

# Relativistische Molekülstruktur

—

## Implementierung der Geometrieoptimierung und neuer Dichtefunktionale

Inaugural-Dissertation  
zur Erlangung der  
Doktorwürde der Naturwissenschaften  
(Dr. rer. nat.)

vorgelegt beim Fachbereich Physik  
der Universität Kassel

von  
Dipl.-Phys. Sven Varga  
aus Friedrichroda/Thüringen

Juni 1998

*”Denn die Wissenschaft bedeutet nicht beschauliches Ausruhen im Besitz gewonnener sicherer Erkenntnis, sondern sie bedeutet rastlose Arbeit und stets vorwärtsschreitende Entwicklung nach einem Ziel, das wir wohl dichterisch zu ahnen, aber niemals verstandesmäßig voll zu erfassen vermögen.“*

Max Planck

1. Gutachter: Prof. Dr. B. Fricke
2. Gutachter: Prof. Dr. D. Kolb

weitere Mitglieder der Prüfungskommission:

Prof. Dr. F. Bosch, GSI  
Prof. Dr. K. Röhl

Tag der mündlichen Prüfung: 09. Juli 1998

# Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>1</b>
<b>2</b>	<b>Grundlagen</b>	<b>5</b>
2.1	Der Hamiltonoperator . . . . .	6
2.2	Der Dichtefunktional-Ansatz . . . . .	8
2.3	Relativistische Dichtefunktionaltheorie . . . . .	9
2.4	Die relativistischen Kohn-Sham Gleichungen . . . . .	12
<b>3</b>	<b>Dichtefunktionalanwendungen</b>	<b>15</b>
3.1	Einleitung und Motivation . . . . .	15
3.2	Die Lokaldichte-Näherung . . . . .	16
3.3	Nichtlokale Dichtefunktionale . . . . .	17
3.4	Die relativistischen Energiefunktionale . . . . .	19
3.5	Eine wichtige Approximation – Fit der Ladungsdichte . . . . .	21
3.6	Ergebnisse der Rechnungen . . . . .	23
3.6.1	Die klassischen Beispiele . . . . .	24
3.6.2	Untersuchungen an hochrelativistischen Systemen . . . . .	26
<b>4</b>	<b>Der Energiegradient</b>	<b>29</b>
4.1	Der Funktionalansatz . . . . .	29
4.1.1	Vorbemerkungen und Vereinbarungen . . . . .	30
4.1.2	Funktional ohne Nebenbedingung (NB) . . . . .	31
4.1.3	Funktional mit parameterunabhängiger Nebenbedingung . . . . .	32
4.1.4	Funktional mit parameterabhängiger Nebenbedingung . . . . .	33
4.2	Simultane Lösung zweier Variationsprobleme . . . . .	33
4.2.1	1.Schritt: Die Variationsgleichungen . . . . .	34
4.2.2	2.Schritt: Differentiation . . . . .	37
4.3	Die Gradiententerme explizit . . . . .	40
4.3.1	Variationsgrößen und Geometrieabhängigkeit . . . . .	41
4.3.2	Geometrieabhängigkeit der totalen Energie . . . . .	42
4.3.3	Geometrieabhängigkeit der Nebenbedingungen . . . . .	48
4.3.4	Hellmann-Feynman Gradient . . . . .	49
4.3.5	Korrekturterme . . . . .	50
4.4	Frozen-Core Näherung und Energiegradient . . . . .	53

<b>5</b>	<b>Die Geometrieoptimierung</b>	<b>59</b>
5.1	Die Beschreibung der Minimumsuche . . . . .	60
5.1.1	Die Schrittweitenoptimierung . . . . .	62
5.1.2	Die Neubestimmung der inversen Hessematrix . . . . .	66
5.2	Ergebnisse . . . . .	68
5.2.1	Voruntersuchungen . . . . .	68
5.2.2	Optimierung dreidimensionaler Strukturen . . . . .	71
5.3	Zusammenfassung der Rechnungen . . . . .	74
<b>6</b>	<b>Fazit</b>	<b>77</b>
6.1	Zusammenfassung . . . . .	77
6.2	Ausblick und Entwicklungspotential . . . . .	78
<b>A</b>		<b>81</b>
A.1	Hellmann–Feynman Theorem . . . . .	81
A.2	Gradientenformel . . . . .	82
A.3	Hessematrix-Update . . . . .	84
A.4	$\gamma$ -Matrizen und relativistische Notation . . . . .	85
A.5	Feldtheoretischer Zugang zur Elektron-Elektron Wechselwirkung . . . . .	86
A.6	Die relativistischen Korrekturfunktionen $\Phi_{0/2}^{L/T}$ . . . . .	89

# Kapitel 1

## Einleitung

Bereits in der Vorlesung über Mechanik lernt man als Physikstudent, daß allen physikalischen Erscheinungen fundamentale Prinzipien zugrunde liegen. Da ist beispielsweise das Prinzip der kleinsten Wirkung, was eine Aussage über das Stattfinden und den Verlauf physikalischer Prozesse erlaubt. Weiterhin beobachten wir, daß alle in der Natur ablaufenden Vorgänge ohne äußeres Zutun in den meisten Fällen von sich aus in einen stabilen Zustand übergehen, dessen Gesamtenergie am kleinsten ist. Diese Prinzipien gelten auch im subatomaren Bereich. Atome gehen unter Aussendung eines Photons der Energie  $h\nu$  vom angeregten in den Grundzustand über, viele chemische Reaktionen laufen spontan ab, wobei die Gesamtenergie der Endprodukte kleiner als die der Ausgangsstoffe ist. Das angemessene Handwerkszeug zur Beschreibung der Phänomene in der Welt der Atome stellt die Quantenmechanik bereit.

Mit der vorliegenden Arbeit soll der Versuch unternommen werden, elektronische Eigenschaften und Strukturen von Molekülen, die insbesondere sehr schwere Elemente enthalten, durch die Weiterentwicklung eines relativistischen *all-electron* Programms eingehender zu studieren. Heutzutage gibt es eine ganze Reihe solcher Molekülstrukturprogramme, die in den unterschiedlichsten Bereichen Anwendung gefunden haben. Je nach konkreten Fragestellungen und Anforderungen an die Genauigkeiten stehen die unterschiedlichsten Rechenprogramme zur Verfügung. Dreh- und Angelpunkt der Qualität der Rechnungen ist dabei das Maß, mit dem man sich zufrieden gibt, um die intramolekularen Wechselwirkungen zwischen den positiv geladenen Kernen, den negativ geladenen Elektronen und beiden Teilchenarten untereinander zu beschreiben; mit anderen Worten also, wie man konkret die Wechselwirkungspotentiale zu behandeln gedenkt. Dabei gilt die einfache Regel: je genauer, umso aufwendiger. Am 'physikalischsten' sind reine *ab initio* Beschreibungen, in denen außer den Naturkonstanten keine weiteren Parameter eingehen. Leider ist damit ein immenser rechentechnischer Aufwand verbunden. Um komplexere Systeme handhaben zu können, müssen geeignete Näherungen gemacht werden.

Was sowohl das Bemühen um eine exakte Beschreibung als auch die Entwicklung von Näherungsmethoden angeht, so hat die Theoretische Chemie auf diesem Gebiet bereits viel geleistet. Dabei liegt das Hauptinteresse vor allem an der geometrischen Struktur

eines Moleküls oder einer chemischen Verbindung. Sie bestimmt im wesentlichen deren physikalische und chemische Eigenschaften. In der organischen Chemie beispielsweise bietet sich eine schier unüberschaubare Vielfalt an homologen Verbindungen, die sich nur in ihrer Struktur unterscheiden. Gerade aber die strukturellen Unterschiede können mitunter beträchtliche Auswirkungen auf die chemischen Eigenschaften haben. Eine besondere Stellung nimmt dabei diejenige geometrische Anordnung der Atomkerne ein, deren Energie am kleinsten ist. Dies ist die Gleichgewichtslage.

Hinsichtlich der heute zur Verfügung stehenden Möglichkeiten der theoretischen Beschreibung läßt sich das Gebiet der Atom- und Molekülphysik grob in zwei Kategorien einteilen. Da sind einmal die 'klassischen', auf Hartree-Fock-Niveau basierenden Methoden mit all ihren Weiterentwicklungen wie Multikonfigurationsverfahren, wobei der Aufwand bereits bei Atomen sehr groß ist und die Umsetzung auf molekularer Ebene gerade in den Anfängen steckt. Eine Alternative zu diesen Methoden sind die Dichtefunktionale, die sich innerhalb der letzten dreißig Jahre zunehmender Popularität erfreuen und den Rechenaufwand drastisch verkleinern. Erst durch ihren Einsatz sind auch komplexe Systeme mit bis zu hunderten von Atomen einigermaßen zufriedenstellend beschreibbar. Auch wir benutzen in unserer Arbeitsgruppe Dichtefunktionalmethoden und verwendeten dafür bisher einen recht einfachen Ausdruck für das Dichtefunktional, mit dem man in den allermeisten Fällen schon sehr gute Resultate erzielen und Interpretationen im Sinne von Trendvorhersagen machen konnte. Aufgrund der physikalischen Näherung unseres Ansatzes erhalten wir jedoch zu tiefe Bindungsenergien und zu große Bindungsabstände. Die theoretischen Entwicklungen auf dem Gebiet der Dichtefunktionale haben es ermöglicht, daß heutzutage leistungsfähigere Funktionale zur Verfügung stehen.

Nachdem man in den letzten Jahren die schweren Elemente mit den Kernladungszahlen 111 und 112 nachgewiesen hat, gewinnen diese Elemente zunächst insofern immer mehr an Bedeutung, als daß wir anhand deren physiko-chemischen Eigenschaften unser Grundverständnis vom Aufbau des Periodensystems der Elemente weiter vertiefen können. Da wir mit dem in unserer Gruppe weiterentwickelten Molekülprogramm in der Lage sind, die bei diesen Elementen wichtigen relativistischen Effekte im Prinzip richtig zu berücksichtigen, liegt es nahe, das Programm zur Unterstützung der gegenwärtig stattfindenden Experimente [1] heranzuziehen.

Konkret leiten sich nun zwei klare Aufgabenstellungen ab. Die eingangs erwähnte Weiterentwicklung des Programms soll **erstens** Gleichgewichtsgeometrien kleiner Moleküle mit schweren Atomen liefern und **zweitens** eine Interpretation der Ergebnisse aufgrund der Verwendung verschiedener verbesserter Dichtefunktionale ermöglichen. Diese sind mittlerweile auch auf relativistische Vielteilchen-Coulombsysteme erweitert worden und vermögen die Grundzustandsenergien schwerer Atome mit ähnlicher Güte zu beschreiben wie die 'herkömmlichen' Verfahren.

Beiden Zielsetzungen gemein ist die Verfügbarkeit von Gradienten der molekularen Größen bezüglich der internuklearen Abstände, weswegen die Hauptarbeit der programmtechnischen Umsetzung in der Berechnung dieser Gradienten bestand – für die Dichtefunktionalanwendungen der Gradient der Dichte, für die Gleichgewichtsgeometrie der Gradient der Totalen Energie.

Im nächsten Kapitel wollen wir die Grundlagen darlegen. Wir beginnen dabei mit dem unseren Rechnungen zugrunde liegenden Hamiltonoperator und gehen auf einige Aspekte der relativistischen Dichtefunktionaltheorie (RDFT) ein. Im dritten Kapitel stellen wir die ersten Anwendungen vor. Die meiste Entwicklungsarbeit hat der Energiegradient in Anspruch genommen. Daher ist diesem ein eigenes Kapitel gewidmet. Daran anschließend wird die Geometrieoptimierung mit ersten Resultaten vorgestellt. Abschließend wollen wir zusammenfassen und auf mögliche Weiterentwicklungen eingehen.

In dieser Arbeit verwenden wir atomare Einheiten:

$\hbar = m_e = e = 1$ . Die Lichtgeschwindigkeit ist  $c = 137.0359$ . Einer atomaren Energieeinheit (a.u.) entsprechen 27.21165 eV. An einigen Stellen fassen wir die wichtigsten Ergebnisse einzelner Unterabschnitte in schattiert unterlegten Kästen zusammen.



# Kapitel 2

## Grundlagen

In diesem ersten Kapitel wollen wir die Grundlagen zur Behandlung des molekularen relativistischen Vielteilchenproblems darlegen. Wir beginnen mit dem unseren Rechnungen zugrunde liegenden Hamiltonoperator, der den physikalischen Rahmen unserer Betrachtungen absteckt. Nach einigen motivierenden Bemerkungen zur Idee der Dichtefunktionaltheorie (DFT) wollen wir diese in der relativistischen Formulierung angeben und anschließend das Berechnungsschema der uns am meisten interessierenden physikalischen Größe, der totalen Energie, aufzeigen.

In einer relativistischen Theorie ist man von vornherein mit neuen physikalischen Phänomenen konfrontiert. Aufgrund der Möglichkeit von Paarerzeugungs- und Vernichtungsprozessen ist die **Teilchenzahl**<sup>1</sup> **keine Erhaltungsgröße** mehr, wohingegen sich die **Ladung nicht ändert**, denn der Hamiltonoperator der Quantenelektrodynamik (QED) vertauscht mit dem Ladungsoperator. Soll die Bewegung von Spin  $\frac{1}{2}$ -Teilchen (wie Elektronen) untersucht werden, ist die Dirac-Gleichung die angemessene Gleichung, die es zu lösen gilt. Sie ist ein gekoppeltes System zweier Differentialgleichungen erster Ordnung in der zeitlichen und in den räumlichen Komponenten. Ihr Lösungsspektrum enthält neben den Elektronen auch deren Antiteilchen, die Positronen. Eine Beschreibung dieser relativistischen Effekte kann strenggenommen nur im Rahmen der relativistischen Quantenfeldtheorie erfolgen. Dies gilt ebenso für die von uns in dieser Arbeit benutzten relativistischen Dichtefunktionale. Eine rigorose Ableitung der Gleichungen ist sehr kompliziert und man steht vor der Aufgabe, die dabei auftretenden mathematischen Schwierigkeiten (Divergenzen und deren Eliminierung durch Renormierungen) zu bewältigen. Beschränkt man sich auf chemische Anwendungen, lassen sich die wesentlichen Effekte sehr gut auch mit ungleich weniger Aufwand erfassen. Das setzt voraus, daß die dabei gemachten Näherungen angemessen sind. Eine instruktive und übersichtliche Darstellung dieses Gegenstandes ist beispielsweise in [2, 3] zu finden.

---

<sup>1</sup>Gemeint ist hier nicht die Leptonenzahl.

## 2.1 Der Hamiltonoperator

Die Gesamtenergie eines Moleküls oder Atomclusters, die auch als totale Energie bezeichnet wird, setzt sich jeweils zusammen aus den Wechselwirkungsenergien der negativ geladenen Elektronen und der positiv geladenen Kerne untereinander, den kinetischen Energien beider Teilchenarten sowie der elektrostatischen Wechselwirkungsenergie von Kernen und Elektronen. Haben wir  $N$  Elektronen mit den Koordinaten  $\vec{r}_1, \dots, \vec{r}_N$  und  $\gamma$  Kerne an den Raumpunkten  $\vec{R}_1, \dots, \vec{R}_\gamma$ , lautet der Hamiltonoperator

$$\hat{\mathcal{H}} = \sum_{\alpha=1}^{\gamma} \hat{T}_{\alpha} + \sum_{i=1}^N \hat{t}_i - \sum_{i=1}^N \sum_{\alpha=1}^{\gamma} \frac{Z_{\alpha}}{|\vec{r}_i - \vec{R}_{\alpha}|} + \frac{1}{2} \sum_{\substack{i=1 \\ i \neq j}}^N \frac{1}{|\vec{r}_i - \vec{r}_j|} + \frac{1}{2} \sum_{\substack{\alpha=1 \\ \alpha \neq \beta}}^{\gamma} \frac{Z_{\alpha} Z_{\beta}}{|\vec{R}_{\alpha} - \vec{R}_{\beta}|}. \quad (2.1)$$

Der erste Term ist der Operator der kinetischen Energie der Kerne, der zweite der Kinetische-Energieoperator der Elektronen. Als nächstes haben wir die Wechselwirkungsoperatoren der Elektronen mit den Kernen und der Elektronen untereinander und schließlich noch die Wechselwirkung der Kerne untereinander. Die totale Energie dieses Systems ist durch den Erwartungswert des Hamiltonoperators gegeben,

$$\mathcal{E} = \langle \Psi | \hat{\mathcal{H}} | \Psi \rangle. \quad (2.2)$$

Eine Lösung von (2.2) ist nicht angebar, die exakte Vielteilchenwellenfunktion

$$\Psi = \Psi(\vec{R}_1, \dots, \vec{R}_{\gamma}, \vec{r}_1, \dots, \vec{r}_N) \quad (2.3)$$

nicht zugänglich. Die erste Vereinfachung ist eine Entkopplung von Kern- und Elektronenbewegung mittels der Born-Oppenheimer Näherung (BO) [4]. Die Elektronen werden von nun an im Feld der als fest im Raum fixierten Kerne betrachtet; die kinetische Energie der Kerne ist in der Kernbewegung absepariert, die wir hier nicht weiter betrachten wollen. Das führt uns auf eine neue Wellenfunktion  $\Psi_{el}$  mit rein elektronischen Charakter, in welche die Kernkoordinaten nur noch parametrisch eingehen,

$$\Psi(\vec{R}_1, \dots, \vec{R}_{\gamma}, \vec{r}_1, \dots, \vec{r}_N) \rightarrow \Psi_{el}(\vec{r}_1, \dots, \vec{r}_N; \vec{R}_1, \dots, \vec{R}_{\gamma}) \quad . \quad (2.4)$$

Mit diesem Ansatz wird aus (2.2)

$$\mathcal{E} \rightarrow E = \langle \Psi_{el} | \hat{\mathcal{H}}_{el} | \Psi_{el} \rangle \quad (2.5)$$

wobei  $\hat{\mathcal{H}}_{el}$  nun der elektronische Anteil des Operators (2.1) ist

$$\hat{\mathcal{H}}_{el} = \sum_{i=1}^N \hat{t}_i - \sum_{i=1}^N \sum_{\alpha=1}^{\gamma} \frac{Z_{\alpha}}{|\vec{r}_i - \vec{R}_{\alpha}|} + \frac{1}{2} \sum_{\substack{i=1 \\ i \neq j}}^N \frac{1}{|\vec{r}_i - \vec{r}_j|} + \frac{1}{2} \sum_{\substack{\alpha=1 \\ \alpha \neq \beta}}^{\gamma} \frac{Z_{\alpha} Z_{\beta}}{|\vec{R}_{\alpha} - \vec{R}_{\beta}|}, \quad (2.6)$$

der Ein- und Zweiteilchenwechselwirkungen enthält. In dieser Form können wir die relativistische Kinematik der Elektronen durch den Operator der kinetischen Energie

$$\hat{t}_i = \vec{\alpha}_i \cdot \vec{p}_i c + \beta_i m c^2, \quad (2.7)$$

die Elektron-Kern Wechselwirkung sowie den longitudinalen Coulombanteil der Elektron-Elektron Wechselwirkung erfassen und den elektronischen Hamiltonoperator mit dem Dirac-Coulomb (DC) Operator identifizieren:  $\hat{\mathcal{H}}_{el} = \hat{\mathcal{H}}_{DC}$ . Interelektronische Wechselwirkungen in Form von transversalen Strom-Strom Beiträgen und Retardierungen sind hierin nicht enthalten. Wollten wir dies tun, müßten wir  $\hat{\mathcal{H}}_{DC}$  noch durch den Breit-Operator [5, 6] (hier in der frequenzunabhängigen Form)

$$\hat{\mathcal{H}}_B = - \sum_{ij} \frac{e^2}{r_{ij}} \left( \vec{\alpha}_i \cdot \vec{\alpha}_j + \frac{(\vec{\alpha}_i \cdot \vec{r}_{ij})(\vec{\alpha}_j \cdot \vec{r}_{ij})}{r_{ij}^2} \right) \quad (2.8)$$

ergänzen. Hier ist  $r_{ij}$  der Abstand zwischen den beiden Elektronen  $i$  und  $j$ . Im Anhang (A.5) gehen wir näher darauf ein, daß sich der so konstruierte Dirac-Coulomb-Breit Operator zwanglos aus der Analyse der niedrigsten nichtverschwindenden <sup>2</sup> Ordnung der Störungsentwicklung der Wechselwirkung des Dirac-Feldes mit dem elektromagnetischen Strahlungsfeld ergibt. Da der Dirac-Operator nach unten unbeschränkt ist, sind die gebundenen Zustände in ein Kontinuum zu Lösungen positiver und negativer Energie eingebettet und energetisch mit diesen entartet, womit eine Minimierung zum Auffinden des Grundzustandes zu unsinnigen Ergebnissen führen würde (Brown-Ravenhall disease [7]). Zur Umgehung dieser Schwierigkeiten hat Sucher die Einführung von Operatoren vorgeschlagen, die auf die gebundenen Zustände im positiven Teil des Dirac-Spektrums projizieren [8]. Dies ist die **no-pair** Näherung. In unseren Rechnungen wird dies implizit durch die Wahl der Randbedingungen bei der Generierung der atomaren Einteilchenorbitale gewährleistet, sodaß Beiträge des negativen Kontinuums auch für chemische Abstände im wesentlichen ausgeschlossen sind.

---

<sup>2</sup>Aus Gründen der Energieerhaltung ist dies die zweite Ordnung, siehe Anmerkungen in [102], Seite 110.

## 2.2 Der Dichtefunktional–Ansatz

Einen anderen Zugang zur Lösung des quantenmechanischen Vielteilchenproblems bietet die Dichtefunktionaltheorie (DFT). Wegen ihrer drastischen Aufwandreduzierung stellt sie eine brauchbare und recht elegante Alternative zu den 'klassischen' Methoden der Behandlung von Vielteilchensystemen, wie etwa Dirac–Fock Verfahren, dar. Nicht nur bei Atomen, sondern auch im molekularen Bereich, in der Festkörperphysik und zur Beschreibung von Phänomenen in der Oberflächenphysik hat die DFT breite Anwendung gefunden und ist dort zum unentbehrlichen Werkzeug geworden [9, 10, 11, 12, 13]. Man muß sich jedoch darüber im klaren sein, daß es das Dichtefunktional schlechthin (noch) nicht gibt. Bestand die ursprüngliche Intention der Dichtefunktionaltheorie darin, einen allgemeingültigen Ausdruck eines Funktionals abzuleiten, welches universell für alle Systeme anwendbar ist und alle physikalischen Effekte richtig beschreibt, so ist man wegen der Schwierigkeiten bei der Umsetzung dieses Vorhabens von diesem hehren Anspruch abgekommen. Dennoch ist die Theorie weiterhin bemüht, verbesserte Funktionale bereitzustellen.

Die Dichtefunktionaltheorie ist eine 'first principle' Theorie. Ausgangspunkt ist das Hohenberg–Kohn Theorem [14]. Dieses zeigt, daß zu jedem externen Potential  $v^{ext}(\vec{r})$  **genau eine** zu diesem Potential gehörende Grundzustandsdichte  $n(\vec{r})$  korrespondiert, aus welcher der energetische Grundzustand und im Prinzip weitere Observablen berechnet werden können. Dies gilt auch für wechselwirkende Vielteilchensysteme. Dementsprechend läßt sich die Energie als ein Funktional  $\mathcal{E}[n]$  der Dichte  $n$  auffassen:

$$\mathcal{E}[n] = \int n(\vec{r})v^{ext}(\vec{r})d\vec{r} + \mathcal{F}[n] \quad . \quad (2.9)$$

Desweiteren gilt für  $\mathcal{E}[n]$  ein Variationsprinzip:  $\delta_n \mathcal{E}[n] \doteq 0$ , dessen Lösung diejenige Dichte  $\bar{n}(\vec{r})$  liefert, welche die zu dem gegebenen externen Potential  $v^{ext}(\vec{r})$  gehörende Grundzustandsdichte ist. Dies ist eine weitere Aussage des Hohenberg–Kohn Theorems. Die Dichte wird dann über Hilfs-Einteilchenorbitale  $\phi_i$  ausgedrückt:

$$n(\vec{r}) = \sum_i^N \phi_i^*(\vec{r})\phi_i(\vec{r}) \quad . \quad (2.10)$$

Diese Darstellung der Dichte erlaubt nun eine entsprechende Aufteilung von  $\mathcal{F}[n]$ :

$$\mathcal{F}[n] = \frac{1}{2} \int \int \frac{n(\vec{r})n(\vec{r}')}{|\vec{r} - \vec{r}'|} d\vec{r} d\vec{r}' + \mathcal{T}_s[n] + \mathcal{E}^{xc}[n] \quad . \quad (2.11)$$

Der erste Term ist das Hartree-Energiefunktional  $\mathcal{E}^h[n]$ .  $\mathcal{T}_s[n]$  stellt das Funktional der kinetischen Energie **nichtwechselwirkendener**, zu  $\phi_i(\vec{r})$  korrespondierenden Teilchen dar. Dieses ist definiert als

$$\mathcal{T}_s[n] = \sum_i^N \int \phi_i^*(\vec{r}) \left( -\frac{\nabla_i^2}{2m} \right) \phi_i(\vec{r}) \quad . \quad (2.12)$$

Der verbleibende Energiebeitrag  $\mathcal{E}^{xc}[n]$  in (2.11) wird als Austausch-Korrelationsenergie bezeichnet. Die Ausführung der Variation bezüglich der  $\phi_i$  anstelle der ursprünglichen Dichte  $n$  führt schließlich auf die Kohn–Sham (KS) Gleichungen [15]:

$$\left( -\frac{1}{2m} \nabla^2 + v^{KS}(\vec{r}) \right) \phi_i(\vec{r}) = \epsilon_i \phi_i(\vec{r}) \quad . \quad (2.13)$$

Hier ist  $v^{KS}$  ein effektives Einteilchenpotential, in dem sich die 'Kohn-Sham' Teilchen bewegen. Es setzt sich zusammen aus dem externen Potential  $v^{ext}(\vec{r})$ , dem Hartreepotential  $v^h(\vec{r})$  und dem Austausch-Korrelationspotential  $v^{xc}(\vec{r})$ . Dieses ist definiert als Funktionalableitung der Austausch-Korrelationsenergie  $E^{xc}$  bezüglich der Dichte

$$v^{xc}(\vec{r}) := \frac{\delta \mathcal{E}^{xc}[n]}{\delta n(\vec{r})} \quad . \quad (2.14)$$

Trotz der Existenz – und dieser Beweis ist mehrfach niedergeschrieben (siehe z.B. [16, 17]) – ist die explizite Gestalt von  $\mathcal{E}^{xc}[n]$  nicht bekannt. Die Schwierigkeit besteht darin, für das Austausch-Korrelationsenergiefunktional einen berechenbaren Ausdruck herzuleiten. In einer systematischen Weise kann dies zunächst für Austausch und Korrelation getrennt angegangen werden [18, 19].

Das Energiefunktional schreiben wir nun als eine Summe der Energiekomponenten:

$$\mathcal{E}[n] = \mathcal{T}_s[n] + \mathcal{E}^{ext}[n] + \mathcal{E}^h[n] + \mathcal{E}^{xc}[n] + E_{\alpha\beta} \quad . \quad (2.15)$$

Hier ist  $E_{\alpha\beta}$  der (triviale) Beitrag der Kern-Kern Wechselwirkung.

In diesem Abschnitt haben wir die Idee der DFT zu motivieren versucht. Zu diesem Zweck haben wir uns zunächst auf den nichtrelativistischen Fall beschränkt. Jetzt wollen wir das DFT Konzept auf den relativistischen Fall erweitern.

## 2.3 Relativistische Dichtefunktionaltheorie

Die erste Formulierung der relativistischen Dichtefunktionaltheorie (RDFT) geht auf Rajagopal und Callaway [20] zurück. Sie haben gezeigt, daß das Hohenberg-Kohn Theorem auch im relativistischen Regime gilt. Wie auch im nichtrelativistischen Falle läßt sich auf dieses Funktional ein Variationsprinzip anwenden. Die wesentlichsten Unterschiede zur nichtrelativistischen DFT sind:

- die Viererstromdichte  $j^\nu$  als grundlegende Variable, die das Funktional bestimmt und
- die relativistischen Kohn–Sham Gleichungen (RKS)

Die folgenden Betrachtungen wollen wir innerhalb der im ersten Abschnitt erwähnten **no-pair** Näherung verstanden wissen. Dementsprechend sollen in den Gleichungen auch alle weiteren quantenelektrodynamischen Prozesse höherer Ordnung wie Vakuumpolarisation und Strahlungskorrekturen vernachlässigt sein. Die Stromdichte ist dann gegeben durch

$$j^\nu(\vec{r}) = \sum_{-mc^2 < \epsilon_k \leq \epsilon_F} \bar{\psi}_k(\vec{r}) \gamma^\nu \psi_k(\vec{r}) \quad . \quad (2.16)$$

und die  $\epsilon_k$  entsprechen den Einteilchenenergien der besetzten Zustände zwischen  $-mc^2 \leq \epsilon_k \leq \epsilon_F$  im Energiespektrum der Diracgleichung. Aus den Ein-Teilchen-Spinoren definieren wir zunächst wieder das Funktional der kinetischen Energie  $\mathcal{T}_s$ :

$$\mathcal{T}_s[j^\nu] = \sum_{-mc^2 < \epsilon_k \leq \epsilon_F} \int \psi_k^\dagger(\vec{r}) [c\vec{\alpha} \cdot \vec{p} + \beta m c^2] \psi_k(\vec{r}) d\vec{r} \quad . \quad (2.17)$$

Im Sinne des ersten Punktes geht Gleichung (2.15) dann über in die Form

$$\mathcal{E}[j^\nu] = \mathcal{T}_s[j^\nu] + \mathcal{E}^{ext}[j^\nu] + \mathcal{E}^h[j^\nu] + \mathcal{E}^{xc}[j^\nu] + E_{\alpha\beta} \quad . \quad (2.18)$$

In der Feldtheorie wird nun die Wechselwirkung von Elektronen (Fermionen) durch den Austausch virtueller Photonen beschrieben. So läßt sich das Funktional der direkten Elektron-Elektron Wechselwirkung kovariant durch

$$\mathcal{E}^h = -\frac{1}{2} \int \int d^3x d^4y j^\nu(\vec{x}) D_{\nu\mu}(x-y) j^\mu(\vec{y}) \quad (2.19)$$

formulieren, wobei hier der Photonenpropagator  $D_{\nu\mu}$  (siehe Anhang, A.33) den Austausch eines Photons zwischen den beiden Strömen vermittelt. Die Tensorstruktur von  $D_{\nu\mu}$  repräsentiert vier mögliche im Rahmen der kovarianten Quantisierung auftretende Polarisationszustände des Photonenfeldes: jeweils einen longitudinalen, zwei transversale und eine skalare Komponente<sup>3</sup>. Daher läßt er sich in einen longitudinalen (L) und transversalen (T) Anteil zerlegen:

---

<sup>3</sup>Hier muß zwischen kovarianter Quantisierung und der Zerlegung des Photonenpropagators unterschieden werden. Ausführlicher wird darauf in [102], Kapitel 5.3, Seite 92 ff. oder in [103], Kapitel 14.6, Seite 85 ff. eingegangen, wo auch die explizite Gestalt des Photonenpropagators angegeben ist.

$$D_{\mu\nu} = D_{\mu\nu}^L + D_{\mu\nu}^T \quad .$$

Aus dieser Aufteilung resultieren die longitudinalen und transversalen Beiträge der interelektronischen Wechselwirkung. Für (2.19) ergibt sich

$$\mathcal{E}^h[j^\nu] = \mathcal{E}_L^h[j^\nu] + \mathcal{E}_T^h[j^\nu] \quad . \quad (2.20)$$

Da wir es mit fermionischen Systemen zu tun haben, muß eine entsprechende Analyse auch für das Austausch–Korrelationsenergiefunktional gemacht werden, welches sich in analoger Weise wie der direkte Term in einen longitudinalen und transversalen Anteil zerlegen läßt:

$$\mathcal{E}^{xc}[j^\nu] = \mathcal{E}_L^{xc}[j^\nu] + \mathcal{E}_T^{xc}[j^\nu] \quad . \quad (2.21)$$

Hier wurde das Korrelationsenergiefunktional wieder mit hinzugefügt. Die Grundzustandsenergie ergibt sich gemäß dem Hohenberg-Kohn Theorem aus der Variation des Energiefunktionals  $\mathcal{E}[j^\nu]$  und führt auf die relativistischen Kohn-Sham Gleichungen, wie sie beispielsweise von MacDonald und Vosko abgeleitet wurden [21]. Diese lauten in relativistischer Notation (vgl. Anhang, A.4)

$$\gamma^0 \left( -ic\vec{\gamma} \cdot \vec{\nabla} + mc^2 + eV(\vec{r}) + \not{p}_h(\vec{r}) + \not{p}_{xc}(\vec{r}) \right) \psi_k(\vec{r}) = \epsilon_k \psi_k(\vec{r}) \quad . \quad (2.22)$$

Hier stellt  $V^\nu(\vec{r})$  ein externes Potential dar. Die weiteren Terme sind das direkte Hartreepotential und das Austausch-Korrelationspotential:

$$\begin{aligned} v_h^\nu(\vec{r}) &= \int \frac{j^\nu(\vec{r}')}{|\vec{r} - \vec{r}'|} d\vec{r}' \quad , \\ v_{xc}^\nu(\vec{r}) &= \frac{\delta E^{xc}[j^\nu]}{\delta j_\nu} \quad . \end{aligned} \quad (2.23)$$

In Gleichung (2.22) besitzen die Potentiale noch Viererstruktur, eine volle selbstkonsistente Lösung zusammen mit Gleichung (A.28) ist auch auf dem *no-pair* Niveau sehr kompliziert. Betrachten wir die Elektronen im statischen Feld der Atomkerne und nehmen keine weiteren externen Magnetfelder an, so haben wir  $V_\nu = (eV_0, \vec{0}) = (v^{ext}, \vec{0})$ . Dann koppelt der Viererstrom nur an die Nullkomponente des Kernvektorpotentials

$$\sum_k \bar{\psi}_k \gamma^\nu \psi_k V_\nu = \begin{cases} = j^0 v^{ext}(\vec{r}); & \nu = 0 \\ = 0; & \nu = 1, 2, 3 \end{cases} \quad (2.24)$$

und die Grundzustandsenergie ist nun ein Funktional von  $j^0(\vec{r}) = n(\vec{r})$ . Die Stromdichte

$$\vec{j}[n] = \sum_k \psi_k^\dagger \vec{\alpha} \psi_k \quad (2.25)$$

kann im Prinzip aus  $n$  berechnet werden. Dies wird auch als **elektrostatischer Limes** bezeichnet. Alle Energiekomponenten des Funktionals  $\mathcal{E}$  lassen sich dann über  $n$  ausdrücken:

$$\mathcal{E}[j^\nu] = \mathcal{E}[n, \vec{j}[n]] \quad . \quad (2.26)$$

## 2.4 Die relativistischen Kohn–Sham Gleichungen

In diesem Abschnitt wollen wir das Lösungsverfahren für die relativistischen Kohn–Sham Gleichungen (RKS) angeben. Diese lauten in dem oben diskutierten elektrostatischen Limes

$$\left( c\vec{\alpha} \cdot \vec{p} + \beta mc^2 + v^{RKS}(\vec{r}) \right) \psi_k(\vec{r}) = \epsilon_k \psi_k(\vec{r}) \quad . \quad (2.27)$$

Das Potential setzt sich aus folgenden Anteilen zusammen:

$$v^{RKS}(\vec{r}) = v^{ext}(\vec{r}) + v_L^h(\vec{r}) + v_T^h(\vec{r}) + v_L^{xc}(\vec{r}) + v_T^{xc}(\vec{r}) \quad , \quad (2.28)$$

wobei wir in den praktischen Rechnungen den transversalen Anteil des direkten Hartree-Potentials  $v_T^h(\vec{r})$  vernachlässigen. In einigen Fällen, wenn z.B. Systeme mit abgeschlossenen Schalen vorliegen, verschwindet diese direkte transversale Strom-Strom Kopplung ohnehin. Dies impliziert nicht, daß damit auch die transversale Austauschenergie verschwindet! Die RKS-Gleichungen werden nun mit unserem Programm mit einer Basisentwicklungsmethode (LCAO - Linear Combination of Atomic Orbitals) gelöst. Die molekularen Wellenfunktionen  $\psi_i^{MO}$  werden zuerst nach Symmetrieorbitalen  $\Upsilon^{SO}$  und diese wiederum in relativistische atomare Dirac-Spinoren  $\xi^{AO}$  entwickelt,

$$| \psi_i^{MO} \rangle = \sum_m b_{im} | \Upsilon_m^{SO} \rangle = \sum_m \sum_\mu b_{im} c_{m\mu} | \xi_\mu^{AO} \rangle \quad . \quad (2.29)$$

Die atomaren Wellenfunktionen sind gegeben durch

$$\xi_\mu^{n\kappa m}(\vec{r}) = \frac{1}{r} \begin{pmatrix} f_\mu^{n\kappa}(r) \mathcal{Y}_\kappa^m(\Omega) \\ i g_\mu^{n\kappa}(r) \mathcal{Y}_{-\kappa}^m(\Omega) \end{pmatrix} \quad . \quad (2.30)$$

Für das  $\mu$ -te Atomorbital bedeuten hier die Funktionen  $f(r)$  und  $g(r)$  die Radialanteile der großen bzw. kleinen Komponente mit der Hauptquantenzahl  $n$ .  $\mathcal{Y}_m^\kappa$  sind die spinoriellen Kugelflächenfunktionen,  $\kappa$  ist die Dirac-Quantenzahl, für die gilt

$$\kappa = \begin{cases} -(j + 1/2) & = -(l + 1) & \text{für } j = l + 1/2 \\ j + 1/2 & = l & \text{für } j = l - 1/2 \end{cases} . \quad (2.31)$$

Setzen wir die Entwicklung (2.29) in die Kohn-Sham Gleichungen ein, multiplizieren mit  $\langle \psi_k |$  und summieren über alle besetzten Zustände, bekommen wir

$$\sum_k \sum_\mu c_{k\mu}^* c_{k\nu} \underbrace{\langle \xi_\nu | (c\vec{\alpha} \cdot \vec{p} + \beta mc^2 + v^{RKS}(\vec{r})) | \xi_\mu \rangle}_{h_{\nu\mu}} = \sum_k \sum_{\nu\mu} c_{k\mu}^* c_{k\nu} \epsilon_k \underbrace{\langle \xi_\nu | \xi_\mu \rangle}_{S_{\nu\mu}} . \quad (2.32)$$

Das sind die Gleichungen eines verallgemeinerten Matrixeigenwertproblems

$$\mathbf{HC} = \epsilon \mathbf{SC} \quad , \quad (2.33)$$

aus dessen Lösung die Eigenwerte  $\epsilon_k$  und die Eigenvektoren  $\mathbf{C}$  gewonnen werden können.  $\mathbf{H}$  ist die Fock- und  $\mathbf{S}$  die Überlappmatrix. Die Dimension und damit der Aufwand der Diagonalisierung von (2.33) wird hauptsächlich von der Anzahl der verwendeten atomaren Basisfunktionen bestimmt. Durch den Zwischenschritt der Entwicklung der molekularen Orbitale in eine Symmetrieorbitalbasis zerfallen Fock- und Überlappmatrix in kleinere Blöcke entsprechend den irreduziblen Darstellungen der molekularen Punktgruppe. Da die Elemente zu disjunkten irreduziblen Darstellungen orthogonal zueinander sind, können beide Matrizen blockweise diagonalisiert werden. Das bringt im allgemeinen eine erhebliche Verringerung des Rechenaufwandes mit sich.

Zur Berechnung der Matrixelemente werden die auftretenden Integrale in eine gewichtete Summe einer diskreten Punktverteilung überführt:

$$\int f(x) dx \longrightarrow \sum_{i=1}^{N_p} f_i(x_i) w_i \quad . \quad (2.34)$$

Hier steht  $x_i$  für die Koordinaten des  $i$ -ten Integrationspunktes,  $f_i$  ist der Wert des Integranden am Punkt  $x_i$  und  $w_i$  ist das an diesem Punkt gegebene Integrationsgewicht. Die Summe läuft über alle  $N_p$  Punkte. Die Integrationspunkte und -gewichte werden mit der Multizentren-Integrationsmethode von Te Velde und Baerends [22] erzeugt. Das Funktional der totalen Energie ergibt sich dann zu

$$\begin{aligned}
\mathcal{E} &= \sum_k^N \langle \psi_k | \hat{t} + v^{ext}(\vec{r}) + v^h(\vec{r}) + v^{xc}(\vec{r}) | \psi_k \rangle \\
&- \frac{1}{2} \int n(\vec{r}) v^h(\vec{r}) d\vec{r} - \int n(\vec{r}) v^{xc}(\vec{r}) d\vec{r} + E^{xc} + E_{\alpha\beta} \\
&= \sum_k^n \langle \psi_k | \hat{t} | \psi_k \rangle + \int n(\vec{r}) v^{ext}(\vec{r}) d\vec{r} + \frac{1}{2} \int n(\vec{r}) v^h(\vec{r}) d\vec{r} + E^{xc} + E_{\alpha\beta} \\
&= \mathcal{T}_s[n] + \mathcal{E}^{ext}[n] + \mathcal{E}_L^h[n] + \mathcal{E}^{xc}[n] + E_{\alpha\beta} \quad . \tag{2.35}
\end{aligned}$$

# Kapitel 3

## Dichtefunktionalanwendungen

### 3.1 Einleitung und Motivation

Das in unseren Rechnungen bisher verwendete Slater'sche Austauschfunktional gehört wegen seiner Ableitung aus dem homogenen Elektronengas zur Familie der lokalen Dichtefunktionale. Es liefert trotz seiner relativ groben Berücksichtigung von Austauschbeiträgen brauchbare Ergebnisse in ganz unterschiedlichen Bereichen und kann die physikalischen Effekte selbst im Relativistischen qualitativ meist richtig wiedergeben. Zwei weit voneinander entfernt liegende Beispiele sind die Potentialenergiekurven der schweren zweiatomigen Moleküle  $Tl_2$ ,  $Pb_2$  und  $Bi_2$  [23] und das aus dem leichten Element Kohlenstoff bestehende Fulleren  $C_{60}$ , für das in Abhängigkeit des Radius verschiedene Ionisationsstufen und der maximale Ionisationsgrad berechnet wurden [24]. Als genereller Trend werden jedoch Bindungsenergien oftmals zu tief erhalten. Wie man weiß, liegt dies in der physikalischen Näherung des Funktionals begründet. Weitere Beispiele für das Versagen lokaler Dichtefunktionale sind die falsche Vorhersage des magnetischen Moments und verschiedener Grundzustandseigenschaften von Eisen [25, 26] oder des  $H^-$ -Ions [27]. Erst durch die Einführung von Dichtegradienten und das Erreichen des richtigen asymptotischen Verhaltens des Austausch-Korrelationspotentials können die Defizite der lokalen Formulierungen behoben werden.

An dieser Stelle besteht nun die Hoffnung, durch Verwendung gradientenkorrigerter Dichtefunktionale unsere bisherigen Ergebnisse zu verbessern. Die Bedeutung von Gradientenkorrekturen ist unumstritten, deren Einbeziehung im DFT-Formalismus ist in der Literatur sehr ausführlich dokumentiert [28, 29, 30]. In den nächsten Kapiteln soll der Formalismus für die Anwendung der DFT für die für uns relevanten Fragestellungen dargelegt werden. Wir wollen dabei auf Details verzichten.

In ihren Anfängen wurde die DFT zunächst nichtrelativistisch formuliert und umgesetzt. Dadurch ist aber die Erfassung relativistischer Effekte, die mit Zunahme der Kernladungszahl immer größer werden, nur im Sinne der Störungstheorie möglich. So geht mit der relativistischen Massenzunahme eine Verkürzung der Längen und Radien einher, infolgedessen  $s$ - und  $p$  Wellenfunktionen näher am Kern lokalisiert sind. Das ist der direkte relativistische Effekt. Da die äußeren  $d$ - und  $f$  Elektronen nunmehr ein abgeschirmtes

positives Kernpotential verspüren, wird deren Bindung abgeschwächt und sie erfahren eine Expansion zu größeren Abständen hin. Dies ist der indirekte relativistische Effekt. Der direkte relativistische Effekt tritt nun am deutlichsten beim Goldatom zutage. Der radiale Erwartungswert  $\langle r \rangle$  des  $6s$  Orbitals erfährt mit über 15 Prozent die stärkste Kontraktion im Vergleich zu allen anderen leichteren Elementen. Diese besondere Eigenschaft für Gold wird auch als 'Goldmaximum' [31] bezeichnet. Zum Studium der relativistischen Effekte eignen sich somit Goldverbindungen als ideale Kandidaten. Eine ausführliche Darstellung dieser Effekte im Hinblick auf physiko-chemische Fragestellungen hat Pyykkö [32] gegeben.

Die Dirac–Fock–Slater Methode, die wir bisher benutzten, berücksichtigt zwar die relativistische Kinematik der Elektronen, relativistische Beiträge zum Dichtefunktional sind jedoch vernachlässigt bzw. dieser Methode fremd. In den letzten Jahren wurden nun auch relativistische Austausch–Korrelationsfunktionale entwickelt und auf atomarer Ebene [33, 2] sowie in skalar-relativistischen molekularen Rechnungen eingesetzt [34, 35, 36]. Entsprechende Anwendungen für den molekularen Fall unter Berücksichtigung der Relativistik sowohl bezüglich der Kinematik der Elektronen als auch der Photonen stehen aber bis heute noch aus. Ziel dieses Teils der Arbeit nun ist

- die Untersuchung der Bindungseigenschaften von Molekülen aufgrund der Einbeziehung gradientenkorrigerter Funktionale und
- die entwickelten relativistischen Funktionale mit unserem vierkomponentigen Formalismus zu kombinieren und an superschweren Verbindungen zu studieren .

Für die folgenden Betrachtungen werden wir das Hauptaugenmerk auf die Austauschfunktionale legen. Korrelationsfunktionale, obwohl an deren Entwicklung mit nicht minderer Intensität gearbeitet wird [37], wollen wir hier nicht in dieser Ausführlichkeit diskutieren. Gelegentlich werden wir aber auf diese verweisen. Wir wollen noch anmerken, daß man im allgemeinen von der '*exchange-only DFT*' spricht, wenn Korrelationsbeiträge unberücksichtigt gelassen werden.

## 3.2 Die Lokaldichte-Näherung

Das der DFT zugrunde liegende einfachste Modell ist das des homogenen Elektronengases (HEG). In diesem Modell ist die Dichte im gesamten Raumgebiet eine konstante Funktion  $n(\vec{r}) = n_0$ . Für das HEG hat Slater einen exakten Ausdruck für die Austauschenergie  $E^x$  abgeleitet [38] und gezeigt, daß sich die nichtlokale Austauschwechselwirkung lokal formulieren läßt. Diese Näherung für das Elektronengas mit dem daraus resultierenden lokalen Charakter des Funktionals bezeichnet man als **Local Density Approximation - LDA**. Die Austauschenergie ist in diesem Modell gegeben durch

$$E_x^{LDA}[n] = \int e_x^{LDA}(n(\vec{r})) d\vec{r} \quad (3.1)$$

mit der Energiedichte

$$e_x^{LDA}(n) = -\frac{9}{8\pi} X_\alpha (3\pi^2)^{\frac{1}{3}} n^{\frac{4}{3}} \quad . \quad (3.2)$$

Das Austauschpotential folgt aus der Differentiation der letzten Gleichung bezüglich  $n$ :

$$v_x^{LDA}(\vec{r}) = -\frac{3}{2\pi} X_\alpha (3\pi^2)^{\frac{1}{3}} n(\vec{r})^{\frac{1}{3}} \quad . \quad (3.3)$$

Der Parameter  $X_\alpha$  beträgt für die exakte Ableitung  $\frac{2}{3}$ . Die beste Übereinstimmung mit atomaren Hartree-Fock Rechnungen wird dann erzielt, wenn  $X_\alpha = 0.7$  gesetzt wird [39]. Diesen Befund haben wir übernommen und daher unsere bisherigen molekularen Rechnungen mit  $X_\alpha = 0.7$  durchgeführt. Am Ende des Abschnitts 2.2 hatten wir darauf verwiesen, daß der Korrelationsenergiebeitrag mitberücksichtigt werden muß. Eine Möglichkeit, die Korrelation in der LDA zu berücksichtigen, wurde von Vosko, Wilk und Nusair [40] vorgeschlagen. Korrelationsenergien auf dem LDA-Niveau werden etwa doppelt so groß erhalten wie in den nichtlokalen Formulierungen, denen wir uns im folgenden Abschnitt zuwenden wollen.

### 3.3 Nichtlokale Dichtefunktionale

Die Berücksichtigung von Inhomogenitäten kann in der DFT im einfachsten Fall durch eine Entwicklung nach Gradienten der Dichte geschehen [41]. Eine zusammenfassende Darstellung der Herleitung gradientenkorrigerter Funktionale hat Salahub [42] gegeben. Ergebnisse von deren Anwendung auf Atome, Moleküle und Festkörper sind z.B. in [43] zu finden. Zur Familie der nichtlokalen Dichtefunktionale gehören die Gradientenentwicklung (GE) und die generalized gradient approximation (GGA). Gradientenkorrigierte Funktionale erfassen den nichtlokalen Charakter nur bis zu einer gewissen Ordnung in der Störungsentwicklung der Dichte (so gut man diese Entwicklung eben kennt). Die GGA Funktionale können als eine partielle Resummation dieser Gradientenentwicklung betrachtet werden.

Bereits Kohn und Sham [15] haben Gradientenentwicklungen für den Fall einer schwach veränderlichen Dichte vorgeschlagen und diese als Störung des homogenen Elektronengases (HEG) betrachtet. Die Energiedichte  $e_{xc}$  läßt sich dann als eine Funktion von Dichte- und Dichtegradiententermen auffassen:  $e_{xc} = e_{xc}(n, (\nabla n)^2, \nabla^2 n, \dots)$ . Aus Symmetriebetrachtungen folgt, daß Terme mit ungeraden Potenzen von  $\nabla n$  in einer Reihentwicklung der Energiedichte nicht auftreten<sup>1</sup>, allenfalls deren Beträge. Im Sinne dieser Entwicklung ist die Austauschenergie dann gegeben durch

---

<sup>1</sup>In [15] wird argumentiert, daß die Energiedichte als ein universelles Funktional der Dichte eine skalare Größe ist, demzufolge sie beispielsweise unter Drehungen und Inversionen invariant bleiben muß.

$$E_{xc}[n] = A[n(\vec{r})] + B[n(\vec{r}), (\nabla n)^2] + \dots + [\mathcal{O}(\nabla^2 n)] \quad (3.4)$$

Hier sind  $A$  und  $B$  Funktionen von  $n$  und  $\nabla n$ -Termen. Herman et al. [44] haben gezeigt, daß im Falle der 'exchange only' DFT diese Austauschenergie folgende Gradientenentwicklung hat:

$$E_x[n] = E_x^{LDA}(n) - b \int \frac{(\nabla n)^2}{n^{\frac{4}{3}}} d\vec{r}, \quad (3.5)$$

wobei  $b$  eine Konstante und  $E_x^{LDA}$  die Austauschenergie (vgl. 3.2) des HEG ist. Die so gegebene Energie und das daraus resultierende Austauschpotential zeigen nur im Falle kleiner Dichtegradienten gutartiges Verhalten; für exponentiell verschwindende Dichten sind sie divergent [45]. Alternativ dazu hat Becke einen Ausdruck gefunden, der die Austauschenergie im Hochdichtelimes korrekt wiedergibt [46]. Die Kombination mit (3.2) liefert

$$E_x[n] = E_x^{LDA}(n) - \beta \int n^{\frac{4}{3}} \frac{x^2}{(1 + \gamma x^2)} d\vec{r} \quad (3.6)$$

mit der charakteristischen Größe

$$x = \frac{|\nabla n|}{n^{\frac{4}{3}}} \quad (3.7)$$

und den Parametern  $\beta$  und  $\gamma$ , die aus Least Square Fits an atomare Daten erhalten sind. 1988 hat Becke einen anderen Ausdruck für die Austauschenergie gefunden [47]. Dieser hat eine ähnliche Struktur wie (3.6):

$$E_x[n] = E_x^{LDA}[n] - \beta \int n^{\frac{4}{3}} \frac{x^2}{(1 + 6\beta x \sinh^{-1} x)} d\vec{r} \quad (3.8)$$

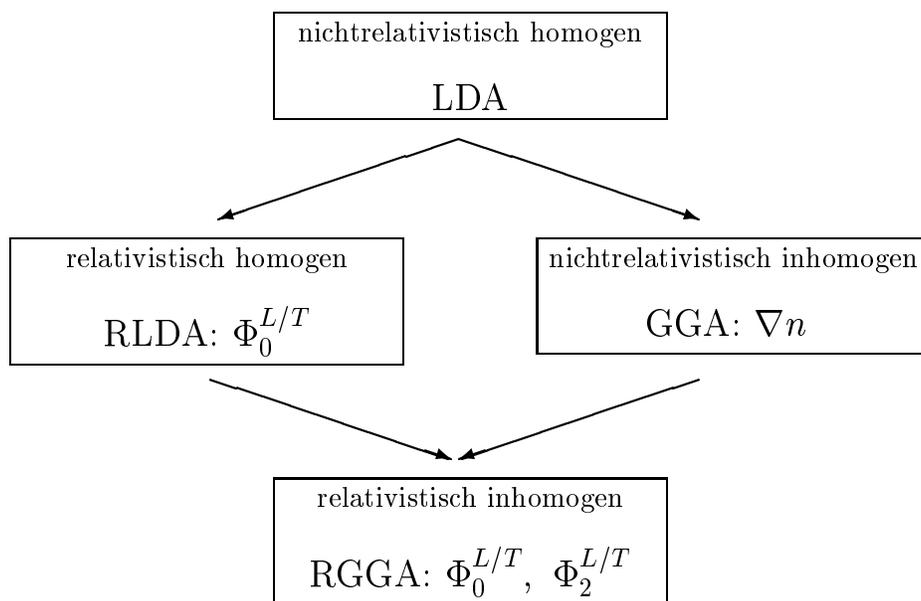
Im Gegensatz zur früheren Versionen des Austauschfunktionals hat dieser Ausdruck nun nur noch einen Parameter  $\beta$ . Er ist an exakte atomare Hartree-Fock Austauschenergien gefittet und sein Wert beträgt  $\beta = 0.0042$ . Weitere Austauschfunktionale wurden von Engel und Vosko [48] und Perdew und Wang [49] vorgeschlagen. Als Nomenklatur hat sich die Angabe des Autors mit der Jahreszahl der Entwicklung des Funktionals eingebürgert. Die gebräuchlichsten Austauschfunktionale sind die von Becke (B88) und Perdew/Wang (PW86 [50, 51] und PW91 [49]). Seitens der Korrelation sind die Funktionale von Perdew (P86 [37]), Perdew und Wang (PW91 [49]) und Lee, Yang und Parr (LYP [52]) zu nennen. Eine ausführliche Studie der Kombination von Austausch- und

Korrelationsfunktionalen durch wahlweises Zuschalten von Gradientenkorrekturen im Vergleich mit quantenchemischen Hartree-Fock- und Møller-Plesset- bzw. CI Verfahren ist von Johnson et al. [53] gemacht worden.

Eine wichtige physikalische Eigenschaft der Austauschenergie ist die Aufhebung der Selbstwechselwirkung in der Hartree-Energie, wie sie z.B. im Hartree-Fock Modell berücksichtigt ist. In der DFT ist dies aufgrund der jeweils gemachten Näherungen aber nicht immer erfüllt. Mit der sogenannten self interaction correction (SIC) wird versucht, die exakte Aufhebung von Coulomb- und Austauschwechselwirkung im Nachhinein wieder in den DFT-Formalismus einzubringen [54].

### 3.4 Die relativistischen Energiefunktionale

In diesem Abschnitt geben wir die expliziten Ausdrücke für die relativistischen Funktionale an (vergleiche auch [2]). Im Vordergrund soll hierbei die Austauschenergie stehen. Grundlage hierfür bildet wieder das homogene nichtwechselwirkende, aber jetzt relativistische Elektronengas der Dichte  $n_0$ , die durch die lokale Dichte  $n(\vec{r})$  am Orte  $\vec{r}$  ersetzt wird. Geht man wieder von der nichtrelativistischen Lokaldichte-Näherung des freien Elektronengases aus, bedarf es einer Modifikation dieses Funktionals, um a) die Relativistik und b) Gradientenkorrekturen zu berücksichtigen. Ansätze zur Beschreibung des relativistischen inhomogenen Elektronengases gehen auf Rajagopal zurück [55]. In der folgenden kleinen Übersicht ist der Weg von der LDA zur relativistischen GGA (RGGA) dargestellt.



**Abbildung 3.1:** Der Weg zu den relativistischen gradientenkorrigenen Energiefunktionalen mit den dafür benötigten Erweiterungen in Abhängigkeit des jeweiligen Elektronengasmodells.

Longitudinale (L) und transversale (T) Austauschenergien lassen sich in systematischer Weise aus der nichtrelativistischen Lokaldichte-Näherung durch eine Modifikation der

nichtrelativistischen LDA-Energiedichte  $e_x^{LDA}(n)$  mit Korrekturfunktionen  $\Phi_{0/2}^{L/T}(\beta)$  von komplizierter Gestalt (siehe Anhang, A.6) gewinnen. Diese Korrekturfunktionen hängen von der charakteristischen relativistischen Variablen  $\beta$  ab, welche den Fermimpuls  $k_f = (3\pi^2 n)^{\frac{1}{3}}$  der Elektronen in Beziehung zum relativistischen Impuls setzt:

$$\beta = \frac{(3\pi^2 n)^{\frac{1}{3}}}{mc} = \frac{k_f}{mc} . \quad (3.9)$$

Für die relativistische 'exchange-only' LDA-Energie bekommt man für longitudinale und transversale Austauschenergien

$$E_x^{L/T,RLDA}[n] = \int e_x^{LDA}(n) \Phi_0^{L/T}(\beta) d\vec{r} . \quad (3.10)$$

Hier ist  $e_x^{LDA}(n)$  die Energiedichte des homogenen nichtrelativistischen Elektronengases. Die Gradientenkorrekturen (GGA's) werden durch einen erweiterten Ansatz über eine Funktion  $g(\xi)$  berücksichtigt, wobei in  $\xi$  der Dichtegradient  $\nabla n$  in der Form

$$\xi = \left( \frac{\nabla n}{2k_f n} \right)^2 \quad (3.11)$$

einght. Zur Berücksichtigung der Relativistik ist die GGA um eine weitere Korrekturfunktion  $\Phi_2$  zu ergänzen. Die relativistische gradientenkorrigierte Austauschenergie lautet dann:

$$E_x^{L/T,RGGA}[n] = \int e_x^{LDA}(n) \left[ \Phi_{x,0}^{L/T}(\beta) + g(\xi) \Phi_{x,2}^{L/T}(\beta) \right] d\vec{r} . \quad (3.12)$$

Auch sind entsprechende Ausdrücke für relativistische Korrelationsenergien formuliert worden [56], deren Weiterentwicklung derzeit Gegenstand aktueller Untersuchungen ist. Wir wollen hier aber nicht weiter darauf eingehen.

Damit sind uns im Prinzip alle Energiebeiträge der totalen Energie (2.35) bekannt. Bevor wir diese aber explizit berechnen können, werden wir im kommenden Abschnitt noch auf eine von uns aus rechentechnischen Gründen gemachte Näherung eingehen, die auch weitreichende Konsequenzen für den Energiegradienten des Kapitels 4 hat.

Im Vorgriff auf die Rechnungen zur Geometrieoptimierung mit dem sich dort abzeichnenden ungleich höheren Rechenaufwand, was die Gradienten angeht, sei hier noch erwähnt, daß sich alle Dichtefunktionalrechnungen auch mit Gradientenkorrekturen unter voller Ausnutzung einer eventuellen Molekülsymmetrie durchführen lassen. Das liegt daran, daß in den Energieausdruck (3.12) nicht der Gradient der Dichte, sondern über (3.11) dessen Betrag einght. Dies ist anders, wenn generell vorzeichenbehaftete Größen in den Rechnungen auftauchen. Das können neben Gradienten beispielsweise auch Dipolmomente sein.

### 3.5 Eine wichtige Approximation – Fit der Ladungsdichte

Die größte Schwierigkeit in den praktischen Rechnungen stellt die Behandlung der Elektron–Elektron Wechselwirkung dar; zur Bestimmung der Coulombenergie  $E^h$  müßten jeweils die sechsdimensionalen Mehrzentrenintegrationen

$$E^h = \frac{1}{2} \int \int \frac{n(\vec{r}) n(\vec{r}')}{|\vec{r} - \vec{r}'|} d\vec{r}' d\vec{r} = \frac{1}{2} \int n(\vec{r}) v^h(\vec{r}) d\vec{r} \quad . \quad (3.13)$$

durchgeführt oder aber die Poisson–Gleichung gelöst werden, was mit wachsender Integrationspunktzahl den Aufwand drastisch in die Höhe treibt und praktisch nicht durchführbar ist. Daher versucht man, wenigstens **eine der Integrationen zu vereinfachen** und den dadurch offensichtlich auftretenden Fehler so klein wie möglich zu halten. Ein Verfahren, das diese beiden Bedingungen erfüllt, resultiert aus einer Kombination verschiedener Least–Square–Fit Methoden [57], [58]. Es wurde in einer modifizierten Variante von Dunlap et al. [59] angewendet und von Baştuğ [60] auf unseren relativistischen MO-Code adaptiert.

Denken wir uns in (3.13) die Dichte  $n$  durch eine Hilfs- oder Modelldichte und einen Rest zusammengesetzt, was eine Aufteilung des Coulombpotentials  $v^h(\vec{r})$  und der Coulombenergie  $E^h$  bewirkt:

$$n(\vec{r}) = \tilde{n}(\vec{r}) + \Delta n(\vec{r}) \quad , \quad v^h(\vec{r}) = \tilde{v}^h(\vec{r}) + \Delta v^h(\vec{r}) \quad . \quad (3.14)$$

Das führt, wenn wir diese Ladungsdichte  $n$  in (3.13) einsetzen, auf die Gleichung

$$\begin{aligned} E^h &= \frac{1}{2} \int \int \left( \frac{\tilde{n}(\vec{r}) \tilde{n}(\vec{r}')}{|\vec{r} - \vec{r}'|} + 2 \frac{\tilde{n}(\vec{r}) n(\vec{r}')}{|\vec{r} - \vec{r}'|} - 2 \frac{\tilde{n}(\vec{r}) \tilde{n}(\vec{r}')}{|\vec{r} - \vec{r}'|} + \frac{\Delta n(\vec{r}) \Delta n(\vec{r}')}{|\vec{r} - \vec{r}'|} \right) d\vec{r} d\vec{r}' \quad (3.15) \\ &= \underbrace{\int \int \frac{\tilde{n}(\vec{r}) n(\vec{r}')}{|\vec{r} - \vec{r}'|} d\vec{r} d\vec{r}' - \frac{1}{2} \int \int \frac{\tilde{n}(\vec{r}) \tilde{n}(\vec{r}')}{|\vec{r} - \vec{r}'|} d\vec{r} d\vec{r}'}_{E_1^h} + \underbrace{\frac{1}{2} \int \int \frac{\Delta n(\vec{r}) \Delta n(\vec{r}')}{|\vec{r} - \vec{r}'|} d\vec{r} d\vec{r}'}_{E_{rest}} \end{aligned}$$

Dies können wir in Formel (2.35) einsetzen. Der letzte Summand von (3.15) wird von nun an bei der Berechnung der totalen Energie fortgelassen, da er die exakte Berechnung des Coulombpotentials voraussetzt. Physikalisch stellt er die Selbstwechselwirkungsenergie einer Differenzladungsdichte  $(n - \tilde{n})$  dar. Wir bezeichnen diesen Term mit  $E_{rest}$ . Dies bedeutet aber einen Fehler zweiter Ordnung in der Energie, den es zu minimieren gilt. Eine einfache Forderung, die wir zunächst an die Modelldichte stellen können ist die, daß auch diese die korrekte Elektronenzahl erhält:

$$\int \tilde{n}(\vec{r}) d\vec{r} = N \quad . \quad (3.16)$$

Nach Ellis und Delley [57] läßt sich eine Modelldichte folgendermaßen konstruieren:

$$\tilde{n}(\vec{r}) = \sum_{\sigma} D_{\sigma} k_{\sigma}(\vec{r}) = \sum_{\sigma} D_{\sigma} \underbrace{w_{\sigma}(r) Y_{lm}(\Omega)}_{k_{\sigma}(\vec{r})} . \quad (3.17)$$

Hier sind die Funktionen  $w_{\sigma}(r)$  die Summen der Betragsquadrate der radialen atomaren Wellenfunktionen; die Modelldichte wird somit gut an die physikalischen Gegebenheiten angepaßt sein. Das Auftreten der Kugelflächenfunktionen entspricht einer Multipolentwicklung; in unseren Rechnungen wurde bisher nur der Monopolanteil verwendet. Da die Funktionen  $w_{\sigma}(r)$  im Prinzip festliegen, bleiben für die Minimierung von  $E_{rest}$  die Koeffizienten  $D_{\sigma}$  zu bestimmen. Von Dunlap et al. [59] wurde vorgeschlagen, die Restenergie  $E_{rest}$  als Funktional  $\mathcal{L}_2[\tilde{n}] = \mathcal{L}_2[D_{\sigma}]$  der Modelldichte aufzufassen und (3.16) als Nebenbedingung in die Variation einzuarbeiten. Das führt auf folgende Variationsgleichungen

$$\delta_{D_{\sigma}} \left\{ \mathcal{L}_2[D_{\sigma}] - \epsilon_2 \underbrace{\left( N - \int \tilde{n}[D_{\sigma}] d\vec{r} \right)}_{g_2[D_{\sigma}]} \right\} \doteq 0 \quad . \quad (3.18)$$

$\epsilon_2$  ist hier der Lagrangeparameter einer zweiten Variation<sup>2</sup>. Wir wollen die Variation hier nicht weiter ausführen; die explizite Gestalt der Variationsgleichungen findet sich in [60], (Kapitel 4.1). Wir bekommen ein Gleichungssystem zur Bestimmung der Koeffizienten  $D_{\sigma}$ . Gleichzeitig kann gezeigt werden, daß damit  $E_{rest}$  minimiert ist. Dies bezeichnet man auch als variationale Konsistenz des Energiefunktional. Im Kapitel 4 bei der Herleitung des Energiegradienten für die Geometrieoptimierung werden wir einige in diesem Abschnitt gemachte Ausführungen nochmals diskutieren. Für eine Beschränkung auf den Monopolanteil der Entwicklung (3.17) kann das aus  $\tilde{n}$  resultierende Potential  $\tilde{v}^h$  mittels des hier vorgestellten Verfahrens aus **eindimensionalen Integralen** gebildet werden. **Dies ist die eigentliche Vereinfachung bei der Berechnung der Coulombwechselwirkung.**

---

<sup>2</sup>Die erste Variation bezieht sich auf die Ableitung der Kohn-Sham Gleichungen aus dem Energiefunktional  $\mathcal{E}[n]$ .

Zusammenfassung:

Die Berechnung der totalen Energie setzt unter anderem die exakte Behandlung der Elektron–Elektron Wechselwirkung voraus. Weil wir aus rechentechnischen Gründen die dreidimensionale Poissongleichung zur Berechnung des Potentials  $v^h(\vec{r})$  aus der exakten Dichte nicht lösen können, nähern wir dieses Potential durch eine bezüglich der atomaren Zentren sphärischen Modelldichte  $\tilde{n}$  an. Der entstehende Fehler kann (nur dadurch) minimiert werden, wenn wir ein spezielles Verfahren zur Bestimmung dieser Modelldichte anwenden.

## 3.6 Ergebnisse der Rechnungen

In diesem Kapitel wollen wir die Ergebnisse der Dichtefunktionaluntersuchungen darlegen. Bevor wir dies ausführlicher tun, seien noch einige Bemerkungen zur Berechnungsmethode angegeben. Sodann wollen wir anhand der 'klassischen' Beispiele unsere Rechnungen mit anderen vergleichen und diskutieren. Anschließend geben wir die Ergebnisse von Untersuchungen zu superschweren Verbindungen an.

Selbstkonsistente GGA-Rechnungen erfordern in den Potentialen Dichtegradienten nicht nur erster, sondern auch zweiter Ordnungen. Bereits zweite Ableitungen der (exakten) Dichte (2.10) stehen programmtechnisch bisher nicht zur Verfügung; der Aufwand für deren Berechnung im dreidimensionalen Integrationsgebiet ist zu groß. Da die Austausch-Korrelationsenergie ihrerseits jedoch nur Funktionen der Dichte und deren ersten Gradienten enthält (vergleiche letzten Abschnitt), lassen sich die Gradientenkorrekturen relativ leicht störungstheoretisch berücksichtigen. Alle nun folgenden Ergebnisse wurden erhalten, indem die Kohn-Sham Gleichungen iterativ mit LDA/RLDA Potential gelöst wurden und die GGA/RGGA anschließend perturbativ in einer 'post-scf'-Rechnung ausgewertet wurde. Ein analoge Vorgehensweise wurde bereits von Ziegler [61] und von Becke [62] auf nichtrelativistischem Niveau praktiziert. Wir wollen noch darauf verweisen, daß sich die GGA/RGGA-Potentiale dann selbstkonsistent mitnehmen lassen, wenn statt der exakten Dichte die Modelldichte, wie im letzten Abschnitts diskutiert, verwendet wird. Da wir bisher nur den Monopolanteil mitnehmen, sind zweite Ableitungen mit weit weniger Aufwand zugänglich. Man erkaufte sich jedoch dadurch konzeptionelle Schwierigkeiten in dem Sinne, daß dadurch das Energiefunktional abgeändert wird und die variationale Konsistenz nicht mehr gewährleistet ist.

Zur Untersuchung der ursprünglichen Fragestellung, ob relativistische Beiträge in den Austausch-Korrelationsenergien Auswirkungen auf die molekularen Bindungseigenschaften haben, müssen zwei Rechnungen immer unter Beibehaltung der relativistischen Kinematik der Elektronen durchgeführt werden, eine mit relativistischem und eine zweite mit dem entsprechenden nichtrelativistischen Austausch-Korrelationsfunktional. Letzteres läßt sich dadurch simulieren, daß die Lichtgeschwindigkeit einfach auf einen sehr großen Wert gesetzt wird. Die transversalen Energiekomponenten und die relativistischen Korrekturen in der longitudinalen Austausch-Korrelationsenergie verschwinden dann.

### 3.6.1 Die klassischen Beispiele

Bestens untersucht sind die relativistischen Effekte bei den homologen Dimeren Kupfer ( $\text{Cu}_2$ ), Silber ( $\text{Ag}_2$ ) und Gold ( $\text{Au}_2$ ) [65, 66]. Wir wollen nun sehen, wie sich unsere Rechnungen hier einordnen.

Die Abkürzungen in den Tabellen haben folgende Bedeutung:

Die erste Spalte bezieht sich auf die Behandlung der kinetischen Energie. Hier steht R für den voll-relativistischen und SR für den schwach-relativistischen <sup>3</sup> Fall. Die zweite und dritte Spalte gibt an, welches Austausch (x)- bzw. Korrelationsfunktional (c) in den jeweiligen Rechnungen benutzt wurde. Dabei gelten folgende Nomenklaturen für

- die Austauschfunktionale:  
LDA – Lokaldichtenäherung; B88 – Becke88; PW91 – Perdew-Wang91
- und Korrelationsfunktionale:  
VWN – Vosko-Wilk-Nusair; P86 – Perdew86; PW91 – Perdew-Wang91.

Diese Funktionale sind nichtrelativistische Dichtefunktionale. Ist dem Funktional noch ein R vorangestellt, so wollen wir damit andeuten, daß wir die entsprechenden relativistischen Erweiterungen benutzt haben. In die Klasse der GGA bzw. RGGA Funktionale gehören B88, PW91 (sowohl für Austausch als auch für Korrelation) und P86, da diese mit Gradientenkorrekturen ausgewertet werden. Zur Erinnerung: Das VWN Funktional wird auch oft unter LDA für die Korrelation geführt. Alle mit unserem Programm erhaltenen Werte sind spin-unpolarisierte Rechnungen (up). Das gilt auch für die atomaren Referenzenergien. Zum Vergleich sind schwach-relativistische, spinpolarisierte Rechnungen (sp) von Rösch et al. [35] und von Baerends und Mitarbeitern [36] angegeben.

t	Funktional		$R_e$ (a.u.)	$D_e(\text{up})$ (eV)	$D_e(\text{sp})$ (eV)	$\omega$ ( $\text{cm}^{-1}$ )	Ref.
	x	c					
R	RLDA	RVWN	4.17	3.23		285	unser
SR	RLDA	RVWN	4.05		2.85	306	[35]
SR	RLDA	RVWN	4.06		2.77	298	[36]
R	RB88	RVWN	4.35	2.38		252	unser
R	B88	VWN	4.35	2.38		252	unser
R	RPW91	RPW91	4.28	2.86		264	unser
R	PW91	PW91	4.28	2.86		264	unser
R	RB88	P86	4.27	2.82		269	unser
SR	B88	P86	4.16		2.27	285	[35]
SR	B88	P86	4.18		2.19	272	[36]
	exp.		4.20	2.05		265	[63]

**Tabelle 3.1:** spektroskopische Parameter von  $\text{Cu}_2$ ; Basis:  $[\text{Ar}] 3d^{10} 4s^1 4p^0$

Die Ergebnisse der Rechnungen lassen sich folgendermaßen zusammenfassen:

<sup>3</sup>Diese mitunter auch als skalar-relativistisch bezeichnete Methode beruht auf einer Entkopplung der großen und kleinen Spinorkomponenten durch eine Douglas-Kroll Transformation [67, 68].

t	Funktional		$R_e$ (a.u.)	$D_e(\text{up})$ (eV)	$D_e(\text{sp})$ (eV)	$\omega$ ( $\text{cm}^{-1}$ )	Ref.
	x	c					
R	RLDA	RVWN	4.81	2.65		204	unser
SR	RLDA	RVWN	4.73		2.25	209	[35]
SR	RLDA	RVWN	4.69		2.28	208	[36]
R	RB88	RVWN	5.04	1.93		175	unser
R	B88	VWN	5.04	1.93		175	unser
R	RB88	P86	4.94	2.32		188	unser
R	RPW91	RPW91	4.93	2.27		189	unser
R	PW91	PW91	4.93	2.27		189	unser
SR	B88	P86	4.87		1.67	185	[35]
SR	B88	P86	4.84		1.71	183	[36]
exp.			4.78	1.66		192	[63]

**Tabelle 3.2:** spektroskopische Parameter von  $\text{Ag}_2$ ; Basis:  $\text{Kr } 4d^{10} 5s^1 5p^0$

1. Die RLDA überschätzt in allen Fällen die Bindungsenergie; beim  $\text{Cu}_2$  um 1.21 eV, beim  $\text{Ag}_2$  um 0.98 eV und beim  $\text{Au}_2$  um 1.07 eV. Die Schwingungsfrequenzen sind generell zu groß, die Bindungsabstände werden vergleichsweise gut wiedergegeben.
2. Gradientenkorrekturen ergeben deutlich bessere Werte für die Bindungsenergien, liefern aber tendenziell größere Bindungsabstände. Während sich bei den Rechnungen, in denen auch Gradientenkorrekturen im Korrelationsfunktional berücksichtigt wurden, die Werte aller spektroskopischen Parameter um jeweils etwa denselben Wert gruppieren, shiftet im Falle der LDA für die Korrelation der Bindungsabstand zu größeren Werten hin, als es die GGA bereits tut. Gleichzeitig nimmt die Bindungsenergie weiter ab und kommt den experimentellen und den Literaturwerten am nächsten.
3. Die relativistischen Beiträge im Austausch-Korrelationsfunktional haben keinen merklichen Einfluß auf die Bindungsverhältnisse. Beim Gold ist dieser Effekt in der Bindungsenergie 0.01 eV und im Bindungsabstand 0.01 a.u., vgl. (RPW91/RPW91 – PW91/PW91).

Aus diesen Befunden lassen sich folgende Schlußfolgerungen ziehen:

Bei den hier untersuchten Systemen im mittleren und höheren Z-Bereich haben Gradientenkorrekturen eine ebenso wichtige Bedeutung wie bei Molekülen mit leichten Konstituenten. Da die unempfindlichste Größe in den spektroskopischen Parametern der Bindungsabstand <sup>4</sup> ist, läßt sich daran in etwa die Güte des verwendeten Funktionals festmachen. Daher sollten in den GGA-Rechnungen auch Gradientenkorrekturen im Korrelationsfunktional berücksichtigt werden.

<sup>4</sup>Ein Problem stellt die korrekte Berechnung der Energiedifferenzen dar. Hier sollte man konsistenterweise atomare Energien abziehen, die aus Rechnungen mit demselben Gitter stammen. Da dies für jeden Abstand gemacht werden muß, sind derartige Rechnungen mitunter recht mühsam. In unserem Falle hat sich aber diesbezüglich kein merklicher Unterschied ergeben. Die atomaren Referenzenergien wurden aus molekularen Einzentrerenrechnungen am Minimum der Potentialkurve erhalten.

t	Funktional		$R_e$ (a.u.)	$D_e(\text{up})$ (eV)	$D_e(\text{sp})$ (eV)	$\omega$ ( $\text{cm}^{-1}$ )	Ref.
	x	c					
R	RLDA	RVWN	4.73	3.37		194	unser
SR	RLDA	RVWN	4.65		2.92	191	[35]
SR	RLDA	RVWN	4.64		2.92	193	[36]
R	RB88	RVWN	4.93	2.34		157	unser
R	B88	VWN	4.92	2.33		157	unser
R	RB88	P86	4.83	2.84		171	unser
R	RPW91	RPW91	4.82	2.89		180	unser
R	PW91	PW91	4.81	2.88		181	unser
SR	B88	P86	4.76		2.27	177	[35]
SR	B88	P86	4.75		2.26	174	[36]
exp.			4.67	2.30		191	[63]

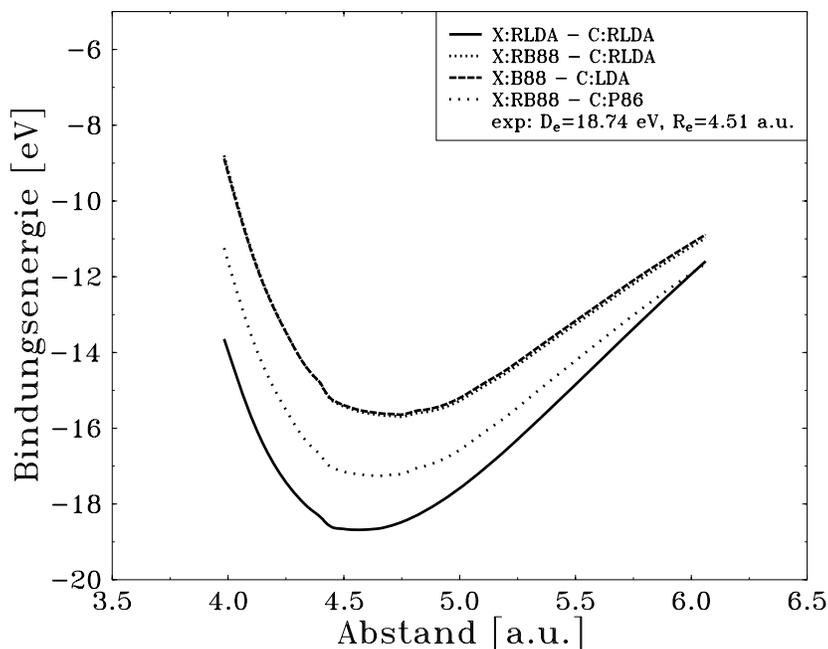
**Tabelle 3.3:** spektroskopische Parameter von  $\text{Au}_2$ ; Basis:  $[\text{Xe}] 4f^{14} 5d^{10} 6s^1 6p^0$

### 3.6.2 Untersuchungen an hochrelativistischen Systemen

Im letzten Abschnitt dieses Kapitels wollen wir die ersten Anwendungen der relativistischen Dichtefunktionaltheorie auf Moleküle mit superschweren Konstituenten präsentieren (Tabelle (3.4)). In der ersten Zeile sind die Ergebnisse der Rechnungen auf reinem RLDA-Niveau angegeben (das Korrelationsfunktional ist hier das VWN Funktional). In den folgenden beiden Zeilen wurden Gradientenkorrekturen für den Austausch hinzugenommen, zusätzlich wurden hier nochmals die relativistischen und nichtrelativistischen Funktionale gegenübergestellt. Selbst bei diesen sehr schweren Systemen ist diesbezüglich kein Unterschied im Bindungsverhalten erkennbar. In der vorletzten Zeile schließlich wurde auch das Korrelationsfunktional mit Gradientenkorrekturen behandelt. Hier sind die Rechnungen nur am System  $\text{RfCl}_4$  vorgenommen worden.

Funktional		$\text{HfCl}_4$		$\text{RfCl}_4$	
x	c	$R_e[\text{a.u.}]$	$D_e(\text{eV})$	$R_e[\text{a.u.}]$	$D_e(\text{eV})$
RLDA	RVWN	4.43	-20.44	4.53	-18.68
RB88	RVWN	4.54	-17.14	4.75	-15.69
B88	VWN	4.53	-17.12	4.75	-15.64
RB88	P86	-	-	4.64	-17.25
exp. [64]		4.37	-20.55	4.51	-18.74

**Tabelle 3.4:** Bindungsenergien und -längen der Tetrachloride von Hafnium ( $Z=72$ , Basis:  $[\text{Xe}] 4f^{14} 5d^2 6s^2 6p^0$ ) und Rutherfordium ( $Z=104$ , Basis:  $[\text{Rn}] 5f^{14} 6s^2 6p^6 6d^2 7s^2 7p^0$ ). Es wurde jeweils eine  $T_d$ -Symmetrie zugrunde gelegt.



**Abbildung 3.2:** Erste voll-relativistische Rechnungen mit gradientenkorierten Dichtefunktionalen am System  $\text{RfCl}_4$ . Als Symmetrie wurde hier eine tetraedrische Struktur ( $T_d$ ) angenommen. Als Basis wurden die atomaren 1s-7p Wellenfunktionen für Rf und 1s-3p Wellenfunktionen für Cl mitgenommen.

Wie die Tabelle zeigt, werden in Bezug auf die experimentellen Werte hier die besten Ergebnisse mit den RLDA-Rechnungen erzielt. In Abbildung (3.2) sind nochmals die Trends in den Bindungsabständen erkennbar. Im Falle der GGA's gibt es einen generellen Shift zu größeren Werten, unabhängig davon, ob Gradientenkorrekturen im Korrelationsfunktional mitberücksichtigt werden oder nicht. Tut man dies nicht, so erfährt der Bindungsabstand eine weitere Vergrößerung. Dies ist in Übereinstimmung mit den Resultaten der Rechnungen an den dimeren Systemen  $\text{Cu}_2$ ,  $\text{Ag}_2$  und  $\text{Au}_2$  sowie mit anderen DFT-Untersuchungen im Bereich der leichten Elemente [53].



# Kapitel 4

## Der Energiegradient

Gegenstand dieses Kapitels ist die Entwicklung des Formelwerks für den Gradient der totalen Energie. Er ist die entscheidende Größe für eine erfolgreiche Geometrieoptimierung, mit der wir uns im Kapitel 5 beschäftigen wollen. Historisch gesehen sind die Arbeiten von Hellmann [69] und Feynman [70] über intramolekulare elektrostatische Kräfte sowie von Pulay [71] über sogenannte 'artificial forces', die wir weiter unten näher erläutern, wichtige 'Meilensteine' der Gradiententheorie quantenmechanischer molekularer Vielteilchensysteme. Die Komplexität und der Aufwand einer Gradientenberechnung hängt unmittelbar von der konkreten Art der Energiebestimmung ab. Ausführliche Zusammenstellungen von Energie- und Energiegradientenformeln sind z.B. in [72, 73, 74] zu finden. Mit der Implementierung der Art von Gradienten des folgenden Kapitels in nichtrelativistische Dichtefunktionalprogramme wurde Anfang der achtziger Jahre von Satoko begonnen [75, 76]. Auch andere Gruppen haben dies ebenfalls angegangen und umgesetzt [77, 78, 79].

### 4.1 Der Funktionalansatz

Den meisten Methoden der modernen Elektronenstrukturrechnungen ist eigen, daß sie die Anwendbarkeit eines Variationsverfahrens zur Berechnung der totalen Energie voraussetzen. Dies gilt auch in der Dichtefunktionaltheorie, wo das Energiefunktional bezüglich der (Grundzustands)dichte variiert wird. Der stationäre Wert des Funktionals ist dann die totale Energie. Sie ist eine Funktion, in welche die Kernkoordinaten parametrisch eingehen (Born-Oppenheimer Näherung, vgl. Abschnitt 2.1). Von dieser Funktion suchen wir das energetische Minimum; es handelt sich, wenn wir so wollen, um die Verknüpfung eines Variationsverfahrens mit einem Extremalproblem, welche wir nacheinander in zwei Schritten abarbeiten können:

1. Lösung der Variationsaufgabe, wodurch wir aus dem Funktional eine parameterabhängige Funktion gewinnen;
2. Differenzieren dieser Funktion nach dem Parameter  $\alpha$ .

Stellen wir uns vor, wir haben ein Funktional  $\mathcal{L} = \mathcal{L}[C]$ , welches von gewissen Variationsgrößen  $C$  abhängt. Zunächst besteht die Variationsaufgabe darin, diejenigen  $C$  zu finden, die unser Funktional gegebenenfalls auch unter Nebenbedingungen stationär machen. Zusätzlich wollen wir noch eine parametrische Abhängigkeit des Funktionals von einem beliebigen äußeren Parameter  $\lambda$ , z.B. einem externen Feld oder Kernkoordinaten zulassen. Die Frage ist nun, wie sich der stationäre Wert des Funktionals bezüglich dieses Parameters ändert. In den nächsten Abschnitten werden wir sehen, daß schon das Ausformulieren und Aufschreiben der Vorgehensweise der Gradientenbestimmung zu besonderer Sorgfalt gemahnt. Um die Herleitungen nicht unnötig aufzublähen, wollen wir zunächst auf physikalische Größen verzichten und die formale mathematische Theorie angeben und später die entwickelten Formeln mit 'physikalischem Leben' erfüllen. Am Schluß wollen wir die Gradiententerme explizit angeben und auf eventuelle Modifikationen eingehen, die sich durch Anwendung der Frozen-Core Näherung (FC) ergeben.

### 4.1.1 Vorbemerkungen und Vereinbarungen

Vorab noch einige Bemerkungen zur Konvention. Die in den kommenden Abschnitten auftretenden Variationsgrößen  $C(\lambda)$  (bzw. auch  $D$ ), von ihnen abhängige Größen  $F$  und deren Ableitungen  $F_\lambda$  können unterschiedliche Bedeutung haben. Wir stellen das tabellarisch zusammen:

$C = C(\lambda)$	F	Bedeutung	Ableitung bzgl. $\lambda$
1. Variable	$F(C)$	Funktion einer Veränderlichen	$F_\lambda = \frac{dF}{dC} \frac{dC}{d\lambda}$
2. Vektor $C = (C_1, \dots, C_n)$	$F(C_1, \dots, C_n)$	Funktion mehrerer Veränderlicher	$F_\lambda = \sum_{i=1}^n \frac{\partial F}{\partial C_i} \frac{\partial C_i}{\partial \lambda}$
3. Funktion $C(x)$	$\mathcal{F}[C]$	Funktional	$\mathcal{F}_\lambda = \int \frac{\delta \mathcal{F}}{\delta C} \frac{\partial C}{\partial \lambda} dx$
4. Vektorfunktionen $C(x) = C_1(x), \dots, C_n(x)$	$\mathcal{F}[C_1, \dots, C_n]$	Funktional	$\mathcal{F}_\lambda = \sum_i^n \int \frac{\delta \mathcal{F}}{\delta C_i} \frac{\partial C_i}{\partial \lambda} dx$

**Tabelle 4.1:** zur Konvention der in den folgenden Abschnitten verwendeten Größen

Die Lösungen (bzw. die Lösungsfunktionen) der Variationsgleichungen kennzeichnen wir durch einen Querstrich über dem entsprechenden Buchstaben: z.B. bedeute  $\bar{F} = F(\bar{C})$  für die von  $C$  abhängige Größe  $F$ , daß wir die gesuchte Lösung  $\bar{C}$  in  $F$  einzusetzen haben.

Im Abschnitt (4.3) werden uns komplexe Größen, z.B. Wellenfunktionen begegnen, die aus Real- und Imaginärteil bestehen:  $C = \Re(C) + \Im(C)$ . Die Größe  $F$  hängt von beiden Anteilen ab;  $F[C] = F[\Re(C), \Im(C)]$ . Hierbei läßt sich  $F$  so umformulieren, daß es von  $C$

und dessen konjugiert komplexen Anteil  $C^*$  abhängt, also  $F[C] = F[C, C^*]$ . Die Variationen umfassen dann sowohl deren Real- als auch Imaginärteil. Oftmals genügt es, eine der beiden Variationen anzugeben, da die Variation bezüglich des anderen Anteils auf dasselbe Ergebnis führt. Davon werden wir Gebrauch machen und die Behandlung des konjugiert Komplexen durch *c.c.* symbolisieren.

$$\begin{aligned} \frac{\delta \mathcal{F}}{\delta C} \frac{\partial C}{\partial \lambda} &= \frac{\delta \mathcal{F}}{\delta \Re(C)} \frac{\partial \Re(C)}{\partial \lambda} + \frac{\delta \mathcal{F}}{\delta \Im(C)} \frac{\partial \Im(C)}{\partial \lambda} \\ &= \frac{\delta \mathcal{F}}{\delta C} \frac{\partial C}{\partial \lambda} + \frac{\delta \mathcal{F}}{\delta C^*} \frac{\partial C^*}{\partial \lambda} \\ &:= \frac{\delta \mathcal{F}}{\delta C} \frac{\partial C}{\partial \lambda} + c.c. \end{aligned} \quad (4.1)$$

Die Betrachtungen wollen wir wie folgt vornehmen. Zunächst untersuchen wir ausschließlich das Funktional, danach die Variationsaufgabe unter Berücksichtigung einer Nebenbedingung (die auch von unseren Variationsgrößen  $C$  abhängt) untersuchen. Drittens schließlich soll auch diese Nebenbedingung vom externen Parameter abhängen.

### 4.1.2 Funktional ohne Nebenbedingung (NB)

Betrachten wir zunächst das Funktional. Wir haben

$$\mathcal{L} = \mathcal{L}[C; \lambda]. \quad (4.2)$$

Die stationären Punkte des Funktionals erhalten wir aus der Forderung

$$\delta_C \mathcal{L} = \frac{\delta \mathcal{L}}{\delta C} \doteq 0, \quad (4.3)$$

die Lösungen  $\bar{C}$  erfüllen *per constructionem* die Bedingung (4.3). Jetzt ist aus dem Funktional eine parameterabhängige Funktion geworden und wir können schreiben

$$\bar{\mathcal{L}}(\lambda) := \mathcal{L}[\bar{C}(\lambda); \lambda]. \quad (4.4)$$

Die stationäre Punkte bezüglich dieser Funktion bekommen wir aus der Differentiation von (4.4) nach  $\lambda$  :

$$\frac{d\bar{\mathcal{L}}}{d\lambda} = \underbrace{\frac{\delta \mathcal{L}}{\delta C} \Big|_{\bar{C}}}_{=0} \frac{\partial \bar{C}}{\partial \lambda} + \frac{\partial \mathcal{L}}{\partial \lambda} \Big|_{\bar{C}} \quad (4.5)$$

Wegen (4.3) verschwindet der erste Summand. Von der Differentiation bleibt damit

$$\frac{d\bar{\mathcal{L}}(\lambda)}{d\lambda} = \frac{\partial \mathcal{L}(\lambda)}{\partial \lambda} \Big|_{\bar{C}} \quad (4.6)$$

übrig. Diese Aussage firmiert in der Literatur unter dem Begriff des Prinzips der formalen Differentiation eines Extremums [80].

### 4.1.3 Funktional mit parameterunabhängiger Nebenbedingung

Die bisherige Formulierung galt für den Fall einer freien Variation. Berücksichtigen wir, daß die zur Variation zugelassenen Größen  $C$  gewissen Restriktionen genügen sollen, müssen wir dies auch im Variationsansatz beachten. Die Variation erfolgt dann unter Nebenbedingungen, welche wir explizit durch Einarbeitung Lagrange'scher Multiplikatoren in die Variationsgleichungen berücksichtigen. Mit  $g[C] \doteq 0$  sei im folgenden die Nebenbedingung, mit  $\epsilon$  der Lagrangemultiplikator bezeichnet. Die Variationsgleichung zur Bestimmung der  $C$  lautet statt (4.3) jetzt

$$\delta_C \left\{ \mathcal{L}[C; \lambda] - \epsilon g[C] \right\} \doteq 0. \quad (4.7)$$

Als Lösung bekommen wir wiederum  $C = \bar{C}$ . Diese machen nicht nur das Funktional stationär, sondern erfüllen auch unsere Nebenbedingung, für die ja  $\bar{g} = g[\bar{C}] \equiv 0$  gilt. Da die Nebenbedingung immer Null ist, gilt dies erst recht, wenn wir  $\bar{C}$  in diese einsetzen. Aus (4.7) folgt, daß wir wieder dieselbe Funktion  $\mathcal{L}$  wie in Gleichung (4.4) bekommen. Somit gilt auch die Folgerung (4.5) einschließlich der Formel (4.6).

Dieses Resultat können wir uns auch folgendermaßen klarmachen: Aus (4.7) folgt

$$\frac{\delta \mathcal{L}}{\delta C} \Big|_{\bar{C}} = \epsilon \frac{\delta g}{\delta C} \Big|_{\bar{C}}. \quad (4.8)$$

Multiplizieren wir Gleichung (4.8) mit  $\frac{\partial \bar{C}}{\partial \lambda}$ , läßt sich der erste Summand in Gleichung (4.5) durch die rechte Seite von (4.8) ersetzen und wir erhalten

$$\frac{d\bar{\mathcal{L}}(\lambda)}{d\lambda} = \epsilon \frac{\partial g}{\partial C} \Big|_{\bar{C}} \frac{\partial \bar{C}}{\partial \lambda} + \frac{\partial \mathcal{L}(\lambda)}{\partial \lambda} \Big|_{\bar{C}}. \quad (4.9)$$

Der erste Summand verschwindet aus folgendem Grunde: Die Nebenbedingung ist immer Null, insbesondere auch dann, wenn  $\bar{C}$  in diese eingesetzt wird,  $\bar{g} \equiv 0$ . Zudem soll  $g$  nach Voraussetzung unabhängig von  $\lambda$  sein. Somit gilt

$$\frac{d\bar{g}}{d\lambda} = \frac{\delta g}{\delta C} \Big|_{\bar{C}} \frac{d\bar{C}}{d\lambda} = 0; \quad (4.10)$$

die linke Seite von (4.8) ist damit Null und wir bekommen wiederum Gleichung (4.6).

#### 4.1.4 Funktional mit parameterabhängiger Nebenbedingung

In einem dritten Schritt wollen wir sowohl eine Parameterabhängigkeit des Funktionals als auch der Nebenbedingung zulassen. Dann haben wir anknüpfend an (4.7)

$$\delta_C \left\{ \mathcal{L}[C; \lambda] - \epsilon g[C; \lambda] \right\} \doteq 0. \quad (4.11)$$

Dies liefert wieder Gleichung (4.8) und erneutes Einsetzen der rechten Seite in die Gleichung (4.5) führt wiederum auf Gleichung (4.9). Im Unterschied zum letzten Abschnitt haben wir jetzt  $\bar{g}(\lambda) := g[\bar{C}; \lambda]$ , womit wir andeuten wollen, daß  $\bar{g}$  nach Einsetzen der Lösungen des Variationsproblems eine Funktion von  $\lambda$  ist. Die Differentiation der Nebenbedingung liefert jetzt im Unterschied zu (4.10)

$$\frac{d\bar{g}}{d\lambda} = \left. \frac{\delta g}{\delta C} \right|_{\bar{C}} \frac{d\bar{C}}{d\lambda} + \left. \frac{\partial g}{\partial \lambda} \right|_{\bar{C}}. \quad (4.12)$$

Der erste Summand der rechten Seite ist (bis auf  $\epsilon$ ) identisch mit dem selbigen in (4.9). Er läßt sich durch den zweiten Summanden ersetzen. Wir bekommen damit

$$\frac{d\bar{\mathcal{L}}(\lambda)}{d\lambda} = \left. \frac{\partial \mathcal{L}(\lambda)}{\partial \lambda} \right|_{\bar{C}} - \epsilon \left. \frac{\partial g}{\partial \lambda} \right|_{\bar{C}}. \quad (4.13)$$

Wie wir sehen, ist durch die parameterabhängige Nebenbedingung ein Zusatzterm zum Gradienten entstanden. Was haben wir nun erreicht?

Das wichtigste Ergebnis der bisherigen Formulierung ist, daß wir eine **formale** (noch abstrakte) **Vorschrift zur Berechnung des Energiegradienten unter Ausnutzung eines Variationsproblems** gefunden haben.

## 4.2 Simultane Lösung zweier Variationsprobleme

Nun wird es ein wenig komplizierter. In unserer Ableitung ist bisher keinerlei Näherung gemacht worden und alles ist noch exakt. Bei der praktischen Umsetzung sind aber gewisse Näherungen unumgänglich, wie wir im Abschnitt (3.5) bei der Behandlung der Elektron-Elektron Wechselwirkung gesehen haben. Eine Folge dieser Näherung ist, daß wir in Wirklichkeit vor der Aufgabe stehen, **zwei** gekoppelte Variationsprobleme simultan zu lösen. Das wollen wir im folgenden näher diskutieren. Aus Gründen der Übersichtlichkeit werden wir weiterhin die abstrakte Formulierung beibehalten.

Ausgangspunkt unserer Überlegungen war ein Funktional  $\mathcal{L} = \mathcal{L}[C(\lambda); \lambda]$ . Aufgrund unserer Näherung sind wir gezwungen, unseren Funktionalansatz abzuändern, der beide Variationen der Abschnitte (2.4) und (3.5) beinhaltet. Wir berücksichtigen dies in Form

eines zweiten Satzes von Variationsgrößen  $D$ . Das ursprüngliche Funktional teilen wir in zwei Summanden auf:

$$\mathcal{L} = \mathcal{L}[C; \lambda] = \mathcal{L}_1[C, D; \lambda] + \mathcal{L}_2[C, D; \lambda] \quad . \quad (4.14)$$

Formal führt  $\mathcal{L}_1$  auf den ersten und  $\mathcal{L}_2$  den zweiten unterstrichenen Ausdruck in Gleichung (3.15) aus Abschnitt (3.5). Dort wurde auch darauf hingewiesen, daß der Ausdruck  $E_{rest}$  vernachlässigt werden kann, da er durch die spezielle Dichtefitprozedur des Abschnitts (3.5) minimiert ist. Physikalisch entspricht dieser Term zweiter Ordnung der Selbstenergie einer Differenzladungsdichte  $(n(\vec{r}) - \tilde{n}(\vec{r}))$ . Für die folgenden Ausführungen ist dieser Punkt wichtig, da wir die Kleinheit dieses Terms im Vergleich zu  $\mathcal{L}_1$  als Voraussetzung in die Ableitung des Gradienten hineinstecken müssen. In diesem Sinne teilen wir unser Funktional in einen dominanten und einen für die explizite Berechnung vernachlässigbaren Term auf.

Wie wir in den letzten Abschnitten gesehen haben, bekommt man etwas andere Ausdrücke für die stationären Werte von  $\bar{\mathcal{L}}(\lambda)$ , je nach dem, ob eventuelle Nebenbedingungen parameterabhängig sind oder nicht. Wir wollen die verschiedenen Fälle nicht nochmal gesondert diskutieren sondern gleich den für uns später in Frage kommenden allgemeinen Fall parameterabhängiger Nebenbedingungen behandeln. Vor der Differentiation bezüglich  $\lambda$  geben wir zuerst die Variationsgleichungen an. Doch bevor wir dies im Detail aufschreiben, fällt an Gleichung (4.14) auf, daß trotz der künstlichen Aufteilung  $\mathcal{L}$  nicht von  $D$  abhängt,  $\delta_D \mathcal{L} \equiv 0$ . Das hat zur Folge, daß

$$\left. \frac{\delta \mathcal{L}_1[C, D; \lambda]}{\delta D} \right|_C = - \left. \frac{\delta \mathcal{L}_2[C, D; \lambda]}{\delta D} \right|_C \quad . \quad (4.15)$$

Diese Beziehung werden wir im nun folgenden Abschnitt ausnutzen.

### 4.2.1 1.Schritt: Die Variationsgleichungen

Wir nehmen dazu als Ausgangspunkt die Gleichung (4.11),

$$\delta_C \left\{ \mathcal{L}[C; \lambda] - \epsilon g[C; \lambda] \right\} = \delta_C \left\{ \mathcal{L}_1[C, D; \lambda] + \mathcal{L}_2[C, D; \lambda] - \epsilon g_1[C; D, \lambda] \right\} \doteq 0. \quad (4.16)$$

Durch die Aufteilung von  $\mathcal{L}$  haben wir am Funktional selbst nichts abgeändert; es ist nach wie vor nur von einem Satz Variationsgrößen  $C$  abhängig. Daher geht auch in die Nebenbedingung <sup>1</sup> nur  $C$  zur Variation ein. Diese liefert uns  $\bar{C}$ . Dennoch haben wir uns eine  $D$ -Abhängigkeit eingehandelt und es scheint, daß wir unser Variationsproblem nur

---

<sup>1</sup>Das wir in (4.16)  $g$  mit einem Index 1 versehen haben liegt daran, daß sich diese Nebenbedingung auf die Variationsgleichungen zur Bestimmung der totalen Energie des Abschnitts (2.4) bezieht. Eine weitere Nebenbedingung  $g_2$  ist für den Dichtefit (Abschnitt (3.5)) eingeführt worden.

komplizierter gemacht haben. Obige Gleichung (4.16) ist eine Bestimmungsgleichung für  $\bar{C}$  und muß für jedes zunächst beliebige  $D$  gelten, weil  $\mathcal{L}$  nicht von  $D$  abhängt. Deshalb stellen wir uns vor, daß die  $D = \tilde{D}$  Lösungen eines anderen, hier noch nicht genauer spezifizierten Variationsverfahrens sind und in irgendeiner Weise von den  $C$  abhängen können;  $\tilde{D} = \tilde{D}(C)$ . Dann können wir (4.16) auch schreiben als

$$\delta_C \left\{ \mathcal{L}[C; \lambda] - \epsilon g[C; \lambda] \right\} = \delta_C \left\{ \mathcal{L}_1[C, \tilde{D}(C); \lambda] + \mathcal{L}_2[C, \tilde{D}(C); \lambda] - \epsilon g_1[C, \tilde{D}(C); \lambda] \right\} \doteq 0. \quad (4.17)$$

An dieser Stelle setzt nun die Näherung ein. Der Term  $\mathcal{L}_2$  liefert, wie im Abschnitt (3.5) dargelegt, einen Restenergiebeitrag zur totalen Energie in zweiter Ordnung. Die Vernachlässigung dieses Terms ist physikalisch durch die Annahme motiviert, daß unsere Modelldichte nahezu der exakten Dichte entspricht. Mathematisch gesehen hat das für das Funktional  $\mathcal{L}_2$  zur Folge, daß die  $\tilde{D}$  so beschaffen sein müssen, daß  $\mathcal{L}_2$  in Abhängigkeit von  $C$  klein sein muß. Unter dieser Voraussetzung geht dann Gleichung (4.17) über in

$$\delta_C \left\{ \mathcal{L}[C; \lambda] - \epsilon g[C; \lambda] \right\} \approx \delta_C \left\{ \mathcal{L}_1[C, \tilde{D}(C); \lambda] - \epsilon g_1[C, \tilde{D}(C); \lambda] \right\} \doteq 0. \quad (4.18)$$

Variieren wir nun nach  $C$ , erhalten wir

$$\left. \frac{\delta \mathcal{L}_1}{\delta C} \right|_{\tilde{D}} + \left. \frac{\delta \mathcal{L}_1}{\delta D} \right|_{C, \tilde{D}} \frac{\partial \tilde{D}}{\partial C} - \epsilon_1 \left. \frac{\delta g_1}{\delta C} \right|_{\tilde{D}} - \epsilon_1 \left. \frac{\delta g_1}{\delta D} \right|_{C, \tilde{D}} \frac{\partial \tilde{D}}{\partial C} \doteq 0. \quad (4.19)$$

Diese Gleichung hat formale Ähnlichkeit mit der ursprünglichen Variationsgleichung (4.8), wenn der zweite und vierte Summand nicht auftreten würden. Wir können versuchen, diese Terme zu eliminieren. Dazu erinnern wir uns der Tatsache, daß unser Funktional unabhängig von  $D$  ist. Dann können wir nämlich die Beziehung (4.15) ausnutzen und  $\frac{\delta \mathcal{L}_1}{\delta D}$  im zweiten Summanden von Gleichung (4.19) durch  $-\frac{\delta \mathcal{L}_2}{\delta D}$  ersetzen. Daraus folgt

$$\left. \frac{\delta \mathcal{L}_1}{\delta C} \right|_{\tilde{D}} - \left. \frac{\delta \mathcal{L}_2}{\delta D} \right|_{C, \tilde{D}} \frac{\partial \tilde{D}}{\partial C} - \epsilon_1 \left. \frac{\delta g_1}{\delta C} \right|_{\tilde{D}} - \epsilon_1 \left. \frac{\delta g_1}{\delta D} \right|_{C, \tilde{D}} \frac{\partial \tilde{D}}{\partial C} \doteq 0. \quad (4.20)$$

Jetzt konkretisieren wir die Abhängigkeit  $\tilde{D}(C)$ . Wir fordern, daß wir die in der Variation (4.19) eingesetzten Lösungen  $\tilde{D}$  aus dem Minimum von  $\mathcal{L}_2$  gewinnen,

$$\delta_D \mathcal{L}_2 \Big|_C \doteq 0 \quad . \quad (4.21)$$

Diese Gleichung beinhaltet die Forderung der Stationarität des Funktionals  $\mathcal{L}_2$  bezüglich  $D$ . Zusammen mit der Variationsgleichung (4.19) haben wir somit ein Gleichungssystem für  $C$  und  $D$  vorliegen. Bezüglich der  $D$ -Variation haben wir nun die Option, entweder frei zu variieren oder auch hier den Variationsraum durch eine Nebenbedingung einzuschränken. Wir wollen hier beide möglichen Varianten angeben. Das ändert natürlich die Gleichungen unseres System ab.

• **erste Variante – freie Variation**

Stellen wir keine Forderungen an die Variationsparameter  $D$ , können wir die letzte Gleichung direkt benutzen und für die sich daraus ergebenden Lösungen  $\tilde{D}$  gilt

$$\left. \frac{\delta \mathcal{L}_2}{\delta D} \right|_{C, D=\tilde{D}} = 0 \quad . \quad (4.22)$$

Dies definiert ein spezielles Funktional  $\tilde{D}(C)$  und der zweite Summand aus (4.20) verschwindet. Es bleibt

$$\left. \frac{\delta \mathcal{L}_1}{\delta C} \right|_{\tilde{D}} - \epsilon_1 \left. \frac{\delta g_1}{\delta C} \right|_{\tilde{D}} - \epsilon_1 \left. \frac{\delta g_1}{\delta D} \right|_{\tilde{D}} \frac{\partial \tilde{D}}{\partial C} \doteq 0 \quad . \quad (4.23)$$

Das ist eine Möglichkeit, unser Gleichungssystem zu lösen.

• **zweite Variante – Variation mit NB**

Verlangen wir, daß die Variationsparameter  $D$  nicht beliebig sind, können wir dies wieder durch eine Nebenbedingung  $g_2[C, D] = 0$  in den Variationsgleichungen berücksichtigen. Der Variationsraum für  $D$  ist nun eingeschränkt. Statt (4.22) haben wir dann

$$\left. \frac{\delta \mathcal{L}_2}{\delta D} \right|_{C, D=\tilde{D}} - \epsilon_2 \left. \frac{\delta g_2}{\delta D} \right|_{C, D=\tilde{D}} \doteq 0 \quad , \quad (4.24)$$

wenn  $\epsilon_2$  der zugehörige Lagrangeparameter ist. Anstelle von (4.23) folgt nun für (4.20)

$$\left. \frac{\delta \mathcal{L}_1}{\delta C} \right|_{\tilde{D}} - \epsilon_2 \left. \frac{\delta g_2}{\delta D} \right|_{\tilde{D}} \frac{\partial \tilde{D}}{\partial C} - \epsilon_1 \left. \frac{\delta g_1}{\delta C} \right|_{\tilde{D}} - \epsilon_1 \left. \frac{\delta g_1}{\delta D} \right|_{\tilde{D}} \frac{\partial \tilde{D}}{\partial C} \doteq 0 \quad . \quad (4.25)$$

Mit (4.24) ist  $g_2$  bezüglich  $D$  festgelegt,  $g_2[C; D = \tilde{D}] := \bar{g}_2[C] = 0$ . Wenn wir das totale Differential bilden, erhalten wir

$$\frac{d\bar{g}_2}{dC} = \left. \frac{\delta g_2}{\delta C} \right|_D + \left. \frac{\delta g_2}{\delta D} \right|_C = 0 \quad . \quad (4.26)$$

Damit können den zweiten Summanden in Gleichung (4.25) umschreiben und wir erhalten

$$\frac{\delta \mathcal{L}_1}{\delta C} \Big|_{\bar{D}} + \epsilon_2 \frac{\delta g_2}{\delta C} \Big|_{\bar{D}} \frac{\partial \tilde{D}}{\partial C} - \epsilon_1 \frac{\delta g_1}{\delta C} \Big|_{\bar{D}} - \epsilon_1 \frac{\delta g_1}{\delta D} \Big|_{\bar{D}} \frac{\partial \tilde{D}}{\partial C} \doteq 0 \quad . \quad (4.27)$$

Mit (4.24) und (4.27) haben wir jetzt ein Gleichungssystem vorliegen, in dem zwei Nebenbedingungen für zwei Variationsprobleme enthalten sind und das uns die gesuchten Lösungen  $\bar{C}, \bar{D}$  liefert:

$$\begin{aligned} \frac{\delta \mathcal{L}_2}{\delta D} \Big|_{C, D=\bar{D}} - \epsilon_2 \frac{\delta g_2}{\delta D} \Big|_{C, D=\bar{D}} &\doteq 0 & ; & \quad g_2 = g_2[C, D]; \\ \frac{\delta \mathcal{L}_1}{\delta C} \Big|_{\bar{D}} + \epsilon_2 \frac{\delta g_2}{\delta D} \Big|_{\bar{D}} \frac{\partial \tilde{D}}{\partial C} - \epsilon_1 \frac{\delta g_1}{\delta C} \Big|_{\bar{D}} - \epsilon_1 \frac{\delta g_1}{\delta D} \Big|_{\bar{D}} \frac{\partial \tilde{D}}{\partial C} &\doteq 0 & ; & \quad g_1 = g_1[C, D] \quad . \end{aligned} \quad (4.28)$$

An dieser Stelle wollen wir nun die **gegenseitigen Abhängigkeiten der Nebenbedingungen** von  $C$  und  $D$  **aufgeben** und unsere Gleichungen diskutieren, wenn  $g_1$  nur von  $C$  und  $g_2$  nur von  $D$  abhängt. Die Beibehaltung der gegenseitigen Abhängigkeiten führt auf andere Variationsgleichungen und andere Terme für den Gradienten, den wir im folgenden Abschnitt angeben wollen. Den allgemeinen Fall können wir in einem späteren Ausblick andiskutieren. Mit dieser vereinfachenden Annahme schreibt sich unser Gleichungssystem nun in einer symmetrischen Form

$$\begin{aligned} \frac{\delta \mathcal{L}_1}{\delta C} \Big|_{\bar{D}} - \epsilon_1 \frac{\delta g_1}{\delta C} &\doteq 0 & ; & \quad g_1 = g_1[C] \\ \frac{\delta \mathcal{L}_2}{\delta D} \Big|_{\bar{C}} - \epsilon_2 \frac{\delta g_2}{\delta D} &\doteq 0 & ; & \quad g_2 = g_2[D] \quad . \end{aligned} \quad (4.29)$$

## 4.2.2 2.Schritt: Differentiation

Nachdem wir die Variationsgleichungen angegeben haben, können wir nun zur Differentiation unseres Funktional übergehen. Im Abschnitt (4.1.4) haben wir gesehen, daß wir Zusatzterme zum Gradienten bekommen, wenn die Nebenbedingung parameterabhängig ist. Setzen wir die Lösungen beider Variationen ins ursprüngliche Funktional ein, erhalten wir die zu differenzierende Funktion <sup>2</sup>

$$\bar{\bar{\mathcal{L}}}(\lambda) = \mathcal{L}_1(\bar{C}, \bar{D}; \lambda) + \mathcal{L}_2(\bar{C}, \bar{D}; \lambda) = \bar{\bar{\mathcal{L}}}_1(\lambda) + \bar{\bar{\mathcal{L}}}_2(\lambda) \quad (4.30)$$

---

<sup>2</sup>Das stimmt nicht exakt, denn wir haben ja bei der Variation bzgl.  $C$  in (4.18)  $\mathcal{L}_2$  vernachlässigt. Die folgende Funktion  $\bar{\bar{\mathcal{L}}}(\lambda)$  unterscheidet sich von der wahren Funktion durch die Unterlassung der Variation von  $\mathcal{L}_2$  bzgl.  $C$ .

mit  $\bar{D} = \tilde{D}(\bar{C})$ . Analog der Vorgehensweise bei der Herleitung der Variationsgleichungen (4.18) vernachlässigen wir auch hier den Term  $\bar{\bar{\mathcal{L}}}_2$ . Die von uns näherungsweise berechnete Energie ist damit  $\bar{\bar{\mathcal{L}}} \approx \bar{\bar{\mathcal{L}}}_1$ , deren Ableitung wir nun bilden wollen:

$$\frac{d\bar{\bar{\mathcal{L}}}_1(\lambda)}{d\lambda} = \left. \frac{\partial \mathcal{L}_1(\lambda)}{\partial \lambda} \right|_{\bar{D}, \bar{C}} + \left. \frac{\delta \mathcal{L}_1}{\delta C} \right|_{\bar{D}, \bar{C}} \frac{\partial \bar{C}}{\partial \lambda} + \left. \frac{\delta \mathcal{L}_1}{\delta D} \right|_{\bar{D}, \bar{C}} \frac{\partial \bar{D}}{\partial \lambda} . \quad (4.31)$$

Im letzten Summanden ersetzen wir  $\left. \frac{\delta \mathcal{L}_1}{\delta D} \right|_{\bar{D}, \bar{C}}$  gemäß (4.15) durch  $-\frac{\delta \mathcal{L}_2}{\delta D}$  und dies wiederum wegen (4.24) durch  $-\epsilon_2 \frac{\delta g_2}{\delta D}$ , sodaß wir nun

$$\frac{d\bar{\bar{\mathcal{L}}}_1(\lambda)}{d\lambda} = \left. \frac{\partial \mathcal{L}_1(\lambda)}{\partial \lambda} \right|_{\bar{D}, \bar{C}} + \left. \frac{\delta \mathcal{L}_1}{\delta C} \right|_{\bar{D}, \bar{C}} \frac{\partial \bar{C}}{\partial \lambda} - \epsilon_2 \left. \frac{\delta g_2}{\delta D} \right|_{\bar{D}, \bar{C}} \frac{\partial \bar{D}}{\partial \lambda} \quad (4.32)$$

erhalten. Den ersten Faktor des zweiten Summanden können wir wegen Formel (4.28) durch die Terme mit den Ableitungen der Nebenbedingungen  $g_1$  und  $g_2$  ersetzen. Das führt auf

$$\begin{aligned} \frac{d\bar{\bar{\mathcal{L}}}_1(\lambda)}{d\lambda} &= \left. \frac{\partial \mathcal{L}_1(\lambda)}{\partial \lambda} \right|_{\bar{D}, \bar{C}} \\ &+ \epsilon_1 \left( \left. \frac{\delta g_1}{\delta C} \right|_{\bar{D}} \frac{\partial \bar{C}}{\partial \lambda} + \left. \frac{\delta g_1}{\delta D} \right|_{\bar{C}} \frac{\partial \bar{D}}{\partial C} \frac{\partial \bar{C}}{\partial \lambda} \right) \\ &- \underbrace{\left. \frac{\delta g_2}{\delta D} \right|_{\bar{D}, \bar{C}} \frac{\partial \bar{D}}{\partial C} \frac{\partial \bar{C}}{\partial \lambda}}_{\epsilon_2 \left. \frac{\delta g_2}{\delta D} \right|_{\bar{D}, \bar{C}} \frac{\partial \bar{D}}{\partial \lambda}} - \epsilon_2 \left. \frac{\delta g_2}{\delta D} \right|_{\bar{D}, \bar{C}} \frac{\partial \bar{D}}{\partial \lambda} . \end{aligned} \quad (4.33)$$

Die beiden Ausdrücke der letzten Zeile lassen sich zusammenfassen. Den zweiten Summanden können wir wieder umschreiben. Da wir aus Abschnitt (4.1.4) wissen, daß an der Lösung  $(\bar{C}, \bar{D})$  die Nebenbedingung identisch verschwindet:  $\bar{\bar{g}}_1(\lambda) := g_1[\bar{C}(\lambda), \bar{D}(\bar{C}); \lambda] \equiv 0$ , ist auch deren Ableitung nach  $\lambda$  Null;

$$\frac{d\bar{\bar{g}}_1}{d\lambda} = \left. \frac{\delta g_1}{\delta C} \right|_{\bar{D}} \frac{\partial \bar{C}}{\partial \lambda} + \left. \frac{\delta g_1}{\delta D} \right|_{\bar{C}} \frac{\partial \bar{D}}{\partial C} \frac{\partial \bar{C}}{\partial \lambda} + \left. \frac{\partial g_1}{\partial \lambda} \right|_{\bar{C}, \bar{D}} \equiv 0 . \quad (4.34)$$

Die ersten beiden Summanden können wir hier durch den letzten ersetzen und diesen in Gleichung (4.33) für den dortigen zweiten Summanden einsetzen. Das liefert uns

$$\frac{d\bar{\bar{\mathcal{L}}}_1(\lambda)}{d\lambda} = \left. \frac{\partial \mathcal{L}_1(\lambda)}{\partial \lambda} \right|_{\bar{D}, \bar{C}} - \epsilon_1 \left. \frac{\partial g_1}{\partial \lambda} \right|_{\bar{C}, \bar{D}} - \epsilon_2 \left. \frac{\delta g_2}{\delta D} \right|_{\bar{D}, \bar{C}} \frac{\partial \bar{D}}{\partial \lambda} - \epsilon_2 \left. \frac{\delta g_2}{\delta D} \right|_{\bar{D}, \bar{C}} \frac{\partial \bar{D}}{\partial \lambda} . \quad (4.35)$$

Durch dieses Umschreiben der Ableitungen tritt nun mit dem zweiten Summanden neben dem ersten von Gleichung (4.33) ein weiterer Term mit direkter  $\lambda$  Ableitung auf. Wir versuchen nun, auch die letzten beiden Summanden von (4.35) zugunsten der partiellen Ableitung der zweiten Nebenbedingungen bezüglich  $\lambda$  umzuformulieren. Wie bei der ersten Nebenbedingungen gilt auch hier, daß sie an der Lösung  $\bar{C}, \bar{D}$  identisch verschwindet,  $\bar{g}_2(\lambda) := g_2[\bar{C}(\lambda), \bar{D}(\bar{C}); \lambda] \equiv 0$ , sodaß

$$\frac{d\bar{g}_2}{d\lambda} = \left. \frac{\delta g_2}{\delta D} \right|_{\bar{C}} \frac{\partial \bar{D}}{\partial \lambda} + \left. \frac{\delta g_2}{\delta C} \right|_{\bar{D}} \frac{\partial \bar{C}}{\partial \lambda} + \left. \frac{\partial g_2}{\partial \lambda} \right|_{\bar{C}, \bar{D}} \equiv 0 \quad . \quad (4.36)$$

Der letzte Summand  $\epsilon_2 \frac{\delta g_2}{\delta D} \frac{\partial \bar{D}}{\partial \lambda}$  von (4.35) kann jetzt mit der eben aufgeschriebenen Beziehung durch die partielle Ableitung der Nebenbedingung  $g_2$  ersetzt werden und wir erhalten

$$\frac{d\bar{\mathcal{L}}_1(\lambda)}{d\lambda} = \left. \frac{\partial \mathcal{L}_1(\lambda)}{\partial \lambda} \right|_{\bar{D}, \bar{C}} - \epsilon_1 \left. \frac{\partial g_1}{\partial \lambda} \right|_{\bar{C}, \bar{D}} + 2\epsilon_2 \left. \frac{\partial g_2}{\partial \lambda} \right|_{\bar{C}, \bar{D}} + 2\epsilon_2 \left. \frac{\delta g_2}{\delta C} \right|_{\bar{D}} \frac{\partial \bar{C}}{\partial \lambda} \quad . \quad (4.37)$$

Diese Beziehung gilt allgemein, wenn die Nebenbedingungen von beiden Variationsgrößen  $C$  und  $D$  abhängen. Diese gegenseitigen Abhängigkeiten hatten wir ja zur Vereinfachung unseres Variationsgleichungssystems (4.29) aufgegeben. Aus der letzten Formel (4.37) können wir folgern, daß es aufgrund der Beziehung (4.34) keinen Unterschied macht, ob die erste Nebenbedingungen von den Variationsgrößen  $D$  abhängt oder nicht da dieser gegebenenfalls auftretende Term immer zusammen mit  $\frac{\delta g_1}{\delta C} \frac{\partial \bar{C}}{\partial \lambda}$  in der partiellen  $\lambda$ -Ableitung der ersten Nebenbedingung absorbiert ist. Etwas anders hingegen gestaltet sich das bezüglich des letzten Terms von Gleichung (4.37). Er verschwindet nur dann, wenn wir verlangen, daß die zweite Nebenbedingung nicht von den Variationsgrößen  $C$  der ersten Variation abhängt. Diese Asymmetrie liegt in unserer Näherung begründet, die eine der beiden Variationen (und zwar die bezüglich  $C$ ) auszeichnet. Unter der Annahme, daß  $g_2$  nicht von  $C$  abhängt, folgt für den Energiegradient

$$\frac{d\bar{\mathcal{L}}_1(\lambda)}{d\lambda} = \left. \frac{\partial \mathcal{L}_1(\lambda)}{\partial \lambda} \right|_{\bar{D}, \bar{C}} - \epsilon_1 \left. \frac{\partial g_1}{\partial \lambda} \right|_{\bar{C}, \bar{D}} + 2\epsilon_2 \left. \frac{\partial g_2}{\partial \lambda} \right|_{\bar{C}, \bar{D}} \quad . \quad (4.38)$$

Der Bedeutung dieser Gleichung bleibt zunächst noch verborgen; sie wird sich erst am Ende dieses Kapitels offenbaren, wenn es darum geht, die Gradiententerme explizit zu berechnen. Folgende Tabelle soll die bisher benutzten Symbole mit den zuzuordnenden physikalischen Größen im Überblick festhalten. Wir wollen diese in unsere allgemeine Gradientenformel einsetzen und die hier gewonnenen Erkenntnisse beim Aufschreiben mit einarbeiten.

Symbol	physikalische Größe	Bedeutung
$\mathcal{C}$	$\psi_i$ oder $c_{ij}$	Wellenfunktionen oder Entwicklungskoeffizienten
$D$	$\tilde{n}$ oder $D_\sigma$	Modelldichte, Dichtefit-Koeffizienten
$\mathcal{L}$	$\mathcal{E}$	Energiefunktional
$\mathcal{L}_1$	$\mathcal{E}_1$	Energiefunktional der Variation bzgl. C
$\mathcal{L}_2$	$\mathcal{E}_2$	Energiefunktional der Variation bzgl. D
$g_1$	$\sum_{ij} (\langle \psi_i   \psi_j \rangle - \delta_{ij}) = 0$	1. Nebenbedingung; orthonormierte molekulare Wellenfunktionen
$g_2$	$N - \int \sum_j D_j k_j d\vec{r} = 0$	2. Nebenbedingung; N: Anzahl der Elektronen $k_j$ sind fest vorgegebene Funktionen
$\lambda$	$\vec{R}$ oder $\vec{R}_\alpha$ ; $R_{\alpha_j}, j = 1, 2, 3$	Kernanordnung oder speziell Koordinaten des $\alpha$ -ten Kernes bzw. $x$ -, $y$ - oder $z$ Komponente
$\overline{\mathcal{L}}$	$E_{tot}$	totale Energie
$\overline{\mathcal{L}}_1$	$E$	näherungsweise berechnete totale Energie
$\overline{\mathcal{L}}_2$	$E_{rest}$	Restenergiefunktion, wird vernachlässigt

**Tabelle 4.2:** Zuordnung der in diesem Abschnitt verwendeten Variablen zu den physikalischen Größen, die uns im folgenden Abschnitt begegnen werden.

Zusammenfassung:

Die Berechnung des Energiegradienten bezüglich eines externen Parameters können wir in zwei Schritten abarbeiten. Dazu ist zuerst das Variationsproblem für die totale Energie zu lösen, was die Minimierung eines Energiefunktionals durch Bestimmung der Variationsgrößen (hier Wellenfunktionen, Dichte) beinhaltet. Aus rechentechnischen Gründen müssen wir diese Variation abändern. Daraus resultiert ein weiteres zweites Variationsproblem und erst die simultane Lösung beider Variationsprobleme macht aus dem Energiefunktional eine parameterabhängige Funktion, die totale Energie. Der Gradient ergibt sich sodann durch Differentiation nach dem Parameter. Weil die Variationen zusätzlich unter Nebenbedingungen erfolgen, die ebenfalls parameterabhängig sind, ergeben sich dadurch Zusatzterme zum Gradienten.

### 4.3 Die Gradiententerme explizit

Im letzten Abschnitt haben wir die Variationsgleichungen abgeleitet. In diesem Abschnitt nun erfolgt der zweite Schritt, die Differentiation. Wie wir dabei prinzipiell vorgehen

müssen, haben wir im vorherigen Abschnitt ausführlich dargelegt. Jetzt wollen wir physikalischen Größen in unsere Formel (4.38) konkret einsetzen und jeden der drei Summanden detailliert angeben. Daher leiten wir nochmals die totale Energie her und geben insbesondere deren Geometrieabhängigkeit an. Dann können wir den ersten Term aus (4.38) aufschreiben. Anschließend diskutieren wir die abzuleitenden Nebenbedingungen und fassen am Schluß zusammen.

### 4.3.1 Variationsgrößen und Geometrieabhängigkeit

Die allgemeinen Variationsgrößen <sup>3</sup>  $C$  unserer formalen Herleitung sind die Entwicklungskoeffizienten  $c_{i\mu}$ , die über den LCAO-Ansatz die molekularen Wellenfunktionen  $\psi_i$  mit atomaren Wellenfunktionen  $\xi_\mu$  verknüpfen,

$$\psi_i^{MO}(\vec{r}) = \sum_{\nu} c_{i\nu} \xi_{\nu}^{AO}(\vec{r}) \quad . \quad (4.39)$$

Die Wellenfunktionen  $\xi_{\nu}^{AO}$  sind in unserem Falle Lösungen der atomaren Diracgleichung und an den Atomen des Moleküls oder Clusters zentriert. Dadurch werden die molekularen Wellenfunktionen geometrieabhängig. Aus (4.39) wird, wenn wir eine Basisentwicklung um die atomaren Zentren vornehmen,

$$\psi_i^{MO}(\vec{r}) = \sum_{\nu=1}^m c_{i\nu} \xi_{\nu}^{\alpha}(\vec{r} - \vec{R}_{\alpha[\nu]}) = \psi_i^{MO}(\vec{r}; C, \vec{R}_{\alpha[\nu]}) \quad . \quad (4.40)$$

Dabei sei  $\alpha[\nu]$  dasjenige atomare Zentrum  $\alpha$ , an welchem die Basisfunktionen  $\nu$  dieses Atoms zentriert sind. Die Summe über  $\nu$  erstreckt sich über  $m$  Stück (im Prinzip über alle AO's). Die molekulare Dichte ist jetzt geometrieabhängig:

$$\begin{aligned} n(\vec{r}) &= \sum_i^N \psi_i^{\dagger}(\vec{r}, C, \vec{R}) \psi_i(\vec{r}, C, \vec{R}) \\ &= n[\psi_i[\vec{r}, C, \vec{R}]] \quad . \end{aligned} \quad (4.41)$$

wobei  $\vec{R}$  hier für die Koordinaten aller atomaren Zentren steht.  $N$  ist die Anzahl der besetzten Orbitale.

---

<sup>3</sup>Hier steht  $C$  abkürzend für einen Koeffizientenvektor, die Variation bezüglich  $C$  bezieht sich also immer auf die einzelnen Komponenten, siehe auch Bemerkungen am Ende des Abschnitts (4.1.1).

Der zweite Satz von Variationsgrößen  $D$  sind die aus unserem speziellen Dichteansatz des Abschnitts (3.5) erhaltenen Koeffizienten, die die Modelldichte, die wir bereits in (3.17) angegeben hatten, der exakten Dichte  $n$  möglichst gut annähern sollen:

$$\tilde{n}(\vec{r}) = \sum_{\sigma} D_{\sigma} k_{\sigma}(\vec{r}, \vec{R}) = \tilde{n}[D, k_{\sigma}[\vec{r}, \vec{R}]] \quad . \quad (4.42)$$

Die Funktionen  $k_{\sigma}$  enthalten die Summen der Betragsquadrate der radialen atomaren Wellenfunktionen. Dadurch wird  $\tilde{n}$  ebenfalls geometrieabhängig.

### 4.3.2 Geometrieabhängigkeit der totalen Energie

Im Dichtefunktionalformalismus wird die totale Energie  $E_{tot}$  als ein Funktional der Dichte  $n$  betrachtet,  $E_{tot} = \mathcal{E}[n]$ . Aus der Variation von  $\mathcal{E}$  bezüglich der Dichte folgt eine spezielle Dichte  $n = \bar{n}$ , die  $\mathcal{E}$  stationär macht. Das ist im Prinzip die schon von uns im Kapitel (2.2) betrachtete Vorgehensweise im Sinne des Hohenberg-Kohn Theorems. Dort ließ sich die Dichte über Hilfs-Einteilchenorbitale  $\psi_i$  ausdrücken. Nichtrelativistisch kann dann die kinetische Energie als Operator der zweiten Ableitung wirkend auf  $\psi_i$  geschrieben werden. Weil die Dichte sich aus der Summe von Betragsquadraten der Wellenfunktionen ergibt, ist der Zusammenhang zwischen Dichte und Wellenfunktionen eindeutig. Potentielle Energien werden weiterhin als Funktional der Dichte behandelt. Das Funktional der kinetischen Energie ist durch

$$\begin{aligned} \mathcal{T} &= \sum_{i=1}^N \langle \psi_i^{MO} | \hat{t} | \psi_i^{MO} \rangle \\ &= \sum_{i=1}^N \sum_{\mu\nu} c_{i\mu}^* c_{i\nu} \langle \xi_{\mu}^{AO} | \hat{t} | \xi_{\nu}^{AO} \rangle \end{aligned} \quad (4.43)$$

gegeben, wenn wir hier der Einfachheit halber den Zwischenschritt über die Symmetrie-orbitalbasiszerlegung übergehen.

In diesem Sinne ist die Abhängigkeit des Energiefunktionals  $\mathcal{E}$  von den Variationsgrößen  $C$  über die Dichte und die Wellenfunktionen gegeben. Über die Ansätze (4.40) und (4.42) wird  $\mathcal{E}$  geometrieabhängig:

$$\begin{aligned} \mathcal{E} &= \mathcal{E}[\psi_i[C, \vec{R}], n[\psi_i[C, \vec{R}]]; \vec{R}] \\ &= \mathcal{T}[\psi_i[\vec{r}, \vec{R}]] + \mathcal{W}[n[\psi_i[\vec{r}, \vec{R}]], \vec{R}] \\ &= \mathcal{T}[\psi_i[\vec{r}, \vec{R}]] + \mathcal{W}^h[n[\psi_i[C, \vec{R}]]] + \mathcal{W}^{xc}[n[\psi_i[C, \vec{R}]]] + \mathcal{W}^{ext}[n[\psi_i[C, \vec{R}]]; \vec{R}] \quad . \end{aligned} \quad (4.44)$$

$\mathcal{T}$  ist das Funktional der kinetischen und  $\mathcal{W}$  das der potentiellen Energie. Letzteres setzt sich aus mehreren Anteilen zusammen; dem Funktional  $\mathcal{W}^h$  der Coulombenergie der

Elektronen untereinander, dem Austausch-Korrelationsfunktional  $\mathcal{W}^{xc}$  und dem durch die Atomkerne gegebenen externen Funktional  $\mathcal{W}^{ext}$ , welches von den Kernkoordinaten  $\vec{R}$  über die Kern-Kern Abstoßungsenergie und das Kernpotential

$$v^{ext}(\vec{r}) = - \sum_{\alpha}^{\gamma} \frac{Z_{\alpha}}{|\vec{r} - \vec{R}_{\alpha}|} \quad (4.45)$$

abhängt. Die durch Variation von  $\mathcal{E}$  erhaltene totale Energie  $E_{tot}$  als Funktion der geometrischen Anordnung der Kerne schreibt sich dann

$$E = \left( \mathcal{T} + \mathcal{W}^h + \mathcal{W}^{xc} + \mathcal{W}^{ext} \right) \Big|_{\bar{C}} \quad (4.46)$$

$$\begin{aligned} &= \sum_i \langle \psi_i | t | \psi_i \rangle + \int n(\vec{r}) v(\vec{r}') d\vec{r}' + \int v^{xc} dn + \underbrace{\int n v^{ext} d\vec{r} + \sum_{\alpha \neq \beta} \frac{Z_{\alpha} Z_{\beta}}{|\vec{R}_{\alpha} - \vec{R}_{\beta}|}}_{E^{ext}} \\ &= T + E^h + E^{xc} + E^{ext}. \end{aligned} \quad (4.47)$$

Der hier einzige explizit von der Geometrie abhängige Anteil ist im externen Potential enthalten. Bedingt durch unsere Näherung des elektronischen Coulombpotentials zerlegen wir nun das Funktional  $\mathcal{W}^h$  gemäß (4.14) in zwei Anteile:

$$\mathcal{W}^h[n[\psi_i[C, \vec{R}]]] = \mathcal{W}_1^h[n[\psi_i[C, \vec{R}]], \tilde{n}[D, \vec{R}]] + \mathcal{W}_2^h[n[\psi_i[C, \vec{R}]], \tilde{n}[D, \vec{R}]] \quad (4.48)$$

Der Ansatz (3.14),  $n(\vec{r}) = \tilde{n}(\vec{r}) + \Delta n(\vec{r})$  liefert dann <sup>4</sup>

$$\begin{aligned} \mathcal{W}^h[n] &= \frac{1}{2} \int \int \frac{(\tilde{n}(\vec{r}) + \Delta n(\vec{r})) (\tilde{n}(\vec{r}') + \Delta n(\vec{r}'))}{|\vec{r} - \vec{r}'|} d\vec{r} d\vec{r}' \quad (4.49) \\ &= \frac{1}{2} \int \int \left( \frac{\tilde{n}(\vec{r}) \tilde{n}(\vec{r}')}{|\vec{r} - \vec{r}'|} + 2 \frac{\tilde{n}(\vec{r}) \Delta n(\vec{r}')}{|\vec{r} - \vec{r}'|} - 2 \frac{\tilde{n}(\vec{r}') \Delta n(\vec{r})}{|\vec{r} - \vec{r}'|} + \frac{\Delta n(\vec{r}) \Delta n(\vec{r}')}{|\vec{r} - \vec{r}'|} \right) d\vec{r} d\vec{r}' \\ &= \underbrace{\int \int \frac{\tilde{n}(\vec{r}) \tilde{n}(\vec{r}')}{|\vec{r} - \vec{r}'|} d\vec{r} d\vec{r}' - \frac{1}{2} \int \int \frac{\tilde{n}(\vec{r}) \tilde{n}(\vec{r}')}{|\vec{r} - \vec{r}'|} d\vec{r} d\vec{r}'}_{=: \mathcal{W}_1^h[n, \tilde{n}]} + \underbrace{\frac{1}{2} \int \int \frac{\Delta n(\vec{r}) \Delta n(\vec{r}')}{|\vec{r} - \vec{r}'|} d\vec{r} d\vec{r}'}_{=: \mathcal{W}_2^h[n, \tilde{n}]} \end{aligned}$$

$\mathcal{W}_2^h[n, \tilde{n}]$  ist das Funktional der Selbstenergie der positiv definiten Differenzladungsdichte  $(n - \tilde{n})$  und führt auf die von uns vernachlässigte Restenergie  $E_{rest}$  (siehe Formel (3.15) aus Abschnitt (3.5)). Für das Gesamtfunktional  $\mathcal{E}$  folgt damit

$$\mathcal{E} = \mathcal{E}_1[\psi_i[C, \vec{R}], n[\psi_i[C, \vec{R}]], \tilde{n}[D, \vec{R}]; \vec{R}] + \mathcal{E}_2[n[\psi_i[C, \vec{R}]], \tilde{n}[D, \vec{R}]] \quad (4.50)$$

<sup>4</sup>Beim Ausmultiplizieren benutzen wir  $\tilde{n} \cdot \Delta n = \tilde{n}(n - \tilde{n})$ .

Das entspricht der Aufteilung unseres ursprünglich nur von einem Satz von Variationsgrößen abhängigen Funktionals in zwei Anteile (siehe auch Gleichung (4.14)). Diese sind durch  $\mathcal{E}_1 = \mathcal{T} + \mathcal{W}_1^h + \mathcal{W}^{xc} + \mathcal{W}^{ext}$  und  $\mathcal{E}_2 = \mathcal{W}_2^h$  gegeben. Setzen wir nun unsere Ansätze (4.40) und (4.42) in die Nebenbedingungen  $g_1$  und  $g_2$  aus Tabelle (4.2.2) ein. Die erste lautet

$$g_1 = \sum_{ij} \underbrace{\left( \langle \psi_i | \psi_j \rangle - \delta_{ij} \right)}_{g_1^{ij}} \equiv 0 \quad (4.51)$$

und umfaßt im eigentlichen Sinne  $N^2$  Terme, die aber aufgrund der Orthogonalitätsbedingung der  $\psi_i$  auf  $N$  reduziert sind;  $g_1^{ij} \rightarrow g_1^{ii} = g_1^i$ . Setzen wir hier unseren LCAO-Ansatz ein, folgt

$$g_1^i = \sum_{\mu\nu} c_{i\mu}^* c_{i\nu} \underbrace{\langle \xi_\mu(\vec{r}, \vec{R}) | \xi_\nu(\vec{r}, \vec{R}) \rangle}_{S_{\mu\nu}(\vec{r}, \vec{R})} - 1 = g_1^i(C, \vec{R}) \quad (4.52)$$

Die Komponenten  $g_1^i$  und folglich  $g_1$  sind jetzt geometrieabhängig; in kompakterer Form  $g_1 = g_1[\psi_i[C, \vec{R}]]$ . Die zweite Nebenbedingung hatten wir bereits im Kapitel (3.5) notiert. Sie ist durch

$$g_2 = \int \tilde{n}(\vec{r}) d\vec{r} - N \quad (4.53)$$

gegeben, wobei  $\tilde{n}(\vec{r})$  unsere Hilfsdichte ist. Für diese setzten wir entsprechend (4.42) ein und bekommen

$$g_2 = \int \left( \sum_{\sigma} D_{\sigma} k_{\sigma}(\vec{r}, \vec{R}) \right) d\vec{r} - N = g_2(D_{\sigma}, k_{\sigma}, \vec{R}) \quad . \quad (4.54)$$

$N$  ist wiederum die Gesamtzahl aller Elektronen des Moleküls,  $D_{\sigma}$  sind die durch die zweite Variation (4.24) bestimmten Fitkoeffizienten;  $k_{\sigma}$  beinhaltet über (3.17) die Betragsquadrate der atomaren Wellenfunktionen. Speziell für  $k_{\sigma}(\vec{R}_{\alpha})$  ist das die Dichte, gebildet aus der Summe der Betragsquadrate der großen und kleinen Komponente aus (2.30) des  $\sigma$ -ten Orbitals, welche am  $\alpha$ -ten Kern zentriert sind. Die Geometrieabhängigkeit, wieder kompakt geschrieben, ist  $g_2 = g_2[D_{\sigma}, k_{\sigma}[\vec{R}]]$ . Unsere Energiegradientenformel (4.38) läßt sich nun direkt 'in die Physik übersetzen'. Übertragen auf die physikalischen Größen schreibt sie sich

$$\begin{aligned} \frac{d\bar{\mathcal{L}}_1(\lambda)}{d\lambda} &= \left. \frac{\partial \mathcal{L}_1(\lambda)}{\partial \lambda} \right|_{\bar{D}, \bar{C}} - \epsilon_1 \left. \frac{\partial g_1}{\partial \lambda} \right|_{\bar{C}, \bar{D}} + 2\epsilon_2 \left. \frac{\partial g_2}{\partial \lambda} \right|_{\bar{C}, \bar{D}} \\ &\Downarrow \\ \frac{dE(\vec{R})}{d\vec{R}} &= \left. \frac{\partial \mathcal{E}_1(\vec{R})}{\partial \vec{R}} \right|_{\bar{D}, \bar{C}} - \epsilon_1 \left. \frac{\partial g_1}{\partial \vec{R}} \right|_{\bar{C}, \bar{D}} + 2\epsilon_2 \left. \frac{\partial g_2}{\partial \vec{R}} \right|_{\bar{C}, \bar{D}} . \end{aligned} \quad (4.55)$$

Anhand der Gleichung (4.55) wollen wir nun die einzelnen Gradiententerme diskutieren. Der erste Summand der rechten Seite lautet ausführlicher

$$\left. \frac{\partial \mathcal{E}_1}{\partial \vec{R}} \right|_{\bar{D}, \bar{C}} = \underbrace{\sum_i \left. \frac{\partial \mathcal{E}_1}{\partial \psi_i} \frac{\partial \psi_i}{\partial \vec{R}} \right|_{\bar{C}, \bar{D}}}_{\text{Typ1}} + \underbrace{\sum_i \left. \frac{\partial \mathcal{E}_1}{\partial n} \frac{\partial n}{\partial \psi_i} \frac{\partial \psi_i}{\partial \vec{R}} \right|_{\bar{C}, \bar{D}}}_{\text{Typ2}} + \underbrace{\sum_\sigma \left. \frac{\partial \mathcal{E}_1}{\partial \tilde{n}} \frac{\partial \tilde{n}}{\partial k_\sigma} \frac{\partial k_\sigma}{\partial \vec{R}} \right|_{\bar{C}, \bar{D}}}_{\text{Typ3}} + \underbrace{\left. \frac{\partial E}{\partial \vec{R}} \right|_{\bar{\psi}, \bar{n}, \bar{\tilde{n}}}}_{\text{Typ4}} . \quad (4.56)$$

Machen wir uns diese Gleichung klar. Der letzte Summand (Typ 4) enthält eine explizite  $R$ -Abhängigkeit<sup>5</sup> über das äußere Kernpotential. Er tritt also immer auf, unabhängig von der Art und Weise, wie die totale Energie berechnet wird. Es ist dies der sogenannte Hellmann-Feynman Term, den wir im Abschnitt (4.3.4) ausführlich angeben. Alle anderen Terme resultieren daraus, daß, bedingt durch unsere Ansätze für die Variationsgrößen, das Funktional  $\mathcal{E}$  zusätzlich geometrieabhängig sind ( $\psi = \psi(C, \vec{R})$ ,  $\tilde{n} = \tilde{n}(D, \vec{R})$ ).

Kommen wir nun zur ausführlichen Darstellung der einzelnen Terme. Die Abhängigkeiten entnehmen wir direkt der Gleichung (4.44). Alle nun folgenden Ableitungen der atomaren Wellenfunktionen nach  $\vec{R}$  beziehen sich auf diejenigen  $m$  Basisfunktionen, die an dem Atom zentriert sind, daß sich am Ort  $\vec{R}$  befindet. Die Summationsindizes  $\mu, \nu$  laufen dann über alle  $m$  Basisfunktionen dieses atomaren Zentrums (siehe nochmals (4.40)). Die Ableitungen der Wellenfunktionen geben wir hier bezüglich  $\psi^\dagger$  an. Die nicht aufgeschriebenen konjugiert komplexen Beiträge berücksichtigen wir aber später.

### Typ1

Dieser Term besteht aus der Ableitung der kinetischen Energie. Dessen Anteile sind

$$\begin{aligned} \frac{\partial \mathcal{E}_1}{\partial \psi_i^\dagger} &= \frac{\partial \mathcal{T}}{\partial \psi_i^\dagger} = \frac{\partial}{\partial \psi_i^\dagger} \left( \sum_j \langle \psi_j | \hat{t} | \psi_j \rangle \right) = \hat{t} | \psi_i \rangle \\ \frac{\partial \psi_i^\dagger}{\partial \vec{R}} &= \sum_\mu c_{i\mu}^* \frac{\partial}{\partial \vec{R}} \xi_\mu^\dagger(\vec{r} - \vec{R}) . \end{aligned} \quad (4.57)$$

---

<sup>5</sup> $\bar{\psi} = \psi(\bar{C})$ ,  $\bar{n} = n(\bar{C})$ ,  $\bar{\tilde{n}} = \tilde{n}(\bar{D})$

Wir erhalten somit

$$\sum_i \frac{\partial \mathcal{E}_1}{\partial \psi_i^\dagger} \frac{\partial \psi_i^\dagger}{\partial \vec{R}} = \sum_i \sum_{\mu\nu} c_{i\mu}^* c_{i\nu} \int \frac{\partial \xi_\mu^\dagger}{\partial \vec{R}} \hat{t} \xi_\nu d\vec{r} \quad . \quad (4.58)$$

### Typ2

Die Faktoren des mit *Typ2* bezeichneten Summanden lauten

$$\begin{aligned} \frac{\partial \mathcal{E}_1}{\partial n} &= \frac{\partial \mathcal{W}_1^h}{\partial n} + \frac{\partial \mathcal{W}_{xc}}{\partial n} + \frac{\partial \mathcal{W}^{ext}}{\partial n} \\ &= \int \frac{\tilde{n}(\vec{r}')}{|\vec{r} - \vec{r}'|} d\vec{r}' + \frac{\delta E^{xc}}{\delta n} - \sum_\alpha \frac{Z_\alpha}{|\vec{r} - \vec{R}_\alpha|} \quad ; \\ &= \tilde{v}^h(\vec{r}) + v^{xc}(\vec{r}) + v^{ext}(\vec{r}) \\ \frac{\partial n}{\partial \psi_i^\dagger} &= \frac{\partial}{\partial \psi_i^\dagger} \left( \sum_j \psi_j^\dagger \psi_j \right) = \psi_j \quad ; \\ \frac{\partial \psi_i^\dagger}{\partial \vec{R}} &= \sum_\mu c_{i\mu}^* \frac{\partial}{\partial \vec{R}} \xi_\mu^\dagger(\vec{r} - \vec{R}) \quad . \end{aligned} \quad (4.59)$$

Es folgt

$$\sum_i \frac{\partial \mathcal{E}_1}{\partial n} \frac{\partial n}{\partial \psi_i^\dagger} \frac{\partial \psi_i^\dagger}{\partial \vec{R}} = \sum_i \sum_{\mu\nu} c_{i\mu}^* c_{i\nu} \int \frac{\partial \xi_\mu^\dagger}{\partial \vec{R}} \left( v^{ext} + \tilde{v}^h + v^{xc} \right) \xi_\nu d\vec{r} \quad . \quad (4.60)$$

Dieser Ausdruck hat dieselbe Struktur wie (4.58); wir fassen beide zusammen:

$$\sum_i \left( \frac{\partial \mathcal{E}_1}{\partial \psi_i^\dagger} + \frac{\partial \mathcal{E}_1}{\partial n} \frac{\partial n}{\partial \psi_i^\dagger} \right) \frac{\partial \psi_i^\dagger}{\partial \vec{R}} = \sum_i \sum_{\mu\nu} c_{i\mu}^* c_{i\nu} \int \frac{\partial \xi_\mu^\dagger}{\partial \vec{R}} \left( t + \tilde{v}^h + v^{xc} + v^{ext} \right) \xi_\nu d\vec{r} \quad . \quad (4.61)$$

### Typ3

Als nächstes betrachten wir den dritten Summanden in (4.56), in welchem eine Ableitung der Modelldichte auftritt. Diese ist über die Fitkoeffizienten  $D_\sigma$  mit der zweiten Variation verweben. Die einzelnen Faktoren lauten

$$\begin{aligned}
\frac{\partial \mathcal{E}_1}{\partial \tilde{n}} &= \frac{\partial \mathcal{W}_1^h}{\partial \tilde{n}} = \int \frac{n}{|\vec{r} - \vec{r}'|} d\vec{r}' - \int \frac{\tilde{n}}{|\vec{r} - \vec{r}'|} d\vec{r}' \quad ; \\
\frac{\partial \tilde{n}}{\partial k_\sigma} &= D_\sigma \quad ; \\
\frac{\partial k_\sigma}{\partial \vec{R}} &= \nabla_{\vec{R}} k_\sigma \quad .
\end{aligned} \tag{4.62}$$

Die Terme in der ersten Zeile folgen unmittelbar aus (4.49). Wir bekommen:

$$\begin{aligned}
\sum_\sigma \frac{\partial E_1}{\partial \tilde{n}} \frac{\partial \tilde{n}}{\partial k_\sigma} \frac{\partial k_\sigma}{\partial \vec{R}} &= \sum_\sigma \int (n(\vec{r}) - \tilde{n}(\vec{r})) \left( \int \frac{1}{|\vec{r}' - \vec{r}|} \sum_\sigma D_\sigma \nabla_{\vec{R}} k_\sigma d\vec{r}' \right) d\vec{r} \\
&= \int (n(\vec{r}) - \tilde{n}(\vec{r})) \frac{\partial \tilde{v}^h}{\partial \vec{R}} d\vec{r} \quad .
\end{aligned} \tag{4.63}$$

Der Ausdruck  $\frac{\partial \tilde{v}^h}{\partial \vec{R}}$  ist die Ableitung des Modell-Hartreepotentials  $\tilde{v}^h(\vec{r})$ . Dieses ergibt sich durch Integration der Modelldichte zu

$$\tilde{v}^h(\vec{r}) = \int \frac{\tilde{n}(\vec{r}')}{|\vec{r} - \vec{r}'|} d\vec{r}' \quad . \tag{4.64}$$

#### Typ 4

Schließlich bleibt noch der letzte Summand. Dieser beinhaltet die explizite Abhängigkeit der totalen Energie in Form des externen Potentialoperators und der elektrostatischen abstandsabhängigen Kern-Kern Wechselwirkungsenergie  $E_{\alpha\beta}$

$$\frac{\partial E}{\partial \vec{R}} = \int n(\vec{r}) \frac{\partial v^{ext}}{\partial \vec{R}} d\vec{r} + \nabla_{\vec{R}} E_{\alpha\beta} \quad . \tag{4.65}$$

Im Anschluß an Gleichung (4.56) hatten wir darauf verwiesen, daß dieser Gradiententerm immer auftritt.

### 4.3.3 Geometrieabhängigkeit der Nebenbedingungen

Nachdem wir den ersten Summanden der Energiegradientenformel abgeleitet haben, bleiben noch die Terme mit den Ableitungen der Nebenbedingungen, denen wir uns nun zuwenden wollen. Der zweite Summand der Energiegradientenformel (4.38) beinhaltet die partielle Ableitung der ersten Nebenbedingung an der Lösung  $\bar{C}, \bar{D}$  unseres Variationsgleichungssystems (4.29). Die Übertragung des formal abgeleiteten Ausdrucks auf physikalische Größen ist

$$\frac{\partial g_1}{\partial \lambda} \Big|_{\bar{D}, \bar{C}} \longrightarrow \sum_i \frac{\partial g_1}{\partial \psi_i} \frac{\partial \psi_i}{\partial \vec{R}} \quad (4.66)$$

mit (wiederum Ableitung nach  $\psi_i^\dagger$ )

$$\frac{\partial g_1}{\partial \psi_i^\dagger} = \frac{\partial}{\partial \psi_i^\dagger} \left( \sum_k \langle \psi_k | \psi_k \rangle \right) = | \psi_i \rangle \quad . \quad (4.67)$$

Nehmen wir noch den konjugiert komplexen Anteil hinzu, bekommen wir

$$\begin{aligned} \sum_i \frac{\partial g_1}{\partial \psi_i^\dagger} \frac{\partial \psi_i^\dagger}{\partial \vec{R}} + c.c. &= \sum_i \left( \int \frac{\partial \psi_i^\dagger}{\partial \vec{R}} \psi_i d\vec{r} + \int \psi_i^\dagger \frac{\partial \psi_i}{\partial \vec{R}} d\vec{r} \right) \\ &= \sum_i \sum_{\mu\nu} c_{i\mu}^* c_{i\nu} \int \left( \frac{\partial \xi_\mu^\dagger}{\partial \vec{R}} \xi_\nu + \xi_\mu^\dagger \frac{\partial \xi_\nu}{\partial \vec{R}} \right) d\vec{r} \\ &= \sum_i \sum_{\mu\nu} c_{i\mu}^* c_{i\nu} \frac{\partial S_{\mu\nu}(\vec{R})}{\partial \vec{R}} \quad . \end{aligned} \quad (4.68)$$

$S_{\mu\nu}(\vec{R})$  ist die atomare Überlappmatrix. Zur Komplettierung des zweiten Gradiententermes schließt sich nun noch eine Multiplikation mit  $\epsilon_1$  an, was in der bisherigen Formulierung abkürzend für eine ganze Matrix mit Elementen  $\epsilon_1^{ij}$  stand. Durch eine kanonische Transformation [81] innerhalb des Raumes der besetzten Zustände wird sie auf Diagonalform gebracht (siehe auch Gleichung (2.27)), woraus  $N$  Komponenten  $\epsilon_1^1, \dots, \epsilon_1^N$  resultieren. Wir erhalten

$$\sum_i \epsilon_i \frac{\partial g_1}{\partial \psi_i^\dagger} \frac{\partial \psi_i^\dagger}{\partial \vec{R}} + c.c. = \sum_i \epsilon_i \sum_{\mu\nu} c_{i\mu}^* c_{i\nu} \int \left( \frac{\partial \xi_\mu^\dagger}{\partial \vec{R}} \xi_\nu + \xi_\mu^\dagger \frac{\partial \xi_\nu}{\partial \vec{R}} \right) d\vec{r} \quad (4.69)$$

Dieser Term hat dieselbe Struktur wie (4.61). Wir kommen auf beide Ausdrücke nochmal im Abschnitt (4.3.5) zurück.

Kommen wir zum letzten Gradiententerm von Gleichung (4.38). Er lautete

$$2\epsilon_2 \frac{\partial g_2}{\partial \lambda} \Big|_{\bar{C}, \bar{D}} . \quad (4.70)$$

Aus dem formal abgeleiteten Term wird

$$\frac{\partial g_2}{\partial \lambda} \Big|_{\bar{D}, \bar{C}} \longrightarrow \sum_{\sigma} \frac{\partial g_2}{\partial k_{\sigma}} \frac{\partial k_{\sigma}}{\partial \vec{R}} . \quad (4.71)$$

Wenn wir konkret einsetzen, liefert das

$$\begin{aligned} \sum_{\sigma} \frac{\partial g_2}{\partial k_{\sigma}} \frac{\partial k_{\sigma}}{\partial \vec{R}} &= \sum_{\sigma} D_{\sigma} \int \frac{\partial k_{\sigma}}{\partial \vec{R}} d\vec{r} \\ &= \frac{\partial}{\partial \vec{R}} \int \underbrace{\left( \sum_{\sigma} D_{\sigma} k_{\sigma} \right)}_{\approx N ; \text{ normiert}} d\vec{r} = 0 . \end{aligned} \quad (4.72)$$

Bilden wir beispielsweise die Ableitung bezüglich der Koordinaten des  $\alpha$ -ten Atoms, ist in der letzten Formel  $\sigma$  nur über alle Zustände zu summieren, deren Wellenfunktionen an diesem Atom zentriert sind. Das liefert aber gerade die Elektronenanzahl dieses Atoms, also eine konstante normierte Größe, deren Ableitung Null ist. Infolgedessen trägt dieser dritte Gradiententerm nicht bei.

#### 4.3.4 Hellmann-Feynman Gradient

Die in molekularen Systemen immer vorhandene elektrostatische Wechselwirkung der positiv geladenen Kerne untereinander und deren Wechselwirkung mit der durch alle Elektronen erzeugten negativen Ladungsverteilung wird auch als Hellmann-Feynman (HF) Kraft bezeichnet [69, 70]. Sie sich läßt aus der Differentiation der explizit von den Kernkoordinaten abhängigen Energieterme ableiten. Das sind das externe Potential und die Kern-Kern Wechselwirkungsenergie. Wir hatten diesen Term in (4.65) schon aufgeschrieben. Er lautet für die Ableitung bezüglich des  $\alpha$ -ten Kernes

$$\begin{aligned} \frac{dE}{d\vec{R}_{\alpha}} \Big|_{HF} &:= \frac{\partial E}{\partial \vec{R}_{\alpha}} \\ &= \int n(\vec{r}) \frac{\partial v^{ext}}{\partial \vec{R}_{\alpha}} d\vec{r} + \frac{\partial E_{\alpha\beta}}{\partial \vec{R}_{\alpha}} \\ &= \int \frac{n(\vec{r}) Z_{\alpha}}{|\vec{R}_{\alpha} - \vec{r}|^3} (\vec{R}_{\alpha} - \vec{r}) d^3r - \sum_{\beta \neq \alpha} \frac{Z_{\alpha} Z_{\beta}}{|\vec{R}_{\alpha} - \vec{R}_{\beta}|^3} (\vec{R}_{\alpha} - \vec{R}_{\beta}) . \end{aligned} \quad (4.73)$$

Für die weiteren Betrachtungen wollen diesen Term mit  $\mathcal{G}^{HF}$  bezeichnen,

$$\mathcal{G}^{HF} := \left. \frac{dE}{d\vec{R}_\alpha} \right|_{HF} . \quad (4.74)$$

### 4.3.5 Korrekturterme

Den Hellmann-Feynman Term können wir als den physikalischen Anteil des Gradienten ansehen. Wir haben aber gesehen, daß das nicht der einzige Gradiententerm ist. Die verbleibenden Terme sind Korrekturen. Wir können sie als ein Maß dafür betrachten, wie gut unser Ansatz zur Berechnung der Energie ist und wie gut die dabei gemachten Näherungen sind. Im Verlaufe der bisherigen Betrachtungen ist uns schon mehrmals die Möglichkeit begegnet, Terme von gleicher Struktur zusammenzufassen, wovon wir nun nochmals Gebrauch machen wollen.

#### Die Basissatzkorrektur

In der sogenannten Basissatzkorrektur (OBC) <sup>6</sup> können all die Terme zusammenfaßt werden, in denen Ableitungen der Wellenfunktionen  $\psi_i$  vorkommen. Wir hatten das bereits in Gleichung (4.61) mit den Anteilen der kinetischen und potentiellen Energie getan. Ein weiterer Term mit derselben Struktur kam in der Ableitung der ersten Nebenbedingung (4.69) vor. Für die Basissatzkorrektur folgt dann (wiederum für die Ableitung nach  $\vec{R}_\alpha$ )

$$\begin{aligned} \left. \frac{dE}{d\vec{R}_\alpha} \right|_{OBC} &:= \sum_i \left( \frac{\partial \mathcal{E}_1}{\partial \psi_i} + \frac{\partial \mathcal{E}_1}{\partial n} \frac{\partial n}{\partial \psi_i} - \epsilon_i \frac{\partial g_1}{\partial \psi_i} \right) \frac{\partial \psi_i}{\partial \vec{R}} + c.c. \\ &= \sum_i^N \sum_{\nu\mu} c_{i\nu}^* c_{i\mu} \int \frac{\partial \xi_\nu^\dagger}{\partial \vec{R}_\alpha} \left\{ t + v^{ext} + \tilde{v}^h + v^{xc} - \epsilon_i \right\} \xi_\mu d\vec{r} + c.c. \\ &= \sum_i^N \sum_{\nu\mu} c_{i\nu}^* c_{i\mu} \left\langle \frac{\partial \xi_\nu}{\partial \vec{R}_\alpha} \left| t + \underbrace{v^{ext} + \tilde{v}^h + v^{xc}}_{\tilde{v}^{RKS}} - \epsilon_i \right| \xi_\mu \right\rangle + c.c. \quad (4.75) \end{aligned}$$

Dieser Ausdruck berücksichtigt die Mitbewegung des Basissatzes, der am  $\alpha$ -ten Atom zentriert ist. Daher ist die Summation nur über all diejenigen Basisfunktionen  $\xi_\nu$  zu bilden, die zum  $\alpha$ -ten Atom gehören. In den praktischen Rechnungen können wir nur eine **endliche** Anzahl von Basisfunktionen mitnehmen. Es sei hier noch erwähnt, daß dieser Term im Falle einer **vollständigen Basis** verschwinden würde. Um das einzusehen, erinnern wir uns des Lösungsverfahrens der Kohn-Sham (RKS) Gleichungen. Diese lauteten

---

<sup>6</sup>Die Nomenklaturen der Korrekturterme sind aus [78] entnommen. Dort ist die Basissatzkorrektur als Orbital basis correction bezeichnet

$$(\hat{t} + v^{RKS}(\vec{r})) | \psi_i \rangle = \epsilon_i | \psi_i \rangle \quad . \quad (4.76)$$

Entwickeln wir die  $\psi_i$  nach atomaren Wellenfunktionen  $\xi_\nu$ , multiplizieren wiederum mit  $\langle \xi_\mu |$  und summieren über alle  $N$  besetzten Zustände, erhalten wir wieder die Matrixdarstellung (2.33) der RKS-Gleichungen in Form des Gleichungssystems  $\mathbf{HC} = \epsilon \mathbf{SC}$ , dessen Lösung die Energieeigenwerte  $\epsilon_i = \bar{\epsilon}_i$  und Eigenvektoren  $c_{i\nu} = \bar{c}_{i\nu}$  liefert. Mit diesen Lösungen <sup>7</sup> gilt

$$\sum_i^N \langle \psi_i | (\hat{t} + \tilde{v}^{RKS} - \epsilon_i) | \psi_i \rangle = 0 \quad , \quad (4.77)$$

oder, wenn wir unseren Ansatz (4.40) benutzen und den konjugiert komplexen Anteil hier explizit hinzufügen

$$\sum_i^N \sum_{\nu\mu} c_{i\nu}^* c_{i\mu} \langle \xi_\nu | \hat{t} + \tilde{v}^{RKS} - \epsilon_i | \xi_\mu \rangle + c.c. = 0 \quad . \quad (4.78)$$

Diese Gleichung hat dieselbe Struktur wie Gleichung (4.75) mit Ausnahme der dort vorkommenden Basisableitung. Nehmen wir an, wir haben einen vollständigen Basissatz vorliegen. Dann spannen die Basisvektoren einen vollständigen Hilbertraum auf und eine jede beliebige Funktion kann durch deren Linearkombination dargestellt werden, so auch die Funktion  $\frac{\partial \xi_\nu}{\partial \vec{R}_\alpha}$ :

$$\left| \frac{\partial \xi_\nu}{\partial \vec{R}_\alpha} \right\rangle = \sum_\kappa^\infty q_\kappa | \xi_\kappa \rangle \quad . \quad (4.79)$$

Wenn wir dies in Gleichung (4.75) einsetzen, haben wir eine identische Form wie Gleichung (4.78) und folglich verschwindet der Basiskorrekturterm. In den praktischen Rechnungen sind wir jedoch gezwungen, eine endliche Basisentwicklung zu machen, wobei dann im allgemeinen

$$\left| \frac{\partial \xi_\nu}{\partial \vec{R}_\alpha} \right\rangle \neq \sum_\kappa^m q_\kappa | \xi_\kappa \rangle \quad (4.80)$$

ist. Gleichung (4.75) läßt sich dann nicht in die Form (4.78) bringen. Der Basiskorrekturterm stellt somit ein Maß für die Vollständigkeit der Basis dar. Für die explizite

---

<sup>7</sup>Strenggenommen müßten wir unsere Lösungen gemäß unserer Vereinbarung wieder mit einem Querstrich versehen, z.B.  $\bar{\psi}_i = \sum_\nu \bar{c}_{i\nu} \xi_\nu$ , wir lassen ihn aber hier der Übersicht halber weg.

Berechnung der Ableitungen der atomaren Wellenfunktionen verweisen wir auf den Anhang (A.2).

$$\mathcal{G}^{OBC} := \left. \frac{dE}{d\vec{R}_\alpha} \right|_{OBC} . \quad (4.81)$$

### Die Fitkorrektur

Den dritten Gradiententyp hatten wir in (4.63) bereits notiert. In ihn gehen Terme ein, die mit unserer speziellen Fitprozedur zu tun haben, was durch das Auftauchen der Ableitung der Energie bezüglich der Modelldichte dort nochmals verdeutlicht ist. Entsprechend wird dieser Gradiententerm auch als (Dichte)Fitkorrektur (DFC) bezeichnet:

$$\left. \frac{dE}{d\vec{R}_\alpha} \right|_{DFC} := \int \left( n(\vec{r}) - \tilde{n}(\vec{r}) \right) \frac{\partial \tilde{v}^h}{\partial \vec{R}_\alpha} d\vec{r} . \quad (4.82)$$

Speziell für den  $\alpha$ -ten Kern bezieht sich die Ableitung wiederum nur auf den zu diesem Atom gehörenden Anteil des Potentials.

$$\mathcal{G}^{DFC} := \left. \frac{dE}{d\vec{R}_\alpha} \right|_{DFC} . \quad (4.83)$$

Soweit haben wir einen Ausdruck für den Energiegradienten hergeleitet. Das dafür etwas umständlich anmutende Konzept der Ausnutzung der Variationsgleichungen hat jedoch den entscheidenden Vorteil, daß dadurch eine vollnumerische Differentiation der molekularen Entwicklungskoeffizienten umgangen ist. Diese läßt sich durch die Ableitung der Nebenbedingungen ausdrücken. Für alle höheren Ableitungen der Energie nach den Kernkoordinaten, insbesondere für die zweiten, die im nächsten Kapitel noch eine wichtige Rolle spielen werden, ist dies nicht mehr möglich [72, 73, 74].

Zusammenfassung:

Durch die von uns benutzte Basissatzmethode zur Darstellung der molekularen Orbitale aus endlich vielen atomzentrierten Wellenfunktionen treten neben den expliziten Geometrieabhängigkeiten des externen Potentials weitere abstandsabhängigen Terme auf. Daraus resultieren Korrekturterme zum physikalischen Hellmann-Feynman Gradient, zu denen wegen der Ausnutzung der Variationsgleichungen auch die Nebenbedingungen beitragen. Bei spezieller Wahl können aber deren Ableitungen verschwinden. Der Energiegradient ist

$$\frac{dE}{d\vec{R}} = \mathcal{G}^{HF} + \mathcal{G}^{OBC} + \mathcal{G}^{DFC} .$$

## 4.4 Frozen-Core Näherung und Energiegradient

Bei chemischen Abständen werden die Bindungseigenschaften im Molekülverband im wesentlichen durch die Elektronen in den Valenzbereichen der Atome bestimmt. Weil es bei der Berechnung der Bindungsenergie nur auf die Differenz zwischen der totalen Energie des Moleküls und der Summe der totalen Energien der beteiligten Atome ankommt ( $E_b = E(AB) - (E(A) + E(B))$ ), heben sich die Core-Anteile, da sie zur Bindung nicht beitragen, gänzlich weg. Deshalb ist es gerechtfertigt, die Core-Elektronen bei der Berechnung der uns interessierenden spektroskopischen Parameter garnicht erst explizit mitzunehmen und nur die aktiv an der Bindung beteiligten äußeren Valenzelektronen zu berücksichtigen. Bei großen Molekülen und solchen mit schweren Elementen ist das ein vielfach angewendetes Verfahren [82, 83]. Entsprechende Überlegungen sind von Baştuğ innerhalb der Dirac-Fock-Slater Methode in unser Programm umgesetzt und angewendet worden [60].

Andererseits hat sich in dieser Arbeit gezeigt, daß die Gradientenberechnungen für den allgemeinsten Fall, wenn keine Molekülsymmetrie angenommen werden kann, den Aufwand hinsichtlich der Rechenzeit derart in die Höhe treiben, daß wir selbst schon bei kleinen Systemen an die Grenzen des Machbaren gestoßen sind (siehe Kapitel 5). Ein wesentlicher Einflußfaktor ist dabei die Anzahl der verwendeten Basisfunktionen oder genauer, die Anzahl der aktiven Elektronen bzw. Orbitale. Ein möglicher Ausweg aus diesem Dilemma könnte hier die Frozen-Core FC-Näherung zu sein. Dies motiviert uns, an dieser Stelle einige Überlegungen zur Integration dieser Option in die Geometrieoptimierung anzustellen, wofür der Energiegradient in der FC-Näherung formuliert werden muß. Diese sind längst nicht ausgereift und bedürfen einer genaueren Studie, falls ernsthaft damit begonnen werden sollte, dieses Projekt zu realisieren. Zu diesem Zweck muß auf die Grundgleichungen zu Beginn dieses Kapitels zurückgegriffen werden.

Bei der Frozen-Core Theorie gehen wir davon aus, daß der Hilbertraum in einen Core- und einen Valenzraum partitioniert werden kann,

$$\hat{\mathcal{H}} = \hat{\mathcal{H}}_{core} \oplus \hat{\mathcal{H}}_{valenz} \quad . \quad (4.84)$$

Die molekularen Wellenfunktionen  $\psi_i$  setzen sich aus Core- und Valenzorbitalen zusammen:

$$\psi_i = \begin{cases} \psi_{core} = \xi^{AO} & i \in Core \\ \psi_{val} & i \in Valenz \end{cases} \quad . \quad (4.85)$$

Daraus resultiert die Aufteilung der elektronischen Dichte  $n$  und folglich auch des Coulombpotentials in einen Core- und Valenzanteil:

$$\begin{aligned} n &= n_{core} + n_{val} \quad , \\ v^h &= v_{core}^h + v_{val}^h \quad . \end{aligned} \quad (4.86)$$

Mit diesem Ansatz können wir uns die Wechselwirkungen im Molekül in einen Core-Core (CC) einen Core-Valenz (CV) und einen Valenz-Valenz (VV) Beitrag aufgeteilt vorstellen. Das bedingt das Auftauchen von CC-, CV- und VV- Matricelementen in Fock- und Überlappmatrix. **Froze-Core Näherung** heißt nun, die **Core-Valenz Wechselwirkung unberücksichtigt** zu lassen. Dies wird durch Nullsetzen der Core-Valenz Matricelemente in der Fockmatrix erreicht. Es muß aber beachtet werden, daß die Valenzzustände vor ihrer Diagonalisierung (Gleichung (2.33)) zuerst auf den Core-zuständen orthogonalisiert werden müssen. Dies geschieht durch die Anwendung einer Core-Valenz Orthogonalisierungsmatrix. Ausführlich ist dies in [60], Kapitel 6 beschrieben. Dort ist auch die Herleitung der totalen Energie in der FC-Näherung gezeigt. Wir übernehmen hier die Endformel (Gleichung 6.25)<sup>8</sup> nach Abspaltung der reinen Coreanteile in einer leicht modifizierten Form, die die Austausch-Korrelationswechselwirkung in allgemeinerer Form beinhaltet:

$$\begin{aligned} \tilde{E}_{FC}^{tot} &= \sum_{val} \tilde{\epsilon}_{val} - \frac{1}{2} \int \tilde{n}_{val} \tilde{v}_{val}^h d\vec{r} + \sum_{\sigma \neq \sigma'} \int n_{core,\sigma} \left( v_{\sigma'}^{ext} + \frac{1}{2} \tilde{v}_{core,\sigma'}^h \right) d\vec{r} \quad (4.87) \\ &- \sum_{\sigma} \int n_{val,\sigma} v^{xc} d\vec{r} + E^{xc}[n] + E_{\alpha\beta} \quad . \end{aligned}$$

Der Index  $\sigma$  läuft hierbei über alle Atome,  $\tilde{n}_{val}$  und  $\tilde{v}_{val}$  sind Valenzdichte und Valenz-Hartreepotential, die wieder aus der Modelldichte berechnet sind. Der erste Term in Gleichung (4.87) ist der Valenzanteil der Einteilchenenergien

$$\tilde{\epsilon}_{val} = \langle \psi_{val} | t + v^{ext} + \tilde{v}_{core}^h + \tilde{v}_{val}^h + v^{xc} | \psi_{val} \rangle \quad . \quad (4.88)$$

Der zweite Summand stellt die Coulombwechselwirkung aller Valenzelektronen dar. Diese hat wie schon  $\sum \tilde{\epsilon}_{val}$  reinen Valenzcharakter. Der dritte Term beschreibt die Wechselwirkung der Elektronen des  $\sigma$ -ten Cores mit dem Kernpotential und dem Valenzanteil des elektronischen Coulombpotentials des  $\sigma'$ -ten Zentrums. Er kann in guter Näherung auch als Wechselwirkungsenergie einer effektiven Punktladung  $Q_{core,\sigma}$  des Atoms  $\sigma$  mit einem abgeschirmten Potential der Ladung  $(Z_{\sigma'} - \frac{1}{2}Q_{core,\sigma'})$  im Abstand  $R_{\sigma,\sigma'}$  aufgefaßt werden, sodaß wir dafür auch schreiben können

$$\sum_{\sigma \neq \sigma'} \int n_{core,\sigma} \left( v_{\sigma'}^{ext} + \frac{1}{2} \tilde{v}_{core,\sigma'}^h \right) d\vec{r} = \sum_{\sigma \neq \sigma'} \frac{Q_{core,\sigma} \left( Z_{\sigma'} - \frac{1}{2} Q_{core,\sigma'} \right)}{R_{\sigma,\sigma'}} \quad . \quad (4.89)$$

Der vorletzte Term von Gleichung (4.87) korrigiert den Austausch-Korrelationsbeitrag der Einteilchenenergien  $\tilde{\epsilon}_{val}$  und es bleibt somit noch die Austausch-Korrelationsenergie zu berechnen. Mit der Umformung (4.89) können wir für die totale Energie schreiben:

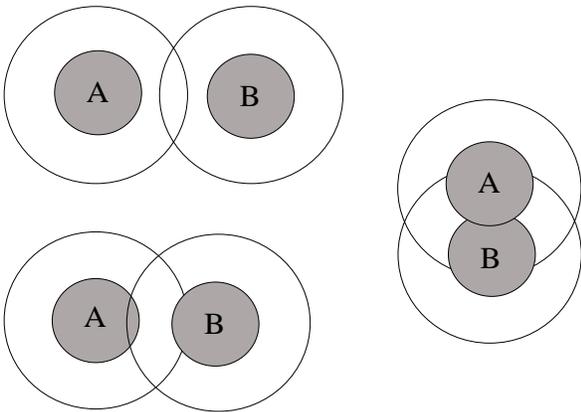
<sup>8</sup>Dort ist diese Gleichung im Rahmen des Slater'schen  $X_{\alpha}$ -Modells abgeleitet worden. Wir merken noch an, daß die in dem Term  $\frac{1}{4} \int \tilde{\rho} V^{ex} d^3r$  auftretende Dichte  $\tilde{\rho}$  (offenbar durch einen Tippfehler) durch  $\tilde{\rho}_{val}$  zu ersetzen ist.

$$\begin{aligned} \tilde{E}_{FC}^{tot} &= \sum_{val} \tilde{\epsilon}_{val} - \frac{1}{2} \int \tilde{n}_{val} \tilde{v}^h d\vec{r} + \sum_{\sigma \neq \sigma'} \frac{Q_{core,\sigma} \left( Z_{\sigma'} - \frac{1}{2} Q_{core,\sigma'} \right)}{R_{\sigma,\sigma'}} \\ &- \sum_{\sigma} \int n_{val,\sigma} v^{xc} d\vec{r} + E^{xc} + E_{\alpha\beta} . \end{aligned} \quad (4.90)$$

Gleichung (4.90) wollen wir nun hernehmen, um den Gradient abzuleiten. Erinnern wir uns dreier unterschiedlicher Typen von Gradiententermen, so fällt zunächst auf, daß wir uns durch die Umformung (4.89) einen zusätzlichen explizit von den Kernkoordinaten abhängigen Hellmann-Feynman-artigen Term eingehandelt haben. Der gesamte Hellmann-Feynman Gradient ist jetzt (bzgl. Atom  $\alpha$ ):

$$\begin{aligned} \mathcal{G}_{FC}^{HF} &= \int \frac{n_{val}(\vec{r}) Z_{\alpha}}{|\vec{r} - \vec{R}_{\alpha}|^3} (\vec{r} - \vec{R}_{\alpha}) d\vec{r} - \sum_{\beta \neq \alpha} \frac{Z_{\alpha} Z_{\beta}}{|\vec{R}_{\alpha} - \vec{R}_{\beta}|^3} (\vec{R}_{\alpha} - \vec{R}_{\beta}) \\ &+ \sum_{\beta \neq \alpha} \frac{Q_{core,\alpha} \left( Z_{\beta} - \frac{1}{2} Q_{core,\beta} \right)}{|\vec{R}_{\alpha} - \vec{R}_{\beta}|^3} (\vec{R}_{\alpha} - \vec{R}_{\beta}) . \end{aligned} \quad (4.91)$$

Schwieriger als im Allelektronen-Fall wird sich die Behandlung der Basissatzkorrektur gestalten. Als erster Schritt müssen nämlich wie schon die Valenzorbitale auch deren Ableitungen auf der Core-Basis orthogonalisiert werden. Dafür kann im Prinzip diesselbe Core-Valenz Orthogonalisierungsmatrix verwendet werden. Ein größeres Problem wird vermutlich die bei der Optimierung vorkommende Verschiebung der Atome darstellen.



**Abbildung 4.1:** Frozen-Core Näherung: Im oberen linken Bild überlappen nur die Valenzzustände der Atome A und B; darunter ragen die Valenzen in den Core (grau) des jeweils anderen Atoms. In der Mitte rechts überlappen sogar die Corezustände.

In Abbildung (4.1) sind einige denkbare Fälle angegeben. Im linken oberen Bild überlappen nur die Valenzen der Atome A und B. Die Corezustände (grau) der Atome A und B sind aufeinander orthogonal. Im Bild links unten überlappen die Valenzen mit den Corezuständen des jeweils anderen Atoms. Hier muß die Orthogonalisierungsprozedur

angewendet werden. Die Corezustände der Atome können auch hier in guter Näherung als orthogonal angenommen werden. Diese Situation entspricht der FC-Näherung, wie sie oben beschrieben ist. Anders ist dies, wenn sich die Abstände derart ändern, daß sogar die Corezustände einander überlappen. Dann muß eine Umdefinition der Core- und Valenzräume gemacht werden, mit anderen Worten, der Coreraum muß verkleinert und der Valenzraum vergrößert werden. Die Matrixelemente der Basissatzkorrektur haben eine analoge Struktur wie in der Allelektronen-Formulierung (vgl. (4.75)). Zu summieren ist hier aber über alle Beiträge der Valenzorbitale. Dadurch wird die Verschiebung der Valenzbasis simuliert:

$$\mathcal{G}_{FC}^{OBC} = \sum_{i \in \text{val}} \sum_{\nu}^{occ} \sum_{\mu}^{n(\alpha)} \sum_{\mu}^m c_{i\nu}^* c_{i\mu} \int \frac{\partial \xi_{\nu}^{\dagger}}{\partial \vec{R}_{\alpha}} \left\{ t + v^{ext} + \tilde{v}^h + v^{xc} - \epsilon_i \right\} \xi_{\mu} d\vec{r} + c.c. \quad (4.92)$$

Der dritte Gradiententerm beinhaltet die Ableitungen aller Energiekomponenten, in die Modelldichte  $\tilde{n}$  eingeht. In Gleichung (4.90) ist dies der zweite Summand. Ein weiterer Beitrag steckt in den Einteilchenenergien  $\tilde{\epsilon}_{val}$ :

$$[\tilde{\epsilon}_{val}]_{\tilde{n}} = \int n_{val} (\tilde{v}_{core}^h[\tilde{n}] + \tilde{v}_{val}^h[\tilde{n}]) d\vec{r} = \int n_{val}(\vec{r}) \tilde{v}^h(\vec{r}) d\vec{r} \quad . \quad (4.93)$$

Wenn wir die beiden Anteile zusammenfassen, folgt

$$\left[ \tilde{E}_{FC}^{tot} \right]_{\tilde{n}} = \int n_{val}(\vec{r}) \tilde{v}^h(\vec{r}) d\vec{r} - \frac{1}{2} \int \tilde{n}_{val}(\vec{r}) \tilde{v}_{val}^h(\vec{r}) d\vec{r} \quad . \quad (4.94)$$

Im ersten Summanden ist das volle, im zweiten nur das Valenz-Hartreepotential enthalten. Diese Asymmetrie kommt daher, daß der Term  $\frac{1}{2} \int \tilde{n}_{val} \tilde{v}_{core}^h d\vec{r}$  mit in den abstandsabhängigen Ausdruck (4.89) aufgenommen wurde. Bilden die Ableitung (z.B. wieder bezüglich  $\vec{R}_{\alpha}$ ), erhalten wir:

$$\begin{aligned} \mathcal{G}_{FC}^{DFC} &= \int n_{val} \frac{\partial \tilde{v}^h}{\partial \vec{R}_{\alpha}} d\vec{r} - \frac{1}{2} \int \left( \tilde{n}_{val} \frac{\partial \tilde{v}_{val}^h}{\partial \vec{R}_{\alpha}} + \frac{\partial \tilde{n}_{val}}{\partial \vec{R}_{\alpha}} \tilde{v}_{val}^h \right) d\vec{r} \quad (4.95) \\ &= \int n_{val} \frac{\partial \tilde{v}^h}{\partial \vec{R}_{\alpha}} d\vec{r} - \int \tilde{n}_{val} \frac{\partial \tilde{v}_{val}^h}{\partial \vec{R}_{\alpha}} d\vec{r} \quad . \end{aligned}$$

Der Gesamtgradient (bzgl. Atom  $\alpha$ ) setzt sich aus der Summe aller drei Gradiententypen zusammen:

$$\left. \frac{dE}{d\vec{R}_{\alpha}} \right|_{FC} = \mathcal{G}_{FC}^{HF} + \mathcal{G}_{FC}^{OBC} + \mathcal{G}_{FC}^{DFC} \quad . \quad (4.96)$$

Am Schluß dieses Abschnitts wollen wir zu bedenken geben, daß die Richtigkeit der Energiegradientenformel (4.96) nur durch eine nochmalige Herleitung, jetzt aber auf der formaleren Ebenen der Abschnitte (4.1.1) bis (4.2.2) überprüft werden kann. Hierfür sollte wieder bei dem ursprünglichen Funktional  $\mathcal{L}[C]$  begonnen und jeder Schritt in der Ableitung sehr sorgfältig formuliert werden. Die eigentliche FC-Näherung, also die Nullsetzung der Core-Valenz Wechselwirkung in der Fockmatrix führt dazu, daß der Ausdruck (4.87) für die totale Energie **nicht** mit der Lösung der Variationsgleichungen übereinstimmt. Eine mögliche Herangehensweise wäre eine zusätzliche Aufteilung des Funktionals  $\mathcal{L}[C]$  (4.14) in einen Core- und Valenzanteil:

$$\mathcal{L}[C] = \mathcal{L}_{core}[C] + \mathcal{L}_{val}[C] \quad . \quad (4.97)$$



# Kapitel 5

## Die Geometrieoptimierung

Moleküle können wir als eine Ansammlung von Atomen ansehen, die auf chemische Weise zusammenhalten. Atome bestehen aus einem positiv geladenen Kern und negativ geladenen Elektronen. Betrachten wir Moleküle als abgeschlossene Systeme, d.h. sind sie beispielsweise keinen äußeren elektrischen und magnetischen Feldern ausgesetzt, wird die Lage der Kerne allein durch die intramolekularen elektrostatischen Kräfte zwischen den Elektronen und Atomkernen bestimmt. Kerne und Elektronen ziehen einander an, während sich die Kerne und Elektronen untereinander jeweils abstoßen. Wir haben sowohl attraktive als auch abstoßende Kräfte und das Zusammenspiel beider Kräfte gibt Kernen und Elektronen Anlaß, sich in einer ganz bestimmten geometrischen Anordnung zu arrangieren. Die Gesamtheit aller dieser Kernanordnungen bildet die Energiehyperfläche. Diese ist eine komplizierte Funktion aller Kernkoordinaten. Wir können sie uns als ein Gebirge mit Bergen und Tälern vorstellen. Die Täler entsprechen dabei denjenigen Anordnungen, bei denen sich die elektrostatischen Kräfte gerade die Waage halten; in unserem konkreten Fall sind das die Minima der Energiefläche. Im allgemeinen gibt es deren mehrere, die als lokale Minima bezeichnet werden. Das energetisch tiefstliegende ist das globale Minimum. Lokale Minima sind durch Übergangszustände (Berge, Sattelpunkte) voneinander getrennt. Um von einem in ein anderes zu gelangen, ist entsprechend Energie aufzuwenden.

Die Geometrieoptimierung ist die Anwendung eines mathematischen Rechenkalküls aus der Optimierungstheorie. Dort ist die Zielstellung, eine Funktion, die im allgemeinen von mehreren Variablen abhängen kann, zu einem Extremum zu machen. In Abhängigkeit davon, ob die betrachtete Funktion noch Nebenbedingungen erfüllen soll, sind verschiedene Verfahren zu deren Minimierung möglich. Zur Funktionsminimierung mit Nebenbedingungen wird häufig das Simplexverfahren herangezogen. In unserem konkreten Falle werden für die Energieminimierung keine Nebenbedingungen benötigt, weswegen dieses Problem in die Klasse der sogenannten 'unconstrained optimization methods' gehört.

Eine weitere Untergliederung der Art der Minimierung resultiert aus der Menge der möglichen Informationen über die Funktion selbst, also konkret, ob nur die Funktion oder aber auch deren erste und Ableitungen höherer Ordnung bekannt sind. Je nach

dem Verhältnis, wie schnell Funktionswerte im Vergleich zu den Ableitungen berechnet werden, ist gegebenenfalls eine Variante der anderen vorzuziehen. Für die Geometrieoptimierung wurde von uns ein Verfahren gewählt, das sowohl die Funktion als auch deren erste Ableitung (Gradient) benötigt, den wir im letzten Kapitel behandelt haben. Diese Optimierungsprozedur soll hier nicht in aller Ausführlichkeit dargelegt werden; dafür sei auf Werke der numerischen Mathematik verwiesen [84, 85, 86]. Zum Verständnis geben wir aber die wesentlichsten Grundgedanken hier wieder.

In diesem Kapitel wollen wir im ersten Abschnitt die Minimumsuche darlegen. Gelegentlich werden wir hier der Begriffe Funktionswert und Energiewert nebeneinander benutzen. Gemeint ist natürlich damit stets die uns interessierende totale Energie, deren Minimum wir ja letztlich berechnen wollen. Danach gehen wir auf konvergenzbeschleunigende Aspekte ein und wollen hernach die Ergebnisse von Vorstudien und ersten Anwendungen präsentieren.

## 5.1 Die Beschreibung der Minimumsuche

Ausgangspunkt der Betrachtungen ist die Annahme, daß die überaus komplizierte Energiefläche lokal durch eine Fläche zweiter Ordnung approximiert werden kann. Die Entwicklung bis zum Term zweiter Ordnung lautet

$$E(\vec{\mathbf{X}}) = E(\vec{\mathbf{X}}_0) + \vec{\mathbf{g}}^T|_{\vec{\mathbf{x}}=\vec{\mathbf{x}}_0} (\vec{\mathbf{X}} - \vec{\mathbf{X}}_0) + \frac{1}{2} (\vec{\mathbf{X}} - \vec{\mathbf{X}}_0)^T \mathbf{B}|_{\vec{\mathbf{x}}=\vec{\mathbf{x}}_0} (\vec{\mathbf{X}} - \vec{\mathbf{X}}_0) + \dots \quad (5.1)$$

Hier bedeuten  $E(\vec{\mathbf{X}}_0)$  die Energie an einer vorgegebenen Kernanordnung  $\vec{\mathbf{X}}_0$ ,  $\vec{\mathbf{g}}|_{\vec{\mathbf{x}}=\vec{\mathbf{x}}_0}$  die erste Ableitung der Energie an dieser Stelle, deren Bestimmung Gegenstand des vorigen Kapitels war und  $\mathbf{B}$  die Hessematrix der zweiten Ableitungen in  $\vec{\mathbf{X}}_0$ . Sie enthält die Information über die Krümmung der Fläche an der Stelle  $\vec{\mathbf{X}}_0$ . Die unbekannte Energiefunktion, von welcher wir das Minimum suchen, wird somit lokal durch eine Modellfunktion zweiter Ordnung approximiert. Minima, Maxima und Sattelpunkte der Energiefläche sind dadurch charakterisiert, daß die erste Ableitung der Energie verschwindet. Wenn wir dies fordern, führt das auf die Gleichung

$$\nabla_{\vec{\mathbf{x}}} E \doteq 0 = \vec{\mathbf{g}} + \mathbf{B}(\vec{\mathbf{X}} - \vec{\mathbf{X}}_0) \quad (5.2)$$

In dieser quadratischen Näherung erhalten wir nach Umstellung von (5.2) eine neue Kernanordnung

$$\vec{\mathbf{X}} = \vec{\mathbf{X}}_0 - \mathbf{B}^{-1} \vec{\mathbf{g}} \quad (5.3)$$

Nehmen wir für einen Moment an, daß die Energiefläche tatsächlich eine quadratische Funktion ist und wir die ersten und auch die zweiten Ableitungen exakt zur Verfügung

haben, dann ist  $\vec{X}$  die Lösung und  $E$  minimiert. In unserem konkreten Falle nun sind diese zwei Voraussetzungen dafür aber nicht erfüllt:

1. die Energie ist keine quadratische Funktion,
2. die zweiten Ableitungen sind uns direkt nicht zugänglich .

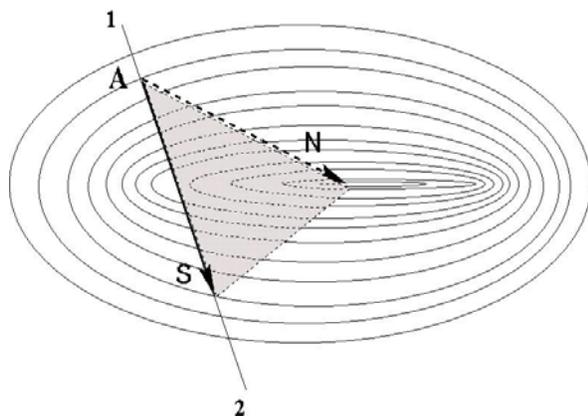
Dem ersten Punkt wird dadurch Genüge getan, daß durch die Entwicklung (5.1) die Minimierung von der wahren Energiefunktion auf die Minimierung der Modellfunktion zweiter Ordnung zurückgeführt wird. Das Minimum dieser Modellfunktion liefert eine neue Stelle  $\vec{X}'$ , an welcher dann der nächste Energiewert zu berechnen ist, womit sich bereits andeutet, daß das Minimierungsverfahren auf einen iterativen Prozeß hinausläuft. Aus (5.3) wird dann

$$\vec{X}_{k+1} = \vec{X}_k - \mathbf{B}_k^{-1} \vec{g}_k \quad . \quad (5.4)$$

Definiert man noch für die Inverse der Hessematrix die Matrix  $\mathbf{H}_k := \mathbf{B}_k^{-1}$ , so folgt für die Suchrichtung  $\vec{p}_k$  im k-ten Optimierungsschritt

$$\vec{p}_k := \vec{X}_{k+1} - \vec{X}_k = -\mathbf{H}_k \vec{g}_k \quad . \quad (5.5)$$

Die Richtung wird nun im wesentlichen durch den Gradient  $\vec{g}_k$  festgelegt. Diese und auch der Betrag von  $\vec{p}_k$  können aber im allgemeinen durch die dem Gradienten vorgeschaltete Matrix  $\mathbf{H}_k$  beeinflußt werden. Um dies einzusehen, wollen wir Gleichung (5.5) anhand der Abbildung (5.1) eingehender studieren.



**Abbildung 5.1:** Mögliche Wahl der Richtung des Suchvektors  $\vec{p}_k$  in Abhängigkeit der Hessematrix (siehe Text). Der Gradient liegt auf der Geraden 1-2, die im Punkt A senkrecht auf der Äquipotentiallinie stehe.

Dort ist die Projektion unserer Modellpotentialfläche (5.1) mit den Äquipotentiallinien dargestellt. Stets bekannt sind uns die Koordinaten der aktuellen Position der Atomkerne  $\vec{X}_0$  und der Gradient. Die einzig unbestimmte Größe ist die Matrix  $\mathbf{H}_k$ . Wir wollen nun zwei Extremfälle diskutieren und zuerst annehmen, daß wir die zweiten Ableitungen (und damit  $\mathbf{H}_k$ ) exakt kennen würden. Dann würde uns Gleichung (5.5) sofort

zum Minimum unserer Modellfunktion (5.1) führen (Pfeil N in Abbildung (5.1)). Diese Situation entspräche dem Newton-Verfahren.

Im anderen Extremfall wollen wir annehmen, daß  $\mathbf{H}_k = \mathbf{E}$  sei, wobei  $\mathbf{E}$  die Einheitsmatrix ist. Dann bestimmt ausschließlich der Gradient die Suchrichtung und diese liegt somit dann in der Richtung des stärksten Gefälles, im Bild auf der Linie 1–2, Pfeil S. Daher können wir annehmen, daß es solche Matrizen  $\mathbf{H}_k$  geben wird, die eine Suchrichtung definieren, welche in den durch N und S begrenzten schattierten Zwischenbereich liegt. *In realita* scheidet der erste Fall ja aus, da wir  $\mathbf{H}_k$  exakt nicht kennen. Wie wir uns diese Matrizen konstruieren können, wollen wir im übernächsten Abschnitt behandeln.

Haben wir uns eine solche Matrix beschafft, ist zwar die Suchrichtung festgelegt, ein grundsätzliches Problem ist jedoch die Bestimmung einer optimalen Schrittweite, also wie weit man in Richtung des Gradienten gehen muß. Aufgrund von deren Ungewißheit verfährt man nun üblicherweise so, daß man versucht, eine verbesserte Geometrie entlang einer durch einen Schrittlängenparameter  $\lambda_k$  charakterisierten Suchrichtung zu finden,  $\vec{\mathbf{p}}_k \rightarrow \lambda \vec{\mathbf{p}}_k$ . Dabei ist  $\lambda_k \geq 0$  ein Element eines Paramterraumes  $\Lambda$ ,  $\lambda_k \in \Lambda$  und kann im Prinzip für jeden Optimierungsschritt andere Werte annehmen. Welche das sind und wie dieses Verfahren im Detail aussieht, wollen wir im folgenden Abschnitt erläutern.

### 5.1.1 Die Schrittweitenoptimierung

Der Übersicht halber lassen wir in diesem Abschnitt den Index  $k$  bei  $\lambda_k$  fort, behalten ihn aber in den Koordinaten bei, um zu erinnern, daß wir im  $k$ -ten Optimierungsschritt sind. In diesem sind uns folgende Werte zugänglich: die geometrische Anordnung  $\vec{\mathbf{X}}_k$  der Atomkerne, der Wert der totalen Energie  $E(\vec{\mathbf{X}}_k)$  und deren Gradient  $\vec{\mathbf{g}}_k = \vec{\mathbf{g}}(\vec{\mathbf{X}}_k)$  sowie eine Näherungslösung für die Matrix  $\mathbf{H}_k$ . Mit Hilfe dieser Größen läßt sich gemäß Formel (5.5) unter Beachtung der parametrisierten Suchrichtung eine neue Kernanordnung  $\vec{\mathbf{X}}_t$  bestimmen <sup>1</sup>,

$$\vec{\mathbf{X}}_t(\lambda) = \vec{\mathbf{X}}_k - \lambda \mathbf{H} \vec{\mathbf{g}}_k = \vec{\mathbf{X}}_k - \lambda \vec{\mathbf{p}}_k \quad . \quad (5.6)$$

Wie im letzten Abschnitt gezeigt, ist die Schrittweite entlang der Suchrichtung aufgrund unserer Näherungen mit einer gewissen Unsicherheit behaftet. Daher wird nun ein ganzer Schritt ( $\lambda_1 = 1$ ) gewählt. Dies scheint zunächst etwas willkürlich. Den Schrittlängenparameter können wir als eine Art Skalierungsfaktor des Gradienten betrachten. Letzterer ist ein Maß dafür, wie stark sich die Energie auf der von uns betrachteten Längenskala ändert. Ganz grob können wir annehmen, daß diese Änderung größenordnungsmäßig im Bereich von etwa  $1eV/a.u.$  liegt. Von daher ist die Wahl von  $\lambda_1 = 1$  offenbar sinnvoll. In den praktischen Rechnungen, deren Ergebnisse wir später im Abschnitt (5.2.1) angeben werden, ist auch der Einfluß der Wahl von  $\lambda_1$  untersucht worden. Allgemeiner ist also

---

<sup>1</sup>t bedeute trial point

$\lambda_1 = \lambda_{start}$  anzusetzen. Die folgenden Betrachtungen diskutieren wir mit  $\lambda_1 = 1$ .

Die Energie ist nun zunächst an der Stelle  $\vec{\mathbf{X}}_t(\mathbf{1}) = \vec{\mathbf{X}}_0 - \mathbf{1}\mathbf{H}_k\vec{\mathbf{g}}_k$  zu berechnen. Ist deren Wert größer als an der Ausgangsgeometrie, wird der Schrittlängenparameter solange verkleinert, bis bei einem gewissen Wert  $\bar{\lambda}_k$  ein niedrigerer Energiewert erreicht ist. Die dazugehörige Kernanordnung  $\vec{\mathbf{X}}_t(\bar{\lambda}_k)$  kann nun als verbesserte Geometrie  $\vec{\mathbf{X}}_{k+1}$  betrachtet werden und als Ausgangspunkt für den nächsten Optimierungsschritt dienen. Jedoch garantiert ein kleiner werdender Energiewert nicht unbedingt das Erreichen des Minimums<sup>2</sup>. Daher muß der Energiewert  $E(\vec{\mathbf{X}}_t(\lambda)_1)$  noch zusätzliche Bedingungen erfüllen. Die Energiewerte  $E(\lambda = 0) = E_0(\vec{\mathbf{X}}_k)$  und  $E(\lambda_1)$  sowie der Gradient  $\vec{\mathbf{g}}(\vec{\mathbf{X}}_k)$  lassen sich dazu verwenden, um ein Akzeptanzkriterium für den berechneten Energiewert  $E(\vec{\mathbf{X}}_t)$  aufzustellen. Bei diesem Kriterium wird gefordert, daß für den neuen Funktionswert  $E(\vec{\mathbf{X}}_t; \lambda)$  gelten soll:

$$E(\vec{\mathbf{X}}_t) = E(\vec{\mathbf{X}}_k + \lambda\vec{\mathbf{p}}_k) \leq E(\vec{\mathbf{X}}_k) + \alpha\lambda\vec{\mathbf{g}}_k^T\vec{\mathbf{p}}_k, \quad (\alpha \geq 0) \quad . \quad (5.7)$$

Im Parameterraum  $\Lambda$  können wir dies als Gleichung einer Geraden mit dem Anstieg  $\alpha\vec{\mathbf{g}}_k^T\vec{\mathbf{p}}_k$  und dem absoluten Glied  $E(\vec{\mathbf{X}}_k)$  auffassen. In Abbildung (5.2) entspricht dies der Geraden  $l(\lambda)$  mit dem negativen Anstieg  $\alpha\vec{\mathbf{g}}_k^T\vec{\mathbf{p}}_k$  ( $\alpha \in R^+$ ), denn die Richtungsableitung  $\vec{\mathbf{g}}_k^T\vec{\mathbf{p}}_k$  kann immer so gewählt werden, daß  $\vec{\mathbf{g}}_k^T\vec{\mathbf{p}}_k \leq 0$ . Damit ist (5.7) erfüllt und  $E(\vec{\mathbf{X}}_k + \lambda\vec{\mathbf{p}}_k)$  ist nicht nur kleiner als  $E(\vec{\mathbf{X}}_k)$  ist, sondern liegt zusätzlich unterhalb der Linie A–B.

Eine weitere Einschränkung der zulässigen Energiewerte ergibt sich aus der Forderung, daß der Betrag der Richtungsableitung an der Versuchsgeometrie  $\vec{\mathbf{X}}_t$  (Punkt D) stets um einen Faktor  $\beta$  kleiner sein soll als in  $\vec{\mathbf{X}}_k$  (Punkt A),

$$|\vec{\mathbf{g}}_k^T(\vec{\mathbf{X}}_k + \lambda\vec{\mathbf{p}}_k)\vec{\mathbf{p}}_k| \leq \beta |\vec{\mathbf{g}}_k^T(\vec{\mathbf{X}}_k)\vec{\mathbf{p}}_k| \quad . \quad (5.8)$$

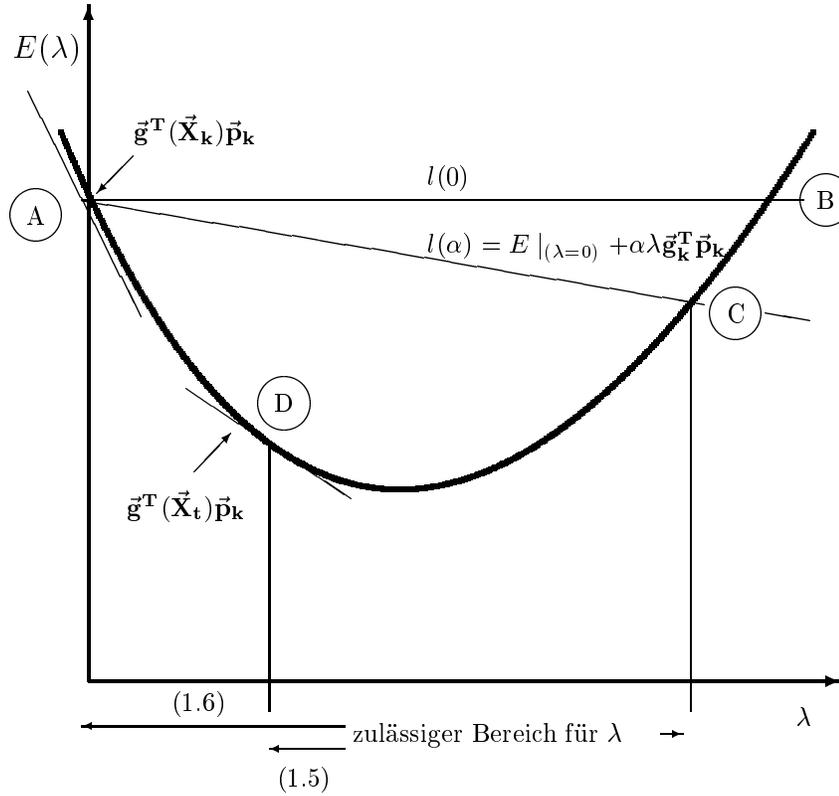
Für die Richtungsableitung selbst folgt daraus:

$$\vec{\mathbf{g}}_k^T(\vec{\mathbf{X}}_k + \lambda\vec{\mathbf{p}}_k)\vec{\mathbf{p}}_k \geq \beta\vec{\mathbf{g}}_k^T(\vec{\mathbf{X}}_k)\vec{\mathbf{p}}_k \quad , \quad (5.9)$$

da ja  $\vec{\mathbf{g}}_k^T\vec{\mathbf{p}}_k \leq 0$  ist. Dies entspricht einem flacheren Funktionenverlauf (Punkt D) als in (A) und schränkt den zulässigen Parameterbereich für  $\lambda$  weiter ein. Wie aus der letzten Gleichung hervorgeht, muß dafür der Gradient auch an der Stelle  $\vec{\mathbf{X}}_t$  berechnet werden. Da im allgemeinen die Situation eintreten kann, daß mehrere Versuchsgeometrien  $\vec{\mathbf{X}}_{t1}, \vec{\mathbf{X}}_{t2}, \dots$  notwendig sind, um die beiden Akzeptanzkriterien (5.7,5.8) zu erfüllen, müßten mitunter sehr viele Gradientenberechnungen durchgeführt werden. Da

---

<sup>2</sup>Beispielsweise kann eine monotone Folge immer tieferer Funktions(=Energie)werte zwar konvergieren, jedoch nicht unbedingt zum Minimum, siehe z.B. [85], Seite 118.



**Abbildung 5.2:** Bestimmung einer optimalen Schrittweite: Die Tangenten an die parametrisierte Funktion (5.6) sind die Richtungsableitungen für  $\lambda_k = 0$  (Punkt A) und für  $\lambda_k \neq 0$  (Punkt D). Die Gerade  $l(\alpha)$  schneidet die Funktion (5.6) in Abhängigkeit von  $\lambda$  an verschiedenen Stellen (Punkt C).

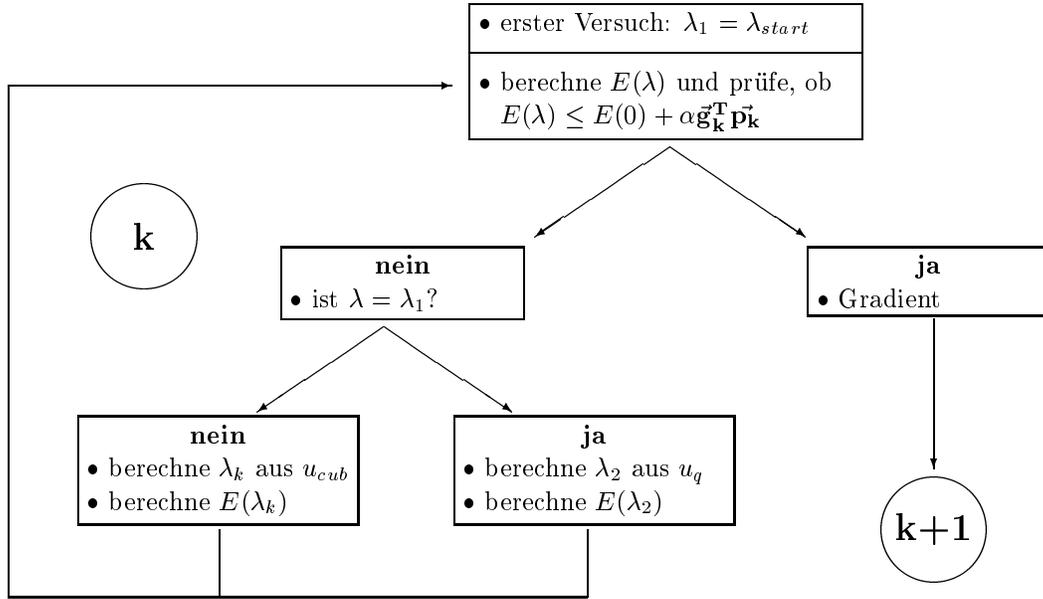
sich herausgestellt hat, daß die Gradientenberechnungen deutlich mehr Rechenzeit als eine Energieberechnung benötigen, ist das zweite Kriterium (5.8) für unsere Zwecke un- zweckmäßig und wir müssen uns mit dem abgeschwächten Kriterium (5.7) behelfen. Für den Fall, daß für den ersten Versuchspunkt  $\vec{X}_t(1)$  die Bedingung (5.7) nicht erfüllt ist, also

$$E(1) > E(0) + \alpha \vec{g}_k^T \vec{p}_k \quad , \quad (5.10)$$

kann folgende Vorgehensweise angewendet werden, um den Schrittweitenparameter  $\lambda$  zu optimieren. Aus der totalen Energie  $E(0) := E(\vec{X}_k) = E(\lambda = 0)$ , dem Gradienten  $\vec{g}(\lambda = 0)$  und der Energie  $E(1) := E(\lambda_1 = 1)$  im ersten Versuchsschritt läßt sich eine quadratische Hilfsfunktion  $u_q$  folgender Gestalt modellieren:

$$u_q(\lambda) = \left( E(1) - E(0) - \vec{g}_k^T \vec{p}_k \right) \lambda^2 + \vec{g}_k^T \vec{p}_k \lambda + E(0) \quad . \quad (5.11)$$

Für  $\lambda = 1$  folgt  $u_q(1) = E(1) = E(\vec{X}_t)$  und für  $\lambda = 0$  erhalten wir  $u_q(0) = E(0) = E(\vec{X}_k)$ . Durch Differentiation von  $u_q$  bezüglich  $\lambda$  folgt sodann ein Minimum für  $\lambda_2 = \bar{\lambda}$ :



**Abbildung 5.3:** Schrittlengthenoptimierung in Abhängigkeit der vorrätigen Informationen im  $k$ -ten Optimierungsschritt: Die Schrittlengthe innerhalb dieses Zyklus wird solange verkleinert, bis ein akzeptabler Wert für die Energie (5.7) erreicht ist. An der dazugehörigen Geometrie wird nun der Gradient berechnet und zum nächsten Zyklus ( $k+1$ ) übergegangen.

$$\bar{\lambda} = \frac{-\vec{g}_k^T \vec{p}_k}{2(E(1) - E(0) - \vec{g}_k^T \vec{p}_k)} . \quad (5.12)$$

Der auf diese Weise gewonnene Parameter  $\bar{\lambda}$  ist größer Null, da  $\vec{g}_k^T \vec{p}_k \leq \mathbf{0}$  und weiterhin ein Minimum, da die zweite Ableitung der Modellfunktion  $u_q$  wegen (5.10)

$$u_q'' = 2(E(1) - E(0) - \vec{g}_k^T \vec{p}_k) \geq 0 . \quad (5.13)$$

Falls nun  $E(1) > E(0) + \alpha \vec{g}_k^T \vec{p}_k$ , folgt für  $\bar{\lambda}$

$$\lambda_2 = \bar{\lambda} < \frac{1}{2(1 - \alpha)} . \quad (5.14)$$

Dadurch ist eine obere Schranke für die Schrittlengthe gegeben. Sollte der Fall eintreten, daß  $E(1) \gg E(0)$  ist, kann  $\bar{\lambda}$  sehr klein werden, wodurch der Optimierungsalgorithmus ineffizient wird. Um dies abzufangen, wird eine untere Schranke für  $\lambda_2$  angesetzt:  $\lambda_2 \geq 0.1\lambda_1$ . An der jetzt durch  $\lambda_2$  festgelegten Schrittlengthe wird nun erneut eine Energierechnung durchgeführt und geprüft, ob  $E(\lambda_2)$  Kriterium (5.7) erfüllt. Ist dies der Fall,

wird die durch  $\lambda_2$  erhaltene Geometrie als verbessert akzeptiert und an dieser kann nun wieder der Gradient berechnet werden. Sollte  $E(\lambda_2)$  jedoch noch immer nicht Kriterium (5.7) erfüllen, so muß die Schrittlänge weiter verkleinert werden. Da nun aus der Energierechnung  $E(\lambda_2)$  neben  $E(0), E(1)$  und  $\vec{\mathbf{g}}(\mathbf{0})$  ein vierter Zahlenwert zur Verfügung steht, läßt sich daraus eine kubische Funktion

$$u_{\text{cub}}(\lambda) = a\lambda^3 + b\lambda^2 + \vec{\mathbf{g}}_{\mathbf{k}}^T \vec{\mathbf{p}}_{\mathbf{k}} \lambda + E(0) \quad (5.15)$$

bilden, die ein Minimum bei

$$\bar{\lambda} = \frac{-b}{3a} + \frac{\sqrt{b^2 - 3a\vec{\mathbf{g}}_{\mathbf{k}}^T \vec{\mathbf{p}}_{\mathbf{k}}}}{3a} \quad (5.16)$$

hat. Die Zahlen  $a$  und  $b$  ( $a, b \in R^+$ ) sind so bestimmt, daß  $u_{\text{cub}}$  die Werte von  $E(\lambda_1)$  bzw.  $E(\lambda_2)$  annimmt (siehe z.B. [86], S. 377 f.). In der Abbildung (5.3) ist die Schrittlängenbestimmung nochmals in einer grafischen Übersicht verdeutlicht.

### 5.1.2 Die Neubestimmung der inversen Hessematrix

Um die Gleichung (5.3) zum Auffinden einer verbesserten Geometrie anwenden zu können, ist die Matrix  $\mathbf{B}_{\mathbf{k}}$  der zweiten Ableitungen der Energie nach den Kernkoordinaten erforderlich. Diese stehen aber wegen ihres immensen rechentechnischen Aufwandes nicht direkt zur Verfügung. An dieser Stelle können wir uns auf zweierlei Weise behelfen.

Die naheliegendste Möglichkeit ist die, für  $\mathbf{B}_{\mathbf{k}}$  die Einheitsmatrix anzusetzen, was wir bereits im Abschnitt (5.1.1) bei der Auswahl einer akzeptablen Suchrichtung  $\vec{\mathbf{p}}_{\mathbf{k}}$  diskutiert hatten. Die Einheitsmatrix hat die Eigenschaft, daß sie zu sich selbst invers ist, woraus folgt, daß gemäß Gleichung (5.5) die inverse Hessematrix  $\mathbf{H}_{\mathbf{k}}$  und damit eine Suchrichtung  $\vec{\mathbf{p}}_{\mathbf{k}}$  existiert. Die Bedingung, daß  $\mathbf{B}_{\mathbf{k}}$  invertierbar ist, ist eine wichtige Forderung. Multiplizieren wir nämlich Gleichung (5.2) mit der Suchrichtung  $\vec{\mathbf{p}}_{\mathbf{k}}$  und fordern für die Richtungsableitung  $\vec{\mathbf{g}}_{\mathbf{k}}^T \vec{\mathbf{p}}_{\mathbf{k}} < 0$ , folgt

$$\vec{\mathbf{g}}_{\mathbf{k}}^T \vec{\mathbf{p}}_{\mathbf{k}} = -(\vec{\mathbf{X}}_{\mathbf{k}} - \vec{\mathbf{X}}_0)^T \mathbf{B}_{\mathbf{k}} (\vec{\mathbf{X}}_{\mathbf{k}} - \vec{\mathbf{X}}_0) < 0 \quad . \quad (5.17)$$

Dies ist nur erfüllt, wenn  $\mathbf{B}_{\mathbf{k}}$  positiv definit ist. Somit ist die Wahl der Einheitsmatrix  $\mathbf{E}$  als einfachste Approximation gerechtfertigt. Da  $\mathbf{E}$  jedoch die Krümmung der Energiefläche im allgemeinen sehr schlecht wiedergibt, führt die Wahl  $\mathbf{B}_{\mathbf{k}} = \mathbf{E}$  auf eine ineffektive Optimierung, wie wir im Abschnitt (5.2.1) anhand konkreter Beispiele sehen werden. Wir wollen noch anmerken, daß im Falle der Verwendung der Einheitsmatrix das Optimierungsverfahren als 'steepest-descent method' bezeichnet wird. Dieses ist linear konvergent.

Eine andere Möglichkeit der Wahl von  $\mathbf{B}_k$  bzw.  $\mathbf{H}_k$  besteht darin, die im Verlauf einer Optimierung anfallenden Informationen zur Konstruktion von  $\mathbf{B}_k$  zu verwenden. Denn im  $k$ -ten Optimierungsschritt sind der Optimierungsroutine die Werte der Koordinaten und der Gradienten des vorherigen und des aktuellen Optimierungszyklus zugänglich. Die Werte der Koordinatendifferenzen und Gradienten zweier aufeinanderfolgender Optimierungszyklen lassen sich daher zur schrittweisen Aufstellung der  $\mathbf{B}_k$  Matrizen ausnutzen. Wir können uns dies anhand folgender Überlegungen plausibel machen: So, wie die Differenz von Funktionswerten  $f(x + \delta x) - f(x)$  die erste Ableitung der Funktion  $y = f(x)$  in  $x$  approximiert, läßt sich die zweite Ableitung  $f''(x)$  durch den Differenzenquotienten der ersten Ableitungen  $f'(x + \delta x) - f'(x)$  annähern (Sekantenverfahren).

$$f'' \propto \frac{f'(x + \delta x) - f'(x)}{\delta x} \quad . \quad (5.18)$$

Genau diese Werte stehen uns ja zur Verfügung und wegen des eindeutigen Zusammenhanges zwischen  $\mathbf{B}_k$  und  $\mathbf{H}_k$  scheint es plausibel, daß ebene Informationen in die Update-Formeln eingehen, die wir im Anhang (A.3) explizit angeben werden. Zum Abschluß dieses Abschnitts wollen wir noch anmerken, daß es zu Beginn einer Optimierung, wenn man noch weit entfernt von Minimum ist, offenbar keinen Sinn macht, allzu viele Überlegungen in die Konstruktion einer Startmatrix  $\mathbf{H}_0$  zu stecken. Wählen wir nämlich als 'initial guess' die Einheitsmatrix, so sind wir auf der sicheren Seite, denn erstens liegt die neue Geometrie  $\vec{\mathbf{X}}_{k+1}$  auf jeden Fall in Richtung des Gradienten und obendrein ist die Richtungsableitung in diesem Falle immer kleiner Null:  $\vec{\mathbf{g}}_1^T \vec{\mathbf{p}}_1 = -\vec{\mathbf{g}}_1^T \mathbf{B}_1 \vec{\mathbf{g}}_1 = -\vec{\mathbf{g}}_1^T \mathbf{E} \vec{\mathbf{g}}_1 \leq \mathbf{0}$ . Da der Gradient aber im allgemeinen nicht direkt zum Minimum zeigt (siehe nochmals Abbildung (5.1)), wird dieses Verfahren im Verlaufe einer Optimierung ineffizient werden und zunehmend unnötig viele Schrittlängenoptimierungen gemäß Abschnitt (5.1.1) durchführen. Deshalb scheint es angebracht, beim Start einer Optimierung zunächst das steepest-descent Verfahren anzuwenden, weil dann die Suchrichtungen  $\vec{\mathbf{p}}_k$  auf jeden Fall zunächst in Richtung des Gradienten zeigen. Im weiteren Verlauf sollte dann das (superlinear konvergente) 'variable-metric' Verfahren angewendet werden. In diesem Sinne können wir die 'variable-metric' Methode als eine Art weiches Umschalten zwischen dem steepest descent Verfahren und dem Newtonverfahren verstehen. Für den Begriff 'variable-metric' wird auch häufig die Bezeichnung Quasi-Newton Methode benutzt.

Für den Update der  $\mathbf{H}_k$  Matrizen sind unterschiedliche Implementierungen vorgeschlagen worden, die sich geringfügig in der Art und Weise der Neubestimmung unterscheiden. Dies ist ausführlich in Büchern der numerischen Mathematik [84, 85, 86] dokumentiert.

Zusammenfassung:

Die Potentialenergiefläche wird lokal durch eine Fläche zweiter Ordnung approximiert und die Minimierung der 'wahren' Energiefunktion auf die dieser Modellfunktion zurückgeführt. Die Richtung des stärksten Gefälles auf der Potentialenergiefläche wird durch den Energiegradienten festgelegt. Die zugehörige Suchrichtung muß wegen der Unsicherheiten ihres Betrages und ihrer Richtung parametrisiert werden. Zur Konvergenzbeschleunigung läßt sich die Matrix der zweiten Ableitungen im Verlaufe der Optimierung sukzessive verbessern und an die Topologie der Energiefläche anpassen. Deren Matrixelemente stellen im Minimum der totalen Energie gerade die Kraftkonstanten dar.

## 5.2 Ergebnisse

### 5.2.1 Voruntersuchungen

Für alle Rechnungen zur Geometrieoptimierung ist stets das Slater'sche Austauschfunktional mit  $X_\alpha = 0.7$  benutzt worden.

In diesem Abschnitt wollen wir die Resultate von Voruntersuchungen darlegen. Als erstes wollen wir den Energiegradienten untersuchen. Um zu testen, ob dieser mit unserer Methode<sup>3</sup> richtig berechnet wird, wurde die totale Energie des zweiatomigen  $\text{Cu}_2$  Moleküls numerisch differenziert. Der numerische Gradient ergibt sich in der allereinfachsten Form durch die Berechnung des Differenzenquotienten

$$D(R) = \frac{E(R + \delta R) - E(R)}{\delta R} \quad (5.19)$$

an allen in der folgenden Tabelle aufgeführten internuklearen Abständen, wobei  $\delta R \sim 10^{-4} a.u.$  angesetzt wurde. Dadurch ist es möglich, den semianalytisch<sup>3</sup> berechneten Gradienten mit dem numerischen direkt zu vergleichen. Das Ergebnis dieser Untersuchung ist in der folgenden Tabelle dargestellt.

---

<sup>3</sup>Diese Art des Gradienten, die wir im Kapitel (4.1) abgeleitet haben und für die Geometrieoptimierung benutzen, wird auch als 'analytischer' Gradient bezeichnet. Diese Bezeichnung stammt aus der Quantenchemie, wo sich die Gradiententerme durch die Verwendung von Gauß-Basen echt analytisch berechnen lassen. Da wir jedoch beispielsweise für die Basissatzkorrektur die Radialableitungen numerisch berechnen müssen, trifft diese Bezeichnungswiese so nicht zu. Wie wir gesehen haben, lassen sich jedoch die Winkelanteile dieses Gradiententyps analytisch behandeln, sodaß wir unseren Energiegradienten als 'semianalytisch' bezeichnen wollen.

R (a.u.)	$E_b$ (eV)	Gradiententyp		HF	OBC	DFC
		numerisch	semianalytisch			
4.0	-2.545	-0.029	-0.029	-0.063	0.540	-0.477
4.1	-2.607	-0.016	-0.016	0.035	0.475	-0.527
4.2	-2.637	-0.007	-0.006	0.118	0.418	-0.536
4.3	-2.642	0.002	0.003	0.187	0.368	-0.553
4.4	-2.628	0.009	0.009	0.245	0.326	-0.562
4.5	-2.599	0.014	0.014	0.293	0.288	-0.567
4.6	-2.557	0.017	0.018	0.333	0.256	-0.571

**Tabelle 5.1:** Numerischer und semianalytisch berechneter Energiegradient im Vergleich. Die rechten drei Spalten bezeichnen den Hellmann-Feynman Term sowie die beiden Korrekturterme der Basissatzkorrektur (OBC) und Dichtefitkorrektur (DFC). Man beachte deren Größenordnung!

Das Verhalten des Gradienten ist konsistent bezüglich des Verlaufes der Energie (hier wurde die Bindungsenergie eingetragen); am Minimum der Potentialkurve hat der Gradient einen Nulldurchgang. Links davon haben wir ein repulsives Verhalten (Minuszeichen), rechts davon erfahren die beiden Atome eine Anziehung. Der semianalytische Gradient ergibt sich aus der Summe von Hellmann-Feynman Gradient (HF), der Basissatzkorrektur (OBC) und der Dichtefitkorrektur (DFC). Es wird hier eine sehr gute Übereinstimmung mit dem numerisch berechneten Gradienten erzielt. Wie die Tabelle weiterhin verdeutlicht, liegen die Gradientenbeiträge der Korrekturterme in derselben Größenordnung wie der eigentlich physikalische Hellmann-Feynman Gradient. Sie sind daher nicht vernachlässigbar!

In einem nächsten Schritt wurde die Optimierungsroutine getestet. Zu diesem Zweck wurde von jeweils von verschiedenen Ausgangsgeometrien gestartet; einmal bei  $R=4.0$  a.u. (links) und zum anderen bei  $R=6.0$  a.u. (rechts) vom Minimum. Gleichzeitig wurde der Einfluß der Anpassung der Hessematrix an die Potentialkurve untersucht, indem einmal mit (variable-metric) und einmal ohne (steepest-descent) Update optimiert wurde.

Schritt	variable-metric			steepest-descent		
	R(a.u.)	$E_{tot}$ (eV)	Gradient	R(a.u.)	$E_{tot}$ (eV)	Gradient
0	4.000	-89958.014613	0.04036	4.000	-89958.014613	0.04036
1	4.057	-89958.053624	0.02925	4.057	-89958.053624	0.02925
2	4.207	-89958.106969	0.00631	4.098	-89958.074937	0.02235
3	4.249	-89958.111238	0.00141	4.130	-89958.087636	0.01749
4	4.260	-89958.111753	0.00008	4.155	-89958.095569	0.01391
				:	:	:

**Tabelle 5.2:** Annäherung an das Minimum von links mit und ohne Update der Hessematrix. Die Optimierung erfolgte hierbei in kartesischen Koordinaten. Der Schrittlängenparameter betrug  $\lambda = 1.0$ .

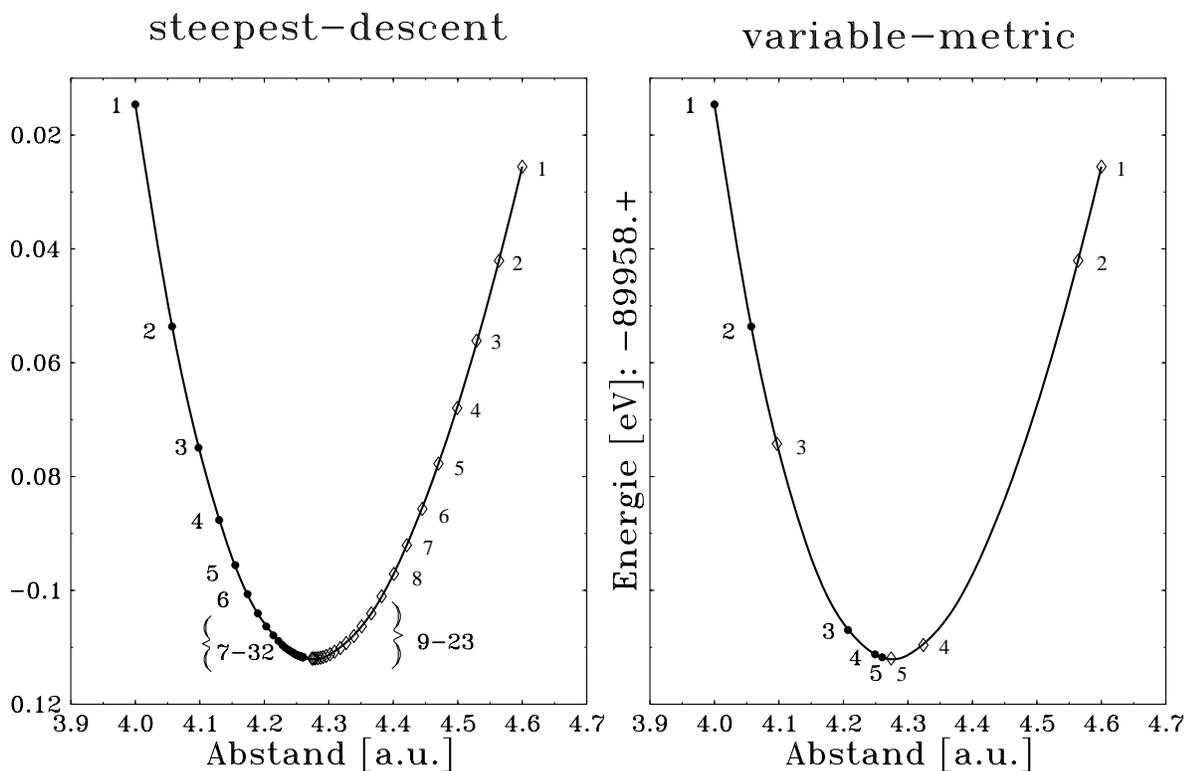
Zum Auffinden des Energieminimums waren bei vorgegebener Wahl des Schrittlängenparameters mit der variable-metric Methode insgesamt fünf Optimierungsschritte notwendig. Dieselbe Rechnung wurde nochmals durchgeführt, diesmal aber die Hessematrix unverändert gelassen (steepest-descent Verfahren). Insgesamt wurden (32!) Optimierungszyklen benötigt, um zum Energieminimum zu gelangen. In der Tabelle (5.2) sind daher nur die ersten fünf Schritte angegeben. Die ersten beiden Zeilen sind für beide Verfahren identisch, da auch bei der variable-metric Methode mit der Einheitsmatrix gestartet wurde. Diese ist beim steepest-descent Verfahren für alle Optimierungszyklen beibehalten worden. Dieselben Untersuchungen wurden nun für eine Ausgangsgeometrie bei einem internuklearen Abstand von  $R=4.6$  a.u. durchgeführt. Das Ergebnis ist in Tabelle (5.3) dargestellt.

Schritt	variable-metric			steepest-descent		
	R(a.u.)	$E_{tot}$ (eV)	Gradient	R(a.u.)	$E_{tot}$ (eV)	Gradient
0	4.600	-89958.025549	0.02571	4.600	-89958.025548	0.04036
1	4.564	-89958.042062	0.02385	4.564	-89958.042062	0.02925
2	4.097	-89958.074241	0.02259	4.529	-89958.056146	0.02235
3	4.324	-89958.109596	0.00644	4.499	-89958.067968	0.01749
4	4.274	-89958.111073	0.00153	4.470	-89958.077734	0.01391
				:	:	:

**Tabelle 5.3:** Annäherung an das Minimum von rechts, wiederum mit und ohne Update der Hessematrix. Auch hier wurde mit einem Schrittlängenparameter  $\lambda = 1.0$  optimiert.

Auch bei einem internuklearen Abstand, der größer als der Gleichgewichtsabstand ist, wird mit Update der Hessematrix das Minimum mit deutlich weniger Optimierungsschritten (fünf) im Vergleich der Optimierung mit starrer Metrik gefunden. Für letztere waren in diesem Fall 23 Zyklen notwendig. Aus den Tabellen ist ersichtlich, daß sich geringfügige Unterschiede im Gleichgewichtsabstand für die beiden unterschiedlichen Ausgangsgeometrien ergeben, obwohl in beiden Fällen dieselbe Genauigkeitsschranke für das Abbrechen der Optimierung benutzt wurde. Als Maß wurde ein bestimmter Wert der Gradientennorm zugrunde gelegt. Wie eine genauere Analyse der Energie und des Gradienten in unmittelbarer Umgebung des Minimums zeigen, stimmt das Verschwinden des Gradienten nicht exakt mit dem Minimum der totalen Energie überein. Im Rahmen unserer Genauigkeit ist diese Diskrepanz jedoch tolerierbar. Es scheint sich hier anzudeuten, daß wir nunmehr an die Grenzen der numerischen Genauigkeit gestoßen sind. Zusammenfassend sind die Ergebnisse der bisherigen Rechnungen nochmals in Abbildung (5.4) grafisch veranschaulicht.

Um zu testen, ob unser Optimierungsverfahren auch auf echt relativistische Systeme anwendbar ist, wurde in einer weiteren Rechnung das Gold-Dimer  $Au_2$  optimiert (Abbildung 5.5). Dazu wurde zunächst die Potentialenergiekurve 'zu Fuß' für einen Abstandsbereich von  $R=4.2$  a.u. bis  $R=5.4$  a.u. berechnet. Anschließend wurde die Optimierung bei einem internuklearen Abstand von  $R=4.2$  a.u. mit dem variable-metric Verfahren gestartet und das Minimum in sechs Schritten gefunden. In einer weiteren Rechnung ab  $R=5.4$  a.u. wurde nochmals das steepest-descent Verfahren benutzt, jetzt aber mit



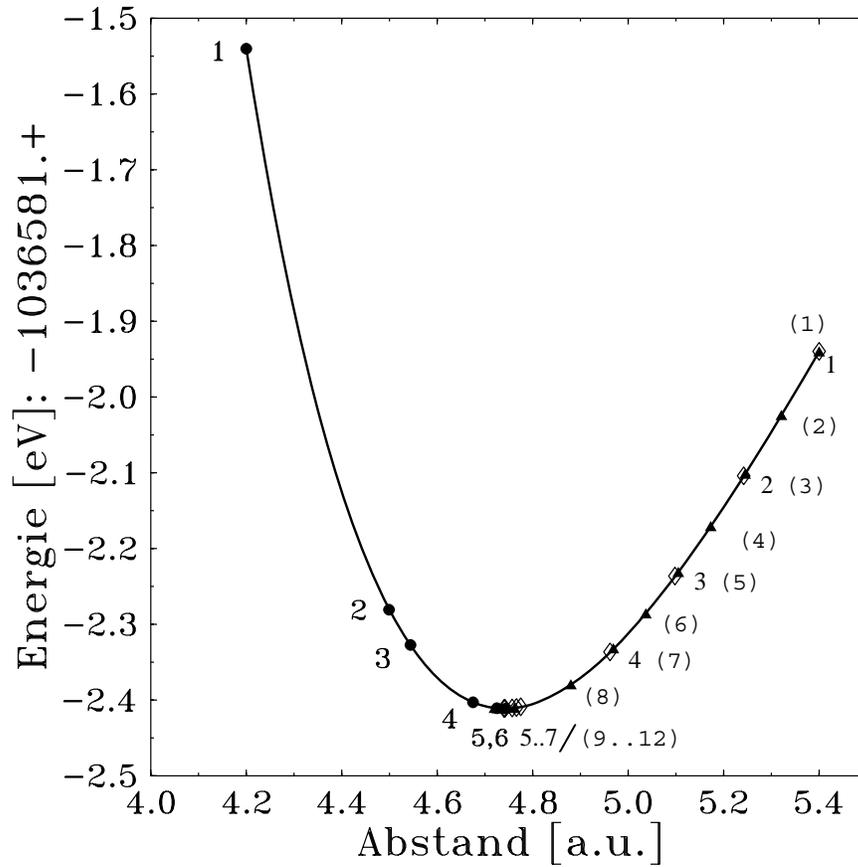
**Abbildung 5.4:** Optimierungen von  $\text{Cu}_2$ : im linken Bild ist die Abfolge der Zyklen für den Fall der 'starren' Metrik angegeben. Das Verfahren wird nach einer anfänglich guten Konvergenz mit zunehmender Annäherung an das Minimum ineffizient. Das rechte Bild zeigt den Verlauf der Optimierung unter Verwendung der variable-metric Methode. Die Energieskalen sind in beiden Fällen identisch.

unterschiedlichen Schrittlängenparametern. Im ersten Fall wurde  $\lambda_1 = 1.0$  verwendet. Die Abfolge der Optimierungsschritte ist durch die Zahlen in den runden Klammern gekennzeichnet. Insgesamt wurden zwölf Zyklen durchlaufen. Sodann wurde nochmals an derselben Ausgangsgeometrie gestartet, nun aber mit dem Schrittlängenparameter  $\lambda_1 = 2.0$ . Das Minimum wird jetzt in nur sieben Schritten gefunden.

## 5.2.2 Optimierung dreidimensionaler Strukturen

Der letzte Abschnitt befaßte sich hauptsächlich mit dem Test der Gradiententerme und dem korrekten Ablauf einer Optimierung. Dazu haben wir Untersuchungen an zweiatomigen Molekülen vorgenommen. In diesem Falle lassen sich, wie geschehen, die semi-analytischen Gradienten mit den numerischen leicht vergleichen. In diesem Abschnitt wollen wir den eigentlich interessanten dreidimensionalen Fall betrachten.

Zunächst diskutieren wir die Geometrieoptimierung für den Fall, daß sich die Atome nur



**Abbildung 5.5:** Optimierung von  $\text{Au}_2$ : links des Minimums mit Hessematrix-Update. Zur Veranschaulichung der Abhängigkeit vom Schrittweitenparameter  $\lambda_1$  sind für die Optimierungen rechts des Minimums für diesen unterschiedliche Werte verwendet worden:  $\lambda_1 = 1.0$  (schwarze Dreiecke) und  $\lambda_1 = 2.0$  (Diamonds; Zahlen in runden Klammern).

innerhalb einer vorgegebenen Symmetrie bewegen können. Das vermindert den Aufwand der Berechnungen insofern, als das dadurch die Zerlegung der Fock- und Überlappmatrizen in kleinere disjunkte Unterräume entsprechend den irreduziblen Darstellungen der vorgegebenen molekularen Punktgruppe zerfallen. Dadurch verringert sich der Aufwand für die Diagonalisierung dieser Matrizen.

Als erstes Testmolekül wurde ein  $\text{Na}_3$ -Cluster in einer  $D_{3h}$  Symmetrie untersucht. In dieser Konfiguration liegen die Atome in den Ecken eines gleichseitigen Dreiecks mit jeweils einem Winkel von  $60^\circ$  zwischen jeweils zwei Atomen. Der einzige zur Optimierung freigegebene Parameter ist die Länge des Abstands eines der Atome zum Massenschwerpunkt. Aufgrund der Molekülsymmetrie ist dieser Abstand für alle drei Atome gleich. Völlig gleichwertig damit ist der internukleare Abstand. Für diesen wurde als Ausgangswert  $R=3.46$  a.u. angesetzt und sodann die Optimierung gestartet. Das Ergebnis ist in der Tabelle (5.4) angegeben. Die Ausgangskernanordnung war offenbar so beschaffen, daß die Atome eine sehr starke Abstoßung erfahren haben. Die internuklearen Abstände haben sich etwa verdoppelt.

R[a.u.]	$E_{tot}$ (eV)	$\ \vec{g}\ $
3.464	-13177.453915	0.29384
3.757	-13179.423275	0.20530
4.439	-13182.105318	0.10026
5.089	-13183.379892	0.04968
5.728	-13183.977831	0.02153
6.217	-13184.180396	0.00862
6.543	-13184.239461	0.00218
6.654	-13184.250407	0.00108
6.762	-13184.254628	0.00031

**Tabelle 5.4:** Optimierung des  $\text{Na}_3$  Moleküls in  $D_{3h}$ -Symmetrie: Die Gradientennorm  $\|\vec{g}\|$  nimmt mit zunehmender Annäherung an das Minimum ab. Basis für diese und alle weiteren Na-Rechnungen:  $[\text{Ne}] 3s^1 3p^0$ .

In einer weiteren Rechnung wurde die Molekülsymmetrie aufgegeben und eine beliebige Ausgangskernanordnung angenommen. Als Ergebnis der Optimierung wird vom Programm eine  $C_{2v}$  Symmetrie (gleichschenkliges Dreieck) gefunden. In der Tabelle (5.5) ist dies dargestellt. Wir erzielen eine gute Übereinstimmung mit anderen Rechnungen.

Ein anderes, wenn auch ähnliches System ist ein  $\text{Al}_3$ -Cluster. Die Ausgangsgeometrie war wieder beliebig vorgegeben. In der Tabelle (5.6) sind die Ergebnisse der Optimierung dargestellt. Wir finden eine  $C_{2v}$  Struktur mit den Bindungslängen von etwa 5.2 a.u. und einem Bindungswinkel von  $38.8^\circ$ . Eine CI-Rechnung<sup>4</sup> dieses Systems von Petterson et al. [89] ergibt eine Bindungslänge von 5.12 a.u. und einen Winkel von  $55.6^\circ$ . Hier wird keine gute Übereinstimmung im Bindungswinkel erzielt. Sicherlich liegt dies am verwendeten Slater'schen Dichtefunktional, welches 'winkelartige' Bindungen oftmals nicht gut wiedergeben kann.

<sup>4</sup>mit zusätzlichem Zwei- (Lennard-Jones) und Dreikörperwechselwirkungspotential

$R_1$ [a.u.]	$R_2$ [a.u.]	Referenz
5.7	7.1	LSD-PP [87]
6.46	9.05	AE-HF [88]
6.43	8.86	ECP-HF [88]
6.43	8.67	ECP-CI [88]
6.53	8.37	unser

**Tabelle 5.5:** Optimierung des  $\text{Na}_3$  Moleküls ohne Symmetrie: Die Gleichgewichtslage hat eine  $C_{2v}$  Symmetrie.  $R_1$  und  $R_2$  sind die Seitenlängen eines gleichschenkligen Dreiecks. Angegeben sind Vergleichsrechnungen mit einer Pseudopotentialmethode (PP), all-electron Hartree-Fock (AE-HF) sowie Effective-Core-Potential (ECP) Hartree-Fock (ECP-HF) und Effective-Core-Potential CI (ECP-CI).

$R_{12}$	$R_{13}$	$R_{23}$	$E_{tot}(eV)$	$\ g\ $
3.17	3.91	5.38	-19720.20405	0.63376
4.02	4.16	5.82	-19726.44871	0.19625
4.34	4.42	6.03	-19727.51103	0.10932
4.72	4.78	6.29	-19728.29194	0.06396
5.08	5.05	6.44	-19728.48592	0.01751
5.24	5.18	6.50	-19728.51383	0.00992

**Tabelle 5.6:** Optimierung des  $\text{Al}_3$  Moleküls: es wurde keine Symmetrie angenommen. In den ersten drei Spalten der ersten Zeile sind die internuklearen Abstände an der Ausgangsgeometrie angegeben. Die Abstände an der Endgeometrie deuten eine  $C_{2v}$ -Symmetrie an. In der letzten Spalte ist die mit Annäherung an das Minimum abnehmende Gradientennorm erkennbar. Basis war in diesem Falle:  $[\text{Ne}] 3s^2 3p(\frac{1}{2})^1 3p(\frac{3}{2})^0$ .

### 5.3 Zusammenfassung der Rechnungen

Die Ergebnisse aller bisher durchgeführten Untersuchungen lassen folgende Schlussfolgerung zu:

Die für eine erfolgreiche Optimierung notwendigen Gradiententerme sind richtig implementiert; das Verhalten des Gradienten in Bezug auf das Minimum der Energie ist konsistent. Es ist ratsam, zu Beginn einer Optimierung mit dem steepest-descent Verfahren zu starten und im Verlaufe der Optimierung auf die variable-metric Methode umzuschalten, wodurch die Konvergenz erheblich verbessert werden kann.

Das Ziel einer Geometrieoptimierung besteht ja darin, von einer unbekannt geometrischen Anordnung die Gleichgewichtslage zu finden. Da man im allgemeinen nur eine ungefähre Vermutung von der möglichen Anfangsgeometrie – und diese muß so beschaffen sein, daß unsere quadratische Näherung noch halbwegs gut ist – hat, kann *a priori*

**keine Symmetrie** gefordert und in die Rechnungen eingebracht werden. Für eine akzeptable Genauigkeit treibt dies die Anzahl der Integrationspunkte jedoch derart in die Höhe, daß wir damit sehr schnell an die rechentechnischen Grenzen stoßen. So wurden für die Na<sub>3</sub>- und Al<sub>3</sub>-Rechnungen ca. 70000 Integrationspunkte benötigt. In einer versuchsweise begonnenen Rechnung des RfCl<sub>4</sub> Moleküls, an dem wir im Abschnitt (3.6.2) einige Dichtefunktionaluntersuchungen vorgenommen hatten, wurden von der Integrationsroutine ca. 250000 (!) Integrationspunkte erzeugt. Zeitmessungen in einer Rechnung am Na<sub>3</sub>-Molekül mit ca. 83000 Integrationspunkten haben ergeben, daß mit fast 37% der gesamten CPU-Zeit (ca. 20 Stunden) die mit Abstand meiste Rechenzeit für die Berechnung des Gradienten der Basissatzkorrektur (4.75) benötigt wird. Eine entsprechende Analyse am Cu<sub>2</sub>-Dimer ergab 26%. Daher können wir vermuten, daß sich dieser Anteil deutlich vergrößern wird, wenn man zu größeren und schwereren Systemen übergeht. Letzters ist offensichtlich, denn in die Berechnung der Matrixelemente für die Basissatzkorrektur geht unmittelbar die Anzahl der atomaren Basisfunktionen ein.

Auch dann, wenn man eine gewisse Symmetrie des Moleküls zuläßt, ergibt sich keine Aufwandsreduktion hinsichtlich der Anzahl der Integrationspunkte, denn die Gradienten sind vorzeichenbehaftete Größen (siehe auch Bemerkungen am Ende des Abschnitts (3.4)). Dies geht nur dann, wenn die Optimierung in Symmetriekoordinaten formuliert wird. Dafür notwendig ist aber die Adaptierung der Gradienten auf diese Koordinaten. Die entsprechende Implementierung der Transformationsmatrizen [90] in das Integrationspaket hat sich als zu kompliziert herausgestellt.

Das die Gradientenkorrekturterme etwa in derselben Größenordnung wie der physikalische Hellmann-Feynman Term liegen, mag neben den von uns gemachten Näherungen (endliche Basissätze, Ladungsdichtefit) auch an der pathologischen Struktur des Operators der Elektron-Kern Wechselwirkung liegen. Dies wird durch die Differentiation noch verstärkt:

$$\frac{d}{dR} \left( \frac{1}{|\vec{r} - \vec{R}|} \right) \rightarrow -\frac{1}{|\vec{r} - \vec{R}|^3} (\vec{r} - \vec{R}) \quad .$$

Schließlich und endlich bedarf folgender Punkt noch einer Beachtung: Die Diskussion der Geometrieabhängigkeit der einzelnen Energiekomponenten und Nebenbedingungen hatten wir ohne Rücksicht auf deren konkrete Berechnungsvorschrift vorgenommen. Im Abschnitt (2.4) des Kapitels 2 wurde darauf hingewiesen, daß die Berechnung der Matrixelemente mit einer numerischen Integrationsmethode durchgeführt wird. Strenggenommen sind auch die bei dieser Methode benutzten Integrationsgewichte geometrieabhängig. Ändert man nämlich die Kernanordnung ab, so werden Integrationspunkte und -gewichte anders im Integrationsgebiet verteilt. Wie in [53] gezeigt wurde, spielt diese Abhängigkeit keine wesentliche Rolle mehr, wenn man mit hinreichend vielen Integrationspunkten arbeitet. Dies ist in unseren Rechnungen sicher erfüllt, sodaß wir keine Anstalten machen wollen, zusätzliche Ableitungen der Integrationsgewichte zu programmieren.



# Kapitel 6

## Fazit

### 6.1 Zusammenfassung

In dieser Arbeit haben wir unser relativistisches Dichtefunktionalprogramm um ein Optimierungsverfahren erweitert, daß es erlaubt, Gleichgewichtsgeometrien von kleinen Molekülen und Atomclustern aufzufinden. Dafür notwendig war die Implementierung des Gradienten der totalen Energie, der sich aus mehreren Anteilen zusammensetzt. Wichtig war insbesondere die Einbeziehung der Korrekturterme zum physikalischen Hellmann-Feynman Gradient, da diese Beiträge etwa in derselben Größenordnung liegen. Nach reichlichem Programmieraufwand ist es gelungen, den Gradienten innerhalb einer akzeptablen Genauigkeit richtig zu berechnen. Ausführliche Untersuchungen an dimeren Systemen zeigen ein konsistentes Verhalten des Gradienten in Bezug auf die totale Energie. Auch in den für uns interessanten relativistischen Systemen ist dieses Verfahren untersucht worden. Erste Rechnungen an dreidimensionalen Atomclustern haben gezeigt, daß dieses Optimierungsverfahren allgemein anwendbar ist. Trotz intensiver Bemühungen ist es bisher leider nicht gelungen, Moleküle, welche aus mehr als drei Atomen bestehen, die zudem noch große Kernladungszahlen haben, zu optimieren. Die Gründe hierfür haben wir am Ende des letzten Kapitels erläutert.

Eine qualitative Verbesserung der Güte unserer bisherigen Rechnungen konnte durch die Verwendung der gradientenkorrigierten Austausch-Korrelationsfunktionale erreicht werden. Weiterhin haben die Ergebnisse der Rechnungen ergeben, daß die Relativistik in diesen Funktionalen selbst bei sehr schweren Molekülen keinen Einfluß auf die Bindungsverhältnisse hat. Wir wollen noch betonen, daß die durchgehende Verwendung der relativistischen Varianten der Funktionale keinen Effizienzverlust des Programmablaufs mit sich bringt. Daher können sie standardmäßig immer benutzt werden.

Wenn die experimentellen Werte der spektroskopischen Parameter einigermaßen zuverlässig sind, so muß die Bedeutung von Gradientenkorrekturen im hochrelativistischen Bereich neu bedacht werden und es scheint, als sei hier die Theorie gefordert.

Im Kapitel 4 bei der Herleitung der Formeln für den Energiegradienten haben wir gesehen, daß sich im Prinzip richtige physikalische Annahmen am Ende mitunter als trügerisch erweisen und in eine Sackgasse führen können. Wir hatten dort gesehen, daß die

Berechnung der Korrekturterme – sofern man überhaupt von Korrekturen sprechen kann; liegen sie doch in derselben Größenordnung wie der eigentlich von der Physik erwartete Effekt – unbedingt notwendig ist.

Alle diese Befunde geben Anlaß für mehrere mögliche Fortsetzungsprojekte. Darauf wollen wir im letzten Abschnitt dieser Arbeit eingehen.

## 6.2 Ausblick und Entwicklungspotential

Nach der zunächst getrennten Umsetzung der ursprünglichen Vorhaben – Geometrieoptimierung und Dichtefunktionalanwendungen – ist der konsequente nächste Schritt deren Kombination. Da sich zeigte, daß trotz der qualitativen Verbesserung der Ergebnisse wie z.B. bei den spektroskopischen Konstanten immer noch Diskrepanzen zu Vergleichswerten bestehen, die nicht allein durch 'Basiseffekte' oder Einflüsse anderer Natur erklärbar sind, sollte über die von uns aus rechentechnischen Gründen gemachte Monopolnäherung im elektronischen Coulombpotential neu nachgedacht werden. Sehr wahrscheinlich lohnt es sich, hier auch höhere Multipole bereitzustellen.

Der große Aufwand der Rechnungen zur Geometrieoptimierung wird hauptsächlich durch die Anzahl der verwendeten Integrationspunkte bestimmt. Die bisher von uns verwendete sehr genaue Integrationsmethode beruhte auf einer komplizierten Zerlegung des Integrationsgebietes, kann aber bei Vorliegen einer Molekülsymmetrie diese erkennen. Läßt man jedoch die Symmetrie außen vor, wie es bei der Geometrieoptimierung der Fall ist, steigt diese Anzahl drastisch an. Daher haben wir die Implementierung einer neuen Integrationsmethode angegangen, die alle auftretenden Ein- und Zweizentren-Matrixelemente in gleicher Weise behandelt [91, 92] und zudem noch mindestens dieselbe Genauigkeit wie das bisher benutzte Integrationsverfahren haben soll.

Im letzten Abschnitt des Kapitels (4) hatten wir einige Überlegungen zur Adaption der Geometrieoptimierung für den Frozen-Core Fall angestellt und angedeutet, daß die Umsetzung mit großer Wahrscheinlichkeit zu einigen Komplikationen führen wird. Die Formulierung des entsprechenden Variationsverfahrens kann sich als noch subtiler erweisen, als es für die Allelektronen-Methode der Fall ist. Eine abgeschwächte und möglicherweise leichter handhabbare Variante wäre aber die Implementierung der Möglichkeit, nur bestimmte Atome zur Optimierung freizugeben. Ein Beispiel wäre die automatische Variation des Abstandes eines Adatoms vor einer Oberfläche, wobei die Positionen der Oberflächenatome festhalten werden. Wir haben bereits damit begonnen, diese physikalische Situation mit unserem Programm zu realisieren und es liegen erste ermutigende Ergebnisse vor [93].

Das Programm wurde bisher erfolgreich bei der Vorhersage von physiko-chemischen Eigenschaften im Bereich der Transaktinide eingesetzt [94]. Mit der nunmehr verbesserten Version, mit der die Möglichkeit besteht, die spektroskopischen Parameter genauer zu berechnen, wollen wir diese Rechnungen fortsetzen.

Sollte dieses Programm eines Tages umgesetzt sein und Optimierungen für atomare Cluster und Moleküle mit etwa 50 Atomen in endlichen Zeiten noch akzeptable Ergeb-

nisse liefern, so wäre eine Adaption bzw. eine Kombination mit molekulardynamischen Methoden [95] oder '*simulated annealing*' Verfahren [96] denkbar.

Abschließend sei noch betont, daß sich die Weiterentwicklung des Programm in allen Fällen lohnt, solange relativistische CI-Rechnungen nicht zum Standard geworden sind.



# Anhang A

## A.1 Hellmann–Feynman Theorem

Die ursprünglichen Formulierungen über die Wirkung zwischenmolekularer Kräfte in quantenmechanischen Vielteilchensystemen gehen auf Hellmann [69] und Feynman [70] zurück. Sie untersuchten die parametrische Abhängigkeiten von Erwartungswerten der Form  $\langle \psi(\lambda) | \hat{H}(\lambda) | \psi(\lambda) \rangle$ , wobei  $\lambda$  irgendein Parameter, in unserem Falle die Anordnung der Atomkerne sein kann. Für ein quantenmechanisches molekulares System gilt  $\hat{H}|\psi\rangle = E|\psi\rangle$ , wenn  $|\psi\rangle$  die exakte Wellenfunktion ist. In diesem Falle ist die Vollständigkeitsrelation  $\langle \psi | \psi \rangle = 1$  erfüllt. Für den Erwartungswert der molekularen Kraft gilt dann

$$\mathcal{F}_\lambda = -\langle \psi | \frac{\partial \hat{H}}{\partial \lambda} | \psi \rangle - \left( \langle \frac{\partial \psi}{\partial \lambda} | \hat{H} | \psi \rangle + \langle \psi | \hat{H} | \frac{\partial \psi}{\partial \lambda} \rangle \right) \quad (\text{A.1})$$

Wegen der Hermitezität von  $\hat{H}$  und  $\hat{H}|\psi\rangle = E|\psi\rangle$  läßt sich (A.1) umschreiben zu

$$E \frac{\partial}{\partial \lambda} \langle \psi | \psi \rangle = E \frac{\partial}{\partial \lambda} (1) = 0 \quad . \quad (\text{A.2})$$

Damit folgt

$$\mathcal{F}_\lambda = -\langle \psi | \frac{\partial \hat{H}}{\partial \lambda} | \psi \rangle \quad . \quad (\text{A.3})$$

Dieses einfache Ergebnis besagt, daß sich die zwischenmolekularen Kräfte allein aus der expliziten Abhängigkeit des Operators ergeben. Wie Slater gezeigt hat, gilt das Hellmann-Feynman Theorem auch in der Dichtefunktionaltheorie [97].

## A.2 Gradientenformel

Hier seien die Ableitungen dreidimensionaler Funktionen  $\phi(\vec{r})$  angegeben, die zur Berechnung der Korrekturterme des Energiegradienten notwendig sind. Dies sind zwei Arten von Funktionen; einmal die Ableitungen der Dirac–Spinoren und andererseits die der Modelldichte. Die Gradientenformel, wie sie schließlich implementiert ist, ist gültig für Funktionen der Gestalt:

$$\phi(\vec{r}) = u(r)Y_{lm}(\Omega) \quad . \quad (\text{A.4})$$

Die Anwendung des Nablakalküls auf Funktionen dieses Typus führt auf die Vektorkugelfunktionen, die unter anderem zur Darstellung transversaler und longitudinaler Vektorfelder, elektrischer und magnetischer Multipolmomente oder Strahlungsintensitäten bei Multipolstrahlung herangezogen werden. Die Vektorkugelfunktionen ihrerseits sind Linearkombinationen der Kugelflächenfunktionen. Die auftretenden Entwicklungskoeffizienten sind gerade die Clebsch–Gordan Koeffizienten zur Spinquantenzahl 1. In der Basis der sphärischen Einheitsensoren sind die Vektorkugelfunktionen durch

$$\begin{aligned} \mathcal{Y}_{jm}^l(\Omega) &= \sum_{m'\sigma} C_{lm'1\sigma}^{jm} Y_{m'}^l(\Omega) \vec{e}_\sigma \\ &= \sum_{m'} \sum_{\sigma} \left( \begin{array}{cc|c} l & 1 & j \\ m' & \sigma & m \end{array} \right) Y_{m'}^l(\Omega) \vec{e}_\sigma \end{aligned} \quad (\text{A.5})$$

gegeben. Das Kopplungsschema ist  $j = l \pm 1$ ;  $\sigma = m - m'$ . Die Einheitsvektoren sind in der sphärischen Basis gegeben durch

$$\vec{e}_0 = \vec{e}_z; \quad \vec{e}_{\pm 1} = \mp \frac{1}{\sqrt{2}} (\vec{e}_x \pm i\vec{e}_y) \quad (\text{A.6})$$

Wegen der Eigenschaften der Clebsch–Gordan Koeffizienten tragen jeweils nur drei Summanden für eine Kombination von  $l$  zu  $j$  (und entsprechend  $m, \sigma, j$ ) zu einer der Vektorkugelfunktionen bei. Der Gradient von (A.4) ist <sup>1</sup>

$$\begin{aligned} \nabla \left( u(r) Y_m^l(\Omega) \right) &= \sqrt{\frac{l}{2l+1}} \left( \frac{d}{dr} + \frac{l+1}{r} \right) u(r) \mathcal{Y}_{lm}^{l-1}(\Omega) \\ &\quad - \sqrt{\frac{l+1}{2l+1}} \left( \frac{d}{dr} - \frac{l}{r} \right) u(r) \mathcal{Y}_{lm}^{l+1}(\Omega) \quad . \end{aligned} \quad (\text{A.7})$$

---

<sup>1</sup>In der Literatur gibt es unterschiedliche Darstellungen der  $\mathcal{Y}_{JM}^L(\Omega)$  und der sphärischen Einheitsvektoren. Wegen des Verhaltens bei Zeitumkehr tritt gelegentlich ein Faktor  $i^l$  auf. Wir halten uns streng an die Referenz [98].

Durch Übergang von der sphärischen zur kartesischen Basis lassen sich die Gradienten bzgl. der kartesischen Koordinaten bilden. Für die einzelnen Komponenten bekommt man:

$$\begin{aligned}
 [\mathcal{Y}_{jm}^l(\Omega)]_x &= \frac{1}{\sqrt{2}} \left\{ \begin{pmatrix} l & 1 & | & j \\ m+1 & -1 & | & m \end{pmatrix} Y_{m+1}^l - \begin{pmatrix} l & 1 & | & j \\ m-1 & 1 & | & m \end{pmatrix} Y_{m-1}^l \right\} \\
 [\mathcal{Y}_{jm}^l(\Omega)]_y &= \frac{i}{\sqrt{2}} \left\{ \begin{pmatrix} l & 1 & | & j \\ m+1 & -1 & | & m \end{pmatrix} Y_{m+1}^l + \begin{pmatrix} l & 1 & | & j \\ m-1 & 1 & | & m \end{pmatrix} Y_{m-1}^l \right\} \quad (\text{A.8}) \\
 [\mathcal{Y}_{jm}^l(\Omega)]_z &= \begin{pmatrix} l & 1 & | & j \\ m & 0 & | & m \end{pmatrix} Y_m^l \quad (\text{A.9})
 \end{aligned}$$

Setzt man die entsprechenden Kombinationen in (A.7) ein, erhält man die Winkelanteile der kartesischen Komponenten des Gradientenoperators. Die verbleibenden Radialableitungen für beide Funktionenklassen lassen sich folgendermaßen beschaffen:

- Radialableitungen der Dirac–Spinoren

Trotz des im Vergleich zur Schrödinger’schen Theorie größeren Aufwandes bietet die relativistischen Behandlung teilweise Vorteile. Ein Beispiel hierfür sind die Radialableitungen der (atomaren) Wellenfunktionen. Zur Bereitstellung der Basis für die Molekülorbitale wird zunächst die Diracgleichung für die beteiligten Atome numerisch an logarithmisch verteilten radialen Stützstellen gelöst. Sie stellt ein Differentialgleichungssystem dar, welches die große ( $f$ ) und kleine ( $g$ ) Komponente sowie deren Ableitungen miteinander verkoppelt <sup>2</sup>.

$$\begin{aligned}
 \frac{df(r)}{dr} &= -\frac{\kappa+1}{r} f(r) - \frac{1}{c} \left( 2m c^2 + \epsilon - \frac{V}{r} \right) g(r) \\
 \frac{dg(r)}{dr} &= \frac{1}{c} \left( \epsilon - \frac{V}{r} \right) f(r) + \frac{\kappa-1}{r} g(r) \quad (\text{A.10})
 \end{aligned}$$

Hierbei ist  $V$  ein zentralsymmetrisches Potential,  $\epsilon$  die Energie und  $c$  die Lichtgeschwindigkeit.  $\kappa$  ist die Dirac-Quantenzahl (vgl. (2.31)). Die linken Seiten in (A.10) sind gerade die gesuchten radialen Ableitungen.

---

<sup>2</sup>Je nach Referenz ist die Diracgleichung unterschiedlich aufgeschrieben, z.B. [99], S. 196 oder [100], S. 65. Sie ist hier so angegeben, wie sie auch im Programm implementiert ist.

- Radialableitungen der Dichte

Für die Berechnung der Dichtefitkorrekturterme zum Energiegradienten werden die Ableitungen der Modelldichte  $\tilde{n}$  benötigt. Da diese aus Radialfunktionen der atomzentrierten Wellenfunktionen gebildet wird, können die eben angegebenen Ableitungen der Radialfunktionen  $f$  und  $g$  dazu benutzt werden, den Radialanteil der Ableitung der Modelldichte zu bestimmen. Dieser ergibt sich dann aus

$$\begin{aligned}\frac{\partial}{\partial r}n(r) &= \frac{\partial}{\partial r} (|f|^2 + |g|^2) \\ &= 2 \left( \frac{\partial}{\partial r} f^\dagger \right) f + 2 \left( \frac{\partial}{\partial r} g^\dagger \right) g.\end{aligned}\quad (\text{A.11})$$

### A.3 Hessematrix-Update

Im Abschnitt (5.2.1) hatten wir gesehen, daß durch Anpassung der Hessematrix an die Potentialenergiefläche die Konvergenz zum Energieminimum hin deutlich beschleunigt werden kann. Das sind die Vorteile des variable-metric Verfahrens, daß wir im Abschnitt (5.1.2) diskutiert hatten. Dort hatten wir auch motiviert, warum in die Update-Formeln die Koordinaten- und Gradientendifferenzen eingehen. Wir geben nun hier die expliziten Formeln zum Update der Hessematrix an [84, 101].

$$\mathbf{H}_k = \mathbf{H}_{k-1} + \frac{\Delta \vec{x}_k \vec{x}_k^T}{\Delta \vec{x}_k^T \cdot \Delta \vec{g}_k} - \frac{\mathbf{H}_{k-1} \Delta \vec{g}_k \vec{g}_k^T \mathbf{H}_{k-1}}{\Delta \vec{g}_k^T \mathbf{H}_{k-1} \Delta \vec{g}_k} + \pi \left( \Delta \vec{g}_k^T \mathbf{H}_{k-1} \Delta \vec{g}_k \Delta \vec{g}_k \vec{w}_k \vec{w}_k^T \right).\quad (\text{A.12})$$

Hierbei bedeuten  $\Delta \vec{x}_k$  und  $\Delta \vec{g}_k$  die Differenzen in den Koordinaten bzw. Gradienten des aktuellen und vorangegangenen Optimierungsschrittes,

$$\Delta \vec{x}_k = \vec{x}_k - \vec{x}_{k-1} \quad , \quad \Delta \vec{g}_k = \vec{g}_k - \vec{g}_{k-1}.\quad (\text{A.13})$$

Die ersten drei Summanden der rechten Seite von (A.12) bilden den sogenannte Davidon-Fletcher-Powell (DFP) Update. Durch Zuschalten des letzten Summand ( $\pi = 1$ ) erhält man den Broyden-Fletcher-Goldfarb-Shanno (BFGS) Update. Die dort zusätzlich auftretenden Faktoren lauten explizit

$$\vec{w}_k = \frac{\Delta \vec{x}_k}{\Delta \vec{x}_k^T \Delta \vec{g}_k} - \frac{\mathbf{H}_{k-1} \Delta \vec{g}_k}{\Delta \vec{g}_k^T \mathbf{H}_{k-1} \Delta \vec{g}_k}.\quad (\text{A.14})$$

In unserem Programmpaket **RELMOS**<sup>3</sup> sind gegenwärtig diese beiden Update-Verfahren implementiert. In den Untersuchungen an den dimeren Systemen ist kein Unterschied

<sup>3</sup>RELativistic MOlecular Structure

in der Abfolge der Optimierungen festgestellt worden.

Steepest-descent, Quasi-Newton und Newton Verfahren zeigen jeweils unterschiedliches Konvergenzverhalten, welches durch die Konvergenzordnung beschrieben werden kann. Eine Folge von Werten  $x_k$  konvergiert gegen  $x^*$ , wenn

$$\lim_{k \rightarrow \infty} \|x_k - x^*\| = 0 \quad . \quad (\text{A.15})$$

Die Folge  $x_k$  konvergiert gegen  $x^*$  mit der Ordnung  $p$ , wenn  $p$  die größte nichtnegative Zahl ist, für die ein endlicher Wert  $\beta$  dergestalt existiert, daß

$$0 \leq \beta \leq \lim_{k \rightarrow \infty} \frac{\|x_{k+1} - x^*\|}{\|x_k - x^*\|^p} \quad . \quad (\text{A.16})$$

Lineare Konvergenz:  $p=1, \beta \leq 1$

Superlineare Konvergenz:  $p=1, \beta = 0$

Quadratische Konvergenz:  $p=2, \beta = 0$  .

## A.4 $\gamma$ -Matrizen und relativistische Notation

Die  $\gamma$  Matrizen sind gegeben durch  $\gamma^i = \beta \alpha^i = \gamma_0 \alpha^i$

$$\gamma_0 = \beta = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} ; \quad \alpha_k = \begin{pmatrix} 0 & \sigma_k \\ \sigma_k & 0 \end{pmatrix} \quad (\text{A.17})$$

$\sigma_k$  sind die Pauli'schen Spinmatrizen:

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} ; \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} ; \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (\text{A.18})$$

Der im nächsten Abschnitt auftretende metrische Tensor  $g_{\mu\nu}$  hat die (kovarianten) Komponenten

$$g_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} ; \quad (\text{A.19})$$

Die Slash-Schreibweise der Vierervektoren hat folgende Bedeutung:

$$\gamma^0 \not{A} = \gamma^\nu A_\nu \quad . \quad (\text{A.20})$$

## A.5 Feldtheoretischer Zugang zur Elektron-Elektron Wechselwirkung

Ziel dieses Abschnittes ist es, den für unsere Rechnungen verwendeten Operator aus Kapitel 2.1 abzuleiten und die fundamentale Elektron-Elektron Wechselwirkung aus der Wechselwirkung des Diracfeldes (welches durch die Elektronen beschrieben wird) mit dem Photonen- (oder Strahlungs)feld zu begründen. Wir legen dabei weniger Wert auf mathematische Strenge; dies kann beispielsweise bei [3] nachvollzogen werden.

Auf sehr fundamentaler Ebene läßt sich jede gegebene physikalische Situation üblicherweise durch eine Lagrangedichte beschreiben. In der klassischen Mechanik geht man von einer Lagrangefunktion aus und gelangt über die Euler-Lagrangegleichungen direkt zu den Bewegungsgleichungen. Desweiteren läßt sich aus der Lagrangefunktion mit einer Legendretransformation die Hamiltonfunktion ableiten, die ebenjenen Bewegungsgleichungen zugrunde liegt. In der Feldtheorie ist dieses Vorgehen völlig analog – jedoch mit dem Unterschied, daß die Feldfunktionen nun Operatoren sind, für die wohlbestimmte Kommutationsrelationen gelten. Die Elektronenwellenfunktionen sind dann durch Feldoperatoren darzustellen.

Betrachten wir unsere physikalische Situation: Elektronen bewegen sich im statischen Feld der Atomkerne. Nehmen wir zuerst freie Elektronen an. Die Lagrangedichte des freien Dirac-Feldes lautet

$$\mathcal{L}_D = \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi \quad , \quad (\text{A.21})$$

aus der sich nach Anwendung des Lagrangeformalismus die Dirac-Gleichung ergibt

$$(i\gamma^\mu \partial_\mu - m)\psi = 0 \quad . \quad (\text{A.22})$$

Eingebettet ist dieses Materiefeld in das (freie) elektromagnetische Strahlungsfeld, dessen Lagrangedichte in Feynman-Eichung <sup>4</sup> durch

$$\mathcal{L}_\gamma = -\frac{1}{4} \hat{F}_{\mu\nu} \hat{F}^{\mu\nu} - \frac{1}{2} \left( \partial_\nu \hat{A}^\nu \right) \quad (\text{A.23})$$

---

<sup>4</sup>Zum Problem der Eichfreiheit sei auf [104], Kapitel 7.1 verwiesen.

mit dem elektromagnetischen Feldstärketensor  $\hat{F}_{\mu\nu} = \partial_\mu \hat{A}_\nu - \partial_\nu \hat{A}_\mu$  gegeben ist. Die Wechselwirkung des Diracfeldes mit dem Photonenfeld  $\hat{A}^\nu$  vollzieht sich mathematisch durch die minimale Substitution

$$\partial_\mu \rightarrow \partial_\mu + ie\hat{A}_\mu \quad . \quad (\text{A.24})$$

Geraten die freien Elektronen in den Einflußbereich eines äußeren Feldes wie beispielsweise dem Kernfeld der Atome  $V_\nu$ , ist zu (A.21) noch der entsprechende Potentialterm einzufügen:

$$\mathcal{L}_D = \hat{\psi}(i\gamma^\mu \partial_\mu - m + \gamma^\mu V_\mu)\hat{\psi} \quad . \quad (\text{A.25})$$

Wir bekommen dann

$$\begin{aligned} & \hat{\psi}(i\gamma^\mu(\partial_\mu + iq\hat{A}_\mu) - m + \gamma^\mu V_\mu)\hat{\psi} \\ = & \hat{\psi}(i\gamma^\mu \partial_\mu - m)\hat{\psi} + \hat{\psi}(\gamma^\mu e\hat{A}_\mu)\hat{\psi} + \hat{\psi}\gamma^\mu V_\mu\hat{\psi} \end{aligned} \quad (\text{A.26})$$

mit der Wechselwirkungsichte

$$\mathcal{L}_{int} = -e\hat{\psi}\gamma^\mu\hat{\psi}(\hat{A}_\mu + V_\mu) \quad (\text{A.27})$$

und der Stromdichte

$$\hat{j}^\nu = -e\hat{\psi}\gamma^\nu\hat{\psi} \quad (\text{A.28})$$

Die Wechselwirkung mit dem Kernpotential kann im klassischen Sinne behandelt werden. Hat zudem das Kernvektorpotential keine raumartigen Anteile ( $V_\nu = (eV_{ext}, \vec{0})$ ), bleibt von dieser Wechselwirkung  $e \int j^0 V_{ext}$  übrig. Der Wechselwirkungs-Hamiltonoperator ist dann

$$\hat{\mathcal{H}}_{int} = e \int \hat{j}^\mu(\vec{r})\hat{A}_\mu(\vec{r})d\vec{r} \quad (\text{A.29})$$

Als Standardverfahren läßt sich die Wechselwirkung aufgrund der Kleinheit der Kopplungskonstanten (in der QED ist dies die Feinstrukturkonstante  $\alpha \propto e^2$ ) störungstheoretisch durch die Streumatrix analog der Aufstellung der Born'schen Reihe bei Streuproblemen im Wechselwirkungsbild behandeln. Diese lautet

$$S = \sum_{n=0}^{\infty} S^{(n)} = \sum_{n=0}^{\infty} \frac{(-i)^n}{n!} \int \dots \int d^4x_1 \dots d^4x_n T\{\hat{H}_{int}(x_1) \dots \hat{H}_{int}(x_n)\} \quad . \quad (\text{A.30})$$

Dabei ist die zeitliche Abfolge in der Anwendung des Störtermes zu beachten, wobei immer die 'früheren' Faktoren rechts von den 'späteren' stehen. Dies wird durch das zeitgeordnete Produkt erreicht, was durch die symbolische Schreibweise  $T\{\hat{H}_{int}(x'_0)\hat{H}_{int}(x''_0)\}$  angedeutet ist. Für die zweite Ordnung <sup>5</sup> folgt daraus

$$S^{(2)} = \frac{(-i)^2}{2!} \int \int d^4x d^4y \left[ \left( (e\hat{\psi}(x)\gamma^\mu\hat{\psi}(x)\hat{A}_\mu(x)) \right) \left( (e\hat{\psi}(y)\gamma^\mu\hat{\psi}(y)\hat{A}_\mu(y)) \right) \right] \quad . \quad (\text{A.31})$$

Zur Auswertung dieses Ausdrucks sind nach Bildung des Erwartungswertes  $\langle f | S | i \rangle$  ( $f$  – final state,  $i$  – initial state) unter Beachtung des Wick-Theorems bestimmte Regeln einzuhalten. Jeder Operator in (A.31) kann durch eine Überlagerung von Erzeugern und Vernichtern dargestellt werden – z.B.  $\psi(x_1) \propto \psi^{(+)}(x_1) + \psi^{(-)}(x_1)$  –, wie es beispielsweise bei der Einführung der Leiteroperatoren beim quantenmechanischen harmonischen Oszillator getan wird. Wegen des Operatorcharakters führen unter allen möglichen Kombinationen in der Reihenfolge ihrer Anwendung nur diejenigen zu nichtverschwindenden Beiträgen, bei denen alle Erzeugungs- und Vernichtungsoperatoren abgepaart sind. Die verschiedenen Kombinationen werden auch als Kontraktionen bezeichnet. Stehen alle Vernichtungsoperatoren rechts von allen Erzeugungsoperatoren, so nennt man diese spezielle Anordnung das normalgeordnete Produkt <sup>6</sup>. Eine dieser Kontraktionen, nämlich die über die Photonen  $A_\mu$ , liefert als Ergebnis gerade den Photonpropagator (0) bezeichne hier den Vakuumzustand)

$$\langle 0 | T(\hat{A}_\mu(x)\hat{A}_\nu(y)) | 0 \rangle = iD_{\mu\nu}(x-y) \quad . \quad (\text{A.32})$$

Dieser ist (in Feynman-Eichung) gegeben durch folgenden Ausdruck:

$$D_{\mu\nu}(x-y) = -g_{\mu\nu} \int \frac{d^4k}{(2\pi)^4} \frac{e^{-ik(x-y)}}{k^2 + i\epsilon} \quad (\text{A.33})$$

Hier ist  $k^2 = k_0^2 - \vec{k}^2$ . Setzt man (A.33) in (A.31) ein, so bekommt man

$$S^{(2)} = \frac{(-i)^2}{2!} \int \int d^4x d^4y j^\nu(x) D_{\mu\nu}(x-y) j^\mu(y) \quad . \quad (\text{A.34})$$

Die explizite Ausführung der (etwas länglichen) Rechnung wollen wir hier nicht vornehmen; dies kann beispielsweise in [105], S.417 ff. nachvollzogen werden. Als Ergebnis ergibt sich ein Ausdruck, der den  $\frac{1}{|x-y|}$ -Operator der Coulombwechselwirkung und den im Kapitel (2.1) angegebenen Breit-Operator enthält.

<sup>5</sup>siehe auch Fußnote 2 des Abschnitts (2.1)

<sup>6</sup>Ausführlich ist dies z.B. in [102], Kapitel 6.3, Seite 104 ff. beschrieben.

## A.6 Die relativistischen Korrekturfunktionen $\Phi_{0/2}^{L/T}$

Hier geben wir die relativistischen Korrekturfunktionen  $\Phi_{0/2}^{L/T}(\beta)$  an, wie sie zur Auswertung der RLDA und RGGGA benötigt werden. Die Funktionen  $\Phi_0^{L/T}$  sind relativistische Korrekturen zur nichtrelativistischen LDA [106, 107]:

$$\Phi_0^L(\beta) = \frac{5}{6} + \frac{1}{3\beta^2} + \frac{2\eta}{3\beta} \operatorname{arcsinh}(\beta) - \frac{2\eta^4}{3\beta^4} \ln(\eta) - \frac{1}{2} \left( \frac{\eta}{\beta} - \frac{\operatorname{arcsinh}(\beta)}{\beta^2} \right)^2 \quad (\text{A.35})$$

$$\Phi_0^T(\beta) = \frac{1}{6} - \frac{1}{3\beta^2} - \frac{2\eta}{3\beta} \operatorname{arcsinh}(\beta) + \frac{2\eta^4}{3\beta^4} \ln(\eta) - \left( \frac{\eta}{\beta} - \frac{\operatorname{arcsinh}(\beta)}{\beta^2} \right)^2 \quad (\text{A.36})$$

Die Variablen  $\eta$  und  $\beta$  sind gegeben durch

$$\beta = \frac{(3\pi^2 n)^{\frac{1}{3}}}{mc}, \quad \eta = (1 + \beta^2)^{\frac{1}{2}} \quad (\text{A.37})$$

wo  $n$  hier die Dichte ist. Die zusätzlich zur Berücksichtigung der Relativistik in der GGA benötigten Korrekturfunktionen  $\Phi_2^{L/T}$  sind durch eine [2/2]-Padé Approximierung gegeben <sup>7</sup>:

$$\Phi_2^L(\beta) = \frac{1 + a_1^L \beta^2 + a_2^L \beta^4}{1 + b_1^L \beta^2 + b_2^L \beta^4}, \quad \Phi_2^T(\beta) = \frac{a_1^T \beta^2 + a_2^T \beta^4}{1 + b_1^T \beta^2 + b_2^T \beta^4},$$

wobei die Koeffizienten  $a, b$  an atomare longitudinale und transversale Energiekomponenten angepaßt sind [33].

Der Vollständigkeit halber geben wir noch die gradientenkorrigierten relativistischen Austauschpotentiale an, die aber in unseren Rechnungen wegen der dafür notwendigen zweiten Ableitungen der Dichte, die programmtechnisch bisher nicht zur Verfügung stehen, nicht benutzt worden sind. Der longitudinale und transversale Anteil des Potentials  $v$  ergibt sich durch entsprechendes Einsetzen von  $\Phi_{0/2}^{L/T}$ :

---

<sup>7</sup>Ein sogenannter Padé-Approximant ist eine rationale Funktion, deren Reihenentwicklung mit einer gegebenen Reihe bis zu deren höchsten Ordnung übereinstimmt. So ist  $F(x)$  eine solche Approximierung zur Reihe  $f(x) = \sum_{k=0}^{\infty} c_k x^k$ , wenn

$$F(x) = \frac{\sum_{k=0}^M a_k x^k}{1 + \sum_{k=1}^N b_k x^k},$$

ausführlicher z.B. in [86], Kapitel 5.12, Seite 194 ff.

$$\begin{aligned}
v_x^{RGGA} = v_x^{LDA}(n) & \left\{ \Phi_0(\beta) + \frac{\beta}{4} \frac{d\Phi_0}{d\beta}(\beta) \right. \\
& + \frac{\beta}{4} \frac{d\Phi_2}{d\beta}(\beta) \left[ g(\xi) - 2\xi \frac{dg}{d\xi} \xi \right] \\
& \left. + \Phi_2(\beta) \left[ g(\xi) - \frac{3}{2}\mu \frac{dg}{d\xi}(\xi) - \frac{3}{2}\tau \frac{d^2g}{d\xi^2}(\xi) \right] \right\}
\end{aligned} \tag{A.38}$$

Hier ist  $v_x^{LDA}(n)$  das nichtrelativistische LDA-Austauschpotential

$$v_x^{LDA}(n) = -\frac{(3\pi^2 n)^{\frac{1}{3}}}{\pi} \tag{A.39}$$

und

$$\mu = \frac{\nabla^2 n}{4(3\pi^2 n)^{\frac{2}{3}} n} \quad ; \quad \tau = \frac{\nabla n \cdot \nabla \xi}{4(3\pi^2 n)^{\frac{2}{3}} n} \quad . \tag{A.40}$$

# Literaturverzeichnis

- [1] M.Schädel, J.V.Kratz, *Physikalische Blätter*, Heft 9, 53.Jhg., Sept. 1997
- [2] E.Engel, S.Keller, A. Facco Bonetti, H.Müller, R.M.Dreizler, *Phys.Rev.A* **52(4)**, 2750, (1995)
- [3] E.Engel, R.M.Dreizler, *Top.Curr.Chem.* **181**,1 (1996)
- [4] M.Born, J.R.Oppenheimer, *Annalen der Physik*, **84**, 457 (1927)
- [5] G.Breit, *Phys.Rev* **34(4)**, 553 (1929)
- [6] H.A.Bethe, E.E.Salpeter, *Quantum Mechanics of One- and Two-Electron Atoms*, Plenum Publishing Corporation, New York, 1977, S.170 ff.
- [7] G.E.Brown, D.G.Ravenhall, *Proc.Roy.Soc.A* **208**, 552 (1951)
- [8] J.Sucher,*Phys.Rev.A*, **22(2)**, 348 (1980)
- [9] E.Wimmer, H.Krakauer,M.Weinert, A.J.Freeman *Phys.Rev.B* **24(2)**, 864 (1981)
- [10] M.Weinert, E.Wimmer, A.J.Freeman *Phys.Rev.B* **26(8)**, 4571 (1982)
- [11] N.D.Lang, *Surface Science* **299/300**, 284 (1993)
- [12] S. Wilke, D. Henning, R. Löber, *Phys.Rev.B* **50(4)**, 2548 (1994)
- [13] J.Bormet, J.Neugebauer, M.Scheffler *Phys.Rev.B* **49(24)**, 17242 (1994)
- [14] P.Hohenberg, W.Kohn, *Phys.Rev.B* **136(3B)**,864 (1964)
- [15] W. Kohn. L.J.Sham, *Phys.Rev.A* **140(4A)**, 1133 (1965)
- [16] R.M.Dreizler, E.K.U. Gross, *Density Functional Theory*, Springer Verlag, 1990
- [17] H.Eschrig, *The Fundamentals of Density Functional Theory*, B.G.Teubner Verlagsgesellschaft, Stuttgart, Leipzig 1996
- [18] E.Engel, S.H.Vosko, *Phys.Rev.B* **47(20)**, 13164 (1993)
- [19] P. Jemmer, P.J.Knowles, *Phys.Rev.A* **51(5)**, 3571 (1995)
- [20] A.K.Rajagopal, J.Callaway, *Phys.Rev.B* **10**, 1 (1973)

- [21] A.H.MacDonald, S.H.Vosko *J.Phys.C* **12**, 2977 (1979)
- [22] P.M.Boerritger, G.TeVelde, E.J.Baerends *Int.J.Quant.Chem.* **33**,87(1988)
- [23] T.Baştuğ, K.Rhasid, W.-D.Sepp, D.Kolb, B.Fricke *Phys.Rev.A***55**, 1760 (1997)
- [24] T.Baştuğ, P.Kürpick, J.Meyer, W.-D.Sepp, B.Fricke, A.Rosén *Phys.Rev.B* **55**, 5015 (1997)
- [25] C.S.Wang, B.M.Klein, H.Krakauer, *Phys.Rev.Lett.* **54**, 1852 (1985)
- [26] P.Bagno, O.Jespen, O.Gunnarson, *Phys.Rev.B* **40(3)**, 1997 (1989)
- [27] H.B.Shore, J.H.Rose, E.Zaremba, *Phys.Rev.B* **15(6)**, 2858 (1977)
- [28] *Density Functional Methods in Physics*; NATO ASI Series B, Vol. 123, edited by R.M.Dreizler and J.da Providencia (Plenum Press, New York, 1985)
- [29] D.C.Langreth, M.J.Mehl, *Phys.Rev.B* **28(4)**, 1809 (1983)
- [30] E.K.U. Gross, R.M.Dreizler, *Density Functional Theory*, Plenum Press, New York, 1995
- [31] P.Pyykkö, *Chem.Rev* **97**, 597 (1997)
- [32] P.Pyykkö, *Chem.Rev* **88**, 563–594 (1988)
- [33] E.Engel, S.Keller, R.M.Dreizler, *Phys.Rev.A* **53(3)**, 1367 (1996)
- [34] S.Chung, S.Krüger, G.Pacchioni, N.Rösch, *J.Chem.Phys.* **102(9)**, 3695 (1995)
- [35] M.Meyer, O.D.Häberlen, N.Rösch, *Phys.Rev.A* **54(6)**, 4775 (1996)
- [36] E. van Lenthe, E.J.Baerends, J.G.Snijders, *J.Chem.Phys.* **101**, 1272 (1994)
- [37] J.P.Perdew, *Phys.Rev.B* **22(12)**, 8822 (1986)
- [38] J.C. Slater, *Phys.Rev.*, **35**, 210 (1930)
- [39] K.Schwarz, *Phys.Rev.B* **5**, 2466 (1972)
- [40] S.H.Vosko, L. Wilk, M.Nusair, *Can.J.Phys.* **58**, 1200 (1980)
- [41] V.Shani, J.Gruenebaum, J.P.Perdew, *Phys.Rev.B* **26(8)**, 4371 (1982)
- [42] P. Mlynarski, D.R.Salahub, *Phys.Rev.B* **43(2)**, 1399 (1991)
- [43] J.P.Perdew, J.A.Chevary, S.H.Vosko, K.A.Jackson, M.R.Pederson, D.J.Singh, C. Fiolhais, *Phys.Rev.B* **46(11)**, 6671 (1992)
- [44] F.Herman, J.P.Van Dyke, I.B.Ortenburger *Phys.Rev.A* **22(16)**, 807 (1969)
- [45] A.E.DePristo, J.D.Kress *J.Chem.Phys.* **86(3)**, 1425 (1987)

- [46] A.D.Becke, *J.Chem.Phys.* **85(12)**, 7184 (1986)
- [47] A.D.Becke, *Phys.Rev.A.* **38(6)**, 3098 (1988)
- [48] E.Engel, S.H.Vosko, *Phys.Rev.B* **47(20)**, 13164 (1993)
- [49] Electronic Structure of Solids 1991,  
ed. P.Ziesche, H.Eschrig, Akademie-Verlag Berlin (1991)
- [50] J.P.Perdew, *Phys.Rev.Lett.* **55**, 1665 (1985)
- [51] J.P.Perdew, Wang Yue, *Phys.Rev.B* **33(12)**, 8800 (1986)
- [52] C.Lee, W.Yang, R.G.Parr, *Phys.Rev.B* **37**, 785 (1988)
- [53] B.G.Johnson, P.M.W.Gill, J.A.Pople *J.Chem.Phys.* **98(7)**, 5612 (1993)
- [54] J.P.Perdew, A.Zunger, *Phys.Rev.B* **23(10)**, 5048 (1981)
- [55] A.K.Rajagopal, *J.Phys.C* **11**, L943 (1978)
- [56] M.V.Ramana, A.K.Rajagopal, *Phys.Rev.A* **24(4)**, 1689 (1981)
- [57] B.Delley, D. E. Ellis, *J.Chem.Phys.* **76(4)**, 1949 (1982)
- [58] J.W.Mintmire, B.I.Dunlap, *Phys.Rev.A.* **25(1)**, 88 (1982)
- [59] B.I.Dunlap, J.W.D.Conolly, J.R.Sabin, *J.Chem.Phys.* **71(12)**, 4993 (1979)
- [60] Turgut Baştuğ, Dissertation, Universität Kassel, 1994
- [61] L.Fan, T.Ziegler, *J.Chem.Phys* **95(10)**, 7401 (1991)
- [62] A.D.Becke, *J.Chem.Phys.* **96(3)**, 2155 (1992)
- [63] K.P.Huber, G.Herzberg, *Molecular spectra and molecular structure I*, (Van Nostrand Reinhold, New York 1950)
- [64] V.Pershina, private Mitteilungen
- [65] P.Schwerdtfeger, M.Dolg, W.H.E. Schwarz, *J.Chem.Phys* **91(3)**, 1762 (1989)
- [66] P.Schwerdtfeger, *Chem.Phys.Lett* **183(5)**, 457 (1991)
- [67] M.Douglas, N.M.Kroll, *Ann.Phys.* **82**, 89 (1974)
- [68] B.A.Hess, *Phys.Rev.A.* **33**, 3742 (1986)
- [69] H.Hellmann, *Einführung in die Quantenchemie*, Deuticke, Leipzig 1937
- [70] R.Feynman, *Phys.Rev.B*, **36**, 340 (1939)
- [71] P.Pulay *Mol.Phys.*, **17**, 197 (1969)

- [72] P.Pulay *J.Chem.Phys.*, **78**, 5044 (1983)
- [73] P.Jørgensen, J.Simons, *J.Chem.Phys.*, **79**, 350 (1983)
- [74] Y.Yamaguchi, Y.Osamura, J.D. Goddard, H.F. Schaefer III, *A New Dimension to Quantum Chemistry*, Oxford Univ. Press 1994
- [75] C. Satoko, *Chem.Phys.Lett.* **83**, 111 (1981)
- [76] C. Satoko, *Phys.Rev.B* **30(4)**, 1754 (1984)
- [77] L.Versluis, T.Ziegler, *J. Chem. Phys.*, **88**, 322 (1988)
- [78] R.Fournier, J.Andzelm, D.R.Salahub, *J.Chem.Phys.* **90(11)**,6371 (1989)
- [79] B. Delley, *J.Chem.Phys* **94(11)**, 7245 (1991)
- [80] P. Funk, *Variationsrechnung und ihre Anwendung*, (2. Aufl.), Springer-Verlag, 1970, Seite 493 ff.
- [81] P.O.Löwdin, *Adv. Phys.* **5**, 51 (1956)
- [82] T.Ziegler, V.Tschinke, E.J.Baerends, J.G.Snijders, W.Raveneck, *J.Phys.Chem.* **93**, 3050 (1989)
- [83] Abschnitt 3.2.1 (Kapitel 8) aus *Methods and Techniques in Computational Chemistry - METECC95*, edited by E.Clementi & G.Corongiu, STEF Cagliari 1995
- [84] K.B.Lipkowitz, D.B.Boyd (Hrsg.) *Reviews in Computational Chemistry*, **Vol.3** (1993)
- [85] J.E.Dennis Jr., R.B.Schnabel, *Numerical Methods for Unconstrained Optimization ana Nonlinear Equations*,  
Classics in Applied Mathematics 16,  
Prentice-Hall, Inc.,Englewood Cliffs,NJ 1983
- [86] J.L. Mohamed, J. Walsh (Hrsg.), *Numerical Algorithms*, Oxford Science Publications
- [87] J.L.Martins, J. Buttet, R.Car, *Phys.Rev.B* **31(4)**, 1804 (1985)
- [88] V.Bonačić-Koutecký, P.Fantucci, J.Koutecký, *Phys.Rev.B* **37(9)**, 4369 (1988)
- [89] L.G.M.Petterson, C.W.Bauschlicher Jr., T.Halicioglu *J.Chem.Phys.* **87(4)**, 2205 (1987)
- [90] A.Fadini, *Molekülkraftkonstanten*, wissenschaftliche Forschungsberichte, Reihe 1A, Band 75, Dr.Dietrich Steinkopff Verlag, Darmstadt 1976
- [91] A.D. Becke, *J.Chem.Phys.* **88(4)**, 2547 (1988)
- [92] J.M.Pérez-Jordá, A.D. Becke, E.San-Fabián, *J.Chem.Phys.* **100(9)**, 6520 (1994)

- [93] D.Geschke, private Mitteilungen
- [94] V.Pershina, B.Fricke, *J.Chem.Phys.* **99(12)**, 9720 (1993)
- [95] R.Car, M.Parrinello, *Phys.Rev.Lett.* **55**, 2471 (1985)
- [96] S.Kirkpatrick, C.D.Gellat, M.P.Vecchi, *Science* **220**, 671 (1983)
- [97] J.C. Slater, *Phys.Rev.* **81(3)**, 385 (1951)
- [98] D.A.Varshalovich, A.N.Moskalev, V.K.Khersonskii *Quantum Theory of Angular Momentum*, World Scietific
- [99] E.M. Rose, *Relativistische Elektronentheorie*, Bibliogr. Institut Mannheim/Wien/Zürich, Band 554a\*, 1964
- [100] J.D. Bjorken, S.D. Drell, *Relativistische Quantenmechanik* Bibliogr. Institut Mannheim/Wien/Zürich, 98/98a\*, 1964
- [101] H.B.Schlegel K.P. Lawley (Ed.), *Ab initio Methods in Quantum Chemistry I*, 249 (1987)
- [102] Franz Mandl, Graham Shaw *Quantenfeldtheorie* (Aula-Verlag Wiesbaden, 1. Aufl. 1993)
- [103] J.D. Bjorken, S.D. Drell, *Relativistische Quantenfeldtheorie* Bibliogr. Institut Mannheim/Wien/Zürich, 101/101a\*, 1964
- [104] W. Greiner, J. Reinhardt *Theoretische Physik Bd. 7a – Feldquantisierung* (Verlag Harri Deutsch, 1. Aufl. 1993)
- [105] W. Greiner, J. Reinhardt *Theoretische Physik Bd. 7 – Quantenelektrodynamik* (Verlag Harri Deutsch, 2. Aufl. 1995)
- [106] B.Jancovici, *Nuovo Cim.***25**, 428 (1962)
- [107] I.A.Akhiezer, S.V.Peletminskii, *Zh. Eksp.Teor.Fiz.***38**, 1829 (1960) (Sov.Phys.JETP **11**, 1316 (1960) )



Für die Anregung zu dieser Arbeit, die Möglichkeit zur freien Gestaltung ihrer Anfertigung und die aufgebrachte Geduld möchte ich meinem Doktorvater Prof. Dr. Burkhard Fricke ganz herzlich danken. Bei ihm habe ich gelernt, daß es neben der fachlichen Komponente auch wesentlich darauf ankommt, die erzielten Ergebnisse in einer angemessenen und klaren Form darzustellen.

Speziell danke ich Herrn Dr. Wolf-Dieter Sepp. Die zahlreichen, oft bis in späte Nachtstunden hinein andauernden erhellenden Diskussionen haben mir eine fundamentalere Herangehensweise an die physikalischen Prinzipien und Probleme erschlossen.

Mein besonderer Dank gilt Herrn Dr. Eberhard Engel. Seine wertvollen Ratschläge, Hilfestellungen und sein Engagement haben den Fortgang meiner Arbeit wesentlich beeinflußt. Das spiegelt sich nicht zuletzt in den unzähligen e-mails wieder, die zwischen Kassel und Frankfurt ausgetauscht wurden.

Herrn Dr. Turgut Baştuğ sei für den Ideenaustausch und die Zusammenarbeit bei der programmtechnischen Umsetzung gedankt.

Vergessen will ich nicht alle weiteren Mitglieder der Arbeitsgruppe Theoretische Physik. Die lockere und unkomplizierte Atmosphäre haben geholfen, so manche frustrierenden Momente mit Gelassenheit zu ertragen.

Gedankt sei dem Hochschulrechenzentrum der Uni Kassel sowie den Rechenzentren der Universität Frankfurt/M., der Technischen Universität Darmstadt und Universität Marburg für die Bereitstellung der Rechenmöglichkeiten innerhalb des Hessischen Hoch- und Höchstleistungsrechnerverbundes (HHLR). Dank auch dem Rechenzentrum der Gesellschaft für Schwerionenforschung (GSI), Darmstadt.

Der Deutschen Forschungsgemeinschaft (DFG) danke ich für die finanzielle Unterstützung.



Hiermit versichere ich, daß ich die vorliegende Dissertation selbständig und ohne unerlaubte Hilfe angefertigt und andere als die in der Dissertation angegebenen Hilfsmittel nicht benutzt habe. Alle Stellen, die wörtlich oder sinngemäß aus veröffentlichten oder unveröffentlichten Schriften entnommen sind, habe ich als solche kenntlich gemacht. Kein Teil dieser Arbeit ist in einem anderen Promotions- oder Habilitationsverfahren verwendet worden.



## Lebenslauf

**Name** : Sven Varga

**Geboren** : 23.01.1970 in Friedrichroda/Thür.

**Eltern** : Erich Varga  
Sieglinde Varga, geb. Loth

### Schulbildung

1976–1986 : Besuch der zehnklassigen allgemeinbildenden Oberschule Hörselgau

1986–1988 : Abitur Salzmannschule Schnepfenthal/Thür.

Nov. 1988–Juli 1989 : Wehrdienst

### Studium

Sep. 1989–Sep. 1991 : Physikstudium an der Technischen Hochschule 'Carl Schorlemmer' Leuna-Merseburg

Sep. 1991 : Vordiplom

Okt. 1991–Aug. 1994 : Studium der Physik an der Universität Kassel

26. August 1994 : Diplomhauptprüfung

### Berufstätigkeit

seit September 1994 : wissenschaftlicher Mitarbeiter in der Arbeitsgruppe Theoretische Physik der Universität Kassel