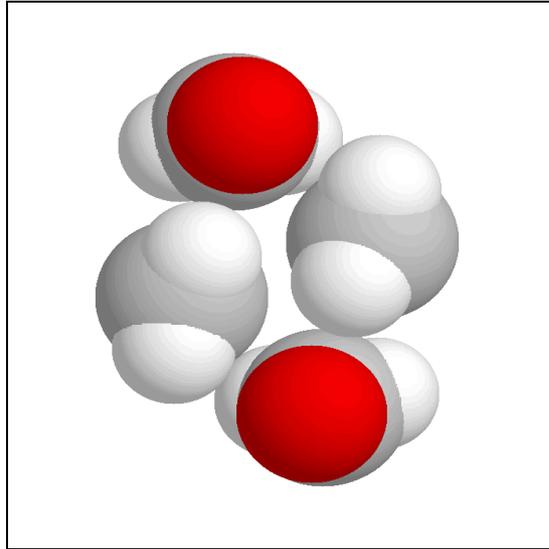


Institut für Numerische
und Instrumentelle Mathematik der
Westfälischen Wilhelms-Universität Münster



Die Gaußtransformation als Homotopiemethode zur globalen
Optimierung von Molekulpotentialfunktionen

vorgelegt von
Marcus Weber
aus Emsdetten

Diplomarbeit bei Herrn Prof. Dr. H. Maurer
November 1998

Inhaltsverzeichnis

1	Einleitung	4
2	Globale Optimierung	6
2.1	Problemformulierung	6
2.2	Begründung für unzuverlässige Algorithmen	7
2.3	Begründung für nichtdeterministische Algorithmen	9
2.3.1	Komplexitätstheorie	9
2.3.2	Multi-Start zur Lösung \mathcal{NP} -harter Probleme	11
2.4	Konvexe Funktionen	12
2.5	Globale Optimierung konvexer Funktionen	13
3	Die Homotopiemethode	16
3.1	Allgemeine Beschreibung	16
3.2	Diffusion als Homotopiemethode	19
3.3	Numerische Realisierung der Homotopiemethode	21
3.4	Ein eindimensionales Beispiel	22
3.5	Die Diffusion als Gaußtransformation	23
3.6	Ein zweidimensionales Beispiel	26
3.7	Konvexe Minoranten und Strukturen	29
4	Distanzgeometrie	35
4.1	Spezielle Funktionstypen	35
4.2	Eine spezielle Distanzgeometriefunktion	37
4.3	Die numerische Umsetzung	39
4.3.1	Die Simplex-Methode	40
4.3.2	Eindeutige kartesische Koordinaten	41

4.4	Der Sensitivitätsschritt	45
5	Molekülpotentialfunktionen	48
5.1	Quantenchemischer Ansatz	48
5.2	Das Coulomb Potential	51
5.3	Die Strafmethode	53
5.4	Transformation des Potentials	55
5.4.1	Numerische Berechnung der Gaußtransformation	55
5.4.2	Approximation der Zielfunktion	60
5.4.3	Approximation des Integrales	62
5.4.4	Approximation der Diffusionsgleichung	64
6	Aggregate	67
6.1	Probleme in der globalen Optimierung	67
6.2	Rechnungen	68
6.2.1	Acrolein und Wasser	68
6.2.2	Formaldehyd	71
6.3	Resultat	73
7	Isolierte Moleküle	76
7.1	Trichlordifluorbutan	76
7.2	Verminderung der Freiheitsgrade	80
7.2.1	United-Atom Modellierung	80
7.2.2	Torsionspotentiale	81
7.3	Vergleich: Homotopiemethode und Multi-Start	83
7.3.1	Effizienz	83
7.3.2	Van-der-Waals Abstoßungen	85
7.3.3	Dichlordifluorpentan	87
7.4	Interpretationen der Optimierungsmethode	90
7.4.1	Die Gaußtransformation als Erwartungswert	90
7.4.2	Krümmung und Diffusion	90
8	Rückschau auf die Diplomarbeit	92

Abbildungsverzeichnis

3.1	Eine zweidimensionale Funktion	18
3.2	Diffusion	18
3.3	Glättung eines Polynoms	22
3.4	Glättung der Griewankfunktion	28
4.1	Glättung einer Distanzgeometriefunktion	38
5.1	Das Lennard-Jones Potential	50
5.2	Numerische Glättung	58
5.3	Gauß-Hermite Integration	64
6.1	Acrolein und Wasser	69
6.2	Acrolein- und Wasser: Ergebnisstruktur	70
6.3	Formaldehyd	72
6.4	Formaldehyd: Optimalstruktur	73
6.5	Wirkungsweise der Homotopiefunktion	74
7.1	1,2,3-Trichlor-1,4-difluorbutan	76
7.2	1,2,3-Trichlor-1,4-difluorbutan: Optimalstruktur	78
7.3	Hexan: vereinfachtes Kraftfeld	81
7.4	Ethan: gestaffelte Konformation	82
7.5	Dichlordifluorpentan: Optimalstruktur	88
7.6	Dichlordifluorpentan: AMBER-Struktur	89

Kapitel 1

Einleitung

Thema meiner Diplomarbeit ist die globale Optimierung von Molekülpotentialfunktionen.

Nach einer kurzen Einführung in die Problematik der globalen Optimierung, in der bereits deutlich wird, daß entsprechende Optimierungsalgorithmen von einer *globalen Struktur* der Zielfunktion Gebrauch machen müssen, werde ich die Homotopiemethode als möglichen Ansatz zur Lösung dieser komplexen Problemklasse erklären.

Schwerpunkt wird die Kombination der Diffusionsmethode (zum Optimieren von sogenannten Lennard-Jones Clustern) mit der Gaußtransformation (als Homotopiemethode zum Lösen von Distanzgeometrie Problemen) sein, die es ermöglicht durch ein quantenchemisch motiviertes, vereinfachtes Kraftfeld, Aggregate und isolierte Moleküle global zu optimieren. Neu wird der verwendete Torsionsterm sein, der die Abstoßung der Bindungselektronendichten repräsentiert. Er liefert selbst im Rahmen der United-Atom Näherung gute Ergebnisse im Vergleich zu denen, die man mit vollständigen Kraftfeldern durch zeitraubendes Ausprobieren erhält.

Im Vordergrund stehen auch heuristische Überlegungen zur chemischen Relevanz der verwendeten Homotopiemethode, sowie die Bewertung mathematischer Approximationsansätze zur Berechnung der Gaußtransformation aus der Sicht der Optimierungsproblematik.

Besonders hinweisen möchte ich schon auf die Optimierung von Formaldehyd, das auch auf der Titelseite zu sehen ist und letztendlich die Stärken der Homotopiemethode demonstrieren wird.

Ein immer wiederkehrendes Problem ist die eindeutige Darstellung von Mo-

lekülen in kartesischen Koordinaten, es tritt in der lokalen und besonders in der globalen Optimierung von Molekülen auf. Da Rotationen und Translationen den Wert der Molekülpotentialfunktionen unverändert lassen, sind bestimmte lokale und globale Optimierungsroutinen auf eine eindeutige Koordinatendarstellung angewiesen. Insbesondere Branch-And-Bound Routinen sind sonst nicht in der Lage, ein globales Minimum effektiv zu bestimmen. In dieser Arbeit wird eine alternative, numerisch besser konditionierte Methode der Transformation in eindeutige Koordinaten vorgestellt, die man zukünftig anstelle der allgemeinen Drehmatrizen verwenden kann. Zudem werden Vorzeichenbeschränkungen erklärt, die die Eindeutigkeit der Koordinatendarstellung gewährleisten und wiederum für die globale Optimierung mittels deterministischer Algorithmen von Bedeutung sind.

Kapitel 2

Globale Optimierung

2.1 Problemformulierung

Das allgemeine globale Optimierungsproblem lautet:

Definition 1 Gegeben sei eine nicht-konvexe Funktion (Zielfunktion) $f : \mathbb{R}^n \rightarrow \mathbb{R}$. Ein globales Optimum von f auf D ist ein Punkt $x^* \in D \subseteq \mathbb{R}^n$, für den gilt:

$$f^* := f(x^*) \leq f(x) \quad \forall x \in D. \quad (2.1)$$

Das globale Optimierungsproblem wird als Minimierungsaufgabe eingeführt, jedoch kann auch die Maximierung einer Funktion f als Minimierung von $-f$ geschrieben werden.

Aus der Formulierung (2.1) geht nicht hervor, ob eine Lösung x^* existiert. Bekannt ist, daß im Falle einer stetigen Zielfunktion f und einer kompakten Menge D die Existenz eines solchen Punktes x^* vorausgesetzt werden kann.

Ein weiteres Problem ist die Eindeutigkeit von x^* . So wird im allgemeinen x^* Element einer Menge $M \subseteq \mathbb{R}^n$ sein, so daß alle $x \in M$ globale Minima von f sind. Eine notwendige Bedingung für globale Optimalität eines Punktes $x^* \in D$ einer stetig differenzierbaren Zielfunktion f ergibt sich aus der Variationsungleichung; vgl. Maurer [22].

Satz 1 (notwendige Bedingung für Optimalität) Sei \bar{x} lokale Minimalstelle des Problems (2.1) und sei f differenzierbar in \bar{x} , dann gilt:

$$\nabla f(\bar{x})^T v \geq 0 \quad \forall v \in T(D, \bar{x}), \quad (2.2)$$

wobei $T(D, \bar{x})$ der Tangentialkegel von D im Punkt \bar{x} bedeutet.

Dabei heißen Punkte, die die Variationsungleichung (2.2) erfüllen, *kritische Punkte*.

Der Optimalwert f^* ist, im Falle der Existenz, eindeutig, so daß sich als hinreichende Bedingung für die globale Optimalität eines Punktes $x^* \in \mathbb{R}^n$ ergibt:

Satz 2 (hinreichende Bedingung für Optimalität) Ist $f : \mathbb{R}^n \rightarrow \mathbb{R}$ durch $f^* \in \mathbb{R}$ nach unten beschränkt, dann gilt für $x^* \in \mathbb{R}^n$

$$f(x^*) = f^* \quad \Rightarrow \quad x^* \text{ globales Optimum von } f.$$

2.2 Begründung für unzuverlässige Algorithmen

Es läßt sich neben Satz 2 bisher keine andere, hinreichende Bedingung für globale Optimalität von x^* in der allgemeinen Formulierung von (2.1) finden. Ist der Optimalwert f^* also unbekannt, dann gibt es kein Testkriterium dafür, daß ein bestimmter Punkt globales Optimum ist. Rinnooy Kan und Timmer [30] schließen daraus, daß das globale Optimierungsproblem, hier für eine m -mal stetig differenzierbare Zielfunktion f und eine offene zulässige Menge $D \neq \{\}$, nicht in endlich vielen Schritten gelöst werden kann.

Das Argument: Sei f eine m -mal stetig differenzierbare Funktion, \bar{x} sei ein Punkt der zulässigen Menge D und B eine Umgebung von \bar{x} , dann gibt es eine Funktion \tilde{f} , so daß $f + \tilde{f}$ m -mal stetig differenzierbar, $\tilde{f}(x) = 0$ für alle $x \notin B$ und \bar{x} ein globales Minimum von $f + \tilde{f}$ ist.

Falls der Satz 2 nicht angewendet werden kann, wird also ein endliches, numerisches Verfahren zum Lösen eines globalen Optimierungsproblems nicht garantieren können, daß ein Ergebnispunkt x^* tatsächlich ein globales Optimum ist, wenn nur punktuelle Eigenschaften der Funktion zur Verfügung stehen. Diese können in einer nicht untersuchten Umgebung eines Punktes \bar{x} so abgeändert sein, daß \bar{x} globales Optimum ist. Um zu einer Lösung zu gelangen, müssen also globale Eigenschaften von f ausgenutzt werden.

Bei dem obigen Argument ist zudem nicht berücksichtigt, daß ein Computer nur endlich viele Dezimalstellen vergleichen kann und auch schon deshalb keine exakte Lösung von (2.1) garantiert werden kann.

Das Problem (2.1) wird deshalb auch oft in einer abgeschwächten Variante formuliert:

Definition 2 Gegeben sei eine Funktion $f : \mathbb{R}^n \rightarrow \mathbb{R}$ mit unbekanntem, optimalem Funktionswert f^* . Ferner sei $\epsilon > 0$ beliebig, aber fest gewählt und $D \subseteq \mathbb{R}^n$.

Gesucht wird ein Punkt $\tilde{x} \in D$ mit

$$f(\tilde{x}) < f^* + \epsilon. \quad (2.3)$$

Je nachdem, ob ein Algorithmus das Problem (2.3), das auch als ϵ -globales Optimierungsproblem bezeichnet wird, lösen kann oder nicht, wird er zuverlässig oder unzuverlässig genannt.

Das obige Argument bedeutet, daß man bei Funktionen, deren globale Eigenschaften unbekannt sind, in jedem Fall auf unzuverlässige, globale Optimierungsalgorithmen zurückgreifen muß.

Bisher konnten für nicht-konvexe Funktionen mit bekannten globalen Eigenschaften nur zuverlässige Optimierungsalgorithmen gefunden werden, deren

Rechenaufwand exponentiell mit der Zahl n der Eingabedaten wächst. Alle höherdimensionalen Optimierungsprobleme sind bei der Lösung daher bisher auch auf unzuverlässige Algorithmen angewiesen.

2.3 Begründung für nichtdeterministische Algorithmen

2.3.1 Komplexitätstheorie

Die Lösung globaler Optimierungsprobleme ist aus numerischer Sicht bisher nur für bestimmte Funktionentypen praktikierbar. Selbst bei diesen Funktionen führt der Aufwand, der mit der numerischen Lösung verbunden ist, bei höherdimensionalen Problemen zu unannehmbaren Laufzeiten. Um zu beurteilen, mit welchem Aufwand die Lösung eines gegebenen Problems verbunden ist, wurde die sogenannte *Komplexitätstheorie* entwickelt, mit deren Hilfe die Probleme in bestimmte *Komplexitätsklassen* eingeteilt werden können, wenn man bei der Laufzeit immer vom ungünstigsten Fall ausgeht. Läßt sich die Komplexität eines Algorithmus durch ein Polynom in n , der Anzahl der Eingabedaten, abschätzen, so heißt der Algorithmus *polynomial*. Die Klasse von Problemen, für deren Lösung es einen polynomialen, deterministischen Algorithmus gibt, wird mit dem Symbol \mathcal{P} versehen. Ein *deterministischer* Algorithmus ist dadurch gekennzeichnet, daß nur die Eingabedaten darüber entscheiden, welche Befehlsreihenfolge in dem Algorithmus angewendet wird.

Dann gibt es Probleme, bei denen die Lösung mit Hilfe eines nichtdeterministischen, polynomialen Algorithmus gefunden werden kann. Die Klasse dieser Probleme erhält das Symbol \mathcal{NP} . Ein *nichtdeterministischer* Algorithmus hat ein weiteres Sprachelement: Gegeben seien n Anweisungen a_1, a_2, \dots, a_n . Dann bedeutet die Anweisung $(a_1|a_2|\dots|a_n)$, daß der Computer nichtdeterministisch entweder die Anweisung a_1 oder die Anweisung a_2 oder ... oder die Anweisung a_n ausführen kann. Ein und dieselbe Eingabe kann zu einer Vielzahl von Berechnungsfolgen führen; vgl. z.B. Goos [9] oder Großmann [11], Seiten 213 bis 229.

Das tatsächliche Auffinden der Lösung von \mathcal{NP} -Problemen (deren Existenz vorausgesetzt werden muß) kann also beliebig komplex sein, weshalb nach Horst [19], S.32, $\mathcal{P} \subseteq \mathcal{NP}$ gilt und $\mathcal{P} \neq \mathcal{NP}$ vermutet wird.

Andererseits ist bei \mathcal{NP} -Problemen der Test, ob ein vorgeschlagenes, erratenes Ergebnis (man spricht von einem *Orakelschritt*) das Problem löst, mit polynomialen Aufwand zu bewältigen.

Definition 3 Ein Problem π_1 ist polynomial transformierbar in ein Problem π_2 (Symbol: $\pi_1 \rightarrow\rightarrow \pi_2$), wenn ein polynomialer, deterministischer Algorithmus zur Lösung von π_2 einen polynomialen, deterministischen Algorithmus zur Lösung von π_1 impliziert.

Definition 4 Ein Problem $\pi_2 \in \mathcal{NP}$ heißt \mathcal{NP} -vollständig, wenn für alle Probleme $\pi_1 \in \mathcal{NP}$ gilt: $\pi_1 \rightarrow\rightarrow \pi_2$.

\mathcal{NP} -vollständige Probleme repräsentieren also die Klasse \mathcal{NP} . Sie sind “mindestens so schwer zu lösen”, wie die Gesamtheit aller \mathcal{NP} -Probleme. Eine Stufe “höher” sind die \mathcal{NP} -harten Probleme anzusiedeln:

Definition 5 Ein Problem π_2 heißt \mathcal{NP} -hart, wenn es ein \mathcal{NP} -vollständiges Problem π_1 gibt, mit $\pi_1 \rightarrow\rightarrow \pi_2$.

Um die Komplexität \mathcal{NP} -harter Probleme zu verstehen, führt man sich nun folgendes vor Augen: Angenommen ein \mathcal{NP} -hartes Problem π_2 wäre durch einen polynomialen, deterministischen Algorithmus lösbar. Weiter sei $\pi_1 \in \mathcal{NP}$ das Problem, welches sich nach Definition 5 polynomial in π_2 transformieren läßt. Dann wäre also, nach Definition 3, π_1 mit Hilfe eines polynomialen, deterministischen Algorithmus lösbar. Weil aber π_1 \mathcal{NP} -vollständig ist, kann, nach Definition 4, auch jedes weitere Problem aus \mathcal{NP} polynomial gelöst werden. Schließlich hätte man $\mathcal{P} = \mathcal{NP}$, was zwar nicht widerlegt, aber unwahrscheinlich ist.

Im Klartext: Die Annahme, ein \mathcal{NP} -hartes Problem lasse sich mit einem

polynomialen, deterministischen Algorithmus lösen, ist wenig erfolgsversprechend.

Im Beispiel der globalen Optimierung von Molekülpotentialfunktionen wird sich ein Unterproblem, ein spezielles Distanzgeometrieproblem, bereits als \mathcal{NP} -hart erweisen. In diesem Fall ist sogar zu vermuten, daß dieses globale Optimierungsproblem nicht einmal zur Menge \mathcal{NP} gehört. Um einen entsprechenden Beweis führen zu können, müßte jedoch ein hinreichendes Optimalitätskriterium für allgemeine Molekülpotentialfunktionen bekannt sein, das nachweislich nicht mit einem polynomialen Algorithmus getestet werden könnte, oder es müßte das Nichtvorhandensein eines solchen prüfbareren Kriteriums gezeigt werden. Beides ist bisher noch nicht gelungen.

2.3.2 Multi-Start zur Lösung \mathcal{NP} -harter Probleme

Es gibt zuverlässige, deterministische, nicht-polynomiale, globale Optimierungsroutinen; vergleiche Horst [19]. Diese sind jedoch bei höherdimensionalen Problemen aufgrund der intensiven Rechenzeit weniger praktikabel, wie der vorangehende Abschnitt gezeigt hat.

Eine Möglichkeit, einen nichtdeterministischen Algorithmus zur Lösung von (2.1) zu formulieren, besteht darin, die Funktion f von verschiedenen, zufällig gewählten Startpunkten aus lokal zu optimieren.

Diese Methode heißt *Multi-Start Optimierung* und ist eine durchaus praktikable und häufig angewandte Art, eine Funktion global zu optimieren.

Es stellt sich jedoch die Frage, ob die Suche nach geeigneten Startpunkten effizienter zu gestalten ist, als die Auswahl dem Zufallszahlengenerator zu überlassen.

Einen entsprechenden Algorithmus zeigten Moré und Wu [23, 24] auf. Die Grundlagen der von ihnen verwendeten Homotopiemethode, die sich nicht nur auf die globale Optimierung von Distanzgeometriefunktionen, sondern auch auf allgemeinere Molekülpotentialfunktionen ausweiten läßt, möchte ich (nach dem Spezialfall der konvexen Funktionen) im nächsten Kapitel behandeln.

2.4 Konvexe Funktionen

Das allgemeine globale Optimierungsproblem (2.1) wurde mit einer nicht-konvexen Funktion formuliert, da konvexe Funktionen auf konvexen Mengen in der globalen Optimierung eine Sonderstellung einnehmen.

Konvexe Mengen $C \subseteq \mathbb{R}^n$ enthalten zu je zwei Punkten $x_1, x_2 \in C$ auch deren Verbindungslinie, also für $\lambda \in [0, 1]$ gilt: $\lambda x_1 + (1 - \lambda)x_2 \in C$.

Konvexe Mengen spielen bei der Optimierung eine entscheidende Rolle. Ihre Eigenschaften sind daher schon früh und umfangreich studiert worden.

Neben den konvexen Mengen haben, gerade in der globalen Optimierung, die *konvexen Funktionen* eine große Bedeutung. Konvexe Funktionen f haben die Eigenschaft, daß eine stückweise lineare Approximation von $f : C \rightarrow \mathbb{R}$ die Funktion stets überschätzt. Also für beliebige Punkte x_1, x_2 aus der konvexen Menge C und $\lambda \in [0, 1]$ gilt:

$$f(\lambda x_1 + (1 - \lambda)x_2) \leq \lambda f(x_1) + (1 - \lambda)f(x_2).$$

Eine Reihe von Stetigkeits- und Differenzierbarkeitseigenschaften kann aus dieser Definition abgeleitet werden. So ist eine konvexe Funktion f z.B. im Inneren ihres Definitionsbereiches stetig.

Ist f zweimal differenzierbar, so ist die Hessematrix $H(x)$ konvexer Funktionen für jedes $x \in C$ positiv semi-definit. Hier gilt auch die Umkehrung: Ist die Hessematrix $H(x)$ für alle $x \in C$ positiv semi-definit, so ist f konvex.

Die Eigenschaften konvexer Funktionen sind für die Optimierung so wichtig, daß man eine Reihe von Definitionen für sogenannte *verallgemeinert konvexe* Funktionen aufgestellt hat. Ziel ist es jeweils, Eigenschaften für Funktionen zu finden, so daß folgender Satz noch Gültigkeit besitzt; aus [19] Seite 14.

Satz 3 (Konvexe Funktionen) Sei $f : \mathbb{R}^n \rightarrow \mathbb{R}$ eine stetig differenzierbare, konvexe Funktion. Ein Punkt $x^* \in D$ einer konvexen Menge D ist genau dann ein globales Optimum der Aufgabe (2.1), wenn er ein kritischer Punkt ist.

2.5 Globale Optimierung konvexer Funktionen

Hier soll kurz skizziert werden, wie man zu einem bekannten, superlinear konvergenten, numerischen Verfahren zur unrestringierten, globalen Optimierung konvexer Funktionen gelangt. Dieses ist ein Beispiel für einen zuverlässigen, deterministischen, globalen Optimierungsalgorithmus, der Bestandteil vieler anderer Algorithmen ist und hier die Grundlage für die vorgestellte Homotopiemethode sein wird.

Sei nun $f : \mathbb{R}^n \rightarrow \mathbb{R}$ eine konvexe, zweimal differenzierbare Funktion mit positiv definiten Hessematrix.

Um f global zu optimieren, reicht es nach Satz 3, eine lokale Optimierung durchzuführen.

Dieses versetzt uns in die Lage eine hinreichende Bedingung für globale Optimalität eines Punktes x^* einer konvexen Funktion zu formulieren, nämlich $\nabla f(x^*) = 0$.

Führen wir die Gradientenfunktion $G : \mathbb{R}^n \rightarrow \mathbb{R}^n$ mit $G(x) := \nabla f(x)$ ein, so geht das globale Optimierungsproblem in ein Gleichungssystem über:

$$G(x^*) = 0, \quad x^* \in \mathbb{R}^n.$$

Numerisch läßt sich dieses mit Hilfe des quadratisch konvergenten Newton-Verfahrens lösen. Die Funktion G wird dabei linear approximiert:

$$0 = G(x^*) \approx G(x) + H(x)(x^* - x),$$

wobei $H(x)$ die Hessematrix von f im Punkte x ist. Durch Umformung erhalten wir:

$$x^* \approx x - H^{-1}(x) G(x),$$

was für quadratische, konvexe Funktionen f mit positiv definiten Hessematrix exakt ist und nun als allgemeiner Ansatz für ein Iterationsverfahren dient.

Ausgehend von einem Startpunkt $x^k \in \mathbb{R}^n$ gelangt man zum nachfolgenden Iterationspunkt durch:

$$x^{k+1} := x^k - H^{-1}(x^k) G(x^k). \tag{2.4}$$

Es hat sich in der Praxis bewährt, die Richtung $d^k := H^{-1}(x^k)G(x^k)$ als Suchrichtung zu interpretieren. Mit Hilfe einer eindimensionalen lokalen Optimierungsroutine wird daher im sogenannten Newton-Verfahren mit variabler Metrik der Punkt x^{k+1} folgendermaßen definiert:

$$x^{k+1} := x^k - \lambda_k d^k,$$

wobei man λ_k so wählt, daß die Funktion $F_k(\lambda) := f(x^k - \lambda d^k)$ in λ_k ein lokales Minimum annimmt. Dieses gewährleistet bei konvexen Zielfunktionen f eine Konvergenz des Verfahrens, da in jedem Iterationsschritt ein sicherer Abstieg erfolgt.

Ein Problem des Newton-Verfahrens tritt bei nicht-konvexen Funktionen auf: Während des Verfahrens können Iterationspunkte berechnet werden, an denen die Hessematrix indefinit oder gar negativ-definit ist.

Deshalb wird in den sogenannten Quasi-Newton-Verfahren vor der Berechnung der Suchrichtung d^k eine Approximation der inversen Hessematrix $H_k \approx H^{-1}(x^k)$ aus den bisher erhaltenen Daten hergeleitet. Dieses läßt sich z.B. im BFGS-Verfahren nach Broyden, Fletcher, Goldfarb und Shanno so durchführen, daß die Symmetrie und positive Definitheit der approximierten, inversen Hessematrix erhalten bleibt und daß eine quadratische, konvexe Funktion $f : \mathbb{R}^n \rightarrow \mathbb{R}$ in spätestens $n + 1$ Iterationsschritten global optimiert wird, da für diesen Funktionstyp das BFGS-Verfahren in ein Verfahren der konjugierten Richtungen übergeht. Zur Theorie der Quasi-Newton-Verfahren siehe [7].

Die Schritte des BFGS-Verfahrens lauten also:

1. Setze einen Startpunkt $x^0 \in \mathbb{R}^n$ fest. Setze $H_0 := E_n$ auf die n -te Einheitsmatrix. $k := 0$
2. Bestimme die Suchrichtung $d^k := H_k G(x^k)$.
3. Führe die eindimensionale Optimierung von $F_k(\lambda)$ mit lokalem Minimum bei λ_k durch.
4. Setze $x^{k+1} := x^k - \lambda_k d^k$ und prüfe ein gegebenes Stoppkriterium (z.B. Norm des Gradienten hinreichend klein).
5. Falls Stoppkriterium nicht erfüllt: Führe ein Update der inversen Hessematrix H_{k+1} durch. Setze $k \rightarrow k + 1$, gehe zu Schritt 2.

Die Update-Formel, bei der eine Korrekturmatrix mit Rang 2 addiert wird, lautet dabei

$$H_{k+1} = H_k + \frac{1}{v^{kT}v^k} \left[\left(1 + \frac{w^{kT}H_k w^k}{v^{kT}w^k} \right) v^k v^{kT} - (H_k w^k v^{kT} + v^k w^{kT} H_k) \right],$$

wobei $v^k := x^{k+1} - x^k$ und $w^k := G(x^{k+1}) - G(x^k)$ ist.

Kapitel 3

Die Homotopiemethode

3.1 Allgemeine Beschreibung

Ausgangsproblem sei wieder das nicht-konvexe, unbeschränkte, globale Optimierungsproblem (2.1) mit einer stetigen Zielfunktion $f : \mathbb{R}^n \rightarrow \mathbb{R}$. Es wird im folgenden stets die Existenz einer Lösung $x^* \in \mathbb{R}^n$ vorausgesetzt.

Definition 6 Eine Funktion $F : \mathbb{R}^n \times [0, T] \rightarrow \mathbb{R}$ heißt Homotopiefunktion zu der stetigen Zielfunktion $f : \mathbb{R}^n \rightarrow \mathbb{R}$, wenn gilt:

(1) F ist stetig im ersten Argument.

(2) Es gilt für alle $x \in \mathbb{R}^n$ die Stetigkeit in 0 im zweiten Argument, also

$$\lim_{t \rightarrow 0} F(x, t) = f(x). \quad (3.1)$$

Jede dieser Funktionen F impliziert in Abhängigkeit von dem Scharparameter t ein eigenes, globales Optimierungsproblem:

$$\mathbf{HP}(t) : \text{globmin}_{x \in \mathbb{R}^n} F(x, t). \quad (3.2)$$

Satz 4 (Homotopiemethode) Sei nun mit obigen Bezeichnungen $x = x(t)$ für $t \in]0, T]$ eine Lösung von (3.2) und gelte

$$\lim_{t \rightarrow 0} x(t) = x^0,$$

dann ist x^0 eine Lösung des Ausgangsproblems (2.1).

Beweis: Da $x(t)$ ein globales Minimum von $F(\cdot, t)$ ist, gilt für $t \in]0, T]$

$$F(x(t), t) \leq F(y, t) \quad \forall y \in \mathbb{R}^n,$$

so daß

$$\lim_{t \rightarrow 0} F(x(t), t) \leq \lim_{t \rightarrow 0} F(y, t) \quad \forall y \in \mathbb{R}^n.$$

Aufgrund der Stetigkeit von F , der Bedingung des Satzes und der Stetigkeit der Einbettung (3.1) folgt damit

$$f(x^0) \leq f(y) \quad \forall y \in \mathbb{R}^n$$

und somit die Optimalität von x^0 . \square

Die Frage ist nun, was man mit diesem Ansatz erreicht hat. Das globale Optimierungsproblem (2.1) geht nämlich in eine Schar globaler Optimierungsprobleme (3.2) über. Dadurch scheint das Ausgangsproblem nicht vereinfacht worden zu sein.

Nimmt man jedoch an, der Zielfunktion f läge eine einfache, globale Struktur zugrunde, z.B. eine konvexe Funktion f_g , die durch eine Rauschfunktion f_r gestört wird (ein Beispiel dazu ist die Griewankfunktion, Seite 27), dann liegt die Vermutung nahe, daß die Lage des globalen Optimums von $f = f_g + f_r$ maßgeblich durch die Funktion f_g bestimmt wird. Eine mögliche Homotopie, mit $t \in [0, 1]$, wäre dann:

$$F(x, t) := f_g(x) + (1 - t)f_r(x). \quad (3.3)$$

Die Homotopiemethode versucht also, eine vorhandene, globale Struktur der Zielfunktion f zu bestimmen, die u.U. weniger Minima aufweist und von

daher einfacher (z.B. durch einen Multi-Start Algorithmus) global optimiert werden kann. Ideal geeignet sind also Funktionen, die durch eine konvexe, globale Struktur beschrieben werden können, wie in (3.3).

Wie gelangt man nun zu der globalen Struktur einer Funktion?

Bild 3.1 zeigt den Ausschnitt aus dem Graphen einer zweidimensionalen Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, wobei der Wert der Funktion f durch unterschiedlich dicht angeordnete Punkte dargestellt wird.

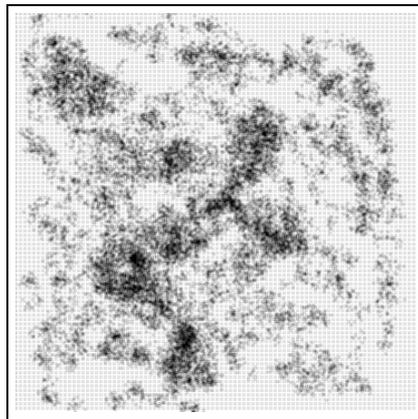


Abbildung 3.1: Eine zweidimensionale Funktion

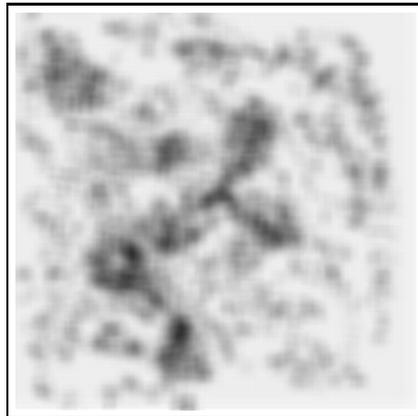


Abbildung 3.2: Diffusion

Zu einer globalen Struktur der Funktion gelangt man, indem man eine Diffusion durchführt. Illustriert wird dieses durch den Übergang von Bild 3.1 zu Bild 3.2. Diese Methode wurde von Scheraga [20] entwickelt und nun als eine mögliche Homotopie eingeführt.

3.2 Diffusion als Homotopiemethode

Diffusionsvorgänge werden durch das zweite Ficksche Gesetz [1] beschrieben. Gegeben sei eine zeitabhängige Funktion $F : \mathbb{R}^n \times [0, \infty[\rightarrow \mathbb{R}$, die in einem Raum jedem Punkt $x \in \mathbb{R}^n$ zum Zeitpunkt $t \in [0, \infty[$ die meßbare Eigenschaft (z.B. Temperatur) $F(x, t)$ zuordnet. Ausgehend von einem stetigen Zustand $f(x)$ wird die Diffusion durch folgendes Anfangswertproblem beschrieben:

$$\frac{\partial F}{\partial t} = \Delta F, \quad F(x, 0) = f(x), \quad (3.4)$$

wobei der Laplace-Operator Δ folgender Differentialoperator ist:

$$\Delta := \frac{\partial^2}{\partial x_1^2} + \dots + \frac{\partial^2}{\partial x_n^2}.$$

Das Anfangswertproblem (3.4) ist vernünftig gestellt, da die Charakteristiken der Diffusionsgleichung die Hyperebenen $H \in \{(x, t) \in \mathbb{R}^{n+1}, t = \text{const}\}$ sind.

Eine Grundlösung der Diffusionsgleichung lautet:

$$\gamma_n(x, t) = \frac{e^{-\|x\|^2/4t}}{(4\pi t)^{n/2}}.$$

Mit dieser lassen sich folgende Sätze [26] über die Existenz, Form und Eindeutigkeit der Lösung des Anfangswertproblems (3.4) formulieren. Bezeichne $C(\mathbb{R}^n)$ die Menge der stetigen und $C^2(\mathbb{R}^n)$ die Menge der auf \mathbb{R}^n zweimal stetig differenzierbaren Funktionen.

Satz 5 (Lösung der Diffusionsgleichung) Sei $f \in C(\mathbb{R}^n)$, und es gelte mit Konstanten $A, M > 0$

$$|f(x)| \leq M e^{A\|x\|^2}.$$

Dann ist für $T < \frac{1}{4A}$

$$F(x, t) = \int_{\mathbb{R}^n} \gamma_n(x - y, t) f(y) dy \quad (3.5)$$

eine Lösung $F \in C^2(\mathbb{R}^n \times (0, T)) \cap C(\mathbb{R}^n \times [0, T])$ des Anfangswertproblems (3.4) und es gibt von x unabhängige Konstanten $A_1(t), M_1(t)$ mit:

$$|F(x, t)| \leq M_1(t) e^{A_1(t)\|x\|^2}.$$

Satz 6 (Eindeutigkeit der Lösung) Das Anfangswertproblem (3.4) ist in der Klasse der Funktionen $F \in C^2(\mathbb{R}^n \times (0, T)) \cap C(\mathbb{R}^n \times [0, T])$, welche einer Abschätzung der Form

$$|F(x, t)| \leq M e^{A\|x\|^2}, \quad x \in \mathbb{R}^n, \quad t \in [0, T]$$

mit $A, M > 0$ genügen, eindeutig lösbar.

Die Lösung (3.5) erfüllt die geforderten Stetigkeitseigenschaften von (3.1). Daher läßt sich die Diffusion als Homotopiemethode auffassen, auf die der Satz 4 anwendbar ist. Wie die Homotopiemethode dabei numerisch angegangen wird, zeigt der nächste Abschnitt.

3.3 Numerische Realisierung der Homotopiemethode

Zunächst fordert der Satz 4 einen stetigen Übergang $t \rightarrow 0$, dieses wird numerisch durch eine Diskretisierung ersetzt, bei der man t schrittweise von T nach 0 konvergieren läßt.

Man bestimmt dazu eine Menge von Zeitpunkten t_i , $i = 0, \dots, m$

$$0 = t_0 < \dots < t_m = T.$$

für die man jeweils die Funktionen $F(x, t_i)$ mit Hilfe der Lösungsformel (3.5) berechnet.

Es kann dabei nicht vorausgesetzt werden, daß $F(\cdot, T)$ konvex ist. Jedoch wird die Zahl der lokalen Minima für größer werdendes t_i abnehmen, so daß man die globale Optimierung von $F(\cdot, T)$, die in Satz 4 gefordert wird, gut durch einen Multi-Start Ansatz realisieren kann. Dieses Vorgehen bewirkt, daß das numerische Verfahren nichtdeterministisch wird.

Satz 4 fordert weiter, daß $F(\cdot, t)$ für $0 < t < T$ ebenfalls global optimiert wird. Nimmt man jedoch an, daß das globale Minimum x^{i+1} von $F(\cdot, t_{i+1})$ ungefähr mit einem globalen Minimum von $F(\cdot, t_i)$ übereinstimmt, dann reicht es, $F(\cdot, t_i)$ ausgehend von x^{i+1} lokal zu minimieren. Dadurch wird der Algorithmus unzuverlässig, denn diese Annahme ist nicht immer zutreffend.

Der vorgestellte Optimierungsalgorithmus besteht also aus drei Phasen, einer Diffusionsphase, einer sukzessiven lokalen Optimierung der geglätteten Funktionen ausgehend von den lokalen Minima der Diffusionsphase und einer lokalen Optimierung der Zielfunktion ausgehend von dem erhaltenen Ergebnispunkt der ersten beiden Phasen.

Der genaue Algorithmus wird später in Abschnitt 4.3 wiedergegeben. Zu bemerken ist nur noch, daß der Algorithmus zwar lokale Optimierungen aufgrund punktueller Informationen, jedoch in gewisser Weise auch globale Eigenschaften durch die analytische Berechnung der Homotopiefunktionen enthält, was das Argument auf Seite 7 für die Unzuverlässigkeit dieser Methode relativiert.

3.4 Ein eindimensionales Beispiel

Die Arbeitsweise des Algorithmus soll an einer eindimensionalen Funktion demonstriert werden.

Als Zielfunktion dient hier folgendes Polynom $f : \mathbb{R} \rightarrow \mathbb{R}$:

$$f(x) = 8x^4 - \frac{44}{3}x^3 + 2x^2 + \frac{11}{3}x + 1.$$

Mit Hilfe von (3.5) erhält man als Homotopiefunktion:

$$F(x, t) = f(x) + 96t^2 + (4 - 88x + 96x^2)t.$$

F ist für alle $x, t \in \mathbb{R}$ eine Lösung der Anfangswertaufgabe (3.4).

Die zweite Ableitung der Homotopiefunktion nach dem ersten Argument x lautet $F''(x, t) = 4 - 88x + 96x^2 + 192t$ und ist für $t \geq 97/1152 \approx 0.0842$ unabhängig von x positiv. Das bedeutet, daß $F(\cdot, t)$ für $t \geq 97/1152$ konvex ist.

Wählen wir also für unseren Algorithmus die vier Zeitpunkte

$$t_0 = 0, t_1 = 0.05, t_2 = 0.08, t_3 = 0.12,$$

so erhalten wir für F die Funktionsgraphen in Abbildung 3.3.

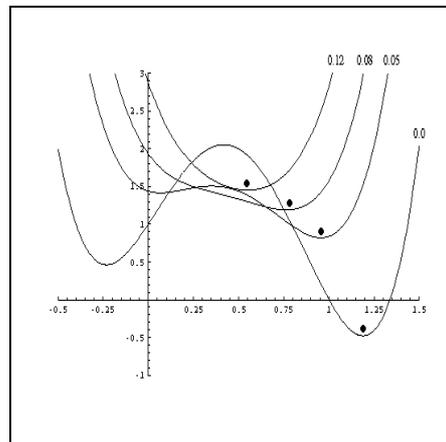


Abbildung 3.3: Glättung eines Polynoms

$F(\cdot, t_3)$ ist konvex und hat ein eindeutiges Minimum x^3 . Minimiert man nun für $i = 2, 1, 0$ sukzessive die Funktionen $F(\cdot, t_i)$ vom Minimum x^{i+1} der Funktion $F(\cdot, t_{i+1})$ aus, so gelangt man tatsächlich zum globalen Minimum der Zielfunktion f .

Diese Vorgehensweise wird in der Abbildung 3.3 durch kleine Punkte dargestellt, die jeweils das Minimum angeben, das im Laufe der Homotopiemethode von dem Algorithmus aufgesucht wird.

Bei allgemeinen Zielfunktionen kann man nicht davon ausgehen, daß bei einem bestimmten Wert von t eine konvexe Homotopiefunktion $F(\cdot, t)$ vorliegt, weshalb mehrere Durchläufe des Verfahrens nötig werden.

Dieses erklärt, warum der Homotopieansatz zur Lösung globaler Optimierungsprobleme als Multi-Start Routine formuliert werden muß.

3.5 Die Diffusion als Gaußtransformation

Die Lösung der Diffusionsgleichung ist nicht in jedem Fall eindeutig bestimmt.

Der Integralausdruck (3.5), der eine mögliche Lösung der Diffusionsgleichung darstellt, geht durch eine Substitution in eine bekannte Transformationsformel über. Setzt man $4t = \lambda^2$, ergibt sich:

$$[f]_\lambda(x) := F(x, \lambda) = \frac{1}{\pi^{n/2} \lambda^n} \int_{\mathbb{R}^n} f(y) e^{-\frac{\|y-x\|^2}{\lambda^2}} dy. \quad (3.6)$$

Und eine weitere Substitution $v = (y - x)/\lambda$ vereinfacht diesen Ausdruck zu:

$$[f]_\lambda(x) = \frac{1}{\pi^{n/2}} \int_{\mathbb{R}^n} f(x + \lambda v) e^{-\|v\|^2} dv. \quad (3.7)$$

Die Funktion $[f]_\lambda$ ist dabei die sogenannte *Gaußtransformation* von f zum Parameter λ . Die Funktionswerte $[f]_\lambda(x)$ sind dabei mit der Normalverteilung $N(0, \lambda^2/4)$ gewichtete Mittelwerte von f zentriert um den Punkt x .

Anmerkung: Es lassen sich verallgemeinernd auch andere Verteilungsfunktionen wählen, so daß neben der Diffusion noch andere Homotopiemethoden darstellbar sind.

Moré und Wu [23, 24] haben die Eigenschaften der Gaußtransformation im Sinne der globalen Optimierung untersucht. Auch haben sie den Nachweis für die Homotopieeigenschaft (3.1) gebracht. Der nächste Satz formuliert das

Resultat von Moré und Wu [23] jedoch entsprechend der Abschätzung aus Satz 5. Die Modifikation des Beweises, der diese Abschätzung betrifft, geht in (3.8):

Satz 7 (Gaußtransformation als Homotopiemethode) *Erfüllt eine stetige Funktion $f : \mathbb{R}^m \rightarrow \mathbb{R}$ die Bedingungen aus Satz 5, und konvergiert $(x^k)_{k \in \mathbb{N}}$ gegen x^* , sowie $(\lambda_k)_{k \in \mathbb{N}}$ gegen 0, so gilt:*

$$\lim_{k \rightarrow \infty} [f]_{\lambda_k}(x^k) = f(x^*).$$

Beweis: Wegen der Konvergenz der Folgen $(x^k)_{k \in \mathbb{N}}$ und $(\lambda_k)_{k \in \mathbb{N}}$ kann man $k_0 \in \mathbb{N}$ so wählen, daß für alle $k \geq k_0$ mit Konstanten $B, C > 0$ und $A, M > 0$ aus der Abschätzungsformel gilt:

- a) $0 \leq \lambda_k \leq 1$,
- b) $M \exp(A\|x^k\|^2) \leq B$,
- c) $2\|x^k\| + 1 \leq C/A$,
- d) $C\lambda_k \leq 0.5$.

Dieses führt mit Hilfe der Bedingungen aus Satz 5 und für $f^* := |f(x^*)|$, $\|u\| \geq 1$ und $\mu := B + f^*$ zu folgender Abschätzung:

$$\begin{aligned} |f(x^k + \lambda_k u) - f(x^*)| &\leq |f(x^k + \lambda_k u)| + f^* \\ &\leq M \exp(A\|x^k + \lambda_k u\|^2) + f^* \\ &\leq M \exp(A(\|x^k\| + \lambda_k \|u\|)^2) + f^* \\ &= M \exp(A(\|x^k\|^2 + 2\lambda_k \|x^k\| \|u\| + \lambda_k^2 \|u\|^2)) + f^* \\ &\leq M \exp(A(\|x^k\|^2 + \lambda_k (2\|x^k\| + 1) \|u\|^2)) + f^* \end{aligned}$$

$$\begin{aligned}
&= M \exp(A\|x^k\|^2) \exp(A(2\|x^k\| + 1)\lambda_k \|u\|^2) + f^* \\
&\leq B \exp(C\lambda_k \|u\|^2) + f^* \\
&\leq \mu \exp(0.5 \|u\|^2)
\end{aligned} \tag{3.8}$$

Gegeben sei die Menge

$$C_r := \{x \in \mathbb{R}^m : \|x\| > r\}.$$

Es gilt also wegen (3.8) für $r > 1$:

$$\int_{C_r} |f(x^k + \lambda_k u) - f(x^*)| \exp(-\|u\|^2) du \leq \mu \int_{C_r} \exp(-0.5\|u\|^2) du$$

Da das Integral auf der rechten Seite für genügend großes r beliebig klein werden kann, gibt es für jedes $\epsilon > 0$ ein $r > 0$, so daß gilt:

$$\int_{C_r} |f(x^k + \lambda_k u) - f(x^*)| \exp(-\|u\|^2) du \leq \epsilon, \quad k \geq k_0 \tag{3.9}$$

Wegen der Stetigkeit von f in x^* und der Konvergenz der Folgen $(x^k)_{k \in \mathbb{N}}$ und $(\lambda_k)_{k \in \mathbb{N}}$, sowie der Beschränktheit der Menge $\mathbb{R}^n - C_r$ gibt es zu jedem r , k_0 und $\epsilon > 0$, ein $k_1 \geq k_0$, so daß

$$\int_{\mathbb{R}^n - C_r} |f(x^k + \lambda_k u) - f(x^*)| \exp(-\|u\|^2) du \leq \epsilon, \quad k \geq k_1.$$

Addition dieser Ungleichung mit (3.9) liefert für $k \geq k_1$:

$$\begin{aligned}
\frac{2}{\pi^{n/2}} \epsilon &\geq \frac{1}{\pi^{n/2}} \int_{\mathbb{R}^n} |(f(x^k + \lambda_k u) - f(x^*)) \exp(-\|u\|^2)| du \\
&\geq \left| \frac{1}{\pi^{n/2}} \int_{\mathbb{R}^n} (f(x^k + \lambda_k u) - f(x^*)) \exp(-\|u\|^2) du \right| \\
&= |[f]_{\lambda_k}(x^k) - f(x^*)|
\end{aligned} \tag{3.10}$$

Dieses zeigt die geforderte Konvergenzeigenschaft. \square

Die Gaußtransformation ist zudem als Operator linear und isoton, d.h. für Funktionen $f_1, f_2 : \mathbb{R}^n \rightarrow \mathbb{R}$, und einem Skalar $\alpha \in \mathbb{R}$ gelten folgende Aussagen:

1. $[\alpha f]_\lambda = \alpha[f]_\lambda, \quad [f_1 + f_2]_\lambda = [f_1]_\lambda + [f_2]_\lambda$
2. $f_1 \leq f_2 \Rightarrow [f_1]_\lambda \leq [f_2]_\lambda$

Aus der Darstellung (3.7), der Isotonie der Gausstransformation und der Tatsache, daß Konvexität von $f : \mathbb{R}^n \rightarrow \mathbb{R}$ für $x_1, x_2 \in \mathbb{R}^n$ und $\alpha \in [0, 1]$

$$f(\alpha x_1 + (1 - \alpha)x_2 + \lambda u) \leq \alpha f(x_1 + \lambda u) + (1 - \alpha)f(x_2 + \lambda u)$$

impliziert, folgt direkt, daß konvexe Funktionen f durch Gaußtransformation wieder in konvexe Funktionen $[f]_\lambda$ übergehen.

Dieses bedeutet, daß ein konvexes, globales Optimierungsproblem durch Diffusion als Homotopiemethode nicht komplizierter wird.

Die Isotonie der Gaußtransformation garantiert auch, daß eine durch $\mu \in \mathbb{R}$ nach unten beschränkte Funktion $f : \mathbb{R}^n \rightarrow \mathbb{R}, f(x) \geq \mu$ wieder in eine nach unten beschränkte Funktion $[f]_\lambda : \mathbb{R}^n \rightarrow \mathbb{R}, [f]_\lambda(x) \geq \mu$ übergeht.

Bemerkung: Die Gaußtransformation überführt eine stetige Ausgangsfunktion in eine Schar von zweimal stetig differenzierbaren Funktionen. Sie kann deshalb nicht nur zur globalen Glättung einer Zielfunktion, sondern auch zur lokalen Minimierung nichtglatter Funktionen im Sinne der parametrischen Einbettung in eine Schar lokal geglätteter Funktionen verstanden werden.

3.6 Ein zweidimensionales Beispiel

Wu [24] hat die analytische Form der Gaußtransformation von sogenannten *zerlegbaren* Funktionen untersucht.

Dieses sind höherdimensionale Funktionen, für die man die Gaußtransformation auf den eindimensionalen Fall zurückführen kann.

Definition 7 Eine Funktion $f : \mathbb{R}^n \rightarrow \mathbb{R}$ heißt zerlegbar, wenn f geschrieben werden kann als:

$$f(x) = \sum_{k=1}^m f_k(x), \quad f_k(x) = \prod_{j=1}^n f_{k,j}(x_j),$$

wobei $f_{k,j} : \mathbb{R} \rightarrow \mathbb{R}$.

Die Gaußtransformation zerlegbarer Funktionen läßt sich durch die Transformation eindimensionaler Funktionen ausdrücken:

$$[f]_\lambda(x) = \sum_{k=1}^m \left(\prod_{j=1}^n [f_{k,j}]_\lambda(x_j) \right).$$

Zusammen mit der Linearität der Gaußtransformationen und einiger elementarer Funktionen lassen sich so viele globale Optimierungsbeispiele rechnen. Der entsprechende Algorithmus ist bereits im Abschnitt 3.3 erläutert worden. Einzige Änderung ist, daß man statt des Parameters t , der die Zeit in der Diffusionsgleichung darstellte, den Glättungsparameter λ verwendet.

Eine wichtige elementare Funktion ist $f(x) = x^k$, für die man erhält:

$$[f]_\lambda(x) = \sum_{l=0}^{\lfloor k/2 \rfloor} \left(\frac{k!}{(k-2l)! l!} \right) \left(\frac{\lambda}{2} \right)^{2l} x^{k-2l}. \quad (3.11)$$

Als Beispiel möchte ich hierzu die Optimierung der zweidimensionalen Griewankfunktion [10] aufgreifen. Diese besitzt eine konvexe, globale Struktur und wird durch eine periodische "Rauschfunktion" gestört. Die Form entspricht also (3.3) und ist als Optimierungsbeispiel deshalb gut geeignet.

Die Griewankfunktion $f : \mathbb{R}^n \rightarrow \mathbb{R}$ lautet:

$$f(x) = 1 + \sum_{i=1}^n \left(\frac{x_i^2}{200} \right) - \prod_{i=1}^n \cos \left(\frac{x_i}{\sqrt{i}} \right).$$

Als zerlegbare Funktion mit $[\cos]_\lambda(x) = \cos(x) \exp(-0.25\lambda^2)$ ergibt sich für die Gaußtransformierte zusammen mit (3.11):

$$[f]_\lambda(x) = 1 + \sum_{i=1}^n \left(\frac{x_i^2}{200} + \frac{\lambda^2}{400} \right) - \prod_{i=1}^n \exp \left(-\frac{\lambda^2}{4i} \right) \cos \left(\frac{x_i}{\sqrt{i}} \right),$$

Folgende Bilder zeigen den Funktionsgraphen von $[f]_\lambda$ für $\lambda = 0, 1, 2, 4$:

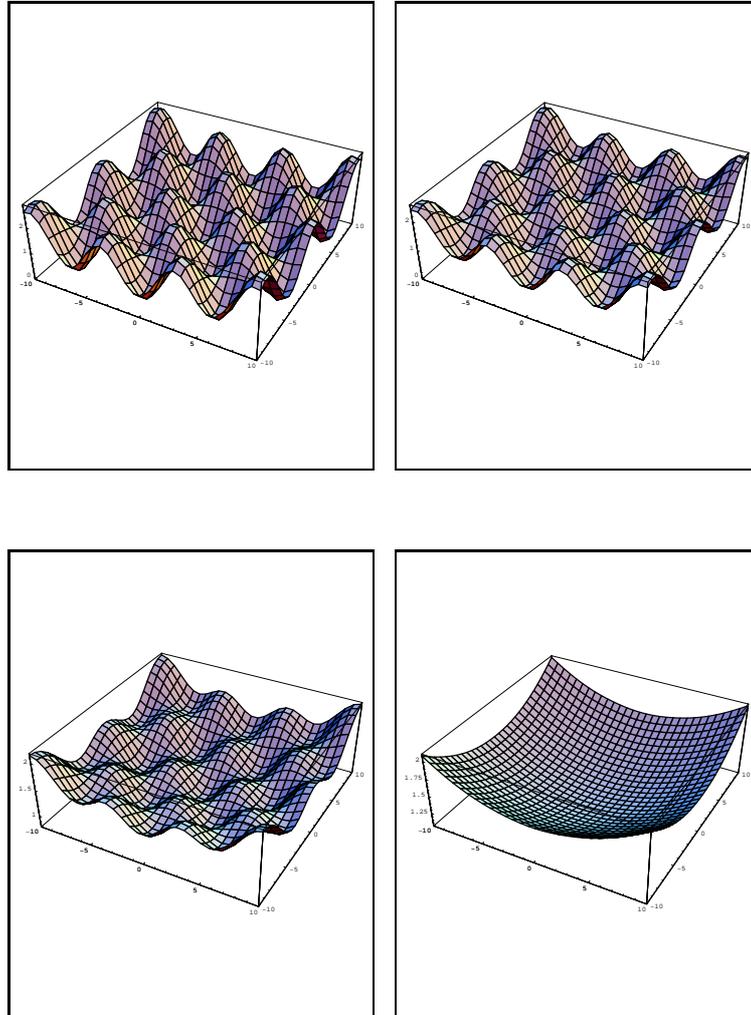


Abbildung 3.4: Glättung der Griewankfunktion

Die Glättung konvergiert für wachsendes λ gegen die konvexe, globale Struktur der Griewankfunktion.

Eine Berechnung der Hessematrix zeigt, daß die geglättete Funktion bereits bei $\lambda \geq 4$ konvex ist.

Interessanterweise hat jede der Funktionen $[f]_\lambda$ ein globales Minimum in $(0, 0)^T$. Der Punkt $(0, 0)^T$ ist zudem einziges lokales Minimum von $[f]_4$, so daß im Homotopiealgorithmus schon im ersten Schritt die Lösung des Problems gefunden wird.

3.7 Konvexe Minoranten und Strukturen

Die grundlegende Idee vieler globaler Optimierungsverfahren ist die Konstruktion von konvexen Minoranten.

Definition 8 Eine Funktion $Q : \mathbb{R}^n \rightarrow \mathbb{R}$ heißt konvexe Minorante der stetigen Zielfunktion $f : \mathbb{R}^n \rightarrow \mathbb{R}$, wenn gilt:

(1) Q ist konvex.

(2) Es gilt für alle $x \in \mathbb{R}^n$ die Ungleichung $Q(x) \leq f(x)$.

Gilt zusätzlich für jede weitere konvexe Minorante q von f , daß $Q(x) \geq q(x)$ für alle $x \in \mathbb{R}^n$, so ist Q die beste konvexe Minorante oder einfach beste Minorante von f .

Eine Besonderheit der besten Minorante, die aufgrund der universellen Eigenschaft der Definition stets eindeutig bestimmt ist, wurde bereits in Satz 3 beschrieben, der eine einfache Methode zur globalen Optimierung konvexer Funktionen lieferte.

Ein weiterer Satz (Theorem 1.20 aus [19]) zeigt eine schöne Besonderheit der besten Minorante auf, die dadurch zu einem wichtigen Werkzeug der globalen Optimierung geworden ist.

Satz 8 (Beste Minorante) *Betrachte das Problem*

$$\text{global } \min_{x \in S} f(x)$$

mit einer konvexen, kompakten Menge $S \subset \mathbb{R}^n$. Sei $Q(x)$ die beste Minorante von f . Dann gilt für den Optimalwert:

$$f^* := \min\{f(x) : x \in S\} = \min\{Q(x) : x \in S\}$$

und

$$\{y \in S : f(y) = f^*\} \subseteq \{y \in S : Q(y) = f^*\}.$$

Der Beweis dieser Aussage stammt ebenfalls aus [19]. Ich führe ihn hier auf, da er auch für die Theorie der Strukturen gilt.

Beweis: Die Übereinstimmung der Optimalwerte von Q und f ergibt sich aus der Bedingung $Q(x) \leq f(x)$ und der Tatsache, daß die konstante Funktion $h(x) := f^*$ ebenfalls eine konvexe Minorante von f ist.

Den zweiten Teil sieht man durch einen Widerspruchsbeweis ein: Angenommen x^* wäre ein globales Minimum von f auf S , jedoch nicht von Q . Sei weiter y^* ein globales Minimum von Q , dann gilt:

$$Q(y^*) < Q(x^*) \leq f^*.$$

Dieses widerspricht jedoch der Aussage $Q(y^*) = f^*$. \square

Das nicht-konvexe, globale Optimierungsproblem (2.1) geht also durch Einführung der besten Minoranten in ein konvexes Optimierungsproblem über. Dieses Optimierungsproblem kann gemäß Satz 3 auf einfache Weise gelöst werden. Ist Q zudem streng konvex, dann besitzt es ein eindeutiges Minimum, das wegen des obigen Satzes mit dem globalen Minimum von f übereinstimmt.

Die Schwierigkeit besteht nun aber darin, die beste Minorante eines Optimierungsproblems ausfindig zu machen. Dies ist aber nicht einfacher, als das

Ausgangsproblem zu lösen, wie der Abschnitt über die Komplexität bereits vermuten läßt.

Hier soll nun die Bedeutung der Gaußtransformation zur Bestimmung der globalen Struktur einer Funktion erläutert werden. Schränkt man die Definition der konvexen Minorante nämlich auf bestimmte konvexe Funktionen ein, so gelangt man zu dem Begriff der Struktur einer Funktion. Zum Beispiel soll die quadratische Struktur folgendermaßen definiert sein:

Definition 9 Eine Funktion $Q : \mathbb{R}^n \rightarrow \mathbb{R}$ heißt quadratische, konvexe Minorante der stetigen Zielfunktion $f : \mathbb{R}^n \rightarrow \mathbb{R}$, wenn gilt:

- (1) Q ist konvex und quadratisch.
- (2) Es gilt für alle $x \in \mathbb{R}^n$ die Ungleichung $Q(x) \leq f(x)$.

Gilt zusätzlich für jede weitere quadratische, konvexe Minorante q von f , daß $Q(x) \geq q(x)$ für alle $x \in \mathbb{R}^n$, so ist Q die quadratische Struktur von f .

Eine Struktur ist wieder wegen der universellen Eigenschaft eindeutig definiert.

Die Existenz einer quadratischen Struktur ist nicht immer gesichert. Jedoch gilt mit dem obigen Beweis für die quadratische Strukturen ebenfalls der Satz 8, da die konstante Funktion quadratisch und konvex ist. Die Menge der globalen Minima einer Struktur enthält die Menge der globalen Minima der zugehörigen Zielfunktion. Ein Beispiel lieferte die Griewankfunktion.

Folgender Satz zeigt den Zusammenhang zwischen der Gaußtransformation und der quadratischen Struktur einer Zielfunktion:

Satz 9 (quadratische Struktur und Gaußtransformation) Sei $f : \mathbb{R}^n \rightarrow \mathbb{R}$ eine stetige Funktion mit einer quadratischen Struktur $P : \mathbb{R}^n \rightarrow \mathbb{R}$. Sei $Q : \mathbb{R}^n \rightarrow \mathbb{R}$ eine quadratische, konvexe Minorante von f und $R_Q := f - Q$. Für ein $a \in \mathbb{R}$ gelte:

$$\lim_{\lambda \rightarrow \infty} [R_Q]_\lambda(x) = a, \quad \forall x \in \mathbb{R}^n.$$

Dann gilt bereits für ein $\xi \in [0, a]$, daß $P(x) = Q(x) + \xi$ für alle $x \in \mathbb{R}^n$. Insbesondere gilt für $Q(x^*) = f^*$ nach Satz 8 auch $P(x^*) = Q(x^*)$, also $P(x) = Q(x)$ für alle $x \in \mathbb{R}^n$.

Im Beweis dieses Satzes wird der Limes superior \limsup verwendet, er ist wie folgt definiert:

$$\limsup_{\lambda \rightarrow \infty} F(\lambda) = \lim_{\lambda \rightarrow \infty} (\sup\{F(\bar{\lambda}); \bar{\lambda} \geq \lambda\}).$$

Dieser Grenzwert existiert, da die obige Supremumsbildung monoton fallend ist, für eine beschränkte Funktion. Analog definiert man den Limes inferior.

Beweis zu Satz 9: Schreibe:

$$\begin{aligned} P(x) &= P_1 + P_2^T x + \frac{1}{2} x^T P_3 x, \\ Q(x) &= Q_1 + Q_2^T x + \frac{1}{2} x^T Q_3 x, \end{aligned} \tag{3.12}$$

mit Konstanten $P_1, Q_1 \in \mathbb{R}$, Vektoren $P_2, Q_2 \in \mathbb{R}^n$ und positiv semidefiniten $n \times n$ -Matrizen P_3, Q_3 .

Da P die quadratische Struktur von f ist, gilt

$$f(x) \geq P(x) \geq Q(x) \quad \forall x \in \mathbb{R}^n.$$

Und mit $R_P := f - P$ erhält man die Abschätzung:

$$0 \leq R_P(x) \leq R_Q(x) \quad \forall x \in \mathbb{R}^n.$$

Aufgrund der Isotonie der Gaußtransformation für $\lambda \geq 0$ existiert für alle $x \in \mathbb{R}^n$ folgender Grenzwert:

$$\begin{aligned}
0 &\leq \limsup_{\lambda \rightarrow \infty} [R_P]_\lambda \\
&\leq \limsup_{\lambda \rightarrow \infty} [R_Q]_\lambda \\
&= a.
\end{aligned} \tag{3.13}$$

Mit den Voraussetzungen des Satzes und der Formel (3.11) schreibt man für alle $x \in \mathbb{R}^n$ mit den Ausdrücken aus (3.12):

$$\begin{aligned}
-a &\leq \limsup_{\lambda \rightarrow \infty} ([R_P]_\lambda(x) - [R_Q]_\lambda(x)) \\
&= \limsup_{\lambda \rightarrow \infty} ((Q - P)(x) - \frac{1}{4} \text{Spur}(Q_3 - P_3) \lambda^2) \\
&\leq 0.
\end{aligned} \tag{3.14}$$

Dasselbe Argument läßt sich mit \liminf anstelle von \limsup durchführen. Dieses liefert aus der Beschränktheit der gegebenen Limiten in (3.14), daß die Spur der Hessematrizen der Funktionen P und Q übereinstimmt und weiter für alle $x \in \mathbb{R}^n$ die Bedingung $-a \leq (Q - P)(x) \leq 0$. Da P und Q quadratisch sind, kann diese Ungleichung nur erfüllt sein, wenn es ein $\xi \in [0, a]$ gibt, mit $(Q - P)(x) = -\xi$ für alle $x \in \mathbb{R}^n$. \square

Der Satz besagt, daß unter der Voraussetzung der Existenz einer streng konvexen, quadratischen Struktur P eine Funktion Q mit der beschriebenen Grenzwerteigenschaft ihr Optimum an derselben Stelle annimmt wie P , also nach Satz 8 an derselben Stelle wie f .

Es gilt dabei weiter:

$$Q(x) = \lim_{\lambda \rightarrow \infty} [f]_\lambda(x) - \frac{\text{Spur}(Q_3)}{4} \lambda^2 - a.$$

Das bedeutet also, daß für hinreichend große λ das Minimum von $[f]_\lambda$ dem Minimum von Q beliebig nahe kommt und somit dem globalen Minimum von f .

So findet die Gaußtransformation im Falle quadratischer Strukturen die globale Eigenschaft der Zielfunktion.

Leider lassen sich diese Aussagen nicht ohne weiteres auf Molekülpotentialfunktionen anwenden. Eine Voraussetzung für die Existenz einer streng konvexen, quadratischen Struktur der Zielfunktion ist die Eindeutigkeit des globalen Minimums. Andererseits hängt die Sythese dieser Struktur von der Grenzwerteigenschaft aus Satz 9 ab, die bei Molekülpotentialfunktionen ebenfalls nicht vorausgesetzt werden kann.

Kapitel 4

Distanzgeometrie

4.1 Spezielle Funktionstypen

Ziel dieser Arbeit wird es sein, Molekülpotentialfunktionen zu optimieren, deren Energiewert von den paarweisen Abständen der Atome abhängt.

Sei dabei $x \in \mathbb{R}^n$ der Vektor der $n := 3N$ kartesischen Koordinaten eines Moleküls mit N Atomen X_1, \dots, X_N . $r_{ij} : \mathbb{R}^6 \rightarrow \mathbb{R}$ bezeichne den euklidischen Abstand zwischen Atom X_i und Atom X_j , d.h. ist $x_{kl} := x_{3(k-1)+l}$ für $k = 1, \dots, N$ und $l = 1, 2, 3$ die l -te Koordinate von Atom X_k , dann gilt:

$$r_{ij}(x) = \sqrt{(x_{i1} - x_{j1})^2 + (x_{i2} - x_{j2})^2 + (x_{i3} - x_{j3})^2}.$$

Mit "Abstandspotentialen" $H_{ij} : \mathbb{R} \rightarrow \mathbb{R}$ hat eine entsprechende Potentialfunktion $f : \mathbb{R}^n \rightarrow \mathbb{R}$ die Gestalt:

$$f(x) = \sum_{i>j}^N f_{ij}(x) = \sum_{i>j}^N H_{ij}(r_{ij}(x)). \quad (4.1)$$

Aus der Linearität der Gaußtransformation folgt:

$$[f]_\lambda(x) = \sum_{i>j}^N [H_{ij} \circ r_{ij}]_\lambda(x).$$

Führt man den Differenzvektor $y^{ij} \in \mathbb{R}^3$ ein, so gilt mit einer entsprechenden Funktion $h_{ij} : \mathbb{R}^3 \rightarrow \mathbb{R}$ und $y_l^{ij} := x_{il} - x_{jl}$ für $l = 1, 2, 3$

$$H_{ij}(r_{ij}(x)) = H_{ij}(\|y^{ij}\|) = h_{ij}(y^{ij}). \quad (4.2)$$

Für diesen speziellen Funktionstyp haben Moré und Wu [23, 24] eine Transformationsvorschrift hergeleitet:

Satz 10 (Transformationsformel) *Es gilt mit obigen Bezeichnungen (4.2):*

$$[H_{ij} \circ r_{ij}]_{\lambda}(x) = [h_{ij}]_{\sqrt{2\lambda}}(y^{ij}).$$

Und weiter für gerade Funktionen H_{ij} :

$$[h_{ij}]_{\sqrt{2\lambda}}(y^{ij}) = \frac{1}{\sqrt{\pi} \|y^{ij}\|} \int_{-\infty}^{\infty} (\|y^{ij}\| + \sqrt{2\lambda}s) H_{ij}(\|y^{ij}\| + \sqrt{2\lambda}s) e^{-s^2} ds.$$

Das Resultat dieses Satzes ist eine Transformationsformel für die Funktion f der Form (4.1):

$$[f]_{\lambda}(x) = \sum_{i>j}^N \frac{1}{\sqrt{\pi} r_{ij}(x)} \int_{-\infty}^{\infty} (r_{ij}(x) + \sqrt{2\lambda}s) H_{ij}(r_{ij}(x) + \sqrt{2\lambda}s) e^{-s^2} ds. \quad (4.3)$$

Beweis von Satz 10: Der Beweis der ersten Aussage des Satzes ist ein Resultat von Theorem 4.1 in [24], das besagt, daß für Funktionen $f : \mathbb{R}^n \rightarrow \mathbb{R}$ und $h : \mathbb{R}^p \rightarrow \mathbb{R}$ mit

$$f(x) = h(P^T x)$$

und einer Matrix $P \in \mathbb{R}^{n \times p}$ mit $P^T P = \sigma^2 E_p$, E_p ist die p -te Einheitsmatrix, gilt:

$$[f]_{\lambda}(x) = [h]_{\sigma\lambda}(P^T x).$$

Ebenso erhält man die zweite Aussage des Satzes durch direktes Rechnen, indem man y auf sphärische Koordinaten

$$y = \begin{pmatrix} s \cos \theta \\ s \sin \theta \\ s \cos \phi \end{pmatrix}$$

transformiert, entsprechend Theorem 3.3 und Corollar 3.5 in [24]. \square

4.2 Eine spezielle Distanzgeometriefunktion

Wie bereits in der Einleitung erwähnt, basiert die vorgestellte Methode zum Lösen globaler Optimierungsprobleme von Molekülpotentialfunktionen auf einem Algorithmus zum Lösen von *Distanzgeometrie*problemen.

Hier soll nun dieses Problem vorgestellt werden, das bei der Strukturaufklärung von Molekülgeometrien auftritt.

Moleküle sind vereinfacht betrachtet eine Anordnung von N Punkten im dreidimensionalen Raum. Eine bestimmte Molekülgeometrie läßt sich daher durch $n := 3N$ kartesische Koordinaten beschreiben bzw. durch einen Vektor $x \in \mathbb{R}^n$.

Durch NMR-Experimente ist es möglich, bestimmte intermolekulare Distanzen zu messen [3, 21, 5, 36]. Das bedeutet, man erhält Aufschluß darüber, wie weit jeweils ein Atom X_i von einem anderen Atom X_j entfernt ist. Dieser Abstand $r_{ij}(x)$ ist entweder genau bekannt oder durch einen Bereich $a_{ij} \leq r_{ij}(x) \leq b_{ij}$ gegeben, wobei für b_{ij} auch das Symbol ∞ zugelassen ist. Bezeichnet man mit \mathfrak{R} die Menge aller Paare (i, j) , $i < j$, deren Distanzbereich bekannt ist, so lautet das Distanzgeometrieproblem:

Finde einen Punkt $x^* \in \mathbb{R}^n$, so daß

$$a_{ij} \leq r_{ij}(x^*) \leq b_{ij}, \quad \forall (i, j) \in \mathfrak{R}, \quad (4.4)$$

wobei $0 \leq a_{ij} \leq b_{ij} \leq \infty$.

Moré und Wu haben zunächst ein spezielles Distanzgeometrieproblem untersucht. Es sollen dabei bestimmte intermolekulare Distanzen exakt vorgegeben sein, d.h. zu einer Menge von Paaren $(i, j) \in \mathfrak{R}$, $i > j$, ist der Abstand $r_{ij}(x) = a_{ij}$ genau bekannt.

Cripen und Havel [6] haben bewiesen, daß dieser spezielle Problemtyp im eindimensionalen Raum \mathcal{NP} -hart ist, indem sie das \mathcal{NP} -vollständige Knapsackproblem polynomial in ein solches Distanzgeometrieproblem transformierten. Saxe [32] hat dieses Resultat auf höherdimensionalen Räumen gezeigt. Das wiederum bedeutet, daß das globale Optimierungsproblem, in das man das spezielle Distanzgeometrieproblem umwandelt, \mathcal{NP} -hart ist. Löst man dieses Problem also über einen globalen Optimierungsalgorithmus, so wird dieser nicht gleichzeitig polynomial und deterministisch sein können (Anmerkung

auf Seite 10). Eine Distanzgeometriefunktion für obiges Problem läßt sich in der Form (4.1) schreiben, wobei $H_{ij} : \mathbb{R} \rightarrow \mathbb{R}$ folgende Gestalt hat:

$$H_{ij}(r) = (r^2 - a_{ij}^2)^2. \quad (4.5)$$

Jeder Punkt $x \in \mathbb{R}^n$ mit $f(x) = 0$ löst das Distanzgeometrieproblem und ist nach Satz 2 globales Optimum der Funktion. Gilt andererseits für das globale Optimum x^* von f , daß $f(x^*) > 0$, so hat das zugehörige Distanzgeometrieproblem keine Lösung.

Die Gaußtransformation von $f : \mathbb{R}^n \rightarrow \mathbb{R}$ mit den geraden Funktionen H_{ij} ergibt sich aus Formel (4.3) und der Tatsache, daß $r \cdot H_{ij}(r)$ ein Polynom 5.Grades in r ist und somit nach Formel (3.11) transformiert werden kann, d.h.

$$[f]_\lambda(x) = \sum_{(i,j) \in \mathfrak{R}} \left((r_{ij}(x)^2 - a_{ij}^2)^2 + 10r_{ij}(x)^2\lambda^2 - 6\lambda^2 a_{ij}^2 + 15\lambda^4 \right). \quad (4.6)$$

Nehmen wir ein zweiatomiges Molekül an, bei dem die Atome den Abstand 1 voneinander haben sollen. Folgendes Bild zeigt die Glättung (4.6) der entsprechenden Distanzgeometriefunktion in Abhängigkeit von dem Abstand der beiden Atome bei verschiedenen Werten für λ :

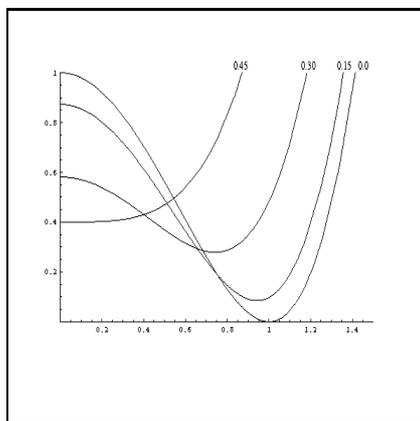


Abbildung 4.1: Glättung einer Distanzgeometriefunktion

Aus der Funktion (4.6) kann man sehen, daß $[f]_\lambda(x)$ für $\lambda > a_{ij}\sqrt{0.2}$ in $r_{ij} = 0$ ein Minimum und für $\lambda < a_{ij}\sqrt{0.2}$ in $r_{ij} = 0$ ein Maximum hat.

Für $\lambda = a_{ij}\sqrt{0.2}$ verzweigt sich im Homotopiealgorithmus die Trajektorie der Minima; vgl. auch [17].

Dadurch hängt das Verhalten der Optimierungsroutine entscheidend von der gewählten Diskretisierung und der Wahl der lokalen Optimierungsmethode ab.

Wie das Beispiel aus Abbildung 3.3 zeigt, führt nicht jede Anwendung des Homotopiealgorithmus zu Verzweigungspunkten und somit zu chaotischem Verhalten.

Eine vernünftige Wahl des Startparameters für die Homotopiemethode zur Vermeidung dieser (sicheren) Verzweigungspunkte liefert

$$L < \min\{a_{ij}\}\sqrt{0.2}. \quad (4.7)$$

4.3 Die numerische Umsetzung

Wir haben bereits die drei Phasen der numerischen Homotopiemethode in Abschnitt 3.3 kennengelernt. Hier soll sie nochmal mit der Gaußtransformation formuliert werden. L bezeichnet dabei den Startwert für den Glättungsparameter, m die Zahl an Diskretisierungspunkten und p die Zahl der Starts:

0. Wähle $L > 0$, $m \in \mathbb{N}$ und $p \in \mathbb{N}$, sowie eine Diskretisierung:

$$L = \lambda_m > \dots > \lambda_0 = 0.$$

1. Führe die Schritte 2 bis 4 p -mal durch. Danach wähle aus allen erhaltenen lokalen Minima von f das mit kleinstem Funktionswert als mögliches globales Minimum x^* aus.
2. Wähle einen Zufallsvektor x und optimiere $[f]_L$ ausgehend von x lokal (Diffusionsphase). Ergebnis: x^m .
3. Optimiere für $k = m - 1, \dots, 1$ die Funktionen $[f]_{\lambda,k}$ ausgehend von x^{k+1} lokal mit Ergebnisvektor x^k .
4. Optimiere f ausgehend von x^1 lokal (Verfeinerung). Ergebnis: x^0 . Füge x^0 zur Menge der erhaltenen lokalen Minima hinzu.

4.3.1 Die Simplex-Methode

Das BFGS-Verfahren zur lokalen Optimierung konvexer Funktionen wurde bereits in Abschnitt 2.5 behandelt, diese und andere Optimierungsmethoden liegen als Programmcode in [28] vor. Ihr Verhalten wurde bereits an Molekülpotentialfunktionen getestet; vgl. Kunz [21].

Bei der Homotopiemethode spielt das *Erreichen* eines lokalen Optimums eine entscheidende Rolle. Dieses kann eine BFGS-Routine nicht in hundert Prozent der Fälle garantieren, zumal die Routine bei kleinen Gradienten die Update-Formel nicht aufruft. Eine Optimierungsroutine, die in solchen Fällen sicher zum Ziel führt ist die sogenannte Simplex-Methode.

Die Downhill Simplex-Methode von Nelder und Mead [27] benötigt zur lokalen Minimierung einer multidimensionalen Funktion $f : \mathbf{R}^n \rightarrow \mathbf{R}$ nur Funktionswerte und keine Gradienteninformationen. Sie ist deshalb nicht sehr effizient und wird nur eingesetzt, falls die BFGS-Routine nach einer bestimmten Zahl von Schritten (n^2) nicht zu einem lokalen Minimum gelangen konnte.

Ihre Arbeitsweise: Die n -dimensionale Funktion f wird zunächst an $n + 1$ Stellen ausgewertet, die die Ecken eines sogenannten *Simplex* bilden. Der Punkt mit dem höchsten Funktionswert wird entfernt und an der Fläche, die die übrigen n Punkte bilden, gespiegelt. Dadurch entsteht ein neues Simplex.

Um ein Stoppkriterium (Differenz der Funktionswerte an den Ecken des Simplex) zu erreichen, kontrahiert der Algorithmus das Simplex, wenn der reflektierte Punkt einen höheren Funktionswert als den zweitschlechtesten Wert aufweist.

Ist der reflektierte Punkt jedoch besser als alle anderen Punkte, so wird das Simplex in Richtung der vollzogenen Spiegelung expandiert, damit der Algorithmus nach Durchlaufen eines langen schmalen "Tales" in einer anschließenden "Ebene" mit größeren Simplexes arbeitet.

Auf diese Weise kriecht das Simplex wie eine "Amöbe" über die Hyperfläche des Funktionsgraphen und bewegt sich dabei sicher von Maxima und Sattelpunkten weg, weshalb die Routine auch AMOEBA [28] genannt wird.

Zur Nutzung dieser Routine muß das Molekül eine bestimmte Koordinatendarstellung besitzen, die in dem folgenden Abschnitt erklärt wird.

4.3.2 Eindeutige kartesische Koordinaten

Die $n := 3N$ kartesischen Koordinaten eines Moleküls mit N Atomen, beschreiben zwar dessen Molekülgeometrie, jedoch kann ein und dieselbe Molekülgeometrie durch Translation und Rotation durch unendlich viele Vektoren $x \in \mathbb{R}^n$ repräsentiert werden.

Dieses hat zur Folge, daß die Hessematrix einer Molekülpotentialfunktion $f : \mathbb{R}^n \rightarrow \mathbb{R}$ stets sechs triviale (zum Eigenwert 0 gehörende) Eigenvektoren besitzt, die die translatorischen und rotatorischen Eigenschwingungsrichtungen des Moleküls repräsentieren; vergleiche Höweler [15].

Durch zusätzliche Nebenbedingungen kann man diese trivialen Eigenschwingungen des Moleküls unterbinden. Beim BFGS-Verfahren ist dieses nicht nötig, mehr noch: Es vergrößert die Zahl an Iterationsschritten, wenn man in den sog. "eindeutigen kartesischen Koordinaten" rechnet; vergleiche Bollweg [4].

Beim Newton-Verfahren muß man jedoch aufgrund der Invertierbarkeit der Hessematrix Nebenbedingungen in die Minimierung aufnehmen.

Auch beim Simplex-Algorithmus sollte man in eindeutigen kartesischen Koordinaten rechnen, da andernfalls das Molekül während der Optimierung zu rotieren oder zu translätieren beginnt, weil keine Gradienteninformationen zur Verfügung stehen.

Sei mit der alten Notation wieder $x_{3(k-1)+l}$ die l -te Koordinate von Atom X_k , dann lauten die Nebenbedingungen zur Unterdrückung der trivialen Eigenschwingungen:

$$x_1 = x_2 = x_3 = x_5 = x_6 = x_9 = 0.$$

Normalerweise wird eine beliebige Molekülgeometrie $x \in \mathbb{R}^n$ durch Translation und allgemeine Drehmatrizen in eindeutige kartesische Koordinaten überführt. In diesen Matrizen treten Winkelfunktionen auf, die durch Abbruchfehler zu kleinen Ungenauigkeiten führen können. Der Beweis zu folgendem Satz beinhaltet eine Transformationsmethode, die numerisch besser konditioniert ist und deshalb in meinem Programm verwendet wird. Gleichzeitig werden noch zwei zusätzliche Nebenbedingungen eingehalten, die 180°-Drehsymmetrien unterbinden, so daß zwei Moleküle auch wirklich nur dann kongruent und nicht enantiomorph (spiegelsymmetrisch) sind, wenn sie dieselben eindeutigen kartesischen Koordinaten besitzen.

Wir definieren eine Menge M , wobei das vorliegende Molekül mindestens 3

Zentren besitzen soll:

$$M := \{x \in \mathbb{R}^n; x_1, x_2, x_3, x_5, x_6, x_9 = 0 \wedge x_4, x_8 \geq 0\}.$$

Satz 11 (Transformation der Koordinaten) *Alle Molekülgeometrien $x \in \mathbb{R}^n$ können durch eine affin-orthogonale, spiefelfreie Symmetrioperation in äquivalente Geometrien $y \in M$ transformiert werden.*

Im folgenden Beweis ist $\langle x, y \rangle$ das kanonische Skalarprodukt zweier Vektoren und $x \times y$ das Vektorprodukt.

Beweis: In folgender Matrix stehen die kartesischen Koordinaten der einzelnen Atome einer beliebigen Geometrie jeweils spaltenweise aufgeführt:

$$A_0 := \begin{pmatrix} x_1 & x_4 & x_7 & x_{10} & \cdots & x_{3n-2} \\ x_2 & x_5 & x_8 & x_{11} & \cdots & x_{3n-1} \\ x_3 & x_6 & x_9 & x_{12} & \cdots & x_{3n} \end{pmatrix}. \quad (4.8)$$

Man translatiere diese Geometrie so, daß das Atom X_1 in den Ursprung des Koordinatensystems fällt. Mathematisch bedeutet dieses, daß in der j -ten Zeile der Matrix A_0 jeweils der Wert x_j abgezogen wird ($j = 1, 2, 3$). Wir erhalten die Matrix A_1 , wobei für $i = 1, \dots, n$ und $j = 1, 2, 3$ gilt: $\tilde{x}_{3(i-1)+j} := x_{3(i-1)+j} - x_j$.

$$A_1 = \begin{pmatrix} 0 & \tilde{x}_4 & \tilde{x}_7 & \tilde{x}_{10} & \cdots & \tilde{x}_{3n-2} \\ 0 & \tilde{x}_5 & \tilde{x}_8 & \tilde{x}_{11} & \cdots & \tilde{x}_{3n-1} \\ 0 & \tilde{x}_6 & \tilde{x}_9 & \tilde{x}_{12} & \cdots & \tilde{x}_{3n} \end{pmatrix}. \quad (4.9)$$

Nun wird das Molekül so gedreht, daß sämtliche Gleichungsbedingungen in M erfüllt sind. Mathematisch entspricht dieses einer Multiplikation von links mit einer orthogonalen Matrix Q , deren Determinante 1 ist:

$$A_2 = QA_1 = \begin{pmatrix} 0 & y_4 & y_7 & y_{10} & \cdots & y_{3n-2} \\ 0 & 0 & y_8 & y_{11} & \cdots & y_{3n-1} \\ 0 & 0 & 0 & y_{12} & \cdots & y_{3n} \end{pmatrix}. \quad (4.10)$$

Bezeichne X_i nun den Vektor $(\tilde{x}_{i1}, \tilde{x}_{i2}, \tilde{x}_{i3})^T$, also den Vektor der kartesischen Koordinaten des Zentrums X_i . Die Matrix Q bestehe aus den Zeilenvektoren q_1, q_2, q_3 . Dann liefern folgende Festlegungen das gewünschte Ergebnis aus (4.10), wie man durch eine kurze Rechnung zeigen kann:

$$\begin{aligned} q_1 &:= \frac{X_2^T}{\|X_2\|}, \\ q_2 &:= \frac{X_3^T - \langle q_1^T, X_3 \rangle q_1}{\|X_3 - \langle q_1^T, X_3 \rangle q_1\|}, \\ q_3 &:= (q_1^T \times q_2^T)^T. \end{aligned} \tag{4.11}$$

Anmerkung: Es ist $X_2 \neq 0$ zu wählen oder es muß X_2 durch einen Vektor $\tilde{X}_2 \neq 0$ ersetzt werden. Anstelle q_2^T wählt man einen beliebigen, zu X_2 orthogonalen und normierten Vektor, falls X_2 und X_3 kollinear sind.

Die Normierung der Zeilenvektoren q_1, q_2, q_3 und die paarweise Orthogonalität zueinander bewirken, daß Q orthogonal ist. Da das Vektorprodukt in (4.11) ein Rechtssystem bildet, hat die Determinante von Q den Wert 1. Q repräsentiert also eine spiefelfreie Drehung.

Weiterhin ist durch die Gleichheit $y_4 = \|X_2\|$ stets $y_4 \geq 0$ gewährleistet.

Die Tatsache, daß $\langle X_3, X_3 \rangle \langle X_2, X_2 \rangle$ stets größer oder gleich $\langle X_3, X_2 \rangle^2$ ist, bewirkt, daß $y_8 \geq 0$ gilt. \square

Diese Möglichkeit, die zulässige Menge einzuschränken, um eine Invertierbarkeit der Hessematrix zu erzwingen, kann durch Aufaddieren der *Straffunktion*

$$\Phi(x) = x_1^2 + x_2^2 + x_3^2 + x_5^2 + x_6^2 + x_9^2 + \min^2\{0, x_4\} + \min^2\{0, x_8\}$$

geschehen.

Satz 11 besagt zwar, daß alle Geometrien $x \in \mathbb{R}^n$ in kongruente Geometrien $y \in M$ transformiert werden können, nicht jedoch daß diese Abbildung eindeutig ist. Dazu sind weitere Voraussetzungen nötig.

Satz 12 (Eindeutigkeit der Koordinatentransformation) Sei $x \in M$ eine Molekülgeometrie mit $x_4, x_8 \neq 0$ (d.h. die ersten drei Atome von x spannen eine Ebene auf). Sei $y \in M$ eine zu $x \neq y$ kongruente Geometrie, dann ist y ein Spiegelbild von x .

Dieser Satz besagt, daß durch die Straffunktion Φ alle rotatorischen und translatorischen Freiheitsgrade unterdrückt werden, solange die ersten drei Atome des Moleküls eine Ebene aufspannen. Der sechsfache Eigenwert 0 der Hessematrix tritt also nicht mehr auf.

Beweis: Bezeichne wieder X_i bzw. Y_i den Vektor der kartesischen Koordinaten von Atom i in der Geometrie x bzw. y . Die ersten drei Atome von x spannen eine Ebene auf. Das Dreieck $\Delta(X_1, X_2, X_3)$ muß also kongruent zu dem Dreieck $\Delta(Y_1, Y_2, Y_3)$ sein. Insbesondere müssen die Seitenlängen $a, b, c \neq 0$

$$\begin{aligned} a^2 &= \langle X_2 - X_3, X_2 - X_3 \rangle \\ b^2 &= \langle X_1 - X_3, X_1 - X_3 \rangle \\ c^2 &= \langle X_1 - X_2, X_1 - X_2 \rangle \end{aligned} \quad (4.12)$$

mit denen in y übereinstimmen.

Das Gleichungssystem (4.12), für y formuliert, führt zu:

$$\begin{aligned} y_4^2 &= c^2 \\ y_7^2 + y_8^2 &= b^2 \\ (y_4 - y_7)^2 + y_8^2 &= a^2. \end{aligned} \quad (4.13)$$

Da $y_4 \geq 0$ gilt, ist $y_4 = c$. Zieht man in (4.13) die dritte von der zweiten Gleichung ab, so erhält man

$$y_7 = \frac{b^2 - a^2 + c^2}{2c}.$$

Eingesetzt in die zweite Gleichung ergibt sich mit $y_8 \geq 0$:

$$y_8 = \sqrt{b^2 - \left(\frac{b^2 - a^2 + c^2}{2c}\right)^2}.$$

Damit sind die Koordinaten der ersten drei Zentren in y fixiert und stimmen mit den Koordinaten der ersten drei Zentren in x überein.

Ist y nun kongruent zu x , so bleibt das Dreieck $\Delta(X_1, X_2, X_3)$ bei der Kongruenzabbildung fixiert. Die einzige Abbildung, die dieses ermöglicht, ist eine Spiegelung an der Ebene, die durch die drei Atome X_1, X_2, X_3 aufgespannt wird. \square

Bemerkung: Eine solche Spiegelung der Geometrie $x \in M$ ist gleichbedeutend mit einem Vorzeichenwechsel aller z -Koordinaten in x . Durch eine weitere Nebenbedingung (z.B. $x_{12} > 0$) kann man auch diese Symmetrieoperation unterbinden.

Eine Molekülpotentialfunktion kann nur dann ein eindeutiges, globales Minimum haben, wenn man die zulässige Menge auf die oben beschriebene Weise einschränkt. In jedem anderen Fall führen Symmetrieoperationen zu äquivalenten Minima und somit zum sicheren Fehlen einer streng konvexen, besten Minoranten, die die Grundlage vieler Optimierungsalgorithmen (z.B. CGU-Verfahren) bildet.

4.4 Der Sensitivitätsschritt

Durch die Homotopiefunktion $[f]_\lambda$ wird eine Trajektorie der lokalen Minima $x(\lambda)$ definiert. Bei einer *dynamischen Betrachtung* der Optimierungsroutine ist daher der *Übergang* von einem Parameter zu kleineren Werten von Bedeutung. Wie sich nun zeigen wird, entspricht dieser Übergang zu kleineren Werten für λ einer Warmstartoption für die BFGS-Routine.

Unter der Voraussetzung, daß die Hessematrix stets positiv definit ist, also stets strenge lokale Minima vorliegen, läßt sich aus der Theorie der Sensitivitätsanalyse unbeschränkter Optimierungsprobleme eine Dynamik

$$x : \mathbb{R} \rightarrow \mathbb{R}^n$$

herleiten, die die Lage der lokalen Minima in Abhängigkeit von dem Glättungsparameter λ beschreibt. Einfachheitshalber schreiben wir $f(x, \lambda) := [f]_\lambda(x)$ für die Gaußtransformierte.

Es gilt:

$$\frac{d}{d\lambda}x(\lambda) = -f_{xx}^{-1}(x(\lambda), \lambda) f_{x\lambda}(x(\lambda), \lambda),$$

$$x(L) := \min_{x \in \mathbb{R}^n} f(x, L), \quad \lambda \in [0, L], \quad (4.14)$$

wobei die partiellen Ableitungen als untere Indizes an die Funktion geschrieben worden sind.

Diese Dynamik (4.14) läßt sich nur dann so beschreiben, wenn die Hessematrix f_{xx} stets invertierbar, also positiv definit ist. Deshalb muß man die sechs trivialen Eigenschwingungen eines Moleküles, zum Beispiel durch die Transformation in eindeutige kartesische Koordinaten und durch Addition der Straffunktion

$$\Phi(x) := x_1^2 + x_2^2 + x_3^2 + x_5^2 + x_6^2 + x_9^2$$

zu der Zielfunktion, eliminieren. Weiter dürfen keine Verzweigungspunkte auftreten, was wiederum die Wahl der Anfangsglättung durch die Formel (4.7) beschränkt.

Approximiert man die Dynamik (4.14) durch eine Diskretisierungstechnik, so nimmt man eine Zerlegung des Intervalles $[0, L]$ in Teilintervalle der Länge h vor und schreibt

$$\frac{x(\lambda) - x(\lambda - h)}{h} \approx -f_{xx}^{-1}(x(\lambda), \lambda) f_{x\lambda}(x(\lambda), \lambda). \quad (4.15)$$

Man erhält also den Punkt $x(\lambda - h)$ aus dem gegebenen Punkt $x(\lambda)$ mit Hilfe der Gleichung (4.15).

Es fehlt jetzt noch die Berechnung der Ausdrücke f_{xx} und $f_{x\lambda}$.

Ableitung nach dem Störparameter: Die Dynamik (4.14) setzt voraus, daß in jedem Punkt ein lokales Minimum der jeweiligen Funktionen $f(\cdot, \lambda)$ vorliegt, also der Gradient f_x verschwindet. Schreiben wir auch $f_{x\lambda}$ als Differenzenquotient, so ergibt sich aus Gleichung (4.15) zunächst

$$\frac{x(\lambda) - x(\lambda - h)}{h} \approx -f_{xx}^{-1}(x(\lambda), \lambda) \frac{f_x(x(\lambda), \lambda) - f_x(x(\lambda), \lambda - h)}{h}.$$

Und mit $f_x(x(\lambda), \lambda) = 0$ schließlich:

$$x(\lambda - h) \approx x(\lambda) - f_{xx}^{-1}(x(\lambda), \lambda) f_x(x(\lambda), \lambda - h). \quad (4.16)$$

Inverse Hessematrix: Um die Rekursion (4.16) berechnen zu können, benötigt man eine Approximation der inversen Hessematrix.

Aus Abschnitt 2.5 ist bekannt, daß diese durch den Aufruf eines BFGS-Algorithmus als lokale Optimierungsmethode erhalten werden kann.

Gleichung (4.16) stellt dann lediglich eine Warmstartoption für den BFGS-Algorithmus dar, wie man aus einem Vergleich mit der Formel (2.4) leicht sieht. Sie besagt nämlich, daß bei dem Übergang von dem Glättungsparameter λ zu $\lambda - h$ die genäherte inverse Hessematrix aus der Optimierung auf dem Niveau λ für einen ersten Optimierungsschritt auf dem Niveau $\lambda - h$ übernommen wird.

Der Algorithmus aus Abschnitt 4.3 läßt sich auf diese Weise leicht um eine dynamische Variante erweitern. Dazu benötigt man

- eine äquidistante Einteilung des Intervalls $[0, L]$,
- positiv definite Hessematrizen und
- eine Warmstartoption für den BFGS-Algorithmus.

Kapitel 5

Molekülpotentialfunktionen

5.1 Quantenchemischer Ansatz

Die Stabilität eines Moleküles ergibt sich vor allem aus den Wechselwirkungen ungebundener Atome. Diese können *Gleichgewichtsabstände* einnehmen, so daß das Molekül potentielle Energie an die Umgebung verliert und sich somit stabilisiert.

Die Stabilität einer Molekel wird durch eine Potentialfunktion beschrieben, die von der Anordnung der Zentren im Molekül abhängt.

Seit der Heisenbergschen Unbestimmtheitsrelation wissen wir, daß Ort und Impuls eines Teilchens niemals gleichzeitig scharf zu bestimmen sind. Vielmehr können wir nur bestimmte Wahrscheinlichkeiten für den Aufenthalt und den Impuls der Teilchen angeben, die in der Schrödingerdarstellung in einer Zustandsfunktion Ψ zusammengefaßt sind.

Die Funktionen stationärer Zustände sind Lösungen des folgenden Eigenwertproblems, das sich unter gewissen Randbedingungen aus der zeitunabhängigen Schrödingergleichung ergibt:

$$\mathbf{H}\Psi_\nu = E_\nu\Psi_\nu, \quad (5.1)$$

mit einem Hamiltonoperator (ein spezieller Differentialoperator) \mathbf{H} , den Energieeigenwerten $E_\nu \in \mathbb{R}$ und den zu den Energiewerten gehörenden, im Lebesgueschen Sinne quadratisch integrierbaren Eigenzustandsfunktionen $\Psi_\nu : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{C}$.

Der Hamiltonoperator setzt sich dabei aus verschiedenen Summanden zu-

sammen, die die einzelnen Energiebeiträge zu einer Gesamtenergie E_ν darstellen. Diese Summanden sind wiederum Operatoren, die sich aus den Vertauschungsrelationen von Ort und Impuls ergeben.

Wirken keine äußeren Käfte auf ein Molekül ein, so setzt sich die Gesamtenergie des Moleküls aus der kinetischen Energie der Kerne T_K und der Elektronen T_E , sowie einer potentiellen Energie V zusammen, die vorrangig Coulomb-Wechselwirkungen zwischen den Kernladungen und den Elektronen bedeutet, also Kern-Kern, Kern-Elektron, Elektron-Elektron Wechselwirkungen.

$$\mathbb{H} = \mathbb{T}_K + \mathbb{T}_E + V.$$

Bei der Born-Oppenheimer Näherung geht man davon aus, daß sich die Kerne aufgrund ihrer Masse viel langsamer bewegen als die Elektronen. Es ist also möglich, die Kerne als starre Punkte anzusehen ($\mathbb{T}_K = 0$). Man gibt den Ort der Kerne in dem Vektor $x \in \mathbf{R}^n$ fest vor. Dieses Vorgehen widerspricht der Heisenbergschen Unbestimmtheitsrelation, die besagt, daß Ort x und Impuls T_K nicht gleichzeitig scharf zu bestimmen sind. Jedoch hat sich die Born-Oppenheimer Näherung in der Praxis als eine Möglichkeit etabliert, Potentialfunktionen formulieren zu können, die die Realität genau genug approximieren.

Die potentielle Energie V hängt nur von den Kernkoordinaten x und Elektronenkoordinaten y ab. Der Operator V ist dabei ein einfacher Multiplikationsoperator $V(x, y)$; vergleiche Primas und Müller-Herold [29], Seiten 153 bis 159.

Die Schrödingergleichung (5.1) vereinfacht sich abhängig von der Anordnung der Kerne (also für festes x) jeweils zu:

$$[\mathbb{T}_E + V(x, \mathbf{y})]\Psi_\nu(x, \mathbf{y}) = E_\nu(x)\Psi_\nu(x, \mathbf{y}),$$

mit \mathbf{y} , der Anordnung der Elektronen, als freie Variable.

Zwei Punkte sind hier zu nennen:

1. Zu jeder Kernanordnung gibt es diskrete Eigenwerte $E_\nu(x)$. Die Funktion

$$f(x) := \min_{\nu} \{E_\nu(x)\}$$

wird als Molekülpotentialfunktion (oder einfach nur "Potentialfunktion") bezeichnet.

2. Die Molekülpotentialfunktion beschreibt also die potentielle Energie des Gesamtsystems in seinem Grundzustand bei einer theoretischen Temperatur von Null Kelvin, da die kinetische Energie der Kerne T_K durch die Born-Oppenheimer Näherung auf Null gesetzt wurde.

Der Anteil $V(x, y)$ im vereinfachten Hamilton-Operator umfaßt im wesentlichen Coulomb-Wechselwirkungen, deren Energiebeitrag nur von den paarweisen Abständen der wechselwirkenden Zentren abhängt.

Die Idee liegt also nahe, die Wechselwirkung ungebundener Zentren im Molekül durch abstandsabhängige Potentialterme zu beschreiben, wohingegen die Wechselwirkungen gebundener Zentren zunächst als konstant gesehen werden.

Quantenchemische Rechnungen haben dabei ein ganz bestimmtes Verhalten ungebundener Atome festgestellt (Van-der-Waals Wechselwirkung). Bis zu einem Gleichgewichtsabstand r_0 ziehen sich ungebundene Atome in einem Molekül gegenseitig an. Kommen sich zwei Atome jedoch näher als dieser Abstand, so tritt eine starke Van-der-Waals Abstoßung ein.

Der Verlauf einer solchen Potentialkurve in Abhängigkeit von dem Abstand zweier Atome ist im nächsten Bild dargestellt, als Energiewert Null definiert man dabei die potentielle Energie bei unendlich großem Abstand der Atome.

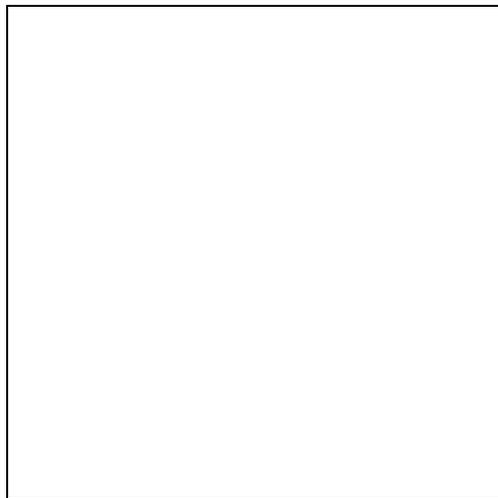


Abbildung 5.1: Das Lennard-Jones Potential

Die Kurve beschreibt also den Energiegewinn durch stabilisierende Wechselwirkung gegenüber nicht wechselwirkenden Atomen. Eine globale Minimierung der Summe dieser Energiegewinne ist gleichbedeutend mit maximaler Stabilität des Moleküls. Der Verlauf dieser Kurve wird in den meisten Kraftfeldern (das sind Molekülpotentialfunktionen, die nicht quantenchemisch berechnet werden) durch ein Lennard-Jones Potential approximiert; vergleiche Kunze [21], Seite 92.

$$H(r) = \frac{n K_{VdW}}{n - m} \left[\frac{m}{n} \left(\frac{r_0}{r} \right)^n - \left(\frac{r_0}{r} \right)^m \right],$$

wobei meistens $n = 6$ und $m = 12$ (also der 6-12-Term) verwendet wird. Als Lennard-Jones Potential zwischen zwei Atomen X_i und X_j ergibt sich mit $A_{ij} = K_{ij} r_{0,ij}^{12}$ und $B_{ij} = 2 K_{ij} r_{0,ij}^6$:

$$H_{ij}(r_{ij}) = \frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6}. \quad (5.2)$$

5.2 Das Coulomb Potential

Das Lennard-Jones Potential berücksichtigt Wechselwirkungen ungebundener Atome im Molekül, wenn sich diese in einer dem Atomtypus entsprechenden, chemischen Umgebung befinden.

Jedoch gibt es in der Natur bestimmte Atomtypen oder Gruppen von Atomen, die benachbarte Elektronen besonders stark anziehen oder abstoßen. Dadurch wird die Verteilung der Elektronen im Molekül verändert und es treten sogenannte Partiaalladungen durch lokalen Elektronenüberfluß oder Elektronenmangel auf, die die chemische Umgebung anderer Atome verändern. Um die sogenannte *Populationsanalyse* durchzuführen, werden zunächst quantenchemische Rechnungen an der vorliegenden Molekülstruktur vorgenommen. Dabei entspricht das Quadrat der komplexwertigen Zustandsfunktion $|\Psi|^2$ der lokalen Aufenthaltswahrscheinlichkeit der Elektronen bzw. der Elektronendichte. Aus dieser Größe kann man den Atomen des Moleküls jeweils eine bestimmte Partiaalladung zuordnen. Aufgrund der Born-Oppenheimer Näherung ergibt sich für jede Kernanordnung $x \in \mathbf{R}^n$ eine andere Zustandsfunktion $\Psi(x, \cdot)$, so daß die Populationsanalyse nur bei vollständig bekannten Molekülteilgeometrien durchgeführt werden kann.

Zu der Van-der-Waals Wechselwirkung erfahren ungebundene Atome somit zusätzlich eine elektrostatische Coulomb Wechselwirkung. Dieser Wechselwirkungstyp ist besonders weitreichend. Die Energie, die aus dieser resultiert, nimmt nur proportional zu $1/r$ ab (unter Vernachlässigung einer abstandsabhängigen Abschirmung).

Um Formel (4.3) verwenden zu können, schreibt man als paarweise Coulomb Wechselwirkung zwischen den Atomen X_i und X_j :

$$H_{ij}(r) = \frac{q_i q_j}{4\pi\epsilon_0} \frac{1}{|r|}, \quad (5.3)$$

wobei q_i, q_j die Ladungen der Atome X_i bzw. X_j und ϵ_0 die sogenannte Dielektrizitätskonstante des Vakuums sind.

Mit $C_{ij} := q_i q_j / (4\pi\epsilon_0)$, $s^* := r_{ij}(x) / (\sqrt{2}\lambda)$ und der Formel (4.3) erhält man für $\lambda > 0$:

$$[f_{ij}]_\lambda(x) = \frac{C_{ij}}{r_{ij}(x)} \frac{2}{\sqrt{\pi}} \int_0^{s^*} \exp(-s^2) ds.$$

Der letzte Term entspricht dabei der Fehlerfunktion $\text{erf}(\cdot)$:

$$[f_{ij}]_\lambda(x) = \frac{C_{ij}}{r_{ij}(x)} \text{erf}\left(\frac{r_{ij}(x)}{\sqrt{2}\lambda}\right). \quad (5.4)$$

Eine numerische Approximation der Fehlerfunktion ist Bestandteil der Mathematikbibliothek der Programmiersprache C [2]. Zu bemerken bleibt, daß (5.4) für $\lambda \rightarrow \infty$ gegen die Ausgangsfunktion konvergiert, da $\text{erf}(x) \rightarrow 1$ für $x \rightarrow \infty$. Nun noch ein Wort zu den verwendeten Einheiten: Die Ladung eines Atomes wird üblicherweise in e angegeben (Anzahl an Elementarladungen), dabei entspricht $1 e = 1.60219 \cdot 10^{-19} C$, die Abstände der Atome werden in Å also in $10^{-10} m$ angegeben. Der Dielektrizitätsfaktor hat den Wert $4\pi\epsilon_0 = 1.11265 \cdot 10^{-10} J^{-1} C^2 m^{-1}$.

Um nun auf die Einheit $kcal/mol$ umzurechnen, muß man noch verwenden, daß $1 kcal = 4184 J$ und $1 mol = 6.02205 \cdot 10^{23}$ Teilchen entsprechen.

Insgesamt ergibt sich für C_{ij} in (5.4):

$$C_{ij} = 332.0643 q_i q_j [kcal/mol]$$

5.3 Die Strafmethode

Die Wahl der Parameter des Lennard-Jones Potentials hängt entscheidend von der Konnektivität des Moleküles ab.

Die Konnektivitätsmatrix eines Moleküles läßt bei gebundenen Atomen nur bestimmte *Bindungsabstände* und *Bindungswinkel* zu, die z.B. durch ein Distanzgeometrieproblem in der Form (4.5) beschrieben werden können.

Das eigentliche Problem lautet also: Minimiere die Summe der Lennard-Jones und Coulomb-Potentiale $g : \mathbf{R}^n \rightarrow \mathbf{R}$ unter Einhaltung gewisser intermolekularer Distanzen.

$$\begin{aligned} \min_{x \in \mathbf{R}^n} \quad & g(x) \\ \text{unter} \quad & g_{ij}(x) = 0 \quad (i, j) \in \mathfrak{R}, \end{aligned} \tag{5.5}$$

wobei

$$g_{ij}(x) = r_{ij}(x)^2 - a_{ij}^2.$$

Um die bekannten Optimierungsroutinen verwenden zu können, ist es wünschenswert obiges Problem (5.5) in ein unrestringiertes Optimierungsproblem zu überführen. Im Falle stetig differenzierbarer Funktionen in (5.5) bietet sich die Methode der Straffunktionen an.

Satz 13 (Kuhn-Tucker Bedingungen) *Sei \bar{x} eine lokale Minimalstelle von (5.5) und die Gradienten $\nabla g_{ij}(\bar{x})$ linear unabhängig, dann gibt es Zahlen $\lambda_{ij} \in \mathbf{R}$, so daß folgende Bedingungen gelten:*

$$KT1 \quad \nabla g(\bar{x}) + \sum_{ij} \lambda_{ij} \nabla g_{ij}(\bar{x}) = 0,$$

$$KT2 \quad g_{ij}(\bar{x}) = 0, \forall (i,j) \in \mathfrak{R}.$$

Die Kuhn-Tucker Bedingungen sind notwendig für das Vorliegen lokaler Minima. Jedoch ist das Auffinden von Punkten, die den obigen Bedingungen genügen, meistens schwierig.

Eine Möglichkeit bietet die sogenannte Strafmethode. Anstelle von (5.5) löst man folgende, unrestringierte Optimierungsprobleme:

$$(P^k) \quad \min_{x \in \mathbb{R}^n} T_k(x) := g(x) + \sum_{(i,j) \in \mathfrak{R}} \sigma_{ij}^k g_{ij}(x)^2, \quad (5.6)$$

wobei $\sigma_{ij}^k \rightarrow \infty$ für $k \rightarrow \infty$.

Diese Form der Straffunktion T_k liefert im Grenzfall $k \rightarrow \infty$ einen Kuhn-Tucker Punkt; vgl. Großmann [11], Seite 177.

Satz 14 (Strafmethode) Sei x^k eine Lösung des Problems (P^k) in (5.6). Definiere

$$\lambda_{ij}^k := 2\sigma_{ij}^k g_{ij}(x^k), \quad (i,j) \in \mathfrak{R}.$$

Dann liefert ein Häufungspunkt $(\bar{x}, \bar{\lambda})$ der Folge (x^k, λ^k) einen Kuhn-Tucker Punkt des Ausgangsproblems (5.5).

Beweis: Notwendig für das Vorliegen eines lokalen Minimums der Straffunktion T^k ist

$$\begin{aligned} 0 &= \nabla T^k(x^k) \\ &= \nabla g(x^k) + \sum_{(i,j) \in \mathfrak{R}} 2\sigma_{ij}^k g_{ij}(x^k) \nabla g_{ij}(x^k) \\ &= \nabla g(x^k) + \sum_{(i,j) \in \mathfrak{R}} \lambda_{ij}^k \nabla g_{ij}(x^k). \end{aligned} \quad (5.7)$$

Für einen Häufungspunkt $(\bar{x}, \bar{\lambda})$ gilt also wegen der Stetigkeit der Funktionen die Bedingung **KT1**.

Da für $g_{ij}(\bar{x}) \neq 0$ die Straffunktion $T^k(\bar{x})$ für unbeschränktes σ^k nicht konvergiert, muß auch die Bedingung **KT2** in \bar{x} erfüllt sein. \square

In der Chemie ist es üblich die Strafparameter σ_{ij}^k konstant zu wählen. Unter vereinfachten, physikalischen Annahmen erlangen sie die Bedeutung von Kraftkonstanten (z.B. durch das Federgesetz von Hook), so wird im Kraftfeld auch eine Veränderung der potentiellen Energie durch Deformationen

beschrieben. Es werden die Nebenbedingungen des Optimierungsproblem es zwar nicht exakt eingehalten, jedoch wird dagegen eine natürliche Flexibilität der Molekülstruktur modelliert, die entscheidend zur Lage des globalen Optimums, gerade bei langkettigen Molekülen, beiträgt.

Die zu minimierende Zielfunktion

$$f(x) := g(x) + \sum_{(i,j) \in \mathfrak{R}} K_{ij} g_{ij}^2(x)$$

setzt sich also aus den Lennard-Jones Wechselwirkungen und gewichteten Distanzgeometriefunktionen zusammen.

5.4 Transformation des Potentials

Um für die Zielfunktion die bekannte Homotopiemethode verwenden zu können, muß noch das Lennard-Jones Potential transformiert werden. Man ist dabei auf Approximationen angewiesen. Hier sollen die einzelnen Möglichkeiten kurz beschrieben werden. Darunter befinden sich auch die Methode der Approximation des Integrales (4.3) von Zhijun Wu [23, 24], die für bestimmte Distanzgeometriefunktionen anwendbar ist, und die Methode der Approximation der Zielfunktion von H. A. Scheraga [20], die nur für ausgesuchte Lennard-Jones Cluster funktioniert. Zunächst soll jedoch als Vergleichsbasis eine möglichst exakte numerische Berechnung der Gaußtransformation vorgestellt werden, die in der Praxis jedoch aufgrund ihrer ungünstigen Rechenzeit noch nicht angewendet wurde. Es steht natürlich immer zur Diskussion, ob man einer exakten Berechnung der Lösung der Diffusionsgleichung den Vorrang geben soll oder einer Approximation, die bei der globalen Optimierung zuverlässige Ergebnisse liefert, da es letztendlich nicht um eine Lösung der Diffusionsgleichung, sondern um die Optimierung einer Molekülstruktur geht.

5.4.1 Numerische Berechnung der Gaußtransformation

Man betrachte im Folgenden nur das Atompaar (X_i, X_j) und lasse zur Vereinfachung die explizite Indizierung ij der Funktionen in den besprochenen Formeln weg.

Um ein abstandsabhängiges Potential $f(x) = H(r(x))$ (4.1) in $[f]_\lambda(x)$ zu transformieren, kann man wieder die Formel (4.3) verwenden:
 Sei nun $\tilde{H} : \mathbb{R}^2 \rightarrow \mathbb{R}$ eine Lösung der Diffusionsgleichung:

$$\frac{\partial^2 \tilde{H}}{\partial r^2}(r, t) = \frac{\partial \tilde{H}}{\partial t}(r, t), \quad \tilde{H}(r, 0) = r H(r). \quad (5.8)$$

Dann gilt mit der Lösungsformel (3.5)

$$\tilde{H}(r, t) = \frac{1}{\sqrt{4\pi t}} \int_{\mathbb{R}} y H(y) e^{-\frac{(y-r)^2}{4t}} dy.$$

Ersetzt man $\lambda^2 = 2t$, so erhält man:

$$\tilde{H}\left(r, \frac{\lambda^2}{2}\right) = \frac{1}{\sqrt{2\pi\lambda}} \int_{\mathbb{R}} y H(y) e^{-\frac{(y-r)^2}{2\lambda^2}} dy.$$

Nun substituiert man $y = r + \sqrt{2}\lambda s$:

$$\tilde{H}\left(r, \frac{\lambda^2}{2}\right) = \frac{1}{\sqrt{\pi}} \int_{\mathbb{R}} (r + \sqrt{2}\lambda s) H(r + \sqrt{2}\lambda s) e^{-s^2} ds.$$

Schließlich wendet man nun Formel (4.3) an, um einen Ausdruck für die Gaußtransformierte des Lennard-Jones Potentials (5.2) zu erhalten:

$$[f]_\lambda(x) = \frac{1}{r(x)} \cdot \tilde{H}\left(r(x), \frac{\lambda^2}{2}\right), \quad (5.9)$$

wobei \tilde{H} eine Lösung der eindimensionalen Diffusionsgleichung (5.8) ist.
 Ein einfaches, explizites Verfahren zum Lösen der Gleichung (5.8) erhält man durch Diskretisierung in den Variablen r und t [34].
 Dabei wird das Zeitintervall $[0, T]$ in konstante Teilintervalle mit der Länge Δt und $[r_a, r_e]$ in Intervalle mit der Länge Δr zerlegt.
 Definiere:

$$\tilde{H}_i^j := \tilde{H}(r_a + i\Delta r, j\Delta t), \quad i = 0, \dots, I \quad j = 0, \dots, J.$$

Eine Approximation der Diffusionsgleichung durch Differenzenquotienten lautet dann mit $s = \Delta t / (\Delta r)^2$:

$$\tilde{H}_i^{j+1} \approx s(\tilde{H}_{i+1}^j + \tilde{H}_{i-1}^j) + (1 - 2s)\tilde{H}_i^j. \quad (5.10)$$

Für $j = 1$ können mit Hilfe der Formel (5.10) aus der Anfangsbedingung (5.8) somit die Werte $\tilde{H}_1^1, \dots, \tilde{H}_{I-1}^1$ berechnet werden, und aus diesen für $j = 2$ die Werte $\tilde{H}_2^2, \dots, \tilde{H}_{I-2}^2$.

Rekursiv können also allgemein für den Zeitpunkt j aus (5.10) die Werte für $\tilde{H}_j^j, \dots, \tilde{H}_{I-j}^j$ berechnet werden.

Es ergibt sich also folgender Algorithmus zur Berechnung des transformierten Lennard-Jones Potentials:

0. Gesucht ist $[f]_\lambda(x)$ bei gegebener Anzahl J an zeitlichen Diskretisierungspunkten (siehe (5.14)) für die numerische Lösung der Differentialgleichung (5.8) und

$$I := 2J$$

Diskretisierungspunkten für r .

1. Setze:

$$T := \lambda^2/2 \quad \Delta t := T/J.$$

Bestimme ein geeignetes Δr , so daß

$$r_a := r(x) - J\Delta r > 0$$

gilt. Berechne die Anfangswerte

$$\tilde{H}_i^0 := (r_a + i\Delta r) \cdot H(r_a + i\Delta r)$$

für $i = 0, \dots, I$ aus dem Lennard-Jones Potential (5.2).

2. Für $j = 1, \dots, J$ berechne jeweils die Werte \tilde{H}_i^j für $i = j, \dots, I - j$ mit Hilfe von (5.10).

3. Gemäß (5.9) gilt :

$$[f]_\lambda(x) \approx \frac{1}{r(x)} \tilde{H}_J^J.$$

Durch dieses Verfahren liegt das transformierte Lennard-Jones Potential nur als Rekursionsformel vor. Zwar können so Funktionswert $[f]_\lambda(x)$ und Gradient $\nabla[f]_\lambda(x)$ numerisch genähert werden, jedoch wird dieses nur durch einen enormen Rechenaufwand möglich. Das nächste Bild zeigt die numerisch ermittelte Glättung des Lennard-Jones Potentials mit $A_{ij} = 3036676.096$ und $B_{ij} = 1482.752$ in Abhängigkeit vom Abstand der Atome und für verschiedene λ .

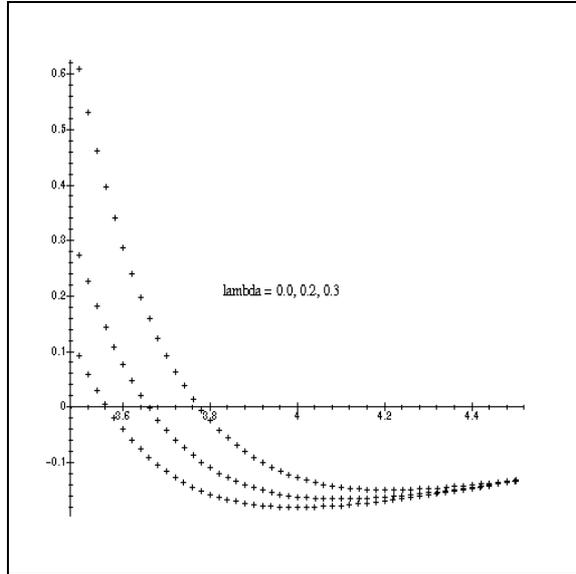


Abbildung 5.2: Numerische Glättung

Ein Punkt, der bei der numerischen Berechnung der Diffusionsgleichung (5.8) mittels (5.10) beachtet werden muß, ist die Stabilität des Verfahrens. Das allgemeine Vorgehen, die Stabilität des Verfahrens (5.10) abzuschätzen, besteht darin, in der Differenzgleichung eine Variablentrennung vorzunehmen [34]. Man geht davon aus, daß die Näherung \tilde{H}_i^j als Produkt zweier Funktionen $R(i)$ und $T(j)$ geschrieben werden kann. Mit Hilfe von (5.10) erhält man dann:

$$\frac{T(j+1)}{T(j)} = 1 - 2s + s \frac{R(i+1) + R(i-1)}{R(i)}. \quad (5.11)$$

Die rechte und die linke Seite müssen hierbei mit einer von i und j unabhängigen Konstanten ζ übereinstimmen. Es ergibt sich:

$$T(j) = \zeta^j T(0),$$

$$\zeta = s \frac{R(i+1) + R(i-1)}{R(i)} + 1 - 2s. \quad (5.12)$$

Mit der Annahme

$$R(i) = (e^{ik\Delta r})^i$$

ergibt sich schließlich für den “Wachstumsfaktor” ζ aus (5.12):

$$\zeta = 1 - 2s + s(e^{ik\Delta r} + e^{-ik\Delta r}). \quad (5.13)$$

Die “von Neumannsche Stabilitätsbedingung” [34] fordert nun für alle Wellenzahlen k :

$$|\zeta(k)| \leq 1 + O(\Delta t).$$

Eine abgeschwächte Form dieser Bedingung ist

$$|\zeta(k)| \leq 1,$$

die ein unkontrolliertes Anwachsen des Faktors $T(j)$ in (5.12) unterbindet. Es gilt bereits $\zeta(k) \leq 1$, aus $\zeta(k) \geq -1$ folgt mit (5.13) sogleich

$$s \leq \frac{1}{2},$$

was gleichbedeutend mit der Forderung ist, daß in (5.10) alle Koeffizienten positiv sind.

Gibt man also im obigen Algorithmus λ und $r(x)$ vor, um $[f]_\lambda(x)$ zu berechnen, so ergibt sich aus den Gleichungen

$$\begin{aligned} \Delta t &= \frac{\lambda^2}{2J}, \\ 0 &< r(x) - J\Delta r, \\ \frac{1}{2} &\geq \frac{\Delta t}{(\Delta r)^2} \end{aligned} \quad (5.14)$$

die Bedingung für die maximale Anzahl J der Diskretisierungspunkte:

$$\left(\frac{r(x)}{\lambda}\right)^2 > J,$$

was die Wahl von J einschränkt und somit auch die Genauigkeit der approximierten Lösung.

Neben den expliziten Verfahren zur Lösung von (5.8) gibt es auch uneingeschränkt stabile, implizite Verfahren, z.B. das Crank-Nicolson Verfahren. Jedoch müssen hierzu in jedem Diskretisierungsschritt Gleichungssysteme gelöst werden, was den Rechenaufwand enorm vergrößert und die Berechnung der Gradienten erschwert.

Es ist vielleicht ratsam, eine analytische Funktion zu suchen, die zumindest approximativ eine Gaußtransformierte des Lennard-Jones Potentials darstellt, so daß der Rechenaufwand bei der Auswertung der Potentialfunktion weniger ins Gewicht fällt.

5.4.2 Approximation der Zielfunktion

Kostrowicki, Piela, Cherayil und Scheraga [20] haben die Homotopiemethode mittels Diffusion bei Lennard-Jones Clustern angewendet. Diese Cluster sind dadurch gekennzeichnet, daß sie durch eine Molekülpotentialfunktion

$$f(x) = \sum_{i>j} H(r_{ij}(x))$$

mit einem vorgegebenen, für alle Atompaaare identischen Lennard-Jones Potential $H : \mathbb{R} \rightarrow \mathbb{R}$ beschrieben werden.

Den “chemisch relevanten” Ausschnitt $[r_a, r_e]$ des Definitionsbereiches dieses Potentials haben sie durch d Gaußglockenkurven angenähert. Die Werte r_a und r_e konnten dabei aus vorangehenden Untersuchungen des Clusters bestimmt werden. Als Approximation ergab sich:

$$\bar{f}(x) = \sum_{i>j} \sum_{k=1}^d a_k e^{-b_k r_{ij}(x)^2}.$$

Diese Art der Approximation ist gut geeignet, um einen analytischen Ausdruck für die Lösung der Diffusionsgleichung (3.4) des Potentials \bar{f} zu erhalten:

$$\bar{F}(x, t) = \sum_{i>j} \sum_{k=1}^d a_k (1 + 8b_k t)^{-3/2} \exp\left(\frac{-b_k r_{ij}(x)^2}{1 + 8b_k t}\right).$$

Dabei werden die Parameter a_k und b_k , so gewählt, daß für genügend großes t_z folgende Gleichung gilt:

$$\sum_{k=1}^d a_k b_k (1 + 8b_k t_z)^{-5/2} = 0, \quad (5.15)$$

so daß die Ableitung der geglätteten Einzelpotentiale nach r_{ij} im Punkt $r_{ij} = 0$ verschwindet.

Diese Bedingung ist gleichbedeutend damit, daß $\bar{F}(x, t_z)$ ein Minimum in $x = 0$ hat. Ausgehend von diesem Punkt wird nun die Homotopiemethode gestartet.

Die Resultate, die mit dieser Methode erzielt wurden, sind sehr gut. Jedoch erschweren drei Probleme die Automatisierung dieser Approximationsmethode bei allgemeinen Molekülen:

- Anders als bei Clustern, gibt es bei Molekülen nicht von vorneherein einen "chemisch relevanten" Bereich $[r_a, r_e]$ für den Abstand zweier ungebundener Atome.
- Bei der Approximation von Lennard-Jones Potentialen durch Gaußglockenkurven, eignen sich nicht alle Parameterpaare A_{ij}, B_{ij} gleich gut. Es garantieren nur solche Koeffizienten a_k, b_k die Existenz eines Minimums in $r_{ij} = 0$ bei t_z , für die

$$\sum_{k=1}^d a_k (b_k)^{-3/2} < 0$$

gilt; vergleiche Scheraga [20].

- Der stetige Übergang von $\bar{F}(x, t)$ in die Funktion $f(x)$ für $t \rightarrow 0$, der für die Homotopiemethode wichtig war, geht durch die Approximation der Zielfunktion verloren, es gilt $\lim_{t \rightarrow 0} \bar{F}(x, t) = \bar{f}(x)$

Neben diesen drei Problemen, die bei der Behandlung von Clustern keine Rolle spielten, jedoch bei allgemeinen Molekülpotentialfunktionen auftreten, gibt es noch einen anderen erwähnenswerten Aspekt.

Normalerweise würde man bei der Glättung eines Lennard-Jones Potentials erwarten, daß das Minimum zu größeren Abständen verschoben wird, wie es in Abbildung 5.2 der Fall ist. Bei der vorgeschlagenen Approximation der Zielfunktion jedoch verlagert sich das Minimum nach $r_{ij} = 0$.

Aufgrund der geschilderten Probleme, besonders bei der Automatisierung des Verfahrens, habe ich mich dazu entschieden, eine Glättungsstrategie für Lennard-Jones Potentiale zu finden, die unabhängig von speziellen Werten für A_{ij} und B_{ij} einen analytischen Ausdruck für eine genäherte Gaußtransformation liefert.

5.4.3 Approximation des Integrales

Schauen wir uns die Formel (4.3) für festes λ und festes $r(x)$ an, dann hat sie folgende allgemeine Gestalt:

$$G_{r,\lambda} = K_r \int_{-\infty}^{\infty} g_{r,\lambda}(s) e^{-s^2} ds.$$

Solche Integrale lassen sich mit Hilfe der Gaußschen Quadraturformel berechnen [12, 25]. Als Integrationsansatz dient eine endliche Summe:

$$\int_a^b g_{r,\lambda}(s) w(s) dx \approx \sum_{\nu=1}^q \omega_\nu g_{r,\lambda}(s_\nu).$$

Da man die Gewichte ω_ν und die Stützstellen s_ν frei wählen kann, ist es möglich, zu fordern, daß alle Polynome mit Höchstgrad $2q-1$ exakt integriert werden.

Führt man eine positive Wichtungsfunktion $w(s)$ in das Integral ein, also hier $w(s) = \exp(-s^2)$, dann ergeben sich aus der obigen Forderung Stützstellen, die Nullstellen eines bestimmten Systems orthogonaler Polynome sind (hier: Hermite'sche Polynome), und als *positive* Gewichte ω_ν erhält man die bestimmten Integrale über gewichtete Legendre-Koeffizienten. Bei symmetrischer Wichtungsfunktion sind die Stützstellen zudem ebenfalls symmetrisch um 0 verteilt.

Insgesamt erhält man:

$$[f]_\lambda(x) \approx [[f]]_\lambda(x) := \frac{1}{\sqrt{\pi} r(x)} \sum_{\nu=1}^q \omega_\nu (r(x) + \sqrt{2}\lambda s_\nu) H(r(x) + \sqrt{2}\lambda s_\nu). \quad (5.16)$$

Satz 15 (Gaußintegration) Sei mit obigen Bezeichnungen $f^{(l)}$ stückweise stetig auf \mathbb{R} für ein $l \leq q$, dann gibt es eine von f unabhängige Konstante μ_l mit

$$|[f]_\lambda(x) - [[f]]_\lambda(x)| \leq \mu_l \lambda^l \sigma(x),$$

wobei

$$\sigma(x) = \sqrt{\int_{-\infty}^{\infty} \exp(-s^2) |f^{(l)}(x + \lambda s)|^2 ds}.$$

Der obige Satz folgt aus einer einfachen Integralabschätzung; vgl. Wu [24] oder Hoffmann [12]. Für die Transformation des Lennard-Jones Potentials ergeben sich bei der Fehlerabschätzung Probleme, da $\sigma(x)$ für diesen Funktionentyp nicht beschränkt ist.

Folgende Tabelle [25] gibt für verschiedene q und für $w(s) = \exp(-s^2)$, $a = -\infty$, $b = \infty$ die Stützstellen und die Gewichte auf 8 Nachkommastellen gerundet an:

	$(-)_s\nu$	ω_ν
$q = 2$	0.70710678	0.88622692
$q = 3$	0.00000000	1.18163590
	1.22474487	0.29540897
$q = 4$	0.52464762	0.80491409
	1.65068012	0.08131284
$q = 5$	0.00000000	0.94530872
	0.95857246	0.39361932
	2.02018287	0.01995324

Abbildung 5.3 stellt das Lennard-Jones Potential mit $A = 3036676.096$ und $B = 1482.752$ und dessen Glättung für $\lambda = 0.3$ und $q = 5$ dar.

Minimum (numerisch)	Wert	“Minimum” (Gauß-Her.)	Wert
4.210	-0.14869	4.215	-0.14790

Nachteil: Die Isotonieeigenschaft der Gaußtransformation geht verloren. Durch die Art der Quadratur werden zusätzliche Polstellen x_0 in die Funktion $[[f]]_\lambda$ eingeführt, so daß es stets eine gegen x_0 konvergierende Folge $(x_n)_{n \in \mathbb{N}}$ gibt mit:

$$\lim_{n \rightarrow \infty} [[f]]_\lambda(x_n) = -\infty.$$

$[[f]]_\lambda$ besitzt kein globales Minimum mehr. Dieses Phänomen wird nur dann numerisch problematisch, wenn $r(x) + \sqrt{2}\lambda s_\nu < 0$ gilt, was für $q = 5$ und $\lambda < 0.7$ gleichbedeutend ist mit: $r(x) < 2.00$ (der Van-der-Waals Abstand zweier aromatischer Wasserstoffatome beträgt 1.9 Å).

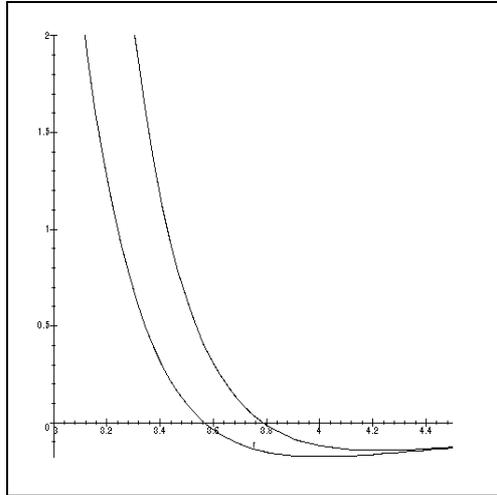


Abbildung 5.3: Gauß-Hermite Integration

Ich habe daher versucht, einen weiteren Approximationsansatz zu finden, der die Isotonieeigenschaft der Gaußtransformation bewahrt. Meine Idee dazu war es, den Ansatz der numerischen Lösung (5.9) zu verwenden, um mit Hilfe einer Approximation der Diffusionsgleichung (5.8) für kleine λ einen analytischen Ausdruck für \tilde{H} zu erhalten.

5.4.4 Approximation der Diffusionsgleichung

Betrachten wir noch einmal den Zusammenhang (5.9) zwischen der Gaußtransformation eines abstandsabhängigen Potentials und einer eindimensionalen Diffusionsgleichung.

Die numerische Lösung der Gaußtransformation in Abbildung 5.2, die sich daraus ergeben hatte, entspricht wiederum als Funktionsbild einem Lennard-Jones Potential, ein möglicher Ansatz für \tilde{H} in (5.8) ist demnach eine Variation der Lennard-Jones Parameter:

$$\tilde{H}(r, t) = \frac{A(r, t)}{r^{11}} - \frac{B(r, t)}{r^5}, \quad A(r, 0) \equiv A, \quad B(r, 0) \equiv B.$$

Mit der Linearität der Gaußtransformation und der Gleichung (5.8) ergeben sich für die Funktionen $A : \mathbb{R}^2 \rightarrow \mathbb{R}$ und $B : \mathbb{R}^2 \rightarrow \mathbb{R}$ folgende partielle Differentialgleichungen:

$$\begin{aligned}\frac{\partial A}{\partial t}(r, t) &= \frac{\partial^2 A}{\partial r^2}(r, t) - \frac{22}{r} \frac{\partial A}{\partial r}(r, t) + \frac{132}{r^2} A(r, t), \\ \frac{\partial B}{\partial t}(r, t) &= \frac{\partial^2 B}{\partial r^2}(r, t) - \frac{10}{r} \frac{\partial B}{\partial r}(r, t) + \frac{30}{r^2} B(r, t).\end{aligned}\quad (5.17)$$

Für kleine t und $r \gg 0$ können wir diese Differentialgleichungen aufgrund der Anfangsbedingungen für A und B vereinfachen, indem wir die partiellen Ableitungen nach r vernachlässigen. Es ergibt sich:

$$\begin{aligned}\frac{\partial A}{\partial t}(r, t) &\approx \frac{132}{r^2} A(r, t), \\ \frac{\partial B}{\partial t}(r, t) &\approx \frac{30}{r^2} B(r, t),\end{aligned}$$

d.h.:

$$\begin{aligned}A(r, t) &\approx A e^{\frac{132}{r^2} t}, \\ B(r, t) &\approx B e^{\frac{30}{r^2} t}.\end{aligned}$$

Zusammen mit (5.9) und $t = \lambda^2/2$ ergibt sich somit für die Gaußtransformierte eines Lennard-Jones Potentials:

$$[f]_\lambda(x) \approx \frac{A e^{66\lambda^2/r(x)^2}}{r(x)^{12}} - \frac{B e^{15\lambda^2/r(x)^2}}{r(x)^6}. \quad (5.18)$$

Da die Exponentialfunktion für kleine Werte von r numerisch nicht mehr zu berechnen ist, sollte man einen Cut-Off Wert, einen kleinsten Abstand, zum Beispiel $r > 1.52$, wählen. Dieser Cut-Off ist numerisch motiviert, anders als bei der Approximation des Integrales, in dem ein theoretischer Cut-Off gesetzt werden muß.

Wie gut ist die Approximation (5.18)?

Um diese Frage zu beantworten, muß man zwei Aspekte untersuchen:

- a) Wie gut ist (5.18) gegenüber den bisher behandelten Approximationsmethoden?

- b) In welchem Rahmen liegt die Abweichung zur tatsächlichen Gaußtransformierten?

Die Frage a) ist leicht zu beantworten. Gegenüber der numerischen Approximation ist die Methode (5.18) weniger rechenintensiv, auch die Berechnung des Gradienten ist weniger aufwendig.

Gegenüber der Approximation der Zielfunktion garantiert (5.18) einen stetigen Übergang zur Ausgangsfunktion und eine Automatisierungsmöglichkeit für beliebige Parameterpaare A, B . Außerdem spiegelt diese Methode selbst für große λ einen richtigen Funktionsverlauf der Glättung wider.

Gegenüber der Approximation des Integrales wird durch (5.18) keine weitere Polstelle eingeschleppt, die Existenz eines globalen Minimums auf jedem Glättungsniveau bleibt gesichert.

Um Frage b) zu beantworten, müßte eine analytische Lösung der Transformaten vorliegen, jedoch soll in diesem Fall ein Vergleich der Approximation mit der numerischen Transformation einen Hinweis auf die Brauchbarkeit liefern.

Ein Beispiel: Folgende Tabelle zeigt für verschiedene Werte von λ die Lage des Minimums r_0 und den Wert des geglätteten Lennard-Jones Potentials mit $A = 3036676.096$ und $B = 1482.752$ einmal für die numerische Approximation mit $J_{max} = 10$ und einmal für die Approximation der Diffusionsgleichung an.

λ	Minimum (num.)	Wert	Minimum (DGL.)	Wert
0.00	4.000	-0.18100	4.000	-0.18100
0.05	4.006	-0.17999	4.006	-0.17999
0.10	4.024	-0.17699	4.024	-0.17702
0.15	4.052	-0.17213	4.052	-0.17228
0.20	4.094	-0.16562	4.092	-0.16605
0.25	4.146	-0.15771	4.140	-0.15867
0.30	4.210	-0.14869	4.198	-0.15048

Es ist deutlich zu sehen, daß diese Approximationsmethode, gemäß ihrer Konstruktion, die Form der Gaußtransformation für kleine Werte λ richtig widerspiegelt. Diese Approximationsmethode hat sich bei den behandelten Beispielen am besten bewährt. Sie wird also nachträglich durch die guten Rechenergebnisse bestätigt.

Kapitel 6

Aggregate

6.1 Probleme in der globalen Optimierung

Besondere Probleme in der globalen Optimierung von Molekülpotentialfunktionen bereiten bis heute die Aggregate, Ansammlungen von einzelnen Molekülen.

Die Optimierung dieser Art von Potentialfunktionen ist jedoch unumgänglich, will man zukünftig Wirkstoffe am Computer designen. Denn hier spielen nicht strukturinterne Größen eine Rolle, sondern vielmehr das Wechselspiel zwischen den einzelnen Bausteinen, den Molekülen.

Ebenso kann meiner Meinung nach die vorgeschlagene Einführung sogenannter *Pseudobindungen*, um die Aggregate künstlich in isolierte Moleküle zu überführen, nicht die Vielfalt an Möglichkeiten ausschöpfen, die frei bewegliche Moleküle in einem Aggregat bieten.

Die folgenden zwei Beispiele befassen sich deshalb mit Aggregaten, ohne die Einführung von Pseudobindungen. Wenn auch die gebrachten Beispiele aufgrund der Komplexität von Proteinen, noch relativ einfache Aggregate behandeln, sind doch schon die grundlegenden Fragestellungen der globalen Optimierung im Wirkstoffdesign impliziert.

Ich möchte auch darauf hinweisen, daß die beiden Beispiele mit 51 bzw. 42 optimierbaren Freiheitsgraden (eindeutige kartesische Koordinaten) eine hohe Anzahl von Variablen behandeln, die eine Betrachtung mit Hilfe von Branch-and-Bound Techniken nicht erlauben.

6.2 Rechnungen

6.2.1 Acrolein und Wasser

Das erste Rechenbeispiel betrifft eine Gruppe von Molekülen. Zwei Acroleinmoleküle sollen dabei mit einem Wassermolekül wechselwirken. Es soll festgestellt werden, wie sich diese Moleküle gegeneinander ausrichten können, so daß möglichst wenig potentielle Energie im System verbleibt.

Folgende Tabelle gibt die Partiaalladungen der beiden Acrolein- und des Wassermoleküls an. Die Partiaalladungen entnimmt man einer AM1 Rechnung mit der sogenannten Coulson Ladungsdichte, die nachträglich so skaliert wird, daß das empirisch ermittelte Dipolmoment der Fragmente mit dem errechneten übereinstimmt.

Molekül	Atom	Ladung [e]
Acrolein	O	-0.2905
	C1	0.2088
	C2	-0.2616
	C3	-0.1383
	H1	0.0875
	H2	0.1517
	H3	0.1236
	H4	0.1187
Wasser	O	-0.6600
	H	0.3300

Gegeben ist auch die metastabile, planare Struktur aus Abbildung 6.1. Die Aufgabe des Optimierers besteht darin, dieses relative Minimum zu verlassen und eine bessere Struktur zu errechnen. Die abgebildete Geometrie wurde im MOBY-Eingabeformat von Herrn Höweler zu Testzwecken zur Verfügung gestellt.

Aus den Atomtypen geht hervor, daß Acrolein ein System konjugierter Doppelbindungen enthält, was eine relativ starre Molekülgeometrie verursacht. Unter diesen Umständen kann man die Geometrie der beiden Acroleinmoleküle und die des Wassers jeweils durch ein Distanzgeometrieproblem beschreiben. Aus der MOBY-Datei werden dabei alle paarweisen Abstände der Atome eines Moleküles extrahiert und in eine Distanzgeometriefunktion der Form (4.5) umgewandelt.

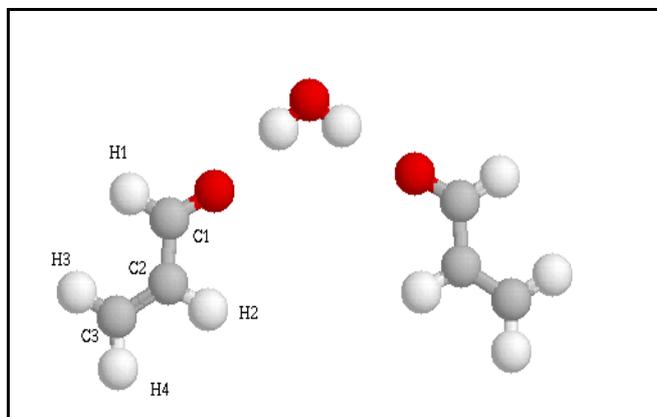


Abbildung 6.1: Acrolein und Wasser

Die paarweisen Wechselwirkungen zwischen den drei Molekülen werden durch Lennard-Jones und Coulomb Potentiale berechnet.

Um das Lennard-Jones Potential in der Form (5.2) verwenden zu können, benutzt man die Relationen: $A_{ij} = K_{ij}r_{0,ij}^{12}$ und $B_{ij} = 2K_{ij}r_{0,ij}^6$.

Der Gleichgewichtsabstand $r_{0,ij}$ ist die Summe der Van-der-Waals Radien und die Stabilisierungsenergie K_{ij} ist das geometrische Mittel aus den Van-der-Waals Energien der Atome X_i und X_j .

Folgende Tabelle gibt die erforderlichen Größen aus dem AMBER-Kraftfeld [16] an:

Atom	VdW Radius [Å]	VdW Energie [kcal/mol]
O (Acrolein)	1.600	0.2000
O (Wasser)	1.650	0.1500
H (Acrolein)	1.375	0.0380
H (Wasser)	1.000	0.0200
C	1.850	0.1200

Um den Algorithmus aus Abschnitt 4.3 verwenden zu können, benötigen wir einen Startwert L für den maximalen Wert von λ . Aus Formel (4.7) erhält man z.B. $L := 0.4$, da die Bindungslängen im Wasser 0.96Å betragen.

In 100 gleichgroßen Intervallen läßt man nun λ_k , für $k = 0, \dots, 100$ von $\lambda_0 := L$ nach $\lambda_{100} = 0$ streben und wendet zur Glättung der Funktionen die

Formeln (4.6), (5.18) und (5.4) an. Als Startgeometrie bieten sich die in der obigen Datei gegebenen Testkoordinaten an.

Das Ergebnis ist eine Schichtstruktur. Folgende Abbildung zeigt diese Struktur. Sie ist von dem Programm RasMol [33] erzeugt worden und reflektiert die ungefähre Raumauffüllung der einzelnen Atomsorten bezogen auf den Van-der-Waals Radius.

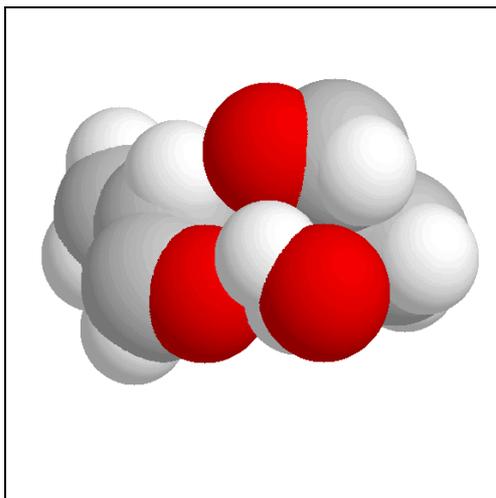


Abbildung 6.2: Acrolein- und Wasser: Ergebnisstruktur

Bei dieser Struktur lagern sich die negativ geladenen Sauerstoffatome jeweils zwischen zwei positiv geladenen Wasserstoffatomen an. Außerdem bildet sich ein “Molekülknubbel”, eine dichte Anordnung der einzelnen Moleküle, so daß 21 Van-der-Waals Kontakte zwischen den Molekülen gemessen werden konnten.

Dieses gibt einen Hinweis darauf, wie sich die Bausteine möglicherweise anordnen könnten. Das Beispiel “Acrolein- und Wasser” ermutigt zu weiteren Computorexperimenten mit der Homotopiemethode.

Die Struktur aus Abbildung 6.2 wurde mit einem Energiewert von $-11.764[kcal/mol]$ im vereinfachten Kraftfeld berechnet: Im Gegensatz zum AMBER-Kraftfeld fehlen hier Terme für Wasserstoffbrückenbindungen, Coulombabschirmung, Torsions- und Out-of-plane Spannungen.

Zwar reicht das vereinfachte Kraftfeld aus, um Hinweise auf die Optimalstruktur zu geben, jedoch müssen die Rechnungen nicht mit anderen Kraftfeldern übereinstimmen.

Diese Tatsache gilt auch bei dem Vergleich von dem AMBER-Kraftfeld mit anderen Potentialfunktionen. Es ist also stets entscheidend anzugeben, mit welcher Potentialfunktion man die Optimalstruktur einer Molekel berechnet hat.

Gleichzeitig wird ein Vorteil der Homotopiemethode deutlich:

Die Homotopiemethode kommt ohne den Begriff der *Pseudobindungen* aus, die man in diesem Beispiel rein intuitiv zwischen den Wasserstoffatomen des Wassers und den Sauerstoffatomen der Acroleinmoleküle setzen würde; vgl. Bollweg [4]. Das Ergebnis der Optimierung zeigt in Abbildung 6.2 jedoch ein von der Intuition abweichendes Verhalten auf. Hier findet ein besonders günstiger Ausgleich der Ladungstrennung durch die Wechselwirkung von je einem negativ geladenen Sauerstoffatom mit zwei positiv geladenen Wasserstoffatomen statt. Eine solche Möglichkeit konnte man aus der Startstruktur nicht ersehen. Das Festlegen einer erwarteten Kontaktstelle oder gar einer erwarteten Symmetrie hätte hier also das Modell zu stark eingeschränkt. Für die Homotopiemethode gilt: Die große Vielfalt möglicher Anordnungen der einzelnen Bausteine zueinander bleibt erhalten und muß nicht intuitiv durch das Festlegen von Pseudobindungen geraten werden.

Bemerkung Acrolein und Wasser ist ein einfaches Beispiel für die Fragestellungen des Wirkstoffsdesigns, bei dem es um die geometrische Anordnung wechselwirkender Moleküle geht. Zum einen müssen die beteiligten Moleküle besonders viele stabilisierende Van-der-Waals Wechselwirkungen eingehen und gleichzeitig entgegengesetzt geladene Atome zueinander anordnen.

6.2.2 Formaldehyd

Ein weiteres, schönes Beispiel stellt eine Anordnung aus vier Formaldehydmolekülen CH_2O dar. Abbildung 6.3 zeigt ein solches Molekül, das wiederum planar ist.

Auch hier kann wieder eine Populationsanalyse vorgenommen werden.

Kunz [21] (Seite 142) errechnete nach Mulliken mit GAMESS und einem 6-31G* Basisansatz (die Abkürzung gibt an, auf welche Weise die Zustandsfunktionen in der Schödingergleichung approximiert werden [29]) Werte für die Partialladungen des Formaldehyds in [e], die in der nachstehenden Tabelle zusammengefaßt sind.

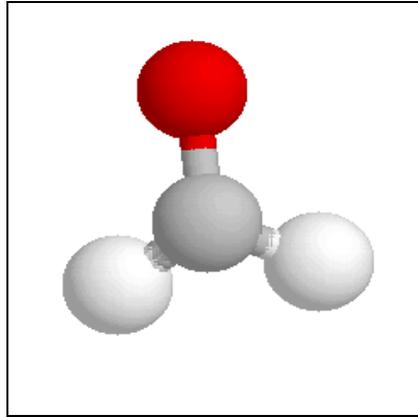


Abbildung 6.3: Formaldehyd

Atom	Ladung
O	-0.207548
C	0.092179
H	0.057685

Mit Hilfe dieser Ladungsverteilung kann man analog zu dem Beispiel “Acrolein und Wasser” ein globales Optimierungsproblem formulieren.

Die Distanzen innerhalb der Formaldehydmoleküle werden als Distanzgeometrieproblem formuliert, die Lennard-Jones Parameter errechnet man für die Atomtypen des Acroleins aus der Tabelle Seite 69.

Als Eingabestruktur wurde eine Geometrie gewählt, in der die Formaldehydmoleküle in gleiche Richtung zeigen und in einer Ebene liegen.

Eine Optimierung in 200 Schritten mit dem Startwert $L := 0.48$ liefert jedoch mit dieser Ausgangsstruktur folgendes Ergebnis:

Es ergibt sich eine hochsymmetrische, ringförmige Anordnung der Moleküle. Wieder ist ein dichter Molekülknubbel optimal. Die Sauerstoffatome der Formaldehydmoleküle weisen jeweils im Wechsel in die eine und in die entgegengesetzte Richtung, so daß auch die Coulomb Anziehung optimal ausgeglichen wird.

Die errechnete Energie beträgt $-7.296[kcal/mol]$.

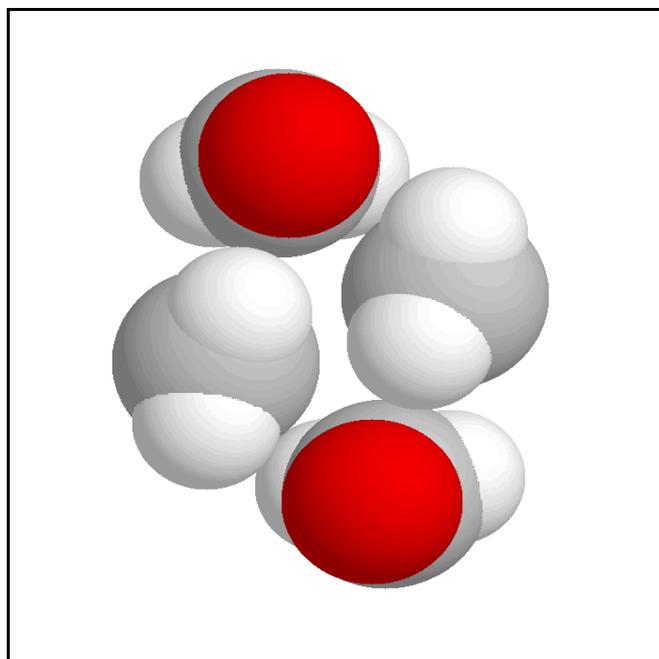


Abbildung 6.4: Formaldehyd: Optimalstruktur

6.3 Resultat

Die obigen Beispiele zeigen die Möglichkeiten der globalen Optimierung mittels der Gaußtransformation als Homotopiemethode auf.

Im ersten Beispiel gelangte man von einer metastabilen Molekülgeometrie, die ein lokales Minimum auf der Potentialfläche darstellt, zu einer Struktur, die besonders viele stabilisierende Wechselwirkungen zeigt.

Im Gegensatz zur lokalen Optimierung kann also die Homotopiemethode aus einem gegebenen Minimum heraus zu anderen Minima finden.

Das zweite Beispiel zeigt, daß sich auch vorgegebene Ausrichtungen der Moleküle während des Algorithmus ändern können. Die hohe Symmetrie der Ergebnisstruktur liefert einen starken Hinweis auf globale Optimalität.

Formaldehyd stellt eine große Herausforderung an einen Optimierer dar, einerseits kann die Optimierung nicht in begrenzten internen Koordinaten geschehen, da es sich um vier frei bewegliche Moleküle handelt, andererseits ist die vorgegebene Struktur "weit" von dem globalen Optimum entfernt, so

daß Optimierer, die auf der Grundlage “punktuelle Informationen aus der Zielfunktion” arbeiten, Schwierigkeiten haben werden, bis zu dieser Struktur zu gelangen.

Was nun im Laufe der Homotopiemethode passiert, ist leicht zu verstehen: Die Form der Glättung läßt die einzelnen Moleküle der Aggregate “schrumpfen”, gleichzeitig wachsen deren Van-der-Waals Oberflächen und eine vorgegebene Ladungstrennung wird relativiert.

Für große Glättungsparameter ergeben sich also ungeladene “Kugeln” anstelle der Moleküle. Diese ordnen sich im ersten Schritt der Optimierung zu einem dichten Haufen an. Erst nach und nach bekommen die “Kugeln” eine definierte Struktur und eine definierte Ladungsverteilung, so daß sie sich dann optimal ausrichten können.

Dieses gibt der Homotopiemethode eine gewisse, chemische Relevanz, wenn auch die geglättete Funktion nicht als eine Potentialfunktion gedeutet werden darf.

Der beschriebene Sachverhalt wird in Abbildung 6.5 schematisch dargestellt, als Beispiel dient ein zweiatomiges Molekül mit entgegengesetzt geladenen Atomen.

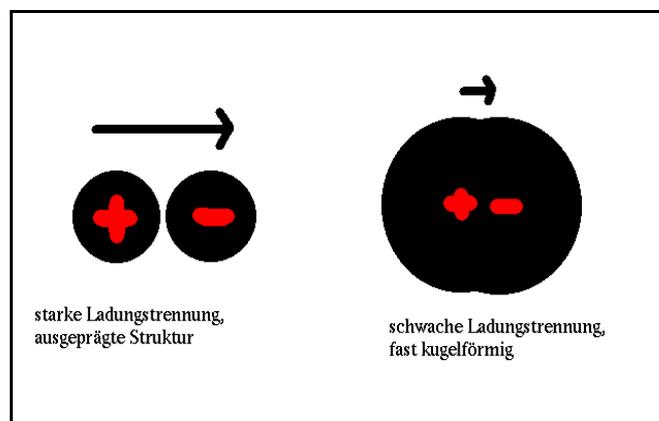


Abbildung 6.5: Wirkungsweise der Homotopiefunktion

Durch die Glättung des Coulomb-Potentials wird die Ladungstrennung in den einzelnen Molekülen relativiert.

Das Coulomb-Potential ist ein Beispiel für die Mehrdeutigkeit der Lösungen der Diffusionsgleichung: Für eine Zielfunktion $f : \mathbb{R}^n \rightarrow \mathbb{R}$, die aus einer

Summe von Coulomb Potentialen besteht, gilt nämlich für alle $x \in \mathbb{R}^n$ die Beziehung $\Delta f(x) = 0$. Bei einer Diffusionsmethode müßte man die Coulombanteile gar nicht glätten. Jedoch habe ich mich aus den oben genannten Gründen für eine Glättung dieser Potentialterme entschieden.

Kapitel 7

Isolierte Moleküle

Dieses Kapitel zeigt, welche Probleme auftreten, wenn man die Ergebnisse der Optimierung von Aggregaten mit festen Bausteinen auf flexible Molekülstrukturen übertragen will.

Auch diese Problematik tritt im Wirkstoffdesign auf. Hier wird als Lösungsansatz eine neue Form des Torsionspotentials vorgestellt.

7.1 Trichlordifluorbutan

Während man bei starren Molekülgeometrien wie im Beispiel des Acroleins und des Wassers eine Populationsanalyse zur Errechnung der Partialladungen in den Molekülen durchführen kann, ist dieses bei flexiblen Molekülen nicht möglich, da sich die relative Kernanordnung erst nach der Optimierung ergibt und die Zustandsfunktion Ψ zur Errechnung der Partialladungen von diesen Koordinaten abhängt.

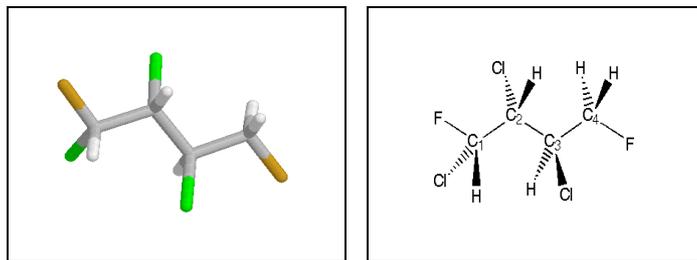


Abbildung 7.1: 1,2,3-Trichlor-1,4-difluorbutan

Aus diesem Grunde werden Optimierungen einer flexiblen Molekülgeometrie häufig ohne die Beachtung stabilisierender oder destabilisierender Coulomb-Wechselwirkungen durchgeführt. Ein einfaches Beispiel eines flexiblen Moleküles ist 1,2,3-Trichlor-1,4-difluorbutan.

In diesem Molekül sind nur die Bindungsabstände und die Bindungswinkel jeweils als konstant anzusehen. Sie können daher wieder durch eine Distanzgeometriefunktion der Form (4.5) ausgedrückt werden. Da die vier Kohlenstoffatome einen tetraedrischen Bindungspolyeder bevorzugen, beträgt der Bindungswinkel jeweils 109.5° .

Die Bindungsabstände ergeben sich aus den Kovalenzradien der beteiligten Atome und ihren Elektronegativitäten, sowie der Bindungsordnung, die in diesem Beispiel stets einfach ist.

Folgende Tabelle [31] gibt die entsprechenden Größen für die beteiligten Atomsorten wieder:

Atom	Kov.rad. r_X [pm]	Elektroneg. χ_X
H	32	2.20
C	77	2.55
Cl	99	3.16
F	71	3.98

Mit Hilfe der Formel [31]:

$$a_{XY} = r_X + r_Y - 8|\chi_X - \chi_Y| [pm]$$

ergeben sich somit folgende Bindungsabstände:

Bindung	Länge [Å]
C - C	1.54
C - H	1.06
C - Cl	1.71
C - F	1.37

Die Lennard-Jones Parameter werden diesmal der Literatur [18] entnommen:

(A,B)	H	C	Cl	F
H	7220, 76.0	37430, 127.4	80740, 272.8	9361, 79.8
C		285800, 372.5	646400, 759.3	100700, 223.2
Cl			2005000, 1562	296700, 457.6
F				36520, 134.6

Diese Parameter sind mit dem Faktor 0.05 zu multiplizieren, wenn die entsprechenden Atome in einer sogenannten 1,4-Beziehung zueinander stehen, wie dieses im Butan z.B. bei den beiden äußeren Kohlenstoffen der Fall ist. Dieser Faktor repräsentiert die Abschirmung der Van-der-Waals Wechselwirkung durch die Elektronen, die sich zwischen den Atomen befinden.

Es wird wieder der Algorithmus aus Abschnitt 4.3 verwendet, wobei diesmal der Distanzgeometrieanteil mit dem Faktor 50 gewichtet wird, um die Starrheit der Bindungslängen und Bindungswinkel stärker gegenüber den Lennard-Jones Potentialen zu berücksichtigen.

Eine Multi-Start Optimierung ergibt folgendes globales Optimum:

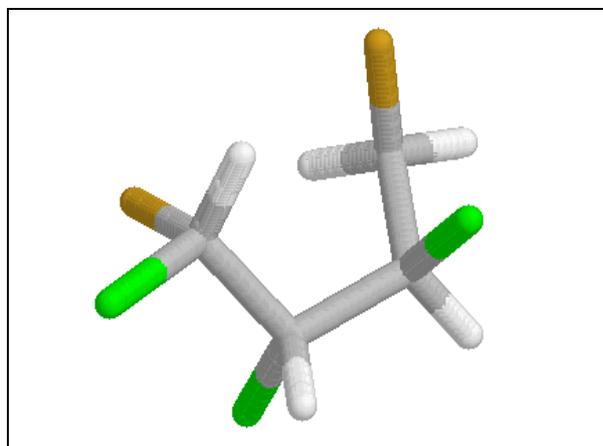


Abbildung 7.2: 1,2,3-Trichlor-1,4-difluorbutan: Optimalstruktur

Hier zeigt sich ein entscheidender Unterschied der Homotopiemethode zu den Optimierern, die in internen Koordinaten rechnen.

Betrachtet man nur den Bindungspolyeder des zweiten Kohlenstoffatoms, einmal in Abbildung 7.1 und einmal in 7.2, dann sieht man, daß das Chlor- und das Wasserstoffatom "ihre Plätze getauscht" haben. Der Bindungspolyeder in Abbildung 7.2 ist ein Spiegelbild des Bindungspolyeders in Abbildung 7.1.

Die einzigen Freiheiten, die das Molekül jedoch hat, sind die Drehungen der Bindungspolyeder um die C-C-Bindungsachsen. Durch Variation dieser sogenannten *Torsionswinkel* wird die globale Optimierung in internen Koordinaten bewerkstelligt.

In der Homotopiemethode können jedoch nur intermolekulare Abstände kontrolliert und bewertet werden; es sind Spiegelsymmetrien möglich. Das Molekül kann dadurch in Geometrien überführt werden, die durch Variation seiner eigentlichen Freiheitsgrade physikalisch nicht erreicht werden.

Das Molekül aus Abbildung 7.2 ist zwar energetisch günstiger als das Molekül in Abbildung 7.1, jedoch kann die eine Form nicht durch Variation der Torsionswinkel in die andere überführt werden.

Die Homotopiemethode ist nicht geeignet, bei diesem Molekül die Frage nach dem globalen Minimum zu beantworten, welches auch tatsächlich ausgehend von der Startstruktur aus Abbildung 7.1 physikalisch erreicht werden kann. Dieses Ergebnis ist auf alle Moleküle übertragbar, deren Spiegelbild nicht durch eine Variation ihrer Freiheitsgrade erzeugt werden kann.

Bemerkung: Beim Acrolein und beim Wasser kann jeweils das Spiegelbild durch eine 180° Drehung erzeugt werden, da diese Moleküle planar sind und somit in der Bindungsebene eine Spiegelfläche enthalten.

Cluster, das sind “Moleküle”, in denen nur Lennard-Jones und Coulomb Wechselwirkungen auftreten, sind ein Beispiel für eine Klasse optimierbarer Systeme. Sie wurden bereits von Scheraga [20] untersucht. Eine andere Klasse bilden die sogenannten Alkane.

Der Chemiker nennt den Effekt, der beim Trichlordifluorbutan auftritt *Stereoisomerie*, der am häufigsten durch ein sogenanntes *asymmetrisches* Kohlenstoffatom (ein C-Atom mit vier verschiedenen Bindungspartnern) verursacht wird.

Eine Variation der stereoisomerischen Eigenschaften eines isolierten Moleküles läßt eine weitere Fragestellung an ein gegebenes Molekül zu: Welches Stereoisomer von Trichlordifluorbutan ist energetisch zu bevorzugen?

Betrachten wir jedoch zunächst noch einmal die Optimalstruktur. Eine optimale Molekülstruktur zeichnet sich nicht nur durch ihre stabilisierenden Wechselwirkungen ungebundener Atome aus, sondern auch durch möglichst geringe innerer Spannungen, die durch die Verletzung der optimalen Bindungswinkel und Bindungslängen erzeugt werden. Im Falle der Struktur aus Abbildung 7.2 liegen die Werte für die C-C-Bindungen jeweils bei 1.54 \AA . Die H-C-Bindungen haben Längen zwischen 1.06 und 1.07 \AA , während die Cl-C- und die F-C-Bindungen wiederum den Sollwerten entsprechen. Auch die

Bindungswinkel liegen mit 109.27° bis 109.86° in dem vorgegebenen Rahmen. Dennoch wird ein Chemiker schnell sehen, daß diese Molekülstruktur nicht optimal sein wird. Es ist ein weiterer Potentialterm im Falle der isolierten Moleküle nötig, der die Abstoßung der Bindungselektronen berücksichtigt.

7.2 Verminderung der Freiheitsgrade

7.2.1 United-Atom Modellierung

Möchte man die Gaußtransformation als Homotopiefunktion zur globalen Optimierung von Molekülpotentialfunktionen nutzen, so wird man alle kartesischen Koordinaten eines Moleküles zur Modellierung der Molekülgeometrie verwenden.

Eine Verminderung an Freiheitsgraden kann nur durch eine Verminderung der Zentrenzahl geschehen.

Bei diesem Vorgehen nutzt man die sogenannte *United-Atom* Modellierung. Eine Gruppe von Atomen wird dabei durch ein einziges "Atom" repräsentiert. Die Van-der-Waals Oberfläche der Atomgruppe werden approximiert durch eine Kugel; die drei kartesischen Koordinaten des Mittelpunktes geben die räumliche Position dieser Gruppe an.

Alkane C_nH_{2n+2} sind gesättigte Kohlenwasserstoffketten. An eine Kette von Kohlenstoffatomen werden so viele Wasserstoffatome gebunden, daß jedes Kohlenstoffatom vier Bindungspartner besitzt. In der United-Atom Näherung wird dabei nur die Kette von Kohlenstoffatomen berücksichtigt.

Die C-C Bindungslänge beträgt jeweils 1.562\AA , der Bindungswinkel 109.47° ; vgl. Bollweg [4]. Nur die Bindungslängen und Bindungswinkel werden durch ein Distanzgeometrieproblem beschrieben, wobei dieses gegenüber den Lennard-Jones Wechselwirkungen bei isolierten Molekülen mit dem Faktor 50.0 aufgewertet wird.

Der Van-der-Waals Radius beträgt 4.0\AA und die Gleichgewichtsenergie beträgt $0.181[\text{kcal/mol}]$. Daraus ergeben sich als Lennard-Jones Parameter:

$$A = 3036676.096 \quad B = 1482.752,$$

diese werden wiederum mit dem Faktor 0.05 multipliziert, falls die entsprechenden Atome in einer 1,4-Beziehung zueinander stehen.

Eine Multi-Start Auswertung der entsprechenden Potentialfunktion für Hexan mit sechs C-Atomen, liefert einen Optimalwert von $-0.537[kcal/mol]$.

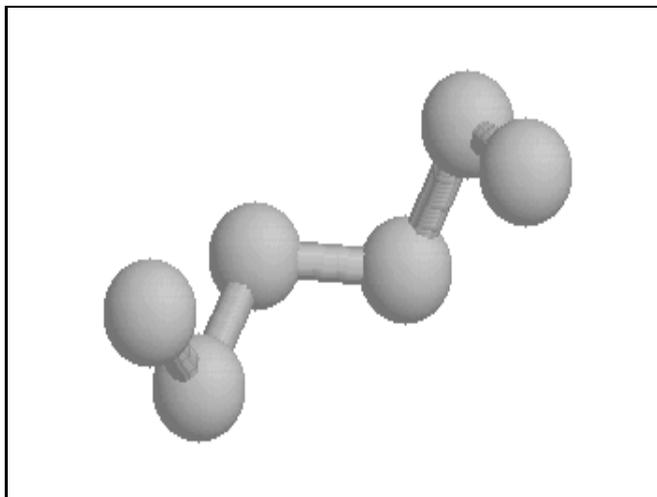


Abbildung 7.3: Hexan: vereinfachtes Kraftfeld

Das zugehörige Bild zeigt eine “Haarnadel”-Struktur. Ausgehend von der zentralen C-C Bindung laufen zwei Stränge von C-Atomen nebeneinander in eine Richtung. Diese Struktur bildet sich normalerweise erst bei längeren Alkanketten aus (ab $C_{18}H_{38}$); vgl. Goodman [8].

Der Grund für die Fehleinschätzung liegt wieder in der vernachlässigten Abstoßung der Elektronenwolken des Bindungsgerüsts.

Diese Abstoßung muß durch weitere Summanden in dem vereinfachten Kraftfeld berücksichtigt werden.

7.2.2 Torsionspotentiale

Bei Experimenten hat sich gezeigt, daß *vollständig freie Rotationen* eines Bindungspolyeders um eine C-C Einfachbindung möglich sind; vgl. Hart [13]. Die einzelnen räumlichen Anordnungen, die dabei eingenommen werden, heißen *Konformationen* des betreffenden Moleküles.

Obwohl die Umwandlung der Konformationen leicht (bei Raumtemperatur) vor sich geht, sind diese energetisch unterscheidbar.

Die energieärmste Konformation ist die sogenannte *gestaffelte*.

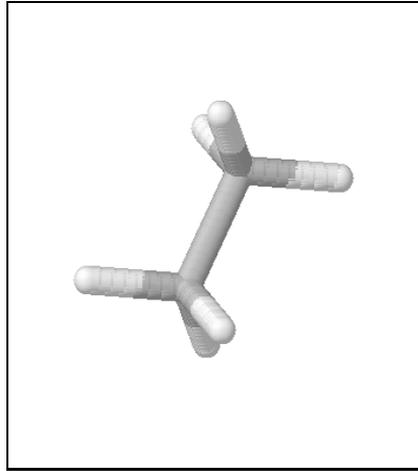


Abbildung 7.4: Ethan: gestaffelte Konformation

Bei der gestaffelten Konformation sind die Bindungstetraeder der Kohlenstoffatome um 60° beziehungsweise 180° gegeneinander verdreht.

Dieser Winkel wird als *Torsionswinkel* bezeichnet.

Verschiedene Ansätze wurden probiert, um diese energetischen Unterschiede in einer mathematischen Form als zusätzlichen Term in die Kraftfelder aufzunehmen; vgl. zum Beispiel das AMBER-Kraftfeld [35].

Diese Ansätze sind jedoch für die Gaußtransformation nicht geeignet, da die entsprechenden Potentialterme nicht den Abstand zweier Atome, sondern den Wert des Torsionswinkels beinhalten.

Aus Abbildung 7.4 ist jedoch zu erkennen, daß ein Wasserstoffatom, das an eines der Kohlenstoffatome gebunden ist, jeweils zwei verschiedene Abstände zu einem Wasserstoffatom, das an das andere Kohlenstoffatom gebunden ist, einnehmen kann.

Seien diese Abstände mit d_1 und d_2 bezeichnet, dann ist ein möglicher Strafterm für die Einhaltung der gestaffelten Konformation:

$$H_{ij}(r_{ij}(x)) = (r_{ij}(x)^2 - d_1^2)^2 (r_{ij}(x)^2 - d_2^2)^2, \quad (7.1)$$

wobei $(i, j), i > j$ jeweils die Paare von Atomen durchläuft, die in einer 1,4-Beziehung zueinander stehen.

Da H_{ij} ein Polynom achten Grades in r_{ij} ist, kann man $r_{ij} \cdot H_{ij}(r_{ij})$ mit Hilfe der Gauß-Hermite Integration mit $q = 5$ Stützstellen transformieren; siehe

dazu Seite 62. Nach den vorangehenden Überlegungen sind wir jetzt in der Lage, das gesamte, vereinfachte Kraftfeld für die Optimierung eines isolierten Moleküles zu formulieren.

$$\begin{aligned}
 f(x) = & \sum_{\text{Bindungen, Winkel}} K_{ij} (r_{ij}^2(x) - a_{ij}^2)^2 + \\
 & \sum_{\text{Lennard Jones}} \left(\frac{A_{ij}}{r_{ij}(x)^{12}} - \frac{B_{ij}}{r_{ij}(x)^6} \right) + \\
 & \sum_{\text{Torsionspotentiale}} T_{ij} (r_{ij}(x)^2 - d_{1,ij}^2)^2 (r_{ij}(x)^2 - d_{2,ij}^2)^2 \quad (7.2)
 \end{aligned}$$

7.3 Vergleich: Homotopiemethode und Multi-Start

7.3.1 Effizienz

Als erstes Beispiel rechnet man wieder Hexan mit dem zusätzlichen Potentialterm (7.1) für die Torsionswinkel. Aus Formel (4.7) bestimmt man als Startwert L für den Glättungsparameter $L \leq 0.68$. In 200 Schritten wird die Trajektorie der Minima zurückverfolgt.

Folgende Tabelle zeigt den maximalen Glättungsparameter L , die Anzahl energetisch unterscheidbarer Ergebnisstrukturen nach 20 Durchläufen des Homotopiealgorithmus von jeweils 20 zuvor zufällig erzeugten Startstrukturen, den tiefsten Energiewert der Ergebnisstrukturen und die Anzahl von Treffern auf das globale Minimum.

Startwert L	Anzahl der Minima	bester Wert [kcal/mol]	Treffer
0.68	4	-0.196	10
0.54	4	-0.196	9
0.40	4	0.029	0
0.27	6	-0.108	0
0.14	7	-0.108	0
0.00	8	-0.196	1

In der optimalen Struktur mit dem Energiewert von $-0.196[kcal/mol]$ sind alle Torsionswinkel auf 180° eingestellt, wie es auch der Literaturangabe [8] für die Optimalstruktur von Hexan entspricht. Die Homotopiemethode führt also im Falle des Hexans zu einem überprüfbar Ergebnis und zeigt, daß das spezielle Torsionspotential im Falle eines kurzkettigen Alkanmoleküles brauchbar ist.

Wie zu erwarten, steigt die Zahl der erreichten Minima für kleiner werdende Startparameter L an. Jedes Minimum wird daher im Durchschnitt seltener aufgesucht.

Voraussetzung für eine hohe Zuverlässigkeit der Homotopiemethode ist es also, daß einerseits nur wenige Minima der Potentialfunktion nach der Glättung verbleiben und daß das globale Minimum in der Menge der erreichbaren Minima enthalten ist.

Im Schnitt führt jeder zweite Durchlauf der Routine zum globalen Minimum, wenn man $L = 0.68$ wählt. Bei Werten über 0.68 ist das Verhalten der Routine, wahrscheinlich aufgrund der Ausführungen zu (4.7), nicht so gut zu interpretieren. Die Zahl der unterscheidbaren Minima nimmt teilweise wieder zu ($L = 0.99$, Zahl der Minima: 5), die Zahl der Treffer auf das globale Minimum schwankt ebenfalls.

Da die 20 Startstrukturen in der obigen Tabelle für jeden Wert von L übereinstimmen, zeigt diese wiederum, daß die Homotopiemethode in der Lage ist, anders als die lokale Optimierung, aus dem Einzugsgebiet eines bestimmten Minimums in ein anderes Minimum zu gelangen. In diesem Fall ist das durch Homotopie erreichte Minimum fast ausschließlich besser.

Ein Vergleich der Rechenzeiten für $L_1 := 0.68$ und $L_2 := 0.0$ (also zwischen der Homotopiemethode und dem reinen Multi-Start) ergibt auf einer LINUX-Workstation mit einem Pentiumchip und 100MHz Systemfrequenz eine Rechenzeit für 20 Durchläufe von 209.7 Sekunden bzw. 95.2 Sekunden mit den behandelten (nicht zeitoptimalen) lokalen Minimierungsmethoden. Der reine Multi-Start Algorithmus ist also nur etwas mehr als doppelt so schnell, jedoch weniger effizient.

Der Homotopiealgorithmus ist eine effizientere Optimierungsmethode. Auch zu beachten ist, daß durch die große Anzahl von Optimierungsvariablen neben den *weichen* Freiheitsgraden des Moleküles (den Torsionswinkeln) durch die Betrachtung des Problems in kartesischen Koordinaten zusätzlich auch eine gewisse Flexibilität des Moleküles in seinen *harten* Freiheitsgraden (den Bindungslängen und -winkeln) erhalten wird. Zur globalen Optimierung von

Molekülpotentialfunktionen werden also nicht nur *starre Rotationen* herangezogen, sondern auch Energiegewinne durch leichte Deformationen berücksichtigt. Diese können bei größeren Molekülen für das globale Optimum entscheidend sein.

7.3.2 Van-der-Waals Abstoßungen

Als zweites Beispiel für die United-Atom Näherung möchte ich nocheinmal auf die Optimierung von Trichlordifluorbutan zurückgreifen.

Wir sind nun in der Lage, die fehlende Abstoßung der Bindungselektronen zu modellieren, denn Trichlordifluorbutan ist ein Derivat eines kurzkettigen Alkanmoleküles. Auch kann man durch die United-Atom Näherung die fünf Wasserstoffatome des Moleküles vernachlässigen, indem man die entsprechenden Parameter aus dem Alkanmodell übernimmt, d.h den Van-der-Waals Radius und die entsprechende Gleichgewichtsenergie der Kohlenstoffatome anpaßt und die geänderten C-C Bindungslängen und C-C-C Bindungswinkel übernimmt.

Führt man wieder eine vergleichende Minimierung mit der reinen Multi-Start Optimierung und mit der Homotopiemethode durch, so ergibt sich ein anderes Bild als bei den obigen Beispielen.

Der Grund liegt in den Van-der-Waals Abstoßungen. Diese Abstoßung tritt dann auf, wenn der Wert des Lennard-Jones Potentials für ein Atompaar positiv wird.

Folgende Tabelle zeigt beispielhaft 10 Durchläufe des reinen Multi-Starts und die entsprechenden Ergebnisse mit der Homotopiemethode, wenn man die gleichen Startstrukturen wählt.

Die Spalte "VdW" zeigt die Anzahl der Van-der-Waals Kontakte an: Ein solcher Kontakt liegt vor, wenn die Distanz der beiden entsprechenden Atome kleiner als das 1.01-fache ihres Gleichgewichtsabstandes beträgt.

Die Spalte "davon +" gibt an, wieviele der Kontakte abstoßende Wechselwirkungen darstellen.

In diesem Fall liefert die Homotopiemethode in den meisten Fällen schlechtere Ergebnisse als der reine Multi-Start (Ausnahmen werden jeweils mit * markiert).

Der Grund dafür zeigt sich in der Auswertung der Van-der-Waals Kontakte: In den Ergebnisstrukturen der Homotopiemethode liegen stets mehr Kontakte vor, als bei den Multi-Start Ergebnissen. Damit steigt auch die Zahl der

abstoßenden Wechselwirkungen und der Zielfunktionswert wird verschlechtert.

Ergebnisse für den Multi-Start:

Start	Gradient [$< 10^{-4}$]	Zielfunktion	VdW	davon +
1	0.0005	133.879	7	2
2		320.693	11	6
3		41.536	11	6
4		319.041	8	4
5		90.041	8	4
6		311.725	11	6
7		21.943	10	7
8		198.160	9	7
9		114.228	10	7
10		0.439	9	5

Und für die Homotopiemethode:

Start	Gradient [$< 10^{-3}$]	Zielfunktion	VdW	davon +	
1	0.001	484.730	12	7	
2		577.974	14	9	
3		435.243	11	7	
4		0.00009	5.595*	11	6
5		321.552	12	5	
6		484.730	12	7	
7		51.459	10	8	
8		7.677*	12	6*	
9		0.001	307.950	11	8
10		265.034	7*	3	

Ausnahme: In "Start 8" wird die Zahl der Van-der-Waals Kontakte zwar größer, jedoch sinkt die Zahl der abstoßenden Wechselwirkungen, dadurch wird der Funktionswert besser.

Daß die Homotopiemethode zu einer größeren Anzahl von Van-der-Waals Kontakten tendiert wird deutlich, wenn man sich vor Augen hält, wie die Zielfunktion geglättet wird. Das Bindungsgerüst wird einerseits gestaucht, andererseits jedoch werden die Van-der-Waals Radien gedehnt. Dadurch kommt es

in der Diffusionsphase zu einer größeren Zahl von Van-der-Waals Wechselwirkungen, die ihm Laufe der Minimierung durch "Potentialberge" beibehalten werden.

Dieses Verhalten der Homotopiemethode wirkt sich besonders positiv auf die Minimierung von Aggregaten aus und erklärt somit die guten Ergebnisse: In diesem Fall ist eine Maximierung der Zahl von Van-der-Waals Kontakten stets wünschenswert.

Anders ist der Fall bei Trichlordifluorbutan, denn hier führen die Van-der-Waals Kontakte häufig zu Abstoßungen, die eigentlich zu vermeiden sind.

Klares Fazit: Die Homotopiemethode eignet sich nicht für jedes Molekül. Man kann sogar genauer formulieren: Es sind besonders diejenigen Moleküle für die Minimierung schlecht geeignet, in denen Van-der-Waals Abstoßungen kritisch sind. Sind Derivate der Alkane mit wenigen Substituenten also besser geeignet?

7.3.3 Dichlordifluorpentan

Ein Beispiel, in dem die Abstoßungsterme weniger ins Gewicht fallen, liefert 1,4-Dichlor-1,5-difluorpentan ($C_5H_8Cl_2F_2$).

Diese Struktur soll nun wieder nach dem United-Atom Modell minimiert werden. Die Atome mit dem kleinsten Van-der-Waals Radius (Wasserstoff) werden also wieder eliminiert und der Van-der-Waals Radius der Kohlenstoffatome entsprechend angepaßt.

Damit in diesem Fall auch die 1,4-Abstoßung der Kohlenstoffatome nicht stört, setzt man für die 1,4-Terme (Lennard-Jones und Torsionspotential) im Kraftfeld eine Distanzgeometriefunktion an, so daß der Gleichgewichtsabstand einem Torsionswinkel von 180° entspricht, in diesem Fall 3.93\AA .

Zur Struktur einer Eingabedatei siehe Seite 98.

Vergleicht man hier wieder Multi-Start mit der Homotopiemethode ($L = 0.61$) und die Zahl der Van-der-Waals Kontakte bei beiden Verfahren, so ergibt sich diesmal folgendes Bild, das man durch Anwenden der beiden Methoden auf zehn vorgegebene Startstrukturen erhält:

Wieder strebt die Homotopiemethode eine möglichst große Anzahl von Van-der-Waals Kontakte an. Der Unterschied zu Trichlordifluorbutan liegt jedoch nun darin, daß dieses Vorgehen bei Dichlordifluorpentan mit besseren Funktionswerten einhergeht.

Nr.	Wert (Multi-Start)	VdW	Wert (Homotopie)	VdW
1	-0.528*	5	-0.482	5
2	51.498	3	-0.450	4
3	51.772	2	-0.489	5
4	-0.450	4	-0.489	5
5	54.309	3	-0.485	5
6	-0.450	4	-0.489	5
7	-0.515	4	-0.528	5
8	-0.450	4	-0.489	5
9	41.704	2	-0.528	5
10	-0.490	3	-0.528	5

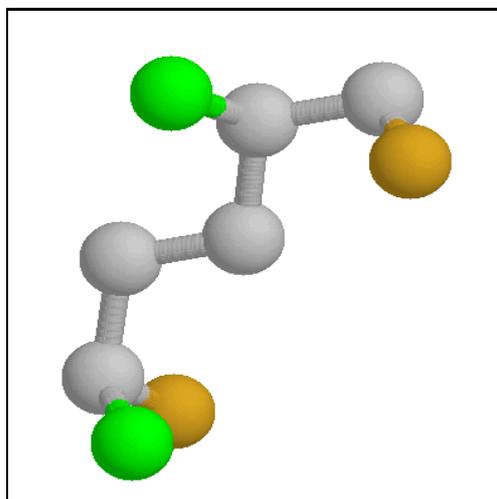


Abbildung 7.5: Dichlordifluorpentan: Optimalstruktur

Verallgemeinert man das Vorgehen dieses Beispiels einer globalen Optimierung kurzkettiger Alkanderivate, so erhält man folgende Punkte:

- Das Molekül darf nur an wenigen Stellen substituiert sein.
- Die Form der Alkankette wird durch Festlegen der Torsionswinkel vorgegeben. Entsprechende 1,4-Terme werden im Kraftfeld durch Distanzgeometriefunktionen ersetzt. Nach Goodman [8] ist die Annahme einer linearen Kohlenstoffkette bis C_{18} eine gute Näherung.

- Die Wasserstoffatome werden aus der Verbindung eliminiert und die Van-der-Waals Radien der Kohlenstoffatome durch die United-Atom Modellierung angepaßt, wobei lediglich die 1,5-Wechselwirkungen der substituieren Atome mit den vernachlässigten Wasserstoffatomen zu schwach wiedergegeben wird.

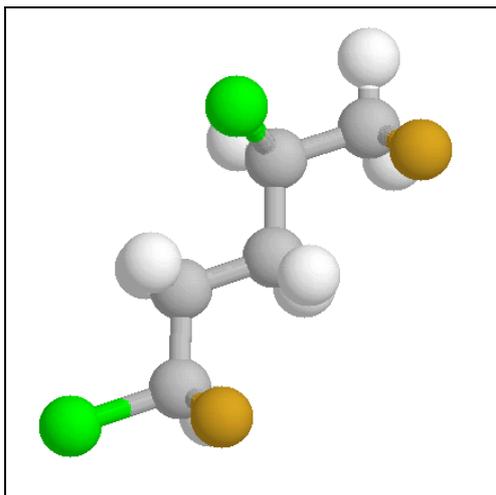


Abbildung 7.6: Dichlordifluorpentan: AMBER-Struktur

Bemerkung: Die in Abbildung 7.5 gezeigte Struktur ist optimal für das stark vereinfachte Kraftfeld (7.2). Der Optimierung muß also wieder eine kritische Betrachtung des Ergebnisses folgen.

Es ist zu erwarten, daß durch die United-Atom Modellierung die 1,5-Wechselwirkungen zwischen den Substituenten und den vernachlässigten Wasserstoffatomen zu schwach wiedergegeben wird.

Tatsächlich liefert (zeitaufwendiges) systematisches Probieren durch Variation von Stereoisomerie und Torsionswinkeln für Dichlordifluorpentan im AMBER-Kraftfeld in Abbildung 7.6 folgendes Ergebnis:

Die asymmetrischen Kohlenstoffatome (C1 und C4) weisen im Optimum des vollständigen Kraftfeldes diesselbe Stereoisomerie wie in Abbildung 7.5 auf. Einziger Unterschied besteht in dem Bindungspolyeder des ersten Kohlenstoffatoms, der im AMBER-Kraftfeld um 120° gegenüber Abbildung 7.5 verdreht ist.

Wie zu erwarten, liegt der Grund in der Modellierung der 1,5-Wechselwirkung des Chloratoms mit dem Wasserstoffatom am C3.

Dennoch ist es bemerkenswert, daß ein vereinfachtes Kraftfeld mit einer ad hoc Parametrisierung so gute Voraussagen ermöglicht.

7.4 Interpretationen der Optimierungsmethode

Das Verhalten der Optimierungsroutine wurde an Beispielen demonstriert. Hervorstechende Eigenschaft war die Maximierung der Zahl von Wechselwirkungen ungebundener Zentren durch die Homotopiefunktion. Daß dieses Verhalten durch die Theorie gestützt wird, zeigen folgende zwei Sichtweisen.

7.4.1 Die Gaußtransformation als Erwartungswert

Die Gaußtransformation (3.7) stellt eine bestimmte Erwartungswertbildung dar.

Bei dieser Erwartungswertbildung werden abhängig von dem Parameter λ die Funktionswerte der Zielfunktion in einer Umgebung um den betrachteten Punkt gemittelt.

Die Gaußtransformation als Homotopiemethode eignet sich also für Funktionen, deren globales Minimum in einem Bereich tiefer Energie liegt.

Bei Molekülen zeichnen sich solche Bereiche dadurch aus, daß eine entspannte Geometrie (also mit den vorgegebenen Sollabständen) in einem Bereich anziehender Van-der-Waals Wechselwirkungen liegt.

Dieses erklärt das Verhalten der Optimierungsroutine bei den behandelten Beispielen, in denen die Geometrien mit möglichst vielen Van-der-Waals Kontakten bevorzugt wurden.

7.4.2 Krümmung und Diffusion

Eine weitere Sichtweise der Homotopiemethode ergibt sich aus der Diffusionsgleichung:

$$\frac{\partial F}{\partial t} = \Delta F. \quad (7.3)$$

Liegt in einem Punkt $x^0 \in \mathbb{R}^n$ ein Minimum der Funktion $F(\cdot, t)$ vor, so gilt $\Delta F(x^0, t) \geq 0$, der Wert der Homotopiefunktion F wird wegen der Diffusionsgleichung für wachsendes t ansteigen. Analog sinkt der Funktionswert für relative Maxima.

Durch diesen Effekt wird die globale Struktur der Zielfunktion durch Glättung erreicht.

Quantitativ wird das Wachstumsverhalten der Minima jedoch nicht durch ihren Funktionswert, sondern durch die Krümmung im Minimum beeinflusst. Minima mit starker Krümmung werden schneller geglättet als andere.

In einem isolierten Molekül läßt sich die Krümmung der Potentialfunktion in einem Minimum durch das Schwingungsverhalten der Molekel interpretieren. Eine hohe Schwingungsfrequenz ist gleichbedeutend mit einer starken Krümmung, da die Frequenzen der Eigenschwingungen der Molekel dem Wert der Wurzel der Eigenwerte der Hessematrix im lokalen Minimum proportional sind. Den entsprechenden quantenchemischen Ansatz findet man bei Primas und Müller-Herold [29], Seiten 191 bis 195.

Hohe Schwingungsfrequenzen besitzen Minima, in denen besonders viele abstoßende Van-der-Waals Kontakte vorliegen, da hier der steile Ast des Lennard-Jones Potentials und der steile Ast der Strafterme für die Distanzbedingungen gegenüberstehen.

Dieser Effekt ist zwar bei dem Beispiel Trichlordifluorbutan nicht offensichtlich, aber es macht den besonderen Charme der Diffusionsmethode deutlich: Einerseits werden, aufgrund der Sichtweise im vorangehenden Abschnitt, viele Van-der-Waals Kontakte bevorzugt, andererseits ist zu erwarten, daß Minima mit starker Van-der-Waals Abstoßung stärker geglättet werden und eventuell elimiert worden sind.

Diese Argumente untermauern den Einsatz der Homotopiemethode besonders im Bereich der Optimierung von Molekülstrukturen und von Aggregaten.

Kapitel 8

Rückschau auf die Diplomarbeit

Die große Schwierigkeit bei meiner Diplomarbeit bestand darin, einen funktionierenden Algorithmus zur Lösung von globalen Optimierungsproblemen mit Molekülpotentialfunktionen zu *finden*.

Es war nicht von vorneherein klar, daß ein nichtdeterministischer Weg eingeschlagen werden mußte, um die Probleme in brauchbarer Zeit zu lösen.

Jedoch nach einer Betrachtung der Komplexität von Optimierungsproblemen und nach vielen Computereperimenten mit zuverlässigen Algorithmen, bei denen man im schlimmsten Fall (Aufsuchen aller Minima der Funktion) eine Laufzeit von mehreren Jahren erwarten konnte, war ich positiv überrascht, daß die Homotopiemethode so viel schneller war und mit nur wenigen Starts bereits das globale Minimum findet.

Noch viel überraschender war es dann, daß die Methode mit kleinen Werten für λ_{max} bereits brauchbare Ergebnisse lieferte. In den Ausführungen von Scheraga [20] werden Anfangsparameter gewählt, die zwei bis drei Größenordnungen über meinen liegen. Versuche mit der gewählten Zielfunktion meiner Arbeit haben sogar so etwas wie eine *magische, obere Grenze* für die Wahl der Anfangsglättung ergeben, unterhalb der die Zahl der erreichbaren Minima mit kleiner werdendem λ_{max} monoton steigt. Eine genaue Analyse der geglätteten Funktionen unter dem Aspekt der Verzweigungstheorie lieferte eine mögliche Erklärung für dieses Phänomen.

Weitere Schwierigkeiten (Konvergenz des BFGS-Verfahrens, rotatorische Freiheitsgrade der Moleküle,...) und entsprechende Lösungsansätze habe ich be-

reits besprochen. Deshalb möchte ich hier nocheinmal kurz skizzieren, welchen Reiz die zuverlässigen, deterministischen Algorithmen auf mich ausgeübt haben und daß es nicht leicht war, sich von ihnen zu trennen.

Zunächst einmal stehen einem Mathematiker zur Lösung von globalen Optimierungsproblemen die sogenannten Branch-and-Bound Techniken zur Verfügung. Um diese nutzen zu können, braucht man lediglich eine Möglichkeit, eine untere Schranke der Zielfunktion auf einer gegebenen Teilmenge des zulässigen Bereichs zu bestimmen. Diese Schranke muß für kleiner werdende Partitionierung des zulässigen Bereichs immer besser werden.

Auf diesem Weg nähert man sich durch sukzessives Partitionieren und Verwerfen unbrauchbarer Teilmengen, der Menge der globalen Minima einer Zielfunktion.

Am Anfang der Überlegungen standen die Distanzgeometrie probleme, die die Grundlage für die globale Optimierung von Molekülen bilden. Die Distanzgeometriefunktionen sind sogenannte *d.c.*-Funktionen, eine Differenz zweier konvexer Funktionen p, q . Im Falle der Distanzgeometriefunktionen

$$\begin{aligned} p(x) &= \sum_{(i,j) \in \mathfrak{R}} r_{ij}^2(x)^2 + a_{ij}^4, \\ q(x) &= \sum_{(i,j) \in \mathfrak{R}} 2r_{ij}^2(x)a_{ij}^2. \end{aligned} \quad (8.1)$$

Man gelangt zu einer unteren Schranke von *d.c.*-Funktionen $p(x) - q(x)$, auf Simplexes S , indem man jeweils von p und von $-q$ die beste Minorante auf S berechnet.

Die beste Minorante von p ist p selber aufgrund der Konvexität. Die beste Minorante der konkaven Funktion $-q$ ist eine lineare Funktion, und zwar genau die lineare Funktion $l^T x + b$, für die $l^T v_i + b = -q(v_i)$ gilt, wobei v_1, \dots, v_{n+1} die Ecken von S sind [19].

Man gelangt so zu folgendem, konvexen Optimierungsproblem auf einem konvexen Simplex in den $n+1$ Variablen α_i , indem man x als Konvexkombination der Ecken von S darstellt:

$$\begin{aligned} \min \quad & p\left(\sum_{i=1}^{n+1} \alpha_i v_i\right) - \sum_{i=1}^{n+1} \alpha_i q(v_i) \\ \text{unter} \quad & \sum_{i=1}^{n+1} \alpha_i = 1, \quad \alpha_i \geq 0, i \in \{1, \dots, n\}. \end{aligned} \quad (8.2)$$

Die Lösung dieses Optimierungsproblems gibt eine untere Schranke $u(S)$ von $f = p - q$ auf dem Simplex S an. Die Differenz $f(x^*) - u(S)$ mit dem optimalen Punkt x^* gibt die maximale Abweichung zum tatsächlichen Minimum von f auf S an. Wird dieser Wert kleiner als ein vorgegebenes ϵ , so ist x^* eine Lösung des ϵ -globalen Optimierungsproblems. Der entsprechende Algorithmus ist also zuverlässig und deterministisch.

Die Sache wird komplizierter, wenn man Lennard-Jones Wechselwirkungen hinzunimmt. Hier kann man eine *d.c.*-Zerlegung nur dann angeben, wenn man den Pol der Funktion abschneidet, so daß die Hessematrix beschränkt bleibt. In diesem Fall kann man, wie beim α -Branch-and-Bound durch einen Shift der Eigenwerte erreichen, daß sich die Funktion durch eine *d.c.*-Dekomposition darstellen läßt.

Diese Möglichkeiten, einen deterministischen Algorithmus anzugeben, der immer zuverlässige Ergebnisse liefert und dabei die gesamte Menge der ϵ -globalen Minima approximiert, sind sehr verlockend, zumal sie mathematisch exakt durchführbar sind. Das einzige Problem, die Darstellung in eindeutigen, kartesischen Koordinaten, wurde bereits in dieser Arbeit gelöst. Dieses Problem stellt sich, da die approximierten Mengen beschränkt sein sollten.

Führt man sich jetzt vor Augen, was es bedeutet, z.B. eine 20-dimensionale Menge sukzessiv immer feiner zu partitionieren, dann wird einem klar, mit welchem (exponentiell steigenden) Aufwand dieses geschieht. Und weiter: Oft werden sämtliche, lokalen Minima der zulässigen Menge aufgesucht, da diese bei Molekülen sehr dicht beieinander liegen können.

Also, habe ich die Theorie der deterministischen Algorithmen verlassen und versucht einen nichtdeterministischen Ansatz zu finden, bei dem der Chemiker bei Aufrufen des Programmes nicht Tage lang auf das Ergebnis warten muß.

Es fiel schnell auf, daß nichtdeterministische Routinen Gebrauch von globalen Strukturen der Funktion machen. Für den Homotopiealgorithmus habe ich mich letztendlich deshalb entschieden, da er nicht eine *bestimmte* globale Struktur voraussetzt.

Es ist klar, daß diese Art von Optimierungsroutinen immer angreifbare Schwachpunkte besitzen, sie sind ganz einfach unzuverlässig. Auch kann man bei der Homotopiemethode nicht sicher sein, daß das globale Minimum in der Menge der erreichbaren Minima vorhanden ist.

Aber andererseits überwiegen meiner Meinung nach die Stärken dieser Methode (eine große Anzahl behandelbarer Freiheitsgrade und damit eine Mög-

lichkeit, Aggregate zu optimieren) deren Nachteile, was ich hoffentlich dem Leser meiner Diplomarbeit zeigen konnte.

Die Besonderheit bei der Themenwahl dieser Diplomarbeit war es, daß ein chemisches Problem zu lösen war, für das eine mathematische Methode, d.h. eine Optimierungsroutine und ein abgewandeltes Kraftfeld entwickelt werden mußten.

Der Ausgangspunkt war also nicht die Vorgabe einer Methode und die anschließende Suche nach funktionsfähigen Beispielen, sondern eine gegebene Funktionenklasse, die viele Schwierigkeiten (z.B. Polstellen und das daraus resultierende Nichtvorhandensein einer *d.c.*-Zerlegung) für die globale Optimierung beinhaltete.

Anhang

Die Eingabedateien

Die globale Optimierungsroutine *minim* liest bei dem Aufruf zunächst eine Datei *param.lst* auf, in der die notwendigen Parameter zur Steuerung der Optimierung zu finden sind.

- Der erste Parameter gibt die Anzahl der Gewichte an, die bei einer numerischen Integration bei der Glättung der Torsionspotentiale verwendet werden. In jedem Beispiel wurde diese Zahl auf 5 gesetzt.
- Der zweite Parameter gibt die Anzahl an Diskretisierungspunkten für die Homotopiemethode an, bei Acrolein und Wasser wurde 100, sonst 200 gewählt.
- Der dritte Parameter gibt den negativen dekadischen Logarithmus für *tolX* (BFGS-Algorithmus) an: Für Acrolein und Wasser war er 10, sonst 14.
- Der vierte Parameter gibt den negativen dekadischen Logarithmus für die relative Gradientennorm (BFGS-Algorithmus) an: Für Acrolein und Wasser und Formaldehyd war er 3, sonst 4.
- Der fünfte Parameter gibt die Anzahl an Optimierungszyklen an.
- Der sechste Parameter gibt den Cut-off Wert für den minimalen Abstand zweier Atome an. Er betrug jeweils 1.52, im Falle der Alkane konnte 0.5 gewählt werden.
- Der letzte Parameter gibt das hundertfache des Wert L_{max} an, eine zweistellige Größe.

Nachdem diese Datei gelesen wurde, wird eine Datei *Input.dat* aufgerufen, die alle Informationen über das zu minimierende Molekül enthält.

- Zunächst deutet die Abkürzung *GO* in der ersten Zeile auf die globale Optimierung hin, die man mit dem Molekül durchführt. Dann folgt die Anzahl der Zentren und die Zentrentypen im RasMol-Format.
- Nach den Zentrentypen werden die Bindungen initialisiert. Dabei werden jeweils die beiden Zentren, der Verstärkungsfaktor für die Gewichtung der Bindung gegenüber den Lennard-Jones Wechselwirkungen und die Bindungslänge aufgeführt.
- Nach den Bindungen werden die Bindungswinkel beschrieben, zunächst wieder die drei beteiligten Zentren, der Verstärkungsfaktor und dann der Winkel in Grad.
- Nach den Bindungswinkeln werden die Torsionswinkel aufgeführt. Wieder zunächst die beteiligten Zentren, der Verstärkungsfaktor und die Zähligkeit der Drehung um die zentrale Bindung.
- Dann folgen noch die Wechselwirkungen ungebundener Zentren. Hier werden die beiden Atomnummern, die beiden Lennard-Jones Parameter und der Coulomb Parameter aufgeführt.
- Ganz am Ende kann man noch Distanznebenbedingungen setzen, was in dieser Diplomarbeit nicht beschrieben wurde, aber in [23] bereits ausführlich behandelt worden ist.

Hier nun eine Beispieldatei für Hexan mit dem Torsionspotential:

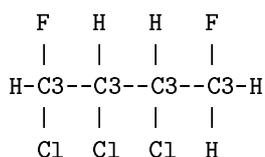
```
GO
6
C
C
C
C
C
C
C
5
1 2 50.000000 1.562000
2 3 50.000000 1.562000
```

```

3  4  50.000000  1.562000
4  5  50.000000  1.562000
5  6  50.000000  1.562000
4
1  2  3  50.000000  109.470000
2  3  4  50.000000  109.470000
3  4  5  50.000000  109.470000
4  5  6  50.000000  109.470000
3
1  2  3  4  1.000000  3
2  3  4  5  1.000000  3
3  4  5  6  1.000000  3
6
1  4  151833.805000  74.137600  0.000000
1  5  3036676.100000  1482.752000  0.000000
1  6  3036676.100000  1482.752000  0.000000
2  5  151833.805000  74.137600  0.000000
2  6  3036676.100000  1482.752000  0.000000
3  6  151833.805000  74.137600  0.000000
0

```

Die Eingabedateien wurden mit von einem TURBO-Vision C++ Programm erzeugt, in dem der Benutzer ein Molekül auf einer menügesteuerten Oberfläche in Form eines Bildes eingeben und in der Datei *molec.pic* ablegen kann. Aus den vorher bestimmten Atomdaten und dem Molekülbild erzeugt das Programm eine Eingabedatei im obigen Format, wobei die Atome des Molekülbildes bei der Nummerierung zeilenweise übernommen werden. Ein Beispiel für ein Molekülbild, in dem die Atomtypen nicht im RasMol-Format stehen müssen, liefert Trichlordifluorbutan:



Im Beispiel Acrolein und Wasser gibt es auch noch die Möglichkeit, eine Datei im MOBY-Format einzulesen. Aus dieser Datei werden die Distanzgeometriefunktionen, die Startkoordinaten und die Partiaalladungen der Atome extrahiert. Der Benutzer muß nur noch die Lennard-Jones Parameter vordefinieren, damit das Programm eine Datei *Input.dat* erzeugen kann.

Die globale Optimierungsroutine ruft also zunächst *param.lst* auf, danach *Input.dat* zur Festlegung der Zielfunktion. Das Bild des Moleküls wird, falls vorhanden, ausgegeben und die Startkoordinaten aus der Datei *start.car* entnommen oder zufällig gewählt, falls eine solche Datei nicht gefunden wurde. Danach beginnt die eigentliche Optimierung. Die kartesischen Koordinaten des Ergebnisses werden in der Datei *optimal.xyz* ausgegeben. Diese kann mit dem Aufruf *rasmol -xyz optimal.xyz* in ein Molekülbild verwandelt werden. Während das Programm zur Erzeugung der Eingabedatei nur unter DOS zur Verfügung steht. Kann die Minimierung unter DOS oder LINUX durchgeführt werden.

Die Minimierungsroutine kann auch aus dem TURBO-Vision Programm heraus gestartet werden, es wird dann ein Fenster geöffnet, in dem die Optimierungsparameter für die Datei *param.lst* eingestellt werden.

Die maximale Anzahl an Atomen wurde zunächst auf 30 vorgegeben, das entspricht 90 Optimierungsvariablen.

Berechnung der Gleichgewichtsabstände

Zur Definition des Distanzgeometrieproblems aus vorgegebenen Bindungslängen, Bindungswinkeln und Torsionswinkeln müssen die erwarteten Winkelgrößen in Gleichgewichtsabstände umgerechnet werden.

Liegt ein dreiatomiges Molekül mit den Bindungslängen a und b vor. Schließen weiter diese Bindungen einen Winkel ϕ ein, so errechnet sich der sogenannte *1,3-Abstand* aus dem Kosinussatz:

$$d = \sqrt{a^2 + b^2 - 2ab \cos(\phi)}.$$

Mit ähnlichen geometrischen Überlegungen errechnet man für eine Kette von vier gebundenen Atomen mit den Bindungslängen b_{12}, b_{23}, b_{34} und den Bindungswinkeln $\phi_{123}, \phi_{234} > 90^\circ$ für einen Torsionswinkel ω einen *1,4-Abstand* d_ω .

$$d_{180} = \sqrt{b_{12}^2 + 2b_{12}b_{34} \cos(\phi_{123} - \phi_{234}) + b_{24}^2 - 2b_{12}b_{23} \cos(\phi_{123})},$$

$$d_{60} = (b_{12}^2 + 2b_{12}b_{34}(\cos(\phi_{123}) \cos(\phi_{234}) - 1/2 \sin(\phi_{123}) \sin(\phi_{234})) + b_{24}^2 - 2b_{12}b_{23} \cos(\phi_{123}))^{1/2},$$

$$d_0 = \sqrt{b_{12}^2 + 2b_{12}b_{34} \cos(\phi_{123} + \phi_{234}) + b_{24}^2 - 2b_{12}b_{23} \cos(\phi_{123})}. \quad (8.3)$$

Bei einer dreizähligen Torsionsachse nehmen die endständigen Atome den Abstand d_{180} bzw. d_{60} ein. Bei einer zweizähligen Achse (z.B. Alkene) berechnet man mit d_{180} und d_0 die zugehörigen Gleichgewichtsabstände.

Praktische Durchführung

Aggregat

Möchte man ein Aggregat aus vorgegebenen Molekülgeometrien im Raum optimal ausrichten, so errechnet man zunächst mit quantenchemischen Methoden die Partialladungen der beteiligten Atome, z.B. mit dem Programm *gaussian94*.

Mit Hilfe des Programms *xmoby* erzeugt man nach Eingabe der Partialladungen eine lokale Minimumsgeometrie des Molekülverbundes und gibt diese in einer Datei *mobyout.opt* aus.

In dieser Datei befinden sich Angaben zu den beteiligten Atomtypen. In meinem Programm *smooth* gibt man nun unter dem Menüpunkt *Add new atoms* die erforderlichen Parameter der Atomtypen ein. Diese können dem AMBER-Kraftfeld oder anderen Literaturangaben entnommen werden.

Danach wird unter dem Menüpunkt *mobyout.opt* die MOBY-Datei eingelesen und eine Datei *Input.dat* erzeugt. Das Programm zeigt die Zahl der fixierten Distanzen an.

Unter dem Menüpunkt *Start minimization* ruft man die Parameterdatei auf, die man nun beliebig ändern kann, und durch Anklicken der *OK*-Option wird schließlich die Minimierung gestartet.

Isolierte Moleküle

Möchte man ein isoliertes Molekül optimieren (Stereoisomerie, Torsionswinkel), so zeichnet man mit meinem Programm *smooth* unter dem Menüpunkt *Draw molecule* zunächst eine Molekülgeometrie wie auf Seite 98, dabei dürfen die Atomtypen durch nicht mehr als zwei Zeichen definiert werden.

Unter dem Menüpunkt *Add new atoms* und *Add new angles* werden die benötigten Parameter gesetzt, die mit Hilfe des Menüpunktes *Create molecule datas* in eine Datei *Input.dat* verwandelt werden.

Unter dem Menüpunkt *Input.dat* kann nun noch die Zähligkeit der Torsionsbindungen von 3 auf 2 herabgesetzt werden oder bestimmte Eingabedaten anders manipuliert werden.

Distanzgeometrie probleme

Möchte man reine Distanzgeometrie probleme lösen, so kann man die bekannten Distanzen in der Datei *Input.dat* als Bindungslängen definieren.

Wie bereits erwähnt kann die Datei *Input.dat* auch noch Distanznebenbedingungen erhalten, die einen unteren und einen oberen Wert akzeptieren.

Das zugehörige Format ist:

Zentrum1	Zentrum2	Gewicht	untere Schranke	obere Schranke
----------	----------	---------	-----------------	----------------

Literaturverzeichnis

- [1] Peter W. Atkins. *Physikalische Chemie*, VCH Publishers, Vol. III, 1992.
- [2] Copyright (c) 1992, 1993. The Regents of the University of California. Version 8.1 (Berkeley) 6/4/93
- [3] Jeffrey M. Blaney und J.Scott Dixon. *Distance Geometry in Molecular Modeling*, S. 299-335, Rewies in Computational Chemistry, Vol. V, VCH Publishers, Inc. New York, 1994.
- [4] W.Bollweg. *Globale Optimierungsmethoden zur Bestimmung von kristallinen und molekularen Strukturen*, Inaugural-Dissertation, Westfälische Wilhelms-Universität Münster, 1998.
- [5] W.Braun. *Distance Geometry and Related Methods for Protein Structure Determination from NMR Data*, S. 115, Vol. 19, Q. Rev. Biophys., 1987.
- [6] G.M. Crippen und T.F. Havel. *Distance Geometry and Molecular Conformation*, Jhon Wiley and Sons, 1988.
- [7] J.E. Dennis Jr. und R.B. Schnabel. *A New Derivation of Symmetric Positive Definite Secant Updates*, S. 167 - 199, Nonlinear Programming 4, O.L. Mangasarian, R.R. Meyer und S.M. Robinson, (eds). Academic Press, London and New York, 1981.
- [8] J.M. Goodman. *J. Chem. Inf. Comput. Sci.*, S. 876 - 878, Vol. 37, 1997.
- [9] G.Goos. *Vorlesungen über Informatik*, S. 30 ff., Bd. 1, 2. Aufl., Springer, 1997.
- [10] A. Griewank. *Generalised descend for global optimization*, S. 11 - 39, J. Optim. Theory Appl., Vol. 34, 1981.

- [11] Ch.Großmann, J. Terno. *Numerik der Optimierung*, Teubner Studienbücher Mathematik, 1997.
- [12] G. Hämmerlin und K.H. Hoffmann. *Numerische Mathematik*, Springer Verlag, 1994.
- [13] H.Hart. *Organische Chemie*, S.46, VCH, 1989.
- [14] U. Höweler. *Einführung ins Molecular Modelling*, Vorlesung WS 94/95.
- [15] U. Höweler. *Fortgeschrittene Methoden des Molecular Modelling*, Vorlesung SS 97.
- [16] Udo Höweler. *xmoby*, Version 2.0M 8.96, 1996.
Spezielle Datei: "mobyambr.par"
- [17] A. Hoffjan, A. Reuter und M. Schilke. *Seminar: Stabilität, Verzweigung, Chaos*, WS 1990/91.
- [18] A.J. Hopfinger. *Conformational Properties in Macromolecules*, Academic Press, New York and London, 1995.
- [19] Reiner Horst, Panos M. Pardalos und NguyenV. Thoai. *Introduction to Global Optimization*, Kluwer Academic Publishers, 1995.
- [20] J. Kostrowicki, L.Piela, B.J. Cherayil und H.A. Scheraga. *Performance of the Diffusion Equation Method in Searches for Optimum Structures of Cluster of Lennard-Jones Atoms*, S. 4113 - 4119, J. Phys. Chem, American Chemical Society, Vol 95, 1991.
- [21] R.W. Kunz. *Molekular Modelling für Anwender*, Teubner Studienbücher Chemie, 1991.
- [22] H. Maurer. *Vorlesungsskript über nichtlineare Optimierungsprobleme*, Satz (3.5), interne Quellen.
- [23] Jorge J. Moré und Zhijun Wu. *ϵ -Optimal Solutions to Distance Geometry Problems via Global Continuation*, Preprint MCS-P520-0595, Mai 1995. — *Issues in Large-Scale Global Molecular Optimization*, Preprint MCS-P539-1095, März 1996. — *Distance Geometry Optimization for Protein Structures*, Preprint MCS-P628-1296, Dezember 1996.

- [24] Jorge J. Moré und Zhijun Wu. *Global Continuation for Distance Geometry Problems*, S. 814 - 836, SIAM Journal of Optimization, Vol. 7, No. 3, 1997.
- [25] Soichiro Nakamura. *Applied numerical methods in C*, PTR Prentice Hall, 1993.
- [26] F. Natterer. *Vorlesung in partiellen Differentialgleichungen*, S. 114, Vorlesungsskript WS 95/96.
- [27] J.A.Nelder und R.Mead. *Computer Journal*, S. 308-313, vol. 7, 1965.
- [28] William H. Press. *Numerical recipes in C: the art of scientific computing*, Press Syndicate of the University of Camebridge, 2nd ed., 1992.
- [29] H. Primas und U. Müller-Herold. *Elementare Quantenchemie*, Teubener Studienbücher, 2. Aufl., 1990.
- [30] A.H.G. RinnooyKan und G.T.Timmer. *New Methods in Optimization and their Industrial Uses*, 133-155. Birkhäuser Verlag, Basel, 1989.
- [31] P.Rademacher. *Strukturen organischer Moleküle*, VHC, 1987.
- [32] James B. Saxe. *Embeddability of weighted graphes in k -space is strongly \mathcal{NP} -hard*, S. 480 - 489, Proc. 17th Allerton Conference in Communications, Control and Computing, Monticello, Il, 1979
- [33] Roger Sayle. *RasMol Molecular Renderer*, Version 2.6, [8bit version], August 1995.
- [34] Walter A. Strauss. *Partial Differential Equations: An Introduction*, John Wiley and Sons, 1992.
- [35] S.J. Weiner. *A New Force Field for Molecular Mechanical Simulation of Nucleic Acids and Proteins*, S. 765 - 784, J. Am. Soc., Vol. 106, 1982.
- [36] K.Wuethrich *Protein Structure Determination in Solution by Nuclear Magnetic Resonance Spectroscopy*, S. 243, Siecne, Vol 45, 1989.

Die vorliegende Arbeit wurde im Institut für Numerische Mathematik der Westfälischen Wilhelms-Universität Münster unter Anleitung von Herrn Prof. Dr. H. Maurer ausgeführt. Ich versichere, daß ich die hier vorliegende Arbeit selbständig angefertigt und keine weiteren Hilfsmittel als die im Literaturverzeichnis aufgeführten Quellen verwendet habe.

Münster, 20. November 1998