Å]	N-H Distanz [Å]			
Mg—H(2x)	2.10	[V _{Mg} -H]	1.04		
Mg—H(2x)	2.46	[V _{Mg} -2H]	1.03	1.03	
Mg—H(2x)	2.48	[V _{Mg} -3H]	1.04	1.04	1.04
Mg—H(6x)	2.36	[V _{Mg} -4H]	1.03	1.03	1.03 1.03

Tab. 7.6:

Berechnete Abstände der Wasserstoffstörstelle in Mg₃N₂ zum nächstgelegenen Nachbaratom.

ben von sechs Magnesiumatomen, welche von einem negativ geladenen Wasserstoffion besetzt werden kann. Die Magnesiumvakanz ihrerseits lässt vier Stickstoffatome als potentielle Bindungspartner für den Wasserstoff zurück. Diese Adsorptionszentren sind auf ihre energetischen und geometrischen Eigenschaften untersucht worden. Die berechneten Abstände nächster Nachbarn gibt Tabelle 7.6 wieder. Während der Wasserstoff in der Stickstoffvakanz an der Stickstoffposition verharret, bindet er in der Magnesiumvakanz stark an die umgebenen Stickstoffatome und formt [N—H]-Komplexe. Definiert man die Bindungsenergie gemäß

$$E^b = E^f[V_{\text{Mg}}^{-2}] + nE^f[\text{H}_{\text{N}_2}^+] - E^f[[V_{\text{Mg}}+n\text{H}]^{n-2}], \quad (7.6)$$

so ergibt sich für das erste Wasserstoffatom eine Bindungsenergie von 2.35 eV. Das zweite Wasserstoffatom bindet mit 1.33 eV, das dritte mit 0.82 eV und das vierte noch mit 0.02 eV. Der [V_N+H]-Defekt bietet dagegen eine Bindungsenergie

$$E^b = E^f[V_{\text{N}}^+] + E^f[\text{H}_{\text{N}_2}^+] - E^f[[V_{\text{N}}+\text{H}]^{2+}] \quad (7.7)$$

von 0.33 eV. Aufgrund der hohen Bindungsenergie im [V_{Mg}+H]⁻¹-Defektkomplex ist es zu erwarten, dass die Bildung von Magnesiumvakanz in Mg₃N₂ in einer wasserstoffhaltigen Atmosphäre erheblich gefördert wird.

7.3.5. Mg₃N₂ Einschlüsse in GaN

Um nun aus den im vorherigen Kapitel dargestellten Ergebnissen Hinweise auf das Verhalten der Magnesiumnitrideinschlüsse in Galliumnitrid zu gewinnen, bedarf es zunächst einer Umreferenzierung der chemischen Potentiale. Diese waren auf die Referenzphasen (Mg_{bulk}, Mg₃N₂ und H₂ für T = 0 K) für pures Mg₃N₂ im magnesiumreichen Wachstum gesetzt worden. Da sich die Magnesiumnitrideinschlüsse unter GaN Wachstumsbedingungen bilden, sind die entsprechende Referenzsysteme für galliumreiches Wachstum

$$\mu(\text{Ga}) = E_{\text{tot}}[\text{Ga, bulk}] \quad (7.8)$$

	Bildungsenergie [eV]		
	Defekt	<i>bcc</i> -Zelle 40 Atome	
		DFT	+ E_{align}
[V _{Mg} + H]	-1	2.48	2.95
	0	2.72	2.72
	+1	2.95	3.11
[V _{Mg} + 2H]	0	0.81	0.81
	+1	0.73	0.90
	+2	1.05	2.08
[V _{Mg} + 3H]	+1	-1.04	-0.82
	+2	-0.87	0.19
	+3	-0.58	2.08
[V _{Mg} + 4H]	+1	-0.52	-0.21
	+2	-3.01	-1.65
	+3	-2.77	0.51
	+4	-2.44	2.52
[V _{N₁} + H]	+1	0.41	0.59
	+2	-1.96	-1.02
	+3	-1.89	0.40
[V _{N₂} + H]	+1	0.37	0.57
	+2	-2.06	-1.06
	+3	-1.95	0.48

Tab. 7.7:

Berechnete Bildungsenergie der Wasserstoffdefektkomplexe in Mg₃N₂ als unmittelbares Resultat der PBE-DFT Superzellenrechnung sowie nach Addition der Superzellenkorrektur für geladenen Defekte (E_{align}).

Wachstum ohne Wasserstoff		Wachstum in Wasserstoffatmosphäre	
Defekt	c [cm ⁻³]	Defekt	c [cm ⁻³]
V _N	1.6 · 10 ¹⁴	V _N	1.7 · 10 ¹⁴
V _{Mg}	8.8 · 10 ¹¹	V _{Mg}	7.8 · 10 ¹¹
I _N	4.9 · 10 ⁸	I _N	4.8 · 10 ⁸
I _{Mg}	4.2 · 10 ⁸	I _{Mg}	4.6 · 10 ⁸
		H _{N,ab}	1.7 · 10 ¹⁵
		V _N + H	7.5 · 10 ⁸
		H _{okt}	2.0 · 10 ¹⁴
		[V _{Mg} +H]	1.7 · 10 ¹⁵
		[V _{Mg} +2H]	1.8 · 10 ¹³
e ⁺	1.6 · 10 ¹⁴	e ⁺	1.6 · 10 ¹³
Fermi Energie	1.74 eV	Fermi Energie	1.73 eV

Tab. 7.8:

Konzentrationen der ausgewählten Punktdefekte in Mg₃N₂, sowie Lage des Fermi-niveaus beim Wachstum ohne (links) und mit (rechts) Wasserstoffumgebung. Die Wachstumsbedingungen sind im Text genauer spezifiziert ($T = 1275$ K, $p = 1$ bar).

eine realistischere Wahl für die Simulation. Das chemische Stickstoffpotential ist demnach in Gleichgewicht zu dem chemischen Potential von GaN

$$\mu(\text{N}) = E_{\text{tot}}[\text{GaN, bulk}] - \mu(\text{Ga}) \quad (7.9)$$

zu wählen und das chemische Magnesiumpotential entsprechend im Gleichgewicht mit dem Magnesiumnitrideinschluss

$$3\mu(\text{Mg}) = E_{\text{tot}}[\text{Mg}_3\text{N}_2, \text{bulk}] - 2\mu(\text{N}). \quad (7.10)$$

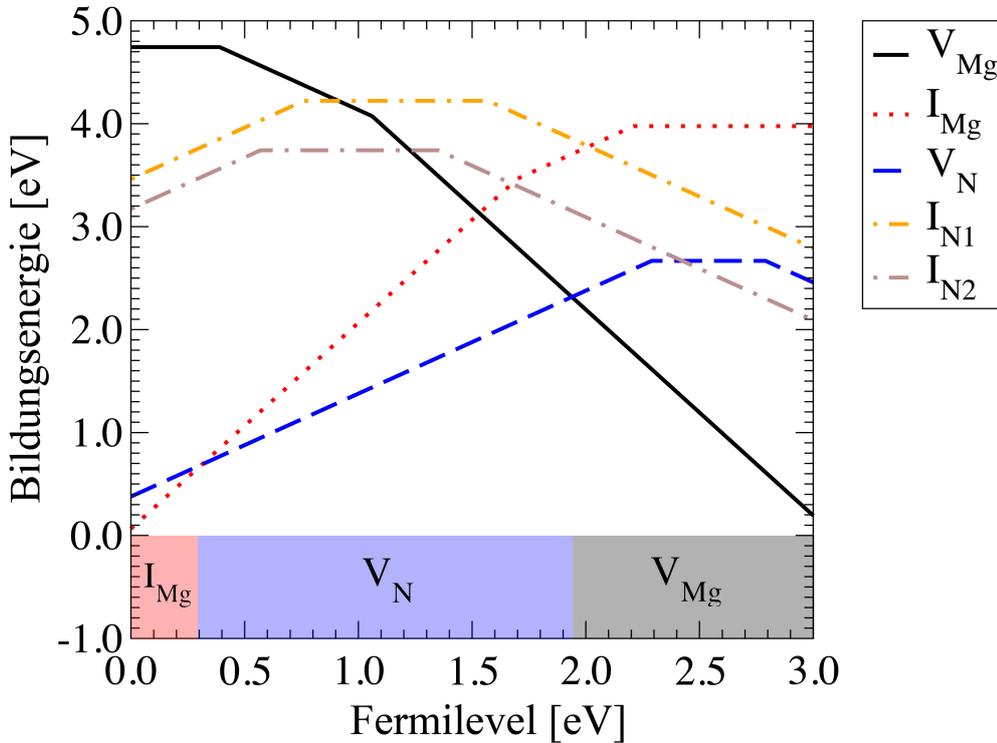
Das chemische Wasserstoffpotential wird dann im thermischen Gleichgewicht mit den Wachstumsbedingungen zu

$$2\mu(\text{H}) = E_{\text{tot}}[\text{H}_2, T = 0 \text{ K}] + k_b T \ln \left(\frac{pV_q}{k_b T} \right) - k_b T \ln \left(\frac{k_b T}{2\Theta} \right) \quad (7.11)$$

bestimmt[20]. V_q bezeichnet hierbei das Phasenraumvolumen

$$V_q = \left(\frac{2\pi\hbar^2}{m_H k_b T} \right)^{3/2}. \quad (7.12)$$

Die Rotationskonstante des Wasserstoff ist der experimentellen Literatur entnommen und beträgt 59.339 cm^{-1} [100]. Zur Simulation realistischer Wachstumsbedingungen wird eine Temperatur von 1275 K und ein Wasserstoffgasdruck von 1 bar angenommen. Diese Werte repräsentieren eine obere Schranke zu den normalerweise in der metallorganischen Gasphasenepitaxie vorherrschenden Bedingungen[9].

**Abb. 7.8:**

Bildungsenergien der intrinsischen Punktdefekte in Mg_3N_2 aufgetragen über der Position des Fermi-niveaus. Grundlage bilden die in PBE berechneten Gesamtenergien der Defekte in der bcc -Einheitszelle mit 40 Atomen. Die Superszellenkorrektur für geladene Defekte ist angewandt worden.

Die Bildungsenergien wurden nun für diese experimentellen Randbedingungen berechnet und in Abbildung 7.8 für die intrinsischen Defekte graphisch dargestellt. Gezeigt ist in der Auftragung die Bildungsenergie in Abhängigkeit der Lage des Fermi-niveaus. Der Defekt niedrigster Bildungsenergie ist dann der dominant vorherrschende Defekt in diesem Bereich. Der Abbildung ist zu entnehmen, dass über einen weiten Bereich die Vakanzbildung favorisiert wird. Für n -leitende Bedingungen bilden sich vermehrt Magnesiumvakanzen, für p -leitende Bedingungen spielen Stickstoffvakanzen eine bedeutende Rolle. Im Falle eines niedrigen Fermi-levels ($E_{Fermi} < 0.3$ eV) wird die Auffüllung der Strukturvakanzen mit Mg^{2+} relevant.

Wenn während des Wachstums Wasserstoff in signifikanten Mengen vorhanden ist, bildet der positiv geladene Wasserstoffdefekt am Stickstoff in der anti-bindenden Position den energetisch stabilsten Defekt für p -leitende Bedingungen (vgl. Abbildung 7.9). Für n -leitenden Bedingungen spielen die negativ geladenen Wasserstoffdefekte, sowie die pure Magnesiumvakanz eine wichtige Rolle. Betrachtet

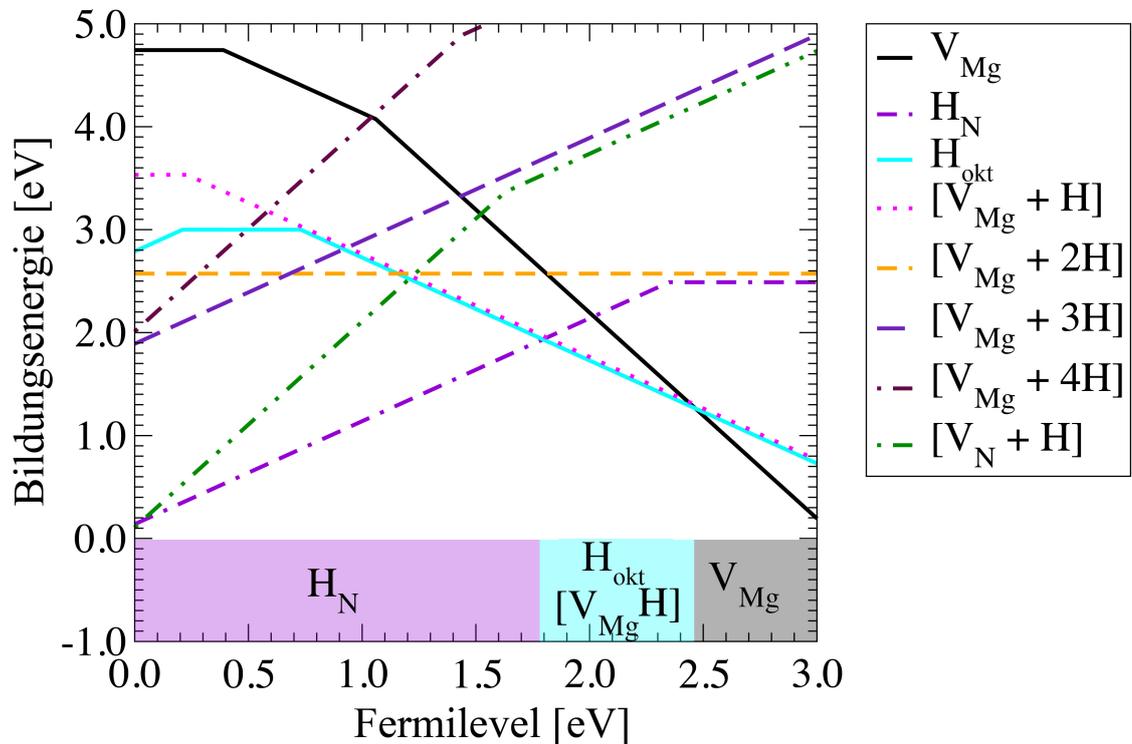
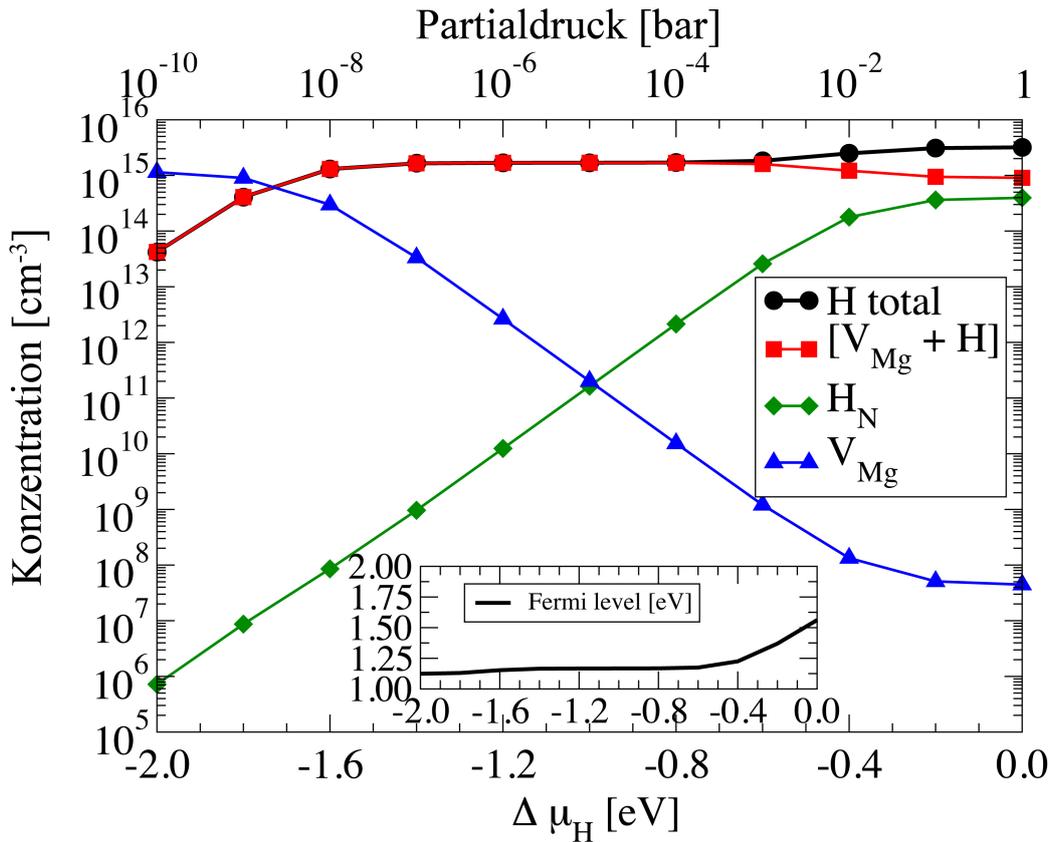


Abb. 7.9:

Bildungsenergien der wasserstoffbeinhaltenden Punktdefekte in Mg_3N_2 aufgetragen über der Position des Fermienergielevels. Zusätzlich ist noch die Bildungsenergie der reinen Magnesiumvakanz als stabilster Defekt für ein hohes Fermienergielevel eingezeichnet. Grundlage bilden die in PBE berechneten Gesamtenergien der Defekte in der bcc -Einheitszelle mit 40 Atomen. Die Superszellenkorrektur für geladene Defekte ist angewandt worden.

man das System im thermischen Gleichgewicht und berechnet die Konzentrationen wie in Abschnitt 5.5 diskutiert, ergeben sich die Zahlenwerte der Tabelle 7.8.

Im wasserstofffreien Wachstum ist die einfach positiv geladene Stickstoffvakanz mit $1,6 \cdot 10^{14} \text{ cm}^{-3}$ der Defekt größter Konzentration. Die Fermienergie liegt mit 1,74 eV knapp unterhalb des $V_{\text{N}}/V_{\text{Mg}}$ Schnittpunktes bei ca. 1,9 eV. Allerdings wird die Ladungsneutralität nicht von der zweifach negativ geladenen Stickstoffvakanz übernommen. Die Konzentration ist mit $8,8 \cdot 10^{11} \text{ cm}^{-3}$ hier um Größenordnungen zu klein. Es zeigt sich, dass die Ladungskompensation von freien Löchern bewerkstelligt wird. Dies ist möglich, da die auftretenden Konzentrationen mit der Größenordnung 10^{14} cm^{-3} deutlich geringer sind, als typische Konzentrationen in dotierten Materialien (typischerweise $10^{18} - 10^{21} \text{ cm}^{-3}$). Die Störstellen spielen generell eine untergeordnete Rolle.

**Abb. 7.10:**

Verlauf der Konzentrationen dominanter Defekte in Mg_3N_2 im Hinblick auf das chemische Potential des Wasserstoffs zur Simulation der thermischen Austreibung. Die Absenkung des chemischen Wasserstoffpotential von 0 eV auf -2 eV bedeuten eine Änderung des H_2 -Partialdruckes von 1 bar auf 10^{-10} bar bei einer konstanten Temperatur von 900 K.

Im Wachstum innerhalb einer vorhandenen wasserstoffhaltigen Atmosphäre ändert sich das Bild. Zwar weisen die intrinsischen Defekte keine signifikanten Änderungen in ihrer Konzentration auf, aber der Einbau von Wasserstoff begünstigt Magnesiumvakancen um den negativ geladenen $[\text{V}_{\text{Mg}} + \text{H}]$ -Defektkomplex zu bilden, so dass ein Anstieg der Vakancen um nahezu drei Größenordnungen zu verzeichnen ist. Der kompensierende Defekt bezüglich Ladung ist der positiv geladene Wasserstoff an der Stickstoff anti-bindenden Position. Der Schnittpunkt dieser beiden Defekte liegt nahe des im intrinsischen Fall bestimmten Fermilevels von 1.74 eV, so dass hier keine signifikante Änderung des Fermi-niveaus zu erwarten ist.

7.3.6. Thermische Austreibung

Die thermische Austreibung von Wasserstoff ist der Schlüsselprozess zur Aktivierung der Akzeptorniveaus im p -GaN. Die exakte theoretische Modellierung dieses Prozesses ist nicht trivial, da eine Vielzahl von Effekten die thermische

Austreibung begleiten. Eine Abschätzung kann gewonnen werden, indem man vereinfachte Annahmen macht. So wird im weiteren angenommen, dass sich mit Ausnahme des Wasserstoffes keine Atome durch den Kristall bewegen und somit die Magnesium- und Stickstoffdefekte in ihrer Position verharren und in ihrer Konzentration konstant bleiben. Die Austreibung kann dann durch Absenken des chemischen Potentials für Wasserstoff für eine feste Temperatur modelliert werden.

Im folgenden wird eine Temperatur von 900 K angenommen. Diese Temperatur ist charakteristisch für den so genannten *rapid thermal annealing* Prozess[9]. Weiterhin wird angenommen, dass sich die Wasserstoffatome frei durch den Kristall bewegen können, sobald sie sich aus ihren Defektzentren gelöst haben. Die sich so ergebende Wasserstoffkonzentration in Abhängigkeit des Wasserstoffpartialdruckes ist in Abbildung 7.10 wiedergegeben. Der Abbildung ist zu entnehmen, dass der Wasserstoff in einem Zweischnittverfahren ausgetrieben wird. Zunächst findet eine geringe Austreibung statt. Diese resultiert in der Entvölkerung der $H_{N,ab}$ und H_{okt} Defekte, sowie aus der Reduktion des $[V_{Mg} + 2H]$ Defektkomplexes zu dem $[V_{Mg}+H]$ Defektkomplex. Der einzelne Wasserstoff in der Magnesiumvakanz bleibt durch seine hohe Bindungsenergie von 2.35 eV stabil, bis der Partialdruck 10^{-8} bar unterschreitet.

Es sind folglich drastische Bedingungen von Nöten, um eine Austreibung von Wasserstoff aus Magnesiumnitrid zu realisieren.

7.3.7. Zusammenfassung

In diesem Kapitel sind die intrinsischen und wasserstoffinduzierten Punktdefekte in Magnesiumnitrid unter Anwendung der Dichtefunktionaltheorie identifiziert und modelliert worden. Basierend auf den Bildungsenergien wurden Defektkonzentrationen unter der Annahme galliumreicher Wachstumsbedingungen bestimmt, um eine Beziehung zu dem Wachstum von Galliumnitrid herzustellen. Im thermodynamischen Gleichgewicht stellt sich heraus, dass ein einkristallines Magnesiumnitrid ein nahezu defektfreier Isolator ist. Die Stickstoffvakanz ist mit einer Defektkonzentration von $2 \times 10^{14} \text{ cm}^{-3}$ der vornehmlich vorkommende Defekt. Unter Einfluss einer Wasserstoffumgebung wird die Bildung von stark gebundenen $[N-H]$ -Komplexen begünstigt. Dabei bildet der Wasserstoff eine starke Bindung zu vorhandenen Stickstoffatomen aus, oder belegt negativ geladene Magnesiumvakanzen. Die maximale Konzentration von Wasserstoff ist nichtsdestotrotz mit $4 \times 10^{15} \text{ cm}^{-3}$ als klein zu bezeichnen. Insbesondere ein Vergleich mit den typischen Konzentrationen des $[MgH]$ -Komplexes in GaN:Mg (10^{18} cm^{-3} und mehr) macht dies deutlich. Dies zusammen mit der Tatsache, dass das Magnesiumnitridvorkommen in den Inversionsdomänengrenzen in GaN nur einen äußerst geringen Volumenbruchteil ausmachen, lässt die Schlussfolgerung zu, dass eine signifikante Modulation der Wasserstoffverteilung in aktuellen Proben nicht durch die Magnesiumnitridein-

schlüsse hervorgerufen werden kann.

Die Modellierung einer Inversionsdomäne wäre im Zusammenhang der Untersuchung der verminderten Wasserstoffkonzentration sicherlich interessant, jedoch würde dies die simultane Beschreibung von einigen 1000 Atomen bedeuten. Für solche Rechnungen ist die Dichtefunktionaltheorie in ihrer Ebenen-Wellen-Beschreibung nicht ausgelegt. Typische Algorithmen skalieren kubisch mit der Elektronenzahl. Die Verwendung von Realraumansätzen, die im idealen Falle linear skalieren, wären hier eher Methode der Wahl. Mit den QUAMOLs sind geeignete Basisätze vorhanden um solche Systemgrößen akkurat und effizient zu modellieren. Dies setzt jedoch noch die Entwicklung eines effizienten LCAO Codes oder die Erzeugung geeigneter Tight-Binding-Hamiltonmatrizen voraus, deren Entwicklung zeitintensiv gewesen wäre und nicht im zentralen Fokus dieser Dissertation liegen. Entscheidender ist es, sich dem eigentlichen Materialsystem zu widmen, um nun weitere Gründe für die experimentell beobachtete Limitierung der p -Dotierbarkeit zu finden. Zugang zu modernen physikalischen Ansätzen, wie der Ladungskorrektur und die Modellierung gängiger Defekte auf HSE Niveau erlauben nun eine akkurate Beschreibung dieser Defekte unabhängig von dem Bandlückenproblem in einer LDA oder PBE Modellierung.

7.4. Hohe Mg-Akzeptorkonzentration: Gratwanderung zwischen Kompensation und Phasentrennung

7.4.1. Einleitung

In Kapitel 7.3.5 sind die Eigenschaften von Magnesiumnitrid hinsichtlich des Wasserstoffeinbaus untersucht worden. Es wurde gezeigt, dass reines Magnesiumnitrid Wasserstoff nur im geringen Maße bindet. Eine deutliche Steigerung des Einbaus durch die geometrischen Veränderungen, die sich bei Magnesiumnitrid in Form von Inversionsdomänengrenzen ergeben, ist nicht zu erwarten. Weiterhin machen die Inversionsdomänengrenzen nur einen geringen Teil des gewachsenen Galliumnitrids aus, so dass eine signifikante Störung des Wasserstoffeinbaus nicht zu erwarten ist. Deswegen wird nun im Folgenden die Modifikation der Wasserstoffkonzentration mit Hinblick auf die Defektbildung im Galliumnitrid untersucht.

Die Defektenergetik ist in vielen Arbeiten[5, 6, 7, 4, 101] diskutiert worden und gilt als verstanden. Jedoch können die mit unterschiedlichen Methoden berechneten Bildungsenergien stark voneinander abweichen, wie Abbildung 7.11 illustriert. Bei solchen Differenzen, die je nach Defekt und Methode einige eV umfassen können, stellt sich natürlich die Frage nach der Aussagekraft und der Berechtigung des Vertrauens in diese Zahlenwerte. Das folgende Kapitel wird sich — nach einer Einführung der wichtigsten Defekte des mit Magnesium und Wasserstoff dotierten GaN — der Frage widmen wie unterschiedlich die berechneten Bildungsenergien tatsächlich sind und welche Ursachen diese eklatanten Abweichungen haben. Die Kenntnis

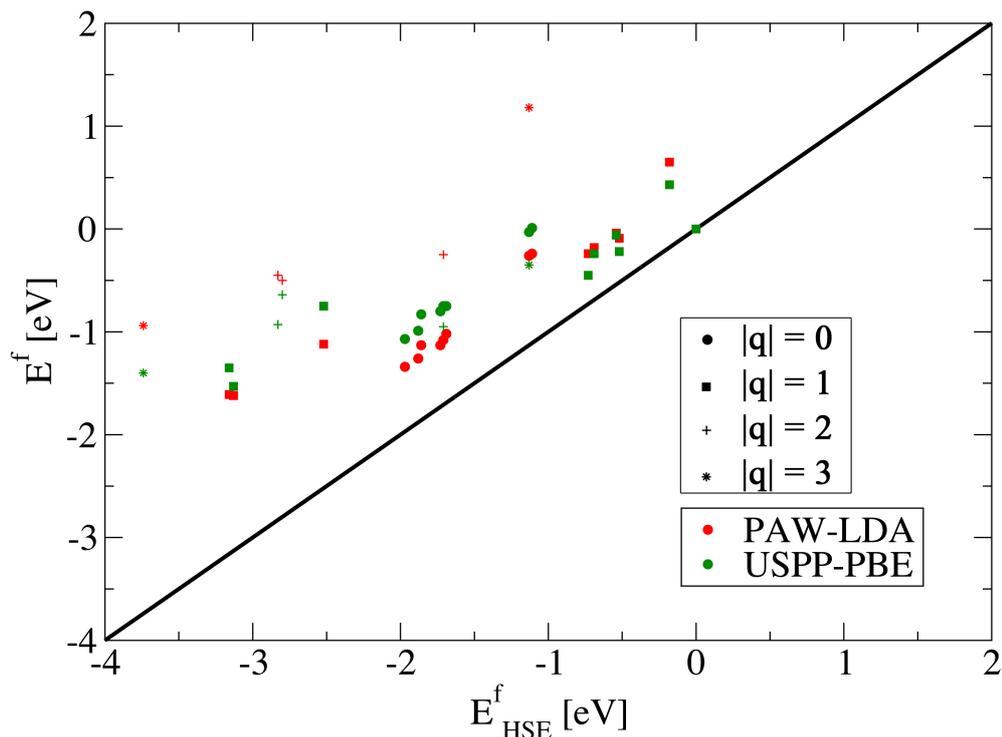


Abb. 7.11:

Unterschiede der Defektenergetik wichtiger Defekte in GaN:Mg von LDA und PBE zu HSE. Die schwarze Linie gibt den Fall perfekter Übereinstimmung an. Die roten Symbole geben den Vergleich der LDA Daten, die grünen den Vergleich der PBE Daten[101] an. Die unterschiedlichen Symbole verweisen auf den jeweilige Ladungszustand des Defektes (ohne Vorzeichen).

der Ursachen erlaubt hier erstmals die Einführung geeigneter Korrekturen, welche zu einer besseren Übereinstimmung der Defektenergetik (bis auf wenige Zehntel eV) führen. Dies wird dabei am Beispiel der Datensätze demonstriert. Aus den Bildungsenergien werden dann Defektkonzentrationen im Falle des GaN Wachstumsprozesses berechnet, welche dann den Grund des verminderten Wasserstoffeinbaus bei hohen Mg-Konzentrationen und den damit verbundenen wechselnden Kompensationsmechanismus aufzeigen.

7.4.2. Defekte in GaN

Dieser Abschnitt widmet sich der Nomenklatur der betrachteten Defekte in GaN und soll durch Abbildung 7.12 ein visuelles Verständnis der Defekte ermöglichen.

Die intrinsische Stickstoffvakanz V_{N} wird durch das Entfernen eines Stickstoffatoms aus dem atomaren Gitter erzeugt. Der bevorzugte Ladungszustand ist der einfach positiv geladene Zustand. Für p -leitende Bedingungen wird zudem noch der dreifach positiv geladene Zustand vorhergesagt[4, 101]. Das Umladungsniveau liegt dabei ca.

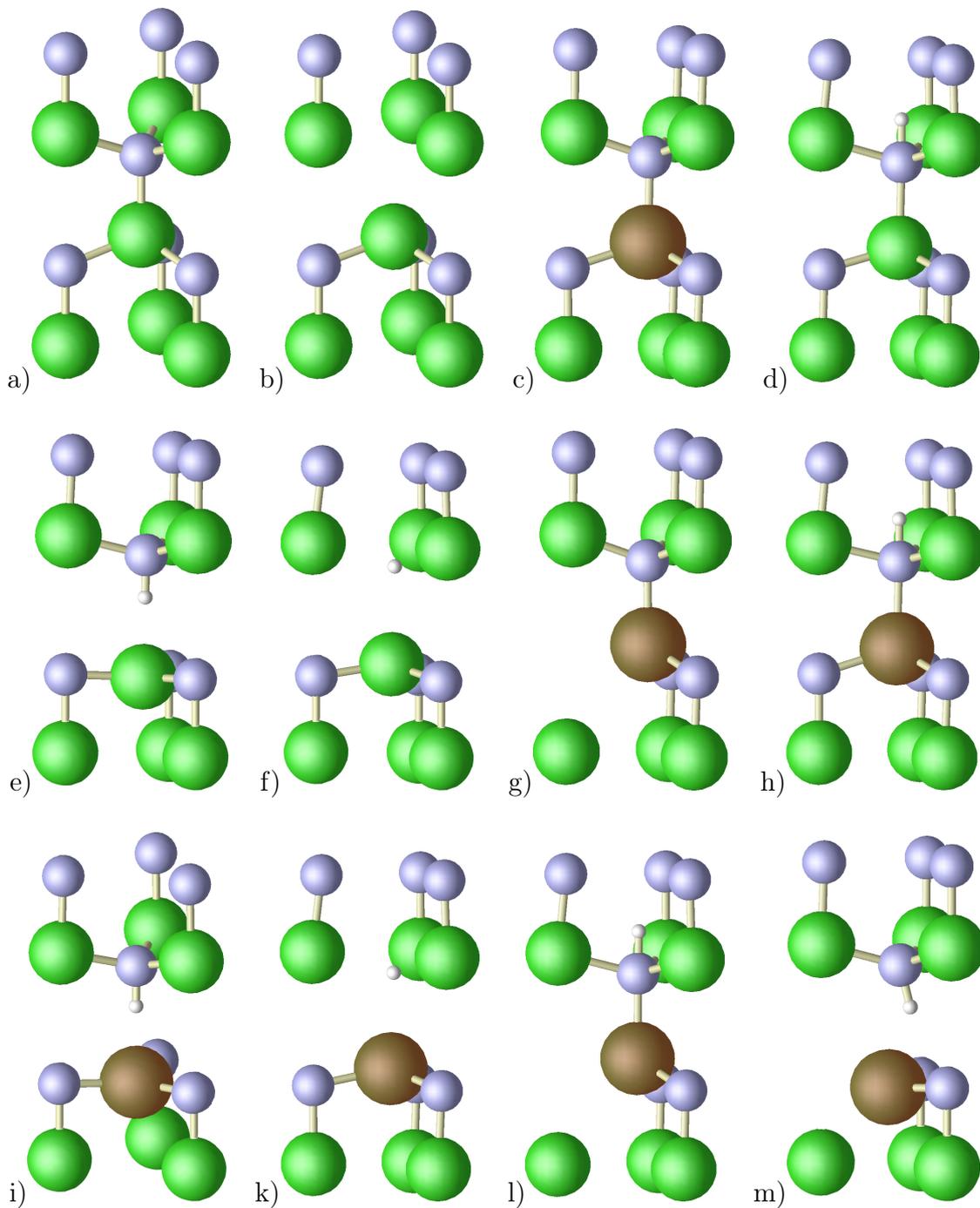


Abb. 7.12:

Darstellung einiger wichtiger Defekte in GaN: (Ga grün, N blau, Mg braun)

a) bulk b) V_N c) Mg_{Ga} d) $H_{\parallel}(ab)$ e) $H_{\parallel}(bc)$ f) H_N g) MgV_N h) $[MgH]_{\parallel}(ab)$

i) $[Mg_{Ga}H]_{\parallel}(bc)$ k) $[Mg_{Ga}H_N]$ l) $[Mg_{Ga}V_NH]_{\parallel}(ab)$ m) $[Mg_{Ga}V_NH]_{\parallel}(bc)$

0.4-0.5 eV oberhalb der Valenzbandkante. Ein weiterer Punktdefekt von Interesse in GaN:Mg ist das substitutionelle Magnesium auf einen Galliumgitterplatz Mg_{Ga} . Dieser Defekt tritt hauptsächlich in einem einfach negativ geladenen Zustand auf. Experimentelle Befunde geben ein Umladungsniveau von 0.16 eV oberhalb des Valenzbandmaximums an, unter welchem substitutionelles Magnesium in einem neutralen Ladungszustand vorliegt[102]. Wasserstoff besetzt Zwischengitterplätze. Experimentell wird er in das Material gebracht um den negativ geladenen Magnesiumdefekt zu kompensieren, weswegen hier nur der positiv geladene Wasserstoffdefekt H^+ diskutiert wird. Der negativ geladene Wasserstoffdefekt, sowie der neutrale, spielen für n -leitende Bedingungen eine Rolle, wo sich das Fermi-niveau näher am Leitungsbandminimum befindet. Aufgrund seiner positiven Ladung bindet der Wasserstoff an den elektronegativeren Stickstoff im Kristall. Dabei unterscheidet man zwischen der *anti-bindenden* (AB) Position, in welcher der Wasserstoff in einer Ga-N-H Kette bindet, und der sogenannten *bindungs-zentrierten* (BC) Position, in welcher der Wasserstoff die Ga-N Bindung bricht und sich anstelle des Galliums als Bindungspartner anbietet. Aufgrund der ausgezeichneten Kristallrichtungen in der Wurtzitstruktur werden diese Wasserstoffdefekte noch weiter unterschieden. So existieren solche deren formale Bindungsachse parallel zur c -Richtung des Kristalls liegt (\parallel) und solche deren formale Bindungsachse einen deutlichen Anteil senkrecht zur c -Achse besitzt (\perp).

Weiter ist noch substitutioneller Wasserstoff an der Stickstoffposition H_{N} denkbar, welcher formal schon als erster Defektkomplex einer Stickstoffvakanz mit Wasserstoff aufgefasst werden kann. Aufgrund der zwei positiven Ladungszustände der Stickstoffvakanz sind zwei Ladungszustände von H_{N} interessant. Im 2-fach positiven Ladungszustand paart sich eine 3-fach positiv geladene Stickstoffvakanz mit einem negativ geladenen Wasserstoff. Im neutralen Zustand paart sich V_{N}^+ mit H^- . Unsere Rechnungen zeigen, dass das Umladungsniveau mit 2.01 eV im energetisch höheren Bereich der Bandlücke sitzt. Der H_{N}^0 -Defekt spielt also für die hier gemachten Betrachtungen keine Rolle.

Neben dem eher formellen Defektkomplex H_{N} gibt es eine Reihe tatsächlicher Defektkomplexe, die beim Wachstum von Magnesium dotiertem GaN eine Rolle spielen. Zu den Komplexen bestehend aus zwei Punktdefekten zählen hierbei $[\text{MgV}_{\text{N}}]$ im Ladungszustand +2 und 0, sowie der neutrale $[\text{MgH}]$ -Komplex. Neben den Defektkomplexen, welche aus zwei Punktdefekten bestehen, diskutieren wir noch solche, die aus drei Punktdefekten bestehen, die sogenannten $[\text{MgHV}_{\text{N}}]$ -Komplexe. Die Anzahl möglicher Komplexkonfigurationen nimmt hier schon deutlich zu. Die Nomenklatur wird dabei mit den eingeführten Größen fortgesetzt.

Nach der Darstellung der wichtigsten Punktdefekte und ihrer Komplexe folgt nun in den nächsten Abschnitten die Suche nach den Ursachen der offensichtlichen

Defekt	Bildungsenergie [eV]		
	PAW-LDA	PAW-HSE	USPP-PBE[101]
Mg_{Ga}^-	0.00	0.00	0.00
V_{N}^+	0.65	-0.18	0.43
V_{N}^{3+}	1.18	-1.13	-0.35
$\text{H}_{\parallel}^+(\text{bc})$	-0.24	-0.95	-0.45
$[\text{Mg}_{\text{Ga}}\text{H}_{\perp}]^0(\text{ab Mg}_{\perp})$	-1.34	-1.97	-0.75
$[\text{Mg}_{\text{Ga}}\text{V}_{\text{N}}]_{\perp}^0$	-0.26	-1.13	0.29
$[\text{Mg}_{\text{Ga}}\text{V}_{\text{N}}]_{\perp}^{2+}$	-0.45	-2.83	-0.61
H_{N}^{2+}	-0.25	-1.71	-0.95
$[\text{Mg}_{\text{Ga}}\text{H}_{\text{N}}]_{\parallel}^+$	-1.61	-3.16	-1.21
$[\text{Mg}_{\text{Ga}}\text{H}_{\text{N}}]_{\perp}^+$	-1.62	-3.13	-1.03
$[\text{Mg}_{\text{Ga}}\text{HV}_{\text{N}}]_{\parallel}^+(\text{ab Mg})$	-1.12	-2.52	-0.75
$[\text{Mg}_{\text{Ga}}\text{HV}_{\text{N}}]_{\parallel}^{3+}(\text{ab Mg})$	-0.94	-3.74	-1.40

Tab. 7.9:

Bildungsenergien einiger ausgewählter Defekte im wasserstoffkompensierten GaN:Mg. Als Referenzsystem ist das GaN, Galliummetall, der negativgeladene Mg^- Defekt und das Wasserstoffmolekül gewählt worden. Die Fermienergie ist auf das Valenzbandmaximum referenziert.

Diskrepanzen der mit den unterschiedlichen Funktionalen LDA, PBE und HSE berechneten Defektbildungsenergien. Dabei sind die PBE Daten der Literatur[101] entnommen worden. Die HSE Daten entstammen einer Kooperation mit der Universität Californien — Santa Barbara.

7.4.3. Vergleich der Methoden

Betrachtet man die in der Tabelle 7.9 vorgestellte Defektenergetiken für die verschiedenen Funktionale genauer, so fällt auf, dass die relativen Unterschiede innerhalb einzelner Defektklassen mit gleicher Ladung gar nicht so eklatant sind, wie

Defekt	Umladungsniveau [eV]		
	PAW-LDA	PAW-HSE	USPP-PBE[101]
$\text{V}_{\text{N}}^{3+/+}$	-0.27	0.48	0.39
$\text{MgV}_{\text{N}}^{2+/0}$	0.12	0.85	0.45
$\text{MgHV}_{\text{N}}^{3+/+}$	-0.09	0.33	-0.34

Tab. 7.10:

Vergleich der Umladungsniveaus wichtiger Defekte in GaN:Mg referenziert auf das Valenzbandmaximum für LDA, PBE und HSE.

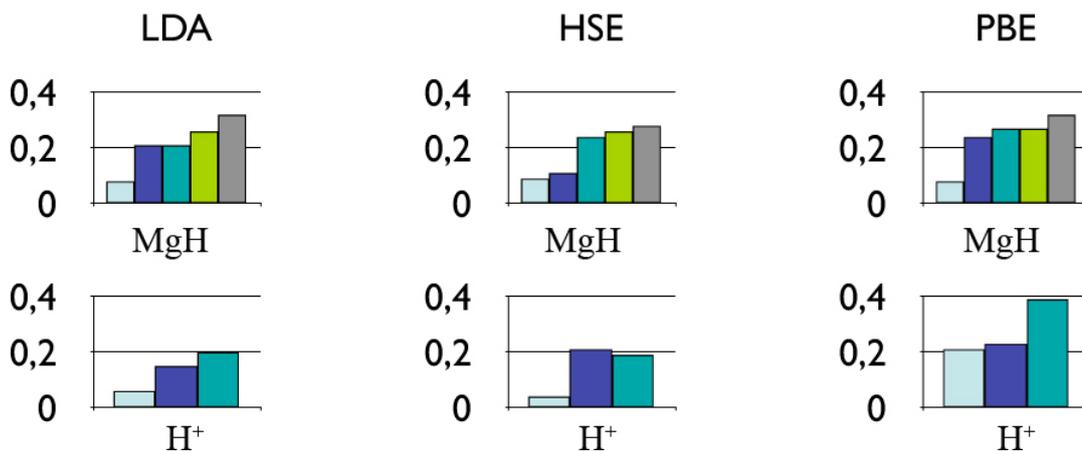


Abb. 7.13:

Unterschiede der Defektenergetik in LDA, HSE und PBE für MgH und H⁺. Jede Säule steht für eine berechnete Defektconfiguration innerhalb der Defektklasse. Die Höhen der Säulen gibt die Bildungsenergie in eV relativ zum stabilsten Defekt dieser Klasse an.

zunächst zu vermuten ist. Im Falle des neutralen MgH Komplexes und der positiv geladenen Wasserstoffstörstelle stimmen gemäß Abbildung 7.13 sowohl Reihenfolge der Stabilität, als auch die Bildungsenergie bis auf 0.2 eV überein. Vergleicht man jedoch unterschiedlich geladene Defekte, zum Beispiel anhand des Umladungsniveaus wie in Tabelle 7.10, ergibt sich ein geradezu ernüchterndes Bild. Die Umladung findet an derart qualitativ unterschiedlichen Positionen im Energiespektrum statt, dass von einem vertrauenswürdigen Datensatz nicht gesprochen werden kann. Zwar zeigt das Umladungsniveau der Stickstoffvakanz zwischen HSE und PBE eine gute Übereinstimmung, jedoch ist die erwartete Vorhersagegenauigkeit zwischen HSE und PBE nicht gleichwertig. Diese Übereinstimmung ist — auch mit Sicht auf die Diskrepanz der anderen Niveaus — eher als zufällig zu bewerten.

Die Betrachtung lässt den Schluss zu, dass die Unterschiede der einzelnen Defektenergetiken vornehmlich mit der Ladung korrelieren. An der Gleichung der Bildungsenergie geladener Defekte

$$E^f[X^q] = E_{\text{tot}}[\text{bulk}:X^q] - E_{\text{tot}}[\text{bulk}] - \sum_i n_i \mu_i + E_{\text{Corr}}[q] + q(\epsilon_{\text{VBM}} + \epsilon_{\text{Fermi}}) \quad (7.13)$$

erkennt man, dass die Ladung nur in der Superzellenkorrektur und in der Positionierung des Fermi-niveaus als chemisches Potential der Elektronen eine Rolle spielt. Aber auch die Referenzen der chemischen Potentiale können in den einzelnen Methoden unterschiedlich beschrieben werden und somit zu quantitativen Differenzen führen. So ist bekannt, dass gerade Moleküle in der LDA, der PBE und nicht zuletzt auch auf dem Niveau von HSE recht unterschiedlich beschrieben werden, was

Ladung	Ladungskorrektur [eV]		
	$E_{\text{Madelung}}^{\text{PBE}}$	$E_{\text{Align}}^{\text{LDA}}$	
		min	max
1	0.21	-0.09	0.00
2	0.86	-0.30	-0.23
3	1.93	-0.77	-0.49

Tab. 7.11:

Ladungskorrekturen für Defekte in GaN aufgeteilt in den Madelungbeitrag und Alignmentbeitrag. Der Madelungbeitrag wurde für die in PBE verwendete 72 Atomzelle explizit berechnet. Der Alignmentbeitrag kann ohne Neuberechnung der Defektgeometrie nur abgeschätzt werden. Dafür dienen die Alignmentkorrekturen der LDA Rechnungen, welche gemäss $1/L^3$ von der 96 Atomzelle auf die 72 Atomzelle skaliert wurden. Da das Alignment defektspezifisch ist, sind hier der jeweils kleinste und größte Wert für die Defekte der jeweiligen Ladungsklasse angegeben.

zu der Diskrepanz in der Vorhersage der Bildungsenthalpie von GaN führt[5]. Diese Überlegungen werden nun im Folgenden weiter ausgearbeitet. Basierend auf den Erkenntnissen werden Korrekturen entworfen, um den Unzulänglichkeiten der Funktionale entgegenzuwirken.

7.4.4. Einfluss der Superzellenkorrektur

Laaksonen hat gezeigt, dass das Umladungsniveau der Stickstoffvakanz empfindlich von der Superzellengröße abhängt, wenn keine superzellenspezifische Ladungskorrektur angewendet wird[103]. Ihre theoretischen Betrachtungen der Stickstoffvakanz auf LDA Niveau und die Interpolation zu unendlich ausgedehnten Superzellen ergeben ein Umladungsniveau unterhalb der Valenzbandkante. Das bedeutet, dass der V_{N}^{3+} -Defekt bei einem Fermilevel oberhalb der Valenzbandkante kein stabiler Defektzustand ist. Dies steht im Widerspruch zum experimentellen Befund, aber auch zu früheren theoretischen Betrachtungen. Auch die hier berechneten LDA Ergebnisse zeigen diesen Befund. Dies lässt den Schluss zu, dass die Position des Umladungsniveaus früherer Arbeiten auf die Superzellengröße und die Wechselwirkung des Defektes mit seinen periodischen Bildern zurückzuführen ist. Die Superzellenkorrektur für geladene Defekte ist folglich zwingend erforderlich um konsistente Ergebnisse in den theoretischen Methoden zu erhalten. Da die PBE Daten der Literatur keine Ladungskorrektur erfahren haben, ist diese nun nachträglich anzuwenden. Die hier in dieser Dissertation verwendete Superzellenkorrektur kann gemäß Kapitel 5.4.1 als

$$E_{\text{Corr}} = \frac{\alpha q^2}{2\epsilon L} + E_{\text{Align}}, \quad (7.14)$$

dargestellt werden. Der typischerweise dominante Madelungsterm hängt dabei nur von der Geometrie der Superzelle ab und kann leicht berechnet werden. Das

Defekt	Bildungsenergie [eV]		
	USPP-PBE	PAW-LDA	
	Ref. [101]	corr.	corr.
Mg_{Ga}^-	0.00	0.00	0.00
V_{N}^+	0.43	0.55	0.65
V_{N}^{3+}	-0.35	0.98	1.18
$\text{H}_{\parallel}^+(\text{bc})$	-0.45	-0.25	-0.24
$\text{H}_{\perp}^+(\text{bc})$	-0.22	-0.02	-0.09
$\text{H}_{\parallel}^+(\text{ab})$	-0.06	0.16	-0.04
$\text{H}_{\perp}^+(\text{ab})$	-0.24	-0.03	-0.18
H_{N}^{2+}	-0.95	-0.33	-0.25
$[\text{Mg}_{\text{Ga}}\text{H}]_{\parallel}^0(\text{bc})$	-0.99	-1.17	-1.26
$[\text{Mg}_{\text{Ga}}\text{H}]_{\parallel}^0(\text{ab})$	-0.80	-0.98	-1.13
$[\text{Mg}_{\text{Ga}}\text{H}]_{\perp}^0(\text{bc Mg}_{\perp})$	-0.75	-0.93	-1.02
$[\text{Mg}_{\text{Ga}}\text{H}]_{\perp}^0(\text{ab Mg}_{\perp})$	-1.07	-1.25	-1.34
$[\text{Mg}_{\text{Ga}}\text{H}]_{\perp}^0(\text{ab Ga}_{\parallel})$	-0.75	-0.93	-1.08
$[\text{Mg}_{\text{Ga}}\text{H}]_{\perp}^0(\text{ab Ga}_{\perp})$	-0.83	-1.01	-1.13
$[\text{Mg}_{\text{Ga}}\text{V}_{\text{N}}]_{\parallel}^0$	0.01	-0.17	-0.24
$[\text{Mg}_{\text{Ga}}\text{V}_{\text{N}}]_{\perp}^0$	-0.03	-0.21	-0.26
$[\text{Mg}_{\text{Ga}}\text{V}_{\text{N}}]_{\parallel}^{2+}$	-0.64	-0.26	-0.50
$[\text{Mg}_{\text{Ga}}\text{V}_{\text{N}}]_{\perp}^{2+}$	-0.93	-0.55	-0.45
$[\text{Mg}_{\text{Ga}}\text{H}_{\text{N}}]_{\parallel}^+$	-1.35	-1.39	-1.61
$[\text{Mg}_{\text{Ga}}\text{H}_{\text{N}}]_{\perp}^+$	-1.53	-1.56	-1.62
$[\text{Mg}_{\text{Ga}}\text{HV}_{\text{N}}]_{\parallel}^+(\text{ab Mg})$	-0.75	-0.77	-1.17
$[\text{Mg}_{\text{Ga}}\text{HV}_{\text{N}}]_{\parallel}^{3+}(\text{ab Mg})$	-1.40	-0.42	-0.94

Tab. 7.12:

Vergleich der Bildungsenergien in GaN:Mg in PBE mit LDA vor und nach Anwendung der Ladungskorrektur.

Alignment hingegen benötigt die elektronische Dichteverteilung des Defektes. Nun ist es nicht praktikabel diese zu erzeugen, indem alle Defekte der Literatur nachgerechnet werden. Wir verwenden deswegen hier als Abschätzung die typischen Größen, die in der LDA Behandlung berechnet wurden und skalieren diese entsprechend mit der Superzellengröße. Ziel ist es an dieser Stelle nicht möglichst akkurate Defektenergetiken für PBE zu erzeugen, sondern vielmehr zu zeigen, dass die Diskrepanzen zwischen LDA und PBE vornehmlich auf der nicht verwendeten Ladungskorrektur der PBE Daten beruhen. Die berechneten Zahlenwerte für die Ladungskorrektur sind ladungsspezifisch in Tabelle 7.11 zusammengefasst. Die damit korrigierten PBE Daten sind in Tabelle 7.12 zusammengefasst.

Die Anwendung der Ladungskorrekturen verbessert die Übereinstimmung zwischen den LDA und den PBE signifikant. Die meisten Bildungsenergien weisen nur noch Differenzen von unter 0.2 eV auf. Allerdings gibt es auch einige Defekte, die trotz der Ladungskorrektur größere Differenzen aufweisen. Zunächst einmal fällt auf, dass die Übereinstimmung der Wasserstoffstörstellen in der anti-bindenden Position sich mit der Ladungskorrektur verschlechtert haben. Eigene Berechnungen des Defektes mittels PAW-PBE zeigen hingegen eine zu LDA vergleichbare Energetik, so dass das Funktional nicht Ursache dieser Diskrepanz sein kann. Vielmehr scheint der Unterschied in der Behandlung der Elektron-Kern-Wechselwirkung zu bestehen. Diese Arbeit verwendet den PAW Ansatz im Unterschied zu den Literaturdaten, die mit ultrasofthen Pseudopotentialen erzeugt worden sind. Während PAW als genäherte Allelektronenrechnung physikalisch korrekte Wellenfunktionen in Kernnähe liefert, ist das ultrasofte Pseudopotential eher „trickreich“ in der Beschreibung des Wasserstoffes mit möglichst wenigen ebenen Wellen. Die Reaktion dieses Potentials auf die sich hier drastisch ändernde chemische Umgebung ist nicht bekannt und kann durchaus die Ursache der Diskrepanz sein. Einen ähnlichen Grund kann auch die unterschiedliche Beschreibung des MgV_N^{2+} und des MgH_N^+ haben. In der PAW-LDA zeigen sich diese Defekte nahezu entartet, während in der USPP-PBE Variante die Niveaus um 0.25–0.3 eV aufspalten. Hierbei ist die Aufspaltung des MgH_N^+ sicherlich auf die des MgV_N^{2+} zurückzuführen, da dieser Bestandteil jenes Defektkomplexes ist.

Auch die Diskrepanzen der ausgedehnten $[\text{Mg}_{\text{Ga}}\text{HV}_\text{N}]$ -Komplexe könnten auf die Pseudoisierung von Stickstoff und Wasserstoff zurückzuführen sein, jedoch möchte ich an dieser Stelle anmerken, dass die Ausdehnung des Defektes hier schon eine Größe erreicht, welche die Beschreibung in der aktuellen Superzelle, sowie die Verwendung der Ladungskorrektur fragwürdig erscheinen lassen. Zudem findet die Behandlung in Superzellen unterschiedlicher Größe statt, so dass die endliche Größe der Superzelle sicherlich quantitativ spürbar wird. Für eine akkurate Beschreibung dieses Defektes ist eine Behandlung in einer größeren Defektzelle anzuraten, jedoch ist diese erheblich rechenintensiver weswegen an dieser Stelle darauf verzichtet wird.

Trotz dieser kleinen verbleibenden Diskrepanzen ist belegt, dass der Unterschied zwischen LDA und PBE hauptsächlich auf die fehlende Ladungskorrektur der PBE Daten zurückgeführt werden kann. Die Umladungsniveaus zeigen nun eine deutlich bessere Übereinstimmung zwischen beiden Ansätzen, so dass nun lediglich noch die Unterschiede zwischen LDA/PBE und HSE zu betrachten sind.

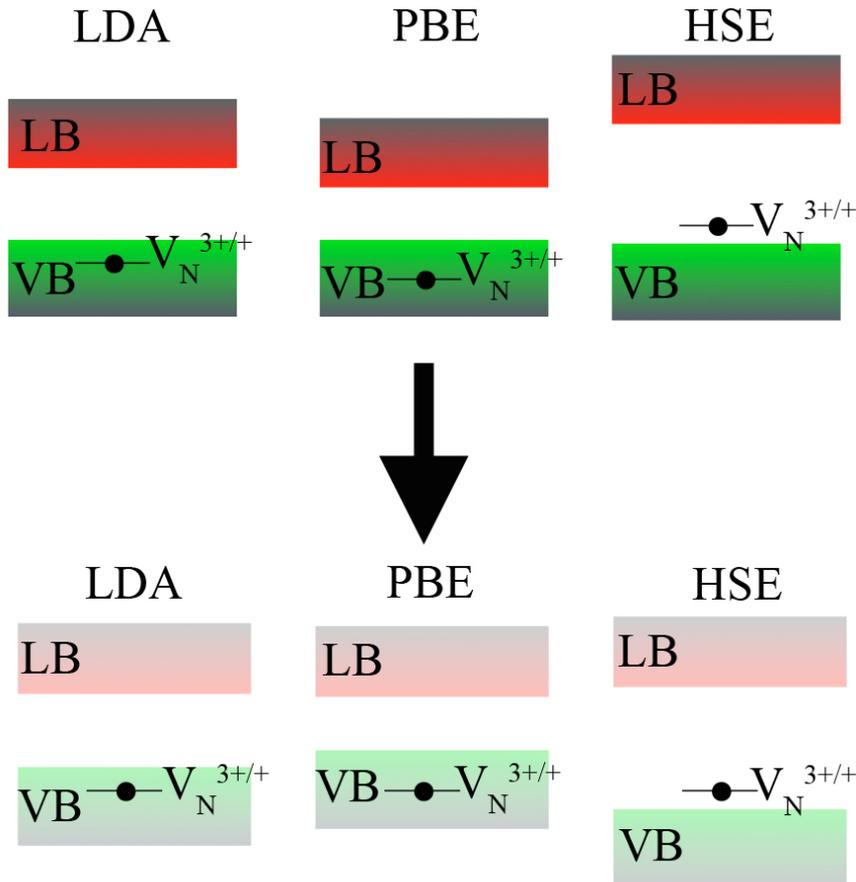
7.4.5. Das Bandlückenproblem — LDA/PBE versus HSE

Der Unterschied zwischen den „klassischen“ Funktionalen LDA/PBE und dem Hybridfunktional HSE liegt in der Beimischung des exakten Hartree-Fock-Austausches für den kurzreichweitigen Teil der Coulombwechselwirkung. In Folge dessen deloka-

Defekt	Bildungsenergie [eV]	
	PAW-LDA	PAW-HSE
Mg_{Ga}^-	0.00	0.00
V_{N}^+	0.38	0.29
V_{N}^{3+}	0.38	0.29
H_{\parallel}^+ (bc)	-0.51	-0.26
H_{\perp}^+ (bc)	-0.36	-0.05
H_{\parallel}^+ (ab)	-0.31	-0.07
H_{\perp}^+ (ab)	-0.45	-0.22
H_{N}^{2+}	-0.79	-0.76
$[\text{Mg}_{\text{Ga}}\text{H}]_{\parallel}^0$ (bc)	-1.52	-1.41
$[\text{Mg}_{\text{Ga}}\text{H}]_{\parallel}^0$ (ab)	-1.39	-1.26
$[\text{Mg}_{\text{Ga}}\text{H}]_{\perp}^0$ (bc Mg_{\perp})	-1.29	-1.22
$[\text{Mg}_{\text{Ga}}\text{H}]_{\perp}^0$ (ab Mg_{\perp})	-1.60	-1.50
$[\text{Mg}_{\text{Ga}}\text{H}]_{\perp}^0$ (ab Ga_{\parallel})	-1.34	-1.24
$[\text{Mg}_{\text{Ga}}\text{H}]_{\perp}^0$ (ab Ga_{\perp})	-1.40	-1.39
$[\text{Mg}_{\text{Ga}}\text{V}_{\text{N}}]_{\parallel}^0$	-0.50	-0.64
$[\text{Mg}_{\text{Ga}}\text{V}_{\text{N}}]_{\perp}^0$	-0.53	-0.66
$[\text{Mg}_{\text{Ga}}\text{V}_{\text{N}}]_{\parallel}^{2+}$	-1.30	-1.38
$[\text{Mg}_{\text{Ga}}\text{V}_{\text{N}}]_{\perp}^{2+}$	-1.25	-1.41
$[\text{Mg}_{\text{Ga}}\text{H}_{\text{N}}]_{\parallel}^+$	-2.14	-2.21
$[\text{Mg}_{\text{Ga}}\text{H}_{\text{N}}]_{\perp}^+$	-2.15	-2.18
$[\text{Mg}_{\text{Ga}}\text{HV}_{\text{N}}]_{\parallel}^+$ (ab Mg)	-1.65	-1.58
$[\text{Mg}_{\text{Ga}}\text{HV}_{\text{N}}]_{\perp}^+$ (bc Mg_{\parallel})	-1.58	-1.55
$[\text{Mg}_{\text{Ga}}\text{HV}_{\text{N}}]_{\perp}^+$ (ab Mg_{\parallel})	-1.41	-1.33
$[\text{Mg}_{\text{Ga}}\text{HV}_{\text{N}}]_{\parallel}^{3+}$ (ab Mg)	-2.01	-1.85
$[\text{Mg}_{\text{Ga}}\text{HV}_{\text{N}}]_{\perp}^{3+}$ (bc Mg_{\parallel})	-1.50	-1.55
$[\text{Mg}_{\text{Ga}}\text{HV}_{\text{N}}]_{\perp}^{3+}$ (ab Mg_{\parallel})	-1.52	-1.61

Tab. 7.13:

Mit LDA und HSE berechnete Bildungsenergien der vornehmlich vorkommenden Defekte in Wasserstoff kompensierten GaN:Mg. Die Referenzzustände sind Ga Metall, GaN, Mg_{Ga}^- , sowie das halbe Wasserstoffdimer. Das Fermi-niveau ist auf das $\text{V}_{\text{N}}^{3+/+}$ Umladungsniveau gesetzt.

**Abb. 7.14:**

Wechsel der Fermilevel Referenz vom Valenzbandmaximum zum Umladungsniveau der Stickstoffvakanz für LDA, PBE und HSE. Die roten Bereiche symbolisieren die Leitungsbander, die grünen Bereiche die Valenzbander. Dargestellt ist die Position des Umladungsniveaus der Stickstoffvakanz.

lisiert das Austauschloch mehr als in der LDA/PBE Behandlung, was zu einer Aufspaltung der Bandlücke in halbleitenden Materialien führt. Durchforstet man die Literatur trifft man auf einige Arbeiten, welche die gängige Ansicht vertreten, dass die Aufspaltung durch eine Verschiebung des Leitungsbandminimums in einen höheren Energiebereich erreicht wird[104, 105, 106]. In Folge dessen sind Defekte prozentual gemäß ihres leitungsbandähnlichen Charakter zu korrigieren. Heutzutage hat man jedoch einige berechtigte Zweifel an diesem Bild. So haben beispielsweise Alkauskas und Pasquarello anhand der Sauerstoffvakanz in ZnO gezeigt, dass die mit verschiedenen Methoden berechneten tiefen Umladungsniveaus eine deutlich geringere Diskrepanz haben, wenn sie auf das mittlere elektrostatische Potential referenziert werden[40]. Auch für andere Materialien ist dies eine geeignete Wahl[41]. Nun stellt das mittlere elektrostatische Potential aus experimenteller Sicht keine gut messbare Größe dar. Wir wählen hier den pragmatischen Weg das Fermi-niveau direkt auf das Umladungsniveau der Stickstoffvakanz zu referenzieren (Abbildung 7.14). Zwar ist

das Umladungsniveau der Stickstoffvakanz nach der Ladungskorrektur kein tiefes Defektniveau mehr, allerdings ist es in der Modellierung in der endlichen Superzelle auf LDA Niveau vor der Korrektur durchaus als solches zu sehen und erfüllt somit das Kriterium von Alkauskas und Pasquarello. Tabelle 7.13 zeigt die Bildungsenergien von LDA und HSE in dieser neuen Referenzierung. Die Diskrepanzen belaufen sich nun nur noch auf 0.2 eV für den Grossteil der Defekte. Eine Ausnahme stellt hier wieder der Wasserstoff auf den Zwischengitterplätzen dar, der aber auch innerhalb von 0.3 eV eine gute Übereinstimmung zeigt. Generell scheint die LDA die Bildungsenergie dieser Wasserstoffdefekte zu unterschätzen.

7.4.6. Feinjustierung — Die chemischen Potentiale

Die vorangegangenen Kapitel haben gezeigt, dass die Diskrepanzen zwischen LDA, PBE und HSE hauptsächlich auf die nicht vorgenommene Ladungskorrektur der PBE-Daten und der Referenzierung des Fermilevels auf das Valenzbandmaximum, welches zwischen LDA/PBE und HSE unterschiedlich beschrieben wird, zurückzuführen sind. Vollzieht man eine Umreferenzierung des Fermilevels auf das Umladungsniveau der Stickstoffvakanz lassen sich die Defektenergetiken bis auf 0.3 eV in Einklang bringen, wobei die reinen Wasserstoffdefekte die größte Diskrepanz aufweisen. Eine letzte Stellschraube stellen hier die chemischen Potentiale der einzelnen Atomsorten dar. Die Referenzierung auf molekulare Systeme, wie im Falle des Wasserstoffes, kann hier noch zu Diskrepanzen zwischen den mit unterschiedlichen Funktionalen berechneten Defektenergetiken führen. Schon Myers nutzt zur Referenzierung des chemischen Potentials für Magnesium den Magnesiumdefekt[101]. An dieser Stelle soll nun untersucht werden, in wie fern sich eine komplette Referenzierung aller chemischen Potentiale auf Defektzustände positiv auf die Defektenergetiken auswirkt. Im folgenden werden nun das substitutionelle Magnesium auf einem Galliumgitterplatz, die Stickstoffvakanz und der stabilste MgH-Defektkomplex als Referenzsysteme für die chemischen Potentiale eingeführt. Die Stickstoffvakanz begründet sich mit der schon vollzogenen Referenzierung auf ihr Umladungsniveau für das Fermilevel. Der MgH-Komplex stellt experimentell den bedeutendsten Defekt dar um die Löslichkeit von Magnesium in GaN zu erhöhen und scheint eine geeignete Wahl zu sein. Die so referenzierte Defektenergetik ist für LDA und HSE in Tabelle 7.14 zusammengefasst. Die Diskrepanzen zwischen den Wasserstoffstörstellen sind kleiner geworden. Insgesamt stimmt nun die Defektenergetik bis auf maximal 0.2 eV überein.

Damit ist eine Validierung der Defektbildungsenergien gelungen, wenn entsprechende Korrekturterme mit einbezogen werden. Allerdings bietet nur HSE die Möglichkeit, die hier gewählten Referenzdefekte konsistent mit experimentell verfügbaren Referenzen, wie dem Valenzbandmaximum, zu beschreiben. Die mit LDA und PBE berechneten Bildungsenergien können aber einen gute Überblick liefern, um interessante Defekte und Mechanismen zu identifizieren, die dann auf dem akkuraterem

Defekt	Bildungsenergie [eV]	
	PAW-LDA	PAW-HSE
Mg_{Ga}^-	0.00	0.00
V_{N}^+	0.00	0.00
V_{N}^{3+}	0.00	0.00
$[\text{Mg}_{\text{Ga}}\text{H}]_{\perp}^0$ (ab Mg_{\perp})	0.00	0.00
H_{\parallel}^+ (bc)	1.09	1.24
H_{\perp}^+ (bc)	1.25	1.45
H_{\parallel}^+ (ab)	1.29	1.43
H_{\perp}^+ (ab)	1.16	1.28
H_{N}^{2+}	0.43	0.44
$[\text{Mg}_{\text{Ga}}\text{H}]_{\parallel}^0$ (bc)	0.08	0.09
$[\text{Mg}_{\text{Ga}}\text{H}]_{\parallel}^0$ (ab)	0.21	0.24
$[\text{Mg}_{\text{Ga}}\text{H}]_{\perp}^0$ (bc Mg_{\perp})	0.32	0.28
$[\text{Mg}_{\text{Ga}}\text{H}]_{\perp}^0$ (ab Ga_{\parallel})	0.26	0.26
$[\text{Mg}_{\text{Ga}}\text{H}]_{\perp}^0$ (ab Ga_{\perp})	0.20	0.11
$[\text{Mg}_{\text{Ga}}\text{V}_{\text{N}}]_{\parallel}^0$	-0.88	-0.93
$[\text{Mg}_{\text{Ga}}\text{V}_{\text{N}}]_{\perp}^0$	-0.91	-0.95
$[\text{Mg}_{\text{Ga}}\text{V}_{\text{N}}]_{\parallel}^{2+}$	-1.68	-1.67
$[\text{Mg}_{\text{Ga}}\text{V}_{\text{N}}]_{\perp}^{2+}$	-1.63	-1.70
$[\text{Mg}_{\text{Ga}}\text{H}_{\text{N}}]_{\parallel}^+$	-0.92	-1.01
$[\text{Mg}_{\text{Ga}}\text{H}_{\text{N}}]_{\perp}^+$	-0.93	-0.97
$[\text{Mg}_{\text{Ga}}\text{HV}_{\text{N}}]_{\parallel}^+$ (ab Mg)	-0.43	-0.37
$[\text{Mg}_{\text{Ga}}\text{HV}_{\text{N}}]_{\perp}^+$ (bc Mg_{\parallel})	-0.35	-0.34
$[\text{Mg}_{\text{Ga}}\text{HV}_{\text{N}}]_{\perp}^+$ (ab Mg_{\parallel})	-0.19	-0.12
$[\text{Mg}_{\text{Ga}}\text{HV}_{\text{N}}]_{\parallel}^{3+}$ (ab Mg)	-0.79	-0.64
$[\text{Mg}_{\text{Ga}}\text{HV}_{\text{N}}]_{\perp}^{3+}$ (bc Mg_{\parallel})	-0.28	-0.34
$[\text{Mg}_{\text{Ga}}\text{HV}_{\text{N}}]_{\perp}^{3+}$ (ab Mg_{\parallel})	-0.29	-0.40

Tab. 7.14:

Mit LDA und HSE gerechnete Bildungsenergien der hauptsächlich vorkommenden Defekte in Wasserstoff kompensierten GaN:Mg. Die Referenzzustände sind Ga Metall, V_{N}^+ , Mg_{Ga}^- , sowie $[\text{Mg}_{\text{Ga}}\text{H}]_{\perp}^0$ (ab Mg_{\perp}). Das Fermi-niveau ist auf das $\text{V}_{\text{N}}^{3+/+}$ Umladungs-niveau gesetzt.

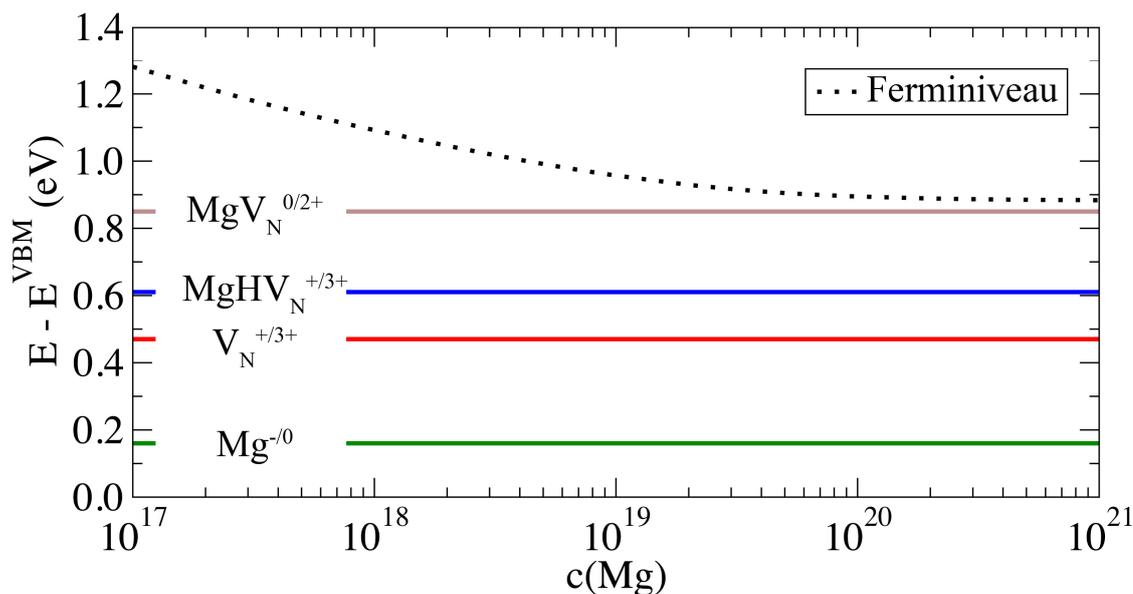


Abb. 7.15:

Position des Fermi-niveaus in GaN:Mg basierend auf der HSE Defektenergetik in Abhängigkeit der Magnesiumkonzentration. Die farbigen Linien zeigen die Umladungsniveaus der entsprechenden Defekte.

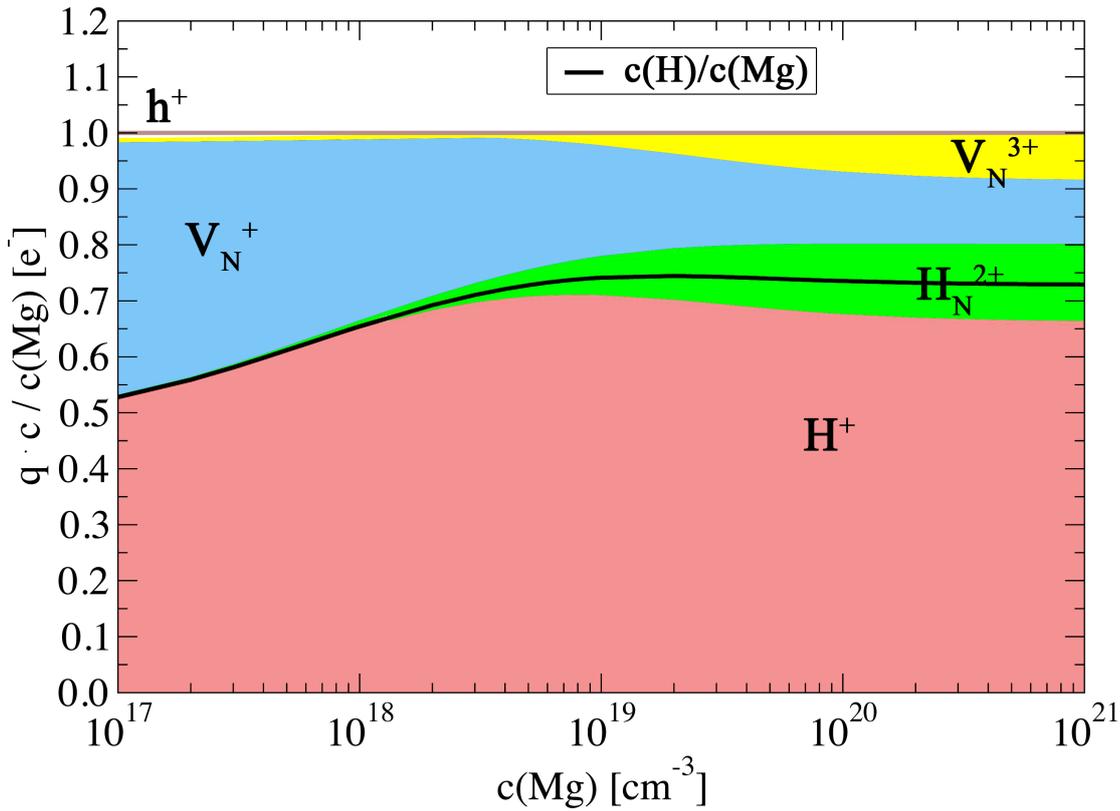
— aber auch erheblich aufwendigerem — HSE Niveau nachgerechnet werden sollten.

7.4.7. Die Kompensationsmechanismen in p -GaN

Der Wechsel in ein adäquates Referenzsystem hat es ermöglicht zu zeigen, dass die berechneten Defektenergetiken nahezu methodenunabhängig sind und somit das Vertrauen in diese Zahlenwerte deutlich gestärkt. Um nun jedoch die Effekte der hohen Magnesiumdotierung zu studieren ist der Wechsel zurück in das Standardreferenzsystem unumgänglich. An dieser Stelle ist es nun möglich experimentelle Befunde oder Resultate der vertrauenswürdigsten Methode bezüglich Bandstruktur und molekularer Systeme zur Referenzierung zu verwenden. In dieser Arbeit stehe nun die HSE Daten im Vordergrund.

Die Bestimmung der Defektkonzentrationen geschieht über die in Abschnitt 5.5 vorgestellte Methode. Als MOVPE-Wachstumstemperatur sind hierbei 1275 K angesetzt worden[9]. Das chemische Potential des Wasserstoffes entspricht dieser Wachstumstemperatur bei einem Partialdruck von 0.1 bar. Das chemische Potential des Stickstoffs ist so justiert worden, dass die gemessene Wasserstoffkonzentration für eine Magnesiumkonzentration von $[0.7 - - - 0.9] \times 10^{18} \text{ cm}^{-3}$ reproduziert wird.

Die Ergebnisse werden unter der Annahme des thermodynamischen Gleichgewichtes diskutiert. Anhand der Boltzmannverteilung

**Abb. 7.16:**

Relative Defektkonzentration basierend auf der HSE Defektenergetik in GaN:Mg bei unterschiedlichen Mg Konzentrationen. Die einzelnen Defektkonzentrationen wurden mit ihrer Ladung multipliziert und werden im Plot aufsummiert dargestellt. Der rote Bereich gibt hierbei den Störstellenwasserstoff an. Der grüne Bereich repräsentiert den Wasserstoff, der an die Stickstoffvakanz bindet. Der blaue und gelbe Bereich zeigt die ladungsmultiplizierte Konzentration freier Stickstoffvakancen im Ladungszustand +1 und +3. Der mit h^+ indizierte braune Bereich repräsentiert die freien Löcher. Die schwarze Linie zeigt die sich ergebene relative Wasserstoffkonzentration.

$$c(X^q) = N_0 N_{\text{conf}} \exp\left(-\frac{E^0[X^q] + q\epsilon_{\text{Fermi}}}{k_b T}\right) \quad (7.15)$$

lässt sich schon eine erste Aussage über den Effekt steigender Magnesiumkonzentration treffen: Damit die Anzahl der negativen Magnesiumdefekte steigen kann muss zur Kompensation auch die Konzentration der positiven Defekte (oder auch freien Löcher) steigen. Dies ist jedoch nur durch eine Absenkung des Fermi-niveaus ϵ_{Fermi} im System möglich. Exakt diesen Trend kann man in Abbildung 7.15 erkennen. Das Fermi-niveau sinkt kontinuierlich bis es bei einer Magnesiumkonzentration von ca. $5 \times 10^{19} \text{ cm}^{-3}$ das $[\text{MgV}_\text{N}]^{0/2+}$ Umladungsniveau erreicht und dort fixiert wird.

Die relative Konzentration von Wasserstoff, der Stickstoffvakancen und der Löcher kann der Abbildung 7.16 entnommen werden. Um die Kompensation zu zeigen,

ist dort die relative Konzentration mit der Ladung des entsprechenden Defektes multipliziert worden. Um zwischen Stickstoffvakanz und Störstellen leicht zu unterscheiden bezeichnen wir an dieser Stelle jeglichen Störstellenwasserstoff — das sind alle H^+ -Defekte sowie $[Mg_{Ga}H]$ -Komplexe — mit H^+ , alle Stickstoffvakanz — allein oder gepaart mit Magnesium — mit $V_N^{+/3+}$, sowie die gemischten mit H_N^{2+} . Es ist zu erkennen, dass über den gesamten Konzentrationsbereich die dominante Kompensation durch die Wasserstoffstörstellen geleistet wird. Nur ein geringer Anteil (ca. 20–30%) der Magnesiumakzeptoren werden durch andere Defekte kompensiert. Freie Ladungsträger in Form von Löchern spielen keine Rolle.

Im Konzentrationsbereich von 10^{17} – 10^{18} cm^{-3} wird diese Sekundärkompensation von den einfach geladenen Stickstoffvakanz übernommen. Ab einer Konzentration von 10^{19} cm^{-3} beginnen die dreifach geladenen Stickstoffvakanz etwas mehr Bedeutung zu gewinnen. Es bilden sich zudem die Defekte der H_N^{2+} -Klasse. Dieser Defekt ist zweifach positiv geladen und vermindert die Notwendigkeit freien Wasserstoff in das Material einzubauen um die Akzeptoren zu passivieren. Allerdings besitzt dieser Defekt Komplex eine um ca. 0.5 eV höhere Barriere um den Wasserstoff freizusetzen[101] und sollte unter diesem Gesichtspunkt thermisch schwieriger auszutreiben sein. Im Bereich hoher Magnesiumkonzentrationen werden fast 35% der Akzeptoren durch stickstoffvakanzartige Defekte kompensiert. Die reinen Wasserstoffstörstellen weichen von ihrem Maximum von knapp 70% der Kompensation zurück auf 65%. Die Gesamtwasserstoffkonzentration sinkt aber kaum merklich.

Die Rechnungen liefern eine qualitative Reproduktion des experimentellen Befundes und liefern einen möglichen Mechanismus für den verminderten Wasserstoffeinbau. Durch das absinkende Fermi-niveau findet eine Umladung der Stickstoffvakanz aus den einfach geladenen Zustand in den dreifach geladenen statt. Dadurch verdrängen die Vakanz den Wasserstoff geringfügig aus seiner Rolle als Passivator der Magnesiumakzeptoren zur Erreichung der Ladungsneutralität. Im Gegensatz zu den Rechnungen zeigt sich der experimentelle Übergang aber deutlich schärfer und dominanter, sodass nun im folgenden eine weitere Untersuchung der Ergebnisse stattfindet um das Modell zu erweitern und die experimentellen Befunde besser zu repräsentieren.

7.4.8. Kopplung der chemischen Potentiale

In den im vorherigen Abschnitt vollzogenen Überlegungen waren keinerlei Einschränkungen bezüglich des Wertebereiches der chemischen Potentiale enthalten. Damit vernachlässigt man jedoch die Bildung neuer stabiler Phasen, welche bei der Justierung der Potentiale auftreten können. In Abbildung 7.17 ist die Variation des chemischen Potentials für Magnesium im Hinblick auf die Magnesiumkonzentration gegeben. Um die Magnesiumkonzentration in GaN zu erhöhen, muss das chemische Potential sukzessive erhöht werden. Dabei durchschreitet es jedoch

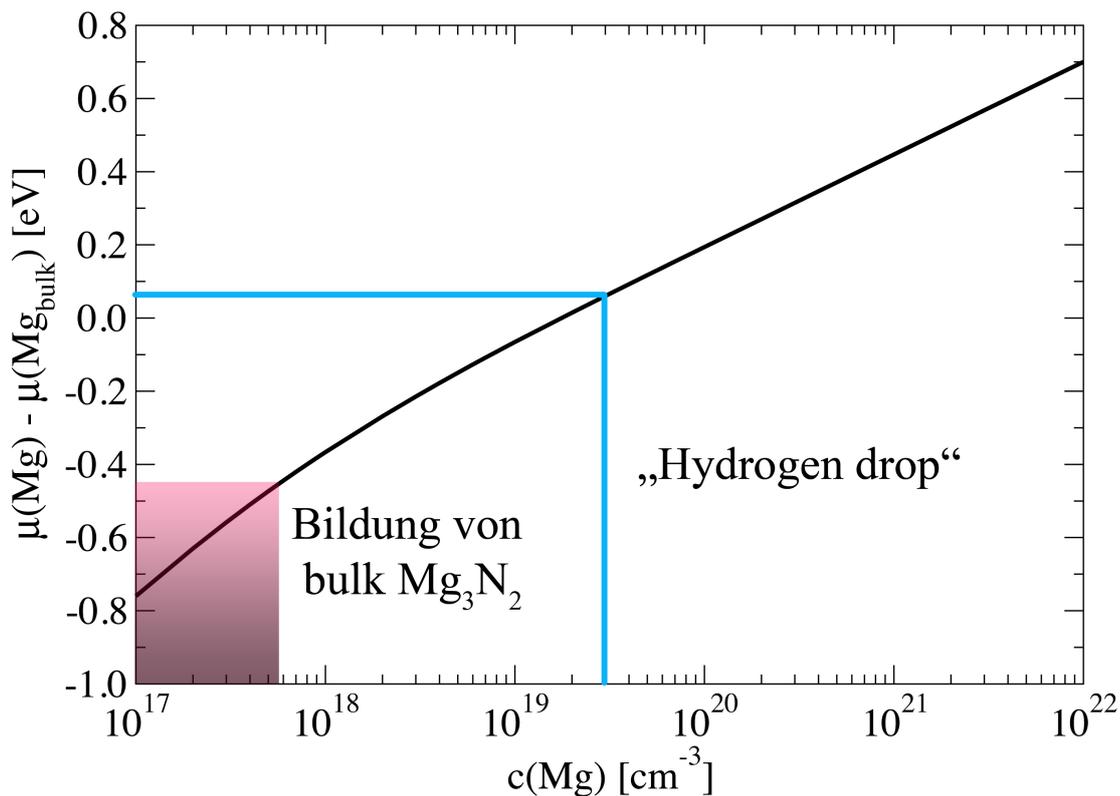


Abb. 7.17:

Notwendiges chemisches Potential zum Einstellen der Magnesiumkonzentration in GaN basierend auf der HSE Defektenergetik. Dargestellt sind weiter der theoretisch berechnete Bildungsbereich von Mg_3N_2 , sowie das chemische Potential, bei welchem der „Hydrogen Drop“ auftaucht.

die Stabilitätsgrenzen von Mg_3N_2 und Mg (vgl. Tabelle 7.15). An dieser Stelle erkennt man deutlich die Schwäche des rein thermodynamischen Bildes. Nach dieser Aussage würde Magnesium ab einer Konzentration von $7 \cdot 10^{17} \text{ cm}^{-3}$ in der magnesiumnitridartigen Phase vorliegen und ab einer Konzentration von $2 \cdot 10^{19} \text{ cm}^{-3}$ als reine Metallphase. Temperatureffekte können eine mögliche Erklärung sein, weswegen GaN dennoch die stabile Phase bleibt. Eine Abschätzung der thermischen Energie (Phasenumwandlung und Wärmekapazität) gibt hier einen Spielraum von bis zu 0.4 eV vor. Experimentell zersetzt sich reines Magnesiumnitrid bei den Wachstumstemperaturen und Magnesium liegt in einer flüssigen Phase vor. Auch der Einfluss kinetischer Effekte kann hier eine Rolle spielen. An dieser Stelle lassen wir uns folglich von den experimentellen Befunden führen um einen interessanten Aspekt zu motivieren.

Die Bildung eines magnesiumnitridartigen Filmes ist experimentell anhand der Inversionsdomänen (vgl. Abschnitt 7.1) belegt. Unter der Berücksichtigung von kinetischen und thermischen Effekten ist es durchaus möglich, dass die Bildungsenergie dieser Inversionsdomänengrenzphase (IDB- Mg_3N_2) vergleichbar ist zu der

	Bildungsenergie [eV]	
	theoretisch	experimentell
GaN	-1.533	-1.625 ^a
Mg ₃ N ₂	-4.429	-4.773 ^b
$\mu(\text{Mg}_3\text{N}_2) - \mu(\text{Mg})$	-0.454	-0.508

^a[107], ^b[108]

Tab. 7.15:

Bildungsenergien von GaN und Mg₃N₂ bezogen auf das entsprechende Metall(bulk) und N₂ in der Gasphase. Die theoretischen Werte wurden mittels PAW-LDA bestimmt. Aus den Bildungsenergien lässt sich die Differenz des chemischen Potentials für Mg gemäß $\mu(\text{Mg}_3\text{N}_2) - \mu(\text{Mg}) = (E^f[\text{Mg}_3\text{N}_2] - 2E^f[\text{GaN}])/3$ berechnen.

Energie, an welcher die Wasserstoffkonzentration zusammenbricht. Im weiteren identifizieren wir deswegen die Bildungsenergie von IDB-Mg₃N₂ mit der Energie des Zusammenbruchs der Wasserstoffkonzentration. Diese Inversionsdomänen sind im Wachstum zu vermeiden. Dies kann bei steigenden Magnesiumpotential nur durch eine Absenkung des Stickstoffpotentials geschehen. Es ist folglich notwendig das chemische Potential des Stickstoffes an das des Magnesiums zu koppeln, um der Bildung der Inversionsdomänen entgegenzuwirken. Die Wirkung dieser Kopplung wird nun untersucht und mit dem experimentellen Befund verglichen.

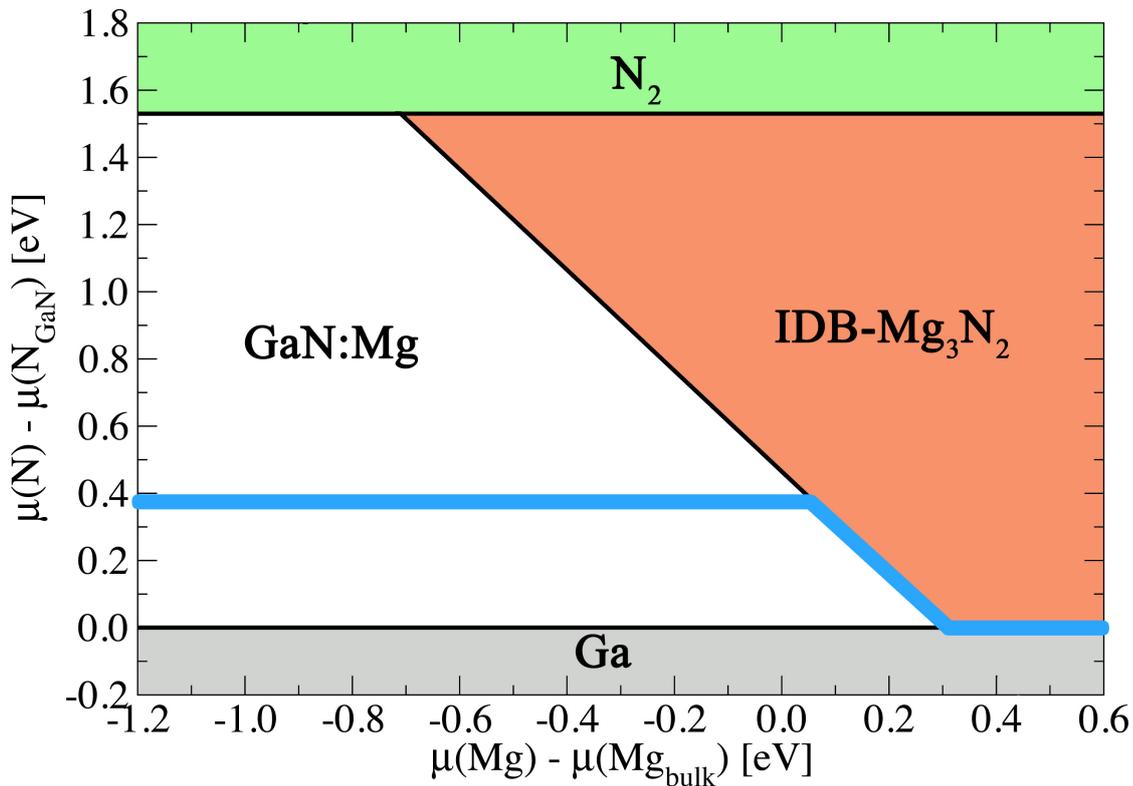
Im Referenzsystem GaN und Magnesiumnitrid sind die gekoppelten Relationen

$$\mu(\text{Ga}) \leq 0 \text{ Bildung von Ga (bulk)} \quad (7.16)$$

$$\mu(\text{Ga}) + \mu(\text{N}) \geq 0 \text{ Bildung von GaN} \quad (7.17)$$

$$3\mu(\text{Mg}) + 2\mu(\text{N}) \leq 0 \text{ Bildung von Mg}_3\text{N}_2 \quad (7.18)$$

für die chemischen Potentiale zu lösen. Damit ist eine Abhängigkeit des chemischen Potentials für Stickstoff von dem chemischen Potential des Magnesiums wie in Abbildung 7.18 gegeben. Für geringe Magnesiumkonzentrationen stünde dem Experimentator die gesamte Variationsbreite des chemischen Stickstoffpotentials zur Verfügung. Mit steigender Magnesiumkonzentration muss die Bildung von dem am Rand der Inversionsdomänen auftauchenden filmartigen Magnesiumnitrid unterdrückt werden, so dass die Variationsbreite mehr und mehr eingeschränkt wird und das chemische Potential des Stickstoffes abgesenkt werden muss. Dies hat deutliche Folgen für die Verteilung der Konzentrationen wie Abbildung 7.19 zeigt. Bei einer Magnesiumkonzentration von $3 \times 10^{19} \text{ cm}^{-3}$ bricht nun die relative Wasserstoffkonzentration zu Gunsten der Bildung von Stickstoffvakanz deutlich ein. Die Übereinstimmung mit dem Experiment ist qualitativ und quantitativ gut. Der Grund für dieses Verhalten ist die Begünstigung des Stickstoffvakanz durch das sinkende chemische Stickstoffpotential. Der Abbildung 7.19 ist ein

**Abb. 7.18:**

Phasendiagramm des GaN Wachstums in Abhängigkeit der chemischen Potentiale für Stickstoff und Magnesium. Die jeweiligen Bereiche geben die jeweils stabile Phase an. Die hellblaue Linie zeigt einen möglichen Einstellungspfad der chemischen Potentiale für das Wachstum von Magnesium dotierten GaN mit unterschiedlichen Magnesiumkonzentrationen.

klarer Zugewinn der positiv geladenen Stickstoffvakanz und des Wasserstoff-Stickstoffvakanzkomplexes zu entnehmen.

7.4.9. Zusammenfassung

In diesem Kapitel ist die Limitierung der p -Dotierbarkeit von GaN mit Magnesium als Dotiermaterial auf Basis von DFT Berechnungen von Defektbildungsenergien untersucht worden. Ausgangspunkt der Untersuchung stellt das anerkannte Bild dar, welches die gute Löslichkeit von Magnesium in GaN mit der Bildung von [MgH]-Defektkomplexen erklärt. Experimentelle Befunde aber zeigen, dass mit zunehmender Magnesiumdotierung die relative Wasserstoffkonzentration abnimmt. Ist diese Konzentration für moderate Dotierungen immerhin noch nahe bei 70%, so fällt sie für hoch dotierte Proben auf einen Wert von ca. 20% ab. Eine thermische Austreibung des Wasserstoffes und somit eine Aktivierung der passivierten Magnesiumakzeptoren ist dann nicht mehr möglich. Dies steht im klaren Widerspruch zu

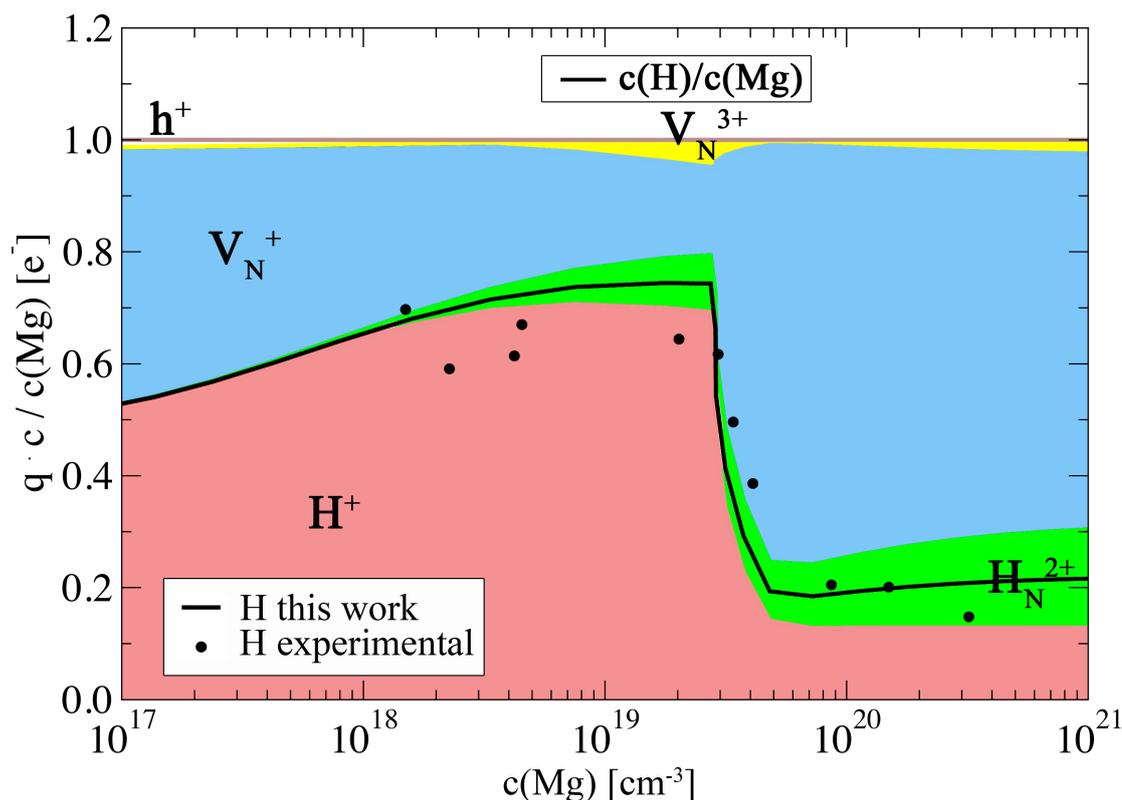


Abb. 7.19:

Relative Defektkonzentration basierend auf der HSE Defektenergetik in GaN:Mg bei unterschiedlichen Mg-Konzentrationen und variierenden chemischen Stickstoffpotential. Die einzelnen Defektkonzentrationen sind mit ihrer Ladung multipliziert und werden im Plot aufsummiert dargestellt. Der rote Bereich gibt hierbei den Störstellenwasserstoff an. Der grüne Bereich repräsentiert den Wasserstoff, der an die Stickstoffvakanz bindet. Der blaue und gelbe Bereich zeigt die ladungsmultiplizierte Konzentration freier Stickstoffvakancen im Ladungszustand +1 und +3. Der mit h^+ indizierte braune Bereich repräsentiert die freien Löcher. Die schwarze Linie zeigt die sich ergebene relative Wasserstoffkonzentration.

dem allgemein anerkanntem Bild, dass Magnesium und Wasserstoff vornehmlich als [MgH]-Defektkomplex im GaN-Wirtssystem eingebaut werden. Dieses Bild erfährt durch die in dieser Dissertation getätigten Überlegungen eine Revision. Dazu wurden zunächst die bei hohen Magnesiumkonzentrationen experimentell nachgewiesenen Inversionsdomänen diskutiert. In den Inversionsdomänengrenzen bildet sich eine magnesiumnitridartige Zwischenschicht, die auf ihre Eigenschaften der Wasserstoffbindung untersucht wurde. Da eine vollständige Modellierung der Inversionsdomänen auf DFT-Niveau sehr aufwendig und rechenintensiv ist, wurden zunächst die Eigenschaften von reinem Magnesiumnitrid in der bulk-Phase bestimmt, um auf chemische Trends für die Inversionsdomänengrenze zu schließen. Die Eigenschaften der Wasserstoffadsorption sind hier jedoch minimal im Vergleich zu den typischen Wasserstoffkonzentrationen, die in magnesiumdotiertem GaN

während des Wachstums herrschen. Deswegen konnten die magnesiumnitridartigen Inversionsdomänengrenzen als relevante Modifikatoren der Wasserstoffkonzentration ausgeschlossen werden.

Da die Inversionsdomänengrenzen nur einen schwachen direkten Einfluss auf die Wasserstoffkonzentration ausüben, ist der Ursprung des Phänomens in der Defektenergetik des Wirtssystems zu suchen. Zwar sind in der Literatur schon komplette Datensätze zur GaN-Defektenergetik vorhanden, allerdings haben jüngste Entwicklungen es möglich gemacht die Genauigkeit der berechneten Bildungsenergien deutlich zu verbessern. Hier ist die Superzellenkorrektur für geladene Defekte zu nennen, welche die im Superzellenansatz der DFT ungewollte Defekt-Defekt-Wechselwirkung geladener Defekte korrigiert, sowie die Verfügbarkeit besserer Austausch- und Korrelationsfunktionale wie HSE, welche in einem weitaus geringerem Maße vom Bandlückenproblem betroffen sind. Eine vergleichende Defektenergetik wichtiger Defekte in p -GaN zwischen den Literaturdaten und den in dieser Dissertation getätigten Rechnungen mit LDA und HSE zeigt deutliche Abweichungen, die für sich genommen das Vertrauen in die Berechnung von Defektenergetiken mittels DFT zunächst einmal erschüttern. In dieser Dissertation ist es aber gelungen Korrekturen zu entwickeln, welche die Literaturdaten, sowie die hier mit LDA berechneten Defektenergien mit dem HSE Datensatz in Einklang bringen. Dafür ist eine konsequente Anwendung der Superzellenkorrektur geladener Defekte notwendig, sowie die Umreferenzierung der Fermienergie von der Valenzbandkante hin zu den Umladungsniveaus tiefer Defekte. Mit diesen Korrekturen variieren die Bildungsenergien der betrachteten Defekte in einem Bereich von 0.2 eV je nach Defektart und liefern somit vergleichbare Ergebnisse.

Für die Untersuchung der Defektkonzentrationen in Abhängigkeit zur Magnesiumkonzentration sind die HSE Ergebnisse verwendet worden. Bei steigender Magnesiumkonzentration sinkt das Fermienergielevel bis es durch das Umladungsniveau des $[\text{MgV}_\text{N}]$ -Defektkomplexes fixiert wird. Hier laden V_N^+ Defekte in V_N^{3+} Defekte um. Durch die Erzeugung der dreifach positiv geladenen Defekte verliert der Wasserstoff als positiver Defekt geringfügig an Notwendigkeit. Die relative Wasserstoffkonzentration sinkt minimal. Aus Sicht der Defektenergetik allein kann die Verminderung der relativen Wasserstoffkonzentration folglich nicht erklärt werden. Die Defektenergetik ist in Einklang mit dem anerkannten Bild des Magnesiumeinbaus und den anschließenden thermischen Austreibung des Wasserstoffes.

Eine detaillierte Analyse des chemischen Magnesiumpotentialverlaufes hingegen zeigt, dass zur Einbringung des Magnesiums der Bereich der Bildung von Magnesiumnitrid überschritten wird. Experimentell wird tatsächlich Magnesiumnitrid in Form von Inversionsdomänengrenzen beobachtet, deren Bildung generell entgegengewirkt werden muss. Dies kann in der theoretischen Modellierung durch eine Kopp-

lung des chemischen Stickstoffpotentials an das chemische Magnesiumpotential geschehen, um in der Summe unterhalb der Bildungsenergie von Magnesiumnitrid und insbesondere des magnesiumnitridartigen Filmes in den Inversionsdomänengrenzen zu bleiben. Da die Bildungsenergie der magnesiumnitridartigen Grenzschicht nicht bekannt ist, wurde diese über das chemische Magnesiumpotential an der experimentell bestimmten Stelle des Wasserstoffeinbruchs berechnet. Dabei wird angenommen, dass der dieser relative Konzentrationseinbruch eine indirekte Konsequenz der Inversionsdomänenbildung ist. Reines Magnesiumnitrid besitzt jedoch eine theoretische Bildungsenergie die ca. 0.45 eV unter der hier angenommenen liegt. Diese Differenz kann jedoch zum einen in strukturellen Verspannungen begründet sein, da GaN das eigentliche Wirtssystem bildet. Weiter gehen in die Betrachtung keine Temperatureffekte oder kinetischen Effekte ein, die eine Verschiebung des Stabilitätsbereiches von Magnesiumnitrid in höhere Konzentrationsbereiche durchaus zulassen. Eine Berechnung der Defektkonzentrationen unter diesen Randbedingungen offenbart eine scharfe und signifikante Verminderung der relativen Wasserstoffkonzentration vergleichbar zu den experimentellen Daten. Die Passivierung der Magnesiumakzeptoren wird nun durch Stickstoffvakanzan übernommen, welche durch das Absenken des chemischen Stickstoffpotentials dem Wasserstoff gegenüber bevorzugt sind. Der Aspekt der Vermeidung der Inversionsdomänen erweitert folglich das ursprüngliche Bild der Magnesiumdotierung:

Die *p*-Dotierbarkeit von GaN mittels Magnesium lässt sich nicht mittels Wasserstoffkodotierung in beliebige Höhen treiben, sondern wird durch die Gratwanderung zwischen Stickstoffvakanzbildung und Inversionsdomänenbildung limitiert.

8. Zusammenfassung

Die Wissenschaft fängt eigentlich erst da an, interessant zu werden, wo sie aufhört.

Justus von Liebig, deutscher Chemiker, 1803 - 1873

Die vorliegende Dissertation hat es sich zum Ziel gesetzt Ursachen der aktuellen Limitierungen in der p -Dotierbarkeit von Galliumnitrid (GaN) mit theoretischen Mitteln zu ergründen und zu erklären. Ausgangspunkt stellte dabei der experimentell beobachtete Einbruch der relativen Wasserstoffkonzentration in hoch mit Magnesium dotierten Proben dar. Dieser Befund steht im Widerspruch zu dem anerkannten Bild, dass sich die Magnesiumlöslichkeit in GaN deutlich erhöht, wenn mit Wasserstoff kodotiert wird und Magnesium als neutraler [MgH]-Defektkomplex im Wirtssystem eingebaut wird. Der Einbruch der relativen Wasserstoffkonzentration deutet auf einen weiteren Kompensationsmechanismus hin. Messungen mittels Sekundärionenspektroskopie konnten aber keinen geeigneten Kandidaten feststellen, der die Rolle des Wasserstoffes übernimmt.

An dieser Stelle erlaubte es das computergestützte Materialdesign durch die Verwendung von *ab initio* Methoden, wie der Dichtefunktionaltheorie (DFT), über die Berechnung von Defektenergetiken die Informationen über das Materialsystem zu gewinnen, welche dem Experimentator in der Regel verschlossen bleiben. Auch wenn die in dieser Dissertation durchgehend verwendete Annahme eines thermodynamischen Gleichgewichtes die kinetischen Effekte vernachlässigt, so erlaubt sie dennoch wichtige Trends abzuleiten und dem Experimentator neue Zielsetzungen für notwendige weitere Experimente aufzuzeigen.

Da der beobachtete Einbruch der relativen Wasserstoffkonzentration mit der Bildung von Inversionsdomänen begleitet wird, rückten diese zunächst in den Fokus der theoretischen Betrachtung. Eine Inversionsdomäne kann Ausmaße von einigen Kubiknanometern und mehr annehmen. Dies bedeutet eine herausfordernde Modellierung eines Systems von 10.000 bis 100.000 Atomen und entsprechender Rechenzeit. Um relativ zügig erste Aussagen zu erhalten, ist ein zweigleisiger Ansatz gewählt worden: Eine effiziente Modellierung von Superzellen in dieser Größe ist auf Basis der DFT-Verfahren nicht möglich, deswegen sind sie durch sogenannte ‘coarse-grained’ Verfahren zu ersetzen. Dabei soll aber nach Möglichkeit die Genauigkeit der Aussagen mit denen der DFT konkurrieren können, weswegen der Blick auf die Konstruktion akkurater atomzentrierter Basissätze, den Quamols, fiel. Auf der anderen Seite war über das Material, welches die Inversionsdomänengrenze bildet — Magnesiumnitrid — in der Literatur recht wenig bekannt. Eine komplette Defektenergetik dieses Materials in der Antibixbyitephase sollte Aufschluss darüber geben, inwiefern Magnesiumnitrid die Wasserstoffkonzentration

in GaN modifizieren kann.

Atomzentrierte Basissätze leiden unter dem Manko, dass sie im Vergleich zu ebenen Wellen keine systematisch kontrollierbare Genauigkeit im Sinne einer Konvergenz aufweisen. Die Quamols wurden mit dem Anspruch entwickelt akkurat und transferabel zu sein und zudem durch die Verbindung zu einer zu Grunde liegenden Ebenen-Wellen-Basis die systematisch kontrollierbare Genauigkeit dieser zu besitzen. Um besonders die Transferabilität zu gewährleisten wurde die Beibehaltung der sphärischen Symmetrie gefordert. Da zudem nicht abzuschätzen ist, inwieweit ein zu Grunde liegender Basissatz, in welchen Atomorbitale generell entwickelt werden können, Einschränkungen in der Flexibilität bedeutet, sind die Quamols basissatzfrei als numerische Funktionen definiert worden. Als Optimierungskriterium wurde die von Sanchez-Portal 1995 eingeführte Spillage gewählt: die Basisraumdifferenz zwischen Zuständen in der Ebenen-Wellen-Darstellung und den Zuständen in der Atomorbital-Darstellung. Die physikalischen Vorteile der basissatzfreien Repräsentation stehen hierbei jedoch einer deutlich komplexeren Numerik gegenüber, so dass die in allen früheren Arbeiten genutzten analytische Parametrisierungen — die allerdings nur bedingt flexibel sind — nicht zur Verfügung standen. Die effiziente Berechnung von Überlappintegralen zwischen Zuständen in der Ebenen-Wellen-Darstellung und der Quamol Repräsentation stellte eine Herausforderung dar, die mittels des geschickten Einsatzes von Splinefunktionen gelöst werden konnte. Den Daten dieser Dissertation kann entnommen werden, dass mit den Quamols ein akkurater und transferabler Basissatz geschaffen wurde, der im Gegensatz zu allen früheren Methoden vollkommen flexibel ist.

Die Identifizierung und Berechnung der Defektenergetik wichtiger Punktdefekte in Magnesiumnitrid erfolgte erstmals auf DFT-Niveau unter Verwendung normerhaltender Pseudopotentiale und des PBE Korrelation- und Austauschfunktionals. Unter Anwendung des thermodynamischen Gleichgewichtes konnten Defektkonzentrationen bestimmt werden, welche, zusammen mit der geringen Konzentration von dem in Inversionsdomänen gebundenen Magnesiumnitrid, eine signifikante Modifikation Wasserstoffkonzentration in GaN:Mg nicht erklären können. Von einer kompletten Modellierung der Inversionsdomäne wurde somit Abstand genommen, da auch die strukturellen Änderungen von der Antibixbyitephase zur Inversionsdomänengrenzphase keine Änderung der Wasserstoffkonzentration um mehrere Größenordnungen bedeutet hätten. Die Ursache der modifizierten Wasserstoffkonzentration muss also im Galliumnitrid selbst liegen.

GaN ist ein deutlich intensiver diskutiertes Materialsystem als Magnesiumnitrid. Hier lassen sich schon komplette Defektenergetiken in der Literatur finden. Jüngste Entwicklungen, wie die Ladungskorrektur für geladenen Defekte in Superzellen, sowie die bessere Verfügbarkeit des Hybridfunktionals HSE, welches Bandlücken

in Übereinstimmung mit den experimentellen Befunden berechnet und somit die DFT von ihrem Bandlückenproblem befreit, motivierten die Frage, inwiefern deren Anwendung die Genauigkeit der Defektenergetiken weiter steigern kann. Die Gegenüberstellung der mit diesen neuen Methoden berechneten Defektenergetiken mit den Literaturdaten war jedoch sehr ernüchternd. Weder quantitativ noch qualitativ ließ sich eine Übereinstimmung ausmachen. Dies stellte natürlich die Verlässlichkeit von mit DFT berechneten Bildungsenergien generell in Frage, die für eine akkurate Beschreibung von Defektkonzentrationen unumgänglich sind. Eine Validierung der Defektenergetiken musste vorgenommen werden. Dazu wurden die Formeln zur Bestimmung von Bildungsenergien analysiert, um Korrekturterme zu entwickeln, welche die Schwächen der alten Funktionale abmildern sollten. In dieser Dissertation ist es gelungen die qualitative und innerhalb von 0.3 eV auch quantitative Übereinstimmung der mit den unterschiedlichen Methoden berechneten Defektenergien zu zeigen. Wichtig hierfür ist die konsequente Anwendung der Ladungskorrektur für geladene Defekte in Superzellen, sowie eine physikalisch motivierte und transparente Umreferenzierung des Fermi-niveaus von der, durch das Bandlückenproblem fehlerhaften, Valenzbandkante zu dem funktionalübergreifenden gut beschriebenen Umladungsniveau der Stickstoffvakanz. Mit diesen Erkenntnissen wird es möglich, einen Großteil der Fehler in der DFT zu identifizieren und durch geeignete Transformationsvorschriften zu beheben. Durch die Verwendung neuester Techniken im Bereich der Dichtefunktionaltheorie wie der Verwendung des HSE Hybridfunktionals, sowie die Anwendung der Superzellenkorrektur für geladene Defekte konnte zudem die Genauigkeit signifikant verbessert werden und eine verlässliche Datenbasis geschaffen werden um die Fragestellung zu untersuchen.

Mit den validierten Defekt-daten wurden unter Annahme des thermodynamischen Gleichgewichtes Konzentrationsprofile erstellt. Diese Profile zeigen, dass Stickstoffvakanz in vergleichbaren Konzentrationen zum Wasserstoff auftreten und auf Grund ihrer Ladung die Kompensation des Magnesiums übernehmen können. Allerdings sind die berechneten Konzentrationsänderungen nach wie vor zu gering, um die experimentell beobachteten Verhältnisse zu erklären. Erst die Berücksichtigung der Phasenseparation GaN/Mg₃N₂ und eine entsprechende Kopplung der chemischen Potentiale untereinander, führte zu einer quantitativen Übereinstimmung der theoretischen Resultate mit den experimentellen Daten. Die Gradwanderung zwischen einem ausreichendem Stickstoffangebot zur Vermeidung der Vakanzbildung und einem Unterangebot an Stickstoff um die Phasenseparation Galliumnitrid/Magnesiumnitrid bei hohen Magnesiumkonzentrationen zu verhindern, ist schlussendlich als eigentliche Limitierung der *p*-Dotierbarkeit identifiziert worden.

Die in dieser Dissertation geschaffenen Quamols sind in ihrem Potential bei weitem

noch nicht erschöpft. Ihre Möglichkeiten bei der Verwendung von *coarse-grained* Methoden sind dabei nur ein möglicher weiterer Entwicklungsstrang der besprochen werden könnte. Schon jetzt haben sich die Quamols zu analytischen Zwecken bewährt. Sei es bei der Mullikenschen Populationsanalyse oder der orbitalaufgelösten Bandstrukturanalyse. Dadurch, dass die Quamols den entsprechenden Unterraum praktisch verlustfrei aufspannen, bilden sie einen systemnäheren Basissatz, als es die Orbitalfunktionen der Pseudoatome tun, die oftmals zur Analyse von Ebenen-Wellen-Rechnungen eingesetzt werden. Ein solcher Basissatz erlaubt weiter die orts-aufgelöste Analyse von Energiebeiträgen zur Gesamtenergie eines Systems. Eine Vielzahl von Fragestellungen kann mit Hilfe dieses Basissatzes angegangen werden und wird Bestandteil weiterer Forschung mit spannender Physik sein.

A. Mg_3N_2 in der idealen Antixbytestruktur

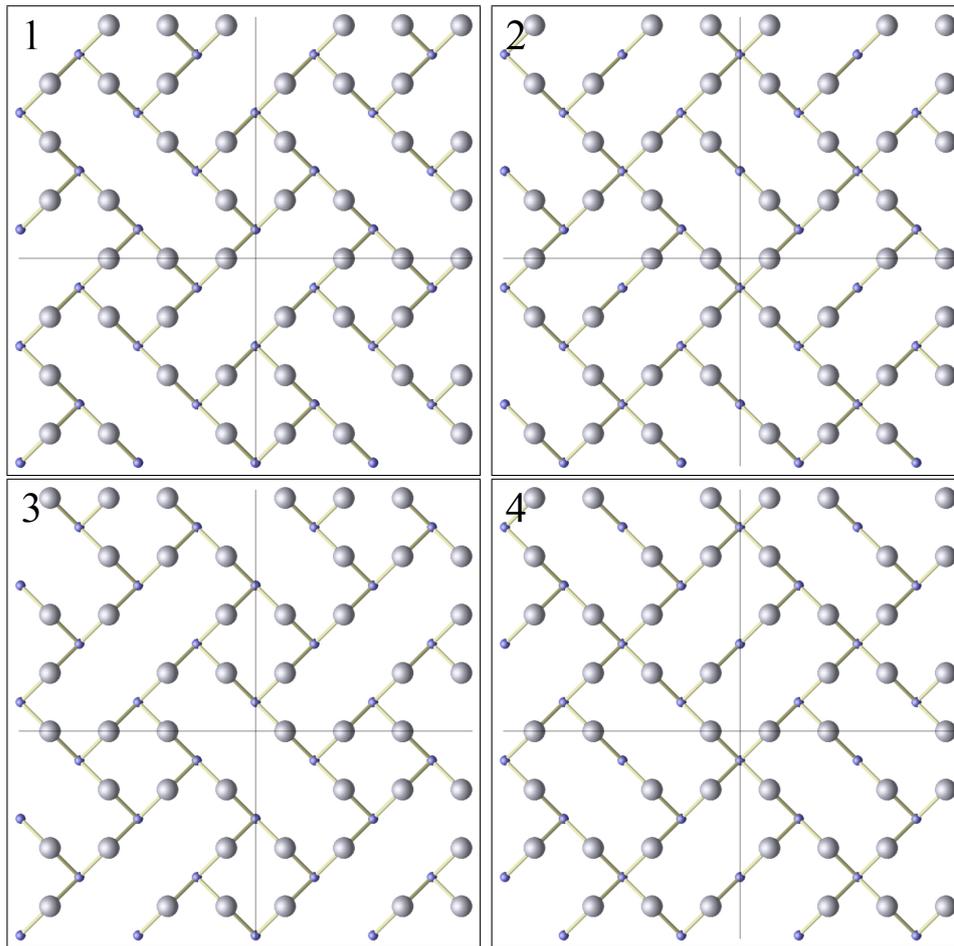


Abb. A.1:

Magnesiumnitrid in der idealen, nichtrelaxierten Antixbytestruktur. Die Struktur ist in vier Atomlage aufgeteilt, die die Superzelle bilden. Die kleinen Kugeln stehen dabei für die Positionen der Stickstoffatome, die großen Kugeln symbolisieren die Positionen der Magnesiumatome. Die einzelnen Lagen sind entlang der (100) und der (010) Richtung verdoppelt worden um ein besseres Verständnis für die Symmetrie der einzelnen Lagen zu geben.

Magnesiumnitrid kristallisiert in Antixbytestruktur. Diese Struktur besitzt in ihrer idealen Form eine grosse Ähnlichkeit zur Kalziumfluoridstruktur. Den Unterschied bilden die zur Erhaltung der richtigen Stöchiometrie einzubringenden Strukturvakanten. Abbildung A.1 zeigt die einzelnen Lagen der kubischen Superzelle. Wie hier zu sehen ist, ergeben sich die Lagen 3 und 4 aus Translationen der Lagen 1 und 2, so dass diese Superzelle auf eine primitive *fcc*-Zelle reduziert werden kann, die nur noch 40 Atome enthält. Eine entsprechende Struktur ist in [109] gegeben. Die $\mathbf{A}_{1...3}$ bezeichnen die Gittervektoren der kubischen Zelle. Die 32 Stickstoffatome $N_{1...32}$ bilden ein *fcc*-Untergitter. Die 48 Magnesiumatome $Mg_{1...48}$ besetzen die Tetraederlücken, wobei in jeder Lage vier dieser Lücken unbesetzt bleiben und die

sogenannten Strukturvakanz $V_{1\dots 16}$ bilden. Die Koordinaten dieser Strukturvakanz in der idealen Antixbytestruktur für Magnesiumnitrid lauten

$$\begin{aligned}
V_1 &= \frac{5}{8}\mathbf{A}_1 + \frac{1}{8}\mathbf{A}_2 + \frac{1}{8}\mathbf{A}_3 & V_2 &= \frac{3}{8}\mathbf{A}_1 + \frac{3}{8}\mathbf{A}_2 + \frac{1}{8}\mathbf{A}_3 \\
V_3 &= \frac{7}{8}\mathbf{A}_1 + \frac{5}{8}\mathbf{A}_2 + \frac{1}{8}\mathbf{A}_3 & V_4 &= \frac{1}{8}\mathbf{A}_1 + \frac{7}{8}\mathbf{A}_2 + \frac{1}{8}\mathbf{A}_3 \\
V_5 &= \frac{7}{8}\mathbf{A}_1 + \frac{1}{8}\mathbf{A}_2 + \frac{3}{8}\mathbf{A}_3 & V_6 &= \frac{1}{8}\mathbf{A}_1 + \frac{3}{8}\mathbf{A}_2 + \frac{3}{8}\mathbf{A}_3 \\
V_7 &= \frac{5}{8}\mathbf{A}_1 + \frac{5}{8}\mathbf{A}_2 + \frac{3}{8}\mathbf{A}_3 & V_8 &= \frac{3}{8}\mathbf{A}_1 + \frac{7}{8}\mathbf{A}_2 + \frac{3}{8}\mathbf{A}_3 \\
V_9 &= \frac{3}{8}\mathbf{A}_1 + \frac{1}{8}\mathbf{A}_2 + \frac{5}{8}\mathbf{A}_3 & V_{10} &= \frac{5}{8}\mathbf{A}_1 + \frac{3}{8}\mathbf{A}_2 + \frac{5}{8}\mathbf{A}_3 \\
V_{11} &= \frac{1}{8}\mathbf{A}_1 + \frac{5}{8}\mathbf{A}_2 + \frac{5}{8}\mathbf{A}_3 & V_{12} &= \frac{7}{8}\mathbf{A}_1 + \frac{7}{8}\mathbf{A}_2 + \frac{5}{8}\mathbf{A}_3 \\
V_{13} &= \frac{1}{8}\mathbf{A}_1 + \frac{1}{8}\mathbf{A}_2 + \frac{7}{8}\mathbf{A}_3 & V_{14} &= \frac{7}{8}\mathbf{A}_1 + \frac{3}{8}\mathbf{A}_2 + \frac{7}{8}\mathbf{A}_3 \\
V_{15} &= \frac{3}{8}\mathbf{A}_1 + \frac{5}{8}\mathbf{A}_2 + \frac{7}{8}\mathbf{A}_3 & V_{16} &= \frac{5}{8}\mathbf{A}_1 + \frac{7}{8}\mathbf{A}_2 + \frac{7}{8}\mathbf{A}_3.
\end{aligned}$$

B. Effektivmassenberechnung

Die effektiven Massen von Elektronen und Löchern sind notwendige Größen zur Berechnung freier Ladungsträgerkonzentrationen wie sie zum Beispiel im Lehrbuch von Kittel[26] angegeben werden

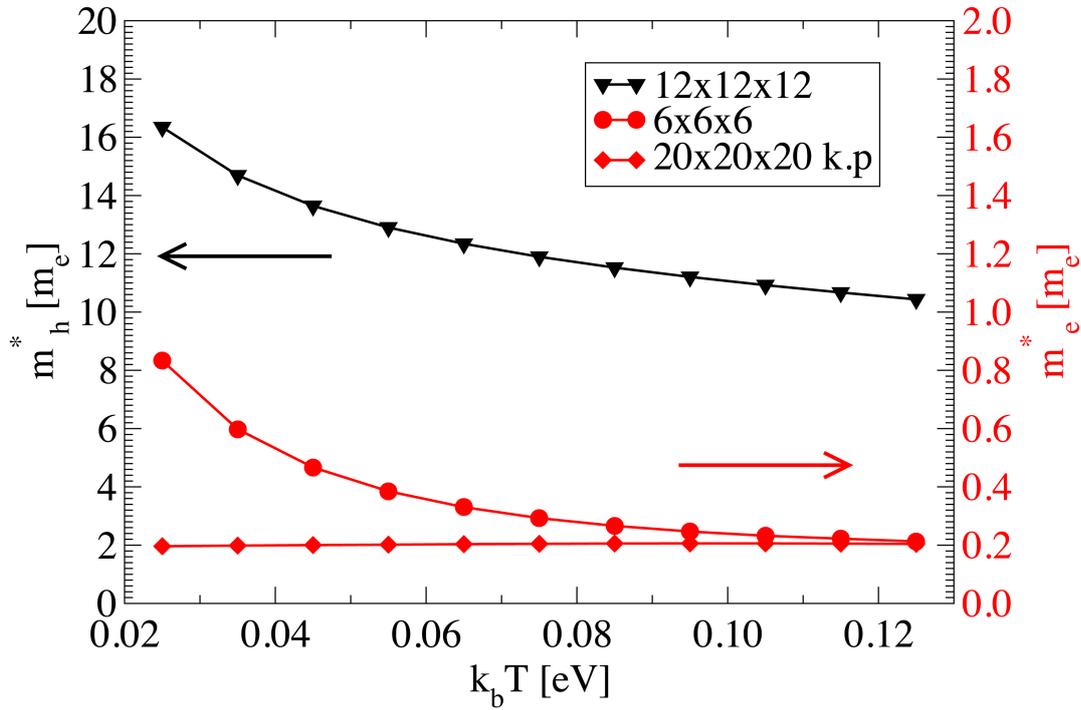
$$c_e = 2 \left(\frac{m_e^* k_b T}{2\pi \hbar^2} \right)^{3/2} e^{(E_{\text{Fermi}} - E_{\text{CBM}})/k_b T}, \quad (\text{B.1})$$

$$c_h = 2 \left(\frac{m_h^* k_b T}{2\pi \hbar^2} \right)^{3/2} e^{(E_{\text{VBM}} - E_{\text{Fermi}})/k_b T}. \quad (\text{B.2})$$

Die effektiven Massen können über Hallmessungen experimentell bestimmt werden. An dieser Stelle wird ein Verfahren vorgestellt, welches den Wert der effektiven Massen theoretisch abschätzt. Im Allgemeinen verwendet man als Definition der effektiven Masse m^*

$$\frac{1}{m^*} = \frac{1}{\hbar^2} \Delta_{\mathbf{k}} \epsilon(\mathbf{k}). \quad (\text{B.3})$$

$\epsilon(\mathbf{k})$ bezeichnet hierbei die Banddispersion, die erhalten wird, indem in der Umgebung des VBM oder des CBM ein parabolischer Fit vollzogen wird. Mit $\Delta_{\mathbf{k}}$ ist


Abb. B.1:

Effektive Massen der Löcher und der Elektronen in Magnesiumnitrid in Abhängigkeit der thermischen Energie. Die linke Skala bezieht die Lochmassen, die rechte die Elektronenmassen. Die Legende gibt das für die jeweilige Berechnung verwendete \mathbf{k} -Punkt Gitter wieder.

der Laplaceoperator bezüglich der \mathbf{k} -Vektoren bezeichnet. Über einen parabolischen Fit an die Bandstruktur kann die effektive Masse des Bandes am entsprechenden \mathbf{k} -Punkt abgeschätzt werden. Allerdings ist dieses Verfahren problematisch im Falle vieler flacher Bänder, wo der Fit individueller Bänder zu einer falschen Zustandsdichte führt. Besser ist es in einer solchen Situation die Zustandsdichte als Eingangsgröße zu verwenden, um die effektiven Massen abzuschätzen. Die energieabhängige Zustandsdichte $D(E)$ ist über die Beziehung

$$n(E_{\text{Fermi}}) = \int_{-\infty}^{\infty} f(E - E_{\text{Fermi}}) D(E) dE \quad (\text{B.4})$$

mit der Teilchenzahldichte n verknüpft. Die energieabhängige Funktion f stellt hierbei die elektronische Besetzung der Zustände dar und wird über die Fermi-Dirac Verteilung modelliert. In der numerischen Betrachtung wird die Zustandsdichte über

$$D(E) = \sum_{n,\mathbf{k}} w_{\mathbf{k}} \delta(E - \epsilon_{n\mathbf{k}}) \quad (\text{B.5})$$

genähert, wobei $\epsilon_{n\mathbf{k}}$ die Eigenenergien des Systems sind und die $w_{\mathbf{k}}$ die \mathbf{k} -Punkt Gewichtung der diskretisierten Brillouin Zone bezeichnet. Damit ergibt sich die Teilchenzahldichte numerisch zu

$$n(E_{\text{Fermi}}) = \sum_{n,\mathbf{k}} w_{\mathbf{k}} f(\epsilon_{n\mathbf{k}} - E_{\text{Fermi}}). \quad (\text{B.6})$$

Die folgenden Betrachtungen gelten für die Elektronen. Sie lassen sich aber in analoger Weise auf die Löcher übertragen. Unter der Annahme, dass die Energiedifferenz zwischen Fermienergie und Leitungsbandminimum deutlich größer ist als die thermische Energie

$$\frac{E_{\text{CBM}} - E_{\text{Fermi}}}{k_b T} \gg 1, \quad (\text{B.7})$$

kann die Fermi-Dirac-Verteilungsfunktion über die Boltzmannverteilung abgeschätzt werden. Nimmt man weiter an, dass elektronischen Anregungen hauptsächlich von Defektelektronen vollzogen werden, so ist die Anzahl der Valenzelektronen N_V konstant. Dies ist erfüllt wenn die Energiedifferenz von der Valenzbandkante zur Fermienergie deutlich größer ist als die Energiedifferenz der Fermienergie zum Leitungsbandminimum

$$E_{\text{Fermi}} - E_{\text{VBM}} \gg E_{\text{CBM}} - E_{\text{Fermi}}. \quad (\text{B.8})$$

Beide Bedingungen zusammen definieren einen sinnvollen Bereich für die Wahl des Fermienergie

$$\frac{1}{2}(E_{\text{VBM}} + E_{\text{CBM}}) \gg E_{\text{Fermi}} \gg E_{\text{CBM}} - k_b T. \quad (\text{B.9})$$

Auch die Wahl der thermischen Energie ist damit sinnvoll eingegrenzt. Zum einen muss sie gering genug sein um obige Bedingung zu erfüllen, zum anderen groß genug um ein grobes \mathbf{k} -Punkt Gitter zu rechtfertigen.

Setzt man $E_{\text{CBM}} = 0$ als Referenzenergie an, so ergibt sich die Anzahl der Leitungselektronen pro Volumenelement zu

$$n_c \approx e^{E_{\text{Fermi}}/k_b T} \underbrace{\int_0^{\infty} e^{-\epsilon/k_b T} D(\epsilon) d\epsilon}_{D_{\text{eff}}(k_b T)}. \quad (\text{B.10})$$

Für ein freies Elektronengas mit der effektiven Masse m_e^* kann die effektive Zustandsdichte D_{eff} analytisch zu

$$D_{\text{eff}}(k_b T) = \frac{2}{h^3} (2\pi m_e^* k_b T)^{3/2} \quad (\text{B.11})$$

berechnet werden. Mit dieser Beziehung und der Leitungselektronenzahl $N_c = n_c \Omega$ — mit Ω als Zellvolumen — ergibt sich die effektive Masse zu

$$m_e^* = \frac{2^{1/3} \hbar^2}{\Omega^{2/3} (k_b T)^{5/3}} (k_b T \ln N_c - E_{\text{Fermi}})^{2/3}. \quad (\text{B.12})$$

Die Fermienergie geht in die obige Gleichung als Parameter ein und ist zum jetzigen Zeitpunkt der Berechnungen nicht bekannt. Um sinnvolle Werte für die effektive Masse zu erhalten, wird in der Praxis das exponentielle Verhalten von Gleichung (B.10) über die Konstanz von $k_b T \ln N_c - E_{\text{Fermi}}$ abgefragt und die Fermienergie in diesem Bereich gesetzt.

Wir demonstrieren diesen Ansatz an Magnesiumnitrid, dessen Bandstruktur sich durch eine Vielzahl flacher Bänder nahe der Valenzbandkante auszeichnet. Abbildung B.1 zeigt die berechneten effektiven Massen in Abhängigkeit der thermischen Energie. Die Qualität der effektiven Masse hängt stark von der zu Grunde liegenden Zustandsdichte ab. Eine Approximation der Zustandsdichte über δ -Funktionen resultiert in Artefakten, wenn die Peakabstände vergleichbar oder sogar größer als $k_b T$ werden. Dies kann durch eine höhere Temperatur oder ein höheres \mathbf{k} -Punkt Gitter vermieden werden. Dieses Verhalten sieht man insbesondere bei der elektronischen effektiven Masse, wo das \mathbf{k} -Punkt Gitter ausgehend von einem $6 \times 6 \times 6$ zu einem $20 \times 20 \times 20$ Gitter mittels $\mathbf{k} \cdot \mathbf{p}$ Methode[110] interpoliert wurde. Die effektiven Lochmassen hingegen konvergieren schneller. Eine Erhöhung der Auflösung jenseits eines $12 \times 12 \times 12$ Gitters liefert keinerlei signifikante Änderung der Masse, was eine Folge der Dispersionsfreiheit dieser Zustände nahe der Valenzbandkante ist. Die starke Temperaturabhängigkeit ist eine Folge des Freien-Loch-Modells. Hier wird die wahre Zustandsdichte über eine Quadratwurzelfunktion modelliert. Im Falle von Magnesiumnitrid repräsentiert eine Stufenfunktion die Zustandsdichte in diesem Energiebereich besser. Ein solcher Ansatz liefert für die effektive Zustandsdichte $D_{\text{eff}} \propto k_b T$, was dann für die effektive Masse eine thermische Abhängigkeit $m_h^* \propto (k_b T)^{-1/3}$ bedeutet. Genau dieses Verhalten ist in Abbildung B.1 zu sehen. An dieser Stelle gehen wir dann den pragmatischen Weg: Für Berechnungen an spezifischen Temperaturen wählen wir die effektive Masse, die für die gegebene Temperatur die Zustandsdichte am besten repräsentiert.

C. Pulay Kräfte

Durch den Übergang zu lokalisierten Basisfunktionen bedarf das Hellmann-Feynman Theorem einer Erweiterung, welche Pulay[111] erstmals für die Hartree-Fock Theorie einführte und die nun im Folgenden diskutiert wird.

Gemäß dem Ehrenfest'schen Theorem lässt sich der Erwartungswert der Kraft als Ableitung des Erwartungswertes der Energie im Bezug auf die Atomkoordinaten

$$-\mathbf{F}_i = \nabla_{\mathbf{R}_i} E = \langle \phi | \nabla_{\mathbf{R}_i} H | \phi \rangle + 2\Re \langle \nabla_{\mathbf{R}_i} \phi | H | \phi \rangle. \quad (\text{C.1})$$

beschreiben. Der erste Ausdruck in der rechten Seite der Gleichung bezeichnet hierbei die Hellmann-Feynman Kraft. Der zweite Ausdruck wird als Pulay Kraft be-

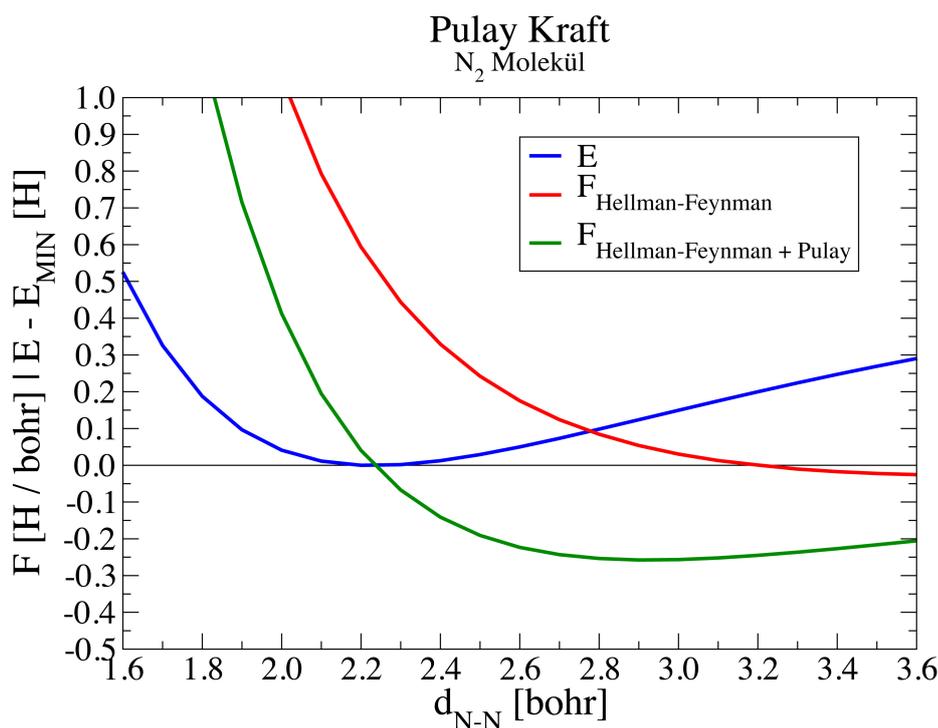


Abb. C.1:

Dissoziationskurve des Stickstoffdimers. In blau ist die abstandsabhängige Gesamtenergie aufgetragen. Die rote und die grüne Kurve zeigt den Betrag der berechneten Kraft einmal nach dem Hellman-Feynman Theorem und einmal unter Berücksichtigung der Pulaybeiträge. Als atomarer Basissatz sind das atomare s - und das atomare p -Orbital des Potentials verwendet worden. Da diese Orbitale nicht auf das Problem angepasst sind erhält man einen deutlichen Beitrag in der Pulay Kraft.

zeichnet und resultiert aus der Verschiebung der Basisfunktionen mit den Atomen. Diese Kraft ist eine Pseudokraft, aber notwendig zur akkuraten Beschreibung von Kräften in elektronischen Strukturberechnungen mit lokalisierten Basisfunktionen. Satoko[112] hat den obigen Ausdruck für die Pulay Kraft speziell für die Kohn-Sham Gleichungen der Dichtefunktionaltheorie zu

$$F_i^{\text{Pulay}} = -2\Re \left[\sum_{n,\tau} f_n c_{n\tau}^* \left\langle \frac{d\chi_\tau}{dR_i} \middle| \hat{H} - \epsilon_n \hat{S} \middle| \psi_n \right\rangle \right] \quad (\text{C.2})$$

berechnet. An dieser Gleichung lassen sich zwei Eigenschaften der Pulay-Kräfte sehen:

1. Die Pulay-Kräfte verschwinden, wenn das Residuum $\hat{H} - \epsilon_n \hat{S} | \psi_n \rangle$ verschwindet.
2. Die Pulay-Kräfte verschwinden, wenn die Basisfunktionen unabhängig von den Atomkoordinaten sind und somit $\frac{d\chi_\tau}{dR_i} = 0$ gilt.

Das erste Argument gilt für die exakten Molekül- oder Festkörperorbitalfunktionen und bildet somit ein Kriterium zur Bewertung der Qualität der Orbitalentwicklung. Das zweite Argument hingegen gilt insbesondere für den Ebenen-Wellenansatz und verdeutlicht noch einmal, warum für diese Methode die Pulay Kräfte keine Rolle spielen.

Die Relevanz der Pulay Kräfte lässt sich am Beispiel des Stickstoffdimers wunderbar illustrieren. Abbildung C.1 zeigt die Dissociationskurve des Moleküls, berechnet mittels selbstkonsistenter LCAO und dem potentialeigenen s - und p -Orbital als Basisfunktionen. Dieser nichtoptimierte Basissatz ist verwendet worden um einen möglichst großen Beitrag zur Pulay Kraft zu bekommen und dennoch ein einigermaßen physikalisches Bild der Bindung zu erhalten. Eine Strukturoptimierung ohne die Verwendung der Pulay Kräfte resultiert in einer fehlerhaften Bestimmung der Gleichgewichtsgeometrie. Anstelle einer Bindungslänge von 2.233 Bohr liefert die Kraftoptimierung eine Bindungslänge von 3.195 Bohr. Erst unter Berücksichtigung der Pulay Kräfte lässt sich das Energieminimum bestimmen und somit die Bindungslänge korrekt vorhersagen.

D. Variation der inversen Überlappmatrix

Die Variation der inversen Überlappmatrix S^{-1} in Abhängigkeit der Orbitalfunktionen lässt sich über die Matrix-Identität $SS^{-1} = \mathbf{1}$ bestimmen. Mit

$$0 = \sum_{\beta} \frac{dS_{\alpha\beta}}{d\langle\mu_{\tau}|} S_{\beta\gamma}^{-1} + S_{\alpha\beta} \frac{dS_{\beta\gamma}^{-1}}{d\langle\mu_{\tau}|} \quad (\text{D.1})$$

$$= \sum_{\alpha,\beta} S_{\epsilon\alpha}^{-1} \frac{dS_{\alpha\beta}}{d\langle\mu_{\tau}|} S_{\beta\gamma}^{-1} + S_{\epsilon\alpha}^{-1} S_{\alpha\beta} \frac{dS_{\beta\gamma}^{-1}}{d\langle\mu_{\tau}|} \quad (\text{D.2})$$

$$= \sum_{\alpha,\beta} \left(S_{\epsilon\alpha}^{-1} \frac{dS_{\alpha\beta}}{d\langle\mu_{\tau}|} S_{\beta\gamma}^{-1} \right) + \frac{dS_{\epsilon\gamma}^{-1}}{d\langle\mu_{\tau}|} \quad (\text{D.3})$$

folgt daraus

$$\frac{dS_{\alpha\beta}^{-1}}{d\langle\mu_{\tau}|} = - \sum_{\gamma,\epsilon} S_{\alpha\gamma}^{-1} \frac{dS_{\gamma\epsilon}}{d\langle\mu_{\tau}|} S_{\epsilon\beta}^{-1} \quad (\text{D.4})$$

$$= -S_{\alpha\tau}^{-1} \sum_{\epsilon} |\mu_{\epsilon}\rangle S_{\epsilon\beta}^{-1} \quad (\text{D.5})$$

E. Spline-Interpolation

Sei im folgenden die Funktion ϕ im Intervall $[a, b]$ durch ein Gitter $\{g_x\}_X$ mit den Datenpunkten $\phi_x = \phi(g_x)$ beschrieben. Ein weiteres Gitter $\{g_{\xi}\}_{\Xi}$ spanne dasselbe

Intervall auf, jedoch mit einer größeren Menge an Datenpunkten. Es gilt folglich $X < \Xi$. Gesucht werden nun Datenpunkte $\phi_\xi = \phi(g_\xi)$. Diese sollen durch eine kubische Spline-Interpolation berechnet werden. Dabei zerlegt man das Intervall $[a, b]$ in $X - 1$ Teilintervalle auf denen kubische Spline-Funktionen

$$S_x(g_\xi) = \sum_{i=0}^3 \chi_{x,i}(g_\xi - g_x)^i \quad \text{mit } g_x \leq g_\xi \leq g_{x+1} \quad (\text{E.1})$$

definiert werden. Eine kubische Spline-Funktion besitzt nach [113] die folgenden Eigenschaften:

$$S_x(g_x) = \phi_x \quad \forall \quad 0 \leq x < X - 1 \quad (\text{E.2})$$

$$S_x(g_{x+1}) = S_{x+1}(g_{x+1}) \quad \forall \quad 0 \leq x < X - 1 \quad (\text{E.3})$$

$$S'_x(g_{x+1}) = S'_{x+1}(g_{x+1}) \quad \forall \quad 0 \leq x < X - 1 \quad (\text{E.4})$$

$$S''_x(g_{x+1}) = S''_{x+1}(g_{x+1}) \quad \forall \quad 0 \leq x < X - 1 \quad (\text{E.5})$$

Zudem wird noch eine der folgenden Randbedingungen erfüllt:

Natürliche Randbedingung:

$$S''_0(g_0) = S''_{X-1}(g_X) = 0 \quad (\text{E.6})$$

Hermite Randbedingung:

$$\begin{aligned} S'_0(g_0) &= \phi'_0 \\ &\wedge \\ S'_{X-1}(g_X) &= \phi'_X \end{aligned} \quad (\text{E.7})$$

Aus diesen Eigenschaften lassen sich gemäß [113] die Spline-Koeffizienten $\chi_{x,i}$ berechnen. Aus der Bedingung (E.2) folgt sofort, dass sich die $i = 0$ Koeffizienten zu

$$\chi_{x,0} = \phi_x \quad (\text{E.8})$$

ergeben. Weiter lässt sich über die Bedingungen (E.3)-(E.5) und der Hilfsbezeichnung

$$h_x = g_{x+1} - g_x \quad (\text{E.9})$$

die Beziehungen

$$\chi_{x+1,0} = \chi_{x,0} + h_x \chi_{x,1} + h_x^2 \chi_{x,2} + h_x^3 \chi_{x,3} \quad (\text{E.10})$$

$$\chi_{x+1,1} = \chi_{x,1} + 2h_x \chi_{x,2} + 3h_x^2 \chi_{x,3} \quad (\text{E.11})$$

$$\chi_{x+1,2} = 3h_x \chi_{x,3} + \chi_{x,2} \quad (\text{E.12})$$

herleiten, welche die Koeffizienten aneinander koppeln. Eine Verknüpfung von Beziehung (E.12) in (E.11) liefert weiter

$$\chi_{x+1,1} = \chi_{x,1} + h_x(\chi_{x+1,2} + \chi_{x,2}) \quad (\text{E.13})$$

und Beziehung (E.11) in (E.10) liefert

$$\chi_{x+1,0} = \chi_{x,0} + h_x\chi_{x,1} + \frac{h_x^2}{3}(\chi_{x+1,2} + 2\chi_{x,2}) \quad (\text{E.14})$$

Aus Gleichung (E.14) lassen sich nun Bestimmungsgleichungen für

$$\chi_{x,1} = \frac{1}{h_x}(\chi_{x+1,0} - \chi_{x,0}) - \frac{h_x}{3}(\chi_{x+1,2} + 2\chi_{x,2}) \quad (\text{E.15})$$

und - durch Indexverschiebung - für

$$\chi_{x+1,1} = \frac{1}{h_{x+1}}(\chi_{x+2,0} - \chi_{x+1,0}) - \frac{h_{x+1}}{3}(\chi_{x+2,2} + 2\chi_{x+1,2}) \quad (\text{E.16})$$

angeben. Diese in Gleichung (E.13) eingesetzt, liefert als Bestimmungsgleichung für die Gesamtheit der $\{\chi_{x,2}\}_{x=1}^{X-1}$, wenn der Index nachträglich noch um eins reduziert wird

$$h_{x-1}\chi_{x-1,2} + 2(h_{x-1} + h_x)\chi_{x,2} + h_x\chi_{x+1,2} = \frac{3}{h_x}(\chi_{x+1,0} - \chi_{x,0}) - \frac{3}{h_{x-1}}(\chi_{x,0} - \chi_{x-1,0}) \quad (\text{E.17})$$

Für $\chi_{0,2}$ und $\chi_{X,2}$ muss nun eine der Nebenbedingungen beachtet werden. Die natürliche Nebenbedingung (E.6) ihrerseits führt zu

$$\chi_{0,2} = 0 \quad (\text{E.18})$$

$$\chi_{X,2} = 0, \quad (\text{E.19})$$

während die Hermite Nebenbedingung (E.7) auf

$$\frac{2}{3}h_0\chi_{0,2} + \frac{1}{3}h_0\chi_{1,2} = \frac{1}{h_0}(\chi_{1,0} - \chi_{0,0}) - \phi'_0 \quad (\text{E.20})$$

$$\frac{1}{3}h_{X-1}\chi_{X-1,2} + \frac{2}{3}h_{X-1}\chi_{X,2} = \phi'_X - \frac{1}{h_{X-1}}(\chi_{X,0} - \chi_{X-1,0}) \quad (\text{E.21})$$

schliessen lässt. Diese Gleichungen bilden ein geschlossenes lineares Gleichungssystem der Form

$$\underline{T} \underline{\chi}_2 = \underline{\beta}, \quad (\text{E.22})$$

dessen Lösung die Gesamtheit der $\{\chi_{x,2}\}_{x=0}^X$ festlegt. Aus diesen wiederum lassen sich dann die übrigen Spline-Koeffizienten bestimmen. Die benötigten Bestimmungsgleichungen seien an dieser Stelle noch mal zusammengefasst:

$$\chi_{x,0} = \phi_x \quad (\text{E.23})$$

$$\chi_{x,2} = (\underline{T}^{-1} \underline{\beta})_x \quad (\text{E.24})$$

$$\chi_{x,1} = \frac{1}{h_x}(\chi_{x+1,0} - \chi_{x,0}) - \frac{h_x}{3}(\chi_{x+1,2} + 2\chi_{x,2}) \quad (\text{E.25})$$

$$\chi_{x,3} = \frac{1}{3h_x}(\chi_{x+1,2} - \chi_{x,2}) \quad (\text{E.26})$$

Die Werte für $\chi_{x,1}$ und $\chi_{x,3}$ lassen sich aus den obigen Gleichungen nicht berechnen, besitzen aber auch für die Spline-Interpolation keine Relevanz, da ihr zugehöriger Spline nicht im Intervall $[a, b]$ liegt. Es sei an dieser Stelle angemerkt, dass alle Spline-Koeffizienten linear von den erzeugenden Knotenpunkten ϕ_x abhängen. Eine Taylorentwicklung

$$\chi_{x,\{0..3\}} = \sum_i \frac{\partial \chi_{x,\{0..3\}}}{\partial \phi_i} \phi_i \quad (\text{E.27})$$

ist somit vollständig.

Mittels des oben beschriebenen Verfahrens lassen sich auf einem Intervall $[a, b]$ gegebenen Gitter beliebige Punkte in diesem Intervall interpolieren. Davon wird in dem Quamolverfahren bei der Auffaltung der radialen Funktion im reziproken Raum auf das dreidimensionale reziproke Gitter der Wellenfunktionen exzessiv Gebrauch gemacht. Die Zusammenfaltung könnte mit dem selben Ansatz von statten gehen. Da sich aber im reziproken Raum der Wellenfunktionen das Signal durch die Projektionen und Rechenoperationen mit einem numerischen Rauschen überlagert, ist dies kein sinnvoller Weg, wie Abb. F.1 eindrucksvoll zeigt. Durch die Spline-Interpolation pflanzt sich der Fehler im radialen Raum fort und stört dort die Optimierung erheblich. Um diesen Fehler zu minimieren bietet sich das Verfahren des Spline-Fittens an, welches im Folgenden beschrieben wird.

F. Spline-Fitten

Die Ausgangslage bei diesem Verfahren ist umgekehrt zur Spline-Interpolation. Die Funktion ϕ sei nun eine gestörte Funktion, die auf einem dichten Gitter g_ξ im Intervall $[a, b]$ über $\phi_\xi = \phi(g_\xi)$ gegeben ist. Gesucht werden nun die Knotenpunkte $f_x = f(g_x)$, welche auf dem weniger dichten Gitter g_x einen Spline der Art erzeugen, so dass das Residuum

$$R = \sum_\xi [\phi_\xi - S(g_\xi)]^2 \quad (\text{F.1})$$

minimal wird. Die Variation des Residuums bezüglich eines Knotenpunktes f_i lässt sich zu

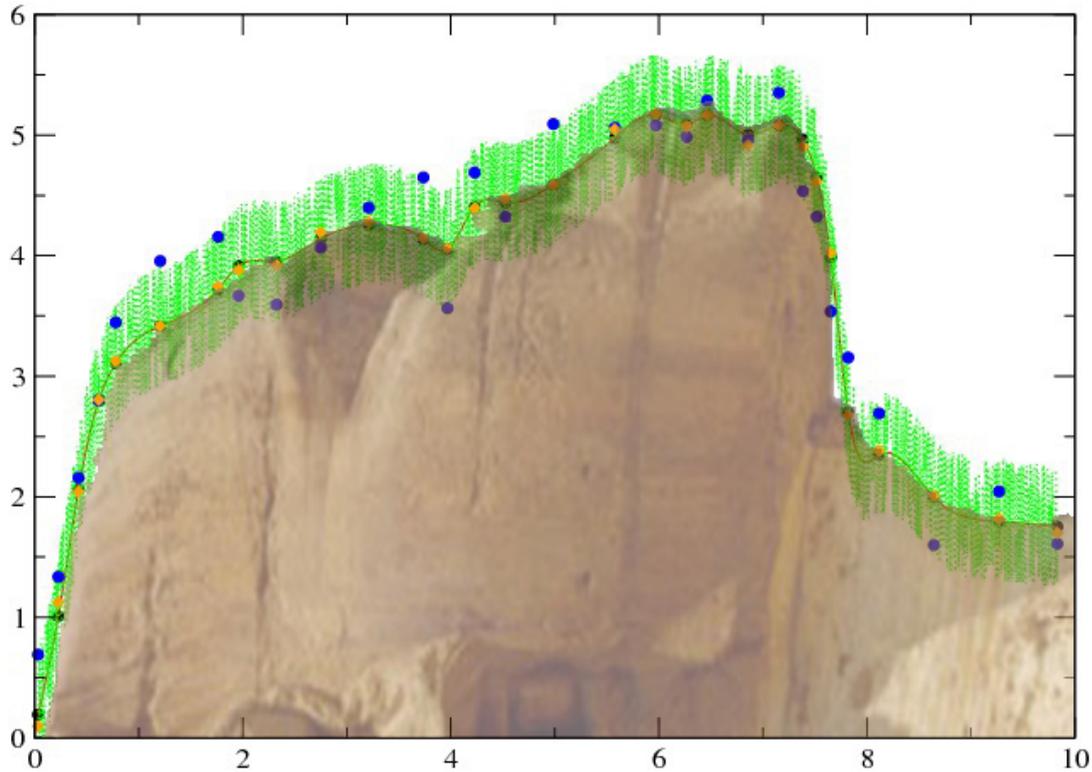


Abb. F.1:

Spline-Interpolation und Spline-Fitten am Profil der Sphinx. Die Abbildung der Sphinx ist mit einem Stützstellengitter belegt worden, dessen Spline-Interpolation ihr Profil reproduziert. Die Interpolation ist dann mit einem Rauschen (grün) überlagert worden. Die blauen Punkte zeigen die sich aus diesem verrauschten Signal ergebenden neuen Stützstellen. Die orangenen Punkte geben die Stützstellen wieder, die das Verfahren des Spline-Fittens liefert.

$$\begin{aligned} \frac{\partial R}{\partial f_i} = & 2 \sum_{\xi} \sum_{p,q=0}^3 \chi_{x,p} \frac{\partial \chi_{x,q}}{\partial f_i} (g_{\xi} - g_x)^{p+q} \\ & - 2 \sum_{\xi} \sum_{p=0}^3 \phi_{\xi} \frac{\partial \chi_{x,p}}{\partial f_i} (g_{\xi} - g_x)^p \end{aligned} \quad (\text{F.2})$$

berechnen. Da nun die Spline-Koeffizienten linear von den Knotenpunkten abhängen (Gl. (E.27)) und das lokale Minimum des Residuums über die hinreichende Bedingung $\frac{\partial R}{\partial f_i} = 0$ gegeben ist, ergibt sich das folgende lineare Gleichungssystem zur Bestimmung der f_x

Simulationsschritt	Mittelwert	Varianz
Rauschen	0.0082	0.1243
Splineinterpolation	0.0646	0.1064
Splinefitting	0.0008	0.0018

Tab. F.1:

Der durch das Profil der Sphinx gelegte Spline wurde mit einem Rauschen überlagert. Vollzieht man nun an diesem Signal eine Spline-Interpolation und berechnet Mittelwert und Varianz der Differenz zum rauschfreien Ursprungssignal, so erkennt man einen deutlichen Fehler. Durch das Spline-Fitting wird dieser Fehler reduziert und die Effekte des Rauschens minimiert.

$$\sum_k f_k A_{ki} = b_i \quad (\text{F.3})$$

$$\text{mit } A_{ki} = \sum_{\xi} \sum_{p,q=0}^3 \frac{\partial \chi_{x,p}}{\partial f_k} \frac{\partial \chi_{x,q}}{\partial f_i} (g_{\xi} - g_x)^{p+q}$$

$$\text{und } b_i = \sum_{\xi} \sum_{p=0}^3 \phi_{\xi} \frac{\partial \chi_{x,p}}{\partial f_i} (g_{\xi} - g_x)^p.$$

Die entsprechenden Ableitungen der Spline-Koeffizienten nach den Knotenpunkten lassen sich über

$$\frac{d\chi_{x,0}}{df_i} = \delta_{x,i} \quad (\text{F.4})$$

$$\frac{d\chi_{x,1}}{df_i} = \frac{1}{h_x} (\delta_{x+1,i} - \delta_{x,i}) - \frac{h_x}{3} \left(\frac{d\chi_{x+1,2}}{df_i} + 2 \frac{d\chi_{x,2}}{df_i} \right) \quad (\text{F.5})$$

$$\frac{d\chi_{x,2}}{df_i} = \left(\underline{\underline{T}}^{-1} \frac{d\beta}{df_i} \right)_x \quad (\text{F.6})$$

$$\frac{d\chi_{x,3}}{df_i} = \frac{1}{3h_x} \left(\frac{d\chi_{x+1,2}}{df_i} - \frac{d\chi_{x,2}}{df_i} \right) \quad (\text{F.7})$$

berechnen. Die Mächtigkeit dieses Verfahrens lässt sich an Abbildung F.1 demonstrieren. Über das Profil der Sphinx ist ein Stützstellengitter gelegt worden, welches mittels Spline-Interpolation das Profil nachbildet. Diese Spline-Interpolation wurde dann mit einem Rauschen überlagert. Die direkte Verwendung dieses Signals zur Rückrechnung der Stützstellen resultiert in einem deformierten Profil. Durch die Verwendung des Spline-Fittens lassen sich die Stützstellen nahezu perfekt rekonstruieren. Tabelle F.1 zeigt die berechneten Varianzen der durch Spline-Fittens und Spline-Interpolation rekonstruierten Stützstellen und den Initialstützstellen. Durch die Verwendung des Spline-Fittens kann der Fehler und die Varianz des Rauschens signifikant reduziert werden.

Literatur

- [1] OSRAM Semiconductors. Lebensdauer von led-modulen, 12 2011. URL http://www.osram.de/osram_de/LED/Alles_ueber_LED/Lebensdauer_von_LED-Modulen/index.html.
- [2] Harold M. Manasevit. Single-crystal gallium arsenide on insulating substrates. *Applied Physics Letters*, 12(4):156–159, 1968. doi: 10.1063/1.1651934. URL <http://link.aip.org/link/?APL/12/156/1>.
- [3] D. L. Rode and D. K. Gaskill. Electron hall mobility of n-gan. *Applied Physics Letters*, 66(15):1972–1973, 1995. doi: 10.1063/1.113294. URL <http://link.aip.org/link/?APL/66/1972/1>.
- [4] Chris G. Van de Walle and Jörg Neugebauer. First-principles calculations for defects and impurities: Applications to iii-nitrides. *J. Appl. Phys.*, 95(8):3851–3879, 2004. doi: 10.1063/1.1682673.
- [5] Jörg Neugebauer and Chris G. Van de Walle. Role of hydrogen in doping of gan. *Appl. Phys. Lett.*, 68(13):1829–1831, 1996. doi: 10.1063/1.116027.
- [6] A. F. Wright, C. H. Seager, S. M. Myers, D. D. Koleske, and A. A. Allerman. Hydrogen configurations, formation energies, and migration barriers in gan. *J. Appl. Phys.*, 94(4):2311–2318, 2003. doi: 10.1063/1.1592607.
- [7] S. M. Myers and A. F. Wright. Configurations, energies, and thermodynamics of the neutral mgh complex in gan. *J. Appl. Phys.*, 94(8):4918–4922, 2003. doi: 10.1063/1.1610232.
- [8] W. R. Wampler and S. M. Myers. Hydrogen release from magnesium-doped gan with clean ordered surfaces. *J. Appl. Phys.*, 94(9):5682–5687, 2003. doi: 10.1063/1.1616986.
- [9] Norbert Linder and Martin Strassburg. project 03x0512g (vektra). private communication, 2010.
- [10] M. Leroux, P. Vennéguès, S. Dalmaso, M. Benaissa, E. Feltin, P. De Mierry, B. Beaumont, B. Damilano, N. Grandjean, and P. Gibart. Structural defects and relation with optoelectronic properties in highly mg-doped gan. *Phys. Stat. Sol. (a)*, 192(2):394–400, 2002. doi: 10.1002/1521-396X.
- [11] M. Leroux, P. Vennegues, S. Dalmaso, P. de Mierry, P. Lorenzini, B. Damilano, B. Beaumont, P. Gibart, and J. Massies. Pyramidal defects in highly mg-doped gan: atomic structure and influence on optoelectronic properties. *The European Physical Journal - Applied Physics*, 27(1-3):259–262, 2004. doi: 10.1051/epjap:2004119-2. URL <http://dx.doi.org/10.1051/epjap:2004119-2>.

- [12] P. Vennéguès, M. Benaïssa, B. Beaumont, E. Feltin, P. De Mierry, S. Dalmaso, M. Leroux, and P. Gibart. Pyramidal defects in metalorganic vapor phase epitaxial mg doped gan. *Appl. Phys. Lett.*, 77(6):880–882, 2000. doi: 10.1063/1.1306421.
- [13] P. Vennéguès, M. Leroux, S. Dalmaso, M. Benaïssa, P. De Mierry, P. Lorenzini, B. Damilano, B. Beaumont, J. Massies, and P. Gibart. Atomic structure of pyramidal defects in mg-doped gan. *Phys. Rev. B*, 68(23):235214, Dec 2003. doi: 10.1103/PhysRevB.68.235214.
- [14] John E. Northrup. Magnesium incorporation at (0001) inversion domain boundaries in gan. *Appl. Phys. Lett.*, 82(14):2278–2280, 2003. doi: 10.1063/1.1565707.
- [15] A. F. Wright and T. R. Mattsson. N vacancy diffusion and trapping in mg-doped wurtzite gan. *Journal of Applied Physics*, 96(4):2015–2022, 2004. doi: 10.1063/1.1767981. URL <http://link.aip.org/link/?JAP/96/2015/1>.
- [16] Chris G. Van de Walle and Jörg Neugebauer. Universal alignment of hydrogen levels in semiconductores, insulators and solutions. *Nature*, 423(6940):626–628, Jun 2003. doi: 10.1038/nature01665.
- [17] G. Kresse and J. Furthmüller. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Phys. Rev. B*, 54(16):11169–11186, Oct 1996. doi: 10.1103/PhysRevB.54.11169.
- [18] Sixten Boeck, Christoph Freysoldt, Alexej Dick, Lars Ismer, and Jörg Neugebauer. The object-oriented dft program library s/phi/nx. *Computer Physics Communications*, 182(3):543 – 554, 2011.
- [19] Leslie E. Ballentine. *Quantum mechanics - A modern development*. World Scientific Singapore, New Jersey, London, Hong Kong, 6 edition, 2006. ISBN 981-02-4105-4.
- [20] Frank Jensen. *Introduction to Computational Chemistry*. John Wiley & Sons Ltd, Chichester, New York, Weinheim, Brisbane, Singapore, Toronto, 1999. ISBN 0-471-98055-4.
- [21] M. Born and R. Oppenheimer. Zur quantentheorie der molekeln. *Annalen der Physik*, 389(20):457–484, 1927. ISSN 1521-3889. doi: 10.1002/andp.19273892002. URL <http://dx.doi.org/10.1002/andp.19273892002>.
- [22] P. Hohenberg and W. Kohn. Inhomogeneous electron gas. *Phys. Rev.*, 136(3B):B864–B871, Nov 1964. doi: 10.1103/PhysRev.136.B864.
- [23] Klaus Capelle. A bird’s-eye view of density-functional theory. *Proceeding of the VIII’th Brazilian Summer School on Electronic Structure*, 10:1, 2003.

-
- [24] W. Kohn and L. J. Sham. Self-consistent equations including exchange and correlation effects. *Phys. Rev.*, 140(4A):A1133–A1138, Nov 1965. doi: 10.1103/PhysRev.140.A1133.
- [25] W. Kohn, A.D. Becke, and R.G. Parr. Density functional theory of electronic structure. *Journal of Physical Chemistry*, 100(31):12974–12980, 1996. ISSN 0022-3654.
- [26] Charles Kittel. *Einführung in die Festkörperphysik*. Oldenburg Verlag München Wien, 14 edition, 2006.
- [27] CRC. *Handbook of Chemistry and Physics*. CRC Press, Taylor & Francis Group, 6000 Broken Sound Parkway NW, Suite 300, 86 edition, 2005-2006.
- [28] Robert G. Parr and Weitao Yang. *Density-Functional Theory of Atoms and Molecules*. Oxford University Press, 1989.
- [29] A. D. Becke. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A*, 38(6):3098–3100, Sep 1988. doi: 10.1103/PhysRevA.38.3098.
- [30] D. M. Ceperley and B. J. Alder. Ground state of the electron gas by a stochastic method. *Phys. Rev. Lett.*, 45(7):566–569, Aug 1980. doi: 10.1103/PhysRevLett.45.566.
- [31] J. P. Perdew and Alex Zunger. Self-interaction correction to density-functional approximations for many-electron systems. *Phys. Rev. B*, 23(10):5048–5079, May 1981. doi: 10.1103/PhysRevB.23.5048.
- [32] Frank Herman, John P. Van Dyke, and Irene B. Ortenburger. Improved statistical exchange approximation for inhomogeneous many-electron systems. *Phys. Rev. Lett.*, 22(16):807–811, Apr 1969. doi: 10.1103/PhysRevLett.22.807.
- [33] John P. Perdew. Density-functional approximation for the correlation energy of the inhomogeneous electron gas. *Phys. Rev. B*, 33(12):8822–8824, Jun 1986. doi: 10.1103/PhysRevB.33.8822.
- [34] John P. Perdew and Wang Yue. Accurate and simple density functional for the electronic exchange energy: Generalized gradient approximation. *Phys. Rev. B*, 33(12):8800–8802, Jun 1986. doi: 10.1103/PhysRevB.33.8800.
- [35] John P. Perdew, Kieron Burke, and Matthias Ernzerhof. Generalized gradient approximation made simple. *Phys. Rev. Lett.*, 77(18):3865–3868, Oct 1996. doi: 10.1103/PhysRevLett.77.3865.
- [36] B. Monemar. Fundamental energy gap of gan from photoluminescence excitation spectra. *Phys. Rev. B*, 10:676–681, Jul 1974. doi: 10.1103/PhysRevB.10.676. URL <http://link.aps.org/doi/10.1103/PhysRevB.10.676>.
-

- [37] Olaf Reckeweg, Cora Lind, Arndt Simon, and Francis J. DiSalvo. Rietveld refinement of the crystal structure of α - be_3n_2 and the experimental determination of optical band gaps for mg_3n_2 , ca_3n_2 and camg_2n_2 . *Z. Naturforsch.*, 58b:159 – 162, 2003.
- [38] Kazuaki Toyoura, Takuya Goto, Kan Hachiya, and Rika Hagiwara. Structural and optical properties of magnesium nitride formed by a novel electrochemical process. *Elec. Acta*, 51(1):56 – 60, 2005. ISSN 0013-4686. doi: DOI:10.1016/j.electacta.2005.04.004.
- [39] R. Shaltaf, G.-M. Rignanese, X. Gonze, Feliciano Giustino, and Alfredo Pasquarello. Band offsets at the Si/ siO_2 interface from many-body perturbation theory. *Phys. Rev. Lett.*, 100:186401, May 2008. doi: 10.1103/PhysRevLett.100.186401. URL <http://link.aps.org/doi/10.1103/PhysRevLett.100.186401>.
- [40] Audrius Alkauskas and Alfredo Pasquarello. Band-edge problem in the theoretical determination of defect energy levels: The o vacancy in zno as a benchmark case. *Phys. Rev. B*, 84:125206, Sep 2011. doi: 10.1103/PhysRevB.84.125206. URL <http://link.aps.org/doi/10.1103/PhysRevB.84.125206>.
- [41] R. Ramprasad, H. Zhu, Patrick Rinke, and Matthias Scheffler. New perspective on formation energies and energy levels of point defects in nonmetals. *Phys. Rev. Lett.*, 108:066404, Feb 2012. doi: 10.1103/PhysRevLett.108.066404. URL <http://link.aps.org/doi/10.1103/PhysRevLett.108.066404>.
- [42] Axel D. Becke. Density-functional thermochemistry. iii. the role of exact exchange. *The Journal of Chemical Physics*, 98(7):5648–5652, 1993. doi: 10.1063/1.464913. URL <http://link.aip.org/link/?JCP/98/5648/1>.
- [43] Jochen Heyd, Gustavo E. Scuseria, and Matthias Ernzerhof. Hybrid functionals based on a screened coulomb potential. *The Journal of Chemical Physics*, 118(18):8207–8215, 2003.
- [44] Jochen Heyd, Gustavo E. Scuseria, and Matthias Ernzerhof. Erratum: Hybrid functionals based on a screened coulomb potential [j. chem. phys. 118, 8207 (2003)]. *The Journal of Chemical Physics*, 124(21):219906, 2006.
- [45] John P. Perdew, Matthias Ernzerhof, and Kieron Burke. Rationale for mixing exact exchange with density functional approximations. *The Journal of Chemical Physics*, 105(22):9982–9985, 1996. doi: 10.1063/1.472933. URL <http://link.aip.org/link/?JCP/105/9982/1>.
- [46] Matthias Ernzerhof, John P. Perdew, and Kieron Burke. Coupling-constant dependence of atomization energies. *International Journal of Quantum Chemistry*, 64(3):285–295, 1997. ISSN 1097-461X. doi: 10.1002/(SICI)

-
- 1097-461X(1997)64:3<285::AID-QUA2>3.0.CO;2-S. URL [http://dx.doi.org/10.1002/\(SICI\)1097-461X\(1997\)64:3<285::AID-QUA2>3.0.CO;2-S](http://dx.doi.org/10.1002/(SICI)1097-461X(1997)64:3<285::AID-QUA2>3.0.CO;2-S).
- [47] Matthias Ernzerhof and Gustavo E. Scuseria. Assessment of the perdew–burke–ernzerhof exchange–correlation functional. *The Journal of Chemical Physics*, 110(11):5029–5036, 1999. doi: 10.1063/1.478401. URL <http://link.aip.org/link/?JCP/110/5029/1>.
- [48] Carlo Adamo and Vincenzo Barone. Toward reliable density functional methods without adjustable parameters: The pbe0 model. *The Journal of Chemical Physics*, 110(13):6158–6170, 1999. doi: 10.1063/1.478522. URL <http://link.aip.org/link/?JCP/110/6158/1>.
- [49] Lars Hedin. New method for calculating the one-particle green’s function with application to the electron-gas problem. *Phys. Rev.*, 139:A796–A823, Aug 1965. doi: 10.1103/PhysRev.139.A796. URL <http://link.aps.org/doi/10.1103/PhysRev.139.A796>.
- [50] Patrick Rinke, Abdallah Qteish, Jörg Neugebauer, Christoph Freysoldt, and Matthias Scheffler. Combining gw calculations with exact-exchange density-functional theory: an analysis of valence-band photoemission for compound semiconductors. *New Journal of Physics*, 7(1):126, 2005. URL <http://stacks.iop.org/1367-2630/7/i=1/a=126>.
- [51] Max Born and Theodore von Kármán. On fluctuations in spatial grids. *Physikalische Zeitschrift der Sowjetunion*, 13:297 – 309, 1912.
- [52] Max Born and Theodore von Kármán. On the distribution of natural vibrations of spot-lattices. *Physikalische Zeitschrift der Sowjetunion*, 14:65 – 71, 1913.
- [53] A. Baldereschi. Mean-value point in the brillouin zone. *Phys. Rev. B*, 7(12):5212–5215, Jun 1973. doi: 10.1103/PhysRevB.7.5212.
- [54] D. J. Chadi and Marvin L. Cohen. Special points in the brillouin zone. *Phys. Rev. B*, 8(12):5747–5753, Dec 1973. doi: 10.1103/PhysRevB.8.5747.
- [55] Hendrik J. Monkhorst and James D. Pack. Special points for brillouin-zone integrations. *Phys. Rev. B*, 13(12):5188–5192, Jun 1976. doi: 10.1103/PhysRevB.13.5188.
- [56] U. von Barth and C. D. Gelatt. Validity of the frozen-core approximation and pseudopotential theory for cohesive energy calculations. *Phys. Rev. B*, 21(6):2222–2228, Mar 1980. doi: 10.1103/PhysRevB.21.2222.
-

- [57] D. R. Hamann, M. Schlüter, and C. Chiang. Norm-conserving pseudopotentials. *Phys. Rev. Lett.*, 43:1494–1497, Nov 1979. doi: 10.1103/PhysRevLett.43.1494. URL <http://link.aps.org/doi/10.1103/PhysRevLett.43.1494>.
- [58] Alex Zunger and Mark A. Ratner. On the first principles hartree-fock and local density pseudopotentials. *Chemical Physics*, 30(3):423 – 443, 1978. ISSN 0301-0104. doi: 10.1016/0301-0104(78)87014-1. URL <http://www.sciencedirect.com/science/article/pii/0301010478870141>.
- [59] G. B. Bachelet, D. R. Hamann, and M. Schlüter. Pseudopotentials that work: From h to pu. *Phys. Rev. B*, 26:4199–4228, Oct 1982. doi: 10.1103/PhysRevB.26.4199. URL <http://link.aps.org/doi/10.1103/PhysRevB.26.4199>.
- [60] David Vanderbilt. Optimally smooth norm-conserving pseudopotentials. *Phys. Rev. B*, 32:8412–8415, Dec 1985. doi: 10.1103/PhysRevB.32.8412. URL <http://link.aps.org/doi/10.1103/PhysRevB.32.8412>.
- [61] Warren E. Pickett. Pseudopotential methods in condensed matter applications. *Computer Physics reports*, 9(3):115 – 197, 1989. ISSN 0167-7977. doi: DOI:10.1016/0167-7977(89)90002-6. URL <http://www.sciencedirect.com/science/article/pii/0167797789900026>.
- [62] N. Troullier and José Luriaas Martins. Efficient pseudopotentials for plane-wave calculations. *Phys. Rev. B*, 43(3):1993–2006, Jan 1991. doi: 10.1103/PhysRevB.43.1993.
- [63] P. E. Blöchl. Projector augmented-wave method. *Phys. Rev. B*, 50(24):17953–17979, 1994.
- [64] Jörg Neugebauer and Chris G. Van de Walle. Tight-binding initialization for generating high-quality initial wave functions: Application to defects and impurities in gan. *MRS Proceedings*, 408:43 – 48, 1995.
- [65] G. Makov and M. C. Payne. Periodic boundary conditions in ab initio calculations. *Phys. Rev. B*, 51(7):4014–4022, Feb 1995. doi: 10.1103/PhysRevB.51.4014.
- [66] Christoph Freysoldt, Jörg Neugebauer, and Chris G. Van de Walle. Fully ab initio finite-size corrections for charged-defect supercell calculations. *Phys. Rev. Lett.*, 102(1):016402, Jan 2009. doi: 10.1103/PhysRevLett.102.016402.
- [67] Christoph Freysoldt, Jörg Neugebauer, and Chris G. Van de Walle. Electrostatic interactions between charged defects in supercells. *physica status solidi (b)*, 248(5):1067–1076, 2011. ISSN 1521-3951. doi: 10.1002/pssb.201046289. URL <http://dx.doi.org/10.1002/pssb.201046289>.

-
- [68] A. S. Barker and M. Ilegems. Infrared lattice vibrations and free-electron dispersion in gan. *Phys. Rev. B*, 7(2):743–750, Jan 1973. doi: 10.1103/PhysRevB.7.743.
- [69] D. D. Manchon, A. S. Barker, P. J. Dean, and R. B. Zetterstrom. Optical studies of the phonons and electrons in gallium nitride. *Solid State Communications*, 8(15):1227 – 1231, 1970. ISSN 0038-1098. doi: DOI:10.1016/0038-1098(70)90365-0. URL <http://www.sciencedirect.com/science/article/pii/0038109870903650>.
- [70] www.chemicaldictionary.org. Magnesium nitride. online dictionary entry, 2009. URL http://www.chemicaldictionary.org/dic/M/Magnesium-nitride_987.html.
- [71] J. C. Slater. Atomic shielding constants. *Phys. Rev.*, 36:57–64, Jul 1930. doi: 10.1103/PhysRev.36.57. URL <http://link.aps.org/doi/10.1103/PhysRev.36.57>.
- [72] S.F. Boys. Electronic wave functions. i. a general method of calculation for the stationary states of any molecular system. *Proc. R. Soc. Lond. A*, 200: 542–554, 1950.
- [73] G.H. Wannier. The structure of electronic excitation levels in insulating crystals. *Phys. Rev.*, 52:191, 1937.
- [74] Nicola Marzari and David Vanderbilt. Maximally localized generalized wannier functions for composite energy bands. *Phys. Rev. B*, 56(20):12847–12865, Nov 1997. doi: 10.1103/PhysRevB.56.12847.
- [75] M. Wahn and J. Neugebauer. Generalized wannier functions: An efficient way to construct ab-initio tight-binding parameters for group-iii nitrides. *Phys. stat. sol. (b)*, 243:1583–1587, 2006.
- [76] Daniel Sánchez-Portal, Emilio Artacho, and José M Soler. Projection of plane-wave calculations into atomic orbitals. *Sol. Stat. Comm.*, 95(10):685, 1995.
- [77] Daniel Sánchez-Portal, Emilio Artacho, and José M Soler. Analysis of atomic orbital basis sets from the projection of plane-wave results. *J. Phys.: Cond. Mat.*, 8(21):3859, 1996.
- [78] S. D. Kenny, A. P. Horsfield, and Hideaki Fujitani. Transferable atomic-type orbital basis sets for solids. *Phys. Rev. B*, 62(8):4899–4905, Aug 2000. doi: 10.1103/PhysRevB.62.4899.
- [79] Michele Gusso. Study on the maximum accuracy of the pseudopotential density functional method with localized atomic orbitals versus plane-wave basis sets. *J. Chem. Phys.*, 128(4):044102, 2008. doi: 10.1063/1.2821023.
-

- [80] Mohan Chen, G-C Guo, and Lixin He. Systematically improvable optimized atomic basis sets for ab initio calculations. *Journal of Physics: Condensed Matter*, 22(44):445501, 2010. URL <http://stacks.iop.org/0953-8984/22/i=44/a=445501>.
- [81] Mohan Chen, G-C Guo, and Lixin He. Electronic structure interpolation via atomic orbitals. *Journal of Physics: Condensed Matter*, 23(32):325501, 2011. URL <http://stacks.iop.org/0953-8984/23/i=32/a=325501>.
- [82] K. L. Schuchardt, B. T. Didier, T. Elsethagen and L. Sun, V. Gurumoorthi, J. Chase, J. Li, and T. L. Windus. Basis set exchange: A community database for computational sciences. *J. Chem. Inf. Model.*, 47(3):1045–1052, 2007.
- [83] Delano P. Chong. Completeness profiles of one-electron basis sets. *Can. J. Chem.*, 73(1):79–83, 1995.
- [84] G. Kresse and D. Joubert. From ultrasoft pseudopotentials to the projector augmented-wave method. *Phys. Rev. B*, 59(3):1758–1775, Jan 1999. doi: 10.1103/PhysRevB.59.1758.
- [85] F. D. Murnaghan. The compressibility of media under extreme pressures. *Proc. Nat. Acad. Sci. USA*, 30:244–247, 1944.
- [86] R. S. Mulliken. Electronic population analysis on lcao[single bond]mo molecular wave functions. i. *The Journal of Chemical Physics*, 23(10):1833–1840, 1955. doi: 10.1063/1.1740588. URL <http://link.aip.org/link/?JCP/23/1833/1>.
- [87] F. A Ponce and D. P. Bour. Nitride-based semiconductors for blue and green light-emitting devices. *Nature*, 386(6623):351–359, Mar 1997. doi: 10.1038/386351a0.
- [88] Deodatta V. Shenai-Khatkhate, Randall J. Goyette, Ronald L. DiCarlo Jr., and Gregory Dripps. Environment, health and safety issues for sources used in movpe growth of compound semiconductors. *Journal of Crystal Growth*, 272:816 – 821, 2004. ISSN 0022-0248. doi: 10.1016/j.jcrysgro.2004.09.007. URL <http://www.sciencedirect.com/science/article/pii/S0022024804011121>. <ce:title>The Twelfth International Conference on Metalorganic Vapor Phase Epitaxy</ce:title>.
- [89] B. Kuhn, M. Welsch, M. Kessler, and F. Scholz. The influence of growth conditions on the electrical properties of magnesium-doped gan grown by movpe. *physica status solidi (a)*, 176(1):787–792, 1999. ISSN 1521-396X. doi: 10.1002/(SICI)1521-396X(199911)176:1<787::AID-PSSA787>3.0.CO;2-J. URL [http://dx.doi.org/10.1002/\(SICI\)1521-396X\(199911\)176:1<787::AID-PSSA787>3.0.CO;2-J](http://dx.doi.org/10.1002/(SICI)1521-396X(199911)176:1<787::AID-PSSA787>3.0.CO;2-J).

-
- [90] E. R. Glaser, W. E. Carlos, G. C. B. Braga, J. A. Freitas, W. J. Moore, B. V. Shanabrook, R. L. Henry, A. E. Wickenden, D. D. Koleske, H. Obloh, P. Kozodoy, S. P. DenBaars, and U. K. Mishra. Magnetic resonance studies of mg-doped gan epitaxial layers grown by organometallic chemical vapor deposition. *Phys. Rev. B*, 65:085312, Feb 2002. doi: 10.1103/PhysRevB.65.085312. URL <http://link.aps.org/doi/10.1103/PhysRevB.65.085312>.
- [91] M. E. Zvanut, Y. Uprety, J. Dashdorj, M. Moseley, and W. Alan Doolittle. Passivation and activation of mg acceptors in heavily doped gan. *Journal of Applied Physics*, 110(4):044508, 2011. doi: 10.1063/1.3626461. URL <http://link.aip.org/link/?JAP/110/044508/1>.
- [92] Hazem Abu-Farsakh and Jörg Neugebauer. Enhancing nitrogen solubility in gaas and inas by surface kinetics: An *ab initio* study. *Phys. Rev. B*, 79:155311, Apr 2009. doi: 10.1103/PhysRevB.79.155311. URL <http://link.aps.org/doi/10.1103/PhysRevB.79.155311>.
- [93] Ping Chen, Zhitao Xiong, Jizhong Luo, Jianyi Lin, and Kuang Lee Tan. Interaction of hydrogen with metal nitrides and imides. *Nature*, 420(6913):302–304, Nov 2002. doi: 10.1038/nature01210.
- [94] Y. Nakamori, G. Kitahara, K. Miwa, and S. Towata. Reversible hydrogen-storage functions for mixtures of li_3n and mg_3n_2 . *Appl. Phys. A*, 80(1):1 – 3, 2005. doi: DOI:10.1007/s00339-004-3002-6.
- [95] Yoshitsugu Kojima, Yasuaki Kawai, and Nobuko Ohba. Hydrogen storage of metal nitrides by a mechanochemical reaction. *J. Pow. Sour.*, 159(1):81 – 87, 2006. ISSN 0378-7753. doi: DOI:10.1016/j.jpowsour.2006.04.024.
- [96] D. E. Partin, D. J. Williams, and M. O’Keeffe. The crystal structures of mg_3n_2 and zn_3n_2 . *J. Sol. Stat. Chem.*, 132(1):56 – 59, 1997. ISSN 0022-4596. doi: DOI:10.1006/jssc.1997.7407.
- [97] Ma. Guadalupe Moreno Armenta, Armando Reyes-Serrato, and Miguel Avalos Borja. Ab initio determination of the electronic structure of beryllium-, aluminum-, and magnesium-nitrides: A comparative study. *Phys. Rev. B*, 62(8):4890–4898, Aug 2000. doi: 10.1103/PhysRevB.62.4890.
- [98] A. Mokhtari and H. Akbarzadeh. Ab initio calculations of the electronic and structural properties of beryllium-, magnesium- and calcium-nitrides. *Physica B*, 337(1-4):122 – 129, 2003. ISSN 0921-4526. doi: DOI:10.1016/S0921-4526(03)00387-9.
- [99] Martin Fuchs and Matthias Scheffler. Ab initio pseudopotentials for electronic structure calculations of poly-atomic systems using density-functional theory.
-

- Comp. Phys. Comm.*, 119(1):67 – 98, 1999. ISSN 0010-4655. doi: DOI:10.1016/S0010-4655(98)00201-X.
- [100] B. P. Stoicheff. High resolution raman spectroscopy of gases: Ix. spectra of h_2 , hd , and d_2 . *Can. J. Phys.*, 37(6):730, 1957.
- [101] S. M. Myers, A. F. Wright, M. Sanati, and S. K. Estreicher. Theoretical properties of the n vacancy in p-type gan(mg,h) at elevated temperatures. *J. Appl. Phys.*, 99(11):113506, 2006. doi: 10.1063/1.2195894.
- [102] Peter Kozodoy, Huili Xing, Steven P. DenBaars, Umesh K. Mishra, A. Saxler, R. Perrin, S. Elhamri, and W. C. Mitchel. Heavy doping effects in mg-doped gan. *J. Appl. Phys.*, 87(4):1832–1835, 2000. ISSN 00218979. doi: DOI:10.1063/1.372098. URL <http://dx.doi.org/10.1063/1.372098>.
- [103] K Laaksonen, M G Ganchenkova, and R M Nieminen. Vacancies in wurtzite gan and aln. *Journal of Physics: Condensed Matter*, 21(1):015803, 2009. URL <http://stacks.iop.org/0953-8984/21/i=1/a=015803>.
- [104] G. A. Baraff and M. Schlüter. Migration of interstitials in silicon. *Phys. Rev. B*, 30:3460–3469, Sep 1984. doi: 10.1103/PhysRevB.30.3460. URL <http://link.aps.org/doi/10.1103/PhysRevB.30.3460>.
- [105] P. Bogusławski, E. L. Briggs, and J. Bernholc. Native defects in gallium nitride. *Phys. Rev. B*, 51:17255–17258, Jun 1995. doi: 10.1103/PhysRevB.51.17255. URL <http://link.aps.org/doi/10.1103/PhysRevB.51.17255>.
- [106] S B Zhang. The microscopic origin of the doping limits in semiconductors and wide-gap materials and recent developments in overcoming these limits: a review. *Journal of Physics: Condensed Matter*, 14(34):R881, 2002. URL <http://stacks.iop.org/0953-8984/14/i=34/a=201>.
- [107] J. Karpinski and S. Porowski. High pressure thermodynamics of gan. *Journal of Crystal Growth*, 66(1):11 – 20, 1984. ISSN 0022-0248. doi: 10.1016/0022-0248(84)90071-X. URL <http://www.sciencedirect.com/science/article/pii/002202488490071X>.
- [108] D. W. Mitchell. Heat contents and heat of formation of magnesium nitride. *Industrial & Engineering Chemistry*, 41(9):2027–2031, 1949. doi: 10.1021/ie50477a046. URL <http://pubs.acs.org/doi/abs/10.1021/ie50477a046>.
- [109] Naval Research Laboratory. Crystal lattice structures web page, 10 2008. URL <http://cst-www.nrl.navy.mil/lattice>.
- [110] Clas Persson and Claudia Ambrosh-Draxl. A full-band fplapw + $\mathbf{k} \cdot \mathbf{p}$ -method for solving the kohn-sham equation. *Comp. Phys. Comm.*, 177:280–287, 2007.

- [111] P. Pulay. Ab initio calculation of force constants and equilibrium geometries in polyatomic molecules – i. theory. *Molecular Physics: An International Journal at the Interface Between Chemistry and Physics*, 17:197–204, 1969. doi: 10.1080/00268976900100941. URL <http://www.informaworld.com/10.1080/00268976900100941>.
- [112] C. Satoko. Direct force calculation in the x[alpha] method and its application to chemisorption of an oxygen atom on the al(111) surface. *Chemical Physics Letters*, 83(1):111 – 115, 1981. ISSN 0009-2614. doi: DOI:10.1016/0009-2614(81)80300-4. URL <http://www.sciencedirect.com/science/article/pii/0009261481803004>.
- [113] Richard L. Burden and J. Douglas Faires. *Numerische Methoden: Näherungsverfahren und ihre praktische Anwendung*. Spektrum Akademischer Verlag Heidelberg, Berlin, Oxford, 1993. ISBN 3-8274-0596-3.