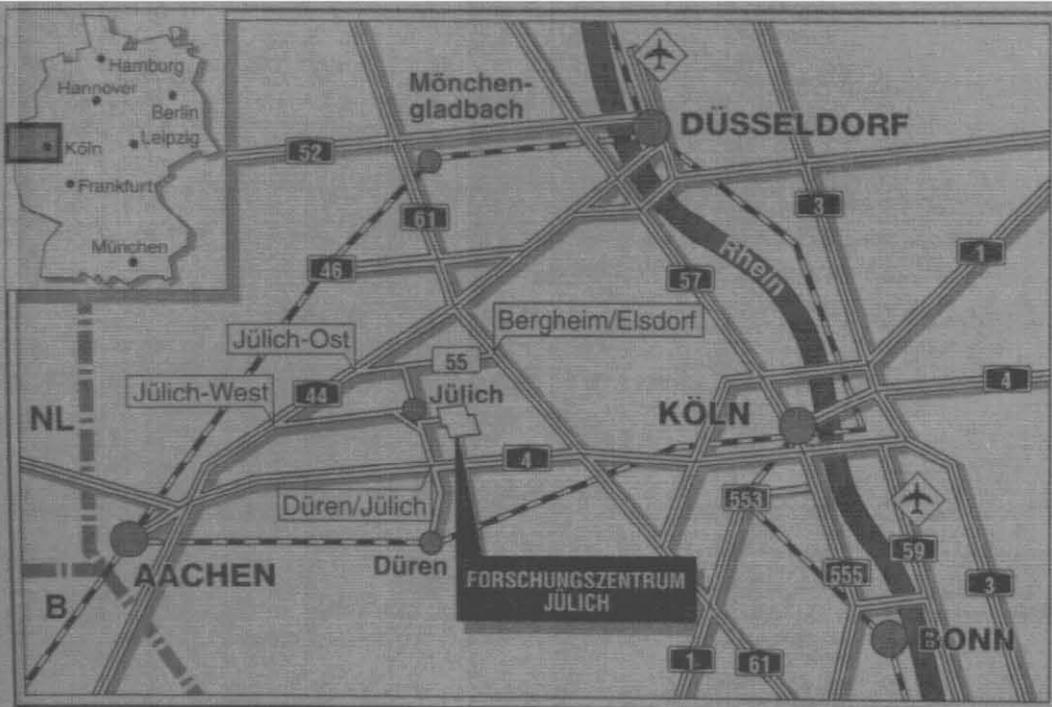


Institut für Festkörperforschung

***Ab initio* Berechnung der (110) Oberfläche von III-V Halbleitern:
Simulation von Rastertunnelmikroskopie-Aufnahmen**

Benedikt Engels



Berichte des Forschungszentrums Jülich ; 3215

ISSN 0944-2952

Institut für Festkörperforschung Jül-3215

D82 (Diss. RWTH Aachen)

Zu beziehen durch : Forschungszentrum Jülich GmbH · Zentralbibliothek
D-52425 Jülich · Bundesrepublik Deutschland

Telefon: 02461/61-6102 · Telefax: 02461/61-6103 · Telex: 833556-70 kfa d

***Ab initio* Berechnung der (110) Oberfläche von III-V Halbleitern:
Simulation von Rastertunnelmikroskopie-Aufnahmen**

Benedikt Engels

Kurzfassung

Verbindungshalbleiter sind vor allem wegen ihrer Verwendbarkeit in optoelektronischen Bauelementen von besonderem technologischen Interesse. Voraussetzung für Herstellung und zielgerichtetes Design elektronischer Bauelemente ist ein genaues Verständnis der atomaren und elektronischen Struktur von Halbleiter-Schichtsystemen und Oberflächen. Eines der wichtigsten experimentellen Werkzeuge für die Untersuchung von Oberflächen ist die Rastertunnelmikroskopie (RTM). Obwohl die RTM auf einzigartige Weise eine Abbildung der elektronischen Oberflächenstruktur mit atomarer Auflösung ermöglicht, so kann sie doch die atomare Oberflächengeometrie nicht immer direkt und zweifelsfrei klären. Die hierzu fehlende Verbindung zwischen elektronischer Struktur (lokale Zustandsdichte) und atomarer Geometrie (Atompositionen) kann durch “*ab initio*”-Rechnungen hergestellt werden.

In der hier vorgelegten Doktorarbeit wird eine solche “*ab initio*” Berechnung der (110) Oberfläche von III-V Halbleitern durchgeführt und der berechnete Tunnelstrom mit experimentellen RTM-Bildern verglichen:

- Die atomare und elektronische Struktur der (110) Oberfläche von III-V Halbleitern (Superzelle, Slab-Geometrie) wird berechnet. Für die Relaxation der Atome in die atomare Gleichgewichtsstruktur werden die Hellmann-Feynman-Kräfte auf die Ionen in Kombination mit einem niedrig-dimensionalen Quasi-Newton-Verfahren (Broyden-Fletcher-Goldfarb-Shano-Jakobian) verwendet. Das gängige Strukturmodell (“bond rotation model”) wird für mehrere Systeme bestätigt. Die theoretisch bestimmten Verschiebungen, atomare Abstände und Bindungswinkel, stimmen mit den Experimenten überein ($\Delta d \sim 0.05 \text{ \AA}$, $\Delta \Theta \sim 5^\circ$). Dies entspricht der experimentellen Genauigkeit und liegt innerhalb der Schwankungen verschiedener anderer Rechnungen.
- Eine detaillierte Analyse der Oberflächenzustände und Oberflächenresonanzen wird durchgeführt. Die lokale Zustandsdichte in einiger Entfernung zur Oberfläche wird im Rahmen der Tersoff-Hamann-Näherung für die “*ab in-*

itio” Simulation von RTM–Aufnahmen verwendet. Die so berechneten RTM–Bilder werden mit hochauflösenden Aufnahmen aus RTM–Experimenten verglichen, die von Dr. Philipp Ebert am gleichen Institut durchgeführt werden. Der Vergleich ergibt eine unerwartet gute Übereinstimmung zwischen Theorie und Experiment einschließlich feiner Details. Dies erlaubt, die Spannungsabhängigkeit der RTM–Aufnahmen auf die elektronische Oberflächenstruktur zurückzuführen, i.e. die Ursache morphologischer Details zu klären. In diesem Zusammenhang konnten wir z.B. zeigen, daß die spannungsabhängige Reihenumkehr, die in RTM–Aufnahmen der unbesetzten elektronischen Zustände beobachtet wird, sowie die spannungsabhängige Reihenverschiebung, die in Aufnahmen der besetzten Zustände beobachtet wird, auf Oberflächenresonanzen zurückzuführen ist. Dies kann in dem bis dato benutzten konventionellen Modell reiner lokalisierter Oberflächenzustände nicht verstanden werden.

Die Rechnungen wurden mit einem Car–Parrinello–ähnlichen Computerprogramm durchgeführt, welches in der Arbeitsgruppe, Dr. Stefan Blügel, Dipl. Phys. Benedikt Engels¹, Dipl. Phys. Peter Richard und Prof. Dr. Kurt Schroeder entwickelt wurde. Die elektronische Struktur wird im Rahmen der Dichtefunktionaltheorie (DFT) in der Lokalen–Dichte–Näherung (LDA) in Kombination mit der Pseudopotentialmethode (nichtlokale, normerhaltende Pseudopotentiale) berechnet. Das elektronische Eigenwertproblem wird iterativ im Impulsraum (Superzellen, ebene Wellen Basis) gelöst. Verwendet wird dazu der Davidson–Kosugi–Algorithmus. Die elektronische Selbstkonsistenz wird sehr effizient mittels eines hochdimensionalen Quasi–Newton–Verfahrens (Broyden–Mixing, verallgemeinertes Anderson–Mixing) erreicht.

Die Kombination von Theorie und Experiment liefert wertvolle Impulse im Hinblick auf die Zielstellung, aus RTM–Experimenten möglichst vollständige Information über die elektronische und atomare Struktur von Oberflächen zu gewinnen. So ist es zur Zeit experimentell z.B. nicht möglich, verschiedene Atomsorten auf der Oberfläche in RTM–Aufnahmen chemisch zu unterscheiden. Die hier vorgestellte Methode kann dazu beitragen, diese Lücke zu schließen und mit steigender Rechnerleistung, z.B. durch den Einsatz von Parallelcomputern, in nächster Zukunft auch auf komplexere Systeme wie z.B. kompliziertere Oberflächenrekonstruktionen, Stufen oder Defekte angewendet werden.

¹Autor der hier vorgelegten Arbeit

Inhaltsverzeichnis

Kurzfassung	iii
1 Die Dichtefunktionaltheorie	1
1.1 Die Energie als Dichtefunktional	2
1.2 Die Kohn–Sham–Gleichungen	4
1.3 Die Spindichtefunktionaltheorie	7
1.4 Die Lokale–Dichte–Näherung	8
1.5 Die Gesamtenergie des elektronischen Systems	9
2 Die Pseudopotentialmethode	13
2.1 Idee der Pseudopotential–Methode	14
2.2 Semilokale, normerhaltende Pseudopotentiale	17
2.3 Separable, “Kleinman–Bylander”–Pseudopotentiale	20
2.4 Partielle Core–Korrektur	22
3 Explizite Darstellung der Gleichungen	25
3.1 Der Hamiltonian in der Impulsdarstellung	26
3.2 Superzellen und ebene Wellen	27
3.3 Gittersummen und Konvolutionsgleichungen	29
3.3.1 Gittersummen	29
3.3.2 Konvolutionsgleichungen	30
3.4 Wellenfunktionen und Ladungsdichte	31
3.5 Die kinetische Energie	33
3.6 Der Hartree–Term	33
3.7 Der lokale Anteil des Hamiltonian	35
3.7.1 Berechnung der $V_{lok}^{\alpha}(\vec{G})$	37
3.8 Der nichtlokale Anteil des Hamiltonian	39
3.8.1 Die separablen KB–Potentiale	40

3.8.2	Die semilokalen BHS-Potentiale	42
3.9	Austausch und Korrelation	44
3.10	Die Ewaldsumme	46
3.11	Die divergenten Terme	48
3.12	Energie und Kraft im Computerprogramm	48
4	Lösung der Kohn-Sham-Gleichungen	51
4.1	Traditionelle Bandstrukturmethode	52
4.2	Iterative Diagonalisierung	53
4.3	Der Davidson-Kosugi-Algorithmus	55
5	Minimierung des Energiefunktional	61
5.1	“ <i>simple mixing</i> ”	63
5.2	Die Newton-Raphson-Methode	64
5.3	Quasi-Newton-Methoden	64
5.3.1	Broydens erste Formel	66
5.3.2	Broydens zweite Formel	66
5.3.3	Der verallgemeinerte Anderson-Jacobian	67
5.3.4	Der DFP-Jacobian	67
5.3.5	Der BFGS-Jacobian	67
5.4	Die Relaxation der elektronischen Struktur	68
5.5	Relaxation der atomaren Struktur	70
5.5.1	Der erste Quasi-Newton-Schritt	70
5.5.2	Negative Definitheit des Jacobian	71
5.5.3	Konvergenzanforderung an die Kraft	71
6	(110) Oberflächen von III-V Halbleitern	73
6.1	Motivation	73
6.2	Das Strukturmodell	74
6.3	Vorgehensweise bei der Berechnung	77
6.3.1	Die Pseudopotentiale	77
6.3.2	Die Oberflächen-Superzelle	78
6.4	Berechnung der Bulkstruktur	82
6.4.1	Bestimmung der theoretischen Gitterkonstanten	82
6.4.2	Die elektronische Struktur des Bulks	84
6.5	Die atomare Relaxation der Oberfläche	88
6.5.1	Globale Konvergenz	88
6.5.2	Notation	89

6.5.3	Konvergenz mit der Systemgröße	91
6.5.4	Konvergenz mit der Anzahl der relaxierten Lagen	91
6.5.5	Konvergenz mit der Anzahl der ebenen Wellen	92
6.5.6	Vergleich mit der Literatur	92
6.6	Die elektronische Struktur der Oberfläche	100
6.6.1	Quantentrogzustände	101
6.6.2	Oberflächenzustände und Oberflächenresonanzen	110
6.6.2.1	Methodik	111
6.6.2.2	Zustandsdichten und Oberflächenzustandsdichten	114
6.6.2.3	Die besetzten Zustände	121
6.6.2.4	Die unbesetzten Zustände	132
7	Vergleich mit dem Tunnelexperiment	141
7.1	Theorie des Tunnelstroms	141
7.2	Experimentelle Ergebnisse	145
7.3	Vergleich zwischen Theorie und Experiment	152
8	Resümee und Ausblick	161
A	Formelsammlung	163
	Literaturverzeichnis	169
	Danksagung	179

Kapitel 1

Die Dichtefunktionaltheorie

Ein wichtiges Thema in der modernen Physik ist die quantenmechanische Beschreibung von Vielteilchensystemen, insbesondere die Beschreibung von Elektronen in kondensierter Materie. Analytische oder numerisch exakte Lösungen der Schrödingergleichung sind nur für wenige einfache Systeme möglich. Mit der Entwicklung der Dichtefunktionaltheorie (density functional theory: DFT) [JG89] gelang ein bedeutender Durchbruch. Grundlage der DFT bilden die Arbeiten von Hohenberg und Kohn [HK64] sowie Kohn und Sham [KS65] und Levy [Lev79, Lev82].

Die DFT nutzt die Tatsache, daß die Energie eines Vielelektronensystem sich als Funktional der Elektronendichte auffassen läßt. Dabei ist die Grundzustandsenergie das Minimum der Energie bezüglich aller (mit der Antisymmetrie der Vielelektronwellenfunktionen verträglichen) Elektronendichten. Die Grundzustandselektronendichte bestimmt alle Eigenschaften des Vielelektronensystem.

Die DFT ist heute Grundlage für eine realistische, quantitative "*ab initio*" Beschreibung des Grundzustands von Vielelektronensystemen. Dichtefunktionalrechnungen, welche nur auf quantenmechanischen Prinzipien beruhen, also keine anpaßbaren Parameter benötigen, tragen heute erheblich zu einem besseren Verständnis der mikroskopischen Eigenschaften von Atomen, Molekülen und Festkörpern bei. In der lokalen Dichte-Näherung (local density approximation: LDA) wird die Austausch- und Korrelationswechselwirkung der Elektronen durch ein universelles Funktional dargestellt, welches nur lokal von der totalen Elektronendichte abhängt. Semiempirische, analytische Formeln werden benutzt, um dieses Funktional darzustellen. Die DFT leistet eine Abbildung des komplizierten Vielteilchenproblems auf ein System von Einteilchenschrödingergleichungen in einem effektiven Potential. Dieses Gleichungssystem, die Kohn-Sham-Gleichungen, muß selbstkonsistent gelöst werden.

1.1 Die Energie als Dichtefunktional

Ein System von N Elektronen in einem äußeren Potential wird durch folgenden Hamiltonoperator \mathcal{H} beschrieben:

$$\mathcal{H} = T + V_{ee} + \sum_{i=1}^N V_{ext}(X_i) \quad (1.1)$$

$$T = -\frac{\hbar^2}{2m_e} \sum_{i=1}^N \partial_{\vec{r}_i}^2 \quad (1.2)$$

$$V_{ee} = \frac{1}{2} \sum_{i \neq j}^N \frac{e^2}{|\vec{r}_i - \vec{r}_j|} \quad (1.3)$$

T und V_{ee} sind die Operatoren der kinetischen Energie und der Elektron–Elektron–Coulomb–Wechselwirkung. Ein externes Potential V_{ext} ist z.B. das Coulomb–Potential der Atomrümpfe, welches im Pseudopotentialformalismus durch die Pseudopotentiale der Ionenrümpfe ersetzt wird.

In dieser Arbeit werden atomare Einheiten benutzt. Energien werden in Rydberg angegeben und Längen in Bohrschen Radien:

$$\hbar = 1, \quad m_e = \frac{1}{2}, \quad e^2 = 2, \quad c = \frac{2}{\alpha} = 274.074, \\ 1\text{Ry} \cong 13.6058 \text{ eV}, \quad a_B \cong 0.529177 \text{ \AA} = 0.529177 \times 10^{-10} \text{ m}$$

\mathcal{H} wirkt auf die Vielteilchenwellenfunktionen

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

Die meßbaren Eigenschaften des Vielelektronensystems können im Prinzip durch Erwartungswerte des Hamiltonian mit den Vielteilchenwellenfunktionen berechnet werden. Das Auffinden von Vielteilcheneigenfunktionen führt in Ortsraumdarstellung auf eine $3N$ -dimensionale Eigenwert-Differentialgleichung. Dabei sind physikalische Nebenbedingungen wie Normierbarkeit und Pauliprinzip zu berücksichtigen. Das Pauliprinzip besagt, daß die Vielteilchenwellenfunktionen aufgrund der quantenmechanischen Ununterscheidbarkeit der Elektronen vollständig antisymmetrisch gegen Vertauschung zweier Teilchen–Koordinaten sein müssen. Diese mathematische Aufgabenstellung wird bereits für kleine Systeme so komplex, daß eine analytische oder numerisch exakte Lösung in der Regel nicht möglich ist.

Glücklicherweise kommt man, zumindest für den nichtentarteten Grundzustand, mit der Dichtefunktionaltheorie zu einer sehr viel einfacheren Beschreibung des wechselwirkenden Elektronengases. Hohenberg und Kohn [HK64] zeigten, daß in diesem

Fall die Elektronendichte als Systemvariable genügt:

$$n(\vec{r}) := \langle \Psi | \sum_{i=1}^N \delta(\vec{r} - \vec{r}_i) | \Psi \rangle = \int \prod_{i=1}^N dX_i \sum_{i=1}^N \delta(\vec{r} - \vec{r}_i) |\Psi(X_1, \dots, X_N)|^2 \quad (1.4)$$

Das Theorem von Hohenberg und Kohn besteht aus zwei Teilen und besagt:

- Die Grundzustandsdichte $n_0(\vec{r})$ eines Vielelektronensystems enthält alle Information über das System, d.h. alle elektronischen Eigenschaften des Systems sind Funktionale der Grundzustandsdichte $n_0(\vec{r})$.

Insbesondere existiert zu jeder möglichen Grundzustandsdichte $n_0(\vec{r})$, abgesehen von einer additiven Konstanten, umkehrbar eindeutig ein äußeres Potential V_{ext} . Somit sind auch der Hamiltonoperator und alle Grundzustandseigenschaften, wie z.B. die Grundzustandsenergie E_0 , durch die Grundzustandsdichte $n_0(\vec{r})$ eindeutig festgelegt.

- Es existiert ein Energiefunktional $E[n]$, welches unter der Nebenbedingung der Ladungserhaltung ($N = \int d^3r n(\vec{r})$) durch die Grundzustandsdichte $n_0(\vec{r})$ minimiert wird. Für die Grundzustandsdichte $n_0(\vec{r})$ liefert es die Grundzustandsenergie E_0 :

$$E[n] > E[n_0] = E_0 \text{ für alle } n(\vec{r}) \neq n_0(\vec{r}) \quad (1.5)$$

E_0 und $n_0(\vec{r})$ können also durch Variation des Energiefunktionals nach $n(\vec{r})$ bzw. durch Lösen der daraus resultierenden Euler-Lagrange-Gleichung bestimmt werden.

Während Hohenberg und Kohn den Beweis dieses Theorems indirekt führten, konnte Levy 1979 einen direkten Beweis mit Hilfe eines Variationsverfahren angeben. Das Energiefunktional $E[n]$ ist nach Levy [Lev79] wie folgt definiert:

$$E[n] := \min_{\{\Psi\}} \langle \Psi | \hat{H} | \Psi \rangle \quad (1.6)$$

Dabei ist der Erwartungswert des Hamiltonoperators bezüglich aller normierten, vollständig antisymmetrischen Vielteilchenwellenfunktionen Ψ zu minimieren, die eine vorgegebene Teilchendichte $n(\vec{r})$ reproduzieren.

Die Grundzustandsdichte $n_0(\vec{r})$ und die zugehörige Grundzustandsenergie E_0 werden durch Minimierung des Energiefunktionals $E[n]$ bestimmt:

$$\delta E[n] = 0 \quad (1.7)$$

Die Nebenbedingung der Teilchenzahlerhaltung

$$\int d^3r n(\vec{r}) = N \quad (1.8)$$

kann durch einen Lagrange-Parameter μ berücksichtigt werden. Dies führt auf die Minimierung eines modifizierten Energiefunktionals

$$E[n] - \mu \left(\int d^3r n(\vec{r}) - N \right), \quad (1.9)$$

bzw. auf die folgende Euler–Lagrange–Gleichung:

$$\left. \frac{\delta E[n(\vec{r})]}{\delta n(\vec{r})} \right|_{n_0(\vec{r})} = \mu \quad (1.10)$$

Leider konnte jedoch aus der Definition von Levy keine explizite Darstellung des Energiefunktionals hergeleitet werden. Die genaue Form des Energiefunktionals ist also unbekannt und man ist daher auf sinnvolle Näherungen angewiesen.

1.2 Die Kohn–Sham–Gleichungen

Um die explizite Darstellung des Energiefunktionals in Gleichung (1.10) zu umgehen, erweist sich die Idee von Kohn und Sham [KS65] als nützlich, das Problem auf ein einfacheres System nicht wechselwirkender Elektronen in einem effektiven Potential $V_{eff}(\vec{r})$ abzubilden. Dazu ist es zweckmäßig, einzelne Energiebeiträge abzuspalten und exakt zu behandeln:

$$E[n] = T_0[n] + E_{ext}[n] + E_H[n] + E_{xc}[n] \quad (1.11)$$

$T_0[n]$ ist das Funktional der kinetischen Energie eines wechselwirkungsfreien Elektronengases. Die klassische potentielle Energie der Ladungsdichte $n(\vec{r})$ in einem externen Potential, bzw. die Wechselwirkung mit den Atomkernen oder Ionenrümpfen wird durch $E_{ext}[n]$ beschrieben:

$$E_{ext}[n] = \int d^3r n(\vec{r}) V_{ext}(\vec{r}) \quad (1.12)$$

Die Coulomb-Wechselwirkungsenergie in der Hartree–Näherung ist durch

$$E_H[n] = \frac{1}{2} e^2 \int d^3r d^3r' \frac{n(\vec{r})n(\vec{r}')}{|\vec{r} - \vec{r}'|} \quad (1.13)$$

gegeben. Das Austausch–Korrelations–Funktional $E_{xc}[n]$ enthält nun alle unbekanntesten Anteile zum Energiefunktional. Dies sind alle Korrekturen zu den vorhergegangenen Termen, also alle klassischen und quantenmechanischen Vielteilcheneffekte:

- Die Austauschwechselwirkung rührt von der Antisymmetrie der Vielteilchenwellenfunktionen her (Pauli-Prinzip). Dies bewirkt, daß sich Elektronen mit parallelen Spins nicht beliebig nahe kommen können. Dieser Korrelationseffekt führt somit zu einer Absenkung der elektrostatischen Energie. Die Austauschwechselwirkung beruht rein auf der quantenmechanischen Ununterscheidbarkeit der Elektronen und hat somit kein klassisches Analogon. Es handelt sich hier also um einen rein quantenmechanischen Effekt.
- Die klassische Coulomb-Korrelation beruht darauf, daß sich gleichnamig geladene Teilchen wegen der Coulomb-Wechselwirkung soweit wie möglich aus dem Weg gehen. Auch dies führt zu einer Absenkung der elektrostatischen Energie. Hierbei handelt es sich um einen rein klassischen Effekt. Er ist somit auch unabhängig vom Spin der Teilchen.
- Auch die kinetische Energie eines wechselwirkenden Elektronengases hängt von Korrelationseffekten ab, welche die Bewegungsfreiheit der Elektronen einschränken. Dies wiederum ist nach der Heisenbergschen Unschärferelation mit einem Anstieg der kinetischen Energie verknüpft. Korrekturen zur kinetischen Energie des wechselwirkungsfreien Elektronengases sind ebenfalls in $E_{xc}[n]$ enthalten.

Die Kohn-Sham-Gleichungen erhält man mit Hilfe eines speziellen Ansatzes für die Dichte $n(\vec{r})$. Im Falle wechselwirkungsfreier Elektronen erhält man die Dichte als Summe von Quadraten von Einteilchenwellenfunktionen, die mit den Besetzungszahlen f_i gewichtet werden.

$$n(\vec{r}) = 2 \sum_i f_i |\psi_i(\vec{r})|^2 \quad (1.14)$$

Die Besetzungszahlen genügen einer Fermiverteilung und sind auf die Teilchenzahl N normiert:

$$N = 2 \sum_i f_i \quad (1.15)$$

Die Summe über den Index i läuft über alle auf eins normierten Einteilchenfunktionen $\psi_i(\vec{r})$ mit Besetzungszahlen $0 \leq f_i \leq 1$. Der Faktor zwei berücksichtigt die Spinartung.

Zerlegt man die Ladungsdichte in (fiktive) Einteilchenfunktionen, so nimmt die kinetische Energie folgende Gestalt an:

$$T_0[n] = 2 \sum_i f_i \int d^3r \psi_i^*(\vec{r}) \left(-\frac{\hbar^2}{2m_e} \partial_{\vec{r}_i}^2 \right) \psi_i(\vec{r}) \quad (1.16)$$

Durch die Wiedereinführung der Einteilchenwellenfunktionen wird die kinetische Energie nicht wechselwirkender Elektronen exakt behandelt.

Um das Energiefunktional $E[n]$ zu minimieren, wird nicht nach der Dichte $n(\vec{r})$, sondern nach den (komplex konjugierten) Einteilchenwellenfunktionen $\psi_i^*(\vec{r})$ variiert. Dabei wird die Nebenbedingung der Ladungserhaltung $\int d^3r n(\vec{r}) = N$ durch die Normierung der Wellenfunktionen erfüllt:

$$\int d^3r |\psi_i(\vec{r})|^2 = 1 \quad (1.17)$$

Die Normierungsbedingung für $\psi_i(\vec{r})$ wird als Lagrange-Bedingung mit Lagrange-Parameter ε_i im Energiefunktional berücksichtigt. Die Variation der Gesamtenergie führt so auf ein System von gekoppelten Eigenwertgleichungen, welche für die Einteilchenfunktionen $\psi_i(\vec{r})$ die Form einer Schrödinger-Gleichung haben. Dies sind die sogenannten Kohn-Sham-Gleichungen:

$$\frac{1}{2f_i} \frac{\delta E[n(\vec{r})]}{\delta n(\vec{r})} \frac{\delta n[\dots, \psi_i, \dots]}{\delta \psi_i^*(\vec{r})} = \left(-\frac{\hbar^2}{2m_e} \partial_{\vec{r}}^2 + V_{eff}[n(\vec{r})] \right) \psi_i(\vec{r}) = \varepsilon_i \psi_i(\vec{r}) \quad (1.18)$$

bzw.

$$\left(-\frac{\hbar^2}{2m_e} \partial_{\vec{r}}^2 + V_{ext}(\vec{r}) + V_H[n(\vec{r})] + V_{xc}[n(\vec{r})] \right) \psi_i(\vec{r}) = \varepsilon_i \psi_i(\vec{r}) \quad (1.19)$$

Die Kohn-Sham-Gleichungen sind selbstkonsistent zu lösen.

Das effektive Potential V_{eff} setzt sich aus drei Beiträgen zusammen, dem externen Potential V_{ext} , dem Hartree-Potential

$$V_H[n(\vec{r})] = \frac{\delta}{\delta n(\vec{r})} \left(\frac{1}{2} e^2 \int d^3r d^3r' \frac{n(\vec{r}) n(\vec{r}')}{|\vec{r} - \vec{r}'|} \right) \quad (1.20)$$

$$= e^2 \int d^3r' \frac{n(\vec{r}')}{|\vec{r} - \vec{r}'|} \quad (1.21)$$

und dem Austausch-Korrelations-Potential

$$V_{xc}[n(\vec{r})] = \frac{\delta E_{xc}[n(\vec{r})]}{\delta n(\vec{r})} \quad (1.22)$$

Das effektive Potential ist somit ein Einteilchenpotential, welches seinerseits wieder ein Funktional der Dichte ist.

Die Eigenwerte der Kohn-Sham-Gleichungen (1.19) sind formal Einteilchenenergien, haben aber strenggenommen in der DFT als Lagrange-Parameter keine Rechtfertigung als solche. Ebenso handelt es sich bei den Einteilchenwellenfunktionen ψ_i strenggenommen nur um mathematische Hilfsgrößen, welche formal keine physikalische Bedeutung haben. Bei der Interpretation der Wellenfunktionen als Einteilchenfunktionen sollte man also immer die erforderliche Sorgfalt und Vorsicht walten

lassen. Ähnlich dem Koopman'schen Theorem in der Hartree-Fock-Theorie ist es trotzdem im allgemeinen möglich, die ε_i und ψ_i als Eigenwerte und Einteilchenwellenfunktionen zu interpretieren [JG89]. In der Tat stimmen bei delokalisierten Zuständen, bei denen die Korrelation gegenüber dem Austausch vernachlässigt werden kann, die spektroskopisch gemessenen Anregungsenergien mit den in der DFT berechneten erstaunlich gut überein. Dies gilt nicht für stark lokalisierte wie z.B. atomare Zustände. Die in dieser Arbeit berechneten Oberflächenzustände sind hinreichend delokalisiert und können getrost als physikalische Zustände interpretiert werden.

1.3 Die Spindichtefunktionaltheorie

Die Behandlung magnetischer Systeme erfordert eine Erweiterung der Dichtefunktionaltheorie, die Spindichtefunktionaltheorie. Diese ist immer dann relevant, wenn der Grundzustand magnetisch wird, wie z.B. beim Ferro- oder Antiferromagnetismus von Festkörpern, beim Oberflächenmagnetismus, bei den meisten Atomen oder bei der Anwendung externen Magnetfelder.

Diese Verallgemeinerung der Dichtefunktionaltheorie kann leicht erreicht werden, indem anstatt der Ladungsdichte $n(\vec{r})$ die Spindichten $n^\uparrow(\vec{r})$ und $n^\downarrow(\vec{r})$ für die Elektronen mit Spin auf ($\sigma = \uparrow$) und Spin ab ($\sigma = \downarrow$) als grundlegende Systemvariablen eingeführt werden. Die Teilchen- und Magnetisierungsdichten lauten dann:

$$n(\vec{r}) = n^\uparrow(\vec{r}) + n^\downarrow(\vec{r}) \quad (1.23)$$

$$m(\vec{r}) = n^\uparrow(\vec{r}) - n^\downarrow(\vec{r}) \quad (1.24)$$

Das Theorem von Hohenberg und Kohn geht dann über in ein Spindichtefunktionaltheorem [vBH72]:

$$E[n^\uparrow, n^\downarrow] \geq E[n_0^\uparrow, n_0^\downarrow] = E_0 \quad (1.25)$$

Analog zum nicht spinpolarisierten Fall ergeben sich zugehörigen Kohn-Sham-Gleichungen durch Variation nach spinabhängigen Einteilchenfunktionen.

$$\frac{1}{f_i^\sigma} \frac{\delta E}{\delta \psi_i^{\sigma*}} = \left(-\frac{\hbar^2}{2m_e} \partial_{\vec{r}}^2 + V_{eff}^\sigma[n^\uparrow(\vec{r}), n^\downarrow(\vec{r})] \right) \psi_i^\sigma(\vec{r}) = \varepsilon_i^\sigma \psi_i^\sigma \quad (1.26)$$

$$n^\sigma(\vec{r}) = \sum_i f_i^\sigma |\psi_i^\sigma(\vec{r})|^2 \quad (1.27)$$

Ein verallgemeinertes Austausch-Korrelations-Funktional führt dabei zu einem spinabhängigen effektiven Potential.

1.4 Die Lokale–Dichte–Näherung

Alle Schwierigkeiten der DFT stecken in dem in seiner genauen Form unbekanntem Austausch–Korrelations–Funktional $E_{xc}[n]$. In diesem sind alle komplizierten Vielteilcheneffekte enthalten. Eine explizite Darstellung ist Voraussetzung für den sinnvollen Einsatz der DFT. Dies ist die einzige Stelle der ansonsten exakten Theorie, an der man auf eine Näherung angewiesen ist.

Eine sehr einfache, jedoch überraschend gute Näherung ist die lokale Dichtenäherung bzw. lokale Spindichtenäherung. Die LDA bzw. LSDA behandelt ein inhomogenes System wechselwirkender Elektronen lokal wie ein homogenes System. Ausgangspunkt ist die Austausch–Korrelationsenergiedichte $\epsilon_{xc}[n]$ pro Teilchen eines homogenen Elektronengases. Mit dieser Größe wird das unbekannte Austausch–Korrelationsfunktional genähert.

$$E_{xc}^{LDA}[n] = \int d^3r n(\vec{r}) \epsilon_{xc}[n(\vec{r})] \quad (1.28)$$

$$V_{xc}^{LDA}[n] = \epsilon_{xc}[n(\vec{r})] + n(\vec{r}) \frac{\partial \epsilon_{xc}}{\partial n}(n(\vec{r})) \quad (1.29)$$

Obwohl diese Näherung nur für langsam variierende Dichten offensichtlich ist, erhält man auch für stark inhomogene Systeme gute Ergebnisse, was durch die Mittelung der LDA über das Austausch–Korrelationsloch erklärt werden kann [JG89].

Für die Austausch–Korrelations–Energie sind in unserem Computerprogramm vier Approximationen implementiert. Die älteste geht dabei auf E. Wigner zurück [Wig34]. Ferner sind es das Funktional von v. Barth und Hedin in den Parametrisierungen von v. Barth und Hedin [vBH72] und Moruzzi, Janak, Williams [MJW78] sowie die Parametrisierung von Vosko, Wilk, Nusair [VWN80], die auf eine Monte–Carlo–Simulation des homogenen Elektronengases von Ceperley und Alder [CA80] zurückgeht. Da die Monte–Carlo–Resultate von Ceperley und Alder die bisher besten lokalen Näherungen darstellen, wird, falls nicht anders vermerkt, die Parametrisierung von Vosko, Wilk und Nusair benutzt. Alle Parametrisierungen, bis auf die von Wigner, sind optional mit relativistischer Korrektur von MacDonald und Vosko [MV79] implementiert. Es zeigt sich in der Praxis, daß Grundzustandseigenschaften wie Gitterkonstanten und elastische Module von Rechnungen mit unterschiedlichen Austausch–Korrelations–Funktionalen sich nur wenig unterscheiden. Allerdings dürfen aus Konsistenzgründen nur Rechnungen mit der gleichen Näherung für $\epsilon_{xc}[n]$ verglichen werden.

Die Stärke der LDA liegt in der Fähigkeit physikalische Trends gut zu beschreiben, während absolute Zahlen nur eine relativ begrenzte Aussagekraft haben.

1.5 Die Gesamtenergie des elektronischen Systems

Bei der Berechnung der Gesamtenergie $E[n]$ nach (1.11) ist es numerisch von Vorteil, die kinetische Energie nicht explizit aus (1.16) auszurechnen, sondern indirekt unter Verwendung der Kohn–Sham–Gleichungen (1.19) bzw. (1.26). Multiplikation dieser Gleichung mit $2f_i\psi_i^*$ bzw. $f_i^\sigma\psi_i^{\sigma*}$, anschließende Ortsintegration und Summation über alle besetzten Zustände i und σ ergibt:

$$T_0[n] = \sum_i 2f_i\varepsilon_i - \int d^3r n(\vec{r}) V_{eff}[n(\vec{r})] \quad (1.30)$$

$$T_0[n^\uparrow, n^\downarrow] = \sum_{i\sigma} f_i^\sigma \varepsilon_i^\sigma - \sum_\sigma \int d^3r n^\sigma(\vec{r}) V_{eff}^\sigma[n^\uparrow(\vec{r}), n^\downarrow(\vec{r})] \quad (1.31)$$

Diese Gleichungen gelten natürlich nur im selbstkonsistenten Fall, denn bei jedem iterativen Lösungsverfahren des Selbstkonsistenzproblems wird das effektive Potential für die Wellenfunktionen des (j)-ten Selbstkonsistenzschritts von der Ladungsdichte des vorhergehenden Selbstkonsistenzschritts $n^{(j-1)}(\vec{r})$ aufgebaut. Diese unterscheidet sich von der Ladungsdichte $n^{(j)}(\vec{r})$, die aus den aktuellen Wellenfunktionen dieses Selbstkonsistenzschritts aufgebaut wird. Damit ergibt sich für die kinetische Energie der Wellenfunktionen des (j)-ten Selbstkonsistenzschritts:

$$T_0[n^{(j)}] = \sum_i 2f_i\varepsilon_i - \int d^3r n^{(j)}(\vec{r}) V_{eff}[n^{(j-1)}(\vec{r})] \quad (1.32)$$

$$T_0[n^{\uparrow(j)}, n^{\downarrow(j)}] = \sum_{i\sigma} f_i^\sigma \varepsilon_i^\sigma - \sum_\sigma \int d^3r n^{\sigma(j)}(\vec{r}) V_{eff}^\sigma[n^{\uparrow(j-1)}(\vec{r}), n^{\downarrow(j-1)}(\vec{r})] \quad (1.33)$$

Einen Ausdruck für die totale elektronische Energie erhält man durch Einsetzen der kinetischen Energie (1.32) in Gleichung (1.11). Für den nicht spinpolarisierten Fall ergibt sich somit:

$$E[n^{(j)}] = 2 \sum_i f_i \varepsilon_i + \int d^3r n^{(j)}(\vec{r}) \left(\frac{1}{2} V_H [n^{(j)}(\vec{r})] - V_H [n^{(j-1)}(\vec{r})] \right) + E_{xc} [n^{(j)}(\vec{r})] - \int d^3r n^{(j)}(\vec{r}) V_{xc} [n^{(j-1)}(\vec{r})] \quad (1.34)$$

$$E^{LDA}[n^{(j)}] = 2 \sum_i f_i \varepsilon_i + \int d^3r n^{(j)}(\vec{r}) \left(\frac{1}{2} V_H [n^{(j)}(\vec{r})] - V_H [n^{(j-1)}(\vec{r})] \right) + \int d^3r n^{(j)}(\vec{r}) \left(\varepsilon_{xc} [n^{(j)}(\vec{r})] - V_{xc} [n^{(j-1)}(\vec{r})] \right) \quad (1.35)$$

Analog ergibt sich im spinpolarisierten Fall aus Gleichung (1.33):

$$E [n^{\uparrow(j)}, n^{\downarrow(j)}] = \sum_{i\sigma} f_i^\sigma \varepsilon_i^\sigma + \int d^3r n^{(j)}(\vec{r}) \left(\frac{1}{2} V_H [n^{(j)}(\vec{r})] - V_H [n^{(j-1)}(\vec{r})] \right) + E_{xc} [n^{\uparrow(j)}, n^{\downarrow(j)}]$$

$$- \sum_{\sigma} \int d^3 r n^{\sigma(j)} V_{xc}^{\sigma} [n^{\uparrow(j-1)}(\vec{r}), n^{\downarrow(j-1)}(\vec{r})] \quad (1.36)$$

$$\begin{aligned} E^{LSDA}[n^{\uparrow}, n^{\downarrow}] &= \sum_{i\sigma} f_i^{\sigma} \epsilon_i^{\sigma} + \int d^3 r n^{(j)}(\vec{r}) \left(\frac{1}{2} V_H [n^{(j)}(\vec{r})] - V_H [n^{(j-1)}(\vec{r})] \right) \\ &\quad + \sum_{\sigma} \int d^3 r n^{\sigma(j)}(\vec{r}) \epsilon_{xc} [n^{\uparrow(j)}(\vec{r}), n^{\downarrow(j)}(\vec{r})] \\ &\quad - \sum_{\sigma} \int d^3 r n^{\sigma(j)}(\vec{r}) V_{xc}^{\sigma} [n^{\uparrow(j-1)}(\vec{r}), n^{\downarrow(j-1)}(\vec{r})] \end{aligned} \quad (1.37)$$

Im Fall von perfekter Selbstkonsistenz vereinfachen sich die Gleichungen:

$$E[n] = 2 \sum_i f_i \epsilon_i - E_H[n] + E_{xc}[n] - \int d^3 r n(\vec{r}) V_{xc}[n(\vec{r})] \quad (1.38)$$

$$E^{LDA}[n] = 2 \sum_i f_i \epsilon_i - E_H[n] - \int d^3 r (n(\vec{r}))^2 \frac{\partial \epsilon_{xc}[n]}{\partial n}(\vec{r}) \quad (1.39)$$

$$E[n^{\uparrow}, n^{\downarrow}] = \sum_{i\sigma} f_i^{\sigma} \epsilon_i^{\sigma} - E_H[n] + E_{xc}[n^{\uparrow}, n^{\downarrow}] - \sum_{\sigma} \int d^3 r n^{\sigma}(\vec{r}) V_{xc}^{\sigma}[n(\vec{r})] \quad (1.40)$$

$$E^{LSDA}[n^{\uparrow}, n^{\downarrow}] = \sum_{i\sigma} f_i^{\sigma} \epsilon_i^{\sigma} - E_H[n] - \int d^3 r n(\vec{r}) \sum_{\sigma} \left(n^{\sigma}(\vec{r}) \frac{\partial \epsilon_{xc}[n^{\uparrow}, n^{\downarrow}]}{\partial n^{\sigma}}(\vec{r}) \right) \quad (1.41)$$

Der numerische Vorteil liegt darin, daß alle explizit dichteunabhängigen Anteile der totalen Energie in der Summation über die Eigenwerte enthalten sind.

Im Verbindung mit der Pseudopotentialtheorie, in der $n(\vec{r})$ ja die Valenzladungsdichte ist, wird in der Austauschkorrelationsenergie häufig eine feste, partielle Rumpfladungsdichte $n_{pc}(\vec{r})$ berücksichtigt (siehe Kapitel 2.4):

$$\left(E_{xc}[n], V_{xc}[n] \right) \longrightarrow \left(E_{xc}[n+n_{pc}], V_{xc}[n+n_{pc}] \right) \quad (1.42)$$

$$E[n, n_{pc}] = T_0[n] + E_{ext}[n] + E_H[n] + E_{xc}[n+n_{pc}] \quad (1.43)$$

Damit ändern sich die Gleichungen für die Gesamtenergie des elektronischen Systems für den nicht spinpolarisierten Fall wie folgt:

$$\begin{aligned} E[n^{(j)}] &= 2 \sum_i f_i \epsilon_i + \int d^3 r n^{(j)}(\vec{r}) \left(\frac{1}{2} V_H [n^{(j)}(\vec{r})] - V_H [n^{(j-1)}(\vec{r})] \right) \\ &\quad + E_{xc} [n^{(j)} + n_{pc}] - \int d^3 r n^{(j)}(\vec{r}) V_{xc} [n^{(j-1)}(\vec{r}) + n_{pc}(\vec{r})] \end{aligned} \quad (1.44)$$

$$\begin{aligned} E^{LDA}[n^{(j)}] &= 2 \sum_i f_i \epsilon_i + \int d^3 r n^{(j)}(\vec{r}) \left(\frac{1}{2} V_H [n^{(j)}(\vec{r})] - V_H [n^{(j-1)}(\vec{r})] \right) \\ &\quad + \int d^3 r \left(n^{(j)}(\vec{r}) + n_{pc}(\vec{r}) \right) \epsilon_{xc} [n^{(j)}(\vec{r}) + n_{pc}(\vec{r})] \\ &\quad - \int d^3 r n^{(j)}(\vec{r}) V_{xc} [n^{(j-1)}(\vec{r}) + n_{pc}(\vec{r})] \end{aligned} \quad (1.45)$$

Im Computerprogramm wird für die Berechnung der totalen Energie die Gleichung (1.45) benutzt. Ohne Berücksichtigung der partiellen Rumpfladung vereinfacht sich Gleichung (1.45) zu Gleichung (1.35).

Kapitel 2

Die Pseudopotentialmethode

Für die Berechnung derjenigen Eigenschaften von Festkörpern und Molekülen, die von der chemischen Bindung bestimmt sind, hat sich die Pseudopotentialmethode bewährt. Sie weist eine lange Tradition auf [Hei70, HW70]. Wurden früher vorwiegend Modellpseudopotentiale oder empirische Pseudopotentiale, die an spektroskopische Daten angefügt werden, verwendet [CH70], so markieren die Entwicklung semilokaler normerhaltender Pseudopotentiale durch Hamann, Schlüter und Chiang [HSC79] bzw. Bachelet, Hamann und Schlüter [BHS82] und deren nichtlokale, separable Darstellung durch Kleinman und Bylander [KB82] einen Meilenstein für die Anwendung von Pseudopotentialen. Durch sie wurden *ab initio* Strukturrechnungen mittels Pseudopotentialen im Rahmen der vorher besprochenen Dichtefunktionaltheorie überhaupt erst möglich.

Die Darstellung von Hamiltonian und Wellenfunktionen mittels einer Basis aus ebenen Wellen [IZC79], die Verwendung der Fast-Fourier-Transformation sowie die Entwicklung neuer iterativer Verfahren zur Minimierung des Energiefunktionals bzw. zur Diagonalisierung großer Matrizen [CP85, Dav75, Kos84, TPA89, PTA⁺92] haben in Verbindung mit der Pseudopotentialmethode in den letzten 15 Jahren zu sehr leistungsfähigen Methoden für die Behandlung struktureller Fragestellungen geführt [Ihm88, Pic89, PTA⁺92]. Das Bild wird komplettiert durch die Entwicklung glatter bzw. weicherer Pseudopotentiale [Ker80, Van85, RRKJ90, TM90, TM91] und verallgemeinerter separabler Potentiale [Ham89, GVM89, SAMJ89, Blö90, Van90, Cho92, BK92, SSO92], sowie durch die Beseitigung des "Ghost-state"-Problems [BK90, GKS90, GSS91, SSO92], so daß die Pseudopotentialmethode im Prinzip auf alle Elemente des Periodensystems und auf die Berechnung großer Systeme angewendet werden kann.

Uns steht ein Generierungsprogramm für Pseudopotentiale zur Verfügung, welches

im Rahmen einer Diplomarbeit [Eng92] entwickelt wurde. Implementiert sind das Schema von Bachelet, Hamann und Schlüter, "BHS"-Schema [BHS82], sowie dessen Modifikationen nach Vanderbilt, "Van"-Schema [Van85], und das Schema nach Rappe et. al., "RRKJ"-Schema [RRKJ90], welches leicht modifiziert wurde. Ebenfalls wurde ein Programm zur Analyse der logarithmischen Ableitungen analog zu [GSS91] entwickelt, um "Ghost-states" zu vermeiden.

Die verschiedenen Pseudopotentiale wurden von uns mit folgendem Ergebnis auf Glattheit und Transferierbarkeit untersucht: Entscheidend für die Glattheit, d.h. für die Konvergenz bezüglich der Basisgröße in einer Basis aus ebenen Wellen, ist die Fourierrepartitionierung der kinetischen Energie der Pseudowellenfunktionen. Diese wird in allen Schemata im wesentlichen durch die Wahl der Abschneideradien¹ R_c festgelegt: Auf der einen Seite darf der Abschneideradius nicht zu groß sein, denn dies verringert die Transferierbarkeit der Pseudopotentiale in andere chemische Umgebungen. Andererseits darf er aber auch nicht zu klein sein, denn dies führt zu tieferen Pseudopotentialen und somit zu einer schlechteren Konvergenz bezüglich Basis aus ebenen Wellen.

Keines der untersuchten Schemata bietet dabei entscheidende Vorteile. Dieses Ergebnis deckt sich mit den Ergebnissen anderer ähnlicher Untersuchungen [KHN92, LQPH93]. Einzig die Idee der sogenannten "ultrasofter" Pseudopotentiale von Vanderbilt [Van90] scheint den anderen Konzepten überlegen zu sein [FHK94]. Dieses Verfahren wurde von uns aber bisher nicht implementiert.

2.1 Idee der Pseudopotential-Methode

Bei der quantenmechanischen Beschreibung von Elektronen in kondensierter Materie kann man die räumliche Verteilung der Elektronen erfahrungsgemäß in zwei Bereiche unterteilen:

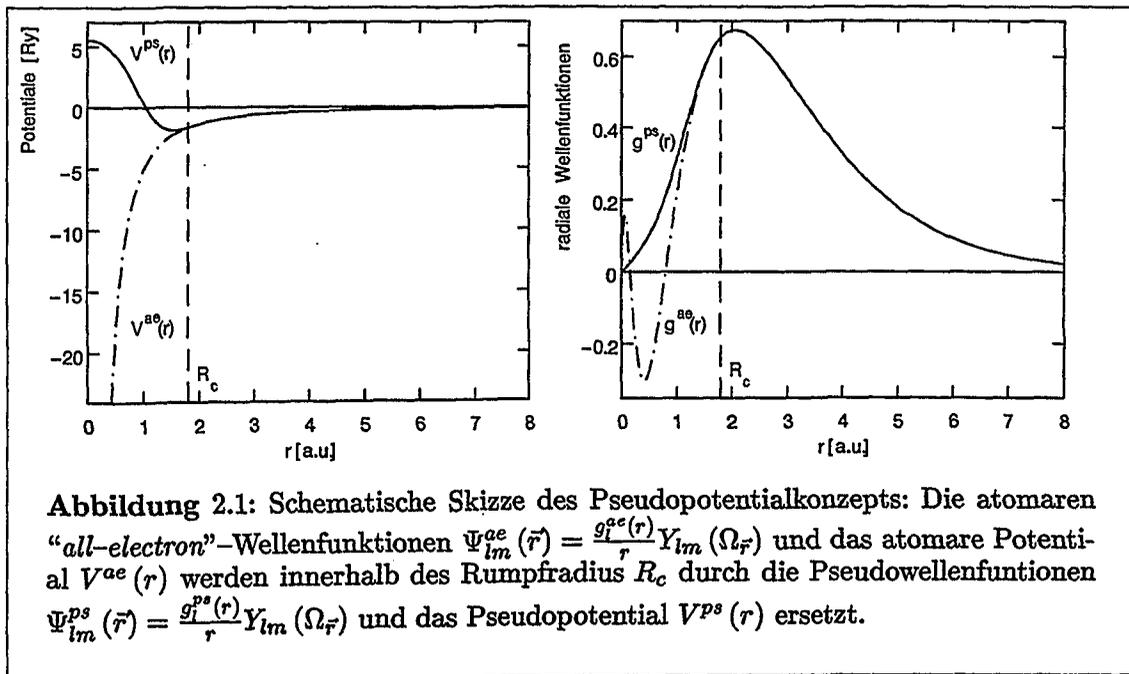
- Zum einen in den *Rumpfbereich* ($r < R_c$) in unmittelbarer Umgebung der Atomkerne, in dem sich die stark gebundenen Elektronen der inneren Schalen, die "Rumpfelektronen", aufhalten. Der Überlapp der Wellenfunktionen der Rumpfelektronen verschiedener Atomrümpfe ist nahezu Null und damit vernachlässigbar. Aus diesem Grund verhalten sich Rumpfelektronen in verschiedenen chemischen Umgebungen nahezu gleich und liefern fast keinen Beitrag zur chemischen Bindung. Dieser Sachverhalt liefert die Grundlage für die

¹Der Index c steht hier für "core" oder "cutoff".

“frozen-core”-Näherung. Der Rumpfbereich ist geprägt durch das stark anziehende Coulombpotential der Atomkerne. Deshalb zeigen Wellenfunktionen und Ladungsdichten dort kurzweilige Rumpfszillationen.

- Zum anderen in den *Valenzbereich*, der den ganzen Zwischengitterraum umfaßt und sich durch abgeschirmte, flache Potentiale auszeichnet. Wellenfunktionen – und somit auch die Ladungsdichten – sind dort, verglichen mit dem Rumpfbereich, glatt. Die Valenzelektronen reagieren sehr empfindlich auf unterschiedliche chemische Umgebungen und sind deshalb besonders wichtig bei strukturellen Fragestellungen.

Die Idee der Pseudopotentialmethode ist es nun, die elektronischen Freiheitsgrade der Rumpfelektronen zu eliminieren und die Valenzelektronen durch möglichst glatte knotenlose Wellenfunktionen ohne Rumpfszillationen in einem glatten nicht-singulären Pseudopotential zu beschreiben (siehe Abbildung 2.1):



In einem ersten Schritt werden die Rumpfelektronen “eingefroren” (“frozen-core”-Näherung). In der DFT wählt man typischerweise die Grundzustandsdichte der Rumpfelektronen eines freien Atoms. Die Valenzelektronen bewegen sich sozusagen in einem effektiven Rumpfpotential $V_{Rumpf}^{eff}(\vec{r})$ der Atomkerne und der Rumpfelektronen. In einem zweiten Schritt sucht man eine Beschreibung der Valenzelektronen

durch knotenlose Pseudowellenfunktionen² $\Psi_{val}^{ps}(\vec{r})$, die im Zwischengitterbereich möglichst exakt ist und die möglichst glatt in dem Rumpfbereich hinein fortgesetzt wird. Dies ist gleichbedeutend mit einer Ersetzung des Rumpfpotentials $V_{Rumpf}^{eff}(\vec{r})$ durch ein glattes Pseudopotential $V^{ps}(\vec{r})$. $V^{ps}(\vec{r})$ wird so angepaßt, daß die Pseudowellenfunktionen $\Psi_{val}^{ps}(\vec{r})$ Eigenfunktionen des Pseudopotentials $V^{ps}(\vec{r})$ zu den exakten Valenzeigenwerten ϵ_{val} der exakten Wellenfunktionen $\Psi_{val}(\vec{r})$ im exakten Rumpfpotential $V_{Rumpf}^{eff}(\vec{r})$ werden. Somit werden also die tiefen Coulombpotentiale durch flachere Pseudopotentiale ersetzt, die Valenzelektronen im Valenzbereich richtig beschreiben.

$$(V_{Rumpf}^{eff}(\vec{r}), \Psi_{val}^{ae}(\vec{r}), \epsilon_{val}) \longrightarrow (V^{ps}(\vec{r}), \Psi_{val}^{ps}(\vec{r}), \epsilon_{val}) \quad (2.1)$$

Bei numerischen Berechnungen bringt die Pseudopotentialmethode somit enorme Vorteile:

- Die Anzahl der elektronischen Freiheitsgrade wird stark reduziert, da elektronische Rump fzustände nicht mehr berechnet werden.
- Ein wesentlicher Vorteil ist die Reduzierung des numerischen Rauschens, da das Verhältnis von typischen Energiedifferenzen und Gesamtenergie in zweierlei Hinsicht reduziert wird:
 - In einer “all-electron”-Rechnung ist der Beitrag der Rumpfelektronen zur Gesamtenergie erheblich größer als der Beitrag der Valenzelektronen. Ein kleiner relativer Fehler bei der Berechnung des Energiebeitrages der Rumpfelektronen könnte somit typische Energiedifferenzen, wie sie z.B. bei chemischen Bindungen auftreten, leicht verdecken.
 - Die Pseudowellenfunktionen sind (nach Konstruktion) glatt und knotenlos. Damit ist der Beitrag der kinetische Energie der Pseudozustände zur Gesamtenergie um Größenordnungen kleiner als der der echten Valenzzustände, da die Rumpfoszillationen beseitigt sind.
- Der wohl bedeutendste Vorteil der Pseudopotentialmethode ist die Darstellbarkeit von Potential und Wellenfunktionen durch nichtlokalisierte Basissätze, wie z.B. durch eine Basis aus ebenen Wellen. Diese trägt einerseits der periodischen Struktur des Festkörpers Rechnung und ist andererseits mathematisch sehr einfach zu handhaben. Wie sich in den nächsten Kapiteln herausstellen wird, ist die Basis aus ebenen Wellen aus diesem Grund besonders geeignet für iterative Verfahren.

²Die Indices “ps” und “ae” stehen für “pseudo” und “all-electron”

2.2 Semilokale, normerhaltende Pseudopotentiale

Mit der Idee und Entwicklung von normerhaltenden Pseudopotentialen durch Hamann, Schlüter und Chiang [HSC79] bzw. Bachelet, Hamann und Schlüter [BHS82] gelang ein bedeutender Durchbruch für *ab initio* Rechnungen. Erst durch das Konzept der Normerhaltung wurden Pseudopotentiale im Rahmen der DFT einsetzbar, da die in dieser Theorie so wichtige Ladungsdichte richtig beschrieben wird.

Normerhaltende Pseudopotentiale werden in einer atomaren Referenzkonfiguration an eine einzige selbstkonsistente “*all-electron*”-Rechnung angepaßt. Bei Transfer eines “Pseudoatoms” in eine andere chemische Umgebung ändern sich Ladungsdichten, Potentiale und Valenzenergieeigenwerte. Der Fehler, der bei dieser Näherung gemacht wird, hängt, ebenso wie schon die “*frozen-core*”-Näherung allein, nur in 2. Ordnung von diesen Änderungen ab [vBG80, YC82]. Ursache dafür ist, daß durch die Normerhaltung die Streueigenschaften des Pseudopotentials an die des “*all-electron*”-Potentials angepaßt werden. Dies wird durch folgende Identität [YC82, Eng92] gesichert:

$$\delta \left(\frac{d}{dr} \ln g(r) \right) \Big|_{r=R_c} = -\frac{1}{g^2(R_c)} \int_0^{R_c} (\delta\varepsilon - \delta V(r)) g^2(r) dr \quad (2.2)$$

Diese Variationsgleichung gilt für alle Lösungen der radialen Schroedingergleichung:

$$\left(-\partial_r^2 + \frac{l(l+1)}{r^2} + V(r) \right) g(r) = \varepsilon g(r) \quad (2.3)$$

Ausgangspunkt zur Konstruktion eines normerhaltendes Pseudopotentials ist eine radialsymmetrische, atomare Dichtefunktionalrechnung der Elektronen in einer Referenzkonfiguration bis zur Selbstkonsistenz, bei der sowohl Valenz- als auch Rumpfelektronen berücksichtigt werden. Als Referenzkonfiguration dient meistens die Grundzustandskonfiguration des Atoms. Der radialsymmetrische Anteil der Wellenfunktionen wird durch die radialsymmetrischen Wellenfunktionen $g_l(r)$ dargestellt, der winkelabhängige Anteil durch die Kugelflächenfunktionen $Y_{lm}(\Omega_{\vec{r}})$:

$$\Psi_{lm}(\vec{r}) = \frac{g_l(r)}{r} Y_{lm}(\Omega_{\vec{r}}) \quad (2.4)$$

Für die Referenzkonfiguration wird nun zu jeder radialen, atomaren Valenzwellenfunktion³ $g_{v,i}^{ae}(r)$ zum Drehimpuls l eine Pseudowellenfunktion $g_i^{ps}(r)$ konstruiert, welche die folgenden Bedingungen erfüllt:

³Die Indices c, v stehen für “core” bzw. “valence”, der Index v wird bei den Pseudowellenfunktionen weggelassen.

- $g_i^{ps}(r)$ ist knotenlos, stetig differenzierbar und stimmt außerhalb eines geeignet zu wählenden "core-Radius" R_c mit $g_i^{ae}(r)$ überein.
- $g_i^{ps}(r)$ erfüllt die Bedingung der Normerhaltung:

$$\int_0^{R_c} |g_i^{ps}(r)|^2 dr = \int_0^{R_c} |g_{v,i}^{ae}(r)|^2 dr \quad (2.5)$$

Diese Pseudowellenfunktionen sind dann im ganzen Raum richtig normiert. Aus diesem Grund gehört zu den Pseudowellenfunktionen $g_i^{ps}(r)$ eine Pseudoladungsdichte $n^{ps}(r) = \sum_l f_l |g_i^{ps}(r)|^2$, die im Valenzbereich (außerhalb von R_c) mit der vollen Valenzladungsdichte $n_v^{ae}(r)$ übereinstimmt und exakt die richtige Valenzladung⁴ $Z_v = \int n^{ps}(r) dr$ enthält. Die Entartung bzgl. Spin ist dabei in den Besetzungszahlen f_l der Valenzzustände der atomaren Referenzkonfiguration berücksichtigt. Somit ist sichergestellt, daß die Pseudoladungsdichte $n^{ps}(r)$ für den Dichtefunktionalformalismus geeignet ist. Diese Pseudoladungsdichte $n^{ps}(r)$ kann man dazu verwenden, Austauschkorrelationspotentiale zu berechnen, die ebenfalls außerhalb von R_c mit den ae-Austauschkorrelationspotentialen übereinstimmen. Im folgenden sind die radialen ps-Gleichungen den entsprechenden ae-Gleichungen gegenübergestellt:

$$\left(-\partial_r^2 + \frac{l(l+1)}{r^2} - \frac{e^2 Z}{r} + V_{Hxc}[n^{ae}(r)] \right) g_{[c,v],l}^{ae}(r) = \varepsilon_{[c,v],l} g_{[c,v],l}^{ae}(r) \quad (2.6)$$

$$n^{ae}(r) = n_c^{ae}(r) + n_v^{ae}(r)$$

$$n_c^{ae}(r) = \sum_{l,c} f_{c,l} |g_{c,l}^{ae}(r)|^2$$

$$n_v^{ae}(r) = \sum_l f_{v,l} |g_{v,l}^{ae}(r)|^2$$

$$\left(-\partial_r^2 + \frac{l(l+1)}{r^2} + V_{ion,l}^{ps}(r) + V_{Hxc}[n^{ps}(r)] \right) g_i^{ps}(r) = \varepsilon_{v,l} g_i^{ps}(r) \quad (2.7)$$

$$n^{ps}(r) = \sum_l f_{v,l} |g_i^{ps}(r)|^2$$

Das Potential $V_{ion,l}^{ps}(r)$ ist das l -abhängige, ionische Pseudopotential, welches für die Valenzelektronen das effektive Potential von Atomkern und Rumpfelektronen ersetzt. $V_{Hxc}[n^{ps}(r)]$ ist die Summe aus Hartreepotential und Austauschkorrelationspotential.

Die verschiedenen Konstruktionsverfahren für normerhaltende Pseudopotentiale unterscheiden sich nur durch die Konstruktionsvorschrift für die Pseudowellenfunktionen $g_i^{ps}(r)$. In allen modernen Verfahren werden analytische Ansätze für die $g_i^{ps}(r)$ benutzt.

⁴Der Faktor $4\pi r^2$ aus Winkel- und Ortsintegration ist in $n(r)$ bereits enthalten.

Zunächst bestimmt man das volle abgeschirmte Potential

$$V_{scr,l}^{ps}(r) [n^{ps}(r)] = V_{ion,l}^{ps}(r) + V_{Hxc} [n^{ps}(r)]$$

durch direktes Invertieren der Schrödingergleichung (2.7):

$$V_{scr,l}^{ps}(r) = \varepsilon_{v,l} - \frac{l(l+1)}{r^2} + \frac{\partial_r^2 g_l^{ps}(r)}{g_l^{ps}(r)} \quad (2.8)$$

Dann werden Hartreepotential und Austauschkorrelationspotential der Pseudovalenzelektronen abgezogen, so daß sich ein drehimpulsabhängiges, ionisches Pseudopotential ergibt:

$$V_{ion,l}^{ps}(r) = V_{scr,l}^{ps}(r) - (V_H [n^{ps}(r)] + V_{xc} [n^{ps}(r)]) \quad (2.9)$$

Diesen Schritt bezeichnet man auch als lineares "unscreening". Durch Summation über alle Drehimpulse l erhält man schließlich ein semilokales (d.h. lokal im Radialteil, aber nichtlokal im Winkelanteil), normerhaltendes Pseudopotential:

$$\mathcal{V}^{ps} = \sum_l V_{ion,l}^{ps}(r) \mathcal{P}_l \quad (2.10)$$

Dabei ist \mathcal{P}_l der Projektionsoperator auf die Drehimpulskomponente l . Dieses Pseudopotential hat nun alle wichtigen Eigenschaften: Es reproduziert per Konstruktion in einer selbstkonsistenten Rechnung des Referenzatoms die atomaren Valenzeigenwerte $\varepsilon_{v,l}$; außerdem sind die zugehörigen Pseudowellenfunktionen knotenlos, richtig normiert und stimmen im Valenzbereich mit den richtigen Wellenfunktionen überein.

Das Verfahren läßt sich problemlos relativistisch oder skalarrelativistisch verallgemeinern [Eng92]. Unser Generierungsprogramm startet mit einer selbstkonsistenten, skalarrelativistischen Berechnung des Referenzatoms. So erzeugte Pseudopotentiale werden genauso benutzt wie nichtrelativistisch erzeugte Potentiale, enthalten aber relativistische Korrekturen aus dem Rumpfbereich.

Bei semilokalen Pseudopotentialen "sieht" jede Drehimpulskomponente einer Wellenfunktion ein anderes Potential $V_{ion,l}^{ps}(r)$. Alle $V_{ion,l}^{ps}(r)$ zeigen das gleiche langreichweitige Coulombverhalten $\sim -\frac{e^2 Z_v}{r}$. Aus diesem Grund ist es üblich, ein im Prinzip frei wählbares gemeinsames, langreichweitiges lokales Potential $V_{loc}(r)$ abzuspalten:

$$\begin{aligned} \mathcal{V}^{ps} &= \sum_l V_{ion,l}^{ps}(r) \mathcal{P}_l \\ &= V_{loc}(r) + \sum_l \Delta V_l^{SL}(r) \mathcal{P}_l \end{aligned} \quad (2.11)$$

$$\Delta V_l^{SL}(r) = V_{ion,l}^{ps}(r) - V_{loc}(r) \quad (2.12)$$

Die neuen semilokalen, l -abhängigen Potentiale $\Delta V_l^{SL}(r) \mathcal{P}_l$ sind dann kurzreichweitig, d.h. sie sind nur im Rumpfbereich für $r < R_c$ von Null verschieden. Die Drehimpulsentwicklung wird üblicherweise bei einem maximalen $l = l_{max}$ abgebrochen, wobei als lokales Potential $V_{ion, l_{max}+1}^{ps}(r)$ gewählt wird. Diese Näherung ist gleichbedeutend mit $V_{ion, l}^{ps}(r) = V_{ion, l_{max}+1}^{ps}(r)$ für alle $l > l_{max}$:

$$\mathcal{V}^{ps} = V_{loc}(r) + \sum_{l=0}^{l_{max}} \Delta V_l^{SL}(r) \mathcal{P}_l \quad (2.13)$$

$$V_{loc}(r) = V_{ion, l_{max}+1}^{ps}(r) \quad (2.14)$$

Das Pseudopotential setzt sich also aus einem lokalen Anteil $V_{loc}(r)$ und einem semilokalen Anteil $\mathcal{V}^{SL} = \sum_{l=0}^{l_{max}} \Delta V_l^{SL}(r) \mathcal{P}_l$ zusammen.

In unseren Rechnungen mit III-V Halbleitern ist $l_{max} = 1$, so daß das d -Potential noch exakt berücksichtigt wird. Bei Übergangsmetallen wird üblicherweise $l_{max} = 2$ gewählt, für die $2p$ -Reihe des Periodensystems reicht üblicherweise $l_{max} = 0$. Diese Näherung kann durch Berücksichtigung weiterer Drehimpulsbestandteile l auf einfache Weise verbessert werden.

Pseudopotentiale zu verschiedenen Drehimpulsbestandteilen l , die in unterschiedlichen Referenzkonfigurationen erzeugt worden sind, lassen sich problemlos miteinander kombinieren. Dies wird z.B. dann notwendig, wenn in der atomaren Grundzustandskonfiguration für höhere l -Komponenten keine gebundenen Valenzzustände existieren.

2.3 Separable, "Kleinman-Bylander"-Pseudopotentiale

Bei der Darstellung der Kohn-Scham-Gleichungen im Impulsraum müssen die Pseudopotentialmatrixelemente bezüglich der Basis aus ebenen Wellen ausgerechnet werden. Darüberhinaus müssen die Zustandsvektoren der Wellenfunktionen häufig mit der Matrix des Pseudopotentials multipliziert werden. Dies ist für semilokale Potentiale sehr aufwendig:

$$\langle \vec{k} + \vec{G}' | \Delta V_l^{SL}(r) \mathcal{P}_l | \vec{k} + \vec{G} \rangle = \frac{4\pi}{\Omega} (2l + 1) P_l(\cos \gamma_{\vec{k} + \vec{G}', \vec{k} + \vec{G}}) \quad (2.15)$$

$$\begin{aligned} & \times \int_{\infty}^{\infty} r^2 dr j_l(|\vec{k} + \vec{G}'| r) \Delta V_l^{SL}(r) j_l(|\vec{k} + \vec{G}| r) \\ \langle \vec{r} | \vec{k} + \vec{G} \rangle & = \frac{1}{\sqrt{\Omega}} e^{-i(\vec{k} + \vec{G}) \cdot \vec{r}} \end{aligned} \quad (2.16)$$

Ω ist das Volumen, in dem die ebenen Wellen auf 1 normiert sind (in der Anwendung ist das die Superzelle). Das Integral aus Gleichung (2.15) ist über den ganzen Raum auszuführen.

Es sind also im Computerprogramm $\frac{N_{\vec{G}}(N_{\vec{G}}+1)}{2}$ Matrixelemente auszurechnen und zu speichern, wobei $N_{\vec{G}}$ die Anzahl der ebenen Wellen ist. Die Matrixvektormultiplikation skaliert mit $N_{\vec{G}}^2$. Diese ungünstige Skalierung des Problems mit der Größe der Basis kann nach einer Idee von Kleinman und Bylander [KB82] durch eine vollständig nichtlokale, d.h. separable Darstellung der Pseudopotentiale vermieden werden:

$$\mathcal{V}^{ps} = V_{loc}(r) + \sum_{l=0}^{l_{max}} \Delta \mathcal{V}_l^{KB} \quad (2.17)$$

$$\Delta \mathcal{V}_l^{KB} = E_l^{KB} \sum_{m=-l}^{m=+l} |t_{lm}\rangle \langle t_{lm}| \quad (2.18)$$

$$E_l^{KB} = \frac{\langle \tilde{t}_{lm} | \tilde{t}_{lm} \rangle}{\langle \Psi_{lm}^{ps,ref}(\vec{r}) | \tilde{t}_{lm} \rangle} \quad (2.19)$$

$$\langle \vec{r} | t_{lm} \rangle = \frac{\langle \vec{r} | \tilde{t}_{lm} \rangle}{\sqrt{\langle \tilde{t}_{lm} | \tilde{t}_{lm} \rangle}} = \frac{t_l(r)}{r} Y_{lm}(\Omega_{\vec{r}}) \quad (2.20)$$

$$\langle \vec{r} | \tilde{t}_{lm} \rangle = \Delta V_l^{SL}(r) \Psi_{lm}^{ps,ref}(\vec{r}) \quad (2.21)$$

$$\Psi_{lm}^{ps,ref}(\vec{r}) = \frac{g_l^{ps,ref}(r)}{r} Y_{lm}(\Omega_{\vec{r}}) \quad (2.22)$$

Die $\Psi_{lm}^{ps,ref}(\vec{r})$ sind dabei die pseudisierten Wellenfunktionen der jeweiligen Referenzkonfiguration. Unmittelbar durch Einsetzen und unter Berücksichtigung der Orthogonalität der Kugelflächenfunktionen $Y_{lm}(\Omega_{\vec{r}})$ ergibt sich:

$$\langle t_{l'm'} | \Psi_{lm}^{ps,ref}(\vec{r}) \rangle = \delta_{l',l} \delta_{m',m} \langle t_{lm} | \Psi_{lm}^{ps,ref}(\vec{r}) \rangle \quad (2.23)$$

$$\Delta \mathcal{V}_l^{KB} \Psi_{lm}^{ps,ref}(\vec{r}) \equiv \Delta V_l^{SL}(r) \Psi_{lm}^{ps,ref}(\vec{r}) \quad (2.24)$$

Die neue separable Darstellung wirkt also auf die Referenzwellenfunktionen exakt genauso wie die alte semilokale Darstellung. Damit sind genauso alle Anforderungen an ein normerhaltendes Pseudopotential erfüllt. Für die Matrixelemente des "Kleinman-Bylander"-Potentials im Impulsraum ergibt sich ebenfalls eine separable Darstellung:

$$\langle \vec{k} + \vec{G} | \Delta \mathcal{V}_l^{KB} | \vec{k} + \vec{G}' \rangle = E_l^{KB} \sum_{m=-l}^l \langle \vec{k} + \vec{G} | t_{lm} \rangle \langle t_{lm} | \vec{k} + \vec{G}' \rangle \quad (2.25)$$

$$\langle \vec{k} + \vec{G} | t_{lm} \rangle = \frac{4\pi i^l}{\sqrt{\Omega}} Y_{lm}(\Omega_{\vec{k}+\vec{G}}) \int_{\infty}^0 j_l(|\vec{k} + \vec{G}|r) t_l(r) r dr \quad (2.26)$$

Entscheidender Vorteil dieser Transformation ist die günstigere Skalierung mit $N_{\bar{G}}$. So sind im Computerprogramm für die Potentialmatrix nur $2l + 1$ Vektoren der Länge $N_{\bar{G}}$ zu speichern. Die Matrixvektormultiplikationen reduzieren sich zu Vektorvektormultiplikationen der Länge $N_{\bar{G}}$.

Prinzipiell läßt sich so jedes semilokale Pseudopotential $\Delta V_l^{SL}(r)\mathcal{P}_l$ in ein separables Potential ΔV_l^{KB} transformieren. Obwohl ΔV_l^{KB} auf die Referenzwellenfunktionen identisch wirkt wie $\Delta V_l^{SL}(r)\mathcal{P}_l$, so ist die Wirkung auf beliebige Wellenfunktionen i.a. unterschiedlich. Bei der Konstruktion von separablen Pseudopotentialen ist besondere Vorsicht geboten, denn sie haben unter Umständen ein sogenanntes “ghost-state”-Problem: “ghost-states” sind zusätzliche Zustände mit niedrigeren Energien als den Valenzenergien aber mit Knoten im Rumpfbereich. Solche zusätzlichen Zustände sind unphysikalisch, und ein separables Pseudopotential, welches “ghost-states” erzeugt, ist unbrauchbar. Dieses Problem ist aber gelöst [BK90, GKS90, GSS91, SSO92].

Ergänzend sei noch erwähnt, daß man auch die semilokalen Pseudopotentiale auf Gridpunkten in eine separable Form “zwingen” kann [GVM89]. Diese Näherung wird z.B. verwendet, wenn das zugehörige Kleinman–Bylander–Pseudopotential Ghost–States erzeugt. Die Anzahl der benötigten Gridpunkte liegt dabei typischerweise bei zehn bis zwanzig, wobei die Genauigkeit dieser Methode von der Anzahl der verwendeten Gridpunkte abhängt.

2.4 Partielle Core–Korrektur

In der Ableitung des Konzepts normerhaltender Pseudopotentiale in Kapitel 2.2 geht man implizit davon aus, daß sich die Rumpfladungsdichte $n_c^{ae}(r)$ und die Pseudovalenzladungsdichte $n^{ps}(r)$ räumlich nicht überlappen, denn nur dann kann man das nichtlineare Austauschkorrelationspotential linearisieren:

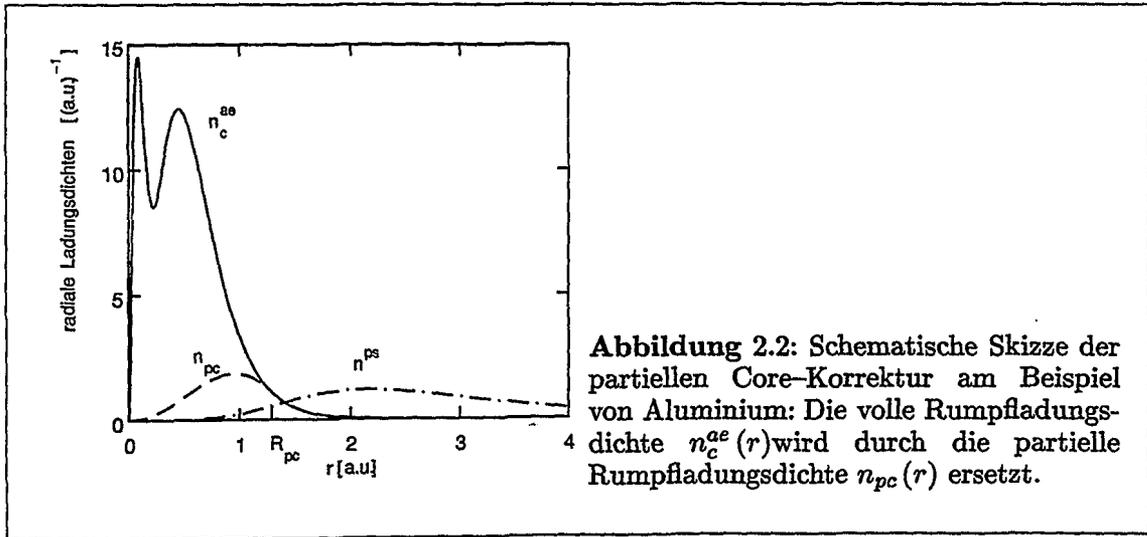
$$V_{xc} [n_c^{ae}(r) + n^{ps}(r)] \approx V_{xc} [n_c^{ae}(r)] + V_{xc} [n^{ps}(r)] \quad (2.27)$$

Dies ist Grundlage des linearen “*unscreenings*” (Gleichung (2.9)). Durch diese implizite Linearisierung wird die Bindung überschätzt, d.h. Bindungsenergien werden überschätzt und Bindungslängen werden unterschätzt [Eng92]. Nun kann man unmöglich die volle Rumpfladungsdichte berücksichtigen, da deren Beschreibung im Impulsraum zu viele ebene Wellen erfordern würde. Nach einem Vorschlag von Louie, Froyen und Cohen [LFC82] reicht eine partielle Rumpfladungsdichte $n_{pc}(r)$ bereits aus, den nichtlinearen Beitrag der vollen Rumpfladungsdichte zum Austauschkorre-

lationspotential zu berücksichtigen.

$$n_{pc}(r) = \begin{cases} A r \sin(Br), & \text{für } (n_c^{ae}(r) \geq \alpha_{pc} n^{ps}(r)) \text{ bzw. } (r < R_{pc}) \\ n_c^{ae}(r), & \text{für } (n_c^{ae}(r) < \alpha_{pc} n^{ps}(r)) \text{ bzw. } (r > R_{pc}) \end{cases} \quad (2.28)$$

Aus der Forderung nach stetiger Differenzierbarkeit von $n_{pc}(r)$ werden die Parameter A und B bestimmt. Mit einem einfachen Newtonverfahren läßt sich so immer eine knotenlose partielle Rumpfladungsdichte finden. Der Parameter α_{pc} bestimmt den Radius R_{pc} für den die Anpassung erfolgt. Ein größerer Wert für α_{pc} bedeutet einen kleineren Radius R_{pc} und damit eine schärfere partielle Rumpfladungsdichte $n_{pc}(r)$, was deren Fourierdarstellbarkeit erschwert. In der Praxis erweist sich ein Wert von 2–3 für α_{pc} als ausreichend.



Für die Konstruktion der Pseudopotentiale bedeutet dies nur, daß der "unscreening"-Schritt (Gleichung 2.9) modifiziert werden muß. Die partielle Rumpfladung ist im Austauschkorrelationspotential zu berücksichtigen:

$$V_{ion,l}^{ps}(r) = V_{scr,l}^{ps}(r) - (V_H [n^{ps}(r)] + V_{xc} [n_{pc}(r) + n^{ps}(r)]) \quad (2.29)$$

Die partielle Rumpfladung muß in einer selbstkonsistenten Anwendung des Pseudopotentials natürlich auch im Energiefunktional berücksichtigt werden, was auch Auswirkung auf die Berechnung der Hellman-Feynman-Kräfte hat. Dabei ist das Austauschkorrelationspotential zu ersetzen:

$$E_{xc} [n^{ps}(r)] \longrightarrow E_{xc} [n_{pc}(r) + n^{ps}(r)] \quad (2.30)$$

Kapitel 3

Explizite Darstellung der Gleichungen im Impulsraum

Das Funktional der totalen Energie $E_{tot}[n]$ setzt sich aus der totalen elektronischen Energie der DFT (Gleichung (1.11) bzw. Gleichung (1.43)) in der LDA und der Coulomb-Energie der Ionenrümpfe, der Ewald-Energie $E_{ion,ion}$ (3.146), zusammen. Die totale elektronische Energie wird innerhalb der elektronischen Selbstkonsistenzschleife über Gleichung (1.35) bzw. über Gleichung (1.45) berechnet, die Coulomb-Energie der Ionenrümpfe mittels Ewald-Summation (3.151). Die Pseudopotentiale der Ionenrümpfe (2.11) bzw. (2.17) bilden das externe Potential $V_{ext}(\vec{r})$. Dieses setzt sich aus dem lokalen Potential $V_{lok}(\vec{r})$ und einem nichtlokalen Anteil \mathcal{V}_{NL} zusammen.

$$E_{tot}[n] = T_0[n] + E_{lok}[n] + E_{NL}[n] + E_H[n] + E_{xc}[n] + E_{ion,ion} \quad (3.1)$$

Die totale Energie ist in unserem Fall die totale Energie einer Superzelle. Die Beiträge $E_{lok}[n]$, $E_H[n]$ und $E_{ion,ion}$ sind jeweils einzeln divergent, ihre Summe hat aber einen wohldefinierten endlichen Wert (siehe unten, Kap. 3.11).

Der Hamiltonian für das Elektronensystem setzt sich demzufolge aus den entsprechenden einzelnen Beiträgen zusammen:

$$\mathcal{H} = -\frac{\hbar^2}{2m}\Delta + V_{lok}(\vec{r}) + \mathcal{V}_{NL} + V_H(\vec{r}) + V_{xc}(\vec{r}) \quad (3.2)$$

Die Ewaldsumme $E_{ion,ion}$ liefert (neben der Kompensation der Divergenz) auch Beiträge zu den Hellman-Feynman Kräften auf die Ionen. Die Hellman-Feynman Kraft auf das μ 'te Ion ergibt sich durch Ableiten des Energiefunktionals nach den Atomkoordinaten \vec{r}_μ [IZC79].

$$\vec{F}^\mu = -\frac{\partial}{\partial \vec{r}_\mu} E_{tot}[n] = \vec{F}_{lok}^\mu + \vec{F}_{NL}^\mu + \vec{F}_{xc}^\mu + \vec{F}_{ion,ion}^\mu \quad (3.3)$$

Obwohl generell eine Formulierung der Gleichungen im Impulsraum benutzt wird, so werden doch einzelne Beiträge im Ortsraum ausgewertet. So werden z.B. bei der Berechnung des Matrixvektorprodukts $\mathcal{H}|\Psi\rangle$ zwar alle Beiträge im Impulsraum benötigt $\langle \vec{k} + \vec{G} | \mathcal{H} | \Psi \rangle_{\Omega}$, manche Beiträge werden aber in der Ortsraumdarstellung $\langle \vec{r} | \mathcal{H} | \Psi \rangle_{\Omega}$ berechnet und anschließend durch Fouriertransformation in die Impulsraumdarstellung gebracht.

Wie die einzelnen Beiträge des Hamiltonian – i.e. die kinetische Energie, die Hartree-Energie, Austausch und Korrelation, die Ewaldsumme sowie lokales und nichtlokales Pseudopotential – zu Gesamtenergie, Matrixvektorprodukt und Kraft im Computerprogramm ausgewertet werden, wird im nun folgenden Abschnitt erklärt. In einem Beziehungsdiagramm am Ende des Kapitels (Seite 50) wird skizziert, welche Größen im Orts- bzw. Impulsraum berechnet werden und wie diese über Fouriertransformation zusammenhängen.

3.1 Der Hamiltonian in der Impulsdarstellung

Wie im nächsten Abschnitt noch deutlich wird, besitzt die Hamiltonmatrix im Impulsraum folgende spezielle Struktur:

$$\begin{aligned} \mathcal{H}_{\vec{G}, \vec{G}'}(\vec{k}) &= \langle \vec{k} + \vec{G} | \mathcal{H} | \vec{k} + \vec{G}' \rangle_{\Omega} \\ &= + \frac{\hbar^2}{2m} |\vec{k} + \vec{G}|^2 \delta(\vec{G} - \vec{G}') + V_{Kovv}(\vec{G} - \vec{G}') \\ &\quad + \sum_{\mu} \sum_{lm} E_l^{\alpha(\mu)} T_{lm}^{\mu}(\vec{k}, \vec{G}) T_{lm}^{\mu*}(\vec{k}, \vec{G}') \end{aligned} \quad (3.4)$$

Der erste Summand, die kinetische Energie, ist diagonal im Impulsraum (siehe Kapitel 3.5). Der zweite Summand

$$V_{Kovv}(\vec{G} - \vec{G}') = V_H(\vec{G} - \vec{G}') + V_{lok}(\vec{G} - \vec{G}') + V_{xc}(\vec{G} - \vec{G}') \quad (3.5)$$

ist die Zusammenfassung der Potentialanteile, die im Ortsraum lokal sind, i.e. des Hartreepotentials V_H (siehe Kapitel 3.6), des lokalen Pseudopotentials V_{lok} (siehe Kapitel 3.7), und des Austauschkorrelationspotentials V_{xc} (siehe Kapitel 3.9). Die Multiplikation dieses Anteils mit einem Blochvektor $\Psi_{\vec{k}}$ hat im Impulsraum die Struktur einer dreidimensionalen Faltung (Konvolutionsgleichung). Der dritte Summand, das nichtlokale Pseudopotential (siehe Kapitel 3.8), ist ein dyadisches Produkt im Impulsraum.

Wir werden im folgenden Abschnitt sehen, wie für die Berechnung des Matrixvektor-Produkts $\mathcal{H}|\Psi\rangle$ diese spezielle Form der Hamiltonmatrix im Impulsraum

explizit ausgenutzt wird:

$$\begin{aligned}
 \langle \vec{k} + \vec{G} | \mathcal{H} | \Psi_{\vec{k}} \rangle &= \frac{\hbar^2}{2m} |\vec{k} + \vec{G}|^2 \langle \vec{k} + \vec{G} | \Psi_{\vec{k}} \rangle \\
 &+ \int_{\Omega} \langle \vec{k} + \vec{G} | \vec{r} \rangle \langle \vec{r} | V_{Konv} | \Psi_{\vec{k}} \rangle \\
 &+ \sum_{\mu} \sum_{lm} E_l^{\alpha(\mu)} T_{lm}^{\mu}(\vec{k}, \vec{G}) \sum_{\vec{G}'} \{ T_{lm}^{\mu *}(\vec{k}, \vec{G}') \langle \vec{k} + \vec{G}' | \Psi_{\vec{k}} \rangle \}
 \end{aligned} \tag{3.6}$$

Der erste Term und das dyadische Produkt werden direkt im Impulsraum ausgerechnet. Der zweite Term wird zuerst im Ortsraum ausgerechnet ($\langle \vec{r} | V_{Konv} | \Psi_{\vec{k}} \rangle$) und anschließend mittels Fouriertransformation in den Impulsraum transformiert (siehe auch Kapitel 3.3.2).

3.2 Superzellen und ebene Wellen

Eine Beschreibung der physikalischen Eigenschaften von Festkörpern wäre praktisch unmöglich, wenn nicht die Gleichgewichtsstruktur der meisten Festkörper ein regelmäßiges Kristallgitter wäre. Die kleinste Struktureinheit des Kristalls ist die *Elementarzelle*. Die Aneinanderreihung der Elementarzellen füllt den Raum lückenlos. Dies hat zur Folge, daß das Gitter invariant ist gegenüber Translationen um Strecken, die ein ganzzahliges Vielfaches der Gitterperiode sind. Dies gilt leider nur für einen idealen Kristall. Der unendlich ausgedehnte, ideale Kristall kann durch die Einführung periodischer Randbedingungen künstlich endlich gemacht werden.

In Wirklichkeit ist der unendlich ausgedehnte Kristall nur eine Idealvorstellung. Jede Probe ist zwangsläufig von endlicher Größe, womit die dreidimensionale Periodizität durch Oberflächen gestört ist. Weiterhin haben praktisch alle Festkörper strukturelle Fehler, wie z.B. Verunreinigungen und Defekte, welche die Periodizität stören. Solche Störungen sind die Ursache für viele physikalische Effekte, die man gerne beschreiben möchte. Als Beispiel sei der Transistor genannt, dessen Funktion erst durch die Dotierung ermöglicht wird.

Der gestörte Kristall läßt sich mit *Superzellen* beschreiben: Eine Superzelle enthält einen möglichst großen Bereich des Kristalls inklusive Störung. Durch periodisches Aneinanderreihen dieser Superzellen erhält man wieder einen periodischen Kristall. Die Superzelle ist dabei möglichst groß zu wählen, damit Wechselwirkungen der Defekte mit ihren "Bildern" in benachbarten Superzellen ausgeschlossen werden können. Die Vorteile einer periodischen Beschreibung des Festkörpers werden so künstlich erzwungen.

Die Superzelle wird durch drei linear unabhängige Basisvektoren \vec{a}_1 , \vec{a}_2 und \vec{a}_3 aufgespannt. Das Volumen Ω der Superzelle ergibt sich somit als Spat-Produkt:

$$\Omega = \vec{a}_1 \cdot (\vec{a}_2 \times \vec{a}_3) \quad (3.7)$$

Man hat also die Kohn-Sham-Gleichungen (1.19) für ein zellperiodisches Potential zu lösen:

$$V(\vec{r}) = (\mathcal{T}_{\vec{R}}V)(\vec{r}) = V(\mathcal{T}_{-\vec{R}}\vec{r}) = V(\vec{r} - \vec{R}) \quad (3.8)$$

Der Translationsoperator $\mathcal{T}_{\vec{R}}$ beschreibt eine Verschiebung um einen Gittervektor \vec{R} . Gittervektoren sind alle ganzzahligen Linearkombinationen der drei Basisvektoren \vec{a}_1 , \vec{a}_2 und \vec{a}_3 . Der Einfachheit halber wird eine einzige gemeinsame Schreibweise für den Translationsoperator $\mathcal{T}_{\vec{R}}$ benutzt, gleichgültig, ob er auf Operatoren, Potentiale, Wellenfunktionen oder Ortskoordinaten wirkt.

Die durch die Superzellenbeschreibung künstlich erzeugte Periodizität erlaubt nun die Anwendung des Blochschen Theorems [Blo28]. Dieses besagt, daß sich die Lösungen $\psi_{\vec{k}\nu}(\vec{r})$ einer Einteilchen-Schrödingergleichung zu einem gitterperiodischen Potential als Blochwellenfunktionen darstellen lassen. Eine Blochfunktion ist das Produkt einer gitterperiodischen Funktion $u(\vec{r})$ mit einer ebenen Welle $\frac{1}{\sqrt{\Omega}}e^{-i\vec{k}\vec{r}}$.

$$\begin{aligned} \Psi_{\vec{k}}(\vec{r}) &= e^{-i\vec{k}\vec{r}}u(\vec{r}) \\ \mathcal{T}_{\vec{R}}u(\vec{r}) &= u(\mathcal{T}_{-\vec{R}}\vec{r}) = u(\vec{r} - \vec{R}) = u(\vec{r}) \end{aligned} \quad (3.9)$$

Die Blochvektoren \vec{k} liegen in der ersten Brillouin-Zone, der Einheitszelle des reziproken Gitters, welches durch die reziproken Basisvektoren aufgespannt wird:

$$\vec{b}_1 = 2\pi \frac{(\vec{a}_2 \times \vec{a}_3)}{\vec{a}_1 \cdot (\vec{a}_2 \times \vec{a}_3)}, \quad (3.10)$$

$$\vec{b}_2 = 2\pi \frac{(\vec{a}_3 \times \vec{a}_1)}{\vec{a}_1 \cdot (\vec{a}_2 \times \vec{a}_3)}, \quad (3.11)$$

$$\vec{b}_3 = 2\pi \frac{(\vec{a}_1 \times \vec{a}_2)}{\vec{a}_1 \cdot (\vec{a}_2 \times \vec{a}_3)} \quad (3.12)$$

Die gitterperiodischen Funktionen $u(\vec{r})$ lassen sich in eine Fourier-Reihe von reziproken Gittervektoren entwickeln:

$$u(\vec{r}) = \frac{1}{\sqrt{\Omega}} \sum_{\vec{G}} u(\vec{G}) e^{-i\vec{G}\vec{r}} \quad (3.13)$$

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

Der Raum der Wellenfunktionen ist also der Raum der komplexen, quadratintegrierbaren Blochfunktionen, der durch die Wahl der Superzelle festgelegt wird. Das übliche Skalarprodukt in diesem Raum, d.h. die Integration über eine Superzelle vom Volumen Ω , wird in Dirac-Schreibweise abgekürzt:

$$\langle f | g \rangle_{\Omega} = \int_{\Omega} d^3 r f^*(\vec{r}) g(\vec{r}) \quad (3.14)$$

Für die Darstellung von Eigenfunktionen und Hamiltonian im Impulsraum wird die Basis aus ebenen Wellen mit folgender Normierung gewählt:

$$\begin{aligned} \langle \vec{r} | \vec{k} + \vec{G} \rangle &= \frac{1}{\sqrt{\Omega}} e^{-i(\vec{k} + \vec{G})\vec{r}} \\ \langle \vec{k}' + \vec{G}' | \vec{k} + \vec{G} \rangle_{\Omega} &= \frac{1}{\Omega} \int_{\Omega} d^3 r e^{i(\vec{k}' + \vec{G}')\vec{r}} e^{-i(\vec{k} + \vec{G})\vec{r}} = \delta_{\vec{k}' + \vec{G}', \vec{k} + \vec{G}} \end{aligned} \quad (3.15)$$

Die im Prinzip unendliche Fourierreihe über die Gittervektoren \vec{G} wird bei einer endlichen Wellenzahl G_{max} abgeschnitten. Üblicherweise wird diese in Energieeinheiten angegeben:

$$E_{max} = \frac{\hbar^2}{2m} G_{max}^2 \quad (3.16)$$

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

$$\int_{BZ} d^3 k F(\vec{k}) \rightarrow \sum_{\vec{k}} \omega_{\vec{k}} F(\vec{k}) \quad (3.17)$$

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 $\omega_{\vec{k}}$ nach dem Verfahren von Monkhorst und Pack [PM77] bestimmt werden.

3.3 Gittersummen und Konvolutionsgleichungen

3.3.1 Gittersummen

Mehrere Anteile des Hamiltonian setzen sich aus Beiträgen der einzelnen Atome μ zusammen, die auf den Positionen $\vec{R} + \vec{\tau}_{\mu}$ sitzen und der Atomsorte α (μ) angehören.

Die \vec{r}_μ sind die Koordinaten der Atome in der Superzelle, die zum Gittervektor \vec{R} gehört. Bei der Bildung von Matrixelementen wird dann über das Volumen Ω einer Superzelle integriert. Unabhängig von der speziellen Form der atomaren Beiträge kann die Summation über die Superzellen \vec{R} und Integration über eine einzige Superzelle Ω durch eine Integration über den ganzen Raum ersetzt werden, da Blochfunktionen $\Psi_{\vec{k}}(\vec{r})$ (Gleichung (3.9)) Eigenfunktionen zum Translationsoperator $\mathcal{T}_{\vec{R}}$ sind:

$$\mathcal{T}_{\vec{R}}\Psi_{\vec{k}}(\vec{r}) = \Psi_{\vec{k}}(\mathcal{T}_{-\vec{R}}\vec{r}) = \Psi_{\vec{k}}(\vec{r}-\vec{R}) = e^{i\vec{k}\vec{R}}\Psi_{\vec{k}}(\vec{r}) \quad (3.18)$$

Läßt sich nun ein gittersymmetrischer Potentialoperator \mathcal{V} als Gittersumme $\sum_{\vec{R}} \mathcal{T}_{\vec{R}} \tilde{\mathcal{V}}$ eines Potentialoperators $\tilde{\mathcal{V}}$ darstellen, so gilt unabhängig von der speziellen Struktur des Operators $\tilde{\mathcal{V}}$ für die Matrixelemente mit Blochfunktionen $\Psi_{\vec{k}}$ und $\Phi_{\vec{k}}$:

$$\begin{aligned} \langle \Phi_{\vec{k}} | \mathcal{V} | \Psi_{\vec{k}} \rangle_{\Omega} &= \sum_{\vec{R}} \langle \Phi_{\vec{k}} | \mathcal{T}_{\vec{R}} \tilde{\mathcal{V}} | \Psi_{\vec{k}} \rangle_{\Omega} = \sum_{\vec{R}} \langle \mathcal{T}_{-\vec{R}} \Phi_{\vec{k}} | \tilde{\mathcal{V}} | \mathcal{T}_{-\vec{R}} \Psi_{\vec{k}} \rangle_{\mathcal{T}_{\vec{R}}\Omega} \\ \langle \Phi_{\vec{k}} | \mathcal{V} | \Psi_{\vec{k}} \rangle_{\infty} &= \langle \Phi_{\vec{k}} | \tilde{\mathcal{V}} | \Psi_{\vec{k}} \rangle_{\sum_{\vec{R}} \mathcal{T}_{\vec{R}}\Omega} = \sum_{\vec{R}} \langle \Phi_{\vec{k}} | \tilde{\mathcal{V}} | \Psi_{\vec{k}} \rangle_{\mathcal{T}_{\vec{R}}\Omega} \end{aligned} \quad (3.19)$$

Für die Matrixelemente des Potentialoperators \mathcal{V} mit ebenen Wellen $\Psi_{\vec{k}} = |\vec{k} + \vec{G}\rangle$ und $\Phi_{\vec{k}} = |\vec{k} + \vec{G}'\rangle$ ergibt sich damit:

$$\mathcal{V} = \sum_{\vec{R}} \mathcal{T}_{\vec{R}} \tilde{\mathcal{V}} \quad (3.20)$$

$$\langle \vec{k} + \vec{G} | \mathcal{V} | \vec{k} + \vec{G}' \rangle_{\Omega} = \langle \vec{k} + \vec{G} | \tilde{\mathcal{V}} | \vec{k} + \vec{G}' \rangle_{\infty} \quad (3.21)$$

Durch Einsetzen der Blochfunktionen zum von $\Psi_{\vec{0}} = |\vec{G}\rangle$ und $\Phi_{\vec{0}} \equiv 1$ ergibt sich für die Fouriertransformierte des Potentialoperators \mathcal{V} :

$$\langle \vec{G} | \mathcal{V} \rangle_{\Omega} = \langle \vec{G} | \tilde{\mathcal{V}} \rangle_{\infty} \quad (3.22)$$

Für beliebige Funktionen $f(\vec{r})$ ergibt sich folgende, von der speziellen Form des Integranden $f(\vec{r})$ unabhängige Gleichung:

$$\sum_{\vec{R}} \int_{\Omega} d^3r f(\vec{r}-\vec{R}) = \sum_{\vec{R}} \int_{\Omega} d^3r \mathcal{T}_{\vec{R}} f(\vec{r}) = \sum_{\vec{R}} \int_{\mathcal{T}_{\vec{R}}\Omega} d^3r f(\vec{r}) = \int_{\infty} d^3r f(\vec{r}) \quad (3.23)$$

3.3.2 Konvolutionsgleichungen

Obwohl generell eine Formulierung der Gleichungen im Impulsraum benutzt wird, so werden doch einzelne Beiträge im Ortsraum ausgerechnet, und zwar immer dann,

wenn im Impulsraum Konvolutionsgleichungen auszuwerten wären. Sind $f(\vec{r})$ und $g(\vec{r})$ Blochfunktionen, so nennt man Gleichungen folgenden Typs "Konvolutionsgleichungen":

$$(g \circ f)(\vec{G}) := \sum_{\vec{G}'} g(\vec{G}') f(\vec{G} - \vec{G}') \quad (3.24)$$

$$(g^* \circ f)(\vec{G}) := \sum_{\vec{G}'} g^*(\vec{G}') f(\vec{G} + \vec{G}') \quad (3.25)$$

Der Kringel "o" symbolisiert hier die Konvolutionssummation im Impulsraum. Konvolutionsgleichungen im Impulsraum werden nach Fouriertransformation in den Ortsraum diagonal:

$$(g \circ f)(\vec{r}) = g(\vec{r}) f(\vec{r}) \quad (3.26)$$

$$(g^* \circ f)(\vec{r}) = g^*(\vec{r}) f(\vec{r}) \quad (3.27)$$

Für eine direkte Auswertung eines Konvolutionsintegrals im Impulsraum benötigt man $N_{\vec{G}}^2$ Operationen, wobei $N_{\vec{G}}$ die Anzahl der ebenen Wellen \vec{G} ist. Für die Auswertung der Multiplikation im Ortsraum benötigt man jedoch lediglich $N_{\vec{G}}$ Operationen. Zwischen den beiden Darstellungen vermittelt eine FFT, welche zusätzlich $N_{\vec{G}} \log(N_{\vec{G}})$ Operationen benötigt. Damit ist die Skalierung des Problems mit $N_{\vec{G}}^2$, die für große Systeme untragbar ist, auf eine Skalierung mit $N_{\vec{G}} \log(N_{\vec{G}})$ reduziert. Konvolutionsgleichungen im Impulsraum werden also wie folgt ausgewertet:

$$\begin{array}{c} g(\vec{G}) \\ f(\vec{G}) \end{array} \xrightarrow{\text{FFT}} \begin{array}{c} g(\vec{r}) \\ f(\vec{r}) \end{array} \left\{ \begin{array}{l} \xrightarrow{(3.26)} (g \circ f)(\vec{r}) \\ \xrightarrow{(3.27)} (g^* \circ f)(\vec{r}) \end{array} \right. \xrightarrow{\text{FFT}} \begin{array}{c} (g \circ f)(\vec{G}) \\ (g^* \circ f)(\vec{G}) \end{array} \quad (3.28)$$

Analog werden werden Konvolutionsintegrale im Ortsraum, die hier durch "•" symbolisiert werden, im Impulsraum diagonal:

$$(g \bullet f)(\vec{r}) := \int_{\Omega} d^3r' g(\vec{r}') f(\vec{r} - \vec{r}') \quad (3.29)$$

$$(g^* \bullet f)(\vec{r}) := \int_{\Omega} d^3r' g^*(\vec{r}') f(\vec{r} + \vec{r}') \quad (3.30)$$

$$(g \bullet f)(\vec{G}) = g(\vec{G}) f(\vec{G}) \quad (3.31)$$

$$(g^* \bullet f)(\vec{G}) = g^*(\vec{G}) f(\vec{G}) \quad (3.32)$$

3.4 Wellenfunktionen und Ladungsdichte

Wie bereits erwähnt, gehören die Eigenfunktionen $\Psi_{\vec{k}\nu}(\vec{r})$ dem Raum der Blochfunktionen an, der durch die Wahl der Superzelle festgelegt wird. Für die Darstellung

im Impulsraum wird folgende Konvention gewählt:

$$\begin{aligned}\Psi_{\vec{k}\nu}(\vec{r}) &= \sum_{\vec{G}} \langle \vec{r} | \vec{k} + \vec{G} \rangle \langle \vec{k} + \vec{G} | \Psi_{\vec{k}\nu} \rangle_{\Omega} = \frac{1}{\sqrt{\Omega}} \sum_{\vec{G}} c_{\vec{k}\nu}(\vec{G}) e^{-i(\vec{k}+\vec{G})\vec{r}} \\ c_{\vec{k}\nu}(\vec{G}) &= \langle \vec{k} + \vec{G} | \Psi_{\vec{k}\nu} \rangle_{\Omega} = \frac{1}{\sqrt{\Omega}} \int_{\Omega} d^3r e^{i(\vec{k}+\vec{G})\vec{r}} \Psi_{\vec{k}\nu}(\vec{r})\end{aligned}\quad (3.33)$$

Die Eigenfunktionen $\Psi_{\vec{k}\nu}(\vec{r})$ sind orthogonal und werden im Zellvolumen Ω auf 1 normiert:

$$\langle \Psi_{\vec{k}'\nu'} | \Psi_{\vec{k}\nu} \rangle_{\Omega} = \int_{\Omega} d^3r \Psi_{\vec{k}'\nu'}^*(\vec{r}) \Psi_{\vec{k}\nu}(\vec{r}) = \sum_{\vec{G}} c_{\vec{k}'\nu'}^*(\vec{G}) c_{\vec{k}\nu}(\vec{G}) = \delta_{\vec{k}'\vec{k}} \delta_{\nu'\nu} \quad (3.34)$$

Aus diesen Eigenfunktionen wird die Ladungsdichte zusammengesetzt.

$$n(\vec{r}) = \sum_{\vec{k}\nu} \omega_{\vec{k}} f_{\vec{k}\nu} \Psi_{\vec{k}\nu}^*(\vec{r}) \Psi_{\vec{k}\nu}(\vec{r}) \quad (3.35)$$

$$n(\vec{G}) = \sum_{\vec{G}'} \sum_{\vec{k}\nu} \omega_{\vec{k}} f_{\vec{k}\nu} c_{\vec{k}\nu}^*(\vec{G}') c_{\vec{k}\nu}(\vec{G} + \vec{G}') \quad (3.36)$$

Die $\omega_{\vec{k}}$ sind die Gewichte, die für die \vec{k} -Integration (3.17) notwendig sind. Dabei sind die $f_{\vec{k}\nu}$ die Besetzungszahlen für den Zustand $\Psi_{\vec{k}\nu}$. Die Besetzung erfolgt durch eine Fermiverteilung:

$$f(\varepsilon_{\vec{k}\nu}, T) = \frac{1}{e^{\frac{\varepsilon_{\vec{k}\nu} - \varepsilon_F}{kT}} + 1} \quad (3.37)$$

Bei Systemen mit Bandgap wird in der Regel $T = 0$ gewählt. Damit reduziert sich die Fermiverteilung zu einer Stufenfunktion. Bei Halbleitern sind somit alle Zustände bis zur Fermienergie ε_F (die Valenzbänder) besetzt, alle Zustände oberhalb von ε_F (die Leitungsbänder) bleiben unbesetzt.

Damit ergibt sich für die Fourierdarstellung der Ladungsdichte folgende Normierung:

$$n(\vec{r}) = \frac{1}{\Omega} \sum_{\vec{G}} n(\vec{G}) e^{-i\vec{G}\vec{r}} \quad (3.38)$$

$$n(\vec{G}) = \sqrt{\Omega} \langle \vec{G} | n \rangle_{\Omega} = \int_{\Omega} d^3r n(\vec{r}) e^{i\vec{G}\vec{r}} \quad (3.39)$$

$$N_v = \int_{\Omega} d^3r n(\vec{r}) = n(\vec{G}) \Big|_{\vec{G}=\vec{0}} \quad (3.40)$$

Die Ladungsdichte $n(\vec{G})$ wird nicht über die Konvolutionsgleichung (3.36) aus den Wellenfunktionen $c_{\vec{k}\nu}(\vec{G})$ berechnet. Dazu würde man $N_{\vec{G}}^2$ Operationen pro Wellenfunktion $\Psi_{\vec{k}\nu}$ benötigen. Deshalb werden die Wellenfunktionen mittels FFT in den

Ortsraum transformiert. Dort wird mittels Gleichung (3.35) die Ladungsdichte $n(\vec{r})$ berechnet. Diese wird dann mittels FFT in den Impulsraum zurücktransformiert.

$$c_{\vec{k}\nu}(\vec{G}) \xrightarrow{\text{FFT}} \Psi_{\vec{k}\nu}(\vec{r}) \xrightarrow{(3.35)} n(\vec{r}) \xrightarrow{\text{FFT}} n(\vec{G}) \quad (3.41)$$

3.5 Die kinetische Energie

$T_0[n(\vec{r})]$ ist das Funktional der kinetischen Energie wechselwirkungsfreier Elektronen. Es ist im Impulsraum diagonal. Alle Beiträge werden deshalb im Impulsraum ausgewertet. Einen Beitrag zur Kraft gibt es nicht, da die kinetische Energie nicht explizit von den Atomkoordinaten \vec{r}_μ abhängt.

$$T_0[n] = \langle \Psi | T_0 | \Psi \rangle_\Omega = \sum_{\vec{k}\nu} \omega_{\vec{k}} f_{\vec{k}\nu} \langle \Psi_{\vec{k}\nu} | -\frac{\hbar^2}{2m} \Delta | \Psi_{\vec{k}\nu} \rangle_\Omega \quad (3.42)$$

$$= \sum_{\vec{k}\nu} \sum_{\vec{G}} \omega_{\vec{k}} f_{\vec{k}\nu} (\vec{k} + \vec{G})^2 c_{\vec{k}\nu}^*(\vec{G}) c_{\vec{k}\nu}(\vec{G}) \quad (3.43)$$

$$\langle \vec{r} | T_0 | \Psi_{\vec{k}\nu} \rangle = \frac{1}{\sqrt{\Omega}} \sum_{\vec{G}} (\vec{k} + \vec{G})^2 c_{\vec{k}\nu}(\vec{G}) e^{-i(\vec{k} + \vec{G})\vec{r}} \quad (3.44)$$

$$\langle \vec{k} + \vec{G} | T_0 | \Psi_{\vec{k}\nu} \rangle_\Omega = (\vec{k} + \vec{G})^2 c_{\vec{k}\nu}(\vec{G}) \quad (3.45)$$

$$-\frac{d}{d\vec{r}_\mu} \langle \Psi | T_0 | \Psi \rangle_\Omega = \vec{0} \quad (3.46)$$

3.6 Der Hartree-Term

Die Hartree-Energie ist der Beitrag der Coulombwechselwirkung der Elektronen zum Energiefunktional:

$$E_H[n] = \frac{1}{2} e^2 \int_\Omega d^3r \int_\Omega d^3r' \frac{n(\vec{r})n(\vec{r}')}{|\vec{r} - \vec{r}'|} \quad (3.47)$$

$$= \frac{1}{2} \int_\Omega d^3r V_H[n(\vec{r})] n(\vec{r}) \quad (3.48)$$

$$V_H[n(\vec{r})] = \frac{\delta}{\delta n(\vec{r})} E_H[n(\vec{r})] = e^2 \int_\Omega d^3r' \frac{n(\vec{r}')}{|\vec{r} - \vec{r}'|} \quad (3.49)$$

Für die Fourierdarstellung des Hartree-Potentials $V_H(\vec{r})$ wird folgende Normierung gewählt:

$$V_H(\vec{r}) = \sum_{\vec{G}} V_H(\vec{G}) e^{-i\vec{G}\vec{r}} \quad (3.50)$$

$$V_H(\vec{G}) = \frac{1}{\sqrt{\Omega}} \langle \vec{G} | V_H \rangle_\Omega = \frac{1}{\Omega} \int_\Omega d^3r V_H(\vec{r}) e^{i\vec{G}\vec{r}} \quad (3.51)$$

Das Hartree-Potential $V_H(\vec{G})$ wird über die Poisson-Gleichung (A.48) im Impulsraum aus der Ladungsdichte $n(\vec{G})$ berechnet:

$$\Delta V_H(\vec{r}) = -4\pi e^2 n(\vec{r}) \quad (3.52)$$

$$V_H(\vec{G}) = \frac{1}{\Omega} \frac{4\pi e^2}{G^2} n(\vec{G}) \quad (3.53)$$

Die Hartree-Energie kann direkt im Impulsraum ausgerechnet werden:

$$\langle \Psi | V_H | \Psi \rangle_\Omega = \frac{1}{2} \sum_{\vec{G}} n(\vec{G}) V_H(\vec{G}) \quad (3.54)$$

Für den Beitrag zum Matrixvektorprodukt ergibt sich:

$$\langle \vec{r} | V_H | \Psi_{\vec{k}\nu} \rangle = V_H(\vec{r}) \Psi_{\vec{k}\nu}(\vec{r}) \quad (3.55)$$

$$= \frac{1}{\sqrt{\Omega}} \sum_{\vec{G}} \left(\sum_{\vec{G}'} V_H(\vec{G}') c_{\vec{k}\nu}(\vec{G} - \vec{G}') \right) e^{-i(\vec{k} + \vec{G})\vec{r}} \quad (3.56)$$

$$\langle \vec{k} + \vec{G} | V_H | \Psi_{\vec{k}\nu} \rangle = \sum_{\vec{G}'} V_H(\vec{G}') c_{\vec{k}\nu}(\vec{G} - \vec{G}') \quad (3.57)$$

Die Konvolutionsgleichung (3.57) für das Matrixvektorprodukt $\langle \vec{r} | V_H | \Psi_{\vec{k}\nu} \rangle$ wird im Ortsraum ausgerechnet:

$$\begin{pmatrix} V_H(\vec{G}) \\ c_{\vec{k}\nu}(\vec{G}) \end{pmatrix} \xrightarrow{\text{FFT}} \begin{pmatrix} V_H(\vec{r}) \\ \Psi_{\vec{k}\nu}(\vec{r}) \end{pmatrix} \xrightarrow{(3.55)} (V_H \Psi_{\vec{k}\nu})(\vec{r}) \xrightarrow{\text{FFT}} (V_H \Psi_{\vec{k}\nu})(\vec{G}) \quad (3.58)$$

Einen Beitrag zur Kraft gibt es nicht, da die Hartree-Energie nicht von den Atomkoordinaten \vec{r}_μ abhängt:

$$-\frac{\delta}{\delta \vec{r}_\mu} \langle \Psi | V_H | \Psi \rangle_\Omega = \vec{0} \quad (3.59)$$

Hartree-Energie und Hartree-Potential sind divergent, i.e. im Impulsraum ist die $\vec{G} = \vec{0}$ -Komponente divergent:

$$\frac{1}{2} n(\vec{G}) V_H(\vec{G}) \Big|_{\vec{G}=\vec{0}} = \frac{1}{2} \frac{N_v^2}{\Omega} \frac{4\pi e^2}{G^2} \quad (3.60)$$

$$V_H(\vec{G}) \Big|_{\vec{G}=\vec{0}} = \frac{1}{\Omega} \frac{4\pi e^2}{G^2} N_v \quad (3.61)$$

Die $(\vec{G} = \vec{0})$ -Komponente der Hartree-Energie ist eine Konstante, d.h. sie hängt nicht von den Wellenfunktionen und den Atomkoordinaten ab. Sie kann also hier explizit zu 0 gesetzt werden. An dieser Stelle wird somit auch die divergente $(\vec{G} = \vec{0})$ -Komponente des Hartree-Potentials explizit zu Null gesetzt:

$$V_H(\vec{G}) \Big|_{\vec{G}=\vec{0}} \stackrel{!}{\rightarrow} 0 \quad (3.62)$$

Möglich ist dies nur, weil der divergente Beitrag der $(\vec{G} = \vec{0})$ -Komponente durch die entsprechenden divergenten Anteile des lokalen Pseudopotentials und der Ewaldsumme kompensiert wird. Dies wird in Kapitel 3.11 diskutiert.

3.7 Der lokale Anteil des Hamiltonian

Das lokale Potential $V_{lok}(\vec{r})$ setzt sich aus den radialsymmetrischen Beiträgen $V_{lok}^{\alpha(\mu)}$ der einzelnen Atome μ zusammen, die auf den Positionen $\vec{R} + \vec{\tau}_\mu$ sitzen und der Atomsorte $\alpha(\mu)$ angehören. Die $\vec{\tau}_\mu$ sind die Koordinaten der Atome in der Superzelle mit Gittervektor \vec{R} . Gemäß Kapitel 3.3.1 kann unabhängig von der speziellen Form des lokalen Potentials die Summation über die Superzellen \vec{R} und das Integral über eine einzige Superzelle Ω durch eine Integration über den ganzen Raum ersetzt werden. Dabei ist folgender Beitrag zum Energiefunktional auszuwerten:

$$\langle \Psi | V_{lok}(\vec{r}) | \Psi \rangle_\Omega = \sum_{\vec{R}} \sum_{\mu} \sum_{\vec{k}\nu} \omega_{\vec{k}} f_{\vec{k}\nu} \langle \Psi_{\vec{k}\nu} | V_{lok}^{\alpha(\mu)}(|\vec{r} - (\vec{R} + \vec{\tau}_\mu)|) | \Psi_{\vec{k}\nu} \rangle_\Omega \quad (3.63)$$

$$\begin{aligned} & \parallel \\ \langle \Psi | \tilde{V}_{lok}(\vec{r}) | \Psi \rangle_\infty &= \sum_{\mu} \sum_{\vec{k}\nu} \omega_{\vec{k}} f_{\vec{k}\nu} \langle \Psi_{\vec{k}\nu} | V_{lok}^{\alpha(\mu)}(|\vec{r} - \vec{\tau}_\mu|) | \Psi_{\vec{k}\nu} \rangle_\infty \end{aligned} \quad (3.64)$$

$$V_{lok}(\vec{r}) = \sum_{\vec{R}, \mu} V_{lok}^{\alpha(\mu)}(|\vec{r} - (\vec{R} + \vec{\tau}_\mu)|) \quad (3.65)$$

$$\begin{aligned} & \downarrow \\ \tilde{V}_{lok}(\vec{r}) &= \sum_{\mu} V_{lok}^{\alpha(\mu)}(|\vec{r} - \vec{\tau}_\mu|) \end{aligned} \quad (3.66)$$

Für den lokalen Anteil der Pseudopotentiale wird in Orts- und Impulsraum für die Fouriertransformation folgende Darstellung gewählt:

$$V_{lok}(\vec{r}) = \sum_{\vec{G}} V_{lok}(\vec{G}) e^{-i\vec{G}\vec{r}} = \sum_{\vec{G}, \mu} S^\mu(\vec{G}) V_{lok}^{\alpha(\mu)}(\vec{G}) e^{-i\vec{G}\vec{r}} \quad (3.67)$$

$$\begin{aligned} V_{lok}(\vec{G}) &= \frac{1}{\sqrt{\Omega}} \langle \vec{G} | V_{lok} \rangle_\Omega = \frac{1}{\Omega} \int_\infty d^3r \tilde{V}_{lok}(\vec{r}) e^{i\vec{G}\vec{r}} \\ &= \sum_{\mu} \frac{1}{\sqrt{\Omega}} \langle \vec{G} | V_{lok}^{\alpha(\mu)}(|\vec{r} - \vec{\tau}_\mu|) \rangle_\infty = \sum_{\mu} S^\mu(\vec{G}) V_{lok}^{\alpha(\mu)}(\vec{G}) \end{aligned} \quad (3.68)$$

$$V_{lok}^{\alpha(\mu)}(\vec{G}) = \frac{1}{\sqrt{\Omega}} \langle \vec{G} | V_{lok}^{\alpha(\mu)} \rangle_\infty = \frac{1}{\Omega} \int_\infty d^3r V_{lok}^{\alpha(\mu)}(r) e^{i\vec{G}\vec{r}} \quad (3.69)$$

$$S^\mu(\vec{G}) = e^{i\vec{G}\vec{\tau}_\mu} \quad (3.70)$$

Das lokale Potential im Ortsraum wird demnach nicht mit Gleichung (3.65) im Ortsraum ausgerechnet, sondern über Gleichungen (3.68) und (3.69) im Impulsraum

und durch Rücktransformation in den Ortsraum.

$$V_{lok}^\alpha(\vec{r}) \xrightarrow{(3.69)} V_{lok}^\alpha(\vec{G}) \xrightarrow{(3.68)} V_{lok}(\vec{G}) \xrightarrow{\text{FFT}} V_{lok}(\vec{r}) \quad (3.71)$$

$V_{lok}(\vec{G})$ faktorisiert in zwei verschiedene Anteile: Zum einen sind das die Strukturfaktoren $S^\mu(\vec{G})$, die explizit von den Atompositionen \vec{r}_μ abhängen, zum andren die gitterunabhängigen $V_{lok}^\alpha(\vec{G})$. Da $V_{lok}^\alpha(r)$ radialsymmetrisch ist, kann $V_{lok}^\alpha(\vec{G})$ durch ein einfaches radiales Integral ausgerechnet werden und hängt nur von $G = |\vec{G}|$ ab:

$$V_{lok}^\alpha(\vec{G}) = \frac{4\pi}{\Omega} \begin{cases} \int_0^\infty r^2 dr V_{lok}^\alpha(r) j_0(Gr) & \text{für } (\vec{G} \neq \vec{0}) \\ \int_0^\infty r^2 dr V_{lok}^\alpha(r) & \text{für } (\vec{G} = \vec{0}) \end{cases} \quad (3.72)$$

$$j_0(Gr) = \frac{\sin(Gr)}{(Gr)} \quad (3.73)$$

Die ebene Welle $e^{i\vec{G}\vec{r}}$ wurde hier nach sphärischen Besselfunktionen $j_l(Gr)$ entwickelt (siehe Anhang A). Die Nullkomponente von $V_{lok}^\alpha(\vec{G})$ ist das mittlere Zellpotential von $V_{lok}^\alpha(\vec{r})$.

Der lokale Beitrag zum Energiefunktional $\langle \Psi | \mathcal{H} | \Psi \rangle_\Omega$ bestimmt sich dann wie folgt:

$$\begin{aligned} \langle \Psi | \mathcal{H}_{lok} | \Psi \rangle_\Omega &= \sum_{\vec{k}\nu\mu} \omega_{\vec{k}} f_{\vec{k}\nu} \langle \Psi_{\vec{k}\nu} | V_{lok}^{\alpha(\mu)} | \Psi_{\vec{k}\nu} \rangle_\infty \quad (3.74) \\ &= \sum_{\vec{G}''\vec{G}'} \sum_{\vec{k}\nu\mu} \omega_{\vec{k}} f_{\vec{k}\nu} c_{\vec{k}\nu}^*(\vec{G}'') c_{\vec{k}\nu}(\vec{G}') \frac{1}{\Omega} \int_0^\infty d^3r e^{i(\vec{G}''-\vec{G}')\vec{r}} V_{lok}^{\alpha(\mu)}(\vec{r}-\vec{r}_\mu) \\ &= \sum_{\vec{G}''\vec{G}'} \sum_{\vec{k}\nu\mu} \omega_{\vec{k}} f_{\vec{k}\nu} c_{\vec{k}\nu}^*(\vec{G}'') c_{\vec{k}\nu}(\vec{G}') S^\mu(\vec{G}''-\vec{G}') V_{lok}^{\alpha(\mu)}(\vec{G}''-\vec{G}') \\ &= \sum_{\vec{G}\vec{G}'} \sum_{\vec{k}\nu\mu} \omega_{\vec{k}} f_{\vec{k}\nu} c_{\vec{k}\nu}^*(\vec{G}+\vec{G}') c_{\vec{k}\nu}(\vec{G}') S^\mu(\vec{G}) V_{lok}^{\alpha(\mu)}(\vec{G}) \\ &= \sum_{\vec{G}} \sum_{\mu} n^*(\vec{G}) S^\mu(\vec{G}) V_{lok}^{\alpha(\mu)}(\vec{G}) \quad (3.75) \end{aligned}$$

Die lokale Energie kann also mit Gleichung (3.75) direkt im Impulsraum ausgerechnet werden.

Für den lokalen Beitrag zum Matrixvektorprodukt $\mathcal{H} | \Psi_{\vec{k}\nu} \rangle$ ergibt sich:

$$\begin{aligned} \langle \vec{r} | V_{lok} | \Psi_{\vec{k}\nu} \rangle_\Omega &= V_{lok}(\vec{r}) \Psi_{\vec{k}\nu}(\vec{r}) \quad (3.76) \\ &= \sum_{\vec{G}''\vec{G}'} \sum_{\mu} \frac{1}{\sqrt{\Omega}} S^\mu(\vec{G}') V_{lok}^{\alpha(\mu)}(\vec{G}') c_{\vec{k}\nu}(\vec{G}'') e^{-i(\vec{k}+\vec{G}'+\vec{G}'')\vec{r}} \\ &= \sum_{\vec{G}\vec{G}'} \sum_{\mu} \frac{1}{\sqrt{\Omega}} S^\mu(\vec{G}') V_{lok}^{\alpha(\mu)}(\vec{G}') c_{\vec{k}\nu}(\vec{G}-\vec{G}') e^{-i(\vec{k}+\vec{G})\vec{r}} \end{aligned}$$

$$\begin{aligned}
\langle \vec{k} + \vec{G} | V_{lok} | \Psi_{\vec{k}\nu} \rangle_{\Omega} &= \sum_{\vec{G}'} \sum_{\mu} S^{\mu}(\vec{G}') V_{lok}^{\alpha(\mu)}(\vec{G}') c_{\vec{k}\nu}(\vec{G} - \vec{G}') \\
&= \sum_{\vec{G}'} V_{lok}(\vec{G}') c_{\vec{k}\nu}(\vec{G} - \vec{G}') \quad (3.77)
\end{aligned}$$

Die Konvolutionssumme (3.77) für das Matrixvektorprodukt $\mathcal{H}_{lok} |\Psi_{\vec{k}\nu}\rangle$ wird im Ortsraum (3.76) ausgerechnet:

$$\begin{pmatrix} V_{lok}(\vec{G}) \\ c_{\vec{k}\nu}(\vec{G}) \end{pmatrix} \xrightarrow{\text{FFT}} \begin{pmatrix} V_{lok}(\vec{r}) \\ \Psi_{\vec{k}\nu}(\vec{r}) \end{pmatrix} \xrightarrow{(3.76)} (V_{lok} \Psi_{\vec{k}\nu})(\vec{r}) \xrightarrow{\text{FFT}} (V_{lok} \Psi_{\vec{k}\nu})(\vec{G}) \quad (3.78)$$

Der Beitrag der lokalen Kraft zur Hellmann–Feynman–Kraft ergibt sich durch Ableiten der Gleichung (3.75) nach den Atomkoordinaten \vec{r}_{μ} . Dabei hängen nur die Strukturfaktoren $S^{\mu}(\vec{G})$ von den \vec{r}_{μ} ab:

$$\frac{\partial}{\partial \vec{r}_{\mu}} S^{\mu}(\vec{G}) = i \vec{G} S^{\mu}(\vec{G}) \quad (3.79)$$

Für die Kraft auf das Atom mit den Koordinaten \vec{r}_{μ} erhält man so:

$$\begin{aligned}
\vec{F}_{lok}^{\mu} &= -\frac{\partial}{\partial \vec{r}_{\mu}} \langle \Psi | \mathcal{H}_{lok} | \Psi \rangle_{\Omega} \quad (3.80) \\
&= -\frac{\partial}{\partial \vec{r}_{\mu}} \left(\sum_{\vec{G}} \sum_{\mu} n^*(\vec{G}) S^{\mu}(\vec{G}) V_{lok}^{\alpha(\mu)}(\vec{G}) \right) \\
&= \sum_{\vec{G}} -i \vec{G} n^*(\vec{G}) S^{\mu}(\vec{G}) V_{lok}^{\alpha(\mu)}(\vec{G}) \quad (3.81)
\end{aligned}$$

Die lokale Kraft kann also mit Gleichung (3.81) ebenfalls wie die lokale Energie direkt im Impulsraum ausgerechnet werden.

3.7.1 Berechnung der $V_{lok}^{\alpha}(\vec{G})$

Die lokalen Potentiale $V_{lok}^{\alpha}(r)$ haben einen langreichweitigen Limes $\sim -\frac{Z_{\alpha}^2 e^2}{r}$. Daher können die radialen Integrale aus Gleichung (3.72) nicht einfach numerisch ausgeführt werden und die Nullkomponente $V_{lok}^{\alpha}(\vec{G} \rightarrow \vec{0})$ divergiert sogar. Deshalb wird ein gemeinsamer, langreichweitiger Teil $V_{core}^{\alpha}(r)$ abgespalten und analytisch behandelt:

$$V_{lok}^{\alpha}(r) = V_{core}^{\alpha}(r) + \delta V_{lok}^{\alpha}(r) \quad (3.82)$$

Nur für die kurzreichweitigen $\delta V_{lok}^\alpha(r)$ wird die Gleichung (3.72) numerisch mittels einer einfachen Simpson-Integration integriert:

$$\delta V_{lok}^\alpha(\vec{G}) = \frac{4\pi}{\Omega} \begin{cases} \int_{\infty} r^2 dr \delta V_{lok}^\alpha(r) j_0(Gr) & \text{für } (\vec{G} \neq \vec{0}) \\ \int_{\infty} r^2 dr \delta V_{lok}^\alpha(r) & \text{für } (\vec{G} = \vec{0}) \end{cases} \quad (3.83)$$

Für die langreichweitigen Anteile benutzen wir ausschließlich die in [BHS82] tabellierten Core-Potentiale $V_{core}^\alpha(r)$. Sie lassen sich darstellen als die elektrostatischen Potentiale einer positiven Ladungsdichte $\rho_{core}^\alpha(r)$, die sich aus zwei Gaußladungsdichten zusammensetzt, und die positive Gesamtladung N_v^α enthält. Damit lassen sich die langreichweitigen $V_{core}^\alpha(r)$ mit Hilfe der Poissongleichung (A.48) analytisch behandeln:

$$\rho_{core}^\alpha(r) = N_v^\alpha \sum_{i=1}^2 c_i^\alpha \left(\frac{\alpha_i^\alpha}{\pi}\right)^{\frac{3}{2}} e^{-\alpha_i^\alpha r^2} \quad (3.84)$$

$$\rho_{core}^\alpha(\vec{G}) = N_v^\alpha \sum_{i=1}^2 c_i^\alpha e^{-\frac{G^2}{4\alpha_i^\alpha}} \quad (3.85)$$

$$\nabla^2 V_{core}^\alpha(\vec{r}) = 4\pi e^2 \rho_{core}^\alpha(\vec{r}) \quad (3.86)$$

$$V_{core}^\alpha(\vec{G}) = -\frac{4\pi e^2 \rho_{core}^\alpha(\vec{G})}{\Omega G^2} \quad (3.87)$$

$$V_{core}^\alpha(r) = -\frac{e^2 N_v^\alpha}{r} \sum_{i=1}^2 c_i^\alpha \operatorname{erf}(\sqrt{\alpha_i^\alpha} r) \quad (3.88)$$

$$V_{core}^\alpha(\vec{G}) = -\frac{4\pi e^2 N_v^\alpha}{\Omega G^2} \sum_{i=1}^2 c_i^\alpha e^{-\frac{G^2}{4\alpha_i^\alpha}} \quad (3.89)$$

Die $(\vec{G} = \vec{0})$ -Komponente $\lim_{\vec{G} \rightarrow \vec{0}} V_{lok}(\vec{G})$ muß gesondert behandelt werden, da die Grenzwerte $\lim_{\vec{G} \rightarrow \vec{0}} V_{core}^\alpha(\vec{G})$ divergieren:

$$\lim_{\vec{G} \rightarrow \vec{0}} V_{core}^\alpha(\vec{G}) = -\frac{4\pi e^2}{\Omega} N_v^\alpha \left(\frac{1}{G^2} - \sum_{i=1}^2 \frac{c_i^\alpha}{4\alpha_i^\alpha} + \mathcal{O}(G^2) \right) \quad (3.90)$$

Damit ergibt sich für die $(\vec{G} = \vec{0})$ -Komponente des gesamten lokalen Potentials:

$$\begin{aligned} \lim_{\vec{G} \rightarrow \vec{0}} V_{lok}(\vec{G}) &= -\frac{4\pi e^2}{\Omega} \left(\sum_{\mu} N_v^{\alpha(\mu)} \right) \frac{1}{G^2} \\ &+ \sum_{\mu} \left(\delta V_{lok}^{\alpha(\mu)}(\vec{G}) \Big|_{\vec{G}=\vec{0}} + \frac{4\pi e^2}{\Omega} N_v^{\alpha(\mu)} \left(\sum_{i=1}^2 \frac{c_i^{\alpha(\mu)}}{4\alpha_i^{\alpha(\mu)}} \right) \right) \\ &+ \mathcal{O}(G^2) \end{aligned} \quad (3.91)$$

Die divergente ($\vec{G} = \vec{0}$)-Komponente der lokalen Energie ergibt sich dann wie folgt:

$$\begin{aligned} \lim_{\vec{G} \rightarrow \vec{0}} E_{lok}(\vec{G}) &= \lim_{\vec{G} \rightarrow \vec{0}} \left(\sum_{\mu} n^*(\vec{G}) S^{\mu}(\vec{G}) V_{lok}^{\alpha(\mu)}(\vec{G}) \right) \\ &= -\frac{4\pi e^2 N_v^2}{\Omega G^2} \\ &\quad + \sum_{\mu} N_v \left(\delta V_{lok}^{\alpha(\mu)}(\vec{G}) \Big|_{\vec{G}=\vec{0}} + \frac{4\pi e^2}{\Omega} N_v^{\alpha(\mu)} \left(\sum_{i=1}^2 \frac{c_i^{\alpha(\mu)}}{4\alpha_i^{\alpha(\mu)}} \right) \right) \\ &\quad + \mathcal{O}(G^2) \end{aligned} \quad (3.92)$$

Die ($\vec{G} = \vec{0}$)-Komponente der lokalen Energie ist somit eine Konstante, d.h. sie ist von den Wellenfunktionen und Atompositionen unabhängig und kann an dieser Stelle zu Null gesetzt werden. Damit wird auch die ($\vec{G} = \vec{0}$)-Komponente des lokalen Potentials explizit zu Null gesetzt:

$$V_{lok}(\vec{G}) \Big|_{\vec{G}=\vec{0}} \xrightarrow{!} 0 \quad (3.93)$$

Der tatsächliche Beitrag, insbesondere der divergente Anteil, wird in Kapitel 3.11 diskutiert. Er wird durch die ($\vec{G} = \vec{0}$)-Komponenten der Hartree-Energie und der Ewaldsumme kompensiert.

3.8 Der nichtlokale Anteil des Hamiltonian

Der nichtlokale Anteil des Hamiltonian setzt sich ebenfalls aus Beiträgen der einzelnen Atome μ zusammen, die auf den Positionen $\vec{R} + \vec{\tau}_{\mu}$ sitzen und der Atomsorte $\alpha(\mu)$ angehören und kurzreichweitig sind. Die $\vec{\tau}_{\mu}$ sind die Koordinaten der Atome in der Superzelle mit Gittervektor \vec{R} .

Unabhängig von der speziellen Form der Potentialoperatoren $\mathcal{V}_{NL}^{\alpha(\mu)}$ kann die Summation über die Superzellen \vec{R} und das Integral über eine einzige Superzelle Ω nach Gleichung (3.19) durch eine Integration über den ganzen Raum ersetzt werden. Für die nichtlokalen Potentialbeiträge ist folgender Beitrag zum Energiefunktional auszuwerten:

$$\langle \Psi | \mathcal{V}_{NL} | \Psi \rangle_{\Omega} = \sum_{\vec{R}} \sum_{\mu} \sum_{\vec{k}\nu} \omega_{\vec{k}} f_{\vec{k}\nu} \langle \Psi_{\vec{k}\nu} | \mathcal{T}_{\vec{R}+\vec{\tau}_{\mu}} \mathcal{V}_{NL}^{\alpha(\mu)} | \Psi_{\vec{k}\nu} \rangle_{\Omega} \quad (3.94)$$

$$\begin{aligned} &\downarrow \\ \langle \Psi | \tilde{\mathcal{V}}_{NL} | \Psi \rangle_{\infty} &= \sum_{\mu} \sum_{\vec{k}\nu} \omega_{\vec{k}} f_{\vec{k}\nu} \langle \Psi_{\vec{k}\nu} | \mathcal{T}_{\vec{\tau}_{\mu}} \mathcal{V}_{NL}^{\alpha(\mu)} | \Psi_{\vec{k}\nu} \rangle_{\infty} \end{aligned} \quad (3.95)$$

$$\mathcal{V}_{NL} = \sum_{\vec{R}} \sum_{\mu} \mathcal{T}_{\vec{R}+\vec{\tau}_{\mu}} \mathcal{V}_{NL}^{\alpha(\mu)} \quad (3.96)$$

$$\downarrow$$

$$\tilde{\mathcal{V}}_{NL} = \sum_{\mu} \mathcal{T}_{\vec{\tau}_{\mu}} \mathcal{V}_{NL}^{\alpha(\mu)} \quad (3.97)$$

Dabei ist $\mathcal{T}_{\vec{R}+\vec{\tau}_{\mu}}$ der Translationsoperator, der den nachfolgenden Potentialoperator $\mathcal{V}_{NL}^{\alpha(\mu)}$ um den Vektor $\vec{R} + \vec{\tau}_{\mu}$ verschiebt.

Im Computerprogramm sind zwei verschiedene Formen für das nichtlokale Potential implementiert. Dies sind zum einen die separablen KB-Potentiale \mathcal{V}_{KB} und zum anderen semilokale BHS-Potentiale \mathcal{V}_{SL} , die mittels Gridpunktmethode auf eine separable Form gebracht werden.

3.8.1 Die separablen KB-Potentiale

Die KB-Potentiale sind im Ortsraum separabel, d.h. sie lassen sich dort als Linearkombination von Projektionsoperatoren schreiben:

$$\mathcal{V}_{KB} = \sum_{\vec{R}} \sum_{\mu} \sum_{l=0}^{l_{max}} \sum_{m=-l}^{+l} E_l^{\alpha(\mu)} \mathcal{T}_{\vec{R}+\vec{\tau}_{\mu}} |t_{lm}^{\alpha(\mu)}\rangle \langle t_{lm}^{\alpha(\mu)}| \quad (3.98)$$

$$\downarrow$$

$$\tilde{\mathcal{V}}_{KB} = \sum_{\mu} \sum_{l=0}^{l_{max}} \sum_{m=-l}^{+l} E_l^{\alpha(\mu)} \mathcal{T}_{\vec{\tau}_{\mu}} |t_{lm}^{\alpha(\mu)}\rangle \langle t_{lm}^{\alpha(\mu)}| \quad (3.99)$$

$$\mathcal{T}_{\vec{\tau}_{\mu}} t_{lm}^{\alpha(\mu)}(\vec{r}) = \frac{t_l^{\alpha(\mu)}(|\vec{r} - \vec{\tau}_{\mu}|)}{|\vec{r} - \vec{\tau}_{\mu}|} Y_{lm}(\Omega_{\vec{r}-\vec{\tau}_{\mu}}) \quad (3.100)$$

Dabei ist $\mathcal{T}_{\vec{\tau}_{\mu}}$ in Gleichung (3.99) der Translationsoperator, der den nachfolgenden Projektor um $\vec{\tau}_{\mu}$ verschiebt. Auch im Impulsraum sind die die KB-Potentiale separabel. Dies wird unmittelbar einsichtig, wenn man die ebenen Wellen nach Besselfunktionen entwickelt (siehe Gleichung (A.35)):

$$\langle \vec{k} + \vec{G} | \mathcal{T}_{\vec{\tau}_{\mu}} t_{lm}^{\alpha(\mu)} \rangle_{\infty} = i^l S^{\mu}(\vec{k} + \vec{G}) Y_{lm}(\Omega_{\vec{k}+\vec{G}}) t_l^{\alpha(\mu)}(\vec{k} + \vec{G}) \quad (3.101)$$

$$t_l^{\alpha}(\vec{k} + \vec{G}) = \frac{4\pi}{\sqrt{\Omega}} \int_{\infty} dr r t_l^{\alpha}(r) j_l(|\vec{k} + \vec{G}|r) \quad (3.102)$$

$$\langle \vec{k} + \vec{G} | \mathcal{V}_{KB} | \vec{k} + \vec{G}' \rangle_{\Omega} = \sum_{lm} \sum_{\mu} E_l^{\alpha(\mu)} T_{lm}^{\mu}(\vec{k}, \vec{G}) T_{lm}^{\mu*}(\vec{k}, \vec{G}') \quad (3.103)$$

$$T_{lm}^{\mu}(\vec{k}, \vec{G}) = S^{\mu}(\vec{G}) \mathcal{Y}_{lm}(\Omega_{\vec{k}+\vec{G}}) t_l^{\alpha(\mu)}(\vec{k} + \vec{G}) \quad (3.104)$$

In Gleichung (3.104) von der reellen Darstellung \mathcal{Y}_{lm} der Kugelflächenfunktionen Gebrauch gemacht (A.26).

Der separable Beitrag zum Energiefunktional $\langle \Psi | \mathcal{H} | \Psi \rangle_\Omega$ bestimmt sich dann wie folgt:

$$\begin{aligned} \langle \Psi | \mathcal{V}_{KB} | \Psi \rangle_\Omega &= \sum_{\vec{k}\nu} \sum_{lm} \sum_{\mu} \omega_{\vec{k}} f_{\vec{k}\nu} \left(\sum_{\vec{G}} c_{\vec{k}\nu}^* (\vec{G}) T_{lm}^{\mu} (\vec{k}, \vec{G}) \right) E_l^{\alpha(\mu)} \quad (3.105) \\ &\quad \times \left(\sum_{\vec{G}'} T_{lm}^{\mu *} (\vec{k}, \vec{G}') c_{\vec{k}\nu} (\vec{G}') \right) \end{aligned}$$

Für den separablen Anteil zum Matrixvektorprodukt $\mathcal{H} | \Psi_{\vec{k}\nu} \rangle$ ergibt sich:

$$\begin{aligned} \langle \vec{r} | \mathcal{V}_{KB} | \Psi_{\vec{k}\nu} \rangle_\Omega &= \sum_{lm} \sum_{\mu} \left(\sum_{\vec{G}'} T_{lm}^{\mu *} (\vec{k}, \vec{G}') c_{\vec{k}\nu} (\vec{G}') \right) E_l^{\alpha(\mu)} \quad (3.106) \\ &\quad \times \sum_{\vec{G}} \frac{1}{\sqrt{\Omega}} T_{lm}^{\mu} (\vec{k}, \vec{G}) e^{-i(\vec{k}+\vec{G})\vec{r}} \end{aligned}$$

$$\langle \vec{k} + \vec{G} | \mathcal{V}_{KB} | \Psi_{\vec{k}\nu} \rangle_\Omega = \sum_{lm} \sum_{\mu} T_{lm}^{\mu} (\vec{k}, \vec{G}) E_l^{\alpha(\mu)} \left(\sum_{\vec{G}'} T_{lm}^{\mu *} (\vec{k}, \vec{G}') c_{\vec{k}\nu} (\vec{G}') \right) \quad (3.107)$$

Der Beitrag der separablen Kraft zur Hellman–Feynman–Kraft ergibt sich durch Ableiten der Gleichung (3.105) nach den Atomkoordinaten \vec{r}_μ . Dabei hängen die $T_{lm}^{\mu} (\vec{k}, \vec{G})$ nur über die Strukturformfaktoren $S^{\mu} (\vec{G})$ von den \vec{r}_μ ab (siehe Gleichung (3.104)):

$$\begin{aligned} \vec{F}_{KB}^{\mu} &= -\frac{\partial}{\partial \vec{r}_\mu} \langle \Psi | \mathcal{V}_{KB} | \Psi \rangle_\Omega \\ &= -\sum_{\vec{k}\nu} \sum_{lm} \omega_{\vec{k}} f_{\vec{k}\nu} \left(\sum_{\vec{G}} i\vec{G} c_{\vec{k}\nu}^* (\vec{G}) T_{lm}^{\mu} (\vec{k}, \vec{G}) \right) E_l^{\alpha(\mu)} \quad (3.108) \\ &\quad \times \left(\sum_{\vec{G}'} T_{lm}^{\mu *} (\vec{k}, \vec{G}') c_{\vec{k}\nu} (\vec{G}') \right) + cc. \end{aligned}$$

$$\begin{aligned} &= -\sum_{\vec{k}\nu} \sum_{lm} \omega_{\vec{k}} f_{\vec{k}\nu} \left(\sum_{\vec{G}} 2i\vec{G} c_{\vec{k}\nu}^* (\vec{G}) T_{lm}^{\mu} (\vec{k}, \vec{G}) \right) E_l^{\alpha(\mu)} \quad (3.109) \\ &\quad \times \left(\sum_{\vec{G}'} T_{lm}^{\mu *} (\vec{k}, \vec{G}') c_{\vec{k}\nu} (\vec{G}') \right) \end{aligned}$$

Alle separablen Beiträge, i.e. Energiebeitrag, Matrixvektorprodukt und Kraft lassen sich direkt im Impulsraum auswerten. Die Gleichungen lassen sich unter Benutzung eines vierkomponentigen Vektors $\mathcal{D}_{lm\vec{k}\nu}^{\mu}$ wie folgt zusammenfassen:

$$\mathcal{D}_{lm\bar{k}\nu}^\mu = \begin{pmatrix} \overset{\circ}{\mathcal{D}}_{lm\bar{k}\nu}^\mu \\ \vec{\mathcal{D}}_{lm\bar{k}\nu}^\mu \end{pmatrix} = \sum_{\vec{G}} \begin{pmatrix} 1 \\ -2i\vec{G} \end{pmatrix} c_{\bar{k}\nu}^* (\vec{G}) T_{lm}^\mu (\vec{k}, \vec{G}) \quad (3.110)$$

$$T_{lm}^\mu (\vec{k}, \vec{G}) = S^\mu (\vec{G}) \mathcal{Y}_{lm} (\Omega_{\vec{k}+\vec{G}}) t_l^{\alpha(\mu)} (\vec{k} + \vec{G}) \quad (3.111)$$

$$t_l^\alpha (\vec{k} + \vec{G}) = \frac{4\pi}{\sqrt{\Omega}} \int_{\infty}^0 dr r t_l^\alpha (r) j_l (|\vec{k} + \vec{G}| r) \quad (3.112)$$

Damit ergibt sich für die einzelnen Beiträge:

$$\langle \Psi | \mathcal{V}_{KB} | \Psi \rangle_{\Omega} = \sum_{\bar{k}\nu} \sum_{lm} \sum_{\mu} \omega_{\bar{k}} f_{\bar{k}\nu} E_l^{\alpha(\mu)} \overset{\circ}{\mathcal{D}}_{lm\bar{k}\nu}^\mu \overset{\circ}{\mathcal{D}}_{lm\bar{k}\nu}^{\mu*} \quad (3.113)$$

$$\begin{aligned} \langle \vec{r} | \mathcal{V}_{KB} | \Psi_{\bar{k}\nu} \rangle_{\Omega} &= \sum_{lm} \sum_{\mu} E_l^{\alpha(\mu)} \overset{\circ}{\mathcal{D}}_{lm\bar{k}\nu}^{\mu*} \quad (3.114) \\ &\times \sum_{\vec{G}} \frac{1}{\sqrt{\Omega}} T_{lm}^\mu (\vec{k}, \vec{G}) e^{-i(\vec{k}+\vec{G})\vec{r}} \end{aligned}$$

$$\langle \vec{k} + \vec{G} | \mathcal{V}_{KB} | \Psi_{\bar{k}\nu} \rangle_{\Omega} = \sum_{lm} \sum_{\mu} E_l^{\alpha(\mu)} \overset{\circ}{\mathcal{D}}_{lm\bar{k}\nu}^{\mu*} T_{lm}^\mu (\vec{k}, \vec{G}) \quad (3.115)$$

$$\vec{F}_{KB}^\mu = -\frac{\partial}{\partial \vec{r}_\mu} (\langle \Psi | \mathcal{V}_{KB} | \Psi \rangle_{\Omega}) = \sum_{\bar{k}\nu} \sum_{lm} \omega_{\bar{k}} f_{\bar{k}\nu} E_l^{\alpha(\mu)} \vec{\mathcal{D}}_{lm\bar{k}\nu}^\mu \overset{\circ}{\mathcal{D}}_{lm\bar{k}\nu}^{\mu*} \quad (3.116)$$

3.8.2 Die semilokalen BHS–Potentiale

Die BHS–Potentiale sind im Ortsraum semilokal, d.h. lokal in r und nichtlokal in den Winkelkoordinaten $\Omega_{\vec{r}}$.

$$\mathcal{V}_{SL} = \sum_{\vec{R}} \sum_{\mu} \sum_{l=0}^{l_{max}} \mathcal{T}_{\vec{R}+\vec{r}_\mu} (V_l^{\alpha(\mu)} \mathcal{P}_l) \quad (3.117)$$

$$\downarrow$$

$$\tilde{\mathcal{V}}_{SL} = \sum_{\mu} \sum_{l=0}^{l_{max}} \mathcal{T}_{\vec{r}_\mu} (V_l^{\alpha(\mu)} \mathcal{P}_l) \quad (3.118)$$

$$\mathcal{T}_{\vec{r}_\mu} (V_l^{\alpha(\mu)} \mathcal{P}_l) (\vec{r}) = V_l^{\alpha(\mu)} (\vec{r} - \vec{r}_\mu) \mathcal{P}_l (\Omega_{\vec{r}-\vec{r}_\mu}) \quad (3.119)$$

\mathcal{P}_l ist der Projektor auf die Drehimpulskomponente l . Im Impulsraum ergibt sich bei Entwicklung der ebenen Wellen nach Besselfunktionen (siehe Gleichung (A.35)) folgende Darstellung:

$$\langle \vec{k} + \vec{G} | \mathcal{T}_{\vec{r}_\mu} (V_l^{\alpha(\mu)} \mathcal{P}_l) | \vec{k} + \vec{G}' \rangle_{\infty} \quad (3.120)$$

$$= S^\mu (\vec{G} - \vec{G}') \langle \vec{k} + \vec{G} | (V_l^{\alpha(\mu)} \mathcal{P}_l) | \vec{k} + \vec{G}' \rangle_{\infty}$$

$$= \sum_{m=-l}^l S^\mu (\vec{G}) \mathcal{Y}_{lm} (\vec{k} + \vec{G}) S^{\mu*} (\vec{G}') \mathcal{Y}_{lm} (\vec{k} + \vec{G}') \quad (3.121)$$

$$\times \frac{(4\pi)^2}{\Omega} \int_{\infty}^0 r^2 dr j_l (|\vec{k} + \vec{G}| r) V_l^{\alpha(\mu)} (r) j_l (|\vec{k} + \vec{G}'| r)$$

In Gleichung (3.121) von der reellen Darstellung \mathcal{Y}_{lm} der Kugelflächenfunktionen Gebrauch gemacht (A.26).

Das Integral ist leider im Impulsraum nicht separabel. Trotzdem kann der ganze Ausdruck in eine separable Form gezwungen werden, indem man ausnutzt, dass numerische Integration einer Funktion immer eine gewichtete Summation der Funktionswerte an endlich vielen Stützpunkten bedeutet. In unserem Fall ist es wichtig mit möglichst wenig Stützpunkten auszukommen. Wir benutzen deshalb das Gauß-Hermite-Integrationschema (siehe Anhang A):

$$\int_{-\infty}^{\infty} dr e^{-\beta r^2} f(r) = \sum_{i=1}^N \frac{\omega_{i,N}}{\sqrt{\beta}} f\left(\frac{r_{i,N}}{\sqrt{\beta}}\right) \quad (3.122)$$

$$\int_{-\infty}^{\infty} dr f(r) = \sum_{i=1}^N \frac{\omega_{i,N}}{\sqrt{\beta}} e^{r_{i,N}^2} f\left(\frac{r_{i,N}}{\sqrt{\beta}}\right) \quad (3.123)$$

Die Stützpunkte $r_{i,N}$ und die Gewichte $\omega_{i,N}$ liegen in tabellierter Form vor. Der Parameter β ist im Prinzip frei wählbar, wird durch die räumliche Ausdehnung der Potentiale aber festgelegt. Da die Potentiale lokalisiert sind, das Integrationsgebiet also beschränkt ist, sollte auch der äußerste Stützpunkt $\frac{r_{i,N}}{\sqrt{\beta}}$ noch im Integrationsgebiet liegen. Wir benutzen daher räumliche Ausdehnung der tabellierten Pseudopotentiale aus [BHS82], um für jede Atomsorte α ein β^α zu bestimmen. Diese tabellierten Potentiale haben folgende Struktur:

$$V_l^\alpha(r) = \sum_{j=1}^6 (A_{l,j}^\alpha + B_{l,j}^\alpha r^2) e^{-\beta_{l,j}^\alpha r^2} \quad (3.124)$$

$$\beta^\alpha = \min_{l,j} \beta_{l,j}^\alpha \quad (3.125)$$

In der Regel reichen 16–24 Stützpunkte aus, um die Integrale mit numerisch erreichbarer Genauigkeit zu bestimmen. In unserem Fall braucht nur über die positiven Stützpunkte summiert zu werden, da das Integral nur über $r \geq 0$ auszuführen ist und der Integrand für $r = 0$ verschwindet. Dies sind $\hat{N} = \frac{N}{2}$ bzw. $\hat{N} = \frac{N+1}{2}$ Stützstellen, je nachdem ob N gerade oder ungerade ist.

Damit läßt sich das Integral aus Gleichung (3.121) in separabler Form darstellen:

$$\langle \vec{k} + \vec{G} | \mathcal{T}_{\vec{r}_\mu} (V_l^{\alpha(\mu)} \mathcal{P}_l) | \vec{k} + \vec{G}' \rangle_\infty \quad (3.126)$$

$$= \sum_{i=1}^{\hat{N}} \sum_{m=-l}^l W_l^{\alpha(\mu)i} T_{lm}^{\mu i}(\vec{k}, \vec{G}) T_{lm}^{\mu i*}(\vec{k}, \vec{G}') \\ W_l^{\alpha i} = \frac{(4\pi)^2}{\Omega} \frac{\omega_{i,N}}{\sqrt{\beta^\alpha}} e^{r_{i,N}^2} \frac{r_{i,N}^2}{\beta^\alpha} V_l^\alpha\left(\frac{r_{i,N}}{\sqrt{\beta^\alpha}}\right) \quad (3.127)$$

$$T_{lm}^{\mu i}(\vec{k}, \vec{G}) = S^\mu(\vec{G}) \mathcal{Y}_{lm}(\vec{k} + \vec{G}) j_l\left(|\vec{k} + \vec{G}| \frac{r_{i,N}}{\sqrt{\beta^{\alpha(\mu)}}}\right) \quad (3.128)$$

Völlig analog zum separablen Fall der KB-Potentiale lassen sich so die semilokalen Beiträge zu Energiefunktional, Matrixvektorprodukt und Kraft im Impulsraum auswerten. Die Gleichungen lassen sich ebenfalls unter Benutzung eines vierkomponentigen Vektors $\mathcal{D}_{lm\vec{k}\nu}^{\mu i}$ zusammenfassen:

$$\mathcal{D}_{lm\vec{k}\nu}^{\mu i} = \begin{pmatrix} \overset{o}{\mathcal{D}}_{lm\vec{k}\nu}^{\mu i} \\ \vec{\mathcal{D}}_{lm\vec{k}\nu}^{\mu i} \end{pmatrix} = \sum_{\vec{G}} \begin{pmatrix} 1 \\ -2i\vec{G} \end{pmatrix} c_{\vec{k}\nu}^*(\vec{G}) T_{lm}^{\mu i}(\vec{k}, \vec{G}) \quad (3.129)$$

Damit ergibt sich für die einzelnen Beiträge:

$$\langle \Psi | \mathcal{V}_{SL} | \Psi \rangle_{\Omega} = \sum_{\vec{k}\nu} \sum_{lmi} \sum_{\mu} \omega_{\vec{k}} f_{\vec{k}\nu} W_l^{\alpha(\mu)i} \overset{o}{\mathcal{D}}_{lm\vec{k}\nu}^{\mu i} \overset{o}{\mathcal{D}}_{lm\vec{k}\nu}^{\mu i*} \quad (3.130)$$

$$\begin{aligned} \langle \vec{r} | \mathcal{V}_{SL} | \Psi_{\vec{k}\nu} \rangle_{\Omega} &= \sum_{lmi} \sum_{\mu} W_l^{\alpha(\mu)i} \overset{o}{\mathcal{D}}_{lm\vec{k}\nu}^{\mu i*} \quad (3.131) \\ &\times \sum_{\vec{G}} \frac{1}{\sqrt{\Omega}} T_{lm}^{\mu i}(\vec{k}, \vec{G}) e^{-i(\vec{k}+\vec{G})\vec{r}} \end{aligned}$$

$$\langle \vec{k} + \vec{G} | \mathcal{V}_{SL} | \Psi_{\vec{k}\nu} \rangle_{\Omega} = \sum_{lmi} \sum_{\mu} W_l^{\alpha(\mu)i} \overset{o}{\mathcal{D}}_{lm\vec{k}\nu}^{\mu i*} T_{lm}^{\mu i}(\vec{k}, \vec{G}) \quad (3.132)$$

$$\vec{F}_{SL}^{\mu} = -\frac{\partial}{\partial \vec{r}_{\mu}} (\langle \Psi | \mathcal{V}_{SL} | \Psi \rangle_{\Omega}) = \sum_{\vec{k}\nu} \sum_{lmi} \omega_{\vec{k}} f_{\vec{k}\nu} W_l^{\alpha(\mu)i} \vec{\mathcal{D}}_{lm\vec{k}\nu}^{\mu i} \overset{o}{\mathcal{D}}_{lm\vec{k}\nu}^{\mu i*} \quad (3.133)$$

Von der Struktur her unterscheidet sich der semilokale Fall vom separablen nur durch die zusätzliche Summation über die Gridpunkte $r_{i,N}$. Deshalb ist die Benutzung von semilokalen BHS-Potentialen auch \hat{N} mal langsamer als die Anwendung von separablen KB-Potentialen, wobei \hat{N} die Anzahl der Gauß-Hermite-Stützpunkte ist.

3.9 Austausch und Korrelation

Die Austausch- und Korrelationsenergie wird in der lokalen Dichtenäherung beschrieben:

$$\begin{aligned} E_{xc}[n_{pc}(\vec{r}) + n(\vec{r})] &= \int_{\Omega} d^3r (n_{pc}(\vec{r}) + n(\vec{r})) \epsilon_{xc}^{LDA}[n_{pc}(\vec{r}) + n(\vec{r})] \quad (3.134) \\ &= \frac{1}{\Omega} \sum_{\vec{G}} (n_{pc}^*(\vec{G}) + n^*(\vec{G})) \epsilon_{xc}(\vec{G}) \end{aligned}$$

$$\epsilon_{xc}(\vec{G}) = \frac{1}{\Omega} \int_{\Omega} d^3r \epsilon_{xc}^{LDA} [n_{pc}(\vec{r}) + n(\vec{r})] e^{i\vec{G}\vec{r}} \quad (3.135)$$

$$n_{pc}(\vec{G}) + n(\vec{G}) = \int_{\Omega} d^3r (n_{pc}(\vec{r}) + n(\vec{r})) e^{i\vec{G}\vec{r}} \quad (3.136)$$

Die Austauschkorrelationsenergie wird über Gleichung (3.134) im Ortsraum ausgerechnet. $n_{pc}(\vec{r})$ ist die partielle Coreladungsichte. Sie setzt sich aus den Beiträgen der einzelnen Atome $\alpha(\mu)$ an den Positionen $\vec{r}_{\mu} + \vec{R}$ und wird im Impulsraum berechnet:

$$\begin{aligned} n_{pc}(\vec{r}) &= \sum_{\vec{R}} \sum_{\mu} n_{pc}^{\alpha(\mu)}(\vec{r} - (\vec{r}_{\mu} + \vec{R})) \\ &= \frac{1}{\Omega} \sum_{\vec{G}} n_{pc}(\vec{G}) e^{-i\vec{G}\vec{r}} \end{aligned} \quad (3.137)$$

$$\begin{aligned} n_{pc}(\vec{G}) &= \sum_{\mu} \int_{\infty} d^3r n_{pc}^{\alpha(\mu)}(\vec{r} - \vec{r}_{\mu}) e^{i\vec{G}\vec{r}} \\ &= \sum_{\mu} S^{\mu}(\vec{G}) n_{pc}^{\alpha(\mu)}(\vec{G}) \end{aligned} \quad (3.138)$$

$$\frac{d}{d\vec{r}_{\mu}} n_{pc}(\vec{G}) = i\vec{G} S^{\mu}(\vec{G}) n_{pc}^{\alpha(\mu)}(\vec{G}) \quad (3.139)$$

$$\begin{aligned} n_{pc}^{\alpha(\mu)}(\vec{G}) &= \int_{\infty} d^3r n_{pc}^{\alpha(\mu)}(\vec{r}) e^{i\vec{G}\vec{r}} \\ &= 4\pi \int_{\infty} r^2 dr n_{pc}^{\alpha(\mu)}(r) j_0(Gr) \end{aligned} \quad (3.140)$$

Der Beitrag zum Matrixvektorprodukt wird im Ortsraum ausgewertet:

$$\begin{aligned} \langle \vec{r} | V_{xc} | \Psi_{\vec{k}\nu} \rangle_{\Omega} &= \left[\frac{\delta}{\delta n} E_{xc}[n] \right]_{n=n_{pc}(\vec{r})+n(\vec{r})} \langle \vec{r} | \Psi_{\vec{k}\nu} \rangle \\ &= V_{xc}[n_{pc}(\vec{r}) + n(\vec{r})] \Psi_{\vec{k}\nu}(\vec{r}) \end{aligned} \quad (3.141)$$

$$V_{xc}[n(\vec{r})] = \epsilon_{xc}^{LDA}[n(\vec{r})] + n(\vec{r}) \frac{d\epsilon_{xc}^{LDA}}{dn}[n(\vec{r})] \quad (3.142)$$

Der Beitrag zur Hellman–Feynman–Kraft auf das Atom $\alpha(\mu)$ ergibt sich durch Ableiten der Austauschkorrelationsenergie nach der Atomkoordinate \vec{r}_{μ} .

$E_{xc}[n_{pc}(\vec{r}) + n(\vec{r})]$ hängt nur über die partielle Coreladungsichte $n_{pc}(\vec{r})$ von den Atomkoordinaten \vec{r}_{μ} ab. Deshalb gibt es nur dann einen Beitrag zur Kraft, wenn die partielle Coreladungsichte überhaupt berücksichtigt wird. Der Kraftbeitrag kann im Impulsraum ausgewertet werden:

$$\vec{F}_{xc}^{\mu} = -\frac{\partial}{\partial \vec{r}_{\mu}} E_{xc} [n_{pc}(\vec{r}) + n(\vec{r})] \quad (3.143)$$

$$= -\int_{\Omega} d^3r \frac{d}{dn_{pc}} \left\{ \epsilon_{xc}^{LDA} [n_{pc}(\vec{r}) + n(\vec{r})] (n_{pc}(\vec{r}) + n(\vec{r})) \right\} \frac{dn_{pc}}{d\vec{r}_{\mu}}$$

$$= -\int_{\Omega} d^3r V_{xc} [n_{pc}(\vec{r}) + n(\vec{r})] \frac{dn_{pc}}{d\vec{r}_{\mu}}$$

$$= \sum_{\vec{G}} -i\vec{G} V_{xc}^* (\vec{G}) S^{\mu} (\vec{G}) n_{pc}^{\alpha(\mu)} (\vec{G}) \quad (3.144)$$

$$V_{xc} (\vec{G}) = \frac{1}{\Omega} \int_{\Omega} d^3r V_{xc} [n_{pc}(\vec{r}) + n(\vec{r})] e^{i\vec{G}\vec{r}} \quad (3.145)$$

3.10 Die Ewaldsumme

Die Coulombenergie der Ionenrümpfe bezeichnet man als Ewaldsumme:

$$E_{ion,ion} = \frac{1}{2} e^2 \sum_{\vec{R}} \sum_{\mu\mu'} \frac{Z^{\alpha(\mu)} Z^{\alpha(\mu')}}{|\vec{r}_{\mu} - (\vec{R} + \vec{r}_{\mu'})|} \left(1 - \delta_{\vec{r}_{\mu}, \vec{R} + \vec{r}_{\mu'}}\right) \quad (3.146)$$

$$= \frac{1}{2} e^2 \sum_{\vec{G}} \sum_{\mu\mu'} Z^{\alpha(\mu)} Z^{\alpha(\mu')} \frac{1}{\Omega} \frac{4\pi}{G^2} e^{-i\vec{G}(\vec{r}_{\mu} - \vec{r}_{\mu'})} \quad (3.147)$$

Der Faktor $\left(1 - \delta_{\vec{r}_{\mu}, \vec{R} + \vec{r}_{\mu'}}\right)$ schließt identische Ionenrümpfe aus der Doppelsummation aus.

Die Ewaldsumme ist divergent. Dies sieht man unmittelbar an der Divergenz der $(\vec{G} = \vec{0})$ -Komponente, wenn man die Summation über die Gittervektoren \vec{R} durch eine Summation im reziproken Gitter ersetzt (siehe Gleichung (A.41)).

Die $(\vec{G} = \vec{0})$ -Komponente ist allerdings eine Konstante, d.h. sie hängt weder von den Wellenfunktionen noch von den Atomkoordinaten ab. Sie wird durch die divergenten Anteile von Hartree-Energie und lokaler Energie kompensiert und kann an dieser Stelle explizit zu Null gesetzt werden:

$$E_{ion,ion}(\vec{G}) \Big|_{\vec{G}=\vec{0}} = \frac{1}{2} \frac{N_v^2 4\pi e^2}{\Omega G^2} \xrightarrow{!} 0 \quad (3.148)$$

Die Gittersumme wird aus Konvergenzgründen mit dem Hilfsmittel der Ewaldsummation A.43 teilweise im Impulsraum und teilweise im Ortsraum ausgeführt:

$$\begin{aligned} E_{ion,ion} &= \frac{1}{2} e^2 \sum_{\mu\mu'} Z^{\alpha(\mu)} X(\vec{r}_{\mu} - \vec{r}_{\mu'}) Z^{\alpha(\mu')} \\ &= \frac{1}{2} e^2 \sum_{\mu\mu'} Z^{\alpha(\mu)} X_{\mu\mu'} Z^{\alpha(\mu')} \end{aligned} \quad (3.149)$$

$$X_{\mu\mu'} = \sum_{\vec{R}} \frac{(1 - \delta_{\vec{\tau}_{\mu\mu'}, \vec{R}})}{|\vec{\tau}_{\mu\mu'} - \vec{R}|} \Big|_{\vec{\tau}_{\mu\mu'} = \vec{\tau}_\mu - \vec{\tau}_{\mu'}} \quad (3.150)$$

$$= \frac{4\pi}{\Omega} \sum_{\vec{G}} \frac{1}{G^2} \cos(\vec{G}\vec{\tau}_{\mu\mu'}) e^{-\frac{G^2}{4\eta^2}} - \frac{\pi}{\eta^2 \Omega} \quad (3.151)$$

$$+ \sum_{\vec{R}} \frac{\operatorname{erfc}(\eta |\vec{\tau}_{\mu\mu'} - \vec{R}|)}{|\vec{\tau}_{\mu\mu'} - \vec{R}|} (1 - \delta_{\vec{\tau}_{\mu\mu'}, \vec{R}}) - 2 \frac{\eta}{\sqrt{\pi}} \delta_{\vec{\tau}_\mu, \vec{\tau}_{\mu'}}$$

Die Ewaldkraft erhält man durch Ableiten der Ewaldenergie nach den Atomkoordinaten τ_μ :

$$-\frac{d}{d\vec{\tau}_\mu} \left(\frac{1}{2} e^2 \sum_{\mu\mu'} Z^{\alpha(\mu)} Z^{\alpha(\mu')} X_{\mu\mu'} \right) = -e^2 \sum_{\mu'} Z^{\alpha(\mu)} \frac{d}{d\vec{\tau}_\mu} (X_{\mu\mu'}) Z^{\alpha(\mu')} \quad (3.152)$$

$$\frac{d}{d\vec{\tau}_\mu} (X_{\mu\mu'}) = \frac{4\pi}{\Omega} \sum_{\vec{G}} \frac{-\vec{G} \sin(\vec{G}\vec{\tau}_{\mu\mu'})}{G^2} e^{-\frac{G^2}{4\eta^2}} \quad (3.153)$$

$$+ \sum_{\vec{R}} \frac{\vec{\tau}_{\mu\mu'} - \vec{R}}{|\vec{\tau}_{\mu\mu'} - \vec{R}|^3} (1 - \delta_{\vec{\tau}_{\mu\mu'}, \vec{R}})$$

$$\times \left\{ \operatorname{erfc}(\eta |\vec{\tau}_{\mu\mu'} - \vec{R}|) + \frac{2\eta}{\sqrt{\pi}} |\vec{\tau}_{\mu\mu'} - \vec{R}| e^{-(\eta |\vec{\tau}_{\mu\mu'} - \vec{R}|)^2} \right\}$$

Beide Teilsummen konvergieren exponentiell. Die erste Teilsumme konvergiert exponentiell mit $e^{-\frac{G^2}{4\eta^2}}$, die zweite mit $\operatorname{erfc}(\eta R) \leq e^{-\eta^2 R^2}$. Bricht man die Summationen bei einer relative Genauigkeit ϵ_{Ewald} der Summanden ab, so ergibt sich folgendes Abbruchkriterium:

$$\epsilon_{Ewald} = e^{-\frac{G_{max}^2}{4\eta^2}} = e^{-\eta^2 R_{max}^2} \quad (3.154)$$

Die Teilsummationen werden dann nur bis G_{max} bzw. $R_{max} + d(\Omega)$ ausgeführt. $d(\Omega)$ ist der Durchmesser der Superzelle und wird hinzugefügt, um den Einfluss der Koordinaten $\vec{\tau}_{\mu\mu'}$ abzuschätzen:

$$|\vec{\tau}_{\mu\mu'} - \vec{R}| \leq |R - d(\Omega)| \quad (3.155)$$

Der konvergenzerzeugende Faktor η wird so gewählt, daß beide Teilsummen möglichst gleich schnell konvergieren, d.h. ungefähr gleich viele Summanden haben.

3.11 Die divergenten Terme

Die $(\vec{G} = \vec{0})$ -Komponenten der Hartree-Energie (3.60), der lokalen Pseudopotentialenergie (3.92) und der Ewaldenergie (3.148) divergieren jede für sich. Zusammen kompensieren sich die divergenten Anteile. Übrig bleibt eine nur von den Pseudopotentialen abhängige Konstante, die zum Energiefunktional addiert wird:

$$\begin{aligned} E_{lok}^0 &= \frac{N_v}{\Omega} \sum_{\mu} \int d^3r \left(V_{lok}^{\alpha(\mu)}(\vec{r}) + \frac{e^2 N_v^{\alpha(\mu)}}{r} \right) \\ &= \sum_{\mu} N_v \left(\delta V_{lok}^{\alpha(\mu)}(\vec{G}) \Big|_{\vec{G}=\vec{0}} + \frac{4\pi e^2}{\Omega} N_v^{\alpha(\mu)} \left(\sum_{i=1}^2 \frac{c_i^{\alpha(\mu)}}{4\alpha_i^{\alpha(\mu)}} \right) \right) = N_v \alpha_1 \end{aligned} \quad (3.156)$$

Nur wegen dieser Kompensation ist es erlaubt, bei den Teilenergien und Potentialen, die entsprechende $(\vec{G} = \vec{0})$ -Komponente wegzulassen.

3.12 Energie und Kraft im Computerprogramm

Die Gesamtenergie pro Einheitszelle E_{tot} setzt sich nach Gleichung (3.1) aus einem elektronischen Anteil und der Coulombenergie des Ionengitters $E_{ion,ion}$ zusammen. Der elektronische Anteil zur Gesamtenergie

$$T_0[n] + E_{lok}[n] + E_{NL}[n] + E_H[n] + E_{xc}[n]$$

wird im Computerprogramm allerdings nicht durch explizites Ausrechnen der Einzelanteile bestimmt. Vielmehr wird der in Kapitel 1.5 beschriebene Weg gewählt, i.e. der elektronische Anteil der totalen Energie wird im Computerprogramm aus Gleichung (1.45) bzw. (1.35) berechnet. Dazu sind die Ewaldenergie E_{Ewald} (Gleichung (3.149)) und die von den lokalen Pseudopotentialen abhängige Konstante E_{lok}^0 (Gleichung (3.156)) zu addieren.

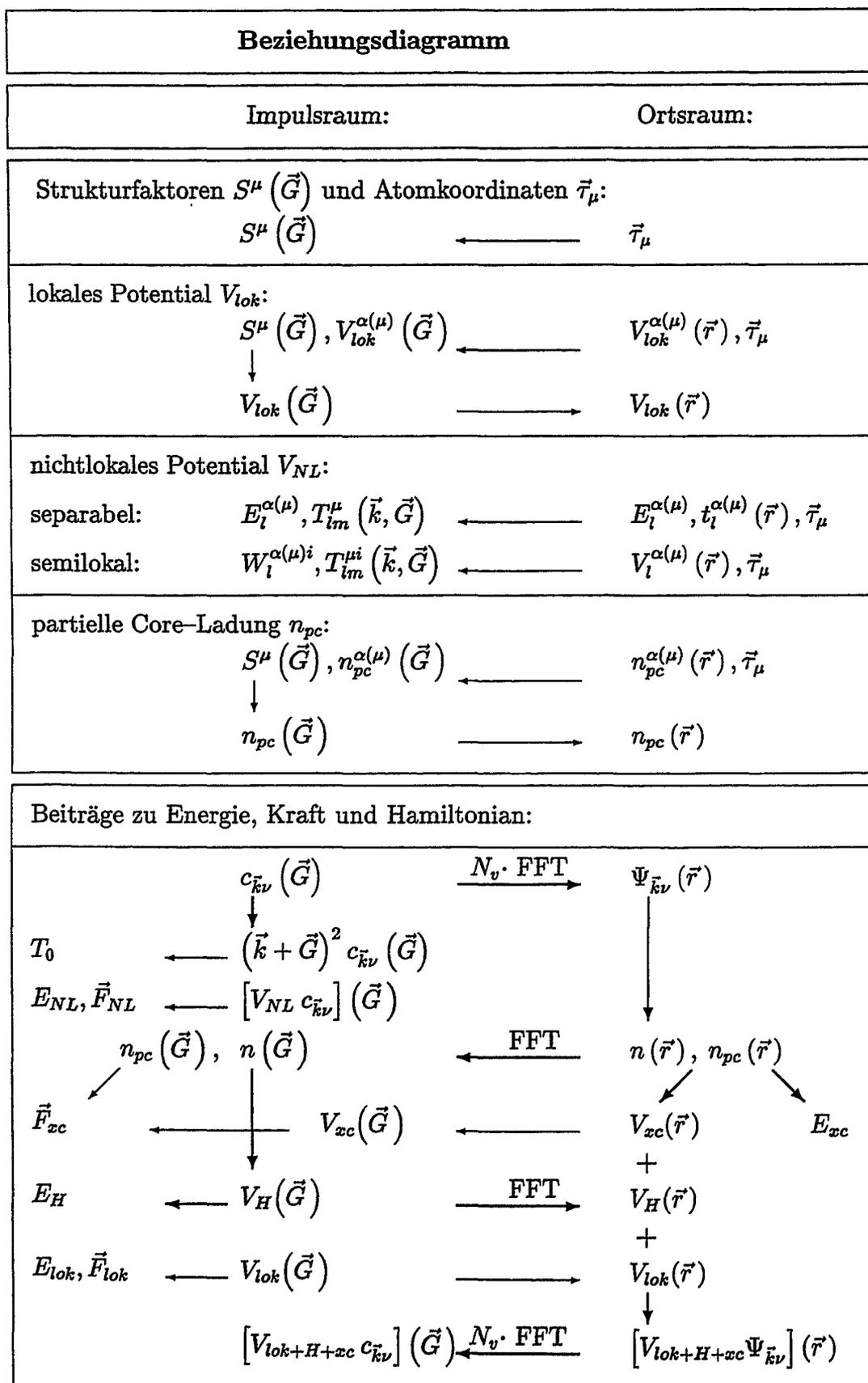
Insgesamt ergibt sich für die totale Energie der (j)-ten Selbstkonsistenziteration:

$$\begin{aligned} E_{tot}[n^{(j)}] &= 2 \sum_i f_i^{(j)} \varepsilon_i^{(j)} + E_{Ewald} + E_{lok}^0 \\ &\quad + \sum_{\vec{G}} n(\vec{G}) \left(\frac{1}{2} V_H^{(j)}(\vec{G}) - V_H^{(j-1)}(\vec{G}) \right) \\ &\quad + \int d\vec{r} \left(n^{(j)}(\vec{r}) + n_{pc}(\vec{r}) \right) \varepsilon_{xc} \left[n^{(j)}(\vec{r}) + n_{pc}(\vec{r}) \right] \\ &\quad - \int d\vec{r} n^{(j)}(\vec{r}) V_{xc} \left[n^{(j-1)}(\vec{r}) + n_{pc}(\vec{r}) \right] \end{aligned} \quad (3.157)$$

Wie in der Gleichung angedeutet, wird der Hartree-Anteil zur Doppelzählkorrektur im Impulsraum ausgerechnet und der Austausch- und Korrelationsanteil im Ortsraum.

Die Kraft \vec{F}^μ ein Atom μ setzt sich aus vier Anteilen zusammen, der Ewaldkraft \vec{F}_{Ewald}^μ (Gleichung (3.152)), der lokalen Pseudopotentialkraft \vec{F}_{lok}^μ (Gleichungen (3.81)), der nichtlokalen Pseudopotentialkraft \vec{F}_{NL}^μ , d.h. entweder \vec{F}_{KB}^μ (3.116) oder \vec{F}_{SL}^μ (3.133) sowie der Kraft \vec{F}_{xc}^μ , die nur durch Berücksichtigung der partiellen Rumpfladung in E_{xc} entsteht (Gleichung (3.144)):

$$\vec{F}^\mu = \vec{F}_{Ewald}^\mu + \vec{F}_{lok}^\mu + \vec{F}_{NL}^\mu + \vec{F}_{xc}^\mu \quad (3.158)$$



Kapitel 4

Lösung der Kohn–Sham–Gleichungen

Möchte man die atomare Gleichgewichtsstruktur eines Systems berechnen, so zerfällt dieses Problem in drei Teile:

- Berechnung der elektronischen Struktur
- Berechnung der Kräfte auf die Atome
- Relaxation der Atome in ihre Gleichgewichtslage

Bei Verwendung der Pseudopotential–Methode in der ebenen Wellen Darstellung benötigt die Berechnung der Kräfte und die Relaxation der Atome vergleichsweise geringen numerischen Aufwand. Der Grund hierfür liegt in der Unabhängigkeit der Basisfunktionen von den Atomkoordinaten.

Der sowohl rechenzeit- als auch speicherintensivste Schritt liegt in der Lösung des elektronischen Eigenwertproblems, i.e. in der Lösung der Kohn–Sham–Gleichungen. Die maximal rechenbare Systemgröße wird deshalb im wesentlichen durch die Anzahl der elektronischen Freiheitsgrade bestimmt. Zur Beschreibung von größeren Systemen mit vielen Atomen pro Einheitszelle oder Atomen mit tiefen Pseudopotentialen benötigt man leicht einige tausend ebene Wellen als Basisfunktionen. Hier stößt man mit traditionellen Methoden zur Lösung der Kohn–Sham–Gleichungen an die Grenzen der Leistungsfähigkeit moderner Computer.

Einen Ausweg zeigten 1985 Car und Parrinello [CP85], als sie gleichzeitig drei Verbesserungen präsentierten, um sowohl Rechenzeit und als auch Speicherplatz zu sparen:

- Einführung von iterativen Methoden zur Lösung des elektronischen Eigenwertproblems
- gleichzeitige Berechnung der elektronischen und ionischen Freiheitsgrade
- Einführung der Fast-Fourier-Transformation (FFT) zur Berechnung der Produkte aus Hamiltonian und Wellenfunktionen

Diese Neuerungen haben der Pseudopotentialmethode völlig neue Problemstellungen zugänglich gemacht.

4.1 Traditionelle Bandstrukturmethode

Bei der klassischen Methode zur selbstkonsistenten Lösung der Kohn-Sham-Gleichungen (4.1) startet man mit einer geratenen Startladungsdichte $n_0(\vec{r})$, um aus ihr das Startpotential $V_{eff}[n_0(\vec{r})]$ zu konstruieren.

$$\left(-\partial_{\vec{r}}^2 + V_{eff}[n(\vec{r})]\right)\psi_{\nu}(\vec{r}) = \varepsilon_{\nu}\psi_{\nu}(\vec{r}) \quad (4.1)$$

Diese Startladungsdichte kann in unserem Programm entweder aus überlagerten Gaußladungsdichten oder aus atomaren Valenzladungsdichten, welche jeweils an den entsprechenden Atomorten lokalisiert sind, konstruiert werden. Nach der Berechnung des effektiven Potentials folgt die Berechnung der Matrixelemente des Hamiltonian in ebener Wellen Darstellung, oder anders ausgedrückt die Berechnung der Hamilton-Matrix im reziproken Raum.

Aus der Hamilton-Matrix gewinnt man durch Diagonalisierung die Einteilchenenergien und Wellenfunktionen im reziproken Raum. Aus den Wellenfunktionen wird wiederum eine neue Ladungsdichte berechnet, aus der man schließlich ein neues effektives Potential bestimmen kann usw.. In der Praxis konvergiert dieses Verfahren jedoch nicht, sodaß entweder die alte und die neue Ladungsdichte, oder alternativ altes und neues Potential, "gemischt" werden müssen (siehe Kapitel 5).

In der traditionellen Bandstrukturmethode werden zur numerischen Bestimmung der Eigenwerte und Eigenvektoren¹ der Hamilton-Matrix Standard-Routinen zur Diagonalisierung beliebiger hermitescher Matrizen verwendet. Diese stehen auf allen Rechnern in Standard-Bibliotheken zur Verfügung (LAPACK, etc.)

¹Die Begriffe Eigenvektoren, Eigenzustände und Wellenfunktionen werden in diesem Kapitel synonym benutzt.

Die traditionellen Methoden haben jedoch den Nachteil, daß sie wegen ihres ungünstigen Skalierungsverhaltens nur für relativ kleine Systeme anwendbar sind. So steigt der benötigte Speicherplatz quadratisch mit der Anzahl $N_{\vec{G}}$ der ebenen Wellen an, da die Hamilton-Matrix explizit aufgebaut werden muß. Die Rechenzeit steigt sogar kubisch mit der Anzahl der ebenen Wellen. Aus diesem Grund sind direkte Verfahren nur für $N_{\vec{G}} \times N_{\vec{G}}$ -Matrizen bis zur Größenordnung $N_{\vec{G}} \sim 1000$ einsetzbar. Der sinnvolle Einsatzbereich für die traditionelle Methode liegt in unserem Programm erfahrungsgemäß bei Matrizen bis zur Größe $N_{\vec{G}} \simeq 400 - 500$.

Um einerseits Systeme mit einigen tausend ebenen Wellen berechnen zu können, andererseits aber sehr wohl die spezielle Struktur des Hamilton-Operators in ebener Wellen Darstellung auszunutzen, bedient man sich iterativer Methoden zur Matrixdiagonalisierung. Die zu Grunde liegende Idee der in dieser Arbeit verwendeten Verfahren soll im folgenden Unterkapitel kurz erläutert werden.

4.2 Iterative Diagonalisierung

Die zentrale Größe der Dichtefunktionaltheorie ist die Ladungsdichte $n(\vec{r})$. Sie wird durch die besetzten Eigenzustände des Hamiltonian aufgebaut. In der vollständig delokalisierten ebene Wellen Basis sind relativ viele Basisfunktionen zur Beschreibung der Wellenfunktionen erforderlich. Tatsächlich werden aber nur sehr wenige Eigenwerte und Eigenfunktionen benötigt. So benötigt man z.B. für die Beschreibung der Valenzelektronen von Si etwa 150 ebene Wellen pro Atom, wohingegen nur die zwei niedrigsten Eigenzustände² pro Atom besetzt werden. Eine vollständige Diagonalisierung des Hamiltonian ist daher nicht notwendig. Man benötigt ein Verfahren, mit dem man auf möglichst schnelle und speicherplatzsparende Art und Weise die untersten 2 – 10% der Eigenwerte ε_ν und Eigenvektoren $|\psi_\nu\rangle$ bestimmen kann. Für diese Problemklasse werden i.a. iterative Methoden zur Matrixdiagonalisierung verwendet.

Die iterativen Verfahren haben folgenden Verlauf:

- "Raten" der Startvektoren $|\psi_\nu\rangle^{(0)}$
- Berechnen der Matrix-Vektor-Produkte $|\mathcal{H}\psi_\nu\rangle^{(j)} = \mathcal{H}|\psi_\nu\rangle^{(j)}$
- Konstruieren von Korrekturvektoren $|\psi_\nu^{korr}\rangle^{(j)}$ aus den $|\mathcal{H}\psi_\nu\rangle^{(j)}$ und Test auf Konvergenz

²in einer nicht spinpolarisierten Rechnung

- Berechnen neuer Startvektoren $|\psi_\nu\rangle^{(j+1)}$ für die nächste Iteration

Die rechenzeitkritischen Schritte dieser Verfahren sind die Matrix-Vektor-Produkte und die Konstruktion der Korrekturvektoren, da hier i.a. eine Orthonormierung der Korrekturvektoren erforderlich ist.

Wie in Kapitel 3 gezeigt wird, zerfällt die Hamiltonmatrix (und damit auch die Matrix-Vektor-Produkte) in drei Teile, den kinetischen Anteil, das lokale und das nichtlokale Potential:

$$\mathcal{H}_{\vec{k}+\vec{G},\vec{k}+\vec{G}'} = (\vec{k} + \vec{G})^2 \delta_{\vec{G}\vec{G}'} + V_{lok}(\vec{G} - \vec{G}') + V_{NL}(\vec{k} + \vec{G}, \vec{k} + \vec{G}') \quad (4.2)$$

Insgesamt sind Matrix-Vektor-Produkte für alle N_{state} Zustände auszuführen. Die einzelnen Beiträge zu jedem Matrix-Vektor-Produkt skalieren zusätzlich wie folgt:

- Die Anwendung des kinetischen Anteils auf eine Wellenfunktion erfordert, da er diagonal im reziproken Raum ist, $N_{\vec{G}}$ Multiplikationen.
- Das lokale Potential V_{lok} ist diagonal im Ortsraum und wird aus diesem Grund dort mit den Wellenfunktionen multipliziert. Der zeitdominierende Schritt ist dabei die FFT des Potentials und der Wellenfunktionen in den Ortsraum und die Rücktransformation der Matrix-Vektor-Produkte in den reziproken Raum. Insgesamt skaliert dieser Teil wie die FFT mit $N_{\vec{G}} \log N_{\vec{G}}$.
- Die Multiplikation des nichtlokalen Pseudopotentials V_{NL} mit den Wellenfunktionen wird im reziproken Raum ausgeführt. Für KB-Potentiale wird die Rechenzeit hierbei proportional zur Anzahl der ebenen Wellen $N_{\vec{G}}$, zur Anzahl der Atome N_{at} pro Einheitszelle und zur Anzahl der Drehimpulskomponenten $(l_{max} + 1)^2$. Für semilokale Pseudopotentiale kommt noch eine zusätzliche Skalierung mit der zur Darstellung des Potentials erforderlichen Anzahl der Gauß-Hermite-Integrationsstützstellen N_{GH} hinzu.

Die Orthonormierung skaliert hingegen mit $N_{state}^2 N_{\vec{G}}$.

Je nach Verhältnis der zu berechnenden Zustände N_{state} zur Anzahl der ebenen Wellen $N_{\vec{G}}$ wird relativ mehr oder weniger Rechenzeit für die Orthonormierung oder für die Berechnung der Matrix-Vektor-Produkte $|\mathcal{H}\psi_\nu\rangle^{(j)} = \mathcal{H}|\psi_\nu\rangle^{(j)}$ benötigt. Mit wachsender Zustandszahl N_{state} dominiert die Orthonormierung, bei vielen ebenen Wellen $N_{\vec{G}}$ dominiert die Berechnung der Matrix-Vektor-Produkte. Dieses unterschiedliche Skalierungsverhalten der Rechenzeit sollte man im Auge behalten, wenn man das zu berechnende System vergrößert.

Eine weitere spezielle Eigenschaft der iterativen Verfahren kommt unserer Problemstellung sehr entgegen. Ausgehend von den "geratenen" Startwellenfunktionen für die verschiedenen Zustände werden die Eigenwerte und Eigenfunktionen sukzessive von Iteration zu Iteration verbessert. Startet man nun mit einer "geratenen" Ladungsdichte, so macht es wenig Sinn, die zu dieser Ladungsdichte gehörigen Eigenfunktionen genau auszukonvergieren, da man i.a. noch weit von der Selbstkonsistenz entfernt ist. Die iterativen Verfahren erlauben es nun, die Wellenfunktionen nur soweit auszukonvergieren, wie es der Güte des Potentials bzw. der Ladungsdichte entspricht. Zusätzlich kann man die approximierten Lösungen der Kohn-Sham-Gleichungen des einen Selbstkonsistenzschritts als gute Startvektoren für die Matrixdiagonalisierung im nächsten Selbstkonsistenzschritt verwenden. Die Genauigkeit der Eigenwertbestimmung und die Selbstkonsistenz der Ladungsdichte werden also sukzessive simultan erreicht. Die Anpassung der Genauigkeit der Matrixdiagonalisierung an die gerade erreichte Selbstkonsistenz erlaubt erhebliche Rechenzeiterparnisse.

Die Einzelheiten des in dieser Arbeit verwendeten Davidson-Kosugi-Algorithmus [Dav75, Kos84, Ric93] werden im nun folgenden Abschnitt erläutert.

4.3 Der Davidson-Kosugi-Algorithmus

Das hermitesche Eigenwertproblem läßt sich auf ein Optimierungsproblem abbilden. Eine zentrale Rolle spielt dabei der Rayleigh-Quotient:

$$\rho[\psi] = \frac{\langle \psi | \mathcal{H} | \psi \rangle}{\langle \psi | \psi \rangle} \quad (4.3)$$

Der Rayleigh-Quotient ist der Erwartungswert des Hamiltonian, i.e. der Erwartungswert der Energie, für den Zustand $|\psi\rangle$ und ist für Eigenlösungen $|\psi_\nu\rangle$ zum Hamiltonoperator \mathcal{H} stationär. Der Rayleigh-Quotient übernimmt somit die Rolle der Zielfunktion des Optimierungsproblems.

Ist man in der Lage den Gradienten einer Zielfunktion zu berechnen, so werden zur Optimierung i.a. Gradientenverfahren verwendet. Der Gradient gibt an einem beliebigen Punkt die Richtung des steilsten Anstiegs an, sodaß man sich auf ein Minimum oder einen Sattelpunkt zubewegt, solange man dem negativen Gradienten folgt. Nach einer kurzen Rechnung erhält man für den Gradienten des Rayleigh-Quotienten:

$$\left| \frac{\delta \rho[\psi]}{\delta \psi} \right\rangle = \frac{2 |\mathcal{R}[\psi]\rangle}{\langle \psi | \psi \rangle} \quad (4.4)$$

$$|\mathcal{R}[\psi]\rangle = (\mathcal{H} - \rho[\psi])|\psi\rangle \quad (4.5)$$

Durch das Residuum $|\mathcal{R}[\psi]\rangle$ ist die Richtung des Gradienten festgelegt und damit die Korrekturrichtung.

Das einfachste Verfahren, immer dem negativen Gradienten zu folgen, um die Zielfunktion zu minimieren, konvergiert sehr schnell für isotrope Minima. Es wird jedoch zunehmend schlechter, je anisotroper das Minimum ist. Anschaulich bedeutet dies für den zweidimensionalen Fall: Das Verfahren konvergiert schlecht in langen, schmalen Tälern. Einen Ausweg bietet die Prädiktionierung:

In der Umgebung eines Eigenvektors $|\psi_{i_0}\rangle$ und Eigenwerts ϵ_{i_0} läßt sich das Residuum linearisieren:

$$|\mathcal{R}[\psi_{i_0} + \delta\psi]\rangle = (\mathcal{H} - \epsilon_{i_0})|\delta\psi\rangle + \mathcal{O}((\delta\psi)^2) \quad (4.6)$$

Wegen Gleichung (4.6) wird das Residuum $|\mathcal{R}[\psi]\rangle$ mit einer geeigneten Approximation der Matrix $(\mathcal{H} - \rho[\psi])^{-1}$, dem "Prädiktionierer", multipliziert, um zu einer besseren Korrekturrichtung $|\delta\psi\rangle$ zu gelangen.

Unser Hamiltonian ist im reziproken Raum diagonal dominant, d.h. mit zunehmendem Abstand zur Diagonalen werden die Matrixelemente kleiner. Dies liegt in erster Linie am quadratischen Anstieg der Diagonalelemente des Hamiltonians $\sim (\vec{k} + \vec{G})^2$, welcher von der kinetischen Energie herrührt. Daher verwendet man die invertierte Diagonale der Matrix $(\mathcal{H} - \rho[\psi])$ als Prädiktionierer. Dieser Prädiktionierer erfordert keinen Mehraufwand an Speicherplatz. Der zusätzliche Rechenzeitbedarf ist vernachlässigbar. Insgesamt ergibt sich somit für die Fourierkoeffizienten des Korrekturvektors die folgende Gleichung:

$$c_{\vec{k}\nu}^{korr}(\vec{G}) = - \sum_{\vec{G}'} \frac{(\rho[\psi_{\vec{k}\nu}] \delta_{\vec{G}\vec{G}'} - \mathcal{H}_{\vec{k},\vec{G}\vec{G}'})}{(\rho[\psi_{\vec{k}\nu}] - \mathcal{H}_{\vec{k},\vec{G}\vec{G}})} c_{\vec{k}\nu}(\vec{G}') \quad (4.7)$$

Der Index für den \vec{k} -Vektor wird aus Gründen der Übersichtlichkeit im folgenden weggelassen, da die Lösung des Eigenwertproblems für jeden \vec{k} -Punkt separat erfolgt.

Hat man einen Korrekturvektor konstruiert, so stellt sich die Frage wie man aus Startvektor und Korrekturvektor zu einer neuen, verbesserten Approximation des Eigenvektors kommt. Die einfachste Methode wäre, den Korrekturvektor auf den Startvektor zu addieren:

$$|\psi_\nu\rangle^{(j+1)} = |\psi_\nu\rangle^{(j)} + \alpha |\psi_\nu^{korr}\rangle^{(j)} \quad (4.8)$$

Dies hat jedoch den Nachteil, daß man hiermit eine willkürliche Schrittweite α einführt.

Eine wesentlich bessere Methode ist es, die Hamiltonmatrix in dem $(2 \times N_{state})$ -dimensionalen Raum zu diagonalisieren, der von den N_{state} Startvektoren und den N_{state} Korrekturvektoren aufgespannt wird. Dies kann einerseits dadurch geschehen, daß man mittels Gram-Schmidt-Verfahren aus den Start- und Korrekturvektoren eine Orthonormalbasis konstruiert, die Matrixelemente des Hamiltonoperators bzgl. dieser Basis berechnet und diese Matrix schließlich diagonalisiert. Andererseits kann der Orthonormierungsschritt auch vermieden werden. Dies führt wegen der fehlenden Orthogonalität der Basisvektoren zu einer Überlappmatrix und somit zu einem verallgemeinerten Eigenwertproblem. In unserem Programm ist die erste Methode implementiert.

Unsere Methode hat ihren Ursprung im Davidson-Algorithmus [Dav75], der jedoch nur für die Berechnung eines einzelnen Eigenwertes und Eigenvektors konzipiert war. Im Davidson-Algorithmus startet man mit einem geratenen Eigenvektor, berechnet den Korrekturvektor und diagonalisiert die Matrix in dem durch diese Vektoren aufgespannten Unterraum. In den folgenden Iterationen wird dem Unterraum sukzessive jeweils ein Korrekturvektor hinzugefügt und die Matrix in dem neuen, größeren Raum diagonalisiert. Spätestens wenn man nach N_G Iterationen den gesamten Vektorraum aufgespannt hat, ist das Eigenwertproblem konvergiert. Die Erfahrung zeigt jedoch, daß die Eigenwerte deutlich schneller konvergieren.

In vielen Problemstellungen benötigt man jedoch nicht nur einen Eigenwert. Aus diesem Grund wurde der Davidson-Algorithmus von Liu [Liu78] zu einem Block-Davidson-Verfahren verallgemeinert. Im Gegensatz zum Davidson-Algorithmus startet man hier mit N_{state} Startvektoren, also für jeden Eigenwert einen. Zusätzlich wird in jeder Iteration für jeden Eigenwert ein Korrekturvektor hinzugefügt. Dies hat zur Folge, daß die Dimension des Unterraums in jeder Iteration um die Zahl N_{state} wächst. Dieses "Aufblähen" des Unterraums führt sehr schnell zu Speicherplatz- und Rechenzeitproblemen.

Einen Ausweg bietet der Kosugi-Algorithmus [Kos84]. Die Größe des Unterraums N_{sub} wird zu Beginn gewählt und anschließend festgehalten. Der Kosugi-Algorithmus ist wie der Liu-Davidson-Algorithmus zur Berechnung von N_{state} Eigenwerten bestimmt. Auch hier startet man mit N_{state} "geratenen" Startvektoren und füllt die N_{sub} -dimensionale Basis mit Einheitsvektoren auf. Während der Iterationen wird zu jedem Startvektor jeweils ein Korrekturvektor berechnet. Somit werden die Einheitsvektoren nach und nach durch Korrekturvektoren ersetzt. Sind alle Einheitsvektoren ersetzt, werden die jeweils ältesten Korrekturvektoren ausgetauscht.

Der in unserem Programm implementierte Kosugi-Algorithmus hat somit folgenden Verlauf:

1. Initialisierung:

- (a) Berechne im ersten elektronischen Selbstkonsistenzschritt die Startwellenfunktionen $|\psi_\nu\rangle^{(0)}$ und Starteigenwerte $\varepsilon_\nu^{(0)}$ durch Diagonalisierung einer niedrigdimensionalen Unterraummatrix des Hamiltonian \mathcal{H} . In unserem Fall sind dies die niedrigsten Fourierkomponenten $\mathcal{H}_{\vec{G}\vec{G}}$, des Hamiltonian – typischerweise $N_{ini} \sim 500$.

Sonst nehme die Wellenfunktionen $|\psi_\nu\rangle$ aus dem letzten Selbstkonsistenzschritt und berechne die Starteigenwerte³ $\varepsilon_\nu^{(0)} = \langle \psi_\nu | \mathcal{H} | \psi_\nu \rangle$

- (b) Besetze die $N_{korr} = N_{sub} - N_{state}$ Korrekturvektoren $|\psi_\mu^{korr}\rangle$ im reziproken Raum mit Einheitsvektoren $(\Psi_\mu^{korr}(\vec{G})) = \delta_{\vec{G}, \vec{G}_{N_{state}+\mu}}$.

2. Iteration:

Die N_{state} Eigenvektoren $|\psi_\nu\rangle^{(j)}$ und die N_{korr} Korrekturvektoren $|\psi_\mu^{korr}\rangle$ bilden eine Orthonormalbasis $\{|e_m\rangle, m = 1 \dots N_{sub}\}$ eines N_{sub} -dimensionalen Unterraum des reziproken Raumes.

- (a) Berechne die Residuen $|\mathcal{R}[\psi_\nu^{(j)}]\rangle = (\mathcal{H} - \varepsilon_\nu^{(j)}) |\psi_\nu\rangle^{(j)}$ und teste auf Konvergenz. Berechne dazu die Norm der Residuen $\|\mathcal{R}[\psi_\nu^{(j)}]\|$. Wenn die jeweilige Norm unter einer vorgegebenen Schranke liegt, ist der zugehörige Eigenvektor konvergiert. Markiere die konvergierten Wellenfunktionen. Falls alle Wellenfunktionen konvergiert sind, fahre fort mit (2e).
- (b) Berechne zu den noch nicht konvergierten Zuständen die neuen Korrekturvektoren $(\mathcal{D}_{\mathcal{H}} - \varepsilon_\nu^{(j)})^{-1} |\mathcal{R}[\psi_\nu^{(j)}]\rangle$, wobei $\mathcal{D}_{\mathcal{H}}$ die Diagonale der Hamiltonmatrix \mathcal{H} ist. Ersetze die ältesten Korrekturvektoren $|\psi_\mu^{korr}\rangle$ durch die neuen.
- (c) Konstruiere aus den Eigenvektoren $|\psi_\nu\rangle^{(j)}$ und den aktuellen Korrekturvektoren $|\psi_\mu^{korr}\rangle$ eine neue Orthonormalbasis $\{|e_m\rangle, m = 1 \dots N_{sub}\}$. Dazu müssen die neuen Korrekturvektoren auf die alten Eigenvektoren $|\psi_\nu\rangle^{(j)}$ und auf die nicht in (2b) ersetzten Korrekturvektoren orthonormiert werden.
- (d) Diagonalisierung die Unterraummatrix $\langle e_n | \mathcal{H} | e_m \rangle$:

$$\sum_{m=1}^{N_{sub}} \langle e_n | \mathcal{H} | e_m \rangle C_{mn} = \varepsilon_n C_{mn}$$

³Die Eigenwerte aus dem letzten Selbstkonsistenzschritt werden nicht genommen, da sich der Hamiltonian \mathcal{H} von Selbstkonsistenzschritt zu Selbstkonsistenzschritt ändert.

Die Koeffizienten bzw. Eigenvektoren $C_{m\nu}$ werden zur Bestimmung der neuen approximierten Eigenvektoren $|\psi_\nu\rangle^{(j+1)}$ benutzt:

$$|\psi_\nu\rangle^{(j+1)} = \sum_{m=1}^{N_{sub}} C_{m\nu} |e_m\rangle, \quad \nu = 1..N_{state}$$

Falls die vorgegebene maximale Anzahl an Iterationen noch nicht erreicht ist, fahre fort mit (2a).

- (e) Normiere die Wellenfunktionen $|\psi_\nu\rangle$, um Ladungsneutralität sicherzustellen.

Für den Kosugi-Algorithmus gibt es in unserem Programm im wesentlichen zwei Einsatzbereiche: Zum einen die Diagonalisierung des Hamiltonians innerhalb der Selbstkonsistenzschleife, und zum anderen die Bestimmung der Eigenwerte und Eigenfunktionen zu einer vorgegebenen festen Ladungsdichte im Rahmen von Bandstrukturrechnungen. Je nach Verwendungszweck haben sich unterschiedliche Strategien für den Kosugi-Algorithmus bewährt:

Innerhalb des Selbstkonsistenzzyklus hat sich $2 \times N_{state}$ bis $3 \times N_{state}$ als optimale Unterraumgröße N_{sub} bewährt. Drei bis fünf Davidson-Iterationen pro Selbstkonsistenzschritt haben sich als sinnvoll herausgestellt, um Konvergenz der Eigenwertbestimmung und Selbstkonsistenz der Ladungsdichte simultan zu erreichen. Diese Parameterwahl liefert i.a. eine – bezogen auf die Rechenzeit – schnelle Konvergenz. Bei der Berechnung der Eigenwerte zu einem fest vorgegebenen Potential ist man an einer schnellen numerischen Konvergenz der Eigenwerte, d.h. an einer Konvergenz mit möglichst wenig Davidson-Iterationen, interessiert. In diesem Fall muß das Eigenwertproblem einmal sehr genau gelöst werden. Möchte man also die Eigenwerte zu einer festen Matrix möglichst genau bestimmen, so ist es vorteilhaft den Unterraum für den Kosugi-Algorithmus möglichst groß zu wählen. Mit Rücksicht auf die wachsenden Speicherplatzanforderungen und die mit der Unterraumdimension kubisch anwachsende Rechenzeit hat sich als Unterraumgröße N_{sub} ein Wert von ca. $2 \times N_{state}$ bis $4 \times N_{state}$ bewährt.

Kapitel 5

Minimierung des Energiefunktional

Eine sehr allgemeine und immer wiederkehrende Problemstellung in der Physik ist die Minimierung eines Funktionals, in unserem Fall die Minimierung des DFT-Energiefunktional in der LDA-Näherung. Das Funktional der totalen Energie hat zwei grundsätzlich verschiedene Arten von Freiheitsgraden:

- die elektronischen Freiheitsgrade, i.e. die Wellenfunktionen $\Psi_{\vec{k}\nu}(\vec{r})$
- die ionischen Freiheitsgrade, i.e. die Atomkoordinaten $\vec{\tau}_\mu$

Wir suchen das Energieminimum in zwei Schritten:

1. Zuerst werden die elektronischen Freiheitsgrade relaxiert, d.h. das Energiefunktional wird bei festen Atompositionen $\vec{\tau}_\mu$ bzgl. der Wellenfunktionen $\Psi_{\vec{k}\nu}(\vec{r})$ minimiert:

$$E_0[\vec{\tau}_\mu] = \min_{\{\Psi_{\vec{k}\nu}(\vec{r})\}} E_{LDA}[\vec{\tau}_\mu, \Psi_{\vec{k}\nu}(\vec{r})] \quad (5.1)$$

Dies ist gleichbedeutend mit der Bestimmung der LDA-Grundzustandsenergie durch selbstkonsistente Berechnung der Grundzustandsladungsdichte $n_0(\vec{r})$ bzw. des Grundzustandspotentials $V_0(\vec{r}) = V[n_0(\vec{r})]$.

Der im Computerprogramm implementierte Algorithmus zur Lösung der Schrödingergleichung definiert eine Abbildung $n(\vec{r}) \rightarrow F\{n(\vec{r})\}$ wie folgt:

$$\begin{pmatrix} n(\vec{r}) \\ V[n(\vec{r})] \end{pmatrix} \xrightarrow{(\mathcal{H}\Psi = \epsilon\Psi)} \{\Psi_{\vec{k}\nu}(\vec{r}), \epsilon_{\vec{k}\nu}\} \longrightarrow \begin{pmatrix} F\{n(\vec{r})\} := \sum_{\vec{k}\nu} \omega_{\vec{k}} f_{\vec{k}\nu} |\Psi_{\vec{k}\nu}(\vec{r})|^2 \\ \hat{F}\{V[n(\vec{r})]\} := V[F\{n(\vec{r})\}] \end{pmatrix} \quad (5.2)$$

$F\{n(\vec{r})\}$ ist die Ladungsdichte, die von den Eigenfunktionen $\{\Psi_{\vec{k}\nu}(\vec{r}), \varepsilon_{\vec{k}\nu}\}$ zum Potential $V[n(\vec{r})]$ aufgebaut wird. $\hat{F}\{V[n(\vec{r})]\}$ ist das zu dieser Ladungsdichte gehörende Potential. Die Grundzustandsdichte $n_0(\vec{r})$ bzw. das Grundzustandspotential $V_0(\vec{r})$ sind dann Fixpunkte der Abbildung $F\{n\}$ bzw. $\hat{F}\{V\}$, die es zu bestimmen gilt. Die Fixpunkte sind auch Nullstellen der Funktionale $\mathcal{F}\{n\}$ bzw. $\hat{\mathcal{F}}\{V\}$:

$$\begin{aligned}\mathcal{F}\{n_0\} &= F\{n_0\} - n_0 = 0 \\ \hat{\mathcal{F}}\{V[n_0]\} &= \hat{F}\{V[n_0]\} - V[n_0] = 0\end{aligned}\quad (5.3)$$

2. Im zweiten Schritt wird eine atomare Strukturrelaxation durchgeführt, d.h. die LDA-Grundzustandsenergie wird bzgl. der Atomkoordinaten $\vec{\tau}_\mu$ minimiert:

$$E_{tot} = \min_{\{\vec{\tau}_\mu\}} E_0[\vec{\tau}_\mu] \quad (5.4)$$

Das Auffinden von Extremwerten bzw. Sattelpunkten des Energiefunktionals bezüglich der Atomkoordinaten $\vec{\tau}_\mu$ ist gleichbedeutend mit dem Auffinden von Nullstellen der Hellman-Feynman-Kräfte:

$$\vec{\mathcal{F}}\{\vec{\tau}\} = -\frac{d}{d\vec{\tau}} E_0[\vec{\tau}] \quad (5.5)$$

$$\vec{\mathcal{F}}\{\vec{\tau}_0\} = \vec{0} \quad (5.6)$$

$\vec{\tau}$ ist eine Abkürzung für die Gesamtheit aller Atomkoordinaten $\{\vec{\tau}_1, \dots, \vec{\tau}_\mu, \dots\}$, $\vec{\tau}_0$ ist der Vektor $\vec{\tau}$, der zum Energieminimum führt, d.h. der Vektor der Gleichgewichtskoordinaten.

$n(\vec{r})$ bzw. $V(\vec{r})$ sind im Computerprogramm auf endlich vielen Stützpunkten definiert. So definiert Gleichung (5.3) ein $N_{\vec{r}}$ -dimensionales nichtlineares Gleichungssystem, dessen Nullstellen es zu bestimmen gilt. Gleichung (5.6) definiert ein nichtlineares Gleichungssystem der Dimension $N_{\vec{\tau}}$, wobei $N_{\vec{\tau}}$ die Anzahl der ionischen Freiheitsgrade, i.e. der Atomkoordinaten, bedeutet. Es sind ebenfalls die Nullstellen dieses Gleichungssystems zu bestimmen.

Von diesem Standpunkt aus betrachtet, sind die Bestimmung des elektronischen Grundzustands $n_0(\vec{r})$ bzw. $V_0(\vec{r})$ und die Bestimmung der atomaren Gleichgewichtsstruktur $\vec{\tau}_0$ gleichartige Probleme. Es sind jeweils die Nullstellen $n_0(\vec{r})$ oder $V_0(\vec{r})$, bzw. $\vec{\tau}_0$ einer hochdimensionalen nichtlinearen Abbildung $\mathcal{F}\{n\}$, $\hat{\mathcal{F}}\{V\}$ bzw. $\vec{\mathcal{F}}\{\vec{\tau}\}$ zu bestimmen. Funktional und Fixpunkt werden im folgenden durch $\mathcal{F}\{x\}$, $F\{x\}$ und x_0 abgekürzt, die Dimension des Gleichungssystems mit N_x .

$$\mathcal{F}\{x\} = F\{x\} - x \quad (5.7)$$

$$\mathcal{F}\{x_0\} = F\{x_0\} - x_0 = 0 \quad (5.8)$$

Die Nullstellen hochdimensionaler nichtlinearer Gleichungssysteme werden in der Regel iterativ bestimmt. Das einfachste Verfahren, stabil aber langsam, ist das sogenannte "*simple mixing*". Es konvergiert linear.

Ist man in der Lage, die Jakobi-Matrix $\mathcal{J}\{x\} = \frac{d}{dx}\mathcal{F}\{x\}$, bzw. die inverse Jakobi-Matrix $[\mathcal{J}\{x\}]^{-1}$ zu berechnen, so bietet sich die Methode nach Newton-Raphson an. Der größte Vorteil dieser Methode ist ihre quadratische Konvergenz. Nachteilig ist der geringe Konvergenzradius und die Tatsache, daß der Jacobian \mathcal{J} invertiert werden muß. Der dafür benötigte numerische Aufwand skaliert mit N_x^3 .

In unserem Fall ist man nicht in der Lage, den Jacobian bzw. den inversen Jacobian mit vertretbarem numerischen Aufwand zu berechnen. Im Fall der elektronischen Selbstkonsistenz ist die Abbildung \mathcal{F} nicht bekannt. Ein numerisches Ausrechnen bzw. Invertieren der Jacobi-Matrix erübrigt sich schon wegen der Hochdimensionalität ($N_{\mathcal{G}} \sim 10^4$) des Jacobian. Im Fall der atomaren Strukturrelaxation ($N_{\mathcal{T}} \sim 10^2$) wäre es zwar prinzipiell möglich, den Jacobian zu invertieren, aber der numerische Aufwand für die Berechnung der einzelnen Matrixelemente ist extrem hoch. Dies ist im allgemeinen nicht mehr tragbar. Daher bieten sich hier Quasi-Newton-Verfahren an, welche die Nullstelle x_0 und den inversen Jacobian \mathcal{J}^{-1} simultan iterativ bestimmen. Auch Quasi-Newton-Verfahren konvergieren wie die Newton-Raphson-Methode superlinear, haben jedoch im allgemeinen einen größeren Konvergenzradius und vermeiden die Invertierung des Jacobian \mathcal{J} .

Lineare, superlineare und quadratische Konvergenz ist wie folgt definiert:

- lineare Konvergenz:
$$\frac{|\delta x^{(m+1)}|}{|\delta x^{(m)}|} \leq \text{const.} \quad (5.9)$$

- superlineare Konvergenz:
$$\frac{|\delta x^{(m+1)}|}{|\delta x^{(m)}|} \rightarrow 0 \quad (5.10)$$

- quadratische Konvergenz:
$$\frac{|\delta x^{(m+1)}|}{|\delta x^{(m)}|^2} \leq \text{const.} \quad (5.11)$$

Dabei ist $\delta x^{(m)} = x^{(m)} - x_0$ der Abstand der (m)-ten Iteration $x^{(m)}$ vom Fixpunkt x_0 .

Einen guten Überblick über verschiedene Optimierungsmethoden findet sich in [Fle87].

5.1 "*simple mixing*"

Die Abbildung $\mathcal{F}\{x\}$ ist so etwas wie eine treibende Kraft in Richtung auf die Nullstelle x_0 . Damit bietet sich zwanglos ein einfaches iteratives Verfahren zur Nullstel-

lenbestimmung an, indem man den Vektor $x^{(m)}$ in Krafrichtung $\mathcal{F}\{x^{(m)}\}$ korrigiert:

$$x^{(m+1)} = x^{(m)} + \mathcal{F}\{x^{(m)}\} = F\{x^{(m)}\} \quad (5.12)$$

Dieses Verfahren ist aber in der Regel nicht konvergent, kann aber sehr leicht konvergent gemacht werden:

$$x^{(m+1)} = x^{(m)} + \alpha \mathcal{F}\{x^{(m)}\} = (1 - \alpha)x^{(m)} + \alpha F\{x^{(m)}\} \quad (5.13)$$

Diese Methode bezeichnet man als “*simple mixing*” und ist für hinreichend kleine Mixingparameter α sicher konvergent. Allerdings konvergiert diese Methode nur linear. Aus dem zweiten Teil von Gleichung (5.13) geht hervor, daß mit dem Mixingparameter α der alte Vektor $x^{(m)}$ und der neue Vektor $F\{x^{(m)}\}$ gemischt werden. Aus dem ersten Teil der Gleichung wird klar, daß es sich bei α um eine Art “*line-search*”-Parameter längst der Krafrichtung $\mathcal{F}\{x^{(m)}\}$ handelt. Beide Sichtweisen sind äquivalent.

Im Fall der elektronischen Selbstkonsistenz, i.e. “Ladungsdichtemischen” oder “Potentialmischen”, sind nur Werte von $(0 < \alpha \leq 1)$ sinnvoll. Im Fall der atomaren Strukturrelaxation ist α eine Schrittweite längst der Krafrichtung \mathcal{F} . Dort sind auch andere Werte sinnvoll.

5.2 Die Newton–Raphson–Methode

Kennt man den Jacobian $\mathcal{J}\{x\}$, so kann man die Abbildung $\mathcal{F}\{x\}$ um den Vektor der (m) -ten Iteration $x^{(m)}$ linearisieren. Ein vernünftiger Ansatz für den Vektor der nächsten Iteration $x^{(m+1)}$ ergibt sich aus der Forderung nach Verschwinden der linearisierten treibenden Kraft $\mathcal{F}\{x^{(m+1)}\}$ an der Stelle $x^{(m+1)}$:

$$\mathcal{F}\{x^{(m+1)}\} \approx \mathcal{F}\{x^{(m)}\} + \mathcal{J}\{x^{(m)}\} (x^{(m+1)} - x^{(m)}) \stackrel{!}{=} \vec{0} \quad (5.14)$$

$$\Rightarrow x^{(m+1)} = x^{(m)} - [\mathcal{J}\{x^{(m)}\}]^{-1} \mathcal{F}\{x^{(m)}\} \quad (5.15)$$

Die Newton–Raphson–Methode konvergiert quadratisch, jedoch mit sehr kleinem Konvergenzradius. Sie ist nicht anwendbar, wenn der inverse Jacobian $[\mathcal{J}\{x\}]^{-1}$ nur mit hohem numerischen Aufwand oder auch gar nicht auszurechnen ist.

5.3 Quasi–Newton–Methoden

Quasi–Newton–Methoden unterscheiden sich von der Newton–Raphson–Methode dadurch, daß zusätzlich zur Nullstelle auch der inverse Jacobian iterativ bestimmt

wird. Aus der Linearisierung der Abbildung $\mathcal{F}\{x\}$ um den Vektor der (m) -ten Iteration $x^{(m)}$ ergibt sich für die iterative Bestimmung des Jacobian folgende Bedingung, die auch Quasi-Newton-Bedingung genannt wird:

$$\mathcal{F}\{x^{(m-1)}\} \approx \mathcal{F}\{x^{(m)}\} + \mathcal{J}\{x^{(m)}\} (x^{(m-1)} - x^{(m)}) \quad (5.16)$$

$$\Rightarrow \Delta x^{(m)} = [\mathcal{J}\{x^{(m)}\}]^{-1} \Delta \mathcal{F}^{(m)} \quad (5.17)$$

$$\Delta x^{(m)} = x^{(m)} - x^{(m-1)}, \quad \Delta \mathcal{F}^{(m)} = \mathcal{F}\{x^{(m)}\} - \mathcal{F}\{x^{(m-1)}\}$$

Die Newton-Raphson-Gleichungen (5.15) und die Quasi-Newton-Bedingung (5.17) für Nullstelle und inversen Jacobian sind simultan iterativ zu bestimmen:

- $[\mathcal{J}^{(m)}]^{-1}$ erfüllt: $\Delta x^{(m)} = [\mathcal{J}^{(m)}]^{-1} \Delta \mathcal{F}^{(m)} \quad (5.18)$

- nächster Lösungsvektor: $x^{(m+1)} = x^{(m)} + \Delta x^{(m+1)} \quad (5.19)$

$$\Delta x^{(m+1)} = -[\mathcal{J}^{(m)}]^{-1} \mathcal{F}\{x^{(m)}\} \quad (5.20)$$

Quasi-Newton-Verfahren konvergieren superlinear und haben in der Regel einen größeren Konvergenzradius als die Newton-Raphson-Methode. Konvergenzgebiet ist das Gebiet um eine Extremwertstelle oder einen Sattelpunkt herum, in dem die Eigenwerte von \mathcal{J} , i.e. die Krümmung des zu minimierenden Energiefunktionals, ihr Vorzeichen nicht wechseln. Ist die Matrix \mathcal{J} positiv definit, liegt ein Energiemaximum vor, ist \mathcal{J} negativ definit, ein Minimum. Wechselt ein Eigenwert genau an der Konvergenzstelle sein Vorzeichen, so handelt es sich um einen Sattelpunkt. Das Verfahren konvergiert auch gegen Sattelpunkte.

Die Quasi-Newton-Bedingung (5.17) bestimmt den inversen Jacobian nicht eindeutig. Deshalb ist es üblich, einen separablen Ansatz für den inversen Jacobian zu machen, der die Quasi-Newton-Bedingung (5.17) erfüllen muß. Damit ergibt sich folgende Iterationsvorschrift:

- nächster Lösungsvektor: $x^{(m+1)} = x^{(m)} + \Delta x^{(m+1)} \quad (5.21)$

- Newton-Raphson-Term: $\Delta x^{(m+1)} = -[\mathcal{J}^{(m)}]^{-1} \mathcal{F}\{x^{(m)}\} \quad (5.22)$

- nächster Jacobian: $[\mathcal{J}^{(m+1)}]^{-1} = [\mathcal{J}^{(m)}]^{-1} + \Delta [\mathcal{J}^{(m+1)}]^{-1} \quad (5.23)$

- Start Jacobian: $[\mathcal{J}^{(1)}]^{-1} = -\alpha 1 \quad (5.24)$

- separabler Term: $\Delta [\mathcal{J}^{(m)}]^{-1} = |u^{(m)}\rangle \langle v^{(m)}| + \dots \quad (5.25)$

Die verschiedenen Quasi-Newton-Verfahren unterscheiden sich nur durch verschiedene Ansätze für den inversen Jacobian. Je nach Anzahl der separablen Korrekturterme unterscheidet man Rang-1- und Rang-2-Verfahren.

Zur Lösung des elektronischen Selbstkonsistenzproblems sind drei verschiedene Verfahren implementiert, in denen ein im allgemeinen nichtsymmetrischer inverser Jacobian aufgebaut wird. Dies sind zwei Rang 1 Verfahren, die auf Vorschläge von Broyden zurückgehen, Broyden erste Formel [Bro65, Blü88] und Broydens zweite Formel [Bro66, Blü88], sowie ein verallgemeinertes Anderson-Verfahren [And64, Blü]. In der Regel wird das verallgemeinerte Anderson-Verfahren benutzt.

Der Jacobian der atomaren Strukturrelaxation ist symmetrisch, da es sich ja bis auf Vorzeichen um die Matrix der zweiten Ableitungen des Energiefunktional nach den Atomkoordinaten handelt. Daher sind für die Strukturrelaxation zwei verschiedene symmetrische Rang 2 Ansätze für den Jacobian implementiert, der DFP-Jacobian von Davidson, Fletcher und Powell und der BFGS-Jacobian von Broyden, Fletcher, Goldfarb und Shanno [Jac77, Fle87, Blü88]. In der Regel wird der BFGS-Jacobian benutzt.

Im folgenden sind die separablen Ausdrücke für die verschiedenen implementierten Jacobian aufgeführt. Dabei wird folgende Abkürzung verwendet:

$$|\Delta\mathcal{F}^{(i)}\rangle = |\mathcal{F}\{x^{(i)}\} - \mathcal{F}\{x^{(i-1)}\}\rangle \quad (5.26)$$

$$|\Delta x^{(i)}\rangle = |x^{(i)} - x^{(i-1)}\rangle \quad (5.27)$$

5.3.1 Broydens erste Formel

$$[\mathcal{J}^{(m)}]^{-1} = -\alpha 1 + \sum_{i=2}^m |u^{(i)}\rangle\langle v^{(i)}| \quad (5.28)$$

$$|u^{(i)}\rangle = |\Delta x^{(i)}\rangle - [\mathcal{J}^{(i-1)}]^{-1} |\Delta\mathcal{F}^{(i)}\rangle \quad (5.29)$$

$$= |\Delta x^{(i)} + \alpha\Delta\mathcal{F}^{(i)}\rangle - \sum_{j=2}^{i-1} |u^{(j)}\rangle\langle v^{(j)}| \Delta\mathcal{F}^{(i)}\rangle$$

$$|v^{(i)}\rangle = \frac{[\mathcal{J}^{(i-1)}]^{-1} |\Delta x^{(i)}\rangle}{\langle \Delta\mathcal{F}^{(i)} | [\mathcal{J}^{(i-1)}]^{-1} | \Delta x^{(i)} \rangle} \quad (5.30)$$

$$= \frac{\alpha|\Delta x^{(i)}\rangle - \sum_{j=2}^{i-1} |u^{(j)}\rangle\langle v^{(j)}| \Delta x^{(i)}\rangle}{\|\Delta x^{(i)}\|^2 - \langle u^{(i)} | \Delta x^{(i)} \rangle}$$

5.3.2 Broydens zweite Formel

$$[\mathcal{J}^{(m)}]^{-1} = -\alpha 1 + \sum_{i=2}^m |u^{(i)}\rangle\langle v^{(i)}| \quad (5.31)$$

$$\begin{aligned} |u^{(i)}\rangle &= |\Delta x^{(i)}\rangle - [\mathcal{J}^{(i-1)}]^{-1} |\Delta \mathcal{F}^{(i)}\rangle \\ &= |\Delta x^{(i)} + \alpha \Delta \mathcal{F}^{(i)}\rangle - \sum_{j=2}^{i-1} |u^{(j)}\rangle \langle v^{(j)} | \Delta \mathcal{F}^{(i)} \rangle \end{aligned} \quad (5.32)$$

$$|v^{(i)}\rangle = \frac{|\Delta \mathcal{F}^{(i)}\rangle}{\|\Delta \mathcal{F}^{(i)}\|^2} \quad (5.33)$$

5.3.3 Der verallgemeinerte Anderson–Jacobian

$$[\mathcal{J}^{(m)}]^{-1} = -\alpha 1 + \sum_{i=2}^m |u^{(i)}\rangle \langle v^{(i)}| \quad (5.34)$$

$$\begin{aligned} |u^{(i)}\rangle &= |\Delta x^{(i)}\rangle - [\mathcal{J}^{(i-1)}]^{-1} |\Delta \mathcal{F}^{(i)}\rangle \\ &= |\Delta x^{(i)} + \alpha \Delta \mathcal{F}^{(i)}\rangle - \sum_{j=2}^{i-1} |u^{(j)}\rangle \langle v^{(j)} | \Delta \mathcal{F}^{(i)} \rangle \end{aligned} \quad (5.35)$$

$$|v^{(i)}\rangle = \frac{|w^{(i)}\rangle}{\langle w^{(i)} | \Delta \mathcal{F}^{(i)} \rangle} \quad (5.36)$$

$$|w^{(i)}\rangle = |\Delta \mathcal{F}^{(i)}\rangle - \sum_{j=2}^{i-1} |w^{(j)}\rangle \langle v^{(j)} | \Delta \mathcal{F}^{(i)} \rangle \quad (5.37)$$

5.3.4 Der DFP–Jacobian

$$[\mathcal{J}^{(m)}]^{-1} = -\alpha 1 + \sum_{i=2}^m \{ |u^{(i)}\rangle \langle v^{(i)}| - |w^{(i)}\rangle \langle y^{(i)}| \} \quad (5.38)$$

$$|u^{(i)}\rangle = |\Delta x^{(i)}\rangle \quad (5.39)$$

$$|v^{(i)}\rangle = \frac{|\Delta x^{(i)}\rangle}{\langle \Delta \mathcal{F}^{(i)} | \Delta x^{(i)} \rangle} \quad (5.40)$$

$$|w^{(i)}\rangle = [\mathcal{J}^{(i-1)}]^{-1} |\Delta \mathcal{F}^{(i)}\rangle \quad (5.41)$$

$$= -\alpha |\Delta \mathcal{F}^{(i)}\rangle + \sum_{j=2}^{i-1} |u^{(j)}\rangle \langle v^{(j)} | \Delta \mathcal{F}^{(i)} \rangle - \sum_{j=2}^{i-1} |w^{(j)}\rangle \langle y^{(j)} | \Delta \mathcal{F}^{(i)} \rangle$$

$$|y^{(i)}\rangle = \frac{|w^{(i)}\rangle}{\langle \Delta \mathcal{F}^{(i)} | w^{(i)} \rangle} \quad (5.42)$$

5.3.5 Der BFGS–Jacobian

$$[\mathcal{J}^{(m)}]^{-1} = -\alpha 1 + \sum_{i=2}^m \{ |u^{(i)}\rangle \langle v^{(i)}| - |y^{(i)}\rangle \langle w^{(i)}| - |w^{(i)}\rangle \langle y^{(i)}| \} \quad (5.43)$$

$$|u^{(i)}\rangle = |\Delta x^{(i)}\rangle \quad (5.44)$$

$$|y^{(i)}\rangle = \frac{|\Delta x^{(i)}\rangle}{\langle \Delta x^{(i)} | \Delta \mathcal{F}^{(i)} \rangle} \quad (5.45)$$

$$|w^{(i)}\rangle = [\mathcal{J}^{(i-1)}]^{-1} |\Delta \mathcal{F}^{(i)}\rangle \quad (5.46)$$

$$\begin{aligned} &= -\alpha |\Delta \mathcal{F}^{(i)}\rangle + \sum_{j=2}^{i-1} |u^{(j)}\rangle \langle v^{(j)} | \Delta \mathcal{F}^{(i)} \rangle \\ &\quad - \sum_{j=2}^{i-1} |y^{(j)}\rangle \langle w^{(j)} | \Delta \mathcal{F}^{(i)} \rangle - \sum_{j=2}^{i-1} |w^{(j)}\rangle \langle y^{(j)} | \Delta \mathcal{F}^{(i)} \rangle \\ |v^{(i)}\rangle &= |y^{(i)}\rangle \left(1 + \frac{\langle \Delta \mathcal{F}^{(i)} | w^{(i)} \rangle}{\langle \Delta x^{(i)} | \Delta \mathcal{F}^{(i)} \rangle} \right) \end{aligned} \quad (5.47)$$

5.4 Die Relaxation der elektronischen Struktur

Im Computerprogramm wird das Mischen von alter und neuer Ladungsdichte im Impulsraum durchgeführt, das Mischen der Potentiale im Ortsraum. Bei kleinen Zellen ist im allgemeinen kein Unterschied zwischen Ladungsdichtemischen und Potentialmischen festzustellen. Beide Verfahren konvergieren problemlos. Die elektronische Selbstkonsistenz wird durch den mittleren quadratischen "RMS-Fehler"¹ in Ladungsdichte bzw. Potential beschrieben. Nach einigen wenigen (2 bis 4) Startiterationen konvergiert das Verfahren superlinear. Ca. alle zwei Iterationen sinkt dann der "RMS-Fehler" um etwa eine Größenordnung. Die Wahl des Mischungsparameters ist unkritisch und meist kann $\alpha \sim 0.5$ gewählt werden.

Größere Zellen mit abrupten Übergängen wie Oberflächen oder Grenzflächen konvergieren manchmal etwas schwieriger. Im allgemeinen konvergiert dann das Potentialmischen problemloser als das Ladungsdichtemischen. Konvergenzschwierigkeiten können dann auftreten, wenn das elektronische System "zu stark" reagiert. Startet man zum Beispiel bei der Berechnung einer Oberflächenstruktur mit der unrelaxierten Oberfläche, so hat die Ladungsdichte die Tendenz, sich nach außen ins Vakuum hinein auszudehnen. Kommt es hier zu einer "Überreaktion", so zieht sich die Ladungsdichte im nächsten Selbstkonsistenzschritt wieder in den Festkörper zurück. Dies kann zu einem oszillatorischen Verhalten der Ladungsdichte führen und die Konvergenz stark beeinträchtigen. Treten Konvergenzschwierigkeiten dieser Art auf, so muß das Verfahren gedämpft werden. Dabei gibt es im wesentlichen drei Möglichkeiten, die elektronische Selbstkonsistenz zu beschleunigen:

¹RMS: "root mean square"

- Wahl von kleineren Mischungsparametern α : Bei kleineren Mischungsparametern konvergieren die Quasi-Newton-Verfahren allerdings langsamer. Wird der Mischungsparameter zu klein, dann konvergiert das Verfahren praktisch linear. Werte für α unter 0.1 sind für halbleitende Systeme in der Regel unvernünftig.
- Weniger Davidson-Schritte pro Mischungsschritt: Bei der iterativen Diagonalisierung des Hamiltonian bestimmt die Anzahl der Davidson-Schritte die Genauigkeit, mit der Eigenwerte und Eigenzustände ermittelt werden. Weit weg von der Born-Oppenheimer-Oberfläche ist es sinnlos, Eigenwerte und Eigenzustände sehr genau zu ermitteln. Deshalb werden in der Regel drei bis vier Davidson-Schritte pro Mischungsschritt durchgeführt. Konvergenz der Eigenwert- und Eigenzustandsbestimmung und Selbstkonsistenz des Potentials werden so simultan erreicht. Senkt man die Zahl der Davidson-Schritte, so verkleinert man den Abstand zwischen zwei zu mischenden Potentialen. Dies führt effektiv zu einer Dämpfung und kann Konvergenz erzwingen.
- Wahl einer größeren Fermi-Temperatur: Auch Systeme, die halbleitend sind, d.h. eine Bandlücke haben, können auf dem Weg zum Grundzustand, der bei iterativen Verfahren eingeschlagen wird, leitend werden. So sind viele unrelaxierte Halbleiteroberflächen leitend und erst die Strukturrelaxation öffnet die Bandlücke und macht das System zum Halbleiter. Auch bei den "richtigen" Atomkoordinaten kann ein System weit weg von der Born-Oppenheimer-Oberfläche leitend werden, d.h. keine Bandlücke besitzen. Ursache für Konvergenzschwierigkeiten sind dann Zustände im Bereich der Fermienergie, die ihre energetische Reihenfolge von Iteration zu Iteration vertauschen können. So tragen von Iteration zu Iteration andere Zustände zum Aufbau der neuen Ladungsdichte bei. Dies führt zu Ladungsdichteoszillationen und kann die Konvergenz des Verfahrens zerstören. Durch Wahl einer größeren Fermitemperatur wird der Aufbau der Ladungsdichte um die Fermikante verschmiert. Dies führt zu einer effektiven Dämpfung und kann ebenfalls Konvergenz erzwingen.

Bis hierher wurde nur die Geschwindigkeit der elektronischen Selbstkonsistenz diskutiert. Auch bei optimal schneller Konvergenz konvergiert das Verfahren jedoch nicht beliebig genau, sondern nur bis zu einer Genauigkeitsgrenze, die von der Wahl der Konvergenzparameter abhängt und von System zu System unterschiedlich ist. Die maximal erzielbare elektronische Selbstkonsistenz hängt davon ab, wie exakt das elektronische Eigenwertproblem gelöst wird, i.e. von der Genauigkeit der iterativen Matrixdiagonalisierung. Beim Davidson-Kosugi-Algorithmus steuern im

wesentlichen zwei Parameter die Genauigkeit. Dies ist zum einen die Anzahl der Davidsonschröte, zum anderen die Davidsondimension, d.h. die Dimension des Unterraumes, in dem das Eigenwertproblem exakt gelöst wird. Eine sinnvolle untere Schranke für die Davidsondimension ist durch zwei mal die Anzahl der elektronischen Zustände gegeben. Dies ist die Unterraumdimension, mit der das Verfahren weit weg vom elektronischen Minimum gestartet wird. Es kann nun der Fall eintreten, daß die elektronische Selbstkonsistenz zwar optimal schnell konvergiert, die notwendige Zielgenauigkeit jedoch nicht erreicht wird. In diesem Fall ist es notwendig, die Davidsondimension auf drei oder vier mal die Anzahl der Zustände zu erhöhen.

5.5 Relaxation der atomaren Struktur

Im Falle der atomaren Strukturrelaxation sind drei Dinge entscheidend für die Konvergenz:

- Die Wahl der Schrittweite α für den ersten Schritt
- Der separable Jacobian muß immer negativ definit sein.
- Die elektronische Kraft muß hinreichend auskonvergiert sein.

5.5.1 Der erste Quasi-Newton-Schritt

Die Schrittweite für den ersten Schritt wird durch den Startjacobian festgelegt. Dieser wird durch den Mixingparameter α bestimmt.

$$\vec{r}_\mu^{(2)} = \vec{r}_\mu^{(1)} + \Delta\vec{r}_\mu^{(2)} = \vec{r}_\mu^{(1)} + \alpha \vec{F}_\mu^{(1)} \quad (5.48)$$

Optimale Mixingparameter unterscheiden sich stark für verschiedene Systeme. Auf den ersten Blick ist keine Systematik zu erkennen. So kann ein Parameter α , der bei der Relaxation der (110) Oberfläche von InP zu schneller Konvergenz führt, bei der Relaxation der (111) Oberfläche von Si völlig versagen. Die Wahl eines optimalen Mixingparameters scheint so etwas wie schwarze Magie zu sein.

Nach einem Vorschlag von M. Weinert [Wei] wird die experimentell bestimmte Debye-Temperatur Θ_D benutzt, um den Mixingparameter α festzulegen. Θ_D definiert eine maximale Phononfrequenz ω_D . So kann man einen harmonischen Oszillator definieren, der mit der Phononfrequenz ω_D schwingt und diagonal in den Koordinaten $\Delta\vec{r}_\mu$ ist:

$$\hbar \omega_D = k_B \Theta_D \quad (5.49)$$

$$E[\dots, \vec{\tau}_\mu, \dots] = \frac{1}{2} \sum_\mu \langle \Delta \vec{\tau}_\mu | M^{\alpha(\mu)} \omega_D^2 | \Delta \vec{\tau}_\mu \rangle \quad (5.50)$$

$$\vec{\mathcal{F}}_\mu = M^{\alpha(\mu)} \omega_D^2 \Delta \vec{\tau}_\mu = \frac{M^{\alpha(\mu)} k_b^2 \Theta_D^2}{\hbar^2} \Delta \vec{\tau}_\mu \quad (5.51)$$

Man versucht nun, im ersten Schritt ins Energieminimum dieses Oszillators zu springen.

$$\Delta \vec{\tau}_\mu = \frac{\hbar^2}{M^{\alpha(\mu)} k_B^2 \Theta_D^2} \vec{\mathcal{F}}_\mu \quad (5.52)$$

Für die Wahl eines Mixingparameters wird die größte vorkommende Atommasse $M = \max_\alpha M^{(\alpha)}$ gewählt. Außerdem wird die Schrittlänge $\Delta \vec{\tau}_\mu$ nach oben durch einen Bruchteil der Bulk-Gitterkonstanten begrenzt ($\|\Delta \vec{\tau}_\mu\| \leq 0.2a_0$). Damit ergibt sich für den Mixingparameter:

$$\alpha = \min_\mu \left\{ \frac{0.2 a_0}{\|\vec{\mathcal{F}}_\mu\|}, \frac{\hbar^2}{M^{\alpha(\mu)} k_B^2 \Theta_D^2} \right\} \quad (5.53)$$

5.5.2 Negative Definitheit des Jacobian

Bei der Minimumsuche muß der separable inverse Jacobian in jeder Iteration negativ definit sein. Dies wird durch folgende Bedingung gesichert:

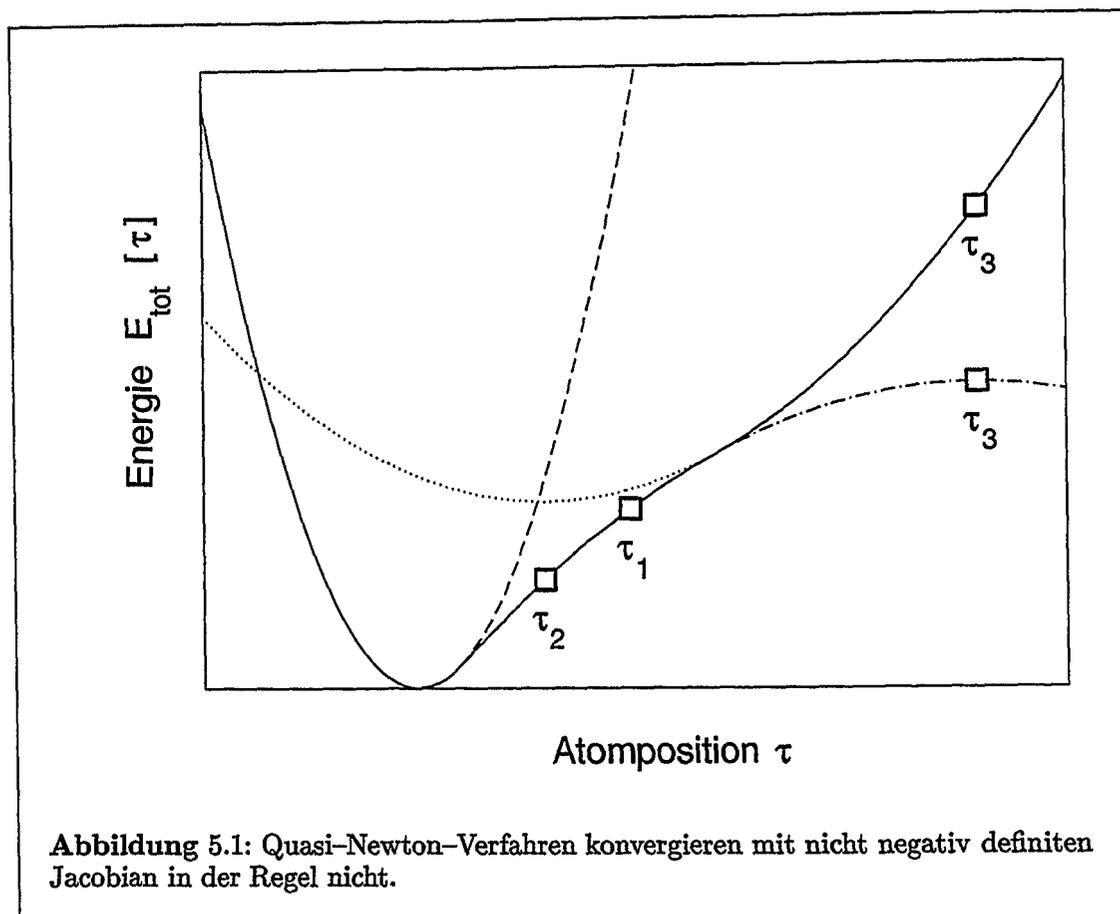
$$\langle \Delta x^{(i)} | \Delta \mathcal{F}^{(i)} \rangle = \sum_\mu \langle \Delta \vec{\tau}_\mu^{(i)} | \Delta \vec{\mathcal{F}}_\mu^{(i)} \rangle < 0 \quad (5.54)$$

Andernfalls kann die Konvergenz des Verfahrens zerstört werden. Dies ist in Abbildung 5.1 anhand eines eindimensionalen Beispiels skizziert.

Der erste Schritt von $\tau^{(1)}$ nach $\tau^{(2)}$ erfolgt längst der Kraftrichtung $\vec{\mathcal{F}}^{(1)}$, also in die richtige Richtung. Dann wird Jacobian aufgebaut, in diesem eindimensionalen Fall die Krümmung, und $\tau^{(3)}$ liegt im Extremwert der vom Quasi-Newton-Algorithmus ausgerechneten Parabel. Dies bedeutet bei nicht negativ definitem Jacobian wie in diesem Fall einen Schritt in die falsche Richtung. Tritt diese Instabilität auf, so muß man das Verfahren mit einem Gradientenschritt neu starten.

5.5.3 Konvergenzanforderung an die Kraft

Die Kraft setzt sich aus mehreren Anteilen zusammen, die sich in der Gleichgewichtskonfiguration kompensieren. Die Absolutwerte der einzelnen Anteile, i.e. Ewaldkraft, lokale und nichtlokale Kraft sind sehr groß ($\sim 10^4 - 10^5 \left[\frac{mRy}{au} \right]$) und kompensieren sich zu einer kleinen Gesamtkraft der Größenordnung ($\sim 10^0 - 10^{-2} \left[\frac{mRy}{au} \right]$).



Dementsprechend müssen die einzelnen Anteile sehr genau ausgerechnet werden. Ohne eine hinreichend genau berechnete Kraft kann ein Quasi-Newton-Verfahren nicht konvergieren. Vor allem am Anfang, bei der Startkonfiguration und bei den ersten relativ großen atomaren Relaxationsschritten, sind pro Relaxationsschritt viele elektronische Selbstkonsistenzschritte (~ 15) notwendig. In der Nähe der Gleichgewichtskonfiguration werden die atomaren Relaxationsschritte immer kleiner. Die Anzahl der benötigten elektronischen Selbstkonsistenzschritte pro Relaxationsschritt reduziert sich dann auf etwa (~ 8).

Im Computerprogramm wird die Genauigkeitsanforderung an die elektronische Selbstkonsistenz über die Konvergenz der lokale Kraft gesteuert. Die lokale Kraft wird immer eine Dezimalstelle genauer ausgerechnet als die totale Kraft der letzten atomaren Konfiguration.

Kapitel 6

(110) Oberflächen von III–V Halbleitern

6.1 Motivation

Obwohl die Herstellung von elektronischen Bauelementen aus Verbindungshalbleitern sehr viel aufwendiger und teurer ist als die Verwendung herkömmlicher Siliziumtechnologie, sind Verbindungshalbleiter doch von besonderem Interesse, vor allem wegen ihrer optoelektronischen Eigenschaften. Voraussetzung für Herstellung und zielgerichtetes Design elektronischer Bauelemente ist ein genaues Verständnis der atomaren und elektronischen Struktur von Halbleiter-Schichtsystemen und Oberflächen.

Die Rastertunnelmikroskopie (RTM) hat sich dabei in den letzten 10 Jahren als überaus leistungsfähiges experimentelles Werkzeug für die Untersuchung von Oberflächen, insbesondere für die Charakterisierung der elektronischen und atomaren Struktur, erwiesen [Che93, Wie94]. Atomare Auflösung ist bei hochauflösender Tunnelmikroskopie mittlerweile Standard, die neuere Methode der Tunnelspektroskopie bietet eine Möglichkeit, elektronische Oberflächenzustände zu detektieren.

Die (110) Oberfläche von III–V Halbleitern ist neben der Si-(111)-Oberfläche die zur Zeit wohl meist untersuchte und wohlbekannteste Halbleiteroberfläche. Sie dient als eine Art Referenzsystem, an der Oberflächeneffekte studiert werden, an der aber auch neue Methoden und theoretische Konzepte getestet werden können. Die Berechnung der (110) Oberfläche von III–V Halbleitern ist daher für den Autor in zweierlei Hinsicht interessant. Zum einen ist die Oberflächenstruktur für verschiedene III–V Halbleiter mit äquivalenten Methoden mehrfach berechnet worden. Somit handelt es sich um ein geeignetes Referenzsystem, welches dazu benutzt werden kann, Er-

fahrung bezüglich Konvergenzverhalten und numerischer Stabilität des Computerprogramms zu sammeln und Vertrauen in das Programm zu entwickeln.

Zum anderen interessiert den Autor die direkte Verbindung zu Tunnelexperimenten, die am gleichen Institut durchgeführt werden. Insbesondere interessiert den Autor eine theoretische Bestätigung der Interpretation der Tunnelexperimente von Ph. Ebert [Ebe93]. Dieser versucht mittels spannungsabhängiger, hochauflösender Tunnelmikroskopie, Informationen über die verschiedenen Oberflächenzustände zu gewinnen. Ph. Ebert untersucht dabei die Spannungsabhängigkeit der Morphologie von hochaufgelösten RTM-Bildern. Hochauflösende RTM wird somit auf eine neue Art und Weise spektroskopisch ausgewertet.

Die Spannungsabhängigkeit der Tunnelbilder, die durch die Beiträge unterschiedlicher Oberflächenzustände und Oberflächenresonanzen zum Tunnelstrom erklärt wird, soll im Detail theoretisch untersucht werden.

Auch wenn bei hochauflösender RTM atomare Auflösung erreicht wird, so heißt das nicht, daß man Atome "sieht". Vielmehr wird das Aussehen von RTM-Bildern durch die elektronischen Zustände bestimmt, die zum Tunnelstrom beitragen. So ist z.B. eine genaue Bestimmung der räumlichen Lage von Atomen, Adsorptionsplätzen oder Punktdefekten aus RTM-Bildern nur dann möglich, wenn die räumliche Ladungsdichtekontur dieser elektronischen Zustände genau bekannt ist. An eine weitere Zusammenarbeit mit den Experimentatoren zur Interpretation von Defektstrukturen ist gedacht.

Eine wichtige Zielsetzung bei der Weiterentwicklung der RTM-Methode ist sicherlich, möglichst vollständige Information über die elektronische Oberflächenstruktur aus RTM-Bildern bzw. Tunnelspektroskopie zu gewinnen. So ist es zur Zeit nicht möglich, Oberflächenatome mittels RTM chemisch zu unterscheiden. Eine theoretische Untersuchung der Ursache morphologischer Details von RTM-Bildern kann in dieser Richtung möglicherweise Beiträge liefern.

6.2 Das Strukturmodell

Die Klasse der III-V Halbleiter besteht aus je einem Element der Gruppe III und der Gruppe V des Periodensystems. In der Zinkblendestruktur ist jedes Atom der einen Sorte tetraedrisch an vier Nachbaratome gebunden (Abbildung 6.2) und die Valenzelektronen beider Elemente hybridisieren zu sp^3 -Orbitalen. Die Bindungen der verschiedenen III-V Halbleiter sind mehr oder weniger polar, da wegen der unterschiedlichen Elektronegativitäten das fünfwertige Anion Elektronen stärker anzieht

als das dreiwertige Kation. Die $\{110\}$ -Ebenen sind stöchiometrisch, d.h. sie enthalten ebenso viele Anionen wie Kationen und sind somit elektrisch neutral. Daher lassen sich III-V Halbleiter leicht längst $\{110\}$ -Ebenen spalten. Bei der Spaltung längst polarer Ebenen wie z.B. der $\{111\}$ -Ebene müßte zusätzliche elektrostatische Arbeit verrichtet werden. Die durch Spaltung im Ultrahochvakuum erzeugten (110) Oberflächen von III-V Halbleitern sind sehr rein und defektarm.

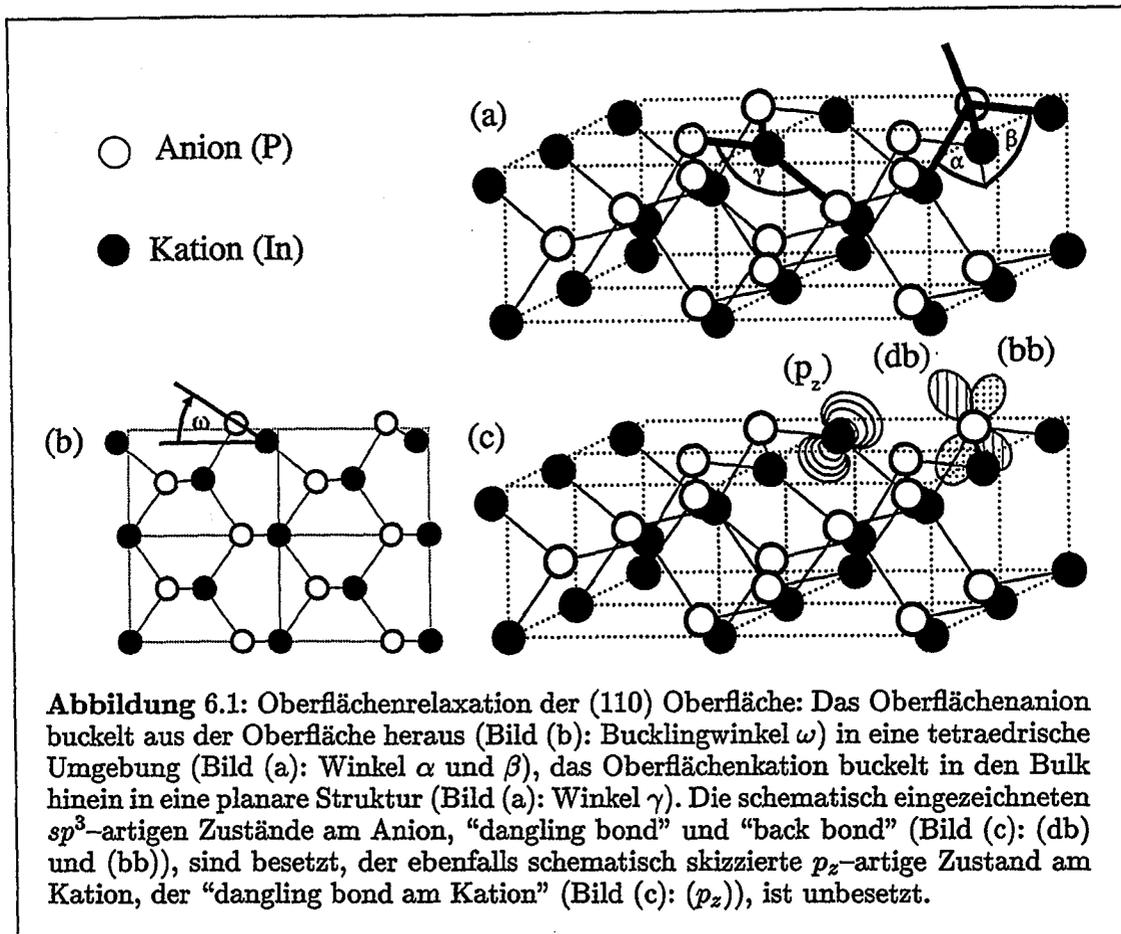
Für die reine (110) Oberfläche von III-V Halbleitern existiert ein experimentell wie theoretisch gut fundiertes Strukturmodell, das sogenannte "*bond rotation model*". Die genauesten experimentellen Strukturdaten [LB93] erhält man zur Zeit aus dynamischen LEED Experimenten [Kah83a, DMP⁺85]. Die besten experimentellen Daten über Oberflächenzustände [BE88, LB94] erhält man aus winkelaufgelösten Photoemissionsexperimenten [R RTP88, CCMS90] und aus der Tunnelspektroskopie [FS87a, MPBH⁺92, Fee94]. Theoretische Strukturdaten erhält man aus verschiedenen, selbstkonsistenten "*ab initio*"-Rechnungen [CC76, Cha79, CC79, Zun80, MBCM81, MMBC82, TMX84, SSMW83, FS87b, AHS91].

Alle (110) Oberflächen der III-V Halbleiter zeigen die gleiche Strukturrelaxation: Es handelt sich um eine 1×1 -Relaxation der Oberflächenatome, die in Abbildung 6.1 schematisch dargestellt ist. Das Oberflächenanion buckelt mit einem Bucklingwinkel ω aus der Oberfläche heraus in eine verzerrt tetraedrische, sp^3 -hybridisierte Umgebung (Winkel α und β). Das Oberflächenkation buckelt in den Bulk hinein in eine annähernd planare sp^2 -hybridisierte Umgebung (Winkel γ).

Von den vier Bindungen der Oberflächenatome der unrelaxierten Oberfläche liegen je zwei in der Oberflächenebene. Je eine Bindung ist auswärts und je eine Bindung ist bulkwärts gerichtet. Diese bezeichnet man als "dangling bond" bzw. "back bond". Durch die Relaxation der Oberflächenstruktur rehybridisieren die Oberflächenzustände und ändern ihren Charakter. Die Energie des "dangling bonds" am Anion senkt sich ab. In der tetraedrischen Umgebung des Anions ist der "dangling bond" ein sp^3 -artiger Zustand. Die Energie des "dangling bonds" am Kation wird angehoben. In der planaren Umgebung des Kations wird aus ihm ein p_z -artiger Zustand, der senkrecht auf der planaren sp^2 -Struktur steht.

Der energetisch tiefer liegende "dangling bond" am Anion wird somit besetzt, ebenso wie der energetisch noch tieferliegende in den Bulk hineingerichtete "back bond". Der aus der Oberfläche herausragende, energetisch aber oberhalb der Fermieenergie liegende p_z -Zustand am Kation bleibt unbesetzt. Insgesamt findet ein Ladungstransfer vom Kation zum Anion statt.

Es handelt sich an dieser Stelle natürlich nur um eine grobe Skizzierung der Oberflächenzustände. Um etwa Tunnelbilder im Detail zu verstehen, benötigt man weit



mehr Information, etwa über die Energiedispersion der Oberflächenzustände in der Oberflächenbrillouinzone, über die genaue Ladungsdichtekontur der Zustände und über deren Abfall ins Vakuum hinein, i.e. die lokale Zustandsdichte (LDOS: "local density of states") über der Oberfläche. Bisher wurden für die Interpretation von Tunnelbildern der (110) Oberflächen von III-V Halbleitern nur der besetzte "dangling bond" des Oberflächenanions und der unbesetzte p_z -Zustand am Oberflächenkation berücksichtigt. Dies erklärt zwar den grundsätzlichen Unterschied zwischen Tunnelbildern, die mit unterschiedlichem Spannungsvorzeichen aufgenommen werden¹, eine weitaus detailliertere Analyse ist aber möglich und erklärt interessante Details der Spannungsabhängigkeit der Tunnelbilder.

¹Für positive Tunnelspannungen wird aus dem besetzten Zustand am Anion heraus getunnelt, für negative Spannung in den unbesetzten Zustand am Kation hinein

6.3 Vorgehensweise bei der Berechnung

Bei der Berechnung der Oberflächenstruktur wird wie folgt vorgegangen:

Zunächst wird eine Strukturoptimierung des Bulksystems vorgenommen, d.h. die theoretische Gitterkonstante wird bestimmt. Die Konvergenz der Bulkrechnungen mit der Anzahl der \vec{k} -Punkte und der Anzahl der ebenen Wellen wird untersucht. Mit der theoretischen Gitterkonstanten werden die Bulkbandstruktur und deren Projektion auf die zweidimensionale Oberflächenbrillouinzone ausgerechnet und mit Literaturwerten verglichen.

Anschließend wird durch "Übereinanderstapeln" von tetragonalen Bulkeinheitenzellen (Abbildung 6.2) und Vakuum eine große Superzelle mit zwei äquivalenten (110) Oberflächen konstruiert. Dabei wird ebenfalls die theoretisch bestimmte Gitterkonstante verwendet. Dies ist notwendig, da die Dichtefunktionaltheorie Bindung überschätzt und somit die berechneten, theoretischen Gitterkonstanten um einige Prozent unter den experimentellen liegen. Die Verwendung der experimentellen Gitterkonstanten für die Oberflächensuperzelle würde zu unerwünschten Verspannungen des Systems führen. Einige Atomlagen der Oberfläche werden jeweils relaxiert und so die atomare Oberflächenstruktur ermittelt.

Die Konvergenz der Oberflächenrechnungen mit Anzahl der \vec{k} -Punkte und der Anzahl der ebenen Wellen wird untersucht. Ebenso wird die Konvergenz der Rechnungen mit der Größe der Superzellen (Anzahl der Atomlagen) und mit der Anzahl der relaxierten Lagen untersucht.

Bei fixierter Oberflächengeometrie wird dann in einer noch größeren Superzelle die elektronische Struktur berechnet, d.h. selbstkonsistente Ladungsdichte, Oberflächenzustände, zweidimensionale Bandstruktur und lokale Zustandsdichte. Diese werden ausgewertet und mit den Ergebnissen aus den Tunnelexperimenten verglichen. Die größere Superzelle ist notwendig, um Oberflächenresonanzen, die sehr langsam in den Bulk hinein abfallen, noch einigermaßen richtig zu beschreiben.

Glücklicherweise hängt die atomare Strukturrelaxation selbst nicht sehr empfindlich von der genauen Beschreibung der Oberflächenresonanzen ab, sondern mehr von der lokalen Ladungsdichteverteilung an der Oberfläche. Diese wird schon mit kleineren Superzellen hinreichend genau beschrieben.

6.3.1 Die Pseudopotentiale

Die in dieser Arbeit verwendeten Pseudopotentiale der III-V Halbleiter sind mit unserem eigenen Generierungsprogramm [Eng92] konstruiert worden. Sie sind alle-

samt normerhaltend und nach dem Schema von Bachelet, Haman, Schlüter "BHS" [BHS82] bzw. dessen Modifikationen nach Vanderbilt "Van" [Van85] konstruiert und werden bis auf das Ga-Pseudopotential in der separablen Kleinman-Bylander-Form "KB" [KB82] angewandt. Referenzatom für die Konstruktion der s - und p -Potentiale ist das freie Atom im Grundzustand. Für die Erzeugung der d -Potentiale wird eine angeregte, atomare Referenzkonfiguration verwendet. Alle Referenzatome werden skalarrelativistisch und nicht spinpolarisiert berechnet. Für die lokale Dichtenäherung wurde die Parametrisierung von Vosko, Wilk und Nusair [VWN80] mit den relativistischen Korrekturen von MacDonald und Vosko [MV79] benutzt. Die für die Konstruktion notwendigen Angaben, Abschneideradien und Referenzatome, sind in Tabelle 6.1 festgehalten.

Atom	r_s [a.u.]	r_p [a.u.]	r_d [a.u.]	Referenzatom für $l = 2$
<i>Al</i>	1.15	1.25	1.25	[Ne] $3s^{1.2}3p^{0.6}3d^{0.2}$
<i>Ga*</i>	1.15	1.35	1.35	[Ar] $4s^{1.5}4p^{0.5}4d^{0.25}$
<i>In</i>	1.15	1.25	1.25	[Kr] $5s^{1.2}5p^{0.6}5d^{0.2}$
<i>P</i>	1.15	1.25	1.25	[Ne] $3s^{1.8}3p^{1.8}3d^{0.4}$
<i>As</i>	1.15	1.25	1.25	[Ar] $4s^{1.8}4p^{1.8}4d^{0.4}$
<i>Sb</i>	1.15	1.25	1.25	[Kr] $5s^{1.8}5p^{1.8}5d^{0.4}$

Tabelle 6.1: Konstruktionsparameter für die Pseudopotentiale: Die Potentiale sind nach dem Schema von D. Vanderbilt "Van" [Van85] konstruiert, außer Ga, welches nach "BHS" [BHS82] konstruiert ist.

6.3.2 Die Oberflächen-Superzelle

Üblicherweise wird die Zinkblendestruktur als fcc-Kristall mit einer Basis aus zwei Atomen² ($Ka(0, 0, 0)$, $An(\frac{1}{4}a, \frac{1}{4}a, \frac{1}{4}a)$) beschrieben. Die drei Vektoren³ $(\frac{1}{2}a, \frac{1}{2}a, 0)$, $(\frac{1}{2}a, 0, \frac{1}{2}a)$ und $(0, \frac{1}{2}a, \frac{1}{2}a)$, die von der $(0, 0, 0)$ -Ecke der kubischen Einheitszelle (EZ) zu drei benachbarten Seitenmitten zeigen, spannen die fcc-Einheitszelle mit Volumen $V_{fcc} = \frac{1}{4}a^3$ auf (Abbildung: 6.2, oben, gestrichelte EZ). Die Koordinatengaben beziehen sich auf die kubische Einheitszelle mit der kubischen Gitterkonstanten a und Volumen $V_{cub} = a^3$, die insgesamt 8 Atomen enthält.

Grundlage unserer Beschreibung ist die tetragonale Einheitszelle (Abbildung: 6.2),

²das dreiwertige Kation (z.B. In) und das fünfwertige Anion (z.B. P)

³Kristallrichtungen [110], [101] und [011].

die 4 Atome enthält und das Volumen $V_{tet} = \frac{1}{2}a^3$ besitzt. Sie wird von den Vektoren⁴ $(\frac{1}{2}a, \frac{1}{2}a, 0)$, $(-\frac{1}{2}a, \frac{1}{2}a, 0)$ und $(0, 0, a)$ aufgespannt. Die tetragonale EZ besitzt die Symmetrie C_{2v} , i.e. zwei Spiegelebenen σ_x und σ_y , deren Schnittlinie die zweizählige Drehachse C_{2z} ist.

Die Oberflächen-Superzelle wird durch Übereinanderstapeln von tetragonalen Einheitszellen aufgebaut (Abbildung: 6.2, rechts unten). Die Atome der Superzelle sind in Lagen angeordnet, deren Abstand $\frac{\sqrt{2}}{4}a$ beträgt. Jede Lage besteht aus einer Zickzackkette von Atomen, in der sich Anionen und Kationen abwechseln. So enthält jede Atomlage der Superzelle genau ein Anion und ein Kation.

Jede Superzelle enthält eine ungerade Anzahl von Atomlagen. So wird sie spiegelsymmetrisch zur σ_x -Ebene. Oben und unten wird die Superzelle durch ein Vakuum abgeschlossen, welches jeweils N_{vak} Atomlagen entspricht. Die orthorhombische Oberflächen-Superzelle enthält somit zwei äquivalente (110) Oberflächen, deren Abstand $\frac{N_{vak}a}{\sqrt{2}}$ beträgt. Sie besitzt ebenso wie die tetragonale Einheitszelle die Symmetrie C_{2v} . Auch die (1×1) -Relaxation der Oberfläche zerstört diese Symmetrie nicht.

Die lange Seite der Superzelle, die [110]-Richtung (senkrecht zur (110) Oberfläche), wird die x-Achse. Die zweidimensionale Oberflächeneinheitszelle ist ein $(a \times \frac{\sqrt{2}}{2}a)$ -Rechteck. Die kurze Kante, die $[\bar{1}10]$ -Richtung, wird die neue y-Achse. Die lange Kante, die [001]-Richtung, wird die z-Achse.

In Tabelle 6.2 sind die Symmetrieelemente der einzelnen Zellen festgehalten. Die Beschreibung der Zinkblendestruktur durch tetragonale Einheitszellen entspricht keiner neuen internationalen Raumgruppe, denn die internationale Nummer einer Raumgruppe ist ja eindeutig und nicht von der Wahl der Beschreibung abhängig.

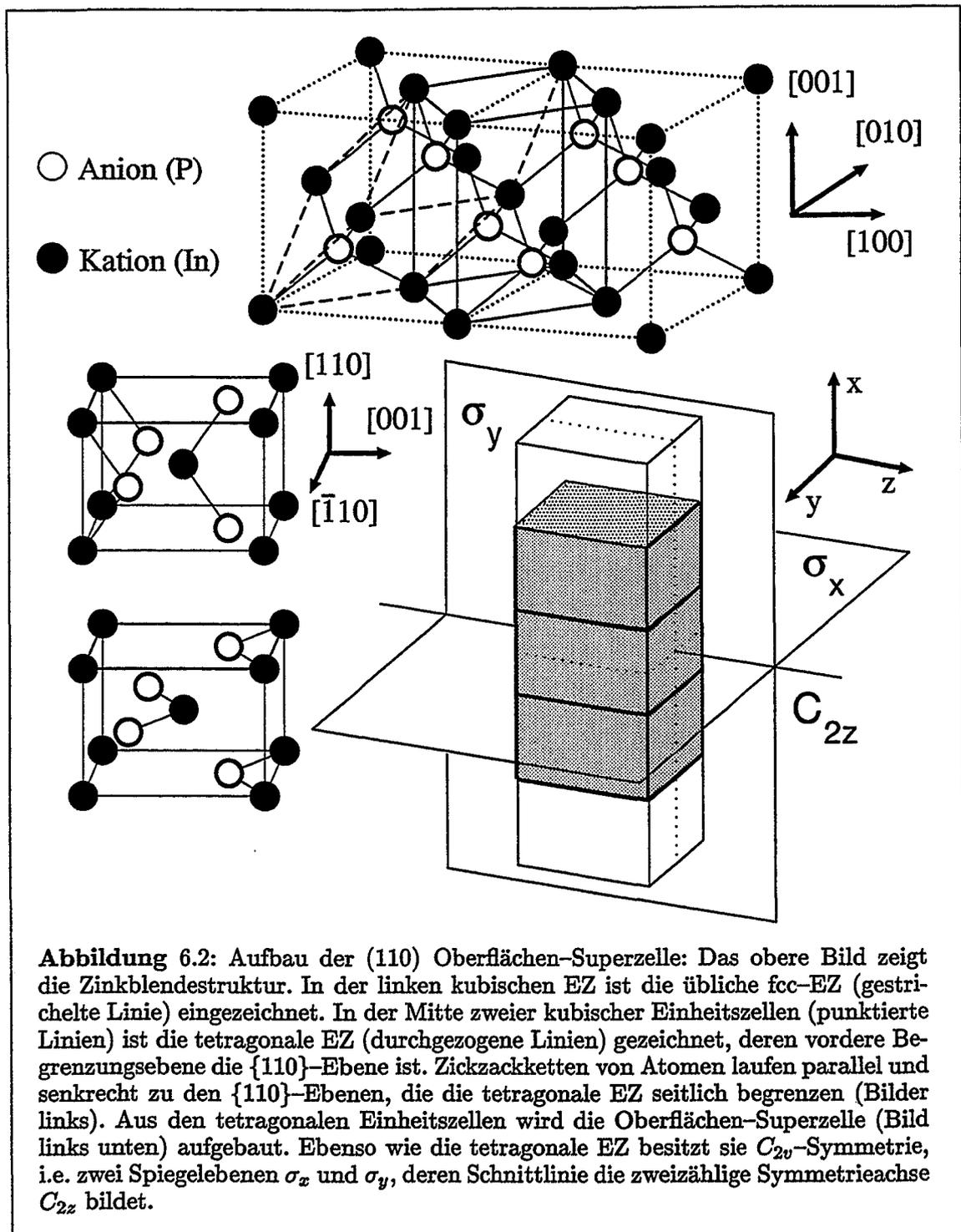
Abbildung 6.3 zeigt die Brillouinzone des reziproken Gitters.

Die Brillouinzone "tet-BZ", die zur tetragonalen Einheitszelle gehört, ist ein Quader, der ganz innerhalb der Brillouinzone "fcc-BZ" liegt, die zur fcc-Einheitszelle gehört. Die tet-BZ wird durch die Vektoren $\pm\vec{k}_x = \pm\frac{\pi}{a}(1, 1, 0)$, $\pm\vec{k}_y = \pm\frac{\pi}{a}(-1, 1, 0)$ und $\pm\vec{k}_z = \pm\frac{\pi}{a}(0, 0, 1)$ aufgespannt. Ihr irreduzibler Anteil "tet-IBZ" ist der Quader, der durch die positiven Vektoren \vec{k}_x , \vec{k}_y und \vec{k}_z aufgespannt wird. Er beträgt genau $\frac{1}{8}$ der totalen tet-BZ.

Die fcc-BZ liegt ganz in einem Würfel der Kantenlänge $\frac{4\pi}{a}$. Ihr Volumen beträgt $\frac{1}{2}\Omega_a$, wobei $\Omega_a = (\frac{4\pi}{a})^3$ das Volumen des obigen Würfels ist.

Bulkbandstrukturen werden längst Hochsymmetrielinien der irreduziblen Brillouinzone $(\Gamma \xrightarrow{(\Delta)} X - U(K) \xrightarrow{(\Sigma)} \Gamma \xrightarrow{(\Lambda)} L - K - W \xrightarrow{(Z)} X)$ ausgerechnet.

⁴Kristallrichtungen [110], $[\bar{1}10]$ und [001]



Das Volumen der tet-BZ ist halb so groß und beträgt $\frac{1}{4}\Omega_a$. Wird nun die tetragonale Einheitszelle im Ortsraum in [110]-Richtung zur orthogonalen Einheitszelle gestreckt, so schrumpft entsprechend die tet-BZ im reziproken Raum in [110]-Richtung

Einheitszelle	intern. Nummer	intern. Symbol	Schönflies-symbol	Symmetrieelemente
fcc	216	$F\bar{4}3m$	$\Gamma_c^f T_D^2$	$\{C_{2z} 000\}$ $\{C_{2x} 000\}$ $\{\sigma_{da} 000\}$ $\{C_{3i}^+ 000\}$
tetragonal	–	$Pmm2$	$\Gamma_q C_{2v}^1$	$\{\sigma_x 000\}$ $\{\sigma_y 000\}$
3d-Superzelle	25	$Pmm2$	$\Gamma_0 C_{2v}^1$	$\{\sigma_x 000\}$ $\{\sigma_y 000\}$
2d-Oberflächenzelle	–	$plm1$	–	$\{m_y 00\}$

Tabelle 6.2: Symmetriegruppen der Einheitszellen: Die ersten beiden Zeilen beschreiben die Zinkblendestruktur, wobei einmal die fcc-Einheitszelle und einmal die tetragonale Einheitszelle Grundlage der Beschreibung ist. Die Superzelle mit zwei äquivalenten (110) Oberflächen besitzt die gleichen Symmetrieelemente wie die tetragonale Einheitszelle.

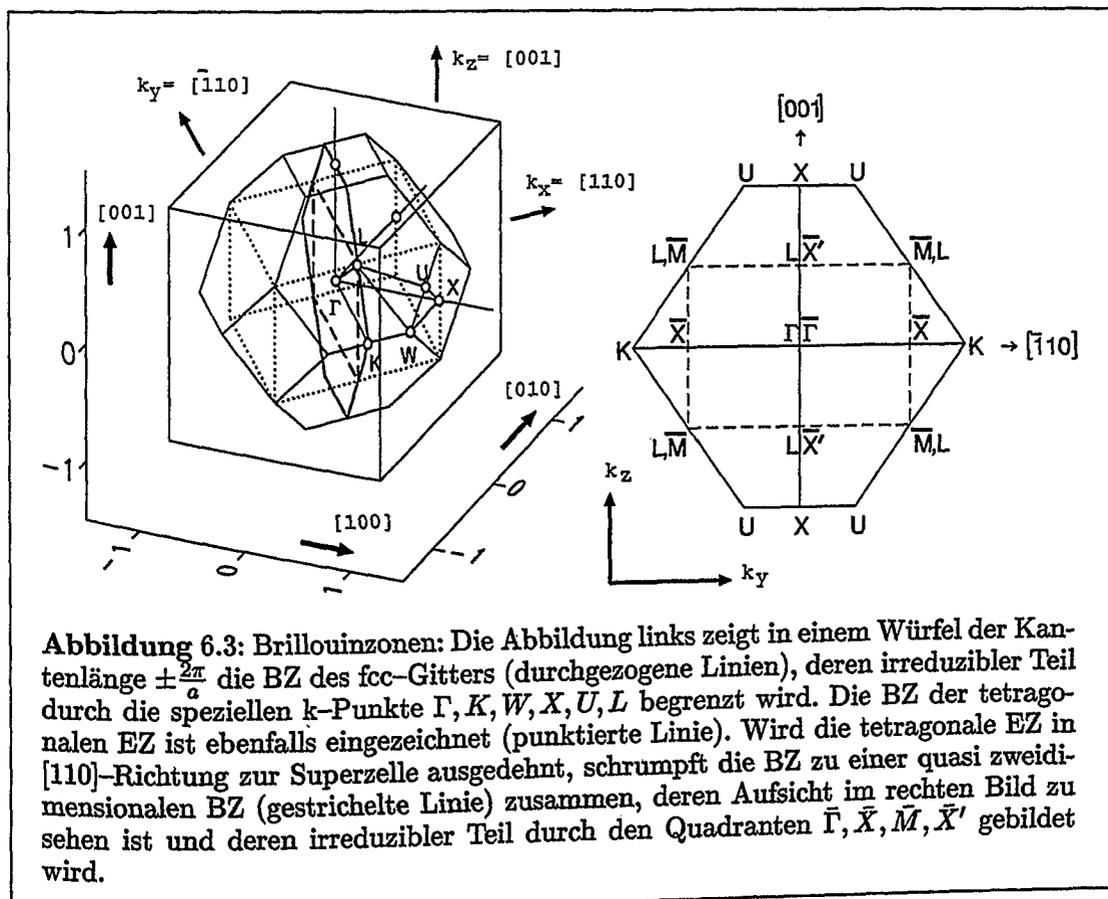


Abbildung 6.3: Brillouinzone: Die Abbildung links zeigt in einem Würfel der Kantenlänge $\pm \frac{2\pi}{a}$ die BZ des fcc-Gitters (durchgezogene Linien), deren irreduzibler Teil durch die speziellen k -Punkte Γ, K, W, X, U, L begrenzt wird. Die BZ der tetragonalen EZ ist ebenfalls eingezeichnet (gestrichelte Linie). Wird die tetragonale EZ in $[110]$ -Richtung zur Superzelle ausgedehnt, schrumpft die BZ zu einer quasi zweidimensionalen BZ (gestrichelte Linie) zusammen, deren Aufsicht im rechten Bild zu sehen ist und deren irreduzibler Teil durch den Quadranten $\bar{\Gamma}, \bar{X}, \bar{M}, \bar{X}'$ gebildet wird.

zu einer quasi zweidimensionalen Oberflächen-BZ zusammen. Die Oberflächen-BZ ist ein Rechteck mit den Kantenlängen $\frac{2\pi}{a}$ und $\sqrt{2}\frac{2\pi}{a}$ und den Eckpunkten \bar{M} . Der Mittelpunkt ist der $\bar{\Gamma}$ -Punkt. Die σ_y -Spiegelebene der Oberflächen-Superzelle wird

zur zweidimensionalen m_y -Spiegelachse, der $[1\bar{1}0]$ -Achse. Da im reziproken Raum wegen der Zeitumkehrinvarianz immer Inversionssymmetrie vorliegt, ist der irreduzible Teil der Oberflächenbrillouinzone genau ein Viertel der totalen Oberflächenbrillouinzone. Oberflächenbandstrukturen werden längst der Hochsymmetrielinien ($\bar{\Gamma} - \bar{X} - \bar{M} - \bar{X}' - \bar{\Gamma}$) gezeichnet.

Im reziproken Raum kann man nun die dreidimensionalen Energiebänder des Bulks auf die zweidimensionale Oberflächenbrillouinzone projizieren. Die Dispersion der dreidimensionalen Bulkenergiebänder senkrecht zur Oberflächenbrillouinzone wird dabei auf eine dichtliegende Folge von zweidimensionalen Energiebändern abgebildet: die projizierte Bulkbandstruktur. Zustände der Oberflächen-Superzelle, die in den Bandlücken oder Taschen der projizierten Bulkbandstruktur liegen, sind lokalisierte Oberflächenzustände. Zustände der Oberflächen-Superzelle, die innerhalb der projizierten Bulkbandstruktur liegen, sind Oberflächenresonanzen oder entsprechen Bulkzuständen.

6.4 Berechnung der Bulkstruktur

6.4.1 Bestimmung der theoretischen Gitterkonstanten

In einer früheren Diplomarbeit [Ric93] wurden die Pseudopotentiale am Beispiel von GaAs und Si auf Konvergenz getestet. Am Beispiel von InAs soll nun die Konvergenz mit der Anzahl von ebenen Wellen $N_{\vec{c}}$ bzw. der Abschneideenergie E_{cut} und mit der \vec{k} -Punkt-Dichte in der Brillouinzone überprüft werden (Tabelle 6.3). Die Strukturoptimierung für den Bulk zur Bestimmung der theoretischen Gitterkonstanten wird in der tetragonalen Einheitszelle (Abbildung 6.2) durchgeführt. Dies ist, was die Rechenzeit betrifft, nicht die günstigste (kleinste) Einheitszelle, sondern sie wird so gewählt, um mit den größeren Oberflächen-Superzellen kompatibel zu sein, die ja aus solchen tetragonalen Einheitszellen aufgebaut werden. Die Konvergenz mit der Anzahl an ebenen Wellen wird durch die "Cutoff-Energie" $E_{cut} = g_{max}^2$ bestimmt, bis zu der die Wellenfunktionen nach ebenen Wellen entwickelt werden.

Die \vec{k} -Punkt-Dichte wird auf die Brillouinzone des fcc-Gitters V_{fcc-BZ} umgerechnet. Da das Volumen der tetragonalen Einheitszelle doppelt so groß ist wie das Volumen der fcc-Einheitszelle, ist umgekehrt das Volumen V_{fcc-BZ} der Brillouinzone des fcc-Gitters doppelt so groß wie das Volumen V_{tet-BZ} der Brillouinzone des tetragonalen Gitters. So entsprechen z.B. 27 \vec{k} -Punkte in der Brillouinzone V_{tet-BZ} , i.e. 3 \vec{k} -Punkte pro Raumrichtung, einer \vec{k} -Punkt-Dichte von $54[V_{fcc-BZ}^{-1}]$. Die verwendeten

$(n \times n \times n)$ - \vec{k} -Punktsätze werden durch folgende Gleichungen definiert:

$$\vec{k}_{ijk} = \frac{1}{2n} \begin{cases} 2i \vec{k}_x + 2j \vec{k}_y + 2k \vec{k}_z \\ \text{sign}(i) (2|i| - 1) \vec{k}_x + \text{sign}(j) (2|j| - 1) \vec{k}_y + \text{sign}(k) (2|k| - 1) \vec{k}_z \end{cases} \quad (6.1)$$

Die obere Gleichung gilt für ungerade n und $(|i|, |j|, |k| < \frac{n}{2})$. Die untere Gleichung gilt für gerade n und $(0 < |i|, |j|, |k| < \frac{n}{2})$.

Insgesamt zeigt Tabelle 6.3 in Übereinstimmung mit [Ric93]:

- Mit der \vec{k} -Punkt-Summation wird über die Energiedispersion der Bandstruktur gemittelt: Eine \vec{k} -Punkt-Dichte von $n_{kpt} = 128 [V_{fcc-BZ}^{-1}]$, das entspricht 4 \vec{k} -Punkten pro Raumrichtung in der Brillouinzone der tetragonalen Einheitszelle, ist ausreichend für eine Konvergenz der totalen Energie bis auf ~ 1 mRy pro Atom.
- Ein ebene Wellen Cutoff von $g_{max} = 3.5 \sqrt{\text{Ry}}$, $E_{cut} = 12.25$ Ry ist ausreichend für eine Konvergenz der totalen Energie bis auf ~ 10 mRy pro Atom. Die Konvergenz mit der Anzahl der ebenen Wellen wird durch die Fourierkomponenten der kinetischen Energie bestimmt. Diese sind im wesentlichen unabhängig von der chemischen Umgebung [RRKJ90, Eng92]. Daher ist die totale Konvergenz der Energie auf ~ 10 mRy pro Atom auf jeden Fall ausreichend für eine relative Genauigkeit besser als ~ 1 mRy pro Atom. Nur diese wird beim Vergleich verschiedener Strukturen benötigt.

Alle weiteren Rechnungen werden mit $(4 \times 4 \times 4)$ \vec{k} -Punkten in der Brillouinzone der tetragonalen Einheitszelle durchgeführt und mit einem Wert für g_{max} von $3.5 \sqrt{\text{Ry}}$. Insgesamt ergibt sich so für die Konvergenz der Berechnung von Strukturparametern wie z.B. der kubischen Gitterkonstanten a_0 ein Fehler von weniger als 0.5%. Die Ergebnisse der Strukturoptimierung für einige III-V Halbleiter sind in Tabelle 6.4 festgehalten.

Ursache der Unterschiede zwischen Theorie und Experiment, i.e. der bis zu einige Prozent zu kleinen Gitterkonstanten und der Abweichungen in den Bulkmoduli von bis zu 10% ist nicht mangelnde Konvergenz der Rechnung. Vielmehr ist dies typisch für Pseudopotentialrechnungen. Zum einen überschätzt die Dichtefunktionaltheorie im allgemeinen die Bindung. Zum anderen werden die d-Elektronen nicht als Valenzelektronen behandelt, sondern sie sind Teil des ionischen ("frozen") Pseudopotentials. Dies erklärt die Unterschiede zwischen Theorie und Experiment.

	$E_{cut} = 12.25\text{Ry}$ $g_{max} = 3.5\sqrt{\text{Ry}}$	$E_{cut} = 20.25\text{Ry}$ $g_{max} = 4.5\sqrt{\text{Ry}}$
$n_{kpt} = 54 [V_{fcc-BZ}^{-1}]$ $(3 \times 3 \times 3) [V_{tet-BZ}^{-1}]$	$E_0 = -8.42183\text{Ry}$ $a_0 = 6.05549\text{Å}$	$E_0 = -8.44529\text{Ry}$ $a_0 = 6.03597\text{Å}$
$n_{kpt} = 128 [V_{fcc-BZ}^{-1}]$ $(4 \times 4 \times 4) [V_{tet-BZ}^{-1}]$	$E_0 = -8.43007\text{Ry}$ $a_0 = 6.03347\text{Å}$	$E_0 = -8.45365\text{Ry}$ $a_0 = 6.01264\text{Å}$
$n_{kpt} = 250 [V_{fcc-BZ}^{-1}]$ $(5 \times 5 \times 5) [V_{tet-BZ}^{-1}]$	$E_0 = -8.42769\text{Ry}$ $a_0 = 6.03729\text{Å}$	$E_0 = -8.45133\text{Ry}$ $a_0 = 6.01873\text{Å}$

Tabelle 6.3: Konvergenz der Bulkstrukturoptimierung mit Dichte der \vec{k} -Punkte und ebene Wellen Cutoff für InAs: totale Energien sind in Ry pro Atom angegeben, kubische Gitterkonstanten in Å. Die Anzahl der \vec{k} -Punkte ist umgerechnet auf das Volumen der entsprechenden Brillouinzone des fcc-Gitters. Tatsächlich benötigt man wegen Symmetrie und größerer tetragonaler Einheitszelle natürlich weniger \vec{k} -Punkte.

System		$a_0[\text{Å}]$	$B[\text{Mbar}]$
InAs	theo.	6.0335	0.679
	exp.	6.0584	0.612
InP	theo.	5.8549	0.747
	exp.	5.8693	0.725
GaAs	theo.	5.5706	0.727
	exp.	5.6533	0.742
AlAs	theo.	5.6290	0.729
	exp.	5.6607	0.763

Tabelle 6.4: Strukturoptimierung III-V Halbleiter: Die experimentellen Werte sind [LB84] entnommen. Die um einige Prozent zu kleinen theoretischen Gitterkonstanten, und die Fehler in den Bulkmoduli um ebenfalls einige Prozent sind methodenspezifisch und bekannt. Die DFT überschätzt im allgemeinen die Bindung, die d-Elektronen werden nicht als Valenzelektronen behandelt.

6.4.2 Die elektronische Struktur des Bulks

Mit der selbstkonsistent berechneten Ladungsdichte für das System mit der theoretischen Gitterkonstanten werden dann für die üblichen Hochsymmetrielinien die Bandstrukturen errechnet. Da die LDA wegen der falschen Beschreibung der Selbstenergie Bandlücken völlig falsch beschreibt, werden diese mit dem sogenannten "Scissors"-Operator korrigiert, d.h. durch Verschieben der Leitungsbänder um einen konstanten Wert wird die Bandlücke am Γ -Punkt an den experimentell ermittelten Wert angepaßt. In Tabelle 6.5 sind die errechneten LDA-Bandlücken am Γ -Punkt und die experimentellen Werte festgehalten. Die experimentellen Werte sind [LB82] entnommen.

System	Δ_{Γ}^{LDA} [eV]	Δ_{Γ}^{exp} [eV]	Δ_{Γ}^{gap} [eV]
InAs	0.000	0.41	+ 0.410
InP	0.847	1.42	+ 0.573
GaAs	1.109	1.52	+ 0.411
AlAs	2.164	3.13	+ 0.996

Tabelle 6.5: Experimentelle und LDA-Bandlücken für verschiedene III-V Halbleiter: Die experimentellen Werte sind [LB82] entnommen. Die LDA-Leitungsbänder müssen um Δ^{gap} verschoben werden.

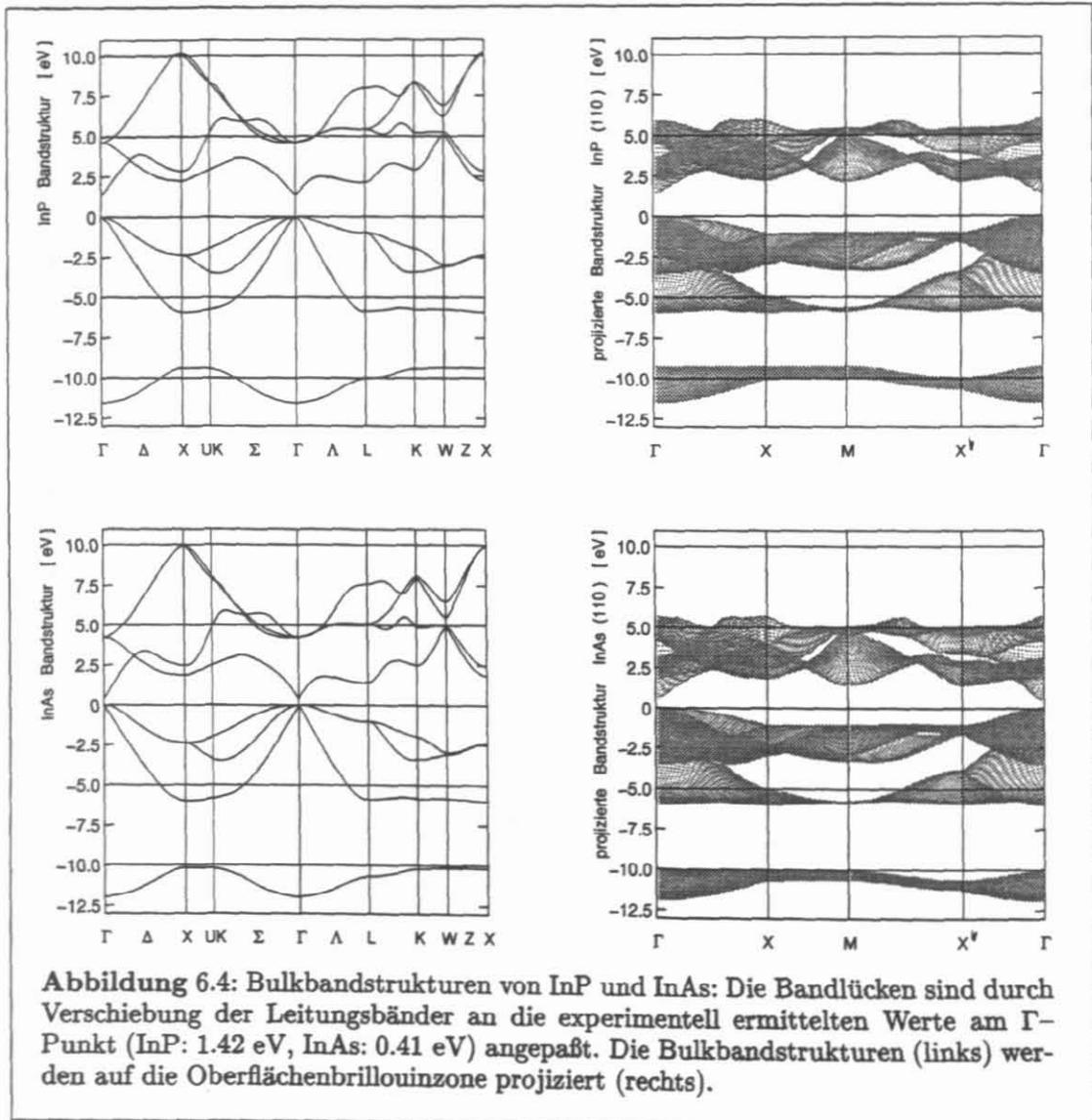
Die so errechneten Bandstrukturen von InP, InAs, GaAs und AlAs sind in Abbildung 6.4 und Abbildung 6.5 abgebildet. Insgesamt werden Details der Bandstrukturen gut wiedergegeben.

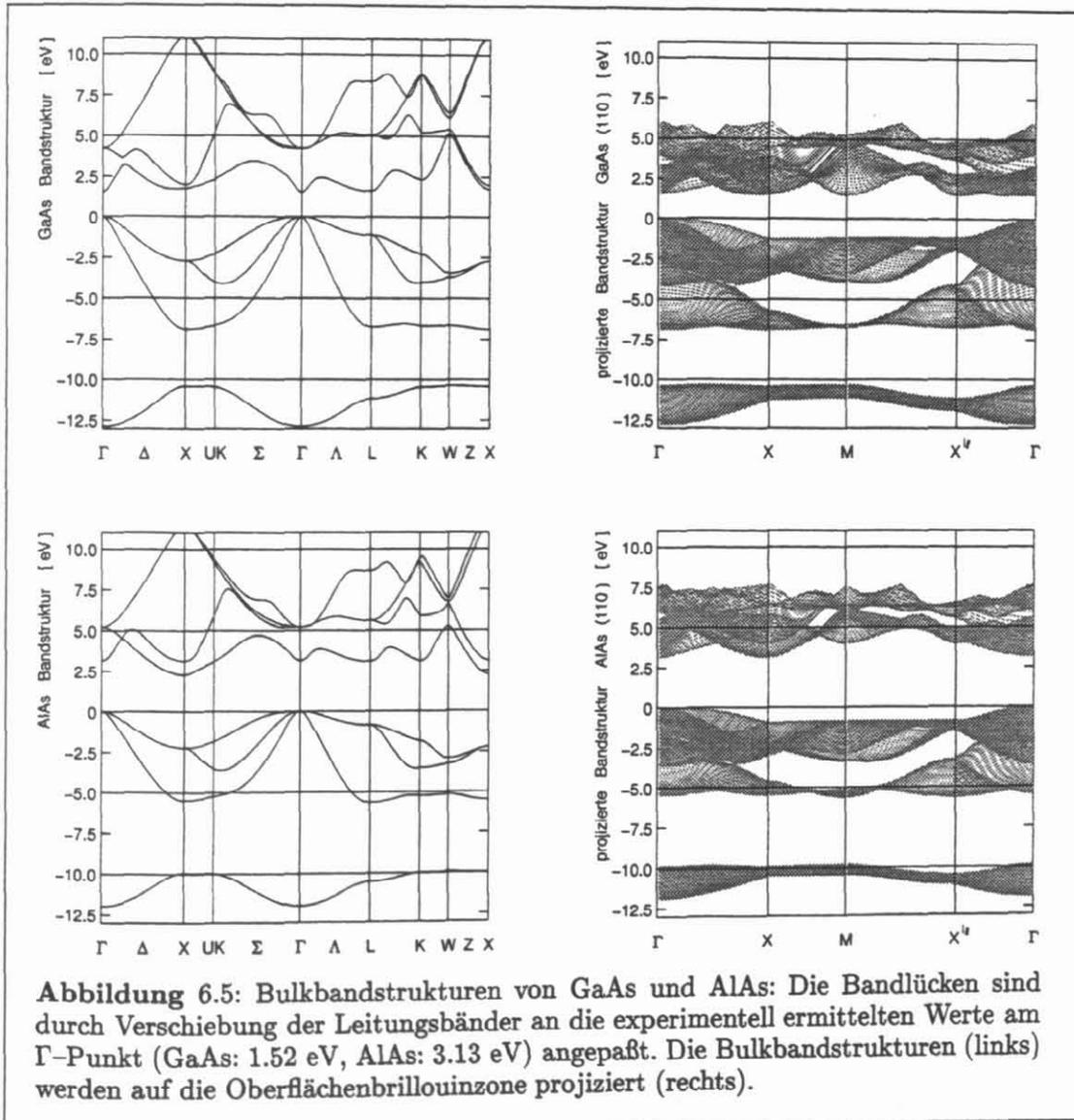
AlAs ist der einzige indirekte Halbleiter. Wird die Bandlücke am Γ -Punkt an den experimentellen Wert 3.13 eV angepaßt, so ergibt sich eine indirekte Bandlücke ($X \rightarrow \Gamma$) von 2.27 eV. Dies stimmt sehr genau mit dem experimentellen Wert 2.23 eV überein.

Exemplarisch sind in Tabelle 6.6 für InP einige explizite Energiewerte der Bänder für die Hochsymmetriepunkte Γ , X und L festgehalten. Unsere errechneten Werte werden mit Literaturwerten aus [O'R91] verglichen. Dabei stammen die experimentellen Daten aus winkelaufgelösten Photoemissionsexperimenten [LRF⁺74, LVHK87], die theoretischen Werte stammen aus Pseudopotentialrechnungen [JM76]. Insgesamt stimmen alle Werte recht gut überein.

	<i>exp.*</i>	<i>wir</i>	<i>theor.*</i>		<i>exp.*</i>	<i>wir</i>	<i>theor.*</i>		<i>exp.*</i>	<i>wir</i>	<i>theor.*</i>
Γ_1^v	-11.00	-11.46	-11.42	X_1^v	-8.90	-9.23	-8.91	L_1^v	-10.00	-9.91	-9.67
Γ_2^v	-0.11	0.00	-0.21	X_2^v	-5.90	-5.92	-6.01	L_2^v	-5.90	-5.87	-5.84
Γ_3^v		0.00		X_3^v	-2.00	-2.37	-2.09	L_3^v	-1.00	-0.99	-1.09
Γ_4^v	0.00	0.00	0.00	X_4^v	-2.00	-2.37	-2.06	L_4^v	-1.00	-0.99	-0.94
Γ_1^c	1.43	1.43	1.50	X_1^c	2.40	2.21	2.44	L_1^c	2.04	2.10	2.19
Γ_2^c	4.78	4.64	4.64	X_2^c		2.85	2.97	L_2^c		5.45	5.58
Γ_3^c	4.85	4.64	4.92					L_3^c		5.45	5.70

Tabelle 6.6: Energien einiger Hochsymmetriepunkte der InP Bandstruktur (in eV): Die Eigenwerte sind von niedrigen zu höheren Energien durchnummeriert. Der Energienullpunkt ist auf Γ_4^v gelegt, das höchste Valenzband am Γ -Punkt. Γ_1^c ist das niedrigste Leitungsband. Die mittleren Spalten sind unsere errechneten Werte, die Literaturdaten "*" stammen aus [O'R91]. In der Rechnung in [O'R91] ist die Spin-Bahn-Kopplung berücksichtigt, in unseren Rechnungen nicht. Daraus erklären sich die Unterschiede. Die Spin-Bahn-Kopplungsenergie ist allerdings für InP sehr klein.





6.5 Die atomare Relaxation der Oberfläche

Für AlAs, InAs und InP wird die Strukturrelaxation für verschieden große Superszellen durchgeführt. Die Anzahl der Atomlagen, die relaxiert werden, wird variiert. Die Ergebnisse der Rechnungen mit verschiedenen \vec{k} -Punkt-Sätzen, und mit einer Basis aus ebenen Wellen bis zu $g_{max} = 3.5$ bis $4.5\sqrt{Ry}$ werden verglichen.

6.5.1 Globale Konvergenz

Die elektronische Selbstkonsistenz wird durch Potentialmischen mit Mixingparameter $\alpha = 0.5$ erreicht. Es werden vier Davidson-Kosugi-Schritte pro Selbstkonsistenzschritt durchgeführt. Die Dimension des Davidsonunterraums wird auf $3 \times N_{state}$ (drei mal die Anzahl der elektronischen Zustände) festgelegt.

Startpositionen für die atomare Relaxation sind die idealen Bulkklagen. Der erste Quasi-Newton-Schritt wird durch eine Debye-Temperatur von 330 K festgelegt, wobei die maximal zulässige Schrittweite 0.15 au. beträgt.

In Tabelle 6.7 ist skizziert, wie die Konvergenz von Kraft δF , Atompositionen δr und Energie δE_{tot} zusammenhängen. Dies sind natürlich nur grobe Richtwerte, die von System zu System schwanken.

N_{Mdyn}	δF	δr	δr	$\frac{\delta E_{tot}}{Atom}$
	$\frac{mRy}{au.}$	[au.]	[a_0]	[mRy]
5	10^1	5×10^{-2}	5×10^{-3}	10^{-1}
10	10^0	5×10^{-3}	5×10^{-4}	10^{-2}
15	10^{-1}	5×10^{-4}	5×10^{-5}	10^{-4}
20	10^{-2}	5×10^{-5}	5×10^{-6}	10^{-8}

Tabelle 6.7: Konvergenz der Strukturrelaxation: Die Kraft konvergiert in etwa quadratisch mit der Anzahl der Quasi-Newton-Schritte N_{Mdyn} , die Atompositionen ebenso. Die totale Energie konvergiert quadratisch mit der Kraft.

Dieses Konvergenzverhalten wird strenggenommen nur im Einzugsbereich eines quadratischen Energieminimums erwartet. Für die Relaxation der (110) Oberfläche aller von uns untersuchten III-V Halbleiter stellt sich heraus, daß sich die unrelaxierte Oberfläche bereits im Einzugsbereich der Minimumkonfiguration befindet. Bei der atomaren Relaxation anderer Systeme ist dies nicht immer der Fall.

Realistisch gesehen kann die Pseudopotentialmethode verschiedene atomare Konfigurationen, die in der totalen Energie $\sim mRy$ pro Atom auseinanderliegen, energie-

III-V (110) (1 × 1)-Relaxation						
Gitterkonstanten: $a = a_0$, $b = \frac{a}{\sqrt{2}}$, $\varphi = 90^\circ$						
Lage	Atom	δz	δx	y	z	x
		[a]	[b]	[b]	[a]	[b]
		Superzelle		unrelaxierte Oberfläche		
1	Anion	$\delta_{a }^1$	$\delta_{a\perp}^1$	0.0	1.0	0.0
	Kation	$\delta_{c }^1$	$\delta_{c\perp}^1$	0.5	0.25	0.0
2	Kation	$\delta_{c }^2$	$\delta_{c\perp}^2$	0.0	0.75	-0.5
	Anion	$\delta_{a }^2$	$\delta_{a\perp}^2$	0.5	0.5	-0.5
3	Anion	$\delta_{a }^3$	$\delta_{a\perp}^3$	0.0	1.0	-1.0
	Kation	$\delta_{c }^3$	$\delta_{c\perp}^3$	0.5	0.25	-1.0
4	Kation	$\delta_{c }^4$	$\delta_{c\perp}^4$	0.0	0.75	-1.5
	Anion	$\delta_{a }^4$	$\delta_{a\perp}^4$	0.5	0.5	-1.5

Tabelle 6.8: III-V (110) Strukturrelaxation: Die absoluten Verschiebungen der Atome aus ihren unrelaxierten Positionen werden in relativen Einheiten angegeben.

III-V (110) (1 × 1)-Relaxation									
Gitterkonstanten: $a = a_0$, $b = \frac{a}{\sqrt{2}}$, $\varphi = 90^\circ$									
Lage	$\Delta_{i,\perp}$	$\Delta_{i,x}$	$d_{i,i+1,\perp}$	$d_{i,i+1,x}$	$\Delta_{i,\perp}$	$\Delta_{i,x}$	$d_{i,i+1,\perp}$	$d_{i,i+1,x}$	
	[b]	[0.75 × a]	[0.5 × b]	[0.5 × a]	[b]	[0.75 × a]	[0.5 × b]	[0.5 × a]	
		Superzelle				unrelaxierte Oberfläche			
1	$\Delta_{1,\perp}$	$\Delta_{1,x}$	$d_{12,\perp}$	$d_{12,x}$	0.0	1.0	1.0	1.0	
2	$\Delta_{2,\perp}$	$\Delta_{2,x}$	$d_{23,\perp}$	$d_{23,x}$	0.0	1.0	1.0	1.0	
3	$\Delta_{3,\perp}$	$\Delta_{3,x}$	$d_{34,\perp}$	$d_{34,x}$	0.0	1.0	1.0	1.0	
4	$\Delta_{4,\perp}$	$\Delta_{4,x}$	$d_{45,\perp}$	$d_{45,x}$	0.0	1.0	1.0	1.0	

Tabelle 6.9: III-V (110) Strukturrelaxation: Die Abstände der relaxierten Atome untereinander werden in relativen Einheiten angegeben.

entsprechenden unrelaxierten Bulkabstände aufgeschrieben. Sind die unrelaxierten Bulkabstände 0, so werden wie oben die relativen Einheiten der tetragonalen Einheitszelle benutzt.

Der Grund für die zwei unterschiedlichen Notationen ist folgender:

- Die relativen Abstände der Oberflächenatome konvergieren mit zunehmender Zellgröße schneller als die absoluten Verschiebungen relativ zu den unrelaxierten Bulkpositionen.
- In der Literatur wird oft nur ein unvollständiger Satz von relativen Abständen angegeben, aus dem sich die absoluten Verschiebungen nicht ausrechnen lassen. Diese enthalten aber die vollständige Information, auf die hier nicht verzichtet werden soll.

6.5.3 Konvergenz mit der Systemgröße

Zunächst wird am Beispiel von AlAs die Konvergenz der Rechnungen mit der Systemgröße untersucht. Dabei werden zunächst die Größe der Superzelle und die Größe des \vec{k} -Punkt-Satzes (Gleichung (6.1)) variiert. Die Größe des Vakuums $\frac{3.5a_0}{\sqrt{2}}$ bleibt fest und entspricht genau 3.5 Atomlagen über jeder Oberfläche, also insgesamt 7 Atomlagen. Dies ist ausreichend, da die Wellenfunktionen sehr schnell ins Vakuum hinein abfallen. Die Rechnungen werden mit $g_{max} = 3.5\sqrt{Ry}$ durchgeführt. Die Ergebnisse sind in Tabelle 6.10 festgehalten.

Festzustellen ist, daß es kaum Abhängigkeiten von der Anzahl der Atomlagen gibt. Bereits beim kleinsten möglichen System mit 5 Atomlagen sind die Verschiebungen der Atome aus ihren unrelaxierten Bulkpositionen bis auf eine relative Genauigkeit von $\delta\tau < 5 \times 10^{-4}$ auskonvergiert. Zwischen den Zellen mit 7 und 9 Atomlagen ist nur noch ein Unterschied $\delta\tau < 1 \times 10^{-4}$ erkennbar.

Auch der kleinste untersuchte \vec{k} -Punkt-Satz ist mit $2 \times 4 \times 4$ \vec{k} -Punkten in der Brillouinzone der Superzelle auf vergleichbare Genauigkeit auskonvergiert. Nutzt man die Symmetrie der Zelle aus, so entspricht dies 4 \vec{k} -Punkten in der irreduziblen Brillouinzone. In Einheiten der reziproken Gittervektoren sind dies: $\{ (\frac{1}{4}, \frac{1}{8}, \frac{1}{8}), (\frac{1}{4}, \frac{3}{8}, \frac{1}{8}), (\frac{1}{4}, \frac{1}{8}, \frac{3}{8}), (\frac{1}{4}, \frac{3}{8}, \frac{3}{8}) \}$. Im folgenden wird nur noch dieser \vec{k} -Punkt-Satz verwendet.

6.5.4 Konvergenz mit der Anzahl der relaxierten Lagen

Auch die Konvergenz der Rechnungen mit der Anzahl der Lagen wird untersucht. Dabei werden drei verschiedene Zellgrößen verglichen, Superzellen mit 5, 7 und 9 Atomlagen. Wieder werden die beiden äquivalenten (110) Oberflächen der Superzelle durch Vakuum getrennt, welches insgesamt 6 Atomlagen entspricht. In jeder Superzelle wird nur die mittlere Lage festgehalten. Das bedeutet, daß in den drei Zellen 4,6 bzw. 8 Lagen relaxiert werden, je 2,3 bzw. 4 auf jeder Seite. Auch diese Rechnungen werden mit $g_{max} = 3.5\sqrt{Ry}$ durchgeführt. Die Ergebnisse sind in Tabelle 6.11

festgehalten.

Die Ergebnisse hängen signifikant von der Anzahl der relaxierten Lagen ab. Relaxationen der unteren Atomlagen pflanzen sich bis zur Oberfläche fort, die relativen Abstände der oberen Lagen ändern sich jedoch nur wenig. Daher werden für diese Testreihe auch die relativen Abstände der relaxierten Atome untersucht (Tabelle 6.12). Diese konvergieren für die äußeren Atomlagen sehr viel schneller als die absoluten Verschiebungen aus den unrelaxierten Bulktagen. Insgesamt reicht schon die kleinste Zelle (5 Atomlagen) aus, die Oberflächenrelaxation genau zu beschreiben.

6.5.5 Konvergenz mit der Anzahl der ebenen Wellen

Für InAs wird am Beispiel der kleinsten Superzelle (Zelle mit 5 Atomlagen) die Konvergenz mit der Größe der ebene Wellen Basis untersucht. Dazu wurden Strukturrelaxationen mit verschiedenen ebene Wellen Cutoffs durchgeführt ($g_{max} = 3.5, 4.0$ und $4.5 \sqrt{Ry}$). Zwei verschiedene theoretische Gitterkonstanten wurden verwendet, die mit $g_{max} = 3.5 \sqrt{Ry}$ bzw. $4.5 \sqrt{Ry}$ berechnet wurden (siehe Tabelle 6.3). Die Ergebnisse sind in Tabelle 6.13 festgehalten. Die Rechnungen sind bei $g_{max} = 3.5 \sqrt{Ry}$ schon sehr schön auskonvergiert. Dies zeigt der Vergleich der Ergebnisse der Strukturrelaxationen bei $g_{max} = 3.5$ bzw. $4.5 \sqrt{Ry}$, die mit den passenden Gitterkonstanten 6.03347 \AA bzw. 6.01264 \AA durchgeführt wurden (Tabelle 6.13 links oben bzw. rechts unten).

6.5.6 Vergleich mit der Literatur

Unsere Strukturparameter der (110) Oberfläche werden mit Literaturwerten verglichen. Die experimentellen Daten aus LEED-Messungen sind den Tabellen aus [Kah83b, BE88, AHS91] entnommen und auf die hier tabellierten Größen (Tabellen 6.14 und 6.15) umgerechnet. Die theoretischen Literaturwerte stammen aus Pseudopotentialrechnungen [AHS91].

Insgesamt zeigen alle hier untersuchten III-V Halbleiter die gleiche Strukturrelaxation. Alle experimentell und theoretisch ermittelten Werte stimmen recht gut überein:

- Haupteffekt der Relaxation ist die senkrechte Verschiebung der Oberflächeatome $\Delta_{1,\perp}$ und $d_{12,\perp}$ sowie die Parallelverschiebung $d_{12,\parallel}$ der ersten Oberflächenlage gegenüber der zweiten in [001] Richtung. Diese Verschiebungen sind $\sim 1 \text{ \AA} - 2 \text{ \AA}$ groß und alle experimentellen und theoretischen Werte stimmen auf $\sim 0.05 \text{ \AA}$ genau überein.

- Verschiebungen der Größenordnung $\sim 0.1\text{\AA}$ werden noch in Richtung und Größenordnung richtig beschrieben.
- Alle Relaxationswinkel stimmen auf wenige Prozent genau überein.
- Die Bindungslängen der Oberflächenatome bleiben bis auf einige Prozent gegenüber dem Bulk unverändert. Diese Aussage läßt sich experimentell wie theoretisch bestätigen. Die genaue Änderung der Bindungslängen stimmen mit den experimentellen Resultaten nicht überein. Dies liegt nicht an der mangelnden numerischen Konvergenz. Die Unterschiede zu anderen Pseudopotentialrechnungen und die Unterschiede verschiedener experimentell ermittelter Werte lassen vermuten, daß sowohl die Pseudopotentialmethode als auch die Experimente dafür nicht genau genug sind.

AlAs (110) (1 × 1)-Relaxation							
berechnete Gitterkonstanten: $a = 5.6290\text{Å}$, $b = \frac{a}{\sqrt{2}}$, $\varphi = 90^\circ$							
experimentelle Gitterkonstanten: $a = 5.6607\text{Å}$, $b = \frac{a}{\sqrt{2}}$, $\varphi = 90^\circ$							
Lage	Atom	δz [a]	δx [b]	δz [a]	δx [b]	δz [a]	δx [b]
Konvergenz mit der Anzahl der Atomlagen: (2 × 4 × 4) \vec{k} -Punkte							
		5 Lagen		7 Lagen		9 Lagen	
1	As	- 0.028374	0.058172	- 0.028279	0.058101	- 0.028265	0.057917
	Al	- 0.072075	- 0.115138	- 0.071730	- 0.114695	- 0.071717	- 0.114754
2	Al	0.009506	0.021082	0.009640	0.020976	0.009580	0.020880
	As	0.006821	- 0.000709	0.006754	- 0.001075	0.006755	- 0.001108
3	As	0.0	0.0	0.0	0.0	0.0	0.0
	Al	0.0	0.0	0.0	0.0	0.0	0.0
\vec{k} -Punkt-Konvergenz: (5 Atomlagen)							
		(2 × 4 × 4) \vec{k} -Punkte		(4 × 4 × 4) \vec{k} -Punkte		(2 × 6 × 6) \vec{k} -Punkte	
1	As	- 0.028374	0.058172	- 0.028380	0.058287	- 0.028329	0.056997
	Al	- 0.072075	- 0.115138	- 0.072091	- 0.115104	- 0.072074	- 0.115569
2	Al	0.009506	0.021082	0.009517	0.021154	0.009325	0.020504
	As	0.006821	- 0.000709	0.006818	- 0.000674	0.006954	- 0.000944
3	As	0.0	0.0	0.0	0.0	0.0	0.0
	Al	0.0	0.0	0.0	0.0	0.0	0.0
\vec{k} -Punkt-Konvergenz: (9 Atomlagen)							
		(1 × 4 × 4) \vec{k} -Punkte		(2 × 4 × 4) \vec{k} -Punkte		experimentell [Kah83b](ELEED)	
1	As	- 0.028282	0.057963	- 0.028265	0.057917	- 0.058	0.010
	Al	- 0.071709	- 0.114639	- 0.071717	- 0.114754	- 0.087	- 0.152
2	Al	0.009602	0.020905	0.009580	0.020880	0.0	0.015
	As	0.006732	- 0.001057	0.006755	- 0.001108	0.0	- 0.015
3	As	0.0	0.0	0.0	0.0	0.0	0.0
	Al	0.0	0.0	0.0	0.0	0.0	0.0

Tabelle 6.10: AlAs (110) Strukturrelaxation: Konvergenz mit der Anzahl der Atomlagen in der Superzelle und mit der Anzahl der \vec{k} -Punkte: Es werden jeweils die beiden obersten Atomlagen relaxiert. Der ebene Wellen Cutoff beträgt $g_{max} = 3.5\sqrt{\text{Ry}}$.

III-V (110) (1 × 1)-Relaxation							
Konvergenz mit der Anzahl der relaxierten Lagen							
Lage	Atom	δz [a]	δx [b]	δz [a]	δx [b]	δz [a]	δx [b]
		AlAs $a = 5.6290\text{Å}, b = \frac{a}{\sqrt{2}}$		InAs $a = 6.0335\text{Å}, b = \frac{a}{\sqrt{2}}$		InP $a = 5.8549\text{Å}, b = \frac{a}{\sqrt{2}}$	
Supercelle: 5 Atomlagen, 4 relaxiert							
1	An	- 0.028374	0.058172	- 0.032466	0.059214	- 0.029588	0.055102
	Ka	- 0.072075	- 0.115138	- 0.070975	- 0.113138	- 0.068057	- 0.107651
2	Ka	0.009506	0.021082	0.009070	0.017495	0.008023	0.016739
	An	0.006821	- 0.000709	0.006901	- 0.003872	0.006637	- 0.003382
Supercelle: 7 Atomlagen, 6 relaxiert							
1	An	- 0.031832	0.063741	- 0.035664	0.062204	- 0.033397	0.057441
	Ka	- 0.076436	- 0.112298	- 0.075252	- 0.113637	- 0.072899	- 0.108603
2	Ka	0.004162	0.028739	0.003278	0.023079	0.001745	0.021665
	An	0.002130	0.001137	0.001905	- 0.005820	0.001124	- 0.005365
3	An	- 0.005060	0.008959	- 0.004904	0.006770	- 0.004961	0.006411
	Ka	- 0.001496	- 0.001804	- 0.002760	- 0.004955	- 0.003097	- 0.005397
Supercelle: 9 Atomlagen, 8 relaxiert							
1	An	- 0.029025	0.067205	- 0.032207	0.063082	- 0.029991	0.058064
	Ka	- 0.073909	- 0.109636	- 0.072149	- 0.113863	- 0.069787	- 0.108993
2	Ka	0.006546	0.032666	0.006169	0.024550	0.004659	0.022792
	An	0.004644	0.003574	0.004896	- 0.006376	0.004111	- 0.006081
3	An	- 0.002259	0.013493	- 0.001515	0.008814	- 0.001638	0.008014
	Ka	0.001270	0.000160	0.000603	- 0.006254	0.000192	- 0.006742
4	Ka	0.002252	0.005999	0.002204	0.004148	0.002075	0.003695
	An	0.001025	0.000207	0.002010	- 0.002752	0.001991	- 0.002700

Tabelle 6.11: III-V (110) Strukturrelaxation: Konvergenz Atomverschiebungen mit der Anzahl der relaxierten Atomlagen: In jeder Zelle bleibt nur die mittlere Atomlage unrelaxiert. Der ebene Wellen Cutoff beträgt $g_{max} = 3.5\sqrt{\text{Ry}}$.

III-V (110) (1 × 1)-Relaxation								
Konvergenz mit der Anzahl der relaxierten Lagen								
Lage	$\Delta_{i,\perp}$ [b]	$\Delta_{i,x}$ [$\frac{3}{4}a$]	$d_{i,i+1,\perp}$ [$\frac{1}{2}b$]	$d_{i,i+1,x}$ [$\frac{1}{2}a$]	$\Delta_{i,\perp}$ [b]	$\Delta_{i,x}$ [$\frac{3}{4}a$]	$d_{i,i+1,\perp}$ [$\frac{1}{2}b$]	$d_{i,i+1,x}$ [$\frac{1}{2}a$]
AlAs				InAs				
Gitterkonstanten: $a = 5.6290$, $b = \frac{a}{\sqrt{2}}$				Gitterkonstanten: $a = 6.0335$, $b = \frac{a}{\sqrt{2}}$				
Superzelle: 5 Atomlagen, 4 relaxiert								
1	0.173310	1.058268	0.727560	1.163162	0.172352	1.051346	0.738734	1.16009
2	0.021791	0.996420	0.998582	1.019012	0.021367	0.997108	0.992256	1.01814
Superzelle: 7 Atomlagen, 6 relaxiert								
1	0.176039	1.059473	0.717926	1.161196	0.175840	1.052784	0.726568	1.157060
2	0.027601	0.997292	0.984358	1.011315	0.028899	0.998169	0.974819	1.012077
3	0.010763	0.995247	0.996392	1.002991	0.011725	0.997141	0.990090	1.005520
Superzelle: 9 Atomlagen, 8 relaxiert								
1	0.176843	1.059845	0.715392	1.160910	0.176945	1.053256	0.723173	1.156636
2	0.029093	0.997464	0.980162	1.010551	0.030926	0.998302	0.969619	1.011132
3	0.013333	0.995294	0.988321	1.001963	0.015069	0.997176	0.979196	1.003202
4	0.005792	0.998364	1.000414	1.004504	0.006900	0.999741	0.994497	1.004409
InP								
Gitterkonstanten: $a = 5.8549$, $b = \frac{a}{\sqrt{2}}$								
Superzelle: 5 Atomlagen, 4 relaxiert								
1	0.162753	1.051293	0.751219	1.152161				
2	0.020121	0.998152	0.993237	1.016047				
Superzelle: 7 Atomlagen, 6 relaxiert								
1	0.166044	1.052669	0.739465	1.149288				
2	0.027030	0.999172	0.976449	1.009685				
3	0.011808	0.997515	0.989205	1.006194				
Superzelle: 9 Atomlagen, 8 relaxiert								
1	0.167057	1.053060	0.736431	1.148891				
2	0.028873	0.999270	0.971811	1.008934				
3	0.014756	0.997559	0.979126	1.003765				
4	0.006395	0.999889	0.994600	1.004149				

Tabelle 6.12: III-V (110) Strukturrelaxation: Konvergenz der Atomabstände mit der Anzahl der relaxierten Atomlagen: In jeder Zelle bleibt nur die mittlere Atomlage unrelaxiert. Der ebene Wellen Cutoff beträgt $g_{max} = 3.5\sqrt{Ry}$.

InAs (110) (1 × 1)-Relaxation							
Konvergenz mit der Anzahl der ebenen Wellen							
Lage	Atom	δz [a]	δx [b]	δz [a]	δx [b]	δz [a]	δx [b]
mit $g_{max} = 3.5\sqrt{Ry}$ berechnete Gitterkonstanten: $a = 6.03347\text{\AA}$, $b = \frac{a}{\sqrt{2}}$, $\varphi = 90^\circ$							
		$g_{max} = 3.5\sqrt{Ry}$		$g_{max} = 4.0\sqrt{Ry}$		$g_{max} = 4.5\sqrt{Ry}$	
1	As	- 0.032466	0.059214	- 0.031893	0.057945	- 0.031705	0.055701
	In	- 0.070975	- 0.113138	- 0.071062	- 0.115864	- 0.071283	- 0.117743
2	In	0.009070	0.017495	0.009176	0.016909	0.008843	0.015758
	As	0.006901	- 0.003872	0.007292	- 0.004672	0.007356	- 0.005374
3	As	0.0	0.0	0.0	0.0	0.0	0.0
	In	0.0	0.0	0.0	0.0	0.0	0.0
mit $g_{max} = 4.5\sqrt{Ry}$ berechnete Gitterkonstanten: $a = 6.01264\text{\AA}$, $b = \frac{a}{\sqrt{2}}$, $\varphi = 90^\circ$							
		$g_{max} = 3.5\sqrt{Ry}$		$g_{max} = 4.0\sqrt{Ry}$		$g_{max} = 4.5\sqrt{Ry}$	
1	As	- 0.033504	0.066337	- 0.032800	0.065099	- 0.032608	0.062761
	In	- 0.070468	- 0.106821	- 0.070450	- 0.109559	- 0.070748	- 0.111548
2	In	0.009970	0.020731	0.010105	0.020222	0.009702	0.019012
	As	0.006538	- 0.001568	0.007000	- 0.002416	0.007038	- 0.003132
3	As	0.0	0.0	0.0	0.0	0.0	0.0
	In	0.0	0.0	0.0	0.0	0.0	0.0

Tabelle 6.13: InAs (110) Strukturrelaxation: Konvergenz mit der Anzahl der ebenen Wellen: Es wurden jeweils die beiden obersten Atomlagen relaxiert. Strukturrelaxationen werden mit verschiedenen g_{max} durchgeführt. Mit $g_{max} = 3.5\sqrt{Ry}$ und $g_{max} = 4.5\sqrt{Ry}$ berechnete Gitterkonstanten werden verwendet.

III-V	Ref.	a_0 [Å]	$\Delta_{1,\perp}$ [b]	$\Delta_{1,x}$ [$\frac{3}{4}a$]	$d_{12,\perp}$ [$\frac{1}{2}b$]	$d_{12,x}$ [$\frac{1}{2}a$]	$\Delta_{2,\perp}$ [b]	$\Delta_{2,x}$ [$\frac{3}{4}a$]
Bulk			0.0000	1.0000	1.0000	1.0000	0.0000	1.0000
InAs	[Kah83b]	6.0585	0.182	1.097	0.706	1.188	0.033	–
	[MDC85]	6.0585	0.184	1.055	0.622	1.116	–	–
	[AHS91]	6.036	0.183	1.101	0.701	1.192	0.033	–
	wir,12.25Ry	6.0335	0.176	1.052	0.727	1.153	0.029	0.998
	wir,20.25Ry	6.0126	0.174	1.051	0.739	1.161	0.022	0.996
	[AHS91],8Ry	5.844	0.169	1.062	0.708	1.150	0.030	–
	[AHS91],18Ry	5.861	0.181	1.061	0.697	1.159	0.031	–
InP	[Kah83b]	5.8695	0.166	1.041	0.663	1.177	0.034	–
	[Cha79]	5.8695	0.157	1.041	0.740	1.136	0.027	–
	[MDC85]	5.8695	0.195	1.034	0.624	1.119	–	–
	[AHS91]	5.869	0.176	1.045	0.747	1.152	0.034	–
	wir,12.25Ry	5.8549	0.166	1.053	0.740	1.149	0.027	0.999
	[AHS91],8Ry	5.678	0.159	1.063	0.721	1.142	0.027	–
	[AHS91],18Ry	5.662	0.167	1.064	0.717	1.136	0.029	–
AlAs	[Kah83b]	5.66	0.162	1.038	0.665	1.173	0.030	–
	wir,12.25Ry	5.6290	0.176	1.059	0.718	1.161	0.028	0.997

Tabelle 6.14: Atomare Abstände der Oberflächenatome: Verglichen werden experimentelle Werte (obere Zeilen) aus LEED-Experimenten mit unseren und anderen Pseudopotentialrechnungen (untere Zeilen)

III-V	Ref.	ω [°]	α [°]	β [°]	γ [°]	$\delta c_1 a_1$ [%]	$\delta c_2 a_1$ [%]	$\delta c_1 a_2$ [%]
bulk		0.0	109.5	109.5	109.5	0.00	0.00	0.00
InAs	[Kah83b]	31.0	87.5	116.4	121.8	- 3.9	2.1	1.4
	[MDC85]	32.0	91.0	110.3	124.8	- 0.5	- 2.8	-
	[AHS91]	36.6	87.3	117.0	121.5	-4.2	2.0	1.5
	wir,12.25Ry	30.6	89.1	110.7	124.3	- 0.74	0.63	- 0.95
	wir,20.25Ry	30.2	89.0	110.5	124.3	- 0.65	0.92	- 0.91
	[AHS91],8Ry	30.7	90.1	112.5	123.3	- 1.80	- 0.22	- 2.00
	[AHS91],18Ry	32.0	88.9	111.4	123.9	- 1.18	- 0.18	- 1.82
InP	[Kah83b]	28.1	90.3	109.7	124.8	- 0.2	- 7.4	- 1.7
	[Cha79]	26.5	91.7	110.4	123.8	- 0.6	- 1.0	- 2.1
	[MDC85]	31.8	89.7	106.6	126.6	1.8	- 3.3	-
	[AHS91]	29.9	89.0	109.5	124.7	- 0.0	1.6	0.5
	wir,12.25Ry	29.2	90.2	111.4	123.7	-1.18	0.58	-1.02
	[AHS91],8Ry	29.2	91.4	113.7	122.4	-2.47	-0.04	-2.28
	[AHS91],18Ry	30.1	90.4	112.5	123.2	-1.78	-0.23	-1.89
AlAs	[Kah83b]	27.3	90.8	109.5	124.9	-0.04	-7.84	-2.29
	wir,12.25Ry	31.2	89.0	111.6	123.9	-1.29	0.38	-1.30

Tabelle 6.15: Relaxationswinkel und Änderung der Bindungslängen: Verglichen werden experimentelle Werte (obere Zeilen) aus LEED-Experimenten mit unseren und anderen Pseudopotentialrechnungen (untere Zeilen) $\delta c_i a_j$ ist die Bindungslängenänderung zwischen dem Kation der i'ten Lage und dem Anion der j'ten Lage.

6.6 Die elektronische Struktur der Oberfläche

Am Beispiel von InP wird die elektronische Struktur der (110) Oberfläche im Detail untersucht. Supercellen mit 9, 13 und 17 Atomlagen werden verwandt. Die beiden obersten Atomlagen einer jeden Oberfläche sind relaxiert, für die inneren 5, 9 bzw. 13 Atomlagen werden die idealen Bulkpositionen angenommen. Abgeschlossen wird die Supercelle auf jeder Seite mit 4 Atomlagen entsprechendem Vakuum. Die Strukturrelaxation wurde in der 9-Lagen-Zelle durchgeführt. Die relaxierten Atompositionen dieser Zelle werden fixiert und für die größeren Zellen übernommen.

Da die Supercelle endlich ist, enthält sie eine diskrete, endliche Anzahl von elektronischen Zuständen. Diese werden parallel zur Oberfläche durch den zweidimensionalen Wellenvektor \vec{k}_{\parallel} und den Bandindex ν charakterisiert.

Für die Berechnung des selbstkonsistenten Potentials bzw. der selbstkonsistenten Ladungsdichte wird die Energiedispersion der elektronischen Zustände bezüglich \vec{k}_{\parallel} genauso wie bei der atomaren Strukturrelaxation durch vier spezielle \vec{k} -Punkte (siehe Kapitel 3.2) in der irreduziblen Oberflächenbrillouinzone (irr-OBZ) beschrieben. Dies entspricht 16 \vec{k} -Punkten in der totalen OBZ.

$$\vec{k}_{ij} = \frac{i}{8}\vec{k}_y + \frac{j}{8}\vec{k}_z, \quad i = 1, 3, \quad j = 1, 3 \quad (6.2)$$

Mit diesem Potential werden die Größen, die die elektronische Struktur der Oberfläche beschreiben, ausgerechnet. Dies sind die Bandstruktur $\epsilon_{\vec{k}_{\parallel}, \nu}$, die Ladungsdichtekonturen $|\Psi_{\vec{k}_{\parallel}, \nu}(\vec{r})|^2$ der einzelnen Zustände $\Psi_{\vec{k}_{\parallel}, \nu}(\vec{r})$, sowie die aus diesen berechnete lokale Zustandsdichte $\rho_{lok}(\vec{r}, E)$.

$$\rho_{lok}(\vec{r}, E) = \sum_{\nu} \int_{\text{irr-OBZ}} d\vec{k}_{\parallel} |\Psi_{\vec{k}_{\parallel}, \nu}(\vec{r})|^2 \delta(E - \epsilon_{\vec{k}_{\parallel}, \nu}) \quad (6.3)$$

Die \vec{k}_{\parallel} -Dispersion der elektronischen Struktur wird an dieser Stelle durch 6×6 gleichverteilte \vec{k} -Punkte in der irr-OBZ beschrieben.

$$\vec{k}_{ij} = \frac{i}{10}\vec{k}_y + \frac{j}{10}\vec{k}_z, \quad i = 0..5, \quad j = 0..5 \quad (6.4)$$

Dies entspricht 100 \vec{k} -Punkten in der totalen OBZ. Die \vec{k} -Punkt-Integration wird durch gewichtete Summation ausgeführt, wobei die Eckpunkte das Gewicht 0.01, die Randpunkte 0.02 und die Punkte im Inneren der irr-OBZ das Gewicht 0.04 bekommen.

Es existieren mehrere prinzipielle Schwierigkeiten, die $\Psi_{\vec{k}_{\parallel}, \nu}(\vec{r})$ und die $\epsilon_{\vec{k}_{\parallel}, \nu}$ als "wahre" Wellenfunktionen und Energiebänder zu interpretieren:

- Strenggenommen sind die Eigenwerte und die Wellenfunktionen in der Dichtefunktionaltheorie lediglich mathematische Hilfsgrößen. Ähnlich dem Koopman'schen Theorem der Hartree-Fock-Theorie ist es trotzdem im allgemeinen möglich, Eigenwerte und Einteilchenwellenfunktionen als solche zu interpretieren [JG89]. Viele Jahre praktischer Erfahrung mit der DFT zeigen, daß dies immer der Fall ist, solange die Zustände nicht zu sehr lokalisiert sind. Da sogar die im Vergleich zu Oberflächenzuständen viel stärker lokalisierten d-Zustände der Übergangsmetalle noch vernünftig beschrieben werden, ist dies sicher auch für alle Zustände unseres Systems der Fall.
- Unsere Oberflächen-Superzelle ist senkrecht zur Oberfläche endlich. Insbesondere besteht sie aus endlich vielen Lagen Bulk. Völlig analog zum Problem wechselwirkungsfreier Elektronen in einem eindimensionalen Potentialkasten stellen sich Quantentrogzustände ("Quantum-well-states") ein, deren einhüllende Kontur von der Länge des Bulks (bzw. der Länge des Kastens) abhängt. Dies ist völlig unphysikalisch. Trotzdem wird die elektronische Struktur des Bulks in gewisser Weise "richtig" durch diese Quantentrogzustände beschrieben (siehe Kapitel 6.6.1).
- Die Superzelle ist spiegelsymmetrisch zur Mittelebene σ_x . Wegen dieser Spiegelsymmetrie wird für alle Zustände Symmetrie bzw. Antisymmetrie bezüglich dieser Spiegelebene erzwungen. Oberflächenzustände, deren Überlapp im Bulk vernachlässigbar ist, sind somit zweifach entartet (symmetrisch und antisymmetrisch). Oberflächenresonanzen, deren Überlapp nicht vernachlässigbar ist, spalten energetisch auf. Auch dies ist unphysikalisch. Nun ist es aber so, daß die unphysikalische Anschlußbedingung von der Mitte der Superzelle aus zum Rand hin "ausheilt", d.h. die Ladungsdichtekonturen der Zustände im Bereich der Oberfläche und darüber hinaus deren Abfall ins Vakuum hinein sind sehr wohl physikalisch interpretierbar und damit auch die lokalen Zustandsdichten (siehe Kapitel 6.6.2).

6.6.1 Quantentrogzustände

Die Variation der Bulkzustände der Oberflächen-Superzelle in [110]-Richtung (\perp zur Oberfläche) ähnelt sehr stark den Zuständen eines eindimensionalen (1d-)Potentialkastens:

$$V_L(x) = \begin{cases} 0 & 0 \leq x \leq L \\ \infty & x > L \end{cases}, \quad \Phi_{E_{Ln}}(x) = \sqrt{\frac{2}{L}} \sin(\sqrt{E_n} x), \quad E_{Ln} = \left(\frac{\pi n}{L}\right)^2 \quad (6.5)$$

Dies ist in Abbildung 6.8 festgehalten. Links sind die Zustände des 1d-Potentialkastens $|\Phi_{E_{Ln}}(x)|^2$ skizziert, rechts jeweils die Zustände der großen InP-Oberflächen-Superzelle (17 Atomlagen). Die Ladungsdichtekontur der Superzellenzustände wurde dabei über die (110) Ebenen integriert, um die Quantentrogstruktur zu verdeutlichen:

$$\Psi_{\vec{k}_{\parallel}, \nu}(\vec{r}) \longrightarrow \Phi_{\vec{k}_{\parallel}, \nu}(x) = \sqrt{\int dy dz |\Psi_{\vec{k}_{\parallel}, \nu}(x, y, z)|^2} \quad (6.6)$$

Die Einhüllenden der niedrigsten besetzten Bulkzustände sehen, abgesehen von der durch die Atomlagen bestimmten Feinstruktur, den Zuständen des Kastenpotentials sehr ähnlich. Die obersten besetzten Bulkzustände zeigen ebenfalls deutlich das Verhalten von Quantentrogzuständen.

Die Frage, inwieweit die Quantentrogzustände Bulkeigenschaften repräsentieren, läßt sich am einfachen Modell des 1d-Potentialkastens analytisch studieren.

Zunächst wird die diskrete Summation über die Zustände $\Phi_{E_{Ln}}$ durch eine Energieintegration ersetzt:

$$\sum_n \rightarrow \frac{1}{\Delta n} \int \frac{dn}{dE_{Ln}} \Big|_{E_{Ln}=E} dE \rightarrow \int dE \frac{L}{2\pi\sqrt{E}} \quad (6.7)$$

Der Kontinuumsimes ($L \rightarrow \infty$) der diskreten lokalen Zustandsdichte $\rho_{lok}^{(L,1d)}$ läßt sich mit (6.7) und (A.49) leicht analytisch vornehmen:

$$\rho_{lok}^{(L,1d)}(x, E) = \sum_n |\Phi_{E_{Ln}}(x)|^2 \delta(E - E_{Ln}) \quad (6.8)$$

$$\downarrow$$

$$\rho_{lok}^{(\infty,1d)}(x, E) = \frac{L}{2\pi\sqrt{E}} |\Phi_E(x)|^2 = \frac{1}{2\pi\sqrt{E}} (1 - \cos(2\sqrt{E}x)) \quad (6.9)$$

Durch diesen Kontinuumsimes wird das halbunendliche 1d-Elektronengas mit der ($x=0$)-Grenzfläche beschrieben. Dieser Limes oszilliert mit $q = 2\sqrt{E}$ um die räumlich konstante lokale Zustandsdichte $\frac{1}{2\pi\sqrt{E}}$ des freien 1d-Elektronengases. Zu erwarten ist aber, daß der Kontinuumsimes für große x , i.e. ($0 \ll x \ll L$) in irgendeiner Form gegen das unendlich ausgedehnte System konvergiert.

Die lokale Zustandsdichte ist als Distribution zu verstehen, d.h. nur integrierte Größen sind physikalisch von Bedeutung. Dies sind z.B. die Ladungsdichte, aber auch die integrierte lokale Zustandsdichte $\mathcal{Z}_{lok}^{(1d)}(x, E)$.

$$\mathcal{Z}_{lok}(x, E) = \int_0^E dE' \rho_{lok}(x, E') \quad (6.10)$$

Unter Zuhilfenahme einiger trigonometrischer Formeln (siehe Anhang A: Seite 167)

ergibt sich:

$$\begin{aligned} \mathcal{Z}_{lok}^{(L,1d)}(x, N) &= \sum_{n=1}^N |\Phi_{E_{Ln}}(x)|^2 = \frac{N}{L} \left(1 - \frac{\sin\left(\frac{N\pi}{L}x\right)}{N \sin\left(\frac{\pi}{L}x\right)} \cos\left(\left(N+1\right)\frac{\pi}{L}x\right) \right) \\ &= \frac{N}{L} \left(1 + \frac{1}{2N} - \frac{\sin\left(\left(2N+1\right)\frac{\pi}{L}x\right)}{2N \sin\left(\frac{\pi}{L}x\right)} \right) \end{aligned} \quad (6.11)$$

$$\begin{aligned} \mathcal{Z}_{lok}^{(\infty,1d)}(x, E) &= \int_0^E dE \rho_{lok}^{(\infty,1d)}(x, E) = \frac{\sqrt{E}}{\pi} \left(1 - \frac{\sin(\sqrt{E}x)}{\sqrt{E}x} \cos(\sqrt{E}x) \right) \\ &= \frac{\sqrt{E}}{\pi} \left(1 - \frac{\sin(2\sqrt{E}x)}{2\sqrt{E}x} \right) \end{aligned} \quad (6.12)$$

Für große Zahlen N und mit $x \ll L$ läßt sich der $\sin\left(\frac{\pi}{L}x\right)$ im Nenner entwickeln und die diskrete Gleichung (6.11) geht mit Ordnung $\mathcal{O}\left(\frac{1}{N}\right)$ in den Kontinuumsimes (6.12) über:

$$\mathcal{Z}_{lok}^{(L,1d)}(x, N) \rightarrow \mathcal{Z}_{lok}^{(\infty,1d)}(x, E = \left(\frac{N\pi}{L}\right)^2) \quad (6.13)$$

Für große x konvergiert $\mathcal{Z}_{lok}^{(\infty,1d)}(x, E)$ mit Ordnung $\mathcal{O}\left(\frac{1}{x}\right)$ gegen den räumlich konstanten Wert des freien 1d-Elektronengases $\frac{\sqrt{E}}{\pi}$. Die diskrete $\mathcal{Z}_{lok}^{(L,1d)}(x, N)$ konvergiert ebenfalls für große x , i.e. ($0 \ll x \ll L$), mit Ordnung $\mathcal{O}\left(\frac{1}{x}, \frac{1}{N}\right)$ gegen den räumlich konstanten Wert des freien 1d-Elektronengases $\frac{\sqrt{E}}{\pi}$.

Kontinuumseigenschaften werden also immer dann vernünftig beschrieben, wenn über genügend viele diskrete Zustände $\Phi_{E_{Ln}}$ summiert (gemittelt, integriert) wird. Für große Entfernungen x von der Oberfläche werden nur diejenigen Eigenschaften des unendlich ausgedehnten Systems erreicht, die durch Integration über genügend viele Zustände beschrieben werden. Dies ist z.B. bei der eindimensionalen lokalen Zustandsdichte $\rho_{lok}^{(\infty,1d)}(x, E)$ nicht der Fall.

Im dreidimensionalen Festkörper hingegen mit zweidimensionaler Oberfläche konvergiert weit weg von der Oberfläche auch die lokale Zustandsdichte gegen die des Bulks. Der dafür notwendige Mittelungsprozess wird durch die \vec{k}_{\parallel} -Dispersion erreicht.

Um nur die Diskretisierung senkrecht zur Oberfläche zu analysieren, kann die \vec{k}_{\parallel} -Dispersion durch ein freies 2d-Elektronengas in der yz -Ebene simuliert werden. Die Energie eines Zustands setzt sich also aus dem diskreten Quantentrog-Beitrag E_{Ln} und dem kontinuierlichen Beitrag E_{\parallel} des 2d-Elektronengases zusammen. Die Zustandsdichte des 2d-Elektronengases ist gleichverteilt in E_{\parallel} . Für die lokale Zustandsdichte $\rho_{lok}^{(L,3d)}(x, E)$ ist also über alle E_{Ln} zu summieren (im Limes ($L \rightarrow \infty$))

zu integrieren), für die es ein positives E_{\parallel} gibt mit $E = E_{\parallel} + E_{Ln}$, i.e. über alle ($E_{Ln} \leq E$). Unter Zuhilfenahme einiger trigonometrischer Formeln (siehe Anhang A: Seite 167) ergibt sich:

$$\begin{aligned} \rho_{lok}^{(L,3d)}(x, E) &\sim \sum_{n=1}^{N_E} |\Phi_{E_{Ln}}(x)|^2 = \frac{N_E}{L} \left(1 - \frac{\sin\left(\frac{N_E \pi x}{L}\right)}{N_E \sin\left(\frac{\pi x}{L}\right)} \cos\left(\frac{(N_E + 1)\pi x}{L}\right) \right) \\ &= \frac{N_E}{L} \left(1 + \frac{1}{2N_E} - \frac{\sin\left(\frac{(2N_E + 1)\pi x}{L}\right)}{2N_E \sin\left(\frac{\pi x}{L}\right)} \right) \end{aligned} \quad (6.14)$$

$$\begin{aligned} \rho_{lok}^{(\infty,3d)}(x, E) &\sim \int_0^E dE' \rho_{lok}^{(\infty,1d)}(x, E') = \frac{\sqrt{E}}{\pi} \left(1 - \frac{\sin(\sqrt{E}x)}{\sqrt{E}x} \cos(\sqrt{E}x) \right) \\ &= \frac{\sqrt{E}}{\pi} \left(1 - \frac{\sin(2\sqrt{E}x)}{2\sqrt{E}x} \right) \end{aligned} \quad (6.15)$$

Dabei ist N_E die größte natürliche Zahl mit $\left(\frac{\pi}{L} N_E \leq \sqrt{E}\right)$.

Für große Zahlen N_E und für $x \ll L$ geht die diskrete Gleichung (6.14) mit Ordnung $\mathcal{O}\left(\frac{1}{N_E}\right)$ in den Kontinuumsfall (6.15) über:

$$\rho_{lok}^{(L,3d)}(x, E) \rightarrow \rho_{lok}^{(\infty,3d)}(x, E) \quad (6.16)$$

Für große x konvergiert $\rho_{lok}^{(\infty,3d)}(x, E)$ mit Ordnung $\mathcal{O}\left(\frac{1}{x}\right)$ gegen den räumlich konstanten Wert des freien 3d-Elektronengases $\sim \frac{\sqrt{E}}{\pi}$. Die diskrete $\rho_{lok}^{(L,3d)}(x, E)$ konvergiert ebenfalls für große x , i.e. ($0 \ll x \ll L$), mit Ordnung $\mathcal{O}\left(\frac{1}{x}, \frac{1}{N}\right)$ gegen den räumlich konstanten Wert des freien 3d-Elektronengases $\sim \frac{\sqrt{E}}{\pi}$.

Für die integrierte lokale Zustandsdichte $\mathcal{Z}_{lok}^{(3d)}(x, E)$ ergibt sich unter Zuhilfenahme einiger trigonometrischer Formeln (siehe Anhang A: Seite 167):

$$\begin{aligned} \mathcal{Z}_{lok}^{(L,3d)}(x, E) &= \int_0^E dE \rho_{lok}^{(L,3d)}(x, E) = \sum_{n=1}^{N_E} (E - E_n) |\Phi_{E_{Ln}}(x)|^2 \quad (6.17) \\ &= \frac{\sqrt{E_{N_E}}}{3\pi} \left\{ 3E \left(1 + \frac{1}{2N_E} \right) - \sqrt{E_{N_E + \frac{1}{2}} E_{N_E + 1}} \right\} \\ &\quad + \frac{\sin\left(2\sqrt{E_{N_E + \frac{1}{2}}}x\right)}{2\pi \left(\frac{L}{\pi} \sin \frac{\pi x}{L}\right)} \left\{ \sqrt{E_{N_E + \frac{1}{2}} E_{N_E + 1}} - E \right\} \\ &\quad + \frac{\sqrt{E_{N_E + 1}}}{8\pi \left(\frac{L}{\pi} \sin \frac{\pi x}{L}\right)^2} \left\{ 3 \cos\left(2\sqrt{E_{N_E + 1}}x\right) + \cos\left(2\sqrt{E_{N_E}}x\right) \right\} \\ &\quad - \frac{1}{8\pi \left(\frac{L}{\pi} \sin \frac{\pi x}{L}\right)^3} \left\{ \sin\left(2\sqrt{E_{N_E + \frac{1}{2}}}x\right) + \sin\left(2\sqrt{E_{N_E + \frac{3}{2}}}x\right) \right\} \end{aligned}$$

$$\begin{aligned} \mathcal{Z}_{lok}^{(\infty,3d)}(x, E) &= \int_0^E dE \rho_{lok}^{(\infty,3d)}(x, E) \\ &= \frac{2}{3\pi} (E)^{\frac{3}{2}} \times \left(1 + \frac{3 \cos(2\sqrt{E}x)}{(2\sqrt{E}x)^2} - \frac{3 \sin(2\sqrt{E}x)}{(2\sqrt{E}x)^3} \right) \end{aligned} \quad (6.18)$$

Wieder ist N_E die größte natürliche Zahl mit $(\frac{\pi}{2}N_E \leq \sqrt{E})$.

Für große Zahlen N_E und mit $x \ll L$ geht die diskrete Gleichung (6.17) geht mit Ordnung $\mathcal{O}\left(\frac{1}{N_E}\right)$ in den Kontinuumsimes (6.18) über:

$$\mathcal{Z}_{lok}^{(L,3d)}(x, E) \rightarrow \mathcal{Z}_{lok}^{(\infty,3d)}(x, E) \quad (6.19)$$

Für große x konvergiert $\mathcal{Z}_{lok}^{(\infty,3d)}(x, E)$ sogar quadratisch mit Ordnung $\mathcal{O}\left(\frac{1}{x^2}\right)$ gegen den räumlich konstanten Wert des freien 3d-Elektronengases $\sim E^{\frac{3}{2}}$. Die diskrete $\mathcal{Z}_{lok}^{(L,3d)}(x, E)$ konvergiert ebenfalls für große x , i.e. $(0 \ll x \ll L)$, mit Ordnung $\mathcal{O}\left(\frac{1}{x^2}, \frac{1}{N}, \frac{1}{Nx}\right)$ gegen den räumlich konstanten Wert des freien 3d-Elektronengases $\sim E^{\frac{3}{2}}$.

Für die Relaxation der atomaren Struktur ist entscheidend, wie gut die Bulkladungsdichte in der Mitte der Superzelle erreicht wird. Die schnelle Konvergenz der Gesamtladungsdichte $n^{(\infty,3d)}(x) = \mathcal{Z}_{lok}^{(\infty,3d)}(x, E_F)$ gegen das freie 3d-Elektronengas mit Ordnung $\mathcal{O}\left(\frac{1}{x^2}\right)$ läßt sich auf die Superzelle übertragen. Die Relaxation der atomaren Oberflächenstruktur konvergiert quadratisch mit der Dicke des Slabs. Die elektronische Struktur hingegen konvergiert nur linear mit der Slabdicke. Auch hier läßt sich die Konvergenz von $\rho_{lok}^{(\infty,3d)}(x, E)$ mit Ordnung $\mathcal{O}\left(\frac{1}{x}\right)$ gegen das freie 3d-Elektronengas auf die Superzelle übertragen.

Der Vergleich zwischen den freien Elektronen im Potentialkasten und der Superzellenergebnisse ist in Abbildung 6.7 auf Seite 107 festgehalten. Alle Größen des Potentialkastens sind auf die entsprechenden Größen des freien 3d-Elektronengases normiert. Ein Wert von 1 bedeutet also freies 3d-Elektronengas.

Das Bild oben links zeigt $\mathcal{Z}_{lok}^{(\infty,3d)}(x, E)$ und $\rho_{lok}^{(\infty,3d)}(x, E)$ bei einer Energie von $(E = 0.5\text{Ry})$. Das kleine eingefügte Bild zeigt die Abschätzungen $(2\sqrt{E}x)^1$ und $(2\sqrt{E}x)^2$ des absoluten Fehlers für $\rho_{lok}^{(\infty,3d)}(x, E)$ bzw. $\mathcal{Z}_{lok}^{(\infty,3d)}(x, E)$. So erreicht z.B. $\mathcal{Z}_{lok}^{(\infty,3d)}(x, E)$ bei $x = 20$ a.u. das freie 3d-Elektronengas schon auf 1% genau.

Die mittlere und die rechte Spalte zeigen, ebenfalls bei $(E = 0.5\text{Ry})$, die diskreten $\rho_{lok}^{(L,3d)}(x, E)$ und $\mathcal{Z}_{lok}^{(L,3d)}(x, E)$ für verschiedene Kastenlängen⁵ L . Die unteren Bilder dieser Spalten zeigen die über das Energieintervall von 0.45 Ry bis 0.5 Ry integrierten lokalen Zustandsdichten, ebenfalls für die verschiedenen Kastenlängen L . Während

⁵.....: $L=20$ a.u., ...---...: $L=36$ a.u., - - -: $L=68$ a.u., —: $\lim L \rightarrow \infty$

die Bilder der mittleren Spalte die absoluten Größen zeigen, zeigen die Bilder der rechten Seite die Abweichungen vom Kontinuumslimites.

Die beiden Bilder links mitte und links unten zeigen die selbstkonsistent berechnete Ladungsdichten von Oberflächen-Superzellen mit verschiedenen Slabdicken⁶. Das Bild links mitte zeigt die absoluten Ladungsdichten $\rho_{surf}(x)$. Sie liegen so genau übereinander, daß man die verschiedenen Graphen nicht trennen kann. Das Bild links unten zeigt deshalb die Differenzen $\Delta\rho_{surf}(x)$ zum System mit der größten Slabdicke (17 Lagen). Schon die Ladungsdichte der 9-Lagen-Superzelle ist bis auf wenige Promille auskonvergiert.

Die Abbildung 6.9 zeigt für die verschieden großen Superzellen die Ladungsdichtekontur 6 a.u. über dem Oberflächenanion. Die Ladungsdichtekontur ist sehr schön auskonvergiert. Mit dem Auge sind keinerlei Unterschiede mehr zu erkennen.

Insgesamt lassen sich die wichtigsten Ergebnisse wie folgt zusammenfassen:

- Die Ergebnisse des Potentialkastens lassen sich qualitativ auf die Oberflächen-Superzelle übertragen: Die Ladungsdichte des Bulks und damit das Bulkpotential werden mit zunehmendem Abstand zur Oberfläche quadratisch mit Ordnung $\mathcal{O}\left(\frac{1}{x^2}\right)$ erreicht. Die elektronische Struktur, i.e. die lokale Zustandsdichte konvergiert nur linear mit Ordnung $\mathcal{O}\left(\frac{1}{x}\right)$ gegen die lokale Zustandsdichte des Bulks. Die atomare Relaxation der Oberfläche wird also schon mit relativ kleinen Superzellen hinreichend genau beschrieben. Für die Beschreibung der elektronischen Struktur der Oberfläche sind größere Superzellen notwendig.
- Aus dem gleichen Grund wird der Abfall der Ladungsdichte ins Vakuum schon für relativ kleine Superzellgrößen recht genau beschrieben. Denn genauso wie die Ladungsdichte mit zunehmendem Abstand zur "künstlichen" Randbedingung der Oberfläche quadratisch gegen die Bulkladungsdichte konvergiert, so sollte die Ladungsdichte an der Oberfläche von der "künstlichen" Anschlußbedingung, Spiegelsymmetrie in der Mitte der Superzelle, nur wenig abhängen. Auch der Abfall der lokalen Zustandsdichte ins Vakuum hinein ist im Falle der großen Superzelle ausreichend, um etwa Tunnelbilder zu erklären. Dies wird im folgenden Kapitel 6.6.2 noch genau diskutiert.

⁶.....: 9 Lagen (~ 36 a.u.), ...—...: 13 Lagen (~ 52 a.u.), - - -: 17 Lagen (~ 68 a.u.)

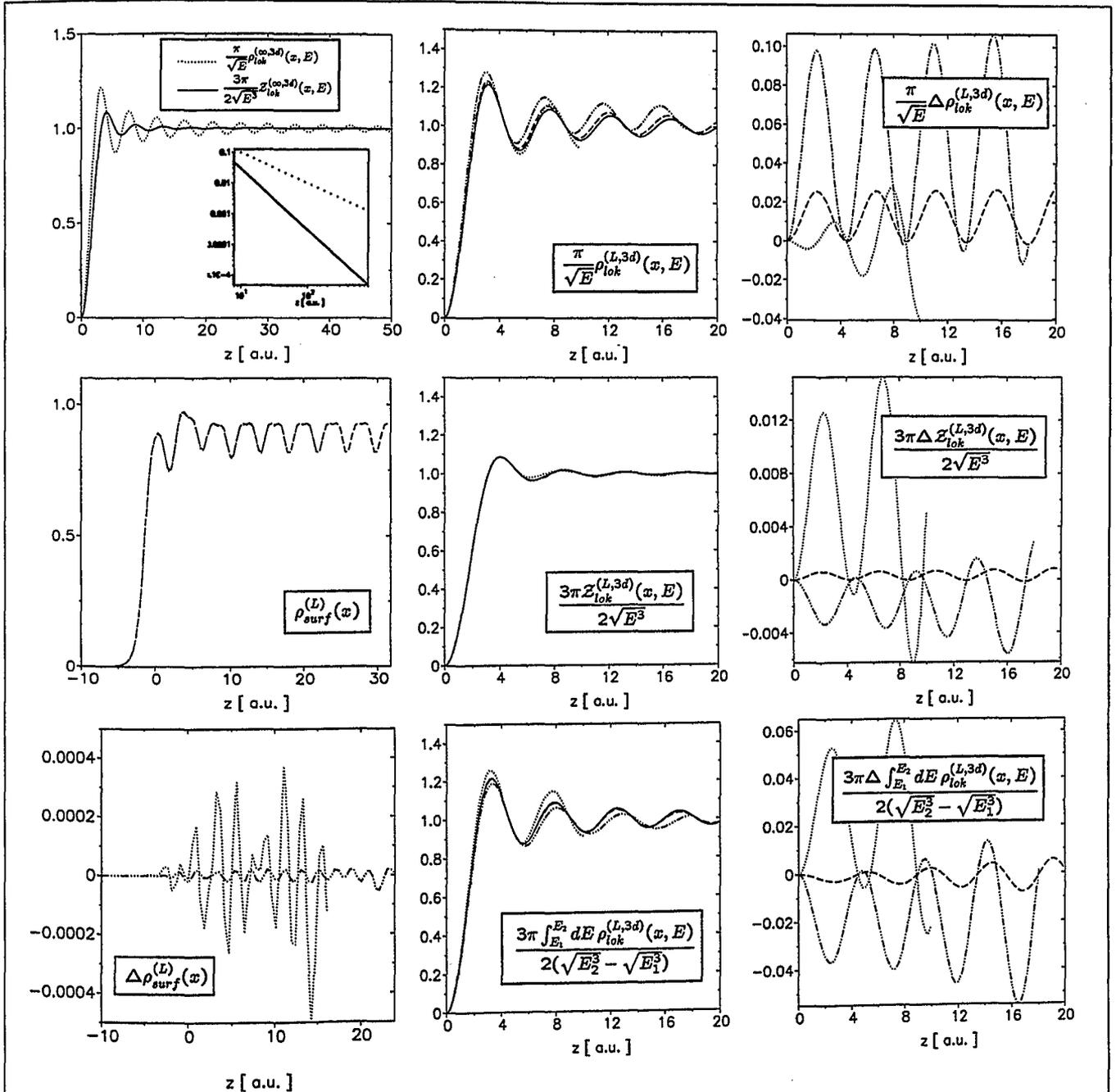


Abbildung 6.7: Vergleich der elektronischen Struktur der Oberflächen-Superzelle mit der von wechselwirkungsfreien Elektronen in einem endlichen Potentialkasten: Die Energien sind $E = E_2 = 0.50\text{Ry}$ und $E_1 = 0.45\text{Ry}$. Δf^L ist eine Abkürzung für die Differenz zwischen diskreter Funktion und Kontinuumslimites $f^{(L)} - f^{(\infty)}$, wobei f für die verschiedenen Funktionen ρ , \mathcal{Z} bzw. $\int dE \mathcal{Z}(E)$ steht. Die genaue Diskussion erfolgt im Text auf Seite 105ff. Die Abbildungen links mitte und links unten zeigen die selbstkonsistenten Ladungsdichten $\rho_{surf}(x)$ aus verschiedenen großen Superzellen. Die verschiedenen Ladungsdichten liegen genau übereinander (Bild links mitte). Daher sind in Bild links unten die Differenzen $\Delta \rho_{surf}(x)$ zur Ladungsdichte der 17-Lagen-Zelle aufgetragen.

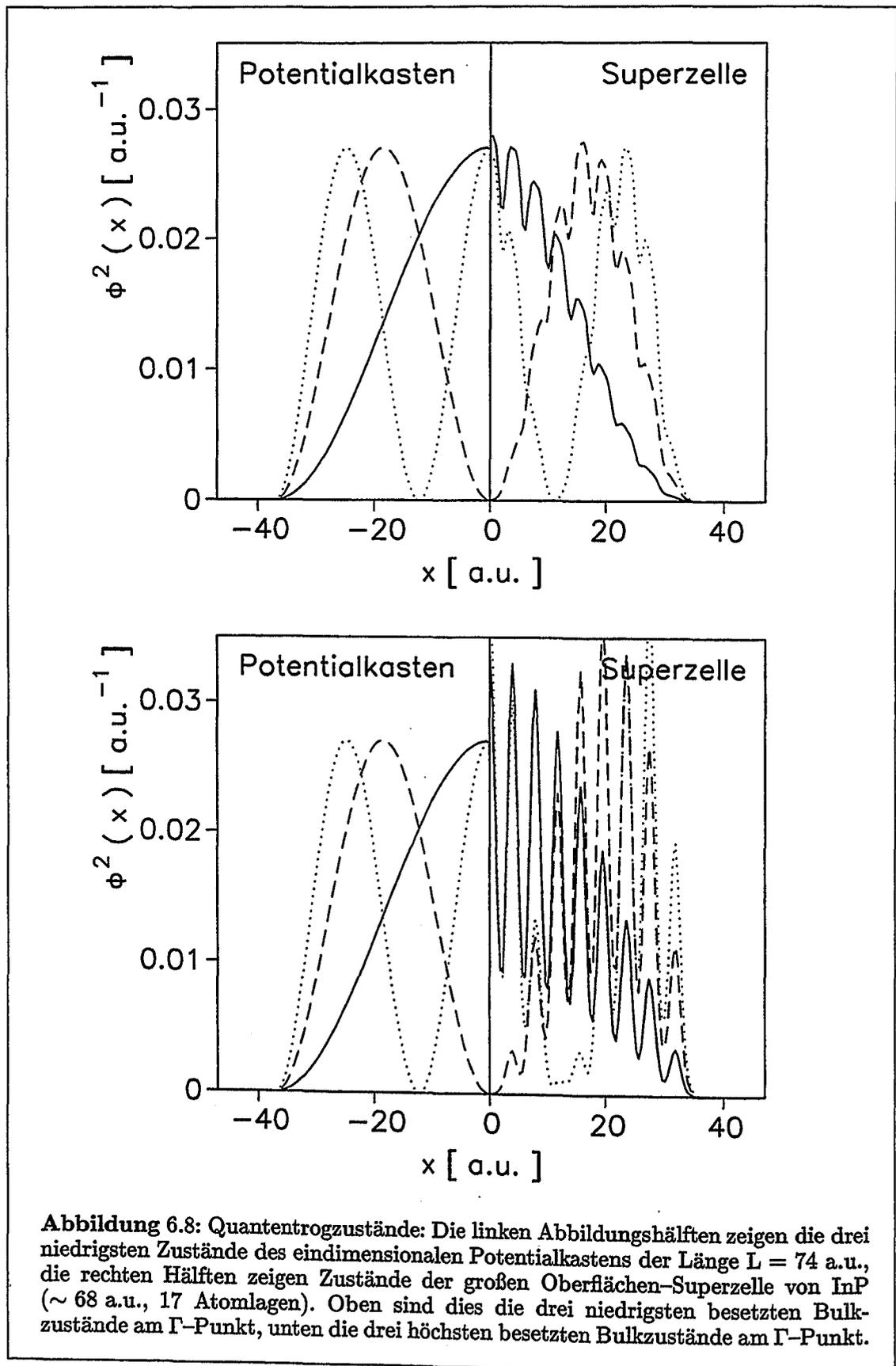
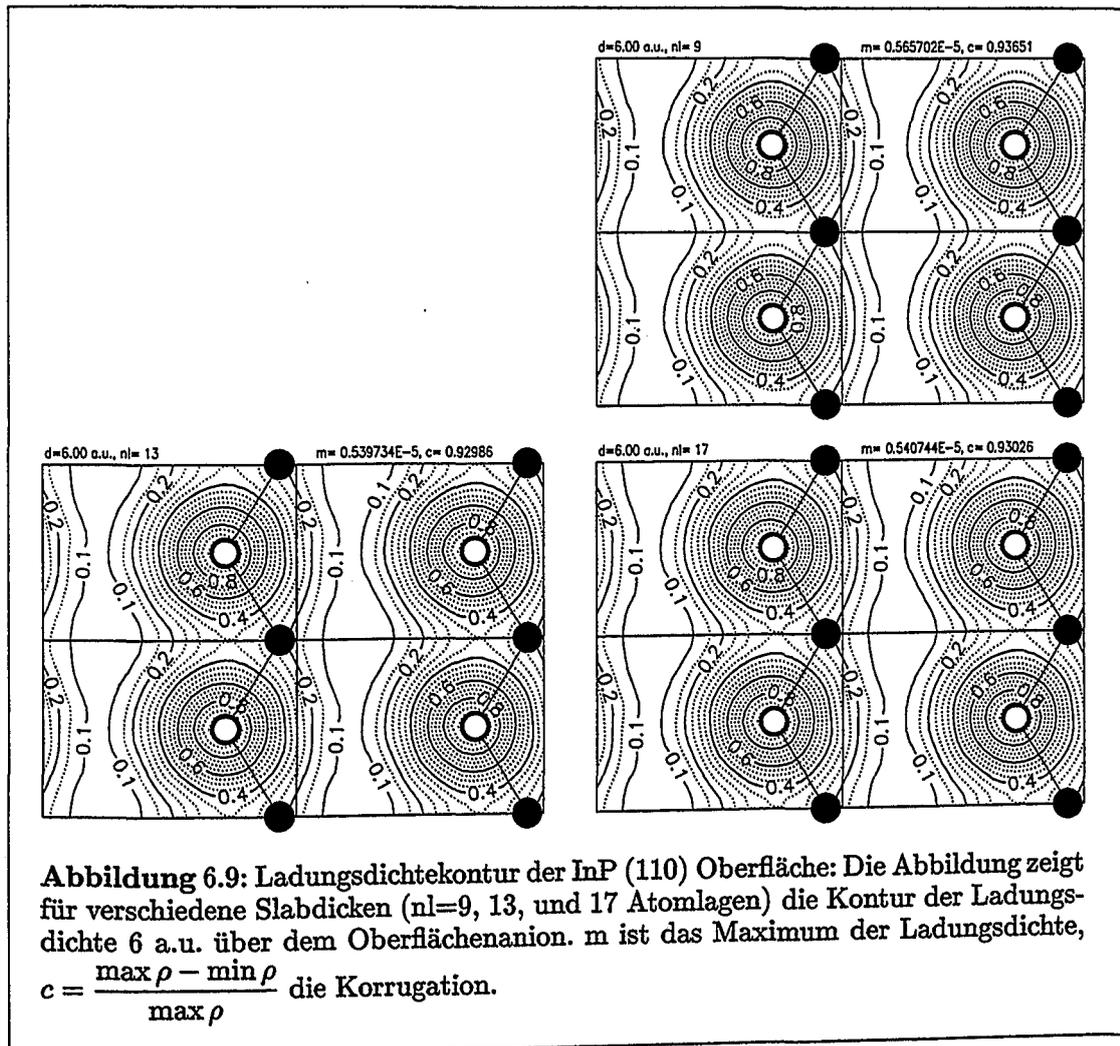


Abbildung 6.8: Quantentrogzustände: Die linken Abbildungshälften zeigen die drei niedrigsten Zustände des eindimensionalen Potentialkastens der Länge $L = 74$ a.u., die rechten Hälften zeigen Zustände der großen Oberflächen-Superzelle von InP (~ 68 a.u., 17 Atomlagen). Oben sind dies die drei niedrigsten besetzten Bulkzustände am Γ -Punkt, unten die drei höchsten besetzten Bulkzustände am Γ -Punkt.



6.6.2 Oberflächenzustände und Oberflächenresonanzen

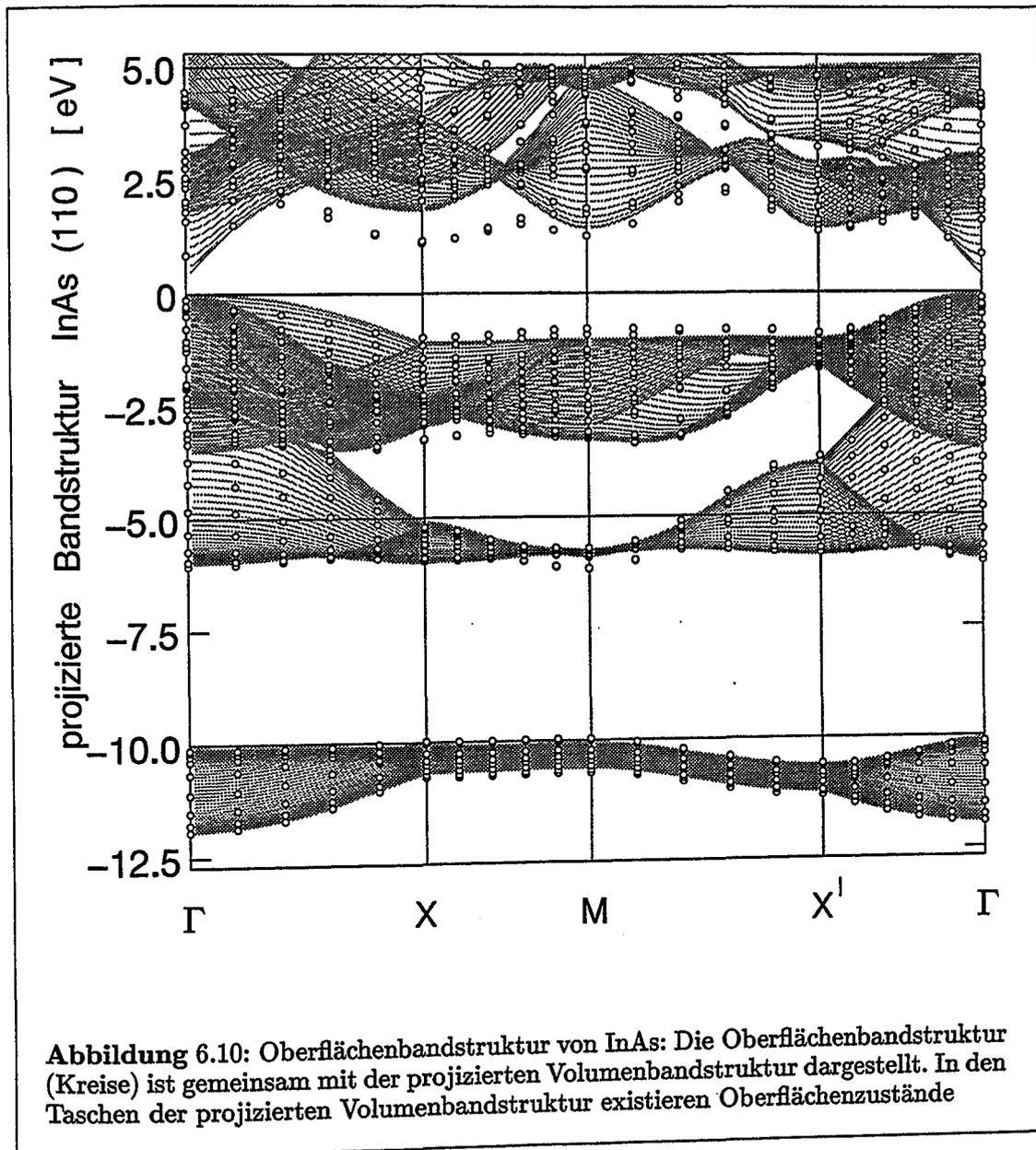
Wie im vorhergehenden Kapitel beschrieben, werden in der Oberflächen-Superzelle die Bulkzustände durch Quantentrogzustände repräsentiert. Startend von den Bandkanten sind diese Zustände nach Knotenzahl sortiert, d.h. die Einhüllende des Quantentrogzustands an der Bandkante hat keinen Knoten innerhalb des Slabs, die Einhüllende des nächsten Zustands hat einen Knoten, usw.. Es existieren aber auch zusätzliche Zustände, die nicht in dieses Schema passen. In den Taschen der projizierten Bulkbandstruktur sind dies Oberflächenzustände. Innerhalb der Bulkbandstruktur rehybridisieren potentielle Oberflächenzustände mit den Bulkzuständen zu Oberflächenresonanzen.

In den Abbildungen 6.10 und 6.11 sind die berechneten Oberflächenbandstrukturen von InAs und InP gemeinsam mit den projizierten Bandstrukturen der jeweiligen Bulkmaterialien aufgetragen. Einige Bänder lokalisierter Oberflächenzustände sind in den Taschen deutlich erkennbar.

Insgesamt existieren also drei verschiedene Arten von Zuständen. Die gitterperiodischen Blochzustände des Bulks werden durch die Oberfläche begrenzt und klingen exponentiell ins Vakuum hinein ab. Oberflächenzustände haben eine große Amplitude an der Oberfläche und klingen exponentiell in den Bulk hinein ab. Oberflächenresonanzen haben ebenso wie Oberflächenzustände eine große Amplitude in der Nähe der Oberfläche, verhalten sich im Bulk aber wie Blochwellen. Wie alle gebundenen Zustände klingen auch Oberflächenzustände und Oberflächenresonanzen exponentiell ins Vakuum hinein ab.

Wegen der Symmetrie der Superzelle (siehe Abb.6.2) – es existieren zwei äquivalente, spiegelsymmetrische (110) Oberflächen – sind alle Oberflächenzustände doppelt vertreten. Oberflächenzustände, deren Überlapp im Bulk vernachlässigbar ist, sind zweifach entartet, je einer dieser Zustände ist symmetrisch bzw. antisymmetrisch bezüglich der σ_x -Ebene. Dies gilt nicht für Oberflächenresonanzen. Diese spalten energetisch auf. In diesem Kapitel sollen die Oberflächenzustände und -resonanzen der (110) Oberfläche von InP untersucht werden.

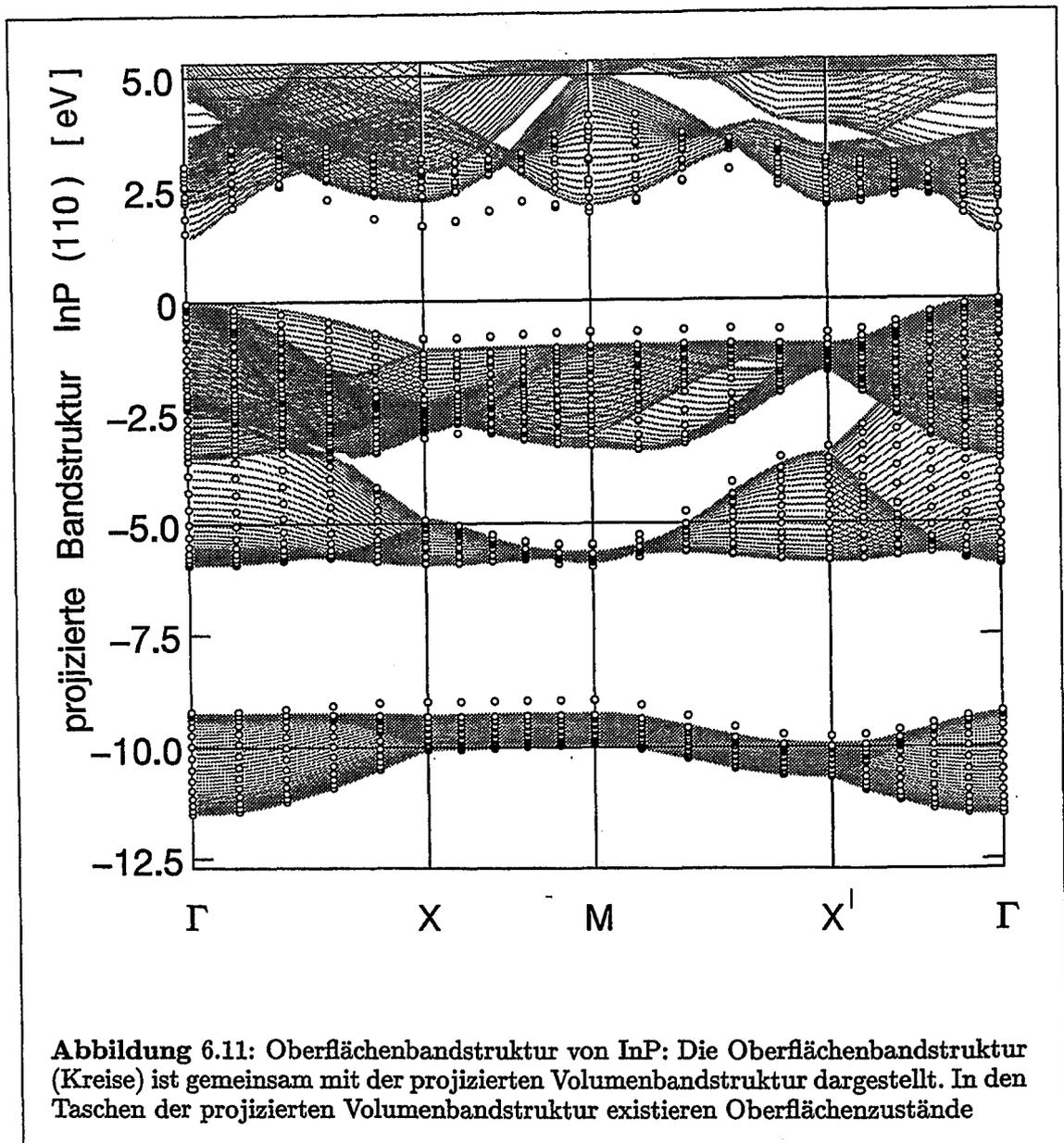
Da dies vor allem im Hinblick auf den späteren Vergleich mit der RTM geschieht, wird vor allem der Abfall der Zustände ins Vakuum hinein untersucht. Zustände, die für die RTM wichtig sind, i.e. die Zustände in einem Energieintervall von einigen eV um das Valenzbandmaximum herum, werden genauer analysiert.



6.6.2.1 Methodik

Im Computerprogramm liegen alle Wellenfunktionen auf einem gleichverteilten, äquidistanten Ortsraumgitter vor, welches durch die Fouriertransformation bestimmt wird. Bei der hier verwendeten Superzelle⁷ ($12\sqrt{2}a_0 \times \sqrt{2}a_0 \times a_0$) und bei einem ebene Wellen Cutoff von $g_{max} = 3.5\sqrt{Ry}$ ergeben sich $216 \times 18 \times 24$ gleichverteilte Stützpunkte im Ortsraum. Alle Wellenfunktionen $\Psi_{\vec{k}\nu}$ sind auf dem diskreten

⁷ $a_0 = 5.854941 \text{ \AA}$



Ortsraumgitter auf 1 normiert:

$$\sum_{x,y,z} |\Psi_{\vec{k}\nu}(x,y,z)|^2 = 1 \quad (6.20)$$

Um die Struktur der einzelnen Zustände zu verdeutlichen, wird deren Ladungsdichtekontur $|\Psi_{\vec{k}\nu}|^2$ in verschiedenen Schnittebenen der Superzelle abgebildet. Dies wird anhand der Prinzipskizze 6.12 dargestellt, welche die Abbildungen der Zustände (6.18 bis 6.23 und 6.26 bis 6.29) erklärt:

Da der Abfall der Zustände ins Vakuum von besonderem Interesse ist, wird die

Ladungsdichtekontur einer (110) Ebene über der Oberfläche gezeigt. Senkrecht dazu werden Schnitte durch zwei verschiedene ($\bar{1}10$)-Ebenen und Schnitte durch zwei verschiedene (001)-Ebenen gezeigt. Das Maximum jedes Konturplots ist einheitlich auf 1 normiert. Zahlenwerte andere Größen wie Maximum, Gewicht und Korrugation der Ebene, sowie \vec{k} -Punkt und Energieeigenwert $\epsilon_{\vec{k}\nu}$ und die Lage der Ebene lassen sich am Rand ablesen.

Die Ladungsdichte einer Ebene hängt von zwei diskreten Koordinaten aus $\{x,y,z\}$ ab. Stellvertretend für beliebige Ebenen wird in den folgenden Gleichungen eine (110) Ebene ($x=\text{konst.}$) genommen, die von y und z abhängt. Dabei werden folgende Größen ausgerechnet, um die Ladungsdichtekontur einer Ebene zu charakterisieren:

$$\text{Maximum:} \quad m = \max_{y,z} |\Psi_{\vec{k}\nu}(x, y, z)|^2 \quad (6.21)$$

$$\text{Gewicht:} \quad s = \sum_{y,z} |\Psi_{\vec{k}\nu}(x, y, z)|^2 \quad (6.22)$$

$$\text{relative Korrugation:} \quad c = \frac{\max_{y,z} |\Psi_{\vec{k}\nu}(x, y, z)|^2 - \min_{y,z} |\Psi_{\vec{k}\nu}(x, y, z)|^2}{\max_{y,z} |\Psi_{\vec{k}\nu}(x, y, z)|^2} \quad (6.23)$$

$$\text{Richtungs-Korrugation:} \quad c_y = \frac{\max_y \sum_z |\Psi_{\vec{k}\nu}(x, y, z)|^2 - \min_y \sum_z |\Psi_{\vec{k}\nu}(x, y, z)|^2}{\max_y \sum_z |\Psi_{\vec{k}\nu}(x, y, z)|^2} \quad (6.24)$$

$$c_z = \frac{\max_z \sum_y |\Psi_{\vec{k}\nu}(x, y, z)|^2 - \min_z \sum_y |\Psi_{\vec{k}\nu}(x, y, z)|^2}{\max_z \sum_y |\Psi_{\vec{k}\nu}(x, y, z)|^2} \quad (6.25)$$

$$\text{Korrugationsverhältnis:} \quad c_{yz} = \frac{c_y}{c_z} \quad (6.26)$$

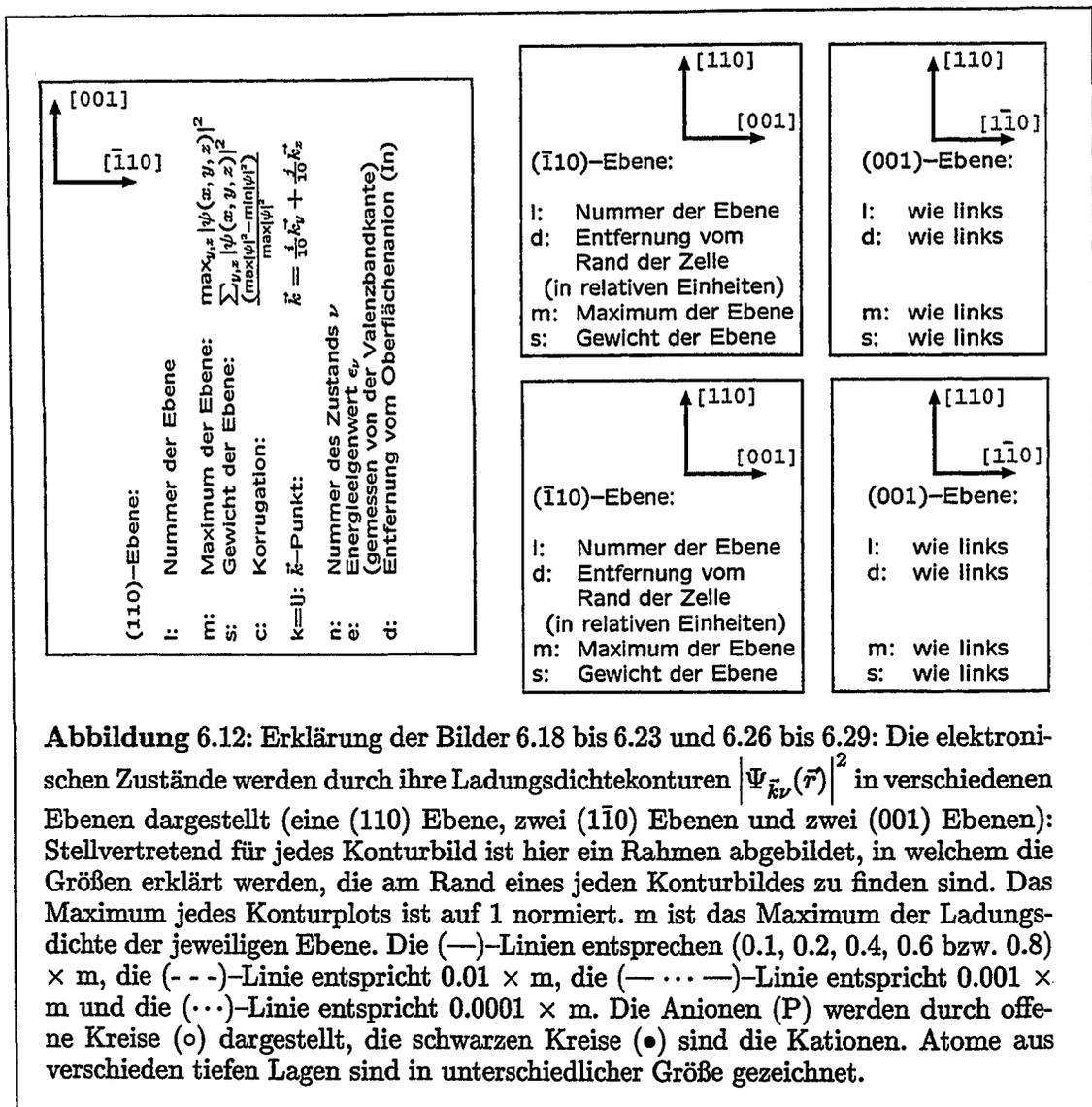
Um den Abfall von Wellenfunktionen ins Vakuum zu analysieren, werden die Betragsquadrate der Wellenfunktionen über Ebenen parallel zur (110) Oberfläche summiert.

$$\Phi_{\vec{k}\nu}^2(x) = \sum_{y,z} |\Psi(x, y, z)|^2 \quad ; \quad (6.27)$$

In genügendem Abstand zur Oberfläche, wo das konstante Vakuumpotential erreicht ist, haben die Wellenfunktionen eine exponentiell abfallende Asymptotik.

$$\Phi_{\vec{k}\nu}(x) \approx C_{\vec{k}\nu} e^{-\kappa_{\vec{k}\nu} x} \quad (6.28)$$

Je nach Abfall ins Vakuum erhalten so unterschiedliche Zustände auch unterschiedliches Gewicht in der lokalen Zustandsdichte über der Oberfläche. Zustände können



wenig Gewicht haben, weil sie schnell abfallen (großes $\kappa_{\vec{k}\nu}$), sie können aber auch wenig Gewicht haben, wenn ihre Amplitude $C_{\vec{k}\nu}$ klein ist.

6.6.2.2 Zustandsdichten und Oberflächenzustandsdichten

Die Zustandsdichten⁸ (DOS(E)), die sich durch Summation über die berechneten, diskreten Eigenzustände ergeben, werden durch Gaußverbreiterung in der Energie

⁸DOS: density of states

der einzelnen Zustände kontinuierlich gemacht:

$$\rho_g(E) = \sum_{\{\vec{k}\nu\}} \omega_{\vec{k}} g_{\vec{k}\nu} \delta(E - \epsilon_{\vec{k}\nu}) \quad (6.29)$$

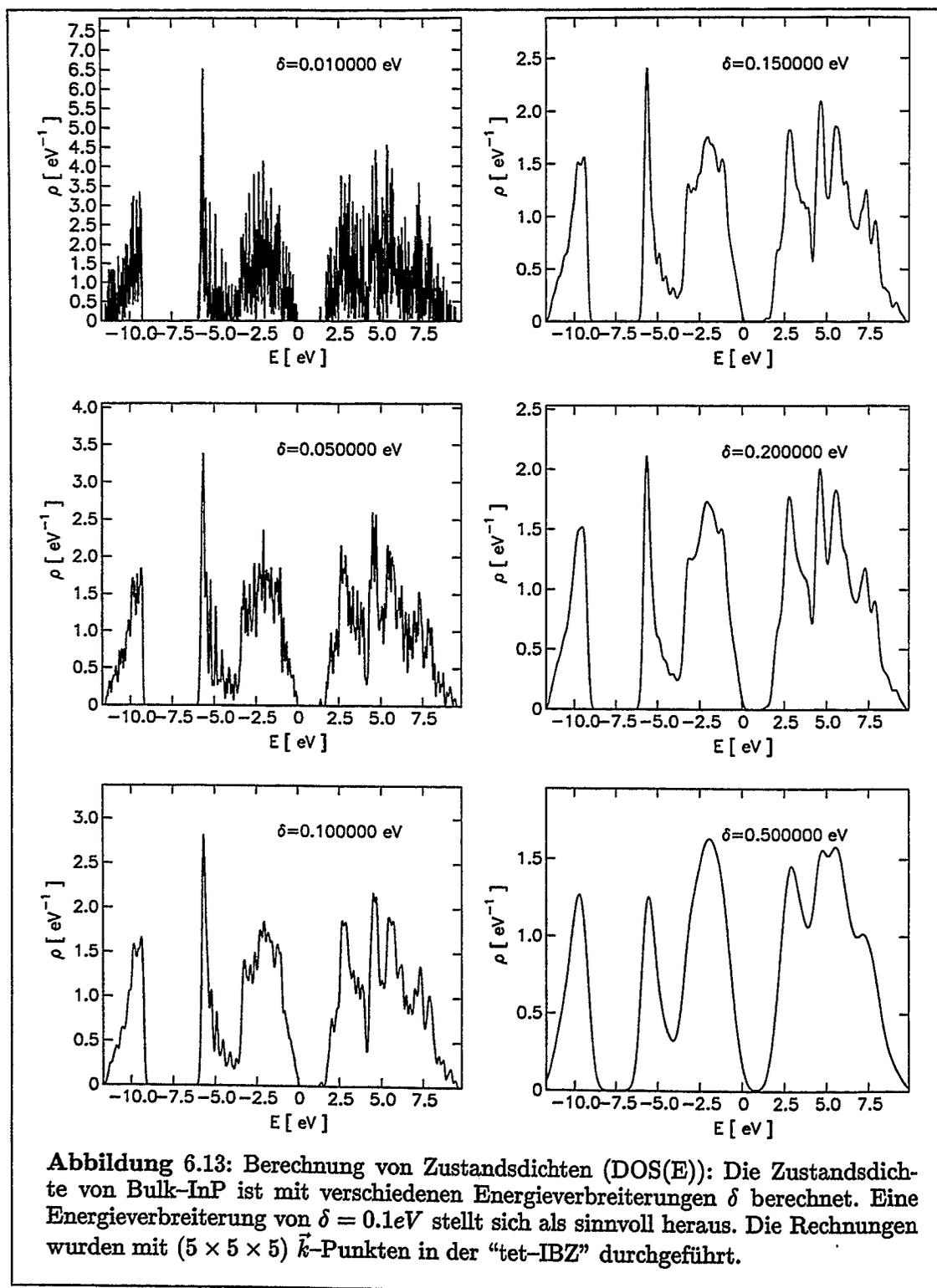
$$\rightarrow \sum_{\{\vec{k}\nu\}} \omega_{\vec{k}} g_{\vec{k}\nu} \frac{1}{\sqrt{\pi\delta}} e^{-\left(\frac{E - \epsilon_{\vec{k}\nu}}{\delta}\right)^2} \quad (6.30)$$

Dabei müssen die unterschiedlichen Gewichte $\omega_{\vec{k}}$ der \vec{k} -Punkte berücksichtigt werden. Die $g_{\vec{k}\nu}$ sind zusätzliche Gewichte. Sie können dazu dienen, Information über die Energieverteilung bestimmter Zustände zu gewinnen. Für die Berechnung von totalen Zustandsdichten sind die $g_{\vec{k}\nu}$ identisch gleich 1. Um Information über die Energieverteilung derjenigen Zustände zu erhalten, welche besonders langsam ins Vakuum hinein abfallen, kann man z.B. die Zustände mit dem Gewicht einer ihrer (110) Ebenen ($x=\text{konst.}$) im Vakuum wichten, i.e. $g_{\vec{k}\nu} = \Phi_{\vec{k}\nu}^2(x)$ (Gleichung (6.27)). Wegen der Diskretheit des \vec{k} -Punkt-Satzes sind die berechneten Zustandsdichten unstetig. Die Energieverbreiterung ist ein Kompromiß zwischen Glattheit und Auflösung von Feinstruktur. Einerseits soll die Unstetigkeit beseitigt werden, andererseits soll aber die physikalisch sinnvolle Feinstruktur noch aufgelöst werden. Die gewählte Energieverbreiterung beträgt $\delta = 0.1$ eV.

Die Verwendung endlicher \vec{k} -Punktsätze für die \vec{k} -Punktintegration hat Einfluß auf die Genauigkeit der Berechnung der selbstkonsistenten Ladungsdichte und der Berechnung der Zustandsdichten. Die unterschiedliche Glattheit der mit verschiedenen großen \vec{k} -Punkt-Sätzen berechneten DOS (Abbildungen 6.13, 6.14 und 6.15) vermittelt einen guten Eindruck von der Genauigkeit bzw. dem Diskretisierungsfehler, der bei der Ersetzung der \vec{k} -Punkt-Integration durch eine Summation über spezielle \vec{k} -Punkte gemacht wird.

Abbildung 6.13 zeigt totale Zustandsdichte (tot-DOS(E)) des Bulks, berechnet mit 125 gleichverteilten \vec{k} -Punkten ($\{\vec{k}_{ijk} = \frac{2i-1}{20}\vec{k}_x + \frac{2j-1}{20}\vec{k}_y + \frac{2k-1}{20}\vec{k}_z, i, j, k = 1..5\}$) in der irreduziblen Brillouinzone der tetragonalen Einheitszelle (tet-IBZ). Die Bulk-tot-DOS ist mit verschiedenen Energieverbreiterungen berechnet. Eine Verbreiterung von 0.01 eV ergibt keine glatte DOS, eine Verbreiterung von 0.5 eV zeigt jedes Band nur noch als Gaußkurve. $\delta = 0.1$ eV stellt sich als sinnvoll heraus.

Die Abbildung 6.14 zeigt ebenfalls die totale Zustandsdichte des Bulks. Die selbstkonsistente Berechnung der Ladungsdichte wurde wie oben mit 125 \vec{k} -Punkten in der tet-IBZ durchgeführt. Zu dieser Ladungsdichte wird die Zustandsdichte nicht selbstkonsistent mit $(10 \times 10 \times 10)$ \vec{k} -Punkten ($\{\vec{k}_{ijk} = \frac{2i-1}{40}\vec{k}_x + \frac{2j-1}{40}\vec{k}_y + \frac{2k-1}{40}\vec{k}_z, i, j, k = 1..10\}$) berechnet.



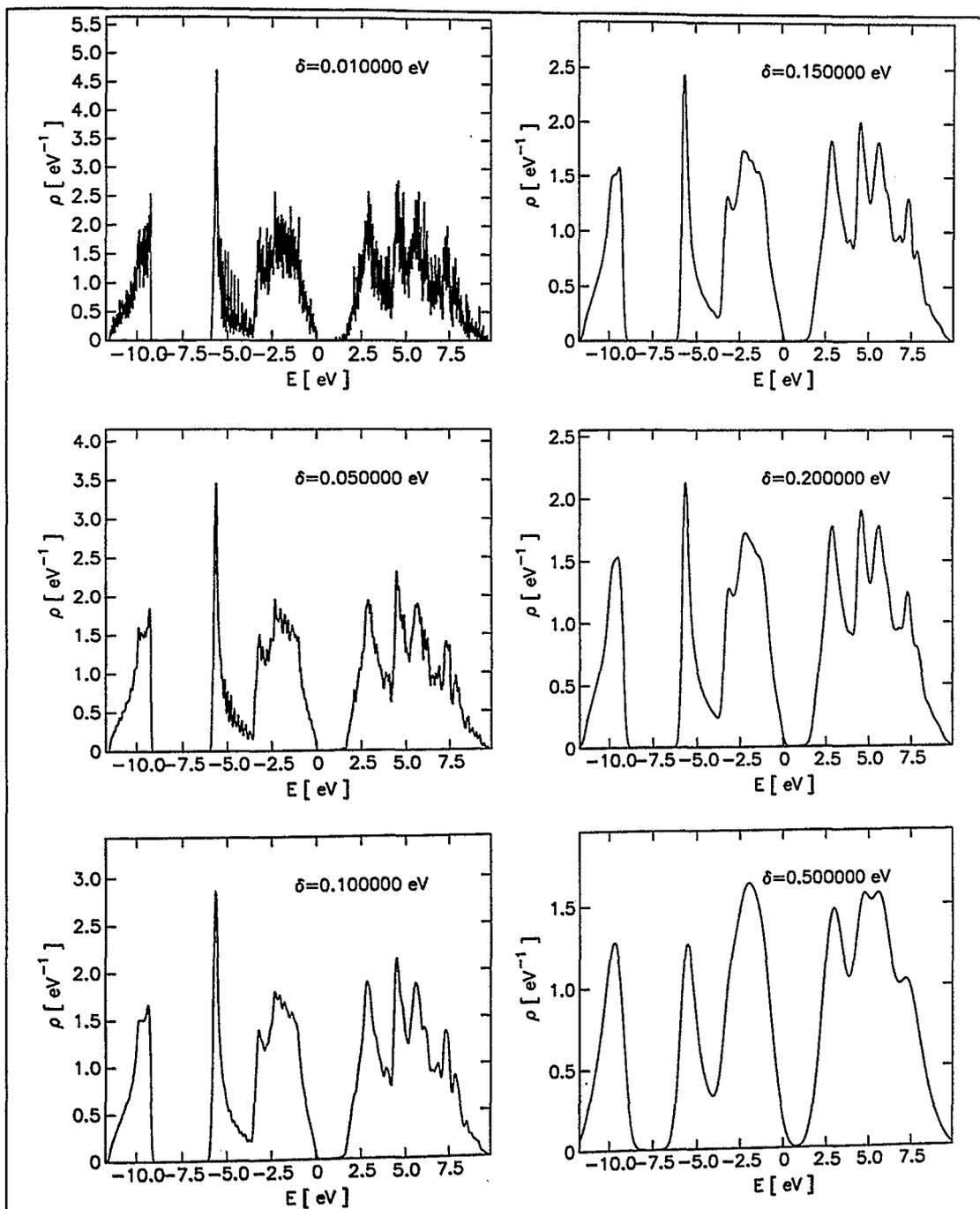


Abbildung 6.14: Berechnung von Zustandsdichten (DOS(E)): Die Zustandsdichte von Bulk-InP ist mit verschiedenen Energieverbreiterungen δ berechnet. Eine Energieverbreiterung von $\delta = 0.1\text{eV}$ stellt sich als sinnvoll heraus. Die Rechnungen wurden mit $(10 \times 10 \times 10)$ \vec{k} -Punkten in der "tet-IBZ" durchgeführt.

Abbildung 6.15 zeigt mit verschiedenen \vec{k} -Punktsätzen berechnete Zustandsdichten. Für die Selbstkonsistenz reichen 3 \vec{k} -Punkte pro Richtung auf jeden Fall aus, die mit 5 \vec{k} -Punkten berechneten Zustandsdichten sind schon recht ordentlich.

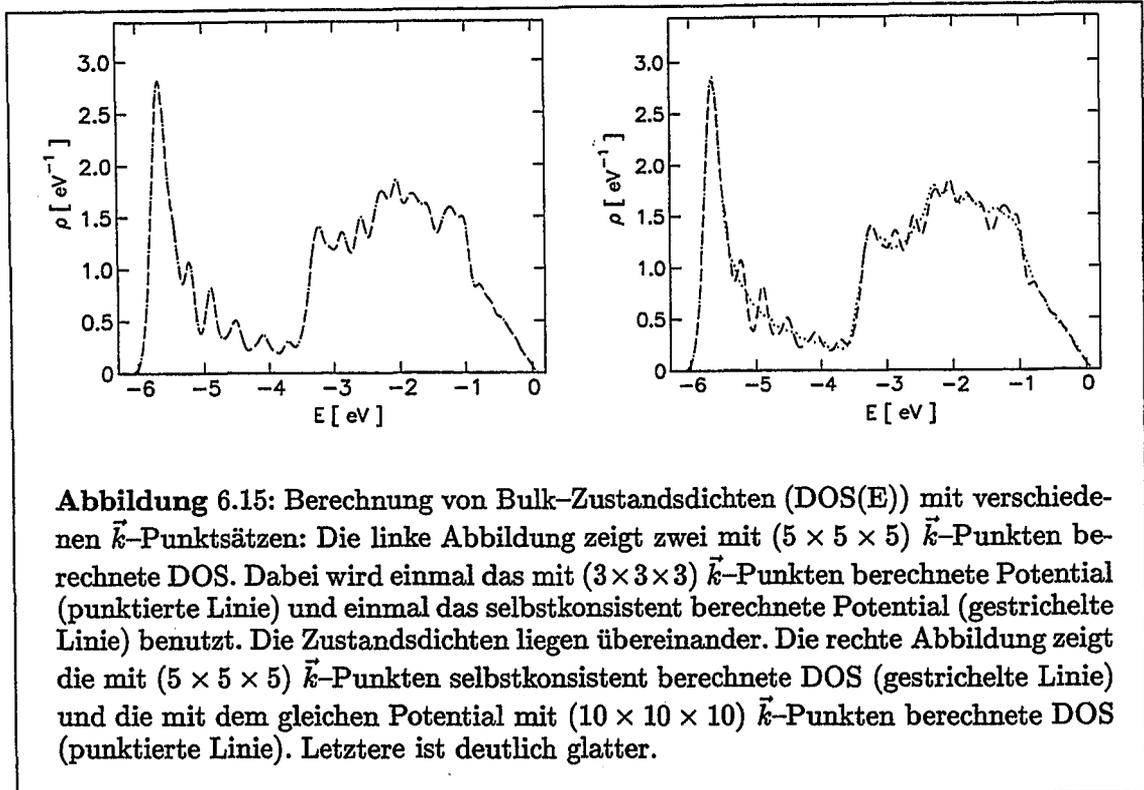


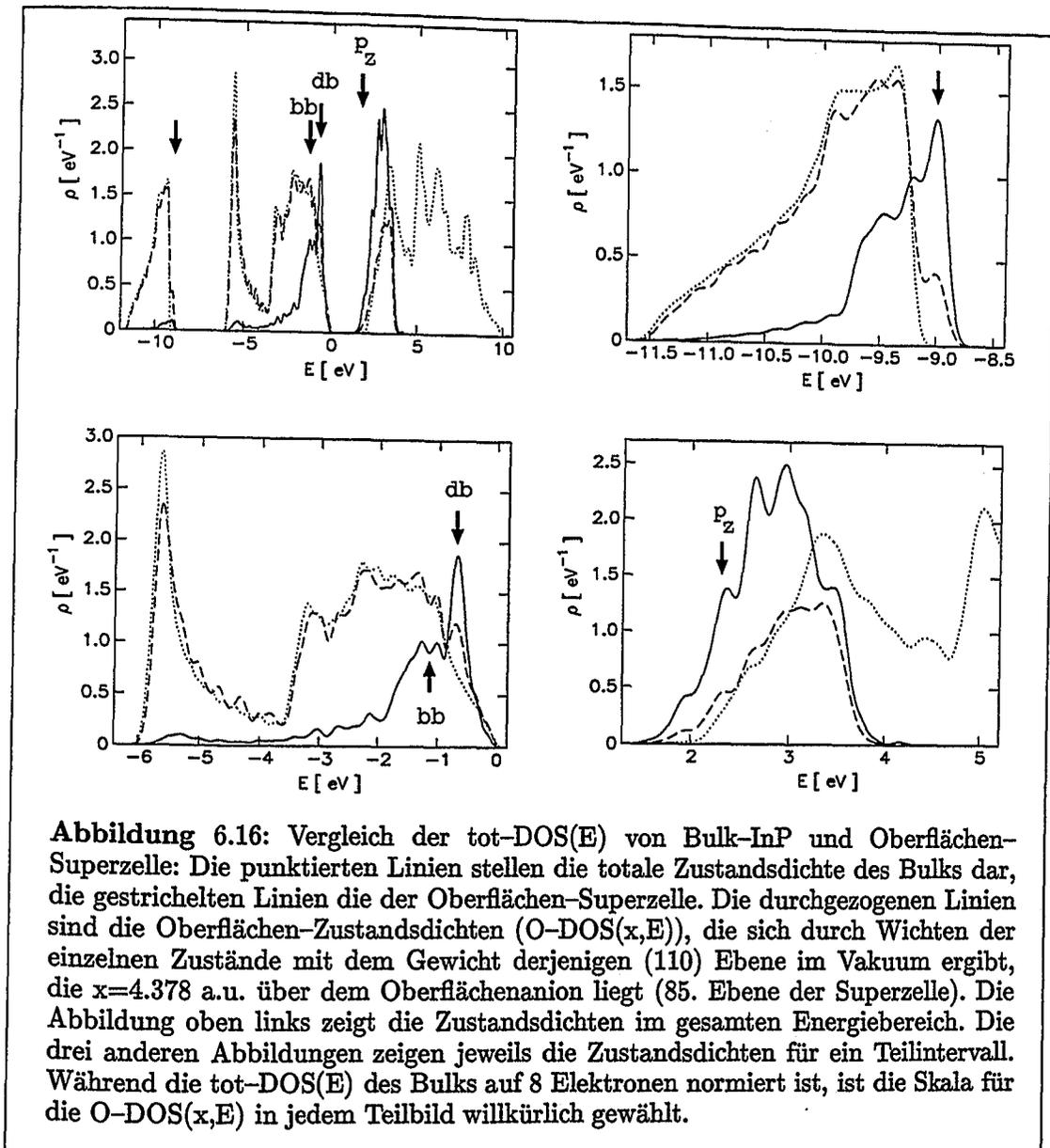
Abbildung 6.15: Berechnung von Bulk-Zustandsdichten (DOS(E)) mit verschiedenen \vec{k} -Punktsätzen: Die linke Abbildung zeigt zwei mit $(5 \times 5 \times 5)$ \vec{k} -Punkten berechnete DOS. Dabei wird einmal das mit $(3 \times 3 \times 3)$ \vec{k} -Punkten berechnete Potential (punktierte Linie) und einmal das selbstkonsistent berechnete Potential (gestrichelte Linie) benutzt. Die Zustandsdichten liegen übereinander. Die rechte Abbildung zeigt die mit $(5 \times 5 \times 5)$ \vec{k} -Punkten selbstkonsistent berechnete DOS (gestrichelte Linie) und die mit dem gleichen Potential mit $(10 \times 10 \times 10)$ \vec{k} -Punkten berechnete DOS (punktierte Linie). Letztere ist deutlich glatter.

In Abbildung 6.16 werden die totale Bulk-DOS und die totale DOS der Oberflächen-Superzelle verglichen. Gleichzeitig ist auch die mit dem Gewicht einer Vakuumlage⁹ gewichtete Zustandsdichte eingezeichnet, um diejenigen Zustände zu markieren, die besonders langsam ins Vakuum hinein abfallen. Diese so gewichteten Zustandsdichten nennen wir zur Abkürzung Oberflächen-DOS (O-DOS(x,E)).

Da bei der Berechnung der Zustandsdichte der Oberflächen-Superzelle nur eine feste Anzahl von 12 unbesetzten Zuständen pro \vec{k} -Punkt berücksichtigt wurde, verschwindet diese oberhalb von 4 eV völlig. Oberhalb von 3 eV sind nicht mehr alle Zustände berücksichtigt.

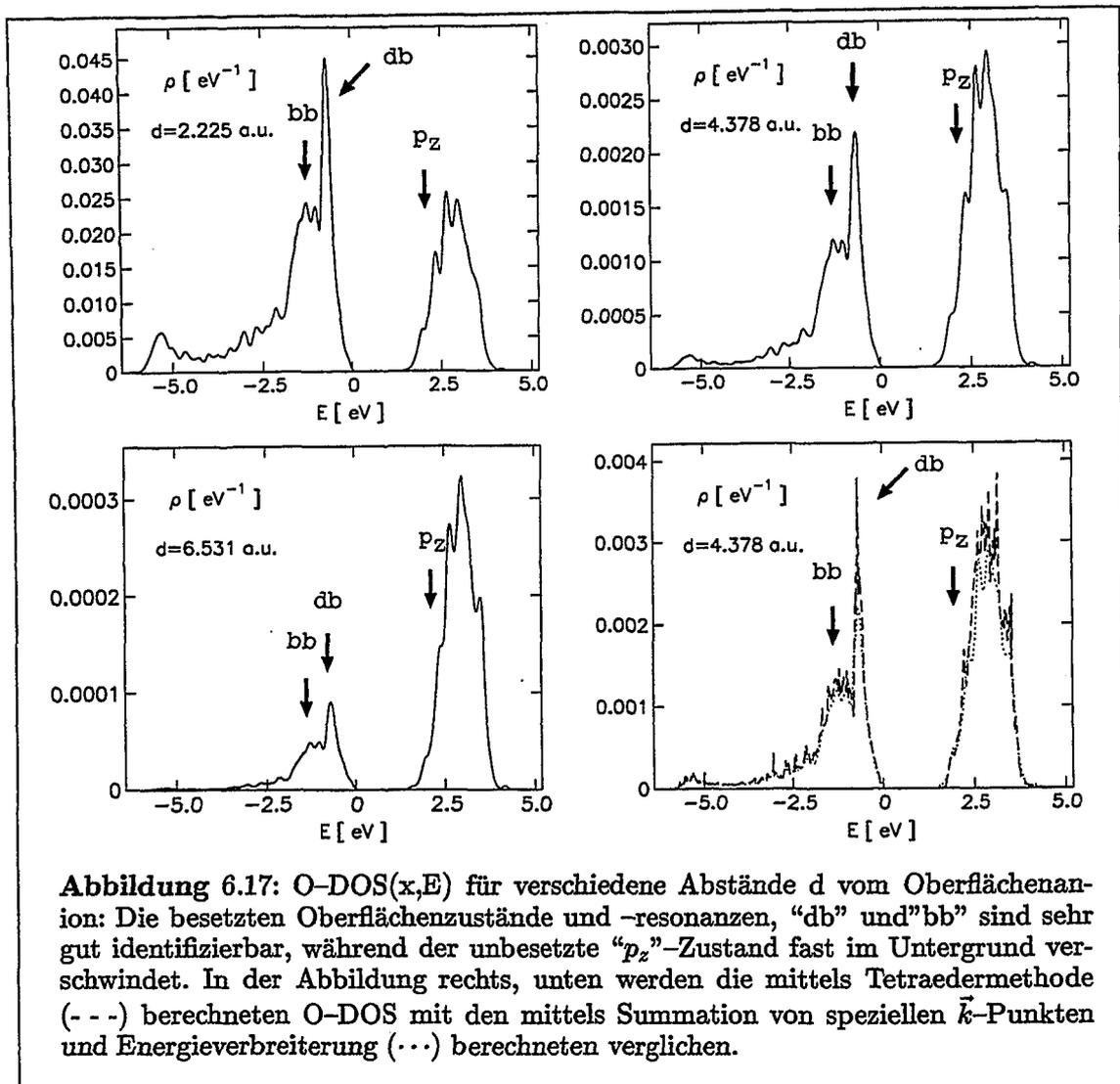
Die Zustandsdichten von Bulk und Oberflächen-Superzelle unterscheiden sich besonders in der Nähe von Bandkanten. Dort sind Oberflächenzustände und Oberflächenresonanzen lokalisiert. Sie zeigen sich als mehr oder weniger scharfe Peaks in den Zustandsdichten. Die Zustände im Bereich von Valenzbandkante und Leitungsbandkante werden in den nachfolgenden Kapiteln identifiziert und analysiert.

⁹85. (110) Ebene der Superzelle, 4.378 a.u. über dem Oberflächenanion



Was den Abfall ins Vakuum betrifft, so unterscheiden sich besetzte und unbesetzte Zustände wesentlich. Die langreichweitigen Valenzbandzustände sind unterhalb der Valenzbandkante lokalisiert. Sie werden von zwei verschiedenartigen Zuständen dominiert (siehe Abb. 6.16: unten, links), dem "dangling bond", der sich als scharfer Oberflächenzustand darstellt, und einer breiten Oberflächenresonanz, dem "back bond". Durch diese werden auch die RTM-Bilder dominiert (siehe unten).

Im Gegensatz dazu existieren im Leitungsband viele verschiedenartige Zustände. Der " p_z "-Zustand (siehe Abb. 6.16: unten, rechts), der die RTM-Bilder bei kleinen Tunnelspannungen bestimmt, ist in der Zustandsdichte fast nicht erkennbar.



Die Abstandsabhängigkeit der O-DOS(x,E) ist in Abbildung 6.17 festgehalten. Drei verschiedene Abstände d vom Oberflächenanion wurden ausgewertet. Mit zunehmendem Abstand von der Oberfläche verschieben sich die relativen Gewichte von den besetzten zu den unbesetzten Zuständen. Die unbesetzten Zustände fallen sehr viel langsamer ins Vakuum hinein ab als die besetzten. Die Struktur der O-DOS der besetzten Zustände ändert sich mit zunehmendem Abstand kaum. Die besetzten "db"- und "bb"-Zustände sind auch bei einem Abstand von 6 a.u. vom Oberflächenanion sehr gut in der O-DOS identifizierbar, während der unbesetzte " p_z "-Zustand im Untergrund verschwindet.

Der Vergleich zwischen der Methode der speziellen \vec{k} -Punkte und der zweidimensionalen Tetraedermethode ist in Abbildung 6.17, rechts, unten festgehalten. Mit der

Methode der speziellen \vec{k} -Punkte werden sehr viel glattere Zustandsdichten erzielt. Auch dies vermittelt einen Eindruck von dem Fehler, der bei der \vec{k} -Punktintegration gemacht wird.

6.6.2.3 Die besetzten Zustände

Bei den besetzten Zuständen lassen sich, wenn man sich vom Vakuum aus der Oberfläche nähert, drei Klassen von Zuständen unterscheiden: Dies sind sogenannte "dangling bonds" (DB), "back bonds" (BB) und "bridge bonds" (BR). Typische Vertreter von DB, BB, und BR sind für zwei \vec{k} -Punkte¹⁰ in den Abbildungen Abb. 6.18 bis Abb. 6.23 dargestellt.

In der Oberflächenatomlage sind DB und BB jeweils als Doppelkeule am Oberflächenanion lokalisiert. Je eine Keule jeder Doppelkeule ist vakuumwärts gerichtet, die andere entgegengesetzt bulkwärts. Die vakuumwärts gerichtete Keule des DB ist in $[00\bar{1}]$ -Richtung gekippt, die des BB in $[001]$ -Richtung. Die bulkwärts gerichtete Keule des BB ist in Richtung des Kations der zweiten Atomlage gerichtet und stabilisiert diese Bindung. Die bulkwärts gerichtete Keule des DB liegt innerhalb des Tetraeders, welches von dem Oberflächenanion, zwei Oberflächenkationen sowie einem Kation der nächsten Atomlage aufgespannt wird (siehe auch Abbildung 6.1). Die BR sind längst der Bindungen zwischen Oberflächenanionen und -kationen lokalisiert. Die BR und die BB existieren nur als Bulkzustände und Oberflächenresonanzen. Die DB existieren auch als Oberflächenzustände.

Am X' -Punkt ($\vec{k} = \frac{1}{2}\vec{k}_z$) sind die beiden entarteten DB-Zustände ($n=67,68$) dargestellt (Abb. 6.18). Sie sind lokalisierte Oberflächenzustände. Ebenfalls dargestellt sind zwei BB-Zustände ($n=52,56$) (Abb. 6.19) und zwei BR-Zustände ($n=65,66$) (Abb. 6.20). Diese Zustände sind Oberflächenresonanzen. Die DB und die BR fallen sehr schnell in den Bulk hinein ab (Abb. 6.24, oben). Die Wellenfunktionen haben daher wenig Überlapp und die Eigenwerte sind daher fast entartet ($\Delta\epsilon = 0.7$ meV (0.3 meV)). Die beiden BB-Zustände fallen nicht so schnell in den Bulk hinein ab und sind auch nicht entartet ($\Delta\epsilon = 130$ meV). Die Energieaufspaltung der BB-Zustände ist möglicherweise ein Artefakt der endlichen Superzelle. Dies könnte nur durch Rechnungen mit einer größeren Superzelle entschieden werden.

Die Abbildungen 6.21 bis 6.23 zeigen Zustände am \vec{k} -Punkt $\vec{k} = \frac{3}{10}\vec{k}_y + \frac{4}{10}\vec{k}_z$. Es sind wieder zwei DB (Abb. 6.21) und zwei BB (Abb. 6.22), sowie zwei Mischzustände als typische Vertreter von Bulkzuständen (Abb. 6.23). Auch hier fallen die DB-

¹⁰ $\vec{k} = \frac{1}{2}\vec{k}_z$ (X' -Punkt am Rand der irreduziblen BZ) und $\vec{k} = \frac{3}{10}\vec{k}_y + \frac{4}{10}\vec{k}_z$ (willkürlich ausgewählter \vec{k} -Punkt im Inneren der irreduziblen BZ)

Zustände – weil lokalisierte Oberflächenzustände – schneller in den Bulk hinein ab als die BB-Zustände, die auch hier Oberflächenresonanzen sind (Abb. 6.24, unten). Die Energieaufspaltung der DB beträgt $\Delta\epsilon = 6.9$ meV, die der BB beträgt $\Delta\epsilon = 176$ meV.

Obwohl alle Zustände exponentiell ins Vakuum hinein abfallen (6.28), unterscheiden sich die verschiedenen Zustände in Amplitude und Abfallkonstante. Dies ist in der Abbildung 6.24 skizziert: Als Oberflächenzustände bzw. Oberflächenresonanzen haben die DB und die BB ein deutlich größeres Gewicht im Vakuum als die verschiedenen Bulkzustände. Dabei haben die Bulkzustände ein niedriges Gewicht im Vakuum, weil ihre Amplitude $C_{\vec{k}\nu}$ klein ist gegenüber den Amplituden von BB und DB. Obwohl auch die BR Oberflächenresonanzen sind, haben sie ein geringes Gewicht im Vakuum, da sie ein vergleichsweise größeres $\kappa_{\vec{k}\nu}$ haben, also schneller abfallen als die DB und BB. Im Bereich weit außerhalb des Kristalls, übersteigen die Gewichte der DB und der BB die aller anderen Zustände um ein bis zwei Größenordnungen.

Die lokale Zustandsdichte im Vakuum wird also von DB- und BB-Zuständen dominiert. Wertvolle Information, gerade im Hinblick auf die Interpretation von RTM-Bildern, steckt in der Energiedispersion dieser Zustände. Diese wird wie folgt analysiert:

Alle Wellenfunktionen $\Psi_{\vec{k}\nu}$ werden auf der 90'sten (110) Ebene der Ortsraumstützpunkte betrachtet. Diese Ebene liegt 6.531 a.u. über dem Oberflächenanion. Das totale Gewicht $\Phi_{\vec{k}\nu}^2(x)$ (6.27) der Wellenfunktionen in dieser Ebene wird ausgerechnet, d.h. das Betragsquadrat der Wellenfunktionen wird über alle Stützpunkte dieser Ebene summiert. Alle Wellenfunktionen, deren Gewicht¹¹ 0.5×10^{-5} übersteigt, werden ausgewählt und weiter analysiert. Aus insgesamt 2448 besetzten Zuständen werden so die 202 langreichweitigsten ausgewählt. Es handelt sich hierbei ausschließlich um DB- und BB-Zustände. Die [001]-Position des Maximums der jeweiligen Ladungsverteilung wird ausgewertet, um zwischen DB und BB zu unterscheiden. Liegt das Maximum relativ zur Position des Oberflächenanions in [001]-Richtung verschoben, so handelt es sich um einen BB, im andren Fall um einen DB.

Nachdem die Zustände so getrennt sind, werden ihre Oberflächenzustandsdichten ausgewertet.

$$\rho_{90}^{DB(BB)}(E) = \sum_{\{\vec{k}\nu \in DB(BB)\}} \omega_{\vec{k}} \Phi_{\vec{k}\nu}^2(x_{90}) \frac{1}{\sqrt{\pi\delta}} e^{-\left(\frac{E - \epsilon_{\vec{k}\nu}}{\delta}\right)^2} \quad (6.31)$$

¹¹Die Wellenfunktionen sind auf dem diskreten Ortsraumgitter auf 1 normiert: $\sum_{i,j,k} |\Psi_{\vec{k}\nu}(x_i, y_j, z_k)|^2 = 1$

Die gewählte Energieverbreiterung beträgt wie im vorhergehenden Kapitel $\delta = 0.1$ eV.

Die Zustandsdichten sind in Abb. 6.25 festgehalten:

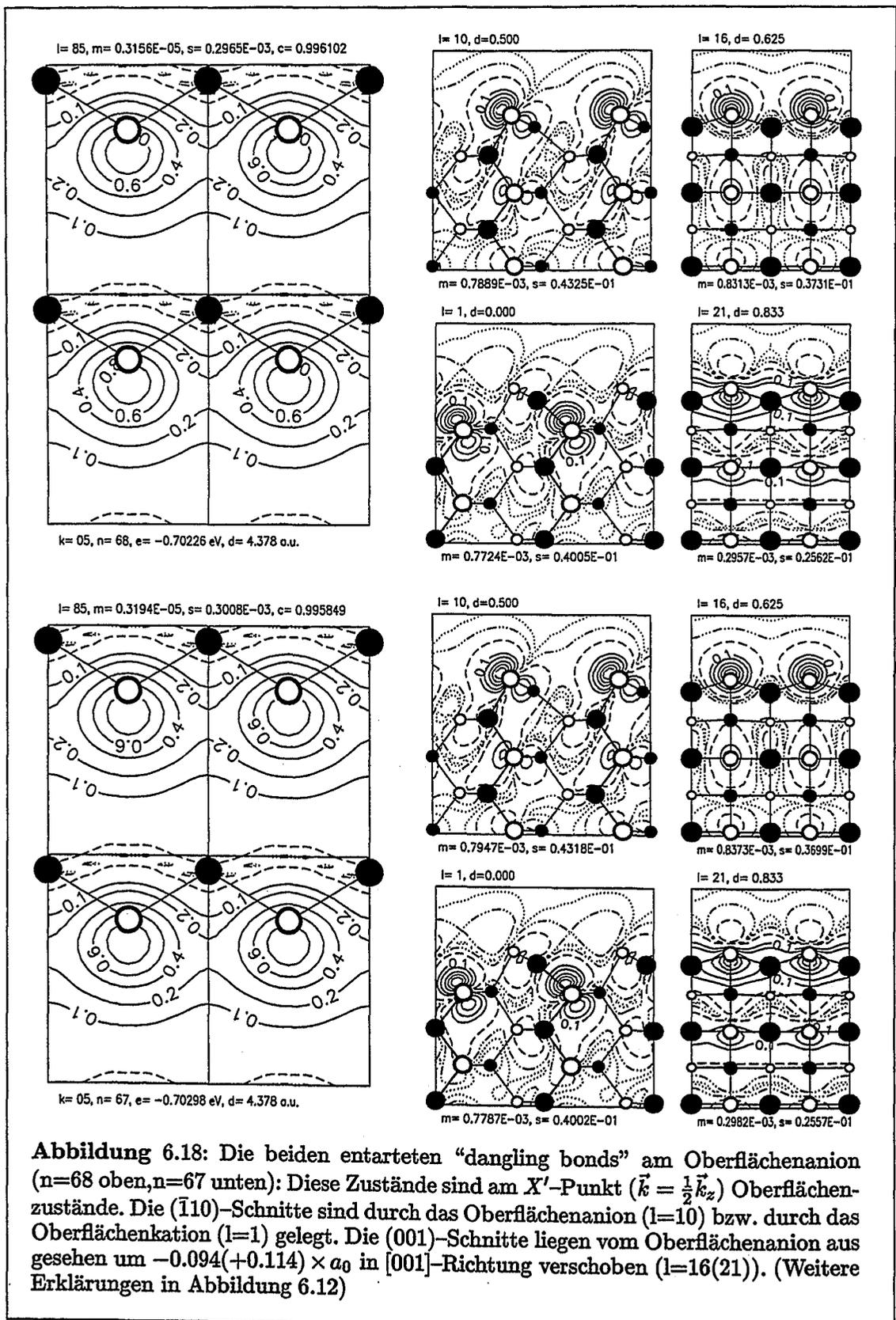
Das obere Bild zeigt die Anteile von DB und BB zur gesamten Zustandsdichte. Verglichen mit der breiten Resonanz der BB-Zustände zeigen sich die DB-Zustände als scharfes Oberflächenband. Das DB-Band hat sein Maximum bei -0.66 eV, gemessen von der Valenzbandkante. Seine Energieverteilung¹² reicht von -0.91 eV bis 0.39 eV. Die BB-Zustände erscheinen als zwei Oberflächenresonanzbänder, deren Energieverteilungen überlappen. Die Maxima liegen bei -1.26 und -1.01 eV. Die Energieverteilung erstreckt sich von -1.58 eV bis -0.91 eV. Bei einem "least-square"-Fit der Oberflächenzustandsdichte der besetzten Zustände mit 8 Gaußfunktionen ergeben sich für die Peaks folgende Parameter:

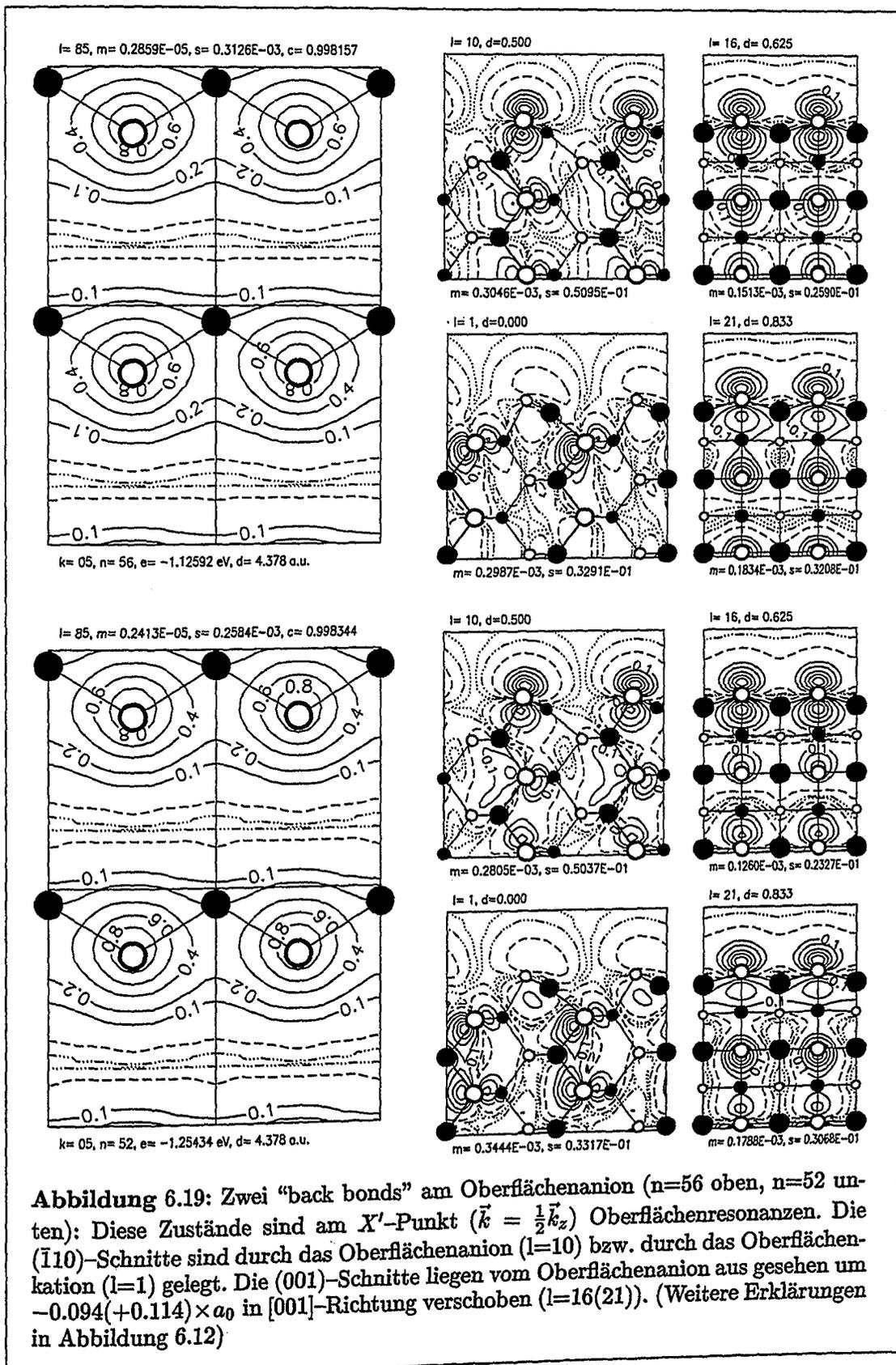
Zustand	E [eV]	δ [eV]	C
DB	- 0.66	0.19	$0.84 \cdot 10^{-2}$
BB(1)	- 1.01	0.13	$0.24 \cdot 10^{-2}$
BB(2)	- 1.27	0.22	$0.35 \cdot 10^{-2}$

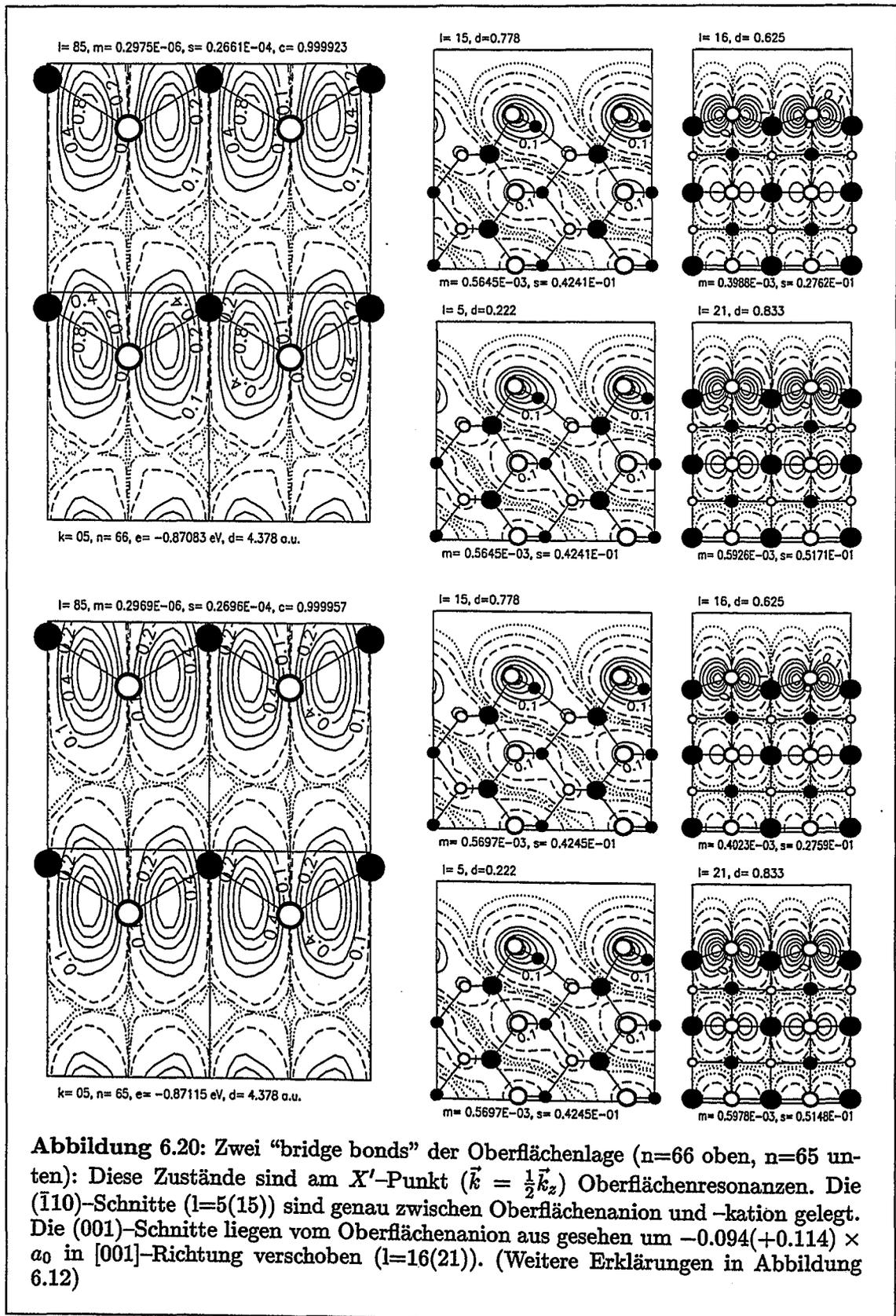
Tabelle 6.16: Gaußfit der Oberflächenzustandsdichte: Parameter Energie E, Breite δ und Amplitude C der besetzten "DB" und "BB".

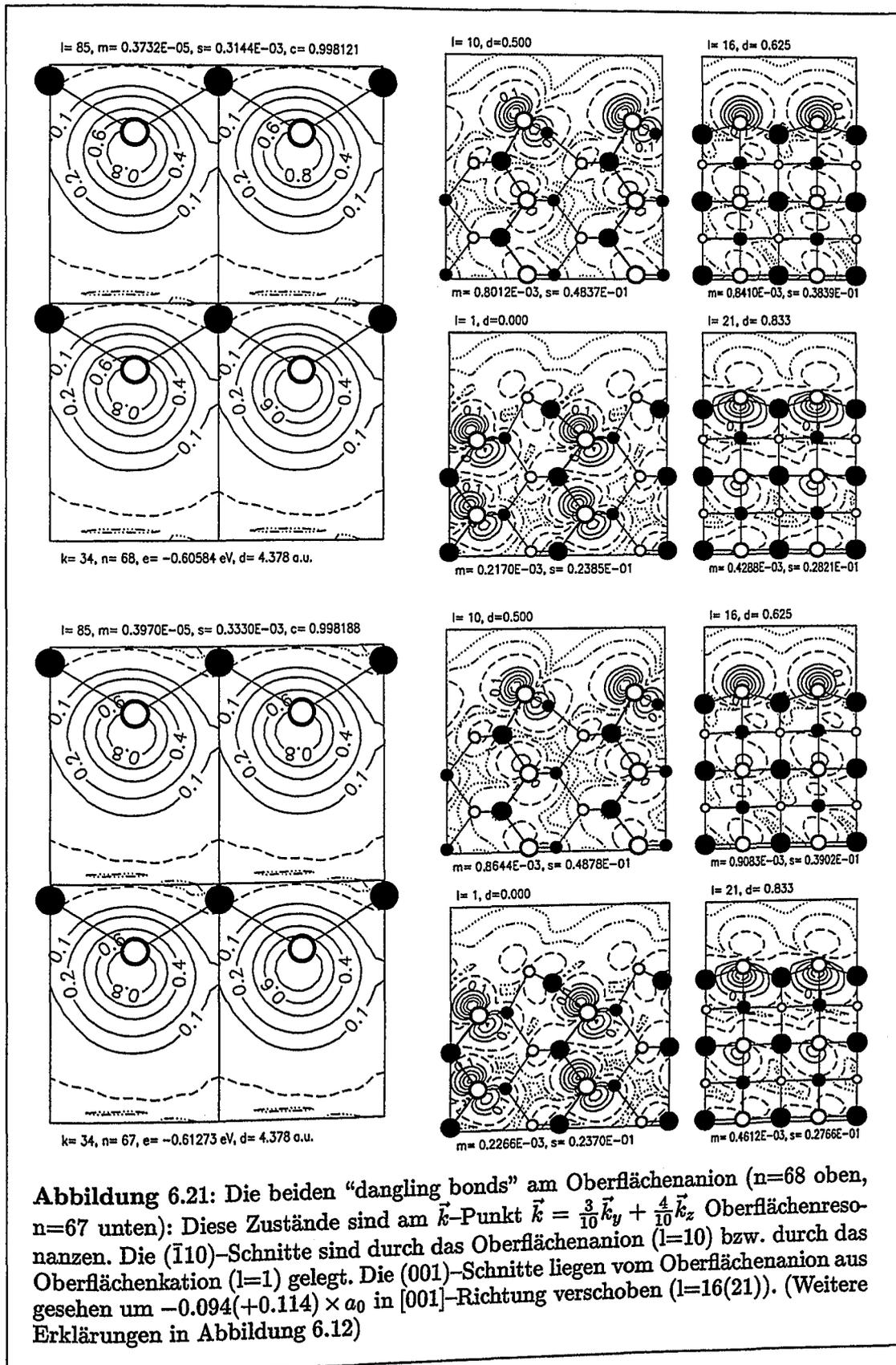
Im unteren Bild sind die einzelnen Beiträge nach der [001]-Position des Ladungsdichtemaximums sortiert. Die verschiedenen Zustandsdichten fassen jeweils all die Zustände zusammen, die ihr Ladungsdichtemaximum an der gleichen diskreten [001]-Position des Ortsgitters haben. Die [001]-Position ist relativ zum Oberflächenanion in Einheiten der Gitterkonstanten $a_0 = 5.85494 \text{ \AA}$ angegeben. Je weiter das Ladungsdichtemaximum eines Zustands in Richtung [001] (von der "DB"-Position zur "BB"-Position) verschoben ist, umso energetisch tiefer liegen die Zustände. Qualitativ ergibt sich folgendes Bild: Der Anschluß zwischen Vakuum und Bulk wird durch eine Oberflächenatomlage vermittelt. Die elektronische Zustandsdichte der (110) Oberfläche ist geprägt durch die tetraedrische Umgebung des Oberflächenanions und die planare Umgebung des Oberflächenkations. Nähert man sich störungstheoretisch, so sind die sp^3 -Zustände am Anion und die sp^2 -Zustände am Kation eine gute Ausgangsbasis, um Wellenfunktionen im Bereich der Oberfläche zu entwickeln. So entstehen DB-, BB- und BR-Zustände.

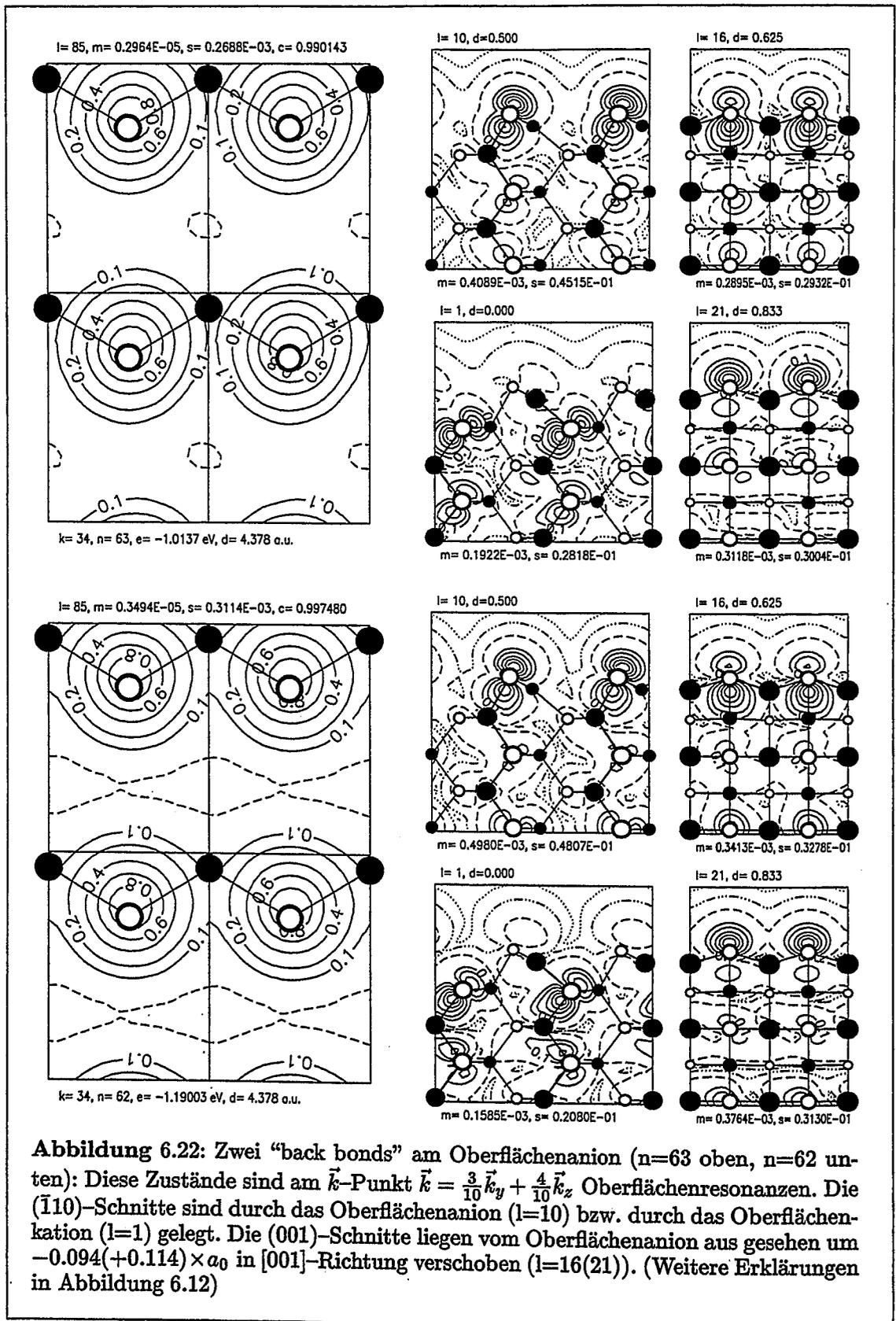
¹²Die Breite der Verteilung wird durch den Abfall der Zustandsdichte auf $\frac{1}{e}$ ihres Maximalwerts festgelegt.

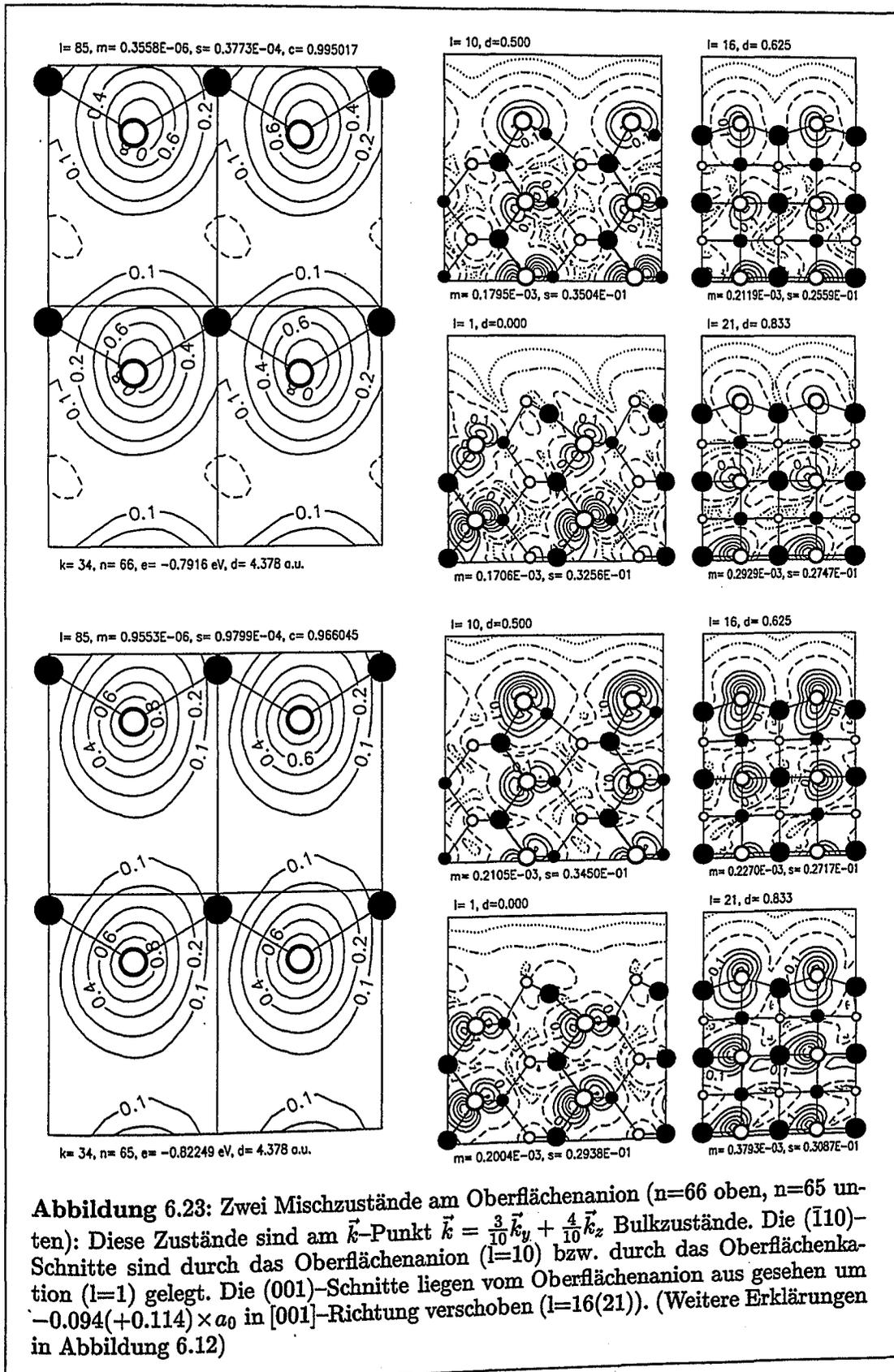












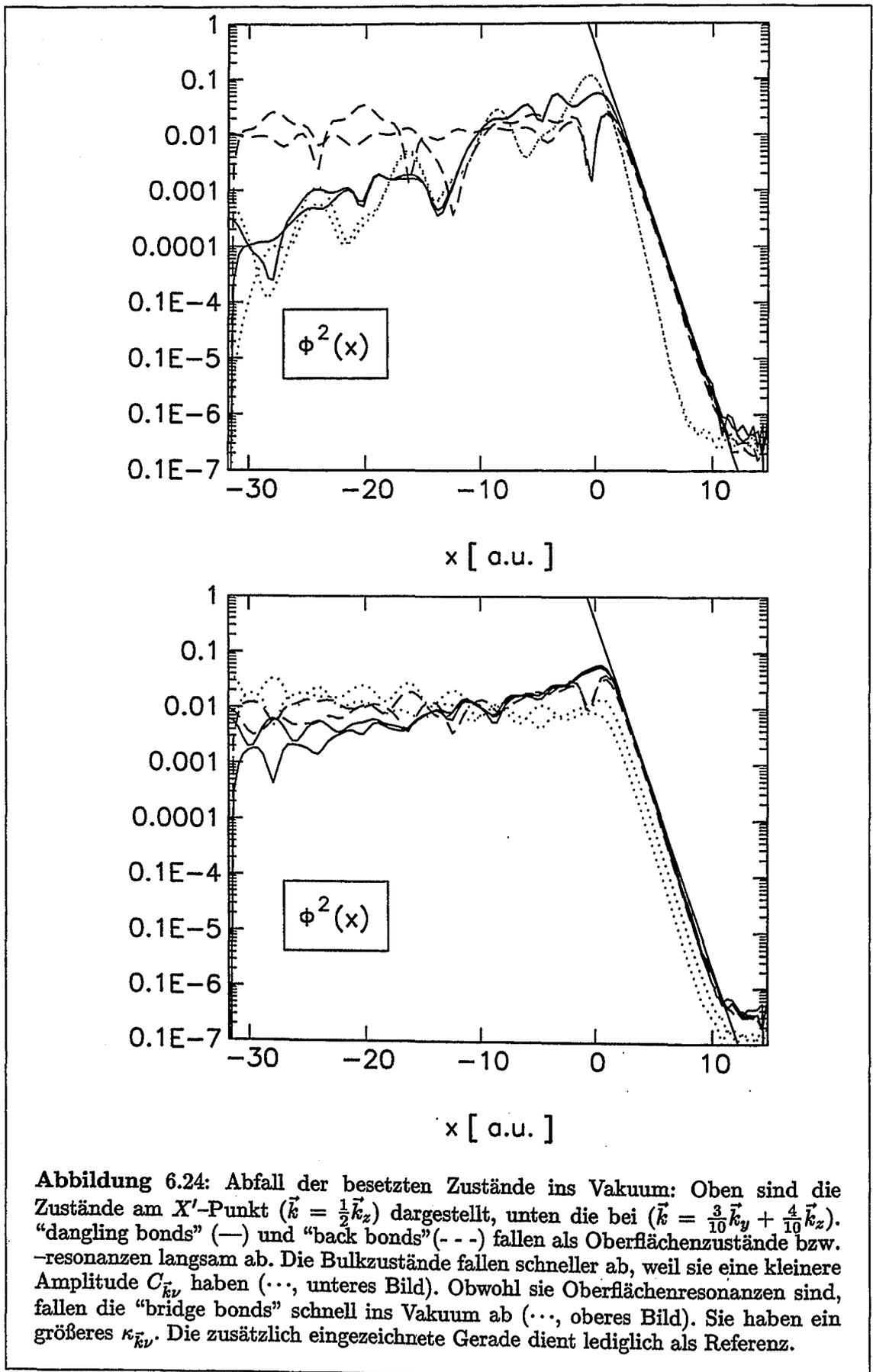


Abbildung 6.24: Abfall der besetzten Zustände ins Vakuum: Oben sind die Zustände am X' -Punkt ($\vec{k} = \frac{1}{2}\vec{k}_z$) dargestellt, unten die bei ($\vec{k} = \frac{3}{10}\vec{k}_y + \frac{4}{10}\vec{k}_z$). "dangling bonds" (—) und "back bonds" (---) fallen als Oberflächenzustände bzw. -resonanzen langsam ab. Die Bulkzustände fallen schneller ab, weil sie eine kleinere Amplitude $C_{\vec{k}\nu}$ haben (\dots , unteres Bild). Obwohl sie Oberflächenresonanzen sind, fallen die "bridge bonds" schnell ins Vakuum ab (\dots , oberes Bild). Sie haben ein größeres $\kappa_{\vec{k}\nu}$. Die zusätzlich eingezeichnete Gerade dient lediglich als Referenz.

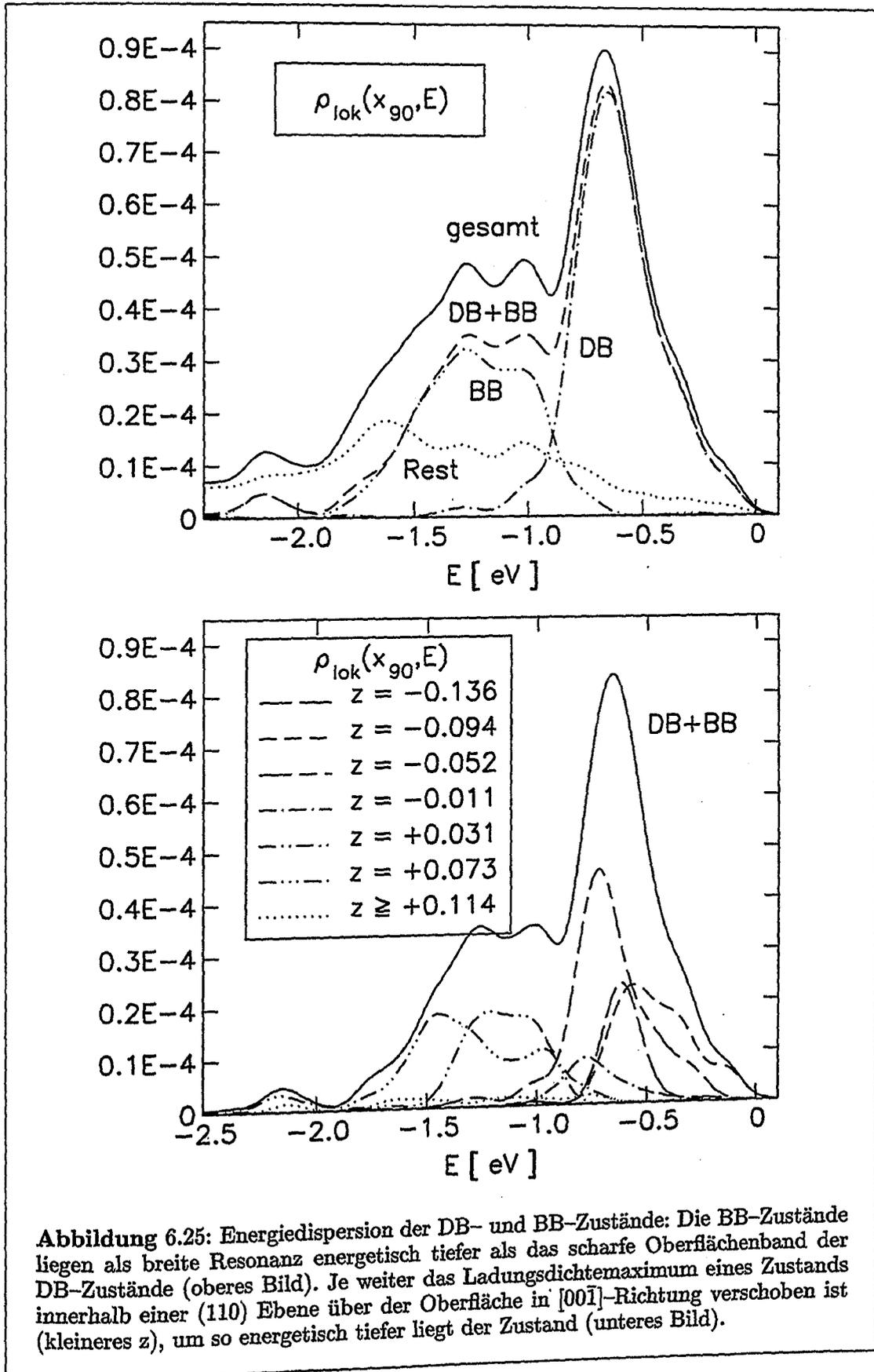


Abbildung 6.25: Energiedispersion der DB- und BB-Zustände: Die BB-Zustände liegen als breite Resonanz energetisch tiefer als das scharfe Oberflächenband der DB-Zustände (oberes Bild). Je weiter das Ladungsdichtemaximum eines Zustands innerhalb einer (110) Ebene über der Oberfläche in $[00\bar{1}]$ -Richtung verschoben ist (kleineres z), um so energetisch tiefer liegt der Zustand (unteres Bild).

6.6.2.4 Die unbesetzten Zustände

Die unbesetzten Zustände sind, was ihren Abfall ins Vakuum angeht, nicht so einfach in verschiedene Klassen einzuteilen, wie es bei den besetzten Zuständen möglich war. Es existieren viele verschiedenartige Zustände, die allesamt ähnliches Gewicht im Vakuum haben. Darunter sind Zustände mit Ladung über dem Oberflächenkation, Zustände mit Ladung über dem Oberflächenanion und Zustände mit Ladungsdichte über beiden Oberflächenatomen. In den Abbildungen 6.26 bis 6.29 sind einige Zustände stellvertretend für viele andere aufgeführt.

Abbildung 6.26 zeigt den " p_z "-Zustand am Kation. Dieser Zustand existiert als scharfer Oberflächenzustand zwischen X- und M-Punkt. In einer (110) Ebene im Vakuum erscheint seine Ladungsdichtekontur als Reihe in [001]-Richtung. Die Energiedispersion dieses Oberflächenbandes liegt zwischen 1.70 eV und 2.00 eV über dem Valenzbandmaximum. An diesem Oberflächenband lassen sich auch sehr schön Hybridisierungseffekte studieren. So ist die " p_z "-Doppelkeule am X-Punkt (Abbildung 6.26, oberes Bild) komplett ausgebildet, während am M-Punkt (Abbildung 6.26, unteres Bild) lediglich die Vakuumkeule ausgeprägt ist.

Abbildung 6.27 zeigt zwei Hybrid-Zustände, welche ebenso wie der " p_z "-Zustand viel Ladungsdichte über dem Oberflächenkation besitzen. Sie liegen energetisch höher 2.34 eV bzw. 2.93 eV über dem Valenzbandmaximum und stehen exemplarisch für eine ganze Reihe von ähnlichen Zuständen. Im Gegensatz zu den " p_z "-Zuständen erscheint ihre Ladungsdichtekontur in einer (110) Ebene im Vakuum als Reihe in $[\bar{1}10]$ -Richtung und nicht als Reihe in [001]-Richtung.

Abbildung 6.28 zeigt zwei Zustände, die im Vakuum über dem Anion viel Ladungsdichte haben. Auch diese stehen exemplarisch für eine ganze Reihe von ähnlichen Zuständen. Diese Zustände zeigen schwach ausgeprägte "db"-artige Keulen am Oberflächenanion.

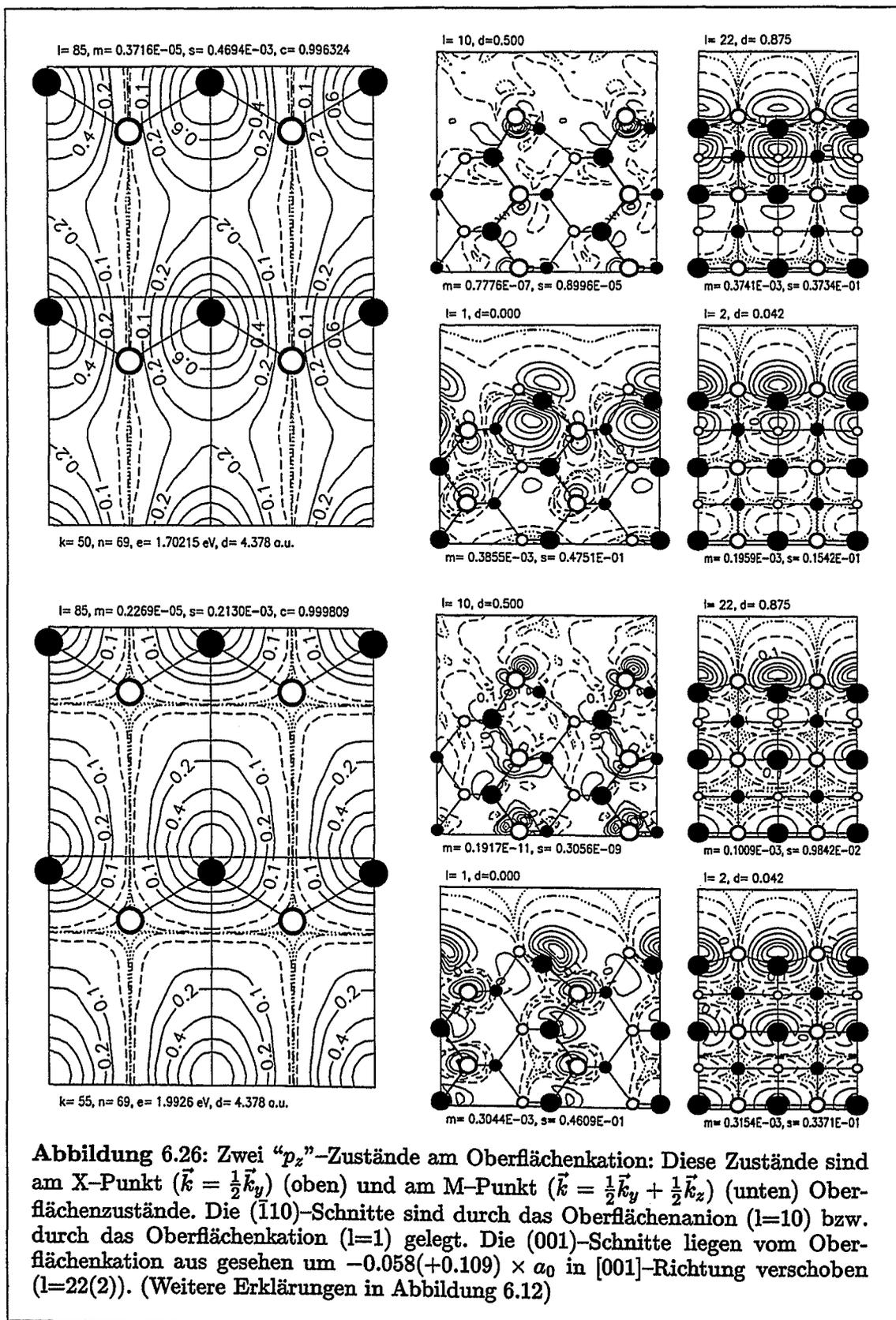
Die Abbildung 6.29 schließlich zeigt zwei Zustände mit Ladungsverteilungen im Vakuum, die mehr als ein Maximum in einer (110) Ebene besitzen. Auch diese Zustände stehen exemplarisch für eine ganze Reihe von ähnlichen Zuständen.

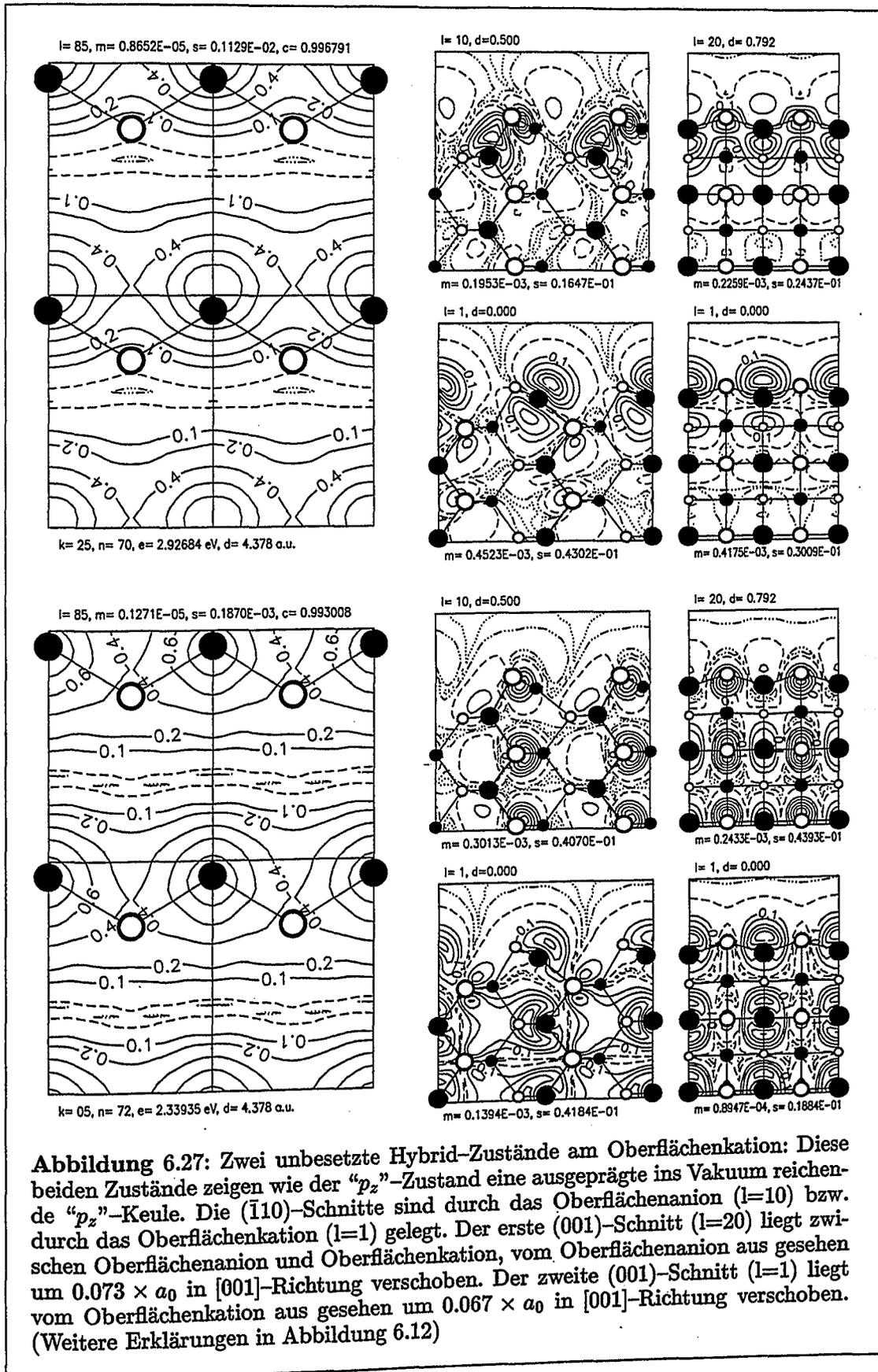
Die Abbildung 6.30 zeigt den Abfall der ausgewählten Zustände ins Vakuum. Die eingezeichnete Referenzgerade ist identisch mit der in 6.24 eingezeichneten Referenzgeraden, die in etwa den Abfall der besetzten "DB"- und "BB"-Zustände wiedergibt. Die obere Abbildung zeigt den Abfall der Zustände mit Ladungsdichteanhäufung über dem Oberflächenkation, i.e. die beiden " p_z "-Zustände aus Abbildung 6.26 ((—) und (- - -)) sowie die beiden Hybrid-Zustände aus Abbildung 6.27 ((- · -) und (· · ·)). Jeweils der obere der beiden Zustandstypen ((—) und (- · -)) ist ein stark

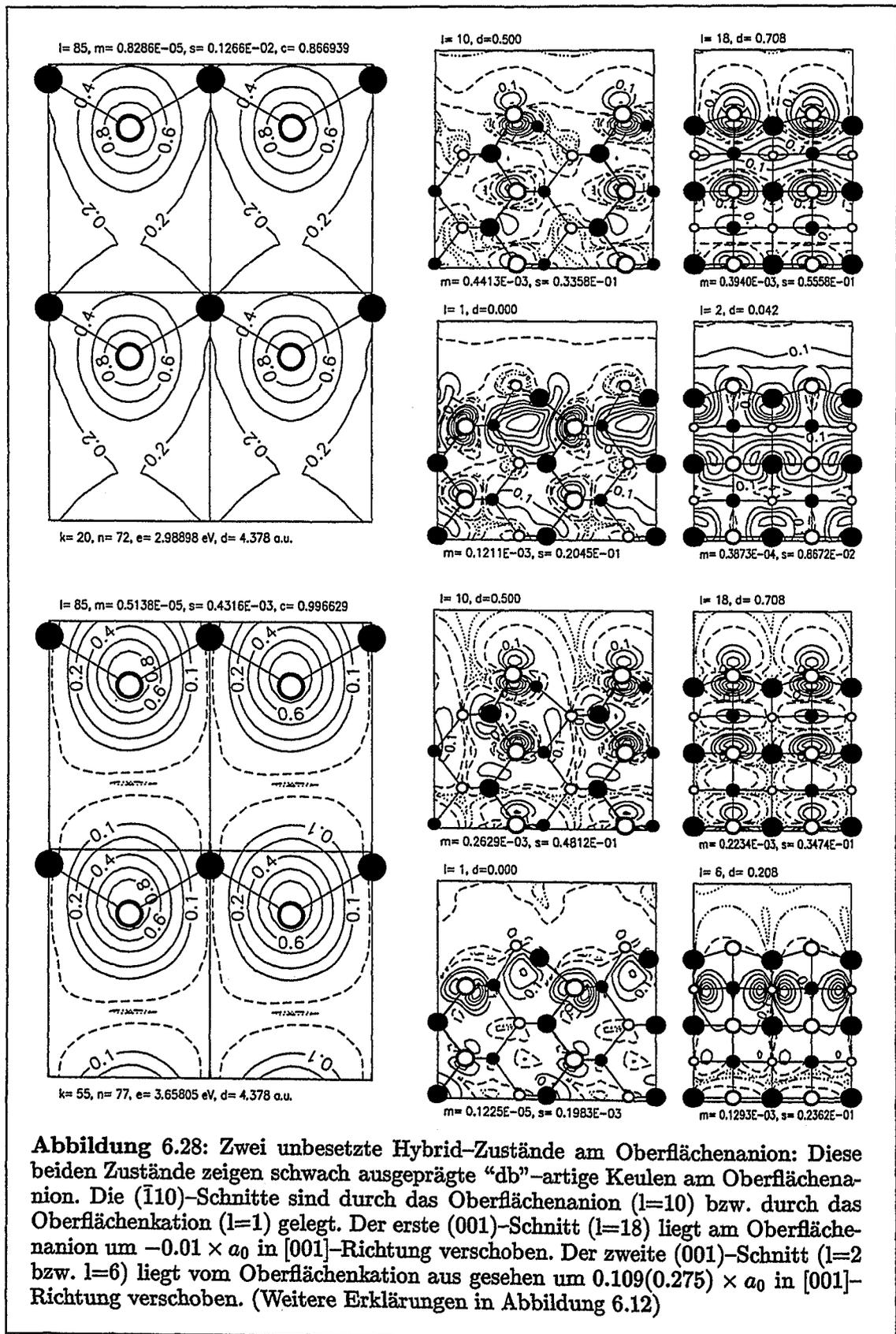
lokalisierter Oberflächenzustand.

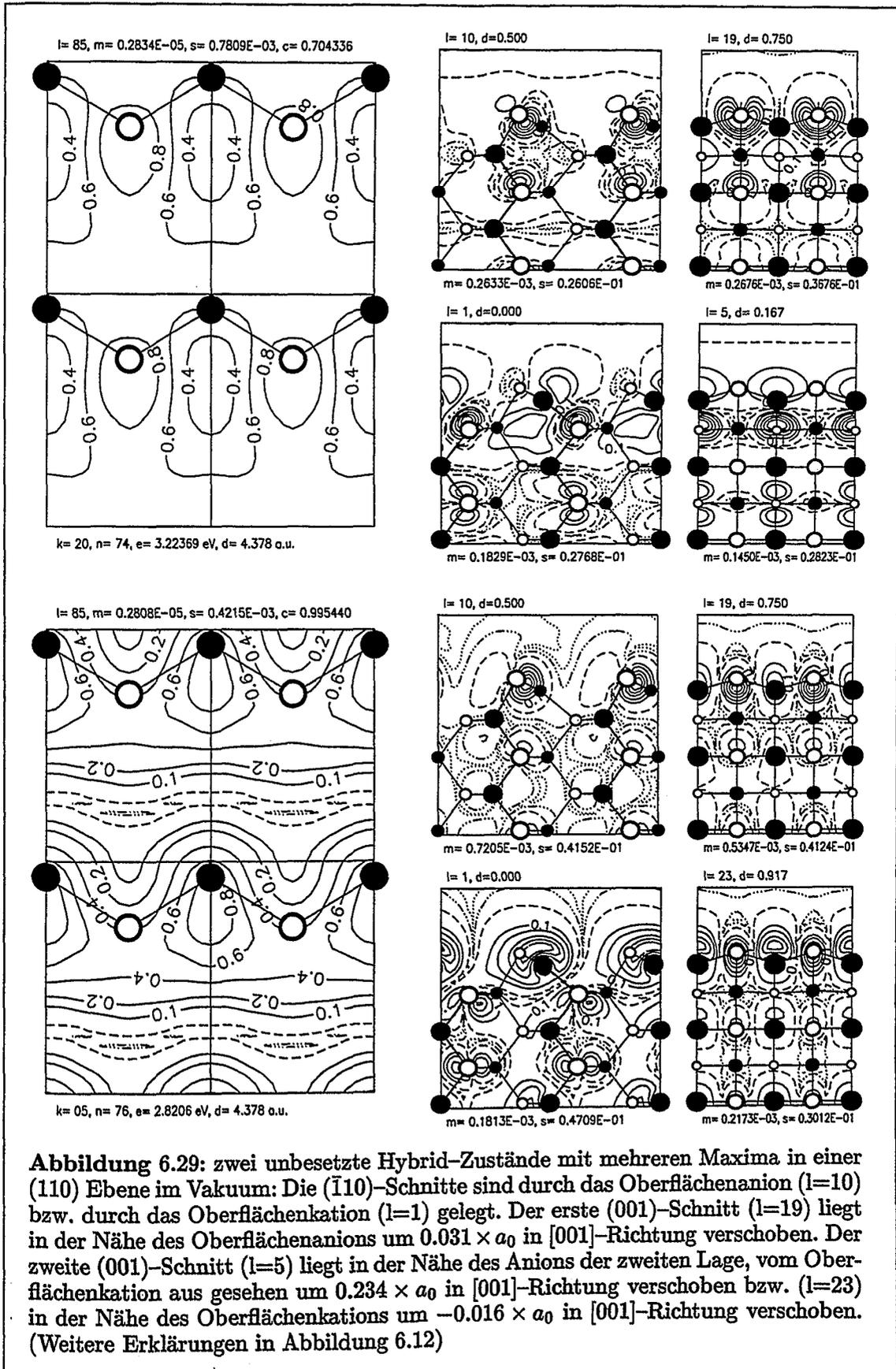
Die untere Abbildung zeigt den Abfall der Zustände mit Ladungsdichteanhäufung über dem Oberflächenanion aus Abbildung 6.28 ((—) und (- - -)) sowie den Abfall der Zustände aus Abbildung 6.29 ((- · -) und (· · ·)). Obwohl diese Zustände keine gut lokalisierten Oberflächenzustände sind, so fallen sie doch qualitativ ähnlich ab, wie die Oberflächenzustände mit Ladungsdichteanhäufung über dem Oberflächenkation. Ähnlich wie bei den besetzten Zuständen wird auch hier die Energiedispersion der einzelnen Zustandsarten untersucht. Die einzelnen Zustände werden nach verschiedenen Kriterien sortiert und deren O-DOS (vergleiche Gleichung 6.31) über der Oberfläche getrennt ausgewertet. Dies ist in Abbildung 6.31 festgehalten. Die obere Abbildung unterscheidet zwischen Zuständen mit einem Ladungsdichtemaximum in der (110) Ebene und Zuständen mit mehreren Ladungsdichtemaxima. Die Zustände mit einem Ladungsdichtemaximum sind zusätzlich danach sortiert, ob das Maximum über dem Oberflächenanion, -kation oder anderswo liegt. In der unteren Abbildung werden die Zustände unabhängig von der Zahl der Maxima danach unterschieden, ob das jeweils stärkste Maximum über dem Oberflächenanion, -kation oder anderswo liegt.

Insgesamt ergibt sich folgendes Bild: Die Zustände mit Ladungsdichteanhäufung über dem Oberflächenkation (Maximum bei 2.6 eV über der Valenzbandkante) liegen energetisch näher an der Leitungsbandkante als diejenigen mit Ladungsdichteanhäufung über dem Oberflächenanion (Maximum bei 3.1 eV über der Valenzbandkante). Beide Zustandsarten überlappen. Im Zwischenbereich existieren sogar viele Zustände mit Ladungsdichteanhäufungen sowohl über dem Kation, als auch über dem Anion.









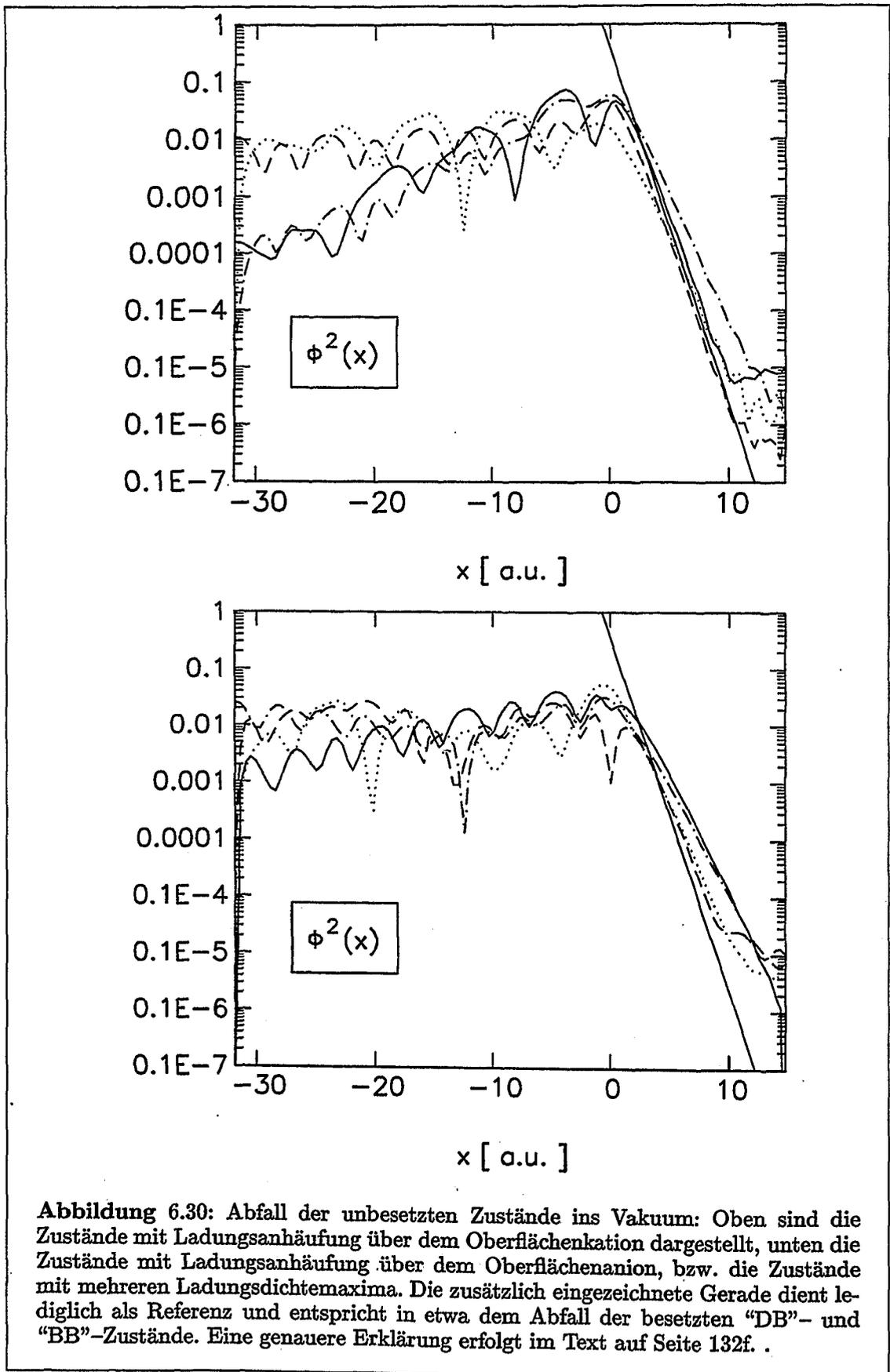


Abbildung 6.30: Abfall der unbesetzten Zustände ins Vakuum: Oben sind die Zustände mit Ladungsanhäufung über dem Oberflächenkation dargestellt, unten die Zustände mit Ladungsanhäufung über dem Oberflächenanion, bzw. die Zustände mit mehreren Ladungsdichtemaxima. Die zusätzlich eingezeichnete Gerade dient lediglich als Referenz und entspricht in etwa dem Abfall der besetzten "DB"- und "BB"-Zustände. Eine genauere Erklärung erfolgt im Text auf Seite 132f. .

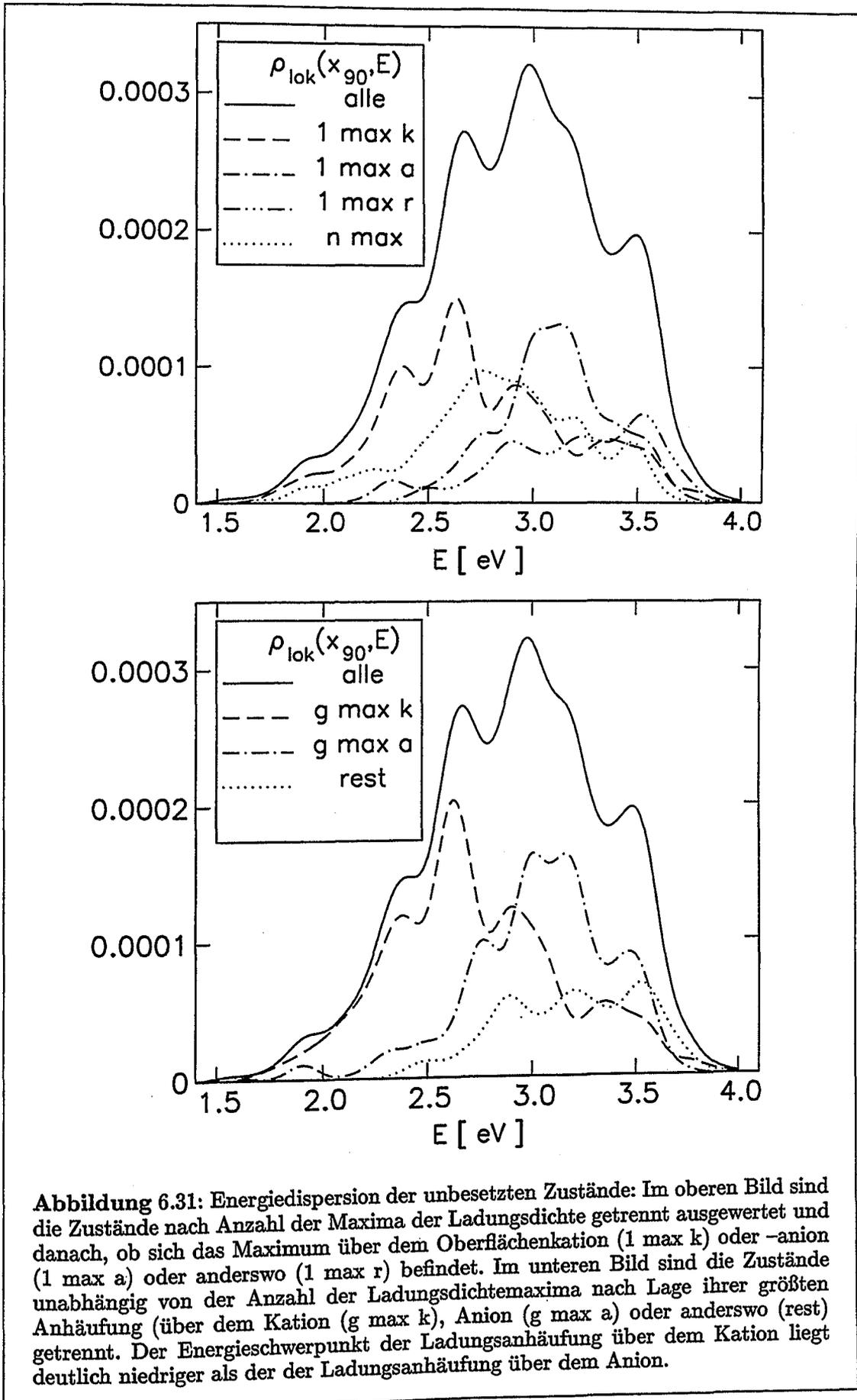


Abbildung 6.31: Energiedispersion der unbesetzten Zustände: Im oberen Bild sind die Zustände nach Anzahl der Maxima der Ladungsdichte getrennt ausgewertet und danach, ob sich das Maximum über dem Oberflächenkation (1 max k) oder -anion (1 max a) oder anderswo (1 max r) befindet. Im unteren Bild sind die Zustände unabhängig von der Anzahl der Ladungsdichtemaxima nach Lage ihrer größten Anhäufung (über dem Kation (g max k), Anion (g max a) oder anderswo (rest)) getrennt. Der Energieschwerpunkt der Ladungsanhäufung über dem Kation liegt deutlich niedriger als der der Ladungsanhäufung über dem Anion.

Kapitel 7

Vergleich mit dem Tunnelexperiment

Mit der berechneten elektronischen Struktur der (110) Oberfläche ist es möglich, die Interpretation der RTM-Experimente von Philipp Ebert [Ebe93] theoretisch zu testen. Er untersucht in seiner Doktorarbeit unter anderem die Spannungsabhängigkeit morphologischer Details von STM-Bildern am Beispiel von InP und GaP. Diese Details werden erklärt durch den Beitrag verschiedener Oberflächenzustände.

Ziel ist es, theoretisch zu klären, inwieweit diese spannungsabhängigen Details tatsächlich durch die elektronische Struktur der Oberfläche bestimmt sind, bzw., die verschiedenen Beiträge zum STM-Bild genau zu identifizieren.

Nach einer kurzen Erklärung der theoretischen Grundlagen auf der Basis der Bardeenschen Transferhamiltonian-Methode [TKIK91, Che93] unter Vernachlässigung von Spitzeneffekten [Bar61, TH85], (Tersoff-Hamann-Näherung: s-Wellenfunktionen für die Spitze, strukturlose Zustandsdichte der Spitze) werden im folgenden zunächst die experimentellen Ergebnisse dargestellt. Diese werden mit der berechneten lokalen Zustandsdichte verglichen.

7.1 Theorie des Tunnelstroms

Bei der Erklärung der theoretischen Grundlagen des Tunnelstroms folgen wir [Che93]: Das Potential $U(\vec{r})$ des aus Probe und Tunnelspitze zusammengesetzten Systems kann künstlich in zwei räumlich getrennte Beiträge

- $U_S(\vec{r})$, das Potential der Oberfläche (“surface”: S), und
- $U_T(\vec{r})$, das Potential der Tunnelspitze (“tip”: T),

aufgespalten werden:

$$U(\vec{r}) = U_S(\vec{r}) + U_T(\vec{r}), \quad U_S(\vec{r}) U_T(\vec{r}) = 0 \quad (7.1)$$

Die beiden Raumgebiete, \mathcal{G}_T im Bereich der Spitze und \mathcal{G}_S im Bereich der Oberfläche, sind durch eine Grenzfläche Σ getrennt. U_S und U_T werden so gewählt, daß sie jeweils im gegenüberliegenden Raumgebiet identisch verschwinden:

$$U_S(\vec{r}) = 0 \quad \text{in } \mathcal{G}_T, \quad U_T(\vec{r}) = 0 \quad \text{in } \mathcal{G}_S \quad (7.2)$$

Das Vakuumpotential dient als Referenzpotential ($U_{\text{Vak}} = 0$). Mit diesen beiden Potentialen werden im Folgenden nun Oberfläche und Spitze als getrennte Systeme behandelt:

$$(T + U_S) \Psi_\mu^S = E_\mu^S \Psi_\mu^S \quad (7.3)$$

$$(T + U_T) \Psi_\nu^T = E_\nu^T \Psi_\nu^T \quad (7.4)$$

Das zusammengesetzte System wird aus den Eigenvektoren der getrennten Systeme¹ durch zeitabhängige Störungstheorie berechnet. Das Potential der Spitze U_T sei für Zeiten $t < 0$ ausgeschaltet und wird für Zeiten $t > 0$ adiabatisch eingeschaltet. Für $t < 0$ befinde sich das System im Eigenzustand $\Psi_\mu^S(\vec{r})$ der Oberfläche. Die Änderung der zeitabhängigen Wellenfunktion $\Psi(\vec{r}, t)$ durch diese Störung kann nach Eigenfunktionen der Spitze $\Psi_\nu^T(\vec{r})$ entwickelt werden.

$$\Psi(\vec{r}, t) = \Psi_\mu^S(\vec{r}) e^{-\frac{i}{\hbar} E_\mu^S t} + \sum_\nu c_\nu(t) \Psi_\nu^T(\vec{r}) e^{-\frac{i}{\hbar} E_\nu^T t} \quad (7.5)$$

Für die Entwicklungskoeffizienten $c_\nu(t)$ ergibt sich durch Einsetzen in die Schrödingergleichung folgende Differentialgleichung:

$$i \hbar \frac{d}{dt} c_\nu(t) = \langle \Psi_\nu^T | U_T | \Psi_\mu^S \rangle e^{-\frac{i}{\hbar} (E_\mu^S - E_\nu^T) t} + \sum_\lambda c_\lambda(t) \langle \Psi_\nu^T | U_S | \Psi_\lambda^T \rangle e^{-\frac{i}{\hbar} (E_\lambda^T - E_\nu^T) t} \quad (7.6)$$

In 1. Ordnung Störungstheorie und mit der Darstellung der δ -Funktion mittels Gleichung (A.68) ergibt sich:

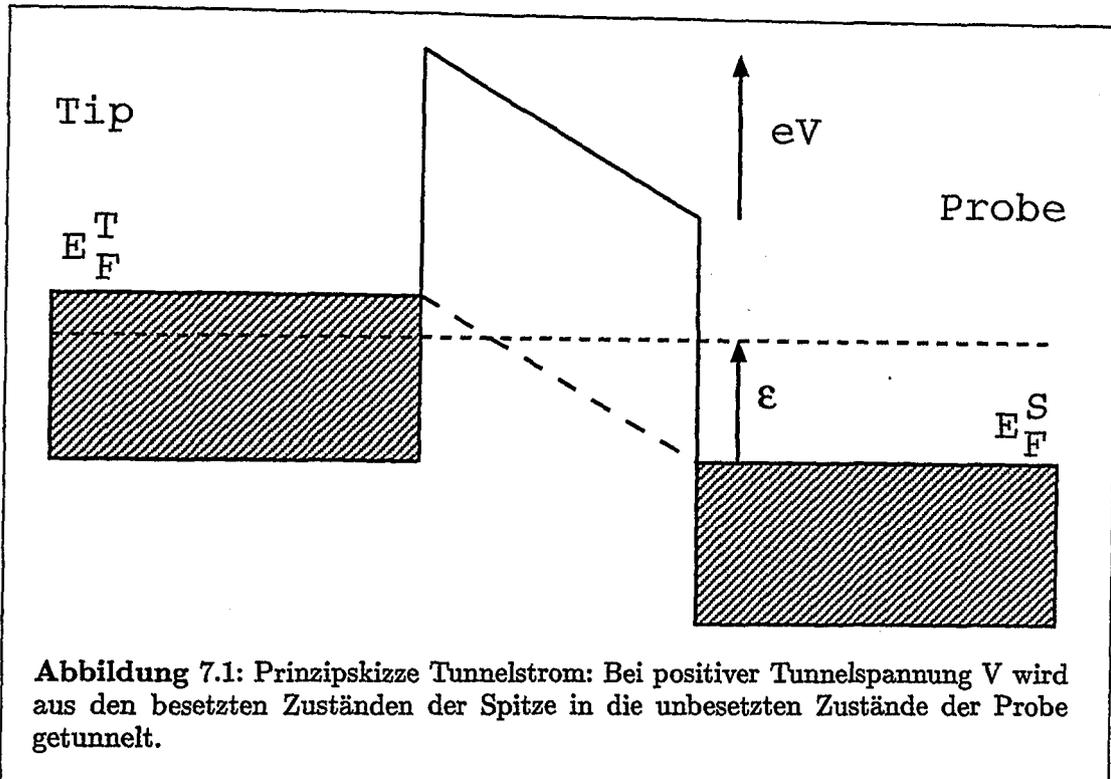
$$c_\nu(t) = \frac{1}{i} \int_0^{\frac{t}{\hbar}} dt' e^{-i(E_\mu^S - E_\nu^T)t'} \langle \Psi_\nu^T | U_T | \Psi_\mu^S \rangle \quad (7.7)$$

$$|c_\nu(t)|^2 = 2\pi \frac{t}{\hbar} \delta(E_\nu^T - E_\mu^S) \left| \langle \Psi_\nu^T | U_T | \Psi_\mu^S \rangle \right|^2 \quad (7.8)$$

$$w_{\mu\nu} = \frac{|c_\nu(t)|^2}{t} = \frac{2\pi}{\hbar} \delta(E_\nu^T - E_\mu^S) |M_{\mu\nu}|^2 \quad (7.9)$$

$$M_{\mu\nu} = \langle \Psi_\nu^T | U_T | \Psi_\mu^S \rangle \quad (7.10)$$

¹Die Potentiale U_S und U_T sind nicht die ungestörten Potentiale U_S^0 und U_T^0 der ursprünglichen Bardeen'schen Theorie [Bar61].



$|c_\nu(t)|^2$ ist die Wahrscheinlichkeit, ein Elektron, welches sich zur Zeit $t = 0$ im Zustand Ψ_μ^S befunden hat, zur Zeit t im Zustand Ψ_ν^T zu finden. $w_{\mu\nu}$ ist entsprechend die Übergangsrate.

U_T ist nur im Gebiet um die Spitze \mathcal{G}_T von 0 verschieden. Durch Einsetzen der Schrödingergleichung (7.4) in Gleichung (7.10) kann das Potential U_T entfernt werden. Das Integral (7.10) über \mathcal{G}_T kann dann mit dem Greenschen Integraltheorem (A.69) in ein Integral über die Grenzfläche Σ zwischen Oberfläche und Spitze umgeformt werden:

$$M_{\mu\nu} = -\frac{\hbar}{2m} \int_{\Sigma} d\vec{S} \cdot (\Psi_\nu^{T*} \nabla \Psi_\mu^S - \Psi_\mu^S \nabla \Psi_\nu^{T*}) \quad (7.11)$$

Ein Elektron kann natürlich nur dann von Ψ_μ^S nach Ψ_ν^T wechseln, wenn Ψ_μ^S besetzt ist und Ψ_ν^T unbesetzt. Summation über alle Übergangsraten unter Berücksichtigung der Besetzung bzw. Nichtbesetzung der Zustände Ψ_μ^S und Ψ_ν^T liefert einen Ausdruck, der proportional zum Strom ist, der von der Probe zur Spitze fließt:

$$I = \frac{4\pi e}{\hbar} \sum_{\nu\mu} [f(E_\mu^S - E_F^S) - f(E_\nu^T - E_F^T)] |M_{\mu\nu}|^2 \delta(E_\nu^T - E_\mu^S) \quad (7.12)$$

Bei Anlegen einer Spannung V (siehe Abbildung 7.1) ergibt sich entsprechend:

$$I = \frac{4\pi e}{\hbar} \sum_{\nu\mu} [f(E_\mu^S - E_F^S) - f(E_\nu^T - E_F^T)] |M_{\mu\nu}|^2 \delta(E_\nu^T - E_\mu^S - eV) \quad (7.13)$$

Mit Ersetzung der Summationen durch Integrationen

$$\sum_{\mu} \rightarrow \int \rho^S(E_{\mu}^S) dE_{\mu}^S, \quad \sum_{\nu} \rightarrow \int \rho^T(E_{\nu}^T) dE_{\nu}^T$$

kann wegen der δ -Funktion eine Integration ausgeführt werden:

$$I = \frac{4\pi e}{\hbar} \int d\epsilon \left[f(E_F^T - eV + \epsilon) - f(E_F^S + \epsilon) \right] \times \rho^T(E_F^T - eV + \epsilon) \rho^S(E_F^S + \epsilon) \left| M(E_F^S + \epsilon, E_F^T - eV + \epsilon) \right|^2 \quad (7.14)$$

Dabei sind ρ^S und ρ^T die Zustandsdichten (DOS) von Oberfläche und Spitze, $f(E)$ ist die Fermiverteilung. Im Limes $k_B T \rightarrow 0$ ergibt sich:

$$I = \frac{4\pi e}{\hbar} \int_0^{eV} d\epsilon \rho^T(E_F^T - eV + \epsilon) \rho^S(E_F^S + \epsilon) |M|^2 \quad (7.15)$$

Die Ortsabhängigkeit des Tunnelstroms I ist im Übergangsmatrixelement M enthalten. Die einfachste Näherung ist, s -Wellenfunktionen für die Zustände der Spitze anzunehmen (Tersoff-Hamann-Näherung [TH85]).

$$\Psi_s^T(\vec{r}) = C \frac{1}{\kappa r} e^{-\kappa r}, \quad r = |\vec{r} - \vec{R}| \quad (7.16)$$

Dabei ist \vec{R} der Ort der Tunnelspitze. Durch Einsetzen von (7.16) in (7.11) ergibt sich somit für das Matrixelement:

$$M_{\mu s} = \frac{2\pi C \hbar^2}{\kappa m} \Psi_{\mu}^S(\vec{R}) \quad (7.17)$$

Nimmt man weiter an, daß die Zustandsdichte der Spitze gegenüber der Zustandsdichte der Oberfläche weitgehend strukturlos ist ($\rho^T = \text{konst.}$), so ergibt sich für den Tunnelstrom:

$$I(\vec{R}, V) = \frac{16\pi^3 C^2 \hbar^3 e}{\kappa^2 m^2} \rho^T \int_0^{eV} d\epsilon \rho_{\text{lok}}^S(\vec{R}, E_F^S + \epsilon) \quad (7.18)$$

$$\frac{dI}{dV} = \frac{16\pi^3 C^2 \hbar^3 e^2}{\kappa^2 m^2} \rho^T \rho_{\text{lok}}^S(\vec{R}, E_F^S + eV) \quad (7.19)$$

Die Ableitung des Tunnelstroms I nach der Tunnelspannung V ist also proportional zur lokalen Zustandsdichte $\rho_{\text{lok}}^S(\vec{R}, E_F^S + eV)$ am Ort der Tunnelspitze \vec{R} zur Energie $E_F + eV$. Für kleine Tunnelspannungen V wird Gleichung (7.19) zur Tersoff-Hamann-Formel [TH85]. Gleichung (7.19) ist die Grundlage unserer Berechnung von STM-Aufnahmen. Die Formulierung für diskrete Wellenfunktionen lautet:

$$I(\vec{R}, V) \sim \sum_{E_{\mu} \in (E_F, E_F + eV)} \left| \Psi_{\mu}(\vec{R}) \right|^2 \quad (7.20)$$

Die Ψ_μ sind dabei die diskreten, selbstkonsistent berechneten Zustände der Oberflächen-Superzelle.

Die hier verwendete Näherung bricht zusammen, wenn die Tunnelspannung eV nicht klein ist gegenüber der "work function" $\Phi = V_{\text{vakuum}} - E_F$. Das Potential des zusammengesetzten Systems aus Spitze und Potential $U(\vec{r})$ kann zwar immer nach Gleichung (7.1) zerlegt werden, jedoch sind die Potentiale $U_S(\vec{r})$ und $U_T(\vec{r})$ dann nicht die Potentiale der beiden ungestörten Teilsysteme. Der Fehler kann in der WKB-Näherung abgeschätzt werden [Che93]. Die rechte Seite von (7.19) muß nach dieser Abschätzung mit einem spannungsabhängigen Faktor $T(V)$ multipliziert werden.

Der Einfluß der Spitze wird auch von M. Tsukada und N. Shima [TS87] diskutiert. Dort wird gezeigt, daß sich der Einfluß der Spitze auf $\frac{dI}{dV}$ wie folgt abschätzen läßt:

$$\frac{dI}{dV} \sim \rho^T(E_F - eV) \rho_{\text{lok}}^S(\vec{R}(V), E_F - eV) \quad (7.21)$$

Demnach ist der Tunnelstrom proportional zu einer effektiven Zustandsdichte der Spitze und der lokalen Zustandsdichte der Probe, allerdings nicht am Ort der Tunnelspitze, sondern an einem spannungsabhängig leicht verschobenen Ort $\vec{R}(V)$.

In einigen Arbeiten wird die Zustandsdichte der Spitze durch selbstkonsistente Clusterrechnungen berücksichtigt [KT89, TKO90, OT90, KT90, THKK92, BM94, BM95]. Auf diese Weise können auch sogenannte anomale STM-Aufnahmen simuliert werden. Beschränkt man sich auf die Simulation normaler STM-Aufnahmen, so sollte Gleichung (7.20) eine brauchbare Näherung darstellen.

7.2 Experimentelle Ergebnisse

Konstantstrom-Bilder der (110) Oberfläche mit atomarer Auflösung werden für InP und GaP gemessen. Bei Anlegung einer negativen Tunnelspannung U wird aus den besetzten Zuständen am Oberflächenanion (P) herausgetunnelt. Das Untergitter der P-Oberflächenatome wird sichtbar. Bei Anlegen einer positiven Tunnelspannung wird in die unbesetzten Zustände des Oberflächenkations (In, Ga) getunnelt. Das Untergitter der In(Ga)-Oberflächenatome wird sichtbar. In Abbildung 7.2 werden zwei verschiedene Tunnelbilder ($U = -2.4 \text{ V}, +2.3 \text{ V}$) überlagert. Für das Verhältnis $\frac{\Delta a}{a}$, welches die relative Verschiebung in [001]-Richtung der Maxima der beiden Untergitter zueinander beschreibt, ergibt sich ein experimenteller Wert von 0.36 ± 0.07 . Interpretiert man die Maxima der STM-Bilder in erster Näherung als die Positionen der P und der In-Atome, so ergibt die STM-Messung von $\frac{\Delta a}{a}$ einen

deutlich zu großen Wert. Die Zustandsdichtemaxima der beteiligten Zustände, die das Tunnelbild aufbauen, liegen eben nicht direkt über den Oberflächenatomen, sondern in $[00\bar{1}]$ - bzw. $[001]$ -Richtung verschoben.

Unsere Strukturrelaxation ergibt für $\frac{\Delta a}{a}$ einen Wert von 0.21, andere äquivalente Rechnungen ergeben einen Wert von 0.20, experimentelle LEED-Messungen liefern Werte zwischen 0.216 und 0.225 (vergleiche Tabelle 6.14, Seite 98). Ohne Strukturrelaxation wäre das Verhältnis $\frac{\Delta a}{a} = 0.25$.

Bisher wurden STM-Messungen nur im Hinblick auf zwei Zustände interpretiert, einen besetzten "dangling bond" am Anion, der die STM-Bilder bei negativen Tunnelspannungen dominiert und einen unbesetzten "dangling bond" am Kation, der die STM-Bilder bei positiven Tunnelspannungen dominiert. Um den Einfluß weiterer Zustände auf STM-Bilder zu untersuchen, sind sehr genaue Messungen bei verschiedenen und vor allen Dingen auch relativ hohen Tunnelspannungen erforderlich. Ausgewertet werden nur diejenigen Bilder, die eine sehr gute laterale und vertikale Auflösung aufweisen, und bei denen sicher ist, daß die Morphologie nicht durch Spitzenartefakte verfälscht wird. Dies wird an Punktdefekten und Stufen sowie an der Symmetrie der atomaren Auflösung überprüft.

Zur Auswertung der Bilder werden Profilschnitte entlang der $[1\bar{1}0]$ -Richtungen und entlang der $[001]$ -Richtungen ausgelesen und die Korrugationsamplituden berechnet. Als relativ auflösungsunabhängiges Maß für die Morphologie stellt sich das Verhältnis der Amplituden der Korrugation in $[1\bar{1}0]$ -Richtung und in $[001]$ -Richtung heraus. Dieses Verhältnis hängt nur wenig vom Absolutwert der Korrugation ab. Reihen in $[1\bar{1}0]$ -Richtung bedeuten ein Korrugationsverhältnis < 1 , Reihen in $[001]$ -Richtung bedeuten ein Korrugationsverhältnis > 1 .

Für alle untersuchten III-V Halbleiter werden ähnliche Spannungsabhängigkeiten gefunden.

- Besetzte Zustände (negative Tunnelspannung): Eine Erhöhung der Tunnelspannung ändert nur wenig an der Morphologie der STM-Bilder. Das Korrugationsverhältnis steigt mit wachsender Tunnelspannung leicht an (siehe Abbildung 7.3). Das Verhältnis bewegt sich zwischen 0.2 und 0.5. Die STM-Bilder zeigen somit eine klare Reihenbildung in $[1\bar{1}0]$ -Richtung. Die Änderung des Korrugationsverhältnisses wird durch Verminderung der Korrugation in $[001]$ -Richtung mit wachsender Tunnelspannung verursacht, bei mehr oder weniger konstanter Korrugation in $[1\bar{1}0]$ -Richtung. Der Effekt ist aber sehr klein. Die Ergebnisse werden dahingehend interpretiert, daß bei kleinen Tunnelspannungen nur ein Oberflächenzustand am Tunnelprozess teilnimmt, der "dangling

bond" am Anion, bei größeren Spannungen kommt ein weiterer Zustand hinzu, der "back bond" am Anion. Die räumliche Verteilung der Zustandsdichte des "back bonds" wird im Differenzbild zweier Aufnahmen mit großer negativer Tunnelspannung und kleiner negativer Tunnelspannung sichtbar (Abbildung 7.4 und Abbildung 7.5). Ihr Maximum findet sich in $[001]$ -Richtung gegenüber dem Maximum des "dangling bonds" verschoben. Quantitativ ergibt sich für InP eine Verschiebung von 0.17 nm, für GaP eine Verschiebung von 0.22 nm.

- Unbesetzte Zustände (positive Tunnelspannungen): Für die unbesetzten Zustände ist es sehr viel schwieriger, atomare Auflösung zu erzielen. Mit zunehmender Tunnelspannung zeigen sie eine starke Veränderung der Morphologie. Der Haupteffekt ist eine Drehung der Reihenrichtung von $[001]$ -Reihen hin zu $[1\bar{1}0]$ -Reihen. Das Korrugationsverhältnis sinkt zwischen 2 V und 3 V für InP und zwischen 1.5 V und 2.5 V für GaP von Werten deutlich über 1 ($[001]$ -Reihen) auf Werte um 0.5 ($[1\bar{1}0]$ -Reihen) (Abbildung 7.3). Der Effekt wird hauptsächlich durch eine Verminderung der Korrugation in $[1\bar{1}0]$ -Richtung verursacht, während die Korrugation in $[001]$ -Richtung im wesentlichen unverändert bleibt. Abbildung 7.6 zeigt die STM-Bilder für verschiedene positive Spannungen. Die Reihenumkehr ist deutlich erkennbar. Diese Bilder werden wie folgt interpretiert: Für kleine positive Tunnelspannungen bestimmt ein Zustand das Tunnelbild, der "dangling bond" am Kation. Für größere Tunnelspannungen kommt ein weiterer hinzu. Im Gegensatz zu den Bildern der besetzten Zuständen ist dies ein sehr großer Effekt.

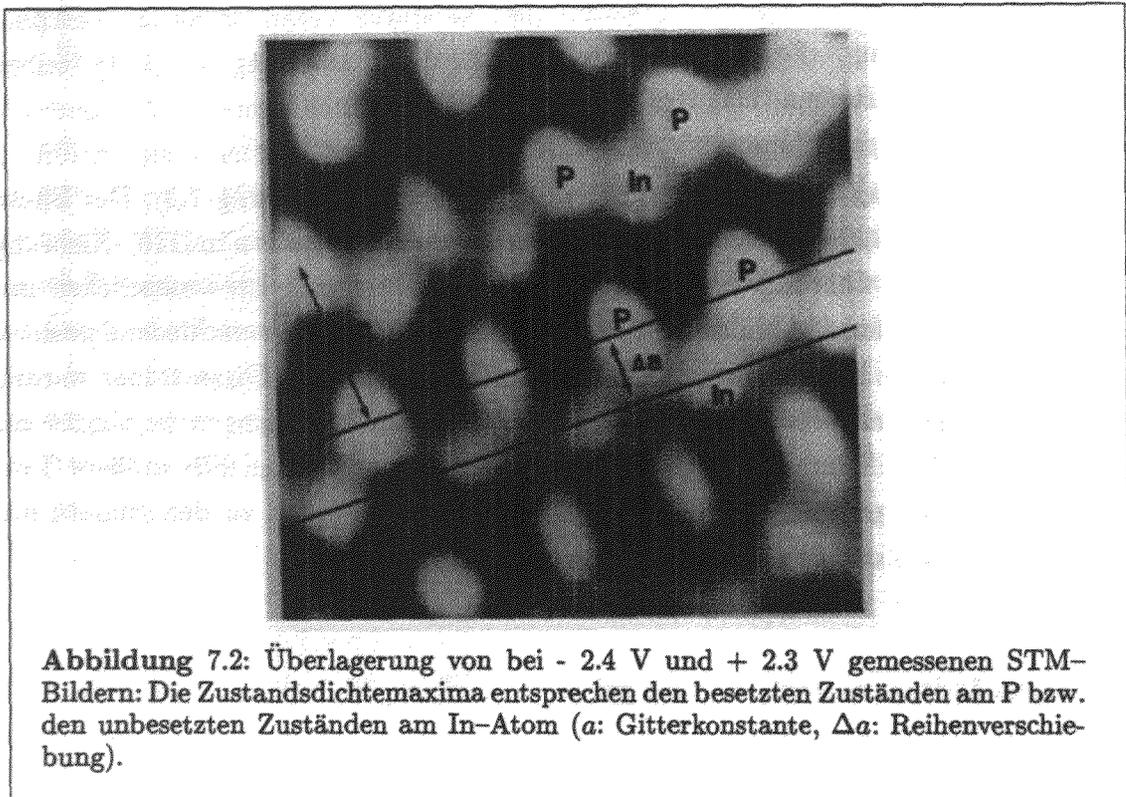


Abbildung 7.2: Überlagerung von bei - 2.4 V und + 2.3 V gemessenen STM-Bildern: Die Zustandsdichtemaxima entsprechen den besetzten Zuständen am P bzw. den unbesetzten Zuständen am In-Atom (a : Gitterkonstante, Δa : Reihenverschiebung).

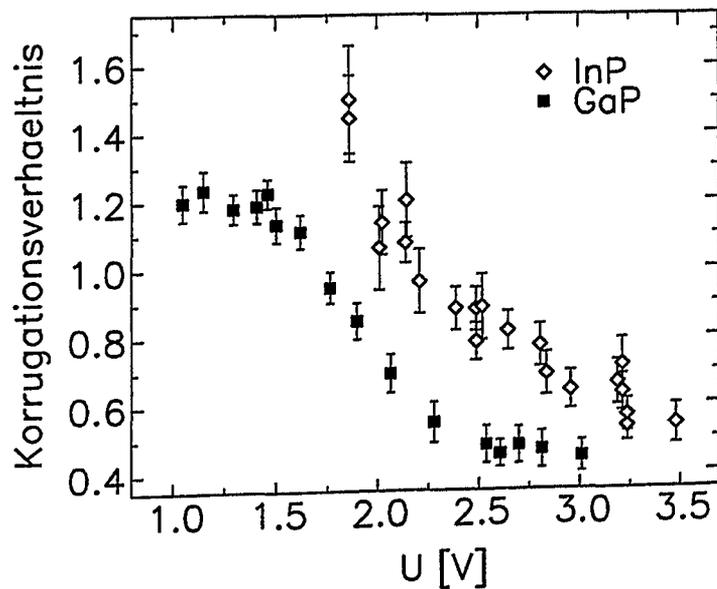
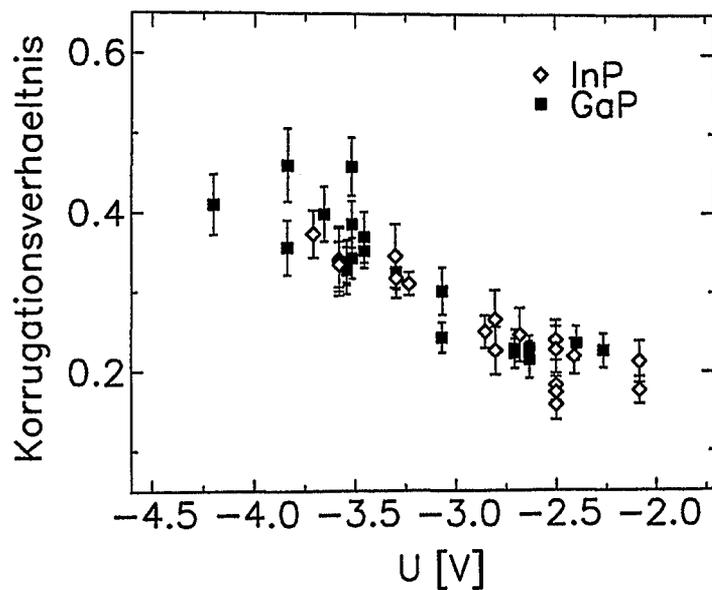


Abbildung 7.3: Korrugationsamplituden: Die obere Abbildung zeigt das Korrugationsverhältnis der Korrugationen in $[1\bar{1}0]$ -Richtung und $[001]$ -Richtung für die besetzten Zustände als Funktion der Spannung für p-GaP(110) und p-InP(110). Die untere Abbildung zeigt das Korrugationsverhältnis für die unbesetzten Zustände für n-GaP(110) und n-InP(110).

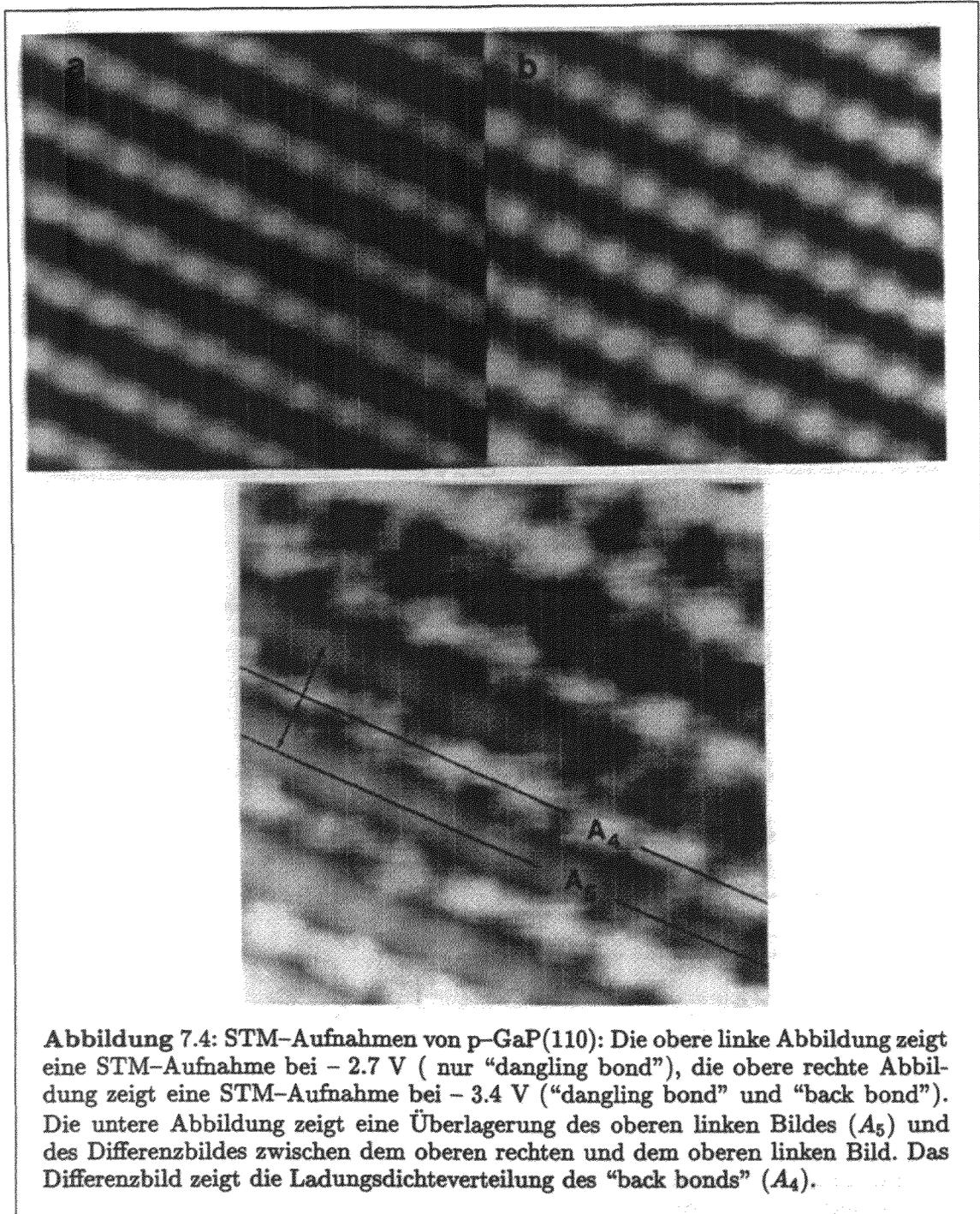


Abbildung 7.4: STM-Aufnahmen von p-GaP(110): Die obere linke Abbildung zeigt eine STM-Aufnahme bei -2.7 V (nur "dangling bond"), die obere rechte Abbildung zeigt eine STM-Aufnahme bei -3.4 V ("dangling bond" und "back bond"). Die untere Abbildung zeigt eine Überlagerung des oberen linken Bildes (A_5) und des Differenzbildes zwischen dem oberen rechten und dem oberen linken Bild. Das Differenzbild zeigt die Ladungsdichteverteilung des "back bonds" (A_4).

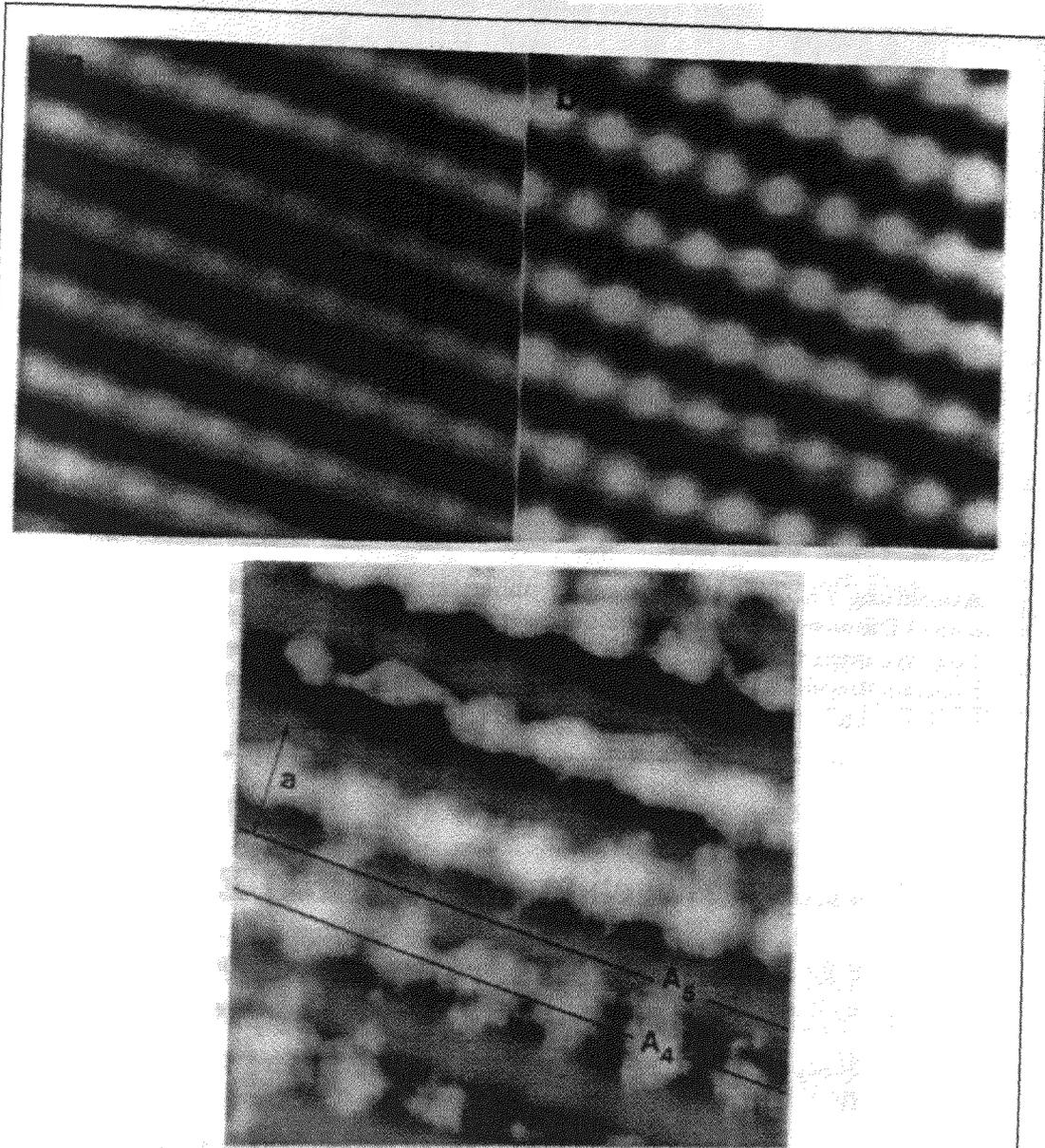
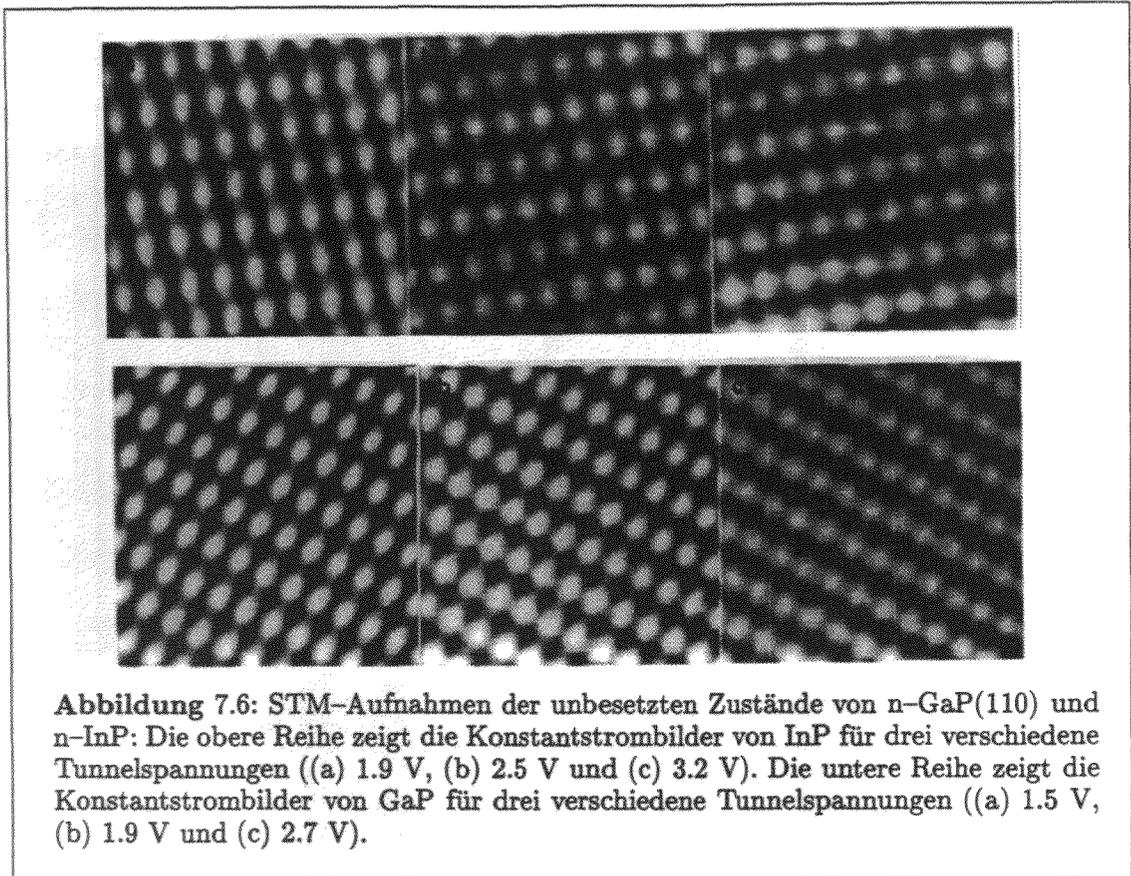


Abbildung 7.5: STM-Aufnahmen von $p\text{-InP}(110)$: Die obere linke Abbildung zeigt eine STM-Aufnahme bei -2.2 V (nur "dangling bond"), die obere rechte Abbildung zeigt eine STM-Aufnahme bei -2.7 V ("dangling bond" und "back bond"). Die untere Abbildung zeigt eine Überlagerung des oberen linken Bildes (A_5) und des Differenzbildes zwischen dem oberen rechten und dem oberen linken Bild. Das Differenzbild zeigt die Ladungsdichteverteilung des "back bonds" (A_4).



7.3 Vergleich zwischen Theorie und Experiment

Grundlage für die Simulation von STM-Aufnahmen ist die Proportionalität von $\frac{dI}{dV}$ und der lokalen Zustandsdichte $\rho_{lok}(\vec{R}, E_F + eV)$ am Ort der Tunnelspitze $\vec{R} = (\vec{r}_{||}, d)$:

$$\frac{dI}{dV}(\vec{r}_{||}, d, V) \sim \rho_{lok}(\vec{r}_{||}, d, E_F - eV) \quad (7.22)$$

$$\rightarrow \sum_{\vec{k}\nu} w_{\vec{k}} |\Psi_{\vec{k}\nu}(\vec{r}_{||}, d)|^2 \delta(E_{\vec{k}\nu} - (E_F + eV)) \quad (7.23)$$

Dabei ist $\vec{r}_{||}$ der Koordinatenvektor parallel zur Probenoberfläche, d der Abstand zwischen Spitze und Probe. In unserem Fall ist dies der Koordinatenvektor $\vec{r}_{||}$ in einer (110) Ebene mit Abstand d zum Oberflächenanion. STM-Aufnahmen und Differenzbilder von STM-Aufnahmen können so direkt durch das Integral über $\frac{dI}{dV}$ bzw. durch Summation über alle beteiligten Zustände simuliert werden. Die \vec{k} -Punkt-Integration wird durch Summation über 100 spezielle \vec{k} -Punkte durchgeführt (Gleichung (6.4), Seite 100).

STM-Aufnahmen sind Konstantstromaufnahmen, i.e. Δd -Abbildungen. Die berechneten Konturbilder zeigen Zustandsdichtekonturen bezüglich \bar{r}_{\parallel} bei konstantem Abstand d . Sie entsprechen also ΔI -Abbildungen. In hinreichender Entfernung zur Oberfläche besteht für den Tunnelstrom folgender funktionaler Zusammenhang:

$$I(\bar{r}_{\parallel}, d) \approx f(\bar{r}_{\parallel}) e^{-\kappa d} \quad (7.24)$$

Da das Verhältnis $\frac{\Delta I}{I}$ bzw. $\frac{\Delta d}{d}$ sehr klein ist, kann die Gleichung linearisiert werden.

$$\frac{\Delta I}{I} = \frac{\Delta f}{f} - \kappa \Delta d \quad (7.25)$$

Sowohl Aufnahmen mit konstantem Strom I , als auch Aufnahmen mit konstantem Abstand d sind also proportional zu Δf . Daher sind Konstantstromaufnahmen unmittelbar mit unseren Konturbildern vergleichbar.

Zwei unbekannte Faktoren machen den quantitativen Vergleich trotzdem schwierig:

- STM-Aufnahmen werden in großem Abstand zur Oberfläche aufgenommen ($\sim 5 - 10 \text{ \AA}$). Die genaue Entfernung ist jedoch im allgemeinen unbekannt. In unseren Rechnungen kann die lokale Zustandsdichte jedoch nur bis maximal 4.5 \AA ausgewertet werden. Darüberhinaus sind die Wellenfunktionen schon so stark abgefallen, daß unser Computerprogramm nur noch numerisches Rauschen liefert (siehe Abbildungen 6.24, Seite 130 und 6.30, Seite 138).
- Wegen der Bandverbiegungen an der Oberfläche entspricht die angelegte Tunnelspannung nicht dem Abstand eV zur Fermienergie, bei der die lokale Zustandsdichte ausgewertet werden muß. Die experimentelle Spannungsskala ist viel größer als der korrespondierende Abstand zur Fermienergie. Alle weiteren Spannungs- bzw. Energieangaben sind von der Valenzbandkante aus gerechnet. Die berechneten Leitungsbänder wurden um 0.50 eV nach oben verschoben, um die berechnete Bandlücke² an das experimentelle Bandgap anzupassen.

Analog zu Gleichung (6.26) (Seite 113) wird zunächst die Spannungsabhängigkeit des Korrugationsverhältnisses für verschiedene Abstände von Spitze und Probe berechnet. Dies ist in Abbildung 7.7 festgehalten. Das Korrugationsverhältnis ist stark abstandsabhängig. Die Abbildung 7.7 ist mit den experimentellen InP-Werten aus Abbildung 7.3 zu vergleichen. Ignoriert man die unterschiedlichen Spannungsskalen, so werden die spannungsabhängigen Trends richtig wiedergegeben. Korrugationsverhältnisse < 1 bei negativen Tunnelspannungen bedeuten $[1\bar{1}0]$ -Reihen. Der

²Im Rahmen der LDA berechnete Bandlücken sind falsch (siehe Tabelle 6.5, Seite 85).

Wechsel von Werten > 1 zu Werten < 1 bei positiven Tunnelspannungen bedeutet eine Reihenumkehr von $[001]$ -Reihen zu $[1\bar{1}0]$ -Reihen. Mit wachsendem Abstand zur Oberfläche wird auch die quantitative Übereinstimmung immer besser. Der "dangling bond" am Oberflächenanion ist bei -0.66 eV (siehe Kapitel 6.6.2.3) als starke Änderung des Korrugationsverhältnisses sehr gut identifizierbar, der "back bond" jedoch nicht. Nach unserer Analyse der Oberflächenzustände kann man auch keinen isolierten "back bond" Zustand mit fester Energie erwarten, sondern nur eine Gruppe von mehr oder weniger verbreiterten Resonanzen mit "back bond" Charakter. Die Abbildungen 7.8, 7.9 und 7.10 zeigen schließlich die simulierten Tunnelbilder für drei verschiedene Abstände $d = 2.23, 4.38$ und 6.53 a.u. zwischen Spitze und Oberfläche. Die linken Abbildungen zeigen die besetzten Zustände, die rechten Abbildungen die unbesetzten Zustände. Die oberen und die unteren Bilder zeigen die besetzten Zustände bis zu einer Spannung V_1 bzw. V_2 von der Valenzbandkante aus gemessen, i.e. $\int_0^{eV_1} dE \rho_{tok}(\vec{r}_{||}, d, E)$, bzw. $\int_0^{eV_2} dE \rho_{tok}(\vec{r}_{||}, d, E)$. Die mittleren Abbildungen zeigen die Differenzbilder $\int_{eV_1}^{eV_2} dE \rho_{tok}(\vec{r}_{||}, d, E)$.

Die Spannungen V_1 und V_2 werden anhand der Energiedispersion der beteiligten Zustände ausgewählt (siehe Abbildung 6.25, Seite 131 und Abbildung 6.31, Seite 139). Für die besetzten Zustände werden die Spannungen $V_1 = -0.9$ eV und $V_2 = -1.5$ eV so gewählt, daß das erste Bild im wesentlichen vom "dangling bond" am Oberflächenanion bestimmt wird, das Differenzbild hingegen im wesentlichen von den "back bond"-Zuständen. Für die unbesetzten Zustände werden die Spannungen $V_1 = +2.2$ eV und $V_2 = +2.8$ eV so gewählt, daß das erste Bild im wesentlichen vom " p_z "-Zustand am Oberflächenkation bestimmt wird und die anderen im Abschnitt 6.6.2.4 beschriebenen Zustände zum Differenzbild beitragen.

Die simulierten STM-Aufnahmen der besetzten Zustände zeigen nur wenig Abstandsabhängigkeit. Die Interpretation der experimentellen Resultate können bestätigt werden. Bei niedrigen Tunnelspannungen bestimmt der "dangling bond" die STM-Aufnahmen, bei höheren Spannungen kommt die breite Resonanz der "back bond"-Zustände hinzu. Die Verschiebung in $[001]$ -Richtung zwischen "dangling bond" (Abbildung 7.10: Bild oben links) und "back bond" (Abbildung 7.10: Bild mitte links) ergibt sich bei $d = 6.53$ a.u. zu $\Delta a = 0.63$ Å. Dies ist wesentlich niedriger als der experimentelle Wert von 1.7 Å. Eine quantitative Übereinstimmung ist aber auch nicht zu erwarten, da die simulierten Bilder bei anderen Abständen d gerechnet sind als die Experimente. Da die Zustandsdichte des "dangling bonds" groß ist im Vergleich zu der Zustandsdichte der "back bond"-Zustände, dominiert der "dangling bond" das STM-Bild.

Die simulierten STM-Aufnahmen der unbesetzten Zustände zeigen eine sehr starke

Abstandsabhängigkeit. Auch hier kann die Interpretation der experimentellen Resultate im wesentlichen bestätigt werden. Zwar existieren nahe an der Oberfläche ($d = 2.23$ a.u.) auch Beiträge zum STM-Bild mit Zustandsdichte über dem Oberflächenanion, doch fallen diese schneller ins Vakuum ab bzw. deren Korrugation fällt schneller ins Vakuum ab als die der Zustände mit Ladungsdichteanhäufung über dem Oberflächenkation. STM-Aufnahmen mit solch kleinen Abständen zwischen Spitze und Probe sind allerdings nicht möglich, da die Spitze wegen des großen Tunnelstroms durch Überhitzung zerstört würde. Bei realistischen Abständen bleibt der Effekt der Reihenumkehr, der sich auch in den simulierten Aufnahmen zeigt. In Übereinstimmung mit dem Experiment wird mit zunehmender Tunnelspannung das Verhältnis $\frac{\Delta I}{I}$ bzw. $\frac{\Delta d}{d}$ schnell kleiner. Insgesamt ergibt sich für die unbesetzten Zustände folgendes in sich konsistente Bild: Bei sehr niedriger Tunnelspannung bestimmt der " p_z "-Zustand das STM-Bild. Bei höheren Tunnelspannungen kommen viele verschiedenartige Zustände hinzu mit unterschiedlichster Morphologie. Obwohl diese Zustände insgesamt einen großen Beitrag zur Zustandsdichte bringen, kann man sie kaum im Tunnelbild erkennen, da sich die unterschiedlichen Morphologien gegenseitig wegmitteln. Zusammen verursachen sie die Reihenumkehr und die Verminderung von $\frac{\Delta I}{I}$.

Die Bilder oben links und unten rechts aus Abbildung 7.10 können noch ausgewertet werden, um den relativen Abstand $\frac{\Delta a}{a}$ der Maxima zu bestimmen, der sich bei Vorzeichenumkehr der Tunnelspannung ergibt. Der berechnete Wert beträgt 0.45, der experimentelle Wert beträgt 0.36 ± 0.07 .

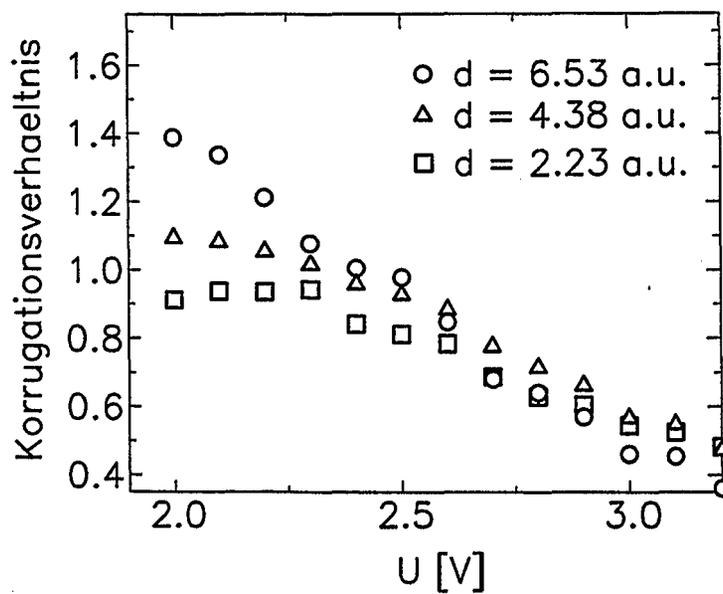
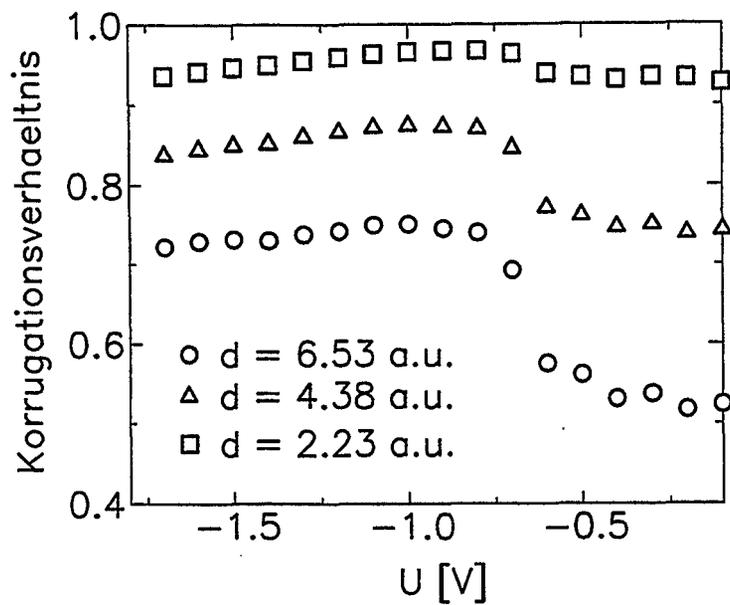
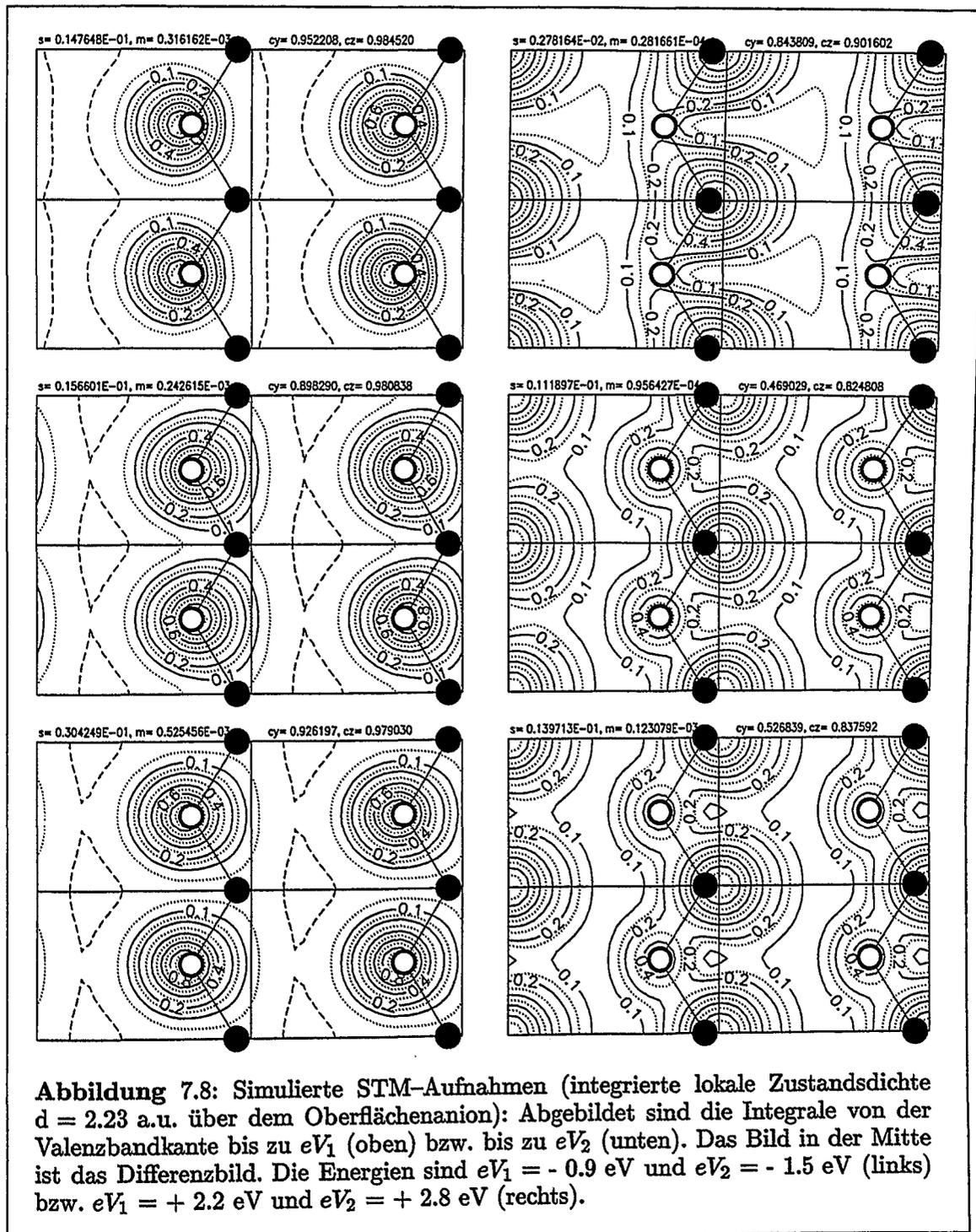
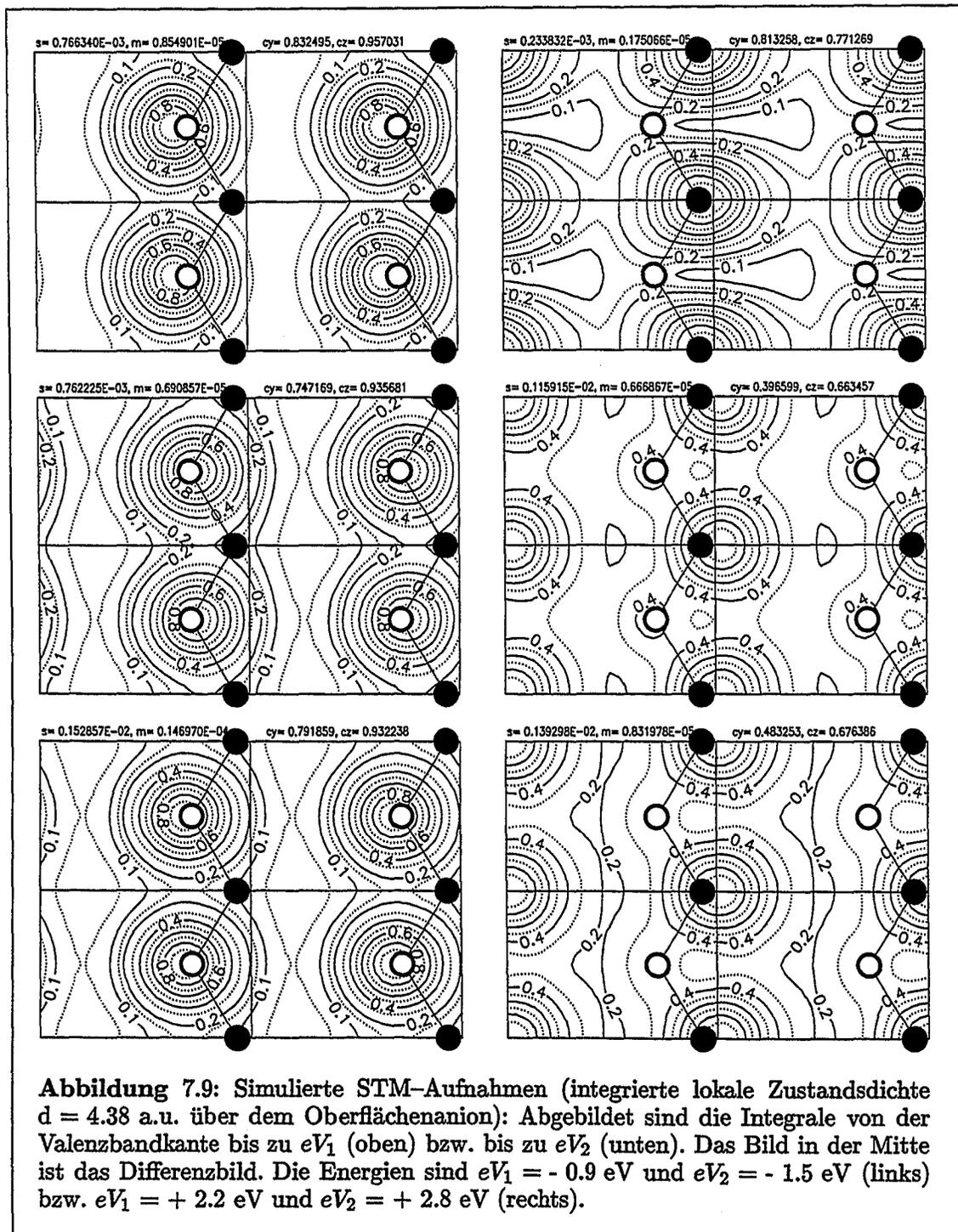
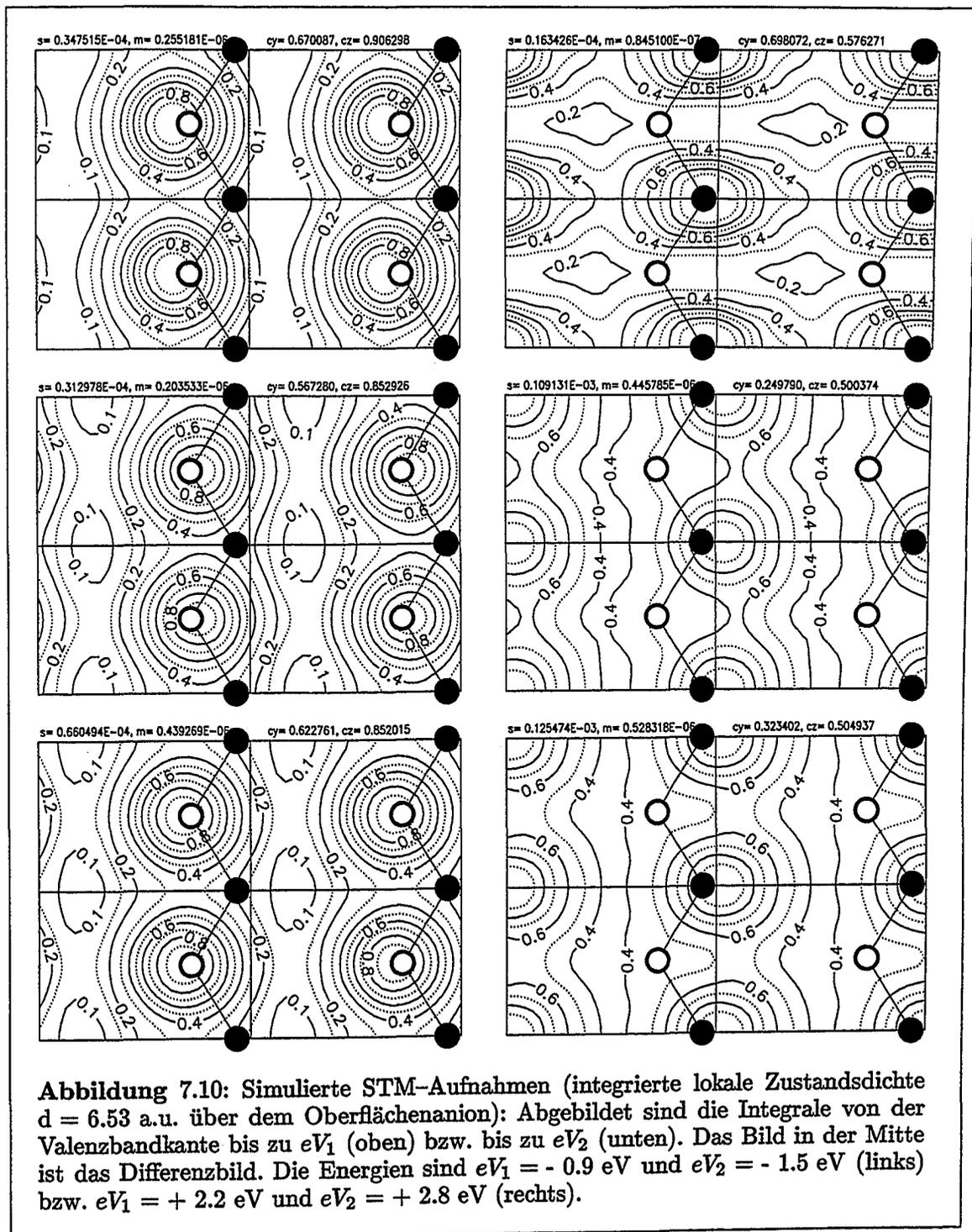


Abbildung 7.7: Berechnete Korrugationsamplituden für drei verschiedene Abstände d der Spitze vom Oberflächenanion: Die obere Abbildung zeigt das Korrugationsverhältnis der Korrugationen in $[1\bar{1}0]$ -Richtung und $[001]$ -Richtung für die besetzten Zustände als Funktion der Spannung. Die untere Abbildung zeigt das Korrugationsverhältnis für die unbesetzten Zustände.







Kapitel 8

Resümee und Ausblick

Das hier vorgestellte Computerprogramm erlaubt es, die atomare und elektronische Struktur von Halbleiteroberflächen ohne anpaßbare Parameter effizient und numerisch stabil zu berechnen. Bei entsprechender Rechnerausstattung sollten 30–100 Atome pro Einheitszelle berücksichtigt werden können. Unser Hauptinteresse galt nicht der Formulierung einer Car–Parrinello–Dynamik, sondern einer effizienten Formulierung der Molekularstatik, i.e. der schnellen Relaxation der Atomkoordinaten in lokale Minima, wie sie bei Halbleiteroberflächen vorwiegend vorkommen. Deshalb haben wir den BFGS Algorithmus implementiert. Die Oberflächenzustände, die der Halbleiteroberfläche einen teils metallischen Charakter verleihen können, stellen eine besondere Herausforderung an den Algorithmus zur Lösung des Eigenwertproblems dar. Der in dieser Arbeit vorgestellte Davidson–Kosugi–Algorithmus in Verbindung mit der Quasi–Newton–Methode erweist sich diesbezüglich als sehr stabil und sehr effizient.

Wir haben die Struktur der (110) Oberflächen der III–V Halbleiter AlAs, InP und InAs bestimmt und gute Übereinstimmung mit den gemessenen Strukturdaten bzw. früheren Rechnungen gefunden. Wir finden für alle III–V Halbleiter die in der Literatur bekannte Struktur der (110) Oberfläche. Ferner ist die Übereinstimmung der atomaren Gleichgewichtsstruktur mit den experimentellen Daten typischerweise ~ 0.05 a.u. und liegt im Streubereich der Experimente.

Speziell haben wir in dieser Arbeit den Zusammenhang zwischen atomarer Struktur, elektronischer Struktur und den daraus zu erwartenden Tunnelbildern erforscht. Weltweit besteht ein großer Bedarf für eine atomistische Interpretation der lateralen elektronischen Struktur, die das RTM leisten kann, aber bisher wurden nur ganz wenige theoretische Versuche unternommen, die Morphologie von RTM–Aufnahmen theoretisch zu erfassen. In dieser Arbeit haben wir die RTM–Aufnahmen der InP (110) Oberfläche spannungsabhängig simuliert. Ausgangspunkt sind sehr einfache

Annahmen über die Tunnelspitze. In der Tersoff–Hamann–Näherung ergibt sich die laterale Verteilung des Tunnelstroms dann aus den elektronischen Zuständen der Oberfläche durch Energieintegration der lokalen Zustandsdichte. Diese einfache Näherung ist erstaunlich leistungsfähig. Ein Vergleich zwischen Theorie und Experiment ergibt sehr gute Übereinstimmung. Sogar Details der RTM–Aufnahmen werden qualitativ richtig wiedergegeben. Dies erlaubt, die Bilder auf die elektronische Struktur der Oberfläche zurückzuführen. Wir haben damit gezeigt, daß spannungsabhängige Details der Morphologie hochaufgelöster RTM–Aufnahmen theoretisch beschrieben und erklärt werden können.

Die Rastertunnelmikroskopie alleine liefert zwar Messungen der Elektronendichte mit atomarer Auflösung, aber diese Meßergebnisse erlauben nicht ohne weiteres Rückschlüsse auf die Positionen der Atome. Dies behindert insbesondere die Interpretation von RTM–Aufnahmen von Defekten, da es schwierig ist, die Anwesenheit oder Abwesenheit bzw. den chemischen Charakter einzelner Atome zu identifizieren. Dieses Problem stellt sich bei allen mehrkomponentigen Kristallen mit kovalentem Bindungsanteil. Durch den kombinierten Einsatz von Theorie und Experiment sollte es in Zukunft möglich sein, verschiedenartigste Defekte auf Halbleiteroberflächen mit Hilfe des Rastertunnelmikroskops sicher zu identifizieren. Dies bildet z.B. die Grundlage für eine quantitative Messung der Konzentrationen spezieller Defekte. Die im folgenden dargestellte Problemstellung, die zur Zeit im IFF verfolgt wird, ist ein gutes Beispiel für die potentiellen Möglichkeiten der Verbindung von Theorie und Experiment:

Die durch die Kombination von Theorie und Experiment erhaltene Informationsfülle eröffnet auch erstmalig die Möglichkeit, Bulkdefekte atomar auf dem Umweg über die Oberfläche zu untersuchen. Für diese Untersuchungen würde man III–V Halbleiterkristalle spalten, die Defekte auf der Spaltoberfläche identifizieren und deren Konzentrationen bestimmen. Nach Separation von Oberflächeneffekten kann dann auf die Konzentration der identifizierten Defekte im Volumen zurückgeschlossen werden. Diese Technik würde es z.B. erlauben, die mikroskopischen Mechanismen der Kompensation von Si–Dotieratomen in GaAs zu bestimmen und damit die Dotiereffektivität für optoelektronische Anwendungen zu optimieren.

Eine andere wichtige Anwendung wäre die chemische Unterscheidung verschiedener Fremdatome in III–V Halbleitern sowie ternären III–V Halbleitern mittels RTM. Dieser Aspekt ist insbesondere in der Herstellung und Anwendung von Heterostrukturen und Halbleiterübergittern von Bedeutung. Für die quantitative Untersuchung solcher Problemstellungen kann die Kombination von Theorie und Experiment nicht nur hilfreich, sondern häufig auch entscheidend sein.

Anhang A

Formelsammlung

sphärische Besselfunktionen $j_l(x)$:

Differentialgleichung:

$$\left[- \left(\frac{d}{dr} \right)^2 + \frac{l(l+1)}{r^2} \right] (r j_l(qr)) = q^2 (r j_l(qr)) \quad (\text{A.1})$$

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:

$$j_l(x) \xrightarrow{x \rightarrow \infty} \frac{\sin\left(x - \frac{1}{2}l\pi\right)}{x} \quad (\text{A.9})$$

Orthogonalitätsrelation: $(q_1, q_2 > 0)$

$$q_1^2 \int_0^\infty dr r^2 j_l(q_1 r) j_l(q_2 r) = \frac{\pi}{2} \delta(q_1 - q_2) \quad (\text{A.10})$$

Legendre Polynome $P_l(x)$:

Differentialgleichung:

$$\left[(1-x^2) \left(\frac{d}{dx} \right)^2 - 2x \left(\frac{d}{dx} \right) + l(l+1) \right] P_l(x) = 0 \quad (\text{A.11})$$

Formel von Rodrigues:

$$P_l(x) = \frac{1}{2^l l!} \left(\frac{d}{dx} \right)^l (x^2 - 1)^l \quad (\text{A.12})$$

$$P_0(x) = 1 \quad (\text{A.13})$$

$$P_1(x) = x \quad (\text{A.14})$$

$$P_2(x) = \frac{1}{2} (3x^2 - 1) \quad (\text{A.15})$$

$$P_3(x) = \frac{1}{2} (5x^3 - 3x) \quad (\text{A.16})$$

$$P_4(x) = \frac{1}{2 \cdot 4} (5 \cdot 7x^4 - 2 \cdot 3 \cdot 5x^2 + 1 \cdot 3) \quad (\text{A.17})$$

Rekursionsformel:

$$(l+1)P_{l+1}(x) = (2l+1)xP_l(x) - lP_{l-1}(x) \quad (\text{A.18})$$

Orthogonalitätsrelation:

$$\int_{-1}^{+1} P_l(x) P_l(x) dx = \frac{2\delta_{ll}}{2l+1} \quad (\text{A.19})$$

zugeordnete Legendre Polynome $P_l^m(x)$:

Differentialgleichung:

$$\left[(1-x^2) \left(\frac{d}{dx} \right)^2 - 2x \left(\frac{d}{dx} \right) + l(l+1) - \frac{m^2}{1-x^2} \right] P_l^m(x) = 0 \quad (\text{A.20})$$

Formel von Rodrigues:

$$P_l^m(x) = \frac{(1-x^2)^{\frac{m}{2}}}{2^l l!} \left(\frac{d}{dx} \right)^{l+m} (x^2 - 1)^l \quad (\text{A.21})$$

Orthogonalitätsrelation:

$$\int_{-1}^{+1} P_l^m(x) P_l^m(x) dx = \frac{2\delta_{ll} (l+m)!}{(2l+1) (l-m)!} \quad (\text{A.22})$$

$$\int_{-1}^{+1} P_l^m(x) P_l^{m'}(x) \frac{dx}{1-x^2} = \frac{\delta_{mm'} (l+m)!}{m (l-m)!} \quad (\text{A.23})$$

komplexe Kugelflächenfunktionen $Y_{lm}(\theta, \varphi)$:

Definitionsformel:

$$Y_{lm}(\theta, \varphi) = (-1)^m \sqrt{\frac{(2l+1)(l-m)!}{4\pi(l+m)!}} P_l^m(\cos\theta) e^{im\varphi} \quad (\text{A.24})$$

Orthogonalitätsrelation:

$$\int_0^{2\pi} d\varphi \int_0^\pi \sin\theta d\theta Y_{lm}^*(\theta, \varphi) Y_{l'm'}(\theta, \varphi) = \delta_{ll'} \delta_{mm'} \quad (\text{A.25})$$

reelle Kugelflächenfunktionen $\mathcal{Y}_{lm}(\theta, \varphi)$:

Definitionsformel:

$$\begin{aligned} \mathcal{Y}_{lm} &= \frac{(-1)^m}{\sqrt{2}} (Y_{lm} + Y_{lm}^*) = \sqrt{2} (-1)^m \operatorname{Re}(Y_{lm}) && \text{für } (m > 0) \\ \mathcal{Y}_{lm} &= Y_{l0} && \text{für } (m = 0) \\ \mathcal{Y}_{l-m} &= \frac{1}{\sqrt{2i}} (Y_{lm} - Y_{lm}^*) = \sqrt{2} \operatorname{Im}(Y_{lm}) && \text{für } (m > 0) \end{aligned} \quad (\text{A.26})$$

Orthogonalitätsrelation:

$$\int_0^{2\pi} d\varphi \int_0^\pi \sin\theta d\theta \mathcal{Y}_{lm}(\theta, \varphi) \mathcal{Y}_{l'm'}(\theta, \varphi) = \delta_{ll'} \delta_{mm'} \quad (\text{A.27})$$

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	$\frac{1}{2} \sqrt{5} (3z^2 - 1)$	$\frac{1}{2} \sqrt{5} (3z^2 - 1)$	
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	$\mp \frac{1}{4} \sqrt{21} (5z^2 - 1) (x \pm iy)$	$\frac{1}{4} \sqrt{42} (5z^2 - 1) x$	$-\frac{1}{4} \sqrt{42} (5z^2 - 1) y$
3	0	$\frac{1}{2} \sqrt{7} (5z^2 - 3) z$	$\frac{1}{2} \sqrt{7} (5z^2 - 3) z$	

Vermischtes zu P_l , P_l^m und Y_{lm} :

$$P_l(-x) = (-1)^l P_l(x) \quad (\text{A.29})$$

$$P_l^m(-x) = (-1)^{l+m} P_l^m(x) \quad (\text{A.30})$$

$$Y_{lm}(\pi - \theta, \pi + \varphi) = (-1)^l Y_{lm}(\theta, \varphi) \quad (\text{A.31})$$

$$P_l^{-m}(x) = (-1)^m \frac{(l-m)!}{(l+m)!} P_l^m(x) \quad (\text{A.32})$$

$$Y_{l-m}(\theta, \varphi) = (-1)^m Y_{lm}^*(\theta, \varphi) \quad (\text{A.33})$$

$$(2l+1) P_l(\cos \gamma_{\vec{x}, \vec{x}'}) = \begin{cases} \sum_m 4\pi Y_{lm}^*(\Omega_{\vec{x}}) Y_{lm}(\Omega_{\vec{x}'}) \\ \sum_m 4\pi Y_{lm}(\Omega_{\vec{x}}) Y_{lm}^*(\Omega_{\vec{x}'}) \\ \sum_m 4\pi \mathcal{Y}_{lm}(\Omega_{\vec{x}}) \mathcal{Y}_{lm}(\Omega_{\vec{x}'}) \end{cases} \quad (\text{A.34})$$

Darstellung von ebenen Wellen durch Besselfunktionen:

$$e^{i\vec{k}\vec{r}} = 4\pi \sum_{l,m} i^l Y_{lm}^*(\Omega_{\vec{k}}) Y_{lm}(\Omega_{\vec{r}}) j_l(kr) \quad (\text{A.35})$$

$$= 4\pi \sum_{l,m} i^l Y_{lm}(\Omega_{\vec{k}}) Y_{lm}^*(\Omega_{\vec{r}}) j_l(kr) \quad (\text{A.36})$$

Gauß-Hermite-Integration:

$$\int_{-\infty}^{\infty} dr e^{-r^2} f(r) = \sum_{i=1}^N \omega_{i,N} f(r_{i,N}) \quad (\text{A.37})$$

Hermite-Polynome H_N :

$$H_N(r) = (-1)^N e^{r^2} \left(\frac{d}{dr} \right)^N \left[e^{-r^2} \right] \quad (\text{A.38})$$

Nullstellen $r_{i,N}$:

$$H_N(r_{i,N}) = 0 \quad (\text{A.39})$$

Gewichte $\omega_{i,N}$:

$$\omega_{i,N} = \frac{2^N \sqrt{\pi}}{N!} (2 - \delta_{r_{i,N}, 0}) \left[\frac{d}{dr} H_N(r) \Big|_{r=r_{i,N}} \right]^{-2} \quad (\text{A.40})$$

Die Ewaldsumme:

$$\sum_{\vec{R}} \frac{(1 - \delta_{\vec{r}, \vec{R}})}{|\vec{r} - \vec{R}|} = \sum_{\vec{G}} \frac{1}{\vec{G}} \frac{4\pi}{G^2} e^{-i\vec{G}\vec{r}} \quad (\text{A.41})$$

$$\begin{aligned} \frac{1}{\sqrt{\Omega}} \sum_{\vec{R}} \left\langle \vec{G} \left| \frac{(1 - \delta_{\vec{r}, \vec{R}})}{|\vec{r} - \vec{R}|} \right. \right\rangle_{\Omega} &= \frac{1}{\sqrt{\Omega}} \left\langle \vec{G} \left| \frac{1}{r} \right. \right\rangle \quad (\text{A.42}) \\ &= \frac{1}{\Omega} \int d^3r \frac{1}{r} e^{i\vec{G}\vec{r}} = \frac{1}{\Omega} \frac{4\pi}{G^2} \end{aligned}$$

$$\sum_{\vec{R}} \frac{(1 - \delta_{\vec{r}, \vec{R}})}{|\vec{r} - \vec{R}|} = \frac{4\pi}{\Omega} \sum_{\vec{G}} \frac{1}{G^2} \cos(\vec{G}\vec{r}) e^{-\frac{G^2}{4\eta^2}} - \frac{\pi}{\eta^2 \Omega}$$

$$+ \sum_{\vec{R}} \frac{\operatorname{erfc}(\eta |\vec{r} - \vec{R}|)}{|\vec{r} - \vec{R}|} (1 - \delta_{\vec{r}, \vec{R}}) - 2 \frac{\eta}{\sqrt{\pi}} \delta_{\vec{r}, \vec{0}} \quad (\text{A.43})$$

Die Fehlerfunktion:

$$\operatorname{erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z dt e^{-t^2} \quad (\text{A.44})$$

$$\operatorname{erfc}(z) = \frac{2}{\sqrt{\pi}} \int_z^\infty dt e^{-t^2} = 1 - \operatorname{erf}(z) \quad (\text{A.45})$$

$$e^{z^2} \operatorname{erfc}(z) \leq 1 \quad (\text{A.46})$$

$$\lim_{z \rightarrow \infty} e^{z^2} \operatorname{erfc}(z) = 0 \quad (\text{A.47})$$

Poisson-Gleichung:

$$u(\vec{r}) = \int d^3 r' \frac{f(\vec{r}')}{|\vec{r} - \vec{r}'|} \Leftrightarrow \Delta u(\vec{r}) = -4\pi f(\vec{r}) \quad (\text{A.48})$$

trigonometrische Formeln:

$$\cos x = \frac{1}{2} (e^{ix} + e^{-ix}) \quad (\text{A.49})$$

$$\sin x = \frac{1}{2i} (e^{ix} - e^{-ix}) \quad (\text{A.50})$$

$$\sin^2 x = \frac{1}{2} (1 - \cos 2x) \quad (\text{A.51})$$

$$\sin x \sin y = \frac{1}{2} (\cos(x - y) - \cos(x + y)) \quad (\text{A.52})$$

$$\sin x \cos y = \frac{1}{2} (\sin(x - y) + \sin(x + y)) \quad (\text{A.53})$$

$$\cos x \sin y = \frac{1}{2} (\sin(x + y) - \sin(x - y)) \quad (\text{A.54})$$

$$\cos x \cos y = \frac{1}{2} (\cos(x - y) + \cos(x + y)) \quad (\text{A.55})$$

$$\sin x - \sin y = 2 \sin \frac{x - y}{2} \cos \frac{x + y}{2} \quad (\text{A.56})$$

$$\sin x + \sin y = 2 \sin \frac{x + y}{2} \cos \frac{x - y}{2} \quad (\text{A.57})$$

$$\sin 2x = 2 \sin x \cos x \quad (\text{A.58})$$

trigonometrische Reihen:

$$\sum_{n=1}^N \sin(2nx) = \frac{\sin((N+1)x) \sin(Nx)}{\sin x} \quad (\text{A.59})$$

$$\sum_{n=1}^N \cos(2nx) = \frac{\sin((N+1)x) \cos(Nx)}{\sin x} = \frac{1}{2} \left(\frac{\sin((2N+1)x)}{\sin x} - 1 \right) \quad (\text{A.60})$$

$$\sum_{n=1}^N \sin^2(nx) = \frac{N}{2} \left(1 - \frac{\cos((N+1)x) \sin(Nx)}{N \sin(x)} \right) \quad (\text{A.61})$$

$$= \frac{N}{2} \left(1 + \frac{1}{2N} - \frac{\sin((2N+1)x)}{2N \sin(x)} \right)$$

$$\sum_{n=1}^N n \sin(2nx) = -\frac{1}{2} \frac{d}{dx} \left(\sum_{n=1}^N \cos(2nx) \right) \quad (\text{A.62})$$

$$= \frac{\sin(2(N+1)x)}{4 \sin^2 x} - \frac{(N+1) \cos((2N+1)x)}{2 \sin x}$$

$$\sum_{n=1}^N n^2 \cos(2nx) = -\frac{1}{4} \left(\frac{d}{dx} \right)^2 \left(\sum_{n=1}^N \cos(2nx) \right) \quad (\text{A.63})$$

$$= \frac{(n+1)(2n+1) \sin((2n+1)x)}{4 \sin x} \\ + \frac{(n+1)}{8 \sin^2 x} (3 \cos((2N+2)x) + \cos(2Nx)) \\ - \frac{1}{8 \sin^3 x} (\sin((2N+1)x) + \sin((2N+3)x))$$

$$\sum_{n=1}^N n^2 \sin^2(nx) = \frac{1}{2} \sum_{n=1}^N n^2 (1 - \cos(2nx)) \quad (\text{A.64})$$

$$= \frac{1}{12} (N(N+1)(2N+1)) - \frac{1}{2} \sum_{n=1}^N n^2 \cos(2nx)$$

trigonometrische Integrale:

$$\int dE \frac{\cos(2\sqrt{E}x)}{2\sqrt{E}} = \int d\sqrt{E} \cos(2\sqrt{E}x) = \frac{\sin(2\sqrt{E}x)}{2x} \quad (\text{A.65})$$

$$\int dE \frac{\sin(2\sqrt{E}x)}{2\sqrt{E}x} = \int d\sqrt{E} \frac{2\sqrt{E} \sin(2\sqrt{E}x)}{2x} \quad (\text{A.66}) \\ = 2 \left(\frac{\sin(2\sqrt{E}x)}{(2x)^3} - \frac{\sqrt{E} \cos(2\sqrt{E}x)}{(2x)^2} \right)$$

geometrische Reihe:

$$\sum_{n=n_1}^{n_2-1} y^n = \frac{y^{n_2} - y^{n_1}}{y - 1} \quad (\text{A.67})$$

eine Darstellung der δ -Funktion:

$$\left| \int_0^t dt' e^{i\Delta t'} \right|^2 = \frac{4}{\Delta^2} \sin^2 \frac{\Delta t}{2} \rightarrow 2\pi t \delta(\Delta) \quad (\text{A.68})$$

die zweite Green'sche Formel:

$$\iiint_G d^3x (P\Delta Q - Q\Delta P) = \iint_{\partial G} d\vec{S} \cdot (P\nabla Q - Q\nabla P) \quad (\text{A.69})$$

Literaturverzeichnis

- [AHS91] José Luiz A. Alves, Jörk Hebenstreit, and Matthias Scheffler. Calculated atomic structures and electronic properties of GaP, InP, GaAs, and InAs (110) surfaces. *Phys. Rev. B*, 44(12):6188, 1991.
- [And64] D.G. Anderson. *J. Assoc. Comput. Mach.*, 12:547, 1964.
- [Bar61] J. Bardeen. Tunneling from a many-particle point of view. *Phys. Rev. Lett.*, 6:57, 1961.
- [BE88] F. Bechstedt and R. Enderlein. *Semiconductor Surfaces and Interfaces*. Physical Research. Akademie Verlag, Berlin, 1988.
- [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.
- [BK92] D.M. Bylander and L. Kleinman. Exact Phillips-Kleinman-like pseudopotentials for transition metals. *Phys. Rev. B*, 46(15):9837, 1992.
- [Blo28] F. Bloch. *Z. Physik*, 52:555, 1928.
- [Blö90] P.E. Blöchl. Generalized separable potentials for electronic-structure calculations. *Phys. Rev. B*, 41(8):5414, 1990.
- [Blü] S. Blügel. generalized anderson method. private communication.
- [Blü88] S. Blügel. *First Principles Calculations of the Electronic Structure of Magnetic Overlayers on Transition Metal Surfaces*. PhD thesis, RWTH Aachen, 1988.

- [BM94] J.M. Bass and C.C. Matthai. Theoretical simulation of scanning-tunneling-microscopy images of the $\text{GaAs}(001)$ $\beta(2 \times 4)$ and $\beta(4 \times 2)$ surfaces. *Phys. Rev. B*, 50(15):11212, 1994.
- [BM95] J.M. Bass and C.C. Matthai. Scanning-tunneling-microscopy and spectroscopy simulation of the $\text{GaAs}(110)$ surface. *Phys. Rev. B*, 52(7):4712, 1995.
- [Bro65] C.G. Broyden. *Math. comput.*, 19:577, 1965.
- [Bro66] C.G. Broyden. *Math. comput.*, 21:368, 1966.
- [CA80] D. M. Ceperley and B. J. Alder. quantum monte carlo. *Phys. Rev. Lett.*, 45:566, 1980.
- [CC76] J.R. Chelikowsky and M.L. Cohen. (110) surface states in III-V and II-VI zinc-blende semiconductors. *Phys. Rev. B*, 13(2):826, 1976.
- [CC79] J.R. Chelikowsky and M.L. Cohen. Self-consistent pseudopotential calculation for the relaxed (110) surface of GaAs. *Phys. Rev. B*, 20(10):4150, 1979.
- [CCMS90] H. Carstensen, R. Claessen, R. Manzke, and M. Skibowski. Direct determination of III-V semiconductor surface band gaps. *Phys. Rev. B*, 41:9880, 1990.
- [CH70] M.L. Cohen and V. Heine. The fitting of pseudopotentials to experimental data and their subsequent application. In H.Ehrenreich, F. Seitz, and D. Turnbull, editors, *Solid State Physics*, volume 24, pages 38–248. Academic Press, New York, 1970.
- [Cha79] D. J. Chadi. (110) surface atomic structures of covalent and ionic semiconductors. *Phys. Rev. B*, 19(4):2074, 1979.
- [Che93] C. Julian Chen. *Introduction to Scanning Tunneling Microscopy*. Oxford series in optical and imaging sciences. Oxford University Press, New York, 1993.
- [Cho92] N.Y. Chou. Reformulation of generalized separable pseudopotentials. *Phys. Rev. B*, 45(20):11465, 1992.
- [CP85] R. Car and M. Parrinello. *Phys. Rev. Lett.*, 55:2471, 1985.

- [Dav75] W.E. Davidson. *J. Comput. Phys.*, 17:97, 1975.
- [DMP⁺85] C. B. Duke, C. Mailhot, A. Paton, D. J. Chadi, and A. Kahn. Atomic geometries of the (110) surfaces of III-V compound semiconductors: Determination by total-energy minimization and elastic low-energy electron diffraction. *J. Vac. Sci. Technol. B*, 3:1087, 1985.
- [Ebe93] Ph. Ebert. Quantitative Untersuchung von Defekten auf Oberflächen von III-V-Verbindungshalbleitern mit dem Rastertunnelmikroskop. Berichte des Forschungszentrums Jülich 2785, Forschungszentrum Jülich, Jülich, 1993.
- [Eng92] B. Engels. Konstruktion von glatten normerhaltenden Pseudopotentialen. Master's thesis, RWTH Aachen, 1992.
- [Fee94] R. M. Feenstra. Tunneling spectroscopy of the (110) surface of direct-gap III-V semiconductors. *Phys. Rev. B*, 50(7):4561, 1994.
- [FHK94] J. Furthmüller, J. Hafner, and G. Kresse. *ab initio* calculation of the structural and electronic properties of carbon and boron nitride using ultrasoft pseudopotentials. *Phys. Rev. B*, 50(21):15606, 1994.
- [Fle87] R. Fletcher. *Practical Methods of Optimization*. John Wiley & Sons, New York, 1987.
- [FS87a] R. M. Feenstra and J. A. Stroscio. Tunneling spectroscopy of the GaAs(110) surface. *J. Vac. Sci. Technol. B*, 5:923, 1987.
- [FS87b] A. C. Ferraz and G. P. Srivastava. Determination of the surface geometry of GaAs (110) by the total energy and force methods. *Surf. Sci.*, 182:161-178, 1987.
- [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.
- [GVM89] X. Gonze, J.P. Vigneron, and J.-P. Michenaud. Fast diagonalisation of non-local pseudopotential hamiltonians. *J. Phys.: Condens. Matter*, 1:525, 1989.

- [Ham89] D.R. Hamann. Generalized norm-conserving pseudopotentials. *Phys. Rev. B*, 40(5):2980, 1989.
- [Hei70] V. Heine. The pseudopotential concept. In H.Ehrenreich, F. Seitz, and D. Turnbull, editors, *Solid State Physics*, volume 24, pages 1–36. Academic Press, New York, 1970.
- [HK64] P. Hohenberg and W. Kohn. *Phys. Rev.*, 136:B864, 1964.
- [HSC79] D.R. Hamann, M. Schlüter, and C. Chiang. Norm-conserving pseudopotentials. *Phys. Rev. Lett.*, 43(20):1494, 1979.
- [HW70] V. Heine and D. Weaire. Pseudopotential theory of cohesion and structure. In H.Ehrenreich, F. Seitz, and D. Turnbull, editors, *Solid State Physics*, volume 24, pages 249–463. Academic Press, New York, 1970.
- [Ihm88] J. Ihm. Total energy calculations in solid state physics. *Rep. Prog. Phys.*, 51:105–142, 1988.
- [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.
- [Jac77] D. Jacobs. *The State of the Art in Numerical Analysis*. Academic Press, New York, 1977.
- [JG89] R. O. Jones and O. Gunnarsson. The density functional formalism, its applications and prospects. *Rev. Mod. Phys.*, 61:689, 1989.
- [JM76] J.R.Chelikowsky and M.L.Cohen. *Phys. Rev. B*, 14(2):556–582, 1976.
- [Kah83a] A. Kahn. *J. Vac. Sci. Technol. B*, 1:511, 1983.
- [Kah83b] A. Kahn. Semiconductor surface structures. *Surf. Sci. Rep.*, 3:193ff, 1983.
- [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.
- [KHN92] G. Kresse, J. Hafner, and R. J. Needs. Optimized norm-conserving pseudopotentials. *J. Phys.: Condens. Matter*, 4:7451–7468, 1992.

- [Kos84] N. Kosugi. *J. Comput. Phys.*, 55:426, 1984.
- [KS65] W. Kohn and L. J. Sham. *Phys. Rev.*, 140:A1133, 1965.
- [KT89] Katsuyoshi Kobayashi and Masaru Tsukada. Effect of microscopic tip electronic state on STM image of graphite. *J. Phys. Soc. Jpn.*, 58:2238–2241, 1989.
- [KT90] Katsuyoshi Kobayashi and Masaru Tsukada. Simulation of scanning tunneling microscope image based on electronic states of surface/tip system. *J. Vac. Sci. Technol. A*, 8:170, 1990.
- [LB82] *Semiconductors: Physics of Group IV Elements and III–V Compounds*, volume 17a of *Landolt–Börnstein: Numerical Data and Functional Relationships in Science and Technology (New Series), Group III: Crystal and Solid State Physics*, chapter 2. Springer–Verlag, Berlin, 1982.
- [LB84] *Semiconductors: Technology of III–V, II–VI and Non–Tetrahedrally Bounded Compounds*, volume 17d of *Landolt–Börnstein: Numerical Data and Functional Relationships in Science and Technology (New Series), Group III: Crystal and Solid State Physics*, chapter 6.3, page 13. Springer–Verlag, Berlin, 1984.
- [LB93] *Physics of Solid Surfaces: Structure*, volume 24a of *Landolt–Börnstein: Numerical Data and Functional Relationships in Science and Technology (New Series), Group III: Crystal and Solid State Physics*, chapter 2.2 Surface reconstruction and relaxation, page 196. Springer–Verlag, Berlin, 1993.
- [LB94] *Physics of Solid Surfaces: Electronic and Vibrational Properties*, volume 24b of *Landolt–Börnstein: Numerical Data and Functional Relationships in Science and Technology (New Series), Group III: Crystal and Solid State Physics*, chapter 3.2 Electronic structure of surfaces: semiconductors, page 352ff. Springer–Verlag, Berlin, 1994.
- [Lev79] M. Levy. *Proc. Natl. Acad. Sci. USA*, 76:6062, 1979.
- [Lev82] M. Levy. *Phys. Rev. A*, 26:1200, 1982.
- [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.

- [Liu78] B. Liu. In *Report on Workshop: Numerical Algorithms in Chemistry: Algebraic Methods*, page 49. Lawrence Berkley Lab., Univ. of California, 1978.
- [LQPH93] J. S. Lin, A. Qteish, M. C. Payne, and V. Heine. Optimized and transferable nonlocal separable *ab initio* pseudopotentials. *Phys. Rev. B*, 47(8):4174, 1993.
- [LRF⁺74] L.Ley, R.A.Pollak, F.R.McFeely, S.P.Kowalczyk, and D.A.Shirley. *Phys. Rev. B*, 9(2):600–621, 1974.
- [LVHK87] L.Sorba, V.Hinkel, H.U.Middelmann, and K.Horn. *Phys. Rev. B*, 36(15):8075–8081, 1987.
- [MBCM81] F. Manghi, C.M. Bertoni, C. Calandra, and E. Molinari. Theoretical study of the electronic structure of GaP (110). *Phys. Rev. B*, 24(10):6029, 1981.
- [MDC85] C. Mailhot, C. B. Duke, and D. J. Chadi. *Surf. Sci.*, 149:366, 1985.
- [MJW78] V. L. Moruzzi, J. F. Janak, and A. R. Williams. Pergamon, New York, 1978.
- [MMBC82] F. Manghi, E. Molinari, C.M. Bertoni, and C. Calandra. Self-consistent pseudopotential calculation of the electronic properties of the InP (110) surface. *J. Phys. C.: Solid State Phys.*, 15:1099–1109, 1982.
- [MPBH⁺92] R. Maboudian, K. Pond, V. Bressler-Hill, M. Wassermeier, P.M. Petroff, G.A.D. Briggs, and W.H. Weinberg. Tunneling spectroscopy on the GaAs(110) surface: effect of dopant concentration. *Surf. Sci. Lett.*, 275:L662–L668, 1992.
- [MV79] A.H. MacDonald and S.H. Vosko. relativistic extension of LSDA. *J. Phys. C.*, 12:2977, 1979.
- [O'R91] E.P. O'Reilly. *Properties of Indium Phosphide*, volume 6 of *EMIS Datareviews Series*, chapter 6.1, page 93ff. Inspec, 1991.
- [OT90] S. Ohnishi and M. Tsukada. Effect of the microscopic electronic states of the tip on the scanning tunneling microscopy image. *J. Vac. Sci. Technol. A*, 8:174, 1990.

- [Pic89] W.E. Pickett. Pseudopotential methods in condensed matter applications. *Computer Physics Reports*, 9:115–198, 1989.
- [PM77] J.D. Pack and H.J. Monkhorst. *Phys. Rev. B*, 16:1748, 1977.
- [PTA⁺92] M.C. Payne, M.P. Teter, D.C. Allan, T.A. Arias, and J.D. Joannopoulos. Iterative minimization techniques for *ab initio* total-energy calculations: molecular dynamics and conjugate gradients. *Computer Physics Reports*, 64(4):1045–1097, 1992.
- [Ric93] P. Richard. Iterative Lösung der Schrödinger Gleichung für große Systeme. Master's thesis, RWTH Aachen, 1993.
- [RRKJ90] A.M. Rappe, K.M. Rabe, E. Kaxiras, and J.D. Joannopoulos. Optimized pseudopotentials. *Phys. Rev. B*, 41(2):1227, 1990.
- [RRTP88] B. Reihl, T. Riesterer, M. Tschudy, and P. Perfetti. Surface electronic bands of GaAs (110) determined by angle-resolved inverse photoemission. *Phys. Rev. B*, 38:13456, 1988.
- [SAMJ89] E.L. Shirley, D.C. Allan, R.M. Martin, and J.D. Joannopoulos. Extended norm-conserving pseudopotentials. *Phys. Rev. B*, 40(6):3652, 1989.
- [SSMW83] G.P. Srivastava, I. Singh, V. Montgomery, and R.H. Williams. The electronic structure of cleaved indium phosphide (110) surfaces: experiment and theory. *J. Phys. C.: Solid State Phys.*, 16:3627–3640, 1983.
- [SSO92] M. Saito, O. Sugino, and A. Oshiyama. Real-space-partitioned separable pseudopotential. *Phys. Rev. B*, 46(4):2606, 1992.
- [TH85] T. Tersoff and D. R. Hamann. Theory of the scanning tunneling microscope. *Phys. Rev. B*, 31:805, 1985.
- [THKK92] Masaru Tsukada, Nobuyuki Isshiki Hiroyuki Kageshima, and Katsuyoshi Kobayashi. Connected vacuum tail method and its application to scanning tunneling microscopy. *Surf. Sci.*, 266:253–258, 1992.
- [TKIK91] Masaru Tsukada, Katsuyoshi Kobayashi, Nobuyuki Isshiki, and Hiroyuki Kageshima. First-principles theory of scanning tunneling microscopy. *Surf. Sci. Rep.*, 13:265–304, 1991.

- [TKO90] Masaru Tsukada, Katsuyoshi Kobayashi, and Shuhei Ohnishi. First-principles theory of the scanning tunneling microscopy simulation. *J. Vac. Sci. Technol. A*, 8:160, 1990.
- [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.
- [TMX84] S.Y. Tong, W.N. Mei, and G. Xu. *J. Vac. Sci. Technol. B*, 2:393, 1984.
- [TPA89] M.P. Teter, M.C. Payne, and D.C. Allan. *Phys. Rev. B*, 40:12255, 1989.
- [TS87] Masaru Tsukada and Nobuyuki Shima. Theory of electronic processes of scanning tunneling microscopy. *Jap. Jour. Phys. Soc.*, 56:2875–2885, 1987.
- [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.
- [vBG80] U. von Barth and C.D. Gelatt. Validity of the frozen-core approximation and pseudopotential theory for cohesive energy calculations. *Phys. Rev. B*, 21(6):2222, 1980.
- [vBH72] U. von Barth and L. Hedin. local density approximation. *J. Phys. C.*, 5:1629, 1972.
- [VWN80] S. H. Vosko, L. Wilk, and N. Nussair. local density approximation. *Can. J. Phys.*, 58:1200, 1980.
- [Wei] M. Weinert. usage of the debeye temperature for initial quasi newton steps. private communication.
- [Wie94] R. Wiesendanger. *Scanning probe microscopy and spectroscopy: methods and applications*. Cambridge University Press, 1994.
- [Wig34] E. Wigner. *Phys. Rev.*, 46:1002, 1934.

- [YC82] M.T. Yin and M.L. Cohen. Theory of *ab initio* pseudopotential calculations. *Phys. Rev. B*, 25(12):7403, 1982.
- [Zun80] A. Zunger. Nonlocal pseudopotential calculation electronic properties of relaxed GaAs (110) surface. *Phys. Rev. B*, 22(2):959, 1980.

Danksagung

An dieser Stelle möchte ich allen danken, die das Gelingen dieser Arbeit ermöglicht haben:

Ich bedanke mich bei Herrn Prof. Dr. K.W.Schroeder und Herrn Dr. Stefan Blügel für die interessante Themenstellung und für die fortwährende Betreuung und Unterstützung auch in den schwierigen Phasen dieser Arbeit.

Herrn Dr. Philipp Ebert und Peter Richard möchte ich besonders danken für die gute Zusammenarbeit.

Für sorgfältiges Korrekturlesen und Hilfestellung bei der Fertigstellung der Arbeit danke ich besonders Peter Richard, Philipp Ebert und Markus Aretz.

Herrn Professor Dr. H. Müller-Krumbhaar, unserem Institutsleiter, danke ich für die Möglichkeit, an seinem Institut diese Promotionarbeit erstellen zu dürfen.

Ohne sie namentlich aufzuzählen möchte ich mich bei allen Mitarbeitern des Instituts Theorie III für die Hilfsbereitschaft und für die angenehme Arbeitsatmosphäre bedanken, bei den Damen und Herren, die die Rechner betreuen und bei Rainer Schmitz, den man praktisch jederzeit über alles fragen kann, was Rechner, Software etc. betrifft.

Schließlich möchte ich mich besonders bei meinen Eltern bedanken, die mir alles ermöglicht haben.

Jül-3215
März 1996

ISSN 0944-2952