

Matthias Wahn

Implementierung und Test des Variationellen EXX-Verfahrens

IMPLEMENTIERUNG UND TEST
DES VARIATIONELLEN EXX-VERFAHRENS

Matthias Wahn

Dissertation

zur Erlangung des akademischen Grades

DOCTOR RERUM NATURALIUM

vorgelegt dem

Department Physik der Fakultät für Naturwissenschaften
an der Universität Paderborn

Tag der Einreichung: 15.5.2009

Tag der mündlichen Prüfung: 22.6.2009

Promotionskommission:

Professor Dr. Klaus Lischka (Vorsitzender)

Professor Dr. Wolf Gero Schmidt (1. Gutachter)

Professor Dr. Jörg Neugebauer (2. Gutachter)

Dr. Harald Herrmann

Kurzbeschreibung

In der vorliegenden Arbeit wurde das variationelle EXX-Verfahren (varEXX) [78, 79, 80] in einen Pseudopotential-Ebene-Wellen-Code (S/PHI/nX [81, 82]) implementiert, und die Methode — zuvor nur an lokalisierten Systemen getestet — so erstmalig auch für periodische Strukturen verfügbar gemacht. Anschließend durchgeführte Tests an verschiedenen Halbleitermaterialien zeigten, dass varEXX die benötigte Rechenzeit im Vergleich zur herkömmlichen, numerisch äußerst aufwendigen Formulierung des EXX-Verfahrens drastisch reduziert, dabei jedoch nur einen Bruchteil des Speicherplatzes benötigt.

Als Beispiel für die auf diese Weise deutlich erweiterten Anwendungsmöglichkeiten des EXX-Formalismus wurde untersucht, ob sich exakte Austauschpotentiale annähernd linear in Abhängigkeit von der Ladungsdichte verhalten, was von großer Bedeutung für die Verwendung von EXX-Pseudopotentialen ist. Es konnten erstmals DFT-Berechnungen mit exaktem Austauschpotential für GaN und InN durchgeführt werden, bei denen die komplette *semi-core*-Schale der Kationen (3. Schale bei Ga, 4. Schale bei In) explizit in der Valenz berücksichtigt wird. Diese wurden anschließend mit Rechnungen, die nur die 3*d*- bzw. 4*d*-Elektronen als Valenzelektronen behandeln, verglichen.

Ein weiterer Fokus der Arbeit lag auf der Entwicklung eines Konstruktionsschemas für *ab-initio*-Parameter im Modell starker Bindung (*Tight-Binding*-Modell) [122]. Die Methode basiert auf dem Konzept der Verallgemeinerten Wannier-Funktionen von Marzari und Vanderbilt [116], und erlaubt es, die elektronische Struktur (Eigenenergien und Eigenzustände) sehr großer Systeme mit mehreren Tausend Atomen in der Genauigkeit atomistischer Berechnungen auszuwerten. Dies kann etwa zur Berechnung der optischen Eigenschaften von Quantenpunkten eingesetzt werden, wobei dann auf DFT-EXX-Berechnungen zurückgegriffen werden kann, um eine korrekte Beschreibung der Bandlücke zu gewährleisten. Im Rahmen dieser Arbeit wurden effektive *ab-initio*-Parametersätze für das GaN- und InN-Volumensystem entwickelt und hinsichtlich ihrer Eignung zur Beschreibung der Bandstruktur sorgfältig getestet.

Abstract

In the present thesis, the variational EXX method (varEXX) [78, 79, 80] has been implemented into a pseudopotential plane-wave code (S/PHI/nX [81, 82]). So far this method has been applied only for localized systems. By integrating it into the plane-wave framework it has been made applicable to periodic systems for the first time. The conventional approach of EXX is computationally very demanding. Subsequently performed tests of the new method indicate that varEXX allows a tremendous performance speed up while only a fraction of the memory consumption is required compared to the conventional EXX formulation.

The efficient formulation of varEXX allows for a wide range of new applications. As an example it has been investigated whether exact exchange potentials can be considered as approximate linear functionals of the electron density. This is of significant importance concerning the practical use of EXX pseudopotentials. For the first time, DFT calculations using exact exchange pseudopotentials could be performed for GaN and InN with the complete semi-core-shell of the cations (3rd shell for Ga, 4th shell for In) being treated explicitly in the valency. Thereafter the results have been compared with those of calculations considering only the semi-core $3d$ - and $4d$ -electrons in the valency, respectively.

A second focus of the thesis is the development of a generation scheme for *ab-initio* tight-binding parameters [122]. The method employs the formalism of Generalized Wannier-Functions introduced by Marzari und Vanderbilt [116] and allows for a calculation of the electronic structure (eigenenergies and eigenstates) of huge systems with several thousands of atoms within the accuracy of the atomistic scale. This can be used for a computation of the optical properties of quantum dots based on foregoing DFT-EXX calculations in order to guarantee a correct description of the bandgap. Within the scope of this work effective *ab-initio* parametrizations for the GaN and InN bulk system have been developed and carefully tested with regard to their ability to an accurate description of the bandstructure.

Inhaltsverzeichnis

Einleitung	1
I Exakter Austausch	7
1 Das Vielteilchenproblem	9
1.1 Born-Oppenheimer-Näherung	10
1.2 Das Variationsprinzip	11
1.3 Das Hartree-Fock-Verfahren	12
1.4 Dichtefunktional-Theorie	16
1.4.1 Die Hohenberg-Kohn-Theoreme	17
1.4.2 Die Kohn-Sham-Gleichungen	18
1.4.3 Levy's <i>Constraint-Search</i> -Formulierung der DFT	21
1.5 Das Austausch-Korrelations-Funktional	22
1.5.1 Definition von Austausch- und Korrelationsenergie in der DFT	22
1.5.2 Lokale Dichtenäherung und verwandte Funktionale	23
1.5.3 Selbstwechselwirkungskorrektur	24
1.5.4 Exakte Darstellung des Austausch-Korrelationsfunktionals als Störreihe	25
1.6 Exakte Austauschpotentiale	28
1.6.1 Das Standard-EXX-Verfahren	28
1.6.2 Das OEP-Verfahren	31
1.6.3 Störungstheoretische Herleitung der OEP/EXX-Gleichung	32
2 Das variationelle EXX-Verfahren	35
2.1 Idee und mathematische Formulierung	36
2.2 Äquivalenz des variationellen Ansatzes zu OEP und Standard-EXX	38
2.3 Variation an Stelle direkter Inversion	39
2.4 Variationelles EXX und Hartree-Fock-Verfahren	40
3 Methoden zur numerischen Beschreibung periodischer Systeme	43
3.1 Symmetrien und periodische Randbedingungen	44
3.2 Numerische Integration in der Brillouin-Zone	45
3.3 Entwicklung nach ebenen Wellen	46
3.4 Näherung der starren Rumpfe	49
3.5 Pseudopotentiale	50
3.5.1 Konstruktion	51
3.5.2 Kleinman-Bylander-Darstellung	53

3.5.3	Nicht-lineare Rumpfkorrektur	53
4	Implementierung	55
4.1	Die Sternheimer-Gleichung	55
4.1.1	Numerische Lösung der Sternheimer-Gleichung: Umformulierung in ein Minimierungs-Problem	56
4.1.2	Minimierungsverfahren	59
4.1.2.1	Methode des steilsten Abstiegs (<i>Steepest Descent</i>)	60
4.1.2.2	Methode der konjugierten Gradienten (<i>Conjugate Gradient</i>)	63
4.1.2.3	Präkonditionierung	65
4.2	Induzierte Ladungsdichte	69
4.3	Minimierung der induzierten Ladungsdichte	69
4.3.1	Mixing, Selbstkonsistenz und Minimierung	72
4.3.2	Pulay-Mixing	74
4.3.3	Präkonditionierung	76
4.4	Der nicht-lokale Austauschoperator in ebenen Wellen	78
4.4.1	Behandlung der Singularität	78
4.4.2	Symmetrisierung	80
4.5	Variationelles EXX als selbstkonsistenter Zyklus	81
4.5.1	Optimierung des Algorithmus	82
4.5.2	Variationelles und Standard-EXX: Ein Vergleich des Iterationsprozesses	85
4.6	Symmetriegenäherte EXX-Methode	86
4.6.1	Exakte Symmetrisierung beim Standard-EXX-Verfahren	87
4.6.2	Symmetriegenähertes Standard-EXX-Verfahren	88
4.6.3	Übertragung auf das variationelle EXX-Verfahren	90
5	Variationelles und Standard-EXX: Ein Leistungsvergleich	93
6	Physikalische Bedeutung der induzierten Ladungsdichte	97
7	EXX-Pseudopotentiale: Linearisierung der Austauschwechselwirkung	101
II	Wannier-Funktionen	109
8	Verallgemeinerte Wannier-Funktionen als lokalisierte <i>ab-initio</i>-Basis im Modell starker Bindung	111
8.1	Verallgemeinerte Wannier-Funktionen	112
8.1.1	Definition	112
8.1.2	Lokalisierung	114
8.1.3	Maximale Lokalisierung	115
8.1.4	Initialisierung	117
8.2	<i>Ab-initio</i> -Parametrisierung im Modell starker Bindung	118
8.3	Lokalisierung <i>kein</i> Maß für die Güte der Basis	119
8.4	Optimierte Wannier-Basen für GaN und InN	124
	Zusammenfassung und Ausblick	129

Anhänge	133
A Fourier-Entwicklungen einiger wichtiger Größen	135
A.1 Das Coulomb-Potential	135
A.2 Der nicht-lokale Austauschoperator	135
A.3 Die lineare Antwort	138
B Die Greensche Funktion	140
C Die Sternheimer-Gleichung	143
C.1 Die Sternheimer-Gleichung aus dem Variationsprinzip von Ritz	143
C.2 Entartung	145
C.3 Die Sternheimer-Gleichung mit nicht-hermiteschem H_0	146
D Tensordarstellung der EXX-Methode	149
E Verallgemeinerung des EXX-Formalismus auf N-darstellbare Ladungsdichten	154
F Symmetriegenähertes EXX: Herleitung der variationellen Formulierung	156
G Die OEP/EXX-Gleichung aus dem Vielteilchenansatz von Görling und Levy	158
H Verallgemeinerte Wannier-Funktionen	164
H.1 Translation um einen Gittervektor	164
H.2 Aktualisierung der Überlapp-Matrix	164
H.3 Der Hamilton-Operator in der Darstellung Verallgemeinerter Wannier-Funktionen	165
H.4 Berechnung der Ein-Teilchen-Energien aus der Wannier-Darstellung	166
Literaturverzeichnis	169
Danksagung	177

Einleitung

Warum „exakte Austauschpotentiale“?

Der Kohn-Sham-Formalismus der Dichtefunktional-Theorie (DFT) [1, 2] ist die zurzeit wohl am weitesten verbreitete Methode für *ab-initio*-Berechnungen¹ im Bereich der Festkörperphysik. Obwohl er eine *prinzipiell* exakte Methode darstellt, ist man in der Praxis jedoch bislang² auf Näherungen angewiesen, wie etwa der Lokalen Dichte-Approximation (LDA) oder den verschiedenen Varianten der Generalisierten Gradientenkorrekturen (GGA). Zwar gestatten diese Ansätze sehr effiziente Berechnungen der Gesamtenergie eines Systems im Grundzustand und anderer damit zusammenhängender Größen — darauf begründet sich der große Erfolg der DFT —, allerdings besitzen sie auch eine Reihe schwerwiegender Nachteile. Konzeptionell ist die DFT eine reine Grundzustandstheorie; die mit ihr berechneten Ein-Teilchen-Energien lassen sich streng genommen nicht als reale Größen interpretieren. So werden etwa in LDA die Bandlücken von Halbleitern und Isolatoren um bis zu über 100% unterschätzt (vgl. Abb. 1 und 2). Ein weiterer Punkt ist die falsche Asymptotik des Kohn-Sham-Potentials endlicher Systeme, als Folge daraus wiederum erscheinen z.B. eine Reihe experimentell stabiler, negativ geladener Ionen (H^- , O^- , F^- , ...) in LDA als instabil [3, 4, 5]. Für einige Übergangsmetalle mit $d^{n-2}s^2$ -Konfiguration liefert ferner die Lokale Spin-Dichte-Approximation (LSD) fälschlicherweise die $d^{n-1}s^1$ -Konfiguration [6]. Schließlich sei aufgeführt, dass das exakte Kohn-Sham-Potential (genauer: dessen Austauschanteil, siehe unten) in Abhängigkeit von der Besetzungszahl eine so genannte „Ganzzahl-Diskontinuität“ aufweist, die in LDA/GGA nicht beschrieben werden kann [7, 8].

Mit Ausnahme des letzten Punktes ist die Ursache für die erwähnten Schwierigkeiten in der elektronischen Selbstwechselwirkung zu suchen [5]: Im Kohn-Sham-Formalismus wird das reale System zunächst auf ein fiktives System wechselwirkungsfreier Teilchen abgebildet. Sie bewegen sich in einem effektiven, lokalen Potential (dem Kohn-Sham-Potential), das alle elektronischen Wechselwirkungen des realen Systems in geeigneter Weise berücksichtigt. Eine korrekte Formulierung muss dabei beachten, dass ein einzelnes Elektron nur mit der Ladungsverteilung der anderen Elektronen, jedoch nicht mit sich selbst wechselwirkt. So enthält das effektive Potential zusätzlich zum Hartree-Term, der die elektrostatische Energie der gesamten Ladungsdichte beschreibt, noch das so genannte *Austausch-Korrelationspotential*, dessen Austauschanteil bei exakter Beschreibung die in der elektrostatischen Energie enthaltene elektronische Selbstwechselwirkung wieder kompensiert. Diese Kompensation gelingt jedoch mit den in LDA und GGA gemachten

¹Es werden nur die Naturkonstanten und die Geometrie des Systems zu Grunde gelegt, d.h., man verwendet dabei keine empirischen oder angepassten Parameter.

²Tatsächlich besteht nur sehr wenig Aussicht darauf, bei praktischen Anwendungen der DFT jemals auf Näherungsansätze verzichten zu können.

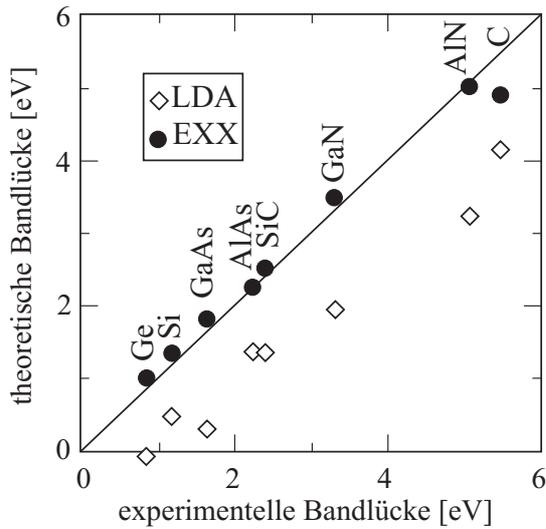


Abbildung 1: Vergleich der LDA- und EXX-Bandlücken verschiedener Halbleiter mit den experimentellen Werten. Nach Städele *et al.* [52].

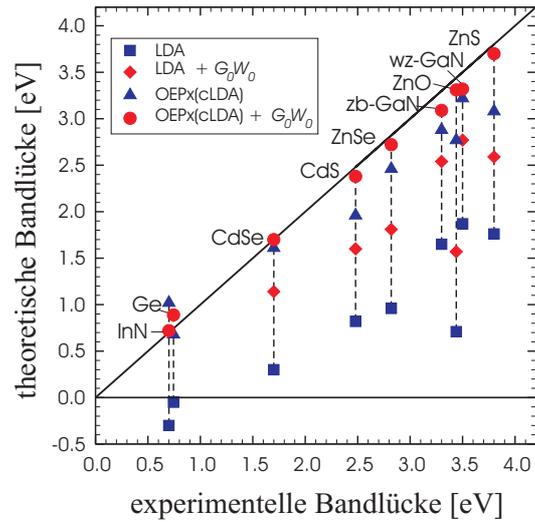


Abbildung 2: Vergleich theoretischer und experimenteller Bandlücken. OEPx(cLDA) bezeichnet hier EXX-Berechnungen mit LDA-Korrelation. Die LDA- und OEPx(cLDA)-Ergebnisse dienen zudem als Startpunkt einer GW-Rechnung (+GW). Nach Rinke *et al.* [171, 175].

Näherungsannahmen nur unvollständig.

Aus diesem Grund wurden verschiedene, über LDA/GGA hinausgehende Methoden entwickelt. Bei Selbstwechselwirkungskorrektur-Methoden (SIC) [5] etwa wird versucht, die unvollständige Kompensation der Selbstwechselwirkung durch zusätzliche Terme wieder auszugleichen. Vor allem bei lokalisierten Systemen führt SIC häufig zu einer deutlich verbesserten Beschreibung, der Zugang ist jedoch nicht invariant gegenüber unitären Transformationen der Wellenfunktionen, d.h. die Systembeschreibung ist abhängig von der Wahl der Basisfunktionen.

Ein sehr erfolgreiches Instrument zur Berechnung von Bandstrukturen ist ferner die Quasiteilchen-Theorie in Gestalt der GW-Näherung von Hedin [150, 151], die es erlaubt, Ergebnisse einer DFT-Näherungsrealisierung mit störungstheoretischen Methoden zu verbessern. Die Methode ist rechentechnisch sehr aufwendig, liefert aber wesentlich genauere Bandlücken als LDA/GGA. Da GW in der Regel nicht selbst-konsistent betrieben wird, hängt die Genauigkeit der Resultate von der Güte der Eingangsdaten ab. Es ist daher sinnvoll, elektronische Strukturen bereits innerhalb der DFT möglichst genau bestimmen zu können. Eine geeignete Möglichkeit dazu eröffnete sich mit der Entwicklung des *Exakten-Austausch-Verfahrens* (EXX) [41, 47, 51, 52], das auch den Schwerpunkt der vorliegenden Arbeit bildet. Die Kombination von EXX und GW erweist sich als äußerst präzise Strategie zur Bestimmung von Bandstrukturen (vgl. Abb. 2) und ist seit jüngster Zeit Gegenstand intensiver Forschung [170, 171, 172, 173, 174, 175].

Das EXX-Verfahren ermöglicht die Konstruktion des exakten DFT-Austauschpotentials. Dies hat gegenüber der Verwendung von Näherungspotentialen zwei entscheidende Vorteile: (i) Innerhalb der Austausch-Korrelationswechselwirkung kommt dem Austauschanteil die dominierende Rolle zu: er übertrifft die Korrelationsenergie um typischerweise eine Größenordnung [44, 45, 46]. (ii) Die exakte Kenntnis des Austauschpotentials genügt,

um die elektronische Selbstwechselwirkung vollständig zu eliminieren. Formal greift der Formalismus dabei zurück auf das Austauschfunktional der Hartree-Fock-Theorie [12, 13], in der die Selbstwechselwirkung konzeptionell in völlig korrekter Weise gehandhabt wird. Während jedoch das Hartree-Fock-Verfahren auf ein nicht-lokales, orbitalabhängiges Austauschpotential führt, ist das EXX-Potential lokal, in Übereinstimmung mit dem Ansatz des Kohn-Sham-Formalismus. Anders als beim Hartree-Fock-Verfahren, das Bandlücken typischerweise um einen Faktor 2–3 überschätzt, findet man zudem eine sehr gute Übereinstimmung der in EXX berechneten Bandlücken mit den experimentellen Werten [52] (vgl. Abb. 1). Als entscheidender Nachteil des EXX-Formalismus erweist sich allerdings der enorme technische Aufwand, der zu seiner numerischen Realisierung erforderlich ist, sowohl hinsichtlich der benötigten Speicherkapazität als auch in Bezug auf die Rechenzeit (im Vergleich mit entsprechenden LDA/GGA-Rechnungen beträgt der Mehraufwand 2–3 Größenordnungen). So ist u.a. die Greensche Funktion und der lineare Antwortoperator zu berechnen, wozu neben den letztlich interessierenden besetzten Zuständen auch unbesetzte Zustände berücksichtigt werden müssen. Bei realistischen Berechnungen von Halbleiterkristallen, etwa GaAs oder Gruppe-III-Nitriden, findet man dabei typischerweise, dass auf jeden besetzten Zustand ca. 100 unbesetzte Zustände kommen. Dies beschränkt die Anwendungsmöglichkeiten dieses Verfahrens auf einfachste Kristallstrukturen (mit weniger als etwa 5 Atomen pro Einheitszelle).

Es wurde jedoch 2002 von Wahn und Neugebauer [78] sowie 2003 von Kümmel und Perdew [79, 80] eine alternative Formulierung des EXX-Formalismus vorgeschlagen, die gänzlich ohne unbesetzte Zustände auskommt. Das EXX-Potential wird dabei mit Hilfe variationeller Algorithmen bestimmt, weshalb dieser Zugang im Folgenden als *variationalles EXX-Verfahren* (varEXX) bezeichnet werden soll. Im Vergleich zur ursprünglichen Fassung der EXX-Methode ist varEXX nicht nur vollkommen konsistent mit dem Konzept der DFT als einer reinen Grundzustandstheorie, sondern erlaubt eine wesentliche Reduktion des numerischen Aufwandes; die Konstruktion der Greenschen Funktion und des linearen Antwortoperators entfällt völlig. Darüber hinaus lässt sich die Schlüsselgröße des varEXX-Konzepts, die „induzierte Ladungsdichte“, in einfacher und anschaulicher Weise physikalisch interpretieren. Das neue Verfahren ist bisher erfolgreich im Falle von Atomen [78, 79] und Clustern [80] getestet worden.

Ziele der Arbeit

Hauptanliegen der vorliegenden Arbeit war es nun, varEXX in den Pseudopotential-Ebene-Wellen-Code S/PHI/nX [81, 82] zu implementieren, um die Methode so erstmalig auch für die Anwendung auf periodische Systeme verfügbar zu machen. Der Algorithmus wurde dann an einer Reihe von Halbleitermaterialien getestet und seine Effizienz (CPU-Zeit, Speicheranforderungen) mit der ursprünglichen EXX-Formulierung verglichen. Dabei konnte auf eine bereits in S/PHI/nX vorhandene Implementierung des Standard-EXX-Verfahrens³ zurückgegriffen werden.

Durch den neuen Code eröffnet sich nun die Möglichkeit, Fragestellungen anzugehen, die über das bisher mit dem Standard-EXX-Verfahren Machbare hinausgehen. Als ein Beispiel dafür wurde in dieser Arbeit untersucht, ob das exakte Austauschpotential ein (zumindest annähernd) lineares Funktional der Ladungsdichte ist. Dies ist von großer Bedeutung insbesondere im Hinblick auf die Verwendung von EXX-Pseudopotentialen,

³Es handelt sich dabei allerdings um eine Näherung des Standard-EXX-Verfahrens, die numerisch effizienter als das exakte Verfahren realisiert werden kann.

bei denen ein Teil der so genannten *semi-core*-Elektronen als Valenzelektronen behandelt wird, der andere Teil jedoch im Rumpf verbleibt.

Ein weiterer Fokus dieser Arbeit lag auf der Entwicklung einer Methode zur Konstruktion von *ab-initio*-Parametern im *Tight-Binding*-Modell. Der Zugang erlaubt es, die Genauigkeit von Berechnungen auf atomistischer Skala (etwa Berechnungen in EXX oder EXX+GW⁴) auf die Behandlung von Systemen mesoskopischer Größenordnung zu übertragen. So lassen sich etwa die optischen Eigenschaften von Quantenpunkten untersuchen, die wegen der großen Anzahl beteiligter Atome (typischerweise 10^3 bis 10^5) einerseits nicht mehr in einem vollständigen Basissatz (ebene Wellen, LAPW, ...) gehandhabt werden können, für die andererseits aber auch eine genaue Beschreibung der elektronischen Struktur (d.h., der Eigenwerte und Eigenzustände) vonnöten ist.

Aufbau der Arbeit

Der erste Teil der Arbeit ist der variationellen Formulierung des Exakten-Austausch-Verfahrens (varEXX) sowie dessen Implementierung und Anwendung gewidmet.

In Kapitel 1 wird zunächst die Austauschenergie sowie die elektronische Selbstwechselwirkung im Rahmen des Hartree-Fock-Formalismus definiert. Anschließend wird eine kurze Einführung in die Dichtefunktional-Theorie gegeben. Es wird gezeigt, wie sich das quantenmechanische Vielteilchenproblem mit Hilfe der Hohenberg-Kohn-Theoreme in Form einer effektiven Ein-Teilchen-Gleichung (der Kohn-Sham-Gleichung) schreiben lässt, und es werden die gängigsten Näherungsrealisierungen des Austausch-Korrelationsfunctionals (LDA und GGA) vorgestellt. Einen mathematisch strengeren, dabei aber sogar anschaulicheren Zugang zur DFT bietet die *Constraint-Search*-Formulierung von Levy; sie ermöglicht darüber hinaus die Ableitung einer exakten Reihenentwicklung des Austausch-Korrelationspotentials (Görling-Levy-Störungstheorie), die allerdings außerordentlich kompliziert ist. Schließlich wird gezeigt, wie ein Abbruch dieser Reihe nach dem linearen Glied auf den (Standard-)EXX-Formalismus führt.

In Kapitel 2 wird das variationelle EXX-Verfahren eingeführt. Es wird die zentrale Größe des Konzepts, die „induzierte Ladungsdichte“ definiert und ihr Zusammenhang zum EXX-Potential über ein Variationsprinzip formuliert. Überdies wird erläutert, wie es mit vergleichsweise einfachen variationellen Methoden gelingen kann, den komplizierten Zusammenhang zwischen Austauschpotential und Ladungsdichte zu erfassen, ohne dabei auf unbesetzte Zustände und aufwendig zu berechnende Größen wie der Greenschen Funktion oder der linearen Antwort (siehe oben) zurückgreifen zu müssen.

Kapitel 3 stellt das notwendige Rüstzeug zur numerischen Beschreibung periodischer Systeme zur Verfügung, insbesondere mit Hinblick auf die Verwendung einer Ebene-Wellen-Basis. Dabei wird auch das Konzept der Pseudopotentiale kurz skizziert.

Kapitel 4 beschreibt detailliert die Implementierung des variationellen EXX-Verfahrens in das DFT-Programm-Paket S/PHI/nX. Zunächst wird die Sternheimer-Gleichung, d.i. die Grundgleichung der Rayleigh-Schrödingerschen Störungstheorie in 1. Ordnung, als Minimierungsaufgabe formuliert, und es werden geeignete numerische Minimierungsverfahren sowie Prädiktionierungen diskutiert. Dies ermöglicht die effiziente Berechnung der induzierten Ladungsdichte. Als elegante Möglichkeit, aus dieser Größe das EXX-Potential zu erhalten, werden ferner *Mixing*-Verfahren diskutiert. Schließlich wird die Optimierung

⁴Es ist dabei allerdings zu beachten, dass die *GW*-Methode nur Quasiteilchen-Energien, aber keine Zustände liefert.

des Codes beschrieben und bereits auf Möglichkeiten zur weiteren Steigerung der Effizienz hingewiesen.

In Kapitel 5 folgt ein tabellarischer Vergleich zwischen der Leistung (CPU-Zeit, Speicherbedarf) des varEXX-Verfahrens und der des Standard-EXX-Verfahrens.

Eine anschauliche Interpretation der induzierten Ladungsdichte und ihr Bezug zur elektronischen Selbstwechselwirkung wird in Kapitel 6 anhand der Beispiele GaAs und Si aufgezeigt.

Kapitel 7 untersucht, ob das exakte Austauschpotential näherungsweise eine lineare Abhängigkeit von der Ladungsdichte aufweist (siehe oben). Die Resultate werden mit den Ergebnissen von LDA+GW-Rechnungen nach Rohlfing *et al.* [152] verglichen und hinsichtlich ihrer Bedeutung für die Verwendung von EXX-Pseudopotentialen in zukünftigen Berechnungen diskutiert.

Der zweite Teil der Arbeit beschäftigt sich mit der Konstruktion von *ab-initio*-Parametern im Rahmen des *Tight-Binding*-Modells (siehe oben). Dazu wird zunächst eine kurze Einführung in die Theorie der Verallgemeinerten Wannier-Funktionen nach Marzari und Vanderbilt [116] gegeben, um darauf aufbauend das Konstruktionsschema für die *ab-initio*-Parametrisierung gemäß Wahn und Neugebauer [122] herzuleiten. Das Verfahren wird an den Halbleitern GaAs, GaN und InN getestet, die Ergebnisse werden kritisch diskutiert.

Es schließen sich eine Zusammenfassung der Arbeit und ein Ausblick an.

Der Anhang enthält die schrittweise Herleitung einer Reihe von im Haupttext verwendeter Formeln, um so ein wirkliches Nachvollziehen der Zusammenhänge zu vereinfachen. Darüber hinaus finden sich dort Ergänzungen zu einigen Abschnitten, die für deren Verständnis zwar nicht unmittelbar benötigt werden, jedoch die Thematik abrunden (z.B. der Zusammenhang zwischen EXX und Sham-Schlüter-Gleichung). Es wird zudem eine mathematisch etwas abstraktere Formulierung des EXX-Verfahrens gegeben (mit Hilfe einer Verallgemeinerung des linearen Antwortoperators), die zwar nicht für eine Implementierung geeignet ist, aber auf Grund ihrer Kürze äußerst hilfreich etwa bei der Diskussion des Potential-Ladungsdichte-*Mixings* ist sowie einige Sachverhalte, z.B. den linearen Zusammenhang zwischen Austauschpotential und induzierter Ladungsdichte, unmittelbar verdeutlicht.

In der gesamten Arbeit werden *atomare Einheiten* verwendet, d.h. also

$$\hbar = 1, \quad m_e = 1, \quad \frac{e^2}{4\pi\epsilon_0} = 1.$$

Energie und Länge werden daher, wenn nicht anders vermerkt, in der Einheit *Hartree* bzw. in *Bohr* angegeben. Um Verwirrungen zu vermeiden, wird ferner der Laplace-Operator stets mit ∇^2 bezeichnet und das Symbol Δ ausschließlich zur Kennzeichnung von Differenzen und der induzierten Ladungsdichte ($\Delta\rho$) verwendet.

Teil I

Exakter Austausch

Kapitel 1

Das Vielteilchenproblem

Um den Aufbau der Materie physikalisch-mathematisch beschreiben zu können, bedarf es zunächst eines geeigneten Modells. Das den folgenden Betrachtungen zu Grunde liegende Modell beruht auf den in der Chemie üblichen Vorstellungen: Die Materie wird als System von Elektronen und Kernen (oder Ionen) aufgefasst — der innere Aufbau der Kerne aus Protonen und Neutronen sowie noch feinere Strukturen, etwa Quarks oder Strings, werden jedoch gemeinhin nicht in die Überlegungen mit einbezogen. Besteht das betrachtete Modellsystem nun aus N Elektronen und M Kernen, und vernachlässigt man zudem relativistische Effekte, so wird es im Prinzip vollständig durch die Viel-Teilchen-Schrödinger-Gleichung

$$H \Phi(\mathbf{r}_1, \dots, \mathbf{r}_N, \mathbf{R}_1, \dots, \mathbf{R}_M, t) = i \frac{\partial \Phi(\mathbf{r}_1, \dots, \mathbf{r}_N, \mathbf{R}_1, \dots, \mathbf{R}_M, t)}{\partial t} \quad (1.1)$$

beschrieben, wobei natürlich noch geeignete Rand- oder Anfangsbedingungen beachtet werden müssen. Die Wellenfunktion Φ des Gesamtsystems hängt dabei von den Ortskoordinaten der Elektronen \mathbf{r}_i , denen der Kerne \mathbf{R}_j und der Zeit t ab. Wirken ferner keine äußeren Kräfte auf das System, setzt sich der Hamilton-Operator

$$H := T_e + T_n + V_{ee} + V_{en} + V_{nn} \quad (1.2)$$

zusammen aus den Operatoren der kinetischen Energie der Elektronen und Kerne, in Ortsdarstellung gegeben durch

$$T_e := \sum_{i=1}^N -\frac{1}{2} \nabla_{\mathbf{r}_i}^2, \quad T_n := \sum_{I=1}^M -\frac{1}{2m_I} \nabla_{\mathbf{R}_I}^2$$

mit den Kernmassen m_I , sowie den Coulomb-Wechselwirkungen der Teilchen

$$V_{ee} := \frac{1}{2} \sum_{\substack{i,j=1 \\ i \neq j}}^N \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}, \quad V_{en} := \sum_{i=1}^N \sum_{I=1}^M \frac{-Z_I}{|\mathbf{r}_i - \mathbf{R}_I|}, \quad V_{nn} := \frac{1}{2} \sum_{\substack{I,J=1 \\ I \neq J}}^M \frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|}$$

mit den Kernladungszahlen Z_I . Da (1.2) nicht explizit von der Zeit abhängt, lassen sich die Lösungen Φ der Schrödinger-Gleichung (1.1) als Linearkombinationen stationärer Zustände der Form $\Psi_\alpha(\mathbf{r}_1, \dots, \mathbf{r}_N, \mathbf{R}_1, \dots, \mathbf{R}_M) \exp(-iE_\alpha t)$ schreiben, die sich aus der Eigenwertgleichung

$$H \Psi_\alpha = E_\alpha \Psi_\alpha, \quad (1.3)$$

der zeitunabhängigen Schrödinger-Gleichung, ergeben. Als Eigenwertgleichung wird (1.3) im Allgemeinen durch verschiedene Paare von Eigenfunktionen Ψ_α und Eigenenergien E_α , jeweils unterschieden durch den Index α , gelöst. Für den Hamilton-Operator H gibt es eine kleinste Eigenenergie, die als *Grundzustandsenergie* bezeichnet wird. Die zugehörige Eigenfunktion ist der *Grundzustand* (vgl. auch Abschnitt 1.2). Bei Eigenfunktionen zu höheren Energien spricht man von *angeregten Zuständen*.

Gleichung (1.3) kann jedoch nur für einfachste Systeme wie dem freien Wasserstoffatom [9] analytisch gelöst werden. Für sehr große Systeme, z.B. einen Festkörper, für den die Anzahl der zu betrachtenden Teilchen in der Größenordnung der Avogadro-Konstante $N_A = 6.023 \cdot 10^{23}$ liegt, lässt sich (1.3) beim derzeitigen Stand der Computertechnik jedoch auch numerisch nicht mehr direkt lösen. Es müssen daher (weitere) vereinfachende Näherungsannahmen gemacht werden. Solche Näherungsansätze können ggf. auch mit einer (exakten) Transformation des Problems — etwa auf ein System wechselwirkungsfreier Teilchen, vgl. Abschnitt 1.4.2 — kombiniert werden. Eine wesentliche Vereinfachung besteht in der Entkopplung von Elektronen- und Kernbewegung, die im folgenden Abschnitt skizziert werden soll.

1.1 Born-Oppenheimer-Näherung

Ausgangspunkt der *Born-Oppenheimer-Näherung* oder *adiabatischen Näherung* [10] ist der große Unterschied zwischen der Elektronenmasse m_e und den Kernmassen M_I : Das Verhältnis m_e/M_I liegt typischerweise in der Größenordnung $10^{-4} - 10^{-5}$. Klassisch betrachtet besitzen die Kerne damit eine wesentlich größere Trägheit als die Elektronen. Daher können Letztere nahezu „instantan“ auf die Bewegung der Kerne reagieren, d.h., ihr Zustand wird der Kernbewegung „adiabatisch“ folgen. Der elektronische Zustand ist folglich näherungsweise nur durch die momentane Konfiguration der Kerne bestimmt, die zu dem betrachteten Zeitpunkt ein festes Potential darstellen. Unter diesen Annahmen lässt sich die Wellenfunktion Ψ des Gesamtsystems in einen elektronischen Anteil Ψ_e und einen nuklearen Anteil Ψ_n separieren,

$$\Psi(\mathbf{r}, \mathbf{R}) = \Psi_e(\mathbf{r}, \mathbf{R})\Psi_n(\mathbf{R}), \quad (1.4)$$

wobei hier mit $\mathbf{r} := (\mathbf{r}_1, \dots, \mathbf{r}_N)$ die Elektronenkoordinaten und mit $\mathbf{R} := (\mathbf{R}_1, \dots, \mathbf{R}_M)$ die Kernkoordinaten zusammengefasst sind. Der elektronische Anteil hängt dabei nur parametrisch von \mathbf{R} ab und ist Lösung des Eigenwertproblems

$$H_e\Psi_e(\mathbf{r}, \mathbf{R}) = E_e\Psi_e(\mathbf{r}, \mathbf{R}) \quad (\text{für } \mathbf{R} \in \mathbb{R}^{3M} \text{ fest}). \quad (1.5)$$

Der Hamilton-Operator ist hier durch

$$H_e(\mathbf{r}, \mathbf{R}) := \sum_{i=1}^N \left[-\frac{1}{2}\nabla_{\mathbf{r}_i}^2 + v_{\text{ext}}(\mathbf{r}_i, \mathbf{R}) \right] + \frac{1}{2} \sum_{\substack{i,j \\ i \neq j}}^{1 \dots N} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \quad (1.6)$$

gegeben, wobei der Einfluss der Kerne in dem Potential

$$v_{\text{ext}}(\mathbf{r}_i, \mathbf{R}) := \sum_{I=1}^M \frac{-Z_I}{|\mathbf{r}_i - \mathbf{R}_I|} \quad (1.7)$$

zusammengefasst ist. Die Bezeichnung v_{ext} wurde hier im Vorgriff auf die Abschnitte 1.3 und 1.4 gewählt, in denen (1.7) das „äußere“ oder „externe“ Potential darstellt.

Eine ähnliche Gleichung wie (1.5) lässt sich nun auch für die Bewegung der Kerne aufstellen. Dazu beachte man, dass die Kerne auf Grund ihrer vergleichsweise hohen Trägheit die Elektronenbewegung nur noch in gemittelter Form „wahrnehmen“ können: Das von den Elektronen auf die Kerne wirkende Potential ergibt sich im Rahmen der Born-Oppenheimer-Näherung zu $\langle H_e(\mathbf{R}) \rangle$, wobei die Mittelung über die elektronische Wellenfunktion Ψ_e erfolgt. Der Hamilton-Operator der Kernbewegung ist dann gegeben durch

$$H_n(\mathbf{R}) := \sum_{I=1}^M -\frac{1}{2m_I} \nabla_{\mathbf{R}_I}^2 + \frac{1}{2} \sum_{\substack{I,J=1 \\ I \neq J}}^M \frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|} + \langle H_e(\mathbf{R}) \rangle. \quad (1.8)$$

Die Born-Oppenheimer-Näherung funktioniert dann, wenn die Elektron-Phonon-Kopplung vernachlässigt werden kann. Für die Fragestellungen dieser Arbeit, die sich auf die Berechnung des elektronischen Grundzustandes von Halbleiter-Volumensystemen beziehen, ist sie daher eine exzellente Näherung. Beispiele für Phänomene, bei denen die Elektron-Phonon-Kopplung wesentlich ist, und für die sich die Born-Oppenheimer-Näherung folglich nicht eignet, sind hochenergetische Streuprozesse von Atomen an Festkörperoberflächen oder die BCS-Theorie der Supraleitung [11].

1.2 Das Variationsprinzip

Ausgangspunkt für viele Methoden zur näherungsweisen Beschreibung eines quantenmechanischen Viel-Teilchen-Systems ist oft nicht die Schrödinger-Gleichung selbst in Gestalt eines Eigenwertproblems (1.3) oder (1.5), sondern eine äquivalente Formulierung des Problems als „Variationsaufgabe“. Präziser ausgedrückt besteht die Aufgabe darin, diejenigen Funktionen Ψ_α (oder allgemeiner Vektoren eines geeigneten Hilbertraums \mathcal{H}) zu finden, die das Funktional¹

$$E[\Psi] := \frac{\langle \Psi | H | \Psi \rangle}{\langle \Psi | \Psi \rangle} \quad (\Psi \in \mathcal{H} \setminus \{0\}) \quad (1.9)$$

stationär werden lassen, für die es also ein Extremum oder einen Sattelpunkt besitzt:

$$\left. \frac{\delta}{\delta \Psi^*} E[\Psi] \right|_{\Psi = \Psi_\alpha} \stackrel{!}{=} 0. \quad (1.10)$$

Ausführen der Variationsableitung² liefert unmittelbar die Eigenwertgleichung (1.3) mit $E_\alpha = E[\Psi_\alpha]$, vgl. z.B. Ref. 92. Wegen des Nenners in (1.9) braucht hier keine Normierung der Ψ vorausgesetzt werden. Alternativ kann man auch einfacher $E[\Psi] := \langle \Psi | H | \Psi \rangle$ definieren, muss dann aber explizit normierte Ψ fordern:

$$\frac{\delta}{\delta \Psi^*} \left[\langle \Psi | H | \Psi \rangle - \varepsilon \cdot (\langle \Psi | \Psi \rangle - const.) \right] \stackrel{!}{=} 0. \quad (1.11)$$

Auch hier führt Ableiten direkt auf (1.3), wobei die Gleichung aber nur dann erfüllt werden kann, wenn der Lagrange-Parameter ε einem Eigenwert E_α entspricht. Ist H ein nach unten

¹In (1.9) wird keine Normierung der Vektoren vorausgesetzt. Um eine Division durch Null zu vermeiden, muss daher der Nullvektor im Definitionsbereich ausgeschlossen werden.

²Zur Definition der Ableitung im Komplexen vgl. Ref. 126.

beschränkter³ Operator, besitzt (1.9) ein globales Minimum Ψ_0 . Im Falle des Hamilton-Operators spricht man dann vom Grundzustand des Systems, der zugehörige Eigenwert

$$E_0 = E[\Psi_0] \leq \min_{\Psi} E[\Psi] \quad (1.12)$$

ist die Grundzustandsenergie.

Eine einfache Näherungsmethode besteht darin, nicht alle $\Psi \in \mathcal{H} \setminus \{0\}$ zur Variation zuzulassen, sondern nur in einem bestimmten Unterraum zu suchen. Es ist klar, dass man dann im Allgemeinen die tatsächlichen stationären Punkte „verpassen“ wird; im Falle des Grundzustandes etwa wird man nur eine obere Schranke für E_0 erhalten. Diese Methode wird auch *Ritzsches Variationsverfahren* genannt. Je größer der betrachtete Unterraum, desto besser wird die Näherung. Im Grenzfall einer vollständigen Basis geht das Ritzsche Verfahren in das exakte Variationsprinzip (1.10) oder (1.11) über.

Auf dem Ritzschen Ansatz basiert auch die im folgenden Abschnitt beschriebene Hartree-Fock-Methode: hier wird der Hilbertraum aller antisymmetrischen N -Elektronen-Funktionen eingeschränkt auf die Teilmenge der einfachen Slater-Determinanten. Die in Abschnitt 1.4 eingeführte Dichtefunktional-Theorie kann dagegen als spezielle Formulierung des exakten Variationsprinzips aufgefasst werden (vgl. Abschnitt 1.4.3), die für praktische Anwendungen notwendigen Näherungen erfolgen dort an anderer Stelle (Abschnitt 1.5.2). Schließlich lässt sich auch die Umformulierung der im variationellen EXX-Verfahren verwendeten Sternheimer-Gleichung in ein Minimierungsproblem (Abschnitt 4.1.1) über das Variationsprinzip verstehen (vgl. Anhang C.1).

1.3 Das Hartree-Fock-Verfahren

Eine der ersten Methoden zur näherungsweise Beschreibung des quantenmechanischen Vielteilchenproblems (1.5) wurde um 1930 von Hartree und Fock entwickelt [12, 13]. Dem *Hartree-Fock-Verfahren* kommt in dieser Arbeit eine besondere Bedeutung zu, da hier bereits der Ausdruck für die Austauschenergie abgeleitet (definiert) wird, der später bei der Konstruktion exakter Austauschpotentiale im Rahmen der Dichtefunktional-Theorie eine zentrale Rolle spielt. Es sei daher an dieser Stelle kurz umrissen, eine detailliertere Darstellung findet sich z.B. in den Referenzen 21, 92 und 127.

Eine Folge des Paulischen Prinzips der Ununterscheidbarkeit identischer Teilchen ist, dass eine Wellenfunktion, die dem Zustand eines N -Fermionen-Systems entspricht, antisymmetrisch bzgl. der Vertauschung zweier Teilchenkoordinaten sein muss [92, 127]. Ein geeigneter Hilbertraum wird dann von den *Slater-Determinanten* aufgespannt. Ist mit $\{\phi_1, \phi_2, \dots\}$ eine vollständige Orthonormalbasis in einem Ein-Teilchen-Hilbertraum gegeben, lässt sich aus jedem Satz $\{\phi_{k_i}\}_{i \in \{1, \dots, N\}}$ von N paarweise verschiedenen Basisfunktionen eine Slater-Determinante konstruieren, die dann durch

$$\Psi_{k_1 \dots k_N}^{\text{SD}}(\mathbf{x}_1, \dots, \mathbf{x}_N) := \frac{1}{\sqrt{N!}} \begin{vmatrix} \phi_{k_1}(\mathbf{x}_1) & \phi_{k_2}(\mathbf{x}_1) & \cdots & \phi_{k_N}(\mathbf{x}_1) \\ \phi_{k_1}(\mathbf{x}_2) & \phi_{k_2}(\mathbf{x}_2) & \cdots & \phi_{k_N}(\mathbf{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_{k_1}(\mathbf{x}_N) & \phi_{k_2}(\mathbf{x}_N) & \cdots & \phi_{k_N}(\mathbf{x}_N) \end{vmatrix} \quad (1.13)$$

³Ein nach unten beschränkter linearer Operator besitzt einen kleinsten Eigenwert. Hamilton-Operatoren — in der Mathematik auch häufig Schrödinger-Operatoren genannt — zu physikalisch sinnvollen Systemen sind gewöhnlich nach unten beschränkt, weil ein physikalisches System im Allgemeinen einen Grundzustand besitzt.

gegeben ist. Dabei sollen die Ein-Teilchen-Zustände aus einem Orts- und einem Spinanteil zusammengesetzt sein, also $\phi_k(\mathbf{x}_i) := \psi_k(\mathbf{r}_i)\chi_{\sigma_k}(s_i)$ mit $\mathbf{x}_i := (\mathbf{r}_i, s_i)$ und der Spinquantenzahl σ_k des k -ten Zustandes. Der Faktor vor der Determinante sorgt für die richtige Normierung $\langle \Psi^{\text{SD}} | \Psi^{\text{SD}} \rangle = 1$. Die Slater-Determinanten bilden ein vollständiges Orthonormalsystem, die Eigenfunktionen des Hamilton-Operators (1.6) lassen sich daher nach ihnen entwickeln. Die Näherung des Hartree-Fock-Formalismus besteht nun darin, dass man als Ansatz für diese Eigenfunktionen nur *eine* Slater-Determinante verwendet. Dazu setzt man (1.13) in das Energiefunktional (1.9) ein. Dann hat man

$$E_{\text{HF}}[\Psi^{\text{SD}}] = E_{\text{T}}[\Psi^{\text{SD}}] + E_{\text{ext}}[\Psi^{\text{SD}}] + E_{\text{H}}[\Psi^{\text{SD}}] + E_{\text{x}}[\Psi^{\text{SD}}] \quad (1.14)$$

mit der kinetischen Energie

$$E_{\text{T}}[\Psi^{\text{SD}}] = \left\langle \Psi^{\text{SD}} \left| \sum_{i=1}^N -\frac{1}{2} \nabla_i^2 \right| \Psi^{\text{SD}} \right\rangle = \sum_{i=1}^N \int \psi_{k_i}^*(\mathbf{r}_i) \left(-\frac{1}{2} \nabla_i^2 \right) \psi_{k_i}(\mathbf{r}_i) d^3 r_i \quad (1.15)$$

und der potentiellen Energie

$$E_{\text{ext}}[\Psi^{\text{SD}}] = \sum_{i=1}^N \int \psi_{k_i}^*(\mathbf{r}_i) v_{\text{ext}}(\mathbf{r}_i) \psi_{k_i}(\mathbf{r}_i) d^3 r_i. \quad (1.16)$$

Die Energie der Elektron-Elektron-Wechselwirkung ergibt sich hingegen komplizierter als Summe des *Hartree-Funktional*s

$$E_{\text{H}}[\Psi^{\text{SD}}] = \frac{1}{2} \sum_{i,j}^{1\dots N} \int \frac{\psi_{k_i}^*(\mathbf{r}) \psi_{k_j}^*(\mathbf{r}') \psi_{k_j}(\mathbf{r}') \psi_{k_i}(\mathbf{r})}{|\mathbf{r} - \mathbf{r}'|} d^3 r d^3 r' \quad (1.17)$$

und des so genannten *Austauschfunktional*s

$$E_{\text{x}}[\Psi^{\text{SD}}] = -\frac{1}{2} \sum_{i,j}^{1\dots N} \delta_{\sigma_i \sigma_j} \int \frac{\psi_{k_i}^*(\mathbf{r}) \psi_{k_j}^*(\mathbf{r}') \psi_{k_i}(\mathbf{r}') \psi_{k_j}(\mathbf{r})}{|\mathbf{r} - \mathbf{r}'|} d^3 r d^3 r'. \quad (1.18)$$

Die nach dem Pauli-Prinzip notwendige Antisymmetrie der N -Elektronen-Wellenfunktionen hat hier das Funktional E_{x} ins Spiel gebracht, das die so genannte *Austauschwechselwirkung* beschreibt, zu der kein klassisches Analogon existiert. Der Name rührt daher, dass die Koordinaten \mathbf{r} und \mathbf{r}' , von denen die Funktionen ψ_{k_i} bzw. ψ_{k_j} abhängen, in E_{x} gegenüber dem Hartree-Funktional E_{H} gerade „ausgetauscht“ worden sind. Dieser „Koordinatenaustausch“ findet natürlich auch im Spinanteil der Wellenfunktionen statt und führt dann zu dem in (1.18) zusätzlich auftretenden Kronecker-Symbol $\delta_{\sigma_i \sigma_j}$, womit E_{x} nur bei parallelen Spins eine Rolle spielt.

Bei der anschließenden Minimierung gemäß (1.10) ist zu beachten, dass die Orthonormalität der Ein-Teilchen-Funktionen erhalten bleibt. Dies kann mit Hilfe von Lagrange-Multiplikatoren geschehen:

$$\frac{\delta}{\delta \psi_k^*(\mathbf{r})} \left[E_{\text{HF}}[\Psi^{\text{SD}}] - \sum_{i,j}^{1\dots N} \lambda_{k_i k_j} \int \psi_{k_i}^*(\mathbf{r}) \psi_{k_j}(\mathbf{r}) d^3 r \right] \stackrel{!}{=} 0. \quad (1.19)$$

Hier sollen die ψ_k jeweils einen der besetzten Zustände $\psi_{k_1}, \dots, \psi_{k_N}$ kennzeichnen. Die Variationsableitungen der Energiefunktionale (1.15) und (1.16) lauten dann zusammengefasst

$$\frac{\delta(E_{\text{T}} + E_{\text{ext}})}{\delta \psi_k^*(\mathbf{r})} = \left[-\frac{1}{2} \nabla^2 + v_{\text{ext}}(\mathbf{r}) \right] \psi_k(\mathbf{r}). \quad (1.20)$$

Bei der Hartree-Energie E_H ergibt sich

$$\frac{\delta E_H}{\delta \psi_k^*(\mathbf{r})} = \sum_{i=1}^N \int \frac{\psi_{k_i}^*(\mathbf{r}') \psi_{k_i}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \psi_k(\mathbf{r}) d^3 r', \quad (1.21)$$

und für die Austauschenergie E_x findet man

$$\frac{\delta E_x}{\delta \psi_k^*(\mathbf{r})} = - \sum_{i=1}^N \delta_{\sigma\sigma_i} \int \frac{\psi_{k_i}^*(\mathbf{r}') \psi_k(\mathbf{r}') \psi_{k_i}(\mathbf{r})}{|\mathbf{r} - \mathbf{r}'|} d^3 r', \quad (1.22)$$

wobei σ die Spinrichtung von ψ_k bezeichnen soll. Schließlich muss auch der Term für die Nebenbedingungen in (1.19) berücksichtigt werden:

$$\frac{\delta}{\delta \psi_k^*(\mathbf{r})} \sum_{i,j}^{1\dots N} \lambda_{k_i k_j} \int \psi_{k_i}^*(\mathbf{r}) \psi_{k_j}(\mathbf{r}) d^3 r = \sum_{i=1}^N \lambda_{k k_i} \psi_{k_i}(\mathbf{r}). \quad (1.23)$$

Durch eine unitäre Transformation der Ein-Teilchen-Wellenfunktionen,⁴ die so gewählt wird, dass die Matrix $(\lambda_{k k_i})$ diagonal wird, entfällt die Summation über die i 's in der letzten Gleichung. Die Ausdrücke (1.20) – (1.23) liefern dann zusammengesetzt die Hartree-Fock-Gleichungen in ihrer „kanonischen Form“:

$$\left[-\frac{1}{2} \nabla^2 + v_{\text{ext}} + v_H[\rho] + V_{\text{Fock}}[\{\psi_i\}] \right] \psi_k = \varepsilon_k \psi_k \quad (k \in \{1, \dots, N\}). \quad (1.24)$$

Hier wurde das *Hartree-Potential* oder *elektrostatistische Potential*

$$v_H[\rho](\mathbf{r}) := \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3 r' \quad (1.25)$$

als von der Ladungsdichte

$$\rho(\mathbf{r}) := \sum_{i=1}^N |\phi_{k_i}(\mathbf{r})|^2 \quad (1.26)$$

abhängiges Funktional und der nicht-lokale Austauschoperator $V_{\text{Fock}}[\{\psi_i\}]$ mit

$$\left(V_{\text{Fock}}[\{\psi_i\}] \psi_k \right)(\mathbf{r}) := - \int \sum_{i=1}^N \delta_{\sigma\sigma_i} \frac{\psi_{k_i}(\mathbf{r}) \psi_{k_i}^*(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \psi_k(\mathbf{r}') d^3 r' \quad (1.27)$$

eingeführt, wobei σ wieder die Spinrichtung des Zustandes ψ_k kennzeichnen soll. Um hervorzuheben, dass es sich bei $V_{\text{Fock}}[\{\psi_i\}]$ um einen Integral-Operator handelt, wurde es hier im Gegensatz zu den multiplikativ wirkenden Potentialen v_{ext} und v_H mit einem Großbuchstaben bezeichnet.

Gemäß des Variationsprinzips (1.12) findet man mit dem Hartree-Fock-Verfahren eine obere Schranke für die tatsächliche Grundzustandsenergie des Systems, die sich aus (1.5) ergeben würde, wenn diese Gleichung lösbar wäre. Die Abweichung der Hartree-Fock-Energie zu der tatsächlichen Grundzustandsenergie wird als *Korrelationsenergie* bezeichnet. Weiterführende auf der Berechnung der Wellenfunktionen basierende Ansätze, die auch die Korrelation näherungsweise miteinbeziehen, sind etwa die Møller-Plesset-Störungstheorien 2. bis 4. Ordnung (MP2, MP4) [14] sowie *Configuration-Interaction* (CI) oder *Coupled-Cluster-Methoden* (CC) [15].

⁴Die Slater-Determinante (1.13) ändert sich bis auf einen Phasenfaktor nicht bei einer unitären Transformation der Ein-Teilchen-Funktionen, und auch das nachfolgend definierte Hartree-Potential v_H sowie der nicht-lokale Austauschoperator $V_{\text{Fock}}[\{\psi_i\}]$ sind invariant gegenüber einer solchen Transformation [21].

Der nicht-lokale Austauschoperator als Funktional der Dichtematrix

Im Vorgriff auf die im folgenden Abschnitt beschriebene Dichtefunktional-Theorie wurde der Zähler im Integranden von (1.25) hier bereits zur *Ladungsdichte* ρ zusammengefasst. Auch im Hinblick auf die Untersuchungen von Kapitel 7 sei erwähnt, dass Ähnliches für den nicht-lokalen Austauschoperator ebenfalls gelingt, wenn man auf die ursprünglichen Wellenfunktionen $\phi_k(\mathbf{x}_i) := \psi_k(\mathbf{r}_i)\chi_{\sigma_k}(s_i)$ zurückgreift. Die *Dichtematrix*⁵ der Slater-Determinante (1.13) ist dann gegeben durch

$$\gamma(\mathbf{x}, \mathbf{x}') = \sum_{i=1}^N \phi_{k_i}(\mathbf{x})\phi_{k_i}^*(\mathbf{x}') . \quad (1.28)$$

Der Kern des Austauschoperators kann damit aber geschrieben werden als

$$v_{\text{Fock}}[\gamma](\mathbf{x}, \mathbf{x}') := -\frac{\gamma(\mathbf{x}, \mathbf{x}')}{|\mathbf{r} - \mathbf{r}'|} . \quad (1.29)$$

Anwendung von $v_{\text{Fock}}[\gamma]$ auf ϕ_k liefert für den Ortsanteil dann sofort (1.27) wenn man beachtet, dass die Integration über \mathbf{x}' die Summation über die Spinkomponenten miteinschließt.

Sehr ähnlich, wie im nächsten Abschnitt die Kohn-Sham-Gleichungen aus einem Variationsprinzip für die Dichte abgeleitet werden (siehe Gleichungen (1.39) und (1.48)), lassen sich interessanterweise die Hartree-Fock-Gleichungen (1.24) aus einer Variation für die *Dichtematrix* herleiten. Wie man zeigt, lässt sich nicht nur der Austauschoperator und damit auch die Austauschenergie (1.18), sondern die gesamte Hartree-Fock-Energie E_{HF} von (1.14) als Funktional von γ ausdrücken [21]. Die Variationsableitung von E_{HF} nach γ erfolgt dann unter der Nebenbedingung, dass γ die Dichtematrix einer einfachen N -Teilchen-Determinante ist. Daraus ergibt sich das Variationsprinzip

$$E_{\text{HF}}[\gamma_{\text{HF}}] \leq E_{\text{HF}}[\gamma] , \quad (1.30)$$

wobei γ_{HF} die Dichtematrix ist, die aus den Lösungen von (1.24) zusammengesetzt ist.

Nicht-lokaler Austauschoperator und elektronische Selbstwechselwirkung

Eine wesentliche Eigenschaft der Hartree-Fock-Methode ist die korrekte Handhabung der *elektronischen Selbstwechselwirkung*. Dabei handelt es sich um eine unphysikalische Größe, die durch die Beiträge der Summanden mit $i = j$ zur Hartree-Energie (1.17) und zur Austauschenergie (1.18) gegeben ist. Wegen des unterschiedlichen Vorzeichens in beiden Ausdrücken kürzt sich die Selbstwechselwirkung jedoch in Gleichung (1.24) exakt heraus und ist daher in diesem Formalismus unproblematisch. Betrachtet man aber etwa die Austauschwechselwirkung allein, so zeigt sich, dass die Selbstwechselwirkung den dominierenden Anteil im Energiefunktional (1.18) darstellt [5, 16]. So ist beispielsweise beim freien Wasserstoffatom die Austauschenergie (und — mit entgegengesetztem Vorzeichen — ebenso die Hartree-Energie) zu 100% durch die Selbstwechselwirkung gegeben. Aus diesem Grund kann eine unvollständige Eliminierung der Selbstwechselwirkung infolge einer nur näherungsweisen Berechnung des Austausch-Terms erhebliche Probleme verursachen. Dies betrifft auch die nachfolgend vorgestellte Dichtefunktional-Theorie bei Verwendung von LDA- oder GGA-Austausch-Korrelationsfunktionalen (vgl. Abschnitt 1.5.2); auf eine Reihe damit verbundener Schwierigkeiten wurde bereits in der Einleitung verwiesen.

⁵genauer: die reduzierte Dichtematrix 1. Ordnung

Im Gegensatz zu den in der Dichtefunktional-Theorie verwendeten lokalen Austauschpotentialen (auch dem exakten „EXX-Potential“, vgl. Abschnitt 1.6), die ein auf alle Elektronen identisch wirkendes Kraftfeld repräsentieren, stellt das nicht-lokale Austauschpotential (1.27) der Hartree-Fock-Gleichungen eine auf jedes Elektron „individuell“ wirkende Kraft dar. Insbesondere werden im Hartree-Fock-Formalismus besetzte und unbesetzte Zustände verschieden gehandhabt, und zwar so, dass formal auf unbesetzte Zustände ein von N Elektronen verursachtes elektrostatisches Potential wirkt, besetzte Zustände hingegen ein nur von den $N - 1$ „anderen“ Elektronen bewirktes Potential „sehen“. Dies erkennt man wie folgt: Multipliziert man die rechte Seite von (1.27) mit der „geschickten Eins“ $\psi_k(\mathbf{r})/\psi_k(\mathbf{r}) = 1$, lässt sich der Austauschoperator mit

$$\rho_k(\mathbf{r}, \mathbf{r}') := \sum_{i=1}^N \delta_{\sigma\sigma_i} \frac{\psi_{k_i}^*(\mathbf{r}')\psi_k(\mathbf{r}')\psi_{k_i}(\mathbf{r})}{\psi_k(\mathbf{r})}, \quad (1.31)$$

zusammenfassen zu

$$V_k^{\text{Fock}}[\{\psi_i\}](\mathbf{r}) = - \int \frac{\rho_k(\mathbf{r}, \mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3r'. \quad (1.32)$$

Dies ist nun eine lokale Größe, allerdings ist sie für jeden Zustand k verschieden definiert. Fasst man das Austauschpotential in dieser Form mit dem Hartree-Potential (1.25) zusammen, ergibt sich ein *formales* elektrostatisches Potential

$$\int \frac{\rho(\mathbf{r}') - \rho_k(\mathbf{r}, \mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3r', \quad (1.33)$$

wobei streng genommen auch dieser Ausdruck nur sinnvoll ist, wenn er vor die Wellenfunktion ψ_k geschrieben wird. Mit

$$\int \rho_k(\mathbf{r}, \mathbf{r}') d^3r' = \begin{cases} 1 & \text{mit } \psi_k \text{ besetzter Zustand} \\ 0 & \text{mit } \psi_k \text{ unbesetzter Zustand} \end{cases} \quad (1.34)$$

lässt sich (1.33) so deuten, dass durch das Austauschpotential die in v_{H} berücksichtigte Ladungsverteilung „um ein Elektron“ reduziert wird, falls ψ_k einen besetzten Zustand bezeichnet. Wie man zeigt (vgl. etwa Ref. 89, 92), findet diese Reduktion von Ladungsdichte gerade in der Umgebung des k -ten Elektrons statt und betrifft nur Teilchen mit der selben Spinrichtung σ . Man spricht von einem so genannten „Austauschloch“. Die Verdrängung von Teilchen gleicher Spinrichtung vom Ort des k -ten Teilchens ist letztlich eine Folge des Pauli-Prinzips. Ein nicht besetzter Zustand ψ_k hingegen „sieht“ das gesamte, von allen N Valenzelektronen verursachte Potential; in diesem Falle beschreibt ρ_k zwar eventuell eine Umverteilung der Ladung, die Gesamtladung jedoch wird davon nicht beeinflusst.

1.4 Dichtefunktional-Theorie

Die Grundidee der Dichtefunktional-Theorie (DFT) ist, das Variationsprinzip (1.12) so umzuformulieren, dass nicht mehr der Vielteilchen-Wellenfunktion Ψ , sondern der viel „handlicheren“ Ladungsdichte ρ die Schlüsselrolle bei der Berechnung der elektronischen Struktur zufällt. Statt der $3N$ Koordinaten, die für ein N -Elektronengas bei Betrachtung der Wellenfunktionen zu berücksichtigen wären, hat man mit der Ladungsdichte eine Größe, die sich unabhängig von der Anzahl der betrachteten Teilchen durch 3 Koordinaten beschreiben lässt. Insbesondere bei Verwendung von Minimierungstechniken (vgl. Abschnitt 4.1.2) stellt dies numerisch eine enorme Vereinfachung des Problems dar.

Einen ersten Ansatz für ein allein von der Ladungsdichte abhängendes Energiefunktional bieten die Arbeiten von Thomas und Fermi aus dem Jahre 1927 [17, 18, 19, 20]. Diese von vorneherein als Näherung konzipierte Methode liefert jedoch vielfach keine befriedigenden Ergebnisse⁶ und ist heute eher von historischer Bedeutung. Eine Darstellung des Thomas-Fermi-Modells findet sich z.B. in Ref. 21.

Die Grundlage für eine — im Prinzip — exakte Theorie, die ein gegebenes elektronisches System im Grundzustand mit Hilfe eines Funktionals der Ladungsdichte beschreibt, wurde 1964 von Hohenberg und Kohn durch den Beweis zweier Sätze, den berühmten Hohenberg-Kohn-Theoremen [1], geschaffen. Sowohl die traditionelle Herleitung dieser Dichtefunktional-Theorie als auch — im Hinblick auf Abschnitt 1.5.4 — ein mathematisch etwas strengerer Zugang sollen nun kurz umrissen werden, wobei die folgende Darstellung an der von Parr und Yang [21] orientiert ist.

1.4.1 Die Hohenberg-Kohn-Theoreme

Der Hamilton-Operator des elektronischen Systems (1.5) ist mit der Angabe der Elektronenzahl N und des äußeren Potentials v_{ext} eindeutig definiert.⁷ Unter der Annahme, dass der Grundzustand des Systems nicht entartet ist, existiert daher zu jedem äußeren Potential v_{ext} genau eine N -Teilchen-Wellenfunktion Ψ_0 , die diesen Grundzustand beschreibt. Mit dem Operator der Teilchendichte $\hat{\rho}$ lässt sich Ψ_0 über die Beziehung

$$\rho_0(\mathbf{r}) = \langle \Psi_0 | \hat{\rho}(\mathbf{r}) | \Psi_0 \rangle = \left\langle \Psi_0 \left| \sum_{i=1}^N \delta(\mathbf{r} - \mathbf{r}_i) \right| \Psi_0 \right\rangle \quad (1.35)$$

wiederum eindeutig die Grundzustandsladungsdichte zuordnen. Insgesamt ergibt sich also eine eindeutige Zuordnung $v_{\text{ext}} \mapsto \rho_0$. Der Wertebereich dieser Zuordnung ist die Menge der *v-darstellbaren* Ladungsdichten, also solcher Dichten, die sich mit (1.35) überhaupt einem Potential zuordnen lassen, vgl. Abschnitt 1.4.3. Für den weiteren Verlauf dieses Abschnitts sei die *v-Darstellbarkeit* von ρ_0 stets vorausgesetzt. Der Beweis des *Hohenberg-Kohn-Theorems I* [1] liefert nun die Eindeutigkeit der umgekehrten Zuordnung $\rho_0 \mapsto v_{\text{ext}}$: Das äußere Potential v_{ext} ist (bis auf eine additive Konstante) ein Funktional der Grundzustandsladungsdichte. Da die Grundzustandsenergie E_0 über die Wellenfunktion Ψ_0 als Funktional von v_{ext} aufgefasst werden kann, folgt die eigentliche Aussage des Theorems: Die Grundzustandsenergie E_0 des N -Teilchen-Systems ist ein Funktional der Grundzustandsladungsdichte ρ_0 .

Es ergeben sich also sowohl v_{ext} als auch — durch Integration — die Teilchenzahl N aus ρ_0 ; daher bestimmt ρ_0 den gesamten Hamilton-Operator und alle Eigenschaften des Systems, wie z.B. die kinetische oder die potentielle Energie. Für die Gesamtenergie des Grundzustandes ergibt sich

$$\begin{aligned} E_0[\rho_0] &= T[\rho_0] + E_{\text{ext}}[\rho_0] + E_{\text{ee}}[\rho_0] \\ &= \int v_{\text{ext}}[\rho_0](\mathbf{r})\rho_0(\mathbf{r}) \, d^3r + F[\rho_0], \end{aligned} \quad (1.36)$$

wobei im zweiten Schritt das Funktional F mit

$$F[\rho] := T[\rho] + E_{\text{ee}}[\rho] \quad (1.37)$$

⁶Das Thomas-Fermi-Modell erweist sich beispielsweise als ungeeignet für die Vorhersage molekularer Bindungen [22].

⁷Die Kernkoordinaten seien ab jetzt als fest betrachtet.

benutzt wurde, das nicht explizit vom äußeren Potential abhängt und deshalb auch als „universelles Potential“ bezeichnet wird⁸.

Wie lässt sich nun aber zu einem *gegebenen* äußeren Potential v_{ext} die entsprechende Grundzustandsladungsdichte bestimmen, ohne die N -Teilchen-Wellenfunktion Ψ_0 bestimmen zu müssen? Die Antwort hierauf gibt das *Hohenberg-Kohn-Theorem II*, das ein Variationsprinzip für die Elektronendichte konstituiert: Bei Variation von ρ nimmt das Energiefunktional

$$E_{v_{\text{ext}}}[\rho] := \int v_{\text{ext}}(\mathbf{r})\rho(\mathbf{r}) d^3r + F[\rho] \quad (1.38)$$

sein Minimum gerade für die Grundzustandselektronendichte ρ_0 an; für geeignete Ladungsdichten ρ gilt also

$$E_0 = E_{v_{\text{ext}}}[\rho_0] \leq E_{v_{\text{ext}}}[\rho]. \quad (1.39)$$

Dies ist das Analogon zum Variationsprinzip (1.12) für die Vielteilchen-Wellenfunktion Ψ_0 . Berücksichtigt man die Erhaltung der Teilchenzahl N über einen Lagrange-Parameter μ , folgt die Bedingung

$$\frac{\delta}{\delta\rho} \left\{ E_{v_{\text{ext}}}[\rho] - \mu \cdot \left[\int \rho(\mathbf{r}) d^3r - N \right] \right\} \stackrel{!}{=} 0 \quad (1.40)$$

und damit die Euler-Lagrange-Gleichung

$$\mu = \frac{\delta E_{v_{\text{ext}}}[\rho]}{\delta\rho(\mathbf{r})} = v_{\text{ext}}(\mathbf{r}) + \frac{\delta F[\rho]}{\delta\rho(\mathbf{r})}. \quad (1.41)$$

Beide Theoreme geben allerdings keinen Hinweis auf eine explizite Gestalt des Funktionals F . Daher ist die DFT bislang nur eine *im Prinzip* exakte Theorie, die in der Praxis jedoch nicht ohne Näherungen auskommt. Ihre Formulierung durch Kohn und Sham [2], die Gegenstand des folgenden Abschnitts sein soll, gestattet allerdings eine Reihe von Näherungen, die für viele Anwendungen hinreichende Genauigkeit aufweisen und letztlich den großen Erfolg der DFT begründen.

1.4.2 Die Kohn-Sham-Gleichungen

In der Kohn-Sham-Formulierung der Dichtefunktional-Theorie [2] wird dem realen elektronischen System ein System fiktiver „Elektronen“ gegenübergestellt, für die die Coulomb-Abstoßung ausgeschaltet ist, die aber die gleiche Ladungsdichte besitzen wie das reale System. Für das „universelle Funktional“ F aus (1.37) lassen sich im Falle dieses *nicht-wechselwirkenden Referenzsystems* jedoch viel bessere Näherungen finden als beim tatsächlichen System (vgl. dazu auch Kap. 1.5).

Das fiktive System kann durch folgende Eigenwertgleichung beschrieben werden:

$$H_s\psi_i = \left[-\frac{1}{2}\nabla^2 + v_s(\mathbf{r})\right]\psi_i = \varepsilon_i\psi_i, \quad (1.42)$$

also durch eine formale Einteilchen-Schrödinger-Gleichung. Die Kohn-Sham-Teilchen „sehen“ sich zwar nicht durch die Coulomb-Wechselwirkung, dem Pauli-Prinzip müssen sie aber gehorchen. Täten sie es nämlich nicht, würden ja alle in denselben Zustand, den

⁸ F ist — formal gesehen — genauso auf das Wasserstoffatom wie auch auf ein kompliziertes Molekül oder einen Festkörper anwendbar.

Grundzustand des Einteilchensystems, rutschen. Also können wir die Vielteilchen-Wellenfunktion des Grundzustandes als Slater-Determinante ansetzen,

$$\Psi_s = \frac{1}{\sqrt{N!}} \det[\psi_1, \dots, \psi_N], \quad (1.43)$$

wobei die Determinante aus den N niedrigsten Einteilchenzuständen gebildet werde. Im Gegensatz zur kinetischen Energie T des tatsächlichen Systems ergibt sich die kinetische Energie T_s des Kohn-Sham-Systems damit einfach zu

$$T_s[\rho] = \sum_{i=1}^N \langle \psi_i | -\frac{1}{2} \nabla^2 | \psi_i \rangle. \quad (1.44)$$

Auch auf das Referenzsystem lassen sich die beiden Theoreme von Hohenberg und Kohn [1] anwenden; deshalb wurde T_s hier bereits gleich als Funktional von ρ aufgefasst. Entsprechend zu (1.41) lautet die Euler-Langrange-Gleichung des Kohn-Sham-Systems⁹

$$\mu = v_s(\mathbf{r}) + \frac{\delta T_s[\rho]}{\delta \rho(\mathbf{r})}. \quad (1.45)$$

Auf eine formal identische Gleichung gelangt man jedoch auch für das reale System, wenn man für F folgenden Ansatz macht:

$$F[\rho] = T_s[\rho] + E_H[\rho] + E_{xc}[\rho]. \quad (1.46)$$

Hier ist E_H die Hartree-Energie, die bereits in (1.17) als Funktional von ρ definiert wurde, und E_{xc} bezeichnet die so genannte *Austausch-Korrelationsenergie*, definiert durch

$$E_{xc}[\rho] := T[\rho] - T_s[\rho] + E_{ee}[\rho] - E_H[\rho]. \quad (1.47)$$

Die Austausch-Korrelationsenergie zerfällt also in einen kinetischen Anteil $T - T_s$, der, wie zumindest für Atome gezeigt wurde [21, 23], sehr klein ist, und den nicht-klassischen Anteil der Elektron-Elektron-Wechselwirkung. Die Euler-Langrange-Gleichung (1.41) wird damit im Falle des realen Systems zu

$$\begin{aligned} \mu &= v_{\text{ext}}(\mathbf{r}) + \frac{\delta F[\rho]}{\delta \rho(\mathbf{r})} \\ &= v_{\text{ext}}(\mathbf{r}) + \frac{\delta E_H[\rho]}{\delta \rho(\mathbf{r})} + \frac{\delta E_{xc}[\rho]}{\delta \rho(\mathbf{r})} + \frac{\delta T_s[\rho]}{\delta \rho(\mathbf{r})} \\ &= v_{\text{eff}}[\rho](\mathbf{r}) + \frac{\delta T_s[\rho]}{\delta \rho(\mathbf{r})}, \end{aligned} \quad (1.48)$$

wobei das effektive Potential durch

$$v_{\text{eff}}[\rho](\mathbf{r}) := v_{\text{ext}}(\mathbf{r}) + v_H[\rho](\mathbf{r}) + v_{xc}[\rho](\mathbf{r}) \quad (1.49)$$

mit dem *Hartree-Potential*

$$v_H[\rho](\mathbf{r}) := \frac{\delta E_H[\rho]}{\delta \rho(\mathbf{r})} = \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3r' \quad (1.50)$$

⁹Bereits hier sei darauf hingewiesen, dass das Funktional F — trotz der Bezeichnung „universell“ — für das Kohn-Sham-System anders definiert ist, als für das reale System! Da es im Kohn-Sham-System keine Coulomb-Abstoßung gibt, gilt in (1.37) nämlich $E_{ee}[\rho] = 0$ und folglich $F[\rho] = T_s[\rho]$.

und dem (lokalen!) *Austausch-Korrelations-Potential*

$$v_{\text{xc}}[\rho](\mathbf{r}) := \frac{\delta E_{\text{xc}}[\rho]}{\delta \rho(\mathbf{r})} \quad (1.51)$$

gegeben ist. Der „Trick“ des Ansatzes von Kohn und Sham ist also, bis auf T_s alle Anteile des Funktionals F von (1.46) in das effektive Potential $v_{\text{eff}}[\rho]$ „herüberzuschieben“. Damit nehmen die Euler-Lagrange-Gleichungen des realen und des nicht-wechselwirkenden Referenzsystems (1.48) genau die gleiche Gestalt an, wenn man $v_s = v_{\text{eff}}[\rho]$ setzt. Da v_{eff} allerdings ein von der Dichte ρ abhängiges Funktional ist, muss die entsprechende Eigenwert-Gleichung

$$\left[-\frac{1}{2}\nabla^2 + v_{\text{ext}}(\mathbf{r}) + v_{\text{H}}[\rho](\mathbf{r}) + v_{\text{xc}}[\rho](\mathbf{r}) \right] \psi_i(\mathbf{r}) = \varepsilon_i \psi_i(\mathbf{r}) \quad (1.52)$$

selbstkonsistent zusammen mit der Gleichung für die Ladungsdichte gelöst werden, die für eine einfache Slater-Determinante nach (1.35) durch

$$\rho(\mathbf{r}) = \sum_{i=1}^N |\psi_i(\mathbf{r})|^2 \quad (1.53)$$

gegeben ist. Die beiden vorangehenden Gleichungen (1.52) und (1.53) stellen zusammen die *Kohn-Sham-Gleichungen* dar. Benutzt man (1.46) in (1.38), erhält man die Gesamtenergie des realen Systems:

$$E_{\text{real}}[\rho] = \int v_{\text{ext}}(\mathbf{r})\rho(\mathbf{r}) \, d^3r + T_s[\rho] + E_{\text{H}}[\rho] + E_{\text{xc}}[\rho]. \quad (1.54)$$

Eine alternative Form findet man, indem $T_s[\rho]$ mit Hilfe von (1.52) durch die Summe der Kohn-Sham-Eigenwerte ε_i substituiert wird¹⁰:

$$E_{\text{real}}[\rho] = \sum_i^N \varepsilon_i - E_{\text{H}}[\rho] + E_{\text{xc}}[\rho] - \int v_{\text{xc}}[\rho](\mathbf{r})\rho(\mathbf{r}) \, d^3r. \quad (1.55)$$

Die Gesamtenergie des Kohn-Sham-Systems unterscheidet sich von der des realen Systems, was später in Abschnitt 1.5.4 benutzt wird. Da das Kohn-Sham-System wechselwirkungsfrei ist, gilt $F[\rho] = T_s[\rho]$ (vgl. Fußnote auf Seite 19), und das äußere Potential ist durch das effektive Potential von (1.49) gegeben. Aus (1.38) folgt damit

$$\begin{aligned} E_{\text{KS}}[\rho] &= \int v_{\text{eff}}[\rho](\mathbf{r})\rho(\mathbf{r}) \, d^3r + T_s[\rho] \\ &= E_{\text{real}}[\rho] + E_{\text{H}}[\rho] - E_{\text{xc}}[\rho] + \int v_{\text{xc}}[\rho](\mathbf{r})\rho(\mathbf{r}) \, d^3r. \end{aligned} \quad (1.56)$$

Die Euler-Lagrange-Gleichungen (1.41) und (1.48) folgen direkt aus dem Variationsprinzip für die Ladungsdichte (1.38). Deshalb liefert die Lösung der Kohn-Sham-Gleichungen zwar die korrekte Ladungsdichte des realen Systems — die Wellenfunktionen ψ_i und die Energieniveaus ε_i des Kohn-Sham-Systems erlauben jedoch nicht mehr ohne Weiteres eine direkte physikalische Interpretation. Für finite Systeme und Oberflächen lässt sich

¹⁰Die Substitution gelingt mit $\sum_i^N \varepsilon_i = \sum_i^N \langle \psi_i | -\frac{1}{2}\nabla^2 + v_{\text{eff}}[\rho] | \psi_i \rangle = T_s[\rho] + \int v_{\text{eff}}[\rho](\mathbf{r})\rho(\mathbf{r}) \, d^3r$. Dies ist mit (1.56) auch gerade die Gesamtenergie des Kohn-Sham-Systems.

allerdings zeigen, dass der höchste besetzte Eigenwert ε_{\max} des Kohn-Sham-Systems „exakt“ ist in dem Sinne, dass er der (negativen) Ionisierungsenergie I und dem chemischen Potential μ entspricht [2, 25], das als Lagrange-Parameter bereits in (1.41) aufgetaucht ist:

$$\varepsilon_{\max} = -I = \mu . \quad (1.57)$$

Lässt man nicht-ganzzahlige Besetzungszahlen n_i der Kohn-Sham-Zustände zu, ergeben sich die Einteilchen-Energien ferner auch über das *Janak-Theorem* [26] aus der Gesamtenergie durch Ableitung nach der Besetzungszahl:

$$\frac{\partial E[\rho]}{\partial n_i} = \varepsilon_i . \quad (1.58)$$

1.4.3 Levy's *Constraint-Search*-Formulierung der DFT

Zum besseren Verständnis der Göring-Levy-Störungstheorie [40, 41], auf die in Abschnitt 1.5.4 eingegangen wird, soll an dieser Stelle kurz eine zu der Hohenberg-Kohn-Formulierung [1] alternative Darstellung der Dichtefunktional-Theorie angesprochen werden, die als *Constraint-Search*-Formulierung bezeichnet wird. Diese 1979 von Levy [27] vorgeschlagene Herangehensweise hat zudem den Vorteil, dass die von Hohenberg und Kohn gemachten Annahmen (Nicht-Entartung des Grundzustandes, v -Darstellbarkeit der Ladungsdichte, vgl. Abschnitt 1.4.1) abgeschwächt werden können. Sind die Annahmen jedoch erfüllt, stimmen beide Formulierungen überein; es handelt sich also um eine echte Verallgemeinerung der ursprünglichen Hohenberg-Kohn-Formulierung.

Bei der *Constraint-Search*-Formulierung startet man zunächst mit dem „normalen“ Variationsprinzip (1.12) für die Vielteilchen-Wellenfunktion Ψ , führt dann aber die Minimierung in zwei Stufen aus: Einmal wird über alle (antisymmetrischen) Ψ minimiert, die auf eine bestimmte Ladungsdichte ρ führen, das ist die *Constraint Search*. Da nach dem Variationsprinzip die Suche aber über alle (normierten) N -Teilchen-Zustände durchgeführt werden muss, wird in einem zweiten Variationsschritt über alle geeigneten *Ladungsdichten* minimiert:

$$\begin{aligned} E_0 &= \min_{\Psi} \langle \Psi | T + V_{\text{ee}} + \sum_i^N v_{\text{ext}}(\mathbf{r}_i) | \Psi \rangle \\ &= \min_{\rho} \left\{ \min_{\Psi \rightarrow \rho} \langle \Psi | T + V_{\text{ee}} + \sum_i^N v_{\text{ext}}(\mathbf{r}_i) | \Psi \rangle \right\} \\ &= \min_{\rho} \left\{ \min_{\Psi \rightarrow \rho} \left[\langle \Psi | T + V_{\text{ee}} | \Psi \rangle + \int v_{\text{ext}}(\mathbf{r}) \rho(\mathbf{r}) d^3r \right] \right\} . \end{aligned} \quad (1.59)$$

Für die Minimierung wird hier nicht vorausgesetzt, dass die Ladungsdichte v -darstellbar ist (vgl. Abschnitt 1.4.1), allerdings muss sie „ N -darstellbar“ sein. Alle physikalisch sinnvollen Ladungsdichten sind N -darstellbar — die Bedingungen dafür lauten [28]

$$\rho(\mathbf{r}) \geq 0 , \quad \int \rho(\mathbf{r}) d^3r = N , \quad \int |\sqrt{\rho(\mathbf{r})} \nabla^2 \sqrt{\rho(\mathbf{r})}| d^3r < \infty . \quad (1.60)$$

Durch die letzte Bedingung wird sichergestellt, dass die kinetische Energie endlich bleibt (Abschnitt IV in Ref. 28). Vergleicht man die letzte Zeile von (1.59) mit dem aus den Hohenberg-Kohn-Theoremen folgenden Variationsprinzip (1.39) für die Ladungsdichte,

$$E_0 = \min_{\rho} \left\{ F[\rho] + \int v_{\text{ext}}(\mathbf{r}) \rho(\mathbf{r}) d^3r \right\} \quad (\text{mit } \rho \text{ } v\text{-darstellbar}), \quad (1.61)$$

so findet man für das „universelle Funktional“ F einen Ausdruck, der im Falle v -darstellbarer Ladungsdichten genau dasselbe Ergebnis liefern muss wie Ausdruck (1.37) der Hohenberg-Kohn-Theorie, aber allgemeiner für N -darstellbare Ladungsdichten definiert ist:

$$F[\rho] = \min_{\Psi \rightarrow \rho} \langle \Psi | T + V_{ee} | \Psi \rangle \quad (\text{mit } \rho \text{ } N\text{-darstellbar}). \quad (1.62)$$

Auch die Beschränkung auf nicht-entartete Grundzustände entfällt hier; es wird dasjenige Ψ gewählt, das die Klammer minimiert. Diese Beziehung stellt also „eine Art“ funktionalen Zusammenhang zwischen der Ladungsdichte und den Zuständen dar, der aber — anders als bei den Hohenberg-Kohn-Theoremen — wegen der möglichen Entartung nicht mehr ein-eindeutig ist.

Für die Gesamtenergien (1.54) und (1.56) schreibt man daher in der *Constraint-Search*-Formulierung auch

$$E_{\text{nww}}[\rho] = \int v_{\text{eff}}[\rho](\mathbf{r}) \rho(\mathbf{r}) d^3r + F_{\text{nww}}[\rho], \quad (1.63)$$

$$E_{\text{real}}[\rho] = \int v_{\text{ext}}(\mathbf{r}) \rho(\mathbf{r}) d^3r + F_{\text{real}}[\rho] \quad (1.64)$$

mit den universellen Potentialen

$$F_{\text{nww}}[\rho] := \min_{\Psi \rightarrow \rho} \langle \Psi | T | \Psi \rangle = T_{\text{s}}[\rho] \stackrel{!}{=} \langle \Psi^0 | T | \Psi^0 \rangle, \quad (1.65)$$

$$F_{\text{real}}[\rho] := \min_{\Psi \rightarrow \rho} \langle \Psi | T + V_{ee} | \Psi \rangle \stackrel{!}{=} \langle \Psi^1 | T + V_{ee} | \Psi^1 \rangle. \quad (1.66)$$

Die Abkürzung „nww“ steht hier für „nicht-wechselwirkendes System“. Dessen Viel-Teilchen-Zustände Ψ^0 und die des realen Systems Ψ^1 ergeben sich für eine *gegebene* Ladungsdichte ρ damit (formal) nur aus der Minimierung der spitzen Klammern — völlig unabhängig vom äußeren bzw. vom effektiven Potential. Es sei jedoch nochmals darauf hingewiesen, dass Ψ^0 hier im Allgemeinen *nicht* mehr, wie noch im Kohn-Sham-Formalismus, durch eine einfache Slater-Determinante gegeben ist.

1.5 Das Austausch-Korrelations-Funktional

1.5.1 Definition von Austausch- und Korrelationsenergie in der DFT

Üblicherweise wird die Austausch-Korrelationsenergie aus (1.47) in einen Austauschanteil E_{x} und einen Korrelationsanteil E_{c} untergliedert:

$$E_{\text{xc}}[\rho] = E_{\text{x}}[\rho] + E_{\text{c}}[\rho]. \quad (1.67)$$

Eine analoge Aufspaltung folgt dann auch für das Austausch-Korrelationspotential:

$$v_{\text{xc}}[\rho] = v_{\text{x}}[\rho] + v_{\text{c}}[\rho] \quad \text{mit} \quad v_{\text{x}}[\rho] := \frac{\delta E_{\text{x}}[\rho]}{\delta \rho}, \quad v_{\text{c}}[\rho] := \frac{\delta E_{\text{c}}[\rho]}{\delta \rho}. \quad (1.68)$$

Zur Definition der beiden Anteile sei wieder die Notation von Abschnitt 1.4.3 benutzt: Für eine Ladungsdichte ρ bezeichne also Ψ^0 den N -Teilchen-Zustand des Kohn-Sham-Systems und Ψ^1 den des realen Systems. Der Zusammenhang zwischen den Zuständen und der Dichte ist durch die Gleichungen (1.65) und (1.66) gegeben. Ist ρ eine v -darstellbare Dichte,

ist der Kohn-Sham-Zustand nach Konstruktion des nicht-wechselwirkenden Systems durch die Slater-Determinante $\Psi^0 = \det[\psi_1, \dots, \psi_N] / \sqrt{N!}$ gegeben (vgl. Abschnitt 1.4.2), wobei die Einteilchen-Zustände ψ_i die N niedrigsten Eigenfunktionen in (1.52) sind. Damit lässt sich die Austauschenergie in der DFT analog zu der im Hartree-Fock-Formalismus (1.18) definieren:¹¹

$$E_x[\rho] := -\frac{1}{2} \sum_{i,j}^{1\dots N} \delta_{\sigma_i \sigma_j} \int \frac{\psi_i^*(\mathbf{r}) \psi_j^*(\mathbf{r}') \psi_i(\mathbf{r}') \psi_j(\mathbf{r})}{|\mathbf{r} - \mathbf{r}'|} d^3r d^3r' \quad (1.69)$$

$$= \langle \Psi^0 | V_{ee} | \Psi^0 \rangle - \langle \Psi^0 | v_H[\rho] | \Psi^0 \rangle. \quad (1.70)$$

In der *Constraint-Search*-Formulierung wird gleich die zweite Zeile als Definition verwendet (vgl. z.B. [29, 30, 31, 40]) mit dem Hartree-Potential v_H von (1.50); sie ist allgemeiner, denn ihr Definitionsbereich lässt sich auf *alle* antisymmetrischen Wellenfunktionen Ψ^0 ausweiten, wie in den Variationsprinzipien (1.12) und (1.59) auch gefordert.

Die Korrelationsenergie ergibt sich einfach als „Rest“, der übrig bleibt, wenn man die Definition von E_x zusammen mit (1.67) und (1.65), (1.66) in Gleichung (1.56) benutzt:

$$E_c[\rho] = \langle \Psi^1 | T + V_{ee} | \Psi^1 \rangle - \langle \Psi^0 | T + V_{ee} | \Psi^0 \rangle. \quad (1.71)$$

Geschlossene Ausdrücke für $E_{xc}[\rho]$ und $v_{xc}[\rho]$ lassen sich, bis auf wenige Ausnahmen wie dem dissoziierten H_2 -Molekül [32] oder He und He-artigen Ionen [44], bislang nicht angeben, jedoch ist es formal möglich, sie mit Hilfe einer Reihenentwicklung darzustellen. Dies soll im übernächsten Abschnitt 1.5.4 skizziert werden. Formalismen zur exakten Ableitung des lokalen Austauschpotentials v_x (bei Annahme der v -Darstellbarkeit der Dichte) werden in Abschnitt 1.6 vorgestellt. Zuvor soll nun aber noch auf die gängigsten Näherungen eingegangen werden.

1.5.2 Lokale Dichtenäherung und verwandte Funktionale

Als effizientes Verfahren zur Untersuchung praktischer Fragestellungen wurde die Dichtefunktional-Theorie zunächst in der Physik durch die Einführung der *Lokalen Dichtenäherung* (LDA) [2] und ihrer Erweiterung für spin-polarisierte Systeme, der *Lokalen Spin-Dichte-Näherung* (LSD) [33], populär. Ihre weite Verbreitung auch in der Quantenchemie erfolgte dann etwas später und geht mit der Entwicklung einer Reihe so genannter *Verallgemeinerter Gradientennäherungen* (GGA) [35, 36, 37, 38, 39] einher, die sich als Weiterentwicklungen der LDA auffassen lassen.

Ausgangspunkt der LDA ist das homogene Elektronen-Gas, also ein System von N Elektronen in einem konstanten äußeren Potential. In diesem Fall lassen sich die Hartree-Fock-Gleichungen exakt lösen, vgl. z.B. Ref. 92, und man erhält den folgenden Zusammenhang zwischen der Austauschenergie pro Elektron $\varepsilon_x^{\text{hom}}$ und der konstanten Ladungsdichte ρ^{hom} :

$$\varepsilon_x^{\text{hom}}[\rho^{\text{hom}}] = -\frac{3}{4} \left(\frac{3}{\pi} \rho^{\text{hom}} \right)^{\frac{1}{3}}. \quad (1.72)$$

Die exakte Gestalt des entsprechenden Korrelationsanteils $\varepsilon_c^{\text{hom}}$ ist dagegen unbekannt. Deshalb wird hier in der Regel auf die Ergebnisse von Quanten-Monte-Carlo-Rechnungen für das homogene Elektronen-Gas nach Ceperley und Alder [34] zurückgegriffen. Für die in

¹¹Der einzige Unterschied zwischen der Kohn-Sham- und der Hartree-Fock-Austauschenergie besteht dann darin, dass im ersten Fall Kohn-Sham-, im zweiten aber Hartree-Fock-Orbitale verwendet werden.

dieser Arbeit durchgeführten Kohn-Sham-LDA-Rechnungen werden sie in der parametrisierten Form von Perdew und Zunger [5] benutzt. Ein näherungsweise Ausdruck für die Austausch-Korrelationsenergie (1.67) ist dann durch

$$E_{xc}^{\text{LDA}}[\rho] := \int \varepsilon_{xc}^{\text{hom}}[\rho(\mathbf{r})]\rho(\mathbf{r}) d^3r \quad (1.73)$$

gegeben, wobei Austausch- und Korrelationsenergie pro Elektron in $\varepsilon_{xc}^{\text{hom}} := \varepsilon_x^{\text{hom}} + \varepsilon_c^{\text{hom}}$ zusammengefasst worden sind. Damit ergibt sich unmittelbar das Austausch-Korrelationspotential in LDA:

$$v_{xc}^{\text{LDA}}[\rho](\mathbf{r}) := \frac{\delta E_{xc}^{\text{LDA}}[\rho]}{\delta \rho(\mathbf{r})} = \varepsilon_{xc}^{\text{hom}}[\rho(\mathbf{r})] + \frac{\delta \varepsilon_{xc}^{\text{hom}}[\rho]}{\delta \rho(\mathbf{r})}\rho(\mathbf{r}). \quad (1.74)$$

Trotz seiner Einfachheit liefert das LDA-Funktional für eine Vielzahl realistischer Fragestellungen quantitativ gute Resultate, vgl. z.B. Ref. 24. Per Konstruktion ist LDA exakt für den Grenzfall eines homogenen Elektronen-Gases. Folglich ist zu erwarten, dass sich die Näherung am besten für Systeme mit nur leicht variierender Dichte eignet. Weist die Dichte dagegen starke Oszillationen auf, scheint die Verwendung des LDA-Funktionals nicht mehr unmittelbar gerechtfertigt zu sein. Bei den verallgemeinerten Gradientennäherungen (GGA) wird daher die Dichte nicht mehr nur lokal wie in (1.73) berücksichtigt, sondern auch ihr Gradienten mit einbezogen. GGA-Energiefunktionale sind also von der Form

$$E_{xc}^{\text{GGA}}[\rho] := \int f_{xc}[\rho(\mathbf{r}), \nabla \rho(\mathbf{r})] d^3r. \quad (1.75)$$

Verschiedene Realisierungen des GGA-Funktionals sind das BP-Funktional von Becke [35] (Austauschanteil) und Perdew [36] (Korrelationsanteil), das BLYP-Funktional von Becke [35] (Austauschanteil), Lee, Yang und Parr [37] (Korrelationsanteil), sowie das PW91-Funktional von Perdew und Wang [38] und das PBE-Funktional von Perdew, Burke und Ernzerhof [39].

Interessanterweise führt die Verwendung von GGA-Funktionalen nicht unbedingt zu exakteren Beschreibungen gegenüber der LDA [24]; vielmehr muss häufig im Einzelfall überprüft werden, welches Funktional für eine gegebene Fragestellung am geeignetsten ist. Auf eine Reihe von Phänomenen, die weder durch LDA noch mit GGAs befriedigend beschrieben werden können, wurde bereits in der Einleitung eingegangen. Hier seien noch einmal das falsche asymptotische Abklingverhalten des Austauschpotentials endlicher Systeme (exponentiell statt $-1/r$ für $r \rightarrow \infty$) und die Unterschätzung der Bandlücke bei Halbleitern um bis zu über 100% erwähnt, beides Konsequenzen aus der nur näherungsweise Handhabung der elektronischen Selbstwechselwirkung (vgl. auch Abschnitt 1.3), sowie ferner das Versagen bei der Beschreibung der so genannten „Ganzzahl-Diskontinuität“.

1.5.3 Selbstwechselwirkungskorrektur

Ein Ansatz, die Handhabung der elektronischen Selbstwechselwirkung bei Verwendung genäherter Austausch-Korrelationspotentiale wie LDA oder GGA zu verbessern, ist die so genannte *Selbstwechselwirkungskorrektur* (SIC, von engl.: *self-interaction correction*). Im Schema von Perdew und Zunger [5] beispielsweise wird dabei ein gegebener Näherungsausdruck E_{xc}^{approx} für die Austausch-Korrelationsenergie (1.67) wie folgt modifiziert:

$$E_{xc}^{\text{SIC}}[\{\rho_k\}] := E_{xc}^{\text{approx}}[\rho] - \sum_k^{\text{occ.}} \left(E_H[\rho_k] + E_{xc}^{\text{approx}}[\rho_k] \right), \quad (1.76)$$

wobei $\rho_k := |\psi_k|^2$ die Dichte des k -ten Orbitals bezeichnet und E_H das Hartree-Energiefunktional (1.17) ist. Für die Kohn-Sham-Gleichung (1.52) erhält man damit ein *orbitalabhängiges* effektives Potential der Form

$$v_{\text{eff}}^k[\{\rho_i\}] := v_{\text{ext}} + v_H[\rho] + v_{\text{xc}}^{\text{approx}}[\rho] - (v_H[\rho_k] + v_{\text{xc}}^{\text{approx}}[\rho_k]). \quad (1.77)$$

Bei lokalisierten Systemen führt SIC häufig zu einer deutlich verbesserten Beschreibung. So lassen sich beispielsweise die Bindungsenergien negativ geladener Ionen sehr genau bestimmen. Zudem erhält man im Gegensatz zu LDA/GGA das korrekte asymptotische Abklingverhalten des effektiven Potentials und der Ladungsdichte.

SIC steht jedoch nicht im Einklang mit dem Kohn-Sham-Formalismus, denn auf Grund der Orbitalabhängigkeit des effektiven Potentials wird durch die Kohn-Sham-Gleichung (1.52) kein wechselwirkungsfreies System mehr beschrieben. Auch sind die Wellenfunktionen nicht mehr notwendigerweise orthogonal, wodurch die Teilchenzahlerhaltung verletzt werden kann. Zudem ist der Kohn-Sham-Operator in SIC nicht invariant gegenüber unitären Transformationen der Wellenfunktionen,¹² d.h., es ist im Allgemeinen $H_{\text{KS}}^{\text{SIC}}[\{\psi_i\}] \neq H_{\text{KS}}^{\text{SIC}}[\{U\psi_i\}]$ mit einer unitären Transformation U im Raum der besetzten Zustände.

1.5.4 Exakte Darstellung des Austausch-Korrelationsfunktionals als Störreihe

Abschnitt 1.5.2 stellte einige Methoden zur näherungsweise Berechnung des DFT-Austausch-Korrelationspotentials (1.51) vor. Kennzeichnend für die LDA- und GGA-Potentiale ist dabei, dass sie keinen Ansatz zu einer *systematischen* Verbesserung des Austausch-Korrelationspotentials aufzeigen. Tatsächlich existiert aber ein Schema, das — zumindest im Prinzip — dessen beliebig genaue Beschreibung zuließe. Dieses 1993 von Görling und Levy vorgeschlagene Verfahren [40, 41, 42], das nun, den genannten Referenzen folgend, skizziert werden soll, basiert auf einem störungstheoretischen Ansatz, und sei deshalb hier als „Görling-Levy-Störungstheorie“ bezeichnet. Mit ihr lassen sich Austausch-Korrelationsenergie E_{xc} und Austausch-Korrelationspotential v_{xc} über eine Reihenentwicklung exakt darstellen. Die Anforderungen an Rechnerkapazität und -leistung, um diese Reihe auszukonvergieren, übersteigen das heute technisch Mögliche allerdings bei weitem. Möglich ist jedoch ein Abbrechen der Entwicklung bereits nach dem ersten Glied; man erhält dann gerade den Austauschanteil des Austausch-Korrelationspotentials. In diesem Sinne kann man die Görling-Levy-Störungstheorie als Verallgemeinerung oder Weiterführung des EXX-Formalismus (vgl. Abschnitt 1.6 und Kapitel 2) bezeichnen.

Für das Folgende sei zunächst die N -Darstellbarkeit der Ladungsdichte (1.60) vorausgesetzt. Um nun eine Reihe für die Austausch-Korrelationswechselwirkung abzuleiten, wird der Ausdruck für die Korrelationsenergie (1.71) zunächst verallgemeinert:

$$E_c^\alpha[\rho] := \underbrace{[\langle \Psi^\alpha | T | \Psi^\alpha \rangle - \langle \Psi^0 | T | \Psi^0 \rangle]}_{=: T_c^\alpha[\rho]} + \alpha \cdot \underbrace{[\langle \Psi^\alpha | V_{\text{ee}} | \Psi^\alpha \rangle - \langle \Psi^0 | V_{\text{ee}} | \Psi^0 \rangle]}_{=: V_c^\alpha[\rho]}, \quad (1.78)$$

¹²Der korrekte Kohn-Sham-Operator ist invariant gegenüber unitären Transformationen der Wellenfunktionen ψ_i , da er allein von der Ladungsdichte abhängt, für die bei Isolatoren und Halbleitern $\rho = \sum_i^{\text{occ}} |\psi_i|^2 = \sum_i^{\text{occ}} |U\psi_i|^2$ mit einer beliebigen unitären Transformation U im Raum der besetzten Zustände gilt. Es sei im Hinblick auf die folgenden Abschnitte darauf hingewiesen, dass dann ähnlich auch für den nicht-lokalen Austauschoperator $V_{\text{Fock}}[\{\psi_i\}] = V_{\text{Fock}}[\{U\psi_i\}]$ gilt.

wobei Ψ^α wieder die gegebene Dichte ρ liefere und über das Minimum von $\langle \Psi | T + \alpha V_{ee} | \Psi \rangle$ definiert sei. T_c^α und αV_c^α bezeichnen den kinetischen Anteil und den Anteil der Elektron-Elektron-Abstoßung an der Korrelationsenergie E_c^α .

Mit dieser Definition lässt sich ein Hamilton-Operator konstruieren,

$$H^\alpha[\rho] := T + \alpha V_{ee} + \sum_{i=1}^N \left(v_{\text{ext}}(\mathbf{r}_i) + (1 - \alpha) v_H[\rho](\mathbf{r}_i) + (1 - \alpha) v_x[\rho](\mathbf{r}_i) + \frac{\delta E_c^1[\rho]}{\delta \rho(\mathbf{r}_i)} - \frac{\delta E_c^\alpha[\rho]}{\delta \rho(\mathbf{r}_i)} \right), \quad (1.79)$$

bei dem mit dem Parameter $\alpha \in [0, 1]$ zwischen dem Hamilton-Operator des Kohn-Sham-Systems ($\alpha = 0$) und dem des realen Systems ($\alpha = 1$) stetig hin- und hergeschaltet werden kann. Während sich die Wellenfunktionen, die sich aus dem entsprechenden Eigenwertproblem ergeben, im Allgemeinen mit α ändern, bleibt die Ladungsdichte dabei stets konstant. Aus diesem Grund wird (1.79) auch „adiabatische Verbindung“ zwischen dem Kohn-Sham- und dem realen System genannt.

Für die Korrelation, ihren kinetischen Anteil und ihren Wechselwirkungsanteil wird weiter angenommen, dass sie sich in Taylorreihen entwickeln lassen, deren Glieder noch zu bestimmen sind, und schreibt

$$E_c^\alpha[\rho] = \sum_{j=1}^{\infty} e_{c,j}[\rho] \alpha^j, \quad T_c^\alpha[\rho] = \sum_{j=1}^{\infty} t_{c,j}[\rho] \alpha^j, \quad V_c^\alpha[\rho] = \sum_{j=1}^{\infty} v_{c,j}[\rho] \alpha^j. \quad (1.80)$$

Unter Ausnutzung von $\partial T_c^\alpha[\rho]/\partial \alpha = -\alpha \cdot \partial V_c^\alpha[\rho]/\partial \alpha$ (Gl. (35) in Ref. 40), erhält man daraus

$$e_{c,j}[\rho] = \begin{cases} 0 & \text{für } j = 1 \\ -\frac{1}{j-1} t_{c,j}[\rho] = \frac{1}{j} v_{c,j-1}[\rho] & \text{für } j \geq 2 \end{cases}, \quad (1.81)$$

womit sich (1.79) auf folgende Gestalt bringen lässt:

$$H^\alpha[\rho] = \underbrace{H^0[\rho]}_{=: H_0} + \alpha \cdot \underbrace{\left[V_{ee} - \sum_{i=1}^N \left(v_H[\rho](\mathbf{r}_i) + v_x[\rho](\mathbf{r}_i) \right) \right]}_{=: H_1} + \sum_{j=2}^{\infty} \alpha^j \cdot \underbrace{\left[- \sum_{i=1}^N \frac{\delta e_{c,j}[\rho]}{\delta \rho(\mathbf{r}_i)} \right]}_{=: H_j}. \quad (1.82)$$

H^α kann also aufgefasst werden als Hamilton-Operator des Kohn-Sham-Systems, das sich unter dem Einfluss der Störungen $\alpha^j H_j$ ($j \geq 1$) befindet. Man beachte, dass die *unteren* Indizes an den Operatoren sich hier auf die Ordnung der Störung beziehen, im Gegensatz zu dem oberen Parameter α , der die Stärke der Störung beschreibt. Da die Störung bereits selbst eine Reihenentwicklung ist, wird das entsprechende Eigenwertproblem in einer etwas modifizierten Version der Störungstheorie von Rayleigh-Schrödinger behandelt,

$$\begin{aligned} (H_0 + \alpha H_1 + \alpha^2 H_2 + \dots) (\Psi_0 + \alpha \Psi_1 + \alpha^2 \Psi_2 + \dots) &= \\ &= (E_0 + \alpha E_1 + \alpha^2 E_2 + \dots) (\Psi_0 + \alpha \Psi_1 + \alpha^2 \Psi_2 + \dots). \end{aligned} \quad (1.83)$$

Die 0-te Ordnung beschreibt das Kohn-Sham-System selbst, es gilt also

$$E_0 = E_{\text{KS}}[\rho] \quad \text{mit} \quad \Psi_0 = \Psi^0. \quad (1.84)$$

Für die höheren Ordnungen erhält man durch Sortieren nach den Potenzen von α die Energiekorrekturen

$$E_j = \sum_{p=1}^j \langle \Psi_0 | H_p | \Psi_{j-p} \rangle \quad (j \in \mathbb{N}) . \quad (1.85)$$

Verwendet man weiter (1.82) in der Skalierungsbeziehung

$$E_c[\rho_\lambda] = \frac{1}{\alpha^2} [\langle \Psi^\alpha | H^\alpha | \Psi^\alpha \rangle - \langle \Psi^0 | H^\alpha | \Psi^0 \rangle] \quad (1.86)$$

mit $\lambda = \alpha^{-1}$ und $\rho_\lambda(x, y, z) := \lambda^3 \rho(\lambda x, \lambda y, \lambda z)$ [40], findet man durch Koeffizientenvergleich mit (1.80) schließlich eine Rekursionsformel für die Koeffizienten mit $j \geq 2$ der Taylorentwicklung:

$$\begin{aligned} e_{c,j}[\rho] &= E_j - \langle \Psi_0 | H_j | \Psi_0 \rangle \\ &\stackrel{(1.85)}{=} \sum_{p=1}^j \langle \Psi_0 | H_p | \Psi_{j-p} \rangle - \langle \Psi_0 | H_j | \Psi_0 \rangle \\ &= \sum_{p=1}^{j-1} \langle \Psi_0 | H_p | \Psi_{j-p} \rangle \\ &\stackrel{(1.82)}{=} \langle \Psi_0 | H_1 | \Psi_{j-1} \rangle + \sum_{p=2}^{j-1} \left\langle \Psi_0 \left| - \sum_{i=1}^N \frac{\delta e_{c,p}[\rho]}{\delta \rho(\mathbf{r}_i)} \right| \Psi_{j-p} \right\rangle , \end{aligned} \quad (1.87)$$

wobei im letzten Schritt die Definition der H_p aus (1.82) eingesetzt worden ist. Die Auswertung der Taylorreihe für die Austausch-Korrelationsenergie (1.80) bis zur Ordnung j verlangt somit nur die Kenntnis der Funktionalableitungen ihrer Koeffizienten von kleinerer Ordnung als j und kann formal in beliebiger Genauigkeit erfolgen! Wie in Ref. 41 gezeigt wird, sind die jeweils benötigten Funktionalableitungen auch tatsächlich analytisch zugänglich.¹³ Damit ergibt sich dann auch das (verallgemeinerte) Austausch-Korrelationspotential

$$v_{xc}^\alpha[\rho] := \frac{\delta}{\delta \rho} (E_x[\rho] + E_c^\alpha[\rho]) = v_x[\rho] + \sum_{j=1}^{\infty} \alpha^j v_{c,j}[\rho] \quad \text{mit} \quad v_{c,j}[\rho] := \frac{\delta e_{c,j}[\rho]}{\delta \rho} , \quad (1.88)$$

das für $\alpha = 1$ dem exakten Austausch-Korrelationspotential entspricht. Wegen (1.81) ergibt sich in erster Ordnung das exakte Austauschpotential $v_x[\rho]$ von (1.68). Dessen Berechnung — und im Prinzip über die $v_{c,j}[\rho]$ auch die sukzessive Berechnung des exakten *Korrelationspotentials* — gelingt unter bestimmten Voraussetzungen über zwei Wege: (i) direktes Ableiten der Austauschenergie $E_x[\rho]$ bzw. der Energieterme $e_{c,j}[\rho]$ nach der Ladungsdichte ρ , (ii) Auswerten des gestörten Eigenwertproblems (1.82) für die verschiedenen Ordnungen von α . Dies soll im folgenden Abschnitt 1.6 beschrieben werden, allerdings nur für v_x — eine weiterführende Auswertung der Störreihe ist, wie eingangs bereits angedeutet, numerisch noch nicht sinnvoll durchführbar und daher rein formaler Natur. Weg (ii) führt dabei direkt auf eine implizite Gleichung für v_x , der so genannten OEP-Gleichung, die den Ausgangspunkt für die Entwicklung des in Kapitel 2 beschriebenen variationellen EXX-Verfahrens [78, 79, 80] darstellt.

¹³ An dieser Stelle wird jedoch v -Darstellbarkeit vorausgesetzt, da eine ein-eindeutige Beziehung zwischen effektiven Potential und Ladungsdichte (vgl. Hohenberg-Kohn-Theorem I [1]) benutzt wird. Formel (1.87) benötigt hingegen nur die schwächere Annahme der N -Darstellbarkeit.

1.6 Exakte Austauschpotentiale

In voller Allgemeinheit wäre das exakte Austauschpotential als Funktionalableitung von (1.70) gegeben,

$$v_x[\rho] := \frac{\delta E_x[\rho]}{\delta \rho} = \frac{\delta}{\delta \rho} \langle \Psi^0 | (V_{ee} - v_H[\rho]) | \Psi^0 \rangle, \quad (1.89)$$

wobei alle N -darstellbaren Ladungsdichten zugelassen sind und die N -Teilchen-Wellenfunktionen Ψ^0 die Beziehung (1.65) erfüllen. In dieser Allgemeinheit lässt sich die Funktionalableitung allerdings (bislang) nicht auswerten. Dies gelingt jedoch, wenn man annimmt, dass die Kohn-Sham-Zustände einfache Slater-Determinanten der besetzten Ein-Teilchen-Zustände ψ_i sind, und die Austauschenergie $E_x[\rho]$ damit wie in (1.69) ausgedrückt werden kann. Hinreichende Bedingung dafür ist die v -Darstellbarkeit der Ladungsdichten [21, 43], die daher im weiteren Verlauf dieser Arbeit stets, wenn nicht ausdrücklich anders gesagt, vorausgesetzt wird. Eine formale Erweiterung des Konzepts auf N -darstellbare Dichten konnte jedoch im Rahmen der Arbeit entwickelt werden und ist in Anhang E beschrieben.

Die zur Hartree-Fock-Theorie analoge Gestalt von E_x hat nun den Vorteil, dass damit auch in der DFT die exakte Behandlung des *Austauschanteils* bereits ausreicht, um die unphysikalische elektronische Selbstwechselwirkung vollständig zu eliminieren, vgl. Abschnitt 1.3. Es ist daher zu erwarten, dass Fehler, die auf deren unvollständige Eliminierung auf Grund einer nur näherungsweise Beschreibung des Austauschs zurückzuführen sind (vgl. Einleitung), bei Verwendung des exakten $v_x[\rho]$ korrigiert werden. Da der Austauschanteil zudem die dominierende Rolle bei der Austausch-Korrelationswechselwirkung spielt — typischerweise übertrifft er die Korrelation um eine Größenordnung [44, 45, 46] —, kommt seiner genauen Kenntnis besondere Relevanz zu.

1.6.1 Das Standard-EXX-Verfahren

Ausgehend von der Definition der Austauschenergie (1.69) lässt sich das Austauschpotential durch zweimaliges Anwenden der Kettenregel für Variationsableitungen finden, wie zuerst Görling und Levy zeigten [41], vgl. aber auch [47, 48, 52, 53, 54]:

$$v_x[\rho](\mathbf{r}) = \sum_i^{\text{occ}} \int d^3r' \int d^3r'' \left(\frac{\delta E_x[\{\psi_k\}]}{\delta \psi_i^*(\mathbf{r}')} \frac{\delta \psi_i^*(\mathbf{r}')}{\delta v_{\text{eff}}(\mathbf{r}'')} + \text{c.c.} \right) \frac{\delta v_{\text{eff}}(\mathbf{r}'')}{\delta \rho(\mathbf{r})}. \quad (1.90)$$

Die Abkürzung „occ.“ (*occupied*) soll hier wie immer im Folgenden andeuten, dass nur über die besetzten Zustände zu summieren ist. Die Methode, das lokale Austauschpotential v_x durch direktes Auswerten der rechten Seite dieser Gleichung zu bestimmen, wird als *Exact-Exchange-Verfahren* (EXX) bezeichnet. Diese Terminologie geht auf Langreth und Mehl [29], Kotani [57], sowie auf Städele *et al.* [52] zurück. Zur Unterscheidung von der in Kapitel 2 vorgestellten variationellen Bestimmung von v_x sei dieses Verfahren in der vorliegenden Arbeit auch als „Standard-EXX-Verfahren“ (stEXX) bezeichnet. Alternativ wird auch der Begriff *Optimized-Effektive-Potential-Methode* (zwei Abkürzungen, OEP und OPM, sind gebräuchlich) verwendet, der aber ursprünglich nicht aus der „Tradition“ der DFT herrührt. Genau genommen versteht man unter OEP [63, 65, 66, 67] auch nicht das direkte Berechnen von v_x gemäß (1.90), sondern das Lösen einer äquivalenten Integralgleichung, vgl. Abschnitt 1.6.2.

Um Gleichung (1.90) benutzen zu können, werden noch die drei Funktionalableitungen der rechten Seite benötigt. Die erste berechnet sich analog zu Gleichung (1.22) aus

Abschnitt 1.3 und entspricht der Anwendung des Fock-Operators (1.27) auf die i -te Wellenfunktion:

$$\frac{\delta E_x[\{\psi_k\}]}{\delta \psi_i^*(\mathbf{r})} = - \sum_j^{\text{occ}} \delta_{\sigma_i \sigma_j} \int \frac{\psi_j^*(\mathbf{r}') \psi_i(\mathbf{r}') \psi_j(\mathbf{r})}{|\mathbf{r} - \mathbf{r}'|} d^3 r' = \left(V_{\text{Fock}}[\{\psi_j\}] \psi_i \right)(\mathbf{r}). \quad (1.91)$$

Die Variationsableitung der Zustände nach dem effektiven Potential findet man mit Hilfe der Störungstheorie in 1. Ordnung, vgl. Anhang B:

$$\frac{\delta \psi_i(\mathbf{r}')}{\delta v_{\text{eff}}(\mathbf{r})} = \sum_{\substack{j \\ \varepsilon_j \neq \varepsilon_i}} \psi_j(\mathbf{r}') \frac{\psi_j^*(\mathbf{r}) \psi_i(\mathbf{r})}{\varepsilon_i - \varepsilon_j} = G_i(\mathbf{r}', \mathbf{r}) \psi_i(\mathbf{r}). \quad (1.92)$$

Die Summe geht hier über *alle* Kohn-Sham-Orbitale mit Ausnahme jener mit $\varepsilon_j = \varepsilon_i$, d.h., auch die unbesetzten Zustände müssen berücksichtigt werden. Ferner bezeichnet G_i die in (B.11) definierte Greensche Funktion der Eigenwert-Gleichung (1.52) im zum Eigenraum von ε_i orthogonalen Unter-Hilbertraum.

Die Handhabung des letzten Faktors in (1.90) ist etwas schwieriger. Er kann als Inverse des linearen Antwortoperators χ_0 aufgefasst werden, der die Reaktion der Ladungsdichte auf eine „infinitesimal kleine“ Änderung des effektiven Potentials beschreibt. Es ergibt sich:

$$\chi_0(\mathbf{r}, \mathbf{r}') := \frac{\delta \rho(\mathbf{r})}{\delta v_{\text{eff}}(\mathbf{r}')} = \sum_i^{\text{occ.}} \sum_j^{\text{unocc.}} \frac{\psi_i^*(\mathbf{r}) \psi_j(\mathbf{r}) \psi_j^*(\mathbf{r}') \psi_i(\mathbf{r}') + \text{c.c.}}{\varepsilon_i - \varepsilon_j}. \quad (1.93)$$

Für die in (1.90) benötigte Umkehrung $\chi_0^{-1} = \delta v_{\text{eff}} / \delta \rho$ existiert nun allerdings keine analytisch geschlossene Darstellung mehr. Streng genommen ist sie auch gar nicht definiert, da das Potential v_{eff} nur bis auf eine Konstante bestimmt ist, und einer Änderung der Dichte deshalb unendlich viele Potentialänderungen zugeordnet werden können. In geeigneten Basissatzdarstellungen von χ_0^{-1} , die z.B. in [41, 52, 55, 56] diskutiert werden, lassen sich konstante Verschiebungen jedoch explizit ausschließen. Wie Görling und Levy in Anhang B von Ref. 41 zeigen, korrespondiert dies im Definitionsbereich der Dichten auch gerade mit der physikalisch sinnvollen Forderung $\int \delta \rho(\mathbf{r}) d^3 r = 0$, also der Erhaltung der Ladung im System. Im Falle periodischer Systeme lässt sich dies z.B. einfach durch die Herausnahme der ($\mathbf{G} = 0$)-Komponente im reziproken Raum erreichen [52]. Es sei an dieser Stelle auch bemerkt, dass hier wiederum die v -Darstellbarkeit benötigt wird: für ein nicht- v -darstellbares ρ ist $v_{\text{eff}}[\rho]$ und damit auch $v_{\text{eff}}[\rho] / \delta \rho$ nicht mehr ohne Weiteres¹⁴ wohldefiniert, vgl. auch Ref. 47.

Durch Einsetzen der obigen Ausdrücke (1.91) und (1.92) in Gleichung (1.90) hat man dann

$$v_x[\rho](\mathbf{r}) = \int d^3 r' \chi_0^{-1}(\mathbf{r}, \mathbf{r}') \underbrace{\sum_i^{\text{occ}} \int d^3 r'' \left[\psi_i^*(\mathbf{r}') G_i^\perp(\mathbf{r}', \mathbf{r}'') V_{\text{Fock}}[\{\psi_j\}] \psi_i(\mathbf{r}'') + \text{c.c.} \right]}_{=: \Delta \rho_{\text{Fock}}}, \quad (1.94)$$

wobei die G_i^\perp die auf den Raum der unbesetzten Zustände projizierten Greenschen Funktionen (B.12) sind. Bereits mit Blick auf den weiteren Verlauf der Arbeit sei Folgendes angemerkt: Nach der Rayleigh-Schrödingerschen Störungstheorie kann der Ausdruck

¹⁴Eine Erweiterung der Definition ist jedoch formal möglich, vgl. Anhang E.

$G_i V_{\text{Fock}}[\{\psi_j\}] \psi_i$ als lineare Korrekturfunktion interpretiert werden, die sich bei Störung des Kohn-Sham-Systems durch $V_{\text{Fock}}[\{\psi_j\}]$ ergibt. In (1.94) besitzt aber damit die Summe über i die Gestalt einer später in (2.4) definierten „induzierten Ladungsdichte“. Im Unterschied zu dort wird sie hier allerdings durch den nicht-lokalen Austauschoperator erzeugt und sei deshalb mit $\Delta\rho_{\text{Fock}}$ bezeichnet.

Ähnlich wie das Funktional (1.44) der kinetischen Energie T_s ist also auch v_x nur über den „Umweg“ der Kohn-Sham-Orbitale ψ_i zugänglich, eine direkte parametrische Abhängigkeit von ρ in geschlossener Form gibt der EXX-Formalismus nicht. Gleichung (1.94) liefert dann das Austauschpotential der zu den ψ_i gehörende Dichte, die jedoch nicht notwendigerweise die Grundzustandsdichte ρ_0 ist. Um ρ_0 und das zugehörige $v_x[\rho_0]$ zu finden, muss (1.94) daher selbstkonsistent zusammen mit den Kohn-Sham-Gleichungen (1.52) und (1.53) gelöst werden.

Beim Lösen der Kohn-Sham-Gleichungen können dann Korrelationseffekte entweder näherungsweise berücksichtigt werden, beispielsweise durch den Korrelationsteil des LDA-Potentials (1.74), oder völlig vernachlässigt werden. Letzteres wird häufig, so auch in der vorliegenden Arbeit, mit der Bezeichnung „*exchange-only*“ gekennzeichnet.

Während die Methode — allerdings in einer modifizierten Form (OEP, vgl. den folgenden Abschnitt) — auf atomare Systeme bereits lange erfolgreich angewandt wird, stößt man im Falle größerer Systeme, wie beispielsweise Festkörpern, schnell an Grenzen. Diese resultieren daraus, dass (i) bei der Berechnung der letzten beiden Faktoren in (1.90) unbesetzte Zustände ins Spiel kommen (Greensche Funktion), (ii) aus der Notwendigkeit, den linearen Antwortoperator χ_0 zu invertieren und (iii) aus dem Aufwand bei der Berechnung des nicht-lokalen Austauschoperators V_{Fock} (vgl. Abschnitt 4.4). Das Umgehen von zumindest der ersten beiden der genannten Schwierigkeiten ist ein zentrales Anliegen der vorliegenden Arbeit und wird in Kapitel 2 ausführlich dargelegt.

Abschließend noch ein paar Bemerkungen zur Unterscheidung des DFT-EXX-Schemas vom Hartree-Fock-Formalismus:

- (i) Obwohl die Austauschenergien in beiden Fällen formal über den selben Ausdruck gegeben sind, vgl. (1.18) und (1.69), sind sie verschieden, da sich die Kohn-Sham-Orbitale von den Hartree-Fock-Orbitalen (um mehr als eine unitäre Transformation) unterscheiden, siehe auch Punkt (iv). Vergleichende Berechnungen, die bisher für Atome und kleine Moleküle durchgeführt worden sind [49, 58, 59, 60, 61, 68, 78], vgl. auch Ref. 62 mit Ref. 39, legen jedoch nahe, dass sowohl die Wellenfunktionen als auch die Austauschenergien in beiden Formalismen nur sehr gering voneinander abweichen. Dieser Punkt wird noch einmal in Abschnitt 2.4 aufgegriffen.
- (ii) Im Hartree-Fock-Schema wird der Austausch über ein *nicht-lokales*, d.h., auf jedes Hartree-Fock-Orbital anders wirkendes Potential berücksichtigt (vgl. Abschnitt 1.3). Im Gegensatz dazu ist das Austauschpotential der Kohn-Sham-Gleichungen *lokal* und beschreibt daher eine Kraft, die auf alle Kohn-Sham-Teilchen gleich wirkt (diese können sich gegenseitig nicht „sehen“). Zusammen mit dem Hartree- und dem Korrelationspotential beschreibt es die gesamte elektronische Wechselwirkung des realen Systems gewissermaßen in einer gemittelten Form.
- (iii) Diese Beschreibung kann im Kohn-Sham-Schema — zumindest im Prinzip — völlig ohne Näherung erfolgen, wenn das exakte Austausch-Korrelationspotential verwendet würde. Das lokale Austauschpotential lässt sich dabei nach Gleichung (1.88) des vorangehenden Abschnitts als erstes Glied einer entsprechenden Taylorentwicklung

auffassen. Im Gegensatz dazu ist Hartree-Fock jedoch bereits konzeptuell als Näherung angelegt, wobei allerdings auch hier eine systematische Verbesserung in Form der *Configuration Interaction* existiert.

- (iv) Im Hartree-Fock-Formalismus ist der Grundzustand durch diejenige Slater-Determinante definiert, die die Gesamtenergie des Systems $\langle \Psi | H | \Psi \rangle$ minimiert. Für die Minimierung werden dabei außer der Orthonormalität der Ein-Teilchen-Orbitale keine Einschränkungen angenommen. Im “exchange-only” Kohn-Sham-Formalismus wird ebenfalls die Gesamtenergie $\langle \Psi | H | \Psi \rangle$ minimiert, zur Minimierung werden jedoch hier nur Slater-Determinanten aus normierten Ein-Teilchen-Zuständen zugelassen, die sich als Grundzustand eines nicht-wechselwirkenden Referenzsystems mit *irgendeinem* lokalen Potential (dem effektiven “exchange-only” Kohn-Sham-Potential) darstellen lassen. Ganz deutlich erkennt man dies anhand von Gleichung (1.99) im nachfolgend besprochenen OEP-Formalismus. Die Orthogonalität der Orbitale folgt übrigens daraus, denn ein reelles lokales Potential liefert immer einen hermiteschen Hamilton-Operator.

Daran lassen sich Schranken für die Gesamtenergien E_{HF} , $E_{\text{KS(EXX)}}$ und $\tilde{E}_{\text{KS(X)}}$ des Systemgrundzustands im Hartree-Fock-, im “exchange-only” Kohn-Sham-EXX- bzw. in einem genäherten “exchange-only” Kohn-Sham-Schema ablesen: Da die Minimierung im Hartree-Fock-Schema weniger Nebenbedingungen unterliegt als bei Kohn-Sham, folgt, dass $E_{\text{KS(EXX)}}$ nicht kleiner als E_{HF} sein kann. Jedes *genäherte* effektive “exchange-only” Kohn-Sham-Potential kann wiederum nicht auf die niedrigste “exchange-only” Kohn-Sham-Gesamtenergie (also die Gesamtenergie in EXX) führen [87]. Allgemein gilt also

$$E_{\text{HF}} \leq E_{\text{KS(EXX)}} < \tilde{E}_{\text{KS(X)}} .$$

1.6.2 Das OEP-Verfahren

Bereits sehr früh, und zunächst sogar völlig unabsichtlich, fand man exakte, “exchange-only” Kohn-Sham-Potentiale für atomare, radialsymmetrische Probleme [65, 66, 67]. Dabei wurde v_{x} nicht direkt berechnet wie im vorangehenden Abschnitt, sondern durch Lösen der mit

$$\int \chi_0^{-1}(\mathbf{r}', \mathbf{r}) \chi_0(\mathbf{r}, \mathbf{r}'') d^3r = \delta(\mathbf{r}' - \mathbf{r}'') \quad (1.95)$$

zu (1.90) äquivalenten Integralgleichung

$$\int d^3r' v_{\text{x}}[\rho](\mathbf{r}') \chi_0(\mathbf{r}', \mathbf{r}) = \sum_i^{\text{occ}} \int \left(\frac{\delta E_{\text{x}}[\{\psi_i\}]}{\delta \psi_i^*(\mathbf{r}')} \frac{\delta \psi_i^*(\mathbf{r}')}{\delta v_{\text{eff}}(\mathbf{r})} + \text{c.c.} \right) d^3r' . \quad (1.96)$$

Einsetzen der Ausdrücke (1.91) – (1.93) für die Funktionalableitungen liefert dann

$$\sum_i^{\text{occ}} \int d^3r' \left[\psi_i^*(\mathbf{r}) G_i^\perp(\mathbf{r}, \mathbf{r}') \left(V_{\text{Fock}}[\{\psi_j\}] \psi_i(\mathbf{r}') - v_{\text{x}}(\mathbf{r}') \cdot \psi_i(\mathbf{r}') \right) + \text{c.c.} \right] = 0 \quad (1.97)$$

mit den auf die unbesetzten Zustände projizierten Greenschen Funktionen G_i^\perp von (B.12). Dies ist die so genannte OEP-Gleichung. Interessanterweise wurde sie von Sharp und Horton bereits 1953 angegeben [63], d.h., vor Entwicklung der DFT! Die beiden Autoren folgten dabei einer Idee Slaters [64], eine einfache Näherung für die Hartree-Fock-Methode zu

finden, indem man dasjenige *lokale* Potential konstruiert, das das *nicht-lokale* Austauschpotential (1.27) der Hartree-Fock-Gleichungen am besten annähert. Die Bezeichnung OEP (*optimized effective potential*) wurde später von Talman, Shadwick, Aashamar und Luke geprägt, die einen ähnlichen Ansatz verfolgten [65, 66, 67]. Hier wird versucht, ein lokales Potential v_{OEP} der formalen Schrödinger-Gleichung

$$\left[-\frac{1}{2}\nabla^2 + v_{\text{OEP}}(\mathbf{r}) \right] \psi_i(\mathbf{r}) = \varepsilon_i \psi_i(\mathbf{r}) \quad (1.98)$$

zu finden, so dass die Slater-Determinante Ψ der besetzten Einteilchen-Wellenfunktionen ψ_i das Energiefunktional $\langle H \rangle := \langle \Psi | H | \Psi \rangle$ stationär werden lässt:

$$\frac{\delta \langle H \rangle}{\delta v_{\text{OEP}}(\mathbf{r})} = \sum_i^{\text{occ}} \int \frac{\delta \langle H \rangle}{\delta \psi_i^*(\mathbf{r}')} \frac{\delta \psi_i^*(\mathbf{r}')}{\delta v_{\text{OEP}}(\mathbf{r})} d^3 r' + \text{c.c.} \stackrel{!}{=} 0. \quad (1.99)$$

Zunächst ebenfalls als Näherung für Hartree-Fock gedacht, kann dieses Variationsprinzip für das lokale Potential einer effektiven Ein-Teilchen-Gleichung im Rahmen des Kohn-Sham-Formalismus als Schema zur Konstruktion des exakten, “exchange-only” Kohn-Sham-Potentials interpretiert werden. Darauf wurde zuerst von Perdew und Mitarbeitern hingewiesen [68, 69]. Tatsächlich ist dieses Variationsprinzip äquivalent zu Gleichung (1.97), wenn man beachtet, dass der erste Faktor der rechten Seite, $\delta \langle H \rangle / \delta \psi_i^*$, formal gerade der linken Seite der Hartree-Fock-Gleichung (1.24) entspricht und der zweite Faktor analog zu (1.92) gegeben ist. Einsetzen, Umformen und Vergleichen mit (1.97) führt schließlich auf

$$v_{\text{OEP}}(\mathbf{r}) = v_{\text{ext}}(\mathbf{r}) + v_{\text{H}}(\mathbf{r}) + v_{\text{x}}(\mathbf{r}). \quad (1.100)$$

Dass OEP anfangs nur auf radialsymmetrische Systeme Anwendung fand [65, 66, 67], ist wohl im Wesentlichen auf die damals zur Verfügung stehenden Rechnerkapazitäten zurückzuführen. Zum einen muss Gleichung (1.97) in diesem Fall nur in einer Dimension für den Radialteil gelöst werden, zum anderen ist auch die Randbedingung $\lim_{r \rightarrow \infty} v_{\text{OEP}}(r) = 0$ recht leicht zu handhaben. Letztlich ist aber das numerische Lösen der Integralgleichung (1.96) mit Hilfe eines linearen Gleichungssystems äquivalent zur Inversion des linearen Antwortoperators χ_0 von (1.93) und führt auf dieselben Probleme: auch die Integralgleichung bestimmt v_{x} bzw. v_{OEP} nur bis auf eine additive Konstante; das entsprechende Gleichungssystem ist also unterbestimmt.

Eine sehr genaue und daher wichtige Näherung für OEP ist die von Krieger, Li und Iafrate entwickelte und nach ihnen benannte KLI-Methode [70, 71, 72, 73]. Sie beruht auf einer ebenfalls bereits in der oben zitierten Veröffentlichung von Sharp und Horton [63] vorgeschlagenen Ersetzung der Eigenwertdifferenzen $\varepsilon_j - \varepsilon_i$ im Nenner der Greenschen Funktionen G_{ij}^{\pm} durch einen geeigneten konstanten Wert. Ausführliche Darstellungen der Methode finden sich z.B. in den Referenzen 48, 49, 50. KLI lässt sich vergleichsweise einfach implementieren und auch auf Halbleiter und Isolatoren anwenden. Dies wurde beispielsweise eingehend von Bylander und Kleinman untersucht [74, 75, 76, 77].

1.6.3 Störungstheoretische Herleitung der OEP/EXX-Gleichung

In Abschnitt 1.6.1 wurde in Gleichung (1.90) das exakte Austauschpotential $v_{\text{x}}[\rho]$ mit Hilfe der Kettenregel direkt aus der Definition der Austauschenergie (1.69) abgeleitet. Alternativ lässt sich diese Gleichung — bzw. die zu ihr äquivalente OEP-Gleichung (1.97) — aber auch elegant aus der in Abschnitt 1.5.4 vorgestellten störungstheoretischen Entwicklung des

Austausch-Korrelationsfunctionals erhalten, wie Görling und Levy in Ref. 42 zeigen. Diese Vorgehensweise ist hier insofern von Bedeutung, als sie auf der „Ebene“ der *Viel-Teilchen*-Wellenfunktionen das genaue Analogon zu der in dieser Arbeit untersuchten variationellen EXX-Methode [78, 79, 80] darstellt, die auf einer störungstheoretischen Behandlung der *Ein-Teilchen*-Wellenfunktionen beruht.

Der in der Görling-Levy-Störungstheorie verwendete Ansatz (1.83) liefert in linearer Ordnung für Energie- und Zustandskorrektur noch Ausdrücke, die mit denen der Rayleigh-Schrödinger-Störungstheorie identisch sind. Mit dem in (1.82) definierten Störoperator H_1 ergibt sich die lineare Zustandskorrektur

$$\Psi_1 = \sum_{k=1}^{\infty} \Phi_k \frac{\langle \Phi_k | V_{ee} - v_H[\rho] - v_x[\rho] | \Phi_0 \rangle}{E_0 - E_k}. \quad (1.101)$$

Hier bezeichne $\Phi_0 := \Psi_0$ den Grundzustand des ungestörten Kohn-Sham-Systems, und die Φ_k mit $k \geq 1$ seien dessen virtuell¹⁵ angeregte Zustände. Alle Φ_k ($k \in \mathbb{N}_0$) sind dabei durch einfache Slater-Determinanten gegeben, die aus den Kohn-Sham-Orbitalen ψ_i von Gleichung (1.52) aufgebaut sind. Ferner sollen die E_k die zu den Determinanten Φ_k gehörenden Kohn-Sham-Energien bezeichnen.

Alle Zustände $\Psi^\alpha = \Psi_0 + \alpha\Psi_1 + \dots$, die sich in (1.82) entlang der „adiabatischen Verbindung“, also für alle $\alpha \in [0, 1]$, ergeben, führen auf die gleiche Ladungsdichte ρ . Anders, und bereits im Hinblick auf Kapitel 2 formuliert: das Aufschalten der Störung durch „Hochdrehen“ von α darf keine Änderung der Ladungsdichte bewirken. Mit dem Dichteoperator $\hat{\rho}(\mathbf{r})$ folgt also

$$\begin{aligned} \langle \Psi_0 | \hat{\rho}(\mathbf{r}) | \Psi_0 \rangle &= \langle \Psi^\alpha | \hat{\rho}(\mathbf{r}) | \Psi^\alpha \rangle \\ &= \langle \Psi_0 | \hat{\rho}(\mathbf{r}) | \Psi_0 \rangle + \alpha \langle \Psi_0 | \hat{\rho}(\mathbf{r}) | \Psi_1 \rangle + \alpha \langle \Psi_1 | \hat{\rho}(\mathbf{r}) | \Psi_0 \rangle + \\ &\quad + \text{Terme der Ordnung 2 oder höher in } \alpha. \end{aligned}$$

Damit die Gleichheit für jedes $\alpha \in [0, 1]$ erfüllt ist, müssen alle Entwicklungskoeffizienten bis auf den der 0. Ordnung verschwinden. In linearer Ordnung ergibt sich damit

$$\langle \Psi_0 | \hat{\rho}(\mathbf{r}) | \Psi_1 \rangle + \langle \Psi_1 | \hat{\rho}(\mathbf{r}) | \Psi_0 \rangle = 0. \quad (1.102)$$

Setzt man hier den Ausdruck (1.101) für die 1. Zustandskorrektur ein und nutzt aus, dass nur einfach angeregte Slater-Determinanten Φ_k einen Beitrag zur linken Seite von (1.102) liefern, erhält man daraus eine Gleichung in den Ein-Teilchen-Wellenfunktionen ψ_i des Kohn-Sham-Systems (schrittweise in Anhang G vorgeführt):

$$\sum_i^{\text{occ.}} \sum_j^{\text{unocc.}} \left[\psi_i^*(\mathbf{r}) \psi_j(\mathbf{r}) \frac{\langle \psi_j | V_{\text{Fock}}[\{\psi_k\}] - v_x[\rho] | \psi_i \rangle}{\varepsilon_i - \varepsilon_j} + \text{c.c.} \right] = 0, \quad (1.103)$$

wobei $V_{\text{Fock}}[\{\psi_k\}]$ wieder von der Form des Fock-Operators (1.27) ist. Diese Gleichung ist, wenn auch etwas anders aufgeschrieben, völlig identisch mit der OEP-Gleichung (1.97). Damit ist sie auch äquivalent zur EXX-Gleichung (1.94), wenn man von den in den vorangehenden beiden Abschnitten beschriebenen Schwierigkeiten bei dieser Umformung einmal absieht. Die linke Seite von (1.102) wird im variationellen EXX-Verfahren dann über die Ein-Teilchen-Zustände formuliert und stellt in diesem Konzept die zentrale Größe der „induzierten Ladungsdichte“ dar. Wegen der grundlegenden Bedeutung dieser Methode im Rahmen der vorliegenden Arbeit sei ihr im Folgenden ein eigenes Kapitel gewidmet.

¹⁵Es sei unterschieden zwischen den *virtuellen* Anregungen, die sich aus den unbesetzten Zuständen einer Berechnung des N -Elektronen-Systems ergeben, und den Anregungen, die über die Elektronenaffinität A und die Ionisierungsenergie I durch $(I - A)$ definiert sind.

Kapitel 2

Das variationelle EXX-Verfahren

In der Einleitung wurde auf eine Reihe von Nachteilen verwiesen, die die Verwendung genäherter Austausch-Korrelationsfunktionale wie beispielsweise LDA oder GGA mit sich bringt und die auf eine unvollständige Eliminierung der elektronischen Selbstwechselwirkung zurückzuführen sind. Diese Probleme lassen sich mit dem im Vorangehenden vorgestellten EXX/OEP-Formalismus überwinden, da hier die Eliminierung ganz analog zum Hartree-Fock-Schema vollständig gelingt. Der Preis, den man dafür zu zahlen hat, ist ein immenser numerischer Aufwand. OEP hat deshalb fast nur bei radialsymmetrischen, effektiv eindimensionalen Problemen Anwendung finden können, vgl. Abschnitt 1.6.2. Jedoch auch in der breiter anwendbaren EXX-Formulierung (1.94) skaliert allein die Inversion des linearen Antwortoperators χ_0 kubisch mit der Abschneideenergie E_{cut} (kritisch bzgl. der CPU-Zeit), und die Anzahl der benötigten unbesetzten Zustände ist beispielsweise bei Verwendung einer Ebene-Wellen-Basis (vgl. Abschnitt 3.3) typischerweise um zwei Größenordnungen höher als die der besetzten (kritisch bzgl. des Speicherbedarfs). Eine drastische Restriktion der Anwendungsmöglichkeiten ist die Folge.

Mehr theoretischer Natur ist Frage, weshalb *überhaupt* unbesetzte Orbitale benötigt werden. In der störungstheoretischen Herleitung der OEP-Gleichung nach Abschnitt 1.6.3 kamen diese durch Slater-Determinanten ins Spiel, die (virtuelle¹) einfach angeregte Zustände des Kohn-Sham-Systems repräsentieren. Steht dies nicht aber im Widerspruch zum Charakter der Kohn-Sham-DFT als einer reinen Grundzustandstheorie?

Um die Antwort gleich vorweg zu nehmen: Nein, denn es zeigt sich, dass die unbesetzten Zustände zur Berechnung des exakten Austauschpotentials nur als Hilfsgrößen benutzt werden und auch völlig umgangen werden können. Dies wurde unabhängig voneinander zunächst von Wahn und Neugebauer [78] und wenige Zeit später von Kümmel und Perdew [79, 80] nachgewiesen. Bei geschickter Formulierung des Problems wird ferner auch die explizite Berechnung von χ_0 und seiner Inversen vermieden, so dass sowohl die speicher- als auch die zeitkritischen Aspekte des EXX-Formalismus völlig eliminiert werden können. Die genannten Autoren entwickelten daraus eine Methode, die ausschließlich auf Grundzustandsgrößen basiert und in der das exakte Austauschpotential durch Minimierung eines v_x -abhängigen Funktionals, der (Norm der) „induzierten Ladungsdichte“ berechnet wird. Daher wird dieses Verfahren im Folgenden als *variationelles EXX-Verfahren* (varEXX) bezeichnet. Es wurde bisher erfolgreich im Falle von Atomen [78, 79] und Clustern [80] getestet. In der vorliegenden Arbeit wurde es erstmalig in einen Pseudopotential-Ebene-Wellen-Code [81, 82] implementiert (vgl. Kapitel 4), womit eine Erweiterung seines Anwendungsbereichs auch auf den Fall periodischer Systeme ermöglicht wurde (vgl. Kapitel 5

¹Vgl. Fußnote auf Seite 33.

und 7). Das Verfahren sei im Folgenden vorgestellt.

2.1 Idee und mathematische Formulierung

Die von Wahn und Neugebauer gegebene Herleitung des variationellen EXX-Verfahrens (vgl. Kapitel 2 in Ref. 78) ähnelt der Vorgehensweise, die OEP-Gleichung aus der Göring-Levy-Störungstheorie [40, 41, 42] zu entwickeln (vgl. Abschnitt 1.6.3). Ferner werden derartige Ansätze auch schon seit längerer Zeit in der Dichtefunktional-Störungstheorie [83, 84] verwendet, etwa zur Berechnung von Phononenspektren, wie z.B. in dem Übersichtsartikel von Baroni *et al.* [85] beschrieben.

Ausgangspunkt des variationellen EXX ist die formal gleiche Gestalt der Austauschenergien sowohl im Hartree-Fock- als auch im Kohn-Sham-Schema. Der Grundgedanke besteht darin, die Eigenschaften des Fock-Operators als exaktes, aber nicht-lokales Austauschpotential der Hartree-Fock-Gleichung (1.24) mittels einer Störung auf das lokale Austauschpotential der Kohn-Sham-Gleichung (1.52) „herüberzuschieben“. Mit dem Ansatz

$$H^\alpha := \underbrace{-\frac{1}{2}\nabla^2 + v_{\text{ext}} + v_{\text{H}}[\rho] + v_{\text{x}}[\rho] + v_{\text{c}}[\rho]}_{=: H_0} + \alpha \cdot \underbrace{\left(V_{\text{Fock}}[\{\psi_i\}] - v_{\text{x}}[\rho] \right)}_{=: H_1} \quad (2.1)$$

lässt sich der Operator H^0 des Kohn-Sham-Systems durch „Hochdrehen“ von $\alpha \in [0, 1]$ stetig in den Operator $H^1 = H^\alpha|_{\alpha=1}$ des Hartree-Fock-Systems² überführen, wobei Letzterer hier allerdings noch um den Korrelationsanteil der Kohn-Sham-Gleichung ergänzt ist. Man beachte, dass der obere, stetige Index α sich hier auf die Stärke der Störung bezieht und der untere Index für den Kohn-Sham-Operator bzw. die Störung steht. Weiter wird nun gefordert, dass die Ladungsdichte für alle $\alpha \in [0, 1]$ in 1. Näherung konstant bleibt. (Was genau hier mit „1. Näherung“ gemeint ist, wird unten mit Gleichung (2.4) klar.) Damit etabliert (2.1) eine „adiabatische Verbindung“ ähnlich zu der im Störungsansatz (1.79) von Göring und Levy. Dort allerdings war H^α im N -Teilchen-Hilbertraum definiert, wohingegen die Operatoren in (2.1) auf die Ein-Teilchen-Wellenfunktionen ψ_i wirken. Der Ansatz (2.1) ist daher bereits konzeptuell viel näher an der numerischen Praxis.

Um die Forderung nach konstanter Dichte formulieren zu können, werden zunächst die Orbitale des durch αH_1 gestörten Systems benötigt. Sie sind nach der Rayleigh-Schrödingerschen Störungstheorie in 1. Ordnung durch $\psi_i^0 + \alpha \psi_i^1$ gegeben, wobei die ψ_i^0 die Eigenfunktionen des ungestörten Problems (1.52) sind, und sich die Korrekturfunktionen ψ_i^1 aus der so genannten *Sternheimer-Gleichung*

$$(H_0 - \varepsilon_i^0)\psi_i^1 + (H_1 - \varepsilon_i^1)\psi_i^0 = 0 \quad (2.2)$$

ergeben [86]. Die Energiekorrekturen sind dabei durch $\varepsilon_i^1 = \langle \psi_i^0 | H_1 | \psi_i^0 \rangle$ gegeben. Wie in Kapitel 4 noch näher erläutert, beruhen die Vorteile des variationellen EXX-Verfahrens gegenüber dem Standard-EXX-Verfahren entscheidend auf den effizienten Lösungsmöglichkeiten für diese Gleichung.

²Genau genommen handelt es sich bei H^1 nur *formal* um den Operator des Hartree-Fock-Systems, da $V_{\text{Fock}}[\{\psi_i\}]$ aus den Kohn-Sham-Orbitalen aufgebaut ist. Dies geschieht in völliger Analogie zur Definition der Kohn-Sham-Austauschenergie (1.69) mit Hilfe des Ausdrucks der Hartree-Fock-Austauschenergie.

Mit den Korrekturen ψ_i^1 findet man nun die zu H^α gehörende Dichte

$$\begin{aligned} \rho_\alpha &= \sum_i^{\text{occ}} |\psi_i^0 + \alpha \psi_i^1|^2 \\ &= \underbrace{\sum_i^{\text{occ}} |\psi_i^0|^2}_{=\rho_0} + \alpha \cdot \underbrace{\sum_i^{\text{occ}} \left((\psi_i^0)^* \psi_i^1 + (\psi_i^1)^* \psi_i^0 \right)}{=: \Delta\rho} + \text{Ausdruck der Ordnung 2 in } \alpha. \end{aligned} \quad (2.3)$$

In linearer Ordnung, d.h., für $\alpha \rightarrow 0$, ist die Dichteänderung also durch $\alpha \Delta\rho$ gegeben. Da $\alpha \Delta\rho$ die (lineare) Ladungsverschiebung beschreibt, die durch den Einfluss der Störung αH_1 hervorgerufen wird, sei diese Größe hier, Ref. 78 folgend, als *induzierte Ladungsdichte* bezeichnet. Ist die Dichte konstant, folgt für die 1. Ordnung

$$\Delta\rho := \sum_i^{\text{occ}} \left((\psi_i^0)^* \psi_i^1 + (\psi_i^1)^* \psi_i^0 \right) = 0 \quad (2.4)$$

oder, äquivalent dazu,

$$\|\Delta\rho[v_x]\| = 0. \quad (2.5)$$

In der letzten Beziehung wurde $\Delta\rho$ gleich als Funktional des Austauschpotentials v_x geschrieben. Dies ist möglich, da die induzierte Ladungsdichte nach (2.4) über die direkt von v_x abhängenden ψ_i^0 und die durch die Störung $H_1[\{\psi_i^0\}, v_x]$ zumindest indirekt von v_x abhängenden Korrekturen ψ_i^1 gegeben ist. Man beachte, dass (2.5) nur für das exakte v_x gilt. Für ein nur angenähertes Probe-Potential v_x^{Probe} ist (2.5) hingegen niemals³ erfüllt, so dass sich daraus insgesamt ein *Variationsprinzip* für das exakte Austauschpotential ergibt:

$$\|\Delta\rho[v_x^{\text{Probe}}]\| \geq 0, \quad (2.6)$$

wobei die Gleichheit nur im exakten Fall gilt. Für eine effiziente numerische Minimierung der linken Seite sei auf Abschnitt 4.3 verwiesen.

In das Variationsprinzip gehen an keiner Stelle unbesetzte Zustände ein, und die lineare Antwortfunktion χ_0 wird ebenfalls nirgends gebraucht. Es ist daher plausibel, dass dieser Ansatz bei der Entwicklung von Computeralgorithmen die Möglichkeit großer technischer Vorteile gegenüber dem Standard-EXX-Verfahren bietet. Eine tiefere Einsicht, auf welche Weise in diesem Ansatz die Schwierigkeiten des Standard-Verfahrens umgangen werden können, wird im übernächsten Abschnitt 2.3 zu geben versucht.

Vergleicht man (2.4) mit (1.102) aus Abschnitt 1.6.3, fällt auf, dass beide Gleichungen eine ähnliche Struktur besitzen. Tatsächlich sind sie *identisch*. Darüber hinaus stellen sie eine besonders knappe Formulierung der OEP-Gleichung dar, was im Folgenden gezeigt werden soll.

³In diesem Kapitel werden v -darstellbare Dichten vorausgesetzt. Folglich wäre das Hohenberg-Kohn-Theorem I verletzt, falls es für eine gegebene Dichte ρ ein zweites v'_x gäbe, das die Bedingung (2.5) ebenfalls erfüllt.

2.2 Äquivalenz des variationellen Ansatzes zu OEP und Standard-EXX

Im Ansatz (1.79) von Görling und Levy wurde das gesamte lokale Austausch-Korrelations-Potential v_{xc} des Kohn-Sham-Schemas in den Austausch-Korrelations-Operator $V_{ee} - v_H[\rho]$ des *realen* Systems überführt. Die Konstanz der Dichte folgt dabei letztlich aus der Definition des Kohn-Sham-Systems, das ja gerade so konstruiert ist, dass es die reale Dichte repliziert (vgl. Abschnitt 1.4.2).

Im Gegensatz dazu wird im vorangehend beschriebenen Ansatz nur das lokale Kohn-Sham-Austauschpotential in ein nicht-lokales von der Form des Fock-Operators überführt. Nun liegen jedoch den Herleitungen des Fock-Operators und des Kohn-Sham-Austauschpotentials gänzlich verschiedene Ideen zu Grunde; einmal wird die Austauschenergie E_x nach der Ein-Teilchen-Funktion ψ_i abgeleitet (vgl. Abschnitt 1.3), im anderen Fall jedoch nach der Dichte ρ (vgl. Abschnitt 1.6.1). Die Einführung von $V_{\text{Fock}}[\{\psi_i\}]$ in das Kohn-Sham-Schema gemäß (2.1) ist daher zunächst ein heuristischer Ansatz, und es ist nötig, die Äquivalenz des Variationsprinzips (2.6) mit der OEP-Gleichung (1.97) noch einmal streng zu beweisen.

Das Variationsprinzip (2.6) ist genau dann erfüllt, wenn man ein v_x^{Probe} gefunden hat, das, in (2.1) eingesetzt, auf Korrekturfunktionen ψ_i^1 führt, die die induzierte Ladungsdichte $\Delta\rho$ verschwinden lassen.⁴ Gemäß der Störungstheorie lassen sich die Lösungen der Sternheimer-Gleichung (2.2) auch als Reihenentwicklung mit Hilfe der Greenschen Funktion (B.11) darstellen,

$$\psi_i^1(\mathbf{r}) = \sum_{\substack{j \\ \varepsilon_j \neq \varepsilon_i}} \psi_j^0(\mathbf{r}) \frac{\langle \psi_j^0 | H_1 | \psi_i^0 \rangle}{\varepsilon_i^0 - \varepsilon_j^0} = \int G_i(\mathbf{r}, \mathbf{r}') H_1 \psi_i^0(\mathbf{r}') d^3r', \quad (2.7)$$

wobei H_1 in (2.1) definiert wurde und über *alle* Eigenfunktionen ψ_j^0 des Kohn-Sham-Operators von (1.52) summiert wird mit Ausnahme derjenigen aus dem Eigenraum zu ε_i^0 . Setzt man dies in Gleichung (2.4) ein,

$$\begin{aligned} \Delta\rho(\mathbf{r}) &= \sum_i^{\text{occ}} \sum_{\substack{j \\ \varepsilon_j \neq \varepsilon_i}} \psi_i^0(\mathbf{r})^* \psi_j^0(\mathbf{r}) \frac{\langle \psi_j^0 | H_1 | \psi_i^0 \rangle}{\varepsilon_i^0 - \varepsilon_j^0} + \text{c.c.} \\ &= \sum_i^{\text{occ}} \int d^3r' \left[\psi_i^0(\mathbf{r})^* G_i(\mathbf{r}, \mathbf{r}') H_1 \psi_i^0(\mathbf{r}') + \text{c.c.} \right] = 0, \end{aligned} \quad (2.8)$$

und beachtet ferner, dass Summanden mit besetzten Zuständen ψ_j^0 sich herauskürzen (Vertauschen von i und j liefert das negative Vorzeichen), ergibt sich daraus unmittelbar die OEP-Gleichung in der Form (1.103) bzw. (1.97). Damit ist auch die Äquivalenz zur EXX-Gleichung (1.90) gezeigt, wobei aber die in Kapitel 2 besprochenen Schwierigkeiten bei der Umstellung der Formel berücksichtigt werden müssen.

Perdew und Kümmel [79, 80] verfolgten bei der Herleitung des variationellen EXX-Verfahrens gewissermaßen den umgekehrten Weg: Sie starteten von der OEP-Gleichung

⁴Eigentlich ist die induzierte Ladungsdichte in erster Näherung durch $\alpha \Delta\rho$ gegeben. Da aber mit (2.4) nur das (punktweise) Verschwinden des Koeffizienten $\Delta\rho$ entscheidend ist, sei im Folgenden stets $\alpha = 1$ gesetzt.

in der kompakten Gestalt (2.5), die bereits durch Krieger, Li und Iafrate [72] sowie von Grabo *et al.* [48, 49] angegeben wurde, und konstruierten daraus — ebenfalls über die Sternheimer-Gleichung (2.2) — das Störungsproblem (2.1).

2.3 Variation an Stelle direkter Inversion

Wenn das variationelle Verfahren ohne die unbesetzten Zustände und ohne χ_0 und dessen Inversion auskommt, müssen deren „Aufgaben“ an anderer Stelle übernommen werden. Dies soll hier kurz diskutiert werden.

Zunächst zu den unbesetzten Zuständen: Bei der Berechnung von $v_x[\rho]$ mit dem Standard-Verfahren erscheinen sie gemäß (1.94) an zwei Stellen: in der Greenschen Funktion (B.12) und zur Berechnung der linearen Antwort χ_0 , wobei nach (F.2) bei Letzterer die unbesetzten Zustände ebenfalls über die Greensche Funktion ins Spiel kommen. Die Greensche Funktion dient dabei der Beschreibung einer Störung des Kohn-Sham-Systems, und könnte — wie (2.7) zeigt — auch im variationellen Verfahren zur Berechnung der linearen Korrekturen ψ_i^1 benutzt werden. Die Berechnung der ψ_i^1 gelingt jedoch viel effizienter über die Sternheimer-Gleichung (2.2). Die Idee dabei ist sehr ähnlich zu der, die (direkte) Inversion von χ_0 zu vermeiden und sei deshalb weiter unten damit gemeinsam diskutiert.

Zur Eliminierung von χ_0 : Die induzierte Ladungsdichte $\alpha \Delta\rho$ kann als *Verschiebung* der Ladungsverteilung weder Ladung hinzu-, noch abführen. Mathematisch wird dies durch die Forderung der Orthogonalität $\psi_i^0 \perp \psi_i^1$ in der Sternheimer-Gleichung erreicht, vgl. Abschnitt 4.1.1. Es gilt also

$$\int \Delta\rho(\mathbf{r}) d^3r = \sum_i^{\text{occ}} \int \psi_i^0(\mathbf{r})^* \psi_i^1(\mathbf{r}) d^3r + \text{c.c.} = 0. \quad (2.9)$$

Wie von Göring und Levy in Ref. 41 bzw. hier im Anschluss an Gleichung (1.93) diskutiert, charakterisiert diese Bedingung gerade den Wertebereich der linearen Antwort χ_0 . Dies ist der Untervektorraum (der Ladungsdichteänderungen), in dem χ_0 invertierbar ist, also die EXX-Gleichung (1.90) definiert ist. Ferner geht die induzierte Ladungsdichte $\alpha \Delta\rho$ für „infinitesimal kleine“ α formal in den „Zähler“ von $\chi_0 = \delta\rho/\delta v_{\text{eff}}$ über:

$$\begin{aligned} \alpha \Delta\rho &= \sum_i^{\text{occ}} (\psi_i^0)^* (\alpha \psi_i^1) + \text{c.c.} \\ &\longrightarrow \sum_i^{\text{occ}} (\psi_i^0)^* \delta\psi_i^0 + \text{c.c.} = \delta \sum_i^{\text{occ}} |\psi_i^0|^2 = \delta\rho \quad (\alpha \longrightarrow 0). \end{aligned} \quad (2.10)$$

Die in Abschnitt 2.1 verlangte Forderung, dass die Ladungsdichte in linearer Näherung konstant bleibt, wenn man das Potential um $\Delta v_{\text{eff}} := \alpha H_1$ ändert, könnte man nun sehr einfach formulieren, wenn der Störoperator lokal wäre:

$$\int \chi_0[v_{\text{eff}}](\mathbf{r}, \mathbf{r}') \Delta v_{\text{eff}}(\mathbf{r}') d^3r' = \int \frac{\delta\rho[v_{\text{eff}}](\mathbf{r})}{\delta v_{\text{eff}}(\mathbf{r}')} \Delta v_{\text{eff}}(\mathbf{r}') d^3r' \stackrel{!}{=} 0. \quad (2.11)$$

So ist z.B. die linke Seite der OEP-Gleichung (1.96) von dieser Form, wenn man $\Delta v_{\text{eff}} := v_x$ setzt. Das Problem ist nun aber, dass der Störoperator H_1 von (2.1) nicht-lokal ist, da er den Fock-Operator enthält. Damit ist χ_0 allerdings „überfordert“. Wie in Anhang D gezeigt, lässt sich die lineare Antwort jedoch auch so erweitern, dass sie auf diese Form

nicht-lokaler Potentialdifferenzen anwendbar ist, und die Idee, lineare Dichteänderungen gleich Null zu setzen, kann auf ähnliche Weise wie in obiger Gleichung (2.11) formuliert werden. Dort wird auch auf den engen Bezug dieser Interpretation zur Sham-Schlüter-Gleichung und damit letztlich auch zur *GW*-Methode [150, 151] verwiesen. Die induzierte Ladungsdichte (2.4) kann demnach als Ergebnis einer Anwendung der linearen Antwort auf die Potentialänderung aufgefasst werden; eine explizite Berechnung von χ_0 wird nicht mehr benötigt.

Durch welchen Mechanismus wird im variationellen Verfahren nun aber die im Standard-Formalismus benötigte *Inversion* von χ_0 ersetzt? Dies geschieht auf ganz ähnliche Weise, wie sich die Greensche Funktion mit Hilfe der Sternheimer-Gleichung umgehen lässt. Der Greensche Operator ist gemäß (B.8) in einem geeigneten Unterraum die Umkehrung des Operators $H_0 - \varepsilon_i^0$, mit deren Hilfe die Sternheimer-Gleichung in (2.7) explizit nach ψ_i^1 aufgelöst werden konnte. Analog wird auch χ_0^{-1} in (1.94) nur zur *expliziten* Auflösung nach dem Austauschpotential v_x benötigt. In den *impliziten* Darstellungen von v_x und ψ_i^1 mit Hilfe der OEP-Gleichung in Gestalt von (1.96) oder (2.4) bzw. über die Sternheimer-Gleichung (2.2) werden nur die Operatoren χ_0 bzw. $H_0 - \varepsilon_i^0$ selbst, nicht aber deren Umkehrungen benötigt. Solche impliziten Gleichungen lassen sich aber als Variationsprinzip für die gesuchte Variable verstehen: sie muss „nur“ so lange variiert werden, bis die Gleichung erfüllt ist. Es erweist sich dabei häufig als bequem, das Variationsprinzip als Extremalprinzip für eine skalare Größe zu formulieren, die dann zum Funktional der Variablen wird. Dies lässt sich stets — wie in (2.6) für die OEP-Gleichung — durch Betragsbildung erreichen. Es gibt aber auch andere Möglichkeiten, solche Funktionale zu konstruieren, wie z.B. in Abschnitt 4.1.1 für den Fall der Sternheimer-Gleichung ausgeführt. Wie die Minimierung dieser Funktionale im Einzelnen gelingt, wird in den Abschnitten 4.1.2 und 4.3 ausführlich diskutiert.

Noch einmal kurz: Die Vorteile der variationellen Formulierung des EXX-Verfahrens beruhen letztlich auf dem Trick, komplizierte Operator-Inversionen mit Hilfe impliziter Gleichungen in numerisch einfachere Extremalwertprobleme umzuformulieren. Dieser Kunstgriff lässt sich dabei gleich zweimal anwenden, sowohl zur Vermeidung des Greenschen Operators $\hat{G}_i = (H_0 - \varepsilon_i^0)^{-1}$, als auch zur Vermeidung von χ_0^{-1} (beide in geeigneten Unterräumen definiert).

2.4 Variationelles EXX und Hartree-Fock-Verfahren

Gleichung (2.1) stellt einen engen Bezug zwischen dem “exchange-only” ($v_c := 0$) Kohn-Sham-Formalismus (H^α mit $\alpha = 0$) und dem Hartree-Fock-Schema (H^α mit $\alpha = 1$) her. Dennoch lassen sich die Ergebnisse einer Kohn-Sham-EXX-Rechnung *nicht* in die Ergebnisse einer Hartree-Fock-Rechnung überführen — auch nicht in 1. Näherung —, indem man einfach die durch die Störung H_1 hervorgerufenen Korrekturen ψ_i^1 berechnet und auf die Kohn-Sham-Funktionen addiert. Der Grund ist, dass der nicht-lokale Austauschoperator $V_{\text{Fock}}[\{\psi_i\}]$ hier aus den Kohn-Sham-, nicht aber aus den selbst-konsistenten Hartree-Fock-Orbitalen aufgebaut ist.

Dennoch macht Ansatz (2.1) die große Übereinstimmung der Ladungsdichte, der Gesamtenergie [49, 60, 61, 62, 68, 78] und sogar der Orbitale [58, 59] im Hartree-Fock- und dem “exchange-only” Kohn-Sham-Schema plausibel.⁵ Denn in (2.1) wird mit Hilfe der

⁵Es sei allerdings an dieser Stelle auf die großen qualitativen Unterschiede der beiden Methoden hinsichtlich der Ein-Teilchen-Energien hingewiesen: Wie in der Einleitung bereits ausgeführt, steht der sehr guten Übereinstimmung der theoretischen Bandlücken in EXX mit den experimentellen Werten eine

Minimierung (2.6) das lokale Potential v_x gewissermaßen so „gefittet“ oder „gemittelt“, dass es sich optimal an V_{Fock} anpasst. Anders ausgedrückt, das EXX-Potential ist die *beste lokale Näherung* für V_{Fock} . Dies war — wenngleich in anderer Formulierung — ja auch der Leitgedanke beim OEP-Ansatz (1.99); es bestätigt sich hier also wiederum die Äquivalenz von EXX und OEP. Nimmt man weiter an, dass v_{EXX} eine gute Näherung für V_{Fock} ist, so sollte man erwarten, dass sich die Orbitale und die (aus ihnen aufgebauten) Gesamtenergien in beiden Formalismen nur sehr wenig voneinander unterscheiden. Dass sie gut ist, bestätigen die oben genannten Referenzen. Vollständig kann der Unterschied zwischen Hartree-Fock- und Kohn-Sham-Orbitalen jedoch im Allgemeinen nicht verschwinden, da H_1 als Differenz zwischen einem lokalen und nicht-lokalen Operator im Allgemeinen nicht 0 werden kann — außer in einfachsten Fällen wie dem Wasserstoffatom, für das auch das Fock-Potential als lokaler Operator darstellbar ist (und sich gerade gegen das Hartree-Potential weghebt). Es sei noch bemerkt, dass die Vernachlässigung der Orbitaldifferenzen zum so genannten *lokalisierten Hartree-Fock-Verfahren* [87] führt, das als Verallgemeinerung des KLI-Verfahrens [70, 71, 72, 73] angesehen werden kann und gegenüber Letzterem den Vorteil besitzt, wie der Fock-Operator invariant bzgl. unitärer Transformationen der Wellenfunktionen zu sein.

Überschätzung der Bandlücken bei Hartree-Fock um typischerweise einen Faktor 2–3 gegenüber.

Kapitel 3

Methoden zur numerischen Beschreibung periodischer Systeme

Zentrale Größen bei der mathematischen Beschreibung quantenmechanischer Systeme sind in Gestalt linearer Operatoren und Funktionen gegeben, die als Elemente eines geeigneten Vektorraums, z.B. eines Hilbertraums, bzw. als lineare Abbildungen darin definiert sind. Zur konkreten Darstellung dieser Objekte benötigt man Koordinatensysteme, auch Basissysteme genannt. Ihre Wahl ist beliebig (solange sie „vollständig“ sind) und kann daher nach Zweckmäßigkeit erfolgen. Da dieser Vektorraum jedoch in der Regel ∞ -dimensional ist, besteht eine erste numerische Vereinfachung¹ darin, nicht mehr die vollständige Basis zu betrachten, sondern nur noch eine endliche, von der gewünschten Genauigkeit abhängende Anzahl n von Basisvektoren („Diskretisierung“). Damit wird das Problem überhaupt erst mit dem Rechner handhabbar; Operatoren werden zu $(n \times n)$ -Matrizen und Funktionen zu n -dimensionalen Vektoren.

Je nach Zweckmäßigkeit verwendet man in der Praxis verschiedene Typen von Basisfunktionen. Halbleiterkristalle, wie sie in dieser Arbeit behandelt werden, lassen sich als (näherungsweise) periodische Systeme besonders gut mit ebenen Wellen beschreiben (Abschnitt 3.3). Effektiv ist dieser Ansatz, wenn er mit dem durch die Theorie der chemischen Bindung motivierten Pseudopotentialkonzept kombiniert wird (Abschnitte 3.4 und 3.5). Weitere Reduktionen des numerischen Aufwandes ergeben sich durch Ausnutzung vorhandener Symmetrien; statt den (im Idealfall unendlich ausgedehnten) Kristall zu betrachten, beschränkt man sich de facto auf eine kleine „Einheitszelle“ mit nur wenigen Kernen/Ionen (Abschnitte 3.1 und 3.2).

Anders dagegen bei Systemen, in denen etwa mehrere Tausend Atome explizit zu berücksichtigen sind, z.B. Quantenpunkte: Hier sind ebene Wellen im Allgemeinen nicht mehr praktikabel, und ist es besser, lokalisierte *Tight-Binding*-Basen zu verwenden. In diesem Zusammenhang sei auf Teil II dieser Arbeit verwiesen, in dem eine spezielle *Tight-Binding*-Basis, konstruiert aus „Verallgemeinerten Wannier-Funktionen“, vorgestellt wird.

¹Die in diesem Kapitel besprochenen Näherungen sind jedoch nicht die einzigen: durch die Einführung der Born-Oppenheimer-Näherung, der Verwendung genäherter Austausch-Korrelationspotentiale wie LDA oder GGA sowie dem Abbrechen der Görling-Levy-Entwicklung nach dem ersten Glied („exchange-only“) u.Ä. wurden bereits im Vorangehenden erhebliche Vereinfachungen vorgenommen.

3.1 Symmetrien und periodische Randbedingungen

Ein idealer Kristall besteht aus einer periodischen Anordnung von Kernen oder — bei Verwendung von Pseudopotentialen, vgl. Abschnitt 3.5 — Ionen, in deren Potential sich eine unendliche Anzahl von Elektronen bewegt. Je nach Struktur gibt es Sätze linear unabhängiger Vektoren $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3 \in \mathbb{R}^3$, deren ganzzahlige Linearkombinationen \mathbf{R} den Kristall in sich selbst überführen (Translationssymmetrie):

$$\mathbf{R} = n_1 \mathbf{a}_1 + n_2 \mathbf{a}_2 + n_3 \mathbf{a}_3 \quad (\text{mit } n_1, n_2, n_3 \in \mathbb{Z}). \quad (3.1)$$

Das durch die \mathbf{a}_i aufgespannte Parallelepiped wird *Einheitszelle* genannt und besitzt das Volumen $\Omega = |\mathbf{a}_1 \cdot (\mathbf{a}_2 \times \mathbf{a}_3)|$. Eine Wahl der \mathbf{a}_i , die auf das kleinste Volumen Ω führt, das periodisch fortgeführt, den ganzen Kristall ergibt, definiert die *primitive Zelle*. Tatsächlich lassen sich durch einen solchen Ansatz nicht nur Volumensysteme, sondern auch Oberflächen, Cluster, Moleküle und isolierte Atome beschreiben. Dabei setzt man diese Strukturen periodisch fort, wählt aber die Abstände zwischen den Einzelsystemen so groß, dass die Überlappung ihrer Wellenfunktionen so gering wie möglich gehalten wird. Da die Einheitszelle im Rechner natürlich nicht beliebig gedehnt werden kann, ist hier ein sinnvoller Kompromiss zu finden.

Analog zu den Gittervektoren (3.1) definiert man ein *reziprokes Gitter* durch die Vektoren

$$\mathbf{G} = n_1 \mathbf{b}_1 + n_2 \mathbf{b}_2 + n_3 \mathbf{b}_3 \quad (\text{mit } n_1, n_2, n_3 \in \mathbb{Z}), \quad (3.2)$$

wobei die \mathbf{b}_i eine zu den \mathbf{a}_i reziproke Basis des \mathbb{R}^3 sind:

$$\mathbf{b}_i := 2\pi \frac{\mathbf{a}_j \times \mathbf{a}_k}{\Omega} \quad (\text{für } i, j, k \text{ zyklisch}). \quad (3.3)$$

Der durch die \mathbf{b}_i aufgespannte Bereich mit dem Volumen $8\pi^3/\Omega$ heißt *reduzierter Bereich*. Häufig wird er jedoch durch die volumensgleiche *1. Brillouin-Zone* (BZ), einer Wigner-Seitz-Zelle im reziproken Raum, ersetzt. Die Konstruktion von Wigner-Seitz-Zellen wird in vielen Standardwerken der Theoretischen Festkörperphysik, z.B. in Ref. 88 oder 89, beschrieben.

Im Einteilchen-Hamilton-Operator $H(\mathbf{r}) = -\frac{1}{2}\nabla^2 + V(\mathbf{r})$ kommt die Translationssymmetrie des Systems durch ein translationsinvariantes Potential $V(\mathbf{r}) = V(\mathbf{r} + \mathbf{R})$ zum Ausdruck. Als Eigenfunktionen $\psi_{n\mathbf{k}}$ ergeben sich damit so genannte *Blochfunktionen*, die in eine ebene Welle und eine gitterperiodische Modulationsfunktion $u_{n\mathbf{k}}$ separieren:

$$\psi_{n\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}} u_{n\mathbf{k}}(\mathbf{r}) \quad \text{mit} \quad u_{n\mathbf{k}}(\mathbf{r}) = u_{n\mathbf{k}}(\mathbf{r} + \mathbf{R}). \quad (3.4)$$

Dabei kann der „Quasi-Wellenvektor“ \mathbf{k} stets als Element der 1. Brillouin-Zone gewählt werden, denn es gilt $\psi_{n,\mathbf{k}} = \psi_{n,\mathbf{k}+\mathbf{G}}$ für beliebige reziproke Gittervektoren aus (3.2). Die Dispersion, d.h., die \mathbf{k} -Abhängigkeit der zugehörigen Eigenwerte $\varepsilon_{n\mathbf{k}}$ lässt sich in Form einer *Bandstruktur* darstellen, in der die Energiebänder durch den Bandindex n abgezählt werden. Blochfunktionen genügen der *Blochbedingung*

$$\psi_{n\mathbf{k}}(\mathbf{r} + \mathbf{R}) = e^{i\mathbf{k}\cdot\mathbf{R}} \psi_{n\mathbf{k}}(\mathbf{r}), \quad (3.5)$$

d.h., sie sind im Allgemeinen nicht gitterperiodisch (nur dann, wenn \mathbf{k} ein reziproker Gittervektor gemäß (3.2) ist).

Neben der Translationssymmetrie weist das System in der Regel eine Anzahl N_{PS} von Punktsymmetrien auf, d.h., für das Potential des Hamilton-Operators gilt $V(\mathbf{r}) = V(S_\nu \mathbf{r})$

mit $\nu = 1, \dots, N_{\text{PS}}$. Die S_ν sind dabei durch unitäre (3×3) -Matrizen gegeben, die eine Gruppe, die *Punktgruppe* des Systems, bilden. Aus der Blochbedingung erhält man

$$\psi_{n,S_\nu \mathbf{k}}(\mathbf{r}) = \psi_{n,\mathbf{k}}(S_\nu^{-1} \mathbf{r}), \quad \varepsilon_{n,S_\nu \mathbf{k}} = \varepsilon_{n,\mathbf{k}}, \quad (3.6)$$

was beispielsweise in Ref. 90 oder 91 vorgeführt wird. Unabhängig von der speziellen Punktgruppensymmetrie des Systems gilt im reziproken Raum stets die so genannte *Zeitumkehrsymmetrie*. Sie ergibt sich aus der konjugiert-komplexen Schrödinger-Gleichung und der Blochbedingung (vgl. z.B. Ref. 92):

$$\psi_{n,-\mathbf{k}}(\mathbf{r}) = e^{i\alpha_{n,\mathbf{k}}} \psi_{n,\mathbf{k}}^*(\mathbf{r}), \quad \varepsilon_{n,-\mathbf{k}} = \varepsilon_{n,\mathbf{k}}, \quad (3.7)$$

wobei $\alpha_{n,\mathbf{k}}$ eine beliebige Phase ist.

Zu der praktischen Unmöglichkeit, alle unendlich vielen Elektronen des idealen Kristalls numerisch zu berücksichtigen, tritt das konzeptuelle Problem, dass Blochfunktionen auf dem \mathbb{R}^3 nicht normierbar und damit zur quantenmechanischen Beschreibung zunächst nicht geeignet sind. Beide Schwierigkeiten lassen durch Verwendung *periodischer Randbedingungen*, auch *Born-von-Kármán-Randbedingungen* genannt, umgehen. Die Näherung besteht hier darin, nur einen Teil des periodischen Systems zu betrachten, der durch ein ganzzahliges Vielfaches der Einheitszelle, d.h. also, durch ein Parallelepipet mit den Kanten $L_i \mathbf{a}_i$ ($L_i \in \mathbb{N}$ geeignet) gegeben ist. Die Blochfunktionen werden dann nur auf diesem *periodischen Grundgebiet* normiert, und man nimmt weiter an, dass sie mit diesem Gebiet periodisch sind:

$$\psi_{n\mathbf{k}}(\mathbf{r} + L_i \mathbf{a}_i) = \psi_{n\mathbf{k}}(\mathbf{r}). \quad (3.8)$$

Diese Annahme führt zu einer Diskretisierung des reziproken Raums: Der Quasi-Wellenvektor \mathbf{k} kann nur noch die Werte

$$\mathbf{k} = \frac{m_1}{L_1} \mathbf{b}_1 + \frac{m_2}{L_2} \mathbf{b}_2 + \frac{m_3}{L_3} \mathbf{b}_3 \quad (\text{mit } m_1, m_2, m_3 \in \mathbb{Z}) \quad (3.9)$$

annehmen, wobei der reduzierte Bereich durch $0 \leq m_i < L_i$ abgetastet wird. Für genügend große L_i kann der Abstand benachbarter \mathbf{k} jedoch beliebig klein gemacht werden, daher bezeichnet man diese äquidistante Verteilung der \mathbf{k} -Punkte auch als „quasikontinuierlich“. Man beachte, dass die Anzahl $L_1 L_2 L_3$ der \mathbf{k} -Punkte gerade der Anzahl der Einheitszellen im periodischen Grundgebiet entspricht. Etwas plakativ kann man also formulieren: der Ortsraum-Variablen \mathbf{R} entspricht im reziproken Raum die Variable \mathbf{k} . Die Beschreibung des Kristalls wird umso genauer, je größer das periodische Grundgebiet ist, d.h., je dichter das durch (3.9) gegebene Gitter gewählt wurde. Im Grenzfall gehen Summen über \mathbf{k} -Punkte daher in Integrale über. Solche Integrale sollen im folgenden Abschnitt 3.2 behandelt werden.

3.2 Numerische Integration in der Brillouin-Zone

Für gitterperiodischen Systeme findet man mit Hilfe der Bloch-Bedingung (3.5), dass die Ladungsdichte und eine Reihe anderer Größen, wie etwa der nicht-lokale Austauschoperator (vgl. Abschnitt 4.4), durch Integrale über die 1. Brillouin-Zone (BZ) gegeben sind. Im Folgenden sei hier exemplarisch die Ladungsdichte betrachtet, die sich dann gemäß (1.53) zu

$$\rho(\mathbf{r}) = \frac{\Omega}{8\pi^3} \sum_n \int_{\text{BZ}} f_{\text{occ}}(n, \mathbf{k}) \cdot |\psi_{n\mathbf{k}}(\mathbf{r})|^2 d^3k \quad (3.10)$$

ergibt. Hier ist

$$f_{\text{occ}}(n, \mathbf{k}) := \frac{1}{1 + \exp\left(\frac{\varepsilon_{n\mathbf{k}} - \varepsilon_{\text{F}}}{k_{\text{B}}T}\right)} \quad (3.11)$$

die Fermi-Funktion mit der Fermi-Energie ε_{F} , der Boltzmann-Konstante k_{B} und der Temperatur T , ferner beinhaltet der Index n dabei die Spinquantenzahl σ . Für die numerische Behandlung muss die Integration durch eine Summe über einen diskreten Satz an \mathbf{k} -Punkten approximiert werden. Dabei wird ausgenutzt, dass sich die Wellenfunktionen dicht beieinander liegender \mathbf{k} -Punkte nur wenig voneinander unterscheiden, und damit ein ganzer Bereich im reziproken Raum durch einen einzigen \mathbf{k} -Punkt repräsentiert werden kann. Wie hoch die Auflösung tatsächlich gewählt werden muss, hängt vom betrachteten System ab und muss durch Konvergenztests herausgefunden werden.

Zur Generierung geeigneter \mathbf{k} -Punkt-Sätze sind verschiedene Methoden entwickelt worden [93, 94, 95, 96, 97]. Eine der gebräuchlichsten ist das auch in dieser Arbeit benutzte Verfahren von Monkhorst und Pack [96], bei dem ein regelmäßiges Gitter durch den reduzierten Bereich (vgl. Abschnitt 3.1) gelegt wird. Unterteilt man den reduzierten Bereich dabei in Richtung der \mathbf{b}_i -Vektoren je L_i mal, spricht man von einem $(L_1 \times L_2 \times L_3)$ -Monkhorst-Pack-Gitter, vgl. Abb 3.1.

Weiter reduzieren lässt sich die Anzahl der benötigten \mathbf{k} -Punkte durch Ausnutzung der Punktsymmetrie des Systems. Punkte, die durch Punkttransformationen ineinander überführt werden können, bilden eine Äquivalenzklasse. Die Symmetrie-Relationen (3.6) und (3.7) erlauben es nun, pro Äquivalenzklasse jeweils nur einen \mathbf{k} -Punkt, gewichtet mit der Mächtigkeit der Klasse, in der Summe zu berücksichtigen und das Ergebnis dafür anschließend zu „symmetrisieren“. Die Dichte (3.10) wird dann annähernd durch

$$\rho(\mathbf{r}) \approx \rho_{\text{sym}}(\mathbf{r}) := \frac{1}{N_{\text{PS}}} \sum_{\nu=1}^{N_{\text{PS}}} \rho_{\text{unsym}}(S_{\nu}\mathbf{r}) \quad (3.12)$$

beschrieben, wobei N_{PS} die Anzahl der Punkttransformationen ist, und ρ_{unsym} als Summe über einen *reduzierten* Satz N_{red} nicht-äquivalenter \mathbf{k} -Punkte gegeben ist:

$$\rho_{\text{unsym}}(\mathbf{r}) := \sum_n \sum_{s=1}^{N_{\text{red}}} w_{\mathbf{k}_s} f_{\text{occ}}(n, \mathbf{k}_s) |\psi_{n\mathbf{k}_s}(\mathbf{r})|^2. \quad (3.13)$$

Die Gewichte $w_{\mathbf{k}_s}$ mit $\sum_{s=1}^{N_{\text{red}}} w_{\mathbf{k}_s} = 1$ sind dabei proportional zur Mächtigkeit der durch die \mathbf{k}_s repräsentierten Klassen.

Es sei darauf hingewiesen, dass die Symmetrisierung (3.12) im Allgemeinen einer Summation über *mehr* \mathbf{k} -Punkte als den $L_1 L_2 L_3$ Punkten des ursprünglichen Monkhorst-Pack-Gitters entspricht. Dies rührt daher, dass das Monkhorst-Pack-Gitter im Allgemeinen nicht invariant gegenüber Punkttransformationen des Bravais-Gitters (und der Zeitinversionssymmetrie) ist, wie Abbildung 3.1 demonstriert. Eine Summation explizit über das gesamte Monkhorst-Pack-Gitter, aber ohne anschließende Symmetrisierung, liefert deshalb nicht notwendigerweise ein symmetrisches Ergebnis.

3.3 Entwicklung nach ebenen Wellen

Die Separierbarkeit der Blochfunktionen (3.4) gestattet es, den Ebene-Wellen-Faktor $e^{i\mathbf{k}\cdot\mathbf{r}}$ aus dem Eigenwertproblem $H\psi_{n\mathbf{k}} = \varepsilon_{n\mathbf{k}}\psi_{n\mathbf{k}}$ herauszukürzen und eine Differentialgleichung nur für die Modulationsfunktion $u_{n\mathbf{k}}$ aufzustellen:

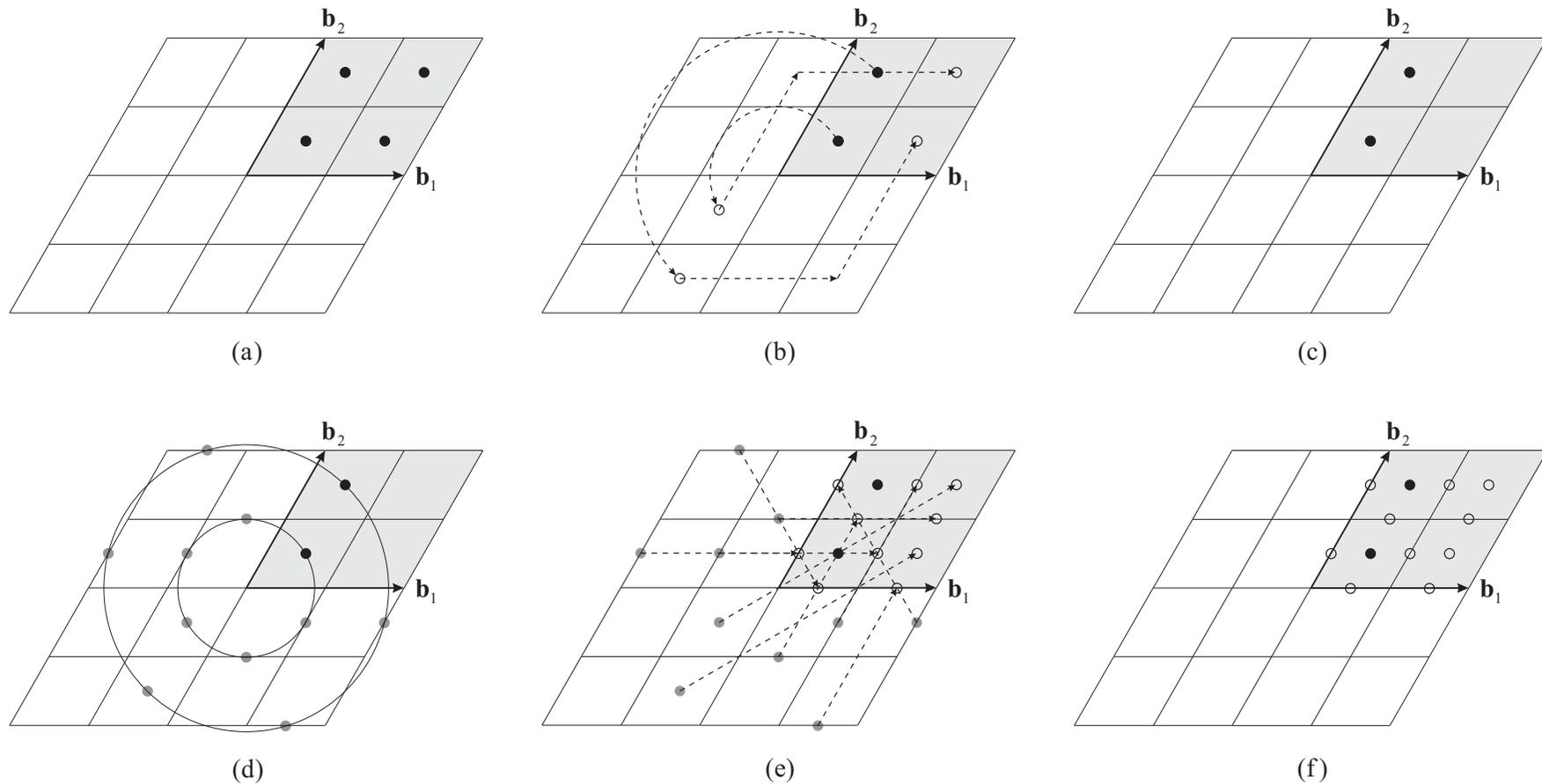


Abbildung 3.1: Symmetrien beim Monkhorst-Pack-Schema am Beispiel eines 2-dimensionalen Gitters mit hexagonaler Struktur. Die reziproken Gittervektoren \mathbf{b}_1 und \mathbf{b}_2 spannen den reduzierten Bereich (grau) auf. (a) \mathbf{k} -Punkt-Satz einer (2×2) -Monkhorst-Pack-Zerlegung im reduzierten Bereich. (b) Der \mathbf{k} -Punkt-Satz kann auf die schwarz gezeichneten Punkte reduziert werden, denn diese lassen sich durch Symmetrietransformationen (hier: Inversion, Verschiebung um reziproke Gittervektoren) in die mit leeren Kreisen markierten \mathbf{k} -Punkte überführen. (c) Der reduzierte \mathbf{k} -Punkt-Satz. (d) Durch Anwenden aller Punktsymmetrien (hier: Drehung um 0° , 60° , 120° , 180° , 240° , 300° und die Inversion) auf den reduzierten \mathbf{k} -Punkt-Satz ergeben sich die so genannten „Sterne“. (e) Translation aller Punkte der Sterne um geeignete reziproke Gittervektoren in den reduzierten Bereich. (f) Dieser \mathbf{k} -Punkt-Satz ist dem in (d) äquivalent. Der Vergleich mit (a) zeigt, dass die ursprüngliche Monkhorst-Pack-Zerlegung hier nicht invariant gegenüber Punkttransformationen des Bravais-Gitters ist. Summiert man jeweils über die markierten Punkte im grauen Feld, erhält man in (a) folglich nicht die volle Symmetrie, in (f) hingegen schon.

$$\left[-\frac{1}{2} \left(\nabla^2 - 2\mathbf{k} \cdot \frac{1}{i} \nabla - \mathbf{k}^2 \right) + V(\mathbf{r}) \right] u_{n\mathbf{k}}(\mathbf{r}) = \varepsilon_{n\mathbf{k}} u_{n\mathbf{k}}(\mathbf{r}). \quad (3.14)$$

Wegen der Gitterperiodizität des Potentials V und der Modulationen $u_{n\mathbf{k}}$ kann diese Gleichung fouriertransformiert werden:

$$\sum_{\mathbf{G}'} \left[+\frac{1}{2} |\mathbf{G} + \mathbf{k}|^2 \delta_{\mathbf{G}, \mathbf{G}'} + \hat{V}(\mathbf{G} - \mathbf{G}') \right] c_{n\mathbf{k}}(\mathbf{G}') = \varepsilon_{n\mathbf{k}} c_{n\mathbf{k}}(\mathbf{G}). \quad (3.15)$$

Die

$$c_{n\mathbf{k}}(\mathbf{G}) = \frac{1}{\sqrt{\Omega}} \int_{\Omega} u_{n\mathbf{k}}(\mathbf{r}) e^{-i\mathbf{G} \cdot \mathbf{r}} d^3r \quad (3.16)$$

sind dabei die Koeffizienten der Fourierentwicklung

$$u_{n\mathbf{k}}(\mathbf{r}) = \frac{1}{\sqrt{\Omega}} \sum_{\mathbf{G}} c_{n\mathbf{k}}(\mathbf{G}) e^{i\mathbf{G} \cdot \mathbf{r}}. \quad (3.17)$$

Offensichtlich wird beim Übergang von der Orts- zur Impulsdarstellung aus dem nicht-lokalen Operator der kinetischen Energie ein lokaler Multiplikationsoperator. Umgekehrt erhält man jedoch aus einem — wie im Falle der Kohn-Sham-Gleichungen (1.52) — lokalen Potential $V(\mathbf{r})$ einen nicht-lokalen Operator. Daher empfiehlt es sich, die Operationen jeweils in der Basis auszuführen, in der sie lokal dargestellt werden können. Dabei kann mit Hilfe schneller Fouriertransformationen (FFT) effizient zwischen den beiden Darstellungen hin- und hergewechselt werden.

Aus der Rücktransformation (3.17) ergibt sich zusammen mit (3.4) auch die Fourierdarstellung der Blochfunktionen:

$$\psi_{n\mathbf{k}}(\mathbf{r}) = \frac{1}{\sqrt{\Omega}} \sum_{\mathbf{G}} c_{n\mathbf{k}}(\mathbf{G}) e^{i(\mathbf{G} + \mathbf{k}) \cdot \mathbf{r}}. \quad (3.18)$$

Die Basis hier ist die so genannte „ $|\mathbf{G} + \mathbf{k}$ “-Basis“ mit $e^{i(\mathbf{G} + \mathbf{k}) \cdot \mathbf{r}} = \langle \mathbf{r} | \mathbf{G} + \mathbf{k} \rangle$.

Wie zu Beginn des Kapitels bereits erwähnt, muss die Menge der Basisfunktionen in der Praxis auf eine endliche Anzahl reduziert werden. Dies geschieht hier durch Einführung einer *Abschneideenergie* oder *cutoff-Energie* E_{cut} . Sie stellt eine Obergrenze für die kinetische Energie der „schnellsten“ Basisfunktionen dar. Für gegebenes \mathbf{k} wird die Summation in (3.17) bzw. (3.18) also auf alle \mathbf{G} mit $\frac{1}{2} |\mathbf{G} + \mathbf{k}|^2 < E_{\text{cut}}$ eingeschränkt. Die Abschneideenergie ist somit ein leicht kontrollierbares Maß für die Qualität (die „Vollständigkeit“) der Basis. Bei praktischen Anwendungen muss mit Hilfe von Konvergenztests sichergestellt sein, dass die Rechnungen bzgl. E_{cut} auskonvergieren sind. Aus dem Variationsprinzip von Ritz (vgl. Abschnitt 1.2) folgt eine Konvergenz von oben gegen die tatsächliche Gesamtenergie.

Mit dem Vorangehenden wird klar, dass die Verwendung einer Ebene-Wellen-Basis umso effizienter ist, je kleiner die benötigte Abschneideenergie ist. Die Anzahl der erforderlichen \mathbf{G} -Vektoren lässt sich dabei durch die Relation $N_{\mathbf{G}} \sim E_{\text{cut}}^{3/2}$ abschätzen. Die Fourierentwicklung stark lokalisierter und/oder heftig oszillierender Wellenfunktionen, wie sie z.B. durch die Singularitäten des äußeren Potentials an den Kernorten hervorgerufen werden, verlangt jedoch auch sehr hohe Frequenzen. Dem kann mit der Einführung so genannter Pseudopotentiale begegnet werden, die die Orbitale innerhalb eines für die

chemische Bindung nicht relevanten Rumpfbereichs durch geglättete Pseudoorbitale ersetzen und damit die erforderliche Abschneideenergie drastisch reduzieren. Andererseits wurden in den vergangenen Jahren auch „gemischte“ Basissätze entwickelt, in denen ebene Wellen mit lokalisierten Basisfunktionen kombiniert werden. Erwähnt seien hier die *linearized augmented plane-waves* (LAPW) [100], die *augmented plane-waves plus local orbitals* (APW+lo) [101, 102] sowie die *projector augmented plane-waves* (PAW) [103]. In dieser Arbeit werden ebene Wellen in Verbindung mit nicht-lokalen, normerhaltenden Pseudopotentialen benutzt, auf deren Konstruktion genauer in Abschnitt 3.5 eingegangen wird. Mit dieser Methode lassen sich die hier untersuchten Halbleitermaterialien (GaAs, Gruppe-III-Nitride) gut beschreiben.

3.4 Näherung der starren Rumpfe

Chemische Bindungen werden fast ausschließlich von den äußeren, so genannten *Valenzelektronen* der beteiligten Atome bestimmt. Ferner ist die Hybridisierung zwischen den Valenzelektronen und den inneren, so genannten Rumpfelektronen häufig vernachlässigbar. Daher ist es in der Regel ausreichend, die inneren Schalen zusammen mit den Kernen als „unveränderliche Ionen“ oder „starre Rumpfe“ (engl.: „*frozen cores*“) aufzufassen, deren Wechselwirkung mit den Valenzelektronen genähert über ein effektives Potential v_{Ion} beschrieben wird.

Von den Ergebnissen einer Allelektronen-Rechnung für ein freies Atom mit Kernladungszahl Z ausgehend, wird dafür die atomare Ladungsdichte $\rho_{\text{AE}} = \rho_{\text{R}} + \rho_{\text{V}}$ in einen Anteil $\rho_{\text{R}} = \sum_{i=1}^{N_{\text{R}}} |\psi_i|^2$ der N_{R} Rumpfelektronen und einen Anteil $\rho_{\text{V}} = \sum_{i=N_{\text{R}}+1}^Z |\psi_i|^2$ der $Z - N_{\text{R}}$ Valenzelektronen aufgeteilt. Von nun an wird die atomare Rumpfladungsdichte ρ_{R} als konstant vorausgesetzt, also nicht mehr zur Variation zugelassen. Damit erhält man den nur von der Valenzladungsdichte ρ_{V} abhängenden atomaren Kohn-Sham-Operator

$$-\frac{1}{2}\nabla^2 + v_{\text{Ion}}(\mathbf{r}) + v_{\text{H}}[\rho_{\text{V}}](\mathbf{r}) + v_{\text{xc}}[\rho_{\text{V}}](\mathbf{r}). \quad (3.19)$$

Alle ρ_{R} -abhängigen Terme sind hier durch das Ionenpotential

$$v_{\text{Ion}}(\mathbf{r}) := -\frac{Z}{|\mathbf{r}|} + v_{\text{H}}[\rho_{\text{R}}](\mathbf{r}) + \Delta v_{\text{xc}}[\rho_{\text{R}}, \rho_{\text{V}}](\mathbf{r}) \quad (3.20)$$

berücksichtigt. Es enthält neben dem Kernpotential den Rumpfanteil des Hartree-Potentials, der auf Grund der Linearität in der Ladungsdichte $v_{\text{H}}[\rho_1 + \rho_2] = v_{\text{H}}[\rho_1] + v_{\text{H}}[\rho_2]$ exakt vom Valenzanteil abgespalten werden kann. Das Austausch-Korrelationspotential hingegen ist nicht linear in der Dichte; daher ist v_{Ion} über den Ausgleichsterm

$$\Delta v_{\text{xc}}[\rho_{\text{R}}, \rho_{\text{V}}] := v_{\text{xc}}[\rho_{\text{R}} + \rho_{\text{V}}] - v_{\text{xc}}[\rho_{\text{V}}] \quad (3.21)$$

streng genommen noch von der Valenzladungsdichte ρ_{V} abhängig. Sie wird aber näherungsweise auch bei der Betrachtung mehratomiger Systeme gleich der atomaren Valenzladungsdichte ρ_{V} gesetzt und festgehalten, so dass (3.20) das Potential eines „starren Rumpfs“ darstellt. Dies entspricht letztlich einer „Linearisierung“ des Austausch-Korrelationspotentials, $v_{\text{xc}}[\rho_{\text{R}}] \approx v_{\text{xc}}[\rho_{\text{R}} + \rho_{\text{V}}] - v_{\text{xc}}[\rho_{\text{V}}]$, die damit gerechtfertigt wird, dass die Überlappung von Rumpf- und Valenzzuständen des Allelektronen-Atoms in der Regel vernachlässigt werden kann. Falls diese Vorgehensweise keine befriedigenden Ergebnisse liefert, lässt sich eine „nicht-lineare Rumpfkorrektur“ [104, 105, 115] durchführen, die im Rahmen der Näherung der starren Rumpfe exakt ist, vgl. Abschnitt 3.5.3.

Um schließlich zu einer Kohn-Sham-Gleichung für M Atome zu gelangen, wird das externe Potential v_{ext} in (1.52) durch eine geeignete Superposition von Ionenpotentialen der Form (3.20) definiert und zur Berechnung der Dichte in (1.53) über alle Valenzelektronen im System summiert.

Die Entscheidung, auf welche Weise die Einteilung der Elektronen in Rumpf- und Valenzbereich erfolgen soll, hängt vom betrachteten Element und der gewünschten Genauigkeit ab. Bei den Elementen der III-V-Verbindungen werden oft die abgeschlossenen inneren Schalen dem Rumpf-, die s - und p -Elektronen der äußeren Schale hingegen dem Valenzbereich zugeordnet, so z.B. bei GaAs oder InAs. In einigen Fällen sind jedoch auch tiefer liegende Zustände an der chemischen Bindung beteiligt: So wechselwirken etwa bei GaN und InN die kationischen d -Zustände der so genannten *semi-core*-Schale (der Schale unterhalb der eigentlichen Valenzschale) mit den energetisch nah liegenden $2s$ -Zuständen des Stickstoffs und müssen daher bei einer genaueren Rechnung explizit als Valenzelektronen berücksichtigt werden. Im Rahmen dieser Arbeit wurde erstmalig die Gültigkeit dieser Vorgehensweise im Zusammenhang mit EXX-Pseudopotentialen untersucht, vgl. dazu Kapitel 7.

3.5 Pseudopotentiale

Durch den im vorangehenden Abschnitt beschriebenen Ansatz der unveränderlichen Ionen lässt sich zwar die Anzahl der effektiv zu berechnenden Ein-Teilchen-Zustände wesentlich reduzieren, dennoch würde auch die Beschreibung der verbleibenden Valenzelektronen noch einen immensen numerischen Aufwand verlangen: In der Nähe der Ionenrümpfe weisen die Valenzorbitale starke Oszillationen auf, die von der Singularität der Ionenpotentiale (3.20) sowie von der Orthogonalität zu den Rumpforbitalen herrühren — große Basissätze wären also von Nöten. Vor allem bei Verwendung ebener Wellen (vgl. Abschnitt 3.3) erweist sich daher die Einführung von „Pseudopotentialen“ als äußerst hilfreiche Weiterführung des Konzepts der starren Rümpfe.

Die zu Grunde liegende Idee ist, dass die Gestalt der Valenzorbitale innerhalb eines durch einen Abschneideradius r_{nl}^{cut} bestimmten Rumpfbereichs um die Kernorte die Eigenschaften einer chemischen Bindung des Atoms kaum beeinflusst. r_{nl}^{cut} soll dabei zumindest so groß gewählt sein, dass sich alle Knoten (Nulldurchgänge) der betrachteten Orbitale darin befinden, wobei im Allgemeinen für verschiedene Hauptquantenzahlen n und Drehimpulse l unterschiedliche Radien zugelassen sind. Gesucht ist nun ein Potential, das die tatsächlichen Wellenfunktionen außerhalb des Rumpfbereichs mit genügender Genauigkeit reproduziert, die numerisch problematischen Oszillationen innerhalb jedoch durch einen möglichst weichen, d.h., durch möglichst wenige ebene Wellen beschreibbaren Verlauf ersetzt. Ein anderer Aspekt ist die *Übertragbarkeit*: das ideale Pseudopotential soll unabhängig von der chemischen Umgebung die richtigen, d.h., mit einer Allelektronen-Rechnung übereinstimmenden Resultate liefern. Während das Pseudopotential mit wachsendem r_{nl}^{cut} immer weicher gewählt werden kann, wird seine Übertragbarkeit geringer. Zwischen diesen beiden Anforderungen muss daher ein sinnvoller Kompromiss gefunden werden. Eng mit der Übertragbarkeit hängt ferner die so genannte *Normerhaltung* zusammen; bei normerhaltenden Pseudopotentialen besitzt das „pseudoisierte“ Atom innerhalb des Abschneideradius die gleiche Ladung wie sein Allelektronen-Gegenstück.

Moderne, normerhaltende *ab-initio* Pseudopotentiale erfüllen eine Reihe weiterer Bedingungen [106, 107, 108, 109], die natürlich nicht alle zugleich exakt erfüllt werden können. In den genannten Referenzen findet sich eine eingehende Diskussion dieser Bedingungen,

die zusammengenommen lauten:

- (i) Die Orbitale des Pseudoatoms besitzen die gleichen Eigenwerte wie die entsprechenden Valenzorbitale des Allelektronen-Atoms.
- (ii) Die Orbitale des Pseudoatoms stimmen außerhalb des Abschneideradius r_{nl}^{cut} mit den entsprechenden Orbitalen des Allelektronen-Atoms überein. Zudem müssen die Pseudoorbitale mindestens zweimal differenzierbar sein und dürfen keine Nulldurchgänge besitzen, damit sich die Schrödinger-Gleichung nach dem Pseudopotential auflösen lässt.
- (iii) Innerhalb des Rumpfbereichs enthalten die Orbitale des Pseudoatoms die gleiche Ladung wie die entsprechenden Allelektronen-Orbitale (Normerhaltung).
- (iv) Die logarithmische Ableitung nach r und wiederum deren Energieableitung stimmt für $r > r_{nl}^{\text{cut}}$ bei den jeweiligen Orbitalen von Pseudo- und Allelektronen-Atom überein. Durch diese Bedingung soll zusammen mit Forderung (i) erreicht werden, dass die Streueigenschaften des Potentials bei der Pseudoisierung in 1. Ordnung erhalten bleiben [108]. Dies ist eine Voraussetzung für die Übertragbarkeit des Pseudopotentials. Da sich die Eigenwerte je nach chemischer Umgebung ändern, ist die Übereinstimmung der Energieableitung als Erweiterung von (i) gedacht.

Streng genommen muss (iv) hier nicht explizit gefordert werden, da dieser Punkt durch die Bedingung (iii) bereits implizit abgedeckt ist, vgl. den folgenden Abschnitt.

3.5.1 Konstruktion

Die Generierung des Pseudopotentials erfolgt iterativ in einem selbstkonsistenten Zyklus. Dabei wird der Radialanteil R_{nl}^{Ps} eines Pseudoorbitals im Rumpfbereich durch eine geeignete Funktion parametrisiert, die genügend Freiheitsgrade sowohl zur Anpassung an die oben genannten Forderungen (i) – (iv) als auch zur Modellierung eines möglichst weichen Pseudopotentials besitzt. Hierfür wurden verschiedene Ansätze entwickelt, vgl. [106, 107, 109, 110] und die darin enthaltenen Literaturverweise. Bei dem Verfahren von Troullier und Martins [109] etwa wird ein gerades Polynom 12. Grades $p(r) = \sum_{i=0}^6 a_{2i} r^{2i}$ zur Beschreibung des Radialteils verwendet, so dass — zusätzlich zum Abschneideradius — sieben Parameter variiert werden können:

$$R_{nl}^{\text{Ps}}(r) := \begin{cases} r^l \exp\{p(r)\} & \text{für } r \leq r_{nl}^{\text{cut}} \\ R_{nl}^{\text{AE}}(r) & \text{für } r \geq r_{nl}^{\text{cut}} \end{cases} \quad (3.22)$$

Diese Funktion kann in die radiale Schrödinger-Gleichung eingesetzt werden, die sich dann wegen Punkt (ii) nach dem *abgeschirmten Pseudopotential*

$$v_{nl}^{\text{Ps, abg}}(r) := \varepsilon_{nl} + \frac{1}{2} \frac{1}{u_{nl}^{\text{Ps}}(r)} \frac{d^2 u_{nl}^{\text{Ps}}(r)}{dr^2} - \frac{l \cdot (l + 1)}{2r^2} \quad (3.23)$$

aauflösen lässt, wobei $u_{nl}^{\text{Ps}}(r) := r \cdot R_{nl}^{\text{Ps}}(r)$ gesetzt wurde. Daraus ergibt sich das *ionische Pseudopotential* durch Abziehen des elektrostatischen (Hartree) Potentials und des Austausch-Korrelationspotentials:

$$v_{nl}^{\text{Ps, ion}}(r) := v_{nl}^{\text{Ps, abg}}(r) - v_{\text{H}}[\rho^{\text{Ps}}](r) - v_{\text{xc}}[\rho^{\text{Ps}}](r) \quad (3.24)$$

Dabei ist ρ^{Ps} die Ladungsdichte des Pseudoatoms. Mit Gleichung (3.24) knüpft das Pseudopotentialkonzept an die in Abschnitt 3.4 beschriebene „Näherung der starren Rumpfe“ an, womit aber auch die dort gemachten Näherungsannahmen, insbesondere die Linearität des Austausch-Korrelationspotentials, übernommen werden (vgl. Abschnitt 3.5.3). Abbildung 3.2 vergleicht schematisch den Verlauf einer Pseudowellenfunktion mit dem der tatsächlichen Wellenfunktion.

Zwischen Norm und logarithmischer Ableitung der Pseudoorbitale besteht folgender Zusammenhang:

$$-\frac{1}{2} \frac{\partial}{\partial \varepsilon} \left[\frac{\partial}{\partial r} \ln R_l(r, \varepsilon) \right] \Bigg|_{\substack{r=r^{\text{cut}} \\ \varepsilon=\varepsilon'}} = \frac{1}{R_l^2(r^{\text{cut}}, \varepsilon')} \int_0^{r^{\text{cut}}} R_l^2(r, \varepsilon') dr', \quad (3.25)$$

wobei $r \cdot R_l(r, \varepsilon)$ die Lösung der radialen Schrödinger-Gleichung des Allelektronen- oder des Pseudoatoms für eine gegebene Energie ε ist (nicht notwendigerweise eine Eigenfunktion). In der Nähe der atomaren Eigenenergien besitzt daher ein normerhaltendes Pseudopotential die gleichen Streueigenschaften (logarithmische Ableitungen) wie das Allelektronen-Potential, was eine wichtige Voraussetzung für die Übertragbarkeit des Pseudopotentials ist.

Auch die in dieser Arbeit verwendeten EXX/OEP-Pseudopotentiale wurden auf diese Weise generiert. Dabei wurde das bereits erwähnte Troullier-Martins-Schema [109] angewandt. Da allerdings im EXX/OEP-Formalismus der Austauschanteil von $v_{\text{xc}}[\rho]$ nicht einfach parametrisch von ρ abhängt, wird bei der Konstruktion der ionischen EXX/OEP-Pseudopotentiale die OEP-Gleichung (1.97) für das Pseudoatom gelöst (vgl. S. 62ff. in Ref. 111).

Das Pseudopotential (3.24) hängt von n und l ab; es gilt daher bei Anwendung auf eine beliebige Funktion nur für jene Komponenten dieser Funktion, die im Raum der atomaren Pseudoorbitale $|nlm\rangle$ mit $m \in \{-l, \dots, +l\}$ liegen. Das gesamte ionische Pseudopotential kann als Entwicklung nach den $|nlm\rangle$ geschrieben werden und ist dann durch die nicht-lokale² Darstellung

$$V^{\text{Ps, ion}} := \sum_{n,l,m} |nlm\rangle v_{nl}^{\text{Ps, ion}}(r) \langle nlm| \quad (3.26)$$

gegeben³, wobei $\langle \mathbf{r} | nlm \rangle = R_{nl}^{\text{Ps}}(r) Y_{lm}(\theta, \phi)$. Um den Rechenaufwand zu verringern, beschränkt man die Summe über n in (3.26) häufig auf die Valenzschale(n). Sollen auch angeregte Zustände sehr genau beschrieben werden, müssen entsprechend höhere n berücksichtigt werden. Weiter reduzieren lässt sich der Rechenaufwand mit der nachfolgend skizzierten Methode von Kleinman und Bylander [113, 114].

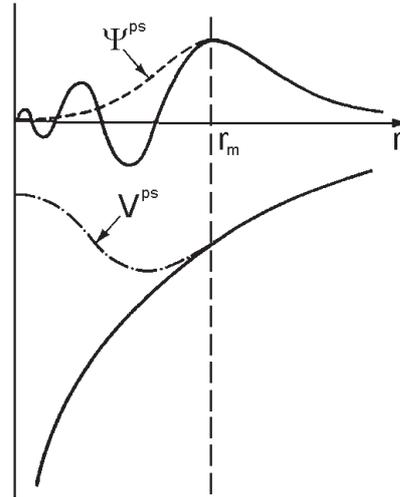


Abbildung 3.2: Schematische Darstellung eines tatsächlichen Atomorbitals (durchgezogenen) und eines Pseudoorbitals (gestrichelt) sowie den zugehörigen Potentialen. r_m bezeichnet den Abschneideradius. (Quelle: Ref. 82)

²Wenn in (3.26) nur über eine Hauptquantenzahl summiert wird, lässt sich das Pseudopotential auch als *semi-lokales* Potential $V^{\text{Ps, ion}} := \sum_{l,m} |Y_{lm}\rangle v_{nl}^{\text{Ps, ion}}(r) \langle Y_{lm}|$ formulieren. „Semi-lokal“ soll hier bedeuten: nicht-lokal im Winkelanteil (Integraloperator), aber lokal im Radialteil (Multiplikationsoperator).

³Präziser formuliert: $V^{\text{Ps, ion}}$ ist ein Integraloperator, dessen Kern im Ortsraum gegeben ist durch $\langle \mathbf{r} | V^{\text{Ps, ion}} | \mathbf{r}' \rangle := \sum_{n,l,m} \langle \mathbf{r} | nlm \rangle v_{nl}^{\text{Ps, ion}}(r) \langle nlm | \mathbf{r}' \rangle$.

3.5.2 Kleinman-Bylander-Darstellung

Soll das Pseudopotential (3.26) in einer Ebene-Wellen-Basis direkt dargestellt werden, ist der Rechenaufwand proportional zu $N_{\mathbf{G}} \cdot (N_{\mathbf{G}} + 1)/2$, wenn $N_{\mathbf{G}}$ die Anzahl der Basisfunktionen bezeichnet [113]. Auf eine mit $N_{\mathbf{G}}$ annähernd lineare Skalierung lässt sich der Aufwand jedoch vermindern, wenn man (3.26) auf eine *separable* Gestalt bringt, d.h., durch eine Linearkombination geeigneter Projektoren $\sum |\xi_{nlm}\rangle f_{nlm} \langle \xi_{nlm}|$ darstellt. Kleinman und Bylander [113] konstruierten eine solche Darstellung folgendermaßen:

Zunächst wird (3.26) in einen langreichweitigen lokalen Anteil und eine kurzreichweitige nicht-lokale Korrektur aufgespalten:

$$V^{\text{Ps, ion}} = v_{\text{lok}}^{\text{lang}}(r) + \Delta V_{\text{nllok}}^{\text{kurz}} = v_{\text{lok}}^{\text{lang}}(r) + \sum_{n,l,m} |nlm\rangle \Delta v_{nl}^{\text{kurz}}(r) \langle nlm|. \quad (3.27)$$

Hierbei wird ausgenutzt, dass alle Pseudopotentiale (3.24) für große r in die gemeinsame Asymptote $-Z_{\text{val}}/r$ münden. Ab einem bestimmten Abschneideradius r_{nl}^{cut} ist $V^{\text{Ps, ion}}$ daher (nahezu) allein durch den lokalen Anteil gegeben. Der nicht-lokale Anteil ist entsprechend nur für $r < r_{\text{cut}}$ von Null verschieden und durch $\Delta v_{nl}^{\text{kurz}}(r) := v_{nl}^{\text{Ps, ion}}(r) - v_{\text{lok}}^{\text{lang}}(r)$ definiert. Auf Grund der Zentrifugalbarriere $l \cdot (l+1)/2r^2$ der radialen Schrödinger-Gleichung werden Elektronen mit hohen Drehimpulsen l „nach außen gedrückt“ und also allein durch den lokalen Anteil des Pseudopotentials beschrieben, für den eine geeignete Komponente $v_{n_{\text{lok}} l_{\text{lok}}}^{\text{Ps, ion}}$ der Entwicklung (3.26) benutzt wird. Im nicht-lokalen Anteil kann daher die Summe auf einige wenige $l \leq l_{\text{max}}$ beschränkt werden. Da $v_{\text{lok}}^{\text{lang}}$ die Streueigenschaften der höheren Drehimpulskomponenten reproduzieren soll, wird vorzugsweise $l_{\text{lok}} = l_{\text{max}}$ gesetzt.

Nach dieser Vorbereitung wird $\Delta V_{\text{nllok}}^{\text{kurz}}$ durch das separable Kleinman-Bylander-Pseudopotential

$$\Delta V^{\text{KB}} := \sum_{n,l,m} \frac{\Delta v_{nl}^{\text{kurz}} |nlm\rangle \langle nlm| \Delta v_{nl}^{\text{kurz}}}{\langle nlm| \Delta v_{nl}^{\text{kurz}} |nlm\rangle} \quad (3.28)$$

ersetzt, das gerade so konstruiert ist, dass gilt: $\Delta V^{\text{KB}} |nlm\rangle = \Delta V_{\text{nllok}}^{\text{kurz}} |nlm\rangle$. Dieser Übergang kann jedoch zu unphysikalischen „Geisterzuständen“ führen, die Energien unterhalb oder nahe denen der physikalischen Valenzzustände besitzen [114]. Durch eine geeignete Wahl der Abschneideradien r_{nl}^{cut} und der lokalen Komponente l_{lok} kann deren Auftreten jedoch vermieden werden [115].

3.5.3 Nicht-lineare Rumpfkorrektur

Um im Rahmen der „Näherung der starren Rumpfe“ (Abschnitt 3.4) Ausdrücke für die Gesamtenergie sowie das Kohn-Sham-Potential zu erhalten, die allein von der Dichte der Valenzelektronen abhängen, wird u.a. folgende Näherung vorgenommen, die sich mit Gleichung (3.24) auf das Pseudopotentialkonzept überträgt: Es wird unterstellt, dass der Anteil der elektronischen Wechselwirkung im Kohn-Sham-Potential des betrachteten Atoms linear in der Dichte ist und daher für Rumpf- und Valenzladungsdichte getrennt berechnet werden kann. Dies ist exakt jedoch nur für den elektrostatischen Anteil erfüllt:

$$v_{\text{H}}[\rho_1 + \rho_2](\mathbf{r}) = \int \frac{\rho_1(\mathbf{r}') + \rho_2(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3 r' = v_{\text{H}}[\rho_1](\mathbf{r}) + v_{\text{H}}[\rho_2](\mathbf{r}). \quad (3.29)$$

Sowohl für das lokale Austauschpotential als auch für die Korrelation ist eine analoge Relation, streng formal gesehen, nicht gültig. Die Potentiale beider Wechselwirkungen sind nicht-linear; im Allgemeinen gilt daher

$$v_x[\rho_1 + \rho_2](\mathbf{r}) \neq v_x[\rho_1](\mathbf{r}) + v_x[\rho_2](\mathbf{r}) \quad (3.30)$$

$$v_c[\rho_1 + \rho_2](\mathbf{r}) \neq v_c[\rho_1](\mathbf{r}) + v_c[\rho_2](\mathbf{r}) . \quad (3.31)$$

Diese Nicht-Linearität kann zu Fehlern führen, wenn die Ladungsdichten der Rumpf- und der Valenzzustände räumlich nicht mehr hinreichend voneinander getrennt sind, wie es beispielsweise bei Elementen mit nur wenigen Elektronen in der Valenzschale oft der Fall ist (Alkalimetalle). Man kann dann allerdings versuchen, eine so genannte „nicht-lineare Rumpfkorrektur“ (engl.: *non-linear core correction* — „nlcc“) [104, 105, 115] vorzunehmen. Die dahinter stehende Idee ist, im Austausch-Korrelationspotential nun wieder die gesamte Ladungsdichte zu benutzen, und nicht nur — wie in Gleichung (3.19) — die Valenzladungsdichte. Die Gesamtladungsdichte ist dabei näherungsweise durch die Summe der (atomaren) Rumpfladungsdichten und der Valenzladungsdichte gegeben. Da die Rumpfladungsdichte allerdings stark lokalisiert und oszillierend ist, lässt sie sich nicht effektiv im Rahmen einer Ebene-Wellen-Darstellung beschreiben. Sie wird daher durch eine so genannte „partielle Rumpfladungsdichte“ (engl.: *partial core density*) $\tilde{\rho}_R$ ersetzt, die im chemisch relevanten Bereich außerhalb eines geeigneten Radius r_{nlcc} die echte Rumpfladungsdichte ρ_R reproduziert, innerhalb von r_{nlcc} jedoch durch eine leichter fouriertransformierbare Funktion gegeben ist:

$$\tilde{\rho}_R(r) := \begin{cases} \rho_R(r) & \text{für } r \geq r_{\text{nlcc}} \\ b(r)\rho_R(r) & \text{für } r < r_{\text{nlcc}}, \text{ mit geeignetem } b(r) \leq 1 . \end{cases} \quad (3.32)$$

Damit wird das Austausch-Korrelationspotential wie folgt „linearisiert“:

$$v_{\text{xc}}[\rho_R + \rho_V] \approx v_{\text{xc}}[\tilde{\rho}_R + \rho_V] + \underbrace{(v_{\text{xc}}[\rho_R + \rho_V] - v_{\text{xc}}[\tilde{\rho}_R + \rho_V])}_{=: \tilde{\Delta}v_{\text{xc}}[\rho_V, \rho_R, \tilde{\rho}_R]} . \quad (3.33)$$

In der Kohn-Sham-Gleichung für mehrere Atome wird dann $v_{\text{xc}}[\tilde{\rho}_R^{\text{SP}} + \rho]$ anstelle von $v_{\text{xc}}[\rho]$ verwendet (wobei $\tilde{\rho}_R^{\text{SP}}$ die Superposition der festen atomaren partiellen Rumpfladungsdichten bezeichnen soll und ρ die Dichte aller explizit berücksichtigten Elektronen sei), und die Ionenpotentiale sind gegeben, indem in Gleichung (3.20) entsprechend $\Delta v_{\text{xc}}[\rho_R, \rho_V]$ durch $\tilde{\Delta}v_{\text{xc}}[\rho_V, \rho_R, \tilde{\rho}_R]$ ersetzt wird. Es lässt sich $\tilde{\rho}_R$ also als Kontrollparameter auffassen [105], der im Falle $\tilde{\rho}_R := \rho_R$ die Rumpf-Valenz-Austauschwechselwirkung exakt beschreibt (zumindest für das Pseudoatom), im Falle $\tilde{\rho}_R := 0$ hingegen in die in Abschnitt 3.4 beschriebene Näherung der starren Rumpfe *ohne* die nicht-lineare Rumpfkorrektur übergeht. Im Zusammenhang mit dem Pseudopotentialkonzept folgt dann, dass die Gleichung (3.24) für den radialen Anteil des ionischen Pseudopotentials in

$$v_{nl}^{\text{Ps, ion}}(r) := v_{nl}^{\text{Ps, abg}}(r) - v_{\text{H}}[\rho^{\text{Ps}}](r) - v_{\text{xc}}[\tilde{\rho}_R + \rho^{\text{Ps}}](r) \quad (3.34)$$

abgeändert werden muss.

Kapitel 4

Implementierung

Im vorangehenden Kapitel wurden grundlegende Techniken zur Beschreibung periodischer Systeme, insbesondere mit Hilfe ebener Wellen in Verbindung mit Pseudopotentialen, dargestellt. Darauf aufbauend soll nun die im Rahmen dieser Arbeit durchgeführte numerische Realisierung der variationellen EXX-Methode (varEXX) mit dem Ebene-Wellen-Pseudopotential-Code S/PHI/nX [81, 82] dokumentiert werden. Gemäß Kapitel 2 sind die beiden Schlüsselpunkte der Methode: (i) die Lösung der Sternheimer-Gleichung (2.2) und (ii) die Minimierung der induzierten Ladungsdichte (2.4). In den folgenden drei Abschnitten 4.1 bis 4.3 werden daher zunächst numerische Verfahren (Minimierungs- und Mixing-Techniken) zur effizienten Behandlung dieser Aufgaben vorgestellt und — wo notwendig — an die speziellen Erfordernisse von varEXX angepasst. Eine weitere Schwierigkeit stellt die Auswertung des nicht-lokalen Austauschoperators (1.27) dar, die in Abschnitt 4.4 besprochen wird. Im Anschluss daran werden die Einzelschritte dann zu einem iterativen Zyklus zusammengesetzt, der aus einer „inneren Schleife“ besteht, die das exakte Austauschpotential einer gegebenen Dichte berechnet, und einer „äußeren“, in der die Kohn-Sham-Gleichung selbstkonsistent gelöst wird (Abschnitt 4.5). Abschließend wird ein Näherungsansatz für das EXX-Verfahren vorgestellt, der sowohl bei der ursprünglichen Formulierung (stEXX) als auch bei der variationellen Formulierung (varEXX) angewandt werden kann und in beiden Fällen eine drastische Reduktion des numerischen Aufwands ermöglicht, zugleich aber Bandstrukturen liefert, die sehr gut mit jenen einer exakten Rechnung übereinstimmen (Abschnitt 4.6).

4.1 Die Sternheimer-Gleichung

Wie in Abschnitt 2.3 bereits ausgeführt, lässt sich die Berechnung der Greenschen Funktion G_i , also die explizite Umkehrung von $H_0 - \varepsilon_i^0$ im Orthogonalraum zu ε_i^0 , durch die Verwendung der Sternheimer-Gleichung (2.2) ersetzen, die dann „variationell“ gelöst wird. Der Vorteil ist, dass zur Darstellung der linearen Korrekturfunktionen ψ_i^1 keine unbesetzten Zustände mehr benötigt werden. Der Bedarf an Speicherplatz zur Berechnung der induzierten Ladungsdichte liegt damit in der gleichen Größenordnung wie der zur Berechnung der Grundzustandsladungsdichte, wenn man von der Störung, im Wesentlichen also dem nicht-lokalen Austauschoperator, einmal absieht. Andererseits gewinnt man auch hinsichtlich der Rechenzeit: Wie das Folgende zeigt, lässt sich mit geeigneten Methoden ein erheblich besseres Skalierungsverhalten der CPU-Zeit erzielen als beim Greensfunktion-Ansatz.

Als Variationsmethoden eignen sich Minimierungsverfahren, zwei der gebräuchlichsten

wurden dazu in dieser Arbeit implementiert: das *Verfahren des steilsten Abstiegs* und das zumeist effizientere *Verfahren der konjugierten Gradienten*. Die Effizienz beider lässt sich mit Hilfe einer so genannten *Präkonditionierung* nochmals steigern. Jedoch lassen sich diese Methoden nicht unmittelbar auf die Sternheimer-Gleichung anwenden — dafür sind zunächst einige Vorbetrachtungen notwendig, die im Folgenden diskutiert werden sollen. Recht detaillierte Ergänzungen dazu sowohl in praktischer als auch in theoretischer Hinsicht finden sich in Anhang C.

4.1.1 Numerische Lösung der Sternheimer-Gleichung: Umformulierung in ein Minimierungs-Problem

Die Sternheimer-Gleichung ist die Grundgleichung der Rayleigh-Schrödingerschen Störungstheorie in 1. Ordnung:

$$(H_0 - \varepsilon_i^0)|\psi_i^1\rangle + (H_1 - \varepsilon_i^1)|\psi_i^0\rangle = 0. \quad (4.1)$$

Hier bezeichne H_0 den Hamilton-Operator eines ungestörten Systems mit den bekannten Eigenenergien ε_i^0 und Eigenfunktionen ψ_i^0 , dem eine durch H_1 beschriebene Störung aufgeschaltet sei. Die linearen Korrekturfunktionen seien durch die ψ_i^1 beschrieben. Es erweist sich als bequem, ψ_i^1 orthogonal zum Eigenraum von ε_i^0 anzusetzen.¹ Die Energiekorrektur ε_i^1 ergibt sich, indem die Sternheimer-Gleichung von links mit $\langle\psi_i^0|$ multipliziert wird:

$$\varepsilon_i^1 = \langle\psi_i^0|H_1|\psi_i^0\rangle. \quad (4.2)$$

Numerisch, d.h., in einer endlichen Basisdarstellung, ist die Sternheimer-Gleichung ein unterbestimmtes² System m gekoppelter linearer Gleichungen in m Variablen. Bei der Beschreibung der hier untersuchten Materialien mit ebenen Wellen (vgl. Abschnitt 3.3) liegt die Anzahl m der benötigten „ $(\mathbf{G} + \mathbf{k})$ -Vektoren“ typischerweise in einer Größenordnung zwischen 10^2 und 10^4 . Bei einer direkten Diagonalisierung von $(H_0 - \varepsilon_i^0)$, was der Bestimmung der Greenschen Funktion (B.11) entspricht, skaliert die Anzahl der benötigten Fließkommaoperationen proportional zu m^3 . Eine geeignete Strategie bei so aufwendigen Gleichungssystemen ist, das Problem als Minimierungsaufgabe zu formulieren. Für deren Lösung gibt es dann Verfahren, die proportional zu m^2 skalieren. Sie werden in Abschnitt 4.1.2 besprochen.

Gesucht ist also (für jedes i) ein Funktional, das eine (Probe-)Korrekturfunktion auf eine reelle Zahl abbildet, also ein $E_i : \mathcal{H} \rightarrow \mathbb{R}; \psi_i^1 \mapsto E_i[\psi_i^1]$ (\mathcal{H} : Hilbertraum der Einzelchenzustände), und das genau für jenes ψ_i^1 minimal wird, das Lösung der Sternheimer-Gleichung ist. Eine Möglichkeit, ein solches Funktional zu konstruieren, ist, auf beiden Seiten von (4.1) das Betragsquadrat zu bilden. Auf der rechten Seite steht dann die (reelle Zahl) 0 und auf der linken ergibt sich unmittelbar ein Funktional von ψ_i^1 in der gesuchten Form. Dies funktioniert für alle auf \mathcal{H} definierten Operatoren H_0 , also auch für solche, die

¹Die Sternheimer-Gleichung (4.1) besitzt, falls sie lösbar ist, im Allgemeinen keine eindeutige Lösung. Denn ist ψ_i^1 Lösung der Sternheimer-Gleichung, so ist z.B. auch $\psi_i^1 + \lambda\psi_i^0$ mit jedem $\lambda \in \mathbb{C}$ Lösung. Durch die zusätzliche Forderung der Orthogonalität von ψ_i^1 zum Eigenraum von ε_i^0 wird aus dem Lösungsraum ein eindeutiges Element ausgewählt. Man beachte, dass (4.2) nicht von der speziellen Wahl des Winkels zwischen Original- und Korrekturfunktion abhängt.

² $(H_0 - \varepsilon_i^0)$ besitzt einen Nullraum, den Eigenraum zu ε_i^0 , und kann folglich in m -dimensionaler Basisdarstellung höchstens vom Rang $m - 1$ sein.

nicht hermitesch sind.³ Für den in dieser Arbeit relevanten Fall eines hermiteschen Operators H_0 lässt sich das Funktional aber noch auf elegantere Weise konstruieren, nämlich so, dass der für die Minimierung benötigte Gradient gerade durch die linke Seite der Sternheimer-Gleichung gegeben ist. Das Funktional ist dann gewissermaßen die Stammfunktion von (4.1):

$$E_i[\psi_i^1] := \frac{1}{2} \langle \psi_i^1 | (H_0 - \varepsilon_i^0) | \psi_i^1 \rangle + \operatorname{Re} \langle \psi_i^1 | (H_1 - \varepsilon_i^1) | \psi_i^0 \rangle . \quad (4.3)$$

Die Richtung des steilsten Anstiegs ist dabei durch

$$2 \frac{\delta E_i[\psi_i^1]}{\delta (\psi_i^1)^*} = (H_0 - \varepsilon_i^0) | \psi_i^1 \rangle + (H_1 - \varepsilon_i^1) | \psi_i^0 \rangle \quad (4.4)$$

gegeben. Hier wurde benutzt, dass dazu gemäß der (Variations-)Ableitung im Komplexen ψ_i^1 und $(\psi_i^1)^*$ formal als unabhängige Vektoren aufzufassen sind. Die Differentiation erfolgt nur nach der konjugiert-komplexen Funktion $(\psi_i^1)^*$, dafür ist allerdings noch mit einem Faktor 2 zu multiplizieren, vgl. z.B. Ref. 126. Die Sternheimer-Gleichung lässt sich damit auch so formulieren:

$$\frac{\delta E_i[\psi_i^1]}{\delta (\psi_i^1)^*} = 0 , \quad (4.5)$$

d.h., Lösungen der Sternheimer-Gleichung lassen das in (4.3) angegebene Funktional stationär werden. Das Funktional hat die Dimension einer Energie und stellt eine quadratische Form in ψ_i^1 dar. Die tiefere Begründung dafür ist, dass es sich als eine Art Ritzsches Variationsprinzip für die Korrekturfunktion auffassen lässt. Dies ist in Anhang C.1 ausgeführt. Bevor die oben erwähnten numerischen Algorithmen eingesetzt werden können, um stationäre Punkte von E_i aufzuspüren, muss allerdings zunächst klar sein, ob es sich dort um ein Minimum, ein Maximum oder einen Sattelpunkt handelt. Wie in Ref. 78 gezeigt, lässt sich der Hilbertraum \mathcal{H} in drei Bereiche \mathcal{H}^- , \mathcal{H}^0 und \mathcal{H}^+ untergliedern, für die gilt:

$$\langle \psi | (H_0 - \varepsilon_i^0) | \psi \rangle < 0 \quad \text{für } \psi \in \mathcal{H}^- \setminus \{0\} \quad \text{mit } \mathcal{H}^- := \operatorname{span}\{\psi_j^0 | \varepsilon_j^0 < \varepsilon_i^0\} , \quad (4.6)$$

$$\langle \psi | (H_0 - \varepsilon_i^0) | \psi \rangle = 0 \quad \text{für } \psi \in \mathcal{H}^0 \quad \text{mit } \mathcal{H}^0 := \operatorname{span}\{\psi_j^0 | \varepsilon_j^0 = \varepsilon_i^0\} , \quad (4.7)$$

$$\langle \psi | (H_0 - \varepsilon_i^0) | \psi \rangle > 0 \quad \text{für } \psi \in \mathcal{H}^+ \setminus \{0\} \quad \text{mit } \mathcal{H}^+ := \operatorname{span}\{\psi_j^0 | \varepsilon_j^0 > \varepsilon_i^0\} . \quad (4.8)$$

Im linearen Teilraum \mathcal{H}^+ ist $(H_0 - \varepsilon_i^0)$ also positiv definit⁴, und E_i besitzt dort folglich ein eindeutiges Minimum. In \mathcal{H}^- ist $(H_0 - \varepsilon_i^0)$ hingegen negativ definit, so dass E_i dort ein eindeutiges Maximum aufweist. Falls die Sternheimer-Gleichung eine Lösung besitzt, wird es sich dabei also im Allgemeinen — außer, wenn ε_i^0 die tiefste Eigenenergie von H_0 ist — um einen Sattelpunkt von E_i handeln, vgl. Abb. 4.1. Der Teilraum \mathcal{H}^0 schließlich ist der Eigenraum von ε_i^0 . Hier reduziert sich E_i wegen (4.7) auf ein lineares Funktional von ψ_i^1 , dessen Gradient (4.4) durch den konstanten Vektor $(H_1 - \varepsilon_i^1) | \psi_i^0 \rangle$ gegeben ist. Je nachdem, ob dieser Vektor verschwindet oder nicht, ist die Sternheimer-Gleichung (4.5)

³Dieser Fall ist realistisch: Ist H_0 beispielsweise der Hamilton-Operator eines kugelsymmetrischen Systems (Atom) in Ortsdarstellung, wobei die Radialanteile der Wellenfunktionen mit Hilfe einer Basis *logarithmischer* Stützstellen beschrieben werden sollen, ist das (diskretisierte) H_0 nicht mehr hermitesch. Dies wird — quasi als Nachtrag zur Diplomarbeit [78] — in Anhang C.3 diskutiert.

⁴Ein auf einem Hilbertraum \mathcal{H} erklärter linearer Operator A heißt genau dann „positiv definit“, wenn gilt: $\langle \psi | A | \psi \rangle > 0$ für alle $\psi \in \mathcal{H} \setminus \{0\}$. Analog ist „negativ definit“ definiert, indem „>“ durch „<“ ersetzt wird. Ist die 2. (Variations-)Ableitung eines Funktionals an einer Stelle durch einen positiv/negativ definiten Operator gegeben, besitzt es dort ein Minimum/Maximum. Die 2. Ableitung von E_i ist $(H_0 - \varepsilon_i^0)$.

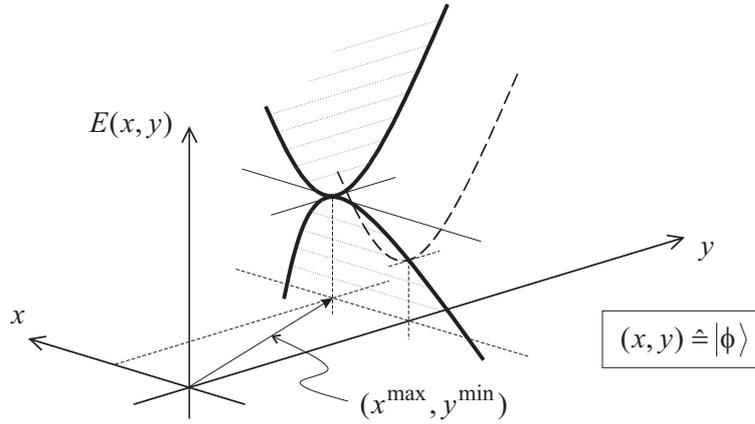


Abbildung 4.1: Dargestellt ist ein Paraboloid zur Gleichung $E(x, y) := a_{11}x^2 + a_{22}y^2 + b_1x + b_2y + const.$ mit $a_{11} < 0$ und $a_{22} > 0$. Bei (x^{max}, y^{min}) hat E einen Sattelpunkt. Die x -Achse entspricht hier dem linearen Teilraum \mathcal{H}^- und die y -Achse dem Teilraum \mathcal{H}^+ in (4.6) bzw. (4.8). Aus der Zeichnung erkennt man, dass wegen $\mathcal{H}^- \perp \mathcal{H}^+$ die Extremwertbestimmung in den beiden Teilräumen separat durchgeführt werden kann. Zur Bestimmung der induzierten Ladungsdichte ist jedoch nur der Teilraum \mathcal{H}^+ relevant.

also in ganz $\mathcal{H} = \mathcal{H}^- \oplus \mathcal{H}^0 \oplus \mathcal{H}^+$ lösbar oder nicht. Ist sie lösbar, ist der Lösungsraum durch alle

$$\psi_i^1 = \phi_{\max} + \phi_0 + \phi_{\min} \quad \text{mit} \quad \begin{cases} \phi_{\max} \in \mathcal{H}^-, & E_i[\phi_{\max}] = \max_{\mathcal{H}^-} \{E_i[\phi]\} \\ \phi_{\min} \in \mathcal{H}^+, & E_i[\phi_{\min}] = \min_{\mathcal{H}^+} \{E_i[\phi]\} \\ \phi_0 \in \mathcal{H}^0 \end{cases} \quad (4.9)$$

gegeben. ϕ_{\max} und ϕ_{\min} lassen sich analytisch über die Projektion der Greenschen Funktion (B.11) auf \mathcal{H}^- bzw. \mathcal{H}^+ darstellen:

$$\phi_{\max} = \sum_{\substack{j \\ \varepsilon_j^0 < \varepsilon_i^0}} \psi_j^0 \frac{\langle \psi_j^0 | H_1 | \psi_i^0 \rangle}{\varepsilon_i^0 - \varepsilon_j^0}, \quad \phi_{\min} = \sum_{\substack{j \\ \varepsilon_j^0 > \varepsilon_i^0}} \psi_j^0 \frac{\langle \psi_j^0 | H_1 | \psi_i^0 \rangle}{\varepsilon_i^0 - \varepsilon_j^0}. \quad (4.10)$$

Wird die Sternheimer-Gleichung nur für die Valenzzustände gelöst — dies ist im variationalen EXX-Verfahren der Fall (vgl. Abschnitt 2.1) —, gehen im Gegensatz zur Darstellung von ϕ_{\min} bei ϕ_{\max} nur Zustände und Energien von Valenzelektronen ein. Die Berechnung von ϕ_{\max} über (4.10) wird daher in der Regel effizienter sein als mit einem Variationsverfahren.

Die Frage nach der Lösbarkeit der Sternheimer-Gleichung hängt unmittelbar damit zusammen, ob das ungestörte System Entartungen aufweist oder nicht. Dies ist in Anhang C.2 genauer ausgeführt. Dabei zeigt sich: Ist der Eigenwert ε_i^0 nicht entartet, ist die Sternheimer-Gleichung (4.1) stets lösbar. Im Falle einer Entartung des ungestörten Systems gilt dies jedoch nicht, und entsprechend besitzt das Funktional (4.3) im Allgemeinen keinen stationären Punkt! Tatsächlich hat man es in der Praxis aber vielfach mit entarteten Systemen zu tun. So ist bei vielen Halbleitern (etwa bei den hier untersuchten Materialien GaAs, GaN, InN) das oberste Valenzband am Γ -Punkt dreifach entartet. Ist H_0 der Kohn-Sham-Operator eines LDA-Systems, verschwindet zudem in einigen Fällen (z.B. InN, Ge) fälschlicherweise die Bandlücke, wodurch sich der Entartungsgrad nochmals

erhöhen kann.⁵ Typischerweise weisen aber nicht nur Hochsymmetriepunkte, sondern ganze Bereiche der Brillouin-Zone Entartungen auf. Daher lässt sich das Problem auch in einer geeignet diskretisierten Brillouin-Zone (z.B. mit dem Monkhorst-Pack-Verfahren [96], vgl. Abschnitt 3.2), in der Hochsymmetriepunkte üblicherweise vermieden werden, nicht umgehen.

Im Prinzip ließe sich nun diese Schwierigkeit auf analoge Weise behandeln wie in der entarteten Störungstheorie, nämlich durch Diagonalisierung von H_1 (vgl. Anhang C.2). Es zeigt sich jedoch, dass im variationellen EXX-Verfahren die gesamte Lösung ψ_i^1 der Sternheimer-Gleichung gar nicht benötigt wird: Zur Berechnung der induzierten Ladungsdichte ist nämlich nur der Anteil von ψ_i^1 relevant, der im Hilbertraum \mathcal{H}^c der unbesetzten Zustände liegt. Alle Komponenten in Richtung besetzter Zustände können dagegen beliebig gewählt werden, also etwa durch Projektion gleich Null gesetzt werden. Bezeichnet P_c den Projektionsoperator auf \mathcal{H}^c , kann statt (2.4) also auch

$$\Delta\rho = \sum_i^{\text{occ}} (\psi_i^0)^* P_c \psi_i^1 + \text{c.c.} \quad (4.11)$$

verwendet werden. Dies liegt daran, dass sich in $\Delta\rho$ der Anteil der Korrektur in Richtung der Valenzzustände gerade herauskürzt, wie es sich analog auch im Standard-EXX-Verfahren ergibt, vgl. Abschnitt 2.2. Physikalisch bedeutet dies, dass die Ladungsdichte zumindest in linearer Ordnung nicht auf eine Störung reagiert, die nur besetzte Zustände untereinander koppelt [85]. $P_c \psi_i^1$ ist Lösung der „vereinfachten“ Sternheimer-Gleichung

$$P_c [(H_0 - \varepsilon_i^0)|\psi_i^1\rangle + (H_1 - \varepsilon_i^1)|\psi_i^0\rangle] = (H_0 - \varepsilon_i^0)P_c|\psi_i^1\rangle + P_c H_1|\psi_i^0\rangle = 0. \quad (4.12)$$

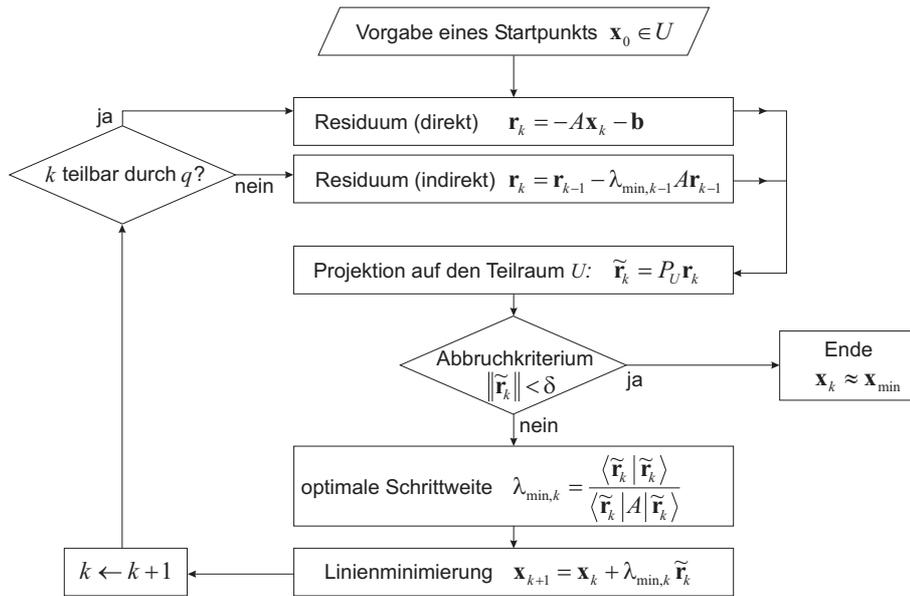
Hier wurde $[P_c, H_0] = 0$ und $P_c \psi_i^0 = 0$ für $i \in \{1, \dots, N\}$ benutzt. Gegenüber (4.1) hat diese Gleichung numerisch einige Vorteile: (i) Es ist nicht mehr notwendig, die Energiekorrektur ε_i^1 zu berechnen. (ii) $P_c \psi_i^1$ ist das eindeutige Minimum von (4.3) in \mathcal{H}^c , da $\mathcal{H}^c \subset \mathcal{H}^+$, d.h., die aufwendige Kombination von Maximierung und Minimierung zur Bestimmung des Sattelpunkts ψ_i^1 entfällt. Im Falle einer Entartung schließlich muss (iii) nicht mehr der Eigenraum zu ε_i^0 ermittelt werden, (iv) entfällt die sonst notwendige Diagonalisierung von H_1 , vgl. Anhang C.2, und (v) ist nicht mehr zu *entscheiden*, wann ein Eigenwert überhaupt entartet ist oder nicht, was numerisch nur durch Vergleichen der Differenz zweier Eigenwerte mit einem kleinen, aber letztlich willkürlich festgelegten Parameter δ möglich wäre.

4.1.2 Minimierungsverfahren

Die beiden im Folgenden vorgestellten Verfahren, *Steepest Descent* und *Conjugate Gradient*, sind für die Minimierung einer quadratischen Form konzipiert und damit für die Anwendung auf das Energiefunktional (4.3) besonders geeignet. Besitzt eine (mindestens 2-mal stetig differenzierbare, reellwertige) Funktion ein Minimum bei \mathbf{x} , kann jedoch die Umgebung von \mathbf{x} in vielen Fällen⁶ durch eine quadratische Form approximiert werden (Taylorentwicklung); daher sind *Steepest Descent* und *Conjugate Gradient* ganz generell Standardmethoden zur numerischen Minimierung.

⁵In LDA wird die Bandlücke also in einigen Fällen so stark unterschätzt, dass das System metallisch erscheint (vgl. auch Abb. 1 und 2). Der EXX-Formalismus wäre dann durch Einführung einer Fermi-Funktion geeignet zu verallgemeinern.

⁶2-malige stetige Differenzierbarkeit ist notwendige Voraussetzung für eine Taylorentwicklung der Funktion bis zum 2. Grad, sie ist allerdings nicht hinreichend, vgl. z.B. Ref. 123. Die meisten in der Physik relevanten Funktionen sind jedoch taylorentwickelbar.



Flussdiagramm 4.1: Die Methode des steilsten Abstiegs (optimiert). Der Iterationszyklus wird beendet, sobald das Residuum kleiner als δ ist. Der Projektionsschritt und die Startbedingung $\mathbf{x}_0 \in U$ entfallen, wenn die Minimierung nicht auf den linearen Teilraum U beschränkt sein soll. Werden viele Iterationen benötigt, empfiehlt es sich, alle q (z.B. $q > 10$) Iterationen das Residuum direkt über den Gradienten zu berechnen, um eine Akkumulation numerischer Rundungsfehler zu vermeiden.

Eine Wellenfunktion $\psi \in \mathcal{H}$ wird numerisch im Allgemeinen durch einen Vektor $\mathbf{x} \in \mathbb{C}^m$ beschrieben, wobei m die Größe des verwendeten Basissatzes ist. In ebenen Wellen beispielsweise entsprechen die Elemente von \mathbf{x} den Fourierkoeffizienten $c_{n\mathbf{k}}(\mathbf{G})$ von Gleichung (3.16). Deshalb muss die Minimierung hier im \mathbb{C}^m betrachtet werden. Sei also A eine hermitesche, positiv definite $(m \times m)$ -Matrix und \mathbf{b} ein Vektor in \mathbb{C}^m . Dann ist durch

$$F : \mathbb{C}^m \longrightarrow \mathbb{R} ; \mathbf{x} \longmapsto \frac{1}{2} \langle \mathbf{x} | A | \mathbf{x} \rangle + \frac{1}{2} \langle \mathbf{x} | \mathbf{b} \rangle + \frac{1}{2} \langle \mathbf{b} | \mathbf{x} \rangle \quad (4.13)$$

oder ausgeschrieben

$$F(x_1, \dots, x_m) := \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^m x_i^* a_{ij}^* x_j + \frac{1}{2} \sum_{i=1}^m x_i^* b_i + \frac{1}{2} \sum_{i=1}^m b_i^* x_i \quad (4.14)$$

eine quadratische Form definiert, deren Gradient durch

$$\nabla F(\mathbf{x}) = A \mathbf{x} + \mathbf{b} \quad \text{mit} \quad \nabla_i F(\mathbf{x}) := 2 \frac{\partial}{\partial x_i^*} F(\mathbf{x}) = \sum_{j=1}^m a_{ij}^* x_j + b_i \quad (4.15)$$

gegeben ist. Zur Definition der Ableitung im Komplexen vgl. Ref. 126. Ein intuitives Verständnis der Methoden kann man erhalten, wenn man sich F als einen von zwei reellen Variablen abhängenden Paraboloiden vorstellt, der in der Ebene durch elliptische Höhenlinien veranschaulicht werden kann. Dies wird im Folgenden zur Veranschaulichung der Methoden benutzt. Die Darstellung orientiert sich dabei an der von Shewchuk [131].

4.1.2.1 Methode des steilsten Abstiegs (*Steepest Descent*)

Die Idee dieser Methode ist, von einem beliebigen Startpunkt $\mathbf{x}_0 \in \mathbb{C}^m$ aus so lange der Richtung des steilsten Abstiegs, also $-\nabla F(\mathbf{x}_0)$, zu folgen, bis man entlang dieser

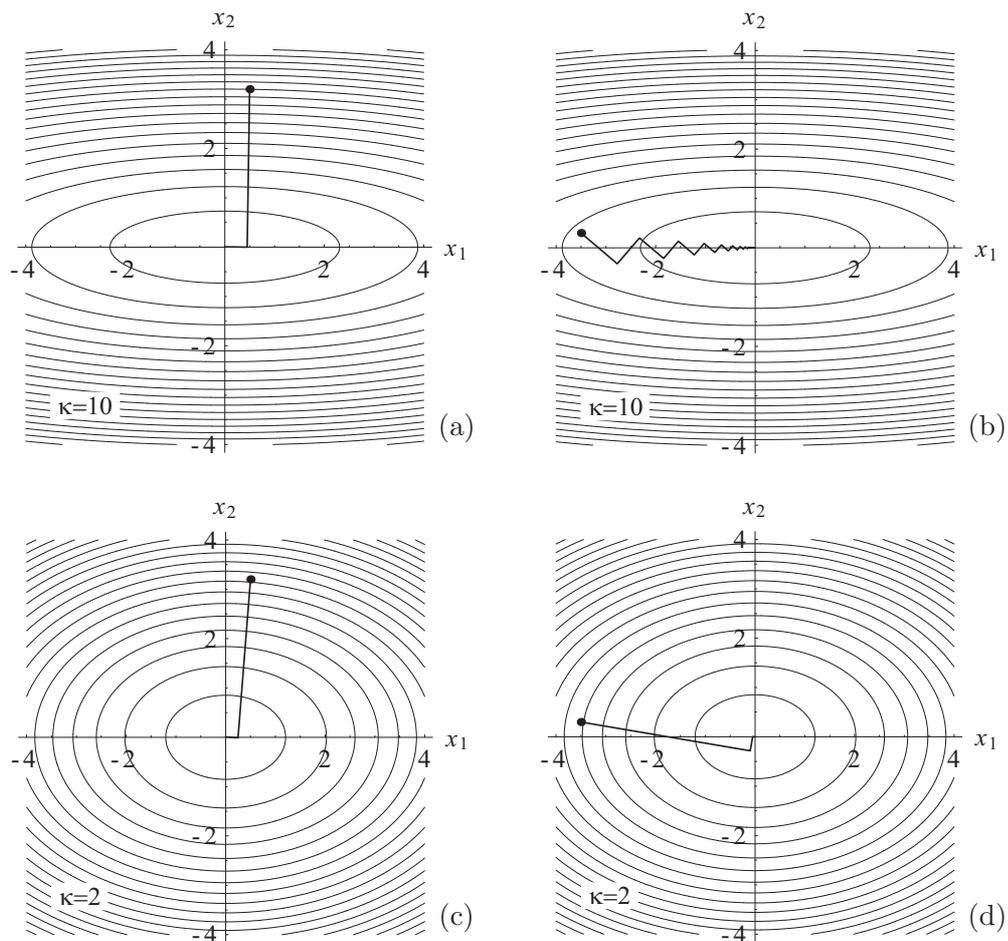


Abbildung 4.2: Konvergenzverhalten der Methode des steilsten Abstiegs in Abhängigkeit von der Konditionierung κ und dem Startpunkt. (a) Großes κ , günstige Wahl des Startpunkts. (b) Großes κ , ungünstiger Startpunkt: Die Methode benötigt in dem schmalen Tal viele Schritte, um das Minimum zu finden. In (c) und (d) sind die Startpunkte jeweils wie darüberliegend gewählt, aber κ ist klein. Auch mit einem schlechten Startpunkt ist die Konvergenz schnell. Bei der idealen Konditionierung $\kappa = 1$ wären die Höhenlinien Kreise, und das Minimum wäre in einem einzigen Schritt erreicht. Nach Ref. 131.

Linie das Minimum \mathbf{x}_1 erreicht hat. In \mathbf{x}_1 bestimmt man erneut den Gradienten und wiederholt diesen Prozess iterativ so oft, bis man dicht genug ans Minimum \mathbf{x}_{\min} gelangt ist. Am Minimum selbst verschwindet der Gradient, so dass „dicht genug“ z.B. über ein Abbruchkriterium definiert werden kann, das den Prozess beendet, sobald die Steigung eine bestimmte Untergrenze δ erreicht hat. Die Punkte \mathbf{x}_k definieren eine Folge, die im Prinzip ins exakte Minimum konvergiert:

$$\lim_{k \rightarrow \infty} \mathbf{x}_k \longrightarrow \mathbf{x}_{\min} . \quad (4.16)$$

Der Fehler des k -ten Schrittes ist durch $\mathbf{e}_k := \mathbf{x}_k - \mathbf{x}_{\min}$ gegeben. Damit kann der negative Gradient, in diesem Kontext auch als *Residuum* (Rest) bezeichnet, auch als Anwendung von A auf den Fehler interpretiert werden:

$$\mathbf{r}_k := -\nabla F(\mathbf{x}_k) = -A \mathbf{x}_k - \mathbf{b} = -A \mathbf{x}_k + A \mathbf{x}_{\min} = -A \mathbf{e}_k . \quad (4.17)$$

Steepest Descent gibt es in zwei Varianten: Beim „einfachen“ *Steepest Descent* erhält man die Folge der \mathbf{x}_k über die Rekursion

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \lambda \mathbf{r}_k \quad (4.18)$$

mit einer konstanten Schrittweite λ . Dies konvergiert für

$$0 < \lambda < \frac{\alpha_{\min}}{2} \quad (4.19)$$

mit dem kleinsten Eigenwert α_{\min} von A , was aus (4.18) und der Forderung $\|\mathbf{x}_{k+1} - \mathbf{x}_{\min}\| < \|\mathbf{x}_k - \mathbf{x}_{\min}\|$ folgt.

Schneller konvergiert das „optimierte“ *Steepest-Descent*-Verfahren: Hier wird zunächst die optimale Schrittweite $\lambda_{\min,k}$ bestimmt, die F entlang der Linie $\mathbf{x} = \mathbf{x}_k + \lambda \mathbf{r}_k$ ($\lambda \in \mathbb{R}$) minimiert. Sie ergibt sich, indem die Ableitung von F entlang dieser Linie zu 0 gesetzt wird,

$$\left. \frac{d}{d\lambda} F(\mathbf{x}_k + \lambda \mathbf{r}_k) \right|_{\lambda=\lambda_{\min,k}} \stackrel{!}{=} 0,$$

woraus nach Anwendung der Kettenregel und Einsetzen von (4.17)

$$\lambda_{\min,k} = \frac{\langle \mathbf{r}_k | \mathbf{r}_k \rangle}{\langle \mathbf{r}_k | A | \mathbf{r}_k \rangle} \quad (4.20)$$

folgt. Dies ist reell, da ein hermitesches A vorausgesetzt wurde. Die Folge der \mathbf{x}_k ergibt sich dann aus

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \lambda_{\min,k} \mathbf{r}_k. \quad (4.21)$$

In dieser Form benötigt das optimierte *Steepest Descent* jedoch noch zwei Matrix-Vektor-Multiplikationen pro Iteration, nämlich $A \mathbf{x}_k$ in (4.17) und $A \mathbf{r}_k$ in (4.20). Die erste davon lässt sich eliminieren, indem man (4.21) durch Multiplikation mit $-A$ und Subtraktion von \mathbf{b} etwas umformt:

$$\mathbf{r}_{k+1} = \mathbf{r}_k - \lambda_{\min,k} A \mathbf{r}_k. \quad (4.22)$$

Das Skalierungsverhalten, dominiert von den m^2 Fließkommaoperationen der Matrix-Vektor-Multiplikation, verbessert sich so um einen Faktor 2. Da das Residuum dann allerdings nicht mehr direkt berechnet wird, können Rundungsfehler auftreten, so dass es sich empfiehlt, nach einer bestimmten Anzahl von Iterationen (4.21) zu benutzen.

Bei der Anwendung von *Steepest Descent* auf die Sternheimer-Gleichung ist zu beachten, dass die Minimierung nur in einem bestimmten Unterraum $U \subset \mathbb{C}^m$ durchgeführt wird, der dem Teilraum \mathcal{H}^c aus Abschnitt 4.1.1 entspricht. Dies lässt sich erreichen, indem (i) ein Startpunkt $\mathbf{x}_0 \in U$ gewählt wird und (ii) in (4.21) nur der Anteil des Residuums in U benutzt wird, also $\tilde{\mathbf{r}}_k := P_U \mathbf{r}_k = -P_U \nabla F(\mathbf{x}_k)$ anstelle von \mathbf{r}_k , wobei P_U der Projektor auf U ist. Die Projektion ist numerisch unproblematisch (skaliert linear mit m), da sie gemäß (4.12) eine Orthogonalisierung zu den besetzten Zuständen darstellt. Flussdiagramm 4.1 zeigt schematisch den Ablauf des Algorithmus.

Das Verfahren ist jedoch nur für gut konditionierte Funktionen geeignet. Dabei bestimmt das Verhältnis des größten Eigenwerts α_{\max} zum kleinsten Eigenwert α_{\min} der Matrix A die Konditionierungszahl $\kappa := \alpha_{\max}/\alpha_{\min}$. Ist κ sehr groß, hat F die Gestalt eines schmalen Tals, und *Steepest Descent* konvergiert bei ungünstiger Wahl des Startpunkts

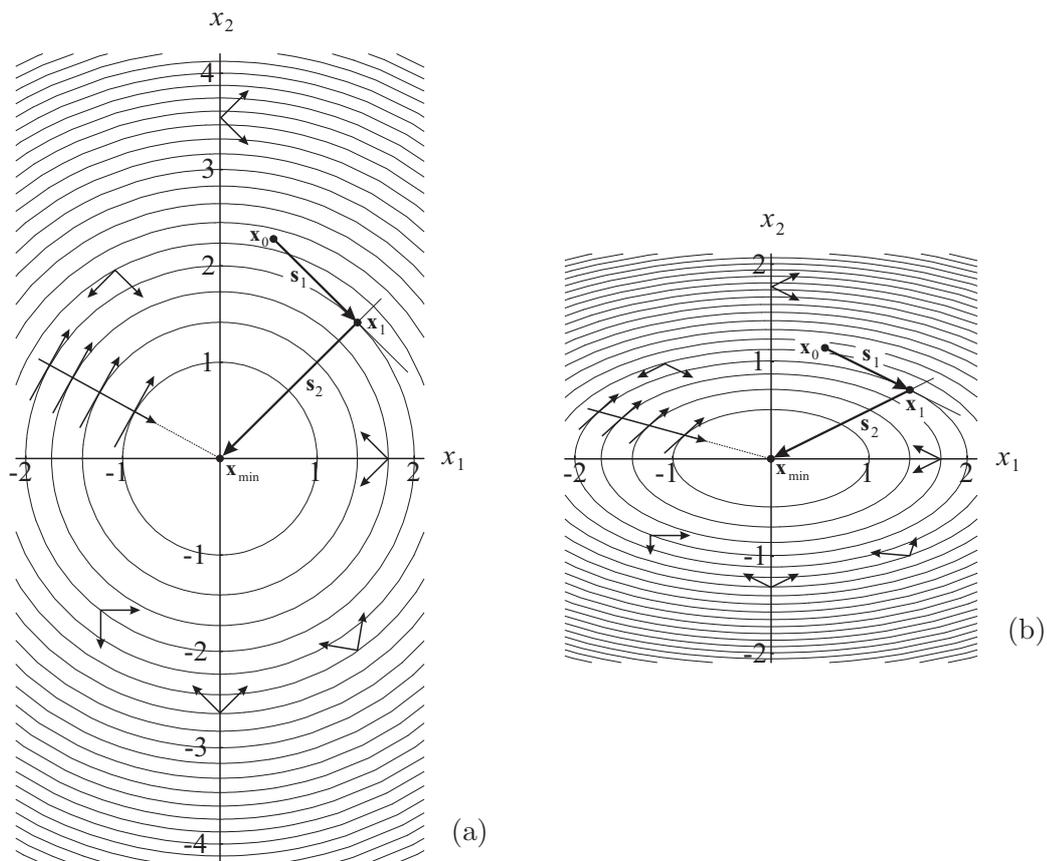


Abbildung 4.3: (a) Paraboloid mit kreisförmigen Höhenlinien: $F_1(\mathbf{x}) := \mathbf{x}^2 = \mathbf{x}^T \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \mathbf{x}$. Führt man Linienminimierungen entlang zweier orthogonaler, aber sonst beliebiger Richtungen $\mathbf{s}_1, \mathbf{s}_2$ aus, gelangt man ins Minimum. (b) zeigt eine um den Faktor 2 gestauchte Kopie von (a); die Ellipsen kennzeichnen jetzt den Paraboloiden $F_2(\mathbf{x}) := \mathbf{x}^T A \mathbf{x}$ mit $A = \begin{pmatrix} 1 & 0 \\ 0 & 1/4 \end{pmatrix}$. Alle rechten Winkel in (a) sind dabei in A -orthogonale Winkel überführt worden, zur Illustration sind einige Pfeilpaare mit eingezeichnet. Nun führen Linienminimierungen entlang zweier beliebiger A -orthogonaler Richtungen ins Ziel. Man beachte, dass beim *Conjugate-Gradient*-Verfahren im Unterschied zu hier die erste Linienminimierung immer in Richtung des Gradienten erfolgt, vgl. Abb. 4.4 (b). Vgl. auch Ref. 131.

nur sehr langsam. Abbildung 4.2 illustriert das Konvergenzverhalten in Abhängigkeit von der Konditionierung und dem Startpunkt.

Im Rahmen der Arbeit wurde die Anwendung der hier vorgestellten Minimierungsmethoden für verschiedene Halbleitermaterialien auf das Energiefunktional (4.3) getestet (vgl. Abb. 4.5). Tatsächlich zeigt dabei *Steepest Descent* ein nur sehr langsames Konvergenzverhalten. Im Falle von InN etwa werden bei einer Abschneideenergie $E_{\text{cut}} = 80 \text{ Ry}$ und dem Startpunkt 0 für die Korrekturfunktionen ψ_i^1 mehr als 600 Iterationen benötigt, um ins Minimum zu gelangen. Eine wesentliche Verbesserung erzielt man mit dem im Folgenden skizzierten *Conjugate-Gradient*-Verfahren.

4.1.2.2 Methode der konjugierten Gradienten (*Conjugate Gradient*)

Der Ansatz dieser Methode nutzt aus, dass die Isooberflächen $F(\mathbf{x}) = \text{const.}$ des Hyper-Paraboloiden (4.13) Hyper-Ellipsoide im \mathbb{C}^m sind, die durch geeignete Streckungen/Stauchungen entlang ihrer Hauptachsen in die Gestalt einer Hyper-Kugel gebracht werden

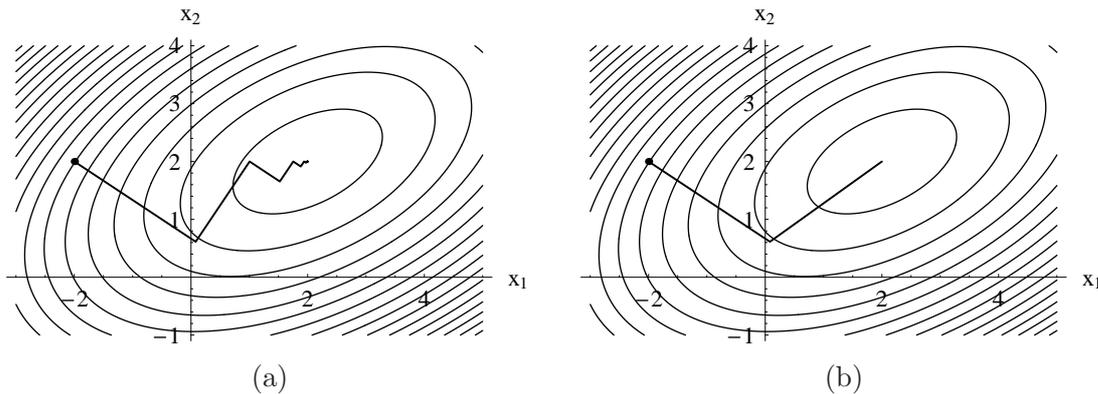


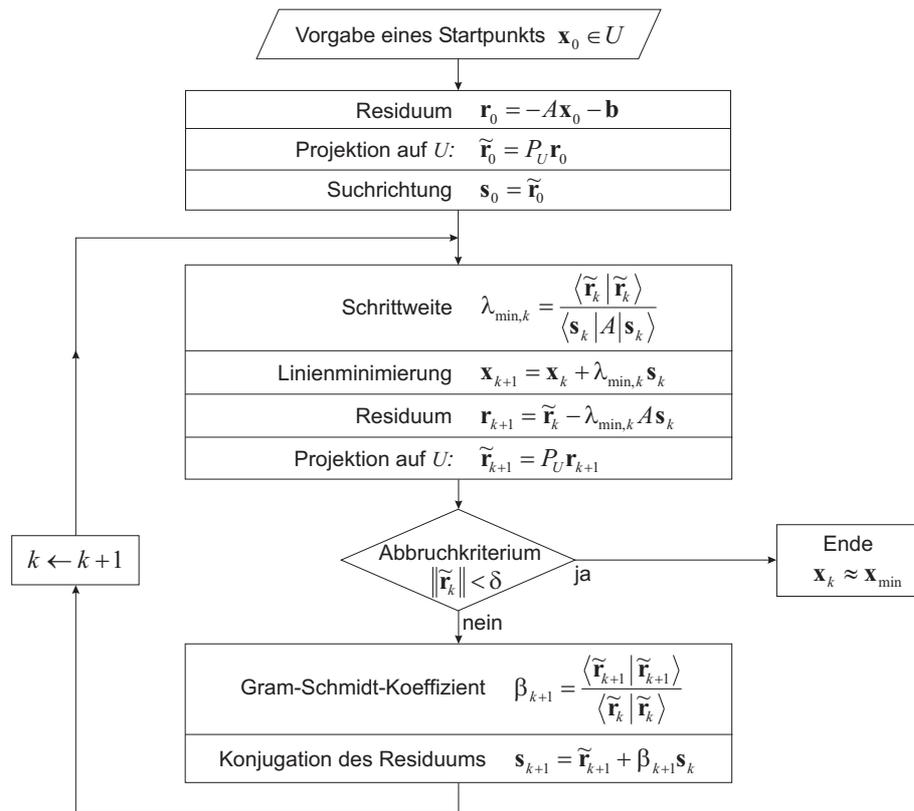
Abbildung 4.4: Minimierung einer quadratischen Funktion im Vergleich: (a) *Steepest Descent*, (b) *Conjugate Gradient*. Im Gegensatz zu *Steepest Descent* findet *Conjugate Gradient* bei einem m -dimensionalen Problem in m Schritten ins Minimum (hier $m = 2$). Der erste Schritt ist in beiden Verfahren durch den Gradienten gegeben und daher identisch. Nach Ref. 131.

können. Legt man nun eine kartesische Basis $\{\mathbf{s}_1, \dots, \mathbf{s}_m\}$ durch den \mathbb{C}^m , gelangt man mit maximal m Schritten ins Minimum \mathbf{x}_{\min} der Kugel, wenn für jede Richtung \mathbf{s}_k genau eine Linienminimierung durchgeführt wird. Mit einem beliebigen Startpunkt \mathbf{x}_0 und m -maliger Anwendung der Rekursionsformel

$$\mathbf{x}_{k+1} := \mathbf{x}_k + \lambda_{\min,k} \mathbf{s}_k \quad (k \in \{0, \dots, m-1\}) \quad (4.23)$$

ist also $\mathbf{x}_{\min} = \mathbf{x}_m$ gefunden, wobei durch $\lambda = \lambda_{\min,k}$ jeweils der Ausdruck $F(\mathbf{x}_k + \lambda \mathbf{s}_k)$ minimiert wird. Dies ist für den 2-dimensionalen reellen Fall in Abb. 4.3 (a) veranschaulicht. Der Vorteil dieser Strategie, die im Falle kugelförmiger Isoflächen (ideale Konditionierung, $\kappa = 1$) natürlich sehr uneffizient wäre, liegt in ihrer Verallgemeinerbarkeit: Ähnlich, wie sich eine Kugel $\|\mathbf{x}\|^2 = \langle \mathbf{x} | \mathbf{x} \rangle = \text{const.}$ als Spezialfall der Ellipse $\langle \mathbf{x} | A | \mathbf{x} \rangle = \text{const.}$ für $A = \mathbf{1}$ auffassen lässt, kann auch die Orthogonalitätsrelation $\langle \mathbf{x} | \mathbf{x}' \rangle = 0$ zur „ A -Orthogonalität“ $\langle \mathbf{x} | A | \mathbf{x}' \rangle = 0$ verallgemeinert werden. A -orthogonale Vektoren werden in diesem Kontext auch als zueinander *konjugiert* bezeichnet. Durch eine entsprechende Verallgemeinerung des Gram-Schmidtschen Orthogonalisierungsverfahrens kann dann aus m beliebigen linear unabhängigen Vektoren im \mathbb{C}^m eine Basis erzeugt werden, deren Vektoren paarweis A -orthogonal zueinander stehen [131]. Benutzt man nun für die Suchrichtungen \mathbf{s}_k in (4.23) anstelle eines rechtwinkligen ein solches A -orthogonales System, findet man, völlig unabhängig von der Konditionierung der Matrix A , das Minimum von (4.13) in maximal m Schritten. Dies ist in Abb. 4.3 (b) demonstriert.

Die Schwierigkeit dieser Vorgehensweise ist, dass das Gram-Schmidt-Verfahren kubisch mit der Dimension m skaliert und zudem alle Suchrichtungen $\{\mathbf{s}_1, \dots, \mathbf{s}_m\}$ im Speicher gehalten werden müssen. Numerisch interessant werden konjugierte Richtungen erst dadurch, dass in jedem Schritt ein zu allen vorherigen Suchrichtungen konjugierter Vektor auch aus dem Gradienten von F (dem Residuum (4.17)) konstruiert werden kann, indem dieser allein zur *letzten* Suchrichtung A -orthogonalisiert wird [128]. Daraus ergibt sich das in Flussdiagramm 4.2 gezeigte Schema des Algorithmus. Abbildung 4.4 veranschaulicht den Vorteil der Methode gegenüber dem Verfahren des steilsten Abstiegs. Für eine tiefere Betrachtung der Thematik sei auf die Referenzen 128, 129, 130, 131 verwiesen. Die zur Lösung der Sternheimer-Gleichung benötigte Restriktion der Minimierung auf einen Teilraum von \mathbb{C}^m kann ebenso leicht wie beim *Steepest-Descent*-Verfahren erreicht werden. Sie ist in Flussdiagramm 4.2 mit angegeben.



Flussdiagramm 4.2: Die Methode der konjugierten Gradienten. Bei diesem Verfahren ergibt sich die Suchrichtung durch Konjugation (A -Orthogonalisierung, vgl. Abb. 4.3) des Gradienten zur vorangehenden Suchrichtung. Die Projektionen und die Startbedingung $\mathbf{x}_0 \in U$ sind nur notwendig, wenn die Minimierung auf einen linearen Teilraum U beschränkt sein soll.

In Abb. 4.5 ist das Konvergenzverhalten des Verfahrens bei Anwendung auf das Energiefunktional (4.3) dargestellt. Es zeigt sich eine deutlich Verbesserung im Vergleich zu *Steepest Descent*. Je nach Abschneideenergie wurden bei den getesteten Materialien (GaAs, GaN und InN mit Abschneideenergien von 15, 70 bzw. 80 Ry) zwischen 20 und 60 Iterationen für die Minimierung benötigt. Eine weitere Steigerung der Effizienz lässt sich durch Präkonditionierung erreichen, was im folgenden Abschnitt kurz umrissen wird.

4.1.2.3 Präkonditionierung

Der große Vorteil des Konjugierte-Gradienten-Verfahrens, dass nämlich der neue Suchvektor nur durch Konjugation des Gradienten zur letzten Suchrichtung erzeugt wird, macht das Verfahren in der Praxis aber auch anfällig für numerische Rundungsfehler, die sich bei hochdimensionalen Problemen im Laufe der Zeit akkumulieren können. Gegebenenfalls werden dadurch mehr als m Schritte für die Minimierung benötigt. Tatsächlich möchte man bei vielen Anwendungen aber sogar mit deutlich weniger als m Schritten auskommen. Aus diesem Grund wird das Verfahren im Allgemeinen mit einer so genannten *Präkonditionierung* betrieben. Ziel dabei ist, durch eine Transformation von (4.13) und (4.15) die Konditionierung κ (vgl. Abschnitt 4.1.2.1) des Problems zu verbessern. Mit einer hermite-

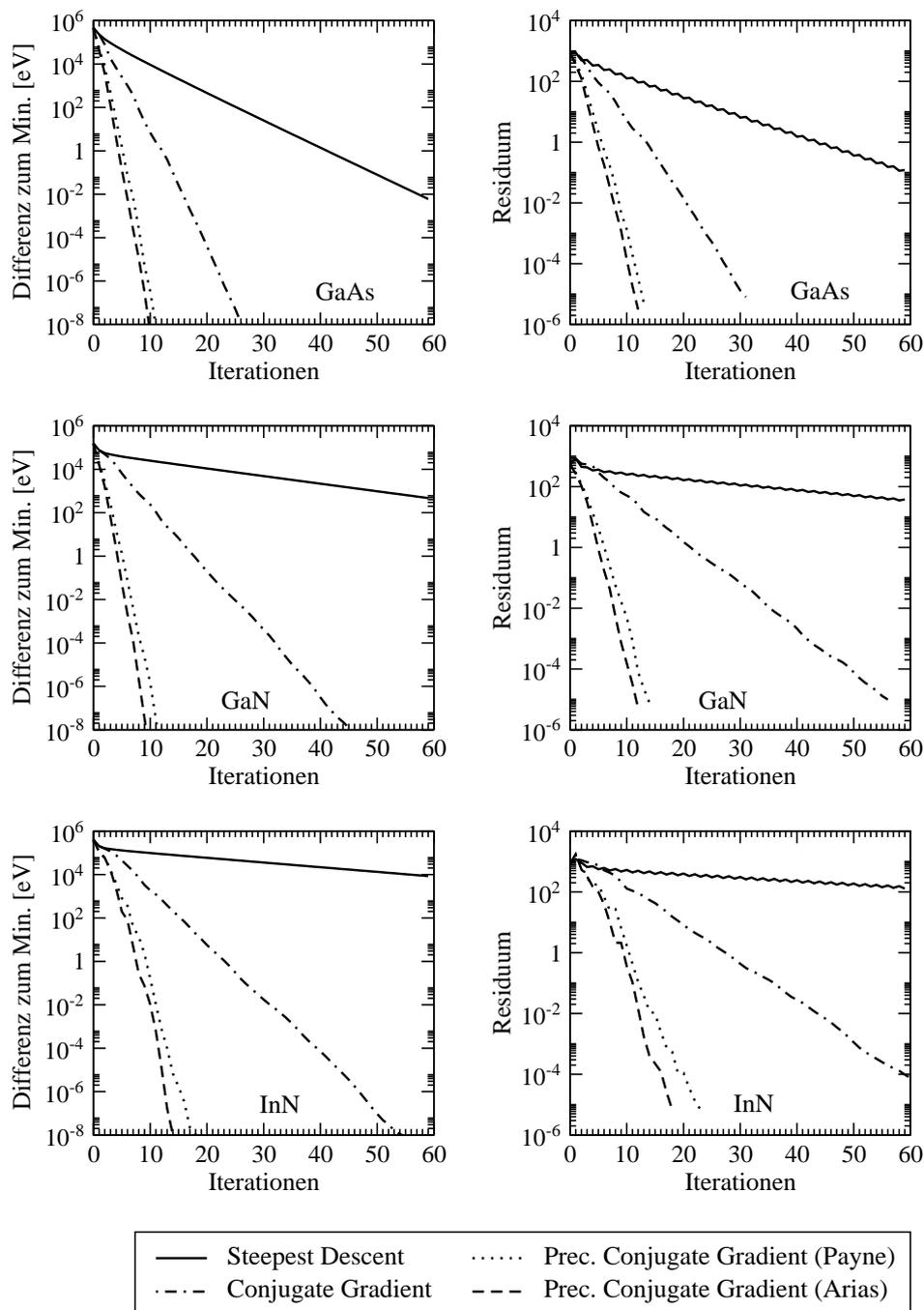
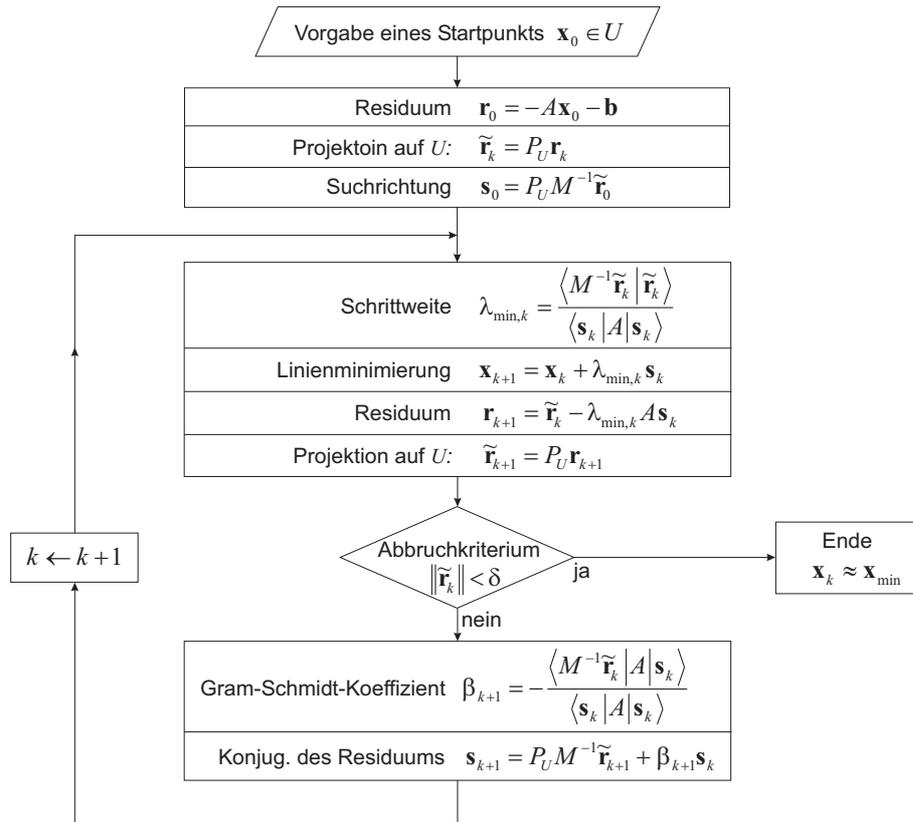


Abbildung 4.5: Iterative Minimierungsmethoden im Vergleich. Die Graphen links zeigen die Abweichung des Energiefunktionals (4.3) von dessen Minimum als Funktion der Iterationsschritte. Rechts daneben ist, ebenfalls über die Anzahl der Iterationen, das Residuum (die linke Seite der Sternheimer-Gleichung (4.1) im Teilraum der unbesetzten Zustände) aufgetragen. Die Abschneideenergie betrug 15 Ry für GaAs, 70 Ry für InN und 80 Ry für InN, und Startpunkt für die Korrekturfunktionen war 0. In allen Fällen ergibt sich die gleiche Reihenfolge bei der Qualität der Methoden: *Steepest Descent* zeigt ein sehr langsames Konvergenzverhalten. Eine wesentliche Verbesserung erreicht man mit dem *Conjugate-Gradient*-Verfahren, dessen Effizienz sich durch geeignete Präkonditionierungen nochmals deutlich steigern lässt.



Flussdiagramm 4.3: Die Methode der Konjugierten Gradienten mit Prädiktionierung. M^{-1} ist eine Näherung für die Inverse von A . Im Algorithmus wird nur die Anwendung auf (die Projektion des) Residuums benötigt. Wie in den Flussdiagrammen 4.1 und 4.2 sind Projektionen P_U eingefügt, um die Minimierung auf einen Unterraum U zu beschränken.

schen, positiv definiten Matrix M gilt⁷

$$A \mathbf{x} + \mathbf{b} = 0 \iff M^{-1} A \mathbf{x} + M^{-1} \mathbf{b} = 0. \quad (4.24)$$

Dabei soll M eine geeignete „Näherung“ von A mit der Eigenschaft $\kappa(M^{-1}A) \ll \kappa(A)$ sein, die sich aber deutlich leichter als A selbst invertieren lässt, z.B. die Matrix der Diagonalelemente von A . Die Schwierigkeit besteht darin, dass die linke Seite dieser Gleichung im Allgemeinen *nicht* als Ableitung eines Funktionals analog zu (4.13) dargestellt werden kann, da $M^{-1}A$ nicht notwendigerweise hermitesch oder positiv definit ist. Als hermitesche und positiv definite Matrix kann jedoch M in der Form $M = EE^*$ faktorisiert werden, wobei sich E z.B. aus einer Cholesky-Zerlegung [129] ergibt. Nun haben $M^{-1}A$ und $E^{-1}A(E^*)^{-1}$ die gleichen Eigenwerte [131] und sind damit gleich gut konditioniert. Da aber $E^{-1}A(E^*)^{-1}$ hermitesch und positiv definit ist, benutzt man statt (4.24) die Transformation

$$A \mathbf{x} + \mathbf{b} = 0 \iff E^{-1}A(E^*)^{-1} \hat{\mathbf{x}} + E^{-1}\mathbf{b} = 0 \quad \text{mit} \quad \hat{\mathbf{x}} := E^* \mathbf{x}. \quad (4.25)$$

Die Lösung dieser Gleichung ist das Minimum der quadratischen Form

$$\tilde{F}(\mathbf{x}) := \frac{1}{2} \langle \hat{\mathbf{x}} | E^{-1}A(E^*)^{-1} | \hat{\mathbf{x}} \rangle + \frac{1}{2} \langle \hat{\mathbf{x}} | E^{-1}\mathbf{b} \rangle + \frac{1}{2} \langle E^{-1}\mathbf{b} | \hat{\mathbf{x}} \rangle, \quad (4.26)$$

⁷Eine positiv definite Matrix ist invertierbar, da alle ihre Eigenwerte positiv sind. Invertierbare Matrizen wiederum sind bijektive Abbildungen. Damit folgt, dass (4.24) eine Äquivalenzumformung ist.

deren Isoflächen — anschaulich gesprochen — „kugelförmiger“ als die des ursprünglichen Funktionals (4.13) sind. Die Anwendung von *Steepest Descent* oder *Conjugate Gradient* auf \tilde{F} konvergiert folglich schneller als die auf F . In der Praxis muss letztlich weder die Invertierung von M noch die Faktorisierung $M = EE^*$ explizit ausgeführt werden, da alle auftretenden Ausdrücke so umgestellt werden können, dass nur $M^{-1}\mathbf{r}_k$ benötigt wird [131]. Es muss folglich noch nicht einmal M selbst bekannt sein, sondern nur die Anwendung seiner Inversen auf das Residuum. Der Ablauf des Algorithmus ist in Flussdiagramm 4.3 demonstriert.

In dieser Arbeit wurden zwei Präkonditionierungen getestet, die ursprünglich für die Diagonalisierung der Kohn-Sham-Gleichung (1.52) in Ebene-Wellen-Darstellung mit Hilfe des Konjugierte-Gradienten-Verfahrens entwickelt wurden. Beim Ansatz von Teter, Payne und Allan [130, 132] ist M^{-1} im reziproken Raum durch eine Diagonalmatrix gegeben,

$$\begin{aligned} \langle \mathbf{G} + \mathbf{k} | M^{-1} | \mathbf{G}' + \mathbf{k} \rangle &:= \\ &:= \delta_{\mathbf{G}\mathbf{G}'} \frac{27 + 18x + 12x^2 + 8x^3}{27 + 18x + 12x^2 + 8x^3 + 16x^4} \quad \text{mit} \quad x := \frac{1}{2} \frac{|\mathbf{G} + \mathbf{k}|^2}{E_{\text{kin}}^{n\mathbf{k}}}, \end{aligned} \quad (4.27)$$

wobei $E_{\text{kin}}^{n\mathbf{k}} := \langle \psi_{n\mathbf{k}}^0 | \frac{1}{2} \nabla^2 | \psi_{n\mathbf{k}}^0 \rangle$ die kinetische Energie des ungestörten Zustandes $\psi_{n\mathbf{k}}^0$ ist. Für $x \rightarrow 0$ strebt dieser Ausdruck gegen 1, so dass sich M^{-1} also im Bereich kleiner Wellenvektoren annähernd wie die Einheitsmatrix verhält. Für große x hingegen gehen die Diagonalelemente asymptotisch in $1/(2x)$ über, d.h., sie werden proportional zu $|\mathbf{G} + \mathbf{k}|^{-2}$. Die ursprüngliche Intention dabei ist, die in der reziproken Kohn-Sham-Gleichung (3.15) für hohe Frequenzen gegebene Dominanz der kinetischen Energie wieder auszugleichen. Empirisch zeigt sich, dass diese Vorgehensweise auch bei der Lösung der Sternheimer-Gleichung (4.1) eine erhebliche Steigerung der Effizienz gegenüber dem Konjugierte-Gradienten-Verfahren ohne Präkonditionierung ermöglicht, obwohl der Störoperator H_1 im Unterschied zu H_0 keinen Beitrag der kinetischen Energie enthält. Bei den untersuchten Systemen konnte die Anzahl der benötigten Iterationen so um mehr als 60% reduziert werden, vgl. Abb. 4.5. Der Vorteil ist dabei umso größer, je höher die verwendete Abschneideenergie ist.

Als geringfügig schneller als (4.27) erweist sich der Präkonditionierer von Ismail-Beigi und Arias [133]. Die Idee ist dabei ganz ähnlich wie eben; wieder ist M^{-1} im reziproken Raum durch eine Diagonalmatrix gegeben, deren Diagonalelemente für kleine Wellenvektoren näherungsweise 1, für große jedoch proportional zu $|\mathbf{G} + \mathbf{k}|^{-2}$ sind:

$$\langle \mathbf{G} + \mathbf{k} | M^{-1} | \mathbf{G}' + \mathbf{k} \rangle := \delta_{\mathbf{G}\mathbf{G}'} \frac{p(x)}{1 + x p(x)} \quad \text{mit} \quad x := \frac{|\mathbf{G} + \mathbf{k}|^2}{E_{\text{kin}}/n_{\text{max}}}, \quad (4.28)$$

wobei

$$E_{\text{kin}} = \sum_{n,\mathbf{k}}^{\text{occ.}} \langle \psi_{n\mathbf{k}}^0 | \frac{1}{2} \nabla^2 | \psi_{n\mathbf{k}}^0 \rangle$$

die kinetische Energie des ungestörten Systems,

$$p(x) := 1 + x (1 + x))))))$$

ein Polynom 8. Grades und n_{max} die Anzahl der besetzten Bänder ist. Der Vorteil dieses Präkonditionierers gegenüber dem von Teter *et al.* beträgt bei den hier untersuchten Materialien allerdings nur sehr wenige Iterationsschritte (1 bis 4), vgl. Abb. 4.5. Mit beiden Methoden lässt sich das Energiefunktional (4.3) in 10 bis 20 Iterationsschritten minimieren. Für das variationelle EXX-Verfahren ist es daher nicht entscheidend, ob (4.27) oder (4.28) eingesetzt wird.

4.2 Induzierte Ladungsdichte

Im letzten Abschnitt wurde detailliert die Lösung der Sternheimer-Gleichung, also die Berechnung der linearen Korrekturfunktionen ψ_i^1 , mit Hilfe schneller, iterativer Minimierungsverfahren besprochen. Der nächste Schritt des variationellen EXX-Verfahrens besteht nun in der Berechnung der induzierten Ladungsdichte $\Delta\rho$ gemäß Gleichung (2.4). Zur numerischen Auswertung wird sie dabei zweckmäßigerweise in der Form

$$\Delta\rho := 2 \operatorname{Re} \sum_i^{\text{occ}} (\psi_i^0)^* \psi_i^1 \quad (4.29)$$

geschrieben. Wird eine Ebene-Wellen-Basis verwendet und die Brillouin-Zonen-Integration durch eine Summe über die \mathbf{k} -Punkte eines Monkhorst-Pack-Gitters genähert, können ganz analog zu der in Abschnitt 3.2 für die Ladungsdichte ρ beschriebenen Vorgehensweise auch hier die im System vorhandenen Punktsymmetrien zur Reduktion des \mathbf{k} -Punkt-Satzes ausgenutzt werden. Die Sternheimer-Gleichung (4.12) wird also nur an den N_{red} Punkten eines reduzierten \mathbf{k} -Punkt-Satzes ausgewertet, und mit den Korrekturfunktionen wird anschließend

$$\Delta\rho_{\text{unsym}} := 2 \operatorname{Re} \sum_n^{\text{occ}} \sum_{s=1}^{N_{\text{red}}} w_{\mathbf{k}_s} (\psi_{n\mathbf{k}_s}^0)^* \psi_{n\mathbf{k}_s}^1 \quad (4.30)$$

berechnet. Dabei sollen die $w_{\mathbf{k}_s}$ geeignete Wichtungen der \mathbf{k}_s sein. Bezeichnet weiter N_{PS} die Anzahl der Punkttransformationen S_i im System, ist die induzierte Dichte in dieser Basis durch

$$\Delta\rho_{\text{sym}}(\mathbf{r}) := \frac{1}{N_{\text{PS}}} \sum_{i=1}^{N_{\text{PS}}} \Delta\rho_{\text{unsym}}(S_i \mathbf{r}) \quad (4.31)$$

gegeben. Gemäß dem in Abschnitt 2.1 formulierten Variationsprinzip (2.6) ist es nun das Ziel, diese Größe durch Variieren des lokalen Austauschpotentials zum Verschwinden zu bringen. Dies gelingt mit den im folgenden Abschnitt beschriebenen „Mixing“-Verfahren. Die (nicht-verschwindende) induzierte Ladungsdichte, die sich aus einem nur genäherten Austauschpotential wie LDA ergibt, kann dabei physikalisch sehr anschaulich interpretiert werden, vgl. dazu Abschnitt 6.

4.3 Minimierung der induzierten Ladungsdichte

Der letzte Schritt im variationellen EXX-Verfahren ist die numerische Auswertung des Variationsprinzips (2.6) von Abschnitt 2.1, also die Minimierung der (Norm der) induzierten Ladungsdichte in Abhängigkeit des lokalen Austauschpotentials:

$$\|\Delta\rho[v_x]\| \longrightarrow 0 \quad \text{für} \quad v_x \longrightarrow v_{\text{EXX}}, \quad (4.32)$$

wobei v_{EXX} das exakte Austauschpotential bezeichnen soll. Die Minimierung kann dabei auf zwei verschiedene Arten geschehen: (i) Bei festgehaltenen Kohn-Sham-Funktionen ψ_i^0 und -energien ε_i^0 oder aber (ii) bei expliziter Berücksichtigung der v_x -Abhängigkeit von ψ_i^0 und ε_i^0 . Beide Möglichkeiten werden in dem in dieser Arbeit entwickelten Algorithmus verwendet, haben aber verschiedene physikalische Bedeutungen. Um den Unterschied zu verdeutlichen, soll die Minimierung zunächst kurz theoretisch analysiert werden, im Anschluss daran werden dann geeignete numerische Methoden erörtert.

Zu Punkt (i): Der Übersichtlichkeit halber sei die induzierte Ladungsdichte für die folgenden Betrachtungen — natürlich nicht in der Implementierung — in der knappen Form (D.12) aus Anhang D ausgedrückt:

$$\begin{aligned}\Delta\rho[v_x] &= \hat{X}_0[\{\psi_i^0\}, \{\varepsilon_i^0\}] (V_{\text{Fock}}[\{\psi_i^0\}] - v_x) \\ &= \hat{X}_0[\{\psi_i^0\}, \{\varepsilon_i^0\}] V_{\text{Fock}}[\{\psi_i^0\}] - \hat{\chi}_0[\{\psi_i^0\}, \{\varepsilon_i^0\}] v_x.\end{aligned}\quad (4.33)$$

Hier ist \hat{X}_0 der auf nicht-lokale Störungen erweiterte⁸ lineare Antwortoperator (D.5), und ferner seien ψ_i^0 und ε_i^0 Eigenfunktionen und -Eigenwerte eines Kohn-Sham-Systems mit einem effektiven Potential v_{eff} . Dass sich v_{eff} — und damit eben auch ψ_i^0 und ε_i^0 — bei Änderung von v_x ebenfalls ändert, sei dabei *nicht* berücksichtigt. $\Delta\rho$ ist dann eine lineare (genauer: affine) Funktion von v_x , und die Aufgabe, $\Delta\rho[v_x]$ verschwinden zu lassen, läuft damit auf die Lösung eines linearen Gleichungssystems (der Inversion von χ_0) heraus. Dies steht in Einklang damit, dass nach der Görling-Levy-Störungstheorie (vgl. die Abschnitte 1.5.4 und 1.6.3) die Austauschwechselwirkung durch eine Taylorentwicklung in 1. Ordnung gegeben ist. Löst man $\Delta\rho[v_x] = 0$ durch Inversion von χ_0 , ergibt sich ausgeschrieben gerade die Standard-EXX-Gleichung (1.90), d.h., *die Minimierung von (4.33) liefert das exakte Austauschpotential $v_{\text{EXX}}[\rho]$ der zu den ψ_i^0 gehörenden Dichte ρ .*

Interessant ist, dass damit die Bestimmungsgleichung für v_{EXX} eine formal analoge Gestalt aufweist wie die Sternheimer-Gleichung; wie diese könnte man sie also mit den in Abschnitt 4.1.2 besprochenen Minimierungsverfahren (z.B. *Steepest Descent*) lösen: es ist in (4.13) nur $A = -\chi_0$ und $\mathbf{b} = \hat{X}_0 V_{\text{Fock}}$ zu setzen. Das zu minimierende Funktional lautet dann:

$$S[v_x] := \frac{1}{2} \langle v_x | (-\hat{\chi}_0) | v_x \rangle + \langle v_x | \hat{X}_0 V_{\text{Fock}} \rangle \quad (4.34)$$

(man beachte, dass χ_0 negativ semi-definit ist [134, 148], und es sich hier um ein rein reelles Problem handelt). Dieser Zusammenhang zwischen der Sternheimer-Gleichung und der Minimierung von $\Delta\rho$ wurde bereits in Abschnitt 2.3 angedeutet. Im nächsten Abschnitt wird die Minimierung von $\Delta\rho$ mit Hilfe von Selbstkonsistenzverfahren vorgestellt. Dabei wird sich die einfachste Variante dieser Methoden — das „einfache Mixen“ — in der Tat als Anwendung des *Steepest-Descent*-Verfahrens auf (4.34) entpuppen.

Zunächst aber noch zu Punkt (ii): In Gleichung (4.33) werde nun berücksichtigt, dass \hat{X}_0 und V_{Fock} implizit — über die Kohn-Sham-Wellenfunktionen und -Eigenwerte — vom effektiven Potential und damit von v_x abhängen:

$$\Delta\rho[v_x] = \hat{X}_0[\{\psi_i^0[v_x]\}, \{\varepsilon_i^0[v_x]\}] (V_{\text{Fock}}[\{\psi_i^0[v_x]\}] - v_x). \quad (4.35)$$

Dabei sind $\psi_i^0[v_x]$ und $\varepsilon_i^0[v_x]$ als Lösungen des Eigenwertproblems

$$\left[-\frac{1}{2} \nabla^2 + v_{\text{ext}}(\mathbf{r}) + v_{\text{H}}[\rho](\mathbf{r}) + v_x(\mathbf{r}) + v_{\text{c}}[\rho](\mathbf{r}) \right] \psi_i(\mathbf{r}) = \varepsilon_i \psi_i(\mathbf{r}) \quad (4.36)$$

gegeben. Wie die Kohn-Sham-Gleichung (1.52) soll diese Gleichung selbstkonsistent zusammen mit der Formel für die Ladungsdichte (1.53) gelöst werden, im Unterschied zu (1.52) ist hier das Austauschpotential aber *keine* Funktion der Dichte. Berechnet man also die induzierte Dichte gemäß (4.35) für ein spezielles v_x , wird zunächst (4.36) gelöst, und

⁸Wie in Anhang D sei bei Anwendung auf ein lokales Potential v_{lok} wieder $\hat{X}_0 v_{\text{lok}} := \hat{\chi}_0 v_{\text{lok}}$ mit dem „normalen“ Antwortoperator $\hat{\chi}_0$ definiert, um die Schreibweise nicht zu sehr aufzublähen.

anschließend werden mit den daraus resultierenden Eigenfunktionen $\psi_i^0[v_x]$ und Eigenwerten $\varepsilon_i^0[v_x]$ der Fock-Operator sowie der (erweiterte) Antwortoperator (D.5) berechnet, woraus sich schließlich $\Delta\rho[v_x]$ ergibt.

Die selbstkonsistente Lösung von (4.36) ist aber gleichbedeutend mit der Minimierung des Energiefunktionals $E[\rho] := \int v_{\text{ext}}(\mathbf{r})\rho(\mathbf{r}) d^3r + T_s[\rho] + E_H[\rho] + E_x[\rho] + E_c[\rho]$ mit $E_x[\rho] := \int v_x(\mathbf{r})\rho(\mathbf{r}) d^3r$, was dem Ausdruck (1.54) für die Gesamtenergie des realen Systems bei festgehaltenem v_x entspricht. Im Gegensatz zu der in (4.33) definierten Induktionsdichte führt die Definition (4.35) in (4.32) damit auf ein v_{EXX} , das zwei Variationsprinzipien gemeinsam erfüllt: Zum einen gilt $\|\Delta\rho[v_{\text{EXX}}]\| = 0$, und zum anderen minimiert das aus den $\psi_i^0[v_{\text{EXX}}]$ gebildete $\rho_{v_{\text{EXX}}}$ die Energie $E[\rho]$. Da in Punkt (i) Wellenfunktionen und Ein-Teilchen-Energien eines beliebigen effektiven Potentials vorausgesetzt wurden, können dort auch die so konstruierten $\psi_i^0[v_{\text{EXX}}]$ und $\varepsilon_i^0[v_{\text{EXX}}]$ eingesetzt werden. Aus der in (i) kursiv hervorgehobenen Aussage folgt dann, dass v_{EXX} das exakte Austauschpotential der zu den $\psi_i^0[v_{\text{EXX}}]$ gehörenden Dichte $\rho_{v_{\text{EXX}}}$ ist. Diese ergibt sich aber aus dem effektiven Potential $v_{\text{eff}}[\rho_{v_{\text{EXX}}}] := v_{\text{ext}} + v_H[\rho_{v_{\text{EXX}}}] + v_{\text{EXX}} + v_c[\rho_{v_{\text{EXX}}}]$, denn (4.36) ist ja selbstkonsistent gelöst. Damit folgt sofort $v_{\text{eff}}[\rho_{v_{\text{EXX}}}] = v_{\text{ext}} + v_H[\rho_{v_{\text{EXX}}}] + v_{\text{EXX}}[\rho_{v_{\text{EXX}}}] + v_c[\rho_{v_{\text{EXX}}}]$, d.h., v_{EXX} und $\rho_{v_{\text{EXX}}}$ erfüllen nicht nur das Eigenwertproblem (4.36), sondern auch die Kohn-Sham-Gleichung (1.52). Äquivalent dazu: $\rho_{v_{\text{EXX}}}$ minimiert nicht nur das oben definierte $E[\rho]$, sondern auch die Gesamtenergie des realen Systems (1.54), und es gilt $\rho_{v_{\text{EXX}}} = \rho_0$ mit der Grundzustandsladungsdichte ρ_0 des Systems. *Die Minimierung von (4.35) beinhaltet damit neben der Berechnung des exakten Austauschpotentials $v_{\text{EXX}}[\rho_0]$ der Grundzustandsdichte ρ_0 auch zugleich die Lösung des gesamten Kohn-Sham-EXX-Problems!*

Daher ist (4.35) — im Gegensatz zu (4.33) — auch keine lineare (affine) Funktion in v_x mehr. Es sei darauf hingewiesen, dass aber auch die Bestimmung des exakten Austauschpotentials zu einer vorgegebenen Ladungsdichte ρ nach (i) nur dadurch als lineares Problem aufgefasst werden konnte, weil das zu ρ gehörende (und unter der Bedingung der v -Darstellbarkeit eindeutig bestimmte) effektive Potential und damit auch die entsprechenden Eigenfunktionen und -werte schon als bekannt vorausgesetzt wurden. Dies wird auch in der Göring-Levy-Störungstheorie getan, vgl. (1.101). Von ρ ausgehend ist die Berechnung von $v_{\text{EXX}}[\rho]$ hingegen ein nicht-lineares Problem — die Nichtlinearität liegt dabei jedoch einzig in der Zuordnung $\rho \longleftrightarrow v_{\text{eff}} \longleftrightarrow \{\psi_i^0\}, \{\varepsilon_i^0\}$, die für sich alleine genommen nichts mit der Austauschwechselwirkung zu tun hat. Ergänzend sei noch bemerkt, dass sich hierin wiederum der im Sinne von Abschnitt 1.4 „universelle“ Charakter des exakten Austauschpotentials manifestiert, da in diese Zuordnung das äußere Potential nicht explizit⁹ einfließt.

Mit dem Vorgegangenen ist klar, dass man bei praktischen Anwendungen in der Regel an der Minimierung der induzierten Ladungsdichte nach (ii), also bei Berücksichtigung auch der impliziten v_x -Abhängigkeiten, interessiert sein wird. Mit iterativen Verfahren wie dem nachfolgend vorgestellten *Mixing* ließe sich dies auch ohne Weiteres bewerkstelligen. Später wird sich jedoch bei der Optimierung des variationellen EXX-Verfahrens (Abschnitt 4.5.1) herausstellen, dass eine Kombination der Minimierungen nach (i) und (ii) eine technisch weitaus bessere Strategie ist. Der Algorithmus besteht dann aus einer inneren Schleife, in der gemäß (i) vorgegangen wird, und einer äußeren, die (ii) realisiert. Da sich das *Mixing* dabei vorwiegend in der inneren Schleife abspielt, wird es im Folgenden vor allem mit Blick auf Punkt (i) diskutiert.

⁹Mit den Sätzen von Hohenberg und Kohn ist natürlich ρ eine Funktion von v_{ext} , und damit ist das Austauschpotential ebenso wie das universelle Funktional $F[\rho]$ in (1.37) eine implizite Funktion von v_{ext} .

4.3.1 Mixing, Selbstkonsistenz und Minimierung

Mixing-Verfahren sind zentral für die Behandlung von Problemen, die „selbstkonsistent“ gelöst werden müssen, d.h., bei denen die Größen, aus denen die Lösung ermittelt werden soll, bereits ihrerseits von der (im Allgemeinen noch nicht bekannten) Lösung abhängen. Beispiele sind die Hartree-Fock-Gleichung (1.24) und die Kohn-Sham-Gleichung (1.52): in beiden Fällen müssen die Ein-Teilchen-Wellenfunktionen bekannt sein, damit aus ihnen die Eigenwertgleichung erst formuliert werden kann. Da bei Kohn-Sham die Wellenfunktionen nur zusammengefasst zur Ladungsdichte benötigt werden, lässt sich die Selbstkonsistenz hier in der Form

$$\rho_* = F[\rho_*] \quad (4.37)$$

schreiben [134, 135], wobei F eine nicht-lineare Funktion ist, die die Lösung des Eigenwertproblems (1.52) in Abhängigkeit einer Eingangsdichte ρ_{in} und die anschließende Berechnung der daraus resultierenden Ausgangsdichte $\rho_{\text{out}} = F[\rho_{\text{in}}]$ über (1.53) beinhaltet.¹⁰ Mathematisch ist ρ_* ein Fixpunkt von F , und gemäß dem Fixpunktsatz von Banach (vgl. z.B. Ref. 123 oder 124) erhält man ρ_* durch (unendlich oftmaliges) rekursives Anwenden von F auf ein beliebiges ρ_{in} , wenn F „kontrahierend“¹¹ ist. Diese Voraussetzung ist aber im Allgemeinen nicht gegeben, und häufig divergiert daher die Folge der so erzeugten Dichten in der Praxis. Ein einfacher Ansatz, dennoch das selbstkonsistente ρ_* durch Rekursion finden zu können, ist nun, in jeder Iteration Eingangs- und Ausgangsdichte miteinander zu „mischen“ und den Rekursionsprozess auf diese Weise zu dämpfen. Mit dem Residuum

$$R[\rho_{\text{in}}] := \rho_{\text{out}} - \rho_{\text{in}} = F[\rho_{\text{in}}] - \rho_{\text{in}} \quad (4.38)$$

erhält man etwa beim „einfachen Mixen“ die neue Eingangsladungsdichte jeweils aus

$$\rho_{\text{in}}^{m+1} = (1 - \lambda)\rho_{\text{in}}^m + \lambda\rho_{\text{out}}^m = \rho_{\text{in}}^m + \lambda R[\rho_{\text{in}}^m], \quad (4.39)$$

wobei für λ typischerweise Werte im Bereich zwischen 0.1 und 0.5, oft aber auch wesentlich kleiner, angesetzt werden [134, 137, 148].

Durch das Residuum kommt nun ein Bezug zur Minimierung ins Spiel. Denn aus (4.37) und (4.38) folgt, dass das Residuum für die selbstkonsistente Dichte ρ_* verschwindet; die Suche nach ρ_* ist also äquivalent zu der Aufgabe, $\|R[\rho]\|$ zu minimieren. Dies ermöglicht nun, Selbstkonsistenz-Methoden im Rahmen des variationellen EXX-Verfahrens auch auf das Minimierungsproblem (4.32) anzuwenden. Indem die (negative) induzierte Ladungsdichte als Residuum betrachtet wird,

$$R[v_{\text{x}}] := -\Delta\rho[v_{\text{x}}], \quad (4.40)$$

lässt sich das einfache *Mixing* dabei analog zu (4.39) durch

$$v_{\text{x}}^{m+1} := v_{\text{x}}^m + \lambda R[v_{\text{x}}^m] = v_{\text{x}}^m - \lambda\Delta\rho[v_{\text{x}}^m] \quad (4.41)$$

mit $\lambda > 0$ realisieren. Hier ist allerdings zu beachten, dass im Gegensatz zur Vorgehensweise bei der Kohn-Sham- oder der Hartree-Fock-Gleichung zwei unterschiedliche physikalische Größen — Potential und Ladungsdichte — miteinander „gemischt“ werden.

¹⁰Alternativ lässt sich dies auch über das effektive Kohn-Sham-Potential formulieren: $\tilde{F}[v_{\text{x}}^{\text{eff}}] = v_{\text{x}}^{\text{eff}}$, wobei \tilde{F} für ein gegebenes $v_{\text{in}}^{\text{eff}}$ zunächst die Ladungsdichte und damit anschließend $v_{\text{out}}^{\text{eff}}$ berechnet.

¹¹Ein in einem normierten Raum X definierter Operator F ist „kontrahierend“, wenn für $x, y \in X$ mit einem $L < 1$ gilt: $\|F(x) - F(y)\| \leq L\|x - y\|$.

Benutzt man in (4.41) die Definition (4.33) von $\Delta\rho[v_x]$ — erfolgt also keine Aktualisierung des Eigenwertproblems (4.36) in jedem Schritt —, so fungiert der skalare Parameter λ dabei als einfache Näherung für die Inverse der linearen Antwort. Dies erkennt man unter Beachtung von $\Delta\rho[v_{\text{EXX}}] = 0$, indem man (4.33) für beliebiges v_x sowie für das exakte v_{EXX} betrachtet und anschließend die beiden so entstandenen Gleichungen voneinander subtrahiert:

$$v_{\text{EXX}} = v_x + \hat{\chi}_0^{-1} \Delta\rho[v_x]. \quad (4.42)$$

Gleichung (4.41) ist also eine einfache Approximation von (4.42). Damit erklärt sich auch das negative Vorzeichen von λ in der rechten Seite von (4.41) bzw. von $\Delta\rho$ bei der Definition des Residuums (4.40): $\hat{\chi}_0$ ist ein negativ definiten Operator [134, 148] und $\hat{\chi}_0^{-1}$ folglich ebenfalls.¹² Abbildung 4.6 zeigt die Minimierung der induzierten Dichte gemäß (4.41) in Abhängigkeit von der Iteration bei GaAs.

Zwischen *Mixing*-Verfahren und der in Abschnitt 4.1.2 beschriebenen Minimierung eines Funktionals bestehen einige Analogien, die hier besonders deutlich werden. Ist nämlich die 1. Variationsableitung (diskret: Jakobi-Matrix) des Residuums R ein selbstadjungierter Operator (diskret: eine hermitesche Matrix), existiert ein reellwertiges Funktional S , dessen 1. Variationsableitung (diskret: Gradient ∇S) durch $-R$ gegeben ist. Die stationären Punkte von S sind dann gerade die gesuchten Nullstellen von R — ganz ähnlich, wie sich die Lösungen von (4.15) als stationäre Punkte von (4.13) darstellen lassen. Für $R = -\Delta\rho$ mit der induzierten Ladungsdichte (4.33) ist dies tatsächlich der Fall: es gilt $\delta R[v_x]/\delta v_x = \chi_0$, und χ_0 ist hermitesch. Insbesondere ergibt sich eine Übereinstimmung von einfachem *Mixing* und einfachem *Steepest Descent*: die Rekursionsformeln (4.18) und (4.41) beider Verfahren besitzen die gleiche Gestalt. In der Tat ist das einfache Mixen (4.41) von Potential und induzierter Dichte nichts anderes als die Anwendung des *Steepest-Descent*-Verfahrens auf das Funktional $S[v_x]$ von (4.34): es ist $\delta S[v_x]/\delta v_x = \Delta\rho[v_x] = -R[v_x]$, und die Rekursionsformel (4.18) für *Steepest Descent* liefert unmittelbar (4.41) mit $\mathbf{x}_i = v_x^i$ und $\mathbf{r}_i = R[v_x^i]$. Der einzige Unterschied zur Vorgehensweise von Abschnitt 4.1.2.1 besteht darin, das Residuum nicht direkt nach (4.33) zu bestimmen, sondern — um die direkte Berechnung von χ_0 zu vermeiden — indirekt über (4.29). Dies ist auch der Grund, weshalb im variationellen EXX-Verfahren die Minimierung von $\Delta\rho$ überhaupt durch *Mixing* erfolgt und nicht beispielsweise mit *Conjugate Gradient* — dazu wäre eine explizite Kenntnis von χ_0 nötig, die bei *Mixing*-Methoden wie dem nachfolgend vorgestellten Pulay-Verfahren nicht notwendig ist.

Es sei noch darauf hingewiesen, dass nicht immer eine solche „Stammfunktion“ S konstruiert werden kann. Dies ist etwa der Fall für das Residuum (4.38) des Ladungsdichte-*Mixings* beim Lösen der Kohn-Sham-Gleichung, denn seine Ableitung $J[\rho] := -\delta R[\rho]/\delta\rho$ ist nicht hermitesch (vgl. Fußnote auf Seite 76). Das einfache *Mixing* (4.39) der Dichte kann also nicht als *Steepest-Descent*-Verfahren interpretiert werden. Dennoch gibt es starke Ähnlichkeiten: Zum einen sind auch hier die Rekursionsformeln beider Verfahren, (4.39) und (4.18), von gleicher Gestalt. Zum anderen besitzt J nur reelle, positive (rechte) Eigenwerte α_i [134, 146, 148], was sich als Verallgemeinerung der positiven Definitheit von A in (4.13) auffassen lässt. Hätten die Eigenwerte unterschiedliche Vorzeichen, so würde das *Mixing* — ähnlich wie die Minimierungsmethoden in Abschnitt 4.1.2 — nicht konvergieren [134]. Schließlich ist auch der in den Referenzen 134 und 135 angegebene Bereich

¹²Wie in Abschnitt 1.6.1 wird hier ein auf die nicht-konstanten Funktionen beschränkter Definitions- (und Werte-)bereich von $\hat{\chi}_0$ angenommen.

für λ , für den das einfache Mixing konvergent ist, identisch mit der Konvergenzbedingung (4.19) des einfachen *Steepest-Descent*-Verfahrens.

Die Verbindung von Selbstkonsistenz- und Minimierungsproblemen basiert letztlich darauf, dass erstere über das Residuum immer in der Form einer homogenen Gleichung geschrieben werden können, hier etwa als $R[\rho] = 0$ bzw. $R[v_x] = 0$. Numerisch, also in einer geeigneten endlichen Basisdarstellung, entspricht dies einem System gekoppelter (im Allgemeinen nicht-linearer) Gleichungen. Damit kann eine Vielzahl von in der numerischen Mathematik entwickelten Methoden zur Lösung solcher Systeme auf das Selbstkonsistenzproblem übertragen werden mit dem Ziel, die Konvergenz gegenüber dem einfachen *Mixing* zu steigern. Für das Dichte-*Mixing* im Rahmen der Dichtefunktionaltheorie wurde eine Reihe von Ansätzen entwickelt [137, 138, 139, 140, 141, 144], die darauf beruhen, in aufeinander folgenden Schritten immer bessere Näherungen für die (Inverse der) Jakobi-Matrix von R zu konstruieren. Wie Kresse und Furthmüller [142, 143] ausführen, handelt es sich bei allen diesen Verfahren um verschiedene Realisierungen oder Weiterentwicklungen eines von Broyden [136] vorgeschlagenen Quasi-Newton-Raphson-Schemas. Ausgangspunkt dabei ist eine „Linearisierung“ des Residuums (4.38) an der Stelle der selbstkonsistenten Dichte ρ_* , woraus sich mit der Forderung $R[\rho_*] = 0$ die Gleichung

$$\rho_* = \rho + J^{-1}R[\rho] \quad (4.43)$$

mit der Jakobi-Matrix (der 1. Ableitung) J des Residuums ergibt. Die Struktur dieser Gleichung ist mit der von Gleichung (4.42) zur Berechnung des EXX-Potentials völlig identisch. Alle in der Literatur zur effizienten Lösung der Kohn-Sham-Gleichung vorgeschlagenen Algorithmen können folglich unmittelbar auf die Minimierung von $\Delta\rho$ übertragen werden! Da (4.42) im Gegensatz zu (4.43) eine exakte Gleichung ist, stellt sie sogar einen Idealfall für die Anwendung von Quasi-Newton-Raphson-Verfahren dar. Die verschiedenen Algorithmen unterscheiden sich in der Art und Weise, wie J^{-1} im Laufe des Iterationsprozesses aktualisiert und approximiert wird. In dieser Arbeit wurde die Methode von Pulay [140, 141] verwendet, und zwar in einer von Kresse und Furthmüller [142] zur Verbesserung der numerischen Stabilität leicht modifizierten Fassung. Sie soll nun, in einer für das *Mixing* von Potential und induzierter Dichte adaptierten Version, skizziert werden.

4.3.2 Pulay-Mixing

Während beim einfachen *Mixing* (4.41) in jedem Schritt vom Potential v_x^m der jeweils vorangehenden Iteration ausgegangen wird, ist die Idee des Pulay-Mixers, aus allen (oder einer bestimmten Anzahl der) vorangehenden Iterationen die beste Linearkombination zu benutzen:

$$v_x^{m+1} := v_x^{\text{opt}} - \lambda \Delta\rho[v_x^{\text{opt}}] \quad (4.44)$$

mit

$$v_x^{\text{opt}} = \sum_i \alpha_i v_x^i. \quad (4.45)$$

Die Aufgabe besteht dann in der Konstruktion von v_x^{opt} , d.h., in der Bestimmung der Koeffizienten α_i . Da in der ursprünglichen Formulierung des Verfahrens nicht Potentiale, sondern Ladungsdichten gemischt werden, wird an dieser Stelle die zusätzliche Bedingung

$$\sum_i \alpha_i \stackrel{!}{=} 1 \quad (4.46)$$

eingeführt, um die Ladungserhaltung zu gewährleisten. Hier entspricht sie der Erhaltung der konstanten Komponente ($\mathbf{G} = \mathbf{0}$) im Startpotential v_x^1 . Obwohl nun konstante Verschiebungen des Potentials die Physik eines Systems nicht beeinflussen, wird diese Forderung hier jedoch ebenfalls benötigt; (4.45) soll nämlich analog auch für die induzierten Dichten formuliert werden. Mit (4.33) ergibt sich dann:

$$\begin{aligned} \Delta\rho[v_x^{\text{opt}}] &= \Delta\rho\left[\sum_i \alpha_i v_x^i\right] = \hat{X}_0 V_{\text{Fock}} - \hat{\chi}_0 \left(\sum_i \alpha_i v_x^i\right) = \\ &\stackrel{(4.46)}{=} \sum_i \alpha_i \left(\hat{X}_0 V_{\text{Fock}} - \hat{\chi}_0 v_x^i\right) = \sum_i \alpha_i \Delta\rho[v_x^i]. \end{aligned} \quad (4.47)$$

Die Koeffizienten der „besten“ Linearkombination (4.45) sind nun dadurch definiert, dass mit ihnen die Norm der induzierten Dichte

$$\langle \Delta\rho[v_x^{\text{opt}}] | \Delta\rho[v_x^{\text{opt}}] \rangle \quad (4.48)$$

minimal wird. Daraus und mit der Nebenbedingung (4.46) folgt

$$\alpha_i = \frac{\sum_j A_{ji}^{-1}}{\sum_{kj} A_{kj}^{-1}} \quad \text{mit} \quad A_{ij} := \langle \Delta\rho[v_x^j] | \Delta\rho[v_x^i] \rangle. \quad (4.49)$$

Dies entspricht der originalen Vorgehensweise von Pulay [140, 141]. Während dabei die aufeinander folgenden Iterationen des Potentials und der induzierten Dichte durch (4.45) und (4.49) explizit berücksichtigt werden, geht deren Verlauf in der Variante von Kresse und Furthmüller [142] nur in Form der Differenzen

$$\Delta v_x^i := v_x^{i+1} - v_x^i, \quad \Delta\Delta\rho^i := \Delta\rho[v_x^{i+1}] - \Delta\rho[v_x^i] \quad \text{für} \quad i < m \quad (4.50)$$

ein. Zunächst wird die Nebenbedingung (4.46) benutzt, um die Anzahl der Koeffizienten in (4.45) um Eins zu reduzieren:

$$v_x^{\text{opt}} = v_x^m + \sum_{i=1}^{m-1} \bar{\alpha}_i \Delta v_x^i. \quad (4.51)$$

Für die $\bar{\alpha}_i$ ist dann keine Nebenbedingung mehr nötig, und es gilt

$$\bar{\alpha}_i = - \sum_{j=1}^{m-1} \bar{A}_{ji}^{-1} \langle \Delta\Delta\rho^j | \Delta\rho[v_x^m] \rangle \quad \text{mit} \quad \bar{A}_{ij} := \langle \Delta\Delta\rho^j | \Delta\Delta\rho^i \rangle. \quad (4.52)$$

In der Praxis wird meist nicht der gesamte Verlauf von $i = 1$ an betrachtet, vielmehr werden nur die letzten n Schritte miteinbezogen. Die Summation startet also bei $i = m - n + 1$. Für $n = 1$ geht das Schema dann in den einfachen Mixer (4.41) über. Der Algorithmus ist in Flussdiagramm 4.4 veranschaulicht. Abbildung 4.6 zeigt das Konvergenzverhalten im Vergleich mit dem einfachen Mixer.

Mit der selben Argumentation wie in Ref. 142 und 143 im Falle der Jakobi-Matrix von (4.43) vorgeführt, lässt sich nun auch hier zeigen, wie der Pulay-Mixer sukzessive immer bessere Näherungen für $\hat{\chi}_0^{-1}$ liefert. Einsetzen von (4.51) in (4.44) führt auf

$$v_x^{m+1} = v_x^m + G^m \Delta\rho[v_x^m], \quad (4.53)$$

wobei die Matrix

$$G^m := -\lambda \mathbf{1} + \sum_{k,n=1}^{m-1} \bar{A}_{kn}^{-1} (\lambda |\Delta\Delta\rho^n\rangle - |\Delta v_x^n\rangle) \langle \Delta\Delta\rho^k | \quad (4.54)$$

die Relation

$$|\Delta v_x^i\rangle + G^m |\Delta\Delta\rho^i\rangle = 0 \quad \text{für } i < m \quad (4.55)$$

erfüllt. Wäre G^m das exakte $\hat{\chi}_0^{-1}$, wäre die Beziehung ebenfalls erfüllt, allerdings ohne die Beschränkung $i < m$. Dies ergibt sich durch Subtraktion von Gleichung (4.42) für zwei verschiedene Iterationen. G^m ist folglich die exakte Inversion von $\hat{\chi}_0$ in dem Unterraum, der durch Δv_x^1 bis Δv_x^{m-1} aufgespannt wird. Dies veranschaulicht sehr schön die auf Pulay zurückgehende Bezeichnung *Direct Inversion in the Iterative Subspace* (DIIS) der Methode.

4.3.3 Präkonditionierung

Beim selbstkonsistenten Lösen der Kohn-Sham-Gleichung wird das *Mixing* meist mit Präkonditionierung betrieben. Als einfacher Ansatz für die Inverse der Jakobi-Matrix J von (4.43) wird dafür häufig das Schema von Kerker verwendet, in dem J^{-1} näherungsweise durch die im reziproken Raum diagonale Matrix

$$K(\mathbf{G}, \mathbf{G}') = A \frac{G^2}{G^2 + G_0^2} \delta_{\mathbf{G}\mathbf{G}'} \quad (4.56)$$

mit einem A nahe bei 1 und einem $G_0 \approx 1.5 \text{ \AA}^{-1}$ gegeben ist [142, 143, 145]. Im Rahmen des Thomas-Fermi-Modells [17, 18, 19, 20] lässt sich K als Inverse der dielektrischen Matrix in Jellium interpretieren. Ersetzt man etwa beim einfachen *Mixing* (4.39) den skalaren Parameter λ durch K , erkennt man, dass die Präkonditionierung als eine Art Filter wirkt: kleine \mathbf{G} -Vektoren werden im Selbstkonsistenzzyklus stärker gebremst als stark oszillierende Komponenten, für die bei $A \approx 1$ fast keine Dämpfung mehr erfolgt.

In dieser Arbeit durchgeführte Tests zeigten jedoch, dass das Kerker-Schema für das Mischen von Potential und induzierter Dichte gemäß (4.41) und (4.44) nicht geeignet ist: Der Selbstkonsistenzzyklus konvergiert schlechter als ohne Präkonditionierung oder divergiert sogar. Nun ist K allerdings auch als Näherungsansatz für J^{-1} gedacht, man kann also nicht unbedingt erwarten, dass er für χ_0^{-1} ebenfalls funktioniert. Es sei nun gezeigt, dass aber ein skalarer Ansatz für χ_0^{-1} bereits dem „Niveau“ der Kerker-Präkonditionierung bei J^{-1} entspricht, in dem Sinne, dass aus (4.56) eine skalare Gestalt von χ_0^{-1} folgt. Dafür wird eine Relation zwischen J und χ_0 benötigt. Sie ist im reziproken Raum näherungsweise durch

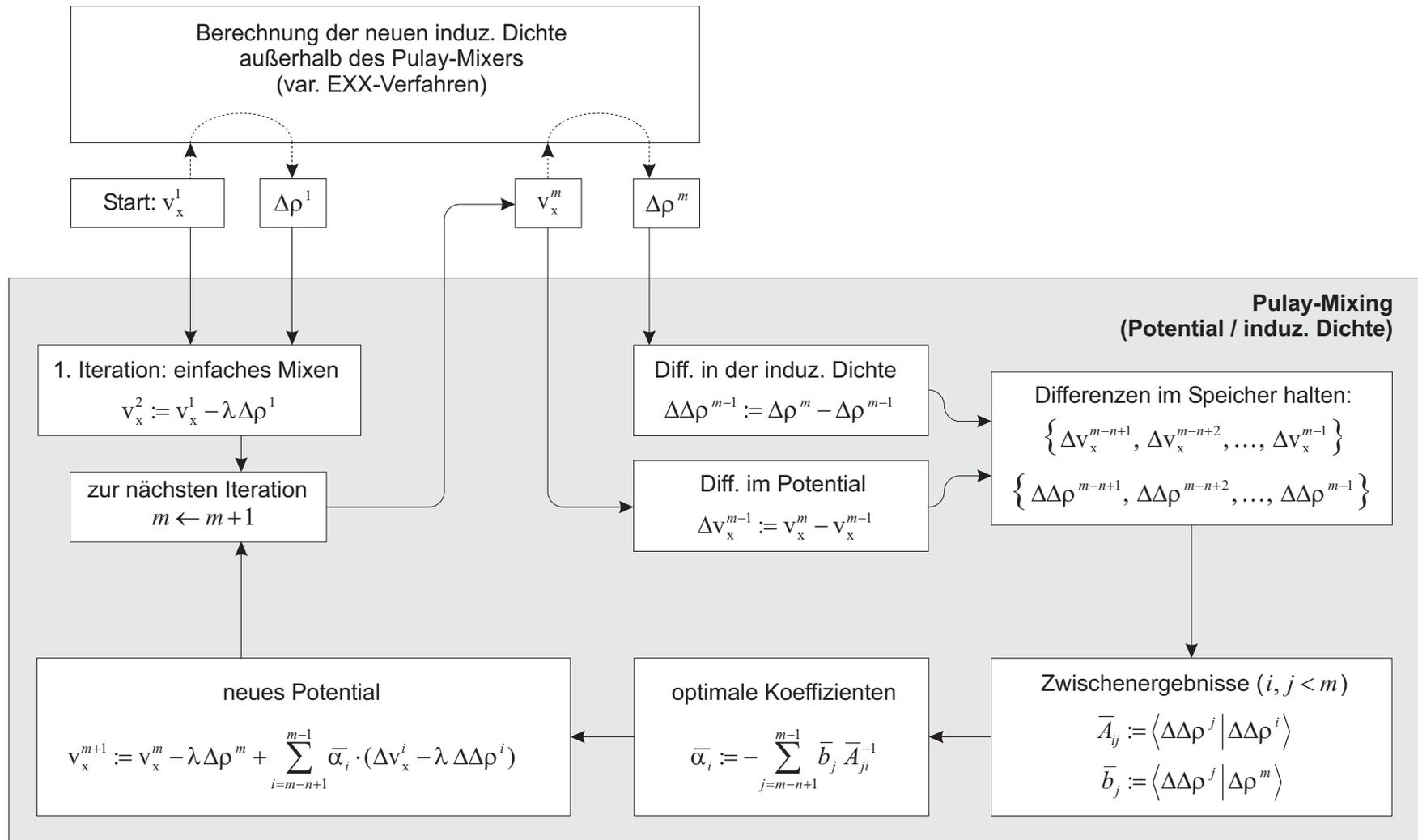
$$J(\mathbf{G}, \mathbf{G}') = \delta_{\mathbf{G}\mathbf{G}'} - \chi_0(\mathbf{G}, \mathbf{G}') v_{\text{Coul}}(\mathbf{G}') \quad (4.57)$$

gegeben, wobei v_{Coul} das Coulomb-Potential ist [139, 143].^{13,14} Benutzt man nun hier für J die diagonale Näherung K^{-1} mit $A = 1$, also

$$\delta_{\mathbf{G}\mathbf{G}'} \frac{G^2 + G_0^2}{G^2} \stackrel{!}{=} \delta_{\mathbf{G}\mathbf{G}'} - \chi_0(\mathbf{G}, \mathbf{G}') v_{\text{Coul}}(\mathbf{G}'), \quad (4.58)$$

¹³Vgl. auch [134, 146, 147, 148]. Dort wird die Beziehung zwischen der dielektrischen Matrix ϵ und der linearen Antwort diskutiert. Analog zu J beim Mischen von Ladungsdichten ist ϵ die Jakobi-Determinante beim Mischen von Potentialen. Die dort angestellten Überlegungen können aber direkt auf die Herleitung von (4.57) übertragen werden, indem man nämlich die Rollen von Ladungsdichte und effektivem Kohn-Sham-Potential vertauscht.

¹⁴An (4.57) wird auch klar, dass J im Allgemeinen keine symmetrische/hermitesche Matrix ist (vgl. Abschnitt 4.3.1), weil das hermitesche χ_0 hier noch mit dem Coulomb-Potential multipliziert wird.



Flussdiagramm 4.4: Pulay-Mixing von Austauschpotential und induzierter Ladungsdichte im variationellen EXX-Verfahren. Der erste Schritt entspricht dem einfachen *Mixing* (4.41), ab dem zweiten wird der Verlauf der bisherigen Ergebnisse bis zu maximal n Schritten rückwärts berücksichtigt. In jeder Iteration wird außer dem aktuellen v_x^m auch das zugehörige $\Delta\rho^m$ (Residuum) benötigt, dessen Berechnung außerhalb des Mixers erfolgt (gestrichelte Linien, vgl. innere Schleife in Flussdiagramm 4.5).

folgt zunächst, dass χ_0 in dieser Darstellung ebenfalls diagonal sein muss. Mit der Fourierentwicklung des Coulomb-Potentials (A.1) gilt dann für die Diagonalelemente

$$\frac{G^2 + G_0^2}{G^2} = 1 - \chi_0(\mathbf{G}, \mathbf{G}) \frac{4\pi}{G^2} \quad (4.59)$$

oder äquivalent dazu

$$\chi_0(\mathbf{G}, \mathbf{G}) = -\frac{G_0^2}{4\pi} = \text{const.} \quad (4.60)$$

In dieser Arbeit wurde daher sowohl beim einfachen als auch beim Pulay-*Mixing* ein skalarer Parameter benutzt, d.h., es wurden (4.41) und (4.44) ohne weitere Modifikationen verwendet. Empirisch zeigte sich, dass dabei Werte zwischen 0.01 und 0.04 eine gute Wahl für λ sind. Einsetzen des oben genannten Wertes für G_0 in atomaren Einheiten liefert $\lambda \approx 0.06$, führt also auf eine vergleichbare Größenordnung.

4.4 Der nicht-lokale Austauschoperator in ebenen Wellen

Notwendig für die Entwicklung eines EXX-Codes sowohl nach der Standard-Methode (Abschnitt 1.6.1) wie auch nach der variationellen Methode (Kapitel 2) ist die Implementierung des nicht-lokalen Austauschoperators (1.27). Seine numerische Handhabung ist allerdings mit zwei Schwierigkeiten verbunden: (i) die Berechnung der Singularität im Nenner und (ii) die Symmetrisierung, wenn mit einem reduzierten \mathbf{k} -Punkt-Satz gearbeitet werden soll. In der Tat wird sich der Operator als die gegenwärtig zeitaufwendigste Größe im Algorithmus des variationellen EXX-Verfahrens erweisen, vgl. Kapitel 5.

In dieser Arbeit wurde die Berechnung des Austauschoperators in der Ebene-Wellen-Basis ausgeführt, dabei wurde auf eine bereits im Rahmen eines früheren Projekts (der Implementierung des Standard-EXX-Verfahrens in S/PHI/nX [81, 82]) angefertigte Realisierung zurückgegriffen. Allerdings legen unten angestellte Überlegungen die Vermutung nahe, dass eine Auswertung im Ortsraum die günstigere Alternative sein könnte. Es wäre daher interessant, in einem zukünftigen Projekt das variationelle EXX-Verfahren mit einer Ortsraumdarstellung des nicht-lokalen Austauschoperators zu kombinieren.

4.4.1 Behandlung der Singularität

Im reziproken Raum wird die Austauschwechselwirkung eines spin-kompensierten Systems näherungsweise durch einen Satz nicht-lokaler Ausdrücke der Form

$$\hat{V}_{\text{Fock}}(\mathbf{k}, \mathbf{G}, \mathbf{G}') = -\frac{4\pi}{\Omega} \sum_{n, \mathbf{q}}^{\text{occ}} \sum_{\mathbf{G}''} \frac{c_{n\mathbf{q}}(\mathbf{G} + \mathbf{G}'') c_{n\mathbf{q}}^*(\mathbf{G}' + \mathbf{G}'')}{|\mathbf{k} - \mathbf{q} - \mathbf{G}''|^2} \quad (4.61)$$

beschrieben [52, 98], wie in Anhang A.2 explizit vorgeführt ist. Sie sind, ähnlich wie die kinetische Energie in (3.15), jeweils für einen festen Punkt \mathbf{k} in der Brillouin-Zone (BZ) erklärt.

In der Summe treten dabei integrable Polstellen 2. Ordnung auf, wenn $\mathbf{k} - \mathbf{q}$ ein reziproker Gittervektor ist. Deshalb kann die Summation nicht ohne Weiteres mit Hilfe von wenigen, diskreten Punkten in der BZ ausgeführt werden, wie es etwa bei der Ladungsdichte möglich ist (vgl. Abschnitt 3.2). Eine Lösung dieses Problems wurde von Gygi und Baldereschi [98] vorgeschlagen. Durch Einführung einer Hilfsfunktion F wird die numerische Integration dabei an den Polstellen durch eine exakte Integration ersetzt.

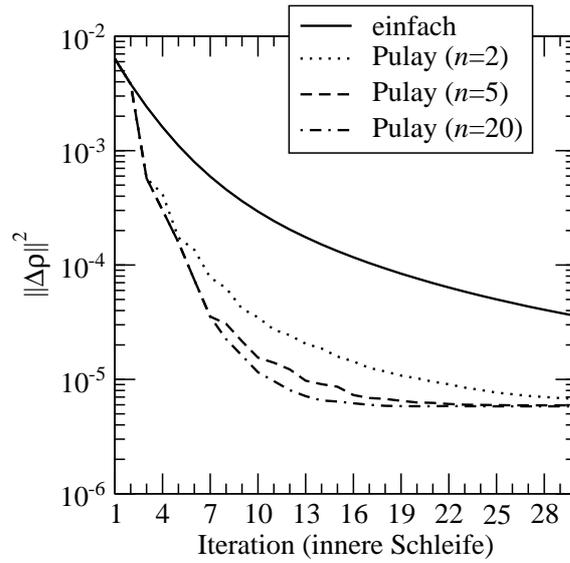


Abbildung 4.6: Minimierung der induzierten Ladungsdichte in GaAs. Gegenüber dem einfachen Mixer (4.41) lässt sich mit dem Pulay-Ansatz (4.44) eine deutlich schnellere Konvergenz erreichen. Das optimale v_x^{opt} wurde dabei jeweils aus einer Linearkombination der letzten n Iterationen bestimmt. Da der Pulay-Mixer für $n = 1$ in den einfachen übergeht, stimmen alle Kurven im ersten Schritt überein. Startpotential war das LDA-Austauschpotential. Für die Ladungsdichte wurde (4.33) benutzt, d.h., es wurde nur die explizite v_x -Abhängigkeit berücksichtigt. In Flussdiagramm 4.5 entspricht dies der inneren Schleife. Da der Grenzwert der induzierten Ladungsdichte ja exakt bekannt ist, nämlich $\Delta\rho \rightarrow 0$, ließ sich hier — im Gegensatz etwa zu Abb. 4.5, wo nur die Differenz zum besten erreichten Wert aufgetragen werden kann — der tatsächliche Fehler in jedem Schritt angeben. Die obige Kurve strebt einem Sättigungswert zu oder konvergiert zumindest nach einer gewissen Zeit sehr langsam; dann ist offensichtlich die Grenze der numerischen Genauigkeit erreicht.

Für festes \mathbf{k} seien also $\mathbf{q}_0, \mathbf{G}_0$ so, dass der Integrand in (4.61) singularär wird: $\mathbf{q}_0 = \mathbf{k} - \mathbf{G}_0$. Wegen (A.5) ist die Singularität damit für $\mathbf{q} \rightarrow \mathbf{q}_0$ asymptotisch durch

$$\frac{c_{n,\mathbf{k}-\mathbf{G}_0}(\mathbf{G} + \mathbf{G}_0) c_{n,\mathbf{k}-\mathbf{G}_0}^*(\mathbf{G}' + \mathbf{G}_0)}{|\mathbf{k} - \mathbf{q} - \mathbf{G}_0|^2} = \frac{c_{n,\mathbf{k}}(\mathbf{G}) c_{n,\mathbf{k}}^*(\mathbf{G}')}{|\mathbf{k} - \mathbf{q} - \mathbf{G}_0|^2} \quad (4.62)$$

gegeben. Dies kann für $\mathbf{q} \approx \mathbf{q}_0$ wiederum durch $c_{n,\mathbf{k}}(\mathbf{G}) c_{n,\mathbf{k}}^*(\mathbf{G}') F(\mathbf{p})$ angenähert werden, wenn $F(\mathbf{p})$ bei $\mathbf{p} \rightarrow 0$ die gleiche Asymptotik wie $1/p^2$ besitzt. Fordert man zusätzlich, dass F im reziproken Gitter (3.2) periodisch ist, lässt sich damit zu (4.61) folgendermaßen eine „geschickte Null“ hinzufügen:

$$\begin{aligned} \hat{V}_{\text{Fock}}(\mathbf{k}, \mathbf{G}, \mathbf{G}') = & \quad (4.63) \\ & -\frac{4\pi}{\Omega} \sum_n \left[\sum_{\mathbf{q}}' \left(\sum_{\mathbf{G}''} \frac{c_{n\mathbf{q}}(\mathbf{G} + \mathbf{G}'') c_{n\mathbf{q}}^*(\mathbf{G}' + \mathbf{G}'')}{|\mathbf{k} - \mathbf{q} - \mathbf{G}''|^2} - c_{n,\mathbf{k}}(\mathbf{G}) c_{n,\mathbf{k}}^*(\mathbf{G}') F(\mathbf{k} - \mathbf{q}) \right) \right. \\ & \left. + c_{n,\mathbf{k}}(\mathbf{G}) c_{n,\mathbf{k}}^*(\mathbf{G}') \sum_{\mathbf{q}} F(\mathbf{k} - \mathbf{q}) \right]. \end{aligned}$$

Hier heben sich die Singularitäten in den großen runden Klammern gerade auf. Alle Terme, bei denen $\mathbf{k} - \mathbf{q}$ ein reziproker Gittervektor ist, werden also ausgelassen, was durch den Strich an der ersten Summe über \mathbf{q} angedeutet ist. Im Argument von F konnte der jeweils kritische \mathbf{G}'' -Vektor wegen der vorausgesetzten Periodizität weggelassen werden. Ebenfalls

auf Grund der Periodizität von F ist die zweite Summe über \mathbf{q} unabhängig von \mathbf{k} . Dort kann F daher durch den mittleren Wert $\Omega/(2\pi)^3 \int_{\text{BZ}} F(\mathbf{p}) d^3p$ in der BZ ersetzt werden, falls das Integral bekannt ist.

Die Gestalt der Hilfsfunktion und ihr Integral hängen von der Struktur des betrachteten Systems ab. In Ref. 98 findet man indes beides nur für den Fall der kubisch-flächenzentrierten Gitterstruktur (fcc) angegeben. Offensichtlich aus diesem Grund sind EXX-Berechnungen periodischer Systeme¹⁵ bis 2005 nur für diese Struktur veröffentlicht worden. Wenzien, Cappellini und Bechstedt leiteten jedoch entsprechende Hilfsfunktionen auch für einfach-kubische (sc), kubisch-innenzentrierte (bcc), hexagonale sowie für einfach-orthorhombische Bravaisgitter ab [99]. Die benötigten Integrale liegen allerdings nur für die kubischen Gitter in analytisch-geschlossener Form vor. Für den hexagonalen Fall wurde daher das Ergebnis einer numerischen Integration in Ref. 99 tabelliert als Funktion $I(c/a)$ verschiedener Verhältnisse der Gitterkonstanten angegeben. Hier sei auf einen Druckfehler hingewiesen: Die erste Spalte von Tabelle II genannter Referenz listet c/a in Einheiten von $\sqrt{2/3}$ auf. Irrtümlich wurde aber der Kehrwert angegeben. Da I dort nur für wenige Seitenverhältnisse gegeben ist und sich besonders bei kleinen Werten c/a nur schwer linear approximieren lässt, wurde die Tabelle für die vorliegende Arbeit noch einmal neu und etwas feinmaschiger berechnet. Zusätzlich wurden die Integrale auch für den tetragonalen Fall (einfach-orthorhombisch mit $a = b$) ausgewertet. Die Ergebnisse zeigt Tab. 4.1.

Die Methode wurde im Rahmen dieser Arbeit für alle oben genannten Bravaisgitter¹⁶ in den Code S/PHI/nX [81, 82] implementiert. Dadurch konnten beispielsweise erstmals EXX- und darauf aufbauende EXX+ G_0W_0 -Berechnungen von GaN und InN in der gegenüber der Zinkblende-Struktur stabileren und daher technologisch relevanteren Wurtzit-Phase ermöglicht werden [172].

4.4.2 Symmetrisierung

Die zweite Schwierigkeit bei der Berechnung des Austauschoperators betrifft die Symmetrisierung, die sich im Vergleich mit der einer lokalen Größe wie der Dichte (vgl. Abschnitt 3.2) sehr aufwendig gestaltet. Wird in (4.61) die Summation über \mathbf{q} in eine Summe über einen irreduziblen Bereich der BZ sowie einer weiteren über alle Punkttransformationen aufgespalten, ergibt sich, wie in (A.8) gezeigt,

$$\hat{V}_{\text{Fock}}(\mathbf{k}, \mathbf{G}, \mathbf{G}') = -\frac{4\pi}{\Omega} \sum_n^{\text{occ}} \sum_{s=1}^{N_{\text{red}}} \sum_{i=1}^{N_{\text{PS}}} \sum_{\mathbf{G}''} w_s \frac{c_{n,\mathbf{q}_s}(S_i^{-1}\mathbf{G} + \mathbf{G}'') c_{n,\mathbf{q}_s}^*(S_i^{-1}\mathbf{G}' + \mathbf{G}'')}{|S_i^{-1}\mathbf{k} - \mathbf{q}_s - \mathbf{G}''|^2}. \quad (4.64)$$

Hier ist N_{red} die Anzahl der Punkte \mathbf{q}_s im irreduziblen Bereich der BZ mit den zugehörigen Wichtungen w_s , und N_{PS} bezeichnet die Anzahl der Punkttransformationen S_i . Die numerische Auswertung dieses Ausdrucks wird durch dessen Abhängigkeit von \mathbf{k} erschwert. Zum einen kommt dadurch eine Symmetrieoperation im Nenner ins Spiel. Wäre dies nicht der Fall, ließe sich zunächst ohne Einbeziehung der Symmetrien eine Matrix $\hat{V}_{\text{Fock}}^{\text{unsym}}(\mathbf{k}, \mathbf{G}, \mathbf{G}')$ konstruieren, die nur durch Umordnung ihrer Einträge (Ganzzahlarithmetik) und eine anschließende Summation symmetrisiert werden könnte — sehr ähnlich, wie in (3.12) die symmetrisierte Dichte aus der unsymmetrisierten gewonnen wird. Zum anderen aber muss

¹⁵Für radialsymmetrische Systeme wurde die EXX-Methode unter der Bezeichnung OEP bereits seit 1976 angewandt, vgl. Abschnitt 1.6.2.

¹⁶Für das orthorhombische Gitter wurde allerdings nur der tetragonale Spezialfall berücksichtigt. Eine Verallgemeinerung wäre jedoch nicht schwer zu implementieren; die Integrale hängen dann jeweils von zwei Seitenverhältnissen, z.B. b/a und c/a , ab.

hexagonal		tetragonal	
c/a	$I(c/a)$	c/a	$I(c/a)$
in Einh. $\sqrt{2/3}$	in Einh. $(a/2\pi)^3$	in Einh. 1	in Einh. $(a/2\pi)^3$
0	0	0	0
0.25	2.41198	0.25	3.04513
0.5	4.6399	0.5	5.77592
1	8.1361	1	9.8293
1.5	10.502	1.5	12.5214
2	12.4069	2	14.4238
3	14.8682	3	17.2456
4	16.4539	4	19.27
5	17.6677	5	20.436
6	18.7455	6	21.9869
7	19.8621	7	22.6149
8	20.4293	8	23.4241
10	21.7925	10	24.956
12	22.548	12	26.3131
24	26.4598	24	30.5836

Tabelle 4.1: Integrale der Hilfsfunktion über die Brillouin-Zone für verschiedene Verhältnisse c/a der Gitterkonstanten im hexagonalen und tetragonalen Bravaisgitter (vgl. Abschnitt 4.4.1). Die entsprechenden Hilfsfunktionen F sind Referenz 99 entnommen. Im tetragonalen Fall wurden in der dort angegebenen Funktion des orthorhombischen Gitters die Gitterkonstanten a und b gleichgesetzt.

(4.64) für jeden der N_{red} Punkte aus dem reduzierten \mathbf{k} -Punkt-Satz separat ausgerechnet werden. Die Größe skaliert damit äußerst ungünstig: bestehen die $|\mathbf{G} + \mathbf{k}$ -Basen jeweils aus etwa m ebenen Wellen, ist der erforderliche zeitliche Aufwand insgesamt durch $N_{\text{red}}N_{\text{PS}}n_{\text{val}}m^3$ gegeben (n_{val} : Anzahl der Valenzbänder).

Demgegenüber könnte die Symmetrisierung des Operators im Ortsraum numerisch einfacher zu gelingen, denn aus (1.27) folgt (für ein spin-kompensiertes System) mit den selben Bezeichnungen wie oben

$$V_{\text{Fock}}(\mathbf{r}, \mathbf{r}') = - \sum_n^{\text{occ}} \sum_{s=1}^{N_{\text{red}}} \sum_{i=1}^{N_{\text{PS}}} w_s \frac{\psi_{n,\mathbf{q}_s}(S_i^{-1}\mathbf{r})\psi_{n,\mathbf{q}_s}^*(S_i^{-1}\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}. \quad (4.65)$$

Die Vorteile dieser Formulierung sind: (i) Es muss nur noch *eine* Matrix berechnet werden. (ii) Wegen $|\mathbf{r} - \mathbf{r}'| = |S_i^{-1}\mathbf{r} - S_i^{-1}\mathbf{r}'|$ lässt sich im Gegensatz zur Darstellung im reziproken Raum zunächst durch Weglassen der Summe über die S_i eine unsymmetrisierte Version $V_{\text{Fock}}^{\text{unsym}}(\mathbf{r}, \mathbf{r}')$ konstruieren, die anschließend durch Umordnen ihrer Einträge symmetrisiert werden kann. (iii) Es entfällt die durch die Fourier-Entwicklung ins Spiel gekommene Faltung (Summe über \mathbf{G}'' in (4.64)).

4.5 Variationelles EXX als selbstkonsistenter Zyklus

Nachdem in den vorangehenden Abschnitten alle für die numerische Implementierung der variationellen EXX-Methode notwendigen Einzelschritte vorgestellt wurden, können sie jetzt zusammengefügt werden. Flussdiagramm 4.5 gibt schematisch einen Überblick über den Programmablauf, wie er in dieser Arbeit realisiert wurde. Er besteht aus mehreren, ineinander verschachtelten Iterationszyklen, deren Funktion und Zusammenspiel direkt aus

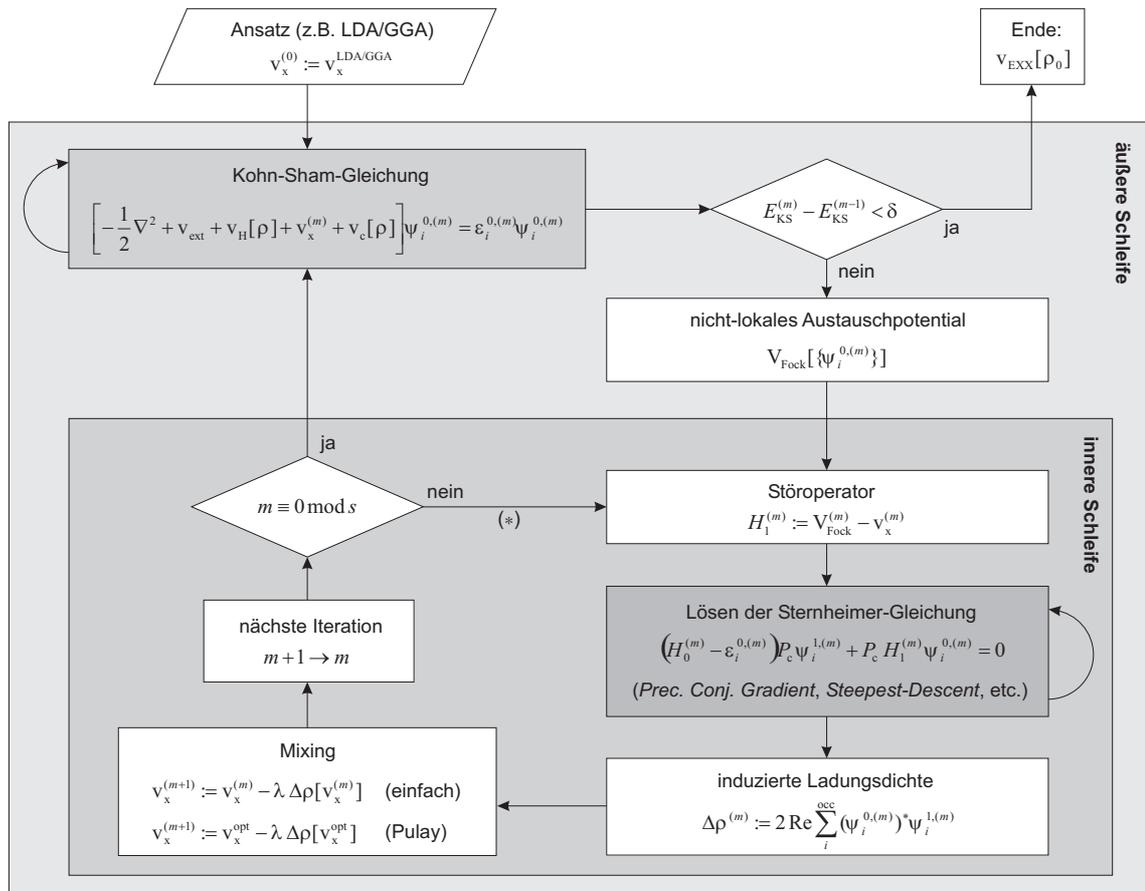
den einleitenden Überlegungen von Abschnitt 4.3 folgt. Dort wurde diskutiert, dass die induzierte Ladungsdichte $\Delta\rho$ entweder (i) mit *festgehaltenen* Kohn-Sham-Orbitalen ψ_i^0 erfolgen kann — dies liefert das EXX-Potential der zu diesen ψ_i^0 gehörigen Ladungsdichte — oder aber (ii) mit einer *Aktualisierung* der Orbitale — dann erhält man das EXX-Potential der Grundzustandsladungsdichte ρ_0 . Da man gewöhnlich an Letzterem interessiert ist, könnte man nun direkt Strategie (ii) verfolgen: sie ist in Diagramm 4.5 gegeben, wenn die mit (*) markierte Querverbindung nicht benutzt wird (d.h., es wird dort $s = 1$ gesetzt). Im Prinzip würde der Algorithmus so bereits funktionieren. Er besteht dann aus folgenden Schritten:

- (1) Vorgabe eines geeigneten Start-Austauschpotentials $v_x^{(0)}$, etwa aus einer vorangehenden LDA- oder GGA-Rechnung (Abschnitt 1.5.2).
- (2) Berechnung der Wellenfunktionen $\psi_i^{0,(m)}$ durch Lösen der Kohn-Sham-Gleichung zu festgehaltenem Probe-Austauschpotential $v_x^{(m)}$ (Gleichung (4.36)).
- (3) Berechnen des nicht-lokalen Austauschoperators $V_{\text{Fock}}^{(m)}$ aus den $\psi_i^{0,(m)}$ (Abschnitt 4.4).
- (4) Erzeugen des Störoperators $H_1^{(m)} := V_{\text{Fock}}^{(m)} - v_x^{(m)}$.
- (5) Lösen der Sternheimer-Gleichung mit Hilfe eines geeigneten Minimierungsverfahrens (Abschnitt 4.1), um die linearen Korrekturen $\psi_i^{1,(m)}$ zu erhalten.
- (6) Berechnen der induzierten Ladungsdichte $\Delta\rho^{(m)}$ aus den Orbitalen $\psi_i^{0,(m)}$ und den Korrekturen $\psi_i^{1,(m)}$ (Abschnitt 4.2).
- (7) Erzeugen des neuen Probe-Austauschpotentials $v_x^{(m+1)}$ durch einfaches Mischen oder Pulay-Mischen von $\Delta\rho^{(m)}$ und $v_x^{(m)}$ (Abschnitt 4.3).
- (8) Wiederholen der Schritte (2)–(7) solange, bis Selbstkonsistenz erreicht ist.

Die Schritte (2) und (5) sind dabei selbst iterative Zyklen. Es sei auch noch einmal betont, dass Punkt (2) nur die Lösung der Kohn-Sham-Gleichung zu einem fest vorgegebenen Probe-Austauschpotential beinhaltet und folglich nicht aufwendiger ist als eine gewöhnliche LDA-Kalkulation. Es zeigt sich jedoch im Folgenden, dass es technisch viel günstiger ist, auch bei der Berechnung von $v_{\text{EXX}}[\rho_0]$ so viel wie möglich auf die Minimierung gemäß (i) zurückzugreifen.

4.5.1 Optimierung des Algorithmus

Auf Grund der schlechten Skalierung des Austauschoperators in Abhängigkeit von der Größe des Basissatzes (vgl. Abschnitt 4.4.2 und Tabelle 5.1) wäre eine Implementierung der variationellen EXX-Methode gemäß der obigen Punkte (1)–(8), wobei in jedem Iterationszyklus der Austauschoperator neu generiert würde, nicht effizient. Der Aufwand kann aber auf einfache Weise erheblich reduziert werden: Die Idee ist, nicht für jede Aktualisierung $v_x^{(m)}$ des lokalen Austauschpotentials die Wellenfunktionen neu zu berechnen und damit anschließend die teure Aktualisierung von V_{Fock} durchzuführen, sondern dies erst dann zu tun, wenn v_x an die aktuell *gegebenen* Orbitale hinreichend angepasst ist. Mit anderen Worten: man minimiert zunächst $\Delta\rho$, *ohne* die Abhängigkeit der Orbitale von v_x zu berücksichtigen, verfährt also genau wie in Punkt (i) des Abschnitts 4.3 beschrieben.



Flussdiagramm 4.5: Schematische Darstellung des varEXX-Algorithmus. Der iterative Prozess besteht aus vier ineinander verschachtelten Selbstkonsistenzzyklen, die hier durch verschiedene Graustufen hervorgehoben sind. In der äußeren Schleife wird die Kohn-Sham-Gleichung (1.52) in EXX-Genauigkeit gelöst. Da v_x nur über die Orbitale zugänglich ist, werden diese zunächst in einer Unterschleife für festes $v_x^{(m)}$ erzeugt und dann — nach Berechnung von V_{Fock} — an eine zweite, „innere“ Schleife übergeben. Dort findet die Minimierung der induzierten Dichte statt, wobei in jedem Schritt noch die Sternheimer-Gleichung iterativ gelöst wird. Das so erzeugte Austauschpotential dient dann als Startpunkt der nächsten (äußeren) Iteration.

Erst, *nachdem* dies vollbracht ist,¹⁷ werden Wellenfunktionen, nicht-lokaler Austauschoperator und Störoperator aktualisiert. Solange keine Selbstkonsistenz erreicht ist, sind diese Größen dann jedoch nicht mehr mit dem zuletzt berechneten v_x konsistent. $\Delta \rho$ ist daher wieder verschieden von 0 und wird aufs Neue minimiert, natürlich ebenfalls, ohne die Orbitale zu aktualisieren. Dieses Vorgehen lässt sich zusammenfassen, indem in obiger Aufzählung Schritt (8) durch die folgenden beiden Schritte ersetzt wird:

- (8') Wiederholen der Schritte (4)–(7) solange, bis $v_x^{(m)}$ hinreichend (siehe unten) konvergiert ist.
- (9) Wiederholen der Schritte (2)–(8') solange, bis Selbstkonsistenz erreicht ist.

Durch (8') wird also eine weitere Unterschleife (im Folgenden „innere Schleife“ genannt)

¹⁷In der Praxis muss die Minimierung von $\Delta \rho$ dabei nicht bis an die Grenze der numerischen Genauigkeit erfolgen, da dies insgesamt nicht zu einer schnelleren Konvergenz führt (siehe unten).

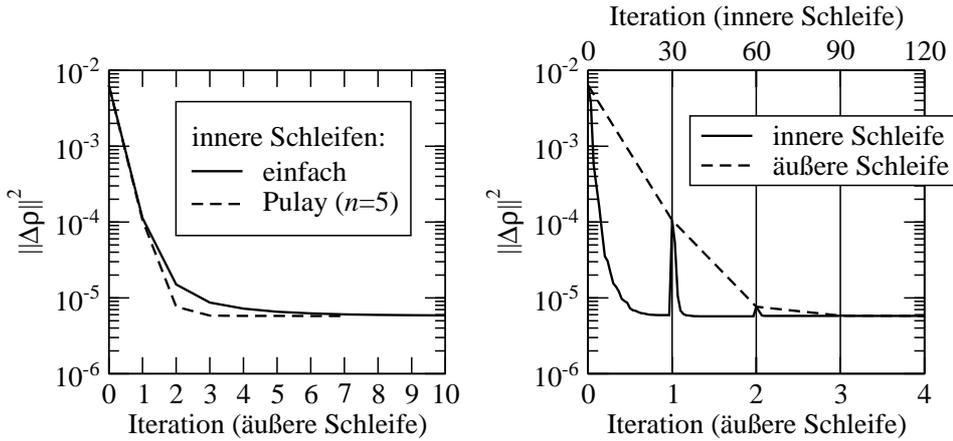


Abbildung 4.7: Links: Minimierung der induzierten Ladungsdichte in GaAs gemäß der äußeren Schleife in Flussdiagramm 4.5. Die für jeden Einzelschritt benötigte Minimierung von $\Delta\rho$ gemäß der inneren Schleife (vgl. Abb. 4.6) erfolgte dabei mittels einfachem *Mixing* (durchgezogene Linie) oder mit dem Pulay-Verfahren (gestrichelte Linie). Letzteres ist hier bereits nach 3 Iterationen bis zur Grenze der numerischen Genauigkeit konvergiert. Rechts: Norm der induzierten Ladungsdichte $\Delta\rho$ in Abhängigkeit von allen Aktualisierungen des Austauschpotentials v_x . Nach jeder Minimierung von $\|\Delta\rho\|$ in der inneren Schleife (hier jeweils 30 Schritte) werden die Wellenfunktionen an v_x angepasst (äußere Schleife). Dadurch springt $\|\Delta\rho\|$ zu Beginn jeder äußeren Schleife wieder nach oben, mit zunehmender Selbstkonsistenz verschwinden diese Ausschläge jedoch. Zum Konvergenzverhalten der beiden Kurven vgl. Abb. 4.6.

in den Gesamtzyklus (im Folgenden als „äußere Schleife“ bezeichnet) eingefügt. In Diagramm 4.5 wird dies durch Einschalten der Querverbindung (*) erreicht, wobei s die Anzahl der für (8') benötigten Schritte ist. Bei den in dieser Arbeit untersuchten Materialien stellt $10 \leq s \leq 30$ eine geeignete Wahl dar. In Abbildung 4.7 ist die Abhängigkeit der (Norm der) induzierten Dichte von der Iteration der inneren bzw. äußeren Schleife illustriert. Tatsächlich braucht in (8') nicht stets bis an die Grenze der numerischen Genauigkeit konvergiert zu werden; da die Konsistenz aller beteiligten Größen mit jedem äußeren Schritt zunimmt, verbleibt in den jeweiligen Unterschritten nämlich immer weniger „zu tun“. Anders formuliert: mit der Zeit werden die Vorgaben für die Unterzyklen so gut, dass die Konvergenz dort schnell erreicht ist. Dies gilt insbesondere auch für die Sternheimer-Gleichung, die sich innerhalb der inneren Schleife befindet und deshalb in jedem — und nicht nur in jedem s -ten — Schritt gelöst werden muss. Dazu benutzt man sinnvollerweise als Ansatz für die Korrekturfunktionen $\psi_i^{1,(m)}$ das Ergebnis der jeweils vorangehenden Iteration $\psi_i^{1,(m-1)}$. Ähnliches gilt auch für die Ansatzfunktion des EXX-Potentials selbst; zwar erwies sich der Algorithmus — wie an GaAs, GaN und InN getestet — auch für das Startpotential $v_x^{(0)} = 0$ als stabil, aber effizienter ist es natürlich, bereits von einem Näherungsansatz wie etwa LDA oder GGA aus zu starten.

Obwohl durch die Einführung der „innere Schleife“ die Anzahl notwendiger Aktualisierungen des nicht-lokalen Austauschoperators schon stark reduziert werden konnte, ist es dennoch sinnvoll, noch nach weiteren Möglichkeiten zu suchen, den mit V_{Fock} verbundenen Aufwand zu vermindern. So lässt sich beispielsweise auszunutzen, dass V_{Fock} nur benötigt wird, um die linearen Korrekturfunktionen der Valenzbandorbitale zu berechnen. Wie Hybertsen und Louie diskutieren [176], benötigt man zu deren Beschreibung aber weit weniger ebene Wellen als für die Funktionen des Leitungsbandes. Daher ist es sinnvoll, für V_{Fock} eine Extra-Abschneideenergie $E_{\text{cut}}^{\text{Fock}}$ einzuführen, die sich unabhängig

von der Abschneideenergie E_{cut} des vollen Hamiltonoperators einstellen lässt und kleiner gewählt werden kann. Wie E_{cut} muss auch $E_{\text{cut}}^{\text{Fock}}$ letztlich in Konvergenztests bestimmt werden; in der Regel ist dabei $E_{\text{cut}}^{\text{Fock}} \approx 2/3 E_{\text{cut}}$ ausreichend [168].

Wie in Abschnitt 4.4.2 diskutiert, trägt auch die Symmetrisierung des Austauschoperators erheblich zu dessen schlechter Skalierung bei. Praktisch wäre es daher, die Symmetrisierung des Operators einfach weglassen zu können. Tatsächlich lässt sich nun die EXX-Methode dahingehend nähern, dass nur noch das *lokale* Austauschpotential symmetrisiert werden muss. Dies soll in Abschnitt 4.6 erörtert werden. Es ist dazu allerdings hilfreich, sich die Unterschiede sowie die Parallelen zwischen den numerischen Algorithmen des variationellen EXX-Verfahrens (Flussdiagramm 4.5) und der Standard-Methode (Abschnitt 1.6.1) genau zu verdeutlichen.

4.5.2 Variationelles und Standard-EXX: Ein Vergleich des Iterationsprozesses

Im Unterschied zu den Abschnitten 2.2 und 2.3, in denen variationelle und Standard-Methode rein formal-analytisch miteinander verglichen wurden, soll hier untersucht werden, inwieweit sich der *iterative Prozess*, also die Entwicklung von der Ansatzfunktion $v_{\text{x}}^{(0)}$ bis zum auskonvergierten EXX-Potential, bei beiden Methoden unterscheidet. Die Kenntnis der Gemeinsamkeiten kann besonders beim Programmieren nützlich sein, wenn eines der Verfahren bereits im Code vorhanden ist.

Grob umrissen basiert der Algorithmus der Standard-Methode (stEXX) auf folgenden Punkten:

- (1a) Vorgabe eines geeigneten Start-Austauschpotentials $v_{\text{x}}^{(0)}$, etwa aus einer vorangehenden LDA- oder GGA-Rechnung (Abschnitt 1.5.2).
- (2a) Selbstkonsistentes Lösen der Kohn-Sham-Gleichung zu festgehaltenem Probe-Austauschpotential $v_{\text{x}}^{(m)}$ (Gleichung (4.36)).
- (3a) Vollständige Diagonalisierung der Hamilton-Matrix (Kohn-Sham-Matrix), um auch die unbesetzten Zustände zu erhalten.
- (4a) Berechnen des nicht-lokalen Austausch-Operators $V_{\text{Fock}}^{(m)}$, der Greenschen Funktionen $G_i^{(m)}$ und der (Inversen der) linearen Antwort $\chi_0^{(m)}$.
- (5a) Berechnen der von $V_{\text{Fock}}^{(m)}$ induzierten Ladungsdichte $\Delta\rho_{\text{Fock}}^{(m)}$ mit Hilfe der Greenschen Funktionen (eckige Klammer in Gleichung (1.94)).
- (6a) Erzeugen des neuen Probe-Austauschpotentials $v_{\text{x}}^{(m+1)}$ durch Anwenden von $(\chi_0^{(m)})^{-1}$ auf $\Delta\rho_{\text{Fock}}^{(m)}$ (Gleichung (1.94)).
- (7a) Wiederholen der Schritte (2a)–(6a) solange, bis Selbstkonsistenz erreicht ist.

Um die folgende Gegenüberstellung mit dem oben (Seiten 82/83) beschriebenen Ablauf der variationellen Methode (varEXX) zu erleichtern, sind die stEXX-Schritte mit hier einem „a“ markiert.

Sei nun beiden Methoden ein gleiches Startpotential $v_{\text{x}}^{(0)}$ vorgegeben und zunächst der „nullte“ Durchlauf (also $m = 0$) betrachtet. Damit sind auch die in den varEXX-Schritten (1)–(3) bzw. den stEXX-Schritten (1a)–(4a) berechneten unbesetzten Zustände $\psi_i^{0,(0)}$

und $V_{\text{Fock}}^{(0)}$ völlig identisch. Der Unterschied besteht bis jetzt also lediglich darin, dass in stEXX eine Reihe (teurer) Größen zusätzlich berechnet wird: unbesetzte Zustände durch Diagonalisierung der Kohn-Sham-Matrix, die Greensche Funktion sowie die lineare Antwort und deren Inverse.

Anschließend wird nun sowohl in varEXX (4)–(6) als auch in stEXX (5a) eine lineare Dichteänderung betrachtet — ab hier allerdings unterscheiden sich die Berechnungen: Während bei varEXX die Änderung durch den Operator $H_1 := V_{\text{Fock}} - v_x$ induziert wird, ist die Störung bei stEXX durch V_{Fock} allein gegeben (vgl. die Bemerkungen nach Gleichung (1.94)). Noch größer die Unterschiede bei der Aktualisierung des Austauschpotentials: varEXX erzeugt $v_x^{(1)}$ durch Mixen von $\Delta\rho^{(0)}$ und $v_x^{(1)}$, stEXX hingegen durch Multiplizieren von $\Delta\rho_{\text{Fock}}^{(0)}$ mit $(\chi_0^{(0)})^{-1}$. Tatsächlich weichen die in (7) bzw. (6a) generierten Austauschpotentiale $v_x^{(1)}$ im Allgemeinen stark voneinander ab, und auf den ersten Blick scheinen damit auch alle weiteren Iterationen bis zum Erreichen der Selbstkonsistenz nicht mehr miteinander vergleichbar zu sein. Durch die in (8') eigentlich zur Optimierung der varEXX-Methode eingeführte „innere Schleife“ ist dies aber dennoch möglich: Da die stEXX-Schritte (5a) und (6a) der Auswertung der EXX-Gleichung (1.94) entsprechen, gilt dort $v_x^{(1)} = v_{\text{EXX}}[\rho^{(0)}]$, d.h., es wird das *exakte* Austauschpotential der zu den $\psi_i^{0,(0)}$ gehörenden Dichte $\rho^{(0)}$ erzeugt. Nun sei andererseits die innere Schleife des varEXX-Algorithmus betrachtet: Hier wird, ebenfalls ausgehend von den $\psi_i^{0,(0)}$, die induzierte Dichte minimiert, *ohne* zwischenzeitlich die Wellenfunktionen zu aktualisieren. Auch dies führt aber gemäß Punkt (i) der Einleitung von Abschnitt 4.3 gerade auf $v_{\text{EXX}}[\rho^{(0)}]$. Lässt man die innere Schleife also vollständig auskonvergieren, wird auf diese Weise das stEXX-Verfahren „wieder eingeholt“. Somit startet der nächste Durchlauf, beginnend in (2) bzw. (2a), in beiden Verfahren wiederum vom gleichen Ausgangs-Potential und alle weiteren ebenso. Es lässt sich also zusammenfassen: Die Aktualisierungen des lokalen Austausch-Potentials im stEXX-Verfahren entsprechen jenen der *äußeren Schleife* im varEXX-Verfahren.

Die Äquivalenz beider Methoden geht damit über das in Abschnitt 2.2 Gezeigte hinaus: Dort wurde — streng genommen — nur die Äquivalenz der Formulierungen für irgendeine feste Dichte, die nicht notwendigerweise die Grundzustandsladungsdichte¹⁸ des betrachteten Systems ist, bewiesen. In stEXX und varEXX wird aber nicht nur das EXX-Potential einer Dichte berechnet, sondern zugleich die Kohn-Sham-Gleichung 1.52 selbstkonsistent gelöst. Dass dies, trotz aller technischen Verschiedenheit im Umgang mit dem Austauschpotential, auf ganz analoge Weise geschieht, wurde hier gezeigt. Die Kenntnis der Analogien zwischen beiden Methoden wird nun im folgenden Abschnitt benutzt, um eine zunächst für die Standard-Methode formulierte Näherung bei der Behandlung der Symmetrien auch auf die variationelle Formulierung übertragen zu können.

4.6 Symmetriegenäherte EXX-Methode

Die Hauptschwierigkeit dieser Arbeit bestand darin, dass während der Implementierung der variationellen EXX-Methode in S/PHI/nX [81, 82] zunächst auf Teile (insbesondere den nicht-lokalen Austauschoperator) des bereits zuvor in diesen Code implementierten Standard-EXX-Verfahrens zurückgegriffen wurde, wobei allerdings nicht bekannt war, dass es sich bei dessen Realisierung um eine sehr effektive *Näherung* handelte. Diese Näherung

¹⁸Sie ist aber durch die vorausgesetzte v -Darstellbarkeit und die Hohenberg-Kohn-Theoreme eindeutig einem Satz Wellenfunktionen zugeordnet.

besteht darin, alle im Formalismus auftretenden Brillouin-Zonen-Integrationen nur über dem *irreduziblen* Bereich der Brillouin-Zone auszuführen, *ohne* sie im Anschluss daran zu symmetrisieren. Symmetrisiert wird dann einzig das daraus resultierende *lokale* Austauschpotential. Daher sei diese Vorgehensweise im Folgenden als „symmetriegenäherte EXX-Methode“ bezeichnet.

Tatsächlich ist die so erreichte Reduktion der benötigten CPU-Zeit enorm, was vor allem auf die Vereinfachungen bei der Auswertung des nicht-linearen Austauschoperators zurückzuführen ist: Wird etwa die Integration über die Brillouin-Zone einer Zinkblende-Struktur mit Hilfe einer $(4 \times 4 \times 4)$ -Zerlegung nach Monkhorst und Pack [96] ausgeführt, so benötigt man bei Weglassen der Summe über die Symmetrien S_i in Ausdruck (4.64) nur rund 4% der Zeit, die bei voller Symmetrisierung des nicht-linearen Austauschoperators erforderlich wäre. Streng formal lässt sich diese Näherung allerdings nicht begründen, ihre Rechtfertigung liegt vielmehr in der empirisch gefundenen guten Übereinstimmung der so berechneten Bandlücken mit den Resultaten einer exakten Rechnung (Abweichungen kleiner als wenige Zehntel eV, vgl. Tab. 4.2). In gewisser Weise ist die Situation somit ähnlich wie beim LDA-Funktional: auch dessen Erfolg konnte aus dem zu Grunde liegenden ad-hoc-Modell (Abschnitt 1.5.2) kaum begründet werden und musste daher zunächst¹⁹ als eine praktische Erfahrungstatsache gewertet werden.

Die Idee der symmetriegenäherten EXX-Methode basiert auf einer Modifikation der für periodische Randbedingungen formulierten EXX-Gleichung (1.94) und sei deshalb zuerst für das Standard-EXX-Verfahren beschrieben (Abschnitt 4.6.2). Da eine komplette Symmetrisierung des nicht-lokalen Austauschoperators aber auch die Anwendungsmöglichkeiten des variationellen Ansatzes stark beschränkt (V_{Fock} erweist sich als *die* Engstelle im variationellen EXX-Verfahren, vgl. Kapitel 5), ist es lohnenswert, die Näherung auch hierauf zu übertragen (Abschnitt 4.6.3). Zuvor jedoch scheint es hilfreich, die Gleichungen einmal explizit so aufzuschreiben, wie sie sich bei exakter Symmetrisierung ergeben (nächster Abschnitt). Zur Vereinfachung wird dazu im Folgenden ein nicht spin-polarisierter Halbleiter oder Isolator vorausgesetzt.

4.6.1 Exakte Symmetrisierung beim Standard-EXX-Verfahren

Formuliert man Gleichung (1.94) speziell für periodische Randbedingungen und spaltet dabei gemäß Abschnitt 3.2 alle Brillouin-Zonen-Integrale auf in eine Summe über den reduzierten \mathbf{k} -Punkt-Satz $\{\mathbf{k}_1, \dots, \mathbf{k}_{N_{\text{red}}}\}$ und eine zweite Summe über die Punkttransformationen $\{S_1, \dots, S_{N_{\text{PS}}}\}$ im reziproken Raum (die von den räumlichen Punkttransformationen unabhängige Zeitumkehrsymmetrie (3.7) sei also darin enthalten), erhält man

$$\begin{aligned} v_x[\rho](\mathbf{r}) &= \int d^3r' \chi_0^{-1}(\mathbf{r}, \mathbf{r}') \Delta\rho_{\text{Fock}}(\mathbf{r}') \\ &= \int d^3r' \left(\sum_{\nu=1}^{N_{\text{PS}}} \sum_{s=1}^{N_{\text{red}}} w_s \chi_{S_\nu \mathbf{k}_s}^0 \right)^{-1}(\mathbf{r}, \mathbf{r}') \left(\sum_{\nu=1}^{N_{\text{PS}}} \sum_{s=1}^{N_{\text{red}}} w_s \Delta\rho_{S_\nu \mathbf{k}_s}^{\text{Fock}}(\mathbf{r}') \right). \end{aligned} \quad (4.66)$$

Hier ist

$$\Delta\rho_{\text{Fock}}(\mathbf{r}') := \sum_{\mathbf{k}} \Delta\rho_{\mathbf{k}}^{\text{Fock}}(\mathbf{r}') \quad (4.67)$$

¹⁹Es wurden allerdings später Begründungen für das LDA-Funktional in Form von Erhaltungssätzen gefunden, vgl. z.B. die Referenzen 5 und 149.

$$= \sum_{\mathbf{k}} \sum_i^{\text{occ.}} 2 \int d^3 r'' \underbrace{\left[\psi_{i\mathbf{k}}^*(\mathbf{r}') G_{i\mathbf{k}}^\perp(\mathbf{r}', \mathbf{r}'') V_{\text{Fock}}[\{\psi_l\}] \psi_{i\mathbf{k}}(\mathbf{r}'') + \text{c.c.} \right]}_{=: \Delta \rho_{\mathbf{k}}^{\text{Fock}}(\mathbf{r}')}$$

mit der in (F.1) definierten Greenschen Funktion $G_{i\mathbf{k}}^\perp$ die vom nicht-lokalen Austauschoperator V_{Fock} induzierte lineare Dichteänderung (vgl. Abschnitt 1.6.1), und die Summation über \mathbf{k} wurde in (4.66) dann mit Hilfe der „Symmetrisierung“ als

$$\Delta \rho_{\text{Fock}}(\mathbf{r}') := \sum_{\mathbf{k}} \Delta \rho_{\mathbf{k}}^{\text{Fock}}(\mathbf{r}') = \sum_{\nu=1}^{N_{\text{PS}}} \sum_{s=1}^{N_{\text{red}}} w_s \Delta \rho_{S_\nu \mathbf{k}_s}^{\text{Fock}}(\mathbf{r}')$$

ausgeführt. Der Austauschoperator selbst ist dabei durch den Kern

$$V_{\text{Fock}}(\mathbf{r}, \mathbf{r}') := - \sum_{i, \mathbf{k}}^{\text{occ.}} \frac{\psi_{i\mathbf{k}}(\mathbf{r}) \psi_{i\mathbf{k}}^*(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} = - \sum_{\nu=1}^{N_{\text{PS}}} \sum_{s=1}^{N_{\text{red}}} \sum_i^{\text{occ.}} w_s \frac{\psi_{i, S_\nu \mathbf{k}_s}(\mathbf{r}) \psi_{i, S_\nu \mathbf{k}_s}^*(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \quad (4.68)$$

gegeben. Ferner wird die lineare Antwort im hier betrachteten Unterraum gitterperiodischer Ladungsdichten und Potentiale zu

$$\begin{aligned} \chi_0(\mathbf{r}, \mathbf{r}') &= \sum_{\mathbf{k}} \sum_i^{\text{occ.}} \sum_j^{\text{unocc.}} 2 \frac{\psi_{i\mathbf{k}}^*(\mathbf{r}) \psi_{j\mathbf{k}}(\mathbf{r}) \psi_{j\mathbf{k}}^*(\mathbf{r}') \psi_{i\mathbf{k}}(\mathbf{r}') + \psi_{i\mathbf{k}}(\mathbf{r}) \psi_{j\mathbf{k}}^*(\mathbf{r}) \psi_{j\mathbf{k}}(\mathbf{r}') \psi_{i\mathbf{k}}^*(\mathbf{r}')}{\varepsilon_{i\mathbf{k}} - \varepsilon_{j\mathbf{k}}} \\ &= \sum_{\mathbf{k}} \sum_i^{\text{occ.}} \sum_j^{\text{unocc.}} 4 \frac{\psi_{i\mathbf{k}}^*(\mathbf{r}) \psi_{j\mathbf{k}}(\mathbf{r}) \psi_{j\mathbf{k}}^*(\mathbf{r}') \psi_{i\mathbf{k}}(\mathbf{r}')}{\varepsilon_{i\mathbf{k}} - \varepsilon_{j\mathbf{k}}}, \quad (4.69) \\ &=: \chi_{\mathbf{k}}^0(\mathbf{r}, \mathbf{r}') \end{aligned}$$

wie den Referenzen 52 und 176 entnommen werden kann. Im letzten Schritt wurde dabei die Zeitumkehrsymmetrie (3.7) ausgenutzt, was für das Folgende entscheidend sein wird. Der in (4.67) sowie (4.69) auftretende Faktor 2 resultiert aus der Summe über den Spin.²⁰

Insgesamt treten in (4.66) also drei Brillouin-Zonen-Integrationen auf — in $\Delta \rho_{\text{Fock}}$, in V_{Fock} und in χ_0 —, die bei exakter Auswertung der Gleichung jeweils unabhängig voneinander ausgeführt werden müssen. Insbesondere muss also jeder dieser Ausdrücke „extra“ symmetrisiert werden, und erst anschließend können sie gemäß (4.66) zu v_x zusammengesetzt werden, das selbst *nicht* mehr symmetrisiert zu werden braucht.

4.6.2 Symmetriegenähertes Standard-EXX-Verfahren

Beim symmetriegenäherten EXX wird nun die Auswertung der vorangehenden Gleichungen wie folgt modifiziert: Sowohl in (4.66) als auch in der rechten Seite von (4.68) wird nur noch die Summe über den reduzierten \mathbf{k} -Punkt-Satz berücksichtigt, die Summation über die Punkttransformationen S_ν hingegen entfällt. Im Allgemeinen besitzen aber dann die Ausdrücke (4.67) bis (4.69) für sich selbst genommen nicht mehr die korrekte Symmetrie, d.h., sie sind nicht mehr invariant gegenüber den (räumlichen) Punkttransformationen des Kristalls. Dies überträgt sich auf das in (4.66) berechnete v_x , das deswegen jetzt — im Gegensatz zur exakten Rechnung — nachträglich symmetrisiert wird. Man beachte, dass diese Symmetrisierung aber keiner \mathbf{k} -Punkt-Summation mehr zugeordnet werden kann; es

²⁰In den Gleichungen (8) und (11) von Ref. 52 wurde dieser Faktor irrtümlicherweise nicht mit angegeben, im späteren Verlauf jedoch implizit berücksichtigt.

werden also nur die räumlichen Punkttransformationen des Bravais-Gitters berücksichtigt. Ein weiteres Problem ist, wie sich gleich zeigen wird, dass das so berechnete v_x komplexwertig in der Ortsdarstellung ist. Um zu einem rein reellwertigen Potential zu gelangen, werden daher die Imaginäranteile abgeschnitten, d.h., künstlich zu Null gesetzt. Insgesamt hat man also fünf Änderungen zur exakten Vorgehensweise:

- (i) keine Symmetrisierung von V_{Fock}
- (ii) keine Symmetrisierung von $\Delta\rho_{\text{Fock}}$
- (iii) keine Symmetrisierung von χ_0
- (iv) Wegschneiden des Imaginäranteils im aus (i) – (iii) berechneten v_x
- (v) Symmetrisierung von v_x

Dass die Punkte (i)–(iii) nicht nur auf ein räumlich nicht der Kristallsymmetrie entsprechendes, sondern auch auf ein komplexwertiges Potential führen, folgt aus der Tatsache, dass das Weglassen der Symmetrisierung in (4.66) und (4.68) auch die Inversionssymmetrie, also die Zeitumkehrinvarianz (3.7), betrifft. In einer „gewöhnlichen“ DFT-Rechnung, in der ein von der Dichte parametrisch abhängendes Austausch-(Korrelations-)Potential verwendet wird, kann man die Zeitumkehrsymmetrie bei der Symmetrisierung vernachlässigen: wegen $|\psi_{i,-\mathbf{k}}|^2 = |\psi_{i,\mathbf{k}}|^2$ ist es egal, ob sie in (3.12) explizit mit einbezogen wird oder nicht, wenn nur die Gewichte w_s entsprechend gesetzt werden. Auf Grund der *Orbital*abhängigkeit des EXX-Formalismus ist diese Vereinfachung jetzt allerdings nicht mehr zulässig. Tut man es — wie hier — dennoch, ändern sich grundlegende Eigenschaften der betrachteten Ausdrücke. So erkennt man leicht, dass die Einträge des Austauschoperators (4.68) in Ortsdarstellung rein reell sind, wenn zu jedem \mathbf{k} auch $-\mathbf{k}$ berücksichtigt wird. Geht dagegen nur noch einer dieser Punkte in die Summe ein, wird der Operator komplexwertig. Die von ihm bewirkte Dichteänderung (4.67) bleibt jedoch weiterhin eine reellwertige Größe, da sich der Imaginärteil darin explizit wegekürzt.

Nun zur linearen Antwort: Hier hängt es davon ab, ob in (4.69) die obere oder die untere Zeile benutzt wird. Enthält die Summe \mathbf{k} sowohl als $-\mathbf{k}$, sind beide Zeilen äquivalent und reell. Entfällt aber einer dieser Punkte, ist nur noch die obere reell, die untere nicht mehr. Zudem liefert die untere Zeile dann auch keine symmetrische Matrix mehr (im reziproken Raum ist sie entsprechend nicht mehr hermitesch, vgl. (A.10) und (A.12) in Anhang A.3). Aus der Anwendung der so manipulierten linearen Antwort auf den — wie gezeigt — reellwertigen Ausdruck (4.67) folgt dann unmittelbar die Komplexwertigkeit von v_x .

Tatsächlich ist es wichtig, beim symmetriegenäherten EXX die zweite Zeile von (4.69) zu benutzen. Würde man die erste verwenden, und damit also die Zeitinversionssymmetrie aus der Näherung herausnehmen, funktioniert die Methode nicht mehr. Dies gilt auch umgekehrt, wenn etwa V_{Fock} korrekt symmetrisiert wird, $\Delta\rho_{\text{Fock}}$ und/oder χ_0 aber nicht. Die Ursache für das Funktionieren der Näherung liegt also offensichtlich in einer *Kompensation* der bei der Modifikation der Einzelausdrücke (4.67) bis (4.69) gemachten Fehler, vgl. Tab. 4.2. Abbildung 4.8 zeigt die gute Übereinstimmung der so generierten EXX-Potentiale mit denen einer exakten Rechnung.

Bei der nun folgenden Übertragung der Methode auf den variationellen Ansatz ist die Frage also vor allem: Wie können die obigen Punkte (i) – (v) in dem iterativen Zyklus von Abschnitt 4.5 „simuliert“ werden?

	E_{gap}^{Γ} [eV]			
			V_{Fock} symm.,	V_{Fock} nicht symm.,
	symm.	nicht symm.	$\chi_0, \Delta\rho_{\text{Fock}}$ nicht symm.	$\chi_0, \Delta\rho_{\text{Fock}}$ symm.
AlN	5.71	5.74	5.98	5.35
Ge	0.99	0.83	1.06	0.72
Si	3.23	3.06	3.32	2.97
GaAs	1.55	1.38	1.70	1.21

Tabelle 4.2: EXX-Bandlücken verschiedener Halbleiter am Γ -Punkt bei exakter Handhabung der Symmetrisierung (Spalte 2) und mit dem symmetriegenäherten Verfahren (Spalte 3). Wird V_{Fock} symmetrisiert, χ_0 und $\Delta\rho_{\text{Fock}}$ hingegen nicht, sind die Bandlücken stets etwas zu groß (Spalte 4), im umgekehrten Fall zu klein (Spalte 5). Es wurden KLI-Pseudopotentiale benutzt und ferner der Korrelationsanteil des LDA-Funktionalen verwendet (außer für GaAs, wo keine Korrelation berücksichtigt wurde).

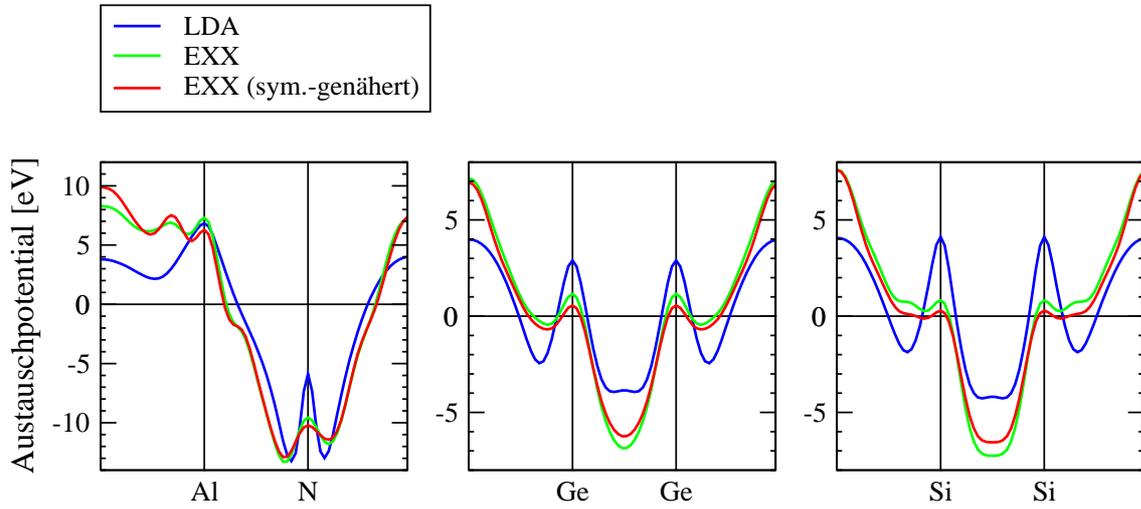


Abbildung 4.8: EXX-Potentiale von AlN (links), Ge (mitte) und Si (rechts) bei exakter Handhabung der Symmetrie (grün) sowie bei Verwendung des symmetriegenäherten Verfahrens (rot). Zum Vergleich ist auch das LDA-Potential mit eingezeichnet (blau).

4.6.3 Übertragung auf das variationelle EXX-Verfahren

Das Minimierungsprinzip (2.6) — der Ausgangspunkt der variationellen EXX-Formulierung — stellt letztlich nur eine sehr verknappte Form der OEP-Gleichung (1.96) dar (vgl. Abschnitt 2.2). Um nun ein analoges Prinzip für das symmetriegenährte EXX zu entwerfen, erscheint es daher am einfachsten, auch hier nicht von der expliziten EXX-Gleichung (4.66) auszugehen, sondern die äquivalente OEP-Gleichung zu betrachten. Dies hat insbesondere den Vorteil, dass χ_0 dabei nicht invertiert wird, und dessen Änderungen daher mit denen der anderen Größen zusammengefasst werden können, wie detailliert in Anhang F beschrieben. So findet man schließlich, dass die in (2.4) definierte induzierte Ladungsdichte $\Delta\rho$ im Variationsprinzip (2.6) durch folgende „modifizierte“ induzierte Ladungsdichte ersetzt werden muss:

$$\widetilde{\Delta\rho}(\mathbf{r}) := \sum_{s=1}^{N_{\text{red}}} w_s \left(\psi_{i,\mathbf{k}_s}^0(\mathbf{r})^* \psi_{i,\mathbf{k}_s}^1(\mathbf{r}) + \psi_{i,\mathbf{k}_s}^0(\mathbf{r}) \widetilde{\psi_{i,\mathbf{k}_s}^1(\mathbf{r})^*} \right). \quad (4.70)$$

Wie oben bezeichnen die \mathbf{k}_s mit den zugehörigen Gewichten w_s einen geeigneten \mathbf{k} -Punktsatz zur numerischen Integration über einen irreduziblen Bereich der Brillouin-Zone. Abgesehen davon, dass $\widetilde{\Delta\rho}$ nicht symmetrisiert ist, sind die linearen Korrekturen entscheidend abgeändert: Statt wie eigentlich im variationellen EXX-Verfahren *eine* Störung der Form $H_1 := V_{\text{Fock}} - v_x$ zu betrachten, sind die Korrekturfunktionen nun durch *zwei verschiedene* Störungen definiert. Die ψ_{i,\mathbf{k}_s}^1 ergeben sich dabei aus einer Störung des Kohn-Sham-Systems mit $H_1 := V_{\text{Fock}} - 2v_x$, und die $\widetilde{\psi}_{i,\mathbf{k}_s}^1$ erhält man aus Störung mit $\widetilde{H}_1 := V_{\text{Fock}}$. Daher sind die beiden Summanden in (4.70) im Allgemeinen nicht komplex-konjugiert zueinander, und $\widetilde{\Delta\rho}$ ist insgesamt eine komplexwertige Größe. Wird $\widetilde{\Delta\rho}$ nun im weiteren Verlauf der Rechnung mit dem Austauschpotential gemischt (vgl. Abschnitt 4.3) erhält man in der Folge ein komplexwertiges und nicht der Kristallsymmetrie entsprechendes Potential in Übereinstimmung mit den Betrachtungen des vorangehenden Abschnitts.

Dort allerdings erfolgt gemäß den Punkten (iv) und (v) in jedem Schritt eine explizite Symmetrisierung und ein Wegschneiden der Imaginäranteile des Potentials. Wie in Abschnitt 4.5.2 gezeigt, entsprechen allerdings die Iterationen von v_x beim Standard-EXX nicht unmittelbar den Aktualisierungen von v_x beim variationellen Verfahren, sondern nur jenen der „äußeren Schleife“. Damit gelingt es, die Punkte (i)–(v) des vorangehenden Abschnitts genau auf die variationelle Vorgehensweise zu übertragen: in der *inneren Schleife* von Flussdiagramm 4.5 wird dabei $\widetilde{\Delta\rho}$ aus (4.70) minimiert (es sind dazu natürlich *zwei* Sternheimer-Gleichungen nötig), wobei ein komplexes Austauschpotential zugelassen ist. Dies entspricht genau dem im vorigen Abschnitt mit (i)–(iii) erzeugten v_x . Bevor dann allerdings mit dem Lösen der Kohn-Sham-Gleichung die nächste Iteration der *äußeren Schleife* beginnt, wird dieses Austauschpotential entsprechend den obigen Punkten (iv) und (v) zu einem Potential mit physikalisch „vernünftiger“ Gestalt (reell und symmetrisch) modifiziert.

Kapitel 5

Variationelles und Standard-EXX: Ein Leistungsvergleich

Wie bereits in der Einleitung formuliert, bestand das Hauptanliegen der vorliegenden Arbeit in der Implementierung des variationellen EXX-Verfahrens [78, 79, 80] in das Ebene-Wellen-DFT-Paket S/PHI/nX [81, 82], so dass es erstmalig auch an periodischen Systemen getestet werden kann. Hier soll nun abgeschätzt werden, wie groß der daraus resultierende Vorteil gegenüber der Standard-Methode (Abschnitt 1.6.1) in Bezug auf die benötigten technischen Anforderungen (CPU-Zeit, Speicherbedarf) tatsächlich ist. Das Ergebnis einer Anwendung beider Methoden auf eine Reihe von Halbleitern ist in Tab. 5.1 zusammengefasst. In allen Rechnungen wurde dabei die Symmetrie mit der in Abschnitt 4.6 besprochenen Näherung behandelt.

Schon bei vergleichsweise geringen Abschneideenergien (etwa bei GaAs mit $E_{\text{cut}} = 15$ Ry und $E_{\text{cut}}^{\text{Fock}} = 12$ Ry) findet man bzgl. der pro Schritt benötigten CPU-Zeit eine Reduktion um den Faktor 0.75. Der Speicherbedarf reduziert sich sogar ca. um die Hälfte. Der Vorsprung wird noch größer bei hohen Abschneideenergien: So konnte für GaN (mit $E_{\text{cut}} = 70$ Ry und $E_{\text{cut}}^{\text{Fock}} = 52$ Ry) die Zeit etwa halbiert und der Speicherbedarf auf ein Drittel reduziert werden.

Der Aufwand zur Berechnung des nicht-lokalen Austauschoperators V_{Fock} ist in beiden Verfahren identisch. Im Standard-Verfahren macht dies ca. 50% der CPU-Zeit pro Schritt aus, im variationellen Verfahren hingegen sind dies typischerweise zwischen 80% und 90%. Rechnet man den Anteil von V_{Fock} heraus, ist der variationelle Ansatz damit sogar bis zu einem Faktor 10 mal schneller als die Standard-Methode. V_{Fock} ist damit zugleich als entscheidender Engpass der neuen Methode identifiziert.

Die Zahlen von Tab. 5.1 lassen sich also wie folgt zusammenfassen: (i) Der numerische Aufwand für EXX lässt sich durch Verwendung des variationellen Ansatzes deutlich reduzieren. (ii) Eine weitere Reduktion des Aufwandes kann nur durch eine Optimierung in der Berechnung des nicht-lokalen Austauschoperators erfolgen (vgl. dazu auch Abschnitt 4.4.2).

GaAs (15 Ry / 12 Ry)	Zeit / Schritt	Zeit V_{Fock}	Speicher
LDA	2 s	–	13 MB
Std. EXX	120 s = 2 min	50 s	35 MB
Var. EXX	90 s = 1.5 min	50 s	19 MB
Speed-Up: mit (ohne) V_{Fock} Reduktionsfaktor	1.33 (1.75)		0.54
Si (30 Ry / 23 Ry)	Zeit / Schritt	Zeit V_{Fock}	Speicher
LDA	2.6 s	–	53 MB
Std. EXX	1020 s = 17 min	480 s = 8 min	180 MB
Var. EXX	600 s = 10 min	480 s = 8 min	99 MB
Speed-Up: mit (ohne) V_{Fock} Reduktionsfaktor	1.69 (4.55)		0.55
Ge (40 Ry / 30 Ry)	Zeit / Schritt	Zeit V_{Fock}	Speicher
LDA	4.2 s	–	66 MB
Std. EXX	3900 s = 66 min	2040 s = 34 min	466 MB
Var. EXX	2280 s = 38 min	2040 s = 34 min	196 MB
Speed-Up: mit (ohne) V_{Fock} Reduktionsfaktor	1.72 (7.14)		0.42
AlN (60 Ry / 45 Ry)	Zeit / Schritt	Zeit V_{Fock}	Speicher
LDA	3.3 s	–	60 MB
Std. EXX	2280 s = 38 min	1200 s = 20 min	293 MB
Var. EXX	1380 s = 23 min	1200 s = 20 min	145 MB
Speed-Up: mit (ohne) V_{Fock} Reduktionsfaktor	1.64 (5.88)		0.49
GaN* (70 Ry / 52 Ry)	Zeit / Schritt	Zeit V_{Fock}	Speicher
LDA	3.7 s	–	78 MB
Std. EXX	6480 s = 1.8 h	3000 s = 50 min	710 MB
Var. EXX	3360 s = 56 min	3000 s = 50 min	227 MB
Speed-Up: mit (ohne) V_{Fock} Reduktionsfaktor	1.92 (10.0)		0.32

**Semi-core*-Schale über das Pseudopotential berücksichtigt

(→ S. 94)

Tabelle 5.1: Leistungsvergleich zwischen variationellem und Standard-EXX-Verfahren für einige ausgewählte Volumensysteme. Die verwendeten Abschneideenergien E_{cut} und $E_{\text{cut}}^{\text{Fock}}$ sind jeweils in Klammern angegeben. Aufgelistet sind die CPU-Zeit pro Schritt, die pro Schritt für V_{Fock} aufgewendete CPU-Zeit (in beiden Verfahren identisch) sowie der maximal benötigte Speicherplatz. Ferner ist für die Zeit der Speed-Up-Faktor sowie für den Speicherbedarf der Reduktionsfaktor im Vergleich zum Standard-Verfahren angegeben. Ebenfalls sind die entsprechenden Werte einer LDA-Rechnung aufgeführt.

Kapitel 6

Physikalische Bedeutung der induzierten Ladungsdichte

Vergleicht man Ladungsdichteverteilungen miteinander, die sich aus einer DFT-LDA-Rechnung und einer DFT-EXX-Rechnung ergeben, so zeigt sich, wie bereits in Ref. 52 beschrieben, ein allgemeiner Trend: Die Ladungsdichte ist in EXX stärker lokalisiert als in LDA. Beim Halbleiterkristall ist dies besonders deutlich auf den Bindungen zu erkennen (vgl. Abb. 6.2 oben, 6.3 und 6.4 auf den folgenden Seiten), für Atome ist ein ähnliches Verhalten bekannt [44, 45, 72, 73]. Diese Beobachtung lässt sich leicht deuten, wenn man bedenkt, dass im EXX/OEP-Formalismus die elektronische Selbstwechselwirkung vollständig eliminiert ist, vgl. Abschnitt 1.6. In LDA, wo dies nicht der Fall ist, stellt die Selbstwechselwirkung ein (unphysikalisches) Kraftfeld dar, durch das die Elektronen in Regionen hoher Ladungsdichte zu weit voneinander „abgedrängt“ werden: die Ladungsdichte erscheint dadurch homogener. Folglich führt eine Herausnahme dieser Kraft durch EXX/OEP dann umgekehrt zu einer stärkeren Lokalisierung.

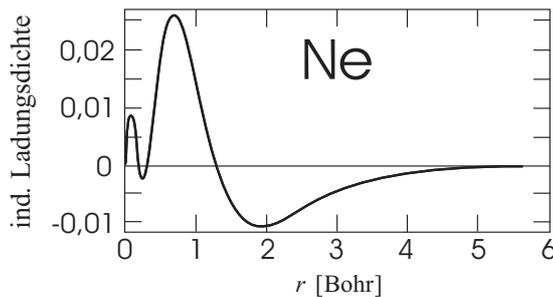


Abbildung 6.1: Radialanteil $r^2\Delta\rho[v_x^{\text{LDA}}]$ der induzierten Ladungsdichte beim Ne-Atom. Die Dichte konzentriert sich stärker in Richtung des Kerns. (Nach Wahn und Neugebauer [78])

Gleichung (2.4) die Ladungsverschiebung beim Übergang $v_x^{\text{LDA}} \rightarrow v_{\text{EXX}}$ nur in 1. Ordnung beschreibt, ist darin dennoch bereits alle Information für eine *vollständige* Herausnahme der Selbstwechselwirkung enthalten. Die Begründung liegt in der Linearität des EXX-Formalismus: mit Gleichung (4.42) hat man $v_{\text{EXX}}[\rho_{\text{LDA}}] = v_x^{\text{LDA}} + \hat{\chi}_0^{-1}\Delta\rho[v_x^{\text{LDA}}]$, wobei die Berechnung von $\hat{\chi}_0^{-1}$ im variationellen EXX-Verfahren natürlich durch einen geeigneten *Mixing*-Prozess ersetzt wird, vgl. Abschnitt 4.3. Im daraus konstruierten Kohn-Sham-Potential $v_{\text{ext}} + v_{\text{H}}[\rho_{\text{LDA}}] + v_{\text{EXX}}[\rho_{\text{LDA}}] + v_{\text{c}}[\rho_{\text{LDA}}]$ ist dann die Selbstwechselwirkung exakt

Als die Größe, die im variationellen EXX-Verfahren für die Herausnahme der Selbstwechselwirkung „sorgt“, lässt sich die induzierte Ladungsdichte $\Delta\rho[v_x]$ identifizieren. Nach dem Variationsprinzip (2.6) verschwindet sie nur für $v_x = v_{\text{EXX}}$. Setzt man dagegen $v_x = v_x^{\text{LDA}}$, so illustriert nebenstehende Abb. 6.1 recht anschaulich den verursachten Effekt am Beispiel des Ne-Atoms: $\Delta\rho[v_x^{\text{LDA}}]$ verschiebt Ladungsdichte von äußeren Bereichen hin zum Kern, die Lokalisierung der Dichte nimmt dadurch insgesamt zu. Obwohl nun $\Delta\rho[v_x^{\text{LDA}}]$ gemäß ihrer Konstruktion nach

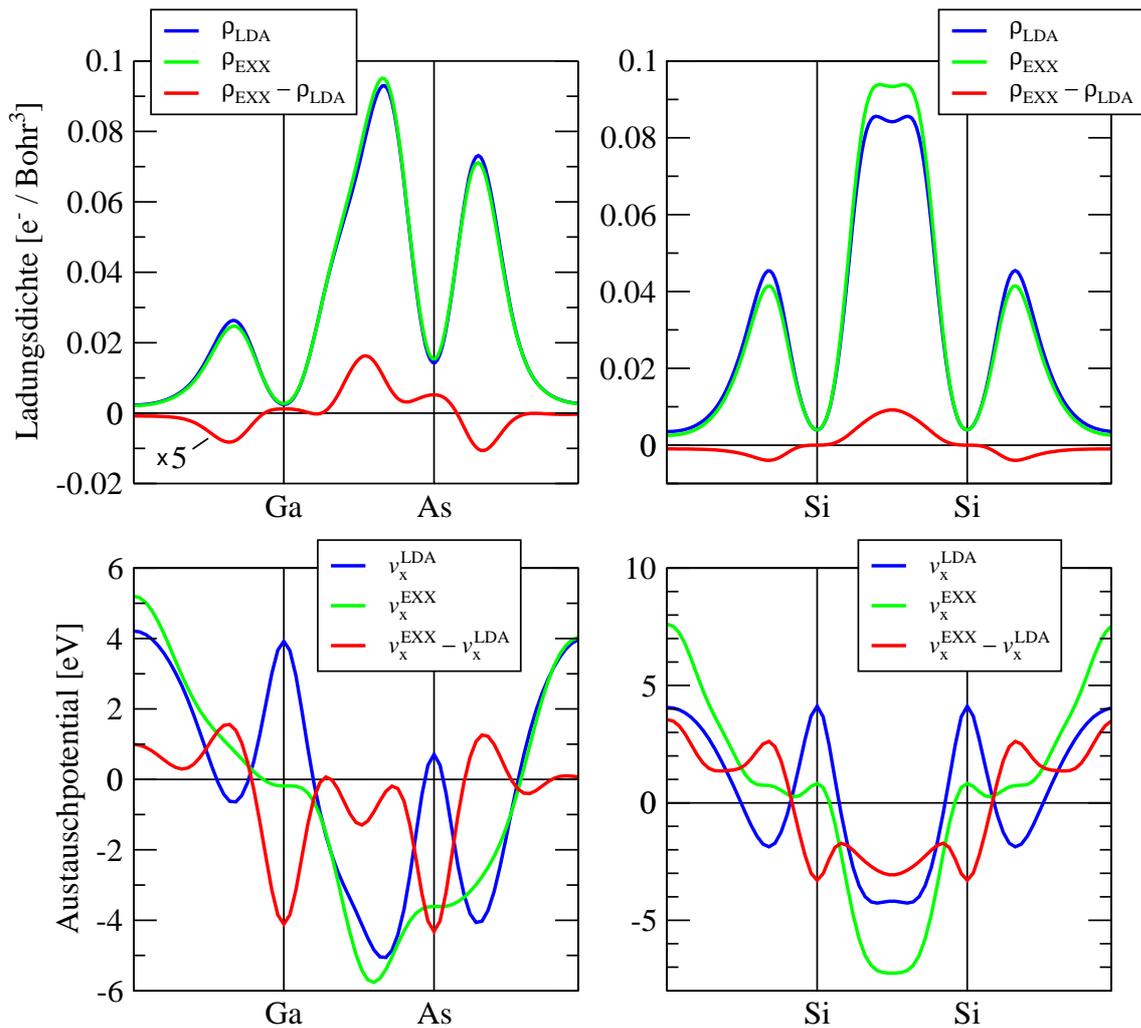


Abbildung 6.2: Schnitt entlang der (1,1,1)-Richtung durch die Einheitszelle von GaAs (*links*) und Si (*rechts*). *Oben:* Ladungsdichten in LDA und EXX sowie deren Differenzen (Differenz bei GaAs um den Faktor 5 vergrößert dargestellt). *Unten:* Austauschpotentiale in LDA und EXX sowie deren Differenzen.

eliminiert,¹ wengleich auch die Kohn-Sham-Gleichung noch nicht selbstkonsistent gelöst ist.

Durch das selbstkonsistente Lösen der Kohn-Sham-Gleichung wird die Dichte nun noch in nicht-linearer Weise modifiziert (vgl. Abschnitt 4.3), diese „Nachjustierungen“ sind jedoch bei den hier untersuchten Systemen im Verhältnis zu $\Delta\rho[v_x^{\text{LDA}}$] um etwa eine Größenordnung kleiner. Abbildung 6.2 (oben) zeigt die selbstkonsistente EXX-Ladungsdichte (grün) im Vergleich mit der LDA-Ladungsdichte (blau) für die Halbleiter GaAs und Si. Bemerkenswert ist, dass die Ladungsverschiebung (rot) beim Übergang LDA \rightarrow EXX äußerst klein gegenüber der eigentlichen Ladungsverteilung ist, während die entsprechende Differenz zwischen EXX- und LDA-Austauschpotential etwa von gleicher Größenordnung wie das Austauschpotential selbst ist (Abb. 6.2 unten). Der gleiche Trend wurde in Ref. 78 für Atome beschrieben. Die geringe Größenordnung der induzierten Ladungsdichte

¹Unter der Bedingung, dass die Korrelation keine zusätzlichen Fehler ins Spiel bringt, also etwa beim „exchange-only“ EXX-Verfahren.

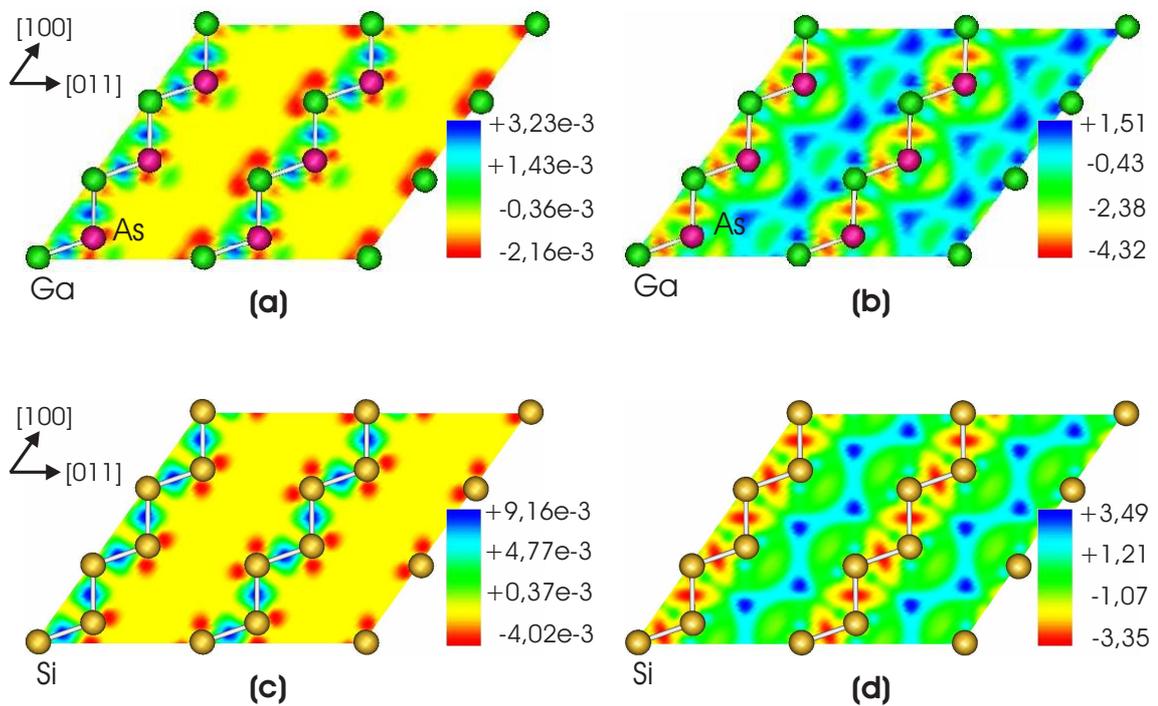


Abbildung 6.3: Differenz zwischen EXX- und LDA-Ladungsdichte in e^-/Bohr^3 (*links*) sowie zwischen EXX- und LDA-Austauschpotential in eV (*rechts*) für GaAs (*oben*) und Si (*unten*). In EXX ergibt sich vor allem auf den Bindungen eine stärkere Lokalisierung der Dichte als in LDA, entsprechend ist das Austauschpotential in EXX gegenüber dem einer LDA-Rechnung an diesen Stellen abgesenkt.

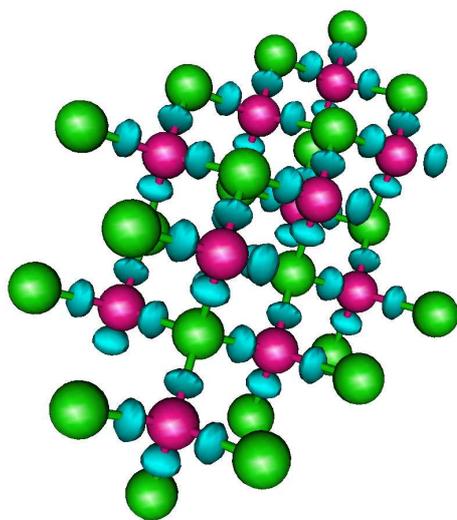


Abbildung 6.4: Isoberfläche der Differenz zwischen EXX- und LDA-Ladungsdichte in GaAs: $\rho_{\text{EXX}}(\mathbf{r}) - \rho_{\text{LDA}}(\mathbf{r}) = +1.4 e^-/\text{Bohr}^3$. Wie in den vorangehenden Abbildungen 6.2 und 6.3 erkennt man auch hier die Ladungverschiebung hin zu den Bindungen.

te ist konsistent mit der allgemein anerkannten Aussage, dass LDA- und Hartree-Fock-Wellenfunktionen sehr ähnlich sind, vgl. z.B. Ref. 58 und 59.

Ferner wird an Abb. 6.2 auch der negativ-definite Charakter von χ_0 erkennbar: Die Differenz der Austauschpotentiale (unten, rot) ergibt sich grob, außer in der Nähe der Kerne, aus der entsprechenden Ladungsdichtedifferenz (oben, rot) durch Spiegelung an der Abszisse und Streckung in Richtung der Ordinate. Auch Abbildung 6.3 illustriert dieses Verhalten. Natürlich liegen gerade in den Abweichungen von einer exakten Spiegelung die wesentlichen Eigenschaften — und Schwierigkeiten — des EXX-Formalismus, denn χ_0 kann eben nicht einfach durch eine geeignete Konstante genähert werden.

Kapitel 7

EXX-Pseudopotentiale: Linearisierung der Austauschwechselwirkung

Die gegenüber der Standard-EXX-Methode numerisch (hinsichtlich der CPU-Zeit und des Speicherbedarfs) deutlich effizientere variationelle EXX-Methode erlaubt die Untersuchung von Fragestellungen, die bisher nicht angegangen werden konnten. Als Beispiel soll hier untersucht werden, ob sich das EXX-Potential in Bezug auf die Ladungsdichte als annähernd lineares Funktional betrachten lässt. Von großer Bedeutung ist dies speziell im Hinblick auf die Verwendung von EXX-Pseudopotentialen, bei denen ein Teil der *semi-core*-Elektronen explizit als Valenzelektronen behandelt wird, der andere Teil jedoch im Rumpf verbleibt.

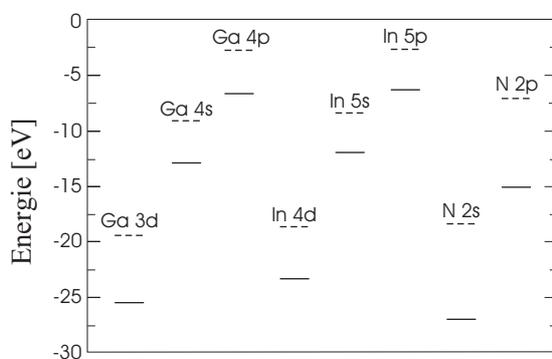


Abbildung 7.1: Eigenwerte von Ga-, In- und N-Pseudoatomen. *Getrichelte Linien:* LDA-Berechnungen mit LDA-Pseudopotentialen. *Durchgezogene Linien:* EXX-Berechnungen mit EXX-Pseudopotentialen. Nach Ref. 169.

So verfährt man üblicherweise etwa bei den Materialien GaN und InN, bei denen die Zuordnung der kationischen *d*-Elektronen zum Valenzbereich auf Grund ihrer Hybridisierung mit den *2s*-Zuständen des Stickstoffs notwendig ist (vgl. Abbildung 7.1 und Abschnitt 3.4). Bei einem „Schnitt“ durch die *semi-core*-Schale, wobei deren *s*- und *p*-Orbitale dem Rumpf, die *d*-Orbitale hingegen der Valenz zugeordnet werden, gibt es jedoch eine starke Überlappung von Rumpf- und Valenzladungsdichte.

Es ist daher notwendig, Gültigkeit und/oder Grenzen der mit der Pseudoisierung einhergehenden Linearisierung des verwendeten Austausch-Korrelationspotentials bezüglich Rumpf- und Valenzladungsdichte zu testen. Alternativ ließe sich im Prinzip auch die gesamte *semi-core*-Schale den Valenzzuständen zuordnen. Dies ist aber unter numerischen Gesichtspunkten wenig wünschenswert; insbesondere in EXX stellt die dann notwendige Erhöhung der Abschneideenergie wegen der ungünstigen Skalierung des nicht-lokalen Austauschoperators (vgl. Abschnitt 4.4) ein Problem dar. So lassen sich zwar Berechnungen von GaN und InN bei expliziter Berücksichtigung der *semi-core-d*-Elektronen mit

		EXX ^a		EXX ^b	
E_{cut}	$E_{\text{cut}}^{\text{Fock}}$	$E_{\text{gap}}^{\text{spd}}$	E_{gap}^d	$E_{\text{gap}}^{\text{spd}}$	E_{gap}^d
60	50	1.05	0.80	0.92	0.69
80	60	1.05	0.80	0.93	0.69
80	70		0.80	0.94	
100	60	1.04	0.80	0.93	0.69
100	70		0.80	0.94	
100	80	1.04	0.80	0.94	0.69
120	60	1.03	0.80	0.91	
120	80	1.02	0.80	0.91	0.69

Tabelle 7.1: Konvergenztest für InN. Die Bandlücken sind in der Einheit eV angegeben, die Abschneideenergien für die Wellenfunktionen E_{cut} sowie den nicht-lokalen Austauschoperator $E_{\text{cut}}^{\text{Fock}}$ in Ry. Die *semi-core*-Schale wurde einmal vollständig ($E_{\text{gap}}^{\text{spd}}$) und einmal nur mit den d -Elektronen (E_{gap}^d) in die Valenz miteinbezogen. EXX^a: LDA-Korrelation sowohl im Pseudopotential als auch in der expliziten Berechnung. EXX^b: LDA-Korrelation nur im Pseudopotential, nicht in der expliziten Berechnung.

dem Standard-EXX-Verfahren noch bewältigen (vgl. Ref. 169, 171), die Hinzunahme der s - und p -Zustände übersteigt bei vertretbarem technischen Aufwand jedoch das derzeit Machbare.

Der genannte „Schnitt“ durch den *semi-core*-Bereich wäre allerdings gerechtfertigt, wenn sich das Austausch-Korrelationspotential, oder zumindest dessen dominierender Austauschanteil, hinreichend linear bzgl. der Ladungsdichte verhält, also die Relation (3.30) durch

$$v_{\text{x}}[\rho_1 + \rho_2](\mathbf{r}) \approx v_{\text{x}}[\rho_1](\mathbf{r}) + v_{\text{x}}[\rho_2](\mathbf{r}) \quad (7.1)$$

ersetzt werden kann. Für ρ_1 und ρ_2 sind hier natürlich speziell die Rumpf- und die Valenzladungsdichte einzusetzen. Die Vermutung, dass das EXX-Potential (7.1) erfüllt, wird durch folgende Beobachtungen gestützt: (i) Die EXX/OEP-Gesamtenergie E_{EXX} stimmt stets in etwa mit der Hartree-Fock-Gesamtenergie E_{HF} überein, es gilt also $E_{\text{HF}} \lesssim E_{\text{EXX}}$ [50, 52, 60, 61]. (ii) Der nicht-lokale Austauschoperator kann ähnlich wie das elektrostatische Potential als lineares Funktional aufgefasst werden. Tatsächlich besteht der Unterschied formal nur darin, dass das elektrostatische Potential ein Funktional der Dichte ist, während der Kern des nicht-lokalen Austauschoperators vom Dichteoperator abhängt (siehe Gleichung (1.29)). In völliger Analogie zu (3.29) ergibt sich daher mit zwei Dichteoperatoren γ_1 und γ_2 :

$$v_{\text{Fock}}[\gamma_1 + \gamma_2](\mathbf{x}, \mathbf{x}') = -\frac{\gamma_1(\mathbf{x}, \mathbf{x}') + \gamma_2(\mathbf{x}, \mathbf{x}')}{|\mathbf{r} - \mathbf{r}'|} = v_{\text{Fock}}[\gamma_1](\mathbf{x}, \mathbf{x}') + v_{\text{Fock}}[\gamma_2](\mathbf{x}, \mathbf{x}') . \quad (7.2)$$

Hierbei beachte man, dass der Kern des nicht-lokalen Austauschoperators spin-abhängig ist, Orts- und Spinkoordinate sind in der Variablen \mathbf{x} zusammengefasst worden. Da sich nun ferner das (“exchange-only”) EXX/OEP-Potential als beste lokale Näherung an das nicht-lokale Austauschpotential des Hartree-Fock-Formalismus auffassen lässt (vgl. Abschnitt 1.6.2), so ist zumindest naheliegend, dass sich dessen Linearität auf das EXX-Potential „überträgt“. Wie gut die Beziehung (7.1) für ein gegebenes Austauschpotential

		EXX ^a		EXX ^b	
E_{cut}	$E_{\text{cut}}^{\text{Fock}}$	$E_{\text{gap}}^{\text{spd}}$	E_{gap}^d	$E_{\text{gap}}^{\text{spd}}$	E_{gap}^d
80	60		2.75	1.57	2.65
100	60	2.54	2.75	2.43	2.65
100	80	2.53	2.75	2.44	2.65
120	60	3.21	2.75	3.03	2.65
120	70			2.81	2.65
120	80	2.84	2.75	2.71	2.65

Tabelle 7.2: Konvergenztest für GaN. Bezeichnungen und Einheiten wie in Tab. 7.1. Die Abschneideenergie des nicht-linearen Austauschoperators $E_{\text{cut}}^{\text{Fock}}$ konnte auf Grund begrenzter Speicherkapazität nicht ganz auskonvergieren werden.

tatsächlich erfüllt ist, konnte mit der bisherigen Implementierung (stEXX) nicht getestet werden, da die Mitnahme der energetisch tief liegenden und damit stark lokalisierten s - und p -Orbitale der kationischen *semi-core*-Schale mit den zur Verfügung stehenden Computerclustern zu aufwendig war. Mit dem im Rahmen dieser Arbeit implementierten variationellen EXX-Verfahren hingegen ist dies nun möglich, indem zwei Rechnungen durchgeführt werden: einmal wird die *semi-core*-Schale nur bis zu den d -Elektronen, ein anderes Mal insgesamt in der Valenz berücksichtigt (im Weiteren seien dafür abkürzend die Bezeichnungen „ d -Rechnung“ bzw. „ spd -Rechnung“ verwendet). Ein solcher Vergleich liefert für das LDA-Potential eine gute Übereinstimmung der berechneten Bandlücken, obwohl LDA (sowie auch GGA) ein nicht-lineares Funktional in Bezug auf die Dichte darstellt (bei GaN findet man eine Differenz von ca. 0.15 eV, und InN ist sowohl für eine d - als auch eine spd -Rechnung metallisch, vgl. Tab. 7.3). Anders die GW -Methode [150, 151], deren Selbstenergieoperator ebenfalls nicht-linear ist. Diese Methode (aufgesetzt auf eine selbstkonsistente LDA-Berechnung) liefert für GaN gemäß Rohlfing *et al.* [152] einen deutlichen Unterschied in der Bandlücke: 0.71 eV. Mit dem variationellen EXX-Verfahren lässt sich ein solcher Vergleich nun erstmalig auch für das exakte Austauschpotential durchführen; die Ergebnisse sind in den Tabellen 7.1 und 7.2 in Form von Konvergenztests zusammengefasst und in Tab. 7.3 im Vergleich mit den Resultaten anderer Methoden sowie dem Experiment dargestellt.

Durchführung der Rechnung

Die Bandstrukturen wurden mit der in dieser Arbeit durchgeführten und in Kapitel 4 detailliert beschriebenen Implementierung des variationellen EXX-Verfahrens in den DFT-Code S/PHI/nX [81, 82] berechnet. Dabei wurde allerdings auf das in Abschnitt 4.6 beschriebene „symmetriegenäherte EXX-Verfahren“ zurückgegriffen, da bei den hier untersuchten Gruppe-III-Nitriden eine Symmetrisierung des nicht-lokalen Austauschoperators wegen seines Skalierungsverhaltens im reziproken Raum (vgl. Abschnitt 4.4) nicht mit vertretbarem Aufwand bewältigt werden kann, wenn die *semi-core*-Schale ganz oder teilweise in die Valenz miteinbezogen werden soll. Die Berechnungen wurden für die metastabile [153] Zinkblende-Phase durchgeführt. Frühere Berechnungen in LDA [154, 155] sowie EXX+ GW [169, 172] (s - und p -Elektronen im Rumpfbereich) legen jedoch nahe, dass sich die entsprechenden Bandlücken der stabilen Wurtzit-Phase daraus ablei-

Method	<i>semi-core</i>	InN	GaN	
LDA	<i>d</i>	-0.44 ^b	1.51 ^a	^a vorliegende Arbeit
LDA	<i>spd</i>	0.0 ^a	1.66 ^a	^b Ref. 168
EXX(cLDA)	–	1.49 ^k	3.52 ^k	^c Ref. 160, 161
EXX(cLDA)	<i>d</i>	0.8 ^a	2.75 ^a	^d Ref. 152
EXX(cLDA)	<i>spd</i>	1.02 ^a	2.84 ^a	^e Ref. 172
LDA+ <i>GW</i>	<i>d</i>	≈ 0.0 ^c (W)	3.59 ^d	^f Ref. 157, 158, 159
LDA+ <i>GW</i>	<i>spd</i>		2.88 ^d	^g Ref. 156
EXX(cLDA)+ <i>GW</i>	<i>d</i>	0.53 ^e	3.07 ^e	^h Ref. 162, 163, 164
Experiment		0.7–1.0 ^{f,j} (W)	3.2 ^h	ⁱ Ref. 165, 166
		1.9 ^g (W)	3.3 ^{i,j}	^j Ref. 167
				^k Ref. 169

Tabelle 7.3: Bandlücken von InN und GaN in DFT-EXX(cLDA), DFT-LDA und mit der Quasiteilchen-Methode (*GW*) im Vergleich mit dem Experiment. Einige Daten sind bei InN nur für die Wurtzite-Phase verfügbar, insbesondere auch die experimentelle Bandlücke; diese Werte sind hier durch ein W gekennzeichnet.

ten lassen, indem ≈ 0.2 eV auf die Werte der Zinkblende-Phase addiert wird. Um eine bessere Vergleichbarkeit mit dem Experiment und den Daten der meisten anderen EXX-Berechnungen zu ermöglichen, wurde jeweils die experimentelle Gitterkonstante angenommen (4.53 Å für GaN und 4.98 Å für InN). Es wurden normerhaltende relativistische EXX-Pseudopotentiale benutzt [111, 112], die einen LDA-Korrelationsanteil besitzen. Die selbst-konsistente Rechnung wurde sowohl mit als auch ohne LDA-Korrelation durchgeführt. Die Brillouin-Zonen-Integration erfolgte dabei auf einer ($4 \times 4 \times 4$)-Zerlegung nach Monkhorst und Pack [96].

Orbitale, Ladungsdichten, etc. wurden mit Hilfe von ebenen Wellen dargestellt, dabei konnte die Abschneideenergie $E_{\text{cut}}^{\text{Fock}}$ für den Austauschoperator unabhängig von der Abschneideenergie E_{cut} der Wellenfunktionen variiert werden. Mit einem zur Verfügung stehenden Arbeitsspeicher von maximal 4 GB ergab sich für $E_{\text{cut}}^{\text{Fock}}$ eine Obergrenze von etwa 80 Ry. Während sich die *spd*-Rechnungen für InN damit auf ca. 0.04 eV auskonvergieren lassen (Tab. 7.1), gelingt dies für GaN nicht mehr vollständig (Tab. 7.2). Insbesondere scheint in dem betrachteten Bereich noch eine $E_{\text{cut}}^{\text{Fock}}$ -Abhängigkeit zu bestehen: mit wachsendem $E_{\text{cut}}^{\text{Fock}}$ wird die Bandlücke geringer. Entsprechende *spd*-Rechnungen in LDA zeigen, dass bei GaN mindestens $E_{\text{cut}} \approx 140$ Ry notwendig ist; damit wäre gemäß Abschnitt 4.5.1 ein $E_{\text{cut}}^{\text{Fock}} \gtrsim 100$ Ry erforderlich. Die anschließende Diskussion soll in Bezug auf *GaN* daher als Hinweis verstanden werden, der unter dem Vorbehalt steht, dass endgültige Klarheit erst durch eine zukünftige genauere Rechnung erlangt werden kann. Dies lässt sich nach gegenwärtigem Ermessen am schnellsten durch einen Wechsel der Basisdarstellung des nicht-linearen Austausch-Operators erreichen (vgl. Abschnitt 4.4.2).

Ergebnisse und Diskussion

Aus Tab. 7.3 ist ersichtlich, dass sich die Bandlücken einer *d*-Rechnung und einer *spd*-Rechnung bei InN um etwa 0.2 eV voneinander unterscheiden. Für GaN kann der Differenzbetrag auf Grund des oben Genannten nicht präzise angegeben werden, die besten verfügbaren Rechnungen ($E_{\text{cut}} = 120$, $E_{\text{cut}}^{\text{Fock}} = 80$, vgl. Tab. 7.2) lassen jedoch hier sogar auf eine Differenz von weniger als 0.1 eV schließen. Beachtet man, dass für Bandstruktur-

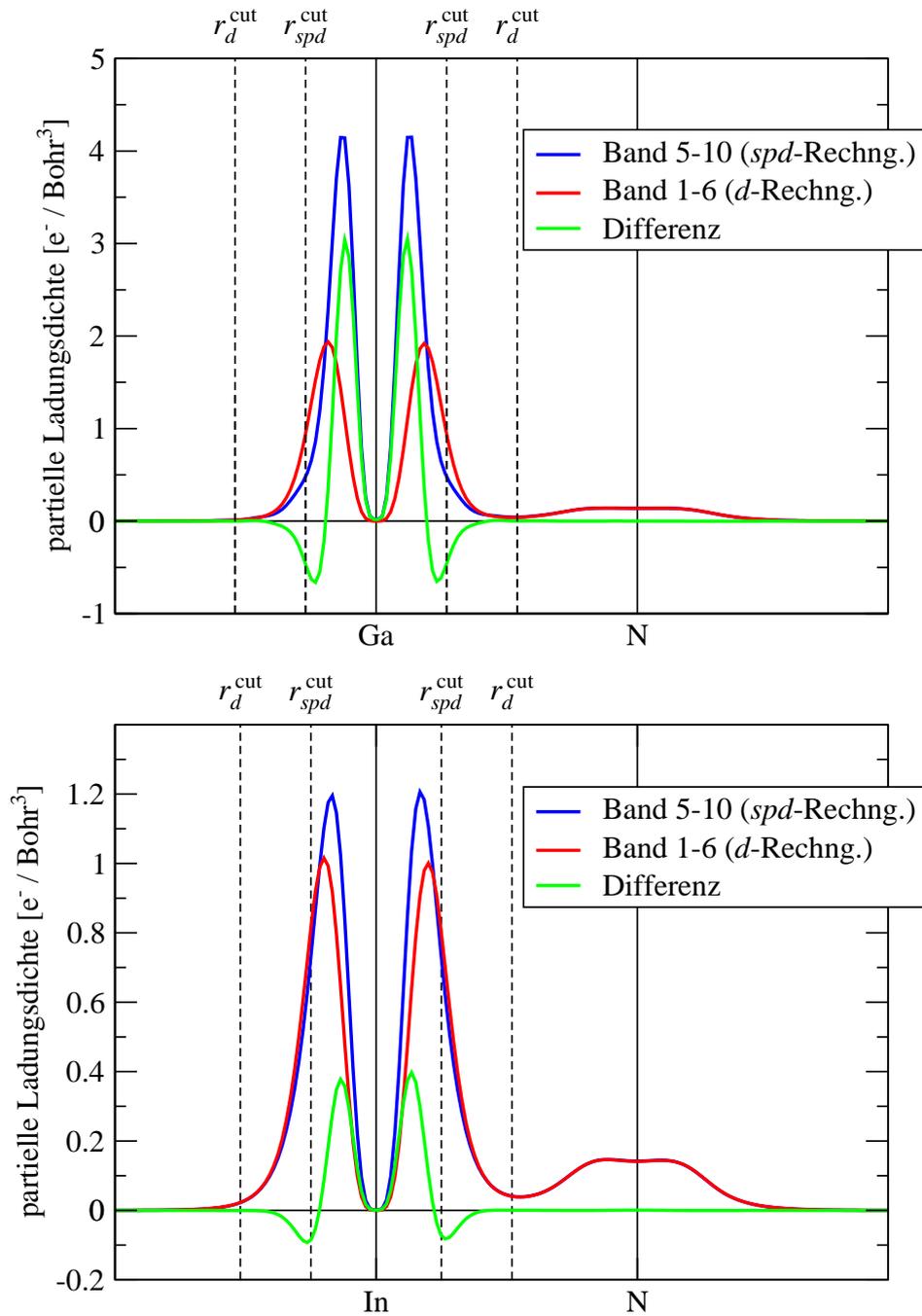


Abbildung 7.2: Partielle Ladungsdichten entlang der $(1,1,1)$ -Richtung durch die Einheitszelle von GaN (*oben*) und InN (*unten*). Im Bereich der Bindung und des N-Atoms stimmen die partiellen Ladungsdichten einer d - und einer spd -Rechnung weitgehend überein. Die Abweichungen im Bereich der Kationen werden durch Pseudopotentialeffekte verursacht, vgl. Abb. 7.3. Die Abschneideradien r_d^{cut} und r_{spd}^{cut} kennzeichnen den Bereich der kationischen Pseudopotentiale, die für die d - bzw. spd -Rechnung verwendet worden sind.

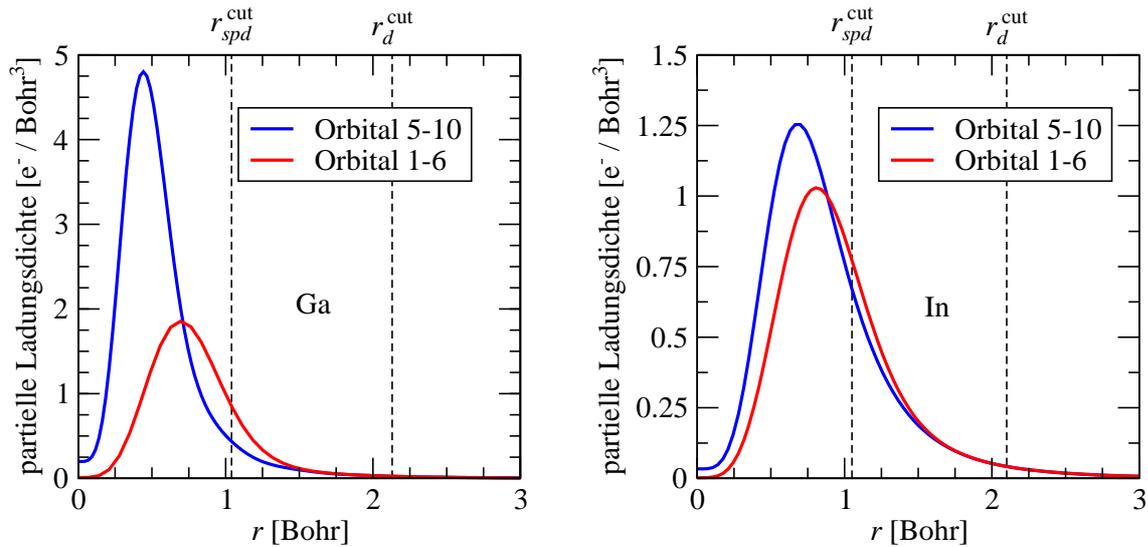


Abbildung 7.3: Partielle Ladungsdichten der $3d$ - und $4s$ -Elektronen des Ga-Atoms (*links*) bzw. der $4d$ - und $5s$ -Elektronen des In-Atoms (*rechts*). *Blaue Kurven:* Explizite Berücksichtigung aller *semi-core*-Elektronen. *Rote Kurven:* Die s - und p -Elektronen der *semi-core*-Schale wurden nur implizit über das Pseudopotential berücksichtigt. Wie in Abb. 7.2 kennzeichnen r_d^{cut} und r_{spd}^{cut} die Bereiche der Pseudopotentiale, die für die d - bzw. die spd -Rechnung benutzt worden sind. Innerhalb der Abschneideradien ergeben sich Unterschiede der berechneten Ladungsdichten, die in etwa den Differenzen entsprechen, die in Abb. 7.2 um die Kationen entstehen. Dies zeigt, dass es sich bei den Differenzen um reine Pseudopotentialeffekte handelt.

berechnungen mit Hilfe von Pseudopotentialen 0.2 eV ohnehin als Toleranz anzunehmen sind, wird die Vermutung eines linearen EXX-Potentials (7.1) durch diese Untersuchungen sehr gut bestätigt. Sieht man von dieser Toleranz ab, liefern die spd -Rechnungen in beiden Fällen eine geringe Öffnung der Bandlücke im Vergleich zu den entsprechenden d -Rechnungen. Im Falle von GaN ergibt sich damit eine leichte Verbesserung hin zum experimentellen Resultat (vgl. Tab. 7.3), für InN wird die Bandlücke (nach „Umrechnung“ auf die Wurtzitstruktur durch Addieren von 0.2 eV , siehe oben) im Vergleich zu den experimentellen Werten zwischen 0.7 und 1 eV [157, 158, 159] etwas überschätzt.

Die Linearität von v_{EXX} sollte sich andererseits auch an der Ladungsdichteverteilung überprüfen lassen, insbesondere an dem Anteil, der durch die Wellenfunktionen der d -Bänder gegeben ist. Wäre nämlich das EXX-Potential nicht linear, sollte die dann unzulässige Separation in Valenz- und Rumpfelektronen zu einer Verzerrung der tatsächlichen Gestalt dieser partiellen Dichte führen. Die s - und p -Bänder der *semi-core*-Schale sind sowohl für GaN als auch InN weit von den darüberliegenden d -Bändern entfernt; ihr Anteil kann daher leicht von der Gesamtladungsdichte einer spd -Rechnung abgezogen werden. Eine saubere Trennung der d -Bänder vom $2s$ -Band des Stickstoffs gelingt dagegen auf Grund der Hybridisierung der beteiligten Orbitale nicht. In Abb. 7.2 wird daher die partielle Ladungsdichte betrachtet, die sich aus den kationischen *semi-core-d*-Bändern und dem $N-2s$ -Band ergibt. Tatsächlich erweist sich diese Dichte, wie die Abbildung illustriert, für beide untersuchten Halbleiter sowohl im Bereich der Bindung als auch im Bereich des Stickstoff-Atoms als weitgehend unabhängig davon, ob sie mit einer d - oder einer spd -Rechnung erzeugt worden ist. In einem Bereich um das jeweilige Kation allerdings ergeben sich Differenzen. Gibt es also doch Gebiete, in denen sich v_{EXX} nicht linear verhält? Dazu beachte man, dass die beobachteten Differenzen auch von Pseudopotential-

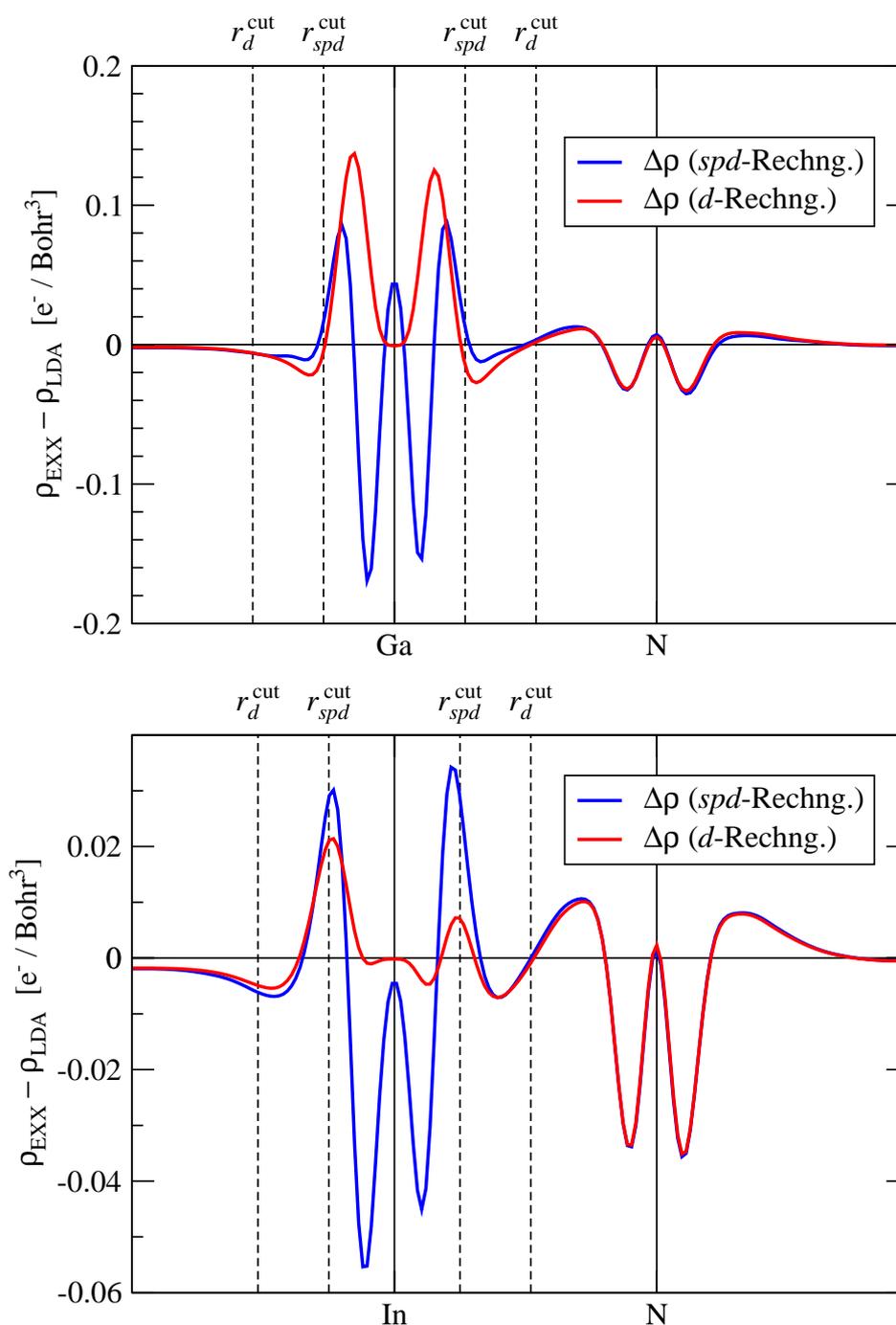


Abbildung 7.4: Differenz zwischen der EXX-Ladungsverteilung und der LDA-Ladungsverteilung entlang der (1,1,1)-Richtung durch die Einheitszelle von GaN (*oben*) und InN (*oben*). Ähnlich wie in Abb. 7.2 für die Dichten ergeben sich auch für die induzierten Ladungsdichten nur im Bereich der kationischen Pseudopotentiale (gekennzeichnet wie in Abb. 7.2) nennenswerte Unterschiede.

effekten herrühren können, sich also zumindest zu einem Teil auch als numerische Artefakte erklären lassen. Um dies zu verdeutlichen, wurden in Abb. 7.2 die Abschneideradien der verwendeten Pseudopotentiale mit eingezeichnet. Innerhalb der so markierten Pseudopotentialbereiche weichen aber nun die berechneten Dichteverteilungen bekanntermaßen von den tatsächlichen Dichten beträchtlich ab (vgl. Abb. 3.2). Folglich sind insbesondere bei Verwendung unterschiedlicher Pseudopotentiale (wie hier bei einer d - und einer spd -Rechnung) die berechneten (partiellen) Dichten innerhalb dieser Bereiche nicht miteinander vergleichbar. Dieser Effekt ergibt sich bereits bei der Berechnung des freien Atoms, wie Abb. 7.3 zeigt. Dies ist ein deutlicher Hinweis darauf, dass für die in Abb. 7.2 beobachteten Differenzen in der Tat zum großen Teil Pseudopotentialeffekte verantwortlich sind und daher nicht auf eine Nichtlinearität des EXX-Potentials in diesem Bereich geschlossen werden kann.

Ein ähnliches Bild ergibt sich, wenn man die Ladungsdichtedifferenzen zwischen einer EXX- und einer LDA-Rechnung betrachtet, vgl. Abb. 7.4. Außerhalb des Pseudopotentialbereichs findet man wiederum eine weitgehende Übereinstimmung zwischen der induzierten Ladungsdichte einer d -Rechnung und einer spd -Rechnung. Dabei zeigt sich wie bereits in den Resultaten von Kapitel 6, dass Ladungsdichte aus der Umgebung des Anions verschoben wird in den Bereich der Bindung, was als Herausnahme der elektronischen Selbstwechselwirkung interpretiert werden kann. Es ist zu erwarten, dass die Herausnahme der Selbstwechselwirkung auch zu einer stärkeren Lokalisierung der *semi-core*-Schale selbst führt, wie es etwa die induzierten Ladungsdichten der d -Rechnungen nahelegen (Abb. 7.4, rote Kurven). Da sich dieser Effekt jedoch im Bereich der hier benutzten Pseudopotentiale abspielt, wie abermals durch die eingezeichneten Abschneideradien verdeutlicht, lässt er sich in diesem Rahmen nicht hinreichend genau analysieren — dies ist allerdings für die untersuchte Fragestellung auch nicht erforderlich.

Zusammenfassend kann Folgendes gesagt werden: (i) Die vorliegenden Ergebnisse zeigen, dass es sich beim EXX-Potential um ein zumindest näherungsweise linear von der Dichte abhängendes Funktional handelt; dies folgt sowohl aus einem Vergleich der aus einer d - bzw. einer spd -Rechnung resultierenden Bandlücken (Differenz ca. 0.2 eV) sowie auch aus einem Vergleich der entsprechenden Ladungsdichten und induzierten Ladungsdichten. Die Gültigkeit von EXX-Rechnungen, die die s - und p -Zustände nur über das Pseudopotential berücksichtigen, wird damit bestätigt, was entscheidend für die künftigen Anwendungsmöglichkeiten der (trotz der variationellen Formulierung ja immer noch sehr teuren) EXX-Methode ist. (ii) Diese näherungsweise Linearität des EXX-Potentials führt bei der Verwendung von EXX-Pseudopotentialen zu einem gänzlich anderen Verhalten als die (nicht-lineare) GW -Methode, denn letztere erweist sich als stark abhängig davon, wie die Zuordnung der Elektronen zu Rumpf- und Valenzbereich erfolgt: je nach Zuordnung der kationischen s - und p -Zustände der *semi-core*-Schale ergibt sich bei GaN eine Differenz von 0.71 eV in der Bandlücke (GW basierend auf einer LDA-Rechnung) [152]. Für die Zukunft verbleibt zu untersuchen, wie sich eine GW -Korrektur auf die hier berechneten EXX-Bandlücken auswirkt.

Teil II

Wannier-Funktionen

Kapitel 8

Verallgemeinerte Wannier-Funktionen als lokalisierte *ab-initio*-Basis im Modell starker Bindung

In Kapitel 3 wurden — in Verbindung mit normerhaltenden Pseudopotentialen — die ebenen Wellen als eine geeignete Basis zur Beschreibung periodischer Systeme vorgestellt. Ebenen Wellen sind, ebenso wie die anderen in Abschnitt 3.3 erwähnten Basissätze (z.B. den LAPWs), vollständig in dem Sinne, dass sich die Qualität einer Rechnung systematisch durch Hinzunahme von mehr Basisfunktionen verbessern lässt. Für die Berechnung sehr großer Systeme, so z.B. von Quantenpunkten mit einer Anzahl von typischerweise 10^3 bis 10^5 Atomen, erweist sich die Verwendung einer vollständigen Basis jedoch als zu aufwendig; eine Behandlung des Problems mit der heute zur Verfügung stehenden Computertechnik wäre nicht machbar, insbesondere, wenn man an den optischen Eigenschaften interessiert ist und daher zur genauen Beschreibung der elektronischen Struktur der EXX-Formalismus herangezogen werden soll. In dieser Arbeit wurde daher eine *ab-initio*-Parametrisierung im Modell starker Bindung (engl.: *tight-binding approach*, TB) entwickelt [122] und in die DFT-Bibliothek S/PHI/nX [81, 82] eingebaut. Dieser Zugang erlaubt es, die Genauigkeit von Berechnungen auf atomistischer Skala auf Systeme mesoskopischer Größenordnung zu übertragen. Ferner sind die hier durchgeführten Untersuchungen auch als erste Vorarbeiten dahingehend zu verstehen, die Berechnung des nicht-lokalen Austauschoperators V_{Fock} — und damit die bei Verwendung des variationellen EXX-Verfahrens (Teil I) letzte größere Schwierigkeit der EXX-Methode — effizienter zu gestalten. Die Methode soll nun in diesem Kapitel vorgestellt werden.

Wie in der Einleitung bereits erwähnt, eröffnen den Ansatz dazu die von Marzari und Vanderbilt entwickelten *Verallgemeinerten Wannier-Funktionen* (GWFs, von amer.: *Generalized Wannier functions*) [116, 117], die im nächsten Abschnitt vorgestellt werden: Über eine unitäre Transformation mit der Bloch-Basis einer bestimmten, linearen Teilmenge $\mathcal{T} \subseteq \mathcal{H}$ des Hilbertraums \mathcal{H} verbunden¹, spannen die Verallgemeinerten Wannier-

¹Ist man beispielsweise — wie in dieser Arbeit — an den optischen Eigenschaften eines Halbleiterkristalls interessiert, wird \mathcal{T} sinnvollerweise aus den Bloch-Zuständen der (oberen) Valenzbänder und (mindestens) des untersten Leitungsbands aufgespannt.

Funktionen ebenfalls genau den Teilraum \mathcal{T} auf. Zur Darstellung von Funktionen oder Operatoren in \mathcal{T} sind Bloch- und Wannier-Basis also völlig gleichwertig, ein Wechsel zwischen ihnen stellt *keine Näherung* dar!

Dies beinhaltet aber auch, dass beide Basen gleich mächtig sind; wollte man also alle Matrixelemente eines Operators berechnen, hätte man durch den Basiswechsel noch nichts erreicht. Jedoch ist die unitäre Transformation nicht eindeutig festgelegt. Daraus ergeben sich Freiheiten, die zu unterschiedlichen Lokalisierungen der Verallgemeinerten Wannier-Funktionen führen. Wählt man die Basistransformation daher so, dass die Lokalisierung maximal ist, ist umgekehrt die räumliche Überlappung zweier ausreichend entfernt liegenden Funktionen minimal. Ein Operator in der Darstellung solcher *maximal lokalisierten Wannier-Funktionen* [116] sollte also nur für Matrixelemente benachbarter Basisfunktionen eine (wesentlich) von Null verschiedene Größenordnung aufweisen, alle anderen Elemente hingegen könnten vernachlässigt werden. Anders ausgedrückt, Verallgemeinerte Wannier-Funktionen sollten sich für die Verwendung als Basisfunktionen im Modell der starken Bindung eignen. Ob und wie gut dies gelingt, wird in den Abschnitten 8.2 bis 8.4 anhand von GaAs, GaN und InN aufgezeigt.

8.1 Verallgemeinerte Wannier-Funktionen

8.1.1 Definition

Die Verallgemeinerten Wannier-Funktionen ergeben sich aus einer Erweiterung des Konzepts der „klassischen“, von G. H. Wannier 1937 zur Beschreibung angeregter Zustände in kristallinen Isolatoren eingeführten und nach ihm benannten *Wannier-Funktionen* [118]. Letztere sind definiert als Fourier-Transformation eines Energiebandes der Bloch-Zustände $\psi_{n\mathbf{k}}$ in der Brillouin-Zone, wobei der Quasi-Wellenvektor \mathbf{k} in den Gittervektor \mathbf{R} überführt wird:²

$$|n\mathbf{R}\rangle := \frac{\Omega}{(2\pi)^3} \int_{\text{BZ}} e^{-i\mathbf{R}\cdot\mathbf{k}} |\psi_{n\mathbf{k}}\rangle d^3k, \quad (8.1)$$

wobei Ω das Volumen der Einheitszelle ist und das Integral über die Brillouin-Zone (BZ) ausgeführt wird. Tatsächlich birgt diese Definition Schwierigkeiten hinsichtlich der Normierung, vgl. Anhang H.3, um Analogie zu den hier zitierten Veröffentlichungen zu wahren, sei sie jedoch auch hier verwendet. Die inverse Relation ergibt sich daraus durch Multiplikation mit $e^{i\mathbf{R}\cdot\mathbf{k}'}$ und anschließende Summation über die Gittervektoren:

$$|\psi_{n\mathbf{k}}\rangle = \sum_{\mathbf{R}} e^{i\mathbf{R}\cdot\mathbf{k}} |n\mathbf{R}\rangle. \quad (8.2)$$

Da die Fourier-Transformation eine unitäre Transformation ist, bilden die Wannier-Funktionen ein vollständiges Orthonormalsystem im Hilbertraum \mathcal{H} :

$$\langle n\mathbf{R}|n'\mathbf{R}'\rangle = \delta_{n,n'}\delta_{\mathbf{R},\mathbf{R}'}, \quad \sum_{n=1}^{\infty} \sum_{\mathbf{R}} |n\mathbf{R}\rangle\langle n\mathbf{R}| = \mathbf{1}. \quad (8.3)$$

Nun sind (normierte) Eigenzustände eines Operators nur bis auf Phasenfaktoren genau bestimmt, daraus ergeben sich Freiheiten bei der Wahl der Bloch-Funktionen:

$$|\psi_{n\mathbf{k}}\rangle \longrightarrow e^{i\theta(n,\mathbf{k})} |\psi_{n\mathbf{k}}\rangle, \quad (8.4)$$

²Man beachte: Es wird nicht ein Blochzustand $|\psi_{n\mathbf{k}}\rangle$ fouriertransformiert, wie etwa in (3.18), sondern tatsächlich das *Band* in dem Sinne, dass (8.1) eine Linearkombination (Superposition) aus den Blochzuständen dieses Bandes darstellt.

wobei θ eine beliebige reelle Funktion von n und \mathbf{k} ist. Während $|e^{i\theta(n,\mathbf{k})}\psi_{n\mathbf{k}}(\mathbf{r})|^2$ allerdings invariant gegenüber der Wahl von θ ist, trifft dies für das Quadrat der resultierenden Wannier-Funktionen nicht mehr zu, vielmehr lässt sich ihre Form und räumliche Ausdehnung durch θ beeinflussen.

Eine praktische Schwierigkeit bei der Bestimmung der Wannier-Funktionen besteht darin, dass die Bänder einer elektronischen Struktur im Allgemeinen nicht isoliert sind, sondern an vielen Stellen Kreuzungen und Berührungspunkte (Entartungen) mit anderen Bändern aufweisen. Eine Gruppe³ solcher sich überlagernder Bänder stellen z.B. die oberen drei Valenzbänder von GaAs dar, während das untere Valenzband keine Entartungen enthält und daher ein isoliertes Band ist (vgl. Abb. 8.5). Nur auf Letzteres lässt sich daher (8.1) direkt anwenden, für die anderen stellt sich die Frage, auf welche Weise den Energien die Bandindizes zugeordnet werden sollen.

Durch eine Verallgemeinerung von (8.4) lässt sich dieses Problem jedoch umgehen. Sei dazu eine Gruppe von N Bändern betrachtet, die in der Bandstruktur keine gemeinsamen Punkte mit energetisch höher oder tiefer liegenden Bändern besitzt. O.B.d.A. sei der Bandindex m ohne Rücksicht auf die Gestalt der Bänder an jedem \mathbf{k} -Punkt in energetisch aufsteigender Reihenfolge festgelegt, wobei die Numerierung bei 1 beginne. Die Menge der entsprechenden Blochfunktionen spannt im Hilbertraum \mathcal{H} einen Unterraum $\mathcal{T} := \text{span}\{|\psi_{n\mathbf{k}}\rangle \mid n \in \{1, \dots, N\}, \mathbf{k} \in \text{BZ}\}$ auf. Zusätzlich zur Freiheit bei der Wahl der Phasenfaktoren sei nun auch eine *Durchmischung* der Bänder an einem \mathbf{k} -Punkt zugelassen: (8.4) wird erweitert zu

$$|\psi_{n\mathbf{k}}\rangle \longrightarrow \sum_{m=1}^N U_{mn}^{\mathbf{k}} |\psi_{m\mathbf{k}}\rangle, \quad (8.5)$$

wobei die $U_{mn}^{\mathbf{k}}$ einen Satz unitärer $(N \times N)$ -Matrizen darstellen. Durch diese so genannte *Eichtransformation* werden die Blochfunktionen also zunächst bzgl. des Bandindex transformiert. Anschließend erfolgt genau wie in (8.1) noch eine Fouriertransformation bzgl. \mathbf{k} . Damit ergeben sich insgesamt die *Verallgemeinerten Wannier-Funktionen* [116] zu

$$|n\mathbf{R}\rangle := \frac{\Omega}{(2\pi)^3} \int_{\text{BZ}} e^{-i\mathbf{R}\cdot\mathbf{k}} \sum_{m=1}^N U_{mn}^{\mathbf{k}} |\psi_{m\mathbf{k}}\rangle. \quad (8.6)$$

Die Hintereinanderschaltung unitärer Transformationen ist wiederum unitär, daher bleibt bei einem Basiswechsel gemäß (8.6) der Teilraum \mathcal{T} invariant. Unterteilt man die gesamte Bandstruktur geeignet in solche Gruppen von Bändern, gilt für die Verallgemeinerten Wannier-Funktionen daher ebenfalls (8.3).

Der Einfachheit halber soll sich der Begriff „Wannier-Funktion“ im weiteren Kontext dieser Arbeit immer auf diese Definition (8.6) beziehen. Funktionen im Sinne von (8.1) seien als „klassische Wannier-Funktionen“ bezeichnet. Natürlich sind sie ein Spezialfall von (8.6), wenn nämlich alle $U_{mn}^{\mathbf{k}}$ diagonal gewählt werden.

In Ortsdarstellung können alle Wannier-Funktionen zu festem n durch Translation ineinander überführt werden:

$$\langle \mathbf{r} | n\mathbf{R} \rangle = \langle \mathbf{r} - \mathbf{R} | n\mathbf{0} \rangle. \quad (8.7)$$

Daher reicht es häufig aus, nur die der „Heimatzelle“ $\mathbf{R} = \mathbf{0}$ zugeordnete Wannier-Funktion $|n\mathbf{0}\rangle$ zu betrachten. Der Beweis dieser Relation findet sich in Anhang H. Ferner ist

³Um Verwirrungen zu vermeiden: Der Begriff „Gruppe“ steht hier nicht mit dem gleichnamigen Begriff aus der Gruppentheorie in Zusammenhang.

die Ladungsdichte in einem Halbleiter- oder Isolatorokristall invariant gegenüber unitären Transformationen der Ein-Teilchen-Wellenfunktionen, es gilt also

$$\rho(\mathbf{r}) = \frac{\Omega}{(2\pi)^3} \sum_n^{\text{occ}} \int_{\text{BZ}} |\psi_{n\mathbf{k}}(\mathbf{r})|^2 = \sum_n^{\text{occ}} \sum_{\mathbf{R}} |\langle \mathbf{r} | n\mathbf{R} \rangle|^2 . \quad (8.8)$$

Lokalisierte Wannier-Funktionen lassen sich als Orbitale interpretieren, die denen des in der Chemie gebräuchlichen Bildes zur Beschreibung kovalenter oder ionischer Bindungen entsprechen. Verallgemeinerte wie auch klassische Wannier-Funktionen sind aber im Allgemeinen *keine* Eigenfunktionen des Hamilton-Operators mehr.

Die hier beschriebene Vorgehensweise funktioniert zunächst nur für eine Gruppe von Bändern, die durch eine Energielücke von anderen Bändern isoliert ist. Sind die betrachteten Bänder jedoch Teil einer größeren Gruppe zusammenhängender Bänder — z.B. bei der Betrachtung eines Metalls —, lässt sich mit einem in Ref. 117 vorgeschlagenen Algorithmus eine physikalisch sinnvolle Zuordnung des Bandindex erreichen: Beim Übergang von einem \mathbf{k} -Punkt zum nächsten wird die Zuordnung stets so gewählt, dass die Bänder dabei „möglichst stetig“ verlaufen. Der Schwerpunkt der hier durchgeführten Untersuchungen lag jedoch auf einer möglichst guten Beschreibung der Bandlücke in der Umgebung des Γ -Punktes, die Wahl der Bänder konnte dabei leicht manuell erfolgen.

8.1.2 Lokalisierung

Die Transformation (8.4) bzw. ihre Verallgemeinerung (8.5) beeinflussen die Gestalt und damit die Lokalisierung der Wannier-Funktionen. Um zur Generierung einer TB-Basis möglichst lokalisierte Wannier-Funktionen aufzufinden, muss „Lokalisierung“ zunächst geeignet definiert werden. In Ref. 116 wird dazu das Quadrat einer Wannier-Funktion im Ortsraum als Verteilung aufgefasst, deren Varianz bzgl. der Variablen \mathbf{r} dann zur Definition der Streuung

$$S(\{U^{\mathbf{k}}\}) := \sum_{n=1}^N [\langle n\mathbf{R} | r^2 | n\mathbf{R} \rangle - \langle n\mathbf{R} | \mathbf{r} | n\mathbf{R} \rangle^2] \quad (8.9)$$

verwendet wird. Dies ist die Summe der Varianzen aller N Wannier-Funktionen in der Einheitszelle. Je kleiner die Streuung S , umso größer ist die *Lokalisierung* der Wannier-Funktionen. Die Streuung wurde hier gleich als Funktion der $U^{\mathbf{k}}$ geschrieben, da sie über die Wannier-Funktionen implizit von der Wahl der Basistransformation (8.6) abhängt. Wegen (8.7) genügt es, diesen Ausdruck für $\mathbf{R} = \mathbf{0}$ auszuwerten.

Für das Weitere erweist es sich als vorteilhaft, die in (8.9) benötigten ersten beiden Momente der „Wannier-Verteilung“ im reziproken Raum darstellen zu können. Man findet für das erste Moment

$$\langle \mathbf{r} \rangle_n := \langle n\mathbf{0} | \mathbf{r} | n\mathbf{0} \rangle = i \frac{\Omega}{(2\pi)^3} \int_{\text{BZ}} \langle u_{n\mathbf{k}} | \nabla_{\mathbf{k}} | u_{n\mathbf{k}} \rangle d^3k \quad (8.10)$$

und für das zweite Moment

$$\langle r^2 \rangle_n := \langle n\mathbf{0} | r^2 | n\mathbf{0} \rangle = - \frac{\Omega}{(2\pi)^3} \int_{\text{BZ}} \langle u_{n\mathbf{k}} | \nabla_{\mathbf{k}}^2 | u_{n\mathbf{k}} \rangle d^3k , \quad (8.11)$$

wobei die $u_{n\mathbf{k}}(\mathbf{r}) = e^{-i\mathbf{k}\cdot\mathbf{r}} \psi_{n\mathbf{k}}(\mathbf{r})$ die periodischen Anteile der nach (8.5) transformierten Blochzustände bezeichnen [119]. $\langle \mathbf{r} \rangle_n$ kennzeichnet den „Schwerpunkt“ der Wannier-Funktion $|n\mathbf{0}\rangle$.

Um diese beiden Ausdrücke auch numerisch auf einem diskreten \mathbf{k} -Punkt-Gitter auswerten zu können, wird eine einfache Finite-Differenzen-Darstellung des Nabla-Operators $\nabla_{\mathbf{k}}$ benutzt, bei der für einen Punkt \mathbf{k} die Differenzen $\Delta\mathbf{k}_i$ zu den nächsten Nachbarstützstellen berücksichtigt werden. Eine solche Darstellung ist in Ref. 116 für kubische Systeme explizit angegeben. Setzt man diese Finite-Differenzen-Formeln in (8.10) und (8.11) ein, erhält man zwar im Limes immer dichter Auflösung der Brillouin-Zone ($\Delta\mathbf{k}_i \rightarrow \mathbf{0}$) wieder die exakten Ausdrücke, für diskrete Gitter jedoch erhält man bei einer Verschiebung der Blochfunktionen um einen Gittervektor \mathbf{R}_0 ,

$$\psi_{n\mathbf{k}}(\mathbf{r}) \longrightarrow \psi_{n\mathbf{k}}(\mathbf{r} - \mathbf{R}_0) = e^{-i\mathbf{R}_0 \cdot \mathbf{k}} \psi_{n\mathbf{k}}(\mathbf{r}),$$

nicht mehr die erwartete analoge Translation der Wannier-Funktionen

$$\langle \mathbf{r} | n\mathbf{R} \rangle \longrightarrow \langle \mathbf{r} - \mathbf{R}_0 | n\mathbf{R} \rangle. \quad (8.12)$$

Korrektes Transformationsverhalten kann aber auch bei den diskretisierten Ausdrücken erreicht werden, indem sie nämlich in höherer Ordnung der $\Delta\mathbf{k}_i$ modifiziert werden. Der Grenzübergang $\Delta\mathbf{k}_i \rightarrow \mathbf{0}$ zur exakten Differentiation wird dadurch nicht geändert. Diskretisierte Versionen von (8.10) und (8.11), die die Relation (8.12) erfüllen, sind gegeben mit

$$\langle \mathbf{r} \rangle_n \approx -\frac{1}{N} \sum_{i,\mathbf{k}} w_i \Delta\mathbf{k}_i \operatorname{Im} \ln M_{nn}^{\mathbf{k},\Delta\mathbf{k}_i} \quad (8.13)$$

und

$$\langle r^2 \rangle_n \approx \frac{1}{N} \sum_{i,\mathbf{k}} w_i \left\{ (1 - |M_{nn}^{\mathbf{k},\Delta\mathbf{k}_i}|^2) + (\operatorname{Im} \ln M_{nn}^{\mathbf{k},\Delta\mathbf{k}_i})^2 \right\}, \quad (8.14)$$

wobei i für jedes \mathbf{k} die nächsten Nachbarstützstellen abzählt, die jeweils noch mit einem Faktor w_i gewichtet werden, und die $M_{nn}^{\mathbf{k},\Delta\mathbf{k}_i}$ als Diagonalelemente der durch

$$M_{mn}^{\mathbf{k},\Delta\mathbf{k}_i} := \langle u_{m\mathbf{k}} | u_{n,\mathbf{k}+\Delta\mathbf{k}_i} \rangle \quad (8.15)$$

definierten *Überlapp-Matrix* gegeben sind. Sie sind ein Maß für die Unterschiedlichkeit des periodischen Anteils der Blochfunktionen beim Übergang von einem Punkt \mathbf{k} zum Nachbarpunkt $\mathbf{k} + \Delta\mathbf{k}_i$.

8.1.3 Maximale Lokalisierung

Um nun über ein Variationsverfahren zu maximal lokalisierten Wannier-Funktionen zu gelangen, muss der Gradient des Streufunktional (8.9) bzgl. der Transformationsmatrizen $U^{\mathbf{k}}$ bekannt sein. Für ein diskretes \mathbf{k} -Punkt-Gitter sei dessen Herleitung hier grob skizziert. Die Änderung der Streuung hängt über die ersten beiden Momente (8.13) und (8.14) von der Änderung der Überlapp-Matrix (8.15) ab. Sie lässt sich berechnen, indem zunächst „infinitesimal kleine“ Transformationen

$$\exp\{\Delta W^{\mathbf{k}}\} = \mathbf{1} + \Delta W^{\mathbf{k}} + \text{höhere Ordnungen in } \Delta W^{\mathbf{k}} \quad (8.16)$$

betrachtet werden. Damit die Transformationen unitär sind, müssen die $\Delta W^{\mathbf{k}}$ antihermitesch sein; es soll also $(\Delta W^{\mathbf{k}})^+ = -\Delta W^{\mathbf{k}}$ gelten, vgl. z.B. Ref. 125. In linearer Ordnung wird (8.5) damit zu

$$|\psi_{n\mathbf{k}}\rangle \longrightarrow |\psi_{n\mathbf{k}}\rangle + \sum_{m=1}^N \Delta W_{mn}^{\mathbf{k}} |\psi_{m\mathbf{k}}\rangle, \quad (8.17)$$

woraus sich die Änderung der benötigten Überlapp-Matrixelemente ergibt:

$$\begin{aligned} & \left(\langle u_{nk} | - \sum_{m=1}^N (\Delta W_{nm}^{\mathbf{k}})^* \langle u_{mk} | \right) \left(|u_{n,\mathbf{k}+\Delta\mathbf{k}} \rangle + \sum_{m=1}^N \Delta W_{mn}^{\mathbf{k}+\Delta\mathbf{k}} |u_{m,\mathbf{k}+\Delta\mathbf{k}} \rangle \right) \\ &= M_{nn}^{\mathbf{k},\Delta\mathbf{k}} + \underbrace{\sum_{m=1}^N M_{nm}^{\mathbf{k},\Delta\mathbf{k}} \Delta W_{mn}^{\mathbf{k}+\Delta\mathbf{k}} - \sum_{m=1}^N (\Delta W_{nm}^{\mathbf{k}})^* M_{mn}^{\mathbf{k},\Delta\mathbf{k}}}_{= \Delta M_{nn}^{\mathbf{k},\Delta\mathbf{k}}} + \text{quadr. Term.} \end{aligned} \quad (8.18)$$

Benutzt man dies in (8.13) und (8.14), erhält man schließlich, nach einigen algebraischen Umformungen und dem Grenzübergang $\Delta W^{\mathbf{k}} \rightarrow 0$, einen Ausdruck für den Gradienten der Streuung:

$$G^{\mathbf{k}} := \frac{dS}{dW^{\mathbf{k}}} = 4 \sum_i w_i \left[\mathcal{A}(R^{\mathbf{k},\Delta\mathbf{k}_i}) - \mathcal{S}(T^{\mathbf{k},\Delta\mathbf{k}_i}) \right]. \quad (8.19)$$

Dabei bezeichnen $\mathcal{S}(B) := (B + B^+)/2i$ und $\mathcal{A}(B) := (B - B^+)/2$ den symmetrischen bzw. antisymmetrischen Anteil einer Matrix B . Ferner gilt

$$R_{nm}^{\mathbf{k},\Delta\mathbf{k}} := M_{mn}^{\mathbf{k},\Delta\mathbf{k}} (M_{nn}^{\mathbf{k},\Delta\mathbf{k}})^*$$

und

$$T_{nm}^{\mathbf{k},\Delta\mathbf{k}} := \frac{M_{mn}^{\mathbf{k},\Delta\mathbf{k}}}{M_{nn}^{\mathbf{k},\Delta\mathbf{k}}} (\text{Im} \ln M_{nn}^{\mathbf{k},\Delta\mathbf{k}} + \Delta\mathbf{k} \cdot \langle \mathbf{r} \rangle_n).$$

Streng genommen ist (8.19) nicht der Gradient von S bzgl. der unitären Transformationen $U^{\mathbf{k}}$, sondern der Gradient im Raum ihrer (durch i dividierten) Erzeugenden⁴. Die Transformationen selbst sind dann analog zu (8.16) durch $\exp\{G^{\mathbf{k}}\}$ gegeben. Im Raum der Erzeugenden lässt sich nun zur Minimierung des Streufunktional das *Steepest-Descent*-Verfahren (vgl. Abschnitt 4.1.2.1) anwenden. Dazu folgt man iterativ dem Gradienten in kleinen Schritten $\varepsilon G_s^{\mathbf{k}}$. Im Raum der unitären Transformationen wiederum ergeben sich die Schritte dann rekursiv durch

$$U_{s+1}^{\mathbf{k}} := U_s^{\mathbf{k}} \exp\{\varepsilon G_s^{\mathbf{k}}\}. \quad (8.20)$$

Die Exponentialfunktion der antihermiteschen Matrizen $G_s^{\mathbf{k}}$ kann durch Diagonalisierung berechnet werden kann, vgl. z.B. Ref. 123.

Interessanterweise müssen während der Minimierung keine Aktualisierungen der Basisfunktionen vorgenommen werden: im Iterationszyklus werden nur die Größen $U_s^{\mathbf{k}}$, $G_s^{\mathbf{k}}$ und $M_s^{\mathbf{k},\Delta\mathbf{k}}$ benötigt. Anschließend können gemäß (8.6) die Wannier-Funktionen in einem Schritt aus den selbstkonsistenten $U^{\mathbf{k}}$ berechnet werden. Aktualisierungen der Überlapp-Matrix sind dabei durch

$$M_s^{\mathbf{k},\Delta\mathbf{k}} = (U_s^{\mathbf{k}})^+ M_0^{\mathbf{k},\Delta\mathbf{k}} U_s^{\mathbf{k}+\Delta\mathbf{k}} \quad (8.21)$$

gegeben. Diese letzte Beziehung macht das Verfahren sehr effizient, da — bis auf den Initialisierungsschritt — die aufwendige Berechnung nach (8.15) durch eine einfache Matrizenmultiplikation ersetzt wird. Eine Herleitung von (8.21) findet sich in Anhang H.2.

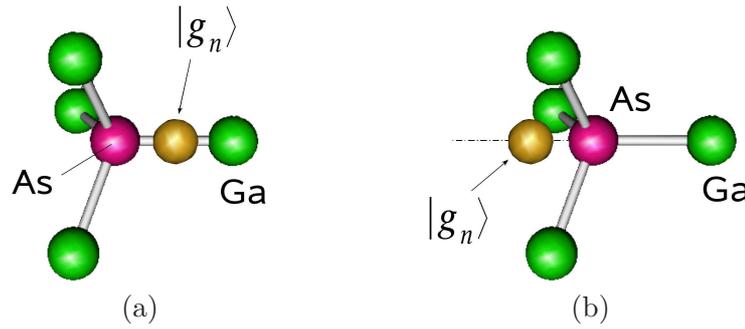


Abbildung 8.1: Das As-Atom und seine Bindungen zu den vier benachbarten Ga-Atomen im idealen GaAs-Kristall. Gauß-förmige Ansatzfunktionen $|g_n\rangle$ für die sp^3 -hybrid-artigen Wannier-Funktionen sind (a) auf der Bindung genau in der Mitte zwischen zwei Nachbaratomen bzw. (b) auf der der Bindung gegenüberliegenden Seite des Anions positioniert, die Positionen (a) und (b) gehen dabei gerade durch Spiegelung am As-Atom auseinander hervor. Es ist jeweils nur eine Ansatzfunktion gezeigt; bei der Initialisierung sind die anderen entsprechend der tetraedrischen Symmetrie angeordnet. Die sich aus diesen Ansätzen ergebenden Wannier-Funktionen sind in Abb. 8.2 gezeigt.

8.1.4 Initialisierung

Um die Konvergenz des Verfahrens zu beschleunigen und um Nebenminima des Streuungsfunktionals (8.9) zu vermeiden, sollte in (8.20) mit einem Startpunkt $U_0^{\mathbf{k}}$ begonnen werden, der sich in der Nähe des globalen Minimums befindet. Dazu macht man, geleitet beispielsweise durch die chemische Intuition, zunächst grobe Ansätze für die Wannier-Funktionen, die etwa durch geeignet positionierte, linear unabhängige Gauß-Funktionen g_n realisiert werden können, vgl. Abb. 8.1. Diese Ansatzfunktionen werden dann in zwei Schritten in eine Orthonormalbasis des interessierenden Unterhilbertraums $\mathcal{T} := \text{span}\{|\psi_{n\mathbf{k}}\rangle | n \in \{1, \dots, N\}, \mathbf{k} \in \text{BZ}\}$ (vgl. Abschnitt 8.1.1) überführt. Zunächst erfolgt dazu eine Projektion auf die Bloch-Basis:

$$|\phi_{n\mathbf{k}}\rangle = \sum_{m=1}^N |\psi_{m\mathbf{k}}\rangle \langle \psi_{m\mathbf{k}} | g_n \rangle. \quad (8.22)$$

Dies wird anschließend noch mit Hilfe einer Löwdin-Transformation orthonormalisiert:⁵

$$|\tilde{\psi}_{n\mathbf{k}}\rangle = \sum_{m=1}^N \left(\frac{1}{\sqrt{P}} \right)_{mn} |\phi_{m\mathbf{k}}\rangle \quad \text{mit} \quad P_{mn} := \langle \phi_{m\mathbf{k}} | \phi_{n\mathbf{k}} \rangle. \quad (8.23)$$

Da sowohl (8.22) als auch (8.23) Basistransformationen in \mathcal{T} sind und das Ergebnis eine Orthonormalbasis darstellt, folgt, dass ihre Hintereinanderschaltung unitär ist. Die Startpunkte in (8.20) sind damit durch

$$(U_0)_{mn}^{\mathbf{k}} := \sum_{\mu=1}^N \langle \psi_{m\mathbf{k}} | g_{\mu} \rangle \left(\frac{1}{\sqrt{P}} \right)_{\mu n} \quad (8.24)$$

⁴Unter der „Erzeugenden“ einer unitären Transformation U sei hier ein hermitescher Operator T mit $U = \exp\{iT\}$ verstanden. Der Operator iT ist dann antihermitesch, vgl. Gleichung (8.16).

⁵Die Löwdin-Transformation [120, 121] ist ein Orthonormalisierungsverfahren, bei dem die Symmetrieeigenschaften des ursprünglichen Basissatzes erhalten bleiben. Die Transformation (8.22) geschieht ganz analog dazu, wie etwa die (nicht-orthonormalen, aber in der Symmetrie des Bravais-Gitters angeordneten) atomaren Orbitale (AO) mittels der Löwdin-Transformation in LCAO-Orbitale überführt werden können. Physikalisch sinnvolle Ansätze $\{g_n\}$ für die Wannier-Basis werden im Allgemeinen ebenfalls die Symmetrie des Bravais-Gitters widerspiegeln.

gegeben. In der Praxis erhält man bereits allein durch Anwendung dieser $U_0^{\mathbf{k}}$ in (8.6) sehr lokalisierte Wannier-Funktionen. Der anschließende Iterationsprozess gemäß (8.20) „schiebt“ dann gewissermaßen nur noch kleinere Ausläufer der Wannier-Funktion in den Hauptbereich der Lokalisierung. Völlig lassen sich solche Nebenausläufer jedoch nicht vermeiden (vgl. z.B. Abb. 8.2, 8.6 und 8.7); sie sorgen für die Orthogonalität der Wannier-Funktionen untereinander.

8.2 *Ab-initio*-Parametrisierung im Modell starker Bindung

Als Parameter im Modell starker Bindung seien hier die Matrixelemente des Hamilton-Operators in der Darstellung einer stark lokalisierten Wannier-Basis bezeichnet. Mit der Definition der Verallgemeinerten Wannier-Funktionen (8.6) findet man die Matrixelemente

$$\langle n\mathbf{R}|H|n'\mathbf{R}'\rangle = \frac{\Omega^2}{(2\pi)^6} \sum_{m=1}^N \int e^{i\mathbf{k}\cdot(\mathbf{R}-\mathbf{R}')} (U_{mn}^{\mathbf{k}})^* U_{mn'}^{\mathbf{k}} \varepsilon_{m\mathbf{k}} d^3k, \quad (8.25)$$

wobei die $\varepsilon_{m\mathbf{k}}$ die (mit den Blochfunktionen konsistenten) Ein-Teilchen-Energien sind. Für $|\mathbf{R} - \mathbf{R}'| = 0$ erhält man Wechselwirkungen zwischen Wannier-Orbitalen, die der selben Einheitszelle \mathbf{R} zugeordnet sind, $|\mathbf{R} - \mathbf{R}'| > 0$ beschreibt Wechselwirkungen zwischen Atomen verschiedener Einheitszellen. Im Gegensatz zur Darstellung in klassischen Wannier-Funktionen verschwinden hier Elemente mit $n \neq n'$ im Allgemeinen nicht. Dies ist eine Folge der Durchmischung der Bänder gemäß (8.5) (vgl. Anhang H.3, wo auch die Herleitung von (8.25) angegeben ist).

Die Effizienz der Wannier-Funktionen als TB-Basis hängt nun davon ab, wieviele Wechselwirkungen bei einer Diagonalisierung des (Wannier-)Hamilton-Operators explizit berücksichtigt werden müssen, oder anders formuliert, ab welcher Distanz R_{\max} Wechselwirkungen zwischen verschiedenen Zellen vernachlässigt werden können. Ein geeignetes Kriterium, um die Genauigkeit des Parametersatzes in Abhängigkeit von R_{\max} zu testen, ergibt sich aus einem Vergleich der originalen Bandstruktur mit der aus den TB-Parametern berechneten. Letztere ergibt sich dabei aus der Diagonalisierung einer \mathbf{k} -abhängigen Matrix $E^{\mathbf{k}}$ mit den Elementen

$$E_{nn'}^{\mathbf{k}} \approx N_{\mathbf{R}} \sum_{|\mathbf{R}'| < R_{\max}} e^{-i\mathbf{k}\cdot\mathbf{R}} \langle n\mathbf{R}|H|n'\mathbf{0}\rangle, \quad (8.26)$$

wobei $N_{\mathbf{R}}$ die Anzahl der Einheitszellen im periodischen Grundgebiet ist [122].⁶ Auf Grund der Gitterperiodizität des Hamilton-Operators und der Translationseigenschaft der Wannier-Funktionen (8.7) gehen in der rechten Seite nur Wechselwirkungen zu einer beliebigen Referenzzelle (hier $\mathbf{R}' = \mathbf{0}$) ein. Die Matrixelemente $\langle n\mathbf{R}|H|n'\mathbf{0}\rangle$ lassen sich dabei als Koeffizienten einer Fourierentwicklung der Bandstruktur auffassen: Die grobe Gestalt der Bänder wird bereits durch kleine $|\mathbf{R}|$ festgelegt, zur Beschreibung von Details — wie etwa den Bereich um den Γ -Punkt — sind aber auch Wechselwirkungen mit weiter entfernten Wannier-Funktionen zu berücksichtigen. Es sei bemerkt, dass die Gleichung exakt wird, wenn bei der Summation alle Gittervektoren aus dem periodischen Grundgebiet zugelassen werden. Natürlich ist es jedoch das Ziel bei der Konstruktion einer TB-Basis, die $U^{\mathbf{k}}$ in (8.25) so zu wählen, dass die Approximation auch bei möglichst kleinem R_{\max} noch hinreichend genau ist. Gleichung (8.26) ist ferner auch insofern interessant, als sie ein schnelles

⁶Eine schrittweise Herleitung der Formel findet sich in Anhang H.4.

Sampling der BZ ermöglicht, ohne dabei an jedem Punkt \mathbf{k} den vollen Hamilton-Operator diagonalisieren zu müssen.

Ähnlich, wie in diesem Abschnitt für den Hamilton-Operator vorgestellt, ließe sich nun auch der nicht-lokale Austauschoperator V_{Fock} in der Basis der Wannier-Orbitale darstellen. Wiederum wäre zu erwarten, dass Elemente mit genügend großem $|\mathbf{R} - \mathbf{R}'|$ vernachlässigt werden könnten. Allerdings ist zu erwarten, dass die Darstellung des Austauschoperators in der Wannier-Basis nicht so einfach gelingt wie die des Hamilton-Operators (8.25), bei der sich ausnutzen ließ, dass die Wannier-Funktionen unitäre Transformationen seiner Eigenfunktionen sind (vgl. Anhang H.3). Im Rahmen dieser Arbeit wurde die *ab-initio*-Parametrisierung daher nur für den Hamilton-Operator getestet.

Schema zur Konstruktion des Parametersatzes

Mit Hilfe der beiden vorangehenden Formeln lässt sich nun ein Schema zur Generierung von *ab-initio*-TB-Parametersätzen angeben [122]. Es kann in fünf Schritten zusammengefasst werden:

- (i) Berechnung der Blochzustände und Ein-Teilchen-Energien mit Hilfe eines geeigneten Näherungsverfahrens (DFT-LDA, DFT-EXX, etc.)
- (ii) Wahl einer geeigneten Startkonfiguration für die Wannier-Funktionen (vgl. Abschnitte 8.1.4 und 8.3)
- (iii) Konstruktion einer lokalisierten Wannier-Basis gemäß dem Variationsschema von Marzari und Vanderbilt [116] (vgl. Abschnitt 8.1.3)
- (iv) Berechnung der Hamilton-Matrix in der Wannier-Darstellung nach Gleichung (8.25)
- (v) Mit Hilfe einer Rekonstruktion der Bandstruktur gemäß (8.26) überprüfen, wieviele Wechselwirkungen in dem Parametersatz berücksichtigt werden müssen. Hier muss ein sinnvoller Kompromiss zwischen der Effizienz und der Genauigkeit eingegangen werden. Ist dies nicht möglich, muss ggf. in (ii) nach einer besseren Anfangskonfiguration gesucht werden, vgl. den folgenden Abschnitt 8.3.

8.3 Lokalisierung *kein* Maß für die Güte der Basis

In Abschnitt 8.1.4 wurde beschrieben, wie sich mit Hilfe geschickt gewählter Ansatzfunktionen g_n geeignete Startpunkte $U_0^{\mathbf{k}}$ für den Maximierungsprozess der Lokalisierung (vgl. Abschnitt 8.1.3) konstruieren lassen. Wie aber gelangt man nun zu solchen g_n ? Hier wird man sich im Allgemeinen zunächst von der chemischen Vorstellung leiten lassen und Ansätze für die Wannier-Funktionen dort machen, wo man sie erwartet, beispielsweise auf den atomaren Bindungen wie in Abbildung 8.1 (a) dargestellt. Im Rahmen dieser Arbeit durchgeführte Tests [122] zeigten, dass die Qualität einer aus Wannier-Funktionen konstruierten TB-Basis entscheidend von der Initialisierung abhängen kann. Zugleich offenbarten sie aber auch, dass die chemisch intuitive Wahl *nicht* in jedem Fall die optimalen TB-Parametersätze liefert — es erweist sich vielmehr als notwendig, die Resultate verschiedener Startkonfigurationen hinsichtlich ihrer Eignung als TB-Basis sorgfältig miteinander zu vergleichen. Im Folgenden sei daher exemplarisch die Situation beschrieben, die sich bei der Untersuchung des GaAs-Volumensystems ergab:

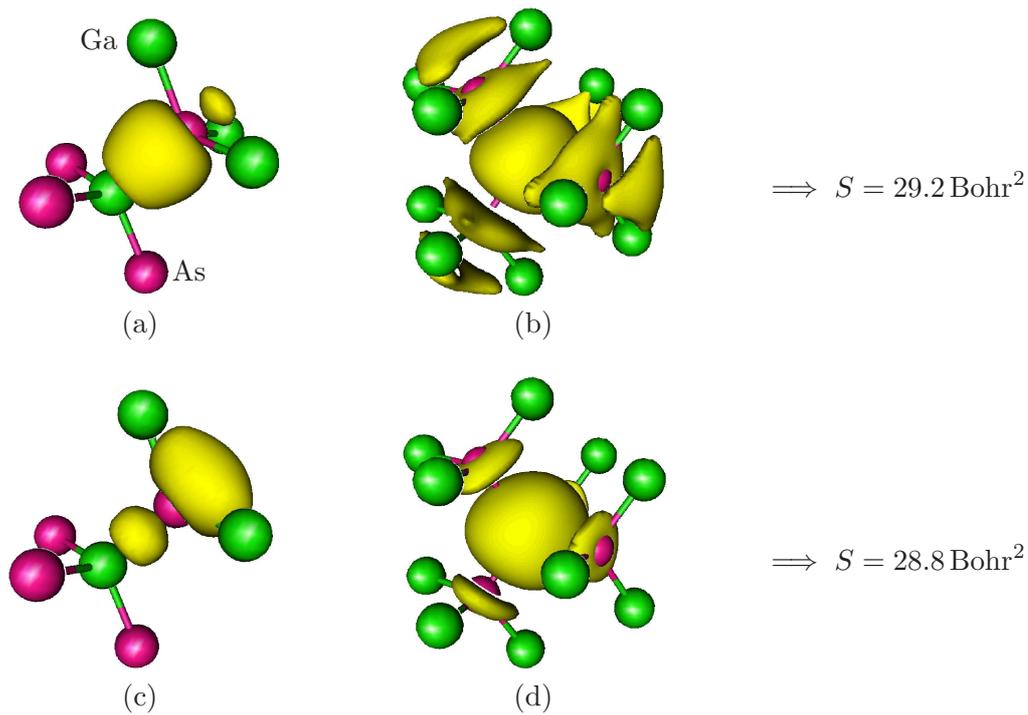


Abbildung 8.2: Zwei verschiedene Wannier-Basen in GaAs. *Oben:* Für diesen Satz erfolgte die Initialisierung der sp^3 -artigen Wannier-Funktionen (a) auf der Bindung (vgl. Abb. 8.1 (a)), während die s^* -artige Funktion (b) direkt auf dem Ga-Atom initialisiert wurde. *Unten:* (c) Bei diesem Satz haben die Wannier-Funktionen, die gemäß Abb. 8.1 (b) initialisiert wurden, nur noch bedingt einen sp^3 -artigen Charakter, da sich ihr Ladungszentrum nun auf der der Bindung gegenüberliegenden Seite befindet. (d) Die s^* -artige Funktion wurde wiederum auf dem Ga-Atom initialisiert. Beide Basissätze weisen in etwa die gleiche Lokalisierung auf. Von den sp^3 -artigen Funktionen (links) ist jeweils nur eine der vier tetraedrisch angeordneten Funktionen dargestellt. In Abb. 8.3 wird ihre Eignung als *tight-binding*-Basis untersucht.

Gemäß dem in Abschnitt 8.2 vorgestellten Schema zur Konstruktion einer *ab-initio*-TB-Basis müssen zuerst die Bloch-Funktionen $|\psi_{m\mathbf{k}}\rangle$ erzeugt werden. Dazu wurde in einer selbstkonsistenten DFT-Rechnung zunächst die Ladungsdichte berechnet, wobei die Austausch-Korrelationswechselwirkung über das LDA-Funktional berücksichtigt wurde. Die \mathbf{k} -Punkt-Integration erfolgte mit Hilfe eines symmetriereduzierten $(4 \times 4 \times 4)$ -Monkhorst-Pack-Gitters, die Abschneideenergie der Ebene-Wellen-Basis betrug 15 Ry, und es wurde die experimentelle Gitterkonstante von 10.68 Bohr benutzt. Ferner wurden mit dem Programm fhi98PP [115] nach dem Hamann-Verfahren [110] generierte Pseudopotentiale verwendet, wobei die Elektronen der 4. Schale des Ga-Atoms als Valenz-, alle anderen als Rumpfelektronen behandelt wurden. Um später mit der TB-Basis die Bandlücke beschreiben zu können, wurden anschließend in einer Bandstrukturrechnung Blochfunktionen und Ein-Teilchen-Energien der vier Valenzbänder und des untersten Leitungsbandes bestimmt. Dies geschah für die Punkte eines nicht symmetriereduzierten Monkhorst-Pack-Gitters, das den Γ -Punkt miteinschließt.

Nun wurde zunächst die Initialisierung gemäß dem chemisch intuitiven Bild getestet, das folgende Konfiguration nahe legt: Die Beschreibung der besetzten Zustände, also der Valenzbänder, sollte durch vier Wannier-Funktionen pro Einheitszelle gelingen, die eine Gestalt ähnlich der von sp^3 -Hybriden aufweisen und am (elektronegativeren) Anion lokalisiert sind. Da die Zentren der Ladungswolken dabei jeweils auf einer der tetraedrisch

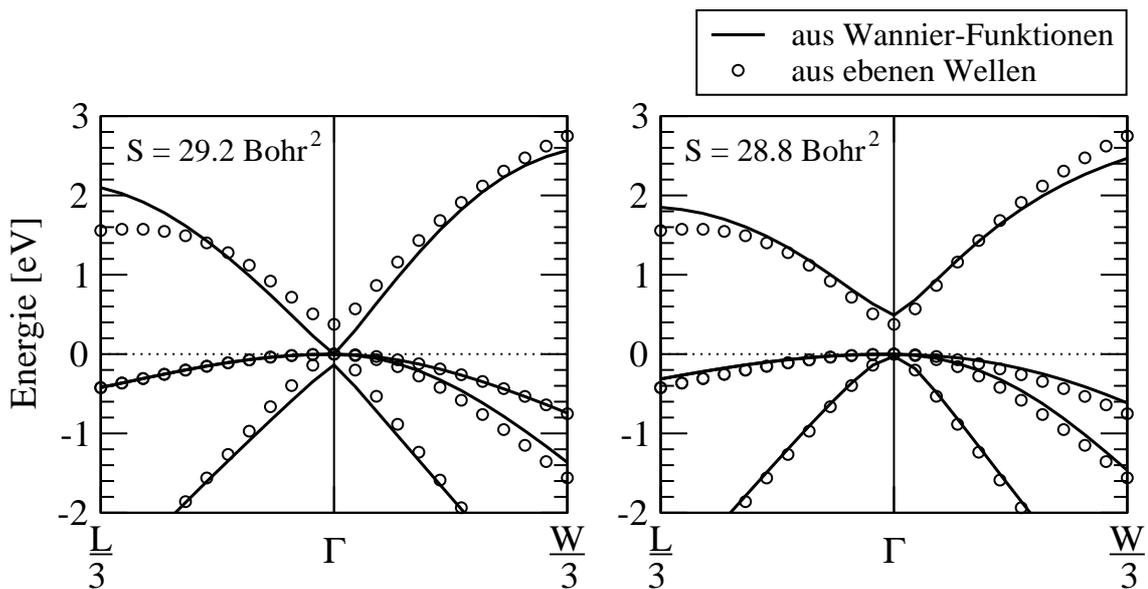


Abbildung 8.3: Bandstruktur von GaAs in der Umgebung des Γ -Punktes. Die durch Kreise dargestellte Bandstruktur (links und rechts identisch) wurde aus einer Ebene-Wellen-Darstellung berechnet (LDA), die durchgezogenen Kurven ergeben sich aus der Wannier-Darstellung. *Links:* Die auf den Bindungen initialisierten Wannier-Funktionen (vgl. Abb. 8.1 (a) und Abb. 8.2 oben) versagen völlig bei der Beschreibung der Bandlücke. *Rechts:* Werden stattdessen Wannier-Funktionen mit vergleichbarer Lokalisierung verwendet, die aber jeweils auf der der Bindung gegenüberliegenden Seite initialisiert wurden (vgl. Abb. 8.1 (b) und Abb. 8.2 unten), kann eine deutlich bessere Reproduktion der originalen Bandstruktur erreicht werden. (Anmerkung: Die Knicke einiger durchgezogener Kurven am Γ -Punkt sind auf den gewählten Abstand der Stützstellen zurückzuführen, sind also rein numerisch bedingt und stellen keine durch die Wannier-Basis ins Spiel kommenden Besonderheiten dar.)

angeordneten Bindungen zu erwarten sind, erfolgte die Positionierung der entsprechenden Ansatzorbitale (Gauß-Funktionen) wie in Abbildung 8.1 (a) demonstriert. Für das Leitungsband steht damit noch eine Funktion pro Einheitszelle zur Verfügung, die — sie soll ja unbesetzte Zustände beschreiben — auf dem Kation mit einer s^* -artigen Gestalt zu erwarten wäre. Die Initialisierung erfolgte demgemäß durch eine Gauß-Funktion direkt auf dem Ga-Atom. Nach Projektion auf die Blochfunktionen betrug die Streuung der so konstruierten Startkonfiguration 107.1 Bohr^2 . Sie wurde in der anschließenden Minimierung des Funktional (8.9) um ca. 70% auf 29.2 Bohr^2 reduziert, wobei die in Abbildung 8.2 (a) und (b) dargestellten lokalisierten Wannier-Funktionen entstanden. Tatsächlich entspricht deren Form und Position in etwa der im Vorangehenden beschriebenen „naiven“ Erwartung, wobei um das kationische s^* -Orbital noch einige an den benachbarten N-Atomen lokalisierte Ausläufer mit sp -Charakter gruppiert sind, die die Orthogonalität der Basisfunktionen untereinander gewährleisten (vgl. Abschnitt 8.1.4).

Eine weitergehende Untersuchung machte allerdings deutlich, dass die Rekonstruktion der Bandstruktur mit Hilfe dieses Basissatzes nur äußerst langsam hinsichtlich der Anzahl der benötigten Matrixelemente gegen die originale, d.h., direkt in der Ebene-Wellen-Basis erzeugten Bandstruktur konvergiert. Selbst wenn R_{\max} in Gleichung (8.26) so groß eingestellt wird, dass noch Wechselwirkungen zwischen Wannier-Funktionen viert-nächster Nachbarzellen erfasst werden, misslingt die Beschreibung des für die optischen Eigenschaften relevanten Bereichs der BZ um den Γ -Punkt völlig: Die Bandlücke verschwindet,

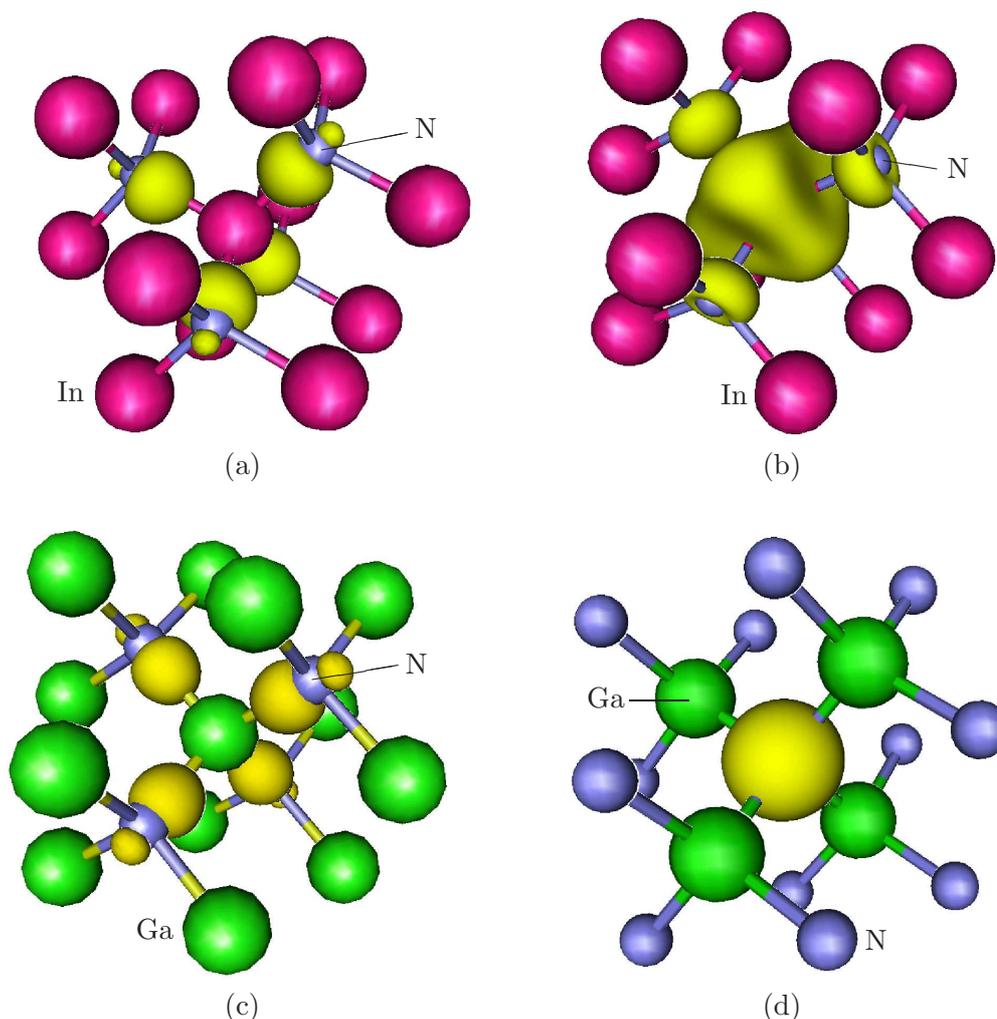


Abbildung 8.4: Oben: Optimierte Wannier-Basis in InN. (a) Vier Wannier-Funktionen in der Form von sp^3 -Hybriden sind an den Anionen lokalisiert. (b) Eine weitere Funktion befindet sich auf dem Kation, weist aber keine klar s^* -artige Gestalt auf. Unten: Optimierte Wannier-Basis in GaN. (c) Vier Wannier-Funktionen mit sp^3 -Gestalt sind wiederum an den Anionen lokalisiert. (d) Die fünfte Basisfunktion besitzt hier klaren s^* -artigen Charakter, befindet sich jedoch im Gegensatz zu InN ebenfalls auf dem Anion.

wohingegen die Entartung bei Γ anstelle der oberen drei nur zwei Valenzbänder umfasst (vgl. linke Bandstruktur in Abb. 8.3). Somit erweist sich diese Basis für die Konstruktion eines TB-Parametersatzes als ungeeignet.

In einer Reihe weiterer Tests wurden nun von der intuitiven Vorstellung abweichende Anfangskonfigurationen untersucht. Es zeigte sich, dass eine effiziente TB-Basis (d.h., eine Basis, in der sich hinreichende Genauigkeit bei Berücksichtigung möglichst weniger Wechselwirkungen erreichen lässt) mit folgender Initialisierung konstruiert werden kann: Während wie zuvor eine Gauß-Funktion für das s^* -Orbital auf dem Kation angesetzt wird, werden die anderen vier nicht mehr auf den Bindungslinien platziert, sondern jeweils auf der der Bindung gegenüberliegenden Seite des Anions, genauer: sie gehen aus den alten Ansätzen für die sp^3 -Hybride durch Spiegelung am As-Kern hervor. Dies ist in Abbildung 8.1 (b) veranschaulicht. Anschließende Projektion dieser Konfiguration auf die Blochfunktionen und Minimierung liefert die in Abbildung 8.2 (c) und (d) gezeigten Wannier-

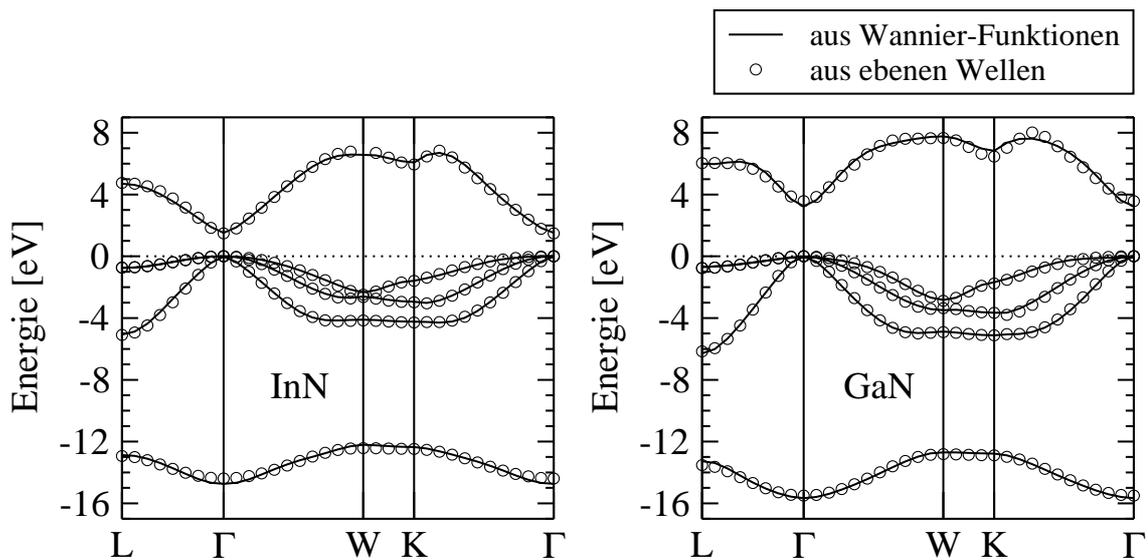


Abbildung 8.5: Bandstrukturen von InN (links) und GaN (rechts). Es ist jeweils das Resultat der originalen DFT-EXX-Berechnung in ebenen Wellen (Kreise) im Vergleich mit der Rekonstruktion aus den TB-Parametern (durchgezogene Linien) dargestellt. In den Parametersätzen werden Wechselwirkungen zwischen Wannier-Funktionen bis zu den viert-nächsten Nachbarzellen berücksichtigt. Für die Bandlücke ergibt sich eine Abweichung von ca. 0.04 eV bei InN und ca. 0.3 eV bei GaN.

Funktionen. Die Streuung wurde dabei um ca. 80% auf 28.9 Bohr^2 reduziert, d.h., der intuitive und der hier konstruierte Ansatz führen zwar zu verschiedenen Minima des Streufunktional (8.9), in beiden Fällen besitzen die entsprechenden Wannier-Basen aber eine vergleichbare Lokalisierung. Die Rekonstruktion der Bandstruktur offenbart nun jedoch den Vorteil der zweiten Wannier-Basis für die Konstruktion eines TB-Parametersatzes: Zumindest in der interessierenden Region um den Γ -Punkt lässt sich das Resultat der Ebene-Wellen-Rechnung deutlich besser reproduzieren (die maximale Abweichung im Bereich des Γ -Punkts beträgt 0.11 eV), wenn in Gleichung (8.26) Wechselwirkungen noch bis zu den viert-nächsten Nachbarzellen miteinbezogen werden (vgl. rechte Bandstruktur in Abb. 8.3).

Dieses Beispiel zeigt, dass die Lokalisierung alleine noch kein ausreichendes Kriterium zur Generierung effizienter TB-Basen darstellt. Tatsächlich wäre auch zu erwarten, dass ein zu beschreibender Operator X selbst die Gestalt einer optimalen Basis beeinflusst. Definiert man „Effizienz“ in diesem Zusammenhang über die Anzahl der (für eine gegebene Genauigkeit) benötigten Matrixelemente, ließe sich bei der Variation der unitären Transformationsmatrizen $U^{\mathbf{k}}$ in (8.6) beispielsweise fordern, dass $\langle n\mathbf{R}|X|n'\mathbf{R}'\rangle$ als Funktion der Gittervektoren \mathbf{R} und \mathbf{R}' ein möglichst starkes Abklingverhalten bzgl. des Abstandes $|\mathbf{R} - \mathbf{R}'|$ der Einheitszellen aufweist. Damit wäre etwa (8.26) automatisch optimiert, eine Nachjustierung „per Hand“ wie oben geschildert wäre nicht mehr notwendig. Diese Idee ist hier als Ausblick zu verstehen, eine Weiterentwicklung des Ansatzes geht über die Zielstellung dieser Arbeit hinaus.

8.4 Optimierte Wannier-Basen für GaN und InN

Der in Abschnitt 8.2 vorgestellte Formalismus wurde nun auf GaN und InN in Zinkblendestruktur angewandt [122]. Dafür wurden in einem ersten Schritt wieder die Blochfunktionen erzeugt. Die Vorgehensweise war dabei völlig analog wie im vorangehenden Abschnitt für GaAs beschrieben, allerdings wurden die DFT-Berechnungen nun mit Hilfe des EXX-Formalismus durchgeführt. Wiederum wurde für die \mathbf{k} -Punkt-Integration ein $(4 \times 4 \times 4)$ -Monkhorst-Pack-Gitter verwendet, die Abschneideenergie der Ebene-Wellen-Berechnung betrug 50 Ry für InN und 60 Ry im Falle von GaN. Es wurden relativistische OEP-Pseudopotentiale benutzt [111, 112], die d -Elektronen der Kationen wurden als Rumpfelektronen betrachtet. Damit ergeben sich Bandlücken von 1.48 eV bei InN und 3.57 eV bei GaN. Die entsprechenden Bandstrukturen sind in Abbildung 8.5 gezeigt.

Für die Optimierung der TB-Basen wurden nun wieder verschiedene Startkonfigurationen für die Wannier-Orbitale untersucht. Anders als bei GaAs erwies sich für die betrachteten Gruppe-III-Nitride eine Initialisierung von jeweils vier Orbitalen auf den Bindungen (analog zu Abb. 8.1 (a)) als vorteilhafter als die Initialisierung auf den gegenüberliegenden Seiten. Nach der Minimierung wurden so in beiden Materialien zu den Anionen (N) hin orientierte Bindungsorbitale mit deutlichen sp^3 -Charakter ausgeprägt, vgl. Abb. 8.4. In Übereinstimmung mit der chemischen Vorstellung befindet sich in InN die Position des fünften Orbitals auf dem Kation (In). Seine Gestalt weicht dabei allerdings stark von der für ein s^* -Orbital typischen Kugelform ab (Abb. 8.4 (b)). Deutlichen s^* -Charakter trägt hingegen das fünfte Orbital der für GaN optimierten Wannier-Basis. Seine Position befindet sich hier jedoch entgegen der Erwartung auf dem Anion (Abb. 8.4 (d)).

Abbildung 8.5 zeigt eine sehr gute Übereinstimmung der originalen EXX-Bandstruktur, die mit Hilfe einer Ebene-Wellen-Rechnung erzeugt wurde, und ihrer Rekonstruktion aus den *ab-initio*-TB-Parametern. Insbesondere zeigt sich, dass auch die Bandlücke mit Hilfe des erzeugten Parametersatzes korrekt reproduziert werden konnte. Dies ist plausibel: Denn während die Konstruktion der Wannier-Orbitale gemäß (8.6) nur von den Wellenfunktionen abhängt, die in EXX und LDA/GGA nur unwesentlich voneinander verschieden sind, fließen in die Berechnung der *ab-initio*-Parameter gemäß (8.25) auch die Einteilchen-Energien ein. Der Zugang kann damit sensibel auf die durch EXX erreichte deutliche Verbesserung der Bandstruktur reagieren, was mit den in Abb. 8.5 gezeigten Ergebnissen gut bestätigt wird.

Es sei an dieser Stelle noch einmal explizit darauf hingewiesen, dass die Verwendung von im Vergleich zu LDA/GGA sehr komplizierten und numerisch aufwendigen Funktionalen wie hier dem EXX-Funktional für die Konstruktion der *ab-initio*-TB-Parameter keinerlei zusätzlichen Aufwand darstellt. Sind einmal die Wellenfunktionen und die Einteilchen-Energien bekannt, geschieht die Konstruktion der Parameter stets auf die gleiche Art und Weise (vgl. Abschnitt 8.2), völlig unabhängig von den Anforderungen, die zur Berechnung eben dieser Wellenfunktionen und Einteilchen-Energien vonnöten sind. Die verwendeten Eingangsdaten müssen noch nicht einmal aus einer DFT-Rechnung herrühren, so können beispielsweise Ergebnisse einer Hartree-Fock-Rechnung ebenso als Grundlage zur Konstruktion Verallgemeinerter Wannier-Orbitale und entsprechender TB-Parametersätze dienen, einzige Bedingung ist dabei, dass sich die Wellenfunktionen der vorausgehenden Rechnung als Blochwellen (vgl. Abschnitt 3.1) darstellen lassen.

Ein interessanter Ausgangspunkt für weiterführende Untersuchungen wäre zudem, die Methode auch auf Ergebnisse von EXX+GW-Berechnungen aufzusetzen, eine der zurzeit genauesten Methoden zur Berechnung von Bandlücken [170, 171, 172, 173, 174, 175]. Hier

ist allerdings zu beachten, dass GW nur Korrekturen zu den Einteilchen-Energien liefert, nicht aber zu den Wellenfunktionen. Bei der Konstruktion der ab-initio-Parameter gemäß (8.25) würden die GW -Korrekturen folglich nur die Einteilchen-Energien beeinflussen, die Wellenfunktionen wären weiterhin einer EXX-Rechnung zu entnehmen.

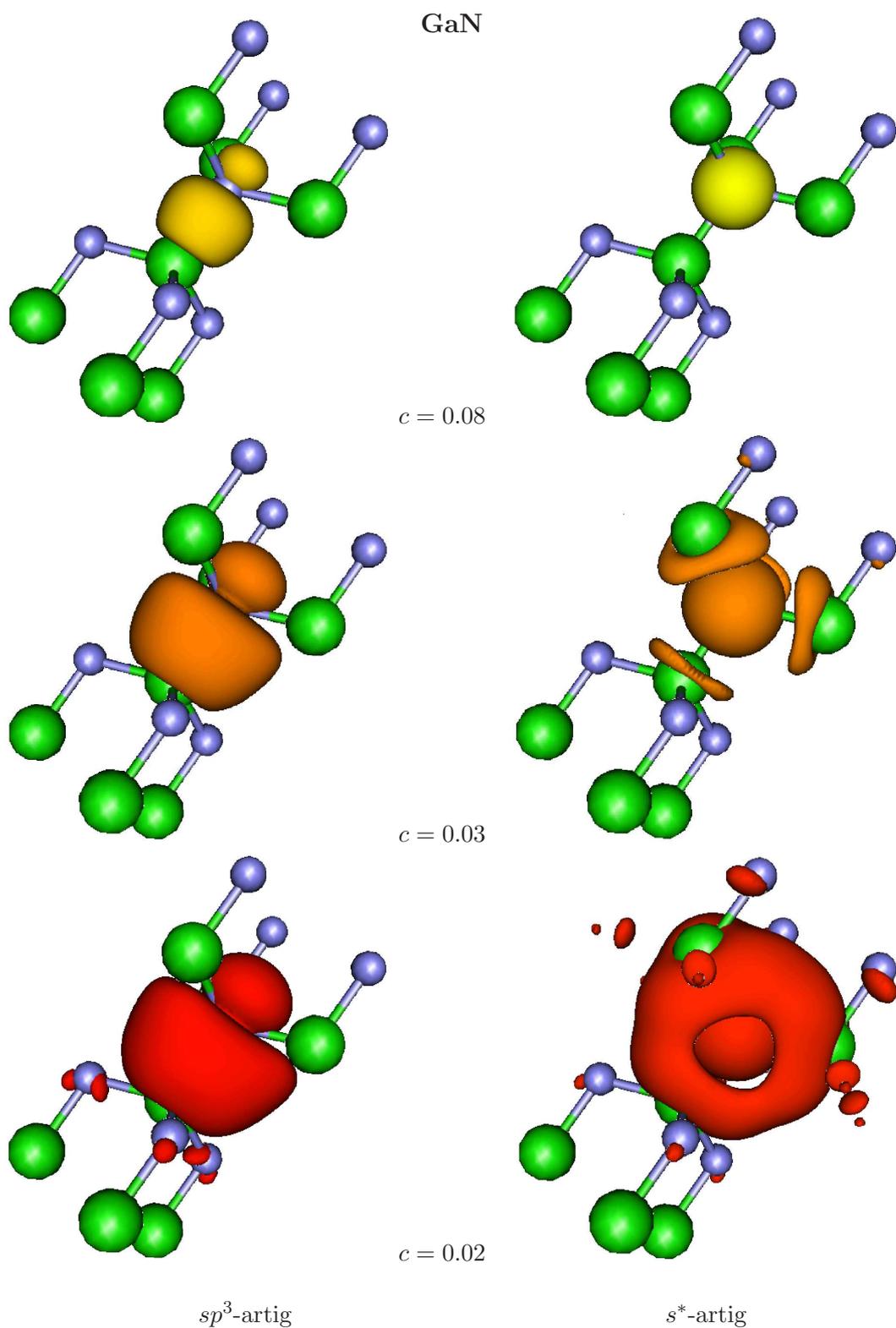


Abbildung 8.6: Verallgemeinerte Wannier-Orbitale in GaN (Ga: grün, N: blau) für verschiedene Isoflächenparameter c . Linke Spalte: sp^3 -artiges, rechte Spalte: s^* -artiges Wannier-Orbital. Bei kleinen c erkennt man, dass die Funktionen neben dem Hauptlokalisierungsgebiet noch zahlreiche kleinere Ausläufer besitzen.

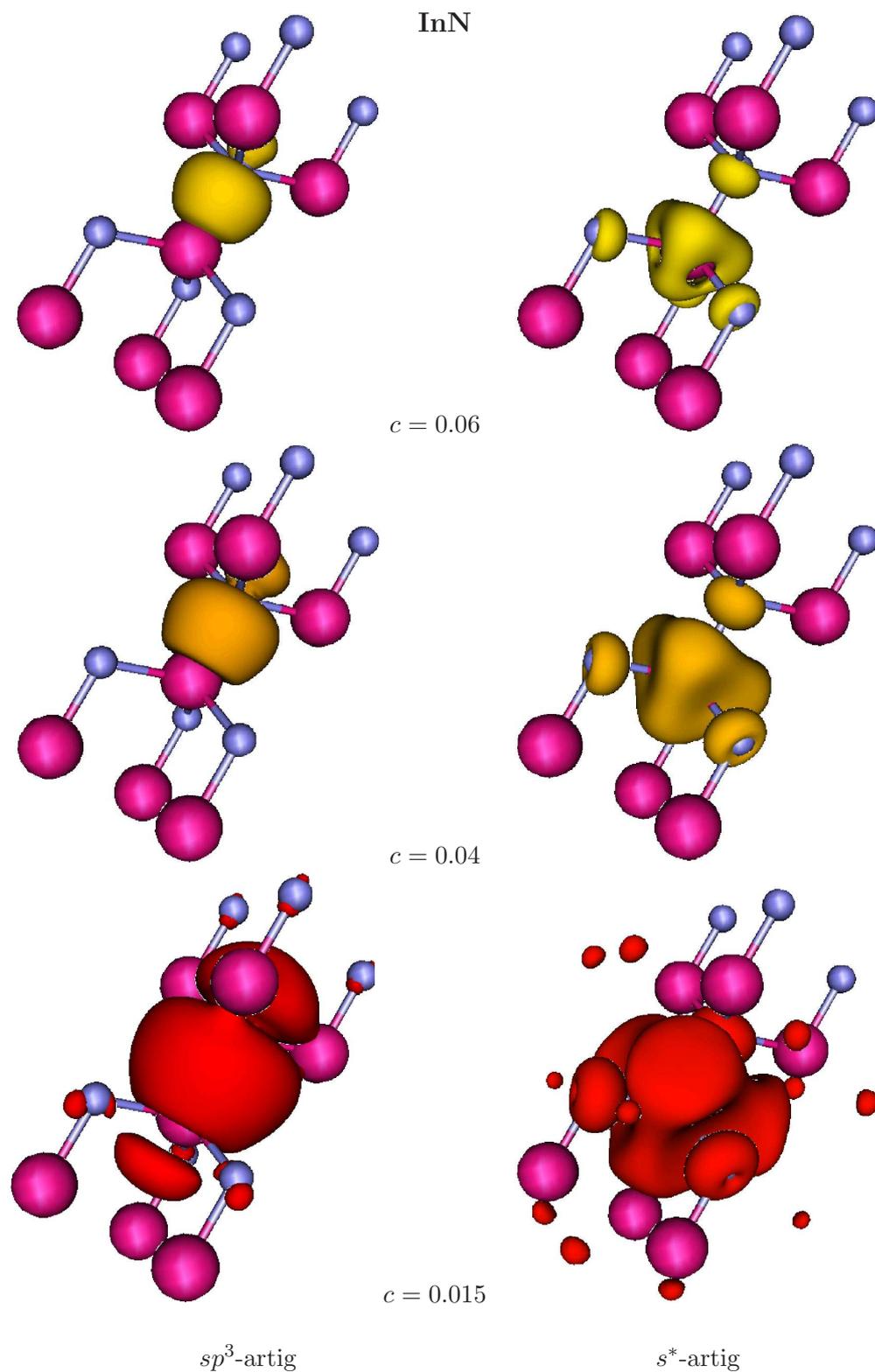


Abbildung 8.7: Verallgemeinerte Wannier-Orbitale in InN (In: rot, N: blau) für verschiedene Isoflächenparameter c . Linke Spalte: sp^3 -artiges, rechte Spalte: s^* -artiges Wannier-Orbital. Bei kleinen c erkennt man, dass die Funktionen neben dem Hauptlokalisierungsgebiet noch zahlreiche kleinere Ausläufer besitzen.

Zusammenfassung und Ausblick

Es sei zuerst eine kurze Zusammenfassung der beiden Teile der vorliegenden Arbeit — Variationelle EXX-Methode (Teil I), Verallgemeinerte Wannier-Funktionen als lokalisierte *ab-initio*-Basis (Teil II) — gegeben. Anschließend wird eine kurze Einordnung dieser Arbeit in die aktuelle wissenschaftliche Entwicklung auf dem Gebiet der theoretischen Festkörperphysik versucht.

Zu Teil I (Variationelles EXX-Verfahren)

Schwerpunkt der vorliegenden Arbeit war die Entwicklung und Implementierung des variationellen EXX-Verfahrens (varEXX) [78, 79, 80] in einen Pseudopotential-Ebene-Wellen-Code (S/PHI/nX [81, 82]), um diese Methode, die bislang nur für Atome und Cluster getestet worden ist, so erstmalig auch für periodische Systeme verfügbar zu machen.

Die Programmierung des varEXX-Verfahrens umfasste im Wesentlichen vier Schritte:

- (i) Die Implementierung eines geeigneten Minimierungsverfahrens zur Lösung der Sternheimer-Gleichung (Abschnitt 4.1). Hierfür wurden verschiedene Algorithmen (*Steepest-Descent*, *Conjugate-Gradient*, *Preconditioned Conjugate-Gradient*) getestet.
- (ii) Die Erzeugung der induzierten Ladungsdichte (Abschnitt 4.2).
- (iii) Ein geeignetes Verfahren für das *Mixing* von Potential und induzierter Ladungsdichte (Abschnitt 4.3). Die prinzipielle Vorgehensweise wurde dabei zunächst mit dem einfachen *Mixing* untersucht. Anschließend wurde eine optimierte Version des Pulay-Mixers an die Erfordernisse im varEXX-Verfahren angepasst. Interessanterweise zeigte sich dabei, dass eine Präkonditionierung des Mixers nach dem Kerker-Schema, wie es etwa beim selbstkonsistenten Lösen der Kohn-Sham-Gleichung für das Mischen der Ladungsdichte verwendet wird, für das Mixen von Potential und induzierter Ladungsdichte ungeeignet ist. Stattdessen findet man, dass durch die Besonderheit des Mischens dieser unterschiedlichen Größen der Kerker-Ansatz bereits implizit erfüllt ist.
- (iv) Die Berechnung des nicht-lokalen Austauschoperators (Abschnitt 4.4). Die numerische Auswertung dieser Größe erweist sich aus zweierlei Gründen als sehr aufwendig: zum einen muss die Singularität des Nenners in geeigneter Weise gehandhabt werden, und zum anderen gestaltet sich die Symmetrisierung außerordentlich schwierig. Dieser Schritt ist daher der problematischste des Verfahrens (siehe unten).

Das Verfahren wurde anschließend auf eine Reihe von Halbleitern angewandt und seine Leistung mit der einer bereits im Code vorhandenen Implementierung des Standard-EXX-Verfahrens verglichen. Mit dem variationellen Ansatz gelang es, die benötigte CPU-Zeit

dabei ungefähr zu halbieren und den Speicherbedarf sogar auf rund 1/3 zu reduzieren. Es zeigte sich zudem, dass varEXX umso effektiver gegenüber der Standard-Methode ist, je höher die Abschneideenergie der Ebene-Wellen-Basis angesetzt wird. Rechnet man ferner den Anteil der vom nicht-lokalen Austauschoperator¹ benötigten CPU-Zeit heraus, erlaubt der variationelle Zugang sogar eine Reduktion der Rechenzeit um rund einen Faktor 10. Zudem erwies sich der Algorithmus als sehr stabil: In der Regel wurde ein LDA-Austauschpotential als Start für den selbstkonsistenten DFT-EXX-Zyklus verwendet, jedoch zeigten Tests an GaAs, GaN und InN, dass das Verfahren auch mit der Nullfunktion (also keinem Ansatz, der sich bereits in der Nähe des exakten EXX-Potentials befindet) als Startpunkt konvergiert.

Der Engpass des Verfahrens (vor allem hinsichtlich der CPU-Zeit) ist der nicht-lokale Austauschoperator; zukünftige Versuche einer weiteren Effizienzsteigerung des EXX-Verfahrens müssen an dieser Stelle ansetzen. Wie in Abschnitt 4.4.2 bereits angedeutet, erscheint dazu die Auswertung des nicht-lokalen Austauschoperators in der Ortsdarstellung eine geeignete Maßnahme. Interessant wäre es auch, zu untersuchen, inwieweit sich durch die Verwendung einer Basis Verallgemeinerter Wannier-Funktionen (siehe unten) Vorteile in Bezug auf die Handhabung dieses Operators erzielen lassen.

Der erste Teil der Arbeit schließt mit einer Untersuchung der Linearität des exakten Austauschpotentials als Funktional der Ladungsdichte. Speziell wurde am Beispiel von GaN und InN untersucht, ob die bei der Verwendung von Pseudopotentialen bei diesen Materialien übliche Aufspaltung in Rumpf- und Valenzelektronen im Bereich der so genannten *semi-core*-Schale (*s*- und *p*-Schale in den Rumpf, *d*-Schale in die Valenz) hinsichtlich des exakten Austauschpotentials zulässig ist. Die Ergebnisse entsprechender DFT-EXX-Rechnungen wurden dafür mit denen von Berechnungen verglichen, die die gesamte *semi-core*-Schale des Kations in der Valenz berücksichtigen. Diese für eine Abschätzung der Leistungsfähigkeit der EXX-Methode insgesamt sehr wichtigen Untersuchungen wurden erstmalig durch den im Rahmen dieser Arbeit implementierten varEXX-Formalismus ermöglicht. Durch die errechneten Daten wird die Annahme gestützt, dass es sich beim EXX-Potential um ein zumindest annähernd lineares Funktional der Dichte handelt. Zwar ergeben sich kleine Differenzen in der Bandlücke (ca. 0.2 eV bei InN), jedoch muss bei Verwendung von Pseudopotentialen ohnehin von einer Fehlertoleranz in dieser Größenordnung ausgegangen werden. Insbesondere wurde sowohl mit als auch ohne Berücksichtigung der *d*-Elektronen in der Valenz eine weitgehende Übereinstimmung der Ladungsdichteverteilung gefunden (abgesehen von Bereichen in unmittelbarer Nähe der Kationen-Kerne, in denen man wegen der durch die Pseudopotentiale verursachten Verzerrungen auch keine Vergleichbarkeit erwarten kann).

Die gefundenen Ergebnisse zeigen, dass sich EXX-Berechnungen damit anders verhalten als die nicht-linear konzipierte *GW*-Methode: Entsprechende *GW*-Resultate auf der Basis von DFT-LDA-Berechnungen (LDA+*GW*) für GaN zeigen, dass sich bei dieser Methode die Bandlücke dieses Halbleiters bei Hinzunahme der *d*-Zustände in die Valenz um 0.71 eV verringert (vgl. Rohlffing *et al.* [152]).

¹Der nicht-lokale Austauschoperator wird sowohl im Standard- als auch im variatonellen EXX-Verfahren benötigt.

Zu Teil II (Verallg. Wannier-Funktionen als lokalisierte *ab-initio*-Basis)

Hier wurde eine im Rahmen dieser Arbeit entwickelte Konstruktionsmethode für *ab-initio*-Basissätze im *Tight-Binding*-Modell [122] vorgestellt. Die Methode benutzt die Ergebnisse (Wellenfunktionen und Einteilchen-Energien) einer vorausgehenden Berechnung des entsprechenden (periodischen) Systems als Eingangsdaten. Daraus werden mit Hilfe einer systematischen Konstruktionsvorschrift zunächst Verallgemeinerte Wannier-Orbitale und anschließend TB-Parameter erzeugt. Hierbei werden keinerlei empirischen Anpassungen o.Ä. vorgenommen, d.h., die erzeugten TB-Parameter sind reine *ab-initio*-Größen (unter der Voraussetzung, dass die vorausgehende Berechnung der Wellenfunktionen und Einteilchen-Energien ebenfalls *ab-initio* erfolgte). Als großer Vorteil der Methode kann weiterhin gewertet werden, dass die Konstruktion der *ab-initio*-Parameter völlig unabhängig von der speziellen Wahl des Verfahrens zur Erzeugung der Eingangsdaten ist: So bedeutet etwa die Konstruktion eines *ab-initio*-Parametersatzes auf Grundlage einer DFT-EXX-Rechnung keinen Mehraufwand gegenüber einer entsprechenden Konstruktion aus DFT-LDA-Daten. Neuartige Verfahren, die über Standard-DFT-Funktionale hinausgehen, lassen sich also sehr einfach in diesen Formalismus integrieren. Schließlich ist zu betonen, dass durch die *ab-initio*-Konstruktion die Genauigkeit der jeweils zu Grunde liegenden Daten in den Parametersätzen „konserviert“ wird. Mit Hilfe von TB-Parametern, die etwa aus einer EXX-Rechnung gewonnen wurden, lassen sich somit wieder die EXX-Bandlücken reproduzieren. Die Bedeutung der in dieser Arbeit entwickelten TB-Methode liegt nun darin, dass sie eine Möglichkeit eröffnet, die Genauigkeit von Berechnungen über EXX- und andere komplizierte Funktionale, deren direkte Anwendung mit der derzeitigen Computerleistung nur auf Systeme mit wenigen Atomen beschränkt ist, auf mesoskopische Skalen mit ca. 10 000 Atomen (z.B. Quantenpunkte) übertragen zu können.

Das Verfahren wurde im Rahmen der Arbeit an GaAs, GaN und InN in der Zinkblende-Phase getestet. Mit einer DFT-Rechnung wurden dazu in einem ersten Schritt Ein-Teilchen-Zustände und -Energien erzeugt. Durch die Verwendung von EXX-Funktionalen erhält man dabei einen TB-Parametersatz, der für die Berechnung optischer Matrixelemente geeignet ist (siehe oben). Anschließend wurde die Basis der Eigenfunktionen (oder ein geeigneter Unterraum) durch eine unitäre Transformation in eine Basis Verallgemeinerter Wannier-Funktionen überführt. Bei dieser Transformation entsteht kein Verlust an Genauigkeit. Die Lokalisierung der Wannier-Funktionen lässt sich über ihre Varianz (Streuung) im Ortsraum messen; eine Minimierung dieses Funktionals führt auf Wannier-Funktionen mit maximaler Lokalisierung, die daher als *Tight-Binding*-Orbitale einsetzbar sind. Eine Darstellung des Hamilton-Operators in dieser Basis liefert *Tight-Binding*-Parameter, deren Genauigkeit sich leicht durch die Anzahl der berücksichtigten Nächste-Nachbar-Wechselwirkungen kontrollieren lässt.

Erste Tests an GaAs offenbarten jedoch, dass die Minimierung der Varianz kein ausreichendes Kriterium für die Erzeugung effizienter *Tight-Binding*-Parametersätze darstellt. Verschiedene Sätze Verallgemeinerter Wannier-Funktionen mit vergleichbarer minimaler Streuung, jedoch sehr unterschiedlicher Konvergenz bzgl. der Nächsten-Nachbar-Wechselwirkung wurden gefunden, indem die Minimierung von verschiedenen Anfangskonfigurationen aus gestartet wurde. Dies zeigt die Notwendigkeit sorgfältiger Tests der Anfangskonfiguration bei der Konstruktion der Orbitale.

Die optimalen Wannier-Orbitale, die dabei für InN gefunden wurden, stehen im Einklang mit der chemischen Intuition: es ergeben sich vier sp^3 -artige Orbitale, die jeweils am Anion lokalisiert sind, und ein s^* -artiges Orbital auf dem Kation. Für GaN ergeben sich zwar ebenfalls vier auf den Anionen platzierte sp^3 -artige Orbitale, die s^* -artigen Or-

bitale sind hier aber, im Gegensatz zur chemischen Intuition, ebenfalls auf den Anionen lokalisiert.

Eine Weiterentwicklung der Methode könnte darin bestehen, nicht die Lokalisierung der Wannier-Funktionen zu minimieren, sondern etwa bei der unitären Transformation auf die Wannier-Basis zu fordern, dass die Matrixelemente des betrachteten Operators (z.B. des Hamilton-Operators) in dieser Darstellung ein möglichst starkes Abklingverhalten bzgl. des Abstandes $|\mathbf{R} - \mathbf{R}'|$ der Gittervektoren aufweisen. Diese Strategie würde automatisch auf einen optimierten *Tight-Binding*-Parametersatz führen.

Einordnung der Arbeit in die aktuelle Forschung

Die Situation zu Beginn der vorliegenden Untersuchungen lässt sich etwa wie folgt umreißen: In der Kohn-Sham-Dichtefunktionaltheorie wurden eine Vielzahl von Austausch-Korrelationsfunktionalen entwickelt (LDA, GGAs), die zwar effiziente Berechnungen erlauben, aber für Fragestellungen, die eine genaue Beschreibung der Bandlücke von Isolatoren und Halbleitern benötigen (bspw. optische Eigenschaften des Materials), zu ungenau sind. Demgegenüber steht das EXX-Funktional, das insbesondere in Kombination mit einer anschließenden *GW*-Korrektur äußerst genaue Vorhersagen der Bandlücke ermöglicht, das andererseits aber so hohe numerische Anforderungen stellt, dass die Anwendung auf einfachste Systeme beschränkt bleibt (siehe Einleitung), und damit ein Einsatz bei komplexeren Untersuchungen nicht möglich ist.

Hier setzt nun die vorliegende Arbeit an: Es konnte gezeigt werden, dass sich durch die Verwendung einer alternativen, variationellen Formulierung des EXX-Verfahrens der benötigte numerische Aufwand auch für periodische Systeme drastisch reduzieren lässt. Durch den variationellen Ansatz gelingt es, zwei der drei Engpässe der originalen Formulierung (Greensche Funktion, Inversion der linearen Antwort, nicht-lokaler Austauschoperator) durch numerisch einfache Minimierungsverfahren zu ersetzen, was ca. eine Verdopplung der Rechengeschwindigkeit bei weniger als einem Drittel des benötigten Speicherplatzes ermöglicht. Ohne die für den Austauschoperator — des nun einzig verbleibenden Engpasses — benötigte Zeit gemessen, lässt sich sogar eine bis zu ca. 10 mal schnellere Rechenzeit erzielen (siehe oben).

Die Möglichkeiten für den praktischen Einsatz von EXX-Potentialen wurden damit entscheidend erweitert, so dass nun neue und deutlich komplexere Fragestellungen angegangen werden können. Im Rahmen dieser Arbeit wurde dies am Beispiel von EXX-Pseudopotential-Berechnungen für GaN und InN demonstriert, die sich erstmalig unter Mitnahme der kationischen *s*- und *p-semi-core*-Schalen in der Valenz durchführen ließen. Vor allem aber sei an dieser Stelle hervorgehoben, dass sich durch die variationelle EXX-Methode zugleich auch der Anwendungsbereich von EXX+*GW* vergrößert — eines der zurzeit präzisesten Verfahren zur Bestimmung von Bandstrukturen und Gegenstand aktiver Forschung [170, 171, 172, 173, 174, 175]. Nicht zuletzt zeigt auch der im zweiten Teil der Arbeit entwickelte *ab-initio*-TB-Ansatz eine Richtung auf, wie sich die Genauigkeit von EXX- und anderen über herkömmliche DFT-Funktionale hinausgehenden Potentialen effizient sogar auf Systeme mesoskopischer Größenordnung übertragen lässt.

Anhänge

Anhang A

Fourier-Entwicklungen einiger wichtiger Größen

Um Unübersichtlichkeiten durch zu viele Klammern zu vermeiden, sei das Konjugiert-Komplexe einer Variablen in den folgenden Abschnitten durch einen Querstrich (z.B. \bar{z}) bezeichnet und nicht wie sonst in dieser Arbeit üblich mit einem Sternchen (z^*).

A.1 Das Coulomb-Potential

An einigen Stellen im Text wird die Fouriertransformierte des Coulomb-Potentials benutzt. Ihre Herleitung wird in vielen Lehrbüchern, z.B. in Ref. [92](#), explizit vorgeführt, daher sei hier nur das Ergebnis zusammengefasst:

$$v_{\text{Coul}}(\mathbf{r}) := \frac{1}{r} \rightsquigarrow \hat{v}_{\text{Coul}}(\mathbf{G}) := \frac{1}{\sqrt{\Omega}} \int \frac{1}{r} e^{-i\mathbf{G}\mathbf{r}} d^3r \approx \frac{1}{\sqrt{\Omega}} \frac{4\pi}{\mathbf{G}^2}, \quad (\text{A.1})$$

wobei Ω das Volumen der durch die Vektoren $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ aufgespannten Einheitszelle ist, vgl. Abschnitt [3.1](#). Es sind zwei Schwierigkeiten zu beachten: (i) Das Coulomb-Potential v_{Coul} ist keine periodische Funktion; die Voraussetzung für eine Fourierreihen-Darstellung ist somit nicht erfüllt. Wendet man den Formalismus trotzdem an, ergibt sich eine künstliche Periodisierung des Potentials, die im Ortsraum der Superposition $\sum_{\mathbf{R}} v_{\text{Coul}}(\mathbf{r} + \mathbf{R})$ über alle Gittervektoren \mathbf{R} entsprechen würde. (ii) Auf Grund des schwachen Abfallverhaltens von $v_{\text{Coul}}(\mathbf{r})$ geht diese Summe jedoch gegen ∞ für alle $\mathbf{r} \in \mathbb{R}^3$, was sich in der Singularität von [\(A.1\)](#) für die ($\mathbf{G} = \mathbf{0}$)-Komponente widerspiegelt. Man behilft sich daher meist mit der Vorstellung, dass das Volumen Ω insgesamt ladungsneutral ist, dass also zu jeder Ladung eine entsprechende Ladung mit entgegengesetztem Vorzeichen existiert. Dann heben sich die Singularitäten gegenseitig weg, und es werden nur noch die Komponenten mit $\mathbf{G} \neq \mathbf{0}$ berücksichtigt. Sollen allgemeine Ergebnisse speziell für ebene Wellen formuliert werden — wie im Folgenden der Austauschoperator —, substituiert man daher näherungsweise

$$\frac{1}{r} \longrightarrow \frac{4\pi}{\Omega} \sum_{\mathbf{G} \neq \mathbf{0}} \frac{1}{\mathbf{G}^2} e^{+i\mathbf{G}\mathbf{r}}. \quad (\text{A.2})$$

A.2 Der nicht-lokale Austauschoperator

Es sei zunächst bemerkt, dass der Austauschoperator [\(1.27\)](#) bei einem gitterperiodischen System zwar invariant gegenüber einer *gemeinsamen* Transformation von \mathbf{r} und \mathbf{r}' nach

$\mathbf{r}+\mathbf{R}$ bzw. $\mathbf{r}'+\mathbf{R}$ ist, nicht jedoch gegenüber einer *getrennten* Transformation dieser Koordinaten. Er ist also *nicht* gitterperiodisch in dem Sinne, dass $V_{\text{Fock}}(\mathbf{r}, \mathbf{r}') = V_{\text{Fock}}(\mathbf{r}+\mathbf{R}, \mathbf{r}') = V_{\text{Fock}}(\mathbf{r}, \mathbf{r}'+\mathbf{R})$ für alle Gittervektoren \mathbf{R} gilt, und kann daher streng genommen nicht über der Einheitszelle fourierentwickelt werden. Ebensovienig ist eine Fourierentwicklung über dem periodischen Grundgebiet (vgl. Abschnitt 3.1) möglich. Soll er auf eine Blochfunktion $\psi_{n\mathbf{k}}$ angewandt werden, sind allerdings nur jene Komponenten von V_{Fock} von Interesse, die sich aus seiner Projektion auf die $|\mathbf{G}+\mathbf{k}$ -Basis ergeben. Diese lassen sich, wie nun gezeigt, mit Hilfe von (A.2) näherungsweise bestimmen. Für jeden \mathbf{k} -Punkt ergibt sich dabei ein eigener nicht-lokaler Ausdruck.¹ Es ist:

$$\begin{aligned}
\langle \mathbf{G}+\mathbf{k} | V_{\text{Fock}} | \mathbf{G}'+\mathbf{k} \rangle &= \tag{A.3} \\
&= \frac{1}{\Omega} \iint e^{-i(\mathbf{G}+\mathbf{k})\cdot\mathbf{r}} \left(- \sum_{n,\mathbf{q}} \frac{\psi_{n\mathbf{q}}(\mathbf{r}) \overline{\psi_{n\mathbf{q}}(\mathbf{r}')}}{|\mathbf{r}-\mathbf{r}'|} \right) e^{i(\mathbf{G}'+\mathbf{k})\cdot\mathbf{r}'} d^3r d^3r' \\
&= \frac{1}{\Omega} \iint e^{-i(\mathbf{G}+\mathbf{k})\cdot\mathbf{r}} \left(- \sum_{n,\mathbf{q}} \frac{\frac{1}{\Omega} \left[\sum_{\mathbf{G}_1} c_{n\mathbf{q}}(\mathbf{G}_1) e^{i(\mathbf{G}_1+\mathbf{q})\cdot\mathbf{r}} \right] \overline{\left[\sum_{\mathbf{G}_2} c_{n\mathbf{q}}(\mathbf{G}_2) e^{i(\mathbf{G}_2+\mathbf{q})\cdot\mathbf{r}'} \right]}}{|\mathbf{r}-\mathbf{r}'|} \right) \\
&\quad \cdot e^{i(\mathbf{G}'+\mathbf{k})\cdot\mathbf{r}'} d^3r d^3r'.
\end{aligned}$$

Nun wird das Coulomb-Potential $v_{\text{Coul}}(\mathbf{r}-\mathbf{r}') = |\mathbf{r}-\mathbf{r}'|^{-1}$ durch die periodisierte Version (A.2) ersetzt. Die Entwicklung erfolgt dabei aber — abhängig vom jeweiligen Summanden — in der $|\mathbf{G}_3+(\mathbf{k}-\mathbf{q})$ -Basis, um sich den mit \mathbf{k}, \mathbf{q} jeweils verschiedenen Perioden der Zähler anzupassen:

$$\begin{aligned}
\langle \mathbf{G}+\mathbf{k} | V_{\text{Fock}} | \mathbf{G}'+\mathbf{k} \rangle &\approx \\
&\approx -\frac{1}{\Omega^2} \sum_{n,\mathbf{q}} \iint_{\Omega} \sum_{\mathbf{G}_1} \sum_{\mathbf{G}_2} c_{n\mathbf{q}}(\mathbf{G}_1) \overline{c_{n\mathbf{q}}(\mathbf{G}_2)} \cdot \\
&\quad \cdot e^{-i(\mathbf{G}+\mathbf{k})\cdot\mathbf{r}} e^{i(\mathbf{G}_1+\mathbf{q})\cdot\mathbf{r}} e^{-i(\mathbf{G}_2+\mathbf{q})\cdot\mathbf{r}'} e^{i(\mathbf{G}'+\mathbf{k})\cdot\mathbf{r}'} \cdot \\
&\quad \cdot \left[\frac{4\pi}{\Omega} \sum_{\mathbf{G}_3}^{\mathbf{G}_3+\mathbf{k}-\mathbf{q} \neq 0} \frac{1}{|\mathbf{G}_3+\mathbf{k}-\mathbf{q}|^2} e^{i(\mathbf{G}_3+\mathbf{k}-\mathbf{q})\cdot(\mathbf{r}-\mathbf{r}')} \right] d^3r d^3r' \\
&= -\frac{4\pi}{\Omega^3} \sum_{n,\mathbf{q}} \iint_{\Omega} \sum_{\mathbf{G}_1} \sum_{\mathbf{G}_2} \sum_{\mathbf{G}_3}^{\mathbf{G}_3+\mathbf{k}-\mathbf{q} \neq 0} c_{n\mathbf{q}}(\mathbf{G}_1) \overline{c_{n\mathbf{q}}(\mathbf{G}_2)} \cdot \\
&\quad \cdot e^{-i(\mathbf{G}-\mathbf{G}_1-\mathbf{G}_3)\cdot\mathbf{r}} e^{i(\mathbf{G}'-\mathbf{G}_2-\mathbf{G}_3)\cdot\mathbf{r}'} \frac{1}{|\mathbf{G}_3+\mathbf{k}-\mathbf{q}|^2} d^3r d^3r'.
\end{aligned}$$

Substituiert man $\mathbf{G}'_1 := \mathbf{G} - \mathbf{G}_1$ und $\mathbf{G}'_2 := \mathbf{G}' - \mathbf{G}_2$, wodurch jeweils der Summationsbereich (alle reziproken Gittervektoren) nicht geändert wird, ergibt sich daraus

¹Bei radialsymmetrischen Systemen führt die Darstellung des Austauschoperators mit Hilfe der Radialanteile der Wellenfunktionen in völlig analoger Weise auf einen Satz vom Drehimpuls l abhängender Operatoren [78].

$$\begin{aligned}
\langle \mathbf{G} + \mathbf{k} | V_{\text{Fock}} | \mathbf{G}' + \mathbf{k} \rangle &\approx \\
&\approx -\frac{4\pi}{\Omega^3} \sum_{n, \mathbf{q}} \sum_{\mathbf{G}'_1} \sum_{\mathbf{G}'_2} \sum_{\mathbf{G}_3}^{\mathbf{G}_3 + \mathbf{k} - \mathbf{q} \neq 0} c_{n\mathbf{q}}(\mathbf{G} - \mathbf{G}'_1) \overline{c_{n\mathbf{q}}(\mathbf{G}' - \mathbf{G}'_2)} \cdot \\
&\quad \cdot \underbrace{\int_{\Omega} e^{-i(\mathbf{G}'_1 - \mathbf{G}_3) \cdot \mathbf{r}} d^3 r}_{= \Omega \delta_{\mathbf{G}'_1 \mathbf{G}_3}} \underbrace{\int_{\Omega} e^{i(\mathbf{G}'_2 - \mathbf{G}_3) \cdot \mathbf{r}'} d^3 r'}_{= \Omega \delta_{\mathbf{G}'_2 \mathbf{G}_3}} \frac{1}{|\mathbf{G}_3 + \mathbf{k} - \mathbf{q}|^2} \\
&= -\frac{4\pi}{\Omega} \sum_{n, \mathbf{q}} \sum_{\mathbf{G}_3}^{\mathbf{G}_3 + \mathbf{k} - \mathbf{q} \neq 0} \frac{c_{n\mathbf{q}}(\mathbf{G} - \mathbf{G}_3) \overline{c_{n\mathbf{q}}(\mathbf{G}' - \mathbf{G}_3)}}{|\mathbf{G}_3 + \mathbf{k} - \mathbf{q}|^2}.
\end{aligned}$$

Mit der Substitution $\mathbf{G}'' := -\mathbf{G}_3$ erhält man schließlich

$$\langle \mathbf{G} + \mathbf{k} | V_{\text{Fock}} | \mathbf{G}' + \mathbf{k} \rangle \approx -\frac{4\pi}{\Omega} \sum_{n, \mathbf{q}} \sum_{\mathbf{G}''}^{\mathbf{k} - \mathbf{q} - \mathbf{G}'' \neq 0} \frac{c_{n\mathbf{q}}(\mathbf{G} + \mathbf{G}'') \overline{c_{n\mathbf{q}}(\mathbf{G}' + \mathbf{G}'')}}{|\mathbf{k} - \mathbf{q} - \mathbf{G}''|^2}, \quad (\text{A.4})$$

was dem in den Referenzen 52 und 98 angegebenen Ausdruck entspricht. Die Summation über \mathbf{q} stellt dabei eigentlich eine Integration über die 1. Brillouin-Zone dar, wobei über einen Pol 2. Ordnung hinwegintegriert wird. Eine von Gygi und Baldereschi [98] entwickelte numerische Behandlung dieser Schwierigkeit ist in Abschnitt 4.4.1 ausgeführt. Dort wird die Relation

$$c_{n, \mathbf{k} + \mathbf{G}_0}(\mathbf{G}) = c_{n, \mathbf{k}}(\mathbf{G} + \mathbf{G}_0) \quad (\text{A.5})$$

verwendet, die der Vollständigkeit halber hier bewiesen werden soll. Es ist

$$\begin{aligned}
\psi_{n, \mathbf{k}}(\mathbf{r}) &= \psi_{n, \mathbf{k} + \mathbf{G}}(\mathbf{r}) \\
\iff u_{n, \mathbf{k}}(\mathbf{r}) e^{i\mathbf{k} \cdot \mathbf{r}} &= u_{n, \mathbf{k} + \mathbf{G}}(\mathbf{r}) e^{i(\mathbf{k} + \mathbf{G}) \cdot \mathbf{r}} \\
\iff u_{n, \mathbf{k} + \mathbf{G}}(\mathbf{r}) &= u_{n, \mathbf{k}}(\mathbf{r}) e^{-i\mathbf{G} \cdot \mathbf{r}}.
\end{aligned} \quad (\text{A.6})$$

Mit der Fourierentwicklung (3.16) ergibt sich dann

$$\begin{aligned}
c_{n, \mathbf{k} + \mathbf{G}_0}(\mathbf{G}) &= \frac{1}{\sqrt{\Omega}} \int_{\Omega} u_{n, \mathbf{k} + \mathbf{G}_0}(\mathbf{r}) e^{-i\mathbf{G} \cdot \mathbf{r}} d^3 r \\
&\stackrel{(\text{A.6})}{=} \frac{1}{\sqrt{\Omega}} \int_{\Omega} \left(u_{n, \mathbf{k}}(\mathbf{r}) e^{-i\mathbf{G}_0 \cdot \mathbf{r}} \right) e^{-i\mathbf{G} \cdot \mathbf{r}} d^3 r \\
&= \frac{1}{\sqrt{\Omega}} \int_{\Omega} u_{n, \mathbf{k}}(\mathbf{r}) e^{-i(\mathbf{G} + \mathbf{G}_0) \cdot \mathbf{r}} d^3 r = c_{n, \mathbf{k}}(\mathbf{G} + \mathbf{G}_0),
\end{aligned} \quad (\text{A.7})$$

was zu beweisen war.

Symmetrisierung des Austauschoperators

Benötigt wird noch die Symmetrisierung des Austauschoperators, die sich aus (A.4) mit einer Aufspaltung der Summe über \mathbf{q} in eine gewichtete Summe über einen irreduziblen Bereich der Brillouin-Zone sowie einer weiteren über alle Punkttransformationen (vgl.

Abschnitt 3.2) ergibt:

$$\begin{aligned}
\hat{V}_{\text{Fock}}(\mathbf{k}, \mathbf{G}, \mathbf{G}') &= -\frac{4\pi}{\Omega} \sum_n^{\text{occ}} \sum_{s=1}^{N_{\text{red}}} \sum_{i=1}^{N_{\text{PS}}} \sum_{\mathbf{G}''} w_s \frac{c_{n, S_i \mathbf{q}_s}(\mathbf{G} + \mathbf{G}'') \overline{c_{n, S_i \mathbf{q}_s}(\mathbf{G}' + \mathbf{G}'')}}{|\mathbf{k} - S_i \mathbf{q}_s - \mathbf{G}''|^2} \quad (\text{A.8}) \\
&= -\frac{4\pi}{\Omega} \sum_n^{\text{occ}} \sum_{s=1}^{N_{\text{red}}} \sum_{i=1}^{N_{\text{PS}}} \sum_{\mathbf{G}''} w_s \frac{c_{n, \mathbf{q}_s}(S_i^{-1} \mathbf{G} + S_i^{-1} \mathbf{G}'') \overline{c_{n, \mathbf{q}_s}(S_i^{-1} \mathbf{G}' + S_i^{-1} \mathbf{G}'')}}{|S_i^{-1} \mathbf{k} - \mathbf{q}_s - S_i^{-1} \mathbf{G}''|^2} \\
&= -\frac{4\pi}{\Omega} \sum_n^{\text{occ}} \sum_{s=1}^{N_{\text{red}}} \sum_{i=1}^{N_{\text{PS}}} \sum_{\mathbf{G}''} w_s \frac{c_{n, \mathbf{q}_s}(S_i^{-1} \mathbf{G} + \mathbf{G}'') \overline{c_{n, \mathbf{q}_s}(S_i^{-1} \mathbf{G}' + \mathbf{G}'')}}{|S_i^{-1} \mathbf{k} - \mathbf{q}_s - \mathbf{G}''|^2}.
\end{aligned}$$

Die hier benutzte Relation $c_{n, S_i \mathbf{q}}(\mathbf{G}) = c_{n, \mathbf{q}}(S_i^{-1} \mathbf{G})$, also das Gegenstück zu (3.6) im reziproken Raum, ist in Ref. 91 hergeleitet.

A.3 Die lineare Antwort

Sollen nur Potential- und Ladungsdichteänderungen betrachtet werden, die periodisch mit der Einheitszelle sind, ist die lineare Antwort für einen nicht spin-polarisierten Halbleiter oder Isolator in der Ortsdarstellung durch

$$\chi_0(\mathbf{r}, \mathbf{r}') = 2 \sum_{v\mathbf{c}\mathbf{k}} \frac{\overline{\psi_{v\mathbf{k}}(\mathbf{r})} \psi_{c\mathbf{k}}(\mathbf{r}) \overline{\psi_{c\mathbf{k}}(\mathbf{r}')} \psi_{v\mathbf{k}}(\mathbf{r}') + \text{c.c.}}{\varepsilon_{v\mathbf{k}} - \varepsilon_{c\mathbf{k}}} \quad (\text{A.9})$$

gegeben [52, 176]. Dabei ist v der Index der Valenzbänder und c der der Leitungsbänder. Im Gegensatz zum nicht-lokalen Austausch-Operator kann dieser Ausdruck exakt in den reziproken Raum transformiert werden, da die beiden Eingänge unabhängig voneinander Gitterperiodizität besitzen, d.h., es gilt $\chi_0(\mathbf{r}, \mathbf{r}') = \chi_0(\mathbf{r} + \mathbf{R}, \mathbf{r}') = \chi_0(\mathbf{r}, \mathbf{r}' + \mathbf{R})$ für beliebige Gittervektoren \mathbf{R} . Die Fourier-Entwicklung ist dann gegeben durch

$$\begin{aligned}
\hat{\chi}_0(\mathbf{G}, \mathbf{G}') &= \quad (\text{A.10}) \\
&= \left(\frac{1}{\sqrt{\Omega}} \right)^2 \iint_{\Omega} e^{-i\mathbf{G} \cdot \mathbf{r}} \chi_0(\mathbf{r}, \mathbf{r}') e^{+i\mathbf{G}' \cdot \mathbf{r}'} d^3 r d^3 r' \\
&= \frac{1}{\Omega} \iint_{\Omega} 2 \sum_{v\mathbf{c}\mathbf{k}} \left[\frac{\overline{\psi_{v\mathbf{k}}(\mathbf{r})} e^{-i\mathbf{G} \cdot \mathbf{r}} \psi_{c\mathbf{k}}(\mathbf{r}) \cdot \overline{\psi_{c\mathbf{k}}(\mathbf{r}')} e^{+i\mathbf{G}' \cdot \mathbf{r}'} \psi_{v\mathbf{k}}(\mathbf{r}')}{\varepsilon_{v\mathbf{k}} - \varepsilon_{c\mathbf{k}}} + \right. \\
&\quad \left. + \frac{\psi_{v\mathbf{k}}(\mathbf{r}) e^{-i\mathbf{G} \cdot \mathbf{r}} \overline{\psi_{c\mathbf{k}}(\mathbf{r})} \cdot \psi_{c\mathbf{k}}(\mathbf{r}') e^{+i\mathbf{G}' \cdot \mathbf{r}'} \overline{\psi_{v\mathbf{k}}(\mathbf{r}')}}{\varepsilon_{v\mathbf{k}} - \varepsilon_{c\mathbf{k}}} \right] d^3 r d^3 r' \\
&= \frac{2}{\Omega} \sum_{v\mathbf{c}\mathbf{k}} \left[\frac{\left(\int \overline{\psi_{v\mathbf{k}}(\mathbf{r})} e^{-i\mathbf{G} \cdot \mathbf{r}} \psi_{c\mathbf{k}}(\mathbf{r}) d^3 r \right) \left(\int \overline{\psi_{c\mathbf{k}}(\mathbf{r}')} e^{+i\mathbf{G}' \cdot \mathbf{r}'} \psi_{v\mathbf{k}}(\mathbf{r}') d^3 r' \right)}{\varepsilon_{v\mathbf{k}} - \varepsilon_{c\mathbf{k}}} + \right. \\
&\quad \left. + \frac{\left(\int \psi_{v\mathbf{k}}(\mathbf{r}) e^{-i\mathbf{G} \cdot \mathbf{r}} \overline{\psi_{c\mathbf{k}}(\mathbf{r})} d^3 r \right) \left(\int \psi_{c\mathbf{k}}(\mathbf{r}') e^{+i\mathbf{G}' \cdot \mathbf{r}'} \overline{\psi_{v\mathbf{k}}(\mathbf{r}')} d^3 r' \right)}{\varepsilon_{v\mathbf{k}} - \varepsilon_{c\mathbf{k}}} \right].
\end{aligned}$$

Dieses Ergebnis lässt sich mit der Zeitinversionssymmetrie (3.7) noch vereinfachen. Im zweiten Summanden wird dazu die Summation über \mathbf{k} durch die Summation über $-\mathbf{k}$ ersetzt (was erlaubt ist wegen $\mathbf{k} \in \text{BZ} \Rightarrow -\mathbf{k} \in \text{BZ}$):

$$\begin{aligned}
\hat{\chi}_0(\mathbf{G}, \mathbf{G}') &= \tag{A.11} \\
&= \frac{2}{\Omega} \sum_{vc} \left[\sum_{\mathbf{k} \in \text{BZ}} \frac{\left(\int \overline{\psi_{v\mathbf{k}}(\mathbf{r})} e^{-i\mathbf{G}\cdot\mathbf{r}} \psi_{c\mathbf{k}}(\mathbf{r}) d^3r \right) \left(\int \overline{\psi_{c\mathbf{k}}(\mathbf{r}')} e^{+i\mathbf{G}'\cdot\mathbf{r}'} \psi_{v\mathbf{k}}(\mathbf{r}') d^3r' \right)}{\varepsilon_{v\mathbf{k}} - \varepsilon_{c\mathbf{k}}} + \right. \\
&\quad \left. + \sum_{\mathbf{k} \in \text{BZ}} \frac{\left(\int \psi_{v,-\mathbf{k}}(\mathbf{r}) e^{-i\mathbf{G}\cdot\mathbf{r}} \overline{\psi_{c,-\mathbf{k}}(\mathbf{r})} d^3r \right) \left(\int \psi_{c,-\mathbf{k}}(\mathbf{r}') e^{+i\mathbf{G}'\cdot\mathbf{r}'} \overline{\psi_{v,-\mathbf{k}}(\mathbf{r}')} d^3r' \right)}{\varepsilon_{v,-\mathbf{k}} - \varepsilon_{c,-\mathbf{k}}} \right] \\
&\stackrel{(3.7)}{=} \frac{2}{\Omega} \sum_{vc} \left[\sum_{\mathbf{k} \in \text{BZ}} \frac{\left(\int \overline{\psi_{v\mathbf{k}}(\mathbf{r})} e^{-i\mathbf{G}\cdot\mathbf{r}} \psi_{c\mathbf{k}}(\mathbf{r}) d^3r \right) \left(\int \overline{\psi_{c\mathbf{k}}(\mathbf{r}')} e^{+i\mathbf{G}'\cdot\mathbf{r}'} \psi_{v\mathbf{k}}(\mathbf{r}') d^3r' \right)}{\varepsilon_{v\mathbf{k}} - \varepsilon_{c\mathbf{k}}} + \right. \\
&\quad \left. + \sum_{\mathbf{k} \in \text{BZ}} \frac{\left(\int e^{i\alpha_{v\mathbf{k}}} \overline{\psi_{v\mathbf{k}}(\mathbf{r})} e^{-i\mathbf{G}\cdot\mathbf{r}} e^{i\alpha_{c\mathbf{k}}} \overline{\psi_{c\mathbf{k}}(\mathbf{r})} d^3r \right) \left(\int e^{i\alpha_{c\mathbf{k}}} \overline{\psi_{c\mathbf{k}}(\mathbf{r}')} e^{+i\mathbf{G}'\cdot\mathbf{r}'} e^{i\alpha_{v\mathbf{k}}} \overline{\psi_{v\mathbf{k}}(\mathbf{r}')} d^3r' \right)}{\varepsilon_{v\mathbf{k}} - \varepsilon_{c\mathbf{k}}} \right].
\end{aligned}$$

Die durch die Zeitumkehrsymmetrie ins Spiel gekommenen Phasenfaktoren kürzen sich gegenseitig gerade heraus; beide Summanden sind also identisch, und man erhält schließlich

$$\begin{aligned}
\hat{\chi}_0(\mathbf{G}, \mathbf{G}') &= \tag{A.12} \\
&= \frac{4}{\Omega} \sum_{v\mathbf{c}\mathbf{k}} \frac{\left(\int \overline{\psi_{v\mathbf{k}}(\mathbf{r})} e^{-i\mathbf{G}\cdot\mathbf{r}} \psi_{c\mathbf{k}}(\mathbf{r}) d^3r \right) \left(\int \overline{\psi_{c\mathbf{k}}(\mathbf{r}')} e^{+i\mathbf{G}'\cdot\mathbf{r}'} \psi_{v\mathbf{k}}(\mathbf{r}') d^3r' \right)}{\varepsilon_{v\mathbf{k}} - \varepsilon_{c\mathbf{k}}}
\end{aligned}$$

oder, in einer üblichen abkürzenden, jedoch nicht ganz strengen Schreibweise,

$$\hat{\chi}_0(\mathbf{G}, \mathbf{G}') = \frac{4}{\Omega} \sum_{v\mathbf{c}\mathbf{k}} \frac{\langle v\mathbf{k} | e^{-i\mathbf{G}\cdot\mathbf{r}} | c\mathbf{k} \rangle \langle c\mathbf{k} | e^{+i\mathbf{G}'\cdot\mathbf{r}} | v\mathbf{k} \rangle}{\varepsilon_{v\mathbf{k}} - \varepsilon_{c\mathbf{k}}}. \tag{A.13}$$

Anhang B

Die Greensche Funktion

Greensche Funktionen werden über so genannte Fundamentallösungen linearer inhomogener Differentialgleichungen definiert. Dies sei hier nur an einem Beispiel illustriert:

Es soll etwa die Lösung u der Differentialgleichung

$$\left(\frac{d^2}{dx^2} + 1\right) u(x) = f(x) \quad (\text{B.1})$$

gefunden werden. Dazu bestimme man zunächst die Lösung E der Gleichung, die aus (B.1) entsteht, indem man die Inhomogenität f durch die δ -Distribution ersetzt:

$$\left(\frac{d^2}{dx^2} + 1\right) E(x) = \delta(x) . \quad (\text{B.2})$$

E wird als *Fundamentallösung* bezeichnet. Man erhält sie z.B. aus einer Fourier-Transformation von Gleichung (B.2):

$$(-k^2 + 1) \hat{E}(k) = \hat{\delta}(k) = \frac{1}{\sqrt{2\pi}} . \quad (\text{B.3})$$

Umstellen nach $\hat{E}(k)$ und inverses Fourier-Transformieren liefert dann E . Ohne Beweis sei nun angegeben, dass man die Lösung der inhomogenen Gleichung (B.1) durch Faltung der Fundamentallösung mit der Inhomogenität f erhält:

$$u(x) = (E * f)(x) := \int E(x - y) f(y) dy = \int G(x, y) f(y) dy . \quad (\text{B.4})$$

Der auf der linken Seite eingeführte Integralkern $G(x, y) := E(x - y)$ heißt *Greensche Funktion*. Genauer entnehmen man der Spezialliteratur, z.B. Ref. 123.

In der Quantenmechanik ist man häufig an der Greenschen Funktion G_ε interessiert, die sich als Lösung der Gleichung

$$-(H(\mathbf{r}) - \varepsilon) G_\varepsilon(\mathbf{r}, \mathbf{r}') = \delta(\mathbf{r} - \mathbf{r}') \quad (\text{B.5})$$

ergibt, wobei H der Hamilton-Operator in Ortsdarstellung sein soll. Die δ -Distribution lässt sich formal als kontinuierliches Analogon zur Einheitsmatrix im Endlich-Dimensionalen auffassen, womit die Greensche Funktion in gewisser Weise zur Inversen des Hamilton-Operators wird:

$$G_\varepsilon = -(H - \varepsilon \cdot \mathbf{1})^{-1} \quad (\text{für } \varepsilon \text{ nicht aus dem Spektrum von } H) . \quad (\text{B.6})$$

Sind die Eigenwerte (das Spektrum) von H durch $\{\varepsilon_n\}$ gegeben, und stellt man sich H als „echte“ (diskrete) Matrix vor, erkennt man sofort, dass G_{ε_n} strenggenommen nicht definiert ist: Mit $\det(H - \varepsilon_n \cdot \mathbf{1}) = 0$ kann keine Inverse zu $(H - \varepsilon_n \cdot \mathbf{1})$ existieren. Allerdings besteht in diesem Fall die Möglichkeit, zunächst ε_n durch $\varepsilon_n + i\eta$ zu ersetzen, um dann später den Grenzübergang $\eta \rightarrow 0$ vollziehen zu können, vgl. z.B. Ref. 92.

Das System $\{|n\rangle\}$ der Eigenfunktionen von H bildet ein vollständiges Orthonormalsystem im Hilbertraum \mathcal{H} . Mit der Spektraldarstellung des Hamilton-Operators

$$H = \sum_n |n\rangle \varepsilon_n \langle n| \quad (\text{B.7})$$

und dem *Greenschen Operator*

$$G_\varepsilon := \sum_n \frac{|n\rangle \langle n|}{\varepsilon - \varepsilon_n} \quad (\varepsilon \notin \text{Spektrum von } H) \quad (\text{B.8})$$

findet man

$$-G_\varepsilon (H - \varepsilon \cdot \mathbf{1}) = \sum_{n,m} \frac{|n\rangle \langle n|m\rangle (\varepsilon_n - \varepsilon) \langle m|}{\varepsilon - \varepsilon_n} = \mathbf{1} = -(H - \varepsilon \cdot \mathbf{1}) G_\varepsilon. \quad (\text{B.9})$$

Schreibt man den Greenschen Operator (B.8) in Ortsdarstellung, wird daraus die Greensche Funktion

$$G_\varepsilon(\mathbf{r}, \mathbf{r}') = \langle \mathbf{r} | G_\varepsilon | \mathbf{r}' \rangle = \sum_n \frac{\phi_n(\mathbf{r}) \phi_n^*(\mathbf{r}')}{\varepsilon - \varepsilon_n} \quad (\varepsilon \notin \text{Spektrum von } H), \quad (\text{B.10})$$

wobei $\phi_n(\mathbf{r}) := \langle \mathbf{r} | n \rangle$ gesetzt wurde. Für $\varepsilon = \varepsilon_m$ ist die Greensche Funktion nur im Unterraum \mathcal{H}_m^\perp definierbar, der von allen Elementen aus \mathcal{H} gebildet wird, die auf dem Eigenraum von ε_m senkrecht stehen:

$$G_m(\mathbf{r}, \mathbf{r}') = \sum_{\substack{n \\ \varepsilon_n \neq \varepsilon_m}} \frac{\phi_n(\mathbf{r}) \phi_n^*(\mathbf{r}')}{\varepsilon_m - \varepsilon_n}. \quad (\text{B.11})$$

An vielen Stellen im Text wird die Projektion der Greenschen Funktion G_m auf den Teilraum \mathcal{H}^c der unbesetzten Zustände verwendet, wobei m einem besetzten Zustand zugeordnet ist. Bezeichnet $P_c := \sum_i^{\text{unocc}} |n\rangle \langle n|$ den Projektionsoperator auf \mathcal{H}^c , ist sie durch

$$G_m^\perp(\mathbf{r}, \mathbf{r}') := P_c G_m(\mathbf{r}, \mathbf{r}') = \sum_n \frac{\phi_n(\mathbf{r}) \phi_n^*(\mathbf{r}')}{\varepsilon_m - \varepsilon_n} \quad (\text{B.12})$$

gegeben.

Berechnung von Variationsableitungen mit Hilfe der Störungstheorie in 1. Ordnung

An dieser Stelle sei auch kurz auf den Zusammenhang zwischen der Störungstheorie 1. Ordnung und der Variationsableitung einer Eigenfunktion ϕ_i eines Operators der Form $H := -\frac{1}{2}\nabla^2 + v$ eingegangen, wobei v ein lokaler, also multiplikativ wirkender Operator sei. Dazu führt man eine kleine Variation Δv von v ein, die sich als Störung auffassen lässt. Die Eigenwert-Gleichung des gestörten Systems schreibt sich dann

$$\left[-\frac{1}{2}\nabla^2 + v + \Delta v\right](\phi_i + \Delta\phi_i) = (\varepsilon_i + \Delta\varepsilon_i)(\phi_i + \Delta\phi_i). \quad (\text{B.13})$$

Die der Variation Δv entsprechende Änderung $\Delta\phi_i$ der Eigenfunktion ϕ_i ergibt sich also als „Zustandskorrektur“, für die die Störungstheorie in *linearer* Ordnung die folgende näherungsweise Darstellung liefert:

$$\Delta\phi_i(\mathbf{r}) \approx \sum_{\substack{j \\ \varepsilon_j \neq \varepsilon_i}} \phi_j(\mathbf{r}) \frac{\langle \phi_j | \Delta v | \phi_i \rangle}{\varepsilon_i - \varepsilon_j} = \int_{\mathbf{r}' \in \mathbb{R}^3} \left(\sum_{\substack{j \\ \varepsilon_j \neq \varepsilon_i}} \phi_j(\mathbf{r}) \frac{\phi_j^*(\mathbf{r}') \phi_i(\mathbf{r}')}{\varepsilon_i - \varepsilon_j} \right) \Delta v(\mathbf{r}') d^3 r'. \quad (\text{B.14})$$

Die Summation erfolgt über alle zum Eigenraum von ε_i orthogonalen Zustände. Die Approximation (B.14) wird dabei immer besser, je kleiner die Variation Δv ist. Die gesuchte Variationsableitung entspricht daher gerade dem Ausdruck in den großen runden Klammern und kann mit Hilfe der Greenschen Funktion (B.11) geschrieben werden:

$$\frac{\delta\phi_i(\mathbf{r})}{\delta v(\mathbf{r}')} = \sum_{\substack{j \\ \varepsilon_j \neq \varepsilon_i}} \phi_j(\mathbf{r}) \frac{\phi_j^*(\mathbf{r}') \phi_i(\mathbf{r}')}{\varepsilon_i - \varepsilon_j} = G_i(\mathbf{r}, \mathbf{r}') \phi_i(\mathbf{r}'). \quad (\text{B.15})$$

Man beachte, dass die Variationsableitung $\delta\phi_i/\delta v$ ein Ausdruck ist, der von zwei Variablen — hier \mathbf{r} und \mathbf{r}' — abhängt. Dies ist in vollkommener Analogie zu der Ableitung einer Funktion mehrerer Variablen, die sich bekanntlich als Matrix darstellen lässt (vgl. z.B. Ref. 123).

Anhang C

Die Sternheimer-Gleichung

C.1 Die Sternheimer-Gleichung aus dem Variationsprinzip von Ritz

In Abschnitt 4.1 wurde zur numerisch effizienten Lösung der Sternheimer-Gleichung (4.1) auf dem Hilbertraum \mathcal{H} ein Funktional $E_i : \mathcal{H} \rightarrow \mathbb{R}; \psi_i^1 \mapsto E_i[\psi_i^1]$ eingeführt, das die Einheit einer Energie hat und eine quadratische Form in der linearen Korrekturfunktion ψ_i^1 darstellt. Einheit und mathematische Gestalt ergaben sich dabei rein abstrakt, indem E_i in (4.3) als reellwertige Stammfunktion der linken Seite der Sternheimer-Gleichung konstruiert wurde.

Hier möchte ich jedoch noch eine weitere Herleitung dieses Funktionals geben, die, obschon sie etwas aufwendiger ist, eine physikalisch anschaulichere Interpretation erlaubt. Es sei gezeigt, dass E_i — und damit die Sternheimer-Gleichung selbst ebenfalls — auch direkt aus dem Variationsprinzip von Ritz (vgl. Abschnitt 1.2) hergeleitet werden kann, wenn dort die in der Rayleigh-Schrödingerschen Störungstheorie benutzten Reihenentwicklungen eingesetzt werden.

Die Idee dabei ist folgende: Gemäß dem Variationsprinzip (1.11) stellen sich die (normierten) Eigenfunktionen eines Systems gerade so ein, dass das Energiefunktional $\langle \psi | H | \psi \rangle$ minimal ist, wenn $|\psi\rangle$ einem Grundzustand entspricht, und zumindest stationär ist, wenn $|\psi\rangle$ einen angeregten Zustand beschreibt. Dies gilt natürlich für das ungestörte System ebenso wie für das gestörte. Wenn aber die Größen des ungestörten Systems bereits bekannt sind, könnten diese bei der Formulierung des gestörten Problems fest eingesetzt werden, und die Variation brauchte nur noch über die Korrekturen der Wellenfunktionen zu erfolgen. Die Variation der Grundzustands*korrektur* würde ein Minimum des für das gestörte Problem formulierten $\langle \psi | H | \psi \rangle$ liefern (wobei ψ die normierte Summe aus ungestörter Wellenfunktion und Korrektur ist), und die Variation von *Korrekturen* höherer Zustände würde auf stationäre Punkte führen — dies entspricht genau dem Verhalten des in (4.3) für die lineare Korrektur ψ_i^1 gefundenen Energiefunktional! Tatsächlich ergibt sich (4.3) direkt aus der hier skizzierten Idee, wenn man nur Korrekturfunktionen bis zur 1. Ordnung, d.h. also, im Energiefunktional $\langle \psi | H | \psi \rangle$ Ausdrücke noch bis zur 2. Ordnung berücksichtigt.

Dazu sei zunächst das gestörte System unter denselben Annahmen wie bei der Rayleigh-Schrödingerschen Störungstheorie nach einem reellen Parameter λ entwickelt:

$$\begin{aligned} H &:= H_0 + \lambda H_1 \\ \varepsilon_i &:= \varepsilon_i^0 + \lambda \varepsilon_i^1 + \lambda^2 \varepsilon_i^2 + \dots, \end{aligned} \tag{C.1}$$

$$|\psi_i\rangle := |\psi_i^0\rangle + \lambda |\psi_i^1\rangle + \lambda^2 |\psi_i^2\rangle + \dots .$$

Hier bezeichne wie üblich H_0 den Hamilton-Operator des ungestörten Systems, dem eine Störung λH_1 aufgeschaltet sei. Eigenwerte ε_i^0 und Eigenfunktionen $|\psi_i^0\rangle$ des ungestörten Systems seien bekannt. Nach dem Variationsprinzip (1.11) wird das Funktional $\langle\psi|(H_0 + H_1)|\psi\rangle$ stationär für alle Eigenfunktionen $|\psi_i\rangle$ des gestörten Systems, wobei die Variation unter der Nebenbedingung der Normierung erfolgt, d.h. also, wenn nur Wellenfunktionen auf der „Kugel“ $\langle\psi|\psi\rangle = \|\psi\|^2 = 1$ im Hilbertraum \mathcal{H} zugelassen sind.¹ Wird die Nebenbedingung über einen Lagrange-Multiplikator ε berücksichtigt, hat man

$$\left. \frac{\delta E_\varepsilon[\psi]}{\delta\psi^*} \right|_{\substack{\varepsilon=\varepsilon_i \\ \psi=\psi_i}} = 0 \quad \text{mit} \quad E_\varepsilon[\psi] := \langle\psi|(H_0 + \lambda H_1)|\psi\rangle - \varepsilon \cdot [\langle\psi|\psi\rangle - 1] \quad (\text{C.2})$$

für alle Eigenenergien ε_i und -funktionen $|\psi_i\rangle$ des gestörten Systems. Einsetzen von (C.1) in (C.2) liefert

$$\begin{aligned} E_\varepsilon[\psi] &= (\langle\psi^0| + \lambda\langle\psi^1| + \lambda^2\langle\psi^2| + \dots)(H_0 + \lambda H_1)(|\psi^0\rangle + \lambda|\psi^1\rangle + \lambda^2|\psi^2\rangle + \dots) \\ &\quad - (\varepsilon^0 + \lambda\varepsilon^1 + \lambda^2\varepsilon^2 + \dots)(\langle\psi^0| + \lambda\langle\psi^1| + \lambda^2\langle\psi^2| + \dots) \\ &\quad \quad \quad (|\psi^0\rangle + \lambda|\psi^1\rangle + \lambda^2|\psi^2\rangle + \dots) \\ &\quad + (\varepsilon^0 + \lambda\varepsilon^1 + \lambda^2\varepsilon^2 + \dots) \\ &= \langle\psi^0|H_0|\psi^0\rangle + \lambda\langle\psi^0|H_0|\psi^1\rangle + \lambda\langle\psi^1|H_0|\psi^0\rangle + \lambda\langle\psi^0|H_1|\psi^0\rangle \\ &\quad + \lambda^2\langle\psi^0|H_0|\psi^2\rangle + \lambda^2\langle\psi^1|H_0|\psi^1\rangle + \lambda^2\langle\psi^2|H_0|\psi^0\rangle \\ &\quad + \lambda^2\langle\psi^0|H_1|\psi^1\rangle + \lambda^2\langle\psi^1|H_1|\psi^0\rangle \\ &\quad - \varepsilon^0\langle\psi^0|\psi^0\rangle - \lambda\varepsilon^0\langle\psi^0|\psi^1\rangle - \lambda\varepsilon^0\langle\psi^1|\psi^0\rangle - \lambda\varepsilon^1\langle\psi^0|\psi^0\rangle \\ &\quad - \lambda^2\varepsilon^0\langle\psi^0|\psi^2\rangle - \lambda^2\varepsilon^0\langle\psi^1|\psi^1\rangle - \lambda^2\varepsilon^0\langle\psi^2|\psi^0\rangle \\ &\quad - \lambda^2\varepsilon^1\langle\psi^0|\psi^1\rangle - \lambda^2\varepsilon^1\langle\psi^1|\psi^0\rangle \\ &\quad + (\varepsilon^0 + \lambda\varepsilon^1 + \lambda^2\varepsilon^2) + O(\lambda^3) . \end{aligned} \quad (\text{C.3})$$

Hier wurde der Übersichtlichkeit halber der Index i an den Wellenfunktionen und Energien unterdrückt. Im Restglied $O(\lambda^3)$ sind alle Terme kubischer oder höherer Ordnung in λ zusammengefasst. Unter Beachtung von $H_0\psi^0 = \varepsilon^0\psi^0$ lassen sich eine Reihe von Gliedern herauskürzen. Vernachlässigt man ferner das Restglied $O(\lambda^3)$, erhält man in quadratischer Näherung:

$$\begin{aligned} E_\varepsilon[\psi] &\approx \lambda\langle\psi^0|(H_1 - \varepsilon^1)|\psi^0\rangle + \lambda^2\langle\psi^1|(H_0 - \varepsilon^0)|\psi^1\rangle \\ &\quad + \lambda^2\langle\psi^0|(H_1 - \varepsilon^1)|\psi^1\rangle + \lambda^2\langle\psi^1|(H_1 - \varepsilon^1)|\psi^0\rangle + (\varepsilon^0 + \lambda\varepsilon^1 + \lambda^2\varepsilon^2) . \end{aligned}$$

Hier liefern alle Glieder, die ψ^1 weder als Bra- noch als Ket-Vektor explizit enthalten, nur einen konstanten Beitrag zum Energiefunktional. Die Lage der stationären Punkte in \mathcal{H} ist aber invariant gegenüber einer Eichung der Energieskala; deshalb können die konstanten

¹In Abschnitt 4.1 wurde die Sternheimer-Gleichung für ein effektives Ein-Teilchen-Problem (Kohn-Sham-System) formuliert, in diesem Falle bezeichnet ψ also eine Ein-Teilchen-Wellenfunktion. In der folgenden Diskussion kann es sich jedoch auch allgemein um eine N -Teilchen-Wellenfunktion handeln.

Glieder ignoriert werden, und man darf das Funktional noch mit einem beliebigen Faktor, z.B. $1/(2\lambda^2)$, multiplizieren. Insgesamt kann man statt $E_\varepsilon[\psi]$ also auch

$$\begin{aligned}\tilde{E}_\varepsilon[\psi^1] &:= \frac{1}{2}\langle\psi^1|(H_0 - \varepsilon^0)|\psi^1\rangle + \frac{1}{2}\langle\psi^0|(H_1 - \varepsilon^1)|\psi^1\rangle + \frac{1}{2}\langle\psi^1|(H_1 - \varepsilon^1)|\psi^0\rangle \\ &= \frac{1}{2}\langle\psi^1|(H_0 - \varepsilon^0)|\psi^1\rangle + \operatorname{Re}\langle\psi^1|(H_1 - \varepsilon^1)|\psi^0\rangle\end{aligned}\quad (\text{C.4})$$

verwenden. Da auf der rechten Seite nur die lineare Korrekturfunktion variabel ist, wurde als Argument von \tilde{E}_ε hier gleich ψ^1 benutzt. Fügt man den Index i an Energien und Zustände wieder an, ist dies in der Tat identisch mit dem in Abschnitt 4.1 angegebenen Energiefunktional (4.3), was zu zeigen war.

Noch eine Bemerkung zur Normierung: Die hier benutzte Nebenbedingung $\langle\psi|\psi\rangle = 1$ ist im Allgemeinen *nicht* mit der üblicherweise in der Störungstheorie angesetzten Normierung $\langle\psi^0|\psi\rangle = N$ konsistent, aus der man die Orthogonalität der Korrekturfunktionen zu ψ^0 erhält, vgl. z.B. Ref. 127. Deshalb wurde bei der Vereinfachung des obigen Ausdrucks (C.3) auch nicht mit der (nicht gegebenen) Orthogonalität zwischen der Korrekturfunktion $|\psi^1\rangle$ und der ungestörten Funktion $|\psi^0\rangle$ argumentiert! Fordert man Orthogonalität zwischen der Korrektur $|\Delta\psi\rangle := |\psi\rangle - |\psi^0\rangle$ und der ungestörten Funktion $|\psi^0\rangle$, kann die Normierung im Allgemeinen nicht exakt erfüllt werden (es muss „Nachnormiert“ werden), denn es gilt dann

$$\begin{aligned}\langle\psi|\psi\rangle &= (\langle\psi^0| + \langle\Delta\psi|)(|\psi^0\rangle + |\Delta\psi\rangle) \\ &= \langle\psi^0|\psi^0\rangle + \langle\psi^0|\Delta\psi\rangle + \langle\Delta\psi|\psi^0\rangle + \langle\Delta\psi|\Delta\psi\rangle \\ &= \underbrace{\langle\psi^0|\psi^0\rangle}_{=1} + 2\operatorname{Re}\underbrace{\langle\psi^0|\Delta\psi\rangle}_{=0} + \langle\Delta\psi|\Delta\psi\rangle \\ &= 1 + \langle\Delta\psi|\Delta\psi\rangle.\end{aligned}\quad (\text{C.5})$$

Wegen $|\Delta\psi\rangle = \lambda|\psi^1\rangle + \lambda^2|\psi^2\rangle + \dots$ enthält die Abweichung $\langle\Delta\psi|\Delta\psi\rangle$ allerdings nur quadratische und höhere Ordnungen in λ , oder anders formuliert: in *linearer Näherung* lässt sich die Normierungsbedingung zugleich mit der Orthogonalitätsforderung erfüllen. Genau dies wurde in Kapitel 2 bei der Definition (2.4) der induzierten Ladungsdichte $\Delta\rho$ ausgenutzt: $\Delta\rho$ beschreibt den Einfluss von H_1 nur in *linearer* Ordnung, daher ist die Ladungserhaltung (2.9) exakt erfüllt, obwohl zur Berechnung der induzierten Ladungsdichte die Sternheimer-Gleichung zusammen mit der Orthogonalitätsforderung gelöst wurde, vgl. Abschnitt 4.1. Es sei an dieser Stelle nochmals darauf hingewiesen, dass dies aber keine Näherung der EXX-Methode darstellt, vielmehr ist die lineare Näherung der induzierten Ladungsdichte eine wesentliche Eigenschaft der EXX-Methode selbst.

C.2 Entartung

Am Ende des Abschnitts 4.1 wurde gezeigt, dass die Sternheimer-Gleichung im Rahmen der variationellen EXX-Methode nur in einem „verkleinerten“ Hilbertraum, nämlich dem der unbesetzten Zustände, gelöst zu werden braucht, und sich das Problem der Entartung auf diese Weise eliminieren lässt. Um jedoch die Diskussion der Sternheimer-Gleichung abzurunden, sei an dieser Stelle kurz beschrieben, wie im Falle einer Entartung auch die vollständigen Korrekturfunktionen berechnet werden können. Der zu Grunde liegende Formalismus ist natürlich die entartete Störungstheorie von Rayleigh-Schrödinger. Deren Ergebnisse sind aber üblicherweise mit Hilfe von Greenschen Funktionen formuliert. Das

Ziel hier besteht also darin, eine entsprechende Verallgemeinerung der in Abschnitt 4.1 entwickelten Methoden zu finden, um wieder ohne die Greensche Funktion auskommen zu können.

Eng mit der Entartung — gemeint ist eine mögliche Entartung eines Eigenwertes ε_i von H_0 , also dem Hamilton-Operator des *ungestörten* Systems — verknüpft ist die Frage, ob die Sternheimer-Gleichung (4.1) überhaupt eine Lösung besitzt, oder, äquivalent dazu, ob das in (4.3) definierte Energiefunktional E_i einen stationären Punkt (Minimum, Sattelpunkt oder Maximum) besitzt. Seien dazu noch einmal die drei in (4.6)–(4.8) definierten Teilräume \mathcal{H}^- , \mathcal{H}^0 und \mathcal{H}^+ betrachtet, von denen \mathcal{H}^0 gerade dem Eigenraum von ε_i entspricht. Gemäß Abschnitt 4.1 besitzt E_i in \mathcal{H}^- ein Maximum und in \mathcal{H}^+ ein Minimum, zu klären bleibt also nur noch das Verhalten in \mathcal{H}^0 . Hier reduziert sich das Funktional auf den linearen Anteil

$$E_i[\psi_i^1] = \frac{1}{2}\langle\psi_i^1|(H_1 - \varepsilon_i^1)|\psi_i^0\rangle + \frac{1}{2}\langle\psi_i^0|(H_1 - \varepsilon_i^1)|\psi_i^1\rangle \quad (\text{für } \psi_i^1 \in \mathcal{H}^0), \quad (\text{C.6})$$

dessen Gradient durch den konstanten Vektor

$$2\frac{\delta E_i[\psi_i^1]}{\delta(\psi_i^1)^*} = (H_1 - \varepsilon_i^1)|\psi_i^0\rangle \quad (\text{für } \psi_i^1 \in \mathcal{H}^0) \quad (\text{C.7})$$

gegeben ist. Eine Lösung der Sternheimer-Gleichung im gesamten Hilbertraum existiert also genau dann, wenn (C.7) identisch verschwindet. In \mathcal{H}^0 ergibt sich für das Energiefunktional in diesem Fall $E_i = 0$. Ist (C.7) jedoch von 0 verschieden, definiert (C.6) eine „schiefe (Hyper-)Ebene“ über \mathcal{H}^0 , d.h., es kann kein stationärer Punkt existieren, und darauf angesetzte Algorithmen zur Extremwertsuche können nicht konvergieren (sie laufen „ins Unendliche“).

Für nicht-entartetes ε_i^0 ergibt sich nun, dass (C.7) stets verschwindet, und die Sternheimer-Gleichung immer lösbar ist. Da dann nämlich der Eigenraum \mathcal{H}^0 eindimensional ist und nur durch den zugehörigen Eigenvektor ψ_i^0 aufgespannt wird, reduziert sich der Störoperator auf die (1×1) -Matrix $\langle\psi_i^0|H_1|\psi_i^0\rangle$ und entspricht der in (4.2) gegebenen Energiekorrektur ε_i^1 .

Für entartetes ε_i^0 ist der Fall etwas komplizierter: Nach der entarteten Störungstheorie ist nämlich nun die Basis in \mathcal{H}^0 nicht mehr frei, sondern muss so gewählt werden, dass die Störung H_1 darin diagonal wird. Dies entspricht gerade einer (möglichen) Aufhebung der Entartung durch die Störung. Für die ψ_j^0 mit $\varepsilon_j^0 = \varepsilon_i^0$ gilt dann $\langle\psi_j^0|H_1|\psi_i^0\rangle = \delta_{ij}\varepsilon_i^1$ mit der üblichen Definition $\varepsilon_i^1 := \langle\psi_i^0|H_1|\psi_i^0\rangle$. Damit gilt insbesondere $H_1\psi_i^0 = \varepsilon_i^1\psi_i^0$, womit schließlich (C.7) verschwindet.

Anders formuliert: Im entarteten Fall muss zunächst die „richtige“ Sternheimer-Gleichung gefunden werden, d.h. präziser: es muss eine Basis des entarteten Eigenraums so bestimmt werden, dass die damit aufgestellte Sternheimer-Gleichung lösbar ist. Die entsprechenden Basisfunktionen sind zugleich die Eigenfunktionen des Gesamtsystems $H_0 + H_1$ in \mathcal{H}^0 . Die Korrektur besteht in \mathcal{H}^0 also allein in der Justierung der Basis. Anschließend kann, wie in Abschnitt 4.1 beschrieben, der Sattelpunkt in $\mathcal{H}^- \oplus \mathcal{H}^+$ über geeignete Maximierungs-/Minimierungsverfahren berechnet werden.

C.3 Die Sternheimer-Gleichung mit nicht-hermiteschem H_0

Dieser Abschnitt lässt sich in gewissen Sinne als Ergänzung zu Ref. 78 verstehen, in der das variationelle EXX-Verfahren an kugelsymmetrischen Systemen (Be-Atom, Ne-Atom)

getestet worden ist. Beschreibt man solche Systeme in Kugelkoordinaten, muss die Schrödinger-Gleichung (bzw. hier die Kohn-Sham-Gleichung als effektive Ein-Teilchen-Schrödinger-Gleichung) nur noch in einer Dimension für den Radialteil der Wellenfunktionen gelöst werden. Üblicherweise verwendet man dabei eine logarithmische Zerlegung (Basis) der Radialachse, da die starken Oszillationen nahe des Ursprungs eine wesentlich stärkere Auflösung verlangen als die im Vergleich dazu „ruhigen“ Schwänze der Funktionen. Der Operator der kinetischen Energie (Laplace-Operator) und damit der gesamte Hamilton-Operator werden in dieser Darstellung jedoch zu *nicht-hermiteschen* Matrizen. Die in Abschnitt 4.1 beschriebene Lösung der Sternheimer-Gleichung mit Hilfe der Minimierung ihrer „Stammfunktion“ (4.3) kann daher nicht mehr ohne Weiteres verwendet werden, wie in Ref. 78 gezeigt. Mit dem nun vorgestellten Trick kann die Minimierungsmethode jedoch auch auf einen solchen Fall übertragen werden. Wie in Abschnitt 4.1.2 sei abkürzend $A = H_0 - \varepsilon_i^0$, $\mathbf{x} = \psi_i^1$ und $\mathbf{b} = (H_1 - \varepsilon_i^1)\psi_i^0$ gesetzt, und wie dort sei nur der diskrete Fall diskutiert.

Die Idee ist einfach: Ist A eine beliebige n -quadratische Matrix, ist das Produkt A^+A , wobei A^+ die Adjungierte von A sein soll, in jedem Fall eine hermitesche Matrix, denn aus $(AB)^+ = B^+A^+$ (mit A, B n -quadratisch) folgt sofort $(A^+A)^+ = A^+(A^+)^+ = A^+A$. Zudem ist A^+A eine positiv semi-definite Matrix, denn für alle $\mathbf{x} \in \mathbb{C}^n$ gilt $\langle \mathbf{x} | A^+A \mathbf{x} \rangle = \langle A \mathbf{x} | A \mathbf{x} \rangle = \|A \mathbf{x}\|^2 \geq 0$. Multipliziert man also die Gleichung

$$A \mathbf{x} + \mathbf{b} = \mathbf{0} \quad \text{mit } \mathbf{b} \in \mathbb{C}^n \quad \text{fest} \quad (\text{C.8})$$

mit von links A^+ , erfüllt die neue Gleichung die in (4.13) geforderten Voraussetzungen, und alle in Abschnitt 4.1.2 beschriebenen Methoden können auf sie angewandt werden:

$$A^+A \mathbf{x} + A^+ \mathbf{b} = \mathbf{0}. \quad (\text{C.9})$$

Bevor man diese Gleichung jedoch zur Lösung von (C.8) benutzen kann, muss klar sein, dass beide Gleichungen auch wirklich äquivalent sind. Dies lässt sich leicht beweisen, wenn (i) A^+ eine ein-eindeutige (bijektive) lineare Abbildung auf dem Wertebereich $\mathcal{W}(A)$ der Abbildung A ist und (ii) Gleichung (C.8) überhaupt eine Lösung besitzt. Dies soll hier kurz gezeigt werden:

Beweis der Bijektivität von A^+ auf $\mathcal{W}(A)$: Ist $\mathcal{N}(A)$ der Nullraum von A , hat $\mathcal{W}(A)$ die Dimension $p := n - \dim \mathcal{N}(A)$. Sei dann $\{\mathbf{w}_1, \dots, \mathbf{w}_p\}$ eine Basis von $\mathcal{W}(A)$. Nun lässt sich \mathbb{C}^n in den Wertebereich von A und dessen Orthogonalraum² zerlegen: $\mathbb{C}^n = \mathcal{W}(A) \oplus \mathcal{W}(A)^\perp$. Mit einem Satz der linearen Algebra³ gilt $\mathcal{N}(A^+) = \mathcal{W}(A)^\perp$, womit hier $\mathbb{C}^n = \mathcal{W}(A) \oplus \mathcal{N}(A^+)$ folgt. Dann gilt für jedes $\mathbf{x} \in \mathbb{C}^n$: es gibt $\mathbf{x}_1 \in \mathcal{W}(A)$, $\mathbf{x}_2 \in \mathcal{N}(A^+)$ so, dass $\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2$ und folglich $A^+\mathbf{x} = A^+\mathbf{x}_1 + A^+\mathbf{x}_2 = A^+\mathbf{x}_1 + \mathbf{0}$. Der Wertebereich von A^+ ist daher durch $\mathcal{W}(A^+) = \text{span}\{A^+\mathbf{x} | \mathbf{x} \in \mathcal{W}(A)\}$ gegeben. Daraus folgt, dass $\{A^+\mathbf{w}_1, \dots, A^+\mathbf{w}_p\}$ eine Basis von $\mathcal{W}(A^+)$ ist, denn als adjungierte Matrizen besitzen A und A^+ den gleichen Rang, d.h., die Dimensionen ihrer Wertebereiche sind gleich.

Damit folgt die Injektivität: Seien $\mathbf{x}, \mathbf{x}' \in \mathcal{W}(A)$. Es ist zu zeigen: Aus $A^+\mathbf{x} = A^+\mathbf{x}'$ folgt $\mathbf{x} = \mathbf{x}'$. Sei also $A^+\mathbf{x} = A^+\mathbf{x}'$. Dann hat man $A^+(\mathbf{x} - \mathbf{x}') = \mathbf{0}$ mit $\mathbf{x} - \mathbf{x}' \in \mathcal{W}(A)$.

²Der Orthogonalraum einer Menge $U \subset \mathbb{C}^n$ ist definiert als die Menge aller Elemente, die senkrecht auf U stehen: $U^\perp := \{\mathbf{x} \in \mathbb{C}^n | \langle \mathbf{x} | \mathbf{u} \rangle = 0 \text{ für } \mathbf{u} \in U\}$. Die direkte Summe zweier linearer Teilräume U und V mit $U \cap V = \{0\}$ ist definiert als $U \oplus V := \{\mathbf{x} \in \mathbb{C}^n | \text{es gibt } \mathbf{u} \in U, \mathbf{v} \in V \text{ mit } \mathbf{x} = \mathbf{u} + \mathbf{v}\}$. Für einen linearen Teilraum $U \subset \mathbb{C}^n$ gilt $\mathbb{C}^n = U \oplus U^\perp$. Vgl. dazu z.B. Ref. 123, Band 1, Kapitel 8.

³Vgl. z.B. Ref. 123, Band 1, Satz 10.14.

Mit einer Basis $\{\mathbf{w}_1, \dots, \mathbf{w}_p\}$ von $\mathcal{W}(A)$ existieren also $\alpha_1, \dots, \alpha_p \in \mathbb{C}$ so, dass $\mathbf{x} - \mathbf{x}' = \sum_{i=1}^p \alpha_i \mathbf{w}_i$. Da aber $A^+(\mathbf{x} - \mathbf{x}') = \sum_{i=1}^p \alpha_i A^+ \mathbf{w}_i = \mathbf{0}$ und $\{A^+ \mathbf{w}_1, \dots, A^+ \mathbf{w}_p\}$ gemäß dem Vorangehenden Basis von $\mathcal{W}(A^+)$, folgt $\alpha_1 = \dots = \alpha_p = 0$. Damit ist $\mathbf{x} - \mathbf{x}' = \mathbf{0}$, d.h., $\mathbf{x} = \mathbf{x}'$, womit die Injektivität gezeigt ist.

Nun zur Surjektivität: Sei $\mathbf{y} \in \mathcal{W}(A^+)$. Zu zeigen: Es gibt $\mathbf{x} \in \mathcal{W}(A)$ mit $\mathbf{y} = A^+ \mathbf{x}$. Mit der Basis von $\mathcal{W}(A^+)$ wie oben existieren $\beta_1, \dots, \beta_p \in \mathbb{C}$ so, dass $\mathbf{y} = \sum_{i=1}^p \beta_i A^+ \mathbf{w}_i = A^+ \sum_{i=1}^p \beta_i \mathbf{w}_i$. Definiert man $\mathbf{x} := \sum_{i=1}^p \beta_i \mathbf{w}_i$, folgt unmittelbar $\mathbf{y} = A^+ \mathbf{x}$ mit $\mathbf{x} \in \mathcal{W}(A)$, womit auch die Surjektivität gezeigt ist. Damit gelingt nun der

Beweis der Äquivalenz der Gleichungen (C.8) und (C.9): Es ist klar, dass alle Lösungen $\mathbf{x} \in \mathbb{C}^n$ von (C.8) auch Lösungen von (C.9) sind. Bleibt zu zeigen, dass auch umgekehrt alle Lösungen von (C.9) Lösungen von (C.8) sind, wenn (C.8) überhaupt eine Lösung $\mathbf{x}_0 \in \mathbb{C}^n$ besitzt (vgl. den vorangehenden Abschnitt C.2). Unter dieser Annahme folgt, dass \mathbf{b} Element des Wertebereichs $\mathcal{W}(A)$ der Abbildung A ist, da $\mathbf{b} = -A \mathbf{x}_0 \in \mathcal{W}(A)$. Es ist also $A \mathbf{x} + \mathbf{b} \in \mathcal{W}(A)$ für alle $\mathbf{x} \in \mathbb{C}^n$, d.h., der Wertebereich der linken Seite von (C.8) ist ebenfalls $\mathcal{W}(A)$. Nun ist aber — wie oben gezeigt — A^+ auf $\mathcal{W}(A)$ eine bijektive lineare Abbildung. Aus $A^+ A \mathbf{x} + A^+ \mathbf{b} = A^+(A \mathbf{x} + \mathbf{b}) = \mathbf{0}$ folgt daher unmittelbar Gleichung (C.8), was zu zeigen war.

Interessanterweise lässt sich dieses Ergebnis auch auf eine ganz andere Art ableiten. Bildet man nämlich das Betragsquadrat auf beiden Seiten von (C.8), erhält man

$$\begin{aligned} \|A \mathbf{x} + \mathbf{b}\|^2 &= \langle \mathbf{x} A + \mathbf{b} | A \mathbf{x} + \mathbf{b} \rangle & (C.10) \\ &= \langle \mathbf{x} A | A \mathbf{x} \rangle + \langle \mathbf{x} A | \mathbf{b} \rangle + \langle \mathbf{b} | A \mathbf{x} \rangle + \langle \mathbf{b} | \mathbf{b} \rangle \\ &= \langle \mathbf{x} | A^+ A \mathbf{x} \rangle + \langle \mathbf{x} | A^+ \mathbf{b} \rangle + \langle \mathbf{b} | A \mathbf{x} \rangle + \|\mathbf{b}\|^2 = 0 . \end{aligned}$$

Das Betragsquadrat ist *genau* dann 0, wenn sein Argument $\mathbf{0}$ ist, womit sofort die Äquivalenz dieser Gleichung zu (C.8) folgt. Zudem ist 0 auch das (eindeutige) Minimum des Betragsquadrats. Ferner ist das Betragsquadrat — im Gegensatz zur Betragsfunktion (!) — überall stetig differenzierbar, besitzt also einen Gradienten, der genau bei $\mathbf{0}$ verschwindet. Dies liefert unmittelbar Gleichung (C.9):

$$\nabla \|A \mathbf{x} + \mathbf{b}\|^2 = 2 \frac{d}{d\mathbf{x}^*} \|A \mathbf{x} + \mathbf{b}\|^2 = 2 A^+ A \mathbf{x} + 2 A^+ \mathbf{b} = 0 . \quad (C.11)$$

Dabei wurde (C.10) nur nach dem konjugiert-komplexen Vektor \mathbf{x}^* abgeleitet (also dem Bra-Vektor), dafür aber noch mit 2 multipliziert.⁴ Mit diesem Weg findet man also zusätzlich zu (C.9) auch noch „automatisch“ das passende Funktional, mit dem die Minimierung durchgeführt werden kann, nämlich die linke Seite von (C.10). Man beachte aber, dass das Ableiten im Allgemeinen *keine* Äquivalenzrelation darstellt; so besitzt (C.11) in jedem Fall eine Lösung, auch wenn (C.8) nicht lösbar ist. Notwendig und hinreichend für die Äquivalenz ist, dass (C.8) lösbar ist, dass also $\mathbf{b} \in \mathcal{W}(A)$, was auch in dem zuerst geführten Beweis vorausgesetzt worden ist.

⁴Zur Definition des Gradienten im Komplexen vgl. Ref. 126.

Anhang D

Tensor Darstellung der EXX-Methode

Die linke Seite der Gleichung (2.11) in Abschnitt 2.3 beschreibt mit Hilfe der linearen Antwort-Funktion χ_0 , wie sich die Dichte in linearer Ordnung ändert, wenn auf das effektive Potential der Kohn-Sham-Gleichung (1.52) eine lokale Änderung $\Delta v : \mathbb{R}^3 \rightarrow \mathbb{R}$ addiert wird. Zur Handhabung einer nicht-lokalen Potentialdifferenz, wie sie im variationellen EXX-Verfahren gemäß (2.1) auftritt, wird jedoch ein „mächtigerer“ linearer Antwort-Operator X_0 benötigt, da Δv nun durch eine $(\mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{C})$ -Funktion repräsentiert wird, nämlich den Kern des Integraloperators H_1 von Gleichung (2.1).

Dieses verallgemeinerte X_0 lässt sich sehr leicht finden, wenn man noch einmal dem Gedankengang von Abschnitt 2.1 folgt, diesmal allerdings für eine beliebige lineare Störung H_1 , deren Anwendung auf eine Wellenfunktion ψ sich durch

$$(H_1\psi)(\mathbf{r}) := \int \Delta v(\mathbf{r}, \mathbf{r}') \psi(\mathbf{r}') d^3 r' \quad (\text{D.1})$$

mit einem geeigneten Δv beschreiben lässt. Die Herleitung der induzierten Ladungsdichte, d.h., der linearen Dichteänderung $\alpha \Delta \rho$ mit $\Delta \rho$ gemäß (2.4) geschah dort völlig unabhängig von der speziellen Gestalt der Störung, deshalb bleibt sie auch für das allgemeiner definierte H_1 aus (D.1) gültig. Auch sind die Korrekturfunktionen ψ_i^1 wieder über die Sternheimer-Gleichung (2.2) oder mit Hilfe der Greenschen Funktion wie in (2.7) gegeben. Dies, eingesetzt in (2.4), liefert in Abhängigkeit von Δv die lineare Dichteänderung

$$\begin{aligned} \Delta \rho(\mathbf{r}) &= \sum_i^{\text{occ}} \int d^3 r' \left[\psi_i^0(\mathbf{r})^* G_i(\mathbf{r}, \mathbf{r}') (H_1 \psi_i^0)(\mathbf{r}') + \text{c.c.} \right] \\ &= \sum_i^{\text{occ}} \int d^3 r' \left[\psi_i^0(\mathbf{r})^* \sum_{\substack{j \\ \varepsilon_j \neq \varepsilon_i}} \frac{\psi_j^0(\mathbf{r}) \psi_j^0(\mathbf{r}')^*}{\varepsilon_i - \varepsilon_j} \int \Delta v(\mathbf{r}', \mathbf{r}'') \psi_i^0(\mathbf{r}'') d^3 r'' + \text{c.c.} \right] \\ &= \iint T(\mathbf{r}, \mathbf{r}', \mathbf{r}'') \Delta v(\mathbf{r}', \mathbf{r}'') d^3 r' d^3 r'' + \text{c.c.} , \end{aligned} \quad (\text{D.2})$$

wobei

$$T(\mathbf{r}, \mathbf{r}', \mathbf{r}'') := \sum_i^{\text{occ}} \sum_{\substack{j \\ \varepsilon_j \neq \varepsilon_i}} \frac{\psi_i^0(\mathbf{r})^* \psi_j^0(\mathbf{r}) \psi_j^0(\mathbf{r}')^* \psi_i^0(\mathbf{r}'')}{\varepsilon_i - \varepsilon_j} . \quad (\text{D.3})$$

Ausdruck (D.2) läßt sich noch vereinfachen, wenn man eine hermitesche Störung voraussetzt, also $\Delta v(\mathbf{r}', \mathbf{r}'')^* = \Delta v(\mathbf{r}'', \mathbf{r}')$ annimmt:

$$\Delta\rho(\mathbf{r}) = \iint T(\mathbf{r}, \mathbf{r}', \mathbf{r}'') \Delta v(\mathbf{r}', \mathbf{r}'') d^3r' d^3r'' + \iint T(\mathbf{r}, \mathbf{r}', \mathbf{r}'')^* \Delta v(\mathbf{r}'', \mathbf{r}') d^3r' d^3r'' .$$

Umbenennung der Variablen im zweiten Summanden liefert dann

$$\begin{aligned} \Delta\rho(\mathbf{r}) &= \iint \left[T(\mathbf{r}, \mathbf{r}', \mathbf{r}'') + T(\mathbf{r}, \mathbf{r}'', \mathbf{r}')^* \right] \Delta v(\mathbf{r}', \mathbf{r}'') d^3r' d^3r'' & (D.4) \\ &= \iint X_0(\mathbf{r}, \mathbf{r}', \mathbf{r}'') \Delta v(\mathbf{r}', \mathbf{r}'') d^3r' d^3r'' \end{aligned}$$

mit dem Kern

$$\begin{aligned} X_0(\mathbf{r}, \mathbf{r}', \mathbf{r}'') &:= T(\mathbf{r}, \mathbf{r}', \mathbf{r}'') + T(\mathbf{r}, \mathbf{r}'', \mathbf{r}')^* & (D.5) \\ &= \sum_i^{\text{occ.}} \sum_j^{\text{unocc.}} \frac{\psi_i^0(\mathbf{r})^* \psi_j^0(\mathbf{r}) \psi_j^0(\mathbf{r}')^* \psi_i^0(\mathbf{r}'') + \psi_i^0(\mathbf{r}) \psi_j^0(\mathbf{r})^* \psi_j^0(\mathbf{r}'') \psi_i^0(\mathbf{r}')^*}{\varepsilon_i - \varepsilon_j} . \end{aligned}$$

Hier wurden bei der Summierung über j nur noch die unbesetzten Zustände berücksichtigt, da alle anderen Summanden sich — analog wie bei der Herleitung der „normalen“ linearen Antwort (1.93) — gerade herauskürzen. Würde man X_0 näherungsweise in einem endlichen Basissatz darstellen, wäre dies ein Tensor dritter Stufe — ganz analog, wie kontinuierliche Operatoren bei „Diskretisierung“ in Matrizen übergehen.

Während der Nullraum von χ_0 allein durch die Menge der konstanten Funktionen gegeben ist [56], ist allerdings der von X_0 wesentlich größer: Für alle hermiteschen Störungen Δv mit $\langle \psi_i^0 | \Delta v | \psi_j^0 \rangle = 0$, also solche, die besetzte und unbesetzte Zustände nur jeweils *untereinander* koppeln¹, folgt aus (D.4) und (D.5) sofort $\Delta\rho = 0$. Konstante Potentialverschiebungen bilden natürlich einen Teilraum dieses Nullraums. Damit ist X_0 auch dann nicht invertierbar, wenn man den Definitionsbereich auf diejenigen Operatoren beschränkt, die ihrerseits wieder nur auf dem Teilraum nicht-konstanter Funktionen definiert sind. In endlicher Darstellung ist dies intuitiv klar: Für die Störung (Matrix) gibt es natürlich mehr Variationsmöglichkeiten als für die induzierte Dichte (Vektor).

Interessanterweise läßt sich X_0 zu einem noch allgemeineren Operator \tilde{X}_0 erweitern, der zu einer gegebenen, nicht-lokalen Potentialänderung nicht die lineare Änderung der Dichte, sondern die des Dichteoperators liefert. Dieser ist (für ein nicht spin-polarisiertes Kohn-Sham-System) in Ortsdarstellung durch

$$\rho(\mathbf{r}_1, \mathbf{r}_2) := \sum_i^{\text{occ.}} \psi_i^0(\mathbf{r}_1) \psi_i^0(\mathbf{r}_2)^* \quad (D.6)$$

gegeben und seine lineare Änderung infolge von Änderungen der Wellenfunktionen ψ_i^1 folglich durch

$$\Delta\rho(\mathbf{r}_1, \mathbf{r}_2) := \sum_i^{\text{occ.}} \psi_i^0(\mathbf{r}_1) \psi_i^1(\mathbf{r}_2)^* + \psi_i^1(\mathbf{r}_1) \psi_i^0(\mathbf{r}_2)^* . \quad (D.7)$$

¹Dieser Nullraum ist die tiefere Begründung dafür, dass bei der Lösung der Sternheimer-Gleichung die so überaus nützliche vereinfachte Version der Sternheimer-Gleichung (4.12) benutzt werden konnte.

Gleichsetzen von \mathbf{r}_1 und \mathbf{r}_2 liefert für (D.6) dann die gewöhnliche Dichte und für (D.7) die induzierte Dichte (2.4). Eine ganz analoge Überlegung wie oben in (D.2) bis (D.5) durchgeführt liefert in diesem Fall

$$\begin{aligned} \tilde{X}_0(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \mathbf{r}_4) &:= & (D.8) \\ & \sum_i^{\text{occ.}} \sum_j^{\text{unocc.}} \frac{\psi_i^0(\mathbf{r}_1) \psi_j^0(\mathbf{r}_2)^* \psi_j^0(\mathbf{r}_3) \psi_i^0(\mathbf{r}_4)^* + \psi_i^0(\mathbf{r}_2)^* \psi_j^0(\mathbf{r}_1) \psi_j^0(\mathbf{r}_4)^* \psi_i^0(\mathbf{r}_3)}{\varepsilon_i - \varepsilon_j} . \end{aligned}$$

EXX-Formalismus — kurzgefasst

Aus (D.4) folgt nun die EXX/OEP-Gleichung, indem gemäß (2.1) speziell

$$\Delta v(\mathbf{r}', \mathbf{r}'') := \langle \mathbf{r}' | H_1 | \mathbf{r}'' \rangle = V_{\text{Fock}}[\{\psi_i\}](\mathbf{r}', \mathbf{r}'') - \delta(\mathbf{r}' - \mathbf{r}'') v_x[\rho](\mathbf{r}') \quad (D.9)$$

gesetzt wird. Damit ergibt sich zunächst die induzierte Ladungsdichte

$$\begin{aligned} \Delta \rho(\mathbf{r}) &= \iint X_0(\mathbf{r}, \mathbf{r}', \mathbf{r}'') \left[V_{\text{Fock}}(\mathbf{r}', \mathbf{r}'') - \delta(\mathbf{r}' - \mathbf{r}'') v_x(\mathbf{r}') \right] d^3 r' d^3 r'' & (D.10) \\ &= \iint X_0(\mathbf{r}, \mathbf{r}', \mathbf{r}'') V_{\text{Fock}}(\mathbf{r}', \mathbf{r}'') d^3 r' d^3 r'' - \int \chi_0(\mathbf{r}, \mathbf{r}') v_x(\mathbf{r}') d^3 r' . \end{aligned}$$

Einsetzen von (D.5), Vereinfachen und Zu-Null-Setzen dieses Ausdrucks führt dann unmittelbar auf die OEP-Gleichung (1.97). Im Vergleich zu dieser oder zur EXX-Gleichung (1.90) erscheint dem Verfasser die Form (D.10) allerdings anschaulicher; insbesondere der lineare Charakter des EXX-Formalismus wird hier unmittelbar klar, denn der Zusammenhang zwischen Störung und Änderung der Ladungsdichte ist nun formal sehr einfach durch einen Integraloperator gegeben.

Als praktische „Merkformel“ sei der EXX-Formalismus nun auf eine noch knappere — aber dafür äußerst prägnante! — Gestalt gebracht. Definiert man dazu

$$\begin{aligned} \hat{X}_0 f &: \mathbb{R}^3 \mapsto \mathbb{R} ; & (D.11) \\ (\hat{X}_0 f)(\mathbf{r}) &:= \begin{cases} \iint X_0(\mathbf{r}, \mathbf{r}', \mathbf{r}'') f(\mathbf{r}', \mathbf{r}'') d^3 r' d^3 r'' & (f \text{ nicht-lokal, hermitesch}) \\ \int \chi_0(\mathbf{r}, \mathbf{r}') f(\mathbf{r}') d^3 r' & (f \text{ lokal, multiplikativ}) \end{cases} \end{aligned}$$

mit X_0 gemäß (D.5) und χ_0 nach (1.93), kann man (D.10) formal

$$\Delta \rho := \hat{X}_0(V_{\text{Fock}} - v_x) = \hat{X}_0 V_{\text{Fock}} - \hat{\chi}_0 v_x \quad (D.12)$$

schreiben. Damit aber lässt sich der gesamte EXX-Formalismus in folgender einfacher Gleichung zusammenfassen:

$$\hat{X}_0(V_{\text{Fock}} - v_x) \stackrel{!}{=} 0 . \quad (D.13)$$

Natürlich ist dabei zu beachten, dass alle Größen der linken Seite — ggf. über die Dichte — noch von den Wellenfunktionen ψ_i abhängen.

Zusammenhang zur Sham-Schlüter-Gleichung

Die Form von (D.10) mit der darin enthaltenen Doppelintegration legt einen Zusammenhang zur Sham-Schlüter-Gleichung [177] nahe. Diese beschreibt allgemein die Beziehung zwischen dem Austausch-Korrelationsanteil der nicht-lokalen, energieabhängigen Selbstenergie Σ_{xc} und dem lokalen Austausch-Korrelationspotential v_{xc} des Kohn-Sham-Systems und ist durch

$$\int d^3r \int d^3r' \int d\varepsilon G_{KS}(\mathbf{r}, \mathbf{r}', \varepsilon) G(\mathbf{r}'', \mathbf{r}, \varepsilon) \left[\Sigma_{xc}(\mathbf{r}', \mathbf{r}'', \varepsilon) - \delta(\mathbf{r}' - \mathbf{r}'') v_{xc}(\mathbf{r}') \right] = 0 \quad (\text{D.14})$$

gegeben. Hier ist G die voll wechselwirkende Greensfunktion [177] und

$$G_{KS}(\mathbf{r}, \mathbf{r}', \varepsilon) := \lim_{\eta \rightarrow 0^+} \sum_j \frac{\psi_j^0(\mathbf{r}) \psi_j^0(\mathbf{r}')^*}{\varepsilon - (\varepsilon_j + i\eta \operatorname{sgn}(E_F - \varepsilon_j))} \quad (\text{D.15})$$

die Greensfunktion des Kohn-Sham-Systems mit Fermi-Energie E_F , wobei der Limes so zu verstehen ist, dass er erst *nach* Anwendung von G_{KS} auf einen anderen Ausdruck ausgeführt wird. Es ist ferner bekannt [171, 177, 178], dass die Sham-Schlüter-Gleichung näherungsweise in die OEP-Gleichung (1.97) überführt werden kann, wenn man (i) die Selbstenergie nur durch ihren Austauschanteil berücksichtigt, der mit dem nicht-lokalen Austauschoperator (1.27) identisch ist, $\Sigma_x = V_{\text{Fock}}$, sowie entsprechend v_{xc} durch v_x ersetzt, und (ii) G durch G_{KS} genähert wird. Dann entspricht die eckige Klammer in (D.14) gerade der Störung Δv aus (D.9) und ist insbesondere unabhängig von ε .

Hier soll nun speziell gezeigt werden, dass sich unter diesen Bedingungen die beiden Greensfunktionen gerade zu dem oben definierten verallgemeinerten Lineare-Antwort-Operator (D.5) zusammenfassen lassen:

$$X_0(\mathbf{r}, \mathbf{r}', \mathbf{r}'') = \frac{1}{2\pi i} \int G_{KS}(\mathbf{r}, \mathbf{r}', \varepsilon) G_{KS}(\mathbf{r}'', \mathbf{r}, \varepsilon) d\varepsilon. \quad (\text{D.16})$$

Da diese Beziehung nicht unmittelbar ersichtlich ist (und auch die natürlich in ähnlicher Weise erfolgende Herleitung der OEP-Gleichung aus der Sham-Schlüter-Gleichung in der Literatur nur in sehr „gestraffter“ Form vorliegt), soll sie hier etwas detaillierter vorgeführt werden. Sei dafür zunächst

$$F(\varepsilon) := \sum_{j,k} \frac{\psi_j^0(\mathbf{r}) \psi_j^0(\mathbf{r}')^*}{\varepsilon - (\varepsilon_j + i\eta \operatorname{sgn}(E_F - \varepsilon_j))} \frac{\psi_k^0(\mathbf{r}'') \psi_k^0(\mathbf{r})^*}{\varepsilon - (\varepsilon_k + i\eta \operatorname{sgn}(E_F - \varepsilon_k))} \quad (\text{D.17})$$

definiert. Ziel ist es dann, F über die reelle Achse zu integrieren, was am einfachsten mit Hilfe der Funktionentheorie gelingt. Sei dazu die Doppelsumme in (D.17) so sortiert, dass alle Glieder mit $\varepsilon_j = \varepsilon_k$ und alle mit $\varepsilon_j \neq \varepsilon_k$ zusammengefasst sind. Für den nicht entarteten Fall ließe sich das auch zu $j = k$ bzw. $j \neq k$ vereinfachen. Dann hat man

$$F(\varepsilon) = \sum_{\substack{j,k \\ \varepsilon_j = \varepsilon_k}} \frac{\psi_j^0(\mathbf{r}) \psi_j^0(\mathbf{r}')^* \psi_k^0(\mathbf{r}'') \psi_k^0(\mathbf{r})^*}{[\varepsilon - (\varepsilon_j + i\eta \operatorname{sgn}(E_F - \varepsilon_j))]^2} \quad (\text{D.18})$$

$$+ \sum_{\substack{j,k \\ \varepsilon_j \neq \varepsilon_k}} \frac{\psi_j^0(\mathbf{r}) \psi_j^0(\mathbf{r}')^* \psi_k^0(\mathbf{r}'') \psi_k^0(\mathbf{r})^*}{[\varepsilon - (\varepsilon_j + i\eta \operatorname{sgn}(E_F - \varepsilon_j))][\varepsilon - (\varepsilon_k + i\eta \operatorname{sgn}(E_F - \varepsilon_k))]}.$$

Zur Integration dieses Ausdrucks lässt sich der Residuensatz anwenden:

$$\int_{-\infty}^{\infty} F(\varepsilon) d\varepsilon = 2\pi i \sum_i \text{Res}(F, p_i) , \quad (\text{D.19})$$

wobei die p_i die in der oberen Hälfte der komplexen Ebene \mathbb{C} liegenden Punkte bezeichnen, für die F nicht holomorph ist. Sie sind hier durch die Polstellen $\varepsilon_i + i\eta \text{sgn}(E_F - \varepsilon_i)$ gegeben, die zu einem *besetzten* Zustand i gehören. Um nun

$$\text{Res}_i := \text{Res} \left(F, \varepsilon_i + i\eta \text{sgn}(E_F - \varepsilon_i) \right) := \frac{1}{2\pi i} \oint_{|\varepsilon - \varepsilon_i| = \delta} F(\varepsilon) d\varepsilon \quad (\text{D.20})$$

zu finden, beachte man, dass die erste Summe in (D.18) nur Pole 2. Ordnung enthält, für die das Residuum verschwindet. Es muss also nur noch die zweite Summe betrachtet werden. Dort verschwinden wiederum alle Beiträge mit $\varepsilon_j \neq \varepsilon_i \wedge \varepsilon_k \neq \varepsilon_i$. Damit verbleibt

$$\begin{aligned} \text{Res}_i = \frac{1}{2\pi i} \oint_{|\varepsilon - \varepsilon_i| = \delta} & \left[\sum_{\substack{j \\ \varepsilon_j \neq \varepsilon_i}} \frac{\psi_j^0(\mathbf{r}) \psi_j^0(\mathbf{r}')^* \psi_i^0(\mathbf{r}'') \psi_i^0(\mathbf{r})^*}{[\varepsilon - (\varepsilon_j + i\eta \text{sgn}(E_F - \varepsilon_j))][\varepsilon - (\varepsilon_i + i\eta)]} \right. \\ & \left. + \sum_{\substack{k \\ \varepsilon_k \neq \varepsilon_i}} \frac{\psi_i^0(\mathbf{r}) \psi_i^0(\mathbf{r}')^* \psi_k^0(\mathbf{r}'') \psi_k^0(\mathbf{r})^*}{[\varepsilon - (\varepsilon_i + i\eta)][\varepsilon - (\varepsilon_k + i\eta \text{sgn}(E_F - \varepsilon_k))]} \right] d\varepsilon . \end{aligned} \quad (\text{D.21})$$

Dies lässt sich weiter vereinfachen mit der Cauchy-Formel

$$f(z) = \frac{1}{2\pi i} \oint_{\partial A} \frac{f(\zeta)}{\zeta - z} d\zeta \quad (z \in A) , \quad (\text{D.22})$$

wobei f eine geeignete holomorphe Funktion in einem zulässigen Gebiet $A \subset \mathbb{C}$ und ∂A der Rand dieses Gebiets ist.² Mit $\zeta = \varepsilon$ und $z = \varepsilon_i + i\eta$ wird (D.21) damit zu

$$\begin{aligned} \text{Res}_i = \sum_{\substack{j \\ \varepsilon_j \neq \varepsilon_i}} & \frac{\psi_j^0(\mathbf{r}) \psi_j^0(\mathbf{r}')^* \psi_i^0(\mathbf{r}'') \psi_i^0(\mathbf{r})^*}{(\varepsilon_i + i\eta) - (\varepsilon_j + i\eta \text{sgn}(E_F - \varepsilon_j))} \\ & + \sum_{\substack{k \\ \varepsilon_k \neq \varepsilon_i}} \frac{\psi_i^0(\mathbf{r}) \psi_i^0(\mathbf{r}')^* \psi_k^0(\mathbf{r}'') \psi_k^0(\mathbf{r})^*}{(\varepsilon_i + i\eta) - (\varepsilon_k + i\eta \text{sgn}(E_F - \varepsilon_k))} . \end{aligned} \quad (\text{D.23})$$

Benennt man in der zweiten Summe k in j um, lassen sich die Summen zusammenfassen. Anschließendes Einsetzen in den Residuensatz (D.19) und Vollziehen des Grenzübergangs $\eta \rightarrow 0^+$ liefert

$$\begin{aligned} \lim_{\eta \rightarrow 0^+} \frac{1}{2\pi i} \int F(\varepsilon) d\varepsilon & = \sum_i^{\text{occ.}} \lim_{\eta \rightarrow 0^+} \text{Res}_i = \\ & = \sum_i^{\text{occ.}} \sum_{\substack{j \\ \varepsilon_j \neq \varepsilon_i}} \left[\frac{\psi_j^0(\mathbf{r}) \psi_j^0(\mathbf{r}')^* \psi_i^0(\mathbf{r}'') \psi_i^0(\mathbf{r})^*}{\varepsilon_i - \varepsilon_j} + \frac{\psi_i^0(\mathbf{r}) \psi_i^0(\mathbf{r}')^* \psi_j^0(\mathbf{r}'') \psi_j^0(\mathbf{r})^*}{\varepsilon_i - \varepsilon_j} \right] . \end{aligned} \quad (\text{D.24})$$

Alle Glieder für besetzte Zustände j kürzen sich heraus (wie man durch Vertauschen von i und j erkennt), und ein Vergleich mit (D.5) liefert unmittelbar die Beziehung (D.16), was zu beweisen war.

²Die genauen Bedingungen für die Gültigkeit der Cauchyschen Integralformel sowie des Residuensatzes lassen sich nicht „kurz“ zusammenfassen. Man entnehme sie der Spezialliteratur, z.B. Ref. 123. Sie sind aber für den hier untersuchten Fall erfüllt.

Anhang E

Verallgemeinerung des EXX-Formalismus auf N -darstellbare Ladungsdichten

Es sei hier ein Ansatz vorgeschlagen, wie der nach Abschnitt 1.6.1 nur im Falle v -darstellbarer Ladungsdichten anwendbare EXX/OEP-Formalismus zumindest formal auf N -darstellbare Dichten verallgemeinert werden kann.

Bei der in Abschnitt 1.6.3 gezeigten Herleitung der EXX/OEP-Gleichung mag folgende Frage aufgetaucht sein: Obwohl nur für v -darstellbare Ladungsdichten definiert, wurde die Gleichung aus der für alle N -darstellbaren Dichten gültigen Göring-Levy-Störungstheorie (Abschnitt 1.5.4) entwickelt. An welcher Stelle ging nun diese Allgemeinheit verloren? Die Antwort ist folgende: In (1.101) wurde davon ausgegangen, dass die zur Ladungsdichte ρ gehörende Slater-Determinante Φ_0 aus den N energetisch niedrigsten Orbitalen der Kohn-Sham-Gleichung (1.52) aufgebaut ist. Dies gilt jedoch nur im Falle der v -Darstellbarkeit.

Wie Levy und Perdew [43] zeigen (vgl. auch Ref. 21), existiert jedoch auch für *beliebige* N -darstellbare Dichten ρ stets ein effektives Potential der Kohn-Sham-Gleichung, so dass eine einfache, aus N geeigneten Kohn-Sham-Orbitalen aufgebaute Slater-Determinante gerade auf ρ führt. Die benutzten Orbitale sind dann aber nicht mehr notwendigerweise die N niedrigsten; die Determinante ist also im Allgemeinen ein *angeregter* Zustand des Kohn-Sham-Systems. Diese Tatsache ermöglicht nun, die Austauschenergie wiederum wie in (1.69) formal über den Fock-Operator zu definieren; nur dass in der Summe jetzt nicht mehr von 1 bis N , sondern über die Ein-Teilchen-Zustände der Determinante, d.h., über die besetzten Orbitale, gezählt wird.

Mit dieser Definition lassen sich ebenfalls beide Strategien zur Berechnung von $v_x[\rho]$, die direkte (Abschnitt 1.6.1) und die störungstheoretische (Abschnitt 1.6.3), verfolgen: Sei die geeignete, auf ρ führende Determinante mit Φ_{k_ρ} bezeichnet. Dann muss in (1.101) nur Φ_0 durch das entsprechende Φ_{k_ρ} ersetzt werden, und umgekehrt wird in der Summe natürlich auch Φ_{k_ρ} durch Φ_0 ersetzt. Die resultierende Gleichung hat *genau dieselbe Gestalt* wie (1.103) oder (1.97), nur dass in der Summe über die N besetzten Orbitale jetzt die Orbitale von Φ_{k_ρ} gezählt werden, und die Summierung der unbesetzten Orbitale ebenfalls entsprechend umgestaltet werden muss. Auch die direkte Berechnung des exakten Austauschpotentials über (1.90) liefert dieses Resultat: in χ_0 von (1.93) werden besetzte und unbesetzte Orbitale analog umgruppiert.

Dieses χ_0 zeigt deutlich, was hier geschieht: es vermittelt zwischen einer kleinen Änderung des effektiven Potentials und der entsprechenden linearen Änderung einer Ladungs-

dichte, die zu einem *angeregten Zustand* gehört. Die Inverse — wiederum „lebt“ sie nur auf dem kleineren, um konstante Funktionen reduzierten Definitionsbereich — kann daher zwischen der *nicht- v -darstellbaren Dichte* ρ und dem Potential vermitteln, was der „normale“ lineare Antwortoperator, zum Grundzustand Φ_0 gehörend, nicht leisten konnte.

Auf zwei Schwierigkeiten im Zusammenhang mit dieser Definition sei allerdings hingewiesen: (i) Man muss von vorneherein wissen, aus welchen Kohn-Sham-Orbitalen sich eine gegebene Ladungsdichte ρ zusammensetzt. Dafür gibt es bislang allerdings kein (praktisch anwendbares) Kriterium, was den formalen Charakter obiger Definition ausmacht. (ii) Bei dieser Vorgehensweise wird die Suche in (1.65), die eigentlich über alle antisymmetrischen Viel-Teilchen-Zustände erfolgt, auf einfache Slater-Determinanten reduziert. Diese Definition des universellen Funktionals ist nicht mehr ganz so allgemein wie die ursprüngliche, aber ebenfalls anwendbar; sie wird z.B. in Ref. 21 und 179 diskutiert. Andererseits ist die Einschränkung auf Determinanten auch von Vorteil: sie garantiert eine Austauschenergie in einer Form analog zur Hartree-Fock-Theorie und folglich, dass bereits die Austauschwechselwirkung zur vollständigen Aufhebung der elektronischen Selbstwechselwirkung ausreichend ist.

Anhang F

Symmetriegenähertes EXX: Herleitung der variationellen Formulierung

Um ein Variationsprinzip für das symmetriegenäherte EXX-Verfahren (Abschnitt 4.6) herzuleiten, wird im Folgenden zunächst die OEP-Gleichung (1.96) in periodischen Randbedingungen formuliert (für ein spin-kompensiertes System) und anschließend auf eine Gestalt gebracht, die sich leicht als induzierte Ladungsdichte interpretieren lässt. Wie in Anhang A sei das Konjugiert-Komplexe zur Vermeidung zu vieler Klammern wieder durch einen Querstrich anstelle eines Sterns gekennzeichnet.

Analog zu (B.12) ist die auf die unbesetzten Zustände projizierte Greensche Funktion des Hamilton-Operators in (3.14) gegeben durch

$$G_{i\mathbf{k}}^\perp(\mathbf{r}', \mathbf{r}) := \sum_j^{\text{unocc.}} \frac{\psi_{j\mathbf{k}}^0(\mathbf{r}') \overline{\psi_{j\mathbf{k}}^0(\mathbf{r})}}{\varepsilon_{i\mathbf{k}}^0 - \varepsilon_{j\mathbf{k}}^0} \implies \overline{G_{i\mathbf{k}}^\perp(\mathbf{r}', \mathbf{r})} = G_{i\mathbf{k}}^\perp(\mathbf{r}, \mathbf{r}'), \quad (\text{F.1})$$

wobei noch die nachfolgend benutzte Symmetrieeigenschaft mit angegeben wurde. Damit kann der lineare Antwortoperator (4.69) in der Form

$$\begin{aligned} \chi_0(\mathbf{r}', \mathbf{r}) &= \sum_{\mathbf{k}} \sum_i^{\text{occ.}} \sum_j^{\text{unocc.}} 4 \frac{\overline{\psi_{i\mathbf{k}}^0(\mathbf{r}') \psi_{j\mathbf{k}}^0(\mathbf{r}') \psi_{j\mathbf{k}}^0(\mathbf{r}) \psi_{i\mathbf{k}}^0(\mathbf{r})}}{\varepsilon_{i\mathbf{k}}^0 - \varepsilon_{j\mathbf{k}}^0} \\ &= \sum_{\mathbf{k}} \sum_i^{\text{occ.}} 4 \overline{\psi_{i\mathbf{k}}^0(\mathbf{r}')} G_{i\mathbf{k}}^\perp(\mathbf{r}', \mathbf{r}) \psi_{i\mathbf{k}}^0(\mathbf{r}) \end{aligned} \quad (\text{F.2})$$

geschrieben werden. Einsetzen in die OEP-Integrodifferentialgleichung (1.96) liefert dann

$$\begin{aligned} \sum_{\mathbf{k}} \sum_i^{\text{occ.}} \int \left\{ v_x(\mathbf{r}') \cdot 4 \overline{\psi_{i\mathbf{k}}^0(\mathbf{r}') G_{i\mathbf{k}}^\perp(\mathbf{r}', \mathbf{r}) \psi_{i\mathbf{k}}^0(\mathbf{r})} \right\} d^3r' &= \\ = \sum_{\mathbf{k}} \sum_i^{\text{occ.}} \int \left\{ 2 \frac{\delta E_x[\rho]}{\delta \psi_{i\mathbf{k}}^0(\mathbf{r}')} \overline{G_{i\mathbf{k}}^\perp(\mathbf{r}', \mathbf{r}) \psi_{i\mathbf{k}}^0(\mathbf{r})} + 2 \frac{\delta E_x[\rho]}{\delta \psi_{i\mathbf{k}}^0(\mathbf{r}')} G_{i\mathbf{k}}^\perp(\mathbf{r}', \mathbf{r}) \psi_{i\mathbf{k}}^0(\mathbf{r}) \right\} d^3r', \end{aligned} \quad (\text{F.3})$$

wobei der Faktor 2 auf der rechten Seite aus der Summe über den Spin resultiert. Um dies auf eine Gestalt zu bringen, die sich in gewohnter Weise störungstheoretisch interpretieren

lässt (d.h., die Greensche Funktion soll von *links* auf eine Störung wirken), werden beide Seiten der Gleichung komplex konjugiert:

$$\begin{aligned}
& \sum_{\mathbf{k}} \sum_i^{\text{occ.}} \int \left\{ \overline{\psi_{i\mathbf{k}}^0(\mathbf{r})} G_{i\mathbf{k}}^\perp(\mathbf{r}, \mathbf{r}') \cdot 2v_x(\mathbf{r}') \psi_{i\mathbf{k}}^0(\mathbf{r}') \right\} d^3r' = \\
& = \sum_{\mathbf{k}} \sum_i^{\text{occ.}} \int \left\{ \psi_{i\mathbf{k}}^0(\mathbf{r}) \overline{G_{i\mathbf{k}}^\perp(\mathbf{r}, \mathbf{r}') \frac{\delta E_x[\rho]}{\delta \psi_{i\mathbf{k}}^0(\mathbf{r}')}} + \overline{\psi_{i\mathbf{k}}^0(\mathbf{r})} G_{i\mathbf{k}}^\perp(\mathbf{r}, \mathbf{r}') \frac{\delta E_x[\rho]}{\delta \psi_{i\mathbf{k}}^0(\mathbf{r}')} \right\} d^3r' \\
& = \sum_{\mathbf{k}} \sum_i^{\text{occ.}} \int \left\{ \psi_{i\mathbf{k}}^0(\mathbf{r}) \overline{G_{i\mathbf{k}}^\perp(\mathbf{r}, \mathbf{r}') V_{\text{Fock}} \psi_{i\mathbf{k}}^0(\mathbf{r}')} + \overline{\psi_{i\mathbf{k}}^0(\mathbf{r})} G_{i\mathbf{k}}^\perp(\mathbf{r}, \mathbf{r}') V_{\text{Fock}} \psi_{i\mathbf{k}}^0(\mathbf{r}') \right\} d^3r'.
\end{aligned} \tag{F.4}$$

Dabei wurde die Symmetrierelation (F.1) der Greenschen Funktion verwendet und in der zweiten Umformung noch (1.22) sowie (1.27) benutzt. Umstellen liefert schließlich

$$\begin{aligned}
& \sum_{\mathbf{k}} \sum_i^{\text{occ.}} \left\{ \overline{\psi_{i\mathbf{k}}^0(\mathbf{r})} \underbrace{\int G_{i\mathbf{k}}^\perp(\mathbf{r}, \mathbf{r}') (V_{\text{Fock}} - 2\hat{v}_x) \psi_{i\mathbf{k}}^0(\mathbf{r}') d^3r'}_{=: \psi_{i\mathbf{k}}^1(\mathbf{r})} + \right. \\
& \quad \left. + \psi_{i\mathbf{k}}^0(\mathbf{r}) \underbrace{\int \overline{G_{i\mathbf{k}}^\perp(\mathbf{r}, \mathbf{r}') V_{\text{Fock}} \psi_{i\mathbf{k}}^0(\mathbf{r}')} d^3r'}_{=: \widetilde{\psi_{i\mathbf{k}}^1(\mathbf{r})}} \right\} = 0.
\end{aligned} \tag{F.5}$$

Durch die Wirkung der Greenschen Funktion auf die jeweils rechts von ihr stehenden Ausdrücke können die Integrale als lineare Korrekturfunktionen $\psi_{i\mathbf{k}}^1$ bzw. $\widetilde{\psi_{i\mathbf{k}}^1}$ aufgefasst werden, mit denen sich die linke Seite der Gleichung in ähnlicher Form wie die induzierte Ladungsdichte (2.4) zusammenfassen lässt:

$$\widetilde{\Delta\rho}(\mathbf{r}) := \sum_{\mathbf{k}} \sum_i^{\text{occ.}} \left(\overline{\psi_{i\mathbf{k}}^0(\mathbf{r})} \psi_{i\mathbf{k}}^1(\mathbf{r}) + \psi_{i\mathbf{k}}^0(\mathbf{r}) \overline{\widetilde{\psi_{i\mathbf{k}}^1(\mathbf{r})}} \right) = 0. \tag{F.6}$$

Die Korrekturfunktionen $\psi_{i\mathbf{k}}^1$ und $\widetilde{\psi_{i\mathbf{k}}^1}$ können dann natürlich wieder über die Sternheimer-Gleichung berechnet werden. Analog zu (4.12) ergibt sich:

$$(H_0 - \varepsilon_i^0) P_c |\psi_i^1\rangle + P_c H_1 |\psi_i^0\rangle = 0 \quad \text{mit} \quad H_1 := V_{\text{Fock}} - 2\hat{v}_x, \tag{F.7}$$

$$(H_0 - \varepsilon_i^0) P_c |\widetilde{\psi_i^1}\rangle + P_c \widetilde{H}_1 |\psi_i^0\rangle = 0 \quad \text{mit} \quad \widetilde{H}_1 := V_{\text{Fock}}. \tag{F.8}$$

Dies ist das in Abschnitt 4.6.3 verwendete Ergebnis.

Mit der Zeitinversionssymmetrie (3.7), angewandt etwa auf die linke Seite von (F.4), erkennt man leicht, dass für reellwertige Potentiale v_x die Definition der induzierten Ladungsdichte (F.6) mit der ursprünglichen Definition (2.4) identisch ist. Für komplexwertige v_x , wie sie beim symmetriegenäherten EXX als Zwischenergebnisse auftreten (vgl. Abschnitte 4.6.2 und 4.6.3), weichen die Definitionen hingegen voneinander ab.

Anhang G

Die OEP/EXX-Gleichung aus dem Vielteilchenansatz von Görling und Levy

In Abschnitt 1.6.3 wurde, Görling und Levy [42] folgend, skizziert, wie sich die OEP/EXX-Gleichung in eleganter Weise aus einem störungstheoretischen Ansatz im Hilbertraum der N -Teilchen-Funktionen herleiten lässt. Dieser Ansatz führt — ganz analog zur variationalen Formulierung von Kapitel 2 — auf die Forderung, die durch die Störung induzierte Ladungsdichte in erster Ordnung verschwinden zu lassen (Gleichung (1.102)), wobei die Beschreibung allerdings komplett im Viel-Teilchen-Hilbertraum erfolgt. Die Äquivalenz zur bekannten Formulierung der OEP/EXX-Gleichung in Ein-Teilchen-Funktionen, d.h., die Umformung von Gleichung (1.102) zu Gleichung (1.103), wurde allerdings in Ref. 42 nicht näher erläutert. Obwohl es sich dabei um eine rein algebraische Aufgabe handelt, ist diese Umformulierung jedoch keinesfalls trivial. Um ein echtes Nachvollziehen zu ermöglichen, sei sie daher an dieser Stelle schrittweise vorgeführt.

Setzt man dazu (1.101) in die linke Seite von Gleichung (1.102) ein, so muss zunächst der Ausdruck $\langle \Phi_0 | \hat{\rho}(\mathbf{r}) | \Phi_j \rangle$ für alle $j \geq 1$ ausgewertet werden. Dabei sind sowohl der Grundzustand Φ_0 als auch die angeregten Zustände Φ_j ($j \geq 1$) durch Slater-Determinanten gegeben, die aus den Ein-Teilchen-Zuständen ϕ_i der Kohn-Sham-Gleichung (1.52) aufgebaut sind:

$$\begin{aligned}\Phi_0(\mathbf{x}_1, \dots, \mathbf{x}_N) &= \frac{1}{\sqrt{N!}} \sum_{P \in \mathcal{S}} (-1)^P \phi_{p_1}(\mathbf{x}_1) \cdots \phi_{p_N}(\mathbf{x}_N), \\ \Phi_j(\mathbf{x}_1, \dots, \mathbf{x}_N) &= \frac{1}{\sqrt{N!}} \sum_{Q \in \mathcal{S}} (-1)^Q \phi_{\nu_{q_1}}(\mathbf{x}_1) \cdots \phi_{\nu_{q_N}}(\mathbf{x}_N).\end{aligned}$$

Es seien ϕ_1, \dots, ϕ_N die besetzten Kohn-Sham-Orbitale, und $\phi_{\nu_1}, \dots, \phi_{\nu_N}$ seien paarweis verschiedene besetzte oder unbesetzte Kohn-Sham-Orbitale, von denen aber mindestens eines unbesetzt sein soll.¹ Die ganzzahligen Indizes stehen jeweils für einen geeigneten Satz von Quantenzahlen und werden durch die Elemente der Permutationsgruppe \mathcal{S} untereinander vertauscht.² Für jedes $P \in \mathcal{S}$ sei dabei $P(1, \dots, N) = (p_1, \dots, p_N)$. p gibt an, ob

¹Genau genommen müsste ν_i^j statt ν_i geschrieben werden, denn jedes Φ_j ist natürlich aus anderen Kohn-Sham-Orbitalen aufgebaut. Der Übersichtlichkeit halber wird der Index j hier jedoch unterdrückt.

²Anstatt der Quantenzahlen könnte man auch die Koordinatenindizes permutieren. Bei der folgenden Auswertung von $\langle \Phi_j | V_{\text{ee}} | \Phi_0 \rangle$ erweist es sich jedoch als einfacher, die Quantenzahlen zu vertauschen.

P eine gerade oder ungerade Permutation ist. Die Permutationen wurden beim angeregten Zustand mit Q bezeichnet, um im Nachfolgenden Verwirrungen zu vermeiden. Ferner sollen die Ein-Teilchen-Zustände wie in (1.13) aus Spin- und Ortsanteil zusammengesetzt sein, also $\phi_k(\mathbf{x}_i) := \psi_k(\mathbf{r}_i)\chi_{\sigma_k}(s_i)$ mit $\mathbf{x}_i := (\mathbf{r}_i, s_i)$ und der Spinquantenzahl σ_k des k -ten Zustandes. Abkürzend sei noch die Notation

$$\int dx_i \cdots := \sum_{s_i=-\frac{1}{2}}^{+\frac{1}{2}} \int d^3 r_i \cdots \quad (\text{G.1})$$

für die Integration über den Ort und die Summe über die beiden Spinkanäle eingeführt. Damit ergibt sich:

$$\begin{aligned} \langle \Phi_0 | \hat{\rho}(\mathbf{r}) | \Phi_j \rangle &= \langle \Phi_0 | \sum_{m=1}^N \delta(\mathbf{r} - \mathbf{r}_m) | \Phi_j \rangle \quad (\text{G.2}) \\ &= \frac{1}{N!} \sum_{P, Q \in \mathcal{S}} (-1)^{p+q} \sum_{m=1}^N \int dx_1 \cdots dx_N \cdot \\ &\quad \cdot \phi_{p_1}^*(\mathbf{x}_1) \cdots \phi_{p_N}^*(\mathbf{x}_N) \delta(\mathbf{r} - \mathbf{r}_m) \phi_{\nu_{q_1}}(\mathbf{x}_1) \cdots \phi_{\nu_{q_N}}(\mathbf{x}_N) \\ &= \frac{1}{N!} \sum_{P, Q \in \mathcal{S}} (-1)^{p+q} \cdot \\ &\quad \cdot \sum_{m=1}^N \langle \phi_{p_1} | \phi_{\nu_{q_1}} \rangle \cdots \langle \phi_{p_m} | \delta(\mathbf{r} - \mathbf{r}_m) | \phi_{\nu_{q_m}} \rangle \cdots \langle \phi_{p_N} | \phi_{\nu_{q_N}} \rangle . \end{aligned}$$

Ist Φ_j ein mehrfach angeregter Zustand, enthält also mehrere Ein-Teilchen-Zustände ϕ_{ν_i} mit $\nu_i > N$, werden in allen Summanden orthogonale Zustände direkt miteinander gekoppelt, und der gesamte Ausdruck verschwindet. Sei also in Φ_j nur der α -te Ein-Teilchen-Zustand angeregt, $\phi_\alpha \rightarrow \phi_{\nu_\alpha}$, und für die anderen Ein-Teilchen-Zustände gelte o.B.d.A. $\phi_{\nu_i} = \phi_i$. Für das Matrixelement $\langle \phi_{p_i} | \phi_{\nu_{q_i}} \rangle$ gilt dann: Es verschwindet, wenn $q_i = \alpha$, denn in jedem Fall ist $\langle \phi_{p_i} | \phi_{\nu_\alpha} \rangle = 0$. Für $q_i \neq \alpha$ erhält man $\langle \phi_{p_i} | \phi_{\nu_{q_i}} \rangle = \langle \phi_{p_i} | \phi_{q_i} \rangle = \delta_{p_i q_i}$. Definiert man ein „umgekehrtes Kronecker-Symbol“ durch

$$\tilde{\delta}_{\alpha\beta} := 1 - \delta_{\alpha\beta} = \begin{cases} 0, & \alpha = \beta \\ 1, & \alpha \neq \beta \end{cases} , \quad (\text{G.3})$$

lässt sich dies mit $\langle \phi_{p_i} | \phi_{\nu_{q_i}} \rangle = \delta_{p_i q_i} \tilde{\delta}_{\alpha, q_i}$ zusammenfassen, und (G.2) wird zu

$$\begin{aligned} \langle \Phi_0 | \hat{\rho}(\mathbf{r}) | \Phi_j \rangle &= \frac{1}{N!} \sum_{P, Q \in \mathcal{S}} (-1)^{p+q} \sum_{m=1}^N \delta_{p_1, q_1} \tilde{\delta}_{\alpha, q_1} \cdots \delta_{p_{m-1}, q_{m-1}} \tilde{\delta}_{\alpha, q_{m-1}} \cdot \\ &\quad \cdot \langle \phi_{p_m} | \delta(\mathbf{r} - \mathbf{r}_m) | \phi_{\nu_{q_m}} \rangle \delta_{p_{m+1}, q_{m+1}} \tilde{\delta}_{\alpha, q_{m+1}} \cdots \delta_{p_N, q_N} \tilde{\delta}_{\alpha, q_N} . \end{aligned}$$

Auf Grund der δ_{p_i, q_i} verschwinden alle Summanden mit $P \neq Q$, es gilt also

$$\begin{aligned} \langle \Phi_0 | \hat{\rho}(\mathbf{r}) | \Phi_j \rangle &= \\ &= \frac{1}{N!} \sum_{Q \in \mathcal{S}} (-1)^{2q} \sum_{m=1}^N \tilde{\delta}_{\alpha, q_1} \cdots \tilde{\delta}_{\alpha, q_{m-1}} \langle \phi_{q_m} | \delta(\mathbf{r} - \mathbf{r}') | \phi_{\nu_{q_m}} \rangle \tilde{\delta}_{\alpha, q_{m+1}} \cdots \tilde{\delta}_{\alpha, q_N} , \end{aligned}$$

wobei noch benutzt wurde, dass \mathbf{r}_m eine reine Integrationsvariable bei der Berechnung des Skalarprodukts ist, die in \mathbf{r}' umbenannt werden kann.³ Bei der Summation über m von 1 bis N durchlaufen auch die q_i jeweils alle Werte von 1 bis N . Daher kann man die Ersetzung $q_i \rightarrow i$ durchführen, und man hat

$$\langle \Phi_0 | \hat{\rho}(\mathbf{r}) | \Phi_j \rangle = \frac{1}{N!} \sum_{Q \in \mathcal{S}} \sum_{m=1}^N \tilde{\delta}_{\alpha,1} \cdots \tilde{\delta}_{\alpha,m-1} \langle \phi_m | \delta(\mathbf{r} - \mathbf{r}') | \phi_{\nu_m} \rangle \tilde{\delta}_{\alpha,m+1} \cdots \tilde{\delta}_{\alpha,N} .$$

Da die Summanden nun von q_i unabhängig sind, liefert die Summation über Q genau $N!$ mal dasselbe, womit sich der Normierungsfaktor wegekürzt. Auf Grund der $\tilde{\delta}_{\alpha,i}$ kann des Weiteren in der Summe über m nur der Summand $m = \alpha$ von 0 verschieden sein. Damit ergibt sich schließlich

$$\begin{aligned} \langle \Phi_0 | \hat{\rho}(\mathbf{r}) | \Phi_j \rangle &= \langle \phi_\alpha | \delta(\mathbf{r} - \mathbf{r}') | \phi_{\nu_\alpha} \rangle \\ &= \int dx' \psi_\alpha^*(\mathbf{r}') \chi_{\sigma_\alpha}^*(s') \delta(\mathbf{r} - \mathbf{r}') \psi_{\nu_\alpha}(\mathbf{r}') \chi_{\sigma_{\nu_\alpha}}(s') \\ &= \psi_\alpha^*(\mathbf{r}) \psi_{\nu_\alpha}(\mathbf{r}) \delta_{\sigma_\alpha \sigma_{\nu_\alpha}} . \end{aligned}$$

Das Ergebnis ist 0, wenn die Spinquantenzahlen σ_α und σ_{ν_α} des besetzten bzw. unbesetzten Zustandes nicht übereinstimmen. Insgesamt erhält man also

$$\langle \Phi_0 | \hat{\rho}(\mathbf{r}) | \Phi_j \rangle = \begin{cases} \psi_\alpha^*(\mathbf{r}) \psi_{\nu_\alpha}(\mathbf{r}) \delta_{\sigma_\alpha \sigma_{\nu_\alpha}} & \text{für } \Phi_j \text{ einfach angeregt mit } \phi_\alpha \rightarrow \phi_{\nu_\alpha} \\ 0 & \text{für } \Phi_j \text{ mehrfach angeregt .} \end{cases} \quad (\text{G.4})$$

Es sei bemerkt, dass das Matricelement hier eine \mathbf{r} -abhängige Funktion und kein Skalar ist, weil der Dichteoperator $\hat{\rho}$ in (G.2) außer von den Teilchenkoordinaten \mathbf{r}_m noch von der Koordinate \mathbf{r} abhängt, über die im Skalarprodukt nicht integriert wird.

Multipliziert man nun (1.101) von links mit $\langle \Psi_0 | \hat{\rho}(\mathbf{r}) |$, wobei gemäß Abschnitt 1.6.3 $\Psi_0 = \Phi_0$, liefern folglich nur die Summanden mit einfach angeregten Slater-Determinanten einen Beitrag. Zur Berechnung ihrer entsprechenden Koeffizienten werden nun noch die Matricelemente des lokalen Potentials $v_{\text{lok}}[\rho] := v_{\text{H}}[\rho] + v_{\text{x}}[\rho]$ und des Operators der gesamten elektronischen Wechselwirkung V_{ee} benötigt.

Sei also Φ_j eine einfach angeregte Determinante, in der wieder wie oben ϕ_α durch ϕ_{ν_α} ersetzt worden ist. Für das lokale Potential erhält man dann mit einer zur Auswertung von (G.2) analogen Überlegung:

$$\langle \Phi_j | v_{\text{lok}}[\rho] | \Phi_0 \rangle = \langle \phi_{\nu_\alpha} | v_{\text{lok}}[\rho] | \phi_\alpha \rangle \delta_{\sigma_\alpha \sigma_{\nu_\alpha}} . \quad (\text{G.5})$$

Für den Operator der elektronischen Wechselwirkung ergibt sich:

$$\begin{aligned} \langle \Phi_j | V_{\text{ee}} | \Phi_0 \rangle &= \frac{1}{N!} \sum_{P,Q \in \mathcal{S}} (-1)^{p+q} \frac{1}{2} \sum_{\substack{m,n=1 \\ m \neq n}}^N \int dx_1 \cdots dx_N \cdot \\ &\cdot \phi_{\nu_{q_1}}^*(\mathbf{x}_1) \cdots \phi_{\nu_{q_m}}^*(\mathbf{x}_m) \cdots \phi_{\nu_{q_n}}^*(\mathbf{x}_n) \cdots \phi_{\nu_{q_N}}^*(\mathbf{x}_N) \cdot \\ &\cdot \frac{1}{|\mathbf{r}_m - \mathbf{r}_n|} \phi_{p_1}(\mathbf{x}_1) \cdots \phi_{p_m}(\mathbf{x}_m) \cdots \phi_{p_n}(\mathbf{x}_n) \cdots \phi_{p_N}(\mathbf{x}_N) . \end{aligned} \quad (\text{G.6})$$

³Formal streng korrekt wäre es, das Skalarprodukt durch $\langle \phi_{q_m} | \delta(\mathbf{r} - \cdot) | \phi_{\nu_{q_m}} \rangle$ auszudrücken.

Da bis auf ϕ_{ν_α} alle ϕ_{ν_i} besetzte Zustände bezeichnen, sei also o.B.d.A. wieder $\phi_{\nu_i} = \phi_i$ für $i \neq \alpha$ gesetzt. Damit können bei der Summation über Q höchstens die Permutation $Q = P$, also insbesondere $q_m = p_m$ und $q_n = p_n$, und die Permutation der Vertauschung dieser beiden Zustände, also $q_1 = p_1, \dots, q_m = p_n, \dots, q_n = p_m, \dots, q_N = p_N$, einen Beitrag liefern. Ganz analog zur Berechnung der Wechselwirkungsenergie einer Slater-Determinante ergibt sich

$$\begin{aligned} \langle \Phi_j | V_{\text{ee}} | \Phi_0 \rangle &= \frac{1}{N!} \sum_{P \in \mathcal{S}} \frac{1}{2} \sum_{\substack{m,n=1 \\ m \neq n}}^N \int dx_1 \cdots dx_N \cdot \\ &\cdot \left\{ (-1)^{2p} \left[\phi_{\nu_{p_1}}^*(\mathbf{x}_1) \cdots \phi_{\nu_{p_m}}^*(\mathbf{x}_m) \cdots \phi_{\nu_{p_n}}^*(\mathbf{x}_n) \cdots \phi_{\nu_{p_N}}^*(\mathbf{x}_N) \cdot \right. \right. \\ &\quad \left. \left. \cdot \frac{1}{|\mathbf{r}_m - \mathbf{r}_n|} \phi_{p_1}(\mathbf{x}_1) \cdots \phi_{p_m}(\mathbf{x}_m) \cdots \phi_{p_n}(\mathbf{x}_n) \cdots \phi_{p_N}(\mathbf{x}_N) \right] + \right. \\ &+ (-1)^{2p+1} \left[\phi_{\nu_{p_1}}^*(\mathbf{x}_1) \cdots \phi_{\nu_{p_n}}^*(\mathbf{x}_n) \cdots \phi_{\nu_{p_m}}^*(\mathbf{x}_m) \cdots \phi_{\nu_{p_N}}^*(\mathbf{x}_N) \cdot \right. \\ &\quad \left. \left. \cdot \frac{1}{|\mathbf{r}_m - \mathbf{r}_n|} \phi_{p_1}(\mathbf{x}_1) \cdots \phi_{p_m}(\mathbf{x}_m) \cdots \phi_{p_n}(\mathbf{x}_n) \cdots \phi_{p_N}(\mathbf{x}_N) \right] \right\} . \end{aligned}$$

Wenn $\phi_{\nu_i} = \phi_i$ für *alle* $i \in \{1, \dots, N\}$ gelten würde, entspräche die erste eckige Klammer exakt dem Ausdruck für die Hartree-Energie und die zweite wäre mit der Austausch-Energie identisch, wenn Summen und Integrale jeweils hineingezogen würden. Nun ist aber ϕ_{ν_α} orthogonal zu jedem besetzten Zustand. Summanden, in denen ϕ_{ν_α} direkt an ϕ_α gekoppelt wird, verschwinden daher beim Ausführen der Integration:

$$\begin{aligned} \langle \Phi_j | V_{\text{ee}} | \Phi_0 \rangle &= \\ &= \frac{1}{N!} \sum_{P \in \mathcal{S}} \frac{1}{2} \sum_{\substack{m,n=1 \\ m \neq n}}^N \left\{ (+1) \left[\tilde{\delta}_{\alpha,p_1} \tilde{\delta}_{\alpha,p_2} \cdots \tilde{\delta}_{\alpha,p_{n-1}} \tilde{\delta}_{\alpha,p_{n+1}} \cdots \tilde{\delta}_{\alpha,p_{m-1}} \tilde{\delta}_{\alpha,p_{m+1}} \cdots \tilde{\delta}_{\alpha,p_N} \cdot \right. \right. \\ &\quad \left. \left. \cdot \int \frac{\phi_{\nu_{p_m}}^*(\mathbf{x}_m) \phi_{\nu_{p_n}}^*(\mathbf{x}_n) \phi_{p_m}(\mathbf{x}_m) \phi_{p_n}(\mathbf{x}_n)}{|\mathbf{r}_m - \mathbf{r}_n|} dx_m dx_n \right] + \right. \\ &+ (-1) \left[\tilde{\delta}_{\alpha,p_1} \tilde{\delta}_{\alpha,p_2} \cdots \tilde{\delta}_{\alpha,p_{n-1}} \tilde{\delta}_{\alpha,p_{n+1}} \cdots \tilde{\delta}_{\alpha,p_{m-1}} \tilde{\delta}_{\alpha,p_{m+1}} \cdots \tilde{\delta}_{\alpha,p_N} \cdot \right. \\ &\quad \left. \left. \cdot \int \frac{\phi_{\nu_{p_n}}^*(\mathbf{x}_m) \phi_{\nu_{p_m}}^*(\mathbf{x}_n) \phi_{p_m}(\mathbf{x}_m) \phi_{p_n}(\mathbf{x}_n)}{|\mathbf{r}_m - \mathbf{r}_n|} dx_m dx_n \right] \right\} , \end{aligned}$$

wobei $\tilde{\delta}_{\alpha\beta}$ wieder wie in (G.3) definiert sei. Die Variablen \mathbf{x}_m und \mathbf{x}_n werden nur zur Integration benötigt und können daher auch in \mathbf{x} bzw. \mathbf{x}' umbenannt werden. Ferner durchlaufen bei der Summation über m und n jeweils von 1 bis N auch die Variablen p_m bzw. p_n je alle Werte von 1 bis N , und es ist $p_m \neq p_n$ genau dann, wenn $m \neq n$. Daher kann man ersetzen: $p_m \rightarrow m$ und $p_n \rightarrow n$. Insgesamt erhält man so:

$$\begin{aligned}
\langle \Phi_j | V_{ee} | \Phi_0 \rangle &= \\
&= \frac{1}{N!} \sum_{P \in \mathcal{S}} \frac{1}{2} \sum_{\substack{m,n=1 \\ (m \neq n)}}^N \left\{ \tilde{\delta}_{\alpha,1} \tilde{\delta}_{\alpha,2} \cdots \tilde{\delta}_{\alpha,n-1} \tilde{\delta}_{\alpha,n+1} \cdots \tilde{\delta}_{\alpha,m-1} \tilde{\delta}_{\alpha,m+1} \cdots \tilde{\delta}_{\alpha,N} \cdot \right. \\
&\quad \left. \left[\int \frac{\phi_{\nu_m}^*(\mathbf{x}) \phi_{\nu_n}^*(\mathbf{x}') \phi_m(\mathbf{x}) \phi_n(\mathbf{x}')}{|\mathbf{r} - \mathbf{r}'|} dx dx' - \int \frac{\phi_{\nu_n}^*(\mathbf{x}) \phi_{\nu_m}^*(\mathbf{x}') \phi_m(\mathbf{x}) \phi_n(\mathbf{x}')}{|\mathbf{r} - \mathbf{r}'|} dx dx' \right] \right\}.
\end{aligned}$$

Der Fall $m = n$ muss hier nicht mehr explizit ausgeschlossen werden, da sich die entsprechenden Ausdrücke gegenseitig wegheben. Ferner ist die geschweifte Klammer von der Permutation P unabhängig, die Summe über P liefert also $N!$ mal dasselbe, und der Normierungsfaktor kürzt sich heraus. Auf Grund der $\tilde{\delta}_{\alpha,i}$ tragen weiter zur Summe über m und n nur Glieder bei, für die $m = \alpha$ oder $n = \alpha$ gilt. Es ist daher

$$\begin{aligned}
\langle \Phi_j | V_{ee} | \Phi_0 \rangle &= \\
&= \frac{1}{2} \sum_{\substack{n=1 \\ (m=\alpha)}}^N \left[\int \frac{\phi_{\nu_\alpha}^*(\mathbf{x}) \phi_{\nu_n}^*(\mathbf{x}') \phi_\alpha(\mathbf{x}) \phi_n(\mathbf{x}')}{|\mathbf{r} - \mathbf{r}'|} dx dx' - \int \frac{\phi_{\nu_n}^*(\mathbf{x}) \phi_{\nu_\alpha}^*(\mathbf{x}') \phi_\alpha(\mathbf{x}) \phi_n(\mathbf{x}')}{|\mathbf{r} - \mathbf{r}'|} dx dx' \right] \\
&+ \frac{1}{2} \sum_{\substack{m=1 \\ (n=\alpha)}}^N \left[\int \frac{\phi_{\nu_m}^*(\mathbf{x}) \phi_{\nu_\alpha}^*(\mathbf{x}') \phi_m(\mathbf{x}) \phi_\alpha(\mathbf{x}')}{|\mathbf{r} - \mathbf{r}'|} dx dx' - \int \frac{\phi_{\nu_\alpha}^*(\mathbf{x}) \phi_{\nu_m}^*(\mathbf{x}') \phi_m(\mathbf{x}) \phi_\alpha(\mathbf{x}')}{|\mathbf{r} - \mathbf{r}'|} dx dx' \right].
\end{aligned}$$

Beide Klammern sind identisch, wenn man in der oberen Zeile die Laufvariable n in m umbenennt und die Namen der Integrationsvariablen \mathbf{x} und \mathbf{x}' vertauscht. So ergibt sich

$$\begin{aligned}
\langle \Phi_j | V_{ee} | \Phi_0 \rangle &= \\
&= \sum_{m=1}^N \left[\int \frac{\phi_{\nu_\alpha}^*(\mathbf{x}) \phi_{\nu_m}^*(\mathbf{x}') \phi_m(\mathbf{x}') \phi_\alpha(\mathbf{x})}{|\mathbf{r} - \mathbf{r}'|} dx dx' - \int \frac{\phi_{\nu_\alpha}^*(\mathbf{x}) \phi_{\nu_m}^*(\mathbf{x}') \phi_m(\mathbf{x}) \phi_\alpha(\mathbf{x}')}{|\mathbf{r} - \mathbf{r}'|} dx dx' \right].
\end{aligned}$$

Nun war nach Voraussetzung $\phi_{\nu_m} = \phi_m$ für $m \neq \alpha$. Der Summand mit $m = \alpha$ jedoch ist 0, da die beiden Integrale identisch sind. Er bleibt auch 0, wenn in der Formel $\phi_{\nu_m}^*(\mathbf{x}')$ für alle $m \in \{1, \dots, N\}$ durch $\phi_m^*(\mathbf{x}')$ ersetzt wird. Damit folgt

$$\begin{aligned}
\langle \Phi_j | V_{ee} | \Phi_0 \rangle &= \\
&= \sum_{m=1}^N \left[\int \frac{\phi_{\nu_\alpha}^*(\mathbf{x}) \phi_m^*(\mathbf{x}') \phi_m(\mathbf{x}') \phi_\alpha(\mathbf{x})}{|\mathbf{r} - \mathbf{r}'|} dx dx' - \int \frac{\phi_{\nu_\alpha}^*(\mathbf{x}) \phi_m^*(\mathbf{x}') \phi_m(\mathbf{x}) \phi_\alpha(\mathbf{x}')}{|\mathbf{r} - \mathbf{r}'|} dx dx' \right] \\
&= \int \phi_{\nu_\alpha}^*(\mathbf{x}) \left(\int \frac{\sum_{m=1}^N \phi_m^*(\mathbf{x}') \phi_m(\mathbf{x}')}{|\mathbf{r} - \mathbf{r}'|} dx' \right) \phi_\alpha(\mathbf{x}) dx - \\
&\quad - \int \phi_{\nu_\alpha}^*(\mathbf{x}) \left(\int \sum_{m=1}^N \frac{\phi_m^*(\mathbf{x}') \phi_m(\mathbf{x})}{|\mathbf{r} - \mathbf{r}'|} \phi_\alpha(\mathbf{x}') dx' \right) dx \\
&= \delta_{\sigma_\alpha \sigma_{\nu_\alpha}} \int \psi_{\nu_\alpha}^*(\mathbf{r}) \left(\int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3 r' \right) \psi_\alpha(\mathbf{r}) d^3 r + \\
&\quad + \delta_{\sigma_\alpha \sigma_{\nu_\alpha}} \int \psi_{\nu_\alpha}^*(\mathbf{r}) \left(- \int \sum_{m=1}^N \delta_{\sigma_m \sigma_\alpha} \frac{\psi_m^*(\mathbf{r}') \psi_m(\mathbf{r})}{|\mathbf{r} - \mathbf{r}'|} \psi_\alpha(\mathbf{r}') d^3 r' \right) d^3 r.
\end{aligned}$$

Im letzten Schritt wurde benutzt, dass die Integration gemäß (G.1) die Summation über die Spinkanäle miteinschließt. Wird diese Summation ausgeführt, verschwinden alle Terme, in denen Wellenfunktionen mit unterschiedlichen Spinquantenzahlen gekoppelt werden. Die Ortsintegration erfolgt dann nur noch über den Ortanteil ψ_i der Wellenfunktion. Ferner wurde die Definition der Ladungsdichte (1.53) eingesetzt. In den Klammern stehen damit gerade die Definitionen des Hartree-Potentials (1.25) und des Austausch-Potentials (1.27). Es ist also schließlich

$$\begin{aligned} \langle \Phi_j | V_{ee} | \Phi_0 \rangle &= \tag{G.7} \\ &= \delta_{\sigma_\alpha \sigma_{\nu_\alpha}} \int \psi_{\nu_\alpha}^*(\mathbf{r}) v_H[\rho](\mathbf{r}) \psi_\alpha(\mathbf{r}) d^3r + \delta_{\sigma_\alpha \sigma_{\nu_\alpha}} \int \psi_{\nu_\alpha}^*(\mathbf{r}) \left(V_{\text{Fock}}[\{\psi_i\}] \psi_\alpha \right)(\mathbf{r}) d^3r \\ &= \delta_{\sigma_\alpha \sigma_{\nu_\alpha}} \langle \psi_{\nu_\alpha} | v_H[\rho] | \psi_\alpha \rangle + \delta_{\sigma_\alpha \sigma_{\nu_\alpha}} \langle \psi_{\nu_\alpha} | V_{\text{Fock}}[\{\psi_i\}] | \psi_\alpha \rangle . \end{aligned}$$

Bevor Ausdruck (1.102) endgültig ausgewertet werden kann, wird noch die Differenz $E_0 - E_j$ der Energien der Slater-Determinanten benötigt. Dies ist jedoch einfach: Da Φ_0 und Φ_j Determinanten eines Kohn-Sham-Systems sind, sind die zugehörigen Energien durch die Summe der Energien der Ein-Teilchen-Zustände gegeben, aus denen sie jeweils aufgebaut sind, vgl. die Fußnote auf Seite 20. Unterscheiden sich Φ_0 und Φ_j nur darin, dass der besetzte Zustand ϕ_i durch den unbesetzten Zustand ϕ_j ersetzt worden ist, folgt unmittelbar $E_0 - E_j = \varepsilon_i - \varepsilon_j$.

Damit gelingt nun die Umformulierung der Viel-Teilchen- in die Ein-Teilchen-Darstellung. Setzt man (1.101) unter Beachtung von $\Psi_0 = \Phi_0$ in (1.102) ein und benutzt (G.4), (G.5) und (G.7), hat man

$$\begin{aligned} \langle \Psi_0 | \hat{\rho}(\mathbf{r}) | \Psi_1 \rangle &= \tag{G.8} \\ &= \sum_{j=1}^{\infty} \langle \Phi_0 | \hat{\rho}(\mathbf{r}) | \Phi_j \rangle \frac{\langle \Phi_j | V_{ee} - v_H[\rho] - v_x[\rho] | \Phi_0 \rangle}{E_0 - E_j} \\ &= \sum_{\alpha=1}^N \sum_{\nu_\alpha=N+1}^{\infty} \delta_{\sigma_\alpha \sigma_{\nu_\alpha}} \psi_\alpha^*(\mathbf{r}) \psi_{\nu_\alpha}(\mathbf{r}) \frac{\langle \psi_{\nu_\alpha} | v_H[\rho] + V_{\text{Fock}}[\{\psi_k\}] - v_H[\rho] - v_x[\rho] | \psi_\alpha \rangle}{\varepsilon_\alpha - \varepsilon_{\nu_\alpha}} . \end{aligned}$$

Die Summe über die besetzten Zustände kommt ins Spiel, da *jeder* besetzte Zustand ψ_α ($\alpha \in \{1, \dots, N\}$) durch jeden unbesetzten Zustand ψ_{ν_α} ($\nu_\alpha \in \{N+1, \dots\}$) ersetzt werden muss. Bei einem nicht-spinpolarisierten System ist stets $\delta_{\sigma_\alpha \sigma_{\nu_\alpha}} = 1$, dafür wären die Summanden aber noch mit geeigneten Besetzungszahlen zu multiplizieren. Die Umbenennung $\alpha \rightarrow i$, $\nu_\alpha \rightarrow j$ liefert schließlich in der Tat die linke Seite der OEP/EXX-Gleichung (1.103), was zu zeigen war.

Mit (2.8) erkennt man, dass es sich bei der in Kapitel 2 definierten induzierten Ladungsdichte (2.4) und der von Levy und Göring hergeleiteten Dichteänderung (1.102) um völlig identische Ausdrücke handelt:

$$\Delta\rho := \sum_i^{\text{occ.}} (\psi_i^0)^* \psi_i^1 + \text{c.c.} = \langle \Psi_0 | \hat{\rho}(\mathbf{r}) | \Psi_1 \rangle + \text{c.c.} , \tag{G.9}$$

wobei die ψ_i^1 die linearen Korrekturen der *Ein*-Teilchen-Wellenfunktionen ψ_i^0 sind und Ψ_1 die lineare Korrektur des *Viel*-Teilchen-Grundzustandes Ψ_0 ist.

Anhang H

Verallgemeinerte Wannier-Funktionen

H.1 Translation um einen Gittervektor

Verallgemeinerte Wannier-Funktionen zu festem n können durch Translation ineinander überführt werden. Setzt man die Definition der Blochfunktionen (3.4) nämlich in (8.6) ein, folgt

$$\begin{aligned}\langle \mathbf{r} | n \mathbf{R} \rangle &= \frac{\Omega}{(2\pi)^3} \int_{\text{BZ}} e^{-i\mathbf{R}\cdot\mathbf{k}} \sum_{m=1}^N U_{mn}^{\mathbf{k}} \psi_{m\mathbf{k}}(\mathbf{r}) \\ &= \frac{\Omega}{(2\pi)^3} \int_{\text{BZ}} e^{-i\mathbf{R}\cdot\mathbf{k}} \sum_{m=1}^N U_{mn}^{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{r}} u_{m\mathbf{k}}(\mathbf{r}) \\ &= \frac{\Omega}{(2\pi)^3} \int_{\text{BZ}} e^{-i\mathbf{0}\cdot\mathbf{k}} \sum_{m=1}^N U_{mn}^{\mathbf{k}} e^{i\mathbf{k}\cdot(\mathbf{r}-\mathbf{R})} u_{m\mathbf{k}}(\mathbf{r}-\mathbf{R}) \\ &= \langle \mathbf{r}-\mathbf{R} | n \mathbf{0} \rangle ,\end{aligned}\tag{H.1}$$

d.h., $|n\mathbf{R}\rangle$ ist in Ortsdarstellung die um den Gittervektor \mathbf{R} verschobene Funktion $|n\mathbf{0}\rangle$.

H.2 Aktualisierung der Überlapp-Matrix

Zentrale Bedeutung bei der Implementierung des in Abschnitt 8.1 vorgestellten Verfahrens zur Erzeugung von Wannier-Funktionen von Marzari und Vanderbilt [116, 117] kommt den Überlapp-Matrizen (8.15) zu. Aus ihnen ergibt sich über die Momente der „Wannier-Verteilung“ (8.13) und (8.14) sowohl die Lokalisierung (8.9) der Wannier-Funktionen als auch der Gradient der Lokalisierung (8.19). Andererseits stellt die Berechnung der $N_{\mathbf{k}} \cdot N^2$ Matrixelemente ($N_{\mathbf{k}}$: Anzahl der \mathbf{k} -Punkte, N : Anzahl der Bänder) aus den Blochfunktionen gemäß Definition (8.15) aber auch den teuersten Schritt in dem Verfahren dar.

Die Effizienz der Methode beruht letztlich darauf, dass — außer bei der Initialisierung — die Überlapp-Matrizen mittels einer numerisch trivialen Matrizenmultiplikation auch aus den unitären Transformationen erhalten werden können, wie Gleichung (8.21) zeigt. Da die Elemente der Überlapp-Matrizen zudem die einzige Stelle im Iterationsprozess sind, an der die (durch vorangehende Schritte im Allgemeinen bereits transformierten)

Blochfunktionen explizit eingehen, entfällt damit auch die Notwendigkeit, die Durchmischung der Bänder gemäß (8.5) explizit ausführen zu müssen.

Die Beziehung (8.21) lässt sich leicht herleiten: Multipliziert man (8.5) mit $e^{-i\mathbf{k}\cdot\mathbf{r}}$, folgt

$$|u_{n\mathbf{k}}\rangle \longrightarrow \sum_{m=1}^N U_{mn}^{\mathbf{k}} |u_{m\mathbf{k}}\rangle .$$

Für die Transformation der Überlapp-Matrixelemente ergibt sich damit

$$\begin{aligned} M_{mn}^{\mathbf{k},\Delta\mathbf{k}} &:= \langle u_{m\mathbf{k}} | u_{n,\mathbf{k}+\Delta\mathbf{k}} \rangle \longrightarrow \left(\sum_{\mu=1}^N (U_{\mu m}^{\mathbf{k}})^* \langle u_{\mu\mathbf{k}} | \right) \left(\sum_{\nu=1}^N |u_{\nu,\mathbf{k}+\Delta\mathbf{k}}\rangle U_{\nu n}^{\mathbf{k}+\Delta\mathbf{k}} \right) \\ &= \sum_{\mu,\nu=1}^N (U_{\mu m}^{\mathbf{k}})^* \langle u_{\mu\mathbf{k}} | u_{\nu,\mathbf{k}+\Delta\mathbf{k}} \rangle U_{\nu n}^{\mathbf{k}+\Delta\mathbf{k}} \\ &= \left((U^{\mathbf{k}})^+ M^{\mathbf{k},\Delta\mathbf{k}} U^{\mathbf{k}+\Delta\mathbf{k}} \right)_{mn} , \end{aligned}$$

was zu beweisen war.

H.3 Der Hamilton-Operator in der Darstellung Verallgemeinerter Wannier-Funktionen

Die Definition der Verallgemeinerten Wannier-Funktionen (8.6) liefert für die Matrixelemente des Hamilton-Operators

$$\begin{aligned} \langle n\mathbf{R} | H | n'\mathbf{R}' \rangle &= \tag{H.2} \\ &= \frac{\Omega^2}{(2\pi)^6} \left(\sum_{m=1}^N \int_{\text{BZ}} d^3k \langle \psi_{m\mathbf{k}} | e^{+i\mathbf{k}\cdot\mathbf{R}} (U_{mn}^{\mathbf{k}})^* \right) H \left(\sum_{m'=1}^N \int_{\text{BZ}} d^3k' U_{m'n'}^{\mathbf{k}'} e^{-i\mathbf{k}'\cdot\mathbf{R}'} | \psi_{m'\mathbf{k}'} \rangle \right) \\ &= \frac{\Omega^2}{(2\pi)^6} \sum_{m,m'} \int d^3k d^3k' e^{i(\mathbf{k}\mathbf{R}-\mathbf{k}'\mathbf{R}')} (U_{mn}^{\mathbf{k}})^* U_{m'n'}^{\mathbf{k}'} \underbrace{\langle \psi_{m\mathbf{k}} | H | \psi_{m'\mathbf{k}'} \rangle}_{= \varepsilon_{m\mathbf{k}} \delta_{mm'} \delta(\mathbf{k}-\mathbf{k}')} \\ &= \frac{\Omega^2}{(2\pi)^6} \sum_m \int d^3k e^{i\mathbf{k}\cdot(\mathbf{R}-\mathbf{R}')} (U_{mn}^{\mathbf{k}})^* U_{mn'}^{\mathbf{k}} \varepsilon_{m\mathbf{k}} . \end{aligned}$$

Hierbei wurde angenommen, dass die Blochfunktionen auf das *periodische Grundgebiet* normiert sind (vgl. Abschnitt 3.1). Sind sie hingegen auf die Einheitszelle normiert, muss die rechte Seite von (H.2) noch mit der Anzahl $N_{\mathbf{R}}$ der darin enthaltenen Einheitszellen multipliziert werden, da im Kontext mit Wannier-Funktionen das Skalarprodukt $\langle \dots | \dots \rangle$ im Ortsraum sinnvollerweise als Integral über das periodische Grundgebiet¹ definiert ist.

¹Hierin liegt eine potentielle Fehlerquelle bei der Implementierung in einen Ebene-Wellen-Code: Dort wird meist nur der periodische Anteil der Blochfunktionen betrachtet, der dann auf die Einheitszelle normiert ist, vgl. Abschnitt 3.3.

Spezialfall klassischer Wannier-Funktionen

Für den Spezialfall klassischer Wannier-Funktionen, wenn also die Durchmischung der Bänder (8.5) abgeschaltet wird, lässt sich obige Formel (H.2) vereinfachen. Setzt man dazu alle $U^{\mathbf{k}} := \mathbf{1}$, erhält man

$$\begin{aligned} \langle n\mathbf{R}|H|n'\mathbf{R}' \rangle &= \frac{\Omega^2}{(2\pi)^6} \sum_m \int d^3k e^{i\mathbf{k}\cdot(\mathbf{R}-\mathbf{R}')} \delta_{mn} \delta_{mn'} \varepsilon_{m\mathbf{k}} \\ &= \delta_{nn'} \frac{\Omega^2}{(2\pi)^6} \int d^3k e^{i\mathbf{k}\cdot(\mathbf{R}-\mathbf{R}')} \varepsilon_{n\mathbf{k}} . \end{aligned} \quad (\text{H.3})$$

Matrixelemente zwischen klassischen Wannier-Funktionen verschiedener Bänder (hier sind n und n' ja wirklich Bandindizes) verschwinden also. Dies ist klar, weil die Teilhilberträume der zu den (isolierten) Bändern gehörenden Blochfunktionen orthogonal zueinander stehen. Ferner sind die nicht-verschwindenden Elemente direkt über eine Fouriertransformation des Bandes $\varepsilon_{n\mathbf{k}}$ ($\mathbf{k} \in \text{BZ}$) gegeben.

H.4 Berechnung der Ein-Teilchen-Energien aus der Wannier-Darstellung

In (H.2) gehen in der rechten Seite die aus einer vorangegangenen Diagonalisierung von H in ebenen Wellen bereits bekannten Eigenwerte $\varepsilon_{m\mathbf{k}}$ ein. Umgekehrt erhält man die Ein-Teilchen-Energien aus der Wannier-Darstellung von H über Gleichung (8.26) [122]. Sie soll hier bewiesen werden.

Dazu seien zunächst die Projektionen der Bloch- auf die Wannier-Funktionen mit Hilfe der Wannier-Funktionen der „Heimatzelle“ bei $\mathbf{R} = \mathbf{0}$ ausgedrückt:

$$\begin{aligned} \langle n\mathbf{R}|\psi_{m\mathbf{k}} \rangle &= \int_{\text{PG}} d^3r \langle n\mathbf{R}|\mathbf{r} \rangle \langle \mathbf{r}|\psi_{m\mathbf{k}} \rangle \\ &\stackrel{\mathbf{r} \rightarrow \mathbf{r}+\mathbf{R}}{=} \int_{\text{PG}} d^3r \langle n\mathbf{R}|\mathbf{r} + \mathbf{R} \rangle \langle \mathbf{r} + \mathbf{R}|\psi_{m\mathbf{k}} \rangle \\ &\stackrel{(\text{H.1})}{=} \int_{\text{PG}} d^3r \langle n\mathbf{0}|\mathbf{r} \rangle \psi_{m\mathbf{k}}(\mathbf{r} + \mathbf{R}) \\ &\stackrel{\text{Bloch}}{=} \int_{\text{PG}} d^3r \langle n\mathbf{0}|\mathbf{r} \rangle \psi_{m\mathbf{k}}(\mathbf{r}) e^{i\mathbf{k}\cdot\mathbf{R}} \\ &= \langle n\mathbf{0}|\psi_{m\mathbf{k}} \rangle e^{i\mathbf{k}\cdot\mathbf{R}} . \end{aligned}$$

Für die Substitution im zweiten Schritt wurde ausgenutzt, dass sowohl Wannier- wie auch Blochfunktionen auf dem periodischen Grundgebiet (PG) periodisch sind. Im dritten Schritt wurde die Translationseigenschaft (H.1) der Wannier-Funktionen benutzt und im vierten noch die Blochbedingung (3.5). Mit der Vollständigkeitsrelation² (8.3) folgt dann

$$|\psi_{m\mathbf{k}} \rangle = \sum_{n, \mathbf{R}} |n\mathbf{R} \rangle \langle n\mathbf{R}|\psi_{m\mathbf{k}} \rangle = \sum_{n, \mathbf{R}} |n\mathbf{R} \rangle \langle n\mathbf{0}|\psi_{m\mathbf{k}} \rangle e^{i\mathbf{k}\cdot\mathbf{R}} .$$

Diese Relation gilt auch dann, wenn in der Summe nicht über alle Bänder, sondern nur über die Bänder $n \in \{1, \dots, N\}$ einer „isolierten Gruppe“ (vgl. Abschnitt 8.1.1) summiert wird,

²Die Vollständigkeitsrelation (8.3) ist dort für klassische Wannier-Funktionen formuliert. Sie gilt aber gemäß Abschnitt 8.1.1 auch für Verallgemeinerte Wannier-Funktionen.

und $|\psi_{m\mathbf{k}}\rangle$ in dieser Gruppe enthalten ist.³ Damit gilt an jedem Punkt \mathbf{k} des reziproken Raums für die Ein-Teilchen-Energien:

$$\begin{aligned}\varepsilon_{m\mathbf{k}}\delta_{mm'} &= \langle\psi_{m\mathbf{k}}|H|\psi_{m'\mathbf{k}}\rangle \\ &= \left(\sum_{n,\mathbf{R}} e^{-i\mathbf{k}\cdot\mathbf{R}}\langle\psi_{m\mathbf{k}}|n\mathbf{0}\rangle\langle n\mathbf{R}| \right) H \left(\sum_{n',\mathbf{R}'} |n'\mathbf{R}'\rangle\langle n'\mathbf{0}|\psi_{m'\mathbf{k}}\rangle e^{i\mathbf{k}\cdot\mathbf{R}'}\right) \\ &= \sum_{n,n'} \langle\psi_{m\mathbf{k}}|n\mathbf{0}\rangle \left(\sum_{\mathbf{R},\mathbf{R}'} e^{-i\mathbf{k}\cdot(\mathbf{R}-\mathbf{R}')} \langle n\mathbf{R}|H|n'\mathbf{R}'\rangle\right) \langle n'\mathbf{0}|\psi_{m'\mathbf{k}}\rangle.\end{aligned}$$

Dies ist aber — komponentenweise geschrieben — gerade die unitäre Transformation beim Wechsel von der Wannier- zur Energiedarstellung von H , so dass die $\varepsilon_{m\mathbf{k}}$ also auch die Eigenwerte der Matrix $E^{\mathbf{k}}$ mit

$$\begin{aligned}E_{nm'}^{\mathbf{k}} &:= \sum_{\mathbf{R},\mathbf{R}'\in\text{PG}} e^{-i\mathbf{k}\cdot(\mathbf{R}-\mathbf{R}')} \langle n\mathbf{R}|H|n'\mathbf{R}'\rangle \\ &= \sum_{\mathbf{R},\mathbf{R}'\in\text{PG}} e^{-i\mathbf{k}\cdot(\mathbf{R}-\mathbf{R}')} \langle n,\mathbf{R}-\mathbf{R}'|H|n'\mathbf{0}\rangle \\ &= N_{\mathbf{R}} \sum_{\mathbf{R}\in\text{PG}} e^{-i\mathbf{k}\cdot\mathbf{R}} \langle n\mathbf{R}|H|n'\mathbf{0}\rangle\end{aligned}\tag{H.4}$$

sind. Im zweiten Schritt wurde die Gitterperiodizität des Hamilton-Operators verwendet, und für den letzten Schritt beachte man, dass für beliebiges festgehaltenes \mathbf{R}' die Summe über \mathbf{R} wegen der mit dem Grundgebiet PG periodischen Wannier-Funktionen identisch ist. Q.e.d.

³Es gilt nämlich $\sum_{n=1}^N \sum_{\mathbf{R}} |n\mathbf{R}\rangle\langle n\mathbf{R}| = \sum_{m=1}^N \int_{\text{BZ}} d^3k |\psi_{m\mathbf{k}}\rangle\langle\psi_{m\mathbf{k}}|$ wegen der Unitarität der Transformation (8.6). Dies ist aber gerade die 1 im Teilhilbertraum \mathcal{T} der isolierten Gruppe.

Literaturverzeichnis

- [1] P. Hohenberg und W. Kohn, Phys. Rev. **136**, B 864 (1964)
- [2] W. Kohn und L.J. Sham, Phys. Rev. **140**, A 1133 (1965)
- [3] H.B. Shore, J.H. Rose und E. Zaremba, Phys. Rev. B **15**, 2858 (1977)
- [4] K. Schwarz, Chem. Phys. Lett. **57**, 605 (1978)
- [5] J.P. Perdew und A. Zunger, Phys. Rev. B **23**, 5048 (1981)
- [6] J. Harris und R.O. Jones, J. Chem. Phys. **68**, 3316 (1978)
- [7] J.P. Perdew und M. Levy, Phys. Rev. Lett. **20**, 1884 (1983)
- [8] L.J. Sham und M. Schlüter, Phys. Rev. B **32**, 3883 (1985)
- [9] E. Schrödinger, Ann. d. Physik **79**, 361 (1926)
- [10] M. Born und J.R. Oppenheimer, Ann. d. Physik **84**, 457 (1927)
- [11] J. Bardeen, L.N. Cooper und J.R. Schrieffer, Phys. Rev. **108**, 1175 (1957)
- [12] D.R. Hartree, Proc. Camb. Phil. Soc. **24**, 328 (1928)
- [13] V.A. Fock, Z. Phys. **61**, 126 (1930)
- [14] C. Møller und M.S. Plesset, Phys. Rev. **46**, 618 (1934)
- [15] A. Szabo und N.S. Ostlund, *Modern Quantum Chemistry*, McGraw-Hill, New York (1989)
- [16] J.C. Slater und J.H. Wood, Int. J. Quantum Chem. **4**, 3 (1971)
- [17] L.H. Thomas, Proc. Camb. Phil. Soc. **23**, 542 (1927)
- [18] E. Fermi, Rend. Accad. Lincei **6**, 602 (1927)
- [19] E. Fermi, Z. Phys. **48**, 73–79 (1928)
- [20] E. Fermi, Rend. Accad. Lincei **7**, 342 (1928)
- [21] R.G. Parr und W. Yang, *Density-functional theory of atoms and molecules*, Oxford University Press (1989)
- [22] E. Teller, Rev. Mod. Phys. **34**, 627 (1962)

- [23] C.-O. Almbladh und A.C. Pedroza, *Phys. Rev. A* **29**, 2322 (1984)
- [24] R.M. Dreizler und E.K.U. Gross, *Density Functional Theory*, Springer, Berlin (1990)
- [25] C.-O. Almbladh und U. von Barth, *Phys. Rev. B* **31**, 3231 (1985)
- [26] F.J. Janak, *Phys. Rev. B* **18**, 7165 (1978)
- [27] M. Levy, *Proc. Natl. Acad. Sci. USA* **76**, 6062 (1979)
- [28] T.L. Gilbert, *Phys. Rev. B* **12**, 2111 (1975)
- [29] D.C. Langreth und M.J. Mehl, *Phys. Rev. B* **28**, 1809 (1983)
- [30] M. Levy und J.P. Perdew, *Phys. Rev. A* **32**, 2010 (1985)
- [31] M. Levy, *Phys. Rev. A* **43**, 4637 (1991)
- [32] E.J. Baerends, *Phys. Rev. Lett.* **87**, 133004-1 (2001)
- [33] U. von Barth und L. Hedin, *J. Phys. C* **5**, 1629 (1972)
- [34] D.M. Ceperley und B.L. Alder, *Phys. Rev. Lett.* **45**, 566 (1980)
- [35] A.D. Becke, *Phys. Rev. A* **38**, 3098 (1988)
- [36] J.P. Perdew, *Phys. Rev. B* **33**, 8822 (1986)
- [37] C.L. Lee, W. Yang und R.G. Parr, *Phys. Rev. B* **37**, 785 (1988)
- [38] J.P. Perdew und Y. Wang, *Phys. Rev. B* **45**, 13244 (1992)
- [39] J.P. Perdew, K. Burke und M. Ernzerhof, *Phys. Rev. Lett.* **77**, 3865 (1996)
- [40] A. Görling und M. Levy, *Phys. Rev. B* **47**, 13 105 (1993)
- [41] A. Görling und M. Levy, *Phys. Rev. A* **50**, 196 (1994)
- [42] A. Görling und M. Levy, *Int. J. Quantum Chem., Symp.* **29**, 93 (1995)
- [43] M. Levy und J.P. Perdew in *Density Functional Methods In Physics* (Hrsg.: R.M. Dreizler und J. Providencia), Plenum Press, New York (1985)
- [44] C.J. Umrigar und X. Gonze, *Phys. Rev. A* **50**, 3827 (1994)
- [45] C. Filippi, X. Gonze und C.J. Umrigar in *Recent Developments and Applications of Density Functional Theory* (Hrsg.: J.M. Seminario), Elsevier, Amsterdam (1996)
- [46] In-Ho Lee und R.M. Martin, *Phys. Rev. A* **56**, 7197 (1997)
- [47] A. Görling, *Phys. Rev. B* **53**, 7024 (1996)
- [48] T. Grabo, T. Kreibich und E.K.U. Gross, *Mol. Eng.* **7**, 27 (1997)
- [49] T. Grabo, T. Kreibich, S. Kurth und E.K.U. Gross in *Strong Coulomb Correlation in Electronic Structure* (Hrsg.: V.I. Anisimov), Gordon & Breach Science Publishers (2000)

- [50] E. Engel und R.M. Dreizler, *J. Comput. Chem.* **20**, 31 (1999)
- [51] M. Städele, J.A. Majewski, P. Vogl und A. Görling, *Phys. Rev. Lett.* **79**, 2089 (1997)
- [52] M. Städele, M. Moukara, J.A. Majewski, P. Vogl und A. Görling, *Phys. Rev. B*, **59**, 10 031 (1999)
- [53] W.G. Aulbur, M. Städele und A. Görling, *Phys. Rev. B* **62**, 7121 (2000)
- [54] T. Kreibich, E.K.U. Gross und E. Engel, *Phys. Rev. A* **57**, 138 (1998)
- [55] A. Görling, *Phys. Rev. A* **46**, 3753 (1992)
- [56] S. Hirata, S. Ivanov, I. Grabowski, R.J. Bartlett, K. Burke und J.D. Talman, *J. Chem. Phys.* **115**, 1635 (2001)
- [57] T. Kotani, *Phys. Rev. B* **50**, 14 816 (1994)
- [58] Q. Zhao und R.G. Parr, *Phys. Rev. A* **46**, 2337 (1992)
- [59] Q. Zhao und R.G. Parr, *J. Chem. Phys.* **98**, 543 (1992)
- [60] E. Engel und S.H. Vosko, *Phys. Rev. A* **47**, 2800 (1993)
- [61] S. Ivanov, S. Hirata und R.J. Bartlett, *Phys. Rev. Lett.* **83**, 5455 (1999)
- [62] A. Görling, *Phys. Rev. Lett.* **83**, 5459 (1999)
- [63] R.T. Sharp and G.K. Horton, *Phys. Rev.* **90**, 317 (1953)
- [64] J.C. Slater, *Phys. Rev.* **81**, 385, (1951)
- [65] J.D. Talman und W.F. Shadwick, *Phys. Rev. A* **14**, 36 (1976)
- [66] K. Aashamar, T.M. Luke und J.D. Talman, *At. Data and Nucl. Data Tables* **22**, 443 (1978)
- [67] K. Aashamar, T.M. Luke und J.D. Talman, *Phys. Rev. A* **19**, 6 (1979)
- [68] V. Sahni, J. Gruenebaum und J.P. Perdew, *Phys. Rev. B* **26**, 4371 (1982)
- [69] J.P. Perdew und M.R. Norman, *Phys. Rev. B* **26**, 5445 (1982)
- [70] J.B. Krieger, Y. Li und G.J. Iafrate, *Phys. Lett. A* **146**, 256 (1990)
- [71] J.B. Krieger, Y. Li und G.J. Iafrate, *Phys. Rev. A* **45**, 101 (1992)
- [72] J.B. Krieger, Y. Li und G.J. Iafrate, *Phys. Rev. A* **46**, 5453 (1992)
- [73] J.B. Krieger, Y. Li und G.J. Iafrate, *Phys. Rev. A* **47**, 165 (1993)
- [74] D.M. Bylander und L. Kleinman, *Phys. Rev. Lett.* **74**, 3660 (1995)
- [75] D.M. Bylander und L. Kleinman, *Phys. Rev. B* **52**, 14 566 (1995)
- [76] D.M. Bylander und L. Kleinman, *Phys. Rev. B* **54**, 7891 (1996)
- [77] D.M. Bylander und L. Kleinman, *Phys. Rev. B* **55**, 9432, (1997)

-
- [78] M. Wahn, *Implementation und Analyse des exakten Austauschfunctionals in der Dichtefunktionaltheorie*, Diplomarbeit, Technische Universität Berlin (2002)
- [79] S. Kümmel und J.P. Perdew, Phys. Rev. Lett. **90**, 043 004 (2003)
- [80] S. Kümmel und J.P. Perdew, Phys. Rev. B **68**, 035 103 (2003)
- [81] S/PHI/nX, auf C++ basierte modulare DFT-Bibliothek (<http://sphinxlib.de>)
- [82] S. Boeck, A. Dick, Ch. Freysoldt, F. Grzegorzewski, T. Hammerschmidt, L. Ismer, L. Lymperakis und M. Wahn, *SFHingX User's Guide* (Hrsg.: S. Boeck und J. Neugebauer), Berlin (2003), online siehe Ref. [81](#)
- [83] E.N. Zein, Fiz. Tverd. Tela (Leningrad) **26**, 3024 (1984) [englische Übersetzung in Sov. Phys. Solid State **26**, 1825 (1984)]
- [84] S. Baroni, P. Giannozzi und A. Testa, Phys. Rev. Lett. **58**, 1861 (1987)
- [85] S. Baroni, S. de Gironcoli, A. dal Corso und P. Giannozzi, Rev. Mod. Phys. **73**, 515 (2001)
- [86] R.M. Sternheimer führte mit Hilfe dieser Gleichung numerische Berechnungen zur Polarisierbarkeit von Ionen durch, Phys. Rev. **96**, 951 (1954)
- [87] F. della Sala und A. Görling, J. Chem. Phys. **115**, 5718 (2001)
- [88] C. Kittel, *Introduction to Solid State Physics*, 7th ed., Wiley, New York (1996)
- [89] O. Madelung, *Introduction to Solid-State Theory*, Springer, Berlin (1978)
- [90] R. Enderlein und N.J.M. Horing, *Fundamentals of Semiconductor Physics*, World Scientific, Singapore (1997)
- [91] P.J.H. Denteneer und W. van Haeringen, J. Phys. C **18**, 4127 (1985)
- [92] U. Scherz, *Quantenmechanik*, Teubner, Stuttgart (1999)
- [93] A. Baldereschi, Phys. Rev. B **7**, 5212 (1973)
- [94] D.J. Chadi und M.L. Cohen, Phys. Rev. B **8**, 5747 (1973)
- [95] J.D. Joannopoulos und M.L. Cohen, J. Phys. C **6**, 1572 (1973)
- [96] H.J. Monkhorst und J.D. Pack, Phys. Rev. B **13**, 5188 (1976)
- [97] R.A. Evarestov und V.P. Smirnov, Phys. Stat. Sol. B **119**, 9 (1983)
- [98] F. Gygi und A. Baldereschi, Phys. Rev. B **34**, 4405 (1986)
- [99] B. Wenzien, G. Cappellini und F. Bechstedt, Phys. Rev. B **51**, 14 701 (1995)
- [100] D.J. Singh, *Planewaves, pseudopotentials and the LAPW method*, Kluwer, Boston (1994)
- [101] E. Sjöstedt, L. Nordström und D. J. Singh, Solid State Comm. **114**, 15 (2000)

- [102] G.K.H. Madsen, P. Blaha, K. Schwarz, E. Sjöstedt und L. Nordström, Phys. Rev. B **64**, 195 134 (2001)
- [103] P.E. Blöchl, Phys. Rev. B **50**, 17 953 (1994)
- [104] S.G. Louie, S. Froyen und M.L. Cohen, Phys. Rev. B **26**, 1738 (1982)
- [105] M. Fuchs, M. Bockstedte, E. Pehlke und M. Scheffler, Phys. Rev. B **57**, 2134 (1998)
- [106] D.R. Hamann, M. Schlüter und C. Chiang, Phys. Rev. Lett. **43**, 1494 (1979)
- [107] G. Kerker, J. Phys. C **13**, L189 (1980)
- [108] G.B. Bachelet, D.R. Hamann und M. Schlüter, Phys. Rev. B **26**, 4199 (1982)
- [109] N. Troullier und J.L. Martins, Phys. Rev. B **43**, 1993 (1991)
- [110] D.R. Hamann, Phys. Rev. B **40**, 2980 (1989)
- [111] M. Moukara, *Pseudopotentiale mit exaktem Austausch*, Dissertation, Technische Universität München (1998)
- [112] M. Moukara, M. Städele, J.A. Majewski, P. Vogl und A. Görling, J. Phys.: Condens. Matter **12**, 6783 (2000)
- [113] L. Kleinman und D.M. Bylander, Phys. Rev. Lett. **48**, 1425 (1982)
- [114] X. Gonze, R. Stumpf und M. Scheffler, Phys. Rev. B **44**, 8503 (1991)
- [115] M. Fuchs und M. Scheffler, Comput. Phys. Commun. **119**, 67 (1999)
- [116] N. Marzari und D. Vanderbilt, Phys. Rev. B **56**, 12 847 (1997)
- [117] I. Souza, N. Marzari und D. Vanderbilt, Phys. Rev. B **65**, 035 109 (2001)
- [118] G.H. Wannier, Phys. Rev. **52**, 191 (1937)
- [119] E.I. Blount, Solid State Phys. **13**, 305 (1962)
- [120] P.-O. Löwdin, J. Chem. Phys. **18**, 365 (1950)
- [121] P.-O. Löwdin, Adv. Phys. **5**, 172 (1956)
- [122] M. Wahn und J. Neugebauer, Phys. Stat. Sol. B **243**, No. 7, 1583 (2006)
- [123] R. Wüst, *Höhere Mathematik für Physiker* (Band 1 und 2), de Gruyter, Berlin (1995)
- [124] L. Collatz, *Differentialgleichungen*, Teubner, Stuttgart (1990)
- [125] W. Nolting, *Grundkurs Theoretische Physik*, Bd. 5, Teil 1, Abschnitt 3.2.7, Zimmermann-Neufang, Ulmen (1989)
- [126] M.R. Spiegel, *Komplexe Variablen*, McGraw-Hill, London (1977)
- [127] F. Schwabl, *Quantenmechanik*, Springer, Berlin (1988)
- [128] E. Polak, *Computational Methods in Optimization*, Academic Press, New York (1971)

- [129] W.H. Press, S.A. Teulosky, W.T. Vetterling, und B.P. Flannery, *Numerical Recipes in C: the art of scientific computing*, Cambridge University Press (1992)
- [130] M.C. Payne, M.P. Teter, D.C. Allen, T.A. Arias, und J.D. Joannopoulos, *Rev. Mod. Phys.* **64**, 1045 (1992)
- [131] J.R. Shewchuk, *An Introduction to the Conjugate Gradient Method Without the Agonizing Pain*, Edition 1 $\frac{1}{4}$, Pittsburgh, USA (1994),
<http://www.cs.cmu.edu/~quake-papers/painless-conjugate-gradient.pdf>
- [132] M.P. Teter, M.C. Payne und D.C. Allan, *Phys. Rev. B* **40**, 12 255 (1989)
- [133] S. Ismail-Beigi und T.A. Arias, *Comput. Phys. Commun.* **128**, 1–45 (2000)
- [134] P.H. Dederichs und R. Zeller, *Phys. Rev. B* **28**, 5462 (1983)
- [135] H. Akai und P.H. Dederichs, *J. Phys. C* **18**, 2455 (1985)
- [136] C.G. Broyden, *Math. Comput.* **19**, 577 (1965)
- [137] D.D. Johnson, *Phys. Rev. B* **38**, 12 807 (1988)
- [138] G.P. Srivastava, *J. Phys. A* **17**, L317 (1984)
- [139] D. Vanderbilt und S.G. Louie, *Phys. Rev. B* **30**, 6118 (1984)
- [140] P. Pulay, *Chem. Phys. Lett.* **73**, 393 (1980)
- [141] P. Pulay, *J. Comp. Chem.* **3**, 556 (1982)
- [142] G. Kresse und J. Furthmüller, *Comput. Mat. Sci.* **6**, 15 (1996)
- [143] G. Kresse und J. Furthmüller, *Phys. Rev. B* **54**, 11 169 (1996)
- [144] D.R. Bowler und M.J. Gillan, *Chem. Phys. Lett.* **325**, 473 (2000)
- [145] G.P. Kerker, *Phys. Rev. B* **23**, 3082 (1981)
- [146] R. Car, E. Tosatti, S. Baroni und S. Leelaprute, *Phys. Rev. B* **24**, 985 (1981)
- [147] K.-M. Ho, J. Ihm und J.D. Joannopoulos, *Phys. Rev. B* **25**, 4260 (1982)
- [148] J.F. Annett, *Comput. Mat. Sci.* **4**, 23 (1995)
- [149] O. Gunnarsson und B.I. Lundqvist, *Phys. Rev. B* **13**, 4274 (1976)
- [150] L. Hedin, *Phys. Rev.* **139**, A796 (1965)
- [151] L. Hedin und S. Lundqvist in *Solid State Physics, Advances in Research and Application* (Hrsg.: F. Seitz, D. Turnbull und H. Ehrenreich), Academic, New York (1969), Band 23, S. 1
- [152] M. Rohlfing, P. Krüger und J. Pollmann, *Phys. Rev. B* **57**, 6485 (1998)
- [153] Vgl. den Übersichtsartikel von S.C. Jain, M. Willander, J. Narayan und R. van Overstraeten: *J. Appl. Phys.* **87**, 965 (2000)

- [154] C. Stampfl und C.G. Van de Walle, *Phys. Rev. B* **59**, 5521 (1999)
- [155] F. Bechstedt und J. Furthmüller, *J. Cryst. Growth* **246**, 315 (2002)
- [156] *Intrinsic Properties of Group IV Elements and III-V, II-VI, and I-VII Compounds* (Hrsg.: K.H. Hellwege und O. Madelung), Landolt-Börnstein, New Series, Group III, Vol. 22, Pt. A, Springer, Berlin (1987)
- [157] J. Wu, W.W. Walukiewicz, K.M. Yu, J.W. Ager III, E.E. Haller, H. Lu, W.J. Schaff, Y. Saito und Y. Nanishi, *Appl. Phys. Lett.* **80**, 3967 (2002)
- [158] T. Matsuoka, H. Okamoto, M. Nakao, H. Harima und E. Kurimoto, *Appl. Phys. Lett.* **81**, 1246 (2002)
- [159] V.Y. Davydov, A.A. Klochikhin, V.V. Emtsev, S.V. Ivanov, V.V. Vekshin, F. Bechstedt, J. Furthmüller, H. Harima, A.V. Mudryi, A. Hashimoto, Y. Yamamoto, J. Aderhold, J. Graul, E.E. Haller, *Phys. Stat. Sol. B* **230**, R4 (2002)
- [160] T. Kotani und M. van Schilfgaarde, *Solid State Commun.* **121**, 461 (2002)
- [161] M. Usuda, N. Hamada, K. Shiraishi und A. Oshiyama, *Jpn. J. Appl. Phys., Part 2*, **43**, L407 (2004)
- [162] T. Lei, M. Fanciulli, R.J. Molnar, T.D. Moustakas, R.J. Graham, und J. Scanlon, *Appl. Phys. Lett.* **59**, 944 (1992)
- [163] T. Lei, T.D. Moustakas, R.J. Graham, Y. He und S.J. Berkowitz, *J. Appl. Phys.* **71**, 4933 (1992)
- [164] C. R. Eddy, T. Moustakas und J. Scanlon, *J. Appl. Phys.* **73**, 448 (1993)
- [165] M.J. Paisley, Z. Sitar, J.B. Posthill und R.F. Davis, *J. Vac. Sci. Technol. A* **7**, 701 (1989)
- [166] Z. Sitar, M.J. Paisley, J. Ruan, J.W. Choyke und R.F. Davis, *J. Mater. Sci. Lett.* **11**, 261 (1992)
- [167] I. Vurgaftman und J.R. Meyer, *J. Appl. Phys.* **94**, 3675 (2003)
- [168] A. Qteish, private Korrespondenz
- [169] A. Qteish, A.I. Al-Sharif, M. Fuchs, M. Scheffler, S. Boeck und J. Neugebauer, *Phys. Rev. B* **72**, 155 317 (2005)
- [170] A. Qteish, P. Rinke, M. Scheffler und J. Neugebauer, *Phys. Rev. B* **74**, 245 208 (2006)
- [171] P. Rinke, A. Qteish, J. Neugebauer, C. Freysoldt und M. Scheffler, *New J. Phys.* **7**, 126 (2005)
- [172] P. Rinke, M. Scheffler, A. Qteish, M. Winkelnkemper, D. Bimberg und J. Neugebauer, *Appl. Phys. Lett.* **89**, 161 919 (2006)
- [173] P. Rinke, A. Qteish, J. Neugebauer und M. Scheffler, Ψ_k Newsletter **79**, 163 (2007)

- [174] P. Rinke, M. Winkelkemper, A. Qteish, D. Bimberg, J. Neugebauer und M. Scheffler, *Consistent set of band parameters for the group-III nitrides AlN, GaN, and InN*. Phys. Rev. B, akzeptiert
- [175] P. Rinke, A. Qteish, J. Neugebauer und M. Scheffler, *Exciting prospects for solids: Exact-exchange based functionals meet quasiparticle energy calculations*. Phys. Stat. Sol. B, akzeptiert
- [176] M.S. Hybertsen und S.G. Louie, Phys. Rev. B **35**, 5585 (1987)
- [177] L.J. Sham und M. Schlüter, Phys. Rev. Lett. **51**, 1888 (1983)
- [178] M.E. Casida, Phys. Rev. A **51**, 2005 (1995)
- [179] E.H. Lieb, Int. J. Quantum Chem. **24**, 243 (1983)

Danksagung

Mein Dank gilt natürlich an erster Stelle Professor Dr. Jörg Neugebauer, der mir die Gelegenheit zur Durchführung dieser Doktorarbeit gegeben hat. Ohne seine Betreuung, sein stetes Interesse am Projekt und die unzähligen fruchtbaren Diskussionen mit ihm wäre diese Arbeit kaum zu verwirklichen gewesen.

Herrn Professor Dr. Scheffler danke ich dafür, dass er mir die Möglichkeit gab, diese Arbeit am Fritz-Haber-Institut der Max-Planck-Gesellschaft (im vielleicht schönsten Arbeitszimmer des Instituts) anzufertigen.

Sixten Boeck und Christoph Freysoldt möchte ich vor allem für ihre unermüdliche Hilfe bei der Einarbeitung in das Programmieren mit C++ danken, aber auch für die vielen anderen, nicht nur wissenschaftlichen, Diskussionen. Patrick Rinke möchte ich ebenfalls für viele kreative Hinweise und Gespräche danken und Professor Dr. Qteish für die Bereitstellung der OEP-Pseudopotentiale. Ein herzlicher Dank geht an Liverios Lymperakis, mit dem ich lange Zeit das Büro geteilt habe.

In sehr angenehmer Erinnerung bleibt mir die Zusammenarbeit mit Stefan Schulz aus der Gruppe von Professor Dr. Czycholl an der Uni Bremen.

Außerdem gilt mein Dank allen Mitgliedern der Abteilung Computergestütztes Material-Design am MPI für Eisenforschung sowie der Theorie-Abteilung des Fritz-Haber-Instituts. Die gute Atmosphäre in diesen beiden Gruppen hat insgesamt sehr zum Gelingen der Arbeit beigetragen.

Ebenfalls möchte ich meinen besten Freunden danken, die mich auf dem nicht immer völlig mühelosen Weg immer wieder motiviert haben. Meiner lieben Julia danke ich für die Kraft und das nötige Durchhaltevermögen, das sie mir in den letzten Monaten dieser Arbeit gegeben hat.

Ganz besonders aber möchte ich an dieser Stelle meinen Eltern Dank sagen! Sie haben mir Entscheidendes ermöglicht, ohne das ich diese Seiten hier nicht hätte schreiben können!