Diplomarbeit

Optimierungsmethoden zur multiplen flexiblen Ligandüberlagerung

angefertigt am Institut für Angewandte Mathematik
vorgelegt der
Mathematisch-Naturwissenschaftlichen Fakultät der
Universität Bonn

Februar 2001

von
Markus Lilienthal aus Gelsenkirchen

Inhaltsverzeichnis

Kapitel 1

Einleitung

Pharmazeutische Wirkstoffe erzielen ihre spezifische Wirkung in der Regel durch die Bindung an einen *Rezeptor*. Rezeptoren sind makromolekulare Substanzen, häufig Proteine, während die Wirkstoffe, die allgemeiner als *Liganden* bezeichnet werden, in der Regel niedermolekular und hochflexibel sind.

Die Bindung eines Liganden an einen Rezeptor beruht in erster Näherung auf dem „Schlüssel-Schloß-Prinzip“[1]. Die Bindungsstelle des Rezeptors (*Bindungstasche*) stellt das Schloß dar, in das der Ligand wie ein Schlüssel exakt hineinpassen muss. Im Gegensatz zu einem herkömmlichen Schlüssel spielt bei Liganden neben der geometrischen Struktur auch die physikochemischen Eigenschaften[2] eine entscheidende Rolle. Die Menge der Schlüssel, die in ein und dasselbe Schloß passen, kann dem Augenschein nach recht heterogen sein. Die Liganden können sich durchaus strukturell unterscheiden und dennoch aufgrund ihrer geometrischen und physikochemischen Gemeinsamkeiten in der Lage sein, an denselben Rezeptor zu binden.

Die Suche nach neuen pharmazeutischen Wirkstoffen (*Leitstrukturen*) erfordert einen großen zeitlichen und finanziellen Aufwand, da im Experiment oft eine sehr große Zahl von Molekülen auf ihre Bindungsstärke (*Affinität*) zu einem gegebenen Rezeptor untersucht werden müssen. Ein Forschungsgegenstand ist die Entwicklung rechnergestützter Methoden, die es erlauben, aus einer großen Menge von Molekülen einige wenige Kandidaten herauszufiltern, die mit großer Wahrscheinlichkeit im Experiment eine hohe Affinität zum Rezeptor aufweisen. Damit soll der experimentelle Aufwand bei der Leitstruktursuche drastisch verringert werden.

Häufig sind zu einem Rezeptor eine ganze Reihe von Liganden bekannt, während die genaue Struktur der Bindungstasche unbekannt ist. In dieser Situation können die physikochemischen und geometrischen Gemeinsamkeiten der bekannten Liganden wertvolle Hinweise für die Suche nach neuen Wirkstoffen liefern. Insbesondere kann häufig auf dieser Basis ein Modell des Bindungsmechanismus entwickelt werden, welches es erlaubt, Vorhersagen über die Bindungseigenschaften noch nicht experimentell untersuchter Moleküle zu treffen. Eine wichtige Methode

[1] Das Schlüssel-Schloß-Prinzip wurde bereits 1894 von Emil Fischer formuliert [15]

[2] Der Begriff physikochemische Eigenschaft bedeutet in diesem Zusammenhang eine lokale chemische Eigenschaft, wie z.B. die Ladung, die im allgemeinen an verschiedenen Raumpunkten unterschiedlich ausgeprägt ist.

in diesem Zusammenhang ist die quantitative Analyse von Struktur-Wirkungsbeziehungen (*3D Quantitative Structure Activity Relationship*, 3D QSAR, [38]). 3D QSAR beruht auf der Korrelation zwischen den physikochemischen Eigenschaften einer Menge von Liganden und deren Affinität bezüglich eines Rezeptors. Dabei wird von einem Alignment der Liganden ausgegangen und es wird vorausgesetzt, daß die Liganden alle in der Orientierung des Alignments mit dem Rezeptor in Interaktion treten. In früheren QSAR Anwendungen mußten deshalb die Moleküle eine gemeinsame und einfach zu erkennende Basisstruktur besitzen, durch die auch das Alignment bestimmt war. Die Schlüssel mußten sozusagen alle ungefähr die gleiche Struktur besitzen. Die Erfahrung in der Wirkstoffentwicklung hat jedoch gezeigt, daß gerade die Betrachtung verschiedener Basisstrukturen zu einer ganzen Reihe neuer wirksamer Liganden führen kann. Die Betrachtung von Liganden mit verschiedenartigen geometrischen Strukturen erfordert Alignmentmethoden, die die physikochemischen Eigenschaften der Liganden berücksichtigen.

Das *Alignment-Problem* kann als Optimierungsproblem der Relativorientierung zweier Liganden bezüglich einer auf den physikochemischen Eigenschaften der Moleküle basierenden Bewertungsfunktion aufgefaßt werden. Das Resultat einer Berechnung ist zum einen die Relativorientierung der Liganden und zum anderen der Wert einer Gütefunktion des berechneten Alignments.

Im Rahmen des RELIWE-Projekts[3] ist ein Verfahren (FlexS, [41], [42]) zur Lösung des Alignment-Problems entwickelt worden. Das Verfahren ist für die Anwendung in 3D QSAR-Modellen und der Rezeptormodellierung konzipiert.

Das Verfahren FlexS berechnet zu einem gegebenen Paar Liganden eine Anzahl von Konformationsvorschlägen für das Alignment, welche die im Experiment beobachtete Bindungsgeometrie approximieren. Dabei wird einer der Liganden (*Referenzligand*) starr behandelt, während der Andere (*Testligand*) flexibel auf den Referenzliganden gefittet wird. Das zugrunde ligende Optimierungsproblem weist also sowohl Freiheitsgrade der Transformation (3 Rotations- und 3 Translations-Freiheitsgrade) als auch interne (*Konformations-*) Freiheitsgrade auf. Die internen Freiheitsgrade werden zusammen mit den externen mittels heuristischer Methoden der kombinatorischen Optimierung behandelt. Zu diesem Zweck wird ein diskretes Konformationsmodell bemüht. Soll anstelle eines Ligandenpaars eine großere Anzahl von Liganden berücksichtigt werden, so geschieht dies sequentiell.

Ziel der vorliegenden Arbeit ist es, eine Methode zur multidimensionalen Nachoptimierung zu konzipieren, welches es erlaubt vom diskreten Konformationsmodell abzuweichen und so eine Relaxation der Strukturen zu erreichen. Durch simultane Berücksichtigung einer ganzen Anzahl von Liganden sollen zusätzliche Informationen, die bei der paarweisen Überlagerung[4] unentdeckt bleiben, während der Nachoptimierung ausgenutzt werden, um so die Qualität des Alignments zu verbessern. Die Nachoptimierung ist als Bestandteil des Verfahrens FlexS implementiert worden.

Kapitel 2 formuliert das Alignment-Problem als mathematisches Optimierungsproblem und entwickelt die Grundlagen, um die Bewertungsfunktionen der verschiedenen Optimierungsmodelle

[3] RELIWE (*Rechnergestützte Analyse und Vorhersage von Rezeptor-Ligand-Wechselwirkungen*) ist ein vom Bundesministerium für Bildung, Wissenschaft, Forschung und Technologie Projekte (BMBF) gefördertes Verbundprojekt (Förderkennzeichen 01 IB 302 A) am Forschungszentrum Informationstechnik (GMD).

[4] Die Begriffe *Überlagerung* und *Alignment* werden im folgenden synonym verwendet

aufzustellen. Die Bewertungsfunktionen bilden damit die Grundlage für die zu konzipierende Nachoptimierung. Kapitel 3 behandelt die Bewertungsfunktionen und deren Eigenschaften. In Kapitel 4 werden die zur numerischen Optimierung der Bewertungsfunktion verwendeten Verfahren behandelt. In Kapitel 5 wird die konkrete Realisierung verschiedener für die Effizienz des Verfahrens wesentlicher Details der Implementierung beschrieben. Die Diskussion der Ergebnisse der Anwendungen des Verfahrens bilden den Abschluß der Arbeit. In Anhang A werden die Moleküle des Testdatensatzes beschrieben. In Anhang B werden alle wichtigen Symbole aufgeführt, die in dieser Arbeit benutzt werden. Und Anhang C enthält alle mathematischen Definitionen und Sätze, die benutzt worden sind.

Kapitel 2

Das mathematische Modell

2.1 Das Alignment-Problem und seine mathematische Formulierung

Gegeben seien die 3D-Strukturen zweier Ligand-Moleküle. Gesucht ist die räumliche Orientierung (Überlagerung/Alignment) der Liganden zueinander, die sie einnehmen, wenn sie in dieselbe Rezeptortasche binden (siehe Abbildung 2.1) Die Berechnung der Überlagerung der Liganden soll allerdings unabhängig von der im allgemeinen unbekannten Struktur der Rezeptortasche geschehen. Liganden, die in dieselbe Rezeptortasche binden, müssen Ähnlichkeiten in ihren räumlichen Eigenschaften aufweisen. Ziel ist es nun, das obige Alignment-Problem als mathematisches Optimierungsproblem zu formulieren (siehe hierzu auch die Arbeiten von C. Lemmen [41], C. Hiller [32] und M. Hauswald [30]).

Zu diesem Zwecke betrachtet man nun die folgenden formalen Definitionen:

Definition 2.1.1 (Konformation eines Moleküls)
Die Konformation eines Moleküls $\mathcal{M}$, d.h. die 3-dimensionale Struktur $\mathsf{C}(\mathcal{M})$, ist definiert durch die Atombindungen B, die Bindungslängen l_B, die Bindungswinkel α_B und durch die Torsionswinkel θ_B: θ_{ft}

$$\mathsf{C}(\mathcal{M}) \ := \ f_{\mathsf{C}}(\mathcal{M}, B, l_B, \alpha_B, \theta_B) \ \in \ \mathbb{R}^{3n_a}.$$

Dabei ist n_a die Anzahl der Atome. Der Konformationsraum eines Moleküls, d.h. die Menge aller Konformationen, ist gegeben durch

$$\mathcal{C}(\mathcal{M}) \ := \ \{ \ \mathsf{C}(\mathcal{M}) = f_{\mathsf{C}}(\mathcal{M}, B, l_B, \alpha_B, \theta_B) \mid l_B, \alpha_B, \theta_B \} \ \subset \ \mathbb{R}^{3n_a}.$$

Eine Konformation heißt zulässig, falls sie nicht selbst überlappend ist. $\diamond$

Um eine Aussage darüber zu treffen, wie gut eine Überlagerung bzw. das Alignment der Konformationen zweier Liganden zueinander ist, muß man eine Bewertungsfunktion definieren. Dazu

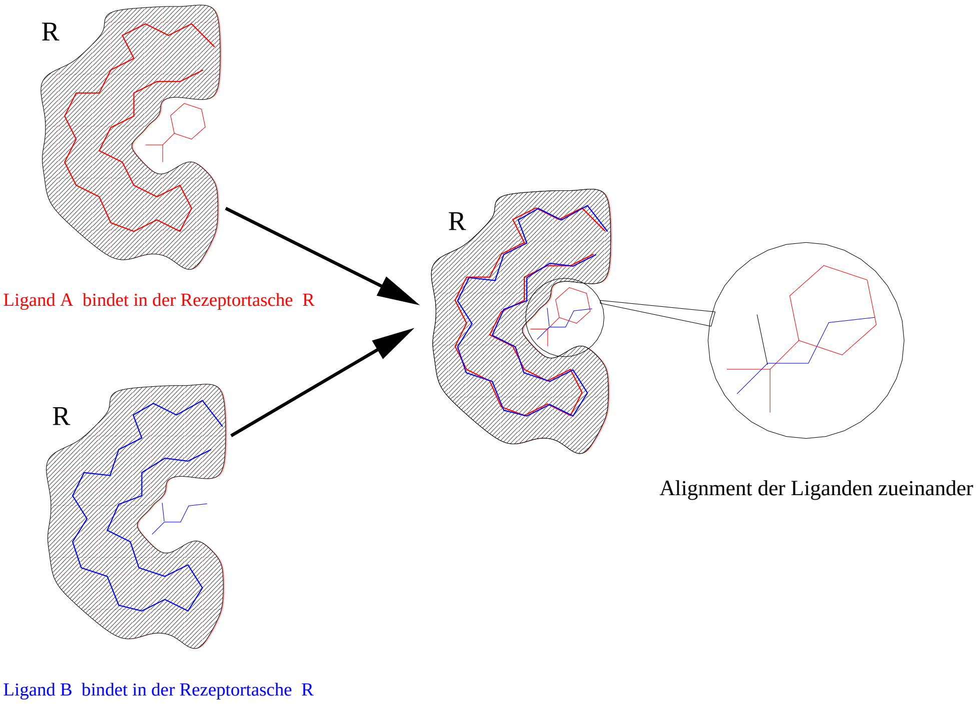


Abbildung 2.1: Die Bindung zweier Liganden in ein und dieselbe Rezeptortasche R (links) und deren räumliche Orientierung (Alignment) zueinander (rechts). Die räumliche Orientierung der Liganden zueinander ist durch die Überlagerung des Rezeptors mit sich selbst bestimmt.

betrachtet man zunächst die folgenden Indizes, die die räumliche Ähnlichkeit zweier Moleküle quantifizieren. Für Moleküle $\mathcal{A}$ und $\mathcal{B}$ mit zugehörigen Elektronendichtefunktionen $\rho_{\mathcal{A}}(\mu)$ und $\rho_{\mathcal{B}}(\mu)$ lautet der von Carbó et.al ([8]) vorgeschlagene sogenannte Carbó-Index

$$\mathsf{C}_{\mathcal{AB}} := \frac{\int_{\mathbb{R}} \rho_{\mathcal{A}}(\mu)\rho_{\mathcal{B}}(\mu)\ \mathsf{d}\mu}{\left(\sqrt{\int_{\mathbb{R}} (\rho_{\mathcal{A}}(\mu))^2\ \mathsf{d}\mu}\right)\left(\sqrt{\int_{\mathbb{R}} (\rho_{\mathcal{B}}(\mu))^2\ \mathsf{d}\mu}\right)}.$$

Hodgkin et.al ([33], [24]) schlugen dagegen einen modifizierten Index, den sogenannten Hodgkin-Index

$$\mathsf{H}_{\mathcal{AB}} := \frac{2\ \int_{\mathbb{R}} \rho_{\mathcal{A}}(\mu)\rho_{\mathcal{B}}(\mu)\ \mathsf{d}\mu}{\int_{\mathbb{R}} (\rho_{\mathcal{A}}(\mu))^2\ \mathsf{d}\mu\ +\ \int_{\mathbb{R}} (\rho_{\mathcal{B}}(\mu))^2\ \mathsf{d}\mu}$$

vor. Einen Überblick über weitere Ähnlichkeitsindizes findet man bei A.C. Good ([23]). Die vorgestellten Indizes basieren auf der Elektronendichte. Das bedeutet, daß hauptsächlich die räumliche Gestalt der Moleküle berücksichtigt wird. Tatsächlich wird die Bindung eines Liganden an einen Rezeptor von einer ganzen Reihe physikochemischer Eigenschaften bestimmt (siehe [6]). Dabei bedeutet der Begriff physikochemische Eigenschaft in diesem Zusammenhang eine lokale chemische Eigenschaft des Moleküls, wie beispielsweise die Ladung, die im allgemeinen an verschiedenen Raumpunkten unterschiedlich ausgeprägt ist. Außer der Elektronendichte kann für das Bindungsverhalten eines Liganden beispielsweise die Ladungsverteilung, die Verteilung von Hydrophobizität und die Lage von Wasserstoffdonatoren und -akzeptoren relevant sein.

Der vorliegende Ansatz basiert darauf, den relevanten physikochemischen Eigenschaften eines Moleküls eine räumlich lokalisierte Dichte[1] zuzuordnen. D.h. der Wirkungsbereich dieser Eigenschaft wird als räumliche Gestalt modelliert. Siehe hierzu Abbildung 2.2. Diese räumliche

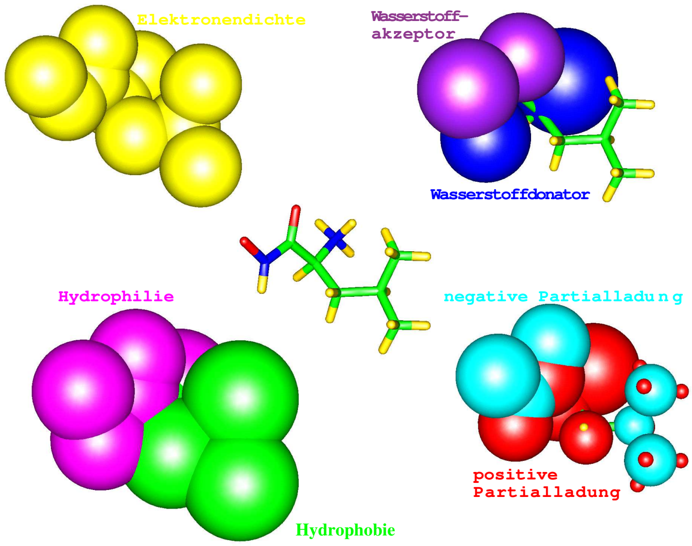


Abbildung 2.2: Räumliche Modellierung der physikochemischen Eigenschaften: L-leucylhydroxylamin, wie es im Komplex mit Thermolysin (PDB-Code für den Komplex: 4tln) beobachtet wird, mit der Darstellung der verschiedenen Eigenschaften durch Gaussfunktionen. Um die jeweiligen Gaussfunktionen zu illustrieren, nimmt man die Höhe h der entsprechenden Gaussfunktion und bildet dann eine Kugel mit Radius h. Entnommen aus [41]

Qualifizierung der zugrundeliegenden Eigenschaften soll es nun ermöglichen, ein Alignment der Konformationen zweier Moleküle zu bewerten. Dazu wird ein Alignmentindex definiert.

Definition 2.1.2 (Alignmentindex)
Gegeben seien die Moleküle $\mathcal{A}$ und $\mathcal{B}$ mit zulässigen Konformationen $\mathsf{C}_{\mathcal{A}}$ und $\mathsf{C}_{\mathcal{B}}$. Es sei $\mathcal{I}_{\mathsf{E}}$ die Indexmenge der relevanten physikochemischen Eigenschaften. Für jedes $i \in \mathcal{I}_{\mathsf{E}}$ sei $f_i^{\mathcal{M}}(\mathsf{C}_{\mathcal{M}})$, $\mathcal{M} = \mathcal{A}, \mathcal{B}$ die Funktion, die die räumlich lokalisierte Dichte der entsprechenden Eigenschaft, in Abhängigkeit der gegebenen Konformation des Moleküls modelliert. Der Alignmentindex bezüglich des Hodgkin-Indexes lautet nun

$$\mathsf{A}^{\mathsf{H}}_{\mathcal{A},\mathcal{B}} := \sum_{i \in \mathcal{I}_{\mathsf{E}}} \lambda_i \, \mathsf{H}_{\mathcal{A}\mathcal{B}} \, \left(f_i^{\mathcal{A}}(\mathsf{C}_{\mathcal{A}}), f_i^{\mathcal{B}}(\mathsf{C}_{\mathcal{B}}) \right), \quad \sum_{i \in \mathcal{I}_{\mathsf{E}}} \lambda_i = 1.$$

$\diamond$

[1] Siehe Definition 2.2.2.

Bemerkung: Der Alignmentindex bewertet mittels einer normierten Summe von Hodgkin-Indizes die Überlagerung zweier Molekülkonformationen anhand der räumlich lokalisierten Dichten der physikochemischen Eigenschaften. Je größer der Alignmentindex zweier Konformationen ist, um so besser ist die Überlagerung. Und je grösser der Alignmentindex ist, um so ähnlicher sind die Konformationen der beteiligten Moleküle bezüglich des zugrundeliegenden Modells, wobei das Modell auf der räumlichen Qualifizierung der physikochemischen Eigenschaften beruht.

Die Optimierungsaufgabe muß also lauten:
Zu je zwei gegebenen Molekülen finde man zulässige Konformationen, so daß der Alignmentindex dieser Molekülkonformationen maximal ist.
Hat man zu gegebenen Liganden $\mathcal{A}$ und $\mathcal{B}$ Molekülkonformationen gefunden, so daß der Alignmentindex optimal ist, kann man vielleicht mit Hilfe des zugrunde liegenden Modells abschätzen, ob gegebenenfalls Molekül $\mathcal{B}$ in die selbe Rezeptortasche bindet wie Molekül $\mathcal{A}$. Sollte das Molekül $\mathcal{B}$ nicht in die selbe Rezeptortasche wie das Molekül $\mathcal{A}$ binden, so sollte sich dies im Wert des Alignmentindex entsprechend ausdrücken.

Nachdem nun die Optimierungsaufgabe formuliert wurde, stellt sich die Frage, wie findet man die optimalen zulässigen Konformationen. Dazu wird nun der Transformationsraum definiert. Anstatt die Positionen der Atome direkt zu optimieren, wie zum Beispiel in der Moleküldynamik, wo die Positionen durch externe wie interne Anziehungs- bzw. Abstoßungskräfte direkt berechnet werden, werden die Atomkoordinaten durch globale wie lokale Transformationen (siehe Kapitel 2.4) bestimmt. Die direkte Optimierung der Atomkoordinaten ist ein $3n$-dimensionales Problem, wobei n die Anzahl der Atome im Liganden ist. Weiter könnten durch die direkte Optimierung der Atomkoordinaten Bindungen und damit z.B. Ringkonformationen im Molekül zerstört werden, oder es könnten Bindungslängen oder Bindungswinkel verändert werden. Um also zu verhindern, daß die Molekülkonformation zerstört wird, muß die Energiefunktion um Energieterme erweitert werden, die Abweichungen von den equilibrierten Bindungswinkeln und Bindungslängen erschweren (siehe hierzu z.B [45]). Da dies die Optimierungsaufgabe erheblich erschweren und verkomplizieren würde, wird in dieser Arbeit die globale und lokale Transformation einer gegebenen Molekülkonformation optimiert. Außerdem bleiben dabei Charakteristika wie Bindungen, Bindungslängen, Bindungswinkel und Ringkonformationen erhalten. Dabei ist unter der globalen Transformation die Rotation um ein Rotationszentrum sowie die Translation der Atomkoordinaten zu verstehen (siehe Kapitel 2.4.1). Würde man nur die globale Translation betrachten (wie z.B. in [32]), bliebe der zu optimierende Ligand starr. Um ein gewissen Maß an Flexibilität zu erreichen, werden die „flexiblen" Torsionswinkel opimiert. Ein Torsionswinkel heißt flexibel, wenn der Winkel variiert werden kann.[2] Die lokale Transformation beschreibt die lokale Drehung von einzelnen Molekülfragmente um molekülinterne Drehachsen (siehe Kapitel 2.4.2). Die formale Definition lautet daher:

Definition 2.1.3
Es sei $\mathcal{A}$ ein Molekül mit n_a Atomen und n_{ft} flexiblen Torsionswinkel. Weiter sei $\mathsf{C}_{\mathcal{A}} \in \mathbb{R}^{3n_a}$

[2]Siehe hierzu Kapitel 2.2.

eine zulässige Konformation des Moleküls $\mathcal{A}$. Die Abbildung

$$\begin{aligned} \mathsf{T}^{\mathcal{A}}_{global} : \mathbb{R}^3 \times \mathsf{SO}_3 \times \mathbb{R}^3 \times \mathbb{R}^{3n_a} &\rightarrow \mathbb{R}^{3n_a} \\ (t,\ S,\ c,\ \mathsf{C}_{\mathcal{A}}) &\rightarrow \mathsf{T}^{\mathcal{A}}_{global}\ (t,\ S,\ c,\ \mathsf{C}_{\mathcal{A}}) \end{aligned}$$

heißt globale Transformation des Moleküls $\mathcal{A}$, wobei t der Translationsvektor, S die Rotationsmatrix und c das Rotationszentrum ist. Die Abbildung

$$\begin{aligned} \mathsf{T}^{\mathcal{A}}_{lokal}\ : ([0;\ 2\pi))^{n_{ft}} \times \mathbb{R}^{3n_a} &\rightarrow \mathbb{R}^{3n_a} \\ (\theta^{\mathcal{A}}_{ft},\ \mathsf{C}_{\mathcal{A}}) &\rightarrow \mathsf{T}^{\mathcal{A}}_{lokal}\ (\theta^{\mathcal{A}}_{ft},\ \mathsf{C}_{\mathcal{A}}) \end{aligned}$$

heißt lokale Transformation des Moleküls $\mathcal{A}$, wobei $\theta^{\mathcal{A}}_{ft} := (\theta^{\mathcal{A}}_1, \ldots, \theta^{\mathcal{A}}_{n_{ft}})$ der Vektor der internen Rotationswinkel des Moleküls $\mathcal{A}$ ist. Damit ist der Transformationsraum des Moleküls $\mathcal{A}$ gegeben durch

$$\mathcal{TR}_{\mathcal{A}}\ :=\ \mathbb{R}^3 \times \mathsf{SO}_3 \times \mathbb{R}^3 \times ([0;\ 2\pi))^{n_{ft}}.$$

Der Transformationsoperator des Moleküls $\mathcal{A}$ ist definiert durch

$$\begin{aligned} &\mathsf{T}_{\mathcal{A}}\ : \mathcal{TR}_{\mathcal{A}} \times \mathbb{R}^{3n_a} \rightarrow \mathbb{R}^{3n_a} \\ &\left(t,\ S,\ c,\ \theta^{\mathcal{A}}_{ft},\ \mathsf{C}_{\mathcal{A}}\right) \rightarrow \mathsf{T}_{\mathcal{A}}\left(t,\ S,\ c,\ \theta^{\mathcal{A}}_{ft},\ \mathsf{C}_{\mathcal{A}}\right) \end{aligned}$$

mit

$$\mathsf{T}_{\mathcal{A}}\ \left(t,\ S,\ c,\ \theta^{\mathcal{A}}_{ft},\ \mathsf{C}_{\mathcal{A}}\right)\ =\ \mathsf{T}^{\mathcal{A}}_{global}\ \left(t,\ S,\ c,\ \mathsf{T}^{\mathcal{A}}_{lokal}\ (\theta^{\mathcal{A}}_{ft},\ \mathsf{C}_{\mathcal{A}})\right).$$

Zunächst werden die molekülinternen Rotationen ausgeführt. Erst danach wird das Molekül global bewegt. ◇

Damit läst sich das obige Optimierungsproblem wie folgt beschreiben:

Definition 2.1.4 (Optimierungsaufgabe)
Gegeben seien die Moleküle $\mathcal{A}$ und $\mathcal{B}$ mit zulässigen Konformationen $\mathsf{C}_{\mathcal{A}}$ und $\mathsf{C}_{\mathcal{B}}$. Es sei $\mathcal{I}_{\mathsf{E}}$ die Indexmenge der relevanten physikochemischen Eigenschaften. Für jedes $i \in \mathcal{I}_{\mathsf{E}}$ sei $f^{\mathcal{M}}_i(\mathsf{C}_{\mathcal{M}})$, $\mathcal{M} = \mathcal{A}, \mathcal{B}$ die Funktion, die die räumlich lokalisierte Dichte der entsprechenden Eigenschaft, in Abhängigkeit der gegebenen Konformation des Moleküls modelliert. Weiter seien $\lambda_i \in \mathbb{R}_{\geq 0}$, $i \in \mathcal{I}_{\mathsf{E}}$ Koeffizienten um die physikochemischen Eigenschaften zu gewichten mit der Eigenschaft, daß $\sum_{i \in \mathcal{I}_{\mathsf{E}}} \lambda_i = 1$ gilt. Es sei $\mathsf{T}_{\mathcal{M}}$ der zum Liganden $\mathcal{M} = \mathcal{A}$,$\mathcal{B}$ gehörende Transformationsoperator und es sei $\mathcal{TR}_{\mathcal{M}}$, $\mathcal{M} = \mathcal{A}$,$\mathcal{B}$ der entsprechende Transformationsraum. Nun lassen sich zwei Modelle unterscheiden. Dabei werden die Bezeichnungen aus Definition 2.1.3 verwendet.

Modell 2.1 (Überlagerung auf eine Referenzstruktur)
Die Konformation $\mathsf{C}_{\mathcal{A}}$ des Liganden $\mathcal{A}$ ist eine Referenzstruktur, d.h. der Ligand bindet in dieser Konformation in die entsprechende Rezeptortasche. Die Konformation des Liganden $\mathcal{B}$ soll auf

die rigide gehaltende Konformation $\mathsf{C}_{\mathcal{A}}$ überlagert werden. Die mathematische Optimierungsaufgabe lautet daher:

$$\max_{\left(t,\ S,\ c,\ \theta_{ft}^{\mathcal{B}}\right)\ \in\ \mathcal{TR}_{\mathcal{B}}} \left[\mathsf{A}^{\mathsf{H}}_{\mathcal{A},\mathcal{B}}\ \left\{\left(f_i^{\mathcal{A}}\ (\mathsf{C}_{\mathcal{A}})\right)_{i\in\mathcal{I}_{\mathsf{E}}},\ \left(f_i^{\mathcal{B}}\ \left(\mathsf{T}_{\mathcal{B}}\ \left(t,\ S,\ c,\ \theta_{ft}^{\mathcal{A}},\ \mathsf{C}_{\mathcal{B}}\right)\right)\right)_{i\in\mathcal{I}_{\mathsf{E}}}\right\}\right]$$

$$\textbf{\textit{s.t.}}\ \ \mathsf{T}_{\mathcal{B}}\ \left(t,\ S,\ c,\ \theta_{ft}^{\mathcal{A}},\ \mathsf{C}_{\mathcal{B}}\right)\ \textit{ist zulässig.}$$

Modell 2.2 (Gegenseitige Überlagerung zweier Liganden)
Die Konformationen der Liganden $\mathcal{A}$ und $\mathcal{B}$ sollen gegenseitig überlagert werden. Dies ist z.B. dann sinnvoll, wenn von $\mathcal{A}$ keine Referenzstruktur bekannt ist. Die mathematische Optimierungsaufgabe lautet daher:

$$\max_{\substack{\left(t_{\mathcal{A}},\ S_{\mathcal{A}},\ c_{\mathcal{A}},\ \theta_{ft}^{\mathcal{A}}\right)\ \in\ \mathcal{TR}_{\mathcal{A}},\\ \left(t_{\mathcal{B}},\ S_{\mathcal{B}},\ c_{\mathcal{B}},\ \theta_{ft}^{\mathcal{B}}\right)\ \in\ \mathcal{TR}_{\mathcal{B}}}} \left[\mathsf{A}^{\mathsf{H}}_{\mathcal{A},\mathcal{B}}\ \left\{\ \left(f_i^{\mathcal{A}}\ \left(\mathsf{T}_{\mathcal{A}}\ \left(t_{\mathcal{A}},\ S_{\mathcal{A}},\ c_{\mathcal{A}},\ \theta_{ft}^{\mathcal{A}},\ \mathsf{C}_{\mathcal{A}}\right)\right)\right)_{i\in\mathcal{I}_{\mathsf{E}}},\right.\right.$$

$$\left.\left.\left(f_i^{\mathcal{B}}\ \left(\mathsf{T}_{\mathcal{B}}\ \left(t_{\mathcal{B}},\ S_{\mathcal{B}},\ c_{\mathcal{B}},\ \theta_{ft}^{\mathcal{B}},\ \mathsf{C}_{\mathcal{B}}\right)\right)\right)_{i\in\mathcal{I}_{\mathsf{E}}}\right\}\right]$$

$$\textbf{\textit{s.t.}}\ \mathsf{T}_{\mathcal{A}}\ \left(t_{\mathcal{A}},\ S_{\mathcal{A}},\ c_{\mathcal{A}},\ \theta_{ft}^{\mathcal{A}},\ \mathsf{C}_{\mathcal{A}}\right),\ \mathsf{T}_{\mathcal{B}}\ \left(t_{\mathcal{B}},\ S_{\mathcal{B}},\ c_{\mathcal{B}},\ \theta_{ft}^{\mathcal{B}},\ \mathsf{C}_{\mathcal{B}}\right)\ \textit{sind zulässig.}$$

Das Modell 2.2 kann dahingehend verallgemeinert werden, daß eine variable Anzahl von Molekülen gegenseitig überlagert wird. Werden die Cayley-Klein-Parametrisierung der SO_3 (Quaternionen) verwendet, so ist die Optimierungsaufgabe 2.1 ein $10+n_{ft}$-dimensionales Problem und die Optimierungsaufgabe 2.2 ein $20+n_{ft}^{\mathcal{A}}+n_{ft}^{\mathcal{B}}$-dimensionales Problem. d.h. die Größenordnung des Problems hängt im wesentlichen von der Anzahl der internen Rotationsparameter bzw. von der Anzahl der flexiblen Moleküle und der Anzahl aller internen Rotationsparameter ab. $\diamond$

In Kapitel 2.2 werden die Daten, auf denen die zugrundeliegende Molekülstruktur und die Modellierung der physikochemischen Eigenschaften beruhen, beschrieben und daraus weitere Informationen, die für die numerische Optimierung bedeutend sind, hergeleitet. Die mathematische Definitionen der hier betrachteten Indizes und deren Eigenschaften werden in Kapitel 2.3 behandelt. In Kapitel 2.4 wird der Transformationsoperator entwickelt und beschrieben. Die Gaussfunktionen, die die Grundlage für die Modellierung der physikochemischen Eigenschaften sind, werden in Kapitel 2.5 vorgestellt. 2.6 beschreibt das Lennard-Jones-Potential, das dazu verwendet wird, um die Nebenbedingung zu erfüllen, daß eine gültige Konformation nicht selbst überlappend ist. In Kapitel 2.7 wird eine Bewertungsfunktion für die Torsionswinkel entwickelt.

2.2 Die zu Verfügung stehenden Daten eines Moleküls

Es sei $\mathcal{A}$ ein Molekül mit n_a Atomen und zulässiger Konformation $\mathsf{C}_{\mathcal{A}} \subset \mathbb{R}^{3n_a}$. Weiter sei $K_{\mathcal{A}} := K(\mathsf{C}_{\mathcal{A}}) := \{a_1, \ldots, a_{n_a}\} \subset \mathbb{R}^3,\ a_i \neq a_j,\ i \neq j$, die Menge der n_a Atomkoordinaten bzgl. der gegebenen Molekülkonformation $\mathsf{C}_{\mathcal{A}}$.[3] Für das Atom $a_i,\ \ i = 1, \ldots, n_a$, bezeichnet

[3] Mit a_i werden sowohl das Atom als auch die zugehörigen Atomkoordinaten bezeichnet .

$\sigma_i \in \mathbb{R}_{>0}$ den Van-der-Waals-Radius des Atoms a_i (siehe [41]). Dann ist der Vektor der Van-der-Waals-Radien durch $\Sigma_{\mathcal{A}} := \{\sigma_1, \ldots, \sigma_{n_a}\} \in \mathbb{R}^{n_a}_{>0}$ gegeben. Die Van-der-Waals-Radien werden für die Berechnung des Lennard-Jones-Potentials benötigt.[4] Es sei $B_{\mathcal{A}} := \{e_{ij} = \{a_i,\ a_j\} \mid$ die Elektronenhüllen der Atome a_i und $a_j,\ i \neq j$, haben mindestens ein Elektronpaar gemeinsam $\} \subset \mathcal{P}_2(K_{\mathcal{A}})$ die Menge der Atombindungen des Moleküls $\mathcal{A}$.[5] Dann bildet $G_{\mathcal{A}} = (K_{\mathcal{A}},\ B_{\mathcal{A}})$ einen zusammenhängenden Graphen[6] in $\mathbb{R}^3$. Weiter sei $\Phi^{\mathcal{A}}_{EB} \colon B_{\mathcal{A}} \to \{0,\ 1\}$ eine Abbildung auf $B_{\mathcal{A}}$, die die Einfach-Bindungen des Moleküls beschreibt, d.h.

$$\Phi^{\mathcal{A}}_{EB}(e_{ij}) \;=\; \begin{cases} 1 & \text{, falls die Elektronenhüllen der Atome } a_i \text{ und } a_j \text{ sich ein} \\ & \text{Elektronenpaar teilen} \\ 0 & \text{, sonst.} \end{cases}$$

Mit der Abbildung $\Phi^{\mathcal{A}}_{FT} \colon B_{\mathcal{A}} \to \{0,\ 1\}$ auf $B_{\mathcal{A}}$ werden diejenigen Bindungen charakterisiert, an denen *flexible* Torsionswinkel betrachtet werden.

Es sei die Atomkonfiguration A-B-C-D gegeben. Wird die Bindung B-C kontrahiert, so ist der Torsionswinkel definiert als der von den Bindungen A-B und C-D eingeschlossene Winkel, d.h.

$$\alpha \;=\; \measuredangle(A-B, C-D) \;:=\; \arccos\left(\frac{< \vec{v}(A-B),\ \vec{v}(C-D) >_2}{\|\vec{v}(A-B)\|_2 \|\vec{v}(C-D)\|_2}\right).$$

Dabei ist $\vec{v}(A-B) \in \mathbb{R}^3$ ein Vektor, der die Bindung A-B repräsentiert. An der Konfiguration in Abbildung 2.3 kann man sehen, wie der Torsionswinkel definiert ist. Dabei wird die folgende Konvention eingeführt: *der Torsionswinkel α wird an der kontrahierten bzw. an der mittleren Atombindung der zugrundeliegenden Konfiguration gemessen.* Damit ist gemeint, daß der Winkel α der Bindung B-C zugeordnet wird. Die Bindung B-C wird zu einer Drehachse für lokale Rotationen, wie sie Kapitel 2.4.2 beschrieben werden. Nun werden Bindungen, an denen Torsionswinkel *gemessen* werden können, wie folgt charakterisiert: an $e = \{a_k,\ a_l\} \in B_{\mathcal{A}}$ kann genau dann, ein Torsionswinkel *gemessen* werden, wenn $\mathsf{d}_{G_{\mathcal{A}}}(a_k) > 1$ und $\mathsf{d}_{G_{\mathcal{A}}}(a_l) > 1$ gilt. Dabei ist $\mathsf{d}_{G_{\mathcal{A}}}(a_k)$ die Anzahl der mit a_k inzidenten Kanten aus $B_{\mathcal{A}}$, d.h. der Grad von a_k.[7]
Flexibel heißt in diesem Zusammenhang nun, daß der Torsionswinkel variiert werden kann. Dabei kommen flexible Torsionswinkel nur an Einfach-Bindungen vor. Aber bestimmte Atomkonfigurationen (wie z.B. die Konfiguration C-S-N-C in Abbildung 2.3)) sind von der Natur aus rigide, d.h., obwohl die Bindung einfach ist, nimmt der Torsinswinkel nur einen festen Wert an. Andere Torsionswinkelwinkel müssen künstlich rigide gehalten werden, um die Molekülkonformation nicht zu zerstören. In Abbildung 2.4 sieht man eine Ringkonformation. Würde man an den Einfach-Bindungen andere Torsionswinkel einstellen, könnte dadurch die Ringkonformation zerstört werden. Insbesondere kann die Ringkonformation durch lokale Rotationen, wie sie in Kapitel 2.4.2 beschrieben werden, aufgebrochen werden. Daher nehmen auch diese Winkel nur einen festen Wert an. Damit sind insbesondere alle Torsionswinkel, die an Einfach-Bindungen in Ringkonformationen *gemessen* werden können, rigide. Formal bedeutet dies: Ist die Bindung

[4] Siehe Kapitel 2.6.
[5] Definition: $\mathcal{P}_2(V) := \{X \mid X \subseteq V,\ |X| = 2\}$.
[6] Zu Graphen, Bäume, Wege u.a. siehe Definition C.1 in Anhang C.
[7] Siehe Definition C.1.

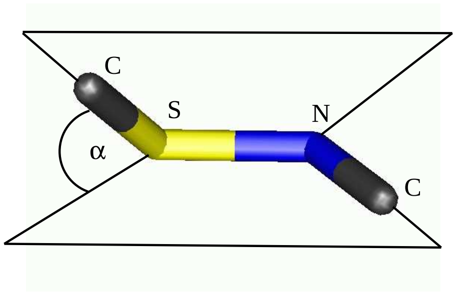


Abbildung 2.3: An der Bindung S-N wird der Torsionswinkel α *gemesen.*

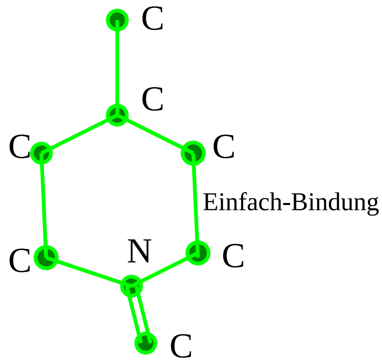


Abbildung 2.4: Rigide Torsionswinkel in einer Ringkonformation

$e \in B_{\mathcal{A}}$, mit $\Phi^{\mathcal{A}}_{EB}(e) = 1$, Kante eines Kreises C in $G_{\mathcal{A}}$, so wird der entsprechende Torsionswinkel α_e gegebenenfalls künstlich fest gehalten. D.h. α_e nimmt nur den Wert an, der aus der Konformation $\mathsf{C}_{\mathcal{A}}$ unmittelbar abgeleitet werden kann. Insbesondere ist also eine Einfach-Bindung $e \in B_{\mathcal{A}}$, an der flexible Torsionswinkel betrachtet werden, dadurch charakterisiert, daß der Graph $G_{\mathcal{A}} \setminus e := (K_{\mathcal{A}},\ B_{\mathcal{A}} \setminus e)$ aus zwei Zusammenhangskomponenten besteht. Damit erhält man nun die folgende Definition für die Abbildung $\Phi^{\mathcal{A}}_{FT}$:

$$\Phi^{\mathcal{A}}_{FT}(e) \;:=\; \begin{cases} 1 & ,\text{falls: } \left(\Phi^{\mathcal{A}}_{EB}(e) = 1\right) \;\wedge \\ & (\text{ein Torsionswinkel kann an der Bindung } e = \{a_k,\ a_l\} \text{ \textit{gemessen}} \\ & \text{werden, d.h. } \mathsf{d}_{G_{\mathcal{A}}}(a_k) > 1 \;\wedge\; \mathsf{d}_{G_{\mathcal{A}}}(a_l) > 1) \;\wedge \\ & (G_{\mathcal{A}} \setminus e \text{ besteht aus zwei Zusammenhangskomponenten})\,, \\ 0 & ,\text{sonst.} \end{cases}$$

Nun setzt man $B^{\mathcal{A}}_{FT} := \{e \in B_{\mathcal{A}} \mid \Phi^{\mathcal{A}}_{FT}(e) = 1\}$. Damit ist $B^{\mathcal{A}}_{FT}$ die Menge der Einfach-Bindungen, an denen der Torsionswinkel variiert werden kann. Mit $n_{ft} := \left|B^{\mathcal{A}}_{FT}\right|$ wird somit die Anzahl der flexiblen Torsionswinkel bezeichnet.

Betrachtet man nun $G'_{\mathcal{A}} = (K_{\mathcal{A}},\ E_{\mathcal{A}})$, $E_{\mathcal{A}} := B_{\mathcal{A}} \setminus B^{\mathcal{A}}_{FT}$, so ist dies ein Graph mit $n_{ft}+1$ Zusammenhangskomponenten[8], den Molekülfragmenten. Die Zusammenhangskomponenten von $G'_{\mathcal{A}}$ werden mit $F_i = (V_i, E(F_i))$, $i = 0, \ldots, n_{ft}$, bezeichnet, wobei für $i = 0, \ldots, n_{ft}$, $V_i := V(F_i) \subset K_{\mathcal{A}}$ und $E(F_i) \subset E_{\mathcal{A}}$ gilt. Dann ist $\bigcup_{i=0}^{n_{ft}} V_i$ eine disjunkte Zerlegung von $K_{\mathcal{A}}$. Indem man die Zusammenhangskomponenten F_i zu je einem Knoten f_i kontrahiert, erhält man einen ungerichteten Graphen $T_{\mathcal{A}}$, mit dessen Hilfe man dann die lokalen Drehstrukturen des Moleküls beschreiben kann. Die Knotenmenge des Graphen $T_{\mathcal{A}}$ ist durch $F_{\mathcal{A}} := \{f_0,\ \ldots,\ f_{n_{ft}}\}$, gegeben. Die folgende Abbildung ordnet jeder Atomkoordinate $a_i \in K_{\mathcal{A}}$ den Knoten f_j zu, falls a_i zum

[8]Dies folgt aus der Definition von $\Phi^{\mathcal{A}}_{FT}$ mittels Induktion über n_{ft}.

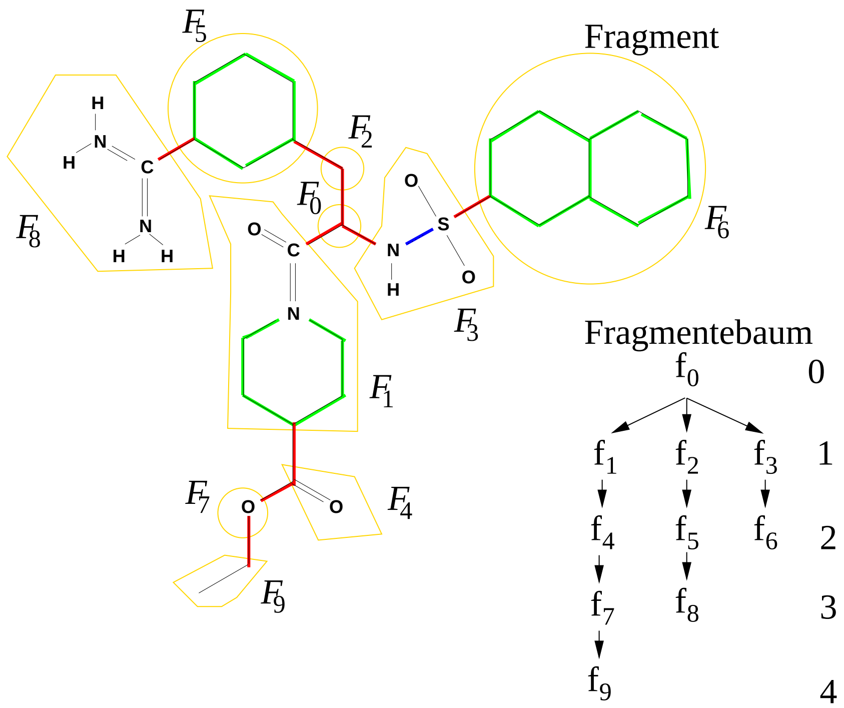


Abbildung 2.5: Das Molekül Napap (PDB-Code: 1dwd) mit seinen Fragmenten und dem zugehörigem Fragmentebaum. Für die roten Bindungen gilt $\Phi^{\mathcal{A}}_{FT}(e) = 1$, d.h. diesen Bindungen werden flexible Torsionswinkel zugeordnet. Blaue bzw. grüne Bindungen sind Bindungen mit rigidem Torsionswinkel (natürlich bzw. künstlich).

Molekülfragment F_j gehört. Dazu setzt man

$$\Psi_{K_\mathcal{A},F_\mathcal{A}} : K_\mathcal{A} \to F_\mathcal{A}$$
$$\Psi_{K_\mathcal{A},F_\mathcal{A}}(a_i) := f_j, \text{ falls } a_i \in V_j := V(F_j) \subset K_\mathcal{A}.$$

Dann ist der Graph durch $T_\mathcal{A} := (F_\mathcal{A}, E^\mathcal{A}_{FT})$, mit

$$\begin{aligned} E^\mathcal{A}_{FT} &:= \Psi_{K_\mathcal{A},F_\mathcal{A}}\left(B^\mathcal{A}_{FT}\right) \\ &:= \left\{\{\Psi_{K_\mathcal{A},F_\mathcal{A}}(a_i),\ \Psi_{K_\mathcal{A},F_\mathcal{A}}(a_j)\} \mid \{a_i,\ a_j\} \in B^\mathcal{A}_{FT}\right\} \subset \mathcal{P}_2(F_\mathcal{A}), \end{aligned}$$

gegeben. Man beachte, das Mehrfachkanten auf Grund der Definition von $B^\mathcal{A}_{FT}$ nicht vorkommen können.

Bemerkung: $T_\mathcal{A}$ ist zusammenhängend. Es seien $f_i,\ f_j \in F_\mathcal{A},\ f_i \neq f_j$. Dann existieren $a_k \in (\Psi_{K_\mathcal{A},F_\mathcal{A}})^{-1}(f_i) = V_i$ und $a_l \in (\Psi_{K_\mathcal{A},F_\mathcal{A}})^{-1}(f_j) = V_j$.[9] Da $G_\mathcal{A}$ zusammenhängend ist, existiert ein a_k-a_l−Weg $P_{a_k,a_l} := a_k \dots a_l$ in $G_\mathcal{A}$. Dann ist $\Psi_{K_\mathcal{A},F_\mathcal{A}}(P_{a_k,a_l}) := \Psi_{K_\mathcal{A},F_\mathcal{A}}(a_k)\dots\Psi_{K_\mathcal{A},F_\mathcal{A}}(a_l)$ ein f_i-f_j−Weg in $T_\mathcal{A}$. Aus $\left|E^\mathcal{A}_{FT}\right| = n_{ft} = |F_\mathcal{A}| - 1$ folgt mit Satz C.2 4) aus Anhang C, daß $T_\mathcal{A}$ ein Baum ist. $T_\mathcal{A}$ heißt Fragmentebaum. Mit

$$\mathcal{B}_{K_\mathcal{A}}(T_\mathcal{A}) = \left\{ \bigcup_{i=0}^{n_{ft}} V_i,\ T_\mathcal{A},\ \Psi_{K_\mathcal{A},F_\mathcal{A}} \right\} \tag{2.1}$$

[9] $(\Psi_{K_\mathcal{A},F_\mathcal{A}})^{-1}(f_k) := \{a_l \in K_\mathcal{A} \mid \Psi_{K_\mathcal{A},F_\mathcal{A}}(a_l) = f_k\}$ Urbild von f_k in $K_\mathcal{A}$.

ist eine Baumstruktur auf $K_{\mathcal{A}}$ bzgl. $T_{\mathcal{A}}$ gegeben.[10] Analog zur Definition 2.4.13 werden nun die Kanten von $T_{\mathcal{A}}$ gerichtet: Nach Satz C.2 gibt es für je zwei Knoten aus $F_{\mathcal{A}}$ genau einen Weg in $T_{\mathcal{A}}$, der die beiden Knoten verbindet. Für $f_i \in F_{\mathcal{A}} \setminus f_0$ sei $P_{f_0,f_i} = (V(P_{f_0,f_i}), E(P_{f_0,f_i}))$ der eindeutig bestimmte f_0-f_i−Weg in $T_{\mathcal{A}}$. Die Kanten aus $E(P_{f_0,f_i})$ werden so gerichtet, daß $\vec{P}_{f_0,f_i}$ ein gerichteter f_0-f_i−Weg ist. Der so entstehende gerichtete Baum $\vec{T}_{\mathcal{A}} = (F_{\mathcal{A}},\ \vec{E}^{\mathcal{A}}_{FT})$ enthält für jeden Knoten $f_i \in F_{\mathcal{A}},\ f_i \neq f_0$, einen gerichteten f_0-f_i−Weg. Damit ist f_0 die Wurzel des gerichteten Baumes $\vec{T}_{\mathcal{A}}$. Ersetzt man nun in Gleichung (2.1) $T_{\mathcal{A}}$ durch $\vec{T}_{\mathcal{A}}$, so erhält man eine gerichtete Baumstruktur auf $K_{\mathcal{A}}$.

In Abbildung 2.5 sieht man das Molekül Napap (PDB-Code: 1dwd) mit seinen Fragmenten und dem zugehörigen gerichteten Fragmentebaum $\vec{T}_{\mathcal{A}}$ (der Baum hat die Tiefe 4).[11] Mit Hilfe des gerichteten Fragmentebaumes $\vec{T}_{\mathcal{A}}$ können nun die lokalen Drehungen in einem Molekül beschrieben werden (siehe dazu Kapitel 2.4.2 und 2.4.3). Dazu muß man aber noch festlegen, welches Molekülfragment mit F_0 bezeichnet wird. Das Fragment F_0 ist die einzige Komponente des Moleküls, auf welche nur die globale Transformation wirkt. Der zugehörige Knoten f_0 bezeichnet dann die Wurzel des Baumes $\vec{T}_{\mathcal{A}}$. Die übrigen Komponenten werden durch einen BFS-Algorithmus (siehe hierzu [9]), der auf den Baum $T_{\mathcal{A}}$ angewendet wird, entsprechend durchnumeriert. F_0 wird als Hauptfragment bezeichnet, während die anderen Komponenten Nebenfragmente heißen. Das Hauptfragment sollte so gewählt werden, daß der spätere gerichtete Baum $\vec{T}_{\mathcal{A}}$ balanciert ist. D.h. die gerichteten Wege $\vec{P}_{f_0,f_i}$ in $\vec{T}_{\mathcal{A}}$ mit $f_i \neq f_0$ sind möglichst kurz. Hat man das Hauptfragment festgelegt, werden die Kanten von $T_{\mathcal{A}}$ wie oben beschrieben gerichtet.

Wie bereits kurz erwähnt worden ist, werden die Einfach-Bindungen, an denen flexible Torsionswinkel *gemessen* werden, als Drehachsen verwendet. Dazu sei $\{a_k,\ a_l\} \in B^{\mathcal{A}}_{FT}$ eine Bindung mit $\Phi^{\mathcal{A}}_{FT}(\{a_k,\ a_l\}) = 1$ und es gelte $(\Psi_{K_{\mathcal{A}},F_{\mathcal{A}}}(a_k),\ \Psi_{K_{\mathcal{A}},F_{\mathcal{A}}}(a_l)) \in \vec{E}^{\mathcal{A}}_{FT}$. Dann werden die Atome, die zum Nebenfragment F_i mit $\Psi_{K_{\mathcal{A}},F_{\mathcal{A}}}(a_l) = f_i$ gehören, zu dem auch das Atom a_l gehört, um die lokale Gerade $g_{f_i}(\lambda) := a_k + \lambda(a_l - a_k)$, $\lambda \in \mathbb{R}$ gedreht. Die lokale Rotation kann dann wie folgt beschrieben werden:[12]

$$a_j^{rot} \;=\; a_k \;+\; \mathsf{B}\left(\frac{d_{f_i}}{\|d_{f_i}\|_2},\ \varphi_{f_i}\right)\ (a_j - a_k)\,,\ \forall a_j \in V(F_i), \tag{2.2}$$

wobei $\mathsf{B}(\cdot,\ \cdot)$ die Rotationsmatrix aus Satz 2.4.3 ist, a_k das lokale Rotationszentrum, $d_{f_i} = \lambda(a_l - a_k) \neq 0$, eine Rotationsachse in Richtung der Geraden g_{f_i} und $\varphi_{f_i} \in \mathbb{R}$ der Rotationsparameter (Rotationswinkel) ist. Um also lokale Rotationen innerhalb des Moleküls zu beschreiben, braucht man lokale Rotationszentren und Drehachsen. Die Abbildungen $\mathsf{R}_{F_{\mathcal{A}},K_{\mathcal{A}}} : F_{\mathcal{A}} \setminus f_0 \to K_{\mathcal{A}}$, $\mathsf{S}_{F_{\mathcal{A}},K_{\mathcal{A}}} : F_{\mathcal{A}} \setminus f_0 \to K_{\mathcal{A}}$ und $\mathsf{D}_{F_{\mathcal{A}},\mathbb{R}^3\setminus\{0\}} : F_{\mathcal{A}} \setminus f_0 \to \mathbb{R}^3 \setminus 0$, die analog zu Definition 2.4.14 gebildet werden, liefern die gesuchten Informationen. Dazu geht man wie folgt vor: Für jedes $f_i \in F_{\mathcal{A}} \setminus f_0$ existiert genau eine gerichtete Kante $(f_{k_i}, f_i) \in \vec{E}^{\mathcal{A}}_{FT}$, da f_0 Wurzel des Baumes $\vec{T}_{\mathcal{A}}$ ist. Nach Definition von $E^{\mathcal{A}}_{FT}$ und insbesondere von $B^{\mathcal{A}}_{FT}$ existiert genau eine Kante $\{a_q,\ a_p\} \in B^{\mathcal{A}}_{FT}$, mit $\Psi_{K_{\mathcal{A}},F_{\mathcal{A}}}(a_q) = f_{k_i}$ und $\Psi_{K_{\mathcal{A}},F_{\mathcal{A}}}(a_p) = f_i$. Dabei gehört a_q zum Fragment

[10] Siehe Definition 2.4.13

[11] 1dwd ist eine Kurzbezeichnung aus der Protein Data Bank (PDB, [5]).

[12] Siehe hierzu Kapitel 2.4.2.

F_{k_i} und a_p zum Fragment F_i. Nun setzt man

$$\mathsf{R}_{F_{\mathcal{A}},K_{\mathcal{A}}}(f_i) := a_q, \quad \mathsf{S}_{F_{\mathcal{A}},K_{\mathcal{A}}}(f_i) := a_p, \quad \mathsf{D}_{F_{\mathcal{A}},\mathbb{R}^3\setminus\{0\}}(f_i) := \mathsf{S}_{F_{\mathcal{A}},K_{\mathcal{A}}}(f_i) - \mathsf{R}_{F_{\mathcal{A}},K_{\mathcal{A}}}(f_i).$$

Nach Lemma 2.4.15 sind die Abbildungen $\mathsf{R}_{F_{\mathcal{A}},K_{\mathcal{A}}}$, $\mathsf{S}_{F_{\mathcal{A}},K_{\mathcal{A}}}$ und $\mathsf{D}_{F_{\mathcal{A}},\mathbb{R}^3\setminus\{0\}}$ wohldefiniert. Mit den Abbildungen $\mathsf{R}_{F_{\mathcal{A}},K_{\mathcal{A}}}$ und $\mathsf{S}_{F_{\mathcal{A}},K_{\mathcal{A}}}$ läßt sich die natürliche Beziehung zwischen den Nebenfragmenten und den Bindungen mit flexiblen Torsionswinkel durch die folgende Abbildung $\mathcal{I}^{\mathcal{A}}_{F,B}\colon F_{\mathcal{A}} \setminus \{f_0\} \to B^{\mathcal{A}}_{FT}$ charakterisieren, indem man für $f_i \neq f_0$

$$\mathcal{I}^{\mathcal{A}}_{F,B}(f_i) := \{\mathsf{R}_{F_{\mathcal{A}},K_{\mathcal{A}}}(f_i),\ \mathsf{S}_{F_{\mathcal{A}},K_{\mathcal{A}}}(f_i)\}$$

setzt. Auf Grund der Definition von $B^{\mathcal{A}}_{FT}$, $\mathsf{R}_{F_{\mathcal{A}},K_{\mathcal{A}}}$ und $\mathsf{S}_{F_{\mathcal{A}},K_{\mathcal{A}}}$ ist diese Abbildung injektiv und surjektiv.

Nun läßt sich Gleichung (2.2) wie folgt darstellen:

$$a_j^{rot} \;=\; \mathsf{R}_{F_{\mathcal{A}},K_{\mathcal{A}}}(f_i) \;+\; \mathsf{B}\left(\frac{\mathsf{D}_{F_{\mathcal{A}},\mathbb{R}^3\setminus\{0\}}(f_i)}{\|\mathsf{D}_{F_{\mathcal{A}},\mathbb{R}^3\setminus\{0\}}(f_i)\|_2},\ \varphi_{f_i}\right)\ (a_j - \mathsf{R}_{F_{\mathcal{A}},K_{\mathcal{A}}}(f_i))\,,\ \forall a_j \in V(F_i).$$

Damit ist durch

$$\mathcal{S}(K_{\mathcal{A}},\vec{T}_{\mathcal{A}}) \;=\; \left\{\ \bigcup_{i=0}^{n} V_i,\ \vec{T}_{\mathcal{A}},\ \Psi_{K_{\mathcal{A}},F_{\mathcal{A}}},\ \mathsf{R}_{F_{\mathcal{A}},K_{\mathcal{A}}},\ \mathsf{S}_{F_{\mathcal{A}},K_{\mathcal{A}}},\ \mathsf{D}_{F_{\mathcal{A}},\mathbb{R}^3\setminus\{0\}}\ \right\}$$

eine Transformationsstruktur[13] auf $K_{\mathcal{A}}$ bzgl. $\vec{T}_{\mathcal{A}}$ gegeben.

Bemerkung: Die Transformationsstruktur auf $K_{\mathcal{A}}$ und damit der lokale Rotationsoperator[14] $\mathcal{R}[\mathcal{S}(K_{\mathcal{A}},\vec{T}_{\mathcal{A}})]$ von $K_{\mathcal{A}}$ hängen ganz entscheidend von der Wahl des Hauptfragmentes F_0 ab. Wählt man ein anderes Hauptfragment, so ändert sich die Transformationsstruktur und damit auch der lokale Rotationsoperator (siehe Kapitel 2.4.2). Mittels der Transformationsstruktur $\mathcal{S}(K_{\mathcal{A}},\vec{T}_{\mathcal{A}})$ kann man nun einen Transformationsoperator[15] $\mathcal{T}[\mathcal{S}(K_{\mathcal{A}},\vec{T}_{\mathcal{A}})]$ auf $K_{\mathcal{A}}$ definieren, der sowohl die globalen wie auch die lokalen Transformationen der Atomkoordinaten und damit der Molekülkonformation beschreibt.

Nun werden die Daten beschrieben, die verwendet werden, um die relevanten physikochemischen Eigenschaften eines Moleküls eine räumlich lokalisierte Dichte zuzuordnen. Dazu wird zuerst definiert, welche relevanten physikochemischen Eigenschaften überhaupt betrachtet werden.

Definition 2.2.1 (Physikochemische Eigenschaften)
Es sei $\mathcal{I}_{PE}$ die Indexmenge der physikochemischen Eigenschaften {Elektronendichte, Ladung, Hydrophobizität, Wasserstoffdonatoren, Wasserstoffakzeptoren}. $\diamond$

Um nun die räumlich lokalisierte Dichte einer physikochemischen Eigenschaft zu modellieren, wird zunächst formal definiert, was eine räumlich lokalisierte Dichte sein soll.

[13]Siehe Definition 2.4.14.
[14]Siehe Definition 2.4.18.
[15]Siehe Definition 2.4.27.

Definition 2.2.2 (Räumlich lokalisierte Dichte)
Es sei $\mathcal{A}$ ein Molekül mit zulässiger Konformation $\mathsf{C}_{\mathcal{A}}$. Die räumlich lokalisierte Dichte einer physikochemischen Eigenschaften ist eine Abbildung, die von der Konformation abhängt und jedem Raumpunkt einen quantitativen Wert zuordnet. Für $i \in \mathcal{I}_{PE}$ ist

$$Q_i[\mathsf{C}_{\mathcal{A}}] \colon \mathbb{R}^3 \to \mathbb{R}$$

eine räumlich lokalisierte Dichte bzgl. der physikochemischen Eigenschaft i und der Konformation $\mathsf{C}_{\mathcal{A}}$. $\diamond$

Der Ansatz in dieser Arbeit besteht darin, die räumlich lokalisierte Dichte mittels Gaussfunktionen[16] bzw. mittels Summen von Gaussfunktionen darzustellen.[17] Zum Ansatz mittels Gaussfunktionen siehe auch die Arbeiten von C.Lemmen [41], G.Klebe u.a [35], A.C.Good u.a [24], M. Hauswald [30], sowie J.A.Grant u.a. [25]

Um die räumlich lokalisierte Dichte eines Moleküls bzgl. einer physikochemischen Eigenschaft zu modellieren, werden zunächst die physikochemischen Eigenschaften der einzelnen Atome entsprechend dem obigen Ansatz räumlich dargestellt. Um nun die physikochemische Eigenschaft $q \in \mathcal{I}_{PE}$ eines Atom a_l räumlich zu modellieren, wird dem entsprechenden Atom die Folge $((h_i^{l,q}, b_i^{l,q}, z_i^{l,q}))_{i=1}^{k_q^l} \subset \mathbb{R} \times \mathbb{R}_{>0} \times \mathbb{R}^3$, $k_q^l \in \mathbb{N}_0$, zugeordnet. Dann ist die räumlich lokalisierte Dichte des Atoms a_l bzgl. der Eigenschaft q durch die Summe

$$Q_{a_l}^q(x) \;:=\; \sum_{i=1}^{k_q^l} \mathsf{G}(h_i^{l,q}, b_i^{l,q}, z_i^{l,q})(x)$$

gegeben. Dabei ist $\mathsf{G}(h_i^{l,q}, b_i^{l,q}, z_i^{l,q})(x)$ eine durch die Höhe $h_i^{l,q}$, die Breite $b_i^{l,q}$ und das Zentrum $z_i^{q,l}$ eindeutig bestimmte Gaussfunktion[18]. Mit k_q^l wird die Anzahl der Gaussfunktionen bezeichnet, die verwendet werden. Dabei bedeutet $k_q^l = 0$, daß das Atom die betreffende Eigenschaft q nicht besitzt. Zum Beispiel a_l ist kein Wasserstoffdonator. Nun setzt man für $q \in \mathcal{I}_{PE}$

$$GC_{a_l}^q \;:=\; \begin{cases} \{z_1^{l,q}, \ldots, z_{k_q^l}^{l,q}\} & \text{, falls } k_q^l > 0 \\ \emptyset & , \textit{sonst}, \end{cases}$$

Bemerkung: Das Gausszentrum wird in der Region des Raumes plaziert, in der man die jeweilige Eigenschaft erwartet (vergleiche hierzu Abbildung 2.2). Als Default-Wert nimmt man die Atomkoordinate a_l. Die Default-Werte für die Höhe bzw. die Breite bzgl. der jeweiligen physikochemischen Eigenschaft sind in der Arbeit von C.Lemmen [41] aufgeführt.

Die räumlich lokalisierte Dichte des Moleküls bzgl. der physikochemische Eigenschaft q ist dann

[16] Zu Definition von Gaussfunktionen siehe Kapitel 2.5.
[17] Der Ansatz mittels Gaussfunktionen ist eine Vorgabe des Programmpakets FlexS.
[18] Definition 2.5.1

durch

$$
\begin{aligned}
Q^q_{\mathcal{A}}(x) \;&:=\; \sum_{l=1}^{n_a} Q^q_{a_l}(x) \;=\; \sum_{l=1}^{n_a} \sum_{i=1}^{k^l_q} \mathsf{G}(h^{l,q}_i, b^{l,q}_i, z^{l,q}_i)(x) \\
&=: \sum_{j=1}^{n^q_g} \mathsf{G}(h^q_j, b^q_j, z^q_j)(x)
\end{aligned}
$$

gegeben, wobei $n^q_g := \sum_{l=1}^{n_a} k^l_q$ die Gesamtzahl der Gaussfunktionen bzgl. der Eigenschaft $q \in \mathcal{I}_{PE}$ ist und h^q_j, b^q_j und z^q_j geeignet indiziert worden sind. Somit wird die physikochemische Eigenschaft $q \in \mathcal{I}_{PE}$ des Moleküls $\mathcal{A}$ mittels der Folge $((h^q_j, b^q_j, z^q_j))_{j=1}^{n^q_g}$ räumlich modelliert. Nun setzt man für $q \in \mathcal{I}_{PE}$

$$
GC^q_{\mathcal{A}} \;:=\; \begin{cases} \{z^q_1, \ldots, z^q_{n^q_g}\} & \text{, falls } n^q_g > 0 \\ \emptyset & , sonst. \end{cases}
$$

Im folgenden sei $GC^q_{\mathcal{A}} \neq \emptyset$:
$GC^q_{\mathcal{A}}$ ist eine geordnete Menge und die Ordnung ist durch folgende Definition $z^q_i \le z^q_j :\Leftrightarrow i \le j$, $i,j = 1, \ldots, n^q_g$ gegeben. Als Koordinaten aus $\mathbb{R}^3$ brauchen die Elemente von $GC^q_{\mathcal{A}}$, $q \in \mathcal{I}_{PE}$, nicht paarweise verschieden zu sein. Die Ordnung garantiert jedoch, daß sie formal paarweise verschieden sind, d.h. $z^q_i \neq z^q_j \Leftrightarrow i \neq j$ für z^q_i, $z^q_j \in GC^q_{\mathcal{A}}$, $q \in \mathcal{I}_{PE}$. Die Abbildung $\Gamma^q_{\mathcal{A}} \colon GC^q_{\mathcal{A}} \to K_{\mathcal{A}}$, $q \in \mathcal{I}_{PE}$ ordnet nun jeder Gausskoordinate die entsprechende Atomkoordinate zu, d.h. $\Gamma^q_{\mathcal{A}}(z^q_k) := a_{l_k}$, falls $z^q_k \in GC^q_{a_{l_k}}$, $k = 1, \ldots, n^q_g$.
Mit Hilfe der Ordnung auf $GC^q_{\mathcal{A}}$, $q \in \mathcal{I}_{PE}$, kann man den Transformationsoperator $\mathcal{T}[\mathcal{S}(K_{\mathcal{A}}, \vec{T}_{\mathcal{A}})]$ von $K_{\mathcal{A}}$ nun auch auf die Gausskordinaten übertragen. Wird der Transformationsoperator $\mathcal{T}[\mathcal{S}(K_{\mathcal{A}}, \vec{T}_{\mathcal{A}})]$ auf die Atomkoordinaten aus $K_{\mathcal{A}}$ angewandt, so ändert sich die Molkülkonformation von $\mathcal{A}$. Da die räumlich lokalisierten Dichten der physikochemischen Eigenschaften von der Konformation abhängen, müssen sich auch die räumlich lokalisierten Dichten entsprechend der Transformation der Atomkoordinaten ändern. Also muß der Transformationsoperator $\mathcal{T}[\mathcal{S}(K_{\mathcal{A}}, \vec{T}_{\mathcal{A}})]$ auch auf die Gausskoodinaten aus $GC^q_{\mathcal{A}}$, $q \in \mathcal{I}_{PE}$ angewendet werden.

Während die praktische Anwendung des Transformationsoperators auf die Gausskoordinaten unproblematisch ist, ist die Wohldefiniertheit der Übertragung nicht ganz unkritisch, da die Gausskordinaten nicht unbedingt paarweise verschieden sein müssen. Die Übertragung läßt sich wie folgt definieren. Dazu sei zunächst $K^*_{\mathcal{A}} := \{\mathsf{R}_{F_{\mathcal{A}},K_{\mathcal{A}}}(f_i), \mathsf{S}_{F_{\mathcal{A}},K_{\mathcal{A}}}(f_i) \mid f_i \in F_{\mathcal{A}} \setminus \{f_0\}\} \subset K_{\mathcal{A}}$. Nun setzt man für $a_i \in K^*_{\mathcal{A}}$ $\Psi_{K^*_{\mathcal{A}},F_{\mathcal{A}}}(a_i) := \Psi_{K_{\mathcal{A}},F_{\mathcal{A}}}(a_i)$. Dann gilt für $i = 0, \ldots, n_{ft}$ $V^*_i := \left(\Psi_{K^*_{\mathcal{A}},F_{\mathcal{A}}}\right)^{-1}(f_i) \neq \emptyset$ Da $(f_0, f_1) \in \vec{E}^{\mathcal{A}}_{FT}$ gilt, folgt $\mathsf{R}_{F_{\mathcal{A}},K_{\mathcal{A}}}(f_1) \in V^*_0$. Für $i = 1, \ldots, n_{ft}$ gilt $\mathsf{S}_{F_{\mathcal{A}},K_{\mathcal{A}}}(f_i) \in V^*_i$. Insbesondere bilden die V^*_i, $i = 0, \ldots, n_{ft}$, eine disjunkte Zerlegung von $K^*_{\mathcal{A}}$, d.h. $K^*_{\mathcal{A}} = \bigcup V^*_i$. Weiter setzt man für $i = 1, \ldots, n_{ft}$ $\mathsf{R}_{F_{\mathcal{A}},K^*_{\mathcal{A}}}(f_i) := \mathsf{R}_{F_{\mathcal{A}},K_{\mathcal{A}}}(f_i)$ und $\mathsf{S}_{F_{\mathcal{A}},K^*_{\mathcal{A}}}(f_i) := \mathsf{S}_{F_{\mathcal{A}},K_{\mathcal{A}}}(f_i)$. Dann ist

$$
\mathcal{S}(K^*_{\mathcal{A}}, \vec{T}_{\mathcal{A}}) \;=\; \left\{ \bigcup_{i=0}^{n} V^*_i,\ \vec{T}_{\mathcal{A}},\ \Psi_{K^*_{\mathcal{A}},F_{\mathcal{A}}},\ \mathsf{R}_{F_{\mathcal{A}},K^*_{\mathcal{A}}},\ \mathsf{S}_{F_{\mathcal{A}},K^*_{\mathcal{A}}},\ \mathsf{D}_{F_{\mathcal{A}},\mathbb{R}^3\setminus\{0\}} \right\} \tag{2.3}
$$

eine Transformationsstruktur auf $K^*_{\mathcal{A}}$ bzgl. $\vec{T}_{\mathcal{A}}$. Diese Transformationsstruktur ist mit der Transformationsstruktur $\mathcal{S}(K_{\mathcal{A}}, \vec{T}_{\mathcal{A}})$ identisch, da nur noch die für die Struktur $\mathcal{S}(K_{\mathcal{A}}, \vec{T}_{\mathcal{A}})$

wichtigen Koordinaten aus $K_{\mathcal{A}}$ betrachtet werden. Also ist insbesondere der zu $\mathcal{S}(K^*_{\mathcal{A}}, \vec{T}_{\mathcal{A}})$ gehörende Transformationsoperator $\mathcal{T}[\mathcal{S}(K^*_{\mathcal{A}}, \vec{T}_{\mathcal{A}})]$ mit $\mathcal{T}[\mathcal{S}(K_{\mathcal{A}}, \vec{T}_{\mathcal{A}})]$ identisch, denn es gilt $\mathcal{T}[\mathcal{S}(K^*_{\mathcal{A}}, \vec{T}_{\mathcal{A}})] = \mathcal{T}[\mathcal{S}(K_{\mathcal{A}}, \vec{T}_{\mathcal{A}})]|_{K^*_{\mathcal{A}}}$. Die Transformationsstruktur $\mathcal{S}(K^*_{\mathcal{A}}, \vec{T}_{\mathcal{A}})$ wird nun verwendet, um den Transformationsoperator $\mathcal{T}[\mathcal{S}(K^*_{\mathcal{A}}, \vec{T}_{\mathcal{A}})]$ auf die entsprechenden Gausskoordinaten anzuwenden. Dazu setzt man $K^q_{\mathcal{A}} := K^*_{\mathcal{A}} \cup GC^q_{\mathcal{A}}$, $q \in \mathcal{I}_{PE}$. Die Elemente von $K^q_{\mathcal{A}}$, $q \in \mathcal{I}_{PE}$, brauchen als Koordinaten des $\mathbb{R}^3$ nicht mehr unbedingt paarweise verschieden zu sein. Jedoch sind die Atomkoordinaten nach Voraussetzung paarweise verschieden und die Gausskoordinaten sind formal paarweise verschieden. Um zu garantieren, daß Atomkoordinaten und Gausskoordinaten stets formal paarweise verschieden sind, definiert man auf $K^q_{\mathcal{A}}$, $q \in \mathcal{I}_{PE}$, die folgende Ordnung $\mathsf{R}_{F_{\mathcal{A}},K_{\mathcal{A}}}(f_1) < \mathsf{S}_{F_{\mathcal{A}},K_{\mathcal{A}}}(f_1) < \mathsf{R}_{F_{\mathcal{A}},K_{\mathcal{A}}}(f_2) < \ldots < \mathsf{S}_{F_{\mathcal{A}},K_{\mathcal{A}}}(f_{n_{ft}}) < z^q_1 < \ldots z^q_{n^q_g}$. Damit gilt formal $a_k \neq z^q_l$ für alle $a_k \in K^*_{\mathcal{A}}$ und für alle $z^q_l \in GC^q_{\mathcal{A}}$ bzgl. der obigen Ordnung.
Die Abbildung $\Psi^q_{K^q_{\mathcal{A}},F_{\mathcal{A}}} : K^q_{\mathcal{A}} \to F_{\mathcal{A}}$, $q \in \mathcal{I}_{PE}$, ist wie folgt definiert: für die Koordinaten aus $GC^q_{\mathcal{A}}$ setzt man $\Psi^q_{K^q_{\mathcal{A}},F_{\mathcal{A}}} := \Psi_{K_{\mathcal{A}},F_{\mathcal{A}}} \circ \Gamma^q_{\mathcal{A}}$ und für die Koordinaten aus $K^*_{\mathcal{A}}$ setzt man $\Psi^q_{K^q_{\mathcal{A}},F_{\mathcal{A}}} := \Psi_{K^*_{\mathcal{A}},F_{\mathcal{A}}}$. Da $K^q_{\mathcal{A}}$ eine geordnete Menge ist, ist die obige Abbildung eindeutig und wohldefiniert. Für $q\mathcal{I}_{PE}$ und $i = 0, \ldots, n_{ft}$ ist das Urbild von $f_i \in F_{\mathcal{A}}$ in $GC^q_{\mathcal{A}}$ definiert durch $GC^q_i := \left(\Psi_{K_{\mathcal{A}},F_{\mathcal{A}}} \circ \Gamma^q_{\mathcal{A}}\right)^{-1}(f_i)$. Dann bilden $V^q_i := V^*_i \cup GC^q_i$, $i = 0, \ldots, n_{ft}$, eine formal disjunkte Zerlegung von $K^q_{\mathcal{A}}$. Ersetzt man nun in der Transformationsstruktur in Gleichung (2.3) V^*_i durch V^q_i und $\Psi_{K^*_{\mathcal{A}},F_{\mathcal{A}}}$ durch $\Psi_{K^q_{\mathcal{A}},F_{\mathcal{A}}}$, so erhält man eine formale Transformationsstruktur $\mathcal{S}(K^q_{\mathcal{A}}, \vec{T}_{\mathcal{A}})$ auf $K^q_{\mathcal{A}}$, $q \in \mathcal{I}_{PE}$, bzgl. $\vec{T}_{\mathcal{A}}$. Formal bedeutet hier, daß die Zerlegung von $K^q_{\mathcal{A}}$, $q \in \mathcal{I}_{PE}$, nur formal disjunkt sein muß. Da aber die zugrundliegende Strukutur $\mathcal{S}(K^*_{\mathcal{A}}, \vec{T}_{\mathcal{A}})$ mathematisch eindeutig und wohldefiniert ist, sind auch diese formalen Strukturen eindeutig und wohldefiniert. Insbesondere sind die zu $\mathcal{S}(K^q_{\mathcal{A}}, \vec{T}_{\mathcal{A}})$, $q \in \mathcal{I}_{PE}$, gehörenden Transformationsoperatoren $\mathcal{T}[\mathcal{S}(K^q_{\mathcal{A}}, \vec{T}_{\mathcal{A}})]$ mit dem Operator $\mathcal{T}[\mathcal{S}(K^*_{\mathcal{A}}, \vec{T}_{\mathcal{A}})]$ bis auf die Verwendung der Abbildung $\Psi^q_{K^q_{\mathcal{A}},F_{\mathcal{A}}}$ anstatt der Abbildung $\Psi_{K^*_{\mathcal{A}},F_{\mathcal{A}}}$ identisch und damit eindeutig und wohldefiniert. Die zu den formalen Transformationsstrukturen gehörenden Operatoren beschreiben die globalen wie lokalen Transformationen der Gausskoordinaten und damit der räumlichen Dichten.

Bemerkung: Bei der praktischen Berechnung nimmt man den durch die Transformationsstruktur $\mathcal{S}(K_{\mathcal{A}}, \vec{T}_{\mathcal{A}})$ eindeutig und wohldefinierten Transformationsoperator $\mathcal{T}[\mathcal{S}(K_{\mathcal{A}}, \vec{T}_{\mathcal{A}})]$ und wendet diesen zusätzlich auf die Gausskoordinaten an, die zuvor auf die Knoten $f_i \in F_{\mathcal{A}}$ entsprechend aufgeteilt worden sind. Denn $\mathcal{S}(K^*_{\mathcal{A}}, \vec{T}_{\mathcal{A}})$ ist nach den obigen Definition die Grundstruktur für die Transformationsstrukturen $\mathcal{S}(K_{\mathcal{A}}, \vec{T}_{\mathcal{A}})$ und $\mathcal{S}(K^q_{\mathcal{A}}, \vec{T}_{\mathcal{A}})$, $q \in \mathcal{I}_{PE}$ und bestimmt nach Definition 2.4.27 die Transformationsopertoren eindeutig. Daher gilt

$$\mathcal{T}[\mathcal{S}(K^*_{\mathcal{A}}, \vec{T}_{\mathcal{A}})](x) = \mathcal{T}[\mathcal{S}(K_{\mathcal{A}}, \vec{T}_{\mathcal{A}})](x) = \mathcal{T}[\mathcal{S}(K^q_{\mathcal{A}}, \vec{T}_{\mathcal{A}})](x) \quad \forall x \in \mathbb{R}^3,\ q \in \mathcal{I}_{PE}.$$

Ein weiterer wichtiger Punkt ist die Bewertung der einzelnen flexiblen Torsionswinkel, genauer, der einzelnen aktuellen Torsionswinkel. Dazu setzt man für $e \in B^{\mathcal{A}}_{FT}$ zunächst $\Theta_e(\mathsf{C}_{\mathcal{A}}) := \alpha_e(\mathsf{C}_{\mathcal{A}}) \in [0;\ 2\pi)$, wobei $\alpha_e(\mathsf{C}_{\mathcal{A}})$ der Winkel ist, der bzgl. der Startkonformation $\mathsf{C}_{\mathcal{A}}$ berechnet und der Einfach-Bindung e zugeordnet worden ist. Nun gibt es zu jedem $f_i \in F_{\mathcal{A}} \setminus \{f_0\}$ genau eine Kante $e_i \in B^{\mathcal{A}}_{FT}$ mit $e_i = \mathcal{I}^{\mathcal{A}}_{F,B}(f_i)$. Dann setzt man $\Theta^{tor}(\mathsf{C}_{\mathcal{A}}) := (\Theta_1(\mathsf{C}_{\mathcal{A}}), \ldots, \Theta_{n_{ft}}(\mathsf{C}_{\mathcal{A}}))$, $\Theta_i(\mathsf{C}_{\mathcal{A}}) := \Theta_{e_i}(\mathsf{C}_{\mathcal{A}})$. Dieser Vektor enthält die Torsionswinkel der Startkonformation. Es sei $\varphi^{akt}_i \in \mathbb{R}$ der aktuelle Rotationswinkel bzgl. der Drehachse $\mathsf{D}_{F_{\mathcal{A}},\mathbb{R}^3\setminus\{0\}}(f_i)$. Dann werden die aktuellen Torsionswinkel wie folgt bestimmt:

$$\theta_i^{akt} \; = \; \Theta_i(\mathsf{C}_{\mathcal{A}}) \; + \; \varphi_i^{akt}, \; i = 1, \ldots, n_{ft}. \tag{2.4}$$

Die aktuellen Torsionswinkel sollen nun bewertet werden. Dazu wird in Kapitel 2.7 eine Bewertungsfunktion definiert. Bei der Optimierung sollen diejenigen Torsionswinkel bevorzugt werden, die zu geringeren intermolekularen Abstoßungskräften führen. In der Abbildung 2.6 sieht man den Torsionswinkel α bzgl. den Atomen C-C-C-C. Die den Winkel α bewertenden Abstoßungskräfte werden zwischen den Atomen C (C.ar) und C (C.any) und weiteren Atomen aus der Umgebung gemessen. Wie aus der Abbildung zu ersehen ist, ist ein Winkel von 90 bzw. 270 Grad günstiger als ein Winkel von 0 oder 180 Grad (vergleiche die Bewertungstabelle in Abbildung 2.7). Da die Potentialkurve eines flexiblen Torsionswinkel, die die intermolekularen Abstoßungskräfte darstellt, meist unbekannt ist, kann man folgendermaßen vorgehen (siehe hierzu auch [7], aber auch [48]): Dazu nimmt man ein Molekülfragment, das die zugrundelie-

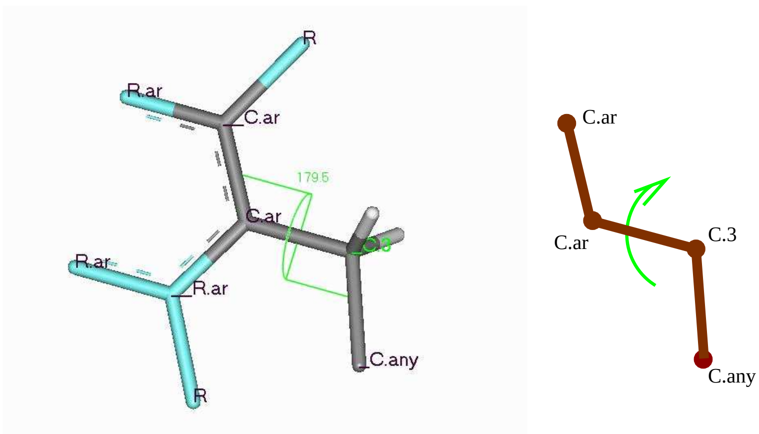


Abbildung 2.6: Eine Unterkonformation, die die Atomkonfiguration C-C-C-C enthält

gende Atomkonfiguration A-B-C-D enthält. Dann erstellt man eine Häufigkeitstabelle für den Torsionswinkel $\measuredangle(A-B, C-D)$, indem man eine möglichst große Anzahl von Molekülekonformationen betrachtet, die das entsprechende Fragment mit der Atomkonfiguration A-B-C-D enthält. Der gemessene Torsionswinkel wird in die Häufigkeitstabelle eingetragen. Meßdaten über die Molekülkonformation liegen häufig für Kristallstrukturen vor.[19] Dabei sollte man jedoch beachten, daß die Konformation eines flexiblen Moleküls in seiner Kristallstruktur oft nicht identisch oder auch nur ähnlich zur Geometrie des Moleküls in der Bindetasche eines Proteins ist. Nach [7] können flexible Moleküle, je nach Umgebung, deutlich unterschiedliche Konformationen einnehmen. Eine von diesen wird die rezeptorgebundene Konformation sein. Diese ist nicht notwendig die energetisch günstigste Konformation, sie wird aber in einem energetisch günstigen Bereich liegen. Die Hoffnung ist also, daß sich die Torsionswinkel energetisch günstiger Konformationen in den Häufigkeitstabellen entsprechend wiederfinden. Deshalb sollten so viele

[19] Die Kristallstrukturen einer Vielzahl organischer Moleküle sind bekannt und in einer Datenbank hinterlegt. Die größte Datenbank mit Strukturinformationen ist die *Cambridge Structure Database* (CSD) [2].

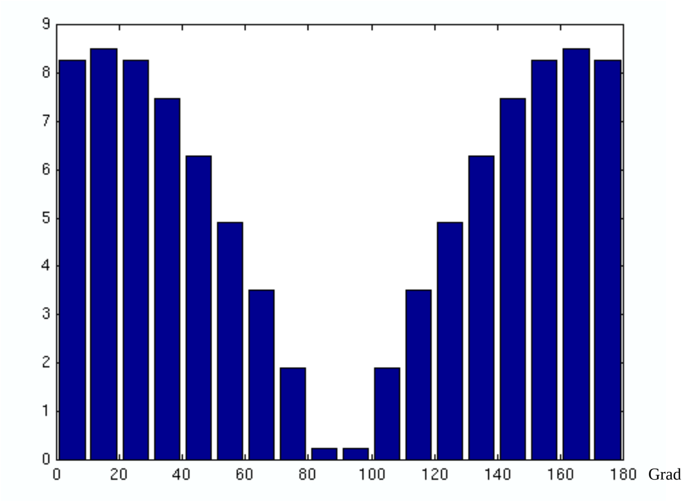


Abbildung 2.7: Bewertungsdiagramm für den Torsionswinkel α bzgl. der Konfiguration C-C-C-C

Molekülkonformationen wie möglich betrachtet werden, damit das Spektrum für den Torsionswinkel möglichst groß ist. Nun erstellt man mit Hilfe der Häufigkeitstabelle eine heuristische Bewertungstabelle. Dabei wird natürlich ein Winkel, der sehr häufig auftritt, günstiger bewertet als ein Winkel, der kaum angenommen wird. In Abbildung 2.7 ist ein Bewertungsdiagramm für den Torsionswinkel in der Atomkonfiguration C-C-C-C dargestellt. Die Bewertungstabelle hat bei 90 Grad eine Symmetrieachse und hat die Periode 180 Grad. Der günstigste Torsionswinkel liegt ungefähr bei 90 Grad. Diese heuristischen Bewertungstabellen bilden die Grundlage für die Bewertungsfunktion der Torsionswinkel (siehe Kapitel 2.7). Jedoch muß nicht für alle flexiblen Torsionswinkel solch eine Bewertungstabelle vorliegen. Die hier verwendeten Bewertungstabelllen sind mit Hilfe der CSD-Datenbank erstellt worden.[20]

Formal sei für alle $e_i \in B_{FT}^{\mathcal{A}}$, $e_i = \mathcal{I}_{F,B}^{\mathcal{A}}(f_i)$, $\Omega_{e_i} := \{x_k\}_{k=0}^{n(e_i)}$ eine äquidistante Zerlegung des Intervalls $[x_0;\ p(e_i) + x_0] \subset [0;\ 2\pi]$, mit $0 \leq x_0 < x_1 < \ldots < x_{n(e_i)} = p(e_i) + x_0$, wobei $p(e_i) = j(e_i)\ \frac{\pi}{18}$ mit $j(e_i) \in \{1, 2, \ldots, 36\}$ die Periode und $x_k := x_0 + k \cdot h$, für $k = 0, \ldots, n(e_i)$ ist.[21] Falls für die Bindung e_i eine Bewertungstabelle gegeben ist, so werden die entsprechenden Energiewerte f_k^i, $k = 0, \ldots, n(e_i)$, gesetzt. Da $x_{n(e_i)} = p(e_i) + x_0$ ist und die Energiewerte periodisch sein müssen, dilt $f_0^i = f_{n(e_i)}^i$. Ist jedoch keine Bewertungstabelle gegeben, so wird $f_k^i := 0$ für $k = 0, \ldots, n(e_i)$ gesetzt. Der *Datensatz der Knotenzerlegung und der Energiewerte* für die Bindung e_i ist durch $\triangle_{e_i} := \{\Omega_{e_i},\ \{f_k^i\}_{k=0}^{n(e_i)}\}$ gegeben. Durch geeignete Interpolation der Energiewerte soll eine analytische Funktion bestimmt werden, die den entsprechenden aktuelllen Torsionswinkel bewertet. Die Summe dieser analytischen Funktionen bewertet dann alle aktuellen Torsionswinkel des Moleküls $\mathcal{A}$, die betrachtet werden.

Zum Abschluß dieses Kapitel werden die Daten, die hier beschrieben worden sind, in einer Datenstruktur zusammengefaßt. Diese Datenstruktur enthält somit alle Informationen, die notwendig sind, um die Energiefunktionen für die Modelle 2.1 und 2.2 aus Kapitel 2.1 zu beschreiben.

[20] *Cambridge Structure Database* [2].

[21] In dieser Arbeit sind die Werte $x_0 := \frac{\pi}{36}$ ($\equiv 5°$) und $h := \frac{\pi}{18}$ ($\equiv 10°$) verwendet worden.

Definition 2.2.3 (Datenstruktur eines Moleküls)
Es sei $\mathcal{A}$ ein Molekül mit n_a Atomen und $\mathsf{C}_\mathcal{A} \subset \mathbb{R}^{3n_a}$ sei eine zulässige Konformation. Dann ist

$$\mathcal{DS}_\mathcal{A}(\mathsf{C}_\mathcal{A}) := \Big\{K_\mathcal{A} := K(\mathsf{C}_\mathcal{A}),\ \Sigma_\mathcal{A},\ B_\mathcal{A},\ B^\mathcal{A}_{FT},\ F_0,\ F_\mathcal{A},\ \Psi_{K_\mathcal{A},F_\mathcal{A}},\ \vec{E}^\mathcal{A}_{FT},\ \mathsf{R}_{F_\mathcal{A},K_\mathcal{A}},\ \mathsf{S}_{F_\mathcal{A},K_\mathcal{A}},$$
$$\mathsf{D}_{F_\mathcal{A},\mathbb{R}^3\setminus\{0\}},\ \Big\{\{(h^q_i, b^q_i, z^q_i)\}_{i=1}^{n^q_g}\Big\}_{q\in\mathcal{I}_{PE}},\ \{\Gamma^q_\mathcal{A}\}_{q\in\mathcal{I}_{PE}},\ \Theta^{tor}(\mathsf{C}_\mathcal{A}),\ \{\triangle_{e_i}\}_{e_i\in B^\mathcal{A}_{FT}}\Big\}$$

eine Datenstruktur für das Molekül $\mathcal{A}$ bzgl. der Molekülkonformation $\mathsf{C}_\mathcal{A}$. Aus der Datenstruktur $\mathcal{DS}_\mathcal{A}(\mathsf{C}_\mathcal{A})$ kann man alle Informationen ableiten, die notwendig sind, um das Molekül, den Fragmentebaum, die Transformationsstruktur, den Transformationsoperator, die räumlich lokalisierten Dichten bzgl. der physikochemischen Eigenschaften, die Torsionswinkel und um die Datensätze der Knotenzerlegung und der Energiewerte für die Bindungen e_i zu beschreiben. Mit diesen Informationen kann man die Energiefunktionen aufstellen. ◇

2.3 Ähnlichkeitsindex

In diesem Abschnitt werden die in Kapitel 2.1 eingeführten Indizes als mathematische Abbildungen definiert. Zunächst werden der Carbo- und Hodgkin-Index für Funktionen aus dem Hilbertraum $\mathsf{L}^2(\mathbb{R}^3)$ eingeführt. Mit Hilfe der Eigenschaften dieser Abbildungen soll dann entschieden werden, welche Abbildung sich für die Definition des Alignmentindex eignet. Die Normen und die Skalarprodukte in disem Abschnitt beziehen sich auf den Hilbertraum $\mathsf{L}^2(\mathbb{R}^3)$.

Definition 2.3.1 (Carbo- und Hodgkin-Index)
Es seien $f, g \in \mathsf{L}^2(\mathbb{R}^3)$.

a) Die Abbildung $\mathsf{H}\colon \mathsf{L}^2(\mathbb{R}^3) \times \mathsf{L}^2(\mathbb{R}^3) \to \mathbb{R}$ mit

$$\mathsf{H}(f,g) := \frac{2\int_{\mathbb{R}^3} f(x)g(x)\,\mathsf{d}x}{\int_{\mathbb{R}^3} f(x)^2\,\mathsf{d}x + \int_{\mathbb{R}^3} g(x)^2\,\mathsf{d}x} = \frac{2 < f,\ g >}{< f,\ f > + < g,\ g >},$$

falls

$$< f,\ f > + < g,\ g > = \|f\|^2 + \|g\|^2 > 0 \tag{2.5}$$

gilt, heißt Hodgkin-Index der Funktionen f, g;

b) Die Abbildung $\mathsf{C}\colon \mathsf{L}^2(\mathbb{R}^3) \times \mathsf{L}^2(\mathbb{R}^3) \to \mathbb{R}$ mit

$$\mathsf{C}(f,g) := \frac{\int_{\mathbb{R}^3} f(x)g(x)\,\mathsf{d}x}{\left(\sqrt{\int_{\mathbb{R}^3} f(x)^2\,\mathsf{d}x}\right)\left(\sqrt{\int_{\mathbb{R}^3} g(x)^2\,\mathsf{d}x}\right)} = \frac{< f,\ g >}{(\sqrt{< f,\ f >})\ (\sqrt{< g,\ g >})},$$

falls

$$\sqrt{< f,\ f >}\sqrt{< g,\ g >} = \|f\|\,\|g\| \neq 0 \tag{2.6}$$

gilt, heißt Carbo-Index der Funktionen f, g. ◇

Die Eigenschaften der obigen Abbildungen sind im folgenden Lemma zusammengefaßt.

Lemma 2.3.2
Es seien $f, g \in \mathsf{L}^2(\mathbb{R}^3)$, $\lambda \in \mathbb{R}$. Gilt (2.5), so gilt

a) $|\mathsf{H}(f,\ g)| \ \leq\ 1.$

b) $\mathsf{H}(\lambda f,\ f) \ =\ \frac{2\ \lambda}{|\lambda|^2\ +\ 1}.$

c) H *ist symmetrisch, d.h.* $\mathsf{H}(f,\ g) = \mathsf{H}(g,\ f)$.

Gilt (2.6), so gilt

d) $|\mathsf{C}(f,\ g)| \ \leq\ 1.$

e) $\mathsf{C}(\lambda f,\ f) \ =\ sig(\lambda).$

f) C *ist symmetrisch, d.h.* $\mathsf{C}(f,\ g) = \mathsf{C}(g,\ f)$.

Beweis:

a)

$$|\mathsf{H}(f,g)| \ =\ \left|\frac{2\ < f,\ g >}{\|f\|^2\ +\ \|g\|^2}\right| \ \overset{C.4}{\leq}\ \frac{2\ \|f\|\ \cdot\ \|g\|}{\|f\|^2\ +\ \|g\|^2} \ \overset{C.3.}{\leq}\ 1.$$

Das erste Ungleichheitszeichen beruht auf der Cauchy-Schwarz-Ungleichung (Satz C.4 in Anhang C) und das zweite Ungleichheitszeichen ist eine Folgerung aus der Ungleichung zwischen geometrischen und arithmetischen Mittel (Satz C.3).

d) Anwendung der Cauchy-Schwarze-Ungleichung.

b),c),e),f) Klar. ■

Bemerkung: Für $f \in \mathsf{L}^2(\mathbb{R}^3)$, $\|f\| \neq 0$ gilt $\mathsf{H}(f, 0) = 0$, während der Carbo-Index nicht definiert ist. Ist die räumlich lokalisierte Dichte einer physikochemische Eigenschaft die Nullfunktion, so kann ein Alignmentindex, der auf dem Hodgkin-Index beruht, dies entsprechend berücksichtigen. Außerdem scheint der Hodgkin-Index wegen Lemma 2.3.2 b) besser geeignet zu sein, den qualitativen Unterschied zweier Funktionen aufzuzeigen.

Daher wird nun der Hodgkin-Index verwendet, um einen Alignmentindex zu definieren.

Definition 2.3.3 (Alignmentindex)
Es $\mathcal{I}$ eine abzählbare Indexmenge. Weiter seien für alle $i \in \mathcal{I}$ Funktionen $f_i, g_i \in \mathsf{L}^2(\mathbb{R}^3)$, mit der Eigenschaft $\|f_i\|^2 + \|g_i\|^2 > 0$ und Zahlen $\lambda_i \in \mathbb{R}_+$ mit der Eigenschaft $\sum_{i\in\mathcal{I}} \lambda_i \ =\ 1$ gegeben. Dann heißt die Abbildung

$$\mathsf{A}^{\mathsf{H}}_{\mathcal{I},(\lambda_i)_{i\in\mathcal{I}}} : (\mathsf{L}^2(\mathbb{R}^3))^{|\mathcal{I}|} \times (\mathsf{L}^2(\mathbb{R}^3))^{|\mathcal{I}|} \to \mathbb{R}$$

mit

$$\mathsf{A}^{\mathsf{H}}_{\mathcal{I},(\lambda_i)_{i\in\mathcal{I}}}((f_i)_{i\in\mathcal{I}}, (g_i)_{i\in\mathcal{I}}) \ :=\ \sum_{i\in\mathcal{I}} \lambda_i\ \mathsf{H}(f_i, g_i)$$

Alignmentindex bezüglich des Hodgkin-Index zum Paar $(\mathcal{I}, (\lambda_i)_{i\in\mathcal{I}})$ mit zugehörigen Funktionen $f_i, g_i, i \in \mathcal{I}$.

$\diamond$

Lemma 2.3.4

Es sei $\mathcal{I}$ eine abzählbare Indexmenge und $(f_i)_{i\in\mathcal{I}}, (g_i)_{i\in\mathcal{I}} \subset \mathsf{L}^2(\mathbb{R}^3), (\lambda_i)_{i\in\mathcal{I}} \subset \mathbb{R}_+$ wie in Definition 2.3.3. Dann ist der Alignmentindex $\mathsf{A}^{\mathsf{H}}_{\mathcal{I},\,(\lambda_i)_{i\in\mathcal{I}}}$ wohldefiniert, symmetrisch und es gilt

$$\left| \mathsf{A}^{\mathsf{H}}_{\mathcal{I},\,(\lambda_i)_{i\in\mathcal{I}}}\left((f_i)_{i\in\mathcal{I}},\ (g_i)_{i\in\mathcal{I}}\right) \right| \leq 1.$$

Beweis:

Folgt direkt aus den Voraussetzungen und Lemma 2.3.2 a) und c). ■

2.4 Affine Transformation eines Moleküls

Da die Atomkoordinaten eines gegebenen Moleküls nicht direkt optimiert werden, sondern indirekt durch die Optimierung der globalen wie internen Transformation des Moleküls, müssen diese Transformationen beschrieben werden. Dabei wird das Molekül durch seine Atomkoordinaten geometrisch dargestellt. Dazu sei $K \subset \mathbb{R}^3$ die Menge der Atomkoordinaten. In dem folgenden Abschnitt wird die Translation sowie die globale Rotation von $K \subset \mathbb{R}^3$ beschrieben. In einem weiteren Abschnitt werden dann die internen Rotationen bzgl. einer Baumstruktur betrachtet. Im letzten Abschnitt wird dann der Transformationssoperator eines Moleküls zusammengesetzt.

Bevor die Transformationen mathematisch definiert werden, werden zwei Definitionen gegeben, die in den folgenden Unterkapiteln immer wieder verwendet werden. Die erste Definition beschreibt einen kanonischen Isomorphismus zwischen den Matrixraum $\mathsf{M}(3\times 3;\mathbb{R})$ und dem Vektorraum $\mathbb{R}^9$. Dieser Isomorphismus wird dazu verwendet, um das Differential einer parameterabhängigen Matrix zu beschreiben. Das Differential einer parameterabhängigen Matrix wird in der zweiten Definition mathematisch gegeben.

Definition 2.4.1

Es seien $\mathcal{I}_{\mathsf{M,V}}\colon \mathsf{M}(3\times 3;\mathbb{R}) \to \mathbb{R}^9$ und $\mathcal{I}_{\mathsf{V,M}}\colon \mathbb{R}^9 \to \mathsf{M}(3\times 3;\mathbb{R})$ lineare Abbildungen mit

$$\mathcal{I}_{\mathsf{M,V}}\left(\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}\right) := \begin{pmatrix} a_{11} \\ a_{12} \\ \vdots \\ a_{33} \end{pmatrix}, \quad \mathcal{I}_{\mathsf{V,M}}\left(\begin{pmatrix} a_1 \\ \vdots \\ a_9 \end{pmatrix}\right) := \begin{pmatrix} a_1 & a_2 & a_3 \\ a_4 & a_5 & a_6 \\ a_7 & a_8 & a_9 \end{pmatrix}.$$

$\mathcal{I}_{\mathsf{M,V}}$ und $\mathcal{I}_{\mathsf{V,M}}$ sind bijektiv und es gilt $(\mathcal{I}_{\mathsf{M,V}})^{-1} = \mathcal{I}_{\mathsf{V,M}}$. $\mathcal{I}_{\mathsf{M,V}}$ ist ein kanonischer Isomorphismus von $\mathsf{M}(3\times 3;\mathbb{R})$ nach $\mathbb{R}^9$, d.h. $\mathsf{M}(3\times 3;\mathbb{R})$ und $\mathbb{R}^9$ werden miteinander identifiziert, d.h. $(\mathsf{M}(3\times 3;\mathbb{R}) \simeq \mathbb{R}^9)$.

■

Definition 2.4.2

Es sei $M\colon \mathbb{R}^n \to \mathsf{M}(3\times 3;\mathbb{R}),\ x \to M(x),\ n \geq 1.$

1. *$M(x)$ heißt in $x \in \mathbb{R}^n$ total differenzierbar, wenn der Abbildung $\mathcal{I}_{\mathsf{M},\mathsf{V}}(M(x))\colon \mathbb{R}^n \to \mathbb{R}^9$ in $x \in \mathbb{R}^n$ total differenzierbar ist. Man setzt dann*

$$\mathsf{D}M(x) := \mathsf{D}\mathcal{I}_{\mathsf{M},\mathsf{V}}(M(x)) = \begin{pmatrix} \frac{\partial}{\partial x_1} m_{11}(x) & \dots & \frac{\partial}{\partial x_n} m_{11}(x) \\ \frac{\partial}{\partial x_1} m_{12}(x) & \dots & \frac{\partial}{\partial x_n} m_{12}(x) \\ \vdots & \vdots & \vdots \\ \frac{\partial}{\partial x_1} m_{33}(x) & \dots & \frac{\partial}{\partial x_n} m_{33}(x) \end{pmatrix} \in \mathsf{M}(9 \times n; \mathbb{R}).$$

2. *$M(x)$ heißt eine $\mathsf{C}^1(\mathbb{R}^n)$-Funktion, wenn die Abbildung $\mathcal{I}_{\mathsf{M},\mathsf{V}}(M(x))$ eine $\mathsf{C}^1(\mathbb{R}^n)$-Funktion ist.*

Falls $n = 1$ gilt, so ist $M(x)$ eine Kurve und $\mathsf{D}M(x)$ der Tangentialvektor der Kurve in $x \in \mathbb{R}$. $\diamond$

Bemerkung: Die wichtigsten Definitionen und Sätze zur partiellen und totalen Differenzierbarkeit sind im Anhang C (C.12 bis C.19) zusammengefaßt.

2.4.1 Translation und globale Rotation

Gegeben sei $K \subset \mathbb{R}^3$. Es werden Operatoren für die Translation und die globale Rotation definiert, die auf alle Elemente von K angewendet werden.

Der Translationsoperator τ wird wie folgt definiert

$$\begin{aligned} \tau : \mathbb{R}^3 \times \mathbb{R}^3 &\to \mathbb{R}^3 \\ \tau(t, x) &:= x + t, \qquad \forall t, x \in \mathbb{R}^3. \end{aligned}$$

Bemerkung: Durch $\tau|_K(t, \cdot)$ wird K um den Vektor $t \in \mathbb{R}^3$ im Raum verschoben.

Die Rotation der Menge K kann wie folgt beschrieben werden

$$S\ (x - c) + c, \qquad \forall\ x \in K.$$

Dabei ist $S \in \mathsf{SO}_3$ die Rotationsmatrix und $c \in \mathbb{R}^3$ das Rotationszentrum. Ziel ist es nun, die Rotationsmatrix S zu parametrisieren. Die Parametrisierung mittels *Euler Winkel* (siehe z.B. [22]) verwendet die folgende parameterabhängige Rotationsmatrix

$$A(\phi,\ \theta,\ \psi) = \begin{pmatrix} \cos\psi & \sin\psi & 0 \\ -\sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} \cos\psi\ \cos\phi - \cos\theta\ \sin\phi\ \sin\psi & \cos\psi\ \sin\phi + \cos\theta\ \cos\phi\ \sin\psi & \sin\psi\ \sin\theta \\ -\sin\psi\ \cos\phi - \cos\theta\ \sin\phi\ \cos\psi & -\sin\psi\ \sin\phi + \cos\theta\ \cos\phi\ \cos\psi & \cos\psi\ \sin\theta \\ \sin\theta\ \sin\phi & -\sin\theta\ \cos\phi & \cos\theta \end{pmatrix}.$$

Dabei sind $(\phi,\ \theta,\ \psi)^{\mathsf{T}} \in ([0;\ 2\pi))^3$ die Euler Winkel. Der Vorteil der Euler Winkel liegt darin, daß sie unabhängig voneinander sind. Die Parametrisierung mittels Euler Winkel kann jedoch zu erheblichen Effizienzverlusten führen (siehe hierzu die Arbeiten von A.O. Griewank u.a [27], C. Hiller [32] und E. Lattman [39]). In der Arbeit von A.O. Griewank u.a. [27] wird gezeigt, daß die Optimierung mittels Euler Winkel zu künstlichen Optima führen kann. Insbesondere dann, wenn θ in einer Umgebung von Null liegt. Weiter wird dargelegt, daß alle Parametrisierungen, die nur drei Variablen verwenden, mit ähnlichen Problemen behaftet sind und daß jede numerisch akzeptable Parametrisierung mit weniger als fünf Variablen äquivalent zur Parametrisierung mittels *Quaternionen* ist. Aus diesem Grund werden die Quaternionen verwendet, um die obige Rotationsmatrix S zu parametrisieren. Quaternionen sind Cayley-Kleinsche Parameter der Drehung. Die Cayley-Klein-Parametrisierung der SO_3 wird nun hergeleitet, denn einige Schritte der Herleitung werden später in Kapitel 2.4.2 verwendet, um die Hintereinanderschaltung lokaler Rotationen zu beschreiben. Die Grundlage für die Herleitung bildet die Darstellung bei B.Schmidt [51] (man vergleiche aber auch z.B. Allen/Tildesley [3] oder [27]).
Es sei $\mathsf{A} \in \mathsf{SO}_3$, $a \in \mathbb{R}^3$ ein Einheitsvektor in Richtung der zu A gehörenden orientierten Drehachse d, d.h. $a = \frac{d}{\|d\|_2}$, und $\varphi \in \mathbb{R}$ der zu A gehörende Drehwinkel. Für den Drehwinkel φ gilt: $\varphi = \arccos(\frac{1}{2}\ (\mathrm{tr}\ \mathsf{A} - 1))$ Ist A nicht symmetrisch, so erhält man eine für die "Rechte-Hand-Regel" orientierte Drehachse $\mathsf{d} = (d_1, d_2, d_3)^T$ aus

$$\mathsf{A} - \mathsf{A}^T =: \begin{pmatrix} 0 & -d_3 & d_2 \\ d_3 & 0 & -d_1 \\ -d_2 & d_1 & 0 \end{pmatrix}.$$

Nach Satz C.9 in Anhang C und den Rechenregeln des Kreuzproduktes C.10 erhält man die folgende Parametrisierung für die Rotationsmatrix A

$$\begin{aligned}
\mathsf{A}y &\overset{C.9}{=} y + (\sin\varphi)\ a \times y + (1-\cos\varphi)\ a \times (a \times y) \\
&\overset{C.10\ 3.)}{=} y + (\sin\varphi)\ a \times y + (1-\cos\varphi)\ (a < a,\ y >_2 - y < a,\ a >_2) \\
&= y + (\sin\varphi)\ a \times y + (1-\cos\varphi)\ (a < a,\ y >_2 - y) \\
&= (\cos\varphi)\ y + (\sin\varphi)\ a \times y + (1-\cos\varphi) < a, y >_2\ a \\
&=: \mathsf{B}(a,\ \varphi)\ y, \quad \forall\, y \in \mathbb{R}^3 \qquad (\Rightarrow\ \mathsf{A} = \mathsf{B}(a,\ \varphi))
\end{aligned}$$

mit $\mathsf{B}(a,\ \varphi) :=$

$$\begin{pmatrix} (1-\cos\varphi)a_1^2 + \cos\varphi & (1-\cos\varphi)a_1a_2 - a_3\sin\varphi & (1-\cos\varphi)a_1a_3 + a_2\sin\varphi \\ (1-\cos\varphi)a_1a_2 + a_3\sin\varphi & (1-\cos\varphi)a_2^2 + \cos\varphi & (1-\cos\varphi)a_2a_3 - a_1\sin\varphi \\ (1-\cos\varphi)a_1a_3 - a_2\sin\varphi & (1-\cos\varphi)a_2a_3 + a_1\sin\varphi & (1-\cos\varphi)a_3^2 + \cos\varphi \end{pmatrix}. \tag{2.7}$$

Wobei $a = (a_1, a_2, a_3)^{\mathsf{T}}$ gilt.

Satz 2.4.3
Die Abbildung $\mathsf{B}\colon \mathbb{R}^3 \mathrm{x} \mathbb{R} \to \mathsf{M}(3\times 3; \mathbb{R})$ *mit* $(a,\ \varphi) \to \mathsf{B}(a,\ \varphi)$, *wobei* $\mathsf{B}(a,\ \varphi)$ *wie in Gleichung (2.7) definiert ist, ist wohldefiniert. Und es gilt*

$$a \in \mathsf{S}^2 \Rightarrow \mathsf{B}(a,\ \varphi) \in \mathsf{SO}_3, \quad \forall\, \varphi \in \mathbb{R}.$$

Beweis:
Es gilt

$$\mathsf{B}\,(a,\ \varphi) * \mathsf{B}\,(a,\ \varphi)^{\mathsf{T}} \ =$$

$$\begin{pmatrix} \left((a_1^2-1)\cos^2\varphi - 2a_1^2\cos\varphi\right)\left(\|a\|_2^2-1\right) + a_1^2\|a\|_2^2 + a_2^2 + a_3^2 \\ a_1a_2\,(\cos\varphi-1)^2\left(\|a\|_2^2-1\right) \\ a_1a_3\,(\cos\varphi-1)^2\left(\|a\|_2^2-1\right) \end{pmatrix.$$

$$\begin{matrix} a_1a_2\,(\cos\varphi-1)^2\left(\|a\|_2^2-1\right) \\ \left((a_2^2-1)\cos^2\varphi - 2a_2^2\cos\varphi\right)\left(\|a\|_2^2-1\right) + a_2^2\|a\|_2^2 + a_1^2 + a_3^2 \\ a_2a_3\,(\cos\varphi-1)^2\left(\|a\|_2^2-1\right) \end{matrix}$$

$$\begin{matrix} a_1a_3\,(\cos\varphi-1)^2\left(\|a\|_2^2-1\right) \\ a_2a_3\,(\cos\varphi-1)^2\left(\|a\|_2^2-1\right) \\ \left((a_3^2-1)\cos^2\varphi - 2a_3^2\cos\varphi\right)\left(\|a\|^2-1\right) + a_3^2\|a\|_2^2 + a_1^2 + a_2^2 \end{matrix}\Bigg)$$

und

$$\begin{aligned} \det\mathsf{B}\,(a,\ \varphi) \ &= \|a\|_2^2 + \left(\|a\|_2^2-1\right) \\ &\quad \left[\ (a_1^2-1)\cos^3\varphi - \|a\|_2^2\cos^2\varphi\ - \left((a_2^2+a_3^2)\sin^2\varphi - a_1^2\right)\ \cos\varphi\ \right]. \end{aligned}$$

Aus $\|a\|_2 = 1$ folgt $\Big(\mathsf{B}\,(a,\ \varphi) * \mathsf{B}\,(a,\ \varphi)^{\mathsf{T}} \ = \mathsf{E}_3\Big) \wedge (\det\mathsf{B}\,(a,\ \varphi) = 1)$ und damit die Behauptung.
■

Korollar 2.4.4
Es sei $\mathsf{B}\,(a,\ \varphi)$ *wie in Gleichung (2.7) definiert, dann gilt* $\mathsf{B}\,(a,\ 0) \ = \ \mathsf{E}_3\ \forall a \in \mathbb{R}^3$.
■

Mittels der normierten Drehachse und des Drehwinkels kann man eine Rotationsmatrix parametrisieren. Setzt man nun

$$\omega_1 \ := \ a_1\,\frac{\sqrt{1-\cos\varphi}}{\sqrt{2}}, \quad \omega_2 \ := \ a_2\,\frac{\sqrt{1-\cos\varphi}}{\sqrt{2}}, \qquad \omega_3 \ := \ a_3\,\frac{\sqrt{1-\cos\varphi}}{\sqrt{2}},$$

und

$$\omega_4 \ := \ \frac{\sqrt{1+\cos\varphi}}{\sqrt{2}},$$

so ist $\omega := (\omega_1, \omega_2, \omega_3, \omega_4)^{\mathsf{T}}$ ein Einheitsvektor des $\mathbb{R}^4$, wenn $a = (a_1, a_2, a_3)^{\mathsf{T}}$ ein Einheitsvektor des $\mathbb{R}^3$ ist. Man erhält nun für die Rotationsmatrix A die Parametrisierung $\mathsf{A} \ = \ \mathsf{B}\,(a,\ \varphi) \ = \ \hat{\mathsf{B}}(\omega)$ mit

$$\hat{\mathsf{B}}(\omega)\ :\ = \begin{pmatrix} \omega_1^2-\omega_2^2-\omega_3^2+\omega_4^2 & 2(\omega_1\omega_2-\omega_3\omega_4) & 2(\omega_1\omega_3+\omega_2\omega_4) \\ 2(\omega_1\omega_2+\omega_3\omega_4) & -\omega_1^2+\omega_2^2-\omega_3^2+\omega_4^2 & 2(\omega_2\omega_3-\omega_1\omega_4) \\ 2(\omega_1\omega_3-\omega_2\omega_4) & 2(\omega_2\omega_3+\omega_1\omega_4) & -\omega_1^2-\omega_2^2+\omega_3^2+\omega_4^2 \end{pmatrix}. \tag{2.8}$$

Satz 2.4.5
Für die Matrix in (2.8) gilt:

$$\omega \in \mathsf{S}^3 \Leftrightarrow \hat{\mathsf{B}}(\omega) \in \mathsf{SO}_3.$$

Beweis:
Denn es gilt $\hat{\mathsf{B}}(\omega) * \hat{\mathsf{B}}(\omega)^{\mathsf{T}} = \|\omega\|_2^4\, \mathsf{E}_3$ und $\det \hat{\mathsf{B}}(\omega) = \|\omega\|_2^6$. ■

Die Komponenten von ω werden die Cayley-Kleinschen Parameter der Drehung genannt. Die Matrix $\hat{\mathsf{B}}(\omega)$ ändert sich nicht, wenn ω durch $-\omega$ ersetzt wird. Deshalb ist jeder Matrix aus SO_3 ein Antipodenpaar $\pm\omega$ aus der Einheitssphäre S^3 des $\mathbb{R}^4$ zugeordnet. Umgekehrt liegt die dem Antipodenpaar $\pm\nu \in \mathsf{S}^3 \subset \mathbb{R}^4$ zugeordnete Matrix

$$\mathsf{C}(\nu) := \begin{pmatrix} \nu_1^2 - \nu_2^2 - \nu_3^2 + \nu_4^2 & 2(\nu_1\nu_2 - \nu_3\nu_4) & 2(\nu_1\nu_3 + \nu_2\nu_4) \\ 2(\nu_1\nu_2 + \nu_3\nu_4) & -\nu_1^2 + \nu_2^2 - \nu_3^2 + \nu_4^2 & 2(\nu_2\nu_3 - \nu_1\nu_4) \\ 2(\nu_1\nu_3 - \nu_2\nu_4) & 2(\nu_2\nu_3 + \nu_1\nu_4) & -\nu_1^2 - \nu_2^2 + \nu_3^2 + \nu_4^2 \end{pmatrix}$$

in SO_3, wie aus Satz 2.4.5 folgt. Die dadurch definierte surjektive Abbildung $\mathsf{S}^3 \subset \mathbb{R}^4 \longrightarrow \mathsf{SO}_3$ heißt Cayley-Klein-Parametrisierung der SO_3.

Mit Hilfe der Cayley-Klein-Parametrisierung und Satz 2.4.5 definiert man nun die folgenden Abbildungen

$$R\colon \mathbb{R}^4 \to \mathsf{M}(3 \times 3; \mathbb{R}) \simeq \mathbb{R}^9$$

$$q \to R(q) := \begin{pmatrix} q_1^2 - q_2^2 - q_3^2 + q_4^2 & 2(q_1q_2 - q_3q_4) & 2(q_1q_3 + q_2q_4) \\ 2(q_1q_2 + q_3q_4) & -q_1^2 + q_2^2 - q_3^2 + q_4^2 & 2(q_2q_3 - q_1q_4) \\ 2(q_1q_3 - q_2q_4) & 2(q_2q_3 + q_1q_4) & -q_1^2 - q_2^2 + q_3^2 + q_4^2 \end{pmatrix} \tag{2.9}$$

und

$$A : \mathbb{R}^4 \setminus \{0\} \to \mathsf{SO}_3 \text{ mit } q \to A(q) := R(\frac{q}{\|q\|_2}) \tag{2.10}$$

Für $q \in \mathbb{R}^4 \setminus \{0\}$ ist $\hat{q} := \frac{q}{\|q\|_2}$ ein vierdimensionaler Quaternionenvektor bzw. die Komponenten von $\hat{q}$ sind die Cayley-Klein-Parameter der Rotationsmatrix $A(q)$.
Die Eigenschaften der in den Gleichungen (2.9) und (2.10) definierten Abbildungen sind in dem folgenden Lemma zusammmengefaßt.

Lemma 2.4.6
Es sei $q, u \in \mathbb{R}^4$, mit $u \neq 0$ und $\lambda,\ \tilde{\lambda} \in \mathbb{R}$, mit $\tilde{\lambda} \neq 0$. Dann gilt

1) $R(\lambda q) = \lambda^2\, R(q)$, $\quad A(u) = \frac{1}{\|u\|_2^2}\, R(u)$;

2) $A(\tilde{\lambda} u) = A(u)$;

3) $R(q)^{\mathsf{T}} * R(q) = \|q\|_2^4\, \mathsf{E}_3$, $\quad A(u)^{\mathsf{T}} * A(u) = \mathsf{E}_3$;

4) $\det R(q) = \|q\|_2^6$, $\quad \det A(u) = 1$;

5) $\|q\|_2 = 1 \;\Rightarrow\; R(q) \in \mathsf{SO}_3$;

6) $R(q)$ *ist in* $q \in \mathbb{R}^4$ *total differenzierbar.* [22] $\mathsf{D}R(q) := \mathsf{D}\mathcal{I}_{\mathsf{M,V}}(R(q)) \in \mathsf{M}(9 \times 4; \mathbb{R})$ *mit*

$$\mathsf{D}\mathcal{I}_{\mathsf{M,V}}(R(q)) \;=\; \begin{pmatrix} 2q_1 & -2q_2 & -2q_3 & 2q_4 \\ 2q_2 & 2q_1 & -2q_4 & -2q_3 \\ 2q_3 & 2q_4 & 2q_1 & 2q_2 \\ 2q_2 & 2q_1 & 2q_4 & 2q_3 \\ -2q_1 & 2q_2 & -2q_3 & 2q_4 \\ -2q_4 & 2q_3 & 2q_2 & -2q_1 \\ 2q_3 & -2q_4 & 2q_1 & -2q_2 \\ 2q_4 & 2q_3 & 2q_2 & 2q_1 \\ -2q_1 & -2q_2 & 2q_3 & 2q_4 \end{pmatrix};$$

7) $q = (0,0,0,1)^{\mathsf{T}} \;\Rightarrow\; R(q) \;=\; A(q) \;=\; \mathsf{E}_3$.

Beweis:

1), 2) und 7) folgen unmittelbar aus der Definition von R in Gleichung (2.9) bzw. von A in Gleichung (2.10).
3) bis 5) folgen aus der Definition von R bzw. A und Satz 2.4.5. Die Differenzierbarkeit folgt aus Satz C.17, da alle Komponentenfunktionen $R(q)_{ij},\ i,j = 1,2,3$, in $q \in \mathbb{R}^4$ stetig partiell differenzierbar sind. ■

Nachdem nun die Rotationsmatrizen parametrisiert worden sind, kann man den globalen Rotationsoperator wie folgt definieren:

$$\mathsf{R}_g : U \times \mathbb{R}^3 \;\to\; \mathbb{R}^3 \quad \text{mit} \quad \mathsf{R}_g(q,c,x) \;:=\; A(q)(x-c) \;+\; c,$$

wobei $U := \mathbb{R}^4 \setminus \{0\} \times \mathbb{R}^3$, U offen gilt. Dabei ist $A(q) \in \mathsf{SO}_3$ wie in Gleichung (2.10) definiert, $c \in \mathbb{R}^3$ das Rotationszentrum und $x \in \mathbb{R}^3$ der zu rotierende Vektor. $\mathsf{R}_g|_K(q,c,.)$ rotiert die Elemente aus K entsprechend der Variablen $(q,c) \in U$. Der Operator R_g ist in $(q,c) \in U$ eine $\mathsf{C}^1(U)$-Funktion und die Funktionalmatrix $\mathsf{D}_{q,c}\mathsf{R}_g(q,c,x) \in \mathsf{M}(3 \times 7; \mathbb{R})$ ist durch $\mathsf{D}_{q,c}\mathsf{R}_g(q,c,x) \;=$

$$\begin{pmatrix} \frac{1}{\|q\|_2^2} \left(\sum\limits_{i=1}^{3} \left[(\frac{\partial}{\partial q_1} R(q)_{1i})\ \mathsf{P}_{i,3}(x-c) \right] \;-\; 2q_1\ \mathsf{P}_{1,3}\left(A(q)(x-c)\right) \right) & \dots \\ \frac{1}{\|q\|_2^2} \left(\sum\limits_{i=1}^{3} \left[(\frac{\partial}{\partial q_1} R(q)_{2i})\ \mathsf{P}_{i,3}(x-c) \right] \;-\; 2q_1\ \mathsf{P}_{2,3}\left(A(q)(x-c)\right) \right) & \dots \\ \frac{1}{\|q\|_2^2} \left(\sum\limits_{i=1}^{3} \left[(\frac{\partial}{\partial q_1} R(q)_{3i})\ \mathsf{P}_{i,3}(x-c) \right] \;-\; 2q_1\ \mathsf{P}_{3,3}\left(A(q)(x-c)\right) \right) & \dots \end{pmatrix}$$

$$\begin{matrix} \frac{1}{\|q\|_2^2} \left(\sum\limits_{i=1}^{3} \left[(\frac{\partial}{\partial q_4} R(q)_{1i})\ \mathsf{P}_{i,3}(x-c) \right] \;-\; 2q_4\ \mathsf{P}_{1,3}\left(A(q)(x-c)\right) \right) \\ \frac{1}{\|q\|_2^2} \left(\sum\limits_{i=1}^{3} \left[(\frac{\partial}{\partial q_4} R(q)_{2i})\ \mathsf{P}_{i,3}(x-c) \right] \;-\; 2q_4\ \mathsf{P}_{2,3}\left(A(q)(x-c)\right) \right) \\ \frac{1}{\|q\|_2^2} \left(\sum\limits_{i=1}^{3} \left[(\frac{\partial}{\partial q_4} R(q)_{3i})\ \mathsf{P}_{i,3}(x-c) \right] \;-\; 2q_4\ \mathsf{P}_{3,3}\left(A(q)(x-c)\right) \right) \end{matrix}$$

---

[22] $R(q)$ ist in $q \in \mathbb{R}^4$ differenzierbar, falls $\mathcal{I}_{\mathsf{M,V}}(R(q))$ in q differenzierbar ist (Definition 2.4.2), wobei $\mathcal{I}_{\mathsf{M,V}}$ ein kanonischer Isomorphismus von $\mathsf{M}(3 \times 3; \mathbb{R})$ nach $\mathbb{R}^9$ ist (Definition 2.4.1).

$$\left.\begin{matrix} 1-A(q)_{11} & -A(q)_{12} & -A(q)_{13} \\ -A(q)_{21} & 1-A(q)_{22} & -A(q)_{23} \\ -A(q)_{31} & -A(q)_{32} & 1-A(q)_{33} \end{matrix}\right) \tag{2.11}$$

gegeben. Dabei ist $\mathsf{P}_{i,3} : \mathbb{R}^3 \to \mathbb{R},\ i = 1,2,3$, mit $\mathsf{P}_{i,3}(y) = y_i$, die Projektion auf die $i-$te Komponente des Vektors y. Der Operator $\mathsf{P}_{i,3}$ wird zur Vereinfachung der Darstellung verwendet. Die Komponenten $\frac{\partial}{\partial q_k} R(q)_{ij},\ k = 1,\ldots,4,\ i,j = 1,2,3$, sind in der Matrix in Lemma 2.4.6 6) dargestellt. Die Differenzierbarkeit folgt aus Satz C.17, da die Komponentenfunktionen

$$(\mathsf{R}_g(q,c,x))_i = \left(\sum_{j=1}^{3} A(q)_{ij}(x_j - c_j)\right) + c_i = \frac{1}{\|q\|_2^2}\left(\sum_{j=1}^{3} R(q)_{ij}(x_j - c_j)\right) + c_i, \tag{2.12}$$

$i = 1,2,3$, in $(q,c) \in U$ stetig partiell differenzierbar sind.

Nun kann man die globale Transformationsabbildung T_g definieren.

Definition 2.4.7 (Globale Transformationsabbildung)
Es sei $V := \mathbb{R}^3 \times \mathbb{R}^4 \setminus \{0\} \times \mathbb{R}^3$, V *offen. Weiter sei* $\mathsf{T}_g : V \times \mathbb{R}^3 \to \mathbb{R}^3,\ (t,\ q,\ c,\ x) \to \mathsf{T}_g(t,\ q,\ c,\ x)$, *mit*

$$\mathsf{T}_g(t,q,c,x) \ := \ \tau\,(t,\ R_g(q,c,x)) \ = \ A(q)(x-c) + c + t.$$

$\diamond$

Bemerkung: $\mathsf{T}_g|_K(t,q,c,\cdot)$ transformiert die Elemente der Menge K entsprechend den Variablen $(t,q,c) \in V$.

Satz 2.4.8
Die Abbildung T_g *ist eine* $\mathsf{C}^1(V \times \mathbb{R}^3)$*-Funktion und die Jacobi-Matrix (das Differential) ist durch*

$$\mathsf{DT}_g(t,q,c,x) \ = \ (\ E_3 \quad \mathsf{D}_{q,c}\mathsf{R}_g(q,c,x) \quad A(q)\ ) \ \in \ \mathsf{M}(3 \times 13; \mathbb{R})$$

gegeben. Dabei ist $\mathsf{D}_{q,c}\mathsf{R}_g(q,c,x)$ *wie in Gleichung (2.11) definiert. Insbesondere gilt*

$$\mathsf{D}_{t,q,c}\mathsf{T}_g(t,q,c,x) \ = \ (\ \mathsf{E}_3 \quad \mathsf{D}_{q,c}\mathsf{R}_g(q,c,x)\ ) \ \in \ \mathsf{M}(3 \times 10; \mathbb{R})$$

und

$$\mathsf{D}_x\mathsf{T}_g(t,q,c,x) \ = \ A(q) \ \in \ \mathsf{SO}_3 \ \subset \ \mathsf{M}(3 \times 3; \mathbb{R}).$$

Beweis:
Da die Komponentenfunktionen

$$(\mathsf{T}_g(t,q,c,x))_i \ = \ \left(\sum_{j=1}^{n} A(q)_{ij}(x_j - c_j)\right) + c_i + t_i,\ i = 1,2,3,$$

in $(t,q,c,x) \in V \times \mathbb{R}^3$ stetig partiell differenzierbar sind, ist T_g eine $\mathsf{C}^1(V \times \mathbb{R}^3)$-Funktion. Aus Satz C.17 folgt, daß T_g in $V \times \mathbb{R}^3$ total differenzierbar ist. Das Differential erhält man, indem man die partiellen Ableitungen der Komponentenfunktionen $(\mathsf{T}_g(t,q,c,x))_i,\ i = 1,2,3$, bzgl. (t,q,c,x) bildet. ■

Im folgenden Lemma werden einige Eigenschaften von T_g dargestellt, die später noch gebraucht werden.

Lemma 2.4.9

a) Es sei $p := (t,q,c,x) \in V \mathrm{x} \mathbb{R}^3$ beliebig aber fest. Für $\lambda \in \mathbb{R}$ mit $\lambda \neq 0$ setzt man $p_\lambda := (t,\lambda q,c,x)$. Dann gilt $\mathsf{T}_g(p_\lambda) = \mathsf{T}_g(p)\ \forall \lambda \neq 0$. Insbesondere gilt sogar $\mathsf{DT}_g(p_\lambda) = \mathsf{DT}_g(p)\ \forall \lambda \neq 0$.

b) Nun seien $x,y \in \mathbb{R}^3$. Dann gilt $\|\mathsf{T}_g(t,q,c,x) - \mathsf{T}_g(t,q,c,y)\|_2 = \|x-y\|_2\ \forall\ (t,q,c) \in V$.

c) Es sei $t = (0,0,0)^\mathsf{T}$ und $q = (0,0,0,1)^\mathsf{T}$. Dann gilt $\mathsf{T}_g(t,q,c,x) = x\ \forall x,c \in \mathbb{R}^3$.

Beweis:

a) Aus Lemma 2.4.6 2) folgt $\mathsf{T}_g(p_\lambda) = \mathsf{T}_g(p)\ \forall \lambda \neq 0$.
Um $\mathsf{DT}_g(p_\lambda) = \mathsf{DT}_g(p)$ zu zeigen, bleibt noch zu zeigen, daß $\mathsf{DR}_g(\lambda q,c,x) = \mathsf{DR}_g(q,c,x)$ für alle $\lambda \neq 0$ gilt. Dabei ist $(q,c,x) \in \mathbb{R}^4 \setminus 0 \mathrm{x} \mathbb{R}^3 \mathrm{x} \mathbb{R}^3$ beliebig aber fest. Dazu betrachtet man die Komponentenfunktionen von R_g aus Gleichung (2.12). Dann gilt

$$\begin{aligned} \frac{\partial}{\partial q_k}\left(R_g(\lambda q,c,x)\right)_i &= \frac{-2q_k}{\lambda^2\|q\|_2^2}\left(\sum_{j=1}^3 R(\lambda q)_{ij}(x_j-c_j)\right) + \frac{1}{\|\lambda q\|_2^2}\left(\sum_{j=1}^3 \frac{\partial}{\partial q_k}\left(\lambda^2 R(q)_{ij}\right)(x_j-c_j)\right) \\ &= \frac{-2q_k}{\|q\|_2^2}\left(\sum_{j=1}^3 R(q)_{ij}(x_j-c_j)\right) + \frac{1}{\|q\|_2^2}\left(\sum_{j=1}^3 \frac{\partial}{\partial q_k}\left(R(q)_{ij}\right)(x_j-c_j)\right) \end{aligned}$$

für $k = 1,2,3,4$ und $i = 1,2,3$. Aus Lemma 2.4.6 2) folgt $\frac{\partial}{\partial c_k}\left(R_g(\lambda q,c,x)\right)_i = \frac{\partial}{\partial c_k}\left(R_g(q,c,x)\right)_i$. Damit folgt insgesamt die Behauptung.

b) Folgt aus Satz C.8, da $A(q) \in \mathsf{SO}_3$.

c) Folgt mit Lemma 2.4.6 7). ■

2.4.2 Lokale Rotationen

Um nachher die internen Rotationen in einer Transformationsstruktur zu beschreiben, werden vorher die folgenden Abbildungen betrachtet.

Definition 2.4.10
Es sei

$$\begin{aligned} \mathcal{B}^n\colon\ \left(\mathbb{R}^3 \setminus \{0\}\right)^n \ \mathrm{x}\ \mathbb{R}^n &\rightarrow \mathsf{SO}_3,\ n \geq 1 \\ ((a_1,\ldots,a_n),(\varphi_1,\ldots,\varphi_n)) &\rightarrow \mathcal{B}^n\left((a_1,\ldots,a_n),(\varphi_1,\ldots,\varphi_n)\right) \end{aligned}$$

mit

$$\begin{aligned}
\mathcal{B}^1\left(a_1, \varphi_1\right) &:= \mathsf{B}\left(\frac{a_1}{\|a_1\|_2},\ \varphi_1\right), \\
\mathcal{B}^n\left((a_1, \ldots, a_n), (\varphi_1, \ldots, \varphi_n)\right) &:= \mathsf{B}\left(\mathcal{B}^{n-1}\left((a_1, \ldots, a_{n-1}), (\varphi_1, \ldots, \varphi_{n-1})\right)\ \frac{a_n}{\|a_n\|_2},\ \varphi_n\right) \\
&\quad * \ \mathcal{B}^{n-1}\left((a_1, \ldots, a_{n-1}), (\varphi_1, \ldots, \varphi_{n-1})\right),\ n > 1,
\end{aligned}$$

wobei $\mathsf{B}\left(\frac{a}{\|a\|_2},\ \varphi\right)$ *die in Satz 2.4.3 definierte Rotationsmatrix ist.*

Satz 2.4.11

Die Abbildung $\mathcal{B}^n$ *aus Definition 2.4.10 ist für alle* $n \in \mathbb{N},\ n \geq 1$ *wohldefiniert.*

Beweis:

Beweis durch Induktion nach n :

$n = 1$: Nach Satz 2.4.3 ist $\mathcal{B}^1 \in \mathsf{SO}_3$ wohldefiniert.

Induktionsgbehauptung: Für alle k mit $1 \leq k < n$ ist $\mathcal{B}^k \in \mathsf{SO}_3$ wohldefiniert.

Induktionsschritt: $n - 1 \ \rightarrow \ n$

Nach Satz C.8 gilt $\mathcal{B}^{n-1}\left((a_1, \ldots, a_{n-1}), (\varphi_1, \ldots, \varphi_{n-1})\right)\ \frac{a_n}{\|a_n\|_2} \in \mathsf{S}^2$, da nach Induktionsbehauptung $\mathcal{B}^{n-1} \in \mathsf{SO}_3$ wohldefiniert ist. $\mathsf{B}\left(\mathcal{B}^{n-1}\left((a_1, \ldots, a_{n-1}), (\varphi_1, \ldots, \varphi_{n-1})\right)\ \frac{a_n}{\|a_n\|_2},\ \varphi_n\right)$ ist damit nach Satz 2.4.3 eine Rotationsmatrix. Da das Produkt zweier Rotationsmatrizen wiederum eine Rotationsmatrix ist, ist $\mathcal{B}^n \in \mathsf{SO}_3$ wohldefiniert. ■

Lemma 2.4.12

Es sei $\mathcal{B}^n$ *die Abbildung aus Definition 2.4.10. Dann gilt* $\mathcal{B}^n\left((a_1, \ldots, a_n), (0, \ldots, 0)\right) = \mathsf{E}_3$ *für alle* $(a_1, \ldots, a_n) \in (\mathbb{R}^3 \setminus \{0\})^n$.

Beweis:

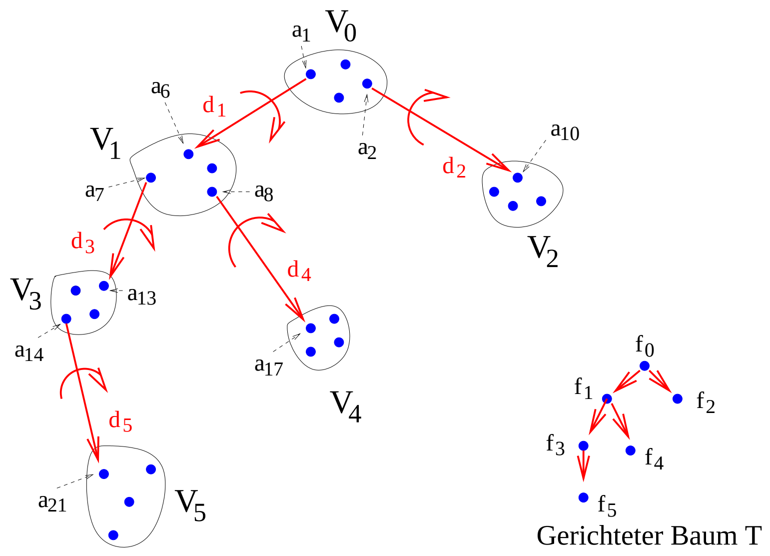
Folgt mittels Induktion nach n unmittelbar aus Korollar 2.4.4. ■

Anhand des folgenden Beispiels soll die Darstellung der lokalen Rotationen und insbesondere deren Hintereinanderschaltung hergeleitet werden.

Dazu sei $V := \{\ a_1,\ \ldots,\ a_{24}\ \} \subset \mathbb{R}^3$ mit $a_i \neq a_k, i \neq k$. Weiter sei $V_i := \{\ a_j \in V \mid 4i < j \leq 4(i+1)\ \} \neq \emptyset,\ i = 0, \ldots, 5$. Dann bilden die V_i eine disjunkte Zerlegung von V. Es sei $T = (F,\ \vec{E})$, mit $F := \{f_0, \ldots, f_5\}$ und $\vec{E} := \{(f_0, f_1), (f_0, f_2), (f_1, f_3), (f_1, f_4), (f_3, f_5)\} \subset F \times F$, ein gerichteter Baum mit der Wurzel f_0.[23] Weiter sei $\Psi_{V,F} \colon V \rightarrow F$, mit $\Psi_{V,F}(a_k) := f_j$, falls $a_k \in V_j$. D.h. die Menge $V_j \subset V$ wird dem Knoten f_j zugeordnet. Dann bildet $\mathcal{B}_V(T) = (\bigcup V_i, T = (F, \vec{E}), \Psi_{V,F})$ eine gerichtete Baumstruktur auf V (Definition 2.4.13). Abbildung 2.8 illustriert diese Baumstruktur. Die lokalen Drehachsen sind, wie folgt definiert: Für $(f_i, f_j) \in \vec{E}$ existieren $a_{k_i} \in (\Psi_{V,F})^{-1}(f_i) = V_i$ und $a_{k_j} \in (\Psi_{V,F})^{-1}(f_j) = V_j$.[24] Dann ist die lokale Drehachse, um die die Elemente aus $V_j \cup \tilde{V}_j,\ \tilde{V}_j := \{a \in V \mid \Psi_{V,F}(a)$ ist ein Nachfahre von f_j in $T\ \}$, gedreht werden, durch $\vec{d}_j := a_{k_j} - a_{k_i}$ gegeben.

[23] In Definition C.1 und Satz C.2 in Anhang C sind die hier benötigten Grundlagen aus der Graphentheorie aufgeführt.

[24] $(\Psi_{V,F})^{-1}(f_k) := \{a_l \in V \mid \Psi_{V,F}(a_l) = f_k\}$ Urbild von f_k in V.



Gerichtete Baumstruktur von V und die lokalen Drehachsen

Abbildung 2.8: Lokale Rotationen der Teilmengen V_i bzgl. lokaler Drehachsen

Damit ist die Gerade

gegeben.

Die lokale Drehung um die Gerade $g_j(\lambda) = a_{k_i} + \lambda \vec{d}_j$, $\lambda \in \mathbb{R}$, bzw. um die Drehachse d_j (die Drehachse wird mit der Geraden identifiziert) kann dann wie folgt beschrieben werden:

$$x^{\{j\}} = a_{k_i} \; + \; \mathsf{B}\left(\frac{\vec{d}_j}{\|\vec{d}_j\|_2},\; \varphi_j\right) \; (x - a_{k_i}) \;\; \forall x \in V_j \cup \tilde{V}_j,$$

wobei $\mathsf{B}\left(\frac{\vec{d}_j}{\|\vec{d}_j\|_2},\; \varphi_j\right) \in \mathsf{SO}_3$ die in Satz 2.4.3 definierte Matrix und $\varphi_j \in [0; 2\pi)$ der Rotationswinkel ist. Der Index $\{j\}$ soll angeben, daß x bzgl. der Geraden g_j bzw. der Drehachse d_j um den Rotationswinkel φ_j gedreht worden ist.
Es seien $\vec{d}_i \in \mathbb{R}^3 \backslash \{0\}, i = 1, \ldots, 5$, die oben definierten lokalen Drehachsen und $\varphi_i \in [0; 2\pi)$, $i = 1, \ldots, 5$, die zugehörigen Rotationswinkel.
Dann lassen sich die lokalen Rotationen bzgl. des ersten Levels im zugrunde liegenden Baum T wie folgt darstellen (siehe Abbildung 2.8):

$$x^{\{1\}} \;\; = a_1 + \mathsf{B}\left(\frac{\vec{d}_1}{\|\vec{d}_1\|_2},\; \varphi_1\right)(x - a_1) \;\; = \;\; a_1 + \mathcal{B}^1\left(\vec{d}_1, \varphi_1\right)(x - a_1)\,, \; \forall x \in V_1 \cup \tilde{V}_1$$

und

$$y^{\{2\}} \;\; = a_2 + \mathcal{B}^1\left(\vec{d}_2, \varphi_1\right)(y - a_2)\,, \; \forall y \in V_2.$$

Der Index $\{1\}$ soll angeben, daß x bzgl. der Geraden g_1 bzw. der Drehachse d_1 um den Rotationswinkel φ_1 gedreht wurde. Entsprechendes gilt für den Index $\{2\}$. Diese Indexschreibweise wird im folgenden weiter verwendet verwendet. Bevor man die lokalen Rotationen bzgl. des zweiten Level im Baum darstellt, muß man beachten, daß sich die lokalen Drehachsen $\vec{d}_3, \vec{d}_4$ und $\vec{d}_5$ auch gedreht haben. Denn es gilt

$$\vec{d}_3^{\{1\}} \;=\; a_{13}^{\{1\}} - a_7^{\{1\}} \;=\; \mathsf{B}\left(\frac{\vec{d}_1}{\|\vec{d}_1\|_2},\ \varphi_1\right)(a_{13} - a_7) \;=\; \mathcal{B}^1\left(\vec{d}_1, \varphi_1\right)\vec{d}_3$$

Ebenso folgt, daß für $i = 4, 5$ $\vec{d}_i^{\{1\}} = \mathcal{B}^1\left(\vec{d}_1, \varphi_1\right)\vec{d}_i$ gilt. Wegen $\mathcal{B}^1\left(\vec{d}_1, \varphi_1\right) \in \mathsf{SO}_3$ (nach Satz 2.4.11), gilt nach Satz C.8 $\|d_i^{\{1\}}\|_2 = \|d_i\|_2$ für $i = 3, 4, 5$. Nun lassen sich die lokalen Rotationen bzgl. des zweiten Levels wie folgt darstellen:

$$\begin{aligned}
x^{\{1,3\}} &= a_7^{\{1\}} + \mathsf{B}\left(\frac{\vec{d}_3^{\{1\}}}{\|\vec{d}_3^{\{1\}}\|_2},\ \varphi_3\right)\left(x^{\{1\}} - a_7^{\{1\}}\right) \\
&= a_7^{\{1\}} + \mathsf{B}\left(\mathcal{B}^1\left(\vec{d}_1, \varphi_1\right)\frac{\vec{d}_3}{\|\vec{d}_3\|_2},\ \varphi_3\right) * \mathcal{B}^1\left(\vec{d}_1, \varphi_1\right)(x - a_7) \\
&= a_7^{\{1\}} + \mathcal{B}^2\left(\vec{d}_1, \vec{d}_3, \varphi_1, \varphi_3\right)(x - a_7)\,,\ \forall x \in V_3 \cup \tilde{V}_3
\end{aligned}$$

und

$$y^{\{1,4\}} = a_8^{\{1\}} + \mathcal{B}^2\left(\vec{d}_1, \vec{d}_4, \varphi_1, \varphi_4\right)(y - a_8)\,,\ \forall y \in V_4$$

Die lokale Drehachse d_5 hat sich wiederum mitgedreht. Es gilt

$$\vec{d}_5^{\{1,3\}} \;=\; a_{21}^{\{1,3\}} - a_{14}^{\{1,3\}} \;=\; \mathcal{B}^2\left(\vec{d}_1, \vec{d}_3, \varphi_1, \varphi_3\right)(a_{21} - a_{14}) \;=\; \mathcal{B}^2\left(\vec{d}_1, \vec{d}_3, \varphi_1, \varphi_3\right)\vec{d}_5$$

Wegen $\mathcal{B}^2\left(\vec{d}_1, \vec{d}_3, \varphi_1, \varphi_3\right) \in \mathsf{SO}_3$ (nach Satz 2.4.11), gilt nach Satz C.8 $\|d_5^{\{1,3\}}\|_2 = \|d_5\|_2$. Die lokalen Rotationen bzgl. des dritten Levels lassen sich nun wie folgt darstellen:

$$\begin{aligned}
x^{\{1,3,5\}} &= a_{14}^{\{1,3\}} + \mathsf{B}\left(\mathcal{B}^2\left(\vec{d}_1, \vec{d}_3, \varphi_1, \varphi_3\right)\frac{\vec{d}_5}{\|\vec{d}_5\|_2},\ \varphi_5\right) * \mathcal{B}^2\left(\vec{d}_1, \vec{d}_3, \varphi_1, \varphi_3\right)(x - a_{14}) \\
&= a_{14}^{\{1,3\}} + \mathcal{B}^3\left(\vec{d}_1, \vec{d}_4, \vec{d}_5, \varphi_1, \varphi_4, \varphi_5\right)(x - a_{14})\,,\ \forall x \in V_5
\end{aligned}$$

An Hand des obigen Beispiels ist die Darstellung der lokalen Rotationen und deren Hintereinanderschaltung dargestellt worden. Die lokalen Rotationen werden entlang von gerichteten Wegen in T hintereinander geschaltet. Je länger ein gerichteter Weg ist, um so mehr Rotationen müssen hintereinander ausgeführt werden. Im obigen Beispiel hat der f_0-f_5−Weg die Länge 3, d.h. auf $V_5 = (\Psi_{V,F})^{-1}(f_5)$ wirken drei lokale Rotationen, die hintereinander ausgeführt werden müssen. Im Baum $T' = \left(F, (\vec{E} \setminus (f_0, f_1) \cup (f_1, f_0)\right)$, mit der Wurzel f_1, müssen jedoch maximal nur 2 Rotationen hintereinander geschaltet werden, da nur gerichtete Wege mit der Länge höchstens 2 vorkommen. Lokale Rotationen, die nicht hintereinander geschaltet werden müssen, sind unabhängig voneinander. Daher sollte die Wurzel des zugrundeliedenden gerichteten Baumes so gewählt sein, daß die gerichteten Wege möglichst kurz sind; der gerichtete Baum sollte also möglichst balanciert sein.

Nun folgt die formale Definition des lokalen Rotationsoperators.

Definition 2.4.13 (Baumstruktur)
Es sei $V \subset \mathbb{R}^3$ mit $V = \bigcup_{i=0}^n V_i$, mit $V_i \neq \emptyset$, $i = 0, \ldots, n$, und $V_i \cap V_j = \emptyset$ für $i \neq j$. Weiter sei $T = (F, E)$ ein ungerichteter Baum, d.h zusammenhängend und kreislos, mit $F := \{f_0, \ldots, f_n\}$ und $E := \{ \{f_i, f_j\} \mid f_i, f_j \in F,\ i \neq j \} \subseteq \mathcal{P}_2(F)$. Es sei $\Psi_{V,F} : V \to F$, mit $\Psi_{V,F}(x) := f_i$, falls $x \in V_i$. Damit gilt $V = \bigcup_{i=0}^n (\Psi_{V,F})^{-1}(f_i)$. Dann heißt

$$\mathcal{B}_V(T) \ = \ \left\{ \bigcup_{i=0}^{n} V_i,\ T,\ \Psi_{V,F} \right\} \tag{2.13}$$

eine Baumstruktur auf V bzgl. des Baumes T.
Nach Satz C.2 gibt es für je zwei Knoten aus F genau einen Weg in T, der die beiden Knoten verbindet. Für $f_i \in F \setminus \{f_0\}$ sei $P_{f_0,f_i} = (V(P_{f_0,f_i}), E(P_{f_0,f_i}))$ der eindeutig bestimmte f_0-f_i−Weg in T. Die Kanten aus $E(P_{f_0,f_i})$ werden so gerichtet, daß $\vec{P}_{f_0,f_i}$ ein gerichteter f_0-f_i−Weg ist. Der so entstehende gerichtete Baum $\vec{T} = (F, \vec{E})$ enthält für jeden Knoten $f_i \in F$, $f_i \neq f_0$, einen gerichteten f_0-f_i−Weg. Damit ist f_0 die Wurzel des gerichteten Baumes $\vec{T}$. Ersetzt man nun in Gleichung (2.13) T durch $\vec{T}$, so erhält man eine gerichtete Baumstruktur auf V. ◇

Mit Hilfe der gerichteten Baumstruktur auf V kann man nun eine Transformationsstruktur auf V definieren. Die Transformationsstruktur beschreibt die lokalen Drehachsen.

Definition 2.4.14
Es sei $V = \bigcup_{i=0}^n V_i$, mit $V_i \neq \emptyset$, $i = 0, \ldots, n$, und $V_i \cap V_j = \emptyset$ für $i \neq j$. Weiter sei $\vec{T} = (F, \vec{E})$ mit $F := \{f_0, \ldots, f_n\}$ und $\vec{E} \subset F \times F$ ein gerichteter Baum mit der Wurzel f_0. Die Abbildung $\Psi_{V,F}$ sei wie in Definition 2.4.13. Dann sind die Abbildungen $\mathsf{R}_{F,V} : F \setminus \{f_0\} \to V$, und $\mathsf{S}_{F,V} : F \setminus \{f_0\} \to V$ wie folgt definiert:
für $(f_i, f_j) \in \vec{E} :\ \exists x \in (\Psi_{V,F})^{-1}(f_i) = V_i \ \wedge\ \exists y \in (\Psi_{V,F})^{-1}(f_j) = V_j$ mit $\mathsf{R}_{F,V}(f_j) := x$ und $\mathsf{S}_{F,V}(f_j) := y$.
Die Abbildung $\mathsf{D}_{F,\mathbb{R}^3\setminus\{0\}} : F \setminus \{f_0\} \to \mathbb{R}^3 \setminus \{0\}$ ist definiert durch: $\mathsf{D}_{F,\mathbb{R}^3\setminus\{0\}}(f_j) := \mathsf{S}_{F,V}(f_j) - \mathsf{R}_{F,V}(f_j)$, $\forall f_j \in F \setminus f_0$. Dann heißt

$$\mathcal{S}(V, \vec{T}) \ = \ \left\{ \bigcup_{i=0}^{n} V_i,\ \vec{T},\ \Psi_{V,F},\ \mathsf{R}_{F,V},\ \mathsf{S}_{F,V},\ \mathsf{D}_{F,\mathbb{R}^3\setminus\{0\}} \right\}$$

eine Transformationsstruktur auf V bzgl. des Baumes $\vec{T}$. ◇

Lemma 2.4.15
Die Abbildungen $\mathsf{R}_{F,V}$, $\mathsf{S}_{F,V}$ und $\mathsf{D}_{F,\mathbb{R}^3\setminus\{0\}}$ aus Definition 2.4.14 sind wohldefiniert.

Beweis:
Es sei $f_i \in F \setminus \{f_0\}$. Dann gibt es genau ein $f_k \in F$ mit $(f_k, f_i) \in \vec{E}$. Denn angenommen es existieren f_k, $f_l \in F$, $f_k \neq f_l$, mit (f_k, f_i), $(f_l, f_i) \in \vec{E}$. Da f_0 die Wurzel des Baumes $\vec{T}$, enthält $\vec{T}$ einen gerichteten f_0-f_k−Weg $\vec{P}_{f_0,f_k}$ und einen gerichteten f_0-f_l−Weg $\vec{P}_{f_0,f_l}$. Dann sind $\vec{P}_{f_0,f_k} + (f_k, f_i)$ und $\vec{P}_{f_0,f_l} + (f_l, f_i)$ gerichtete f_0-f_i−Wege in $\vec{T}$. Der gerichtete f_0-f_i−Weg ist jedoch eindeutig bestimmt, da es genau einen f_0-f_i−Weg im zugrundeliegenden ungerichteten Baum T gibt und dieser ist entsprechend gerichtet worden. Also gilt $\vec{P}_{f_0,f_l} + (f_l, f_i) = \vec{P}_{f_0,f_l} + (f_l, f_i)$. Daraus folgt aber $f_k = f_l$. Widerspruch zur Annahme.

Für $f_i \in F\backslash\{f_0\}$ existiert genau ein $f_k \in F$ mit $(f_k, f_i) \in \vec{E}$. Damit existieren $x \in (\Psi_{V,F})^{-1}(f_k) = V_k \neq \emptyset$ und $y \in (\Psi_{V,F})^{-1}(f_i) = V_i \neq \emptyset$. Damit ist $\mathsf{R}_{F,V}(f_i)$ und $\mathsf{S}_{F,V}(f_i)$ bis auf Wahl von x und y eindeutig festgelegt. Und wegen $V_k \cap V_i = \emptyset$ gilt $\mathsf{D}_{F,\mathbb{R}^3\backslash\{0\}}(f_i) \neq 0$. ■

Satz 2.4.16

Es sei $V \subset \mathbb{R}^3$. Weiter sei $\bigcup_{i=0}^{n} V_i$, $V_i \neq \emptyset$, $n \geq 1$ eine disjunkte Zerlegung von V. Es sei $\vec{T} = (F, \vec{E})$ ein gerichteter Baum mit $F = \{f_0, \ldots, f_n\}$, $\vec{E} \subseteq F \mathsf{x} F$ und mit der Wurzel f_0. Da f_0 Wurzel des Baumes $\vec{T}$ ist, enthält $\vec{T}$ für jeden Knoten $f_i \in F \setminus \{f_0\}$ genau einen gerichteten Weg $\vec{P}_{f_0,f_i} = (f_0 f_{k_1} \ldots f_{k_{l_i-1}} f_i)$. Der Weg $\vec{P}_{f_0,f_i}$ hat die Länge $l_i := l(\vec{P}_{f_0,f_i})$. Die Abbildung $\Psi_{V,F}\colon V \to F$ sei wie in Definition 2.4.13. Die Abbildungen $\mathsf{R}_{F,V}\colon F \setminus \{f_0\} \to V$, $\mathsf{S}_{F,V} \setminus \{f_0\}\colon F \to V$ und $\mathsf{D}_{F,\mathbb{R}^3\backslash\{0\}}\colon F \setminus \{f_0\} \to \mathbb{R}^3 \setminus \{0\}$ seien wie in Definition 2.4.14. $\mathcal{S}(V, \vec{T})$ ist die Transformationsstruktur auf V bzgl. $\vec{T}$. Jetzt kann man jeder Kante $(f_p, f_q) \in \vec{E}$ eine eindeutige lokale Drehachse zuordnen (nach Lemma 2.4.15), indem man für den Endknoten $f_q \in F \setminus \{f_0\}$ den Vektor $d_q := \mathsf{D}_{F,\mathbb{R}^3\backslash\{0\}}(f_q) = \mathsf{S}_{F,V}(f_q) - \mathsf{R}_{F,V}(f_q) \in \mathbb{R}^3 \setminus \{0\}$ definiert. Analog ordnet man jeder Kante $(f_p, f_q) \in \vec{E}$ (und damit indirekt der Rotationsachse $d_q := \mathsf{D}_{F,\mathbb{R}^3\backslash\{0\}}(f_q)$) den Rotationsparameter $\varphi_q \in \mathbb{R}$ zu. Damit kann man jedem Weg $\vec{P}_{f_0,f_i} = (f_0 f_{k_1} \ldots f_{k_{l_i-1}} f_i) \subset \vec{T}$, eine eindeutig bestimmte Folge von lokalen Drehachsen $(d_{k_1}, \ldots, d_{k_{l_i-1}}, d_i)$ und von Rotationsparametern $\Phi[i] := (\varphi_{k_1}, \ldots, \varphi_{k_{l_i-1}}, \varphi_i)$ zuordnen. Man definiert zunächst $\mathsf{R}[\mathcal{S}(V, \vec{T}), f_0]\colon (\Psi_{V,F})^{-1}(f_0)(= V_0) \to V_0$ mit $\mathsf{R}[\mathcal{S}(V, \vec{T}), f_0](x) = x$, $\forall x \in V_0$. Dann beschreibt die Abbildung

$$\mathsf{R}[\mathcal{S}(V,\vec{T}), f_i]\colon \mathbb{R}^{l_i} \mathsf{x} (\Psi_{V,F})^{-1}(f_i)(= V_i) \to \mathbb{R}^3,\ (\Phi[i],\ x) \to\ \mathsf{R}[\mathcal{S}(V,\vec{T}), f_i]\,(\Phi[i],\ x)$$

mit

$$\begin{aligned}\mathsf{R}[\mathcal{S}(V,\vec{T}), f_i]\,(\Phi[i],\ x) := {} & \mathsf{R}[\mathcal{S}(V,\vec{T}), f_{k_i}]\,(\Phi[k_{l_i-1}],\ \mathsf{R}_{F,V}(f_i)) \\ & + \mathcal{B}^{l_i}\left(d_{k_1}, \ldots, d_{k_{l_i-1}}, d_i, \Phi[i]\right)(x - \mathsf{R}_{F,V}(f_i)) \end{aligned} \tag{2.14}$$

für $f_i \in F \setminus \{f_0\}$ mit $\vec{P}_{f_0,f_i} \subset \vec{T}$ die Hintereinanderausführung der l_i verschiedenden lokalen Rotationen bzgl. der entsprechenden Drehachsen und Rotationswinkel auf die Menge $(\Psi_{V,F})^{-1}(f_i)(= V_i)$. Die l_i Drehachsen d_{k_j} sind durch den Weg $\vec{P}_{f_0,f_i} \subset \vec{T}$ vorgegeben. Die l_i Rotationswinkel φ_{k_j} sind die den Rotationsachsen zugeordneten Rotationsparameter der Abbildung.

Beweis:

Aus $n \geq 1$ folgt $|F| \geq 2$. Die Tiefe des gerichteten Baumes $\vec{T}$ ist definiert durch $t(\vec{T}) := \max_{f_i \in F\backslash\{f_0\}} l(\vec{P}_{f_0,f_i})$, wobei $\vec{P}_{f_0,f_i} \subset \vec{T}$ gilt. Nun erfolgt der Beweis, daß $\mathsf{R}[\mathcal{S}(V, \vec{T}), f_i]$ die Hintereinanderausführung der verschiedenen lokalen Rotationen beschreibt, durch Induktion nach der Länge des Weges $\vec{P}_{f_0,f_i}$. Dazu setzt man $\mathsf{R}_V[f_i] = \mathsf{R}[\mathcal{S}(V, \vec{T}), f_i]$ für $f_i \in F$.

Induktionsanfang: Es sei f_i in F mit $l(\vec{P}_{f_0,f_i}) = 1$.

Dann gilt $\vec{P}_{f_0,f_i} = (f_0, f_i) \in \vec{E}$. Nach Definition 2.4.14 gilt $\mathsf{R}_{F,V}(f_i) \in V_0$ und $\mathsf{S}_{F,V}(f_i) \in V_i$. Die lokale Drehachse ist definiert durch $d_i := \mathsf{D}_{F,\mathbb{R}^3\backslash\{0\}}(f_i) = \mathsf{S}_{F,V}(f_i) - \mathsf{R}_{F,V}(f_i) \in \mathbb{R}^3 \setminus \{0\}$. Die Rotation von V_i um die Drehachse d_i kann nun wie folgt beschrieben werden (wobei

$\mathsf{B}\left(\frac{d_i}{\|d_i\|_2},\ \varphi_i\right)$ die in Satz 2.4.3 definierte Rotationsmatrix und φ_i der Rotationswinkel ist):

$$\begin{aligned} x^{rot} &= \mathsf{R}_{F,V}(f_i) \;-\; \mathsf{B}\left(\frac{d_i}{\|d_i\|_2},\ \varphi_i\right)\ (x - \mathsf{R}_{F,V}(f_i)) \\ &= \mathsf{R}_V[f_0]\,(\mathsf{R}_{F,V}(f_i)) \;+\; \mathcal{B}^1\,(d_i,\ \varphi_i)\ \ (x - \mathsf{R}_{F,V}(f_i)) \\ &= \mathsf{R}_V[f_i]\,(\varphi_i, x)\,,\ \forall\ x \in V_i. \end{aligned}$$

$\mathcal{B}^1\,(\ldots)$ ist die in Definition 2.4.10 definierte Rotationsmatrix.

Induktionsbehauptung: Für alle $f_j \in F \setminus \{0\}$ mit $1 \leq l_j := l(\vec{P}_{f_0,f_j}) \leq m < t(\vec{T})$, wobei $\vec{P}_{f_0,f_j} = (f_0 f_{k_1} \ldots f_{k_{l_j-1}} f_j) \subset \vec{T}$ gilt, beschreibt $\mathsf{R}_V[f_j]\,(\Phi[j], x)$, $x \in (\Psi_{V,F})^{-1}(f_j) = V_j$, $\Phi[j] := \left(\varphi_{k_1}, \ldots \varphi_{k_{l_j-1}}, \varphi_j\right)$, die Hintereinanderschaltung der l_j Rotationen der Menge V_j bzgl. der lokalen Drehachsen $d_{k_1}, \ldots, d_{k_{l_j-1}}$ und d_j und der entsprechenden Rotationswinkeln $\varphi_{k_1}, \ldots, \varphi_{k_{l_j-1}}$ und φ_j. Die Drehachsen und die Rotationsparameter sind durch den Pfad $\vec{P}_{f_0,f_j}$ eindeutig gegeben.

Induktionsschritt: $m \;\to\; m{+}1$

Es sei $f_p \in F \setminus \{f_0\}$ mit $\vec{P}_{f_0,f_p} = (f_0 f_{k_1} \ldots f_{k_{l_p-1}} f_p) \subset \vec{T}$ und $l_p := l(\vec{P}_{f_0,f_p}) = m + 1$. Mit $f_{k_{l_p-1}} = f_{k_m}$ hat dann der Pfad $\vec{P}_{f_0,f_{k_m}}$ die Länge $l(\vec{P}_{f_0,f_{k_m}}) = m$. Nach Induktionsbehauptung gilt, daß die Abbildung $\mathsf{R}_V[f_{k_m}]\,(\Phi[k_m], x)$, $x \in (\Psi_{V,F})^{-1}(f_{k_m}) = V_{k_m}$, $\Phi[k_m] := (\varphi_{k_1}, \ldots, \varphi_{k_m})$, die Hintereinanderschaltung der m Rotationen der Menge V_{k_m} bzgl. der m lokalen Drehachsen $d_{k_1}, \ldots, d_{k_m}$ und der entsprechenden Rotationswinkeln $\varphi_{k_1}, \ldots, \varphi_{k_m}$ ist. Die Drehachsen sind durch den Pfad $\vec{P}_{f_0,f_{k_m}}$ eindeutig gegeben. Nun wendet man die Abbildung $\mathsf{R}_V[f_{k_m}]$ auch auf $(\Psi_{V,F})^{-1}(f_p) = V_p$ an. Dies ist notwendig, da f_p ein Nachfolger von f_{k_m} in $\vec{T}$ ist. Für $(f_{k_m}, f_p) \in \vec{E}$ gilt nach Definition 2.4.14 $\mathsf{R}_{F,V}(f_p) \in V_{k_m}$ und $\mathsf{S}_{F,V}(f_p) \in V_p$. Es sei $d_p := \mathsf{D}_{F,\mathbb{R}^3\setminus\{0\}}(f_p) = \mathsf{S}_{F,V}(f_p) - \mathsf{R}_{F,V}(f_p) \in \mathbb{R}^3 \setminus \{0\}$. Die lokale Drehachse bzgl. der $m{+}1$.-Rotation von V_p ist durch

$$\begin{aligned} d_p^{\{m+1\}} &:= \mathsf{R}_V[f_{k_m}]\,(\Phi[k_m], \mathsf{S}_{F,V}(f_p)) \;-\; \mathsf{R}_V[f_{k_m}]\,(\Phi[k_m], \mathsf{R}_{F,V}(f_p)) \\ &= \mathcal{B}^m\,(d_{k_1}, \ldots, d_{k_m}, \Phi[k_m])\,(\mathsf{S}_{F,V}(f_p) - \mathsf{R}_{F,V}(f_p)) \\ &= \mathcal{B}^m\,(d_{k_1}, \ldots, d_{k_m}, \Phi[k_m])\ \ d_p \end{aligned}$$

gegeben. Nach Satz 2.4.11 ist $\mathcal{B}^m\,(\ldots)$ eine Rotationsmatrix und mit Satz C.8 gilt $\|d_p^{\{m+1\}}\|_2 = \|d_p\|_2$. Nun läßt sich die $m{+}1$.-Rotation von V_p bzgl. der Drehachse $d_p^{\{m+1\}}$ und des Rotationswinkel φ_p wie folgt darstellen

$$\begin{aligned} x^{rot} := {} & \mathsf{R}_V[f_{k_m}]\,(\Phi[k_m], \mathsf{R}_{F,V}(f_p)) \;+\; \mathsf{B}\left(\mathcal{B}^m\,(d_{k_1}, \ldots, d_{k_m}, \Phi[k_m])\ \frac{d_p}{\|d_p\|_2},\ \varphi_p\right) \\ & \quad [\mathsf{R}_V[f_{k_m}]\,(\Phi[k_m], x) \;-\; \mathsf{R}_V[f_{k_m}]\,(\Phi[k_m], \mathsf{R}_{F,V}(f_p))] \\ = {} & \mathsf{R}_V[f_{k_m}]\,(\Phi[k_m], \mathsf{R}_{F,V}(f_p)) \;+\; \mathsf{B}\left(\mathcal{B}^m\,(d_{k_1}, \ldots, d_{k_m}, \Phi[k_m])\ \frac{d_p}{\|d_p\|_2},\ \varphi_p\right) \\ & \quad *\ \mathcal{B}^m\,(d_{k_1}, \ldots, d_{k_m}, \Phi[k_m])\,(x - \mathsf{R}_{F,V}(f_p)) \\ = {} & \mathsf{R}_V[f_{k_m}]\,(\Phi[k_m], \mathsf{R}_{F,V}(f_p)) + \mathcal{B}^{m+1}\,(d_{k_1}, \ldots, d_{k_m}, d_p, \Phi[k_m], \varphi_p)\ \ (x - \mathsf{R}_{F,V}(f_p)) \\ = {} & \mathsf{R}_V[f_p]\,(\Phi[k_m], \varphi_p, x)\,,\ \forall\ x \in V_p \end{aligned}$$

$\blacksquare$

Korollar 2.4.17
Mit den Bezeichnugnen aus Satz 2.4.16 gilt für alle $x \in V$ *mit* $\Psi_{V,F}(x) = f_i \neq f_0$:

$$\mathsf{R}[\mathcal{S}(V,\vec{T}), f_i]\,(\Phi[i], x) = \mathsf{R}_{F,V}(f_{k_1}) + \sum_{p=1}^{l_i-1} \mathcal{B}^p\left(d_{k_1}, \dots, d_{k_p}, \varphi_{k_1}, \dots, \varphi_{k_p}\right) y_p$$
$$+ \mathcal{B}^{l_i}\left(d_{k_1}, \dots, d_{k_{l_i-1}}, d_i, \Phi[i]\right)(x - \mathsf{R}_{F,V}(f_i)),$$

mit $\Phi[i] := \left(\varphi_{k_1}, \dots, \varphi_{k_{l_i-1}}, \varphi_i\right)$ *und* $l_i := l(\vec{P}_{f_0,f_i})$, $\vec{P}_{f_0,f_i} = (f_0, f_{k_1}, \dots, f_{k_{l_i-1}}, f_i) \subset \vec{T}$, $y_p := \mathsf{R}_{F,V}(f_{k_{p+1}}) - \mathsf{R}_{F,V}(f_{k_p})$, $p = 1, \dots, l_i - 2$, *und* $y_{l_i-1} := \mathsf{R}_{F,V}(f_i) - \mathsf{R}_{F,V}(f_{k_{l_i-1}})$. ■

Bemerkung: Es sei $\mathcal{S}(V,\vec{T} = (F,\vec{E}))$ eine Transformationsstruktur (Bezeichnungnen wie in Satz 2.4.16). Indem man jeder Kante $(f_k, f_l) \in \vec{E}$ die Drehachse $d_l := \mathsf{D}_{F,\mathbb{R}^3\setminus\{0\}}(f_l)$ zuordnet, kann man der Kante (genauer dem Endknoten f_l), den entsprechenden Rotationswinkel φ_l zuordnen. So entsteht eine Hierarchisierung der Rotationswinkel mittels der gerichteten Wege in $\vec{T}$. Die Entfernung des Endknotens f_l (Länge des f_0-f_l−Weges in $\vec{T}$) von der Wurzel f_0 bezeichnet den Rang des entsprechenden Rotationswinkel φ_l. Für alle Knoten $f_{i_1} \in F \setminus \{f_0\}$ mit $(f_0, f_{i_1}) \in \vec{E}$ haben die Rotationswinkel φ_{i_1} den Rang 1. Diese Rotationswinkel sind unabhängig von allen anderen Rotationswinkeln, wie man aus der Definition von $\mathsf{R}[\mathcal{S}(V,\vec{T}), f_{i_1}]$ in Satz 2.4.16 entnehmen kann. Die entsprechenden Rotationswinkel φ_{i_2}, die zu den Knoten $f_{i_2} \in F \setminus \{f_0\}$ mit $(f_0 f_{i_1} f_{i_2}) \subset \vec{T}$ gehören, haben den Rang 2. Aus der Definition von $\mathsf{R}[\mathcal{S}(V,\vec{T}), f_{i_2}]$ sieht man, daß der Winkel φ_{i_2} nur von dem entsprechenden Rotationswinkel φ_{i_1} abhängt. Damit steht der Winkel φ_{i_1} in der Hierarchie höher als der Winkel φ_{i_2}. Induktiv kann man mit dieser Charakterisierung der Rotationswinkel fortfahren. Je länger der f_0-f_j−Weg in $\vec{T}$ ist, um so abhängiger ist der dem Knoten f_j zugeordnete Winkel φ_j von entsprechenden Rotationswinkel höhreren Ranges in der Hierarchie. Die Rotationswinkel höhreren Ranges sind durch den f_0-f_j−Weg eindeutig bestimmt.

Definition 2.4.18 (Lokaler Rotationsoperator)
Mit den Bezeichnungen aus Satz 2.4.16 heißt die Abbildung

$$\mathcal{R}[\mathcal{S}(V,\vec{T})]:\ \mathbb{R}^n \times V \ \to\ \mathbb{R}^3,\ (\varphi_1,\ \dots,\ \varphi_n,\ x)\ \to\ \mathcal{R}[\mathcal{S}(V,\vec{T})]\,(\varphi_1,\ \dots,\ \varphi_n,\ x)$$

$$\mathcal{R}[\mathcal{S}(V,\vec{T})]\,(\varphi_1,\ \dots,\ \varphi_n,\ x) := \begin{cases} \mathsf{R}[\mathcal{S}(V,\vec{T}), f_0](x) & ,\ \Psi_{V,F}(x) = f_0 \\ \mathsf{R}[\mathcal{S}(V,\vec{T}), f_i]\,(\varphi_{k_1},\ \dots,\ \varphi_{k_i},\ \varphi_i,\ x) & ,\ \Psi_{V,F}(x) = f_i \neq f_0 \end{cases}$$

lokaler Rotationsoperator von V *bzgl. der Transformationsstruktur* $\mathcal{S}(V,\vec{T})$. ◇

Lemma 2.4.19
Es sei $\mathcal{R}[\mathcal{S}(V,\vec{T})]$ *die Abbildung aus Definition 2.4.18. Dann gilt* $\mathcal{R}[\mathcal{S}(V,\vec{T})](0,\dots,0,x) = x$ *für alle* $x \in V$.

Beweis:

Für $\Psi_{V,F}(x) = f_0$ gilt $\mathcal{R}[\mathcal{S}(V,\vec{T})](0,\dots,0,x) = \mathsf{R}[\mathcal{S}(V,\vec{T}),f_0]\,(0,\dots,0,x) = x$ nach Definition von $\mathsf{R}[\mathcal{S}(V,\vec{T}),f_0]$ in Satz 2.4.16. Für $\Psi_{V,F}(x) = f_i \neq f_0$ folgt die Aussage aus Lemma 2.4.12, das auf die Darstellung von $\mathsf{R}[\mathcal{S}(V,\vec{T}),f_i]\,(0,\dots,0,x)$ in Korollar 2.4.17 angewendet wird. ■

Um dieses Unterkapitel abzuschließen, fehlt noch das Änderungsverhalten von $\mathcal{R}[\mathcal{S}(V,\vec{T})]$ bzgl. der Rotationswinkel $\varphi_i,\ i = 1,\dots,n$. Im folgenden sei $\mathsf{B}\,(\cdot,\ \cdot)$ die in Satz 2.4.3 definierte Matrix. $\mathcal{I}_{\mathsf{M,V}}$ und $\mathcal{I}_{\mathsf{V,M}}$ sind die kanonischen Isomorphismen bzgl. dem Matrixraum $\mathsf{M}(3\times 3;\mathbb{R})$ und dem Vektorraum $\mathbb{R}^9$ aus Definition 2.4.1. Im folgenden wird die Definition 2.4.2 für die Differenzierbarkeit einer parameterabhängigen Matrix verwendet.

Lemma 2.4.20

a) Es sei $\mathcal{F}_a\colon \mathbb{R} \to \mathsf{M}(3\times 3;\mathbb{R})$ *mit* $\varphi \to \mathcal{F}_a(\varphi) := \mathsf{B}\,(a,\ \varphi)$, *wobei* $a \in \mathbb{R}^3$ *beliebig aber fest ist. Dann ist die Kurve* $\mathcal{F}_a(\varphi)$ *eine* $\mathsf{C}^1(\mathbb{R})$*-Funktion und für den Tangentialvektor (das Differential) der Kurve gilt* $\mathsf{D}\mathcal{F}_a(\varphi) = \mathsf{D}_\varphi\mathsf{B}\,(a,\ \varphi) := \mathsf{D}_\varphi\mathcal{I}_{\mathsf{M,V}}(\mathsf{B}\,(a,\ \varphi)) =$

$$\mathcal{I}_{\mathsf{M,V}}\left(\begin{pmatrix} (a_1^2-1)\sin\varphi & a_1a_2\sin\varphi - a_3\cos\varphi & a_1a_3\sin\varphi + a_2\cos\varphi \\ a_1a_2\sin\varphi + a_3\cos\varphi & (a_2^2-1)\sin\varphi & a_2a_3\sin\varphi - a_1\cos\varphi \\ a_1a_3\sin\varphi - a_2\cos\varphi & a_2a_3\sin\varphi + a_1\cos\varphi & (a_3^2-1)\sin\varphi \end{pmatrix}\right) \in \mathbb{R}^9.$$

b) Es sei $\mathcal{F}_\varphi\colon \mathbb{R}^3 \to \mathsf{M}(3\times 3;\mathbb{R})$ *mit* $a \to \mathcal{F}_\varphi(a) := \mathsf{B}\,(a,\ \varphi)$, *wobei* $\varphi \in \mathbb{R}$ *beliebig aber fest ist. Dann ist die Abbildung* $\mathcal{F}_\varphi(a)$ *eine* $\mathsf{C}^1(\mathbb{R}^3)$*-Funktion und für das Differential gilt* $\mathsf{D}\mathcal{F}_\varphi(a) =$

$$\mathsf{D}_a\mathsf{B}\,(a,\ \varphi) := \begin{pmatrix} 2\,(1-\cos\varphi)\,a_1 & 0 & 0 \\ (1-\cos\varphi)\,a_2 & (1-\cos\varphi)\,a_1 & -\sin\varphi \\ (1-\cos\varphi)\,a_3 & \sin\varphi & (1-\cos\varphi)\,a_1 \\ (1-\cos\varphi)\,a_2 & (1-\cos\varphi)\,a_1 & \sin\varphi \\ 0 & 2\,(1-\cos\varphi)\,a_2 & 0 \\ -\sin\varphi & (1-\cos\varphi)\,a_3 & (1-\cos\varphi)\,a_2 \\ (1-\cos\varphi)\,a_3 & -\sin\varphi & (1-\cos\varphi)\,a_1 \\ \sin\varphi & (1-\cos\varphi)\,a_3 & (1-\cos\varphi)\,a_2 \\ 0 & 0 & 2\,(1-\cos\varphi)\,a_3 \end{pmatrix} \in \mathsf{M}(9\times 3;\mathbb{R}).$$

c) Es sei $\mathcal{F}_{M,a}\colon \mathbb{R} \to \mathsf{M}(3\times 3;\mathbb{R})$ *mit* $\varphi \to \mathcal{F}_{M,a}(\varphi) := \mathsf{B}\,(a,\ \varphi) * M$, *wobei* $a \in \mathbb{R}^3$ *und* $M \in \mathsf{M}(3\times 3;\mathbb{R})$ *beliebig aber fest sind. Dann ist die Kurve* $\mathcal{F}_{M,a}(\varphi)$ *eine* $\mathsf{C}^1(\mathbb{R})$*-Funktion und für den Tangentialvektor der Kurve gilt*

$$\mathsf{D}\mathcal{F}_{M,a}(\varphi) = \mathsf{D}_\varphi\,(\mathsf{B}\,(a,\ \varphi) * M) := \mathcal{I}_{\mathsf{M,V}}\,(\mathcal{I}_{\mathsf{V,M}}\,[\mathsf{D}_\varphi\mathsf{B}\,(a,\ \varphi)] * M)\ \in \mathbb{R}^9.$$

Beweis:

Die Komponentenfunktionen $\mathsf{B}\,(a,\ \varphi)_{ij},\ i,j = 1,2,3$ der Matrix $\mathsf{B}\,(a,\ \varphi)$ sind in $\varphi \in \mathbb{R}$ stetig differenzierbar bzw. in $a \in \mathbb{R}^3$ stetig partiell differenzierbar. Damit ist $\mathcal{I}_{\mathsf{M,V}}(\mathsf{B}\,(a,\ \varphi))$ eine $\mathsf{C}^1(\mathbb{R})$- bzw. $\mathsf{C}^1(\mathbb{R}^3)$-Funktion. Nach Satz C.17 ist $\mathcal{I}_{\mathsf{M,V}}(\mathsf{B}\,(a,\ \varphi))$ in $\varphi \in \mathbb{R}$ bzw. in $a \in \mathbb{R}^3$

differenzierbar. Das Differential ergibt sich nach Satz C.15 b), indem man die partiellen Ableitungen der Komponentenfunktionen von $\mathsf{B}(a,\ \varphi)$ bzgl. a und φ bildet. Damit ist a) und b) bewiesen. Da aber auch $\sum_{i=1}^{3}\ \mathsf{B}(a,\ \varphi)_{ki}\, m_{il},\ k,l = 1,2,3$ in φ stetig differenzierbar sind, folgt analog c). ■

Lemma 2.4.21

Es sei $\mathcal{M}\colon \mathbb{R} \to \mathsf{M}(3\times 3;\mathbb{R})$, mit $\varphi \to \mathcal{M}(\varphi) \in \mathsf{M}(3\times 3;\mathbb{R})$, eine $\mathsf{C}^1(\mathbb{R})$-Funktion. Weiter sei $x \in \mathbb{R}^3$ beliebig, aber fest. Dann ist auch die Kurve $\mathcal{F}_x\colon \mathbb{R}\to\mathbb{R}^3$, mit $\varphi \to \mathcal{F}_x(\varphi) := \mathcal{M}(\varphi)x \in \mathbb{R}^3$ eine $\mathsf{C}^1(\mathbb{R})$-Funktion und der Tagentialvektor der Kurve ist durch

$$\mathsf{D}\mathcal{F}_x(\varphi) = \mathsf{D}_\varphi\left(\mathcal{M}(\varphi)x\right) := \mathcal{I}_{\mathsf{V},\mathsf{M}}\left[\mathsf{D}_\varphi\mathcal{M}(\varphi)\right]\ x\ \in \mathbb{R}^3.$$

gegeben.

Beweis:

Nach Voraussetzung sind die Komponentenfunktionen $m_{ij}(\varphi),\ i,j = 1,2,3,$ von $\mathcal{M}(\varphi)$ in $\varphi \in \mathbb{R}$ stetig differenzierbar. Dann sind aber auch die Komponentenfunktionen $(\mathcal{F}_x(\varphi))_k = \sum_{j=1}^{3}\ m_{kj}(\varphi)x_j,\ k = 1,2,3,$ von $\mathcal{F}_x(\varphi)$ in $\varphi\in\mathbb{R}$ stetig differenzierbar. Damit ist $\mathcal{I}_{\mathsf{M},\mathsf{V}}(\mathcal{F}_x(\varphi))$ eine $\mathsf{C}^1(\mathbb{R})$-Funktion. Nach Satz C.17 ist $\mathcal{I}_{\mathsf{M},\mathsf{V}}(\mathcal{F}_x(\varphi))$ in $\varphi\in\mathbb{R}$ differenzierbar. Das Differential ergibt sich, indem man die Ableitungen der Komponentenfunktionen von $\mathcal{F}_x(\varphi)$ bzgl. φ bildet.
■

Um nun die Differenzierbarkeit des lokalen Rotationsoperators $\mathcal{R}[\mathcal{S}(V,\vec{T})]$ herzuleiten, zeigt man zunächst die Differenzierbarkeit der Abbildung $\mathcal{B}^n(\dots),\ n\geq 1$ bzgl. der Rotationswinkel φ_i. Dazu braucht man noch die folgenden Hilfssätze.

Hilfssatz 2.4.22

$A\ ,B\colon \mathbb{R}\to\mathsf{M}(3\times 3;\mathbb{R})$ seien $\mathsf{C}^1(\mathbb{R})$-Funktionen mit

$$A(\varphi) := \begin{pmatrix} a_{11}(\varphi) & a_{12}(\varphi) & a_{13}(\varphi)\\ a_{21}(\varphi) & a_{22}(\varphi) & a_{23}(\varphi)\\ a_{31}(\varphi) & a_{32}(\varphi) & a_{33}(\varphi)\end{pmatrix},\ B(\varphi) := \begin{pmatrix} b_{11}(\varphi) & b_{12}(\varphi) & b_{13}(\varphi)\\ b_{21}(\varphi) & b_{22}(\varphi) & b_{23}(\varphi)\\ b_{31}(\varphi) & b_{32}(\varphi) & b_{33}(\varphi)\end{pmatrix},$$

d.h. die Komponentenfunktionen $a_{ij}(\varphi), b_{kl}(\varphi)\colon \mathbb{R}\to\mathbb{R},\ i,j,k,l = 1,2,3,$ sind in $\varphi\in\mathbb{R}$ stetig differenzierbar. Dann ist auch die Kurve $C:\mathbb{R}\to\mathsf{M}(3\times 3;\mathbb{R})$, mit $C(\varphi) := A(\varphi) * B(\varphi)$, eine $\mathsf{C}^1(\mathbb{R})$-Funktion und der Tagentialvektor der Kurve ist durch

$$\mathsf{D}C(\varphi) := \mathcal{I}_{\mathsf{M},\mathsf{V}}\left(\mathcal{I}_{\mathsf{V},\mathsf{M}}\left[\mathsf{D}A(\varphi)\right] * B(\varphi)\ +\ A(\varphi) * \mathcal{I}_{\mathsf{V},\mathsf{M}}\left[\mathsf{D}B(\varphi)\right]\right)\ \in \mathbb{R}^9$$

gegeben.

Beweis:

Es gilt $c_{kl}(\varphi) = \left(\sum_{i=1}^{3}\ a_{ki}(\varphi)b_{il}(\varphi)\right),\ k,l = 1,2,3.$ Nach Voraussetzung sind die Komponentenfunktionen $a_{ki}(\varphi)$ und $b_{il}(\varphi)$ stetig differenzierbar. Mit der Produktregel für gewöhnliche Ableitungen folgt nun, daß auch die Komponentenfunktionen

$$c'_{kl}(\varphi) = \left(\sum_{i=1}^{3}\ a'_{ki}(\varphi)\ b_{il}(\varphi)\right)\ +\ \left(\sum_{i=1}^{3}\ a_{ki}(\varphi)\ b'_{il}(\varphi)\right),\ k,l = 1,2,3 \qquad (2.15)$$

in φ stetig differenzierbar sind. Damit ist $C(\varphi)$ eine $\mathsf{C}^1(\mathbb{R})$-Funktion und aus Satz C.17 folgt, daß $C(\varphi)$ in $\varphi \in \mathbb{R}$ differenzierbar ist. Die Darstellung des Tagentialvektors der Kurve folgt aus Gleichung (2.15). ■

Hilfssatz 2.4.23 (Kettenregel für C^1-Funktionen)
Es sei $g\colon U \subset \mathbb{R}^n \to \mathbb{R}^m$ (U offen) eine $\mathsf{C}^1(U)$-Funktion und $f\colon V \subset \mathbb{R}^m \to \mathbb{R}^r$ (V offen) eine $\mathsf{C}^1(V)$-Funktion mit $g(U) \subset V$. Dann ist auch $h := f \circ g\colon U \to \mathbb{R}^r$, mit $h(x) = f(g(x))$, $x \in U$, eine $\mathsf{C}^1(U)$-Funktion. Das Differential von h ist durch $\mathsf{D}h(x) = \mathsf{D}f(g(x)) * \mathsf{D}g(x)$ gegeben.

Beweis:
Nach Satz C.17 ist g in $x \in U$ und f in $g(x) \in V$ differenzierbar. Nach Satz C.19 ist dann h in $x \in U$ differenzierbar und es ist $\mathsf{D}h(x) = \mathsf{D}f(g(x)) * \mathsf{D}g(x) \in \mathsf{M}(r \times n; \mathbb{R})$. Nach Voraussetzung sind die Komponentenfunktionen $g_j\colon U \to \mathbb{R}$, $j = 1, \dots, m$, von g bzgl. $x \in U$ und $f_k\colon V \to \mathbb{R}$, $k = 1, \dots, r$, von f bzgl. $y \in V$ stetig partiell differenzierbar. Es ist noch zu zeigen, daß die Komponentenfunktionen $h_k\colon U \to \mathbb{R}$, $k = 1, \dots, r$, von h bzgl. $x \in U$ stetig partiell differenzierbar sind. Dabei gilt für $k = 1, \dots, r$

$$h_k(x_1, \dots, x_n) = f_k\left(g_1(x_1, \dots, x_n), \dots, g_m(x_1, \dots, x_n)\right)$$

Aus der Darstellung der Funktionalmatrix $\mathsf{D}h(x)$ erhält man

$$\frac{\partial}{\partial x_i} h_k(x) = \sum_{j=1}^{m} \frac{\partial}{\partial y_j} f_k\left(g(x)\right) \; \frac{\partial}{\partial x_i} g_j(x)$$

Da die partiellen Ableitungen von f_k, $k = 1, \dots, r$, in $g(x) = y \in V$ und von g_j, $j = 1, \dots, m$, in $x \in U$ stetig sind, sind auch die partiellen Ableitungen von h_k, $k = 1, \dots, r$, in $x \in U$ stetig. Daraus folgt, daß auch h_k bzgl. $x \in U$ stetig partiell differenzierbar ist. ■

Satz 2.4.24
Es sei $\mathcal{B}^n(\dots)$, $n \geq 1$ die Abbildung aus Definition 2.4.10. Es seien $a_i \in \mathbb{R}^3 \setminus \{0\}$, $i = 1, \dots, n$, beliebig, aber fest. Dann ist Abbildung $\mathcal{B}^n(a_1, \dots, a_n,\ \varphi_1, \dots, \varphi_n) : \mathbb{R}^n \to \mathsf{SO}_3$ für alle $n \geq 1$ eine $\mathsf{C}^1(\mathbb{R}^n)$-Funktion. Nach Satz C.17 ist dann $\mathcal{B}^n(a_1, \dots, a_n,\ \varphi_1, \dots, \varphi_n)$ in $(\varphi_1, \dots, \varphi_n) \in \mathbb{R}^n$ differenzierbar. Für n=1 ist das Differential durch

$$\mathsf{D}_{\varphi_1}\mathcal{B}^1(a_1, \varphi_1) = \mathsf{D}_{\varphi_1}\mathsf{B}\left(\frac{a_1}{\|a_1\|_2},\ \varphi_1\right) \ \in \mathbb{R}^9$$

gegeben. Mit $\mathcal{A}_n := (a_1, \dots, a_n)$ und $\Phi_n := (\varphi_1, \dots, \varphi_n)$ gilt für n>1

$$\mathsf{D}_{\varphi_1, \dots, \varphi_n}\ \mathcal{B}^n(\mathcal{A}_n, \Phi_n) = \left(\mathsf{D}_{\varphi_1}\mathcal{B}^n(\dots)\ \dots\ \mathsf{D}_{\varphi_n}\mathcal{B}^n(\dots)\right) \ \in \mathsf{M}(9 \times n; \mathbb{R}),$$

wobei für $i = 1, \dots, n-1$

$$\begin{aligned} \mathsf{D}_{\varphi_i}\mathcal{B}^n(\mathcal{A}_n, \Phi_n) = \mathcal{I}_{\mathsf{M},\mathsf{V}} \Bigg\{ & \mathcal{I}_{\mathsf{V},\mathsf{M}} \left[\mathsf{D}_a \mathsf{B} \left(\mathcal{B}^{n-1}(\mathcal{A}_{n-1}, \Phi_{n-1}) \frac{a_n}{\|a_n\|_2},\ \varphi_n \right) \right. \\ & \left. * \left(\mathcal{I}_{\mathsf{V},\mathsf{M}} \left[\mathsf{D}_{\varphi_i}\mathcal{B}^{n-1}(\mathcal{A}_{n-1}, \Phi_{n-1}) \right] \frac{a_n}{\|a_n\|_2} \right) \right] * \mathcal{B}^{n-1}(\mathcal{A}_{n-1}, \Phi_{n-1}) \ + \\ & \mathsf{B}\left(\mathcal{B}^{n-1}(\mathcal{A}_{n-1}, \Phi_{n-1}) \frac{a_n}{\|a_n\|_2},\ \varphi_n \right) * \mathcal{I}_{\mathsf{V},\mathsf{M}} \left[\mathsf{D}_{\varphi_i}\mathcal{B}^{n-1}(\mathcal{A}_{n-1}, \Phi_{n-1}) \right] \Bigg\}, \end{aligned} \tag{2.16}$$

und

$$\mathsf{D}_{\varphi_n}\mathcal{B}^n\left(\mathcal{A}_n,\Phi_n\right)=\mathcal{I}_{\mathsf{M,V}}\left(\mathcal{I}_{\mathsf{V,M}}\left[\mathsf{D}_{\varphi_n}\mathsf{B}\left(\mathcal{B}^{n-1}\left(\mathcal{A}_{n-1},\Phi_{n-1}\right)\frac{a_n}{\|a_n\|_2},\ \varphi_n\right)\right]*\mathcal{B}^{n-1}\left(\mathcal{A}_{n-1},\Phi_{n-1}\right)\right) \tag{2.17}$$

gilt.

Beweis:
Der Beweis, daß $\mathcal{B}^n(\dots)$ für $n\geq 1$ eine $\mathsf{C}^1(\mathbb{R}^n)$-Funktion ist, erfolgt mittels Induktion nach n.
Induktionsanfang: Für $n=1$ folgt die Behauptung aus Lemma 2.4.20 a).
Induktionsbehauptung: Für $1\leq k<n$ ist $\mathcal{B}^k\left(a_1,\dots,a_k,\ \varphi_1,\dots,\varphi_k\right)$ bzgl. $(\varphi_1,\dots,\varphi_k)\in\mathbb{R}^k$ eine $\mathsf{C}^1(\mathbb{R}^k)$-Funktion. Nach Satz C.17 ist dann $\mathcal{B}^{n-1}\left(a_1,\dots,a_k,\varphi_1,\dots,\varphi_k\right)$ in $(\varphi_1,\dots,\varphi_k)\in\mathbb{R}^k$ differenzierbar. Das Differential $\mathsf{D}_{\varphi_1,\dots,\varphi_k}\mathcal{B}^k\left(a_1,\dots,a_k,\ \varphi_1,\dots,\varphi_k\right)$ hat die obige Darstellung.
Induktionsschritt: $n-1\ \to\ n$
Für $n>1$ sei $\mathcal{A}_n:=(a_1,\dots,a_n)$ und $\Phi_n:=(\varphi_1,\dots,\varphi_n)$. Nach Definition von $\mathcal{B}^n(\dots)$ gilt

$$\mathcal{B}^n\left(\mathcal{A}_n,\Phi_n\right)=\mathsf{B}\left(\mathcal{B}^{n-1}\left(\mathcal{A}_{n-1},\Phi_{n-1}\right)\ \frac{a_n}{\|a_n\|_2},\ \varphi_n\right)*\mathcal{B}^{n-1}\left(\mathcal{A}_{n-1},\Phi_{n-1}\right).$$

Aus Lemma 2.4.20 c) folgt, daß $\mathcal{B}^n\left(\mathcal{A}_n,\Phi_n\right)$ bzgl. φ_n eine $\mathsf{C}^1(\mathbb{R})$-Funktion ist und das Differential $\mathsf{D}_{\varphi_n}\mathcal{B}^n\left(\mathcal{A}_n,\Phi_n\right)$ die obige Darstellung (Gleichung (2.17)) besitzt.
Nach Induktionsbehauptung ist $\mathcal{B}^{n-1}\left(\mathcal{A}_{n-1},\Phi_{n-1}\right)$ bzgl. $(\varphi_1,\dots,\varphi_{n-1})\in\mathbb{R}^{n-1}$ eine $\mathsf{C}^1(\mathbb{R}^{n-1})$-Funktion. Insbesondere ist $\mathcal{B}^{n-1}\left(a_1,\dots,a_{n-1},\varphi_1,\dots,\varphi_{n-1}\right)$ bzgl. $\varphi_i,\ i=1,\dots,n-1$ eine $\mathsf{C}^1(\mathbb{R})$-Funktion und damit in $\varphi_i\in\mathbb{R}$ differenzierbar.
Es sei $\mathcal{F}_{\varphi_n}(a)\colon\mathbb{R}^3\to\mathsf{M}(3\times 3;\mathbb{R})$, mit $\mathcal{F}_{\varphi_n}(a):=\mathsf{B}\left(a,\ \varphi_n\right)$ und für $i=1\dots,n-1$ sei

$$\mathcal{F}^i_{\mathcal{A}_n,\Phi_{n-1}}\colon\mathbb{R}\to\mathbb{R}^3\setminus\{0\},\ \text{mit}\ \mathcal{F}^i_{\mathcal{A}_n,\Phi_{n-1}}(\varphi_i):=\mathcal{B}^{n-1}\left(\mathcal{A}_{n-1},\varphi_1,\dots,\varphi_i,\dots,\varphi_{n-1}\right)\frac{a_n}{\|a_n\|_2}.$$

Nach Lemma 2.4.20 b) ist $\mathcal{F}_{\varphi_n}(a)$ eine $\mathsf{C}^1(\mathbb{R}^3)$-Funktion und damit in $a\in\mathbb{R}^3$ differenzierbar. Für $i=1,\dots,n-1$ ist nach Lemma 2.4.21 $\mathcal{F}^i_{\mathcal{A}_n,\Phi_{n-1}}(\varphi_i)$ bzgl. φ_i eine $\mathsf{C}^1(\mathbb{R})$-Funktion, da $\mathcal{B}^{n-1}\left(\mathcal{A}_{n-1},\Phi_{n-1}\right)$ bzgl. φ_i eine $\mathsf{C}^1(\mathbb{R})$-Funktion ist. Für $i=1,\dots,n-1$ ist nach Hilfssatz 2.4.23 die Kurve

$$\begin{aligned}
\mathcal{G}^i_{\mathcal{A}_n,\Phi_n}&:=\mathcal{F}_{\varphi_n}\circ\mathcal{F}^i_{\mathcal{A}_n,\Phi_{n-1}}\colon\mathbb{R}\to\mathsf{M}(3\times 3;\mathbb{R}),\\
\mathcal{G}^i_{\mathcal{A}_n,\Phi_n}(\varphi_i)&=\mathcal{F}_{\varphi_n}\left(\mathcal{F}^i_{\mathcal{A}_n,\Phi_{n-1}}(\varphi_i)\right):=\mathsf{B}\left(\mathcal{B}^{n-1}\left(\mathcal{A}_{n-1},\varphi_1,\dots,\varphi_i,\dots,\varphi_{n-1}\right)\frac{a_n}{\|a_n\|_2},\ \varphi_n\right)
\end{aligned}$$

bzgl. $\varphi_i\in\mathbb{R}$ eine $\mathsf{C}^1(\mathbb{R})$-Funktion. Der Tagentialvektor der Kurve ist durch

$$\begin{aligned}
\mathsf{D}_{\varphi_i}\mathcal{G}^i_{\mathcal{A}_n,\Phi_n}(\varphi_i)&\overset{\mathsf{HS.}\ 2.4.23}{=}\mathsf{D}_a\mathcal{F}_{\varphi_n}\left(\mathcal{F}^i_{\mathcal{A}_n,\Phi_n}(\varphi_i)\right)*\mathsf{D}_{\varphi_i}\mathcal{F}^i_{\mathcal{A}_n,\Phi_n}(\varphi_i)\\
&\overset{\mathsf{L.}\ 2.4.21}{=}\ \mathsf{D}_a\mathsf{B}\left(\mathcal{B}^{n-1}\left(\mathcal{A}_{n-1},\Phi_{n-1}\right)\frac{a_n}{\|a_n\|_2},\ \varphi_n\right)\\
&\quad*\ \left(\mathcal{I}_{\mathsf{V,M}}\left[\mathsf{D}_{\varphi_i}\mathcal{B}^{n-1}\left(\mathcal{A}_{n-1},\Phi_{n-1}\right)\right]\frac{a_n}{\|a_n\|_2}\right)\in\mathbb{R}^9
\end{aligned}$$

gegeben. Damit ist nach Hilfssatz 2.4.22 für $i = 1, \dots, n-1$ das Produkt $\mathcal{G}^i_{\mathcal{A}_n,\Phi_n}(\varphi_i) * \mathcal{B}^{n-1}(\mathcal{A}_{n-1}, \Phi_{n-1})$ bzgl. $\varphi_i \in \mathbb{R}$ eine $\mathsf{C}^1(\mathbb{R})$-Funktion und es ist

$$\begin{aligned}\mathsf{D}_\varphi\left(\mathcal{G}^i_{\mathcal{A}_n,\Phi_n}(\varphi_i) * \mathcal{B}^{n-1}(\mathcal{A}_{n-1}, \Phi_{n-1})\right) = &\mathcal{I}_{\mathsf{M,V}}\left\{\mathcal{I}_{\mathsf{V,M}}\left(\mathsf{D}\mathcal{G}^i_{\mathcal{A}_n,\Phi_n}(\varphi_i)\right) * \mathcal{B}^{n-1}(\mathcal{A}_{n-1}, \Phi_{n-1})\right.\\ &\left. + \ \mathcal{G}^i_{\mathcal{A}_n,\Phi_n}(\varphi_i) * \mathcal{I}_{\mathsf{V,M}}\left(\mathsf{D}_{\varphi_i}\mathcal{B}^{n-1}(\mathcal{A}_{n-1}, \Phi_{n-1})\right)\right\} \ \in \mathbb{R}^9,\end{aligned}$$

was Gleichung (2.16) entspricht. Insgesamt folgt, daß $\mathcal{B}^n(\mathcal{A}_n, \Phi_n)$ bzgl. $\varphi_i \in \mathbb{R}$, $i = 1, \dots, n$, stetig (partiell) differenzierbar ist. Damit ist $\mathcal{B}^n(\mathcal{A}_n, \Phi_n)$ bzgl. $(\varphi_1, \dots, \varphi_n)$ eine $\mathsf{C}^1(\mathbb{R}^n)$-Funktion.
■

Lemma 2.4.25
Mit den Bezeichnungnen aus Satz 2.4.16 und Definition 2.4.18 ist der lokale Rotationsoperator von V

$$\mathcal{R}_V := \mathcal{R}[\mathcal{S}(V, \vec{T})]: \ \mathbb{R}^n \times V \ \rightarrow \ \mathbb{R}^3,$$

$$\mathcal{R}_V(\varphi_1, \dots, \varphi_n, x) := \begin{cases} \mathsf{R}_V[f_0](x) & , \ \Psi_{V,F}(x) = f_0 \\ \mathsf{R}_V[f_i]\left(\varphi_{k_1}, \dots, \varphi_{k_{l_i-1}}, \varphi_i, x\right) & , \ \Psi_{V,F}(x) = f_i \neq f_0 \end{cases},$$

bzgl. $(\varphi_1, \dots, \varphi_n) \in \mathbb{R}^n$ eine $\mathsf{C}^1(\mathbb{R}^n)$-Funktion. Dabei ist Dabei ist $\mathsf{R}_V[f_i] := \mathsf{R}[\mathcal{S}(V, \vec{T}), f_i]$ für $f_i \in F$ gesetzt worden. Für das Differential von $\mathcal{R}_V$ gilt:

1) Für alle $x \in V$ mit $\Psi_{V,F}(x) = f_0$ gilt: $\mathsf{D}_{\varphi_i}\mathsf{R}_V[f_0](x) = (0,0,0)^\mathsf{T} \in \mathbb{R}^3$ für $i = 1, \dots, n$. Damit ist das Differential von $\mathcal{R}_V$ in $x \in V_0$ die Nullmatrix.

2) Für $f_i \neq f_0$ sei $\vec{P}_{f_0,f_i} = (f_0 f_{k_1} \dots f_{k_{l_i-1}} f_i) \subset \vec{T}$ der gerichtete f_0-f_i−Weg in $\vec{T}$ mit der Länge $l_i := l(\vec{P}_{f_0,f_i})$. Weiter sei $(d_{k_1}, \dots, d_{k_{l_i-1}}, d_i) \subset \mathbb{R}^3 \setminus 0$ bzw. $(\varphi_{k_1}, \dots, \varphi_{k_{l_i-1}}, \varphi_i) \subset \mathbb{R}$ die dem f_0-f_i−Weg eindeutig zugeordnete Folge von lokalen Drehachsen bzw. Rotationsparameter. Setzt man $\Phi[i] := \left(\varphi_{k_1}, \dots, \varphi_{k_{l_i-1}}, \varphi_i\right)$, so gilt für $x \in (\Psi_{V,F})^{-1}(f_i)$:

$$\mathsf{D}_{\varphi_p}\mathcal{R}_V(\varphi_1, \dots, \varphi_n, x) = \mathsf{D}_{\varphi_p}\mathsf{R}_V[f_i](\Phi[i], x) \in \mathbb{R}^3,$$

wobei $\mathsf{D}_{\varphi_p}\mathsf{R}_V[f_i]$ wie folgt gegeben ist: a) falls $\varphi_p \notin \{\varphi_{k_1}, \dots, \varphi_{k_{l_i-1}}, \varphi_i\}$, so gilt $\mathsf{D}_{\varphi_p}\mathsf{R}_V[f_i](\Phi[i], x) = (0,0,0)^\mathsf{T}$;
b) falls $\varphi_p = \varphi_i$, so gilt

$$\mathsf{D}_{\varphi_p}\mathsf{R}_V[f_i](\Phi[i], x) = \mathcal{I}_{\mathsf{V,M}}\left(\mathsf{D}_{\varphi_i}\mathcal{B}^{l_i}\left(d_{k_1}, \dots, d_{k_{l_i-1}}, d_i, \Phi[i]\right)\right) x;$$

c) falls $\varphi_p = \varphi_{k_m} \in \{\varphi_{k_1}, \dots, \varphi_{k_{l_i-1}}\}$, so gilt

$$\begin{aligned}\mathsf{D}_{\varphi_{k_m}}\mathsf{R}_V[f_i](\Phi[i], x) = &\sum_{j=m}^{l_i-1} \mathcal{I}_{\mathsf{V,M}}\left(\mathsf{D}_{\varphi_{k_m}}\mathcal{B}^j(d_{k_1}, \dots, d_{k_j}, \varphi_{k_1}, \dots, \varphi_{k_j})\right) y_j \\ &+ \ \mathcal{I}_{\mathsf{V,M}}\left(\mathsf{D}_{\varphi_{k_m}}\mathcal{B}^{l_i}\left(d_{k_1}, \dots, d_{k_{l_i-1}}, d_i, \Phi[i]\right)\right) y_{l_i},\end{aligned}$$

wobei $y_j := \mathsf{R}_{F,V}(f_{k_{j+1}}) - \mathsf{R}_{F,V}(f_{k_j})$, $j = m, \dots, l_i - 2$, $y_{l_i-1} := \mathsf{R}_{F,V}(f_i) - \mathsf{R}_{F,V}(f_{k_{l_i-1}})$ und $y_{l_i} := x - \mathsf{R}_{F,V}(f_i)$ ist.

Beweis:
$\mathsf{R}_V[f_0]$ ist bzgl. der Rotationsparameter konstant. Für $\varphi_p \notin \{\varphi_{k_1}, \ldots, \varphi_{k_{l_i-1}}, \varphi_i\}$ ist die Rotationsabbildung $\mathsf{R}_V[f_i]$, $f_i \neq f_0$ konstant. Nach Korollar 2.4.17 kann man $\mathsf{R}[\mathcal{S}(V,\vec{T}), f_i]\,(\Phi[i], x)$ als Summe von Vektoren darstellen. Aus Satz 2.4.24 und Lemma 2.4.21 folgt, daß die Komponentenfunktionen dieser Vektoren bzgl. $\varphi_p \in \mathbb{R}$, $p = 1, \ldots, n$, stetig partiell differenzierbar sind. Die Summe von stetig partiell differenzierbaren Funktionen ist stetig partiell differenzierbar. Es folgt, daß die Komponentenfunktionen von $\mathsf{R}_V[f_i]$ bzgl. den Rotationsparametern $\varphi_p \in \mathbb{R}$, $p = 1, \ldots, n$, stetig partiell differenzierbar ist. Also ist $\mathsf{R}_V[f_i]$ für alle $f_i \neq f_0$ eine $\mathsf{C}^1(\mathbb{R}^n)$-Funktion und daraus folgt die Behauptung.
Die Darstellungen in Punkt 1. und 2. a) sind klar. Die Darstellungen in 2. b) und c) folgen aus Satz 2.4.24 mit Korollar 2.4.17 und Lemma 2.4.21. ∎

Die folgende Abbildung spielt bei der Betrachtung der Energiefunktionen in Kapitel 3.2 und 3.3 eine große Rolle.

Lemma 2.4.26
Mit den Bezeichnungen aus Lemma 2.4.25 sei $\mathsf{R}_V[f_i] := \mathsf{R}[\mathcal{S}(V,\vec{T}), f_i]$ die zu $f_i \in F$ und $\mathcal{S}(V,\vec{T})$ gehörende Rotationsabbildung aus Satz 2.4.16. Weiter sei $\varphi = (\varphi_1, \ldots, \varphi_n) \in \mathbb{R}^n$. Dann ist für $i, j = 0, 1, \ldots, n$, $i \neq j$, die Abbildung $\mathsf{N}_{ij} : \mathbb{R}^n \mathrm{x} V_i \mathrm{x} V_j \to \mathbb{R}$ mit

$$\mathsf{N}_{ij}\,(\varphi, v_1, v_2) := \|\mathsf{R}_V[f_i]\,(\Phi[i], v_1) - \mathsf{R}_V[f_j]\,(\Phi[j], v_2)\|_2^2,$$

wobei $\Phi[i] := (\varphi_{k_1}, \ldots, \varphi_{k_i-1}, \varphi_i)$ und $\Phi[j] := \left(\varphi_{l_1}, \ldots, \varphi_{l_j-1}, \varphi_j\right)$ gesetzt worden ist, in φ_p, $p = 1, \ldots n$, stetig partiell differenzierbar. Im folgenden sei $\mathsf{P}_{k,3}(x) = x_k, \forall x \in \mathbb{R}^3$ die Projektion auf die k-te Komponente. Die Komponenten der Gradienten sind durch

$$\begin{aligned} \frac{\partial}{\partial \varphi_p}\,\mathsf{N}_{ij}\,(\varphi, v_1, v_2) \;=\; 2\,\sum_{k=1}^{3} &\{\mathsf{P}_{k,3}\,(\mathsf{R}_V[f_i]\,(\Phi[i], v_1) - \mathsf{R}_V[f_j]\,(\Phi[j], v_2)) \cdot \\ &\mathsf{P}_{k,3}\left(\mathsf{D}_{\varphi_p}\mathsf{R}_V[f_i]\,(\Phi[i], v_1) - \mathsf{D}_{\varphi_p}\mathsf{R}_V[f_j]\,(\Phi[j], v_2)\right)\},\ p = 1, \ldots, n, \end{aligned}$$

für $i, j = 0, \ldots, n$, mit $i \neq j$ gegeben.

Beweis:
Die Abbildung $\mathsf{N}^*(x) := \|x\|_2^2$, ist in $x \in \mathbb{R}^3$ stetig partiell differenzierbar, d.h. $\mathsf{N}^* \in \mathsf{C}^1(\mathbb{R}^3)$, und es gilt $\frac{\partial}{\partial x_i}\mathsf{N}^*(x) = 2\ \mathsf{P}_{i,3}(x)$, $i = 1, 2, 3$.
Dann gilt

$$\mathsf{N}_{ij}\,(\varphi, v_1, v_2) \;=\; \mathsf{N}^*\,(\mathsf{R}_V[f_i]\,(\Phi[i], v_1) - \mathsf{R}_V[f_j]\,(\Phi[j], v_2))\,.$$

Nach Lemma 2.4.25 sind die Komponentenfunktionen von $\mathsf{R}_V[f_i]\,(\Phi[i], v_1) - \mathsf{R}_V[f_j]\,(\Phi[j], v_2)$ für alle $v_1 \in V_i$ und alle $v_2 \in V_j$, $i, j = 0, \ldots, n$, in φ_p, $p = 1, \ldots, n$ stetig partiell differenzierbar. Also sind diese Abbildungen $\mathsf{C}^1(\mathbb{R}^n)$-Funktionen. Nach Hilfsatz 2.4.23 ist dann N_{ij} für $i, j = 0, \ldots, n$, eine $\mathsf{C}^1(\mathbb{R}^n)$-Funktion. Für für den Gradienten gilt:

$$\begin{aligned} \nabla_\varphi \mathsf{N}_{ij}\,(\varphi, v_1, v_2) \;=\; &\nabla_x \mathsf{N}^*\,(\mathsf{R}_V[f_i]\,(\Phi[i], v_1) - \mathsf{R}_V[f_j]\,(\Phi[j], v_2)) * \\ &\Big(\mathsf{D}_{\varphi_1}\mathsf{R}_V[f_i]\,(\Phi[i], v_1) - \mathsf{D}_{\varphi_1}\mathsf{R}_V[f_j]\,(\Phi[j], v_2) \quad \ldots \quad \mathsf{D}_{\varphi_n}\mathsf{R}_V[f_i]\,(\Phi[i], v_1) - \mathsf{D}_{\varphi_n}\mathsf{R}_V[f_j]\,(\Phi[j], v_2)\Big) \\ &\in \mathsf{M}(1 \times n; \mathbb{R}). \end{aligned}$$

■

Bemerkung: Aus Lemma 2.4.25 1) und 2a) folgt das einige Komponenten des Gradienten aus Lemma 2.4.26 null sein werden.

2.4.3 Der Transformationsoperator

Nachdem nun die globale Transformation und die Hintereinanderschaltung der lokalen Rotationen definiert worden sind, läßt sich nun der Transformationsoperator für eine Menge $V = \bigcup_{i=0}^{n} V_i \subset \mathbb{R}^3$, wobei $V_i \neq \emptyset$, $i = 1, \ldots, n$, und $V_i \cap V_j = \emptyset$, für $i \neq j$, gilt, mittels einer Transformationsstruktur $\mathcal{S}(V, \vec{T})$ definieren. Danach wird der Transformationsoperator auf eine Molekülstruktur übertragen. Im folgenden sei $U_g := \mathbb{R}^3 \times \mathbb{R}^4 \setminus \{0\} \times \mathbb{R}^3$.

Definition 2.4.27 (Transformationsoperator)
Es sei $V \subset \mathbb{R}^3$. Weiter sei $\bigcup_{i=0}^{n} V_i$, $V_i \neq \emptyset$, $n \geq 1$ eine disjunkte Zerlegung von V. Es sei $\vec{T} = (F, \vec{E})$ ein gerichteter Baum mit $F = \{f_0, \ldots, f_n\}$, $\vec{E} \subseteq F \times F$ und der Wurzel f_0. Für $f_i \in F \setminus \{f_0\}$ sei $\vec{P}_{f_0,f_i} = (f_0 f_{k_1} \ldots f_{k_i-1} f_i)$ der eindeutig bestimmte f_0-f_i-Weg in $\vec{T}$. Die Abbildung $\Psi_{V,F}\colon V \to F$ sei wie in Definition 2.4.13. Die Abbildungen $\mathsf{R}_{F,V}\colon F \setminus \{f_0\} \to V$, $\mathsf{S}_{F,V} \setminus \{f_0\}\colon F \to V$ und $\mathsf{D}_{F,\mathbb{R}^3\setminus\{0\}}\colon F \setminus \{f_0\} \to \mathbb{R}^3 \setminus \{0\}$ seien wie in Definition 2.4.14. Dann ist $\mathcal{S}(V, \vec{T})$ eine Transformationsstruktur auf V (Definition 2.4.14). Die lokalen Drehachsen sind durch $d_k := \mathsf{D}_{F,\mathbb{R}^3\setminus\{0\}}(f_k) = \mathsf{S}_{F,V}(f_k) - \mathsf{R}_{F,V}(f_k) \in \mathbb{R}^3 \setminus \{0\}$, $f_k \in F \setminus \{f_0\}$, definiert. Die Abbildung $\mathcal{R}[\mathcal{S}(V, \vec{T})]$ sei wie in Definition 2.4.18 und die Abbildung T_g sei wie in Definition 2.4.7. Weiter sei $U_n := U_g \times \mathbb{R}^n$, U_n offen. Dann heißt die Abbildung

$$\begin{aligned} &\mathcal{T}[\mathcal{S}(V,\vec{T})]\colon U_n \times V \to \mathbb{R}^3, \\ &(t,\ q,\ c,\ \varphi_1,\ \ldots,\ \varphi_n,\ v) \ \to \mathcal{T}[\mathcal{S}(V,\vec{T})]\,(t,\ q,\ c,\ \varphi_1,\ \ldots,\ \varphi_n,\ v) \end{aligned}$$

mit $u = (t, q, c, \varphi_1, \ldots, \varphi_n) \in U_n$ und

$$\begin{aligned} \mathcal{T}[\mathcal{S}(V,\vec{T})]\,(u,v) &:= \mathsf{T}_g\left(t, q, c, \mathcal{R}[\mathcal{S}(V,\vec{T})]\,(\varphi_1,\ \ldots,\ \varphi_n,\ v)\right) \\ &= \begin{cases} A(q)\,(v-c)+c+t & , \Psi_{V,F}(v) = f_0 \\ A(q)\left(\mathsf{R}[\mathcal{S}(V,\vec{T}), f_i]\,(\varphi_{k_1}, \ldots, \varphi_{k_i-1}, \varphi_i, v) - c\right) + c + t & , \Psi_{V,F}(v) = f_i \neq f_0. \end{cases} \end{aligned}$$

Transformationsoperator von V bzgl. der Transformationsstruktur $\mathcal{S}(V, \vec{T})$. ◇

Die Eigenschaften des Transformationsoperator sind im folgenden Lemma zusammengefaßt.

Lemma 2.4.28
Mit den Bezeichnungen aus Definition 2.4.27 sei $\mathcal{T}_V := \mathcal{T}[\mathcal{S}(V, \vec{T})]$ der Transformationsoperator von V bzgl. $\mathcal{D}_V(\vec{T})$. Weiter sei $u = (t, q, c, \varphi) \in U_n$ mit $\varphi = (\varphi_1, \ldots, \varphi_n) \in \mathbb{R}^n$.

a) Für $u \in U_n$ sei $u_\lambda := (t, \lambda q, c, \varphi) \in U_n$ mit $\lambda \in \mathbb{R}$, $\lambda \neq 0$. Für $u \in U_n$, $x \in V$ beliebig aber fest, gilt $\mathcal{T}_V(u_\lambda, x) = \mathcal{T}_V(u, x)\ \forall \lambda \neq 0$.

b) Es seien $x,\ y \in V$ mit $\Psi_{V,F}(x) = f_i$ und $\Psi_{V,F}(y) = f_j$. Dann gilt für alle $u \in U_n$

$$\|\mathcal{T}_V\,(u,x) - \mathcal{T}_V\,(u,y)\|_2 = \begin{cases} \|x-y\|_2 & f_i = f_j \\ \|\mathsf{R}_V[f_i]\,(\Phi[i],x) - \mathsf{R}_V[f_j]\,(\Phi[j],y)\|_2 & f_i \neq f_j \end{cases}$$

mit $\mathsf{R}_V[f_k] := \mathsf{R}[\mathcal{S}(V,\vec{T}), f_k]$ für $f_k \in F$, $\Phi[i] := (\varphi_{k_1}, \dots, \varphi_{k_i-1}, \varphi_i)$ und $\Phi[j] := \left(\varphi_{l_1}, \dots, \varphi_{l_j-1}, \varphi_j\right)$

c) Mit den Bezeichnungnen aus Satz 2.4.16 gilt für alle $u \in U_n$ und für alle $x \in V$ mit $\Psi_{V,F}(x) = f_i \neq f_0$:

$$\mathcal{T}_V\,(u,x) = \mathsf{T}_g\,(t,q,c,\mathsf{R}_{F,V}(f_{k_1})) + \sum_{p=1}^{l_i-1} A(q) * \mathcal{B}^p\left(d_{k_1},\dots,d_{k_p},\varphi_{k_1},\dots,\varphi_{k_p}\right) y_p$$
$$+\ A(q)*\mathcal{B}^{l_i}\left(d_{k_1},\dots,d_{k_{l_i-1}},d_i,\varphi_{k_1},\dots,\varphi_{k_{l_i-1}},\varphi_i\right)(x - \mathsf{R}_{F,V}(f_i))\,,$$

mit $l_i := l(\vec{P}_{f_0,f_i})$, $\vec{P}_{f_0,f_i} = (f_0, f_{k_1}, \dots, f_{k_{l_i-1}}, f_i) \subset \vec{T}$, $y_p := \mathsf{R}_{F,V}(f_{k_{p+1}}) - \mathsf{R}_{F,V}(f_{k_p})$, $p = 1,\dots,l_i-2$, und $y_{l_i-1} := \mathsf{R}_{F,V}(f_i) - \mathsf{R}_{F,V}(f_{k_{l_i-1}})$.

d) Es sei $u(c) = (t,q,c,\varphi) \in U_n$ mit $t = (0,0,0)$, $q = (0,0,0,1)$, $\varphi = (0,\dots,0)$ und $c \in \mathbb{R}^3$. Dann gilt $\mathcal{T}_V\,(u(c),x) = x$ für alle $x \in V$ und $c \in \mathbb{R}^3$.

Beweis:

a) folgt unmittelbar aus Lemma 2.4.9 a).

b) Aus Lemma 2.4.9 b) folgt

$$\|\mathcal{T}_V\,(u,x) - \mathcal{T}_V\,(u,y)\|_2 = \|\mathsf{R}_V[f_i]\,(\Phi[i],x) - \mathsf{R}_V[f_j]\,(\Phi[j],y)\|_2.$$

Wegen $\mathcal{B}^{l_i}(\dots) \in \mathsf{SO}_3,\ l_i := l(\vec{P}_{f_0,f_i})$, gilt für $f_i = f_j$ mit Satz C.8

$$\|\mathsf{R}_V[f_i]\,(\Phi[i],x) - \mathsf{R}_V[f_i]\,(\Phi[i],y)\|_2 = \|x-y\|_2$$

c) folgt direkt aus Korollar 2.4.17.

d) folgt unmittelbar aus Lemma 2.4.9 c) und Lemma 2.4.19 angewendet auf die Definition von $\mathcal{T}_V = \mathcal{T}[\mathcal{S}(V,\vec{T})]$. ■

Bemerkung: Aus Lemma 2.4.28 b) wird deutlich warum, man zuerst die internen (lokalen) Rotationen ausführt und erst dann die globale Trannsformation anwendet. Denn mit diesem Lemma sind sowohl das n-Potential (Definition 2.6.2) als auch der Nenner des Hodgkin-Index (angewendet auf Summen von Gaussfunktionen) (Lemma 2.5.5) unabhängig von den Variablen t,q,c der globalen Transformationsabbildung T_g aus Definition 2.4.7. Die genauen Einzelheiten sind in Kapitel 3.2 dargestellt.

Die folgende Abbildung spielt bei der Betrachtung der Energiefunktionen in Kapitel 3.2 und 3.3 eine große Rolle.

Lemma 2.4.29
Mit den Bezeichnungen aus Definition 2.4.27 sei $\mathcal{T}_V := \mathcal{T}[\mathcal{S}(V,\vec{T})]$ der Transformationsoperator von V bzgl. $\mathcal{D}_V(\vec{T})$. $\mathcal{R}_V := \mathcal{R}[\mathcal{S}(V,\vec{T})]$ ist der zu $\mathcal{T}_V$ gehörende lokale Rotationsoperator aus Definition 2.4.18. Weiter sei $u = (t,q,c,\varphi) \in U_n$ mit $\varphi = (\varphi_1,\dots,\varphi_n) \in \mathbb{R}^n$.

a) *Es sei $y \in \mathbb{R}^3$ beliebig aber fest. Dann ist die Abbildung $\mathsf{N}_y\colon U_n \mathsf{x} V \to \mathbb{R}$, mit*

$$\mathsf{N}_y\,(u,v) := \|\mathcal{T}_V\,(u,v) - y\|_2^2$$

in $u \in U_n$ stetig partiell differenzierbar. Im folgenden sei $\mathsf{P}_{i,3}(x) = x_i,\ \forall x \in \mathbb{R}^3$ die Projektion auf die i-te Komponente. Dann gilt

$$\frac{\partial}{\partial u_k}\mathsf{N}_y\,(u,v) = \frac{\partial}{\partial t_k}\mathsf{N}_y\,(u,v) = \ 2\ \mathsf{P}_{k,3}\,(\mathcal{T}_V\,(u,v) - y)\,, \quad k = 1,2,3.$$

$$\begin{aligned}\frac{\partial}{\partial u_{3+k}}\mathsf{N}_y\,(u,v) = \frac{\partial}{\partial q_k}\mathsf{N}_y\,(u,v) = &\\ 2\sum_{i=1}^{3}\Bigg[\mathsf{P}_{i,3}\,(\mathcal{T}_V\,(u,v) - y)\left\{\frac{1}{\|q\|_2^2}\left(\sum_{l=1}^{3}\left[\left(\frac{\partial}{\partial q_k}R(q)_{il}\right)\mathsf{P}_{l,3}\,(\mathcal{R}_V(\varphi,v) - c)\right]\right.\right.\\ \left.\left. - \,2q_k\ \mathsf{P}_{i,3}\,(A(q)(\mathcal{R}_V(\varphi,v) - c))\right)\right\}\Bigg]\,, \quad k = 1,2,3,4,\end{aligned}$$

wobei $R(q)$ die Matrix aus Gleichung (2.9) ist. Die Komponenten $\frac{\partial}{\partial q_k}R(q)_{il}$, $k = 1,2,3,4$, $i,l = 1,2,3$, entnimmt man der Matrix aus Lemma 2.4.6 6.).

$$\frac{\partial}{\partial u_{7+k}}\mathsf{N}_y\,(u,v) = \frac{\partial}{\partial c_k}\mathsf{N}_y\,(u,v) = 2\sum_{i=1}^{3}\left[\mathsf{P}_{i,3}\,(\mathcal{T}_V\,(u,v) - y)\,(\delta_{ik} - A(q)_{ik})\right]\,, \quad k = 1,2,3,$$

wobei δ_{ik} das Kroneckersymbol ist.

$$\begin{aligned}\frac{\partial}{\partial u_{10+k}}\mathsf{N}_y\,(u,v) \ &= \ \frac{\partial}{\partial \varphi_k}\mathsf{N}_y\,(u,v)\\ &= \ 2\sum_{i=1}^{3}\left[\mathsf{P}_{i,3}\,(\mathcal{T}_V\,(u,v) - y)\,\mathsf{P}_{i,3}\,(A(q)\mathsf{D}_{\varphi_k}\mathsf{R}_V[\Psi_{V,F}(v)](\Phi[k],v))\right] \quad k = 1,\ldots,n,\end{aligned}$$

wobei $\mathsf{D}_{\varphi_k}\mathsf{R}_V[\Psi_{V,F}(v)](\Phi[k],v)$ das Differential aus Lemma 2.4.25 ist und $\Phi[k] := (\varphi_{l_1},\ldots,\varphi_{l_k-1},\varphi_k)$ gesetzt worden ist.

b) *Für $u \in U_n$ sei $u_\lambda := (t,\lambda q,c,\varphi) \in U_n$ mit $\lambda \neq 0$. Für $u \in U_n$, $v \in V$ beliebig aber fest, gilt $\mathsf{N}_y(u_\lambda,v) = \mathsf{N}_y(u,v)\ \forall\lambda \neq 0$. Insbesondere gilt sogar*

$$\nabla_u\mathsf{N}_y(u_\lambda,v) = \nabla_u\mathsf{N}_y(u,v) \quad \forall\lambda \neq 0$$

Beweis:

a) Es sei $y \in \mathbb{R}^3$ beliebig aber fest. Dann ist die Abbildung $\mathsf{N}_y^*(x) := \|x - y\|_2^2$ in $x \in \mathbb{R}^3$ stetig partiell differenzierbar, d.h. $\mathsf{N}_y^* \in \mathsf{C}^1(\mathbb{R}^3)$, und es gilt $\frac{\partial}{\partial x_i}\mathsf{N}_y^*(x) = 2\ \mathsf{P}_{i,3}(x - y)$ für $i = 1,2,3$. Dann gilt

$$\mathsf{N}_y(u,v) \ = \ \mathsf{N}_y^*\,(\mathcal{T}_V\,(u,v)) \ = \ \mathsf{N}_y^*\,(\mathsf{T}_g\,(t,q,c,\mathcal{R}_V(\varphi,v)))\,.$$

Dabei ist T_g die globale Transformationsabbildung aus Definition 2.4.7. Nach Satz 2.4.8 sind die Komponentenfunktionen von T_g in $(t,q,c) \in U_g$ stetig partiell differenzierbar. Damit ist T_g eine $\mathsf{C}^1(U_g)$-Funktion. Nach Hilfsatz 2.4.23 ist N_y in (t,q,c) stetig partiell differenzierbar. Der Gradient von $\mathsf{N}_y(u,v)$ bzgl. der Variablen t,q,c ist dann durch

$$\nabla_{t,q,c}\mathsf{N}_y(u,v) \ = \ \nabla_x\mathsf{N}_y^*\,(\mathcal{T}_V\,(u,v)) * \mathsf{D}_{t,q,c}\mathsf{T}_g\,(t,q,c,\mathcal{R}_V(\varphi,v)) \in \mathsf{M}(1 \times 10;\mathbb{R})$$

gegeben.

Nach Lemma 2.4.25 sind die Komponentenfunktionen von $\mathcal{R}_V(\varphi, v)$ in $\varphi \in \mathbb{R}^n$ stetig partiell differenzierbar. Also ist $\mathcal{R}_V$ eine $\mathsf{C}^1(\mathbb{R}^n)$-Funktion. Nach Hilfsatz 2.4.23 ist N_y in φ stetig partiell differenzierbar. Der Gradient von $\mathsf{N}_y(u,v)$ bzgl. der Variablen φ ist dann durch

$$\nabla_\varphi \mathsf{N}_y(u,v) \;=\; \nabla_x \mathsf{N}_y^* \left(\mathcal{T}_V\,(u,v)\right) * \mathsf{D}_\varphi\left(A(q)\mathcal{R}_V(\varphi,v)\right) \in \mathsf{M}(1 \times n; \mathbb{R})$$

gegeben.

b) Die 1. Behauptung folgt aus Lemma 2.4.28 a). Wegen

$$\nabla_u \mathsf{N}_y(u,v) \;=\; \nabla_x \mathsf{N}_y^* \left(\mathcal{T}_V\,(u,v)\right) * \Big(\mathsf{D}_{t,q,c}\mathsf{T}_g\left(t,q,c,\mathcal{R}_V(\varphi,v)\right) \quad \mathsf{D}_\varphi\left(A(q)\mathcal{R}_V(\varphi,v)\right)\Big)$$

folgt die 2.Behauptung aus 2.4.28 a) und Lemma 2.4.9 a). ■

Nun wird der obige Transforamtionsoperator auf eine Molekülstruktur übertragen. Dazu sei $\mathcal{A}$ ein Molekül mit n_a-Atomen.[25] Weiter sei $\mathsf{C}_\mathcal{A} \subset \mathbb{R}^{3n_a}$ eine zulässige Konformation von $\mathcal{A}$. Dann ist $K_\mathcal{A} := K(\mathsf{C}_\mathcal{A}) \subset \mathbb{R}^3$ die Menge der Atomkoordinaten bzgl. der Konformation $\mathsf{C}_\mathcal{A}$. Weiter sei $B_{FT}^\mathcal{A}$ die Menge der Bindungen des Moleküls, an denen der Torsionswinkel variiert werden kann. Dann sei $\vec{T}_\mathcal{A} = (F_\mathcal{A}, \vec{E}_{FT}^\mathcal{A})$ der zum Molekül $\mathcal{A}$ gehörende gerichtete Fragmentebaum mit der Wurzel $f_0 \in F_\mathcal{A}$ und $\mathcal{S}(\mathcal{A}, \vec{T}_\mathcal{A})$ eine Transformationsstruktur auf $K_\mathcal{A}$ bzgl. $\vec{T}_\mathcal{A}$. Mit der Abbildung $\Psi_{K_\mathcal{A},F_\mathcal{A}}$ ist eine disjunkte Zerlegung der Menge $K_\mathcal{A}$ gegeben. Denn nach Definition dieser Abbildung gilt $K_\mathcal{A} = \bigcup_{i=0}^{n_{ft}} V_i$ mit $V_i := (\Psi_{K_\mathcal{A},F_\mathcal{A}})^{-1}(f_i) \neq \emptyset$, $i = 0, \ldots, n_{ft}$, $n_{ft} := \left|B_{FT}^\mathcal{A}\right|$, und $V_i \cap V_j = \emptyset$, für $i \neq j$. Für $e_i = \mathcal{I}_{F,B}^\mathcal{A}(f_i) \in B_{FT}^\mathcal{A}$, $i = 1, \ldots, n_{ft}$, ist $d_i := \mathsf{D}_{F_\mathcal{A},\mathbb{R}^3\setminus\{0\}}(f_i)$ die zu Bindung e_i gehörende Drehachse und φ_i sei der entsprechende interne Rotationsparameter. Für $f_i \in F_\mathcal{A} \setminus \{f_0\}$ sei $\vec{P}_{f_0,f_i} = (f_0 f_{k_1} \ldots f_{k_i-1} f_i)$ der eindeutig bestimmte f_0-f_i–Weg in $\vec{T}_\mathcal{A}$. Weiter sei die Abbildung $\mathcal{R}_\mathcal{A} := \mathcal{R}[\mathcal{S}(K_\mathcal{A}, \vec{T}_\mathcal{A})]$ wie in Definition 2.4.18 und die Abbildung T_g wie in Definition 2.4.7. Man setzt $U_{n_{ft}} := \mathbb{R}^3 \times \mathbb{R}^4 \setminus \{0\} \times \mathbb{R}^3 \times \mathbb{R}^{n_{ft}}$. $U_{n_{ft}}$ ist offen. Dann läßt sich der Transformationsoperator von $K_\mathcal{A}$ bzgl. der Transformationsstruktur $\mathcal{S}(\mathcal{A}, \vec{T}_\mathcal{A})$ wie folgt definieren

$$\begin{aligned}
&\mathcal{T}_\mathcal{A} := \mathcal{T}[\mathcal{S}(K_\mathcal{A}, \vec{T}_\mathcal{A})] \colon U_{n_{ft}} \times K_\mathcal{A} \to K_\mathcal{A}^{neu}, \\
&\qquad (t,\ q,\ c,\ \varphi_1,\ \ldots,\ \varphi_n,\ a_p) \;\to \mathcal{T}_\mathcal{A}\,(t,\ q,\ c,\ \varphi_1,\ \ldots,\ \varphi_n,\ a_p)\,,\ p = 1, \ldots, n_a,
\end{aligned}$$

mit $u = (t,\ q,\ c,\ \varphi_1,\ \ldots,\ \varphi_n) \in U_{n_{ft}}$ und

$$\begin{aligned}
\mathcal{T}_\mathcal{A}\,(u, a_p) &:= \mathsf{T}_g\left(t, q, c, \mathcal{R}_\mathcal{A}\,(\varphi_1,\ \ldots,\ \varphi_n,\ a_p)\right) \\
&= \begin{cases} A(q)\,(v - c) + c + t & \Psi_{K_\mathcal{A},F_\mathcal{A}}(a_p) = f_0 \\ A(q)\,\Big(\mathsf{R}[\mathcal{S}(K_\mathcal{A}, \vec{T}_\mathcal{A}), f_i]\,(\varphi_{k_1}, \ldots, \varphi_{k_i-1}, \varphi_i, a_p) - c\Big) + c + t & \Psi_{K_\mathcal{A},F_\mathcal{A}}(a_p) = f_i \neq f_0 \end{cases},
\end{aligned}$$

wobei $K_\mathcal{A}^{neu}$ die Menge der transformierten Atomkoordinaten ist.

[25] Es werden die Bezeichnungen aus Kapitel 2.2 verwendet.

2.5 Gaussfunktionen

Wie in den Kapitel 2.1 und 2.2 bereits erwähnt worden ist, besteht der Ansatz, der in dieser Arbeit verwendet wird, darin, den physikochemischen Eigenschaften eine räumlich lokalisierte Dichte[26] zu geben. Dabei wird die räumlich lokalisierte Dichte bzgl. einer Eigenschaft $q \in \mathcal{I}_{PE}$[27] durch eine Summe von Gaussfunktionen dargestellt. Daher werden in diesem Kapitel Gaussfunktionen und deren Eigenschaften untersucht. Insbesondere wird gezeigt werden, daß der Hodgkin-Index[28] und damit auch der Alignmentindex[29] auf Gaussfunktionen bzw. Gausssummen[30] anwendbar ist. Der Alignmentindex ist, wenn er auf Gausssummen angewendet wird, analytisch berechenbar.

Doch zuvor werden die Gaussfunktionen auf ihre Eigenschaften untersucht:

Definition 2.5.1 (Gaussfunktion)
Eine Gaussfunktion ist eine beschränkte Funktion der Gestalt

$$g\colon \mathbb{R}^3 \;\rightarrow\; \mathbb{R} \textit{ mit } g(x) \;:=\; h\;\exp(-b\;\|x-z\|_2^2),$$

wobei $h \in \mathbb{R}$ *die Höhe,* $b \in \mathbb{R}_{\geq 0}$ *die Breite und* $z \in \mathbb{R}^3$ *das Zentrum der Gaussfunktion ist. Insbesondere ist durch das Tripel* $(h,b,z) \in \mathbb{R} \times \mathbb{R}_{\geq 0} \times \mathbb{R}^3$ *eine Gaussfunktion eindeutig charakterisiert. Es sei* $U := \mathbb{R} \times \mathbb{R}_{\geq 0} \times \mathbb{R}^3$. *Dann erhält man folgende injektive Abbildung*

$$\begin{aligned} \mathsf{G}\colon U &\;\rightarrow\; \mathsf{C}^b(\mathbb{R}^3) \cap \mathsf{C}^\infty(\mathbb{R}^3) \\ (h,b,z) &\;\rightarrow\; \mathsf{G}(h,b,z)(x) \;:=\; h\;\exp(-b\;\|x-z\|_2^2). \end{aligned}$$

◇

Als nächstes betrachtet man die Eigenschaften der Abbildung G.

Lemma 2.5.2 (Eigenschaften der Gaussfunktionen)
Es sei $U := \mathbb{R} \times \mathbb{R}_{\geq 0} \times \mathbb{R}^3$. *Weiter sei* $\mathsf{G}\colon U \rightarrow \mathsf{C}^b(\mathbb{R}^3)$ *mit* $\mathsf{G}(h,b,z)(x) := h\;\exp(-b\;\|x-z\|_2^2)$. *Dann gilt*

a) Für $(h,b,z) \in U$ *gilt* $|\mathsf{G}(h,b,z)(x)| \;\leq\; |h|$.

b) Für $(h,b,z) \in \mathbb{R} \times \mathbb{R}_{>0} \times \mathbb{R}^3$ *gilt*

$$\int\limits_{\mathbb{R}^3} \mathsf{G}(h,b,z)(x)\;\mathsf{d}x \;=\; h\;\left(\frac{\pi}{b}\right)^{\frac{3}{2}}$$

c) Für $(h_1,b_1,z_1),(h_2,b_2,z_2) \in U$, *mit* $b_1+b_2>0$ *gilt*

$$<\mathsf{G}(h_1,b_1,z_1),\;\mathsf{G}(h_2,b_2,z_2)>_{\mathsf{L}^2(\mathbb{R}^3)} \;=\; h_1 h_2\;\left(\frac{\pi}{b_1+b_2}\right)^{\frac{3}{2}}\;\exp\left(\frac{-b_1 b_2}{b_1+b_2}\;\|z_1-z_2\|_2^2\right)$$

[26] Siehe Definition 2.2.2.
[27] siehe Definition 2.2.1
[28] Siehe Definition 2.3.1 a).
[29] siehe Definition 2.3.3
[30] Summe von Gaussfunktionen.

d) Für $(h, b, z) \in \mathbb{R} \mathsf{x} \mathbb{R}_{>0} \mathsf{x} \mathbb{R}^3$ gilt

$$\|\mathsf{G}(h,b,z)(x)\|_{\mathsf{L}^2(\mathbb{R}^3)} \;=\; h \;\left(\frac{\pi}{2b}\right)^{\frac{3}{4}}.$$

Beweis:

a) Klar.

b) Es gilt

$$\int\limits_{-\infty}^{\infty} \exp\left(-bx^2\right)\; \mathsf{d}x \;=\; \left(\frac{\pi}{b}\right)^{\frac{1}{2}} \tag{$*$}$$

Denn

$$\left(\int\limits_{-\infty}^{\infty} \exp\left(-bx^2\right)\; \mathsf{d}x\right)^2 \;=\; \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} \exp\left(-b(x^2+y^2)\;\right)\; \mathsf{d}x\mathsf{d}y$$

Transformationsformel: $\Phi:\; \mathbb{R}_{\geq 0} \mathsf{x}\; [0;\; 2\pi) \;\to \mathbb{R}^2$ mit $(r,\; \varphi) \;\to\; (r\cos\varphi,\; r\sin\varphi) \;= (x,\; y)$

$$\stackrel{C.7}{=} \int\limits_0^{\infty}\int\limits_0^{2\pi} \exp\left(-br^2\right)\; r\; \mathsf{d}\varphi \mathsf{d}r \;=\; 2\pi \lim_{z\to\infty} \left[-\frac{1}{2b}\; \exp\left(-br^2\right)\right]_0^z \;=\; \frac{\pi}{b}$$

Damit erhält man

$$\begin{aligned}\int\limits_{\mathbb{R}^3} \mathsf{G}(h,b,z)(x)\; \mathsf{d}x \;&= h \int\limits_{\mathbb{R}^3} \exp\left(-b\|x-z\|_2^2\right)\; \mathsf{d}x \;\stackrel{O.E.\; z=0}{=}\; h \int\limits_{\mathbb{R}^3} \exp\left(-b\|x\|_2^2\right)\; \mathsf{d}x \\ &= h \prod_{i=1}^{3} \left(\int\limits_{-\infty}^{\infty} \exp\left(-bx_i^2\right)\; \mathsf{d}x_i\right) \;\stackrel{*}{=}\; h\; \left(\frac{\pi}{b}\right)^{\frac{3}{2}}\end{aligned}$$

c)

$$< \mathsf{G}(h_1,b_1,z_1),\; \mathsf{G}(h_2,b_2,z_2) >_{\mathsf{L}^2(\mathbb{R}^3)} = \int\limits_{\mathbb{R}^3} \mathsf{G}(h_1,b_1,z_1)(x) * \mathsf{G}(h_2,b_2,z_2)\; \mathsf{d}x$$

$$\begin{aligned} &= h_1h_2 \int\limits_{\mathbb{R}^3} \exp\left(-b_1\|x-z_1\|_2^2 \;-\; b_2\|x-z_2\|_2^2\right) \mathsf{d}x \\ &= h_1h_2\; \exp\left(-\frac{b_1b_2}{b_1+b_2}\|z_1-z_2\|_2^2\right) \int\limits_{\mathbb{R}^2} \exp\left(-(b_1+b_2)\|x-\frac{b_1z_1+b_2z_2}{b_1+b_2}\|_2^2\right) \mathsf{d}x \\ &\stackrel{b)}{=} h_1h_2\; \exp\left(-\frac{b_1b_2}{b_1+b_2}\|z_1-z_2\|_2^2\right)\; \left(\frac{\pi}{b_1+b_2}\right)^{\frac{3}{2}}\end{aligned}$$

d) Folgt aus c): ∎

Bemerkung: Insbesondere bildet also G den Raum $\mathbb{R} \mathsf{x} \mathbb{R}_{>0} \mathsf{x} \mathbb{R}^3$ in den Hilbertraum $\mathsf{L}^2(\mathbb{R}^3)$ ab.

Ein weiterer wichtiger Punkt ist das Verhalten der Gaussfunktionen, wenn sich das Gausszentrum ändert. Der Alignmentindex zweier Moleküle wird dadurch optimiert, indem die Konformationen und damit auch die Gausszentren der Modellfunktionen durch globale und lokale Transformationen optimiert werden. Also ist es insbesondere wichtig, ob die Gaussfunktionen bzgl. des Gausszentrums differenzierbar sind.

Lemma 2.5.3
Es sei G *wie in Lemma 2.5.2 definiert. Weiter sei* $(h,b) \in \mathbb{R}\mathrm{x}\mathbb{R}_{>0}$ *beliebig, aber fest. Dann ist die Gaussfunktion* $\mathsf{G}(h,b,\cdot)\colon \mathbb{R}^3 \to \mathsf{C}^b(\mathbb{R}^3)$ *in* $z \in \mathbb{R}^3$ *differenzierbar und die Fréchetsche Ableitung an der Stelle* z *lautet*

$$\mathsf{D}_z\mathsf{G}(h,b,z) \;=\; -2bh < z-x,\ \cdot >_2\ \exp\left(-b\|z-x\|_2^2\right). \tag{2.18}$$

Beweis:
Formales ableiten nach $z_k,\ k = 1,2,3$ ergibt

$$\frac{\partial}{\partial z_k}\mathsf{G}(h,b,z) \;=\; -2bh(z_k - x_k)\ \exp\left(-b\|z-x\|_2^2\right).$$

Nun zeigt man, daß die so erhaltende Abbildung mit dem Differential aus der Definition C.20 bzw. C.21 übereinstimmt. Zunächst zeigt man, daß $\mathsf{D}_z\mathsf{G}(h,b,z)$ eine lineare und stetige Abbildung ist.

$$\begin{aligned}\mathsf{D}_z\mathsf{G}(h,b,z)\colon \mathbb{R}^3 &\to \mathsf{C}^b(\mathbb{R}^3)\\ y &\to \mathsf{D}_z\mathsf{G}(h,b,z)y = -2bh < z-x,\ y >_2\ \exp\left(-b\|z-x\|_2^2\right)\end{aligned}$$

Es gilt mit Lemma 2.5.2 a) und der Cauchy-Schwarzsche-Ungleichung C.4

$$\|\mathsf{D}_z\mathsf{G}(h,b,z)y\| \le 2b\,|h|\,\|z-x\|_2\|y\|_2 = K_{z,x}\|y\|_2 \quad \forall y \in \mathbb{R}^3,$$

mit $K_{z,x} := 2b\,|h|\,\|z-x\|_2$. Da das Sklarprodukt in der 2.Komponente linear ist und die Abbildung $\mathsf{D}_z\mathsf{G}(h,b,z)$ beschränkt ist, ist $\mathsf{D}_z\mathsf{G}(h,b,z)$ eine lineare und nach Satz C.6 eine stetige Abbildung. Wendet man nun die Definition C.20 an, so erhält zunächst

$$\begin{aligned}\mathsf{G}(h,b,z+\theta) \;-\; \mathsf{G}(h,b,z) \;&=\; h\exp\left(-b\|(z-x)+\theta\|_2^2\right) \;-\; h\exp\left(-b\|z-x\|_2^2\right)\\ &=\; \underbrace{\exp\left(-b\|z-x\|_2^2\right)}_{=:\ a}\ \left\{h\exp\left(-b\left(\|\theta\|_2^2 \;+\; 2 < z-x,\ \theta >\right)\right) \;-\; h\right\} =: R(z,x,\theta)\end{aligned}$$

Dann gilt

$$\begin{aligned}R(z,x,\theta) \;-\; &\mathsf{D}_z\mathsf{G}(h,b,z)\theta\\ &=\; a\left\{\underbrace{h\exp\left(-b(\|\theta\|_2^2 \;+\; 2 < z-x,\theta >_2)\right)}_{=\sum_{k=0}^{\infty}\frac{\left(-b(\|\theta\|_2^2\ +\ 2<z-x,\theta>_2)\right)^k}{k!}} - h \;+\; 2bh\ < x-z,\theta >_2\right\}\\ &=\; -abh\ \|\theta\|_2^2 \;+\; ah\sum_{k=2}^{\infty}\frac{\left(-b(\|\theta\|_2^2 \;+\; 2 < z-x,\theta >_2)\right)^k}{k!} \;=:\; o(\|\theta\|)\end{aligned}$$

Damit ist $\mathsf{G}(h,b,z)$ in z differenzierbar und das Differential ist durch Gleichung (2.18) gegeben.
■

Da die räumlich lokalisierten Dichten der physikochemischen Eigenschaften durch Gausssummen modelliert werden, stellt sich nun die Frage, wie sieht das Ergebnis aus, wenn der Hodgkin-Index aus Definition 2.3.1 a) auf zwei Gausssummen angewendet wird. Zunächst wird jedoch gezeigt, daß Gausssummen $\mathsf{L}^2(\mathbb{R}^3)$-Funktionen sind.

Lemma 2.5.4

Es sei G *wie in Lemma 2.5.2 definiert. Weiter sei* $k \in \mathbb{N}$ *beliebig aber fest sowie* $(h_i, b_i, z_i) \in \mathbb{R}\mathrm{x}\mathbb{R}_{>0}\mathrm{x}\mathbb{R}^3$ *für* $i = 1, \ldots, k$. *Dann gilt*

$$f(x) \; := \; \sum_{i=1}^{k} \mathsf{G}(h_i, b_i, z_i)(x) \;\in \mathsf{L}^2(\mathbb{R}^3).$$

und

$$\|\sum_{i=1}^{k} \mathsf{G}(h_i, b_i, z_i)(x)\|^2_{\mathsf{L}^2(\mathbb{R}^3)} \; = \; \pi^{\frac{3}{2}} \sum_{i,j=1}^{k} \left(\frac{h_i h_j}{(b_i + b_j)^{\frac{3}{2}}} \exp\left(\frac{-b_i b_j}{b_i + b_j} \|z_i - z_j\|_2^2 \right) \right). \qquad (2.19)$$

Beweis:

Da $\mathsf{L}^2(\mathbb{R}^3)$ ein Vektorraum ist, folgt wegen Lemma 2.5.2 d) und der Minkowskischen Ungleichung (Satz C.5 aus Anhang C), daß $f(x) \in \mathsf{L}^2(\mathbb{R}^3)$ gilt. Die Gleichung (2.19) folgt aus der Bilinearität des Skalarproduktes $< \cdot, \cdot >_{\mathsf{L}^2(\mathbb{R}^3)}$ mit Lemma 2.5.2 c). ■

Lemma 2.5.5

Es sei G *wie in Lemma 2.5.2 und* H *wie in Definition 2.3.1 a) definiert. Weiter seien* $k, l \in \mathbb{N}$ *und* $(h_1^X, b_1^X, z_1^X), \ldots, (h_k^X, b_k^X, z_k^X), (h_1^Y, b_1^Y, z_1^Y), \ldots, (h_l^Y, b_l^Y, z_l^Y) \in \mathbb{R}\mathrm{x}\mathbb{R}_{>0}\mathrm{x}\mathbb{R}^3$.
Dann gilt

$$\mathsf{H}\left(\sum_{i=1}^{k} \mathsf{G}\left(h_i^X, b_i^X, z_i^X\right)(x), \; \sum_{j=1}^{l} \mathsf{G}\left(h_j^Y, b_j^Y, z_j^Y\right)(x) \right) \; =$$

$$\frac{2 \sum_{i=1}^{k} \sum_{j=1}^{l} \left(\frac{h_i^X h_j^Y}{(b_i^X + b_j^Y)^{\frac{3}{2}}} \exp\left(\frac{-b_i^X b_j^Y}{b_i^X + b_j^Y} \|z_i^X - z_j^Y\|_2^2 \right) \right)}{\sum_{i,j=1}^{k} \left(\frac{h_i^X h_j^X}{(b_i^X + b_j^X)^{\frac{3}{2}}} \exp\left(\frac{-b_i^X b_j^X}{b_i^X + b_j^X} \|z_i^X - z_j^X\|_2^2 \right) \right) + \sum_{i,j=1}^{l} \left(\frac{h_i^Y h_j^Y}{(b_i^Y + b_j^Y)^{\frac{3}{2}}} \exp\left(\frac{-b_i^Y b_j^Y}{b_i^Y + b_j^Y} \|z_i^Y - z_j^Y\|_2^2 \right) \right)}$$

Beweis:

Folgt aus der Bilinearität des Skalarproduktes $< \cdot, \cdot >_{\mathsf{L}^2(\mathbb{R}^3)}$ mit Lemma 2.5.2 c). ■

Bemerkung: Insbesondere folgt aus Lemma 2.5.5, daß der Hodgkin-Index angewendet auf zwei Gausssummen mit positiven Breiten endlich und analytisch berechenbar ist. Damit ist auch der Alignment-Index, wenn er auf entsprechende Gaussummen angewendet wird, analytisch berechenbar.

Zum Abschluß des Kapitels wird der Aligmentindex bzgl. der physikochemischen Eigenschaften zweier Moleküle betrachtet.

Definition 2.5.6

Es sei $U := \mathbb{R}\mathrm{x}\mathbb{R}_{>0}\mathrm{x}\mathbb{R}^3$. *Weiter seien* $\mathcal{A}$ *und* $\mathcal{B}$ *zwei Moleküle. Für* $q \in \mathcal{I}_{PE}$[31] *ist* $k_{\mathcal{A}}^q \in \mathbb{N}_0$ *bzw.* $k_{\mathcal{B}}^q \in \mathbb{N}_0$ *die Anzahl der Gaussfunktionen, um die räumlich lokalisierte Dichte der physikochemischen Eigenschaft* q *des Moleküls* $\mathcal{A}$ *bzw.* $\mathcal{B}$ *zu modellieren. Die räumlich lokalisierten*

[31] Definition 2.2.1.

Dichten sind für $q \in \mathcal{I}_{PE}$ durch

$$Q^q_{\mathcal{A}}(x) := \begin{cases} \sum_{i=1}^{k^q_{\mathcal{A}}} \mathsf{G}\left(h^{\mathcal{A}}_{q,i}, b^{\mathcal{A}}_{q,i}, z^{\mathcal{A}}_{q,i}\right)(x) & \text{falls } k^q_{\mathcal{A}} > 0 \\ 0 & \text{, sonst,} \end{cases}$$

wobei $(h^{\mathcal{A}}_{q,i}, b^{\mathcal{A}}_{q,i}, z^{\mathcal{A}}_{q,i}) \in U$ für $i = 1, \ldots, k^q_{\mathcal{A}}$, $q \in \mathcal{I}_{PE}$, geeignet gegeben sind, bzw. durch

$$Q^q_{\mathcal{B}}(x) := \begin{cases} \sum_{j=1}^{k^q_{\mathcal{B}}} \mathsf{G}\left(h^{\mathcal{B}}_{q,j}, b^{\mathcal{B}}_{q,j}, z^{\mathcal{B}}_{q,j}\right)(x) & \text{, falls } k^q_{\mathcal{B}} > 0 \\ 0 & \text{, sonst,} \end{cases}$$

wobei $(h^{\mathcal{B}}_{q,j}, b^{\mathcal{B}}_{q,j}, z^{\mathcal{B}}_{q,j}) \in U$ für $j = 1, \ldots, k^q_{\mathcal{B}}$, $q \in \mathcal{I}_{PE}$, ebenfalls geeignet gegeben sind, definiert. Weiter sind für $q \in \mathcal{I}_{PE}$ positive Energiekoeffizienten $\tilde{\lambda}_q \in \mathbb{R}_{>0}$ gegeben, um die einzelnen physikochemischen Eigenschaften im Aligmentindex zu gewichten. Dann setzt man

$$\mathcal{I}^{\mathcal{A},\mathcal{B}}_{PE} := \{q \in \mathcal{I}_{PE} \mid (Q^q_{\mathcal{A}}(x) \neq 0) \ \vee \ (Q^q_{\mathcal{B}}(x) \neq 0)\} \neq \emptyset$$

und $\tilde{\lambda} := \sum_{i \in \mathcal{I}^{\mathcal{A},\mathcal{B}}_{PE}} \tilde{\lambda}_i > 0$. $\mathcal{I}^{\mathcal{A},\mathcal{B}}_{PE}$ ist ungleich der leeren Menge, da wenigstens die räumlich lokalisierte Dichte der Elektronendichte, d.h. die sterische Gestalt, für beide Moleküle gegeben ist. Dann gilt $\sum_{i \in \mathcal{I}^{\mathcal{A},\mathcal{B}}_{PE}} \lambda_i = 1$ mit $\lambda_i := \frac{\tilde{\lambda}_i}{\tilde{\lambda}}$, $i \in \mathcal{I}^{\mathcal{A},\mathcal{B}}_{PE}$. Der Alignmentindex bzgl. des Hodgkin-Indizes für die Moleküle $\mathcal{A}$ und $\mathcal{B}$ ist dann durch

$$\begin{aligned} \mathsf{A}^{\mathsf{H}}_{\mathcal{A},\mathcal{B}}\left((Q^q_{\mathcal{A}}(x))_{q \in \mathcal{I}^{\mathcal{A},\mathcal{B}}_{PE}}, (Q^q_{\mathcal{B}}(x))_{q \in \mathcal{I}^{\mathcal{A},\mathcal{B}}_{PE}}\right) &= \sum_{q \in \mathcal{I}^{\mathcal{A},\mathcal{B}}_{PE}} \lambda_q \ \mathsf{H}(Q^q_{\mathcal{A}}(x), Q^q_{\mathcal{B}}(x)) \\ = \sum_{q \in \mathcal{I}^{\mathcal{A},\mathcal{B}}_{PE}} \lambda_q \ \mathsf{H}\left(\sum_{i=1}^{k^q_{\mathcal{A}}} \mathsf{G}(h^{\mathcal{A}}_{q,i}, b^{\mathcal{A}}_{q,i}, z^{\mathcal{A}}_{q,i})(x), \right. & \left. \sum_{j=1}^{k^q_{\mathcal{B}}} \mathsf{G}(h^{\mathcal{B}}_{q,j}, b^{\mathcal{B}}_{q,j}, z^{\mathcal{B}}_{q,j})(x)\right) \end{aligned}$$

gegeben.

Lemma 2.5.7
Mit den Voraussetzungen und Bezeichnungen aus Definition 2.5.6 gilt

$$\left|\mathsf{A}^{\mathsf{H}}_{\mathcal{A},\mathcal{B}}\left((Q^q_{\mathcal{A}}(x))_{q \in \mathcal{I}^{\mathcal{A},\mathcal{B}}_{PE}}, (Q^q_{\mathcal{B}}(x))_{q \in \mathcal{I}^{\mathcal{A},\mathcal{B}}_{PE}}\right)\right| \leq 1$$

Beweis:
Folgt mit Lemma 2.5.4 direkt aus Lemma 2.3.4. ■

Bemerkung: Der Alignmentindex für die Moleküle $\mathcal{A}$ und $\mathcal{B}$ berechnet den Wert der Überlagerung bzw. des Alignments der beiden Molekülkonformationen mittels der räumlich lokalisierten Dichten der physikochemischen Eigenschaften. Die räumlich lokalisierten Dichten werden durch Gausssummen dargestellt. Die Gausskoordinaten hängen dabei von der entsprechenden Konformationen $\mathsf{C}_{\mathcal{A}}$ bzw. $\mathsf{C}_{\mathcal{B}}$ ab. Je besser die physikochemischen Eigenschaften durch die entsprechenden Folgen $\{(h^{\mathcal{A}}_{q,i}, b^{\mathcal{A}}_{q,i}, z^{\mathcal{A}}_{q,i})\}_{i=1}^{k^q_{\mathcal{A}}} \subset U$ bzw. $\{(h^{\mathcal{B}}_{q,i}, b^{\mathcal{B}}_{q,i}, z^{\mathcal{B}}_{q,i})\}_{i=1}^{k^q_{\mathcal{B}}} \subset U$, $q \in \mathcal{I}_{PE}$, modelliert werden können, um so aussagekräftiger ist dann der Alignmentindex selber. Denn je grösser der Wert des Index ist, um so besser sollte die Überlagerung der Molekülkonformationen sein.

Mit der Überlagerung korrespondiert aber die Ähnlichkeit der beiden Moleküle in ihren räumlichen Eigenschaften. Ein Wert von ungefähr eins würde also bedeuteten, daß die räumlichen Eigenschaften praktisch identisch sind. Ein negativer Wert würde jedoch bedeuten, daß sich die Moleküle zwar überlagern, jedoch entgegengesetzte räumlichen Eigenschaften besitzen (z.B negativ-positiv geladen oder hydrophile-hydrophobe). Dies führt zu der Forderung an die Daten der Gaussfunktionen und an die Kalibrierung der Energieparameter: je grösser der Index ist, um so ähnlicher sollten die räumlichen Eigenschaften sein.
Indem man den Wert des Alignmentindizes vor der Optimierung und danach berechnet, kann man angeben in wie fern eine signifikante Verbesserung des Alignments erreicht worden ist.

2.6 Lennard-Jones-Potential

Wenn Atome einander sehr nahe kommen, so beginnen die Abstoßungskräfte der Kerne und der Elektronen sowie die zunehmende kinetische Energie der Elektronen die Anziehungskräfte zu übersteigen. Mit kleiner werdendem Abstand nehmen die Abstoßungskräfte sehr schnell zu. Ein einfaches Modell, um diese Wechselwirkung von abstoßenden und anziehenden Kräften darzustellen, ist das Lennard-Jones-Potential (siehe hierzu auch [4], [3], [26]), welches eine Fülle von Anwendungen in der Moleküldynamik und der Festkörperphysik hat. Das Lennard-Jones-Potential besteht aus einer anziehenden Komponente $-\left(\frac{\sigma}{r}\right)^m$ und einer abstoßenden Komponente $\left(\frac{\sigma}{r}\right)^n$, die für sehr kleine Abstände r der Atome dominant wird. Mit $m = 6$ (in Übereinstimmung mit der Van-der-Waals Kraft (siehe hierzu [21]) und $n = 12$ erzeugt es eine sehr schnell mit dem Abstand abfallende Kraft.[32] Sehr oft schreibt man das (12,6)-Potential in der Form

$$U_\sigma^\varepsilon(r) \;=\; 4\varepsilon\left(\left(\frac{\sigma}{r}\right)^{12} \;-\; \left(\frac{\sigma}{r}\right)^{6}\right), \quad \sigma,\varepsilon,r>0. \tag{2.20}$$

Das Potential wird durch σ und ε parametrisiert. ε definiert dabei die Tiefe des Potentials und damit die Stärke der Abstoßung- bzw. Anziehungskräfte. σ gibt den Nulldurchgang des Potentials und über $2^{\frac{1}{6}}\sigma$ den Nulldurchgang der aus dem Potential resultierenden Kraft an. In Abbildung 2.9 wird das Lennard-Jones-(12,6)-Potential ($f(r)$) mit seiner abstoßenden und anziehenden Komponente ($g(r)$ und $h(r)$) dargestellt. Die Eigenschaften des (12,6)-Potential werden in dem folgenden Korollar zusammengefaßt.

Korollar 2.6.1
Es sei $\varepsilon, \sigma > 0$ beliebig aber fest. Weiter sei $U_\sigma^\varepsilon(r)$ wie in (2.20). Dann gilt

a) $U_\sigma^\varepsilon(\sigma)\colon \mathbb{R}_{>0} \;\to\; \mathbb{R}$.

b) $\left(U_\sigma^\varepsilon\right)'(r) \;=\; 24\;\varepsilon\;\sigma^6\left(r^{-7} - 2\;\sigma^6 r^{-13}\right)$.

c) $\left(U_\sigma^\varepsilon\right)''(r) \;=\; 24\;\varepsilon\;\sigma^6\left(26\;\sigma^6 r^{-14} - 7\;r^{-8}\right)$.

d) $U_\sigma^\varepsilon(\sigma) \;=\; 0$.

[32] Die Wahl $n = 12$ hat dabei keine fundamentale physikalische Basis, sondern beruht allein auf "mathematischen Einfachheit".

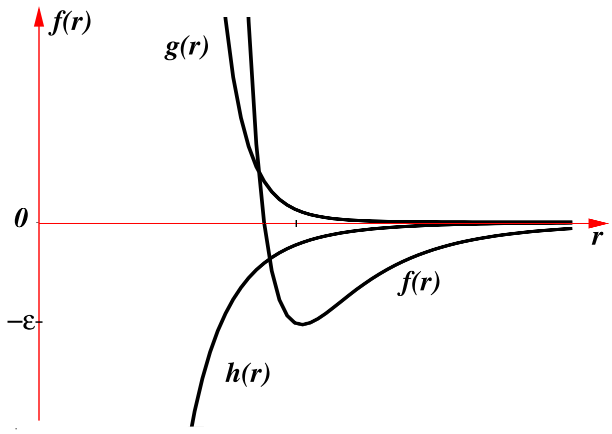


Abbildung 2.9: Das Lennard-Jones-(12,6)-Potential

e) $(U_\sigma^\varepsilon)'(r) \;=\; 0 \;\Leftrightarrow\; r \;=\; 2^{\frac{1}{6}}\sigma.$

f) $U_\sigma^\varepsilon(2^{\frac{1}{6}}\sigma) \;=\; -\varepsilon.$

g) $(U_\sigma^\varepsilon)''(2^{\frac{1}{6}}\sigma) \;=\; 144\,\left(2^{-\frac{4}{3}}\right)\,\varepsilon\,\sigma^{-2} \;>\; 0.$ ■

Bemerkung: Das Lennard-Jones-Potential hat also in $2^{\frac{1}{6}}\sigma$ das einzige Minimum mit dem Funktionswert $-\varepsilon$.

Die aus der Optimierung resultierende Molekülkonformation soll zulässig sein (vergleiche Kapitel 2.1). Daher wird das Lennard-Jones-Potential für alle Atome des Moleküls paarweise verwendet, um diese Nebenbedingung zu erfüllen. Gegeben sei das Molekül $\mathcal{A}$ mit n Atomen. $\mathsf{C}_\mathcal{A}$ sei eine zulässige Konformation des Moleküls $\mathcal{A}$. Dann läßt sich das Potential wie folgt definieren.

Definition 2.6.2 (n-Potential)
Es seien $n \in \mathbb{N}$, und $V_n \;:=\; \{\,(a_1,\dots,a_n) \mid (a_i \in \mathbb{R}^3,\; i=1,\dots,n) \wedge (a_i \neq a_j,\; i \neq j)\,\} \;\subset \mathbb{R}^{3n}$. Dann heißt die Abbildung

$$P_n \;:\; \mathbb{R}_{>0} \times V_n \times \mathbb{R}^n_{>0} \;\rightarrow\; \mathbb{R}$$
$$(\varepsilon, (a_1,\dots,a_n), (\sigma_1,\dots,\sigma_n)) \;\rightarrow\; P_n\,(\varepsilon, (a_1,\dots,a_n), (\sigma_1,\dots,\sigma_n))$$

mit

$$P_n\,(\varepsilon, (a_1,\dots,a_n), (\sigma_1,\dots,\sigma_n)) \;=\; \frac{2}{n(n-1)} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} U^\varepsilon_{\sqrt{\sigma_i\sigma_j}}(\|a_i - a_j\|)$$

$$=\; \frac{8\varepsilon}{n(n-1)} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \left(\left(\frac{\sqrt{\sigma_i\sigma_j}}{\|a_i - a_j\|}\right)^{12} - \left(\frac{\sqrt{\sigma_i\sigma_j}}{\|a_i - a_j\|}\right)^{6} \right)$$

n-Potential zu den Parametern $(\varepsilon, (a_1,\dots,a_n), (\sigma_1,\dots,\sigma_n))$.[33] ◇

[33] $\|.\|$ ist die euklidische Norm des Vektorraumes $\mathbb{R}^3$.

Bemerkung: Die Parameter $\sigma_i,\ i = 1, \dots, n$ sind die zu den Atomen $a_i,\ i = 1, \dots, n$ gehörenden Van-der-Waals-Radien (siehe [41]).[34] Wenn die Molekülkonformation $\mathsf{C}_{\mathcal{A}}$ zulässig ist, so gilt $\mathsf{C}_{\mathcal{A}} \subset V_n$.
Die wichtigsten Eigenschaften der Abbildung P_n sind im nächsten Lemma zusammengefaßt.

Lemma 2.6.3
Es sei $P_n,\ n \in \mathbb{N}$ wie in Definition 2.6.2. Weiter seien $\varepsilon > 0,\ \sigma_1, \dots, \sigma_n > 0$ beliebig aber fest. Dann bildet $P_n\left(\varepsilon, (., \sigma_i)_{i=1}^n\right)$ die offene Menge V_n in das Intervall $[-\varepsilon; \infty)$ ab. Weiter gilt, daß $P_n\left(\varepsilon, (., \dots, .), (\sigma_1, \dots, \sigma_n)\right)$ in $(a_1, \dots, a_n) \in V_n$ zweimal stetig partiell differenzierbar ist.
a) Der Gradient ist gegeben durch

$$\mathsf{D}_{a_1, \dots, a_n}\ P_n\left(\varepsilon, (a_1, \dots, a_n), (\sigma_1, \dots, \sigma_n)\right) =$$

$$\left(\sum_{i=2}^{n} \left(U^{\varepsilon}_{\sqrt{\sigma_1 \sigma_i}}\right)' (\|a_1 - a_i\|)\ \frac{a_1 - a_i}{\|a_1 - a_i\|},\ \dots\ ,\ \sum_{i=1}^{n-1} \left(U^{\varepsilon}_{\sqrt{\sigma_n \sigma_i}}\right)' (\|a_n - a_i\|)\ \frac{a_n - a_i}{\|a_n - a_i\|} \right)^T.$$

b) Die Komponenten der Hesse-Matrix sind gegeben durch

$$\frac{\partial}{\partial a_h^j} \frac{\partial}{\partial a_k^l}\ P_n\left(\varepsilon, (a_1, \dots, a_n), (\sigma_1, \dots, \sigma_n)\right)$$

$$= \frac{(a_h^j - a_k^j)(a_k^l - a_h^l)}{\|a_k - a_h\|^2} \left\{ \left(U^{\varepsilon}_{\sqrt{\sigma_k \sigma_h}}\right)'' (\|a_k - a_h\|)\ -\ \frac{\left(U^{\varepsilon}_{\sqrt{\sigma_k \sigma_h}}\right)' (\|a_k - a_h\|)}{\|a_k - a_h\|} \right\}$$
$$-\ \delta_{jl}\ \frac{\left(U^{\varepsilon}_{\sqrt{\sigma_k \sigma_h}}\right)' (\|a_k - a_h\|)}{\|a_k - a_h\|}$$

für $h,\ k = 1,\ \dots\ , n,\ h \neq k,\ j,\ l = 1,\ 2,\ 3$; und

$$\frac{\partial}{\partial a_k^j} \frac{\partial}{\partial a_k^l}\ P_n\left(\varepsilon, (a_1, \dots, a_n), (\sigma_1, \dots, \sigma_n)\right)$$

$$= \sum_{\substack{i=1 \\ i \neq k}}^{n} \left[\frac{(a_k^j - a_i^j)(a_k^l - a_i^l)}{\|a_k - a_i\|^2} \left\{ \left(U^{\varepsilon}_{\sqrt{\sigma_k \sigma_i}}\right)'' (\|a_k - a_i\|)\ -\ \frac{\left(U^{\varepsilon}_{\sqrt{\sigma_k \sigma_i}}\right)' (\|a_k - a_i\|)}{\|a_k - a_i\|} \right\} \right.$$
$$\left. +\ \delta_{jl} \frac{\left(U^{\varepsilon}_{\sqrt{\sigma_k \sigma_i}}\right)' (\|a_k - a_i\|)}{\|a_k - a_i\|} \right]$$

für $k = 1,\ \dots\ , n,\ j,\ l = 1,\ 2,\ 3$. δ_{jl} ist das Kroneckersymbol.

Beweis:
Das Bild von V_n unter $P_n\left(\varepsilon, (., \dots, .), (\sigma_1, \dots, \sigma_n)\right)$ ergibt sich nach Korollar 2.6.1. Wendet man dann die Kettenregel und Korollar 2.6.1 an, so erhält man für $k = 1, \dots, n$ und $l = 1, 2, 3$

$$\frac{\partial}{\partial a_k^l}\ P_n\left(\varepsilon, (a_1, \dots, a_n), (\sigma_1, \dots, \sigma_n)\right)\ =\ \sum_{\substack{i=1 \\ i \neq k}}^{n} \left(U^{\varepsilon}_{\sqrt{\sigma_k \sigma_i}}\right)' (\|a_k - a_i\|)\ \frac{a_k^l - a_i^l}{\|a_k - a_i\|}.$$

[34] Mit a_i wird sowohl das Atom als auch die Atomkoordinaten bezeichnet.

Wendet man hierauf die Kettenregel, die Produkt- und Quotientenregel der gewöhnlichen Ableitungen aus Satz C.11 an[35], so erhält man die Komponenten der Hessematrix wie sie in b) angegeben sind. ■

2.7 Die Bewertungsfunktion für die Torsionswinkel

Wie in Kapitel 2.2 schon dargelegt worden ist, ist die direkte Bewertung der aktuellen Torsionswinkel ein wichtiger Punkt. Der aktuelle Torsionswinkel ergibt sich aus der Summe des Torsionswinkels, der aus der Molekülkonformation $\mathsf{C}_{\mathcal{A}}$ errechnet worden ist, und des aktuelllen Rotationswinkel (siehe hierzu Gleichung (2.4) in Kapitel 2.2). Daher werden die Torsionswinkel schon indirekt durch das n-Potential[36], das Rotationswinkel unterbinden soll, die zu unzulässigen Molekülkonformationen führen, bewertet. Nun sollen aber die aktuellen Torsionswinkel direkt bewertet werden. Dazu werden Bewertungstabellen verwendet. In Abbildung 2.7 ist eine solche Bewertungstabelle dargestellt. Durch geeignete Interpolation der entsprechenden Energiewerte, die in den Tabellen angegeben sind, soll eine analytische Funktion bestimmt werden, die den entsprechenden aktuellen Torsionswinkel bewertet. Die Summe dieser analytischen Funktionen bewertet dann alle aktuellen Torsionswinkel des Moleküls $\mathcal{A}$. Zur Interpolation werden Spline-Funktionen verwendet. Dazu werden im Kapitel 2.7.1 Spline-Funktionen definiert und ihre wichtigsten Eigenschaften dargestellt, insbesondere wie man Spline-Funktionen mittels Basis-Splines darstellen und auswerten kann. Von besonderem Interesse sind dabei periodische Spline-Funktionen. Im darauf folgenden Kapitel 2.7.2 wird die Bewertungsfunktion als Summe von geeigneten Spline-Funktionen definiert.

2.7.1 Splines

Die Darstellung der Splines folgt bis auf die Definition der B-Splines im wesentlichen der Darstellung im Buch von Hämmerlin und Hoffmann [29]. Man vergleiche aber auch die entsprechenden Kapitel in [54] und [52].

Im folgenden sei $[a,\ b] \subset \mathbb{R},\ a < b$ ein abgeschlossenes Intervall und $\Omega_n := \{x_k\}_{k=0}^{n}$ eine Knotenmenge, die eine *Zerlegung* des Intervalls $[a,\ b]$ bestimmt. Dabei soll $a = x_0 < x_1 < \ldots < x_n = b$ gelten.

Definition 2.7.1 (Polynom-Spline)
Eine Funktion $s\colon [a,\ b] \to \mathbb{R}$ heißt Polynom-Spline vom Grad l, $(l = 0, 1, \ldots)$, wenn sie die folgenden Eigenschaften besitzt

1. $s \in \mathsf{C}_{l-1}([a,\ b])$;
2. $s \in \mathsf{P}(l)$ *für* $x \in [x_k,\ x_{k+1}),\ 0 \leq k \leq n-1.$

[35] Siehe hierzu die Bemerkung zu der Definition C.12.
[36] Siehe Definition 2.6.2.

Dabei bezeichnet $\mathsf{C}_n(U)$ den Vektorraum der auf $U \subset \mathbb{R}$ n-mal stetig differenzierbaren Funktionen, $\mathsf{C}_{-1}(U)$ den Vektorraum der auf U stückweise stetigen Funktionen und $\mathsf{P}(n)$ den Vektoraum der Polynome vom Grad n.[37] *Die Menge aller Polynom-Splines vom Grad l zur Zerlegung Ω_n wird mit $\mathsf{S}_l(\Omega_n)$ bezeichnet.* ◇

Im folgenden werden Polynom-Splines kurz mit *Splines* bezeichnet. Splines sind also Funktionen, die auf den Teilintervallen als Polynome darstellbar sind. Aus der Definition erkennt man, daß $\mathsf{S}_l(\Omega_n)$ ein Vektorraum ist, und zwar ein Unterraum vom $\mathsf{C}_{l-1}([a,\ b])$. Der folgende Satz gibt eine explizite Basis für $\mathsf{S}_l(\Omega_n)$ an. Dazu betrachtet man zunächst die folgenden Funktionen: Für $k = 0, \ldots, n$ setzt man $p_k(x) := x^k$. Weiter setzt man für $j = 1, \ldots, n-1$

$$q_{lj} : [a,\ b] \to \mathbb{R} \quad \text{mit } q_{lj}(x) := \begin{cases} (x - x_j)^l & \text{für } x \geq x_j \\ 0 & \text{für } x < x_j \end{cases}.$$

Offensichtlich gilt $q_{lj} \in \mathsf{S}_l(\Omega_n),\ j = 1, \ldots, n-1$.

Satz 2.7.2
Die Menge $\mathsf{S}_l(\Omega_n)$ bildet einen Vektorraum der Dimension $(n + l)$. Die Elemente $\{p_0, \ldots, p_l, q_{l1}, \ldots, q_{l,n-1}\}$ bilden eine Basis von $\mathsf{S}_l(\Omega_n)$. D.h. jeder Spline $s \in \mathsf{S}_l(\Omega_n)$ besitzt eine eindeutige Darstellung der Form

$$s(x) = \sum_{k=0}^{l} a_k p_k(x) \ + \sum_{j=1}^{n-1} b_j q_{lj}(x),\ x \in [a,\ b],$$

wobei die $a_k \in \mathbb{R},\ k = 0, \ldots, l$ und $b_j \in \mathbb{R},\ j = 1, \ldots, n-1$ gilt.

Referenz: [29]

Von besonderem Interesse ist nun die Interpolation mit periodischen Endbedingungen, wie sie sich auch bei den Bewertungstabellen für die Torsionswinkel stellt. Splines vom Grad $l = 2m-1$ für $m \geq 2$ sind dafür besonders geeignet ([29]). Es sei $[a,\ b] \subset \mathbb{R},\ a < b$ und Ω_n eine Zerlegung von $[a,\ b]$. Weiter sei $\mathsf{S}_{2m-1} := \mathsf{S}_{2m-1}(\Omega_n)$ der Vektorraum der Spline vom Grad $2m-1$. Da $\dim(\mathsf{S}_{2m-1}) = n + 2m - 1$ gilt, können noch $2m-2$ freie Prameter bestimmt werden, wenn man die $n+1$ Interpolationsforderungen in den Punkten $a = x_0, \ldots, x_n = b$ stellt.

Satz 2.7.3 (Interploation mit periodischen Endbedingungen)
Es sei $f \in \mathsf{C}_m([a,\ b])$ und es gelte $f^{(i)}(a) = f^{(i)}(b)$, $i = 0, \ldots m-1$, wobei $f^{(i)}$ die i-te Ableitung von f ist. Dann lautet die Interpolationsaufgabe (IP): man bestimme $s \in \mathsf{S}_{2m-1}$, so daß die Bedingungen

1. $s(x_k) = f(x_k)$ *für* $k = 0, \ldots, n$;
2. $s^{(j)}(a) = s^{(j)}(b)$ *für* $j = 1, \ldots, 2m-2$

erfüllt sind. Die Interpolationsaufgabe (IP) ist stets eindeutig lösbar.

[37] Siehe Anhang B

Referenz: [29]

Die Splines, die der obigen Interpolationsaufgabe genügen, haben folgende Eigenschaft.

Lemma 2.7.4 (Extremaleigenschaft)
f sei wie in Satz 2.7.3. Weiter sei $\tilde{s} \in \mathsf{S}_{2m-1}$ der eindeutig bestimmte Spline, der die Interpolationsaufgabe (IP) löst. Nun sei g eine beliebige Funktion aus $\mathsf{C}_m([a,\ b])$, die periodisch und dieselben Interpolationsforderungen wie $\tilde{s}$ erfüllt. Dann gilt stets

$$\int_a^b \left|\tilde{s}^{(m)}(x)\right|^2 \, \mathsf{d}x \ \leq \ \int_a^b \left|g^{(m)}(x)\right|^2 \, \mathsf{d}x. \tag{2.21}$$

Referenz: [29]

Bemerkung: Die Extremaleigenschaft (2.21) kubischer Splines, d.h. m=2, hat eine geometrische Interpretation[38]:

Die Eigenschaft bedeutet, daß der interpolierende kubische Spline $\tilde{s}$ unter allen Funktionen $g \in \mathsf{C}_2([a,\ b])$, die dieselben Interpolationsforderungen erfüllen, näherungsweise die Norm der Krümmung minimiert. Aus diesem Grund werden zur Lösung der Interpolationsaufgabe (IP) kubische Spline verwendet.

Um nun die Interpolationsaufgabe (IP) praktisch zu lösen, werden zunächst *B-Splines* definiert. Jeder Spline vom Grad l läßt sich als Linearkombination von geeigneten B-Splines vom Grad l darstellen.

Definition 2.7.5 (B-Splines)
Es sei $\Omega_\infty = \{x_k\}_{k\in\mathbb{Z}}$, $x_k < x_{k+1}$ mit $x_k \to -\infty$ für $k \to -\infty$ und $x_k \to \infty$ für $k \to \infty$, eine Zerlegung von $\mathbb{R}$. Nun setzt man

$$\mathsf{B}_{0k}(x) := \begin{cases} 1 & , x_k \ \leq \ x \ < \ x_{k+1} \\ 0 & , \textit{sonst}. \end{cases}$$

Es gilt $\mathsf{B}_{0k}(x) \in \mathsf{S}_0(\Omega_\infty)$, $\forall k \in \mathbb{Z}$, d.h. B_{0k} ist ein Spline vom Grad 0. Nun kann man ein B-Spline vom Grad l, $l \geq 1$, wie folgt definieren

$$\mathsf{B}_{lk}(x) := \frac{x - x_k}{x_{k+l} - x_k}\, \mathsf{B}_{l-1,k}(x) \ + \ \frac{x_{k+l+1} - x}{x_{k+l+1} - x_{k+1}}\, \mathsf{B}_{l-1,k+1}(x).$$

Mittels Induktion nach l, folgt, daß B_{lk} ein Spline vom Grad l ist, d.h. $\mathsf{B}_{lk}(x) \in \mathsf{S}_l(\Omega_\infty)\ \forall k \in \mathbb{Z}$. ◇

Lemma 2.7.6
Mit den Bezeichnungen aus Definition 2.7.5 gilt:

a) Für alle $k \in \mathbb{Z}$ und $l = 1, \ldots$ gilt $\mathsf{B}_{lk}(x) > 0$, $\forall x \in (x_k, x_{k+l+1})$ und $\mathsf{B}_{lk}(x) = 0$, für $x \leq x_k$ oder $x \geq x_{k+l+1}$. Insbesondere hat $\mathsf{B}_{lk}(x)$ den Träger $\overline{\{x \in \mathbb{R} \mid \mathsf{B}_{lk}(x) \neq 0\}} = [x_k,\ x_{k+l+1}]$.

[38] Zu den Einzelheiten siehe Hämmerlin und Hoffmann [29].

b) Rekursionsformel für Ableitungen

$$\mathsf{B}'_{lk}(x) \;=\; l\,\left(\frac{\mathsf{B}_{l-1,k}(x)}{x_{k+l}-x_k} - \frac{\mathsf{B}_{l-1,k+1}(x)}{x_{k+l+1}-x_{k+1}}\right)$$

$$\begin{aligned}\mathsf{B}''_{lk}(x) \;&=\; l(l-1)\,\Bigg(\frac{\mathsf{B}_{l-2,k}(x)}{(x_{k+l}-x_k)(x_{k+l-1}-x_k)} - \frac{\mathsf{B}_{l-2,k+1}(x)}{(x_{k+l}-x_k)(x_{k+l}-x_{k+1})}\\ &\quad+ \frac{\mathsf{B}_{l-2,k+1}(x)}{(x_{k+l+1}-x_{k+1})(x_{k+l}-x_{k+1})} - \frac{\mathsf{B}_{l-2,k+2}(x)}{(x_{k+l+1}-x_{k+1})(x_{k+l+1}-x_{k+2})}\Bigg)\end{aligned}$$

Referenz: [29]

Es sei $[x_0,\ x_n] \subset \mathbb{R}$ und $\Omega_n := \{x_k\}_{k=0}^n$ eine Zerlegung des Intervalls. Aus Lemma 2.7.6 a) folgt, daß nur die B-Splines $\mathsf{B}_{l,-l}, \ldots, \mathsf{B}_{l,n-1}$ in $[x_0,\ x_n]$ von Null verschiedene Werte annehmen. Mittels dieser $n+l$ B-Splines lassen sich nun Splines vom Grad l darstellen. Genauer

Satz 2.7.7 (Darstellungssatz)
Jeder Spline $s \in \mathsf{S}_l(\Omega_n)$ besitzt im Intervall $[x_0,\ x_n]$ eine eindeutige Darstellung durch B-Splines

$$s(x) \;=\; \sum_{k=-l}^{n-1} \alpha_k\ \mathsf{B}_{lk}(x), \quad \alpha_k \in \mathbb{R}. \tag{2.22}$$

Referenz: [29]

Mit Hilfe des Darstellungssatzes kann man nun die Interpolationsaufgabe konkret lösen. Dazu sei $[a,\ b] \subset \mathbb{R}$, $a < b$ ein beliebiges Intervall. Mit $h := \frac{b-a}{n}$ erhält man eine äquidistante Zerlegung $\Omega_n := \{x_k\}_{k=0}^n$, mit $x_k := a + kh$, des Intervalls $[a,\ b]$. Weiter seien $Y_n := \{y_k\}_{k=0}^n$ mit $y_k \in \mathbb{R}$ und $y_0 = y_n$ die zu interpolierenden Energiewerte. Dann läßt sich die Interpolationsaufgabe (IP) wie folgt formulieren: man bestimme $s \in \mathsf{S}_3(\Omega_n)$, so daß die Bedingungen $s(x_k) = y_k$, für $k = 0, \ldots, n$, $s'(a) = s'(b)$ und $s''(a) = s''(b)$ erfüllt. Nach Satz 2.7.2 ist diese Aufgabe eindeutig lösbar und mit Satz 2.7.7 läßt sich diese Lösung sogar konkret berechnen, denn setzt man die obigen Bedingungen in die Gleichung (2.22) ein und wendet Lemma 2.7.6 a) an, so erhält man die folgenden n+3 Gleichungen

$$\begin{aligned}\sum_{k=i-3}^{i-1} \alpha_k\ \mathsf{B}_{3k}(x_i) \;&=\; y_i,\ 0 \le i \le n\\ \sum_{k=-3}^{-1} \alpha_k\ \mathsf{B}'_{3k}(x_0) \;&=\; \sum_{k=-3}^{-1} \alpha_k\ \mathsf{B}'_{3k}(x_n)\\ \sum_{k=-3}^{-1} \alpha_k\ \mathsf{B}''_{3k}(x_0) \;&=\; \sum_{k=-3}^{-1} \alpha_k\ \mathsf{B}''_{3k}(x_n).\end{aligned}$$

Mit Hilfe der Tabelle

	x_k	x_{k+1}	x_{k+2}	x_{k+3}	x_{k+4}
$\mathsf{B}_{3k}(x)$	0	$\frac{1}{6}$	$\frac{2}{3}$	$\frac{1}{6}$	0
$\mathsf{B}'_{3k}(x)$	0	$\frac{1}{2h}$	0	$-\frac{1}{2h}$	0
$\mathsf{B}''_{3k}(x)$	0	$\frac{1}{h^2}$	$-\frac{2}{h^2}$	$\frac{1}{h^2}$	0

kann man ein Gleichungssystem zur Berechnung des Koeffizientenvektors $\alpha = (\alpha_{-3}, \ldots, \alpha_{n-1})^{\mathsf{T}}$ des interpolierenden Splines $s \in \mathsf{S}_3(\Omega_n)$ aufstellen. Das Gleichungssystem ist durch

$$\underbrace{\frac{1}{6}\begin{pmatrix} -\frac{3}{h} & 0 & \frac{3}{h} & 0 & \ldots & 0 & \frac{3}{h} & 0 & -\frac{3}{h} \\ \frac{6}{h^2} & -\frac{12}{h^2} & \frac{6}{h^2} & 0 & \ldots & 0 & -\frac{6}{h^2} & \frac{12}{h^2} & -\frac{6}{h^2} \\ 1 & 4 & 1 & 0 & \ldots & & & & \\ & \ddots & \ddots & \ddots & \ddots & & & & \\ & & \ddots & \ddots & \ddots & \ddots & & 0 & \\ & & & \ddots & \ddots & \ddots & \ddots & & \\ & 0 & & & \ddots & \ddots & \ddots & \ddots & \\ & & & & & 1 & 4 & 1 & 0 \\ & & & & & & 1 & 4 & 1 \end{pmatrix}}_{=:B} \begin{pmatrix} \alpha_{-3} \\ \alpha_{-2} \\ \alpha_{-1} \\ \\ \vdots \\ \\ \\ \alpha_{n-2} \\ \alpha_{n-1} \end{pmatrix} = \underbrace{\begin{pmatrix} 0.0 \\ 0.0 \\ y_0 \\ \\ \vdots \\ \\ \\ y_{n-1} \\ y_0 \end{pmatrix}}_{=:b} \tag{2.23}$$

gegeben.

Bemerkung: Da die Matrix B invertierbar ist, ist das Gleichungssystem $B\alpha = b$ stets lösbar. Die Koeffizienten, um den Spline s als Linearkombination der B-Splines darzustellen, sind durch den Lösungsvektor α eindeutig bestimmt.

2.7.2 Die Bewertungsfunktion

In diesem Kapitel wird eine Bewertungsfunktion für die Torsionswinkel definiert, indem die diskreten Energiewerte der Bewertungstabellen durch kubische Splines interpoliert werden. Mit den Bezeichnungen aus Kapitel 2.2 sei $B_{FT}^{\mathcal{A}} = \{e \in B_{\mathcal{A}} | \; \Phi_{FT}^{\mathcal{A}}(e) = 1\}$ die Menge der Einfach-Bindungen des Moleküls $\mathcal{A}$, an denen flexible Torsionswinkel betrachtet werden. Weiter sei $F_{\mathcal{A}} = \{f_0, \ldots, f_{n_{ft}}\}$, $n_{ft} = |B_{\mathcal{A}}|$, die Knotenmenge des Fragmentebaums $T_{\mathcal{A}}$. Wie in Kapitel 2.2 gezeigt worden ist, gibt es für jedes $f_i \in F_{\mathcal{A}} \setminus \{f_0\}$ genau ein $e_i := \mathcal{I}_{F,B}^{\mathcal{A}}(f_i) \in B_{FT}^{\mathcal{A}}$, mit $\mathcal{I}_{F,B}^{\mathcal{A}}(f_i) = \{\mathsf{R}_{F_{\mathcal{A}},K_{\mathcal{A}}}(f_i), \mathsf{S}_{F_{\mathcal{A}},K_{\mathcal{A}}}(f_i)\}$. Dann sei $\triangle_i := \triangle_{e_i}$, wobei $\triangle_{e_i} = \{\Omega_{e_i}, \{f_k^i\}_{k=0}^{n(e_i)}\}$ der Datensatz der Knotenzerlegung und der Energiewerte für die Bindung e_i ist. Dann lautet die zu $\triangle_i$ gehörende Interpolationsaufgabe (IP($\triangle_i$)): man bestimme $s_i \in \mathsf{S}_3(\Omega_{e_i})$, so daß die Bedingungen

1. $s_i(x_k) = f_k^i$, für $k = 0, \ldots, n(e_i)$,
2. $s_i'(a) = s_i'(b)$ und $s_i''(a) = s_i''(b)$

erfüllt sind. Nach Satz 2.7.3 ist die Interpolationsaufgabe (IP($\triangle_i$)) stets eindeutig lösbar.

Nun sei $\Theta_i(\mathsf{C}_{\mathcal{A}}) = \Theta_{e_i}(\mathsf{C}_{\mathcal{A}})$ der zu Bindung e_i gehörende Torsionswinkel, der aus der Molekülkonformation $\mathsf{C}_{\mathcal{A}}$ abgeleitet worden ist. Weiter sei $\varphi_i \in \mathbb{R}$ der aktuelle Rotationswinkel bzgl. der Drehachse $\mathsf{D}_{F_{\mathcal{A}},\mathbb{R}^3\setminus\{0\}}(f_i)$. Der aktuelle Torsionswinkel bzgl. der Bindung e_i wird wie folgt bestimmt:

$$\theta_i^{akt} \;=\; \Theta_i(\mathsf{C}_{\mathcal{A}}) \;+\; \varphi_i, \; i = 1, \ldots, n_{ft}.$$

Dann läßt sich die Bewertungsfunktion für den aktuellen Torsionswinkel θ_i^{akt} wie folgt definieren:

$$s_{\triangle_i}\left(\Theta_i(\mathsf{C}_{\mathcal{A}})+\varphi_i\right) := \begin{cases} 0 & \text{, falls } f_0^i = \ldots = f_{n(e_i)}^i \\ s_i\left(\Theta_i(\mathsf{C}_{\mathcal{A}})+\varphi_i\right) & \text{, sonst} \end{cases}, \tag{2.24}$$

wobei $s_i \in \mathsf{S}_3(\Omega_{e_i})$ der durch die Interpolationsaufgabe (IP)($\triangle_i$) eindeutig bestimmte periodische kubische Spline ist. Damit ergibt sich für die Bewertungsfunktion des Moleküls:

Definition 2.7.8 (Bewertungsfunktion)
Bezeichnungen und Voraussetzungen wie oben. Weiter sei $N_{BF}^{\mathcal{A}} := \{j \in \{1,\ldots,n_{ft}\} \mid s_{\triangle_j} \neq 0\}$ *und* $n_{bf} := \left|N_{BF}^{\mathcal{A}}\right|$. *Nun heißt die Abbildung*

$$s_{tor}^{\mathcal{A}} \colon \mathbb{R}^{n_{ft}} \times \mathbb{R}^{n_{ft}} \to [-1,\ 1]$$

$$s_{tor}^{\mathcal{A}}\left(\left(\Theta_1(\mathsf{C}_{\mathcal{A}}),\ldots,\Theta_{n_{ft}}(\mathsf{C}_{\mathcal{A}})\right),\left(\varphi_1,\ldots,\varphi_{n_{ft}}\right)\right) := \begin{cases} 0 & \text{, falls } n_{bf}=0 \\ \frac{1}{n_{bf}} \sum\limits_{j\in N_{BF}} \frac{s_{\triangle_j}(\Theta_j(\mathsf{C}_{\mathcal{A}})+\varphi_j)}{\|s_{\triangle_j}\|_\infty} & \text{, sonst} \end{cases}$$

Bewertungsfunktion für die aktuellen Torsionswinkel des Moleküls $\mathcal{A}$ *zur Molekülkonformation* $\mathsf{C}_{\mathcal{A}}$ *bzgl. der Rotationswinkel.*

Lemma 2.7.9
Die Abbildung $s_{tor}^{\mathcal{A}}$ *aus Definition 2.7.8 ist wohldefiniert und bzgl. der Rotationswinkel* φ_j *zweimal stetig partiell differenzierbar.*

Beweis:
Die Wohldefiniertheit und die partielle Differenzierbarkeit folgt unmittelbar aus der Definition von $s_{\triangle_j}$ in Gleichung (2.24). ■

Bemerkung:

a) Falls es Bewertungstabellen für einzelne flexible Torsionswinkel des Molekül $\mathcal{A}$ gibt, so werden die aktuellen Werte dieser Winkel durch die obige Funktion an Hand dieser Tabellen bewertet. Dabei werden nur die Bewertungstabellen betrachtet, die nicht alle möglichen Winkel gleich bewerten.

b) Zur praktischen Berechnung von $s_i\left(\theta_i^{akt}\right)$, $\theta_i^{akt} = \Theta_i(\mathsf{C}_{\mathcal{A}}) + \varphi_i$ wird der Darstellungssatz 2.7.7 verwendet. Indem man für die Bindung e_{f_i} das Gleichungssystem $B_i\alpha_i = b_i$ aus Gleichung (2.23) löst, erhält man die Koeffizienten, um s_i auf dem Intervall $[x_0, x_{n(e_i)}]$ bzgl. der Basissplines darzustellen; d.h.

$$s_i(x) = \sum_{k=-3}^{n-1} \alpha_k^i \; \mathsf{B}_{3k}(x) \quad \forall x \in [x_0, x_{n(e_i)}]. \tag{2.25}$$

Um nun $s_i\left(\theta_i^{akt}\right)$ zu berechnen, muß noch der Winkel θ_i^{akt} auf das Intervall $\mathcal{I}_{e_i} = [x_0, x_{n(e_i)}] = [x_0, p(e_i) + x_0]$ entsprechend der Periode $p(e_i)$ projiziert werden. Dazu sei $\mathcal{P}[\mathcal{I}_{e_i}]\colon \mathbb{R} \to \mathcal{I}_{e_i}$ die entsprechende Projektion. Dann berechnet man $s_i\left(\mathcal{P}[\mathcal{I}_{e_i}]\left(\theta_i^{akt}\right)\right)$ mit Hilfe der Darstellung aus Gleichung (2.25). Da s_i ein $p(e_i)$-periodischer Spline ist und die Projektion $\mathcal{P}[\mathcal{I}_{e_i}]$ entsprechend definiert worden ist, gilt $s_i\left(\theta_i^{akt}\right) = s_i\left(\mathcal{P}[\mathcal{I}_{e_i}]\left(\theta_i^{akt}\right)\right)$. Analog werden die Werte der Ableitungen bestimmt.

Kapitel 3

Die Energiefunktion

Bevor in Kapitel 3.2 die Energiefunktion für das Modell 2.1 der Optimierungsaufgabe 2.1.4 bzw. in Kapitel 3.3 die Energiefunktion für das Modell 2.2 der Optimierungsaufgabe 2.1.4 aufgestellt wird, wird in Kapitel 3.1 das Minimierungsproblem allgemein formuliert. Dazu werden notwendige und hinreichende Kriterien zur Lösung des Minimierungsproblems angegeben.

3.1 Das Minimierungsproblem

Bei der Formulierung des Minimierungsproblems wird im wesentlichen auf das Buch von E. Polak [46] und das Buch von C. Geiger/C. Kanzow [19] zurück gegriffen. Man vergleiche aber auch die entsprechenden Kapitel in [28], [12], [53], [54], [31] und [16].

Ein Minimierungsproblem kann abstrakt als Aufgabe der Form

$$\min_{x \in U} f(x) \tag{3.1}$$

formuliert werden, wobei $f\colon U \to \mathbb{R}$ die *Zielfunktion* und $U \subseteq X$ ein offene Menge ist. Dabei bezeichnet X den zugrunde gelegten Raum, z.B. $X = \mathbb{R}^n$, und U den zulässigen Bereich. Ein Element $x \in X$ heißt *zulässige Lösung* des Problems (3.1), wenn $x \in U$ gilt.

Definition 3.1.1

a) $x^* \in U$ heißt globales Minimum des Problems (3.1), falls

$$f(x^*) \leq f(x), \quad \forall x \in U$$

gilt. Gilt $f(x^*) < f(x)$ $\forall x \in U$, so heißt x^* isoliertes globales Minimum.

b) $x^* \in U$ heißt lokales Minimum des Problems (3.1), falls ein $\varepsilon > 0$ existiert, so daß $B_\varepsilon(x^*) := \{x \in X \mid \|x - x^*\| \leq \varepsilon \} \subset U$ und

$$f(x^*) \leq f(x), \quad \forall x \in B_\varepsilon(x^*)$$

gilt.

c) $x^* \in U$ *heißt isoliertes lokales Minimum des Problems (3.1), falls ein* $\varepsilon > 0$ *existiert, so daß* $B_\varepsilon(x^*) := \{x \in X \mid \|x - x^*\| \leq \varepsilon\} \subset U$ *und* $f(x^*) < f(x)\ \forall x \in B_\varepsilon(x^*)$ *gilt.* ◇

Bemerkung: Es sei $x^* \in U$ ein (isoliertes) globales Minimum. Dann ist x^* für jedes $\varepsilon > 0$ ein (isoliertes) lokales Minimium.

Die nächsten Aussagen charakterisieren lokale Minima.

Korollar 3.1.2
Es sei $U \subseteq \mathbb{R}^n$, *U offen. Die Zielfunktion* $f\colon U \to \mathbb{R}$ *des Minimierungsproblems (3.1) sei stetig partiell differenzierbar. Falls* $x^* \in U$ *ein lokales Minimum von (3.1) ist, so gilt*

$$\nabla f(x^*) \;=\; 0,$$

wobei $\nabla f(x) = \left(\frac{\partial}{\partial x_1} f(x) \ldots \frac{\partial}{\partial x_n} f(x)\right)^{\mathsf{T}}$ *der Gradient von* f *in* x *ist.*

Referenz: [46], [16]

Aus Korollar 3.1.2 folgt: $\nabla f(x^*) = 0$ ist ein notwendiges Kriterium dafür, daß $x^* \in U$ ein lokales Minimum des Problems (3.1) ist. Um hinreichende Kriterien herzuleiten, braucht man die Hesse-Matrix einer Funktion f.

Definition 3.1.3
Es sei $U \subset \mathbb{R}^n$ *eine offene Teilmenge und* $f\colon U \to \mathbb{R}$ *eine zweimal stetig partiell differenzierbare Funktion. Die Hesse-Matrix von* f *in* $x \in U$ *ist die Matrix*

$$\nabla^2 f(x) \;:=\; (\mathit{Hess}\ f)(x) \;:=\; \left(\frac{\partial^2\ f(x)}{\partial x_i\ \partial x_j}\right)_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}} \in \mathsf{M}(n \times n; \mathbb{R}).$$

Die Hesse-Matrix ist symmetrisch, da $\frac{\partial^2\ f(x)}{\partial x_i\ \partial x_j} = \frac{\partial^2\ f(x)}{\partial x_j\ \partial x_i}$ *gilt.* ◇

Der folgende Satz gibt ein weiteres Charakteristikum lokaler Minima an.

Satz 3.1.4
Es sei $U \subseteq \mathbb{R}^n$, *U offen. Die Zielfunktion* $f\colon U \to \mathbb{R}$ *des Minimierungsproblems (3.1) sei zweimal stetig partiell differenzierbar. Falls* $x^* \in U$ *ein lokales Minimum von (3.1) ist, so gilt*

$$< h,\ \nabla^2 f(x^*)\ h >_2\ \geq\ 0, \quad \forall h \in \mathbb{R}^n,$$

wobei $\nabla^2 f(x^*)$ *die Hesse-Matrix von* f *in* x^* *ist.*

Referenz: [46]

Ein hinreichendes Kriterium für ein isoliertes lokales Minimum ist durch den folgenden Satz gegeben.

Satz 3.1.5
Es sei $U \subseteq \mathbb{R}^n$, U offen. Die Zielfunktion $f\colon U \to \mathbb{R}$ des Minimierungsproblems (3.1) sei zweimal stetig partiell differenzierbar und $x^* \in U$ sei ein Punkt mit $\nabla f(x^*) = 0$. Falls

$$< h,\ \nabla^2 f(x^*)\ h >_2\ >\ 0, \quad \forall h \in \mathbb{R}^n \setminus 0$$

gilt, so hat f in x^* ein isoliertes lokales Minimum.

Referenz: [16]

3.2 Modell 1

In Modell 2.1 der Optimierungsaufgabe 2.1.4 wird ein flexibles Molekül auf ein rigides Referenzmolekül überlagert. Es sei also $\mathcal{A}$ ein flexibles und $\mathcal{R}$ ein rigides Molekül mit $n_a := n_a[\mathcal{A}]$ bzw. $n_a[\mathcal{R}]$ Atomen.[1] Weiter sei $\mathsf{C}_\mathcal{A} \subset \mathbb{R}^{3n_a}$ bzw. $\mathsf{C}_\mathcal{R} \subset \mathbb{R}^{3n_a[\mathcal{R}]}$ eine zulässige Konformation für das Molekül $\mathcal{A}$ bzw. $\mathcal{R}$. $\mathcal{DS}_\mathcal{A}(\mathsf{C}_\mathcal{A})$ bzw. $\mathcal{DS}_\mathcal{R}(\mathsf{C}_\mathcal{R})$ sei die Datenstruktur für das entsprechende Molekül bzgl. der gegebenen Konformation. Aus der Datenstruktur kann man alle Informationen, wie sie in Kapitel 2.2 beschrieben worden sind, ableiten. Flexibel bedeutet daher, daß das Molekül Bindungen besitzt, an denen der Torsionswinkel variiert werden kann, d.h. es gilt $B_{FT}^\mathcal{A} \neq \emptyset$. Für das rigide Molekül $\mathcal{R}$ gilt jedoch $B_{FT}^\mathcal{R} = \emptyset$. Insbesondere ist die Molekülkonformation $\mathsf{C}_\mathcal{R}$ unveränderlich. Daher besteht die Energiefunktion aus dem Alignmentindex, dem n-Potential für das Molekül $\mathcal{A}$ und der Bewertungsfunktion für die flexiblen Torsionswinkel des Moleküls $\mathcal{A}$.
Dazu werden für das Molekül $\mathcal{A}$ die folgenden Informationen aus der Datenstruktur $\mathcal{DS}_\mathcal{A}(\mathsf{C}_\mathcal{A})$ benötigt. $K_\mathcal{A} := K(\mathsf{C}_\mathcal{A}) = \{a_1, \dots, a_{n_a}\} \subset \mathbb{R}^3$ sei die Menge der Atomkoordinaten bzgl. der Konformation $\mathsf{C}_\mathcal{A}$. $\Sigma_\mathcal{A} = (\sigma_1, \dots, \sigma_{n_a}) \subset \mathbb{R}^{n_a}$ sei der Vektor der Van-der-Waals-Radien, wobei σ_j zum Atom a_j, $j = 1, \dots, n_a$ gehört.[2] In folgenden sei $n_{ft} := n_{ft}[\mathcal{A}] := \left|B_{FT}^\mathcal{A}\right|$ die Anzahl der Bindungen, an denen der Torsionswinkel variiert werden kann. Es sei $\Theta^{tor}(\mathsf{C}_\mathcal{A}) = (\Theta_1(\mathsf{C}_\mathcal{A}), \dots, \Theta_{n_{ft}}) \subset \mathbb{R}^{n_{ft}}$ der Vektor der Torsionswinkel, wobei $\Theta_i(\mathsf{C}_\mathcal{A})$ der an der Bindung $e_i := \mathcal{I}_{F,B}^\mathcal{A}(f_i) \in B_{FT}^\mathcal{A}$, $i = 1, \dots, n_{ft}$, bzgl. der Konformation $\mathsf{C}_\mathcal{A}$ gemessene Torsionswinkel ist. Für $p \in \mathcal{I}_{PE}$ sei $Q_\mathcal{A}^p(x) = \sum_{j=1}^{n_g^p[\mathcal{A}]} \mathsf{G}\left(h_{p,j}^\mathcal{A}, b_{p,j}^\mathcal{A}, z_{p,j}^\mathcal{A}\right)(x)$ eine Summe von Gaussfunktionen, wobei $\left(h_{p,j}^\mathcal{A}, b_{p,j}^\mathcal{A}, z_{p,j}^\mathcal{A}\right) \in \mathbb{R}\text{x}\mathbb{R}_{>0}\text{x}\mathbb{R}^3$ für $j = 1, \dots, n_g^p[\mathcal{A}]$, und $n_g^p[\mathcal{A}] \in \mathbb{N}_0$ gilt. Für $p \in \mathcal{I}_{PE}$ ist $n_g^p[\mathcal{A}] \in \mathbb{N}_0$ die Anzahl der Gaussfunktionen, die verwendet werden, um die räumlich lokalisierte Dichte des Moleküls $\mathcal{A}$ bzgl. der physikochemische Eigenschaft p darzustellen.[3] Da das Molekül $\mathcal{R}$ rigide ist, ist für $\mathcal{R}$ nur die Darstellung der räumlich lokalisierten Dichte bzgl. der physikochemischen Eigenschaft $p \in \mathcal{I}_{PE}$ interessant. Daher sei für $p \in \mathcal{I}_{PE}$ $Q_\mathcal{R}^p(x) = \sum_{j=1}^{n_g^p[\mathcal{R}]} \mathsf{G}\left(h_{p,j}^\mathcal{R}, b_{p,j}^\mathcal{R}, z_{p,j}^\mathcal{R}\right)(x)$ eine Summe von Gaussfunktionen, wobei $\left(h_{p,j}^\mathcal{R}, b_{p,j}^\mathcal{R}, z_{p,j}^\mathcal{R}\right) \in \mathbb{R}\text{x}\mathbb{R}_{>0}\text{x}\mathbb{R}^3$ für $j = 1, \dots, n_g^p[\mathcal{R}]$, $n_g^p[\mathcal{R}] \in \mathbb{N}_0$ gilt.

[1] Es werden die Bezeichnungen und Definitionen aus Kapitel 2.2 verwendet.
[2] Mit a_j werden sowohl das Atom als auch die zugehörigen Atomkoordinaten bezeichnet.
[3] Siehe die Definitionen 2.2.1 und 2.2.2.

Für $p \in \mathcal{I}_{PE}$ ist $n_g^p[\mathcal{R}] \in \mathbb{N}_0$ die Anzahl der Gaussfunktionen, die verwendet werden, um die räumlich lokalisierte Dichte des Moleküls $\mathcal{R}$ bzgl. der physikochemische Eigenschaft p darzustellen.

$Q_{\mathcal{R}}^p(x)$ ist die räumlich lokalisierte Dichte des Moleküls $\mathcal{R}$ bzgl. der physikochemischen Eigenschaft $p \in \mathcal{I}_{PE}$,
Für $p \in \mathcal{I}_{PE}$ seien reellwertige Energiekoeffizienten $\tilde{\lambda}_p > 0$ gegeben, um die einzelnen physikochemischen Eigenschaften im Alignmentindex zu gewichten. Nun setzt man

$$\mathcal{I}_{PE}^{\mathcal{A},\mathcal{R}} := \{p \in \mathcal{I}_{PE} \mid (Q_{\mathcal{A}}^p(x) \neq 0) \ \vee \ (Q_{\mathcal{R}}^p(x) \neq 0)\} \neq \emptyset$$

und $\tilde{\lambda} := \sum_{i \in \mathcal{I}_{PE}^{\mathcal{A},\mathcal{R}}} \tilde{\lambda}_i > 0$. $\mathcal{I}_{PE}^{\mathcal{A},\mathcal{R}}$ ist ungleich der leeren Menge, da wenigstens die räumlich lokalisierte Dichte der Elektronendichte, d.h. die sterische Gestalt, für beide Moleküle gegeben ist. Dann gilt $\sum_{i \in \mathcal{I}_{PE}^{\mathcal{A},\mathcal{R}}} \lambda_i = 1$ mit $\lambda_i := \tilde{\lambda}_i / \tilde{\lambda}, \ i \in \mathcal{I}_{PE}^{\mathcal{A},\mathcal{R}}$.
Um nun die Atomkoordinaten und die Gausszentren des Moleküls $\mathcal{A}$ zu transformieren, braucht man die entsprechenden Transformationsoperatoren. Es sei $\vec{T}_{\mathcal{A}} = (F_{\mathcal{A}}, \ \vec{E}_{FT}^{\mathcal{A}})$ der zum Molekül $\mathcal{A}$ gehörende und gerichtete Fragmentebaum. Weiter sei $\mathcal{S}(K_{\mathcal{A}}, \vec{T}_{\mathcal{A}})$ eine Transformationsstruktur auf $K_{\mathcal{A}}$ bzgl. $\vec{T}_{\mathcal{A}}$. und $\mathcal{T}_{\mathcal{A}} := \mathcal{T}[\mathcal{S}(K_{\mathcal{A}}, \vec{T}_{\mathcal{A}})]$ der zu $\mathcal{S}(K_{\mathcal{A}}, \vec{T}_{\mathcal{A}})$ gehörende Transformationsoperator aus Definition 2.4.27. Es sei $\mathcal{R}_{\mathcal{A}} := \mathcal{R}[\mathcal{S}(K_{\mathcal{A}}, \vec{T}_{\mathcal{A}})]$ der zu $\mathcal{S}(K_{\mathcal{A}}, \vec{T}_{\mathcal{A}})$ gehörende lokale Rotationsoperator aus Definition 2.4.18.
Für $p \in \mathcal{I}_{PE}^{\mathcal{A},\mathcal{R}}$ sei $\mathcal{S}(K_{\mathcal{A}}^p, \vec{T}_{\mathcal{A}})$ eine formale Transformationsstruktur auf $K_{\mathcal{A}}^p := K_{\mathcal{A}}^* \cup GC_{\mathcal{A}}^p$ bzgl. des Fragmentebaumes $\vec{T}_{\mathcal{A}}$.[4] Für $p \in \mathcal{I}_{PE}^{\mathcal{A},\mathcal{R}}$ ist $GC_{\mathcal{A}}^p$ die formal geordnete Menge der zu räumlich lokalsierten Dichte $Q_{\mathcal{A}}^p$ gehörenden Gausszentren (siehe Kapitel 2.2). Für $p \in \mathcal{I}_{PE}^{\mathcal{A},\mathcal{R}}$ sei $\mathcal{T}_{\mathcal{A}}^p := \mathcal{T}[\mathcal{S}(K_{\mathcal{A}}^p, \vec{T}_{\mathcal{A}})]$ bzw. $\mathcal{R}_{\mathcal{A}}^p := \mathcal{R}[\mathcal{S}(K_{\mathcal{A}}^p, \vec{T}_{\mathcal{A}})]$ der zu $\mathcal{S}(K_{\mathcal{A}}^p, \vec{T}_{\mathcal{A}})$ gehörende Transformationsoperator bzw. lokale Rotationsoperator.[5]

Es sei $U(n_a) := \mathbb{R}^3 \times \mathbb{R}^4 \setminus \{0\} \times \mathbb{R}^3 \times \mathbb{R}^{n_{ft}}$ der Parameterraum der Transformationsoperatoren, die die Koordinaten des Moleküls $\mathcal{A}$ transformieren. Diese Operatoren transformieren die Atom- bzw. Gausskoordinaten des Moleküls $\mathcal{A}$ mittels des Parameters $u_{\mathcal{A}} = (t, q, c, \varphi_{\mathcal{A}}) \in U(n_a)$, wobei t der Translationsvektor, q der Quaternionenvektor, c das globale Rotationszentrum und $\varphi_{\mathcal{A}} = (\varphi_1, \dots, \varphi_{n_{ft}})$ der Vektor der internen Rotationsparameter ist. Dabei ist φ_i, $i = 1, \dots, n_{ft}$ der zu Bindung $e_i = \mathcal{I}_{F,B}^{\mathcal{A}}(f_i) \in B_{FT}^{\mathcal{A}}$ und damit zu Drehachse $\mathsf{D}_{F_{\mathcal{A}}, \mathbb{R}^3 \setminus \{0\}}(f_i)$ gehörende Rotationsparameter. Wendet man die Transformationsoperatoren auf die Gausszentren an, so werden die räumlich lokalisierten Dichten des Moleküls $\mathcal{A}$ vom Parameter $u_{\mathcal{A}}$ abhängig, denn man erhält für $p \in \mathcal{I}_{PE}^{\mathcal{A},\mathcal{R}}$

$$Q_{\mathcal{A}}^p[u_{\mathcal{A}}](x) := \sum_{j=1}^{n_g^p[\mathcal{A}]} \mathsf{G}\left(h_{p,j}^{\mathcal{A}}, b_{p,j}^{\mathcal{A}}, \mathcal{T}_{\mathcal{A}}^p(u_{\mathcal{A}}, \ z_{p,j}^{\mathcal{A}})\right)(x)$$

Die Energiefunktion besteht nun aus dem Alignmentindex $\mathsf{A}_{\mathcal{A},\mathcal{R}}^{\mathsf{H}}$ aus Definition 2.5.6. um die Überlagerung der aktuellen Konformationen zu bewerten, dem n-Potential $P_{n_a[\mathcal{A}]}$ aus Definition 2.6.2, das auf die Atomkoordinaten des Moleküls $\mathcal{A}$ angewendet wird, um unzulässige Konformationen zu vermeiden, und der Bewertungsfunktion $s_{tor}^{\mathcal{A}}$ aus Definition 2.7.8, um die aktuellen

[4]Zu Definition von $K_{\mathcal{A}}^*$ siehe Kapitel 2.2.
[5]Zu Definition der formalen Struktur und deren Operatoren siehe Kapitel 2.2.

Torsionswinkel zu bewerten. Um nun diese Komponenten zu gewichten, nimmt man reellwertige Energiekoeffizienten. Es sei $\lambda_A < 0$ der Energiekoeffizient für den Alignmentindex, $\lambda_P > 0$ der Koeffizient für das n-Potential und $\lambda_T > 0$ der Energiekoeffizient für die Bewertungsfunktion der Torsionswinkel.
Mit den Bezeichnungen und entsprechenden Voraussetzungen aus den einzelnen Definitionen ergibt sich die folgende Zielfunktion

$$\begin{aligned}\mathsf{E}^*_{\mathcal{A},\mathcal{R}}(u_{\mathcal{A}}) := \lambda_A\ \mathsf{A}^{\mathsf{H}}_{\mathcal{A},\mathcal{R}} \Big(\ (Q^p_{\mathcal{A}}[u_{\mathcal{A}}](x))_{p\in\mathcal{I}^{\mathcal{A},\mathcal{R}}_{PE}}\ ,\ (Q^p_{\mathcal{R}}(x))_{p\in\mathcal{I}^{\mathcal{A},\mathcal{R}}_{PE}}\ \Big) \\ +\ P_{n_a}\left(\lambda_P, (\mathcal{T}_{\mathcal{A}}(u_{\mathcal{A}}, a_1), \ldots, \mathcal{T}_{\mathcal{A}}(u_{\mathcal{A}}, a_{n_a}), \Sigma_{\mathcal{A}}\right)\ +\ \lambda_T\ s^{\mathcal{A}}_{tor}\left(\Theta^{tor}(\mathsf{C}_{\mathcal{A}}), \varphi_{\mathcal{A}}\right).\end{aligned}$$

Die Optimierungsaufgabe lautet

$$\min_{u_{\mathcal{A}}\in U(n_a)} \mathsf{E}^*_{\mathcal{A},\mathcal{R}}(u_{\mathcal{A}}). \tag{3.2}$$

Im folgenden soll die Energiefunktion $\mathsf{E}^*_{\mathcal{A},\mathcal{R}}$ untersucht werden. Dazu sei $\Psi_{K_{\mathcal{A}},F_{\mathcal{A}}} : K_{\mathcal{A}} \to F_{\mathcal{A}}$ die Verteilung der Atomkoordinaten auf die zu den Fragmenten gehörenden Knoten im Baum $\vec{T}_{\mathcal{A}}$ und für $p \in \mathcal{I}^{\mathcal{A},\mathcal{R}}_{PE}$ sei $\Psi^p_{K^p_{\mathcal{A}},F_{\mathcal{A}}} : K^p_{\mathcal{A}} \to F_{\mathcal{A}}$ die Verteilung der entsprechenden Gausskoordinaten auf die zu den Fragmenten gehörenden Knoten.
Für den ersten Summanden der Funktion gilt

$$\begin{aligned}&\mathsf{A}^{\mathsf{H}}_{\mathcal{A},\mathcal{R}} \left((Q^p_{\mathcal{A}}[u_{\mathcal{A}}](x))_{p\in\mathcal{I}^{\mathcal{A},\mathcal{R}}_{PE}}\ ,\ (Q^q_{\mathcal{R}}(x))_{p\in\mathcal{I}^{\mathcal{A},\mathcal{R}}_{PE}}\right) \\ &= \sum_{p\in\mathcal{I}^{\mathcal{A},\mathcal{R}}_{PE}} \lambda_p\ \mathsf{H}\left(\sum_{i=1}^{n^p_g[\mathcal{A}]} \mathsf{G}\left(h^{\mathcal{A}}_{p,i}, b^{\mathcal{A}}_{p,i}, \mathcal{T}^p_{\mathcal{A}}(u_{\mathcal{A}},\ z^{\mathcal{A}}_{p,i})\right)(x),\ \sum_{j=1}^{n^p_g[\mathcal{R}]} \mathsf{G}\left(h^{\mathcal{R}}_{p,j}, b^{\mathcal{R}}_{p,j}, z^{\mathcal{R}}_{p,j}\right)(x)\right) \\ &\overset{L.2.5.5}{=} \sum_{p\in\mathcal{I}^{\mathcal{A},\mathcal{R}}_{PE}} \lambda_p\ \frac{2\ \mathcal{Z}\left(\sum_{i=1}^{n^p_g[\mathcal{A}]} \mathsf{G}\left(h^{\mathcal{A}}_{p,i}, b^{\mathcal{A}}_{p,i}, \mathcal{T}^p_{\mathcal{A}}(u_{\mathcal{A}},\ z^{\mathcal{A}}_{p,i})\right)(x),\ \sum_{j=1}^{n^p_g[\mathcal{R}]} \mathsf{G}\left(h^{\mathcal{R}}_{p,j}, b^{\mathcal{R}}_{p,j}, z^{\mathcal{R}}_{p,j}\right)(x)\right)}{\mathcal{N}\left(\sum_{i=1}^{n^p_g[\mathcal{A}]} \mathsf{G}\left(h^{\mathcal{A}}_{p,i}, b^{\mathcal{A}}_{p,i}, \mathcal{T}^p_{\mathcal{A}}(u_{\mathcal{A}},\ z^{\mathcal{A}}_{p,i})\right)(x)\right) + \mathcal{N}\left(\sum_{j=1}^{n^p_g[\mathcal{R}]} \mathsf{G}\left(h^{\mathcal{R}}_{p,j}, b^{\mathcal{R}}_{p,j}, z^{\mathcal{R}}_{p,j}\right)(x)\right)}.\end{aligned}$$

Die Zähler und Nenner der obigen Summanden sind dabei nach Lemma 2.5.5 wie folgt gegeben:

$$\begin{aligned}&\mathcal{Z}\left(\sum_{i=1}^{n^p_g[\mathcal{A}]} \mathsf{G}\left(h^{\mathcal{A}}_{p,i}, b^{\mathcal{A}}_{p,i}, \mathcal{T}^p_{\mathcal{A}}(u_{\mathcal{A}},\ z^{\mathcal{A}}_{p,i})\right)(x),\ \sum_{j=1}^{n^p_g[\mathcal{R}]} \mathsf{G}\left(h^{\mathcal{R}}_{p,j}, b^{\mathcal{R}}_{p,j}, z^{\mathcal{R}}_{p,j}\right)(x)\right) \\ &\quad = \sum_{i=1}^{n^p_g[\mathcal{A}]} \sum_{j=1}^{n^p_g[\mathcal{R}]} \left(\frac{h^{\mathcal{A}}_{p,i} h^{\mathcal{R}}_{p,j}}{(b^{\mathcal{A}}_{p,i} + b^{\mathcal{R}}_{p,j})^{\frac{3}{2}}}\ \exp\left(\frac{-b^{\mathcal{A}}_{p,i} b^{\mathcal{R}}_{p,j}}{b^{\mathcal{A}}_{p,i} + b^{\mathcal{R}}_{p,j}}\ \mathsf{N}^p_{z^{\mathcal{R}}_{p,j}}\left(u_{\mathcal{A}}, z^{\mathcal{A}}_{p,i}\right)\right)\right),\end{aligned}$$

wobei für $p \in \mathcal{I}_{PE}$ $\mathsf{N}^p_y : U(n_a) \times GC^p_{\mathcal{A}} \to \mathbb{R}$ mit $\mathsf{N}_y(u_{\mathcal{A}}, g) := \|\mathcal{T}^p_{\mathcal{A}}(u_{\mathcal{A}}, g) - y\|^2_2$ die Abbildung aus Lemma 2.4.29 a) ist. Dabei ist $y \in \mathbb{R}^3$ beliebig aber fest.

$$\mathcal{N}\left(\sum_{j=1}^{n^p_g[\mathcal{R}]} \mathsf{G}\left(h^{\mathcal{R}}_{p,j}, b^{\mathcal{R}}_{p,j}, z^{\mathcal{R}}_{p,j}\right)(x)\right) = \sum_{k,l=1}^{n^p_g[\mathcal{R}]} \left(\frac{h^{\mathcal{R}}_{p,k} h^{\mathcal{R}}_{p,l}}{(b^{\mathcal{R}}_{p,k} + b^{\mathcal{R}}_{p,l})^{\frac{3}{2}}} \exp\left(\frac{-b^{\mathcal{R}}_{p,k} b^{\mathcal{R}}_{p,l}}{b^{\mathcal{R}}_{p,k} + b^{\mathcal{R}}_{p,l}} \|z^{\mathcal{R}}_{p,k} - z^{\mathcal{R}}_{p,l}\|^2_2\right)\right)$$

und

$$\mathcal{N}\left(\sum_{i=1}^{n_g^p[\mathcal{A}]} \mathsf{G}\left(h_{p,i}^{\mathcal{A}}, b_{p,i}^{\mathcal{A}}, \mathcal{T}_{\mathcal{A}}^p(u_{\mathcal{A}},\ z_{p,i}^{\mathcal{A}})\right)(x)\right)$$
$$= \sum_{k,l=1}^{n_g^p[\mathcal{A}]} \left[\frac{h_{p,k}^{\mathcal{A}} h_{p,l}^{\mathcal{A}}}{(b_{p,k}^{\mathcal{A}}+b_{p,l}^{\mathcal{A}})^{\frac{3}{2}}} \exp\left(\frac{-b_{p,k}^{\mathcal{A}} b_{p,l}^{\mathcal{A}}}{b_{p,k}^{\mathcal{A}}+b_{p,l}^{\mathcal{A}}} \|\mathcal{T}_{\mathcal{A}}^p(u_{\mathcal{A}}, z_{p,k}^{\mathcal{A}}) - \mathcal{T}_{\mathcal{A}}^p(u_{\mathcal{A}}, z_{p,l}^{\mathcal{A}})\|_2^2\right)\right]$$

$$\stackrel{L.2.4.28b)}{=} \sum_{i=0}^{n_{ft}-1} \sum_{j=i+1}^{n_{ft}} \left[\sum_{z_{p,k}^{\mathcal{A}} \in GC_i^p[\mathcal{A}]} \sum_{z_{p,l}^{\mathcal{A}} \in GC_j^p[\mathcal{A}]} \left(\frac{h_{p,k}^{\mathcal{A}} h_{p,l}^{\mathcal{A}}}{(b_{p,k}^{\mathcal{A}}+b_{p,l}^{\mathcal{A}})^{\frac{3}{2}}} \exp\left(\frac{-b_{p,k}^{\mathcal{A}} b_{p,l}^{\mathcal{A}}}{b_{p,k}^{\mathcal{A}}+b_{p,l}^{\mathcal{A}}} \mathsf{N}_{ij}^p\left(\varphi_{\mathcal{A}}, z_{p,k}^{\mathcal{A}}, z_{p,l}^{\mathcal{A}}\right)\right)\right)\right]$$
$$+ \sum_{i=0}^{n_{ft}} \left[\sum_{z_{p,k}^{\mathcal{A}}, z_{p,l}^{\mathcal{A}} \in GC_i^p[\mathcal{A}]} \left(\frac{h_{p,k}^{\mathcal{A}} h_{p,l}^{\mathcal{A}}}{(b_{p,k}^{\mathcal{A}}+b_{p,l}^{\mathcal{A}})^{\frac{3}{2}}} \exp\left(\frac{-b_{p,k}^{\mathcal{A}} b_{p,l}^{\mathcal{A}}}{b_{p,k}^{\mathcal{A}}+b_{p,l}^{\mathcal{A}}} \|z_{p,k}^{\mathcal{A}} - z_{p,l}^{\mathcal{A}}\|_2^2\right)\right)\right].$$

wobei $GC_i^p[\mathcal{A}] = (\Psi_{K_{\mathcal{A}}^p, F_{\mathcal{A}}}^q)^{-1}(f_i) := \{z_{p,k}^{\mathcal{A}} \in GC_{\mathcal{A}}^p \mid \Psi_{K_{\mathcal{A}}^p, F_{\mathcal{A}}}^q(z_{p,k}^{\mathcal{A}}) = f_i\}$, für $p \in \mathcal{I}_{PE}^{\mathcal{A},\mathcal{R}}$ und $i = 0, \ldots, n_{ft}$ das Urbild von f_i in $GC_{\mathcal{A}}^p$ ist. Formal gilt $GC_i^p[\mathcal{A}] \cap GC_j^p[\mathcal{A}] = \emptyset$ für $i \neq j$. Für $p \in \mathcal{I}_{PE}^{\mathcal{A},\mathcal{R}}$ und $i, j = 0, \ldots, n_{ft}$ mit $i \neq j$ ist $\mathsf{N}_{ij}^p : \mathbb{R}^{n_{ft}} \times GC_i^p[\mathcal{A}] \times GC_j^p[\mathcal{A}] \to \mathbb{R}$, mit

$$\mathsf{N}_{ij}^p(\varphi_{\mathcal{A}}, g_1, g_2) := \|\mathsf{R}[\mathcal{S}(K_{\mathcal{A}}^p, \vec{T}_{\mathcal{A}}), f_i](\Phi[i], g_1) - \mathsf{R}[\mathcal{S}(K_{\mathcal{A}}^p, \vec{T}_{\mathcal{A}}), f_j](\Phi[j], g_2)\|_2^2$$

die Abbildung aus Lemma 2.4.26, wobei $\Phi[k]$ der Vektor derjenigen lokalen Rotationsparameter φ_l ist, die an $\mathsf{R}[\mathcal{S}(K_{\mathcal{A}}^p, \vec{T}_{\mathcal{A}}), f_k]$ übergeben werden.

Für den ersten Summanden der Energiefunktion, dem Alignmentindex, folgt insgesamt

$$\mathsf{A}_{\mathcal{A},\mathcal{R}}^{\mathsf{H}}\left(\left(Q_{\mathcal{A}}^p[u_{\mathcal{A}}](x)\right)_{p \in \mathcal{I}_{PE}^{\mathcal{A},\mathcal{R}}},\ \left(Q_{\mathcal{R}}^p(x)\right)_{p \in \mathcal{I}_{PE}^{\mathcal{A},\mathcal{R}}}\right)$$
$$= \sum_{p \in \mathcal{I}_{PE}^{\mathcal{A},\mathcal{R}}} 2\lambda_p \frac{\mathcal{Z}\left(\sum_{i=1}^{n_g^p[\mathcal{A}]} \mathsf{G}\left(h_{p,i}^{\mathcal{A}}, b_{p,i}^{\mathcal{A}}, \mathcal{T}_{\mathcal{A}}^p(u_{\mathcal{A}},\ z_{p,i}^{\mathcal{A}})\right)(x),\ \sum_{j=1}^{n_g^p[\mathcal{R}]} \mathsf{G}\left(h_{p,j}^{\mathcal{R}}, b_{p,j}^{\mathcal{R}}, z_{p,j}^{\mathcal{R}}\right)(x)\right)}{\mathcal{N}\left(\sum_{i=1}^{n_g^p[\mathcal{A}]} \mathsf{G}\left(h_{p,i}^{\mathcal{A}}, b_{p,i}^{\mathcal{A}}, \mathcal{R}_{\mathcal{A}}^p(\varphi_{\mathcal{A}},\ z_{p,i}^{\mathcal{A}})\right)(x)\right) + \mathcal{N}\left(\sum_{j=1}^{n_g^p[\mathcal{R}]} \mathsf{G}\left(h_{p,j}^{\mathcal{R}}, b_{p,j}^{\mathcal{R}}, z_{p,j}^{\mathcal{R}}\right)(x)\right)}.$$

Somit sind nur die Zähler der Hodgkin-Indizes von den Parametern (t, q, c) der globalen Transformation abhängig, während die jeweiligen Nenner nur von den internen Rotationsparametern abhängig sind. Die Nennern sind in einen parameterabhängigen und einen konstanten Teil aufgespalten worden.

Für den zweiten Summanden der Energiefunktion, dem n-Potential, gilt

$$P_{n_a}\left(\lambda_P, (\mathcal{T}_{\mathcal{A}}(u_{\mathcal{A}}, a_1), \ldots, \mathcal{T}_{\mathcal{A}}(u_{\mathcal{A}}, a_{n_a}), \Sigma_{\mathcal{A}}\right)$$
$$= \frac{8\lambda_p}{n_a(n_a-1)} \sum_{i=1}^{n_a-1} \sum_{j=i+1}^{n_a} \left(\left(\frac{\sqrt{\sigma_i \sigma_j}}{\|\mathcal{T}_{\mathcal{A}}(u_{\mathcal{A}}, a_i) - \mathcal{T}_{\mathcal{A}}(u_{\mathcal{A}}, a_j)\|_2}\right)^{12}\right.$$
$$\left. - \left(\frac{\sqrt{\sigma_i \sigma_j}}{\|\mathcal{T}_{\mathcal{A}}(u_{\mathcal{A}}, a_i) - \mathcal{T}_{\mathcal{A}}(u_{\mathcal{A}}, a_j)\|_2}\right)^6\right)$$

$$\overset{L.2.4.28b)}{=} \frac{8\lambda_p}{n_a(n_a-1)} \sum_{i=0}^{n_{ft}-1} \sum_{j=i+1}^{n_{ft}} \left(\sum_{a_k \in V_i} \sum_{a_l \in V_j} \left(\left(\frac{\sigma_k \sigma_l}{\mathsf{N}_{ij}(\varphi_{\mathcal{A}}, a_k, a_l)} \right)^6 - \left(\frac{\sigma_k \sigma_l}{\mathsf{N}_{ij}(\varphi_{\mathcal{A}}, a_k, a_l)} \right)^3 \right) \right)$$
$$+ \frac{8\lambda_p}{n_a(n_a-1)} \sum_{i=0}^{n_{ft}} \left(\sum_{\substack{a_k, a_l \in V_i \\ a_k \neq a_l}} \left(\left(\frac{\sqrt{\sigma_k \sigma_l}}{\|a_k - a_l\|_2} \right)^{12} - \left(\frac{\sqrt{\sigma_k \sigma_l}}{\|a_k - a_l\|_2} \right)^6 \right) \right),$$

wobei für $i = 0, \ldots, n_{ft}$ $V_i = (\Psi_{K_{\mathcal{A}},F_{\mathcal{A}}})^{-1}(f_i) := \{a_k \in K_{\mathcal{A}} \mid \Psi_{K_{\mathcal{A}},F_{\mathcal{A}}}(a_k) = f_i\}$, das Urbild von f_i in $K_{\mathcal{A}}$ ist. Es gilt $V_i \cap V_j = \emptyset$ für $i \neq j$. Für $i, j = 0, \ldots, n_{ft}$ mit $i \neq j$ ist $\mathsf{N}_{ij} : \mathbb{R}^{n_{ft}} \times V_i \times V_j \to \mathbb{R}$ mit

$$\mathsf{N}_{ij}(\varphi_{\mathcal{A}}, v_1, v_2) := \|\mathsf{R}[\mathcal{S}(K_{\mathcal{A}}, \vec{T}_{\mathcal{A}}), f_i](\Phi[i], v_1) - \mathsf{R}[\mathcal{S}(K_{\mathcal{A}}, \vec{T}_{\mathcal{A}}), f_j](\Phi[j], v_2)\|_2^2$$

die Abbildung aus Lemma 2.4.26, wobei $\Phi[k]$ der Vektor derjenigen lokalen Rotationsparameter φ_l ist, die an $\mathsf{R}[\mathcal{S}(K_{\mathcal{A}}, \vec{T}_{\mathcal{A}}), f_k]$ übergeben werden.

Das n-Potential hängt also nur von den lokalen Rotationsparametern φ_i, $i = 1, \ldots, n_{ft}$, ab. Intuitiv ist dies vollkommen richtig, da die internen Abstoßungs- und Anziehungskräfte unabhängig von der Position im Raum sind.[6] Darüberhinaus konnte das n-Potential in einen flexiblen und einen konstanten Teil aufgespalten werden. Für die Optimierung ist nur der flexible Teil von Interesse. Setzt man nun

$$P_{n_a}^F(\lambda_P, (\mathcal{R}_{\mathcal{A}}(\varphi_{\mathcal{A}}, a_1), \ldots, \mathcal{R}_{\mathcal{A}}(\varphi_{\mathcal{A}}, a_{n_a}), \Sigma_{\mathcal{A}}) :=$$
$$\frac{8\lambda_p}{n_a(n_a-1)} \sum_{i=0}^{n_{ft}-1} \sum_{j=i+1}^{n_{ft}} \left(\sum_{a_k \in V_i} \sum_{a_l \in V_j} \left(\left(\frac{\sigma_k \sigma_l}{\mathsf{N}_{ij}(\varphi_{\mathcal{A}}, a_k, a_l)} \right)^6 - \left(\frac{\sigma_k \sigma_l}{\mathsf{N}_{ij}(\varphi_{\mathcal{A}}, a_k, a_l)} \right)^3 \right) \right)$$

so läßt sich die Energiefunktion wie folgt darstellen

$$\mathsf{E}^*_{\mathcal{A},\mathcal{R}}(u_{\mathcal{A}}) = \lambda_A \, \mathsf{A}^{\mathsf{H}}_{\mathcal{A},\mathcal{R}} \left((Q^p_{\mathcal{A}}[u_{\mathcal{A}}](x))_{p \in \mathcal{I}^{\mathcal{A},\mathcal{R}}_{PE}}, \; (Q^p_{\mathcal{R}}(x))_{p \in \mathcal{I}^{\mathcal{A},\mathcal{R}}_{PE}} \right)$$
$$+ P^F_{n_a}(\lambda_P, (\mathcal{R}_{\mathcal{A}}(\varphi_{\mathcal{A}}, a_1), \ldots, \mathcal{R}_{\mathcal{A}}(\varphi_{\mathcal{A}}, a_{n_a}), \Sigma_{\mathcal{A}}) \; + \; \lambda_T \; s^{\mathcal{A}}_{tor}\left(\Theta^{tor}(\mathsf{C}_{\mathcal{A}}), \varphi_{\mathcal{A}}\right) \; + \; C,$$

wobei C eine Konstante ist. Damit folgt insgesamt, daß nur der Alignmentindex von den globalen Translationsparametern (t, q, c) abhängig ist.

Lemma 3.2.1

Die Energiefunktion $\mathsf{E}^*_{\mathcal{A},\mathcal{R}}$ und damit das Optimierungsproblem (3.2) besitzt keine isolierten Minima. Insbesondere besitzt die Funktion keine isolierten globalen Minima.

Beweis:

Man setzt $\mathsf{E}^*(u) := \mathsf{E}^*_{\mathcal{A},\mathcal{R}}(u_{\mathcal{A}})$ mit $u = u_{\mathcal{A}} \in U(n_a)$. Es $u^* \in U(n_a)$ ein lokales Minimum der Optimierungsaufgabe (3.2). D.h. es existiert ein $\varepsilon > 0$, so daß $\mathsf{E}^*(u^*) \leq \mathsf{E}^*(u)$ für alle $u \in B_\varepsilon(u^*)$ $= \{u \in U(n_a) | \|u - u^*\|_2 < \varepsilon\} \subset U(n_a)$ gilt. Man beachte, daß jedes globale Minimum auch ein lokales Minimum ist. Für $u \in U(n_a)$ gilt $u = (t, q, c, \varphi)$. Nach der Definition der Energiefunktion

[6] Dabei wird vorausgesetzt, daß es zu keinen Interaktionen mit anderen Molekülen kommt. Der Raum ist sozusagen leer.

$\mathsf{E}_{\mathcal{A},\mathcal{R}}$ aus Definition 3.2.2 gilt $\mathsf{E}^*(u) = \mathsf{E}_{\mathcal{A},\mathcal{R}}(u) - \lambda_Q(\|q\|_2 - 1)^2 + C$, wobei C der konstante Anteil des n-Potentials ist. Dann gilt für den Gradienten von E^*

$$\nabla\mathsf{E}^*(u) = \nabla\mathsf{E}_{\mathcal{A},\mathcal{R}}(u) - \sum_{k=1}^{4}\left(\frac{2\lambda_Q q_k(\|q\|_2 - 1)}{\|q\|_2^2}\right) e_{3+k},$$

wobei e_i für $i = 4, 5, 6, 7$ der i-te Einheitsvektor des $\mathbb{R}^{10+n_{ft}}$ und $\lambda_Q > 0$ ist. Für $\lambda \neq 0$ und $u = (t, q, c, \varphi) \in U(n_a)$ setzt man $u_\lambda := (t, \lambda q, c, \varphi)$. Dann folgt für $u \in U(n_a)$ aus Lemma 2.4.29 b), daß

$$\mathsf{E}^*(u_\lambda) = \mathsf{E}^*(u) \ \wedge\ \nabla\mathsf{E}^*(u_\lambda) = \nabla\mathsf{E}^*(u) \tag{3.3}$$

für alle $\lambda \neq 0$ gilt. Insbesondere gilt $0 = \nabla E^*(u^*) = \nabla E^*(u^*_\lambda)\ \forall\lambda \neq 0$. Nun wählt man $\lambda \neq 0$ so, daß

$$\|u^*_\lambda - u^*\|_2 = \|\lambda q^* - q^*\|_2 = |1-\lambda|\,\|q^*\|_2 < \varepsilon$$

gilt. Man setzt z.B. $\lambda^* := 1 - \frac{\varepsilon}{2\|q^*\|_2} \neq 0$.[7] Dann gilt $u^*_{\lambda^*} \in B_\varepsilon(u^*)$ und $E^*(u^*_{\lambda^*}) = E^*(u^*)$. Damit ist die Behauptung bewiesen. ■

Aus Gleichung (3.3) folgt insbesondere, daß es für die analytische Lösung des Optimierungsproblems (3.2) keine Rolle spielt, ob $\|q^*\|_2 = 1$ gilt oder nicht. Für die numerische Lösung sollte man jedoch bedenken, daß die Genauigkeit der Zahlendarstellung im Computer beschränkt ist. Um Rundungsfehler zu vermeiden und damit die Eigenschaft $A(q) = R(\frac{q}{\|q\|_2}) \in \mathsf{SO}^3$ möglichst zu erhalten,[8] sollte $\|q\|_2$ möglichst eins sein. Daher wird die folgende Barrierefunktion verwendet

$$P_Q(q) = \lambda_Q\ (\|q\|_2 - 1)^2,$$

wobei $\lambda_Q > 0$ gilt. Diese Funktion soll dafür sorgen, daß in jedem Schritt eines Optimierungsverfahrens $\|q\|_2$ ungefähr eins bleibt.

Mit den obigen Bezeichnungen läßt sich nun die Energiefunktion wie folgt definieren.

Definition 3.2.2
Die Zielfunktion, mit der man das optimale Alignment eines flexiblen Moleküls $\mathcal{A}$ bzgl. eines rigiden Referenzmoleküls $\mathcal{R}$ bestimmen kann, ist durch

$$\begin{aligned}
\mathsf{E}_{\mathcal{A},\mathcal{R}}(u_{\mathcal{A}}) &:= \lambda_A\ \mathsf{A}^{\mathsf{H}}_{\mathcal{A},\mathcal{R}}\left(\left(Q^p_{\mathcal{A}}[u_{\mathcal{A}}](x)\right)_{p\in\mathcal{I}^{\mathcal{A},\mathcal{R}}_{PE}},\ \left(Q^p_{\mathcal{R}}(x)\right)_{p\in\mathcal{I}^{\mathcal{A},\mathcal{R}}_{PE}} \right) + \lambda_Q\ (\|q\|_2 - 1)^2 \\
&\quad + P^F_{n_a}\left(\lambda_P, (\mathcal{R}_{\mathcal{A}}(\varphi_{\mathcal{A}}, a_1), \ldots, \mathcal{R}_{\mathcal{A}}(\varphi_{\mathcal{A}}, a_{n_a}), \Sigma_{\mathcal{A}})\right) + \lambda_T\ s^{\mathcal{A}}_{tor}\left(\Theta^{tor}(\mathsf{C}_{\mathcal{A}}), \varphi_{\mathcal{A}}\right) \\
&= \lambda_A \sum_{p\in\mathcal{I}^{\mathcal{A},\mathcal{R}}_{PE}} 2\lambda_p \frac{\mathcal{Z}\left(\sum_{i=1}^{n^p_g[\mathcal{A}]} \mathsf{G}\left(h^{\mathcal{A}}_{p,i}, b^{\mathcal{A}}_{p,i}, \mathcal{T}^p_{\mathcal{A}}(u_{\mathcal{A}},\ z^{\mathcal{A}}_{p,i})\right)(x),\ \sum_{j=1}^{n^p_g[\mathcal{R}]} \mathsf{G}\left(h^{\mathcal{R}}_{p,j}, b^{\mathcal{R}}_{p,j}, z^{\mathcal{R}}_{p,j}\right)(x)\right)}{\mathcal{N}\left(\sum_{i=1}^{n^p_g[\mathcal{A}]} \mathsf{G}\left(h^{\mathcal{A}}_{p,i}, b^{\mathcal{A}}_{p,i}, \mathcal{R}^p_{\mathcal{A}}(\varphi_{\mathcal{A}},\ z^{\mathcal{A}}_{p,i})\right)(x)\right) + \mathcal{N}\left(\sum_{j=1}^{n^p_g[\mathcal{R}]} \mathsf{G}\left(h^{\mathcal{R}}_{p,j}, b^{\mathcal{R}}_{p,j}, z^{\mathcal{R}}_{p,j}\right)(x)\right)} \\
&\quad + \frac{8\lambda_P}{n_a(n_a-1)} \sum_{i=1}^{n_{ft}-1} \sum_{j=i+1}^{n_{ft}} \left(\sum_{a_k\in V_i} \sum_{a_l\in V_j} \left(\left(\frac{\sigma_k\sigma_l}{\mathsf{N}_{ij}(\varphi_{\mathcal{A}}, a_k, a_l)}\right)^6 - \left(\frac{\sigma_k\sigma_l}{\mathsf{N}_{ij}(\varphi_{\mathcal{A}}, a_k, a_l)}\right)^3 \right)\right) \\
&\quad + \lambda_Q\ (\|q\|_2 - 1)^2 + \lambda_T\ s^{\mathcal{A}}_{tor}\left(\Theta^{tor}(\mathsf{C}_{\mathcal{A}}), \varphi_{\mathcal{A}}\right)
\end{aligned}$$

[7] Falls $\varepsilon = 2\|q^*\|_2$ gilt, ersetzt man ε durch $\varepsilon/2$.
[8] Siehe die Definition von $A(q)$ in Gleichung (2.10) in Kapitel 2.4.1.

gegeben, wobei $u_{\mathcal{A}} = (t, q, c, \varphi_{\mathcal{A}}) \in U(n_a)$ *gilt. Die Optimierungsaufgabe lautet*

$$\min_{u_{\mathcal{A}} \in U(n_a)} \mathsf{E}_{\mathcal{A},\mathcal{R}}(u_{\mathcal{A}}). \tag{3.4}$$

Das Ziel der Optimierung ist es, $u_{\mathcal{A}} \in U(n_a)$ *so zu bestimmen, daß der Index*

$$\mathsf{S}_{\mathcal{A},\mathcal{R}}(u_{\mathcal{A}}) := \mathsf{A}^{\mathsf{H}}_{\mathcal{A},\mathcal{R}} \left((Q^p_{\mathcal{A}}[u_{\mathcal{A}}](x))_{p \in \mathcal{I}^{\mathcal{A},\mathcal{R}}_{PE}} , \ (Q^p_{\mathcal{R}}(x))_{p \in \mathcal{I}^{\mathcal{A},\mathcal{R}}_{PE}} \right) \tag{3.5}$$

maximal ist, wobei die durch $u_{\mathcal{A}}$ *transformierte Molekülkonformation von* $\mathcal{A}$ *zulässig sein muß. Dabei bewertet* $\mathsf{S}_{\mathcal{A},\mathcal{R}}$ *das Alignment des transformierten Moleküls* $\mathcal{A}$ *bzgl. des Moleküls* $\mathcal{R}$. $\diamond$

Lemma 3.2.3

Die Energiefunktion $\mathsf{E}_{\mathcal{A},\mathcal{R}}$ *aus Definition 3.2.2 hat die folgenden Eigenschaften:*

a) Die Energiefunktion $\mathsf{E}_{\mathcal{A},\mathcal{R}}$ *ist nach unten beschränkt und es gilt*

$$\mathsf{E}_{\mathcal{A},\mathcal{R}}(u_{\mathcal{A}}) \geq \lambda_A - \lambda_P - \lambda_T$$

für alle $u_{\mathcal{A}} \in U(n_a)$.

b) Es sei $U_g = \mathbb{R}^3 \times \mathbb{R}^4 \setminus \{0\} \times \mathbb{R}^3$. *Weiter sei*

$$U_\varphi := \left\{ \varphi_{\mathcal{A}} \in \mathbb{R}^{n_{ft}} \mid \left(\mathcal{N}^p_{\mathcal{A},\mathcal{R}}(\varphi_{\mathcal{A}}) \neq 0 \ \forall p \in \mathcal{I}^{\mathcal{A},\mathcal{R}}_{PE} \right) \wedge \right.$$
$$\left. (\mathsf{N}_{ij}(\varphi_{\mathcal{A}}, a_1, a_2) \neq 0 \ \forall a_1 \in V_i, a_2 \in V_j, i, j = 0, \ldots, n_{ft}, i \neq j) \right\}$$

Dabei ist für $p \in \mathcal{I}^{\mathcal{A},\mathcal{R}}_{PE}$

$$\mathcal{Z}^p_{\mathcal{A},\mathcal{R}}(u_{\mathcal{A}}) := \mathcal{Z} \left(\sum_{i=1}^{n^p_g[\mathcal{A}]} \mathsf{G}\left(h^{\mathcal{A}}_{p,i}, b^{\mathcal{A}}_{p,i}, \mathcal{T}^p_{\mathcal{A}}(u_{\mathcal{A}}, \ z^{\mathcal{A}}_{p,i})\right)(x), \ \sum_{j=1}^{n^p_g[\mathcal{R}]} \mathsf{G}\left(h^{\mathcal{R}}_{p,j}, b^{\mathcal{R}}_{p,j}, z^{\mathcal{R}}_{p,j}\right)(x) \right),$$

$$\mathcal{N}^p_{\mathcal{A},\mathcal{R}}(\varphi_{\mathcal{A}}) := \mathcal{N} \left(\sum_{i=1}^{n^p_g[\mathcal{A}]} \mathsf{G}\left(h^{\mathcal{A}}_{p,i}, b^{\mathcal{A}}_{p,i}, \mathcal{R}^p_{\mathcal{A}}(\varphi_{\mathcal{A}}, \ z^{\mathcal{A}}_{p,i})\right)(x) \right) + \mathcal{N} \left(\sum_{j=1}^{n^p_g[\mathcal{R}]} \mathsf{G}\left(h^{\mathcal{R}}_{p,j}, b^{\mathcal{R}}_{p,j}, z^{\mathcal{R}}_{p,j}\right)(x) \right)$$

gesetzt worden. Es gilt $\tilde{U}(n_a) := U_g \times U_\varphi \subset U(n_a)$. *Für alle* $u_{\mathcal{A}} \in \tilde{U}(n_a)$ *ist die Energiefunktion wohldefiniert. Die Energiefunktion ist in* $u_{\mathcal{A}} \in \tilde{U}(n_a)$ *stetig partiell differenzierbar und die Komponenten des Gradienten* $\nabla \mathsf{E}_{\mathcal{A},\mathcal{R}}$ *sind gegeben durch*

$$\frac{\partial}{\partial (u_{\mathcal{A}})_k} \mathsf{E}_{\mathcal{A},\mathcal{R}}(u_{\mathcal{A}}) = \delta_{k,3+l} \frac{2 \ \lambda_Q \ q_l \ (\|q\|_2 - 1)}{\|q\|_2} + 2 \ \lambda_A \sum_{p \in \mathcal{I}^{\mathcal{A},\mathcal{R}}_{PE}} \left\{ \frac{\lambda_p}{\mathcal{N}^p_{\mathcal{A},\mathcal{R}}(\varphi_{\mathcal{A}})} \sum_{i=1}^{n^p_g[\mathcal{A}]} \right.$$
$$\left. \sum_{j=1}^{n^p_g[\mathcal{R}]} \left[\frac{-\left(b^{\mathcal{A}}_{p,i} h^{\mathcal{A}}_{p,i}\right) \left(b^{\mathcal{R}}_{p,j} h^{\mathcal{R}}_{p,j}\right)}{(b^{\mathcal{A}}_{p,i} + b^{\mathcal{R}}_{p,j})^{\frac{5}{2}}} \ \frac{\partial}{\partial (u_{\mathcal{A}})_k} \mathsf{N}^p_{z^{\mathcal{R}}_{p,j}} \left(u_{\mathcal{A}}, z^{\mathcal{A}}_{p,i}\right) \exp\left(\frac{-b^{\mathcal{A}}_{p,i} b^{\mathcal{R}}_{p,j}}{b^{\mathcal{A}}_{p,i} + b^{\mathcal{R}}_{p,j}} \ \mathsf{N}^p_{z^{\mathcal{R}}_{p,j}} \left(u_{\mathcal{A}}, z^{\mathcal{A}}_{p,i}\right) \right) \right] \right\}$$

für $k = 1, \dots, 10$ und $l = 1, 2, 3, 4$, und

$$
\begin{aligned}
&\frac{\partial}{\partial (u_{\mathcal{A}})_{10+k}} \mathsf{E}_{\mathcal{A},\mathcal{R}}(u_{\mathcal{A}}) = \frac{\partial}{\partial \varphi_k} \mathsf{E}_{\mathcal{A},\mathcal{R}}(u_{\mathcal{A}}) = \\
&\quad 2\ \lambda_A \sum_{p \in \mathcal{I}_{PE}^{\mathcal{A},\mathcal{R}}} \left\{ \frac{\lambda_p}{\mathcal{N}_{\mathcal{A},\mathcal{R}}^p(\varphi_{\mathcal{A}})} \sum_{i=1}^{n_g^p[\mathcal{A}]} \sum_{j=1}^{n_g^p[\mathcal{R}]} \left[\frac{-\left(b_{p,i}^{\mathcal{A}} h_{p,i}^{\mathcal{A}}\right)\left(b_{p,j}^{\mathcal{R}} h_{p,j}^{\mathcal{R}}\right)}{(b_{p,i}^{\mathcal{A}} + b_{p,j}^{\mathcal{R}})^{\frac{5}{2}}} \frac{\partial}{\partial (u_{\mathcal{A}})_{10+k}} \mathsf{N}_{z_{p,j}^{\mathcal{R}}}^p \left(u_{\mathcal{A}}, z_{p,i}^{\mathcal{A}}\right) \right.\right. \\
&\qquad \left. \exp\left(\frac{-b_{p,i}^{\mathcal{A}} b_{p,j}^{\mathcal{R}}}{b_{p,i}^{\mathcal{A}} + b_{p,j}^{\mathcal{R}}}\ \mathsf{N}_{z_{p,j}^{\mathcal{R}}}^p \left(u_{\mathcal{A}}, z_{p,i}^{\mathcal{A}}\right) \right) \right] - \frac{\lambda_p\ \mathcal{Z}_{\mathcal{A},\mathcal{R}}^p(u_{\mathcal{A}})}{\left(\mathcal{N}_{\mathcal{A},\mathcal{R}}^p(\varphi_{\mathcal{A}})\right)^2} \sum_{i=0}^{n_{ft}-1} \sum_{j=i+1}^{n_{ft}} \left[\sum_{z_{p,k}^{\mathcal{A}} \in GC_i^p[\mathcal{A}]} \right. \\
&\qquad \sum_{z_{p,l}^{\mathcal{A}} \in GC_j^p[\mathcal{A}]} \left(\frac{-(b_{p,k}^{\mathcal{A}} h_{p,k}^{\mathcal{A}})(b_{p,l}^{\mathcal{A}} h_{p,l}^{\mathcal{A}})}{(b_{p,k}^{\mathcal{A}} + b_{p,l}^{\mathcal{A}})^{\frac{5}{2}}} \frac{\partial}{\partial \varphi_k} \mathsf{N}_{ij}^p \left(\varphi_{\mathcal{A}}, z_{p,k}^{\mathcal{A}}, z_{p,l}^{\mathcal{A}}\right) \right. \\
&\qquad \left.\left.\left. \exp\left(\frac{-b_{p,k}^{\mathcal{A}} b_{p,l}^{\mathcal{A}}}{b_{p,k}^{\mathcal{A}} + b_{p,l}^{\mathcal{A}}} \mathsf{N}_{ij}^p \left(\varphi_{\mathcal{A}}, z_{p,k}^{\mathcal{A}}, z_{p,l}^{\mathcal{A}}\right) \right) \right) \right] \right\} + \frac{8\lambda_p}{n_a(n_a - 1)} \sum_{i=0}^{n_{ft}-1} \sum_{j=i+1}^{n_{ft}} \left[\sum_{a_k \in V_i} \right. \\
&\qquad \left. \sum_{a_l \in V_j} \frac{\partial}{\partial \varphi_k} \mathsf{N}_{ij}\left(\varphi_{\mathcal{A}}, a_k, a_l\right) \left(\frac{-6\ (\sigma_k \sigma_l)^6}{(\mathsf{N}_{ij}(\varphi_{\mathcal{A}}, a_k, a_l))^7} - \frac{-3\ (\sigma_k \sigma_l)^3}{(\mathsf{N}_{ij}(\varphi_{\mathcal{A}}, a_k, a_l))^4} \right) \right] \\
&\quad + \lambda_T\ \frac{\partial}{\partial \varphi_k}\ s_{tor}^{\mathcal{A}} \left(\Theta^{tor}(\mathsf{C}_{\mathcal{A}}), \varphi_{\mathcal{A}}\right) \ \text{für } k = 1, \dots, n_{ft}.
\end{aligned}
$$

Die Ableitungen $\frac{\partial}{\partial (u_{\mathcal{A}})_k} \mathsf{N}_y^p(u_{\mathcal{A}}, .)$, $k = 1, \dots, 10 + n_{ft}$, $p \in \mathcal{I}_{PE}^{\mathcal{A},\mathcal{R}}$ sind in Lemma 2.4.29 a) dargestellt worden, während die Ableitungen $\frac{\partial}{\partial \varphi_k} \mathsf{N}_{ij}(\varphi_{\mathcal{A}}, .)$, $k = 1, \dots, n_{ft}$, $i, j = 0, \dots, n_{ft}, i \neq j$, und die Ableitungen $\frac{\partial}{\partial \varphi_k} \mathsf{N}_{ij}^p(\varphi_{\mathcal{A}}, .)$, $k = 1, \dots, n_{ft}$, $p \in \mathcal{I}_{PE}^{\mathcal{A},\mathcal{R}}$, $i, j = 0, \dots, n_{ft}, i \neq j$, in Lemma 2.4.26 dargestellt worden sind.

c) *Für jedes lokales Minimum $u^* \in U(n_a)$ der Energiefunktion $\mathsf{E}_{\mathcal{A},\mathcal{R}}$ gilt $\|q^*\|_2 = 1$.*

d) *Wenn die Energiefunktion $\mathsf{E}_{\mathcal{A},\mathcal{R}}$ globale Minima hat, so hat sie mindestens abzählbar viele globale Minima.*

Vermutung: In den meisten Anwendungen wird U_φ und damit auch $\tilde{U}(n_a)$ offen und $\mathbb{R}^{n_{ft}} \setminus U_\varphi$ eine Nullmenge von $\mathbb{R}^{n_{ft}}$ sein, wobei U_φ ganz entscheidend vom konkreten Beispiel abhängt. Jedoch wird man sicher Beispiele konstruieren können, wo dies nicht mehr gilt.

Beweis:

a) Folgt aus der Definition von $\mathsf{E}_{\mathcal{A},\mathcal{R}}$ mit Lemma 2.5.7, Lemma 2.6.3 und Definition 2.7.8 mit Lemma 2.7.9.

b) Im folgenden sei $u = (t, q, c, \varphi) \in \tilde{U}(n_a)$ bzw. $U(n_a)$, wobei $\varphi = (\varphi_1, \dots, \varphi_n)$ gilt, und $\mathsf{E} = \mathsf{E}_{\mathcal{A},\mathcal{R}}$. Die Energiefunktion E ist in $u \in \tilde{U}(n_a)$ wohldefiniert, da nach Definition von $\tilde{U}(n_a)$ offensichtlich $|\mathsf{E}(u)| < \infty$ gilt.
Die exp-Funktion ist in $x \in \mathbb{R}$ stetig differenzierbar und es gilt $\exp'(x) = \exp(x)$. Insbesondere ist $\exp(x)$ eine $\mathsf{C}^1(\mathbb{R})$-Funktion. Nach Lemma 2.4.26 sind die Abbildungen N_{ij}^p, mit $p \in \mathcal{I}_{PE}^{\mathcal{A},\mathcal{R}}$ und $i, j = 0, \dots, n_{ft}, i \neq j$, in $\varphi \in \mathbb{R}^{n_{ft}}$ stetig partiell differenzierbar. Insbesondere sind die Abbildungen N_{ij}^p $\mathsf{C}^1(\mathbb{R}^{n_{ft}})$-Funktionen. Nach Lemma 2.4.29 a) sind die Abbildungen N_y^p, mit

$p \in \mathcal{I}_{PE}^{\mathcal{A},\mathcal{R}}$ und $y \in \mathbb{R}^3$ beliebig aber fest, in $u \in U(n_a)$ stetig partiell differenzierbar. Damit sind díe Abbildungen N_y^p $\mathsf{C}^1(U(n_a))$-Funktionen. Nach der Kettenregel 2.4.23 sind dann auch die Abbildungen

$$h_1 \ \exp\left(b_1 \ \mathsf{N}_y^p(u,\cdot)\right) \quad \text{und} \quad h_2 \ \exp\left(b_2 \ \mathsf{N}_{ij}^p(\varphi,\cdot)\right),$$

mit $p \in \mathcal{I}_{PE}^{\mathcal{A},\mathcal{R}}$, $i,j = 0,\ldots,n_{ft}, i \neq j$, und $y \in \mathbb{R}^3$, $h_1,h_2,b_1,b_2 \in \mathbb{R}$ beliebig aber fest, in $u \in U(n_a)$ bzw. in $\varphi \in \mathbb{R}^{n_{ft}}$ stetig partiell differenzierbar. Die entsprechenden Gradienten sind durch

$$\nabla_u \left(h_1 \ \exp\left(b_1 \ \mathsf{N}_y^p(u,\cdot)\right)\right) \ = \ h_1 b_1 \ \exp\left(b_1 \ \mathsf{N}_y^p(u,\cdot)\right) \nabla_u \mathsf{N}_y^p(u,\cdot)$$

und

$$\nabla_\varphi \left(h_2 \ \exp\left(b_2 \ \mathsf{N}_{ij}^p(\varphi,\cdot)\right)\right) \ = \ h_2 b_2 \ \exp\left(b_2 \ \mathsf{N}_{ij}^p(\varphi,\cdot)\right) \nabla_\varphi \mathsf{N}_{ij}^p(\varphi,\cdot)$$

gegeben. Daraus folgt unmittelbar, daß $\mathcal{Z}_{\mathcal{A},\mathcal{R}}^p(u)$, $p \in \mathcal{I}_{PE}^{\mathcal{A},\mathcal{R}}$, in $u \in U(n_a)$ und $\mathcal{N}_{\mathcal{A},\mathcal{R}}^p(\varphi)$, $p \in \mathcal{I}_{PE}^{\mathcal{A},\mathcal{R}}$, in $\varphi \in \mathbb{R}^{n_{ft}}$ stetig partiell differenzierbar sind. Die Komponenten der Gradienten $\nabla \mathcal{Z}_{\mathcal{A},\mathcal{R}}^p$ bzw. $\nabla \mathcal{N}_{\mathcal{A},\mathcal{R}}^p$ sind durch:

$$\frac{\partial}{\partial u_k} \mathcal{Z}_{\mathcal{A},\mathcal{R}}^p(u) \ = \ \sum_{i=1}^{n_g^p[\mathcal{A}]} \sum_{j=1}^{n_g^p[\mathcal{R}]} \left[\frac{-\left(b_{p,i}^{\mathcal{A}} h_{p,i}^{\mathcal{A}}\right)\left(b_{p,j}^{\mathcal{R}} h_{p,j}^{\mathcal{R}}\right)}{\left(b_{p,i}^{\mathcal{A}} + b_{p,j}^{\mathcal{R}}\right)^{\frac{5}{2}}} \ \frac{\partial}{\partial u_k} \mathsf{N}_{z_{p,j}^{\mathcal{R}}}^p \left(u, z_{p,i}^{\mathcal{A}}\right) \exp\left(\frac{-b_{p,i}^{\mathcal{A}} b_{p,j}^{\mathcal{R}}}{b_{p,i}^{\mathcal{A}} + b_{p,j}^{\mathcal{R}}} \ \mathsf{N}_{z_{p,j}^{\mathcal{R}}}^p \left(u, z_{p,i}^{\mathcal{A}}\right)\right)\right]$$

für $k = 1,\ldots,10+n_{ft}$, und

$$\frac{\partial}{\partial \varphi_k} \mathcal{N}_{\mathcal{A},\mathcal{R}}^p(\varphi) =$$
$$\sum_{i=0}^{n_{ft}-1} \sum_{j=i+1}^{n_{ft}} \left[\sum_{z_{p,k}^{\mathcal{A}} \in GC_i^p[\mathcal{A}]} \ \sum_{z_{p,l}^{\mathcal{A}} \in GC_j^p[\mathcal{A}]} \left(\frac{-(b_{p,k}^{\mathcal{A}} h_{p,k}^{\mathcal{A}})(b_{p,l}^{\mathcal{A}} h_{p,l}^{\mathcal{A}})}{(b_{p,k}^{\mathcal{A}} + b_{p,l}^{\mathcal{A}})^{\frac{5}{2}}} \ \frac{\partial}{\partial \varphi_k} \mathsf{N}_{ij}^p \left(\varphi_{\mathcal{A}}, z_{p,k}^{\mathcal{A}}, z_{p,l}^{\mathcal{A}}\right) \right.\right.$$
$$\left.\left. \exp\left(\frac{-b_{p,k}^{\mathcal{A}} b_{p,l}^{\mathcal{A}}}{b_{p,k}^{\mathcal{A}} + b_{p,l}^{\mathcal{A}}} \mathsf{N}_{ij}^p \left(\varphi_{\mathcal{A}}, z_{p,k}^{\mathcal{A}}, z_{p,l}^{\mathcal{A}}\right)\right)\right)\right] \quad \text{für } k = 1,\ldots,n_{ft},$$

gegeben. Daraus folgt insbesondere, daß $\mathcal{Z}_{\mathcal{A},\mathcal{R}}^p$ bzw. $\mathcal{N}_{\mathcal{A},\mathcal{R}}^p$, $p \in \mathcal{I}_{PE}^{\mathcal{A},\mathcal{R}}$, in $\varphi \in U_\varphi$ stetig partiell differenzierbar sind. Mit der Bemerkung zu Definition C.12 und der Quotientenregel aus Satz C.11 folgt, daß auch $\mathcal{Z}_{\mathcal{A},\mathcal{R}}^p(t,q,c,\varphi)/\mathcal{N}_{\mathcal{A},\mathcal{R}}^q(\varphi)$ in $\varphi \in U_\varphi$ stetig partiell differenzierbar ist. Die partiellen Ableitungen von $\mathcal{Z}_{\mathcal{A},\mathcal{R}}^p(t,q,c,\varphi)/\mathcal{N}_{\mathcal{A},\mathcal{R}}^q(\varphi)$ bzgl. φ sind durch

$$\frac{\partial}{\partial \varphi_k} \frac{\mathcal{Z}_{\mathcal{A},\mathcal{R}}^p(t,q,c,\varphi)}{\mathcal{N}_{\mathcal{A},\mathcal{R}}^p(\varphi)} \ = \ \frac{1}{\mathcal{N}_{\mathcal{A},\mathcal{R}}^p(\varphi)} \ \frac{\partial}{\partial \varphi_k} \mathcal{Z}_{\mathcal{A},\mathcal{R}}^p(t,q,c,\varphi) \ - \ \frac{\mathcal{Z}_{\mathcal{A},\mathcal{R}}^p(t,q,c,\varphi)}{\left(\mathcal{N}_{\mathcal{A},\mathcal{R}}^p(\varphi)\right)^2} \ \frac{\partial}{\partial \varphi_k} \mathcal{N}_{\mathcal{A},\mathcal{R}}^p(\varphi),$$

$k = 1,\ldots,n_{ft}$, gegeben.

Nach Lemma 2.4.26 sind die Abbildungen N_{ij}, mit $i,j = 0,\ldots,n_{ft}, i \neq j$, in $\varphi \in \mathbb{R}^{n_{ft}}$ stetig partiell differenzierbar. Insbesondere sind die Abbildungen N_{ij} $\mathsf{C}^1(\mathbb{R}^{n_{ft}})$-Funktionen. Also sind diese Funktionen auch in $\varphi \in U_\varphi \subset \mathbb{R}^{n_{ft}}$ stetig partiell differenzierbar. Mit der Bemerkung zu Definition C.12 und der Quotientenregel aus Satz C.11 folgt, daß auch $F_{ij}: U_\varphi \times V_i \times V_j \to \mathbb{R}$, mit $i,j = 0,\ldots,n_{ft}, i \neq j$, $h \in \mathbb{R}$ beliebig aber fest, und

$$F_{ij}(\varphi,a_1,a_2) \ := \ \frac{h^6}{\left(\mathsf{N}_{ij}(\varphi,a_1,a_2)\right)^6} \ - \ \frac{h^3}{\left(\mathsf{N}_{ij}(\varphi,a_1.a_2)\right)^3}$$

in $\varphi \in U_\varphi$ stetig partiell differenzierbar und die Komponenten des Gradienten $\nabla_\varphi F_{ij}$ sind durch

$$\frac{\partial}{\partial \varphi_k} F_{ij}(\varphi, a_1, a_2) \;=\; \frac{\partial}{\partial \varphi_k} \mathsf{N}_{ij}(\varphi, a_1, a_2) \left[\frac{-6h^6}{(\mathsf{N}_{ij}(\varphi, a_1, a_2))^7} \;-\; \frac{-3h^3}{(\mathsf{N}_{ij}(\varphi, a_1.a_2))^4} \right],$$

$k = 1, \ldots, n_{ft}$, gegeben.

Aus Lemma 2.7.9 folgt, daß $s^{\mathcal{A}}_{tor}\,(\Theta^{tor}(\mathsf{C}_{\mathcal{A}}), \varphi)$ in $\varphi \in \mathbb{R}^{n_{ft}}$ stetig partiell differenzierbar ist. Die Funktion $P(q) = (\|q\|_2 - 1)^2 = \left(\sqrt{\sum_{i=1}^4 \; q_i^2} - 1 \right)^2$ ist nach der Bemerkung zu Definition C.12 und der Kettenregel aus Satz C.11 in $q \in \mathbb{R}^4 \setminus 0$ stetig partiell differenzierbar und es gilt :

$$\frac{\partial}{\partial q_i} P(q) = 2 q_i \; (\|q\|_2 - 1)/\|q\|_2, \; i = 1, 2, 3, 4.$$

Aus den obigen Betrachtungen folgt nun insgesamt, daß die Energiefunktion E in $u \in \tilde{U}(n_a)$ stetig partiell differenzierbar ist und auch die entsprechende Darstellung des Gradienten.

c) Es sei $u^* = (t^*, q^*, c^*, \varphi^*) \in U(n_a)$ ein lokales Minimum der Energiefunktion $\mathsf{E}_{\mathcal{A},\mathcal{R}}$.

Annahme: $\|q^*\|_2 \neq 1$

Nach Definition 3.1.1 existiert ein ε mit $1 > \varepsilon > 0$, so daß $\mathsf{E}_{\mathcal{A},\mathcal{R}}(u^*) \leq \mathsf{E}_{\mathcal{A},\mathcal{R}}(u)$ für alle $u \in B_\varepsilon(u^*)$. Für $\lambda \neq 0$ setzt man $u^*_\lambda := (t^*, \lambda q^*, c^*, \varphi^*) \in U(n_a)$. Dann gilt

$$\|u^*_\lambda - u^*\|_2 \;=\; \|\lambda q^* - q^*\|_2 \;=\; |\lambda - 1| \, \|q^*\|_2.$$

Fallunterscheidung nach $\|q^*\|_2$:

1.Fall: $\|q^*\|_2 > 1$

Es sei $a_1 := \max\left\{ 1 - \frac{\varepsilon}{\|q^*\|_2}, \; \frac{1}{\|q^*\|_2} \right\}$. Es gilt $0 < a_1 < 1$, wegen $\|q^*\|_2 > 1$. Nun wählt man $\lambda \in (a_1, \; 1) \neq \emptyset$. Dann gilt

$$1 - \frac{\varepsilon}{\|q^*\|_2} \;\leq\; a_1 \;<\; \lambda \;\Leftrightarrow\; |1 - \lambda| \;\; \|q^*\|_2 \;=\; (1 - \lambda) \; \|q^*\|_2 \;<\; \varepsilon.$$

Damit gilt $u^*_\lambda \in B_\varepsilon(u^*)$. Weiter gilt

$$\begin{aligned} \frac{1}{\|q^*\|_2} \;\leq\; a_1 \;<\; \lambda \;<\; 1 \;&\Rightarrow\; 0 \;<\; \|\lambda q^*\|_2 - 1 \;=\; \lambda \, \|q^*\|_2 - 1 \;<\; \|q^*\|_2 - 1 \\ &\Rightarrow\; 0 \;<\; (\|\lambda q^*\|_2 - 1)^2 \;<\; (\|q^*\|_2 - 1)^2 \end{aligned}$$

Daraus folgt:

$$\begin{aligned} \mathsf{E}_{\mathcal{A},\mathcal{R}}(u^*_\lambda) &= \lambda_A \sum_{p \in \mathcal{I}^{\mathcal{A},\mathcal{R}}_{PE}} 2\lambda_p \frac{\mathcal{Z}^p_{\mathcal{A},\mathcal{R}}(u^*_\lambda)}{\mathcal{N}^p_{\mathcal{A},\mathcal{R}}(\varphi^*)} \;+\; \lambda_Q \;\; (\|\lambda q^*\|_2 - 1)^2 \;+\; C(\varphi^*) \\ &\overset{L.2.4.28b)}{=} \lambda_A \sum_{p \in \mathcal{I}^{\mathcal{A},\mathcal{R}}_{PE}} 2\lambda_p \frac{\mathcal{Z}^p_{\mathcal{A},\mathcal{R}}(u^*)}{\mathcal{N}^p_{\mathcal{A},\mathcal{R}}(\varphi^*)} \;+\; \lambda_Q \;\; (\|\lambda q^*\|_2 - 1)^2 \;+\; C(\varphi^*) \\ &< \lambda_A \sum_{p \in \mathcal{I}^{\mathcal{A},\mathcal{R}}_{PE}} 2\lambda_p \frac{\mathcal{Z}^p_{\mathcal{A},\mathcal{R}}(u^*)}{\mathcal{N}^p_{\mathcal{A},\mathcal{R}}(\varphi^*)} \;+\; \lambda_Q \;\; (\|q^*\|_2 - 1)^2 \;+\; C(\varphi^*) \;=\; \mathsf{E}_{\mathcal{A},\mathcal{R}}(u^*), \end{aligned}$$

wobei $C(\varphi^*)$ aus $P^F_{n_a}$ und $s^{\mathcal{A}}_{tor}$ besteht. Damit hat man einen Widerspruch zur Voraussetzung erzeugt, daß u^* ein lokales Minimum ist.

2.Fall: $0 < \|q^*\|_2 < 1$

Es sei $a_2 := \min\left\{1 + \frac{\varepsilon}{\|q^*\|_2},\ \frac{1}{\|q^*\|_2}\right\}$. Es gilt $1 < a_2$, wegen $\|q^*\|_2 < 1$. Nun wählt man $\lambda \in (1,\ a_2) \neq \emptyset$. Dann gilt

$$\lambda \ < \ a_2 \ \leq \ 1 + \frac{\varepsilon}{\|q^*\|_2} \Leftrightarrow |\lambda - 1| \ \|q^*\|_2 \ = \ (\lambda - 1) \ \|q^*\|_2 \ < \ \varepsilon.$$

Damit gilt $u^*_\lambda \in B_\varepsilon(u^*)$. Weiter gilt

$$\begin{aligned} 1 \ < \ \lambda \ < \ a_2 \ \leq \ \frac{1}{\|q^*\|_2} &\Rightarrow \ \|q^*\|_2 - 1 < \ \|\lambda q^*\|_2 - 1 \ = \ \lambda \ \|q^*\|_2 - 1 \ < \ 0 \\ &\Rightarrow \ 0 \ < \ (\|\lambda q^*\|_2 - 1)^2 \ < \ (\|q^*\|_2 - 1)^2 \,. \end{aligned}$$

Analog zum ersten Fall gilt dann $\mathsf{E}_{\mathcal{A},\mathcal{R}}(u^*_\lambda) < \mathsf{E}_{\mathcal{A},\mathcal{R}}(u^*)$. Damit hat man einen Widerspruch zur Voraussetzung erzeugt daß u^* ein lokales Minimum ist.

Insgesamt folgt also $q \in \mathsf{S}^3$.

d) Es sei $u^* = (t^*, q^*, c^*, \varphi^*) \in U(n_a)$, mit $\varphi^* = (\varphi^*_1, \dots, \varphi^*_{n_{ft}})$, ein globales Minimum der Funktion $\mathsf{E}_{\mathcal{A},\mathcal{R}}$, d.h. es gilt $\mathsf{E}_{\mathcal{A},\mathcal{R}}(u^*) \leq \mathsf{E}_{\mathcal{A},\mathcal{R}}(u)$ für alle $u \in U(n_a)$.
Setzt man $\tilde{u} := (t^*, -q^*, c^*, \varphi^*) \in U(n_a)$, so gilt offensichtlich $\mathsf{E}_{\mathcal{A},\mathcal{R}}(\tilde{u}) = \mathsf{E}_{\mathcal{A},\mathcal{R}}(u^*)$. Also ist $\tilde{u}$ ein weiteres globales Minimum. Aus der Definition des lokalen Rotationsoperators (Definition 2.4.18) kann man mit Korollar 2.4.17, Definition 2.4.10 und insbesondere mit Gleichung (2.7) in Kapitel 2.4.1 sehen, daß die lokalen Rotationsparameter $\varphi^*_i \in \mathbb{R}$, $i = 1, \dots, n_{ft}$, Parameter von Sinus- und Cosinus-Funktionen sind. Setzt man nun $\varphi^*(k_1, \dots, k_{n_{ft}}) := (\varphi^*_1(k_1), \dots, \varphi^*_{n_{ft}}(k_{n_{ft}})$ mit $\varphi_i(k_i)^* := \varphi^*_i + 2k_i\pi$, $k_i \in \mathbb{Z}$, so folgt aus der 2π-Periodizität der Sinus- und Cosinus-Funktionen, daß $\mathsf{E}_{\mathcal{A},\mathcal{R}}\left((t^*, q^*, c^*, \varphi^*(k_1, \dots, k_{n_{ft}}))\right) = \mathsf{E}_{\mathcal{A},\mathcal{R}}(u^*)$ für alle $(k_1, \dots, k_{n_{ft}}) \in \mathbb{Z}^{n_{ft}}$ gilt. Damit sind $\left(t^*, q^*, c^*, \varphi^*(k_1, \dots, k_{n_{ft}})\right)$ bzw. $\left(t^*, -q^*, c^*, \varphi^*(k_1, \dots, k_{n_{ft}})\right)$ für alle $(k_1, \dots, k_{n_{ft}}) \in \mathbb{Z}^{n_{ft}}$ globale Minima. ∎

Bemerkung: Aus Lemma 3.2.3 c) folgt, daß die lokalen Minima der Funktion $\mathsf{E}_{\mathcal{A},\mathcal{R}}$ mögliche Minima der Funktion $\mathsf{E}^*_{\mathcal{A},\mathcal{R}}$ sind. Es sei $u^* \in U(n_a)$ ein lokales Minimum von $\mathsf{E}_{\mathcal{A},\mathcal{R}}$. Aus Korollar 3.1.2 folgt mit Lemma 3.2.3 c), daß $0 = \nabla\mathsf{E}_{\mathcal{A},\mathcal{R}}(u^*) = \nabla\mathsf{E}^*_{\mathcal{A},\mathcal{R}}(u^*)$ gilt. Insbesondere gilt aber, daß die globalen Minima der Energiefunktion $\mathsf{E}_{\mathcal{A},\mathcal{R}}$ die globalen Minima der Funktion $\mathsf{E}^*_{\mathcal{A},\mathcal{R}}$ sind. Es sei also $u^* \in U(n_a)$ ein globales Minimum von $\mathsf{E}_{\mathcal{A},\mathcal{R}}$, d.h. $\mathsf{E}_{\mathcal{A},\mathcal{R}}(u^*) \leq \mathsf{E}_{\mathcal{A},\mathcal{R}}(u)$ für alle $u \in U(n_a)$. Da ein globales Minimum auch ein lokales Minimum ist, gilt $0 = \nabla\mathsf{E}_{\mathcal{A},\mathcal{R}}(u^*) = \nabla\mathsf{E}^*_{\mathcal{A},\mathcal{R}}(u^*)$. Aus Lemma 3.2.3 c) folgt $\mathsf{E}_{\mathcal{A},\mathcal{R}}(u^*) = \mathsf{E}^*_{\mathcal{A},\mathcal{R}}(u^*) - C$, wobei C der konstante Anteil des n-Potentials ist. Angenommen, es gibt ein $\tilde{u} = (\tilde{t}, \tilde{q}, \tilde{c}, \tilde{\varphi}) \in U(n_a)$ mit $\mathsf{E}^*_{\mathcal{A},\mathcal{R}}(\tilde{u}) < \mathsf{E}^*_{\mathcal{A},\mathcal{R}}(u^*)$. Wegen Gleichung (3.3) kann man ohne Einschränkung vorraussetzten, daß $\|\tilde{q}\|_2 = 1$ gilt. Daraus folgt nun $\mathsf{E}_{\mathcal{A},\mathcal{R}}(\tilde{u}) = \mathsf{E}^*_{\mathcal{A},\mathcal{R}}(\tilde{u}) - C$. Damit gilt

$$\mathsf{E}^*_{\mathcal{A},\mathcal{R}}(u^*) \ = \ \mathsf{E}_{\mathcal{A},\mathcal{R}}(u^*) + C \ \leq \ \mathsf{E}_{\mathcal{A},\mathcal{R}}(\tilde{u}) + C \ = \ \mathsf{E}^*_{\mathcal{A},\mathcal{R}}(\tilde{u}) \ < \ \mathsf{E}^*_{\mathcal{A},\mathcal{R}}(u^*).$$

Dies ist ein Widerspruch. Also gilt $\mathsf{E}^*_{\mathcal{A},\mathcal{R}}(u^*) \leq \mathsf{E}^*_{\mathcal{A},\mathcal{R}}(u)$ für alle $u \in U(n_a)$. Somit ist es gerechtfertigt, anstatt die Energiefunktion $\mathsf{E}^*_{\mathcal{A},\mathcal{R}}$ die Funktion $\mathsf{E}_{\mathcal{A},\mathcal{R}}$ als Zielfunktion verwenden.

Aus Lemma 3.2.3 c) und aus dem Beweis zu Lemma 3.2.3 d) wird deutlich, daß man sich bei der Suche nach der analytischen Lösung der Optimierungsaufgabe (3.4) bzgl. des Quaternionenvektors q und der lokalen Rotationsparameter φ auf den kompakten Parameterraum $\mathsf{S}^3 \times [0,\ 2\pi]^{n_{ft}}$ beschränken kann. Könnte man sich bei der Suche nach dem Translationsvektor t bzw. Rotationszentrum c ebenfalls auf ein entsprechendes Kompaktum $\mathsf{K}_T^{\mathcal{A}} \subset \mathbb{R}^3$ bzw. $\mathsf{K}_R^{\mathcal{A}} \subset \mathbb{R}^3$ beschränken, so würde

$$\min_{u_{\mathcal{A}} \in U(n_a)} \mathsf{E}_{\mathcal{A},\mathcal{R}}(u_{\mathcal{A}}) = \min_{u_{\mathcal{A}} \in \mathsf{K}_{\mathcal{A}}} \mathsf{E}_{\mathcal{A},\mathcal{R}}(u_{\mathcal{A}})$$

gelten, wobei $\mathsf{K}_{\mathcal{A}} := \mathsf{K}_T^{\mathcal{A}} \times \mathsf{S}^3 \times \mathsf{K}_R^{\mathcal{A}} \times [0,\ 2\pi]^{n_{ft}}$ eine kompakte Teilmenge des $U(n_a) \subset \mathbb{R}^{10+n_{ft}}$ ist. Da jede reellwertige Funktion auf einer kompakten Menge des Definitionsbereiches ihr Maximum und ihr Minimum annimmt, gäbe es dann ein $u_K \in \mathsf{K}_{\mathcal{A}}$ mit $\mathsf{E}_{\mathcal{A},\mathcal{R}}(u_K) = \min_{u_{\mathcal{A}} \in U(n_a)} \mathsf{E}_{\mathcal{A},\mathcal{R}}(u_{\mathcal{A}})$.
Wählt man ein festes Rotationszentrum $c_{\mathcal{A}} \in \mathbb{R}^3$, z.B. den Schwerpunkt der Atomkoordinaten, so kann man den Parameterraum bzgl. des Translationsvektors auf ein geeignetes Kompaktum $K_T \subset \mathbb{R}^3$ beschränken. Denn die (Gauss-)Koordinaten des Molküls $\mathcal{A}$ können nur noch um die durch $\vec{n} \in \mathsf{S}^3$ gegebenen Drehachsen $d[\vec{n}](\lambda) := c_{\mathcal{A}} + \lambda\vec{n}$, $\lambda \in \mathbb{R}$, global rotiert werden. Nun muß $\mathsf{K}_T^{\mathcal{A}}$ so groß gewählt werden, daß die rotierten Koordinaten stets in $\mathsf{K}_T^{\mathcal{A}}$ sind. Dies ist aber möglich, da das Rotationszentrum $c_{\mathcal{A}}$ fest gegeben ist. Und die Translation schiebt die Gausskoordinaten des Moleküls $\mathcal{A}$ auf die Gausskoordinaten des Moleküls $\mathcal{R}$, die fest und unveränderlich im Raum vorgegeben sind. Die kompakte Mege $\mathsf{K}_T^{\mathcal{A}}$ kann unter Umstanden sehr groß sein.

Zum Abschluß dieses Kapitels wird die Verallgemeinerung der Optimierungsaufgabe (3.4) betrachtet. Gegeben sind n_m flexible Moleküle $\mathcal{A}_i$ und ein rigides Molekül $\mathcal{R}$.

Definition 3.2.4
Die Zielfunktion, mit der man das optimale Alignment von n_m flexiblen Molekülen $\mathcal{A}_i$ bzgl. eines rigiden Referenzmoleküls $\mathcal{R}$ bestimmen kann, ist durch

$$\mathsf{E}_{\mathcal{A}_1,\dots,\mathcal{A}_{n_m},\mathcal{R}}(u_{\mathcal{A}_1},\dots,u_{\mathcal{A}_{n_m}}) := \frac{1}{n_m} \sum_{i=1}^{n_m} \mathsf{E}_{\mathcal{A}_i,\mathcal{R}}(u_{\mathcal{A}_i})$$

gegeben, wobei $\mathsf{E}_{\mathcal{A}_i,\mathcal{R}}(u_{\mathcal{A}_i})$, $i = 1,\dots,n_m$, die Energiefunktion aus Definition 3.2.2 ist. Es gilt $u_{\mathcal{A}_i} = (t_{\mathcal{A}_i}, q_{\mathcal{A}_i}, c_{\mathcal{A}_i}, \varphi_{\mathcal{A}_i}) \in U(n_a[\mathcal{A}_i])$. Für $i = 1,\dots,n_m$ ist $U(n_a[\mathcal{A}_i])$ der Parameterraum der zum Molekül $\mathcal{A}_i$ gehörenden Transformationsoperatoren. Setzt man nun $U(n_m) := U(n_a[\mathcal{A}_1]) \times \dots \times U(n_a[\mathcal{A}_{n_m}])$ und $u_{n_m} := (u_{\mathcal{A}_1},\dots,u_{\mathcal{A}_{n_m}}) \in U(n_m)$, so lautet die Optimierungsaufgabe

$$\min_{u_{n_m} \in U(n_m)} \mathsf{E}_{\mathcal{A}_1,\dots,\mathcal{A}_{n_m},\mathcal{R}}(u_{n_m})$$

Das Ziel der Optimierung ist es, $(u_{\mathcal{A}_1},\dots,u_{\mathcal{A}_{n_m}}) \in U(n_m)$ so zu bestimmen, daß für $k = 1,\dots,n_m$ der Index $\mathsf{S}_{\mathcal{A}_k,\mathcal{R}}(u_{\mathcal{A}_k})$ maximal ist, wobei die durch $u_{\mathcal{A}_k}$ transformierte Molekülkonformation zulässig sein muß. Dabei ist $\mathsf{S}_{\mathcal{A}_k,\mathcal{R}}$, $k = 1,\dots,n_m$, der Index aus Gleichung (3.5).
$\diamond$

Lemma 3.2.5
Es sei $U(n_m) := U(n_a[\mathcal{A}_1]) \times \ldots \times U(n_a[\mathcal{A}_{n_m}])$ und $\tilde{U}(n_m) := \tilde{U}(n_a[\mathcal{A}_1]) \times \ldots \times \tilde{U}(n_a[\mathcal{A}_{n_m}])$, wobei $\tilde{U}(n_a)[\mathcal{A}_i]$ für $i = 1, \ldots, n_m$ analog zu Lemma 3.2.3 b) gebildet wird. Die Energiefunktion $\mathsf{E}_{\mathcal{A}_1,\ldots,\mathcal{A}_{n_m},\mathcal{R}}$ aus Definition 3.2.4 hat die folgenden Eigenschaften:

a) *Die Energiefunktion $\mathsf{E}_{\mathcal{A}_1,\ldots,\mathcal{A}_{n_m},\mathcal{R}}$ ist nach unten beschränkt und es gilt*

$$\mathsf{E}_{\mathcal{A}_1,\ldots,\mathcal{A}_{n_m},\mathcal{R}}\left(u_{\mathcal{A}_1}, \ldots, u_{\mathcal{A}_{n_m}}\right) \ \geq \ \lambda_A \ - \lambda_P \ - \lambda_T$$

für alle $(u_{\mathcal{A}_1}, \ldots, u_{\mathcal{A}_{n_m}}) \in U(n_m)$.

b) *Die Engiefunktion ist in $(u_{\mathcal{A}_1}, \ldots, u_{\mathcal{A}_{n_m}}) \in \tilde{U}(n_m)$ stetig partiell differenzierbar und die Komponenten des Gradienten $\nabla \mathsf{E}_{\mathcal{A}_1,\ldots,\mathcal{A}_{n_m},\mathcal{R}}$ sind gegeben durch*

$$\nabla_{u_{\mathcal{A}_i}} \mathsf{E}_{\mathcal{A}_1,\ldots,\mathcal{A}_{n_m},\mathcal{R}}(u_{\mathcal{A}_1}, \ldots, u_{\mathcal{A}_{n_m}}) \ = \ \frac{1}{n_m} \nabla \mathsf{E}_{\mathcal{A}_i,\mathcal{R}}(u_{\mathcal{A}_i}), \quad i = 1, \ldots, n_m.$$

c) *Ist $(u^*_{\mathcal{A}_1}, \ldots, u^*_{\mathcal{A}_{n_m}}) \in U(n_m)$ ein lokales Minimum von $\mathsf{E}_{\mathcal{A}_1,\ldots,\mathcal{A}_{n_m},\mathcal{R}}$, so ist für $i = 1, \ldots, n_m$ $u^*_{\mathcal{A}_i}$ ein lokales Minimum der Energiefunktion $\mathsf{E}_{\mathcal{A}_i,\mathcal{R}}$.*
Ist umgekehrt $i = 1, \ldots, n_m$ $u^*_{\mathcal{A}_i} \in U(n_a[\mathcal{A}_i])$ ein lokales Minimum der Energiefunktion $\mathsf{E}_{\mathcal{A}_i,\mathcal{R}}$, so ist $(u^*_{\mathcal{A}_1}, \ldots, u^*_{\mathcal{A}_{n_m}})$ ein lokales Minimum von $\mathsf{E}_{\mathcal{A}_1,\ldots,\mathcal{A}_{n_m},\mathcal{R}}$.

Beweis:

a) Folgt aus der Definiton von $\mathsf{E}_{\mathcal{A}_1,\ldots,\mathcal{A}_{n_m},\mathcal{R}}$ mit Lemma 2.5.7, Lemma 2.6.3 und Definition 2.7.8 mit Lemma 2.7.9.

b) Folgt unmittelbar aus Lemma 3.2.3 b).

c) Im folgenden sei $\mathsf{E} := \mathsf{E}_{\mathcal{A}_1,\ldots,\mathcal{A}_{n_m},\mathcal{R}}$, und $\mathsf{E}_i := \mathsf{E}_{\mathcal{A}_i,\mathcal{R}}$ für $i = 1, \ldots, n_m$. Dann folgt aus b):

$$\nabla \mathsf{E}\left(u_{\mathcal{A}_1}, \ldots, u_{\mathcal{A}_{n_m}}\right) \ = \ 0 \ \Leftrightarrow \ \nabla \mathsf{E}_i\left(u_{\mathcal{A}_i}\right) \ = \ 0 \quad \text{für } i = 1, \ldots, n_m. \tag{3.6}$$

Es sei $u^* := (u^*_{\mathcal{A}_1}, \ldots, u^*_{\mathcal{A}_{n_m}}) \in U(n_m)$ ein lokales Minimum von E. Aus Gleichung (3.6) folgt, daß $u^*_i := u^*_{\mathcal{A}_i}$ für $i = 1, \ldots, n_m$ das notwendige Kriterium für ein Minimum erfüllt (Korollar 3.1.2). Nach Definition 3.1.1 existiert ein $\varepsilon > 0$, so daß $\mathsf{E}(u^*) \leq \mathsf{E}(u)$ für alle $u \in B_\varepsilon(u^*) \subset U(n_m)$ gilt. Angenommen für ein $i \in \{1, \ldots, n_m\}$ existiert ein $\tilde{u}_i \in B_\varepsilon(u^*_i) \subset U(n_a[\mathcal{A}_i])$ mit $\mathsf{E}_i(\tilde{u}_i) < \mathsf{E}_i(u^*_i)$. Nun setzt man $\tilde{u} := (u^*_1, \ldots, u^*_{i-1}, \tilde{u}_i, u^*_{i+1}, \ldots, u^*_{n_m}) \in U(n_m)$. Wegen $\|\tilde{u} - u^*\|_2 = \|\tilde{u}_i - u^*_i\|_2 < \varepsilon$ gilt $\tilde{u} \in B_\varepsilon(u^*)$. Daraus folgt aber:

$$\mathsf{E}(u^*) \leq \mathsf{E}(\tilde{u}) = \frac{1}{n_m}\left(\sum_{\substack{k=1\\k\neq i}}^{n_m} \mathsf{E}_k(u^*_k) \ + \ \mathsf{E}_i(\tilde{u}_i)\right) < \frac{1}{n_m}\left(\sum_{\substack{k=1\\k\neq i}}^{n_m} \mathsf{E}_k(u^*_k) \ + \ \mathsf{E}_i(u^*_i)\right) = \mathsf{E}(u^*),$$

ein Widerspruch. Also ist für $i = 1, \ldots, n_m$ u^*_i ein lokales Minimum der Energiefunktion E_i. Umgekehrt sei nun für $i = 1, \ldots, n_m$ $u^*_i := u^*_{\mathcal{A}_i} \in U(n_a[\mathcal{A}_i])$ ein lokales Minimum der Energiefunktion E_i. Dann setzt man $u^* := (u^*_1, \ldots, u^*_{n_m})$. Aus (3.6) folgt, daß u^* das notwendige Kriterium für ein Minimum erfüllt. Nach Voraussetzung existiert für $i = 1, \ldots, n_m$ $\varepsilon_i > 0$, so daß $\mathsf{E}_i(u_i^*) \leq \mathsf{E}_i(u_i)$ für alle $u_i \in B_{\varepsilon_i}(x_i^*)$. Dann setzt man $\varepsilon := \min\{\varepsilon_1, \dots, \varepsilon_{n_m}\} > 0$. Für alle $u = (u_1, \dots, u_{n_m}) \in B_\varepsilon(u^*)$ gilt:

$$\mathsf{E}(u^*) = \frac{1}{n_m} \left(\sum_{k=1}^{n_m} \mathsf{E}_k(u_k^*) \right) \leq \frac{1}{n_m} \left(\sum_{k=1}^{n_m} \mathsf{E}_k(u_k) \right) = \mathsf{E}(u).$$

Denn wegen $\|u - u^*\|_2^2 = \sum_{k=1}^{n_m} \|u_k - u_k^*\|_2^2 \geq \|u_i - u_i^*\|_2^2$ für $i = 1, \dots, n_m$ gilt:

$$\|u_i - u_i^*\|_2 \leq \|u - u^*\|_2 \; < \; \varepsilon \; \leq \; \varepsilon_i \quad \text{für } i = 1, \dots, n_m.$$

■

Bemerkung: Aus Lemma 3.2.5 c) folgt, daß man die Aussagen aus Lemma 3.2.3 c) und d) entsprechend auf die Energiefunktion $\mathsf{E}_{\mathcal{A}_1, \dots, \mathcal{A}_{n_m}, \mathcal{R}}$ übertragen kann.

3.3 Modell 2

In Modell 2.2 der Optimierungsaufgabe 2.1.4 werden zwei flexible Moleküle gegenseitig überlagert. Es seien also $\mathcal{A}$ und $\mathcal{B}$ flexible Moleküle mit $n_a[\mathcal{A}]$ bzw. $n_a[\mathcal{B}]$ Atomen. Die Bezeichnungen aus Kapitel 3.2 werden entsprechend übernommen und hier nicht nochmal explizit aufgeführt.

Definition 3.3.1
Die Zielfunktion, mit der man das optimale Alignment zweier flexibler Moleküle $\mathcal{A}$ und $\mathcal{B}$ zueinander bestimmen kann, ist durch

$$\mathsf{E}_{\mathcal{A},\mathcal{B}}(u_\mathcal{A}, u_\mathcal{B}) := \lambda_A \; \mathsf{A}^{\mathsf{H}}_{\mathcal{A},\mathcal{B}} \left(\; (Q^p_\mathcal{A}[u_\mathcal{A}](x))_{p \in \mathcal{I}^{\mathcal{A},\mathcal{B}}_{PE}} \, , \; (Q^p_\mathcal{B}[u_\mathcal{B}](x))_{p \in \mathcal{I}^{\mathcal{A},\mathcal{B}}_{PE}} \; \right)$$

$$+ \frac{1}{2} \left[\lambda_Q \left(\|q_\mathcal{A}\|_2 - 1 \right)^2 + P^F_{n_a[\mathcal{A}]} \left(\lambda_P, (\mathcal{R}_\mathcal{A}(\varphi_\mathcal{A}, a_1^\mathcal{A}), \dots, \mathcal{R}_\mathcal{A}(\varphi_\mathcal{A}, a^\mathcal{A}_{n_a[\mathcal{A}]}), \Sigma_\mathcal{A}) \right) + \lambda_T s^\mathcal{A}_{tor} \left(\Theta^{tor}(\mathsf{C}_\mathcal{A}), \varphi_\mathcal{A} \right) \right]$$
$$+ \frac{1}{2} \left[\lambda_Q \left(\|q_\mathcal{B}\|_2 - 1 \right)^2 + P^F_{n_a[\mathcal{B}]} \left(\lambda_P, (\mathcal{R}_\mathcal{B}(\varphi_\mathcal{B}, a_1^\mathcal{B}), \dots, \mathcal{R}_\mathcal{B}(\varphi_\mathcal{B}, a^\mathcal{B}_{n_a[\mathcal{B}]}), \Sigma_\mathcal{B}) \right) + \lambda_T s^\mathcal{B}_{tor} \left(\Theta^{tor}(\mathsf{C}_\mathcal{B}), \varphi_\mathcal{B} \right) \right]$$

$$= \lambda_A \sum_{p \in \mathcal{I}^{\mathcal{A},\mathcal{B}}_{PE}} 2\lambda_p \frac{\mathcal{Z}\left(\sum_{i=1}^{n_g^p[\mathcal{A}]} \mathsf{G}\left(h^\mathcal{A}_{p,i}, b^\mathcal{A}_{p,i}, \mathcal{T}^p_\mathcal{A}(u_\mathcal{A}, \; z^\mathcal{A}_{p,i})\right)(x), \; \sum_{j=1}^{n_g^p[\mathcal{B}]} \mathsf{G}\left(h^\mathcal{B}_{p,j}, b^\mathcal{B}_{p,j}, \mathcal{T}^p_\mathcal{B}(u_\mathcal{B}, \; z^\mathcal{B}_{p,j})\right)(x) \right)}{\mathcal{N}\left(\sum_{i=1}^{n_g^p[\mathcal{A}]} \mathsf{G}\left(h^\mathcal{A}_{p,i}, b^\mathcal{A}_{p,i}, \mathcal{R}^p_\mathcal{A}(\varphi_\mathcal{A}, \; z^\mathcal{A}_{p,i})\right)(x) \right) + \mathcal{N}\left(\sum_{j=1}^{n_g^p[\mathcal{B}]} \mathsf{G}\left(h^\mathcal{B}_{p,j}, b^\mathcal{B}_{p,j}, \mathcal{R}^p_\mathcal{B}(\varphi_\mathcal{B}, \; z^\mathcal{B}_{p,j})\right)(x) \right)}$$
$$+ \frac{4\lambda_P}{n_a[\mathcal{A}](n_a[\mathcal{A}] - 1)} \sum_{i=1}^{n_{ft}[\mathcal{A}]-1} \sum_{j=i+1}^{n_{ft}[\mathcal{A}]} \left(\sum_{a_k^\mathcal{A} \in V_i[\mathcal{A}]} \sum_{a_l^\mathcal{A} \in V_j[\mathcal{A}]} \left(\left(\frac{\sigma_k \sigma_l}{\mathsf{N}^\mathcal{A}_{ij}(\varphi_\mathcal{A}, a_k^\mathcal{A}, a_l^\mathcal{A})} \right)^6 - \left(\frac{\sigma_k \sigma_l}{\mathsf{N}^\mathcal{A}_{ij}(\varphi_\mathcal{A}, a_k, a_l^\mathcal{A})} \right)^3 \right) \right)$$
$$+ \frac{4\lambda_P}{n_a[\mathcal{B}](n_a[\mathcal{B}] - 1)} \sum_{i=1}^{n_{ft}[\mathcal{B}]-1} \sum_{j=i+1}^{n_{ft}[\mathcal{B}]} \left(\sum_{a_k^\mathcal{B} \in V_i[\mathcal{B}]} \sum_{a_l^\mathcal{B} \in V_j[\mathcal{B}]} \left(\left(\frac{\sigma_k \sigma_l}{\mathsf{N}^\mathcal{B}_{ij}(\varphi_\mathcal{B}, a_k^\mathcal{B}, a_l^\mathcal{B})} \right)^6 - \left(\frac{\sigma_k \sigma_l}{\mathsf{N}^\mathcal{B}_{ij}(\varphi_\mathcal{B}, a_k^\mathcal{B}, a_l^\mathcal{B})} \right)^3 \right) \right)$$
$$+ \frac{1}{2} \left\{ \lambda_Q \;\; (\|q_\mathcal{A}\|_2 - 1)^2 + \lambda_T \; s^\mathcal{A}_{tor} \left(\Theta^{tor}(\mathsf{C}_\mathcal{A}), \varphi_\mathcal{A} \right) + \lambda_Q \;\; (\|q_\mathcal{B}\|_2 - 1)^2 + \lambda_T \; s^\mathcal{B}_{tor} \left(\Theta^{tor}(\mathsf{C}_\mathcal{B}), \varphi_\mathcal{B} \right) \right\}$$

gegeben, wobei $u_\mathcal{A} = (t_\mathcal{A}, q_\mathcal{A}, c_\mathcal{A}, \varphi_\mathcal{A}) \in U(n_a[\mathcal{A}])$ und $u_\mathcal{B} = (t_\mathcal{B}, q_\mathcal{B}, c_\mathcal{B}, \varphi_\mathcal{B}) \in U(n_a[\mathcal{B}])$ gilt. $U(n_a[\mathcal{A}])$ bzw. $U(n_a[\mathcal{B}])$ ist der Parameterraum der entsprechenden Transformationsoperatoren, *die die (Gauss-)Koordinaten des Moleküls $\mathcal{A}$ bzw. $\mathcal{B}$ transformieren. Nach Lemma 2.5.5 gilt*

$$\mathcal{Z}\left(\sum_{i=1}^{n_g^p[\mathcal{A}]} \mathsf{G}\left(h_{p,i}^{\mathcal{A}}, b_{p,i}^{\mathcal{A}}, \mathcal{T}_{\mathcal{A}}^p(u_{\mathcal{A}},\ z_{p,i}^{\mathcal{A}})\right)(x),\ \sum_{j=1}^{n_g^p[\mathcal{B}]} \mathsf{G}\left(h_{p,j}^{\mathcal{B}}, b_{p,j}^{\mathcal{B}}, \mathcal{T}_{\mathcal{B}}^p(u_{\mathcal{B}},\ z_{p,j}^{\mathcal{B}})\right)(x)\right)$$
$$= \sum_{i=1}^{n_g^p[\mathcal{A}]} \sum_{j=1}^{n_g^p[\mathcal{B}]} \left(\frac{h_{p,i}^{\mathcal{A}} h_{p,j}^{\mathcal{B}}}{(b_{p,i}^{\mathcal{A}} + b_{p,j}^{\mathcal{B}})^{\frac{3}{2}}} \exp\left(\frac{-b_{p,i}^{\mathcal{A}} b_{p,j}^{\mathcal{B}}}{b_{p,i}^{\mathcal{A}} + b_{p,j}^{\mathcal{B}}} \|\mathcal{T}_{\mathcal{A}}^p(u_{\mathcal{A}},\ z_{p,i}^{\mathcal{A}}) - \mathcal{T}_{\mathcal{B}}^p(u_{\mathcal{B}},\ z_{p,j}^{\mathcal{B}})\|_2^2\right)\right).$$

Nach Lemma 2.5.5 und Lemma 2.4.28 b) gilt

$$\mathcal{N}\left(\sum_{i=1}^{n_g^p[\mathcal{A}]} \mathsf{G}\left(h_{p,i}^{\mathcal{A}}, b_{p,i}^{\mathcal{A}}, \mathcal{R}_{\mathcal{A}}^p(u_{\mathcal{A}},\ z_{p,i}^{\mathcal{A}})\right)(x)\right)$$
$$= \sum_{i=0}^{n_{ft}[\mathcal{A}]-1} \sum_{j=i+1}^{n_{ft}[\mathcal{A}]} \left[\sum_{z_{p,k}^{\mathcal{A}} \in GC_i^p[\mathcal{A}]} \sum_{z_{p,l}^{\mathcal{A}} \in GC_j^p[\mathcal{A}]} \left(\frac{h_{p,k}^{\mathcal{A}} h_{p,l}^{\mathcal{A}}}{(b_{p,k}^{\mathcal{A}} + b_{p,l}^{\mathcal{A}})^{\frac{3}{2}}} \exp\left(\frac{-b_{p,k}^{\mathcal{A}} b_{p,l}^{\mathcal{A}}}{b_{p,k}^{\mathcal{A}} + b_{p,l}^{\mathcal{A}}} \mathsf{N}_{ij}^{\mathcal{A},p}\left(\varphi_{\mathcal{A}}, z_{p,k}^{\mathcal{A}}, z_{p,l}^{\mathcal{A}}\right)\right)\right)\right]$$
$$+ \sum_{i=0}^{n_{ft}[\mathcal{A}]} \left[\sum_{z_{p,k}^{\mathcal{A}}, z_{p,l}^{\mathcal{A}} \in GC_i^p[\mathcal{A}]} \left(\frac{h_{p,k}^{\mathcal{A}} h_{p,l}^{\mathcal{A}}}{(b_{p,k}^{\mathcal{A}} + b_{p,l}^{\mathcal{A}})^{\frac{3}{2}}} \exp\left(\frac{-b_{p,k}^{\mathcal{A}} b_{p,l}^{\mathcal{A}}}{b_{p,k}^{\mathcal{A}} + b_{p,l}^{\mathcal{A}}} \|z_{p,k}^{\mathcal{A}} - z_{p,l}^{\mathcal{A}}\|_2^2\right)\right)\right].$$

und

$$\mathcal{N}\left(\sum_{i=1}^{n_g^p[\mathcal{B}]} \mathsf{G}\left(h_{p,i}^{\mathcal{B}}, b_{p,i}^{\mathcal{B}}, \mathcal{R}_{\mathcal{B}}^p(u_{\mathcal{B}},\ z_{p,i}^{\mathcal{B}})\right)(x)\right)$$
$$= \sum_{i=0}^{n_{ft}[\mathcal{B}]-1} \sum_{j=i+1}^{n_{ft}[\mathcal{B}]} \left[\sum_{z_{p,k}^{\mathcal{B}} \in GC_i^p[\mathcal{B}]} \sum_{z_{p,l}^{\mathcal{B}} \in GC_j^p[\mathcal{B}]} \left(\frac{h_{p,k}^{\mathcal{B}} h_{p,l}^{\mathcal{B}}}{(b_{p,k}^{\mathcal{B}} + b_{p,l}^{\mathcal{B}})^{\frac{3}{2}}} \exp\left(\frac{-b_{p,k}^{\mathcal{B}} b_{p,l}^{\mathcal{B}}}{b_{p,k}^{\mathcal{B}} + b_{p,l}^{\mathcal{B}}} \mathsf{N}_{ij}^{\mathcal{B},p}\left(\varphi_{\mathcal{B}}, z_{p,k}^{\mathcal{B}}, z_{p,l}^{\mathcal{B}}\right)\right)\right)\right]$$
$$+ \sum_{i=0}^{n_{ft}[\mathcal{B}]} \left[\sum_{z_{p,k}^{\mathcal{B}}, z_{p,l}^{\mathcal{A}} \in GC_i^p[\mathcal{B}]} \left(\frac{h_{p,k}^{\mathcal{B}} h_{p,l}^{\mathcal{B}}}{(b_{p,k}^{\mathcal{B}} + b_{p,l}^{\mathcal{B}})^{\frac{3}{2}}} \exp\left(\frac{-b_{p,k}^{\mathcal{B}} b_{p,l}^{\mathcal{B}}}{b_{p,k}^{\mathcal{B}} + b_{p,l}^{\mathcal{B}}} \|z_{p,k}^{\mathcal{B}} - z_{p,l}^{\mathcal{B}}\|_2^2\right)\right)\right].$$

Die Optimierungsaufgabe lautet

$$\min_{(u_{\mathcal{A}}, u_{\mathcal{B}})\ \in\ U(n_a[\mathcal{A}]) \times U(n_a[\mathcal{B}])} \mathsf{E}_{\mathcal{A},\mathcal{B}}(u_{\mathcal{A}}, u_{\mathcal{B}}). \tag{3.7}$$

Das Ziel der Optimierung ist es, $(u_{\mathcal{A}}, u_{\mathcal{B}}) \in U(n_a[\mathcal{A}]) \times U(n_a[\mathcal{B}])$ so zu bestimmen, daß der Index

$$\mathsf{S}_{\mathcal{A},\mathcal{B}}(u_{\mathcal{A}}, u_{\mathcal{B}}) \ :=\ \mathsf{A}_{\mathcal{A},\mathcal{B}}^{\mathsf{H}}\left(\ (Q_{\mathcal{A}}^p[u_{\mathcal{A}}](x))_{p \in \mathcal{I}_{PE}^{\mathcal{A},\mathcal{B}}}\ ,\ (Q_{\mathcal{B}}^p[u_{\mathcal{B}}](x))_{p \in \mathcal{I}_{PE}^{\mathcal{A},\mathcal{B}}}\ \right)$$

maximal ist, wobei die durch $(u_{\mathcal{A}}, u_{\mathcal{B}})$ transformierten Konformationen zulässig sein müssen. Dabei bewertet $\mathsf{S}_{\mathcal{A},\mathcal{B}}$ das gegenseitige Alignment der entsprechenden transformierten Moleküle $\mathcal{A}$ und $\mathcal{B}$. $\diamond$

Lemma 3.3.2

Die Energiefunktion $\mathsf{E}_{\mathcal{A},\mathcal{B}}$ aus Definition 3.3.1 hat die folgenden Eigenschaften:

a) Die Energiefunktion $\mathsf{E}_{\mathcal{A},\mathcal{B}}$ ist nach unten beschränkt und es gilt

$$\mathsf{E}_{\mathcal{A},\mathcal{B}}(u_{\mathcal{A}}, u_{\mathcal{B}})\ \geq\ \lambda_A\ - \lambda_P\ - \lambda_T$$

für alle $(u_{\mathcal{A}}, u_{\mathcal{B}}) \in U(n_a[\mathcal{A}]) \times U(n_a[\mathcal{B}])$.

b) Es sei $U_g = \mathbb{R}^3 \times \mathbb{R}^4 \setminus \{0\} \times \mathbb{R}^3$. *Weiter sei*

$$U_{\mathcal{A}} := \Big\{ \varphi_{\mathcal{A}} \in \mathbb{R}^{n_{ft}[\mathcal{A}]} \mid \Big(N\left(Q^p_{\mathcal{A}}(\varphi_{\mathcal{A}})\right) > 0 \ \forall p \in \mathcal{I}^{\mathcal{A},\mathcal{B}}_{PE} \Big) \ \wedge$$
$$\left(\mathsf{N}^{\mathcal{A}}_{ij}(\varphi_{\mathcal{A}}, a^{\mathcal{A}}_1, a^{\mathcal{A}}_2) \neq 0 \ \forall a^{\mathcal{A}}_1 \in V_i[\mathcal{A}], a^{\mathcal{A}}_2 \in V_j[\mathcal{A}], i,j = 0, \dots, n_{ft}[\mathcal{A}], i \neq j \right) \Big\}$$

Nach Gleichung (2.19) in Lemma 2.5.4 gilt für $p \in \mathcal{I}^{\mathcal{A},\mathcal{B}}_{PE}$:

$$N\left(Q^p_{\mathcal{A}}(\varphi_{\mathcal{A}})\right) := \Big\| \sum_{i=1}^{n^p_g[\mathcal{A}]} G\left(h^{\mathcal{A}}_{p,i}, b^{\mathcal{A}}_{p,i}, \mathcal{R}^p_{\mathcal{A}}(\varphi_{\mathcal{A}}, \ z^{\mathcal{A}}_{p,i})\right)(x) \Big\|^2_{\mathsf{L}^2(\mathbb{R}^3)} \ \pi^{-\frac{3}{2}}$$
$$= \mathcal{N}\left(\sum_{i=1}^{n^p_g[\mathcal{A}]} \mathsf{G}\left(h^{\mathcal{A}}_{p,i}, b^{\mathcal{A}}_{p,i}, \mathcal{R}^p_{\mathcal{A}}(\varphi_{\mathcal{A}}, \ z^{\mathcal{A}}_{p,i})\right)(x) \right).$$

Analog wird $U_{\mathcal{B}}$ *gebildet. Dann gilt* $\tilde{U}(n_a[\mathcal{A}]) := U_g \times U_{\mathcal{A}} \subset U(n_a[\mathcal{A}])$ *und* $\tilde{U}(n_a[\mathcal{B}]) := U_g \times U_{\mathcal{B}} \subset U(n_a[\mathcal{B}])$. *Für alle* $(u_{\mathcal{A}}, u_{\mathcal{B}}) \in \tilde{U}(n_a[\mathcal{A}]) \times \tilde{U}(n_a[\mathcal{B}])$ *ist die Energiefunktion wohldefiniert. Weiter ist die Engiefunktion in* $(u_{\mathcal{A}}, u_{\mathcal{B}}) \in \tilde{U}(n_a[\mathcal{A}]) \times \tilde{U}(n_a[\mathcal{B}])$ *stetig partiell differenzierbar und die Komponenten des Gradienten* $\nabla \mathsf{E}_{\mathcal{A},\mathcal{B}}$ *sind gegeben durch*

$$\frac{\partial}{\partial (u_{\mathcal{A}})_k} \mathsf{E}_{\mathcal{A},\mathcal{B}}(u_{\mathcal{A}}, u_{\mathcal{B}}) = \delta_{k,3+l} \frac{2\ \lambda_Q\ (q_{\mathcal{A}})_l\ (\|q_{\mathcal{A}}\|_2 - 1)}{\|q_{\mathcal{A}}\|_2} + 2\ \lambda_A \sum_{p \in \mathcal{I}^{\mathcal{A},\mathcal{B}}_{PE}}$$
$$\left\{ \frac{\lambda_p}{\mathcal{N}^p_{\mathcal{A},\mathcal{B}}(\varphi_{\mathcal{A}}, \varphi_{\mathcal{B}})} \sum_{i=1}^{n^p_g[\mathcal{A}]} \sum_{j=1}^{n^p_g[\mathcal{B}]} \left[\frac{-\left(b^{\mathcal{A}}_{p,i} h^{\mathcal{A}}_{p,i}\right)\left(b^{\mathcal{B}}_{p,j} h^{\mathcal{B}}_{p,j}\right)}{(b^{\mathcal{A}}_{p,i} + b^{\mathcal{B}}_{p,j})^{\frac{5}{2}}} \frac{\partial}{\partial (u_{\mathcal{A}})_k} \mathsf{N}^{\mathcal{A},p}_{\mathcal{T}^p_{\mathcal{B}}(u_{\mathcal{B}}, z^{\mathcal{B}}_{p,j})}\left(u_{\mathcal{A}}, z^{\mathcal{A}}_{p,i}\right) \right.\right.$$
$$\left.\left. \exp\left(\frac{-b^{\mathcal{A}}_{p,i} b^{\mathcal{B}}_{p,j}}{b^{\mathcal{A}}_{p,i} + b^{\mathcal{B}}_{p,j}} \mathsf{N}^{\mathcal{A},p}_{\mathcal{T}^p_{\mathcal{B}}(u_{\mathcal{B}}, z^{\mathcal{B}}_{p,j})}\left(u_{\mathcal{A}}, z^{\mathcal{A}}_{p,i}\right) \right) \right] \right\}$$

für $k = 1, \dots, 10,\ l = 1,2,3,4,$ *und*

$$\frac{\partial}{\partial (u_{\mathcal{A}})_{10+k}} \mathsf{E}_{\mathcal{A},\mathcal{B}}(u_{\mathcal{A}}, u_{\mathcal{B}}) = \frac{\partial}{\partial (\varphi_{\mathcal{A}})_k} \mathsf{E}_{\mathcal{A},\mathcal{B}}(u_{\mathcal{A}}, u_{\mathcal{B}})$$
$$= 2\ \lambda_A \sum_{p \in \mathcal{I}^{\mathcal{A},\mathcal{B}}_{PE}} \left\{ \frac{\lambda_p}{\mathcal{N}^p_{\mathcal{A},\mathcal{B}}(\varphi_{\mathcal{A}}, \varphi_{\mathcal{B}})} \sum_{i=1}^{n^p_g[\mathcal{A}]} \sum_{j=1}^{n^p_g[\mathcal{B}]} \left[\frac{-\left(b^{\mathcal{A}}_{p,i} h^{\mathcal{A}}_{p,i}\right)\left(b^{\mathcal{B}}_{p,j} h^{\mathcal{B}}_{p,j}\right)}{(b^{\mathcal{A}}_{p,i} + b^{\mathcal{B}}_{p,j})^{\frac{5}{2}}} \right.\right.$$
$$\left. \frac{\partial}{\partial (u_{\mathcal{A}})_{10+k}} \mathsf{N}^{\mathcal{A},p}_{\mathcal{T}^p_{\mathcal{B}}(u_{\mathcal{B}}, z^{\mathcal{B}}_{p,j})}\left(u_{\mathcal{A}}, z^{\mathcal{A}}_{p,i}\right) \exp\left(\frac{-b^{\mathcal{A}}_{p,i} b^{\mathcal{B}}_{p,j}}{b^{\mathcal{A}}_{p,i} + b^{\mathcal{B}}_{p,j}} \mathsf{N}^{\mathcal{A},p}_{\mathcal{T}^p_{\mathcal{B}}(u_{\mathcal{B}}, z^{\mathcal{B}}_{p,j})}\left(u_{\mathcal{A}}, z^{\mathcal{A}}_{p,i}\right) \right) \right]$$
$$- \frac{\lambda_p\ \mathcal{Z}^p_{\mathcal{A},\mathcal{B}}(u_{\mathcal{A}}, u_{\mathcal{B}})}{\left(\mathcal{N}^p_{\mathcal{A},\mathcal{B}}(\varphi_{\mathcal{A}}, \varphi_{\mathcal{B}})\right)^2} \sum_{i=0}^{n_{ft}[\mathcal{A}]-1} \sum_{j=i+1}^{n_{ft}[\mathcal{A}]} \left[\sum_{z^{\mathcal{A}}_{p,k} \in GC^p_i[\mathcal{A}]} \sum_{z^{\mathcal{A}}_{p,l} \in GC^p_j[\mathcal{A}]} \left(\frac{-(b^{\mathcal{A}}_{p,k} h^{\mathcal{A}}_{p,k})(b^{\mathcal{A}}_{p,l} h^{\mathcal{A}}_{p,l})}{(b^{\mathcal{A}}_{p,k} + b^{\mathcal{A}}_{p,l})^{\frac{5}{2}}} \right.\right.$$
$$\left.\left.\left. \frac{\partial}{\partial \varphi_k} \mathsf{N}^{\mathcal{A},p}_{ij}\left(\varphi_{\mathcal{A}}, z^{\mathcal{A}}_{p,k}, z^{\mathcal{A}}_{p,l}\right) \exp\left(\frac{-b^{\mathcal{A}}_{p,k} b^{\mathcal{A}}_{p,l}}{b^{\mathcal{A}}_{p,k} + b^{\mathcal{A}}_{p,l}} \mathsf{N}^{\mathcal{A},p}_{ij}\left(\varphi_{\mathcal{A}}, z^{\mathcal{A}}_{p,k}, z^{\mathcal{A}}_{p,l}\right) \right) \right) \right] \right\}$$
$$+ \frac{8\lambda_p}{n_a[\mathcal{A}](n_a[\mathcal{A}] - 1)} \sum_{i=0}^{n_{ft}[\mathcal{A}]-1} \sum_{j=i+1}^{n_{ft}[\mathcal{A}]} \left[\sum_{a_k \in V_i[\mathcal{A}]} \sum_{a_l \in V_j[\mathcal{A}]} \frac{\partial}{\partial \varphi_k} \mathsf{N}^{\mathcal{A}}_{ij}\left(\varphi_{\mathcal{A}}, a_k, a_l\right) \right.$$
$$\left. \left(\frac{-6\ (\sigma_k \sigma_l)^6}{(\mathsf{N}^{\mathcal{A}}_{ij}(\varphi_{\mathcal{A}}, a_k, a_l))^7} - \frac{-3\ (\sigma_k \sigma_l)^3}{(\mathsf{N}^{\mathcal{A}}_{ij}(\varphi_{\mathcal{A}}, a_k, a_l))^4} \right) \right] + \lambda_T \frac{\partial}{\partial \varphi_k} s^{\mathcal{A}}_{tor}\left(\Theta^{tor}(\mathsf{C}_{\mathcal{A}}), \varphi_{\mathcal{A}}\right)$$

für $k = 1,\dots,n_{ft}[\mathcal{A}]$. Man beachte, daß $\mathcal{T}^p_{\mathcal{B}}(u_{\mathcal{B}}, z^{\mathcal{B}}_{p,j})$, $p \in \mathcal{I}^{\mathcal{A},\mathcal{B}}_{PE}$, $j = 1,\dots,n^p_g[\mathcal{B}]$, zwar beliebig, aber für jedes $u_{\mathcal{B}} \in \tilde{U}(n_a[\mathcal{B}])$ fest ist. Man erhält also eine ganze Schar von partiellen Ableitungen. Für $\mathcal{M} = \mathcal{A},\mathcal{B}$ werden die Ableitungen $\frac{\partial}{\partial(u_{\mathcal{M}})_k}\mathsf{N}^{\mathcal{M},p}_y(u_{\mathcal{M}},.)$, $k = 1,\dots,10+n_{ft}[\mathcal{M}]$, $p \in \mathcal{I}^{\mathcal{A},\mathcal{B}}_{PE}$ in Lemma 2.4.29 a) dargestellt, während die Ableitungen $\frac{\partial}{\partial(\varphi_{\mathcal{M}})_k}\mathsf{N}^{\mathcal{M}}_{ij}(\varphi_{\mathcal{M}},.)$, $k = 1,\dots,n_{ft}[\mathcal{M}]$, $i,j = 0,\dots,n_{ft}[\mathcal{M}], i \neq j$, sowie die Ableitungen $\frac{\partial}{\partial(\varphi_{\mathcal{M}})_k}\mathsf{N}^{\mathcal{M},p}_{ij}(\varphi_{\mathcal{M}},.)$, $k = 1,\dots,n_{ft}[\mathcal{M}]$, $p \in \mathcal{I}^{\mathcal{A},\mathcal{B}}_{PE}$, $i,j = 0,\dots,n_{ft}[\mathcal{M}], i \neq j$, in Lemma 2.4.26 dargestellt werden. Um die Darstellung zu vereinfachen, ist für $p \in \mathcal{I}^{\mathcal{A},\mathcal{B}}_{PE}$

$$\begin{aligned}
\mathcal{Z}^p_{\mathcal{A},\mathcal{B}}(u_{\mathcal{A}}, u_{\mathcal{B}}) &:= \mathcal{Z}\left(\sum_{i=1}^{n^p_g[\mathcal{A}]} \mathsf{G}\left(h^{\mathcal{A}}_{p,i}, b^{\mathcal{A}}_{p,i}, \mathcal{T}^p_{\mathcal{A}}(u_{\mathcal{A}},\ z^{\mathcal{A}}_{p,i})\right)(x),\ \sum_{j=1}^{n^p_g[\mathcal{B}]} \mathsf{G}\left(h^{\mathcal{B}}_{p,j}, b^{\mathcal{B}}_{p,j}, \mathcal{T}^p_{\mathcal{B}}(u_{\mathcal{B}},\ z^{\mathcal{B}}_{p,j})\right)(x)\right), \\
\mathcal{N}^p_{\mathcal{A},\mathcal{B}}(\varphi_{\mathcal{A}}, \varphi_{\mathcal{B}}) &:= \mathcal{N}\left(\sum_{i=1}^{n^p_g[\mathcal{A}]} \mathsf{G}\left(h^{\mathcal{A}}_{p,i}, b^{\mathcal{A}}_{p,i}, \mathcal{R}^p_{\mathcal{A}}(\varphi_{\mathcal{A}},\ z^{\mathcal{A}}_{p,i})\right)(x)\right) \\
&\quad + \mathcal{N}\left(\sum_{j=1}^{n^p_g[\mathcal{B}]} \mathsf{G}\left(h^{\mathcal{B}}_{p,j}, b^{\mathcal{B}}_{p,j}, \mathcal{R}^p_{\mathcal{B}}(\varphi_{\mathcal{B}},\ z^{\mathcal{B}}_{p,j})\right)(x)\right)
\end{aligned}$$

gesetzt worden. Die Ableitungen bzgl. $u_{\mathcal{B}}$ werden analog gebildet.

c) *Für jedes lokales Minimum $(u^*_{\mathcal{A}}, u^*_{\mathcal{B}}) \in U(n_a[\mathcal{A}]) \mathsf{x} U(n_a[\mathcal{B}])$ der Energiefunktion $\mathsf{E}_{\mathcal{A},\mathcal{B}}$ gilt $\|q^*_{\mathcal{A}}\|_2 = 1$ und $\|q^*_{\mathcal{B}}\|_2 = 1$.*

d) *Die Energiefunktion $\mathsf{E}_{\mathcal{A},\mathcal{B}}$ besitzt keine isolierten Minima.*

Beweis:

a) Folgt aus der Definiton von $\mathsf{E}_{\mathcal{A},\mathcal{B}}$ mit Lemma 2.5.7, Lemma 2.6.3 und Definition 2.7.8 mit Lemma 2.7.9.

b) Der Beweis und die Darstellung der partiellen Ableitungen folgen analog zum Beweis von Lemma 3.2.3 b).

c) Der Beweis folgt analog zum Beweis von Lemma 3.2.3 c).

d) Es sei $(u^*_{\mathcal{A}}, u^*_{\mathcal{B}}) \in U(n_a[\mathcal{A}]) \mathsf{x} U(n_a[\mathcal{B}])$ ein lokales Minimum der Energiefunktion $\mathsf{E} := \mathsf{E}_{\mathcal{A},\mathcal{B}}$. Nach Definition 3.1.1 existiert ein $\varepsilon > 0$ so, daß $\mathsf{E}(u^*_{\mathcal{A}}, u^*_{\mathcal{B}}) \leq \mathsf{E}(u_{\mathcal{A}}, u_{\mathcal{B}})$ für alle $(u_{\mathcal{A}}, u_{\mathcal{B}}) \in B_\varepsilon((u^*_{\mathcal{A}}, u^*_{\mathcal{B}})) \subset U(n_a[\mathcal{A}]) \mathsf{x} U(n_a[\mathcal{B}])$.
Nun gilt für $p \in \mathcal{I}^{\mathcal{A},\mathcal{B}}_{PE}$, $i = 1,\dots,n^p_g[\mathcal{A}]$, und $j = 1,\dots,n^p_g[\mathcal{B}]$:

$$\begin{aligned}
&\|\mathcal{T}^p_{\mathcal{A}}(u_{\mathcal{A}},\ z^{\mathcal{A}}_{p,i}) - \mathcal{T}^p_{\mathcal{B}}(u_{\mathcal{B}},\ z^{\mathcal{B}}_{p,j})\|^2_2 \\
&\quad = \|A(q_{\mathcal{A}})\left(\mathcal{R}^p_{\mathcal{A}}(\varphi_{\mathcal{A}}, z^{\mathcal{A}}_{p,i}) - c_{\mathcal{A}}\right) + c_{\mathcal{A}} + t_{\mathcal{A}} - A(q_{\mathcal{B}})\left(\mathcal{R}^p_{\mathcal{B}}(\varphi_{\mathcal{B}}, z^{\mathcal{B}}_{p,j}) - c_{\mathcal{B}}\right) + c_{\mathcal{B}} + t_{\mathcal{B}}\|_2.
\end{aligned}$$

Ersetzt man nun $t_{\mathcal{A}}$ durch $t_{\mathcal{A}} + \gamma e_1$, wobei $\gamma \in \mathbb{R}$ und $e_1 = (1,0,0)^{\mathsf{T}}$ ist,[9] so erhält man für $u_{\mathcal{A}} = (t_{\mathcal{A}}, q_{\mathcal{A}}, c_{\mathcal{A}}, \varphi_{\mathcal{A}}) \in U(n_a[\mathcal{A}])$ $u^\gamma_{\mathcal{A}} := (t_{\mathcal{A}} + \gamma e_1, q_{\mathcal{A}}, c_{\mathcal{A}}, \varphi_{\mathcal{A}}) \in U(n_a[\mathcal{A}])$. Analog wird $u^\gamma_{\mathcal{B}}$ gebildet. Dann gilt:

$$\|(u_{\mathcal{A}}, u_{\mathcal{B}}) - (u^\gamma_{\mathcal{A}}, u^\gamma_{\mathcal{B}})\|_2 = \sqrt{\|u_{\mathcal{A}} - u^\gamma_{\mathcal{A}}\|^2_2 + \|u_{\mathcal{B}} - u^\gamma_{\mathcal{B}}\|^2_2} = \sqrt{2\|\gamma e_1\|^2_2} = \sqrt{2}\,|\gamma|\,.$$

[9]Anstatt e_1 kann man auch $e_2 = (0,1,0)^{\mathsf{T}}$ bzw. $e_3 = (0,0,1)^{\mathsf{T}}$ nehmen.

Offensichtlich gilt nun:

$$\|\mathcal{T}^p_{\mathcal{A}}(u_{\mathcal{A}},\ z^{\mathcal{A}}_{p,i}) - \mathcal{T}^p_{\mathcal{B}}(u_{\mathcal{B}},\ z^{\mathcal{B}}_{p,j})\|_2^2 = \|\mathcal{T}^p_{\mathcal{A}}(u^{\gamma}_{\mathcal{A}},\ z^{\mathcal{A}}_{p,i}) - \mathcal{T}^p_{\mathcal{B}}(u^{\gamma}_{\mathcal{B}},\ z^{\mathcal{B}}_{p,j})\|_2^2\ \forall \lambda \in \mathbb{R} \tag{3.8}$$

für $p \in \mathcal{I}^{\mathcal{A},\mathcal{B}}_{PE}$, $i = 1, \ldots, n^p_g[\mathcal{A}]$, und $j = 1, \ldots, n^p_g[\mathcal{B}]$. Aus der Definition der Energiefunktion folgt nun: $\mathsf{E}(u_{\mathcal{A}}, u_{\mathcal{B}}) = \mathsf{E}(u^{\gamma}_{\mathcal{A}}, u^{\gamma}_{\mathcal{B}})$. Wählt man nun $\gamma < \varepsilon\ \sqrt{2}$, so gilt $(u^{*,\gamma}_{\mathcal{A}}, u^{*,\gamma}_{\mathcal{B}}) \in B_{\varepsilon}((u^*_{\mathcal{A}}, u^*_{\mathcal{B}}))$ und $\mathsf{E}(u^*_{\mathcal{A}}, u^*_{\mathcal{B}}) = \mathsf{E}(u^{*,\gamma}_{\mathcal{A}}, u^{*,\gamma}_{\mathcal{B}})$. Damit ist die Behauptung bewiesen. ∎

Bemerkung: Aus dem Beweis von Lemma 3.3.2 d) folgt, daß die Energiefunktion $\mathsf{E}_{\mathcal{A},\mathcal{B}}$ mindestens überabzählbar viele globale Minima besitzt, falls solche überhaupt existieren. Es sei $(u_{\mathcal{A}}, u_{\mathcal{B}}) \in U(n_a[\mathcal{A}])\text{x}U(n_a[\mathcal{B}])$ ein globales Minimum der Energiefunktion. Dann sind für alle $\gamma \in \mathbb{R}$ nach Gleichung (3.8) auch $(u^{\gamma}_{\mathcal{A}}, u^{\gamma}_{\mathcal{B}}) \in U(n_a[\mathcal{A}])\text{x}U(n_a[\mathcal{B}])$ globale Minima der Energiefunktion. Das bedeutet: wendet man nun $\mathcal{T}_{\mathcal{A}}(u_{\mathcal{A}},\ .)$ auf die Atomkoordinaten des Moleküls $\mathcal{A}$ bzw. $\mathcal{T}_{\mathcal{B}}(u_{\mathcal{B}},\ .)$ auf die Atomkoordinaten des Moleküls $\mathcal{B}$ an, so erhält man die zu $(u_{\mathcal{A}}, u_{\mathcal{B}})$ gehörenden Konformationen der Moleküle und deren relative Überlagerung zueinander. Ersetzt man für $\mathcal{M} = \mathcal{A}, \mathcal{B}$ $u_{\mathcal{M}}$ durch $u^{\gamma}_{\mathcal{M}}$, so erhält man dieselbe relative Überlagerung der Moleküle zueinander nur im Raum verschoben.
Wendet man auf die obigen optimalen Konformationen Rotationsmatrizen an, so erhält man wiederum dieselbe relative Überlagerung der Konformationen zueinander, nur im Raum rotiert. Dazu sei $S \in \mathsf{SO}_3$ eine beliebige Rotationsmatrix. Ersetzt man nun $\mathcal{T}_{\mathcal{M}}(u_{\mathcal{M}},\ \cdot)$ durch $S\mathcal{T}_{\mathcal{M}}(u_{\mathcal{M}},\ \cdot)$, $\mathcal{M}$=$\mathcal{A}$,$\mathcal{B}$, so erhält man die rotierten Konformationen. Die relative Überlagerung der Konformationen zueinander wird dadurch aber offensichtlich nicht geändert.

Die hier beobachteten Merkmale lassen sich auch wie folgt formulieren: zu je zwei gegebenen oder berechneten Konformationen lassen sich mindestens überabzählbar viele Möglichkeiten finden, die Konformationen im Raum zu positionieren, ohne deren relative Überlagerung zueinander zu verändern, indem man auf die betreffenden Koordinaten dieselben globalen Transformationen, wie sie z.B. in Kapitel 2.4.1 beschrieben werden, anwendet. Dies liegt daran, daß hier beide Moleküle frei beweglich sind, anders als im Modell 1, wo das Referenzmolekül fest im Raum positioniert ist. Man sollte diese Eigenschaft der Energiefunktion auf jeden Fall beachten, wenn man die errechneten Konformationen mit den gegebenden Kristallstrukturen der Moleküle vergleichen will. Unter Umständen muß man die Konformationen noch geeignet transformieren, bevor ein aussagekräftiger Vergleich möglich ist, z.B. der *RMS*-Abstand zweier Konformationen.[10]

Nun betrachtet man die Verallgemeinerung der Optimierungsaufgabe (3.7). Es sollen nun n_m, $n_m > 1$, flexible Moleküle gegenseitig überlagert werden. Diese werden mit $\mathcal{A}_1, \ldots, \mathcal{A}_{n_m}$ bezeichnet.

Definition 3.3.3
Die Zielfunktion, mit der man das optimale Alignment n_m flexibler Moleküle $\mathcal{A}_i$ zueinander

[10] Unter dem RMS-Abstand zweier Koordinaten-Mengen versteht man folgendes: es seien zwei endliche Mengen $M_1 = \{a_1, \ldots, a_n\} \subset \mathbb{R}^3$ und $M_2 = \{b_1, \ldots, b_n\} \subset \mathbb{R}^3$ gegeben. Dann ist der RMS-Abstand der beiden Mengen definiert als $RMS(M_1, M_2) := \sqrt{(1/n)\sum_{i=1}^{n} \|a_i - b_i\|_2^2}$.

bestimmen kann, ist durch

$$
\begin{aligned}
&\mathsf{E}_{\mathcal{A}_1,\dots,\mathcal{A}_{n_m}}(u_{\mathcal{A}_1},\dots,u_{\mathcal{A}_{n_m}}) := \\
&\qquad \frac{2\,\lambda_A}{n_m(n_m-1)} \sum_{i=1}^{n_m-1} \sum_{j=i+1}^{n_m} \mathsf{A}^{\mathsf{H}}_{\mathcal{A}_i,\mathcal{A}_j} \Big(\big(Q^p_{\mathcal{A}_i}[u_{\mathcal{A}_i}](x)\big)_{p\in\mathcal{I}^{\mathcal{A}_i,\mathcal{A}_j}_{PE}} ,\ \Big(Q^p_{\mathcal{A}_j}[u_{\mathcal{A}_j}](x)\Big)_{p\in\mathcal{I}^{\mathcal{A}_i,\mathcal{A}_j}_{PE}} \Big) \\
&\qquad + \frac{1}{n_m} \sum_{k=1}^{n_m} \Big\{ P^F_{n_a[\mathcal{A}_k]} \Big(\lambda_P, (\mathcal{R}_{\mathcal{A}_k}(\varphi_{\mathcal{A}_k}, a_1^{\mathcal{A}_k}),\dots,\mathcal{R}_{\mathcal{A}_k}(\varphi_{\mathcal{A}_k}, a^{\mathcal{A}_k}_{n_a[\mathcal{A}_k]}), \Sigma_{\mathcal{A}_k} \Big) \\
&\qquad + \lambda_Q\,(\|q_{\mathcal{A}_k}\|_2 - 1)^2 \ +\ \lambda_T s^{\mathcal{A}_k}_{tor}\left(\Theta^{tor}(\mathsf{C}_{\mathcal{A}_k}),\varphi_{\mathcal{A}_k}\right) \Big\}
\end{aligned}
$$

$$
\begin{aligned}
=&\frac{2\,\lambda_A}{n_m(n_m-1)} \sum_{i=1}^{n_m-1} \sum_{j=i+1}^{n_m} \left\{ \sum_{p\in\mathcal{I}^{\mathcal{A}_i,\mathcal{A}_j}_{PE}} \frac{2\,\lambda_p\ \mathcal{Z}^p_{\mathcal{A}_i,\mathcal{A}_j}(u_{\mathcal{A}_i},u_{\mathcal{A}_j})}{\mathcal{N}^p_{\mathcal{A}_i,\mathcal{A}_j}(\varphi_{\mathcal{A}_i},\varphi_{\mathcal{A}_j})} \right\} + \frac{1}{n_m}\sum_{k=1}^{n_m} \Big\{ \lambda_Q\,(\|q_{\mathcal{A}_k}\|_2-1)^2 \\
&+ P^F_{n_a[\mathcal{A}_k]}\Big(\lambda_P, \mathcal{R}_{\mathcal{A}_k}(\varphi_{\mathcal{A}_k}, a_1^{\mathcal{A}_k}),\dots,\mathcal{R}_{\mathcal{A}_k}(\varphi_{\mathcal{A}_k}, a^{\mathcal{A}_k}_{n_a[\mathcal{A}_k]}), \Sigma_{\mathcal{A}_k}\Big) \ +\ \lambda_T s^{\mathcal{A}_k}_{tor}\left(\Theta^{tor}(\mathsf{C}_{\mathcal{A}_k}),\varphi_{\mathcal{A}_k}\right)\Big\}
\end{aligned}
$$

gegeben, wobei $u_{\mathcal{A}_i} = (t_{\mathcal{A}_i}, q_{\mathcal{A}_i}, c_{\mathcal{A}_i}, \varphi_{\mathcal{A}_i}) \in U(n_a[\mathcal{A}_i])$ *für* $i = 1,\dots,n_m$ *gilt. Für* $i = 1,\dots,n_m$ *ist* $U(n_a[\mathcal{A}_i])$ *der Parameterraum der dem Molekül* $\mathcal{A}_i$ *zugeordneten Transformationsoperatoren.* $\mathcal{Z}^p_{\mathcal{A}_i,\mathcal{A}_j}$ *bzw.* $\mathcal{N}^p_{\mathcal{A}_i,\mathcal{A}_j}$ *werden für* $p \in \mathcal{I}^{\mathcal{A}_i,\mathcal{A}_j}_{PE}$, $i,j = 1,\dots,n_m$, $i \neq j$ *analog zu* $\mathcal{Z}^p_{\mathcal{A},\mathcal{B}}$ *bzw.* $\mathcal{N}^p_{\mathcal{A},\mathcal{B}}$ *aus Lemma 3.3.2 b) definiert. Setzt man nun* $U(n_m) := U(n_a[\mathcal{A}_1]) \times \dots \times U(n_a[\mathcal{A}_{n_m}])$ *und* $u_{n_m} := (u_{\mathcal{A}_1},\dots,u_{\mathcal{A}_{n_m}})$, *so lautet die Optimierungsaufgabe*

$$\min_{u_{n_m}\ \in\ U(n_m)} \ \mathsf{E}_{\mathcal{A}_1,\dots,\mathcal{A}_{n_m}}(u_{n_m})$$

Das Ziel der Optimierung ist es, $(u_{\mathcal{A}_1},\dots,u_{\mathcal{A}_{n_m}}) \in U(n_m)$ *so zu bestimmen, daß die Summe der Alignmentindizes*

$$
\begin{aligned}
\mathsf{S}_{\mathcal{A}_1,\dots,\mathcal{A}_{n_m}}(u_{\mathcal{A}_1},\dots,u_{\mathcal{A}_{n_m}}) \ :=\ & \frac{2}{n_m(n_m-1)} \sum_{i=1}^{n_m-1} \sum_{j=i+1}^{n_m} \\
& \mathsf{A}^{\mathsf{H}}_{\mathcal{A}_i,\mathcal{A}_j} \Big(\big(Q^p_{\mathcal{A}_i}[u_{\mathcal{A}_i}](x)\big)_{p\in\mathcal{I}^{\mathcal{A}_i,\mathcal{A}_j}_{PE}} ,\ \Big(Q^p_{\mathcal{A}_j}[u_{\mathcal{A}_j}](x)\Big)_{p\in\mathcal{I}^{\mathcal{A}_i,\mathcal{A}_j}_{PE}} \Big) \qquad (3.9)
\end{aligned}
$$

maximal ist, wobei die durch $(u_{\mathcal{A}_1},\dots,u_{\mathcal{A}_{n_m}})$ *transformierten Molekülkonformationen zulässig sein müssen. Dabei bewertet* $\mathsf{S}_{\mathcal{A}1,\dots,\mathcal{A}_{n_m}}$ *das gegenseitige Alignment der entsprechenden transformierten Moleküle* $\mathcal{A}_1,\dots\mathcal{A}_{n_m}$. ◇

Lemma 3.3.4

Die Energiefunktion $\mathsf{E}_{\mathcal{A}_1,\dots,\mathcal{A}_{n_m}}$ *aus Definition 3.3.3 hat die folgenden Eigenschaften:*

a) Die Energiefunktion $\mathsf{E}_{\mathcal{A}_1,\dots,\mathcal{A}_{n_m}}$ *ist nach unten beschränkt und es gilt*

$$\mathsf{E}_{\mathcal{A}_1,\dots,\mathcal{A}_{n_m}}(u_{n_m}) \ \geq\ \lambda_A \ - \lambda_P \ - \lambda_T$$

für alle $u_{n_m} \in U(n_m)$.

b) Es sei $U_g = \mathbb{R}^3 \times \mathbb{R}^4 \setminus \{0\} \times \mathbb{R}^3$. *Für* $k = 1,\dots,n_m$ *wird* $U_{\mathcal{A}_k}$ *analog zu Lemma 3.3.2 b) gebildet und es gilt:* $\tilde{U}(n_a[\mathcal{A}_k]) := U_g \times U_{\mathcal{A}_k} \subset U(n_a[\mathcal{A}_k])$. *Für alle* $u_{n_m} \in \tilde{U}(n_m)$ *ist die Energiefunktion wohldefiniert, wobei* $\tilde{U}(n_m) := \tilde{U}(n_a[\mathcal{A}_1]) \times \dots \times \tilde{U}(n_a[\mathcal{A}_{n_m}])$ *gesetzt* *worden ist. Die Energiefunktion ist in $u_{n_m} \in \tilde{U}(n_m)$ stetig partiell differenzierbar und die Komponenten des Gradienten $\nabla \mathsf{E}_{\mathcal{A}_1,\dots,\mathcal{A}_{n_m}}$ sind für $i = 1,\dots,n_m$ gegeben durch:*

$$\begin{aligned}\frac{\partial}{\partial(u_{\mathcal{A}_i})_k}\mathsf{E}_{\mathcal{A}_1,\dots,\mathcal{A}_{n_m}}(u_{n_m}) &= \delta_{k,3+l}\frac{2\,\lambda_Q\,(q_{\mathcal{A}_i})_l\,(\|q_{\mathcal{A}_i}\|_2-1)}{\|q_{\mathcal{A}_i}\|_2}\\ &+ \frac{2\,\lambda_A}{n_m(n_m-1)}\sum_{\substack{j=1\\ i\neq j}}^{n_m}\sum_{p\in\mathcal{I}_{PE}^{\mathcal{A}_i,\mathcal{A}_j}}\left\{\frac{2\,\lambda_p}{\mathcal{N}^p_{\mathcal{A}_i,\mathcal{A}_j}(\varphi_{\mathcal{A}_i},\varphi_{\mathcal{A}_j})}\,\frac{\partial}{\partial(u_{\mathcal{A}_i})_k}\,\mathcal{Z}^p_{\mathcal{A}_i,\mathcal{A}_j}(u_{\mathcal{A}_i},\mathcal{A}_j)\right\}\end{aligned}$$

für $k = 1,\dots,10$, $l = 1,2,3,4$, und

$$\begin{aligned}&\frac{\partial}{\partial(u_{\mathcal{A}_i})_{10+k}}\mathsf{E}_{\mathcal{A}_1,\dots,\mathcal{A}_{n_m}}(u_{n_m})\\ &\quad= \frac{2\,\lambda_A}{n_m(n_m-1)}\sum_{\substack{j=1\\ i\neq j}}^{n_m}\sum_{p\in\mathcal{I}_{PE}^{\mathcal{A}_j,\mathcal{A}_j}}\left\{\frac{2\,\lambda_p}{\mathcal{N}^p_{\mathcal{A}_i,\mathcal{A}_j}(\varphi_{\mathcal{A}_i},\varphi_{\mathcal{A}_j})}\,\frac{\partial}{\partial(\varphi_{\mathcal{A}_i})_k}\,\mathcal{Z}^p_{\mathcal{A}_i,\mathcal{A}_j}(u_{\mathcal{A}_i},u_{\mathcal{A}_j})\right.\\ &\qquad\left. - \frac{\lambda_p\,\mathcal{Z}^p_{\mathcal{A}_i,\mathcal{A}_j}(u_{\mathcal{A}_i},u_{\mathcal{A}_j})}{\left(\mathcal{N}^p_{\mathcal{A}_i,\mathcal{A}_j}(\varphi_{\mathcal{A}_i},\varphi_{\mathcal{A}_j})\right)^2}\,\frac{\partial}{\partial(\varphi_{\mathcal{A}_i})_k}\,\mathcal{N}^p_{\mathcal{A}_i,\mathcal{A}_j}(\varphi_{\mathcal{A}_i},\varphi_{\mathcal{A}_j})\right\}\\ &\qquad + \frac{1}{n_m}\left\{\frac{\partial}{\partial(\varphi_{\mathcal{A}_i})_k}\,P^F_{n_a[\mathcal{A}_i]}\left(\lambda_P,\mathcal{R}_{\mathcal{A}_i}(\varphi_{\mathcal{A}_i},a_1^{\mathcal{A}_i}),\dots,\mathcal{R}_{\mathcal{A}_i}(\varphi_{\mathcal{A}_i},a^{\mathcal{A}_i}_{n_a[\mathcal{A}_i]}),\Sigma_{\mathcal{A}_i}\right)\right.\\ &\qquad\left. + \lambda_T\,\frac{\partial}{\partial(\varphi_{\mathcal{A}_i})_k}\,s^{\mathcal{A}_i}_{tor}\left(\Theta^{tor}(\mathsf{C}_{\mathcal{A}_i}),\varphi_{\mathcal{A}_i}\right)\right\}\end{aligned}$$

für $k = 1,\dots,n_{ft}[\mathcal{A}_i]$. Die genaue Darstellung der einzelnen Komponenten kann man aus Lemma 3.3.2 b) entnehmen.

c) *Für jedes lokale Minimum $u^*_{n_m} \in U(n_m)$ der Energiefunktion $\mathsf{E}_{\mathcal{A}_1,\dots,\mathcal{A}_{n_m}}$ gilt $\|q^*_{\mathcal{A}_k}\|_2 = 1$, $k = 1,\dots,n_m$.*

d) *Die Energiefunktion $\mathsf{E}_{\mathcal{A}_1,\dots,\mathcal{A}_{n_m}}$ besitzt keine isolierten Minima.*

Beweis:

Der Beweis wird analog zum Beweis von Lemma 3.3.2 geführt. ■

Bemerkung: Die Aussagen, die in der Bemerkung zu Lemma 3.3.2 für die Energiefunktion $\mathsf{E}_{\mathcal{A},\mathcal{B}}$ gemacht worden sind, treffen auch auf die Energiefunktion $\mathsf{E}_{\mathcal{A}_1,\dots,\mathcal{A}_{n_m}}$ zu.

Zum Abschluß des Kapitels wird eine Energiefunktion definiert, mit deren Hilfe man einen Gauß-Seidel-artigen Optimierungsalgorithmus für die Energiefunktionen aus Definition 3.3.3 entwickeln kann. Dazu seien $\mathcal{A}$ und $\mathcal{B}_1,\dots,\mathcal{B}_{n_m-1}$ flexible Moleküle. Weiter sei $(u_{\mathcal{B}_1},\dots,u_{\mathcal{B}_{n_m-1}}) \in U(n_a[\mathcal{B}_1])\mathsf{x}\dots\mathsf{x}U(n_a[\mathcal{B}_{n_m-1}])$ so gewählt, daß die transformierten Molekülkonformationen von $\mathcal{B}_1,\dots,\mathcal{B}_{n_m-1}$ zulässig sind, d.h. wendet man für $i = 1,\dots,n_m-1$ den Transformationsoperator $\mathcal{T}_{\mathcal{B}_i}(u_{\mathcal{B}_i},\cdot)$ auf die Atomkoordinaten von $\mathcal{B}_i$ an, so erhält man eine zulässige Konformation. Nun soll das Molekül $\mathcal{A}$ bezüglich der n_m-1 Moleküle $\mathcal{B}_1,\dots,\mathcal{B}_{n_m-1}$ positioniert werden, d.h. die optimale Überlagerung bzw. Alignment des Moleküls $\mathcal{A}$ bezüglich der Moleküle $\mathcal{B}_1,\dots,\mathcal{B}_{n_m-1}$ wird gesucht. Die Gausszentren von $\mathcal{B}_i$, $i = 1,\dots,n_m-1$, sind durch die Transformationsoperatoren $\mathcal{T}^p_{\mathcal{B}_i}(u_{\mathcal{B}_i},\cdot)$, $p \in \mathcal{I}_{PE}$, fest gegeben.

Definition 3.3.5
Die Zielfunktion, mit der man das optimale Alignment eines flexiblen Moleküls $\mathcal{A}$ bezüglich $n_m - 1$ Moleküle bestimmen kann, ist durch

$$\begin{aligned}
&\mathsf{E}_{\mathcal{A}}[u_{\mathcal{B}_1}, \ldots, u_{\mathcal{B}_{n_m-1}}](u_{\mathcal{A}}) := \\
&\quad \frac{2\,\lambda_A}{n_m(n_m-1)} \sum_{k=1}^{n_m-1} \mathsf{A}^{\mathsf{H}}_{\mathcal{A},\mathcal{B}_k} \left(\left(Q^p_{\mathcal{A}}[u_{\mathcal{A}}](x)\right)_{p \in \mathcal{I}^{\mathcal{A},\mathcal{B}_k}_{PE}},\ \left(Q^p_{\mathcal{B}_k}[u_{\mathcal{B}_k}](x)\right)_{p \in \mathcal{I}^{\mathcal{A},\mathcal{B}_k}_{PE}} \right) \\
&\quad + \frac{1}{n_m} \left\{ P^F_{n_a[\mathcal{A}_k]} \left(\lambda_P, (\mathcal{R}_{\mathcal{A}}(\varphi_{\mathcal{A}}, a^{\mathcal{A}}_1), \ldots, \mathcal{R}_{\mathcal{A}}(\varphi_{\mathcal{A}}, a^{\mathcal{A}}_{n_a[\mathcal{A}]}), \Sigma_{\mathcal{A}} \right) \right. \\
&\quad\quad \left. + \lambda_Q \left(\|q_{\mathcal{A}}\|_2 - 1\right)^2 + \lambda_T s^{\mathcal{A}}_{tor}\left(\Theta^{tor}(\mathsf{C}_{\mathcal{A}}), \varphi_{\mathcal{A}}\right) \right\}
\end{aligned}$$

$$\begin{aligned}
= &\frac{2\,\lambda_A}{n_m(n_m-1)} \sum_{k=1}^{n_m-1} \left\{ \sum_{p \in \mathcal{I}^{\mathcal{A},\mathcal{B}_k}_{PE}} \frac{2\,\lambda_p\, \mathcal{Z}^p_{\mathcal{A},\mathcal{B}_k}(u_{\mathcal{A}}, u_{\mathcal{B}_k})}{\mathcal{N}^p_{\mathcal{A},\mathcal{B}_k}(\varphi_{\mathcal{A}}, \varphi_{\mathcal{B}_k})} \right\} + \frac{1}{n_m} \left\{ \lambda_Q \left(\|q_{\mathcal{A}}\|_2 - 1\right)^2 \right. \\
&\left. + P^F_{n_a[\mathcal{A}]} \left(\lambda_P, \mathcal{R}_{\mathcal{A}}(\varphi_{\mathcal{A}}, a^{\mathcal{A}}_1), \ldots, \mathcal{R}_{\mathcal{A}}(\varphi_{\mathcal{A}}, a^{\mathcal{A}}_{n_a[\mathcal{A}]}), \Sigma_{\mathcal{A}} \right) + \lambda_T s^{\mathcal{A}}_{tor}\left(\Theta^{tor}(\mathsf{C}_{\mathcal{A}}), \varphi_{\mathcal{A}}\right) \right\}
\end{aligned}$$

gegeben, wobei $u_{\mathcal{A}} = (t_{\mathcal{A}}, q_{\mathcal{A}}, c_{\mathcal{A}}, \varphi_{\mathcal{A}}) \in U(n_a[\mathcal{A}])$ gilt. Dabei werden $\mathcal{Z}^p_{\mathcal{A},\mathcal{B}_k}$ bzw. $\mathcal{N}^p_{\mathcal{A},\mathcal{B}_k}$ für $p \in \mathcal{I}^{\mathcal{A},\mathcal{B}_i}_{PE}$ und $k = 1, \ldots, n_m - 1$, analog zu $\mathcal{Z}^p_{\mathcal{A},\mathcal{B}}$ bzw. $\mathcal{N}^p_{\mathcal{A},\mathcal{B}}$ in Lemma 3.3.2 b) definiert.
Die Optimierungsaufgabe lautet

$$\min_{u_{\mathcal{A}} \in U(n_a[\mathcal{A}])} \mathsf{E}_{\mathcal{A}}[u_{\mathcal{B}_1}, \ldots, u_{\mathcal{B}_{n_m-1}}](u_{\mathcal{A}}).$$

Das Ziel der Optimierung ist es, $u_{\mathcal{A}} \in U(n_a[\mathcal{A}])$ so zu bestimmen, daß die Summe der Alignmentindizes

$$\begin{aligned}
&\mathsf{S}_{\mathcal{A}}[u_{\mathcal{B}_1}, \ldots, u_{\mathcal{B}_{n_m-1}}](u_{\mathcal{A}}) := \\
&\quad \frac{1}{n_m - 1} \sum_{i=1}^{n_m-1} \mathsf{A}^{\mathsf{H}}_{\mathcal{A},\mathcal{B}_i} \left(\left(Q^p_{\mathcal{A}}[u_{\mathcal{A}}](x)\right)_{p \in \mathcal{I}^{\mathcal{A},\mathcal{B}_i}_{PE}},\ \left(Q^p_{\mathcal{B}_i}[u_{\mathcal{B}_i}](x)\right)_{p \in \mathcal{I}^{\mathcal{A},\mathcal{B}_i}_{PE}} \right) \qquad (3.10)
\end{aligned}$$

maximal ist, wobei die durch $u_{\mathcal{A}}$ transformierte Molekülkonformation zulässig sein muß. Dabei bewertet $\mathsf{S}_{\mathcal{A}}$ das Alignment des transformierten Moleküls $\mathcal{A}$ bzgl. der transformierten Moleküle $\mathcal{B}_1, \ldots, \mathcal{B}_{n_m}$. ◇

Bemerkung: Die Faktoren $2/(n_m(n_m-1))$ und $1/n_m$ in der obigen Energiefunktion werden verwendett, damit die Energiefunktion $\mathsf{E}_{\mathcal{A}}$ mit der Energiefunktion aus Definition 3.3.3 kompatibel ist (siehe Lemma 3.3.6 d)). Wendet man eines der Standardoptimierungsverfahren aus Kapitel 4.2 an, so werden in jedem Schritt die Konformationen aller Moleküle simultan transformiert. Die Idee ist nun, die Moleküle iterativ zu transformieren. D.h. die Konformation eines Moleküls wird optimiert, während die Konformationen der anderen Moleküle rigide gehalten werden. Um alle Moleküle einmal zu transformieren, werden n_m Energiefunktionen vom Typ $\mathsf{E}_{\mathcal{A}}$ nacheinander minimiert. Dieses Verfahren wird dann iteriert. Einzelheiten werden in Kapitel 4.3 beschrieben.
Die Moleküle $\mathcal{B}_1, \ldots, \mathcal{B}_{n_m-1}$ können auch rigide sein. Man wählt dann $u_{\mathcal{B}_k}$, $k = 1, \ldots, n_m - 1$, so, daß der Transformationsoperator die Identität ist.[11]

[11] Siehe Lemma 2.4.28 d).

Lemma 3.3.6

Es sei $u_{\mathcal{B}} := (u_{\mathcal{B}_1}, \dots, u_{\mathcal{B}_{n_m-1}}) \in U(n_a[\mathcal{B}_1]) \mathsf{x} \dots \mathsf{x} U(n_a[\mathcal{B}_{n_m-1}])$ so gewählt, daß die transformierten Molekülkonformationen von $\mathcal{B}_1, \dots, \mathcal{B}_{n_m-1}$ zulässig sind. Dann hat die Energiefunktion $\mathsf{E}_{\mathcal{A}}$ aus Definition 3.3.5 die folgenden Eigenschaften:

a) Die Energiefunktion $\mathsf{E}_{\mathcal{A}}$ ist nach unten beschränkt und es gilt

$$\mathsf{E}_{\mathcal{A}}[u_{\mathcal{B}_1}, \dots, u_{\mathcal{B}_{n_m-1}}](u_{\mathcal{A}}) \;\geq\; \lambda_A \;-\lambda_P \;-\lambda_T$$

für alle $u_{\mathcal{A}} \in U(n_a[\mathcal{A}])$.

b) Es sei $U_g = \mathbb{R}^3 \mathsf{x} \mathbb{R}^4 \setminus \{0\} \mathsf{x} \mathbb{R}^3$. Weiter sei

$$U_{\varphi}[u_{\mathcal{B}_1}, \dots, u_{\mathcal{B}_{n_m-1}}] := \Big\{ \varphi_{\mathcal{A}} \in \mathbb{R}^{n_{ft}[\mathcal{A}]} \mid \Big(\mathcal{N}^p_{\mathcal{A},\mathcal{B}_k}(\varphi_{\mathcal{A}}, \varphi_{\mathcal{B}_k}) \neq 0 \;\forall p \in \mathcal{I}^{\mathcal{A},\mathcal{B}_k}_{PE},\; k = 1, \dots, n_m \Big) \\ \wedge \;\; (\mathsf{N}_{ij}(\varphi_{\mathcal{A}}, a_1, a_2) \neq 0 \;\forall a_1 \in V_i, a_2 \in V_j, i, j = 0, \dots, n_{ft}[\mathcal{A}], i \neq j) \}$$

Setzt man $\tilde{U}(n_a[\mathcal{A}])[u_{\mathcal{B}_1}, \dots, u_{\mathcal{B}_{n_m-1}}] := U_g \mathsf{x} U_{\varphi}[u_{\mathcal{B}_1}, \dots, u_{\mathcal{B}_{n_m-1}}] \subset U(n_a[\mathcal{A}])$, so ist die Energiefunktion für alle $u_{\mathcal{A}} \in \tilde{U}(n_a[\mathcal{A}])[u_{\mathcal{B}_1}, \dots, u_{\mathcal{B}_{n_m-1}}]$ wohldefiniert. Die Engiefunktion ist in $u_{\mathcal{A}} \in \tilde{U}(n_a[\mathcal{A}])[u_{\mathcal{B}_1}, \dots, u_{\mathcal{B}_{n_m-1}}]$ stetig partiell differenzierbar und die Komponenten des Gradienten $\nabla \mathsf{E}_{\mathcal{A}}$ sind gegeben durch

$$\frac{\partial}{\partial (u_{\mathcal{A}})_k} \mathsf{E}_{\mathcal{A}}[u_{\mathcal{B}_1}, \dots, u_{\mathcal{B}_{n_m-1}}](u_{\mathcal{A}}) \;=\; \delta_{k,3+l} \frac{2\,\lambda_Q\;(q_{\mathcal{A}})_l\;(\|q_{\mathcal{A}}\|_2 - 1)}{\|q_{\mathcal{A}}\|_2} \\ +\; \frac{2\,\lambda_A}{n_m(n_m-1)} \sum_{i=1}^{n_m-1} \sum_{p \in \mathcal{I}^{\mathcal{A},\mathcal{B}_i}_{PE}} \left\{ \frac{2\,\lambda_p}{\mathcal{N}^p_{\mathcal{A},\mathcal{B}_i}(\varphi_{\mathcal{A}}, \varphi_{\mathcal{B}_i})} \frac{\partial}{\partial (u_{\mathcal{A}_i})_k} \;\mathcal{Z}^p_{\mathcal{A},\mathcal{B}_i}(u_{\mathcal{A}}, u_{\mathcal{B}_i}) \right\}$$

für $k = 1, \dots, 10,\; l = 1, 2, 3, 4$, und

$$\frac{\partial}{\partial (u_{\mathcal{A}})_{10+k}} \mathsf{E}_{\mathcal{A}}[u_{\mathcal{B}_1}, \dots, u_{\mathcal{B}_{n_m-1}}](u_{\mathcal{A}}) = \\ \frac{2\,\lambda_A}{n_m(n_m-1)} \sum_{i=1}^{n_m-1} \sum_{p \in \mathcal{I}^{\mathcal{A},\mathcal{B}_i}_{PE}} \left\{ \frac{2\,\lambda_p}{\mathcal{N}^p_{\mathcal{A},\mathcal{B}_i}(\varphi_{\mathcal{A}}, \varphi_{\mathcal{B}_i})} \frac{\partial}{\partial (\varphi_{\mathcal{A}})_k} \;\mathcal{Z}^p_{\mathcal{A},\mathcal{B}_i}(u_{\mathcal{A}}, u_{\mathcal{B}_i}) \right. \\ \left. -\; \frac{\lambda_p\;\mathcal{Z}^p_{\mathcal{A},\mathcal{B}_i}(u_{\mathcal{A}}, u_{\mathcal{B}_i})}{\left(\mathcal{N}^p_{\mathcal{A},\mathcal{B}_i}(\varphi_{\mathcal{A}}, \varphi_{\mathcal{B}_i})\right)^2} \frac{\partial}{\partial (\varphi_{\mathcal{A}})_k} \;\mathcal{N}^p_{\mathcal{A},\mathcal{B}_i}(\varphi_{\mathcal{A}}, \varphi_{\mathcal{B}_i}) \right\} \\ +\; \frac{1}{n_m} \left\{ \frac{\partial}{\partial (\varphi_{\mathcal{A}})_k} \;P^F_{n_a[\mathcal{A}]} \left(\lambda_P, \mathcal{R}_{\mathcal{A}}(\varphi_{\mathcal{A}}, a^{\mathcal{A}}_1), \dots, \mathcal{R}_{\mathcal{A}}(\varphi_{\mathcal{A}}, a^{\mathcal{A}}_{n_a[\mathcal{A}]}), \Sigma_{\mathcal{A}}\right) \right. \\ \left. +\; \lambda_T \;\frac{\partial}{\partial (\varphi_{\mathcal{A}})_k} \;s^{\mathcal{A}}_{tor}\left(\Theta^{tor}(\mathsf{C}_{\mathcal{A}}), \varphi_{\mathcal{A}}\right) \right\}$$

für $k = 1, \dots, n_{ft}[\mathcal{A}]$. Die genaue Darstellung der einzelnen Komponenten kann aus Lemma 3.3.2 b) entnehmen kann.

c) Für jedes lokales Minimum $u^*_{\mathcal{A}} \in U(n_a[\mathcal{A}])$ der Energiefunktion $\mathsf{E}_{\mathcal{A}}$ gilt $\|q^*_{\mathcal{A}}\|_2 = 1$.

d) Es sei $\mathsf{E}_{\mathcal{A},\mathcal{B}_1,\dots,\mathcal{B}_{n_m-1}}$ die Energiefunktion aus Definiton 3.3.3, mit der man die gegenseitige Überlagerung der Moleküle $\mathcal{A}$, $\mathcal{B}_1, \dots, \mathcal{B}_{n_m-1}$ optimieren kann. Dann gilt für alle $u_{\mathcal{A}} \in$

$U(n_a[\mathcal{A}])$:

$$\mathsf{E}_{\mathcal{A},\mathcal{B}_1,\dots,\mathcal{B}_{n_m-1}}(u_{\mathcal{A}}, u_{\mathcal{B}_1}, \dots, u_{\mathcal{B}_{n_m-1}}) \; = \mathsf{E}_{\mathcal{A}}[u_{\mathcal{B}_1}, \dots, u_{\mathcal{B}_{n_m-1}}](u_{\mathcal{A}}) \; + \; C_{\mathcal{A}}(u_{\mathcal{B}_1}, \dots, u_{\mathcal{B}_{n_m-1}}),$$

wobei

$$\begin{aligned} C_{\mathcal{A}}(u_{\mathcal{B}_1}, \dots, u_{\mathcal{B}_{n_m-1}}) \; &:= \; \mathsf{E}_{\mathcal{A},\mathcal{B}_1,\dots,\mathcal{B}_{n_m-1}}(u_{\mathcal{A}}, u_{\mathcal{B}_1}, \dots, u_{\mathcal{B}_{n_m-1}}) \; - \; \mathsf{E}_{\mathcal{A}}[u_{\mathcal{B}_1}, \dots, u_{\mathcal{B}_{n_m-1}}](u_{\mathcal{A}}) \\ &= \frac{2\,\lambda_A}{n_m(n_m-1)} \left\{ \sum_{i=1}^{n_m-2} \sum_{j=i+1}^{n_m-1} \mathsf{A}^{\mathsf{H}}_{\mathcal{B}_i,\mathcal{B}_j} \left(\; \left(Q^p_{\mathcal{B}_i}[u_{\mathcal{B}_i}](x)\right)_{p\in\mathcal{I}^{\mathcal{B}_i,\mathcal{B}_j}_{PE}} , \; \left(Q^p_{\mathcal{B}_j}[u_{\mathcal{B}_j}](x)\right)_{p\in\mathcal{I}^{\mathcal{B}_i,\mathcal{B}_j}_{PE}} \; \right) \right\} \\ &= \frac{1}{n_m} \sum_{i=1}^{n_m-1} \Big\{ P^F_{n_a[\mathcal{B}_i]} \left(\lambda_P, \mathcal{R}_{\mathcal{B}_i}(\varphi_{\mathcal{B}_i}, a_1^{\mathcal{B}_i}), \dots, \mathcal{R}_{\mathcal{B}_i}(\varphi_{\mathcal{B}_i}, a^{\mathcal{B}_i}_{n_a[\mathcal{B}_i]}), \Sigma_{\mathcal{B}_i} \right) \; + \; \lambda_Q \left(\|q_{\mathcal{B}_i}\|_2 - 1\right)^2 \\ &\qquad + \; \lambda_T \; s^{\mathcal{B}_i}_{tor} \left(\Theta^{tor}(\mathsf{C}_{\mathcal{B}_i}), \varphi_{\mathcal{B}_i}\right) \Big\} \end{aligned}$$

bzgl. $u_{\mathcal{A}}$ *konstant ist.*

Beweis:

a)-c) Der Beweis wird analog zum Beweis von Lemma 3.3.2 a)-c) geführt.

d) Folgt direkt aus der Definition von $\mathsf{E}_{\mathcal{A}}$. ■

Bemerkung: Es sei $(u_{\mathcal{B}_1}, \dots, u_{\mathcal{B}_{n_m-1}}) \in U(n_a[\mathcal{B}_1]) \mathsf{x} \dots \mathsf{x} U(n_a[\mathcal{B}_{n_m-1}])$ wie in Lemma 3.3.6. Dann folgt aus der Aussage d):

$$\nabla_{u_{\mathcal{A}}} \mathsf{E}_{\mathcal{A},\mathcal{B}_1,\dots,\mathcal{B}_{n_m-1}}(u_{\mathcal{A}}, u_{\mathcal{B}_1}, \dots, u_{\mathcal{B}_{n_m-1}}) \; = \; \nabla \mathsf{E}_{\mathcal{A}}[u_{\mathcal{B}_1}, \dots, u_{\mathcal{B}_{n_m-1}}](u_{\mathcal{A}})$$

für alle $u_{\mathcal{A}} \in \tilde{U}(n_a[\mathcal{A}])[u_{\mathcal{B}_1}, \dots, u_{\mathcal{B}_{n_m-1}}]$.

Kapitel 4

Die Optimierungsalgorithmen

Das Kapitel unterteilt sich in drei Unterkapitel. In Kapitel 4.1 wird der Gesamtalgorithmus vorgestellt, der implementiert worden ist. In Kapitel 4.2 werden die Standardoptimierungsalgorithmen beschrieben, die für die Minimierung der Enegiefunktionen aus Kapitel 3 verwendet worden sind. Im abschließenden Kapitel 4.3 wird die Entwicklung problembezogener Verfahren, die die Minimierung der entsprechenden Energiefunktionen effizienter ausführen sollen, geschildert.

4.1 Gesamtalgorithmus

Der Gesamtalgorithmus ist ein Unterprogramm von FlexS ([41]) und besteht aus drei Komponenten, dem Einlesemodul, dem Minimierungsmodul und dem Ergebnismodul. In der ersten Komponente, dem Einlesemodul, werden die Daten, die das Programm FlexS bereitstellt, verarbeitet. So wird für jedes flexible Molekül $\mathcal{A}$ die Datenstruktur $\mathcal{DS}_{\mathcal{A}}(\mathsf{C}_{\mathcal{A}})$ bestimmt.[1] In der zweiten Komponente, dem Minimierungsmodul, wird die dem Modell entsprechende Energiefunktion aus Kapitel 3.2 bzw. aus Kapitel 3.3 mit einem Optimierungsalgorithmus minimiert. In der dritten Komponente, dem Ergebnismodul, werden die Ergebnisse der Minimierung berechnet und ausgegeben. Eine vierte Komponente, das Steuerungsmodul, die unabhängig von den drei anderen Modulen ist, erlaubt es, einzelne Terme der entsprechenden Energiefunktionen sowie einzelne physikochemische Eigenschaften zu aktivieren bzw. zu deaktivieren. Weiter kann man das Optimierungsverfahren und etwaige Parameter, sowie die Schrittweitenverfahren und deren Parameter festgelegen. Dann können Kriterien festgelegt werden, von denen mindestens eine erfüllt werden muß, bevor der Optimierungsalgorithmus beendet werden kann. Es besteht noch die Möglichkeit, spezielle Algorithmen zu aktivieren, um die Energiefunktionen effizienter zu minimieren.[2]

[1] Siehe Definition 2.2.3.
[2] Siehe Kapitel 4.3.

Im folgenden werden die einzelnen Komponenten etwas detailierter dargestellt. Dabei werden die Notationen und Bezeichnungen aus Kapitel 2.2 verwendet. Modell 1 bezeichnet das Verfahren, wo die Überlagerung eines flexiblen Moleküls bzgl. eines starren Referenzmoleküls optimiert wird, während im Modell 2 das gegenseitige Alignment einer variablen Anzahl von flexiblen Molekülen optimiert wird.[3]

Das Einlesemodul:

Für Modell 1 muß ein Referenzmolekül $\mathcal{R}$ und mindestens ein flexibles Molekül $\mathcal{A}$ ausgewählt werden. Für das Referenzmolekül sei $K_{\mathcal{R}} \subset \mathbb{R}^3$ die Menge der Atomkoordinaten, die der Gesamtalgorithmus von FlexS übernommen hat. Weiter werden die Parameter (Breite, Höhe und Gausszentrum) derjenigen Gaussfunktionen übernommen, deren Summen dann die räumlich lokalisierten Dichten der physikochemischen Eigenschaften aus $\mathcal{I}_{PE}$ darstellen.

Für Modell 2 müssen mindestens zwei flexible Moleküle, $\mathcal{A}_1$ und $\mathcal{A}_2$, ausgewählt werden.

Für jedes flexible Molekül $\mathcal{A}$, das ausgewählt worden ist, wird unabhängig vom Modell die folgende Routine ausgeführt. Dabei ist $\mathsf{C}_{\mathcal{A}}$ die Molekülkonformation von $\mathcal{A}$, die FlexS zu Verfügung stellt. Die Aufgabe dieser Routine ist es, unteranderem für das Molekül $\mathcal{A}$ die Datenstruktur $\mathcal{DS}_{\mathcal{A}}(\mathsf{C}_{\mathcal{A}})$ bzgl. der Konformation $\mathsf{C}_{\mathcal{A}}$ zu bestimmen.

Algorithmus 4.1 (Einleseroutine für das flexible Molekül $\mathcal{A}$)

(S.0) Es besteht die Option, den Transformationsoperator für das Molekül $\mathcal{A}$ zu modifizieren. D.h. man kann die zu otimierenden Tranformationen (Translation, Globale Rotation, interne Rotationen) gezielt aktivieren bzw. deaktivieren. Falls interne Rotationen zugelassen sind, besteht weiter die Option, einzelne Bindungen, an denen der Torsionswinkel variiert werden kann, künstlich rigide zu halten.

(S.1) Von FlexS wird der Molekülgraph $G_{\mathcal{A}} = (K_{\mathcal{A}},\ B_{\mathcal{A}})$ übernommen, wobei $K_{\mathcal{A}} := K(\mathsf{C}_{\mathcal{A}}) \subset \mathbb{R}^3$ die Menge der Atomkoordinaten bzgl. $\mathsf{C}_{\mathcal{A}}$ ist. $B_{\mathcal{A}} \subset \mathcal{P}_2(K_{\mathcal{A}})$ ist die Menge der Atombindungen. Weiter werden die Van-der-Waals-Radien der Atome und die Parameter derjenigen Gaussfunktionen übernommen, deren Summe die räumlich lokalisierten Dichten der physikochemischen Eigenschaften darstellen. Falls die Kristallstruktur FlexS bekannt ist, werden im Modell 1 zusätzlich die Koordinaten der Atome bzgl. der Kristallstruktur übernommen.
Für jede Bindung $e \in B_{\mathcal{A}}$ wird der Wert von $\Phi^{\mathcal{A}}_{FT}(e)$ bestimmt. Falls $\Phi^{\mathcal{A}}_{FT}(e) = 1$ gilt, d.h. an dieser Bindung kann der Torsionswinkel variiert werden, so wird der Torsionswinkel $\alpha_e(\mathsf{C}_{\mathcal{A}})$ berechnet und der Datensatz der Knotenzerlegung und der Energiewerte $\triangle_e$ übernommen. Falls es notwendig ist, wird das Gleichungssystem $B_e \alpha_e = b_e$ aus Gleichung (2.23) in Kapitel 2.7.1 entsprechend dem Datensatz $\triangle_e$ aufgestellt und mit einem Gaußschen Eliminationsverfahren mit Spaltenpivotsuche (siehe hierzu [54]) gelöst, wobei die Matrix B_e equilibriert wird.
Insgesamt stehen nun der Molekülgraph $G_{\mathcal{A}}$, die Menge $B^{\mathcal{A}}_{FT} = \{e \in B_{\mathcal{A}} \mid \Phi^{\mathcal{A}}_{FT}(e) = 1\}$, der Vektor der Van-der-Waals Radien $\Sigma_{\mathcal{A}}$, der Vektor der Torsionswinkel $\Theta^{tor}(\mathsf{C}_{\mathcal{A}})$ und für

[3]Siehe auch die Definition 2.1.4 und Kapitel 3.2 für Modell 1 bzw. Kapitel 3.3 für Modell 2.

$p \in \mathcal{I}_{PE}$ die Menge der Gausszentren $GC^p_{\mathcal{A}}$ sowie die Zuordnungen der Gausszentren $\Gamma^p_{\mathcal{A}}$ zu Verfügung.

(S.2) Es wird ein Knoten $a_i \in K_{\mathcal{A}}$ so bestimmt, daß für alle Knoten $a_j \in K_{\mathcal{A}}$ mit $a_j \neq a_i$ die Anzahl der Kanten aus $B^{\mathcal{A}}_{FT}$, die im kürzesten a_i-a_j-Weg in $G_{\mathcal{A}}$ enthalten sind, möglichst gering ist. Dabei werden die kürzesten a_i-a_j-Wege in $G_{\mathcal{A}}$ mit einem BFS-Algorithmus (siehe hierzu [9]) bestimmt. Die Zusammenhangskomponente des Graphen $G'_{\mathcal{A}} = (K_{\mathcal{A}},\ E_{\mathcal{A}})$, mit $E_{\mathcal{A}} = B_{\mathcal{A}} \setminus B^{\mathcal{A}}_{FT}$, zu der das Atom a_i gehört, wird zum Hauptfragment F_0.

(S.3) Die Zusammenhangskomponenten des Graphen $G'_{\mathcal{A}}$, die Molekülfragmente F_i, werden bestimmt und ausgehend vom Hauptfragment F_0 mit einem BFS-Algorithmus numeriert. Für $i = 0, \dots, n_{ft}$, $n_{ft} := n_{ft}[\mathcal{A}] = \left|B^{\mathcal{A}}_{FT}\right|$, wird dem Fragment F_i der Knoten f_i zugeordnet. Damit wird die Abbildung $\Psi_{K_{\mathcal{A}},F_{\mathcal{A}}}$ festgelegt, wobei $F_{\mathcal{A}} = \{f_0, \dots, f_{n_{ft}}\}$ gilt.

(S.4) Der gerichtete Fragmentebaum $\vec{T}_{\mathcal{A}} = (F_{\mathcal{A}},\ \vec{E}^{\mathcal{A}}_{FT})$ mit der Wurzel f_0 wird aufgebaut. Dabei werden die Kanten aus $E^{\mathcal{A}}_{FT} = \Psi_{K_{\mathcal{A}},F_{\mathcal{A}}}(B^{\mathcal{A}}_{FT})$ so gerichtet, daß für $f_i \neq f_0$ der Baum $\vec{T}_{\mathcal{A}}$ einen gerichteten f_0-f_i-Weg enthält. Der Baum $\vec{T}_{\mathcal{A}}$ ist balanciert, denn das zu f_0 gehörende Hauptfragment F_0 ist so bestimmt worden, daß die f_0-f_i-Wege in $\vec{T}_{\mathcal{A}}$ möglichst kurz sind. Für $i = 0, \dots, n_{ft}$ werden dem Knoten $f_i \in F_{\mathcal{A}}$ die zum Fragment F_i gehörenden Atomkoordinaten $a_{i_l} \in K_{\mathcal{A}}$ und die entsprechenden Parameter der Gaussfunktionen zugeordnet. Für $f_i \in F_{\mathcal{A}}$, $f_i \neq f_0$, wird die Kante $e_i = \{a_{k_j},\ a_{i_l}\} \in B^{\mathcal{A}}_{FT}$ bestimmt, für die $\Psi_{K_{\mathcal{A}},F_{\mathcal{A}}}(e_i) = (f_k,\ f_i) \in \vec{E}^{\mathcal{A}}_{FT}$ gilt. Der Knoten $a_{k_j} \in K_{\mathcal{A}}$ gehört zum Fragment F_k und wird f_i als Referenzkoordinate $\mathsf{R}_{F_{\mathcal{A}},K_{\mathcal{A}}}(f_i)$ zugeordnet. Der Knoten a_{i_l} gehört zum Fragment F_i und wird f_i als $\mathsf{S}_{F_{\mathcal{A}},K_{\mathcal{A}}}(f_i)$ zugeordnet. Die Drehachse $d_i = a_{i_l} - a_{k_j}$ wird f_i als $\mathsf{D}_{F_{\mathcal{A}},\mathbb{R}^3\setminus\{0\}}(f_i)$ zugeordnet. Auch der Torsionswinkel $\alpha_{e_i}(\mathsf{C}_{\mathcal{A}})$, der an der Bindung e_i gemessen worden ist, wird dem Knoten f_i zugeordnet. Schließlich wird dem Knoten f_i der interne Rotationsparameter φ_i zugeordnet, wobei φ_i der Parameter der Rotation um die lokale Drehachse d_i ist. Insgesamt ist für das Molekül $\mathcal{A}$ die Datenstruktur $\mathcal{DS}_{\mathcal{A}}(\mathsf{C}_{\mathcal{A}})$ und die Transformationsstruktur $\mathcal{S}(K_{\mathcal{A}}, \vec{T}_{\mathcal{A}})$ bestimmt worden.

(S.5) Der Schwerpunkt des Moleküls $\mathcal{A}$ $c_{\mathcal{A}} := (1/n_a)\sum_{i=1}^{n_a} a_i \in \mathbb{R}^3$ wird berechnet, wobei n_a die Anzahl der Atome des Moleküls $\mathcal{A}$ ist. Für alle Atomkoordinaten und Gausskoordinaten werden lokale Koordinaten berechnet, d.h. falls die Atomkoordinate a bzw. die Gausskoordinate g dem Knoten f_i zugeordnet worden ist, so wird $a - \mathsf{R}_{F_{\mathcal{A}},K_{\mathcal{A}}}(f_i)$ bzw. $g - \mathsf{R}_{F_{\mathcal{A}},K_{\mathcal{A}}}(f_i)$ berechnet. Aus Korollar 2.4.17 wird deutlich, daß sich die Hintereinanderschaltung von lokalen Rotationen (siehe Satz 2.4.16) als Linearkombination von lokalen Koordinaten, auf die geeignete Rotationsmatrizen angewendet werden, darstellen läßt. Für f_0 wird der Molekülschwerpunkt als Referenzkoordinate verwendet, d.h. $\mathsf{R}_{F_{\mathcal{A}},K_{\mathcal{A}}}(f_0) := c_{\mathcal{A}}$. Dann wird für jede physikochemische Eigenschaft $p \in \mathcal{I}_{PE}$ der entsprechende Hodgkin-Index bzgl. der Überlagerung des Moleküls $\mathcal{A}$ mit sich selbst und insbesondere der konstanten Anteil des Indizes berechnet. Weiter wird das n-Potential für die gegebene Konformation $\mathsf{C}_{\mathcal{A}}$ und, falls $B^{\mathcal{A}}_{FT} \neq \emptyset$ gilt, auch die Bewertungsfunktion für die Torsionswinkel berechnet.

Nachdem die Einleseroutine 4.1 für jedes flexible Moleküle $\mathcal{A}_i$, $i = 1, \dots, n_m$, ausgeführt worden ist, wobei n_m die Anzahl der in Modell 1 bzw. in Modell 2 ausgewählten Moleküle ist, wird in Modell 1 für jedes Molekül $\mathcal{A}_i$, $i = 1, \ldots, n_m$, der Alignmentindex

$$\mathsf{A}^{\mathsf{H}}_{\mathcal{A}_i,\mathcal{R}} \left(\left(Q^p_{\mathcal{A}_i}(x)\right)_{p \in \mathcal{I}^{\mathcal{A}_i,\mathcal{R}}_{PE}}, \left(Q^p_{\mathcal{R}}(x)\right)_{p \in \mathcal{I}^{\mathcal{A}_i,\mathcal{R}}_{PE}} \right) \tag{4.1}$$

berechnet. Die Alignmentindizes bewerten für $i = 1, \ldots, n_m$ die jeweilige Überlagerung des Moleküls $\mathcal{A}_i$ bzgl. des Referenzmoleküls $\mathcal{R}$, wie sie durch die Konformationen $\mathsf{C}_{\mathcal{A}_i}$ und $\mathsf{C}_{\mathcal{R}}$ gegeben ist. Nach der Optimierung wird der Alignmentindex $\mathsf{A}^{\mathsf{H}}_{\mathcal{A}_i,\mathcal{R}}$ für $i = 1, \ldots, n_m$ erneut berechnet, jedoch mit den transformierten (Gauss-)Koordinaten für das Molekül $\mathcal{A}_i$, um zu überprüfen, inwieweit die Optimierung den Alignmentindex verbessern konnte. Das Ziel der Optimierung ist es, die Überlagerung des Moleküls $\mathcal{A}_i$ bzgl. des Moleküls $\mathcal{R}$ zu verbessern. Zum Abschluß wird für jedes flexible Molekül dessen Kristallstruktur FlexS bekannt ist, der *RMS*-Abstand[4] zwischen der Molekülkonformation und der Konformation der Kristallstruktur berechnet. Nach Beendigung der Optimierung wird der *RMS*-Abstand zwischen der transformierten Konformation und der Konformation der Kristallstruktur berechnet, um zu überprüfen, ob der *RMS*-Abstand verbessert werden konnte. Und wenn ja, inwieweit die Konformation, die in der Bindungstasche des Rezeptors bindet und durch die Kristallstruktur gegeben ist, vorhergesagt werden konnte.

In Modell 2 wird hingegen die Summe der Alignmentindizes

$$\frac{2}{n_m(n_m-1)} \sum_{i=1}^{n_m-1} \sum_{j=i+1}^{n_m} \mathsf{A}^{\mathsf{H}}_{\mathcal{A}_i,\mathcal{A}_j} \left(\left(Q^p_{\mathcal{A}_i}(x)\right)_{p \in \mathcal{I}^{\mathcal{A}_i,\mathcal{A}_j}_{PE}}, \left(Q^p_{\mathcal{A}_j}(x)\right)_{p \in \mathcal{I}^{\mathcal{A}_i,\mathcal{A}_j}_{PE}} \right) \tag{4.2}$$

und für $i = 1, \ldots, n_m$ die Summe

$$\frac{1}{n_m-1} \sum_{\substack{j=1 \\ i \neq j}}^{n_m} \mathsf{A}^{\mathsf{H}}_{\mathcal{A}_i,\mathcal{A}_j} \left(\left(Q^p_{\mathcal{A}_i}(x)\right)_{p \in \mathcal{I}^{\mathcal{A}_i,\mathcal{A}_j}_{PE}}, \left(Q^p_{\mathcal{A}_j}(x)\right)_{p \in \mathcal{I}^{\mathcal{A}_i,\mathcal{A}_j}_{PE}} \right) \tag{4.3}$$

berechnet. Während die Summe in Gleichung (4.2) die gegenseitige Überlagerung aller Moleküle $\mathcal{A}_i$, $i = 1, \ldots, n_m$, wie sie durch die Konformationen $\mathsf{C}_{\mathcal{A}_i}$, $i = 1, \ldots, n_m$, gegeben ist, bewertet, entspricht die Summe in Gleichung (4.3) der Überlagerung des Moleküls $\mathcal{A}_i$, $i = 1, \ldots, n_m$, bzgl. der Moleküle $\mathcal{A}_j$, $j = 1, \ldots, n_m$, $j \neq i$. Das Gauß-Seidel-artige Optimierungsverfahren aus Kapitel 4.3 verwendet die Werte aus Gleichung (4.3). Auch hier werden nach der Optimierung die obigen Summen erneut berechnet, wobei die transformierten (Gauss-)Koordinaten verwendet werden, um zu überprüfen inwieweit durch die Optimierung eine Verbesserung erreicht werden konnte. Auch hier besteht das Ziel der Optimierung darin, die gegenseitige Überlagerung zu verbessern.

Das Minimierungsmodul:

Im folgenden seien $\mathcal{A}_1, \ldots, \mathcal{A}_{n_m}$ die im Einlesemodul ausgewählten flexiblen Moleküle.

In Modell 1 wird für $i = 1, \ldots, n_m$ die Energiefunktion $\mathsf{E}_{\mathcal{A}_i,\mathcal{R}}$ aus Definition 3.2.2 minimiert, wobei $\mathcal{R}$ das Referenzmolekül ist. Als Startvektor wird $x_0 := u^0_{\mathcal{A}_i} = (t, q, c, \varphi_{\mathcal{A}_i}) \in U_n(\mathcal{A}_i) := \mathbb{R}^3 \mathrm{x} \mathbb{R}^4 \setminus \{0\} \mathrm{x} \mathbb{R}^3 \mathrm{x} \mathbb{R}^{n_{ft}[\mathcal{A}_i]}$ mit $t = (0,0,0)$, $q = (0,0,0,1)$, $\varphi_{\mathcal{A}_i} = (0, \ldots, 0)$ und $c := c_{\mathcal{A}_i}$ verwendet, wobei $c_{\mathcal{A}_i}$ der Molekülschwerpunkt von $\mathcal{A}_i$ ist. Nach Lemma 2.4.28 d) entspricht der

[4] Siehe Anhang B.

Transformationsoperator mit dem Parameter $u^0_{\mathcal{A}_i}$ der Identität. Dabei wird vorausgesetzt, daß der Transformationsoperator aus der Translation, der globalen Rotation um ein variables Rotationszentrum und dem lokalen Rotationsoperator besteht. Da im Schritt 0 des Algorithmus 4.1 und im Steuerungsmodul der Transformationsoperator modifiziert werden kann, können Komponenten des Startvektors wie der Translationsvektor oder das Rotationszentrum, sowie einzelne Rotationsparameter fehlen.
In Modell 2 wird die Energiefunktion $\mathsf{E}_{\mathcal{A}_1,\dots,\mathcal{A}_{n_m}}$ aus Definition 3.3.3 minimiert. Als Startwert wird $x_0 = (u^0_{\mathcal{A}_1},\dots,u^0_{\mathcal{A}_{n_m}})$ verwendet, wobei $u^0_{\mathcal{A}_i}$ für $i = 1,\dots,n_m$ dem Startwert aus Modell 1 entspricht. In Kapitel 4.3 wird ein Gauß-Seidel-artiger Algorithmus vorgestellt, der die Energiefunktion aus Definition 3.3.3 effizienter minimieren soll. Dieser Algorithmus kann über das Steuerungsmodul ausgewählt werden.

Um die entsprechende Energiefunktion E zu minimieren, wird der folgende Algorithmus verwendet. Die Einzelheiten, wie die Berechnung der Abstiegsrichtung und der Schrittweite, hängen dabei von den in Kapitel 4.2 dargestellten Minimierungs- und Schrittweitenverfahren ab. Im Algorithmus selber werden vor allem die Stoppkriterien dargestellt, die verwendet werden, um den den Algorithmus zu beenden. Bis auf das Gradientenkriterium (Schritt 1) sind die weiteren Stoppkriterien optional.

Algorithmus 4.2 (Minimierungsalgorithmus)

(S.0) Startwert: x_0
Stoppkriterien: $\varepsilon_G > 0$, $\varepsilon_S > 0$, $\varepsilon_E > 0$ *und MAXSTEP.*
Berechne $\mathsf{E}(x_0)$.

(S.1) Berechne $\nabla\mathsf{E}(x_k)$. *Falls* $\|\nabla\mathsf{E}(x_k)\|_\infty < \varepsilon_G$, *STOP.*

(S.2) Berechne Abstiegsrichtung $d_k := d(\nabla\mathsf{E}(x_k))$.

(S.3) Berechne Schrittweite $\alpha_k = \alpha\,(x_k,\ d_k,\ \nabla\mathsf{E}(x_k),\ \mathsf{E}(x_k)) > 0$.
Falls die Schrittweite α_k *nicht berechnet werden konnte, ABRUCH.*

(S.4) Berechne $x_{k+1} = x_k + \alpha_k d_k$. *Falls*

$$\max_{i=1,\dots,n}\left(\frac{|x^i_{k+1} - x^i_k|}{\max\{x^i_{k+1},\ 1.0\}}\right) \ < \ \varepsilon_S,$$

STOP.[5]

(S.5) Falls $|\mathsf{E}(x_{k+1}) - \mathsf{E}(x_k)| < \varepsilon_E$, *STOP.*

(S.6) Falls k=MAXSTEP, STOP. Sonst setze $k = k+1$ *und gehe zu (S.1).*

Es gibt 5 Möglichkeiten für den Minimierungsalgorithmus zu terminieren. Der Anwender hat auf einen Abbruch in Schritt 3 keinen direkten Einfluß, denn die hier implementierten Schrittweitenverfahren versuchen iterativ, ein $\alpha > 0$ so zu bestimmen, daß

$$\mathsf{E}(x_k + \alpha d_k) \ < \ \mathsf{E}(x_k) - s(\alpha, d_k, \nabla\mathsf{E}(x_k)),\ \text{mit}\ s(\alpha, d_k, \nabla\mathsf{E}(x_k)) > 0, \tag{4.4}$$

[5] n bezeichnet die Dimension des Optimierungsproblems.

gilt. Es kann jedoch vorkommen, daß die verwendeten Iterationsverfahren nicht in der Lage sind, solch ein α iterativ zu bestimmen, insbesondere da die Verfahren selbst parameterabhängig sind.[6] Um auch in solch einem Fall zu garantieren, daß das verwendete Interationsverfahren terminiert, wurde die Anzahl der Iterationen beschränkt. Die implementierten Schrittweitenverfahren können maximal 100 Iterationen ausführen, um eine geeignete Schrittweite $\alpha > 0$ zu bestimmen. Wird innerhalb der 100 Iterationen kein geeignetes $\alpha > 0$ berechnet, so bricht das Minimierungsverfahren ab und liefert als Ergebnis x_k und $\mathsf{E}(x_k)$. Ein Abbruch in Schritt 3 bedeutet daher, daß das verwendete Schrittweitenverfahren nicht in der Lage gewesen ist, in einer angemessen Anzahl von Iterationen ein x_{k+1} zu bestimmen, mit $\mathsf{E}(x_{k+1})$ echt kleiner als $\mathsf{E}(x_k)$.

Es stellt sich Frage, in welchen Fällen die Iterationsverfahren keine geeignete Schrittweite bestimmen können. Falls x_k in der Umgebung eines lokalen Minimimums x^* liegt, d.h. falls $x_k \in B_\varepsilon(x^*) = \{y \in U_n \mid \|x^* - y\|_2 < \varepsilon\}$ mit $\varepsilon > 0$, und $\mathsf{E}(x_k + \alpha d_k) \sim \mathsf{E}(x_k)$ für alle $\alpha > 0$ gilt, kann es vorkommen, daß ein Iterationsverfahren kein geeignetes $\alpha > 0$ bestimmen kann, so daß die obige Ungleichung (4.4) erfüllt wird.[7] Insbesondere da die Genauigkeit der Zahlendarstellung im Computer beschränkt ist. Daraus folgt, daß das Ergebnis durchaus akzeptabel sein kann, wenn der Algorithmus 4.2 in Schritt 3 terminiert. Sollte jedoch das Ergebnis nicht zufriedenstellend sein, so muß man entweder andere Parameter für das Schrittweitenverfahren oder aber ein anderes Verfahren wählen. Das Schrittweitenverfahren liefert neben der Schrittweite α_k den Funktionswert $\mathsf{E}(x_k + \alpha_k d_k)$.

Die anderen Stoppkriterien sind das Gradientenkriterium (Schritt 1), das Schrittkriterium (Schritt 4), das Energiekriterium (Schritt 5) und die Vorgabe der maximalen Anzahl von Iterationen (Schritt 6), die der Algorithmus 4.2 ausführen kann. Bis auf das Gradientenkriterium sind die anderen 3 Kriterien optional, d.h. sie können deaktiviert werden. Das Schrittkriterium wird in dieser Form von Dennis und Schnabel in [12] als ein geeignetes Stoppkriterium verwendez. Indem man die maximale Anzahl der Iterationen vorgibt, nämlich MAXSTEP>0, erreicht man, daß der Minimierungsalgorithmus auf jeden Fall terminiert. Sollte jedoch das Ergebnis nach MAXSTEP Iterationen nicht zufriedenstellend sein, kann man die Optimierung fortsetzen. Mit MAXSTEP=0 wird die Abfrage in Schritt 6 deaktiviert. Als Ergebnis liefert der Algorithmus 4.2 x^* und $\mathsf{E}(x^*)$. Nachdem alle Energiefunktionen minimiert worden sind, werden die Ergebnisse an das Ergebnismodul übergeben.

Das Ergebnismodul:

Für $i = 1, \ldots, n_m$ sei $u^*_{\mathcal{A}_i}$ der mit dem Algorithmus 4.2 bestimmte Tranformationsparameter für das flexible Moleküle $\mathcal{A}_i$. Zuerst wird unabhängig vom Modell für $i = 1, \ldots, n_m$ der zum Molekül $\mathcal{A}_i$ gehörende Transformationsoperator mit dem Parameter $u^*_{\mathcal{A}_i}$ auf die Atomkoordinaten von $\mathcal{A}_i$, sowie auf die Zentren der zu $\mathcal{A}_i$ gehörendenden Gaussfunktionen angewendet.

[6] Für die Einzelheiten siehe Kapitel 4.2

[7] Solche Beispiele kommen in den Rechnungen durchaus vor.

In Modell 1 wird dann für $i = 1, \ldots , n_m$ der Alignmentindex

$$\mathsf{A}^{\mathsf{H}}_{\mathcal{A}_i,\mathcal{R}} \left(\ (Q^p_{\mathcal{A}_i}[u^*_{\mathcal{A}_i}](x))_{p\in\mathcal{I}^{\mathcal{A}_i,\mathcal{R}}_{PE}} \ , \ (Q^p_{\mathcal{R}}(x))_{p\in\mathcal{I}^{\mathcal{A}_i,\mathcal{R}}_{PE}} \ \right) \tag{4.5}$$

berechnet, wobei $\mathcal{R}$ das Referenzmolekül ist. Weiter wird das n-Potential bzgl. der transformierten Atomkoordinaten von $\mathcal{A}_i$ und gegebenenfalls die Bewertungsfunktion für die Torsionswinkel berechnet. Dann wird der *RMS*-Abstand zwischen den Koordinaten der Startkonformation $\mathsf{C}_{\mathcal{A}_i}$ und den transformierten Atomkoordinaten berechnet. Falls die Kristallstruktur von $\mathcal{A}_i$ bekannt ist, wird auch der *RMS*-Abstand zwischen den Koordinaten der Kristallstruktur und transformierten Koordinaten von $\mathcal{A}_i$ berechnet. Schließlich werden die Werte $\mathsf{E}_{\mathcal{A}_i,\mathcal{R}}(u^0_{\mathcal{A}_i})$ und $\mathsf{E}_{\mathcal{A}_i,\mathcal{R}}(u^*_{\mathcal{A}_i})$ berechnet. Zum Schluß werden die berechneten Ergebnisse ausgegeben, und zwar jeweils der Wert vor der Optimierung und der Wert nach der Optimierung. So wird der Wert des Alignmentindizes vor der Optimierung (Gleichung (4.1)) und der Wert nach der Optimierung (Gleichung (4.5)) ausgeben. Indem man diese beiden Werte vergleicht, kann man entscheiden, inwieweit die Optimierung die Überlagerung des Moleküls $\mathcal{A}_i$ auf das Molekül $\mathcal{R}$ verbessern konnte. Dann werden einige Statistiken des Minimierungsverfahrens angegeben, wie z.B die Anzahl der Iterationen und die Rechenzeit, die der Algorithmus 4.2 brauchte, um $u^*_{\mathcal{A}_i}$ zu bestimmen, sowie das Stoppkriterium, das den Minimierungsalgorithmus beendete.
In Modell 2 wird zunächst die Summe der Alignmentindizes

$$\frac{2}{n_m(n_m-1)} \sum_{i=1}^{n_m-1} \sum_{j=i+1}^{n_m} \mathsf{A}^{\mathsf{H}}_{\mathcal{A}_i,\mathcal{A}_j} \left(\ (Q^p_{\mathcal{A}_i}[u^*_{\mathcal{A}_i}](x))_{p\in\mathcal{I}^{\mathcal{A}_i,\mathcal{A}_j}_{PE}} \ , \ \left(Q^p_{\mathcal{A}_j}[u^*_{\mathcal{A}_j}](x)\right)_{p\in\mathcal{I}^{\mathcal{A}_i,\mathcal{A}_j}_{PE}} \ \right) \tag{4.6}$$

berechnet. Dann werden für $i = 1, \ldots , n_m$ die Summe

$$\frac{1}{n_m-1} \sum_{\substack{j=1\\ i\neq j}}^{n_m} \mathsf{A}^{\mathsf{H}}_{\mathcal{A}_i,\mathcal{A}_j} \left(\ (Q^p_{\mathcal{A}_i}[u^*_{\mathcal{A}_i}](x))_{p\in\mathcal{I}^{\mathcal{A}_i,\mathcal{A}_j}_{PE}} \ , \ \left(Q^p_{\mathcal{A}_j}[u^*_{\mathcal{A}_j}](x)\right)_{p\in\mathcal{I}^{\mathcal{A}_i,\mathcal{A}_j}_{PE}} \ \right)$$

berechnet. Weiter wird das n-Potential bzgl. der transformierten Atomkoordinaten von $\mathcal{A}_i$ und gegebenenfalls die Bewertungsfunktion für die Torsionswinkel berechnet. Schließlich werden die Werte $\mathsf{E}_{\mathcal{A}_1,\ldots,\mathcal{A}_{n_m}}(u^0_{\mathcal{A}_1}, \ldots , u^0_{\mathcal{A}_{n_m}})$ und $\mathsf{E}_{\mathcal{A}_1,\ldots,\mathcal{A}_{n_m}}(u^*_{\mathcal{A}_1}, \ldots , u^*_{\mathcal{A}_{n_m}})$ berechnet. Zum Schluß erfolgt, wie in Modell 1, die Ausgabe der berechneten Ergebnisse sowie einiger Statistiken des Minimierungsverfahrens. So wird die Summe der Alignmentindizes aus Gleichung (4.2) sowie die Summe aus Gleichung (4.6) ausgegeben. Indem man diese beiden Werte miteinander vergleicht, kann man entscheiden, inwieweit die Optimierung die gegenseitige Überlagerung der Moleküle $\mathcal{A}_i$, $i = 1, \ldots , n_m$ verbessern konnte.

Das Steuerungsmodul:

Das Steuerungsmodul gehört nicht zum eigentlichen Gesamtalgorithmus (Einlesemodul, Minimierungsmodul und Ergebnismodul), sondern ist eine Abfrage nach Parametern, die die Energiefunktionen aus Kapitel 3.2 und 3.3 bzw. den Transformationsoperator aus Kapitel 2.4.3 modifizieren und den Minimierungsalgorithmus 4.2 steuern.

Damit sind die Komponenten des implementierten Gesamtalgorithmus schematisch kurz dargestellt worden. Weitere Aspekte der Implementierung werden in Kapitel 5 angesprochen.

4.2 Die Minimierungsalgorithmen

In diesem Kapitel werden die Minimierungsverfahren, die implementiert worden sind, um die Energiefunktionen aus Kapitel 3.2 und Kapitel 3.3 zu minimieren, kurz vorgestellt. Die hier betrachteten Verfahren erzeugen iterative Folgen $\{x_k\} \subset X$ durch sukzessive Bestimmung von Suchrichtungen d_k und Schrittweiten α_k. Es wird $x_{k+1} = x_k + \alpha_k d_k$ für $k = 0, 1, \dots$ gesetzt, wobei $x_0 \in X$ der Startpunkt ist.[8] Die betrachteten Verfahren unterscheiden sich durch die Bestimmung von $d_k \in X$ und $\alpha_k > 0$. Alle hier vorgestellten Verfahren gehören jedoch zur Klasse der *Abstiegsverfahren*, d.h. für nicht minimale x_k gilt stets $f(x_{k+1}) < f(x_k)$, $k = 0, 1, \dots$, wobei f die Energiefunktion bezeichnet. Ein $d_k \in X$ heißt *Abstiegsrichtung* von f im Punkt x_k, falls ein $\tilde{\alpha} = \tilde{\alpha}(d_k)$ existiert mit $f(x_k + \alpha d_k) < f(x_k)$ für alle $\alpha \in (0,\ \tilde{\alpha})$. Das folgende Lemma charakterisiert ganz allgemein Abstiegsrichtungen.

Lemma 4.2.1
Es sei X ein Hilbertraum und $f: X \to \mathbb{R}$ differenzierbar. Ein Vektor $d \in X$ ist Abstiegsrichtung im Punkt $x \in X$, wenn $< f'(x),\ d >_X < 0$ gilt, wobei $< .,\ . >_X$ das Skalarprodukt des Hilbertraumes ist.

Beweis:
Folgt direkt aus der Definition der Fréchet-Differenzierbarkeit.[9] ■

Um die Abstiegsrichtung d_k zu bestimmen, werden Gradientenverfahren, CG-Verfahren und Quasi-Newton-Verfahren verwendet. Um die Schrittweite α_k zu bestimmen, kann man die *Strahlminimierung* verwenden, d.h. man bestimmt $\alpha_k > 0$ so, daß $f(x_k + \alpha_k d_k) \leq f(x_k + \alpha d_k)$ für alle $\alpha > 0$ gilt. Da dies jedoch ein eindimensionales Minimierungsproblem und nur aufwendig zu lösen ist, verwendet man stattdessen iterative Schrittweitenverfahren. Die Schrittweitenverfahren werden nach den Minimierungsalgorithmen vorgestellt. Alle Verfahren, die hier dargestellt werden, sind auch implementiert worden. Als Grundlage diente die folgende Literatur: [10, 12, 19, 28, 37, 46, 53, 54, 55]. Für jedes Verfahren wird im einzelnen die genaue Quelle angegeben.

Die Minimierungsverfahren:

Das einfachste Abstiegsverfahren ist das *Gradientenverfahren* bzw. das Verfahren des *stärksten Abstiegs*. Zur Herleitung des Verfahrens siehe man Großmann [28] oder Polak [46].

Algorithmus 4.3 (Gradientenverfahren (G))

(S.0) Daten: $x_0 \in X$.

(S.1) Berechne $d_k = -\nabla f(x_k)$, *falls* $\nabla f(x_k) = 0$ *STOP.*

(S.2) Berechne Schrittweite $\alpha_k := \alpha(d_k) > 0$.

(S.3) $x_{k+1} := x_k + \alpha_k d_k$.

[8] X ist der zulässige Bereich (siehe Kapitel 3.1) bzw. der Suchraum, z.B. $X = \mathbb{R}^3 \times \mathbb{R}^4 \setminus \{0\} \times \mathbb{R}^3 \times \mathbb{R}^n$.
[9] Zu Fréchet-Differenzierbarkeit siehe Definition C.20 und C.21 in Anhang C.

(S.4) Setze $k \leftarrow k+1$ und gehe zu (S.1).

Eine weitere Klasse von Abstiegsverfahren, die nur den Gradienten der Energiefunktion verwenden, sind die *CG-Verfahren.* Zur Herleitung der Verfahren siehe man Polak [46], Geiger/Kanzow [19] oder Großmann [28]. Die einzelnen CG-Verfahren unterscheiden sich nur in der Berechnung des Parameters γ_k in (4.7).

Algorithmus 4.4 (CG-Verfahren (CG))

(S.0) Daten: $x_0 \in X$.
Setze $g_0 = \nabla f(x_0)$ und $d_0 = -g_0$.

(S.1) Falls $g_k = \nabla f(x_k) = 0$ STOP, sonst berechne Schrittweite $\alpha_k := \alpha(d_k) > 0$.

(S.2) Berechne:

$$\begin{aligned} x_{k+1} &= x_k + \alpha_k d_k \\ g_{k+1} &= \nabla f(x_{k+1}) \\ \gamma_k & \\ d_{k+1} &= -g_{k+1} + \gamma_k d_k. \end{aligned} \tag{4.7}$$

(S.3) Setze $k \leftarrow k+1$ und gehe zu (S.1).

Das *Fletcher-Reeves-CG-Verfahren* (CG-FR) berechnet γ_k in (4.7) wie folgt:

$$\gamma_k \;:=\; \gamma_k^{FR} \;=\; \frac{< g_{k+1},\; g_{k+1} >_2}{< g_k,\; g_k >_2}.$$

Polak ([46]) (siehe aber auch [28]) empfiehlt eine Variante des Flechter-Reeves Verfahrens, und zwar empfiehlt er für den Parameter γ_k^{FR} die folgende Berechnung:

$$\gamma_k^{FR} \;=\; \begin{cases} \frac{<g_{k+1},\; g_{k+1}>_2}{<g_k,\; g_k>_2} & ,\text{falls } k \;\neq\; i\; n,\ \forall i \in \mathbb{N} \\ 0 & ,\text{falls } \exists i \in \mathbb{N}:\, k = i\; n, \end{cases}$$

wobei n die Dimension des Optimierungsproblems ist. D.h. in jedem n-ten Iterationsschritt erfolgt ein sogenannter Neustart. Man kann auch sagen, daß in jeder n-ten Iteration das CG-Verfahren zu einem Gradientenverfahren wird. Diese Variante wird in den Beispielrechnungen mit CG-FRN bezeichnet.

Ein weiteres CG-Verfahren ist das *Polak-Ribiére-CG-Verfahren* (CG-PR). In diesem Verfahren setzt man:

$$\gamma_k \;:=\; \gamma_k^{PR} \;=\; \frac{< g_{k+1} - g_k,\; g_{k+1} >_2}{< g_k,\; g_k >_2}.$$

Auch für das Polak-Ribiére Verfahren gibt es die Variante mit dem Neustart (siehe [46], [19] und [28]), d.h. der Parameter γ_k^{PR} wird wie folgt berechnet:

$$\gamma_k^{PR} \;=\; \begin{cases} \frac{<g_{k+1}-g_k,\; g_{k+1}>_2}{<g_k,\; g_k>_2} & ,\text{falls } k \;\neq\; i\; n,\ \forall i \in \mathbb{N} \\ 0 & ,\text{falls } \exists i \in \mathbb{N}:\, k = i\; n, \end{cases}$$

wobei n die Dimension des Optimierungsproblems ist. Diese Variante wird in den Beispielrechnungen mit CG-PRN bezeichnet.
Zum Abschluß wird das CG-Verfahren von Gilbert und Nocedal [20] (CG-GN) betrachtet (siehe auch [19]). Der Parameter γ_k wird dabei wie folgt berechnet:

$$\gamma_k := \begin{cases} -\gamma_k^{FR} & \text{, falls } \gamma_k^{PR} < -\gamma_k^{FR} \\ \gamma_k^{PR} & \text{, falls } \left|\gamma_k^{PR}\right| \leq \gamma_k^{FR} \\ \gamma_k^{FR} & \text{, falls } \gamma_k^{PR} > \gamma_k^{FR}, \end{cases}$$

wobei γ_k^{FR} der Parameter des Flechter-Reeves Verfahrens und γ_k^{PR} der Parameter des Polak-Ribiére Verfahrens ist.

Zum Schluß wird die Klasse der *Quasi-Newton* Verfahren betrachtet. Zur Herleitung der Verfahren bzw. der Aufdatierungsformel siehe man Dennis/Schnabel [12], Geiger/Kanzow [19], Großmann [28], Kosmol [37], Polak [46], Stoer [54] und Stoer/Jarre [55]. Die einzelnen Quasi-Newton Verfahren unterscheiden sich in der Aufdatierung der Matrix H_k zur Matrix H_{k+1} in Gleichung (4.8).

Algorithmus 4.5 (Quasi-Newton Verfahren (QN))

(S.0) Daten: $x_0 \in X$ *und* $H_0 \in \mathsf{M}(n \times n; \mathbb{R})$ *eine symmetrische und positiv definite Matrix.*[10]
Setze $g_0 = \nabla f(x_0)$ *und berechne* $d_0 = -H_0\ g_0$.

(S.1) Falls $g_k = \nabla f(x_k) = 0$ *STOP, sonst berechne Schrittweite* $\alpha_k := \alpha(d_k) > 0$.

(S.2) Berechne:

$$\begin{aligned} x_{k+1} &= x_k + \alpha_k d_k \\ g_{k+1} &= \nabla f(x_{k+1}) \\ s_k = x_{k+1} - x_k &= \alpha_k d_k \\ y_k = g_{k+1} - g_k &= \nabla f(x_{k+1}) - \nabla f(x_k). \end{aligned}$$

(S.3) Falls $< s_k,\ y_k >_2 > 0$ *gilt, berechne die Matrix*

$$H_{k+1} = H(H_k,\ s_k,\ y_k). \tag{4.8}$$

Ist die Bedingung jedoch nicht erfüllt, so wird das Verfahren mit einer Matrix H_{k+1}*, die symmetrisch und positiv definit ist, fortgesetzt. Berechne* $d_{k+1} = -H_{k+1}\ g_{k+1}$.

(S.4) Setze $k \leftarrow k+1$ *und gehe zu (S.1)*

Bemerkung: Die Motivation der Quasi-Newton Verfahren besteht darin, die Hesse-Matrix $\nabla^2 f(x_k)$ bzw. $(\nabla^2 f(x_k))^{-1}$ „hinreichend gut“ durch leicht berechenbare Matrizen H_k zu approximieren und in jedem Schritt $d_k = -H_k^{-1}\nabla f(x_k)$ bzw. $d_k = -H_k \nabla f(x_k)$ zu setzen.

[10] $\mathsf{M}(n \times n; \mathbb{R})$ ist der Raum der nxn-Matrizen (siehe Anhang B), wobei n die Dimension des Optimierungsproblems angibt.

Als Startmatrix H_0 schlagen Dennis/Schnabel [12] die Matrix $H_0 := \max\{|f(x_0)|\,,\ 1.0\}E_n$ bzw. $H_0 := (1.0/\max\{|f(x_0)|\,,\ 1.0\})E_n$ für die inverse Aufdatierung vor, wobei $E_n \in \mathsf{M}(n \times n; \mathbb{R})$ die Einheitsmatrix ist. Ist die Bedingung $< s_k,\ y_k >_2 > 0$ in Schritt 3 nicht erfüllt, so wird das Verfahren mit der Wahl von H_{k+1} sozusagen neu gestartet, wobei analog zur Startmatrix H_0 die Matrix $H_{k+1} := \max\{|f(x_{k+1})|\,,\ 1.0\}E_n$ bzw. $H_{k+1} := (1.0/\max\{|f(x_{k+1})|\,,\ 1.0\})E_n$ für die inverse Aufdatierung gewählt wird. Die Abstiegsrichtung erhält man dann durch $d_{k+1} := -(\max\{|f(x_{k+1})|\,,\ 1.0\})^{-1}g_{k+1}$. Die Anzahl der Neustarts, die ein Quasi-Newton Verfahren braucht, wird im Ergebnismodul (siehe Kapitel 4.1) ausgegeben, um das gewählte Verfahren beurteilen zu können.

Im folgenden sei H bzw. $B \in \mathsf{M}(n \times n; \mathbb{R})$ symmetrisch und positiv definit, sowie $s, y \in \mathbb{R}^n$ mit $s^\mathsf{T} y =< s,\ y >_2 > 0$. Dadurch ist gewährleistet, daß die aufdatierte Matrix H_+ bzw. B_+ ebenfalls symmetrisch und positiv definit ist.[11]

Das bekannteste Quasi-Newton Verfahren ist wohl das *Broyden-Fletcher-Goldfarb-Shanno* Verfahren (BFGS). Die BFGS-Aufdatierungsformel lautet mit den obigen Bezeichnungen:

$$H_+^{BFGS} := H + \frac{yy^\mathsf{T}}{< s,\ y >_2} - \frac{Hss^\mathsf{T}H}{s^\mathsf{T}Hs}, \tag{4.9}$$

wobei yy^T bzw. ss^T nxn-Matrizen sind. Anstelle der Aufdatierungsformel (4.9) kann man auch die *inverse* BFGS-Aufdatierungsformel verwenden. Die inverse BFGS-Aufdatierungsformel lautet:

$$\begin{aligned}
B_+^{BFGS} &:= B + \frac{(s-By)s^\mathsf{T} + s(s-By)^\mathsf{T}}{< s,\ y >_2} - \frac{< s-By,\ y >_2\ ss^\mathsf{T}}{< s,\ y >_2^2} && (4.10)\\
&= B + \left(1 + \frac{y^\mathsf{T}By}{< y,\ s >_2}\right)\frac{ss^\mathsf{T}}{< y,\ s >_2} - \frac{1}{< y,\ s >_2}\left(sy^\mathsf{T}B + Bys^\mathsf{T}\right) && (4.11)\\
&= \left(E_n - \frac{sy^\mathsf{T}}{< s,\ y >_2}\right) B \left(E_n - \frac{sy^\mathsf{T}}{< s,\ y >_2}\right)^\mathsf{T} + \frac{ss^\mathsf{T}}{< s,\ y >_2}. && (4.12)
\end{aligned}$$

Die Formel (4.10) (BFGS-INV1) wird von Geiger/Kanzow verwendet, während Großmann die Gleichung (4.11) (BFGS-INV2) angibt.[12] Polak, Kosmol bzw. Stoer/Jarre geben als inverse BFGS-Formel die Gleichung (4.12) (BFGS-INV3) an.

Ein ebenfalls bekanntes Quasi-Newton Verfahren ist das *Davidon-Fletcher-Powell* Verfahren (DFP). Die DFP-Aufdatierungsformel lautet:

$$H_+^{DFP} := H + \left(1 + \frac{s^\mathsf{T}Hs}{< y,\ s >_2}\right)\frac{yy^\mathsf{T}}{< y,\ s >_2} - \frac{1}{< y,\ s >_2}\left(ys^\mathsf{T}H + Hsy^\mathsf{T}\right). \tag{4.13}$$

Die inverse DFP-Aufdatierungsformel ist durch:

$$B_+^{DFP} := B + \frac{ss^\mathsf{T}}{< s,\ y >_2} - \frac{Byy^\mathsf{T}B}{y^\mathsf{T}By} \tag{4.14}$$

gegeben.

In Schritt 3 des Quasi-Newton-(BFGS/DFP)-Verfahrens 4.5 wird die Matrix H_{k+1} bzw. B_{k+1}

[11] Siehe z.B. Geiger/Kanzow [19].

[12] Man beachte, daß in der vorliegenden Auflage von [28] die Bezeichnungen der BFGS- und der DFP-Formeln vertauscht worden sind.

mittels der entsprechenden Aufdatierungsformel berechnet. Die Abstiegsrichtung d_{k+1} erhält man, indem man das Gleichungssystem $H_{k+1}d_{k+1} = -g_{k+1}$ löst bzw. $d_{k+1} = -B_{k+1}g_{k+1}$ berechnet.

Eine weitere Klasse von Quasi-Newton Verfahren sind die Verfahren der *Broyden-Famile*. Großmann [28] gibt als Aufdatierungsformel

$$\begin{aligned} H_+^\phi \; := \; H \; + \; \phi\left\{\frac{s^\top Hs}{<y,\,s>_2}\frac{yy^\top}{<y,\,s>_2} \; - \; \frac{1}{<y,\,s>_2}\left(ys^\top H + Hsy^\top\right)\right\} \\ - \; (1-\phi)\frac{Hss^\top H}{s^\top Hs} \; + \; \frac{yy^\top}{<s,\,y>_2} \end{aligned} \tag{4.15}$$

an, wobei $\phi \in [0,1]$ fest ist. Offensichtlich gilt $H_+^\phi = \phi H_+^{DFP} + (1-\phi)H_+^{BFGS}$, wobei H_+^{DFP} nach Gleichung (4.14) und H_+^{BFGS} nach Gleichung (4.9) gebildet worden sind.

Stoer/Jarre [55] geben eine inverse Aufdatierungsformel für die Verfahren der *Broyden-β-Klasse* an. Diese Formel lautet

$$\begin{aligned} B_+^\phi \; := \; B \; + \; \left(1+\phi\frac{y^\top By}{<y,\,s>_2}\right)\frac{ss^\top}{<y,\,s>_2} \; - \; (1-\phi)\frac{Byy^\top B}{y^\top By} \\ - \; \frac{\phi}{<y,\,s>_2}\left(sy^\top B + Bys^\top\right), \end{aligned} \tag{4.16}$$

wobei $\phi \in [0,1]$ fest ist. Offensichtlich gilt $B_+^\phi = (1-\phi)B_{DFP} + \phi B_{BFGS}$, wobei B_{DFP} nach Gleichung (4.14) und B_{BFGS} nach Gleichung (4.11) gebildet worden sind.

Im im Schritt 3 des Quasi-Newton-(Broyden)-Verfahrens 4.5 wird die Matrix H_{k+1} bzw. B_{k+1} für ein festes $\phi \in [0,1]$ mittels der Aufdatierungsformel (4.15) bzw. (4.16) berechnet. Die Abstiegsrichtung d_{k+1} erhält man, indem man das Gleichungssystem $H_{k+1}d_{k+1} = -g_{k+1}$ löst bzw. $d_{k+1} = -B_{k+1}g_{k+1}$ berechnet.

Eine Verallgemeinerung sind die Verfahren der *Oren-Luenberger-Klasse* bzw. *SSVM-Techniken.*[13] Zur Herleitung dieser Verfahren siehe man zum Beispiel Großmann [28], Geiger/Kanzow [19] und vor allem Stoer [54]. Die Aufdatierungsformel für die Oren-Luenberger-Klasse lautet nach Stoer

$$\begin{aligned} B_+^{\psi,\phi} \; := \; \psi B \; + \; \left(1+\psi\phi\frac{y^\top By}{<y,\,s>_2}\right)\frac{ss^\top}{<y,\,s>_2} \; - \; \psi(1-\phi)\frac{Byy^\top B}{y^\top By} \\ - \; \frac{\psi\phi}{<y,\,s>_2}\left(sy^\top B + Bys^\top\right). \end{aligned} \tag{4.17}$$

wobei $\psi, \phi \in \mathbb{R}$ gilt. Für $\psi = 1$ und $\phi = 0$ gilt $B_+^{1,0} = B_+^{DFP}$, während für $\psi = 1$ und $\phi = 1$ $B_+^{1,1} = B_+^{BFGS}$ gilt. Um ψ und ϕ in jedem Iterationsschritt $k = 0, 1, \ldots$ geeignet zu bestimmen, gibt Stoer das folgende Verfahren an:[14]

$$\psi_k := 1 \quad \text{und} \quad \phi_k := \begin{cases} \left(\sigma(\varepsilon-\sigma)\right)/\left(\varepsilon\tau-\sigma^2\right) & ,\text{falls } \sigma \le (2\varepsilon\tau)/(\varepsilon+\tau) \\ \sigma/(\sigma-\tau) & ,\text{sonst} \end{cases} \tag{4.18}$$

wobei

$$\varepsilon \; := \; s_k^\top B_k^{-1} s_k, \quad \sigma \; := \; <s_k,\, y_k>_2, \quad \tau \; := \; y_k^\top B_k y_k$$

[13] SSVM: self scaling variable metric

[14] Der Vorschlag stammt von Davidon [11].

gesetzt worden sind. Nun wird im Schritt 3 des Oren-Luenberger-Verfahrens 4.5 zunächst ψ_k und ϕ_k mittels der Vorschrift (4.18) bestimmt. Dann wird die Matrix B_{k+1} für ψ_k und ϕ_k mittels der Aufdatierungsformel (4.17) berechnet. Die Abstiegsrichtung d_{k+1} erhält man, indem man $d_{k+1} = -B_{k+1}g_{k+1}$ berechnet.
Zum Abschluß wird das Powell-symmetric-Broyden Verfahren (PSB) betrachtet. Das PSB-Verfahren unterscheidet sich dadurch von den anderen QN-Verfahren, daß die Matrix H_k nur symmetrisch sein muß. Es sei H symmetrisch und $s, y \in \mathbb{R}^n$ mit $s \neq 0$. Dann lautet die PSB-Formel

$$H_+^{PSB} := H + \frac{(y - Hs)\, s^\mathsf{T} + s\,(y - Hs)^\mathsf{T}}{< s,\ s >_2} - \frac{s^\mathsf{T}\,(y - Hs)}{< s,\ s >_2^2}\, ss^\mathsf{T} \tag{4.19}$$

Offensichtlich ist H_+^{PSB} symmetrisch. Falls die Bedingung $< s_k,\ s_k >_2 > 0$ im Schritt 3 des PSB-Verfahren 4.5 erfüllt ist, wird die Matrix H_{k+1} mittels der Aufdatierungsformel (4.19) berechnet. Die Abstiegsrichtung d_k erhält man, indem man das Gleichungssystem $H_{k+1}d_{k+1} = -g_{k+1}$ löst.

Newtonverfahren sind nicht implementiert worden, da es erstens nicht sicher ist, daß die Energiefunktionen aus Kapitel 3.2 und Kapitel 3.3 zweimal stetig differenzierbar sind, und da zweitens aus den Lemmatas 3.2.3 b), 3.3.4 b) und den entsprechenden Lemmatas aus Kapitel 2.4.2 folgt, daß die partiellen Ableitungen der internen Rotationsparameter nur sehr aufwendig zu bestimmen sind.[15] Die zweiten partiellen Ableitungen müßten mindestens ebenso aufwendig bestimmt werden. Und es ist nicht sicher, daß die entsprechende Hesse-Matrix in jedem Iterationsschritt invertierbar bzw. sogar positiv definit ist. Aus diesen Gründen sind Verfahren, die auf der Hesse-Matrix basieren, nicht verwendet worden.

Die hier vorgestellten und implementierten Minimierungsverfahren unterscheiden sich im Speicherbedarf und in der Anzahl der Rechenoperationen, um die Abstiegsrichtung d_k und gegenbenenfalls den Parameter γ_{k-1} bzw. die Matrix H_k zu berechnen. In der Tabelle 4.1 sind die Kosten für die einzelnen Verfahren aufgelistet. Dabei wird der Speicherplatz sowie die Anzahl der Rechenoperationen angegeben, die gebraucht werden, um aus dem Gradienten $\nabla f(x_k)$ die Abstiegsrichtung d_k zu berechnen. Zu Berechnung der Abstiegsrichtung gehört gegebenenfalls die Lösung des Gleichungssystems $H_k d_k = -g_k$. Das Gleichungssystem wird mit einem Gaußschen Eliminationsverfahren mit Spaltenpivotsuche gelöst, wobei die Matrix equilibriert wird.[16] Falls die Bedingung $< s_{k-1},\ y_{k-1} >_2 > 0$ bzw. $< s_{k-1},\ s_{k-1} >_2 > 0$ in Schritt 3 erfüllt ist, so folgt daraus, daß die aufdatierte Matrix H_k symmetrisch ist. Also braucht man nur, die obere Hälfte der Matrix H_k zu berechnen. Die Anzahl der Rechenoperationen, die notwendig sind, um x_{k+1} zu berechnen, wird nicht besonders angegeben, da sie bei allen Verfahren $2n$ (n Additionen und n Multiplikationen) beträgt, wobei n die Dimension des Optimierungsproblems ist. Die Anzahl der Rechenoperationen ist dadurch bestimmt worden, daß die Rechenoperationen in den implementierten Algorithmen gezählt worden sind.
Der Speicherbedarf bzw. die Anzahl der Rechenoperationen beträgt für das Gradientenverfahren und die CG-Verfahren jeweils $O(n)$. Dagegen beträgt der Speicherbedarf der QN-Verfahren

[15] Siehe den entsprechenden Abschitt in Kapitel 5.1.
[16] Siehe Stoer [54] und den Absatz in Kapitel 5.1.

Verfahren	Speicher	Rechenoperationen für		
		d_k	γ_{k-1}	H_k
G	$2n$	-	-	-
CG-FR	$4n$	$2n$	$4n+1$	-
CG-FRN	$4n$	$2n$	$4n+1$	-
CG-PR	$4n$	$2n$	$4n+1$	-
CG-PRN	$4n$	$2n$	$4n+1$	-
CG-DN	$4n$	$2n$	$6n+2$	-
BFGS	$5n+2n^2$	$(1/6)(4n^3+42n^2+14n-12)$	-	$5n^2-3n$
BFGS-INV1	$5n+n^2$	$2n$	-	$6n^2+n+1$
BFGS-INV2	$5n+n^2$	$2n$	-	$(1/2)(15n^2-3n)$
BFGS-INV3	$5n+2n^2$	$2n$	-	$5n^3-5n^2+2n$
DFP	$5n+2n^2$	$(1/6)(4n^3+42n^2+14n-12)$	-	$(1/2)(15n^2-3n)$
DFP-INV	$5n+n^2$	$2n$	-	$5n^2+n$
BROYDEN	$5n+2n^2$	$(1/6)(4n^3+42n^2+14n-12)$	-	$11n^2-5n$
BROYDEN-INV	$5n+n^2$	$2n$	-	$11n^2-5n$
OREN-LUEN	$6n+2n^2$	$2n$	-	$(1/6)(4n^3+44n^2+18n-10)$
PSB	$5n+2n^2$	$(1/6)(4n^3+42n^2+14n-12)$	-	$7n^2-n$

Tabelle 4.1: Kostentabelle für die Minimierungsverfahren

$O(n^2)$. Die einzelnen QN-Verfahren unterscheiden sich in der Anzahl der Rechenoperationen. Die kostengünstigsten Verfahren sind das BFGS-INV1-, BFGS-INV2-, DFP-INV- und BROYDEN-INV-Verfahren mit jeweils $O(n^2)$ Rechenoperationen. Dies sind Verfahren mit inverser Aufdatierungsformel. Das Verfahren BFGS-INV3 braucht wegen der Aufdatierungsformel (4.12) $O(n^3)$ Rechenoperationen. Auch die Verfahren, die das Gleichungssystem $H_k d_k = -g_k$ lösen müssen, brauchen $O(n^3)$ Rechenoperationen.[17] Das Oren-Luenberger-Verfahren muß in jedem Iterationsschritt das Gleichungssystem $B_k h_h = s_k$ lösen, um $\varepsilon = s_k^\top h_k = s_k^\top B_k^{-1} s_k$ in Gleichung (4.18) zu berechnen. Daher beträgt die Anzahl der Rechenoperationen ebenfalls $O(n^3)$. Am Ende des Kapitels wird ein heuristisches Verfahren vorgestellt, um das Minimierungsverfahren zu bestimmen, das am besten geeignet scheint, die Energiefunktionen aus Kapitel 3.2 zu optimieren.

Nun werden die Schrittweitenverfahren dargestellt, die implementiert worden sind, um die Schrittweite $\alpha_k > 0$ zu bestimmen.

Die Schrittweitenverfahren:

Im folgenden werden 5 Schrittweitenverfahren durch ihre Bedingungen, die eine Schrittweite $\alpha > 0$ zu erfüllen hat, vorgestellt. Dazu seien im folgenden $x, d \in \mathbb{R}^n$ mit $< \nabla f(x),\ d >_2 < 0$ und $\beta \in (0,1)$ fest, wobei $< \cdot,\ \cdot >_2$ das euklidische Skalarprodukt ist. Aus Lemma 4.2.1 folgt, daß d eine Abstiegsrichtung ist. f sei die Energiefunktion.
Polak stellt in seinem Buch [46] das Armijo-Gradienten-Verfahren (AG) vor. Das Armijo-

[17] Siehe Kapitel 5.1.

Gradienten-Verfahren entspricht dem Algorithmus 4.3, wobei die Schrittweite $\alpha := \beta^m > 0$, mit $m \in \mathbb{Z}$, so gesetzt wird, daß die beiden Bedingungen

$$\begin{aligned} f(x+\beta^m d) \; - \; f(x) \; &\leq \; \beta^m \lambda < d, \; \nabla f(x) >_2 \\ f(x+\beta^{m-1} d) \; - \; f(x) \; &> \; \beta^{m-1} \lambda < d, \; \nabla f(x) >_2 \end{aligned}$$

erfüllt sind. Dabei ist $\lambda \in (0,1)$ fest gegeben. Das Verfahren, mit dem die Schrittweite α bestimmt wird, wurde dann als Schrittweitenverfahren implementiert und mit AGS bezeichnet. Wie man aus den Bedingungen sehen kann, sind pro Iterationsschritt des AGS-Verfahrens zwei Funktionsauswertungen notwendig. Jedoch kommt man bei geeigneter Implementierung mit einer Funktionsauswertung pro Iterationsschritt aus. Nur im jeweiligen ersten Iterationsschritt des AGS-Verfahrens sind zwei Funktionsauswertungen notwendig. Pro Iterationsschritt beträgt die Anzahl der Rechenoperationen $O(n)$.

Die Armijoregel (AR) setzt die Schrittweite $\alpha := \beta^m > 0$, wobei

$$m := \min_{k \in \mathbb{N}_0} \{ f(x) \; - \; f(x+\beta^k d) \; \geq \; -\lambda \beta^k < \nabla f(x), \; d >_2 \}$$

gilt. Dabei ist $\lambda \in (0, 0.5)$ fest gegeben. Zu den Eigenschaften der Armijoregel siehe zum Beispiel Geiger/Kanzow [19] oder Kosmol [37]. Großmann [28] verwendet die Armijoregel mit $\beta = 0.5$. Eine Variante ist die Armijoregel mit Aufweitung (ARA), die von Kosmol betrachtet wird. Dabei wird die Schrittweite $\alpha := \beta^m > 0$ mit

$$m := \min_{k \in \mathbb{Z}} \{ f(x) \; - \; f(x+\beta^k d) \; \geq \; -\lambda \beta^k < \nabla f(x), \; d >_2 \}$$

gesetzt. Das AR- und das ARA-Verfahren benötigen pro Iterationsschritt eine Funktionsauswertung und $O(n)$ Rechenoperationen.

Die Goldsteinregel (G) setzt die Schrittweite $\alpha := \beta^m > 0$, mit $m \in \mathbb{Z}$, so, daß die beiden Bedingungen

$$\begin{aligned} f(x+\beta^m d) \; &< \; f(x) \; + \; \beta^m \lambda < d, \; \nabla f(x) >_2 \\ f(x+\beta^m d) \; &\geq \; f(x) \; + \; \beta^m (1-\lambda) < d, \; \nabla f(x) >_2 \end{aligned}$$

erfüllt sind. Dabei ist $\lambda \in (0, 0.5)$ fest gegeben. Zu den Eigenschaften der Goldsteinregel siehe Geiger/Kanzow [19] oder Kosmol [37]. Das G-Verfahren benötigen pro Iterationsschritt eine Funktionsauswertung und $O(n)$ Rechenoperationen.

Zum Schluß wird die Powell-Wolfe-Regel (bzw. Wolfe-Powell-Regel) vorgestellt. Die PW-Regel setzt die Schrittweite $\alpha := \beta^m > 0$, mit $m \in \mathbb{Z}$, so, daß die beiden Bedingungen

$$\begin{aligned} f(x+\beta^m d) \; &\leq \; f(x) \; + \; \beta^m \lambda < \nabla f(x), \; d >_2 \\ < \nabla f(x+\beta^m d), \; d >_2 \; &\geq \; \rho < \nabla f(x), \; d >_2 \end{aligned}$$

erfüllt sind. Dabei sind $\lambda \in (0, 0.5)$ und $\rho \in (\lambda, 1)$ fest gegeben. Zu den Eigenschaften der PW-Regel siehe Geiger/Kanzow [19], Kosmol [37] oder Cromme [10]. Das Verfahren ist zwar implementiert worden, aber da pro Iteration neben einer Funktionsauswertung auch der Gradient $\nabla f(x + \beta^m d)$ berechnet werden muß, ist dieses Verfahren bei den Testbeispielen nicht

weiter beachtet worden.
Die obigen Schrittweitenstrategien führen immer zu einem echten Abstieg, d.h. es gilt stets $f(x + \alpha d) < f(x)$, wie aus den entsprechenden Bedindungen folgt. Die Verfahren sind so implementiert worden, daß sie maximal 100 Iterationen ausführen, um die Schrittweite $\alpha > 0$ zu bestimmen. Kann innerhalb von 100 Iterationen kein geeignetes $\alpha > 0$ bestimmt werden, so bricht der Minimierungsalgorithmus ab.[18]

Ein heuristisches Testverfahren

Um nun den Minimierungsalgorithmus 4.2 aus Kapitel 4.1 zu konkretisieren, wählt man eines von jenen Minimierungsverfahren, die oben vorgestellt worden sind, und verwendet die entsprechende Vorschrift in Schritt 2 des Algorithmus 4.2, um die Abstiegsrichtung d_k zu berechnen. Um die Schrittweite $\alpha_k > 0$ zu bestimmen, wählt man eine der obigen Schrittweitenstrategien und wendet das entsprechende Verfahren in Schritt 3 des Algorithmus an. Es stellen sich folgende Probleme: welches Minimierungsverfahren und Schrittweitenverfahren sind am besten geeignet? Und welche Parameter sollte man für die entsprechende Schrittweitenstrategie wählen? Um diese Probleme zu lösen, wurde das folgende heuristische Testverfahren verwendet. In Anhang A werden die 77 Moleküle des FlexS-77 Datensatzes ([42]) beschrieben, die als Beispiele in den Rechnungen dieser Arbeit verwendet worden sind. Am Ende des Anhangs werden 10 Transformationen beschrieben. Wendet man diese Transformationen auf die gegebene Kristallstruktur eines jeden Moleküls an, so erhält man pro Molekül 10 verschiedene Konformationen, insgesamt also 770 Konformationen. Das Testverfahren besteht nun darin, die strukturelle Identität der Moleküle, wie sie durch die Kristallstruktur gegeben ist, wiederzufinden, d.h. die 10 Transformationen aus Anhang A sollten rückgängig gemacht werden. Dazu bestimmt man jeweils das optimale Alignment der 10 transformierten Molekülkonformationen bzgl. der entsprechenden Kristallstruktur mit Hilfe der Energiefunktion aus Definition 3.2.2. Die Energiefunktionen sind dahingehend modifiziert worden, daß die jeweilige Bewertungsfunktion für die Torsionswinkel[19] deaktiviert worden ist. Der zu optimierende Transformationsparameter der Energiefunktionen besteht aus dem Translationsvektor, dem Quaternionenvektor, dem globalen Rotationszentrum und den internen Rotationsparametern.[20]

Um nun die insgesamt 770 entsprechend modifizierten Energiefunktionen zu minimieren, wurde jeweils das Gradientenverfahren, die fünf CG-Verfahren, das BFGS-Verfahren, das inverse BFGS-Verfahren mit der Aufdatierungsformel (4.10), das inverse DFP-Verfahren, das inverse Broyden-β-Verfahren mit den Parametern $\phi = 0.25,\ 0.5,\ 0.75$ und das Oren-Luenberger-Verfahren verwendet. Als Schrittweitenstrategie wurde jeweils das AGS-, AR-, ARA- sowie das G-Verfahren verwendet. Bei den ersten Rechnungen hat sich herausgestellt, daß der Schrittweitenparameter λ keinen Einfluß auf die Ergebnisse und nur einen sehr minimalen Einfluß auf die Rechenzeit hat. Daher wurde für das AGS-Verfahren $\lambda = 0.5$ und für das AR-,ARA- und das G-Verfahren $\lambda = 0.3$ gesetzt. Für den Schrittweitenparameter β wurden jeweils die Werte $0.05,\ 0.1,\ 0.15$ bis 0.95 verwendet. Insgesamt haben damit 13 x 4 Kombinationen von Mini-

[18] Siehe hierzu den Abschnitt über das Minimierungsmodul in Kapitel 4.1.
[19] Siehe Definition 2.7.8.
[20] Siehe die Definitionen 2.4.27 und 3.2.2.

mierungsverfahren sowie von Schrittweitenstrategien mit jeweils 19 Parametern zur Verfügung gestanden. Diese 988 möglichen Kombinationen von Verfahren und Parametern sind dann jeweils auf die 770 Energiefunktionen angewendet worden. Dabei sind alle Rechnungen sind auf Computern vom Typ SUN-Ultra-30 Workstation, die mit einem Einzelprozessor (296-MHz) und 128 MB Hauptspeicher ausgestattet sind, ausgeführt worden. Um nun die beste Kombination auszuwählen, wurde für jede mögliche Kombination der Durchschnitt über alle 770 Ergebnisse gebildet. Für jede der 52 Kombinationen von Minimierungsverfahren und Schrittweitenregel ist der Parameter β ausgewählt worden, der im Vergleich zu den anderen 18 Werten von β das beste durchschnittliche Ergebnis erreicht hat. In der Tabelle D.1 im Anhang D (auf der Seiten 179) sind die durchschnittlichen Ergebnisse für diese 52 Kombinationen mit den entsprechenden Parametern aufgeführt worden.

Die drei besten und auch schnellsten Verfahren sind das inverse BFGS-Verfahren mit der Armijoregel mit Aufweitung (Parameter $\beta = 0.35$), das BFGS-Verfahren mit der Armijoregel (Parameter $\beta = 0.70$) sowie das inverse BFGS-Verfahren mit der Armijoregel (Parameter $\beta = 0.70$). In der Tabelle 4.2 werden die durchschnittlichen Ergebnisse der drei besten und schnellsten

Algorithmus	S_E^0	S_E^*	Schritte	Sec	Stopp	Q	N
BFGS-INV1-ARA(0.35)	0.102	0.654	76.81	31.70	0.34	2.88	1.31
BFGS-AR(0.70)	0.102	0.654	80.28	33.45	0.31	3.50	1.32
BFGS-INV1-AR(0.70)	0.102	0.654	80.43	33.54	0.31	3.54	1.37

Tabelle 4.2: Ergebnisse der besten Optimierungsverfahren

Verfahren dargestellt. Es wird der durchschnittliche Alignmentindex vor und nach den Optimierungen dargestellt, d.h. es gilt

$$\mathsf{S}_\mathsf{E}^0 := \frac{1}{770}\sum_{i=1}^{770} \mathsf{S}_{\mathcal{A}_i,\mathcal{R}_i}(u^0_{\mathcal{A}_i}) \quad \text{und} \quad \mathsf{S}_\mathsf{E}^* := \frac{1}{770}\sum_{i=1}^{770} \mathsf{S}_{\mathcal{A}_i,\mathcal{R}_i}(u^*_{\mathcal{A}_i}), \tag{4.20}$$

wobei $\mathsf{S}_{\mathcal{A}_i,\mathcal{R}_i}$ der Index aus Gleichung (3.5) in Definition 3.2.2 ist. Der Optimierungsalgorithmus verwendet für $i = 1, \dots, 770$ den Transformationsparameter $u^0_{\mathcal{A}_i}$ als Startwert und liefert $u^*_{\mathcal{A}_i}$ als Ergebnis.[21] Weiter wird die durchschnittliche Anzahl der Iterationen und die durchschnittliche Rechenzeit in Sekunden angegeben, die die entsprechende Kombination gebraucht hat, um eine Energiefunktion zu minimieren. In der Spalte Stopp der Tabelle wird die durchschnittliche Bewertung der Stoppkriterien angegeben. Als Stoppkriterien für die Minimierungsverfahren sind das Gradientenkriterium mit $\varepsilon := 10^{-7}$ und das Schrittkriterium mit $\varepsilon_S := eps^{0.9}$ verwendet worden, wobei *eps* die Maschinengenauigkeit des Computers ist. Es sind jedoch nur maximal 100 (=MAXSTEP) Iterationen ausgeführt worden.[22] Die Stoppkriterien sind wie folgt bewertet worden: Hat das Gradientenkriterium das Minimierungsverfahren beendet, so ist dies mit 1.0 bewertet worden. Ist dagegen das Verfahren durch das Schrittkriterium gestoppt worden, so ist dies mit 0.5 bewertet worden. Wenn die maximale Anzahl der Iterationen erreicht worden ist, ohne daß das Gradienten- bzw. das Schrittkriterium das Verfahren beendet hat, so ist

[21] Siehe den Abschnitt über das Minimierungsmodul in Kapitel 4.1.

[22] Siehe 21.

dies mit 0.0 bewertet worden. Konnte jedoch keine Schrittweite $\alpha_k > 0$ bestimmt werden, um $x_{k+1} = x_k + \alpha_k d_k$ zu berechnen, so ist dies mit -0.5 bewertet worden. In der Spalte Q ist die durchschnittliche Anzahl der Iterationen angegeben, die die Schrittweitenregel brauchte, um ein $\alpha_k > 0$ zu bestimmen. In der Spalte N der Tabelle 4.2 wird die durchschnittliche Anzahl der Neustarts angegeben, die bei der Aufdatierung der Matrix H_k zu H_{k+1} ausgeführt worden sind.

In Abbildung 4.1 werden für die BFGS-INV1/ARA(0.35)-, die BFGS/AR(0.7)-, die BFGS-INV1/AR(0.7)- und für die CG-PR/G(0.4)-Kombination (als die schlechteste Kombination aus Tabelle D.1) die durchschnittlichen Ergebnisse der Optimierungen dargestellt. Für jedes der 77 Moleküle wird der durchschnittliche Alignmentindex vor und nach den Optimierungen angegeben, wobei die Moleküle durch ihre Identifikationsnummern dargestellt werden. Dabei ist der durchschnittliche Alignmentindex in Gleichung (D.1) auf Seite 177 definiert worden. Die Identifikationsnummer der Moleküle richtet sich nach der Anzahl der Bindungen, an denen der Torsionswinkel variiert werden kann, und somit nach der Anzahl der internen Rotationsparameter. Die Nummer der einzelnen Moleküle kann man den Beschreibungen aus Anhang A entnehmen. In Abbildung 4.2 wird dagegen die durchschnittliche Rechenzeit in Sekunden angegeben, die die jeweilige Kombination gebraucht hat, um jeweils eine der zehn Energiefunktionen pro Molekül zu minimieren. Wie man aus der Abbildung 4.1 sehen kann, unterscheiden sich die BFGS/AR(0.7)- und die BFGS-INV1/AR(0.7)-Kombination in ihren Ergebnissen nicht besonders. Da die beiden Kombinationen theoretisch ein und daselbe Verfahren darstellen, konnte man dies im Grunde erwarten. Weiter kann man sehen, daß die Ergebnisse der Optimierungen sehr problemabhängig sind. Die BFGS-Verfahren sind aber im Gegensatz zum CG-Verfahren in der Lage gewesen, die durchschnittlichen Alignmentindize erheblich zu verbessern. Die drei besten Verfahren unterscheiden sich kaum in der Rechenzeit, wie man aus der Abbildung 4.2 sehen kann. Die Dimension der Optimierungsaufgaben nimmt mit der Identifikationsnummer zu. Dies korrespondiert mit Abbildung 4.2. Bis Molekül mit der PDB-Nummer 1ivq und mit der Identifikationsnummer 65 beträgt die durchschnittliche Rechenzeit maximal 31.7 sec.[23] Das Optimierungsproblem für dieses Molekül hat die Dimension 35 (10 globale Transformationsparameter und 15 interne Rotationsparamter). Ab diesem Molekül nimmt die durchschnittliche Rechenzeit deutlich zu. Die einzige Außnahme ist das Molekül mit der PDB-Nummer 5hvp.[24] Das Molekül hat die Identifikationsnummer 39 und die Optimierungsaufgabe für diese Molekül hat die Dimension 20. Die durchschnittliche Rechenzeit für diese Molekül beträgt 100 sec.

Als beste und schnellste Kombination von Minimierungsverfahren und Schrittweitenstrategie erscheint das inverse BFGS-Verfahren mit der Aufdatierungsformel (4.10) und der Armijoregel mit Aufweitung, wobei der Parameter $\beta = 0.35$ verwendet wird. Dieses Verfahren konnte den Alignmentindex von durchschnittlich 0.1 auf 0.654 verbessern. Um eine Energiefunktion zu minimieren, hat das inverse BFGS-Verfahren durchschnittlich 31.70 sec bzw. 76.8 Iterationen gebraucht. Die Armijoregel mit Aufweitung hat durchschnittlich 2.9 Iterationen gebraucht, um die Schrittweite $\alpha_k > 0$ zu bestimmen. Das Verfahren der Aufdatierung ist 1.3-mal neu ge-

[23] Protein Data Bank: [5], [49]. Das Molekül 1ivq gehört zur Gruppe der HIV-Inhibitoren (siehe Anhang A).
[24] 5hvp gehört ebenfalls zu den HIV-Inhibitoren.

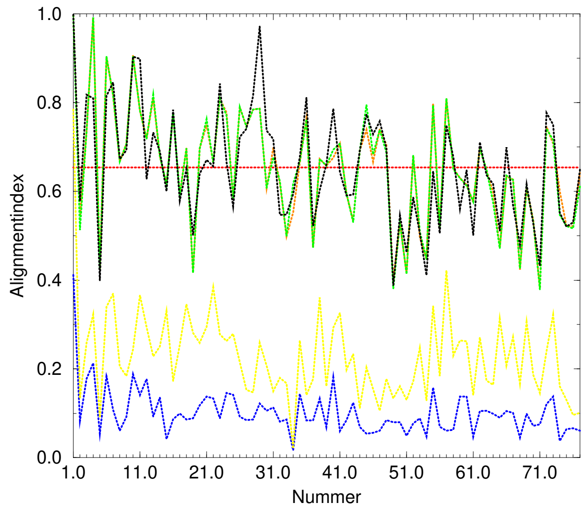


Abbildung 4.1: Durchschnittliche Alignmentindizes für die 77 Moleküle: Roten Punkte: optimaler Durchschitt von 0.654, blauen Punkte: durchschnittliche Alignmentindizes vor den Optimierungen. Durchschnittliche Alignmentindizes nach den Optimierungen: BFGS-INV1/ARA(0.35)-Kombination (schwarz), BFGS/AR(0.7)-Kombination (grün), BFGS-INV1/AR(0.7)-Kombination (orange) und die CG-PR/G(0.4)-Kombination (gelb).

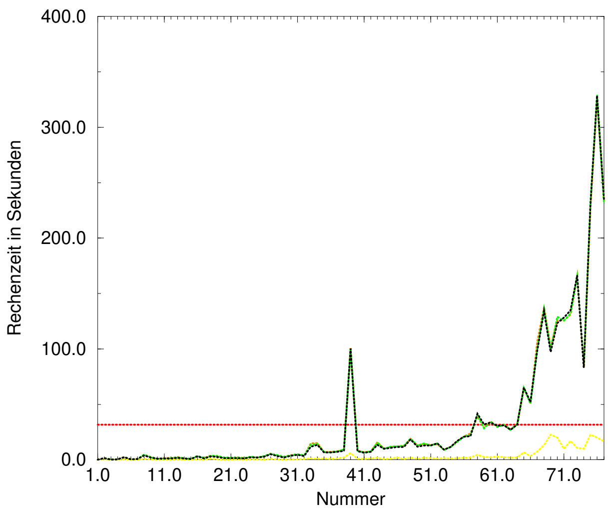


Abbildung 4.2: Durchschnittliche Rechenzeit für die 77 Moleküle: Roten Punkte: optimaler Durchschitt von 31.7 sec. Die durchschnittliche Rechenzeit: BFGS-INV1/ARA(0.35)-Kombination (schwarz), BFGS/AR(0.7)-Kombination (grün), BFGS-INV1/AR(0.7)-Kombination (orange) und die CG-PR/G(0.4)-Kombination (gelb).

startet worden. In durchschnittlich 31.7 sec. kann diese Kombination Optimierungsprobleme lösen, die maximal 35 Freiheitsgrade haben. Höher dimensionale Optimierungsaufgaben benötigen deutlich mehr Rechenzeit. D.h. je flexibler ein Molekül ist, um so mehr Rechenzeit wird die Optimierung der Molekülkonformation benötigen. Die durchschnittlichen Ergebnisse für diese Kombination werden in der Tabelle D.2 dargestellt. Dieses Verfahren wird nun bei allen Rechnungen in Kapitel 5 und 6 angewendet.

Bemerkung: Aufgabe im heuristischen Testverfahren ist es gewesen, die strukturelle Identität wiederzufinden. Somit ist das globale Optimum des Alignmentindizes 1.0. Aber aus der Abbildung 4.1 und der Tabelle D.2 kann man sehen, daß diese Optimum in einem großen Teil der Beispiele nicht erreicht worden ist. Die implementierten Minimierungsverfahren sind lokale Verfahren, d.h. ausgehend vom Startwert (hier: der Startkonformation des Moleküls) wird das nächste lokale Minimum der Energiefunktionen bestimmt. D.h. liegt die Startkonformation „ungünstig“ (man siehe hierzu das Abbild A.1 in Anhang A (auf Seite 165)), so sind die Minimierungsverfahren nicht in der Lage, Transformationsparameter zu bestimmen, damit daß das globale Minimum erreicht wird.

4.3 Spezielle Optimierungstechniken

In diesem Kapitel werden spezielle Verfahren entwickelt und vorgestellt, um die Energiefunktionen aus Kapitel 3.2 und 3.3 effizienter zu optimieren. Zuerst wird ein Gauß-Seidel-artiges Verfahren entwickelt, um die Energiefunktion aus Definition 3.3.3 effizienter zu minimieren. Danach wird ein Verfahren entwickelt, das mit Hilfe der Hierarchie der internen Rotationswinkel[25] die Energiefunktion aus Definition 3.2.2 effizienter optimieren soll. Das Verfahren wird daher Hierarchie-Verfahren genannt.

Das Gauß-Seidel-artige Verfahren für das Modell 2

Ziel ist es, ein Verfahren zu konzipieren, daß die Energiefunktion aus Definition 3.3.3 effizienter minimiert. Die Grundzüge des Verfahrens sind in der Bemerkung zu Definition 3.3.5 schon dargelegt worden. Anhand eines Beispiels soll nun der Algorithmus hergeleitet und motiviert werden. Dazu seien $\mathcal{A}$, $\mathcal{B}$ und $\mathcal{C}$ drei flexible Moleküle, die gegenseitig überlagert werden sollen. Die Startkonformationen der Moleküle seien so gegeben, wie es in Abbbildung 4.3 sym-

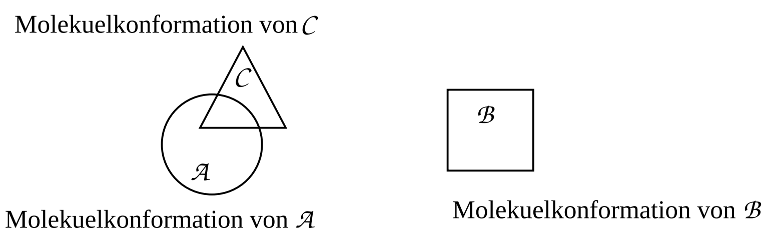


Abbildung 4.3: Schematische Darstellung der Moleküle $\mathcal{A}$, $\mathcal{B}$ und $\mathcal{C}$

[25] Siehe die Bemerkung nach Korollar 2.4.17.

bolisch angedeutet wird. Um die gegenseitige Überlagerung der Moleküle zu bestimmen, wird die Energiefunktion $\mathsf{E}_{\mathcal{A},\mathcal{B},\mathcal{C}}$, die entsprechend Definition 3.3.3 aufgestellt worden ist, minimiert. Die Parameter der Energiefunktion sind die Transformationsparameter $u_{\mathcal{A}}$, $u_{\mathcal{B}}$ und $u_{\mathcal{C}}$. Weiter seien $u^0_{\mathcal{A}}$, $u^0_{\mathcal{B}}$ und $u^0_{\mathcal{C}}$ Startwerte des Optimierungsverfahrens, wie sie im Minimierungsmodul in Kapitel 4.1 beschrieben worden sind. Die Konformationen der Moleküle und damit die gegenseitige Überlagerung hängen von den entsprechenden Transformationsparametern ab, die der Minimierungsalgorithmus 4.2 als Ergebnis liefert. Der Minimierungsalgorithmus bestimmt für $k = 0, 1, \ldots$ eine Abstiegsrichtung d_k und eine Schrittweite α_k und setzt dann

$$\begin{pmatrix} u^{k+1}_{\mathcal{A}} \\ u^{k+1}_{\mathcal{B}} \\ u^{k+1}_{\mathcal{C}} \end{pmatrix} = \begin{pmatrix} u^{k}_{\mathcal{A}} \\ u^{k}_{\mathcal{B}} \\ u^{k}_{\mathcal{C}} \end{pmatrix} + \alpha_k \begin{pmatrix} d^{\mathcal{A}}_k \\ d^{\mathcal{B}}_k \\ d^{\mathcal{C}}_k \end{pmatrix}$$

wobei $d^{\mathcal{M}}_k$, $\mathcal{M}=\mathcal{A}$,$\mathcal{B}$,$\mathcal{C}$, die entsprechenden Komponenten von d_k sind. Daraus folgt, daß in jeder Iteration die Transformationsparameter aller drei Moleküle gleichzeitig berechnet werden. Insbesondere hängt die Bestimmung des Transformationsparameters $u^{\mathcal{M}}_{k+1}$ für $\mathcal{M} = \mathcal{A}, \mathcal{B}, \mathcal{C}$, stets auch von den Transformationsparametern der beiden anderen Moleküle ab, denn die Schrittweite α_k wird mittels aller Komponenten der Vektoren $(u^k_{\mathcal{A}}, u^k_{\mathcal{B}}, u^k_{\mathcal{C}})$ und d_k bestimmt und nicht nur mittels einzelner Komponenten.[26]

Betrachtet man jedoch die Abbildung 4.3, so erscheint es vorteilhafter, zunächst die Molekülkonformation von $\mathcal{B}$ zu transformieren. Denn dadurch wird nicht nur die Überlagerung von $\mathcal{B}$ bzgl. der Moleküle $\mathcal{A}$ und $\mathcal{C}$ verbessert, sondern auch die gegenseitige Überlagerung der Moleküle. Zweitens ist die Bestimmung des Transformationsparameters $u_{\mathcal{B}}$ unabhängig von den anderen beiden Transformationsparametern. Die Konformation von $\mathcal{B}$ wird transformiert, indem man die Überlagerung von $\mathcal{B}$ bzgl. der Moleküle $\mathcal{A}$ und $\mathcal{C}$ optimiert. Dazu wird die folgende Energiefunktion verwendet:

$$\begin{aligned} \mathsf{E}_{\mathcal{B}}[u^0_{\mathcal{A}}, u^0_{\mathcal{C}}](u_{\mathcal{B}}) :=& \frac{\lambda_A}{3} \Big\{ \mathsf{A}^{\mathsf{H}}_{\mathcal{B},\mathcal{A}} \Big(\left(Q^p_{\mathcal{B}}[u_{\mathcal{B}}](x)\right)_{p \in \mathcal{I}^{\mathcal{B},\mathcal{A}}_{PE}},\ \left(Q^p_{\mathcal{A}}[u^0_{\mathcal{A}}](x)\right)_{p \in \mathcal{I}^{\mathcal{B},\mathcal{A}}_{PE}} \Big) \\ & + \mathsf{A}^{\mathsf{H}}_{\mathcal{B},\mathcal{C}} \Big(\left(Q^p_{\mathcal{B}}[u_{\mathcal{B}}](x)\right)_{p \in \mathcal{I}^{\mathcal{B},\mathcal{C}}_{PE}},\ \left(Q^p_{\mathcal{C}}[u^0_{\mathcal{C}}](x)\right)_{p \in \mathcal{I}^{\mathcal{B},\mathcal{C}}_{PE}} \Big) \Big\} \\ & + \frac{1}{3} \Big\{ P^F_{n_a[\mathcal{B}]} \Big(\lambda_P, (\mathcal{R}_{\mathcal{B}}(\varphi_{\mathcal{B}}, a^{\mathcal{B}}_1), \ldots, \mathcal{R}_{\mathcal{B}}(\varphi_{\mathcal{B}}, a^{\mathcal{B}}_{n_a[\mathcal{B}]}), \Sigma_{\mathcal{B}} \Big) \\ & + \lambda_Q \left(\|q_{\mathcal{B}}\|_2 - 1\right)^2 + \lambda_T\, s^{\mathcal{B}}_{tor} \left(\Theta^{tor}(\mathsf{C}_{\mathcal{B}}), \varphi_{\mathcal{B}}\right) \Big\}. \end{aligned} \tag{4.21}$$

Die Energiefunktion $\mathsf{E}_{\mathcal{B}}[u^0_{\mathcal{A}}, u^0_{\mathcal{C}}]$ entspricht der Energiefunktion aus Definition 3.3.5. Diese wird nun mit dem Algorithmus 4.2 minimiert. Dabei wird als Startwert $u^0_{\mathcal{B}}$ verwendet. Das Ergebnis der Minimierung sei $u^*_{\mathcal{B}} =: u^1_{\mathcal{B}}$.

Nun wird die Molekülkonformation von $\mathcal{A}$ transformiert, indem man wie oben die Überlagerung von $\mathcal{A}$ bzgl. der Moleküle $\mathcal{B}$ und $\mathcal{C}$ optimiert. wobei die transformierte Konformation des Moleküls $\mathcal{B}$ verwendet wird. Daher wird die Energiefunktion $\mathsf{E}_{\mathcal{A}}[u^1_{\mathcal{B}}, u^0_{\mathcal{C}}](u_{\mathcal{A}})$ minimiert, wobei $\mathsf{E}_{\mathcal{A}}$ analog zu $\mathsf{E}_{\mathcal{B}}$ in Gleichung (4.21) definiert worden ist. Das Ergebnis der Minimierung sei $u^*_{\mathcal{A}} =: u^1_{\mathcal{A}}$, wobei $u^0_{\mathcal{A}}$ als Startwert verwendet worden ist.

Zum Abschluß wird die Überlagerung von $\mathcal{C}$ bzgl. der Moleküle $\mathcal{A}$ und $\mathcal{B}$ optimiert, wobei die

[26] Siehe den Absatz über Schrittweitenverfahren in Kapitel 4.2.

durch $u^1_{\mathcal{B}}$ und $u^1_{\mathcal{A}}$ transformierten Molekülkonformationen verwendet werden. Dazu wird die Energiefunktion $\mathsf{E}_{\mathcal{C}}[u^1_{\mathcal{A}}, u^1_{\mathcal{B}}](u_{\mathcal{C}})$ minimiert, die analog zu $\mathsf{E}_{\mathcal{B}}$ in Gleichung (4.21) definiert worden ist. Der Algorithmus 4.2, der mit dem Startwert $u^0_{\mathcal{C}}$ gestartet worden ist, liefert als Ergebnis $u^*_{\mathcal{C}} =: u^1_{\mathcal{C}}$.

Indem die drei Energiefunktionen $\mathsf{E}_{\mathcal{A}}$, $\mathsf{E}_{\mathcal{B}}$ und $\mathsf{E}_{\mathcal{C}}$ hintereinander minimiert worden sind, ist auch die Energiefunktion $\mathsf{E}_{\mathcal{A},\mathcal{B},\mathcal{C}}$ sukzessive minimiert worden. Nach Lemma 3.3.6 d) gilt:

$$\begin{aligned} \mathsf{E}_{\mathcal{A},\mathcal{B},\mathcal{C}}(u_{\mathcal{A}}, u_{\mathcal{B}}, u_{\mathcal{C}}) &= \mathsf{E}_{\mathcal{B}}[u_{\mathcal{A}}, u_{\mathcal{C}}](u_{\mathcal{B}}) + C(u_{\mathcal{A}}, u_{\mathcal{C}}) = \mathsf{E}_{\mathcal{A}}[u_{\mathcal{B}}, u_{\mathcal{C}}](u_{\mathcal{A}}) + C(u_{\mathcal{B}}, u_{\mathcal{C}}) \\ &= \mathsf{E}_{\mathcal{C}}[u_{\mathcal{A}}, u_{\mathcal{B}}](u_{\mathcal{C}}) + C(u_{\mathcal{A}}, u_{\mathcal{B}}) \end{aligned} \tag{4.22}$$

für alle $(u_{\mathcal{A}}, u_{\mathcal{B}}, u_{\mathcal{C}}) \in U(n_a[\mathcal{A}]) \mathsf{x} U(n_a[\mathcal{B}]) \mathsf{x} U(n_a[\mathcal{C}])$, wobei

$$\begin{aligned} C_{\mathcal{B}}(u_{\mathcal{A}}, u_{\mathcal{C}}) :=\ & \mathsf{E}_{\mathcal{A},\mathcal{B},\mathcal{C}}(u_{\mathcal{A}}, u_{\mathcal{B}}, u_{\mathcal{C}}) - \mathsf{E}_{\mathcal{B}}[u_{\mathcal{A}}, u_{\mathcal{C}}](u_{\mathcal{B}}) \\ =\ & \frac{\lambda_A}{3}\, \mathsf{A}^{\mathsf{H}}_{\mathcal{A},\mathcal{C}} \left((Q^p_{\mathcal{A}}[u_{\mathcal{A}}](x))_{p \in \mathcal{I}^{\mathcal{A},\mathcal{C}}_{PE}},\ (Q^p_{\mathcal{C}}[u_{\mathcal{C}}](x))_{p \in \mathcal{I}^{\mathcal{A},\mathcal{C}}_{PE}} \right) \\ & + \frac{1}{3} \Big\{ P^F_{n_a[\mathcal{A}]} \Big(\lambda_P, (\mathcal{R}_{\mathcal{A}}(\varphi_{\mathcal{A}}, a^{\mathcal{A}}_1), \ldots, \mathcal{R}_{\mathcal{A}}(\varphi_{\mathcal{A}}, a^{\mathcal{A}}_{n_a[\mathcal{A}]}), \Sigma_{\mathcal{A}} \Big) \\ & \quad + \lambda_Q \left(\|q_{\mathcal{A}}\|_2 - 1\right)^2 + \lambda_T\, s^{\mathcal{A}}_{tor} \left(\Theta^{tor}(\mathsf{C}_{\mathcal{A}}), \varphi_{\mathcal{A}}\right) \\ & \quad + P^F_{n_a[\mathcal{C}]} \Big(\lambda_P, (\mathcal{R}_{\mathcal{C}}(\varphi_{\mathcal{C}}, a^{\mathcal{C}}_1), \ldots, \mathcal{R}_{\mathcal{C}}(\varphi_{\mathcal{C}}, a^{\mathcal{C}}_{n_a[\mathcal{C}]}), \Sigma_{\mathcal{C}} \Big) \\ & \quad + \lambda_Q \left(\|q_{\mathcal{C}}\|_2 - 1\right)^2 + \lambda_T\, s^{\mathcal{C}}_{tor} \left(\Theta^{tor}(\mathsf{C}_{\mathcal{C}}), \varphi_{\mathcal{C}}\right) \Big\} \end{aligned}$$

bzgl. $u_{\mathcal{B}}$ konstant ist. $C_{\mathcal{A}}(u_{\mathcal{B}}, u_{\mathcal{C}})$ und $C_{\mathcal{C}}(u_{\mathcal{A}}, u_{\mathcal{B}})$ werden analog gebildet. Für die Transformationsparameter $u^0_{\mathcal{A}}, u^1_{\mathcal{A}}$ $u^0_{\mathcal{B}}, u^1_{\mathcal{B}}$ und $u^0_{\mathcal{C}}, u^1_{\mathcal{C}}$ gilt dann mit (4.22) nach deren Definition als Startwert bzw. als Ergebnis einer Minimierung:

$$\begin{aligned} \mathsf{E}_{\mathcal{A},\mathcal{B},\mathcal{C}}(u^0_{\mathcal{A}}, u^0_{\mathcal{B}}, u^0_{\mathcal{C}}) &= \mathsf{E}_{\mathcal{B}}[u^0_{\mathcal{A}}, u^0_{\mathcal{C}}](u^0_{\mathcal{B}}) + C_{\mathcal{B}}(u^0_{\mathcal{A}}, u^0_{\mathcal{C}}) \\ &> \mathsf{E}_{\mathcal{B}}[u^0_{\mathcal{A}}, u^0_{\mathcal{C}}](u^1_{\mathcal{B}}) + C_{\mathcal{B}}(u^0_{\mathcal{A}}, u^0_{\mathcal{C}}) = \mathsf{E}_{\mathcal{A},\mathcal{B},\mathcal{C}}(u^0_{\mathcal{A}}, u^1_{\mathcal{B}}, u^0_{\mathcal{C}}) \\ &= \mathsf{E}_{\mathcal{A}}[u^1_{\mathcal{B}}, u^0_{\mathcal{C}}](u^0_{\mathcal{A}}) + C_{\mathcal{A}}(u^1_{\mathcal{B}}, u^0_{\mathcal{C}}) \\ &> \mathsf{E}_{\mathcal{A}}[u^1_{\mathcal{B}}, u^0_{\mathcal{C}}](u^1_{\mathcal{A}}) + C_{\mathcal{A}}(u^1_{\mathcal{B}}, u^0_{\mathcal{C}}) = \mathsf{E}_{\mathcal{A},\mathcal{B},\mathcal{C}}(u^1_{\mathcal{A}}, u^1_{\mathcal{B}}, u^0_{\mathcal{C}}) \\ &= \mathsf{E}_{\mathcal{C}}[u^1_{\mathcal{A}}, u^1_{\mathcal{B}}](u^0_{\mathcal{C}}) + C_{\mathcal{C}}(u^1_{\mathcal{A}}, u^1_{\mathcal{B}}) \\ &> \mathsf{E}_{\mathcal{C}}[u^1_{\mathcal{A}}, u^1_{\mathcal{B}}](u^1_{\mathcal{C}}) + C_{\mathcal{C}}(u^1_{\mathcal{A}}, u^1_{\mathcal{B}}) = \mathsf{E}_{\mathcal{A},\mathcal{B},\mathcal{C}}(u^1_{\mathcal{A}}, u^1_{\mathcal{B}}, u^1_{\mathcal{C}}). \end{aligned}$$

Indem also die drei Energiefunktionen $\mathsf{E}_{\mathcal{B}}$, $\mathsf{E}_{\mathcal{A}}$ und $\mathsf{E}_{\mathcal{C}}$ nacheinander minimiert worden sind, ist auch die Energiefunktion $\mathsf{E}_{\mathcal{A},\mathcal{B},\mathcal{C}}$ minimiert worden. Die Energiefunktion $\mathsf{E}_{\mathcal{M}}$, $\mathcal{M} = \mathcal{A},\mathcal{B},\mathcal{C}$, umfaßt gerade den Teil der Energiefunktion $\mathsf{E}_{\mathcal{A},\mathcal{B},\mathcal{C}}$, der bzgl. $u_{\mathcal{M}}$ variabel ist, während $\mathsf{E}_{\mathcal{A},\mathcal{B},\mathcal{C}} - \mathsf{E}_{\mathcal{M}}$ bzgl. $u_{\mathcal{M}}$ konstant ist.

Indem man nun diesen Vorgang iteriert, erhält man ein Minimierungsverfahren für die Energiefunktion $\mathsf{E}_{\mathcal{A},\mathcal{B},\mathcal{C}}$. Bevor man nun die nächste Iteration ausführt, sollten zuvor globale Stoppkriterien überprüft werden, wie sie im Abschnitt über das Minimierungsmodul in Kapitel 4.1 beschrieben worden sind. Das Gradientenkriterium $\|\nabla \mathsf{E}_{\mathcal{A},\mathcal{B},\mathcal{C}}(u^1_{\mathcal{A}}, u^1_{\mathcal{B}}, u^1_{\mathcal{C}})\|_\infty < \varepsilon_G$ sollte nur dann überprüft werden, wenn die Minimierung der Energiefunktion $\mathsf{E}_{\mathcal{C}}[u^1_{\mathcal{A}}, u^1_{\mathcal{B}}]$ mit dem Gradientenkriterium beendet worden ist, falls also $\|\nabla \mathsf{E}_{\mathcal{C}}[u^1_{\mathcal{A}}, u^1_{\mathcal{B}}](u^1_{\mathcal{C}})\|_\infty < \varepsilon_G$ gilt. Nach Lemma

Lemma 3.3.6 d) gilt:

$$\|\nabla \mathsf{E}_{\mathcal{C}}[u^1_{\mathcal{A}}, u^1_{\mathcal{B}}](u^1_{\mathcal{C}})\|_\infty \;=\; \|\nabla_{u_{\mathcal{C}}} \mathsf{E}_{\mathcal{A},\mathcal{B},\mathcal{C}}(u^1_{\mathcal{A}}, u^1_{\mathcal{B}}, u^1_{\mathcal{C}})\|_\infty \;\leq\; \|\nabla \mathsf{E}_{\mathcal{A},\mathcal{B},\mathcal{C}}(u^1_{\mathcal{A}}, u^1_{\mathcal{B}}, u^1_{\mathcal{C}})\|_\infty.$$

Es sollte auch die Reihenfolge, in der die Energiefunktionen $\mathsf{E}_{\mathcal{A}}$, $\mathsf{E}_{\mathcal{B}}$ und $\mathsf{E}_{\mathcal{C}}$ minimiert werden, erneut festgelegt werden. Aus Gleichung (4.22) folgt zwar, daß die Reihenfolge, in der Energiefunktionen $\mathsf{E}_{\mathcal{A}}$, $\mathsf{E}_{\mathcal{B}}$ und $\mathsf{E}_{\mathcal{C}}$ minimiert werden, keine Rolle spielt, da die Energiefunktion $\mathsf{E}_{\mathcal{A},\mathcal{B},\mathcal{C}}$ in jedem Fall minimiert wird. Damit aber das Verfahren möglichst effizient ist, sollte die Reihenfolge so gewählt sein, daß zuerst die Energiefunktion minimiert wird, die zu dem Molekül gehört, dessen Überlagerung bzgl. der beiden anderen Moleküle am „schlechtesten" ist. Darunter ist folgendes zu verstehen: Man berechnet zunächst $\mathsf{S}_{\mathcal{A}}[u^1_{\mathcal{B}}, u^1_{\mathcal{C}}](u^1_{\mathcal{A}})$, $\mathsf{S}_{\mathcal{B}}[u^1_{\mathcal{A}}, u^1_{\mathcal{C}}](u^1_{\mathcal{B}})$ und $\mathsf{S}_{\mathcal{C}}[u^1_{\mathcal{A}}, u^1_{\mathcal{B}}](u^1_{\mathcal{C}})$, wobei $\mathsf{S}_{\mathcal{A}}$, $\mathsf{S}_{\mathcal{B}}$ und $\mathsf{S}_{\mathcal{C}}$ analog zu der Summe der Alignmentindizes in Gleichung (3.10) in Definition 3.3.5 definiert worden sind. $\mathsf{S}_{\mathcal{A}}$ bewertet die Überlagerung der entsprechend transformierten Molekülkonformation von $\mathcal{A}$ bzgl. der transformierten Konformationen von $\mathcal{B}$ und $\mathcal{C}$. Entsprechendes gilt für $\mathsf{S}_{\mathcal{B}}$ und $\mathsf{S}_{\mathcal{C}}$. Nun werden $\mathsf{S}_{\mathcal{A}}[u^1_{\mathcal{B}}, u^1_{\mathcal{C}}](u^1_{\mathcal{A}})$, $\mathsf{S}_{\mathcal{B}}[u^1_{\mathcal{A}}, u^1_{\mathcal{C}}](u^1_{\mathcal{B}})$ und $\mathsf{S}_{\mathcal{C}}[u^1_{\mathcal{A}}, u^1_{\mathcal{B}}](u^1_{\mathcal{C}})$, mit dem kleinsten Wert beginnend, aufsteigend sotiert. In dieser Reihenfolge werden dann auch die entsprechenden Energiefunktionen minimiert. Es gelte z.B. $\mathsf{S}_{\mathcal{A}}(u^1_{\mathcal{A}}, u^1_{\mathcal{B}}, u^1_{\mathcal{C}}) \leq \mathsf{S}_{\mathcal{C}}(u^1_{\mathcal{A}}, u^1_{\mathcal{B}}, u^1_{\mathcal{C}}) \leq \mathsf{S}_{\mathcal{B}}(u^1_{\mathcal{A}}, u^1_{\mathcal{B}}, u^1_{\mathcal{C}})$. Dann wird zuerst die Energiefunktion $\mathsf{E}_{\mathcal{A}}[u^1_{\mathcal{B}}, u^1_{\mathcal{C}}](u_{\mathcal{A}})$ minimiert, wobei $u^1_{\mathcal{A}}$ als Startwert verwendet wird. $u^*_{\mathcal{A}} =: u^2_{\mathcal{A}}$ sei die Lösung, die der Minimierungsalgorithmus berechnet hat. Anschlieënd wird die Funktion $\mathsf{E}_{\mathcal{C}}[u^2_{\mathcal{A}}, u^1_{\mathcal{B}}](u_{\mathcal{C}})$ minimiert, wobei $u^1_{\mathcal{C}}$ als Startwert verwendet wird. Die Lösung $u^*_{\mathcal{C}} =: u^2_{\mathcal{C}}$ wird verwendet, um nun die Energiefunktion $\mathsf{E}_{\mathcal{B}}[u^2_{\mathcal{A}}, u^2_{\mathcal{C}}](u_{\mathcal{B}})$ zu minimieren. Als Startwert wird $u^1_{\mathcal{B}}$ verwendet. $u^*_{\mathcal{B}} =: u^2_{\mathcal{B}}$ sei die Lösung des Minimierungsverfahrens. Dieses Verfahren wird nun solange iteriert, bis ein globales Stoppkriterium erfüllt ist.

Mittels des obigen Beispiels ist der Aufbau des Gauß-Seidel-artigen Algorithmus hergeleitet worden. Zur Motivation läßt sich noch sagen: Angenommen die Transformationsoperatoren der Moleküle $\mathcal{A}$, $\mathcal{B}$ und $\mathcal{C}$ besitzen jeweils 30 Freiheitsgrade (z.B. 10 globale Transformationsparameter und 20 interne Rotationsparameter). Dann ist die Minimierung der Energiefunktion $\mathsf{E}_{\mathcal{A},\mathcal{B},\mathcal{C}}$ ein 90-dimensionales Optimierungsproblem. Wendet man dagegen das oben beschriebene Verfahren an, indem man die drei Energiefunktion $\mathsf{E}_{\mathcal{A}}$, $\mathsf{E}_{\mathcal{B}}$ und $\mathsf{E}_{\mathcal{C}}$ geeignet nacheinander minimiert, so muß man nur 3 30-dimensionale Optimierungsprobleme lösen. Das Prinzip des Verfahrens läßt sich also wie folgt beschreiben: anstatt ein hochdimensionales Optimierungsproblem zu lösen, werden mehrere geeignete und niedrigdimensionale Teiloptimierungsprobleme nacheinander gelöst. Die Hoffung besteht darin, daß der Gauß-Seidel-artige Algorithmus nicht nur schneller ein Ergebnis, sondern vor allem auch ein besseres Ergebnis liefert.

Nun soll der Gauß-Seidel-artigen Algorithmus allgemein formuliert werden.

Definition 4.3.1 (Gauß-Seidel-artiges Optimierungsverfahren)
Es seien $\mathcal{A}_1, \ldots, \mathcal{A}_{n_m}$ flexible Moleküle mit $n_m > 1$. Das optimale Alignment der n_m flexiblen Moleküle zueinander soll bestimmt werden. Daher lautet die Optimierungsaufgabe

$$\min_{(u_{\mathcal{A}_1}, \ldots, u_{\mathcal{A}_{n_m}}) \,\in\, U(n_m)} \mathsf{E}_{\mathcal{A}_1,\ldots,\mathcal{A}_{n_m}}(u_{n_m}),$$

wobei $\mathsf{E}_{\mathcal{A}_1,\dots,\mathcal{A}_{n_m}}$ die Energiefunktion aus Definition 3.3.3 ist. Der Parameterraum ist durch $U(n_m) := U(n_a[\mathcal{A}_1]) \times \dots \times U(n_a[\mathcal{A}_{n_m}])$ gegeben, wobei $U(n_a[\mathcal{A}_i])$ für $i = 1, \dots, n_m$ der Parameterraum der zum Molekül $\mathcal{A}_i$ gehörenden Transformationsoperatoren und $u_{\mathcal{A}_i} \in U(n_a[\mathcal{A}_i])$ der entsprechende Transformationsparameter ist. Das optimale Alignment von $\mathcal{A}_i$, $i = 1, \dots, n_m$, bzgl. der transformierten Moleküle $\mathcal{A}_1, \dots, \mathcal{A}_{i-1}, \mathcal{A}_{i+1}, \dots, \mathcal{A}_{n_m}$ wird mittels der entsprechenden Energiefunktion $\mathsf{E}_{\mathcal{A}_i}[u_{\mathcal{A}_1}, \dots, u_{\mathcal{A}_{i-1}}, u_{\mathcal{A}_{i+1}}, \dots, u_{\mathcal{A}_{n_m}}]$ aus Definition 3.3.5 bestimmt. Aus Lemma 3.3.6 d) folgt, daß für $i = 1, \dots, n_m$

$$\begin{aligned} \mathsf{E}_{\mathcal{A}_1,\dots,\mathcal{A}_{n_m}}(u_{\mathcal{A}_1}, \dots, u_{\mathcal{A}_{n_m}}) &= \mathsf{E}_{\mathcal{A}_i}[u_{\mathcal{A}_1}, \dots, u_{\mathcal{A}_{i-1}}, u_{\mathcal{A}_{i+1}}, \dots, u_{\mathcal{A}_{n_m}}](u_{\mathcal{A}_i}) \\ &+ C_i(u_{\mathcal{A}_1}, \dots, u_{\mathcal{A}_{i-1}}, u_{\mathcal{A}_{i+1}}, \dots, u_{\mathcal{A}_{n_m}}) \end{aligned} \tag{4.23}$$

für alle $(u_{\mathcal{A}_1}, \dots, u_{\mathcal{A}_{n_m}}) \in U(n_m)$ gilt, wobei

$$\begin{aligned} &C_i(u_{\mathcal{A}_1}, \dots, u_{\mathcal{A}_{i-1}}, u_{\mathcal{A}_{i+1}}, \dots, u_{\mathcal{A}_{n_m}}) := \\ &\quad \mathsf{E}_{\mathcal{A}_1,\dots,\mathcal{A}_{n_m}}(u_{\mathcal{A}_1}, \dots, u_{\mathcal{A}_{n_m}}) - \mathsf{E}_{\mathcal{A}_i}[u_{\mathcal{A}_1}, \dots, u_{\mathcal{A}_{i-1}}, u_{\mathcal{A}_{i+1}}, \dots, u_{\mathcal{A}_{n_m}}](u_{\mathcal{A}_i}) \end{aligned}$$

bzgl. dem Parameter $u_{\mathcal{A}_i}$ konstant ist. Weiter sei $\mathsf{S}_{\mathcal{A}_i}[u_{\mathcal{A}_1}, \dots, u_{\mathcal{A}_{i-1}}, u_{\mathcal{A}_{i+1}}, \dots, u_{\mathcal{A}_{n_m}}](u_{\mathcal{A}_i})$, $i = 1, \dots, n_m$, die Summe der Alignmentindizes aus Gleichung (3.10) in Definition 3.3.5. Die Summe $\mathsf{S}_i := \mathsf{S}_{\mathcal{A}_i}$ bewertet das Alignment des transformierten Moleküls $\mathcal{A}_i$ bzgl. der transformierten Moleküle $\mathcal{A}_1, \dots, \mathcal{A}_{i-1}, \mathcal{A}_{i+1}, \dots, \mathcal{A}_{n_m}$. Für $i = 1, \dots, n_m$ wird $u^0_{\mathcal{A}_i} \in U(n_a[\mathcal{A}_i])$ so gewählt, daß die zugehörigen Transformationsoperatoren $\mathcal{T}_{\mathcal{A}_i}$ die Identität sind.[27] Setzt man nun $\mathsf{E} := \mathsf{E}_{\mathcal{A}_1,\dots,\mathcal{A}_{n_m}}$ und für $i = 1, \dots, n_m$ $\mathsf{E}_i := \mathsf{E}_{\mathcal{A}_i}$ und $x^i_k := (u^k_{\mathcal{A}_1}, \dots, u^k_{\mathcal{A}_{i-1}}, u^k_{\mathcal{A}_{i+1}}, \dots, u^k_{\mathcal{A}_{n_m}})$, $k \in \mathbb{N}_0$, so lautet das Gauß-Seidel-artige Optimierungsverfahren:

Algorithmus 4.6 (Gauß-Seidel-artiger Algorithmus)

(S.0) Startwert: $x_0 := (u^0_{\mathcal{A}_1}, \dots, u^0_{\mathcal{A}_{n_m}}) \in U(n_m)$
Stoppkriterien: $\varepsilon_G > 0$, $\varepsilon_S > 0$, $\varepsilon_E > 0$ und GLOBALMAXSTEP.
Berechne $\mathsf{E}(x_0)$ und für $i = 1, \dots, n_m$ die Werte $\mathsf{S}_i[x^i_0](u^0_{\mathcal{A}_i})$.[28]

(S.1) Berechne $\nabla\mathsf{E}(x_k)$. Falls $\|\nabla\mathsf{E}(x_k)\|_\infty < \varepsilon_G$, STOP.

(S.2) Sortiere $\mathsf{S}_i[x^i_k](u^k_{\mathcal{A}_i})$, $i = 1, \dots, n_m$ so, daß

$$\mathsf{S}_{s_1}[x^{s_1}_k](u^k_{\mathcal{A}_{s_1}}) \leq \dots \leq \mathsf{S}_{s_{n_m}}[x^{s_{n_m}}_k](u^k_{\mathcal{A}_{s_{n_m}}})$$

gilt, wobei $\{s_1, \dots, s_{n_m}\}$ die sortierte Indexmenge $\{1, \dots, n_m\}$ ist.

(S.3) Für $i = 1, \dots, n_m$ löse das Optimierungsproblem

$$\min_{u_{\mathcal{A}_{s_i}} \in U(n_a[\mathcal{A}_{s_i}])} \mathsf{E}_{s_i}[u^{k+1}_{\mathcal{A}_{s_1}}, \dots, u^{k+1}_{\mathcal{A}_{s_{i-1}}}, u^k_{\mathcal{A}_{s_{i+1}}}, \dots, u^k_{\mathcal{A}_{s_{n_m}}}](u_{\mathcal{A}_{s_i}}) \tag{4.24}$$

mit dem Minimierungsalgorithmus 4.2, wobei $u^k_{\mathcal{A}_{s_i}}$ als Startwert verwendet wird. Als Stoppkriterien werden $\varepsilon_G > 0$, $\varepsilon_S > 0$, $\varepsilon_E > 0$ und MAXSTEP=17 verwendet. $u^*_{\mathcal{A}_{s_i}} =: u^{k+1}_{\mathcal{A}_{s_i}}$ sei die vom Minimierungsalgorithmus 4.2 berechnete Lösung des Optimierungsproblem (4.24).

---

[27] Siehe Lemma 2.4.28 d) und den Abschnitt über das Minimierungsmodul in Kapitel 4.1.

[28] Man beachte, daß $\mathsf{S}_i[x^i_0](u^0_{\mathcal{A}_i})$, $i = 1, \dots, n_m$, schon im Einlesemodul (siehe Kapitel 4.1) berechnet worden sind.

(S.4) Setze $x_{k+1} := (u^{k+1}_{\mathcal{A}_1}, \ldots, u^{k+1}_{\mathcal{A}_{n_m}})$. *Falls*

$$\max_{i=1,\ldots,n} \left(\frac{|x_{k+1,i} - x_{k,i}|}{\max\{x_{k+1,i},\ 1.0\}} \right) < \varepsilon_S$$

gilt, STOP.[29]

(S.5) Falls $|\mathsf{E}(x_{k+1}) - \mathsf{E}(x_k)| < \varepsilon_E$, *STOP.*

(S.6) Falls k=GLOBALMAXSTEP, STOP. Sonst berechne $\mathsf{S}_i[x^i_{k+1}](u^{k+1}_{\mathcal{A}_i})$ *für* $i = 1, \ldots, n_m$ *und setze dann* $k = k + 1$ *und gehe zu (S.1).*

Das Ziel der Optimierung ist es, die optimale Überlagerung der Moleküle zueinander zu bestimmen. Dazu wird nacheinander das jeweilige Alignment des entsprechenden Moleküls bzgl. der anderen Moleküle optimiert, wobei es reicht, das entsprechende Aligment „grob“ zu optimieren. Denn es soll nicht das Alignment eines bestimmten Moleküls bzgl. der anderen Moleküle, sondern die Überlagerung aller Moleküle zueinander optimiert werden. Daher reicht es, wenn die Energiefunktionen E_i *nur „grob“ optimiert werden. Der Minimierungsalgorithmus 4.2 führt daher maximal 17 (=MAXSTEP) Iterationen aus.*[30] *Zu den einzelnen Stoppkriterien im Algorithmus vergleiche man den Abschnitt über das Minimierungsmodul in Kapitel 4.1, wobei die Stoppkriterien in den Schritten 4, 5 und 6 optional sind.* ◇

Im folgenden Lemma werden einige Eigenschaften des Algorithmus 4.6 dargestellt.

Lemma 4.3.2
Mit den Bezeichnungen aus Definition 4.3.1 gilt:

a) *Es sei* $x_k = (u^k_{\mathcal{A}_1}, \ldots, u^k_{\mathcal{A}_{n_m}}) \in U(n_m)$. *Dann gilt:*

$$\nabla\mathsf{E}(x_k) \ = \ 0 \ \Leftrightarrow \ \nabla\mathsf{E}_i[x^i_k](u^k_{\mathcal{A}_i}) \ = \ 0, \ i = 1, \ldots, n_m.$$

Insbesondere gilt: $\|\nabla\mathsf{E}(x_k)\|_\infty < \varepsilon_G \Leftrightarrow \|\nabla\mathsf{E}_i[x^i_k](u^k_{\mathcal{A}_i})\|_\infty < \varepsilon_G$ *für* $i = 1, \ldots, n_m$.

b) *Der Algorithmus 4.6 berechnet in jedem Iterationsschritt* $k = 0, 1, \ldots$ *ein* x_{k+1} *mit der Eigenschaft* $\mathsf{E}(x_{k+1}) < \mathsf{E}(x_k)$, *es sei denn, der in Schritt 3 verwendete Minimierungsalgorithmus 4.2 bricht jedesmal im 0-ten Iterationsschritt ab, weil entweder das Gradientenkriterium erfüllt worden ist oder aber das verwendete Schrittweitenverfahren keine geeignete Schrittweite bestimmen konnte (Schritt 3 des Algorithmus 4.2).*

Beweis:

a) Es sei $x_k = (u^k_{\mathcal{A}_1}, \ldots, u^k_{\mathcal{A}_{n_m}}) \in U(n_m)$. Aus Lemma 3.3.6 d) folgt für $i = 1, \ldots, n_m$:

$$\nabla_{u_{\mathcal{A}_i}} \mathsf{E}(u^k_{\mathcal{A}_1}, \ldots, u^k_{\mathcal{A}_{n_m}}) \ = \ \nabla\mathsf{E}_i[u^k_{\mathcal{A}_1}, \ldots, u^k_{\mathcal{A}_{i-1}}, u^k_{\mathcal{A}_{i+1}}, \ldots, u^k_{\mathcal{A}_{n_m}}](u^k_{\mathcal{A}_i}). \qquad (4.25)$$

[29] n bezeichnet die Dimension des globalen Optimierungsproblems. Mit $x_{k,i}$ wird die i-te Komponente des Vektors x_k bezeichnet.

[30] Siehe auch Tabelle 4.3 (Seite 116).

Gilt nun $\nabla\mathsf{E}(x_k) = 0$, so folgt aus Gleichung (4.25) sofort $\nabla\mathsf{E}_i[x_k^i](u_{\mathcal{A}_i}^k) = 0$ für $i = 1, \ldots, n_m$. Gilt nun andererseits $\nabla\mathsf{E}_i[x_k^i](u_{\mathcal{A}_i}^k) = 0$ für $i = 1, \ldots, n_m$ so folgt aus Gleichung (4.25) für $i = 1, \ldots, n_m$ $\nabla_{u_{\mathcal{A}_i}}\mathsf{E}(x_k) = 0$ und damit die Behauptung. Aus Gleichung (4.25) folgt auch:

$$\|\nabla\mathsf{E}_i[x_k^i](u_{\mathcal{A}_i}^k)\|_\infty \;=\; \|\nabla_{u_{\mathcal{A}_i}}\mathsf{E}(x_k)\|_\infty \;\leq\; \max_{i=1,\ldots,n_m} \|\nabla_{u_{\mathcal{A}_i}}\mathsf{E}(x_k)\|_\infty \;=\; \|\nabla\mathsf{E}(x_k)\|_\infty.$$

Daraus folgt nun, daß $\|\nabla\mathsf{E}_i[x_k^i](u_{\mathcal{A}_i}^k)\|_\infty < \varepsilon_G$ für $i = 1, \ldots, n_m$ gilt, falls $\|\nabla\mathsf{E}(x_k)\|_\infty < \varepsilon_G$ erfüllt ist, und daß andererseits $\|\nabla\mathsf{E}(x_k)\|_\infty < \varepsilon_G$ gilt, falls $\|\nabla\mathsf{E}_i[x_k^i](u_{\mathcal{A}_i}^k)\|_\infty < \varepsilon_G$ für $i = 1, \ldots, n_m$ erfüllt ist.

b) Der Algorithmus 4.6 sei im k-ten Iterationsschritt und $x_k = (u_{\mathcal{A}_1}^k, \ldots, u_{\mathcal{A}_{n_m}}^k) \in U(n_m)$ sei der Transformationsvektor. Falls der Algorithmus nicht in Schritt 1 terminiert, so wird durch Schritt 2 und 3 der Transformationsvektor $x_{k+1} = (u_{\mathcal{A}_1}^{k+1}, \ldots, u_{\mathcal{A}_{n_m}}^{k+1}) \in U(n_m)$ berechnet. Für $i = 1, \ldots, n_m$ ist $u_{\mathcal{A}_i}^{k+1}$ die vom Minimierungsalgorithmus 4.2 berechnete Lösung des entsprechenden Optimierungsproblems (4.24). Daraus folgt mit Gleichung (4.23) die Ungleichungskette

$$\begin{aligned}
\mathsf{E}(x_k) &= \mathsf{E}_{s_1}[x_k^{s_1}](u_{\mathcal{A}_{s_1}}^k) + C_{s_1}(x_k^{s_1}) \\
&\geq \mathsf{E}_{s_1}[x_k^{s_1}](u_{\mathcal{A}_{s_1}}^{k+1}) + C_{s_1}(x_k^{s_1}) \;=\; \mathsf{E}(u_{\mathcal{A}_{s_1}}^{k+1}, u_{\mathcal{A}_{s_2}}^k, \ldots, u_{\mathcal{A}_{s_{n_m}}}^k) \\
&= \mathsf{E}_{s_2}[u_{\mathcal{A}_{s_1}}^{k+1}, u_{\mathcal{A}_{s_3}}^k, \ldots, u_{\mathcal{A}_{s_{n_m}}}^k](u_{\mathcal{A}_{s_2}}^k) + C_{s_2}(u_{\mathcal{A}_{s_1}}^{k+1}, u_{\mathcal{A}_{s_3}}^k, \ldots, u_{\mathcal{A}_{s_{n_m}}}^k) \\
&\geq \cdots \\
&= \mathsf{E}(u_{\mathcal{A}_{s_1}}^{k+1}, \ldots, u_{\mathcal{A}_{s_{n_m-1}}}^{k+1}, u_{\mathcal{A}_{s_{n_m}}}^k) \;=\; \mathsf{E}_{s_{n_m}}[x_{k+1}^{s_{n_m}}](u_{\mathcal{A}_{s_{n_m}}}^k) + C_{s_{n_m}}(x_{k+1}^{s_{n_m}}) \\
&\geq \mathsf{E}_{s_{n_m}}[x_{k+1}^{s_{n_m}}](u_{\mathcal{A}_{s_{n_m}}}^{k+1}) + C_{s_{n_m}}(x_{k+1}^{s_{n_m}}) \;=\; \mathsf{E}(u_{\mathcal{A}_{s_1}}^{k+1}, \ldots, u_{\mathcal{A}_{s_{n_m-1}}}^{k+1}, u_{\mathcal{A}_{s_{n_m}}}^{k+1}) \\
&= \mathsf{E}(x_{k+1}),
\end{aligned}$$

wobei $\{s_1, \ldots, s_{n_m}\}$ die durch Schritt 2 im Algorithmus 4.6 sortierte Molekülindexmenge $\{1, \ldots, n_m\}$ ist. Für x_{k+1} gilt also $\mathsf{E}(x_{k+1}) \leq \mathsf{E}(x_k)$.
Um nun $\mathsf{E}(x_{k+1}) < \mathsf{E}(x_k)$ zu zeigen, folgt eine Fallunterscheidung nach möglichen Abbruchkriterien für den Minimierungsalgorithmus 4.2:

1.) Der Algorithmus 4.2 terminiert jedesmal im 0-ten Iterationsschritt mit den Gradientenkriterium, d.h. es gilt $\|\nabla\mathsf{E}_{s_i}[x_k^{s_i}](u_{\mathcal{A}_{s_i}}^k)\|_\infty < \varepsilon_G$ für $i = 1, \ldots, n_m$. Aber aus a) folgt daraus, daß $\|\nabla\mathsf{E}(x_k)\|_\infty < \varepsilon_G$ gilt. Der Algorithmus 4.6 wäre mit dem Gradientenkriterium terminiert und hätte x_{k+1} nicht berechnen können. Also kann dieser Fall nicht eintreten.

2.) Der Algorithmus 4.2 terminiert jedesmal im 0-ten Iterationsschritt mit einem Abbruch in Schritt 3 des Algorithmus, d.h. das verwendete Schrittweitenverfahren konnte keine geeignete Schrittweite $\alpha_0 > 0$ bestimmen. Der Algorithmus liefert jedesmal den Startwert $u_{\mathcal{A}_{s_i}}^k$ als Ergebnis zurück. Damit gilt $x_{k+1} = x_k$.

3.) Der Algorithmus 4.2 terminiert jedesmal im 0-ten Iterationsschritt entweder mit dem Gradientenkriterium oder aber in Schritt 3 des Algorithmus (siehe 2.), jedoch mindestens einmal mit dem Gradientenkriterium und einmal in Schritt 3. Der Algorithmus liefert dann jedesmal den Startwert $u_{\mathcal{A}_{s_i}}^k$ als Ergebnis zurück. Damit gilt $x_{k+1} = x_k$.

4.) Tritt weder der zweite noch der dritte Fall ein, so existiert mindestens ein $i \in \{1, \ldots, n_m\}$, so daß der Minimierungsalgorithmus 4.2, der die entsprechende Energiefunktion $\mathsf{E}_{s_i}[u^{k+1}_{\mathcal{A}_{s_1}}, \ldots, u^{k+1}_{\mathcal{A}_{s_{i-1}}}, u^{k}_{\mathcal{A}_{s_{i+1}}}, \ldots, u^{k}_{\mathcal{A}_{s_{n_m}}}]$ minimiert, entweder im 0-ten Iterationsschritt mit dem Schrittkriterium bzw. Energiekriterium oder aber im l-ten Iterationsschritt mit $l > 0$ terminiert. In jedem Fall gilt:

$$\mathsf{E}_{s_i}[u^{k+1}_{\mathcal{A}_{s_1}}, \ldots, u^{k+1}_{\mathcal{A}_{s_{i-1}}}, u^{k}_{\mathcal{A}_{i+1}}, \ldots, u^{k}_{\mathcal{A}_{n_m}}](u^{k}_{\mathcal{A}_{s_i}}) > \mathsf{E}_{s_i}[u^{k+1}_{\mathcal{A}_{s_1}}, \ldots, u^{k+1}_{\mathcal{A}_{s_{i-1}}}, u^{k}_{\mathcal{A}_{i+1}}, \ldots, u^{k}_{\mathcal{A}_{n_m}}](u^{k+1}_{\mathcal{A}_{s_i}}),$$

da das im Algorithmus 4.2 verwendete Schrittweitenverfahren eine Schrittweite $\alpha_0 > 0$ so berechnet, daß $\mathsf{E}_{s_i}[\ldots](u^{k,0}_{\mathcal{A}_{s_i}}) > \mathsf{E}_{s_i}[\ldots](u^{k,0}_{\mathcal{A}_{s_i}} + \alpha_0\, d_0) \geq \mathsf{E}_{s_i}[\ldots](u^{k+1}_{\mathcal{A}_{s_i}})$ gilt, wobei d_0 die Abstiegsrichtung im 0-Iterationsschritt ist. ∎

Bemerkung:

a) Terminiert der Algorithmus 4.6 in der k-ten Iteration mit dem Gradientenkriterium (Schritt 1), d.h. gilt $\|\nabla \mathsf{E}(x_k)\|_\infty < \varepsilon_G$, so folgt aus Lemma 4.3.2 a), daß $\|\nabla \mathsf{E}_i[x^i_k](u^k_{\mathcal{A}_i})\|_\infty < \varepsilon_G$ für $i = 1, \ldots, n_m$ gilt.

b) Lemma 4.3.2 b) besagt, daß der Algorithmus 4.6 bis auf die zwei pathologischen Fälle ein Abstiegsverfahren ist. Aus dem Beweis von b) folgt, daß das Schrittkriterium in Schritt 4 des Algorithmus 4.6 stets aktiviert sein sollte, um in den beiden pathologischen Fällen (2. und 3. Fall) zu gewährleisten, daß der Algorithmus terminiert.

c) Damit der obige Algorithmus möglichst effizient ist, sollte man bei der Implementierung darauf achten, daß man in Schritt 1 den Gradienten $\nabla \mathsf{E}(x_k)$ nur dann berechnen muß, wenn $\|\nabla \mathsf{E}_{s_{n_m}}[x^{s_{n_m}}_k](u^*_{\mathcal{A}_{s_{n_m}}})\|_\infty < \varepsilon_G$ gilt.

d) Ein weiterer Vorteil des Gauß-Seidel-artigen Optimierungsverfahrens 4.6 gegenüber der direkten Optimierung der Energiefunktion $\mathsf{E}_{\mathcal{A}_1,\ldots,\mathcal{A}_{n_m}}$ liegt darin, daß ein Abbruch in Schritt 3 vom Minimierungsalgorithmus 4.2 nicht zum sofortigen Ende des Verfahrens führt. Siehe hierzu auch Lemma 4.3.2 b).

Um zu überprüfen, ob der Algorithmus 4.6 die Energiefunktion $\mathsf{E}_{\mathcal{A}_1,\ldots,\mathcal{A}_{n_m}}$ effizienter minimiert als die direkte Minimierung mit dem Algorithmus 4.2, ist das schon in Kapitel 4.2 beschriebene heuristische Verfahren mit folgenden Modifikationen angewandt worden. Es sind jeweils pro Molekül die Konformationen der 2., 4., 6., 8., und der 10. sowie die Konformationen der 1., 3., 5., 7. und 9. Transformation aus Anhang A gegenseitig überlagert worden. Um diese 154 Testbeispiele zu berechnen, sind die entsprechenden Energiefunktionen aus Definition 3.3.3 minimiert worden. Die Energiefunktionen sind dahingehend modifiziert worden, daß die jeweiligen Bewertungsfunktionen für die Torsionswinkel[31] deaktiviert worden sind. Der zu optimierende Transformationsparameter der Energiefunktionen besteht pro Molekül aus dem Translationsvektor, dem Quaternionenvektor, dem globalen Rotationszentrum und den internen Rotationsparametern.[32] Um den Algorithmus 4.2 zu konkretisieren, ist das inverse BFGS-Verfahren mit der Aufdatierungsformel (4.10) und der Armijoregel mit Aufweitung (Parameter $\beta = 0.35$)

[31] Siehe die Definition 2.7.8 und 3.3.3.

[32] Siehe die Definitionen 2.4.27 und 3.3.3.

verwendet worden. In Tabelle 4.3 sind die durchschnittlichen Ergebnisse der 154 Optimierungsverfahren dargestellt worden, wobei das Verfahren 4.6 mit MAXSTEP gleich 10, 12, 13, 14, 15, 16, 17, 18 sowie mit MAXSTEP gleich 20 gestartet worden ist. In der Tabelle 4.3 wird die durch-

Algorithmus	$\mathsf{S}^0_{\mathcal{A}_1,\dots,\mathcal{A}_5}$	$\mathsf{S}^*_{\mathcal{A}_1,\dots,\mathcal{A}_5}$	$\mathsf{E}^0_{\mathcal{A}_1,\dots,\mathcal{A}_5}$	$\mathsf{E}^*_{\mathcal{A}_1,\dots,\mathcal{A}_5}$	Schritte	Sec	Stopp
4.2	0.0529	0.3514	438.75	1.02	19.31	197.37	0.22
4.6 (10)	0.0529	0.3667	438.75	0.95	5.68	294.52	0.49
4.6 (12)	0.0529	0.3620	438.75	0.97	5.05	274.05	0.50
4.6 (13)	0.0529	0.3625	438.75	0.97	5.13	237.10	0.50
4.6 (14)	0.0529	0.3620	438.75	0.97	5.29	292.29	0.50
4.6 (15)	0.0529	0.3646	438.75	0.96	5.29	265.59	0.50
4.6 (16)	0.0529	0.3625	438.75	0.97	4.84	275.37	0.50
4.6 (17)	0.0529	0.3630	438.75	0.97	4.79	217.42	0.50
4.6 (18)	0.0529	0.3618	438.75	0.97	5.12	300.60	0.50
4.6 (20)	0.0529	0.3595	438.75	0.98	4.82	292.76	0.50

Tabelle 4.3: Ergebnisse der Algorithmen 4.2 und 4.6

schnittliche Summe der Alignmentindizes vor den Optimierungen und nach den Optimierungen dargestellt, d.h. es gilt

$$\mathsf{S}^0_{\mathcal{A}_1,\dots,\mathcal{A}_5} := \frac{1}{154}\sum_{i=1}^{154} \mathsf{S}_{\mathcal{A}_1^i,\dots,\mathcal{A}_5^i}(u^0_{\mathcal{A}_1^i},\dots,u^0_{\mathcal{A}_5^i})$$

und

$$\mathsf{S}^*_{\mathcal{A}_1,\dots,\mathcal{A}_5} := \frac{1}{154}\sum_{i=1}^{154} \mathsf{S}_{\mathcal{A}_1^i,\dots,\mathcal{A}_5^i}(u^*_{\mathcal{A}_1^i},\dots,u^*_{\mathcal{A}_5^i}),$$

wobei $\mathsf{S}_{\mathcal{A}_1^i,\dots,\mathcal{A}_5^i}$ die Summe der Alignmentindizes aus der Gleichung (3.9) in Definition 3.3.3 ist. Der durchschnittliche Funktionswert der Energiefunktionen vor und nach den Optimierungen, $\mathsf{E}^0_{\mathcal{A}_1,\dots,\mathcal{A}_5}$ und $\mathsf{E}^*_{\mathcal{A}_1,\dots,\mathcal{A}_5}$ ist analog gebildet worden. Die Optimierungsalgorithmen verwenden für $i = 1,\dots,154$ den Transformationsparameter $(u^0_{\mathcal{A}_1^i},\dots,u^0_{\mathcal{A}_5^i})$ als Startwert und liefern $(u^*_{\mathcal{A}_1^i},\dots,u^*_{\mathcal{A}_5^i})$ als Ergebnis. Weiter wird die durchschnittliche Anzahl der Iterationen und die durchschnittliche Rechenzeit in Sekunden angegeben, die das entsprechende Verfahren gebraucht hat, um eine Energiefunktion zu minimieren. Die durchschnittliche Bewertung der Stoppkriterien wird in Spalte Stopp angegeben. Als Stoppkriterien sind das Gradientenkriterium mit $\varepsilon := 10^{-6}$ und das Schrittkriterium mit $\varepsilon_S := eps^{0.9}$ verwendet worden, wobei *eps* die Maschinengenauigkeit des Computers ist. Die Optimierungsverfahren konnten jedoch nur maximal 50 (=GLOBALMAXSTEP) Iterationen ausführen. Zur Bewertung der Stoppkriterien siehe man die Erläuterungen zur Tabelle 4.2 (Seite 105). In den Abbildungen 4.4 bis 4.7 wird jeweilige Wert von $\mathsf{S}_{\mathcal{A}_1^i,\dots,\mathcal{A}_5^i}$, der die gegenseitige Überlagerung bewertet, sowie die Rechenzeit in Sekunden dargestellt, um eine Energiefunktion vom Typ $\mathsf{E}_{\mathcal{A}_1,\dots,\mathcal{A}_5}$ zu minimieren. In den Abbildungen 4.4 und 4.5 sind die Ergebnisse der gegenseitigen Überlagerung der Konformationen von Typ A (Konformationen der 2., 4., 6., 8., und der 10. Transformation aus Anhang A) für den Algorithmus 4.2 sowie für den Algorithmus 4.6 mit MAXSTEP=10 bzw. MAXSTEP=17 dargestellt

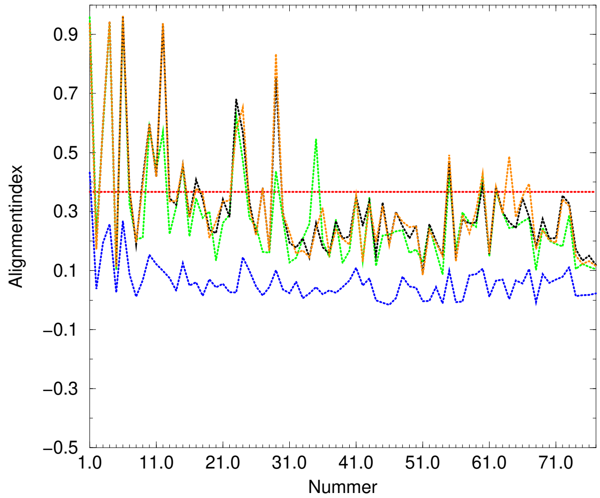


Abbildung 4.4: Ergebnisse der gegenseitigen Überlagerung von Konformationen vom Typ A: Roten Punkte: optimaler Durchschnitt von 0.367, blauen Punkte: $\mathsf{S}_{\mathcal{A}_1^i,\dots,\mathcal{A}_5^i}(u^0_{\mathcal{A}_1^i},\dots,u^0_{\mathcal{A}_5^i})$. $\mathsf{S}_{\mathcal{A}_1^i,\dots,\mathcal{A}_5^i}(u^*_{\mathcal{A}_1^i},\dots,u^*_{\mathcal{A}_5^i})$: Gauß-Seidel-artigen Verfahren 4.6 mit MAXSTEP=10 (orange), Gauß-Seidel-artigen Verfahren 4.6 mit MAXSTEP=17 (schwarz), Optimierungsalgorithmus 4.2 (grün).

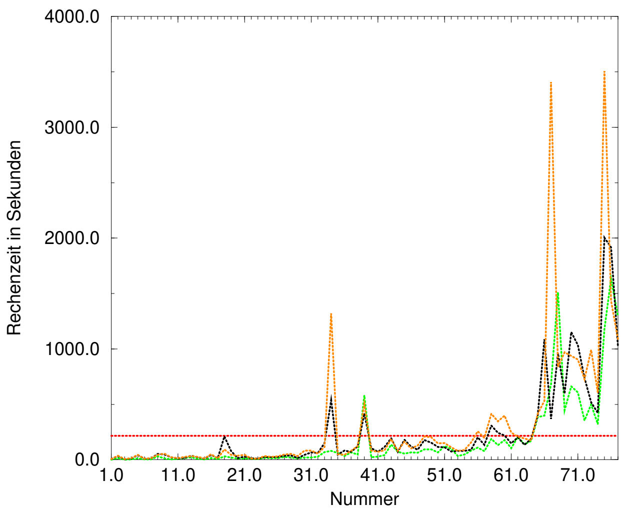


Abbildung 4.5: Rechenzeit für die gegenseitigen Überlagerungen von Konformationen vom Typ A: Roten Punkte: durchschnittliche Rechenzeit 217.44 sec. Rechenzeit der Verfahren: Gauß-Seidel-artigen Verfahren 4.6 mit MAXSTEP=10 (orange), Gauß-Seidel-artigen Verfahren 4.6 mit MAXSTEP=17 (schwarz), Optimierungsalgorithmus 4.2 (grün).

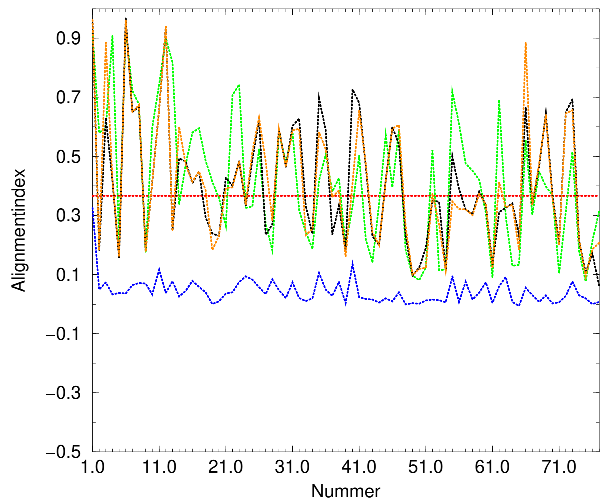


Abbildung 4.6: Ergebnisse der gegenseitigen Überlagerung von Konformationen vom Typ B: Roten Punkte: optimaler Durchschnitt von 0.367, blauen Punkte: $\mathsf{S}_{\mathcal{A}_1^i,\dots,\mathcal{A}_5^i}(u^0_{\mathcal{A}_1^i},\dots,u^0_{\mathcal{A}_5^i})$. $\mathsf{S}_{\mathcal{A}_1^i,\dots,\mathcal{A}_5^i}(u^*_{\mathcal{A}_1^i},\dots,u^*_{\mathcal{A}_5^i})$: Gauß-Seidel-artigen Verfahren 4.6 mit MAXSTEP=10 (orange), Gauß-Seidel-artigen Verfahren 4.6 mit MAXSTEP=17 (schwarz), Optimierungsalgorithmus 4.2 (grün).

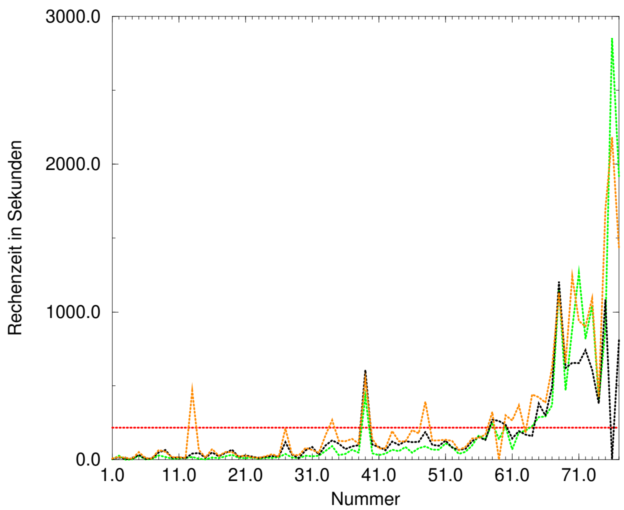


Abbildung 4.7: Rechenzeit für die gegenseitigen Überlagerungen von Konformationen vom Typ B: Roten Punkte: durchschnittliche Rechenzeit 217.44 sec. Rechenzeit der Verfahren: Gauß-Seidel-artigen Verfahren 4.6 mit MAXSTEP=10 (orange), Gauß-Seidel-artigen Verfahren 4.6 mit MAXSTEP=17 (schwarz), Optimierungsalgorithmus 4.2 (grün).

worden, während in den Abbildungen 4.6 und 4.7 die Ergebnisse der gegenseitigen Überlagerung der Konformationen von Typ B (Konformationen der 1., 3., 5., 7. und 9. Transformation) dargestellt worden sind. Dabei werden die Überlagerungen durch die Identifikationsnummern der 77 Moleküle bezeichnet. Für jede Überlagerung vom Typ A bzw. Typ B des Moleküls $\mathcal{A}_i$, $i = 1, \dots, 77$, wird der Wert $\mathsf{S}_{\mathcal{A}_1^i, \dots, \mathcal{A}_5^i}(u^0_{\mathcal{A}_1^i}, \dots, u^0_{\mathcal{A}_5^i})$ und der Wert $\mathsf{S}_{\mathcal{A}_1^i, \dots, \mathcal{A}_5^i}(u^*_{\mathcal{A}_1^i}, \dots, u^*_{\mathcal{A}_5^i})$ angegeben worden, wobei $\mathsf{S}_{\mathcal{A}_1^i, \dots, \mathcal{A}_5^i}$ die gegenseitige Überlagerung der jeweiligen Konformationen bewertet. Die Ergebnisse, die vom Verfahren 4.6 mit MAXSTEP=10 bzw. MAXSTEP=17 berechnet worden sind, unterscheiden sich nicht sehr voneinander. Also wurde für das Gauß-Seidel-artige Verfahren MAXSTEP gleich 17 gesetzt, denn aus den Abbildungen 4.5 und 4.7 folgt, daß sich diese beiden Verfahren in der Rechenzeit erheblich unterscheiden sich. In den Tabellen D.4 und D.5 (Seite 182f.) werden pro Molekül die Ergebnisse des Minimierungsverfahrens 4.2 und des Optimierungsverfahrens 4.6 mit MAXSTEP=17 gegenübergestellt. Die Dimension der Optimierungsaufgaben nimmt mit der Identifikationsnummer zu. Dies korrespondiert im wesentlichen mit den Abbildungen 4.5 und 4.7.

Als Fazit läßt sich sagen, daß der Algorithmus 4.6 mit MAXSTEP=17 die Energiefunktion $\mathsf{E}_{\mathcal{A}_1, \dots, \mathcal{A}_{n_m}}$ im Durchschnitt besser minimiert hat als das normale Verfahren, wie sich aus der Tabelle 4.3 ablesen läßt. In 11 Fällen ist das Gauß-Seidel-artige Verfahren sogar schneller als der Minimierungsalgorithmus 4.2.[33] Aber aus den Abbildungen 4.4 und 4.6 folgt auch, daß es durchaus Beispiele gibt, wo die direkte Minimierung mit dem Algorithmus 4.2 bessere Ergebnisse liefert als der Algorithmus 4.6.

Bemerkung: Das modifizierte heuristische Tesfverfahren hatte die Aufgabe, pro Molekül eine gemeinsame Konformation aller fünf transformierten Konformationen vom Typ A bzw. vom Typ B zu finden. Somit ist für $i = 1, \dots, 77$ das globale Optimum von $\mathsf{S}_{\mathcal{A}_1^i, \dots, \mathcal{A}_5^i}$ 1.0. Zu den Gründen, warum dieses Optimum nicht erreicht worden ist, siehe die Bemerkung am Ende von Kapitel 4.2.

Das Hierarchie-Verfahren

Ziel ist es hier, ein Verfahren zu konzipieren, daß die Energiefunktion aus Definition 3.2.2 effizienter minimiert wird. Wie schon im Gauß-Seidel-artigen Optimierungsverfahren sollen anstatt eines hochdimensionalen Optimierungsproblems geeignete niedrigdimensionale Teiloptimierungsprobleme gelöst werden. Wie in der Bemerkung zu Korollar 2.4.17 in Kapitel 2.4.2 schon erwähnt worden ist, gibt es eine natürliche Hierarchie der internen Rotationsparameter. Man vergleiche hierzu aber auch den zum Molekül Napap (1dwd) gehörenden Fragmentebaum in Abbildung 2.5 auf Seite 15 und den gerichteten Baum in Abbildung 2.8 auf Seite 34. In Abbildung 4.8 werden der Fragmentebaum $\vec{T}_{Napap}$ von Napap sowie die internen Rotationsparameter nochmals dargestellt. Anhand der Abbildung 4.8 soll die Hierarchie der internen Rotationsparameter erläutert werden, wobei die Bezeichnungen aus Kapitel 2.2 verwendet werden. $\vec{T}_{Napap} = (F, \vec{E})$ sei der zum Molekül Napap gehörende Fragmentebaum. Der Knoten

[33]Man vergleiche die Tabelle D.4: Moleküle 39, 52, 63, 67 und 68; und die Tabelle D.5: Moleküle 12, 29, 72, 73 und 74.

$f_i \in F$ des Fragmentebaumes repräsentiert das Molekülfragment F_i.[34] In Schritt 4 des Al-

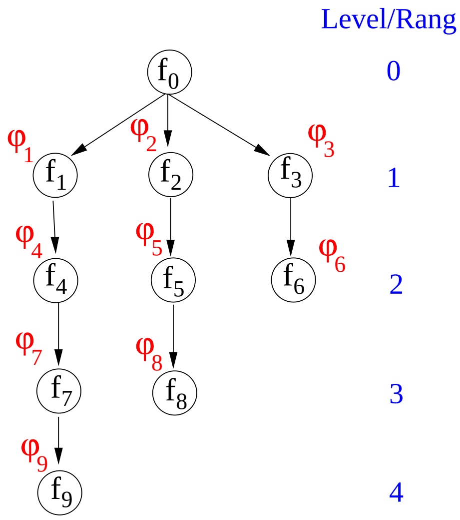


Abbildung 4.8: Der zum Molekül Napap gehörende Fragmentebaum $\vec{T}_{Napap}$ aus Abbildung 2.5

gorithmus 4.1 in Kapitel 4.1 wird jedem Knoten $f_i \neq f_0$ des Fragmentebaumes der interne Rotationsparameter φ_i zugeordnet, wobei $f_0 \in F$ die Wurzel des Baumes ist. Da für $f_i \neq f_0$ die f_0-f_i−Wege in $\vec{T}_{Napap}$ eindeutig bestimmt sind, ist damit auch die Hierarchie der internen Rotationsparameter eindeutig bestimmt. Die Länge des f_0-f_i−Weges bestimmt den Rang des zugehörigen Parameters φ_i. So haben die Rotationsparameter φ_1, φ_2 und φ_3 aus Abbildung 4.8 den Rang 1. Die Rotationsparameter φ_4, φ_5 und φ_6 haben dagegen den Rang 2, während die Parameter φ_7 und φ_8 den Rang 3 haben. Der Rotationsparameter φ_9 hat den Rang 4. Die internen Rotationsparameter φ_{i_1} mit Rang 1 sind von allen anderen Rotationsparametern unabhängig, wie man der Definition der Hintereinanderschaltung von lokalen Rotationen aus Satz 2.4.16 entnehmen kann, während die Parameter φ_{i_2} mit Rang 2 nur von dem entsprechenden Rotationsparameter φ_{i_1} abhängen. Dabei ist φ_{i_1} durch den f_0-f_{i_2}−Weg $(f_0 f_{i_1} f_{i_2})$ eindeutig bestimmt. Damit steht der Winkel φ_{i_1} in der Hierachie höher als der Winkel φ_{i_2}. Induktiv kann man mit dieser Charakterisierung der Rotationsparameter fortfahren. Je länger der f_0-f_j−Weg in $\vec{T}_{Napap}$ ist, um so abhängiger ist der dem Knoten f_j zugeordnete Parameter φ_j von entsprechenden Rotationsparameter höreren Ranges. Diese sind durch den f_0-f_j−Weg eindeutig bestimmt. Somit ist der Rotationsparameter φ_9 von den Parametern φ_1, φ_4 und φ_7 abhängig. Nun soll anhand des Moleküls Napap (1dwd) das Prinzip des Verfahrens erklärt werden. Dazu sei $\mathcal{R}$ ein beliebiges Referenzmolekül. Um das optimale Alignment von Napap bzgl. dem Referenzmolekül $\mathcal{R}$ zu bestimmen, muß man die Optimierungsaufgabe (3.4) aus Definition 3.2.2

$$\min_{u \in U(n_a[Napap])} \mathsf{E}(u)$$

lösen, wobei $\mathsf{E} := \mathsf{E}_{Napap,\mathcal{R}}$ die entsprechende Energiefunktion aus Definition 3.2.2 und $u := (t, q, c, \varphi_1, \ldots, \varphi_9) \in U(n_a[Napap]) := \mathbb{R}^3 \text{x} \mathbb{R}^4 \setminus \{0\} \text{x} \mathbb{R}^3 \text{x} \mathbb{R}^9$ der Transformationsparameter ist. Dabei ist t der Translations-, q der Quaternionenvektor und c das globale Rotationszentrum.

[34]Siehe den entsprechenden Abschnitt in Kapitel 2.2 sowie die Abbildung 2.5.

Die internen Rotationsparameter sind durch $\varphi_1, \ldots, \varphi_9$ gegeben, wobei φ_i für $i = 1, \ldots, 9$ dem Knoten f_i zugeordnet worden ist. Um nun das obige Optimierungsproblem zu lösen, geht man wie folgt vor: zuerst löst man das Optimierungsproblem

$$\min_{(t,q,c)\in \mathbb{R}^3 \mathrm{x} \mathbb{R}^4 \setminus \{0\} \mathrm{x} \mathbb{R}^3} \mathsf{E}(t, q, c), \tag{4.26}$$

wobei die Energiefunktion E dahingegehend modifiziert worden ist, daß die internen Rotationsparameter nicht betrachtet werden. D.h. zuerst wird nur die globale Position von Napap im Raum optimiert oder mit anderen Worten, das Molekül Napap wird rigide auf das Referenzmolekül $\mathcal{R}$ überlagert. Rigide bedeutet in diesem Zusammenhang, daß das Molekül Napap starr gehalten wird. Der Minimierungsalgorithmus 4.2 wird verwendet, um das Optimierungsproblem (4.26) zu lösen. Dazu sei $(t_0, q_0, c_0) \in \mathbb{R}^3 \mathrm{x} \mathbb{R}^4 \setminus \{0\} \mathrm{x} \mathbb{R}^3$ der Startwert, wobei $t_0 = (0, 0, 0)$, $q_0 = (0, 0, 0, 1)$ und c_0 der Molekülschwerpunkt von Napap ist. Das Ergebnis der Minimierung sei $(t_1, q_1, c_1) := (t^*, q^*, c^*) \in \mathbb{R}^3 \mathrm{x} \mathbb{R}^4 \setminus \{0\} \mathrm{x} \mathbb{R}^3$. Nun löst man das Optimierungsproblem

$$\min_{u_{g,1}\in U(g,1)} \mathsf{E}(u_{g,1}), \tag{4.27}$$

wobei $u_{g,1} := (t, q, c, \varphi_1, \varphi_2, \varphi_3) \in U(g, 1) := \mathbb{R}^3 \mathrm{x} \mathbb{R}^4 \setminus \{0\} \mathrm{x} \mathbb{R}^3 \mathrm{x} \mathbb{R}^3$ und die Energiefunktion E dahingegehend modifiziert worden ist, daß nur die internen Rotationsparmeter φ_1, φ_2 und φ_3 betrachtet werden. Die internen Rotationsparameter φ_1, φ_2 und φ_3 haben den Rang 1. Wieder wird der Algorithmus 4.2 verwendet, um das Optimierungsproblem zu lösen. Als Startwert wird diesmal $u^0_{g,1} := (t_1, q_1, c_1, 0, 0, 0)$ verwendet. Das Ergebnis der Minimierung sei $u^1_{g,1} := u^*_{g,1} = (t^{**}, q^{**}, c^{**}, \varphi^*_1, \varphi^*_2, \varphi^*_3)$. Die Hoffnung besteht darin, daß der Algorithmus 4.2 nur wenige Iterationsschritte benötigt, um $u^*_{g,1}$ zu bestimmen, da als Startwerte die optimierten Parameter t_1, q_1 und c_1 verwendet werden. Als nächstes wird das Optimierungsproblem

$$\min_{u_{g,2}\in U(g,2)} \mathsf{E}(u_{g,2}) \tag{4.28}$$

betrachtet, wobei $u_{g,2} := (t, q, c, \varphi_1, \ldots, \varphi_6) \in U(g, 2) := \mathbb{R}^3 \mathrm{x} \mathbb{R}^4 \setminus \{0\} \mathrm{x} \mathbb{R}^3 \mathrm{x} \mathbb{R}^6$ und die Energiefunktion E entsprechend modifiziert worden ist. Die internen Rotationsparameter φ_4, φ_5 und φ_6 haben den Rang 2. Als Startwert für den Minimierungsalgorithmus wird $u^0_{g,2} := (u^1_{g,1}, 0, 0, 0)$ verwendet. Das Ergebnis der Optimierung $u^1_{g,2} := u^*_{g,2} = (u^{**}_{g,1}, \varphi^*_4, \varphi^*_5, \varphi^*_6)$ wird nun verwendet, um das Optimierungsproblem

$$\min_{u_{g,3}\in U(g,3)} \mathsf{E}(u_{g,3}) \tag{4.29}$$

zu lösen. Dabei gilt $u_{g,3} := (t, q, c, \varphi_1, \ldots, \varphi_8) \in U(g, 3) := \mathbb{R}^3 \mathrm{x} \mathbb{R}^4 \setminus \{0\} \mathrm{x} \mathbb{R}^3 \mathrm{x} \mathbb{R}^8$. Die Energiefunktion E ist entsprechend modifiziert worden. Die internen Rotationsparameter φ_7 und φ_8 haben den Rang 3. Als Startwert wird $u^0_{g,3} := (u^1_{g,2}, 0, 0)$ verwendet. $u^1_{g,3} := u^*_{g,3} = (u^{**}_{g,2}, \varphi^*_7, \varphi^*_8)$ sei das Ergebnis der Optimierung. Zum Schluß wird nun das Optimierungsproblem

$$\min_{u_{g,4}\in U(g,4)} \mathsf{E}(u_{g,4}) \tag{4.30}$$

gelöst, wobei $u_{g,4} = u = (t, q, c, \varphi_1, \ldots, \varphi_9) \in U(g, 4) = U(n_a[Napap])$ gilt. Das Optimierungsproblem (4.30) entspricht dem Optimierungsproblem (3.4) aus Definition 3.2.2. Als Startwert für das Minimierungsverfahren wird jedoch $u^0_{g,4} := (u^1_{g,3}, 0)$ verwendet. $u^* = (u^{**}_{g,3}, \varphi^*_9)$ sei das Ergebnis der Optimierung. Offensichtlich ist u^* eine Lösung des Optimierungsproblems (3.4) aus Definition 3.2.2. Das Optimierungsproblem (3.4) aus Definition 3.2.2 wird also mittels der Optimierungsprobleme (4.26) bis (4.30) gelöst. Indem man die Minimierungsprobleme (4.26) bis (4.29) löst, sollen jeweils geeignete Startwerte generiert werden, die es dem Minimierungsalgorithmus ermöglichen, das jeweils aktuelle Optimierungsproblem mit nur wenigen Iterationen zu lösen. Die Reihenfolge der internen Rotationsparameter, die jeweils zum Optimierungsproblem hinzugenommen werden, wird durch den Rang der Parameter in der Hierarchie eindeutig bestimmt. Damit ist das Prinzip des Hierarchie-Verfahrens dargestellt worden.

Nun soll das Hierarchie-Verfahren allgemein formuliert werden.

Definition 4.3.3 (Hierarchie-Verfahren)
Es sei $\mathcal{A}$ ein flexibles Molekül und $\mathcal{R}$ ein rigides Referenzmolekül. Weiter sei $\mathsf{C}_{\mathcal{A}}$ eine zulässige Molekülkonformation von $\mathcal{A}$ und $\mathcal{DS}_{\mathcal{A}}(\mathsf{C}_{\mathcal{A}})$ die Datenstruktur[35] von $\mathcal{A}$ bzgl. $\mathsf{C}_{\mathcal{A}}$. $\vec{T}_{\mathcal{A}} = (F_{\mathcal{A}}, \vec{E}_{\mathcal{A}})$ sei der zum Molekül $\mathcal{A}$ gehörende und gerichtete Fragmentebaum, der aus $\mathcal{DS}_{\mathcal{A}}(\mathsf{C}_{\mathcal{A}})$ abgeleitet worden ist. Um das optimale Alignment des flexiblen Moleküls $\mathcal{A}$ bzgl. des rigiden Referenzmoleküls $\mathcal{R}$ zu bestimmen, wird das Optimierungsproblem

$$\min_{u_{\mathcal{A}} \in U(n_a[\mathcal{A}])} \mathsf{E}_{\mathcal{A},\mathcal{R}}(u_{\mathcal{A}}) \tag{4.31}$$

gelöst, wobei $\mathsf{E}_{\mathcal{A},\mathcal{R}}$ die entsprechende Energiefunktion aus Definition 3.2.2 ist. Der Parameter der Energiefunktion ist durch $u := (t, q, c, \varphi_1, \dots, \varphi_{n_{ft}}) \in U(n_a[\mathcal{A}])$ gegeben, wobei $U(n_a[\mathcal{A}]) := \mathbb{R}^3 \mathrm{x} \mathbb{R}^4 \setminus \{0\} \mathrm{x} \mathbb{R}^3 \mathrm{x} \mathbb{R}^{n_{ft}}$ der Transformationsraum ist. Dabei ist n_{ft} die Anzahl der Bindungen, an denen der Torsionswinkel varriert werden kann. Damit gibt n_{ft} die Anzahl der internen Rotationsparameter an. Die internen Rotationsparameter sind durch $\varphi_1, \dots, \varphi_{n_{ft}}$ gegeben, wobei für $i = 1, \dots, n_{ft}$ φ_i dem Knoten $f_i \in F_{\mathcal{A}}$ zugeordnet worden ist. $f_0 \in F_{\mathcal{A}}$ sei die Wurzel von $\vec{T}_{\mathcal{A}}$. Weiter sei für $f_i \neq f_0$ $l(f_i)$ die Länge des eindeutig bestimmten f_0-f_i−Weges in $\vec{T}_{\mathcal{A}}$. Damit gibt $l(f_i)$ den Rang von φ_i an. Dann setzt man $t(\vec{T}_{\mathcal{A}}) := \max_{f_i \neq f_0} l(f_i)$, wobei $t(\vec{T}_{\mathcal{A}})$ die Länge des längsten Weges in $\vec{T}_{\mathcal{A}}$ angibt. Für $j = 1, \dots, t(\vec{T}_{\mathcal{A}})$ sei $n(j)$ die Anzahl der Knoten f_i mit $l(f_i) = j$. Daraus folgt offensichtlich, daß $n_{ft} = \sum_{j=1}^{t(\vec{T}_{\mathcal{A}})} n(j)$ gilt. Dann lautet das Hierarchie-Verfahren wie folgt:

Algorithmus 4.7 (Hierarchie-Algorithmus)

(S.0) Stoppkriterien: $\varepsilon_G > 0$, $\varepsilon_S > 0$, $\varepsilon_E > 0$ und MAXSTEP.
Falls MAXSTEP / $t(\vec{T}_{\mathcal{A}}) < 10$, setze LOKALSTEP = 10. Falls MAXSTEP / $t(\vec{T}_{\mathcal{A}}) > 25$, setze LOKALSTEP = 25. Sonst setze LOKALSTEP = MAXSTEP / $t(\vec{T}_{\mathcal{A}})$.

(S.1) Löse mit dem Minimierungsalgorithmus 4.2 das Optimierungsproblem

$$\min_{(t,q,c) \in \mathbb{R}^3 \mathrm{x} \mathbb{R}^4 \setminus \{0\} \mathrm{x} \mathbb{R}^3} \mathsf{E}^0_{\mathcal{A},\mathcal{R}}(t, q, c),$$

wobei $\mathsf{E}^0_{\mathcal{A},\mathcal{R}}$ die entsprechend modifizierte Energiefunktion ist. Als Startwert wird (t_0, q_0, c_0) mit $t_0 = (0, 0, 0)$, $q_0 = (0, 0, 0, 1)$ und mit c_0, der Molekülschwerpunkt von $\mathcal{A}$, verwendet. Das

[35] Siehe Definition 2.2.3

Ergebnis der Minimierung sei (t^*, q^*, c^*). Als Stopkriterien werden $\varepsilon_G > 0$, $\varepsilon_S > 0$, $\varepsilon_E > 0$ und LOKALSTEP verwendet.

(S.2) Für $i = 1, \ldots, t(\vec{T}_{\mathcal{A}}) - 1$ löse das Optimierungsproblem

$$\min_{u_{g,i} \in U(g,i)} \mathsf{E}^i_{\mathcal{A},\mathcal{R}}(u_{g,i}),$$

wobei $\mathsf{E}^i_{\mathcal{A},\mathcal{R}}$ die entsprechend modifizierte Energiefunktion ist. Der Parameter ist durch $u_{g,i} := (t, q, c, \varphi_1, \ldots, \varphi_{n(1)+\ldots+n(i)}) \in U(g,i) := \mathbb{R}^3 \mathrm{x} \mathbb{R}^4 \setminus \{0\} \mathrm{x} \mathbb{R}^3 \mathrm{x} \mathbb{R}^{n(1)+\ldots+n(i)}$ gegeben. Als Startwert für den Minimierungsalgorithmus wird $u^0_{g,i} := (u^1_{g,i-1}, 0, \ldots, 0)$ verwendet, wobei $u^1_{g,0} := (t^*, q^*, c^*)$ ist. Das Ergebnis der Minimierung sei $u^1_{g,i} := (u^{**}_{g,i-1}, \varphi^*_{n(1)+\ldots+n(i-1)+1}, \ldots, \varphi^*_{n(1)+\ldots+n(i)})$. Als Stopkriterien werden $\varepsilon_G > 0$, $\varepsilon_S > 0$, $\varepsilon_E > 0$ und LOKALSTEP verwendet.

(S.3) Löse nun das Optimierungsproblem

$$\min_{u \in U(n_a[\mathcal{A}])} \mathsf{E}_{\mathcal{A},\mathcal{R}}(u),$$

wobei als Startwert für den Minimierungsalgorithmus $u^0 := (u^1_{g,t(\vec{T}_{\mathcal{A}})-1}, 0, \ldots, 0)$ verwendet wird. Das Ergebnis der Minimierung sei $u^* := (u^{**}_{g,t(\vec{T}_{\mathcal{A}})-1}, \varphi^*_{n(1)+\ldots+n(t(\vec{T}_{\mathcal{A}})-1)+1}, \ldots, \varphi^*_{n(1)+\ldots+n(t(\vec{T}_{\mathcal{A}}))})$. Als Stopkriterien werden $\varepsilon_G > 0$, $\varepsilon_S > 0$, $\varepsilon_E > 0$ und MAXSTEP verwendet.

Ziel des Verfahrens ist es, das Optimierungsproblem (4.31) effizienter zu minimieren. Dazu werden im Schritt 1 und Schritt 2 für $i = 1, \ldots, t(\vec{T}_{\mathcal{A}}) - 1$ geeignete Startwerte für das darauf folgende Minimierungsverfahren generiert, um schließlich in Schritt 3 das Optimierungsproblem (4.31) mit einem geeigneten Startwert zu lösen. $\diamond$

Bemerkung:

a) Da in Schritt 3 schließlich das Optimierungsproblem (4.31) selbst betrachtet wird, ist das Hierarchie-Verfahren 4.7 letztendlich ein Minimierungsalgorithmus für die Optimierungsaufgabe (4.31). Das Prinzip des Hierarchie-Verfahrens besteht darin, geeignete Startwerte zu generieren. Dazu werden im Schritt 1 und 2 geeignet modifizierte Energiefunktionen optimiert. Diese Energiefunktionen werden jedoch nur „grob" optimiert, da es sich herausgestellt hat, daß sich die Anzahl der Iterationen durchaus in der Größenordnung von $t(\vec{T}_{\mathcal{A}})$ MAXSTEP bewegen kann, wenn LOKALSTEP=MAXSTEP verwendet wird. Daher werden maximal 25 Iterationen ausgeführt. Erst in Schritt 3 wird MAXSTEP, falls es notwendig ist, als Abbruchkriterium verwendet.

b) Man kann das Hierarchie-Verfahren 4.7 dazu verwenden, um das Optimierungsproblem (4.24) in Schritt 3 des Gauß-Seidel-artigem Optimierungsverfahren 4.6 zu lösen.

c) Damit das obige Hierachie-Verfahren wirklich effizient ist, sollte es so implementiert werden, daß nicht nur die generierten Startwerte übernommen werden, sondern auch alle anderen Daten, die bereits berechnet worden sind, wie die Abstiegsrichtung d_k, der CG-Parameter

γ_k bzw. die Quasi-Newton Matrix H_k. Denn sonst würden all diese Informationen ungenutzt verloren gehen.

Um nun zu überprüfen, ob das Hierachie-Verfahren das Optimierungsproblem (4.31) effizienter löst als die direkte Minimierung mit dem Algorithmus 4.2, wird das schon in Kapitel 4.2 beschriebene heuristische Testverfahren angewandt. Um diese 770 Testbeispiele zu berechnen, sind die entsprechenden Energiefunktionen aus Definition 3.2.2 minimiert worden. Die Energiefunktionen sind dahingehend modifiziert worden, daß die Bewertungsfunktion für die Torsionswinkel deaktiviert worden ist. Der zu optimierende Transformationsparameter der Energiefunktionen besteht aus dem Translationsvektor, dem Quaternionenvektor, dem globalen Rotationszentrum und den internen Rotationsparametern. Um den Algorithmus 4.2 zu konkretisieren, ist das inverse BFGS-Verfahren mit der Aufdatierungsformel (4.10) und der Armijoregel mit Aufweitung (Parameter $\beta = 0.35$) verwendet worden. In Tabelle 4.4 sind die durchschnittlichen Ergebnisse der 770 Optimierungsverfahren dargestellt worden. Die einzelnen Spalten sind im Abschnitt über das heuristische Testverfahren in Kapitel 4.2 beschrieben worden. In der Tabelle D.3 in

Algorithmus	S_E^0	S_E^*	Schritte	Sec	Stopp	Q	N
4.2	0.102	0.654	76.81	31.75	0.34	2.88	1.31
4.7	0.102	0.665	136.49	47.63	0.39	3.32	0.00

Tabelle 4.4: Ergebnisse der Algorithmen 4.2 und 4.7

Anhang D werden die durchschnittlichen Ergebnisse der beiden Verfahren pro Molekül gegenübergestellt. In der Abbildung 4.9 werden die durchschnittlichen Ergebnisse der Optimierungen dargestellt. Für jedes der 77 Moleküle wird der durchschnittliche Alignmentindex, der in Gleichung (D.1) auf Seite 177 definiert worden ist, dargestellt. Die Moleküle werden durch ihre Identifikationsnummern dargestell. In Abbildung 4.10 wird die durchschnittliche Rechenzeit in Sekunden für die zwei Verfahren dargestellt.

Als Fazit läst sich sagen, daß der Hierarchie-Algorithmus die Energiefunktion $\mathsf{E}_{\mathcal{A},\mathcal{R}}$ zwar nicht schneller minimieren konnte, aber dafür im Durchschnitt etwas besser. Aus der Abbildungen 4.9 folgt aber, daß es durchaus Beispiele gibt, wo die direkte Minimierung mit dem Algorithmus 4.2 bessere Ergebnisse liefert, als der Algorithmus 4.7.

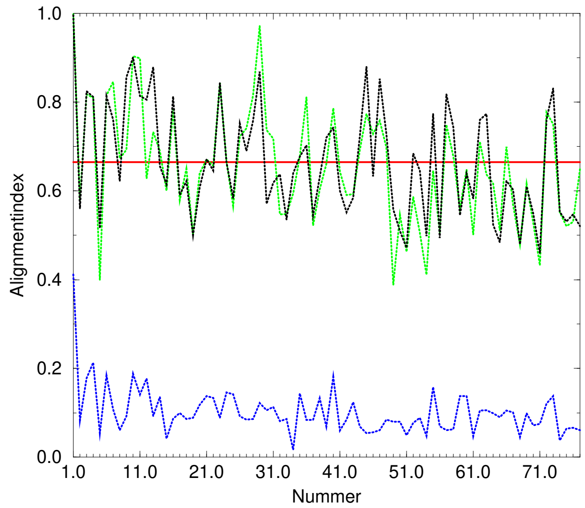


Abbildung 4.9: Durchschnittliche Ergebnisse der Algorithmen 4.2 und 4.7 pro Molekül: Roten Punkte: optimaler Durchschitt von 0.665, blauen Punkte: durchschnittliche Alignmentindizes vor den Optimierungen. Durchschnittliche Alignmentindizes nach den Optimierungen: Minimierungsalgorithmus Algorithmus 4.2 (grün), Hierachie-Verfahren 4.7 (schwarz).

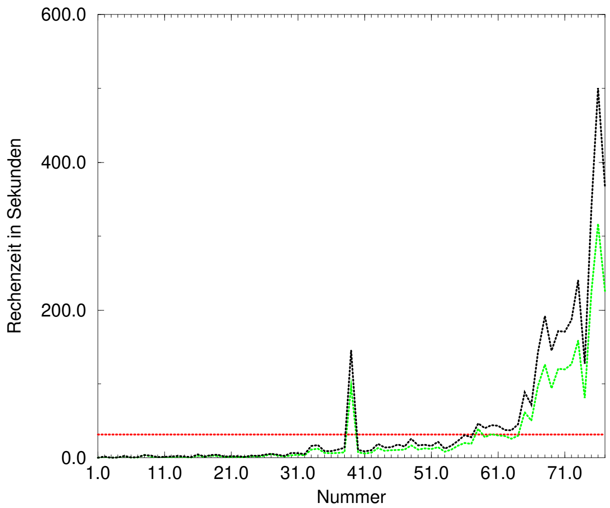


Abbildung 4.10: Durchschnittliche Rechenzeit der Algorithmen 4.2 und 4.7 pro Molekül: Roten Punkte: optimaler Durchschnitt von 31.75 sec. Durchschnittliche Rechenzeit in Sekunden: Minimierungsalgorithmus Algorithmus 4.2 (grün), Hierachie-Verfahren 4.7 (schwarz).

Kapitel 5

Die Implementierung

In Kapitel 5.1 werden einige Aspekte der Implementierung kurz dargestellt. In Kapitel 5.2 wird dargestellt, wie die Energieparameter der Energiefunktionen aus Kapitel 3 kalibriert worden sind.

5.1 Aspekte der Implementierung

In Kapitel 4 sind die Algorithmen beschrieben worden, die implementiert worden sind. In diesem Kapitel werden einzelne Aspekte der Implementierung dargestellt, wie zum Beispiel die verwendeten Datentypen oder die Realisierung des Transformationsoperators. Bei der Beschreibung der einzelnen Aspekte werden unteranderem die Bezeichnungen aus Kapitel 2.2 verwendet. Dazu sei $\mathcal{A}$ ein Molekül mit n_a Atomen und n_{ft} Bindungen, an denen der Torsionswinkel variiert werden kann. Weiter sei $\mathsf{C}_\mathcal{A}$ die zugehörige Konformation, $K_\mathcal{A} \subset \mathbb{R}^3$ die Menge der Atomkoordinaten und $\vec{T}_\mathcal{A}$ der zugehörige Fragmentebaum. Mit $a_i \in K_\mathcal{A}$, $i = 1, \ldots, n_a$, werden sowohl das Atom als auch die entsprechenden Atomkoordinaten bezeichnet. Die Knotenmenge des Baumes $\vec{T}_\mathcal{A}$ ist durch $F_\mathcal{A} := \{f_0, \ldots, f_{n_{ft}}\}$ gegeben, wobei f_0 die Wurzel des Baumes ist.

Die implementierten Datentypen:

Es sind die folgenden Datentypen implementiert worden: *moldaten*, *fragdaten*, *atomdaten*, *gaussdaten* und *optdaten*. Um die Daten des Moleküls zu speichern, wird der Datentyp *moldaten* verwendet. Für jeden Knoten f_j des Fragmentebaums wird der Datentyp *fragdaten* verwendet, um die dem Knoten f_j zugeordneten Daten aufzunehmen. Um die Daten der Atome bzw. der Gaussfunktionen zu speichern, wird für jedes Atom bzw. für jede Gaussfunktion der Datentyp *atomdaten* bzw. *gaussdaten* verwendet. Für das Minimierungsverfahren und die entsprechenden Daten wird der Datentyp *optdaten* verwendet.
Der Datentyp *atomdaten* enthält alle Informationen über das entsprechende Atom a_i, das dem Fragmentknoten $\Psi_{K_\mathcal{A},F_\mathcal{A}}(a_i) =: f_j$ zugeordnet worden ist. Somit enthält die Struktur die Atomnummer i, einen Zeiger auf den Datentyp *fragdaten* für den Fragmentknoten f_j, die Fragment-

nummer j, einen Vektor für die Atomkoordinate $a_i \in \mathbb{R}^3$, einen Vektor für die lokale Koordinate $a_i - \mathsf{R}_{F_\mathcal{A},K_\mathcal{A}}(f_j) \in \mathbb{R}^3$, wobei $\mathsf{R}_{F_\mathcal{A},K_\mathcal{A}}(f_j)$ die Koordinate des dem Knoten f_j zugeordneten Referenzatoms ist.[1] Weiter enthält der Datentyp einen Vektor für die aktuellen Koordinaten des Atoms, einen Vektor für die aktuellen lokalen Koordinaten, einen Zeiger auf den Datentyp *atomdaten* für das Referenzatom $\mathsf{R}_{F_\mathcal{A},K_\mathcal{A}}(f_j)$, den Van-der-Waals-Radius $\sigma_i \in \mathbb{R}$ und einen Vektor für eventuelle Kristallkoordinaten.

Der Datentyp *gaussdaten* enthält alle notwendigen Daten über die Gaussfunktion $g_{l,q}$, die dazu verwendet wird, um die räumlich lokalisierte Dichte der physikochemischen Eigenschaft $q \in \mathcal{I}_{PE}$ darzustellen. Die Struktur enthält daher den Index q, einen Zeiger auf den Datentyp *fragdaten* für den Fragmentknoten $\Psi_{K^q_\mathcal{A},F_\mathcal{A}}(z_{l,q}) =: f_j$, wobei $g_{l,q}$ dem Fragmentknoten f_j zugeordnet worden ist, die Fragmentnummer j, die Höhe $h_{l,q} \in \mathbb{R}$, die Breite $b_{l,q} \in \mathbb{R}_+$ und einen Vektor für das Gausszentrum $z_{l,q} \in \mathbb{R}^3$, einen Vektor für die lokale Koordinate $z_{l,q} - \mathsf{R}_{F_\mathcal{A},K_\mathcal{A}}(f_j) \in \mathbb{R}^3$, wobei $\mathsf{R}_{F_\mathcal{A},K_\mathcal{A}}(f_j)$ die Koordinate des dem Knoten f_j zugeordneten Referenzatoms ist. Weiter enthält der Datentyp je einen Vektor für das aktuelle Zentrum, sowie für die aktuelle lokale Koordinate des Zentrums.

Der Datentyp *fragdaten* enthält alle Daten, die dem entsprechenden Fragmentknoten f_j zugeordnet worden sind. Somit enthält die Struktur die Fragmentnummer j, die Länge l_j des f_0-f_j–Weges im Fragmentebaum $\vec{T}_\mathcal{A}$, wobei die Länge dem Rang des internen Rotationsparameters $\varphi_j \in \mathbb{R}$ entspricht, einen Zeiger[2] auf den internen Rotationsparameter $\varphi_j \in \mathbb{R}$ im Transformationsvektor sowie einen Zeiger auf die entsprechende Position im Gradienten der Energiefunktion, wobei sowohl der Transformationsvektor als auch der Gradient im Datentyp *moldaten* enthalten ist, die Rotationsmatrix $\mathcal{B}^{l_j}(\dots)$ aus Definition 2.4.10, die Drehachse $\mathsf{D}_{F_\mathcal{A},\mathbb{R}^3\setminus\{0\}}(f_j)\mathbb{R}^3$, einen Marker, ob der Rotationsparameter φ_j im Hierarchieverfahren 4.7 festgehalten werden soll, den Torsionswinkel $\Theta_j(\mathsf{C}_\mathcal{A}) \in \mathbb{R}$, der aus der Konformation $\mathsf{C}_\mathcal{A}$ berechnet und dem Knoten f_j zugeordnet worden ist, den aktuellen Torsionswinkel $\Theta_j(\mathsf{C}_\mathcal{A}) + \varphi_j \in \mathbb{R}$, einen Zeiger auf den Datentyp *atomdaten* für das Referenzatom $\mathsf{R}_{F_\mathcal{A},K_\mathcal{A}}(f_j)$, einen Zeiger auf einen Zeiger auf den Datentyp *atomdaten* für die Atome, die dem Fragmentknoten f_j zugeordnet worden sind, sowie deren Anzahl. Weiter enthält die Struktur ein Feld der Länge $|\mathcal{I}_{PE}|$, um die jeweilige Anzahl der entsprechenden Gaussfunktionen zu speichern, die dem Knoten f_j zugeordnet worden sind, sowie ein Feld der Länge $|\mathcal{I}_{PE}|$ von Zeigern, die vom Typ Zeiger auf Zeiger auf den Datentyp *gaussdaten* sind. Diese Zeiger werden verwendet, um auf die Daten der jeweiligen Gaussfunktionen zugreifen zu können. Weiter enthält die Struktur die Daten des kubischen Splines $s_{\triangle_j}$ aus Gleichung (2.24) in Kapitel 2.7.2, wobei $s_{\triangle_j}$ den aktuellen Torsionswinkel $\Theta_j(\mathsf{C}_\mathcal{A}) + \varphi_j$ bewertet. Die Daten des Splines bestehen aus einem Zeiger für den Lösungsvektor des Gleichungssystems (2.23) aus Kapitel 2.7.1, der Länge des Vektors, der Anzahl der Interpolationspunkte, der Schrittweite h der äquidistanten Gitterzerlegung und aus der Periode des Interpolationsintervalls. In Bemerkung b) nach Lemma 2.7.9 ist dargestellt worden, wie man mit diesen Daten und dem Darstellungssatz 2.7.7 den Spline auswerten kann. Eine Liste von Zeigern auf die Söhne des Knoten f_j im Fragmentebaum $\vec{T}_\mathcal{A}$ bilden den Abschluß des

[1] Zum Referenzatom siehe Kapitel 2.2 und insbesondere Schritt 4 im Algorithmus 4.1 in Kapitel 4.1.

[2] Zeiger, die nicht auf die Datentypen *moldaten*, *fragdaten*, *atomdaten*, *gaussdaten* bzw. *optdaten* zeigen, sind vom Typ Zeiger auf „*double*“.

Datentyps *fragdaten*.

Der Datentyp *moldaten* enthält die folgenden Daten über das Molekül $\mathcal{A}$: den Datentyp *atomdaten* für den Molekülschwerpunkt, die Anzahl der Atome n_a, einen Zeiger auf den Datentyp *atomdaten* für die Daten der Atome im Molekül, die Anzahl der Fragmentknoten $n_{ft}+1$, einen Zeiger auf den Datentyp *fragdaten* für die Knoten des Fragmentebaums $\vec{T}_{\mathcal{A}}$, sowie die Tiefe des Baumes, eine Marke, ob eventuell die Atomkoordinaten der Kristallstrukur bekannt sind, und den RMS-Abstand der Konformation $\mathsf{C}_{\mathcal{A}}$ zur Kristallstruktur, falls diese vorhanden ist. Weiter enthält die Struktur ein Feld der Länge $|\mathcal{I}_{PE}|$, um die jeweilige Anzahl der entsprechenden Gaussfunktionen zu speichern, die verwendet werden, um die räumlich lokalisierte Dichte der entsprechenden physikochemischen Eigenschaft $q \in \mathcal{I}_{PE}$ darzustellen, sowie ein Feld der Länge $|\mathcal{I}_{PE}|$ von Zeigern auf den Datentyp *gaussdaten*. Diese Zeiger werden verwendet, um die Daten der jeweiligen Gaussfunktionen zu speichern. Weiter ist im Datentyp *moldaten* ein Feld der Länge $|\mathcal{I}_{PE}|$ enthalten, um den konstanten Anteil im Nenner des jeweiligen Hodgkin-Indizes, der das Alignment bzgl. der physikochemischen Eigenschaft $q \in \mathcal{I}_{PE}$ bewertet, zu speichern.[3]

Der jeweilige Wert des Alignmentindizes, des n-Potentials sowie der heuristischen Bewertungsfunktion für die Torsionswinkel vor der Optimierung werden ebenfalls im Datentyp *moldaten* gespeichert. Weiter enthält der Datentyp die Dimension des Optimierungsproblems, einen Zeiger für den Transformationsvektor sowie für den Gradienten der entsprechenden Energiefunktion. Um nun auf die Werte des Translationsvektors, des Quaternionenvektors, des globalen Rotationszentrums sowie der internen Rotationsparameter im Transformationsvektor bzw. im Gradienten zugreifen zu können, enthält die Struktur jeweils Zeiger, die auf die jeweilige erste Position im Transformationsvektor bzw. im Gradienten zeigen. Der Datentyp *moldaten* enthält die Matrix R aus Gleichung (2.9) in Kapitel 2.4.1 sowie die Funktionalmatrix $\mathsf{D}R$ aus Lemma 2.4.6 6). Der Zeiger auf den Datentyp *optdaten* und ein Zeiger auf die von FlexS verwendete Datenstruktur, aus der die Daten des Moleküls gelesen worden sind, bilden den Abschluß des Datentyps.

Der Datentyp *optdaten* enthält alle Informationen, damit man den Minimierungsalgorithmus 4.2 in Kapitel 4.1 anwenden kann. Somit entählt die Struktur die Angaben, welches Minimierungsverfahren und welche Schrittweitensteuerung aus Kapitel 4.2 verwendet werden sollen, sowie etwaige Parameter des gewählten Minimierungsverfahrens bzw. der Schrittweitensteuerung. Weiter enthält der Datentyp die Angabe über die Dimension des Optimierungsproblems, einen Zeiger auf die Energiefunktion, eine Marke, ob der Grandient analytisch mittels einer Gradientenfunktion oder durch Approximation berechnet werden soll, einen Zeiger auf die Gradientenfunktion, einen Zeiger für den Parametervektor der Energiefunktion sowie für den Gradienten der Energiefunktion. Die einzelnen Minimierungsverfahren aus Kapitel 4.2 verwenden weitere Hilfsvektoren, daher enthält die Struktur für diese Hilfsvektoren entsprechende Zeiger, sowie zwei Zeiger auf Zeiger für die Aufdatierungsmatrix im Quasi-Newton Verfahren und für eine Hilfsmatrix. Die Hilfsmatrix wird zum Lösen von Gleichungssystemen während der Optimierung verwendet. Darüberhinaus enthält der Datentyp *optdaten* den Speicherplatz für die

[3] In Kapitel 3.2 konnte gezeigt werden, das die Nenner der Hodgkin-Indizes im Alignmentindex nur von den internen Rotationsparametern abhängen. Somit konnte der jeweilige Nenner in einen parameterabhängigen und in einen konstanten Teil aufgespalten werden.

Angaben über den aktuellen Wert der Energiefunktion, die Anzahl der Iterationen, die das Minimierungsverfahren ausgeführt hat, die Anzahl der Iterationen, die die Schrittweitensteuerung insgesamt ausgeführt hat, die Anzahl der Neustarts, falls ein Quasi-Newton Verfahren verwendet worden ist, die Maschinengenauigkeit des Computers, sowie über das Stoppkriterium, mit dem der Minimierungsalgorithmus beendet worden ist. Weiter enthält die Struktur die Parameter der Stoppkriterien sowie Marken, ob das jeweilige Kriterium verwendet werden soll.

Transformation von Koordinaten:

Um die Transformation der Koordinaten darzustellen, werden Korollar 2.4.17 und Lemma 2.4.28 c) verwendet. Aus Korollar 2.4.17 folgt, daß sich die Hintereinanderschaltung von lokalen Rotationen als Summe von lokalen Koordinaten darstellen läßt. Es sei $(t, q, c, \varphi_1, \ldots, \varphi_{n_{ft}})$ der Transformationsparameter, wobei $t \in \mathbb{R}^3$ der Translationsvektor, $q \in \mathbb{R}^4 \setminus \{0\}$ der Quaternionenvektor, $c \in \mathbb{R}^3$ das globale Rotationszentrum und $(\varphi_1, \ldots, \varphi_{n_{ft}}) \in \mathbb{R}^{n_{ft}}$ die internen Rotationsparameter sind. Der folgende Algorithmus realisiert den globalen Transformationsoperator $\mathcal{T}_A$ aus Definition 2.4.27:

Algorithmus 5.1

(S.1) Berechne die Rotationsmatrix $A(q)$ aus Gleichung (2.10) in Kapitel 2.4.1.

(S.2) Berechne die aktuellen Koordinaten der Atome bzw. der Gausszentren im Hauptfragment: berechne für alle Atome $a_{i,0}$ und alle Gaussfunktionen $g_{j,0,p}, p \in \mathcal{I}_{PE}$, die dem Fragmentknoten f_0 zugeordnet worden sind, die aktuellen Koordinaten $a^g_{i,0} = A(q)(a_{i,0} - c) + c + t$ bzw. $z^g_{j,0,q} = A(q)(z_{j,0,p} - c) + c + t$, wobei $z_{j,0,q}$ das Zentrum der Gaussfunktion $g_{j,0,p}$ ist.

(S.3) Berechne die aktuellen lokalen Koordinaten der Atome bzw. der Gausszentren in den Nebenfragmenten: für alle Söhne von f_0 im Baum $\vec{T}_{\mathcal{A}}$ führen den Algorithmus 5.2 mit der Matrix $A(q)$ und E_3 aus.

(S.4) Berechne die aktuellen Koordinaten der Atome bzw. der Gausszentren in den Nebenfragmenten: für $k = 1, \ldots, n_{ft}$ berechne $a^g_{i,k} = \mathsf{R}^g_{F_{\mathcal{A}},K_{\mathcal{A}}}(f_k) + a^l_{i,k}$ bzw. $z^g_{j,k,q} = \mathsf{R}^g_{F_{\mathcal{A}},K_{\mathcal{A}}}(f_k) + z^l_{j,k,q}$, wobei $a^l_{i,k}$ bzw. $z^l_{j,k,q}$ die aktuellen lokalen Koordinaten der zum Fragmentknoten f_k gehörenden Atome bzw. Gausszentren sind. Die aktuelle Koordinate des zum Knoten f_k gehörenden Referenzatoms ist durch $\mathsf{R}^g_{F_{\mathcal{A}},K_{\mathcal{A}}}(f_k)$ gegeben.[4]

Algorithmus 5.2

(S.0) Es seien $A(q)$ und R die Rotationsmatrizen, die dem Algorithmus übergeben worden sind. Es werden die aktuellen lokalen Koordinaten der Atome und der Gausszentren berechnet, die dem Knoten f_k zugeordnet worden sind.

(S.1) Es wird die aktuelle Drehachse $d_k = R\mathsf{D}_{F_{\mathcal{A}},\mathbb{R}^3\setminus\{0\}}(f_k)$ berechnet und die Rotationsmatrix $\mathsf{B}(d_k,\ \varphi_k)$ analog zu Gleichung (2.7) aus Kapitel 2.4.1 gebildet, wobei φ_j der zum Knoten f_k gehörende interne Rotationsparameter ist. Berechne dann die Rotationsmatrix $R^* = A(q)\mathsf{B}(d_k,\ \varphi_k)\,R$.

[4]Zum Referenzatom siehe Kapitel 2.2 und insbesondere Schritt 4 im Algorithmus 4.1 in Kapitel 4.1.

(S.2) Berechne für alle Atome $a_{i,k}$ bzw. alle Gaussfunktionen $g_{j,k,p}, p \in \mathcal{I}_{PE}$, die aktuellen lokalen Koordinaten: $a^l_{i,k} = R^* \left(a_{i,k} - \mathsf{R}_{F_\mathcal{A},K_\mathcal{A}}(f_k)\right)$ bzw. $z^l_{j,k,q} = R^* \left(z_{j,k,p} - \mathsf{R}_{F_\mathcal{A},K_\mathcal{A}}(f_k)\right)$, wobei $\mathsf{R}_{F_\mathcal{A},K_\mathcal{A}}(f_k)$ die Koordinate des Referenzatoms ist.

(S.3) Für alle Söhne von f_k im Baum $\vec{T}_\mathcal{A}$ führe den Algorithmus 5.2 mit den Matrizen $A(q)$ und $\mathsf{B}(d_k,\ \varphi_k)\,R$ aus.

Bemerkung: Da die Knoten des Fragmentebaums $\vec{T}_\mathcal{A}$ ausgehend von der Wurzel f_0 mit einem BFS-Algorithmus (siehe hierzu [9]) numeriert worden sind, werden in Schritt 4 des Algorithmus 5.1 die aktuellen Koordinaten der Atome bzw. der Gausszentren berechnet, die zu den Nebenfragmenten gehören.

Lösen vom linearen Gleichungsystemen:

In Schritt 1 des Algorithmus 4.1 aus Kapitel 4.1 wird für jede Bindung e des Moleküls $\mathcal{A}$, für die $\Phi^{\mathcal{A}}_{FT}(e) = 1$ gilt, das lineare Gleichungssystem $B_e\alpha_e = b_e$ aus Gleichung (2.23) in Kapitel 2.7.1 entsprechend dem Datensatz $\triangle_e$ aufgestellt und gelöst. In jeder Iteration der nichtinversen Quasi-Newton Verfahren aus Kapitel 4.2, wie dem *BFGS*-, dem *DFP*-, dem *PSB*-Verfahren, sowie der Verfahren der *Broyden-Familie*, muß zur Berechnung der Abstiegsrichtung d_k das lineare Gleichungssystem $H_k d_k = -g_k$ gelöst werden. In jeder Iteration der *Oren-Luenberger-Klasse* aus Kapitel 4.2 muß das lineare Gleichungssystem $B_k h_k = s_k$ gelöst werden, um den Parameter ε aus Gleichung (4.18) zu berechnen.
Um die jeweiligen Gleichungssysteme zu lösen, wird das Gaußsche Eliminationsverfahren mit Spaltenpivotsuche angewendet, wobei die entsprechenden Matrizen equilibriert werden. Das Gaußsche Eliminationsverfahren mit Spaltenpivotsuche wurde nach den Angaben aus Stoer [54] implementiert. Es sei $A \in \mathsf{M}(n \times n; \mathbb{R})$ die Matrix des jeweiligen Gleichungssystems. Für $i = 1, \ldots, n$ wird

$$s_i := \frac{1}{\sum_{k=1}^n |a_{ik}|}$$

berechnet, wobei die a_{ik} die Komponenten der Matrix A sind. Dann wird die Matrix A durch

$$\tilde{A} = \begin{pmatrix} s_1 & & 0 \\ & \ddots & \\ 0 & & s_n \end{pmatrix} \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \ldots & a_{nn} \end{pmatrix}$$

equilibriert. Es gilt somit $\sum_{k=1}^n |\tilde{a}_{ik}| = 1$ für $i = 1, \ldots, n$. Aber die Matrix $\tilde{A}$ wird nicht explizit berechnet, sondern nur bei der Such nach dem Spaltenpivotelement a_j in der $j-$ten Iteration des Gaußsche Eliminationsverfahrens werden die Komponenten aus der $j-$ten Spalte mit den entsprechenden Elementen s_i multipliziert.[5]
Die Anzahl der Rechenoperationen für das Gaußsche Eliminationsverfahren mit Spaltenpivotsuche und Equilibrierung läßt sich somit in

- Berechnung der Equilibrierungselemente s_i : $n^2 + n$,

[5] Die Einzelheiten findet man in den entsprechenden Kapiteln über das Gaußsche Eliminationsverfahren in Stoer [54].

- Gauß-Elimination mit Spaltenpivotsuche und Equilibrierung: $\frac{(n-1)}{6}\left(4n^2+34n+30\right)$,
- Lösen eine gestaffelten Gleichungssystems: n^2+2n+3,

aufspalten. Damit ist die Anzahl der Rechenoperationen durch $O(n^3)$ gegeben.

Bemerkung:

a) Die Matrix B_e aus Gleichung (2.23) in Kapitel 2.7.1 hat eine besondere Gestalt. Daher ist das Gaußsche Eliminationsverfahren mit Spaltenpivotsuche und Equilibrierung der Gestalt entsprechend modifiziert worden. Um das Gleichungssystem $B_e\alpha_e = b_e$ für das Eliminationsverfahren zu speichern, braucht man nur Speicherplatz in der Größenordnung $7n + 2$: und zwar für die Hauptdiagonale, für die zwei unteren Nebendiagonalen, und für die drei letzten Spalten der Matrix B_e, sowie für die rechte Seite jeweils n Speicherplatz. Und man muß die ersten zwei Komponeten der 3. Spalte speichern. Um das Gleichungssystem zu lösen, braucht das modifizierte Gaußsche Eliminationsverfahren mit Spaltenpivotsuche und Equilibrierung $O(n)$ Rechenoperationen.

b) Die Quasi-Newton Verfahren, die zur der Berechnung der Abstiegsrichtung ein lineares Gleichungssystem lösen müssen, brauchen zusätzlichen Speicher für die Equilibrierungselemente und für die Matrix im Eliminationsverfahren, insgesamt Speicherplatz in der Größenordnung n^2+n. Auch die Verfahren der *Oren-Luenberger-Klasse* brauchen den zusätzlichen Speicher.

Die Berechnung des Gradienten:

Es sei $(t, q, c, \varphi_1, \ldots, \varphi_{n_{ft}})$ der Transformationsparameter der Energiefunktionen, wobei $t \in \mathbb{R}^3$ der Translationsvektor, $q \in \mathbb{R}^4 \setminus \{0\}$ der Quaternionenvektor, $c \in \mathbb{R}^3$ das globale Rotationszentrum und $(\varphi_1, \ldots, \varphi_{n_{ft}}) \in \mathbb{R}^{n_{ft}}$ die internen Rotationsparameter sind. Die Gradienten der Energiefunktionen aus Definition 3.2.2, 3.3.3 und aus Definition 3.3.5 werden wie folgt berechnet:

- Die partiellen Ableitungen bzgl. der Komponenten des Translationsvektors t, des Quaternionenvektors q, und des globalen Rotationszentrums c werden nach den Formeln aus Lemma 3.2.3 b), 3.3.2 b), 3.3.4 b) bzw. aus Lemma 3.3.6 b) analytisch berechnet. Die Ableitungen der heuristischen Bewertungsfunktion für die Torsionswinkel bzgl. der internen Rotationsparameter $(\varphi_1, \ldots, \varphi_{n_{ft}})$ werden mit Hilfe des Darstellungssatzes 2.7.7 und der Bemerkung b) zu Lemma 2.7.9 analytisch berechnet.
- Die partiellen Ableitungen des Alignmentindizes und des n-Potentials bzgl. der internen Rotationsparameter werden approximiert, denn die Berechnung des Differentials für den lokalen Rotationsoperator aus Definiton 2.4.18 ist sehr aufwendig, wie aus Satz 2.4.24 und Lemma 2.4.25 folgt. Um die partiellen Ableitungen der Energifunktionen bzgl. der internen Rotationsparameter darzustellen, wird Lemma 2.4.26 verwendet. Aber in Lemma 2.4.26 wird das Differential des lokalen Rotationsoperators verwendet. Um also die partiellen Ableitungen bzgl. der Rotationsparameter analytisch zu berechnen, müßte man vorher das Differential berechnet haben. Aus diesem Grund werden die partiellen Ableitungen des Alignmentindizes und des n-Potentials bzgl. der internen Rotationsparameter approximiert, und zwar mit

Vorwärtsdifferenzen, d.h. $(E(x + he_i) - E(x))/h$.

5.2 Kalibrierung der Energieparameter

In diesem Kapitel wird kurz beschrieben, wie die Energieparameter λ_A, λ_P, λ_Q und λ_T in den Energiefunktionen aus den Definitionen 3.2.2, 3.3.1 bzw. 3.3.3, sowie 3.3.5 kalibriert worden sind. Es sei $\mathcal{I}_{PE}$ die Indexmenge der physikochemischen Eigenschaften aus Definition 2.2.1. Die Energiekoeffizienten $\tilde{\lambda}_p > 0$ mit $p \in \mathcal{I}_{PE}$, die normiert werden, um dann die Hodgkin-Indizes im Alignmentindex aus Definition 2.3.3 bzw. Definition 2.5.6 zu gewichten, werden vom Programm FlexS vorgegeben.

Eine vernünftige Zielfunktion sollte in der Lage sein, die strukturelle Identität eines Moleküls wiederzufinden. Daher sind die Parameter empirisch so festgelegt worden, daß dieses Ziel möglichst erreicht wird. Während der Implementierungsphase der Algorithmen sind die Energieparameter λ_A, λ_P sowie λ_Q nach dem folgenden Verfahren festgelegt worden. Dazu sind eine Vielzahl von zufälligen Transformationen auf die gegebene Molekülkonformation von Napap (PDB-Code: 1dwd) angewendet worden.[6] Mit Hilfe der entsprechenden Energiefunktionen aus Definition 3.2.2 sollte die strukturelle Identität des Moleküls wiedergefunden werden, d.h. die ausgeführten Transformationen sollten rückgängig gemacht werden.[7] Schließlich ist $\lambda_A = -4.0$, $\lambda_P = 2.0$ sowie $\lambda_Q = 2.75$ gesetzt worden. Es blieb noch, den Energieparameter $\lambda_T > 0$ so zu bestimmen, daß die obige Anforderung an die Energiefunktionen erfüllt wird, wobei durch $\lambda_T > 0$ die heuristische Bewertungsfunktion für die Torsionswinkel gewichtet wird.[8]
Dazu ist das heuristische Testverfahren aus Kapitel 4.2 verwendet worden. Das Verfahren besteht aus 770 Testbeispielen. Um diese 770 Testbeispiele zu berechnen, werden die entsprechenden Energiefunktionen aus Definition 3.2.2 minimiert, wobei das in Kapitel 4.2 ausgewählte Minimierungsverfahren verwendet wird, d.h es wird das inverse BFGS-Verfahren mit der Aufdatierungsformel (4.10) und der Armijoregel mit Aufweitung (Parameter $\beta = 0.35$) verwendet. Der zu optimierende Transformationsparameter der Energiefunktionen besteht aus dem Translationsvektor, dem Quaternionenvektor, dem globalen Rotationszentrum und den internen Rotationsparametern. In Tabelle 5.1 werden die durchschnittlichen Ergebnisse der 770 Optimierungsverfahren dargestellt. Für die Energieparameter

$$\lambda_T = 0.05, 0.10, 0.15, \ldots, 1.00, 1.10, \ldots, 1.90, 2.10, 2.30, 2.50$$

wird der durchschnittliche Alignmentindex sowie die durchschnittliche Bewertung der Torsionswinkel vor und nach den Optimierungen dargestellt, wobei S_E^0 bzw. S_E^* analog zu Gleichung (4.20) in Kapitel 4.2 definiert worden sind. Die durchschnittliche Bewertung der Torsionswinkel

[6] Napap gehört zur Gruppe der Thrombininihibitoren im Anhang A. In Abbildung 2.5 (Seite 15) ist Napap schematisch dargestellt worden.

[7] Man vergleiche hierzu den Abschnitt über das heuristische Testverfahren in Kapitel 4.2.

[8] Zu den Bewertungsfunktionen siehe den entsprechenden Abschnitt in Kapitel 2.2 und das Kapitel 2.7.2.

ist durch

$$s_{tor}^{0} := \frac{1}{770} \sum_{i=1}^{770} s_{tor}^{\mathcal{A}_i} \left(\left(\Theta_1(\mathsf{C}_{\mathcal{A}_i}), \ldots, \Theta_{n_{ft}[\mathcal{A}_i]}(\mathsf{C}_{\mathcal{A}_i}) \right), (0, \ldots, 0) \right),$$

sowie

$$s_{tor}^{*} := \frac{1}{770} \sum_{i=1}^{770} s_{tor}^{\mathcal{A}_i} \left(\left(\Theta_1(\mathsf{C}_{\mathcal{A}_i}), \ldots, \Theta_{n_{ft}[\mathcal{A}_i]}(\mathsf{C}_{\mathcal{A}_i}) \right), \left(\varphi_{i,1}^{*}, \ldots, \varphi_{i,n_{ft}[\mathcal{A}_i]}^{*} \right) \right),$$

definiert, wobei für $i = 1, \ldots, 770$ $s_{tor}^{\mathcal{A}_i}$ die heuristische Bewertungsfunktion für die Torsionswinkel des Moleküls $\mathcal{A}_i$ aus Definition 2.7.8 ist. Für $i = 1, \ldots, 770$ sind $\left(\varphi_{i,1}^{*}, \ldots, \varphi_{i,n_{ft}[\mathcal{A}_i]}^{*} \right)$ die internen Rotationsparameter des optimierten Transformationsparameters $u_{\mathcal{A}_i}^{*}$, den das Minimierungsverfahren als Ergebnis geliefert hat. Weiter wird die durchschnittliche Anzahl von Iterationen und die durchschnittliche Rechenzeit in Sekunden angegeben, die der Minimierungsalgorithmus gebraucht hat, um eine Energiefunktion zu minimieren. Ausserdem wird in Spalte Stopp die durchschnittliche Bewertung der Stoppkriterien angegeben. Als Stoppkriterien sind das Gradientenkriterium mit $\varepsilon := 10^{-7}$ und das Schrittkriterium mit $\varepsilon_S := eps^{0.9}$ verwendet worden, wobei *eps* die Maschinengenauigkeit des Computers ist. Das Optimierungsverfahren konnte jedoch nur maximal 100 Iterationen ausführen. Zur Bewertung der Stoppkriterien siehe man die Erläuterungen zur Tabelle 4.2 (Seite 105).

Auf Grund der Tabelle 5.1 habe ich mich dazu entschieden, den Energieparameter λ_T auf 0.1 zu setzen. Zwar wird der optimale Durchschnitt von 0.654 nicht erreicht, aber im Gegensatz zum Wert 0.05 wird auch die heuristische Bewertungsfunktion für die Torsionswinkel minimiert. Wie man aus der Tabelle 5.1 sehen kann, ist der durchschnittliche Alignmentindex bei allen anderen Werten kleiner als 0.63. Die Energieparameter können jederzeit vom Anwender genauso wie die Energiekoeffizienten modifiziert werden.

λ_T	$\mathsf{S}_{\mathsf{E}}^0$	$\mathsf{S}_{\mathsf{E}}^*$	s_{tor}^0	s_{tor}^*	Schritte	Sec	Stopp
0.00	0.102	0.654	0.000	0.000	76.81	31.69	0.34
0.05	0.102	0.639	0.336	0.368	70.20	30.32	0.34
0.10	0.102	0.631	0.336	0.322	66.10	29.74	0.38
0.15	0.102	0.628	0.336	0.292	64.50	28.98	0.39
0.20	0.102	0.619	0.336	0.269	61.41	28.51	0.41
0.25	0.102	0.615	0.336	0.250	60.37	28.16	0.42
0.30	0.102	0.615	0.336	0.236	59.40	27.89	0.42
0.35	0.102	0.609	0.336	0.225	57.39	27.19	0.44
0.40	0.102	0.612	0.336	0.211	56.74	27.60	0.43
0.45	0.102	0.608	0.336	0.200	55.67	26.64	0.44
0.50	0.102	0.605	0.336	0.194	55.73	27.26	0.46
0.55	0.102	0.604	0.336	0.190	54.69	26.97	0.45
0.60	0.102	0.600	0.336	0.181	53.77	26.54	0.46
0.65	0.102	0.604	0.336	0.174	53.39	26.14	0.46
0.70	0.102	0.605	0.336	0.171	51.78	25.00	0.47
0.75	0.102	0.598	0.336	0.169	51.91	25.83	0.47
0.80	0.102	0.598	0.336	0.161	50.76	21.40	0.47
0.85	0.102	0.606	0.336	0.160	50.98	21.68	0.46
0.90	0.102	0.599	0.336	0.156	50.84	24.87	0.48
0.95	0.102	0.601	0.336	0.152	50.85	24.85	0.48
1.00	0.102	0.606	0.336	0.150	49.27	23.89	0.49
1.10	0.102	0.597	0.336	0.143	48.59	24.24	0.48
1.20	0.102	0.599	0.336	0.135	49.26	23.92	0.48
1.30	0.102	0.599	0.336	0.136	48.52	22.89	0.49
1.40	0.102	0.593	0.336	0.131	48.00	23.28	0.49
1.50	0.102	0.593	0.336	0.129	47.43	22.00	0.50
1.60	0.102	0.590	0.336	0.125	47.84	23.01	0.49
1.70	0.102	0.591	0.336	0.123	46.33	21.82	0.51
1.80	0.102	0.586	0.336	0.121	46.83	22.22	0.49
1.90	0.102	0.590	0.336	0.120	45.63	21.69	0.50
2.10	0.102	0.590	0.336	0.117	44.62	20.71	0.51
2.30	0.102	0.585	0.336	0.113	45.45	21.63	0.49
2.50	0.102	0.579	0.336	0.109	45.59	20.83	0.51

Tabelle 5.1: Kalibrierung des Energieparameters λ_T in der Energiefunktion aus Definition 3.2.2.

Kapitel 6

Anwendungen und Ergebnisse

In diesem Kapitel werden Anwendungen und deren Ergebnisse dargestellt und diskutiert. Eine zentrale Anwendung ist die *flexible* Nachoptimierung von FlexS generierten Konformationen,[1], im Gegensatz zu Konformationen, die nur rigide nachoptimiert werden.[2] Die flexible Nachoptimierung ist der Anstoß für diese Arbeit gewesen.

6.1 Nachoptimierung

Die Liganden der Gruppen Dihydrofolatreduktase, HIV-Protease und Thrombin werden als Beispiele verwendet. Die Liganden werden mit dem PDB-Code bezeichnet.[3] Zu den Aufgaben der Liganden siehe man im einzelnen die Beschreibungen der Gruppen in Anhang A.
Das Programm FlexS verwendet jeden Liganden einer Gruppe einmal als Referenzmolekül und berechnet jeweils für alle Liganden der Gruppe mögliche Konformationen, indem die Moleküle auf das Referenzmolekül überlagert werden. Diese Konformationen werden nun mit Hilfe der entsprechenden Energiefunktionen aus Definition 3.2.2 nachoptimiert. Der zu optimierende Transformationsparameter der Energiefunktionen besteht aus dem Translationsvektor, dem Quaternionenvektor, dem globalen Rotationszentrum und den internen Rotationsparametern.[4] Daher ist die Dimension der jeweiligen Optimierungsaufgabe durch 10 plus Anzahl der Bindungen, an denen der Torsionswinkel variiert werden kann, bestimmt.[5] Um die entsprechenden Energiefunktionen zu minimieren, ist das inverse BFGS-Verfahren mit der Aufdatierungsformel (4.10) und der Armijoregel mit Aufweitung (Parameter $\beta = 0.35$) verwendet worden. Als Stoppkriterien werden dieselben Kriterien verwendet wie im Abschnitt über das heuristische

[1] *Flexibel* meint hier, daß interne Rotationsparameter bzgl. Drehachsen, sprich Bindungen, an denen der Torsionswinkel variiert werden kann, ebenfalls optimiert werden. D.h. $B_{FT}^{A} \neq \emptyset$.

[2] Man vergleiche hierzu die Diplomarbeit von C. Hiller [32]

[3] Protein Data Bank: [5],[49].

[4] Siehe die Definitionen 2.4.27 und 3.2.2.

[5] In Anhang A werden für die Moleküle unter anderem die Anzahl der Bindungen, an denen der Torsionswinkel variiert werden kann, angegeben.

Testverfahren in Kapitel 4.2. Dabei sind alle Rechnungen auf Computern vom Typ SUN-Ultra-30 Workstation, die mit einem Einzelprozessor (296-MHz) und 128 MB Hauptspeicher ausgestattet sind, ausgeführt worden. Die Bilder sind mit dem graphischen Tool FlexsV von FlexS erzeugt worden.
Indem die Testmoleküle (TM) bzgl. der Referenzmoleküle (RM) überlagert werden, berechnet das Programm FlexS mögliche Konformationen der Testmoleküle. Dabei numeriert FlexS diese Konformationsvorschläge, wobei die Reihenfolge durch den von FlexS benutzten Energiescore bestimmt wird. Als Molekülkonformation für das Referenzmolekül wird die Kristallstruktur verwendet, d.h. die Konformation, die man in der jeweiligen Rezeptortasche gemessen hat. Als Molekülkonformation für die Testmoleküle wird eine minimierte Konformation benutzt.[6] Alle Konformationen, die von FlexS berechnet worden sind, sind durch die entsprechenden Energiefunktion aus Definition 3.2.2 nachoptimiert worden. Um die Ergebnisse zu präsentieren, werden pro Gruppe jeweils drei Tabellen angegeben. In der ersten Tabelle werden pro Kombination von Referenzmolekül und Testmolekül die von FlexS berechneten Konformationen angegeben, die jeweils den besten Alignmentindix und den besten RMS-Abstand zu der Kristallstruktur haben. Es seien $M_1 := \{a_1, \ldots, a_n\} \subset \mathbb{R}^3$ und $M_2 := \{b_1, \ldots, b_n\} \subset \mathbb{R}^3$ zwei Koordinatenmengen. Der RMS-Abstand zweier Mengen M_1 und M_2 ist defeniert durch

$$RMS(M_1, M_2) := \sqrt{\frac{1}{n} \sum_{1}^{n} \|a_i - b_i\|_2^2}$$

Der RMS-Abstand ist ein objektives Maß, um zu überprüfen, inwieweit FlexS bzw. die flexible Nachoptimierung in der Lage gewesen sind, die jeweilige Kristallstruktur vorherzusagen. In der zweiten Tabelle werden die Konformationen dargestellt, die nach der Optimierung den besten Alignmentindex haben, während in der dritten Tabelle die Konformationen angeben werden, die nach der Optimierung den besten RMS-Abstand zu der jeweiligen Kristallstruktur besitzen. Dabei sind die Tabellen wie folgt aufgebaut: zuerst wird der Alignmentindex angegeben. Dann kommt der RMS-Abstand zur jeweiligen Kristallstruktur und zum Schluß die Nummer der Konformation. Falls nach der Optimierung die jeweilige Konformation mit dem besten Alignmentindex bzw. RMS-Abstand einen besseren Wert hat als die entsprechende Konformation aus der ersten Tabelle, so wird der jeweilige Wert unterstrichen. In der zweiten und dritten Tabelle werden die Ergebnisse vor (FLEXS) und nach der Optimierung (NOPT) angegeben. Falls FlexS nicht in der Lage gewesen ist, für eine Kombination Konformationen zu berechnen, so wird dies mit einem X in den Tabellen angezeigt.

Dihydrofolatreduktase:

Die Gruppe Dihydrofolatreduktase besteht aus den Molekülen 1dhf und 4dfr. Damit werden insgesamt 4 Kombinationen von Referenzmolekül und Testmolekül betrachtet. Das jeweilige Optimierungsproblem hat die Dimension 22 bzw. 23. Wie man aus den Tabellen 6.1 und 6.2 sehen kann, sind die Konformationen so nachoptimiert worden, daß die jeweilige nachoptimierte Konformation mit dem besten Alignmentindex einen besseren Wert besitzt als die entsprechende Konformation vor der Optimierung. D.h. in allen Kombinationen von Referenz- und

[6] Zu Einzelheiten siehe man die Arbeit von C.Lemmen [41].

RM\TM	1dhf		4dfr	
	AI	RMS	AI	RMS
1dhf	0.922	0.880	0.882	0.678
	1.57	0.60	2.89	2.35
	3	7	1	16
4dfr	0.923	0.470	0.957	0.904
	2.69	2.35	1.56	0.83
	1	22	2	4

Tabelle 6.1: Dihydrofolatreduktase: Beste Ergebnisse von FlexS

RM\TM	1dhf		4dfr	
	FLEXS	NOPT	FLEXS	NOPT
1dhf	0.911	0.985	0.878	0.959
	0.57	0.51	2.88	2.80
	1		2	
4dfr	0.919	0.953	0.904	0.984
	2.72	2.63	0.83	0.77
	4		4	

Tabelle 6.2: Dihydrofolatreduktase: Bester Alignmentindex nach der Optimierung

RM\TM	1dhf		4dfr	
	FLEXS	NOPT	FLEXS	NOPT
1dhf	0.524	0.985	0.678	0.908
	1.19	0.23	2.35	2.16
	19		16	
4dfr	0.470	0.872	0.757	0.984
	2.35	2.56	0.98	0.74
	22		13	

Tabelle 6.3: Dihydrofolatreduktase: Bester RMS-Abstand nach der Optimierung

Testmolekül konnte der Alignmentindex insgesamt verbessert werden. Die Verbesserung des Alignmentindizes in Tabellen 6.2 gegenüber der Tabelle 6.1 liegt im Bereich von 0.027 bis 0.077, wobei der Alignmentindex der Kombination 1dhf/4dfr (RM/TM) um 0.077 verbessert worden ist. Im einzelnen ist der Alignmentindex gegenüber dem Wert vor der Nachoptimierung in allen Kombinationen aus den Tabellen 6.2 und 6.3 verbessert worden. Bis auf die Kombination von 4dfr/1dhf in Tabelle 6.3 konnte auch der RMS-Abstand insgesamt verbessert werden. Die Verbesserung des RMS-Abstands in Tabelle 6.3 gegenüber der Tabelle 6.1 liegt im Bereich von 0.09 bis 0.37 Angstrom (Å), wobei der RMS-Abstand der Kombination 1dfh/1dfh um 0.37 Å verbessert worden ist.[7] In allen Kombinationen aus Tabelle 6.2 ist der RMS-Abstand gegenüber dem Wert vor der Optimierung verbessert worden und die Verbesserung liegt im Bereich von 0.06 bis 0.09 Å. In drei Kombinationen aus Tabelle 6.3 ist der RMS-Abstand gegenüber dem Wert vor der Optimierung verbessert worden. Dabei liegt die Verbesserung im Bereich von 0.19 bis 0.96 Å, wobei der RMS-Abstand der Kombination von 1dhf/1dhf um 0.96 Å gegenüber dem Wert vor der Optimierung verbessert worden ist. In Bild 6.1 und 6.2 wird das Referenzmolekül 1dhf und die Konformation 16 des Testmoleküls 4dfr dargestellt. Die Konformation 16 hatte vor der Optimierung den besten RMS-Abstand zur Kristallstruktur gehabt. Die nachoptimierte Konformation 16 hat immer noch den besten RMS-Abstand. Der RMS-Abstand ist durch die Nachoptimierung um 0.19 Å verbessert worden. Bild 6.1 zeigt die Konformation vor der Nachoptimierung, während Bild 6.2 die nachoptimierte Konformation darstellt. Die Gruppen, wo sich signifikant etwas gegenüber der von FlexS erzeugten Konformation getan hat, sind mit einem Pfeil markiert worden. In Bild 6.2 sieht man, wie die von FlexS erzeugte Konformation

[7] 1 Angstrom (Å) = 10^{-10} Meter.

an dem Referenzmolekül ausgerichtet worden ist. Das Optimierungsproblem des Alignments von 4dfr auf 1dhf hat die Dimension 23.

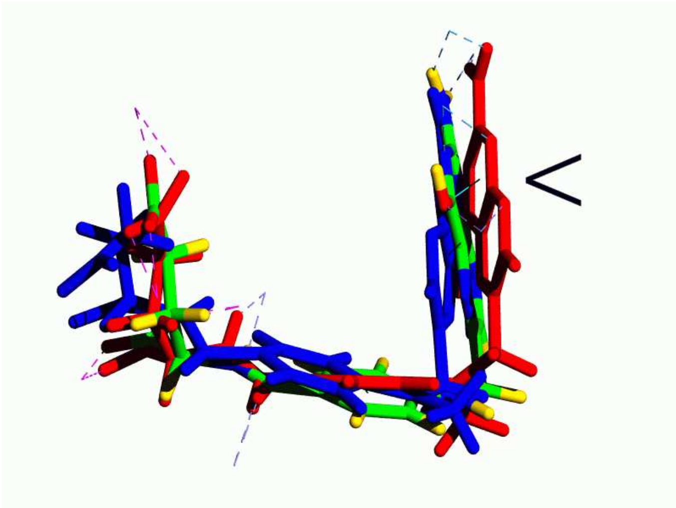


Abbildung 6.1: FlexS: Die Referenzstruktur 1dhf (grün-gelb), die von FlexS erzeugte Konformation 16 des Testmoleküls 4dfr (rot) und die Kristallstruktur von 4dfr (blau).

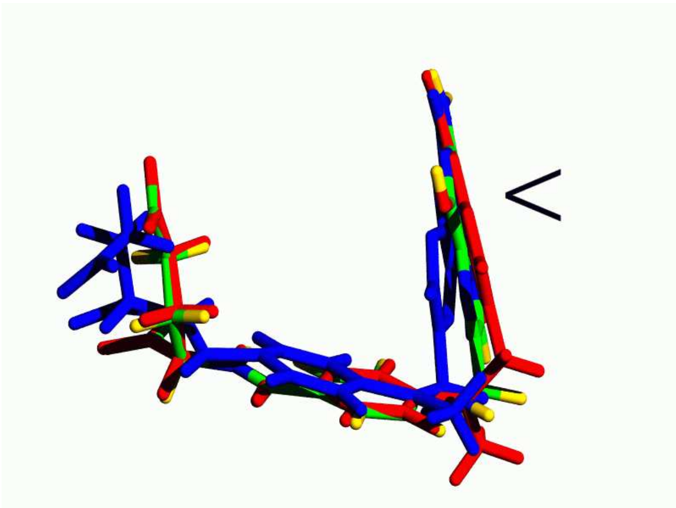


Abbildung 6.2: Nachoptimierung: Die Referenzstruktur 1dhf (grün-gelb), die nachoptimierte Konformation 16 des Testmoleküls 4dfr (rot) und die Kristallstruktur von 4dfr (blau).

Thrombin:

Die Gruppe Thrombin besteht aus den Molekülen 1dwc, 1dwd und 1ett. Damit werden insgesamt 9 Kombinationen von Referenzmolekül und Testmolekül betrachtet. Das jeweilige Optimierungsproblem hat die Dimension 22, 21 bzw. 19. Die Konformationen sind so nachoptimiert worden, daß die jeweilige nachoptimierte Konformation mit dem besten Alignmentindex einen

besseren Wert besitzt als die entsprechende Konformation vor der Optimierung. D.h. in allen Kombinationen von Referenz- und Testmolekül ist der Alignmentindex insgesamt verbessert worden. Die Verbesserung des Alignmentindizes in Tabellen 6.5 gegenüber der Tabelle 6.4 liegt im Bereich von 0.019 bis 0.21, wobei der Alignmentindex der Kombination 1dwd/1dwc um 0.21 verbessert worden ist. Im einzelnen ist der Alignmentindex gegenüber dem Wert vor der Nachoptimierung in alle Kombinationen aus den Tabellen 6.5 und 6.6 verbessert worden. Aber wie man aus den Tabellen 6.4 und 6.6 sehen kann, korrespondiert der RMS-Abstand nicht in allen Fällen mit dem Alignmentindex. So ist nur der RMS-Abstand der Kombinationen 1dwc/1ett, 1dwd/1dwc sowie 1dwd/1dwd insgesamt verbessert worden. Die Verbesserung des RMS-Abstands in Tabellen 6.6 gegenüber der Tabelle 6.4 liegt im Bereich von 0.06 bis 0.2 Å, wobei die Kombination von 1dwd/1dwd um 0.2 Å verbessert worden ist. Jedoch ist der RMS-Abstand gegenüber dem Wert vor der Optimierung in 7 von 9 Kombinationen aus der Tabelle 6.5 verbessert worden und liegt im Bereich von 0.05 bis 0.91 Å, wobei die Kombination von 1ett/1dwd um 0.91 Å verbessert worden ist. In 5 von 9 Kombinationen aus Tabelle 6.6 ist der RMS-Abstand gegenüber dem Wert vor der Optimierung verbessert worden und die Verbesserung liegt im Bereich von 0.03 bis 0.5 Å. Dabei ist der RMS-Abstand der Kombination von 1dwc/1ett um 0.5 Å gegenüber dem Wert vor der Nachoptimierung verbessert worden. In Bild 6.3 und 6.4 wird das Referenzmolekül 1dwc und die Konformation 59 des Testmoleküls 1ett dargestellt. Die nachoptimierte Konformation 59 hat den besten RMS-Abstand zur Kristallstruktur von 1ett. Dabei ist der RMS-Abstand um 0.07 Å besser als der entsprechende Wert der Konformation 60, die vor der Optimierung den besten RMS-Abstand besaß, und um 0.5 Å besser als der entsprechende Wert vor der Nachoptimierung. Bild 6.3 zeigt die Konformation vor der Nachoptimierung, während Bild 6.4 die nachoptimierte Konformation darstellt. Die Gruppen, wo sich signifikant etwas gegenüber der von FlexS erzeugten Konformation getan hat, sind mit einem Pfeil markiert worden. In Bild 6.4 sieht man, wie die von FlexS erzeugte Konformation an dem Referenzmolekül ausgerichtet worden ist. Das Optimierungsproblem des Alignments von 1ett auf 1dwc hat die Dimension 19.
Die Kombination 1dwc/1ett ist ein gutes Beispiel, um darauf hinzuweisen, daß die der RMS-Abstand nicht ganz unproblematisch ist. Denn der RMS-Abstand ist nicht in der Lage, Molekülsymmetrien zu berücksichtigen. Der schwarz Pfeil in Abbildung 6.4 zeigt auf das Ringsystem in Konformation (rot) von 1ett. Würde man diese Ringszstem, um 180 Grad drehen, so bliebe die Konformation und das Alignment erhalten, nur der RMS-Abstand würde seinen Wert ändern. Dies sollte man immer berücksichtigen.

HIV-Protease:

Die Gruppe HIV-Protease besteht aus den Molekülen 1hiv, 1hos, 1ivp, 1ivq, 2mip, 4hvp, 4phv, 5hvp, 8hvp und 9hvp. Damit werden insgesamt 100 Kombinationen von Referenzmolekül und Testmolekül betrachtet. Die Dimension des jeweiligen Optimierungsproblems liegt zwischen 27 und 60, wie man aus der Beschreibung der Gruppe im Anhang A sehen kann. Die Konformationen sind so nachoptimiert worden, daß die jeweilige nachoptimierte Konformation mit dem besten Alignmentindex einen besseren Wert besitzt als die entsprechende Konformation vor der

RM\TM	1dwc		1dwd		1ett	
	AI	RMS	AI	RMS	AI	RMS
1dwc	0.916	0.898	0.375	0.214	0.385	
	0.82	0.69	6.43	3.75	1.82	
	13	1	19	65	60	
1dwd	0.433	0.321	0.932	0.902	0.726	0.508
	4.31	3.49	1.91	0.56	4.16	2.27
	16	79	128	132	105	4
1ett	0.445	0.421	0.530	0.413	0.955	0.834
	3.82	1.88	5.27	4.25	1.47	1.32
	25	24	145	3	1	34

Tabelle 6.4: Thrombin: Beste Ergebnisse von FlexS

RM\TM	1dwc		1dwd		1ett	
	FLEXS	NOPT	FLEXS	NOPT	FLEXS	NOPT
1dwc	0.898	0.963	0.288	0.477	0.229	0.479
	0.69	0.74	6.24	6.35	2.94	2.35
	1		70		5	
1dwd	0.433	0.643	0.888	0.979	0.689	0.797
	4.31	3.92	1.02	0.93	4.13	3.46
	16		5		106	
1ett	0.442	0.587	0.446	0.689	0.955	0.974
	2.58	2.95	6.58	5.67	1.47	1.41
	1		47		1	

Tabelle 6.5: Thrombin: Bester Alignmentindex nach der Optimierung

Optimierung. D.h. in allen Kombinationen von Referenz- und Testmolekül ist der Alignmentindex insgesamt verbessert worden. Die Verbesserung des Alignmentindizes in den Tabellen 6.9 und 6.10 gegebüber den Tabellen 6.7 und 6.8 liegt im Bereich von 0.036 bis 0.385, wobei der Alignmentindex der Kombination 9hvp/4hvp um 0.385 verbessert worden ist. Im einzelnen ist der Alignmentindex gegenüber dem Wert vor der Nachoptimierung in alle Kombinationen aus den Tabellen 6.9 bis 6.12 verbessert worden. Jedoch konnte nur in der Hälfte aller Fällen (39 von 68) der RMS-Abstand insgesamt verbessert werden, wie man aus den Tabellen 6.11 und 6.12 sehen kann. Die Verbesserung des RMS-Abstands in den Tabellen 6.11 und 6.12 gegebüber den Tabellen 6.7 und 6.8 liegt im Bereich von 0.02 bis 1.16 Å, wobei der RMS-Abstand der Kombination 4hvp/4phv um 1.16 Å verbessert worden ist. Im einzelnen ist jedoch der RMS-Abstand gegenüber dem Wert vor der Nachoptimierung in 49 von 68 Kombinationen aus den Tabellen 6.11 und 6.12 verbessert worden. Dabei liegt die Verbesserung des RMS-Abstands gegenüber dem Wert vor der Optimierung im Bereich vom 0.005 bis 1.27 Å, wobei der RMS-Abstand der Kombination 4hvp/4phv um 1.27 Å verbessert worden ist. In 34 von 68 Kombinationen aus den Tabellen 6.9 und 6.10 ist der RMS-Abstand gegenüber dem Wert vor der Optimierung verbessert worden und die Verbesserung liegt im Bereich von 0.005 bis 0.87 Å, wobei die Kombination von 4hvp/1ivp um 0.87 Å verbessert worden ist. In der Kombination von 1hos/2mip hat die Konformation mit dem besten Alignmentindex auch den besten RMS-Abstand zur Kristallstruktur des entsprechenden Testmolekül 2mip. Auch in der Kombination von 9hvp/1hiv hat die Kon-

RM\TM	1dwc		1dwd		1ett	
	FLEXS	NOPT	FLEXS	NOPT	FLEXS	NOPT
	0.821	0.963	0.214	0.281	0.363	0.473
1dwc	0.72	0.74	3.75	4.95	2.25	1.75
	4		39		59	
	0.321	0.443	0.902	0.979	0.592	0.695
1dwd	3.49	3.43	0.56	0.36	2.47	2.44
	79		132		152	
	0.408	0.512	0.413	0.674	0.955	0.974
1ett	2.15	2.56	4.33	4.36	1.47	1.41
	6		4		1	

Tabelle 6.6: Thrombin: Bester RMS-Abstand nach der Optimierung

formation mit dem besten Alignmentindex den besten RMS-Abstand zur Kristallstruktur von 1hiv. In Bild 6.5 und 6.6 wird das Referenzmolekül 4hvp und die Konformation 17 des Testmoleküls 2mip dargestellt. Die nachoptimierte Konformation 17 hat den besten RMS-Abstand zur Kristallstruktur von 2mip. Dabei ist der RMS-Abstand um 0.7 Å besser als der entsprechende Wert der Konformation 32, die vor der Optimierung den besten RMS-Abstand besaß, und um 0.85 Å besser als der entsprechende Wert vor der Nachoptimierung. Die Konformation 32 hat den RMS-Abstand 3.05 Å, während die optimierte Konformation 17 den RMS-Abstand 2.35 Å hat. Bild 6.5 zeigt die Konformation vor der Nachoptimierung, während Bild 6.6 die optimierte Konformation darstellt. Die Gruppen, wo sich signifikant etwas gegenüber der von FlexS erzeugten Konformation getan hat, sind mit einem Pfeil markiert worden. In Bild 6.6 sieht man, wie die von FlexS erzeugte Konformation an dem Referenzmolekül ausgerichtet worden ist. Das Optimierungsproblem des Alignments von 2mip auf 4hvp hat die Dimension 46. In Bild 6.7 und 6.8 wird das Referenzmolekül 4hvp und die Konformation 39 des Testmoleküls 4phv dargestellt. Die optimierte Konformation 39 hat den besten RMS-Abstand zur Kristallstruktur von 4phv. Dabei ist der RMS-Abstand um 1.16 Å besser als der entsprechende Wert der Konformation 39, die vor der Optimierung den besten RMS-Abstand besaß, und um 1.27 Å besser als der entsprechende Wert vor der Nachoptimierung. Bild 6.7 zeigt die Konformation vor der Nachoptimierung, während Bild 6.8 die nachoptimierte Konformation darstellt. Die Gruppen, wo sich signifikant etwas gegenüber der von FlexS erzeugten Konformation getan hat, sind mit einem Pfeil markiert worden. Der Ring der Konformation 39 ist durch die Nachoptimierung auf das entsprechende Ringsystem in der Kristallstruktur des Testmoleküls 4phv zu bewegt worden. In Bild 6.8 sieht man, wie die von FlexS erzeugte Konformation an dem Referenzmolekül ausgerichtet worden ist. Das Optimierungsproblem des Alignments von 4phv auf 4hvp hat die Dimension 27. Ein weiterer Effekt der Nachoptimierung ist es, daß man die Anzahl der Konformationen, die pro Kombination von FlexS berechnet worden sind, verringern kann. Indem das Testmolekül 4phv bzgl. dem Referenzmolekül 4hvp überlagert worden ist, hat FlexS 53 verschiedene Konformationen für 4phv berechnet. Nach der Optimierung blieben 38 verschiedene Konformationen übrig.

RM\TM	1hiv		1hos		1ivp		1ivq		2mip	
	AI	RMS	AI	RMS	AI	RMS	AI	RMS	AI	RMS
1hiv	0.841		0.347	0.333	0.839	0.218	0.584	0.579	0.087	0.087
	0.93		12.09	2.50	12.50	9.53	2.09	1.66	6.74	6.09
	1		19	58	1	50	1	39	2	42
1hos	X		0.945	0.938	0.232	0.230	0.383	0.290	0.134	0.133
			1.65	1.40	11.97	11.39	1.79	1.71	11.10	10.69
			7	4	1	24	16	43	3	12
1ivp	0.916	0.776	0.405	0.370	0.860	0.841	0.538	0.209	0.204	0.115
	12.60	12.33	12.45	2.10	1.81	0.84	11.34	10.04	11.65	7.45
	7	64	1	21	3	2	59	94	62	101
1ivq	X		0.326	0.319	X		0.922	0.856	X	
			11.66	11.11			1.43	1.28		
			1	44			10	37		
2mip	X		0.131	0.099	X		0.152	0.127	0.958	0.953
			5.09	4.46			7.37	6.65	0.69	0.67
			59	2			2	97	1	8
4hvp	0.236	0.224	X		0.224	0.203	0.126	0.125	0.509	0.507
	4.74	4.55			12.37	11.84	10.12	9.79	3.28	3.05
	15	26			4	19	8	16	1	32
4phv	X		X		X		X		X	
5hvp	X		0.261	0.247	X		0.204	0.180	0.222	0.194
			3.25	2.60			9.48	9.09	12.63	12.45
			44	50			8	93	2	54
8hvp	X		0.172	0.146	0.352	0.284	0.319	0.301	0.482	0.467
			2.85	2.26	13.08	12.85	2.81	2.72	6.74	6.47
			47	67	3	1	58	88	1	94
9hvp	0.393		0.654	0.652	0.359	0.339	0.471	0.433	X	
	3.97		2.11	1.84	2.89	2.68	2.73	2.40		
	2		41	43	17	3	42	34		

Tabelle 6.7: HIV-Protease: Beste Ergebnisse von FlexS

RM\TM	4hvp		4phv		5hvp		8hvp		9hvp	
	AI	RMS	AI	RMS	AI	RMS	AI	RMS	AI	RMS
1hiv	0.218	0.213	0.079	0.063	0.293	0.286	X		0.363	
	7.46	7.34	8.18	7.94	11.11	10.92			2.79	
	7	34	27	44	97	70			18	
1hos	X		0.346	0.314	0.286	0.177	0.168		0.658	0.490
			5.27	5.07	12.56	6.27	9.82		2.08	2.00
			6	49	71	106	1		1	64
1ivp	X		0.258	0.039	0.316	0.313	X		0.465	0.329
			10.52	5.93	8.78	8.51			12.33	3.11
			81	71	2	1			2	56
1ivq	X		0.310	0.297	0.366	0.365	X		0.334	0.291
			7.34	7.03	11.86	11.80			11.91	11.51
			63	67	43	41			1	62
2mip	0.513	0.509	0.131	0.075	0.245	0.174	X		0.139	0.136
	2.87	2.48	9.52	9.34	12.93	11.21			2.19	2.10
	14	34	49	72	1	55			22	40
4hvp	0.887	0.886	0.186	0.162	X		0.102	0.101	X	
	0.97	0.97	6.52	6.31			9.70	9.58		
	1	2	21	38			50	47		
4phv	X		0.907		X		X		X	
			1.19							
			61							
5hvp	0.325	0.203	X		0.908	0.908	X		0.374	0.281
	12.61	12.10			0.95	0.90			12.40	4.51
	5	35			1	2			74	87
8hvp	0.169	0.166	0.165	0.064	0.408	0.397	0.902	0.902	0.137	0.106
	11.82	11.68	10.74	9.19	13.15	12.79	0.77	0.75	12.36	8.55
	21	27	43	76	1	54	1	2	33	73
9hvp	0.062	0.056	0.263	0.222	0.378		X		0.890	0.880
	11.39	11.21	6.55	6.29	2.15				1.15	0.91
	1	27	6	8	36				1	18

Tabelle 6.8: HIV-Protease: Beste Ergebnisse von FlexS (Fortsetzung)

RM\TM	1hiv		1hos		1ivp		1ivq		2mip	
	FLEXS	NOPT	FLEXS	NOPT	FLEXS	NOPT	FLEXS	NOPT	FLEXS	NOPT
	0.841	0.976	0.347	0.556	0.835	0.955	0.584	0.675	0.087	0.321
1hiv	0.93	0.65	12.09	12.46	12.52	12.59	2.09	2.16	6.15	6.14
	1		19		20		1		54	
			0.913	0.981	0.231	0.541	0.370	0.492	0.133	0.348
1hos	X		12.35	12.30	11.79	11.34	2.23	3.14	10.97	10.51
			18		7		87		2	
	0.909	0.971	0.370	0.582	0.842	0.969	0.532	0.677	0.201	0.340
1ivp	12.66	12.62	2.10	2.77	0.84	0.72	11.32	11.57	11.71	11.11
	1		21		1		1		75	
			0.321	0.444			0.902	0.976		
1ivq	X		11.58	11.30	X		1.37	1.35	X	
			8				1			
			0.096	0.298			0.134	0.272	0.958	0.993
2mip	X		4.98	5.12	X		7.02	7.77	0.69	0.59
			11				110		1	
	0.224	0.417			0.219	0.419	0.121	0.379	0.508	0.614
4hvp	4.91	4.66	X		12.34	11.47	9.92	10.07	3.48	2.94
	1				2		5		22	
4phv	X		X		X		X		X	
			0.254	0.422			0.202	0.376	0.195	0.394
5hvp	X		3.10	3.72	X		9.59	10.44	13.09	12.51
			53				43		66	
			0.166	0.294	0.336	0.457	0.308	0.392	0.468	0.591
8hvp	X		2.86	2.92	13.12	13.14	3.00	3.29	6.54	6.76
			20		32		7		91	
	0.391	0.602	0.650	0.714	0.331	0.574	0.451	0.611		
9hvp	4.02	3.95	2.32	2.22	2.91	3.07	2.71	3.10	X	
	35		4		16		32			

Tabelle 6.9: HIV-Protease: Bester Alignmentindex nach der Optimierung

RM\TM	4hvp		4phv		5hvp		8hvp		9hvp	
	FLEXS	NOPT	FLEXS	NOPT	FLEXS	NOPT	FLEXS	NOPT	FLEXS	NOPT
	0.213	0.369	0.061	0.346	0.286	0.450			0.356	0.603
1hiv	7.34	7.27	8.14	8.82	10.92	11.00	X		3.10	3.63
	34		18		70				16	
			0.336	0.538	0.269	0.495	0.168	0.536	0.623	0.746
1hos	X		5.32	5.45	12.42	12.02	9.91	9.80	12.46	12.43
			18		27		2		30	
			0.253	0.429	0.312	0.447			0.464	0.630
1ivp	X		10.72	11.26	8.51	8.92	X		12.31	12.31
			82		1				1	
			0.307	0.533	0.362	0.552			0.329	0.542
1ivq	X		7.22	7.59	11.92	11.91	X		11.95	12.01
			66		7				29	
	0.512	0.652	0.127	0.319	0.245	0.572			0.136	0.342
2mip	3.08	3.01	9.49	10.06	12.96	13.05	X		2.11	2.58
	21		58		6				40	
	0.775	0.948	0.150	0.389			0.062	0.422		
4hvp	1.27	0.99	9.07	9.96	X		13.74	13.59	X	
	20		16				3			
			0.875	0.958						
4phv	X		1.23	1.13	X		X		X	
			1							
	0.318	0.510			0.908	0.989			0.365	0.509
5hvp	12.40	12.38	X		0.95	0.52	X		12.86	12.82
	14				1				43	
	0.166	0.443	0.064	0.292	0.397	0.577	0.902	0.980	0.111	0.339
8hvp	11.68	12.55	9.23	9.65	12.90	12.92	0.77	0.66	8.68	8.43
	27		73		59		1		88	
	0.056	0.447	0.259	0.472	0.372	0.514			0.890	0.971
9hvp	11.39	11.04	6.50	6.82	2.32	2.90	X		1.15	1.01
	4		5		48				1	

Tabelle 6.10: HIV-Protease: Bester Alignmentindex nach der Optimierung (Fortsetzung)

RM\TM	1hiv		1hos		1ivp		1ivq		2mip	
	FLEXS	NOPT	FLEXS	NOPT	FLEXS	NOPT	FLEXS	NOPT	FLEXS	NOPT
1hiv	0.838	0.976	0.333	0.524	0.220	0.440	0.581	0.671	0.087	0.295
	0.98	0.61	2.50	2.57	9.59	8.54	1.78	1.81	6.74	5.78
	11		58		372		6		2	
1hos	X		0.931	0.977	0.230	0.479	0.383	0.442	0.133	0.348
			1.41	1.37	11.86	11.31	1.79	1.98	10.97	10.51
			2		18		16		2	
1ivp	0.776	0.963	0.354	0.577	0.841	0.969	0.209	0.405	0.111	0.310
	12.33	12.45	2.28	2.57	0.84	0.72	10.04	9.65	7.46	7.56
	64		18		2		94		102	
1ivq	X		0.322	0.423	X		0.867	0.976	X	
			11.12	11.15			2.18	1.34		
			13				79			
2mip	X		0.105	0.232	X		0.130	0.260	0.953	0.993
			4.55	4.39			6.98	6.33	0.67	0.51
			35				131		8	
4hvp	0.236	0.385	X		0.223	0.415	0.121	0.377	0.509	0.613
	4.72	4.28			12.33	11.37	9.82	9.93	3.20	2.35
	20				10		15		17	
4phv	X		X		X		X		X	
5hvp	X		0.212	0.378	X		0.180	0.274	0.222	0.372
			2.98	2.87			9.36	9.15	12.63	12.23
			85				11		3	
8hvp	X		0.117	0.245	0.279	0.453	0.304	0.390	0.480	0.556
			2.46	2.48	12.87	12.70	2.90	2.90	6.66	6.32
			10		27		9		42	
9hvp	0.391	0.602	0.602	0.703	0.329	0.568	0.470	0.596	X	
	4.02	3.95	2.22	2.14	2.87	2.69	2.61	2.66		
	35		48		21		64			

Tabelle 6.11: HIV-Protease: Bester RMS-Abstand nach der Optimierung

RM\TM	4hvp		4phv		5hvp		8hvp		9hvp	
	FLEXS	NOPT	FLEXS	NOPT	FLEXS	NOPT	FLEXS	NOPT	FLEXS	NOPT
1hiv	0.213	0.361	−0.034	0.288	0.289	0.423	X		0.355	0.516
	7.39	6.86	8.37	7.36	11.26	10.47			2.89	3.06
	13		46		4				34	
1hos	X		0.325	0.536	0.177	0.407	0.163	0.519	0.649	0.737
			5.33	5.31	6.36	5.68	9.99	9.49	2.05	1.83
			20		102		16		21	
1ivp	X		0.039	0.321	0.313	0.443	X		0.326	0.547
			5.93	5.26	8.56	8.65			3.19	4.05
			71		5				52	
1ivq	X		0.306	0.528	0.366	0.486	X		0.292	0.522
			7.15	7.22	11.86	10.96			11.62	10.82
			65		43				43	
2mip	0.508	0.632	0.064	0.229	0.174	0.486	X		0.132	0.285
	2.95	2.55	9.39	9.10	11.21	12.50			2.13	2.04
	28		33		55				37	
4hvp	0.775	0.948	0.178	0.317	X		0.101	0.360	X	
	1.27	0.99	6.42	5.15			9.61	9.46		
	20		39				48			
4phv	X		0.875	0.958	X		X		X	
			1.23	1.13						
			1							
5hvp	0.205	0.438	X		0.908	0.989	X		0.267	0.447
	12.16	11.85			0.90	0.50			4.61	4.22
	36				4				85	
8hvp	0.169	0.410	0.064	0.277	0.397	0.576	0.902	0.980	0.108	0.336
	11.82	12.26	9.22	9.55	12.87	12.81	0.75	0.66	8.67	8.30
	21		74		31		2		84	
9hvp	0.055	0.404	0.222	0.452	0.378	0.506	X		0.876	0.971
	11.49	11.04	6.29	6.59	2.16	2.22			0.94	0.33
	31		8		46				33	

Tabelle 6.12: HIV-Protease: Bester RMS-Abstand nach der Optimierung (Fortsetzung)

Zusammenfassung der Ergebnisse:

Der Alignmentindex ist in allen Rechnungen der obigen Beispiele verbessert worden. Dadurch ist auch die relativ Orientierung (das Alignment) der Moleküle zueinander verbessert worden, wie man in den einzelnen Bilder gut sehen kann. Hier wird die Überlegenheit der flexiblen Nachoptimierung gegenüber der rigiden Nachoptimierung deutlich. Zwar kann auch die rigide Nachoptimierung das Alignment der Moleküle zueinander verbessern,[8] aber die rigide Nachoptimierung ist nicht in der Lage einzelne Molekülfragmente zu optimieren. Die flexible Nachoptimierung kann jedoch einzelne Molekülfragmente optimieren, wie aus den Bildern 6.2, 6.4, 6.6 und 6.8 deutlich wird.

Betrachtet man die Selbsttests in den 3 Gruppen (TeststmoleKül ist identisch mit dem Referenzmolekül), so ist in keinem Fall das mögliche globale Optimum des Alignmentindizes erreicht worden. In keinem Beispiel ist der Wert eins für den Alignmentindex reicht worden. Zum einem kann dies sich darin liegen, daß das verwendete Minimierungsverfahren ein lokales Verfahren ist.[9] Aber es können noch andere Gründe eine Rolle spielen: als Testmolekül wird eine minimierte Molekülkonformation verwendet. Daraus folgt, daß die Bindungswinkel und Bindungslängen sowie die Ringkonformationen in der Testkonformation nicht mit denen aus der Referenzstruktur übereinstimmen müssen. Da die Energiefunktionen aus Kapitel 3.2 und 3.3 Bindungslängen und Bindungswinkel unverändert lassen und Ringsysteme künstlich rigide gehalten werden, kann es durchaus sein, daß die strukturelle Identität durch Nachoptimierung nur approximiert aber nicht nicht wiedergefunden werden kann.

Dies sollte man ebenso beachten, wenn die Ergebnisse der anderen Beispiele untersucht.

Der Alignmentindex konnte in allen Beispielen verbessert werden. Der RMS-Abstand, auch wenn er nicht ganz unproblematisch ist, korrespondiert nur zum Teil mit Alignmentindex. Zwar konnte in der Hälfte aller Beispiele der RMS-Abstand im einzelnen wie insgesamt verbessert werden, doch die Verbesserungen sind meist nur sehr gering. In einigen wenigen Beispielen, wo die Verbesserung über einen Angstrom liegt, wie zum Beispiel 1ett/1dwd (RM/TM) oder 4hvp/4phv, beträgt der RMS-Abstand der entsprechend optimierten Testkonformation zur Kristallstruktur immer noch mindestens 5 Å. In nur einen Beispiel konnte der RMS-Abstand wirklich entscheidend verbessert werden: im Beispiel der Kombination 4hvp/2mip (HIV-Protease). Die beste von FlexS erzeugte Konformation hat einen RMS-Abstand von 3.05 Å, während die beste nachoptimierte Konformation einen RMS-Abstand von 2.35 Å hat, wie der Tabelle 6.11 entnehmen kann. Daraus folgt insgesamt, daß der RMS-Abstand nicht unbedingt mit dem Alignmentindex korrespondieren muß. Es gibt sogar Beispiele, wo sich der Alignmentindex und der RMS-Abstand überhaupt nicht miteinander korrespondieren. DFür die Kombination 1hiv/1ivp (HIV-Protease) aus Tabelle 6.9 gilt, daß die optimierte Konformation 20 des Testmoleküls 1ivp hat bzgl. der Referenzstruktur 1hiv den Alignmentindex 0.955, aber den RMS-Abstand 12.59 Å zur Kristallstruktur besitzt. In Bild 6.9 ist die optimierte Konformation 20 von 1ivp und die Referenzstruktur von 1hiv dargestellt. Das Alignment der optimierten Konformation bzgl. Referenzstruktur ist fast optimal. Nimmt man jedoch die Kristallstruktur hinzu (Bild 6.10), so

[8]Man vergleiche hierzu die Diplomarbeit von C. Hiller [32]

[9]Siehe hierzu die Bemerkung am Ende von Kapitel 4.2.

erkent man, daß sich der RMS-Abstand nur dann verbessern ließe, wenn zugleich das Alignment schlechter würde. Man vergleiche hierzu den entsprechenden Eintrag in Tabelle 6.11. Als Fazit läßt sich sagen, die Energiefunktionen aus Kapitel 3.2 sind in der Lage, das Alignment einer Konformation bzgl. einer Referenzstruktur zu verbessern. Besonders durch die Optimierung der internen Rotationsparmeter können einzelne Molekülfragmente besser orintiert werden. Dieses Ergebnis entspricht damit den Erwartungen, die man bzgl. der flexiblen Nachoptimierung hatte. Der RMS-Abstand korrespondiert nicht in jedem Fall mit dem Alignment. Denn das Modell der Überlagerung hat seine natürlichen Grenzen. Das Modell beruht darauf, daß Liganden, die in dieselbe Rezeptortasche binden, Ähnlichkeiten in ihren räumlichen Eigenschaften aufweisen. Trift dies aber nicht mehr oder nur zum Teil noch zu (siehe hierzu die Abbildung 6.13), so kann man zwar immer noch das Alignment berechnen, aber man kann nicht erwarten, daß die so berechnete Konformation mit der Konformation übereinstimmt, die in die entsprechende Rezeptortasche bindet. An den Tabellen 6.7 und 6.8 kann man sehen, daß auch FlexS mit dieser Grenze Schwierigkeiten hat.

6.2 Berechnung eines Alignments

Man kann die Energiefunktion aus der Definition 3.2.2 direkt dazu verwenden, das Alignment eines Moleküls bzgl. eines Referenzmoleküls zu berechnen. Dies kann z.B. dann notwendig sein, wenn FlexS nicht in der Lage gewesen ist, ein Alignment zu berechnen. In 32 Kombinationen von Referenz- und Testmolekül der Gruppe HIV-Protease konnte FlexS keine Konformationen für das Testmolekül berechnen.
Man wählt dazu ein flexibles Molekül $\mathcal{A}$ und ein Referenzmolekül $\mathcal{R}$ und optimiert das Alignment der Molekülkonformation von $\mathcal{A}$ bzgl. der starren Referenzstruktur von $\mathcal{R}$. Das Problem hierbei ist, eine geeignete Startkonformation des Moleküls $\mathcal{A}$ zu finden, denn die in dieser Arbeit verwendeten Minimierungsverfahren sind lokale Optimierungsverfahren. Lokale Optimierungsverfahren berechnen vom Startwert ausgehend das nächste lokale Optimum der Energiefunktion. D.h. das globale Optimum einer Energiefunktion wird nur dann berechnet, wenn der Startwert entsprechend gut gewählt worden ist.
Wie kann man geeignete Startkonformation für das Molekül $\mathcal{A}$ wählen? Im folgenden wird eine Möglichkeit beschrieben, um dieses Problem zu lösen. Dazu wählt man je k Atome des Moleküls $\mathcal{A}$ und des Moleküls $\mathcal{R}$. Diese Atome bilden paarweise ein Matching, das minimiert wird. D.h. die folgende Abstandsfunktion wird minimiert:

$$\min_{(t,q)\in\mathbb{R}^3\times\mathbb{R}^4\setminus\{0\}} \sum_{j=1}^{k} \|A(q)\left(a_{l_j}-c_s\right)+c_s+t-r_{l_j}\|_2^2,$$

wobei a_{l_j} bzw. r_{l_j} für $j=1,\dots,k$ die Atomkoordinaten der entsprechenden Atome von $\mathcal{A}$ bzw. $\mathcal{R}$ sind.[10] Dabei ist $A(q)$ mit $q\in\mathbb{R}^4\setminus\{0\}$ die Rotationsmatrix aus Gleichung (2.10) in Kapitel 2.4.1. c_s ist der Schwerpunkt der Atomkoordinaten $a_{l_1},\dots,a_{l_k}$, d.h. $c_s := (1/k)\sum_{j=1}^{k} a_{l_j}$. Indem man die obige Abstandsfunktion minimiert, erhält man optimale Parameter $(t^*,q^*)\in$

[10] Mit a_i werden sowohl das Atom als auch die zugehörigen Atomkoordinaten bezeichnet .

$\mathbb{R}^3 \mathrm{x} \mathbb{R}^4 \setminus \{0\}$. Nun wendet man die globale Transformationsabbildung T_g aus Definition 2.4.7 mit den Parametern (t^*, q^*, c_s) auf alle Atomkoordinaten des Moleküls $\mathcal{A}$ an, d.h. die globale Transformation $\mathsf{T}_g(t^*, q^*, c_s, \cdot) = A(q^*)\,(\cdot - c_s) + c_s + t^*$ wird angewendet. Die so transformierte Konformation von $\mathcal{A}$ wird dann als Startkonformation für die entsprechende Energiefunktion aus Definition 3.2.2 verwendet. Das Programm FlexS besitzt Routinen, um diese Abstandsfunktion zu minimieren.
Nun stellt sich die Frage, wie man geeignete Atome paarweise aus den Molekülen $\mathcal{A}$ und $\mathcal{R}$ auswählt. Entweder man besitzt das nötige Wissen, um geeignete Atom auszuwählen, oder man verwendet Hilfsprogramme, die geeignete Atome bzw. Atomgruppen vorschlagen, die zueinander gehören.

Eine andere Möglichkeit, eine geeignete Startkonformation zu bestimmen, besteht darin, die Energiefunktionen aus Definition 3.3.3 geeignet anzuwenden. Es seien $\mathcal{A}$ und $\mathcal{C}$ zwei Moleküle, die FlexS nicht gegenseitig überlagern konnte. Die Idee ist nun, daß man geeignete Molekülen $\mathcal{B}_1$ bis $\mathcal{B}_k$ hinzunimmt, wobei Molekül $\mathcal{A}$ bzgl. dem Molekül $\mathcal{B}_1$ überlagert werden kann. Für $i = 1, \dots, k-1$ kann das Molekül $\mathcal{B}_i$ bzgl. dem Molekül $\mathcal{B}_{i+1}$ überlagert werden. Schließlich kann das Molekül $\mathcal{B}_k$ bzgl. dem Molekül $\mathcal{C}$ überlagert werden. Indem man die entsprechende Energiefunktion aus Definition 3.3.3 minimiert, wird das gegenseitige Alignment der Moleküle $\mathcal{A}$, $\mathcal{B}_1$, $\dots$ $\mathcal{B}_k$ und $\mathcal{C}$ optimiert. Die Hoffnung besteht darin, daß man über diesen Umweg eine geeignete Startkonformatiom für $\mathcal{A}$ erhält.
Das Programm konnte zum Beispiel keine Konformationen von 4phv bzgl. der Kombination von 5hvp/4phv (RM/TM) noch Konformationen von 5hvp bzgl. der Kombination von 4phv/5hvp berechnen. Genauso verhält es sich mit 1hiv und 8hvp bzw. 1ivq und 2mip. Aus den Tabellen 6.7 und 6.8 sieht man, daß das Molokül 1hiv bzgl. dem Molekül 9hvp überlagert werden kann. Weiter kann 9hvp bzgl. 1hos sowie 1hos bzgl. 8hvp überlagert werden. Man nimmt die Konformation 2 von 1hiv, die Konformation 64 von 9hvp sowie die Konformation 67 von 1hos und die Kristallstruktur von 8hvp. In Bild 6.11 sind die von FlexS berechneten Konformationen von 1hiv, 9hvp, 1hos sowie von 8hvp dargestellt. Um nun das gegenseitige Alignment dieser vier Moleküle zu optimieren, wird die entsprechende Energiefunktion aus Definition 3.3.5 minimiert, wobei die Energiefunktion wie folgt modifiziert worden ist: da das Molekül 8hvp gröser als die drei anderen Moleküle ist, ist die globale Transformation für das Molekül 8hvp deaktiviert worden. D.h. der Transformationsoperator aus Definition 2.4.27 für das Molekül 8hvp besteht nur aus dem entsprechenden lokalen Rotationsoperator. Somit hat das Optimierungsproblem die Dimension 168. Um die entsprechende Energiefunktion zu minimieren, wird der Gauß-Seidel-artige Optimierungsalgorithmus 4.6 aus Kapitel 4.3 mit dem Minimierungsalgorithmus aus 4.2 verwendet. Durch die Optimierung ist der Wert von $\mathsf{S}_{\mathcal{A}_1,\dots,\mathcal{A}_4}$ aus der Gleichung (3.9) in Definition 3.3.3 von 0.177 auf 0.367 verbessert worden. In Bild 6.12 sind die optimierten Konformationen der vier Moleküle dargestellt worden. Die Gruppen, wo sich signifikant etwas gegenüber den von FlexS erzeugten Konformationen verändert hat, sind mit einem Pfeil markiert worden. In Abbildung 6.12 kann man sehr gut erkennen, däs alle vier Moleküle sich gegenseitig ausgerichtet haben. Alle vier Moleküle scheinen darüberhinaus eine gemeinsame Grundstruktur zu besitzen. Der schwarze Pfeil in der Mitte von Bild 6.12 zeigt auf diese Grundstruktur.

Die so berechnete Molekülkonformation von 1hiv wird mit $1hiv(2)^*$ bezeichnet und als Startkonformation für die Überlagerung von 1hiv auf 8hvp verwendet. Als Referenzstruktur für 8hvp wird die orientierte Kristallstruktur von 8hvp verwendet und mit $8hvp^*$ bezeichnet. Um die Überlagerung von $1hiv(2)^*$ bzgl. $8hvp^*$ zu berechnen, wird die entsprechende Energiefunktion $\mathsf{E}_{\mathcal{A}(1hiv(2)^*),\mathcal{R}(8hvp^*)}$ aus Definition 3.2.2 minimiert. Die Optimierungsaufgabe hat die Dimension 41. Das Ergebnis der Minimierung ist in der Tabelle 6.13 dargestellt. Der Alignmentindex

Funktion	Vor der Optimierung	Nach der Optimierung
Alignmantindex	0.311	0.403
RMS-Abstand	3.9377 Å	4.1297Å
Energiefunktion	-0.3699	-0.7402

Tabelle 6.13: Ergebnisse der Minimierung der Enerfunktion $\mathsf{E}_{\mathcal{A}(1hiv(2)^*),\mathcal{R}(8hvp^*)}$

der optimierten Konformation $1hiv(2)^*$ bzgl. der Referenzstruktur $8hvp^*$ hat sich verbessert. Dagegen hat sich der RMS-Abstand zur Kristallstruktur von 1hiv verschlechtert. Da die Konformation 2 von 1hiv aus der Kombination 9hvp/1hiv schon in der Tabelle 6.7 den RMS-Abstand von 3.97 Å zur Kristallstruktur gehabt hat, konnte man ncht erwarten daß der RMS-Abstand wirklich verbessert wird. Bild 6.13 zeigt die Konformationen von $1hiv(2)^*$ und $8hvp^*$ vor der Optimierung, sowie die Kristallstruktur von 1hiv, während Bild 6.14 die optimiert Konformation von $1hiv(2)^*$ darstellt. Die Gruppen, wo sich signifikant etwas getan hat, sind mit einem Pfeil markiert worden. Dieses Beispiel zeigt darüberhinaus sehr deutlich die Grenzen des Modells der Überlagerung. Der Doppelt-Secher-Ring der Kristallstruktur von 1hiv zeigt in den freien hinaus Raum. Weder FlexS noch die Energiefunktionen aus Kapitel 3.2 bzw. 3.3 sind in der Lage diese Struktur zu approximieren.

Auch wenn das obige Beispiel kein sehr erfolgreiches Beispiel gewesen ist, so ist an diesem Beispiel hoffentlich klar geworden, wie man die Energiefunktionen aus Kapitel 3.3 verwenden kann, um geeignete Startkonformationen zu generieren.

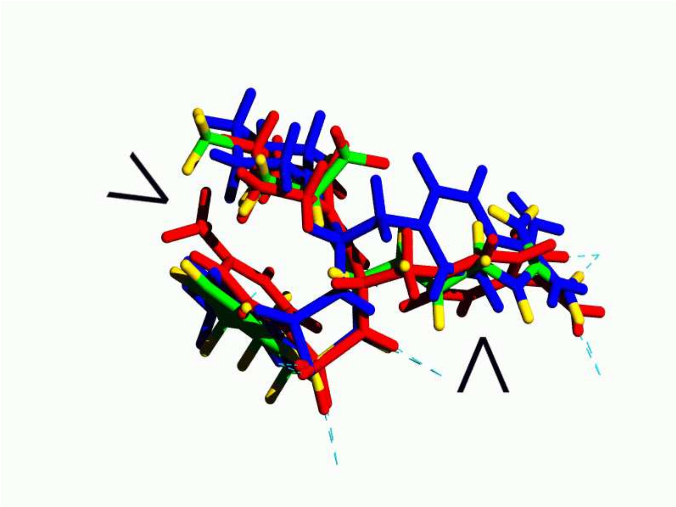


Abbildung 6.3: FlexS: Die Referenzstruktur von 1dwc (grün-gelb), die von FlexS erzeugte Konformation 59 des Testmoleküls 1ett (rot) und die Kristallstruktur von 1ett (blau).

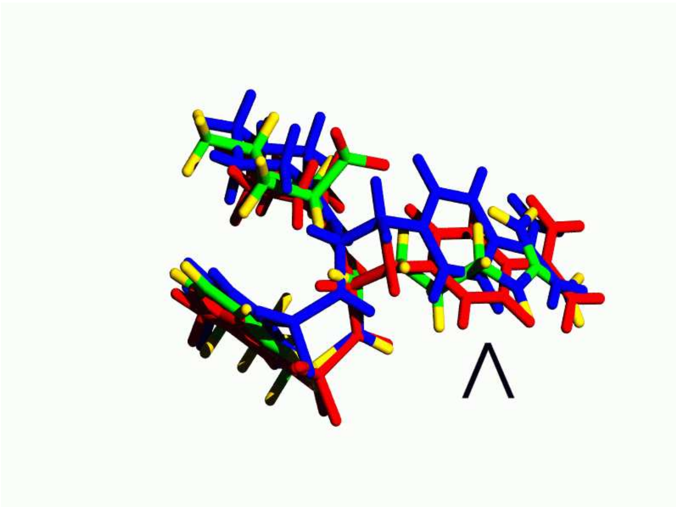


Abbildung 6.4: Nachoptimierung: Die Referenzstruktur von 1dwc (grün-gelb) die nachoptimierte Konformation 59 des Testmoleküls 1ett (rot) und die Kristallstruktur von 1ett (blau).

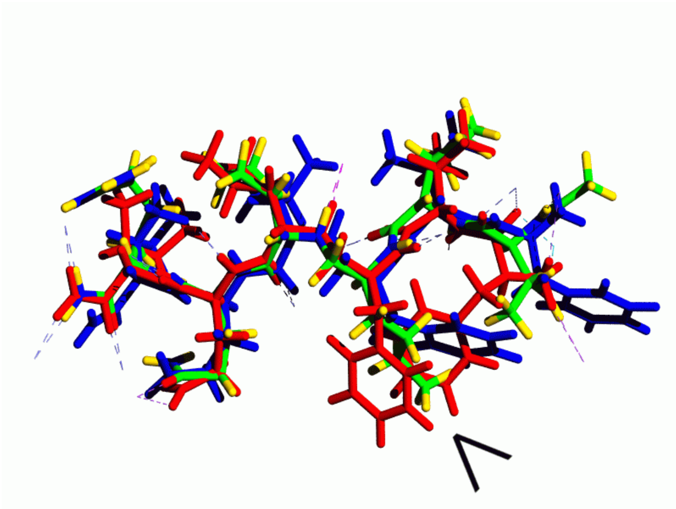


Abbildung 6.5: FlexS: Die Referenzstruktur von 4hvp (grün-gelb), die von FlexS erzeugte Konformation 17 des Testmoleküls 2mip (rot) und die Kristallstruktur von 2mip (blau).

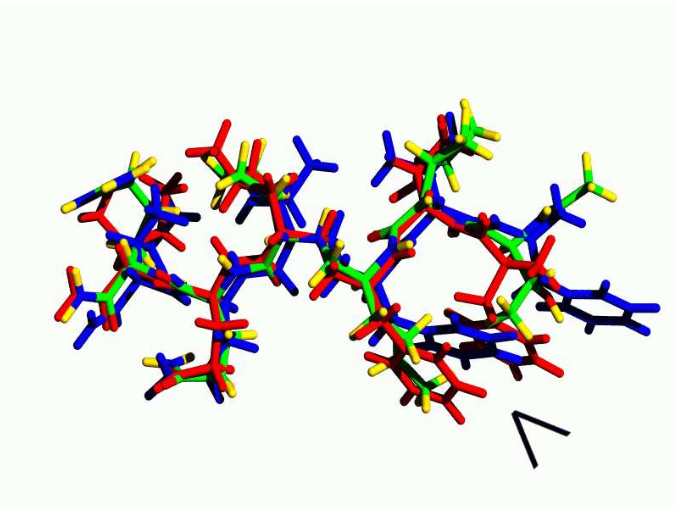


Abbildung 6.6: Nachoptimierung: Die Referenzstruktur von 4hvp (grün-gelb), die nachoptimierte Konformation 17 des Testmoleküls 2mip (rot) und die Kristallstruktur von 2mip (blau).

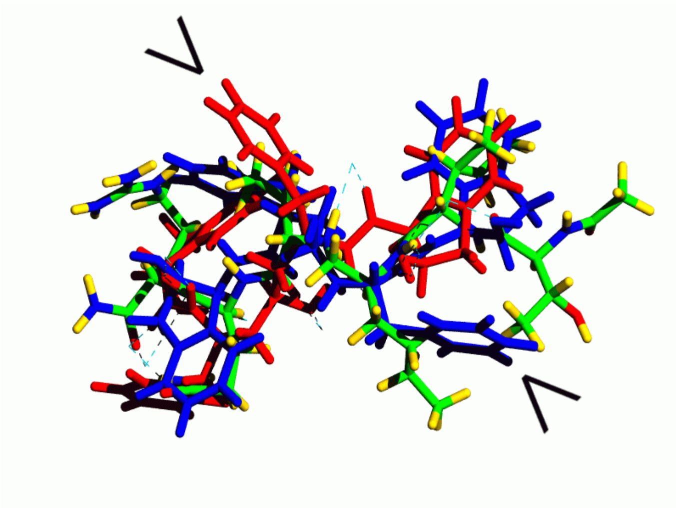


Abbildung 6.7: FlexS: Die Referenzstrukur von 4hvp (grün-gelb), die von FlexS erzeugte Konformation 39 des Testmoleküls 4phv (rot) und die Kristallstruktur von 4phv (blau).

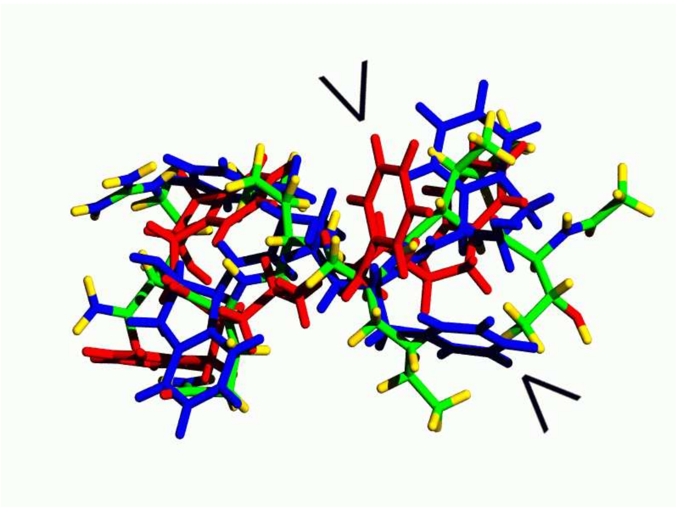


Abbildung 6.8: Nachoptimierung: Die Referenzstruktur von 4hvp (grün-gelb) die nachoptimierte Konformation 39 des Testmoleküls 4phv (rot) und die Kristallstruktur von 4phv (blau).

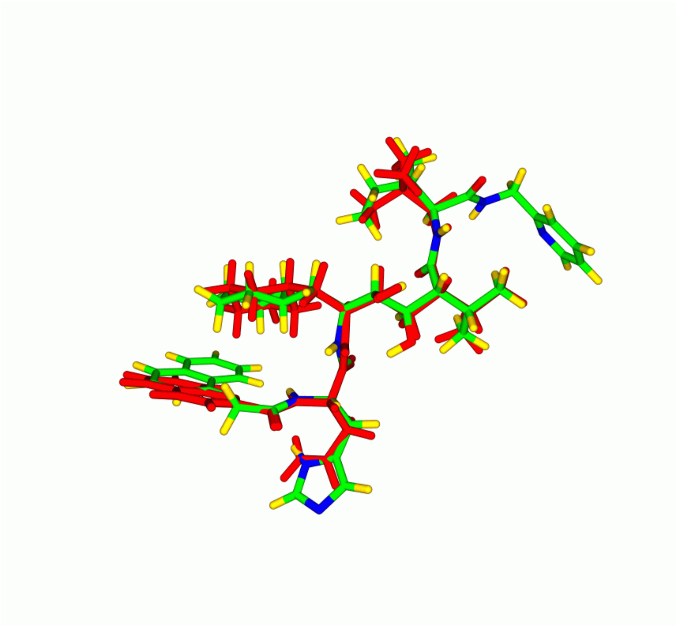


Abbildung 6.9: Die optimierte Konformation 20 von 1ivp (rot) und die Kristallstruktur 1hiv (grün-gelb).

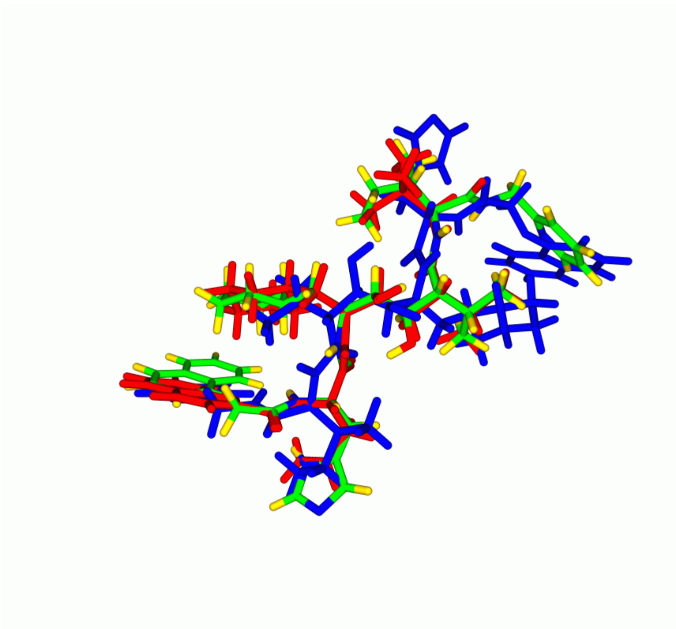


Abbildung 6.10: Die optimierte Konformation 20 von 1ivp (rot), die Kristallstruktur 1hiv (grün-gelb) und die Kristallstruktur von 1ivp (blau).

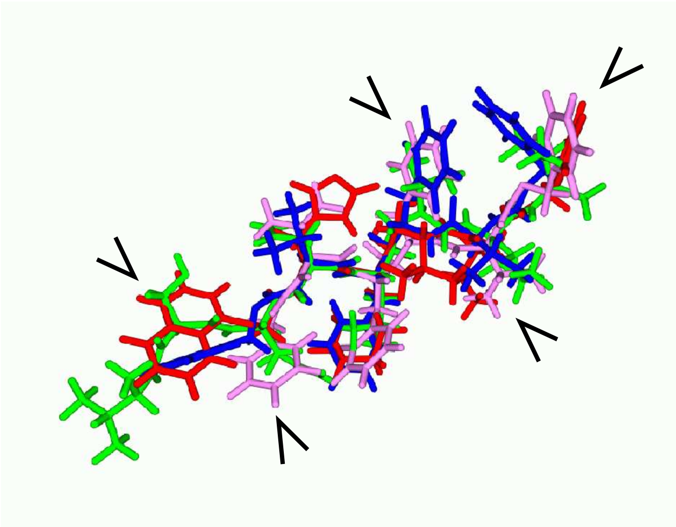


Abbildung 6.11: Die Konformation 2 von 1hiv (rot), die Konformation 64 von 9hvp (blau), die Konformation 67 von 1hos (violett) und die Kristallstruktur von 8hvp (grün).

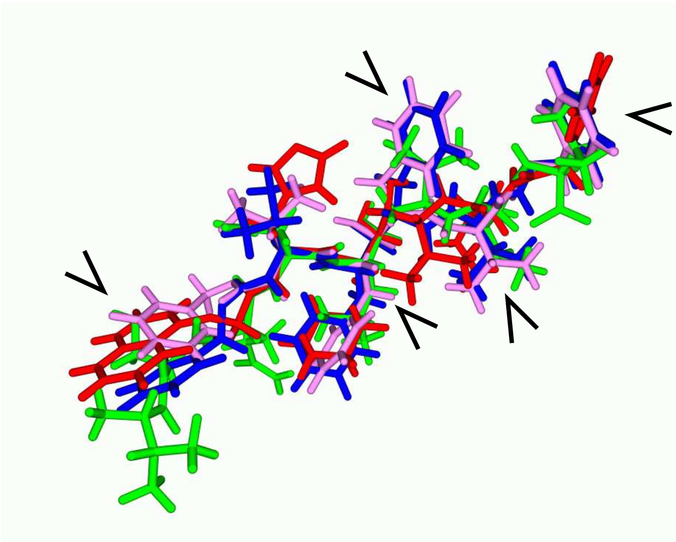


Abbildung 6.12: Die optimierten Strukturen: die Konformation 2 von 1hiv (rot), die Konformation 64 von 9hvp (blau), die Konformation 67 von 1hos (violett) und die Kristallstruktur von 8hvp (grün).

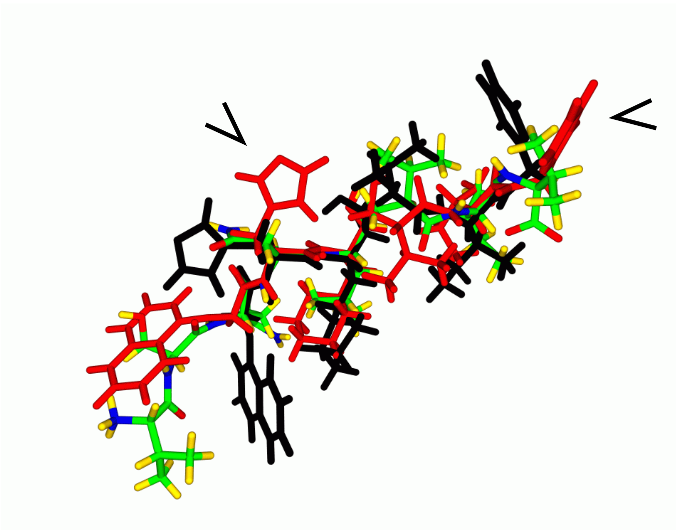


Abbildung 6.13: Die Konformation 2 von 1hiv (rot) und die Referenzstruktur von 8hvp (grün-gelb). Die Kristallstruktur von 1hiv ist schwarz.

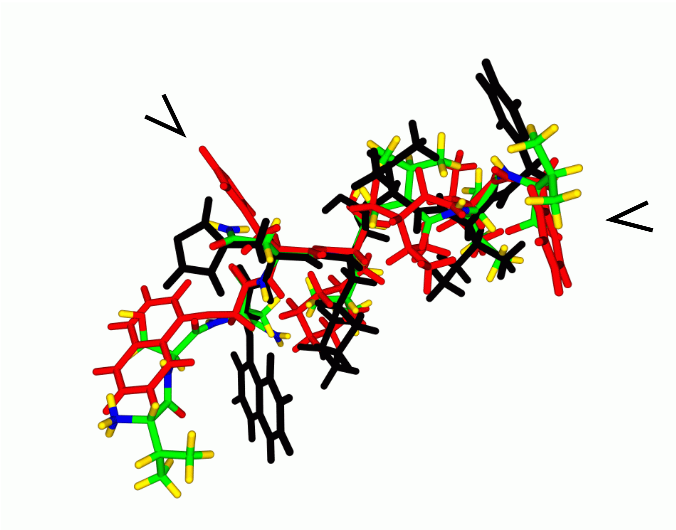


Abbildung 6.14: Die nachoptimierten Konformation 2 von 1hiv (rot) und die Referenzstruktur von 8hvp (grün-gelb). Die Kristallstruktur von 1hiv ist schwarz.

Kapitel 7

Zusammenfassung und Ausblick

Zusammenfassung der Methode:

Um das Alignment zu einem gegebenen Paar von Molekülen zu optimieren, ist eine Methode zur multidimensionalen Nachoptimierung der Molekülstrukturen konzipiert und implementiert worden, wobei die Strukturen die im Experiment beobachtete Bindungsgeometrie approximieren. Dazu ist der Alignmentindex als normierte Summe von Hodgkinindizes definiert worden. Die Hodgkinindizes bewerten die Überlagerung der räumlich modellierten physikochemischen Eigenschaften der beiden Moleküle. Dabei kann man zwei Fälle unterscheiden: die Überlagerung eines flexiblen Moleküls bzgl. eines starren Referenzmoleküls bzw. die gegenseitige Überlagerung zweier flexibler Moleküle. Die gegenseitige Überlagerung zweier Moleküle ist dahingehend erweitert worden, daß die Anzahl der flexiblen Moleküle variiert werden kann. Um die Strukturen der flexiblen Moleküle zu optimieren, ist ein Transformationsoperator definiert worden, der aus der globalen Rotation um ein variables Rotationszentrum, der globalen Translation und einem lokalen Rotationsoperator besteht. Dabei beschreibt der lokale Rotationsoperator hintereinander geschaltete Rotationen um molekülinterne Drehachsen. Die Energiefunktion besteht im ersten Fall aus dem Alignmentindex, einem n-Potential, das molekülinterne Anziehungs- und Abstoßungskräfte der Atome bewertet, und aus einer heuristischen Bewertungsfunktion für diejenigen Torsionswinkel, deren Wert variiert werden kann, während im zweiten Fall die Energiefunktion aus einer normierten Summe von Alignmentindizes, n-Potentialen und von heuristischen Bewertungsfunktionen für die Torsionswinkel besteht. Der Optimierungsparameter der jeweiligen Energiefunktionen besteht aus den Transformationsparametern des jeweiligen Transformationsoperators. Zum Schluß ist ein lokales Optimierungsverfahren ausgewählt worden, um Energiefunktionen zu minimieren.

Zusammenfassung der Resultate:

In Kapitel 4.2 ist ein Optimierungsverfahren bestimmt worden, das in der Lage ist, das Alignmentproblem bzw. die entsprechenden Energiefunktionen in angemessener Zeit zufriedenstellend zu lösen. So konnten bis 35-dimensionale Optimierungsprobleme (10 globale Transformations- und 15 interne Rotationsparameter) in durchschnittlich 31 sec. gelöst werden. Jedoch nimmt

die Rechenzeit ganz erheblich zu, wenn die Dimension des Optimierungsproblems größer als 35 ist. Genauso verhält es sich bei der gegenseitigen Überlagerung einer variablen Anzahl von flexiblen Molekülen. Je gröser die Dimension des Optimierungsproblems ist, um so länger braucht das Minimierungsverfahren zur Lösung. Daher sind zwei Optimierungsalgorithmen konzipiert worden, um die Energiefunktionen effizienter zu minimieren. Nur das Gauß-Seidel-artige Optimierungsverfahren ist teilweise in der Lage zufriedenstellende Lösungen zu produzieren. Das Hierarchie Verfahren ist in dieser Form noch nicht ganz ausgereift.
Die flexible Nachoptimierung mittels der Energiefunktionen aus Kapitel 3.2 ist in der Lage, das Alignment einer Konformation bzgl. einer Referenzstruktur zu verbessern. Insbesondere der lokale Rotationsoperator $\mathcal{R}_{\mathcal{A}}$ ermöglicht es, daß die einzelnen Molekülfragmente bzgl. der Referenzstruktur orientiert werden können. Der RMS-Abstand korrespondiert nicht in jedem Fall mit dem Alignment. Zwar konnte in der Hälfte der Beispielrechnungen, der RMS-Abstand zur Kristallstruktur verbessert werden, aber die meisten Verbesserungen liegen nur in der Grösenordnun von 0.1 bis 0.2 Å.
In einigen Fällen konnte FlexS kein Alignment bestimmen. Anhand eines Beispiels ist gezeigt worden, wie man mit den Energiefunktionen aus Kapitel 3.3 doch noch Alignment berechnen kann.

Ziel der vorliegenden Arbeit ist es gewesen, eine Methode zur multidimensionalen Nachoptimierung zu konzipieren, welches es erlaubt vom diskreten Konformationsmodell abzuweichen und so eine Relaxation der Strukturen zu erreichen. Durch simultane Berücksichtung einer ganzen Anzahl von Liganden sollen zusätzliche Informationen, die bei der paarweisen Überlagerungunentdeckt bleiben, während der Nachoptimierung ausgenutzt werden, um so die Qualität des Alignments zu verbessern.
Als Fazit kann man, daß die Ziele der Arbeit reicht worden sind.

Ausblick:

In Stichpunkten wird kurz dargelegt, was noch von Interesse sein könnte.

- Implementation eines globalen Optimierungsverfahren;
- Entwicklung von Varianten des Hierarchie-Verfahrens;
- Weitere Anwendungen für die Energiefunktionen aus Kapitel 3.3;
- Untersuchung, inwieweit die in Kapitel 6.2 vorgestellten Vorschläge umsetzbar sind;
- Inwieweit können Rechnungen beschleunigt werden?

Anhang A

FlexS-77 Datensatz

In diesem Anhang werden die für die Tests verwendeten Moleküle kurz dargestellt. Die verwendeten Moleküle gehören zum FlexS-77 Datensatz ([42]). Dieser Datensatz besteht aus 77 Molekülen, die nach den folgenden Gruppen unterteilt werden. Diese Gruppen sind Carboxypeptidase A, Concanavalin, Dihydrofolatreduktase, Elastase, Endothiapepsin, Fruktose-Bisphosphatase, Glycogenphosphorylase, HIV-Protease, Human Rhinovirus, Immunoglobulin, Streptavidin, Thermolysin, Thrombin und Trypsin. Die Moleküle werden mit der PDB-Nummer[1] des Rezeptor-Ligand-Komplexes bzw. mit CASP2-Nummer[2] des jeweiligen Rezeptor-Ligand-Komplexes bezeichnet. Bei der Beschreibung der Aufgaben, die die Proteine/Rezeptoren bzw. Moleküle erfüllen, habe ich auf die folgende Literatur zurück gegriffen: Hunnius Pharmazeutisches Wörterbuch [34], Löffler/Petrides, Biochemie [43], Pharmakologie und Toxikologie [18], Pschyrembel Klinisches Wörterbuch [47], Stryer, Biochemistry [56], Taschenatlas der Biochemie [36] und Taschenatlas der Pharmakologie [44].

Neben der Beschreibung der Aufgaben werden die wichtigsten Daten der jeweiligen zu Gruppe gehörenden Moleküle in einer Tabelle aufgelistet. Mit den Bezeichnungen aus Kapitel 2.2 werden für jedes Molekül $\mathcal{A}$ des FlexS-77 Datensatzes die Anzahl der Atome $n_a[\mathcal{A}] = |K_\mathcal{A}|$, die Anzahl der Molekülfragmente $|F_\mathcal{A}|$ und die Anzahl der Bindungen $n_{ft}[\mathcal{A}] = \left|B_{FT}^\mathcal{A}\right|$, an denen der Torsionswinkel variiert werden kann, angegeben. $n_{ft}[\mathcal{A}]$ entspricht der Anzahl der internen Rotationsparameter. Weiter wird die Tiefe des zum Molekül $\mathcal{A}$ gehörenden Fragmentebaums $\vec{T}_\mathcal{A}$ angegeben, wobei die Tiefe $t(\vec{T}_\mathcal{A})$ definiert ist als die Länge des längsten f_0-f_i-Weges in $\vec{T}_\mathcal{A}$: $t(\vec{T}_\mathcal{A}) := \max_{f_i \in F_\mathcal{A} \setminus f_0} l(\vec{P}_{f_0,f_i})$. Dann wird die Anzahl der Gaussfunktionen angegeben, um die räumlich lokalisierte Dichte der Elektronendichte, $\left|GC_\mathcal{A}^1\right|$, der Ladung, $\left|GC_\mathcal{A}^2\right|$, der Hydrophobizität, $\left|GC_\mathcal{A}^3\right|$, der Wasserstoffdonatoren, $\left|GC_\mathcal{A}^4\right|$, und der Wasserstoffakzeptoren, $\left|GC_\mathcal{A}^5\right|$, darzustellen. Mit diesen Daten über die Moleküle kann man sich ungefähr vorstellen, wie die entsprechenden Energiefunktionen aus Kapitel 3.2 bzw. 3.3 aussehen. Zum Schluß wird die Identifikationsnummer des Moleküls angegeben, die benutzt wird, um die Ergebnisse in Kapitel 4.2 und 4.3 bzw. im Anhang A graphisch bzw. in Tabellen darzustellen. Dabei richtet sich die

[1] Protein Data Bank: [5], [49].

[2] Critical Assessment of Techniques for Protein Structure Prediction 2: [14]

Nummer nach der Anzahl der internen Rotationsparameter. Für das Molekül mit der Nummer 1 gilt $n_{ft}[\mathcal{A}] = 1$, während für das Molekül mit der Nummer 77 $n_{ft}[\mathcal{A}] = 50$ gilt.

Bevor nun die Gruppen dargestellt werden, werden noch einige Begriffe erläutert, die häufiger in den Aufgabenbeschreibungen vorkommen. *Aminosäuren* sind organische Säuren, die mindestens eine Carboxyl- und eine Aminogruppe enthalten. Die α-Aminosäuren gehören als Bausteine der Proteine und Peptide zu den wichtigsten organischen Stoffen der lebenden Zelle. Ein *Antikörper* ist ein spezifischer Stoff, der durch Einwirkung einer fremden Substanz, eines Antigens, vom Organismus gebildet wird. Die biologische Bedeutung der Antikörper liegt in ihrer Schutzfunktion gegen Mikroorganismen sowie gegen körperfremde bzw. körpereigene entartete Substanzen, die entfernt oder neutralisiert werden müssen. *Derivate* sind chemische Verbindungen, die aus anderen chemischen Substanzen entstanden sind. *Enzyme* sind Proteine, die als Biokatalysatoren im lebenden tierischen und pflanzlichen Organismus chemische Reaktionen beschleunigen, ohne selbst verbraucht zu werden. Die an der Reaktion beteiligten Substanzen (Substrate) lagern sich an das Aktivitätszentrum des Enzyms an. Einige Enzyme brauchen für ihre Aktivität Kofaktoren, meist Metallionen. *Enzyminhibitoren* sind Moleküle, die die Aktivität einzelner Enzyme hemmen, indem sie dank einer höheren Affinität die Substrate aus den Bindungstaschen der Enzyme verdrängen. *Glykoproteine* sind Eiweißzucker, die zu den zusammengesetzten Proteinen zählen. *Lektine* sind hauptsächlich aus Pflanzen isolierte Glykoproteine, die spezifisch mit membranständigen und gelösten Kohlenhydratstrukturen reagieren und dadurch verschiedene Zellen (z.B. Tumorzellen aber auch bestimmte Bakterien) mit z.T. großer Spezifität agglutinieren (anleimen) bzw. präzipitieren (ausfällen). D.h. man kann sie zum Nachweis bestimmter Zellen verwenden. *Peptide* sind chemische Verbindungen, die aus zwei oder mehreren Aminosäuren aufgebaut sind. Peptide, die aus einer langen Kette von Aminosäuren bestehen (19 bis 100), heißen Polypeptide, während Peptide, die aus 100 und mehr Aminosäuren bestehen, Proteine heißen. *Proteasen* sind Enzyme, die durch Spaltung der Peptidbindung Proteine und Peptide abbauen. Je nach Angriffspunkt an der Peptidkette werden die Proteasen in 2 Gruppen eingeteilt. *Endopeptidasen* katalysieren die Spaltung der Bindung innerhalb einer Peptidkette, während die *Exopeptidasen* immer einzelne Aminosäuren vom Peptidkettenende her abspalten. Proteasen kommen im Organismus weit verbreitet vor, insbesondere im Verdauungstrakt. Außerdem sind sie am System der Blutgerinnung beteiligt. *Proteine* sind vorwiegend aus Aminosäuren aufgebaute Makromoleküle. In ihrer wichtigen Funktion als Enzyme ermöglichen die Proteine den geregelten Ablauf biochemischer Reaktionen im Stoffwechsel der Zelle. Weiter haben die Proteine Bedeutung u.a. als Hormone, als Membranproteine (z.B. Rezeptor, G-Protein, Transporter), als Transportproteine, als Antikörper und als Faktoren der Blutgerinnung. Proteine werden aufgrund ihrer Zusammensetzung in zwei Hauptklassen unterteilt: den einfachen und den zusammengesetzten Proteinen. *Steriode* sind eine umfangreiche Gruppe von Verbindungen, die alle das Grundgerüst des Cyclopentanoperhydrophenanthrens (Gonan), einer Verbindung mit drei Sechs(-Kohlenstoff-)ringen und einem Fünfring, gemeinsam haben. Steriode enthalten 18 bis 30 Kohlenstoffatome. *Viren* sind biologische Strukturen mit den folgenden gemeinsamen Merkmalen. Viren enthalten entweder nur die DNA oder nur die RNA als genetische Informationen. Sie verfügen nicht über die für Teilung und Wachstum erforderlichen Enzyme, sondern benötigen eine Wirtszelle, um sich zu reproduzieren. Viren, die bei ihrer Vermehrung die Wirtszelle schädigen, sind Krankheitserreger. Viren bestehen aus einer Nukleinsäuresequenz (DNA oder RNA), einem Proteinmantel (Kapsid). Viren werden daher in DNA- und RNA-Viren unterteilt. Komplexe Virions (Nukleinsäure und Proteinmantel) werden zusätzlich von einer Hülle umgeben. Die Virushülle stammt z.T. aus der Zellmembran der Wirtszelle, wobei viruseigene Glykoproteine aus der Hülle herausragen, die die Bindung an die Wirtszelle bewirken.

Carboxypeptidase A:

Carboxypeptidase A ist ein zu den Proteasen gehörendes Enzym, das terminale Aminosäuren mit freier Carboxylgruppe von Peptiden abspaltet (Exopeptidasen). Carboxypeptidase A ist eine Zinkprotease, d.h. sie braucht Zink als Kofaktor, um Peptidketten zu spalten. Carboxypeptidasen werden in der Pankreas (Bauchspeicheldrüse) produziert, kommen aber auch in den Nieren, in der Milz und in anderen Organen vor. Carboxypeptidase A ist ein Verdauungsenzym. Die PDB-Strukturen 1cbx, 2ctc, 3cpa, 6cpa und 7cpa sind Peptide und hemmen als Inhibitoren die Fähigkeit von Carboxypeptidase A, Aminosäuren von Peptidketten abzuspalten, indem sie die Substrate aus den Bindungstaschen verdrängen.

Name	$n_a[\mathcal{A}]$	$\lvert F_\mathcal{A}\rvert$	$n_{ft}[\mathcal{A}]$	$t(\vec{T}_\mathcal{A})$	$\lvert GC^1_\mathcal{A}\rvert$	$\lvert GC^2_\mathcal{A}\rvert$	$\lvert GC^3_\mathcal{A}\rvert$	$\lvert GC^4_\mathcal{A}\rvert$	$\lvert GC^5_\mathcal{A}\rvert$	Nr.
1cbx	25	6	5	2	14	24	15	0	4	11
2ctc	21	5	4	2	12	18	12	1	3	7
3cpa	31	9	8	4	17	25	17	3	4	26
6cpa	58	17	16	7	30	51	29	2	8	55
7cpa	74	20	19	7	39	65	38	2	8	60

Concanavalin:

Concanavalin A ist ein Lektin aus Canavalia ensiformis (Schwertbohne) und dient zur Reinigung von Glykoproteinen und zum Nachweis von Membranveränderungen (Krebszellen). Die PDB-Struktur 5cna und die CASP2-Struktur t0013 sind Glukose-Moleküle (Mannose- und Arabinose-Derivate), die an Concanavalin A binden.

Name	$n_a[\mathcal{A}]$	$\lvert F_\mathcal{A}\rvert$	$n_{ft}[\mathcal{A}]$	$t(\vec{T}_\mathcal{A})$	$\lvert GC^1_\mathcal{A}\rvert$	$\lvert GC^2_\mathcal{A}\rvert$	$\lvert GC^3_\mathcal{A}\rvert$	$\lvert GC^4_\mathcal{A}\rvert$	$\lvert GC^5_\mathcal{A}\rvert$	Nr.
5cna	27	8	7	3	13	17	8	4	6	17
t0013	23	7	6	3	11	14	7	3	5	15

Dihydrofolatreduktase:

Dihydrofolatreduktase ist ein Enzym, das in den Mukosazellen des Jejunums Dihydrofolsäure (PDB: 1dhf) zu Tetrahydrofolsäure reduziert. Tetrahydrofolsäure synthetisiert die Bausteine von DNS und RNS. Durch einen Mangel an Tetrahydrofolsäure wird die Zellvermehrung gehemmt. Der Folsäureantagonist Methotrexat (PDB: 4dfr) hat eine viel höhere Affinität zur Dihydrofolatreduktase als das natürliche Substrat, Dihydrofolsäure. Es unterbindet deshalb die Umwandlung von Dihydrofolsäure zur Tetrahydrofolsäure und hemmt damit das Zellwachstum. Methotrexat ist daher ein Inhibitor und wird als Arzneistoff zur Chemotherapie von Tumoren eingesetzt.

Name	$n_a[\mathcal{A}]$	$\lvert F_\mathcal{A}\rvert$	$n_{ft}[\mathcal{A}]$	$t(\vec{T}_\mathcal{A})$	$\lvert GC^1_\mathcal{A}\rvert$	$\lvert GC^2_\mathcal{A}\rvert$	$\lvert GC^3_\mathcal{A}\rvert$	$\lvert GC^4_\mathcal{A}\rvert$	$\lvert GC^5_\mathcal{A}\rvert$	Nr.
1dhf	49	13	12	5	32	45	27	4	10	49
4dfr	53	14	13	5	32	46	28	3	9	50

Elastase:

Elastase ist eine Protease, genauer eine Endopeptidase. Elastase spaltet als Verdauungsenzym Faserproteine (Elastin und Kollagen). Die PDB-Strukturen 1ela, 1elb, 1elc, 1eld, 1ele und die CASP2-Struturen t0035, t0036 sind Peptide und hemmen als Inihibitoren die Verdauungsaktivität von Elastase.

Name	$n_a[\mathcal{A}]$	$\|F_\mathcal{A}\|$	$n_{ft}[\mathcal{A}]$	$t(\vec{T}_\mathcal{A})$	$\|GC^1_\mathcal{A}\|$	$\|GC^2_\mathcal{A}\|$	$\|GC^3_\mathcal{A}\|$	$\|GC^4_\mathcal{A}\|$	$\|GC^5_\mathcal{A}\|$	Nr.
1ela	64	17	16	6	29	57	29	3	3	56
1elb	69	22	21	7	30	64	30	4	3	64
1elc	70	20	19	7	33	63	33	4	3	61
1eld	52	14	13	5	27	48	27	3	3	51
1ele	48	15	14	5	23	46	23	3	3	54
t0035	37	8	7	3	23	34	21	1	5	18
t0036	42	12	11	5	25	39	25	2	5	42

Endothiapepsin:

Endothiapepsin ist eine Protease, genauer eine Endopeptidase. Pepsin ist eine Sammelbezeichnung für eine Gruppe von Endopeptidasen gering unterschiedlicher Spezifität, die im Magensaft Nahrungsproteine spalten. Endothiapepsin ist also ein Verdauungsenzym. Die PDB-Strukturen 2er7, 4er1, 4er2, 5er1 und 5er2 sind Peptide und hemmen als Inhibitoren die Verdauungsaktivität von Endothiapepsin.

Name	$n_a[\mathcal{A}]$	$\|F_\mathcal{A}\|$	$n_{ft}[\mathcal{A}]$	$t(\vec{T}_\mathcal{A})$	$\|GC^1_\mathcal{A}\|$	$\|GC^2_\mathcal{A}\|$	$\|GC^3_\mathcal{A}\|$	$\|GC^4_\mathcal{A}\|$	$\|GC^5_\mathcal{A}\|$	Nr.
2er7	159	49	48	14	79	134	79	10	14	76
4er1	127	30	29	9	60	116	60	5	5	68
4er2	110	41	40	11	48	48	55	7	9	74
5er1	87	29	28	8	37	69	35	5	5	66
5er2	143	42	41	11	68	125	68	10	10	75

Fruktose-Bisphosphatase:

Fruktose-1,6-Bisphosphatase ist ein Schlüsselenzym der Glukoneogenese, d.h. der Neusynthese von Glukose (Traubenzucker) in Leber und Nieren aus Nicht-Kohlenhydratvorstufen wie Aminosäuren, Laktat sowie Glyzerin. Adenosin-Monophosphat (PDB: 4fbp), eine Adenylsäure, hemmt die Ezymaktivität von Fruktose-1,6-Bisphosphatase und ist damit ein Inhibitor. Die CASP2-Struktur t0039 gehört zu den Purinen, die biologisch eine große Rolle spielen.

Name	$n_a[\mathcal{A}]$	$\|F_\mathcal{A}\|$	$n_{ft}[\mathcal{A}]$	$t(\vec{T}_\mathcal{A})$	$\|GC^1_\mathcal{A}\|$	$\|GC^2_\mathcal{A}\|$	$\|GC^3_\mathcal{A}\|$	$\|GC^4_\mathcal{A}\|$	$\|GC^5_\mathcal{A}\|$	Nr.
4fbp	35	8	7	3	23	27	17	3	10	19
t0039	35	10	9	4	22	28	17	4	9	32

Glycogenphosphorylase:

Glycogenphosphorylase ist ein Enzym, das in Leber und Muskulatur Glykogen spaltet. Glykogen ist sog. tierische Stärke und dient im Körper als Reservekohlenhydrat. Die PDB-Strukturen 1gpy, 3gpb, 4gpb und 5gpb sind Glukose-Moleküle (Traubenzucker) und binden an das Enzym Glycogenphosphorylase.

Name	$n_a[\mathcal{A}]$	$\lvert F_{\mathcal{A}}\rvert$	$n_{ft}[\mathcal{A}]$	$t(\vec{T}_{\mathcal{A}})$	$\lvert GC_{\mathcal{A}}^1\rvert$	$\lvert GC_{\mathcal{A}}^2\rvert$	$\lvert GC_{\mathcal{A}}^3\rvert$	$\lvert GC_{\mathcal{A}}^4\rvert$	$\lvert GC_{\mathcal{A}}^5\rvert$	Nr.
1gpy	27	8	7	2	16	20	11	4	9	20
3gpb	27	8	7	3	16	20	11	4	9	21
4gpb	26	7	6	3	15	19	11	3	8	14
5gpb	29	8	7	3	16	20	12	4	8	22

HIV-Protease:

HIV-1 Protease ist eine Protease des humanen Immundefizienz-Virus (HIV). HIV ist ein RNA-Virus. HIV-1 Proteasen spielen eine entscheidende Rolle bei der Infektion und Übernahme von neuen Wirtszellen und damit bei der Vermehrung der HI-Viren. Daher ist die HIV-1 Protease ein attraktives Ziel bei der Forschung nach AIDS-Medikamenten.[3] Die PDB-Strukturen 1hiv, 1hos, 1ivp, 1ivq, 2mip, 4hvp, 4phv, 5hvp, 8hvp und 9hvp sind Peptide und hemmen die Enzymaktivität der HIV-1 Protease. D.h. diese Moleküle sind HIV-1 Protease-Inhibitoren.

Name	$n_a[\mathcal{A}]$	$\lvert F_{\mathcal{A}}\rvert$	$n_{ft}[\mathcal{A}]$	$t(\vec{T}_{\mathcal{A}})$	$\lvert GC_{\mathcal{A}}^1\rvert$	$\lvert GC_{\mathcal{A}}^2\rvert$	$\lvert GC_{\mathcal{A}}^3\rvert$	$\lvert GC_{\mathcal{A}}^4\rvert$	$\lvert GC_{\mathcal{A}}^5\rvert$	Nr.
1hiv	120	32	31	9	59	105	57	7	9	70
1hos	105	29	28	9	55	97	55	4	8	67
1ivp	120	32	31	9	59	105	57	7	9	71
1ivq	95	26	25	7	44	86	43	5	6	65
2mip	116	37	36	10	53	101	54	7	7	72
4hvp	123	44	43	11	52	107	52	12	8	73
4phv	88	18	17	6	46	77	46	5	5	58
5hvp	101	38	37	10	45	87	45	7	9	39
8hvp	134	51	50	13	61	115	60	11	12	77
9hvp	106	30	29	9	54	97	54	5	7	69

Human Rhinovirus:

Rhinoviren sind RNA-Viren ohne Hüllen und lösen beim Menschen die sogenannten „Erkältungskrankheiten“ aus. Der Rhinovirus besteht aus der RNA und dem Kapsid (Proteinmantel), wobei das Kapsid von drei Proteinen gebildet wird. Die PDB-Strukturen 2r04, 2r06, 2r07, 2rm2, 2rr1, 2rs1, 2rs3 und 2rs5 sind antivirale Verbindungen, die an den Proteinen des Kapsid binden.

Name	$n_a[\mathcal{A}]$	$\lvert F_{\mathcal{A}}\rvert$	$n_{ft}[\mathcal{A}]$	$t(\vec{T}_{\mathcal{A}})$	$\lvert GC_{\mathcal{A}}^1\rvert$	$\lvert GC_{\mathcal{A}}^2\rvert$	$\lvert GC_{\mathcal{A}}^3\rvert$	$\lvert GC_{\mathcal{A}}^4\rvert$	$\lvert GC_{\mathcal{A}}^5\rvert$	Nr.
2r04	51	12	11	6	25	45	23	0	5	44
2r06	45	10	9	5	23	39	21	0	5	35
2r07	45	10	9	5	23	39	22	0	5	36
2rm2	54	13	12	6	26	49	25	0	5	45
2rr1	54	13	12	6	26	49	24	0	5	46
2rs1	54	13	12	6	26	49	24	0	5	47
2rs3	57	14	13	7	27	52	25	0	5	52
2rs5	48	11	10	5	24	43	22	0	5	40

Immunoglobuline:

Immunoglobuline sind Glykoproteine mit gemeinsamer Grundstruktur, die nach Kontakt des Organismus mit einem Antigen (einer fremden Substanz) u.a. von den B-Lymphozyten (weiße Blutzellen) gebildet werden und die eine wichtige Rolle als Antikörper im Immunsystem spielen. Antikörper lassen sich durch Enzyme spalten. So entstehen durch *Papain* zwei identische F_{ab}-Fragmente und ein F_c-Fragment. Die F_{ab}-Fragmente sind weiter in der Lage, Antigene zu binden.

[3] Siehe Stryer, Biochemistry [56] und Löffler/Petrides, Biochemie [43].

Die PDB-Strukturen 1dbb, 1dbj, 1dbk, 1dbm und 2dbl binden an dem F_{ab}-Fragment eines monoklonalen Antikörpers. Monoklonale Antikörper werden u.a. aus Zellklonen synthetisiert. Sie werden u.a. zum qualitativen und quantitativen Antigen- und Antikörpernachweis sowie nach radioaktiver Markierung zur Tumorlokalisation verwendet. Die hier verwendeten PDB-Strukturen gehören zu den Sterioden.

Name	$n_a[\mathcal{A}]$	$\|F_{\mathcal{A}}\|$	$n_{ft}[\mathcal{A}]$	$t(\vec{T}_{\mathcal{A}})$	$\|GC^1_{\mathcal{A}}\|$	$\|GC^2_{\mathcal{A}}\|$	$\|GC^3_{\mathcal{A}}\|$	$\|GC^4_{\mathcal{A}}\|$	$\|GC^5_{\mathcal{A}}\|$	Nr.
1dbb	53	5	4	2	23	50	21	0	2	8
1dbj	51	4	3	2	21	48	20	1	2	5
1dbk	49	3	2	2	21	43	19	0	2	2
1dbm	64	10	9	4	31	58	29	0	6	33
2dbl	67	10	9	4	30	63	29	0	5	34

Streptavidin:

Streptavidin ist ein Protein, an das Biotin (Vitamin H) bindet. Biotin ist ein Harnstoffderivat und ist wichtig für die Glukoneogenese sowie beim Abbau einiger Aminosäuren. Streptavidin findet man in den Kulturen des Bakteriums Streptomyces avidinii, welches zu Familie der Streptomyzeten gehört. Streptavidin wird unter anderem als Antigen und zur Reinigung von Proteinen und anderen Antigenen verwendet. Die PDB-Strukturen 1srf, 1srg, 1srh, 1sri und 1srj binden als Reporter-Moleküle an Streptavidin.

Name	$n_a[\mathcal{A}]$	$\|F_{\mathcal{A}}\|$	$n_{ft}[\mathcal{A}]$	$t(\vec{T}_{\mathcal{A}})$	$\|GC^1_{\mathcal{A}}\|$	$\|GC^2_{\mathcal{A}}\|$	$\|GC^3_{\mathcal{A}}\|$	$\|GC^4_{\mathcal{A}}\|$	$\|GC^5_{\mathcal{A}}\|$	Nr.
1srf	39	9	8	3	22	39	20	1	5	27
1srg	30	6	5	3	19	30	17	1	5	13
1srh	35	9	8	3	22	29	17	1	7	28
1sri	33	7	6	3	20	33	18	1	5	16
1srj	33	5	4	3	22	33	20	1	5	9

Thermolysin:

Thermolysin ist ein bakteriologisches Enzym und gehört wie Carboxypeptidase A zu den Zinkproteasen. Die PDB-Strukturen 1tlp, 1tmn, 2tmn, 3tmn, 4tln, 4tmn, 5tln, 5tmn und die Strukturen cbz, ppp, rthior, thior, die in einem Aufsatz von G.Klebe u.a.([35]) veröffentlicht worden sind, sind Peptide und hemmen als Inhibitoren die Fähigkeit von Thermolysin, Peptidketten zu spalten.

Name	$n_a[\mathcal{A}]$	$\|F_{\mathcal{A}}\|$	$n_{ft}[\mathcal{A}]$	$t(\vec{T}_{\mathcal{A}})$	$\|GC^1_{\mathcal{A}}\|$	$\|GC^2_{\mathcal{A}}\|$	$\|GC^3_{\mathcal{A}}\|$	$\|GC^4_{\mathcal{A}}\|$	$\|GC^5_{\mathcal{A}}\|$	Nr.
1tlp	69	19	18	5	37	61	32	6	10	59
1tmn	67	17	16	5	35	57	35	3	5	57
2tmn	26	9	8	3	13	25	13	2	4	29
3tmn	43	11	10	4	22	36	22	3	3	41
4tln	24	8	7	3	10	18	10	2	2	23
4tmn	68	20	19	6	35	63	36	3	7	62
5tln	43	14	13	5	24	38	23	4	5	53
5tmn	64	21	20	7	32	58	32	3	7	63
cbz	38	9	8	4	22	35	22	1	4	30
ppp	40	9	8	4	21	37	22	1	3	31
rthior	30	8	7	3	17	25	16	1	3	24
thior	30	8	7	3	17	25	16	1	3	25

Thrombin:

Thrombin ist eine Protease, und zwar eine Endopeptidase. Thrombin spielt bei der Blutgerinnung eine zentrale Rolle. Thrombin wandelt das lösliche Plasmaprotein Fribrinogen in das unlösliche Protein Fibrin durch Abspaltung kleiner Peptide um. Fibrin polymerisiert dann zu einem Gerinnsel. Die PDB-Strukturen 1dwc, 1dwd und 1ett sind Peptide und hemmen als Inihibitoren die Enzymaktivität von Thrombin und damit die Blutgerinnung.

Name	$n_a[\mathcal{A}]$	$\lvert F_{\mathcal{A}}\rvert$	$n_{ft}[\mathcal{A}]$	$t(\vec{T}_{\mathcal{A}})$	$\lvert GC^1_{\mathcal{A}}\rvert$	$\lvert GC^2_{\mathcal{A}}\rvert$	$\lvert GC^3_{\mathcal{A}}\rvert$	$\lvert GC^4_{\mathcal{A}}\rvert$	$\lvert GC^5_{\mathcal{A}}\rvert$	Nr.
1dwc	71	13	12	4	34	63	31	5	5	48
1dwd	69	12	11	5	36	60	33	4	4	43
1ett	59	10	9	4	29	52	25	3	3	37

Trypsin:

Trypsin ist eine Protease, genauer eine Endopeptidase. Trypsin ist ein Verdauungsenzym, das im Pankreassekret (Bauchspeicheldrüsensekret) vorkommt. Die PDB-Strukturen 1pph, 1tnh, 1tni, 1tnj, 1tnk, 1tnl und 3ptb sind Peptide und hemmen als Inihibitoren die Verdauungsaktivität von Trypsin.

Name	$n_a[\mathcal{A}]$	$\lvert F_{\mathcal{A}}\rvert$	$n_{ft}[\mathcal{A}]$	$t(\vec{T}_{\mathcal{A}})$	$\lvert GC^1_{\mathcal{A}}\rvert$	$\lvert GC^2_{\mathcal{A}}\rvert$	$\lvert GC^3_{\mathcal{A}}\rvert$	$\lvert GC^4_{\mathcal{A}}\rvert$	$\lvert GC^5_{\mathcal{A}}\rvert$	Nr.
1pph	59	10	9	4	29	52	25	3	3	38
1tnh	18	3	2	1	8	15	8	1	0	3
1tni	27	6	5	3	11	24	11	1	0	12
1tnj	21	4	3	2	9	16	9	1	0	6
1tnk	24	5	4	2	10	21	10	1	0	10
1tnl	22	3	2	2	10	16	10	1	0	4
3ptb	18	2	1	1	8	18	8	2	0	1

Transformationen für die heuristischen Testbeispiele:

Im folgenden werden 10 Transformationen vorgestellt. Diese Transformationen sind auf die obigen 77 Moleküle angewendet worden. Die 770 Testbeispiele sind dann in den heuristischen Testverfahren benutzt worden. Im folgenden seien

$$X(\psi_1) := \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\psi_1 & \sin\psi_1 \\ 0 & -\sin\psi_1 & \cos\psi_1 \end{pmatrix}, \quad Y(\psi_2) := \begin{pmatrix} \cos\psi_2 & 0 & \sin\psi_2 \\ 0 & 1 & 0 \\ -\sin\psi_2 & 0 & \cos\psi_2 \end{pmatrix}$$

und

$$Z(\psi_3) := \begin{pmatrix} \cos\psi_3 & \sin\psi_3 & 0 \\ -\sin\psi_3 & \cos\psi_3 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Rotationsmatrizen, wobei $\psi_i \in [0, 360°]$ für $i = 1, 2, 3$ gilt. $X(\psi)$ beschreibt die Rotation um die x-Achse um ψ Grad, während $Y(\psi)$ bzw. $Z(\psi)$ die Rotation um die y- bzw. z-Achse beschreiben. Dann lassen sich die Transformationen für ein Molekül $\mathcal{A}$ wie folgt beschreiben:

$$T_{\mathcal{A}}(D(\psi_1,\psi_2,\psi_3),\varphi_{\mathcal{A}},t,\cdot) \;:=\; D(\psi_1,\psi_2,\psi_3)\Big(\mathcal{R}[\mathcal{S}(K_{\mathcal{A}},\vec{T}_{\mathcal{A}})](\varphi_{\mathcal{A}},\cdot) - c_{\mathcal{A}}\Big) + c_{\mathcal{A}} + t, \qquad \text{(A.1)}$$

wobei die Bezeichnungen aus Kapitel 2.2 und 2.4.2 benutzt werden. Dabei ist $D(\psi_1,\psi_2,\psi_3) \in \mathsf{SO}_3$ eine Rotationsmatrix. $\mathcal{R}[\mathcal{S}(K_{\mathcal{A}},\vec{T}_{\mathcal{A}})]$ ist der lokale Rotationsoperator von $K_{\mathcal{A}}$ bzgl. der

Transformationsstruktur $\mathcal{S}(K_{\mathcal{A}}, \vec{T}_{\mathcal{A}})$, wobei $K_{\mathcal{A}} \subset \mathbb{R}^3$ die Menge der Atomkoordinaten von $\mathcal{A}$, $\vec{T}_{\mathcal{A}}$ der gerichtete Fragmentebaum von $\mathcal{A}$ und $\varphi_{\mathcal{A}} \in [0, 360°]^{n_{ft}[\mathcal{A}]}$ der Vektor der internen Rotationsparameter ist.[4] Dabei ist $n_{ft}[\mathcal{A}]$ die Anzahl der Bindungen, an denen der Torsionswinkel variiert werden kann. Als Rotationszentrum wird der Schwerpunkt des Molekül $\mathcal{A}$ verwendet. Somit gilt $c_{\mathcal{A}} := (1/n_a[\mathcal{A}]) \sum_{i=1}^{n_a[\mathcal{A}]} a_i \in \mathbb{R}^3$, wobei $n_a[\mathcal{A}]$ die Anzahl der Atomkoordinaten ist und $a_i \in K_{\mathcal{A}}$ für $i = 1, \ldots, n_a[\mathcal{A}]$ gilt. Als Translationsvektor wird $t \in \mathbb{R}^3$ verwendet. Wendet man $T_{\mathcal{A}}$ aus Gleichung (A.1) mit den entsprechenden Parametern auf die Atomkoordinaten aus $K_{\mathcal{A}}$ an, so erhält man die entsprechend transformierte Molekülkonformation von $\mathcal{A}$.

In folgenden werden die Parameter der Transformationen angegeben. Man beachte dabei, daß die Anzahl der internen Rotationsparameter von 1 (Trypsin: 3ptb) bis 50 (HIV-Protease: 8hvp) reicht.

1. Transformation:

Es gilt $D(\psi_1, \psi_2, \psi_3) := Z(\psi_3)Y(\psi_2)X(\psi_1)$ mit $\psi_1 = 45°$, $\psi_2 = 35°$ und $\psi_3 = 60°$, $\varphi =$ (5, 95, 10, 105, 15, 115, 20, 125, 25, 135, 30, 145, 35, 155, 40, 165, 45, 175, 50, 185, 55, 195, 60, 205, 65, 190, 70, 175, 75, 160, 80, 145, 85, 130, 90, 115, 95, 100, 100, 95, 105, 90, 110, 85, 115, 80, 120, 75, 125, 70) und $t = (-0.1, 0.2, -0.3)$.

2. Transformation:

Es gilt $D(\psi_1, \psi_2) := Z(\psi_2)X(\psi_1)$ mit $\psi_1 = 12°$ und $\psi_2 = 135°$, $\varphi =$ (34, 0, 43, 0, 52, 0, 61, 0, 70, 0, 89, 0, 98, 0, 107, 0, 116, 0, 125, 0, 134, 0, 143, 0, 142, 0, 151, 0, 160, 0, 80, 0, 75, 0, 70, 0, 65, 0, 60, 0, 55, 0, 50, 0, 45, 0, 40, 0, 35, 0) und $t = (0.2, 0.0, -0.2)$.

3. Transformation:

Es gilt $D(\psi_1, \psi_2) := Z(\psi_2)Y(\psi_1)$ mit $\psi_1 = 67°$ und $\psi_2 = 241°$, $\varphi =$ (0, 13, 0, 17, 0, 25, 0, 32, 0, 37, 0, 47, 0, 58, 0, 71, 0, 79, 0, 95, 0, 108, 0, 117, 0, 126, 0, 135, 0, 144, 0, 153, 0, 162, 0, 171, 0, 180, 0, 90, 0, 45, 0, 40, 0, 35, 0, 30, 0, 25) und $t = (0.0, 0.0, -0.5)$.

4. Transformation:

Es gilt $D(\psi_1, \psi_2) := Z(\psi_2)X(\psi_1)$ mit $\psi_1 = 43°$ und $\psi_2 = 289°$, $\varphi =$ (5, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 105, 110, 115, 120, 125, 130, 135, 140, 145, 150, 155, 160, 165, 170, 175, 180, 18, 36, 72, 144, 134, 124, 114, 104, 94, 84, 74, 64, 54, 44, 34) und $t = (0, 0, 0)$.

5. Transformation:

Es gilt $D(\psi_1) := Y(\psi_1)$ mit $\psi_1 = 156°$, $\varphi =$ (78, 0, 0, 90, 0, 102, 114, 0, 100, 0, 0, 86, 0, 0, 0, 72, 0, 0, 74, 0, 76, 78, 0, 70, 0, 0, 62, 0, 0, 0, 54, 0, 0, 63, 0, 72, 81, 92, 0, 82, 0, 0, 72, 0, 77, 83, 0, 63, 0, 0) und $t = (0, 0, 0)$.

6. Transformation:

Es gilt $D(\psi_1, \psi_2) := Z(\psi_2)Y(\psi_1)$, mit $\psi_1 = 220°$ und $\psi_2 = -45°$, $\varphi =$ (2, 4, 8, 16, 32, 64, 128, 3, 9, 27, 81, 5, 25, 125, 6, 36, 216, 7, 49, 8, 64, 9, 81, 10, 100, 45, 45, 45, 45, 54, 54, 54, 54, 30, 33, 39, 51, 57, 69, 84, 96, 111, 114, 120, 123, 129, 141, 147, 159, 174) und $t = (0, 0, 0)$.

[4] Siehe den entsprechenden Abschnitt in Kapitel 2.2 und die Definitionen 2.4.14 und 2.4.18.

7. Transformation:
Es gilt $D(\psi_1) := X(\psi_1)$ mit $\psi_1 = 121°$, $\varphi = (0, 1, 2, 0, 12, 13, 14, 0, 39, 40, 41, 0, 120, 121, 122, 0, 12, 11, 10, 0, 45, 0, 34, 0, 0, 45, 67, 0, 23, 34, 56, 12, 34, 58, 93, 45, 67, 89, 12, 34, 62, 31, 45, 9, 8, 12, 45, 67, 65, 34)$ und $t = (0, 0, 0)$.

8. Transformation:
Es gilt $D(\psi_1) := Z(\psi_1)$ mit $\psi_1 = 145°$, $\varphi = (180, 175, 170, 165, 160, 155, 150, 145, 140, 135, 130, 125, 120, 115, 110, 105, 100, 95, 90, 85, 80, 75, 70, 65, 60, 55, 50, 45, 40, 35, 30, 25, 20, 15, 10, 20, 40, 80, 160, 40, 10, 30, 90, 270, 27, 3, 15, 35, 46, 76)$ und $t = (0, 0, 0)$.

9. Transformation:
Es gilt $D(\psi_1, \psi_2, \psi_3) := Y(\psi_3)Z(\psi_2)X(\psi_1)$ mit $\psi_1 = 234°$, $\psi_2 = 187°$ und $\psi_3 = 217°$, $\varphi = (5, 0, 0, 15, 0, 0, 0, 25, 0, 0, 0, 0, 35, 0, 0, 0, 0, 0, 45, 0, 0, 0, 0, 0, 0, 55, 0, 0, 0, 0, 0, 0, 0, 65, 60, 55, 50, 45, 40, 0, 0, 35, 30, 0, 60, 70, 80, 90, 0, 0)$ und $t = (0, 0, 0)$.

10. Transformation:
Es gilt $D := \mathsf{E}_3$, $\varphi = (12, 23, 34, 45, 56, 0, 47, 38, 29, 20, 11, 0, 13, 17, 25, 32, 37, 47, 58, 71, 79, 0, 0, 23, 34, 43, 34, 25, 16, 61, 52, 43, 34, 0, 56, 34, 0, 23, 0, 0, 45, 0, 36, 74, 0, 81, 0, 12, 0, 34)$ und $t = (0.5, -0.5, -0.2)$.

Bilder der transformierten Konformationen:
Zum Abschluß werden Bilder der transformierten Konformationen und der Kristallstruktur von HIV-Protease/1ivq ($n_{ft}[\mathcal{A}] = 25$) dargestellt, damit man einen Überblick über die heuristischen Verfahren in Kapitel 4.2 und im Abschnitt über das Gauß-Seidel-artige Optimierungsverfahren in Kapitel 4.3 bekommt.

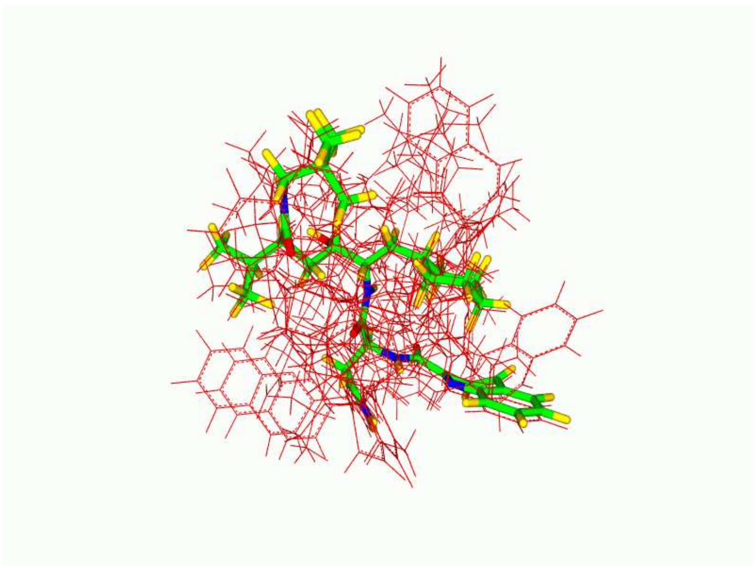


Abbildung A.1: Die 10 transformierten Konformationen (rot) und die Kristallstruktur (grün) des Moleküls 1ivq. Das heuristische Testverfahren in Kapitel 4.2 besteht darin, die 10 roten Konformationen auf die Kristallstruktur zu überlagern.

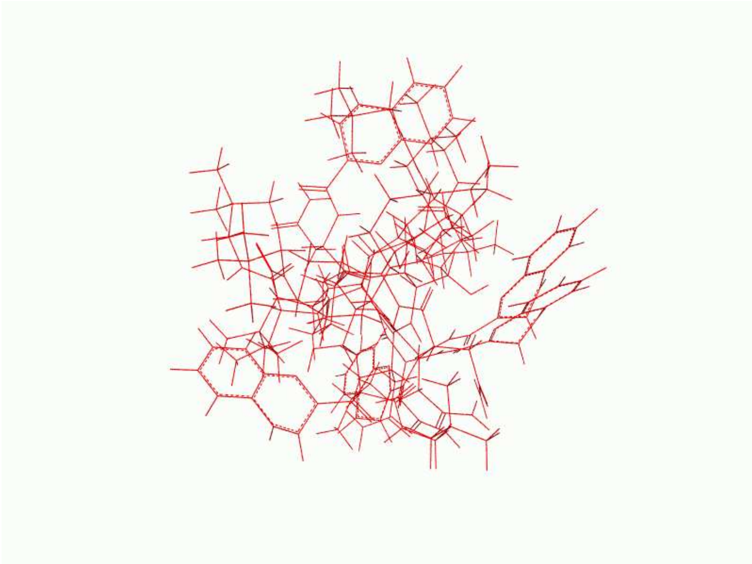


Abbildung A.2: Konformationen vom Typ A des Moleküls 1ivq (2., 4., 6., 8. und 10. transformierte Konformation). Aufgabe in Kapitel 4.3 ist es, die Konformationen gegenseitig zu überlagern.

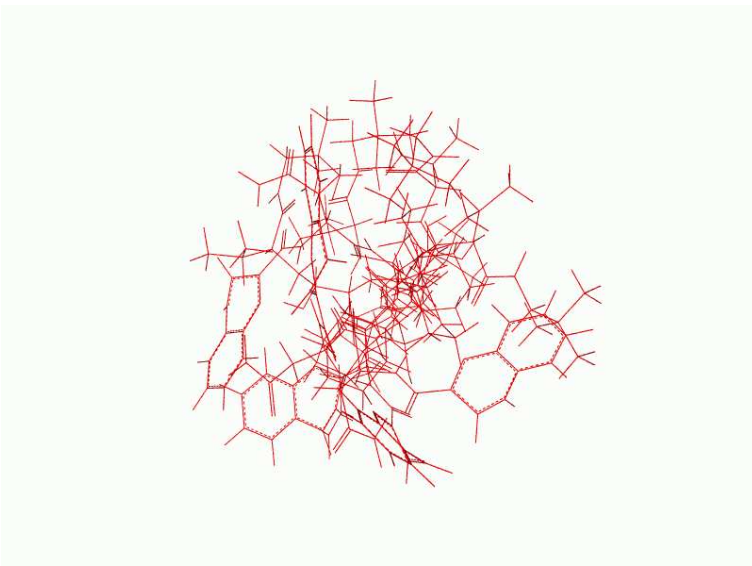


Abbildung A.3: Konformationen vom Typ B des Moleküls 1ivq (1., 3., 5., 7. und 9. transformierte Konformation). Aufgabe in Kapitel 4.3 ist es, die Konformationen gegenseitig zu überlagern.

Anhang B

Symbolverzeichnis

Allgemeine Symbole:

- $\mathsf{S}^{n-1} := \{x \in \mathbb{R}^n \mid \|x\|_2 = 1\}$ Einheitssphäre des $\mathbb{R}^n$.

- $\mathcal{P}_2(V) \ := \ \{\ X \mid X \subseteq V,\ |X| = 2\ \}$ Klasse der 2-elementigen Teilmengen von V.

- Für $x \in \mathbb{R}^n$ ist $\mathsf{P}_{i,n}(x) \ := \ x_i,\ i \ = \ 1,\ \ldots,\ n$ die $i-$te Komponente von x.

- Signum von λ:

$$\mathrm{sig}(\lambda) \ := \ \begin{cases} 1 & \text{falls } \lambda \geq 0 \\ -1 & \text{sonst} \end{cases} .$$

- Kroneckersymbol:

$$\delta_{ij} : \ = \ \begin{cases} 1 & \text{falls } i = j \\ 0 & \text{sonst} \end{cases}$$

- *RMS-Abstand*: Es seien zwei endliche Mengen $M_1 = \{a_1, \ldots, a_n\} \subset \mathbb{R}^3$ und $M_2 = \{b_1, \ldots, b_n\} \subset \mathbb{R}^3$ gegeben. Dann ist der RMS-Abstand der beiden Mengen definiert als $RMS(M_1, M_2) := \sqrt{(1/n)\sum_{i=1}^{n} \|a_i - b_i\|_2^2}$.

- *Umgebung*: Für $\varepsilon > 0$ und $x \in \mathbb{R}^n$ ist die ε-Umgebung von x definiert als $U_\varepsilon(x) := \{y \in \mathbb{R}^n \mid \|x - y\|_2 < \varepsilon\}$.

Norm und Skalarprodukt:

- $\|u\|_2 \ := \ \sqrt{\sum_1^n u_i^2},\ u \in \mathbb{R}^n :$ Euklidische Norm.

- $< u,\ v >_2 := \sum_1^n u_i v_i,\ u, v \in \mathbb{R}^n:$ Euklidisches Skalarprodukt.

- $< f,\ g >_{\mathsf{L}^2(\mathbb{R}^n)} := \int_{\mathbb{R}^n} f(x) g(x)\ \mathsf{d}x,\ f, g \in \mathsf{L}^2(\mathbb{R}^n):$ Skalarprodukt des Hilbertraumes $\mathsf{L}^2(\mathbb{R}^n)$.

- $\|f\|_{\mathsf{L}^2(\mathbb{R}^n)} = \sqrt{< f,\ f >_{\mathsf{L}^2(\mathbb{R}^n)}},\ f \in \mathsf{L}^2(\mathbb{R}^n):$ Norm des Hilbertraumes $\mathsf{L}^2(\mathbb{R}^n)$.

Vektorraum:

- $\mathsf{C}_{-1}(U):$ Vektorraum der auf $U \subset \mathbb{R}$ stückweise stetigen Funktionen nach $\mathbb{R}$.

- $\mathsf{C}_0(U):$ Vektorraum der stetigen Funktionen von $U \subset \mathbb{R}$ nach $\mathbb{R}$.

- $\mathsf{C}_n(U):$ Vektorraum der n-mal stetig differenzierbaren Funtionen von $U \subset \mathbb{R}$ nach $\mathbb{R}$.

- $\mathsf{C}^b(\mathbb{R}^n):$ Vektorraum der stetigen und beschränkten Funtionen von $\mathbb{R}^n$ nach $\mathbb{R}$.

- $\mathsf{C}^m(\mathbb{R}^n):$ Vektorraum der m-mal stetig partiell differenzierbaren Funtionen von $\mathbb{R}^n$ nach $\mathbb{R}$.

- $\mathsf{C}^\infty(\mathbb{R}^n):$ Vektorraum der unendlich oft stetig partiell differenzierbaren Funtionen von $\mathbb{R}^n$ nach $\mathbb{R}$.

- $\mathsf{L}^p(\Omega):$ Vektorraum der Aquivalenzklassen der meßbaren Funktionen $u\colon\ \Omega \subset \mathbb{R}^n \to \mathbb{R}$ mit

$$\int_\Omega |u(x)|^p\ \mathsf{d}x < \infty,\ p \geq 1$$

- $\mathsf{P}(n) := \{\sum_{i=0}^n a_i x^i,\ x \in \mathbb{R} \mid a_i \in \mathbb{R},\ i = 0, \ldots, n,\ \wedge\ a_n \neq 0\}$ Vektorraum der reellwertigen Polynome vom Grad n.

Matrix:

$\mathsf{A}_{ij} = a_{ij}:$ ij-Komponente der Matrix $\mathsf{A} \in \mathsf{M}(n \times m; \mathbb{R})$.

$\mathsf{A}^\mathsf{T} \in \mathsf{M}(m \times n; \mathbb{R}):$ Transponierte Matrix von $\mathsf{A} \in \mathsf{M}(n \times m; \mathbb{R})$.

$\mathsf{E}_n \in \mathsf{M}(n \times n; \mathbb{R}):$ Einheitsmatrix.

$\mathsf{M}(n \times m; \mathbb{R}):$ Vektorraum der $n \times m-$ Matrizen.

$$\mathsf{M}(n \times m; \mathbb{R}) := \left\{ \begin{pmatrix} a_{11} & \ldots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \ldots & a_{nm} \end{pmatrix} \mid a_{ij} \in \mathbb{R},\ i = 1, \ldots, n,\ j = 1, \ldots, m \right\}$$

$\mathsf{SO}_n:$ Spezielle orthogonale Gruppe (Klasse der Rotationsmatrizen).

$$\mathsf{SO}_n := \{\mathsf{A} \in \mathsf{M}(n \times n; \mathbb{R}) \mid \mathsf{A} * \mathsf{A}^\mathsf{T} = \mathsf{E}_n,\ \det \mathsf{A} = 1\}$$

tr $\mathsf{A} := \sum_1^n a_{ii}$: Spur der Matrix $\mathsf{A} \in \mathsf{M}(n \times n; \mathbb{R})$.

Moleküldarstellung: siehe Kapitel 2.2

- $\mathcal{A}$ sei ein Molekül mit n_a Atomen.
- $\mathsf{C}_\mathcal{A} \in \mathbb{R}^{3n_a}$ Molekülkonformation.
- $K_\mathcal{A} \subset \mathbb{R}^3$ Menge der Atomkoordinaten bzgl. der Konformation $\mathsf{C}_\mathcal{A}$.
- $B_\mathcal{A} \subset \mathcal{P}_2(K_\mathcal{A})$ Menge der Atombindungen.
- $G_\mathcal{A} = (K_\mathcal{A},\ B_\mathcal{A})$ Molekülgraph.
- $\Phi_{EB}^\mathcal{A}$ Abbildung auf $B_\mathcal{A}$, die die Einfach-Bindungen charakterisiert.
- $\Phi_{FT}^\mathcal{A}$ Abildung auf $B_\mathcal{A}$, die diejenigen Bindungen charakterisiert, an denen der Torsionswinkel variiert werden kann.
- $B_{FT}^\mathcal{A}$ Menge der Bindungen, an denen der Torsionswinkel variiert werden kann.

Lokale Rotations- und Transformationsabbildungen:

- $\mathcal{B}_V(T)$ eine Baumstruktur auf V bzgl. des Baumes T (Definition 2.4.13).
- $\mathcal{S}(V, \vec{T})$ eine Transformationsstruktur auf V bzgl. des gerichteten Baumes $\vec{T}$. (Definition 2.4.14).
- $\mathcal{R}[\mathcal{S}(V, \vec{T})]$ lokaler Rotationsoperator von V bzgl. der Transformationsstruktur $\mathcal{S}(V, \vec{T})$ (Definition 2.4.18).
- $\mathcal{T}[\mathcal{S}(V, \vec{T})]$ Transformationsoperator von V bzgl. der Transformationsstruktur $\mathcal{S}(V, \vec{T})$ (Definition 2.4.27).
- $\mathsf{B}\,(a,\ \varphi)\,,\ a \in \mathbb{R}^3, \varphi \in \mathbb{R}$ Matrix aus Satz 2.4.3.
- $\mathcal{B}^n\,(\dots)\,,\ n \in \mathbb{N}$ Rotationsmatrix aus Definition 2.4.10.
- $\mathsf{R}[\mathcal{S}(V, \vec{T}), f_i]$ Abbildung aus Satz 2.4.16.

Anhang C

Verwendte mathematische Definitionen und Sätze

Die in diesem Kapitel aufgelisteten Aussagen stammen aus den in der Referenz angegebenen Büchern.

Definition C.1 (Graph, Weg, Kreis, Baum, Digraph)
Ein *Graph* ist ein Paar $G = (V, E)$ disjunkter Mengen mit $E \subseteq \mathcal{P}_2(V) := \{X \mid X \subseteq V,\ |X| = 2\}$; die paarweise verschiedenden Elemente von E sind also 2-elementige Teilmengen von V. Die Elemente von V nennt man *Knoten* des Graphen G, die Elemente von E seine *Kanten*. Für V schreibt man auch $V(G)$, für E auch $E(G)$. Der Graph G heißt *endlich* bzw. *unendlich* je nachdem, ob V endlich oder unendlich ist. Ein Knoten v heißt mit einer Kante e *inzident*, wenn $v \in e$ gilt. Die beiden mit einer Kante e inzidenten Knoten sind ihre Endknoten, und e *verbindet* diese Knoten. Für eine Kante $\{x, y\}$ schreibt man kürzer auch xy (oder yx). Der Grad von $v \in V$ gibt die Anzahl der mit v in G inzidenten Kanten an, d.h. es gilt $\mathsf{d}_G(v) := |\{e \in E | v \in e\}|$. Ein *Weg* ist ein nicht leerer Graph $P = (V, E)$ der Form $V = \{x_0, x_1, \ldots, x_k\}$ und $E = \{x_0x_1, x_1x_2, \ldots, x_{k-1}x_k\}$, wobei die x_i paarweise verschieden sind. Die Anzahl der Kanten eines Weges P ist seine Länge $l(P)$. Man bezeichnet einen Weg oft durch die natürliche Folge seiner Knoten, also z.B. $P = x_0x_1 \ldots x_k$ (oder $P = x_kx_{k-1} \ldots x_0$); man nennt P dann einen x_0-x_k-Weg (bzw. einen x_k-x_0-Weg). Ist $P = x_0 \ldots x_{k-1}$ $(= (V = \{x_0, \ldots, x_{k-1}\}, E = \{x_0x_1, \ldots, x_{k-2}x_{k-1}\})$ ein Weg und $k \geq 3$, so ist $C := P + x_{k-1}x_0$ $(:= (V, E \cup \{x_{k-1}, x_0\}))$ ein *Kreis*. Schreibweise $C = x_0 \ldots x_{k-1}x_0$. Die Länge $l(C)$ eines Kreises C ist die Anzahl seiner Kanten. Es seien $G = (V, E)$ und $G' = (V', E')$ zwei Graphen. G' ist in G *enthalten* ($G' \subseteq G$), wenn $V' \subseteq V$ und $E' \subseteq E$ gilt. G' ist dann ein Teilgraph von G. Ein Graph $G = (V, E)$ heißt *zusammenhängend*, wenn er für je zwei seiner Knoten $x,\ y \in V$ einen x-$y-$Weg enthält. Die nichtleeren und maximal zusammenhängenden Teilgraphen $G_i = (V_i, E_i)$ eines Graphen $G = (V, E)$, $V = \bigcup V_i$, $E = \bigcup E_i$, heißen Zusammenhangskomponenten des Graphen G. Dabei heißt maximal zusammenhängend, daß G für $x_i \in V_i,\ x_j \in V_j,\ i \neq j$, keinen x_i-x_j-Weg enthält. Ein Graph, der keinen Kreis enthält, ist ein

Wald. Ein zusammenhängender Wald ist ein *Baum.*

Ein *gerichteter Graph* bzw. *Digraph* $\vec{G} = (V, \vec{E})$ besteht aus einer disjunkten Knotenmenge V und eine Menge $\vec{E} \subseteq V$xV von geordneten Paaren, die man gerichtete Kanten nennt. Für $(x, y) \in \vec{E}$ heißt x der *Anfangsknoten* und y der *Endknoten.* Ein gerichteter Weg $\vec{P}$ ist eine Folge von paarweise verschiedenen Knoten $x_0, x_1, \ldots, x_n$ mit (x_i, x_{i+1}) für alle i, also $\vec{P} = (\{x_0, \ldots, x_n\}, \{(x_0, x_1), (x_1, x_2), \ldots, (x_{n-1}, x_n)\})$. Schreibweise für den x_0-x_n-Weg: $\vec{P} = (x_0 \ldots x_n)$. Die Länge $l(\vec{P})$ des Weges ist die Anzahl der gerichteten Kanten. Analog wird ein gerichteter Kreis definiert: $\vec{C} = (\{x_0, \ldots, x_n\}, \{(x_0, x_1), \ldots, (x_{n-1}, x_n), (x_n, x_0)\})$, $n \geq 2$. Schreibweise: $\vec{C} = (x_0 \ldots x_n x_0)$. Ein gerichtete Graph $\vec{G}$ heißt *azyklisch*, wenn er keinen gerichteten Kreis $\vec{C}$ enthält. Ein gerichteter Graph $\vec{G} = (V, \vec{E})$ ist *zusammenhängend*, wenn der zugrundeliegende ungerichtete Graph $G = (V, E := \{\{x_i, x_j\} \mid (x_i, x_j) \in \vec{E}\})$ zusammenhängend ist. Ein gerichteter Graph $\vec{T} = (V, \vec{E})$ heißt *gerichteter Baum*, wenn der zugrundeliegende ungerichtete Graph $T = (V, E)$ ein Baum ist. Es sei $\vec{T} = (V, \vec{E})$ ein gerichteter Baum. Weiter seien $x,\ y \in V$. Dann heißt y *Nachfolger* von x in $\vec{T}$, wenn $\vec{T}$ einen gerichteten Weg x-$y-$Weg enthält. Ein Knoten $v_0 \in V$ heißt *Wurzel* des gerichteten Baumes $\vec{T} = (V, \vec{E})$, wenn $\vec{T}$ für alle Knoten $v \in V,\ v \neq v_0$, einen gerichteten v_0-$v-$Weg enthält. Nach Satz C.2 2) ist dieser gerichtete v_0-$v-$Weg, $v \neq v_0$, eindeutig bestimmt. Ein gerichteter Baum $\vec{T}$ kann höchstens eine Wurzel besitzen, da T kreisfrei ist.

Referenz: [13], [1], [9] $\diamond$

Satz C.2 (Eigenschaften von Bäumen)

Es sei $T = (V, E)$ ein ungerichteter Graph. Dann sind die folgenden Aussagen äquivalent.

1) T ist ein Baum.

2) Für je zwei Knoten $x,\ y$ aus V enthält T genau ein x-$y-$Weg.

3) T ist minimal zusammenhängend, d.h. für je zwei Knoten $x,\ y \in V$ mit $\{x, y\} \in E$ gilt: $(V, E \setminus \{x, y\})$ ist nicht zusammenhängend.

4) T ist zusammenhängend und $|E| = |V| - 1$.

5) T ist kreisfrei und $|E| = |V| - 1$.

6) T ist maximal kreisfrei, d.h. für je zwei Knoten $x,\ y \in V$ mit $\{x, y\} \notin E$ enthält $(V, E \cup \{x, y\})$ einen Kreis.

Referenz: [13], [1], [9] ■

Satz C.3 (Ungleichung zwischen geometrischen und arithmetrischen Mittel)

Es seien $a, b \in \mathbb{R}_{\geq 0}$. Dann gilt $\sqrt{ab} \leq \frac{1}{2}(a + b)$. Referenz: [17] ■

Satz C.4 (Cauchy-Schwarzsche Ungleichung)

Es sei $(\mathcal{X}, <.,.>)$ ein Vektorraum mit innerem Produkt $<.,.>$. Weiter sei $\|.\| := \sqrt{<.,.>}$ die durch das innere Produkt induzierte Norm. Dann gilt

$$|<u,\ v>| \ \leq\ \|u\| * \|v\| \quad \forall u, v \in \mathcal{X}.$$

Referenz: [50]

■

Satz C.5 (Minkowskische Ungleichung)
Es sei $f, g \in \mathsf{L}^p(\Omega)$ mit $1 \leq p < \infty$. Dann gilt

$$\left[\int\limits_{\Omega} |f \pm g|^p \, \mathsf{d}x\right]^{\frac{1}{p}} \leq \left[\int\limits_{\Omega} |f|^p \, \mathsf{d}x\right]^{\frac{1}{p}} + \left[\int\limits_{\Omega} |g|^p \, \mathsf{d}x\right]^{\frac{1}{p}}.$$

Referenz: [50]

■

Satz C.6
Es seien $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$ und $(\mathcal{Y}, \|\cdot\|_{\mathcal{Y}})$ zwei normierte Räume. Weiter sei $T\colon \mathcal{X} \to \mathcal{Y}$ eine lineare Abbildung. T ist genau dann in $\mathcal{X}$ stetig, wenn T beschränkt ist, d.h. es existiert ein $K > 0$, so daß $\|Tu\|_{\mathcal{Y}} \leq K\|u\|_{\mathcal{X}}\ \forall u \in \mathcal{X}$ gilt.
Referenz: [50]

■

Satz C.7 (Transformationsformel)
Es seien $U \subset \mathbb{R}^n$ und $V \subset \mathbb{R}^n$ offen, $\varphi\colon\ U \to V$ ein $\mathcal{C}_1$-Diffeomorphismus und $J_\varphi = \varphi'$ die Jacobi-Matrix von φ. Dann ist $f\colon\ V \to \mathbb{R}$ integrierbar genau dann, wenn $(f \circ \varphi) \cdot |\det \varphi'| : U \to \mathbb{R}$ integrierbar ist. Es gilt

$$\int\limits_V f(y)\, \mathrm{d}y = \int\limits_U (f \circ \varphi)(x) \cdot |\det J_\varphi(x)|\, \mathrm{d}x.$$

Referenz: [40]

■

Satz C.8
Es sei $A \in \mathsf{M}(n \times n; \mathbb{R})$ eine orthogonale Matrix, d.h. $A^\mathsf{T} * A = E_n$. Dann gilt $\|Ax\|_2 = \|x\|_2\ \forall x \in \mathbb{R}^n$. Insbesondere gilt diese Eigenschaft für Rotationsmatrizen $A \in \mathsf{SO}_3$, da $A \in \mathsf{SO}_3 \Leftrightarrow (A^\mathsf{T} * A = E_3 \wedge \det A = 1)$.
Referenz: [51]

■

Satz C.9 (Charakerisierung der SO_3)
Für alle $\mathsf{A} \in \mathsf{SO}_3$ gibt es einen Einheitsvektor $a \in \mathbb{R}^3$, so daß

$$\mathsf{A}y = y + (\sin\varphi)\ a \times y + (1 - \cos\varphi)\ a \times (a \times y) \quad \forall y \in \mathbb{R}^3$$

gilt. Dabei ist a ein Einheitsvektor in Richtung der zu A gehörenden Drehachse und $\varphi\ in\ [0;\ 2\pi)$ der zu A gehörende Drehwinkel.
Referenz: [51]

■

Satz C.10 (Rechenregel für das Kreuzprodukt)
Das *Kreuzprodukt* von u und v aus $\mathbb{R}^3$ ist der Vektor

$$u \times v \ := \begin{pmatrix} u_2v_3 & - & u_3v_2 \\ u_3v_1 & - & u_1v_3 \\ u_1v_2 & - & u_2v_1 \end{pmatrix}.$$

Für $u, v, w \in \mathbb{R}^3$ und $\alpha, \beta \in \mathbb{R}$ gilt

1) $(\alpha u)\mathsf{x}w \ = \ u\mathsf{x}(\alpha w) \ = \ \alpha(u\mathsf{x}w) \ =: \ \alpha u\mathsf{x}w$
   $(\alpha u + \beta v)\mathsf{x}w \ = \ (\alpha u\mathsf{x}w) + (\beta v\mathsf{x}w) \ =: \ \alpha u\mathsf{x}w + \beta v\mathsf{x}w$
   $w\mathsf{x}(\alpha u + \beta v) \ = \ \alpha w\mathsf{x}u + \beta w\mathsf{x}v$

2) $u\mathsf{x}v \ = \ -v\mathsf{x}u$ und $u\mathsf{x}u \ = \ 0$

3) Für $a, b, c \in \mathbb{R}^3$ gilt die *„baccab-Regel“* $a\mathsf{x}(b\mathsf{x}c) \ = \ b < a, \ c >_2 - c < a, \ b >_2 = \ (c\mathsf{x}b)\mathsf{x}a$.

Referenz: [51]

■

Satz C.11
Die Funktionen $f, g\colon D \to \mathbb{R}$ seien in $x \in D \subset \mathbb{R}$ differenzierbar. Weiter sei $\lambda \in \mathbb{R}$. Dann sind auch die Funktionen $f + g, \ \lambda f, \ fg\colon D \to \mathbb{R}$ in $x \in D$ differenzierbar und es gilt $(f+g)'(x) = f'(x) + g'(x)$, $(\lambda f)'(x) = \lambda f'(x)$ und $(fg)'(x) = f'(x)g(x) + f(x)g'(x)$ (Produktregel).
Gilt $g(y) \neq 0$ für alle $y \in D$, so ist auch die Funktion $f/g\colon D \to \mathbb{R}$ in $x \in D$ differenzierbar und es gilt

$$\left(\frac{f}{g}\right)'(x) = \frac{f'(x)g(x) - f(x)g'(x)}{(g(x))^2}$$

(Quotientenregel).
Es sei $f\colon D \to \mathbb{R}$ und $g\colon E \to \mathbb{R}$ mit $f(D) \subset E$. Die Funktion f sei im Punkt $x \in D$ und die Funktion g sei in $y := f(x) \in E$ differenzierbar. Dann ist auch die Komposition $g \circ f\colon D \to \mathbb{R}$ im Punkt $x \in D$ differenzierbar und es gilt $(g \circ f)'(x) = g'(f(x))f'(x)$ (Kettenregel).
Referenz: [17]

■

Bemerkung: f sei in $x \in D$ differenzierbar. Ist f' in x stetig, so ist f in x stetig differenzierbar.

Definition C.12 (Partielle Differentation)
Es sei $U \subset \mathbb{R}^n$ eine offene Menge und $f\colon U \to \mathbb{R}$ eine reellwertige Funktion. f heißt im Punkt $x \in U$ partiell differenzierbar bzgl. der i-ten Koordinatenrichtung, falls der Limes

$$\mathsf{D}_i f(x) \ : \ = \ \lim_{h \to 0} \ \frac{f(x + he_i) \ - \ f(x)}{h}$$

existiert. Dabei ist $e_i \in \mathbb{R}^n$ der i-te Einheitsvektor und für den Limes $h \to 0$ hat man sich auf die $h \in \mathbb{R}$ zu beschränken, für die $h \neq 0$ und $x + he_i \in U$ gilt. $\mathsf{D}_i f(x)$ heißt i-te partielle Ableitung von f in x.
Referenz: [16]

◇

Bemerkung: Die partiellen Ableitungen einer Funktion $f\colon U \to \mathbb{R}$ kann man als gewöhnliche Ableitungen von Funktionen einer Veränderlichen interpretieren: Es sei $x = (x_1, \dots, x_n) \in U$ ein fester Punkt. Für $i = 1, \dots, n$ betrachtet man die Funktionen $\xi \to f_i(\mathsf{x}_i) := f(x_1, \dots, x_{i-1}, \xi, x_{i+1}, \dots, x_n)$. Aus der Definition der partiellen Ableitung folgt

$$D_i f(x) = \lim_{h \to 0} \frac{f_i(x_i + h) - f_i(x_i)}{h} = f_i'(x_i).$$

Die partielle Ableitung bzgl. der i-ten Koordinatenrichtung ist also nichts anderes als die gewöhnliche Ableitung bzgl. der i-ten Variable bei Festhaltung der übrigen n-1 Veränderlichen. Deshalb gelten für die partiellen Ableitungen analoge Rechenregeln wie für die gewöhnlichen Ableitungen (siehe Satz C.11).

Definition C.13
Es sei $U \subset \mathbb{R}^n$ eine offene Menge. Eine Funktion $f\colon U \to \mathbb{R}$ heißt partiell differenzierbar, falls $\mathsf{D}_i f(x)$ für alle $x \in U$ und $i = 1, \dots, n$ existiert. f heißt stetig partiell differenzierbar, falls zusätzlich alle partiellen Ableitungen $\mathsf{D}_i f(x)\colon U \to \mathbb{R}$, $i = 1, \dots, n$, stetig sind.
Referenz: [16] ◇

Definition C.14
Es sei $U \subset \mathbb{R}^n$ eine offene Menge und $f\colon U \to \mathbb{R}^m$ eine Abbildung. f heißt im Punkt $x \in U$ *total differenzierbar* (oder *differenzierbar* schlechthin), falls es eine Matrix $A \in \mathsf{M}(m \times n; \mathbb{R})$ gibt, so daß in einer Umgebung von x $f(x + \xi) = f(x) + A\xi + \varphi(\xi)$ gilt, wobei φ eine in einer Umgebung von Null des $\mathbb{R}^n$ definierte Funktion mit Werten in $\mathbb{R}^m$ ist und es gilt $\lim_{\xi \to 0} \frac{\varphi(\xi)}{\|\xi\|} = 0$.
Referenz: [16] ◇

Schreibweise: Ist φ eine in einer Umgebung des Nullpunktes des $\mathbb{R}^n$ definierte Funktion mit $\varphi(0) = 0$ und $\lim_{\xi \to 0, \xi \neq 0} \frac{\varphi(\xi)}{\|\xi\| = 0}$, so verwendet man die symbolische Schreibweise $\varphi(\xi) = o(\|\xi\|)$. Damit schreibt sich dann die Bedingung für die Differenzierbarkeit $f(x+\xi) = f(x) + A\xi + o(\|\xi\|)$.

Satz C.15
Es sei $U \subset \mathbb{R}^n$ eine offene Menge und $f\colon U \to \mathbb{R}^m$ eine Abbildung, die im Punkt $x \in U$ differenzierbar sei und zwar gelte $f(x+\xi) = f(x) + A\xi + o(\|\xi\|)$ mit der Matrix $A \in \mathsf{M}(m \times n; \mathbb{R})$. Dann gilt:

a) f ist im Punkt x stetig.

b) Alle Komponentenfunktionen $f_i\colon U \to \mathbb{R}$, $1 \leq i \leq m$, von f sind in x partiell differenzierbar mit

$$\frac{\partial f_i}{\partial x_j}(x) = a_{ij}.$$

Referenz: [16]

■

Bemerkung: Aus b) folgt insbesondere, daß die Matrix A durch die differenzierbare Abbildung f eindeutig bestimmt ist. Man nennt A das *Differential* oder die *Jacobi-Matrix* (oder auch die *Funktionalmatrix*) von f im Punkt x. Schreibweise:

$$f'(x) := (\mathsf{D}\ f)(x) := \left(\frac{\partial f_i}{\partial x_j}\ (x)\right)_{\substack{1\leq i\leq m\\ 1\leq j\leq n}}.$$

Definition C.16 (C^1–Funktionen)

Es sei $U \subset \mathbb{R}^n$ eine offene Menge. Dann bezeichnet $\mathsf{C}^1(U) := \{h\colon U \to \mathbb{R} \mid h$ ist stetig partiell differenzierbar in $U\}$ die Menge der stetig partiell differenzierbaren Funktionen auf U. Liegt U fest, so nennt man $h \in \mathsf{C}^1(U)$ eine C^1-Funktion.
Nun sei $f\colon U \subset \mathbb{R}^n \to \mathbb{R}^m$, $m > 1$, U offen, eine Abbildung. Es seien $f_1, \dots, f_m\colon U \to \mathbb{R}$ die reellwertigen Komponentenfunktionen von f, d.h. $f(x) = (f_1(x), \dots, f_m(x))^\mathsf{T}$, $x \in U$. Man nennt f eine C^1-Funktion, wenn $f_j \in \mathsf{C}^1(U)$, $j = 1, \dots, m$, gilt.
Referenz: [31] ◇

Satz C.17

Eine C^1-Funktion $f\colon U \subset \mathbb{R}^n \to \mathbb{R}^m$ (U offen) ist in jedem Punkte von U differenzierbar und somit auch stetig.
Referenz: [31]

■

Satz C.18

Die Funktionen $f, g\colon U \subset \mathbb{R}^p \to \mathbb{R}^q$, U offen, seien in $x \in U$ differenzierbar. Dann sind auch $f + g$ und αf, $\forall \alpha \in \mathbb{R}$, in x differenzierbar, und es ist $(f+g)'(x) = f'(x) + g'(x)$ und $(\alpha f)'(x) = \alpha f'(x)$.
Referenz: [31]

■

Satz C.19 (Kettenregel)

Es sei $g\colon U \subset \mathbb{R}^p \to \mathbb{R}^q$, U offen, und $f\colon V \subset \mathbb{R}^q \to \mathbb{R}^r$, V offen, mit $g(U) \subset V$. Weiter sei g in $x \in U$ und f in $g(x) \in V$ differenzierbar. Dann ist $f \circ g$ in $x \in U$ differenzierbar und es gilt $(f \circ g)'(x) = f'(g(x))g'(x)$.
Referenz: [31]

■

Definition C.20 (Differentation in Banachräumen)

Es seien X und Y Banachräume und $G \subset X$ offen. Die Funktion $f: G \to Y$ heißt differenzierbar im Punkt $\xi \in G$, wenn es eine stetige lineare Abbildung $A\colon X \to Y$ gibt, so daß für alle $\xi + h$ in einer ε-Umgebung $U \subset G$ von ξ gilt $f(\xi + h) - f(\xi) = Ah + r(h)$ mit $\lim_{h\to 0} \frac{r(h)}{\|h\|} = 0$.
Referenz: [31] ◇

Definition C.21

Ist die Funktion $f: G \to Y$ (G offen) in $\xi \in G$ differenzierbar, so wird die eindeutig bestimmte Abbildung $A\colon X \to Y$ aus Definition C.20 die *Ableitung* oder auch *Fréchetsche Ableitung* von

f an der Stelle ξ genannt und mit $f'(\xi)$ oder $\mathsf{D}\ f(\xi)$ bezeichnet.
Referenz: [31] $\diamond$

Anhang D

Tabellen

Die Erklärungen zu den einzelnen Spalten der Tabelle D.1 sind im Abschnitt über das heuristische Testverfahren in Kapitel 4.2 aufgeführt und entsprechen den Erklärungen zur Tabelle 4.2 (Seite 105).

In der Tabelle D.2 werden die Ergebnisse der Optimierungen dargestellt, die mit dem inversen BFGS-Verfahren mit der Aufdatierungsformel (4.10) und der Armijoregel mit Aufweitung, (Parameter $\beta = 0.35$) (BFGS-INV1/ARA(0.35)) erzielt worden sind. In der Tabelle D.3 werden die Ergebnisse der Optimierungen, die mit dem BFGS-INV1/ARA(0.35)-Verfahren bzw. mit dem Hierarchie-Verfahrens 4.7 aus Kapitel 4.3 erreicht worden sind, gegenübergestellt. Die 77 Moleküle aus Anhang A sind numeriert worden. Die Numerierung richtet sich dabei nach der Anzahl der Bindungen, an denen der Torsionswinkel variiert werden kann. Für jedes der 77 Moleküle wird der durchschnittliche Alignmentindex (10 Testkonformationen bzgl. der Kristallstruktur) vor den Optimierungen und nach den Optimierungen dargestellt, d.h. es gilt für $i = 1, \ldots, 77$:

$$\mathsf{S}^0_{\mathsf{R}_i} := \frac{1}{10} \sum_{j=1}^{10} \mathsf{S}_{\mathcal{A}_{j,i},\mathcal{R}_i}(u^0_{\mathcal{A}_{j,i}}) \quad \text{und} \quad \mathsf{S}^*_{\mathsf{R}_i} := \frac{1}{10} \sum_{j=1}^{10} \mathsf{S}_{\mathcal{A}_{j,i},\mathcal{R}_i}(u^*_{\mathcal{A}_{j,i}}), \tag{D.1}$$

wobei für $i = 1, \ldots, 77$ und $j = 1, \ldots, 10$ $\mathcal{R}_i$ die Kristallstruktur des Molekül i und $\mathcal{A}_{j,i}$ die j-te transformierte Konformation darstellt. Dabei ist $\mathsf{S}_{\mathcal{A}_{j,i},\mathcal{R}_i}$ der Index aus Gleichung (3.5) in Definition 3.2.2. Das gewählte Optimierungsverfahren verwendet den Transformationsparameter $u^0_{\mathcal{A}_{j,i}}$ als Startwert und liefert $u^*_{\mathcal{A}_{j,i}}$ als Ergebnis. Dann wird die Dimension (Dim) des jeweiligen Optimierungsproblems angegeben, d.h. 10 plus Anzahl der internen Rotationsparameter, wobei sich die 10 durch den Translationsvektor, den Quaternionenvektor und durch das globale Rotationszentrum ergibt. Weiter wird die durchschnittliche Anzahl der Iterationen und die durchschnittliche Rechenzeit in Sekunden angegeben, die das Verfahren gebraucht hat, um eine der 10 Energiefunktion pro Molekül zu minimieren. In Spalte Stopp wird dann die durchschnittliche Bewertung der Stoppkriterien angegeben. Die Stoppkriterien und deren Bewertung sind im Abschnitt über das heuristische Testverfahren in Kapitel 4.2 angegeben worden. In der Spalte Q ist die durchschnittliche Anzahl der Iterationen angegeben, die die Schrittweitenregel brauchte, um ein $\alpha_k > 0$ zu bestimmen. Und in der Spalte N wird die durchschnittliche

Anzahl der Neustarts angegeben, die bei der Aufdatierung der Matrix H_k zu H_{k+1} ausgeführt worden sind. Die Spalten mit den Ergebnissen des Hierarchie-Verfahrens sind mit einem H gekennzeichnet.

In den Tabellen D.4 und D.5 werden die Ergebnisse der Optimierungen, die mit dem BFGS-INV1/ARA(0.35)-Verfahren bzw. mit dem Gauß-Seidel-artigem Optimierungsverfahren 4.6 aus Kapitel 4.3 erreicht worden sind, gegenübergestellt. Das gegenseitige Alignment der Konformationen vom Typ A bzw. vom Typ B ist optimiert worden. Die Konformationen vom Typ A bestehen aus der 2., 4., 6., 8. und 10. transformierten Molekülstruktur des jeweiligen Moleküls, während die Konformationen vom Typ B aus der 1., 3., 5., 7. und 9. transformierten Molekülstruktur bestehen. Die entsprechenden Transformationen sind am Ende von Anhang A beschrieben worden. In der Tabelle D.4 werden die Ergebnisse für die gegenseitige Überlagerung der Konformationen vom Typ A dargestellt, während in der Tabelle D.5 die Ergebnisse für die Überlagerung der Konformationen vom Typ B dargestellt werden. Es werden die Werte vor und nach der jeweiligen Optimierung dargestellt, d.h. es gilt für $i = 1, \ldots, 77$:

$$\mathsf{S}^0_{\mathcal{A}^i_1,\ldots,\mathcal{A}^i_5} := \mathsf{S}_{\mathcal{A}^i_1,\ldots,\mathcal{A}^i_5}(u^0_{\mathcal{A}^i_1},\ldots,u^0_{\mathcal{A}^i_5}) \quad \text{und} \quad \mathsf{S}^*_{\mathcal{A}^i_1,\ldots,\mathcal{A}^i_5} := \mathsf{S}_{\mathcal{A}^i_1,\ldots,\mathcal{A}^i_5}(u^*_{\mathcal{A}^i_1},\ldots,u^*_{\mathcal{A}^i_5}),$$

wobei $\mathsf{S}_{\mathcal{A}^i_1,\ldots,\mathcal{A}^i_5}$ die Summe der Alignmentindizes aus Gleichung (3.9) in Definition 3.3.3 ist. Die beiden Optimierungsverfahren verwenden den Transformationsparameter $(u^0_{\mathcal{A}^i_1},\ldots,u^0_{\mathcal{A}^i_5})$ als Startwert und liefern jeweils $(u^*_{\mathcal{A}^i_1},\ldots,u^*_{\mathcal{A}^i_5})$ als Ergebnis. Dann wird die Dimension (Dim) des jeweiligen Optimierungsproblems angegeben, d.h. 5 mal (10 + Anzahl der internen Rotationsparameter). Weiter wird die Rechenzeit in Sekunden angegeben, die das Verfahren gebraucht hat, um die entsprechende Energiefunktion aus Definition 3.3.3 zu minimieren. In Spalte Stopp wird dann das Stoppkriterium angegeben. Die Stoppkriterien und deren Bewertung sind im Abschnitt über das Gauß-Seidel-artige Optimierungsverfahren aus Kapitel 4.3 beschrieben worden. Die Spalten mit den Ergebnissen des Gauß-Seidel-artigen Verfahrens sind mit GS gekennzeichnet.

Algorithmus	S_E^0	S_E^*	Schritte	Sec	Stopp	Q	N
G-AGS(0.90)	0.102	0.575	98.48	64.32	-0.01	16.39	
G-AR(0.90)	0.102	0.593	99.48	137.55	0.00	39.67	
G-ARA(0.90)	0.102	0.598	99.48	72.56	0.00	39.86	
G-G(0.90)	0.102	0.566	93.83	73.20	-0.03	37.62	
CG-PR-AGS(0.15)	0.102	0.503	79.08	29.13	0.11	2.26	
CG-PR-AR(0.05)	0.102	0.408	39.48	12.65	0.38	3.66	
CG-PR-ARA(0.05)	0.102	0.409	39.51	12.69	0.38	3.77	
CG-PR-G(0.40)	0.102	0.230	6.52	3.19	0.50	18.65	
CG-PRN-AGS(0.15)	0.102	0.516	81.70	62.14	-0.06	2.06	
CG-PRN-AR(0.25)	0.102	0.520	51.80	37.10	-0.19	5.64	
CG-PRN-ARA(0.25)	0.102	0.525	52.50	36.92	-0.19	5.74	
CG-PRN-G(0.40)	0.102	0.242	6.69	7.65	-0.46	10.01	
CG-FR-AGS(0.40)	0.102	0.387	83.10	57.22	-0.06	2.51	
CG-FR-AR(0.55)	0.102	0.426	78.20	64.13	-0.09	13.75	
CG-FR-ARA(0.40)	0.102	0.425	73.19	55.19	-0.13	9.92	
CG-FR-G(0.60)	0.102	0.294	36.10	31.64	-0.34	12.84	
CG-FRN-AGS(0.05)	0.102	0.497	87.05	52.67	-0.03	1.34	
CG-FRN-AR(0.50)	0.102	0.551	80.51	68.54	-0.08	9.75	
CG-FRN-ARA(0.65)	0.102	0.552	78.42	73.30	-0.11	15.21	
CG-FRN-G(0.60)	0.102	0.315	38.11	30.96	-0.32	11.92	
CG-DN-AGS(0.10)	0.102	0.507	83.39	30.08	0.09	1.88	
CG-DN-AR(0.15)	0.102	0.512	78.33	28.35	0.14	3.83	
CG-DN-ARA(0.15)	0.102	0.511	77.42	28.34	0.14	3.94	
CG-DN-G(0.85)	0.102	0.318	38.05	52.64	-0.33	13.97	
BFGS-AGS(0.10)	0.102	0.645	82.58	33.32	0.13	1.32	2.26
BFGS-AR(0.70)	0.102	0.654	80.28	33.45	0.31	3.50	1.32
BFGS-ARA(0.35)	0.102	0.653	73.39	35.05	0.37	5.83	0.87
BFGS-G(0.95)	0.102	0.452	13.75	9.92	-0.49	18.70	0.05
BFGS-INV1-AGS(0.85)	0.102	0.631	67.86	40.25	0.16	7.10	0.65
BFGS-INV1-AR(0.70)	0.102	0.654	80.43	33.54	0.31	3.54	1.37
BFGS-INV1-ARA(0.35)	0.102	0.654	76.81	31.70	0.34	2.88	1.31
BFGS-INV1-G(0.95)	0.102	0.452	13.72	9.83	-0.48	18.67	0.04
DFP-INV-AGS(0.90)	0.102	0.632	80.79	54.90	0.03	12.62	0.43
DFP-INV-AR(0.65)	0.102	0.621	96.84	33.46	0.04	1.54	1.03
DFP-INV-ARA(0.35)	0.102	0.623	95.16	34.81	0.08	2.79	0.86
DFP-INV-G(0.85)	0.102	0.467	71.38	27.32	-0.15	6.12	0.10
OREN-LUEN-AGS(0.90)	0.102	0.632	67.22	47.20	0.03	10.85	0.65
OREN-LUEN-AR(0.75)	0.102	0.636	81.26	32.77	0.29	2.73	1.59
OREN-LUEN-ARA(0.25)	0.102	0.635	84.47	32.99	0.27	2.34	1.70
OREN-LUEN-G(0.70)	0.102	0.465	58.95	37.79	0.34	15.38	0.60
BROY-INV(0.25)-AGS(0.90)	0.102	0.631	65.43	46.04	-0.32	9.93	0.61
BROY-INV(0.25)-AR(0.65)	0.102	0.645	89.93	34.55	0.20	2.19	1.19
BROY-INV(0.25)-ARA(0.40)	0.102	0.641	80.36	35.07	0.31	3.17	1.21
BROY-INV(0.25)-G(0.65)	0.102	0.590	82.64	63.50	-0.20	7.86	0.78
BROY-INV(0.5)-AGS(0.90)	0.102	0.633	64.88	46.60	-0.32	9.49	0.66
BROY-INV(0.5)-AR(0.85)	0.102	0.644	84.76	35.86	-0.08	4.22	1.21
BROY-INV(0.5)-ARA(0.90)	0.102	0.644	79.07	43.26	-0.15	11.40	0.40
BROY-INV(0.5)-G(0.70)	0.102	0.583	82.40	59.64	-0.20	9.43	0.67
BROY-INV(0.75)-AGS(0.90)	0.102	0.630	67.68	46.04	-0.01	9.93	0.61
BROY-INV(0.75)-AR(0.80)	0.102	0.639	82.28	35.38	0.18	4.30	1.33
BROY-INV(0.75)-ARA(0.70)	0.102	0.649	73.86	36.38	0.37	5.89	0.85
BROY-INV(0.75)-G(0.70)	0.102	0.593	80.65	60.63	-0.22	9.45	0.68

Tabelle D.1: Ergebnisse der Optimierungsverfahren

Nr.	Molekül	Dim	$S^0_{\mathcal{R}_i}$	$S^*_{\mathcal{R}_i}$	Schritte	Stopp	Sec	Q	N
11	1cbx	15	0.141	0.899	49.5	0.45	1.16	3.36	1.6
7	2ctc	14	0.108	0.846	43.8	0.60	0.55	3.17	1.3
26	3cpa	18	0.093	0.723	76.2	0.50	2.92	3.10	2.1
55	6cpa	26	0.158	0.646	90.0	0.15	17.09	2.68	1.2
60	7cpa	29	0.138	0.648	94.8	0.20	33.58	2.77	1.2
17	5cna	17	0.101	0.579	69.6	0.40	1.31	3.01	1.1
15	t0013	16	0.042	0.600	60.4	0.50	0.81	3.25	1.0
49	1dhf	22	0.080	0.388	87.1	0.35	11.73	2.67	1.0
50	4dfr	23	0.080	0.546	87.4	0.25	13.00	2.85	1.3
56	1ela	26	0.070	0.506	97.4	0.10	20.76	2.68	1.2
64	1elb	31	0.099	0.617	100.0	0.00	31.75	2.47	0.8
61	1elc	29	0.047	0.502	100.0	0.00	30.70	2.46	0.8
51	1eld	23	0.049	0.463	92.0	0.15	12.86	2.58	0.5
54	1ele	24	0.048	0.411	91.7	0.15	11.48	2.64	0.9
18	t0035	17	0.086	0.650	55.8	0.50	3.13	3.22	0.9
42	t0036	21	0.085	0.591	76.8	0.30	7.23	3.01	1.0
76	2er7	58	0.066	0.530	100.0	0.00	328.58	2.45	1.2
68	4er1	39	0.045	0.480	94.4	0.10	134.33	2.54	1.0
74	4er2	50	0.038	0.552	100.0	0.00	83.23	2.48	1.4
66	5er1	38	0.105	0.700	100.0	0.00	52.16	2.52	1.2
75	5er2	51	0.064	0.521	100.0	0.00	229.90	2.41	0.9
19	4fbp	17	0.089	0.502	59.9	0.55	2.67	3.18	1.4
32	t0039	19	0.082	0.547	67.5	0.55	3.52	3.09	0.9
20	1gpy	17	0.118	0.640	57.3	0.60	1.45	2.88	0.5
21	3gpb	17	0.138	0.671	57.8	0.55	1.47	2.92	1.0
14	4gpb	16	0.136	0.690	51.4	0.55	1.10	3.08	1.2
22	5gpb	17	0.134	0.657	56.2	0.60	1.50	3.28	1.5
70	1hiv	41	0.072	0.528	96.5	0.10	123.33	2.60	0.8
67	1hos	38	0.101	0.569	98.1	0.05	97.42	2.52	0.8
71	1ivp	41	0.075	0.432	98.4	0.05	128.19	2.49	0.7
65	1ivq	35	0.090	0.513	94.9	0.10	65.09	2.54	0.4
72	2mip	46	0.121	0.777	100.0	0.00	134.79	2.50	1.1
73	4hvp	46	0.138	0.751	100.0	0.00	166.35	2.40	0.4
58	4phv	27	0.063	0.678	89.7	0.25	41.51	2.77	1.2
39	5hvp	20	0.069	0.662	99.1	0.05	100.32	2.45	0.5
77	8hvp	60	0.061	0.649	100.0	0.00	234.72	2.51	0.8
69	9hvp	39	0.099	0.616	93.6	0.10	97.36	2.49	0.7
8	1dbb	14	0.061	0.673	64.9	0.50	4.08	3.34	3.0
5	1dbj	13	0.054	0.401	45.0	0.80	2.18	2.96	0.9
2	1dbk	12	0.083	0.578	41.1	0.75	1.42	2.98	1.5
33	1dbm	19	0.087	0.548	81.7	0.40	11.75	3.16	2.1
34	2dbl	19	0.018	0.594	87.0	0.30	13.36	2.97	1.6
44	2r04	21	0.069	0.692	90.6	0.20	9.97	2.73	1.4
35	2r06	19	0.144	0.678	88.2	0.15	6.75	2.69	1.3
36	2r07	19	0.084	0.811	82.2	0.35	6.54	3.16	2.2
45	2rm2	22	0.054	0.774	77.8	0.35	10.79	2.96	1.7
46	2rr1	22	0.056	0.729	83.9	0.25	11.47	2.92	2.4
47	2rs1	22	0.061	0.758	85.8	0.20	11.66	2.90	2.3
52	2rs3	23	0.077	0.587	95.3	0.10	14.83	2.56	1.4
40	2rs5	20	0.181	0.787	81.7	0.20	7.67	2.83	1.7
27	1srf	18	0.085	0.741	76.1	0.40	5.12	3.02	1.4
13	1srg	15	0.093	0.733	53.6	0.60	1.86	3.22	1.9
28	1srh	18	0.086	0.815	67.1	0.40	3.47	3.31	1.9
16	1sri	16	0.087	0.785	67.3	0.55	3.17	3.31	2.3
9	1srj	14	0.092	0.693	45.6	0.70	1.94	3.44	2.6
59	1tlp	28	0.138	0.561	98.7	0.15	31.85	2.59	0.7
57	1tmn	26	0.061	0.749	79.8	0.30	21.33	2.87	0.9
29	2tmn	18	0.122	0.973	63.3	0.50	2.03	3.46	2.3
41	3tmn	20	0.060	0.644	81.0	0.25	6.33	2.89	1.3
23	4tln	17	0.089	0.843	73.3	0.55	1.29	3.16	2.1
62	4tmn	29	0.105	0.711	96.8	0.10	31.28	2.54	1.3
53	5tln	23	0.089	0.506	90.0	0.25	9.21	2.64	0.9
63	5tmn	30	0.106	0.637	93.6	0.15	27.12	2.76	1.5
30	cbz	18	0.106	0.737	59.0	0.55	3.68	3.00	1.5
31	ppp	18	0.113	0.718	69.9	0.40	4.46	3.02	1.8
24	rthior	17	0.146	0.666	74.5	0.50	2.36	2.98	1.3
25	thior	17	0.142	0.566	61.0	0.65	1.91	2.80	1.1
48	1dwc	22	0.085	0.698	81.6	0.30	18.25	3.14	1.5
43	1dwd	21	0.124	0.593	69.1	0.40	13.56	2.93	1.4
37	1ett	19	0.084	0.523	55.3	0.55	7.17	3.07	0.9
38	1pph	19	0.133	0.603	60.4	0.50	7.91	3.43	1.3
3	1tnh	12	0.178	0.818	42.7	0.70	0.26	3.00	1.0
12	1tni	15	0.177	0.628	64.5	0.55	1.22	3.00	1.5
6	1tnj	13	0.184	0.815	50.9	0.60	0.42	3.19	1.3
10	1tnk	14	0.188	0.904	65.4	0.45	0.89	2.99	1.5
4	1tnl	12	0.213	0.810	59.4	0.65	0.48	3.23	2.2
1	3ptb	11	0.412	0.998	21.7	0.95	0.13	2.46	0.6

Tabelle D.2: Durchschnittliche Ergebnisse des inversen BFGS-Verfahrens mit der Aufdatierungsformel (4.10) (BFGS-INV1) sowie mit der Armijo Regel mit Aufweitung (ARA) bzgl. dem Parameter $\beta = 0.35$.

Nr.	Molekül	Dim.	$S^0_{\mathcal{R}_i}$	$S^*_{\mathcal{R}_i}$	Sec	Stopp	$S^*_{\mathcal{R}_i}$(H)	Sec (H)	Stopp (H)
11	1cbx	15	0.141	0.899	1.11	0.45	0.814	1.44	0.65
7	2ctc	14	0.108	0.846	0.57	0.60	0.761	0.82	0.60
26	3cpa	18	0.092	0.723	2.70	0.50	0.752	4.18	0.55
55	6cpa	26	0.158	0.646	16.51	0.15	0.774	22.92	0.40
60	7cpa	29	0.138	0.648	31.88	0.20	0.639	44.41	0.35
17	5cna	17	0.100	0.579	1.28	0.40	0.591	1.68	0.50
15	t0013	16	0.042	0.600	0.76	0.50	0.615	1.21	0.40
49	1dhf	22	0.080	0.388	11.04	0.35	0.558	16.75	0.45
50	4dfr	23	0.080	0.546	12.66	0.25	0.510	17.59	0.40
56	1ela	26	0.070	0.506	20.16	0.10	0.494	30.69	0.10
64	1elb	31	0.099	0.617	30.06	0.00	0.524	45.39	0.25
61	1elc	29	0.047	0.502	30.25	0.00	0.583	43.59	0.15
51	1eld	23	0.049	0.463	12.22	0.15	0.472	16.12	0.45
54	1ele	24	0.048	0.411	11.10	0.15	0.498	16.63	0.30
18	t0035	17	0.086	0.650	2.98	0.50	0.622	4.07	0.50
42	t0036	21	0.085	0.591	6.95	0.30	0.552	10.53	0.60
76	2er7	58	0.066	0.530	316.82	0.00	0.547	500.64	0.15
68	4er1	39	0.045	0.480	126.53	0.10	0.480	191.88	0.10
74	4er2	50	0.038	0.552	80.90	0.00	0.553	127.58	0.00
66	5er1	38	0.105	0.700	50.48	0.00	0.623	71.84	0.10
75	5er2	51	0.064	0.521	224.39	0.00	0.531	338.23	0.10
19	4fbp	17	0.089	0.502	2.53	0.55	0.500	4.15	0.45
32	t0039	19	0.081	0.547	3.41	0.55	0.638	4.93	0.55
20	1gpy	17	0.118	0.640	1.37	0.60	0.608	2.03	0.55
21	3gpb	17	0.138	0.671	1.40	0.55	0.672	2.21	0.55
14	4gpb	16	0.136	0.690	1.06	0.55	0.656	1.81	0.55
22	5gpb	17	0.134	0.657	1.42	0.60	0.646	2.18	0.55
70	1hiv	41	0.072	0.528	120.63	0.10	0.549	171.66	0.15
67	1hos	38	0.101	0.569	97.17	0.05	0.603	143.17	0.25
71	1ivp	41	0.075	0.432	119.66	0.05	0.459	170.64	0.10
65	1ivq	35	0.090	0.512	61.93	0.10	0.485	88.65	0.20
72	2mip	46	0.121	0.777	126.75	0.00	0.760	186.83	0.10
73	4hvp	46	0.138	0.751	158.52	0.00	0.831	240.33	0.10
58	4phv	27	0.063	0.678	39.20	0.25	0.744	46.64	0.45
39	5hvp	20	0.069	0.661	96.87	0.05	0.720	145.33	0.00
77	8hvp	60	0.061	0.649	225.88	0.00	0.521	366.62	0.15
69	9hvp	39	0.099	0.616	94.16	0.10	0.610	145.17	0.20
8	1dbb	14	0.061	0.673	3.87	0.50	0.622	3.79	0.60
5	1dbj	13	0.054	0.400	2.06	0.80	0.516	2.79	0.75
2	1dbk	12	0.083	0.578	1.39	0.75	0.560	1.69	0.70
33	1dbm	19	0.086	0.548	11.15	0.40	0.536	15.97	0.45
34	2dbl	19	0.017	0.594	12.65	0.30	0.640	17.43	0.45
44	2r04	21	0.069	0.692	9.53	0.20	0.723	14.27	0.25
35	2r06	19	0.144	0.677	6.64	0.15	0.677	9.35	0.40
36	2r07	19	0.084	0.811	6.43	0.35	0.703	9.34	0.30
45	2rm2	22	0.054	0.774	10.38	0.35	0.880	14.69	0.35
46	2rr1	22	0.056	0.729	11.00	0.25	0.633	18.08	0.15
47	2rs1	22	0.061	0.758	11.24	0.20	0.853	15.43	0.40
52	2rs3	23	0.077	0.587	14.39	0.10	0.685	21.94	0.10
40	2rs5	20	0.181	0.787	7.46	0.20	0.743	10.88	0.45
27	1srf	18	0.085	0.741	5.02	0.40	0.690	5.44	0.55
13	1srg	15	0.093	0.733	1.79	0.60	0.879	2.75	0.50
28	1srh	18	0.086	0.815	3.11	0.40	0.758	4.19	0.55
16	1sri	16	0.087	0.784	2.87	0.55	0.813	4.52	0.40
9	1srj	14	0.092	0.693	1.73	0.70	0.856	3.23	0.55
59	1tlp	28	0.138	0.561	27.96	0.15	0.545	40.05	0.25
57	1tmn	26	0.061	0.749	18.84	0.30	0.818	27.77	0.35
29	2tmn	18	0.122	0.972	1.81	0.50	0.869	2.71	0.50
41	3tmn	20	0.060	0.644	5.72	0.25	0.600	8.59	0.45
23	4tln	17	0.089	0.843	1.18	0.55	0.843	1.50	0.55
62	4tmn	29	0.105	0.711	29.60	0.10	0.760	37.80	0.30
53	5tln	23	0.089	0.506	8.55	0.25	0.646	12.58	0.35
63	5tmn	30	0.106	0.637	25.88	0.15	0.774	37.16	0.30
30	cbz	18	0.106	0.737	3.42	0.55	0.571	6.74	0.40
31	ppp	18	0.113	0.718	4.20	0.40	0.619	6.21	0.60
24	rthior	17	0.146	0.666	2.20	0.50	0.664	3.12	0.50
25	thior	17	0.142	0.566	1.80	0.65	0.584	2.73	0.55
48	1dwc	22	0.085	0.698	17.03	0.30	0.729	25.74	0.35
43	1dwd	21	0.124	0.593	13.20	0.40	0.585	18.88	0.50
37	1ett	19	0.084	0.522	6.75	0.55	0.545	11.14	0.50
38	1pph	19	0.133	0.603	7.54	0.50	0.631	13.86	0.35
3	1tnh	12	0.178	0.818	0.24	0.70	0.825	0.33	0.65
12	1tni	15	0.177	0.628	1.18	0.55	0.806	1.80	0.55
6	1tnj	13	0.184	0.815	0.43	0.60	0.813	0.48	0.60
10	1tnk	14	0.188	0.904	0.84	0.45	0.901	0.97	0.60
4	1tnl	12	0.213	0.810	0.45	0.65	0.812	0.52	0.55
1	3ptb	11	0.412	0.998	0.12	0.95	0.996	0.33	0.65

Tabelle D.3: Durchschnittliche Ergebnisse des inversen BFGS-Verfahrens mit der Armijo Regel mit Aufweitung (BFGS-INV1/ARA (0.35)) ohne und mit Hierarchie-Verfahren 4.7.

Nr.	Molekül	Dim.	$S^0_{\mathcal{A}_1^i,\dots,\mathcal{A}_5^i}$	$S^*_{\mathcal{A}_1^i,\dots,\mathcal{A}_5^i}$	Stopp	Sec	$S^*_{\mathcal{A}_1^i,\dots,\mathcal{A}_5^i}$ (GS)	Stopp (GS)	Sec (GS)
11	1cbx	75	0.122	0.442	0.5	7.24	0.417	0.5	13.55
7	2ctc	70	0.088	0.324	0.5	3.45	0.375	0.5	13.07
26	3cpa	90	0.046	0.251	-0.5	14.06	0.222	0.5	25.64
55	6cpa	130	0.100	0.427	0.5	83.67	0.468	0.5	92.13
60	7cpa	145	0.107	0.422	0.5	178.21	0.389	0.5	217.74
17	5cna	85	0.061	0.345	-0.5	6.73	0.408	0.5	15.85
15	t0013	80	0.125	0.457	0.5	4.91	0.458	0.5	9.30
49	1dhf	110	0.046	0.160	-0.5	93.58	0.210	0.5	153.08
50	4dfr	115	0.042	0.172	-0.5	66.86	0.253	0.5	114.87
56	1ela	130	-0.007	0.169	0.5	111.06	0.141	0.5	204.03
64	1elb	155	0.004	0.244	0.5	168.71	0.264	0.5	188.15
61	1elc	145	0.012	0.148	0.5	105.21	0.157	0.5	149.22
51	1eld	115	-0.004	0.128	-0.5	123.77	0.093	0.5	112.35
54	1ele	120	-0.011	0.087	0.5	47.15	0.152	0.5	81.05
18	t0035	85	0.015	0.279	0.5	30.94	0.342	0.5	211.46
42	t0036	105	0.050	0.135	-0.5	41.06	0.253	0.5	116.13
76	2er7	290	0.018	0.113	0.5	1634.73	0.150	0.5	1917.63
68	4er1	195	-0.007	0.103	0.0	1512.17	0.182	0.5	960.74
74	4er2	250	0.014	0.106	0.5	323.80	0.173	0.5	422.22
66	5er1	190	0.057	0.266	0.5	399.69	0.346	0.5	1085.73
75	5er2	255	0.017	0.123	0.5	1176.17	0.134	0.5	2002.93
19	4fbp	85	0.071	0.301	0.5	20.90	0.238	0.5	80.36
32	t0039	95	0.063	0.143	0.5	28.36	0.178	0.5	61.49
20	1gpy	85	0.043	0.135	-0.5	9.10	0.227	0.5	15.73
21	3gpb	85	0.056	0.262	0.5	12.93	0.341	0.5	27.43
14	4gpb	80	0.032	0.316	-0.5	7.76	0.323	0.5	22.41
22	5gpb	85	0.027	0.289	0.5	9.84	0.282	0.5	16.02
70	1hiv	205	0.059	0.202	-0.5	663.11	0.209	0.5	1151.47
67	1hos	190	0.106	0.278	0.5	681.08	0.282	0.5	367.60
71	1ivp	205	0.070	0.190	0.5	607.80	0.211	0.5	1036.68
65	1ivq	175	0.067	0.248	0.5	385.57	0.244	0.5	400.72
72	2mip	230	0.079	0.182	0.5	355.39	0.355	0.5	738.81
73	4hvp	230	0.112	0.287	-0.5	499.26	0.327	0.5	517.41
58	4phv	135	0.084	0.263	0.5	189.45	0.260	0.5	307.79
39	5hvp	100	0.044	0.126	-0.5	580.20	0.209	0.5	418.57
77	8hvp	300	0.023	0.105	0.5	1296.30	0.119	0.5	1018.68
69	9hvp	195	0.088	0.245	0.5	447.65	0.275	0.5	597.44
8	1dbb	70	0.012	0.207	0.5	31.62	0.189	0.5	52.68
5	1dbj	65	0.026	0.106	0.5	12.69	0.115	0.5	41.72
2	1dbk	60	0.038	0.226	0.5	13.82	0.178	0.5	33.30
33	1dbm	95	0.007	0.209	0.5	70.53	0.210	0.5	150.79
34	2dbl	95	0.023	0.259	0.5	80.28	0.144	0.5	537.34
44	2r04	105	0.000	0.116	-0.5	71.00	0.154	0.5	73.15
35	2r06	95	0.044	0.545	0.5	60.62	0.261	0.5	47.37
36	2r07	95	0.020	0.198	-0.5	33.21	0.179	0.5	82.91
45	2rm2	110	-0.009	0.219	0.5	56.77	0.331	0.5	181.75
46	2rr1	110	-0.016	0.223	0.5	67.70	0.189	0.5	118.77
47	2rs1	110	0.007	0.233	0.5	62.70	0.299	0.5	91.99
52	2rs3	115	-0.003	0.252	-0.5	99.43	0.258	0.5	73.70
40	2rs5	100	0.067	0.168	-0.5	22.55	0.216	0.5	105.84
27	1srf	90	0.016	0.164	-0.5	24.67	0.380	0.5	37.91
13	1srg	75	0.075	0.224	-0.5	22.45	0.345	0.5	38.09
28	1srh	90	0.044	0.161	0.5	21.31	0.168	0.5	40.76
16	1sri	80	0.050	0.216	0.5	15.97	0.281	0.5	46.02
9	1srj	60	0.068	0.212	0.5	12.48	0.411	0.5	48.33
59	1tlp	140	0.088	0.248	0.5	130.90	0.264	0.5	242.74
57	1tmn	130	-0.005	0.297	0.5	78.74	0.274	0.5	133.65
29	2tmn	90	0.102	0.437	0.5	11.47	0.755	0.5	16.30
41	3tmn	100	0.110	0.350	-0.5	28.42	0.352	0.5	72.77
23	4tln	85	0.026	0.625	0.5	6.60	0.680	0.5	8.54
62	4tmn	145	0.066	0.372	0.5	206.46	0.384	0.5	201.40
53	5tln	115	0.044	0.167	0.5	33.12	0.201	0.5	80.08
63	5tmn	150	0.070	0.296	0.5	139.74	0.299	0.5	134.48
30	cbz	90	0.036	0.276	0.5	19.13	0.298	0.5	45.65
31	ppp	90	0.024	0.128	0.5	19.98	0.194	0.5	67.61
24	rthior	85	0.145	0.473	0.5	17.69	0.569	0.5	24.63
25	thior	85	0.100	0.278	0.5	14.12	0.332	0.5	22.48
48	1dwc	110	0.080	0.237	0.5	93.76	0.251	0.5	178.40
43	1dwd	105	0.074	0.348	-0.5	137.19	0.334	0.5	194.20
37	1ett	95	0.033	0.146	-0.5	66.69	0.158	0.5	73.32
38	1pph	95	0.025	0.272	-0.5	48.64	0.268	0.5	122.40
3	1tnh	60	0.194	0.583	-0.5	1.87	0.599	0.5	5.32
12	1tni	85	0.099	0.570	0.5	14.41	0.937	0.5	18.48
6	1tnj	65	0.268	0.961	-0.5	4.13	0.963	0.5	9.04
10	1tnk	70	0.154	0.582	0.5	4.10	0.598	0.5	22.28
4	1tnl	60	0.258	0.939	0.5	4.18	0.944	0.5	11.75
1	3ptb	55	0.432	0.959	0.5	2.13	0.937	0.5	6.66

Tabelle D.4: Ergebnisse des Minimierungsalgorithmus 4.2 und des Gauß-Seidel-artigem Verfahren 4.6 mit MAXSTEP=17 im Vergleich: Gegenseitige Überlagerung von Konformationen von Typ A.

Nr.	Molekül	Dim.	$S^0_{\mathcal{A}_1^i,\dots,\mathcal{A}_5^i}$	$S^*_{\mathcal{A}_1^i,\dots,\mathcal{A}_5^i}$	Stopp	Sec	$S^*_{\mathcal{A}_1^i,\dots,\mathcal{A}_5^i}$ (GS)	Stopp (GS)	Sec (GS)
11	1cbx	75	0.117	0.744	0.5	8.69	0.663	0.5	15.02
7	2ctc	70	0.065	0.723	0.5	4.35	0.651	0.5	8.24
26	3cpa	90	0.058	0.526	-0.5	17.29	0.624	0.5	21.11
55	6cpa	130	0.094	0.718	-0.5	97.63	0.501	0.5	126.16
60	7cpa	145	0.074	0.309	-0.5	224.55	0.335	0.5	238.38
17	5cna	85	0.059	0.595	-0.5	10.71	0.450	0.5	23.01
15	t0013	80	0.045	0.473	0.5	3.28	0.483	0.5	13.80
49	1dhf	110	0.004	0.097	0.5	69.32	0.097	0.5	101.70
50	4dfr	115	0.002	0.081	0.5	67.87	0.124	0.5	95.83
56	1ela	130	0.008	0.612	-0.5	165.19	0.390	0.5	154.62
64	1elb	155	0.008	0.130	-0.5	228.01	0.341	0.5	160.27
61	1elc	145	0.005	0.090	0.5	71.14	0.133	0.5	143.77
51	1eld	115	0.013	0.125	0.5	112.03	0.190	0.5	130.15
54	1ele	120	0.007	0.117	0.5	53.56	0.116	0.5	75.20
18	t0035	85	0.040	0.485	0.5	23.33	0.293	0.5	46.46
42	t0036	105	0.018	0.219	0.5	42.54	0.466	0.5	64.91
76	2er7	290	0.001	0.213	-0.5	2854.21	0.172	0.5	5.23
68	4er1	195	0.008	0.448	0.5	1164.54	0.469	0.5	1205.99
74	4er2	250	0.030	0.181	0.5	404.90	0.216	0.5	380.65
66	5er1	190	0.056	0.555	0.5	292.31	0.667	0.5	298.49
75	5er2	255	0.020	0.077	0.5	886.25	0.106	0.5	1083.77
19	4fbp	85	0.001	0.415	0.5	33.63	0.239	0.5	67.39
32	t0039	95	0.021	0.319	0.5	27.42	0.628	0.5	34.20
20	1gpy	85	0.010	0.360	0.5	12.38	0.231	0.5	18.60
21	3gpb	85	0.036	0.266	0.5	10.23	0.428	0.5	28.82
14	4gpb	80	0.026	0.337	0.5	7.43	0.493	0.5	45.82
22	5gpb	85	0.040	0.708	0.5	15.80	0.401	0.5	22.39
70	1hiv	205	0.003	0.369	-0.5	886.16	0.368	0.5	656.99
67	1hos	190	0.031	0.304	0.5	368.66	0.336	0.5	490.78
71	1ivp	205	0.007	0.104	0.5	1264.89	0.208	0.5	655.06
65	1ivq	175	-0.006	0.133	0.5	288.30	0.233	0.5	379.65
72	2mip	230	0.028	0.312	0.5	819.91	0.650	0.5	742.64
73	4hvp	230	0.077	0.516	0.5	1044.57	0.693	0.5	608.77
58	4phv	135	0.016	0.452	0.5	249.51	0.305	0.5	272.35
39	5hvp	100	0.004	0.195	-0.5	429.88	0.181	0.5	607.08
77	8hvp	300	0.008	0.316	0.5	1919.70	0.062	0.5	809.91
69	9hvp	195	0.030	0.398	0.5	471.27	0.646	0.5	620.52
8	1dbb	70	0.072	0.676	0.5	29.51	0.668	0.5	49.00
5	1dbj	65	0.038	0.215	-0.5	26.81	0.156	0.5	34.64
2	1dbk	60	0.049	0.580	-0.5	26.74	0.182	0.5	16.44
33	1dbm	95	0.011	0.244	0.5	59.78	0.332	0.5	97.89
34	2dbl	95	0.020	0.187	0.5	93.22	0.238	0.5	132.51
44	2r04	105	0.006	0.340	0.5	59.11	0.201	0.5	101.42
35	2r06	95	0.106	0.419	-0.5	32.61	0.696	0.5	108.55
36	2r07	95	0.049	0.507	0.5	39.04	0.589	0.5	72.04
45	2rm2	110	0.020	0.572	-0.5	86.91	0.419	0.5	125.23
46	2rr1	110	0.010	0.395	0.5	47.85	0.599	0.5	119.37
47	2rs1	110	0.040	0.591	-0.5	79.34	0.541	0.5	120.94
52	2rs3	115	0.016	0.521	0.5	86.15	0.352	0.5	81.92
40	2rs5	100	0.135	0.331	0.5	44.79	0.727	0.5	103.65
27	1srf	90	0.034	0.271	0.5	39.92	0.235	0.5	118.64
13	1srg	75	0.078	0.819	0.5	20.34	0.249	0.5	42.19
28	1srh	90	0.084	0.183	-0.5	15.10	0.270	0.5	34.44
16	1sri	80	0.079	0.581	0.5	15.84	0.410	0.5	50.06
9	1srj	60	0.070	0.176	0.5	18.79	0.182	0.5	67.13
59	1tlp	140	0.039	0.421	0.5	139.36	0.381	0.5	261.50
57	1tmn	130	0.076	0.476	0.5	138.47	0.321	0.5	134.20
29	2tmn	90	0.047	0.584	0.5	11.65	0.594	0.5	11.05
41	3tmn	100	0.024	0.503	0.5	30.97	0.681	0.5	86.09
23	4tln	85	0.075	0.743	-0.5	6.20	0.484	0.5	11.25
62	4tmn	145	0.061	0.691	0.5	183.92	0.312	0.5	195.41
53	5tln	115	0.014	0.116	-0.5	39.49	0.346	0.5	66.26
63	5tmn	150	0.092	0.297	0.5	197.27	0.326	0.5	169.69
30	cbz	90	0.021	0.478	0.5	27.74	0.465	0.5	64.83
31	ppp	90	0.074	0.583	0.5	24.31	0.604	0.5	86.96
24	rthior	85	0.094	0.326	0.5	12.64	0.333	0.5	19.12
25	thior	85	0.083	0.331	0.5	13.71	0.490	0.5	25.31
48	1dwc	110	-0.001	0.183	-0.5	90.99	0.232	0.5	187.90
43	1dwd	105	0.016	0.141	0.5	68.90	0.230	0.5	123.18
37	1ett	95	0.028	0.381	0.5	69.29	0.236	0.5	90.46
38	1pph	95	0.076	0.427	0.5	46.97	0.334	0.5	99.61
3	1tnh	60	0.074	0.598	0.5	1.77	0.631	0.5	6.04
12	1tni	85	0.038	0.902	0.5	11.63	0.940	0.5	10.41
6	1tnj	65	0.036	0.931	0.5	3.40	0.968	0.5	5.71
10	1tnk	70	0.033	0.598	0.5	6.34	0.414	0.5	13.48
4	1tnl	60	0.033	0.909	-0.5	2.92	0.419	0.5	5.04
1	3ptb	55	0.326	0.918	0.5	2.15	0.952	0.5	4.80

Tabelle D.5: Ergebnisse des Minimierungsalgorithmus 4.2 und des Gauß-Seidel-artigem Verfahren 4.6 mit MAXSTEP=17 im Vergleich: Gegenseitige Überlagerung von Konformationen von Typ B.

Literaturverzeichnis

[1] M. Aigner. *Diskrete Mathematik.* Vieweg, Wissbaden, 1993.

[2] F.H. Allen, S. Bellard, M.D. Brice, B.A. Cartwright, A. Doubleday, H. Higgs, T. Hummelink-Peters, O. Kennard, W.D.S. Motherwell, J.R. Rodgers, and D.G. Watson. The Cambridge Crystallographic Data Centre: computer-based search, retrieval, analysis and display of information. *Acta Crystallographica*, B35:2331–2339, 1979.

[3] M.P. Allen and D.J. Tildesley. *Computer Simulation of Liquids.* Clarendon Press, Oxford, 1987.

[4] Peter W. Atkins. *Physikalische Chemie, 3. Auflage.* VCH, Weinheim, 1990.

[5] F.C. Bernstein, G.J.B. Williams, T.T. Koetzle, E.F.Jr. Meyer, M.D. Brice, J.R. Rodgers, O. Kennard, T. Shimanouchi, and M. Tasumi. The protein data bank: A computer based archival file for macromolecular structures. *Journal of Molecular Biology*, 112:535–542, 1977.

[6] H.-J. Böhm, G. Klebe, and H. Kubinyi. *Wirkstoffdesign.* Spektrum Akademischer Verlag, 1996.

[7] H.-J. Böhm, G. Klebe, and H. Kubinyi. *Wirkstoffdesign*, chapter 16 Konformationsanalyse. Spektrum Akademischer Verlag, 1996.

[8] R. Carbó, L. Leyda, and M. Arnau. How similar is a molecule to another? An electron density measure of the similarity between two compounds. *International Journal of Quantum Chemistry*, 17:1185–1189, 1980.

[9] T.H. Cormen, C.E. Leiserson, and R.L. Rivest. *Introduction to algorithms.* The MIT Press, Cambridge, Massachusetts, 1992.

[10] L. Cromme. Verfahren der nichtlinearen Optimierung für Mathematiker, Informatiker und Physiker. Aufgaben und Lösungen zur Vorlesung, Technische Universität Cottbus, 1996.

[11] W.C. Davidon. Optimally conditioned optimization algorithms wihtout line searches. *Mathematical Programming*, 9:1–30, 1975.

[12] J.E. Dennis and R. B. Schnabel. *Numerical Methods for Unconstrained Optimization and Nonlinear Equations.* SIAM, Philadelphia, 1996.

[13] R. Diestel. *Graphentheorie*. Springer, Berlin, 1996.

[14] J.S. Dixon. Evaluation of the CASP2 docking section. *Proteins: Structure, Function, and Genetic*, Suppl. 1:198–204, 1997.

[15] E. Fischer. Einfluß der Konfiguration auf die Wirkung der Enzyme. *Berichte der Deutschen Chemiker Gesellschaft*, 27:2985–2993, 1894.

[16] Otto Forster. *Analysis 2, 5., durchgesehene Auflage*. Vieweg, Braunschweig, 1984.

[17] Otto Forster. *Analysis 1, 4., durchgesehene Auflage*. Vieweg, Braunschweig, 1989.

[18] W. Forth, D. Henschler, and W. Rummel ans K. Starke, editors. *Pharmakologie und Toxikologie. 6., völlig neu bearbeitete Auflage*. BI-Wissenschaftsverlage, Mannheim, 1992.

[19] C. Geiger and C. Kanzow. *Numerische Verfahren zur Lösung unrestringierter Optimierungsaufgaben*. Springer, Berlin, 1999.

[20] J.C. Gilbert and J. Nocedal. Global convergence properties of conjugate gradient methods of optimization. *SIAM Journal on Optimization*, 2:21–42, 1992.

[21] Roland Glaser. *Biophysik, 4. Auflage*. Gustav Fischer, Jena, Stuttgart, 1996.

[22] H. Goldstein. *Klassische Mechanik, 8. Auflage*. AULA-Verlag, Wissbaden, 1970.

[23] A.C. Good. 3D molecular similarity indices and their application in QASR studies. In P.M. Dean, editor, *Molecular Similarity in Drug Design*, pages 24–56. Blackie Academic & Professional, Glasgow, 1995.

[24] A.C. Good, E.E. Hodgkin, and W.G. Richards. Utilization of gaussian functions for the rapid evaluation of molecular similarity. *Journal of Chemical Information and Computer Science*, 32:188–191, 1992.

[25] J.A. Grant, M.A. Gallardo, and B.T. Pickup. A Fast Method of Molecular Shape Comparison: A Simple Application of a Gaussian Description of Molecular Shape. *Journal of Computational Chemistry*, 17:1653–1666, 1996.

[26] M. Griebel, A. Caglar, and S. Knapek. Partikelmethoden und gitterlose Diskretisierungen. Numerisches Praktikum, Universität Berlin, 1999.

[27] A.O. Griewank, B.R. Markey, and D.J. Evans. Singularity-free static lattice energy minimization. *Journal of Chemical Physics*, 71(8):3449–3454, 1979.

[28] C. Großmann and J. Terno. *Numerik der Optimierung. 2. durchgesehene Auflage*. B. G. Teubner, Stuttgart, 1997.

[29] G. Hämmerlin and K.-H. Hoffmann. *Numerische Mathematik, 3. unveränderte Auflage*. Springer, Berlin, 1992.

[30] M. Hauswald. Structural Alignment of Isofunctional Molecules. Dissertation, Freie Universität Berlin, 1998.

[31] H. Heuser. *Lehrbuch der Analysis. Teil 2.* B.G. Teubner, Stuttgart, 1991.

[32] C. Hiller. Optimierungsmethoden zum strukturellen Alignment von Ligandmolekülen. Diplomarbeit, Universität Bonn, 1997.

[33] E.E. Hodgkin and W.G. Richards. Molecular similarity based on electrostatic potential and electric field. *International Journal of Quantum Chemistry, Quantum Biological Symposium*, 14:105–110, 1987.

[34] Hunnius Pharmazeutisches Wörterbuch. 7., völlig neu bearbeitete und start erweiterte auflage. Berlin, 1993.

[35] G. Klebe, T. Mietzner, and F. Weber. Different approaches toward an automatic structural alignment of drug molecules: Application to sterol mimics, thrombin and thermolysin inhibitors. *Journal of Computer-Aided Molecular Design*, 8:751–778, 1994.

[36] J. Koolman and K.-H. Röhm. *Taschenatlas der Biochemie.* Georg Thieme Verlag, Stuttgart, 1994.

[37] P. Kosmol. *Methoden zur numerischen Behandlung nichtlinearer Gleichungen und Optimierungsaufgaben. 2. überarbeitete Auflage.* Teubner, Stuttgart, 1993.

[38] H. Kubinyi, editor. *3D QSAR in Drug Design. Theory, Methods and Applications.* ESCOM Science Publishers, Leiden, The Netherlands, 1993.

[39] E.E. Lattmann. Optimal Sampling of the Rotation Function. *Acta Crystallographica*, B28:1065–1068, 1972.

[40] R. Leis. Infinitesimalrechnung III. Vorlesungsskript, Universität Bonn, 1994.

[41] C. Lemmen. *Computational Methods for the Structural Alignments of Molecules.* GMD, St. Augustin, 1999. Dissertation, Universität Bonn.

[42] C. Lemmen, T. Lengauer, and G. Klebe. FLEXS: a method for fast flexible ligand superposition. *Journal of Medicinal Chemistry*, 41:4502–4520, 1998.

[43] G. Löffler and P.E. Petrides. *Biochemie und Pathobiochemie. 6. korrigierte Auflage.* Springer, Berlin, 1998.

[44] H. Lüllmann, K. Mohr, and A. Ziegler. *Taschenatlas der Pharmakologie. 2. überarbeitete und erweiterte Auflage.* Georg Thieme Verlag, Stuttgart, 1994.

[45] J.C. Meza, T.D. Plantenga, and R.S. Judson. Novel Applications of Optimization to Molecule Design. In L.T. Biegler, Th.F. Coleman, A.R. Conn, and F.N. Santosa, editors, *Large-Scale Optimization with Applications. Part III: Molecular Structure and Optimization*, pages 69–97. Springer, New York, 1997.

[46] E. Polak. *Optimization. Algorithms and Consistent Approximations.* Springer, New York, 1997.

[47] Pschyrembel Klinisches Wörterbuch. 258. neu bearbeitete Auflage. Berlin, 1998.

[48] M. Rarey. *Rechnergestützte Vorhersage von Rezeptor-Ligand-Wechselwirkungen.* R.Oldenbourg Verlag, München/Wien, 1996. Dissertation, Universität Bonn, GMD-Berichte Nr. 268.

[49] RCSB PDB. http://www.rcsb.org/pdb/.

[50] B. Daya Reddy. *Introductory Functional Analysis.* Springer, New York, 1998.

[51] B. Schmidt. *Lineare Algebra. Skriptum zur Vorlesung.* Mathematisches Institut der Universität Bonn, 1993.

[52] H. R. Schwarz. *Numerische Mathematik, 3. überarbeitete und erweiterte Auflage.* B. G. Teubner, Stuttgart, 1993.

[53] P. Spellucci. *Numerische Verfahren der nichtlinearen Optimierung.* Birkhäser, Basel, 1993.

[54] J. Stoer. *Numerische Mathematik 1, 7. neubearbeitete und erweiterte Auflage.* Springer, Berlin, 1994.

[55] J. Stoer and F. Jarre. Optimierung I. Vorlesungsskript, Universität Würzburg, 1999.

[56] L. Stryer. *Biochemistry. 4th Edition.* Freeman and Company, New York, USA, 1995.