

**Die *Projector Augmented Wave*-Methode:
Ein schnelles Allelektronenverfahren
für die *ab-initio*-Molekulardynamik**

Winfried Kromen

Berichte des Forschungszentrums Jülich ; 3887

ISSN 0944-2952

Institut für Festkörperforschung Jül-3887

D82 (Diss., Aachen, RWTH, 2001)

Zu beziehen durch: Forschungszentrum Jülich GmbH · Zentralbibliothek

52425 Jülich · Bundesrepublik Deutschland

☎ 02461/61-5220 · Telefax: 02461/61-6103 · e-mail: zb-publikation@fz-juelich.de

Abstract

The rapid development of information technology is based on the continuing miniaturization of electronic devices as well as on the development of new materials and entirely new concepts. For further development, the understanding of growth and diffusion mechanisms on a microscopic scale is absolutely essential. Valuable insight can be gained using *ab-initio* computer simulations. Probably the most successful theory for solving the quantum mechanical many-particle problem is the density functional theory (DFT).

In this thesis, a new *ab-initio* method for molecular dynamics within the framework of DFT was developed and implemented. It is strongly related to the projector-augmented-wave method (PAW) by P. Blöchl, but also includes elements of a pseudocharge method proposed by M. Weinert for the full-potential linearized augmented plane wave method (FLAPW).

Like the FLAPW-method, the PAW-method is an all-electron method. In the later case, the hamiltonian takes quite a simple form similar to the one obtained when using normconserving Kleinman-Bylander-type pseudopotentials or ultrasoft Vanderbilt pseudopotentials. The PAW-method allows for an efficient calculation of systems containing virtually any element of the periodic table, including those which require high numerical effort when treated using normconserving pseudopotentials (e.g. 2p-elements like oxygen or 3d-transitional metals like copper). In addition, a higher accuracy compared to the pseudopotentialmethod can be obtained due to the improved description of the density in the spacial vicinity of the atoms.

The method was implemented starting from the *ab-initio* molecular dynamics code **ESCoMPP**, which uses normconserving pseudopotentials. The new code allows for physical systems containing pseudopotential-atoms as well as atoms described using the PAW-formalism at the same time.

A program for the generation of the PAW-datasets needed to describe a particular type of atoms was developed as well. Essential ingredients of this datasets are localized atomic wavefunctions and projectors. A procedure was developed and tested which yields projectors leading to a numerically stable representation of the PAW-hamiltonian. Additional tools for checking the generated datasets were implemented, too.

Datasets were generated for carbon, nitrogen and copper. Tests were performed for several existing and non-existing bulk structures. Structural properties such as lattice constants, cohesion energies, bulk moduli as well as electronic

properties (bandstructure, density of states) calculated for materials like diamond, siliconcarbide or aluminumnitride show good agreement with results of other DFT-calculations.

Compared to calculations using normconserving pseudopotentials, the new PAW-code significantly reduces the numerical effort involved (e.g. reduction by 80% for SiC). If comparable conditions are applied, the program was shown to be faster than the FLAPW-code **FLEUR** by a factor of 3-4 in the case of a (4 x 2) Cu(110) surface cell. Possibilities for further computational improvements are indicated.

Like the pseudopotential code it is based on, the PAW-program employs a Car-Parrinello-like scheme. The physical system is represented as a periodical supercell. The electronic structure is determined in the framework of DFT in the local density approximation (LDA) using a basis of plane waves. The Kohn-Sham eigenvalue problem is solved in reciprocal space using an iterative Davidson-Kosugi algorithm. The selfconsistent electronic density is obtained using a quasi-newton algorithm (Broyden mixing). The equilibrium positions of the atoms are obtained by calculating the forces acting on the ions and by moving them according to a quasi-newton algorithm.

Zusammenfassung

Die rasante Entwicklung der Informationstechnologie beruht auf der fortschreitenden Miniaturisierung elektronischer Bauteile, aber ebenso auf der Erschließung neuer Materialien und vollkommen neuartiger Konzepte. Für weitere Fortschritte ist das Verständnis von Wachstums- und Diffusionsprozessen auf mikroskopischer Ebene wesentliche Voraussetzung. Wertvolle Einsichten lassen sich durch ab-initio Computersimulationen gewinnen. Die wahrscheinlich erfolgreichste Theorie zur Lösung des quantenmechanischen Vielteilchenproblems ist die Dichtefunktionaltheorie (DFT).

In dieser Arbeit ist eine neue ab-initio Methode für die Molekulardynamik im Rahmen der DFT entwickelt und implementiert worden. Sie ist eng mit der Projector-Augmented-Wave Methode (PAW) von P. Blöchl verwandt, enthält aber auch Elemente einer Pseudoladungsmethode, die von M. Weinert für die Full-potential Linearized Augmented Plane Wave Methode (FLAPW) vorgeschlagen wurde.

Wie die FLAPW-Methode ist die PAW-Methode eine Allelektronenmethode. Im Falle von PAW nimmt der Hamiltonian eine relativ einfache Form an, ähnlich derjenigen, die man bei Benutzung von normerhaltenden Kleinman-Bylander Pseudopotentialen oder ultrasofthen Pseudopotentialen nach D. Vanderbilt erhält. Die PAW-Methode erlaubt eine effiziente Berechnung von Systemen mit nahezu beliebigen Elementen, auch solchen, die bei Benutzung normerhaltender Pseudopotentiale hohen Rechenaufwand erfordern (z.B. 2p-Elemente wie Sauerstoff oder 3d-Übergangsmetalle wie Kupfer). Zudem läßt sich durch die bessere Beschreibung der Dichte in der Nähe der Atome eine im Vergleich zu normerhaltenden Potentialen höhere Genauigkeit erzielen.

Die Methode wurde ausgehend vom ab-initio Molekulardynamikprogramm **ESiCoMPP** entwickelt, das normerhaltende Pseudopotentiale verwendet. Das neue Programm erlaubt es, physikalische Systeme zu beschreiben, die gleichzeitig Pseudopotential-Atome und durch den PAW-Formalismus beschriebene Atome enthalten.

Ein Programm zur Erzeugung der PAW-Datensätze, die zur Beschreibung jeder Atomsorte benötigt werden, wurde ebenfalls entwickelt. Wesentliche Bestandteile dieser Datensätze sind lokalisierte atomare Wellenfunktionen und Projektoren. Es wurde ein Verfahren entwickelt und erprobt, das Projektoren ergibt, die zu einem numerisch stabilen Hamiltonian führen. Zusätzliche Programme zur Qualitätsprüfung der Datensätze wurden ebenfalls entwickelt.

PAW-Datensätze wurden für Kohlenstoff, Stickstoff und Kupfer generiert. Testrechnungen wurden an verschiedenen existierenden und hypothetischen Volumenstrukturen durchgeführt. Strukturelle Eigenschaften wie Gitterkonstanten, Kohäsionsenergien und bulk modulus und elektronische Eigenschaften wie Bandstrukturen und Zustandsdichten zeigen für Materialien wie Diamant, Siliziumkarbid eine gute Übereinstimmung mit den Ergebnissen anderer DFT-Rechnungen.

Gegenüber Berechnungen, die mit normerhaltenden Pseudopotentialen durchgeführt wurden, wird der numerische Aufwand erheblich reduziert, z.B. im Falle von Siliziumkarbid um 80%. Bei vergleichbaren Bedingungen war das Programm im Falle einer $(4 \times 2)\text{Cu}(110)$ Oberfläche um einen Faktor 3-4 schneller als das FLAPW-Programm **FLEUR**. Möglichkeiten zu weiteren algorithmischen Verbesserungen werden aufgezeigt.

Wie das zugrundeliegende Pseudopotentialprogramm verwendet das PAW-Programm ein Car-Parrinello-artiges Schema. Das physikalische System wird als eine periodische Superzelle dargestellt. Die elektronische Struktur wird im Rahmen der DFT in der lokale-Dichte-Näherung (LDA) mit einer Basis aus Ebenen Wellen bestimmt. Das Kohn-Sham-Eigenwertproblem wird im Impulsraum mit einem iterativen Davidson-Kosugi-Algorithmus gelöst. Die selbstkonsistente elektronische Dichte erhält man mittels einer Quasi-Newton-Methode (Broyden-Mischungsverfahren). Die Gleichgewichtspositionen der Atomrümpfe werden bestimmt, indem die auf sie wirkenden Kräfte berechnet werden und die Ionen gemäß einem Quasi-Newton-Verfahren bewegt werden.

Inhaltsverzeichnis

1	Einleitung	1
2	Dichtefunktionaltheorie und Pseudopotentialmethode	9
2.0	Einheiten	10
2.1	Die Dichte als fundamentale Systemvariable	10
2.2	Die Kohn–Sham Gleichungen	12
2.3	Die Spindichtefunktionaltheorie	15
2.4	Die Lokale–Dichte–Näherung	15
2.5	Die Pseudopotential–Methode	17
3	PAW-Methode	25
3.1	Superzellenmethode und Ebene Wellen	25
3.2	Grundidee	30
3.3	Definition der Transformation	33
3.4	Operatoren	36
3.5	Dichten	37
3.6	Das Energiefunktional	39
3.6.1	Kinetische Energie	39
3.6.2	Austausch-Korrelations-Energie	40
3.6.3	Elektrostatische Energie	41
3.6.4	Resümee	48
3.7	Kohn–Sham–Gleichungen	49
3.8	Überlappmatrix	49
3.9	Ableitung des Hamiltonoperators	50
3.9.1	Formelsammlung	51
3.9.2	Resümee	52
3.10	Berechnung des Hamiltonoperators	53
3.10.1	Kinetische Energie	53

3.10.2	Lokales Potential	53
3.10.3	Nichtlokales Potential	55
3.10.4	Zusammenfassung	58
3.11	Die Pseudoladungsmethode	58
3.12	Die $G=0$ -Komponente	60
3.13	Double counting correction	63
3.14	Kräfte	64
3.15	Symmetrie	67
3.16	Symbolübersicht	72
4	Pseudopotentialerzeugung	75
4.1	Erzeugung der <i>partial waves</i>	76
4.2	Erzeugung der Projektoren	80
4.3	<i>Unscreening</i>	84
4.4	Die praktische Vorgehensweise	85
4.5	Kohlenstoff	87
4.6	Stickstoff	100
4.7	Kupfer	104
5	Minimierung des Energiefunktional	111
5.1	Ionische und elektronische Freiheitsgrade	111
5.2	Lösen der Eigenwertgleichung	114
5.2.1	Traditionelle Diagonalisierung	114
5.2.2	Iterative Diagonalisierung	115
5.2.2.1	Rayleigh-Ritz Verfahren für Standardeigenwertpro- bleme	115
5.2.2.2	Davidson-Kosugi-Verfahren	116
5.2.2.3	Verallgemeinerte Eigenwertprobleme	118
5.2.2.4	Der Algorithmus	119
5.2.2.5	Startvektoren	120
5.2.3	Präkonditionierung	125
5.2.4	Jacobi-Davidson Methode	127
6	Test der PAW-Methode	141
6.1	Kohlenstoffgitter	141
6.2	Molekularer Stickstoff	145
6.3	Kupfer	147

6.3.1	Volumeneigenschaften von fcc-Kupfer	148
6.3.2	Volumeneigenschaften von hcp-Kupfer	151
6.3.3	Effizienztest: Cu(110) Oberfläche	152
7	Halbleiter mit großer Bandlücke	159
7.1	Siliziumkarbid	159
7.2	Aluminiumnitrid und Galliumnitrid	167
8	Zusammenfassung	181
A	Formelsammlung	185
B	Inverse der PAW-Transformation	191
C	Verbesserungsmöglichkeit für das Davidson-Verfahren	193
D	Partielle lokale Zustandsdichte	197
D.1	Weichröntgenspektroskopie SXE	200
	Literaturverzeichnis	203
	Danksagung	213

Kapitel 1

Einleitung

Die rasante Entwicklung in der Informationstechnologie der letzten Jahre beruht auf der fortschreitenden Miniaturisierung elektronischer Bauteile auf der Grundlage bekannter Halbleitermaterialien wie z.B. Silizium, aber ebenso auf der Erschließung neuer Materialien für neuartige Anwendungen und für die Realisierung vollkommen neuer Konzepte. Die Entwicklung der Technologie für Mikroprozessoren geht einher mit ständig steigenden Anforderungen an Speicher- und Datenübertragungstechnik. Immer neue Materialien und Materialkombinationen werden benötigt, um den speziellen chemischen und physikalischen Anforderungen zu genügen.

In diesem Zusammenhang gelten Halbleiter mit großer Bandlücke als vielversprechende Materialien. So eignet sich beispielsweise Siliziumkarbid (SiC) zur Herstellung elektronischer Schaltkreise, welche bei hohen Temperaturen arbeiten können und unempfindlich gegen hochenergetische Strahlung sind. Auch ist es für die Hochfrequenztechnik interessant. Schließlich eignet es sich auch als Substrat für Heterostrukturen mit Nitridhalbleitern. Galliumnitrid (GaN) und Aluminiumnitrid (AlN) eignen sich auf Grund ihrer elektronischen Eigenschaften als Material für blaue Laserdioden. Eine zuverlässige Laserdiode in diesem Bereich des Spektrums billig herstellen zu können, ist beispielsweise für optische Speichermedien sehr interessant, da sich auf Grund der kürzeren Wellenlänge gegenüber den heute üblichen roten Lasern eine höhere Informationsdichte erzielen läßt. Während blaue LED's auf GaN-Basis bereits kommerziell erhältlich sind, sind andere Anwendungsmöglichkeiten z.B. in der Hochfrequenzelektronik, noch weiter von der Marktreife entfernt.

Auch die Miniaturisierung von (flüchtigen) Speicherbausteinen stößt in Gebiete vor, wo herkömmliche Herstellungsweisen an ihre Grenze gelangen. Um die Kondensatoren, welche durch ihren Ladungszustand ein Bit repräsentieren, weiter verkleinern zu

können, untersucht man verstärkt Verbindungen mit hohen Dielektrischen Konstanten – sogenannten Hohepsilonmaterialien wie z.B. die perovskitischen Ferroelektrika BaTiO_3 oder PbZrO_3 .

In naher Zukunft wird aber auch die Miniaturisierung Grenzen (≈ 50 nm) erreichen, unter der Bauelemente bei der jetzigen Bauweise nicht mehr funktionsfähig sein werden. Derzeit verfolgt man zwei neue Konzepte: zum einen die Spinelektronik basierend auf neuen magnetischen Halbleitern wie MnGaN oder MnGaAs , welche Bauelemente mit komplexerer Logik erlauben, und zum anderen die Molekularelektronik. Mit der Entdeckung der Kohlenstoffnanoröhrchen im Jahre 1991 stehen große Moleküle zum Bau von Nanobauelementen zur Verfügung, deren Abmessungen erheblich unter der Grenze von ≈ 50 nm liegen werden.

Aktive und passive Bauelemente sind charakterisiert durch strukturelle Inhomogenitäten. Dies können scharf definierte innere Grenzflächen sein wie zum Beispiel eine Metall-Halbleitergrenzfläche bei der Erzeugung einer Schottky-Barriere oder Halbleiter-Halbleitergrenzflächen im Falle von Halbleiterheterostrukturen; es ist aber auch möglich, gezielt Dichtemodulation von p- oder n-Typ Dotierungen zu verwirklichen. Die Bauelemente sind getrennt durch Passivierungsschichten und werden durch Leiterbahnen kontaktiert. Die Miniaturisierung verlangt eine Beherrschung der Herstellung und der Stabilität von Grenzflächen und Dotierungsschichten auf der Nanometerskala und in Zukunft auf der Subnanometerskala. Dazu ist ein mikroskopisches Verständnis der Wachstums- und Diffusionsmechanismen auf atomarer Ebene nötig. Die Einsicht in atomare Prozesse garantiert das Verständnis der Selektivität chemischer Reaktivität bei den Ätzprozessen für die Strukturierung, die chemische Stabilität und die Temperaturstabilität von Grenzflächen sowie die Modifikation und Kontrolle der kinetischen Wachstumsprozesse.

Angesichts dessen ist zu erwarten, daß in Zukunft die Eigenschaften von Quantenmaterialien noch stärker ins Zentrum des Interesse treten werden. Die Komplexität und die Vielfalt der chemischen Komponenten werden steigen. Ein mikroskopisches Verständnis auf der Basis der Quantennatur der Elektronen ist heute schon eine wichtige Voraussetzung für den weiteren Fortschritt.

Die derzeit wohl erfolgreichste Theorie zur Lösung des quantenmechanischen Vielelektronenproblems realer Materialien ist die Dichtefunktionaltheorie. In dieser Theorie, entwickelt von Hohenberg und Kohn [HK64] und nutzbar gemacht durch Kohn und Sham [KS65], werden die Elektronen quantenmechanisch und die Freiheitsgrade der Atomkerne klassisch behandelt. Anpaßbare Parameter gehen nicht ein, d.h. es ist eine *ab-initio*-Theorie. 1998 wurde die Bedeutung dieser Theorie durch die Ver-

leihung des Nobelpreises an W. Kohn gewürdigt. In der lokalen Dichtenäherung, von der in dieser Arbeit ausschließlich Gebrauch gemacht wird, sind Grundzustandseigenschaften von Halbleitern und Metallen sowie einfachen Oxiden mit guter Vorhersagekraft bestimmbar.

In einer bahnbrechenden Arbeit haben im Jahre 1985 R. Car und M. Parrinello [CP85] die Molekulardynamik mit der Dichtefunktionaltheorie verknüpft. In dieser Methode verursachen die Elektronen Kräfte auf die Atome, unter denen sich die Atome nach den Gesetzen der klassischen Mechanik bewegen. Dadurch ist es heute möglich, in einem eingeschränkten Umfang (für Einheitszellen bis zu etwa 300 Atomen auf einer Zeitskala bis etwa 10 ps) *ab-initio*-Molekulardynamik-Simulationen für reale Systeme mit rein quantenmechanisch bestimmten Wechselwirkungen durchzuführen, d.h. ohne Rückgriff auf empirische oder gefittete Vielteilchenwechselwirkungspotentiale. Die großen Stärken der Methode und die vielleicht am weitesten verbreiteten Anwendungen liegen (i) in der dynamischen Strukturoptimierung komplexer Systeme, (ii) in der Suche nach metastabilen Zuständen wie im Falle von Molekularstatikrechnungen zur Bestimmung von Aktivierungsbarrieren der atomaren Diffusion oder der Reaktionskinetik und (iii) der Bestimmung von Grundzuständen von Systemen durch dynamisches “simulated annealing”¹.

Ein solches *ab initio* Molekulardynamikprogramm, genannt **ESCoMPP** (Electronic Structure Code for Materials Properties and Processes) [BBA⁺00], ist in unserer Arbeitsgruppe im Rahmen früherer Diplom- und Doktorarbeiten [Eng95, Ric96, Ber] speziell für den Anwendungsbereich Volumen-, Oberflächen- und Grenzflächenprobleme der Festkörperforschung vollständig neu erstellt worden. Es ist ein massiv-paralleles *ab-initio*-Molekulardynamikprogramm, das für große Einheitszellen, gedacht ist.

Formuliert wurde die *ab-initio*-Molekulardynamik-Methode mittels der Pseudopotentialtheorie normerhaltender, nichtlokaler Pseudopotentiale [KB82]. Dies ist eine Elektronentheorie-Methode, die es erlaubt, die Wellenfunktionen, die Ladungsdichte und das Potential nach den einfachsten nur denkbaren Basisfunktionen zu entwickeln, nämlich nach Ebenen Wellen. Die Ausnutzung der diskreten Fast-Fourier-Transformation (FFT) erlaubt eine konzeptionell einfache und numerisch schnelle Berechnung der Schrödingergleichung, der Poissonsgleichung und der quantenmechanischen Kräfte, die auf die Atome wirken und unter deren Wirkung die Atome sich bewegen.

¹Kirkpatrick *et al.* [KGV83] hat den Begriff “simulated annealing” im Falle von Monte Carlo vorgeschlagen.

Auf Grund der Verwendung von normerhaltenden Pseudopotentialen ist man in der Auswahl der Systeme allerdings stark eingeschränkt. Sie sind geeignet zur effizienten Beschreibung von Elementhalbleitern (Si, Ge) oder III-V Halbleitern (z.B. GaAs, InAs), einfachen Metallen (z.B. Mg, Al, Ga) oder Alkalimetallen (Li, Na, K, usw.). Die Methode normerhaltender Pseudopotentiale ist allerdings weniger geeignet für Systeme, welche Elemente der zweiten Periode enthalten, wie z.B. Kohlenstoff in Nanoröhrchen, Stickstoff in Halbleiterverbindungen oder Sauerstoff in keramischen Werkstoffen. Ähnliche Schwierigkeiten bereiten auch Übergangsmetalle, speziell die $3d$ -Metalle, aber auch Verbindungen, in denen die $3d$ -Elektronen wesentlich an Valenzbindungen beteiligt sind, wie das z.B. mit den $3d$ -Zuständen der Ga Atome in GaN oder der Mn Atome in MnGaN oder MnGaAs der Fall ist. Generell stellen Oxide, Karbide, Nitride, Fluoride, Edelmetalle, Übergangsmetalle, Lanthanide, Aktinide ein – teilweise erhebliches – Problem dar. Die Basis aus Ebenen Wellen muß bei solchen Elementen eine exorbitante Anzahl von Funktionen umfassen, so daß diese Rechnungen sehr viel Computerzeit verbrauchen; in vielen Fällen ist die Berechnung dadurch praktisch unmöglich.

In jüngster Zeit gab es mehrere erfolgreiche Entwicklungen, um diese Einschränkungen in der Systemauswahl zu überwinden. Zum einen kann man versuchen, die Effizienz bestehender Pseudopotentialprogramme zu verbessern. In diesem Sinne wurden zu Beginn dieser Arbeit einige Algorithmische Veränderungen implementiert. Jedoch bleibt als Grundproblem, daß schon ein einzelnes Atom, zu dessen Darstellung viele Basisfunktionen benötigt werden, die Anzahl der insgesamt erforderlichen Ebenen Wellen für ein großes System auf einen nicht mehr zu bewältigenden Wert steigert. Zum anderen sind heute auch Allelektronenmethoden wie die *Full-Potential Linearized Augmented Plane-Wave* (FLAPW)-Methode [Sin94] in der Lage, effizient die Kräfte [SW90, YSK91] auf die Atome zu berechnen und die Atome zu bewegen. Auch wurden neuartige Methoden entwickelt: z.B. wurde die Pseudopotentialmethode zu einer Ultrasoften Pseudopotentialmethode [Van90] weiterentwickelt. Eine der interessantesten Entwicklungen in diesem Kontext ist die von P. Blöchl 1994 entwickelte *Projector Augmented Wave* (PAW)-Methode. Sie kann sowohl als Weiterentwicklung der FLAPW-Methode als auch der Pseudopotentialmethode aufgefaßt werden. Wie die FLAPW-Methode ist sie im Kern eine Allelektronenmethode.

In dieser Arbeit wurde – auf der Grundlage des vorhandenen Pseudopotentialprogramms EStCoMPP – eine modifizierte Variante der PAW-Methode für den Anwendungsbereich großer Systeme der Festkörperforschung mit Oberflächen und Grenzflächen implementiert, welche insbesondere die volle Kompatibilität mit den vorhan-

denen Pseudopotentialen wahr.

In der PAW-Methode werden zwei Raumbereiche unterschieden: während weit entfernt von den Atomen die Wellenfunktionen allein durch Ebene Wellen beschrieben werden, werden sie in der Nähe der Atome durch lokalisierte Funktionen modifiziert oder *augmentiert*. Der Anteil, der einer atomaren *Pseudowellenfunktion* entspricht, wird aus der Festkörperwellenfunktion entfernt und durch eine *echte* atomare Wellenfunktion ersetzt. Der zu ersetzende Anteil wird durch die Projektion auf die atomare *Pseudowellenfunktion* bestimmt. Die echten atomaren Wellenfunktionen, die daraus durch Unterdrückung von Knoten und scharfen Strukturen bestimmten Pseudowellenfunktionen und die Projektorfunktionen charakterisieren ein Atom einer bestimmten Sorte. Sie werden vor der eigentlichen Festkörperrechnung erzeugt. Im Hamiltonoperator machen sich die Projektorfunktionen in einer Weise bemerkbar, die man als Verallgemeinerung nichtlokaler Pseudopotentiale auffassen kann, welche jedoch von der Dichte abhängen. Auch hat die Augmentation der Wellenfunktionen in Atomnähe zur Folge, daß die Orthogonalität der Basisfunktionen, wie sie bei Ebenen Wellen trivialerweise gegeben ist, verloren geht. Das resultierende verallgemeinerte Eigenwertproblem wird mit der Verallgemeinerung einer iterativen Methode gelöst, die auf W. Davidson [Dav75] zurückgeht. Die größte Hürde jedoch ist die Lösung der Poissongleichung, da zur Bestimmung des Hartree-Potentials die Ladung der *augmentierten* Wellenfunktionen berücksichtigt werden muß. Diese Schwierigkeit wird unter Rückgriff auf eine Pseudoladungsmethode von M. Weinert [Wei81] überwunden, die in leicht veränderter Form auch im Bereich der FLAPW-Methode benutzt wird.

Mit der PAW-Methode steht ein Verfahren zur Verfügung, das eine effiziente Berechnung auch für Systeme erlaubt, welche Elemente enthalten, die in der Pseudopotentialtheorie nur mit hohem Aufwand benutzt werden können; z.B. kann man für SiC-Systeme eine Beschleunigung um einen Faktor fünf gegenüber der Verwendung normerhaltender Pseudopotentiale erreichen, und im Falle einer (4×2) Cu(110) Oberfläche ist sie drei- bis viermal schneller als eine FLAPW-Methode, wobei Effizienzreserven noch vorhanden sind.

Die Anzahl der benötigten Basisfunktionen wird gegenüber einem normerhaltenden Pseudopotential drastisch reduziert und erlaubt so auch die Berechnung großer Systeme, wie es beispielsweise für die Untersuchung von Oberflächen erforderlich ist. War die Pseudopotentialtheorie in der Praxis auf Systeme aus bestimmten Elementen beschränkt, so erlaubt die PAW-Methode nun effiziente Molekulardynamikrechnungen für ein weitaus größeres Spektrum von physikalischen Systemen, welche

nahezu beliebige Elemente enthalten können.

Die Arbeit gliedert sich wie folgt:

Zunächst werden in Kapitel 2 die Grundzüge der Dichtefunktionaltheorie und der Pseudopotentialmethode dargestellt. Die Dichtefunktionaltheorie ermöglicht es, das Lösen des komplizierten Vielteilchensystems der Elektronen auf ein Einteilchenproblem zu reduzieren, dessen effektives Potential allerdings selbstkonsistent bestimmt werden muß. Die Pseudopotentialtheorie eliminiert die core-Elektronen und die Knoten der Valenzwellenfunktionen, welche für die Bindungseigenschaften im Festkörper unwichtig sind. Dadurch ermöglicht sie eine Beschreibung der Wellenfunktionen durch relativ wenige Ebene Wellen.

In Kapitel 3 wird die PAW-Methode beschrieben. Sie beseitigt wie die Pseudopotentialtheorie die core-Elektronen als Freiheitsgrade, berücksichtigt deren Ladungsdichte bei der Berechnung des Potentials im Gegensatz zur Pseudopotentialtheorie jedoch vollständig. Dieses Kapitel erörtert einerseits die Methode; andererseits werden aber auch die Implementation und die dazu benötigten Formeln im Detail dargelegt. Zunächst wird das Funktional der Energie in drei Anteile aufgespalten, welche auf unterschiedliche Weisen numerisch vorteilhaft berechnet werden können. Diese Aufteilung der Energie bildet auch den Ausgangspunkt für die Ableitung des Einteilchen-Hamiltonoperators; die einzelnen Bestandteile müssen in derselben Weise berechnet werden wie in der Energie, damit die numerische Konsistenz gegeben ist. Um die dabei schon eingeflossene Pseudoladungsmethode entsprechend ihrer Bedeutung bei der Lösung der Poissongleichung hervorzuheben, wird sie in einem eigenen Abschnitt zusammenfassend erläutert. Ausführlich wird die Schwierigkeit erläutert, die sich durch das Aufteilen der Energie in mehrere für sich divergierende Anteile mit sich bringt. Schließlich wird die Berechnung der Kräfte erörtert, welche die Benutzung der PAW-Methode zur Molekulardynamik ermöglicht.

In der PAW-Methode wird – ähnlich wie bei der Pseudopotentialmethode – für jeden verwendeten Atomtyp ein Datensatz benötigt, welcher diesen beschreibt. Die Erzeugung dieser Datensätze, die in analoger Namensgebung *Pseudopotentiale* genannt werden, wird in Kapitel 4 erläutert. Ein zentralen Punkt hierbei ist die Konstruktion der Projektorfunktionen, denen die Methode ihren Namen verdankt.

Während in Kapitel drei beschrieben wird, wie die Bestandteile der effektiven Einteilchengleichungen *aufgestellt* werden, befaßt sich Kapitel 5 damit, wie man diese Gleichungen löst. Dort wird zum einen dargestellt, wie die iterative Minimierung bezüglich der ionischen und der elektronischen Freiheitsgrade erfolgt. Zum anderen beschreibt dieses Kapitel auch die iterative Lösung des Eigenwertproblems, das

die Einteilchengleichungen für eine feste Dichte darstellen. Auch werden in diesem Kapitel drei Methoden geschildert, die auf ihr Potential zur Beschleunigung der Konvergenz untersucht wurden. Während die Präkonditionierung und das Jacobi-Davidson-Verfahren ausgesprochen technischer Natur sind, ist die Bestimmung der Startvektoren für die Iteration rein physikalisch motiviert. Analog zu einer LCAO-Methode versucht man, durch eine Linearkombination von atomaren Pseudowellenfunktionen – auch atomare Pseudoorbitale (APSO) genannt – schon zu Beginn der iterativen Diagonalisierung möglichst nahe an die Lösung der Gleichungen zu gelangen.

Kapitel 6 beschreibt erste Tests, welche mit der neuen Implementierung der PAW-Methode durchgeführt wurden. Einerseits soll durch den Vergleich mit bekannten Ergebnissen die Richtigkeit und Zuverlässigkeit des Programmes erprobt werden, andererseits wurden auch Versuche angestellt, um die Geschwindigkeit in Relation zu anderen *ab-initio*-Programmen zu stellen.

Kapitel 7 befaßt sich mit Volumeneigenschaften von Siliziumkarbid, Aluminiumnitrid und Galliumnitrid. Wie zu Anfang geschildert, stellte der Wunsch, auch diese Materialien effizient beschreiben zu können, den Ausgangspunkt dar für die Entwicklung des PAW-Programmes in dieser Arbeit. In diesem Kapitel werden einerseits Ergebnisse präsentiert, welche (vor Vollendung des PAW-Programmes) mit normerhaltenden Pseudopotentialen bestimmt wurden; andererseits bietet dies auch wiederum die Möglichkeit zum direkten Vergleich mit den Resultaten der neuen Methode.

In Kapitel 8 der Arbeit werden die Resultate dieser Arbeit zusammengefaßt.

Kapitel 2

Dichtefunktionaltheorie und Pseudopotentialmethode

Die rasante Entwicklung der Informationstechnologie beruht auf der ständig fortschreitenden Miniaturisierung der verwendeten Bauelemente und der ständigen Verbesserung der dazu nötigen Herstellungsvorgänge. Daher erfreuen sich insbesondere Oberflächen und Grenzflächen regen Interesses.

Einen wertvollen Beitrag zum Verständnis von Oberflächen und Grenzflächen können *ab-initio*-Rechnungen leisten. In solche Rechnungen gehen nur die Kernladungszahlen der Atome und die Regeln der Quantenmechanik ein. Mit dieser Methode können viele physikalische Eigenschaften des betrachteten Festkörpers berechnet werden, z.B. strukturelle Eigenschaften wie Gitterkonstanten und Oberflächenrekonstruktionen oder elektronische Eigenschaften wie Bandstrukturen, elektronische Ladungsdichten, Bindungen usw. Da diese Rechnungen unabhängig von experimentellen Voraussetzungen und Einschränkungen sind, bieten sie vielfach einen zweiten Zugang zum Verständnis von Materialeigenschaften. In ihnen kann man beispielsweise – anders als im Experiment – durch Variation der Parameter auch hypothetische Substanzen untersuchen, und man kann unter Umständen daraus Erkenntnisse über reale Systeme gewinnen. Auch lassen sich experimentell schwer erreichbare Bedingungen verwirklichen – z.B. extrem hohe Drucke oder absolute Abwesenheit von Verunreinigungen.

Jedoch ist die exakte Lösung der Vielteilchen-Schrödingergleichung praktisch unmöglich. Einen Durchbruch für die *ab-initio*-Beschreibung von Vielteilchensystemen stellte die Entwicklung der Dichtefunktionaltheorie [JG89] dar. Deren Grundlagen gehen auf Arbeiten von Hohenberg und Kohn [HK64] sowie Kohn und Sham [KS65]

zurück.

Sie zeigen, daß es möglich ist, statt der Vielteilchenwellenfunktion die Teilchendichte als grundlegende Größe zu benutzen. Die zentrale Aussage ist, daß sich die Gesamtenergie eines Vielelektronensystemes als Funktional der Elektronendichte darstellen läßt, und daß die Grundzustandsenergie das Minimum dieses Funktionals bezüglich der Variation der Dichte ist. Die Variation erfolgt bezüglich aller Elektronendichten, die mit der Antisymmetrie der Vielteilchenwellenfunktionen für Fermiteilchen verträglich sind. Aus der Grundzustandsenergie und der dazugehörigen Elektronendichte kann man alle weiteren Grundzustandseigenschaften des Systems berechnen.

2.0 Einheiten

Zunächst soll kurz das gewählte Einheitensystem erwähnt werden. In dieser Arbeit werden – genauso wie im verwendeten Computerprogramm – atomare Einheiten (a.u.) benutzt. Dadurch nehmen die wichtigsten Naturkonstanten sehr einfache Zahlenwerte an und die Größenskala ist den betrachteten Problemen besser angemessen als SI-Einheiten.

Plancksches Wirkungsquantum	$\hbar = 1$
Elektronenmasse	$m_e = \frac{1}{2}$
Lichtgeschwindigkeit	$c = \frac{2}{\alpha} = 274.074$
Rydbergeinheit	$1\text{Ry} \cong 13.6058 \text{ eV}$
Bohrscher Radius	$a_B \cong 0.529177 \text{ \AA} = 0.529177 \times 10^{-10} \text{ m}$

2.1 Die Dichte als fundamentale Systemvariable

Ein System N wechselwirkender Elektronen, die sich in einem externen Potential, z.B. einem Gitter positiv geladener Atomrümpfe, bewegen, wird durch den folgenden Hamiltonoperator H beschrieben:

$$H = T + V_{ee} + \sum_{i=1}^N V_{ext}(X_i). \quad (2.1)$$

Der Hamiltonian setzt sich zusammen aus kinetischer Energie

$$T = - \sum_{i=1}^N \partial_{\vec{r}_i}^2, \quad (2.2)$$

externem Potential $V_{ext}(X_i)$ und der Elektron–Elektron–Wechselwirkung

$$V_{ee} = \sum_{i \neq j}^N \frac{1}{|\vec{r}_i - \vec{r}_j|}. \quad (2.3)$$

Man hat also in Ortsraumdarstellung eine $3N$ -dimensionale Eigenwertgleichung zu lösen, wobei die Normierung und das Pauli-Prinzip als Randbedingungen zu berücksichtigen sind. Das Pauli-Prinzip besagt, daß die Vielteilchen-Wellenfunktion

$$\Psi(X_1, X_2, \dots, X_N) \text{ mit } X_i = (\vec{r}_i, \sigma_i) = (\text{Orts-, Spinkoordinate})$$

vollständig antisymmetrisch unter der Vertauschung zweier Teilchenkoordinaten ist, d.h. sie erfüllt die Beziehung

$$\Psi(X_1, \dots, X_i, \dots, X_j, \dots, X_N) = -\Psi(X_1, \dots, X_j, \dots, X_i, \dots, X_N). \quad (2.4)$$

Dies ist die Konsequenz der quantenmechanischen Ununterscheidbarkeit der Elektronen (sowie ihres fermionischen Charakters).

Der Zweiteilchenoperator V_{ee} der Coloumb-Wechselwirkung macht eine Entkoppelung der Teilchen unmöglich. Daher ist dieses Problem schon für wenige Teilchen so komplex, daß es weder analytisch noch numerisch zu lösen ist.

Die Dichtefunktionalmethode verwendet anstatt der Vielteilchenwellenfunktion $\Psi(X_1, X_2, \dots, X_N)$ die Elektronendichte

$$n(\vec{r}) := \langle \Psi | \sum_{i=1}^N \delta(\vec{r} - \vec{r}_i) | \Psi \rangle = \int \prod_{i=1}^N dX_i \sum_{i=1}^N \delta(\vec{r} - \vec{r}_i) |\Psi(X_1, \dots, X_N)|^2 \quad (2.5)$$

als fundamentale Systemvariable.

Hohenberg und Kohn [HK64] zeigten 1964 für den nichtentarteten Grundzustand:

- Die Grundzustandsenergie E_0 und die Erwartungswerte sämtlicher Observablen sind eindeutige Funktionale der Grundzustandsdichte $n_0(\vec{r})$.
- Es existiert ein Energiefunktional $E[n]$, welches ein Variationsprinzip erfüllt und unter der Nebenbedingung der Ladungserhaltung ($N = \int d^3r n(\vec{r})$) durch die Grundzustandsdichte $n_0(\vec{r})$ minimiert wird:

$$E[n] > E[n_0] = E_0 \text{ für alle } n(\vec{r}) \neq n_0(\vec{r}). \quad (2.6)$$

Man kann also die Grundzustandsdichte $n_0(\vec{r})$ durch Minimierung des Energiefunktionals $E[n]$ bestimmen.

Die Verallgemeinerung des Dichtefunktional–Formalismus’ auf entartete Grundzustände ist ebenfalls möglich [Koh85, DG90].

Im ursprünglichen Beweis von Hohenberg und Kohn wird nur die Existenz eines solchen Energiefunktional geliefert. Levy [Lev79] dagegen konnte einen konstruktiven Beweis unter Zuhilfenahme eines Variationsverfahrens angeben. Levy definiert das Energiefunktional $E[n]$ wie folgt:

$$E[n] := \min_{\{\Psi\}} \langle \Psi | \mathcal{H} | \Psi \rangle \quad (2.7)$$

Der zugrundeliegende Funktionenraum ist dabei beschränkt auf diejenigen Vielteilchenwellenfunktionen Ψ , welche die vorgegebene Dichte $n(\vec{r})$ reproduzieren (sowie normiert und vollständig antisymmetrisch sind).

Dieser Raum ist sinnvoll definiert, denn Lieb [Lie83] konnte zeigen, daß zu physikalisch vernünftigen Dichten antisymmetrische Vielteilchenwellenfunktionen Ψ existieren, die diese Dichte reproduzieren.

Die Vielteilchen–Eigenwertgleichung wurde also durch ein Minimierungsproblem bzgl. der Dichte ersetzt.

$$\delta E[n] = 0. \quad (2.8)$$

Dabei ist noch die Teilchenzahlerhaltung als Nebenbedingung zu berücksichtigen:

$$\int d^3r n(\vec{r}) = N \quad (2.9)$$

Dies führt auf ein Optimierungsproblem des modifizierten Energiefunktional

$$\tilde{E}[n] = E[n] - \mu \left(\int d^3r n(\vec{r}) - N \right), \quad (2.10)$$

und damit zu folgender Euler–Lagrange–Gleichung:

$$\left. \frac{\delta E[n(\vec{r})]}{\delta n(\vec{r})} \right|_{n_0(\vec{r})} = \mu. \quad (2.11)$$

Als Wermutstropfen bleibt, daß bislang keine explizite Darstellung des Energiefunktional aus der Definition nach Levy gefunden werden konnte.

2.2 Die Kohn–Sham Gleichungen

Da die explizite Form des Funktional $E[n]$ unbekannt ist, ist man auf eine Näherung angewiesen. Kohn und Sham [KS65] führten ein Hilfssystem von nicht wechselwirkenden Elektronen in einem effektiven Potential $V_{eff}(\vec{r})$ ein. Dazu wird das

Energiefunktional in vier Beiträge aufgespalten:

$$E[n] = T_0[n] + E_{ext}[n] + E_H[n] + E_{xc}[n]. \quad (2.12)$$

$T_0[n]$ ist das Funktional der kinetischen Energie *nicht wechselwirkender* Elektronen. Die potentielle Energie $E_{ext}[n]$ beschreibt die Wechselwirkung der Elektronen mit einem äußeren Potential (z.B. die Wechselwirkung mit den Atomrümpfen):

$$E_{ext}[n] = \int d^3r n(\vec{r}) V_{ext}(\vec{r}). \quad (2.13)$$

Die Coulomb–Wechselwirkungsenergie der Elektronen in Hartree–Näherung ist durch

$$E_H[n] = \int d^3r d^3r' \frac{n(\vec{r})n(\vec{r}')}{|\vec{r} - \vec{r}'|} \quad (2.14)$$

gegeben.

Das Austausch–Korrelations–Funktional $E_{xc}[n]$ enthält alle übrigen Energieanteile, nämlich die Austausch– und Korrelationsbeiträge, Korrekturen zur kinetischen Energie infolge der Wechselwirkung der Elektronen und eine Korrektur der in $E_H[n]$ enthaltenen Selbstwechselwirkung. Das Austausch–Korrelations–Funktional ist im Prinzip unbekannt. Jedoch haben sich sehr einfache Näherungen als überaus nützlich erwiesen. Letztlich basiert darauf der Erfolg der Dichtefunktionaltheorie.

Um die Kohn–Sham–Gleichungen zu erhalten, macht man nun einen Ansatz für die Dichte $n(\vec{r})$ in der Form nicht wechselwirkender Elektronen, d.h. als Summe des Absolutquadrates von normierten Einteilchenwellenfunktionen, gewichtet mit ihren Besetzungszahlen f_i .

$$n(\vec{r}) = 2 \sum_i f_i |\psi_i(\vec{r})|^2. \quad (2.15)$$

Summiert werden alle besetzten Einteilchenfunktionen $\psi_i(\vec{r})$. Die Besetzungszahlen ($0 < f_i \leq 1$) genügen einer Fermiverteilung und sind entsprechend der Teilchenzahl N normiert:

$$N = 2 \sum_i f_i \quad (2.16)$$

Der Faktor zwei berücksichtigt die Spinentartung.

Für derartige Ladungsdichten kann man die kinetische Energie wechselwirkungsfreier Elektronen exakt berechnen:

$$T_0[n] = -2 \sum_i f_i \int d^3r \psi_i^*(\vec{r}) \nabla_{\vec{r}}^2 \psi_i(\vec{r}). \quad (2.17)$$

Zur Bestimmung des Minimums von $E[n]$ führt man die Variation statt nach der Dichte $n(\vec{r})$ nach den Einteilchenwellenfunktionen aus. Die als Nebenbedingung geforderte Ladungserhaltung wird in diesem Fall durch die Normierung der Wellenfunktionen

$$\int d^3r |\psi_i(\vec{r})|^2 = 1 \quad (2.18)$$

gewährleistet, welche wiederum durch Einführung von Lagrange-Parametern ε_i im Energiefunktional garantiert wird. Unter Beachtung von

$$\frac{\delta E[n(\vec{r})]}{\delta \psi_i^*(\vec{r})} = \frac{\delta E[n(\vec{r})]}{\delta n(\vec{r})} \frac{\delta n[\dots, \psi_i, \dots]}{\delta \psi_i^*(\vec{r})} \quad (2.19)$$

erhält man die Kohn-Sham Gleichungen:

$$\frac{1}{2f_i} \frac{\delta E[n(\vec{r})]}{\delta \psi_i^*(\vec{r})} = \left(-\partial_{\vec{r}}^2 + V_{eff}[n(\vec{r})] \right) \psi_i(\vec{r}) = \varepsilon_i \psi_i(\vec{r}) \quad (2.20)$$

bzw.

$$\left(-\partial_{\vec{r}}^2 + V_{ext}(\vec{r}) + V_H[n(\vec{r})] + V_{xc}[n(\vec{r})] \right) \psi_i(\vec{r}) = \varepsilon_i \psi_i(\vec{r}) \quad (2.21)$$

Der Form nach entsprechen diese Gleichungen der stationären Schrödingergleichung im wechselwirkungsfreien Fall. Jedoch enthält der Hamiltonoperator hier ein *effektives* Potential, welches sich aus drei Beiträgen zusammensetzt: dem externen Potential V_{ext} , dem Hartree-Potential

$$V_H[n(\vec{r})] = \frac{\delta}{\delta n(\vec{r})} \left(\int d^3r' d^3r'' \frac{n(\vec{r}') n(\vec{r}'')}{|\vec{r}' - \vec{r}''|} \right) \quad (2.22)$$

$$= 2 \int d^3r' \frac{n(\vec{r}')}{|\vec{r} - \vec{r}'|} \quad (2.23)$$

und dem Austausch-Korrelations-Potential

$$V_{xc}[n(\vec{r})] = \frac{\delta E_{xc}[n(\vec{r})]}{\delta n(\vec{r})} . \quad (2.24)$$

Das effektive Potential ist also abhängig von der Dichte $n(\vec{r})$ und damit wiederum von den Einteilchenwellenfunktionen $\psi_i(\vec{r})$. Es ergibt sich also ein Gleichungssystem, welches selbstkonsistent gelöst werden muß. Dies legt es nahe, ein iteratives Verfahren zu benutzen.

Formal handelt es sich bei den Eigenwerten der Kohn-Sham-Gleichungen (2.21) um Einteilchenenergien. Jedoch haben sie zunächst als Lagrange-Parameter keine physikalische Bedeutung. Den obersten besetzten Eigenwert kann man zumindest für

delokalisierte Zustände als Ionisierungsenergie bzw. Austrittsarbeit interpretieren [JG89]. Die Erfahrung zeigt allerdings, daß sich auch die übrigen Eigenwerte sinnvoll verwenden lassen. Beispielweise geben so berechnete Bandstrukturen die Form der Bänder i. Allg. gut wieder. Allerdings sind die Absolutwerte weniger zuverlässig; insbesondere wird die Bandlücke von Halbleitern meist um einen Faktor zwei oder mehr zu klein berechnet. Auch die Interpretation der Einteilchenwellenfunktionen $\psi_i(\vec{r})$, welche eigentlich mathematische Hilfskonstrukte sind, als physikalische Wellenfunktionen liefert vielfach sinnvolle Ergebnisse. Jedoch ist dies nicht notwendig so und entsprechend ist Vorsicht bei entsprechender Deutung geboten.

2.3 Die Spindichtefunktionaltheorie

Für die Untersuchung magnetischer Probleme ist eine Verallgemeinerung der Dichtefunktionaltheorie auf den Fall spinpolarisierter Elektronen möglich. Dies ist immer dann erforderlich, wenn der Grundzustand magnetisch wird. Dieser Fall tritt z.B. beim Ferro- oder Antiferromagnetismus von Festkörpern oder beim Vorliegen externer Magnetfelder auf.

Die Verallgemeinerung besteht darin, daß anstelle der Ladungsdichte $n(\vec{r})$ die Spindichten $n^\uparrow(\vec{r})$ und $n^\downarrow(\vec{r})$ als fundamentale Systemvariablen eingeführt werden.

Analog zum oben Gesagten ergibt sich ein Spindichtefunktionaltheorem [vBH72]. Das effektive Potential ist dann spinabhängig. Eine ausführliche Herleitung findet man in [KV82].

2.4 Die Lokale-Dichte-Näherung

Wie oben schon erwähnt wurde, ist das Austausch-Korrelations-Funktional $E_{xc}[n]$ unbekannt. Hier ist man daher auf Näherungen angewiesen. Davon abgesehen ist die Dichtefunktionaltheorie exakt.

Die einfachste, aber dennoch überaus erfolgreiche Näherung ist die Lokale-Dichte-Näherung (LDA). Für ein homogenes Elektronengas ist die Dichte $\epsilon_{xc}[n(\vec{r})]$ der Austausch-Korrelations-Energie pro Teilchen bekannt. Die LDA beruht auf der Annahme, daß für eine räumlich nicht zu stark variierende Elektronendichte eine (imaginäre) Zerlegung des Gesamtvolumens in voneinander unabhängige Bereiche konstanter Dichte möglich ist. Für diese Bereiche läßt sich dann jeweils die Energiedichte des homogenen Elektronengases mit der jeweiligen lokalen Dichte verwenden.

Das Austausch–Korrelations–Funktional wird also approximiert durch

$$E_{xc}^{LDA}[n] = \int d^3r n(\vec{r}) \epsilon_{xc}(n(\vec{r})) \quad . \quad (2.25)$$

Das möglicherweise nichtlokale Funktional wird also durch ein lokales ersetzt; $\epsilon_{xc}(n(\vec{r}))$ hängt nur von der Dichte am Ort $\vec{r}$ ab. Das entsprechende Austausch–Korrelations–Potential erhält man durch Variation nach der Dichte:

$$V_{xc}^{LDA}[n] = \epsilon_{xc}(n(\vec{r})) + n(\vec{r}) \frac{\partial \epsilon_{xc}(n(\vec{r}))}{\partial n} \quad (2.26)$$

Analytische Ergebnisse für die Austausch–Korrelations–Energie eines homogenen Elektronengases existieren nur für die Grenzfälle hoher und niedriger Dichte. Numerische Resultate von Ceperley und Alder [CA80] für die Korrelationsenergiedichte basieren auf einer Quanten–Monte–Carlo–Simulation. Auf diesen Ergebnissen beruht die Parametrisierung von Vosko, Wilk, Nusair [VWN80]. Eine relativistische Korrektur lieferten MacDonald und Vosko [MV79].

Die Ergebnisse, welche man mit diesen Funktionalen erzielt, unterscheiden sich nur wenig von jenen, die man mit anderen (etwa nach Moruzzi, Janak und Williams [MJW78]) erhält. Allerdings ist ein direkter Vergleich von Ergebnissen nur dann zulässig, wenn sie mit demselben Austausch–Korrelations–Funktional erzielt wurden.

Von ihrer Idee her ist die LDA auf Systeme beschränkt, in denen die Ladungsdichte räumlich nicht stark variiert. In der Praxis jedoch erhält man auch für inhomogene Systeme, wie sie etwa Halbleiteroberflächen bilden, gute Ergebnisse. Dies liegt darin begründet, daß zu einer exakten Berechnung der Austausch–Korrelations–Energie nur das sphärische Mittel des Austausch–Korrelations–Loches korrekt wiedergegeben werden muß. Daraus resultiert eine relative Unempfindlichkeit gegen die Details der Näherung [JG89].

Für magnetische Systeme läßt sich analog eine Lokale–Spindichte–Näherung (LSDA) erfolgreich anwenden.

Einen Schritt über die LDA hinaus geht die Generalized Gradient Approximation (GGA), welche neben der lokalen Dichte auch deren Gradienten berücksichtigt. Dadurch erzielt man insbesondere in stark inhomogenen Systemen wie Molekülen deutlich bessere Resultate [OEK99].

2.5 Die Pseudopotential-Methode

Innerhalb der DFT haben sich eine Reihe verschiedener Methoden etabliert, wie z.B. *Linearized Augmented Plane Wave* (LAPW), Korringa-Kohn-Rostoker (KKR) und die Pseudopotentialmethode (PP) (siehe z.B. [OEK99]).

Während etwa bei der LCAO-Methode Linearkombinationen von atomaren Zuständen als Basisfunktionen dienen, kann man auch eine Basis aus Ebenen Wellen wählen. Wenn man diese in naiver Weise für alle Elektronenzustände benutzen wollte, brauchte man wegen der starken Variation der core-Zustände in Kernnähe eine in der Praxis nicht zu bewältigende Anzahl von Ebenen Wellen. Aber auch die Beschreibung der Valenzfunktionen durch Ebene Wellen führt wegen der Oszillationen durch die Orthogonalität zu den core-Zuständen auf das gleiche Problem. Abhilfe schaffte die Orthogonalized Plane Wave (OPW)-Methode, welche als Vorläufer der PP-Methode aufgefaßt werden kann. Sie verwendet zu den core-Zuständen orthogonalisierte Ebene Wellen als Basis für die Valenzwellenfunktionen. Da diese orthogonal zu den core-Zuständen sind, lassen sie sich gut durch relativ wenige OPW's beschreiben. Die Hamiltonmatrix erfordert hierbei jedoch die Auswertung von Mischtermen (aus Ebenen Wellen und core-Funktionen).

Die Pseudopotentialmethode greift die Idee auf, nur Valenzelektronen zu beschreiben. Dem liegt die Erfahrung zugrunde, daß viele physikalische Eigenschaften eines Festkörpers nicht von den core-Zuständen abhängen, sondern ausschließlich durch die Valenzelektronen bestimmt werden. Daher geht man davon aus, daß die core-Ladungsdichte nicht von der chemischen Umgebung beeinflusst wird, und ersetzt das tiefe Potential von Kern- und core-Ladung durch ein flacheres und weiches *Pseudopotential*. Da durch das Eliminieren der core-Zustände keine Orthogonalitätsoszillationen erforderlich sind, lassen sich die zugehörigen Pseudo-Valenzfunktionen gut durch Ebene Wellen beschreiben.

In den 1970-er Jahren wurden vielfach empirische Modellpotentiale verwendet, deren Parameter angepaßt wurden, um die experimentell gefundenen Daten von Testsystemen möglichst gut wiederzugeben. Ein Durchbruch für parameterfreie sogenannte „*ab-initio*“-Rechnungen gelang Bachelet, Chiang, Hamann und Schlüter [HSC79, BHS82] mit der Entwicklung normerhaltender *ab-initio*-Pseudopotentiale. Die Normerhaltung ermöglichte überhaupt erst eine Verwendung im Rahmen der DFT. Für eine Referenzenergie (nämlich einen Valenzeigenwert) wird die logarithmische Ableitung der Wellenfunktion, welche allein die Streueigenschaften eines Atomes festlegt, korrekt wiedergegeben, und durch die Normerhaltung ist dies auch noch in erster

Ordnung in der Umgebung der Referenzenergie richtig.

Die Berechnung der Wellenfunktionen in einer Basis Ebener Wellen [IZC79] in Kombination mit der von Car und Parrinello [CP85] vorgeschlagenen Benutzung von iterativen Verfahren zur Matrixdiagonalisierung und die dadurch ermöglichte Beschleunigung des Verfahrens durch die Fast Fourier Transform (FFT) haben die Pseudopotentialmethode als ein wichtiges Verfahren für die Berechnung von elektronischen und strukturellen Eigenschaften etabliert.

Die von Kleinman und Bylander [KB82] präsentierte Möglichkeit, diese semilokalen Potentiale in separable Form zu bringen, steigerte die Effizienz von PP-PW-Programmen erheblich. Dabei traten zuweilen Probleme mit *ghost-states* auf, welche jedoch vielfach beseitigt werden können [BK90, GKS90, GSS91, SSO92].

Hamann zeigte, daß man auch andere Energien außer den Eigenwerten zur Referenzenergie wählen, d.h. Pseudopotentiale aus ungebundenen Zuständen konstruieren kann [Ham89].

Die Konstruktion eines Pseudopotentials erfolgt i. Allg. in zwei Schritten: zunächst wird zu jeder atomaren Valenzwellenfunktion eine Pseudowellenfunktion konstruiert, welche außerhalb eines gewissen Abstandes vom Atom mit der echten Wellenfunktion übereinstimmt, innerhalb aber von dieser derart abweicht, daß sie trotz gleicher Normierung erheblich *glatter* (oder *weicher*) ist, d.h. schnell mit der Größe einer Basis aus Ebenen Wellen konvergiert. Im zweiten Schritt wird dann durch Invertierung der Schrödingergleichung dasjenige Potential bestimmt, welches diese Pseudowellenfunktion als Eigenzustand zur zugehörigen Valenzenergie besitzt. Dieses Potential ist das gesuchte Pseudopotential. Neben der Weichheit muß das Pseudopotential gut *transferierbar* sein, d.h. in allen chemischen Umgebungen die Eigenschaften des Atomes möglichst gut wiedergeben. Ein Maß dafür bieten die logarithmischen Ableitungen der Wellenfunktionen (s.u.). Gängige Verfahren konstruieren die Pseudopotentiale so, daß diese Größe für die Energie des jeweiligen Zustandes, welcher zur Konstruktion benutzt wird, korrekt wiedergegeben wird.

Es wurden viele Methoden zur Konstruktion glatter Pseudopotentiale entwickelt [Ker80, Van85, RRKJ90, TM90, TM91, VS92]. Dabei kristallisierten sich als begrenzende Faktoren einerseits für die Konstruktion glatterer Potentiale die zu erfüllende Normerhaltung, andererseits für die Erhöhung der Präzision die Beschränkung auf *eine* Referenzenergie, für die die logarithmische Ableitung korrekt wiedergegeben wird, heraus.

Für beide Probleme schaffte Vanderbilt [Van90] Abhilfe durch die Einführung Ultrasofter Pseudopotentiale. Eine Verallgemeinerung auf mehrere Referenzenergien

veröffentlichte zeitgleich auch Blöchl [Blö90]. Bei Vanderbilts Methode wird der Preis für einen niedrigeren *Cutoff* (i.e. die höchste zur Darstellung benötigte Fourierkomponente) bei den Wellenfunktionen bezahlt durch einen sehr hohen für Ladungsdichte und Potential.

Die *Projector Augmented Wave* (PAW)-Methode von Blöchl [Blö94] ist zwar formal sehr ähnlich zu Pseudopotentialmethoden, geht jedoch eigentlich darüber hinaus und liefert eine eigene, solide theoretische Fundierung. Die Variationsparameter sind glatte Wellenfunktionen, welche – wie bei der Pseudopotentialmethode – keine Orthogonalitätsoszillationen aufweisen. Der Ortsraum wird unterteilt in Augmentationsbereiche um die Atome herum und in das interstitielle Gebiet zwischen diesen. Im interstitiellen Bereich wird die glatte Wellenfunktion durch Ebene Wellen beschrieben. Innerhalb der Augmentationsbereiche jedoch wird die echte Wellenfunktion unter Zuhilfenahme der atomaren Wellenfunktionen dargestellt. Diese sind jedoch nicht zusätzliche Basisfunktionen (wie das in *mixed-basis*-Verfahren der Fall ist), sondern ihre Gewichtung wird mittels eines Projektionsverfahrens aus den glatten Wellenfunktionen gewonnen. Entsprechend werden im Nahbereich der Atome die Ladungsdichte (inklusive *frozen-core*-Ladungsdichte) und das Potential auf einem radialen Gitter bestimmt.

Der Form nach ähnelt die PAW-Methode stark einem generalisierten, dichteabhängigen nichtlokalen Pseudopotential; die unterschiedliche Beschreibung der Funktion in Augmentationsgebiet bzw. interstitiellem Bereich und die demgemäße Konstruktion der Ladungsdichte und des Hartree-Potentials hingegen erinnern an *Augmented Plane Wave* (APW)-Methoden.

In dieser Arbeit wurde eine PAW-Methode implementiert. Aufgrund der hohen Analogie zur Pseudopotentialmethode setzte die Implementierung auf dem bestehenden Pseudopotential-Programm **ESTCoMPP** auf [BBA⁺00]. Daher wird auch die PP-Methode beschrieben.

Das Konzept der Pseudopotential-Methode

Die wesentliche Voraussetzung für das Funktionieren der Pseudopotentialmethode ist, daß bei der quantenmechanischen Beschreibung von Festkörpern die Elektronen räumlich wie energetisch in zwei Klassen zerfallen:

- Die core-Zustände liegen energetisch sehr viel tiefer als die Valenzzustände; ihre Energiebänder sind sehr schmal und weichen nur wenig von der Eigenenergie der atomaren core-Zustände ab. Räumlich sind diese Elektronen sehr nahe

am Atomkern lokalisiert (Rumpfbereich) und haben mit den core-Elektronen der benachbarten Atome überhaupt keinen Überlapp, der mit den Valenzelektronen ist (im Idealfall) gering. Daher sind die core-Elektronen von der chemischen Umgebung ziemlich unabhängig und leisten zu einer chemischen Bindung keinen Beitrag.

- Die Valenzelektronen hingegen liegen energetisch erheblich höher und haben eine größere Ausdehnung, nämlich oft über den ganzen interstitiellen Bereich. Entsprechend ist der Überlapp mit den Elektronen benachbarter Atome groß und daher die Abhängigkeit von der chemischen Umgebung sehr ausgeprägt.

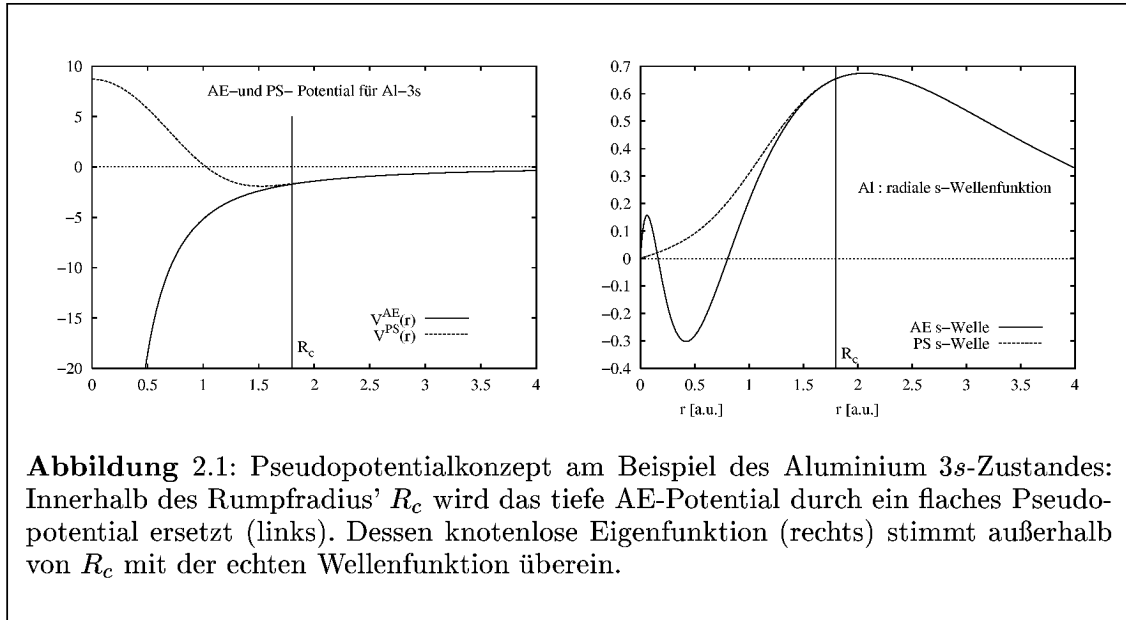
Außerhalb des Rumpfbereiches spüren die Valenzelektronen nur ein durch die core-Elektronen abgeschirmtes, schwächeres Potential. Daher werden sie dort durch relativ glatte Funktionen beschrieben.

Im Rumpfbereich des Atomes hingegen müssen sie die Orthogonalitätsbedingung zu den core-Zuständen erfüllen. Daher zeigen sie dort i. Allg. starke Oszillationen. Das Verhalten in diesem Bereich hat jedoch kaum Einfluß auf die chemische Bindung.

Daher entledigt man sich in der Pseudopotentialmethode der core-Elektronen und ersetzt das echte Potential durch ein Pseudopotential, welches außerhalb des Rumpfbereiches der Atome dieselbe Form hat und folglich – sofern die richtige Normierung beachtet wird – zu denselben Wellenfunktionen führt wie das echte Rumpfpotential (d.h. Kernpotential plus Abschirmpotential der core-Elektronen). Die Valenzelektronen müssen dann nicht mehr im core-Bereich oszillieren, um orthogonal zu den core-Wellenfunktionen zu sein. Die Pseudowellenfunktionen $\tilde{\phi}$ stimmen also im für die Bindung relevanten Bereich mit den echten Wellenfunktionen ϕ überein, sind insgesamt aber viel glatter und daher durch relativ wenige Ebene Wellen zu beschreiben.

Bei der Konstruktion müssen zwei widerstreitende Kriterien erfüllt werden:

- Zum einen sollen die Wellenfunktionen möglichst weich sein, d.h. sich durch möglichst wenige Ebene Wellen darstellen lassen. Dies ist umso leichter möglich, je größer man den Radius R_c des Rumpfbereiches wählt.
- Zum andern sollen die Potentiale optimal transferierbar sein, das heißt in allen chemischen Umgebungen das Bindungsverhalten korrekt wiedergeben. Dies ist natürlich umso besser erfüllbar, je kleiner der Bereich ist, in dem die Pseudowellenfunktion von der echten Wellenfunktion abweicht.



Eine obere Grenze für den Rumpfradius R_c wird ungefähr durch den Kovalenten Radius gegeben.

Logarithmische Ableitungen

Ein Elektron, das sich in einiger Entfernung von einem Atom befindet, „sieht“ von diesem nur dessen Streueigenschaften, welche sich im Rahmen einer Partialwellenentwicklung durch einen Satz von Streuphasen beschreiben lassen (vgl. z.B. [Eng92]). Diese wiederum sind eindeutig festgelegt durch die Logarithmische Ableitung

$$L_l^{AE}(R, \varepsilon) = \left. \frac{\partial \ln \phi_l(r, \varepsilon)}{\partial r} \right|_{r=R} \quad (2.27)$$

der Streuzustände $\phi_l(r, \varepsilon)$ bei der betreffenden Energie ε an einem beliebigen Radius R .

Wenn ein Pseudopotential die Logarithmische Ableitung korrekt reproduziert für den Bereich der Valenzenergien

$$L_l^{AE}(R, \varepsilon) = L_l^{PS}(R, \varepsilon), \quad (2.28)$$

so ist es für ein Elektron außerhalb der Rumpffregion nicht vom echten Potential zu unterscheiden.

Nach der Konstruktion der Pseudopotentiale ist dies für die Referenzenergie $\varepsilon = \varepsilon_{ref}$ exakt erfüllt.

Für die Energieableitung der Logarithmischen Ableitung gilt [YC82]

$$\frac{\partial}{\partial \varepsilon} L_l^{AE}(R, \varepsilon) = -\frac{2}{(\phi_l(R, \varepsilon))^2} \int_0^R dr (\phi_l(r, \varepsilon))^2 \quad (2.29)$$

und entsprechend für die Streuung am Pseudopotential

$$\frac{\partial}{\partial \varepsilon} L_l^{PS}(R, \varepsilon) = -\frac{2}{(\tilde{\phi}_l(R, \varepsilon))^2} \int_0^R dr (\tilde{\phi}_l(r, \varepsilon))^2. \quad (2.30)$$

Da normerhaltende Pseudopotentiale gerade so konstruiert sind, daß für die Referenzenergie ε_{ref} die Norm der Pseudowellenfunktionen innerhalb des Rumpfradius derjenigen der echten Wellenfunktionen entsprechen (und die Funktionen außerhalb übereinstimmen), sind die Energieableitungen bei $\varepsilon = \varepsilon_{ref}$ identisch. Daher stimmen die Logarithmischen Ableitungen nicht nur für die Referenzenergie ε_{ref} selbst überein, sondern dies ist in der ersten Ordnung auch noch um diesen Energiewert herum richtig.

Linearisierung des Austausch-Korrelations-Potentials

Da in der Darstellung durch Ebene Wellen die Rumpfladung wegen ihrer scharfen Strukturierung nicht wiedergegeben werden kann, ist es nötig, das Austausch-Korrelations-Potential zu linearisieren:

$$V_{xc}[n_c + n_v](r) \approx V_{xc}[n_c](r) + V_{xc}[n_v](r) \quad (2.31)$$

So erhält man einen unveränderlichen Anteil $V_{xc}[n_c](r)$, der im Pseudopotential aufgeht und einen additiven Anteil $V_{xc}[n_v](r)$, der nur von der Valenzladungsdichte abhängt.

Im Idealfall sind core-Ladungsdichte $n_c(r)$ und Valenzladungsdichte $n_v(r)$ räumlich voneinander getrennt; dann ist das Vorgehen exakt. Meist gibt es jedoch eine Überlappung, welche zu einem Linearisierungsfehler führt. Dieser kann durch eine partielle core-Ladungsdichte [LFC82] gemildert, jedoch nicht völlig beseitigt werden.

Semilokale Pseudopotentiale

Bei der Konstruktion eines Pseudopotentials geht man von der exakten Lösung der Schrödingergleichung für ein isoliertes Atom aus. Wegen der Rotationssymmetrie kann man die Wellenfunktionen nach Drehimpulsquantenzahlen separieren. Für die Valenzelektronen wird i. Allg. zunächst ein semilokales Potential bestimmt, d.h.

eines, welches lokal in der Radialkomponente, jedoch nichtlokal in den Winkelvariablen ist. Für jede l -Komponente hat man ein anderes radiales Potential $V_l^{SL}(r)$, welches für ein einzelnes Atom eine knotenlose Pseudowellenfunktion $\tilde{\phi}_l$ anstatt der echten Wellenfunktion ϕ_l zur Lösung hat. Der genaue Konstruktionsvorgang wird in Kapitel 4 beschrieben.

$$V^{SL}(\vec{r}) = \sum_l V_l^{SL}(r) \mathcal{P}_l$$

Dabei bezeichnet $\mathcal{P}_l$ die Projektion auf die l -Komponente. Da alle Potentiale dasselbe langreichweitige Verhalten haben, spaltet man ein lokales Potential ab:

$$V^{SL}(\vec{r}) = \sum_l V_{loc}(r) \mathcal{P}_l + \sum_l (V_l^{SL}(r) - V_{loc}(r)) \mathcal{P}_l \quad (2.32)$$

$$= V_{loc}(r) + \sum_l \Delta V_l^{SL}(r) \mathcal{P}_l \quad (2.33)$$

Die Differenzpotentiale $\Delta V_l^{SL}(r)$ sind nur im Rumpfbereich von null verschieden.

Separable Pseudopotentiale

Die semilokale Form ist jedoch sehr rechenaufwendig. Daher überführt man sie gemäß [KB82] in eine separable Form:

$$V^{KB} = V_{loc} + \sum_L |p_L\rangle D_L \langle p_L| \quad (2.34)$$

wobei die Projektoren $|p_L\rangle$ definiert sind durch

$$\langle \vec{r} | p_L \rangle = Y_L(\hat{r}) \Delta V_l^{SL}(r) \frac{\tilde{\phi}_l(r)}{r} \quad (2.35)$$

Dabei ist der Faktor D_l so gewählt, daß die Anwendung des separablen Potentials auf die bei der Konstruktion verwendete Pseudowellenfunktion $\tilde{\phi}_l$ dasselbe Resultat liefert wie das semilokale Potential. Für andere Wellenfunktionen hingegen ist das Resultat nicht notwendig identisch mit dem des semilokalen Potentials; zumindest nahe der Referenzenergie (dem Eigenwert von ϕ) sollte der Effekt jedoch sehr ähnlich sein.

Für semilokale Potentiale garantiert der Knotensatz, daß es keine Eigenzustände mit geringerer Energie gibt als $\tilde{\phi}_l$, da dieser knotenlos konstruiert worden ist. Für Potentiale vom separablen Typ gibt es keine entsprechende Garantie, und so findet

man gelegentlich unerwünschte *ghost-states* mit sehr niedrigen Eigenwerten. Diese lassen sich als zusätzliche Polstellen in den logarithmischen Ableitungen identifizieren und oftmals durch die Wahl eines anderen lokalen Potentials beseitigen [BK90, GKS90, GSS91, SSO92].

Nicht-normerhaltende Pseudopotentiale

Die Pseudisierung der Wellenfunktionen funktioniert sehr gut, solange die AE-Wellenfunktionen Knoten besitzen, wie etwa bei den klassischen Halbleitermaterialien *Si* und *Ge* oder auch wie in Abbildung 2.1 bei *Al*. Weniger günstig ist die Situation, wenn keine Nullstellen der Valenzfunktionen vorhanden sind, so wie das bei den $2p$ - und $3d$ -Elementen der Fall ist. Abbildung 2.2 zeigt die Kupfer $3d$ -Wellenfunktion als Beispiel. Die Normerhaltung ist eine so starke Bedingung, daß die Pseudisie-

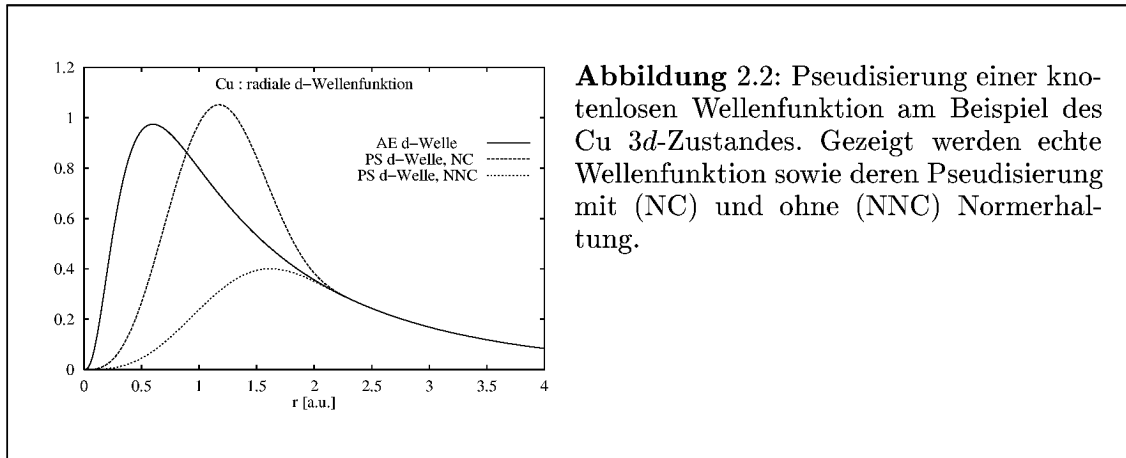


Abbildung 2.2: Pseudisierung einer knotenlosen Wellenfunktion am Beispiel des Cu $3d$ -Zustandes. Gezeigt werden echte Wellenfunktion sowie deren Pseudisierung mit (NC) und ohne (NNC) Normerhaltung.

rung der Wellenfunktion kaum einen Gewinn an Glattheit bringt (man vergleiche die Krümmungsradien). Soweit man aber nicht an die Normerhaltung gebunden ist, läßt sich eine viel glattere Pseudowellenfunktion bestimmen.¹

Die Aufgabe der Normerhaltung (für die Wellenfunktionen) erfordert zusätzliche Maßnahmen, um die vollständige Ladung zu konstruieren, da diese selbstverständlich innerhalb der DFT benötigt wird. Dies leistet die Methode Ultrasofter Pseudopotentiale, aber auch das PAW-Verfahren, welches in dieser Arbeit in einer modifizierten Form umgesetzt worden ist.

¹Der Rumpfradius ist zudem hier zu Vergleichszwecken größer gewählt, als es für normerhaltende Pseudopotentiale sinnvollerweise getan werden könnte.

Kapitel 3

PAW-Methode

Die PAW-Methode wurde 1994 von P. Blöchl [Blö94] veröffentlicht. Sie weist eine große Ähnlichkeit zu Vanderbilts Methode Ultrasofter Pseudopotentiale auf, ist aber selbst keine Pseudopotentialmethode. Vielmehr handelt es sich um eine *all electron*-Methode und zeigt auch eine nahe Verwandtschaft zur FLAPW-Methode (siehe z.B. [Sin94]). Außer Blöchls Programm existieren derzeit meines Wissens zwei weitere PAW-Programme [HMD⁺97, KJ99].

Im folgenden soll einerseits die PAW-Methode präsentiert werden; zum anderen soll aber auch die Umsetzung im Programm detailliert dargestellt werden. Nach Möglichkeit habe ich letzteres jeweils in Unterkapitel abgetrennt, welche teilweise eine Vielzahl von Formeln enthalten, die zum allgemeinen Verständnis der Methode nicht unbedingt erforderlich sind. Um die Voraussetzung dafür zu schaffen, schicke ich einen einführenden Abschnitt über die Benutzung einer Superzelle und die verwendete Normierung voraus, ehe ich zum Kern der PAW-Methode komme.

Da angesichts der Komplexität des Sachverhaltes eine Vielzahl von Symbolen eingeführt werden müssen, sei schon an dieser Stelle auf die Zusammenstellung wichtiger Größen in Abschnitt 3.16 hingewiesen.

3.1 Superzellenmethode und Ebene Wellen

Die Gleichgewichtsstruktur der meisten Festkörper ist die regelmäßige Anordnung von Atomen in einem Kristallgitter. In der Realität ist natürlich jeder Festkörper endlich; zu seiner theoretischen Beschreibung vernachlässigt man aber (zunächst) alle Oberflächen und Störstellen und betrachtet ihn als unendlich ausgedehnt. Der ideale Kristall ergibt sich als unendliche Aneinanderreihung einer kleinen Struktur-

einheit, der sogenannten *Elementarzelle*. Durch die periodische Wiederholung dieser Zelle wird der Raum ohne Zwischenräume ausgefüllt.

Daher überführt eine Verschiebung entlang einer Kante der Elementarzelle um ein ganzzahliges Vielfaches der Kantenlänge den Kristall in sich selbst. Folglich kennt man die Eigenschaften eines idealen Kristalles vollständig, wenn man *eine* Elementarzelle beschreiben kann.

Der unendliche Kristall ist jedoch eine idealisierte Vorstellung. Die endliche Größe oder das Vorhandensein von Verunreinigungen stören die Periodizität. Gerade Oberflächen und Defekte haben aber interessante physikalische Eigenschaften, die man auch beschreiben können möchte.

Zu diesem Zwecke beschreibt man den imperfekten Kristall durch *Superezellen*. Eine Supercelle enthält einen möglichst großen Bereich des Kristalles, welcher auch Störstellen oder Oberflächen enthalten kann. Diese Supercelle wird – ebenso wie die Elementarzelle des idealen Kristalles – in drei Dimensionen periodisch fortgesetzt. Enthält eine Supercelle einen Defekt, so besitzt der Festkörper, der daraus zusammengesetzt wird, unendlich viele Defekte. Daher muß man die Supercelle so groß wählen, daß die Wechselwirkung einer Störstelle mit ihren Replikationen in benachbarten Zellen ausgeschlossen wird.

Auf diese Weise kann man den Vorteil einer periodischen Beschreibung auch für den gestörten Kristall künstlich herbeiführen.

Die Supercelle Ω wird durch drei linear unabhängige Basisvektoren $\vec{a}_1$, $\vec{a}_2$ und $\vec{a}_3$ aufgespannt. Das Volumen der Supercelle, das gleichfalls mit Ω bezeichnet wird, ergibt sich als Spat-Produkt der Basisvektoren:

$$\Omega = \vec{a}_1 \cdot (\vec{a}_2 \times \vec{a}_3) \quad (3.1)$$

Die Kohn–Sham–Gleichungen (2.21) müssen also für ein zellperiodisches Potential gelöst werden:

$$V(\vec{r}) = (\mathcal{T}_{\vec{R}}V)(\vec{r}) = V(\mathcal{T}_{-\vec{R}}\vec{r}) = V(\vec{r} - \vec{R}) \quad (3.2)$$

Der Translationsoperator $\mathcal{T}_{\vec{R}}$ beschreibt eine Verschiebung um einen Gittervektor $\vec{R}$. Gittervektoren sind alle ganzzahligen Linearkombinationen der Basisvektoren $\vec{a}_1$, $\vec{a}_2$ und $\vec{a}_3$.

Die durch die Supercellenbeschreibung erzwungene Periodizität gestattet es, das Bloch'sche Theorem [Blo28] anzuwenden. Dieses besagt, daß sich die Lösungen $\psi_{\vec{k}\nu}(\vec{r})$ einer Einteilchen–Schrödingergleichung zu einem gitterperiodischen Potential als Produkt einer Ebenen Welle $\frac{1}{\sqrt{\Omega}}e^{-i\vec{k}\vec{r}}$ und einer gitterperiodischen Funktion

$u(\vec{r})$ darstellen lassen.

$$\begin{aligned}\Psi_{\vec{k}}(\vec{r}) &= e^{-i\vec{k}\vec{r}} u(\vec{r}) \\ \mathcal{T}_{\vec{R}} u(\vec{r}) &= u(\mathcal{T}_{-\vec{R}} \vec{r}) = u(\vec{r} - \vec{R}) = u(\vec{r})\end{aligned}\quad (3.3)$$

Eine derartige gitterperiodisch modulierte Ebene Welle wird auch als Blochwellenfunktion bezeichnet. Bei Verschiebung um einen Gittervektor $\vec{R}$ ändert sie sich nur um einen Phasenfaktor $e^{i\vec{k}\vec{R}}$.

Die Lösungen der Kohn–Sham Gleichungen (2.21) können nach den Blochvektoren $\vec{k}$ klassifiziert werden. Deren Wertebereich ist die erste Brillouin–Zone, die Einheitszelle des reziproken Gitters, welches durch die reziproken Basisvektoren aufgespannt wird:

$$\vec{b}_1 = 2\pi \frac{(\vec{a}_2 \times \vec{a}_3)}{\vec{a}_1 \cdot (\vec{a}_2 \times \vec{a}_3)}, \text{ und zyklisch.} \quad (3.4)$$

Reziproke Gittervektoren $\vec{G}$ sind alle ganzzahligen Linearkombinationen der drei reziproken Basisvektoren $\vec{b}_1$, $\vec{b}_2$ und $\vec{b}_3$. Nach ihrer Definition gilt

$$e^{i\vec{G}\vec{R}} = 1 \quad (3.5)$$

für alle Gittervektoren $\vec{G}$ des reziproken und $\vec{R}$ des Superzellengitters.

Auf Grund ihrer Gitterperiodizität lassen sich die Funktionen $u(\vec{r})$ in einer diskreten Fourier–Reihe von reziproken Gittervektoren entwickeln.

Für die Darstellung von Wellenfunktionen und Hamiltonian im Impulsraum wird eine Basis aus Ebenen Wellen mit folgender Normierung gewählt:

$$\langle \vec{r} | \vec{k} + \vec{G} \rangle = \frac{1}{\sqrt{\Omega}} e^{-i(\vec{k} + \vec{G})\vec{r}} \quad (3.6)$$

$$\begin{aligned}\langle \vec{k}' + \vec{G}' | \vec{k} + \vec{G} \rangle_{\Omega} &= \frac{1}{\Omega} \int_{\Omega} d^3r e^{i(\vec{k}' + \vec{G}')\vec{r}} e^{-i(\vec{k} + \vec{G})\vec{r}} = \delta_{\vec{k}' + \vec{G}', \vec{k} + \vec{G}} \\ \int_{\Omega} d^3r |\vec{r}\rangle \langle \vec{r}| &= 1 = \sum_{\vec{G}} |\vec{k} + \vec{G}\rangle \langle \vec{k} + \vec{G}| \end{aligned} \quad (3.7)$$

Die Darstellung der Blochfunktionen und der Wellenfunktionen nimmt also folgende

Gestalt an:

$$\begin{aligned}
u_{\vec{k}\nu}(\vec{r}) &= \langle \vec{r} | u_{\vec{k}\nu} \rangle = \langle \vec{r} | \left(\sum_{\vec{G}} |\vec{G}\rangle \langle \vec{G}| \right) | u_{\vec{k}\nu} \rangle = \frac{1}{\sqrt{\Omega}} \sum_{\vec{G}} e^{-i\vec{G}\vec{r}} c_{\vec{k}\nu}(\vec{G}) \\
\tilde{\psi}_{\vec{k}\nu}(\vec{r}) &= \langle \vec{r} | \tilde{\psi}_{\vec{k}\nu} \rangle = e^{-i\vec{k}\vec{r}} u_{\vec{k}\nu}(\vec{r}) = \frac{1}{\sqrt{\Omega}} \sum_{\vec{G}} e^{-i(\vec{k}+\vec{G})\vec{r}} c_{\vec{k}\nu}(\vec{G}) \\
c_{\vec{k}\nu}(\vec{G}) &= \langle \vec{G} | u_{\vec{k}\nu} \rangle = \frac{1}{\sqrt{\Omega}} \int_{\Omega} d^3r e^{i\vec{G}\vec{r}} u_{\vec{k}\nu}(\vec{r}) \\
&= \langle \vec{k} + \vec{G} | \tilde{\psi}_{\vec{k}\nu} \rangle = \frac{1}{\sqrt{\Omega}} \int_{\Omega} d^3r e^{i(\vec{k}+\vec{G})\vec{r}} \tilde{\psi}_{\vec{k}\nu}(\vec{r})
\end{aligned} \tag{3.8}$$

Die Fourierreihe über die Gittervektoren $\vec{G}$ enthält im Prinzip unendlich viele Summanden. Sind jedoch die darzustellenden Funktionen hinreichend glatt, so werden die Koeffizienten $c_{\vec{k}\nu}(\vec{G})$ schnell betraglich sehr klein, so daß man die Summation bei einer endlichen Wellenzahl G_{max} abschneiden kann. Üblicherweise wird diese in Energieeinheiten angegeben:

$$E_{max} = G_{max}^2 \tag{3.9}$$

Die elektronischen Freiheitsgrade des unendlich ausgedehnten periodischen Festkörpers werden über die Translationssymmetrie und die Anwendung des Bloch'schen Theorems auf die Eigenwertspektren für die Blochvektoren $\vec{k}$ innerhalb der ersten Brillouin-Zone abgebildet. Deshalb führt die Bestimmung der Gesamtenergie, der Ladungsdichte oder der Kräfte eines Kristalles zu einer Integration über die Brillouin-Zone. Diese Integration wird numerisch durch eine gewichtete Summation über endlich viele, spezielle $\vec{k}$ -Punkte approximiert:

$$\int_{BZ} d^3k F(\vec{k}) \rightarrow \sum_{\vec{k}} \omega_{\vec{k}} F(\vec{k}) \tag{3.10}$$

Die $\omega_{\vec{k}}$ sind dabei die Gewichte der ausgewählten $\vec{k}$ -Punkte. Dies entspricht einer Art dreidimensionalem Trapezverfahren, wobei die speziellen $\vec{k}$ -Punkte und Gewichte $\omega_{\vec{k}}$ bei uns meist nach dem Verfahren von Monkhorst und Pack [PM77] bestimmt werden.

Das Ortsraumgitter

Durch die Beschränkung der Ebenen Wellen auf solche mit einem Wellenvektor, der kleiner als der Abschneidewert G_{max} ist, wird eine Kugel ($|\vec{G}| \leq G_{max}$) im rezipro-

ken Raum definiert.¹ Für eine dreidimensionale Fouriertransformation wird diese in einen Quader (*FFT-Box*) eingebettet. Durch die Fouriertransformation werden diese $\vec{G}$ -Vektoren auf ein Quader mit derselben Anzahl von Punkten im Ortsraum abgebildet. Diese Punkte sind in gleichmäßigen Abständen im gesamten Volumen der Superzelle angeordnet (ohne Rücksicht auf die Lage der Atome). Die Anzahl der $\vec{G}$ -Vektoren legt also auch gleichzeitig die Anzahl der Gitterpunkte des Ortsraumgitters und damit dessen räumliche Auflösung fest.² Dieses Gitter wird *reguläres* oder *kartesisches* Gitter genannt.

Gittersummen

An mehreren Stellen müssen im folgenden „Gittersummen“ berechnet werden. Das sind Funktionen f , die sich zusammensetzen aus einem Beitrag für jedes Atom, welcher der Wert einer an diesem Atom zentrierten Funktion f_μ ist. Beispielsweise wird das lokale Pseudopotential (s.u.) superponiert aus den lokalen Pseudopotentialen jedes einzelnen Atomes. Um den Wert dieser Gittersummen innerhalb einer Superzelle zu berechnen, müssen auch die „Schwänze“ von den Funktionen summiert werden, die sich in anderen Gitterzellen befinden:

$$f(\vec{r}) = \sum_{\vec{R}} \sum_{\mu} f_{\mu}(\vec{r} - \vec{\tau}_{\mu} - \vec{R}) \quad (3.11)$$

Dabei summiert μ Atome innerhalb *einer* Superzelle (typischerweise diejenige, welche zu $\vec{R} = 0$ gehört). $\vec{\tau}_{\mu}$ ist der Ort des Atomes μ in dieser Superzelle, seine periodischen Replikationen sitzen an den Orten $\vec{R} + \vec{\tau}_{\mu}$, wobei $\vec{R}$ ein beliebiger Gittervektor ist. Bildet man die Fouriertransformierte dieser Funktion, so kann man die Summation über unendlich viele Summanden von Integralen über ein endliches Volumen ersetzen durch eine endliche Summe von Integralen über ein unendliches

¹Natürlich ist mit *Kugel* oder *Quader* nur die Form des einhüllenden Volumens gemeint; es handelt sich aber um diskrete Punkte in diesen Volumina. Im Allgemeinen sind auch die Hauptachsen verschieden, d.h. es handelt sich um Ellipsoide.

²Genauer gesagt entscheidet derjenige Abschneidewert, welcher für Dichte und Potential benutzt wird, über die maximale Auflösung im Ortsraum; dieser ist i. Allg. doppelt so groß wie der für die Wellenfunktionen.

Volumen:

$$f(\vec{G}) = \frac{1}{\Omega} \int_{\Omega} e^{i\vec{G}\vec{r}} \sum_{\vec{R}} \sum_{\mu} f_{\mu}(\vec{r} - \vec{\tau}_{\mu} - \vec{R}) d^3r \quad (3.12)$$

$$= \frac{1}{\Omega} \sum_{\mu} \sum_{\vec{R}} e^{i\vec{G}(\vec{\tau}_{\mu} + \vec{R})} \int_{\Omega} e^{i\vec{G}(\vec{r} - \vec{\tau}_{\mu} - \vec{R})} f_{\mu}(\vec{r} - \vec{\tau}_{\mu} - \vec{R}) d^3r \quad (3.13)$$

$$= \sum_{\mu} e^{i\vec{G}\vec{\tau}_{\mu}} \sum_{\vec{R}} \frac{1}{\Omega} \int_{\Omega + \vec{\tau}_{\mu} + \vec{R}} e^{i\vec{G}\vec{r}'} f_{\mu}(\vec{r}') d^3r' \quad (3.14)$$

$$= \sum_{\mu} e^{i\vec{G}\vec{\tau}_{\mu}} \frac{1}{\Omega} \int_{\infty} e^{i\vec{G}\vec{r}'} f_{\mu}(\vec{r}') d^3r' = \sum_{\mu} S^{\mu}(\vec{G}) f_{\mu}(\vec{G}) \quad (3.15)$$

wobei benutzt wurde, daß $e^{i\vec{G}\vec{R}} = 1$ und daß die Vereinigung aller verschobenen Volumina $\Omega + \vec{\tau}_{\mu} + \vec{R}$ über alle Gittervektoren $\vec{R}$ das unendliche Volumen ergibt (für jedes $\vec{\tau}_{\mu}$).

Eine Gittersumme zerfällt also im reziproken Raum in Strukturfaktoren

$$S^{\mu}(\vec{G}) = e^{i\vec{G}\vec{\tau}_{\mu}}, \quad (3.16)$$

die nur von den Atompositionen abhängen, und Formfaktoren

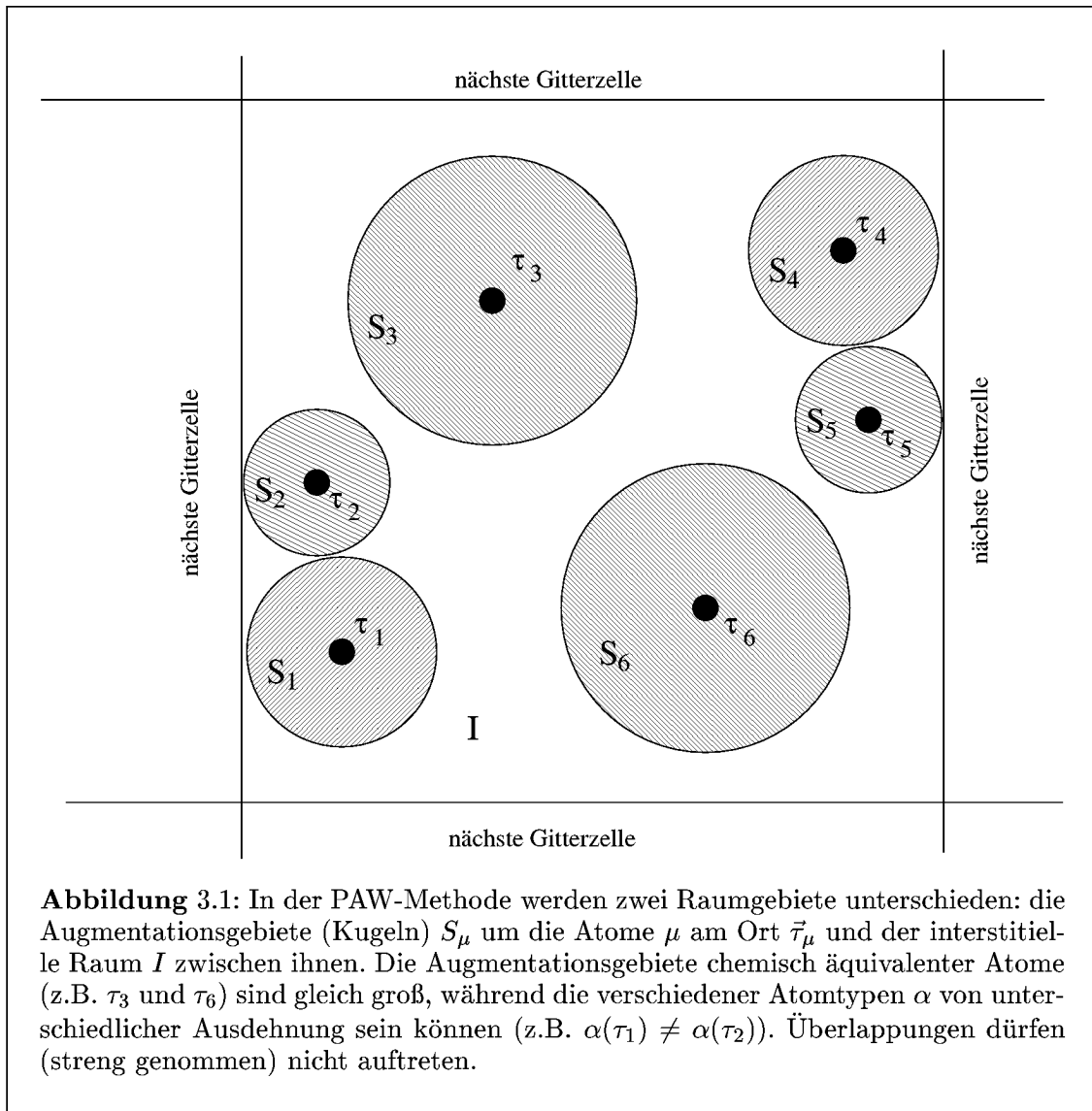
$$f_{\mu}(\vec{G}) = \frac{1}{\Omega} \int_{\infty} e^{i\vec{G}\vec{r}'} f_{\mu}(\vec{r}') d^3r, \quad (3.17)$$

die nur von den Eigenschaften der Atomsorte festgelegt sind. Formfaktoren werden nur einmal zu Beginn des Programmes berechnet, die Strukturfaktoren hingegen werden für jede Atomkonfiguration berechnet, d.h. sie werden nach jedem Verschieben von Atomen aktualisiert. Die Funktion $f(\vec{r})$ im Ortsraum erhält man wiederum durch Fouriertransformation.

3.2 Grundidee

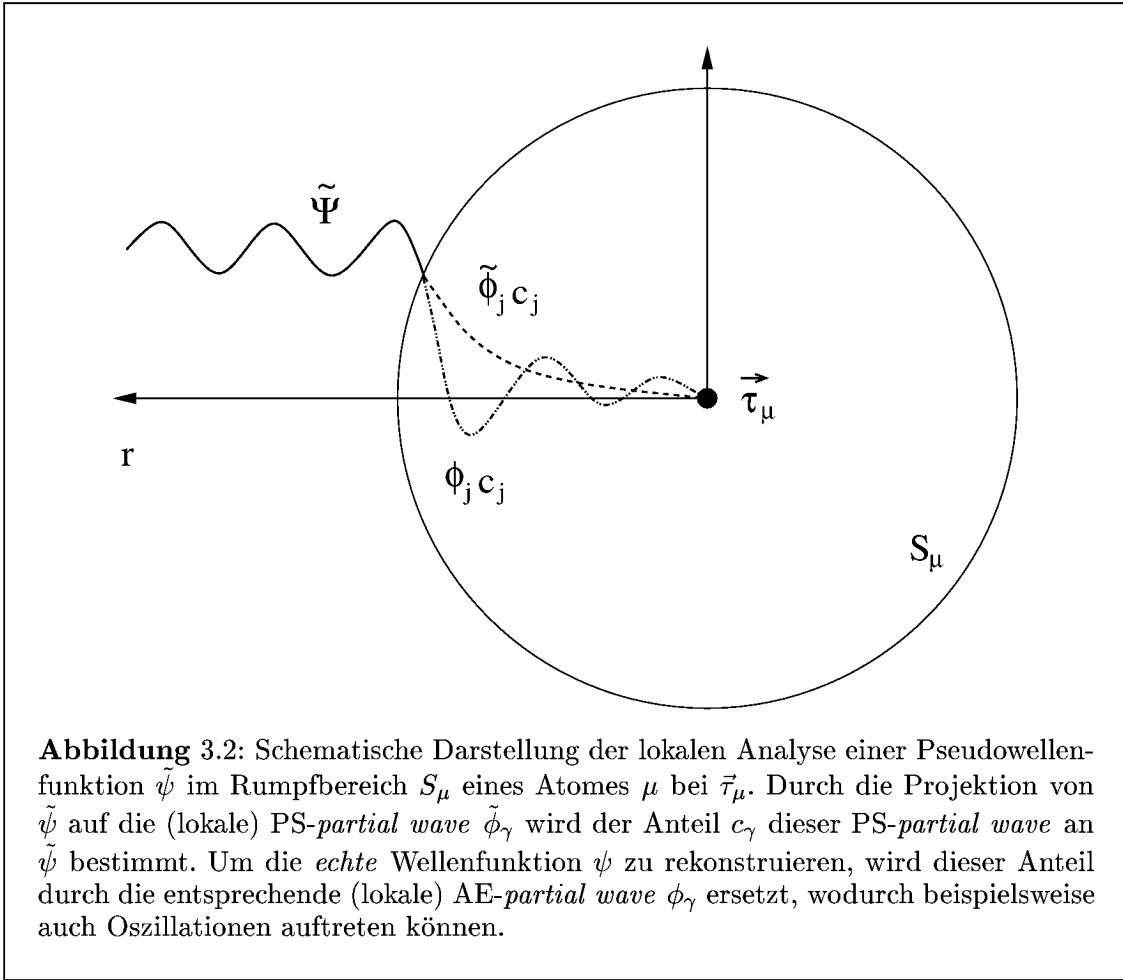
Die PAW-Methode beinhaltet sowohl Elemente der LAPW-Methode als auch solche der Pseudopotentialmethode. Gemeinsam ist den drei Methoden, daß die Wellenfunktionen nach Basisfunktionen entwickelt werden, deren Entwicklungskoeffizienten variationell bestimmt werden. Gemeinsam ist ebenso, daß der Raum in zwei Bereiche aufgeteilt wird: die Rumpfreionen S_{μ} , welche jedes Atom μ kugelförmig umgeben, und den Zwischenrumpfbereich I (siehe Abb. 3.1).

Als Variationsfunktionen werden in der PAW-Methode (wie in der Pseudopotentialmethode) Pseudowellenfunktionen benutzt, die außerhalb des Rumpfbereiches der



Atome mit den echten Wellenfunktionen übereinstimmen. Im Gegensatz zur Pseudopotentialmethode, bei der die Pseudowellenfunktion auch in der Rumpfreion als näherungsweise Beschreibung der *echten* Wellenfunktion benutzt wird, werden in der PAW-Methode innerhalb jedes Rumpfbereiches die *echten* Wellenfunktionen inklusive eventueller Oszillationen mit Hilfe vordefinierter lokalisierter Funktionen aus den Pseudowellenfunktionen *rekonstruiert*. Dazu werden die Pseudowellenfunktionen einer lokalen Analyse unterzogen. Das geschieht durch Projektion der Pseudowellenfunktionen auf vordefinierte, atomspezifische „PS-*partial waves*“. Zur Rekonstruktion der *echten* Wellenfunktion wird der projizierte Anteil durch den atomspezifischen exakten „AE-*partial wave*“-Anteil ersetzt.

Salopp gesprochen wird beispielsweise ein Zustand, der lokal den Charakter einer Pseudo- s -Wellenfunktion eines bestimmten Atomes hat, dort durch die entsprechende *echte* s -Wellenfunktion dieses Atomes ersetzt. Die Wellenfunktion wird also in diesem Bereich (durch die Differenz von echten und glatten *partial waves*) *augmentiert*. Das Rumpfbereich heißt daher auch Augmentationsgebiet. Abbildung 3.2 illustriert diese Rekonstruktion der *echten* Wellenfunktion.



Ein Atom wird also in der PAW-Methode beschrieben durch:

- PS-partial waves, die angeben, Funktionen welchen Charakters im Augmentationsgebiet zu ersetzen sind; diese entsprechen den atomaren Pseudowellenfunktionen eines Pseudopotentials.
- AE-partial waves, welche angeben, wodurch diese zu ersetzen sind; dies sind die *echten* atomaren Wellenfunktionen, aus denen die PS-partial waves generiert wurden.

- Projektoren, welche durch *den Charakter* der zugehörigen PS-*partial wave* festgelegt werden. Sie sind räumlich lokalisiert im Augmentationsgebiet dieses Atomes.

Die Anzahl und Art der *partial waves* bzw. Projektoren bestimmt den Anwendungsbereich und die Genauigkeit der PAW-Methode. Durch Hinzunahme weiterer Funktionen läßt sich die Genauigkeit im Prinzip beliebig erhöhen. Die Konstruktion der Projektoren und *partial waves* wird in Kapitel 4 beschrieben.

Die Berechnung der Elektronendichte, der Potentiale und der Energien wird auf Grundlage der *echten* Wellenfunktionen durchgeführt. Insbesondere kann so das Austausch-Korrelations-Potential mit der vollständigen Elektronendichte berechnet werden.

Auch in der LAPW-Methode werden als Variationsparameter nur die glatten Wellenfunktionen (in einer Basis aus Ebenen Wellen) direkt berücksichtigt. Der Unterschied zwischen PAW- und (L)APW-Methode liegt in der Augmentierungsvorschrift. Bei der (L)APW-Methode wird ebenfalls die glatte Welle innerhalb einer Kugel – dort wird sie *Muffin-Tin*-Kugel genannt – durch eine lokalisierte Funktion ersetzt, welche die Lösung der radialen Schrödingergleichung zu vorgegebenem Allelektronen-Potential ist. Als Augmentierungsvorschrift verlangt man, daß die Lösung innerhalb der Kugel am Kugelrand eine Anschlußbedingung an die Lösung aus Ebenen Wellen im interstitiellen Bereich erfüllt. Dazu werden die Ebenen Wellen nach l -abhängigen Partialwellen entwickelt. Jede Partialwelle muß – je nach Verfahren – stetig oder stetig differenzierbar durch eine lokalisierte radiale Lösung der Schrödingergleichung fortgesetzt werden. Welche der beiden Methoden, die PAW- oder die (L)APW-Methode nun genauer oder effizienter ist, ist a priori unklar. In Kapitel 6 wird versucht, zu einer Antwort darauf zu gelangen.

3.3 Definition der Transformation

Die im vorigen Abschnitt eingeführte Idee soll nun formaler dargestellt werden. Eine echte Wellenfunktion $|\psi\rangle$, welche orthogonal zu den core-Zuständen ist, soll beschrieben werden durch eine glatte Wellenfunktion $|\tilde{\psi}\rangle$, die sich nur in den Augmentationsgebieten von $|\psi\rangle$ unterscheidet.

Diese Zuordnung definiert eine Abbildung $\mathcal{T}$, die linear und umkehrbar eindeutig sein soll:

$$\mathcal{T} : \tilde{\mathcal{H}} \longrightarrow \mathcal{H}, \quad |\tilde{\psi}\rangle \mapsto |\psi\rangle \quad (3.18)$$

$\tilde{\mathcal{H}}$ und $\mathcal{H}$ sind die Hilberträume der glatten bzw. der echten Wellenfunktionen.

Durch diese Abbildung kann man Erwartungswerte $\langle \psi | A | \psi \rangle$ der *echten* Wellenfunktionen durch ebensolche $\langle \tilde{\psi} | A | \tilde{\psi} \rangle$ der Pseudowellenfunktionen und ebenso die Gesamtenergie als Funktional aller $\tilde{\psi}$ ausdrücken. Man erhält also die *echten* Werte, muß als Variationsparameter jedoch nur die viel glatteren Pseudozustände durch Ebene Wellen beschreiben.

Die Abbildung $\mathcal{T}$ kann dargestellt werden als Summe von lokalen Abbildungen $\mathcal{T}_\mu$, welche jeweils die Zuordnungsvorschrift auf dem Rumpfgebiet eines Atomes μ beschreiben und außerhalb dieses Gebietes verschwinden.

$$\mathcal{T} = 1 + \sum_{\mu} \mathcal{T}_\mu \quad \text{mit} \quad \mathcal{T}_\mu \psi | (\Omega \setminus S_\mu) \equiv 0 \quad (3.19)$$

Als AE-*partial waves* wird für jedes Atom eine Menge von Funktionen $|\phi_\gamma^\mu\rangle$ gewählt, welche zusammen eine Basis der *echten* Wellenfunktionen beschränkt auf das Augmentationsgebiet S_μ darstellen. Zu jeder dieser AE-*partial waves* $|\phi_\gamma^\mu\rangle$ wird eine glatte PS-*partial wave* $|\tilde{\phi}_\gamma^\mu\rangle$ angegeben, die auf $|\phi_\gamma^\mu\rangle$ abgebildet werden soll:

$$\mathcal{T}_\mu |\tilde{\phi}_\gamma^\mu\rangle = |\phi_\gamma^\mu\rangle \quad (3.20)$$

Diese Funktionen $|\phi_\gamma^\mu\rangle$ bilden dann eine Basis des Raumes der glatten Wellenfunktionen beschränkt auf das Augmentationsgebiet. Damit ist die Abbildung $\mathcal{T}$ vollständig festgelegt.

Da diese Funktionen Basen der jeweiligen Räume sind, kann man ψ bzw. $\tilde{\psi}$ auf S_μ entwickeln:

$$\begin{aligned} |\psi\rangle &= \sum_{\gamma} |\phi_\gamma^\mu\rangle c_\gamma^\mu \quad \text{auf } S_\mu \\ |\tilde{\psi}\rangle &= \sum_{\gamma} |\tilde{\phi}_\gamma^\mu\rangle c_\gamma^\mu \quad \text{auf } S_\mu \\ |\psi\rangle &= |\tilde{\psi}\rangle \quad \text{außerhalb von } S_\mu. \end{aligned} \quad (3.21)$$

Da $|\tilde{\phi}_\gamma^\mu\rangle$ und $|\phi_\gamma^\mu\rangle$ durch die Abbildung $\mathcal{T}_\mu$ verbunden sind, sind die Koeffizienten c_γ^μ in beiden Entwicklungen nicht unabhängig, sondern für beide Entwicklungen dieselben. Weil es sich um eine lineare Abbildung handelt, kann man diese Koeffizienten als Skalarprodukt von $|\tilde{\psi}\rangle$ mit gewissen noch näher zu diskutierenden, bis dato aber noch unbekannten Projektorfunktionen $\langle \tilde{p}_\gamma^\mu |$ darstellen.³

$$c_\gamma^\mu = \langle \tilde{p}_\gamma^\mu | \tilde{\psi} \rangle \quad (3.22)$$

³Dies folgt aus dem Riesz'schen Darstellungssatz [BS79].

Also gilt:

$$|\psi\rangle = |\tilde{\psi}\rangle + \sum_{\mu,\gamma} \left(|\phi_\gamma^\mu\rangle - |\tilde{\phi}_\gamma^\mu\rangle \right) \langle \tilde{p}_\gamma^\mu | \tilde{\psi} \rangle \quad (3.23)$$

Die Abbildung $\mathcal{T}$ kann man also wie folgt schreiben:

$$\mathcal{T} = 1 + \sum_{\mu,\gamma} \left(|\phi_\gamma^\mu\rangle - |\tilde{\phi}_\gamma^\mu\rangle \right) \langle \tilde{p}_\gamma^\mu | \quad (3.24)$$

Diese Abbildung ist bijektiv (aber nicht unitär). Die Inverse ist (vgl. Anhang B):

$$\mathcal{T}^{-1} = 1 + \sum_{\mu,\gamma} \left(|\tilde{\phi}_\gamma^\mu\rangle - |\phi_\gamma^\mu\rangle \right) \langle p_\gamma^\mu | \quad \text{mit} \quad \langle p_\gamma^\mu | = \sum_{\gamma'} (\langle \tilde{p}^\mu | \phi^\mu \rangle)_{\gamma,\gamma'}^{-1} \langle \tilde{p}_{\gamma'}^\mu | \quad (3.25)$$

Da sich aus Anwendung der Transformation (Glg. (3.24)) auf die PS-*partial waves* wiederum die AE-*partial waves* ergeben sollen, müssen die Projektorfunktionen der Bedingung

$$\langle \tilde{p}_\gamma^\mu | \tilde{\phi}_{\gamma'}^\mu \rangle = \delta_{\gamma,\gamma'} \quad (3.26)$$

gehörten. Daraus folgt:

$$1_{S_\mu} \sum_{\gamma} |\tilde{\phi}_\gamma^\mu\rangle \langle \tilde{p}_\gamma^\mu | 1_{S_\mu} = 1_{S_\mu} \quad (3.27)$$

1_{S_μ} ist entsprechend der Einsoperator auf S_μ , 1_I entsprechend der Einsoperator im Zwischenraum I . Als Funktionen ϕ (AE-*partial waves*) wählt man naheliegenderweise atomare Eigenfunktionen (oder ungebundene Zustände), als Pseudofunktionen $\tilde{\phi}$ (PS-*partial waves*) entsprechend Pseudowellenfunktionen, welche nach einem beliebigen (nicht notwendig normerhaltenden) Konstruktionsverfahren für Pseudopotentiale erzeugt werden. Der Index γ ist dann als Multiindex (l, m, i) zu verstehen, wobei $L = (lm)$ den Drehimpuls charakterisiert und i die Funktionen zum selben Drehimpuls abzählt. In der Praxis ist der Satz der benutzten *partial waves* natürlich endlich. Es ist jedoch etwa aus der LAPW-Methode bekannt, daß man schon mit zwei Funktionen pro Drehimpuls l eine sehr gute Approximation der Wellenfunktionen im relevanten Energiebereich erhält.⁴ Im Prinzip kann man jedoch durch Hinzunahme weiterer Funktionen die Genauigkeit beliebig erhöhen.

Die Projektorfunktionen sind im folgenden als strikt lokalisiert in S_μ angenommen.

⁴Eine ausführliche Diskussion des durch die endliche Anzahl der Basisfunktionen bedingten Fehlers findet sich in [Blö94].

3.4 Operatoren

Erwartungswerte $\langle \psi | A | \psi \rangle$ lassen sich ausgehend von $\tilde{\psi}$ entweder bilden, indem man zunächst $\psi = \mathcal{T} \tilde{\psi}$ berechnet, oder aber indem man zum Operator $\tilde{A}$ übergeht:

$$\langle \psi | A | \psi \rangle = \langle \tilde{\psi} | \mathcal{T}^+ A \mathcal{T} | \tilde{\psi} \rangle = \langle \tilde{\psi} | \tilde{A} | \tilde{\psi} \rangle$$

In diesem Zusammenhang interessieren uns nur lokale Operatoren, nämlich $|\vec{r}\rangle\langle\vec{r}|$, als dessen Erwartungswert sich die Dichte ergibt, und der Operator $-\nabla^2$ der kinetischen Energie.⁵ Unter einem lokalen Operator A versteht man hier einen solchen, der die Eigenschaft $A = 1_S A 1_S + 1_I A 1_I$ hat; z.B. A im gewöhnlichen Sinne lokal.

Unter Beachtung von Glg. (3.24) und Glg. (3.27) kann man $\mathcal{T}$ schreiben als

$$\mathcal{T} = 1 + \sum_{\mu} 1_{S_{\mu}} \mathcal{T}_{\mu} 1_{S_{\mu}} = 1_I + \sum_{\mu} 1_{S_{\mu}} \underbrace{\sum_{\gamma} |\phi_{\gamma}^{\mu}\rangle \langle \tilde{p}_{\gamma}^{\mu}|}_{=: \mathcal{U}_{\mu}} 1_{S_{\mu}} \quad (3.28)$$

Dann gilt

$$\begin{aligned} \tilde{A} &= \mathcal{T}^+ A \mathcal{T} = 1_I A 1_I + \sum_{\mu} \underbrace{1_I A \mathcal{U}_{\mu}}_{=0} + \sum_{\mu} \underbrace{\mathcal{U}_{\mu}^+ A 1_I}_{=0} + \sum_{\mu, \mu'} \mathcal{U}_{\mu}^+ A \mathcal{U}_{\mu'} \\ &= (A - \sum_{\mu} 1_{S_{\mu}} A 1_{S_{\mu}}) + \sum_{\mu} \mathcal{U}_{\mu}^+ A \mathcal{U}_{\mu} \\ &= (A - \sum_{\mu; \gamma, \gamma'} |\tilde{p}_{\gamma}^{\mu}\rangle \langle \tilde{\phi}_{\gamma}^{\mu}| 1_{S_{\mu}} A 1_{S_{\mu}} |\tilde{\phi}_{\gamma'}^{\mu}\rangle \langle \tilde{p}_{\gamma'}^{\mu}|) + \sum_{\mu; \gamma, \gamma'} |\tilde{p}_{\gamma}^{\mu}\rangle \langle \phi_{\gamma}^{\mu}| 1_{S_{\mu}} A 1_{S_{\mu}} |\phi_{\gamma'}^{\mu}\rangle \langle \tilde{p}_{\gamma'}^{\mu}| \\ &= A + \sum_{\mu; \gamma, \gamma'} |\tilde{p}_{\gamma}^{\mu}\rangle \left\{ \langle \phi_{\gamma}^{\mu} | A | \phi_{\gamma'}^{\mu} \rangle - \langle \tilde{\phi}_{\gamma}^{\mu} | A | \tilde{\phi}_{\gamma'}^{\mu} \rangle \right\} \langle \tilde{p}_{\gamma'}^{\mu}| \quad (3.29) \\ &= A + A^1 - \tilde{A}^1 \quad (3.30) \end{aligned}$$

Der transformierte Operator setzt sich also aus drei Teilen zusammen: der erste (A) wirkt unmittelbar auf die Pseudowellenfunktion; der zweite und der dritte Teil enthalten die Projektoren, welche mit den Matrixelementen des Operators mit den echten (A^1) bzw. den pseudisierten ($\tilde{A}^1$) atomaren Wellenfunktionen gewichtet werden. Die Matrixelemente sind diagonal im Atomindex μ , Zweizentrenterme oder gar Dreizentrenterme treten nicht auf.

Operatoren mit der geforderten Lokalitätseigenschaft sind insbesondere der Operator $|\vec{r}\rangle\langle\vec{r}|$ und der Operator der kinetischen Energie, $-\nabla^2$.

⁵Für nichtlokale Operatoren siehe [Blö94].

$$H_{kin} = -\nabla^2 + \sum_{\mu; \gamma, \gamma'} |\tilde{p}_\gamma^\mu\rangle \left(\langle \phi_\gamma^\mu | -\nabla^2 | \phi_{\gamma'}^\mu \rangle - \langle \tilde{\phi}_\gamma^\mu | -\nabla^2 | \tilde{\phi}_{\gamma'}^\mu \rangle \right) \langle \tilde{p}_{\gamma'}^\mu | \quad (3.31)$$

3.5 Dichten

Die Dichte⁶ ist definiert als Erwartungswert des Operators $|\vec{r}\rangle\langle\vec{r}|$. Mithin besteht sie aus drei Anteilen:

$$n(\vec{r}) = \tilde{n}(\vec{r}) + n^1(\vec{r}) - \tilde{n}^1(\vec{r}) \quad (3.32)$$

wobei

$$\tilde{n}(\vec{r}) = \sum_{\vec{k}\nu} f_{\vec{k}\nu} \langle \tilde{\psi}_{\vec{k}\nu} | \vec{r} \rangle \langle \vec{r} | \tilde{\psi}_{\vec{k}\nu} \rangle \quad (3.33)$$

$$n^1(\vec{r}) = \sum_{\mu} n^{1\mu}(\vec{r} - \vec{\tau}_\mu) = \sum_{\mu} n^{1\mu}(\vec{r}_\mu) \quad (3.34)$$

$$n^{1\mu}(\vec{r}_\mu) = \sum_{\vec{k}\nu} f_{\vec{k}\nu} \sum_{\gamma, \gamma'} \langle \tilde{\psi}_{\vec{k}\nu} | \tilde{p}_\gamma^\mu \rangle \langle \phi_\gamma^{\alpha(\mu)} | \vec{r}_\mu \rangle \langle \vec{r}_\mu | \phi_{\gamma'}^{\alpha(\mu)} \rangle \langle \tilde{p}_{\gamma'}^\mu | \tilde{\psi}_{\vec{k}\nu} \rangle \quad (3.35)$$

$$= \sum_{\gamma, \gamma'} \rho_{\gamma, \gamma'}^\mu \langle \phi_\gamma^{\alpha(\mu)} | \vec{r}_\mu \rangle \langle \vec{r}_\mu | \phi_{\gamma'}^{\alpha(\mu)} \rangle \quad (3.36)$$

$$\rho_{\gamma, \gamma'}^\mu = \sum_{\vec{k}\nu} f_{\vec{k}\nu} \langle \tilde{\psi}_{\vec{k}\nu} | \tilde{p}_\gamma^\mu \rangle \langle \tilde{p}_{\gamma'}^\mu | \tilde{\psi}_{\vec{k}\nu} \rangle \quad (3.37)$$

$f_{\vec{k}\nu}$ sind die durch die Fermi-Dirac-Verteilung bestimmten Besetzungszahlen der Zustände (s.u.), gewichtet mit dem Faktor $\omega_{\vec{k}}$ der k-Punkt-Integration in der Brillouinzone.

Völlig analoge Formeln gelten für $\tilde{n}^1(\vec{r})$:

$$\tilde{n}^1(\vec{r}) = \sum_{\mu} \tilde{n}^{1\mu}(\vec{r} - \vec{\tau}_\mu) = \sum_{\mu} \tilde{n}^{1\mu}(\vec{r}_\mu) \quad (3.38)$$

$$\tilde{n}^{1\mu}(\vec{r}_\mu) = \sum_{\vec{k}\nu} f_{\vec{k}\nu} \sum_{\gamma, \gamma'} \langle \tilde{\psi}_{\vec{k}\nu} | \tilde{p}_\gamma^\mu \rangle \langle \tilde{\phi}_\gamma^{\alpha(\mu)} | \vec{r}_\mu \rangle \langle \vec{r}_\mu | \tilde{\phi}_{\gamma'}^{\alpha(\mu)} \rangle \langle \tilde{p}_{\gamma'}^\mu | \tilde{\psi}_{\vec{k}\nu} \rangle \quad (3.39)$$

$$= \sum_{\gamma, \gamma'} \rho_{\gamma, \gamma'}^\mu \langle \tilde{\phi}_\gamma^{\alpha(\mu)} | \vec{r}_\mu \rangle \langle \vec{r}_\mu | \tilde{\phi}_{\gamma'}^{\alpha(\mu)} \rangle \quad (3.40)$$

Die tatsächliche Dichte n ist außerhalb der Kugeln durch die Dichte $\tilde{n}$ gegeben, welche wie üblich durch das Betragsquadrat der Pseudowellenfunktionen $\tilde{\psi}_{\vec{k}\nu}$ bestimmt ist. In den Kugeln jedoch kommt eine Augmentationsladung $n^{1\mu} - \tilde{n}^{1\mu}$ hinzu.

⁶Falls nichts anderes gesagt wird, ist stets nur die Valenzladungsdichte gemeint, ohne daß das durch eine besondere Kennzeichnung verdeutlicht würde.

Es gilt in jeder Kugel S_μ um ein Atom μ :

$$n(\vec{r}) = n^{1\mu}(\vec{r}) \quad ; \quad \tilde{n}(\vec{r}) = \tilde{n}^{1\mu}(\vec{r}) \quad (3.41)$$

und im Zwischenraum I gilt:

$$n(\vec{r}) = \tilde{n}(\vec{r}) \quad ; \quad n^1(\vec{r}) = \tilde{n}^1(\vec{r}) \quad (3.42)$$

Die Dichte $\tilde{n}$ einerseits und die Dichten $n^{1\mu}$ und $\tilde{n}^{1\mu}$ andererseits werden in unterschiedlichen Darstellungen gespeichert; die obigen Identitäten erlauben es, Berechnungen für beide Repräsentationen durchzuführen, ohne eine explizite Umwandlung durchführen zu müssen.

Details und Formeln Die Dichte $\tilde{n}$ wird im Ortsraum aus den Wellenfunktionen $\tilde{\psi}_{\vec{k}\nu}(\vec{r})$ bestimmt. Die Besetzungszahlen werden durch die Fermi-Verteilung gegeben:

$$\tilde{f}(\varepsilon_{\vec{k}\nu}, T) = \frac{1}{e^{\frac{\varepsilon_{\vec{k}\nu} - \varepsilon_F}{kT}} + 1} \quad (3.43)$$

Obwohl die Dichtefunktionaltheorie nur für $T = 0$ gilt, werden aus numerischen Gründen die Besetzungszahlen für eine kleine endliche Temperatur berechnet (vgl. [Ric96]). Der Einfachheit halber werden die Besetzungszahlen zusammengefaßt mit der Gewichtung der k-Punkt Summation und der Spinartung bei nicht-spinpolarisierter Rechnung

$$f_{\vec{k}\nu} = 2 \tilde{f}(\varepsilon_{\vec{k}\nu}, T) \omega_{\vec{k}} \quad (3.44)$$

Die Fouriertransformierte der Ladungsdichte wird definiert durch⁷

$$\tilde{n}(\vec{r}) = \sum_{\vec{G}} \tilde{n}(\vec{G}) e^{-i\vec{G}\vec{r}} \quad (3.45)$$

$$\tilde{n}(\vec{G}) = \frac{1}{\Omega} \int_{\Omega} \tilde{n}(\vec{r}) e^{i\vec{G}\vec{r}} \quad (3.46)$$

Die Dichten $n^{1\mu}$ und $\tilde{n}^{1\mu}$ werden nach Drehimpulskomponenten zerlegt auf einem

⁷Diese Definition deckt sich mit der Implementation, unterscheidet sich aber von [Ric96], [Eng95] durch einen Faktor Ω .

radialen Gitter berechnet:

$$n^{1\mu}(\vec{r}_\mu) = \sum_{L''} n_{L''}^{1\mu}(r_\mu) Y_{L''}(\hat{r}_\mu) \quad (3.47)$$

$$n_{L''}^{1\mu}(r_\mu) = \sum_{\gamma, \gamma'} C(L', L, L'') \rho_{\gamma, \gamma'}^\mu \frac{1}{r_\mu^2} \phi_{l, i}^{*\alpha(\mu)}(r_\mu) \phi_{l', i'}^{\alpha(\mu)}(r_\mu) \quad (3.48)$$

$$C(L, L', L'') = \int d\Omega Y_L Y_{L'} Y_{L''} \quad (3.49)$$

Dabei sind Y_L reelle Kugelflächenfunktionen und $C(L, L', L'')$ die daraus bestimmten *R-Gaunt-Koeffizienten*. Wenn man nur endlich viele *partial waves* benutzt, ist die Anzahl der auftretenden Drehimpulskomponenten L'' der Dichte begrenzt; die höchste auftretende Drehimpulskomponente l'' der Dichte kann höchstens doppelt so groß sein wie die höchste unter den *partial waves*.

Die in der Ladungsdichtematrix $\rho_{\gamma, \gamma'}^\mu$ auftretenden Produkte $\langle \tilde{\psi}_{\vec{k}\nu} | \tilde{p}_\gamma^\mu \rangle$ werden im reziproken Raum berechnet (siehe Gleichung 3.145).

3.6 Das Energiefunktional

Das Energiefunktional setzt sich zusammen aus kinetischer Energie, Austausch-Korrelations-energie und elektrostatischer Energie.

3.6.1 Kinetische Energie

Die kinetische Energie ist der Erwartungswert des Operators $-\nabla^2$, also nach dem oben gesagten:

$$E_{kin} = \sum_{\vec{k}, \nu} f_{\vec{k}, \nu} \langle \tilde{\psi}_{\vec{k}, \nu} | -\nabla^2 | \tilde{\psi}_{\vec{k}, \nu} \rangle + \sum_{\mu, \gamma, \gamma'} \rho_{\gamma, \gamma'}^\mu D_{kin; \gamma, \gamma'}^{\alpha(\mu)} \quad (3.50)$$

wobei

$$D_{kin; \gamma, \gamma'}^{\alpha(\mu)} = \langle \phi_\gamma^\mu | -\nabla^2 | \phi_{\gamma'}^\mu \rangle - \langle \tilde{\phi}_\gamma^\mu | -\nabla^2 | \tilde{\phi}_{\gamma'}^\mu \rangle. \quad (3.51)$$

Details und Formeln Im Prinzip kann man den ersten Anteil der kinetischen Energie im reziproken Raum berechnen

$$\langle \tilde{\psi}_{\vec{k}\nu} | -\nabla^2 | \tilde{\psi}_{\vec{k}\nu} \rangle = \sum_{\vec{G}} \left(\vec{k} + \vec{G} \right)^2 c_{\vec{k}\nu}^*(\vec{G}) c_{\vec{k}\nu}(\vec{G})$$

und die beiden anderen unmittelbar unter Benutzung der Matrix $D_{\gamma, \gamma'}^{kin}$. Praktisch wird jedoch die kinetische Energie ersetzt unter Benutzung der Eigenwerte des Hamiltonoperators (s.u.).

3.6.2 Austausch-Korrelations-Energie

Die Austausch-Korrelations-Energie wird berechnet für die *gesamte* Elektronendichte (einschließlich core-Ladungsdichte n_c). Derzeit wird sie in der Lokale-Dichte-Näherung (LDA) ausgewertet.

$$E_{xc} = \int_{\Omega} \epsilon_{xc}[n_v + n_c](\vec{r}) (n_v(\vec{r}) + n_c(\vec{r})) d^3r \quad (3.52)$$

Diese Energie muß im Ortsraum ausgewertet werden. Dies möchte man auf dem regulären Ortsraumgitter (vgl. Abschnitt 3.1) durchführen. Jedoch ist dieses nicht geeignet, die echte Valenzdichte n_v und die echte core-Dichte n_c mit vertretbarem Aufwand hinreichend genau wiederzugeben, da diese im Rumpfbereich der Atome zu stark strukturiert sind. Im interstitiellen Bereich hingegen ist dies gut möglich. Dort stimmen nach Gleichung (3.42) echte und glatte Valenzdichte n_v und $\tilde{n}_v$ überein. Daher ersetzt man die Dichte n_v auf dem gesamten Volumen durch $\tilde{n}_v$ sowie n_c und $\tilde{n}_c$. Die so berechnete Energie $\tilde{E}_{xc}$ weist dann eine Abweichung durch die Modifikation des Integranden in den Augmentationsgebieten auf; da man auf Grund der Identitäten (3.41) die Dichten in den Augmentationsgebieten auch auf dem jeweiligen radialen Gitter kennt, kann man die nötige Korrektur $E_{xc}^1 - \tilde{E}_{xc}^1$ dort berechnen.

$$E_{xc} = \tilde{E}_{xc} + E_{xc}^1 - \tilde{E}_{xc}^1 \quad (3.53)$$

wobei

$$\tilde{E}_{xc} = \int_{\Omega} \epsilon_{xc}[\tilde{n}_v + \tilde{n}_c] (\tilde{n}_v + \tilde{n}_c) d^3r \quad (3.54)$$

$$E_{xc}^1 = \sum_{\mu} \int_{S_{\mu}} \epsilon_{xc}[n^{1\mu} + n_c^{\mu}] (n^{1\mu} + n_c^{\mu}) d^3r \quad (3.55)$$

$$\tilde{E}_{xc}^1 = \sum_{\mu} \int_{S_{\mu}} \epsilon_{xc}[\tilde{n}^{1\mu} + \tilde{n}_c^{\mu}] (\tilde{n}^{1\mu} + \tilde{n}_c^{\mu}) d^3r \quad (3.56)$$

Dabei kann $\tilde{n}_c^{\mu}(\vec{r})$ eine partielle core-Dichte sein, die außerhalb eines gewählten Radius mit der echten core-Dichte n_c^{μ} übereinstimmt. Die *partial-core*-Dichte kompensiert zum einen den Fehler, der dadurch entsteht, daß die core-Dichte nicht strikt im Augmentationsgebiet lokalisiert ist; zum anderen werden auch numerische Fehler durch die imperfekte Übereinstimmung von $\tilde{n}$ und $\tilde{n}^1$ reduziert.

Details und Formeln Die Austausch-Korrelations-Energie muß im Ortsraum berechnet werden. Das erste Integral (3.54) wird auf dem kartesischen Gitter aus-

geführt, das durch die diskrete Fast-Fourier Transformation vorgegeben ist. Die beiden anderen (3.55 und 3.56) werden auf radialen Gittern ausgewertet; dies geschieht analog zur in Abschnitt 3.9 beschriebenen Bestimmung des Austausch-Korrelations-Potentials.

3.6.3 Elektrostatische Energie

Die elektrostatische Energie bezogen auf eine Einheitszelle Ω setzt sich zusammen aus der Wechselwirkung der Atomkerne, der zwischen diesen und den Elektronen und derjenigen der Elektronen untereinander.

$$\begin{aligned}
 E_{estat} = & \frac{e^2}{2} \sum_{\mu, \mu', \vec{R}'}' \frac{Z^\mu Z^{\mu'}}{|\vec{r}_\mu - (\vec{r}_{\mu'} + \vec{R}')|} \\
 & + e^2 \int_{\Omega} (-n_c(\vec{r}) - n_v(\vec{r})) \sum_{\mu', \vec{R}'} \frac{Z^{\mu'}}{|\vec{r} - (\vec{r}_{\mu'} + \vec{R}')|} d^3r \\
 & + \frac{e^2}{2} \iint_{\Omega \infty} \frac{d^3r d^3r'}{|\vec{r} - \vec{r}'|} (-n_c(\vec{r}) - n_v(\vec{r})) \overline{(-n_c(\vec{r}') - n_v(\vec{r}'))} \quad (3.57)
 \end{aligned}$$

Einige Anmerkungen zur Notation: alle Ladungsdichten sind positiv, $e > 0$ ist der Betrag der Elementarladung; Vorzeichen der Ladung sind demnach explizit angeführt. Das gestrichene Summenzeichen berücksichtigt, daß Atomkerne nicht mit sich selbst wechselwirken, d.h. für $\mu = \mu'$ ist der Term $\vec{R}' = 0$ ausgeschlossen. Der erste Faktor eines Doppelintegrals ist als eine Funktion von $\vec{r}$, der zweite von $\vec{r}'$ zu verstehen. Eine Dichte ohne Atomindex μ (außer n_v) ist die Summe über atomare Dichten aller Atome in der Einheitszelle, z.B. $n_c(\vec{r}) = \sum_{\mu} n_c^{\alpha(\mu)}(\vec{r} - \vec{r}_\mu)$. Zur Vereinfachung mögen alle Atome mit ihren Augmentationsgebieten innerhalb der Einheitszelle Ω liegen (d.h. nicht über ihren Rand hinweg).

Die gitterperiodische Fortsetzung von Funktionen wird durch einen Überstrich markiert (z.B. $\overline{n_c}(\vec{r}) = \sum_{\vec{R}} n_c(\vec{r} - \vec{R})$). Z^μ bezeichnet die Kernladungszahl des Atomes μ . Mit $n_z^\mu = Z^\mu \delta(\vec{r} - \vec{r}_\mu)$ kann man schreiben:

$$E_{estat} = \frac{e^2}{2} \iint_{\Omega \infty} \frac{d^3r d^3r'}{|\vec{r} - \vec{r}'|} \left\{ (n_z - n_c - n_v) \overline{(n_z - n_c - n_v)} - \sum_{\mu} n_z^\mu n_z^\mu \right\} \quad (3.58)$$

So wie Glg. (3.57) eine Aufteilung der elektrostatischen Energie in Anteile darstellt, welche die Wechselwirkung zwischen Atomkernen und *gesamter* Elektronendichte beschreiben, so wird im folgenden eine Aufgliederung angestrebt, deren Terme die

Wechselwirkung zwischen *Atomrümpfen* und *Valenzelektronendichte* beschreiben. Wie in Glg. (3.57) wird auch in einer solchen Unterteilung jeder Term für sich divergent sein, da jeder Ladungsanteil wegen der periodischen Fortsetzung unendlich viel Ladung enthält. Eine Möglichkeit, dieses Problem vermeiden, besteht darin, für PAW-Atome eine innerhalb des Augmentationsgebietes lokalisierte, glatte, radialsymmetrische Referenz(valenz)ladungsdichte $n_{ref}^\mu(\vec{r})$ mit Gesamtladung Z_v^μ einzuführen, welche die positive Rumpfladung kompensiert. Dadurch erreicht man, daß die auftretenden Potentiale von neutralen Ladungsverteilungen erzeugt werden, und vermeidet Divergenzen. Dies funktioniert aber nur, wenn es für *alle* vorhandenen Atome auf diese Weise gehandhabt wird. Da diese Kompensation bei den bisher verwendeten Kleinman-Bylander (KB)-Pseudopotentialen nicht erfolgt, muß man die angesprochenen Divergenzen auf andere Weise behandeln, um Kompatibilität zu erreichen (vgl. Abschnitt 3.12).

Die Benutzung von Referenzladungen wird in den folgenden Herleitungen berücksichtigt; sie kann jedoch auch zwecks KB-Kompatibilität identisch verschwinden.

Die um die Referenzladungsdichte erweiterten Terme $((n_z - n_c - n_{ref}) + (n_{ref} - n_v))$ der elektrostatischen Energie E_{estat} spaltet man nun auf:

$$E_{estat} = \frac{e^2}{2} \iint_{\Omega_\infty} \frac{d^3r d^3r'}{|\vec{r} - \vec{r}'|} \left\{ (n_z - n_c - n_{ref}) \overline{(n_z - n_c - n_{ref})} - \sum_\mu n_z^\mu n_z^\mu \right\} \quad (3.59)$$

$$+ 2 \frac{e^2}{2} \iint_{\Omega_\infty} \frac{d^3r d^3r'}{|\vec{r} - \vec{r}'|} (-1)(n_v - n_{ref}) \overline{(n_z - n_c - n_{ref})} \quad (3.60)$$

$$+ \frac{e^2}{2} \iint_{\Omega_\infty} \frac{d^3r d^3r'}{|\vec{r} - \vec{r}'|} (n_v - n_{ref}) \overline{(n_v - n_{ref})} \quad (3.61)$$

$$= (E_{Ewald} + E_0) + E_{loc} + E_H \quad (3.62)$$

Die gesamte elektrostatische Energie ist damit aufgeteilt in einen Ewald-Term E_{Ewald} , der die Rumpf-Rumpf-Wechselwirkung und gegebenenfalls kompensierende Teile der Elektronenverteilung enthält, einen Hartree-Term E_H , welcher die Wechselwirkung der Valenzelektronen beschreibt und einen Term E_{loc} , welcher die Wechselwirkung zwischen Atomrümpfen und Valenzelektronen wiedergibt.

Die Ewald-Energie Der Term (3.59) hängt nicht von der Valenzdichte, evtl. aber von den Atompositionen ab. Die auftretenden Ladungsdichten sind eine Superposition von radial um die Atome befindlichen Ladungen. Trennt man die on-site Terme vom Rest ab, so handelt es sich um disjunkte, radialsymmetrische Ladungsvertei-

lungen, deren Wechselwirkung durch Punktladungen beschrieben werden. Für PAW-Atome mit Referenzladungsdichte verschwinden diese Punktladungen, wohingegen sie für KB-Atome⁸ einen endlichen Wert haben, nämlich die Valenzladung. Der nomenklatorischen Einfachheit halber werde ich den Wert (d.h. Valenzladung minus Referenzladung) als $\tilde{Z}_v^\mu$ bezeichnen.

$$\begin{aligned}
 (3.59) &= \frac{e^2}{2} \iint_{\Omega_\infty} \frac{d^3r \, d^3r'}{|\vec{r} - \vec{r}'|} \sum_{\mu, \mu', R'}' (n_z^\mu - n_c^\mu - n_{ref}^\mu) (n_z^{\mu' R'} - n_c^{\mu' R'} - n_{ref}^{\mu' R'}) \\
 &+ \frac{e^2}{2} \iint_{\Omega_\infty} \frac{d^3r \, d^3r'}{|\vec{r} - \vec{r}'|} \sum_{\mu} \{ (n_z^\mu - n_c^\mu - n_{ref}^\mu) (n_z^\mu - n_c^\mu - n_{ref}^\mu) - n_z^\mu n_z^\mu \} \\
 &= E_{Ewald} + E_0
 \end{aligned}$$

wobei

$$E_{Ewald} = \frac{e^2}{2} \sum_{\mu, \mu', R'}' \frac{\tilde{Z}_v^\mu \tilde{Z}_v^{\mu'}}{|\vec{r}_\mu - (\vec{r}_{\mu'} + R')|} \quad (3.63)$$

$$E_0 = \sum_{\mu} \frac{e^2}{2} \iint_{\Omega_\infty} \frac{d^3r \, d^3r'}{|\vec{r} - \vec{r}'|} (-2n_z^\mu + n_c^\mu + n_{ref}^\mu) (n_c^\mu + n_{ref}^\mu) = const. \quad (3.64)$$

$$\tilde{Z}_v^\mu = Z_v^\mu - \int_{S_\mu} n_{ref}^\mu(\vec{r}) \, d^3r \quad (3.65)$$

Die Hartree-Energie Auch die Hartree-Energie (Glg. (3.61)) möchte man gerne derart umformen, daß man – wie beim Austausch-Korrelations-Potential – ein Integral über eine glatte Funktion im ganzen Raum auszuführen hat und Korrekturen dazu im Augmentationsgebiet berechnet.

Sie läßt sich schreiben als

$$E_H = \frac{1}{2} \int_{\Omega} d^3r (n_v - n_{ref}) V_H[(n_v - n_{ref})] \quad (3.66)$$

mit dem elektrostatischen Potential

$$V_H[(n_v - n_{ref})](\vec{r}) = e^2 \int_{\infty} \frac{d^3r'}{|\vec{r} - \vec{r}'|} (n_v(\vec{r}') - n_{ref}(\vec{r}')) \quad (3.67)$$

Das Problem besteht darin, daß das elektrostatische Potential, das von einer Dichte erzeugt wird, ein nichtlokales Funktional dieser Dichte ist. Auch wenn man beim Ersetzen von $n(\vec{r})$ durch $\tilde{n}(\vec{r})$ nur die Ladungsdichte innerhalb der Augmentationsgebiete verändert, erzeugen die Dichten dennoch i. Allg. Potentiale, welche sich

⁸Der Begriff KB-Atom steht hier stellvertretend für Atome, die ein lokales Potential mit langreichweitigem Anteil haben.

auch im interstitiellen Gebiet unterscheiden. Jedoch ist das Potential, welches eine Ladungsverteilung in einer Kugel außerhalb derselben erzeugt, allein durch die Multipolmomente der Ladungsverteilung gegeben. Um das Potential im interstitiellen Gebiet nicht zu verändern, ersetzt man daher die echte Valenzdichte n_v durch die Summe aus der glatten Dichte $\tilde{n}_v$ und einer sogenannten Kompensationsladungsdichte $\hat{n}$, welche so konstruiert wird, daß sie in den Kugeln lokalisiert ist und zusammen mit der glatten Dichte $\tilde{n}_v$ bezüglich jeder Kugel dieselben Multipolmomente aufweist wie die echte Valenzdichte.⁹ Die Form dieser Kompensationsladungsdichte $\hat{n}$ kann so gewählt werden, daß sie sehr weich und daher auch gut fouriertransformierbar sind. Die Konstruktion dieser Kompensationsladung wird in Kapitel 3.11 genauer beschrieben. Die benötigten Multipolmomente Q_L^μ kann man leicht aus der Differenz der Dichten $n^{1\mu} - \tilde{n}^{1\mu}$ in den Kugeln gewinnen:

$$Q_L^\mu = \int r_\mu^L Y_L(\hat{r}_\mu) (n_v^{1\mu} - \tilde{n}_v^{1\mu}) d^3r_\mu \quad (3.68)$$

$$\hat{n}^\mu(\vec{r}) = \sum_L Q_L^\mu \hat{n}_L^\alpha(r) Y_L(\hat{r}_\mu) \quad (3.69)$$

Unter Benutzung dieser Kompensationsladungen schreibt man die Valenzdichte als

$$\begin{aligned} n_v &= \tilde{n}_v + \sum_\mu (n^{1\mu} - \tilde{n}^{1\mu}) \\ &= \tilde{n}_v + \hat{n} + \sum_\mu \underbrace{(n^{1\mu} - \tilde{n}^{1\mu} - \hat{n}^\mu)}_{\tilde{n}^\mu} = \tilde{n}_v + \hat{n} + \tilde{n}, \end{aligned} \quad (3.70)$$

wobei die Summe über μ hier nur PAW-Atome bezeichnet. Die Dichte $\tilde{n} + \hat{n}$ stimmt im interstitiellen Raum mit n_v überein und hat zudem bezüglich aller Augmentationsgebiete S_μ dieselben Multipolmomente wie n_v . Die Hilfsgröße $\tilde{n}^\mu$ hingegen ist in S_μ lokalisiert und – das ist nun entscheidend – erzeugt ein Potential, welches ebenfalls nur innerhalb dieses Gebietes von Null verschieden ist.

Damit kann man die Hartree-Energie E_H (Glg. (3.61)) umformen:

$$\begin{aligned} E_H &= \frac{e^2}{2} \iint_{\Omega \infty} \frac{d^3r d^3r'}{|\vec{r} - \vec{r}'|} (\tilde{n}_v + \hat{n} + \tilde{n} - n_{ref}) (\tilde{n}_v + \hat{n} + \tilde{n} - n_{ref}) \\ &= \tilde{E}_H \\ &+ \frac{e^2}{2} \iint_{\Omega \infty} \frac{d^3r d^3r'}{|\vec{r} - \vec{r}'|} 2 (\tilde{n}_v + \hat{n} - n_{ref}) \tilde{n} + \frac{e^2}{2} \iint_{\Omega \infty} \frac{d^3r d^3r'}{|\vec{r} - \vec{r}'|} \tilde{n} \tilde{n} \end{aligned} \quad (3.71)$$

mit

⁹Diese Definition unterscheidet sich (trotz gleicher Notation) von der in [Blö94].

$$\tilde{E}_H = \frac{e^2}{2} \iint_{\Omega_\infty} \frac{d^3r \, d^3r'}{|\vec{r} - \vec{r}'|} (\tilde{n}_v + \hat{n} - n_{ref}) \overline{(\tilde{n}_v + \hat{n} - n_{ref})} \quad (3.72)$$

$$= \frac{1}{2} \int_{\Omega} (\tilde{n}_v + \hat{n} - n_{ref}) \overline{\tilde{V}_H}(\vec{r}), \quad (3.73)$$

wobei das Hartree-Potential wie folgt definiert ist

$$\overline{\tilde{V}_H}[\tilde{n}_v + \hat{n} - n_{ref}](\vec{r}) = e^2 \int_{\infty} \frac{d^3r'}{|\vec{r} - \vec{r}'|} (\tilde{n}(\vec{r}') + \hat{n}(\vec{r}') - n_{ref}(\vec{r}')). \quad (3.74)$$

$\tilde{n}_\mu$ ist lokalisiert in S_μ und hat verschwindende Multipolmomente. Daher verschwindet das von $\tilde{n}_\mu$ erzeugte Potential außerhalb von S_μ ebenfalls. Deswegen reduzieren sich die Integrationsbereiche in Glg. (3.71) auf S_μ

$$(3.71) = \frac{e^2}{2} \sum_{\mu} \iint_{S_\mu S_\mu} \frac{d^3r \, d^3r'}{|\vec{r} - \vec{r}'|} (2(\tilde{n}^{1\mu} + \hat{n}^\mu - n_{ref}^\mu) + (n^{1\mu} - \tilde{n}^{1\mu} - \hat{n}^\mu)) (n^{1\mu} - \tilde{n}^{1\mu} - \hat{n}^\mu)$$

$$= E_H^1 - \tilde{E}_H^1$$

$$E_H^1 = \frac{e^2}{2} \sum_{\mu} \iint_{S_\mu S_\mu} \frac{d^3r \, d^3r'}{|\vec{r} - \vec{r}'|} (n^{1\mu} - n_{ref}^\mu) (n^{1\mu} - n_{ref}^\mu) \quad (3.75)$$

$$\tilde{E}_H^1 = \frac{e^2}{2} \sum_{\mu} \iint_{S_\mu S_\mu} \frac{d^3r \, d^3r'}{|\vec{r} - \vec{r}'|} (\tilde{n}^{1\mu} + \hat{n}^\mu - n_{ref}^\mu) (\tilde{n}^{1\mu} + \hat{n}^\mu - n_{ref}^\mu) \quad (3.76)$$

Insgesamt setzt sich die Hartree-Energie also aus drei Anteilen zusammen

$$E_H = \tilde{E}_H + E_H^1 - \tilde{E}_H^1, \quad (3.77)$$

welche getrennt berechnet werden.

Die Energie des lokalen Pseudopotentials Der Term E_{loc} (Gleichung 3.60) beschreibt die Wechselwirkung der Valenzelektronen mit den Atomrümpfen.

$$E_{loc} = \int_{\Omega} (n_v - n_{ref})(\vec{r}) \overline{V_{at}}(\vec{r}) \, d^3r \quad (3.78)$$

Wie schon bei der Berechnung der Austausch-Korrelationsenergie ersetzt man auch hier im Integranden die *echten* Funktionen, welche im Augmentationsgebiet scharf strukturiert sind, durch solche, die glatter sind, aber im interstitiellen Bereich mit diesen übereinstimmen. Die dadurch entstandene Abweichung kompensiert man durch Integrale, welche sich nur über die Augmentationsgebiete erstrecken.

Das elektrostatische Potential $V_{at}^\alpha(\vec{r})$ der Atomrümpfe der verschiedenen Atomtypen α

$$V_{at}^\alpha(\vec{r}) = -e^2 \int_{\infty} \frac{d^3 r'}{|\vec{r} - \vec{r}'|} (n_z^\alpha - n_c^\alpha - n_{ref}^\alpha) \quad (3.79)$$

ersetzt man zunächst jeweils durch ein glatteres (lokales) Pseudopotential

$$\tilde{V}_{at}^\alpha(\vec{r}) = -e^2 \int_{\infty} \frac{d^3 r'}{|\vec{r} - \vec{r}'|} (\tilde{n}_{at}^\alpha - n_{ref}^\alpha), \quad (3.80)$$

das außerhalb des Augmentationsgebietes mit ersterem übereinstimmt. Dies geschieht einmalig in einem externen Computerprogramm und wird zusammen mit der Generierung von Projektoren, *partial waves* und einigen anderen Größen als *Pseudopotentialerzeugung* bezeichnet (siehe Kapitel 4). Im Allgemeinen wird nur das Potential $\tilde{V}_{at}^\alpha(\vec{r})$ berechnet; die erzeugende Dichte $\tilde{n}_{at}^\alpha$ hingegen wird nicht explizit bestimmt, da sie in der Praxis nicht benötigt wird. Die Gittersumme $\overline{V_{at}}$ der echten Potentiale stimmt mit derjenigen $\overline{\tilde{V}_{at}}$ der glatten dann ebenfalls außerhalb der Augmentationsgebiete überein.

Ebenso ersetzt man in Glg. (3.78) die Dichte n_v , welche mit $\tilde{n} + \hat{n}$ außerhalb des Augmentationsgebietes übereinstimmt.¹⁰

$$E_{loc} = \tilde{E}_{loc} - \int_S (\tilde{n}_v + \hat{n} - n_{ref}) \overline{\tilde{V}_{at}} d^3 r + \int_S (n_v - n_{ref}) \overline{V_{at}} d^3 r \quad (3.81)$$

$$\tilde{E}_{loc} = \int_{\Omega} (\tilde{n}_v + \hat{n} - n_{ref}) \overline{\tilde{V}_{at}} d^3 r \quad (3.82)$$

Arbeitet man für *alle* Atome mit Referenzladungen (d.h. $\int (n_z^\mu - n_c^\mu - n_{ref}^\mu) d^3 r_\mu = 0$), so haben die Potentiale keinen langreichweitigen Anteil. Dann ist das Potential $\tilde{V}_{at}$ in der Nähe eines Atomes nur durch dessen eigenes Potential $\tilde{V}_{at}^\mu(r)$ gegeben (und ebenso für $\overline{V_{at}^\mu}$ und V_{at}^μ).

Auch wenn Atome ohne Referenzladung vorhanden sind, darf man in Glg. (3.81) die Potentiale $\overline{\tilde{V}_{at}}$ und $\overline{V_{at}}$ durch $\tilde{V}_{at}$ und V_{at} ersetzen. Man verändert zwar die einzelnen Terme, ihre Differenz jedoch bleibt gleich. Denn ein Potential in einem Gebiet ist festgelegt durch die Ladungsdichte im Innern und durch die Potentialwerte auf der Randfläche. Die Dichte im Innern ist jeweils gleich (für $\overline{V_{at}}$ und V_{at} bzw. $\overline{\tilde{V}_{at}}$ und $\tilde{V}_{at}$); in einer Kugel hat der Term, welcher den Einfluß der Randwerte beschreibt,

¹⁰Man kann hier n durch $\tilde{n} + \hat{n}$ oder durch $\tilde{n}$ ersetzen; beide Dichten sind im Interstitiellen mit n identisch. Die hier gewählte Variante hat den Vorteil, daß E^1 und $\tilde{E}^1$ invariant unter einer konstanten Verschiebung des elektrostatischen Potentials sind. Vgl. Kap. 3.12.

die Form $\sum_L r^l Y_L(\hat{r})$ (siehe Glg. (3.149)) und ist für glattes und echtes Potential identisch (da sie auf dem Kugelrand übereinstimmen). Aufgrund der Definition von $\hat{n}$ (Glg. (3.68)) heben sich diese Beiträge gegenseitig auf:

$$\int_0^{R_\mu} r_\mu^2 r_\mu^l (n_L^{1\mu}(r_\mu) - \tilde{n}_L^{1\mu}(r_\mu) - \hat{n}_L^{1\mu}(r_\mu)) dr_\mu = 0 \quad (3.83)$$

Also:

$$(E_{loc}^1 - \tilde{E}_{loc}^1) = \sum_\mu \int_{S_\mu} (n_v^1 - n_{ref}^\mu) V_\mu^{at} - (\tilde{n}_v^1 + \hat{n}^\mu - n_{ref}^\mu) \tilde{V}_\mu^{at} d^3 r_\mu \quad (3.84)$$

Die gesamte elektrostatische Energie ist mithin:

$$E_{estat} = E_0 + E_{Ewald} + \tilde{E}_H + (E_H^1 - \tilde{E}_H^1) + \tilde{E}_{loc} + (E_{loc}^1 - \tilde{E}_{loc}^1) \quad (3.85)$$

Details und Formeln Die konstante Energie E_0 wird bei der Erzeugung des Pseudopotentials berechnet.

Die Ewald-Summation wird nach der in [Eng95] beschriebenen Methode durchgeführt.

Die Energien $\tilde{E}_H$ und $\tilde{E}_{loc}$ werden im reziproken Raum berechnet:

$$\tilde{E}_H + \tilde{E}_{loc} = \Omega \sum_G \left(\frac{1}{2} \overline{V_H} [\tilde{n}_v + \hat{n} - n_{ref}] + \overline{V_{at}} \right) (\vec{G}) (\tilde{n}_v + \hat{n} - n_{ref})(\vec{G}) \quad (3.86)$$

Die übrigen Energien werden auf dem radialen Gitter um die Atome unter Benutzung des Hartree-Potentials (Glg. 3.149) bestimmt:

$$E_H^1 = \frac{1}{2} \sum_{\mu,L} \int_0^{R_\mu} V_H^{\mu,L} [n^{1\mu}](r_\mu) r_\mu^2 n_L^{1\mu}(r_\mu) dr_\mu \quad (3.87)$$

$$\tilde{E}_H^1 = \frac{1}{2} \sum_{\mu,L} \int_0^{R_\mu} V_H^{\mu,L} [\tilde{n}^{1\mu} + \hat{n}^{1\mu}](r_\mu) r_\mu^2 (\tilde{n}_L^{1\mu}(r_\mu) + \hat{n}_L^{1\mu}(r_\mu)) dr_\mu \quad (3.88)$$

Die Randbedingung des Potentials ($v_{0,H}^L(R_\mu)$ in Glg. 3.149) hat – wie bei $E_{loc}^1 - \tilde{E}_{loc}^1$ – keinen Einfluß auf $E_H^1 - \tilde{E}_H^1$ (vgl. Glg. (3.83)).

3.6.4 Resümee

Letztendlich hat man also folgenden Ausdruck für die Energie:

$$E = E_0 + E_{Ewald} + \tilde{E} + E^1 - \tilde{E}^1 \quad (3.89)$$

$$E_{Ewald} = \frac{e^2}{2} \sum_{\mu, \mu', \vec{R}'} \frac{\tilde{Z}_v^\mu \tilde{Z}_v^{\mu'}}{|\vec{\tau}_\mu - (\vec{\tau}_{\mu'} + \vec{R}')|} \quad (3.90)$$

$$E_0 = \sum_{\mu} \frac{e^2}{2} \iint_{\Omega} \frac{d^3 r}{|\vec{r} - \vec{r}'|} (-2n_z^\mu + n_c^\mu + n_{ref}^\mu)(n_c^\mu + n_{ref}^\mu) = const. \quad (3.91)$$

$$\begin{aligned} \tilde{E} &= \sum_{\vec{k}, \nu} f_{\vec{k}, \nu} \langle \tilde{\psi}_{\vec{k}, \nu} | -\nabla^2 | \tilde{\psi}_{\vec{k}, \nu} \rangle + \frac{e^2}{2} \iint_{\Omega} \frac{d^3 r}{|\vec{r} - \vec{r}'|} (\tilde{n}_v + \hat{n} - n_{ref}) (\tilde{n}_v + \hat{n} - n_{ref}) \\ &+ \int_{\Omega} (\tilde{n}_v + \hat{n} - n_{ref}) \overline{\tilde{V}_{at}} d^3 r + \int_{\Omega} \epsilon_{xc}[\tilde{n}_v + \tilde{n}_c] (\tilde{n}_v + \tilde{n}_c) d^3 r \end{aligned} \quad (3.92)$$

$$\begin{aligned} E^1 &= \sum_{\mu, \gamma, \gamma'} \rho_{\gamma, \gamma'}^\mu \langle \phi_\gamma^\mu | -\nabla^2 | \phi_{\gamma'}^\mu \rangle + \frac{e^2}{2} \sum_{\mu} \iint_{S_\mu} \frac{d^3 r}{|\vec{r} - \vec{r}'|} (n^{1\mu} - n_{ref}^\mu) (n^{1\mu} - n_{ref}^\mu) \\ &+ \sum_{\mu} \int_{S_\mu} (n^{1\mu} - n_{ref}^\mu) V_{at}^\mu d^3 r + \sum_{\mu} \int_{S_\mu} \epsilon_{xc}[n^{1\mu} + n_c^\mu] (n^{1\mu} + n_c^\mu) d^3 r \end{aligned} \quad (3.93)$$

$$\begin{aligned} \tilde{E}^1 &= \sum_{\mu, \gamma, \gamma'} \rho_{\gamma, \gamma'}^\mu \langle \tilde{\phi}_\gamma^\mu | -\nabla^2 | \tilde{\phi}_{\gamma'}^\mu \rangle \\ &+ \frac{e^2}{2} \sum_{\mu} \iint_{S_\mu} \frac{d^3 r}{|\vec{r} - \vec{r}'|} (\tilde{n}^{1\mu} + \hat{n}^{1\mu} - n_{ref}^\mu) (\tilde{n}^{1\mu} + \hat{n}^{1\mu} - n_{ref}^\mu) \\ &+ \sum_{\mu} \int_{S_\mu} (\tilde{n}^{1\mu} + \hat{n}^{1\mu} - n_{ref}^\mu) \tilde{V}_{at}^\mu d^3 r + \sum_{\mu} \int_{S_\mu} \epsilon_{xc}[\tilde{n}^{1\mu} + \tilde{n}_c^\mu] (\tilde{n}^{1\mu} + \tilde{n}_c^\mu) d^3 r \end{aligned} \quad (3.94)$$

Die Energie $\tilde{E} = \tilde{E}_{kin} + \tilde{E}_H + \tilde{E}_{loc} + \tilde{E}_{xc}$ enthält alle Anteile, die auf dem gesamten Volumen der Superzelle ausgewertet werden (d.h. entweder auf dem regulären Ortsraumgitter oder im reziproken Raum).

Die Energien E^1 und $\tilde{E}^1$ beinhalten die Korrekturen, welche durch die diversen Ersetzungen in den Augmentationsgebieten erforderlich sind. Sie werden auf den radialen Ortsraumgittern um die Atome herum berechnet.

Für die Ableitung des Hamiltonoperators wird diese Aufteilung benutzt.

3.7 Kohn-Sham-Gleichungen

Durch Variation der Energie (Glg. 3.89) nach $\tilde{\psi}_{k\nu}^*(\vec{r})$ erhält man die Kohn-Sham-Gleichungen. Dabei muß als Nebenbedingung die Orthonormierung der *echten* Wellenfunktionen beachtet werden:

$$1 = \langle \Psi_{k\nu} | \Psi_{k\nu} \rangle = \langle \mathcal{T} \tilde{\psi}_{k\nu} | \mathcal{T} \tilde{\psi}_{k\nu} \rangle = \langle \tilde{\psi}_{k\nu} | \mathcal{T}^+ \mathcal{T} | \tilde{\psi}_{k\nu} \rangle = \langle \tilde{\psi}_{k\nu} | \tilde{O} | \tilde{\psi}_{k\nu} \rangle \quad (3.96)$$

Dies geschieht – wie üblich – durch Einführen von Lagrange-Parameter $\varepsilon_{k\nu}$. Es resultiert ein nichttrivialer Überlappoperator $\tilde{O}$ und somit ein verallgemeinertes Eigenwertproblem:

$$\left(H - \varepsilon_{k\nu} \tilde{O} \right) | \tilde{\psi}_{k\nu} \rangle = 0 \quad (3.97)$$

Dessen Eigenzustände $| \tilde{\psi}_{k\nu} \rangle$ gehorchen der Orthogonalitätsrelation

$$\langle \tilde{\psi}_{k\nu} | \tilde{O} | \tilde{\psi}_{k'\nu'} \rangle = \delta_{\vec{k}, \vec{k}'} \delta_{\nu, \nu'} \quad (3.98)$$

Abhängig davon, ob man die Kohn-Sham-Gleichungen durch explizites Aufstellen der Matrizen (bezüglich der Basis aus Ebenen Wellen) löst oder durch ein iteratives Verfahren, benötigt man entweder die Matrixelemente $\langle \vec{k} + \vec{G} | M | \vec{k} + \vec{G}' \rangle$ des Hamiltonoperators und des Überlappoperator oder aber die Komponenten der Operator-Zustand-Produkte $\langle \vec{k} + \vec{G} | M | \tilde{\psi}_{k\nu} \rangle$.

3.8 Überlappmatrix

Der Überlappoperator $\tilde{O} = \mathcal{T}^+ \mathcal{T}$ ist gerade der transformierte Einsoperator, also:

$$\tilde{O} = \mathbf{1} + \sum_{\mu; \gamma, \gamma'} | \tilde{p}_{\gamma}^{\mu} \rangle q_{\gamma, \gamma'}^{\alpha(\mu)} \langle \tilde{p}_{\gamma'}^{\mu} | \quad (3.99)$$

mit

$$q_{\gamma, \gamma'}^{\alpha} = \langle \phi_{\gamma}^{\alpha(\mu)} | \phi_{\gamma'}^{\alpha(\mu)} \rangle - \langle \tilde{\phi}_{\gamma}^{\alpha(\mu)} | \tilde{\phi}_{\gamma'}^{\alpha(\mu)} \rangle = \delta_{L, L'} q_{l; i, i'}^{\alpha(\mu)} \quad (3.100)$$

und

$$q_{l; i, i'}^{\alpha(\mu)} = \int_0^{R_{\mu}} \phi_{l, i}^{\alpha(\mu)} \phi_{l, i'}^{\alpha(\mu)} - \tilde{\phi}_{l, i}^{\alpha(\mu)} \tilde{\phi}_{l, i'}^{\alpha(\mu)} dr_{\mu} \quad (3.101)$$

Die Gewichte $q_{\gamma, \gamma'}^{\alpha}$ des Überlappoperators sind unabhängig von der Dichte ebenso wie von den Positionen der Atome und mithin konstant. Für alle Atome eines Typs sind sie identisch. Sie werden bei der Pseudopotentialerzeugung generiert.

Details und Formeln Die Matrixelemente von $\tilde{O}$ bzw. die Vektorkomponenten von $\tilde{O}|\tilde{\psi}_{\vec{k}\nu}\rangle$ bezüglich der Basis aus Ebenen Wellen erhält man unter Entwicklung Ebener Wellen nach Drehimpulskomponenten (siehe Anhang A):

$$\begin{aligned} \langle \vec{k} + \vec{G} | \tilde{O} | \vec{k} + \vec{G}' \rangle &= \delta_{\vec{G}, \vec{G}'} + 4\pi \sum_{\alpha; l, i, i'} q_{l, i, i'}^{\alpha} \left(\sum_{\mu_{\alpha}} S^{\mu} (\vec{G} - \vec{G}') \right) \\ &\quad * \sum_m Y_L(\vec{k} + \vec{G}) \tilde{p}_{l, i}^{\alpha}(|\vec{k} + \vec{G}|) Y_L(\vec{k} + \vec{G}') \tilde{p}_{l, i'}^{\alpha}(|\vec{k} + \vec{G}'|) \end{aligned} \quad (3.102)$$

$$\langle \vec{k} + \vec{G} | \tilde{O} | \tilde{\psi}_{\vec{k}\nu} \rangle = \langle \vec{k} + \vec{G} | \tilde{\psi}_{\vec{k}\nu} \rangle + \sum_{\mu, L, i} \langle \vec{k} + \vec{G} | \tilde{p}_{L, i}^{\mu} \rangle \left(\sum_{i'} q_{l, i, i'}^{\alpha} \langle \tilde{p}_{L, i'}^{\mu} | \tilde{\psi}_{\vec{k}\nu} \rangle \right) \quad (3.103)$$

Dabei sind Y_L (reelle) Kugelflächenfunktionen.¹¹

Die Berechnung der Größen $\tilde{p}_{l, i'}^{\alpha}(|\vec{k} + \vec{G}'|)$, $\langle \vec{k} + \vec{G} | \tilde{p}_{\gamma}^{\mu} \rangle$ und $\langle \tilde{p}_{L, i'}^{\mu} | \tilde{\psi}_{\vec{k}\nu} \rangle$ erfolgt nach Gleichung (3.143) bis (3.145).

Die Matrixelemente des Überlappoperators ändern sich nur bei Bewegung der Atome; während des elektronischen Selbstkonsistenzprozesses bleiben sie konstant.

3.9 Ableitung des Hamiltonoperators

Den effektiven Einteilchen-Hamiltonoperator der Kohn-Sham-Gleichungen erhält man als Ableitung der Energie nach dem Dichteoperator $\tilde{\rho} = \sum_{\vec{k}, \nu} f_{\vec{k}, \nu} |\tilde{\psi}_{\vec{k}\nu}\rangle \langle \tilde{\psi}_{\vec{k}\nu}|$ (siehe [Blö94]; die Variation der Energie nach $\tilde{\psi}_{\vec{k}\nu}^*(\vec{r})$ führt – weniger elegant – zum selben Ergebnis).

Für die kinetische Energie ergibt sich unmittelbar:

$$\frac{\partial E_{kin}}{\partial \tilde{\rho}} = -\nabla^2 + \sum_{\mu; \gamma, \gamma'} |\tilde{p}_{\gamma}^{\mu}\rangle \left(\langle \phi_{\gamma}^{\mu} | -\nabla^2 | \phi_{\gamma'}^{\mu} \rangle - \langle \tilde{\phi}_{\gamma}^{\mu} | -\nabla^2 | \tilde{\phi}_{\gamma'}^{\mu} \rangle \right) \langle \tilde{p}_{\gamma'}^{\mu} | \quad (3.104)$$

Zur Ableitung der übrigen Terme des Hamiltonoperators betrachtet man $E = \tilde{E} + E^1 - \tilde{E}^1$ als Funktion von $\tilde{n}$, $n^{1\mu}$, $\tilde{n}^{1\mu}$ und Q_L^{μ} . Die Multipolmomente Q_L^{μ} sind ihrerseits

¹¹Die m-Summation der Kugelflächenfunktionen läßt sich mittels Glg. (A.35) auf Legendre-Polynome zurückführen.

gemäß Glg. (3.68) Funktionale von $n^{1\mu}, \tilde{n}^{1\mu}$. Demnach ergibt sich:

$$\frac{\partial(E - E_{kin})}{\partial \tilde{\rho}} = \int_{\Omega} \frac{\partial E}{\partial \tilde{n}} \frac{\partial \tilde{n}}{\partial \tilde{\rho}} d^3 r \quad (3.105)$$

$$+ \sum_{\mu} \int_{S_{\mu}} \left(\frac{\partial E}{\partial n^{1\mu}} + \sum_L \frac{\partial E}{\partial Q_L^{\mu}} \frac{\partial Q_L^{\mu}}{\partial n^{1\mu}} \right) \frac{\partial n^{1\mu}}{\partial \tilde{\rho}} d^3 r \quad (3.106)$$

$$+ \sum_{\mu} \int_{S_{\mu}} \left(\frac{\partial E}{\partial \tilde{n}^{1\mu}} + \sum_L \frac{\partial E}{\partial Q_L^{\mu}} \frac{\partial Q_L^{\mu}}{\partial \tilde{n}^{1\mu}} \right) \frac{\partial \tilde{n}^{1\mu}}{\partial \tilde{\rho}} d^3 r \quad (3.107)$$

Die Rechnung gestaltet sich etwas länglich; daher sind im nächsten Absatz einige Zwischenergebnisse aufgeführt.

3.9.1 Formelsammlung

$$\frac{\partial \tilde{n}}{\partial \tilde{\rho}} = |\vec{r}\rangle \langle \vec{r}| \quad (3.108)$$

$$\begin{aligned} \tilde{V}_{\text{lok}}(\vec{r}) &= \frac{\partial E}{\partial \tilde{n}(\vec{r})} \\ &= \overline{V_H}[\tilde{n}_v + \hat{n} - n_{ref}](\vec{r}) + \overline{\tilde{V}_{at}}(\vec{r}) + V_{xc}[\tilde{n}_v + \tilde{n}_c](\vec{r}) \end{aligned} \quad (3.109)$$

$$\frac{\partial n^{1\mu}(\vec{r})}{\partial \tilde{\rho}} = \sum_{\gamma, \gamma'} |\tilde{p}_{\gamma}^{\mu}\rangle \phi_{\gamma}^{\mu}(\vec{r}) \phi_{\gamma'}^{\mu}(\vec{r}) \langle \tilde{p}_{\gamma'}^{\mu}| \quad (3.110)$$

$$V^{1\mu} = \frac{\partial E}{\partial n_v^{1\mu}(\vec{r})} = V_H[n_v^{1\mu} - n_{ref}^{\mu}](\vec{r}) + V_{at}^{\mu} + V_{xc}[n_v^{1\mu} + n_c^{\mu}](\vec{r}) \quad (3.111)$$

$$-\tilde{V}^{1\mu} = -\frac{\partial E}{\partial \tilde{n}_v^{1\mu}(\vec{r})} = V_H[\tilde{n}_v^{1\mu} + \hat{n}^{\mu} - n_{ref}^{\mu}](\vec{r}) + \tilde{V}_{at}^{\mu} + V_{xc}[\tilde{n}_v^{1\mu} + \tilde{n}_c^{\mu}](\vec{r}) \quad (3.112)$$

Die Ableitung nach den Multipolmomenten Q_L^{μ} liefert einen Term, welcher sich als Randbedingung des elektrostatischen Potentials herausstellt. Er bewirkt, daß sich die Potentiale innerhalb der Augmentationsgebiete an den Grenzflächen glatt an

jenes außerhalb anschließen (siehe Abschnitt 3.11).

$$V_0^\mu(\vec{r}) = \sum_L \frac{\partial E}{\partial Q_L^\mu} \frac{\partial Q_L^\mu}{\partial n^{1\mu}(\vec{r})} = \sum_L \frac{\partial E}{\partial Q_L^\mu} r_\mu^l Y_L(\vec{r}_\mu) \quad (3.113)$$

$$\begin{aligned} \frac{\partial E}{\partial Q_L^\mu} &= \int_\Omega \frac{\partial \tilde{E}}{\partial \hat{n}} \frac{\partial \hat{n}}{\partial Q_L^\mu} d^3r - \int_S \frac{\partial \tilde{E}^1}{\partial \hat{n}^1} \frac{\partial \hat{n}^1}{\partial Q_L^\mu} d^3r \\ &= \int_\Omega d^3r \left(\overline{V_H} [\tilde{n}_v + \hat{n} - n_{ref}] + \overline{\tilde{V}_{at}} \right) \hat{n}_l^\alpha (|\vec{r} - \vec{r}_\mu|) Y_L(\vec{r} - \vec{r}_\mu) \end{aligned} \quad (3.114)$$

$$- \int_{S_\mu} d^3r \left(V_H [\tilde{n}_v^{1\mu} + \hat{n}^\mu - n_{ref}^{1\mu}] (\vec{r}_\mu) - \tilde{V}_{at}^\mu(\vec{r}_\mu) \right) \hat{n}_l^{1\alpha}(r_\mu) Y_L(\hat{r}_\mu) d^3r_\mu \quad (3.115)$$

wobei

$$\frac{\partial \hat{n}}{\partial Q_L^\mu} = \hat{n}_l^\alpha(\vec{r}_\mu) Y_L(\hat{r}_\mu) \quad (3.116)$$

benutzt wurde.

3.9.2 Resümee

Der resultierende Hamiltonoperator setzt sich – im Sinne der PAW-Methode – aus drei Anteilen zusammen, deren erster, welcher unmittelbar auf die Wellenfunktionen $\tilde{\psi}_{\vec{k}\nu}$ wirkt, aus dem Operator der kinetischen Energie und einem lokalen Potential besteht, während die Differenz der beiden anderen, welche auf die *partial waves* wirken, sich im Sinne der Pseudopotentialtheorie als nichtlokales Potential auffassen läßt.

$$\begin{aligned} H &= \tilde{H} + \underbrace{H^1 - \tilde{H}^1}_{= -\nabla^2 + \tilde{V}_{\text{lok}} + V_{NL}} \\ &= \overbrace{-\nabla^2 + \tilde{V}_{\text{lok}}} + V_{NL} \end{aligned} \quad (3.117)$$

wobei

$$\tilde{V}_{\text{lok}} = \overline{\tilde{V}_H} + \overline{\tilde{V}_{at}} + \tilde{V}_{xc} \quad (3.118)$$

$$V_{NL} = \sum_{\mu; \gamma, \gamma'} |\tilde{p}_\gamma^\mu\rangle D_{\gamma, \gamma'}^\mu \langle \tilde{p}_{\gamma'}^\mu| \quad (3.119)$$

$$D_{\gamma, \gamma'}^\mu = \langle \phi_\gamma^\mu | -\nabla^2 + V^{1\mu} | \phi_{\gamma'}^\mu \rangle - \langle \tilde{\phi}_\gamma^\mu | -\nabla^2 + \tilde{V}^{1\mu} | \tilde{\phi}_{\gamma'}^\mu \rangle \quad (3.120)$$

$$V^{1\mu}(\vec{r}) = V_H^\mu(\vec{r}) + V_{xc}^\mu(\vec{r}) + V_{at}^{\alpha(\mu)}(\vec{r}) + V_0^\mu(\vec{r}) \quad (3.121)$$

$$\tilde{V}^{1\mu}(\vec{r}) = \tilde{V}_H^\mu(\vec{r}) + \tilde{V}_{xc}^\mu(\vec{r}) + \tilde{V}_{at}^{\alpha(\mu)}(\vec{r}) + V_0^\mu(\vec{r}) \quad (3.122)$$

Die Matrixelemente der kinetischen Energie $D_{kin;\gamma,\gamma'}^{\alpha(\mu)} = \langle \phi_\gamma^\mu | -\nabla^2 | \phi_{\gamma'}^\mu \rangle - \langle \tilde{\phi}_\gamma^\mu | -\nabla^2 | \tilde{\phi}_{\gamma'}^\mu \rangle$ und die Potentiale $V_{at}^{\alpha(\mu)}(\vec{r})$ und $\tilde{V}_{at}^{\alpha(\mu)}(\vec{r})$ sind konstant und werden zu Programmbeginn für jeden Atomtyp eingelesen. Die übrigen Bestandteile der Potentiale $V^{1\mu}(\vec{r})$ und $\tilde{V}^{1\mu}(\vec{r})$ hängen von der aktuellen Dichte ab; mithin ändern sich auch die Matrixelemente $D_{\gamma,\gamma'}^\mu$ des nichtlokalen Potentials in jedem Selbstkonsistenzschritt. Ebenso verändern sie sich auch bei einer Verschiebung der Atome, da die Dichten $n^{1\mu}$ und $\tilde{n}^{1\mu}$ über die Dichtematrix $\rho_{\gamma,\gamma'}^\mu$ von den Projektoren abhängen, deren Lage durch die Atome bestimmt ist.

Die Bedeutung von $V_0^\mu(\vec{r})$ wird im Abschnitt 3.11 näher beleuchtet.

3.10 Berechnung des Hamiltonoperators

In diesem Abschnitt wird genau beschrieben, wie die Matrixelemente des Hamiltonoperators bestimmt werden. Z.T. wird auch der entsprechende Energiebeitrag aufgelistet, soweit das noch nicht in Abschnitt 3.6 geschehen ist.

3.10.1 Kinetische Energie

Der Operator der kinetischen Energie ist diagonal im reziproken Raum. Daher werden seine Matrixelemente und das Matrix-Vektor-Produkt dort ausgeführt.

$$\langle \vec{k} + \vec{G} | -\nabla^2 | \vec{k} + \vec{G}' \rangle = \left(\vec{k} + \vec{G} \right)^2 \delta_{\vec{G}, \vec{G}'} \quad (3.123)$$

$$\langle \vec{k} + \vec{G} | -\nabla^2 | \tilde{\psi}_{k\nu} \rangle = \left(\vec{k} + \vec{G} \right)^2 \underbrace{\langle \vec{k} + \vec{G} | \tilde{\psi}_{k\nu} \rangle}_{c_{k\nu}(\vec{G})} \quad (3.124)$$

3.10.2 Lokales Potential

Das lokale Potential $\tilde{V}_{\text{lok}}$ besteht aus dem lokalen Pseudopotential, dem Austausch-Korrelations-Potential und dem Hartree-Potential der Valenzelektronen. Es wird zunächst im Ortsraum gebildet.

$$\tilde{V}_{\text{lok}}(\vec{r}) = \overline{\tilde{V}_H}(\vec{r}) + \overline{\tilde{V}_{at}}(\vec{r}) + \tilde{V}_{xc}(\vec{r}) \quad (3.125)$$

Die Berechnung der einzelnen Potentiale im Ortsraum wird in den folgenden Abschnitten beschrieben.

Zum Bilden der Matrixelemente wird das lokale Potential in den reziproken Raum transformiert. Dies geschieht mittels der *Fast Fourier Transform* (FFT):

$$\begin{aligned}
\langle \vec{k} + \vec{G} | \tilde{V}_{\text{lok}} | \vec{k} + \vec{G}' \rangle &= \frac{1}{\Omega} \int_{\Omega} e^{i(\vec{G} - \vec{G}') \cdot \vec{r}} \tilde{V}_{\text{lok}}(\vec{r}) d^3r \\
&= \tilde{V}_{\text{lok}}(\vec{G} - \vec{G}') = FFT \left[\tilde{V}_{\text{lok}}(\vec{r}) \right] (\vec{G} - \vec{G}') \quad (3.126)
\end{aligned}$$

Für die Matrix-Vektor-Produkte hingegen werden zur Vermeidung von aufwendigen Faltungssummen die Wellenfunktionen in den Ortsraum transformiert.

$$\langle \vec{k} + \vec{G} | \tilde{V}_{\text{lok}} | \tilde{\psi}_{\vec{k}\nu} \rangle = FFT \left[\tilde{V}_{\text{lok}}(\vec{r}) FFT^{-1} \left[c_{\vec{k}\nu}(\vec{G}) \right] (\vec{r}) \right] (\vec{k} + \vec{G}) \quad (3.127)$$

Lokales Pseudopotential Das lokale Pseudopotential wird für jede Atomsorte α extern konstruiert (siehe Kapitel 4) und dann als radiale Funktion $\tilde{V}_{at}^{\alpha}(|\vec{r}|)$ im Ortsraum eingelesen. Die Fouriertransformierte des periodisch fortgesetzten Potentials

$$\tilde{V}_{at}^{\mu}(\vec{r}) = \sum_{\vec{R}} \tilde{V}_{at}^{\alpha(\mu)}(|\vec{r} - \vec{r}_{\mu} - \vec{R}|) \quad (3.128)$$

zerfällt als Gittersumme (vgl. Abschnitt 3.1) in Strukturfaktor $S^{\mu}(\vec{G})$ und Formfaktor:

$$\tilde{V}_{at}^{\mu}(\vec{G}) = S^{\mu}(\vec{G}) \overline{\tilde{V}_{at}^{\alpha(\mu)}}(|\vec{G}|) \quad (3.129)$$

wobei der Formfaktor

$$\overline{\tilde{V}_{at}^{\alpha(\mu)}}(|\vec{G}|) = \frac{4\pi}{\Omega} \int_0^{\infty} r^2 j_0(Gr) \tilde{V}_{at}^{\alpha}(r) dr \quad (3.130)$$

nur einmal am Programmanfang berechnet wird.¹² Die Transformation in den Ortsraum geschieht per FFT

$$\overline{\tilde{V}_{at}}(\vec{r}) = FFT^{-1} \left[\sum_{\mu} S^{\mu}(\vec{G}) \overline{\tilde{V}_{at}^{\alpha(\mu)}}(|\vec{G}|) \right] (\vec{r}) \quad (3.131)$$

Austausch-Korrelations-Potential Die Bestimmung des Austausch-Korrelations-Potentials $\tilde{V}_{xc}(\vec{r})$

$$\tilde{V}_{xc}(\vec{r}) = \frac{\delta \tilde{E}_{xc}}{\delta \tilde{n}(\vec{r})} = \epsilon_{xc}[\tilde{n}_v + \tilde{n}_c] + (\tilde{n}_v + \tilde{n}_c) \frac{d\epsilon_{xc}}{d\tilde{n}}[\tilde{n}_v + \tilde{n}_c] \quad (3.132)$$

erfolgt im Ortsraum, gegebenenfalls unter Einschluß einer partiellen Rumpfladungsdichte $\tilde{n}_c$. (Diese wird analog zum lokalen Pseudopotential im Ortsraum für jede Atomsorte eingelesen, im reziproken Raum mit entsprechendem Strukturfaktor superponiert und zurück fouriertransformiert.)

¹²Für Details siehe [Ric96].

Berechnung des Hartree-Potentials Das Hartree-Potential $\overline{\tilde{V}_H}(\vec{r})$ außerhalb der Kugeln ist gegeben durch (vgl. Glg. (3.72)):

$$\overline{\tilde{V}_H}(\vec{r}) = \frac{\delta \tilde{E}_H}{\delta \tilde{n}(\vec{r})} = e^2 \int_{\infty} d^3 r' \frac{\tilde{n}(\vec{r}') + \hat{n}(\vec{r}') - n_{ref}(\vec{r}')}{|\vec{r} - \vec{r}'|} \quad (3.133)$$

Für die Energie gilt also

$$\tilde{E}_H = \frac{1}{2} \int_{\Omega} d^3 r (\tilde{n}(\vec{r}) + \hat{n}(\vec{r}) - n_{ref}(\vec{r})) \overline{\tilde{V}_H}(\vec{r}) \quad (3.134)$$

Die Poissongleichung (A.53)

$$\Delta V_H(\vec{r}) = -4\pi e^2 n(\vec{r}) \quad (3.135)$$

ist im reziproken Raum diagonal. Aus diesem Grund ist es vorteilhaft, das Hartree-Potential $V_H(\vec{G})$ ebenfalls im reziproken Raum zu bestimmen.

Für die Fourierdarstellung des Hartree-Potentials $V_H(\vec{r})$ wird (analog zur Dichte, Glg. (3.45)) folgende Normierung gewählt:

$$\overline{V_H}(\vec{r}) = \sum_{\vec{G}} \overline{V_H}(\vec{G}) e^{-i\vec{G}\vec{r}} \quad (3.136)$$

$$\overline{V_H}(\vec{G}) = \frac{1}{\Omega} \int_{\Omega} d^3 r \overline{V_H}(\vec{r}) e^{i\vec{G}\vec{r}} \quad (3.137)$$

Mit der Ladungsdichte $n(\vec{G})$ erhält man das Hartree-Potential im reziproken Raum

$$\overline{V_H}(\vec{G}) = 4\pi e^2 \frac{n(\vec{G}) + \hat{n}(\vec{G}) - n_{ref}(\vec{G})}{G^2}, \quad (3.138)$$

welches (zusammen mit dem lokalen Pseudopotential) per FFT in den Ortsraum transformiert wird:

$$\overline{\tilde{V}_H}(\vec{r}) = FFT^{-1} \left[\overline{\tilde{V}_H}(\vec{G}) \right] (\vec{r}) \quad (3.139)$$

3.10.3 Nichtlokales Potential

Das nichtlokale Potential ist im reziproken Raum definiert. Es gilt:

$$\langle \vec{k} + \vec{G} | V_{NL} | \vec{k} + \vec{G}' \rangle = \sum_{\mu; \gamma, \gamma'} \langle \vec{k} + \vec{G} | \tilde{p}_{\gamma}^{\mu} \rangle D_{\gamma, \gamma'}^{\mu} \langle \tilde{p}_{\gamma}^{\mu} | \vec{k} + \vec{G}' \rangle \quad (3.140)$$

$$\langle \vec{k} + \vec{G} | V_{NL} | \tilde{\psi}_{\vec{k}\nu} \rangle = \sum_{\mu; \gamma} \langle \vec{k} + \vec{G} | \tilde{p}_{\gamma}^{\mu} \rangle \left(\sum_{\gamma'} D_{\gamma, \gamma'}^{\mu} \langle \tilde{p}_{\gamma'}^{\mu} | \tilde{\psi}_{\vec{k}\nu} \rangle \right) \quad (3.141)$$

$\langle \vec{k} + \vec{G} | \tilde{p}_\gamma^\mu \rangle$ und $\langle \tilde{p}_{\gamma'}^\mu | \tilde{\psi}_{\vec{k}\nu} \rangle$ werden im reziproken Raum berechnet.¹³

$$\tilde{p}_\gamma^{\alpha(\mu)}(\vec{r}) = \frac{1}{r} Y_L(\hat{r}) \tilde{p}_{l,i}^{\alpha(\mu)}(r) \quad (3.142)$$

$$\langle \vec{k} + \vec{G} | \tilde{p}_\gamma^\mu \rangle = i^l e^{i\vec{k}\vec{\tau}_\mu} S^\mu(\vec{G}) \tilde{p}_{l,i}^{\alpha(\mu)}(|\vec{k} + \vec{G}|) \sqrt{4\pi} Y_L(\vec{k} + \vec{G}) \quad (3.143)$$

$$\tilde{p}_{l,i}^{\alpha(\mu)}(|\vec{k} + \vec{G}|) = \sqrt{\frac{4\pi}{\Omega}} \int_0^{R_\mu} r j_l(|\vec{k} + \vec{G}|r) \tilde{p}_{l,i}^{\alpha(\mu)}(r) dr \quad (3.144)$$

$$\langle \tilde{\psi}_{\vec{k}\nu} | \tilde{p}_\gamma^\mu \rangle = \sum_{\vec{G}} \underbrace{\langle \tilde{\psi}_{\vec{k}\nu} | \vec{k} + \vec{G} \rangle}_{c_{\vec{k}\nu}^*(\vec{G})} \langle \vec{k} + \vec{G} | \tilde{p}_\gamma^\mu \rangle \quad (3.145)$$

Im Prinzip lassen sich diese Produkte auch im Ortsraum auswerten [KSPL91]. In den meisten Fällen ergibt sich ein besseres Skalierungsverhalten.

Von KB-Potentialen unterscheidet sich dieses Potential formal dadurch, daß die Matrix $D_{\gamma,\gamma'}^\mu$ nicht konstant, sondern dichteabhängig ist, i. Allg. nicht-diagonal ist (außer im Atomindex μ), von m abhängt und sich für verschiedene Atome desselben Atomtyps unterscheidet (im Gegensatz zur Überlappmatrix). Die Matrixelemente $D_{\gamma,\gamma'}^\mu$ werden berechnet gemäß:

$$\begin{aligned} D_{\gamma,\gamma'}^\mu &= D_{kin;\gamma,\gamma'}^\mu + D_{Hxc;\gamma,\gamma'}^\mu \quad (3.146) \\ D_{Hxc;\gamma,\gamma'}^\mu &= \langle \phi_\gamma^\mu | V^{1\mu} | \phi_{\gamma'}^\mu \rangle - \langle \tilde{\phi}_\gamma^\mu | \tilde{V}^{1\mu} | \tilde{\phi}_{\gamma'}^\mu \rangle \\ &= \sum_{L''} C(L, L', L'') D_{Hxc;l_i, l_{i'}}^{\mu L''} \\ D_{Hxc;l_i, l_{i'}}^{\mu L''} &= \int_0^{R_\mu} \left[\phi_{l,i}^\mu(r) V_{L''}^{1\mu}(r) \phi_{l',i'}^\mu(r) - \tilde{\phi}_{l,i}^\mu(r) \tilde{V}_{L''}^{1\mu}(r) \tilde{\phi}_{l',i'}^\mu(r) \right] dr \quad (3.147) \end{aligned}$$

Man beachte, daß der Integrand jenseits von R_μ verschwindet, da $V^{1\mu}(\vec{r})$ und $\tilde{V}^{1\mu}(\vec{r})$ sowie $\phi_{l,i}^\mu(r)$ und $\tilde{\phi}_{l,i}^\mu(r)$ dort übereinstimmen.

Während die Matrixelemente der kinetischen Energie

$$D_{kin;\gamma,\gamma'}^\mu = \langle \phi_\gamma^\mu | -\nabla^2 | \phi_{\gamma'}^\mu \rangle - \langle \tilde{\phi}_\gamma^\mu | -\nabla^2 | \tilde{\phi}_{\gamma'}^\mu \rangle \quad (3.148)$$

bei der Erzeugung des Pseudopotentials berechnet und dann eingelesen werden, hängt der zweite Anteil $D_{Hxc;\gamma,\gamma'}^\mu$ der Matrixelemente des Nichtlokalen Potentials von den Potentialen $V_{xc}^{\mu L}(r_\mu)$ und $V_H^{\mu L}(r_\mu)$ in den Kugeln ab.

¹³Die hiesige Normierung entspricht der Umsetzung im Programm und ist konsistent mit der Behandlung der KB-Projektoren. Diese deckt sich nicht mit der Beschreibung durch (4.117) in [Ric96] und (3.102) in [Eng95]. Jene erspart jedoch den lästigen Faktor $\sqrt{4\pi}$.

Der Faktor $e^{i\vec{k}\vec{\tau}_\mu}$ in Glg. (3.143) annulliert sich stets mit seinem komplex konjugierten und wird daher nicht ausgerechnet.

Die Bestimmung von Austausch-Korrelations-potential und –Energie in den Kugeln erfordert die Integration über Kugelflächenfunktionen und setzt deren Vollständigkeitsrelation voraus. Das hier zur Integration benutzte Gauss-Legendre Verfahren erfüllt durch die spezielle Wahl der Stützpunkte die Orthogonalitätsrelation der Kugelflächenfunktionen auch auf dem diskreten Gitter für die Winkelvariablen. Die Raumwinkelintegration $\int d\Omega$ geht dann über in eine gewichtete Summation $\sum_j w_j$.

- bestimme Stützstellen (r_i, Ω_j)
Die Punkte $\Omega_j = (\phi_j, \theta_j)$ werden so gewählt, daß die ϕ_j äquidistant sind und die θ_j Nullstellen der Legendre-Polynome.
- berechne $n^{1\mu}(r_i, \Omega_j) = \sum_L n_L^{1\mu}(r_i) Y_L(\Omega_j)$ $\forall(r_i, \Omega_j)$
- berechne $V_{xc}(n^1(r_i, \Omega_j))$ $\forall(r_i, \Omega_j)$
- berechne $V_{xc}^{\mu L}(r_i) = \sum_j w_j V_{xc}(n^1(r_i, \Omega_j)) Y_L(\Omega_j)$ $\forall r_i, L$
- berechne $E_{xc}^{1\mu} = \sum_i r_i^2 \left(\frac{dr}{di}\right)_i \left(\sum_L V_{xc}^L(r_i) n_{\mu,L}^1(r_i)\right)$ Die Radialintegration erfolgt nach der Simpson-Regel. Für lineare Gitter ist $\left(\frac{dr}{di}\right)_i$ konstant und durch den Abstand benachbarter Gitterpunkte gegeben.

Das Hartree-Potential in den Kugeln wird mittels der Methode der Greenschen Funktion berechnet.[Jac75, Wei81]

$$V_H^\mu[n](\vec{r}_\mu) = \sum_L V_H^{\mu,L}[n](r_\mu) Y_L(\hat{r}_\mu) \quad (3.149)$$

$$V_H^{\mu,L}[n](r_\mu) = \frac{4\pi e^2}{2l+1} \frac{1}{r_\mu^{l+1}} \int_0^{r_\mu} r'^2 r'^l n_L^{1\mu}(r') dr' \quad (3.150)$$

$$+ \frac{4\pi e^2}{2l+1} r_\mu^l \int_{r_\mu}^{R_\mu} r'^2 r'^{-(l+1)} n_L^{1\mu}(r') dr' \quad (3.151)$$

$$+ \left(\frac{r_\mu}{R_\mu}\right)^l v_{0,H}^{\mu,L}(R_\mu) \quad (3.152)$$

$$v_{0,H}^{\mu,L}(R_\mu) = V_H^{\mu,L}[n](R_\mu) - \frac{4\pi e^2}{2l+1} \frac{1}{R_\mu^{l+1}} \int_0^{R_\mu} r'^2 r'^l n_L^{1\mu}(r') dr' \quad (3.153)$$

Dabei gehen Randbedingungen über $v_{0,H}^{\mu,L}(R_\mu)$ ein; ist dieser Term null, so verschwindet das Potential für $r = \infty$.

Als radiales Gitter kann ein beliebiges lineares oder nichtlineares Gitter verwendet werden. Dies wird bei der Erzeugung des Pseudopotentials vorgegeben.

3.10.4 Zusammenfassung

Faßt man alle Terme zusammen, so kann man die Matrix des Hamiltonoperators für einen bestimmten k -Punkt schreiben als:

$$\begin{aligned}
 H_{\vec{G}, \vec{G}'}(\vec{k}) &= \langle \vec{k} + \vec{G} | -\nabla^2 + \tilde{V}_{lok} + V_{NL} | \vec{k} + \vec{G}' \rangle \\
 &= (\vec{k} + \vec{G})^2 \delta_{\vec{G}, \vec{G}'} + \tilde{V}_{lok}[n, \{\vec{\tau}_\mu\}](\vec{G} - \vec{G}') \\
 &+ \sum_{\mu; \gamma, \gamma'} \langle \vec{k} + \vec{G} | \tilde{p}_\gamma^\mu \rangle D_{\gamma, \gamma'}^\mu[n, \{\vec{\tau}_\mu\}] \langle \tilde{p}_\gamma^\mu | \vec{k} + \vec{G}' \rangle
 \end{aligned} \tag{3.154}$$

Sowohl das lokale Potential $\tilde{V}_{lok}$ als auch das nichtlokale Potential V_{NL} hängen von der Dichte n wie auch von den Atomkoordinaten $\{\vec{\tau}_\mu\}$ ab.

3.11 Die Pseudoladungsmethode

Das Hartree-Potential, das von einer Ladungsverteilung $n^1 - \tilde{n}^1$ in einer Kugel mit Radius R außerhalb derselben erzeugt wird, hängt nur von den Multipolmomenten der Ladungsverteilung in der Kugel ab.

$$V(\vec{r}) = \sum_L \frac{4\pi e^2}{2l+1} \frac{Q_L}{r^{l+1}} Y_L(\hat{r}) \tag{3.155}$$

Die Multipolmomente Q_L sind definiert als:

$$Q_L = \int_S Y_L(\hat{r}) r^l (n^1 - \tilde{n}^1)(\vec{r}) d^3r \tag{3.156}$$

Daher kann man die echte Ladungsdichte durch eine Pseudoladungsdichte mit denselben Multipolmomenten ersetzen, welche *weicher* ist und dadurch auch gut im reziproken Raum dargestellt werden kann. Die optimale Pseudoladung nach [Wei81] hat im Ortsraum folgende Gestalt:

$$\hat{n}(\vec{r}) = \sum_L Q_L \hat{n}_l(r) Y_L(\hat{r}) \tag{3.157}$$

$$\hat{n}_l(r) = \frac{(2l + 2N_l + 3)!!}{2^{N_l} N_l! (2l + 1)!!} \frac{1}{R^{l+3}} \left(\frac{r}{R}\right)^l \left(1 - \left(\frac{r}{R}\right)^2\right)^{N_l}, \tag{3.158}$$

im reziproken Raum¹⁴

$$\hat{n}(\vec{G}) = Q_L \hat{n}_l(\vec{G}) Y_L(\hat{G}) \quad (3.159)$$

$$\hat{n}_l(\vec{G}) = \frac{4\pi}{\Omega} \sum_L \frac{i^l (2l + 2N_l + 3)!!}{(2l + 1)!!} \frac{1}{R^l} \frac{j_{l+N_l+1}(GR)}{(GR)^{N_l+1}} \quad (3.160)$$

$$\hat{n}_l(\vec{G} = 0) = \delta_{l,0} \frac{4\pi}{\Omega} \quad (3.161)$$

Für $n_l(r)$ ist das l -te Multipolmoment auf eins normiert. Die Zahl N_l wird so bestimmt, daß die bestmögliche Fouriertransformierbarkeit bei gegebenem Kugelradius R und *cutoff*-Parameter G_{max} erreicht wird. Dazu wird das größte N_l gewählt, für das die erste Nullstelle der Besselfunktion $j_{l+N_l+1}(x)$ ungefähr gleich $(G_{max}R)$ ist. Für den interstitiellen Bereich I außerhalb der Kugeln ist die korrekte Fourierentwicklung des Potentials der periodisch fortgesetzten Ladungsverteilung bestimmt durch die Summe aus glatter Ladung $\tilde{n}$ und Pseudoladung $\hat{n}$:

$$\overline{V_H}[\tilde{n} + \hat{n}](\vec{G}) = 4\pi e^2 \frac{\tilde{n}(\vec{G}) + \hat{n}(\vec{G})}{G^2} \quad (3.162)$$

Das Potential der periodisch fortgesetzten Ladungsverteilung in einer Kugel läßt sich (auf einem radialen Gitter) mittels der Greenfunktion für die Kugel berechnen (vgl. Glg. (3.149)). Die Ladungsdichte innerhalb der Kugel ist bekannt, und die Werte des Potentials auf dem Kugelrand erhält man durch Fourierentwicklung von Glg. (3.162), da diese aufgrund ihrer Stetigkeit auch auf der Kugelschale noch gültig ist.

Betrachtet man nun $\frac{\partial E}{\partial Q_L^\mu}$ (Glg. 3.114) erneut und wertet das erste Integral nicht im reziproken, sondern im Ortsraum aus, so reduziert sich der Integrand (vgl. Glg. 3.149) auf

$$\left(\overline{V_H}^{\mu,L} + \overline{\tilde{V}_{at}}^{\mu,L} - V_H^{\mu,L} - \tilde{V}_{at}^{\mu,L} \right) \hat{n}_l(r_\mu) = \left(v_{0,H}^{\mu,L}(R_\mu) + v_{0,at}^{\mu,L}(R_\mu) \right) \left(\frac{r_\mu}{R_\mu} \right)^l \hat{n}_l(r_\mu) \quad (3.163)$$

und wegen $\int_0^{R_\mu} r^2 r^l \hat{n}_l(r) dr = 1$ erhält man

$$v_0^{\mu,L}(r) = \left(v_{0,H}^{\mu,L}(R_\mu) + v_{0,at}^{\mu,L}(R_\mu) \right) \left(\frac{r_\mu}{R_\mu} \right)^l \quad (3.164)$$

¹⁴Formel (28) in [Wei81] ist fehlerhaft abgedruckt.

D.h. der Term V_0 im Hamiltonoperator bewirkt, daß das Potential in der Kugel um jedes Atom μ passend an das Potential im Interstitiellen anschließt, indem er gerade die fehlenden Randterme darstellt.

Es stellt sich jedoch heraus, daß diese Art der Berechnung des Potentials zwar zu guten Gesamtenergien führt, die Eigenwerte jedoch weichen z.T. relativ stark von der Erwartung ab und zeigen zudem ein recht unsystematisches Verhalten bei Variation von Parametern. Berechnet man jedoch das erste Integral von Glg. (3.114) im reziproken Raum – so wie es mit dem entsprechenden Term in der Energie geschieht – so sind auch die Eigenwerte gut.

3.12 Die $G=0$ -Komponente

Hat man Atome ohne Referenzladung (z.B. KB-Atome) im System, so ist die Gesamtladung $\tilde{Z}_v = \int (\tilde{n}_v + \hat{n} - n_{ref}) d^3r$ nicht neutral. Setzt man diese Ladung periodisch fort, so ist die Ladung unendlich und folglich muß das Potential divergieren. Ebenso entspricht das Potential $\overline{\tilde{V}_{at}} = V_H [-\overline{\tilde{n}_{at}}]$ einer nicht neutralen Ladungsdichte $\tilde{n}_{at}$. Eine (formale) Entwicklung der Ladungsdichten im reziproken Raum nach kleinen Argumenten lautet¹⁵

$$\overline{\tilde{n}_v + \hat{n} - n_{ref}} = \frac{\tilde{Z}_v}{\Omega} + \frac{\beta}{4\pi e^2} G^2 + \dots \quad (3.165)$$

$$\overline{\tilde{n}_{at} + n_{ref}} = \frac{\tilde{Z}_v}{\Omega} + \frac{\alpha}{4\pi e^2} G^2 + \dots \quad (3.166)$$

Daher divergieren die $G = 0$ -Komponenten der jeweiligen Potentiale für sich betrachtet (vgl. Glg. 3.138), ihre Summe jedoch ist endlich. Benutzt man nur Atome mit Referenzladung, so sind die $G = 0$ -Komponenten dieser Potentiale zwar endlich, aber unbekannt.

$$\left(\overline{\tilde{V}_H} + \overline{\tilde{v}_{at}} \right) (G = 0) = \beta - \alpha \quad (3.167)$$

Um die Divergenzen der Einzelterme zu vermeiden, setzt man die $G = 0$ -Komponenten beider Potentiale auf Null.

$$\overline{\tilde{V}_H}(\vec{G}) = (1 - \delta_{\vec{G},0}) \overline{\tilde{V}_H}(\vec{G}) \quad , \quad \overline{\tilde{V}_{at}}(\vec{G}) = (1 - \delta_{\vec{G},0}) \overline{\tilde{V}_{at}}(\vec{G})$$

Dies bedeutet eine Verschiebung des gesamten Potentials V_{lok} um eine Konstante $V_{\beta\alpha} = \beta - \alpha$.

¹⁵Die Behandlung folgt [IZC79].

Potentiale in den Kugeln

Für die einzelnen Potentiale, z.B. $\ddot{\tilde{V}}_H$, hat man durch das Annulieren der G=0-Komponente (s.o.) zweierlei verändert: Erstens hat man die Ladung in den Kugeln verändert, d.h. es ist ein Potential, das nicht von $\tilde{n}_v + \hat{n} - n_{ref}$, sondern von $\tilde{n}_v + \hat{n} - n_{ref} - \frac{\tilde{Z}_v^{tot}}{\Omega}$ erzeugt wird.

$$\ddot{\tilde{V}}_H^{\mu L}(r) = \tilde{V}_H^{\mu L}(r) - \delta_{L,0} V_G(r) \quad (3.168)$$

$$\begin{aligned} V_G(r) &= \left(V_H^{\mu L} [-\tilde{Z}_v] - \frac{4\pi e^2}{2l+1} \frac{1}{R^{l+1}} \int_0^R r^2 r^l \sqrt{4\pi} (-\tilde{Z}_v) \right) \\ &= \sqrt{4\pi} 4\pi e^2 \tilde{Z}_v \frac{1}{6} (R^2 - r^2) \end{aligned} \quad (3.169)$$

Entsprechend erhält man

$$\ddot{\tilde{V}}_{at}^{\mu L}(r) = \tilde{V}_{at}^{\mu L}(r) + \delta_{L,0} V_G(r). \quad (3.170)$$

Das heißt aber für die Summe $\ddot{\tilde{V}}_H^{\mu L}(r) + \ddot{\tilde{V}}_{at}^{\mu L}(r)$ der Potentiale, wie sie in $\tilde{V}^1$ auftritt, daß sich diese Veränderungen kompensieren. (Genauso auch in V^1 .)

Zweitens ist das Potential im Interstitiellen um eine Konstante verschoben. Dadurch erfährt der Randbedingungsterm (Glg. 3.114) eine Verschiebung:

$$\frac{\partial \ddot{E}}{\partial Q_L^\mu} = \frac{\partial E}{\partial Q_L^\mu} - \int_\Omega V_{\beta\alpha} Y_{00} \hat{n}_0^{\alpha(\mu)} = \frac{\partial E}{\partial Q_L^\mu} - V_{\beta\alpha} \quad (3.171)$$

also

$$\ddot{V}_0^{\mu L} = V_0^{\mu L} - \delta_{L,0} V_{\beta\alpha} \quad (3.172)$$

und daher erfahren die Potentiale

$$\ddot{V}_L^{1\mu} = V_L^{1\mu} - \delta_{L,0} V_{\beta\alpha} \quad \text{und} \quad \ddot{\tilde{V}}_L^{1\mu} = \tilde{V}_L^{1\mu} - \delta_{L,0} V_{\beta\alpha} \quad (3.173)$$

ebenfalls eine konstante Verschiebung.

Hamiltonoperator

Das lokale Potential im Hamiltonoperator verschiebt sich um einen konstanten Wert $\ddot{\tilde{V}}_{lok} = \tilde{V}_{lok} - V_{\beta\alpha}$. Die Matricelemente des nichtlokalen Potentials V_{NL} verändern sich auf Grund der Verschiebung der Potentiale V^1 und $\tilde{V}^1$ in den Kugeln gemäß

$$\begin{aligned} \ddot{D}_{\gamma,\gamma'}^\mu &= \langle \phi_\gamma^\mu | \ddot{V}^{1\mu} | \phi_{\gamma'}^\mu \rangle - \langle \tilde{\phi}_\gamma^\mu | \ddot{\tilde{V}}^{1\mu} | \tilde{\phi}_{\gamma'}^\mu \rangle \\ &= D_{\gamma,\gamma'}^\mu - V_{\beta\alpha} q_{\gamma,\gamma'}^{\alpha(\mu)} \end{aligned} \quad (3.174)$$

und das Potential wird mithin zu

$$\ddot{V}_{NL} = V_{NL} - V_{\beta\alpha} \sum_{\mu;\gamma,\gamma'} |\tilde{p}_\gamma^\mu\rangle q_{\gamma,\gamma'}^{\alpha(\mu)} \langle \tilde{p}_{\gamma'}^\mu|. \quad (3.175)$$

Ist $|\tilde{\psi}_{\vec{k}\nu}\rangle$ Eigenzustand von $H = T + V_{loc} + V_{NL}$ mit Energie $\varepsilon_{\vec{k}\nu}$,

$$0 = (T + V_{loc} + V_{NL} - \varepsilon_{\vec{k}\nu} \tilde{O}) |\tilde{\psi}_{\vec{k}\nu}\rangle,$$

so ist $|\tilde{\psi}_{\vec{k}\nu}\rangle$ auch Eigenzustand von $\ddot{H}$ mit Energie $\ddot{\varepsilon}_{\vec{k}\nu} = \varepsilon_{\vec{k}\nu} - V_{\beta\alpha}$

$$\begin{aligned} (T + \ddot{V}_{loc} + \ddot{V}_{NL}) |\tilde{\psi}_{\vec{k}\nu}\rangle &= \left(T + V_{loc} + V_{NL} - V_{\beta\alpha} \left(\mathbf{1} + \sum_{\mu;\gamma,\gamma'} |\tilde{p}_\gamma^\mu\rangle q_{\gamma,\gamma'}^{\alpha(\mu)} \langle \tilde{p}_{\gamma'}^\mu| \right) \right) |\tilde{\psi}_{\vec{k}\nu}\rangle \\ &= \varepsilon_{\vec{k}\nu} \tilde{O} |\tilde{\psi}_{\vec{k}\nu}\rangle - V_{\beta\alpha} \tilde{O} |\tilde{\psi}_{\vec{k}\nu}\rangle = (\varepsilon_{\vec{k}\nu} - V_{\beta\alpha}) \tilde{O} |\tilde{\psi}_{\vec{k}\nu}\rangle \end{aligned} \quad (3.176)$$

Man erhält also für $\ddot{H}$ *dieselben* Eigenzustände $|\tilde{\psi}_{\vec{k}\nu}\rangle$, aber verschobene Eigenwerte $\ddot{\varepsilon}_{\vec{k}\nu} = \varepsilon_{\vec{k}\nu} - V_{\beta\alpha}$.

Gesamtenergie

Die Ewald-Energie E_{Ewald} enthält einen divergenten Term (vgl. [Ric96]):

$$E_{Ewald} = \ddot{E}_{Ewald} + \frac{1}{2} \frac{4\pi e^2}{\Omega} \frac{\tilde{Z}_v^2}{G^2}$$

Der $G = 0$ -Anteil der übrigen elektrostatischen Energie kompensiert diese Divergenz und liefert einen endlichen Beitrag.

$$\begin{aligned} \tilde{E}_H + \tilde{E}_{loc} &= \Omega \sum_{\vec{G} \neq 0} (\tilde{n}_v + \hat{n} + n_{ref}) (\vec{G}) \left(\frac{1}{2} \overline{V_H} + \overline{v_{at}} \right) (\vec{G}) \\ &+ \Omega \left(\frac{\tilde{Z}_v}{\Omega} + \frac{\beta G^2}{4\pi e^2} \right) \left[\frac{1}{2} \left(4\pi e^2 \frac{\tilde{Z}_v}{\Omega G^2} + \beta \right) - \left(4\pi e^2 \frac{\tilde{Z}_v}{\Omega G^2} + \alpha \right) \right] \\ &= \ddot{E}_H + \ddot{E}_{loc} + \Omega \left(-\frac{1}{2} \frac{4\pi e^2}{\Omega^2} \frac{\tilde{Z}_v^2}{G^2} - \alpha \frac{\tilde{Z}_v}{\Omega} + O(G^2) \right) \end{aligned} \quad (3.177)$$

Also $E_{Ewald} + \tilde{E}_H + \tilde{E}_{loc} = \ddot{E}_{Ewald} + \ddot{E}_H + \ddot{E}_{loc} - \alpha \tilde{Z}_v$. Die Größe α hängt nur von den lokalen Pseudopotentialen ab und läßt sich leicht berechnen (vgl. [Eng95]).

Für die Gültigkeit von Glg. (3.177) ist es *essentiell*, daß $\tilde{E}_{loc}$ (Gleichung 3.82) die Kompensationsladungsdichte $\hat{n}$ enthält. Die Dichte $\tilde{n}$ alleine beinhaltet nicht dieselbe Ladung wie der Atomrumpf (d.h. wie die Dichte $\tilde{n}_{at}$), und daher tritt die erwünschte gegenseitige Annulierung der Divergenzen nicht ein.

Der Formalismus von [Blö94] ist auf neutrale Potentiale angelegt und daher nicht für Atome ohne Referenzladung (z.B. KB-Atome) geeignet. Da ich die Kompatibilität mit diesen Pseudopotentialen aber für wichtig erachte, habe ich das Energiefunktional reformuliert.

Die Energiedifferenz $E^1 - \tilde{E}^1$ erfährt durch die konstante Verschiebung des Potentials keine Veränderung, da diese nur den Randbedingungsterm V_0 verändert, welcher die Energie nicht beeinflusst (vgl. Glg. 3.83).

3.13 Double counting correction

In der Praxis rechnet man die kinetische Energie nicht direkt aus, sondern nutzt die Kohn-Sham-Gleichungen, um sie mittels der Eigenwerte zu ersetzen. Multipliziert man die Kohn-Sham-Gleichungen (3.97) mit $f_{\vec{k}\nu} \tilde{\psi}_{\vec{k}\nu}^*$, integriert und summiert, so erhält man:

$$\begin{aligned}
\sum_{\vec{k}\nu} \tilde{\epsilon}_{\vec{k}\nu} f_{\vec{k}\nu} &= \sum_{\vec{k}\nu} f_{\vec{k}\nu} \langle \tilde{\psi}_{\vec{k}\nu} | \tilde{\epsilon}_{\vec{k}\nu} \tilde{O} | \tilde{\psi}_{\vec{k}\nu} \rangle = \sum_{\vec{k}\nu} f_{\vec{k}\nu} \langle \tilde{\psi}_{\vec{k}\nu} | \tilde{H} | \tilde{\psi}_{\vec{k}\nu} \rangle \\
&= \sum_{\vec{k}\nu} f_{\vec{k}\nu} \left(\langle \tilde{\psi}_{\vec{k}\nu} | -\nabla^2 | \tilde{\psi}_{\vec{k}\nu} \rangle + \langle \tilde{\psi}_{\vec{k}\nu} | \tilde{V}_{lok} | \tilde{\psi}_{\vec{k}\nu} \rangle + \sum_{\mu;\gamma,\gamma'} \langle \tilde{\psi}_{\vec{k}\nu} | \tilde{p}_{\gamma}^{\mu} \tilde{D}_{\gamma,\gamma'}^{\mu} \tilde{p}_{\gamma'}^{\mu} | \tilde{\psi}_{\vec{k}\nu} \rangle \right) \\
&= \tilde{E}_{kin} + \int_{\Omega} d^3r \tilde{n}_v(\vec{r}) \tilde{\ddot{V}}_{lok}(\vec{r}) + \sum_{\mu;\gamma,\gamma'} \rho_{\gamma,\gamma'}^{\mu} \tilde{D}_{\gamma,\gamma'}^{\mu} \quad (3.178)
\end{aligned}$$

Vom letzten Ausdruck kann man gemäß

$$\langle \phi_{\gamma}^{\mu} | -\nabla^2 + V^{1\mu} | \phi_{\gamma'}^{\mu} \rangle = \langle \phi_{\gamma}^{\mu} | -\nabla^2 | \phi_{\gamma'}^{\mu} \rangle + \int_{S_{\mu}} d^3r \phi_{\gamma}^{\mu}(\vec{r}) \phi_{\gamma'}^{\mu}(\vec{r}) V^{1\mu}(\vec{r}) \quad (3.179)$$

einerseits den Anteil der kinetischen Energie abspalten (vgl. Glg. (3.146)), andererseits den restlichen Anteil als Integral über das Potential in den Augmentationsgebieten schreiben. Mit Glg. (3.36) ergibt sich:

$$\sum_{\mu;\gamma,\gamma'} \rho_{\gamma,\gamma'}^{\mu} \tilde{D}_{\gamma,\gamma'}^{\mu} = E_{kin}^1 - \tilde{E}_{kin}^1 + \sum_{\mu} \int_{S_{\mu}} d^3r \left(n^{1\mu} \tilde{V}^{1\mu} - \tilde{n}^{1\mu} \tilde{\ddot{V}}^{1\mu} \right) \quad (3.180)$$

Mit Hilfe dieser Gleichung ersetzt man E_{kin} in der Gesamtenergie. Dabei wird auch die Energie E_{loc} weitgehend annulliert. Es verbleibt davon jedoch ein Beitrag

$$E_{at} = \int_{\Omega} (\hat{n} - n_{ref}) \tilde{\ddot{V}}_{at} d^3r + \sum_{\mu} \int_{S_{\mu}} (-n_{ref}^{\mu}) V_{at}^{\mu} - (\hat{n}^{1\mu} - n_{ref}^{\mu}) \tilde{V}_{at}^{\mu} d^3r_{\mu}, \quad (3.181)$$

weil in Glg. (3.178) im Gegensatz zu Glg. (3.81) die Dichten $\hat{n}$ und $\hat{n}^1$ nicht auftreten. Im Falle der Hartree-Energie tritt eine teilweise Aufhebung ein, sofern die selbstkonsistente Dichte betrachtet wird. Da aber die Kohn-Sham-Gleichungen iterativ gelöst werden (siehe Kapitel 5), muß man für den Fall, daß die Selbstkonsistenz noch nicht erreicht ist, unterscheiden zwischen der Dichte n_{alt} , aus welcher der Hamiltonoperator berechnet wurde, und der neuen n_{neu} , welche sich aus den Lösungen der Eigenwertgleichung zu diesem Hamiltonoperator ergibt. Während die aktuelle Hartree-Energie E_H und die Austausch-Korrelations-Energie E_{xc} mit der aktuellen Dichte n_{neu} zu bilden sind, enthält Glg. (3.178) nur Potentiale, welche aus den Dichten der vorherigen Iteration generiert sind. Dies wird an den entsprechenden Größen durch den Index *alt* kenntlich gemacht.

Die Gesamtenergie läßt sich demnach schreiben als:

$$E = E_0 + \ddot{E}_{Ewald} + \sum_{\vec{k}\nu} \ddot{\epsilon}_{\vec{k}\nu} f_{\vec{k}\nu} - \alpha \tilde{Z}_v - \ddot{E}_{alt} \quad (3.182)$$

$$+ \left(\ddot{E}_H + E_H^1 - \tilde{E}_H^1 \right) + \left(\tilde{E}_{xc} + E_{xc}^1 - \tilde{E}_{xc}^1 \right) + E_{at} \quad (3.183)$$

wobei

$$\ddot{E}_{alt} = \int_{\Omega} \tilde{n}_v \left(\tilde{V}_{xc,alt} + \ddot{\tilde{V}}_{Halt} \right) \quad (3.184)$$

$$+ \int_{S_\mu} n^{1\mu} \left(V_{H,alt}^{1\mu} + V_{xc,alt}^\mu + \ddot{V}_{0,alt}^\mu \right) - \tilde{n}^{1\mu} \left(\tilde{V}_{H,alt}^{1\mu} + \tilde{V}_{xc,alt}^\mu + \ddot{V}_{0,alt}^\mu \right) \quad (3.185)$$

im Falle der iterativen Lösung die Potentiale enthält, welche aus der *alten* Dichte gebildet werden.

3.14 Kräfte

Die Kräfte auf die Atome sind gegeben durch die totale Ableitung des Energiefunktionals $E[n, \{\vec{\tau}_\mu\}]$ nach den Atomkoordinaten $\vec{\tau}_\mu$. Bei gewöhnlichen Pseudopotentialen benutzt man nun das Hellmann-Feynman-Theorem, welches besagt, daß im elektronischen Grundzustand der Beitrag durch die implizite Abhängigkeit der Dichte von den Atompositionen null ist und die Kräfte durch die *partielle* Ableitung der Energie gegeben sind.

In der PAW-Methode ist die Sache etwas komplizierter, da die Augmentationsgebiete und mithin die Kompensationsdichten $\hat{n}$ sich mit den Atomrümpfen bewegen.

Diese Kräfte werden manchmal Pulay-Kräfte genannt, obwohl dieser Begriff sich eigentlich auf Basisfunktionen bezieht, welche von den Atomkoordinaten abhängen, was hier nicht der Fall ist. Ein weiterer Unterschied zum Falle gewöhnlicher Pseudopotentiale ergibt sich aus der Veränderung der Orthogonalitätsbeziehung durch das Vorhandensein einer Überlappmatrix, welche von den Atomkoordinaten abhängt. Eine adäquate Behandlung der Kräfte erfordert, daß die Kraft definiert wird als partielle Ableitung der Energie nach den Atompositionen bei konstantgehaltenen Variationsparametern $\tilde{\psi}_{\vec{k}\nu}$.

$$\vec{F}^\mu = \left. \frac{\partial E}{\partial \vec{\tau}_\mu} \right|_{\{\tilde{\psi}_{\vec{k}\nu}\}} \quad (3.186)$$

Dabei sind auch die Projektorfunktionen und mithin die Multipolmomente Q_L^μ und die Ladungsdichtematrix $\rho_{\gamma,\gamma'}^\mu$ als explizit abhängige Größen zu betrachten. Daher hängt $\hat{n}$ einerseits explizit und andererseits über die Multipolmomente implizit von den Atomorten ab. Im glatten Anteil $\tilde{E}$ der Energie hängen abgesehen von den Multipolmomenten auch das lokale Pseudopotential sowie die Kompensationsdichte $\hat{n}$ und die partielle core-Dichte $\tilde{n}_c$ von der Position der Atome ab. Deswegen ergeben sich bei festgehaltenen Multipolmomenten daraus drei Kräfte:

$$\left. \frac{\partial \tilde{E}_H}{\partial \vec{\tau}_\mu} \right|_{Q_L^\mu} = \int_{\Omega} \frac{\delta \tilde{E}_H}{\delta \hat{n}} \left. \frac{\partial \hat{n}}{\partial \vec{\tau}_\mu} \right|_{Q_L^\mu} d^3r = \int_{\Omega} \overline{\tilde{V}_H} \left. \frac{\partial \hat{n}}{\partial \vec{\tau}_\mu} \right|_{Q_L^\mu} d^3r \quad (3.187)$$

$$\left. \frac{\partial \tilde{E}_{at}}{\partial \vec{\tau}_\mu} \right|_{Q_L^\mu} = \int_{\Omega} (\tilde{n} + \hat{n}) \frac{\partial \overline{\tilde{V}_{at}}}{\partial \vec{\tau}_\mu} d^3r + \int_{\Omega} \left. \frac{\partial \hat{n}}{\partial \vec{\tau}_\mu} \right|_{Q_L^\mu} \overline{\tilde{V}_{at}} d^3r \quad (3.188)$$

also

$$\vec{F}_{at}^\mu = - \int_{\Omega} (\tilde{n} + \hat{n}) \frac{\partial \overline{\tilde{V}_{at}}}{\partial \vec{\tau}_\mu} d^3r \quad (3.189)$$

$$\vec{F}_{psd}^\mu = - \int_{\Omega} \left(\overline{\tilde{V}_H} + \overline{\tilde{V}_{at}} \right) \left. \frac{\partial \hat{n}}{\partial \vec{\tau}_\mu} \right|_{Q_L^\mu} d^3r \quad (3.190)$$

und

$$\vec{F}_{xc}^\mu = - \frac{d \tilde{E}_{xc}}{d \vec{\tau}_\mu} = - \int_{\Omega} \tilde{V}_{xc} \frac{\partial \tilde{n}_c}{\partial \vec{\tau}_\mu} d^3r \quad (3.191)$$

In den Energien E^1 und $\tilde{E}^1$ werden bei einer Verschiebung des Atomes alle Dichten und alle Potentiale starr verschoben. Daher gibt es keinen Beitrag durch eine explizite Abhängigkeit von den Atomkoordinaten, sondern nur durch die implizite über die

Formänderung der Einzentrendichten, welche sich durch die Verschiebung der Projektorfunktionen ergibt. Diese werden mittels der Dichtematrix $\rho_{\gamma,\gamma'}^\mu$ beschrieben. Wie bei der Ableitung des Hamiltonoperators wird die Energie als Funktional von $\rho_{\gamma,\gamma'}^\mu$, $\tilde{n}$, n^1 , $\tilde{n}^1$ und Q_L^μ betrachtet, wobei die Multipolmomente wiederum Funktionale der Einzentrendichten sind und diese wiederum von der Dichtematrix abhängen.

$$\vec{F}_2^\mu = \sum_{\gamma,\gamma'} \frac{dE}{d\rho_{\gamma,\gamma'}^\mu} \frac{d\rho_{\gamma,\gamma'}^\mu}{d\vec{\tau}_\mu} \quad (3.192)$$

$$\begin{aligned} \frac{dE}{d\rho_{\gamma,\gamma'}^\mu} &= \langle \phi_\gamma^\mu | -\nabla^2 | \phi_{\gamma'}^\mu \rangle - \langle \tilde{\phi}_\gamma^\mu | -\nabla^2 | \tilde{\phi}_{\gamma'}^\mu \rangle \\ &+ \int_{S_\mu} \left(\frac{\partial E}{\partial n^{1\mu}} + \sum_L \frac{\partial E}{\partial Q_L^\mu} \frac{\partial Q_L^\mu}{\partial n^{1\mu}} \right) \frac{\partial n^{1\mu}}{\partial \rho_{\gamma,\gamma'}^\mu} d^3r \\ &+ \int_{S_\mu} \left(\frac{\partial E}{\partial \tilde{n}^{1\mu}} + \sum_L \frac{\partial E}{\partial Q_L^\mu} \frac{\partial Q_L^\mu}{\partial \tilde{n}^{1\mu}} \right) \frac{\partial \tilde{n}^{1\mu}}{\partial \rho_{\gamma,\gamma'}^\mu} d^3r \end{aligned} \quad (3.193)$$

also mit

$$\frac{dn^{1\mu}}{d\rho_{\gamma,\gamma'}^\mu} = \langle \phi_\gamma | \vec{r} \rangle \langle \vec{r} | \phi_{\gamma'} \rangle \quad (3.194)$$

$$\frac{d\tilde{n}^{1\mu}}{d\rho_{\gamma,\gamma'}^\mu} = \langle \tilde{\phi}_\gamma | \vec{r} \rangle \langle \vec{r} | \tilde{\phi}_{\gamma'} \rangle \quad (3.195)$$

$$\begin{aligned} \frac{dE}{d\rho_{\gamma,\gamma'}^\mu} &= \langle \phi_\gamma^\mu | -\nabla^2 + V^{1\mu} + V_0^\mu | \phi_{\gamma'}^\mu \rangle - \langle \tilde{\phi}_\gamma^\mu | -\nabla^2 + \tilde{V}^{1\mu} + V_0^\mu | \tilde{\phi}_{\gamma'}^\mu \rangle \\ &= D_{\gamma,\gamma'}^\mu \end{aligned} \quad (3.196)$$

Dieser Term enthält auch die oben ausgesparte Abhängigkeit des glatten Anteils $\tilde{E}$ der Energie von den Multipolmomenten Q_L^μ .

Ein zusätzlicher Term ergibt sich aus der Orthogonalitätsbeziehung. Die Orthogonalität der echten Wellenfunktionen (bzgl. des gewöhnlichen Skalarproduktes) wird gewährleistet durch diejenige der Pseudowellenfunktionen bezüglich des Überlappoperators (vgl. Abschnitte 3.7 und 3.8). Dieser aber hängt über die Projektoren von der Lage der Atome ab. Aus den Euler-Lagrange-Gleichungen für die Atomkoordinaten ergibt sich unter Berücksichtigung der Nebenbedingungen (vgl. [LPC⁺93])

$$\vec{F}_{ov}^\mu = \sum_{\vec{k}\nu} f_{\vec{k}\nu} \varepsilon_{\vec{k}\nu} \langle \tilde{\psi}_{\vec{k}\nu} | \frac{d\tilde{O}}{d\vec{\tau}_\mu} | \tilde{\psi}_{\vec{k}\nu} \rangle \quad (3.197)$$

$$= \sum_{\vec{k}\nu} f_{\vec{k}\nu} \varepsilon_{\vec{k}\nu} \sum_{\gamma,\gamma'} q_{\gamma,\gamma'}^{\alpha(\mu)} \frac{d}{d\vec{\tau}_\mu} \left(\langle \tilde{\psi}_{\vec{k}\nu} | \tilde{p}_\gamma^\mu \rangle \langle \tilde{p}_{\gamma'}^\mu | \tilde{\psi}_{\vec{k}\nu} \rangle \right) \quad (3.198)$$

Faßt man diesen mit dem vorherigen zusammen, so erhält man

$$\vec{F}_{NL}^\mu = - \sum_{\vec{k}\nu} f_{\vec{k}\nu} \sum_{\gamma, \gamma'} \left(D_{\gamma, \gamma'}^\mu - \varepsilon_{\vec{k}\nu} q_{\gamma, \gamma'}^{\alpha(\mu)} \right) \frac{d}{d\vec{\tau}_\mu} \left(\langle \tilde{\psi}_{\vec{k}\nu} | \tilde{p}_\gamma^\mu \rangle \langle \tilde{p}_{\gamma'}^\mu | \tilde{\psi}_{\vec{k}\nu} \rangle \right) \quad (3.199)$$

Schließlich hat man noch einen Beitrag zur Kraft $\vec{F}_{Ewald}^\mu$ von der Ewald-Energie (vgl. [Eng95]).

Insgesamt ergibt sich die Kraft auf das Atom μ als:

$$\vec{F}^\mu = \vec{F}_{at}^\mu + \vec{F}_{psd}^\mu + \vec{F}_{xc}^\mu + \vec{F}_{NL}^\mu + \vec{F}_{Ewald}^\mu \quad (3.200)$$

Details und Formeln Die verschiedenen Beiträge zur Kraft müssen auf dieselbe Weise berechnet werden wie die Energie, aus der sie hervorgehen. Entsprechend werden die Kräfte $\vec{F}_{at}^\mu$, $\vec{F}_{psd}^\mu$ und $\vec{F}_{xc}^\mu$, welche sich aus dem glatten Anteil $\tilde{E}$ der Energie ergeben, im reziproken Raum berechnet:

$$\vec{F}_{at}^\mu = -\Omega \sum_{\vec{G}} (i \vec{G}) S^\mu(\vec{G}) (\tilde{n} + \hat{n})^* (\vec{G}) \overline{\tilde{V}_{at}^\mu}(\vec{G}) \quad (3.201)$$

$$\vec{F}_{psd}^\mu = -\Omega \sum_{\vec{G}} (i \vec{G}) S^\mu(\vec{G}) \left(\overline{\tilde{V}_H} + \overline{\tilde{V}_{at}} \right)^* (\vec{G}) \hat{n}_l^\mu(\vec{G}) \quad (3.202)$$

$$\vec{F}_{xc}^\mu = -\Omega \sum_{\vec{G}} (i \vec{G}) S^\mu(\vec{G}) \tilde{V}_{xc}^*(\vec{G}) \tilde{n}_c^\mu(\vec{G}) \quad (3.203)$$

Bei der Berechnung der nicht-lokalen Kraft benötigt man die Ableitung der Projektionskoeffizienten. Diese berechnet sich gemäß

$$\begin{aligned} \frac{d}{d\vec{\tau}_\mu} \left(\langle \tilde{\psi}_{\vec{k}\nu} | \tilde{p}_\gamma^\mu \rangle \langle \tilde{p}_{\gamma'}^\mu | \tilde{\psi}_{\vec{k}\nu} \rangle \right) &= \left(\langle \tilde{\psi}_{\vec{k}\nu} | \tilde{p}_\gamma^\mu \rangle e^{-i\vec{k}\vec{\tau}_\mu} \right) \frac{d}{d\vec{\tau}_\mu} \left(e^{i\vec{k}\vec{\tau}_\mu} \langle \tilde{p}_{\gamma'}^\mu | \tilde{\psi}_{\vec{k}\nu} \rangle \right) \\ &+ \frac{d}{d\vec{\tau}_\mu} \left(\langle \tilde{\psi}_{\vec{k}\nu} | \tilde{p}_\gamma^\mu \rangle e^{-i\vec{k}\vec{\tau}_\mu} \right) \left(e^{i\vec{k}\vec{\tau}_\mu} \langle \tilde{p}_{\gamma'}^\mu | \tilde{\psi}_{\vec{k}\nu} \rangle \right) \end{aligned} \quad (3.204)$$

$$\begin{aligned} \frac{d}{d\vec{\tau}_\mu} \left(e^{i\vec{k}\vec{\tau}_\mu} \langle \tilde{p}_{\gamma'}^\mu | \tilde{\psi}_{\vec{k}\nu} \rangle \right) &= 4\pi \sum_{\vec{G}} i^{-l'} (-i \vec{G}) S^{\mu*}(\vec{G}) Y_{L'}(\vec{k} + \vec{G}) \tilde{p}_{\gamma'}^\mu \left(|\vec{k} + \vec{G}| \right) c_{\vec{k}\nu}(\vec{G}) \\ &= e^{i\vec{k}\vec{\tau}_\mu} \sum_{\vec{G}} \langle \tilde{p}_{\gamma'}^\mu | \vec{k} + \vec{G} \rangle (-i \vec{G}) \langle \vec{k} + \vec{G} | \tilde{\psi}_{\vec{k}\nu} \rangle \end{aligned} \quad (3.205)$$

3.15 Symmetrie

Das Vorhandensein von Symmetrien des Kristallgitters läßt sich in dreifacher Hinsicht nutzen. Zum einen induziert die Symmetrie im Ortsraum eine ebensolche im reziproken Raum. Daher stimmen die Eigenwerte zu einem bestimmten k-Punkt

überein mit denen zu k-Punkten, die sich aus diesem durch eine Symmetrietransformation ergeben; auch zwischen den Eigenvektoren besteht eine bestimmte Beziehung. Daher braucht man das Eigenwertproblem nur für einen dieser k-Punkte zu lösen. Die für die Berechnung der Dichte und der Kräfte benötigten Eigenfunktionen der anderen k-Punkte berechnen sich daraus auf einfache Weise. Dies bedeutet eine erhebliche Einsparung an Rechenzeit.

Zum anderen lassen sich bei der Ortsraumdarstellung der Dichte¹⁶ um die Atome herum durch Symmetrien auf zwei Weisen Einsparungen erzielen: Erstens muß man diese lokalisierten Quantitäten nur für ein repräsentatives Atom pro Gruppe von symmetrieverwandten Atomen berechnen und speichern. Die Größen für die anderen Atome erhält man aus denen des Repräsentativatoms durch (evtl. unechte) Drehungen. Zweitens kann eine *globale* Symmetrie eine *lokale* (d.h. bzgl. des Ortes eines Atomes) induzieren. Dies ist immer dann der Fall, wenn es mehrere verschiedene *globale* Symmetrieoperationen gibt, die ein bestimmtes Atom auf ein (bestimmtes) anderes abbilden; die Verkettung ($S_3 = S_2^{-1} \circ S_1$) dieser Operationen ist eine lokale Symmetrie dieses Atomes. Eine lokale Symmetrie hat zur Folge, daß nicht mehr alle L -Komponenten einer Entwicklung der Dichte um dieses Atom herum unabhängig sein können. Gibt es beispielsweise eine Inversionssymmetrie bzgl. eines Atomes, so verschwinden alle Komponenten mit ungeradem l . Man kann also die Anzahl der Funktionen verringern durch das Benutzen geeigneter Linearkombinationen von Kugelflächenfunktionen. Diese sind als *lattice harmonics* bekannt.

Details und Formeln Im folgenden bezeichnet S eine Symmetrieoperation

$$S : \vec{r} \mapsto S\vec{r} = \mathbf{S}\vec{r} + \vec{t}_S \quad (3.206)$$

und $\mathcal{S}$ die Gruppe aller Symmetrieoperationen der vorliegenden Kristallstruktur. Die Anwendung einer Symmetrie S^{-1} auf eine Wellenfunktion $\tilde{\psi}_{\vec{k}\nu}$ ist definiert via

$$\left(S^{-1}\tilde{\psi}_{\vec{k}\nu}\right)(\vec{r}) = \tilde{\psi}_{\vec{k}\nu}(S\vec{r}) \quad (3.207)$$

Die transformierte Funktion ist, wie man sich leicht überzeugt, eine Blochfunktion zum Blochvektor $\vec{k}' = \mathbf{S}^{-1}\vec{k}$:

$$\left(S^{-1}\tilde{\psi}_{\vec{k}\nu}\right)(\vec{r} + \vec{R}) = e^{-i(\mathbf{S}^{-1}\vec{k})\vec{R}} \left(S^{-1}\tilde{\psi}_{\vec{k}\nu}\right)(\vec{r}), \quad (3.208)$$

¹⁶Was hier und im folgenden über die Dichte gesagt wird, gilt auch für die Potentiale, Dichtematrizen $\rho_{\gamma,\gamma'}^\mu$ und Potentialmatrizen $D_{\gamma,\gamma'}^\mu$.

also

$$\left(S^{-1} \tilde{\psi}_{\vec{k}\nu} \right) (\vec{r}) = \tilde{\psi}_{(\mathbf{S}^{-1}\vec{k}),\nu}(\vec{r}) \quad (3.209)$$

Die Fourierkoeffizienten erhält man mittels

$$\begin{aligned} \sqrt{\Omega} \left(S^{-1} \tilde{\psi}_{\vec{k}\nu} \right) (\vec{r}) &= \sum_{\vec{G}} c_{\vec{k}\nu}^{\vec{G}} e^{-i(\vec{k}+\vec{G})(\mathbf{S}\vec{r}+\vec{t}_S)} \\ &= \sum_{\vec{G}} c_{\vec{k}\nu}^{\vec{G}} e^{-i(\mathbf{S}^{-1}\vec{k}+\mathbf{S}^{-1}\vec{G})(\vec{r}+\mathbf{S}^{-1}\vec{t}_S)} \\ &= \sum_{\vec{G}'} c_{\vec{k}\nu}^{\mathbf{S}\vec{G}'} e^{-i(\mathbf{S}^{-1}\vec{k}+\vec{G}')(\mathbf{S}^{-1}\vec{t}_S)} e^{-i(\mathbf{S}^{-1}\vec{k}+\vec{G}')\vec{r}} \end{aligned} \quad (3.210)$$

also

$$c_{\mathbf{S}^{-1}\vec{k},\nu}^{\vec{G}} = c_{\vec{k}\nu}^{\mathbf{S}\vec{G}} e^{-i(\vec{k}+\mathbf{S}\vec{G})\vec{t}_S} \quad (3.211)$$

Gedrehte Kugelflächenfunktionen lassen sich darstellen als (vgl. [Mes61], Anhang A)

$$Y_L(\mathbf{S}^{-1}\hat{r}) = \sum_{m'} Y_{lm'}(\hat{r}) R_{m'm}^{(l)}(S) \quad (3.212)$$

Sei g_L^μ eine am Atom μ lokalisierte Funktion der Form

$$g_L^\mu(\vec{r}) = g_l^{\alpha(\mu)}(|\vec{r} - \vec{\tau}_\mu|) Y_L(\vec{r} + \vec{\tau}_\mu) \quad (3.213)$$

Weil S^{-1} ein Symmetrieelement ist, befindet sich am Ort $S^{-1}\vec{\tau}_\mu$ ein identisches Atom.

$$\vec{\tau}_{S^{-1}\mu} = S^{-1}\vec{\tau}_\mu = \mathbf{S}^{-1}\vec{\tau}_\mu + \vec{t}_{S^{-1}} \quad (3.214)$$

und daher

$$S^{-1}\vec{r} - \vec{\tau}_\mu = \mathbf{S}^{-1}(\vec{r} - \vec{\tau}_{S\mu}) \quad (3.215)$$

Somit erhält man

$$\begin{aligned} \langle \vec{r} | S g_L^\mu \rangle &= \langle S^{-1}\vec{r} | g_L^\mu \rangle = g_L^\mu(S^{-1}\vec{r}) \\ &= g_l^{\alpha(\mu)}(|\mathbf{S}^{-1}(\vec{r} - \vec{\tau}_{S\mu})|) Y_L(\mathbf{S}^{-1}(\vec{r} - \vec{\tau}_{S\mu})) \\ &= \sum_{\tilde{m}} R_{\tilde{m}m}^{(l)}(S) g_L^{\alpha(\mu)}(|\vec{r} - \vec{\tau}_{S\mu}|) Y_{l\tilde{m}}(\vec{r} - \vec{\tau}_{S\mu}) \end{aligned}$$

und schließlich

$$|S \tilde{p}_\gamma^\mu\rangle = \sum_{\tilde{m}} R_{\tilde{m}m}^{(l)}(S) |\tilde{p}_{l\tilde{m}i}^{S\mu}\rangle \quad (3.216)$$

oder

$$|\tilde{p}_\gamma^{S\mu}\rangle = \sum_{\tilde{m}} R_{m\tilde{m}}^{(l)}(S) |S \tilde{p}_{l\tilde{m}i}^\mu\rangle \quad (3.217)$$

Entsprechende Formeln gelten für ϕ und $\tilde{\phi}$. Diese Funktionen werden also unter einer Symmetrie auf ein äquivalentes Atom abgebildet, und das lokale Koordinatensystem wird rotiert. Für die Projektionskoeffizienten ergibt sich damit

$$\langle \tilde{p}_\gamma^\mu | \tilde{\psi}_{\mathbf{S}^{-1}\vec{k},\nu} \rangle = \langle S \tilde{p}_\gamma^\mu | \tilde{\psi}_{\vec{k},\nu} \rangle = \sum_{m'} R_{m'm}^{(l)}(S) \langle \tilde{p}_{lm'i}^{S\vec{\tau}_\mu} | \tilde{\psi}_{\vec{k},\nu} \rangle. \quad (3.218)$$

Wenn man also für einen k-Punkt die Projektionskoeffizienten für alle Atome μ kennt, so kann man unter Benutzung der Symmetrie jene für symmetrieverwandte k-Punkte bestimmen. Statt auf die Wellenfunktion wird die (inverse) Symmetrioperation auf die Projektorfunktionen angewandt.

Dichtematrix Damit kann man die Dichtematrix $\rho_{\gamma,\gamma'}^\mu$ unter Benutzung von k-Punkten nur aus dem irreduziblen Teil der Brillouin-Zone (iBZ) gewinnen mittels

$$\begin{aligned} \rho_{\gamma,\gamma'}^\mu &= \sum_{\vec{k}\nu} f_{\vec{k}\nu} \langle \tilde{\psi}_{\vec{k}\nu} | \tilde{p}_\gamma^\mu \rangle \langle \tilde{p}_{\gamma'}^\mu | \tilde{\psi}_{\vec{k}\nu} \rangle \\ &= \sum_{\substack{\vec{k} \in \text{iBZ} \\ \nu}} f_{\vec{k}\nu} \frac{\#\vec{S}\vec{k}}{\#\mathcal{S}} \sum_{S \in \mathcal{S}} \langle \tilde{\psi}_{\mathbf{S}^{-1}\vec{k},\nu} | \tilde{p}_\gamma^\mu \rangle \langle \tilde{p}_{\gamma'}^\mu | \tilde{\psi}_{\mathbf{S}^{-1}\vec{k},\nu} \rangle \\ &= \sum_{S \in \mathcal{S}} \sum_{\tilde{m}, \tilde{m}'} R_{\tilde{m}m}^{(l)}(S) R_{\tilde{m}'m'}^{(l')}(S) \rho_{l'\tilde{m}'i'}^{S\mu, \text{iBZ}} \end{aligned} \quad (3.219)$$

wobei

$$\rho_{l'\tilde{m}'i'}^{S\mu, \text{iBZ}} = \sum_{\substack{\vec{k} \in \text{iBZ} \\ \nu}} f_{\vec{k}\nu} \frac{\#\vec{S}\vec{k}}{\#\mathcal{S}} \langle \tilde{\psi}_{\vec{k}\nu} | \tilde{p}_{l\tilde{m}i}^{S\vec{\tau}_\mu} \rangle \langle \tilde{p}_{l'\tilde{m}'i'}^{S\vec{\tau}_\mu} | \tilde{\psi}_{\vec{k}\nu} \rangle \quad (3.220)$$

und $\#\mathcal{S}$ die Anzahl der Elemente von $\mathcal{S}$ bezeichnet und $\#\vec{S}\vec{k}$ die Mächtigkeit der Menge, welche durch Anwendung aller Gruppenelemente von $\mathcal{S}$ auf $\vec{k}$ entsteht. Benutzt wurde dabei die Tatsache, daß sich für symmetrieäquivalente k-Punkte dieselben Eigenwerte und folglich dieselben Besetzungszahlen ergeben, und daß die Gewichtung $\omega_{\vec{k}}$ äquivalenter k-Punkte natürlich ebenfalls identisch ist.

Nicht-lokale Kraft Ähnlich wie bei der Dichtematrix kann man auch bei der nicht-lokalen Kraft vorgehen. Wie man leicht aus Glg. 3.145 ersieht, gilt

$$\frac{d}{d\vec{\tau}_\mu} \left(e^{i\vec{k}\vec{\tau}S\mu} \langle \tilde{p}_\gamma^{S\mu} | \tilde{\psi}_{\vec{k}\nu} \rangle \right) = \mathbf{S}^{-1} \frac{d}{d\vec{\tau}S\mu} \left(e^{i\vec{k}\vec{\tau}S\mu} \langle \tilde{p}_\gamma^{S\mu} | \tilde{\psi}_{\vec{k}\nu} \rangle \right) \quad (3.221)$$

Zur Abkürzung definiert man

$$M_{\gamma,\gamma'}^\mu = D_{\gamma,\gamma'}^\mu - \varepsilon_{\vec{k}\nu} q_{\gamma,\gamma'}^{\alpha(\mu)} \quad (3.222)$$

und vermerkt, daß

$$M_{\gamma,\gamma'}^\mu(\vec{k}\nu) = M_{\gamma,\gamma'}^\mu(\mathbf{S}^{-1}\vec{k}\nu) \quad (3.223)$$

Dann ergibt sich die nichtlokale Kraft als

$$\begin{aligned} \vec{F}_{NL}^\mu &= - \sum_{\vec{k}\nu} f_{\vec{k}\nu} \sum_{\gamma,\gamma'} M_{\gamma,\gamma'}^\mu \frac{d}{d\vec{\tau}_\mu} \left(\langle \tilde{\psi}_{\vec{k}\nu} | \tilde{p}_\gamma^\mu \rangle \langle \tilde{p}_{\gamma'}^\mu | \tilde{\psi}_{\vec{k}\nu} \rangle \right) \\ &= - \sum_{\substack{\vec{k} \in iBZ \\ \nu}} f_{\vec{k}\nu} \frac{\#\mathcal{S}\vec{k}}{\#\mathcal{S}} \sum_{\substack{\gamma,\gamma' \\ S \in \mathcal{S}}} M_{\gamma,\gamma'}^\mu \frac{d}{d\vec{\tau}_\mu} \left(\langle \tilde{\psi}_{\mathbf{S}^{-1}\vec{k},\nu} | \tilde{p}_\gamma^\mu \rangle \langle \tilde{p}_{\gamma'}^\mu | \tilde{\psi}_{\mathbf{S}^{-1}\vec{k},\nu} \rangle \right) \end{aligned} \quad (3.224)$$

$$\begin{aligned} &= - \sum_{\substack{\vec{k} \in iBZ \\ \nu}} f_{\vec{k}\nu} \frac{\#\mathcal{S}\vec{k}}{\#\mathcal{S}} \sum_{S \in \mathcal{S}} \sum_{\substack{l \tilde{m} i \\ l' \tilde{m}' i'}} \underbrace{\sum_{m,m'} M_{\gamma,\gamma'}^\mu R_{\tilde{m}m}^{(l)}(S) R_{\tilde{m}'m'}^{(l')}(S)}_{M_{l \tilde{m} i, l' \tilde{m}' i'}^{S\mu}} \times \\ &\times \frac{d}{d\vec{\tau}_\mu} \left(\langle \tilde{\psi}_{\vec{k}\nu} | \tilde{p}_{l\tilde{m}i}^{S\vec{\tau}_\mu} \rangle \langle \tilde{p}_{l'\tilde{m}'i'}^{S\vec{\tau}_\mu} | \tilde{\psi}_{\vec{k}\nu} \rangle \right) \end{aligned} \quad (3.225)$$

$$= \frac{1}{\#\mathcal{S}} \sum_{S \in \mathcal{S}} \mathbf{S}^{-1} \vec{F}_{NL,iBZ}^{S\mu} \quad (3.226)$$

Symmetrieverwandte Atome Werden zwei Atome durch eine Symmetrie aufeinander abgebildet, so muß man nur für eines dieser Atome Dichte- und Potentialmatrix und nicht-lokale Kraft berechnen. Für die übrigen erhält man mittels Glg. (3.218)

$$\rho_{\gamma,\gamma'}^{S\mu} = \sum_{\tilde{m},\tilde{m}'} R_{m\tilde{m}}^{(l)}(S) R_{m'\tilde{m}'}^{(l')}(S) \rho_{l \tilde{m} i, l' \tilde{m}' i'}^\mu \quad (3.227)$$

Ebenso erhält man für $D_{\gamma,\gamma'}^\mu$ wegen

$$\langle \phi_\gamma^{S\mu} | V^1 | \phi_{\gamma'}^{S\mu} \rangle = \sum_{\tilde{m},\tilde{m}'} R_{m\tilde{m}}^{(l)}(S) R_{m'\tilde{m}'}^{(l')}(S) \langle \phi_{l\tilde{m}i}^\mu | \underbrace{S^{-1}V^1S}_{V^1} | \phi_{l'\tilde{m}'i'}^\mu \rangle \quad (3.228)$$

und weil die entsprechende Gleichung für $D_{kin;\gamma,\gamma'}^\mu$ trivialerweise erfüllt ist:

$$D_{\gamma,\gamma'}^{S\mu} = \sum_{\tilde{m},\tilde{m}'} R_{m\tilde{m}}^{(l)}(S) R_{m'\tilde{m}'}^{(l')}(S) D_{l \tilde{m} i, l' \tilde{m}' i'}^\mu \quad (3.229)$$

Für die Kraft erhält man entsprechend

$$\vec{F}_{NL}^\mu = \mathbf{S}^{-1} \vec{F}_{S\mu}^{NL} \quad (3.230)$$

Integrale wie $\int V(\vec{r}) n(\vec{r}) d^3r$ sind für äquivalente Atome identisch, da beide Integranden unter einer Symmetrie auf dieselbe Weise gedreht werden.

3.16 Symbolübersicht

Tabelle 3.1: Zusammenfassung wichtiger Symbole

Symbol	Bedeutung
γ	Multiindex (L, i) : zählt <i>partial waves</i> bzw. Projektoren (für <i>ein</i> Atom)
L	Multiindex (l, m) : Drehimpulskomponenten
i	zählt <i>partial waves</i> bzw. Projektoren gleicher Drehimpulskomponente, aber unterschiedlicher Energie (für <i>ein</i> Atom)
μ	individuelles Atom
$\alpha, \alpha(\mu)$	Atomart (zu der das Atom μ gehört)
μ_α	zählt Atome der Sorte α
$\vec{\tau}_\mu$	Ort, an dem das Atom μ sich befindet
$S^\mu(\vec{G})$	Strukturfaktor $e^{i\vec{G}\vec{\tau}_\mu}$
Ω	Gebiet bzw. Volumen der Superzelle
S_μ	Augmentationsgebiet des Atomes μ („Kugel“)
S	Vereinigung aller Augmentationsgebiete
I	interstitieller Raum (Superzelle ohne alle Augmentationsgebiete)
R_μ	Radius des Augmentationsgebietes des Atomes μ
$\vec{R}$	Translationsvektor des Bravaisgitters
ϕ^μ	AE- <i>partial wave</i> = <i>echte</i> atomare Wellenfunktion; definiert auf radialem Ortsraumgitter im Augmentationsgebiet des Atomes μ ;
$\tilde{\phi}^\mu$	PS- <i>partial wave</i> = atomare Pseudowellenfunktion; definiert auf radialem Ortsraumgitter im Augmentationsgebiet des Atomes μ ;
$\tilde{p}^\mu$	Projektorfunktion; lokalisiert im Augmentationsgebiet des Atomes μ ; definiert im reziproken Raum
$\tilde{\psi}_{\vec{k}\nu}$	Pseudowellenfunktion; definiert im reziproken Raum bzw. auf dem regulären Ortsraumgitter
$\vec{k}$	$\vec{k}$ -Punkt
ν	zählt Zustände (für einen $\vec{k}$ -Punkt)
Y_L	(reelle) Kugelflächenfunktion
j_l	sphärische Besselfunktion

Tabelle 3.1: Zusammenfassung wichtiger Symbole

Symbol	Bedeutung
$Z^\mu = Z^{\alpha(\mu)}$	Kernladung des Atomes μ der Sorte $\alpha(\mu)$
$Z_v^\mu = Z_v^{\alpha(\mu)}$	Valenzladung (= Rumpfladung) des Atomes μ
$\tilde{Z}_v^\mu = \tilde{Z}_v^{\alpha(\mu)}$	=0 bei Benutzung von Referenzladungen, sonst $= Z_v^\mu$
n_z^α	$= Z^\alpha \delta(\vec{r})$ Ladungsdichte des Atomkernes vom Typ α
n_z^μ	$= Z^\mu \delta(\vec{r} - \vec{\tau}_\mu)$ Ladungsdichte des Atomkernes μ am Ort $\vec{\tau}_\mu$
n_z	$= \sum_\mu n_z^\mu$ Ladungsdichte aller Atomkerne in einer Superzelle
$\overline{n_z}$	$= \sum_{\vec{R}} n_z(\vec{r} - \vec{R})$ Ladungsdichte aller Atomkerne in allen Superzellen
$V_H[\cdot](\vec{r})$	elektrostatisches Potential, als Funktional einer Dichte betrachtet
$V_H[n_z](\vec{r})$	elektrostatisches Potential der Atomkerne in <i>einer</i> Superzelle
$\overline{V_H}[n_z](\vec{r})$	$= V_H[\overline{n_z}](\vec{r}) = \sum_{\vec{R}} V_H[n_z](\vec{r} - \vec{R})$ elektrostatisches Potential der Atomkerne in <i>allen</i> Superzellen; dies ist eine Gittersumme der Einzelpotentiale
$\tilde{n}$	glatte Dichte (Betragsquadrat der Pseudowellenfunktionen)
$n^{1\mu}, \tilde{n}^{1\mu}$	echte bzw. glatte Dichte am Atom μ ; diese sind außerhalb von S_μ untereinander identisch; innerhalb von S_μ gilt: $n = n^{1\mu}$ und $\tilde{n} = \tilde{n}^{1\mu}$.
$\hat{n}, \hat{n}^{1\mu}$	Kompensationsladung; $\hat{n}$ ist definiert im gesamten Raum, $\hat{n}^{1\mu}$ dagegen auf S_μ ; dort gilt $\hat{n} = \hat{n}^{1\mu}$; für jedes Atom μ hat $\hat{n}^{1\mu}$ dieselben Multipolmomente Q_l^μ wie $n^{1\mu} - \tilde{n}^{1\mu}$
$\hat{n}_l^\alpha(r_\mu)$	radiale, um das Atom μ lokalisierte Funktion, deren l-tes Multipolmoment eins ist.
Q_l^μ	Multipolmomente Q_l^μ von $n^{1\mu} - \tilde{n}^{1\mu}$
$q_{\gamma,\gamma'}^\alpha$	Matrixelemente des Überlappoperators $\tilde{O}$ für Atome des Typs α ; diagonal in L
$D_{kin;\gamma,\gamma'}^\alpha$	Matrixelemente der kinetischen Energie für Atome des Typs α ; diagonal in L
$D_{Hxc;\gamma,\gamma'}^\mu$	Matrixelemente des elektrostatischen und Austausch-Korrelations-Potentials im Augmentationsgebiet des Atomes μ ; hängt von der aktuellen Dichte ab
$D_{\gamma,\gamma'}^\mu$	$= D_{kin;\gamma,\gamma'}^{\alpha(\mu)} + D_{Hxc;\gamma,\gamma'}^\mu$ effektive Matrixelemente des nichtlokalen Potentials des Atoms μ

Kapitel 4

Pseudopotentialerzeugung

Die PAW-Methode ist, wie schon betont wurde, keine Pseudopotentialmethode. Prinzipiell hat man in der Wahl der *partial waves* und der Projektoren eine große Freiheit. Holzwarth et. al. [HMD98] z.B. schlugen die Benutzung von Orthogonalpolynomen als Projektorfunktionen vor. Hier wird jedoch das von Blöchl angegebene (und cum grano salis auch von Vanderbilt und Kresse benutzte) Verfahren angewendet.

Um die Erfahrung bei (und die vorhandenen Programme zu) der Erzeugung von Pseudopotentialen nutzen zu können, ist es sinnvoll, echte und pseudisierte atomare Wellenfunktionen als *partial waves* zu benutzen und die Projektoren entsprechend zu definieren.

Da die Generierung der benötigten Funktionen daher über weite Strecken analog zur Erstellung eines Pseudopotentials vor sich geht und in Ermangelung eines besseren Sammelbegriffs wird hier der Begriff *Pseudopotentialerzeugung* verwendet.

Für jeden in einer *ab-initio*-Rechnung vorkommenden Atomtyp α müssen die AE- und PS-*partial waves* (ϕ_γ und $\tilde{\phi}_\gamma$), die Projektoren $\tilde{p}_\gamma$, der lokale Anteil des Potentials $\tilde{V}_{at}$ und die Matrizen der kinetischen Energie $\mathbf{D}_{kin}$ sowie des Überlapps $q_{\gamma,\gamma'}^{\alpha(\mu)}$ erzeugt werden.. Außerdem werden die core-Dichte n_c sowie – sofern man diese verwenden möchte – die *partial-core*-Dichte $\tilde{n}_c$ benötigt.

Die Generierung dieser Quantitäten geschieht in mehreren Schritten, wobei sie sich von der eines konventionellen Pseudopotentials vor allem durch einen Orthogonalisierungsschritt und ein anderes *Unscreening* unterscheidet.

4.1 Erzeugung der *partial waves*

Zunächst wird für jeden Drehimpuls l und gegebenenfalls jede Referenzenergie i ein Pseudopotential erzeugt. Dieser Schritt ist analog zur Erzeugung eines semilokalen Potentials (vgl. Kapitel 2.5). Im Falle mehrerer Energien pro Drehimpuls gibt es hier jedoch kein semilokales Potential im eigentlichen Sinne.

1. Zuerst wird das neutrale Atom selbstkonsistent (mit allen Elektronen) gelöst.

$$(T + V^{\text{ae}}(r) - \varepsilon_{nl}) \phi_{nl} = 0 \quad (4.1)$$

Man erhält also $V^{\text{ae}}(r) = V^{\text{ae}}[n[\{\phi_{nl}\}]](r)$ und ε_{nl} sowie ϕ_{nl} .

2. Dann wählt man einen Satz von Referenzenergien $\tilde{\varepsilon}_{l,i}$. Normalerweise sind darunter auch die Eigenwerte ε_{nl} der *Valenzelektronen*.
3. Dann bestimmt man die AE-*partial waves* $\phi_{l,i}(r)$. Für diejenigen $\tilde{\varepsilon}_{l,i}$, die *keine* Eigenwerte sind, bestimmt man die (nicht normierbaren, aber im Ursprung regulären) Lösungen von

$$(T + V^{\text{ae}}(r) - \tilde{\varepsilon}_{l,i}) \phi_{l,i}(r) = 0 \quad (4.2)$$

mit nunmehr fixem Potential $V^{\text{ae}}(r)$. Da diese Funktionen divergieren, löst man die Gleichung nur bis zu einem endlichen Radius $R_{\text{außen}}$. $R_{\text{außen}}$ ist größer oder gleich dem Augmentationradius R . Diese Möglichkeit wurde neu implementiert gemäß der Methode von Hamann [Ham89].

Für die Eigenwerte unter den Referenzenergien wählt man die Eigenzustände ϕ_{nl} als AE-*partial waves*.

4. Nun konstruiert man für alle l, i Pseudopotentiale $\tilde{V}_{l,i}(r)$ und Wellenfunktionen $\tilde{\phi}_{l,i}(r)$, so daß

- (a) die Gleichung

$$(T + \tilde{V}_{l,i}(r) - \tilde{\varepsilon}_{l,i}) \tilde{\phi}_{l,i}(r) = 0 \quad (4.3)$$

erfüllt wird,

- (b) außerhalb des Augmentationsradius R die Funktionen

$$\tilde{V}_{l,i}(r) = V^{\text{ae}}(r) \quad (\forall r > R) \quad (4.4)$$

$$\tilde{\phi}_{l,i}(r) = \phi_{l,i}(r) \quad (\forall r > R) \quad (4.5)$$

übereinstimmen,

- (c) die Funktion $\tilde{\phi}_{l,i}(r)$ keinen Knoten hat und bei $r = 0$ endlich ist,
- (d) das Potential $\tilde{V}_{l,i}(r)$ keine Divergenz bei $r = 0$ hat und
- (e) (für normerhaltende Potentiale) die Funktion $\tilde{\phi}_{l,i}(r)$ die Norm eins hat.

Den letzten Schritt kann man auf verschiedene Weisen durchführen. Kurz beschrieben werden die in unserem Programm implementierten Methoden nach Bachelet, Hamann und Schlüter (BHS) [BHS82] mit Modifikationen nach Vanderbilt (VAN) [Van85] und jene nach Rappe, Rabe, Kaxiras und Joannopoulos (RRKJ) [RRKJ90]. Eine ausführlichere Beschreibung findet man in [Eng92, Fri98].

BHS-Methode

Nach BHS erfolgen (für jedes (l, i)) drei Schritte. Sind zwei Möglichkeiten angegeben, so ist die erste die nach BHS, die zweite jene nach Vanderbilt. Diese Wahlmöglichkeiten werden im Programm durch die angegebenen Kontrollparameter $(k_1^{van}, k_2^{van}, k^{cut})$ gesteuert.

- i. Zuerst wird die $1/r$ -Singularität des Potentials $V(r)$ bei $r = 0$ beseitigt; dies geschieht entweder durch das glatte Anpassen eines geraden Polynoms vierter Ordnung innerhalb eines gewissen Radius' $r'_{cut} < R$ oder durch das Multiplizieren mit einer Abschneidefunktion (geschmolzene Stufe bei $r_{cut} < R$).

$$V_{cut}(r) = \begin{cases} (1 - f(r)) V_{at}(r) + V_{cut}^0 f(r) & (k_1^{van} = 0) \\ \begin{cases} a_0 + a_2 r^2 + a_4 r^4 & (r < r'_{cut}) \\ V(r) & (r > r'_{cut}) \end{cases} & (k_1^{van} = 1) \end{cases}$$

(Der V_{cut}^0 -Term spielt nur dann eine Rolle, wenn man Blöchl folgt und $V_{cut}(r)$ als lokales Potential benutzt.)

- ii. Dann wird das Potential durch Addition eines Vielfachen der Abschneidefunktion so modifiziert, daß es zur Energie $\tilde{\varepsilon}_{l,i}$ eine knotenlose Lösung $\xi(r)$ gibt, die außerhalb von R mit $\phi_{l,i}(r)$ übereinstimmt.

$$V_{cl}(r) = V_{cut}(r) + c_l f(r) \quad (4.6)$$

Für nicht normerhaltende Potentiale ist die Prozedur damit beendet und $\tilde{\phi}_{l,i}(r) = \xi(r)$ ist die PS-*partial wave*.

- iii. Für normerhaltende Potentiale muß die Funktion ξ durch Manipulation innerhalb des Abschneideradius' noch so verändert werden, daß sie die richtige Norm erhält; dies muß so geschehen, daß sie das korrekte Verhalten für $r \rightarrow 0$ hat ($\sim r^{l+1}$).

$$\tilde{\phi}_{l,i}(r) = \xi(r) + \delta f(r) \begin{cases} r^{l+1} & (k_2^{van} = 0) \\ \xi(r) & (k_2^{van} = 1) \end{cases} \quad (4.7)$$

Der Faktor δ wird aus der Normierung $\int_0^\infty (\tilde{\phi}_{l,i}(r))^2 dr = 1$ bestimmt.

Das Potential, das diese knotenlose Funktion als Lösung hat, ergibt sich aus der Invertierung der Schrödingergleichung:

$$\tilde{V}_{l,i}(r) = \tilde{\varepsilon}_{l,i} - \frac{T \tilde{\phi}_{l,i}(r)}{\tilde{\phi}_{l,i}(r)} \quad (4.8)$$

Die Parameter, die dabei zur Verfügung stehen, bestimmen vor allem die Abschneidefunktion $f(r)$, d.h. deren Lage und Form:

$$f(r) = \begin{cases} \exp\left(-\left(\frac{r}{r_{cut}}\right)^\lambda\right) & (k^{cut} = 1) \\ 100^{-\left(\frac{\sinh\left(\frac{b}{1.5} \frac{r}{r_{cut}}\right)}{\sinh(b)}\right)^\lambda} & (k^{cut} = 2) \end{cases} \quad (4.9)$$

Im Prinzip können alle verwendeten Abschneidefunktionen untereinander verschieden sein. Alle r_{cut} müssen so gewählt sein, daß die Abschneidefunktionen bei $r = R$ verschwinden.

RRKJ-Methode

Im Gegensatz zum BHS-Verfahren modifiziert die RRKJ-Methode unmittelbar die Wellenfunktionen. Innerhalb des Abschneideradius' wird für diese ein analytischer Ansatz aus einer Summe von sphärischen Besselfunktionen gemacht:

$$\tilde{\phi}_{l,i}(\vec{r}) = \begin{cases} \sum_{k=1}^K \alpha_{l,i}^k r j_l(q_k r) & (r < r_{cut}) \\ \phi_{l,i}(\vec{r}) & (r > r_{cut}) \end{cases} \quad (4.10)$$

Die $\{q_k\}$ werden so bestimmt, daß jede einzelne Besselfunktion die richtige Logarithmische Ableitung bei $r = r_{cut}$ hat. Die $\{\alpha_{l,i}^k\}$ werden so bestimmt, daß $\tilde{\phi}_{l,i}(\vec{r})$ und deren zweite Ableitung, $\tilde{\phi}_{l,i}''(\vec{r})$, stetig bei $r = r_{cut}$ sind. (Die erste Ableitung ist es nach Konstruktion der $\{q_k\}$ ohnehin.) Wenn man keine Normerhaltung fordert,

so sind diese Bedingungen mit $K = 2$ Besselfunktionen zu erfüllen; fordert man die richtige Normierung, so sind mindestens eine oder manchmal zwei weitere Funktionen erforderlich. Die in [RRKJ90] vorgeschlagene Addition weiterer Funktionen zur Verbesserung der Weichheit hat sich als nicht empfehlenswert herausgestellt (vgl. [Eng92, KH94]). Die zu diesen Pseudowellenfunktionen gehörigen Potentiale erhält man wie im BHS-Verfahren durch Inversion der Schrödingergleichung.

BHS und RRKJ im Vergleich

Die Abschneideradien r_{cut} haben bei BHS- bzw. RRKJ-Verfahren unterschiedliche Bedeutung; die Potentiale sind i. Allg. vergleichbar, wenn der BHS-Radius um einen Faktor 1.5 kleiner ist [ERBS]. Die BHS-Methode liefert manchmal etwas weichere Potentiale; ihre praktische Konstruktion gestaltet sich aber schwierig auf Grund der Vielzahl an (z.T. diskreten) Parametern und der z.T. komplexen Abhängigkeit von ihnen.

Beschränkt man sich bei der RRKJ-Methode auf die minimale Anzahl von Funktionen (d.h. auf zwei für nicht-normerhaltende Potentiale und drei oder vier für normerhaltende), so ist die Anzahl der Parameter gering und man erhält hinreichend weiche Potentiale. Zu wählen ist im Wesentlichen ein Abschneideradius pro Wellenfunktion, und die Abhängigkeit davon ist meist gut durchschaubar; der benötigte *Cutoff* wächst i. Allg. monoton mit schrumpfendem Radius.

Bei RRKJ-Potentialen stimmen ab dem Abschneideradius AE- bzw. PS-Wellenfunktionen und Potentiale *exakt* überein. Für die PAW-Methode ist aus diesem Grunde diese Art der Konstruktion vorzuziehen, da ansonsten die gegenseitige Annullierung von AE- und PS-Funktionen in diversen Integralen nicht exakt gewährleistet ist.

RRKJ-Potentiale sind i. Allg. bei r_{cut} zwar stetig, haben jedoch einen Knick. Damit die Projektoren (s.u.) keinen weiteren Knick innerhalb des Abschneideradius' bekommen, empfiehlt es sich im Regelfalle, denselben Radius für alle Drehimpulse und Energien (l, i) zu verwenden.

Der entscheidende Punkt ist jedoch, daß die mit der BHS-Methode generierten Potentiale sich als problematisch herausstellten, sobald mehr als eine Energie pro Drehimpuls verwendet wurde. Die Ursache ist offenbar, daß die auf diese Weise erzeugten Projektoren dazu neigen, nach der Orthogonalisierung (s.u.) sehr scharfe Oszillationen nahe dem Abschneideradius zu zeigen, welche in der Darstellung durch (relativ wenige) Ebene Wellen auch nicht ansatzweise wiedergegeben werden können.

4.2 Erzeugung der Projektoren

Von den nun bestimmten Potentialen $\tilde{V}_{l,i}$ wird eines als lokaler Potentialanteil ausgewählt. Der erste Schritt der Projektordefinition entspricht der Konstruktion von KB-Potentialen (vgl. Kapitel 2.5). Ein Analogon der nachfolgenden Orthogonalisierung gibt es bei jenen jedoch nicht.

In diesem Abschnitt müssen die Funktionen vor und nach der Orthogonalisierung unterschieden werden. Die vorläufigen Funktionen *vor* der Orthogonalisierung werden durch einen Balken gekennzeichnet (z.B. $\bar{\phi}$, $\bar{\bar{\phi}}$).

Definition vorläufiger Projektoren

Ausgehend von den nunmehr konstruierten PS-*partial waves* definiert man nun vorläufige Projektoren:

$$\bar{\bar{p}}_{l,i}(r) = \left(\tilde{V}_{l,i} - \tilde{V}_{loc} \right) \bar{\phi}_{l,i}(r) \quad (4.11)$$

Da die Potentiale jenseits des Radius' R übereinstimmen, verschwinden diese Funktionen außerhalb der so definierten Kugeln.

Falls man die Potentiale nach dem BHS-Verfahren konstruiert und dieselbe Abschneidefunktion $f(r)$ für $\tilde{V}_{l,i}$ und $\tilde{V}_{loc}$ verwendet, so gilt

$$\bar{\bar{p}}_{l,i} = \text{const. } f(r) \bar{\phi}_{l,i}(r)$$

und die Projektoren besitzen offensichtlich keine Nullstelle.

Verwendet man zur Konstruktion der Potentiale verschiedene atomare Konfigurationen, so verschwinden die Projektoren nicht jenseits von R . Insbesondere gilt dies für ionische Konfigurationen, da sie ein anderes langreichweitiges Verhalten zeigen. Da man im weiteren Verlauf Skalarprodukte der Projektoren auch mit ungebundenen Zuständen bilden muß, die ab einem gewissen Radius exponentiell divergieren, ist dies intolerabel. Zur Behebung der Schwierigkeit könnte man darauf verfallen, die Projektoren mittels der Differenz der *nackten* Potentiale zu bilden (siehe Abschnitt 4.3). Diese stimmen außerhalb von R überein, da der Unterschied im Abschirmpotential enthalten ist. Jedoch führt dies zu einer asymmetrischen Matrix D_{ion} (s.u.), was gänzlich unerwünscht ist. Daher ist es essentiell, daß alle Funktionen ausgehend von *einer* Konfiguration (nämlich im Regelfalle dem Grundzustand des neutralen Atoms) konstruiert werden.

Orthogonalisierung der Projektoren

Nun werden für jeden Drehimpuls l die Projektoren und die Wellenfunktionen zueinander orthogonalisiert, d.h. sie müssen der Orthogonalitätsrelation

$$\langle \tilde{\phi}_i | \tilde{p}_j \rangle = \delta_{i,j} \quad (4.12)$$

gehoren (die Funktionen zu unterschiedlichen Drehimpulsen l erfüllen diese ohnehin). Die *partial waves* und die Projektoren vor der Orthogonalisierung werden hier (unter Auslassung des l -Indexes) mit $\bar{\phi}_i$, $\tilde{\phi}_i$ und $\tilde{p}_i$ bezeichnet. Zur Erfüllung der Orthogonalitätsrelation kann man die endgültigen $\{\tilde{p}_i\}$ als Linearkombinationen der Funktionen $\{\tilde{\phi}_i\}$ wählen und ebenso die $\{\tilde{\phi}_i\}$ im Raum der $\{\tilde{\phi}_i\}$.

Die allgemeinste Transformation dieser Funktionen ist

$$\tilde{p}_i = \sum_{i'} M_{ii'} \tilde{p}_{i'} \quad ; \quad \tilde{\phi}_i = \sum_{i'} N_{ii'} \tilde{\phi}_{i'}, \quad (4.13)$$

wobei $\mathbf{M}$, $\mathbf{N}$ reell und nichtsingulär sind. Die Orthogonalität erlegt n^2 Bedingungen auf:

$$\delta_{i,j} = \sum_{i',j'} N_{ii'} \underbrace{\langle \tilde{\phi}_{i'} | \tilde{p}_{j'} \rangle}_{\bar{B}_{i'j'}} M_{jj'} = (\mathbf{N} \bar{\mathbf{B}} \mathbf{M}^T)_{ij} \quad (4.14)$$

Für den späteren Gebrauch sei hier schon festgehalten, daß sich daraus die Relation

$$\bar{\mathbf{B}}^{-1} = \mathbf{M}^T \mathbf{N} \quad (4.15)$$

ergibt.

Es gibt viele Möglichkeiten, diese Transformation zu definieren.

- Vanderbilt [Van90] wählt $\mathbf{N} = \mathbf{1}$, also $\mathbf{M} = \bar{\mathbf{B}}^{-1}$. Das bedeutet, daß die Wellenfunktionen unverändert bleiben, während i. Allg. alle Projektoren verändert werden.
- Blöchl [Blö94] beschreibt ein Gram-Schmidt-Verfahren; d.h. $\mathbf{M}$ und $\mathbf{N}$ werden als untere Dreiecksmatrizen gewählt; i.e. $\mathbf{N}^{-1} (\mathbf{M}^T)^{-1}$ ist LU-Zerlegung¹ von $\bar{\mathbf{B}}$. Die n freien Bedingungen ermöglichen eine Skalierung

$$\mathbf{N} \mapsto \text{diag}(d_1, \dots, d_n) \mathbf{N}, \quad \mathbf{M} \mapsto \text{diag}(d_1^{-1}, \dots, d_n^{-1}) \mathbf{M}$$

¹Jede Matrix kann in das Produkt einer unteren (L : lower) und einer oberen (U : upper) Dreiecksmatrix zerlegt werden. Die Inverse einer (oberen/unteren) Dreiecksmatrix ist wieder (obere/untere) Dreiecksmatrix. Vgl. [PTVF92].

zur Vermeidung sehr großer oder sehr kleiner Zahlenwerte.²

Bei diesem Verfahren bleiben die erste PS-*partial wave* und der erste Projektor unverändert. Da die PS-*partial waves* und i. Allg. der erste Projektor keine Nullstellen haben, müssen die übrigen Projektoren mindestens einen Vorzeichenwechsel haben, um die Orthogonalitätsbedingung zu erfüllen.

- Ausgangspunkt für eine andere Transformation ist die Überlegung, daß die Projektoren fouriertransformiert werden, die *partial waves* jedoch nur im Ortsraum benutzt werden. Daher schaden Nullstellen der *partial waves* nicht, wohingegen solche der Projektoren i. Allg. eine kleinere charakteristische Längenskala bedeuten und mithin höhere Fourierkomponenten. Also versucht man **M** so zu bestimmen, daß die Projektoren möglichst gut fouriertransformierbar sind.

Die AE-*partial waves* werden genauso transformiert wie die PS-*partial waves*

$$\phi_i = \sum_{i'} N_{ii'} \bar{\phi}_{i'}, \quad (4.16)$$

so daß zueinandergehörige Funktionen auch nach der Orthogonalisierung jenseits des Augmentationsradius³ übereinstimmen.

Matrizen der kinetischen Energie und des Überlapps

Die vorläufige Matrix der kinetischen Energie

$$\bar{D}_{kin;ij} = \langle \bar{\phi}_i | T | \bar{\phi}_j \rangle - \langle \bar{\tilde{\phi}}_i | T | \bar{\tilde{\phi}}_j \rangle \quad (4.17)$$

wird in der Praxis bestimmt nach

$$\bar{D}_{kin;ij} = \langle \bar{\phi}_i | (\tilde{\varepsilon}_j - V_{at}) | \bar{\phi}_j \rangle - \langle \bar{\tilde{\phi}}_i | (\tilde{\varepsilon}_j - \tilde{V}_j) | \bar{\tilde{\phi}}_j \rangle \quad (4.18)$$

und transformiert sich bei der Orthogonalisierung gemäß

$$\mathbf{D}_{kin} = \mathbf{N} \bar{\mathbf{D}}_{kin} \mathbf{N}^T. \quad (4.19)$$

Ebenso transformieren sich die Matrizen des Überlapps

$$\bar{q}_{i,j} = \langle \bar{\phi}_i | \bar{\phi}_j \rangle - \langle \bar{\tilde{\phi}}_i | \bar{\tilde{\phi}}_j \rangle \quad (4.20)$$

und des lokalen Potentials

²*diag*($d_1, \dots, d_n$) bezeichnet die Diagonalmatrix, die aus den Diagonalelementen d_1 bis d_n besteht.

$$\bar{D}_{Hxc;ij} = \langle \bar{\phi}_i | V^{ae} | \bar{\phi}_j \rangle - \langle \tilde{\phi}_i | \tilde{V}_{loc} | \tilde{\phi}_j \rangle, \quad (4.21)$$

so daß letztlich für $\bar{\mathbf{q}}$ und $\bar{\mathbf{D}} = \bar{\mathbf{D}}_{kin} + \bar{\mathbf{D}}_{Hxc;ij}$ gilt:

$$\mathbf{q} = \mathbf{N} \bar{\mathbf{q}} \mathbf{N}^T \quad ; \quad \mathbf{D} = \mathbf{N} \bar{\mathbf{D}} \mathbf{N}^T. \quad (4.22)$$

Zur Definition der Matrix $\mathbf{D}$ vergleiche man Glg. (3.146).

Warum wird die Schrödingergleichung für die ursprünglichen PS-*partial waves* exakt gelöst?

Mit den so konstruierten Funktionen sind die PS-*partial waves* gerade die Lösungen der atomaren Schrödingergleichung

$$\left(T + \tilde{V}_{loc} + \sum_{ij} |\tilde{p}_i\rangle D_{ij} \langle \tilde{p}_j| \right) |\tilde{\phi}_k\rangle = \varepsilon_k \left(\mathbf{1} + \sum_{ij} |\tilde{p}_i\rangle q_{ij} \langle \tilde{p}_j| \right) |\tilde{\phi}_k\rangle, \quad (4.23)$$

wobei $V_{NL} = \sum_{ij} |\tilde{p}_i\rangle D_{ij} \langle \tilde{p}_j|$ das nichtlokale Potential und $O = \mathbf{1} + \sum_{ij} |\tilde{p}_i\rangle q_{ij} \langle \tilde{p}_j|$ der Überlappoperator ist (vgl. Kapitel 3).

Um zu verifizieren, daß Glg. (4.23) erfüllt ist, kann man in folgenden Schritten vorgehen:

1. Zwischen den Matrizen $\bar{\mathbf{D}}$, $\bar{\mathbf{B}}$ und $\bar{\mathbf{q}}$ besteht folgende Beziehung

$$\bar{D}_{ij} = \bar{B}_{ij} + \varepsilon_j \bar{q}_{ij} \quad ,$$

wie man leicht aus Glg. (4.18) und der Definition der vorläufigen Projektoren Glg. (4.11) ersieht.³

2. Nichtlokales Potential und Überlappmatrix lassen sich unter Benutzung von Glg. (4.15) durch die ursprünglichen Funktionen ausdrücken als:

$$\begin{aligned} V_{NL} &= \sum_{ij} |\tilde{p}_i\rangle D_{ij} \langle \tilde{p}_j| = \sum_{ij} |\tilde{p}_i\rangle [\mathbf{M}^T \mathbf{N} \bar{\mathbf{D}} \mathbf{N}^T \mathbf{M}]_{ij} \langle \tilde{p}_j| \\ &= \sum_{ij} |\tilde{p}_i\rangle [\bar{\mathbf{B}}^{-1} \bar{\mathbf{D}} (\bar{\mathbf{B}}^{-1})^T]_{ij} \langle \tilde{p}_j| \\ O - \mathbf{1} &= \sum_{ij} |\tilde{p}_i\rangle q_{ij} \langle \tilde{p}_j| = \sum_{ij} |\tilde{p}_i\rangle [\bar{\mathbf{B}}^{-1} \bar{\mathbf{q}} (\bar{\mathbf{B}}^{-1})^T]_{ij} \langle \tilde{p}_j| \end{aligned}$$

³Bei Vanderbilt wird D_{ij} gerade so definiert. (Wegen $\mathbf{N} = \mathbf{1}$ gilt dort $\mathbf{D} = \bar{\mathbf{D}}$ und $\mathbf{q} = \bar{\mathbf{q}}$).

3. Also erhält man:

$$\begin{aligned}
0 &\doteq \left(T + \tilde{V}_{loc} + V_{NL} - \varepsilon_k O \right) |\tilde{\phi}_k\rangle \\
&= \left(T + \tilde{V}_{loc} - \varepsilon_k \right) |\tilde{\phi}_k\rangle \\
&+ \sum_{ij} |\tilde{p}_i\rangle [\bar{\mathbf{B}}^{-1} (\bar{\mathbf{D}} - \varepsilon_k \bar{\mathbf{q}}) (\bar{\mathbf{B}}^{-1})^T]_{ij} \underbrace{\langle \tilde{p}_j | \tilde{\phi}_k \rangle}_{=(\bar{\mathbf{B}}^T)_{jk}} \\
&= \left(T + \tilde{V}_{loc} - \varepsilon_k \right) |\tilde{\phi}_k\rangle + \left(\tilde{V}_k - \tilde{V}_{loc} \right) |\tilde{\phi}_k\rangle
\end{aligned} \tag{4.24}$$

4.3 *Unscreening*

Das lokale Potential $\tilde{V}_{loc}$ und die Matrix $\mathbf{D}$ enthalten den Beitrag aller Elektronen in der Ladungsverteilung der gewählten Ausgangskonfiguration. In der Festkörperrechnung mit dem Molekuldynamikprogramm wird die Ladungsdichte und das dazugehörige Potential dynamisch erzeugt. Deshalb braucht man als externes Potential das Potential des „nackten“ Ions, d.h. ohne Valenzelektronen. Das wird durch Subtraktion des Abschirmpotentials erreicht (*Unscreening*).

Da im Molekuldynamikprogramm das Hartreepotential für Valenzladung plus Pseudoladung berechnet wird, das Austausch-Korrelations-Potential jedoch für die glatte Ladungsdichte (gegebenenfalls plus *partial-core*-Dichte), definiert man das „nackte“ atomare Pseudopotential entsprechend:

$$\tilde{V}_{at}(r) = \tilde{V}_{loc}(r) - V_H[\tilde{n}_v + \hat{n}](r) - V_{xc}[\tilde{n}_v + \tilde{n}_c](r) \tag{4.25}$$

Da das Potential V^{ae} bei $r = 0$ divergiert, es aber nur in Matricelementen der *echten* Wellenfunktionen auftritt (dort aber immer), ist es numerisch vorteilhaft, statt der Matricelemente $\mathbf{D}_{kin}$ die Größe

$$\mathbf{D} = \langle \phi_i | T + V_{at} | \phi_j \rangle - \langle \tilde{\phi}_i | T + \tilde{V}_{at} | \tilde{\phi}_j \rangle \tag{4.26}$$

im Bandstrukturprogramm einzulesen, wobei das echte atomare Potential

$$V_{at}(r) = V^{ae} - V_H[n_v](r) - V_{xc}[n_v + n_c](r) = \int_{\infty}^r d^3r' \frac{n_z + n_c}{|\vec{r} - \vec{r}'|} \tag{4.27}$$

analog zu Glg. (4.25) gebildet wird.⁴

⁴Trotz identischer Schreibweise unterscheiden sich die Potentiale V_{at} und $\tilde{V}_{at}$ von denen in [Blö94] grundlegend; es handelt sich hier um ionische Potentiale.

Da die Pseudoladung $\hat{n}$ gemäß Kapitel 3.11 optimal für einen gewissen Abschneidewert $G_{\max}$ konstruiert wird, und da im *Unscreening* dasselbe Potential verwendet werden sollte, wie es im Bandstrukturprogramm gebraucht wird, ergibt sich lästigerweise eine Abhängigkeit des Pseudopotentials vom verwendeten $G_{\max}$. Diese Abhängigkeit erwies sich bei Tests als relativ gering; um optimale Ergebnisse zu erzielen, sollte das Potential jedoch mit dem $G_{\max}$ generiert werden, welches in der Bandstrukturechnung benutzt werden soll.

4.4 Die praktische Vorgehensweise

Der Ausgangspunkt für die Konstruktion eines Pseudopotentials ist natürlich das physikalische System, dessen Eigenschaften man zu berechnen gedenkt. Dieses bestimmt einerseits, welche Drehimpulskomponenten zu berücksichtigen sind, d.h. welche Zustände als Valenzzustände zu gelten haben, die dann in der Festkörperrechnung selbstkonsistent neu berechnet werden. Auch kann man die Ausgangskonfiguration so wählen, daß sie der erwarteten chemischen Umgebung im Festkörper möglichst gut angemessen ist. Hier wurde jedoch in allen Fällen der atomare Grundzustand als Ausgangspunkt gewählt.

Der zweite wichtige Gesichtspunkt ist der minimale Abstand, der in diesem System zwischen zwei Atomen auftreten kann. Dieser bestimmt eine Obergrenze der Augmentationsradien. Im Gegensatz zur Pseudopotentialmethode, wo der Kovalente Radius oder das äußere Maximum der echten Wellenfunktion eine Obergrenze für den Rumpfradius darstellen, ist für die PAW-Methode die Begrenzung durch die Bedingung gegeben, daß die Augmentationsregionen der Atome nicht überlappen dürfen. Besteht ein System jedoch z.B. aus einem Atom mit großem und einem mit kleinem Kovalenten Radius, so ist es möglich, für beide Atome einen mittleren Radius zu wählen, um die PS-*partial waves* des weniger ausgedehnten Partners möglichst glatt wählen zu können.

In [FKBK00] wird auch von Versuchen mit Ultrasoften Potentials berichtet, in denen diese Bereiche überlappen, ohne daß dies zu drastischen Verschlechterungen der Ergebnisse führt. Obwohl in der hier vorgestellten Methode die Konstruktion des Hartree-Potentials auf eine Weise geschieht, die streng genommen disjunkte Kugeln voraussetzt, sind auch hier nur kleine Fehler zu erwarten, da der Unterschied zwischen AE- und PS-*partial waves* auf Grund ihrer Konstruktion für Radien knapp unterhalb des Abschneideradius' immer noch sehr klein ist. Jedenfalls aber gibt der kürzeste Abstand einen Richtwert an.

Um eine Vorstellung von der erreichbaren *Cutoff*-Energie zu bekommen, sollte man das Rappe-Kriterium [RRKJ90] für eine Reihe von Radien innerhalb des sinnvollen Bereiches testen. Bei diesem Kriterium wird die Konvergenz der Fourierkomponenten der *kinetischen* Energie der Pseudowellenfunktionen geprüft. Dies hat sich als gutes Maß für die Konvergenz in Festkörperrechnungen herausgestellt. Das Programm hierfür wurde für die PAW-Methode angepaßt.

Bedauerlicherweise ist das Kriterium nur für gebundene Zustände anwendbar. Daher sind – bei Verwendung ungebundener Zustände – die Vorhersagen nicht immer so akkurat wie es bei den bisherigen Pseudopotentialen der Fall war.

Das lokale Potential gilt – wie bei KB-Potentialen – für alle Drehimpulskomponenten, für die keine Projektoren enthalten sind. Im Allg. ist es daher sinnvoll, den höchsten Drehimpuls als lokales Potential zu wählen – es sei denn, dies ist gerade ein problematischer (knotenloser) Fall. Das lokale Potential sollte normerhaltend konstruiert werden. Existiert kein geeigneter gebundener Zustand, so kann es aus einem ungebundenen Zustand konstruiert werden.

Festzulegen ist sodann, wieviele Projektoren für welche Drehimpulse benutzt werden sollen, und zu welchen Referenzenergien. Theoretisch lassen sich beliebig viele Funktionen pro Drehimpuls verwenden; praktisch jedoch führt dies, weil die Zustände dann fast linear abhängig sind, zu numerischen Schwierigkeiten (sehr kleine Zahlen beim Orthogonalisierungsschritt) und im Endeffekt eher zu einer Verschlechterung als einer Verbesserung [Mor]. Praktisch hat man daher die Wahl zwischen einem oder zwei Projektoren pro Drehimpuls. Zur deren Orthogonalisierung wurde im Regelfall das Verfahren von Blöchl benutzt. Versuche mit anderen Orthogonalisierungsverfahren (siehe Abschnitt 4.2) erbrachten keine Verbesserung.

Die Anzahl der Projektoren bestimmt die Streueigenschaften; weicht die Logarithmische Ableitung im relevanten Energiebereich zu stark von derjenigen der echten Wellenfunktion ab, so kann man eine Verbesserung durch eine zweite Referenzenergie in diesem Bereich erzwingen. Eine alternative Betrachtungsweise liefert die Wahl zweier eng benachbarter Referenzenergien; durch die im Prozeß der Orthonormierung enthaltene Differenzbildung ist die zweite Wellenfunktion näherungsweise die Energieableitung $\dot{\phi}$ bei der ersten Referenzenergie. Aus der Erfahrung mit LAPW-Programmen ist bekannt, daß sich mit diesen beiden Funktionen das Verhalten auf einem umgebenden Energieintervall gut beschreiben läßt.

Außerdem bestimmt die Anzahl der Projektoren auch maßgeblich, wie gut sich die Dichte im Augmentationsgebiet beschreiben läßt und damit, wie gut die theoretisch gegebene Übereinstimmung von $\tilde{n}$ und $\tilde{n}^1$ ist.

Auskunft hierüber erhält man aus dem Programm zur Bestimmung der Logarithmischen Ableitung, welches zum Test der PAW-Datensätze entwickelt wurde. Des weiteren wurden im – ebenfalls neu entwickelten – Programm zur Berechnung eines PAW-Atomes zwei Möglichkeiten zur Ladungsberechnung vorgesehen, aus deren Differenz sich Rückschlüsse auf die Güte der Ladungsdarstellung ziehen lassen.

Durch eine gründliche Untersuchung der Logarithmischen Ableitungen des PAW-Atoms kann man sicherstellen, daß keine *ghost-states* auftreten. Man erkennt sie als Polstellen der Logarithmischen Ableitungen; diese können extrem scharf sein und dadurch bei zu geringer Auflösung übersehen werden.

Als Test werden für verschiedene Konfigurationen das Atomproblem in AE-Rechnung und als PAW-Rechnung gelöst. Es sollten zum einen die Eigenwerte jeweils richtig bestimmt werden; zum anderen müssen die Anregungsenergien zwischen den verschiedenen Konfigurationen übereinstimmen.

4.5 Kohlenstoff

Als erstes wurden Pseudopotentiale für Kohlenstoff generiert. Dabei wurden drei mögliche Anwendungsfälle anvisiert: Pseudopotentiale sollen

1. in einem C₂-Dimer als Beispiel eines Moleküles,
2. in Diamant als Beispiel eines Festkörpers, welcher nur aus kleinen Atomen besteht,
3. in SiC als Beispiel eines Festkörpers, in dem nächste Nachbarn jeweils unterschiedliche Kovalente Radien aufweisen,

die Möglichkeiten der PAW-Methode ausloten. Das im ersten Fall untersuchte Molekül weist eine extrem kurze Bindungslänge von $d_{NN} = 2.36$ a.u. auf; im zweiten Falle ist der Abstand mit $d_{NN} = 2.92$ a.u. schon erheblich größer; im dritten Falle schließlich sind nächste Nachbarn bedingt durch die ausladenderen Siliziumatome noch mal deutlich weiter ($d_{NN} = 3.57$ a.u.) voneinander getrennt. Entsprechend werden drei verschiedene Pseudopotentiale konstruiert, welche (in der Reihenfolge ihrer minimalen Bindungsradien) mit C_s , C_m und C_l bezeichnet werden.

Während sich die $2s$ -Wellenfunktion auf Grund ihrer Nullstelle gut pseudisieren läßt, ist es für den $2p$ -Zustand bei Erhaltung der Norm schwierig, eine weiche Pseudowellenfunktion zu generieren.

Für $2p$ - und $2s$ -Zustände werden daher nicht-normerhaltende (NNC) PS-*partial waves* verwendet. Das lokale Potential wird aus einem ungebundenen d -Zustand erzeugt. Als Konstruktionsverfahren für die PS-*partial waves* wird das RRKJ-Verfahren verwendet.

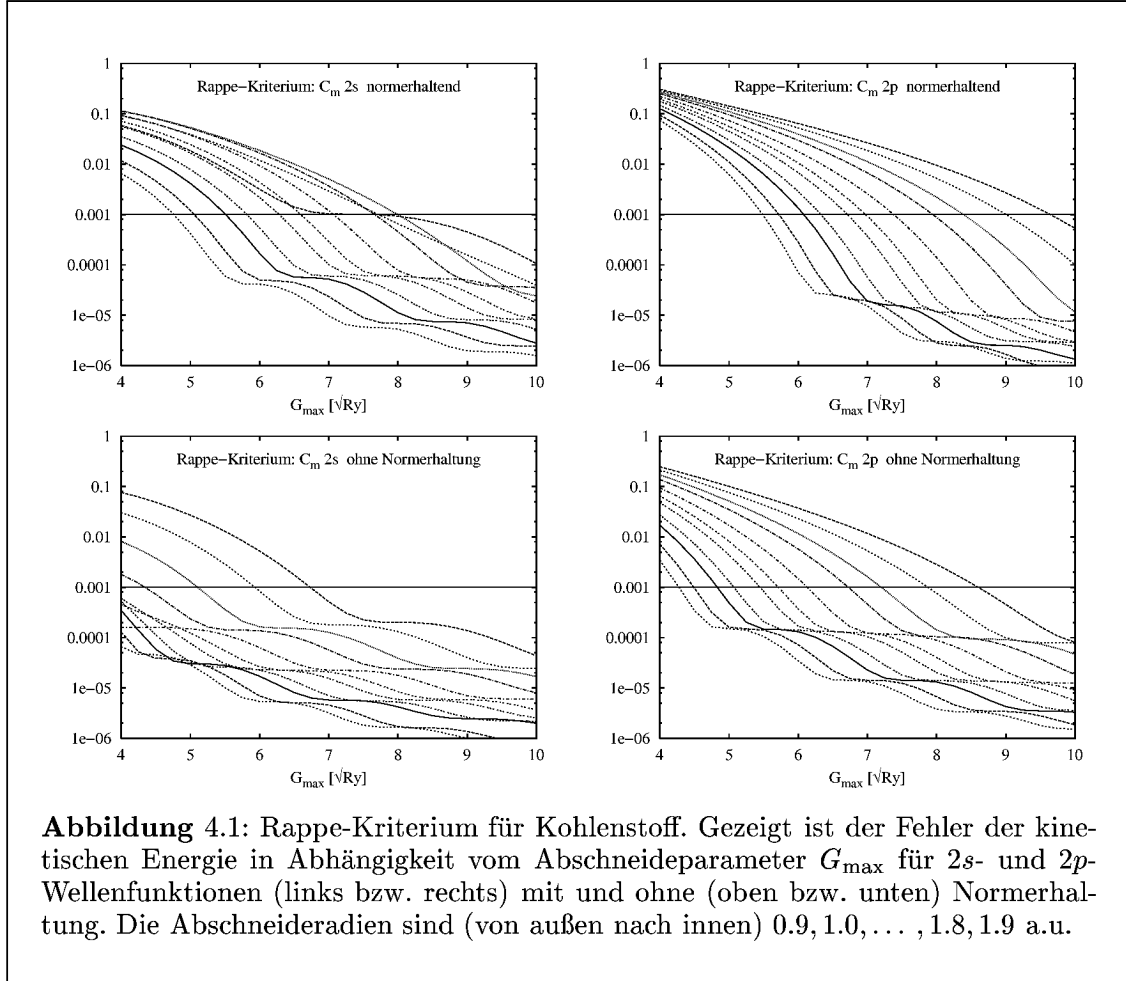


Abbildung 4.1 zeigt die nach dem Rappe-Kriterium zu erwartenden Abschneideenergien für eine Reihe von Rumpfradien zwischen 0.9 und 1.9 a.u. Zunächst erkennt man, daß die $2s$ -Funktion sich im normerhaltenden Fall nicht sonderlich gut pseudisieren läßt, ohne Normerhaltung aber so gut, daß sich alle weiteren Betrachtungen auf die $2p$ -Welle konzentrieren können.

Die für die $2p$ -Zustände benötigte *Cutoff*-Energie fällt monoton mit dem Radius. Das entspricht der Erwartung und ist für RRKJ-Potentiale der Regelfall.

Das üblicherweise angelegte Kriterium für die Genauigkeit ist, daß der Fehler kleiner als 1 mRy pro Atom sein sollte. (Kresse et al. [FKBK00] halten 1 mRy *pro Elektron*

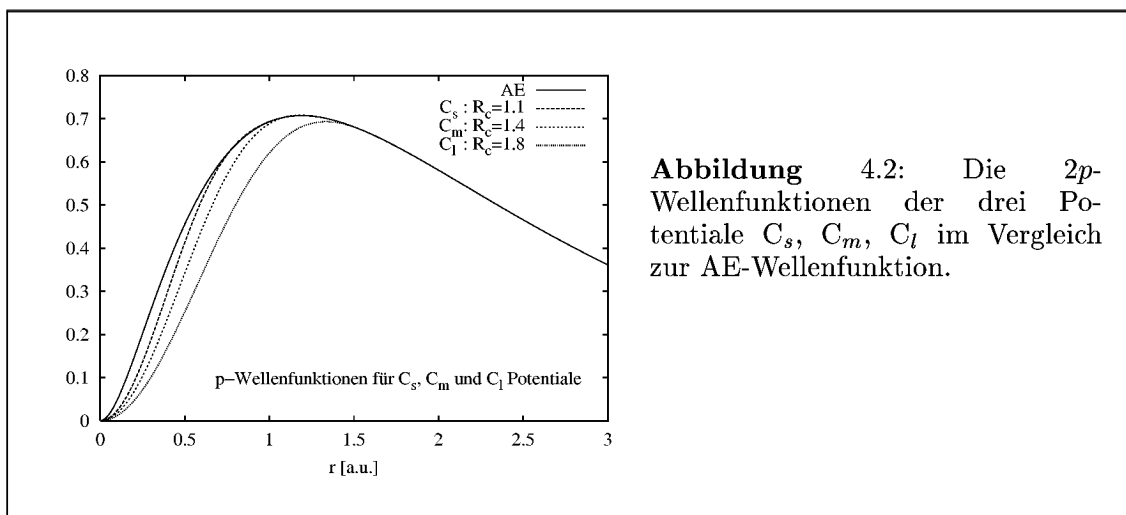
in vielen Fällen für hinreichend). Nimmt man der Einfachheit halber die 1 mRy-Marke als Referenz, so erkennt man, daß die Reduktion des Abschneidewertes $G_{\max}$ bei gleichem Rumpfradius ca. $1 - 1.5 \sqrt{Ry}$ beträgt.

Einen Beitrag zur Weichheit erhält man also erwartungsgemäß dadurch, daß die Pseudowellenfunktion zum *gleichen* Rumpfradius glatter ist. Ein weiterer Gewinn an Glattheit entsteht in vielen Fällen für die PAW-Methode dadurch, daß man im Gegensatz zur Pseudopotentialtheorie nicht an das äußere Maximum der Wellenfunktionen (bei ca. 1.2 a.u.) als Obergrenze des wählbaren Bereiches für den Abschneideradius gebunden ist. Für die C_s - und C_m -Potentiale resultiert die Verkleinerung der *Cutoff*-Energie nur aus ersterer Quelle, während für C_l auch der zweite Effekt wichtig ist.

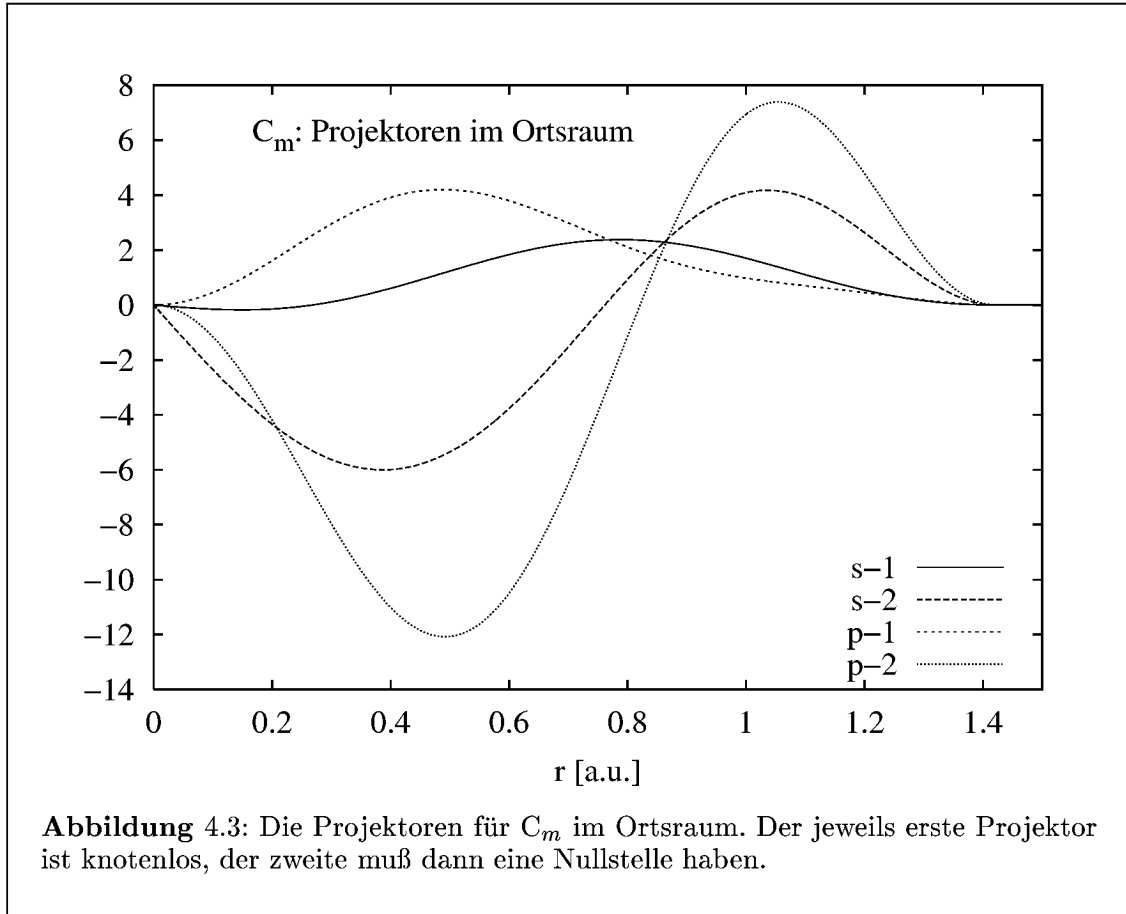
Tabelle 4.1 zeigt die für die Konstruktion entsprechend der geplanten Anwendung gewählten Rumpfradien sowie die nach dem Rappe-Kriterium zu erwartenden Abschneidewerte $G_{\max}$. Als Kriterium wurde eine Genauigkeit von 0.5 mRy pro Elektron gewählt.

Pot.	R_c [a.u.]	$G_{\max}$ [$\sqrt{Ry}$]
C_s	1.1	7.5
C_m	1.4	5.9
C_l	1.8	4.6

Tabelle 4.1: Die nach dem Rappe-Kriterium anzusetzenden Abschneidewerte für Potentiale mit verschiedenen Rumpfradien.



Die verschiedenen $2p$ -Wellenfunktionen sind in Abbildung 4.2 für die drei Potentiale sowie die echte Wellenfunktion wiedergegeben. Wie man sieht, ist in allen Fällen die Übereinstimmung auch noch für einen relativ großen Bereich innerhalb des jeweiligen 'Abschneideradius' außerordentlich gut. Daher kann man davon ausgehen, daß eine geringfügige Überlappung der Augmentationsgebiete verschiedener Atome nur einen minimalen Fehler verursachen sollte (vgl. [FKBK00]).



Das C_m -Potential

Für alle Drehimpulskomponenten des C_m -Potentials wurde der Abschneideradius $R_c = 1.4$ a.u. gewählt. Die ersten Referenzenergien für $2s$ - und $2p$ -Zustände sind die jeweiligen Eigenwerte, die zweite ist jeweils $\varepsilon_{s/p,2} = -0.7$ Ry. Der $3d$ -Zustand ist so ausgedehnt, daß er sich nicht zur Konstruktion des lokalen Potentials eignet. Daher wurde ein ungebundener Zustand zur Energie $\varepsilon_d = -0.4$ Ry benutzt. Eine partielle core-Ladungsdichte $\tilde{n}_c$ wurde nicht konstruiert; es wurde das Austausch-

Korrelations-Potential nach Vosko, Wilk und Nusair [VWN80] benutzt. Die AE-Rechnung wurde nicht-relativistisch durchgeführt.

Abbildung 4.3 zeigt für das C_m -Potential die Projektoren im Ortsraum nach dem Orthogonalisierungsverfahren. Der jeweils erste Projektor ist knotenlos, der zweite muß eine Nullstelle haben, da er orthogonal zur ersten, knotenlosen Wellenfunktion sein muß. Im Gegensatz zu den Projektoren, die mit PS-*partial waves* nach der BHS-Methode konstruiert worden waren, weisen diese Projektoren keine unnötig kleinen Längenskalen auf.

Einen ersten Test bildet die Bestimmung von Eigenwerten für andere Konfigurationen als diejenige, welche als Ausgangspunkt für die Konstruktion des Pseudopotentials diente. Die durch die AE-Rechnung generierten Werte lassen sich mit denen vergleichen, welche mittels der generierten Pseudopotentiale bestimmt werden. Ebenso kann man die Anregungsenergien zwischen dem Grundzustand und den anderen Konfigurationen vergleichen. Tabelle 4.3 zeigt die in der AE-Rechnung gewonnenen Werte und die Differenz der PAW-Werte zu diesen. Die Übereinstimmung ist vorzüglich, d.h. der Fehler der Energie ist jeweils deutlich kleiner als 0.1 mRy.

Konfig.	AE-Eigenwerte		Δ -Eigenwerte		AE- E_{exc}	ΔE_{exc}
	s	p	s	p		
$2s^2 2p^2$	-1.001732	-0.398371	0.000000	0.000000	0	0
$2s^1 2p^3$	-1.033870	-0.427952	-0.000003	-0.000029	0.604654	0.000045
$2s^{1.5} 2p^{2.5}$	-1.018285	-0.413623	0.000031	-0.000029	0.302007	0.000019

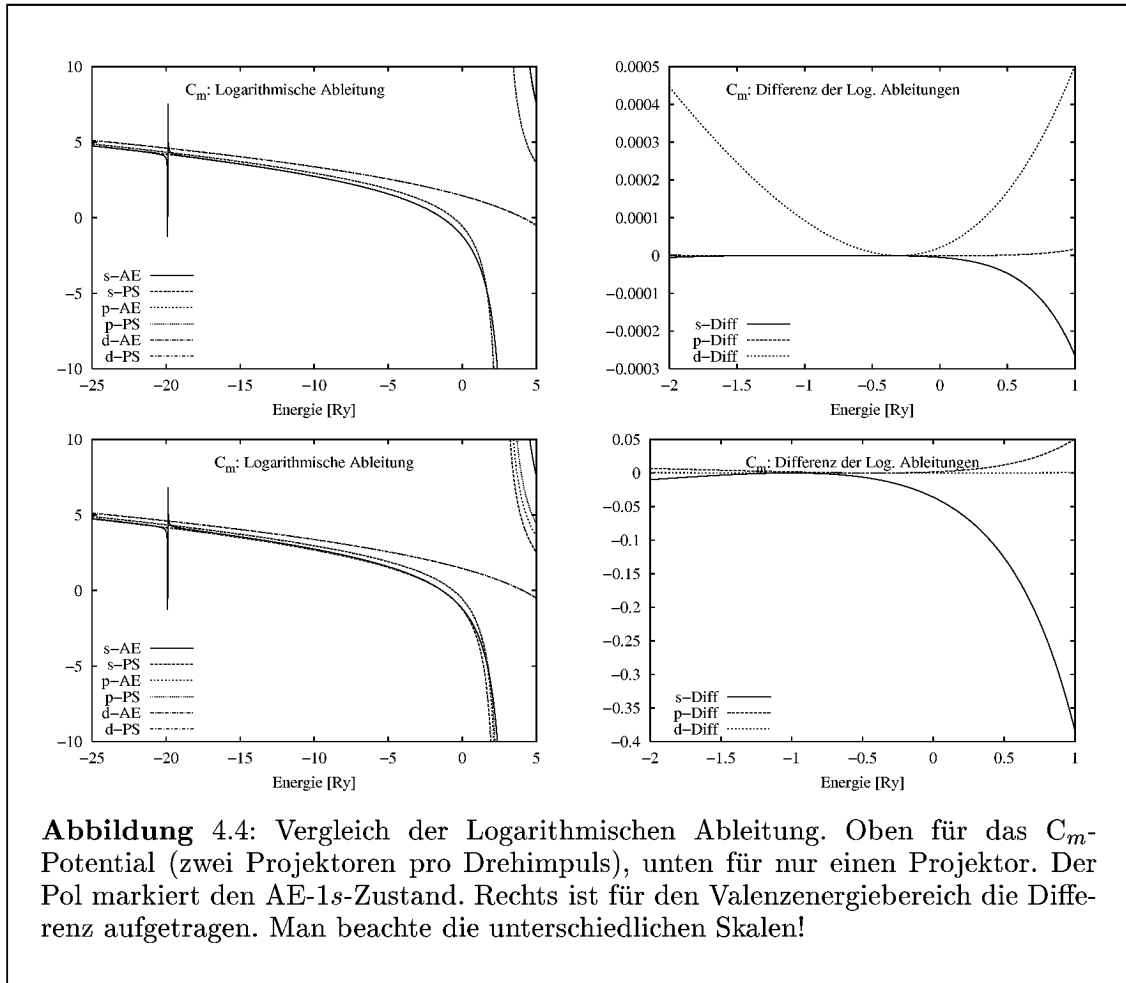
Tabelle 4.2: Vergleich der Eigenwerte für verschiedene Konfigurationen und der Anregungsenergien zwischen diesen. Das C_m -Potential ist mit *zwei* Projektoren pro Drehimpuls konstruiert. Angegeben sind die Werte für die AE-Berechnung und die absoluten Differenzen zur PAW-Rechnung. Alle Energien in Ry.

Um einen Eindruck zu gewinnen, welchen Einfluß der jeweils zweite Projektor auf diese Ergebnisse hat, wurde ein weiteres Potential konstruiert, welches nur jeweils einen Projektor für $2s$ - und $2p$ -Zustände hat bei ansonsten identischen Parametern. Diese Ergebnisse sind in Tabelle 4.3 zusammengestellt. Man sieht, daß die Abweichungen insbesondere für die stark angeregte Konfiguration $2s^1 2p^3$ sehr viel größer sind. Die Anregungsenergie wird um 0.5 mRy falsch wiedergegeben, die Abweichung des p -Eigenwertes ist sogar größer als 1 mRy. Dies ist ein Indiz dafür, daß die Transferierbarkeit des C_m -Potentials durch die jeweils zwei Projektoren erheblich besser ist.

Konfig.	AE-Eigenwerte		Δ -Eigenwerte		AE- E_{exc}	ΔE_{exc}
	s	p	s	p		
$2s^2 2p^2$	-1.001732	-0.398371	0.000000	0.000000	0	0
$2s^1 2p^3$	-1.033870	-0.427952	-0.000598	-0.001125	0.604654	0.000494
$2s^{1.5} 2p^{2.5}$	-1.018285	-0.413623	-0.000205	-0.000813	0.302007	0.000178

Tabelle 4.3: Vergleich der Eigenwerte für verschiedene Konfigurationen und der Anregungsenergien zwischen diesen. Das Potential ist mit *einem* Projektor pro Drehimpuls konstruiert. Angegeben sind die Werte für die AE-Berechnung und die absoluten Differenzen zur PAW-Rechnung. Alle Energien in Ry.

Einen weiteren Anhaltspunkt kann die Bestimmung der Logarithmischen Ableitungen geben. Diese zeigt Abbildung 4.4. Oben links sind AE- und PAW-Berechnung



dieser Größen über einen weiten Energiebereich dargestellt. Abgesehen von der Pol-

stelle der AE-Funktion, welche durch das Vorhandensein des tieferliegenden $1s$ -Zustandes bedingt ist (und im PAW-Fall natürlich fehlt), liegen die Kurven für alle Drehimpulse perfekt übereinander. Das Fehlen weiterer Pole in der PAW-Kurve zeigt, daß dieses Potential keine *ghost-states* aufweist.

Eine Abweichung kann man erst in einem vergrößerten Ausschnitt (rechts oben) erkennen. Hier zeigt sich, daß die Logarithmischen Ableitungen für s - und p -Funktionen über einen weiten Bereich des für Valenzzustände relevanten Energieabschnittes beinahe perfekt übereinstimmen. Für die d -Drehimpulskomponente hingegen, welche mit nur einer Referenzenergie konstruiert ist, ergibt sich – abgesehen von einem sehr kleinen Intervall um die Referenzenergie (-0.4 Ry) – eine deutlich größere Abweichung.

In der unteren Hälfte der Abbildung ist die analoge Berechnung für das Potential mit jeweils einem Projektor zu sehen. Hier kann man bereits in der Übersicht (links unten) deutliche Abweichungen erkennen, welche noch besser im Ausschnitt (rechts unten) zu erkennen sind. Die Abweichungen bei nur einem Projektor sind um Größenordnungen größer. (Man beachte die unterschiedlichen Maßstäbe!)

Bei normerhaltenden Pseudopotentialen zeigt die bisherige Erfahrung, daß die aus dem Rappe-Kriterium gewonnene Abschätzung des benötigten Abschneidewertes für die Ebene Wellen Basis sehr zuverlässig ist. Um die Konvergenz zu testen, wurde die Gesamtenergie für ein einzelnes Kohlenstoffatom in einer großen Superzelle für verschiedene Parameter $G_{\max}$ getestet.

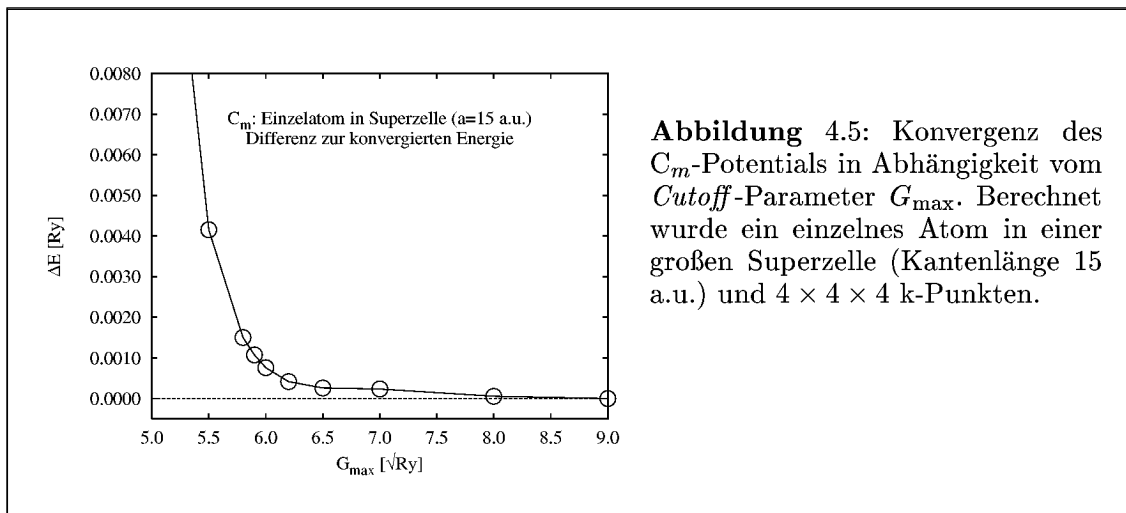


Abbildung 4.5: Konvergenz des C_m -Potentials in Abhängigkeit vom $Cutoff$ -Parameter $G_{\max}$. Berechnet wurde ein einzelnes Atom in einer großen Superzelle (Kantenlänge 15 a.u.) und $4 \times 4 \times 4$ k-Punkten.

Abbildung 4.5 zeigt die Konvergenz der Energie mit der Abschneideenergie. Mit dem aus dem Rappe-Kriterium gewonnenen Wert von $G_{\max} = 5.9\sqrt{\text{Ry}}$ wird das

Kriterium von 1 mRy knapp verfehlt. Für $G_{\max} = 6.0 \sqrt{\text{Ry}}$ jedoch wird dieser Fehler deutlich unterschritten. Dieser Unterschied könnte dadurch bedingt sein, daß man mittels des Rappe-Kriteriums keine Aussage erhält über die ungebundenen Zustände, die bei der Potentialkonstruktion verwendet wurden. Solange deren Parameter (Abschneideradius und Referenzenergie) aber denen des gebundenen Zustandes recht ähnlich sind, ist anzunehmen, daß die Abweichung nicht allzu groß werden kann.

Die Größe der Superzelle für diesen Test kann relativ klein angesetzt werden, da die Differenz der Gesamtenergie zum auskonvergierten Wert oberhalb einer gewissen Untergrenze kaum mit den Ausmaßen der Zelle variiert, wie Tabelle 4.4 zeigt.

Kantenlänge der Superzelle [a.u.]	$E_{\text{tot}}(G_{\max} = 6.0 \sqrt{\text{Ry}})$ $-E_{\text{tot}}(G_{\max} = 9.0 \sqrt{\text{Ry}})$ [mRy]
13.2	0.760
15.1	0.758
22.7	0.756

Tabelle 4.4: Abhängigkeit der Konvergenz von der Zellengröße am Beispiel des C_m -Potentials.

Wie oben erwähnt wurde, hängen die Pseudopotentiale von dem Wert G_{pot} ab, der bei seiner Konstruktion als künftiger Abschneiderwert $G_{\max}$ angenommen wird. Da dies eine gewisse Unbequemlichkeit darstellt, wurde für das einzelne Atom in der Superzelle untersucht, wie stark diese Abhängigkeit ist. Tabelle 4.5 vergleicht

$G_{\max} [\sqrt{\text{Ry}}]$	Energie [Ry] bei $G_{\text{pot}} = G_{\max}$	Energie [Ry] bei $G_{\text{pot}} = 6.0 \sqrt{\text{Ry}}$	Diff. [mRy]
5.0	-17.0535391	-17.0535821	0.04
6.0	-17.0699873	-17.0699873	0
7.0	-17.0705034	-17.0705130	0.009
8.0	-17.0706856	-17.0706903	0.004

Tabelle 4.5: Abhängigkeit der totalen Energie vom bei der Erzeugung des C_m -Potentials benutzten Cutoff G_{pot} .

die Gesamtenergien, welche man für verschiedene $G_{\max}$ erhält, wenn man einerseits ein Potential mit diesem Wert $G_{\text{pot}} = G_{\max}$ konstruiert, andererseits das soeben konstruierte Potential C_m mit $G_{\text{pot}} = 6.0 \sqrt{\text{Ry}}$ verwendet. Die Abweichungen liegen

in diesem Fall unter 0.04 mRy und damit deutlich unter der Genauigkeitsgrenze. Da jedoch nicht klar ist, inwieweit dieser Fall repräsentativ ist, empfiehlt es sich, Potentiale mit dem bei der Anwendung benutzten Abschneidewert zu konstruieren.

Das C_s -Potential

Für das C_s -Potential wurde der sehr kleine Abschneideradius $R_c = 1.1$ a.u. gewählt. Dies erfordert nach dem Rappe-Kriterium für die bei der Konstruktion benutzten gebundenen Zustände einen *Cutoff*-Parameter von etwa $G_{\max} = 7.5 \sqrt{\text{Ry}}$, wie man in Abbildung 4.6 rechts erkennt. Als Referenzenergien wurden für s - und p - wiederum $\tilde{\varepsilon}_{s,2} = \tilde{\varepsilon}_{p,2} = -0.7$ Ry neben den jeweiligen Eigenwerten verwendet. Für das lokale (d -) Potential wurde $\tilde{\varepsilon}_d = -0.3$ Ry gewählt. In diesem Falle wurde eine core-Ladungsdichte $\tilde{n}_c$ benutzt, deren Abschneideradius $r_{pcc} = 1.05$ a.u. betrug.

Konfig.	AE-Eigenwerte		Δ -Eigenwerte		AE- E_{exc}	ΔE_{exc}
	ε_s	ε_p	s	p		
$2s^2 2p^2$	-1.001732	-0.398371	-0.000055	-0.000042	0	0
$2s^1 2p^3$	-1.033870	-0.427952	-0.000117	-0.000043	0.604654	-0.000042
$2s^{1.5} 2p^{2.5}$	-1.018285	-0.413623	-0.000083	-0.000042	0.302007	-0.000013

Tabelle 4.6: Vergleich der Eigenwerte für verschiedene Konfigurationen und der Anregungsenergien zwischen diesen für das C_s -Potential. Angegeben sind die Werte für die AE-Berechnung und die absoluten Differenzen zur PAW-Rechnung. (Alle Energien in Ry.)

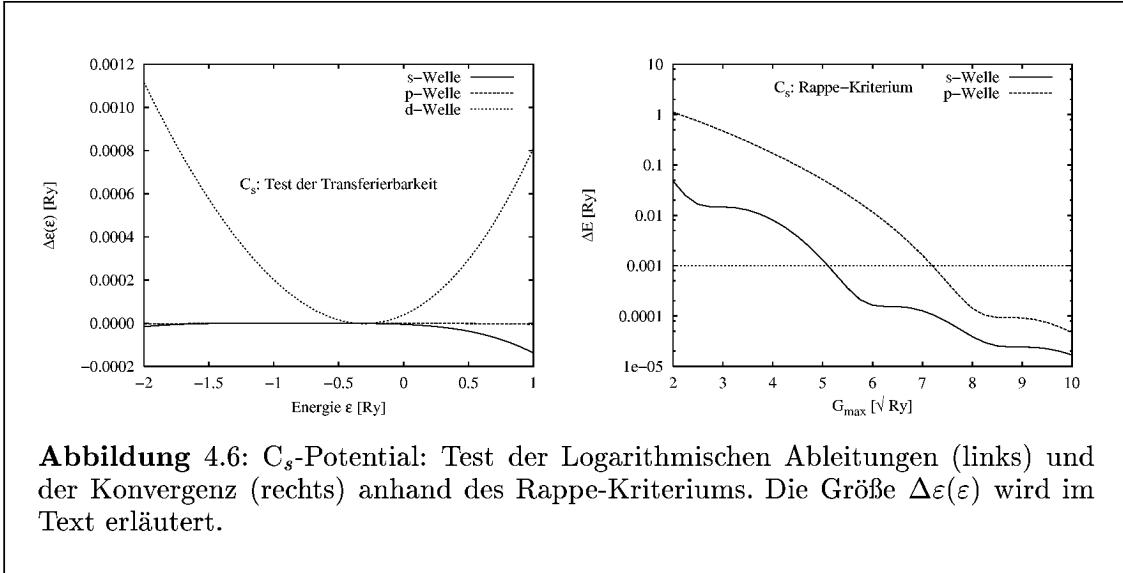
Links in Abbildung 4.6 ist der Fehler der Logarithmischen Ableitungen als Test für die Transferierbarkeit aufgetragen. Gezeigt wird hier nicht unmittelbar die Differenz zwischen den AE- und PS-Funktionen, sondern die Energiedifferenz $\Delta\varepsilon$, die durch

$$L_i^{PS}(\varepsilon + \Delta\varepsilon) = L_i^{AE}(\varepsilon) \quad (4.28)$$

implizit definiert ist (vgl. [ERBS]). Die Größe $\Delta\varepsilon(\varepsilon)$ gibt an, wie weit die Energie, an der die PS-Logarithmische Ableitung denselben Wert annimmt wie die AE-Funktion an der Stelle ε , von dieser entfernt ist. In erster Näherung gilt

$$\Delta\varepsilon(\varepsilon) = \frac{L_i^{AE}(\varepsilon) - L_i^{PS}(\varepsilon)}{\frac{\partial}{\partial \varepsilon} L_i^{PS}(\varepsilon)} \quad (4.29)$$

Die schlichte Differenz der Ableitungen für denselben Energiewert ε liefert – praktisch gesprochen – in der Nähe der Polstellen kein brauchbares Maß der Abweichung.



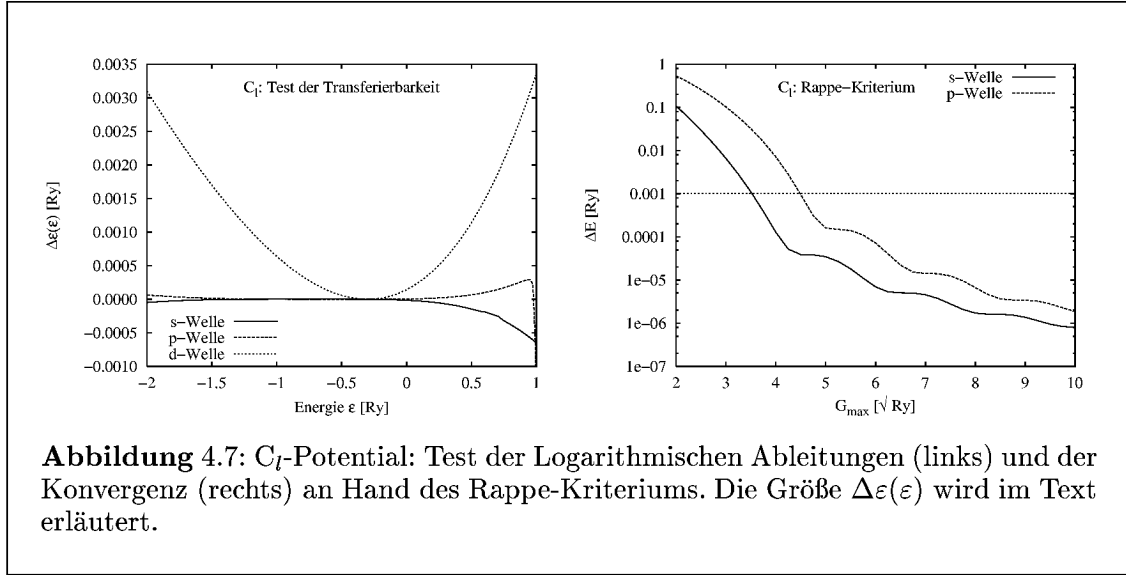
Wie beim C_m -Potential ist auch hier die Übereinstimmung hervorragend für die s - und p -Komponente, für die lokale d -Komponente schlechter, aber immer noch sehr gut.

Das C_l -Potential

Das C_l -Potential zeichnet sich durch einen relativ großen Rumpfradius $R_c = 1.8$ a.u. aus, welcher deutlich größer ist als das äußere Maximum der Wellenfunktionen. Daher erwartet man nach dem Rappe-Kriterium einen Abschneidewert $G_{\max} = 4.5 \sqrt{\text{Ry}}$. Die Referenzenergien sind $\epsilon_{s,2} = -0.55$ Ry, $\epsilon_{p,2} = -0.75$ Ry, $\epsilon_d = -0.3$ Ry. Auf Grund der größeren Ausdehnung wurden diesmal 300 statt 200 Punkte für das radiale Gitter benutzt. Die übrigen Parameter sind identisch mit denen des C_m -Potentials. Das Rappe-Kriterium und der Transferierbarkeitstest sind in Abbildung 4.7 zu sehen. Letzterer vermittelt einen ähnlichen Eindruck wie bei den beiden vorangegangenen Potentialen.

Die quantitative Abweichung der Logarithmischen Ableitung der d -Komponente von der beim C_m -Potential gezeigten rührt daher, daß diese an unterschiedlichen Radien ausgewertet wurden.

Den Vergleich von AE- und PAW-Berechnung des Grundzustandsatoms sowie zweier angeregter Konfigurationen zeigt Tabelle 4.7. Die Wiedergabe der Eigenwerte und der Anregungsenergien ist sehr gut, wenn auch weniger exakt als bei den Potentialen mit kleineren Radien; dies entspricht freilich auch der Erwartung. Die Fehler von weniger als 0.1 mRy sind jedoch immer noch vernachlässigbar.



Konfig.	AE-Eigenwerte		Δ -Eigenwerte		$AE-E_{exc}$	ΔE_{exc}
	ϵ_s	ϵ_p	s	p		
$2s^22p^2$	-1.001732	-0.398371	0.000000	0.000000	0	0
$2s^12p^3$	-1.033870	-0.427952	0.000030	-0.000105	0.604654	0.000121
$2s^{1.5}2p^{2.5}$	-1.018285	-0.413623	0.000059	-0.000089	0.302007	0.000043

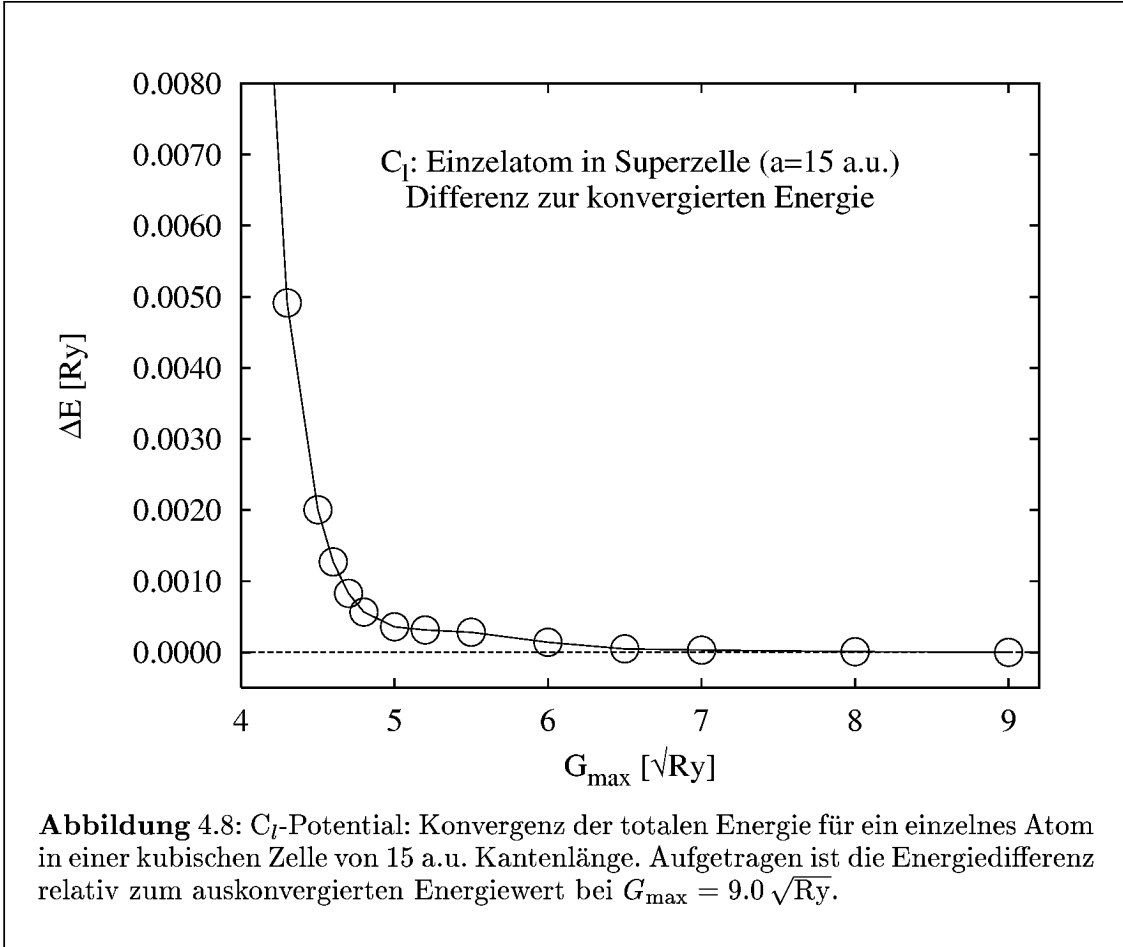
Tabelle 4.7: Vergleich der Eigenwerte für verschiedene Konfigurationen und der Anregungsenergien zwischen diesen für das C_l -Potential. Angegeben sind die Werte für die AE-Berechnung und die absoluten Differenzen zur PAW-Rechnung. (Alle Energien in Ry.)

Beim Konvergenztest für ein einzelnes Atom in einer großen Superzelle, dessen Ergebnisse in Abbildung 4.8 aufgetragen sind, zeigt sich die schon angesprochene Tendenz zu einer etwas schlechteren Konvergenz mit dem *Cutoff* $G_{\max}$ als man dies nach dem Rappe-Kriterium erwartet. In diesem Falle ist die Konvergenz unterhalb die 1 mRy Schwelle erst für $G_{\max} = 4.7 \sqrt{\text{Ry}}$ statt für $G_{\max} = 4.5 \sqrt{\text{Ry}}$ gegeben.

Das C_x -Potential

Um auszuloten, inwieweit die gute Übereinstimmung von AE- und PS-*partial wave* sich nutzen läßt, um die Augmentationsgebiete über die durch deren Disjunktheit gesetzte Grenze hinaus auszudehnen, wurde ein Potential C_x erstellt, dessen Parameter sich von denen des C_l -Potentials allein durch einen größeren Abschneideradius von $R_c = 2.2$ a.u. unterscheiden.

Das Rappe-Kriterium läßt auf eine Abschneideenergie $G_{\max}$ von etwa $3.8 \sqrt{\text{Ry}}$ hof-



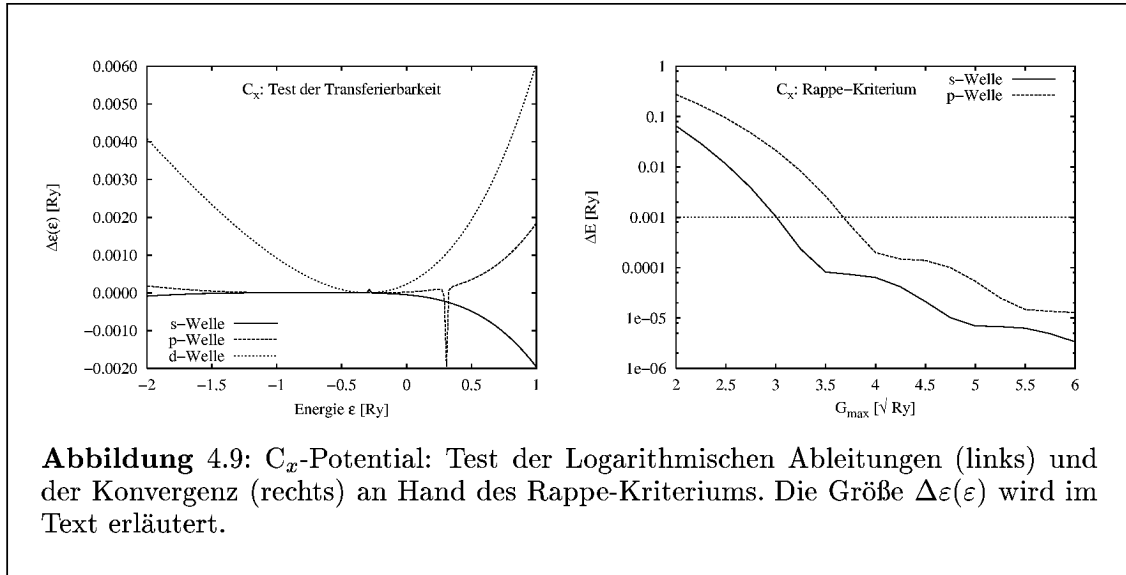
fen (siehe Abbildung 4.9 rechts). Der Test der Logarithmischen Ableitungen (siehe Abbildung 4.9 links) zeigt eine ähnliche Qualität wie für C_l .

Den Vergleich von AE- und PAW-Atom-Rechnung zeigt Tabelle 4.8. Auch hier vermittelt sich ein guter Eindruck.

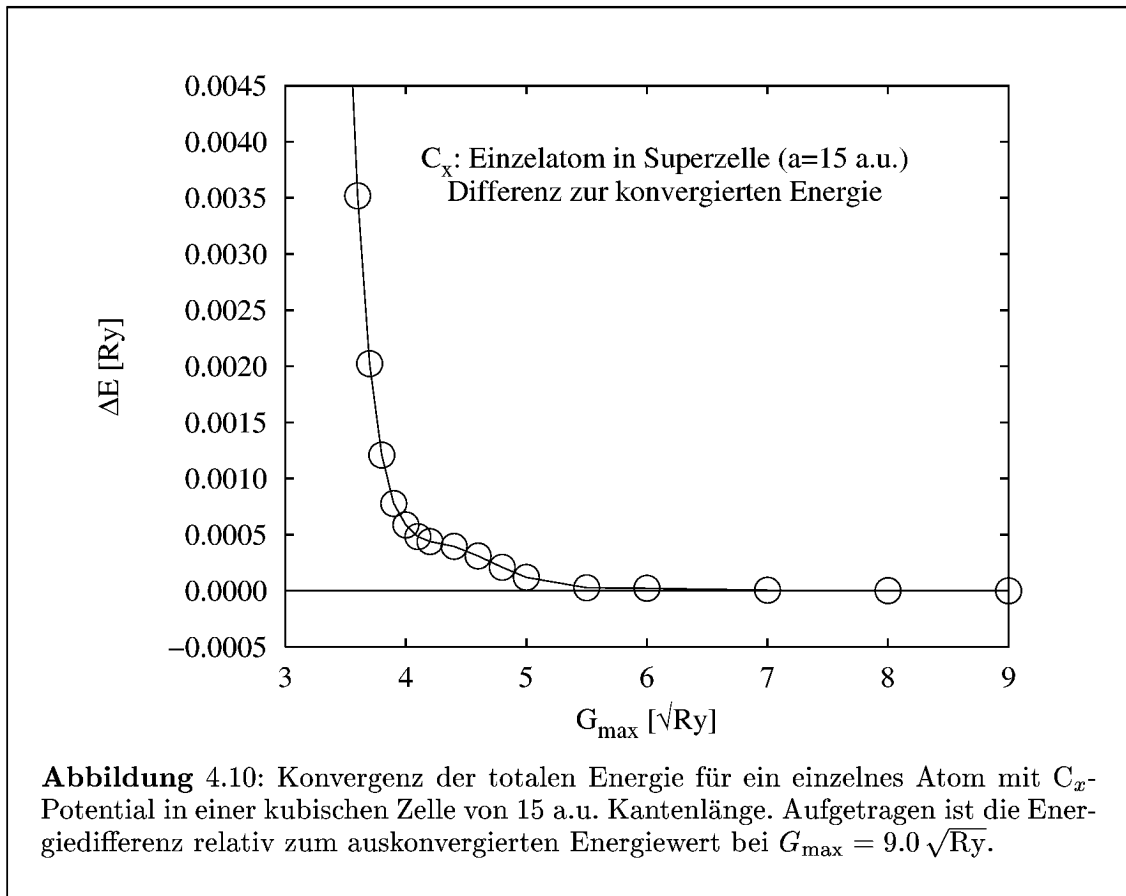
Konfig.	AE-Eigenwerte		Δ -Eigenwerte		AE- E_{exc}	ΔE_{exc}
	ε_s	ε_p	s	p		
$2s^2 2p^2$	-1.001732	-0.398371	0.000000	0.000000	0	0
$2s^1 2p^3$	-1.033870	-0.427952	0.000040	-0.000212	0.604654	0.000206
$2s^{1.5} 2p^{2.5}$	-1.018285	-0.413623	0.000068	-0.000176	0.302007	0.000069

Tabelle 4.8: Vergleich der Eigenwerte für verschiedene Konfigurationen und der Anregungsenergien zwischen diesen für das C_x -Potential. Angegeben sind die Werte für die AE-Berechnung und die absoluten Differenzen zur PAW-Rechnung. (Alle Energien in Ry.)

Beim Konvergenztest des C_x -Potentials in einer großen Superzelle, dessen Ergebnisse



in Abbildung 4.10 aufgetragen sind, wird die 1 mRy Schwelle zum auskonvergierten Fall für einen *Cutoff*-Wert von $G_{\max} = 3.9 \sqrt{\text{Ry}}$ unterschritten.



4.6 Stickstoff

Auch für Stickstoff wurden mehrere Pseudopotentiale konstruiert, die sich vor allem in den Abschneideradien unterscheiden. Diese wurden gewählt im Hinblick auf die Anwendbarkeit für ein Stickstoffmolekül N_2 einerseits und für den Halbleiterkristall AlN andererseits. Diese Potentiale werden – analog zum Kohlenstoff – als N_s und N_l bezeichnet. Ausgangspunkt war – wie bei den übrigen Pseudopotentialen – der Grundzustand des Atoms. Die Berechnung der kinetischen Energie erfolgte nicht-relativistisch. Benutzt wurde das RRKJ-Verfahren. Das Austausch-Korrelations-Potential wurde wiederum nach [VWN80] berechnet. Auch hier wurden – wie beim Kohlenstoff – $2s$ - und $2p$ -Zustände nichtlokal mit jeweils zwei Projektoren behandelt; das lokale Potential wurde von einem ungebundenen d -Zustand abgeleitet.

Das N_l -Potential

Für die Anwendung im Halbleiterkristall mit relativ großen interatomaren Abständen wurde das N_l -Potential konstruiert. Der Abschneideradius für alle Drehimpulskomponenten betrug $R_c = 1.7$ a.u. Als Referenzenergien wurden bei s - und p -Funktionen der jeweilige Eigenwert und $\tilde{\epsilon}_{s/p,2} = -0.95$ Ry verwendet, beim lokalen Potential $\tilde{\epsilon}_d = -0.4$ Ry. Mit diesen Parametern erhält man eine gute Transferierbarkeit (vgl. Tabelle 4.9), und auch die Logarithmischen Ableitungen (Abbildung 4.11 links) sind

Konfig.	AE-Eigenwerte		Δ -Eigenwerte		AE- E_{exc}	ΔE_{exc}
	ϵ_s	ϵ_p	s	p		
$2s^2 2p^3$	-1.352301	-0.532593	0.000000	0.000000	0	0
$2s^{1.5} 2p^{3.5}$	-1.368748	-0.547382	0.000191	-0.000152	0.410269	0.000099
$2s^1 2p^4$	-1.384551	-0.561540	0.000168	-0.000170	0.821364	0.000286

Tabelle 4.9: Vergleich der Eigenwerte für verschiedene Konfigurationen und der Anregungsenergien zwischen diesen für das N_l -Potential. Angegeben sind die Werte für die AE-Berechnung und die absoluten Differenzen zur PAW-Rechnung. (Alle Energien in Ry.)

einwandfrei. (Die Zacken in den Funktionen sind numerische Artefakte; sie indizieren die Lage von Polstellen.)

Nebenbei sei bemerkt, daß natürlich für alle Atome die Logarithmischen Ableitungen über einen weiten Energiebereich mit einem engen Gitter berechnet wurden, um eventuell vorhandene *ghost-states* zu identifizieren. Dieser Test war hier wie in allen übrigen Fällen negativ.

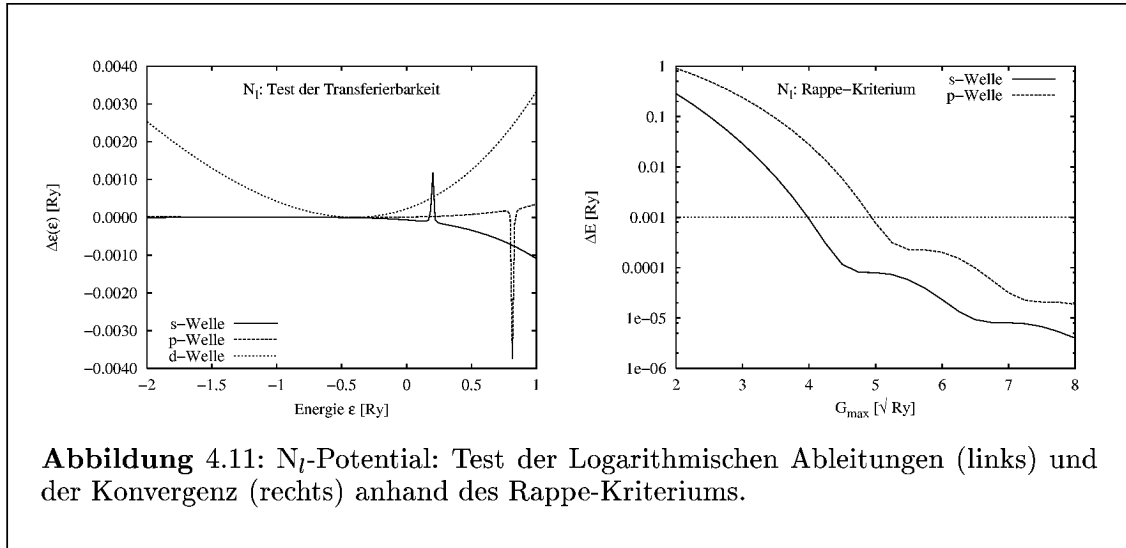


Abbildung 4.11: N_I -Potential: Test der Logarithmischen Ableitungen (links) und der Konvergenz (rechts) anhand des Rappe-Kriteriums.

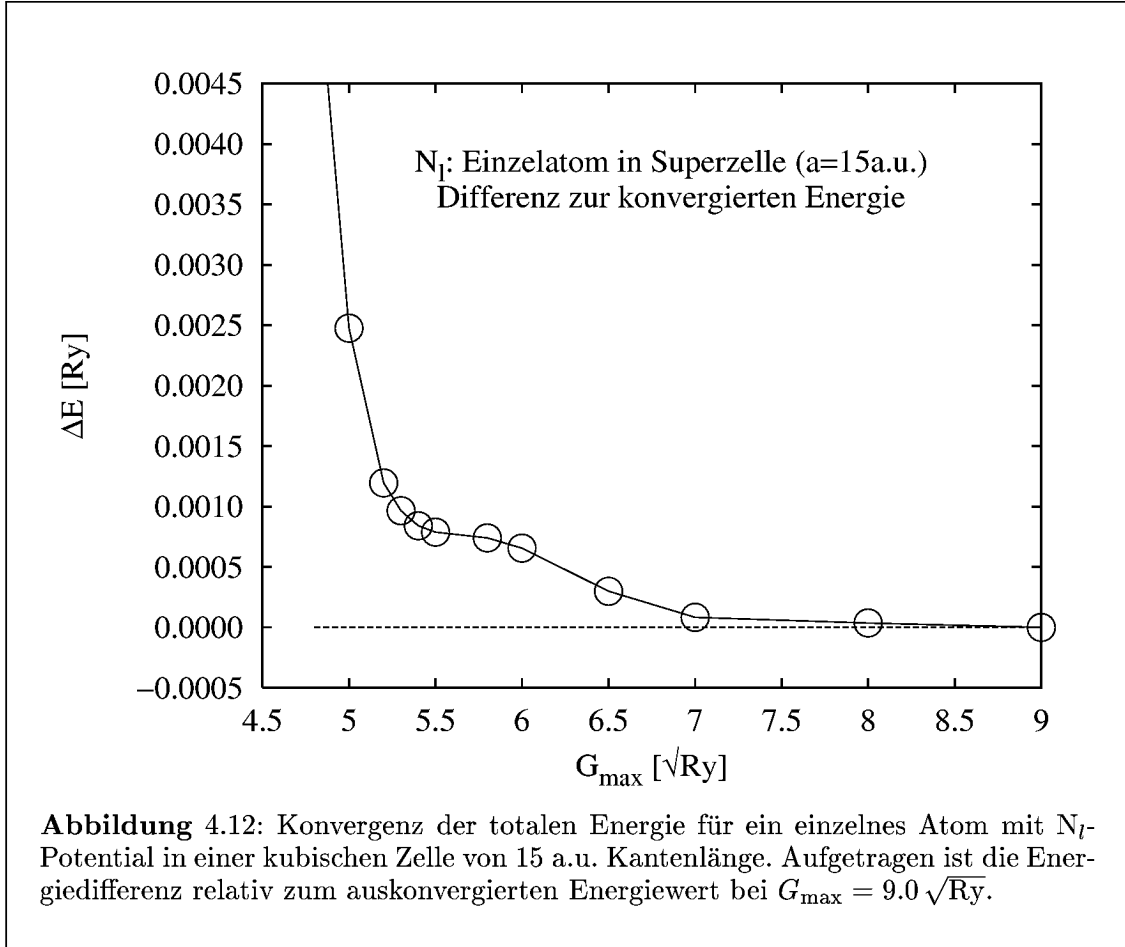
Die Abschätzung der benötigten Abschneideenergie nach dem Rappe-Kriterium ist in Abbildung 4.11 rechts zu sehen. Demnach kann man eine Konvergenz auf 1 mRy bei etwa $G_{\max} = 5.2 \sqrt{\text{Ry}}$ erwarten.

Die tatsächliche Konvergenz bezüglich der Darstellung durch Ebene Wellen ist in Abbildung 4.12 wiedergegeben. Die 1 mRy Marke wird zwischen $5.2 \sqrt{\text{Ry}}$ und $5.3 \sqrt{\text{Ry}}$ unterschritten, etwa in Übereinstimmung mit der obigen Abschätzung.

Das N_s -Potential

Das N_s -Potential ist speziell für die Berechnung kleiner Moleküle, insbesondere des Dimers N_2 konzipiert. Dies erfordert den sehr kleinen Abschneideradius $R_c = 1.0$ a.u. Die übrigen Parameter sind identisch mit denen des N_I -Potentials. Der kleine Radius erfordert eine große Anzahl Ebener Wellen bis zu einer Obergrenze von $G_{\max} = 8.3 \sqrt{\text{Ry}}$ nach dem Rappe-Kriterium (siehe Abbildung 4.13 rechts). Die Wiedergabe atomarer Eigenwerte und Anregungsenergien gelingt sehr gut, wie man in Tabelle 4.10 erkennt.

Allerdings werden hier – wie auch schon beim C_s -Potential die Eigenwerte im Grundzustand, welcher die Grundlage der Potentialgenerierung bildete, nicht exakt reproduziert, was eigentlich durch die Konstruktion garantiert wird (vgl. Abschnitt 4.2). Die Ursache könnte darin liegen, daß die Matrix D_{kin} , welche eigentlich symmetrisch sein sollte, durch numerische Fehler (insbesondere beim Bilden der Ableitung) nicht exakt symmetrisch ist und die Matrix mit ihrer Transponierten gemittelt wird. Der Fehler (0.1 mRy) ist allerdings immer noch sehr klein. Da die Anregungsenergi-



Konfig.	AE-Eigenwerte		Δ -Eigenwerte		AE- E_{exc}	ΔE_{exc}
	ε_s	ε_p	s	p		
$2s^2 2p^3$	-1.352301	-0.532593	-0.000118	0.000000	0	0
$2s^{1.5} 2p^{3.5}$	-1.368748	-0.547382	-0.000144	-0.000007	0.410269	-0.000063
$2s^1 2p^4$	-1.384551	-0.561540	-0.000184	-0.000011	0.821364	-0.000140

Tabelle 4.10: Vergleich der Eigenwerte für verschiedene Konfigurationen und der Anregungsenergien zwischen diesen für das N_s -Potential. Angegeben sind die Werte für die AE-Berechnung und die absoluten Differenzen zur PAW-Rechnung. (Alle Energien in Ry.)

en hinreichend genau berechnet werden, kann man davon ausgehen, daß dies keine Beeinträchtigung des Potentials darstellt.

Die Logarithmischen Ableitungen zeigen wiederum, daß zwei Projektoren zu deutlich geringer Abweichung über einen weiten Energiebereich führen.

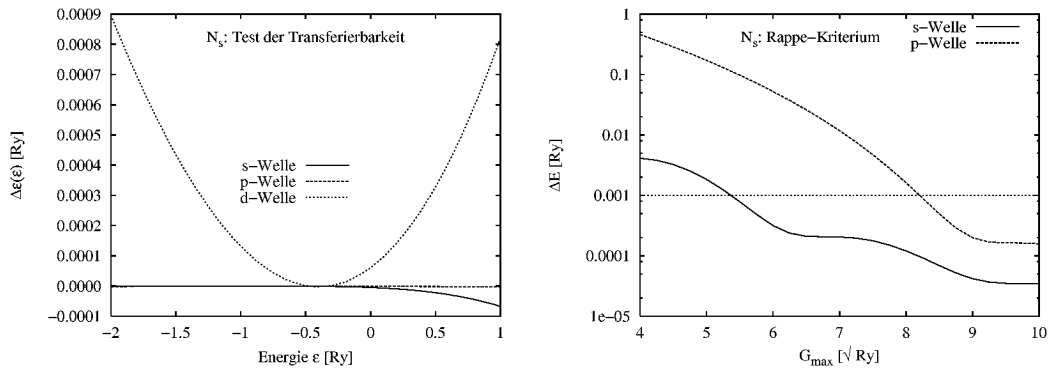


Abbildung 4.13: N_s-Potential: Test der Logarithmischen Ableitungen (links) und der Konvergenz (rechts) an Hand des Rappe-Kriteriums.

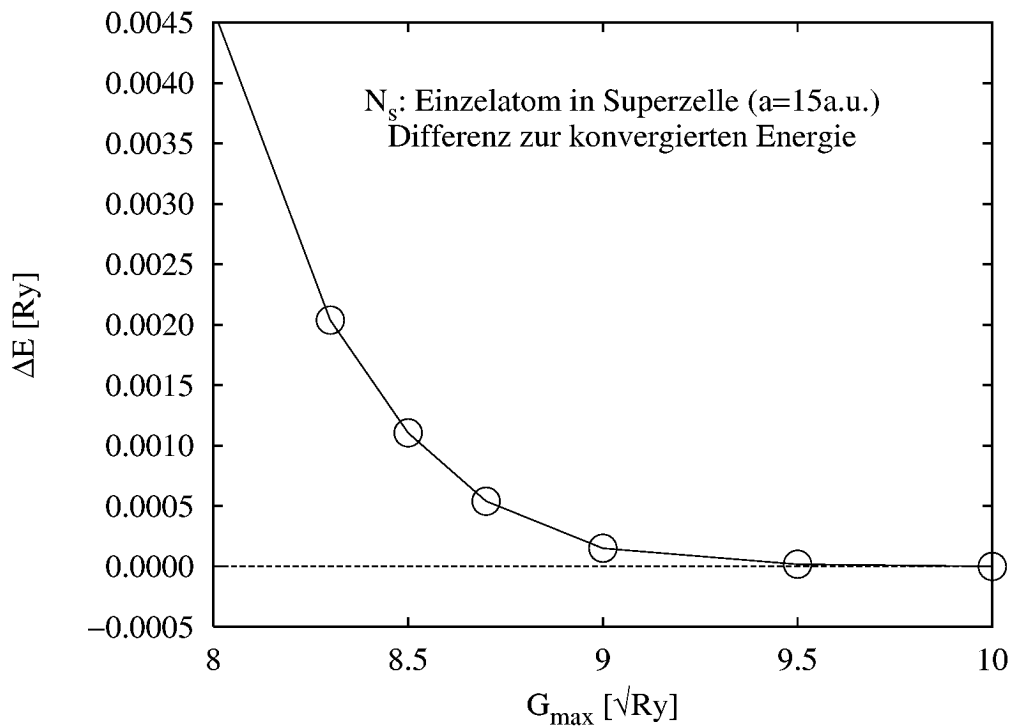
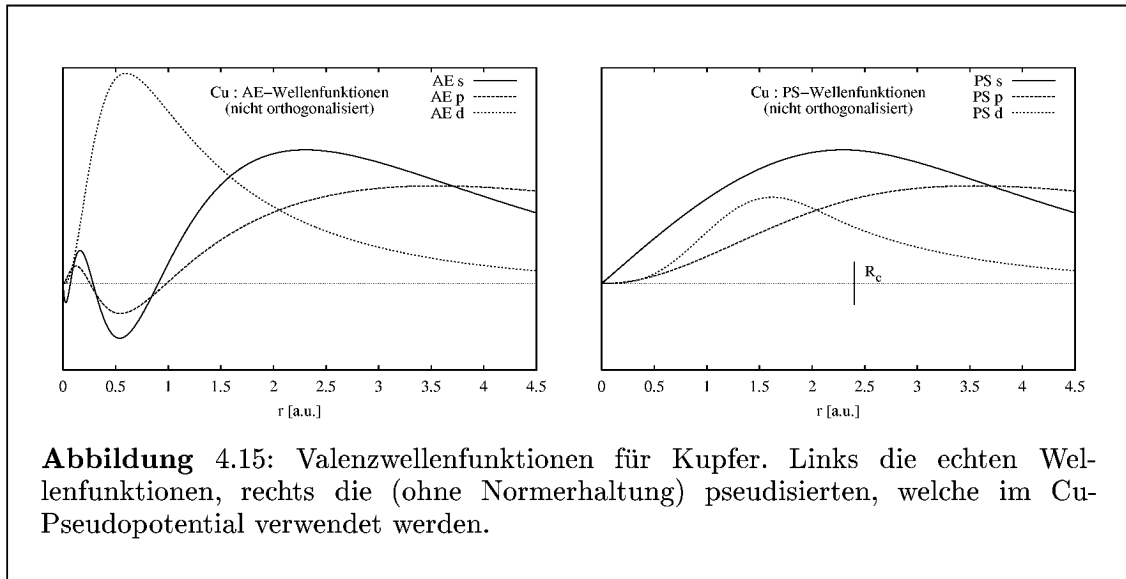


Abbildung 4.14: Konvergenz der totalen Energie für ein einzelnes Atom mit N_s-Potential in einer kubischen Zelle von 15 a.u. Kantenlänge. Aufgetragen ist die Energiedifferenz relativ zum auskonvergierten Energiewert bei $G_{\max} = 9.0 \sqrt{\text{Ry}}$.

4.7 Kupfer

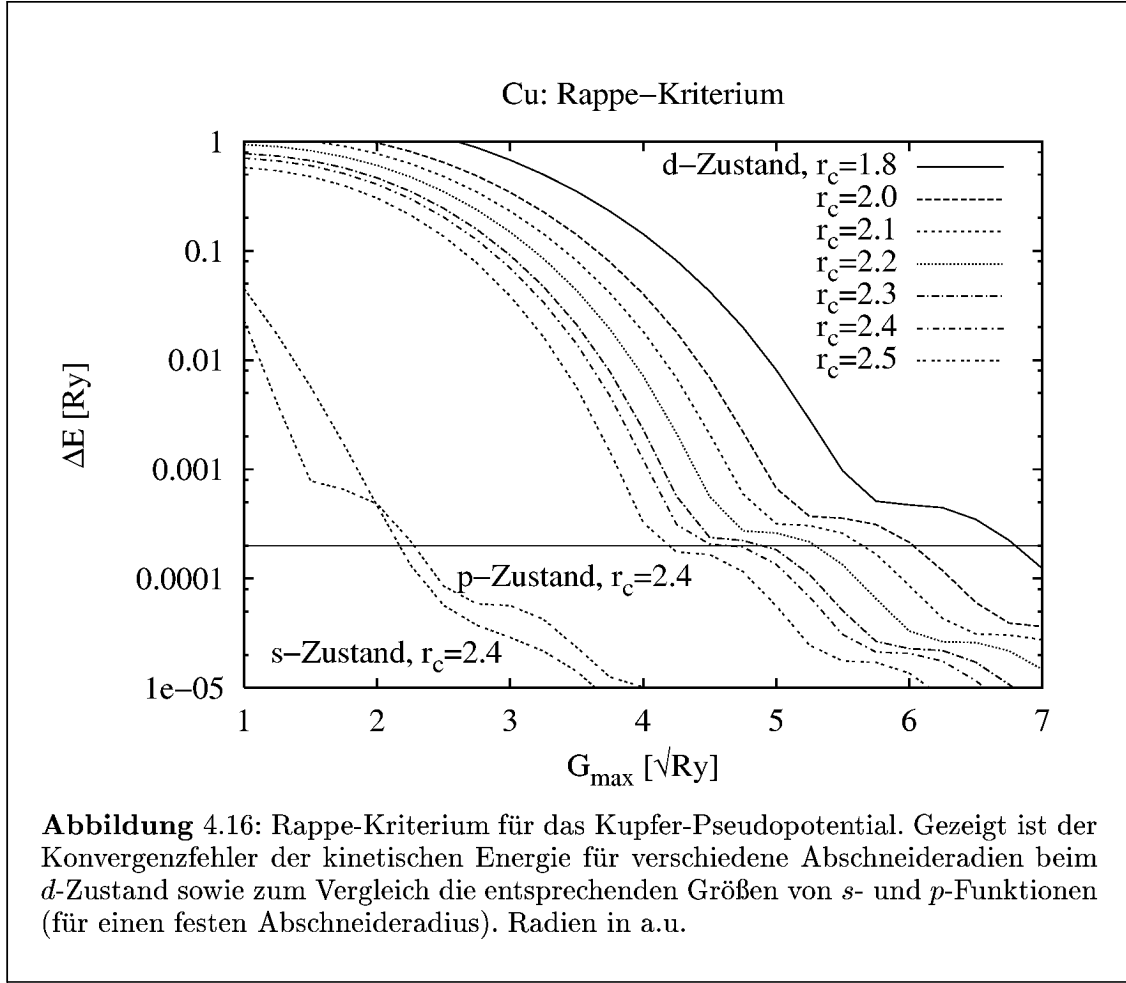
Als Beispiel für ein Übergangsmetall wurde ein Pseudopotential für Kupfer generiert. Dessen $3d$ -Schale ist zwar voll besetzt, liegt jedoch energetisch so nahe an den $4s$ -Zuständen und die Wellenfunktionen haben einen so starken Überlapp, daß eine Beschreibung als Valenzelektronen unerläßlich ist. Gleichzeitig sind die d -Wellenfunktionen aber so stark lokalisiert (vgl. Abbildung 4.15), daß der Abschneideradius für ein konventionelles Pseudopotential extrem klein sein müßte und daher eine sehr große Basis erzwänge.



PAW-Potentiale jedoch sind nicht daran gebunden. Das hier erstellte Pseudopotential ist gedacht für die Berechnung von Volumen- oder Oberflächensuperzellen von metallischem Kupfer. Der Gitterkonstante von fcc-Kupfer beträgt $3.61 \text{ \AA}$ [AM76]. Daraus ergibt sich ein Abstand zweier Kupferatome von 4.83 a.u. . Der Abschneideradius wurde daher mit $R_c = 2.4 \text{ a.u.}$ angesetzt. Die dann nach dem RRKJ-Verfahren erzeugten Pseudowellenfunktionen werden rechts in Abbildung 4.15 gezeigt. Es ist offensichtlich, daß vergleichbar glatte d -Wellenfunktionen unter Einhaltung der Normerhaltung nicht erzielbar sind.

Die nach dem Rappe-Kriterium zu erwartenden *Cutoff*-Energien sind für diesen und einige weitere Radien in Abbildung 4.16 wiedergegeben.

Beim gewählten Abschneideradius erwartet man demnach, daß man mit einer *Cutoff*-Energie von $G_{\max} = 4.5 \sqrt{R_y}$ eine Genauigkeit von 2 mRy pro Atom erzielen kann.



Das Potential wurde mit jeweils zwei Projektoren für s -, p - und d -Zustände erzeugt, und als lokales Potential diente ein ungebundener f -Zustand der Energie $\varepsilon_f = -0.3$ Ry. Die Referenzenergien für die ersten Projektoren waren jeweils die Valenzeigenenergien und als zweite Energien wurden $\varepsilon_{s,2} = -0.1$ Ry, $\varepsilon_{p,2} = -0.4$ Ry und $\varepsilon_{d,2} = -0.3$ Ry gewählt (s.u.). Es wurde eine partielle *core*-Ladungsdichte konstruiert mit einem Abscheideradius von 2.3 a.u.

Der Test der Eigenwerte und Anregungsenergien für verschiedene Zustände (in Tabelle 4.11) liefert eine Überraschung: Die PAW-Rechnung für ein ($s \rightarrow p$) angeregtes Atom $4s^{0.5}4p^{0.5}3d^{10}$ liefert hervorragende Ergebnisse, wohingegen in den Fällen, in denen die Besetzung des d -Zustandes geändert wird, die Abweichungen recht groß sind.

Freilich handelt es sich auch um z.T. recht starke Anregungen; bei kleinen Besetzungsänderungen ($< 0.3e$) der d -Zustände ist der Fehler der Anregungsenergien

Konfig.	AE-Eigenwerte			Δ -Eigenwerte			E_{exc} (AE)	ΔE_{exc}
	ε_s	ε_p	ε_d	s	p	d		
$4s^1 3d^{10}$	-344.11	-58.07	-404.54	0.00	0.01	0.00	0	0
$4s^{0.5} 4p^{0.5} 3d^{10}$	-400.85	-102.66	-485.03	-0.35	-0.12	1.00	145.59	-0.05
$4s^1 4p^{0.1} 3d^{9.9}$	-364.51	-72.60	-451.12	0.32	0.13	-1.11	36.21	-0.06
$4s^1 4p^{0.2} 3d^{9.8}$	-384.36	-85.14	-498.27	0.59	0.25	-2.61	75.78	-0.26
$4s^1 4p^{0.3} 3d^{9.7}$	-403.60	-96.69	-545.86	0.80	0.36	-4.50	118.89	-0.64
$4s^1 4p^{0.4} 3d^{9.6}$	-422.31	-107.60	-593.90	0.96	0.46	-6.83	165.65	-1.25
$4s^1 4p^{0.5} 3d^{9.5}$	-440.57	-118.04	-642.45	1.05	0.56	-9.60	216.18	-2.12
$4s^{1.1} 3d^{9.9}$	-352.41	-61.69	-434.19	0.40	0.17	-1.24	7.10	-0.08
$4s^{1.2} 3d^{9.8}$	-360.89	-65.39	-464.94	0.74	0.34	-2.77	16.39	-0.33
$4s^{1.3} 3d^{9.7}$	-369.54	-69.17	-496.79	1.02	0.50	-4.59	27.94	-0.79
$4s^{1.4} 3d^{9.6}$	-378.34	-73.02	-529.70	1.23	0.66	-6.70	41.87	-1.46
$4s^{1.5} 3d^{9.5}$	-387.28	-76.93	-563.66	1.39	0.81	-9.12	58.25	-2.38
$4s^{2.0} 3d^{9.0}$	-433.66	-97.31	-748.37	1.10	1.71	-25.04	180.11	-11.38
$4s^{1.10} 4p^{0.10} 3d^{9.8}$	-372.81	-75.99	-481.77	0.66	0.28	-2.69	45.99	-0.30
$4s^{1.15} 4p^{0.15} 3d^{9.7}$	-387.02	-84.07	-521.77	0.90	0.41	-4.56	73.16	-0.71
$4s^{1.20} 4p^{0.20} 3d^{9.6}$	-401.12	-91.87	-562.62	1.08	0.53	-6.79	103.27	-1.35

Tabelle 4.11: Vergleich der Eigenwerte für verschiedene Konfigurationen und der Anregungsenergien zwischen diesen für das Cu-Potential. Angegeben sind die Werte für die AE-Berechnung und die absoluten Differenzen zur PAW-Rechnung. (Alle Energien in mRy.)

absolut kleiner als 1 mRy und relativ kleiner als 3%.

Betrachtet man die Eigenwerte der angeregten Atome (Abbildung 4.17), so erkennt man, daß sich in den problematischen Fällen der $3d$ -Eigenwert stark nach unten verschiebt.

Daher wurde versuchsweise die zweite Referenzenergie in diese Region (auf $\varepsilon = -0.8$ Ry) verschoben, um eine gute Beschreibung auch für diese Energiebereiche zu erzwingen. Dieses Potential wird mit Cu_B bezeichnet. Abbildung 4.18 zeigt den Effekt auf die Logarithmischen Ableitungen.

Diese stimmen in beiden Fällen so genau mit der AE-Funktion überein, daß man nur in der Auftragung der Differenzen Unterschiede erkennen kann. Die beiden mittleren Abbildungen zeigen die absoluten Differenzen, die unteren die Abweichung $\Delta\varepsilon(\varepsilon)$ der Energie nach Gleichung 4.28. Man sieht, daß sich der Bereich guter Übereinstimmung

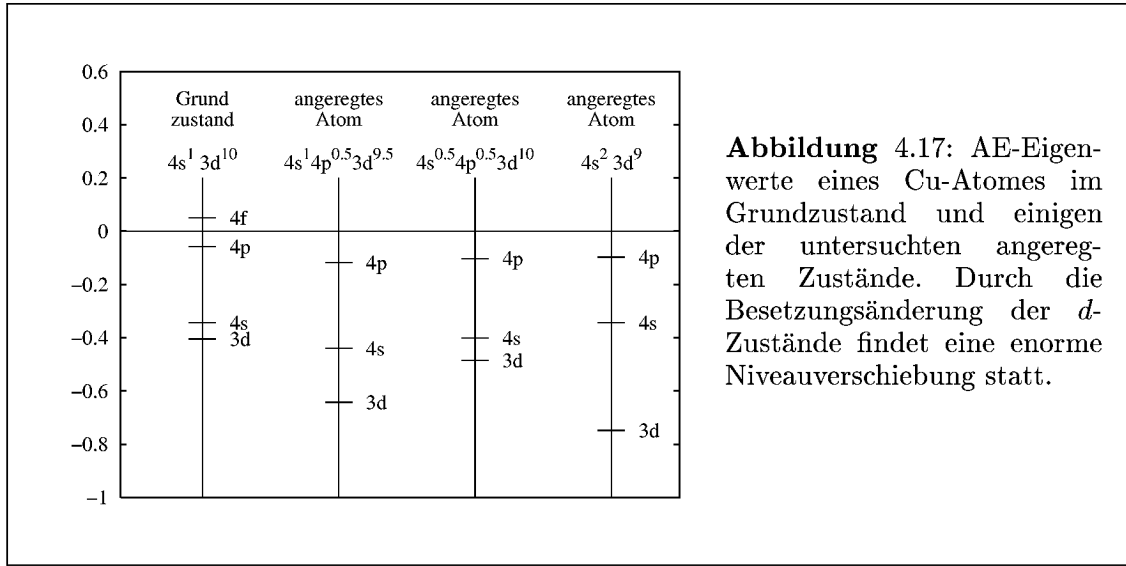


Abbildung 4.17: AE-Eigenwerte eines Cu-Atomes im Grundzustand und einigen der untersuchten angeregten Zustände. Durch die Besetzungsänderung der d -Zustände findet eine enorme Niveaushiftung statt.

für die d -Funktionen deutlich ausweitet. Gleichzeitig läßt sich aber ablesen, daß in dem Bereich, in dem die d -Eigenwerte der angeregten Atome liegen, die Beschreibung auch durch das ursprüngliche Potential Cu_A noch gut sein sollte. Rechts oben erkennt man im vergrößerten Ausschnitt, wie die PS-Logarithmischen Ableitungen bei den Referenzenergien an die AE-Funktionen „angeheftet“ werden.

Die Benutzung einer niedrigeren Referenzenergie beim Cu_B -Potential führt zu einer Verbesserung der Energiewerte für angeregte Atome, wie der Vergleich der Tabellen 4.11 und 4.12 zeigt. Jedoch sind die Fehler immer noch recht ausgeprägt.

Konfig.	AE-Eigenwerte			Δ -Eigenwerte			AE- E_{exc}	ΔE_{exc}
	ε_s	ε_p	ε_d	s	p	d		
$4s^1 3d^{10}$	-344.11	-58.07	-404.54	0.00	0.01	0.00	0	0
$4s^1 4p^{0.5} 3d^{9.5}$	-440.57	-118.04	-642.45	1.04	0.41	7.93	216.18	1.74
$4s^{0.5} 4p^{0.5} 3d^{10}$	-400.85	-102.66	-485.03	0.33	0.11	0.97	145.59	0.05
$4s^2 3d^9$	-433.66	-97.31	-748.37	1.16	1.26	21.62	180.11	9.88

Tabelle 4.12: Vergleich der Eigenwerte für verschiedene Konfigurationen und der Anregungsenergien zwischen diesen für das Cu-Potential. Angegeben sind die Werte für die AE-Berechnung und die absoluten Differenzen zur PAW-Rechnung. (Alle Energien in mRy.)

Bei der Berechnung fiel auf, daß sich die *core*-Eigenwerte der angeregten Atome gegenüber denen des Grundzustandsatoms z.T. relativ stark verändern (d.h. um bis zu 0.5 Ry), insbesondere, wenn die d -Besetzung verändert wird. Daher entstand

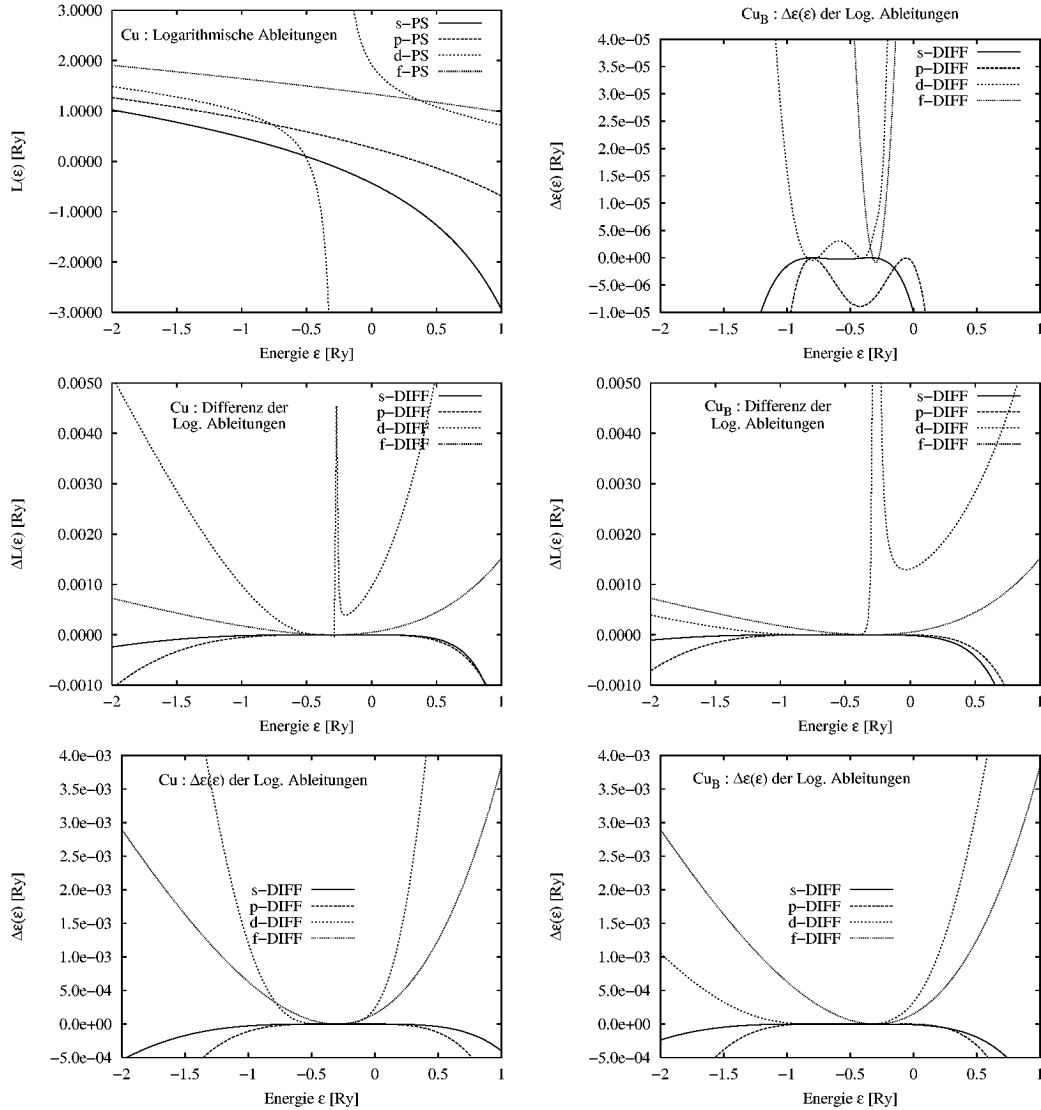


Abbildung 4.18: Logarithmische Ableitungen für Kupfer; links: Cu_A -, rechts Cu_B -Potential. Links oben sind die Logarithmischen Ableitungen selber wiedergegeben. Ein Unterschied zwischen AE- und PS-Funktionen ist nicht erkennbar (weder für Cu_A noch für Cu_B). In der Mitte ist jeweils die Differenz zwischen AE- und PS-Funktionen dargestellt, unten die Energieverschiebung $\Delta\epsilon(\epsilon)$. Rechts oben ist ein vergrößerter Ausschnitt daraus für das Cu_B -Potential gezeigt.

die Vermutung, daß evtl. die Ursache in der *frozen-core*-Näherung zu suchen sei. Eine AE-Rechnung für die angeregten Atome mit der festen *core*-Dichte des Grundzustandsatoms zeigt, daß die Fehler der Eigenwerte in diesem Falle durchaus bis zu 40% derjenigen der PAW-Rechnung betragen (siehe Tabelle 4.13); jedoch werden

hier trotzdem die Anregungsenergien sehr gut wiedergegeben, so daß die *frozen-core*-Näherung als Fehlerquelle ausscheidet.

Konfig.	$\Delta\varepsilon$ -AE-fc			$\Delta\varepsilon$ -PAW			ΔE_{exc} - AE-fc	ΔE_{exc} - PAW
	s	p	d	s	p	d		
$4s^1 4p^{0.5} 3d^{9.5}$	0.3	0.1	3.6	1.1	0.6	9.6	0.08	2.12
$4s^{0.5} 4p^{0.5} 3d^{10}$	0.1	0.03	0.3	0.4	0.1	1.0	0.01	0.05
$4s^2 3d^9$	0.7	0.4	11.0	1.1	1.7	25.0	0.61	11.4

Tabelle 4.13: Fehler ($\Delta\varepsilon$) der Valenzeigenwerte und der Anregungsenergien (ΔE_{exc}) durch die *frozen-core* (fc)-Näherung bei der AE-Rechnung im Vergleich zu denen der PAW-Kalkulation. Wiedergegeben ist jeweils die Differenz zu den selbstkonsistent ohne *frozen-core* berechneten AE-Werten. (Alle Energien in mRy.)

Die Ursache dieser Diskrepanzen bleibt unklar. Da die Abweichungen für kleinere Anregungen relativ gering sind, kann man hoffen, dennoch gute Ergebnisse mit diesem Pseudopotential zu erhalten. Es sei an dieser Stelle vorweggenommen, daß man beispielsweise für fcc-Kupfer eine vernünftige Gitterkonstante erhält. Des weiteren unterscheiden sich die Ergebnisse für die beiden Pseudopotentiale nicht signifikant. Um die benötigte *cutoff*-Energie zu bestimmen (bzw. die Aussage des Rappe-Kriteriums zu verifizieren), wurden wiederum die Gesamtenergie eines Einzelatoms in einer großen kubischen Zelle berechnet. Dies ist in Abbildung 4.19 aufgetragen. Wie

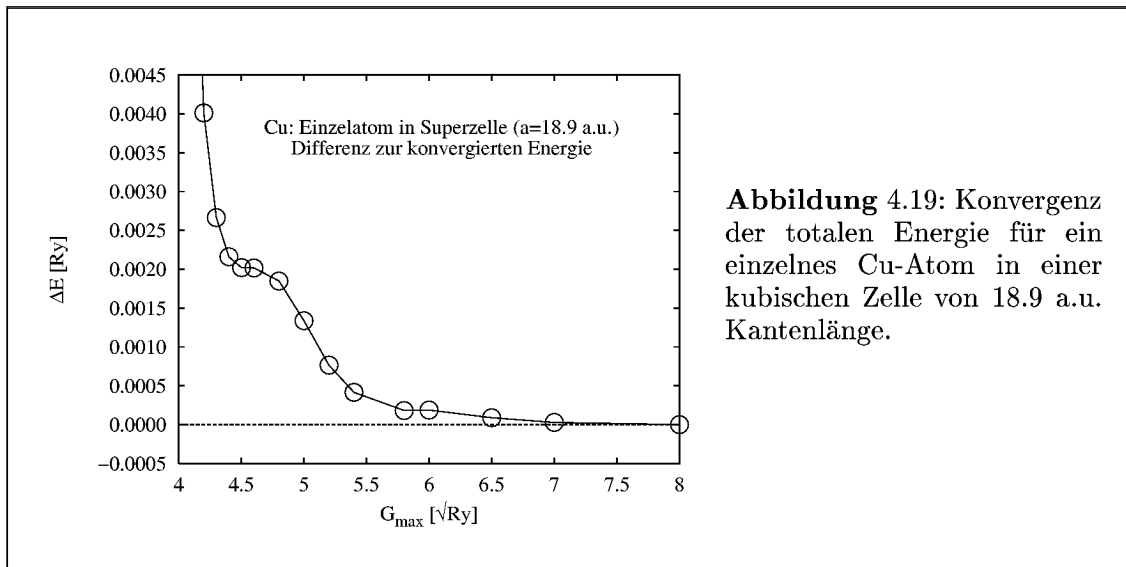


Abbildung 4.19: Konvergenz der totalen Energie für ein einzelnes Cu-Atom in einer kubischen Zelle von 18.9 a.u. Kantenlänge.

man sieht, wird das angestrebte Kriterium von 2 mRy pro Atom in der Tat bei einer Abschneideenergie von $G_{\max} = 4.5 \sqrt{\text{Ry}}$ erreicht.

Kapitel 5

Minimierung des Energiefunktional

5.1 Ionische und elektronische Freiheitsgrade

Im Energiefunktional treten zwei Arten von Freiheitsgraden auf:

- die ionischen Freiheitsgrade, d.h. die Atomkoordinaten $\vec{\tau}_\mu$
- die elektronischen Freiheitsgrade, d.h. die Wellenfunktionen $\tilde{\psi}_{\vec{k}\nu}$. Diese wiederum bestimmen eindeutig eine Ladungsdichte $n(\vec{r})$, bzw. das effektive Potential $V_{eff}[n(\vec{r})]$

Das Energiefunktional soll bzgl. aller Freiheitsgrade minimiert werden. Dazu existieren zwei verschiedene Strategien:

Zum einen schlugen Car und Parrinello [CP85] vor, elektronische und atomare Freiheitsgrade simultan zu minimieren. Dies birgt jedoch den Nachteil, daß die atomare Schrittweite sehr klein bleiben muß, damit die Elektronen in der Nähe des elektronischen Grundzustandes zur jeweiligen atomaren Konfiguration (d.h. der Born–Oppenheimer–Oberfläche) bleiben.

Zum anderen kann man die elektronischen Freiheitsgrade vor jedem Molekulardynamikschritt voll auskonvergieren lassen und dann relativ große atomare Schritte folgen lassen. In unserem Programm wird die zweite Methode umgesetzt. Zwei andere Neuerungen, die Car und Parrinello [CP85] gleichzeitig vorschlugen, werden jedoch sehr wohl berücksichtigt, nämlich die Benutzung iterativer Methoden zur Matrixdiagonalisierung und – dadurch ermöglicht – die Benutzung der *Fast Fourier*

er Transform zur schnelleren Auswertung der Produkte aus Hamiltonoperator und Wellenfunktion.

Die Minimierung des Energiefunktional bzgl. atomarer und elektronischer Freiheitsgrade läßt sich in beiden Fällen als Nullstellensuche eines hochdimensionalen, nichtlinearen Gleichungssystems formulieren. Diese Minimierungsprobleme sind ineinander verschachtelt:

1. In der äußersten Schleife, der atomaren Strukturrelaxation, muß die Grundzustandsenergie hinsichtlich der ionischen Freiheitsgrade $\{\vec{\tau}_\mu\}$ minimiert werden:

$$E_{tot} = \min_{\{\vec{\tau}_\mu\}} E_0 [\{\vec{\tau}_\mu\}] . \quad (5.1)$$

Dabei bezeichnet E_0 jeweils den elektronischen Grundzustand. Im Energieminimum hat die Energieableitung bzgl. der Atomkoordinaten eine Nullstelle $\vec{\tau}_0$:

$$\vec{\mathcal{F}}\{\vec{\tau}\} = -\frac{d}{d\vec{\tau}} E_0 [\vec{\tau}] \quad (5.2)$$

$$\vec{\mathcal{F}}\{\vec{\tau}_0\} = \vec{0} \quad (5.3)$$

Die Ableitung der Energie nach den Atomkoordinaten ist gerade durch die Hellmann–Feynman–Kräfte gegeben. Erwartungsgemäß ist eine Konfiguration nur dann stabil, wenn diese Kräfte verschwinden. Im Prinzip wird durch diese Bedingung nur ein *lokales* Minimum¹ beschrieben; gegebenenfalls muß man für verschiedene Anfangskonfigurationen in der Nähe plausibler Minimumkonfigurationen ausprobieren, ob sie zum selben Minimum führen. Andernfalls ist das lokale Minimum mit der tieferen Energie auszuwählen. Im Gegensatz etwa zu Gläsern ist jedoch die Anzahl metastabiler Konfigurationen für alle bislang berechneten Systeme gering.

2. Innerhalb der Strukturrelaxation muß das Energiefunktional

$$E_0 [\{\vec{\tau}_\mu\}] = \min_{\{\tilde{\psi}_{\vec{k}\nu}\}} E_{LDA} [\{\vec{\tau}_\mu\}, \{\tilde{\psi}_{\vec{k}\nu}\}] \quad (5.4)$$

bzgl. der elektronischen Freiheitsgrade bei festen Atompositionen minimiert werden. Dies führt auf das Problem, die Grundzustandsladungsdichte $n_0(\vec{r})$ bzw. das Grundzustandspotential $V[n_0(\vec{r})]$ selbstkonsistent zu bestimmen.

¹In einem lokalen Maximum verschwindet die Kraft auch; jedoch läuft man aufgrund stets auftretender numerischer Fehler aus einem solchen immer heraus.

Die Kohn–Sham Gleichungen definieren eine iterative Abbildung $n(\vec{r}) \rightarrow F\{n(\vec{r})\}$ bzw. $V(\vec{r}) \rightarrow \hat{F}\{V(\vec{r})\}$ gemäß:

$$\begin{pmatrix} n(\vec{r}) \\ V[n(\vec{r})] \end{pmatrix} \xrightarrow{(\mathcal{H}\tilde{\psi}=\epsilon\mathcal{O}\tilde{\psi})} \left\{ \tilde{\psi}_{\vec{k}\nu}, \epsilon_{\vec{k}\nu} \right\} \longrightarrow \begin{pmatrix} F\{n(\vec{r})\} := F\left\{n[\tilde{\psi}_{\vec{k}\nu}, \epsilon_{\vec{k}\nu}](\vec{r})\right\} \\ \hat{F}\{V[n(\vec{r})]\} := V[F\{n(\vec{r})\}] \end{pmatrix} \quad (5.5)$$

Die Kohn–Sham Gleichungen werden ausgehend von einer Ladungsdichte $n(\vec{r})$ bzw. dem Potential $V[n(\vec{r})]$ gelöst, d.h. die Wellenfunktionen $\tilde{\psi}_{\vec{k}\nu}$ und Energien $\epsilon_{\vec{k}\nu}$ bestimmt. Daraus wird wiederum die neue Ladungsdichte $F\{n(\vec{r})\}$ bestimmt, welche das neue Potential $\hat{F}\{V[n(\vec{r})]\}$ hervorbringt.

Die selbstkonsistente Ladungsdichte n_0 ist ein Fixpunkt der Abbildung $F\{n\}$ und das entsprechende selbstkonsistente Potential V_0 Fixpunkt von $\hat{F}\{V\}$. Die Fixpunkte sind Nullstellen der modifizierten Funktionale $\mathcal{F}\{n\}$ bzw. $\hat{\mathcal{F}}\{V\}$:

$$\begin{aligned} \mathcal{F}\{n_0\} &= F\{n_0\} - n_0 &= 0 \\ \hat{\mathcal{F}}\{V[n_0]\} &= \hat{F}\{V[n_0]\} - V[n_0] &= 0 \end{aligned} \quad (5.6)$$

Auch hier läßt sich $\mathcal{F}$ als eine in Richtung der Nullstelle treibende Kraft interpretieren.

In unserem Programm sind verschiedene Verfahren zur Nullstellensuche implementiert. Diese sind in [Eng95] ausführlich dargestellt. Als Standardverfahren werden für die atomaren und die elektronischen Freiheitsgrade Quasi–Newton–Verfahren verwendet, in denen auch der Jacobian der zu minimierenden Abbildung iterativ genähert wird.

Der Jacobian der atomaren Strukturrelaxation ist – im Gegensatz zu dem der elektronischen Selbstkonsistenz – symmetrisch. Daher wird hier im Regelfalle ein anderes Verfahren verwendet, nämlich ein Rang 2 Verfahren nach Broyden, Fletcher, Goldfarb und Shanno (vgl. [Eng95, Blü88]).

Für die elektronische Selbstkonsistenz bedeuten alle Verfahren letztendlich, daß die neu erzeugte Ladungsdichte (oder das Potential) linear mit ebensolchen aus vorangegangenen Schritten kombiniert wird. Deswegen spricht man auch vom *Mischen* der Ladung oder der Potentiale. Das *Gedächtnis* über die vorherige Iteration hinaus beschleunigt i. Allg. die Konvergenz; unter bestimmten Umständen kann sie diese aber auch beeinträchtigen; dann kann künstlicher Gedächtnisschwund (*Restart*) hilfreich sein. Im Regelfall werden in unserem Programm die Potentiale gemischt, und zwar geschieht dies im Ortsraum nach dem verallgemeinerten Anderson–Verfahren.

Für das PAW-Verfahren muß man in das Mischen der Potentiale (oder der Ladungsdichten) auch die Anteile in den Kugeln um die Atome einbeziehen, da ansonsten mit Konvergenzproblemen zu rechnen ist [Blü]. Dazu wurde eine Metrik eingeführt, die durch die relativen Raumanteile bestimmt ist.

Eine Anfangsladungsdichte muß vorgegeben werden; für PAW-Atome werden hierzu – sofern keine bessere Startladung vorhanden ist – die Ladungsdichten $n^{1\mu}$ und $\tilde{n}^{1\mu}$ eines isolierten Atomes eingelesen und die glatte Dichte $\tilde{n}$ als Superposition der $\tilde{n}^{1\mu}$ -Dichten (und gaußförmiger Dichten für andere Atome) gebildet. Erfahrungsgemäß ist die Konvergenz jedoch viel schneller, wenn man eine der aktuellen Konfiguration besser angepaßte Ladungsdichte vorgeben kann. Innerhalb der Strukturrelaxation ist die Dichte der vorhergehenden Konfiguration stets eine viel bessere Wahl.

5.2 Lösen der Eigenwertgleichung

Im Rahmen der Bestimmung der elektronischen Selbstkonsistenz müssen für jedes neue Potential die Eigenwerte und Eigenzustände der Kohn-Sham-Gleichung bestimmt werden. Die Eigenzustände werden im reziproken Raum bzgl. einer Basis aus Ebenen Wellen bestimmt.

$$H_{\vec{G},\vec{G}'} c_{\nu}^{\vec{G}'} = \varepsilon_{\nu} O_{\vec{G},\vec{G}'} c_{\nu}^{\vec{G}'} \quad (5.7)$$

5.2.1 Traditionelle Diagonalisierung

Für kleine Systeme, zu deren Beschreibung höchstens 1000 Basisfunktionen ausreichen, kann man die Matrizen im Impulsraum explizit aufstellen und das verallgemeinerte Eigenwertproblem durch Aufruf von Funktionen in Standard-Bibliotheken lösen.² Da die Überlappmatrix O positiv definit ist, gibt es eine Cholesky-Zerlegung, mittels derer sich das verallgemeinerte auf ein gewöhnliches hermitesches Eigenproblem zurückführen läßt [PTVF92].

Für größere Systeme ist die explizite Diagonalisierung ungeeignet, da ihr Speicherbedarf quadratisch, die Rechenzeit sogar kubisch mit der Anzahl der Basisfunktionen wächst.

²Das entsprechende *Driver*-Unterprogramm ist zwar erst in LAPACK 3.0 enthalten, welches zum Zeitpunkt der Programmerstellung noch nicht in optimierter Version zu Verfügung steht, aber die darin benötigten Unterprogramme sind verfügbar.

5.2.2 Iterative Diagonalisierung

Bei iterativen Verfahren zur Matrixdiagonalisierung macht man es sich zunutze, daß man nicht an allen Eigenzuständen interessiert ist, sondern nur an einer relativ kleinen Anzahl niedriger Zustände; die untere Grenze für diese Anzahl ist durch die Anzahl besetzter Zustände gegeben, da diese zur Berechnung der Ladungsdichte zwingend erforderlich sind. Dies bedeutet aber üblicherweise, daß man nur die unteren 2 – 10% der Eigenwerte und –vektoren benötigt.

Ein weiterer Vorteil der iterativen Methode rührt daher, daß man nicht ein einzelnes Eigenwertproblem löst, sondern innerhalb der Selbstkonsistenzschleife der Ladungsdichte eine ganze Folge von ebensolchen, welche sich relativ geringfügig von einer Iteration zu nächsten verändern. Bei iterativen Verfahren kann man daher die Lösungen der vorherigen Iteration wiederverwenden, da sie eine ziemlich gute Näherung der Eigenzustände des aktuellen Problems darstellen. Auch kann man die Genauigkeitsanforderung dem Fortgang der Selbstkonsistenz anpassen, da man weit entfernt vom Ladungsfixpunkt auch keine hochakkuraten Eigenwerte benötigt.

Der entscheidende Punkt ist jedoch, daß man bei iterativen Verfahren nicht die komplette Hamiltonmatrix aufstellen muß; die Berechnung großer Systeme wäre sonst unmöglich, da die Anzahl der Basisfunktionen dort ca. 10^5 betragen kann.

Bei iterativen Verfahren löst man sukzessive Eigenwertprobleme auf Unterräumen. Die Eigenwerte die man dort bestimmt, nähern sich den gewünschten Eigenwerten umso mehr an, je besser dieser Unterraum den Raum der Eigenzustände des vollen Eigenwertproblems approximiert.

5.2.2.1 Rayleigh-Ritz Verfahren für Standardeigenwertprobleme

Die bestmögliche Lösung des Eigenwertproblems auf einem Unterraum wird durch das Rayleigh-Ritz Verfahren bestimmt. Der Unterraum $\mathcal{V}$ wird von n orthonormierten Vektoren v_i aufgespannt.³ Zusammen bilden sie ein Bündel

$$\mathbf{V} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\} \quad (5.8)$$

von Vektoren. Die Orthonomierung wird durch

$$\mathbf{V}^+ \mathbf{V} = \mathbf{1} \quad (5.9)$$

³Die Größe des Unterraumes $n = ndvdim$ ist viel kleiner als die Anzahl $N = nbaukp$ der Komponenten des Vektors.

beschrieben. Das Eigenwertproblem auf dem Unterraum ist durch

$$(\mathbf{V}^+ \mathbf{H} \mathbf{V}) \mathbf{u}_i = \lambda_i \mathbf{u}_i \quad (5.10)$$

gegeben. Der Ritzwert λ_i ist dann die beste auf diesem Unterraum erzielbare Approximation des Eigenwertes des vollen Eigenwertproblems und der zugehörige Ritzvektor ist gegeben durch $\mathbf{w}_i = \mathbf{V} \mathbf{u}_i$. Sie sind die Eigenpaare des auf den Unterraum $\mathcal{V}$ projizierten Operators:

$$\mathbf{V}^+ [(\mathbf{H} - \lambda_i \mathbf{1})(\mathbf{V} \mathbf{u}_i)] = 0 \quad (5.11)$$

Ist $\mathcal{V}$ eindimensional, so ist λ_i gerade der Rayleigh-Quotient $\rho(v) = \frac{\langle v | H | v \rangle}{\langle v | v \rangle}$.

Die Ritzvektoren sind - jeweils separat betrachtet - in mehrerer Hinsicht nicht optimal; insbesondere wird *nicht* $\|\mathbf{H} \mathbf{w} - \mathbf{w} \rho(\mathbf{w})\|$ optimiert.⁴ Stattdessen erhält man eine *kollektive* Optimaleigenschaft [Par80].

Wie schon erwähnt braucht man beim Rayleigh-Ritz Verfahren nicht die vollständige Matrix $\mathbf{H}$, sondern muß nur in der Lage sein, für einen Vektor $\mathbf{v}$ das Produkt $\mathbf{H} \mathbf{v}$ zu berechnen.

5.2.2.2 Davidson–Kosugi–Verfahren

Das Problem besteht also darin, eine Sequenz von Unterräumen zu erzeugen, welche die Eigenvektoren der niedrigen Eigenwerte immer besser enthalten.

Eine Variante dazu sind Krylov-Verfahren; in ihnen werden die Unterräume erzeugt durch diejenigen Vektoren, die durch Anwendung von Potenzen eines Operators auf einen Anfangsvektor entstehen.

Beim Davidson–Verfahren [Dav75] dagegen ist der Ausgangspunkt der Rayleigh-Quotient. Dieser wird bekanntlich für Eigenzustände stationär. Daher kann man die Eigenwertsuche als Optimierungsproblem auffassen. Die Richtung des steilsten Anstiegs wird durch das Residuum gegeben

$$|R_i\rangle = (\mathbf{H} - \rho[\psi_i])|\psi_i\rangle \quad (5.12)$$

Man könnte einfach das Residuum als neuen Vektor dem Unterraum hinzufügen. Dieses *steepest-descent*-Verfahren ist zwar sehr stabil, aber meist nur langsam konvergent. Schlecht ist es insbesondere, falls das Minimum in einem stark anisotropen Tal liegt. Das ist der Fall, wenn sich die Eigenwerte stark unterscheiden. Dies kann

⁴Dies geschieht dagegen im RMM-DIIS-Verfahren [KF96, WZ85].

man durch einen Präkonditionierer mildern. In dieser Weise kann man den Korrekturvektor im Davidson-verfahren auffassen:

$$|v^{kor}\rangle = \underbrace{(D - \rho)^{-1} (\mathbf{H} - \rho \mathbf{1})}_{N(\rho)} |\psi\rangle \quad (5.13)$$

wobei

$$\rho = \frac{\langle \psi | \mathbf{H} | \psi \rangle}{\langle \psi | \psi \rangle} \quad (5.14)$$

und

$$D_{\vec{G}, \vec{G}'} = \mathbf{H}_{\vec{G}, \vec{G}} \delta_{\vec{G}, \vec{G}'}$$

die Diagonale von $\mathbf{H}$ ist. Aus Untersuchungen für *Conjugate Gradients*-Verfahren weiß man, daß die Präkonditionierung mit der Diagonalen dazu tendiert, das Eigenwertspektrum um den Wert eins zu komprimieren, wodurch die Anisotropie gemildert wird [MS86].

Da die Matrix $N(\rho)$ nicht konstant ist, handelt es sich nicht um ein Krylov-Verfahren, nähert sich einem solchen aber an, wenn ρ gegen den echten Eigenwert ε konvergiert.⁵

Das Davidson-Verfahren war ursprünglich nur für einen einzelnen Eigenwert konstruiert. Liu verallgemeinerte dies zu einem Block-Algorithmus [Liu78], welcher mehrere Eigenwerte zugleich bestimmt.

Durch das ständige Hinzufügen von Korrekturvektoren wächst der Unterraum schnell an zu einer nicht mehr effizient zu bewältigenden Größe. Daher werden nach einer Idee von Kosugi [Kos84] nach jeder Iteration nur die Ritzvektoren zu den niedrigsten Eigenwerten behalten, die restlichen Vektoren aber durch Korrekturvektoren *ersetzt* (nicht hinzugefügt), so daß eine Folge von Unterräumen identischer Dimension erzeugt wird. Eine sinnvolle untere Grenze dafür ist zweimal die Anzahl der zu bestimmenden Eigenzustände, da zumindest für jeden Ritzvektor ein Korrekturvektor konstruiert werden sollte.

Dieses Verfahren nach Davidson, Kosugi und Liu ist in unserem Programm implementiert. Der Einfachheit halber werde ich es als *Davidson-Verfahren* bezeichnen.

Für eine detaillierte Beschreibung sei auf [Ric96, Ric93] verwiesen. Weitere theoretische Erörterung und Referenzen findet man z.B. in [CPS94, Saa92, GvdV00].

⁵Das ist insofern interessant, als es zumindest plausibel macht, warum einige Eigenschaften von Krylov-Verfahren gelten sollten.

5.2.2.3 Verallgemeinerte Eigenwertprobleme

Das Rayleigh-Ritz Verfahren läßt sich auch auf ein verallgemeinertes Eigenwertproblem anwenden [Mor90]. Dies kann auf zwei Arten geschehen:

- Einerseits kann man die Zustände v_i , welche einen Unterraum V aufspannen, statt bzgl. des gewöhnlichen Skalarproduktes orthonormieren bzgl. eines verallgemeinerten Skalarproduktes

$$\mathbf{V}^+ \mathbf{O} \mathbf{V} = \mathbf{1} \quad (5.15)$$

Dann erhält man durch Lösen eines gewöhnlichen Eigenwertproblems auf dem Unterraum Ritzwerte und -vektoren.

$$\begin{aligned} (\mathbf{V}^+ \mathbf{H} \mathbf{V}) u_i &= \lambda_i u_i \\ \parallel &\quad \parallel \\ \mathbf{V}^+ \mathbf{H} (\mathbf{V} u_i) &\quad \lambda_i (\mathbf{V}^+ \mathbf{O} \mathbf{V}) u_i = \mathbf{V}^+ \mathbf{O} \lambda_i (\mathbf{V} u_i) \end{aligned} \quad (5.16)$$

und damit

$$\mathbf{V}^+ [(\mathbf{H} - \lambda_i \mathbf{O}) (\mathbf{V} u_i)] = 0 \quad (5.17)$$

- Andererseits kann man die Zustände v_i auch bzgl. des gewöhnlichen Skalarproduktes orthonormieren

$$\mathbf{V}^+ \mathbf{V} = \mathbf{1} \quad (5.18)$$

und hat dafür auf dem Unterraum ein verallgemeinertes Eigenwertproblem zu lösen

$$\begin{aligned} (\mathbf{V}^+ \mathbf{H} \mathbf{V}) u_i &= \lambda_i (\mathbf{V}^+ \mathbf{O} \mathbf{V}) u_i \\ \parallel &\quad \parallel \\ \mathbf{V}^+ \mathbf{H} (\mathbf{V} u_i) &\quad \mathbf{V}^+ \mathbf{O} \lambda_i (\mathbf{V} u_i) \end{aligned} \quad (5.19)$$

und erhält wieder Glg. 5.17.

In beiden Fällen muß man zusätzlich zu $\mathbf{H}v_i$ auch $\mathbf{O}v_i$ berechnen und die Skalarprodukte $\langle v_i | \mathbf{O} v_j \rangle$ berechnen. Der Aufwand ist daher vergleichbar. Bei der *traditionellen* Methode erfüllt die Basis notwendig die Orthogonalität bezüglich des normalen Skalarproduktes. Für die iterative Methode hingegen führt die Orthogonalisierung bzgl.

des verallgemeinerten Skalarproduktes zum einfacheren Algorithmus. Jedoch zeigen sich dabei numerische Schwierigkeiten. Diese mögen z.T. dadurch bedingt sein, daß aus Gründen der Effizienz die Produkte von Hamilton- und Überlappmatrix mit den Vektoren gleichzeitig ausgeführt werden müssen. Dies bedingt, daß diese Vektoren bei der Orthogonalisierung mittransformiert werden müssen. Da bei kleinen Eigenwerten der Überlappmatrix hierbei große Zahlen auftreten können, wird hierdurch offenbar eine erhebliche Ungenauigkeit eingeführt. Daher wurde auch für die iterative Lösung die zweite Variante für die Implementierung gewählt.

Die Bestimmung der Korrekturvektoren für das Davidson-Verfahren erfolgt analog zum oben beschriebenen Fall, d.h. das Residuum ist gegeben durch

$$|R_i\rangle = \mathbf{H} \mathbf{v}_i - \varepsilon_i \mathbf{O} \mathbf{v}_i \quad (5.20)$$

und der Korrekturvektor wird als

$$|v^{korr}\rangle = (D^{\mathbf{H}} - \rho D^{\mathbf{O}})^{-1} (\mathbf{H} - \rho \mathbf{O}) |\psi\rangle \quad (5.21)$$

wobei

$$\rho = \frac{\langle \psi | \mathbf{H} | \psi \rangle}{\langle \psi | \mathbf{O} | \psi \rangle} \quad (5.22)$$

und

$$D_{\vec{G}, \vec{G}'}^{\mathbf{H}} = \mathbf{H}_{\vec{G}, \vec{G}} \delta_{\vec{G}, \vec{G}'}, \quad D_{\vec{G}, \vec{G}'}^{\mathbf{O}} = \mathbf{O}_{\vec{G}, \vec{G}} \delta_{\vec{G}, \vec{G}'}$$

die Diagonalen von $\mathbf{H}$ und $\mathbf{O}$ sind.

5.2.2.4 Der Algorithmus

- Initialisierung

- Einlesen der Startvektoren $\{\mathbf{v}_i\}$ ($i=1, \text{nstate}$)
- Orthogonalisiere $\{\mathbf{v}_i\}$ ($i=1, \text{nstate}$) bzgl. Standardskalarprodukt ⁶
- Berechne $\mathbf{H}\mathbf{v}_i, \mathbf{O}\mathbf{v}_i, \lambda_i = \mathbf{v}_i^+ \mathbf{H} \mathbf{v}_i$ ($i=1, \text{nstate}$)
- Berechne Residuen $\mathbf{r}_i = \mathbf{H}\mathbf{v}_i - \lambda_i \mathbf{O}\mathbf{v}_i$

- Iteration

- Teste Konvergenz (Norm der Residuen); falls alle Vektoren konvergiert, beende Iteration

⁶Dies muß nicht bei jedem Aufruf getan werden; bislang wurde es jedoch nie (sic!) gemacht. Dies führt zu einer schleichenden Verschlechterung der Orthogonalität.

- Berechne Korrekturvektoren $\mathbf{v}_{nstate+i} = f(\mathbf{r}_i)$ ($i=1, nstate$) für nicht konvergierte Vektoren; fülle die restlichen Stellen mit früheren Korrekturvektoren noch nicht konvergierter Zustände auf.
- Orthogonalisiere $\{\mathbf{v}_i\}$ ($i=1, ndvdim$) bzgl. Standardskalarprodukt; die ersten $nstate$ Vektoren sind schon orthogonal.
- Berechne $\mathbf{H}\mathbf{v}_i, \mathbf{O}\mathbf{v}_i$ ($i=nstate+1, ndvdim$)
- Berechne Unterraummatrizen $\mathbf{M}_\mathbf{H} = \mathbf{V}^+\mathbf{H}\mathbf{V}$, $\mathbf{M}_\mathbf{O} = \mathbf{V}^+\mathbf{O}\mathbf{V}$ (Dabei kann man Teile der Matrizen aus den Daten der vorherigen Iteration bestimmen; s.u.)
- Löse verallgemeinertes Unterraum-Eigenwertproblem $\mathbf{M}_\mathbf{H}\mathbf{U} = \mathbf{M}_\mathbf{O}\mathbf{U}diag(\lambda_i)$ ($i=1, ndvdim$)
- Rücktransformation

$$\mathbf{v}_i^{neu} = \mathbf{V}_{alt}\mathbf{u}_i \quad (i=1, nstate)$$

$$\mathbf{H}\mathbf{v}_i^{neu} = (\mathbf{H}\mathbf{V}_{alt})\mathbf{u}_i$$

$$\mathbf{O}\mathbf{v}_i^{neu} = (\mathbf{O}\mathbf{V}_{alt})\mathbf{u}_i$$
 Die Vektoren $\mathbf{v}_i$ sind jetzt orthogonal bzgl. $\mathbf{O}$.
- Berechne Residuen $\mathbf{r}_i = \mathbf{H}\mathbf{v}_i - \lambda_i\mathbf{O}\mathbf{v}_i$
- Falls nicht letzter Schritt:

Orthogonalisiere $\mathbf{v}_i$ bzgl. Standardskalarprodukt:
 Bilde Cholesky-Zerlegung $\mathbf{L}^+\mathbf{L} = \mathbf{U}_{nstate}^+\mathbf{U}_{nstate}$
 $\mathbf{v}_i^{orth} = \mathbf{V}_{nstate}^{neu}(\mathbf{L}^+)^{-1}$ (und entsprechend für $\mathbf{H}\mathbf{v}_i^{orth}$ und $\mathbf{O}\mathbf{v}_i^{orth}$)
 Der obere $nstate \times nstate$ -Teil der Unterraummatrizen $\mathbf{M}_\mathbf{O}$ bzw. $\mathbf{M}_\mathbf{H}$ ist dann durch $\mathbf{L}^{-1}(\mathbf{L}^+)^{-1}$ bzw. $\mathbf{L}^{-1}diag(\lambda_i)(\mathbf{L}^+)^{-1}$ gegeben.

5.2.2.5 Startvektoren

Wie oben schon erwähnt, benötigt man für das Davidson-Verfahren Anfangslösungen. Ab dem zweiten Selbstkonsistenzschritt kann man dazu auf die Lösungen des vorhergehenden Diagonalisierungsschrittes zurückgreifen. Beim allerersten Aufruf jedoch muß man anderweitig Startvektoren bestimmen. Diese sollten möglichst gut sein, da die weitere Konvergenz dadurch bestimmt wird.

Wenn man gar nichts über die zu diagonalisierende Matrix weiß, so kann man die Vektoren zufällig besetzen. Diese stellen allerdings i. Allg. ziemlich schlechte Lösungen dar.

Die bislang praktizierte Lösung in unserem Programm war, im ersten Schritt das Eigenwertproblem traditionell zu lösen, d.h. die Hamiltonmatrix explizit aufzustellen und zu diagonalisieren. Dies ist aber nur für Matrizen beschränkter Größe praktikabel; das bedeutet speziell bei großen Systemen, daß man sich auf eine unnatürlich kleine *cutoff*-Energie beschränken muß. Die höheren Komponenten der so gewonnenen Eigenvektoren werden als Null betrachtet. Dabei ist freilich nicht sichergestellt, daß diese Vektoren nahe genug an den Eigenvektoren liegen, um eine schnelle Konvergenz zu ermöglichen.

Daher wurde eine neue Möglichkeit zur Bestimmung von Startvektoren implementiert. Sie basiert auf der Idee, einen Unterraum aufzuspannen aus (periodisch fortgesetzten) atomaren Eigenfunktionen der beteiligten Atome. Auf diesem Raum werden dann nach dem Rayleigh-Ritz-Verfahren die bestmöglichen Lösungen bestimmt. Diese werden demnach als eine Linearkombination atomarer Orbitale (LCAO) gebildet, jedoch nicht aus den echten, sondern den pseudisierten atomaren Wellenfunktionen. Diese Startvektoren wurden APSO-Vektoren getauft (wie atomare Pseudo-Orbitale).

Details und Formeln Der Radialteil der atomaren Wellenfunktion $\phi_l^\alpha(r)$ wird eingelesen.

$$\phi_L^\alpha(\vec{r}) = Y_L(\vec{r}) \frac{1}{r} \phi_l^\alpha(r) \quad (5.23)$$

Daraus konstruiert man einen Blochzustand zum Wellenvektor $\vec{k}$ mittels

$$\varphi_{\vec{k}\nu}^\alpha(\vec{r}) = \sum_{\vec{R}} e^{-i\vec{k}\vec{R}} \varphi_{\vec{k}\nu}^\alpha(\vec{r}). \quad (5.24)$$

Dabei ist in diesem Falle $\nu = (\mu, L)$. Die Darstellung im reziproken Raum ist wieder ein Produkt von Struktur- und Formfaktor

$$\langle \vec{k} + \vec{G} | \varphi_{\vec{k}\nu} \rangle = c S^\mu(\vec{G}) \sqrt{4\pi} Y_L(\vec{k} + \vec{G}) \phi_l^\alpha(|\vec{k} + \vec{G}|) \quad (5.25)$$

wobei c eine unimodulare Konstante ist und

$$\phi_l^\alpha(|\vec{k} + \vec{G}|) = \sqrt{\frac{4\pi}{\Omega}} \int_0^\infty r^2 j_l(|\vec{k} + \vec{G}| r) \phi_l^\alpha(r) dr. \quad (5.26)$$

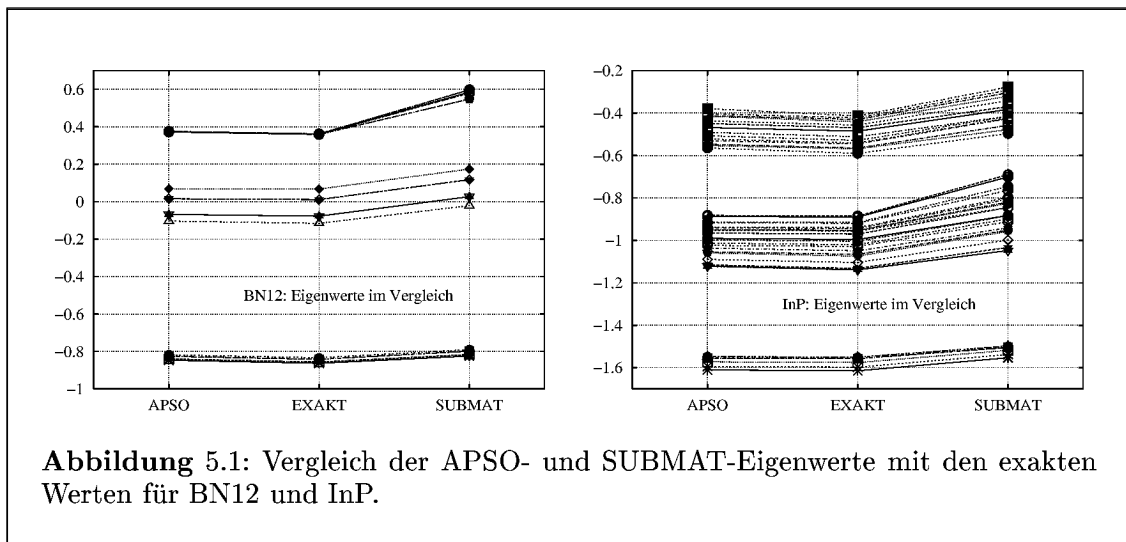
Da diese Wellenfunktionen nicht orthogonal sind, werden sie einem Gram-Schmidt-Verfahren unterzogen. Optional können eine Anzahl Ebene Wellen als zusätzliche Basisfunktionen hinzugefügt werden.

Auf dem durch diese Basisfunktionen aufgespannten Raum werden die Ritzwerte und -vektoren bestimmt und diese als Startvektoren für das nachfolgende Davidson-Verfahren gespeichert.

Dieses Verfahren wurde zunächst für das Standardeigenwertproblem (d.h. normerhaltende Potentiale) implementiert und später auch auf generalisierte Eigenwertprobleme (d.h. für die PAW-Methode) verallgemeinert.

Erfahrungen bei der Anwendung Zum Testen wurden für verschiedene Systeme sowohl bei festgehaltener Ladungsdichte die Eigenwerte berechnet als auch das Verhalten innerhalb der elektronischen Selbstkonsistenz untersucht. Unter anderem wurde eine InP-Oberflächenzelle mit 14 Atomen als typisches Halbleitersystem sowie ein BN-System aus drei übereinandergestapelten Einheitszellen mit insgesamt 12 Atomen (daher BN12) als Beispiel einer weniger vorteilhaft mit Ebenen Wellen zu beschreibenden Anordnung.

Das Verfahren funktioniert sehr gut. Mit sehr wenigen APSO-Funktionen erzielt man Ritzwerte, die sehr nahe bei den tatsächlichen Eigenwerten liegen. Beispielsweise erzielt man für die kleine, reine Siliziumzelle mit 8 APSO-Wellenfunktionen bessere Eigenwerte als mit 80 Ebenen Wellen. Für BN und InP sind in Abbildung 5.1 die Eigenwerte, die mit den APSO-Startvektoren bestimmt werden, verglichen mit den exakten Eigenwerten (zu der Startladungsdichte) und den Eigenwerten, die das bisher verwendete Submatrix-Verfahren (SUBMAT) liefert.



Die Startladungsdichte bestand aus einer Überlagerung gaußförmiger Ladungsverteilungen um die Atome. Die Dimension der Submatrizen betrug 539 (BN12) bzw.

403 (InP) bei einer Basisgröße von 2667 bzw. 3187 Funktionen. Im ersten Falle wurden nur s - und p - APSO-Funktionen verwendet (also 48 Stück), bei InP hingegen s -, p - und d -APSO-Zustände (insgesamt 128). Man erkennt, daß die Eigenwerte trotz deutlich geringerer Basisgröße wesentlich besser durch die APSO-Basis wiedergegeben werden.

Für BN12 wurde derselbe Versuch auch mit der Superposition atomarer Ladungsdichten durchgeführt. Es ergibt sich qualitativ dasselbe Bild. Bei beiden Ladungsdichten werden die Energien der besetzten Zustände sehr gut, die der unbesetzten weniger gut wiedergegeben, was bei einer Beschränkung auf s - und p -artige Zustände auch nicht überrascht. Daher wurde die Basis von 48 APSO-Funktionen um 49 Ebene Wellen erweitert. Damit wurden auch für die unbesetzten Zustände hervorragende Näherungswerte der Eigenenergien gefunden.

Insofern kann man davon ausgehen, daß man auch für sehr große Superzellen durch Benutzung der APSO-Wellenfunktionen (evtl. unter Hinzunahme weniger Ebener Wellen) bei einer bewältigbaren Anzahl von Funktionen gute Anfangslösungen erhält. Untersucht man, welchen Einfluß die APSO-Startvektoren auf die Konvergenz haben, so zeigt sich jedoch (bei den meisten betrachteten Systemen) sowohl bei der bloßen Eigenwertbestimmung (d.h. für eine festgehaltene Ladungsverteilung) wie auch innerhalb der Selbstkonsistenz, daß diese nicht zu einer durchgreifenden Verbesserung der Konvergenz führen.

Tabelle 5.1 faßt die benötigten Iterationsschritte in beiden Fällen für BN12 und InP zusammen.

	BN12				InP			
	APSO		SUBMAT		APSO		SUBMAT	
	Iter. (sc/dv)		Iter. (sc/dv)		Iter. (sc/dv)		Iter. (sc/dv)	
noSC	34		58		95		97	
SC	14	45	18	57	11	36	11	36

Tabelle 5.1: Anzahl der Iterationen APSO-Startvektoren und mit Submatrix-Diagonalisierung für BN12 (links) und InP (rechts) mit (SC) und ohne (noSC) Selbstkonsistenz.

Am Beispiel von InP – und dies ist nach der Erfahrung etwa mit Si, aber auch GaN, der typischere Fall – erkennt man, daß die APSO-Wellenfunktionen keinen (oder im Falle festgehaltener Ladungsdichte bestenfalls einen marginalen) Vorteil bringen.

Um das Verhalten bei BN12 für die SUBMAT-Methode zu verstehen, wurde die

Entwicklung der Residuen über die Iterationen hinweg betrachtet. Diese sind in Abbildung 5.2 links für ausgewählte Zustände aufgetragen. Man erkennt, daß die

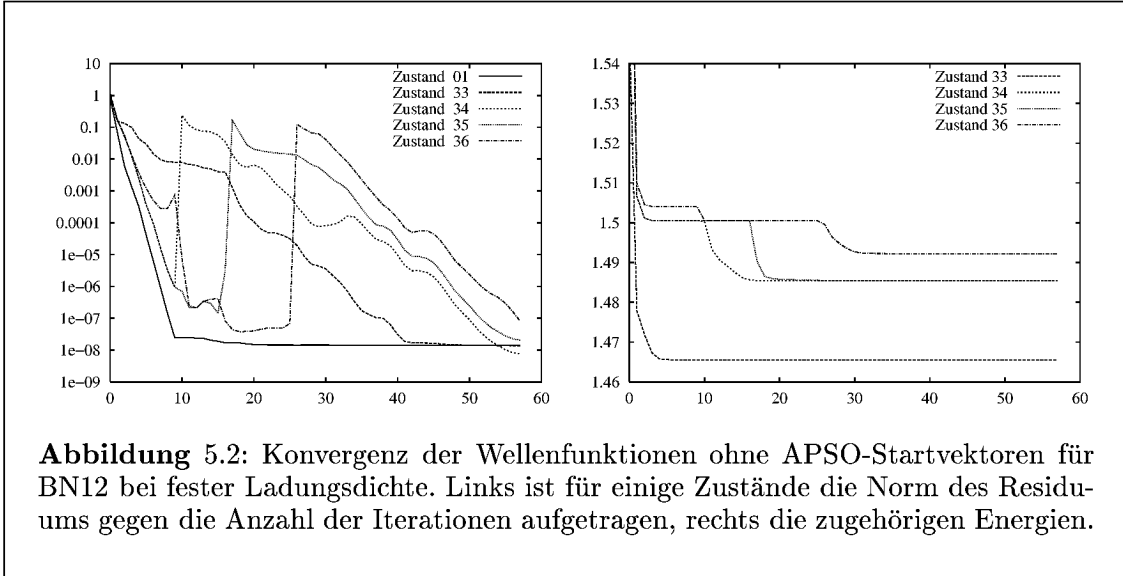


Abbildung 5.2: Konvergenz der Wellenfunktionen ohne APSO-Startvektoren für BN12 bei fester Ladungsdichte. Links ist für einige Zustände die Norm des Residuums gegen die Anzahl der Iterationen aufgetragen, rechts die zugehörigen Energien.

Normen der Residuen Sprünge über viele Größenordnungen aufweisen. Dies indiziert, daß an diesen Stellen plötzlich fehlende Eigenwerte gefunden werden; die rechte Seite der Abbildung belegt dies. Bis dahin konvergiert das Verfahren gegen einen falschen (höheren) Zustand. Das heißt, daß für den letzten Eigenwert die Konvergenz nach ca. 25 Schritten ganz von vorne beginnt. Dies führt zu der erheblich höheren Anzahl von benötigten Iterationen.

Es muß jedoch relativierend hinzugefügt werden, daß im hiesigen Fall nur unbesetzte Zustände betroffen sind; kritisch für die Selbstkonsistenz wäre es, wenn besetzte Zustände betroffen wären, da dann eine falsche Ladungsdichte bestimmt würde. Die höhere Iterationsanzahl im selbstkonsistenten Fall ist hier jedoch nicht durch einen solchen Effekt bedingt; dort konvergiert ein einzelner Vektor schlecht; dies kann man nicht allein der Startlösung zuschreiben; vgl Abschnitt 5.2.4.

Zu bedenken ist bei den Ergebnissen hinsichtlich der Konvergenz jedoch, daß die hier untersuchten Systeme noch so klein sind, daß die Submatrixgröße immer noch einem relativ hohen *Cutoff* entspricht. So entsprechen die 539 Ebenen Wellen bei BN12 immer noch einem Abschneidewert von etwa $G_{max} = 3.6 \sqrt{\text{Ry}}$. Bei sehr großen Supperzellen hingegen zwingt die Beschränkung auf eine handhabbare Größe effektiv zu einem viel kleineren Abschneidewert (relativ zum eigentlich erforderlichen). Insofern entsprechen diese Tests noch nicht den Bedingungen großer Supperzellen.

In der Quintessenz decken sich die hiesigen Beobachtungen mit denen in [Are97]; dort

wurden Startvektoren aus Eigenvektoren von Bandmatrizen verschiedener Breite gewonnen. Eine signifikante langfristige Verbesserung der Konvergenz konnte auch dort nicht festgestellt werden.

Somit stellen diese Möglichkeiten, Startvektoren zu erzeugen, i. Allg. gleichwertige Alternativen dar. Wenn man jedoch – evtl. bei Elementen mit stärker lokalisierten Funktionen – auf den Fall stieße, daß das traditionelle Verfahren keine brauchbare Anfangslösung mehr liefert, oder man zunächst eine größere Anzahl von Davidson-schritten folgen lassen muß, um in den Konvergenzbereich zu gelangen, so bieten die APSO-Vektoren eine Möglichkeit, schneller hinreichend gute Anfangslösungen zu generieren.

LCAO-Ladungsdichte Die Ergebnisse im letzten Abschnitt leiten zu der Schlußfolgerung, daß eine Verbesserung der Konvergenz vielleicht weniger bei den anfänglichen Wellenfunktionen als bei der anfänglichen Ladungsdichte zu suchen ist. Daher liegt es nahe, zunächst eine Ladungsdichte zu bestimmen, die für Wellenfunktionen beschränkt auf den Raum der APSO-Zustände selbstkonsistent ist.

Entsprechende Tests wurden für BN12, InP und Si/InSb durchgeführt. Dabei wurden ca. 5 Selbstkonsistenziterationen mit Beschränkung auf den APSO-Unterraum ausgeführt und die so ermittelte Ladungsdichte und Wellenfunktionen als Startlösungen bei der unbeschränkten Suche verwendet. Zwar ist der Anfangsfehler der Ladungsdichte (oder des Potentials) geringer (typischerweise um einen Faktor 3); jedoch konvergiert die Ladung danach schlechter, so daß sich per saldo eher eine Verschlechterung als eine Verbesserung ergibt.

5.2.3 Präkonditionierung

Hat man zu einem Eigenwertproblem $\mathbf{H}\psi = \epsilon\psi$ eine Näherungslösung $v \approx \psi$ mit Rayleigh-Quotient ρ und Residuum r , so sollte der Korrekturvektor z idealerweise $\psi = v + z$ erfüllen, oder

$$(H - \epsilon)z = -r + (\epsilon - \rho)v \quad (5.27)$$

Da die Eigenwerte schneller konvergieren als die Eigenvektoren, vernachlässigt man den zweiten Term und erhält

$$(H - \epsilon)z = -r \quad (5.28)$$

Diese Gleichung gilt es näherungsweise zu lösen. Präkonditionierung besteht darin, die Matrix $(H - \epsilon)$ durch eine Approximation M zu ersetzen, die sich leicht invertieren läßt. Wie schon erwähnt, wird im Davidson-Verfahren die Diagonale von $H - \rho$ als Präkonditionierer gewählt. Diese Methode sollte insbesondere für diagonal dominante Matrizen gut funktionieren, weil für diese der wesentliche Anteil der Matrix im Präkonditionierer gut berücksichtigt ist.

Dieses Verfahren läßt sich mit anderen Präkonditionierern verallgemeinern, welche der jeweiligen Matrixstruktur besser angepaßt sind [MS86].

Allerdings darf er auch nicht „zu gut“ sein, denn die exakte Lösung von Glg. 5.28 führt dazu, daß der Korrekturvektor gleich dem vorherigen Ritzvektor wird, $v = z$, und das Verfahren zusammenbricht.

In den Hamiltonmatrizen, die in den Kohn–Sham-Gleichungen in unserem Programm auftreten, haben zwar die Ebenen Wellen mit großer kinetischer Energie große Werte als Diagonalelemente, sie sind jedoch nicht (im mathematischen Sinne) diagonal dominant⁷. Obwohl das Davidson-Verfahren im Regelfalle trotzdem gut funktioniert, liegt es nahe zu versuchen, ob sich nicht mit einer Matrix M , welche der Struktur der hier auftretenden Matrizen besser angepaßt ist, eine deutliche Verbesserung der Konvergenzeigenschaften erzielen läßt.

TAP-Präkonditionierer Eine Möglichkeit ist, eine andere Diagonalmatrix als Präkonditionierer zu benutzen. Eine solche wurde in [TPA89] (für die *Conjugate Gradient*-Methode) vorgeschlagen. Die Grundidee dabei ist, hohe Wellenzahlen zu dämpfen, die niedrigen aber unverändert zu lassen. Der Präkonditionierer hat folgende Form:

$$M_{\vec{G}, \vec{G}'} = \delta_{\vec{G}, \vec{G}'} \frac{27 + 18x + 12x^2 + 8x^3}{27 + 18x + 12x^2 + 8x^3 + 16x^4} \quad (5.29)$$

wobei

$$x(\vec{k} + \vec{G}) = \frac{|\vec{k} + \vec{G}|^2}{E_{kin}} \quad (5.30)$$

und

$$E_{kin} = \langle \Psi_{\vec{k}\nu} | -\nabla^2 | \Psi_{\vec{k}\nu} \rangle \quad (5.31)$$

Dieser TAP-Präkonditionierer wird im folgenden Kapitel zusammen mit der Jacobi-Davidson-Methode untersucht.

⁷Eine quadratische Matrix A heißt *diagonal dominant*, wenn für alle Zeilen der Betrag des Diagonalelementes größer ist als die Summe der Beträge aller übrigen Elemente derselben Zeile.

APSO-Präkonditionierer Eine andere Möglichkeit, einen Präkonditionierer zu konstruieren, drängt sich im Zusammenhang der APSO-Startvektoren auf. Da die Eigenlösungen, die mit diesen Funktionen generiert werden, offenbar schon recht gut sind, kann man hoffen, daß die Hamiltonmatrix schon durch wenige APSO-Vektoren gut wiedergegeben wird:

$$H - \theta \mathbf{1} \approx M = \sum_{\gamma} |\phi_{\gamma}\rangle \epsilon_{\gamma}^{APSO} \langle \phi_{\gamma}| - \theta \mathbf{1} \quad (5.32)$$

Ergänzt man nun ϕ_{γ} zu einer Orthonormalbasis, so erhält man leicht die benötigte Inverse Matrix:

$$\theta M^{-1} = -\mathbf{1} + \sum_{\gamma} |\phi_{\gamma}\rangle \frac{\epsilon_{\gamma}^{APSO}}{\epsilon_{\gamma}^{APSO} - \theta} \langle \phi_{\gamma}| \quad (5.33)$$

Als Testsystem für diesen Präkonditionierer diene InP. Erprobt wurde das obige Verfahren; dessen Kombination mit dem Diagonalpräkonditionierer; eine Modifikation, bei der im Präkonditionierer zum Zustand $|\psi_{\gamma'}\rangle$ die Summe in Glg. 5.33 auf $\gamma \neq \gamma'$ beschränkt wurde; eine Modifikation von Glg. 5.33, in der statt des Ritzwertes θ ein konstanter Wert verwendet wurde; die Kombination mit Jacobi-Davidson (s.u.). In keinem Fall wurde eine signifikante Reduktion der Iterationsanzahl festgestellt. Da der Aufwand für diesen Präkonditionierer sehr hoch ist, ist die Rechenzeit entsprechend erheblich größer. In einigen Fällen konvergierten Verfahren gar nicht (d.h. nicht gegen die richtige Lösung). Da diese Untersuchungen keine Ansätze für eine Verfahrensverbesserung lieferten, wird auf eine weitere Darstellung verzichtet. Eine denkbare Erklärung liefert [MS86]. Demnach ist ein notwendiges Kriterium für einen guten Vorkonditionierer, daß die Eigenvektoren von

$$N(\theta) = M^{-1}(\theta)(\mathbf{H} - \theta \mathbf{1}) \quad (5.34)$$

sich nicht stark von denen von $\mathbf{H}$ unterscheiden. Dieses Kriterium jedoch wird hier im Regelfalle verletzt, da – angesichts der hohen Qualität der APSO-Energien – der auftretende kleine Nenner zu einer großen Komponente in Richtung der APSO-Vektoren führt.

5.2.4 Jacobi-Davidson Methode

Das Jacobi-Davidson (JD)-Verfahren [SvdV96] unterscheidet sich vom Davidson-Verfahren dadurch, daß der neue Korrekturvektor z im Orthogonalkomplement des

vorherigen Ritzvektors gesucht wird. An Stelle von Glg. 5.27 tritt dann:

$$(\mathbf{1} - vv^H)(H - \epsilon\mathbf{1})(\mathbf{1} - vv^H)z = -r \text{ und } z \perp v \quad (5.35)$$

Eine Ein-Schritt-Näherung erhält man vermöge

$$z = \alpha M^{-1}v - M^{-1}r, \quad (5.36)$$

wobei

$$\alpha = \frac{v^H M^{-1}r}{v^H M^{-1}v} \quad (5.37)$$

und M ein beliebiger Präkonditionierer ist. In [Are97] wurden schon Ergebnisse für die Implementierung dieses Verfahrens in unser Programm präsentiert; da diese Umsetzung jedoch vom oben beschriebenen Verfahren abwich, wurden nochmals Konvergenztests mit dem korrigierten Verfahren durchgeführt.

Numerische Tests

Numerische Versuche wurden mit drei Testsystemen durchgeführt, nämlich einer InP-Oberflächenzelle und zwei BN-Zellen. Die Indiumphosphid-Zelle ist relativ groß, kann aber durch Pseudopotentiale mit geringer Abschneideenergie (*Cutoff*) beschrieben werden. Dagegen benötigt man für die BN einen großen *Cutoff*, so daß die Anzahl Ebener Wellen trotz des geringen Volumens recht hoch ist. Außerdem ist es ein stark anisotropes Schichtsystem, was eine zusätzliche Schwierigkeit darstellt. Zusätzlich wurde auch eine dreifache BN-Zelle getestet, die sich aus drei übereinandergestapelten einfachen BN-Zellen zusammensetzt (BN12).

Das wesentliche Interesse ist auf das Konvergenzverhalten innerhalb des Ladungs-Selbstkonsistenzzyklus (SC) gerichtet, den man gerne beschleunigen möchte, da dies der häufigste Fall ist. Soweit jedoch eine Analyse der Unterschiede des Konvergenzverhaltens angestrebt wird, ist dieser Fall zu komplex; daher wurde auch das Verhalten für eine feste Matrix (d.h. bei konstant gehaltener Ladungsdichte, noSC) untersucht.

Dabei soll einerseits der Davidson-Algorithmus (Dvd) mit dem Jacobi-Davidson-Algorithmus (JD) und andererseits der Diagonale (Dia) mit dem (TAP)-Präkonditionierer verglichen werden.

Das Konvergenzverhalten wird von einer Vielzahl von Faktoren beeinflußt. Zunächst hängt es von dem betrachteten physikalischen System ab; dieses bestimmt u.a. die Größe der Basis aus Ebenen Wellen, die Anzahl und Anordnung der gesuchten Eigenwerte, insbesondere deren Clusterung oder Entartung. Das Verhalten kann auch für jeden k-Punkt anders sein.

Als algorithmische Parameter spielt die Wahl der Startlösungen eine Rolle, die Größe des Unterraumproblem (ndvdim) ist wichtig, und auch die Angabe, bei Unterschreiten welcher Residuumsnorm (epsdvd) ein Vektor als auskonvergiert betrachtet wird, hat unter Umständen starken Einfluß (jedoch vorwiegend im Fall ohne Selbstkonsistenz). Innerhalb einer Selbstkonsistenzschleife haben auch die Anzahl der Diagonalisierungsschritte innerhalb eines Selbstkonsistenzschrittes einen Einfluß; diese bestimmen indirekt die Stärke des Ladungsmischens.

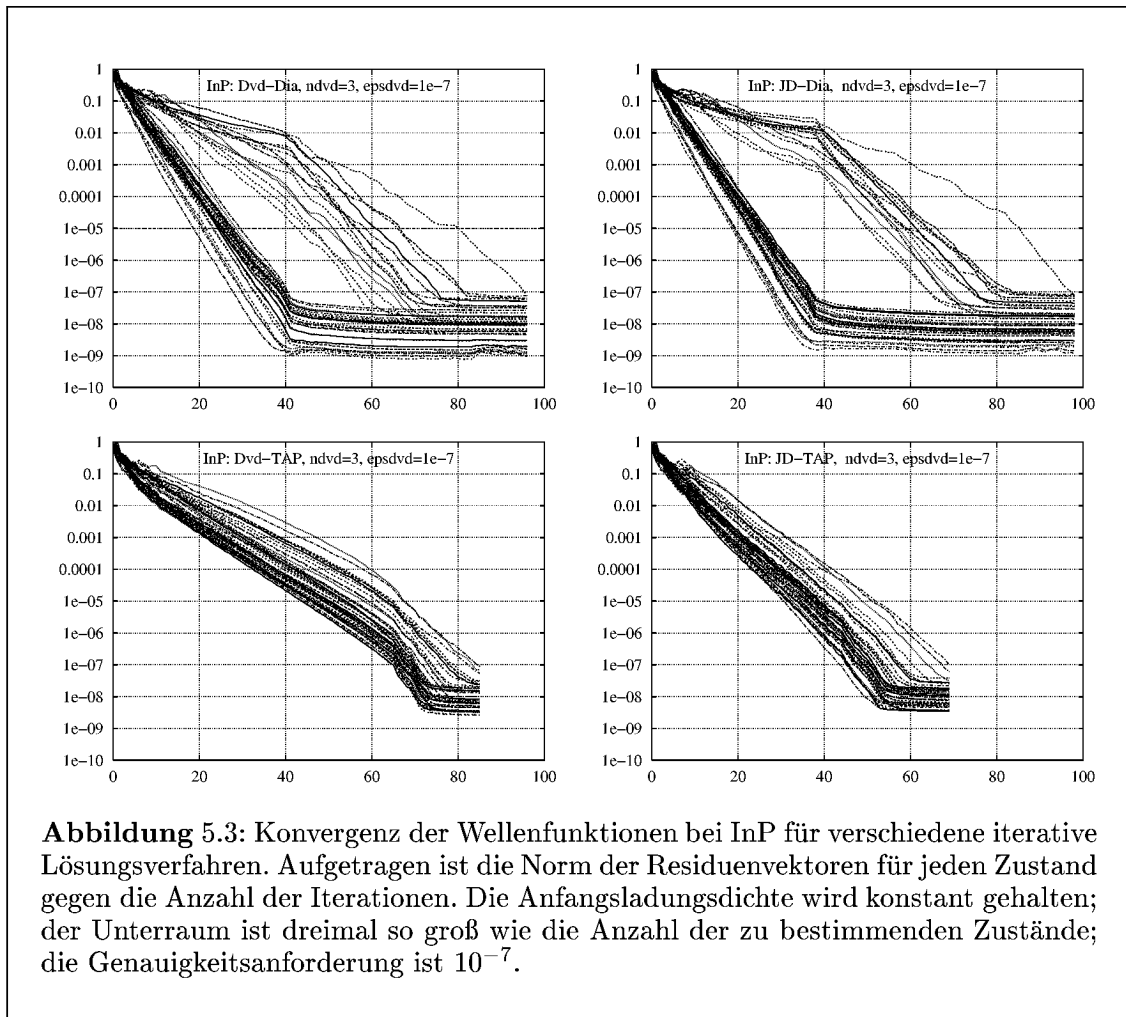
Schließlich zeigt sich auch der Einfluß der endlichen Rechengenauigkeit; bei hoher Optimierung weichen nach vielen Iterationen u.U. die Wege zur Lösung auch von Kompilation zu Kompilation voneinander ab. Das eigentliche Kriterium ist die Rechenzeit bis zur Konvergenz; gerade bei relativ kurzen Zeiten ist diese Angabe (für Multiuser-Systeme) jedoch relativ ungenau. Da die Rechenzeit pro Iterationsschritt zwischen den verschiedenen Verfahren nicht allzu stark voneinander abweicht, kann man (bei gleicher Unterraumgröße) auch die Gesamtzahl der Diagonalisierungsschritte vergleichen. Pro Iterationsschritt ist Dvd schneller als JD und Dia schneller als TAP.

Der TAP-Präkonditionierer wurde in zwei Varianten programmiert; bei der einen wird die kinetische Energie der Residuen benutzt, bei der anderen die des Ritzvektors. Beide Varianten zeigen in ersten Tests ähnliches Verhalten, doch ist die erstere stets etwas günstiger. Daher wurde letztere nicht weiter untersucht.

InP Zunächst wurde eine InP-Oberfläche untersucht. Die Zelle enthielt 14 Atome; für die 56 Elektronen wurden 28 besetzte und 14 unbesetzte Zustände für einen k-Punkt berechnet. Der *Cutoff* betrug $G_{max} = 3.5 \sqrt{Ry}$ und führte zu einer Basisgröße von ca. 3200 Ebenen Wellen. Die Startvektoren wurden in einem Unterraum von 400 Funktionen bestimmt. In selbstkonsistenten Rechnungen wurden je drei Davidsoniterationen pro SC-Schritt ausgeführt. Es wurden Tests für zwei Unterraumgrößen durchgeführt, nämlich zweimal bzw. dreimal die Anzahl der zu bestimmenden Zustände. Dieser Faktor wird mit *ndvd* (=2 bzw. 3) bezeichnet.

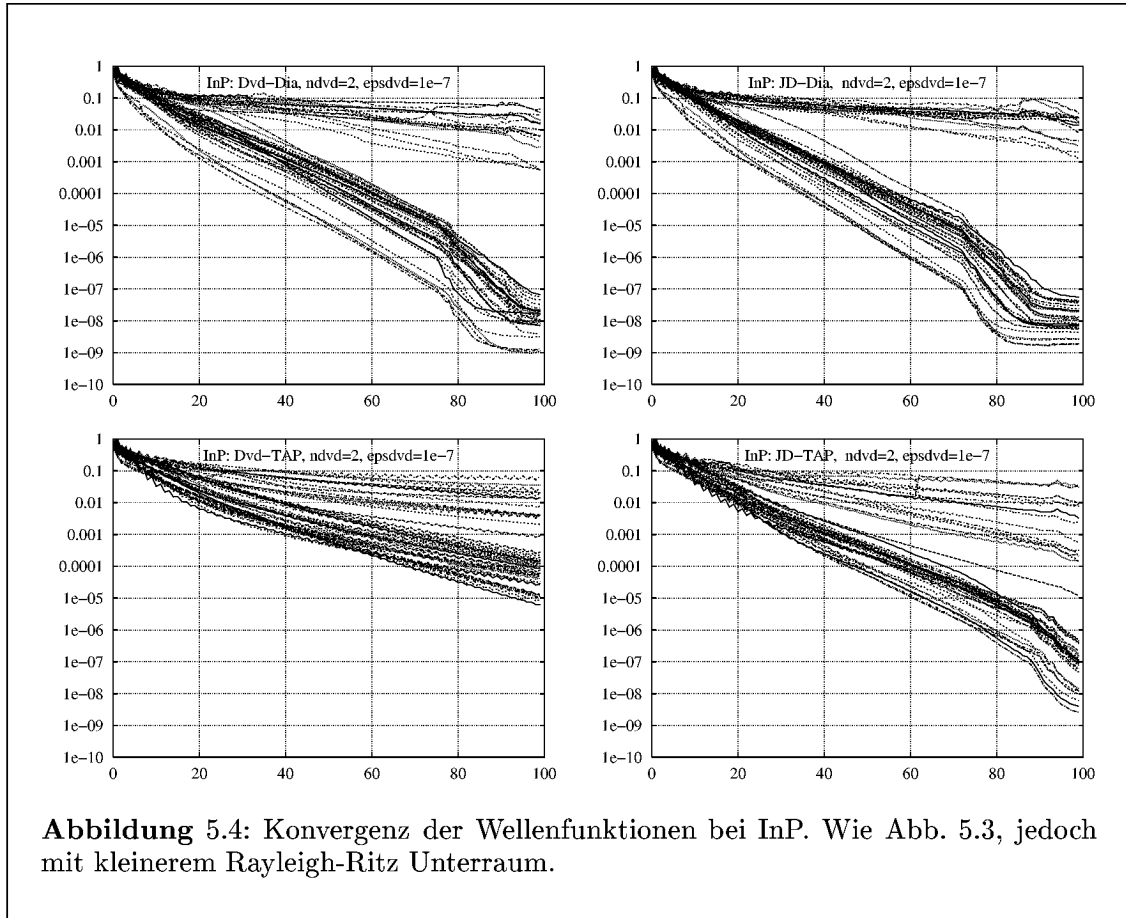
Abbildung 5.3 zeigt den Vergleich der Methoden für eine festgehaltene Anfangsladung. Die Größe des Davidson-Unterraumes ist *ndvd* = 3, das Abbruchkriterium ist *epsdvd* = 10^{-7} . Aufgetragen ist die absolute Größe der Residuenvektoren gegen den Iterationsindex.

Man erkennt, daß die beiden Verfahren mit diagonalem Präkonditionierer (Dvd-Dia und JD-Dia) ein ähnliches Verhalten zeigen insofern, als die besetzten Zustände relativ schnell und gleichmäßig konvergieren, wohingegen die unbesetzten ungleichmäßig



ger konvergieren. Besonders bei (JD-Dia) sieht man, daß die oberen Wellenfunktionen sich erst rapide verbessern, sobald die unteren das Genauigkeitsziel erreicht haben. Ein einzelner schlechter Zustand bestimmt die (etwa gleiche) Anzahl von Iterationen. Die (TAP)-Methoden hingegen zeigen für besetzte und unbesetzte Wellenfunktionen ein ähnliches Verhalten. Dadurch konvergiert der beste Zustand langsamer, der schlechteste jedoch schneller. Unter allen Verfahren benötigt hier (JD-TAP) die wenigsten Iterationen.

Abbildung 5.4 zeigt dieselbe Rechnung wiederholt mit kleinerer Unterraumgröße $ndvd = 2$. Es zeigt sich hier, daß die (TAP)-Methoden offenbar stärker vom größeren Unterraum profitieren. Während die (Dia)-Verfahren zumindest für die besetzten Zustände die angestrebte Präzision erreichen, sind die (TAP)-Varianten davon weit entfernt. Die Konvergenz der besetzten und der unbesetzten Zustände erfolgt nicht mehr so einheitlich, wie es bei $ndvd = 3$ der Fall war. Tabelle 5.2 zeigt einen Vergleich



der Rechenzeiten. (Dvd-Alt wird weiter unten eingeführt.)

Methode	$ndvd = 2$		$ndvd = 3$	
	Zeit	Iter.	Zeit	Iter.
Dvd-Dia	434	$\gg 100$	1001	97
JD-Dia	436	$\gg 100$	1019	99
Dvd-TAP	495	$\gg 100$	908	86
Dvd-Alt	465	$\gg 100$	855	82
JD-TAP	488	$\gg 100$	748	70

Tabelle 5.2: Rechenzeit und Anzahl der Iterationen für verschiedene iterative Methoden bei InP bei fester Ladungsdichte. (Es wurden maximal 100 Iterationen ausgeführt. Entsprechend ist die Rechenzeit aufzufassen.)

Da bei selbstkonsistenten Rechnungen nur die besetzten Zustände einen Einfluß auf die Ladungsdichte haben, erwartet man, daß (Dia)-Varianten im Vorteil sein

sollten. Tabelle 5.3 stellt die Ergebnisse zusammen. Für den größeren Unterraum zeigen sich keine Unterschiede in der Anzahl der Iterationen. Bei $ndvd = 2$ jedoch ist (JD-TAP) in weniger Schritten am Ziel. Da jedoch der TAP-Präkonditionierer teurer zu bestimmen ist, ist die Rechenzeit sogar etwas höher als für (Dvd-Dia).

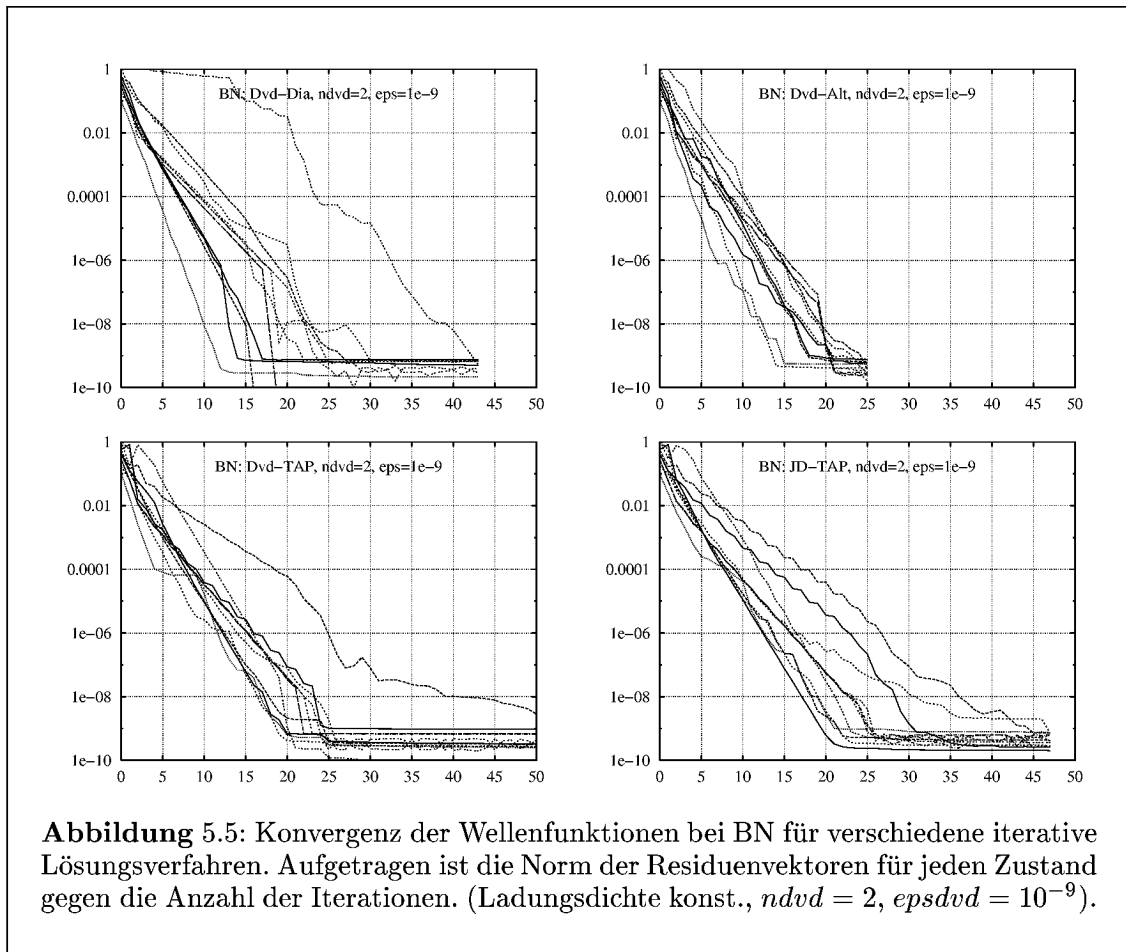
Methode	$ndvd = 2$		$ndvd = 3$	
	Zeit (sec)	Iter. (sc/dv)	Zeit (sec)	Iter. (sc/dv)
Dvd-Dia	292	19 , 60	379	11 , 36
JD-Dia	287	19 , 60	379	11 , 36
Dvd-TAP	361	21 , 66	392	11 , 36
Dvd-Alt	323	20 , 63	382	11 , 36
JD-TAP	298	17 , 54	391	11 , 36

Tabelle 5.3: Rechenzeit und Anzahl der Iterationen für verschiedene iterative Methoden bei InP bei selbstkonsistenter Bestimmung der Ladungsdichte.

BN Analoge Untersuchungen wurden für Bornitrid angestellt. Die Zelle enthält vier Atome und mithin 16 Elektronen. Es wurden 8 besetzte und 4 unbesetzte Zustände berechnet. Zur adäquaten Beschreibung ist ein relativ hoher *Cutoff* von $G_{max} = 6.0 \sqrt{\text{Ry}}$ erforderlich. Dies führt zu ca. 880 Basisfunktionen. Als Startlösungen wurden APSO-Funktionen verwendet. In selbstkonsistenten Rechnungen wurden je drei Davidsoniterationen pro SC-Schritt ausgeführt.

In Abbildung 5.5 ist das Konvergenzverhalten bei festgehaltener Anfangsladungsdichte gezeigt (bei $ndvd = 2$ und $epsdvd = 10^{-9}$). In diesem Falle konvergiert Dvd-Dia etwas schneller als JD-TAP. Die Konvergenz wird für Dvd-Dia wie auch für Dvd-TAP jeweils von einem einzelnen langsamen Zustand bestimmt. Es handelt sich jedoch um unterschiedliche Zustände. Diese Beobachtung inspirierte dazu, beim Davidson-Verfahren die beiden Präkonditionierer *alternierend* zu benutzen (Dvd-Alt). Wie man in der Abbildung sieht, führt dieses Verfahren in der Tat beinahe zu einer Halbierung der Anzahl an Iterationen. Bei einer Genauigkeit von $epsdvd = 10^{-7}$ ergibt sich qualitativ dasselbe Bild, aber der Gewinn durch Dvd-Alt ist geringer. (Bei InP wurden die Ergebnisse für Dvd-Alt schon unkommentiert aufgeführt; sie sind ohne Selbstkonsistenz gut, mit dagegen eher mäßig.)

Abbildung 5.6 zeigt die Konvergenz für einen größeren Unterraum, $ndvd = 3$. Man sieht, daß sich für keines der Verfahren eine dramatische Verbesserung ergibt (im Gegensatz zum InP-System). Dies ist sehr erstaunlich, da man eigentlich mit zuneh-



mender Unterraumdimension eine schnellere Konvergenz (gemessen an der Anzahl der Iterationen) erwartet; der Grund für dieses Verhalten wurde nicht gefunden. Allerdings finden sich in der Literatur [CPS94] (wenn auch für die ungeblockte Variante des Algorithmus) Hinweise darauf, daß bei abnehmender Konvergenzgeschwindigkeit eine Reduktion der Unterraumgröße (*restart*) sinnvoll ist.

Dvd-Dia und JD-TAP sind wiederum etwa gleichauf bzgl. der Anzahl der Iterationen (aber Dvd-Dia ist besser bzgl. der Rechenzeit). Von den drei Zuständen, die bei Dvd-Dia langsam konvergieren, finden sich diesmal auch die beiden, welche bei Dvd-TAP schlecht sind. Dies erklärt, warum hier Dvd-Alt keine Verbesserung bringt; es führt sogar zum Gegenteil. Bei Dvd-TAP fällt auf, daß ein Vektor knapp vor Erreichen der angestrebten Genauigkeit stagniert. Die Ursache hierfür ist unklar.⁸

⁸Es sei an dieser Stelle noch vermerkt, daß diese Tests mit einer veränderten Programmversion ausgeführt wurden. Insbesondere wurde vorher die Unterraumdiagonalisierung nicht mit höchstmöglicher Genauigkeit durchgeführt, und es wurden bei der Rücktransformation Matrixelemente unterhalb eines Schwellenwertes ignoriert. Auch wird die Rücktransformation nun anders

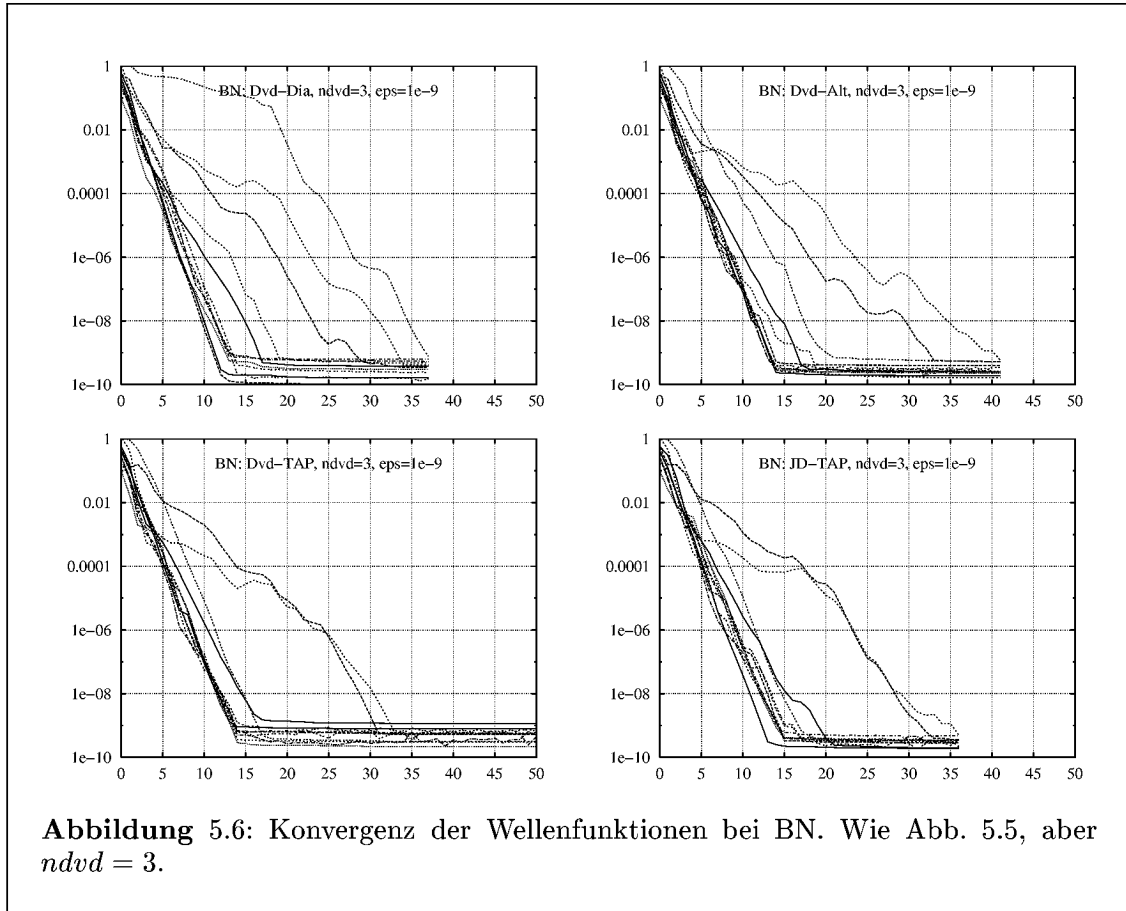


Abbildung 5.6: Konvergenz der Wellenfunktionen bei BN. Wie Abb. 5.5, aber $ndvd = 3$.

Methode	$epsdvd = 10^{-7}$				$epsdvd = 10^{-9}$			
	$ndvd = 2$		$ndvd = 3$		$ndvd = 2$		$ndvd = 3$	
	Zeit (sec)	Iter.	Zeit	Iter.	Zeit	Iter.	Zeit	Iter.
Dvd-Dia	6.4	26	9.2	30	7.8	44	10.6	38
Dvd-TAP	6.9	30	9.4	29	13.5	$\gg 100$	23.7	$\gg 100$
Dvd-Alt	6.1	21	9.5	30	6.5	26	11.8	42
JD-TAP	6.4	26	9.2	28	8.4	48	10.9	37

Tabelle 5.4: Konvergenz der Wellenfunktionen für BN bei fester Ladungsdichte.

Die Resultate sind in Tabelle 5.4 zusammengefaßt. Es ergibt sich hier gegenüber (nämlich per BLAS3) ausgeführt. Bei denselben Tests *vor* diesen Änderungen trat häufiger eine derartige Stagnation auf; z.T. geschah dies auch bei größeren Werten von $epsdvd$, und zwar stagnierten einzelne Vektoren mit Residuen, deren Norm deutlich größer waren. Allein bei Dvd-Dia wurde derlei nicht beobachtet. Diese Effekte lassen sich mit dem jetzigen Programm nicht mehr reproduzieren.

Dvd-Dia insgesamt kein Gewinn durch die neuen Methoden.

Auch für dieses System wurde nun eine selbstkonsistente Rechnung angestellt. Das dabei auftretende Konvergenzverhalten ist in Tabelle 5.5 summiert. Es muß allerdings dazu bemerkt werden, daß hier für alle Verfahren mit $ndvd = 2$ die beiden oberen Eigenwerte falsch bestimmt werden. (Die APSO-Startfunktionen liefern hier gute Eigenwerte für die besetzten Zustände, aber die für die unbesetzten sind nicht so gut und die zugehörigen Vektoren sind offenbar näher an den Eigenvektoren für geringfügig höhere Energien – 1.28 und 1.30 statt 1.21 und 1.22)

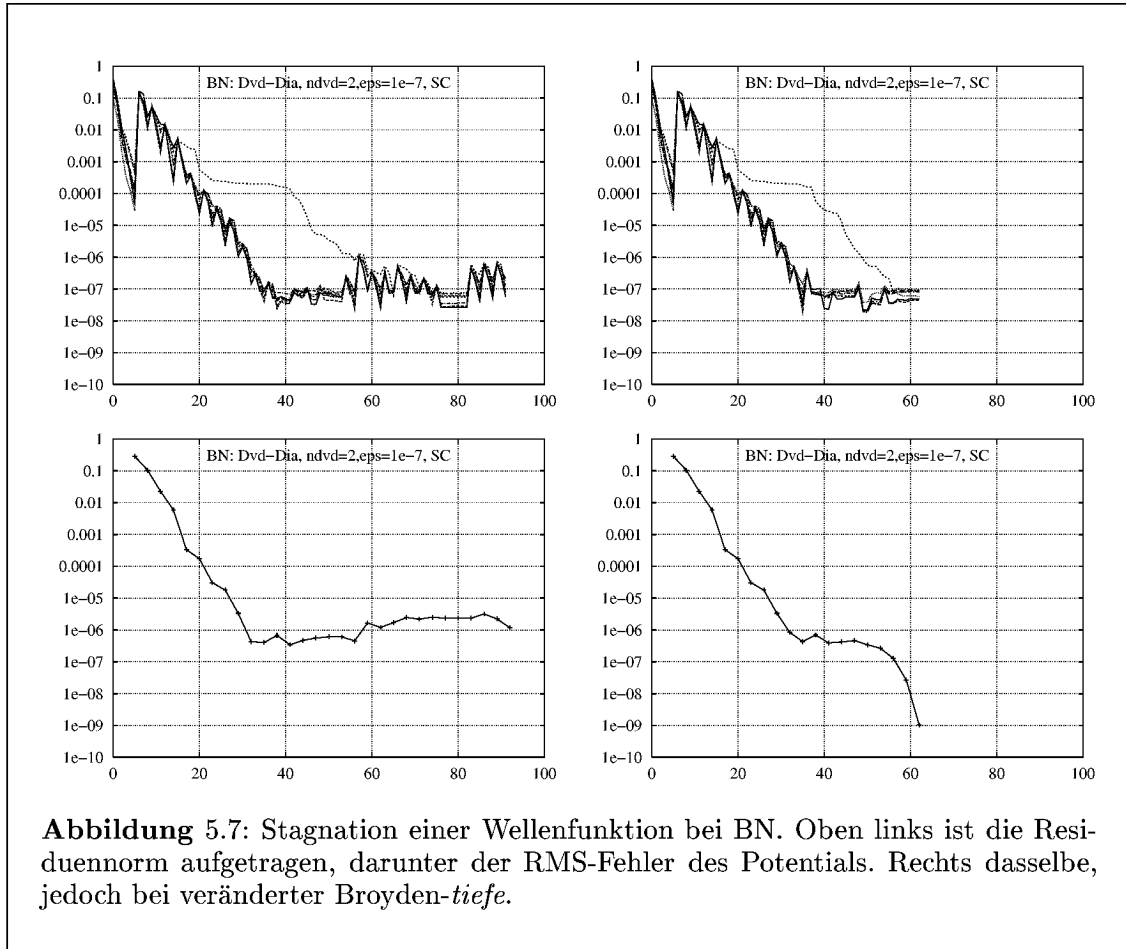
Auffallend ist hier, daß Dvd-Dia besonders für $ndvd = 3$ schlecht abschneidet; die übrigen Verfahren sind fast gleich schnell. Die Anzahl der Selbstkonsistenziterationen wird erwartungsgemäß durch den größeren Unterraum nicht stark reduziert, so daß die Rechenzeit dafür insgesamt höher ist.

Methode	$ndvd = 2$		$ndvd = 3$	
	Zeit (sec)	Iter. (sc/dv)	Zeit (sec)	Iter. (sc/dv)
Dvd-Dia	11.1	16, 51	24.8	$\gg 30, 92$
Dvd-TAP	10.1	13, 42	12.4	11, 36
Dvd-Alt	9.9	13, 41	13.0	12, 39
JD-TAP	9.9	13, 41	13.2	12, 39

Tabelle 5.5: Konvergenz für BN bei selbstkonsistenter Bestimmung der Ladungsdichte. Angegeben sind Rechenzeit und Anzahl der Selbstkonsistenziterationen (sc) sowie die Summe der Davidsonschrte.

Das ungewöhnliche Konvergenzverhalten für Dvd-Dia soll nun näher untersucht werden. In Abbildung 5.7 ist die Konvergenz der besetzten Zustände gegen die Anzahl der Iterationsschritte aufgetragen (links oben). Ein einzelner Zustand stagniert über längere Zeit hinweg (ca. Schritt 20 - 40). Danach setzt die Konvergenz wieder ein. Es handelt sich dabei um einen nicht-entarteten, wohlseparierten Eigenwert. Links unten ist der RMS-Fehler des Potentials aufgetragen. Man erkennt, daß sich dieser nicht wesentlich verbessert, nachdem die Konvergenz des Zustandes eintritt. Dies liegt daran, daß das Broyden-Verfahren nach so vielen Schritten bei einem nahezu unveränderten Potential das Minimum ganz nahe wähnt und daher die erforderliche nachhaltige Korrektur nicht mehr zuläßt.

Auf der rechten Seite der Abbildung sind die Daten für dieselbe Berechnung wiederholt, jedoch diesmal mit einer Maximalanzahl von 8 statt 30 SC-Iterationen für das Broyden-Mischen der Potentiale. Das führt dazu, daß das Broyden-Verfahren



nach 8 SC-Schritten (d.h. bei Iteration 26 und 50 in den Abbildungen) – also kurz nach Ende der Stagnation – die vorherige Konvergenzgeschichte vergißt. Dadurch erlaubt das Mixing in diesem Falle eine relativ starke Änderung des Potentials, mit dem Effekt, daß der RMS-Fehler des Potentials rapide sinkt und binnen weniger Iterationen die Konvergenz erzielt wird.

BN12 Um zu prüfen, inwiefern die Ergebnisse durch die physikalischen Gegebenheiten bedingt oder eher algorithmischen Ursprungs sind, wurden die gleichen Versuche mit einer Superzelle durchgeführt, die aus drei übereinander gestapelten BN-Zellen des obigen Typs besteht. Es ergeben sich dadurch Tripel von jeweils fast entarteten Eigenwerten. Es wurden 24 besetzte und 12 unbesetzte Zustände berechnet. Die Basis umfaßt hier ca 2670 Ebene Wellen.

Die Ergebnisse sind in Tabellen 5.7 und 5.6 zusammengefaßt.

Im Falle festgehaltener Startladung ist Dvd-Dia sehr schlecht für $ndvd = 2$, jedoch

Methode	$epsdvd = 10^{-7}$				$epsdvd = 10^{-9}$			
	$ndvd = 2$		$ndvd = 3$		$ndvd = 2$		$ndvd = 3$	
	Zeit (sec)	Iter.	Zeit	Iter.	Zeit	Iter.	Zeit	Iter.
Dvd-Dia	199	90	185	34	221	$\gg 100$	262	49
Dvd-TAP	124	48	199	32	259	$\gg 100$	268	42
Dvd-Alt	68	26	158	27	96	38	248	42
JD-TAP	112	43	199	32	193	76	247	39

Tabelle 5.6: Konvergenz der Wellenfunktionen für BN12 bei fester Ladungsdichte.

gut für $ndvd = 3$. Dvd-TAP erreicht wiederum nicht eine Genauigkeit von 10^{-9} . Dvd-Alt führt sehr schnell zur Konvergenz. JD-TAP bleibt bei kleinem Unterraum weit dahinter zurück. Ein großer Unterraum bringt hier generell keinen Gewinn.

Methode	$ndvd = 2$		$ndvd = 3$	
	Zeit (sec)	Iter. (sc/dv)	Zeit (sec)	Iter. (sc/dv)
Dvd-Dia	220	$\gg 30, 93$	253	14, 45
Dvd-TAP	123	14, 45	237	12, 39
Dvd-Alt	121	14, 45	241	13, 42
JD-TAP	124	14, 45	254	13, 42

Tabelle 5.7: Konvergenz für BN12 bei selbstkonsistenter Bestimmung der Ladungsdichte. Angegeben sind Rechenzeit und Anzahl der Selbstkonsistenziterationen (sc) sowie die Summe der Davidsonschrte.

Auch im selbstkonsistenten Fall ist Dvd-Dia dieses Mal schlecht für $ndvd = 2$, jedoch besser für $ndvd = 3$; bei BN war es andersherum. Da man auf Grund der physikalischen Ähnlichkeit der Systeme ein ebenso ähnliches Verhalten erwartet hätte, zeigt sich hier eine gewisse Unvorhersehbarkeit des Verhaltens, das eine schlüssige Ursachenanalyse vereitelt. Die anderen Varianten untereinander sind für $ndvd = 2$ etwa gleich schnell, so wie man es ausgehend von BN erwartet hat.

Zusammenfassung und Ausblick Zunächst muß man festhalten, daß sich kein einheitliches Bild ergibt. Es kann nicht die grundsätzliche Überlegenheit eines bestimmten Verfahrens belegt werden – weder im Vergleich von Dvd und JD noch in dem der Präkonditionierer Dia und TAP.

Es stellt sich die Frage, wie man überhaupt ein oder mehrere Testsysteme bestimmen

soll, deren Verhalten man als repräsentativ erachten kann. Wenn man von den hier betrachteten Testsystemen ausgeht, so bleiben Dvd-Dia, Dvd-Alt und JD-TAP in engerer Wahl.

Für Dvd-Dia spricht unter anderem – trotz der hier aufgetretenen Unregelmäßigkeiten – die größere Erfahrung mit diesem Verfahren. Das JD-TAP-Verfahren bringt in einigen Fällen (speziell für die Berechnung *aller* Eigenwerte) Vorteile. Eine sensationelle Verbesserung stellt es jedoch nicht dar.

Eine mögliche Erklärung dafür, daß die Vorteile recht gering sind, könnte sein, daß der oben beschriebene JD-Algorithmus für einen einzelnen Eigenvektor entworfen ist. In unserem Programm ist jedoch ein Block-Verfahren implementiert, wobei für jeden Korrekturvektor *für sich* das obige Verfahren angewandt wird. In [BDD⁺99] wird ein JD-Verfahren für mehrere Eigenwerte beschrieben; diese werden jedoch konsekutiv bestimmt. Die Betrachtung dieses Algorithmus' legt die Vermutung nahe, daß in einem Block-Verfahren der Raum, in dem die neuen Korrekturvektoren bestimmt werden, nicht nur orthogonal zum *jeweiligen* vorherigen Ritzvektor sein sollte, sondern zum Raum *aller* solcher Vektoren. Damit nähme Glg. (5.35) die Form

$$(\mathbf{1} - VV^H)(H - \epsilon\mathbf{1})(\mathbf{1} - VV^H)z_i = -r_i \text{ und } z \perp V \quad (5.38)$$

an, wobei sich die Matrix V aus allen Ritzvektoren v_j zusammensetzt.

Im Gegensatz zu Glg. (5.35) läßt sich hierfür jedoch kein Verfahren angeben, um eine Lösung ohne großen Aufwand anzugeben. Vielmehr muß man (wiederum) zu einem iterativen Verfahren greifen, um die Bestimmungsgleichung für den neuen Korrekturvektor approximativ zu erfüllen.

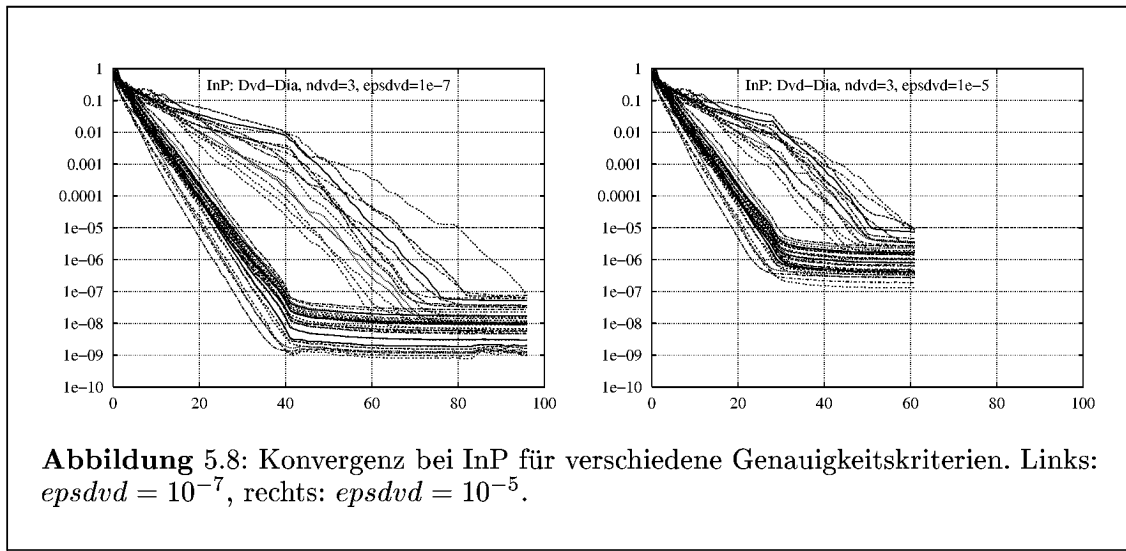
Bei der Dvd-TAP-Methode erweckt vor allem die gelegentliche Stagnation aus ungeklärter Ursache Mißtrauen.

Die Dvd-Alt-Methode ist in einigen Fällen zwar erheblich besser als alle übrigen, bei anderen jedoch auch etwas schlechter. Wichtig erscheint mir jedoch die Beobachtung, daß es (insbesondere für Dvd-Dia) oft nur einen oder einige wenige Vektoren gibt, die die schlechte Konvergenz verursachen. Der teilweise verblüffende Erfolg von Dvd-Alt zeigt, daß es wahrscheinlich sinnvoll ist, dafür zu sorgen, daß kein (besetzter) Zustand weit hinter den übrigen zurückbleibt.

Eine Möglichkeit ist die mit Dvd-Alt eingeschlagene, nämlich zu versuchen, das Auftreten dieses Phänomens von vornherein gänzlich zu verhindern. Dies gelingt freilich nicht immer.

Eine andere wäre, das Konvergenzverhalten aller (besetzten) Zustände separat zu überwachen und gegebenenfalls besondere Maßnahmen zu ergreifen, um speziell für

einzelne Problemvektoren bessere oder eine größere Anzahl von Korrekturvektoren zu erzeugen, beispielsweise mit verschiedenen Präkonditionierern generiert. Denkbar wäre auch, APSO-Vektoren für den entsprechenden Energiebereich hinzuzufügen. Eine dritte Möglichkeit ist, das Konvergenzkriterium der Eigenvektoren dynamisch anzupassen. Da zunächst wenige Vektoren bis zur Genauigkeitsgrenze konvergieren und die übrigen erst dann besser konvergieren, sollte dies automatisch die Gleichmäßigkeit der Konvergenz erhöhen.



In Abbildung 5.8 sieht man, daß bei $\text{epsdvd} = 10^{-5}$ nach 60 Iterationen alle Vektoren unter 10^{-5} konvergiert sind, während dies bei $\text{epsdvd} = 10^{-7}$ erst nach 80 Iterationen der Fall ist. Man könnte das Kriterium für die besten Vektoren relativ zu den schlechtesten setzen, z.B. epsdvd zwei Größenordnungen kleiner als die aktuelle Größe des Residuums des schlechtesten (besetzten) Zustandes. Es wäre auch möglich, die Genauigkeitsanforderung der aktuellen Präzision der Ladungsdichtekonvergenz anzupassen. Dies würde auch für eine gleichmäßige Konvergenz über alle k-Punkte sorgen.

Wichtig ist es, das Auftreten von Sackgassen durch das unglückliche Zusammenspiel von Ladungs- oder Potentialmischen und Stagnation des iterativen Eigenwertlösers zu vermeiden, wie es oben erörtert wurde. Auch dies ließe sich durch eine der Ladungskonvergenz dynamisch angepaßte Genauigkeit der Eigenvektorbestimmung anstelle einer fest vorgegebenen Anzahl von Iterationen pro Selbstkonsistenzschritt gewährleisten. Dies erfordert eine ausgiebige Phase der Tests und der Feinjustage. Eine dahingehende Ausweitung der hiesigen Investigationen könnte die Konvergenz des Verfahrens weiter verbessern.

Kapitel 6

Test der PAW-Methode

6.1 Kohlenstoffgitter

Als erstes Beispiel für ein $2p$ -Element wurde ein Pseudopotential für Kohlenstoff entwickelt, welches für Abstände geeignet ist, wie sie in reinen Kohlenstoffverbindungen auftreten. Das benutzte C_m -Pseudopotential wurde in Kapitel 4.5 eingehend beschrieben.

Um zu testen, welche Genauigkeit sich mit dem PAW-Programm erzielen läßt, wurden verschiedene existierende und hypothetische Gitter von Kohlenstoff berechnet. Zunächst wurden verschiedene kubische Gittervarianten untersucht, nämlich das einfach kubische Gitter (sc), das fcc-Gitter, das bcc-Gitter sowie das Diamantgitter (dia).

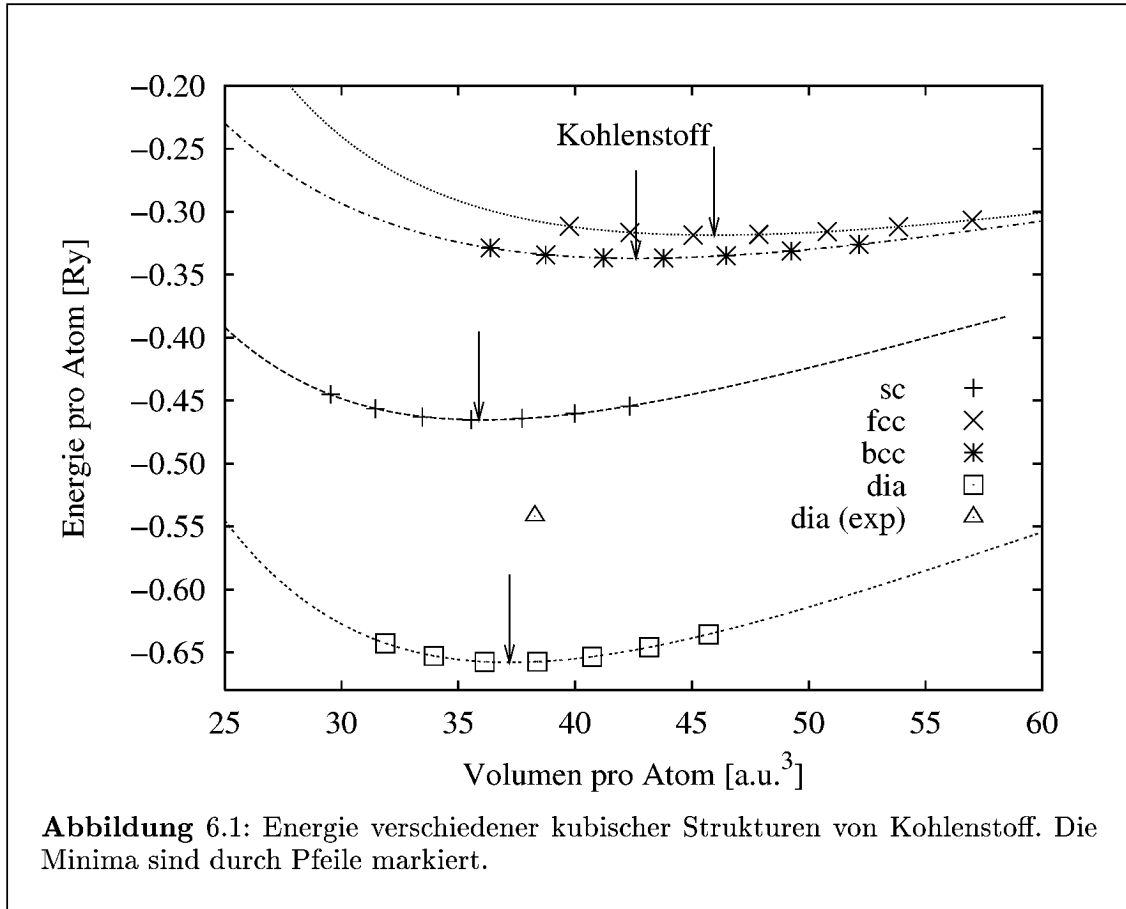
Es wurde jeweils die Gitterkonstante im Bereich um das Minimum variiert und durch Anpassen an Murnaghans Zustandsgleichung das Minimum und der *bulk modulus* bestimmt.

Als Bezugsenergie diente die Energie eines Kohlenstoffatoms in einer kubischen Zelle der Kantenlänge $16 \text{ \AA}$. Diese betrug $E_0 = -17.069995 \text{ Ry}$.¹ Zusätzlich berücksichtigt wurde die Spin-Polarisationsenergie des Atomes, welche eine Energieabsenkung gemäß Hund's Regel von 0.0875 Ry beiträgt.

In Abbildung 6.1 sind die Gesamtenergien (pro Atom) gegen das Volumen pro Atom aufgetragen. In qualitativer Übereinstimmung mit existierenden Berechnungen steigt die Energie der Konfigurationen mit der Anzahl nächster Nachbarn an. Für Diamant zeigt der Vergleich mit dem experimentellen Wert – wie für die LDA üblich – ei-

¹Die bei der Berechnung des Pseudoatoms berechnete Energie läßt sich nur dann nutzen, wenn man dasselbe radiale Gitter benutzt und nicht – wie hier geschehen – auf ein anderes interpoliert.

ne leichte Unterschätzung der Bindungslänge, die Kohäsionsenergie hingegen wird überschätzt.



Den quantitativen Vergleich der strukturellen Daten mit Literaturwerten zeigt Tabelle 6.1. Für Diamant ist die Übereinstimmung der Gitterkonstanten und des *bulk modulus* mit Ergebnissen anderer *ab-initio*-Berechnungen vorzüglich. Die Kohäsionsenergie ist – wenn man in Rechnung stellt, daß bei einigen der angeführten Resultate die Spin-Polarisationsenergie nicht enthalten ist – ebenfalls hervorragend.

Für die hypothetischen Gitter (sc, fcc, bcc) stimmen die Kohäsionsenergien und die Gitterkonstanten bis auf 1-2% mit denen aus [ERBS] überein. Die Abweichung des *bulk modulus* ist mit 3% (bcc) bis 27% (fcc) deutlich größer. Die Energiedifferenzen zur Diamantstruktur zeigt Tabelle 6.2. Diese stimmen gut mit den Resultaten anderer Berechnungen überein.

Während alle obigen Strukturen isotrop in alle Hauptachsenrichtungen sind, handelt es sich bei Graphit um ein stark anisotropes Schichtsystem. Die Atome innerhalb dieser Schicht bilden ein Honigwabengitter mit sehr kurzen Abständen, während der

Struktur Methode		Gitterkonstante a [a.u.]	<i>bulk modulus</i> B [Mbar] $\frac{\partial B}{\partial p}$		Kohäsionsenergie E_{coh} [eV]	
Einfach kubisch						
Diese Arbeit	PAW	3.30	3.64		6.33	
[ERBS]	NC-PP	3.29	3.98		6.40	
[YC83]	NC-PP	3.34				
fcc-Gitter						
Diese Arbeit	PAW	5.69	1.88		4.33	
[ERBS]	NC-PP	5.79	1.48		4.27	
[YC83]	NC-PP	5.82				
bcc-Gitter						
Diese Arbeit	PAW	4.40	2.06		4.58	
[ERBS]	NC-PP	4.44	2.13		4.66	
[YC83]	NC-PP	4.49				
Diamant						
Diese Arbeit	PAW	6.68	4.62	3.71	8.95	10.14
[ERBS]	NC-PP	6.69	4.50	3.68	8.88	
[YC83]	NC-PP	6.71				
[LB63]	exp.	6.74	4.60		7.37	
[KH94]	US-PP	6.67	4.60		9.03	
[FKBK00]	US-PP	6.68	4.62	3.62	9.00	
[KJ99]	US-PP	6.68	4.61			10.12
[KJ99]	PAW	6.68	4.60			10.11
[HMD ⁺ 97]	PAW	6.69	4.60			10.16
[PCV ⁺ 92]	LAPW	6.69	4.70			10.13

Tabelle 6.1: Strukturelle Eigenschaften von Diamant und anderen kubischen Kohlenstoffkristallen. *US-PP* steht für Ultrasofte, *NC-PP* für normerhaltende Pseudopotentiale. Die nach rechts abgesetzten Kohäsionsenergien sind ohne Spin-Polarisationsenergie berechnet. Diese betrug in unserer Rechnung 1.19 eV.

Abstand der Schichten 2.4 mal so groß ist [AM76]. Zwischen diesen Schichten findet vorwiegend eine Van-der-Waals-Wechselwirkung statt.

Für die hexagonale Zelle wurde ein k-Punkt-Satz von $15 \times 15 \times 5$ Punkten in der gesamten Brillouinzone verwendet, was zu 81 Punkten in deren irreduziblen Teil führt. Um einen Fehler durch unterschiedliche Gitterzellen auszuschließen, wurde

Struktur	Energiedifferenz zu Diamant [eV]		
	PAW	[ERBS]	[YC83]
sc	2.62	2.48	2.66
fcc	4.62	4.61	4.59
bcc	4.37	4.21	4.28

Tabelle 6.2: Energiedifferenz zur Diamantstruktur für einfach kubisches (sc), fcc- und bcc-Gitter von Kohlenstoff. Alle Werte in eV.

die Berechnung für Diamant in einer hexagonalen Zelle wiederholt (c-Richtung der hexagonalen Zelle ist die 111-Richtung der kubischen Zelle). Die Ergebnisse sind (in der angegebenen Genauigkeit) identisch mit denen in Tabelle 6.1 für die kubische Zelle.

Für die Berechnung von Graphit wurden die Gitterkonstanten a und c unabhängig voneinander variiert.

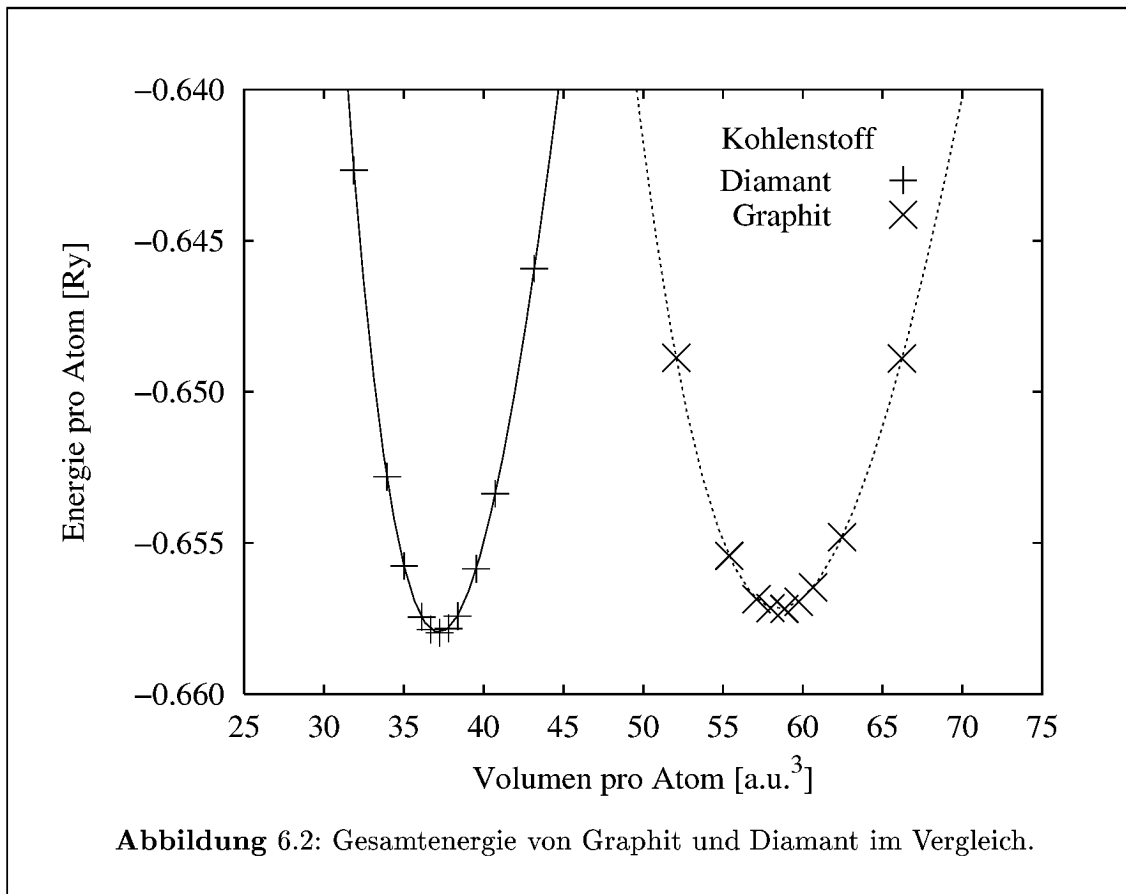


Abbildung 6.2 zeigt die Gesamtenergie (pro Atom) von Graphit in Abhängigkeit

vom atomaren Volumen bei festgehaltenem c/a -Verhältnis. Es wurde der c/a -Wert verwendet, der bei minimaler Gesamtenergie bestimmt wurde.

Die Variation der Energie mit dem Schichtabstand c ist sehr gering. Bei minimaler Gitterkonstante a ändert sich die Gesamtenergie um weniger als 0.1 mRy, wenn man c um $\pm 3\%$ um den Minimumswert variiert.

Man erkennt, daß die Gesamtenergie von Diamant als niedriger berechnet wird denn die von Graphit, was bekanntlich verkehrt ist. Der Unterschied ist jedoch sehr klein; nach einer Untersuchung von Yin und Cohen [YC84] kann das Vorzeichen dieser Energie davon anhängen, welches Austausch-Korrelations-Potential man verwendet. Die berechnete Energie bewegt sich im Rahmen der dort ermittelten Werte.

Die Ursache für die falsche Beschreibung der Gesamtenergie dürfte darin liegen, daß die LDA unterschiedliche Fehler für unterschiedliche Bindungstypen zuläßt. Offensichtlich ist der Fehler bei den kovalenten Bindungen in der Graphitebene und den van-der-Waals-Bindungen zwischen den Ebenen sehr verschieden, was dann einen Fehler von der Größenordnung des Kohäsionsenergieunterschiedes bewirkt. Wir betrachten es als herausfordernde Aufgabe für die Zukunft, Funktionale zu ertwickeln, die auch für diese Systeme Vorhersagekraft besitzen. In der Arbeit von Furthmüller et. al. [FKBK00] liefert eine Rechnung mit einem sehr weichen Ultrasoften Potential ebenfalls das Resultat, daß Graphit die instabile Modifikation sei. Der dort bestimmte Energieunterschied ist jedoch mit 60 meV um einen Faktor einhundert größer als der hier bestimmte. Für das dortige *Referenzpotential* geben die Autoren keinen Zahlenwert an, sondern sprechen nur von „virtually identical energies“. Sie führen den Fehler ebenfalls auf die unzulängliche Beschreibung der Van-der-Waals-Wechselwirkung zwischen den Ebenen zurück, da die Berechnung einer Graphitmonolage gute Ergebnisse liefert.

Tabelle 6.3 zeigt jedoch, daß die in dieser Arbeit ermittelten Werte für die Gitterkonstanten sehr gut mit denen in [FKBK00] übereinstimmen, welche mit Ultrasoften Pseudopotentialen berechnet wurden.

6.2 Molekularer Stickstoff

Als weiterer Test für die Genauigkeit der Methode wurden für ein Stickstoffmolekül N_2 Bindungslänge und Schwingungsfrequenz bestimmt. Auf Grund der geringen Bindungslänge und der im Vergleich zum Festkörper stark ausgeprägten Asymmetrie der Umgebung jedes einzelnen Atomes sollte dies ein guter Test für die Qualität des erstellten PAW-Programmes sein.

		Gitterkonstanten		Energiedifferenz
		a [a.u.]	c/a	$E_{dia} - E_{graph}$ [mRy]
Diese Arbeit	PAW	4.62	2.74	-0.7
[FKBK00] ^a	US-PP	4.61	2.73	„virtually identical“
[YC84]	NC-PP	4.67	2.73 ^b	+1 ··· - 3
	exp.	4.65 ^c	2.73 ^c	+1.8 ^d

Tabelle 6.3: Strukturelle Daten für Graphit. Negatives Vorzeichen in der Energiedifferenz bedeutet, daß Diamant stabiler ist. (US-PP: Ultrasofte Pseudopotentiale; NC-PP: normerhaltende Pseudopotentiale. a: Werte des *Referenzpotentials* in [FKBK00]. b: Der experimentelle c/a-Wert war vorgegeben in [YC84]; Energiewerte für verschiedene Austausch-Korrelations-Potentiale. c: [LB63]; d: Nach [FKBK00].)

Der Test wurde in einer tetragonalen Zelle der Größe $11 \times 11 \times 12.4 \text{ \AA}^3$ durchgeführt. Bei dieser Größe sinkt die Dichte weit entfernt von den Atomen auf einen Plateauwert ab, welcher um mehr als acht Größenordnungen kleiner ist als die Dichte zwischen den Atomen. Die Atome „sehen“ also die „Schwänze“ der Dichten der Atome in benachbarten Superzellen nicht mehr. (Dies stimmt auch etwa überein mit der Größe der Zelle, die Blöchl für entsprechende Tests benutzt hat [Blö94].)

Für das in Kapitel 4.6 beschriebene N_s -Potential wurde eine sehr hohe Abschneideenergie von 75 Ry benutzt. Dies ist sicherlich nicht zur Anwendung im größeren Maßstab gedacht, sondern dient lediglich zu Referenzzwecken.

Es wurden $4 \times 4 \times 4$ k-Punkte in der gesamten Brillouinzone benutzt; erwartungsgemäß findet man aber fast keine Dispersion. Die Bindungslänge wurde zunächst grob bestimmt; dann wurde für sieben Abstände im Bereich von $\pm 1.5\%$ um den erwarteten Gleichgewichtswert die Energie bestimmt. Die gewonnenen Werte wurden an die Universelle Bindungsenergierektion (UBER) nach [BS88] angepaßt. Mit einem kubischen Fit erhält man fast dieselben Werte; die Wahl des Fits ist hier also wohl (bei hinreichend klein gewähltem Intervall der Bindungslänge) unkritisch.

Aus der Fitfunktion wurden dann die minimale Energie, die Bindungslänge und die „Federkonstante“ D als zweite Ableitung der Energie nach der Bindungslänge bestimmt. Aus dieser erhält man die Wellenzahl zur Charakterisierung der Schwingungsfrequenz.

Die Ergebnisse zeigt Tabelle 6.4. Die Bindungslänge stimmt hervorragend überein mit den AE-Werten von Dickson und Becke [DB93], experimentell bestimmten Bindungslängen sowie denjenigen berechneten Werten, bei denen höchste Präzision angestrebt wurde. Kresse und Joubert [KJ99] bestimmten jeweils einen Referenzwert

	Methode	Bindungslänge a [a.u.]	Dissoziationsenergie E_D [eV]	Schwingungsfrequenz k [cm ⁻¹]
Diese Arbeit	PAW	2.068	11.5	2384
[DB93]	NUMOL	2.068	11.47	
[SW93]	APW	2.07	11.3	2380
[KJ99]	PAW	2.068 (2.076)		
	US-PP	2.066 (2.077)		
[Blö94]	PAW	2.09	11.38	2417
experimentell ^a		2.07	9.9	2360

Tabelle 6.4: Vergleich von Bindungslänge, Dissoziationsenergie und Schwingungsfrequenz eines Stickstoffmoleküls N₂. Werte in Klammern bei [KJ99] beziehen sich jeweils auf einen zweiten, weichen Datensatz. Die Ergebnisse von [Blö94] sind mit einem sehr weichen Datensatz erzielt. a: zitiert nach [SW93].

und einen Wert mit einem weichen Potential. Für letzteres wird der Bindungsabstand um 0.01 Å länger. In diesem Sinne läßt sich auch die Abweichung des Wertes von Blöchl [Blö94] deuten, da dies kein Referenzwert ist, sondern mit einem weichen PAW-Potential bei einer Abschneideenergie von 30 Ry bestimmt wurde.

Die Schwingungsfrequenz ist fast identisch mit derjenigen, welche Soler et. al. mit der Augmented Plane Wave (APW)-Methode ermittelt haben. Der Wert von Blöchl weicht wiederum etwas stärker (um ca. 1%) ab. Der experimentelle Wert ist etwas kleiner. Möglicherweise ist dies ein Effekt der LDA-typischen Bindungsüberschätzung (obwohl die Bindungslänge nicht zu klein berechnet wird).

In der Dissoziationsenergie weichen die berechneten Werte untereinander um 1-2% voneinander ab, wobei der Wert von Blöchl diesmal kein Ausreißer ist, sondern in der Mitte liegt. Die Frage danach, welcher dieser Werte *richtiger* ist, erübrigt sich angesichts der Tatsache, daß sie sich alle deutlich (um ca. 15%) von dem experimentellen Wert unterscheiden.

Insgesamt zeigt sich in diesem Test eine sehr gute Übereinstimmung mit anderen Berechnungen.

6.3 Kupfer

Die Beschreibung von Übergangsmetallen wie z.B. Kupfer mit normerhaltenden Pseudopotentialen ist auf Grund der starken Lokalisierung der knotenlosen *d*-Funk-

tionen relativ aufwendig. Die Brauchbarkeit der PAW-Methode soll daher an diesem Beispiel stellvertretend für die 3d-Elemente erprobt werden.

Es sollen die berechneten Volumeneigenschaften mit den Werten aus anderen *ab-initio*-Rechnungen verglichen werden und auch die Effizienz der Methode (bzw. des erstellten Programmes) mit konkurrierenden Umsetzungen verglichen werden.

6.3.1 Volumeneigenschaften von fcc-Kupfer

Die natürliche Volumenstruktur von Kupfer ist ein fcc-Gitter. Die experimentelle Gitterkonstante beträgt 6.84 a.u. [ASH⁺99]. Die LDA überschätzt typischerweise die Bindungsstärke, was sich in einer zu kleinen Gitterkonstante und einem zu großen *bulk modulus* äußert.

Parameter

Das für die Rechnungen benutzte Potential wurde in Kapitel 4.7 beschrieben; es ist für $G_{\max} = 4.5 \sqrt{\text{Ry}}$ auskonvergiert. Das Austausch-Korrelations-Potential nach Vosko, Wilk und Nusair [VWN80] wurde verwendet. Die Rechnungen wurden mit einem Gitter von $13 \times 13 \times 13$ k-Punkten in der gesamten Brillouinzone nach Monkhorst und Pack [PM77] durchgeführt, was 84 Punkten im irreduziblen Teil der Brillouin-Zone entspricht. Zum Vergleich wurde auch ein Satz von $20 \times 20 \times 20$ k-Punkten (262 k-Punkte im irreduziblen Teil) herangezogen.

Zur Bestimmung von Gleichgewichts-Gitterkonstante und *bulk modulus* wurde der Abstand um ca. $\pm 1.5\%$ im Bereich des Minimums variiert. Die Bestimmung der Werte erfolgte durch Anpassen an Murnaghans Zustandsgleichung [Mur44].

Die Ergebnisse sind in Tabelle 6.5 wiedergegeben. Die Gitterkonstante ist um etwa 1% größer als die Werte, welche sich in den zum Vergleich herangezogenen FPKKR- bzw. FLAPW-Rechnungen [ASH⁺99] in LDA-Näherung ergeben. Der Unterschied zum experimentellen Wert ist damit mit etwa 2.1% etwas kleiner als der in den übrigen Rechnungen. Der *bulk modulus* wird zu 1.72 Mbar bestimmt. Dieser Wert ist um etwa 10% kleiner als die Werte, welche die beiden anderen LDA-Methoden ergeben. Die Bindungsstärke wird also als etwas schwächer bestimmt, was einen gewissen Kompensationseffekt mit der LDA-typischen Überschätzung der Bindung ergibt.

Um mögliche Fehler auszuschließen, wurde die Rechnung mit einer größeren Basis (bis zu $G_{\max} = 8.0 \sqrt{\text{Ry}}$) wiederholt, ebenso mit einem umfangreicheren Satz von $20 \times 20 \times 20$ bzw. $30 \times 30 \times 30$ k-Punkten (vgl. Tabelle 6.6).

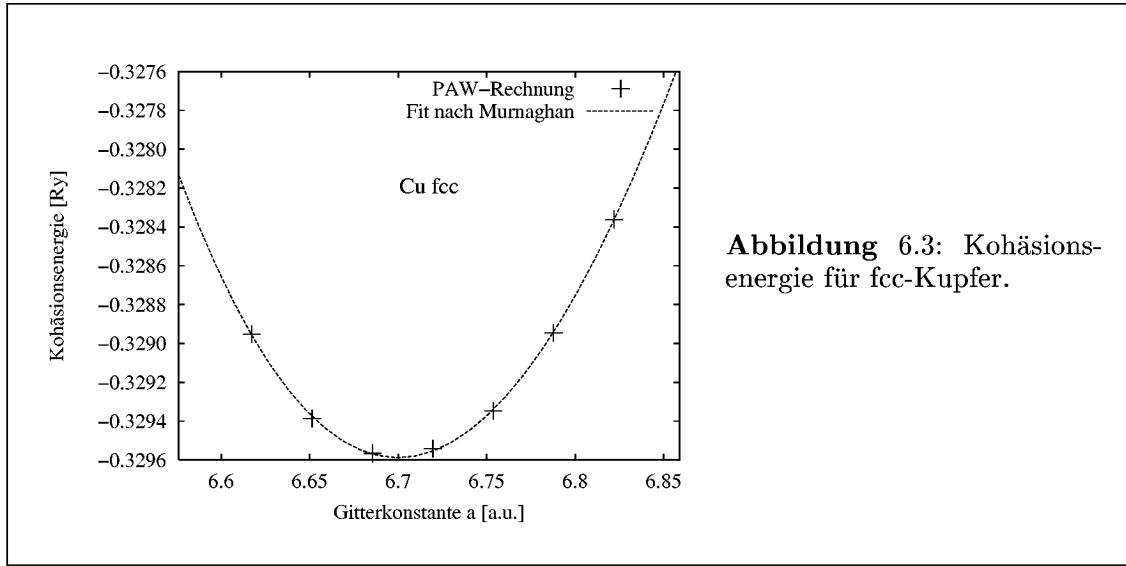


Abbildung 6.3: Kohäsionsenergie für fcc-Kupfer.

		Gitterkonstante a [a.u.]	<i>bulk modulus</i> B [Mbar]	Kohäsionsenergie E_{coh} [eV]
PAW	(NR-LDA)	6.70	1.72	4.48
PW ^a	(NR-LDA)	6.73	1.66	4.37
FLAPW ^a	(SR-LDA)	6.82	1.83	4.42
FLAPW ^a	(NR-LDA)	6.73	1.62	4.14
FLEUR ^b	(SR-LDA)	6.65	1.90	
FLEUR ^b	(NR-LDA)	6.71	1.76	
FPKKR ^c	(SR-LDA)	6.63	1.88	
FPKKR ^c	(SR-GGA)	6.84	1.38	
Exp.		6.84 ^c	1.37 ^c	3.50 ^d

Tabelle 6.5: Gitterkonstante a und *bulk modulus* B für fcc-Kupfer. Verglichen werden die in dieser Arbeit mit der PAW-Methode erzielten Ergebnisse mit anderen theoretischen und experimentellen Werten. *SR* steht für Skalarrelativistische Rechnungen, *NR* für nichtrelativistische. (a: [Jeo96], b: [BB], c: [ASH⁺99], d: [CC88])

Man sieht, daß bei Erhöhen der Abschneideenergie der Absolutwert der Gesamtenergie nur noch um ca. 2.5 mRy sinkt; das entspricht etwa dem erwarteten Wert (siehe Kapitel 4). Die Gitterkonstante und der *bulk modulus* ändern sich aber kaum oberhalb von $G_{\max} = 4.5 \sqrt{\text{Ry}}$. Letzterer hängt aber relativ stark davon ab, für welche Gitterkonstanten man die Energie berechnet und wie man den Fit durchführt. Berechnet man die Energien für Gitterkonstanten, welche bis zu 6% vom Minimumwert abweichen, so erhält man einen *bulk modulus*, welcher ca. 3% kleiner ist

k-Punkte, $G_{\max}$	Gitterkonstante a [a.u.]	Gesamtenergie E_0 [Ry]	<i>bulk modulus</i> B_0 [Mbar]
131313, 3.5	7.01469444	−208.271467	1.41697558
131313, 3.8	6.74550888	−208.413422	2.15796350
131313, 4.0	6.70425418	−208.442443	1.87514107
131313, 4.5	6.70065831	−208.456850	1.67339422
131313, 5.0	6.69729805	−208.458395	1.67567224
131313, 5.5	6.69790137	−208.459115	1.66868727
131313, 6.0	6.69813809	−208.459217	1.66760502
131313, 8.0	6.69802547	−208.459419	1.66680644
131313, 3.5 fine	6.67209022	−208.283408	2.04862424
131313, 3.8 fine	6.77236512	−208.413343	1.82831644
131313, 4.0 fine	6.69081606	−208.442564	1.72748438
131313, 4.5 fine	6.69979195	−208.456915	1.71745355
131313, 4.5 fine	6.69979195	−208.456915	1.71745355
202020, 4.5 fine	6.70103365	−208.456981	1.71385512
303030, 4.5 fine	6.70084426	−208.456937	1.73700468

Tabelle 6.6: Variation der strukturellen Werte von (fcc-) Kupfer mit verschiedenen Parametern. *fine* besagt, daß die Gitterkonstante nur um $\pm 1.5\%$ des Gleichgewichtswertes variiert wurde. Bei den anderen Fits wurde sie um 6% variiert.

als jener, den man bei Berechnung eng benachbarter Gitterkonstanten erhält (Markierung *fine* in Tabelle 6.6). Bei Variation der Größe der k-Punkt Sätze ändert sich die Energie um weniger als 0.1 mRy, der *bulk modulus* hingegen um mehr als 1%.

Auch wurde versuchsweise statt des äquidistanten radialen Gitters für die Augmentationsgebiete das (sehr dichte) exponentielle Gitter verwendet, welches bei der Generierung des Potentials verwendet wurde. Schließlich wurde auch das in Kapitel 4.7 erwähnte Potential Cu_B ausprobiert. In allen Fällen ergab sich bei $G_{\max} = 4.5 \sqrt{\text{Ry}}$ und $13 \times 13 \times 13$ k-Punkten dieselbe Gitterkonstante und die Abweichung des *bulk modulus* vom obigen Wert war kleiner als 1%.

Tabelle 6.5 zeigt, daß die hier bestimmte Gitterkonstante und der *bulk modulus* gut übereinstimmen mit anderen Werten, welche nicht-relativistisch berechnet wurden. Auch hier wurde zur Konstruktion des Pseudopotentials keine skalarrelativistische

Rechnung genutzt. Offensichtlich macht sich dies jedoch deutlich bemerkbar, obwohl Kupfer (Ordnungszahl 29) noch zu den eher leichten Elementen gehört. Daher müssen zur korrekten Beschreibung von Kupfer anscheinend relativistische Effekte bei der Erzeugung des Pseudopotentials berücksichtigt werden. Dies ist in skalarrelativistischer Näherung möglich und bewirkt im wesentlichen eine Veränderung der Matrixelemente $D_{kin;\gamma,\gamma'}^\mu$ der kinetischen Energie (vgl. [Blö94]). Hierzu könnte das Generierungsprogramm modifiziert werden, um noch bessere Ergebnisse zu erzielen.

6.3.2 Volumeneigenschaften von hcp-Kupfer

Zum Vergleich wurde auch die hcp-Struktur von Kupfer berechnet. Für ein Cu-Atom unterscheidet sich diese Struktur von der fcc-Anordnung erst durch die übernächsten Nachbarn. Daher erwartet man, daß der Energieunterschied sehr klein sein wird. Deswegen haben wir dieses Beispiel als Test aufgenommen. Um näherungsweise dieselbe Dichte von k-Punkten wie im Falle der fcc-Struktur zu erreichen, wurden $12 \times 12 \times 7$, $19 \times 19 \times 11$ bzw. $26 \times 26 \times 16$ Punkte verwendet. Die Gitterkonstante a und c wurden unabhängig voneinander in dem Bereich variiert, welcher sich aus der Annahme eines ähnlichen interatomaren Abstandes wie im fcc-Kupfer ergibt. Das Minimum der Energie wurde durch das Anpassen eines kubischen Polynoms ermittelt. Um verlässliche Energiedifferenzen zu garantieren, wurde fcc-Kupfer nochmals in der gleichen Einheitszelle wie hcp-Kupfer berechnet. Die Ergebnisse sind in Tabelle 6.7 zusammengefaßt.

	a [a.u.]	c/a	E_0 [Ry]	$E_0^{hcp} - E_0^{fcc}$ [meV]
PAW, $12 \times 12 \times 7$ k-Punkte	4.76	1.62	-208.45573	16.1
PAW, $19 \times 19 \times 11$ k-Punkte	4.72	1.65	-208.45621	9.7
PAW, $26 \times 26 \times 16$ k-Punkte	4.73	1.64	-208.45628	8.9
FLEUR (SR)	4.69	1.65		3.1
FLEUR (SR)	4.71	1.62 ^a		4.8
FLEUR (NR)	4.74	1.62 ^a		4.0

Tabelle 6.7: Strukturelle und energetische Eigenschaften von (hypothetischem) hcp-Kupfer. $E_0^{hcp} - E_0^{fcc} > 0$ bedeutet, daß die fcc-Struktur energetisch günstiger ist. Die Energien sind pro Atom angegeben. (a: c/a-Wert vorgegeben; SR: skalarrelativistisch; NR: nichtrelativistisch)

Die Rechnung zeigt, daß – wie erwartet – die hcp-Struktur energetisch ungünstiger

ist. Das $\frac{c}{a}$ -Verhältnis ist nahe am *idealen* Wert von $\sqrt{\frac{8}{3}} = 1.633$. Das Volumen pro Atom beträgt 75.5 a.u.^3 im Vergleich zu 75.2 a.u.^3 für fcc-Kupfer. Diese Werte sind allerdings in unterschiedlichen Superzellen berechnet. Berechnet man die Gesamtenergie des fcc-Kupfers in einer hexagonalen Zelle (mit $19 \times 19 \times 8$) k-Punkten, so ergibt sich ein Unterschied von 2 meV gegenüber dem Wert, welcher in der kubischen Zelle (mit $20 \times 20 \times 20$ k-Punkten) berechnet wurde. Offenbar muß man die Ungenauigkeit durch die verschiedenen genauen Konvergenz der k-Punkt Summation in dieser Größenordnung ansetzen. Auch ist das Minimum nicht sehr ausgeprägt, so daß eine gewisse Ungenauigkeit beim Anpassen einer Fit-Funktion auftritt.

Die mit dem FLAPW-Programm FLEUR berechneten Energiedifferenzen fallen zwar auch bei nicht-relativistischer Rechnung geringer aus als die in dieser Arbeit berechneten Werte, aber das Vorzeichen der Energiedifferenz wird korrekt wiedergegeben, und auch die Größenordnung ist richtig.

6.3.3 Effizienztest: Cu(110) Oberfläche

Um eine Vorstellung von der Effizienz des Programmes zu gewinnen, wurden Berechnungen für eine große Kupfer Oberflächenzelle durchgeführt. Dieses System wurde ausgewählt, da hierfür schon Effizienztests anderer Programme verfügbar sind [PWH⁺00, BKB].

Es handelt sich um eine neunlagige (4×2) Cu(110) Zelle. Sie enthält 72 Atome, welche jeweils 11 Valenzelektronen aufweisen. Als Gitterkonstante wurde der LDA-Wert $a_0 = 6.64 \text{ a.u.}$ gewählt. Die Dicke der Vakuumschicht entspricht fünf Substratschichten. Es wurde nur ein einzelner k-Punkt benutzt; alle Methoden skalieren fast linear mit der Anzahl der Punkte. Diese Zelle besitzt die Symmetrie $\Gamma_o D_{2h}^1$. Es wurde hier aber von der Symmetrie nur insofern Gebrauch gemacht, als sie erlaubt, die Anzahl der k-Punkte zu reduzieren; es wurde aber z.B. nicht die Inversionssymmetrie genutzt, um reelle Wellenfunktionen zu erhalten.

Das Pseudopotential erfordert eine Abschneideenergie von $G_{\max} = 4.5 \sqrt{\text{Ry}}$. Die Basis umfaßt dann 12663 Ebene Wellen. Es werden jeweils zwei Projektoren für die *s*-, *p*- und *d*-Komponente benutzt, also 18 pro Atom und somit insgesamt 1296 Projektoren. Die Dichte und das Potential werden entsprechend bis $l = 4$ berechnet. Das radiale Gitter verfügt über 301 Stützstellen.

Für die insgesamt 792 Elektronen wurden 430 Zustände berechnet. Zur Diagonalisierung des Hamiltonian wurde das Davidson-Verfahren (mit diagonalem Präkonditionierer) verwendet. Die Dimension des Unterraumes betrug zweimal die Anzahl

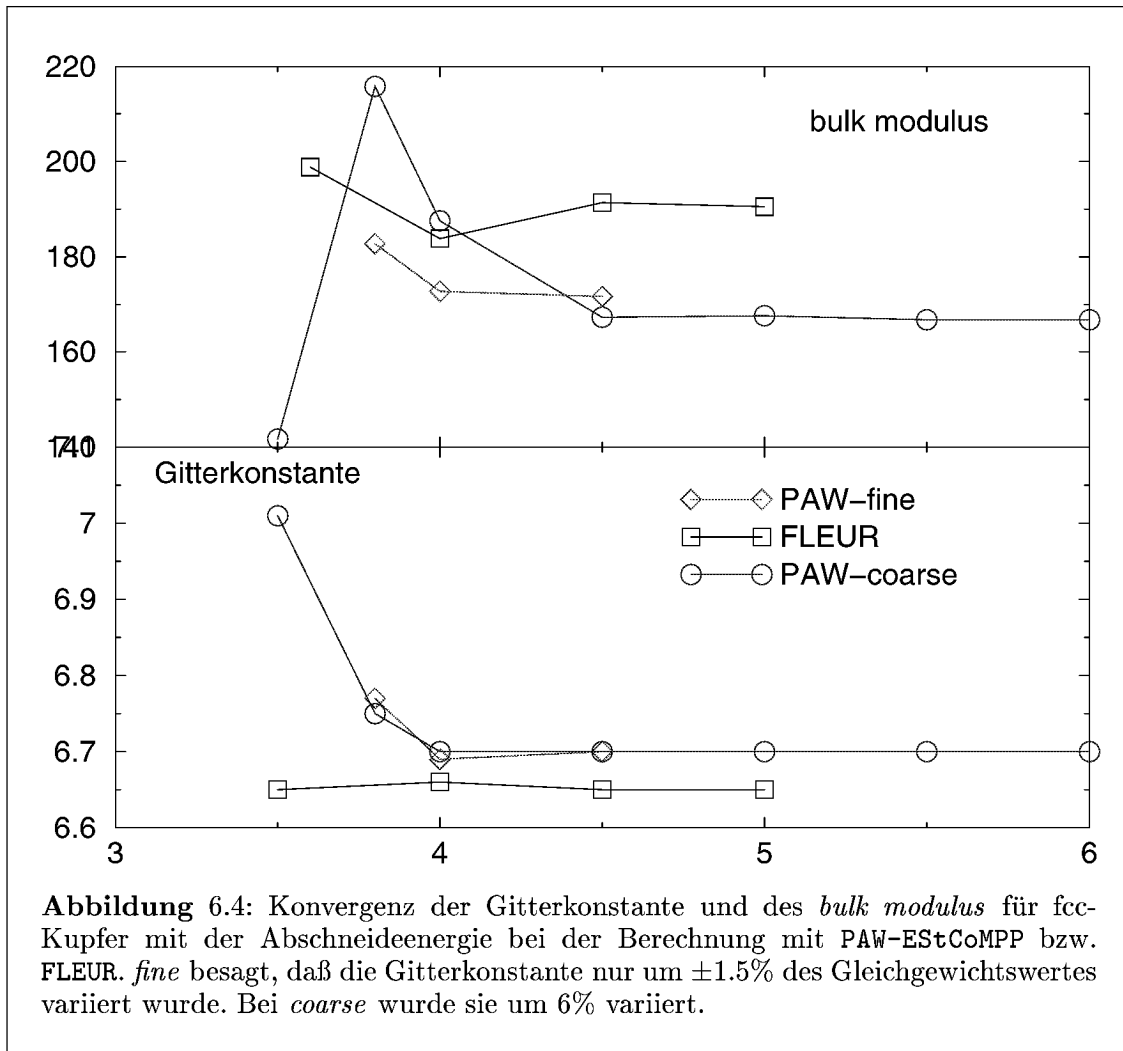
der Zustände. Pro Selbstkonsistenzschritt wurden drei Davidsonschr itte ausgef hrt. Die Startwellenfunktionen wurden nach dem APSO-Schema (vgl. Kapitel 5.2.2.5) mit 9 atomaren Wellenfunktionen pro Atom erzeugt, erg nzt um 153 Ebene Wellen. Die Superposition atomarer Ladungsdichten bildete die Startladungsdichte. Als Konvergenzkriterium wurde ein RMS-Fehler des Potentials von 10^{-6} gew hlt. Die Rechnungen wurden auf einer Cray T90 auf einem Knoten mit 1.8 GFLOPS² Peak-performance ausgef hrt.

Die Konvergenz auf die angestrebte Genauigkeit wurde in 22 Selbstkonsistenzschritten erreicht. Die dazu ben tigte CPU-Zeit betrug 9h 04m. Eine Flop-Rate von 1300 MFLOPS wurde erzielt.

Mit dem FLAPW-Programm FLEUR [BB] hat G. Bihlmayer dasselbe System auf demselben Rechner auf verschiedene Weisen berechnet. Die dort ben tigte Abschneideenergie f r die Wellenfunktionen betr gt $G_{\max}^2 = 13.22$ Ry. Um einsch tzen zu k nnen, ob sich bei den jeweils gew hlten Abschneideenergien ein fairer Vergleich ergibt, wird in Abbildung 6.4 die Konvergenz der Gitterkonstante und des *bulk modulus* f r fcc-Kupfer mit der Abschneideenergie gezeigt. Die Rechnung mit FLEUR erfolgte in der LDA und der skalarrelativistischen N herung f r die kinetische Energie. Man findet, da  bei einer Abschneideenergie von 13.22 Ry f r FLEUR bzw. $G_{\max} = 4.5 \sqrt{\text{Ry}}$ f r PAW-EStCoMPP der *bulk modulus* weitgehend und die Gitterkonstante vollst ndig auskonvergiert ist. Ist man lediglich an Strukturoptimierung interessiert, so gen gt f r die PAW-Methode auch ein Abschneidewert von $G_{\max} = 4.0 \sqrt{\text{Ry}}$. Die Rechenzeit f r einen Selbstkonsistenzschritt ist daher auch f r diesen Wert bestimmt worden; alle weiteren Analysen beziehen sich jedoch auch den Abschneidewert $G_{\max} = 4.5 \sqrt{\text{Ry}}$. Der Quotient der beiden Rechenzeiten von 0.709 stimmt gut mit dem Verh ltnis der Kuben der Abschneidewerte von 0.702  berein; dies entspricht dem Skalierungsverhalten, da  man f r gro e Systeme erwartet.

Den Vergleich der Rechenzeiten pro Iteration zeigt Tabelle 6.8. Wenn man unter gleichen Bedingungen rechnet, d.h. ebenfalls eine Superzelle und mit komplexem Hamiltonoperator, so ist das PAW-Programm beim Abschneidewert $G_{\max} = 4.5 \sqrt{\text{Ry}}$ um einen Faktor 3.0, bei $G_{\max} = 4.0 \sqrt{\text{Ry}}$ sogar um einen Faktor vier schneller als das FLAPW-Programm. Benutzt man jedoch die Film-Version des letzteren, so wird die Berechnung der Vakuumregion sehr viel effizienter als das in einer Superzellenbeschreibung mit Ebenen Wellen der Fall sein kann. Je nachdem, wie stark man die Symmetrie ausnutzt, ist der Geschwindigkeitsunterschied geringer, und bei Ausnutzung der Inversions- und der Spiegelsymmetrie wird FLEUR dann um einen Faktor

²1.8 GFLOPS hei t 1.8×10^9 Gleitkomma-Rechenoperationen pro Sekunde.



1.4 schneller.

Eine Implementierung eines PP-Programmes, welches die Inversionssymmetrie zur Beschränkung auf einen reellen Hamiltonoperator nutzt, ist in Arbeit. Die Änderungen hierbei sind fast deckungsgleich mit den im PAW-Programm erforderlichen Veränderungen, da dieses auf jenem basiert und da fast alle Größen, welche die Augmentationsgebiete beschreiben, ohnehin reell sind.

Tabelle 6.9 zeigt, wie sich die Rechenzeit des PAW-Programmes auf verschiedene Aktivitäten verteilt. Der weitaus größte Teil (93%) wird für die iterative Lösung des Eigenwertproblems benötigt. Daneben benötigt nur die Berechnung der Ladungsdichte noch einen nennenswerten Anteil (3.8%) der Rechenzeit. Diese erfordert pro Selbstkonsistenzschritt etwa 57 Sekunden. Davon entfallen nur etwa 5 Sekunden auf die Erzeugung der glatten Ladungsdichte $\tilde{n}(\vec{r})$; der Rest wird für die Erzeugung

Programm	Zeit pro SC-Iteration [sec]
PAW-EStCoMPP (Superzelle, komplex, $G_{\max} = 4.5 \sqrt{\text{Ry}}$)	1395
PAW-EStCoMPP (Superzelle, komplex, $G_{\max} = 4.0 \sqrt{\text{Ry}}$)	989
FLEUR (Superzelle, komplex)	4132
FLEUR (Film, komplex)	3351
FLEUR (Film, komplex, z-Spiegelsymmetrie)	2146
FLEUR (Film, reell, z-Spiegelsymmetrie)	996

Tabelle 6.8: Vergleich der Rechenzeit für eine Selbstkonsistenziteration bei der Cu(110) 4×2 Zelle zwischen dem PAW-Programm und dem FLAPW-Programm FLEUR.

der Ladungsdichtematrix $\rho_{\gamma,\gamma'}^{\mu}$ benötigt. (Hierzu werden für alle besetzten Zustände die Produkte $\langle p_{\gamma} | \tilde{\psi}_{\vec{k}\nu} \rangle$ gebildet, was den Hauptanteil der Zeit ausmacht.) Die Zeit für die Erzeugung des Potentials verteilt sich etwa zur Hälfte auf die Berechnung des Hartree-Potentials (inklusive Pseudoladung) im reziproken Raum und auf die Bestimmung des Austausch-Korrelations-Potentials in den Augmentationsgebieten. Sie ist jedoch insgesamt vernachlässigbar.

Programmteil	Zeit [sec]	Zeitanteil [%]
Initialisierung	220	0.7
APSO Startvektoren	360	1.1
Davidson	30299	92.9
Erzeugung der Ladungsdichte	1246	3.8
Erzeugung des Potentials	151	0.5
Mischen des Potentials	34	0.1
Kräfte	274	0.8
Insgesamt	32627	100.0

Tabelle 6.9: Laufzeiten verschiedener Programmanteile für die Cu-Oberflächenzelle.

Der größte Anteil der Rechenzeit, nämlich die für das Lösen des Eigenwertproblems, ist in Tabelle 6.10 weiter aufgeschlüsselt. Bei der betrachteten Rechnung wurden drei Davidson-Iterationen pro Ladungsselbstkonsistenzschritt durchgeführt. Die Zeiten der ersten und der letzten Davidson-Iteration unterscheiden sich etwas von der zweiten (bzw. von allen inneren Iterationen, wenn man mehr als drei Davidson-Iterationen pro Selbstkonsistenzschritt durchführen würde). Man sieht, daß die be-

Programmteil	Zeit [sec]	Zeit [sec]
Gram-Schmidt	92	+27
$V_{loc} \times \psi$	10	+10
$V_{NL} \times \psi$	118	+118
Aufstellen der Matrizen	64	
Diagonalisierung	13	
Rücktransformation	84	-38
Insgesamt	381	

Tabelle 6.10: Laufzeiten innerhalb des Davidson-Unterprogrammes. Die zweite Spalte gibt die Zeit für einen mittleren Davidsonschrift an, die dritte die zusätzliche Zeit für den ersten bzw. die ersparte Zeit für den letzten Schritt innerhalb eines Ladungsselbstkonsistenzschrittes.

stimmenden Faktoren die Orthogonalisierung und die Anwendung des nichtlokalen Potentials sind. In beiden Fällen werden über 99% der jeweiligen Ausführungszeit durch hochoptimierte BLAS-Routinen³ benötigt. Bei der Anwendung des nichtlokalen Potentials auf die Wellenfunktionen wird eine Geschwindigkeit von über 1500 MFLOPS erreicht. Eine weitere Verbesserung ist hier wohl nur durch algorithmische Veränderungen, nicht aber durch Optimierung der Implementierung des gewählten Algorithmus⁴ zu erreichen.⁴ Ein Vorschlag für eine algorithmisch vorteilhaftere Variante des Davidson-Verfahrens findet sich in Anhang C.

Weitere Verbesserung ist möglich durch andere Verfahren zur Diagonalisierung. Insbesondere ist hier das RMM-DIIS-Verfahren [KF96] zu nennen; bei diesem wird nicht der (Matrix-)Rayleigh-Quotient minimiert, sondern die Residuen der Vektoren selbst. Da deren Absolutbeträge ein (unbeschränktes) Minimum haben, kann eine Orthogonalisierung hier entfallen, so daß einer der beiden größten Rechenzeitanteile drastisch reduziert wird.

Außerdem ist es möglich, die Anwendung des nichtlokalen Potentials im Ortsraum auszuführen [KSPL91]. Dadurch läßt sich die Rechenzeit hierfür beträchtlich reduzieren. (Laut [PWH⁺00] erreicht man eine Beschleunigung um einen Faktor 2-3).

Die gleiche Kupferoberfläche wurde freundlicherweise auch von Y. Morikawa mit seinem Ultrasoft-Pseudopotentialprogramm (US) berechnet. Der Vergleich der uns zu Verfügung gestellten Rechenzeiten mit den obigen gestaltet sich aber schwierig. Zum

³BLAS steht für *Basic Linear Algebra Subprograms*; diese Sammlung von Unterprogrammen gibt es von fast allen Hardwareherstellern in maschinenspezifisch optimierter Form.

⁴Soweit man die Umsetzung einer reellen Programmversion nicht als solche versteht.

einen benutzt jenes Programm den Davidson-Algorithmus mit *wachsender* statt konstanter Unterraumgröße, nämlich 1-, 2- und 3-mal die Anzahl der Zustände. Am ehesten vergleichen läßt sich ein solcher Dreifachschritt mit zwei Davidson-Iterationen unseres Programmes; in beiden Fällen muß der Hamiltonoperator auf $3 \times N_{state}$ Vektoren angewandt werden. Bei fester Unterraumgröße muß in jedem Schritt eine Rücktransformation der Vektoren durchgeführt werden, was bei wachsendem Unterraum unnötig ist; dagegen ist der Aufwand für die Orthogonalisierung in letzterem Falle höher. Zum anderen kommt erschwerend hinzu, daß ein anderer Computer benutzt wurde, welcher nominell um einen Faktor 4.4 schneller war.

Die Ausführungszeit für zwei Davidson-Schritte in unserem PAW-Programm beträgt ca. 900 Sekunden; dem stehen 240 Sekunden für eine (dreifach-) Schleife im US-Programm von Morikawa gegenüber. Der Faktor 3.75 spiegelt etwa den erwähnten Quotienten der maximalen FLOP-Raten wider.

Die Anzahl der benötigten Iterationen ist fast identisch (22 PAW gegenüber 24 US). Eine Abschätzung der Auswirkung der oben angesprochenen Verbesserungen kann man gewinnen aus der ebenfalls von Y. Morikawa durchgeführten Berechnung unter Benutzung des RMM-DIIS-Verfahrens und der Ortsraumauswertung von Projektor-Wellenfunktion-Produkten. Dies erhöht zwar die Anzahl der Iterationen auf 28, die Gesamtzeit sinkt jedoch um ca. 30%.

Die gesammelten Ergebnisse lassen sich dahingehend resümieren, daß der gewählte Algorithmus effizient implementiert ist. Weiter verbessert werden könnte er durch eine bessere Ausnutzung der Symmetrien, die in vielen der zu untersuchenden physikalischen Systeme vorhanden sind, sowie auch durch algorithmische Varianten wie etwa das RMM-DIIS-Verfahren oder der Ortsraumdarstellung der Projektoren. Berücksichtigen wir diese 30% auch für unser PAW-Programm, so erwarten wir, daß es das Potential hat, ca. 4-5 mal schneller zu sein als das FLAPW-Programm FLEUR.

Kapitel 7

Halbleiter mit großer Bandlücke

7.1 Siliziumkarbid

Siliziumkarbid (SiC) stellt auf Grund seiner physikalischen Eigenschaften ein sehr interessantes Halbleitermaterial dar. Einerseits kann das Material selbst zur Konstruktion elektronischer Bauteile benutzt werden; andererseits eignet es sich wegen seiner Gitterkonstante als Trägermaterial für Heterostrukturen mit Substanzen wie Aluminium- und Galliumnitrid (AlN bzw. GaN). Deren hexagonale Wurtzitstrukturen besitzen eine sehr ähnliche Gitterkonstante.

SiC tritt in einer Vielzahl von Polytypen auf; sie besitzen indirekte Bandgaps, deren Größe zwischen 2.4 eV und 3.3 eV variiert. Neben einer kubischen Variante (β -SiC) gibt es eine Vielzahl von hexagonalen (z.B. 6H-SiC) und rhomboedrischen Typen. Bauelemente aus 6H-SiC weisen im Vergleich zu solchen aus Silizium eine wesentlich höhere Betriebstemperatur, eine höhere thermische Leitfähigkeit und eine größere Strahlungsunempfindlichkeit auf. Durch das relativ große Bandgap ist SiC auch für optische Bauelemente interessant; Polytypen unterschiedlicher Dotierung decken den gesamten sichtbaren Bereich sowie Teile des ultravioletten Spektrums ab.¹

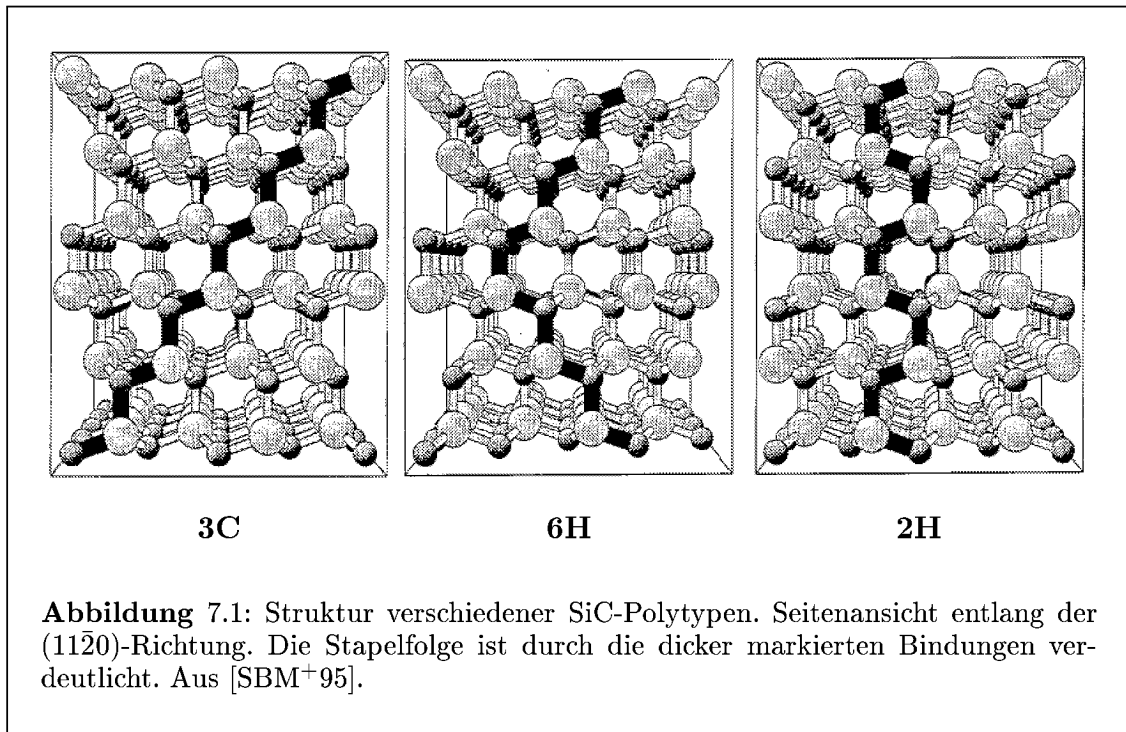
Hier sollen Volumeneigenschaften von zwei Polytypen, nämlich der kubischen β -SiC- (auch 3C-SiC) und der hexagonalen 6H-Variante bestimmt werden. In allen Polytypen sind Si- bzw. C-Atome von vier Atomen der jeweils anderen Sorte tetraedrisch umgeben.

β -SiC ist vom Zinkblende-Typ. Es entsteht wie das Diamantgitter aus zwei ineinandergestellten fcc-Gittern, nur sind die Atome der beiden Gitter hier unterschiedlich. Die Einheitszelle der 6H-Modifikation weist sechs Doppelschichten von Si- und C-

¹Die vorangehenden Angaben wurden [Sab96] entnommen.

Atomen auf. Sie bilden die Schichtungsfolge ABCACB.

Stellt man auch β -SiC als hexagonale Zelle dar (die $[111]$ -Richtung der kubischen Zelle entspricht der $[0001]$ -Richtung der hexagonalen), so weist diese die Schichtfolge ABCABC auf. Die Atome unterscheiden sich also erst in der Anordnung der drittnächsten Nachbarn.



Kubisches 3C-SiC

Die Berechnung der kubischen Variante wurde vergleichshalber sowohl mit der neuen PAW-Methode, als auch mit dem Pseudopotentialprogramm **ESCoMPP** (PP) durchgeführt. Bei ersterer wurde das C_l -Potential benutzt, welches in Kapitel 4 beschrieben wurde; es erfordert eine Abschneideenergie von 21 Ry, um eine Konvergenz der kinetischen Energie nach dem Rappe-Kriterium auf weniger als 1 mRy zu erreichen, während die dafür erforderliche Abschneideenergie für das normerhaltende Kohlenstoff Pseudopotential ca. 60 Ry beträgt [ERBS]. Verlässliche Gitterkonstanten lassen sich in beiden Fällen auch mit einer kleineren Basis erzielen. In beiden Fällen wurde ein normerhaltendes Silizium-Pseudopotential verwendet, dessen Abschneideenergie bei etwa 12.25 Ry liegt. (Die Konstruktionsparameter sind in [Ric96] zu finden.) Es wurden $4 \times 4 \times 4$ k-Punkte verwendet. Die Gitterkonstante wurde um 6% va-

riert und die Parameter von Murnaghans Zustandsgleichung durch Anpassung an die gewonnenen Energien bestimmt.

Die berechneten strukturellen Eigenschaften der kubischen Variante sind in Tabelle 7.1 mit Ergebnissen anderer Berechnungen und experimentellen Werten zusammengestellt.

	Methode	Gitterkonstante a [Å]	<i>bulk modulus</i> B [Mbar]	Kohäsionsenergie E_{coh}/atom [eV]
Diese Arbeit (C_l)	PAW	4.34	2.17	8.12
Diese Arbeit	NC-PP	4.34	2.15	7.53
[PCLC94]	NC-PP	4.35	2.20	
[KPW ⁺ 94]	NC-PP	4.34	2.22	
[Sab96]	GO	4.33	2.22	
[SMV97]	NC-PW	4.33	2.23	
[FKBK00]	US-PP	4.33	2.27	8.33
[HMD ⁺ 97]	PAW	4.32	2.20	8.4
experimentell ^a		4.36	2.24	
Diese Arbeit (C_x)	PAW	4.34	2.22	7.98

Tabelle 7.1: Gitterkonstante a , *bulk modulus* B und Kohäsionsenergie E_{coh} von β -SiC im Vergleich zu Ergebnissen anderer Rechnungen und zum Experiment. *GO* steht für Gauß-Orbitale, *NC-PP* für normerhaltende Pseudopotentiale, *US-PP* für Rechnungen mit Ultrasoftes Pseudopotentialen. C_x bezeichnet ein Kohlenstoff-Potential, das (im Gegensatz zu C_l) die Rumpfreionen benachbarter Siliziumatome schneidet und so eine kleinere Abschneideenergie von 16 Ry erfordert. Keine der Energien enthält eine Spin-Polarisations-Korrektur. (Im Falle von [FKBK00] beruht diese Aussage auf einem Analogieschluß zu [KJ99].) a: zitiert nach [Sab96]

Die Gitterkonstante der PAW-Rechnung ist wie üblich etwas kleiner als der experimentelle Wert, stimmt aber sowohl mit der eigenen Pseudopotentialrechnung als auch mit anderen *ab-initio*-Berechnungen überein.

Der *bulk modulus* ist sowohl für die Pseudopotential- als auch die PAW-Rechnung etwas kleiner als die übrigen Werte. Allerdings hängt dieser Wert bekanntermaßen sehr empfindlich z.B. von der Wahl der Gitterkonstanten, für welche man die Energie bestimmt, von der Art der Fitfunktion und Ähnlichem ab. Abweichungen hierdurch von bis zu 0.03 Mbar wurden festgestellt. Schließlich machen sich speziell bei der PAW-Rechnung beim Vergrößern der Zelle Sprünge in der Basisgröße bemerkbar, da diese so klein ist. Angesichts dieser Effekte kann man die Übereinstimmung als befriedigend bezeichnen.

Der relativ große Unterschied zwischen den Kohäsionsenergien, welche man mit PP- bzw. PAW-Methode bestimmt, könnte darin begründet sein, daß die Bezugsgrößen unterschiedlich bestimmt wurden. Als Gesamtenergie des Kohlenstoffatoms wurde im PP-Falle die bei der Generierung des Potentials bestimmte Energie verwendet; im Falle des PAW-Potentials wurde die Energie des Atomes in einer großen Zelle mit einer Kantenlänge von 16 Å bestimmt unter Benutzung desselben Abschneidewertes $G_{\max}$, wie in der Festkörperrechnung.

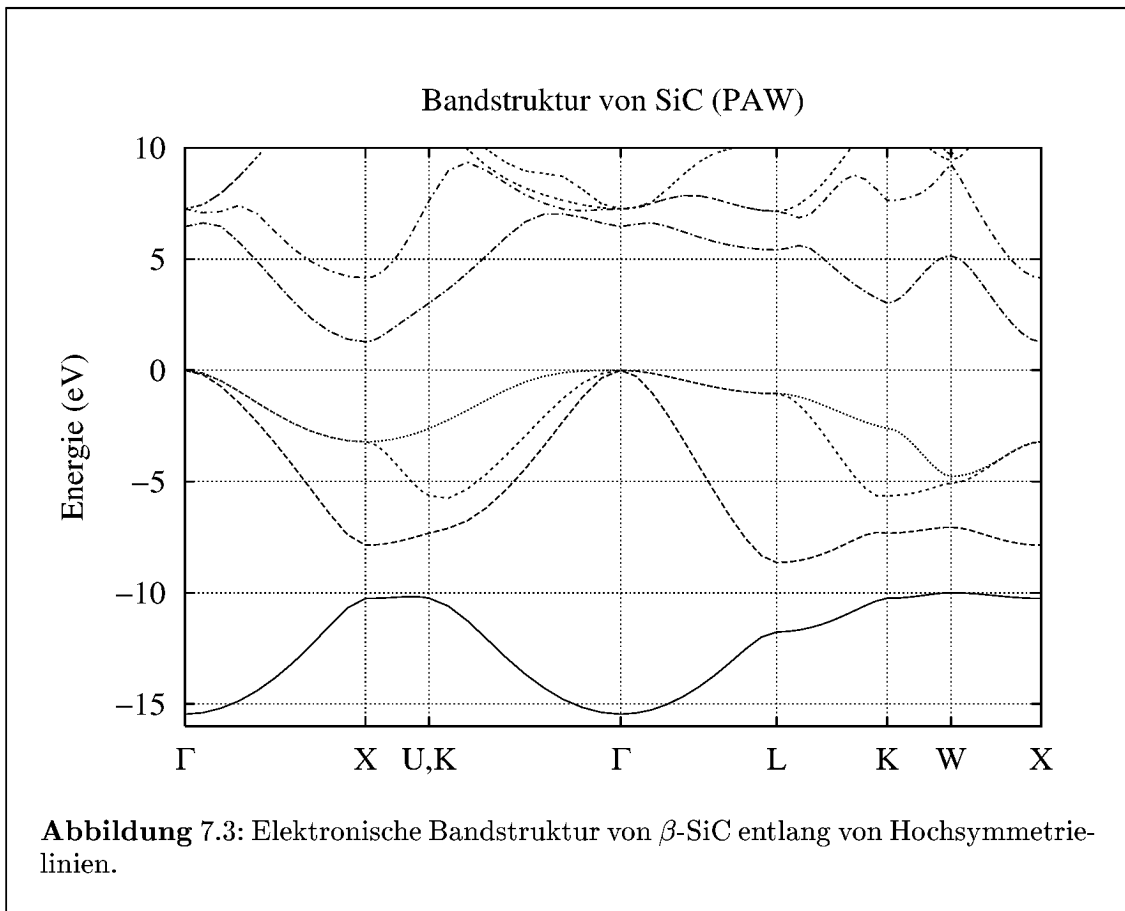
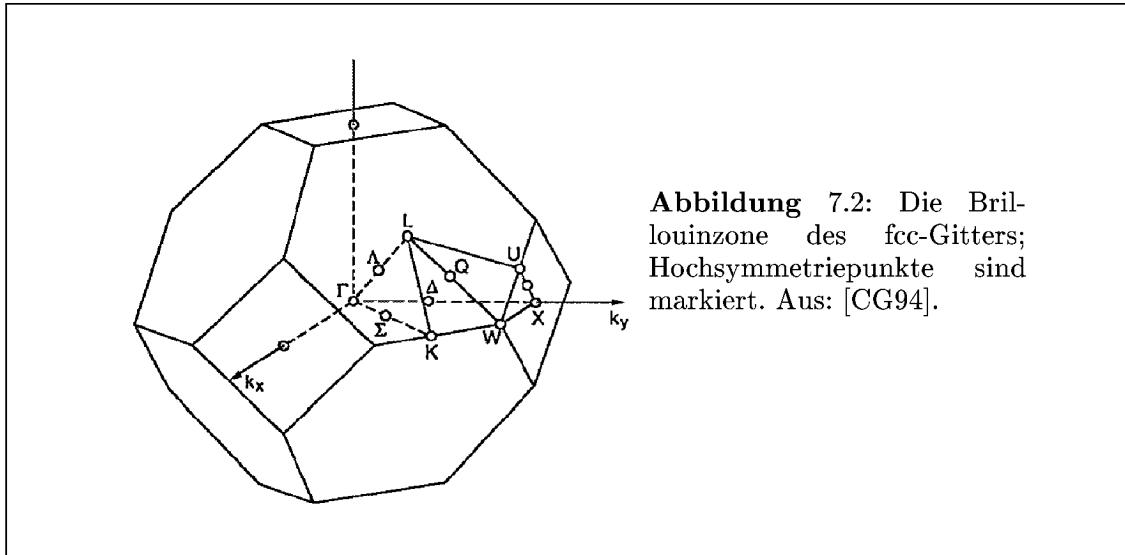
Bei der Erzeugung des C_x -Potentials, die in Kapitel 4.5 beschrieben ist, wurde die Regel mißachtet, daß benachbarte Rumpfbereiche sich nicht überschneiden sollten. Durch den größeren Abschneideradius ist es aber weicher ($G_{\max} = 4.0 \sqrt{\text{Ry}}$ statt $G_{\max} = 4.6 \sqrt{\text{Ry}}$ für das C_l -Potential). Die mit diesem Potential bei $G_{\max} = 4.0 \sqrt{\text{Ry}}$ bestimmte Gitterkonstante und die Kohäsionsenergie stimmen recht gut mit den Werten des C_l -Potentials überein. Eine etwas stärkere Abweichung zeigt der *bulk modulus*; dieser konvergiert für etwas größere Abschneidewerte gegen denselben Wert wie beim C_l -Potential; da dies allerdings mit starken Schwankungen geschieht, läßt sich vermuten, daß dies weniger mit der Konvergenz des Potentials zu tun hat als mit der Ungenauigkeit des Fits auf Grund der sich bei Variation der Gitterkonstante ändernden Basisgröße. Demnach ergeben sich auch bei geringfügig überlappenden Rumpfbereichen noch vernünftige Ergebnisse. Dennoch wurde im folgenden für PAW-Rechnungen das C_l -Potential verwendet.

Während die PAW-Methode mit dem C_l -Potential für die Berechnung *einer* Energie auf einer modernen Workstation weniger als eine Minute benötigt, dauert die entsprechende Rechnung mit KB-Potentialen fast dreimal so lange. Allerdings ist dieser Vergleich wenig aussagekräftig, da das betrachtete System so klein ist. Aus dem Vergleich der Abschneidewerte $G_{\max}$ erwartet man einen Faktor von 5.3. Auch der Unterschied zwischen C_l - und C_x -Potential ist mit einem Faktor von 1.3 weniger deutlich ausgeprägt, als man das im Limes großer Systeme erwartet. Dort erwartet man ein kubisches Skalierungsverhalten der Rechenzeit mit $G_{\max}$.

Auch die elektronische Struktur des β -Siliziumkarbids wurde sowohl mit dem PAW-Programm als auch mit normerhaltenden Pseudopotentialen (PP) untersucht. Dazu wurde die Gleichgewichtsgitterkonstante ($a=4.34$ Å) benutzt. Abbildung 7.2 zeigt die Brillouinzone eines fcc-Gitters für die 3C-Struktur.

Die berechnete Bandstruktur entlang der Hochsymmetrielinien der Brillouinzone zeigt Abbildung 7.3; sie ist auf dieser Skala nicht von der mittels der PP-Methode berechneten Bandstruktur zu unterscheiden.

Den Vergleich der Bandenergien an einigen Hochsymmetriepunkten zeigt Tabelle 7.2.



Die Abweichung zwischen PAW und PW-Rechnung ist speziell für besetzte Zustände sehr gering, aber auch die unbesetzten Bänder stimmen gut überein. Das Bandgap

unterscheidet sich um 0.025 eV, bedingt durch die Abweichung bei X_{1c} , welche vergleichsweise groß ausfällt. Jedoch streuen die Bandlücken, die in anderen *ab-initio*-Methoden bestimmt werden, im Bereich von ca. 1.2 eV bis 1.4 eV (siehe [Sab96]), so daß sich beide Werte durchaus im üblichen Rahmen bewegen.

	PAW	PW	[Sab96]
Γ_{1v}	-15.45	-15.47	-15.44
Γ_{4v}	0.00	0.00	0.00
Γ_{1c}	6.46	6.43	6.33
Γ_{2c}	7.26	7.25	7.14
X_{1v}	-10.25	-10.26	-10.35
X_{2v}	-7.86	-7.86	-7.86
X_{3v}	-3.21	-3.22	-3.20
X_{1c}	1.29	1.26	1.27
X_{2c}	4.15	4.15	4.18
X_{3c}	13.94	13.98	13.80
L_{1v}	-11.76	-11.77	-11.81
L_{2v}	-8.64	-8.65	-8.60
L_{3v}	-1.04	-1.04	-1.06
L_{1c}	5.41	5.40	5.39
L_{2c}	7.15	7.14	7.10
L_{4c}	10.04	10.09	9.96
E_{gap}	1.29	1.26	1.27

Tabelle 7.2: Bandenergien von β -SiC an einigen Hochsymmetriepunkten. Verglichen sind die Ergebnisse (in eV) der PAW-Rechnung ($G_{\max} = 4.6 \sqrt{\text{Ry}}$), die einer PW-Rechnung ($G_{\max} = 8.0 \sqrt{\text{Ry}}$) und die Ergebnisse einer *ab-initio*-Berechnung mit Gaußorbitalen [Sab96]. Die Energien sind von unten durchnummeriert. Der oberste besetzte Zustand Γ_{4v} ist auf Null normiert.

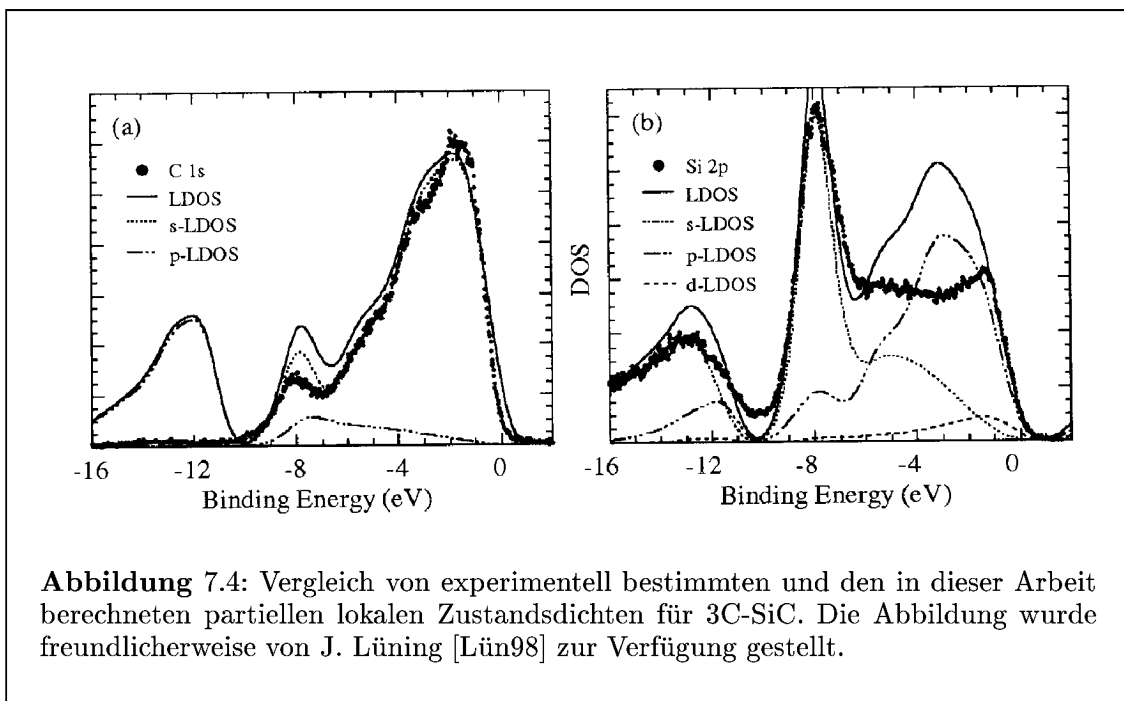
Zustandsdichten von β -SiC Für β -SiC wurden auch die elektronische Zustandsdichte (DOS) sowie partielle lokale Zustandsdichten (PLDOS) berechnet. Die Definition und die Berechnung der PLDOS wird in Anhang D beschrieben.

Die PLDOS wurde mit KB-Potentialen berechnet. Die Dichte, welche bei der theoretisch bestimmten Gitterkonstante vorliegt, wurde benutzt, um für $20 \times 20 \times 20$ k-Punkte die Eigenwerte zu bestimmen. Die Kugel, über die zur Bestimmung der PLDOS integriert wurde, hatte einen Radius von 0.46 Å. Dies ist etwa ein Viertel

des Abstandes nächster Nachbarn in SiC. Die berechneten Zustandsdichten wurden mit einer Gaußfunktion von 1 eV Breite geglättet.

J. Lüning [Lün98] hat die PLDOS experimentell mittels der Weichröntgenspektroskopie (SXE) untersucht (siehe Anhang D.1). Er bestimmte die p -PLDOS für Kohlenstoffatome (entsprechend der K-Kante des Röntgenspektrums) sowie eine Summe aus s - und d -PLDOS des Siliziumatoms (entsprechend der $L_{2,3}$ Kante).

Den Vergleich der berechneten und der gemessenen PLDOS zeigt Abbildung 7.4. Der Nullpunkt ist in beiden Fällen an das Valenzbandmaximum gelegt. Das ges-



sene Spektrum wurde nach Glg. (D.17) mit der dritten Potenz der Photonenenergie $h\nu$ korrigiert. Die berechnete Bandstruktur enthält den in der LDA üblichen Fehler der Eigenwerte. Daher wurde die Energieskala um etwa 10% gestreckt, um die prominentesten Strukturen in Übereinstimmung mit dem experimentellen Spektrum zu bringen. Die Matrixelemente $M(E)$ (vgl. Anhang D.1) sind als konstant über den betrachteten Energiebereich angenommen.

Für die p -symmetrische lokale Zustandsdichte am Kohlenstoffatom ergibt sich eine sehr gute Übereinstimmung der berechneten mit den gemessenen Daten. Im Falle der PLDOS am Siliziumatom werden zwar die im Experiment gefundenen Strukturen bei -8 eV und -12 eV durch die berechnete lokale s -Zustandsdichte am Siliziumatom wiedergegeben. Die flache Schulter jedoch im Bereich von -6 eV bis knapp an

die Oberkante des Valenzbandes findet keine Entsprechung im s -symmetrischen Anteil der LDOS von Silizium. In diesem Bereich wächst jedoch die d -Komponente etwas an. Da die Matrixelemente $M_{s \rightarrow p}$ und $M_{d \rightarrow p}$ verschieden sein können, ist es im Prinzip möglich, durch eine Linearkombination von s - und d -PLDOS eine qualitativ ähnliche Struktur zu erhalten. Quantitative Übereinstimmung wird aber nicht erzielt. Da die Aussagekraft von Matrixelementen, welche core-Zustände enthalten, in einer Pseudopotentialmethode sehr fraglich ist, kann hier keine Aussage über die relative Größe der beiden Matrixelemente gemacht werden. Als zusätzliche Schwierigkeit ergibt sich hier auch, daß die PLDOS vom gewählten Integrationsradius abhängt. Dies macht sich insbesondere beim d -symmetrischen Anteil stark bemerkbar. Je größer man den Radius wählt, desto stärker gewinnen Zustände an Gewicht, welche eigentlich vorwiegend an anderen Atomen lokalisiert sind.

Abschließend sei noch bemerkt, daß die experimentelle Verbreiterung der Lebensdauer nicht konstant ist, wie in der theoretischen Berechnung angenommen wird. Auf Grund der unterschiedlichen Lebensdauer der Endzustände in Abhängigkeit von der Bindungsenergie ist die Verbreiterung nahe am Valenzbandmaximum weniger ausgeprägt als für tiefere Energien. Eine entsprechende Berücksichtigung anstelle der verwendeten mittleren Verbreiterung würde an einigen Stellen die Übereinstimmung vermutlich noch weiter verbessern; z.B. wäre der s -PLDOS Peak von Si bei -12 eV weicher, der Anstieg der p -PLDOS von C an der Valenzbandkante jedoch schärfer.

Hexagonales 6H-SiC

Auch für die hexagonale 6H-Struktur des Siliziumkarbids wurden strukturelle Eigenschaften mit der PAW-Methode bestimmt. Hier wurden die Gitterkonstanten in der hexagonalen Ebene, a , und jene senkrecht dazu, c , unabhängig voneinander variiert und die Position minimaler Energie bestimmt. Allerdings wurden keine Relaxationen der Atome von den idealen Positionen erlaubt. Die Symmetriegruppe C_v^4 läßt bei der 6H-Struktur noch 5 interne Parameter zu, die eine Verschiebung der Ebenen, in welchen die Atome gestapelt sind, parallel zur c -Achse beschreiben. Die Abweichungen von den Idealpositionen sind aber sehr klein [FKBK00].

Die Rechnung wurde mit einem k -Punkt-Satz von $10 \times 10 \times 2$ Punkten durchgeführt; die Abschneideenergie betrug auch hier $G_{\max}^2 = 21$ Ry.

Die Gitterkonstanten, welche zur minimalen Gesamtenergie führen, zeigt Tabelle 7.3. Die Gitterkonstanten wurden zu $a=3.06$ Å bzw. $c=15.05$ Å bestimmt. Daraus ergibt sich ein c/a -Verhältnis von $c/a=4.91$. Alle Werte sind in hervorragender

Übereinstimmung sowohl mit experimentellen Daten als auch mit Ergebnissen der meisten angeführten theoretischen Berechnungen.

		Gitterkonstanten			Kohäsionsenergie
Methode		a [Å]	c [Å]	c/a	E_{coh} [eV]
Diese Arbeit	PAW	3.06	15.05	4.91	8.19
[Sab96]	GO	3.07	15.05	4.90	
[PCLC94]	NC-PP	3.08	15.12	4.91	
[KWB94]	NC-PP	3.03	14.88	4.91	
[FKBK00]	US-PP	3.06	15.02	4.91	8.33
experimentell ^a		3.07	15.08	4.91	

Tabelle 7.3: Gitterkonstanten und Kohäsionsenergie von 6H-SiC im Vergleich zu Ergebnissen anderer Rechnungen und zum Experiment. *GO* steht für Gauß-Orbitale, *NC-PP* für normerhaltende Pseudopotentiale, *US-PP* für Rechnungen mit Ultrasoftes Pseudopotentialen. Die Kohäsionsenergien enthalten nicht die Spin-Polarisations-Energie des Grundzustandsatoms. a: zitiert nach [Sab96]

Berechnet man die 3C-Struktur in derselben Zelle, um Fehler durch unterschiedliche Konvergenz der k-Punkt-Sätze auszuschließen, so ist die Energie der 6H-Struktur pro SiC-Paar um 0.7 meV niedriger als diejenige von β -SiC. Das Vorzeichen dieser Energie ist in Einklang mit anderen theoretischen Untersuchungen (siehe [Sab96], [FKBK00]). Furthmüller et. al. [FKBK00] geben einen Wert von 2.1 meV an für ihr Referenzpotential ($E_{cut} = 66$ Ry) und 3.8 meV für ein extrem weiches Potential. Allerdings ist dort eine Relaxation der Atome bezüglich der internen Freiheitsgrade der 6H-Struktur durchgeführt, ein „extrem dichtes“ k-Punkt Gitter benutzt und ein anderes Verfahren zu Bestimmung der minimalen Gitterkonstanten angewendet worden. Daher ist eine exakte Übereinstimmung hier nicht zu erwarten. Die Werte belegen jedoch daß die PAW-Methode in der Lage ist, kleine Energieunterschiede, wie sie hier bei SiC-Polytypen auftreten, dem Vorzeichen und der Größenordnung nach korrekt wiederzugeben. Die Relaxation der Atome in der 6H-Struktur würde zu einer Absenkung der Energie dieser Konfiguration und damit zu einer Vergrößerung der Energiedifferenz führen. Auch dies steht im Einklang mit den gewonnenen Daten.

7.2 Aluminiumnitrid und Galliumnitrid

Die Nitride GaN und AlN erfreuen sich in den letzten Jahren regen Interesses (siehe z.B. [ZMV99] und die darin enthaltenen Referenzen), da sie wegen ihrer physika-

lischen Eigenschaften vielversprechende Kandidaten für elektronische und optische Bauelemente sind. Beide Verbindungen weisen relativ große Bandlücken von 3.5 eV bzw. 6.2 eV auf. Ihre natürliche Erscheinungsform ist die hexagonale Wurtzitstruktur; beim Aufwachsen auf SiC(001)-Oberflächen entsteht jedoch die kubische Zinkblendestruktur. (Zinkblende- und Wurtzitstruktur entsprechen dem 3C- bzw. 2H-Typ in Abbildung 7.1.) Auf Grund der geringen Gitterfehlانpassung von nur ca. 0.5 % gilt SiC als aussichtsreiches Substrat für das epitaktische Aufwachsen defektarmer AlN-Schichten.

AlN

Für beide Strukturen von AlN wurden die Gleichgewichtsgitterkonstanten durch Minimierung der jeweiligen Gesamtenergie bestimmt.

Die Zinkblende-Variante verfügt nur über einen Parameter, die kubische Gitterkonstante a . Die Rechnung wurde sowohl mit dem PAW-Programm als auch mit normerhaltenden Pseudopotentialen durchgeführt. In beiden Fällen wurde dasselbe normerhaltende Pseudopotential für Al verwendet; dessen Konstruktionsparameter sind in [Ric96] zu finden. In der PAW-Rechnung wird das N_I -Potential aus Kapitel 4.6 verwendet; das normerhaltende Potential ist in [Eng92] beschrieben. In beiden Fällen wurden $6 \times 6 \times 6$ k-Punkte verwendet. Die Tabellen 7.4 bzw. 7.5 zeigen die Konvergenz mit der Anzahl der k-Punkte und der Abschneideenergie für die Rechnung mit dem PAW-Potential bzw. dem normerhaltenden Pseudopotential. Die strukturellen Größen verändern sich kaum noch.

		a [Å]	B [Mbar]	E_{coh} [eV]	B'
$G_{\max} = 5.2 \sqrt{\text{Ry}}$	kpt = $4 \times 4 \times 4$	4.309	2.06	8.262	3.87
	kpt = $5 \times 5 \times 5$	4.309	2.05	8.277	3.85
	kpt = $6 \times 6 \times 6$	4.309	2.05	8.280	3.89
kpt = $6 \times 6 \times 6$	$G_{\max} = 4.5$	4.307	2.01	8.148	3.99
	$G_{\max} = 5.2$	4.309	2.05	8.280	3.89
	$G_{\max} = 5.5$	4.309	2.05	8.284	3.86
	$G_{\max} = 6.0$	4.308	2.05	8.288	3.87

Tabelle 7.4: Konvergenz der Gitterkonstante a , des *bulk modulus* B und seiner Druckableitung B' sowie der Kohäsionsenergie E_{coh} von kubischem AlN. mit der Abschneideenergie $G_{\max}$ und der Anzahl der k-Punkte für das PAW-Potential. Die Kohäsionsenergien enthalten nicht die Spin-Polarisations-Energie des Grundzustandsatoms.

		a [Å]	B [Mbar]	E_{coh} [eV]	B'
$G_{\max} = 8.0 \sqrt{\text{Ry}}$	kpt = $4 \times 4 \times 4$	4.301	1.98	8.128	3.85
	kpt = $5 \times 5 \times 5$	4.301	1.98	8.141	3.86
	kpt = $6 \times 6 \times 6$	4.301	1.97	8.144	3.86
kpt = $6 \times 6 \times 6$	$G_{\max} = 5.0$	4.307	2.05	7.611	4.33
	$G_{\max} = 6.0$	4.305	1.96	7.884	4.13
	$G_{\max} = 7.0$	4.301	1.98	8.100	3.85
	$G_{\max} = 8.0$	4.301	1.97	8.144	3.86

Tabelle 7.5: Konvergenz der Gitterkonstante a , des *bulk modulus* B und seiner Druckableitung B' sowie der Kohäsionsenergie E_{coh} von kubischem AlN. mit der Abschneideenergie $G_{\max}$ und der Anzahl der k-Punkte für das KB-Potential. Die Kohäsionsenergien enthalten nicht die Spin-Polarisations-Energie des Grundzustandsatoms.

Tabelle 7.6 zeigt die Ergebnisse im Vergleich mit denen anderer Rechnungen und mit experimentellen Daten.

Methode		Gitterkonstante a [Å]	<i>bulk modulus</i> B [Mbar]	Kohäsionsenergie E_{coh}/atom [eV]
Diese Arbeit	PAW	4.31	2.06	8.28
Diese Arbeit	NC-PW	4.30	1.98	8.14
[SMV97]	NC-PW	4.35	1.99	
[GFB98]	US-PP	4.34	2.09	
[CG94]	LMTO	4.37	2.15	
[KLS96]	FP-LMTO	4.32	2.03	
experim. ^a		4.38		

Tabelle 7.6: Gitterkonstante a , *bulk modulus* B und Kohäsionsenergie E_{coh} von kubischem AlN (Zinkblendestruktur). *NC-PP* steht für normerhaltende Pseudopotentiale, *US-PP* für Rechnungen mit Ultrasoften Pseudopotentialen, *LMTO* für Linear Muffin Tin Orbital und *FP-LMTO* für die *Full-Potential*-Version davon. Die Kohäsionsenergien enthalten nicht die Spin-Polarisations-Energie des Grundzustandsatoms. a: nach [SMV97].

Die Gitterkonstanten der beiden eigenen Rechnungen stimmen beinahe überein und liegen etwa 2% unter dem experimentellen Wert, was etwa der LDA-üblichen Abweichung entspricht. Die Gitterkonstanten, welche in den anderen Rechnungen ermittelt werden, liegen zwischen diesen Werten. Die berechneten Kohäsionsenergien E_{coh} weichen um 0.14 eV oder 1.7% voneinander ab. Dieser Unterschied könnte

möglicherweise darauf beruhen, daß die PAW-Methode im Gegensatz zur Pseudopotentialmethode das Austausch-Korrelations-Potential der vollen Elektronendichte korrekt berücksichtigt. Dies würde auch erklären, daß der *bulk modulus* näher an dem Wert der FP-LMTO-Rechnung von [KLS96] als an denen der übrigen Pseudopotentialrechnungen liegt.

Die hexagonale Wurtzitstruktur besitzt im Gegensatz zum Zinkblende-Gitter außer der Gitterkonstante a in der hexagonalen Ebene auch noch die Gitterkonstante c senkrecht dazu (oder das c/a -Verhältnis) sowie einen weiteren Parameter u , welcher den interatomaren Abstand senkrecht zur hexagonalen Ebene bestimmt. Die Gitterkonstanten a und c sind hier unabhängig voneinander variiert und das Minimum durch einen polynomialen Fit bestimmt worden. Der Parameter u wurde nicht variiert, sondern der Wert $3/8$ angenommen, welcher bei ideal tetragonaler Koordination angenommen wird. Der k -Punkt-Satz umfaßte $6 \times 6 \times 4$ Elemente in der gesamten Brillouinzone, was zu 28 k -Punkten in deren irreduziblen Teil führt.

Die berechneten Werte der Gitterkonstanten und die entsprechenden Literaturwerte sind in Tabelle 7.7 zusammengetragen.

Methode		Gitterkonstanten			Kohäsionsenergie
		a [Å]	c [Å]	c/a	E_{coh}/atom [eV]
Diese Arbeit	PAW	3.05	4.95	1.619	8.30
Diese Arbeit	PW	3.05	4.94	1.619	8.16
[MSV96]	NC-PP	3.09	4.95	1.60	
[BFV97]	US-PP	3.08	4.98	1.619	
[GFB98]	US-PP	3.08	4.95	1.609	
[CG94]	LMTO	3.08	5.03	1.633 (ideal)	
[KLS96]	FP-LMTO	3.06	4.91	1.60	
[LB63]	exp.	3.11	4.98	1.60	

Tabelle 7.7: Gitterkonstanten und Kohäsionsenergie E_{coh} von hexagonalem AlN (Wurtzitstruktur). *NC-PP* steht für normerhaltende Pseudopotentiale, *US-PP* für Rechnungen mit Ultrasoft-Pseudopotentialen, *LMTO* für Linear Muffin Tin Orbital und *FP-LMTO* für die *Full-Potential*-Version davon. Die Kohäsionsenergien enthalten nicht die Spin-Polarisations-Energie des Grundzustandsatoms.

Wiederum stimmen die Gitterkonstanten aus der eigenen PAW- und Pseudopotentialrechnung hervorragend überein. Auch in diesem Fall sind sie aber etwas kleiner als die Werte aus den anderen angeführten *ab-initio*-Rechnungen. Die Kohäsionsenergien weisen wiederum geringe Unterschiede auf.

Der Energieunterschied zwischen den beiden Strukturen wurde zunächst mit der PAW- und der PP-Methode berechnet, wobei beide Formen in ihren *natürlichen* Gitterzellen berechnet wurden (d.h. Zinkblende in einer kubischen Zelle). Dies führt zwar zu einem systematischen Fehler durch die unterschiedliche Brillouinzoneintegration, doch sollte dieser in beiden Fällen gleich sein. Die Werte (siehe Tabelle 7.8) unterscheiden sich zwar um 5%; absolut jedoch sind dies nur 1.2 meV. Da die

	$E_{ZB} - E_{WZ}$ [meV]
PAW, hexagonale Zelle	19.9
PAW, kubische Zelle	21.1
NC-PP, kubische Zelle	20.0
[YLFZ92]	≈ 18

Tabelle 7.8: Energieunterschied (pro Atom) zwischen Zinkblende (ZB) - und Wurtzitstruktur (WZ) von AlN. Die Zinkblende-Verbindung wurde mit dem PAW-Programm in einer kubischen Zelle wie auch einer hexagonalen Zelle berechnet. Der geringe Unterschied geht auf die Unterschiede in der k-Punkt-Summation zurück.

Potentiale jeweils nur auf ca. 10 meV auskonvergiert sind, ist der relative Fehler akzeptabel. Um zu prüfen, wie groß der Fehler durch die unterschiedliche k-Punkt-Integration ist, wurde mit der PAW-Methode die Zinkblendestruktur auch in einer hexagonalen Zelle berechnet. Dies verändert den Energieunterschied nur um 1.2 meV.

Überraschenderweise konnte trotz des enormen Umfanges der vorhandenen Literatur nur bei Yeh et. al. [YLFZ92] ein Vergleichswert gefunden werden. Die dort von den Autoren bestimmte Energiedifferenz weicht um ca. 10% von den hier bestimmten Werten ab.

Insgesamt lassen sich also die Ergebnisse, welche mit den normerhaltenden Pseudopotentialen bei einem Abschneidewert von $G_{\max} = 8.0 \sqrt{\text{Ry}}$ erzielt wurden, gut mit der PAW-Methode bei einem Wert von $G_{\max} = 5.2 \sqrt{\text{Ry}}$ reproduzieren.

GaN

Analog zu AlN wurde auch GaN in Zinkblende- und Wurtzitstruktur berechnet. Diese Rechnungen wurden allerdings nur mit normerhaltenden Pseudopotentialen durchgeführt, da noch kein PAW-Potential für Gallium erzeugt wurde.

Das normerhaltende Pseudopotential für Gallium ist in [Ric96] beschrieben. Es weist oberhalb von etwa 5 eV einen *ghost-state* auf. Die Valenzbänder etwa von GaAs

werden jedoch angemessen beschrieben (siehe [Ric96]). Der dazu ausreichende Abschneidewert von $G_{\max} = 3.5 \sqrt{\text{Ry}}$ ist viel kleiner als der für das Stickstoffpotential benötigte Wert.

Das Ga-Pseudopotential behandelt die $3d$ -Elektronen als core-Elektronen. Es ist kontrovers diskutiert worden, ob dies eine angemessene Beschreibung liefert, oder ob die d -Elektronen explizit berücksichtigt werden müssen. Vogel et. al. kamen zu dem Ergebnis, daß „Kohäsionseigenschaften“ von GaN auch mit den $3d$ -Elektronen als Teil der eingefrorenen Rumpfdichte sehr gut reproduziert werden können [SMV97]. Furthmüller et. al. [FKBK00] sowie Wright und Nelson [WN95] insistieren, daß eine Behandlung als Valenzelektronen unerlässlich ist.

Die Rechnungen wurden wiederum mit $6 \times 6 \times 6$ k-Punkten und einem Abschneidewert $G_{\max} = 8.0 \sqrt{\text{Ry}}$ durchgeführt. Bei der hexagonalen Struktur wurde für den internen Parameter u , der schon bei AlN eingeführt wurde, auch hier der ideale Wert angenommen. Die Ergebnisse für Zinkblende- bzw. Wurtzitstruktur zeigen die Tabellen 7.9 und 7.10.

	Methode	Gitterkonstante a [Å]	<i>bulk modulus</i> B [Mbar]	Kohäsionsenergie E_{coh}/atom [eV]
Diese Arbeit	NC-PW	4.40	2.38	7.56
[SMV97]	NC-PP	4.50	1.96	
[FKBK00]	US-PP	4.46	2.01	7.01
[GFB98]	US-PP	4.46	1.83	
[CG94]	LMTO	4.46	1.84	
[KLS96]	FP-LMTO	4.46	2.01	
[WN94]	NC-PP	4.46	1.87	
experim. ^a	exp.	4.52		

Tabelle 7.9: Gitterkonstante, *bulk modulus* und Kohäsionsenergie von kubischem GaN (Zinkblendestruktur). Die Kohäsionsenergien enthalten nicht die Spin-Polarisations-Energie des Grundzustandsatoms. a: nach [SMV97].

Wie schon bei Aluminiumnitrid fallen auch hier die Gitterkonstanten um etwa 1% kleiner aus als bei vergleichbaren Rechnungen und ca. 2% kleiner als die experimentellen Werte. Der *bulk modulus* der kubischen Variante weicht noch deutlicher als bei Aluminiumnitrid von den übrigen Werten nach oben ab. Das c/a -Verhältnis der hexagonalen Struktur nimmt fast den idealen Wert (von 1.633) an, was sich mit den übrigen Werten mit Ausnahme von [KLS96] deckt. Auch das Verhältnis von

Methode		Gitterkonstanten			Kohäsionsenergie
		a [Å]	c [Å]	c/a	E_{coh}/atom [eV]
Diese Arbeit	NC-PP	3.11	5.08	1.634	7.55
[SMV97]	NC-PP	3.18	5.17	1.625	
[FKBK00]	US-PP	3.16	5.15	1.630	7.02
[BFV97]	US-PP	3.11	5.08	1.634	
[GFB98]	US-PP	3.15	5.14	1.631	
[CG94]	LMTO	3.13	5.11	1.633 (ideal)	
[KLS96]	FP-LMTO	3.06	4.91	1.60	
[WN94]	NC-PP	3.16	5.14	1.626	
[LB63]	exp.	3.19	5.19	1.628	

Tabelle 7.10: Gitterkonstanten und Kohäsionsenergie von hexagonalen GaN (Wurtzitstruktur). Die Kohäsionsenergien enthalten nicht die Spin-Polarisations-Energie des Grundzustandsatoms.

Zinkblende- zu Wurtzit-Gitterkonstante stimmt mit anderen Werte überein; der in dieser Arbeit erzielte Wert für a_{ZB}/a_{WZ} ist $1.000 \sqrt{2}$, der in [FKBK00] beträgt $0.999 \sqrt{2}$. Die Absolutwerte der Kohäsionsenergien unterscheiden sich von denen in [FKBK00] um ca. 8%. Die Differenz zwischen den Energien von Zinkblende- und Wurtzitstruktur zeigt Tabelle 7.11.

	Methode	ΔE_0 [meV / Atom]
Diese Arbeit	(NC-PP)	3.8
[WN94]	(NC-PP)	11
[FKBK00]	(US-PP)	5.3

Tabelle 7.11: Energieunterschied ΔE_0 zwischen Zinkblende- (ZB) und Wurtzitstruktur (WZ) von GaN. Angegeben ist die Energie pro Formeleinheit GaN.

Der in dieser Arbeit bestimmte Energieunterschied zwischen den beiden Strukturen ist sehr klein und beträgt 3.8 meV pro Atom. Allerdings muß man einkalkulieren, daß auf Grund der Benutzung unterschiedlicher Zellgeometrien und der damit verbundenen unterschiedlichen k-Punkt-Integration ein Fehler von ca. 1 meV (in Analogie zu AlN) und wegen der Ungenauigkeit der Fit-Prozedur nochmals ein Fehler von ca. 10% auftreten kann. Insofern und angesichts des relativ großen Unterschiedes der Gitterkonstanten ist die Übereinstimmung mit dem Wert von Furthmüller et. al.

[FKBK00] überraschend gut.

Wie schon gesagt, konnten Angaben über Energievergleiche für Aluminiumnitrid nicht gefunden werden; aus Abbildung 13 in [CG94] läßt sich jedoch schätzen, daß sich die Energiedifferenzen in AlN und GaN um einen Faktor 3-5 unterscheiden, was etwa in Übereinstimmung mit den in dieser Arbeit gewonnen Daten steht.

Über die Ursache der Abweichung der berechneten Gitterkonstanten können nur Mutmaßungen angestellt werden. Eine mögliche Ursache stellt die Behandlung der *d*-Elektronen als Rumpfelektronen dar. Nach Wright und Nelson [WN94, WN95] führt dieses Vorgehen zu einer kleineren Gitterkonstante für GaN. Allerdings kann das nicht erklären, warum in der hiesigen Arbeit auch die Gitterkonstanten für AlN kleiner ausfallen als in anderen Rechnungen.

Allerdings fällt auch auf, daß die meisten neueren Rechnungen zumindest eine partielle Rumpfladungskorrektur benutzen. Dies ist bei den in dieser Arbeit benutzten Pseudopotentialen für Ga, Al und N nicht der Fall, da bei dieser Rechnung die vorhandenen Pseudopotentiale benutzt wurden, welche ohne diese Korrektur konstruiert worden waren. Gerade angesichts der Ausdehnung der *d*-Wellenfunktionen stellt dies möglicherweise eine gravierende Fehlerquelle dar.

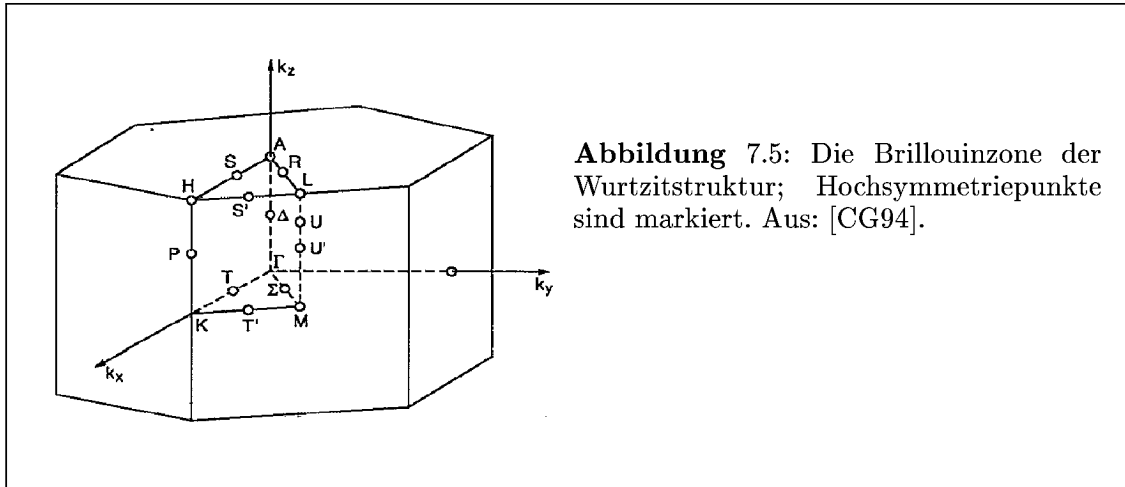
Wright und Nelson [WN94] ermittelten einen Unterschied von 4% zwischen den Gitterkonstanten, die sie mit einem Pseudopotential mit und einem ohne partielle Rumpfladungskorrektur berechneten. Diese Abweichung ist zwar viel größer als die hier aufgetretene, zeigt aber immerhin, *daß* die Behandlung der *d*-Elektronen wichtig für das Erhalten einer korrekten Gitterkonstante ist.

Bandstrukturen

Für die Wurtzitstrukturen von AlN und GaN wurden Bandstrukturen berechnet. Die Brillouinzone des hexagonalen Wurtzitgitters zeigt Abbildung 7.5.

Die Rechnungen wurden mit normerhaltenden Pseudopotentialen durchgeführt. Abbildung 7.6 stellt die Bandstrukturen sowie die jeweiligen Zustandsdichten gegenüber. Beide Halbleiter weisen eine direkte Bandlücke am Γ -Punkt auf. Die Größe der Bandlücke wird mit 4.78 eV für AlN bzw. 3.29 eV für GaN wie üblich in LDA-Rechnungen kleiner als die experimentellen Werte von 6.28 eV bzw. 3.5 eV [LB63]. Während der Wert für AlN in recht guter Übereinstimmung mit anderen LDA-Werten ist (z.B. 4.73 eV in [CG94]), ist der Wert für GaN deutlich größer als andere LDA-Werte (2.45 eV in [CG94]).

Das Leitungsband von GaN zeigt bei etwa 7.5 eV ein sehr flaches Band; dieses ist



ein Artefakt der Kleinman-Bylander-Faktorisierung (d.h. ein *ghost-state*).

Beide Verbindungen weisen zwischen ca. -16 eV und -12 eV ein relativ flaches unteres Valenzband (LVB) auf. Dieses entspricht den *s*-Elektronen des atomaren Stickstoffs, wie man aus der Lage der atomaren Eigenwerte (siehe Tabelle 7.12) vermuten kann. Die atomaren 3*d*-Zustände von Gallium liegen mit den 2*s*-Zuständen des Stickstoff

	ε [Ry]	$\varepsilon - \varepsilon_0$ [eV]
Aluminium		
3s	-0.575	-5.0
3p	-0.205	0.0
Stickstoff		
2s	-1.385	-16.1
2p	-0.562	-4.9
Gallium		
3d	-1.429	-16.7
4s	-0.673	-6.4
4p	-0.201	0.0

Tabelle 7.12: Atomare Valenzeigenwerte von Al, Ga und N. Die Energien sind absolut (in Ry) und relativ zu $\varepsilon_0 = \varepsilon_{4p}^{Ga} \approx \varepsilon_{3p}^{Al}$ (in eV) angegeben.

auf beinahe gleicher Höhe; daher ist zu erwarten, daß eine Hybridisierung zwischen diesen stattfinden wird. Aus diesem Grunde ist, wie schon erwähnt, die Notwendigkeit der expliziten Behandlung dieser Elektronen als Valenzelektronen diskutiert worden. Da Aluminium über keine *d*-Elektronen verfügt, liegt der nächste atomare

Zustand energetisch weit entfernt ($\varepsilon_{2p}^{Al} = -5.1$ Ry). Daher gibt es hier dieses Problem nicht.

Das obere Valenzband (UVB) entsteht durch Hybridisierung der $2p$ -Elektronen von Stickstoff und den sp -Zuständen von Aluminium bzw. Gallium.

Abbildung 7.7 zeigt die lokalen (LDOS) und die nach Drehimpulskomponenten aufgelösten partiellen lokale Zustandsdichten (PLDOS) bezüglich der Al- bzw. Ga- und der N-Atome. Die Definition und die Berechnung der LDOS bzw. PLDOS ist in Anhang D erläutert. Als Integrationsradius wurde hier für jeweils beide Atome der halbe Abstand zum nächsten Nachbarn gewählt.

Zunächst erkennt man, daß sich um die Stickstoffatome herum sehr viel mehr an Ladung befindet als an den Kationen (Ga, Al). Dies ist in qualitativer Übereinstimmung mit anderen *ab-initio*-Rechnungen, z.B. [GFB98].

In Übereinstimmung mit der obigen Überlegung sind die Elektronen des unteren Valenzbandes in beiden Verbindungen fast vollständig in der Nähe der Stickstoffatome anzutreffen und zeigen bezüglich dieser einen nahezu reinen s -Charakter.

Über den gesamten Energiebereich des oberen Valenzbandes dagegen hat die lokale Dichte am Stickstoffatom fast ausschließlich p -Charakter. Die partiellen Zustandsdichten am Kation sind kleiner und zeigen mit Zunahme der Energie eine Abnahme des s -Charakters und eine Zunahme des p -symmetrischen Anteiles; auch dies ist in Übereinstimmung mit der von den atomaren Zuständen ausgehenden Deutung. Für die Leitungsbänder ist ein beträchtlicher Anteil der Elektronen außerhalb der Integrationskugeln um die Atome zu finden; dies entspricht der Vorstellung von weitgehend delokalisierten Zuständen. Beim GaN ist das Valenzband fast völlig von den oben schon erwähnten *p-ghost-states* beherrscht.

Um einen Eindruck von der räumlichen Ladungsverteilung zu gewinnen, wird die Ladungsdichte entlang einer Linie durch einen der hexagonalen Stapelplätze parallel zur c -Achse betrachtet. Abbildung 7.8 illustriert die Lage der Atome. Die Ladungsdichte entlang dieser Linie ist in Abbildung 7.9 für zwei aufeinanderfolgende Gitterzellen aufgetragen. Die Lage der Atome ist markiert. Die Ladung ist zu einem Großteil um die Stickstoffatome zentriert. Bei den Atomen weist die Dichte jeweils ein Minimum auf, da es sich um die *Valenzladungs*dichte handelt; die core-Elektronen werden durch das Pseudopotential repräsentiert und hinterlassen daher eine „Lücke“. Der halbe nächste-Nachbar-Abstand ist durch die Pfeile unter der Achse markiert. Es ist plausibel, daß dies für die Stickstoffatome ein geeigneter Integrationsradius zur Bestimmung der PLDOS ist. Für Aluminium ragt dieser Bereich schon in den Ladungsbereich hinein, den man eher dem Stickstoff zuordnet. Für Al

kann man nur auf einer Seite eine signifikante Struktur ausmachen (Pfeil abwärts); jedoch reicht eine Kugel von diesem Radius in der anderen Richtung schon beinahe bis zum nächsten Atom und ist somit ungeeignet. Insofern ist die Wahl des Integrationsradius' für die Aluminiumatome mit einer gewissen Willkür behaftet; die Bestimmung der PLDOS an diesen Atomen führt aber zumindest qualitativ für etwas kleinere oder größere Radien zu ähnlichen Ergebnissen.

Die Ladungsdichte von GaN zeigt qualitativ ein ähnliches Verhalten wie die von AlN. Jedoch deuten die zu kleine Gitterkonstante und die ungewöhnliche nahe am experimentellen Wert liegende Bandlücke darauf hin, daß die in dieser Pseudopotentialrechnung erzielten Werte überprüft werden sollten. Die große Bandlücke läßt vermuten, daß auch die niedrigen unbesetzten Zustände von dem höher liegenden *ghost-state* beeinflußt werden.

Die Abweichung der Gitterkonstante läßt vermuten, daß in der Tat die 3*d*-Zustände von Gallium als Valenzelektronen behandelt werden müssen, so wie es schon in [FMS93, WN95, FKBK00] (und anderswo) gefordert wird.

Mit der PAW-Methode ist nach Generierung eines entsprechenden Gallium-Pseudopotentials die Voraussetzung dazu nun geschaffen.

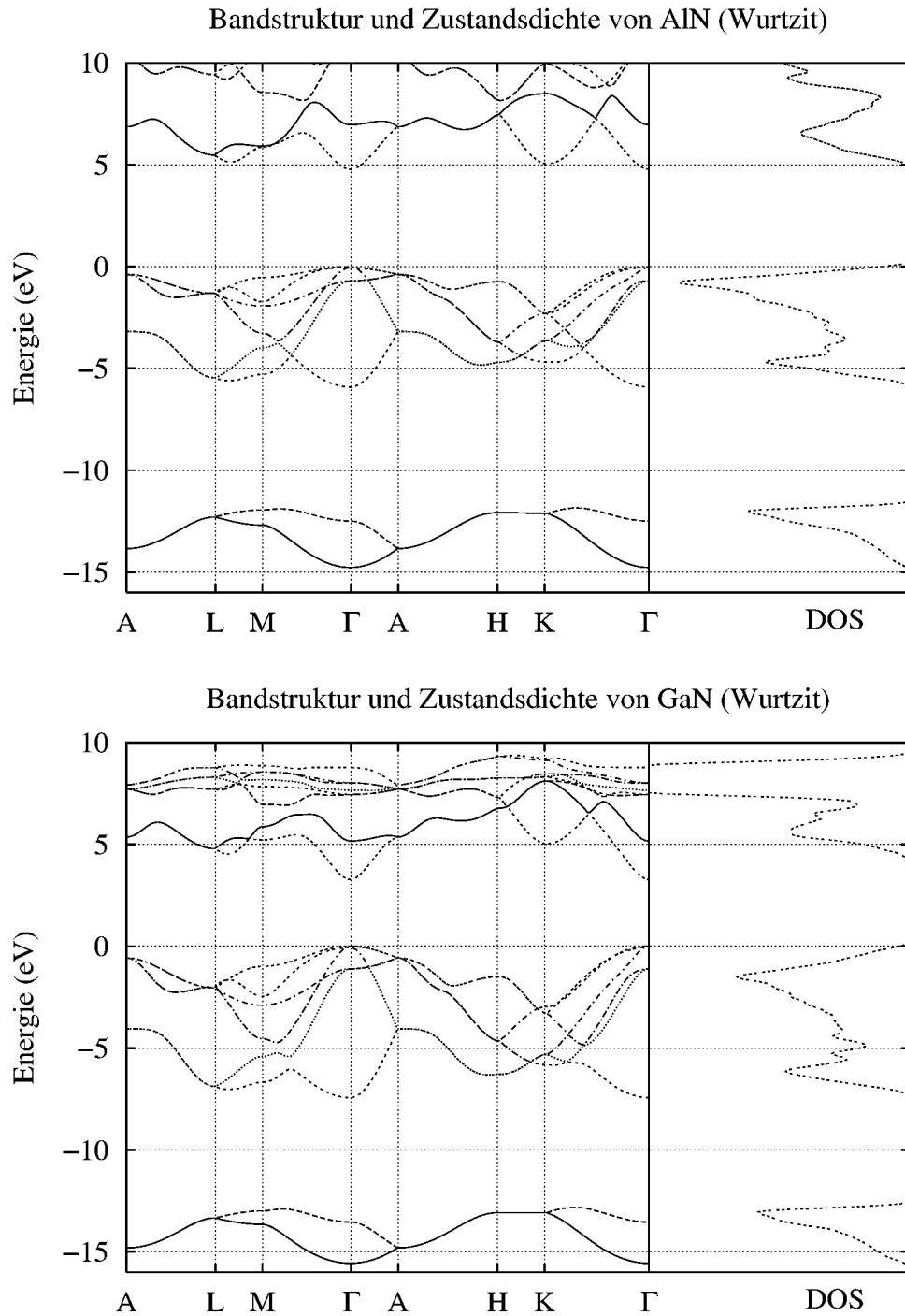


Abbildung 7.6: Bandstrukturen von AlN und GaN entlang der Hochsymmetrielinien der Brillouinzone. Die berechneten Zustandsdichten (DOS) sind im rechten Teil der Abbildungen dargestellt.

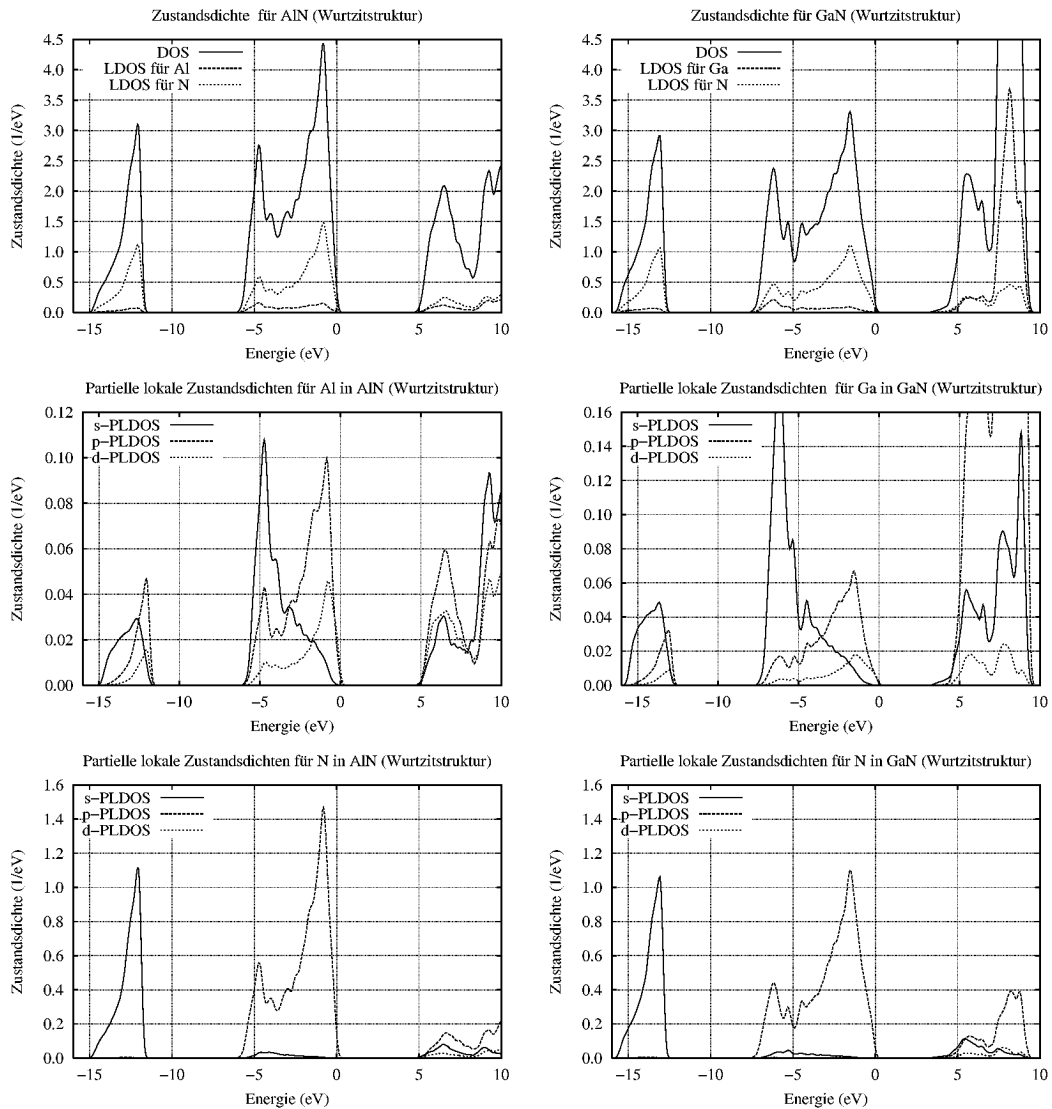


Abbildung 7.7: Lokale und partielle lokale Zustandsdichten von AlN (links) und GaN (rechts). Oben sind die (totale) Zustandsdichten und die lokale Zustandsdichte (für *ein* Atom) dargestellt. In der Mitte und unten sind die *s*-, *p*- und *d*-symmetrischen lokalen Zustandsdichten dargestellt.

	a_1	a_2	a_3
Al	$1/3$	$2/3$	$0/8$
N	$1/3$	$2/3$	$3/8$
Al	$2/3$	$1/3$	$4/8$
N	$2/3$	$1/3$	$7/8$

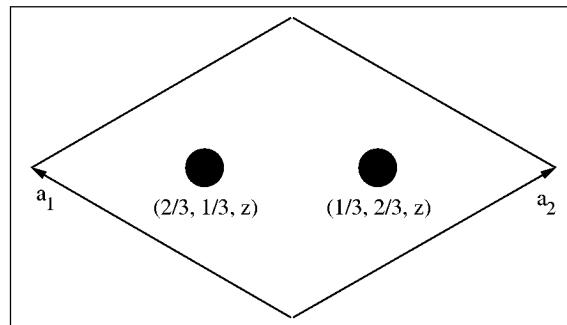


Abbildung 7.8: Koordinaten der Atome im Wurtzitgitter bezüglich der Basisvektor a_i des hexagonalen Gitters.

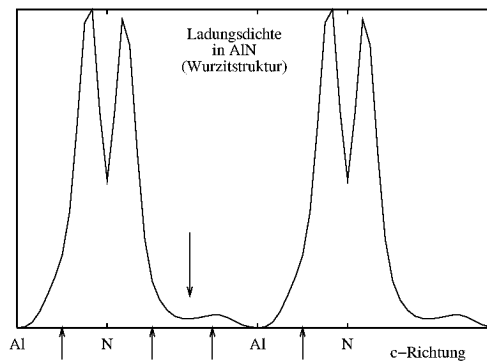


Abbildung 7.9: Ladungsdichte in AlN entlang einer Linie parallel zur c-Achse durch einen der hcp-Stapelpunkte. Die Pfeile unter der Achse markieren die Grenzen des Integrationsgebietes bei der Bestimmung der PLDOS.

Kapitel 8

Zusammenfassung

In dieser Arbeit wurde eine neue *ab-initio*-Methode für die Molekulardynamik entwickelt und implementiert. Sie ist an die PAW (*Projektor-Augmented-Wave*)-Methode von P. Blöchl [Blö94] angelehnt, verarbeitet aber auch Elemente einer Pseudoladungsmethode von M. Weinert [Wei81], welche im Umfeld der FLAPW (*Full-Potential Linearized Augmented Plane-Wave*)-Methode entstanden ist.

Die PAW-Methode ist wie die FLAPW-Methode eine Allelektronenmethode mit dem Charme der einfachen Darstellung des Hamiltonian, der eine zu normerhaltenden Kleinman-Bylander Pseudopotentialen analoge Form annimmt. Die PAW-Methode ermöglicht die effiziente Berechnung von Systemen, die praktisch alle Elemente des Periodensystems enthalten können, auch solche, welche die Berechnung mit normerhaltenden Pseudopotentialen erschweren oder gar unmöglich machen. Gegenüber diesen wird der nötige Aufwand an Rechenzeit drastisch reduziert, so daß Systeme, die *2p*-Elemente oder *3d*-Übergangsmetalle enthalten, mit der PAW-Methode jetzt effizient behandelt werden können. Durch die gegenüber der Pseudopotentialmethode genauere Beschreibung des atomnahen Raumbereiches wird zusätzlich eine höhere Genauigkeit erzielt.

Die Methode wurde ausgehend vom bestehenden *ab-initio*-Molekulardynamikprogramm EStCoMPP [BBA⁺00] implementiert, welches mit normerhaltenden Pseudopotentialen arbeitet. Die neue Implementierung erlaubt die Behandlung von physikalischen Systemen, in denen gleichzeitig Atome durch normerhaltende Pseudopotentiale und andere Atome durch den PAW-Ansatz beschrieben werden.

Des weiteren wurde ein Programm zur Erzeugung der Datensätze entwickelt, die zur Beschreibung eines Atomtyps benötigten werden. Wichtige Bestandteile dieser Datensätze sind lokalisierte atomare Wellenfunktionen und Projektorfunktionen. Es

wurde ein Verfahren entwickelt und erprobt, das Projektoren generiert, die zu einer numerisch stabilen Darstellung der PAW-Hamiltonians führt. Zusätzlich wurden weitere Programme zur Qualitätsanalyse dieser Datensätze erstellt. Das Generierungsverfahren kann möglicherweise bezüglich niedrigerer Abschneideenergien noch geringfügig verbessert werden, es liefert aber einfach und zuverlässig gute Datensätze.

Die zur Lösung der Kohn-Sham-Gleichungen erforderliche Matrixdiagonalisierung wird im PAW-Programm ebenso wie im ursprünglichen Programm **ESTCoMPP** iterativ mittels des Davidson-Verfahrens gelöst. Dazu werden Anfangslösungen benötigt. Da das bisherige Verfahren zur Erzeugung solcher Startvektoren für Systeme mit vielen Atomen oder Atomen, deren Zustände stark lokalisiert sind, an seine Grenzen stößt, wurde ein neues Verfahren entwickelt. Dieses erzeugt die Anfangslösungen als Linearkombination von atomaren Pseudoorbitalen (APSO), d.h. den atomaren Orbitalen von Pseudowellenfunktionen. Tests zeigten, daß dieses Verfahren sehr gute Startvektoren schon bei einer geringen Anzahl von Pseudoorbitalen liefert. Bei großen Einheitszellen oder dem Vorhandensein lokalisierter Zustände ist es damit die Methode der Wahl.

Um das Konvergenzverhalten weiter zu verbessern, wurden Tests mit verschiedenen Präkonditionierern für das Davidson-Verfahren sowie mit dem Davidson-Jacobi-Verfahren durchgeführt. In ihnen konnten Ansätze für künftige Verbesserungen ermittelt werden. Insbesondere erscheint eine bessere Feinabstimmung der ineinander verschachtelten iterativen Verfahren vielversprechend.

Zunächst wurden mit der PAW-Methode diverse kleinere Testsysteme untersucht. Als erster Schritt dazu wurden die erforderlichen Datensätze zur Charakterisierung der Atome erzeugt und getestet, bevor die *eigentlichen* Tests und Anwendungen begonnen werden konnten. Die berechnete Bindungslänge, die Dissoziationsenergie und die Schwingungsfrequenz eines Stickstoffmoleküls stehen in exzellenter Übereinstimmung mit anderen Berechnungen und reproduzieren die experimentellen Daten im Rahmen der üblichen Abweichungen. Ebenso werden Gitterkonstante, *bulk modulus* und Kohäsionsenergie von Diamant in sehr guter Übereinstimmung mit den Ergebnissen anderer aktueller *ab-initio*-Methoden berechnet. Der Vergleich verschiedener hypothetischer Kohlenstoffgitter stehen ebenfalls im Einklang mit früheren Rechnungen. Interessanterweise finden wir durch einen Energievergleich zwischen Diamant und Graphit, daß im Gegensatz zur Wirklichkeit nicht Graphit, sondern Diamant die stabile Modifikation des Kohlenstoffs ist. Wir führen dies darauf zurück, daß die van-der-Waals-Wechselwirkung zwischen den Graphitebenen in der Lokale-

Dichte-Näherung nicht gut genug beschrieben wird und sehen es als zukunftsweisende Herausforderung an, Dichtefunktionale zu entwickeln, die dieses leisten. Die Untersuchung von Siliziumkarbid zeigt, daß nicht nur strukturelle, sondern auch elektronische Eigenschaften, wie z.B. die Bandstruktur oder die Zustandsdichte, mit der PAW-Methode gut beschrieben werden; der numerische Aufwand ist jedoch gegenüber derselben Berechnung mit normerhaltenden Pseudopotentialen um ca. 80% reduziert.

Um die Effizienz des Programms zu testen, wurde auf demselben Computer ein Testsystem sowohl mit dem PAW-Programm als auch mit dem FLAPW-Programm FLEUR berechnet. Es handelte sich dabei um eine (4×2) Cu(110) Oberfläche, die durch eine Superszelle mit 72 Atomen und 792 Valenzelektronen beschrieben wird. Bei vergleichbaren Voraussetzungen ist die PAW-Methode in diesem Fall pro Selbstkonsistenziteration um einen Faktor drei bis vier schneller als FLEUR. Allerdings verfügt FLEUR über erhebliche Leistungsreserven durch konsequente Ausnutzung der Symmetrieeigenschaften und der halbunendlichen Beschreibung des Vakuums. Würden wir diese Symmetrien auch im PAW-Programm ausnutzen, würde sich die Rechenzeit um einen weiteren Faktor vier reduzieren. Ein Vergleich mit der Berechnung desselben Systems mit einem Programm für Ultrasofte Pseudopotentiale durch Y. Morikawa [Mor] zeigt, daß eine weitere Steigerung der Geschwindigkeit bis zu einem Faktor zwei möglich ist, wenn man (i) mit der in Anhang C skizzierten Methode die ähnliche Struktur von nichtlokalen Potential und Überlappmatrix noch besser ausnutzt, um das Davidson-Verfahren weiter zu beschleunigen, oder (ii) statt des Davidson-Verfahrens die RMM-DIIS-Methode zur Bestimmung der Eigenzustände verwendet und dadurch Zeit für die Orthogonalisierung einsparen kann und (iii) die Auswertung der Projektor-Wellenfunktions-Produkte im Ortsraum implementiert. Letzteres führt zu einem linearen Skalierungsverhalten mit der Anzahl der Atome für diesen Teil des Programms.

Es steht also durch diese Arbeit ein Programm bereit, das es ermöglicht, *ab-initio*-Molekulardynamikrechnungen von großen Systemen bei Erfüllung hoher Anforderungen an die Genauigkeit effizient durchzuführen und somit wichtige Beiträge zum mikroskopischen Verständnis von Festkörpereigenschaften und Festkörperprozessen leisten zu können.

Anhang A

Formelsammlung

sphärische Besselfunktionen $j_l(x)$:

Differentialgleichung:

$$\left[- \left(\frac{d}{dr} \right)^2 + \frac{l(l+1)}{r^2} \right] \left(r j_l(q r) \right) = q^2 \left(r j_l(q r) \right) \quad (\text{A.1})$$

Rayleigh's Formel:

$$j_l(x) = x^l \left[-\frac{1}{x} \frac{d}{dx} \right]^l \left(\frac{\sin(x)}{x} \right) \quad (\text{A.2})$$

$$j_0(x) = \frac{\sin(x)}{x} \quad (\text{A.3})$$

$$j_1(x) = \frac{\sin(x)}{x^2} - \frac{\cos(x)}{x} \quad (\text{A.4})$$

$$j_2(x) = \left[\frac{3}{x^3} - \frac{1}{x} \right] \sin(x) - \left[\frac{3}{x^2} \right] \cos(x) \quad (\text{A.5})$$

$$j_3(x) = \left[\frac{15}{x^4} - \frac{6}{x^2} \right] \sin(x) + \left[-\frac{15}{x^3} + \frac{1}{x} \right] \cos(x) \quad (\text{A.6})$$

$$j_4(x) = \left[\frac{105}{x^5} - \frac{45}{x^3} + \frac{1}{x} \right] \sin(x) + \left[-\frac{105}{x^4} + \frac{10}{x^2} \right] \cos(x) \quad (\text{A.7})$$

Rekursionsformel:

$$j_{l+1}(x) = \frac{2l+1}{x} j_l(x) - j_{l-1}(x) \quad (\text{A.8})$$

Asymptotik:

$$j_l(x) \xrightarrow{x \rightarrow \infty} \frac{\sin \left(x - \frac{1}{2} l \pi \right)}{x} \quad (\text{A.9})$$

$$\frac{j_l(x)}{x} \xrightarrow{x \rightarrow 0} \frac{1}{(2l+1)!!} \quad (\text{A.10})$$

Orthogonalitätsrelation: $(q_1, q_2 > 0)$

$$q_1^2 \int_0^\infty dr r^2 j_l(q_1 r) j_l(q_2 r) = \frac{\pi}{2} \delta(q_1 - q_2) \quad (\text{A.11})$$

Legendre Polynome $P_l(x)$:

Differentialgleichung:

$$\left[(1-x^2) \left(\frac{d}{dx} \right)^2 - 2x \left(\frac{d}{dx} \right) + l(l+1) \right] P_l(x) = 0 \quad (\text{A.12})$$

Formel von Rodrigues:

$$P_l(x) = \frac{1}{2^l l!} \left(\frac{d}{dx} \right)^l (x^2 - 1)^l \quad (\text{A.13})$$

$$P_0(x) = 1 \quad (\text{A.14})$$

$$P_1(x) = x \quad (\text{A.15})$$

$$P_2(x) = \frac{1}{2} (3x^2 - 1) \quad (\text{A.16})$$

$$P_3(x) = \frac{1}{2} (5x^3 - 3x) \quad (\text{A.17})$$

$$P_4(x) = \frac{1}{2 \cdot 4} (5 \cdot 7x^4 - 2 \cdot 3 \cdot 5x^2 + 1 \cdot 3) \quad (\text{A.18})$$

Rekursionsformel:

$$(l+1) P_{l+1}(x) = (2l+1) x P_l(x) - l P_{l-1}(x) \quad (\text{A.19})$$

Orthogonalitätsrelation:

$$\int_{-1}^{+1} P_l(x) P_{l'}(x) dx = \frac{2\delta_{ll'}}{2l+1} \quad (\text{A.20})$$

zugeordnete Legendre Polynome $P_l^m(x)$:

Differentialgleichung:

$$\left[(1-x^2) \left(\frac{d}{dx} \right)^2 - 2x \left(\frac{d}{dx} \right) + l(l+1) - \frac{m^2}{1-x^2} \right] P_l^m(x) = 0 \quad (\text{A.21})$$

Formel von Rodrigues:

$$P_l^m(x) = \frac{(1-x^2)^{\frac{m}{2}}}{2^l l!} \left(\frac{d}{dx} \right)^{l+m} (x^2 - 1)^l \quad (\text{A.22})$$

Orthogonalitätsrelation:

$$\int_{-1}^{+1} P_l^m(x) P_{l'}^m(x) dx = \frac{2\delta_{ll'} (l+m)!}{(2l+1) (l-m)!} \quad (\text{A.23})$$

$$\int_{-1}^{+1} P_l^m(x) P_l^{m'}(x) \frac{dx}{1-x^2} = \frac{\delta_{mm'} (l+m)!}{m (l-m)!} \quad (\text{A.24})$$

komplexe Kugelflächenfunktionen $Y_{lm}(\theta, \varphi)$:

Definitionsformel:

$$Y_{lm}(\theta, \varphi) = (-1)^m \sqrt{\frac{(2l+1)(l-m)!}{4\pi(l+m)!}} P_l^m(\cos\theta) e^{im\varphi} \quad (\text{A.25})$$

Orthogonalitätsrelation:

$$\int_0^{2\pi} d\varphi \int_0^\pi \sin\theta d\theta Y_{lm}^*(\theta, \varphi) Y_{l'm'}(\theta, \varphi) = \delta_{ll'} \delta_{mm'} \quad (\text{A.26})$$

reelle Kugelflächenfunktionen $\mathcal{Y}_{lm}(\theta, \varphi)$:

Definitionsformel:

$$\begin{aligned} \mathcal{Y}_{lm} &= \frac{(-1)^m}{\sqrt{2}} (Y_{lm} + Y_{lm}^*) = \sqrt{2} (-1)^m \operatorname{Re}(Y_{lm}) && \text{für } (m > 0) \\ \mathcal{Y}_{lm} &= Y_{l0} && \text{für } (m = 0) \\ \mathcal{Y}_{l-m} &= \frac{1}{\sqrt{2}i} (Y_{lm} - Y_{lm}^*) = \sqrt{2} \operatorname{Im}(Y_{lm}) && \text{für } (m > 0) \end{aligned} \quad (\text{A.27})$$

Orthogonalitätsrelation:

$$\int_0^{2\pi} d\varphi \int_0^\pi \sin\theta d\theta \mathcal{Y}_{lm}(\theta, \varphi) \mathcal{Y}_{l'm'}(\theta, \varphi) = \delta_{ll'} \delta_{mm'} \quad (\text{A.28})$$

Tabelle der Kugelflächenfunktionen für $l \leq 3$:

l	m	$\sqrt{4\pi} r^l Y_{l\pm m}(\vec{r})$	$\sqrt{4\pi} r^l \mathcal{Y}_{lm}(\vec{r})$	$\sqrt{4\pi} r^l \mathcal{Y}_{l-m}(\vec{r})$
0	0	1	1	
1	1	$\mp \sqrt{\frac{3}{2}} (x \pm iy)$	$\sqrt{3}x$	$\sqrt{3}y$
1	0	$\sqrt{3}z$	$\sqrt{3}z$	
2	2	$\frac{1}{2} \sqrt{\frac{15}{2}} (x \pm iy)^2$	$\frac{1}{2} \sqrt{15} (x^2 - y^2)$	$\sqrt{15} (xy)$
2	1	$\mp \sqrt{\frac{15}{2}} z (x \pm iy)$	$\sqrt{15} (xz)$	$\sqrt{15} (yz)$
2	0	$\frac{1}{2} \sqrt{5} (3z^2 - r^2)$	$\frac{1}{2} \sqrt{5} (3z^2 - r^2)$	
3	3	$\mp \frac{1}{4} \sqrt{35} (x \pm iy)^3$	$\frac{1}{4} \sqrt{70} (x^2 - 3y^2) x$	$\frac{1}{4} \sqrt{70} (y^2 - 3x^2) y$
3	2	$\frac{1}{2} \sqrt{\frac{105}{2}} (x \pm iy)^2 z$	$\frac{1}{2} \sqrt{105} (x^2 - y^2) z$	$\sqrt{105} xyz$
3	1	$\mp \frac{1}{4} \sqrt{21} (5z^2 - r^2) (x \pm iy)$	$\frac{1}{4} \sqrt{42} (5z^2 - r^2) x$	$-\frac{1}{4} \sqrt{42} (5z^2 - r^2) y$
3	0	$\frac{1}{2} \sqrt{7} (5z^2 - 3r^2) z$	$\frac{1}{2} \sqrt{7} (5z^2 - 3r^2) z$	

(A.29)

Vermischtes zu P_l , P_l^m und Y_{lm} :

$$P_l(-x) = (-1)^l P_l(x) \quad (\text{A.30})$$

$$P_l^m(-x) = (-1)^{l+m} P_l^m(x) \quad (\text{A.31})$$

$$Y_{lm}(\pi - \theta, \pi + \varphi) = (-1)^l Y_{lm}(\theta, \varphi) \quad (\text{A.32})$$

$$P_l^{-m}(x) = (-1)^m \frac{(l-m)!}{(l+m)!} P_l^m(x) \quad (\text{A.33})$$

$$Y_{l-m}(\theta, \varphi) = (-1)^m Y_{lm}^*(\theta, \varphi) \quad (\text{A.34})$$

$$(2l+1) P_l(\cos \gamma_{\vec{x}, \vec{x}'}) = \begin{cases} \sum 4\pi Y_{lm}^*(\Omega_{\vec{x}}) Y_{lm}(\Omega_{\vec{x}'}) \\ \sum_m^m 4\pi Y_{lm}(\Omega_{\vec{x}}) Y_{lm}^*(\Omega_{\vec{x}'}) \\ \sum_m^m 4\pi \mathcal{Y}_{lm}(\Omega_{\vec{x}}) \mathcal{Y}_{lm}(\Omega_{\vec{x}'}) \end{cases} \quad (\text{A.35})$$

Darstellung von Ebenen Wellen durch Bessel- und komplexe Kugelflächenfunktionen:

$$e^{i \vec{k} \vec{r}} = 4\pi \sum_{l,m} i^l Y_{lm}^*(\Omega_{\vec{k}}) Y_{lm}(\Omega_{\vec{r}}) j_l(kr) \quad (\text{A.36})$$

$$= 4\pi \sum_{l,m} i^l Y_{lm}(\Omega_{\vec{k}}) Y_{lm}^*(\Omega_{\vec{r}}) j_l(kr) \quad (\text{A.37})$$

Drehmatrizen:

$$Y_L(\mathbf{S}^{-1} \hat{r}) = \sum_{m'} Y_{lm'}(\hat{r}) R_{m'm}^{(l)}(S) \quad (\text{A.38})$$

$$R_{m'm}^{(l)}(S) = \int d\Omega Y_{lm'}(\hat{r}) Y_{lm}(\mathbf{S}^{-1} \hat{r}) \quad (\text{A.39})$$

$$R_{m'm}^{(l)}(S) = R_{mm'}^{*(l)}(S^{-1}) \quad (\text{A.40})$$

$$R_{m''m}^{(l)}(S_1 \circ S_2) = \sum_{m'} R_{m''m'}^{(l)}(S_1) R_{m'm}^{(l)}(S_2) \quad (\text{A.41})$$

Gauß-Hermite-Integration:

$$\int_{-\infty}^{\infty} dr e^{-r^2} f(r) = \sum_{i=1}^N \omega_{i,N} f(r_{i,N}) \quad (\text{A.42})$$

Hermite-Polynome H_N :

$$H_N(r) = (-1)^N e^{r^2} \left(\frac{d}{dr} \right)^N \left[e^{-r^2} \right] \quad (\text{A.43})$$

Nullstellen $r_{i,N}$:

$$H_N(r_{i,N}) = 0 \quad (\text{A.44})$$

Gewichte $\omega_{i,N}$:

$$\omega_{i,N} = \frac{2^N \sqrt{\pi}}{N!} (2 - \delta_{r_{i,N},0}) \left[\frac{d}{dr} H_N(r) \Big|_{r=r_{i,N}} \right]^{-2} \quad (\text{A.45})$$

Die Ewaldsumme:

$$\sum_{\vec{R}} \frac{(1 - \delta_{\vec{r}, \vec{R}})}{|\vec{r} - \vec{R}|} = \sum_{\vec{G}} \frac{1}{\Omega} \frac{4\pi}{G^2} e^{-i\vec{G}\vec{r}} \quad (\text{A.46})$$

$$\begin{aligned} \frac{1}{\sqrt{\Omega}} \sum_{\vec{R}} \left\langle \vec{G} \left| \frac{(1 - \delta_{\vec{r}, \vec{R}})}{|\vec{r} - \vec{R}|} \right\rangle_{\Omega} &= \frac{1}{\sqrt{\Omega}} \left\langle \vec{G} \left| \frac{1}{r} \right\rangle \right. \\ &= \frac{1}{\Omega} \int d^3r \frac{1}{r} e^{i\vec{G}\vec{r}} = \frac{1}{\Omega} \frac{4\pi}{G^2} \end{aligned} \quad (\text{A.47})$$

$$\begin{aligned} \sum_{\vec{R}} \frac{(1 - \delta_{\vec{r}, \vec{R}})}{|\vec{r} - \vec{R}|} &= \frac{4\pi}{\Omega} \sum_{\vec{G}} \frac{1}{G^2} \cos(\vec{G}\vec{r}) e^{-\frac{G^2}{4\eta^2}} - \frac{\pi}{\eta^2 \Omega} \\ &+ \sum_{\vec{R}} \frac{\text{erfc}\left(\eta |\vec{r} - \vec{R}|\right)}{|\vec{r} - \vec{R}|} (1 - \delta_{\vec{r}, \vec{R}}) - 2 \frac{\eta}{\sqrt{\pi}} \delta_{\vec{r}, \vec{0}} \quad (\text{A.48}) \end{aligned}$$

Die Fehlerfunktion:

$$\text{erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z dt e^{-t^2} \quad (\text{A.49})$$

$$\text{erfc}(z) = \frac{2}{\sqrt{\pi}} \int_z^\infty dt e^{-t^2} = 1 - \text{erf}(z) \quad (\text{A.50})$$

$$e^{z^2} \text{erfc}(z) \leq 1 \quad (\text{A.51})$$

$$\lim_{z \rightarrow \infty} e^{z^2} \text{erfc}(z) = 0 \quad (\text{A.52})$$

Poisson-Gleichung:

$$u(\vec{r}) = \int d^3r' \frac{f(\vec{r}')}{|\vec{r} - \vec{r}'|} \Leftrightarrow \Delta u(\vec{r}) = -4\pi f(\vec{r}) \quad (\text{A.53})$$

Anhang B

Inverse der PAW-Transformation

Die Abbildung $\mathcal{T}$ ist definiert durch Gleichung 3.24:

$$\mathcal{T} = 1 + \sum_{\mu, \gamma} \left(|\phi_\gamma^\mu\rangle - |\tilde{\phi}_\gamma^\mu\rangle \right) \langle \tilde{p}_\gamma^\mu | \quad (\text{B.1})$$

Die Inverse ist:

$$\mathcal{T}^{-1} = 1 + \sum_{\mu, \gamma} \left(|\tilde{\phi}_\gamma^\mu\rangle - |\phi_\gamma^\mu\rangle \right) \langle p_\gamma^\mu | \quad \text{mit} \quad \langle p_\gamma^\mu | = \sum_{\gamma'} (\langle \tilde{p} | \phi \rangle)_{\gamma, \gamma'}^{-1} \langle \tilde{p}_{\gamma'}^\mu | \quad (\text{B.2})$$

Zum Beweis zeigt man, daß $\mathcal{T}^{-1}\mathcal{T} = \mathbf{1}$ gilt. Zunächst betrachtet man die beiden folgenden Produkte, welche in dem Term auftreten, der quadratisch in den Projektoren ist:

$$\langle p_\gamma^\mu | \tilde{\phi}_{\gamma'}^\mu \rangle = \sum_{\gamma''} (\langle \tilde{p}^\mu | \phi^\mu \rangle)_{\gamma, \gamma''}^{-1} \langle \tilde{p}_{\gamma''}^\mu | \tilde{\phi}_{\gamma'}^\mu \rangle = (\langle \tilde{p}^\mu | \phi^\mu \rangle)_{\gamma, \gamma'}^{-1} \quad (\text{B.3})$$

und

$$\langle p_\gamma^\mu | \phi_{\gamma'}^\mu \rangle = \sum_{\gamma''} (\langle \tilde{p}^\mu | \phi^\mu \rangle)_{\gamma, \gamma''}^{-1} \langle \tilde{p}_{\gamma''}^\mu | \phi_{\gamma'}^\mu \rangle = \delta_{\gamma, \gamma'} \quad (\text{B.4})$$

wobei erinnert sei an

$$\langle \tilde{p}_\gamma^\mu | \tilde{\phi}_{\gamma'}^\mu \rangle = \delta_{\gamma, \gamma'}. \quad (\text{B.5})$$

Mischterme in μ, μ' treten nicht auf, weil einerseits $\tilde{p}_\gamma^\mu$ in S_μ lokalisiert ist, andererseits die Differenz $\phi_\gamma^{\mu'} - \tilde{\phi}_\gamma^{\mu'}$ außerhalb von $S_{\mu'}$ verschwindet.

Damit ergibt sich der quadratische Term zu:

$$\begin{aligned}
& \sum_{\gamma, \gamma', \mu, \mu'} \left[\left(|\tilde{\phi}_\gamma^\mu\rangle - |\phi_\gamma^\mu\rangle \right) \langle p_\gamma^\mu | \right] \left[\left(|\phi_{\gamma'}^{\mu'}\rangle - |\tilde{\phi}_{\gamma'}^{\mu'}\rangle \right) \langle \tilde{p}_{\gamma'}^{\mu'} | \right] \\
&= \sum_{\gamma, \gamma', \mu} \left(|\tilde{\phi}_\gamma^\mu\rangle - |\phi_\gamma^\mu\rangle \right) \left(\delta_{\gamma, \gamma'} - (\langle \tilde{p}^\mu | \phi^\mu \rangle)_{\gamma, \gamma'}^{-1} \right) |\tilde{\phi}_{\gamma'}^\mu\rangle \\
&= - \sum_{\mu, \gamma} \left(|\phi_\gamma^\mu\rangle - |\tilde{\phi}_\gamma^\mu\rangle \right) \langle \tilde{p}_\gamma^\mu | - \sum_{\mu, \gamma} \left(|\tilde{\phi}_\gamma^\mu\rangle - |\phi_\gamma^\mu\rangle \right) \langle p_\gamma^\mu | \quad (\text{B.6})
\end{aligned}$$

Er hebt sich also gerade mit den (in den Projektoren) linearen Termen von $\mathcal{T}^{-1}\mathcal{T}$ auf, so daß als Resultat nur der Einsoperator bleibt, womit die Behauptung bewiesen wäre.

Anhang C

Verbesserungsmöglichkeit für das Davidson-Verfahren

In dem in Kapitel 5.2.2.4 beschriebenen Algorithmus wird der Vektor $\mathbf{O}\mathbf{v}_i$ tatsächlich aufgestellt. Das ist eigentlich nicht nötig. Dieser Weg wurde eingeschlagen, weil es der Struktur des vorhandenen Programmes besser entsprach. Definiert man

$$\mathbf{P}^\mu = \left(\langle \vec{k} + \vec{G} | \tilde{p}_\gamma^\mu \rangle \right)_{\vec{G}, \gamma} \quad (\text{Dimension } N_{\vec{G}} \times N_\gamma^\mu)$$

für alle Atome μ und

$$\mathbf{B}^\mu = \mathbf{P}^{\mu+} \mathbf{V} \quad (\text{Dimension } N_\gamma^\mu \times N_{state})$$

so kann man schreiben:

$$\mathbf{O}\mathbf{V} = \sum_{\mu} \mathbf{P}^\mu \mathbf{q}^\mu \mathbf{B}^\mu \text{ und} \quad (\text{C.1})$$

$$\mathbf{H}\mathbf{V} = \mathbf{H}_0 \mathbf{V} + \sum_{\mu} \mathbf{P}^\mu \mathbf{D}^\mu \mathbf{B}^\mu \quad (\text{C.2})$$

Rechenzeitaufwendig sind vor allem das Bilden von $\mathbf{B}^\mu$ und Multiplikationen mit $\mathbf{P}^\mu$, da die Anzahl der Operationen dabei proportional zur Anzahl $N_{\vec{G}}$ der Ebenen Wellen ist. Es werden also pro Atom drei derartige Multiplikationen ausgeführt. (Im Falle der Cu-Oberfläche aus Kapitel 6.3 machen diese insgesamt 120 Sekunden pro Davidsonschrift aus.)

Hier soll kurz eine Variante skizziert werden, welche mit zwei derartigen Multiplikationen auskommt; sie sollte also ca. 30% schneller sein, wenn die Anzahl der Ebenen Wellen sehr viel größer ist als die Anzahl der Zustände N_{state} und die der Projektoren $\sum_{\mu} N_\gamma^\mu$.

- Berechne $\mathbf{H}_0 \mathbf{V}$, wobei $H_0 = T + V_{\text{lok}}$ das nichtlokale Potential *nicht* enthält
- Berechne $\mathbf{P}^\mu = \left(\langle \vec{k} + \vec{G} | \tilde{p}_\gamma^\mu \rangle \right)_{\vec{G}, \gamma}$ für alle Atome μ .
- Berechne $\mathbf{B}^\mu = \mathbf{P}^{\mu+} \mathbf{V}$ für alle Atome μ .

- Stelle Unterraummatrizen auf:

$$\mathbf{M}_\mathbf{H} = \mathbf{V}^+ (\mathbf{H}_0 \mathbf{V}) + \sum_\mu \mathbf{B}^{\mu+} \mathbf{D}^\mu \mathbf{B}^\mu \text{ und}$$

$$\mathbf{M}_\mathbf{O} = \sum_\mu \mathbf{B}^{\mu+} \mathbf{q}^\mu \mathbf{B}^\mu$$

- Löse verallgemeinertes Unterraum-Eigenwertproblem

$$\mathbf{M}_\mathbf{H} \mathbf{U} = \mathbf{M}_\mathbf{O} \mathbf{U} \Lambda \quad (\text{Dimension } N_{\text{dvd}})$$

- Rücktransformation

$$\mathbf{v}_i^{\text{neu}} = \mathbf{V}_{\text{alt}} \mathbf{u}_i \quad (i=1, \text{nstate})$$

$$\mathbf{H}_0 \mathbf{v}_i^{\text{neu}} = (\mathbf{H}_0 \mathbf{V}_{\text{alt}}) \mathbf{u}_i$$

$$\mathbf{B}^{\mu, \text{neu}} = \mathbf{B}^\mu \mathbf{U}$$

- Berechne Residuen $r_i = \mathbf{H} \mathbf{v}_i^{\text{neu}} - \lambda_i \mathbf{O} \mathbf{v}_i^{\text{neu}} = (\mathbf{H}_0 \mathbf{V}) \mathbf{U} + \sum_\mu \mathbf{P}^\mu \mathbf{C}^\mu$, wobei

$$\mathbf{C}^{\mu}_{\gamma, \nu} = \sum_{\gamma'} (\mathbf{D}^{\mu}_{\gamma, \gamma'} - \lambda_\nu \mathbf{q}^{\mu}_{\gamma, \gamma'}) (\mathbf{B}^\mu \mathbf{U})_{\gamma', \nu}$$

- Falls nicht letzter Schritt:

Orthogonalisiere $\mathbf{v}_i$ bzgl. Standardskalarprodukt:

Bilde Cholesky-Zerlegung $\mathbf{L}^+ \mathbf{L} = \mathbf{U}_{\text{nstate}}^+ \mathbf{U}_{\text{nstate}}$

$$\mathbf{v}_i^{\text{orth}} = \mathbf{V}_{\text{nstate}}^{\text{neu}} (\mathbf{L}^+)^{-1} \quad (\text{und entsprechend für } \mathbf{H}_0 \mathbf{v}_i^{\text{orth}}) \text{ und ebenso } \mathbf{B}^{\mu, \text{orth}} = \mathbf{B}^{\mu, \text{neu}} (\mathbf{L}^+)^{-1}$$

Es ergeben sich mehrere angenehme Nebeneffekte: Zum einen ist mit dieser Methode die Hermitizität (und bei sehr kleinen Eigenwerten vermutlich auch die positive Definitheit) der Überlappmatrix numerisch besser gewährleistet. (D.h. es gibt keinen numerischen Unterschied zwischen $\mathbf{v}_i(\mathbf{O} \mathbf{v}_j)$ und $\mathbf{v}_j(\mathbf{O} \mathbf{v}_i)$ mehr.)

Zum anderen kann man auf diese Weise auch sehr einfach die Produkte $\mathbf{B}^\mu$ zur Erzeugung der Ladungsdichte abspeichern. Dadurch ergibt sich eine zusätzliche Ersparnis. (Im Beispiel der Cu-Oberfläche aus Kapitel 6.3 macht dies immerhin ca. 45 Sekunden pro Selbstkonsistenzschritt aus.)

Schließlich fällt auch die Zeit zum Aufstellen der Überlappmatrix (im Beispiel ca. 30 Sekunden pro Davidsonschrift) fast völlig weg.

Aktion	Anzahl der Operationen
$\mathbf{H}_0 \mathbf{V}$	
$\mathbf{P}^\mu$	$N_\gamma \times N_{\vec{G}}$
$\mathbf{B}^\mu$	$N_\gamma \times N_{\vec{G}} \times N_{state}$
$\mathbf{M}_\mathbf{H}$	$N_{\vec{G}} \times N_{dvd}^2$
$\mathbf{M}_\mathbf{O}$	$N_\gamma \times N_{dvd}^2$
Diag.	N_{dvd}^3
$\mathbf{v}_i^{neu}$	0
$\mathbf{H}_0 \mathbf{v}_i^{neu}$	$N_{state} \times N_{\vec{G}} \times N_{dvd}$
$\mathbf{B}^{\mu, neu}$	$N_\gamma \times N_{state} \times N_{dvd}$
Resid.	$N_\gamma \times N_{\vec{G}} \times N_{state}$
$\mathbf{C}^\mu$	$\sum_\mu (N_\gamma^\mu)^2 \times N_{state}$
Orth. $\mathbf{V}$	$N_{state} \times N_{\vec{G}} \times N_{dvd}$
Orth. $\mathbf{H}_0 \mathbf{V}$	$\frac{1}{2} N_{\vec{G}} \times N_{state}^2$
Orth $\mathbf{B}^\mu$	$\frac{1}{2} N_\gamma \times N_{state}^2$

Tabelle C.1: Anzahl der Multiplikationen. (Für Kupferoberfläche aus Kapitel 6.3.3: $N_{\vec{G}} = 12663$, $N_\gamma = 1296$, $N_{dvd} = 860$, $N_{state} = 430$)

Anhang D

Partielle lokale Zustandsdichte

Die gewöhnliche Zustandsdichte (DOS) gibt an, wieviele Zustände pro Energieintervall vorliegen:

$$\text{DOS: } D(\varepsilon) = \int_{1.BZ} d^3k \sum_{\nu} \delta(\varepsilon - \varepsilon_{\vec{k}\nu}) \times 1 \quad (\text{D.1})$$

Die explizit angefügte Eins läßt sich deuten als

$$1 = \langle \psi_{\vec{k}\nu} | \psi_{\vec{k}\nu} \rangle_{\Omega}, \quad (\text{D.2})$$

d.h. als Skalarprodukt des Zustandes $\psi_{\vec{k}\nu}$ mit sich selbst, wobei das gewöhnliche Skalarprodukt benutzt wird, d.h. die Integration über die ganze Superzelle Ω .

In der *lokalen* Zustandsdichte (LDOS) werden Zustände gewichtet nach ihrer Aufenthaltswahrscheinlichkeit in einer Kugel S_{μ} von Radius R um ein Atom $\vec{r}_{\mu}$.

$$\text{LDOS: } D^{\mu}(\varepsilon) = \int_{1.BZ} d^3k \sum_{\nu} \delta(\varepsilon - \varepsilon_{\vec{k}\nu}) \times \langle \psi_{\vec{k}\nu} | \psi_{\vec{k}\nu} \rangle_{S_{\mu}} \quad (\text{D.3})$$

Das Skalarprodukt $\langle \psi_{\vec{k}\nu} | \psi_{\vec{k}\nu} \rangle_{S_{\mu}}$ bedeutet eine Integration nur im Kugelvolumen S_{μ} . Die Drehimpulsquantenzahlen $L=(l,m)$, welche in einem Atom Erhaltungsgrößen sind, sind für ein Atom in einem Festkörper keine guten Quantenzahlen. Jedoch kann man dennoch formal eine Entwicklung der Wellenfunktionen um ein bestimmtes Atom durchführen. Nahe an einem Atom erwartet man, daß der Einfluß dieses Atomes dominant ist und man daher näherungsweise von „s-, p- und d-Zuständen“ sprechen kann.

Der Operator P_L^{μ} möge die Projektion auf die Drehimpulskomponente L bezüglich des Atomes μ (am Ort $\vec{r}_{\mu}$) bezeichnen.

Die *partielle* lokale Zustandsdichte (PLDOS) ist dann die Zustandsdichte gewichtet nach den jeweiligen Drehimpulsanteilen an einem bestimmten Atom

$$\text{PLDOS: } D_L^\mu(\varepsilon) = \int_{1.BZ} d^3k \sum_\nu \delta(\varepsilon - \varepsilon_{\vec{k}\nu}) \times \langle \psi_{\vec{k}\nu} | P_L^\mu | \psi_{\vec{k}\nu} \rangle_{S_\mu} \quad (\text{D.4})$$

Wegen

$$1_{S_\mu} = \sum_L P_L^\mu \quad (\text{D.5})$$

ist die LDOS an einem Atom also gerade die Summe über alle dortigen PLDOS. Im folgenden werden stets die m-summierten partiellen Dichten betrachtet; die zugehörigen Größen werden mit dem Index l (statt L) bezeichnet.

Zur Abkürzung wird die Größe

$$q_{\vec{k}\nu}^{\mu l}(R) = \langle \psi_{\vec{k}\nu} | P_l^\mu | \psi_{\vec{k}\nu} \rangle_{S_\mu} = \int_{S_\mu} d^3r |\langle \vec{r}_\mu | P_l^\mu | \psi_{\vec{k}\nu} \rangle|^2 \quad (\text{D.6})$$

definiert, wobei benutzt wurde, daß $P_l^\mu = P_l^\mu \times P_l^\mu$, weil P_l^μ ein Projektionsoperator ist.

Durch Entwicklung Ebener Wellen nach Kugelflächenfunktionen (vgl. Glg. (A.36)) erhält man

$$\langle \vec{r}_\mu | \psi_{\vec{k}\nu} \rangle = \frac{e^{i\vec{k}\vec{r}_\mu}}{\sqrt{\Omega}} \sum_{\vec{G}} c_{\vec{k}\nu}^{\vec{G}} e^{-i\vec{G}\vec{r}_\mu} 4\pi \sum_L i^{-l} j_l(|\vec{k}+\vec{G}|r_\mu) Y_L(\vec{k}+\vec{G}) Y_L(\vec{r}_\mu), \quad (\text{D.7})$$

also

$$\langle \vec{r}_\mu | P_l^\mu | \psi_{\vec{k}\nu} \rangle = \frac{e^{i\vec{k}\vec{r}_\mu}}{\sqrt{\Omega}} \sum_{\vec{G}} c_{\vec{k}\nu}^{\vec{G}} e^{-i\vec{G}\vec{r}_\mu} 4\pi \sum_m i^{-l} j_l(|\vec{k}+\vec{G}|r_\mu) Y_L(\vec{k}+\vec{G}) Y_L(\vec{r}_\mu). \quad (\text{D.8})$$

Daraus ergibt sich :

$$q_{\vec{k}\nu}^{\mu l}(R) = \frac{4\pi}{\Omega} \int_0^R dr_\mu \sum_m \left| \sum_{\vec{G}} c_{\vec{k}\nu}^{\vec{G}} e^{-i\vec{G}\vec{r}_\mu} r_\mu j_l(|\vec{k}+\vec{G}|r_\mu) \sqrt{4\pi} Y_L(\vec{k}+\vec{G}) \right|^2 \quad (\text{D.9})$$

Die Symmetrie der Superzellen läßt sich auch hier ausnutzen. Analog zu Kapitel 3.15 findet man

$$q_{\mathbf{S}^{-1}\vec{k},\nu}^{\mu,l} = q_{\vec{k},\nu}^{\mathbf{S}\mu,l}. \quad (\text{D.10})$$

Damit ergibt sich für die k-Punkt Summation

$$D_L^\mu(\varepsilon) = \sum_{\vec{k}\nu} \omega_{\vec{k}} q_{\vec{k},\nu}^{\mu,l} \delta_{\varepsilon,\varepsilon_{\vec{k}\nu}} \quad (\text{D.11})$$

$$= \sum_{\vec{k} \in iBZ, \nu} \frac{\#S\vec{k}}{\#S} \sum_{S \in S} \omega_{S^{-1}\vec{k}} q_{S^{-1}\vec{k},\nu}^{\mu,l} \delta_{\varepsilon,\varepsilon_{\vec{k}\nu}} \quad (\text{D.12})$$

$$= \sum_{\vec{k} \in iBZ, \nu} \omega_{\vec{k}} \frac{\#S\vec{k}}{\#S} \sum_{\mu' \in S\mu} \frac{\#S}{\#S\mu} q_{\vec{k}\nu}^{\mu',l} \delta_{\varepsilon,\varepsilon_{\vec{k}\nu}} \quad , \quad (\text{D.13})$$

wobei iBZ den irreduziblen Teil der Brillouinzone, $\omega_{\vec{k}}$ die Gewichte der Summation, $\#S\vec{k}$ die Anzahl der zu $\vec{k}$ äquivalenten k-Punkte, $\#S\mu$ die Anzahl der zu μ äquivalenten Atome und $\#S$ die Anzahl der Symmetrioperationen in S bezeichnet. Da man nur für eine diskrete Menge von k-Punkten Zustände berechnet, werden die Zustandsdichten ebenfalls diskontinuierlich. Daher ist es nötig, eine Verbreiterung der Peaks einzuführen. Dies geschieht durch Faltung mit einer Gaußfunktion; die Wahl der Breite ist ein Kompromiß zwischen Glattheit einerseits und Auflösungsvermögen andererseits. Sinnvolle Werte für die Verbreiterung hängen stark von der Anzahl der benutzten k-Punkte ab; eine generelle Regel läßt sich daher nicht angeben. Auf jeden Fall ist es aber sinnvoll, möglichst viele verschiedene Punkte zu benutzen.

Die lokalen und die partiellen lokalen Zustandsdichten hängen ab von der Wahl des Kugelradius', über die integriert wird. Auch hier muß ein Kompromiß eingegangen werden; eine Obergrenze stellt das Minimum aus Wigner-Seitz-Radius und halbem Abstand zum nächsten Atom dar. Andererseits sollte der Radius nicht zu klein gewählt werden, da die Pseudopotentialmethode gerade in der Region nahe am Atom das echte Potential und die echte Wellenfunktion nicht wiedergibt. Qualitativ werden die wichtigsten Strukturen einer solchen lokalen Zustandsdichte in den meisten Fällen aber über einen relativ großen Bereich von Radien ähnlich berechnet. Ein Programm, das diese Zustandsdichten berechnet, wurde implementiert für die Pseudopotentialversion des Programmes **ESCoMPP**.

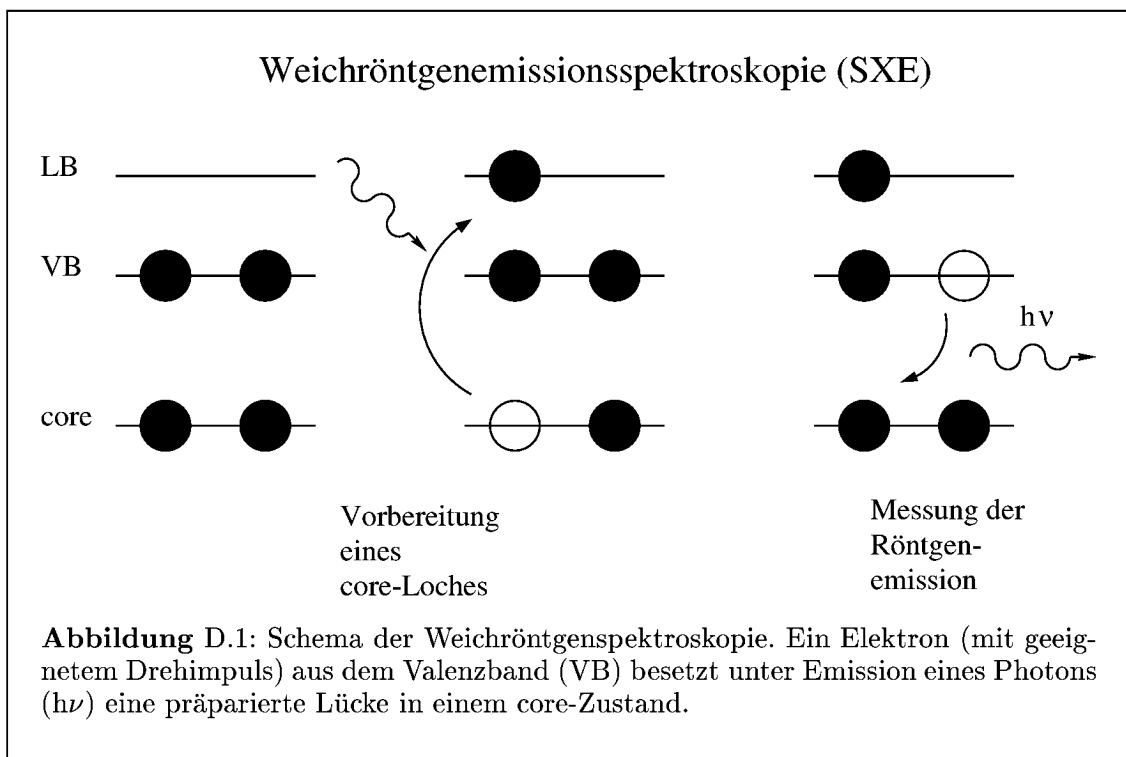
In der PAW-Methode erhält man lokale Zustandsdichten analog zur Drehimpulsentwicklung der *echten* Dichten $n_{\mu L}^1(r)$:

$$D_{L''}^\mu(\varepsilon) = \sum_{\vec{k}\nu} \omega_{\vec{k}} \delta_{\varepsilon,\varepsilon_{\vec{k}\nu}} \int_0^R dr_\mu \sum_{\gamma,\gamma'} \langle \tilde{\psi}_{\vec{k}\nu} | \tilde{p}_\gamma^\mu \rangle \phi_{l',i}^{\alpha(\mu)}(r_\mu) C(L, L', L'') \phi_{l,i}^{\alpha(\mu)}(r_\mu) \langle \tilde{p}_{\gamma'}^\mu | \tilde{\psi}_{\vec{k}\nu} \rangle \quad (\text{D.14})$$

D.1 Weichröntgenspektroskopie SXE

Die Soft X-Ray Emission (SXE) bietet eine Möglichkeit, partielle lokale Zustandsdichten direkt zu messen.¹

In einem vorbereitenden Schritt präpariert man in einem Atom (z.B. durch Einstrahlung geeigneter Photonen) ein core-Loch. Da die dazu benötigten Energien elementspezifisch sind, kann man dies selektiv für einzelne Atomsorten tun. In einem nachfolgenden Schritt kann ein Valenzelektron unter Emission eines Photons ($h\nu$) in diesen core-Zustand übergehen. Die Energie dieser Photonen liegt im Bereich der weichen Röntgenstrahlung. Da der core-Zustand lokalisiert ist, findet dieser Über-



gang an einem bestimmten Atom statt, obwohl wir einen periodischen Festkörper betrachten. Bezüglich dieses Atoms kann man auch den involvierten ausgedehnten Valenzzustand nach seinen Drehimpulskomponenten charakterisieren. Aufgrund der Dipolauswahlregel

$$\Delta l = \pm 1, \quad (\text{D.15})$$

¹Dieser Abschnitt lehnt sich an [Eis95, Ell96, Lün98] an. Eine generelle Behandlung der Theorie von Emission und Absorption findet man z.B. in [Sak67].

welche sich aus Fermis Goldener Regel ergibt (s.u.), können nicht beliebige Valenzelektronen in dieses Niveau übergehen, sondern nur solche, welche bezüglich dieses Atomes eine geeignete Symmetrie aufweisen.

Fermis Goldene Regel gibt die Übergangsrate zwischen Anfangszustand $|i\rangle$ und Endzustand $|f\rangle$ an:

$$\omega_{i \rightarrow f} \propto \left| \langle f | \hat{V} | i \rangle \right|^2 \delta(E_i - E_f - h\nu) \quad (\text{D.16})$$

Der Operator $\hat{V}$ beschreibt die Wechselwirkung von Photon und Elektron. Da die Wellenlänge der weichen Röntgenstrahlen sehr viel größer ist als die Ausdehnung des Atomes, kann man die Dipolnäherung verwenden. Die Winkelintegration des Matrixelementes ergibt die obige Auswahlregel, und es verbleibt ein radiales Matrixelement. Da im Festkörper nicht einzelne Zustände, sondern ein Kontinuum von Valenzzuständen vorliegt, führt die Integration über diese Zustände zum Auftreten der Zustandsdichte. Ein anfängliches s-Loch kann nur mit einem p-Elektron besetzt werden:

$$I_s(h\nu) \propto \nu^3 D_p^\alpha(E) M_{s,p}^2(E) \quad (\text{D.17})$$

und ein anfängliches p-Loch erfordert s- oder d-Elektronen:

$$I_p(h\nu) \propto \nu^3 \left(D_s^\alpha(E) M_{p,s}^2(E) + \frac{2}{5} D_d^\alpha(E) M_{p,d}^2(E) \right) \quad (\text{D.18})$$

Dabei bezeichnet $h\nu$ die Energie des emittierten Photons und E die Energie des Valenzelektrons; diese sind wegen der Energieerhaltung durch $E = E_{\text{core}} + h\nu$ miteinander verbunden. E_{core} ist die Energie des core-Niveaus, wo ein Loch präpariert worden war. D_l^α bezeichnet die partielle lokale Zustandsdichte der Atomsorte α , in welcher man selektiv core-Löcher präpariert hat. $I(h\nu)$ ist die Intensität, die man bei der Photonenergie $h\nu$ erhält. Die radialen Matrixelemente $M(E)$ des betreffenden Übergangs sind oft auf dem zu untersuchenden Energiebereich nahezu konstant, so daß man mit der SXE-Methode unmittelbar partielle lokale Zustandsdichten im Valenzband messen kann. Analog kann man mit *Absorptionsmessungen* (SXA) die PLDOS des *Leitungsbandes* messen.

Abschließend sei darauf hingewiesen, daß bei dieser Messung ein Zustand mit einem core-Loch involviert ist und daher nicht klar ist, ob man die PLDOS des Anfangszustandes (mit core-Loch) oder die des Endzustandes (ohne core-Loch) mißt. Inzwischen gilt jedoch die *final-state-rule* allgemein als anerkannt, obwohl sie sich nur auf Plausibilitätsüberlegungen und experimentelle Daten stützt. Diese besagt, daß

die Zustandsdichte des Endzustandes maßgeblich ist; d.h. bei der Emissionsmessung (SXE) wird die PLDOS des Zustandes *ohne* core-Loch gemessen.

Literaturverzeichnis

- [AM76] N.W. Ashcroft and N.D. Mermin. *Solid State Physics*. Saunders College Publishing, 1976.
- [Are97] M. Aretz. Punktdefekte in oberflächennahen Schichten von III-V-Halbleitern. Master's thesis, Universität zu Köln, 1997.
- [ASH⁺99] M. Asato, A. Settels, T. Hoshino, T. Asada, S. Blügel, R. Zeller, and P.H. Dederichs. Full-potential KKR calculations for metals and semiconductors. *Phys. Rev. B*, 60(8):5202, 1999.
- [BB] S. Blügel and G. Bihlmayer. www.flapw.de.
- [BBA⁺00] R. Berger, S. Blügel, A. Antons, W. Kromen, and K. Schroeder. A parallelized ab-initio Molecular Dynamics Code for the Investigation of Atomistic Growth Processes. In *Proceedings of the NIC-Workshop „Molecular Dynamics on Parallel Computers“, Jülich, 8.-10. Feb. 1999*. World Scientific, 2000. ISBN 981-02-4232-8.
- [BDD⁺99] Z. Bai, J. Demmel, J. Dongarra, A. Ruhe, and H. v.d. Vorst, editors. *Templates for the Solution of Algebraic Eigenvalue Problems: a Practical Guide*. SIAM, Philadelphia, 1999.
- [Ber] R. Berger. Phd thesis. to be published.
- [BFV97] F. Bernardini, V. Fiorentini, and D. Vanderbilt. Offsets and polarization at strained AlN/GaN polar interfaces. *MRS Proceedings*, 449:923, 1997.
- [BHS82] G.B. Bachelet, D.R. Hamann, and M. Schlüter. Pseudopotentials that work: From H to Pu. *Phys. Rev. B*, 26(8):4199, 1982.
- [BK90] D.M. Bylander and L. Kleinman. 4f resonances with norm-conserving pseudopotentials. *Phys. Rev. B*, 41(2):907, 1990.

- [BKB] G. Bihlmayer, P. Kurz, and S. Blügel. An efficient FLAPW program: FLEUR. to be published.
- [Blo28] F. Bloch. Über die Quantenmechanik der Elektronen in Kristallgittern. *Z. Physik*, 52:555, 1928.
- [Blö90] P.E. Blöchl. Generalized separable potentials for electronic-structure calculations. *Phys. Rev. B*, 41(8):5414, 1990.
- [Blö94] P.E. Blöchl. Projector augmented-wave method. *Phys. Rev. B*, 50:17953, 1994.
- [Blü] S. Blügel. private communication.
- [Blü88] S. Blügel. *First Principles Calculations of the Electronic Structure of Magnetic Overlayers on Transition Metal Surfaces*. PhD thesis, RWTH Aachen, 1988.
- [BS79] I.N. Bronstein and K.A. Semendjajew. *Taschenbuch der Mathematik*, volume Ergänzungsband. Verlag Harry Deutsch, 19. Aufl. edition, 1979.
- [BS88] A. Banerjea and J. Smith. Universal Binding Energie Relation. *Phys. Rev. B*, 37:6632, 1988.
- [CA80] D.M. Ceperley and B.J. Alder. Ground State of the Electron Gas by a Stochastic Method. *Phys. Rev. Lett.*, 45:566, 1980.
- [CC88] J.R. Chelikowsky and M.Y. Chou. *Phys. Rev. B*, 38:7966, 1988.
- [CG94] N.E. Christensen and I. Gorczyca. Optical and structural properties of III-V nitrides under pressure. *Phys. Rev. B*, 50:4397, 1994.
- [CP85] R. Car and M. Parrinello. Unified Approach for Molecular Dynamics and Density-Functional Theorie. *Phys. Rev. Lett.*, 55:2471, 1985.
- [CPS94] M. Crouzeix, B. Philippe, and M. Sadkane. The davidson method. *SIAM J. Sci. Comp*, 15:62, 1994.
- [Dav75] W.E. Davidson. The Iterative Calculation of a Few of the Lowest Eigenvalues and Corresponding Eigenvectors of Large Real-Symmetric Matrices. *J. Comput. Phys.*, 17:87, 1975.

- [DB93] R.M. Dickson and A. Becke. *J. Comput. Phys.*, 99:3898, 1993.
- [DG90] R.M. Dreizler and E.K.U. Gross. *Density Functional Theory: An Approach to the Quantum Many-Body Problem*. Springer-Verlag, Berlin, 1990.
- [Eis95] S. Eisebitt. *Electronic and Structural Properties of Interfaces and Semiconductor Nanostructures*. PhD thesis, Universität zu Köln, 1995.
- [Ell96] C. Ellmers. Elektronische Struktur von SiC-Polytypen untersucht mit Weichröntgenspektroskopie. Master's thesis, Universität zu Köln, 1996. Jül-3204.
- [Eng92] B. Engels. Konstruktion von glatten normerhaltenden Pseudopotentialen. Master's thesis, RWTH Aachen, 1992.
- [Eng95] B. Engels. *Ab initio Berechnung der (110) Oberfläche von III-V Halbleitern: Simulation von Rastertunnelmikroskopie- Aufnahmen*. PhD thesis, RWTH Aachen, 1995.
- [ERBS] B. Engels, P. Richard, S. Blügel, and K. Schroeder. Test of different pseudo potential generation schemes with respect to convergence and transferability for $2p$ -elements. unpublished.
- [FKBK00] J. Furthmüller, P. Käckell, F. Bechstedt, and G. Kresse. Extreme softening of Vanderbilt pseudopotentials: General rules and case studies of first row and d -electron elements. *Phys. Rev. B*, 61:4576, 2000.
- [FMS93] V. Fiorentini, M. Methfessel, and M. Scheffler. Electronic and structural properties of GaN by the full-potential LMTO method: The role of the d -electrons. *Phys. Rev. B*, 47:13353, 1993.
- [Fri98] D. Frielinghaus. Ab initio Berechnung von BeSe und BeTe (110) Oberflächen. Master's thesis, RWTH Aachen, 1998.
- [GFB98] U. Grossner, J. Furthmüller, and F. Bechstedt. Bond rotation versus bond-contraction relaxation of (110) surfaces of group-III nitrides. *Phys. Rev. B*, 58:R1722, 1998.
- [GKS90] X. Gonze, P. Käckell, and M. Scheffler. Ghost states for separable, norm-conserving, *ab-initio* pseudopotentials. *Phys. Rev. B*, 41(17):12264, 1990.

- [GSS91] X. Gonze, R. Stumpf, and M. Scheffler. Analysis of separable potentials. *Phys. Rev. B*, 44(16):8503, 1991.
- [GvdV00] G.H. Golub and H. van der Vorst. Eigenvalue computation in the 20th century. to appear, 2000.
- [Ham89] D.R. Hamann. Generalized norm-conserving pseudopotentials. *Phys. Rev. B*, 40(5):2980, 1989.
- [HK64] P. Hohenberg and W. Kohn. Inhomogeneous Electron Gas. *Phys. Rev.*, 136(3B):B864, 1964.
- [HMD⁺97] N.A.W. Holzwarth, G.E. Matthews, R.B. Dunning, A.R. Tackett, and Y. Zeng. Comparison of the projector augmented-wave, pseudopotential, and linearized augmented-plane-wave formalisms for the density-functional calculations of solids. *Phys. Rev. B*, 55:2005, 1997.
- [HMD98] N.A.W. Holzwarth, G.E. Matthews, and R.B. Dunning. Orthogonal polynomial projectors for the projector augmented-wave method of electronic structure calculations. *Phys. Rev. B*, 57:11827, 1998.
- [HSC79] D.R. Hamann, M. Schlüter, and C. Chiang. Norm-conserving Pseudopotentials. *Phys. Rev. Lett.*, 43(20):1494, 1979.
- [IZC79] J. Ihm, A. Zunger, and M.L. Cohen. Momentum-space formalism for the total energy of solids. *J. Phys. C.: Solid State Phys.*, 12:4409, 1979.
- [Jac75] J.D. Jackson. *Classical Electrodynamics*. Wiley, New York, 1975.
- [Jeo96] S. Jeong. Structural Properties of bulk copper: Pseudopotential plane-wave-basis study. *Phys. Rev. B*, 53:13973, 1996.
- [JG89] R.O. Jones and O. Gunnarsson. The density functional formalism, its applications and prospects. *Rev. Mod. Phys.*, 61:689, 1989.
- [KB82] L. Kleinman and D.M. Bylander. Efficacious Forms for Model Pseudopotentials. *Phys. Rev. Lett.*, 48(20):1425, 1982.
- [Ker80] G.P. Kerker. Non-singular atomic pseudopotentials for solid state applications. *J. Phys. C.: Solid State Phys.*, 13:L189–94, 1980.

- [KF96] G. Kresse and J. Furthmüller. Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Comp. Mat. Sci.*, 6:15, 1996.
- [KGV83] S. Kirkpatrick, C.D. Gelatt, and M.P. Vecchi. Optimization by simulated annealing. *Science*, 220:671–680, 1983.
- [KH94] G. Kresse and J. Hafner. Norm-conserving and ultrasoft pseudopotentials for first row and transition elements. *J. Phys.: Condens. Matter*, 6:8245, 1994.
- [KJ99] G. Kresse and D. Joubert. From ultrasoft pseudopotentials to the projector augmented-wave method. *Phys. Rev. B*, 59:1758, 1999.
- [KLS96] K. Kim, W.R.L. Lambrecht, and B. Segall. Elastic constants and related properties of tetrahedrally bonded BN, AlN, GaN, InN. *Phys. Rev. B*, 53:16310, 1996.
- [Koh85] W. Kohn. In M.P. Tosi, F. Bassani, F. Fumi, editor, *Highlights of Condensed Matter Theorie*. North Holland, Amsterdam, 1985.
- [Kos84] N. Kosugi. Modification of the Liu–Davidson Method for Obtaining One or Simultaneously Several Eigensolutions of a Large Real–Symmetric Matrix. *J. Comput. Phys.*, 55:426, 1984.
- [KPW⁺94] K. Karch, P. Pavone, W. Windl, O. Schütt, and D. Strauch. *Phys. Rev. B*, 50:17054, 1994.
- [KS65] W. Kohn and L.J. Sham. Self-Consistent Equations Including Exchange and Correlation Effects. *Phys. Rev.*, 140(4A):A1133, 1965.
- [KSPL91] R.D. King-Smith, M.C. Payne, and J.S. Lin. Real-space implementation of nonlocal pseudopotentials for first-principles total-energy calculations. *Phys. Rev. B*, 44(23):13063, 1991.
- [KV82] W. Kohn and P. Vashishta. *Physics of Solids and Liquids*. Plenum, New York, 1982.
- [KWB94] P. Käckell, B. Wenzien, and F. Bechstedt. *Phys. Rev. B*, 50:10761, 1994.
- [LB63] Landolt-Börnstein. *Numerical Data and Functional Relationships in Science and Technology (New Series)*. Springer–Verlag, Berlin, 1963.

- [Lev79] M. Levy. Universal variational functionals of electron densities, first-order density matrices, and natural spin-orbitals and solutions of the v -representability problem. *Proc. Natl. Acad. Sci. USA*, 76:6062, 1979.
- [LFC82] S.G. Louie, S. Froyen, and M.L. Cohen. Nonlinear ionic pseudopotentials in spin-density-functional calculations. *Phys. Rev. B*, 26(4):1738, 1982.
- [Lie83] E.H. Lieb. Density Functionals for Coulomb Systems. *Int. J. Quantum Chem.*, 24:243, 1983.
- [Liu78] B. Liu. In *Report on Workshop: Numerical Algorithms in Chemistry: Algebraic Methods*, page 49. Lawrence Berkley Lab., Univ. of California, 1978.
- [LPC⁺93] K. Laasonen, A. Pasquarello, R. Car, C. Lee, and D. Vanderbilt. Car-Parrinello molecular dynamics with Vanderbilt ultrasoft pseudopotentials. *Phys. Rev. B*, 47:10142, 1993.
- [Lün98] J. Lüning. *Soft X-Ray Spectroscopy of Broad Band, Size Confined and Correlated Materials*. PhD thesis, Universität zu Köln, 1998.
- [Mes61] Messiah. *Quantum Mechanics*. North Holland, Amsterdam, 1961.
- [MJW78] V.L. Moruzzi, J.F. Janak, and A.R. Williams. *Calculated Electronic Properties of Metals*. Pergamon, New York, 1978.
- [Mor] Y. Morikawa. private communication.
- [Mor90] R.B. Morgan. Davidson's Method and Preconditioning for Generalized Eigenvalue Problems. *J. Comput. Phys.*, 89:241, 1990.
- [MS86] R. B. Morgan and D. S. Scott. Generalizations of Davidson's method for computing eigenvalues of sparse symmetric matrices. *SIAM J. Sci. Statist. Comput.*, 7:817–825, 1986.
- [MSV96] J.A. Majewski, M. Städele, and P. Vogl. Electronic structure of biaxially strained wurtzite crystals GaN, AlN, and InN. *MRS Intern. J. Nitride Semicond. Res.*, 1:30, 1996.
- [Mur44] F.D. Murnaghan. *Proc. Nat. Acad. Sci. U.S.A.*, 30:244, 1944.

- [MV79] A.H. MacDonald and S.H. Vosko. A relativistic density functional formalism. *J. Phys. C.*, 12:2977, 1979.
- [OEK99] K. Ohno, K. Esfarjani, and Y. Kawazoe. *Computational Material Science*. Springer Series in Solid State Sciences. Springer, 1999.
- [Par80] B.N. Parlett. *The symmetric eigenvalue problem*. Prentice Hall series in computational mathematics. Englewood-Cliffs, NJ : Prentice Hall, 1980.
- [PCLC94] C.H. Park, B.H. Cheong, K.H. Lee, and K.J. Chiang. *Phys. Rev. B*, 49:4485, 1994.
- [PCV⁺92] John P. Perdew, J. A. Chevary, S. H. Vosko, K. A. Jackson, M. R. Peder-son, D. J. Singh, and C. Fiolhais. Atoms, molecules, solids, and surfaces: Applications of the generalized gradient approximation for exchange and correlation. *Phys. Rev. B*, 46:6671, 1992.
- [PM77] J.D. Pack and H.J. Monkhorst. “Special points for Brillouine–zone in-TEGRATIONS” – a reply. *Phys. Rev. B*, 16:1748, 1977.
- [PTVF92] W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery. *Nu-merical Recipes in Fortran*. Cambridge University Press, 1992.
- [PWH⁺00] M. Petersen, F. Wagner, L. Hufnagel, M. Scheffler, P. Blaha, and K. Schwarz. Improving the Efficiency of FP-LAPW calculations. *Com-put. Phys. Commun.*, 126:294–309, 2000.
- [Ric93] P. Richard. Iterative Lösung der Schrödinger Gleichung für große Systeme. Master’s thesis, RWTH Aachen, 1993.
- [Ric96] P. Richard. *Banddiskontinuitäten an III-V Halbleitergrenzflächen: Ver-spannungsabhängigkeit und Einfluß von Zwischenschichten*. PhD thesis, RWTH Aachen, 1996.
- [RRKJ90] A.M. Rappe, K.M. Rabe, E. Kaxiras, and J.D. Joannopoulos. Optimized pseudopotentials. *Phys. Rev. B*, 41(2):1227, 1990.
- [Saa92] Y. Saad. *Numerical Methods for large Eigenvalue Problems*. Manchester University Press, 1992.

- [Sab96] M. Sabisch. *Ab-initio Studien elektronischer Eigenschaften von SiC-Kristallen und deren Oberflächen*. PhD thesis, Universität Münster, 1996.
- [Sak67] J.J. Sakurai. *Advanced Quantum Mechanics*. Addison-Wesley, London, 1967.
- [SBM⁺95] J. Schardt, C. Bram, S. Müller, U. Starke, K. Heinz, and K. Müller. *Surf. Sci.*, 337:232, 1995.
- [Sin94] D. Singh. *Planewaves, Pseudopotentials and the LAPW Method*. Kluwer Academic Publishers, Boston/Dordrecht/London, 1994.
- [SMV97] M. Städele, J.A. Majewski, and P. Vogl. Stability and band offsets of polar GaN/SiC(001) and AlN/SiC(001) interfaces. *Phys. Rev. B*, 56:6911, 1997.
- [SSO92] M. Saito, O. Sugino, and A. Oshiyama. Real-space-partitioned separable pseudopotential. *Phys. Rev. B*, 46(4):2606, 1992.
- [SvdV96] G.L.G. Sleijpen and H.A. van der Vorst. A Jacoby-Davidson iteration method for linear eigenvalue problems. *SIAM J. Matrix Anal. Appl.* 17, pages 401–425, 1996.
- [SW90] J.M. Soler and A.R. Williams. Augmented-plane-wave Forces. *Phys. Rev. B*, 42:9728, 1990.
- [SW93] J.M. Soler and A.R. Williams. Comment on „All-electron and pseudopotential force calculations using the linearized-augmented-plane-wave method“. *Phys. Rev. B*, 47:6784, 1993.
- [TM90] N. Troullier and J.L. Martins. A straightforward method for generating soft transferable pseudopotentials. *Solid State Commun.*, 74(7):613, 1990.
- [TM91] N. Troullier and J.L. Martins. Efficient pseudopotentials for plane-wave calculations. *Phys. Rev. B*, 43(3):1993, 1991.
- [TPA89] M.P. Teter, M.C. Payne, and D.C. Allan. Solution of Schrödingers equation for large systems. *Phys. Rev. B*, 40:12255, 1989.

- [Van85] D. Vanderbilt. Optimally smooth norm-conserving pseudopotentials. *Phys. Rev. B*, 32(12):8412, 1985.
- [Van90] D. Vanderbilt. Soft self-consistent pseudopotentials in a generalized eigenvalue formalism. *Phys. Rev. B*, 41(11):7892, 1990.
- [vBH72] U. von Barth and L. Hedin. A local exchange-correlation potential for the spin polarized case: I. *J. Phys. C.*, 5:1629, 1972.
- [VS92] J. Vackar and A. Simunek. *Solid State Commun.*, 81:837, 1992.
- [VWN80] S.H. Vosko, L. Wilk, and N. Nussair. Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis. *Can. J. Phys.*, 58:1200, 1980.
- [Wei81] M. Weinert. *J. Math. Phys.*, 22:2433, 1981.
- [WN94] A.F. Wright and J.S. Nelson. Explicit treatment of the gallium 3d-electrons in GaN using the plane-wave pseudopotential method. *Phys. Rev. B*, 50:2159, 1994.
- [WN95] A.F. Wright and J.S. Nelson. Consistent structural properties for AlN, GaN and InN. *Phys. Rev. B*, 51:7866, 1995.
- [WZ85] J.H. Wood and A. Zunger. *J. Phys. A.*, 18:1343, 1985.
- [YC82] M.T. Yin and M.L. Cohen. Theory of *ab-initio* pseudopotential calculations. *Phys. Rev. B*, 25(12):7403, 1982.
- [YC83] M.T. Yin and M.L. Cohen. Will diamond transform under megabar pressure. *Phys. Rev. Lett.*, 50:2006, 1983.
- [YC84] M.T. Yin and M.L. Cohen. Structural theory of graphite and graphitic silicon. *Phys. Rev. B*, 29:6996, 1984.
- [YLFZ92] C. Yeh, Z. Lu, S. Froyen, and A. Zunger. Predictions and systematizations of the zinc-blende-wurtzite structural energies. *Phys. Rev. B*, 45:12130, 1992.
- [YSK91] R. Yu, D. Singh, and H. Krakauer. Forces in flapw. *Phys. Rev. B*, 43:6411, 1991.

- [ZMV99] G. Zandler, J.A. Majewski, and P. Vogl. Pyroelectronics: Novel device concepts based on nitride interfaces. *J. Vac. Sci. Technol.*, 1999.

Danksagung

An dieser Stelle möchte ich allen danken, die das Gelingen dieser Arbeit ermöglicht haben:

Herrn Prof. Dr. Schroeder danke ich für die große Freiheit, die ich in der Ausrichtung meiner Arbeit hatte, sowie für die Zeit, die er sich stets für mich genommen hat, um meine Arbeit zu unterstützen.

Herrn Dr. Stefan Blügel hat mich nicht nur auf die PAW-Methode aufmerksam gemacht hat, sondern auch meine Arbeit fortwährend unterstützt. In den zahllosen Diskussionen mit ihm habe ich sehr viel gelernt und nicht zuletzt auch von seiner Begeisterung für die Physik gezehrt. Auch für die freundschaftliche Atmosphäre, die er verbreitet hat, möchte ich ihm danken.

Herrn Prof. Dr. P.H. Dederichs danke ich für die freundliche Übernahme des Korreferats, sowie für das Interesse, das er meiner Arbeit entgegengebracht hat.

Herrn Prof. Dr. H. Müller-Krumbhaar danke ich für die Möglichkeit, an seinem Institut diese Doktorarbeit erstellen zu dürfen, und für seine langfristige Unterstützung der gesamten Arbeitsgruppe.

Im Verlaufe dieser Arbeit habe ich die Unterstützung vieler Mitarbeiter des Institutes erfahren, denen allen mein Dank gebührt. Dennoch möchte ich einige besonders hervorheben.

Diedrich „Urmeli“ Frielinghaus war mir im ersten Jahr dieser Arbeit nicht nur ein angenehmer Zimmergenosse, der zu einer entspannten Arbeitsatmosphäre beigetragen hat, sondern er hat mir auch bereitwillig seinen Wissensvorsprung zuteil werden lassen. Durch ihn habe ich die Benutzung der nötigen Programme sehr viel schneller erlernt, als es ohne ihn der Fall gewesen wäre.

Ralf Berger danke ich für viele erhellende Gespräche; neben physikalischen Ratschlägen war er für mich auch *die* Instanz im Hinblick auf das benutzte Pseudopotentialprogramm.

Alexander Settels und Philipp Kurz danke ich für nützliche Informationen über Teile

der von ihnen benutzten KKR- bzw. FLAPW-Programme.

Gustav Bihlmayer danke ich für die FLAPW-Rechnungen, die er zu Vergleichszwecken durchgeführt hat, seine allzeitige Bereitschaft, auf meine Fragen einzugehen, und insbesondere für die Beobachtung, daß die nicht-relativistische Berechnung von Kupfer die Ursache der aufgetretenen Unstimmigkeiten war.

Similarly, I want to express my thankfulnes to Dr. Y. Morikawa from the Joint Research Center for Atom Technology (JRCAT) in Tsukuba, Japan, for performing calculations with his Ultrasoft-Pseudopotential-Program and for discussing with me the differences in the respective implementations.

Jan Lüning danke ich für die angenehme Zusammenarbeit im Umfeld der Siliziumkarbide.

Für die nette Arbeitsatmosphäre gilt mein Dank außerdem Bereshad Nonas, Roland Rzehak, Diego Kienle, Stefan Bröcheler und vielen anderen.

Für sorgfältiges Korrekturlesen bin ich besonders Dr. Peter Wurth und Angela Westdorf sehr dankbar. Auch Andreas Sindermann, Frank Gutheim und Miks Hartmann gebührt Dank dafür.

Diego Dienle bin ich auch für die Unterstützung in den letzten Tagen dieser Arbeit dankbar.

Zuletzt und besonders herzlich danke ich denen, die es mir ermöglicht haben, diese Arbeit zu beginnen, die mir jederzeit bedingungslosen Rückhalt geboten haben, meine Launen ertragen und mir jede erdenkliche Unterstützung geboten haben:

Meinen Eltern Ursula und Eugen Kromen
meinen Geschwistern Bettina und Wolfgang
und
meiner Freundin Angela

