
Entwicklung eines Treibers zur numerischen
Bestimmung von Minima auf CI-Potenzialflächen
und Anwendung auf angeregte Elektronenzustände
des Uracil-Moleküls

Frank Schneider

Diplomarbeit
Bonn 2002

ENTWICKLUNG EINES TREIBERS ZUR NUMERISCHEN BESTIMMUNG
VON MINIMA AUF CI-POTENZIALFLÄCHEN UND ANWENDUNG AUF
ANGEREGTE ELEKTRONENZUSTÄNDE DES URACIL-MOLEKÜLS

Diplomarbeit in Chemie

Angefertigt am Institut für Algorithmen und Wissenschaftliches Rechnen der
Fraunhofer Gesellschaft, Schloss Birlinghoven, St. Augustin

Vorgelegt der Mathematisch-Naturwissenschaftlichen Fakultät
der Rheinischen Friedrich-Wilhelms-Universität Bonn

von
Frank Schneider

Mai 2002

Ich versichere, dass ich diese Arbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt sowie die Zitate kenntlich gemacht habe.

Bonn, 20. Mai 2002

Referentin	Prof. Dr. Christel M. Marian
1. Korreferentin	Prof. Dr. Sigrid D. Peyerimhoff
2. Korreferent	Prof. Dr. Horst Köppel

Für Nana

Inhaltsverzeichnis

Inhaltsverzeichnis	i
1 Einleitung	1
I Theorie	3
2 Energien	5
2.1 Grundlagen	5
2.2 Born/Oppenheimer-Näherg.	7
2.3 Elektronische Energien	10
2.3.1 Das Variationsprinzip und das Verfahren der Lagrangeschen Multiplikatoren	10
2.3.2 Die Slater-Determinante	12
2.3.3 Linearkombinationen von Orbitalen	14
2.3.4 Hartree-Fock	14
2.3.5 Basissätze	17
Standard-Basissätze, Notationen	18
2.3.6 Self-Consistent-Field	19
Angeregte Konfigurationen	22
2.3.7 Dichtefunktionaltheorie	22
2.3.8 Die Theoreme von Hohenberg und Kohn	23
Die Kohn-Sham-Gleichungen	25
Austausch-Korrelations-Funktionale	26
2.3.9 Vergleich zwischen Hartree-Fock und Kohn-Sham	28
2.3.10 Resolution of the Identity	29
2.3.11 Elektronenkorrelation	31
2.3.12 Dichtefunktionale in CI-Methoden	33
DFT/SCI	33
DFT/MRCI	36

3	Potenzialflächen	37
3.1	Kern-Koordinaten	37
3.2	Eigenschaften	39
3.2.1	Der eindimensionale Fall	39
3.2.2	Der zweidimensionale Fall	39
3.2.3	Der n -dimensionale Fall	42
4	Geometrie-Optimierung	45
4.1	Pfadfinder-Alg.	46
4.2	„Function-Only“	48
4.3	Grad.-Verfahren	49
4.3.1	Die Generierung numerischer Gradienten	50
4.3.2	Steepest Descent	53
4.3.3	Conjugate Gradient	54
4.3.4	Quasi-Newton-Algorithmen	56
4.4	Quad. Verfahren	58
4.5	RFO	59
4.6	(G)DIIS	60
4.6.1	DIIS	60
4.6.2	GDIIS	61
4.7	Eingeschr. Opt.	61
4.7.1	Translation und Rotation	61
4.7.2	Symmetrie	62
4.7.3	Fixierung von Koordinaten	62
II	Implementierung	63
5	Das Programm	65
5.1	Die Programmiersprache	66
5.2	Anbindung an Bestehendes	67
5.3	Programmablauf	68
5.3.1	Das Konvergenzkriterium	70
5.3.2	Die Berechnung der Gradienten	71
5.3.3	Bestimmung der Suchrichtung	72
5.3.4	Die Schrittlänge	73
5.4	Eingabe-Optionen	74
5.5	Ausgabe	75
5.6	Geplante Erweiterungen	76

III	Anwendung	77
6	Formaldehyd	79
6.1	Validierung/Leistung	79
6.1.1	Technische Details	80
6.1.2	Diskussion	81
6.2	Angeregte Zustände	84
6.3	Anregungs-Energien	91
7	Uracil	95
7.1	Technische Details	97
7.2	Der elektronische Grundzustand	98
7.3	Vertikale Anregungsenergien	100
7.4	Angeregte Zustände	104
7.5	Diskussion	106
8	Zusammenfassung & Ausblick	109
A	Notationen	111
	Literaturverzeichnis	113

Kapitel 1

Einleitung

In dieser Arbeit werden Motivation, Implementierung und Anwendung eines Programmes beschrieben, das Geometrie-Relaxation von Molekülen auf anspruchsvollem quantenchemischen Niveau ermöglicht.

Mit „anspruchsvoll“ ist hier im Besonderen ein Multireferenz-CI-Verfahren gemeint, das vor kurzem am Bonner Institut für Physikalische und Theoretische Chemie von Grimme & Waletzke entwickelt wurde (DFT/MRCI). Letzteres erlaubt eine hochwertige Einzelpunkt-Berechnung der Energien von Grund- und angeregten Zuständen mittelgroßer Verbindungen. Die Lücke der Nichtexistenz eines Programmes, vermittels dessen entsprechende Geometrie-Optimierungen durchgeführt werden können, wird mit diesen Arbeiten geschlossen: DFT/MRCI-Energien werden bei systematisch verzerrten Kerngeometrien ausgewertet und einem Optimierungsmodul zugänglich gemacht, das diese sammelt und numerisch molekulare Gradienten berechnet. Da die Gradienten bis auf einen Vorfaktor näherungsweise „Newtonschen“ Kräften auf Atome entsprechen, erlaubt dieser Treiber das Herstellen von Kraftfreiheit von Verbindungen, d.h. eine „Optimierung“ der Geometrie.

Typische Anwendungsgebiete für ein solches Programm liegen in der theoretischen Spektroskopie mehratomiger Verbindungen, was z.B. umfasst:

- die Berechnung von stationären Punkten auf Hyperflächen elektronisch angeregter Spezies (Minima und Sattelpunkte)
- generell eine Interpretation/Aufklärung spektroskopischer Fragestellungen und Voraussagen von Messergebnissen
- die Berechnung adiabatischer Anregungsenergien etc.
- die „Vorarbeit“ zur Berechnung von Übergangsdipolmomenten und ähnlichen *properties*

Neben der Anbindung an das DFT/MRCI-Programm von Grimme/Waletzke ist der Treiber in seinem jetzigen Status aus implementatorischen Gründen an die TURBOMOLE-Module angeschlossen.

Die Arbeit geht in einem theoretischen Teil (I) zunächst auf die Grundlagen von Energiebestimmungen, Potenzialflächen und Geometrieoptimierungen ein, dem eine Schilderung der Implementierung und Routinen (Teil II) folgt. Der Test- und Anwendungsteil (III) schildert zu Beginn erste Ergebnisse am Testmolekül Formaldehyd und dokumentiert dann einen zweiten, etwas aufwändigeren Anwendungsfall, der sich mit der theoretischen Spektroskopie am Uracil-Molekül befasst. Da diese Verbindung als Pyrimidin-Base eine Rolle in der lebendigen Welt spielt und kaum experimentelle Fakten über die angeregten Zustände dieser Verbindung bekannt sind, dürften die Ergebnisse von etwas generellem Interesse sein.

Teil I
Theorie

Kapitel 2

Energien molekularer Systeme

2.1 Grundlagen

Theoretische Chemie ist der Versuch, die chemische Wirklichkeit durch ein mathematisches Modell umfassend darzustellen, so dass sie sich unmittelbar erklären und voraussagen lässt.

Wie uns die Erfahrung lehrt, sind die Eigenschaften und das Verhalten eines Systems eng verknüpft mit energetischen Vorgängen. Gesucht ist daher ein allgemein gültiger Formalismus, mit dessen Hilfe man die Energie jedes beliebigen Systems möglichst exakt bestimmen kann.

Aufgrund quantenmechanischer Erkenntnisse geht man davon aus, dass es möglich ist, den Zustand jedes Systems durch eine Wellenfunktion Ψ vollständig zu beschreiben. Dieser Schritt der Natur in die Mathematik impliziert, dass man nunmehr durch die Anwendung geeigneter mathematischer Mittel (Operatoren) in der Lage ist, bestimmte Informationen (Observable) aus dieser Wellenfunktion zu extrahieren. In der **nicht relativistischen zeitunabhängigen Schrödinger-Gleichung**¹

$$\hat{H}\Psi = E\Psi \tag{2.1}$$

erhält man durch die Anwendung des so genannten Hamilton-Operators \hat{H} auf die Wellenfunktion Ψ die Gesamtenergie E des durch Ψ beschriebenen Systems. Die Beschaffenheit der Wellenfunktion Ψ und des Hamilton-Operators \hat{H} sind Gegenstand der Quantenchemie, die Ermittlung der Energie E eines ihrer Ziele.

Der Hamilton-Operator. Die Energie eines Systems hängt vom Zustand und den Wechselwirkungen der konstituierenden Teilchen ab. In erster Näherung lässt

¹Der allgemeine Fall zeitabhängiger und relativistischer Phänomene für Teilchen mit Spin 1/2, wie es die Elektronen sind, wird in den Dirac-Gleichungen behandelt.

sich der molekulare **Hamilton-Operator** auf folgende Weise in artfremde Bestandteile partitionieren:

$$\hat{H} = \hat{T}_k + \hat{T}_e + \hat{V}_{ke} + \hat{V}_{ee} + \hat{V}_{kk} \quad (2.2)$$

Darin repräsentieren die Operatoren verschiedene Arten denkbarer Energiebeiträge (in atomaren Einheiten):

$$\begin{aligned} \hat{T}_k &= - \sum_{\alpha=1}^N \frac{1}{2M_\alpha} \nabla_\alpha^2 && \text{kinetische Energie der } N \text{ Kerne} \\ \hat{T}_e &= - \sum_{i=1}^n \frac{1}{2} \nabla_i^2 && \text{kinetische Energie der } n \text{ Elektronen} \\ \hat{V}_{ke} &= - \sum_{\alpha=1}^N \sum_{i=1}^n \frac{Z_\alpha}{|\mathbf{r}_i - \mathbf{R}_\alpha|} && \text{Kern-Elektron-Wechselwirkungen} \\ \hat{V}_{ee} &= + \sum_{i=1}^n \sum_{j>i}^n \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} && \text{Elektron-Elektron-Repulsion} \\ \hat{V}_{kk} &= + \sum_{\alpha=1}^N \sum_{\beta>\alpha}^N \frac{Z_\alpha Z_\beta}{|\mathbf{R}_\alpha - \mathbf{R}_\beta|} && \text{Kern-Kern-Repulsion} \end{aligned}$$

Die Wellenfunktion. Die exakte nicht relativistische Wellenfunktion Ψ eines Moleküls ist eine Eigenfunktion des Hamilton-Operators (2.2). Damit sie den Zustand eines Systems tatsächlich vollständig beschreibt, muss sie eine Funktion sämtlicher Einflussgrößen des Systems sein. Bei einem zeitunabhängigen System sind dies die Ortskoordinaten aller Systembausteine, d.h. der Kerne $\{\mathbf{R}_\alpha\}$ und der Elektronen $\{\mathbf{r}_i\}$, sowie weitere Freiheitsgrade $\{s_i\}$ und $\{I_\alpha\}$, die den Spin bezeichnen².

$$\Psi = \Psi(\{\mathbf{r}_i\}; \{s_i\}; \{\mathbf{R}_\alpha\}; \{I_\alpha\}) \quad (2.3)$$

Die Koordinaten \mathbf{r}_i und s_i werden häufig zusammengefasst und als \mathbf{x}_i bezeichnet.

In der Praxis zeigt sich, dass die magnetischen Eigenschaften des Kern- und Elektronenspins in leichten Atomen und Molekülen eine untergeordnete Rolle spielen. In die nicht-relativistische Schrödinger-Gleichung hält der Elektronenspin nur Einzug über das Antisymmetrie-Prinzip (Pauli-Prinzip). Die Operatoren hängen nicht explizit vom Spin der Teilchen ab.

Das Vielkörperproblem. Da alle Teilchen eines Kern-Elektronen-Systems sich gegenseitig beeinflussen, ist die Schrödinger-Gleichung aus dem Blickwinkel der Physik ein hoch komplexes $(3N + 3n)$ -dimensionales $(N + n)$ -Teilchenproblem, das sich in dieser Form nur für wenige kleine Systeme wie etwa das Wasserstoffatom oder

² Der Spin ist eine quantenmechanische Eigenschaft ohne klassisches Analogon.

das Helium-Kation analytisch geschlossen lösen lässt³. Es stellt sich also die Frage nach geeigneten Modifikationen, die die Schrödinger-Gleichung in eine lösbare Form überführen und das System dennoch möglichst fehlerfrei beschreiben.

Die Quantenchemie befasst sich daher auch und vor allem mit der Entwicklung geeigneter Modelle und Näherungsverfahren, die im Wesentlichen Zweierlei leisten müssen:

1. die Überführung der Schrödinger-Gleichung in eine lösbare Form bei minimalem technischen Aufwand und
2. eine möglichst exakte Beschreibung der Wirklichkeit.

Die Praxis zeigt, dass diese beiden Ansprüche meist gegenläufig sind: je aufwändiger ein Formalismus, desto genauer vermag er, chemische Phänomene zu quantifizieren, und je schlichter ein Formalismus, desto geringer die Beanspruchung technischer Ressourcen.

Eine Auflösung des Paradoxons ist nicht in Sicht, und so liegt gerade in der Gratwanderung zwischen diesen Antagonisten eine besondere Herausforderung der Theoretischen Chemie. Dabei ist man aufgefordert, die Auswahl der geeigneten Methode nach den Gegebenheiten des zu kalkulierenden chemischen Systems und der jeweiligen Fragestellung zu treffen.

Von den zahlreichen quantenchemischen Lösungsansätzen seien hier einige beschrieben, die in der vorliegenden Arbeit Verwendung finden.

2.2 Die Born-Oppenheimer-Näherung

Im Born-Oppenheimer-Modell betrachtet man ein chemisches System als eine Anzahl Elektronen, die sich im elektrostatischen Feld räumlich fixierter Kerne bewegen. Da Elektronenbewegungen analog zu ihrem Massenverhältnis etwa vier Größenordnungen schneller erfolgen als Kernbewegungen, reagieren Elektronen instantan auf Veränderungen in der Kern-Geometrie, und so erscheint diese Annahme plausibel.

Formalistisch ist die Born-Oppenheimer Näherung die Trennung von Kern- und Elektronenkoordinaten. Durch sie transportiert man den Strukturbegriff der klassischen Chemie in die Quantenchemie: Die Energie eines chemischen Systems mit fest

³ Als Lösungen erhält man **Atomorbital** genannte Wellenfunktionen ϕ , die bei der Ermittlung höherdimensionaler Systeme noch hilfreich sein werden (siehe Abschnitt 2.3.3, S. 14).

stehender Kern-Geometrie wird als Summe aller energetischer Beiträge der Elektronenbewegungen aufgefasst und berechnet.

Die Entkoppelung von Elektronen- und Kernkoordinaten ermöglicht eine Aufteilung der Schrödinger-Gleichung in einen elektronischen und einen Kern-relevanten Anteil. Unter der Maßgabe einer Betrachtung elektronischer Einflüsse verändert sich die Schrödingergleichung in dreierlei Weise:

1. Der Operator für die kinetische Energie der Kerne \hat{T}_k wird im ersten Schritt vernachlässigt und der Operator der Kern-Abstoßungsenergie \hat{V}_{kk} als konstant angesehen,⁴ so dass sich der Hamilton-Operator auf seine elektrodynamischen Komponenten reduziert:

$$\hat{H}_{el}(\{\mathbf{r}_i\}; \{\bar{\mathbf{R}}_\alpha\}) = \hat{T}_e + \hat{V}_{ke} + \hat{V}_{ee} \quad (2.4)$$

Die Kernkoordinaten $\{\bar{\mathbf{R}}_\alpha\}$ gehen hier — durch einen Balken angedeutet — parametrisch ein, und man spricht vom **elektronischen Hamilton-Operator**.

2. In der Wellenfunktion lassen sich die Elektronen- von den Kernkoordinaten durch einen Produktansatz separieren:

$$\Psi(\{\mathbf{r}_i\}; \{\mathbf{R}_\alpha\}) = \Psi_{el}(\{\mathbf{r}_i\}; \{\bar{\mathbf{R}}_\alpha\})\Psi_k(\{\mathbf{R}_\alpha\}) \quad (2.5)$$

Auch die **elektronische Wellenfunktion** Ψ_{el} hängt signifikant, aber nur noch parametrisch von den fixierten Kernkoordinaten $\{\bar{\mathbf{R}}_\alpha\}$ ab, die Kern-Wellenfunktion Ψ_k beschreibt die Kerndynamik.

3. Als Observable interessiert nur noch die **elektronische Energie** E_{el} , die ebenfalls eine Funktion der Kernkoordinaten ist:

$$E_{el} = E_{el}(\{\bar{\mathbf{R}}_\alpha\}) \quad (2.6)$$

Bei fixierter Kern-Geometrie verschwindet zwar der Anteil kinetischer Kern-Energie, nicht allerdings ein konstanter Beitrag der Kern-Kern-Repulsion, so dass sich die Gesamtenergie E unter diesen Bedingungen wie folgt zusammensetzt:

$$\begin{aligned} E &= E_{el}(\{\bar{\mathbf{R}}_\alpha\}) + E_{kk}(\{\bar{\mathbf{R}}_\alpha\}) \\ &= E_{el}(\{\bar{\mathbf{R}}_\alpha\}) + \sum_{\alpha=1}^N \sum_{\beta>\alpha}^N \frac{Z_\alpha Z_\beta}{|\mathbf{R}_\alpha - \mathbf{R}_\beta|} \end{aligned} \quad (2.7)$$

⁴ Eine additive Konstante ändert die Eigenwerte, nicht aber die Eigenfunktionen eines Operators.

Mit Hilfe dieser Modifikationen erhält man die so genannte **elektronische Schrödinger-Gleichung**

$$\hat{H}_{el}(\{\mathbf{r}_i\}; \{\bar{\mathbf{R}}_\alpha\})\Psi_{el}(\{\mathbf{r}_i\}; \{\bar{\mathbf{R}}_\alpha\}) = E_{el}(\{\bar{\mathbf{R}}_\alpha\})\Psi_{el}(\{\mathbf{r}_i\}; \{\bar{\mathbf{R}}_\alpha\}), \quad (2.8)$$

die die mathematische Komplexität auf ein $3n$ -dimensionales n -Teilchen-Problem verringert. Die Lösungen dieser Differenzialgleichung (für jeweils festgelegte Kern-Geometrien) sind elektronische Wellenfunktionen Ψ_{el} und deren Energie-Eigenwerte E_{el} als Funktion der fixierten Kernkoordinaten $\{\bar{\mathbf{R}}_\alpha\}$. Jede dieser Wellenfunktionen repräsentiert einen elektronischen Zustand.

Analog zu einer elektrodynamischen Betrachtung erlaubt die Koordinaten-Separation auch die Formulierung der **Kern-Schrödinger-Gleichung**, die die Kernbewegungen des Systems im nunmehr konstanten Potenzial der elektronischen Energien E_{el} beschreibt:

$$[\hat{T}_k + E_{el}(\{\mathbf{R}_\alpha\}) + \hat{V}_{kk}(\{\mathbf{R}_\alpha\})]\Psi_k(\{\mathbf{R}_\alpha\}) = E_k\Psi_k(\{\mathbf{R}_\alpha\}) \quad (2.9)$$

Als Lösungen dieser Differenzialgleichung erhält man Wellenfunktionen Ψ_k , die das Schwingungs-, Rotations- und Translationsverhalten der Kerne quantifizieren, und deren Energieeigenwerte E_k .

Der durch die Born-Oppenheimer Näherung entstehende Fehler ist im Allgemeinen vernachlässigbar klein. Erst in strukturellen Grenzbereichen, in denen sich die Energien unterschiedlicher elektronischer Zustände gleicher Symmetrie einander nähern, kommt es zu sprunghaften Änderungen der geometrischen Anordnung aufgrund der beachtlichen Kopplung von Elektronen- und Kernkoordinaten. Die Born-Oppenheimer Näherung verliert hier ihre physikalische Grundlage und verursacht unter Umständen erhebliche Fehlkalkulationen.

Die Born-Oppenheimer-Näherung ist weitgehend akzeptiert, so dass die Quantenchemie ihre Anstrengungen in der Regel auf Ansätze zur Lösung der elektronischen Schrödinger-Gleichung konzentriert. Sie erweitert das Spektrum geschlossen lösbarer Systeme von Atomen auf Moleküle mit einem Elektron und ermöglicht so die exakte Berechnung etwa des Wasserstoff-Molekül-Kations H_2^+ . Die daraus berechneten Wellenfunktionen sind präzise Beschreibungen der Bewegungen eines Elektrons im Mehr-Kern-Feld und heißen Molekülorbitale (MO).

2.3 Elektronische Energien

2.3.1 Das Variationsprinzip und das Verfahren der Lagrangeschen Multiplikatoren

Eine zentrale Fragestellung der Quantenchemie ist die Lösung der zeitunabhängigen Schrödingergleichung.

$$\hat{H}|\Phi\rangle = E|\Phi\rangle \quad (2.10)$$

Die Lösung dieser Differenzialgleichung ist ein prinzipielles mathematisches Problem und für die überwiegende Mehrheit aller Fälle nicht linear lösbar. Dennoch existiert unter der Voraussetzung, dass der (hermitesche) Hamilton-Operator \hat{H} bekannt ist, ein unendlicher Satz exakter orthogonaler Wellenfunktionen $\{|\Phi_\alpha\rangle\}$, die die obige Eigenwertgleichung erfüllen:

$$\hat{H}|\Phi_\alpha\rangle = E_\alpha|\Phi_\alpha\rangle \quad \text{mit } \alpha = 0, 1, 2, \dots \quad (2.11)$$

Die zugehörigen Energie-Eigenwerte folgen der Monotonie

$$E_0 \leq E_1 \leq E_2 \leq \dots \leq E_\alpha \leq \dots \quad (2.12)$$

Wichtige Konsequenz dieser Reihe ist, dass es eine untere Grenze E_0 gibt, die man günstigstenfalls erreichen, nicht aber unterschreiten kann. Die entsprechende Wellenfunktion Φ_0 identifiziert man mit dem Grundzustand.

Dies führt zu einer bedeutenden These, die als **Variationsprinzip** bekannt ist: Keine normalisierte Wellenfunktion $|\tilde{\Phi}\rangle$, die die Randbedingungen der Eigenwertgleichung erfüllt, kann mit ihrem Eigenwert \tilde{E} die untere Grenze von E_0 unterschreiten.

$$\langle \tilde{\Phi} | \hat{H} | \tilde{\Phi} \rangle \geq E_0 \quad (2.13)$$

Das bedeutet, dass man von einer (beinahe) beliebigen Testfunktion $|\tilde{\Phi}\rangle$ ausgehen und diese wie auch immer verändern/variieren kann, bis keine weitere Verringerung der Energie mehr möglich erscheint. Die Qualität einer Wellenfunktion wird an der Tiefe ihres Energie-Eigenwertes gemessen.

Das mathematische Mittel der systematischen linearen Variation einer Wellenfunktion unter der Maßgabe der Eigenwertminimierung ist das **Verfahren der Lagrangeschen Multiplikatoren**. Dabei wird die Testfunktion $|\tilde{\Phi}\rangle$ als Linearkombination aus einer festgelegten endlichen Basis n orthonormaler Wellenfunktionen $\{|\Psi_i\rangle\}$ angenähert:

$$|\tilde{\Phi}\rangle = \sum_{i=1}^n c_i |\Psi_i\rangle \quad (2.14)$$

Die Matrix-Darstellung des Hamilton-Operators in dieser Basis führt zu einer symmetrischen $n \times n$ -Matrix mit den Elementen

$$\begin{aligned} H_{ij} &= \langle \Psi_i | \hat{H} | \Psi_j \rangle \\ &= H_{ji} \end{aligned} \quad (2.15)$$

Die Testfunktion ist normalisiert, so dass gilt

$$\langle \tilde{\Phi} | \tilde{\Phi} \rangle = \sum_{ij} c_i c_j \langle \Psi_i | \Psi_j \rangle = \sum_i c_i^2 = 1 \quad (2.16)$$

Der zu minimierende Erwartungswert der Energie ergibt sich demnach als eine Funktion der Expansionskoeffizienten $\{c_i\}$:

$$E(\{c_i\}) = \langle \tilde{\Phi} | \hat{H} | \tilde{\Phi} \rangle = \sum_{ij} c_i \langle \Psi_i | \hat{H} | \Psi_j \rangle c_j = \sum_{ij} c_i c_j H_{ij} \quad (2.17)$$

Unglücklicherweise lässt sich dieser Ausdruck nicht wie gewünscht nach den Koeffizienten differenzieren, da diese nicht voneinander unabhängig sind. Ein formalistischer Trick führt dennoch zum Ziel: Man subtrahiert *de facto* eine Null von dem Erwartungswert aus Gleichung 2.17, schlüsselt diese geschickt auf und erhält so eine modifizierte Funktion L , die außer von den Koeffizienten $\{c_i\}$ noch von dem unbestimmten Multiplikator E abhängt:

$$\begin{aligned} L(\{c_i\}, E) &= \langle \tilde{\Phi} | \hat{H} | \tilde{\Phi} \rangle - E \left(\langle \tilde{\Phi} | \tilde{\Phi} \rangle - 1 \right) \\ &= \sum_{ij} c_i c_j H_{ij} - E \left(\sum_i c_i^2 - 1 \right) \end{aligned} \quad (2.18)$$

Was haben wir dadurch gewonnen? Einen zusätzlichen unabhängigen Parameter E , den wir zur Einflussnahme nutzen können. Wo wollen wir Einfluss nehmen? Bei der Differenziation. Und das geht so:

Ohne Beschränkung der Allgemeinheit nehmen wir an, dass ein beliebiger Expansionskoeffizient c_n von allen anderen Koeffizienten abhängt, diese aber seien voneinander unabhängig. Dann gilt

$$\frac{\partial L}{\partial c_k} = 0 \quad \text{für alle } k = 1, 2, \dots, n-1 \quad (2.19)$$

Die Ableitung von L nach c_n ist nicht notwendigerweise Null, also wählen wir den unbestimmten Lagrangeschen Multiplikator E gerade so, dass auch die Ableitung nach c_n verschwindet, und es gilt

$$\frac{\partial L}{\partial c_k} = 0 \quad \text{für alle } k = 1, 2, \dots, n-1, n \quad (2.20)$$

Dies ist eine linear lösbare Bestimmungsgleichung für die minimierenden Expansionskoeffizienten $\{c_k\}$. Sie führt unmittelbar zu einer Formulierung des Eigenwertproblems der Hamilton-Matrix \mathbf{H} :

$$\mathbf{H}\mathbf{c} = E\mathbf{c} \quad (2.21)$$

mit \mathbf{c} als Koeffizienten-Vektor.

Da die Hamilton-Matrix \mathbf{H} symmetrisch ist, findet man n orthonormale Eigenvektoren \mathbf{c}^α und n zugehörige Eigenwerte E^α . Jeder dieser Vektoren definiert eine andere Linearkombination der Basisfunktionen. Man definiert daher eine Matrix der Koeffizienten \mathbf{C} mit Elementen $C_{i\alpha} = c_i^\alpha$, eine weitere Matrix \mathbf{E} mit den Diagonalelementen E^α und verallgemeinert so Gleichung 2.21 auf die Gesamtheit möglicher Lösungen $|\tilde{\Phi}_\alpha\rangle$:

$$\mathbf{H}\mathbf{C} = \mathbf{E}\mathbf{C} \quad (2.22)$$

Summa summarum reduziert man durch Anwendung des Variationsprinzips und des Verfahrens der Lagrangeschen Multiplikatoren die Lösung einer beliebig komplizierten Differenzialgleichung auf das Lösen eines linearen Gleichungssystems. Die Effizienz der Methode bezahlt man mit einer Näherung, und so hängt die Präzision der Ergebnisse von der Größe und der Qualität der Basis $\{|\Psi_i\rangle\}$ ab. Bei Einsatz einer fiktiven unendlich großen Basis erhält man die theoretische Untergrenze des Algorithmus E_0 .

2.3.2 Die Slater-Determinante

Die Wellenfunktion $\Psi_{el}(\{\mathbf{r}_i\}; \{\bar{\mathbf{R}}_\alpha\})$ der elektronischen Schrödinger-Gleichung ist nur für Systeme mit einem einzelnen Elektron exakt bestimmbar. Sobald ein Atom oder Molekül mehr als ein Elektron enthält, steht man vor dem Problem, die Bewegung eines bestimmten Elektrons in Abhängigkeit von den Bewegungen aller anderen Elektronen beschreiben zu müssen. Bewegungen dreier sich wechselseitig beeinflussenden Körper nennt man Dreikörperproblem und konstatiert, dass es sich nicht analytisch lösen lässt. Um dennoch zu einer angenäherten Wellenfunktion zu gelangen, geht man zunächst von einem recht groben Ansatz aus und versucht dann, diesen schrittweise zu verfeinern.

Ignoriert man in einem ersten Schritt die wechselseitig repulsive Beeinflussung der Elektronen des Systems, so macht es Sinn, die Gesamtwellenfunktion als ein Produkt von Ein-Elektron-Spinorbitalen χ zu formulieren:

$$\Psi_{el}^{\text{HP}} = \chi_i(\mathbf{x}_1)\chi_j(\mathbf{x}_2)\dots\chi_k(\mathbf{x}_n) \quad (2.23)$$

Diese Darstellung einer Wellenfunktion Ψ_{el}^{HP} wird **Hartree-Produkt** genannt.

Das Hartree-Produkt unterscheidet zwischen individuellen Elektronen, was natürlich nicht der physikalischen Wirklichkeit entspricht. Außerdem erfüllt es nicht die Forderung des Antisymmetrie-Prinzips, wonach eine Fermionen-Wellenfunktion beim Austausch zweier Teilchen ihr Vorzeichen ändern muss:

$$\Psi_{el}(\mathbf{x}_1, \dots, \mathbf{x}_i, \dots, \mathbf{x}_j, \dots, \mathbf{x}_n) = -\Psi_{el}(\mathbf{x}_1, \dots, \mathbf{x}_j, \dots, \mathbf{x}_i, \dots, \mathbf{x}_n) \quad (2.24)$$

Das folgende Beispiel einer Wellenfunktion zweier Elektronen ist eine Linearkombination zweier Hartree-Produkte und erfüllt das Antisymmetrie-Prinzip:

$$\begin{aligned} \Psi_{el}(\mathbf{x}_1, \mathbf{x}_2) &= \frac{1}{\sqrt{2}}[\Psi_{12}^{\text{HP}}(\mathbf{x}_1, \mathbf{x}_2) - \Psi_{21}^{\text{HP}}(\mathbf{x}_1, \mathbf{x}_2)] \\ &= \frac{1}{\sqrt{2}}[\chi_i(\mathbf{x}_1)\chi_j(\mathbf{x}_2) - \chi_j(\mathbf{x}_1)\chi_i(\mathbf{x}_2)] \end{aligned} \quad (2.25)$$

Der Faktor $\frac{1}{\sqrt{2}}$ ist der Normalisierungsfaktor für eine Linearkombination zweier Wellenfunktionen.

Die Verallgemeinerung dieser Wellenfunktion gelingt mit den Mitteln der Linearen Algebra in Form einer Determinante, der so genannten **Slater-Determinante**:

$$\Psi_{el}(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n) = \frac{1}{\sqrt{n!}} \begin{vmatrix} \chi_i(\mathbf{x}_1) & \chi_j(\mathbf{x}_1) & \cdots & \chi_k(\mathbf{x}_1) \\ \chi_i(\mathbf{x}_2) & \chi_j(\mathbf{x}_2) & \cdots & \chi_k(\mathbf{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \chi_i(\mathbf{x}_n) & \chi_j(\mathbf{x}_n) & \cdots & \chi_k(\mathbf{x}_n) \end{vmatrix}, \quad (2.26)$$

in der prägnanteren Dirac-Schreibweise:

$$\Psi_{el}(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n) = |\chi_i(\mathbf{x}_1)\chi_j(\mathbf{x}_2)\cdots\chi_k(\mathbf{x}_n)\rangle, \quad (2.27)$$

bei festgelegter Nummerierung der Elektronen:

$$\Psi_{el}(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n) = |\chi_i\chi_j\cdots\chi_k\rangle \quad \text{oder kurz} \quad |ij\cdots k\rangle \quad (2.28)$$

Die Menge der verwendeten Spinorbitale $\{\chi_i\}$ wird Basissatz genannt. In der Slater-Determinante entsprechen die Reihen den Elektronen und die Spalten den Spinorbitalen. Jede mögliche Vertauschung von Koordinaten \mathbf{x} wird korrekt antisymmetrisch berücksichtigt.

Durch die Erfüllung des Antisymmetrie-Prinzips importiert man eine weitere bedeutende Zusatzinformation. Denn wie sich zeigen lässt (Szabo und Ostlund [1996]), ist für Wellenfunktionen vom Slater-Determinanten-Typ das Antreffen zweier Elektronen gleichen Spins am selben Ort unmöglich. Diesen Umstand nennt man Austauschkorrelation. Der Elektronen umgebende Raum, der ein Eindringen eines Elektrons gleichen Spins ausschließt, heißt Fermi-Loch (Kutzelnigg [1975]).

2.3.3 Linearkombinationen von Orbitalen

Die Qualität einer Wellenfunktion, die als Slater-Determinante formuliert wird, hängt von ihrer Vollständigkeit ab. Eine Slater-Determinante mit unendlich großem Basissatz $\{\chi_\infty\}$ wäre im Rahmen ihrer Näherungen eine perfekte Wellenfunktion. Es liegt auf der Hand, dass unendliche Slater-Determinanten unrealisierbar sind, so dass man nach einer ökonomischeren Methode der Basissatz-Konstruktion suchen muss.

Als besonders günstig für die Erzeugung von Molekül-Raum-Orbitalen $\psi_i(\mathbf{r})$ haben sich Linearkombinationen der entsprechenden Atom-Orbitale $\phi_\mu(\mathbf{r})$ erwiesen,

$$\psi_i(\mathbf{r}) = \sum_{\mu} c_{\mu i} \phi_{\mu}(\mathbf{r}), \quad (2.29)$$

da Letztere bei der Bindungsbildung im Wesentlichen erhalten bleiben. Auch hier gilt, dass eine möglichst vollständige Auswahl der zu kombinierenden Funktionen die Qualität der kombinierten Funktion erhöht. Diese Methode nennt man LCAO-MO-Ansatz.⁵ Auf diesen Punkt wird im Zusammenhang mit den Roothaan-Hall-Gleichungen noch zurück gekommen werden (S. 19).

Die entstehenden Molekülorbitale müssen eine orthonormale Basis bilden, eine Bedingung, die man allgemein mit dem Kronecker-Delta δ_{ij} formuliert:⁶

$$\langle \psi_i | \psi_j \rangle = \delta_{ij} \quad , \quad (2.30)$$

Atomorbitale ihrerseits können zwar normiert sein, sind aber i.d.R. nicht orthogonal, da sie an unterschiedlichen Atomzentren lokalisiert sind. Statt dessen ergeben sich so genannte Überlappungsintegrale

$$\langle \phi_{\mu} | \phi_{\nu} \rangle = S_{\mu\nu} \quad . \quad (2.31)$$

2.3.4 Hartree-Fock

Auch der elektronische Hamilton-Operator

$$\hat{H}_{el} = \hat{T}_e + \hat{V}_{ke} + \hat{V}_{ee} \quad (2.32)$$

⁵LCAO-MO = Linear Combination of Atomic Orbitals - Molecular Orbitals

⁶Zur Definition des Kronecker-Deltas: seien Ψ_i und Ψ_j orthonormale Wellenfunktionen, dann gilt:

$$\langle \Psi_i | \Psi_j \rangle = \delta_{ij} = \begin{cases} 0, & \text{falls } i \neq j \\ 1, & \text{falls } i = j \end{cases}$$

enthält einen problematischen Zweikörperterm

$$\hat{V}_{ee} = \sum_{i=1}^n \sum_{j>i}^n \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}, \quad (2.33)$$

der die gegenseitige Abstoßung der Elektronen berücksichtigt.

Um einer geeigneten Approximation näher zu kommen, betrachten wir zunächst ein Ein-Elektron-System. Der Hamilton-Operator lautet in diesem Fall

$$\hat{h}_{el}(1) = \hat{t}_e(1) + \hat{v}_{ke}(1) \quad (2.34)$$

gleichbedeutend mit der Annahme, dass sich die elektronische Energie des Systems additiv aus den Beiträgen kinetischer und potenzieller Energie des Elektrons 1 im elektrostatischen Feld der Kerne zusammensetzt. Bringt man weitere Elektronen in dieses System ein, so kommen Wechselwirkungsbeiträge der Elektronen untereinander hinzu.

Kernaussage der Hartree-Fock-Näherung ist nun, dass der Wechselwirkungseinfluss, unter dem ein beliebiges Elektron i in einem System mit n Elektronen steht, sich als effektiv für dieses Elektron „sichtbares“ statisches Potenzial $v^{\text{HF}}(i)$ darstellen lässt:

$$\hat{v}_{ee}(i, j) := v^{\text{HF}}(i) \quad (2.35)$$

Das Potenzial $v^{\text{HF}}(i)$ heißt **Hartree-Fock-Potenzial**. Addiert man dieses Potenzial in Gleichung 2.34 ein, erhält man den so genannten **Fock-Operator**:

$$\begin{aligned} \hat{f}(i) &= \hat{t}_e(i) + \hat{v}_{ke}(i) + \hat{v}_{ee}^{\text{HF}}(i) \\ &= -\frac{1}{2}\nabla_i^2 - \sum_{\alpha=1}^N \frac{Z_\alpha}{|\mathbf{r}_i - \mathbf{R}_\alpha|} + v^{\text{HF}}(i) \end{aligned} \quad (2.36)$$

Analog formuliert man die so genannte **Hartree-Fock-Gleichung** als Ein-Elektron-Eigenwert-Gleichung:

$$\hat{f}(i)\chi(\mathbf{x}_i) = \epsilon_{el}\chi(\mathbf{x}_i) \quad (2.37)$$

Bei genauer Kenntnis des Hartree-Fock-Potenzials für jedes Elektron i des Systems lässt sich die Hartree-Fock-Gleichung lösen, und man erhält Orbitale χ und Orbitalenergien ϵ . Leider ist das Potenzial selbst-rekursiv definiert und zunächst völlig unbekannt. Es bedarf geeigneter Bestimmungsverfahren, von denen in Abschnitt 2.3.6 (S. 19) noch zu sprechen sein wird. Dennoch erlaubt die Tatsache, dass in Hartree-Fock-basierten Methoden in der Regel Slater-Determinanten zum Einsatz kommen, eine genauere Eingrenzung des Hartree-Fock-Potenzials.

Die Anwendung des Zwei-Elektronen-Operators $\hat{v}_{ee}(i, j)$ aus Gleichung 2.35 auf eine Slater-Determinante führt zu den folgenden Integralen:

$$\begin{aligned} \langle \Psi_0 | r_{ij}^{-1} | \Psi_0 \rangle &= \int d\mathbf{x}_1 d\mathbf{x}_2 \chi_i^*(\mathbf{x}_1) \chi_j^*(\mathbf{x}_2) r_{ij}^{-1} \chi_i(\mathbf{x}_1) \chi_j(\mathbf{x}_2) \\ &\quad - \int d\mathbf{x}_1 d\mathbf{x}_2 \chi_i^*(\mathbf{x}_1) \chi_j^*(\mathbf{x}_2) r_{ij}^{-1} \chi_j(\mathbf{x}_1) \chi_i(\mathbf{x}_2) \\ &= \langle ij | ij \rangle - \langle ij | ji \rangle = \langle ij || ij \rangle \quad (\text{Dirac-Notation}) \end{aligned} \quad (2.38)$$

Das erste Integral $\langle ij | ij \rangle$ heißt Coulomb-Integral, das zweite $\langle ij | ji \rangle$ wird Austausch-Integral genannt. Daraus leitet man entsprechende Operatoren ab und erhält aus den Wechselwirkungen zwischen Elektron 1 in Spinorbital χ_i und Elektron 2 in χ_j den

$$\text{Coulomb-Operator} \quad \hat{J}_j(1) = \int d\mathbf{x}_2 |\chi_b(2)|^2 r_{12}^{-1} \quad (2.39)$$

und den

$$\text{Austausch-Operator} \quad \hat{K}_j(1) \chi_i(\mathbf{x}_1) = \left[\int d\mathbf{x}_2 \chi_j^*(\mathbf{x}_2) r_{12}^{-1} \chi_i(\mathbf{x}_2) \right] \chi_j(\mathbf{x}_1) \quad (2.40)$$

Der Coulomb-Operator quantifiziert die Repulsion der Elektronen, gewichtet mit der Aufenthaltswahrscheinlichkeit $d\mathbf{x}_2 |\chi_b(2)|^2$ des zweiten Elektrons im Raumelement $d\mathbf{x}_2$ bei \mathbf{x}_2 . Der konstruktive Beitrag des Austausch-Operators ist weniger plausibel, denn er ist ein formalistisches Abstraktum, das sich aus der antisymmetrisierten Form der Slater-Determinante beim Platzwechsel zweier Elektronen gleichen Spins ergibt.

Betrachten wir ein einzelnes Elektron, so müssen Coulomb- und Austausch-Operator über alle möglichen paarweisen Wechselwirkungen mit anderen Elektronen summiert werden:

$$\left[\hat{h}(1) + \sum_{j \neq i} \hat{J}_j(1) - \sum_{j \neq i} \hat{K}_j(1) \right] \chi_i(1) = \epsilon_i \chi_i(1) \quad (2.41)$$

Da die Wechselwirkungen eines Elektrons mit sich selbst gemäß

$$\left[\hat{J}_i(1) - \hat{K}_i(1) \right] \chi_i(1) = 0 \quad (2.42)$$

verschwinden, kann man auf die Restriktion $j \neq i$ bei der Summation verzichten. Der effektive Ein-Elektron-Fock-Operator lässt sich also schreiben als

$$\hat{f}(1) = \hat{h}(1) + \sum_j \hat{J}_j(1) - \sum_j \hat{K}_j(1) \quad , \quad (2.43)$$

und auch das Hartree-Fock-Potenzial ist jetzt genauer definiert als

$$v^{\text{HF}}(1) = \sum_j \hat{J}_j(1) - \sum_j \hat{K}_j(1) \quad . \quad (2.44)$$

2.3.5 Basissätze

Unter einem Basissatz versteht man eine Menge von Spin- oder Raum-Wellenfunktionen $\{\chi_i\}$ oder $\{\Psi_i\}$, in die man eine Wellenfunktion entwickelt und die so als Grundlage für quantenchemische Berechnungen dienen. Der übliche Weg führt über Linearkombinationen von Atomorbitalen ϕ_i zu Molekül-Raum-Orbitalen Ψ_i , Spin-Orbitalen χ_i und schließlich via Slater-Determinante zu einer Mehr-Elektronen-Wellenfunktion Φ .

Die Verwendung von Atomorbitalen als „Startpunkt“ liegt nahe, weil sie aus geschlossenen Lösungen der elektronischen Schrödinger-Gleichung bekannt sind. Die Radialteile dieser als **Slater-Typ** (STO) kategorisierten Funktionen haben die allgemeine Gestalt

$$F = Ne^{-\zeta r}. \quad (2.45)$$

Wellenfunktionen dieses Typs bereiten grosse Probleme bei der Berechnung von in quantenchemischen Algorithmen häufig auftretenden Mehrzentren-Integralen. Funktionen vom **Gauß-Typ** (GTO)

$$G = Ne^{-\alpha r^2} \quad (2.46)$$

dagegen sind formalistisch unkomplizierter, letztendlich weil die Ableitung einer Gauß-Funktion wieder eine Gauß-Funktion ist. Bedauerlicherweise sind Gauß-Funktionen nur eine schlechte Annäherung der „realen“ Slater-Typ-Orbitale. Insbesondere die Modellierung des so genannten „cusp“, der Spitze des Grafen der Elektronen-Aufenthaltswahrscheinlichkeit am Kernort, gelingt mit einer Gauß-Funktion nicht. Durch Linearkombinationen von Gauß-Funktionen kann man Slater-Orbitalen jedoch befriedigend nahe kommen:

$$F \approx \sum_i c_i G_i \quad (2.47)$$

Eine solche Linearkombination wird **kontrahierte Gauss-Funktion** (CGTO) genannt, weil man mehrere Funktionen zusammenfasst, um eine Slater-Funktion zu reproduzieren. Die Kontraktionskoeffizienten c_i sind atomspezifisch und gehen parametrisch fixiert in Folgerechnungen ein.

Standard-Basissätze, Notationen

Die Wahl des Basissatzes ist ein sensibler Punkt quantenchemischer Kalkulationen. Kleinere Basissätze beschleunigen die Berechnungen merklich, liefern aber häufig nur quantitative Aussagen, während größere Basissätze zwar meist präzisere Ergebnisse liefern, allerdings zum Preis eines erheblichen Rechenaufwandes. Durch die in vielen Molekülberechnungen gewonnene Erfahrung entwickelten sich einige Standard-Basissätze, die sich dadurch auszeichnen, dass sie aus einer verhältnismäßig kleinen Zahl gut ausgewählter Basisfunktionen bestehen und so bei überschaubarem Rechenaufwand sehr gute Ergebnisse liefern.

Ein erster einfacher Ansatz ist die Verwendung einer so genannten minimalen Basis mit nur einem Satz von Basisfunktionen für jede besetzte Atomshale. (Unter einem Satz von Basisfunktionen versteht man die Funktionen, die sich nur durch den winkelabhängigen Anteil unterscheiden, also bspw. p_x p_y p_z .) Werden n Gauß-Funktionen kontrahiert, um einen Slater-Typ zu reproduzieren, so bezeichnet man den Basissatz als STO- n G (n ist meist 3 oder 6). Basissätze dieses Typs sind relativ „klein“. Dass mit ihnen in manchen Fällen gute Ergebnisse erzielt werden, muss auf eine Kompensation entgegengesetzter Fehler zurückgeführt werden.

Eine deutliche Verfeinerung erfährt ein Basissatz durch die Verwendung mehrerer kontrahierter Gauß-Funktionen mit unterschiedlichen Exponenten ζ . Diese Basissätze nennt man dementsprechend Double-Zeta- bzw. Triple-Zeta-Basissätze (DZ, TZ). Oft beschränkt man diesen Aufwand auf die Valenzelektronen und benutzt so genannte *Split-valence*-Basissätze. Ein Beispiel: die Notation 4-31G bedeutet: man kontrahiert vier GTOs zu einer CGTO für die Rumpforbitale und drei bzw. eine GTO(s) zu zwei CGTOs für die Valenzorbitale. Geometrien, die mit diesen Basissätzen erhalten werden, zeigen schon recht gute Übereinstimmung mit experimentell ermittelten.

Um die Polarisation der Ladungsverteilung bei einer Bindungs-Bildung besser modellieren zu können, hat man zusätzliche Polarisations-Funktionen eingeführt. Dazu eignen sich Funktionen mit höherem Drehimpuls, also d -Funktionen für die Atome der ersten Langperiode (6-31G*) und zusätzliche p -Funktionen für Wasserstoff-Atome (6-31G**). Basissätze dieses Typs werden als DZP, TZP, TZPP, etc. notiert. In Molekülen, die durch Konjugations- und Polarisations-Effekte stark beeinflusst werden, sind Basissätze mit Polarisations-Funktionen für eine korrekte Beschreibung unerlässlich.

Die bisher genannten Notationen für Basissätze sind relativ unsystematisch und oft ungenau, so dass man sie häufig auch präziser dokumentiert findet:

$(11s\ 4p\ 1d/4s\ 1p)[3s\ 2p\ 1d/2s\ 1p]$ bezeichnet bspw. den 6-31G*-Satz (für die Atome der ersten Langperiode werden 11 s -, 4 p - und 1 d -GTO zu 3 s -, 2 p - und 1 d -CGTO kontrahiert, für die Wasserstoff-Atome verwendet man 4 s - und 1 p -GTO, kontrahiert zu 2 s - und einem p -CGTO).

Die Beschreibung des elektronischen Zustandes eines Systems durch besetzte und unbesetzte Molekülorbitale allein ist in vielen Fällen nicht ausreichend. Man weiss inzwischen, dass manche Systeme in der Nähe ihres Ionisierungs-Potenzials eher einem einfach geladenen Kation mit einem noch nicht ganz abstrahierten, also „locker“ gebundenen Elektron gleichen. Einen solchen Zustand nennt man in Anlehnung an das Wasserstoff-Atom Rydberg-Zustand. Dementsprechend ähnelt das von diesem „locker“ gebundenen Elektron besetzte Orbital einem diffusen Atomorbital, das sich über das gesamte Molekül ausdehnt. Häufig kommt es dabei zur Mischung von Rydberg- und Valenz-Zuständen, so dass in vielen Berechnungen Rydberg-Funktionen in die Basissätze mit einbezogen werden müssen.

Ein weiterer Fall eines Systems, dessen Berechnung eine Ergänzung der Standard-Basissätze erforderlich machen kann, ist das Anion. Hier werden weitere, „negative Ionen-Funktionen“ genannte Funktionen in den Basissatz aufgenommen, die diffuser als Valenz- und kompakter als Rydberg-Funktionen sind.

2.3.6 Self-Consistent-Field

Im Roothaan-Hall-Verfahren setzt man die Raum-Orbitale ψ_i als Linearkombinationen eines Satzes von κ Atomorbitalen $\{\phi_\mu\}$ an (LCAO-MO, Abschnitt 2.3.3, S. 14):

$$\psi_i(1) = \sum_{\mu=1}^{\kappa} c_{\mu i} \phi_\mu(1) \quad (2.48)$$

Dies führt zu einer modifizierten Hartree-Fock-Gleichung:

$$\hat{f}(1) \sum_{\mu=1}^{\kappa} c_{\mu i} \phi_\mu(1) = \epsilon_i \sum_{\mu=1}^{\kappa} c_{\mu i} \phi_\mu(1) \quad (2.49)$$

Multipliziert man von links mit $\phi_\nu^*(1)$ und integriert über dr_1 , so gelangt man zu den **Roothaan-Hall-Gleichungen**, die einen wesentlichen Schritt zur Maschinenlösbarkeit des Problems darstellen:

$$\sum_{\mu=1}^{\kappa} \int dr_1 \phi_\nu^*(1) \hat{f}(1) c_{\mu i} \phi_\mu = \epsilon_a \sum_{\mu=1}^{\kappa} c_{\mu i} \int dr_1 \phi_\nu^*(1) \phi_\mu(1) \quad (2.50)$$

Diese wichtige Beziehung wird in Kürze gebraucht werden.

Nach der Definition der $S_{\mu\nu}$ aus Gleichung 2.31 ergeben sich die Elemente Überlappmatrix **S** als:

$$S_{\mu\nu} = \int dr_1 \phi_\mu^*(1) \phi_\nu(1) \quad (2.51)$$

Analog zum Fock-Operator kennt man die Elemente der **Fockmatrix** F , die sich aus einem Einelektronenteil \mathbf{H} und einem Zweielektronenteil \mathbf{G} zusammensetzt:

$$\begin{aligned} F_{\mu\nu} &= \int dr_1 \phi_\mu^*(1) \hat{f} \phi_\nu(1) \\ &= \int dr_1 \phi_\mu^*(1) \hat{h}(r_1) \phi_\nu(1) + \sum_{a=1}^{n/2} \int dr_1 \phi_\mu^*(1) [2J_a(1) - K_a(1)] \phi_\nu(1) \\ &= H_{\mu\nu}^{core} + G_{\mu\nu} \end{aligned} \quad (2.52)$$

Die Iteration. Wie in Abschnitt 2.3.4 (S. 14) dargestellt, entspricht das im Fock-Operator eingesetzte Hartree-Fock-Potenzial v^{HF} Coulomb- und Austausch-Integralen von Spin- oder Raum-Orbitalen. Diese Orbitale sind von einander abhängig, so dass eine Lösung auf direktem Weg unmöglich ist. Daher beginnt man die Rechnung zunächst mit einer mehr oder weniger gut angesetzten Testfunktion $\tilde{\Phi}$, dem so genannten *initial guess*. Dann verbessert man die Testfunktion durch lineare Variation mit Lagrangeschen Multiplikatoren ϵ_{ij} .

$$\sum_i F_{ij} c_j = \sum_j \epsilon_{ij} S_{ij} c_j, \quad (2.53)$$

wobei F_{ij} die Matrixelemente des Fock-Operators darstellen. Betrachtet man die Variation aller HF-Orbitale gleichzeitig, so haben wir es mit einer Matrix von Expansions-Koeffizienten $c_{\mu i}$ zu tun, und wir schreiben die Roothaan-Hall-Gleichungen 2.50 in ihrer Matrix-Form:

$$\mathbf{FC} = \mathbf{SCE}. \quad (2.54)$$

Hier ist \mathbf{c} der Vektor der Koeffizienten der Atomorbitale ϕ_μ in der Entwicklung des Orbitals Ψ_i . \mathbf{E} ist der Vektor der Eigenwerte ϵ .

Diese Prozedur wiederholt man so lange, bis das Hartree-Fock-Potenzial, die Spinorbitale und/oder die Energie-Eigenwerte konvergieren. Man spricht von Selbstkonsistenz (*self consistent field*, SCF) und nennt diesen Zustand stationär.

An dieser Stelle stellt die Auswahl der Atomorbital-Basissätze $\{\phi_i\}$ ein wichtiges Kapitel der Quantenchemie dar, da die Qualität der Basen entscheidenden Einfluss einerseits auf die Präzision der Ergebnisse und andererseits auf den rechen-technischen Aufwand haben (Näheres hierzu wurde bereits im Abschnitt 2.3.5, S. 17 diskutiert).

Das Ergebnis. Als Lösung des Self-Consistent-Field-Verfahrens ergibt sich ein Satz orthonormaler Hartree-Fock Spinorbitale $\{\chi_k\}$ und Lagrangesche Multiplikatoren in Form einer Matrix \mathbf{E} . In dieser Form ist das Ergebnis unanschaulich und nicht eindeutig bestimmt. Diagonalisiert man diese Matrix aber, so entsprechen die Diagonalelemente den Orbitalenergien ϵ der jetzt ebenfalls klar definierten Spinorbitale, die man auch *kanonische HF-Orbitale* nennt.

Die Deutung. Die Spinorbitale lassen sich in der Reihenfolge ihrer Energieeigenwerte ordnen und beispielsweise für die Beschreibung des Hartree-Fock-Grundzustandes $|\Psi_0^{\text{HF}}\rangle$ sukzessiv mit Elektronen besetzen. Man erhält besetzte Orbitale $\chi_{a,b,\dots}$ und unbesetzte, so genannte virtuelle Orbitale $\chi_{r,s,\dots}$, und es gilt für eine Besetzung mit n Elektronen:

$$|\Psi_0^{\text{HF}}\rangle = |\chi_1, \chi_2, \dots, \chi_a, \chi_b, \dots, \chi_n\rangle \quad (2.55)$$

Die Energie E_0 des Grundzustandes ergibt sich als Summe aller Orbitalenergien abzüglich der dabei doppelt gezählten Elektron-Elektron-Wechselwirkungen zu

$$E_0 = \sum_a^n \langle \chi_a | \hat{h} | \chi_a \rangle + \frac{1}{2} \sum_a^n \sum_b^n \langle \chi_a \chi_b | | \chi_a \chi_b \rangle \quad (2.56)$$

Die durch Entfernung oder Addition eines Elektrons berechenbaren Energiedifferenzen sind nach Koopmans Theorem (Szabo und Ostlund [1996]) ein Maß für Ionisationspotenziale und Elektronenaffinitäten.⁷ Im Rahmen von Koopmans Theorem wird somit den Orbitalenergien eine eigentlich nicht vorhandene, physikalische Bedeutung zugeordnet.

Die Varianten. Je nach Besetzungsgrad und Art der verwendeten Orbitale unterscheidet man drei unterschiedliche Hartree-Fock-Verfahren:

- **RHF** = Restricted Hartree Fock
verwendet Raumorbitale, die jeweils doppelt oder nicht besetzt werden dürfen
- **ROHF** = Restricted Open-Shell Hartree Fock
verwendet Raumorbitale, die flexibel besetzt werden dürfen
- **UHF** = Unrestricted Hartree Fock
verwendet Spinorbitale, die flexibel besetzt werden dürfen

Die RHF-Methode ist das meist genutzte Verfahren, denn es liefert bei vergleichsweise geringem Rechenaufwand gute Ergebnisse für Gleichgewichtsgeometrien von Molekülen im elektronischen Grundzustand. Flexibler ist dagegen die UHF-Methode, da sie auch spin-relevante Phänomene qualitativ richtig erklären kann.

⁷Das Koopmans Theorem produziert im Allgemeinen zu hohe Ionenpotenziale und zu niedrige Elektronenaffinitäten, weil es von „eingefrorenen“ Orbitalen ausgeht. Tatsächlich aber relaxieren Orbitale nach Addition oder Subtraktion eines Elektrons. Des Weiteren werden Elektronenkorrelations-Effekte vernachlässigt; das allerdings ist eine prinzipielle Schwäche der Hartree-Fock-Näherung.

Die Grenzen. Das SCF-Verfahren erleidet die Restriktionen der Näherungen, die es benutzt. Eine Slater-Determinante repräsentiert eine Wellenfunktion um so präziser, je größer ihre Basis ist. Des Weiteren berücksichtigt die Hartree-Fock-Näherung keinerlei Elektronenkorrelations-Effekte. Dennoch stellen die hier gefundenen Ergebnisse eine sehr gute Grundlage für weiterführende Verfahren wie Configuration Interaction dar, von dem in Kapitel 2.3.11 (S. 31) zu sprechen sein wird.

Angeregte Konfigurationen

Da das SCF-Verfahren nicht nur einen Satz $\{\chi_a\}$ der im Grundzustand besetzten, sondern auch einen Satz $\{\chi_r\}$ virtueller Spinorbitale ermittelt, ist es möglich, Wellenfunktionen auch für angeregte Zustände zu formulieren. Nach den Regeln der Kombinatorik gibt es $\binom{k}{n}$ Möglichkeiten dafür, n (ununterscheidbare) Elektronen auf k Spinorbitale zu verteilen. Die Darstellung der Wellenfunktion eines angeregten Zustandes im Rahmen des SCF-Formalismus geschieht durch Austausch von besetzten Spinorbitalen gegen virtuelle in der Slater-Determinante des Grundzustandes, der somit als Referenz fungiert:

$$\begin{array}{lll}
 |\Psi_0\rangle & = & |\chi_1, \chi_2, \dots, \chi_a, \chi_b, \dots, \chi_n\rangle \quad \text{Grundzustand} \\
 |\Psi_a^r\rangle & = & |\chi_1, \chi_2, \dots, \chi_r, \chi_b, \dots, \chi_n\rangle \quad \text{einfache Anregung aus Orbital } \chi_a \text{ in } \chi_r \\
 |\Psi_{ab}^{rs}\rangle & = & |\chi_1, \chi_2, \dots, \chi_r, \chi_s, \dots, \chi_n\rangle \quad \text{doppelte Anregung aus den Orbitalen } \chi_{a,b} \text{ in } \chi_{r,s} \\
 \vdots & & \vdots \\
 \dots & \dots & n\text{-fache Anregung}
 \end{array}$$

Da die Zuführung von Anregungsenergie immer zu einer statistischen Verteilung/Mischung angeregter Konfigurationen führt, ist keine dieser Determinanten eine akkurate Darstellung eines „wirklichen“ angeregten Zustandes. Allerdings sind diese Determinanten nützliche Bausteine für einen weiterführenden Formalismus, die Configuration Interaction (siehe Abschnitt 2.3.11, S. 31).

2.3.7 Dichtefunktionaltheorie

Die Suche nach einer alternativen Behandlung der nicht relativistischen zeitunabhängigen Schrödinger-Gleichung, die im Gegensatz zur Hartree-Fock-Näherung Elektronen-Korrelations-Effekte systematisch mit berücksichtigt, führte in den sechziger Jahren zur Dichtefunktionaltheorie (Parr und Yang [1994]). Ihr formalistischer Ansatz ist ein nicht-quantenmechanischer, zumal die zentrale Größe nicht mehr eine Wellenfunktion, sondern die durch sie erzeugte Elektronendichte $\rho(\mathbf{r})$ ist. Sie ist wie folgt definiert:

$$\rho(\mathbf{r}) = n \int |\Phi(\{\mathbf{x}_i\})|^2 d\omega_1 d\mathbf{x}_2, \dots, d\mathbf{x}_n \quad (2.57)$$

Darin entspricht n der Gesamt-Elektronen-Anzahl, Φ ist die Gesamt-Elektronen-Wellenfunktion, ω ist eine nicht näher definierte Spinkoordinate, \mathbf{r} die Raumkoordinate und \mathbf{x} die Gesamtheit beider. Randbedingung ist trivialerweise, dass alle Elektronen des Systems sich irgendwo im Raum befinden müssen:

$$\int \rho(\mathbf{r}) d\mathbf{r} = n \quad (2.58)$$

Von den Konsequenzen dieser Substitution soll in den folgenden Abschnitten die Rede sein.

2.3.8 Die Theoreme von Hohenberg und Kohn

Hohenberg und Kohn formulierten in den sechziger Jahren zwei nach ihnen benannte Theoreme (Hohenberg und Kohn [1964]):

1. Durch jede Elektronendichte wird eindeutig ein externes Potenzial festgelegt:

$$\rho(\mathbf{r}) \longrightarrow v(\mathbf{r}) \quad (2.59)$$

Die lässt sich leicht zeigen: Gehen wir von der Annahme aus, dass die Elektronendichte $\rho(\mathbf{r})$ zwei externe Potentiale festlege, und nennen wir diese Potentiale $v_1(\mathbf{r})$ und $v_2(\mathbf{r})$. Folglich gäbe es zu jedem Potenzial unterschiedliche Hamilton-Operatoren \hat{H}_1 und \hat{H}_2 , sowie unterschiedliche Wellenfunktionen Ψ_1 und Ψ_2 , die jedoch nach Voraussetzung dieselbe Grundzustandsdichte aufweisen:

$$|\Psi_1|^2 = |\Psi_2|^2 \quad (2.60)$$

Nach dem Variationsprinzip (siehe Abschnitt 2.3.1, S. 10) gilt für den Erwartungswert der Energie von Ψ_1 :

$$\begin{aligned} E_1^0 &= \langle \Psi_1 | \hat{H}_1 | \Psi_1 \rangle < \langle \Psi_2 | \hat{H}_1 | \Psi_2 \rangle = \langle \Psi_2 | \hat{H}_2 | \Psi_2 \rangle + \langle \Psi_2 | \hat{H}_1 - \hat{H}_2 | \Psi_2 \rangle \\ &= E_2^0 + \langle \Psi_2 | v_1(\mathbf{r}) - v_2(\mathbf{r}) | \Psi_2 \rangle \end{aligned} \quad (2.61)$$

Ebenso formuliert man für die Wellenfunktion Ψ_2

$$\begin{aligned} E_2^0 &= \langle \Psi_2 | \hat{H}_2 | \Psi_2 \rangle < \langle \Psi_1 | \hat{H}_2 | \Psi_1 \rangle = \langle \Psi_1 | \hat{H}_1 | \Psi_1 \rangle + \langle \Psi_1 | \hat{H}_2 - \hat{H}_1 | \Psi_1 \rangle \\ &= E_1^0 + \langle \Psi_1 | v_2(\mathbf{r}) - v_1(\mathbf{r}) | \Psi_1 \rangle \end{aligned} \quad (2.62)$$

Es entsteht ein Gleichungssystem, das zum Widerspruch führt:

$$\left| \begin{array}{l} E_1^0 < E_2^0 + \int \rho(\mathbf{r}) [v_1(\mathbf{r}) - v_2(\mathbf{r})] d\mathbf{r} \\ E_2^0 < E_1^0 + \int \rho(\mathbf{r}) [v_2(\mathbf{r}) - v_1(\mathbf{r})] d\mathbf{r} \end{array} \right| \iff E_1^0 + E_2^0 < E_2^0 + E_1^0 \quad (2.63)$$

Die Zuordnung eines externen Potentials ist also tatsächlich eindeutig, so dass die folgende Formulierung der Energie als Funktional der Dichte gerechtfertigt ist:

$$\begin{aligned} E[\rho(\mathbf{r})] &= V_{\text{ke}}[\rho(\mathbf{r})] + T[\rho(\mathbf{r})] + V_{\text{ee}}[\rho(\mathbf{r})] \\ &= \int \rho(\mathbf{r})v(\mathbf{r})d\mathbf{r} + T[\rho(\mathbf{r})] + V_{\text{ee}}[\rho(\mathbf{r})] \end{aligned} \quad (2.64)$$

In der potenziellen Elektron-Elektron-Wechselwirkungsenergie enthalten ist die klassische Coulomb-Repulsion J , nunmehr ebenfalls ein Funktional der Dichte:

$$J[\rho(\mathbf{r})] = \frac{1}{2} \iint -\frac{1}{r_{12}} \rho(\mathbf{r}_1)\rho(\mathbf{r}_2)d\mathbf{r}_1d\mathbf{r}_2 \quad (2.65)$$

2. Das zweite HK-Theorem lautet: „Jedem Potenzial $v(\mathbf{r})$ kann man eine räumliche Wellenfunktion $\Psi(\mathbf{r})$ zuordnen“:

$$\rho(\mathbf{r}) \longrightarrow v(\mathbf{r}) \longrightarrow \Psi(\mathbf{r}) \quad (2.66)$$

Dieser scheinbare Rückschritt in das gewohnte Umfeld der Quantenchemie ermöglicht eine variationelle Optimierung der Elektronendichte $\rho(\mathbf{r})$ über den Umweg der Wellenfunktion $\Psi(\mathbf{r})$ unter der Maßgabe der Energie-Minimierung. Nach dem Variationsprinzip gilt für jede „Test-Dichte“ $\tilde{\rho}(\mathbf{r})$ und jede Testfunktion $\tilde{\Psi}(\mathbf{r})$:

$$E[\tilde{\rho}(\mathbf{r})] = E[\tilde{\Psi}(\mathbf{r})] \geq E[\Psi(\mathbf{r})] = E[\rho(\mathbf{r})] \quad (2.67)$$

Da also

$$E[\tilde{\rho}(\mathbf{r})] = \langle \tilde{\Psi}(\mathbf{r}) | \hat{H} | \tilde{\Psi}(\mathbf{r}) \rangle = \int \tilde{\rho}(\mathbf{r})v(\mathbf{r})d\mathbf{r} + T[\tilde{\rho}(\mathbf{r})] + V[\tilde{\rho}(\mathbf{r})] \quad , \quad (2.68)$$

erhält man unter Verwendung der Nebenbedingung 2.58 die optimierte Elektronendichte durch Lösung der folgenden Gleichung mittels Lagrangescher Multiplikatoren λ :

$$\begin{aligned} 0 &= \delta E[\rho(\mathbf{r})] - \lambda \delta \left(n - \int \rho(\mathbf{r}) d\mathbf{r} \right) \\ \Rightarrow \lambda &= v(\mathbf{r}) + \frac{\delta T[\rho(\mathbf{r})]}{\delta \rho(\mathbf{r})} + \frac{\delta V_{ee}[\rho(\mathbf{r})]}{\delta \rho(\mathbf{r})} \end{aligned} \quad (2.69)$$

Wäre die funktionale Abhängigkeit der kinetischen Energie $T[\rho]$ und der potenziellen Energie $V_{ee}[\rho]$ bekannt, dann könnte man Gleichung 2.69 exakt lösen. Statt dessen bedarf es auch im Falle der Dichtefunktionale einer geeigneten Näherung für die kinetische Energie T und die Wechselwirkungsbeiträge der Elektronen V_{ee} .

Die Kohn-Sham-Gleichungen

Bei der Herleitung des Kohn-Sham-Formalismus betrachtet man zunächst ein fiktives, aber geschlossen berechenbares System ohne Elektron-Elektron-Wechselwirkungen, vergleicht die Ergebnisse mit dem nicht berechenbaren realen System und versucht anschließend, den Fehler zu quantifizieren.

Im gedankenexperimentellen System \mathbf{S} ohne Elektron-Elektron-Wechselwirkungen definiert man die so genannte **Kohn-Sham-Determinante** für n Elektronen in n Orbitalen ϕ_i gemäß

$$T_S[\rho(\mathbf{r})] = \sum_i^n \langle \phi_i | -\frac{1}{2} \nabla^2 | \phi_i \rangle \quad \text{mit} \quad \rho(\mathbf{r}) = \sum_i^n |\phi_i(\mathbf{r})|^2 \quad (2.70)$$

Die verwendeten Orbitale $\phi_i(\mathbf{r})$ genügen der Eigenwertgleichung

$$\left[-\frac{1}{2} \nabla^2 + v_S(\mathbf{r}) \right] \phi_i(\mathbf{r}) = \epsilon_i \phi_i(\mathbf{r}) \quad (2.71)$$

Das Potenzial v_S ist ein formalistisches Auxiliar, das gerade so konstruiert wird, dass das nicht-wechselwirkende System dieselbe Elektronendichte erzeugt wie das real wechselwirkende:

$$\rho_{\text{nicht-ww}} = \rho_{\text{ww}} \quad (2.72)$$

Da die Elektronendichten des nicht- und des wechselwirkenden Systems gleich sind, und die Energie ein Funktional dieser Dichte ist, kann man die Systemenergie in Termen des nicht-wechselwirkenden Systems angeben:

$$E[\rho(\mathbf{r})] = T_S[\rho(\mathbf{r})] + \int v_S(\mathbf{r}) \rho(\mathbf{r}) d\mathbf{r} \quad (2.73)$$

Der Vergleich mit Gleichung 2.64 des wechselwirkenden Systems führt zu folgendem Ausdruck:

$$\begin{aligned}
 E[\rho(\mathbf{r})] &= \int v_S(\mathbf{r})\rho(\mathbf{r})d\mathbf{r} + T_S[\rho(\mathbf{r})] + J[\rho(\mathbf{r})] \\
 &+ \underbrace{\left(T[\rho(\mathbf{r})] - T_S[\rho(\mathbf{r})]\right) + \left(V_{ee}[\rho(\mathbf{r})] - J[\rho(\mathbf{r})]\right)}_{= E_{XC}[\rho(\mathbf{r})]} \quad (2.74)
 \end{aligned}$$

Im Austausch-Korrelations-Funktional $E_{XC}[\rho(\mathbf{r})]$ enthalten sind demgemäß ein Korrekturfaktor für die kinetischen Energien, sowie alle Elektron-Elektron-Wechselwirkungsbeiträge, die durch das Coulomb-Funktional $J[\rho(\mathbf{r})]$ entweder gar nicht oder falsch erfasst werden. Dazu zählt insbesondere die so genannte *dynamische Korrelation*, die man vereinfachend als eine zeitabhängige Verringerung des Bewegungsspielraumes eines einzelnen Elektrons durch die Repulsion der zum Zeitpunkt t zufällig benachbarten Elektronen verstehen kann. Das entsprechende Austausch-Korrelations-Potenzial $v_{XC}(\mathbf{r})$ erhält man als Funktionalableitung

$$v_{XC}(\mathbf{r}) = \frac{\delta E_{XC}[\rho(\mathbf{r})]}{\delta \rho(\mathbf{r})} \quad , \quad (2.75)$$

und man formuliert die Kohn-Sham-Gleichungen in Form einer Eigenwert-Gleichung:

$$\left[-\frac{1}{2}\nabla^2 + v(\mathbf{r}) + J[\rho(\mathbf{r})] + v_{XC}(\mathbf{r})\right] \phi_i = \epsilon_i \phi_i \quad (2.76)$$

Als Kohn-Sham-Potenzial v_{KS} definiert man zusammenfassend

$$v_{KS}(\mathbf{r}) = v(\mathbf{r}) + J[\rho(\mathbf{r})] + v_{XC}(\mathbf{r}) \quad , \quad (2.77)$$

die Orbitale ϕ_i werden nach ihrer „Herkunft“ und in Abgrenzung von den in Hartree-Fock-Näherung ermittelten Wellenfunktionen **Kohn-Sham-Orbitale** genannt.

Austausch-Korrelations-Funktionale

Die Qualität der Ergebnisse einer Kohn-Sham-basierten Rechnung ist besonders eng verknüpft mit der Qualität des Austausch-Korrelations-Potenzials $v_{XC}(\mathbf{r})$. Um den enormen Geschwindigkeitsgewinn der DFT nicht mit einem ebenso großen Verlust an Präzision bezahlen zu müssen, wurde in den vergangenen 15 Jahren sehr viel quantenchemisches *Know How* in die Entwicklung geeigneter Funktionale gesteckt (Johnson u. a. [1993]; Handy [1988]; Cohen und Handy [2000]). Dabei hat sich eine grundsätzliche Trennung von Austausch- und Korrelations-Funktionalen durchgesetzt.

Eine häufig genutzte Vorgehensweise in beiden Fällen ist ein physikalisch fundierter *initial guess* und dessen Justierung durch Vergleich mit einem geeigneten Bezugssystem. Als solches wurde in einem frühen nahe liegenden Ansatz das Modell des homogenen (spin-polarisierten) Elektronengases — auch „Jellium“ genannt — als Berechnungsgrundlage heran gezogen. Kennzeichen dieses fiktiven Systems sind eine gleichförmige Verteilung der Elektronen und eine gleichförmig verteilte positive Ladung. Es ist das einfachst denkbare Vielkörpersystem, das dennoch über seine Dichte vollständig definiert ist. Später hat man versucht, dieses Modell zu verfeinern und den tatsächlichen Verhältnissen in Molekülen anzugleichen und gelangte so z.B. zum so genannten „fast homogenen“ Elektronengas (Gill [1998]). Den letzten „Schliff“ bekommen manche Austausch-Korrelations-Funktionale erst durch die Anpassung an einen Satz repräsentativer experimenteller Ergebnisse, weshalb man sie als „semiempirisch“ betrachten kann.

Austausch-Funktionale. Die Austausch-Energie im homogenen Elektronengas wurde von Dirac [1930] angegeben mit

$$E_X^{hom.} = -\frac{3}{2} \left(\frac{3}{4\pi} \right)^{1/3} \sum_{\sigma} \int [\rho_{\sigma}(\mathbf{r})]^{4/3} d\mathbf{r} \quad . \quad (2.78)$$

Die Erweiterung der Betrachtungen auf den inhomogenen Fall führt zur so genannten **Lokalen Spin-Dichte Näherung** (LSDA), die in Austausch- und Korrelations-Anteil gegeben ist durch das Integral

$$E_{XC}^{inhom.} = \int \epsilon_{XC}[\rho_{\alpha}(\mathbf{r}), \rho_{\beta}(\mathbf{r})] d\mathbf{r} \quad (2.79)$$

Man beachte die Separation des Integranden in lokale Spin-Dichten für α - und β -Spin. Bezieht man den Gradienten der Dichte $\nabla\rho(\mathbf{r})$ als Quantifizierung der Inhomogenität mit in Funktional mit ein (von Weizsäcker [1935]), so gelangt man zu Potenzialen des GGA-Typs (GGA - *generalized gradient approximation*)

$$E_{XC}^{GGA} = \int f[\rho_{\alpha}(\mathbf{r}), \rho_{\beta}(\mathbf{r}), \nabla\rho_{\alpha}(\mathbf{r}), \nabla\rho_{\beta}(\mathbf{r})] d\mathbf{r} \quad . \quad (2.80)$$

Formalistische Probleme mit diesem Funktional-Typus überwand Becke [1988] mit der Einführung des korrekten asymptotischen Verhaltens im Austausch-Funktional „B88“:

$$E_X = -\frac{1}{2} \int \frac{\rho_{\sigma}(\mathbf{r})}{r} d\mathbf{r} \quad \text{mit} \quad \lim_{r \rightarrow \infty} E_X = 0 \quad (2.81)$$

Obwohl damit die Qualität der Austausch-Funktionale gesteigert werden konnte, beobachtete Becke, dass die Verwendung von Mischformen aus GGA-Funktionalen und der „exakten“ Hartree-Fock-Austausch-Energie noch deutlich bessere Ergebnisse

hervorbringt. Die Entwicklung führte über ein Hybrid mit gleichen Anteilen GGA und HF im Funktional „Becke (half & half)“ zu einem Funktional mit empirisch ermitteltem Mischungsverhältnis namens „B3“ (Becke [1993b]):

$$E_X = c_0 E_X^{\text{HF}} + (1 - c_0^2) E_X^{\text{GGA}} \quad (2.82)$$

Der Mischungs-Parameter c_0 hat einen Wert von ca. 0.2.

Korrelations-Funktionale. Die Form der Korrelations-Funktionale ist sehr komplex und soll hier nicht im Einzelnen aufgeführt werden. Nichtsdestotrotz sind sie physikalisch fundiert und wurden an Modellsystemen parameterisiert. Besonders zu nennen sind das im Zusammenhang mit der LSDA-Näherung entstandene VWN-Funktional (Vosko u. a. [1980]), die Funktionale von Perdew und Wang [1992] sowie das heute meist verwendete LYP-Funktional (benannt nach Lee u. a. [1988]).

Austausch- und Korrelations-Funktionale lassen sich im Prinzip rekombinieren — beispielsweise zu S-VWN oder B3-LYP. Da ein exakter Ausdruck für v_{XC} bisher unbekannt ist, sucht man auch weiterhin nach *ab initio* Funktionalen, die der Dichtefunktionaltheorie zu weiterem Erfolg verhelfen könnten.

2.3.9 Vergleich zwischen Hartree-Fock und Kohn-Sham

Bei einer Entwicklung der Orbitale ϕ_i in eine Basis von Spinorbitalen $\{\chi_\alpha\}$ gemäß

$$\phi_i = \sum_{\alpha} c_{\alpha i} \chi_{\alpha} \quad (2.83)$$

erhält man die Hartree-Fock-Gleichungen in der Form

$$\sum_{\beta} \left\langle \chi_{\alpha} \left| -\frac{1}{2} \nabla^2 + v(\mathbf{r}) + \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' - \int \sum_j \frac{\phi_j(\mathbf{r}) \phi_j(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' P_{\mathbf{r}\mathbf{r}'} - \epsilon_i \right| \chi_{\beta} \right\rangle c_{\beta i} = 0 \quad (2.84)$$

Eine analoge Formulierung der Kohn-Sham-Gleichungen lautet:

$$\sum_{\beta} \left\langle \chi_{\alpha} \left| -\frac{1}{2} \nabla^2 + v(\mathbf{r}) + \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' + v_{\text{XC}}(\mathbf{r}) - \epsilon_i \right| \chi_{\beta} \right\rangle c_{\beta i} = 0 \quad (2.85)$$

Darin ist $P_{\mathbf{r}\mathbf{r}'}$ die sog. Dichte-Matrix. Man erkennt, dass der entscheidende Unterschied beider Näherungen in der Behandlung der Austausch- bzw. Korrelations-Beiträge liegt:

$$\begin{array}{ccc} \text{HF} & & \text{DFT} \\ - \int \sum_j \frac{\phi_j(\mathbf{r}) \phi_j(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' P_{\mathbf{r}\mathbf{r}'} & \longleftrightarrow & v_{\text{XC}}(\mathbf{r}) \end{array} \quad (2.86)$$

Drei Dinge fallen auf:

1. Die Implementierung der Dichtefunktional-Näherung in bestehende Hartree-Fock-Algorithmen ist durch Austausch der Ausdrücke in Gleichung 2.86 prinzipiell immer möglich.
2. Je nachdem, welches Austausch-Korrelations-Funktional verwendet wird, kann man mit einem erheblichen Geschwindigkeitsgewinn rechnen, weil aufwändige Matrixoperationen entfallen.
3. Über das Austausch-Korrelations-Funktional kann auch die im Umfeld der Hartree-Fock-Näherung vernachlässigte dynamische Korrelations-Energie näherungsweise in ein Eindeterminanten-Verfahren einbezogen werden.

2.3.10 Resolution of the Identity

Ein eher heuristisches Anliegen der Theoretischen Chemie ist die Entwicklung geeigneter Methoden zur Einsparung rechentechnischer Ressourcen. Im Umfeld elektronischer *ab initio* Kalkulationen erzeugen insbesondere die speicherintensiven Vier-Index-Zwei-Elektronen-Integrale Probleme; allgemein in Mulliken-Notation:

$$(ij|kl) = \int \phi_i^*(1)\phi_j(1)\frac{1}{r_{12}}\phi_k^*(2)\phi_l(2)d\tau \quad (2.87)$$

Folglich macht es durchaus Sinn, nach alternativen Darstellungen dieser Integrale zu suchen, die den Speicherbedarf verringern, ohne die Präzision der Ergebnisse wesentlich zu beeinträchtigen.

Gesucht ist eine Methode, die den formalistischen Aufwand auf die Komplexität von höchstens Drei-Index-Integralen reduziert. Mathematisch gelingt diese Reduktion durch Einschub eines Einheitsoperators \hat{I} , gegeben als

$$\hat{I} = \sum_m |m\rangle\langle m| \quad (2.88)$$

mit der Hilfsbasis $|m\rangle$. Dieses Vorgehen nennt man *resolution of the identity*. In Abhängigkeit davon, ob die verwendete Hilfsbasis orthonormal ist oder nicht, unterscheidet man zwei Fälle:

1. $|m\rangle$ ist orthonormal

Durch Einfügen der Hilfsbasis erreicht man die gewünschte Reduktion auf Drei-Index-Integrale:

$$(ij|kl) = \sum_m (ijm)(m|kl) \quad (2.89)$$

Darin steht (ijm) für ein Drei-Index-Ein-Elektron-Integral der Form

$$(ikm) = \int \phi_i^*(1)\phi_j(1)\phi_m(1)d\tau, \quad (2.90)$$

und $(m|kl)$ repräsentiert ein Drei-Index-Zwei-Elektronen-Repulsions-Integral

$$(m|kl) = \int \phi_m^*(1) \frac{1}{r_{12}} \phi_k^*(2) \phi_l(2) d\tau \quad (2.91)$$

Der Einschub führt nur dann wirklich zur Identität, wenn die Basis $|m\rangle$ vollständig ist.

2. $|m\rangle$ ist nicht orthonormal

In diesem Fall lautet der korrekte Identitätsoperator

$$\hat{I} = \sum_m \sum_n \frac{|m\rangle\langle n|}{S_{mn}}, \quad (2.92)$$

und der Einschub führt zu

$$(ij|kl) = \sum_{mn} \frac{(ijm)(n|kl)}{S_{mn}}. \quad (2.93)$$

Da reale Hilfsbasen im Allgemeinen weder orthonormal noch vollständig sind, ist die *resolution of the identity* immer eine Näherung und als solche mit einem gewissen Fehler behaftet.

Dieser Fehler lässt sich allerdings deutlich minimieren, indem man eine zweite Hilfsbasis einfügt und das entstandene Produkt variationell optimiert:

$$\phi_i(\mathbf{r})\phi_j(\mathbf{r}) = |ij\rangle \approx \sum_m \mathbf{C}_m^{ij} |m\rangle = |\Theta_{ij}\rangle \quad (2.94)$$

Darin ist \mathbf{C}_m^{ij} eine Koeffizienten-Matrix. Der dabei gemachte Fehler lässt sich formulieren als eine Restfunktion R_{ij} mit

$$R_{ij}(\mathbf{r}) = \phi_i(\mathbf{r})\phi_j(\mathbf{r}) - \Theta_{ij} \quad (2.95)$$

Die Güte der RI-Näherung hängt also nur noch davon ab, ob es gelingt, den Fehler $R_{ij}(\mathbf{r})$ zu minimieren. Dazu gibt es verschiedene Möglichkeiten. Die nächstliegende ist die der Minimierung der Funktionsnorm $\|(R_{ij}R_{ij})\|$. Sie führt zu der Substitution

$$(ij|kl) = \sum_{tuvw} (ijt) S_{tu}^{-1} V_{uv} S_{vw}^{-1} (wkl) \quad \text{mit} \quad V_{uv} = (u|v) = \langle u|r_{12}^{-1}|v\rangle \quad (2.96)$$

V_{uv} sind Zwei-Index-Zwei-Elektronen-Repulsionsintegrale. Vahtras u. a. [1993] konnten zeigen, dass weder die „S“-Näherung aus Gleichung 2.93, noch die in Gleichung 2.96 formulierte „SVS“-Näherung realistische Geometrieparameter liefern. Alternativ schlagen sie das formalistische Konstrukt einer Restfunktions-Selbstrepulsion

$(R_{ij}|R_{ij})$ als Objekt der Minimierung vor. Man erhält z.B. für ein Vier-Index-Zwei-Elektronen-Integral

$$(ij|kl) = \sum_{tu} (ij|t) V_{tu}^{-1}(kl|u) \quad . \quad (2.97)$$

Die Verwendung der RI-Näherung in seiner „V“-Variante führt zu guten Ergebnissen und wurde daher in der TURBOMOLE-Suite implementiert. Sinn macht sie immer dann, wenn ihr Speicherbedarf, der mit $n^2 \times m$ skaliert, kleiner ist als der Speicherbedarf ohne RI-Näherung, der bei n^4 liegt⁸. Auch die Wahl des Hilfsbasissatzes ist daher eine Gratwanderung zwischen Präzision bei großen und Geschwindigkeit bei kleinen Sätzen $\{|m\}$.

2.3.11 Elektronenkorrelation

Für eine präzisere Erfassung der Elektron-Elektron-Wechselwirkungen ist die Vorstellung eines Elektrons im effektiven statischen Feld der anderen Elektronen nicht ausreichend. Die Hartree-Fock-Näherung ist daher mit einem systematischen Fehler behaftet, der so genannten Korrelationsenergie (Wigner [1934]). Diese ist definiert als die Differenz zwischen der im Rahmen nicht relativistischer Berechnungen exakten Energie und dem Hartree-Fock-Limit:

$$E_{\text{corr}} = E_{\text{exakt}}^{\text{nichtrel.}} - E_0^{\text{HF}} \quad (2.98)$$

Da die Korrelationsenergie häufig in der Größenordnung chemischer Bindungen liegt⁹, stellt sich die Frage nach einer geeigneten Methode, die E_{corr} mit erfassen kann.

Die Tatsache, dass der tatsächliche elektronische Zustand eines Systems selten durch nur eine einzelne elektronische Besetzung vollständig beschrieben wird, führt zum Prinzip der *Configuration Interaction*. Im CI-Verfahren setzt man als Wellenfunktion Φ nicht mehr nur eine einzelne sondern vielmehr eine Linearkombination mehrerer Slater-Determinanten an und bestimmt die Koeffizienten durch lineare Variation (Ritzsches Variationsprinzip). In der Regel benutzt man hierfür die Slater-Determinanten angeregter Zustände in der Spinorbital-Basis $\{\chi_i\}$, die man mit einer SCF-Kalkulation ermittelt hat (siehe hierzu Abschnitt 7, S. 22).

$$|\Phi\rangle = c_0 |\Psi_0\rangle + \sum_{ra} c_a^r |\Psi_a^r\rangle + \sum_{\substack{a<b \\ r<s}} c_{rs}^{ab} |\Psi_{rs}^{ab}\rangle + \sum_{\substack{a<b<c \\ r<s<t}} c_{rst}^{abc} |\Psi_{rst}^{abc}\rangle + \dots \quad (2.99)$$

Die Wellenfunktionen $|\Psi_a^r\rangle$ und $|\Psi_{rs}^{ab}\rangle$ nennt man *Configuration State Function*, oder kurz *CSFs*.¹⁰ (Mit $a > b$ ist Summation über alle a und alle $b > a$ gemeint. Andere

⁸ mit n = Anzahl der Elektronen, und m = Größe der Hilfsbasis

⁹ Sie liegt oft bei ca. 1% der Gesamtenergie.

¹⁰ Eine CSF ist etwas detaillierter gesagt eine Linearkombination von Determinanten mit gleicher Besetzung von räumlichen Orbitalen, aber unterschiedlicher Spin-Einstellung. Die Linearkombina-

Summationen folgen analog.)

An dieser Stelle bietet sich an, den Begriff der „statischen Korrelation“ zu erläutern, der im Zusammenhang mit CI-Verfahren häufig bemüht wird: In (unscharfer) Abgrenzung von der dynamischen Korrelationsenergie (S. 26) bezieht man sich hier auf Effekte, die vor allem durch die Existenz energetisch nahe benachbarter Zustände herrührt. Demgemäß hat man es z.B. bei Dissoziationen von kovalenten Bindungen mit statischen Korrelationseffekten zu tun, die mit der Linearkombination der entsprechenden Wellenfunktionen der beteiligten Zustände in der Gesamtwellenfunktion oft gut erfasst werden kann. Geeignete Verfahren zur Behandlung von solchen Problemen sind daher gerade die CI-Verfahren, CAS-Modelle etc. Statische und dynamische Korrelationsenergie-Beiträge sind stabilisierend.

Die im Falle der (DFT)-MRCI zu lösende Eigenwertgleichung hat die folgende Form:

$$(\mathbf{H} - \mathbf{E}\mathbf{I})\mathbf{C} = 0 \quad (2.100)$$

Darin entspricht \mathbf{H} der Hamilton-Matrix mit den Elementen $\langle \text{CSF}_i | \hat{H} | \text{CSF}_j \rangle$, \mathbf{E} dem Vektor der Eigenwerte, \mathbf{I} der Einheitsmatrix und \mathbf{C} der Koeffizienten-Matrix. Die Lösung erhält man, indem man durch Variation der Koeffizienten die Energie minimiert. Der niedrigste Eigenwert repräsentiert den Grundzustand inklusive der Korrelationsenergie, die sich nun explizit angeben lässt (S. 237 im Buch von Szabo und Ostlund [1996]):

$$E_{\text{corr}} = \sum_{\substack{a < b \\ r < s}} c_{ab}^{rs} \langle \Psi_0 | \hat{H} | \Psi_{ab}^{rs} \rangle \quad (2.101)$$

Die höheren Energie-Eigenwerte sind den angeregten Zuständen zuzuordnen. Man findet, dass insbesondere angeregte Zustände nicht mehr dominant nur durch eine einzelne, formal „angeregte“ SCF-Slater-Determinante, sondern vielmehr als „Hybrid“ mehrerer angeregter Slater-Determinanten beschrieben werden.

Wählt man für die Berechnung sämtliche $\binom{k}{n}$ möglichen Slaterdeterminanten aus, so ist die Qualität der Ergebnisse besonders hoch; der tiefste Energiewert gilt (bei unendlicher Basis) als die exakte nicht relativistische Systemenergie. Unglücklicherweise liegen die rechentechnischen Anforderungen dieses Vorgehens, das Full CI genannt wird, auch bei Verwendung endlicher Basissätze fast immer jenseits verfügbarer Ressourcen, und es erscheint notwendig, die Anzahl der in die Rechnung einbezogenen Determinanten zu reduzieren. Das gelingt mit einer Reihe von „Auswahlregeln“:

1. Nur Determinanten gleichen Gesamt-Spins können miteinander wechselwirken (siehe Fußnote oben, liegt im Verwenden von CSFs begründet)

tion ist dabei so gewählt, dass eine CSF eine Eigenfunktion zum Spinoperator \hat{S}^2 ist, demnach also eine definierte Multiplizität hat. Dies trifft für offenschalige Determinanten i.A. nicht zu.

2. Nach Brillouins Theorem (Szabo und Ostlund [1996]) gibt es keinerlei Kopplungen zwischen dem Grundzustand und einfach angeregten Determinanten.
3. Hamilton-Matrix-Elemente von Determinanten-Paaren, die sich um mehr als zwei Anregungsgrade unterscheiden, sind Null. Dies ist eine Konsequenz der Orthogonalität ihrer Orbitale.

Auch nach der Eliminierung irrelevanter Determinanten bzw. Matrixelemente bleibt der Rechenaufwand für ein CI-Verfahren ohne weitere Maßnahmen sehr groß. Es treten hier zwei wesentliche Probleme auf: Zum Einen ist ein SDCI nicht größenkonsistent (*size consistent*), was zu Problemen führt, wenn man Systeme mit relativ vielen Elektronen behandelt (und was im Übrigen auch der Grund dafür ist, dass kein Programm existiert, das periodische Systeme mit CI-Verfahren berechnet). Zum Anderen ist eine CI-Entwicklung zur Beschreibung von *dynamischer* Korrelation nur sehr langsam konvergent, d.h. man benötigt Millionen von CSFs, was rechnerisch Schwierigkeiten bereitet. Wesentliche Stärken von CI-Methoden liegen in ihrem variationellen Charakter (vgl. S. 10), und darin, dass man einen Zugang zu einem Großteil statischer Korrelationsenergie erhält. Eine Kombination aus CI und DFT-Verfahren sollte sowohl dynamische als auch statische Korrelationsbeiträge erfassen, was uns zum nächsten Kapitel führt.

2.3.12 Dichtefunktionale in CI-Methoden

Die Implementierung der Dichtefunktionaltheorie in Mehrdeterminanten-Verfahren verspricht die Berücksichtigung statischer (CI) und dynamischer (DFT) Elektronen-Korrelation zugleich: Man kombiniert in einem Variationsverfahren die effiziente Auswertung der dynamischen Korrelation über Dichtefunktionale mit dem Einbeziehen höherer Anregungen in einer (MR-)CI-Entwicklung — Letztere erfasst die wesentlichen Anteile statischer Korrelationsenergie (Grimme [1996]; Grimme und Waletzke [1999b]). Die CI-Entwicklung kann so wesentlich kürzer gehalten werden, was massive Zeitersparnis nach sich zieht (Waletzke [2001]).

DFT/SCI

Beschränken wir uns zunächst auf ein Hartree-Fock-basiertes „Singles CI“, also eine CI-Variante, die sämtliche Einfach-Anregungs-Determinanten $|\Psi_a^r\rangle$ aus einem Hartree-Fock-SCF-Verfahren in die Berechnungen mit einbezieht, gemäß

$$|\Phi_{\text{SCI}}\rangle = c_0 |\Psi_0\rangle + \sum_{ar} c_a^r |\Psi_a^r\rangle . \quad (2.102)$$

Die resultierende SCI-Hamilton-Matrix enthält für Singulett- und Triplett-Zustände jeweils drei Klassen unterschiedlicher Matrix-Elemente. Diese sind in Mulliken-No-

tation:

Singulett (Hartree-Fock):

$$\begin{aligned}\langle \Psi_a^r | \hat{H} - E_0 | \Psi_a^r \rangle &= (\varphi_r | \hat{f} | \varphi_r) - (\varphi_a | \hat{f} | \varphi_a) - (\varphi_a \varphi_a | \varphi_r \varphi_r) + 2(\varphi_a \varphi_r | \varphi_a \varphi_r) \\ &= \epsilon_r - \epsilon_a - J_{ar} + 2K_{ar}\end{aligned}\quad (2.103)$$

$$\langle \Psi_0 | \hat{H} | \Psi_a^r \rangle = \sqrt{2}(\varphi_a | \hat{f} | \varphi_r) \quad (2.104)$$

$$\langle \Psi_a^r | \hat{H} | \Psi_b^s \rangle = (\varphi_r | \hat{f} | \varphi_s) \delta_{ab} - (\varphi_a | \hat{f} | \varphi_b) \delta_{rs} - (\varphi_a \varphi_b | \varphi_r \varphi_s) + 2(\varphi_a \varphi_r | \varphi_b \varphi_s) \quad (2.105)$$

Triplet (Hartree-Fock):

$$\begin{aligned}\langle \Psi_a^r | \hat{H} - E_0 | \Psi_a^r \rangle &= (\varphi_r | \hat{f} | \varphi_r) - (\varphi_a | \hat{f} | \varphi_a) - (\varphi_a \varphi_a | \varphi_r \varphi_r) \\ &= \epsilon_r - \epsilon_a - J_{ar}\end{aligned}\quad (2.106)$$

$$\langle \Psi_0 | \hat{H} | \Psi_a^r \rangle = \sqrt{2}(\varphi_a | \hat{f} | \varphi_r) \quad (2.107)$$

$$\langle \Psi_a^r | \hat{H} | \Psi_b^s \rangle = (\varphi_r | \hat{f} | \varphi_s) \delta_{ab} - (\varphi_a | \hat{f} | \varphi_b) \delta_{rs} - (\varphi_a \varphi_b | \varphi_r \varphi_s) \quad (2.108)$$

mit

\hat{H}	Ein- und Zwei-Elektronen-Hamilton-Operator
\hat{f}	Fock-Operator
E_0	Hartree-Fock Grundzustandsenergie-Erwartungswert
J	Coulomb-Integral
K	Austausch-Integral
ϵ	Hartree-Fock-Orbitalenergie
δ	Kronecker-Delta

In den Matrix-Elementen für Triplet-Zustände verschwinden erwartungsgemäß (d.h. wegen der Spin-Integration) die Austausch-Integrale K .

Verwendet man nun Kohn-Sham-Orbitale anstatt der Hartree-Fock-Determinanten, so sind einige Modifikationen notwendig; (man bleibt allerdings bei den HF-basierten Energieausdrücken): Während besetzte HF- und KS-Orbitale einander ähneln, sind virtuelle KS-Orbitale im Allgemeinen kompakter als die korrespondierende Hartree-Fock-Funktion. Dies hat zwei formalistische Konsequenzen:

1. Virtuelle Kohn-Sham-Orbitale sind in der Regel bereits recht gute Beschreibungen angeregter Zustände, so dass deren Energiedifferenzen $\Delta\epsilon_{ar} = \epsilon_r - \epsilon_a$ zweier Orbitale Ψ_r und Ψ_a eher die tatsächliche Anregungsenergie wiedergeben als ihre Hartree-Fock-Pendants. (Als Faustregel lässt sich sagen, dass HF-Differenzen eher zu groß und KS-Differenzen eher zu klein resultieren.) Das hat zur Folge, dass die Coulomb-„Korrektur“ ($-J_{ar}$), die für die wesentlich größeren HF-Orbital-Differenzen noch angemessen ist, KS-Orbitaldifferenzen

deutlich zu stark reduziert. Daher wird die Subtraktion des Coulomb-Integrals durch einen empirischen Faktor c_1 „abgepuffert“:

$$J_{ar} \longrightarrow c_1 J_{ar} \quad (2.109)$$

c_1 wurde anhand einer Reihe repräsentativer Moleküle angepasst und variiert je nach verwendetem Funktional zwischen 0.2 und 0.4. Er beeinflusst in gleicher Weise jedes Matrix-Element mit Coulomb-Termen. Das "Becke Half & Half-Funktional (Becke [1993a]) erwies sich hier als günstigste Wahl.

2. Matrix-Elemente der Form $\langle \Psi_0 | \hat{H} | \Psi_a^r \rangle$ liefern einen Energiebeitrag durch *de-facto*-Relaxation des angeregten Molekülorbitals Ψ_a^r bezüglich des Grundzustandes Ψ_0 . Diese Relaxation ist bei Kohn-Sham-Orbitalen nicht mehr notwendig, so dass man diese Beiträge vernachlässigt. (Im HF-Fall sind solche Matrix-Elemente ohnehin Null wegen des Brillouin-Theorems.) Gleiches gilt für Kopplungen von einfach angeregten Konfigurationen Ψ_a^r untereinander:

$$\langle \Psi_0 | \hat{H} | \Psi_a^r \rangle = 0 \quad (2.110)$$

$$\langle \varphi_r | \hat{f} | \varphi_s \rangle = 0 \quad (2.111)$$

Unzureichende Ergebnisse bei Innerschalen-angeregten Zustände und solchen Zuständen, für die Anregungen aus *lone-pair*- oder σ -Orbitalen in MOs mit kleinem räumlichen Überlapp eine Rolle spielen, veranlassten Grimme zu einer Verfeinerung des Verfahrens durch die Einführung eines weiteren Korrekturterms Δ als Funktion der Orbital-Energie ϵ und des Austausch-Integrals K :

$$\Delta \rightarrow \Delta(\epsilon, K) = -0,025\epsilon + c_2 e^{-c_3 K^4} \quad (2.112)$$

Zusammenfassend erhält man folgende, DFT-adaptierte Matrix-Elemente für SCI-Berechnungen:

$$\langle \Psi_a^r | \hat{H} - E_0 | \Psi_a^r \rangle = \epsilon_r^{\text{KS}} - \epsilon_a^{\text{KS}} - c_1 J_{ar} + 2K_{ar} + \Delta(\epsilon_{ar}, K_{ar}) \quad (2.113)$$

$$\langle \Psi_0 | \hat{H} | \Psi_a^r \rangle = 0 \quad (2.114)$$

$$\langle \Psi_a^r | \hat{H} | \Psi_b^s \rangle = -c_1 (\varphi_a \varphi_b | \varphi_r \varphi_s) + 2(\varphi_a \varphi_r | \varphi_b \varphi_s) \quad (2.115)$$

Das beschriebene Verfahren liefert gute Ergebnisse, aber es stößt an seine Grenzen, wenn es um akkurate Berechnungen angeregter Zustände größerer Systeme geht. Es gab des Weiteren Probleme mit Zuständen, die biradikalischen Charakter bzw. solchen, die Doppelanregungscharakter aufweisen. Das führte 1999 zu der im folgenden Abschnitt erläuterten, methodischen Weiterentwicklung.

DFT/MRCI

Drei wesentliche Veränderungen konnten Präzision, Effizienz und den Anwendungsbereich des DFT/SCI-Verfahrens noch deutlich steigern:

- Durch die Vernachlässigung höherer als Einfach-Anregungs-Determinanten im SCI-Verfahren nimmt man einen systematischen Fehler in Kauf. Da dies erfahrungsgemäß die Genauigkeit der Berechnungen von angeregten Zuständen stark beeinträchtigt, werden prinzipiell auch Doppelt-Anregungs-Determinanten einbezogen:

$$|\Phi\rangle = c_0 |\Psi_0\rangle + \sum_{ra} c_a^r |\Psi_a^r\rangle + \sum_{\substack{a<b \\ r<s}} c_{rs}^{ab} |\Psi_{rs}^{ab}\rangle \quad (2.116)$$

Der zugrunde liegende Algorithmus ist also nunmehr ein **SDCI**.

- Zur genauen Beschreibung elektronisch angeregter Zustände erscheint es sinnvoll, die wichtigsten Konfigurationen als Grundlage für weitere Anregungen zuzulassen, so dass nicht mehr nur die Determinante des Grundzustandes als Referenz dient, sondern vor der Rechnung eine bestimmte Auswahl getroffen werden kann. Ein solcher Ansatz erhält das Präfix „Multireferenz“ **MR**.
- Damit auch große System mit vertretbarem rechnerischen Aufwand kalkulierbar bleiben, reduziert man den stark expandierten Multireferenz-Singles-Doubles-Wechselwirkungsraum auf eine Reihe automatisch anhand ihres zu erwartenden Beitrages ausgewählter Konfigurationen.

Das beschriebene DFT/MRCI-Verfahren empfiehlt sich durch ausgezeichnete Ergebnisse bei vergleichsweise geringem Rechenaufwand auch und gerade für angeregte Zustände. Die Tatsache, dass die Methode ihre Qualität auch bei Berechnungen an Molekülen außerhalb des Testsets unter Beweis stellt, unterstützt die These, dass die Konstanten $c_1 - c_5$ physikalisch fundiert sind, auch wenn sie empirisch ermittelt wurden. Auch ein „Doppelt-Zählen“ dynamischer Korrelationsbeiträge durch Kohn-Sham-Orbitale einerseits und höhere Anregungen andererseits entfällt so.

Kapitel 3

Das Konzept der Potenzialflächen

Auf der Grundlage der Born-Oppenheimer Näherung (Abschnitt 2.2, S. 7) separiert man Kern- und Elektronen-Koordinaten und formuliert die elektronische Schrödinger-Gleichung. Die im voran Gegangenen vorgestellten Näherungen und Berechnungsmethoden widmen sich ausschließlich der Ermittlung elektronischer Energien durch Lösung der elektronischen Schrödinger-Gleichung bei fixierter Kern-Geometrie; man nennt sie daher Einzelpunktberechnungen. Bei Änderung der Kern-Geometrie ändert sich in der Regel auch der berechnete Energie-Eigenwert des Systems, so dass man einen funktionalen Zusammenhang herstellen kann zwischen der System-Metrik, gegeben durch einen Satz Kern-Koordinaten $\{R_\alpha\}$, und der ihr zugeordneten elektronischen Energie E_{elek} :

$$E_{\text{elek}} = E_{\text{elek}}(\{R_\alpha\}) \quad (3.1)$$

Der „Graf“ der mehrdimensionalen Funktion $E_{\text{elek}}(\{R_\alpha\})$ wird Potenzialfläche¹ genannt. Jeder Punkt auf einer Potenzialfläche repräsentiert die vollständige Molekülgeometrie und dessen elektronische Energie.

Die Möglichkeiten zur Definition der Kern-Koordination als Abszisse der Potenzialfläche und charakteristische Eigenschaften der Hyperfläche selbst sind Gegenstand dieses Kapitels.

3.1 Kern-Koordinaten

Die formalistische Beschreibung einer chemischen Struktur erfolgt gewöhnlich durch Kern-Koordinaten. Man unterscheidet globale und lokale Koordinaten:

¹Genau genommen handelt es sich bei drei und mehr Koordinaten um eine Potenzial-Hyperfläche.

Globale Koordinaten beschreiben die Lage von N Teilchen im Weltall. Als Bezugssystem wählt man je nach Problemstellung das orthonormale kartesische Koordinatensystem oder auch — z.B. im Falle translationssymmetrischer Verbindungen — so genannte Kristallkoordinaten, deren Bezugssystem durch die Einheitszelle definiert wird. Da die Lage eines Teilchens im Raum in beiden Bezugssystemen jeweils durch drei Raumkoordinaten eindeutig definiert ist, ergibt sich als Anzahl der Freiheitsgrade die dreifache Anzahl der Teilchen im System: $3N$. Dabei ist zu beachten, dass eine Änderung der Lage im Raum sowohl Translationen als auch Rotationen mit einbezieht, die die Innere Energie eines chemisch interessanten Systems nicht beeinflussen. Die Eliminierung von translatorischen und rotatorischen Freiheitsgraden führt zu einer Reduktion der relevanten Freiheitsgrade auf $(3N - 5)$ im Falle linearer und auf $(3N - 6)$ im Falle nicht-linearer Moleküle.

Interne (Lokale) Koordinaten sind eine Alternative, die auf die Wahl eines Bezugssystems verzichten kann, da nicht ihre absolute, sondern ihre Lage relativ zueinander definiert wird. Die Vorgehensweise der Ermittlung interner Koordinaten ähnelt einem Pfad durch das Molekül, beginnend mit Kern 1 („Aufpunkt“, kartesische Koordinaten \mathbf{R}_1), dessen Distanzen und Winkel relativ zu bereits definierten Kernen in einer Z-Matrix festgehalten werden:

	Kernabstand	Dreieckswinkel	Diederwinkel
Kern 1:	(\mathbf{R}_1)		
Kern 2:	$ \mathbf{R}_2 - \mathbf{R}_1 $		
Kern 3:	$ \mathbf{R}_3 - \mathbf{R}_\alpha $	$\angle_3(1, 2, 3)$	
Kern 4:	$ \mathbf{R}_4 - \mathbf{R}_\alpha $	$\angle_4(\alpha, \beta, 4)$	$\angle_4(1, 2, 3, 4)$
⋮	⋮	⋮	⋮
Kern x:	$ \mathbf{R}_x - \mathbf{R}_\alpha $	$\angle_x(\alpha, \beta, x)$	$\angle_x(\alpha, \beta, \gamma, x)$

Tabelle 3.1: Aufbau einer Z-Matrix

Auf Grund dieses rekursiven Algorithmus ergibt sich eine Eliminierung translatorischer und rotatorischer Redundanzen, so dass sich die Anzahl der Freiheitsgrade auf $(3N - 6)$ für nichtlineare Systeme bzw. $(3N - 5)$ für lineare Systeme reduziert. Der Verzicht auf ein kartesisches Bezugssystem ist Vor- und Nachteil zugleich: Zwar kann man sich auf beliebige vorher definierte Kerne beziehen, allerdings sind interne Koordinaten nicht eindeutig. Darüber hinaus existieren bei Molekülen mit mehr als drei Atomen mehr Möglichkeiten, interne Koordinaten zu definieren als wirkliche Freiheitsgrade.

Eine weitere Verringerung der Freiheitsgrade erreicht man durch die Verwendung **Symmetrie adaptierter Koordinaten**: irreduzible Strukturfragmente werden kartesisch durch eine Z-Matrix dargestellt und die komplette Struktur daraus mit Hilfe der entsprechenden Symmetrioperationen definiert.

3.2 Eigenschaften von Potenzialflächen

3.2.1 Der eindimensionale Fall

Betrachten wir als einfachsten Fall eines zusammengesetzten chemischen Systems ein zweiatomiges Molekül. Da es sich dabei um eine lineare Anordnung handeln muss, hat dieses System genau einen Freiheitsgrad ($3 \cdot 2 - 5 = 1$). Seine potenzielle Energie ist also eine Funktion einer einzigen Variablen, des Abstandes x der beteiligten Atome. Da dieses System schwingungstheoretisch von Bedeutung ist, wurde es schon sehr früh stark vereinfacht als harmonischer Oszillator beschrieben:

$$\begin{aligned} V(x) &= \frac{1}{2}Dx^2 \\ F(x) &= -\frac{dV}{dx} = -Dx \end{aligned} \tag{3.2}$$

Charakteristische Größen sind die Kraftkonstante D und die Kraft F , mit der ein Kern in das Minimum des Potenzialtopfes, seinen Gleichgewichtsabstand, gezogen wird. Wir werden später noch auf diese anschauliche Quantifizierung eines Potentials zurückkommen.

Der anharmonische Oszillator ist eine wesentlich präzisere Beschreibung des zweiatomigen Moleküls. Der Vergleich von harmonischem und anharmonischem Oszillator in Abbildung 3.1 zeigt einerseits, dass sich die potenzielle Energie des Systems in einem Zustand nahe des Minimums sehr gut harmonisch annähern lässt, andererseits aber, dass die energetischen Verhältnisse in Molekülen beliebig kompliziert werden können.

Da das Potenzial des zweiatomigen Moleküls also mehr oder weniger genau bekannt ist, lassen sich damit einige wichtige Daten, wie z.B. Gleichgewichtsgeometrie und Schwingungsniveaus, berechnen.

3.2.2 Der zweidimensionale Fall

Abhängigkeiten der elektronischen Energie von genau zwei Variablen ohne äußere Randbedingungen findet man im Umfeld molekularer Systeme nicht, da die Anzahl

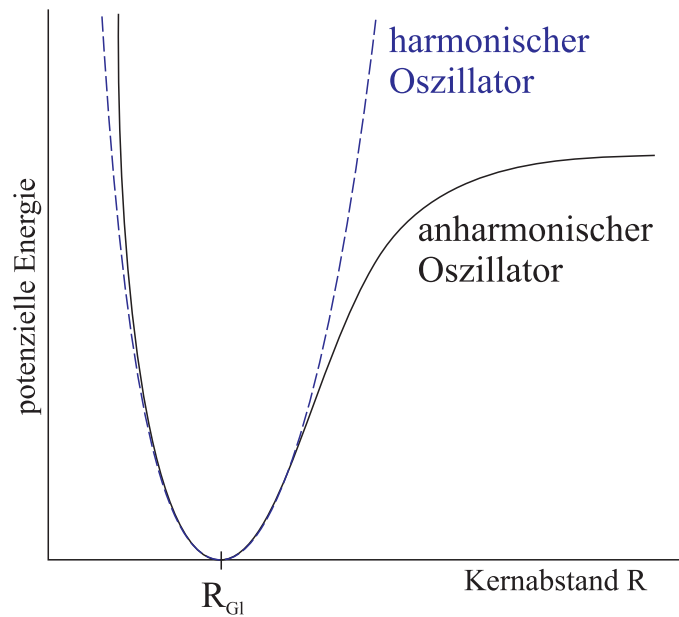


Abbildung 3.1: Schematischer Vergleich zwischen harmonischem und anharmonischem Oszillator

der Freiheitsgrade für keine Kernanzahl N den Wert Zwei annehmen kann². Da aber der zweidimensionale Fall exemplarischen Charakter hat, zumal die hier gewonnenen Erkenntnisse sich unmittelbar auf höhere Dimensionen übertragen lassen, sei er hier kurz behandelt.

Gehen wir vereinfachend von einer analytischen Funktion mit quadratischem Verhalten in Richtung beider Dimensionen aus. Die allgemeine Formulierung einer solchen Funktion ist:

$$f(x_1, x_2) = ax_1^2 + bx_1x_2 + cx_2^2 + dx_1 + ex_2 + g \quad (3.3)$$

In der Schreibweise der Linearen Algebra lautet dieselbe Vorschrift nach Einführung des Variablen-Vektors \mathbf{x} und seiner Transposition \mathbf{x}^T

$$f(\mathbf{x}) = \mathbf{x}^T \begin{pmatrix} a & b \\ b & c \end{pmatrix} \mathbf{x} + \begin{pmatrix} d \\ e \end{pmatrix}^T \mathbf{x} + g \quad \text{mit} \quad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \quad (3.4)$$

²Als Beispiel eines molekularen Systems mit einer Abhängigkeit der elektronischen Energie von zwei Variablen sei hier ein drei-atomiges Molekül genannt, in dem man den Abstand zweier Kerne \mathbf{R}_1 und \mathbf{R}_2 fixiert: $|\mathbf{R}_2 - \mathbf{R}_1| = R_{12} = \text{const}$. In einem solchen System ist der Betrag der Energie eindeutig definiert durch die Abstände des dritten Kernes von den fixierten Kernen, so dass gilt: $E = E(R_{13}, R_{23})$.

Die Matrix der Koeffizienten a , b und c ist bis auf den Faktor $1/2$ identisch mit einer Matrix der zweiten Ableitungen:

$$\mathbf{H} = \begin{pmatrix} \frac{\partial^2 f(\mathbf{x})}{\partial x_1 \partial x_1} & \frac{\partial^2 f(\mathbf{x})}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f(\mathbf{x})}{\partial x_2 \partial x_1} & \frac{\partial^2 f(\mathbf{x})}{\partial x_2 \partial x_2} \end{pmatrix} \quad (3.5)$$

Eine so konstruierte Matrix heißt Hesse-Matrix. Da zweite Ableitungen unabhängig von der Reihenfolge der Ableitungen sind, ist eine Hesse-Matrix symmetrisch, und es gilt allgemein

$$H_{ij} = \frac{\partial^2 f(\mathbf{x})}{\partial x_i \partial x_j} = \frac{\partial^2 f(\mathbf{x})}{\partial x_j \partial x_i} = H_{ji} \quad (3.6)$$

Die grafische Darstellung der Beispielfunktion in Gleichung 3.4 ist eine Oberfläche, die im Prinzip eine der folgenden drei Formen annehmen kann:

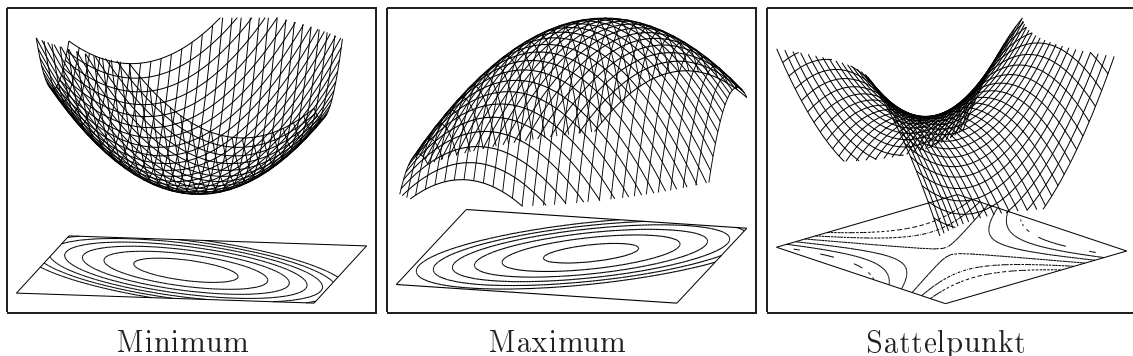
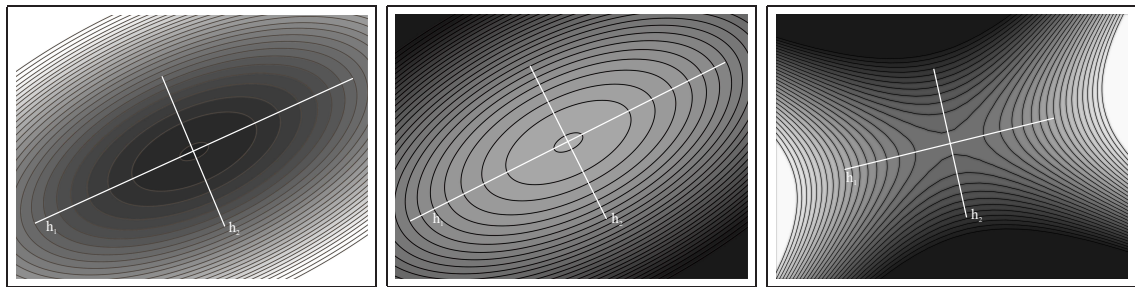


Abbildung 3.2: Stationäre Punkte bei zweidimensional quadratischen Funktionen

Diese Formen entsprechen, vereinfacht dargestellt, zwei Parabeln positiver Krümmung (Typ A \rightarrow Minimum), zwei Parabeln negativer Krümmung (Typ B \rightarrow Maximum) sowie der Kombination aus einer Parabel positiver und einer Parabel negativer Krümmung (Typ C \rightarrow Sattelpunkt). Das Krümmungsverhalten einer zweidimensional quadratischen Funktion wird in Vorzeichen und Betrag durch ihre Hesse-Matrix vollständig beschrieben. Die Kontur-Diagramme in Abbildung 3.3 illustrieren diese Zusammenhänge.

Aus der Perspektive der negativen z -Achse erkennt man deutlich die charakteristischen elliptischen Konturlinien. Dabei entsprechen die Ellipsenachsen \mathbf{h}_1 und \mathbf{h}_2 den Eigenvektoren der Hesse-Matrix, die jeweilige Ausdehnung in Achsenrichtung — eine Projektion der parabolischen Krümmung — ist anti-proportional ihren Eigenwerten: Je „kürzer“ die Achse, desto höher sind Krümmung und Eigenwert. Wie



Minimum

Maximum

Sattelpunkt

Abbildung 3.3: Kontur-Diagramme stationärer Punkte

im eindimensionalen Fall ist die Krümmung der Kurve verknüpft mit der Kraftkonstante D einer harmonischen Schwingung.

3.2.3 Der n -dimensionale Fall

Der zweidimensionale Fall ist der letzte, den man sich als eine Relief-artige Oberfläche im dreidimensionalen Raum vorstellen kann. Dennoch lassen sich alle dort gemachten Aussagen auf mehrdimensionale Fälle übertragen. Die allgemeine Formulierung einer n -dimensional quadratischen Funktion ist:

$$f(\mathbf{x}) = \frac{1}{2} \mathbf{x}^T \mathbf{H} \mathbf{x} + \mathbf{b}^T \mathbf{x} + c \quad (3.7)$$

Darin ist \mathbf{x} der n -dimensionale Abszissen-Vektor, \mathbf{b} ein n -dimensionaler Koeffizienten-Vektor, c ein Skalar und \mathbf{H} die $n \times n$ -dimensionale Hesse-Matrix.

Wie der Vergleich von harmonischem und anharmonischem Oszillator (Abbildung 3.1, S. 40) vermuten lässt, sind die Verhältnisse auf realen Potenzialflächen ungleich komplizierter. Quantenchemische Verfahren liefern keinerlei „Funktionsvorschrift“ in Form von Gleichung 3.7, die eine Potenzialfläche über den gesamten Raum analytisch definieren würde. Um die charakteristischen Punkte auf einer Hyperfläche dennoch ermitteln zu können, benutzt man erste und zweite Ableitungen der Form

$$\mathbf{g}(\mathbf{x}) = \begin{pmatrix} \frac{\partial f(\mathbf{x})}{\partial x_1} \\ \vdots \\ \frac{\partial f(\mathbf{x})}{\partial x_n} \end{pmatrix} = \nabla f(\mathbf{x}), \quad \mathbf{H}(\mathbf{x}) = \begin{pmatrix} \frac{\partial^2 f(\mathbf{x})}{\partial x_1 \partial x_1} & \cdots & \frac{\partial^2 f(\mathbf{x})}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f(\mathbf{x})}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 f(\mathbf{x})}{\partial x_n \partial x_n} \end{pmatrix} \quad (3.8)$$

Der Vektor der ersten Ableitungen $\mathbf{g}(\mathbf{x})$ heißt Gradient. Er ist identisch mit einer Anwendung des Nabla-Operators ∇ auf die Funktion $f(\mathbf{x})$. Wie eine Steigung im

eindimensionalen Fall weist er in jedem Punkt \mathbf{x} in Richtung des steilsten Anstiegs auf der Hyperfläche³. Die Hesse-Matrix ist im allgemeinen Fall nicht-quadratischer Funktionen abhängig von der Stelle \mathbf{x} . Auch hier quantifiziert sie die Krümmungen der Potenzialfläche im Punkt \mathbf{x} . Man klassifiziert sie nach den Vorzeichen ihrer Eigenwerte λ_i :

$$\begin{aligned} \lambda_i > 0 & \quad \text{für alle } i \in \{1, \dots, n\} & \Leftrightarrow & \quad \mathbf{H} \text{ ist positiv definit} \\ \lambda_i < 0 & \quad \text{für alle } i \in \{1, \dots, n\} & \Leftrightarrow & \quad \mathbf{H} \text{ ist negativ definit} \\ \lambda_i \geq 0 & \quad \text{für alle } i \in \{1, \dots, n\} & \Leftrightarrow & \quad \mathbf{H} \text{ ist positiv semidefinit} \\ \lambda_i \leq 0 & \quad \text{für alle } i \in \{1, \dots, n\} & \Leftrightarrow & \quad \mathbf{H} \text{ ist negativ semidefinit} \\ \lambda_i \begin{matrix} \leq \\ > \end{matrix} 0 & \quad \text{für alle } i \in \{1, \dots, n\} & \Leftrightarrow & \quad \mathbf{H} \text{ ist indefinit} \end{aligned}$$

Maxima, Minima und Sattelpunkte sind die markanten Punkte auf einer Potenzialfläche. Man nennt sie **stationäre Punkte**, weil in ihnen der Gradient verschwindet:

$$\mathbf{g}(\mathbf{x}_{min}) = \mathbf{g}(\mathbf{x}_{max}) = \mathbf{g}(\mathbf{x}_{sattel}) = 0 \quad (3.9)$$

Chemisch interessant sind zum einen Minima, weil sie die Gleichgewichtsgeometrie eines Moleküls definieren, und zum anderen Sattelpunkte erster Ordnung, weil man sie mit aktivierten Komplexen im Sinne der Eyring-Theorie identifiziert (Eyring [1935]). Hinreichende Bedingung für ein Minimum ist eine positiv definite Hesse-Matrix (\rightarrow die Potenzialfläche ist in alle Richtungen positiv gekrümmt); ein Sattelpunkt erster Ordnung liegt vor, wenn die Hesse-Matrix einen negativen und ansonsten ausschließlich positive Eigenwerte hat.

³ Man findet auch Definitionen gemäß $\mathbf{g}(\mathbf{x}) = -\nabla f(\mathbf{x})$, wonach der Gradient an jeder Stelle \mathbf{x} die Richtung des steilsten Abstiegs definiert. Wir verwenden im Folgenden die Definition aus Gleichung 3.8: $\mathbf{g}(\mathbf{x}) = +\nabla f(\mathbf{x})$

Kapitel 4

Geometrie-Optimierung

Geometrie-Optimierungen (Press u. a. [1992]) dienen in erster Linie der Ermittlung von Gleichgewichtsgeometrien in Molekülen, indem sie den so genannten **Schrittvektor** \mathbf{x} berechnen, der eine beliebige Start-Geometrie \mathbf{R} in die Gleichgewichtsgeometrie \mathbf{R}_{eq} überführt:

$$\mathbf{R} + \mathbf{x} = \mathbf{R}_{\text{eq}} \quad (4.1)$$

Gleichgewichtsgeometrien identifiziert man anhand ihrer minimalen (potenziellen) Energie. Sei ε ein N -dimensionaler Vektor beliebig kleiner Komponenten endlicher Größe $\varepsilon_i > 0$, dann gilt für ein lokales Minimum E_{eq} :

$$E_{\text{eq}} = E(\mathbf{R}_{\text{eq}}) < E(\mathbf{R} + \varepsilon) \quad (4.2)$$

Das globale Minimum E_{global} wird mit dem Grundzustand des Systems identifiziert:

$$E_{\text{global}} = E(\mathbf{R}_{\text{global}}) < E(\mathbf{R} \neq \mathbf{R}_{\text{global}}) \quad (4.3)$$

Daher sei im Folgenden die Gleichgewichtsgeometrie \mathbf{R}_{eq} als \mathbf{R}_{min} notiert. Notwendige Bedingung für ein Minimum ist der in \mathbf{R}_{min} verschwindende Gradient:

$$\mathbf{g}(\mathbf{R}_{\text{min}}) = 0 \quad (4.4)$$

Hinreichende Bedingung ist eine positiv-definite Hesse-Matrix $\mathbf{H}(\mathbf{R}_{\text{min}})$.

Da eine exakte analytische Potenzialvorschrift für den gesamten Raum nicht zugänglich ist, kann der gesuchte Schrittvektor \mathbf{x} nicht direkt bestimmt werden. Statt dessen geht man iterativ vor:

1. Ermittlung lokaler Informationen am Startpunkt \mathbf{R}_i durch Annäherung und Analyse des lokalen Bereiches der Hyperfläche durch ein geeignetes **Modellpotenzial**
2. Änderung der Geometrie zu \mathbf{R}_{i+1} unter Nutzung der gewonnenen lokalen Informationen

Eine solche schrittweise rekursive Annäherung der Gleichgewichtsgeometrie wird Fixpunktverfahren genannt:

$$\mathbf{R}_{i+1} = \mathbf{R}_i + \mathbf{x}_i \quad (4.5)$$

Nach n Iterationen über $(n + 1)$ Stützpunkte \mathbf{R}_i ist die Geometrie hinreichend konvergiert, und es gilt:

$$\mathbf{R}_n \approx \mathbf{R}_{\min} \Rightarrow \mathbf{x} \approx \sum_{i=1}^n \mathbf{x}_i \quad (4.6)$$

Die gängigsten Optimierungsverfahren unterscheiden sich primär in der Art des verwendeten Modellpotenzials und der Nutzung der lokalen Informationen.

Als Modellpotenzial für eine lokale Approximation der Hyperfläche benutzt man Taylor-Reihen oder rationale Funktionen. Bei einer **Taylor-Reihen-Entwicklung** definiert man zur Ermittlung lokaler Informationen einen Störvektor \mathbf{p} mit den Komponenten $(\Delta R_1, \Delta R_2, \dots, \Delta R_n)$ und setzt an:

$$E(\mathbf{R}_i + \mathbf{p}) = E_i + \mathbf{g}_i^T \mathbf{p} + \frac{1}{2} \mathbf{p}^T \mathbf{H}_i \mathbf{p} + \text{Terme höherer Ordnung} \quad (4.7)$$

$$\text{mit } E_i \equiv E(\mathbf{R}_i), \quad \mathbf{g}_i \equiv \mathbf{g}(\mathbf{R}_i) \quad \text{und} \quad \mathbf{H}_i \equiv \mathbf{H}(\mathbf{R}_i)$$

Diese Näherung ist um so präziser, je kleiner die Norm des Störvektors $\|\mathbf{p}\|$ ist.

Je nach rechnerischem Aufwand am Fixpunkt \mathbf{R}_i unterscheidet man drei Klassen von Optimierungsverfahren:

1. „Function-Only“-Methoden untersuchen den lokalen Bereich des Fixpunktes \mathbf{R}_i nicht. Sie optimieren lediglich auf der Grundlage von Einzelpunkt-Energien E_i .
2. Gradienten-Verfahren ermitteln die erste Ableitung \mathbf{g}_i und nutzen sie auf der Grundlage von Taylor-Reihen erster oder zweiter Ordnung.
3. Quadratische Methoden berechnen zusätzlich zum Gradienten auch die Hesse-Matrix \mathbf{H}_i . Modellpotenzial ist entweder ebenfalls eine Taylor-Reihe zweiter Ordnung oder eine rationale Funktion.

Diese drei Varianten werden in den kommenden Abschnitten erläutert.

4.1 Pfadfinder-Algorithmen

Nicht alle Optimierungs-Verfahren liefern die gesuchten iterativen Schrittvektoren \mathbf{x}_i auf direktem Weg. Statt dessen bestimmen sie zunächst eine Suchrichtung \mathbf{d}_i , den „Pfad“, und ermitteln anschließend die Schrittlänge s_i , die in dieser Richtung zu einem Minimum führt. Diese Vorgehensweise nennt man Pfadfinder-Algorithmus:

$$\mathbf{x}_i = s_i \mathbf{d}_i \quad (4.8)$$

Während die Bestimmung der Suchrichtung je nach Verfahren sehr unterschiedlich sein kann, so folgt die Ermittlung der Schrittlänge s_i im Allgemeinen einer festgelegten Methode, die man **Line Search** nennt. Der *Line Search* ist somit die „Elementarzelle“ von Pfadfinder-Algorithmen und soll hier genauer beleuchtet werden.

Deformiert man ein Molekül beginnend mit einer beliebigen Geometrie \mathbf{R}_0 entlang einer vorgegebenen Richtung \mathbf{d} um die (skalare) Länge s gemäß

$$\mathbf{R}_d(s) = \mathbf{R}_0 + s\mathbf{d}, \quad (4.9)$$

so kommt das einem Schnitt durch eine Potenzial-Hyperfläche gleich. Auf jedem Punkt $\mathbf{R}_d(s)$ des Richtungs-Vektors \mathbf{d} „spürt“ das Molekül ein bestimmtes Potenzial $E[\mathbf{R}_d(s)]$. Infolgedessen kann man die Abszisse durch die lineare Beziehung (4.9) substituieren und das Potenzial $E[\mathbf{R}_d(s)]$ als eindimensionale (verknüpfte) Funktion $E_d(s)$ einer Elongation s auf dem Richtungs-Vektor \mathbf{d} darstellen:

$$E[\mathbf{R}_d(s)] = E \circ \mathbf{R}_d(s) := E_d(s) \quad (4.10)$$

Abbildung 4.1 illustriert die iterativen Schritte eines *Line Searches*: Schwarz dar-

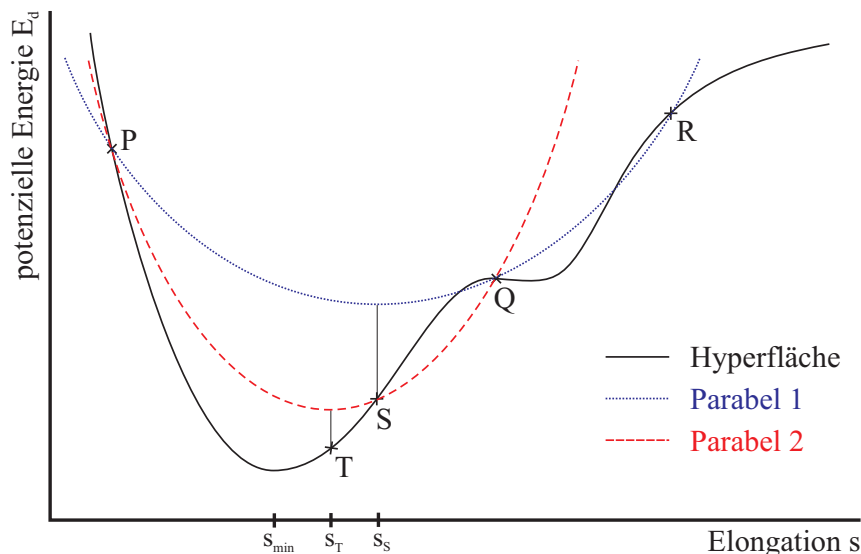


Abbildung 4.1: Projektion einer Potenzial-Hyperfläche und Ermittlung des Minimums durch inverse parabolische Interpolation

gestellt ist die Projektion der Hyperfläche in der vorgegebenen Suchrichtung \mathbf{d} . In einem ersten Schritt (S. 73) berechnet man die Punkte P, Q und R auf der Hyperfläche:

$$\begin{aligned} P: \text{ Geometrie } \mathbf{R}_d(s_P) = \mathbf{R}_0 + s_P\mathbf{d} &\Rightarrow \text{ projizierte Koordinaten: } s_P, E_d(s_P) \\ Q: \text{ Geometrie } \mathbf{R}_d(s_Q) = \mathbf{R}_0 + s_Q\mathbf{d} &\Rightarrow \text{ projizierte Koordinaten: } s_Q, E_d(s_Q) \\ R: \text{ Geometrie } \mathbf{R}_d(s_R) = \mathbf{R}_0 + s_R\mathbf{d} &\Rightarrow \text{ projizierte Koordinaten: } s_R, E_d(s_R) \end{aligned}$$

Im zweiten Schritt approximiert man den Potenzialflächen-Schnitt als quadratische Funktion und berechnet das Minimum s_S einer Parabel durch die drei Punkte P, Q und R mittels **inverser parabolischer Interpolation**:

$$s_S = s_Q - \frac{1}{2} \frac{(s_Q - s_P)^2 [E_d(s_Q) - E_d(s_R)] - (s_Q - s_R)^2 [E_d(s_Q) - E_d(s_P)]}{(s_Q - s_P)[E_d(s_Q) - E_d(s_R)] - (s_Q - s_R)[E_d(s_Q) - E_d(s_P)]} \quad (4.11)$$

Schritt drei ist die Berechnung der Energie bei der gerade ermittelten Geometrie des Minimums $\mathbf{R}_d(s_S)$:

$$S: \text{ Geometrie } \mathbf{R}_d(s_S) = \mathbf{R}_0 + s_S \mathbf{d} \Rightarrow \text{ projizierte Koordinaten: } s_S, E_d(s_S)$$

Nun wiederholt man die inverse parabolische Interpolation für das Punkte-Tripel P, S und Q , erhält als neues Minimum Punkt T , der wiederum als Grundlage für weitere Iterationen dient, solange, bis nach n Schritten die Minimierung als konvergiert angesehen wird:

$$s \approx s_{\min} \quad (4.12)$$

Das Produkt der so ermittelten Schrittlänge s_i mit der Suchrichtung \mathbf{d}_i ist der gesuchte Schrittvektor der i -ten Iteration der Optimierung:

$$\mathbf{x}_i = s_i \mathbf{d}_i \quad (4.13)$$

Der $(i + 1)$ -te Stützpunkt ist folglich

$$\mathbf{R}_{i+1} = \mathbf{R}_i + s_i \mathbf{d}_i, \quad (4.14)$$

und die Rekursion ist vollständig.

4.2 „Function-Only“-Verfahren

Das „Function-Only“-Verfahren ist ein Pfadfinder-Algorithmus, der bei der Bestimmung der Suchrichtung \mathbf{d} auf lokale Informationen wie Gradient oder Hesse-Matrix verzichtet. Minimiert wird jeweils in Richtung der Einheitsvektoren $\{\mathbf{e}_i\}$ solange, bis Konvergenz eintritt.

Abbildung 4.2 zeigt die iterativen Schritte dieses Algorithmus im vereinfachten Fall einer zweidimensional quadratischen Funktion. Von der Startgeometrie aus wird ein *Line Search* in Richtung $\mathbf{d}_1 = \mathbf{e}_y$ durchgeführt.¹ Der resultierende Schrittvektor $\mathbf{x}_1 = s_1 \mathbf{e}_y$ entspannt das System zur Geometrie 1. Der nächste Suchpfad ist der Einheitsvektor in x -Richtung: $\mathbf{d}_2 = \mathbf{e}_x$. Er führt zu Stützpunkt 2 und analog zu 3 und

¹ Man beachte, dass eine erste Suche entlang der x -Achse (gestrichelt dargestellt) die Geometrie zwar entspannen, aber sehr weit vom Minimum entfernen würde.

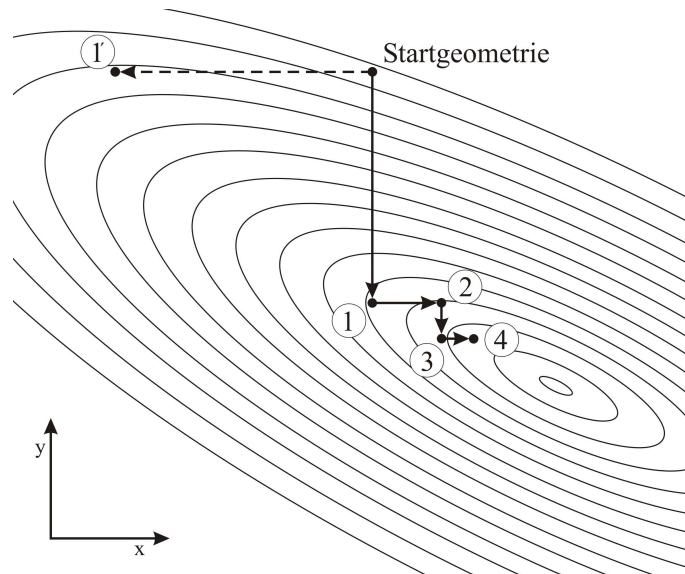


Abbildung 4.2: Iterationen einer „Function-Only“-Optimierung

4. Die weiteren Schritte jenseits Geometrie 4 zeigen ebenfalls eine „treppenstufenartige“ Annäherung an das Minimum mit sinkender Effizienz, d.h. Stufenhöhe und -tiefe werden immer geringer, so dass sie hier nicht mehr dargestellt wurden.

Die Konvergenz dieses Verfahrens ist denkbar schlecht. Insbesondere beim Auftreten „schmaler diagonalen Täler“, d.h. in Fällen stark koppelnder Variablen und stark unterschiedlicher Hesse-Matrix-Eigenwerte, sind bis zu einer befriedigenden Konvergenz unökonomisch viele Schritte notwendig. Dennoch ist das Verfahren sehr robust und zeigt sich unempfindlich gegenüber möglichen Unstetigkeiten auf der Hyperfläche. Auch in Fällen schwer zugänglicher lokaler Informationen kann die „Function-Only“-Optimierung eine sinnvolle Alternative sein.

4.3 Gradienten-Verfahren

In Gradienten-Verfahren berechnet man die erste Ableitung in Fixpunkt \mathbf{R}_i für die Ermittlung des nächsten Schrittes \mathbf{x}_i . Der formalistische Aufwand für eine Schrittvektor-Bestimmung kann sehr unterschiedlich sein. Auffällig dabei ist, dass einfache Suchalgorithmen langsamer konvergieren, dafür aber meist sehr robust sind, während umgekehrt Optimierungs-Methoden, die extensiv Gebrauch von lokalen Informationen machen, häufig sehr empfindlich auf Anharmonizitäten der Potenzialfläche reagieren, dafür aber bei gut modellierter Umgebung sehr schnell ins Minimum führen.

Die Berechnung des Gradienten \mathbf{g}_i im Stützpunkt \mathbf{R}_i kann entweder analytisch oder numerisch erfolgen. Dank der Wegweisenden Arbeiten von Pulay [1969], Handy und Schaefer [1984] und Anderen sehen manche quantenchemischen Programmpakete eine Möglichkeit der analytischen Gradienten-Berechnung vor. Ihre Vorteile sind, dass ihre Ermittlung nur etwa so aufwändig wie eine Einzelpunktberechnung ist, und dass analytische Gradienten im Allgemeinen sehr exakt sind. Falls keine analytischen Gradienten zur Verfügung stehen, kann der Gradient numerisch ermittelt werden.

4.3.1 Die Generierung numerischer Gradienten

Numerische Gradienten sind eine Approximation an analytisch (exakte) Gradienten durch finite Differenzen. Dabei erhält man die k -te Komponente $g_k(\mathbf{R})$ des Gradienten-Vektors $\mathbf{g}(\mathbf{R})$ durch selektive Deformation der Komponente R_k des Geometrie-Vektors \mathbf{R} mit dem Inkrement h :

$$g_k(\mathbf{R}) = \frac{\partial E(\mathbf{R})}{\partial R_k} \approx \frac{\Delta E(\mathbf{R})}{\Delta R_k} = \frac{E(\mathbf{R}') - E(\mathbf{R})}{h}$$

$$\text{mit } \mathbf{R} = \begin{pmatrix} R_1 \\ R_2 \\ \vdots \\ R_k \\ \vdots \\ R_n \end{pmatrix} \quad \text{und} \quad \mathbf{R}' = \begin{pmatrix} R_1 \\ R_2 \\ \vdots \\ R_k + h \\ \vdots \\ R_n \end{pmatrix} \quad (4.15)$$

Da für den vollständigen Gradienten die Energie $E(\mathbf{R})$ und n Energien $E(\mathbf{R}')$ berechnet werden müssen, sind numerische Kalkulationen oft deutlich aufwändiger als analytische. Ihr Vorteil ist, dass man sie auch für quantenchemische Verfahren anwenden kann, für die die Berechnung analytischer Gradienten zu kompliziert oder unmöglich ist.

Die eindimensionale Projektion der Hyperfläche in Richtung des Einheitsvektors \mathbf{e}_k und die Konstruktion des so genannten **Forward-Gradienten** ist in Abbildung 4.3 dargestellt. Ebenso könnte man den **Backward-Gradienten** in $(R_k - h)$ bestimmen, was aber die Übereinstimmung mit dem tatsächlichen Gradienten nicht prinzipiell verbessern würde.

Der Fehler dieser zusammenfassend als **Zwei-Punkt-Gradienten** bezeichneten numerischen Gradienten ist abhängig

1. vom Krümmungsverhalten der Hyperfläche
2. von der Größe des Inkrements h
3. von Fehlerquellen in der Einzelpunkt-Berechnungsmethode selbst.

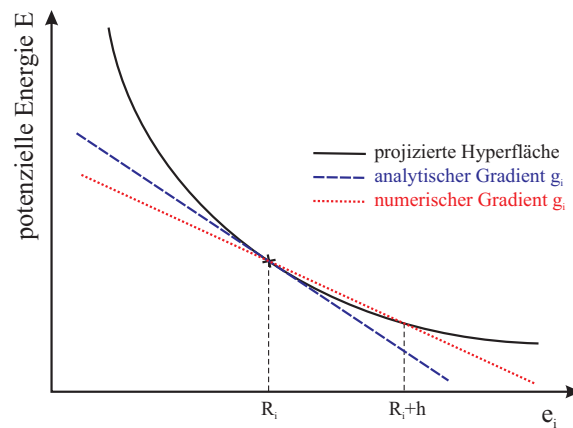


Abbildung 4.3: Konstruktion numerischer *Forward*-Gradienten

Fehler 1: Der durch die Potenzial-Krümmung hervorgerufene Fehler lässt sich verringern, indem man **zentrale** oder auch **Drei-Punkt-Gradienten** ermittelt, wie in Abbildung 4.4 illustriert. Die Komponente $g_k(\mathbf{R})$ eines zentralen Gradienten

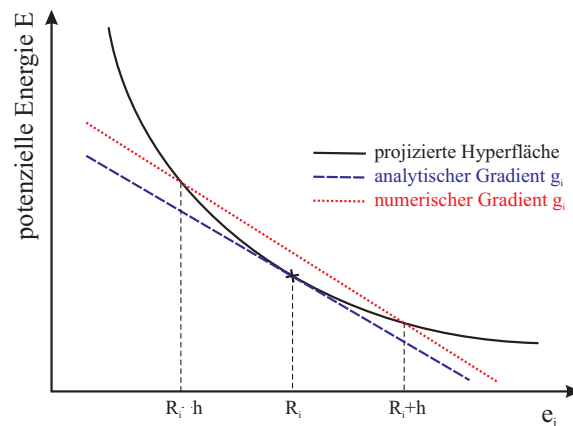


Abbildung 4.4: Konstruktion numerischer zentraler Gradienten

für die Geometrie \mathbf{R} ist

$$g_k(\mathbf{R}) = \frac{E(\mathbf{R}'') - E(\mathbf{R}')}{2h}$$

$$\text{mit } \mathbf{R}' = \begin{pmatrix} R_1 \\ R_2 \\ \vdots \\ R_k + h \\ \vdots \\ R_N \end{pmatrix} \quad \text{und} \quad \mathbf{R}'' = \begin{pmatrix} R_1 \\ R_2 \\ \vdots \\ R_k - h \\ \vdots \\ R_N \end{pmatrix} \quad (4.16)$$

Es handelt sich also gewissermaßen um das arithmetische Mittel aus *Forward*- und *Backward*-Gradient. Wie sich leicht zeigen lässt, kann man die Unempfindlichkeit zentraler Gradienten gegenüber dem Krümmungsverhalten der Hyperfläche dadurch erklären, dass sie sich durch Subtraktion zweier Taylor-Reihen (Gleichung 4.7) für $E(\mathbf{R} + \mathbf{p})$ und $E(\mathbf{R} - \mathbf{p})$ herleiten lassen, wobei der quadratische Term, der die Krümmung enthält, weg fällt.

Da jede Komponente eines zentralen Gradienten zwei Einzelpunkt-Berechnungen erfordert, verdoppelt sich der rechnerische Aufwand, jedoch zeigt Abbildung 4.5, dass er besonders in der Nähe des Minimums unverzichtbar ist. Der Zwei-Punkt-

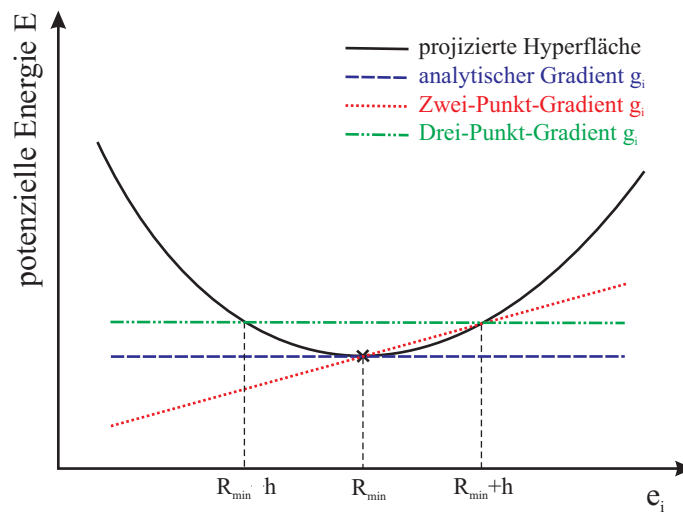


Abbildung 4.5: Die Notwendigkeit zentraler Gradienten in der Nähe eines Minimums

Gradient liefert im Potenzial-Minimum fälschlicherweise einen Gradienten ungleich Null, während der entsprechende zentrale Gradient das Minimum korrekt durch einen verschwindenden Gradienten identifiziert.

Fehler 2: Selbst ein Zwei-Punkt-Gradient ließe sich mit ausreichender Genauigkeit konstruieren, wenn man das Inkrement h beliebig klein wählen könnte:

$$\lim_{h \rightarrow 0} \mathbf{g}_{\text{numerisch}} = \mathbf{g}_{\text{exakt}} \quad (4.17)$$

Bedauerlicherweise überschreitet das die Präzision der verwendeten Einzelpunkt-Methoden. Nach Baker [1987] erhält man das jeweils geeignete Inkrement h in Abhängigkeit vom Energie-Fehler-Intervall Δ_{err} gemäß

$$h = 20 \sqrt{\frac{\Delta_{err}}{1 + |E(\mathbf{R})|}} \quad (4.18)$$

Fehler 3: Quantenchemische Programm-Pakete sehen in aller Regel Möglichkeiten der Einflussnahme auf die interne Genauigkeit der Berechnung vor, bspw. über die Spezifikation von Fehlertoleranzen und/oder Konvergenz-Kriterien. Von entscheidender Bedeutung für einen „stabilen“ numerischen Algorithmus ist dabei nicht der systematische Fehler einer quantenchemischen Methodik, sondern vielmehr das durch wenig akkurate Einzelpunkt-Rechnungen erzeugte numerische „Rauschen“, d.h. der zufällige Fehler.

4.3.2 Steepest Descent

Steepest Descent bezeichnet einen Pfadfinder-Algorithmus, bei dessen iterativen Schritten jeweils die Richtung des steilsten Abstieges auf der Hyperfläche, also der negative Gradient $-\mathbf{g}_i$ in Stützpunkt \mathbf{R}_i , als Suchrichtung angesetzt wird:

$$\mathbf{R}_{i+1} = \mathbf{R}_i - s_i \mathbf{g}_i \quad (4.19)$$

Dabei ermittelt man die Schrittweite s_i wie gewohnt durch einen *Line Search*.

Eine exemplarisch vereinfachte *Steepest-Descent*-Optimierung auf einer zweidimensional quadratischen Oberfläche ist in Abbildung 4.6 dargestellt. Es ist eine

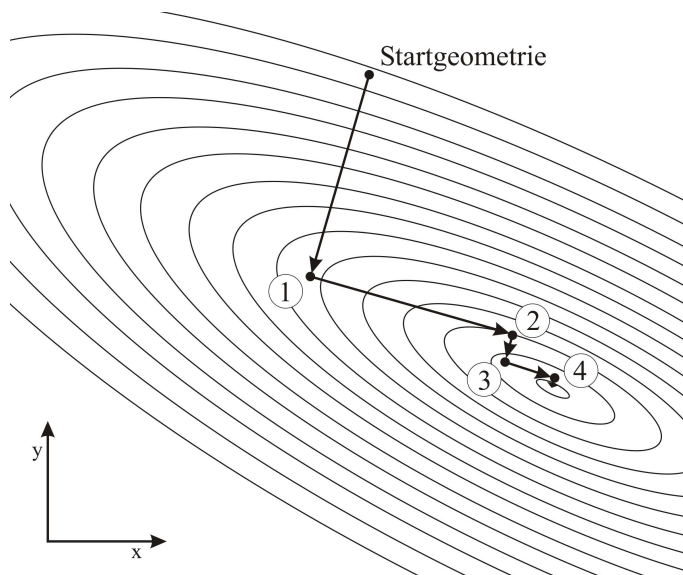


Abbildung 4.6: Iterationen einer *Steepest-Descent*-Optimierung

Folge des Ansatzes, dass aufeinanderfolgende Schritte in orthogonale Richtungen erfolgen. Der im Vergleich zum „*Function-Only*“-Algorithmus erhöhte Aufwand für die zusätzlich durchzuführende Gradienten-Ermittlung wird in der Regel durch das effizientere *lineare* Konvergenz-Verhalten überkompensiert.

4.3.3 Conjugate Gradient

Ein Nachteil des *Steepest Descent* ist, dass koppelnde Variablen („diagonale Täler“ auf der Hyperfläche) mehrfach minimiert werden müssen. Ziel des **Conjugate Gradient** ist es daher, einen Satz „unabhängiger“ Suchrichtungen $\{\mathbf{d}_i\}$ zu finden, also Richtungen, deren Minima im weiteren Verlauf der Optimierung erhalten bleiben.

Sei also \mathbf{d}_i eine beliebige Suchrichtung dieses Satzes, dann lässt sich die Hyperfläche als Taylor-Reihe zweiter Ordnung annähern:

$$E(\mathbf{d}_i) = E + \mathbf{g}^T \mathbf{d}_i + \frac{1}{2} \mathbf{d}_i^T \mathbf{H} \mathbf{d}_i \quad (4.20)$$

Man erhält den Gradienten in Suchrichtung gemäß

$$\nabla E(\mathbf{d}_i) = \mathbf{H} \mathbf{d}_i + \mathbf{g} . \quad (4.21)$$

Bedingung für ein Minimum ist das Verschwinden dieses Gradienten:

$$\begin{aligned} 0 &= \nabla E(\mathbf{d}_i) \\ &= \mathbf{H} \mathbf{d}_i + \mathbf{g} \end{aligned} \quad (4.22)$$

Die Änderung des Gradienten bei Änderung des Vektors \mathbf{d}_i ist demnach gegeben durch

$$\delta(\nabla E(\mathbf{d}_i)) = \mathbf{H}(\delta \mathbf{d}_i) \quad (4.23)$$

Die Forderung, dass sich der Gradient der letzten Suchrichtung \mathbf{d}_i bei einer Suche in Richtung \mathbf{d}_{i+1} nicht ändern darf, entspricht einer Forderung nach Orthogonalität:

$$\begin{aligned} 0 &= \mathbf{d}_{i+1}^T \delta(\nabla E(\mathbf{d}_i)) \\ &= \mathbf{d}_{i+1}^T \mathbf{H}(\delta \mathbf{d}_i) \\ &= \mathbf{d}_{i+1}^T \mathbf{H} \mathbf{d}_i \end{aligned} \quad (4.24)$$

Vektoren, die Bedingung (4.24) erfüllen, nennt man konjugiert. Benutzt man sie zur Konstruktion eines Satzes der Suchrichtungen $\{\mathbf{d}_i\}$ eines Pfadfinder-Verfahrens, so ist sicher gestellt, dass in keiner dieser Richtungen mehrmals minimiert werden muss. Man erwartet daher als Faustregel ebenso viele Iterations-Schritte wie Freiheitsgrade. Voraussetzung dafür sind die hinreichende Nähe zur quadratischen Umgebung des Minimums und möglichst exakte Gradienten. Abbildung 4.7 zeigt exemplarisch den Verlauf einer *Conjugate-Gradient*-Optimierung auf einer „ideal“ quadratischen Oberfläche.

Der *Conjugate-Gradient*-Algorithmus erzeugt also einen Satz Gradienten $\{\mathbf{g}_i\}$ und einen Satz konjugierter Suchrichtungen $\{\mathbf{d}_i\}$, die die folgenden Orthogonalitäts- und Konjugations-Bedingungen erfüllen:

$$\mathbf{g}_i \mathbf{g}_j = 0, \quad \mathbf{g}_i \mathbf{d}_j = 0, \quad \mathbf{d}_i^T \mathbf{H} \mathbf{d}_j = 0 \quad \text{für alle } i \neq j \quad (4.25)$$

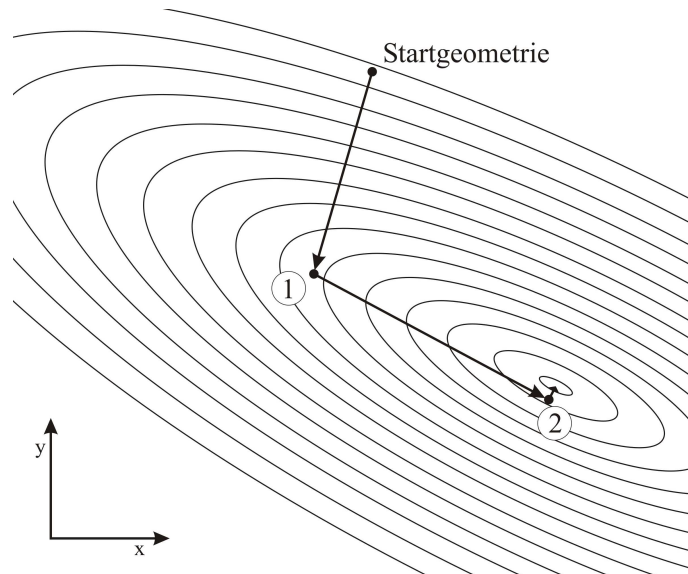


Abbildung 4.7: Optimierung auf einer zweidimensional quadratischen Oberfläche nach dem *Conjugate-Gradient*-Verfahren

Die rekursiven Konstruktionsvorschriften lauten dann:

$$\mathbf{g}_{i+1} = \mathbf{g}_i - \lambda_i \mathbf{H} \mathbf{d}_i \quad \text{und} \quad (4.26)$$

$$\mathbf{d}_{i+1} = \mathbf{g}_{i+1} + \gamma_i \mathbf{d}_i \quad , \quad (4.27)$$

wobei die Skalare λ_i und γ_i gegeben sind als

$$\lambda_i = \frac{\mathbf{g}_i^T \mathbf{g}_i}{\mathbf{d}_i^T \mathbf{H} \mathbf{d}_i} = \frac{\mathbf{g}_i^T \mathbf{d}_i}{\mathbf{d}_i^T \mathbf{H} \mathbf{d}_i} \quad (4.28)$$

$$\gamma_i^{\text{FR}} = \frac{\mathbf{g}_{i+1}^T \mathbf{g}_{i+1}}{\mathbf{g}_i^T \mathbf{g}_i} \quad (4.29)$$

Die letzte Gleichung dient der Bestimmung des Wertes γ_i nach Fletcher und Reeves [1964]). Auf einer quadratischen Hyperfläche mit bekannter Hesse-Matrix \mathbf{H} , bekanntem Start-Gradienten \mathbf{g}_0 und dem Ansatz $\mathbf{d}_0 = \mathbf{g}_0$ ließen sich aufgrund dieser Gleichungen beide Sätze $\{\mathbf{g}_i\}$ und $\{\mathbf{d}_i\}$ berechnen,² und man erreicht in n Schritten das Minimum, wenn n die Zahl der Freiheitsgrade ist. Bedauerlicherweise aber ist die Hesse-Matrix für gewöhnlich unbekannt oder nur mit größerem Rechenaufwand zu ermitteln.

Dennoch erhält man eine Lösung für Gleichung (4.26), indem man den Pfadfinder-Algorithmus einsetzt: *Line Search* in Richtung \mathbf{d}_i führt gerade zur Geometrie

² Mit $\mathbf{d}_0 = \mathbf{g}_0$ entspricht der erste Schritt im *Conjugate-Gradient*-Formalismus also *de facto* einem *Steepest-Descent*-Schritt.

\mathbf{R}_{i+1} , deren Gradient \mathbf{g}_{i+1} sich nun numerisch oder analytisch ermitteln lässt. Die darauf folgende konjugierte Suchrichtung \mathbf{d}_{i+1} wird nach Gleichung (4.27) berechnet, und der iterative Kreis ist geschlossen.

Die Konvergenz des *Conjugate-Gradient*-Algorithmus hängt davon ab, wie exakt die Gradienten bestimmt werden, wie akkurat die *Line Searches* durchgeführt werden und wie gut die Näherung als Taylor-Reihe zweiter Ordnung die lokalen Verhältnisse wiedergibt, d.h. wie nah die Start-Geometrie am Minimum liegt. Polak und Ribière (Polak [1971]) schlugen eine Modifikation der Ermittlung des Skalars γ_i vor, die die Anwendbarkeit des *Conjugate Gradient* auf stark anharmonische Potenzialflächen bzw. größere Anfangs-Entfernungen vom Minimum ausdehnt:

$$\gamma_i^{\text{PR}} = \frac{(\mathbf{g}_{i+1} - \mathbf{g}_i)^T \mathbf{g}_{i+1}}{\mathbf{g}_i^T \mathbf{g}_i} \quad (4.30)$$

Wegen der Orthogonalität von \mathbf{g}_i und \mathbf{g}_{i+1} nach Gleichung (4.25) sind die Bestimmungsgleichungen von Fletcher-Reeves und Polak-Ribière in „ideal“ quadratischer Umgebung äquivalent. Auf realen nicht-harmonischen Oberflächen dagegen kann es vorkommen, dass eine konjugierte Suchrichtung \mathbf{d}_i keine effektive Minimierung der Energie bewirkt, so dass die im *Line Search* ermittelte Schrittlänge s_i beinahe verschwindet. Die Folge-Geometrie \mathbf{R}_{i+1} liegt nahe bei \mathbf{R}_i , folglich sind die Gradienten \mathbf{g}_i und \mathbf{g}_{i+1} sehr ähnlich, und die Suchrichtung \mathbf{d}_{i+1} entspricht dem Gradienten $-\mathbf{g}_{i+1}$. Betrachten wir den Extremfall $s_i = 0$, dann gilt:

$$\begin{aligned} \mathbf{R}_{i+1} &= \mathbf{R}_i \\ \Rightarrow \mathbf{g}_{i+1} &= \mathbf{g}_i \\ \Rightarrow \gamma_i^{\text{PR}} &= \frac{0 \cdot \mathbf{g}_{i+1}}{\mathbf{g}_i^T \mathbf{g}_i} = 0 \\ \Rightarrow \mathbf{d}_{i+1} &= \mathbf{g}_{i+1} + 0 \cdot \mathbf{d}_i = \mathbf{g}_{i+1} \end{aligned} \quad (4.31)$$

Der Einfluss des steilsten Abstieges $-\mathbf{g}_{i+1}$ wird auf diese Weise mit der Ineffektivität der letzten Suchrichtung \mathbf{d}_i gewichtet und resultiert im oben dargestellten Extremfall $s_i = 0$ in einer Neu-Initialisierung des *Conjugate-Gradient*-Algorithmus durch einen *Steepest-Descent*-Schritt ($i+1$). Auf diese Weise werden nicht nur Anharmonizitäten der Hyperfläche, sondern auch Fehler im Gradienten abgefangen. Conjugate-Gradient-Optimierungen vom Polak-Ribière-Typ gelten daher als sehr robust und eignen sich ganz besonders für die Implementierung in numerische Verfahren.

4.3.4 Quasi-Newton-Algorithmen

Die Hyperfläche eines Systems mit n Freiheitsgraden ist im Bereich seines Minimums durch das Modellpotenzial einer n -dimensionalen Taylor-Reihe zweiter Ordnung (siehe Gleichung 4.20) gut approximiert. Prinzipiell enthalten Hesse-Matrix

und Gradient in dieser Darstellung zusammen $n(n+1)/2$ unbekannte Parameter, die die genaue Lage des Minimums festlegen. Daraus folgt, dass die quadratische Näherung nach der Berechnung von n^2 Funktionswerten vollständig definiert ist. Da ein Gradient die Informationen von $(n+1)$ Funktionswerten enthält³, sind nach n Iterationen eines Gradienten-Verfahrens genügend Informationen vorhanden, um sowohl die Hesse-Matrix als auch den Gradienten des Modellpotenzials und somit die Lage des Minimums exakt zu bestimmen. Ein Weg der Akkumulation sämtlicher gewonnener Daten ist das oben beschriebene *Conjugate-Gradient*-Verfahren.

Ein anderer Ansatz ist die sukzessive Aktualisierung einer genäherten bzw. effektiven Start-Hesse-Matrix \mathbf{H}_0 mittels geeigneter Formalismen, denen allesamt die so genannte **Quasi-Newton-Bedingung** zu Grunde liegt:

$$\mathbf{H}_{i+1}^{-1}(\mathbf{g}_{i+1} - \mathbf{g}_i) = (\mathbf{R}_{i+1} - \mathbf{R}_i) \quad (4.32)$$

Ziel dieser „variabel metrischen“ Verfahren ist die iterative Verbesserung der genäherten bis hin zur exakten Hesse-Matrix:

$$\lim_{i \rightarrow \infty} \mathbf{H}_i = \mathbf{H}_{\text{exakt}} \quad (4.33)$$

Einige der gebräuchlichsten Quasi-Newton-Methoden konnten von Oren und Luenberger [1974] als Linearkombination zusammengefasst und verallgemeinert werden:

$$\begin{aligned} \mathbf{H}_{i+1}^{-1} &= \gamma_i \mathbf{H}_i^{-1} + \left(1 + \gamma_i \Theta_i \frac{\mathbf{c}_i^T \mathbf{H}_i^{-1} \mathbf{c}_i}{\mathbf{x}_i^T \mathbf{c}_i}\right) \frac{\mathbf{x}_i \mathbf{x}_i^T}{\mathbf{x}_i^T \mathbf{c}_i} \\ &\quad - \gamma_i (1 - \Theta_i) \frac{\mathbf{H}_i^{-1} \mathbf{c}_i \mathbf{c}_i^T \mathbf{H}_i^{-1}}{\mathbf{c}_i^T \mathbf{H}_i^{-1} \mathbf{c}_i} - \gamma_i \Theta_i \frac{\mathbf{x}_i \mathbf{c}_i^T \mathbf{H}_i^{-1} + \mathbf{H}_i^{-1} \mathbf{c}_i \mathbf{x}_i^T}{\mathbf{x}_i^T \mathbf{c}_i} \end{aligned} \quad (4.34)$$

$$\text{mit } \mathbf{x}_i = \mathbf{R}_{i+1} - \mathbf{R}_i \quad \text{und} \quad \mathbf{c}_i = \mathbf{g}_{i+1} - \mathbf{g}_i$$

Setzt man $\gamma_i, \Theta_i = 1$, so gelangt man zum Ansatz von Broyden [1970], Fletcher [1970], Goldfarb [1970] und Shanno [1970] (BFGS-Verfahren):

$$\mathbf{H}_{i+1}^{-1} = \mathbf{H}_i^{-1} + \left(1 + \frac{\mathbf{c}_i^T \mathbf{H}_i^{-1} \mathbf{c}_i}{\mathbf{x}_i^T \mathbf{c}_i}\right) \frac{\mathbf{x}_i \mathbf{x}_i^T}{\mathbf{x}_i^T \mathbf{c}_i} - \frac{\mathbf{x}_i \mathbf{c}_i^T \mathbf{H}_i^{-1} + \mathbf{H}_i^{-1} \mathbf{c}_i \mathbf{x}_i^T}{\mathbf{x}_i^T \mathbf{c}_i} \quad (4.35)$$

Die Werte $\gamma_i = 1$ und $\Theta_i = 0$ führen zum DFP-Verfahren von Davidon, Fletcher und Powell, das die Hesse-Matrix auf folgende Weise aktualisiert (Davidon [1959]; Fletcher und Powell [1963]):

$$\mathbf{H}_{i+1}^{-1} = \mathbf{H}_i^{-1} + \frac{\mathbf{x}_i \mathbf{x}_i^T}{\mathbf{x}_i^T \mathbf{c}_i} - \frac{\mathbf{H}_i^{-1} \mathbf{c}_i \mathbf{c}_i^T \mathbf{H}_i^{-1}}{\mathbf{c}_i^T \mathbf{H}_i^{-1} \mathbf{c}_i} \quad (4.36)$$

³Bei der Ermittlung numerischer Gradienten über finite Differenzen entspricht dies gerade der Anzahl mindestens notwendiger Einzelpunktrechnungen.

Die Effizienz dieser rekursiven Methoden ist abhängig von der Qualität des *initial guess* für die Start-Matrix \mathbf{H}_0 . Häufig verwendet man Diagonal-Matrizen, allerdings können voroptimierte Hesse-Matrizen, die bereits chemische Informationen z.B. über Bindungsstärken und Kraftkonstanten enthalten, die Optimierung deutlich beschleunigen.

4.4 Quadratische Verfahren

Wenn neben dem Gradienten auch die analytische Hesse-Matrix zur Verfügung steht, gelangt man auf einer quadratischen Hyperfläche mit einem Schritt zum stationären Punkt, der bei positiv definiten Hesse-Matrix das gesuchte Minimum ist:

$$\mathbf{R}_{min} = \mathbf{R}_0 - \mathbf{H}_0^{-1} \mathbf{g}_0 \quad (4.37)$$

Abbildung 4.8 zeigt eine quadratisch modellierte Optimierung auf einer zweidimensionalen quadratischen Oberfläche. Die iterative Optimierung auf nicht-quadratischen

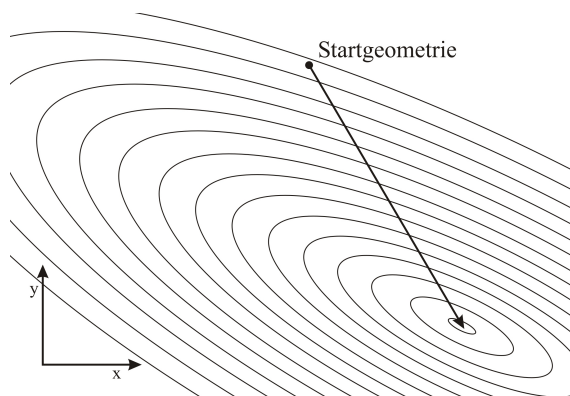


Abbildung 4.8: Newton-Raphson-Optimierung

Potenzialflächen unter Verwendung von Taylor-Reihen 2. Ordnung mit exakten analytischen Hesse-Matrizen nennt man **Newton-Raphson-Verfahren** oder auch **Second-Order Model**:

$$\mathbf{R}_{i+1} = \mathbf{R}_i - \mathbf{H}_i^{-1} \mathbf{g}_i \quad (4.38)$$

Zwar sehen einige quantenchemische Verfahren eine Berechnung analytischer zweiter Ableitungen vor, doch ist der Rechenaufwand im Allgemeinen mindestens 5-10 mal größer als der einer Gradienten-Bestimmung, so dass es sich in der Praxis als effektiver erwiesen hat, die analytische Hesse-Matrix als *initial guess* für aktualisierende Algorithmen (z.B. BFGS) anzusetzen, anstatt sie in jedem Fixpunkt vollständig neu zu berechnen.

In numerischen Verfahren ist die Berechnung von Hesse-Matrizen aus zwei Gründen unpraktikabel:

1. Die Berechnung ist zu aufwändig.
Für die Ermittlung einer (symmetrischen) Hesse-Matrix sind $n(n + 1)/2$ (n : Anzahl der Freiheitsgrade) Einzelpunkt-Rechnungen notwendig.⁴
2. Numerische Hesse-Matrizen sind zu ungenau.
Da erste Ableitungen über finite Differenzen angenähert werden müssen und zweite Ableitungen als finite Differenzen von ersten Ableitungen, potenzieren sich die Fehlerquellen.

4.5 Rational-Function-Optimierungen

Eine der meist genannten Einschränkungen der in den letzten Kapiteln beschriebenen Optimierungsverfahren ist die Nutzung des quadratischen Modellpotenzials (siehe Gleichung 4.20), die den Konvergenzradius auf den lokalen Bereich des Minimums einschränkt. Erweitert man dieses Modell auf rationale Funktionen, so gelangt man zum so genannten *Rational Function Optimization Model*. Darin ist die Energie E_{RFO} gegeben als:

$$E_{\text{RFO}} = E_i + \frac{\mathbf{g}_i^T \mathbf{x} + 1/2 \mathbf{x}^T \mathbf{H}_i \mathbf{x}}{1 + \mathbf{x}^T \mathbf{S} \mathbf{x}} = E_i + \frac{1}{2} \frac{\tilde{\mathbf{x}}^T \tilde{\mathbf{H}} \tilde{\mathbf{x}}}{\tilde{\mathbf{x}}^T \tilde{\mathbf{S}} \tilde{\mathbf{x}}} \quad (4.39)$$

Dabei wurden die folgenden Abkürzungen eingeführt:

$$\tilde{\mathbf{x}} = \begin{pmatrix} \mathbf{x} \\ 1 \end{pmatrix}, \quad \tilde{\mathbf{H}} = \begin{pmatrix} \mathbf{H}_i & \mathbf{g}_i \\ \mathbf{g}_i^T & 1 \end{pmatrix}, \quad \tilde{\mathbf{S}} = \begin{pmatrix} \mathbf{S} & 0 \\ 0 & 1 \end{pmatrix} \quad (4.40)$$

Darin quantifiziert die Diagonalmatrix \mathbf{S} die Anharmonizität des Potenzials. Funktionalableitung und Nullsetzen führt zu der Eigenwertgleichung

$$\tilde{\mathbf{H}} \tilde{\mathbf{x}} = \mu \tilde{\mathbf{S}} \tilde{\mathbf{x}} \quad \text{mit} \quad \mu = \frac{\tilde{\mathbf{x}}^T \tilde{\mathbf{H}} \tilde{\mathbf{x}}}{\tilde{\mathbf{x}}^T \tilde{\mathbf{S}} \tilde{\mathbf{x}}} \quad (4.41)$$

Daraus erhält man den Schrittvektor

$$\mathbf{x} = -(\mathbf{H}_i - \mu \mathbf{S})^{-1} \mathbf{g}_i \quad (4.42)$$

Der Anwendungsbereich einer *Rational-Function-Optimierung* beschränkt sich nicht nur auf quadratische Bedingungen, so dass auch stark deformierte Systeme ohne Vorooptimierung minimiert werden können.

⁴ Bei einem nicht-linearen Molekül mit zwölf Atomen wären demnach bspw. 465 (= (36-6)(36-6+1)/2) Einzelpunkt-Berechnungen nötig, um eine einzelne Hesse-Matrix zu generieren.

4.6 Konvergenzbeschleunigung

Die bisher beschriebenen Verfahren zogen ihren Nutzen aus den Funktionswerten, Gradienten oder Hesse-Matrizen einzelner Stützpunkte \mathbf{R}_i auf Grund der Annahme von quadratischen oder rationalen Modellpotenzialen. Die dabei iterativ erzeugte und gegen das Minimum \mathbf{R}_{min} konvergierende Folge von Geometrien $\{\mathbf{R}_i\}$ enthält jedoch zusätzliche Informationen über die Hyperfläche unabhängig von der Annäherung durch ein Modellpotenzial. Diese Idee führte zu einem alternativen Ansatz der Konvergenzbeschleunigung, dessen Varianten in diesem Abschnitt kurz erläutert werden sollen.

4.6.1 DIIS

Jedes iterative Optimierungs-Verfahren liefert eine Folge von Stützpunkten $\{\mathbf{R}_i\}$ und Schrittvektoren $\{\mathbf{x}_i\}$. Diese Folgen konvergieren mit

$$\lim_{i \rightarrow \infty} \mathbf{R}_i = \mathbf{R}_{min} \Leftrightarrow \lim_{i \rightarrow \infty} \mathbf{x}_i = 0 \quad (4.43)$$

Nach einem Vorschlag von Pulay [1982] werden die Folgeglieder \mathbf{R}_{i+1} und \mathbf{x}_{i+1} als eine Linearkombination der k vorherigen Glieder angesetzt:

$$\mathbf{R}_{i+1} = \sum_{j=i-k}^i c_j \mathbf{R}_j \quad (4.44)$$

$$\mathbf{x}_{i+1} = \sum_{j=i-k}^i c_j \mathbf{x}_j \quad (4.45)$$

Da die Folgen $\{\mathbf{R}_i\}$ und $\{\mathbf{x}_i\}$ bijektiv voneinander abhängen, konvergieren entweder beide oder keine. Es ist daher ausreichend, die Folge der Schrittvektoren zur Konvergenz zu bringen. Man fordert also, dass \mathbf{x}_{i+1} verschwindet und optimiert die Koeffizienten c_j unter der Maßgabe minimaler quadratischer Abweichungen und der Normierungsbedingung

$$\sum_{j=i-k}^i c_j = 1 \quad (4.46)$$

mittels Lagrangescher Multiplikatoren λ :

$$\mathcal{L} = \left(\sum_{j=i-k}^i c_j \mathbf{x}_j \right)^2 + 2\lambda \left(1 - \sum_{j=i-k}^i c_j \right) \quad (4.47)$$

Die Minimierung der Lagrange-Funktion \mathcal{L} liefert einen Satz Koeffizienten $\{c_j\}$, aus dem sich dann über Gleichung (4.44) die Geometrie \mathbf{R}_{i+1} berechnen lässt.

Die hier beschriebene Vorgehensweise, auch DIIS⁵ genannt, eignet sich besonders gut für „sparsame“ Algorithmen wie „*Function-Only*“, aber auch Gradienten-Verfahren werden meist deutlich beschleunigt.

4.6.2 GDIIS

Beim speziell für Geometrie-Optimierungen mit exakten oder approximierten Hesse-Matrizen entwickelten GDIIS (Császár und Pulay [1984]) interpoliert man in einem Zwischenschritt den imaginären Stützpunkt $\tilde{\mathbf{R}}$ und dessen Gradienten $\tilde{\mathbf{g}}$ als Linearkombination der letzten k Schritte:

$$\tilde{\mathbf{R}} = \sum_{j=i-k}^i c_j \mathbf{R}_j \quad (4.48)$$

$$\tilde{\mathbf{g}} = \sum_{j=i-k}^i c_j \mathbf{g}_j \quad (4.49)$$

Die Koeffizienten werden wie im DIIS-Algorithmus optimiert und die resultierenden interpolierten Werte in die Rekursion für Newton-Verfahren eingesetzt:

$$\mathbf{R}_{i+1} = \tilde{\mathbf{R}} - \mathbf{H}_i \tilde{\mathbf{g}} \quad (4.50)$$

Obwohl die GDIIS-Methode durch die Newton-Bedingung auf eine quadratische Näherung zurückgreift, ist sie nicht auf diese beschränkt. Vielmehr wurde in der Praxis eine effektive Beschleunigung der Konvergenz auch in Fällen größerer Entfernung vom Minimum beobachtet. Wichtige Voraussetzung für die Anwendbarkeit der Methode ist, wie beim DIIS-Verfahren, dass der zu Grunde liegende Optimierungsvorgang konvergiert.

4.7 Eingeschränkte Optimierungen

Selbst bei Ausnutzung komplexer formalistischer Grundlagen skaliert der Rechenaufwand von Optimierungen meist in einer Größenordnung von n^2 (n : Anzahl der Freiheitsgrade). Eine Reduktion der Freiheitsgrade sollte daher jedes Optimierungsverfahren beschleunigen.

4.7.1 Translation und Rotation

Der Vollständigkeit halber sei hier auf diese nahe liegende Einsparungsmöglichkeit hingewiesen. Die Eliminierung von Freiheitsgraden der Translation und Rotation ist

⁵≡ Direct Inversion in the Iterative Subspace

bei internen Koordinaten schon vollzogen, bei kartesischen Koordinaten bedarf sie der gesonderten Implementierung. Translationen verhindert man bei kartesischen Koordinaten trivialerweise durch das Festhalten eines beliebigen Kernes A; dieser Kern wird bei der Gradienten-Bestimmung ausgespart. Definiert man zusätzlich einen beliebigen Kern B, der nur in der Verbindungsachse zum fixierten Kern A deformiert werden darf und einen weiteren Kern C, der nicht aus der durch die drei Kerne aufgespannten Ebene herausbewegt werden darf, so hat man auch die Freiheitsgrade der Rotation eliminiert.

4.7.2 Symmetrie

Eine deutliche Reduktion der Freiheitsgrade gelingt durch die Zuordnung von Punktgruppen. Enthält ein System Kerne, die durch Symmetrie-Operationen aufeinander abbildbar sind, so müssen dieser „Kern-Typ“ nur einmal optimiert werden, vorausgesetzt, dass die Optimierung keine Symmetriebrechung erwarten lässt. Das ist insbesondere für angeregte Zustände nicht immer der Fall. Vorsicht ist außerdem beim Wechsel von höherer zu niedriger Symmetrie geboten, da ein Minimum in höherer Symmetrie einem Sattelpunkt in niedriger Symmetrie entsprechen kann.

4.7.3 Fixierung von Koordinaten

Schließt man einzelne oder mehrere Koordinaten von der Optimierung aus, dann spricht man von Randbedingungen/Einschränkungen („Constraints“). So ist es z.B möglich, ganze Struktur-Fragmente zu fixieren, die bereits als optimiert bekannt sind. Besonders vorteilhaft kann die Festlegung von Bindungslängen und/oder -winkeln sein, sie lässt sich sehr anschaulich bei Verwendung interner Koordinaten umsetzen.

Teil II

Implementierung

Kapitel 5

Ein Programm zur numerischen Optimierung

Um Verbindungen auf angeregten Hyperflächen zu optimieren, bedarf es in der Regel als Mindestinformation derjenigen über Gradienten. Analytische Gradienten existieren nicht für alle Energiebestimmungsverfahren, numerische Gradienten sind prinzipiell numerisch bestimmbar. Numerische Optimierungen sind daher immer dann sinnvoll, wenn die Bestimmung analytischer Gradienten entweder nicht möglich oder rechnerisch aufwändiger als der numerische Weg ist. Ein solcher Fall liegt z.B. bei dem in Abschnitt 2.3.12 (S. 36) beschriebenen, kombinierten Dichtefunktional-Multireferenz-CI von Grimme und Waletzke [1999b] vor — einem Paket, das mit hoher Geschwindigkeit und Genauigkeit die Energie von Grund- und angeregten Zuständen ermittelt.

Maßgaben bei der Auswahl des Optimierungsalgorithmus sind a) die Minimierung des rechnerischen Aufwands, b) ein gutes Konvergenzverhalten und c) eine möglichst geringe Fehlerfortpflanzung.

zu a) Newton-Raphson-Verfahren (s. S. 58) sind im Umfeld numerischer Optimierungen auf Grund ihrer hohen rechnerischen Kosten und ihrer Fehleranfälligkeit ausgeschlossen.

zu b) Wünschenswert ist ein Algorithmus mit quadratischem Modellpotenzial, was das Steepest-Descent-Verfahren (S. 53) ausschließt.

zu c) Eine konsequente Nutzung der möglicherweise unpräzisen numerischen Gradienten wie in Quasi-Newton-Verfahren (S. 56) lässt eine höhere Fehlerfortpflanzung erwarten.

Die Berücksichtigung aller hier genannten Argumente führt zum Conjugate-Gradient-Formalismus (S. 54).

Dieses Programm schließt auf pragmatische Weise die bestehende Lücke bei der Bestimmung von Gleichgewichts-Strukturen auf DFT/MRCI-Hyperflächen: Es liest

DFT/MRCI-Energien aus, berechnet aus ihnen numerische Gradienten und verwendet diese innerhalb eines Optimierers, der ohne Zweitableitungsdaten auskommt. So sind bei kleineren Verbindungen innerhalb von Minuten bis zu einigen Stunden entsprechende Ergebnisse zu erhalten. Um auf anderen Energiehyperflächen Verbindungen zu optimieren, muss lediglich die Einleseroutine entsprechend angepasst werden, was dem Code eine hohe Flexibilität verleiht.

5.1 Die Programmiersprache

Obwohl die Programmiersprache FORTRAN 77 bereits alle für ein solches Programm notwendigen Funktionen bereit stellt, fiel die Wahl auf das jüngere FORTRAN 90, weil es ein Reihe profitabler Erweiterungen bietet:

- **Format:** spaltenunabhängige Quellprogramm Schreibweise, Mehrfachanweisungen auf einer Zeile, Quellprogrammeinbettung mittels `INCLUDE`, Datenstrukturen, Felder mit übernommener Gestalt, selbst definierte Operatoren, Kurzschreibweise der Vergleichsoperatoren, selbst definierte Zuweisungsoperationen, maskierte Feldzuweisungen, Zeigerzuweisungen, Attribute in Typ-Vereinbarungsanweisungen, diverse Spezifikationsanweisungen, `CASE`-Struktur, nicht vorrückende E/A (eine Art Datenstrom-E/A), diverse E/A-Parameter, diverse Datenfeldbeschreibungen, Schlüsselwortparameter, private und sichtbare Größen, optionale Formalparameter, Schnittstellenblöcke, diverse Standard-subroutinen, viele neue Standardfunktionen und anderes mehr.
- **Feldverarbeitung.** FORTRAN 90 erlaubt Operationen mit Teil- und kompletten Feldern, während bei FORTRAN 77 Operationen nur mit einzelnen Feldelementen möglich sind. Entsprechende Funktionen für Matrix-Operationen wurden hinzugefügt.
- **Datentypen** können hinsichtlich ihrer internen Speicheranforderungen und ihrer internen Darstellung flexibler definiert werden, so dass der Programmierer in der Lage ist, seine Anwendung an die speziellen Gegebenheiten eines Rechners anzupassen und dafür zu optimieren.
- Mit dem Konzept der **Modularen Definitionen** hält der Entwickler ein sehr viel leistungsfähigeres Werkzeug der globalen Speicherverwaltung in Händen, als es ein `COMMON`-Block in FORTRAN 77 ist. Ein *Modul* ist eine Programmeinheit, die beliebige Typdeklarationen, Spezifikationen, Unterprogrammdefinitionen und/oder Schnittstellendefinitionen enthalten kann. Die Inhalte eines Moduls und damit der entsprechende Speicherbereich stehen den Subroutinen und Funktionen je nach Implementierung komplett oder teilweise zur Verfügung.

- Das **Typkonzept** in FORTRAN 90 sieht vor, dass über die aus FORTRAN 77 hinaus bekannten Datentypen hinaus eigene Datentypen definiert werden können. Danach ist ein Datentyp eine Menge von Werten, er hat einen Namen, eine Menge von Operationen auf diesen Werten und eine Vorschrift über die Schreibweise dieser Werte, die sich in der Schreibweise der Konstanten niederschlägt.
- **Dynamische Speicherverwaltung** ist der Oberbegriff für eine Reihe von FORTRAN 90 Kommandos, die die Belegung und Freigabe von, sowie den Zugriff auf dedizierte Speichersegmente regulieren. Mit ihnen ist es möglich, dynamische Felder während der Laufzeit zu de- und allozieren, Felder automatisch bei der Übernahme durch eine Subroutine allozieren zu lassen und Datentypen zu deklarieren, die auf ein Speichersegment verweisen.

5.2 Anbindung an bestehende Programmteile

Einer der Vorzüge des hier vorgestellten Programmes ist seine Flexibilität. Grundsätzlich kann jedes quantenchemische Verfahren als Grundlage zur Ermittlung einzelner Punkte auf einer Hyperfläche eingesetzt werden. Ebenso können die hier erhaltenen Geometrien in jeder beliebigen Rechenmethode weiterverwendet werden.

Das Einlesen und die Ausgabe der Geometrien erfolgt in Anlehnung an die TURBOMOLE-Suite über die Datei `coord`. Prinzipiell können in zukünftigen Versionen andere Formate implementiert werden.

Die Einbindung alternativer Rechenmethoden erfolgt sehr einfach in zweierlei Weise: Zum Einlesen muss der Name der entsprechenden ausführbaren Datei – gegebenenfalls inklusive Pfad – bekannt sein (z.B. `mrci` bzw. `/usr/bin/mrci`) und in Form eines so genannten *System Calls* in den bestehenden Code des Optimierers eingetragen werden. Bei aufsetzenden Verfahren wie z.B. DFT/MRCI sind dazu mehrere *System Calls* notwendig, die nacheinander die Optimierung der Kohn-Sham-Orbitale (`dscf`), eine RI-Transformation der Vierzentren-Integrale (`ritraf`), die Kalkulation der Ein-Elektron-Integrale (`oneint`) und schließlich die eigentliche MRCI-Rechnung (`mrci`) anstoßen. Geringfügig komplizierter gestaltet sich unter Umständen das Auslesen der berechneten Einzelpunkt-Energien, weil die Rechenmethode und der Optimierer nicht direkt aneinander angeschlossen sind, sondern nur über herausgeschriebene und einzulesende Dateien miteinander kommunizieren. Dieser Umstand macht einen Mechanismus zur Mustererkennung notwendig und stellt den Entwickler vor die Aufgabe, ein Muster zu finden, das den gesuchten Energiewert in einer zu definierenden Datei eindeutig lokalisiert. Dabei ist zu bedenken, dass bei Mehrdeterminanten-Verfahren auch das Auffinden der Energiedaten elektronisch angeregter Zustände eindeutig sein muss.

Im Falle der bereits eingebundenen DFT/MRCI-Routine geschieht das auf folgende Weise: Zunächst wird dem Optimierer die Nummer des zu optimierenden

elektronisch angeregten Zustandes per Eingabedatei mitgeteilt. Dann wird eine erste MRCI-Kalkulation gestartet. Schließlich liest der Optimierer aus der Standard-Ausgabedatei des MRCI, `mrci.sum`, die diesem Zustand zugeordnete dominante elektronische Anregung heraus und speichert sie ab. Das Programm wird diesen Zustand während der Optimierung nunmehr anhand dieser Anregung identifizieren, unabhängig von eventuell auftretenden Kreuzungen.

Es obliegt dem Anwender, die für das jeweils verwendete quantenchemische Verfahren notwendigen Vorbereitungen zu treffen. Das bedeutet z.B. im Fall des DFT/MRCI, dass vor dem Start des Optimierers die folgenden drei Schritte vorgenommen werden müssen:

1. Erzeugung bzw. Änderung der für den `dscf`-Lauf erforderlichen Eingabedatei `control` mit dem interaktiven TURBOMOLE-Skript `define`
2. Modifikation/Ergänzung der für die RI-Näherung unerlässlichen Einträge in der `control`-Datei mit dem interaktiven TURBOMOLE-Skript `rmp2prep`
3. Anlegen/Editieren einer DFT/MRCI-Input-Datei (hier: `mrci.in`)

Die im DFT/MRCI-Verfahren vorgesehene iterative Erzeugung des Referenzraumes wurde in das Programm integriert.

5.3 Programmablauf

Wie bereits in Abschnitt 4 (S. 45) erläutert wurde, löst man ein molekulares Optimierungsproblem durch iterative Wiederholung der folgenden Einzelschritte:

- Vermessung des Systems in Stützpunkt \mathbf{R}_i : Energie E_i und Gradient \mathbf{g}_i
- Prüfung auf Konvergenz
- Bestimmung des Schrittvektors \mathbf{x}_i , d.h. im Falle eines Pfadfinder-Algorithmus:
 - Berechnung der Suchrichtung \mathbf{d}_i
 - Ermittlung der Schrittlänge s_i

$$\implies \mathbf{x}_i = s_i \mathbf{d}_i$$

- Aktualisierung der Geometrie:

$$\mathbf{R}_{i+1} = \mathbf{R}_i + \mathbf{x}_i$$

Der Algorithmus wird beendet, wenn das System bestimmten Konvergenzkriterien genügt oder wenn die energetischen Änderungen von Zyklus zu Zyklus hinreichend gering sind.

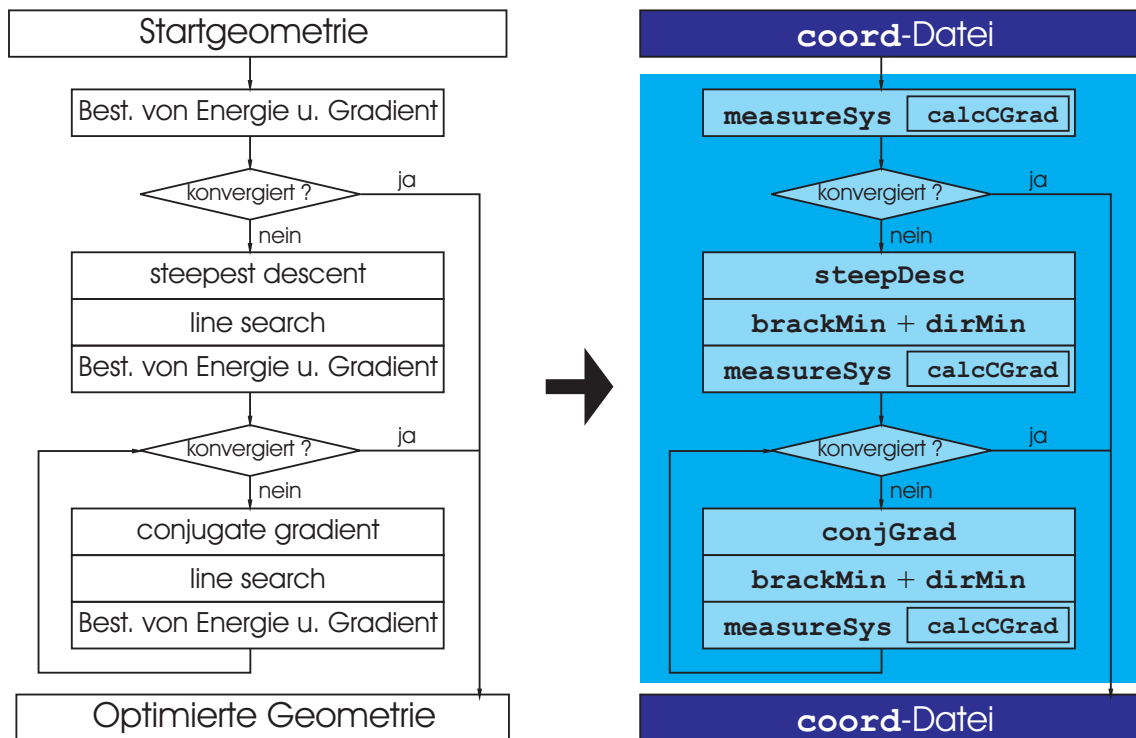


Abbildung 5.1: Programm-Ablauf und Implementierung. Die Begriffe `measureSys`, `calcCGrad`, `steepDesc`, `conjGrad`, `brackMin` und `dirMin` bezeichnen diejenigen Subroutinen, die die auf der linken Seite dargestellten Aufgaben übernehmen.

Die Implementierung des Optimierungskonzeptes wird in Abbildung 5.3 dargestellt. Der Hauptprogrammteil `main` prüft das Vorhandensein aller für den Ablauf notwendigen Informationen wie

- Welche Geometrie hat das zu vermessende System?
- Welche Rechenmethode soll benutzt werden?
- Sind alle notwendigen Dateien zugänglich und mit den korrekten Attributen versehen?
- Welche Parameter wurden per Input-Datei übergeben und wo werden stattdessen die Standardeinstellungen verwendet? (siehe auch Abschnitt 5.4, S. 74)

Danach übergibt `main` das Kommando an die Subroutine `optGeom`, das eigentliche Herzstück der Optimierung. `optGeom` verwaltet die Durchführung folgender Funktionen:

- Allokierung der für die Rechnungen benötigten Felder für Geometrien `cGeom`, Gradienten `cGrad` und konjugierte Richtungen `conjDir`
- Eliminierung der Translations-Freiheitsgrade durch Fixierung eines geeigneten Atoms

- Aufruf der Subroutine `measureSys` zur Berechnung der Systemenergie und Ermittlung des Gradienten mittels `calcCGrad`
- Ansteuerung der Optimierungs-Algorithmen *Steepest Descent* und *Conjugate Gradient*
- Berechnung von Energie und Gradient des aktualisierten Systems
- Prüfung auf Konvergenz und Start eines weiteren Zyklus oder gegebenenfalls Abbruch

Der „Zuständigkeitsbereich“ von `optGeom` ist in Abbildung 5.3 mit mittlerem Blau unterlegt.

Die Subroutine `measureSys` übernimmt während des gesamten Programmablaufs die Vermessung des Systems nach jedem vollständig abgeschlossenem Zyklus; sie bestimmt Energie und Gradient. Nach einer ersten Konvergenz-Abfrage werden Geometrie und Gradient an die Subroutine `steepDesc` weitergegeben. `steepDesc` startet den *Line Search* in Richtung des steilsten Abstiegs, d.h. in Gradienten-Richtung. Der *Line Search*, bestehend aus den Subroutinen `brackMin` und `dirMin`, ermittelt das Richtungsminimum. Daraus berechnet `steepDesc` die neue Geometrie und meldet sie zurück an `optGeom`. `optGeom` prüft erneut auf Konvergenz und führt je nach Schwellwert einen weiteren *Steepest Descent*-Schritt durch oder startet den *Conjugate Gradient*-Algorithmus. `conjGrad` bestimmt die Richtung für nachfolgende *Line Searches* nach dem in Abschnitt 4 (S. 45) dargestellten Formalismus und initiiert ihn. `optGeom` prüft daraufhin wiederum via `measureSys` auf Konvergenz und entscheidet über weitere Iterationen oder einen Abbruch wegen hinreichend optimierter Geometrie.

5.3.1 Das Konvergenzkriterium

Die Frage nach einem geeigneten Konvergenzkriterium ist die Frage nach dem entscheidenden Charakteristikum einer Gleichgewichts-Geometrie. Nach formalistischen Gesichtspunkten ist sie das Minimum eines Potenzialtopfes, d.h. sie ist gekennzeichnet durch den geringsten Energiegehalt und eine verschwindende Energie-Änderung δE bei infinitesimaler Deformation $\delta \mathbf{p}$ entlang des Deformationsvektors \mathbf{p} :

$$\frac{\delta E}{\delta \mathbf{p}} = 0 \quad (5.1)$$

In vielen Fällen wird die Optimierung als konvergiert betrachtet, wenn der letzte Schritt kaum eine Energie-Änderung herbeigeführt hat:

$$E(\mathbf{R}_{i+1}) \approx E(\mathbf{R}_i) \quad (5.2)$$

Die Gefahr dieses Kriteriums ist, dass eine Optimierung auch dann als fälschlicherweise abgeschlossen angesehen wird, wenn der Schrittvektor die Geometrie zufällig gerade auf die gegenüberliegende Seite des Potenzialtopfes deformiert hat, wie in

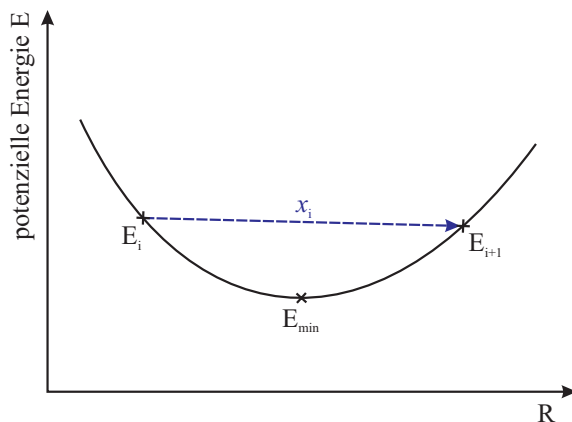


Abbildung 5.2: *Worst case scenario* einer Optimierung mit dem Abbruchkriterium an nähernder Energie-Gleichheit $E_{i+1} \approx E_i$ nach dem i -ten Schritt mit Schrittvektor x_i : Die Optimierung wird als konvergiert abgebrochen, obwohl das Minimum noch nicht erreicht wurde.

Abbildung 5.2 dargestellt. Die Wahrscheinlichkeit für das Eintreffen eines solchen „*worst case scenarios*“ sinkt mit der Größe des Moleküls, da es für jede Dimension mehr oder weniger gut erfüllt sein muss.

Dennoch ist es zuverlässiger und nach Gleichung (5.1) naheliegender, das Konvergenzkriterium mit dem Gradienten zu verknüpfen. Da der Gradient ein Vektor ist, berechnet man seine euklidische Norm

$$|\mathbf{g}| = \sqrt{\sum_i g_i^2} \quad (5.3)$$

und optimiert, bis diese einen bestimmten Schwellwert unterschreitet. Da jedes Atom annähernd additiv zur Gradientennorm beiträgt, ist diese Forderung an große Moleküle ungleich härter und benötigt deutlich mehr Schritte als das bei kleineren Systemen der Fall ist. Das Kriterium skaliert falsch, und man erhält für größere Moleküle verhältnismäßig bessere Geometrien als für kleinere. Aus diesem Grunde werden statt dessen die Normen aller atomarer Gradienten

$$|\mathbf{g}_{atom}| = \sqrt{\sum_{k=1}^3 g_{atom,k}^2} \quad \text{mit } k = \text{Raumkoordinaten-Index} \quad (5.4)$$

ermittelt. Die größte dieser Normen muss einen bestimmten Schwellwert unterschreiten, damit die Optimierung als konvergiert gilt.

5.3.2 Die Berechnung der Gradienten

Die Problematik numerischer Gradienten wurde in Abschnitt 4.3.1 (S. 50) ausführlich beschrieben.

Da die zur Ermittlung des Gradienten notwendige inkrementelle Auslenkung entlang der kartesischen Einheitsvektoren in aller Regel einen Bruch der Symmetrie verursacht, ist der Optimierer grundsätzlich nicht in der Lage, entsprechende Punktgruppen beizubehalten. Zum aktuellen Zeitpunkt wurde jedoch zumindest eine Eliminierung der Translations-Freiheitsgrade integriert. Die Implementierung von internen Koordinaten und Symmetrie-Eigenschaften sind Gegenstand zukünftiger Weiterentwicklungen, die den numerischen Aufwand erheblich senken können.

Es werden zunächst grundsätzlich Zwei-Punkt-Gradienten generiert. Sobald die maximale atomare Gradientennorm einen bestimmten Grenzwert unterschreitet, schaltet das Programm dynamisch auf die Erzeugung exakterer 3-Punkt-Gradienten um. Diese verursachen zwar eine Verdopplung des rechnerischen Aufwandes, doch sind sie für eine zuverlässige Lokalisierung des Minimums unverzichtbar (s. Abb. 4.5, S. 52). Der Anwender hat über die Eingabedatei die Möglichkeit, die Ermittlung der Gradienten während der gesamten Optimierung auf den 2-Punkt- oder den 3-Punkt-Algorithmus fest zu legen.

Ein sensibler Punkt der Gradienten-Berechnung ist das Inkrement h (s. Abschnitt 4.3.1, S. 52). Es wurde daher die in der Publikation von Baker [1987] referenzierte Vorschrift nach Gill u. a. [1981] implementiert:

$$h = 20 \sqrt{\frac{\Delta_{err}}{1 + |E(\mathbf{R})|}} \quad (5.5)$$

Darin wird das Inkrement h in Abhängigkeit von der Fehlertoleranz Δ_{err} der jeweils benutzten quantenchemischen Methode und dem Absolutwert der Energie berechnet.

Da die Ermittlung eines numerischen Gradienten im Grunde einer systematischen Berechnung von $(3N - 3)$ unabhängigen Einzelpunkten auf der Energie-Hyperfläche ist, bietet sich hier die Möglichkeit einer zukünftigen Parallelisierung.

5.3.3 Bestimmung der Suchrichtung

Von den drei Möglichkeiten, Steigungsinformationen für die Bestimmung einer Suchrichtung im Sinne eines Pfadfinder-Algorithmus zu nutzen (*Steepest Descent*, *Conjugate Gradient* und Quasi-Newton), war die Wahl aus verschiedenen Gründen (S. 65) auf ein *Conjugate-Gradient*-Verfahren gefallen. Die *Conjugate-Gradient*-Rekursion (4.26 auf S. 55) erfordert jedoch zwei Gradienten, die typischerweise durch einen initialen *Steepest-Descent*-Schritt erzeugt werden. Die erste Suchrichtung \mathbf{d}_1 entspricht also immer der Gegenrichtung des ersten Gradienten, $-\mathbf{g}_1$:

$$\mathbf{R}_2 = \mathbf{R}_1 - s_1 \mathbf{g}_1 \quad (5.6)$$

Die Schrittlänge s_1 wird beim ersten *Line Search* ermittelt (s.u.). Sobald zwei Gradienten zur Verfügung stehen, hat man die Wahl zwischen weiteren *Steepest-Descent*-

Iterationen und dem Einstieg in den *Conjugate-Gradient*-Formalismus. Der Anwender hat drei Möglichkeiten, auf diese Entscheidung Einfluss zu nehmen:

- Jede Iteration folgt bis zur Konvergenz der Richtung des steilsten Abstiegs (*complete Steepest Descent*).
- Nach dem obligatorischen ersten *Steepest-Descent*-Schritt sollen alle weiteren Schritte in konjugierte Richtungen erfolgen (*complete Conjugate Gradient*).
- Unterschreitet die maximale atomare Gradientennorm einen Grenzwert, so wird dynamisch von *Steepest Descent* auf *Conjugate Gradient* umgeschaltet.

Da der *Steepest-Descent*-Algorithmus die Energie eines Systems unabhängig von quadratischen Modellpotenzialen minimiert, empfiehlt er sich besonders in Fällen großer Entfernung vom Minimum der Hyperfläche. Befindet sich eine Geometrie allerdings hinreichend nah der Gleichgewichts-Geometrie, sollte man das effizientere Konvergenzverhalten einer *Conjugate-Gradient*-Optimierung nutzen. Diese Fakten sprechen für ein dynamisches Umschalten der Methoden. Der Schwellwert des Umschaltens wurde anhand einiger Fallbeispiele auf 0.05 festgelegt (dieser Wert kann durch den Parameter `sd2cd` in der Eingabedatei umdefiniert werden, s.u.).

Wie in Abschnitt 4.3.3 (S. 54) bereits erläutert wurde, kann der für die *Conjugate-Gradient*-Rekursion bedeutsame Parameter γ_i auf zweierlei Weise ermittelt werden. Dabei folgt die Bestimmungsgleichung von Fletcher und Reeves (4.29 auf S. 55) dem strikten, für „ideale“ mehrdimensional parabolische Szenarien gültigen Formalismus, während mit der Formel von Polak und Ribière (4.30, S. 56) gegebenenfalls auftretende numerische Unzulänglichkeiten abgefangen und korrigiert werden. Beide Alternativen wurden implementiert, jedoch hat sich bei den durchgeführten Optimierungen die Berechnung von γ_i nach Polak und Ribière als effizienter herausgestellt.

5.3.4 Die Schrittlänge

Die Bestimmung der Schrittlänge erfolgt iterativ durch eine Reihe inverser parabolischer Interpolationen gemäß Gleichung (4.11) bis zur Konvergenz. Voraussetzung dafür sind drei in Suchrichtung ermittelte Punkte P, Q und R , wie in Abschnitt 4.1 (S. 46) beschrieben. Einzige Forderung an diese Punkte ist, dass sie ein Minimum einrahmen, andernfalls wäre das Ergebnis der parabolischen Interpolation eine quadratische Funktion mit negativer Krümmung und würde folglich ein Maximum ermitteln. Zur Detektion geeigneter Punkte auf der Hyperfläche in der Suchrichtung d_i geht das Programm folgendermaßen vor:

1. Zunächst wird in Abhängigkeit von der Gradientennorm `gradNorm` die Strecke `cPut` ermittelt gemäß

$$\text{cPut} = 0.15/\sqrt{\text{gradNorm}} \quad (5.7)$$

Dieser analytische Ausdruck und der Parameter wurden iterativ mit den Fit-Routinen von `gnuplot` (Williams und Kelley [2000]) bestimmt, wobei eine Reihe von Messdaten das Trainingsset bildeten. Günstigstenfalls entspricht `cPut` der gesuchten Schrittlänge s_i und führt direkt in das Minimum der Suchrichtung \mathbf{d}_i . Ungünstigstenfalls deformiert dieser Wert das Molekül zu stark, so dass der berechnete Wert für `cPut` durch seinen maximal zulässigen Wert¹ ersetzt wird.

2. Im Abstand `cPut` wird ein erster Energiewert bestimmt und mit dem Energiewert der Einstiegsgeometrie verglichen.
3. In Abhängigkeit, ob Elongation um `cPut` entlang der Suchrichtung die Energie erhöht oder verringert hat, werden weitere Punkte auf dem Suchvektor berechnet, und zwar so lange, bis ein Minimum eingerahmt wird.

Das ermittelte Punkte-Tripel dient nun als Grundlage für mehrere inverse parabolische Interpolationen, bis der energetische Unterschied der einrahmenden Punkte zum Minimum eine Energie-Differenz in der 5. Nachkommastelle unterschreitet. Laut Schlick [1992] beschleunigt ein präziser *Line Search* deutlich die Konvergenz eines *Conjugate Gradient*-Verfahrens.

5.4 Eingabe-Optionen

Die Eingabedatei („Input“) gibt dem Anwender die Möglichkeit, den Optimierungsvorgang zu beeinflussen. Alle Optionen werden Schlüsselwort getrieben angegeben, dabei muss jedes Schlüsselwort mit dem Zeichen `$` beginnen, um vom Programm als solches erkannt zu werden. Alle Zeilen, die nicht mit `$` beginnen, werden ignoriert und können zur Kommentierung genutzt werden. In seiner aktuellen Version bietet das Programm die folgenden Einstellmöglichkeiten:

`$method <string>`

Mit diesem Schlüssel wird die zu benutzende Einzelpunkt-Bestimmungsmethode gewählt. Bisherige Einstellmöglichkeiten sind `dscf` und `mrci`, wobei `dscf` die Standardeinstellung ist.

`$state <int>`

Hier legt der Benutzer fest, welcher elektronische Zustand optimiert werden soll. Die Einstellung ist nur relevant, wenn als Methode `mrci` gewählt wurde, bei `dscf`-basierten Optimierungen wird sie ignoriert. Das Programm wird nach einer ersten MRCI-Referenzraum-Erzeugung den `<int>`-ten Zustand auslesen, dessen dominante Anregung abspeichern und den Zustand in allen Folgeschritten anhand dieser Anregung identifizieren. Man bedenke, dass für die Optimierung des S_0 -Grundzustandes gemäß der internen MRCI-Nummerierung eine 1 angegeben werden muss, was auch die Standardeinstellung ist.

¹erlaubt die maximale Verschiebung eines Atoms um 0.01 bohr

`$gradpoints` <int>

An dieser Stelle kann der Anwender die Erzeugung der Gradienten als 2-Punkt- oder 3-Punkt-Verfahren vorschreiben. Standardeinstellung ist ein dynamisches Umschalten zwischen 2-Punkt- und 3-Punkt-Gradient nach Unterschreiten eines gewissen Schwellwertes für die maximale atomare Gradientennorm, wie in Abschnitt 5.3.1 beschrieben.

`$maxcycle` <int>

In manchen Fällen kann es sinnvoll sein, die Anzahl maximal erlaubter Zyklen zu begrenzen oder zu erweitern. Sollte diese Anzahl erreicht werden, bevor das Konvergenzkriterium erfüllt ist, wird eine Datei namens `NOT.CONVERGED` erzeugt. Ohne diese Einflussnahme wird die Optimierung nach 100 Zyklen beendet.

`$convnorm` <real>

Das Konvergenzkriterium ist standardmäßig auf eine maximale atomare Gradientennorm von 0.0005 festgelegt. Unter Umständen sind gerade auch bei rechenintensiven großen Systemen Geometrien weniger hoher Qualität ausreichend, so dass man diesen Wert steigern möchte.

`$inkr` <real>

An dieser Stelle definiert der Anwender die Auslenkung h der numerischen Gradienten-Bestimmung. Diese Werte hängen stark vom Fehler Δ_{err} der jeweils verwendeten Methode ab und werden daher vom Programm komfortabel und zuverlässig nach Vorschrift (5.5), die bei Gill u. a. [1981] zu finden ist, bestimmt. Eine Änderung dieser Einstellung empfiehlt sich nur in besonderen Fällen. Typische Werte liegen zwischen 0.001 und 0.0005.

5.5 Ausgabe

Während der Optimierung informiert das Programm in der Standardausgabe umfassend über den aktuellen Status und berechnete Werte. Dazu zählen die folgenden Daten:

- Parameter der Optimierung, z.B. verwendete quantenchemische Methode, zu optimierender elektronischer Zustand und weitere aus der Eingabedatei übernommene Einstellungen
- aktuelle Stützpunkt-Geometrie
- Status von Gradienten-Berechnungen und *Line Searches*
- Gradienten und konjugierte Suchrichtungen
- ermittelte Schrittlängen

- Informationen über dynamisches Umschalten von *Steepest Descent* zu *Conjugate Gradient*, bzw. von 2-Punkt zu 3-Punkt-Gradienten, etc.
- Angabe der aktuellen molekularen und maximalen atomaren Gradientennorm
- statistische Zusammenfassung der Anzahlen benötigter Steepest-Descent- und Conjugate-Gradient-Schritte sowie Gesamtzahl durchgeführter Einzelpunkt-Rechnungen

5.6 Geplante Erweiterungen

Das hier vorgestellte Programm ist in mancherlei Hinsicht erweiterbar. Wünschenswert scheint jede Maßnahme, die den numerischen Rechenaufwand verringern kann. Dazu zählen vor allem die in Abschnitt 4.7 (S. 61) dargestellten Methoden zur Reduktion der Freiheitsgrade, allen voran die Einführung interner Koordinaten und Symmetrie. In Vorbereitung sind die Parallelisierung der Gradienten-Berechnung inklusive Lastverteilung nach einem Server-Client-Schema und die Möglichkeit eines Neustarts nach Abbruch („*Restart*“).

Teil III
Anwendung

Kapitel 6

Formaldehyd

Als einfachste Carbonylverbindung war und ist Formaldehyd Gegenstand zahlreicher Untersuchungen der Physikalischen und Theoretischen Chemie. Unter Letzteren zu nennen sind vor allem die Pionierarbeiten von Whitten und Hackmeyer [1969] sowie Buenker und Peyerimhoff [1970], aber auch aktuellere Untersuchungen mit Multi-Referenz-CI-Verfahren (Harding und Goddard III [1977]; Hachey u. a. [1995]; Hachey und Grein [1995]; Grein und Hachey [1996]; Perić u. a. [2000]), mit *Complete-active-space* in Kombination mit Störungstheorie zweiter Ordnung (CASPT2, Merchán und Roos [1995]) und mit *Equation-of-motion*-Algorithmen (EOM, Yeager und McKoy [1974]; Gwaltney und Bartlett [1999]). Da Formaldehyd eine der wenigen Spezies ist, für die auch Geometrien elektronisch angeregter Zustände experimentell bestimmt wurden (Raynes [1966]; Taylor u. a. [1982]; Takagi und Oka [1963]; Job u. a. [1969]; Chutjian [1974]; Hodges u. a. [1958]; Birss u. a. [1973]; Jones und Coon [1969] und Referenzen darin), eignet es sich als Prüfstein für ein Programm zur Geometrie-Optimierung auf Potenzial-Hyperflächen dieser Zustände. In diesem Kontext relevante Übersichtsartikel stammen z.B. von Davidson und McMurchie [1982] sowie Clouthier und Ramsay [1983].

6.1 Validierung und Leistungsvergleich

Um die korrekte und effektive Funktionsweise des hier implementierten Optimierers sicher zu stellen, wurde er zunächst mit dem in der TURBOMOLE-Suite implementierten Optimierer verglichen. Da die dabei verwendete quantenchemische Methode ein SCF-Algorithmus ist, Elektronen-Korrelation also unberücksichtigt bleibt, beschränkt sich die Optimierung auf die Grundzustands-Geometrie S_0 , und es wird in diesem ersten Schritt auf einen Vergleich zu experimentellen Werten verzichtet.

6.1.1 Technische Details

Zum Einsatz kam das TURBOMOLE-Modul `dscf` (direktes SCF: Integrale werden nicht auf der Festplatte gespeichert, sondern in jeder Iteration nach Bedarf neu erzeugt), die Energie-Konvergenz der SCF-Prozedur wurde auf `scfconv=10-8` Ha festgesetzt. Die folgenden Optimierungs-Varianten wurden getestet: *Steepest Descent* und Hesse-Matrix-Aktualisierungen zum einen nach dem BFGS-Algorithmus und zum anderen nach dem Ahlrichs-Schema, das die Standard-Einstellung im TURBOMOLE-Paket ist (Karlsruher Quantenchemie-Gruppe [1998]). Dem Ahlrichs-Schema liegen mehrere der in Kapitel 4 besprochenen Algorithmen zu Grunde. Man startet mit einer linearen Zwei-Punkt-Extrapolation der Geometriedaten und schaltet dann auf das Pulaysche GDIIS-Verfahren zur Geometrie-Vorhersage um, kombiniert mit einer Aktualisierung der Diagonalelemente der Hesse-Matrix. In kritischen Fällen, d.h. wenn nach einer vorgegebenen Anzahl von Iterationen die Gradienten-Norm immer noch zwischen 0.01 und 0.001 liegt, wird die Matrix nach dem vollen BFGS-Algorithmus aktualisiert.

Der Optimierer dieser Arbeit wurde in seiner Standard-Einstellung, d.h. dynamischer Umschaltung von *Steepest Descent* zu *Conjugate Gradient* bei einer maximalen atomaren Gradienten-Norm von 0.01, und als vollständige *Steepest-Descent*-Methode durchgeführt, dabei wurden die numerischen Gradienten als Drei-Punkt-Gradienten (siehe Abschnitt 4.3.1, S. 51) ermittelt.

Die Testläufe wurden mit je drei Start-Geometrien durchgeführt, die in Tabelle 6.1 aufgelistet sind. Geometrie 1 wurde als „plausible“ Konformation mit durch-

Koordinate	Geometrie 1	Geometrie 2	Geometrie 3	Referenz
R_{CO}	1.220	2.0109	1.1863	1.1863
R_{CH_1}	1.089	0.5426	1.0961	1.0961
R_{CH_2}	1.089	2.1706	1.0961	1.0961
\angle_{HCH}	120.0	117.97	121.94	116.14
\angle_{OCH}	120.0	121.02	90.0	121.93
\angle_{oop}	0.0	0.0	90.0	0.0

Tabelle 6.1: Start- und Referenz-Geometrien der Optimierungs-Tests in Å und Grad

schnittlichen Bindungslängen und Bindungswinkeln von 120° angesetzt, in Geometrie 2 wurden die Abstände der als Gleichgewichts-Geometrie berechneten Bindungen verändert und die pyramidale Geometrie 3 entspricht einer Winkel-Deformation der Gleichgewichts-Struktur. Aus Gründen der Vergleichbarkeit wurden alle Geometrien in C_1 -Symmetrie berechnet. Die Referenz-Geometrie wurde durch eine TURBOMOLE-Optimierung nach einem Quasi-Newton-Verfahren mit Ahlrichs-Schema bei strengeren Konvergenz-Kriterien (Energie: 10^{-4} mHa, maximale Norm des kartesischen molekularen Gradienten: 10^{-4}) ermittelt. Die Tabellen 6.2 bis 6.4 fassen die Ergebnisse zusammen.

Tabelle 6.2: Vergleich verschiedener Optimierungs-Algorithmen mit dem Optimierer dieser Arbeit für die Geometrie 1

	TM	TM	TM	<i>diese Arbeit</i>	
	SD	BFGS	Ahlichrs	SD	CG
R_{CO}	1.1863	1.1865	1.1860	1.1863	1.1864
R_{CH}	1.0968	1.0956	1.0962	1.0959	1.0963
\angle_{HCH}	116.38	116.22	116.10	116.39	116.18
\angle_{OCH}	121.79	121.89	121.95	121.80	121.91
\angle_{oop}	0.0	0.0	0.0	0.0	0.0
Anzahl der Schritte	13	7	6	11	10
Norm des atomaren Gradienten ($\cdot 10^{-4}$)	7.26	6.45	6.07	4.83	1.70

TM: TURBOMOLE

SD: *Steepest Descent*CG: *Conjugate Gradient*

BFGS: Quasi-Newton-Verfahren nach Broyden-Fletcher-Goldfarb-Shanno

Ahlichrs: Verknüpfung verschiedener Algorithmen, Beschreibung siehe Text

Tabelle 6.3: Vergleich verschiedener Optimierungs-Algorithmen mit dem Optimierer dieser Arbeit für die Geometrie 2 (deformierte Bindungslängen)

	TM	TM	TM	<i>diese Arbeit</i>	
	SD	BFGS	Ahlichrs	SD	CG
R_{CO}	1.1860	1.1864	1.1863	1.1861	1.1861
R_{CH}	1.0961	1.0968	1.0961	1.0962	1.0963
\angle_{HCH}	116.06	116.07	116.11	115.94	115.89
\angle_{OCH}	121.97	121.97	121.94	122.04	122.05
\angle_{oop}	0.0	0.0	0.0	0.0	0.0
Anzahl der Schritte	24	18	18	21	19
Norm des atomaren Gradienten ($\cdot 10^{-4}$)	7.73	6.67	0.62	3.57	3.81

Abkürzungen und Einheiten wie oben in Tab. 6.2

6.1.2 Diskussion

Die TURBOMOLE-Optimierungs-Module arbeiten mit analytischen Gradienten und lassen daher eine deutlich schnellere Konvergenz der Geometrien erwarten, als das bei einem numerischen Optimierer dieser Arbeit der Fall sein kann (siehe Abschnitt

Tabelle 6.4: Vergleich verschiedener Optimierungs-Algorithmen mit dem Optimierer dieser Arbeit für Geometrie 3 (deformierte Bindungswinkel)

	TM	TM	TM	<i>diese Arbeit</i>	
	SD	BFGS	Ahlich	SD	CG
R_{CO}	1.1862	1.1864	1.1862	1.1865	1.1865
R_{CH}	1.0962	1.0959	1.0958	1.0961	1.0959
\angle_{HCH}	116.12	116.12	116.14	116.13	116.12
\angle_{OCH}	121.93	121.92	121.92	121.93	121.94
\angle_{oop}	≈ 0.0	≈ 0.0	≈ 0.0	≈ 0.0	≈ 0.0
Anzahl der Schritte	53	16	15	34	30
Norm des atomaren Gradienten ($\cdot 10^{-4}$)	2.30	2.77	3.59	3.64	4.18

Abkürzungen und Einheiten wie oben in Tab. 6.2

4.3, S. 49). Unglücklicherweise verwenden beide Verfahren unterschiedliche Konvergenz-Kriterien, was die Vergleichbarkeit einschränkt. TURBOMOLE bricht den Vorgang ab, wenn die Norm des kartesischen molekularen Gradienten einen Grenzwert von 10^{-3} unterschreitet und/oder die Änderung der Energie im letzten Schritt unterhalb 10^{-3} mHa lag (TURBOMOLE Manual, Karlsruher Quantenchemie-Gruppe [1998]). In der vorliegenden Arbeit wurde statt dessen die maximale Norm eines *atomaren* Gradienten, also ein Maß für die auf ein einzelnes Atom wirkende Kraft, als Abbruch-Kriterium heran gezogen und auf $5 \cdot 10^{-4}$ festgesetzt (siehe Abschnitt 5.3.1, S. 70)). Um also die Qualität der optimierten Geometrien vergleichen zu können, wurde jeweils diese atomare Gradienten-Norm auch für TURBOMOLE-optimierte Geometrien in den Ergebnis-Tabellen 6.2 bis 6.4 mit aufgeführt. Als Vorteil des numerischen Verfahrens könnte es sich erweisen, dass in jeder Iteration im Sinne eines Pfadfinder-Algorithmus ein *Line Search* durchgeführt wird, so dass jede Iteration sicher ins Richtungsminimum führt, während dies bei Algorithmen mit festem Schrittvektor, wie sie im TURBOMOLE-Optimierer verwendet werden, nicht immer gewährleistet ist. Der bei relativ hoher Präzision (Konvergenz der Energie in der 6. Nachkommastelle) in der Regel zwischen 5 und 7 Einzelpunkt-Berechnungen erfordert. Die TURBOMOLE-Optimierungen verwenden selbst im *Steepest-Descent*-Modus keinen *Line Search*, was aus einem Zeitvergleich zweier Rechnungen (einmal ohne Line Search, einmal Steepest Descent) hervor geht.

Start-Geometrie 1 liegt bereits in der Nähe des Minimums einer DFT/DSCF-Potenzialmulde. Während der Optimierung müssen die CO-Bindung R_{CO} um 3.4 pm verkürzt, die CH-Bindungen R_{CH} um 0.7 pm gestreckt und der Spreizwinkel \angle_{HCH} um etwa 4° verringert werden. Damit kann angenommen werden, dass wir uns im lo-

kalen Bereich der Gleichgewichts-Geometrie befinden, wo Quasi-Newton-Verfahren besonders effektiv arbeiten (siehe Abschnitt 4.3.4, S. 56). Tatsächlich gelangen die Turbomole-Algorithmen mit Hesse-Aktualisierung nach 6 bzw. 7 Schritten ins Minimum. Das TURBOMOLE-*Steepest-Descent*-Verfahren benötigt dagegen etwa doppelt so viele Iterationen. Auch der numerische Optimierer führt verlässlich und schnell zum Minimum: *Steepest-Descent* und *Conjugate-Gradient* ermitteln nach 11 bzw. 10 Schritten Geometrien höherer Qualität, wie aus den Abweichungen der Bindungslängen und den atomaren Gradienten-Normen zu sehen ist. Die schnelle Konvergenz des *Steepest-Descent*-Algorithmus ist darauf zurück zu führen, dass die Kopplungen der Variablen relativ gering sind (siehe Abschnitt 4.3.2, S. 53): Im Wesentlichen wechseln die Optimierungs-Schritte zwischen einer Verkürzung der CO-Bindung und einer Nachrelaxation der CH-Bindungen und des HCH-Winkels, so dass eine neuerliche Verkürzung der CO-Bindung möglich wird, und so fort. Der Vergleich der *Steepest-Descent*-Optimierungen miteinander zeigt, dass der im Allgemeinen ungenauere numerische Gradient kombiniert mit einem Pfadfinder-Algorithmus ein besseres Konvergenzverhalten aufweist.

Start-Geometrie 2 wurde so angesetzt, dass im Wesentlichen die Bindungslängen optimiert werden müssen. Da Streck-schwingungen in der Regel höher frequent schwingen als Winkel-Deformations-Moden, haben sie höhere Kraftkonstanten, gleichbedeutend mit hohen Eigenwerten der Hesse-Matrix und einer starken Krümmung der Potenzialfläche. Eine Streckung oder Stauchung von Bindungen lässt demnach große Gradienten erwarten, wodurch sich Fehler bei numerischen Verfahren verringern. Des Weiteren wurden die Deformationen in Bindungsrichtung jeweils so drastisch gewählt, dass sich der Startpunkt nicht mehr im lokalen Bereich des Potenzial-Minimums befinden dürfte. Insgesamt begünstigen die Start-Bedingungen also den numerischen Optimierer und schwächen die Quasi-Newton-Algorithmen. Tatsächlich können die numerischen Varianten zu den Aktualisierungsmethoden aufschließen; das *Conjugate-Gradient*-Verfahren benötigt lediglich einen Schritt mehr als die „quadratischen“ TURBOMOLE-Routinen, um eine Geometrie sehr guter Qualität zu finden. Es fällt auf, dass die vom quadratischen Modellpotenzial weitgehend unabhängigen *Steepest-Descent*-Methoden ganz besonders von der drastischen Deformation bei verhältnismäßig geringer Variablen-Kopplung profitieren.

Start-Geometrie 3 wurde komplementär zu Geometrie 2 konzipiert. Der Optimierer hatte die Aufgabe, eine pyramidale Verzerrung aufzuheben und das Molekül in seinen planaren Gleichgewichtszustand zurück zu führen. Winkel-Deformationen sind üblicherweise „weichere“ Moden mit niedrigeren Kraftkonstanten als Valenz-Moden. Die entsprechenden Krümmungen auf der Potenzialfläche sind relativ flach und das Risiko einer numerischen Fehlkalkulation hoch. Darüber hinaus sind die Voraussetzungen für *Steepest-Descent*-Algorithmen ungünstiger, da jede Änderung

der Struktur einen Einfluss auf die nun näher liegenden Nachbar-Atome hat, so dass mehr Variablen koppeln. Das zeigt sich überdeutlich in den 53 Iterationsschritten des TURBOMOLE-SD-Testlaufes. Die Quasi-Newton-Methoden der TURBOMOLE-Suite zeigen sich davon vergleichsweise unbeeinträchtigt und minimieren die Struktur in 15 bzw. 16 Schritten. Für die gleichermaßen langsame Konvergenz beider numerischer Verfahren scheint tatsächlich die schlechte Qualität der Gradienten verantwortlich zu sein. Dabei wird ein ähnlich extremer Einbruch der Effizienz beim Steepest-Descent-Verfahren durch die Zuverlässigkeit des *Line Searches* abgefangen, während eine effektive Akkumulation lokaler Daten beim *Conjugate Gradient*-Algorithmus offenbar durch unpräzise Gradienten eingeschränkt wird.

Es sei abschließend darauf hingewiesen, dass die hier kommentierten Ergebnisse aus Zeitgründen nur durch einzelne Testläufe ermittelt wurden und nicht etwa — wie es für eine vollständige statistische Abhandlung notwendig gewesen wäre — durch systematisch angesetzte Testreihen. Dennoch sind die Tendenzen eindeutig.

Da es nicht Sinn und Zweck eines auf numerischen Gradienten-Verfahrens sein kann, mit analytisch basierten Methoden zu konkurrieren, wurde auf eine Laufzeit-Analyse verzichtet. Es soll jedoch nicht verschwiegen werden, dass ein Iterationszyklus des numerischen Optimierers im Durchschnitt etwa 20 Mal so lange dauerte wie ein analytischer Zyklus.

6.2 Optimierungen angeregter Zustände

Das hier vorgestellte numerische Optimierungs-Modul ist grundsätzlich dazu geeignet, jedes beliebige quantenchemische Verfahren zu verwenden. Besonders reiz- und sinnvoll scheint der Anschluss an ein Mehr-Determinanten-Verfahren zu sein, das zwar Ergebnisse hoher Qualität liefert, aber keine Möglichkeit zur Berechnung analytischer Gradienten vorsieht. Ein solches Verfahren ist die in Abschnitt 2.3.12 (S. 36) erläuterte DFT/MRCI-Methode von Grimme und Waletzke [1999a]. Im Folgenden wird die Optimierung angeregter Zustände des Formaldehyds auf DFT/MRCI-Hyperflächen beschrieben. Die für Geometrie-Optimierungen ausgewählten elektronischen Zustände sind in Tabelle 6.5 aufgelistet.

Zustand	Notation in		Charakter
	C_{2v}	C_s	
S ₀	X^1A_1	X^1A'	
S ₁	1A_2	$^1A''$	$n \rightarrow \pi^*$
S ₄	1A_1	$^1A'$	$\pi \rightarrow \pi^*$
T ₁	3A_2	$^3A''$	$n \rightarrow \pi^*$
T ₂	3A_1	$^3A'$	$\pi \rightarrow \pi^*$

Tabelle 6.5: Formaldehyd: Symmetrie, Notation und Charakter von Grundzustand und untersuchten elektronisch angeregten Zuständen

Während der Optimierungen kam ein TZVP-Basissatz aus der TURBOMOLE-Bibliothek zum Einsatz. Da die Einflüsse von Rydberg-Zuständen auf niedrig liegende angeregte Zustände gering sind (Peyerimhoff u. a. [1971]), wurden entsprechende Funktionen nicht in die Basis miteinbezogen. Als einziger der untersuchten Zustände liegt der $^1(\pi \rightarrow \pi^*)$ -Zustand (S_4) energetisch so hoch, dass er starke Rydberg-Valenz-Mischungen mit dem $^1(n \rightarrow p_y)$ -Zustand aufweist (Hachey u. a. [1995]). Das Verfahren beschränkt sich deshalb auf die Optimierung der Geometrie in der Potenzialmulde mit Valenzcharakter. Für die Ermittlung von Anregungs-Energien wurde der Basissatz um Polarisationsfunktionen erweitert (TZVPP), und die Energie der zu Grunde liegenden TURBOMOLE-DFT/SCF-Module wurde mit 10^{-8} mHa als konvergiert angesehen. Als Austausch-Korrelations-Funktional wurde wie oben die BH-LYP (Becke [1993c]; Lee u. a. [1988]) Parameterisierung eingesetzt.

Der Referenzraum der DFT/MRCI-Berechnungen wurde iterativ bestimmt. Im ersten Schritt wurden die 8 Valenzelektronen des Formaldehyds beliebig auf 10 Valenzorbitale verteilt und hiervon Einfach- und Doppelanregungen erzeugt. Die für die Beschreibung der 12 untersten Wurzeln wichtigsten Konfigurationen wurden dann im zweiten Schritt als Referenz-Konfigurationen benutzt. Der gesamte DFT/MRCI-Raum wurde schließlich durch Einfach- und Doppel-Anregungen aus diesen Referenz-Konfigurationen generiert und bestand im Singulett-Fall aus 501 Konfigurationen, entsprechend 931 CSFs, im Triplett-Fall waren es 395 Konfigurationen und 711 CSFs.

Da die planare Konformation oftmals einen Sattelpunkt auf Hyperflächen pyramidalen Moleküle darstellt, wurde für jeden elektronischen Zustand je eine planare und eine pyramidale Start-Geometrie angesetzt. (Tatsächlich sollte sich herausstellen, dass planare Start-Geometrien immer auch planar relaxierten, auch wenn die betreffende Spezies im Potenzial-Minimum pyramidal gebaut ist.) Die Ergebnisse der Optimierungen sind in den Tabellen 6.6 bis 6.9 zusammengefasst und werden dort mit den Literaturwerten und Ergebnissen anderer Berechnungen verglichen.

Welche Änderungen in der Geometrie kann man auf Grund dieser Anregungen erwarten?

Bindungslängen: Alle vier Anregungen besetzen das antibindende π^* -Orbital der Carbonyl-Funktion und verursachen somit eine Verlängerung der CO-Bindung (Buenker und Peyerimhoff [1970]). Das Ausmaß der Verlängerung sollte bei Anregungen aus dem bindenden π -Orbital deutlicher ausfallen als bei Anregungen aus dem am Sauerstoff-Atom lokalisierten nicht bindenden n -Orbital. Beide Annahmen werden durch experimentelle Daten bestätigt (Clouthier und Ramsay [1983]) und sollten von quantenchemischen Verfahren korrekt reproduziert werden.

Bindungswinkel: Bei den vier betrachteten Anregungen erhöht sich die Ladungsdichte am Kohlenstoff-Atom, weil das nunmehr besetzte π^* -Orbital dort den größeren Koeffizienten hat. Da die erhöhte Ladungsdichte im p -Atomorbital eines sp^2 -

Tabelle 6.6: Formaldehyd: Geometrische Daten des Grundzustandes X^1A_1 (S_0)

Koordinate	Wert	Methode, ggf. Basis	Referenz
R_{CO}	1.186	RCIS/6-31+G*	Foresmann u. a. [1992]
	1.243	CI(PCMO)	Buenker und Peyerimhoff [1970]
	1.206	MRD-CI ^b	Hachey u. a. [1995]
	1.206	CCSD(T)/CBS(aDTQ/ e^{-x})	Dixon und Feller [1998]
	1.208	Exp.	Takagi und Oka [1963]
	1.21	Exp. ^a	Raynes [1966]
	1.203	Exp. ^a	Jones und Coon [1969]
	1.199	DFT/MRCI	<i>diese Arbeit</i>
R_{CH}	1.091	RCIS/6-31+G*	Foresmann u. a. [1992]
	1.103	CCSD(T)/CBS(aDTQ/ e^{-x})	Dixon und Feller [1998]
	1.116	Exp.	Takagi und Oka [1963]
	1.12	Exp. ^a	Raynes [1966]
	1.094	DFT/MRCI	<i>diese Arbeit</i>
\angle_{HCH}	116.3	RCIS/6-31+G*	Foresmann u. a. [1992]
	116.4	CCSD(T)/CBS(aDTQ/ e^{-x})	Dixon und Feller [1998]
	116.5	Exp.	Takagi und Oka [1963]
	115.	Exp. ^a	Raynes [1966]
	116.8	DFT/MRCI	<i>diese Arbeit</i>
\angle_{OCH}	121.9	RHF/6-311+G*	Hadad u. a. [1993]
	121.7	Exp.	Takagi und Oka [1963]
	121.6	DFT/MRCI	<i>diese Arbeit</i>
\angle_{oop}	0.0	RCIS/6-31+G*	Foresmann u. a. [1992]
	0.0	Exp.	Takagi und Oka [1963]
	0.0	Exp. ^a	Raynes [1966]
	0.0	DFT/MRCI	<i>diese Arbeit</i>

^a Rotationsspektroskopie^b MRD-CI mit 5mHa Schwelle, Basissatz: TZVP+Ryd, optimiert wurde lediglich der CO-Bindungsabstand

Tabelle 6.7: Formaldehyd: Geometrische Daten des S_1 -Zustandes ${}^1A''$ ($n \rightarrow \pi^*$)

Koordinate	Wert	Methode, ggf. Basis	Referenz
R_{CO}	1.255	RCIS/6-31+G*	Foresmann u. a. [1992]
	1.397	CI(PCMO)	Buenker und Peyerimhoff [1970]
	1.335	MRD-CI ^a	Hachey u. a. [1995]
	1.384	CASSCF(6e-5MO)+SOC	Bataev u. a. [2000]
	1.324	EOM-CCSD	Stanton u. a. [1995]
	1.340	P-EOM-MBPT(2)	Gwaltney und Bartlett [1999]
	1.323	Exp. ^b	Jensen und Bunker [1982]
	1.29	Exp. ^b	Raynes [1966]
	1.328	DFT/MRCI	<i>diese Arbeit</i>
R_{CH}	1.085	RCIS/6-31+G*	Foresmann u. a. [1992]
	1.096	EOM-CCSD	Stanton u. a. [1995]
	1.093	P-EOM-MBPT(2)	Gwaltney und Bartlett [1999]
	1.103	Exp. ^b	Jensen und Bunker [1982]
	1.10	Exp. ^b	Raynes [1966]
	1.088	DFT/MRCI	<i>diese Arbeit</i>
\angle_{HCH}	118.3	RCIS/6-31+G*	Foresmann u. a. [1992]
	118.6	Exp.	Shah und Moule [1978]
	118.1	Exp. ^b	Jensen und Bunker [1982]
	120.	Exp. ^b	Raynes [1966]
	116.2	DFT/MRCI	<i>diese Arbeit</i>
\angle_{OCH}	117.7	CIS/6-311+G*	Hadad u. a. [1993]
	115.8	EOM-CCSD	Stanton u. a. [1995]
	117.0	P-EOM-MBPT(2)	Gwaltney und Bartlett [1999]
	113.4	DFT/MRCI	<i>diese Arbeit</i>
\angle_{oop}	24.9	RCIS/6-31+G*	Foresmann u. a. [1992]
	31.9	CI(PCMO)	Buenker und Peyerimhoff [1970]
	38.3	CASSCF(6e-5MO)+SOC	Bataev u. a. [2000]
	31.2	EOM-CCSD	Stanton u. a. [1995]
	24.6	P-EOM-MBPT(2)	Gwaltney und Bartlett [1999]
	31.0	Exp.	Herzberg [1966]
	26.9	Exp. ^b	Jones und Coon [1969]
	20.5	Exp. ^c	Job u. a. [1969]
	34.0	Exp. ^b	Jensen und Bunker [1982]
	19.5	Exp. ^b	Raynes [1966]
41.1	DFT/MRCI	<i>diese Arbeit</i>	

^a MRD-CI mit 5mHa Schwelle, Basissatz: TZVP+Ryd, optimiert wurde lediglich der CO-Bindungsabstand

^b Rotationspektroskopie

^c Rotations- und Schwingungs-Spektroskopie

Tabelle 6.8: Formaldehyd: Geometrische Daten des S_4 -Zustandes 2^1A_1 ($\pi \rightarrow \pi^*$)

Koordinate	Wert	Methode, ggf. Basis	Referenz
R_{CO}	1.460	RCIS/6-31+G*	Foresmann u. a. [1992]
	1.542	MRD-CI ^a	Hachey u. a. [1995]
	1.459	DFT/MRCI	<i>diese Arbeit</i>
R_{CH}	1.073	RCIS/6-31+G*	Foresmann u. a. [1992]
	1.075	DFT/MRCI	<i>diese Arbeit</i>
\angle_{HCH}	124.4	RCIS/6-31+G*	Foresmann u. a. [1992]
	137.4	DFT/MRCI	<i>diese Arbeit</i>
\angle_{OCH}	117.8	CIS/6-311+G*	Hadad u. a. [1993]
	113.4	DFT/MRCI	<i>diese Arbeit</i>
\angle_{oop}	0.0	RCIS/6-31+G*	Foresmann u. a. [1992]
	0.0	CI(PCMO)	Buenker und Peyerimhoff [1970]
	0.0	DFT/MRCI	<i>diese Arbeit</i>

^a MRD-CI mit 5mHa Schwelle, Basissatz: TZVP+Ryd, optimiert wurde lediglich der CO-Bindungsabstand

hybridisierten Kohlenstoff-Atoms nach den Regeln des von Gillespie und Nyholm [1957] entwickelten VSEPR-Modells¹ elektrostatisch ungünstiger ist als in einem sp^3 -Hybrid-Orbital, verschiebt sich die Bindungssituation des C-Atoms in Richtung einer sp^3 -Hybridisierung, und das Molekül wird pyramidal. Das Ausmaß der Pyramidalisierung ist annähernd proportional dem verursachenden Ladungstransfer.

Im Falle der $n \rightarrow \pi^*$ -Anregungen wurde die Vorhersage durch Experimente (Clouthier und Ramsay [1983]) gestützt: der S_1 - und der T_1 -Zustand sind pyramidal gebaut, was eine Symmetrie-Erniedrigung zur Untergruppe C_s zur Folge hat. Die korrekte Nomenklatur identifiziert den S_1 -Zustand nunmehr als $^1A''$ und den T_1 -Zustand als $^3A''$. Experimentelle Strukturdaten des S_4 -Zustandes sind nicht bekannt, womöglich weil es energetisch in der Nähe des Ionisierungs-Potenzials von 10.87eV liegt und daher spektroskopisch schwer zugänglich ist (Mentall u. a. [1971]). Theoretische Verfahren berechnen es übereinstimmend planar (s. Tabellen).

¹ „Valence shell electron pair repulsion“

Tabelle 6.9: Formaldehyd: Geometrische Daten des T_1 -Zustandes ${}^3A''$ ($n \rightarrow \pi^*$, oben) und des T_2 -Zustandes ${}^3A'$ ($\pi \rightarrow \pi^*$, unten)

Koordinate	Wert	Methode, ggf. Basis	Referenz
<i>T₁-Zustand</i>			
R_{CO}	1.256	RCIS/6-31+G*	Foresmann u. a. [1992]
	1.376	CI(PCMO)	Buenker und Peyerimhoff [1970]
	1.363	CASSCF(6e-5MO)+SOC	Bataev u. a. [2000]
	1.28	Exp. ^a	Raynes [1966]
	1.307	Exp. ^a	Jones und Coon [1969]
	1.269	DFT/MRCI	<i>diese Arbeit</i>
R_{CH}	1.093	RCIS/6-31+G*	Foresmann u. a. [1992]
	1.10	Exp. ^a	Raynes [1966]
	1.091	DFT/MRCI	<i>diese Arbeit</i>
\angle_{HCH}	112.4	RCIS/6-31+G*	Foresmann u. a. [1992]
	116.	Exp. ^a	Raynes [1966]
	114.5	DFT/MRCI	<i>diese Arbeit</i>
\angle_{OCH}	113.8	CIS/6-311+G*	Hadad u. a. [1993]
	114.1	DFT/MRCI	<i>diese Arbeit</i>
\angle_{oop}	43.1	RCIS/6-31+G*	Foresmann u. a. [1992]
	32.2	CI(PCMO)	Buenker und Peyerimhoff [1970]
	40.0	CASSCF(6e-5MO)+SOC	Bataev u. a. [2000]
	35.6	Exp. ^a	Jones und Coon [1969]
	35.0	Exp.	Herzberg [1966]
	38.5	Exp. ^a	Raynes [1966]
	41.1	Exp.	Clouthier und Ramsay [1983]
41.0	DFT/MRCI	<i>diese Arbeit</i>	
<i>T₂-Zustand</i>			
R_{CO}	1.408	RCIS/6-31+G*	Foresmann u. a. [1992]
	1.405	CIS/6-311+G*	Hadad u. a. [1993]
	1.423	Exp. ^b	Taylor u. a. [1982]
	1.441	DFT/MRCI	<i>diese Arbeit</i>
R_{CH}	1.072	RCIS/6-31+G*	Foresmann u. a. [1992]
	1.073	CIS/6-311+G*	Hadad u. a. [1993]
	1.073	DFT/MRCI	<i>diese Arbeit</i>
\angle_{HCH}	119.2	RCIS/6-31+G*	Foresmann u. a. [1992]
	118.7	CIS/6-311+G*	Hadad u. a. [1993]
	122.6	DFT/MRCI	<i>diese Arbeit</i>
\angle_{OCH}	112.9	CIS/6-311+G*	Hadad u. a. [1993]
	116.1	DFT/MRCI	<i>diese Arbeit</i>
\angle_{oop}	39.4	RCIS/6-31+G*	Foresmann u. a. [1992]
	40.2	CIS/6-311+G*	Hadad u. a. [1993]
	24.1	DFT/MRCI	<i>diese Arbeit</i>

^a Elektronen-Energie-Verlust-Spektroskopie (EELS)^b Rotationsspektroskopie

Zur Diskussion:

S₀: Die planare Grundzustands-Geometrie wird gut reproduziert, die Abweichungen der CO-Bindungslänge von experimentell ermittelten Literaturwerten liegen bei 0.4pm (Jones und Coon [1969]) bzw. 0.9pm (Takagi und Oka [1963]). Die CH-Bindung liegt 1.2pm unter dem von Takagi und Oka [1963] spektroskopisch ermittelten Wert. Die Bindungswinkel zeigen die erwartete Abweichung von der Trigonalität. Die Differenzen liegen damit in den üblichen und tolerablen Bereichen. Inkonsistenzen zwischen Experiment und Theorie treten nicht zu Tage.

S₁: Sowohl die antizipierte Verlängerung der CO-Bindung als auch die Pyramidalisierung können gefunden werden, doch während die Bindungslängen eine sehr gute Übereinstimmung mit experimentellen Daten zeigen, liegt der Außerplanar-Winkel \angle_{oop} der CO-Achse aus der HCH-Ebene mit 41.1° deutlich oberhalb des Intervalls der Literaturwerte von 19.5° bis 34.0° . Ebenso ist die berechnete Inversionsbarriere mit 0.086eV doppelt so hoch wie von Jones und Coon [1969] gemessene von 0.043eV. Da auch die Spreizung des \angle_{HCH} -Winkels etwas zu niedrig ermittelt wird, liegt die Vermutung nahe, dass das verwendete Verfahren das Ausmaß des Ladungstransfers vom Sauerstoff- auf das Kohlenstoff-Atom geringfügig überschätzt und das Molekül daher etwas zu stark pyramidal verzerrt.²

S₄: Da die Anregungsenergie des S₄-Zustandes ($\pi \rightarrow \pi^*$), wie bereits angemerkt, im Bereich der ersten Ionisierungsenergie des Formaldehyds liegt (10.87eV, Mentall u. a. [1971]), sind keine experimentellen Daten verfügbar. Die gute Übereinstimmung der Bindungslängen mit den von Foresmann u. a. [1992] sowie Hadad u. a. [1993] mit CIS-Verfahren berechneten Werten ist eine gewisse Bestätigung. Als Einziger in dieser Arbeit berechneter angeregter Zustand relaxiert S₄ während der Optimierung unter Retention der Planarität. Folglich bleiben auch die Symmetrie-Eigenschaften erhalten und der S₄-Zustand ist ein ¹A₁-Zustand.

T₁: Die Übereinstimmung mit dem Experiment ist zufrieden stellend: Die CO-Bindungslänge wird um 3.8pm (Jones und Coon [1969]) bzw. 1.1pm (Raynes [1966]) niedriger ermittelt, Ähnliches gilt für die CH-Bindung. Der Außerplanar-Winkel (41.0°) liegt am oberen Ende des Intervalls der Literaturwerte (35° bis 41.1°). Die Inversionsbarriere von 0.096eV ist identisch mit der bei Jones und Coon [1969] angegebenen.

²Eine TURBOMOLE-DFT/SCF-Optimierung mit dem B3-LYP-Austausch-Korrelations-Funktional führte beim S₀-Grundzustand zu etwas längeren Bindungen, die nach den Regeln des VSEPR-Modells auch eine geringere Pyramidalisierung zur Folge hätten. Da das DFT/MRCI-Programm parametrisch an das BH-LYP-Funktional angepasst wurde, lässt sich der Einfluss des B3-LYP-Funktional auf die Strukturen angeregter Zustände bedauerlicherweise nicht nachprüfen.

T_2 : Auch hier kann die prognostizierte Bindungsverlängerung im Vergleich zum Grundzustand gefunden werden. Der Wert von 1.441\AA liegt zwischen dem experimentellen und denen anderer theoretischer Arbeiten. Darüber hinaus zeigt das Molekül eine leichte Tendenz zur pyramidalen Konformation: der Außerplanar-Winkel wurde mit 24.1° und einer entsprechend niedrigen Inversionsbarriere von 0.014eV berechnet. Experimentelle oder theoretische Vergleichsdaten zur Inversionsbarriere fehlen.

6.3 Anregungs-Energien

Die Reproduktion von Anregungs-Energien kann als ein weiterer Maßstab für die Qualität eines quantenchemischen Optimierungs-Verfahrens fungieren. Abbildung 6.1 illustriert verschiedene Arten der Anregung am Beispiel der Potenzialflächen eines S_0 -Grundzustandes und eines angeregten Singulett-Zustandes S_1 . Unter einer

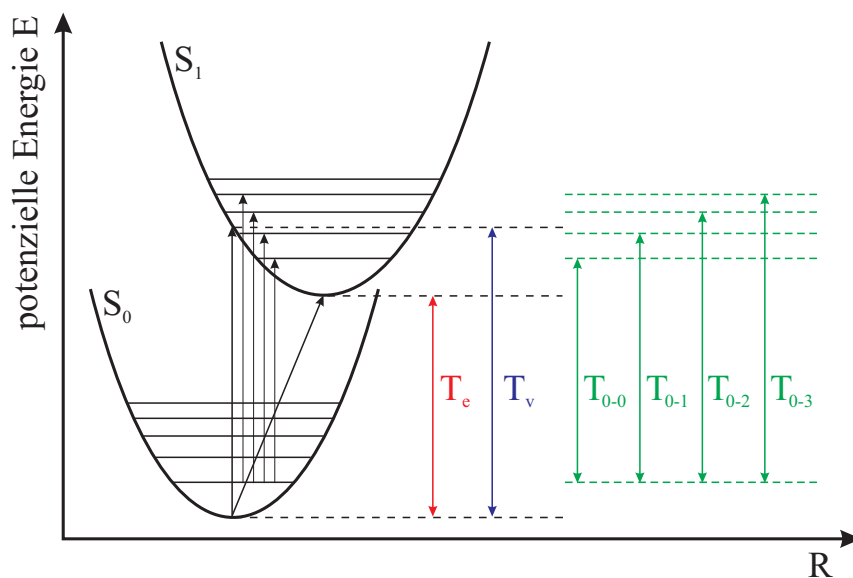


Abbildung 6.1: Verschiedene Anregungs-Arten

vertikalen Anregungsenergie T_v versteht man die Potenzial-Differenz der Kurven bei fixierter Geometrie. Sie wird meist für die Gleichgewichtsgeometrie des Grundzustandes bestimmt. Bei relaxierter Geometrie erhält man die adiabatische Anregungsenergie T_e , definiert als die Potenzial-Differenz zweier Minima. Die Anregungsenergien T_{0-0} , T_{0-1} , T_{0-2} , etc. entsprechen Anregungsenergien aus dem 0. Schwingungsniveau des elektronischen Grundzustandes in das n -te ($0,1,2$, etc.) Schwingungsniveau eines elektronisch angeregten Zustandes. Dabei wird die Anregungsenergie T_{0-0} häufig näherungsweise mit der adiabatischen Anregungsenergie gleichgesetzt.

Vertikale elektronische Spektren, wie man sie z.B. durch Mehrdeterminanten-Verfahren erhalten kann, sind primär ein Maß für die Qualität einer quantenchemischen Methode. Da die Geometrie dabei unverändert bleibt, machen sie keinerlei Aussagen über Optimierungs-Verfahren. Um so aussagekräftiger ist eine Analyse der adiabatischen Anregungsenergien, die in Tabelle 6.10 zusammengestellt sind.

Tabelle 6.10: Formaldehyd: Adiabatische Anregungsenergien der elektronisch angeregten Zustände ${}^1A''$, 1A_1 , ${}^3A''$ und ${}^3A'$

Zustand	Charakter	Wert [Å]	Methode, ggf. Basis	Referenz
$S_1: {}^1A''$	$n \rightarrow \pi^*$	4.40	RCIS/6-31+G*	Foresmann u. a. [1992]
		4.54	CIS/6-311+G*	Hadad u. a. [1993]
		4.33	CIS,MP2/6-311+G*	Hadad u. a. [1993]
		3.81	CI(PCMO)	Peyerimhoff u. a. [1971]
		3.64	MRD-CI ^a	Hachey u. a. [1995]
		3.50	Exp.	Herzberg [1966]
		3.511	Exp. ^b	Brand [1956]
		3.506	Exp. ^c	Taylor u. a. [1982]
		3.423	DFT/MRCI	<i>diese Arbeit</i>
$S_4: {}^1A_1$	$\pi \rightarrow \pi^*$	8.607	RCIS/6-31+G*	Foresmann u. a. [1992]
		8.65	CIS/6-311+G*	Hadad u. a. [1993]
		8.60	CIS,MP2/6-311+G*	Hadad u. a. [1993]
		7.95	MRD-CI ^a	Hachey u. a. [1995]
		8.044	DFT/MRCI	<i>diese Arbeit</i>
$T_1: {}^3A''$	$n \rightarrow \pi^*$	3.501	RCIS/6-31+G*	Foresmann u. a. [1992]
		3.63	CIS/6-311+G*	Hadad u. a. [1993]
		3.41	CI(PCMO)	Peyerimhoff u. a. [1971]
		3.12	Exp.	Herzberg [1966]
		3.124	Exp. ^b	Brand [1956]
		3.128	Exp. ^c	Taylor u. a. [1982]
		3.13	Exp. ^d	Raynes [1966]
		3.093	DFT/MRCI	<i>diese Arbeit</i>
$T_2: {}^3A'$	$\pi \rightarrow \pi^*$	3.525	RCIS/6-31+G*	Foresmann u. a. [1992]
		3.64	CIS/6-311+G*	Hadad u. a. [1993]
		4.72	Exp. ^e	Taylor u. a. [1982]
		4.328	DFT/MRCI	<i>diese Arbeit</i>

^a MRD-CI mit 5mH Schwelle, Basissatz: TZVP+Ryd, optimiert wurde lediglich der CO-Bindungsabstand

^b optisches Spektrum

^c Elektronen-Energie-Verlust-Spektroskopie (EELS)

^d Rotationsspektroskopie

^e extrapoliert aus einem Elektronen-Energie-Verlust-Spektrum in Taylor u. a. [1982]

Die gefundenen Werte liegen meist geringfügig unterhalb der experimentellen Literaturwerte. Dabei weist der T_2 -Zustand mit ca. 0.4eV die größte Abweichung auf, während der „Fehler“ der $n \rightarrow \pi^*$ -Anregungen in einer Größenordnung von weniger als einem Zehntel eV zu Buche schlägt. Zusammenfassend lässt sich die verwendete DFT/MRCI-Methode einmal mehr als hochwertig und der numerische Optimierer als verlässlich einstufen.

Kapitel 7

Uracil

Die Pyrimidinbase Uracil steht vielfach im Interesse spektroskopischer Untersuchungen. Sowohl Uracil als auch das 5-Methyl-Derivat Thymin weisen sehr breite, unstrukturierte Banden im UV-Bereich auf, unabhängig davon, ob man sie in der Gasphase oder in Lösung untersucht (Voet u. a. [1963]; Clark u. a. [1965]; Becker und Kogan [1980]; Callis [1983]; Fujii u. a. [1986]; Tsuchiya u. a. [1988]; Brady u. a. [1988]). Die erste Bande, die bei ca. 4.5eV (Vakuum) beginnt, wurde einem $\pi \rightarrow \pi^*$ Übergang zugeordnet (Clark u. a. [1965]; Brady u. a. [1988]), was durch *ab initio*-Studien gestützt wird (Petke u. a. [1992]; Lorentzon u. a. [1995]). Die Autoren der Arbeiten dokumentieren, dass erst der S_2 -Zustand ein hinreichendes Dipol-Übergangsmoment mit dem S_0 -Grundzustand aufweist. Ein $S_0 \leftarrow S_1$ Übergang sollte den Berechnungen zu Folge lediglich ca. ein Tausendstel der Intensität aufweisen, doch konnten Lorentzon u. a. [1995] die vertikale Anregungsenergie zum S_1 -Zustand mit ca. 4.5eV angeben, was gerade dem Beginn der beobachteten, breiten Bande entspricht, so dass eine Beteiligung wahrscheinlich ist. Die auffällige Breite der Absorptionsbanden bleibt jedoch rätselhaft. Eine Fragmentierung des Uracil-Moleküls nach Anregung im Spektralbereich unterhalb von 5.2eV wurde nicht beobachtet. Schließt man deshalb die Anregung in einen ungebundenen Zustand als Ursache für die Verbreiterung aus, so verbleiben folgende mögliche Ursachen:

1. **Energiedissipation durch interne Konversion:**

Erfährt ein Molekül bei elektronischer Anregung eine drastische Änderung seiner Geometrie, so dass die Potenzialkurven von Gleichgewichts- und angeregtem Zustand stark gegeneinander verschoben sind, so erfolgt die Anregung in der Regel mit größeren Franck-Condon-Faktoren in höhere vibronische Zustände. Da die Schwingungszustandsdichte mit zunehmender Energie immer größer wird, kann es hier zu breiten Banden als Folge einer Überlagerung vieler vibronischer Übergänge ähnlicher Energien und Intensitäten kommen.

2. **Vibronische Kopplung mit einem gebundenen Zustand:**

Unter bestimmten Voraussetzungen kann es zu einer vibronischen Kopplung

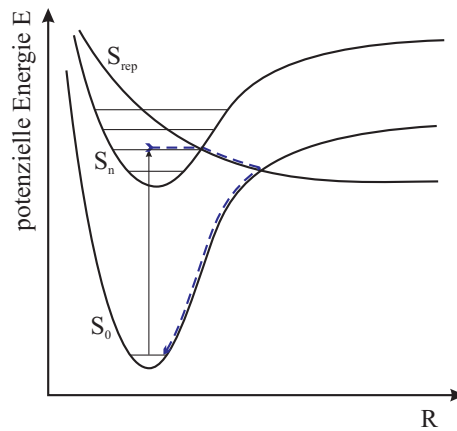


Abbildung 7.1: Anregung in einen repulsiven Zustand S_{rep} (hier: ${}^1\pi\sigma^*$) und möglicher Relaxationspfad durch Prädissociation, konische Überschneidung und geometrische Relaxation

mit einem energetisch nahe liegenden „dunklen“ Zustand kommen. Dazu muss das Produkt der irreduziblen Darstellungen von elektronischer und Schwingungs-Wellenfunktion der koppelnden Zustände (hier mit 1 und 2 indiziert) gleich sein:

$$\Gamma_{\Psi_{el_1}} \otimes \Gamma_{\Psi_{vib_1}} = \Gamma_{\Psi_{el_2}} \otimes \Gamma_{\Psi_{vib_2}} \quad (7.1)$$

3. Konische Durchschneidung mit einem repulsiven Zustand:

Existiert auf der Potenzialhyperfläche des angeregten Zustands im Bereich der Anregung eine konische Durchschneidung mit einem repulsiven Zustand, so kann auch dies die Lebensdauer hinreichend verkürzen, um eine Verbreiterung der Bande zu verursachen. Die Bedingungen sind in Abbildung 7.1 vereinfacht illustriert. Sobolewski u. a. [2002] konnten zeigen, dass diese Konstellation z.B. in Nukleinbasen und aromatischen Aminosäuren auftreten kann. In ihnen existieren repulsive ${}^1(\pi \rightarrow \sigma^*)$ -Zustände, die eine schnelle Prädissociation von NH- oder OH-Bindungen verursachen können. Durch eine zweite konische Überschneidung können sie in den Grundzustand wechseln und dort relaxieren, so dass der Bindungsbruch nicht vollständig vollzogen wird..

4. Intersystem-Crossing (ISC):

Unter einem *Intersystem-crossing* versteht man den strahlungslosen Übergang zwischen elektronischen Zuständen unterschiedlicher Multiplizität. Auch sie können die Lebensdauer einer elektronisch angeregten Spezies verkürzen und so breite Banden verursachen. Voraussetzungen für ein ISC sind die energetische Nähe zweier elektronischer Zustände und eine starke Spin-Bahn-Kopplung.

Um einige dieser Erklärungsansätze bestätigen oder ausschließen zu können, wurden Geometrie-Optimierungen einiger relevanter elektronisch angeregter Zustände durchgeführt, deren Dipolmomente bestimmt sowie vertikale und adiabatische Anregungs-Energien berechnet.

7.1 Technische Details

Die Geometrie-Optimierungen wurden mit Hilfe der Routinen durchgeführt, die im Rahmen dieser Arbeiten implementiert wurden.

Um eine zuverlässige Erfassung von dynamischer und statischer Elektronen-Korrelation zu gewährleisten, kam die von Grimme und Waletzke [1999a] entwickelte kombinierte DFT/MRCI-Methode zum Einsatz. Wie bereits mehrfach ausgeführt, erfordert dies die Benutzung des BHLYP-Funktional und es beschränkt die Berechnungen auf Singulett und Triplett-Zustände.

Die für die Ermittlung der Vertikal-Spektren verwendeten Basissätze sind von TZVP bzw. TZVPP-Qualität und wurden der TURBOMOLE Basissatz-Bibliothek entnommen. Um Rydberg-Beiträge zu erfassen, wurde die TZVPP-Basis um diffuse $3s$, $3p$ und eine d Gaußfunktion erweitert, deren Ursprung an einem Dummy-Zentrum lag, und die die Exponenten 0.05, 0.02, 0.008 (s - und p -Rydberg) und 0.015 (d -Rydberg) trugen. Diese Basis soll im Folgenden mit TZVPP+Ryd benannt werden. Die Lage des Dummy-Zentrums deckte sich mit dem Masseschwerpunkt. Die Integration der Austausch-Korrelations-Wechselwirkung wurde auf einem Cäsium-analogen Gitter vorgenommen. Da die Erweiterung um Polarisations- und Rydberg-Funktionen, wie noch zu sehen sein wird, bei den niedrig liegenden Zuständen keine signifikanten Änderungen der DFT/MRCI-Ergebnisse bewirken, wurden alle Geometrie-Optimierungen mit der TZVP-Basis in C_1 -Symmetrie durchgeführt.

Ein Satz Kohn-Sham-Orbitale, erzeugt durch SCF-Optimierung der Grundzustands-Determinante, diente als Ein-Teilchen-Basis für die aufsetzenden MRCI-Rechnungen. Für die Ermittlung des MRCI-Referenzraumes der Geometrie-Optimierungen wurden 8 Elektronen beliebig auf 8 Valenzorbitale verteilt und jeweils Einfach-Anregungen vorgenommen. Man erhält für beide Multiplizitäten je 17 wichtige Konfigurationen für 8 Wurzeln in C_1 -Symmetrie. Der MRCI-Raum wiederum wurde durch Einfach- und Doppel-Anregungen dieser Konfigurationen gebildet, so dass die Berechnungen der Singulett-Zustände mit beinahe 18.000 Konfigurationen und ca. 35.000 CSFs erfolgten, im Triplett-Fall waren es etwa 12.000 Konfigurationen und über 30.000 CSFs.

Da die vertikalen Anregungs-Energien bei Grundzustands-Geometrie möglichst präzise bestimmt werden sollten, wurden hier 8 Elektronen auf 18 Orbitale verteilt und einfach angeregt, so dass der Referenzraum für die Beschreibung von 12 a' - und 12 a'' -Wurzeln ca. 64 Konfigurationen bzw. etwa 79 CSFs enthielt. Auch hier wurde der gesamte MRCI-Raum daraus durch Einfach- und Doppel-Anregungen erzeugt,

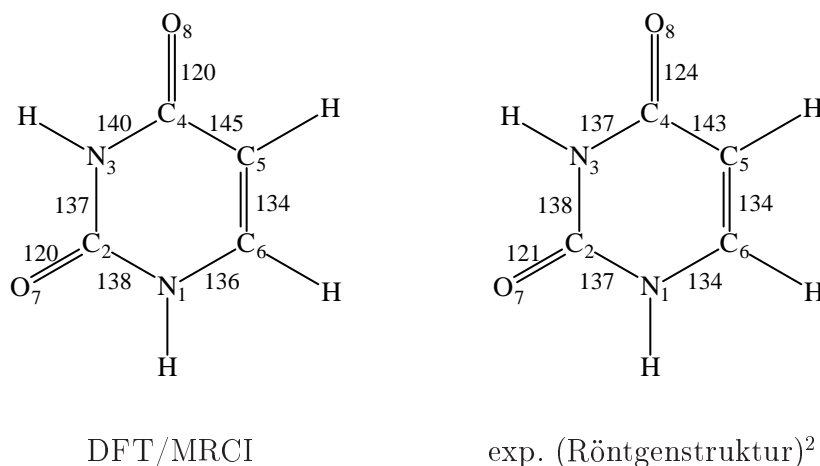
was zu über 49.000 berücksichtigten Konfigurationen und über 114.000 CSFs führte, jeweils für Singulett- und Triplett-Zustände.

Die Bestimmung adiabatischer Anregungs-Energien schließlich wurde in einer TZVP-Basis ohne Rydberg-Funktionen durchgeführt. Der Referenzraum entstand durch eine auf Einfach-Anregungen beschränkte CAS(8,8)-analoge Prozedur und erzeugte zwischen 60 und 80 Konfigurationen. Aus diesen entstand nach analogem Einfach- und Doppel-Anregungsmuster der MRCI-Raum als Grundlage für weitere Berechnungen.

7.2 Der elektronische Grundzustand

Die mit dem hier entwickelten Programm optimierte Geometrie¹ ist in Abbildung 7.2 dargestellt. Die Bindungslängen des Sechsrings stimmen sehr gut mit denen

Abbildung 7.2: Uracil: Grundzustands-Geometrie, numerisch optimiert auf einer DFT/MRCI-Hyperfläche und von Steward und Jensen [1967] röntgen-spektroskopisch bestimmt



einer Röntgen-Struktur-Untersuchung von Steward und Jensen [1967] überein. Die Carbonyl-Bindungen dagegen werden mit je 20pm kürzer ermittelt als im Experiment (C₂-O₇: 121.5pm und C₄-O₈: 124.5pm). Womöglich kann man diesen Effekt zumindest teilweise der im Uracil-Kristall auftretenden N-H...O-Koordination zuschreiben, zumal aktuellen theoretischen Berechnungen zu Folge die Koordination mit Wasser-Molekülen eine Bindungs-Streckung von 2-3pm bewirkt (Marian u. a. [2002]).

¹Auf die Angabe von CH-Bindungslängen wird hier und im Folgenden verzichtet.

²Im Kristall liegt Uracil als Wasserstoffbrücken-gebundenes Dimer vor (beteiligt sind die Gruppen/Atome N₃H und O₈).

Um sich ein qualitatives Bild der noch zu untersuchenden elektronisch angeregten Zustände zu machen, sind die Valenz-Orbitale des Uracils in den Abbildungen 7.3 und 7.4 dargestellt. Die Orbitale sind mit ihrem vorwiegenden Bindungscharakter

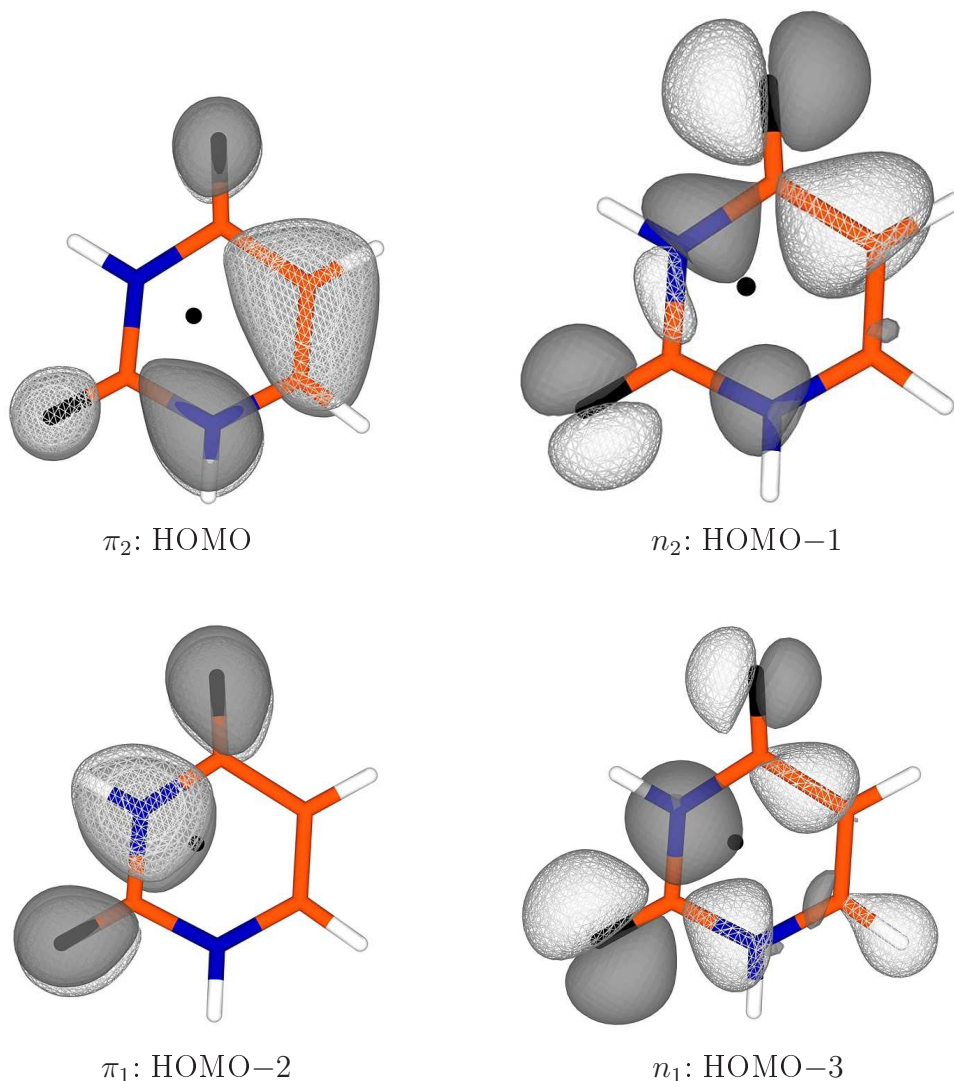


Abbildung 7.3: Höchste besetzte n - und π -Valenzorbitale des Uracils im Grundzustand S_0 . Der schwarze Punkt innerhalb des Sechsrings lokalisiert das Rydberg-Zentrum.

bezeichnet und in ihrer energetischen Reihenfolge indiziert.

HOMOs: Das höchste besetzte Molekülorbital (HOMO) π_2 ist C_5 - C_6 -bindend und an anderen Zentren nicht-bindend. Das zweithöchste besetzte Orbital, HOMO-1, ist ein in der Ring-Ebene liegendes, nicht-bindendes n -Orbital mit den größten Koeffizienten in einem p -Atomorbital des O_8 -Atoms und in der C_4 - C_5 - σ -Bindung. Ein einzelnes Elektronenpaar am Sauerstoff-Atom O_7 dominiert das n_1 -Orbital HOMO-3.

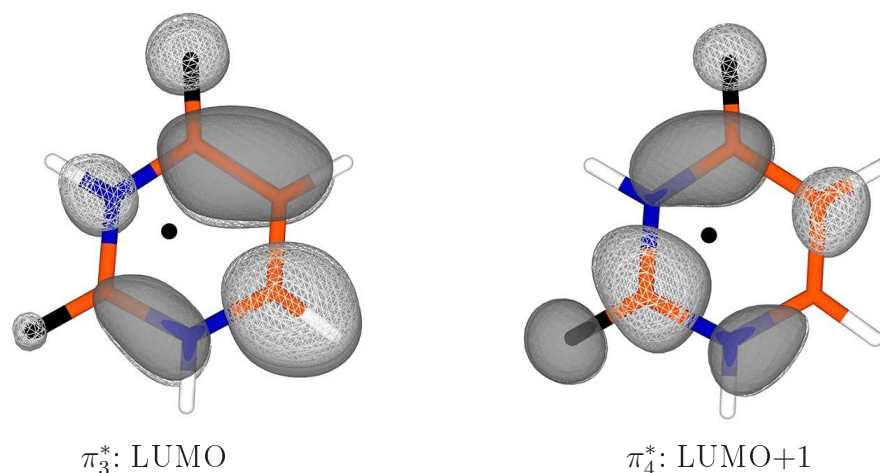


Abbildung 7.4: Tiefste unbesetzte n - und π -Valenzorbitale des Uracils im Grundzustand S_0 . Der schwarze Punkt innerhalb des Sechsrings lokalisiert das Rydberg-Zentrum.

Das dritthöchste besetzte Orbital π_1 (HOMO-2) schließlich hat seine größten Beiträge in p_z -Orbitalen am N_3 -Atom und beiden Sauerstoff-Atomen.

LUMOs: Beide tiefstliegenden unbesetzten Valenz-Orbitale haben π -Charakter; sie sind leicht Carbonyl- C_4 - O_8 -antibindend. Das π_3^* -Orbital (LUMO) weist darüber hinaus vor allem C_4 - C_5 -bindende und C_5 - C_6 -antibindende Anteile auf, während das nächsthöhere Orbital (π_4^* , LUMO+1) zusätzlich durch eine antibindende Carbonyl- C_2 - O_7 -Wechselwirkung gekennzeichnet ist.

In der energetischen Lücke zwischen diesen beiden LUMOs finden sich mehrere Rydberg-Orbitale, die hier nicht einzeln aufgeführt werden sollen. Es fällt auf, dass die Orbitale π_1 (HOMO-2) und π_4^* (LUMO+1) annähernd C_{2v} -symmetrisch sind bezüglich einer zur Ringebene vertikalen Spiegelebene durch die Atome N_3 und C_6 .

7.3 Vertikale Anregungsenergien

Für eine eingehendere Betrachtung der spektroskopischen Gegebenheiten des Uracils macht es Sinn, die energetische Lage und dominante Anregungs-Charaktere der elektronisch angeregten Zustände zuzuordnen. Dazu wurden vertikale Spektren sowohl für Singulett- als auch Triplett-Zustände berechnet und mit Literaturwerten verglichen. Die Daten sind in den Tabellen 7.1 (Singulett) und 7.2 (Triplett) zusammengestellt.

Es zeigt sich, dass für eine korrekte Beschreibung elektronisch angeregter Zustände des Uracils oberhalb von ca. 5.7eV eine Berücksichtigung von Rydberg-

Tabelle 7.1: Uracil: Vertikale Anregungsenergien ΔE [eV] und Oszillator-Stärken $f(r)$ der Singulett-Zustände

Zustand	HF/MRCI; RPA ^a		CASPT2 ^b		DFT/MRCI, diese Arbeit				Absorptionsmaxima	
	DZVP+Ryd		TZVP+Ryd		TZVP	TZVPP+Ryd				
	ΔE_{MRCI}	ΔE_{RPA}	ΔE	$f(r)$	ΔE	ΔE	$f(r)$	Charakter(e)		
S0	1 ¹ A'	0.00	0.00	0.00		0.00	0.00			
S1	1 ¹ A''	5.48	6.04	4.54	0.00	4.61	4.61	0.00	$n_2 \rightarrow \pi_3^*$	4.68 ^c
S2	2 ¹ A'	6.28	6.29	5.00	0.19	5.48	5.44	0.26	$\pi_2 \rightarrow \pi_3^*$	5.08 ^d , 4.81 ^e , 4.77 ^{d,e} , 4.73 ^h
S3	2 ¹ A''							5.83	$\pi_2 \rightarrow R$	
S4	3 ¹ A''	7.15	7.32	6.00	0.00	(6.05)	5.95	0.00	$n_2 \rightarrow \pi_4^*$, $n_1 \rightarrow \{\pi_3^*, \pi_4^*\}$	
S5	3 ¹ A'			5.81	0.08	(6.14)	6.15	0.05	$\pi_1 \rightarrow \pi_3^*$	(6.05) ^d , 6.05 ^g , 6.11 ^e , 6.14 ^f
S6	4 ¹ A'			6.46	0.29	(6.79)	6.53	0.16	$\pi_2 \rightarrow \{\pi_4^*, R\}$, $n_2 \rightarrow R$	
S7	4 ¹ A''			6.37	0.00	(6.65)	6.57	0.00	$n_1 \rightarrow \pi_3^*$, $n_2 \rightarrow \pi_4^*$	
S8	5 ¹ A'					(7.56)	6.65	0.08	$n_2 \rightarrow R$, $\pi_2 \rightarrow \{R, \pi_4^*\}$	
S9	5 ¹ A''						6.71	0.01	$\pi_2 \rightarrow R$	
S10	6 ¹ A''						6.76	0.01	$\pi_2 \rightarrow R$	
S11	7 ¹ A''			6.95	0.00	(7.05)	6.85	0.00	$n_2 \rightarrow \{\pi_4^*, R\}$	
S12	6 ¹ A'						7.07	0.01	$\pi_2 \rightarrow R$	
S13	8 ¹ A''						7.09	0.00	$\pi_1 \rightarrow R$	
S14	7 ¹ A'						7.17	0.05	$n_2 \rightarrow R$	
S15	8 ¹ A'						7.29	0.02	$n_1 \rightarrow R$, $n_2 \rightarrow R$	
S16	9 ¹ A'			7.01	0.76	(7.66)	7.39	0.51	$\pi_1 \rightarrow \{\pi_4^*, R\}$, $n_2 \rightarrow R$	6.97 ^g

^a Petke u. a. [1992]^b Lorentzon u. a. [1995]^c schwacher senkrechter Übergang in orthorhombischen Kristallen von 1-Methyluracil, Eaton und Lewis [1970]^d Gasphasen-Spektrum, Clark u. a. [1965]; (Schulter)^e in Trimethylphosphat, Clark u. a. [1965]^f in Wasser, Voet u. a. [1963]^g in Wasser, Callis [1983]^h in 2-Methyltetrahydrofuran und Ethanol/Methanol, Becker und Kogan [1980]

Tabelle 7.2: Uracil: Vertikale Anregungsenergien ΔE [eV] und Oszillator-Stärken $f(r)$ der Triplet-Zustände

Zustand	HF/MRCI ^a DZVP+Ryd	DFT/MRCI, diese Arbeit				
		TZVP	TZVPP+Ryd			
	ΔE	ΔE	ΔE	$f(r)$	Charakter(e)	
T1	1 ¹ A'	4.00	3.68	3.68		$\pi_2 \rightarrow \pi_3^*$
T2	1 ¹ A''	5.30	4.40	4.39	0.00	$n_2 \rightarrow \pi_3^*$
T3	2 ¹ A'	6.15	5.10	5.08	0.01	$\pi_1 \rightarrow \pi_3^*, \pi_2 \rightarrow \pi_4^*$
T4	2 ¹ A''	6.94		5.73	0.00	$\pi_2 \rightarrow R$
T5	3 ¹ A'	7.28	(5.85)	5.74	0.01	$\pi_2 \rightarrow \pi_4^*, \pi_1 \rightarrow \{\pi_3^*, \pi_4^*\}$
T6	3 ¹ A''		(5.87)	5.76	0.00	$n_2 \rightarrow \pi_4^*, n_1 \rightarrow \{\pi_3^*, \pi_4^*\}$
T7	4 ¹ A'		(6.39)	6.25	0.02	$\pi_1 \rightarrow \{\pi_4^*, R, \pi_3^*\}, \pi_2 \rightarrow \pi_4^*$
T8	4 ¹ A''		(6.59)	6.50	0.00	$n_1 \rightarrow \{\pi_3^*, \pi_4^*\}$
T9	5 ¹ A'			6.55	0.00	$n_2 \rightarrow R$
T10	5 ¹ A''			6.63	0.00	$\pi_2 \rightarrow R$
T11	6 ¹ A''			6.67	0.00	$\pi_2 \rightarrow R$
T12	7 ¹ A''		(6.91)	6.73	0.00	$n_1 \rightarrow \{\pi_4^*, R, \pi_3^*\}, n_2 \rightarrow \pi_3^*$
T13	6 ¹ A'			6.98	0.00	$\pi_2 \rightarrow R$
T14	8 ¹ A''			7.01	0.00	$\pi_1 \rightarrow R$
T15	7 ¹ A'			7.09	0.00	$n_2 \rightarrow R$
T16	8 ¹ A'			7.19	0.00	$n_1 \rightarrow R, n_2 \rightarrow R$
T17	9 ¹ A'		(7.41)	7.38	0.09	$\pi_0 \rightarrow \pi_3^*$

^a Petke u. a. [1992]

Funktionen notwendig ist. Die untersten beiden angeregten Singulett-Zustände (S_1 und S_2) sowie die niedrigsten drei Triplet-Zustände (T_1 , T_2 und T_3) werden dagegen mit dominanten Valenz-Anregungen hinreichend gut charakterisiert:

S_1 :	1 ¹ A''	$n_2 \rightarrow \pi_3^*$	HOMO-1 \rightarrow LUMO
S_2 :	2 ¹ A'	$\pi_2 \rightarrow \pi_3^*$	HOMO \rightarrow LUMO
T_1 :	1 ³ A'	$\pi_2 \rightarrow \pi_3^*$	HOMO \rightarrow LUMO
T_2 :	1 ³ A''	$n_2 \rightarrow \pi_3^*$	HOMO-1 \rightarrow LUMO
T_3 :	2 ³ A'	$\pi_1 \rightarrow \pi_3^*$,	HOMO-2 \rightarrow LUMO,
		$\pi_2 \rightarrow \pi_4^*$	HOMO \rightarrow LUMO+1

S_3 (5.83eV) und T_4 (5.73eV) bezeichnen bereits Anregungen aus dem π_2 - in Rydberg-Orbitale, die in den Tabellen 7.1 und 7.2 allgemein mit R notiert werden. Höhere Anregungen sind meist beträchtlich gemischt.

Wie für eine Carbonyl-Verbindung zu erwarten, haben $n \rightarrow \pi^*$ -Anregungen (z.B. S_1 und T_2) vernachlässigbar kleine Oszillator-Stärken, sie werden daher in Spektren für gewöhnlich nicht beobachtet. Eine Ausnahme ist ein Wert, der von Eaton

Tabelle 7.3: Uracil: Dipol-Momente μ [D] verschiedener Zustände in der S₀- und ihren Gleichgewichts-Geometrien (DFT/MRCI, TZVP Basis)

Zustand	S ₀ Geometrie	Gleichgewichts-Geometrie
S ₀ , 1 ¹ A'	4.58	4.58
S ₁ , 1 ¹ A''	2.54	2.98
S ₂ , 2 ¹ A'	5.85	5.11
T ₁ , 1 ³ A'	3.89	3.75
T ₂ , 1 ³ A''	2.67	3.02

und Lewis [1970] gefunden wurde; der von ihnen gefundene Übergang im 1-Methyl-Uracil ist charakterisiert durch ein senkrechtes Übergangsdipolmoment — er wird hier zum Vergleich mit heran gezogen, weil der Einfluss der Methylgruppe auf die Anregungs-Energie nur marginal sein dürfte. Dagegen ergeben sich in Übereinstimmung mit früheren theoretischen Berechnungen (Petke u. a. [1992]; Lorentzon u. a. [1995]; Broo und Holmén [1997]) und experimentellen Befunden (s.u.) signifikante Absorptions-Wahrscheinlichkeiten für $\pi \rightarrow \pi^*$ -Anregungen.

Alle hier berechneten intensiven Übergänge lassen sich spektroskopisch beobachtbaren Banden zuordnen (siehe Voet u. a. [1963]; Clark u. a. [1965]; Becker und Kogan [1980]; Callis [1983]). Dabei ist zu beachten, dass es durch Lösungsmittel-Effekte zu beträchtlichen Verschiebungen der Banden-Maxima kommen kann: Im Gasphasen-Spektrum finden Eaton und Lewis [1970] den ersten $\pi \rightarrow \pi^*$ -Übergang bei ca. 244nm (5.08eV), Spektren in polaren Lösungsmitteln (Trimethylphosphat, Wasser, 2-Methyl-THF und einem Ethanol/Methanol-Gemisch) lokalisieren ihn ca. 0.3eV tiefer. Da der S₂-Zustand nach unseren Berechnungen (auf TZVPP+Ryd-Basis, siehe Tabelle 7.3) ein Dipolmoment von 5.85D aufweist, wird er in polaren Medien elektrostatisch stärker stabilisiert als der Grundzustand mit einem Dipolmoment von 4.58D.

Die von Clark u. a. [1965] im Gasphasen-Spektrum beobachtete Schulter bei 205nm (6.05eV) lässt sich dem S₀ \rightarrow S₅-Übergang zuordnen. Der S₅-Zustand hat ein kleineres Dipolmoment (2.43D) als der Grundzustand, daher wird diese Bande in polaren Lösungsmitteln um ca 0.05eV blau-verschoben. Die nächsthöhere Bande korrespondiert mit einem S₀ \rightarrow S₆-Übergang bei 187nm (6.63eV, Clark u. a. [1965]); auch sie wird bei polarer Solvatation gegenüber dem Grundzustand blau-verschoben, obwohl deren Dipolmomente annähernd gleich berechnet wurden (S₆: 3.95D, S₀: 4.58D). Der Übergang mit der größten Oszillator-Stärke ist der mit 7.39eV recht hoch angesiedelte Übergang vom S₀- zum S₁₆-Zustand. Er scheint mit der von Callis [1983] dokumentierten Bande bei 177nm (6.97eV) für Uracil in Wasser übereinzuz-

stimmen; dieselbe Zuordnung trafen auch Lorentzon u. a. [1995]. Die vorherrschende elektronische Konfiguration im mit Rydberg-Anteilen gemischten S_{16} -Zustand entsteht durch eine $\pi_1 \rightarrow \pi_4^*$ -Anregung. Diesen beiden Orbitalen wurde oben näherungsweise C_{2v} -Symmetrie zugeordnet (siehe auch Abbildungen 7.3/S. 99 und 7.4/S. 100). Wie bereits Lorentzon u. a. [1995] argumentiert haben, wird das beträchtliche Übergangsmoment durch einen Ladungstransfer vom Stickstoffatom N_3 zu den Carbonylgruppen hervorgerufen.

Die energetische Reihenfolge der Triplett-Zustände ist eine andere. Hier repräsentiert der T_1 -Zustand die HOMO \rightarrow LUMO-Anregung ($\pi_2 \rightarrow \pi_3^*$) und T_2 die im Singulett-Spektrum tiefer ermittelte (HOMO-1) \rightarrow LUMO-Anregung ($n_2 \rightarrow \pi_3^*$). Wie bereits vorweg genommen (s. S. 102, ist der nächsthöhere Zustand, T_3 , eine Mischung aus zwei $\pi \rightarrow \pi^*$ -Anregungen. In den darüber liegenden Zuständen wächst der Einfluss von Rydberg-Anregungen.

Der Großteil dieser Berechnungen stimmt recht gut mit den CASPT2-Kalkulationen von Lorentzon u. a. [1995] überein, die allerdings die Ermittlung von Rydberg- und Triplett-Zuständen aussparten. Wie bereits Grimme und Waletzke [1999a] in einer Testreihe mit DFT/MRCI-Rechnungen feststellten, werden $\pi \rightarrow \pi^*$ -Anregungen durch das Verfahren energetisch meist etwas überschätzt, auch im Vergleich zu experimentellen Daten. Die HF/MRCI- und RPA²-basierten Ergebnisse von Petke u. a. [1992] liegen offensichtlich zu hoch.

7.4 Angeregte Zustände

Drei Geometrien der Zustände S_1 , S_2 , T_1 und T_2 sind in Abbildung 7.5 wiedergegeben. Optimierungen mit pyramidalen Startgeometrien wurden zwar angesetzt,

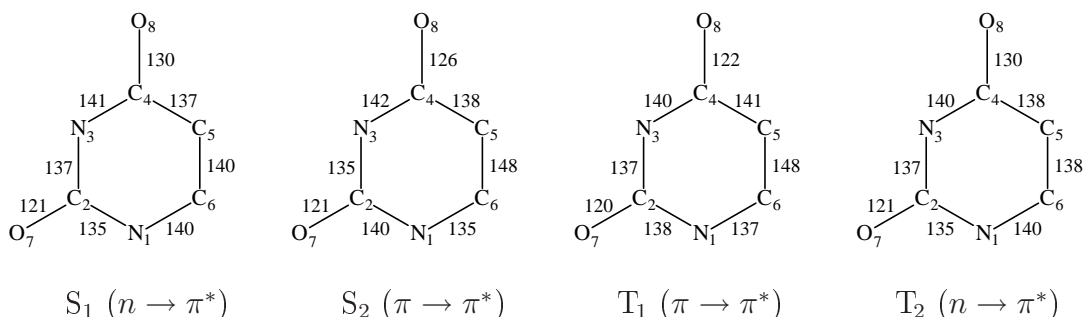


Abbildung 7.5: Uracil: Geometrien ausgewählter angeregter Zustände (C,N,O-Gerüst)

konnten allerdings aus Zeitgründen nicht beendet werden. Deutliche Tendenz der

²Random phase approximation

ersten Optimierungs-Schritte war jedoch übereinstimmend eine Wiederherstellung der Planarität.

Bei keiner der hier betrachteten Anregungen kommt es zu gravierenden Änderungen der Bindungs-Kette $C_4-N_3-C_2-O_7$, da die beteiligten Grenzorbitale n_2 , π_2 und π_3^* hier keine bindenden und kaum antibindende Komponenten haben (siehe Abbildungen 7.3/S. 99 und 7.4/S. 100).

Die $n_2 \rightarrow \pi_3^*$ -Anregung führt bei Singulett- und Triplett-gekoppelter Anregung zu sehr ähnlichen Geometrien. Dabei werden die Tendenzen der Änderung durch die Wechselwirkungen innerhalb des π_3^* -Orbitals festgelegt: Die Carbonyl-Bindung C_4-O_8 , die C_5-C_6 -Doppelbindung und die C_6-N_1 -Bindung werden entsprechend ihres antibindenden Charakters in π_3^* gestreckt; der Effekt ist am größten in der Carbonyl-Gruppe (10pm), gefolgt von der $C-C$ -Doppelbindung (7pm bzw. 5pm) und der $C-N$ -Bindung, die nur relativ wenig beeinflusst wird (3pm). Die C_4-C_5 -Bindung zeigt den für elektronisch angeregte konjugierte Ketene typischen, zunehmenden Doppelbindungscharakter, in Folge dessen wird sie um 8pm bzw. 7pm kürzer. Auch der Doppelbindungs-Anteil der N_1-C_2 -Bindung, die im Grundzustand schon Teil eines konjugierten Systems ist, wird noch leicht erhöht, so dass sich der Kernabstand um 3pm verringert.

Bei den $\pi \rightarrow \pi^*$ -Anregungen müssen zusätzlich zu den konstruktiven und destruktiven Einflüssen des nunmehr besetzten π_3^* -Orbitals auch die Einflüsse des nach der Anregung nur noch einfach besetzten π_2 -Orbitals berücksichtigt werden, und zwar gewissermaßen mit negativem Vorzeichen. Dies führt einerseits dazu, dass sich gegenläufige Tendenzen annähernd kompensieren, so z.B. bei der Carbonyl-Gruppe, andererseits aber verstärken sich gleich gerichtete Einflüsse, so dass bspw. die C_5-C_6 -Bindung um 15pm gestreckt wird. Unterschiede zwischen der S_2 - und T_1 -Geometrie findet man vor allem in der Kette $O_8-C_4-C_5$; hier sind die Bindungslängen im Singulett-Zustand verglichen mit dem Triplett-Zustand einander leicht angeglichen. Dies kann darauf zurückgeführt werden, dass der Singulett-Zustand mit zwitterionischen Konfigurationen mischen kann, während der Triplett-Zustand höchstens Biradikal-Anteile haben kann. Diese Tendenzen machen sich auch im Dipolmoment des S_2 - und T_1 -Zustandes bemerkbar (S_2 : 5.11D; T_1 : 3.75D).

Die Tabelle 7.4 zeigt Anregungs-Energien einiger elektronischer Zustände bei optimierten Geometrien relativ zur Grundzustands-Energie. Auf der Tabellen-Diagonale finden sich die adiabatischen Anregungs-Energien. Die Maxima der beobachteten breiten Absorptionsbanden entsprechen ungefähr den vertikalen Übergängen, während der Bandenursprung in etwa mit der adiabatischen Energie korreliert. Es muss festgestellt werden, dass die Energien des S_2 -Zustandes konsistent um ca. 0.4eV zu hoch berechnet werden. Testrechnungen haben gezeigt, dass die Energie des Uracil-Dimers energetisch auf ca. 5.1eV abgesenkt wird. Je nach apparativen Bedingungen könnte die Bande im Gasphasenspektrum also eventuell vom Uracil-Dimer herrühren.

Tabelle 7.4: Uracil: Relative Energien ΔE [eV] bei verschiedenen, auf DFT/MRCI-Niveau optimierten Geometrien. Die Nomenklatur S_0 , S_1 , S_2 , T_1 , T_2 , and T_4 folgt der energetischen Reihenfolge der Zustände bei der Gleichgewichts-Geometrie des Grundzustandes S_0 .

Zustand	Geometrien					
	S_0	S_1	S_2	T_1	T_2	T_4^a
	$n \rightarrow \pi^*$	$\pi \rightarrow \pi^*$	$\pi \rightarrow \pi^*$	$n \rightarrow \pi^*$	$\pi \rightarrow \sigma^*$	
$S_0, 1^1A'$	0.00	0.59	0.63	0.47	0.52	5.84
$S_1, 1^1A''$	4.61	3.96	4.30	4.41	3.96	4.11
$S_2, 2^1A'$	5.48	5.21	4.99	5.10	5.23	5.88
$T_1, 1^3A'$	3.68	3.51	3.29	3.20	3.52	6.09
$T_2, 1^3A''$	4.40	3.84	4.22	4.32	3.83	4.97
$T_3, 2^3A'$	5.10	4.72	4.99	5.04	4.73	6.52

^a mit TURBOMOLE UDFT-optimiert

7.5 Diskussion

Kehren wir zurück zu der anfänglich gestellten Frage nach möglichen Ursachen für die breite Bande im Ultraviolett-Spektrum des Uracils. Einige der hier ermittelten Daten sollten bei einer Aufklärung von Nutzen sein:

zu Ansatz 1: Voraussetzung für die Verbreiterung einer Bande durch Anregung in vibronisch hoch liegende Zustände ist eine drastische Änderung der Molekül-Geometrie. Die Änderung der Geometrie bei dem im unteren Spektralbereich interessanten $S_0 \rightarrow S_2$ -Übergang ist für ein solches Szenario allerdings vermutlich nicht ausreichend.

zu Ansatz 2: Die Auswahlregeln für eine vibronische Kopplung wurden mittels der Gleichung 7.1 formuliert. Der im unteren Spektralbereich relevante elektronische Zustand S_2 hat die irreduzible Darstellung A' . Der darunter liegende, für eine vibronische Kopplung in Frage kommende, „dunkle“ Zustand, S_1 , ist ein A'' -Zustand, so dass auch die irreduziblen Darstellungen ihrer Schwingungs-Wellenfunktionen Ψ_{vib} entgegengesetzt sein müssen, um eine Kopplung zu ermöglichen.

Einen Hinweis auf den Charakter der gesuchten vibronischen Wellenfunktion des S_2 -Zustandes liefert das Spektrum bei Becker und Kogan [1980]. Darin erscheint die betreffende Bande als eine Schwingungs-Progression mit einem Abstand von ca. 790cm^{-1} . Die harmonische Schwingungs-Analyse des S_0 -Zustandes auf DFT-

Niveau liefert zwei dafür in Frage kommende Moden: Mode 11 mit 798cm^{-1} ist eine Außer-Planar-Schwingung des Kohlenstoff-Atoms C_2 im Dreieck $\text{N}_1\text{N}-\text{C}-\text{O}$ ($\rightarrow A''$), und Mode 12 mit 802cm^{-1} repräsentiert eine Ring-Atmungs-Schwingung ($\rightarrow A'$). Während die Frequenz der ersten Mode sehr gut mit der bei 806cm^{-1} gemessenen IR-Bande übereinstimmt (Szczesniak u. a. [1983]), konnte Mode 12 wohl auf Grund ihrer geringen Intensität von 4km/mol im IR-Spektrum bisher nicht experimentell bestätigt werden. Da die numerische Optimierung des S_2 -Zustandes eine planare Geometrie ergab, ist die Anregung einer Außer-Planar-Schwingung wie in Mode 11 eher unwahrscheinlich. Die Geometrieänderung von S_0 zu S_2 entspricht ziemlich genau der Schwingungskordinate von Mode 12. Die Anregung dieser Mode ist daher sehr wahrscheinlich. Ordnet man also dem A' -Zustand S_2 eine A' -Schwingung (Mode 12) zu, so muss für eine erfolgreiche vibronische Kopplung der S_1 -Zustand mit einer A'' -Mode schwingen. Da die Moden 11 und 12 energetisch sehr nah beieinander liegen, handelt es sich dabei vermutlich um Mode 11, und die Voraussetzungen für eine vibronische Kopplung wären erfüllt.

zu Ansatz 3: Auf den ersten Blick ist keiner der berechneten Zustände S_1 - S_{16} als repulsiv zu erkennen. Anregungen erfolgen entweder in antibindende π^* - oder Rydberg-Orbitale, nicht aber in antibindende σ^* -Orbitale. Um so überraschender ist es, dass der T_4 -Zustand, der einer Anregung aus dem höchsten besetzten Orbital (π_2) in das niedrigste virtuelle a' -Orbital entspricht, bei einer Optimierung mittels TURBOMOLE-UDFT-Module zu einer Dissoziation der N_1 -H-Bindung führt. Während der Optimierung gewinnt das a' -Orbital, das bei S_0 -Gleichgewichts-Geometrie noch diffuse Rydberg-Eigenschaften hat, immer mehr Valenz-Charakter und wird schließlich zu dem gesuchten antibindenden σ^* -Orbital. Der S_3 -Zustand wird ebenfalls durch eine Anregung aus dem π_2 - in das tiefstliegende a' -Orbital dominiert, so dass sich auch dieser bei zunehmender N_1 -H-Abstraktion als $\pi \rightarrow \sigma^*$ -Zustand notieren lässt, der als Singulett-Zustand den in Abbildung 7.1 blau dargestellten Pfad einer NH-Prädissoziation, konischen Überschneidung mit dem Grundzustand und anschließender Relaxation auf der S_0 -Hyperfläche ermöglichen dürfte. Die konische Überschneidung konnte bisher nicht lokalisiert werden, so dass hier keine definitive Aussage möglich ist. Doch spricht die Tatsache, dass der S_3 -Zustand bei T_4 -Geometrie mit 4.11eV energetisch unterhalb des ursprünglichen Grundzustandes (5.84eV) erscheint, deutlich dafür. Der S_3 -Zustand liegt im Vertikal-Spektrum bei berechneten 5.83eV , also gilt dieser Erklärungsversuch ausschließlich für den höheren Spektralbereich.

zu Ansatz 4: Nach dem Prinzip der Energie-Nähe (*energy gap rule*) sind die Voraussetzungen für eine Kopplung immer dann besonders günstig, wenn die kopplenden elektronischen Zustände ähnliche Energien haben. Da die Energien des T_3 - und des S_2 -Zustandes bei S_2 -Geometrie als identisch berechnet wurden, ist die Existenz signifikanter Spin-Bahn-Kopplung hier sehr wahrscheinlich. Wie Marian u. a.

[2002] allerdings unlängst feststellen konnten, ist das Spin-Bahn-Kopplungs-Matrix-Element dieser Zustände nahezu Null, was ein ISC ausschließt. T_2 liegt energetisch nur ca. 0.4eV unterhalb des S_2 -Zustandes und hat ein nicht-verschwindendes Spin-Bahn-Kopplungs-Matrix-Element, doch sprechen die konkreten Werte für eine ISC-Rate von ungefähr 10^6 /sec. Für eine beobachtbare Banden-Verbreiterung wären ca. 10^9 /sec nötig, wie etwa bei konischen Überschneidungen.

Abschließend lässt sich feststellen, dass von den hier untersuchten Ansätzen die vibronische Kopplung (Ansatz 2) im niedrigen Spektralbereich und die konische Durchschneidung mit einem repulsiven Zustand im höheren Spektralbereich als ursächlich für die beobachtete Bandenverbreiterung gelten können.

Kapitel 8

Zusammenfassung & Ausblick

In der vorliegenden Arbeit wurde über die Implementation eines Optimierungsmoduls für DFT/MRCI-Energiehyperflächen berichtet. Das Programm ist zum jetzigen Zeitpunkt in der Lage, die DFT/MRCI-Energien von kartesisch verzerrten Kerngeometrien anzufordern und nachfolgend zu nutzen, um Information über molekulare Gradienten zu akkumulieren und so die Relaxation eines Systems bzw. das Finden von stationären Punkten zu ermöglichen. Um auf anderen Energieflächen zu optimieren, muss lediglich die Einleseroutine angepasst werden, was dem Programm universell einsetzbar macht. Der Optimierer startet zur Relaxation von Verbindungen mit einem *Steepest-Descent*-Verfahren, das nach Kontrolle auf bestimmte Konvergenzkriterien von einem *Conjugate-Gradient*-Algorithmus abgelöst wird, da Letzterer üblicherweise in der Nähe von stationären Punkten verlässlicher arbeitet.

Typische Optimierungen von Formaldehyd konvergieren innerhalb von Minuten bis Stunden auf einem handelsüblichen PC (1GHz Taktung, 1GB RAM), der kompliziertere Anwendungsfall Uracil benötigte typischerweise mehrere Tage, startet man von stärker verzerrten Geometrien. Der Speicherbedarf wird von dem Programm bestimmt, das die elektronische Energie berechnet, da beim Optimierer kaum nennenswert Daten im Hauptspeicher der Maschine gehalten werden muss.

Eine Parallelisierung ist für die nahe Zukunft vorgesehen und sollte hier entscheidende Vorteile mit sich bringen, da die numerische Gradientenbestimmung auf voneinander unabhängigen Einzelpunktberechnungen fußt.

In der Validierungsphase wurde der Treiber erfolgreich mit den Turbomole-Modulen verglichen. Hier diente das Formaldehyd-Molekül, da über diese Verbindung ein umfangreicher Vergleichsdatensatz sowohl auf Seiten der Experimente als auch bei den theoretisch Forschenden existiert, insbesondere auch für elektronisch angeregte Spezies. Beim Test-/Anwendungsfall Uracil konnten zwei mögliche Ursachen für die in Experimenten beobachtete, ungewöhnlichen Bandenbreiten im elektronischen Spektrum gefunden werden. Zwei andere Hypothesen wurden als unwahrscheinlich eingestuft.

Anhang A

Notationen

Die vorliegende Arbeit verwendet folgende Notationen:

n	Anzahl der Elektronen
N	Anzahl der Kerne
Z	Kernladung
\mathbf{r}	Ortskoordinatenvektor eines Elektrons, lateinisch indiziert (i, j, k, \dots)
ω	Spinkoordinate eines Elektrons, lateinisch indiziert (i, j, k, \dots)
\mathbf{x}	Orts- und Spinkoordinaten eines Elektrons, lateinisch indiziert (i, j, k, \dots)
\mathbf{R}	Ortskoordinatenvektor eines Kerns, griechisch indiziert (α, β, \dots)
ϕ	Atom-Raum-Orbitale, griechisch indiziert (μ, ν, λ, \dots)
ψ	Molekül-Raum-Orbitale, lateinisch indiziert (i, j, k, \dots)
χ	Molekül-Spin-Orbitale
Ψ	Mehr-Elektronen Wellenfunktion
a, b, c, \dots	Indizes für besetzte Orbitale
r, s, t, \dots	Indizes für unbesetzte (virtuelle) Orbitale
$\hat{H}, \hat{J}, \hat{K}, \dots$	Mehr-Teilchen-Operatoren
$\hat{h}, \hat{j}, \hat{k}, \dots$	Ein-Teilchen-Operatoren
E	Energie
V	Potenzielle Energie
T	Kinetische Energie
ϵ	Orbitalenergie

Literaturverzeichnis

- [Baker 1987] BAKER, Jon: An Algorithm for Geometry Optimizations without Analytical Gradients. In: *J. Comp. Chem.* 8 (1987), Nr. 5, S. 563–574 52, 72
- [Bataev u. a. 2000] BATAEV, V.A. ; PUPYSHEV, V.I ; ABRAMENKOV, A.V. ; GODUNOV, I.A.: Quantum-Mechanical Calculations of the Structure and Dynamics of Conformationally Nonrigid Carbonyl Compound Molecules in the Ground and Excited States. In: *Russ. J. Phys. Chem.* 74 (2000), S. S279–S291 87, 89
- [Becke 1988] BECKE, A. D.: Density-functional exchange-energy approximation with correct asymptotic behavior. In: *Phys. Rev. A* 38 (1988), S. 3098–3107 27
- [Becke 1993a] BECKE, A. D.: . In: *J. Chem. Phys.* 98 (1993), S. 1372–1377 35
- [Becke 1993b] BECKE, A. D.: Density-functional thermochemistry. III. The role of exact exchange. In: *J. Chem. Phys.* 98 (1993), S. 5648–5652 28
- [Becke 1993c] BECKE, A.D.: A new mixing of Hartree-Fock and local density-functional theories. In: *J. Chem. Phys.* 98 (1993), S. 1372–1377 85
- [Becker und Kogan 1980] BECKER, Ralph ; KOGAN, Gregory: Photophysical Properties of Nucleic Acid Components — 1. The Pyrimidines: Thymine, Uracil, N,N-Dimethyl Derivatives and Thymidine. In: *Photochem. Photobiol.* 31 (1980), S. 5–13 95, 101, 103, 106
- [Birss u. a. 1973] BIRSS, F. W. ; DONG, R. Y. ; RAMSAY, D. A.: In: *Chem. Phys. Lett.* 18 (1973), 11–13 S 79
- [Brady u. a. 1988] BRADY, B. B. ; PETEANU, L. A. ; LEVY, D. H.: The Electronic Spectra of the Pyrimidine Bases Uracil and Thymine in a Supersonic Molecular Beam. In: *Chem. Phys. Lett.* 147 (1988), S. 538–543 95
- [Brand 1956] BRAND, J.C.D.: In: *J. Chem. Soc.* (1956), 858 S 92
- [Broo und Holmén 1997] BROO, A. ; HOLMÉN, A.: Calculations and Characterization of the Electronic Spectra of DNA Bases Based on *ab Initio* MP2 Geometries of Different Tautomeric Forms. In: *J. Phys. Chem.* 101 (1997), S. 3589–3600 103
- [Broyden 1970] BROYDEN, C.G: The Convergence of a Class of Double-Rank Minimization Algorithms. In: *J. Inst. Math. Applic.* 6 (1970), S. 76–90 57

- [Buenker und Peyerimhoff 1970] BUENKER, Robert J. ; PEYERIMHOFF, Sigrid D.: Combined SCF and CI Method for the Calculation of Electronically Excited States of Molecules: Potential Curves for the Low-Lying States of Formaldehyde. In: *J. Chem. Phys.* 53 (1970), Nr. 4, S. 1368–1384 79, 85, 86, 87, 88, 89
- [Callis 1983] CALLIS, P. R.: In: *Annu. Rev. Phys. Chem.* 34 (1983), 329–357 S 95, 101, 103
- [Chutjian 1974] CHUTJIAN, Ara: Electron-impact excitation of the low-lying electronic states of formaldehyde. In: *J. Chem. Phys.* 15 (1974), Nr. 10, S. 4279–4284 79
- [Clark u. a. 1965] CLARK, L. B. ; PESCHEL, G. G. ; TINOCO, JR., I.: Vapor Spectra and Heats of Vaporization of Some Purine and Pyrimidine Bases. In: *J. Phys. Chem.* 69 (1965), S. 3615–3618 95, 101, 103
- [Clouthier und Ramsay 1983] CLOUTHIER, D.J ; RAMSAY, D.A.: The Spectroscopy of Formaldehyde and Thioformaldehyde. In: *Annu. Rev. Phys. Chem.* 34 (1983), S. 31–58 79, 85, 88, 89
- [Cohen und Handy 2000] COHEN, Aron J. ; HANDY, Nicholas C.: Assessment of exchange correlation functionals. In: *Chem. Phys. Lett.* 316 (2000), S. 160–166 26
- [Császár und Pulay 1984] CSÁSZÁR, P. ; PULAY, P.: Geometry optimization by direct inversion in the iterative subspace. In: *J. Mol. Struct.* 114 (1984), S. 31–34 61
- [Davidon 1959] DAVIDON, W. C.: Variable Metric Method for Minimization. In: *AEC Report* (1959), Nr. ANL-5990 (Rev.) 57
- [Davidson und McMurchie 1982] DAVIDSON, Ernest A. ; MCMURCHIE, Larry E.: *Ab Initio Calculations of Excited-State Potential Surfaces of Polyatomic Molecules*. Bd. 5. S. 1–39, Academic Press, 1982 79
- [Dirac 1930] DIRAC, P. A. M.: Note on exchange phenomena in the Thomas atom. In: *Proc. Cambridge Phil. Soc.* 26 (1930), S. 376–385 27
- [Dixon und Feller 1998] DIXON, David A. ; FELLER, David: Heats of Formation of CF₂, FCO, and CF₂O. In: *J. Phys. Chem.* 102 (1998), S. 8209–8216 86
- [Eaton und Lewis 1970] EATON, W. A. ; LEWIS, T. P.: In: *J. Chem. Phys.* 53 (1970), 2164–2172 S 101, 102, 103
- [Eyring 1935] EYRING, Henry: The Activated Complex in Chemical Reactions. In: *J. Chem. Phys.* 3 (1935), S. 107 43
- [Fletcher und Reeves 1964] FLETCHER, R. ; REEVES, C. M.: Function minimisation by conjugate gradients. In: *Comp. J.* (1964), S. 149 55
- [Fletcher 1970] FLETCHER, R.: A New Approach to Variable Metric Algorithms. In: *Comp. J.* 13 (1970), S. 317–322 57

- [Fletcher und Powell 1963] FLETCHER, R. ; POWELL, M. J. D.: A Rapidly Convergent Descent Method for Minimization. In: *Comp. J.* 6 (1963), S. 163–168 57
- [Foresmann u. a. 1992] FORESMANN, James B. ; HEAD-GORDON, Martin ; POPLE, John A. ; FRISCH, Michael J.: Toward a Systematic Molecular Orbital Theory for Excited States. In: *J. Phys. Chem.* 96 (1992), S. 135–149 86, 87, 88, 89, 90, 92
- [Fujii u. a. 1986] FUJII, M. ; TAMURA, T. ; MIKAMI, N. ; ITO, M.: In: *Phys. Chem. Lett.* 126 (1986), S. 583 95
- [Gill u. a. 1981] GILL, P. E. ; MURRAY, W. ; WRIGHT, M. H.: *Practical Optimization*. Academic Press, London, 1981 72, 75
- [Gill 1998] GILL, Peter M. W.: Density Functional Theory (DFT), Hartree-Fock (HF), and the Self-consistent Field. In: SCHLEYER, P. V. R. (Hrsg.) ; ALLINGER, N. L. (Hrsg.) ; CLARK, T. (Hrsg.) ; GASTEIGER, J. (Hrsg.) ; KOLLMAN, P. A. (Hrsg.) ; III., H. S. (Hrsg.) ; SCHREINER, P. R. (Hrsg.): *The Encyclopedia of Computational Chemistry*. John Wiley Sons, 1998, S. 678–689 27
- [Gillespie und Nyholm 1957] GILLESPIE, R. J. ; NYHOLM, R. S.: In: *Quart. Rev.* (1957), Nr. 11, 339 S 88
- [Goldfarb 1970] GOLDFARB, D.: A Family of Variable Metric Updates Derived by Variational Means. In: *Mathematics of Computing* 24 (1970), S. 23–26 57
- [Grein und Hachey 1996] GREIN, Friedrich ; HACHEY, Michael R.: In: *Int. J. Quantum Chem.* S30 (1996), 1661 S 79
- [Grimme 1996] GRIMME, Stefan: Density functional calculations with configuration interaction for the excited states of molecules. In: *Chem. Phys. Lett.* 259 (1996), S. 128–137 33
- [Grimme und Waletzke 1999a] GRIMME, Stefan ; WALETZKE, Mirko: A combination of Kohn-Sham density functional theory and multi-reference configuration interaction methods. In: *J. Chem. Phys.* 111 (1999), Nr. 13, S. 5645–5655 84, 97, 104
- [Grimme und Waletzke 1999b] GRIMME, Stefan ; WALETZKE, Mirko: A combination of Kohn-Sham density functional theory and multi-reference configuration interaction methods. In: *J. Chem. Phys.* 111 (1999), Nr. 13, S. 5645–5655 33, 65
- [Gwaltney und Bartlett 1999] GWALTNEY, Steven R. ; BARTLETT, Rodney J.: Gradients for the partitioned equation-of-motion coupled-cluster method. In: *J. Chem. Phys.* 110 (1999), Nr. 1, S. 62–71 79, 87
- [Hachey u. a. 1995] HACHEY, Michael R. ; BRUNA, Pablo J. ; GREIN, Friedrich: Spectroscopy of Formaldehyde. 1. Ab Initio Studies on Singlet Valence and Rydberg States of Planar H₂CO, with Emphasis on ¹(π, π^*) and ¹(σ, π^*). In: *J. Phys. Chem.* 99 (1995), S. 8050–8057 79, 85, 86, 87, 88, 92

- [Hachey und Grein 1995] HACHEY, Michael R. ; GREIN, Friedrich: The important role of $^1(\pi, \pi^*)$ in the UV absorption spectrum of formaldehyde, as shown by ab initio MR CI studies. In: *J. Phys. Chem.* 99 (1995), S. 8050–8057 79
- [Hadad u. a. 1993] HADAD, Christopher M. ; FORESMANN, James B. ; WIBERG, Kenneth B.: Excited States of Carbonyl Compounds. 1. Formaldehyde and Acetaldehyde. In: *J. Phys. Chem.* 97 (1993), S. 4293–4312 86, 87, 88, 89, 90, 92
- [Handy und Schaefer 1984] HANDY, N. C. ; SCHAEFER, Henry F.: In: *J. Chem. Phys.* 81 (1984), 5031 S 50
- [Handy 1988] HANDY, Nicholas C.: Density Functional Theory. Springer, 1988 (Lecture Notes in Quantum Chemistry II), S. 91–123 26
- [Harding und Goddard III 1977] HARDING, Lawrence B. ; GODDARD III, William A.: Ab Initio Theoretical Studies of the Rydberg States of Formaldehyde. In: *J. Am. Chem. Soc.* 99 (1977), Nr. 3, S. 677–683 79
- [Herzberg 1966] HERZBERG, G.: *Electronic Spectra and Electronic Structure of Polyatomic Molecules*. New York : Van Nostrand, 1966 87, 89, 92
- [Hodges u. a. 1958] HODGES, S. E. ; HENDERSON, J. R. ; COON, J. B.: In: *J. Mol. Spectr.* 2 (1958), 99–102 S 79
- [Hohenberg und Kohn 1964] HOHENBERG, P. ; KOHN, W.: Inhomogeneous Electron Gas. In: *Phys. Rev. B* 136 (1964), S. 864–871 23
- [Jensen und Bunker 1982] JENSEN, Per ; BUNKER, P.R.: The Geometry and the Inversion Potential Function of Formaldehyde in the \tilde{A}^1A_2 and \tilde{a}^3A_2 Electronic States. In: *J. Mol. Spectrosc.* 94 (1982), S. 114–125 87
- [Job u. a. 1969] JOB, V.A. ; SETHURAMAN, V. ; INNES, K.K.: The 3500 \tilde{A}^1A_2 - \tilde{X}^1A_1 Transition of Formaldehyde- h_2 , d_2 , and hd . In: *J. Mol. Spectrosc.* 30 (1969), S. 365–425 79, 87
- [Johnson u. a. 1993] JOHNSON, Benny G. ; GILL, Peter M. W. ; POPLE, John A.: The performance of a family of density functional methods. In: *J. Chem. Phys.* 98 (1993), S. 5612–5626 26
- [Jones und Coon 1969] JONES, V.T. ; COON, J.B.: Rotational Constants and Geometrical Structure of the 1A_2 and 1A_2 States of H_2CO and D_2CO . In: *J. Mol. Spectrosc.* 31 (1969), S. 137–154 79, 86, 87, 89, 90
- [Karlsruher Quantenchemie-Gruppe 1998] KARLSRUHER QUANTENCHEMIE-GRUPPE: *Turbomole User's Manual*. Version 5. Universität Karlsruhe, Germany: Inst. f. Physikalische Chemie, 11 1998 80, 82
- [Kutzelnigg 1975] KUTZELNIGG, Werner: *Einführung in die Theoretische Chemie, Band 1: Quantenmechanische Grundlagen*. Verlag Chemie, Weinheim, 1975 13

- [Lee u. a. 1988] LEE, C. ; YANG, W. ; PARR, R.G.: Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. In: *Phys. Rev. B* 37 (1988), S. 785–788, 85
- [Lorentzon u. a. 1995] LORENTZON, J. ; FÜLSCHER, M. P. ; ROOS, B. O.: Theoretical Study of the Electronic Spectra of Uracil and Thymine. In: *J. Am. Chem. Soc.* 117 (1995), S. 9265–9273 95, 101, 103, 104
- [Marian u. a. 2002] MARIAN, Christel M. ; SCHNEIDER, Frank ; KLEINSCHMIDT, Martin ; TATCHEN, Jörg: Electronic excitation and singlet-triplet coupling in uracil tautomers and uracil-water complexes. In: *Eur. Phys. J.* (2002). – zur Veröffentlichung angenommen 98, 107
- [Mentall u. a. 1971] MENTALL, J.E. ; GENTIEU, E.D. ; KRAUSS, M. ; NEUMANN, D.: Photoionization and Absorption Spectrum of Formaldehyde in the Vacuum Ultraviolet. In: *J. Chem. Phys.* 55 (1971), S. 5471–5479 88, 90
- [Merchán und Roos 1995] MERCHÁN, M. ; ROOS, B.O.: In: *Theor. Chim. Acta* 92 (1995), 277 S 79
- [Oren und Luenberger 1974] OREN, Shmuel S. ; LUENBERGER, David G.: Self-Scaling Variable Metric (SSVM) Algorithms: Part I: Criteria and Sufficient Conditions for Scaling a Class of Algorithms. In: *Man. Sci.* 20 (1974), S. 845–862 57
- [Parr und Yang 1994] PARR, Robert G. ; YANG, Weitao: *Density Functional Theory of Atoms and Molecules*. Oxford University Press, 1994 22
- [Perdew und Wang 1992] PERDEW, J. P. ; WANG, Y.: Accurate and simple analytic representation of the electron-gas correlation energy. In: *Phys. Rev. B* 45 (1992), S. 244–249 28
- [Perić u. a. 2000] PERIĆ, Miljenko ; GREIN, Friedrich ; HACHEY, Michael R.: *Ab initio study of the role of vibronic coupling in the ultraviolet valence/Rydberg spectrum of formaldehyde: Handling of vibronic interaction between three electronic states*. In: *J. Chem. Phys.* 113 (2000), Nr. 20, S. 9011–9021 79
- [Petke u. a. 1992] PETKE, J. D. ; MAGGIORA, G. M. ; CHRISTOFFERSEN, R. E.: Ab Initio Configuration Interaction and Random Phase Approximation Calculations of the Excited Singlet and Triplet States of Uracil and Cytosine. In: *J. Phys. Chem.* 96 (1992), S. 6992–7001 95, 101, 102, 103, 104
- [Peyerimhoff u. a. 1971] PEYERIMHOFF, S.D. ; BUENKER, R.J. ; KAMMER, W.E ; HSU, H.: Calculation of the Electronic Spectrum of Formaldehyde. In: *Chem. Phys. Lett.* 8 (1971), Nr. 1, S. 129–135 85, 92
- [Polak 1971] POLAK, E.: *Computational Methods in Optimisation*. Academic Press, 1971 56

- [Press u. a. 1992] PRESS, William H. ; TEUKOLSKY, Saul A. ; VETTERLING, William T. ; FLANNERY, Brian P.: *Numerical Recipes in FORTRAN*. Cambridge University Press, 1992 45
- [Pulay 1969] PULAY, P.: In: *Mol. Phys.* 17 (1969), 197 S 50
- [Pulay 1982] PULAY, P.: Improved SCF convergence acceleration. In: *J. Comput. Chem.* 3 (1982), S. 556 60
- [Raynes 1966] RAYNES, W.T.: Rotational Analysis of Some Bands of the Triplet←Singlet Transition in Formaldehyde. In: *J. Chem. Phys.* 44 (1966), Nr. 7, S. 2755–2777 79, 86, 87, 89, 90, 92
- [Schlick 1992] SCHLICK, Tamar: *Reviews in Computational Chemistry*. Bd. III. Kap. Optimization Methods in Computational Chemistry, S. 1–71. New York : VCH Publishers, Inc., 1992 74
- [Shah und Moule 1978] SHAH, A.K. ; MOULE, D.C.: In: *Spectrochim. Acta* 34A (1978), 749 S 87
- [Shanno 1970] SHANNO, D. F.: Conditioning of Quasi-Newton Methods for Function Minimization. In: *Mathematics of Computing* 24 (1970), S. 647–656 57
- [Sobolewski u. a. 2002] SOBOLEWSKI, A. L. ; DOMCKE, W. ; DEDONDER-LARDEUX, C. ; JOUVET, C.: In: *Phys. Chem. Chem. Phys.* (2002). – published as an Advance Article on the web 28th February 96
- [Stanton u. a. 1995] STANTON, John F. ; GAUSS, Jürgen ; ISHIKAWA, Naoto ; HEAD-GORDON, Martin: A comparison of single reference methods for characterizing stationary points of excited state potential energy surfaces. In: *J. Chem. Phys.* 103 (1995), S. 4160–4174 87
- [Steward und Jensen 1967] STEWARD, R. F. ; JENSEN, L. H.: In: *Acta Crystallogr.* 23 (1967), S. 1102 98
- [Szabo und Ostlund 1996] SZABO, Attila ; OSTLUND, Neil S.: *Modern Quantum Chemistry*. Dover Publications, Inc., 1996 13, 21, 32, 33
- [Szczesniak u. a. 1983] SZCZESNIAK, M. ; NOVAK, M. J. ; ROSTKOWSKA, H. ; SZCZESNIAK, K.: In: *J. Am. Chem. Soc.* 105 (1983), 5969 S 107
- [Takagi und Oka 1963] TAKAGI, K. ; OKA, T.: In: *J. Phys. Soc. Jpn.* 18 (1963), 1147 S 79, 86, 90
- [Taylor u. a. 1982] TAYLOR, Steven ; WILDEN, David G. ; COMER, John: Electron Energy Loss Spectroscopy of Forbidden Transitions to Valence and Rydberg States of Formaldehyde. In: *Chem. Phys.* 70 (1982), S. 291–298 79, 89, 92
- [Tsuchiya u. a. 1988] TSUCHIYA, Y. ; TAMURA, T. ; FUJII, M. ; ITO, M.: Keto-Enol Tautomer of Uracil and Thymine. In: *J. Phys. Chem.* 92 (1988), S. 1760–1765 95

- [Vahtras u. a. 1993] VAHTRAS, O. ; ALMLÖF, J. ; FEYEREISEN, M. W.: Integral approximations for LCAO-SCF calculations. In: *Chem. Phys. Lett.* 213 (1993), S. 514–518
30
- [Voet u. a. 1963] VOET, D. R. ; GRATZER, W. B. ; COX, R. A. ; DOTY, P.: In: *Biopolymers* 1 (1963), 193 S 95, 101, 103
- [von Weizsäcker 1935] VON WEIZSÄCKER, C. F.: Zur Theorie der Kernmassen. In: *Z. Physik* 96 (1935), S. 431–458 27
- [Vosko u. a. 1980] VOSKO, S. J. ; WILK, L. ; NUSAIR, M.: Accurate spin-dependent electron liquid correlation energies for local spin density calculations: A critical analysis. In: *Can. J. Phys.* 58 (1980), S. 1200–1211 28
- [Waletzke 2001] WALETZKE, Mirko: *Implementation und Anwendung von Multireferenz-DFT und -ab initio Methoden für große Moleküle*. IPCTC Wegelerstrasse 12 : Dissertation, Universität Bonn, 2001 33
- [Whitten und Hackmeyer 1969] WHITTEN, J.L. ; HACKMEYER, M.: In: *J. Chem. Phys.* 51 (1969), 5584 S 79
- [Wigner 1934] WIGNER, E. P.: On the Interaction of Electrons in Metals. In: *Phys. Rev.* 46 (1934), S. 1002–1011 31
- [Williams und Kelley 2000] WILLIAMS, Thomas ; KELLEY, Colin: *gnuplot v. 3.8f*. 2000. – <http://www.gnuplot.org> 74
- [Yeager und McKoy 1974] YEAGER, D.L. ; MCKOY, V.: In: *J. Chem. Phys.* 60 (1974), 2714 S 79

Danksagung

Da diese Arbeit den Endpunkt eines eher ungewöhnlichen Studienweges markiert, ist die Liste derer, die mich mit Worten und Taten unterstützt, aufgebaut und motiviert haben, sehr lang. Auch wenn ich nicht jede Person namentlich nennen kann, so hat doch jede von ihnen zu ihrer Entstehung beigetragen. Danke!

Die beispielhafte und im besten Sinne „mütterlich“ zu nennende Betreuung durch Frau Prof. Dr. Christel M. Marian geht weit über das übliche Maß wissenschaftlicher Professionalität hinaus. Ich danke ihr nicht nur für die unerschöpfliche chemische Kompetenz, sondern auch für die persönliche Basis unserer Zusammenarbeit. Frau Prof. Dr. Sigrid D. Peyerimhoff hat mit viel Wohlwollen zum entscheidenden Zeitpunkt die entscheidende Weichenstellung in meinem Werdegang bewirkt. Das werde ich immer in uneingeschränkt positiver Erinnerung behalten. Prof. Dr. Horst Köppel möchte ich herzlich dafür danken, dass er sich trotz der derzeitigen Doppelbelastung zur Koreferenz bereiterklärt hat.

Die vorliegende Arbeit wäre in dieser Form zum jetzigen Zeitpunkt nicht existent ohne die persönliche und wissenschaftliche Hilfe meines Freundes Dr. Marcus Gastreich. Das Wort „Freundschaftsdienst“ muss wohl neu skaliert werden.

Meiner Mutter danke ich für ihre unendliche Geduld und die eben *nicht* selbstverständliche, umfassende Unterstützung. Wer ausser ihr hätte wohl gewusst, dass „Ameisenkuchen“ ein wirksamer Katalysator für die Beendigung eines Studiums sein kann?