
Bestimmung von Tal-Rücken-Umschlagpunkten auf Potentialenergieflächen mittels eines Variationsansatzes für Newtontrajektorien

Diplomarbeit

Benjamin Schmidt
Mathematisches Institut

Universität Leipzig

1. Gutachter: Dr. Wolfgang Quapp
2. Gutachter: Dr. Anita Kripfganz

15. Oktober 2009

Inhaltsverzeichnis

Einleitung	5
1. Chemische Grundlagen	7
1.1. Theorie der Potentialenergieflächen	7
1.1.1. Schrödingergleichung und Born-Oppenheimer-Näherung	7
1.1.2. Potentialenergieflächen	9
1.2. Reaktionswege auf Potentialenergieflächen	11
1.2.1. Theorie des Übergangszustands	12
1.2.2. Minimum-Energie-Weg und Reaktionskoordinate	13
1.3. Verzweigungen auf Potentialenergieflächen	15
2. Wegeverfolgung auf Potentialenergieflächen	20
2.1. Die Methode des steilsten Abstiegs	20
2.1.1. Die intrinsische Reaktionskoordinate IRC	21
2.1.2. Die String-Methode und ihre Anwendung auf die IRC	23
2.1.3. Krümmung von Kurven des steilsten Abstiegs	27
2.2. Das Gradient-Extremal-Verfahren	27
2.2.1. Pfade im Talgrund	28
2.2.2. Numerische Verfolgung von GE-Kurven	31
2.2.3. Bifurkationen und Umkehrpunkte auf Gradient Extremals	36
3. Die Methode des reduzierten Gradienten	39
3.1. Die Distinguished-Coordinate-Methode	39
3.2. Das desingularisierte kontinuierliche Newtonverfahren	40
3.3. Das RGF-Verfahren und Newtontrajektorien	42
3.3.1. VRI-Punkte als außerordentliche Singularitäten	46
3.3.2. Newtontrajektorien als Reaktionswege	47
3.4. Numerische Verfolgung von RGF-Kurven	48
3.5. Das TASC-Verfahren	51
4. Variationsmethode	55
4.1. Einführung in die Variationsrechnung	55
4.1.1. Hinreichende Bedingung für schwache Minima	56
4.1.2. Hinreichende Bedingung für starke Minima	58
4.2. Variationsansatz für Reaktionswege	60
4.2.1. Die intrinsische Reaktionskoordinate IRC	60

Inhaltsverzeichnis

4.2.2. Gradient Extremals	62
4.2.3. Das RGF-Verfahren	63
5. Bestimmung von VRI-Punkten	66
5.1. Der Algorithmus	66
5.2. Anwendung auf zweidimensionale Modellpotentiale	70
5.2.1. Die PES-Familie (5.5)	70
5.2.2. Das Müller-Brown-Potential	73
5.3. Ein dreidimensionales Beispiel	74
Zusammenfassung	81
A. Verwendete Potentialenergieflächen	83
B. Rechnungen	85
B.1. Eigenschaften der adjunkten Matrix	85
B.2. Sun-Ruedenberg-Algorithmus	86
B.3. Minimaler Abstand zwischen Punkt und Gerade	88
C. Quelltext	89
Abkürzungsverzeichnis	95
Abbildungsverzeichnis	96
Tabellenverzeichnis	97
Literatur	98

Einleitung

Das Konzept des Minimum-Energie-Pfades (MEP) dient als grundlegendes Modell, um den Ablauf chemischer Reaktionen zu verstehen. Es basiert auf der Theorie des Übergangszustandes, derzufolge Ausgangsstoffe und Reaktionsprodukte energetisch stabile Zustände bilden, die durch eine Potentialbarriere voneinander getrennt sind. Beim Übergang der Reaktanten in die Produkte gilt es, diese Barriere zu überwinden.

Aus mathematischer Sicht fassen wir die Anfangs- und Endzustände einer chemischen Reaktion als Minima auf einer Potentialenergiefläche (PES) auf. Den Minimum-Energie-Pfad verstehen wir dann als Kurve im Konfigurationsraum der PES, die jene beiden Minima über einen zur Potentialbarriere gehörigen Sattelpunkt miteinander verbindet. Zudem ist erstens zu fordern, daß entlang des MEP die potentielle Energie vom Reaktantenminimum zum Sattelpunkt streng monoton wächst und anschließend streng monoton zum Produktminimum fällt. Zweitens soll der MEP bis zum Sattelpunkt durch ein Tal verlaufen.

Aufgrund der Unterschiede möglicher Potentialenergieflächen läßt sich der Minimum-Energie-Pfad nicht mathematisch exakt fixieren. Pfade, die den obigen Bedingungen genügen, müssen als gleichwertig betrachtet werden; den *einen* MEP gibt es nicht. Diese Unschärfe führt wiederum zu verschiedenen Reaktionswegmodellen.

Das einfachste, aber nach wie vor bedeutendste dieser Modelle ist die auf Steilst Abstiegs-Methoden basierende intrinsische Reaktionskoordinate (IRC) nach Fukui [13]. Sie erfüllt zwar die Monotonieforderungen, unterscheidet aber nicht zwischen Tal- und Rückenstrukturen auf der PES, wodurch sie der zweiten Bedingung an den MEP nicht immer genügt.

Das zweite Reaktionswegmodell, das Gradient-Extremal-Verfahren (GE), behebt diesen Mißstand, da es Kurven liefert, die im Talgrund verlaufen [23, 47]. Entlang jener GE-Kurven ist der Gradient ein Eigenvektor der Hessematrix. Ihre Bestimmung erfordert allerdings einen relativ hohen Rechenaufwand. Als entscheidendes Problem für das GE-Verfahren erweist sich jedoch die Tatsache, daß es Potentialenergieflächen gibt, auf denen gar kein Talgrund zwischen Minimum und Sattelpunkt existiert. Vielmehr können Täler irgendwo auf der PES enden, ohne den Sattelpunkt je erreicht zu haben.

Einen Ausweg bietet das Verfahren des reduzierten Gradienten (RGF) als drittes Modell [44]. Entlang von RGF-Kurven zeigt der Gradient konstant in eine fixierte Suchrichtung. Ihre Bestimmung ist zwar aufwendiger als bei der IRC, aber leichter als bei den Gradient Extremals. Wegen der engen Verwandtschaft des RGF-Verfahrens zum kontinuierlichen

Einleitung

Newtonverfahren bezeichnen wir die RGF-Kurven auch als Newtontrajektorien. Sie verbinden in der Regel mehrere Minima und Sattelpunkte auf der PES miteinander. Desweiteren spricht das Verhalten bei Verzweigungen für die Verwendung des RGF-Verfahrens als Reaktionswegmodell.

Reaktionswegverzweigungen treten beispielsweise dann auf, wenn aus einem Reaktanten mehrere Produkte entstehen oder sich aus verschiedenen Ausgangsstoffen ein Produkt bildet. Solche Stellen sind ausgezeichnete Punkte auf der PES, die sogenannten Tal-Rücken-Umschlagpunkte (VRI-Punkte). Sie hängen nicht vom speziell gewählten Reaktionswegmodell ab. Andererseits verzweigen sich gerade Newtontrajektorien in VRI-Punkten. Ihre numerische Bestimmung ist allerdings schwierig.

Eine alternative Herangehensweise an die Reaktionswegbestimmung liefert der Variationsansatz. Hier betrachten wir eine Kette von (endlich vielen) Punkten zwischen Reaktanten- und Produktminimum. Diese wird bei fixierten Enden solange variiert bis sie durch den Sattelpunkt der Potentialbarriere verläuft und zudem die Forderungen an den Minimum-Energie-Pfad erfüllt. Aufgrund der Eindeutigkeit des Gradienten lassen sich sowohl die IRC als auch RGF-Kurven auf diese Weise bestimmen. Insbesondere besteht die Möglichkeit mithilfe des Variationsansatzes die sich in VRI-Punkten verzweigenden Newtontrajektorien zu ermitteln. Kern der Arbeit ist die Entwicklung eines Algorithmus zur Bestimmung solcher sich verzweigender Newtontrajektorien, mittels derer direkt auf den zugehörigen VRI-Punkt geschlossen werden kann.

Der Algorithmus wird abschließend an mehreren zwei- und dreidimensionalen Potentialenergieflächen erfolgreich getestet. Inwieweit sich die programmtechnische Umsetzung des Algorithmus für die Bestimmung von VRI-Punkten auf einer typischerweise hochdimensionalen PES eignet, konnte im Rahmen dieser Arbeit nicht mehr geklärt werden. Hierfür bieten sich weitere Untersuchungen an.

1. Chemische Grundlagen

Das Hauptziel dieser Arbeit liegt in der Bestimmung gewisser singulärer Punkte auf einer hochdimensionalen Fläche. Wir wollen diese Problemstellung jedoch nicht als eine mathematisch abstrakte auffassen, sondern eingebettet wissen in die Theorie der chemischen Reaktionswege. Demzufolge dient dieser erste Abschnitt als Einführung in die für unsere Bedürfnisse relevanten Gebiete der Theoretischen Chemie. Wir wollen zunächst das bedeutende Konzept der Potentialenergiefläche vorstellen und anschließend Reaktionswege auf dieser Fläche erklären.

1.1. Theorie der Potentialenergieflächen

Ein Themenschwerpunkt der Theoretischen Chemie ist die Untersuchung der Dynamik chemischer Reaktionen. Hierfür müssen wir die Bewegung der Atomkerne studieren. Wie wir im Folgenden sehen werden, benötigen wir eine Funktion der potentiellen Energie in Abhängigkeit von der Position der Kerne.¹

1.1.1. Schrödingergleichung und Born-Oppenheimer-Näherung

In der Theoretischen Chemie werden die Bewegungen von Atomkernen und Elektronen innerhalb eines Moleküls durch die zeitabhängige Schrödingergleichung beschrieben. Da man den Zeitanteil durch Separation direkt lösen kann, bleibt die Betrachtung der stationären (zeitunabhängigen) Schrödingergleichung.

Gegeben sei eine molekulare Struktur, bestehend aus M Atomkernen und N Elektronen. Mit m_i , \mathbf{R}_i , Z_i und $\Delta_i^{(n)}$ seien Masse, Position, Ladungszahl und Laplace-Operator der Kerne bezeichnet, wobei $i = 1, \dots, M$ gilt. Analog bezeichnen wir mit m_e , \mathbf{r}_j , e und $\Delta_j^{(e)}$ Masse, Position, Ladung und Laplace-Operator der Elektronen für $j = 1, \dots, N$. Dann ist die stationäre Schrödingergleichung von der Gestalt

$$E\psi(\mathbf{R}, \mathbf{r}) = \left[- \sum_{i=1}^M \frac{\hbar^2}{2m_i} \Delta_i^{(n)} - \sum_{j=1}^N \frac{\hbar^2}{2m_e} \Delta_j^{(e)} - \sum_{i=1}^M \sum_{j=1}^N \frac{Z_i e^2}{4\pi\epsilon_0 |\mathbf{R}_i - \mathbf{r}_j|} + \sum_{i=1}^M \sum_{k<i} \frac{Z_i Z_k e^2}{4\pi\epsilon_0 |\mathbf{R}_i - \mathbf{R}_k|} + \sum_{j=1}^N \sum_{k<j} \frac{e^2}{4\pi\epsilon_0 |\mathbf{r}_j - \mathbf{r}_k|} \right] \psi(\mathbf{R}, \mathbf{r}). \quad (1.1)$$

Bei den ersten beiden Termen der rechten Seite handelt es sich um die Operatoren der kinetischen Energie für Kerne und Elektronen. Die drei anderen Terme sind Potentiale,

¹“Chemistry is knowing the energy as a function of the nuclear coordinates.“ [26]

1. Chemische Grundlagen

die durch die Kern-Elektron-Anziehung, Kern-Kern-Abstoßung und Elektron-Elektron-Abstoßung entstehen. Diese Coulomb-Wechselwirkungen zwischen den Teilchen sind Ursache dafür, daß die Schrödingergleichung abgesehen von einfachsten Systemen nicht analytisch gelöst werden kann. Einen ersten Approximationsansatz bietet die Born-Oppenheimer-Näherung.

Wegen der großen Massendifferenz von Kernen und Elektronen ($m_i \gg m_e$), können wir die Bewegung der Kerne gegenüber der der Elektronen in sehr guter Näherung vernachlässigen, da die Kernbewegung bei normalen Molekülschwingungen in Bezug auf die Elektronenbewegung so langsam ist, daß die Elektronen jeder Veränderung der Kernlage instantan folgen.

Trennen wir also Kern- und Elektronenbewegung voneinander, so läßt sich die Wellenfunktion $\psi(\mathbf{R}, \mathbf{r})$ aus (1.1) als Produkt einer elektronischen Wellenfunktion $\psi_e(\mathbf{R}, \mathbf{r})$ und einer nuklearen Wellenfunktion $\psi_n(\mathbf{R})$ schreiben:

$$\psi(\mathbf{R}, \mathbf{r}) \approx \psi_e(\mathbf{R}, \mathbf{r}) \cdot \psi_n(\mathbf{R}) .$$

Aus diesem Produktansatz erhalten wir durch Einsetzen in (1.1) einerseits die elektronische Schrödingergleichung mit fixierten Kernen,

$$E(\mathbf{R})\psi_e(\mathbf{R}, \mathbf{r}) = \left[- \sum_{j=1}^N \frac{\hbar^2}{2m_e} \Delta_j^{(e)} - \sum_{i=1}^M \sum_{j=1}^N \frac{Z_i e^2}{4\pi\epsilon_0 |\mathbf{R}_i - \mathbf{r}_j|} + \sum_{j=1}^N \sum_{k < j} \frac{e^2}{4\pi\epsilon_0 |\mathbf{r}_j - \mathbf{r}_k|} \right] \psi_e(\mathbf{R}, \mathbf{r}) , \quad (1.2)$$

und andererseits die nukleare Schrödingergleichung,

$$\begin{aligned} [E - E(\mathbf{R})]\psi_n(\mathbf{R}) &= \left[- \sum_{i=1}^M \frac{\hbar^2}{2m_i} \Delta_i^{(n)} + \sum_{i=1}^M \sum_{k < i} \frac{Z_i Z_k e^2}{4\pi\epsilon_0 |\mathbf{R}_i - \mathbf{R}_k|} \right] \psi_n(\mathbf{R}) , \\ E\psi_n(\mathbf{R}) &= \left[- \sum_{i=1}^M \frac{\hbar^2}{2m_i} \Delta_i^{(n)} + \sum_{i=1}^M \sum_{k < i} \frac{Z_i Z_k e^2}{4\pi\epsilon_0 |\mathbf{R}_i - \mathbf{R}_k|} + E(\mathbf{R}) \right] \psi_n(\mathbf{R}) . \end{aligned} \quad (1.3)$$

Für Ein-Elektronen-Systeme wie das Wasserstoffatom ist die elektronische Schrödingergleichung (1.2) exakt lösbar. Wegen der Coulomb-Abstoßung der Elektronen erhalten wir in Mehr-Elektronen-Systemen lediglich eine approximierete Lösung, etwa durch Monte-Carlo-Methoden [22].

Mithilfe der bisherigen Überlegungen haben wir Informationen über die Bewegung der Atomkerne erlangt: Die Schrödingergleichung beschreibt die Bewegungen von Kernen und Elektronen. Dank der Born-Oppenheimer-Näherung können wir diese Bewegungen unabhängig voneinander betrachten. Wir lösen zuerst (approximativ) die elektronische Schrödingergleichung und erhalten daraus $E(\mathbf{R})$, die Energie der Elektronen, als Eigenwert der elektronischen Wellenfunktion $\psi_e(\mathbf{R}, \mathbf{r})$. Setzen wir anschließend $E(\mathbf{R})$ in die nukleare Schrödingergleichung ein, so ergibt sich aus deren Lösung die Bewegung der Atomkerne.

1.1.2. Potentialenergieflächen

Die potentielle Energie in der nuklearen Schrödingergleichung (1.3) setzt sich zusammen aus der Potentialenergie der Kerne,

$$V_{nn}(\mathbf{R}) := \sum_{i=1}^M \sum_{k<i} \frac{Z_i Z_k e^2}{4\pi\epsilon_0 |\mathbf{R}_i - \mathbf{R}_k|}$$

und $E(\mathbf{R})$, der Energie der Elektronen. Die Summe $V_{nn}(\mathbf{R}) + E(\mathbf{R})$ bezeichnen wir als *Potentialenergiefläche* (Potential Energy Surface, PES).

Die Geometrie der PES wird durch $3M$ Koordinaten beschrieben. Diese (hohe) Dimensionalität läßt sich durch die Verwendung interner Koordinaten (Bindungslänge, Valenzwinkel) auf $3M - 6$ (bzw. $3M - 5$ bei linearen Molekülen [7]) reduzieren. Eine weitere Verringerung ist etwa bei organischen Reaktionen möglich, wenn man die Bindungslänge zwischen Kohlen- und Wasserstoffatomen in guter Näherung als konstant während des Reaktionsverlaufs annimmt [28]. Trotzdem ist es nicht möglich, die gesamte PES zu berechnen. In der Regel ist die Potentialenergiefläche keine analytische Funktion, sondern mittels numerischer Verfahren Punkt für Punkt gegeben. Wie man leicht einsieht, wächst die Zahl der so zu ermittelnden Punkte mit zunehmender Komplexität der untersuchten Moleküle derart rapide, daß man sich in der Theoretischen Chemie bei der Berechnung der PES auf die „chemisch interessanten“ Bereiche beschränkt, in denen lokale Minima und Sattelpunkte auftreten.

Wie bereits erwähnt, genügt uns für die mathematische Herangehensweise eine abstrakte Vorstellung von der Potentialenergiefläche: Wir nehmen an, daß die PES durch eine skalare Funktion der Koordinaten des betrachteten Moleküls in jedem für uns interessanten Punkt gegeben ist. Daraus ergibt sich folgende Definition:

Definition

Sei $K \subset \mathbb{R}^n$. Dann bezeichnen wir die Abbildung $E : K \rightarrow \mathbb{R}$ als n -dimensionale *Potentialenergiefläche*. K heißt *Konfigurationsraum* der Potentialenergiefläche E .

1. Chemische Grundlagen

Um im weiteren Verlauf numerische Methoden wie beispielsweise das Newtonverfahren auf die PES anwenden zu können, benötigen wir den Gradienten, die Hessematrix und deren Adjunkte. Wir fordern somit insbesondere, daß $E \in C^k(K, \mathbb{R})$ mit $k \geq 2$ gilt.

Definition

Sei $\mathbf{x} = (x_1, \dots, x_n)^T \in K$. Dann heißen der Vektor der ersten Ableitungen $\mathbf{g} : K \rightarrow \mathbb{R}^n$ und die Matrix der zweiten Ableitungen $\mathbf{H} : K \rightarrow \mathbb{R}^{n \times n}$ gemäß

$$\mathbf{g}(\mathbf{x}) = \nabla E(\mathbf{x}) = \left(\frac{\partial E(\mathbf{x})}{\partial x_1}, \dots, \frac{\partial E(\mathbf{x})}{\partial x_n} \right)^T,$$
$$\mathbf{H}(\mathbf{x}) = \nabla \nabla^T E(\mathbf{x}) = \left(\frac{\partial^2 E(\mathbf{x})}{\partial x_i \partial x_j} \right)_{\substack{i=1, \dots, n \\ j=1, \dots, n}}$$

Gradient und Hessematrix von E .

Bemerkung

Die Hessematrix ist symmetrisch; ihre Eigenwerte sind reell. Sind sämtliche Eigenwerte von \mathbf{H} in einem Punkt $\mathbf{x} \in K$ positiv (negativ), so ist die Potentialenergiefläche E konvex (konkav) in \mathbf{x} . Wir definieren den Richtungsvektor $\mathbf{y} \in \mathbb{R}^n$ mit $\|\mathbf{y}\|_2 = 1$ sowie die differenzierbare Funktion $\varphi(t) := E(\mathbf{x} + t\mathbf{y})$. Dann erhalten wir mithilfe der Richtungsableitung folgende Aussage über die Krümmung von E :

$$\frac{\partial^2 E(\mathbf{x})}{\partial \mathbf{y}^2} = \left. \frac{d^2 \varphi(t)}{dt^2} \right|_{t=0} = \left. \frac{d^2 E(\mathbf{x} + t\mathbf{y})}{dt^2} \right|_{t=0} = \mathbf{y}^T \mathbf{H}(\mathbf{x} + t\mathbf{y}) \mathbf{y} \Big|_{t=0} = \mathbf{y}^T \mathbf{H}(\mathbf{x}) \mathbf{y}.$$

Der Term $\mathbf{y}^T \mathbf{H}(\mathbf{x}) \mathbf{y}$ gibt die Krümmung der PES im Punkt $(\mathbf{x}, E(\mathbf{x}))$ entlang der Richtung \mathbf{y} an. Ist \mathbf{y} zudem ein Eigenvektor von $\mathbf{H}(\mathbf{x})$, dann heißt \mathbf{y} *Hauptkrümmungsrichtung*.

Definition

Sei $\mathbf{M} \in \mathbb{R}^{n \times n}$ und bezeichne $\mathbf{M}_{ij} \in \mathbb{R}^{(n-1) \times (n-1)}$ diejenige Matrix, die aus \mathbf{M} durch Streichen der i -ten Zeile und j -ten Spalte entsteht. Bilden wir \hat{m}^{ij} gemäß

$$\hat{m}^{ij} := (-1)^{i+j} \det(\mathbf{M}_{ij}),$$

dann ist $\mathbf{M}^A := (\hat{m}^{ji})$ die *Adjunkte* von \mathbf{M} .

Bemerkung

Die Adjunkte der Hessematrix bezeichnen wir mit \mathbf{A} . Zwischen \mathbf{A} und \mathbf{H} bestehen folgende Zusammenhänge:

- (i) Bezeichnet \mathbf{I}_n die n -dimensionale Einheitsmatrix, so gilt $\mathbf{A}\mathbf{H} = \det(\mathbf{H})\mathbf{I}_n$.
- (ii) Die Adjunkte \mathbf{A} ist genau dann regulär, wenn \mathbf{H} es auch ist.

- (iii) Hessematrix \mathbf{H} und Adjunkte \mathbf{A} haben dieselben Eigenvektoren.
- (iv) Seien die Eigenwerte von \mathbf{H} mit λ_i bezeichnet. Dann gilt für die Eigenwerte von \mathbf{A} gerade $\tilde{\lambda}_i = \prod_{j \neq i} \lambda_j$.

Beweise für die Aussagen (i)–(iv) sind im Anhang B.1 nachzulesen.

Abschließend sei der Begriff der Äquipotentialfläche eingeführt. Im Folgenden werden wir häufig Höhenlinienbilder, also eindimensionale Äquipotentialflächen, zur Visualisierung von Potentialenergieflächen verwenden.

Definition

Die Menge $M_c := \{\mathbf{x} \in K \mid E(\mathbf{x}) = c\}$ heißt *Äquipotentialfläche* zum Niveau c .

Beispiele

Abbildung 1.1 zeigt zwei Beispiele für zweidimensionale Potentialenergieflächen als Höhenlinienbilder. Die in 1.1(a) dargestellte PES ist nach [43] gegeben durch (A.1). Im rechten Bild sehen wir das gemäß (A.2) definierte sogenannte Müller-Brown-Potential [31].

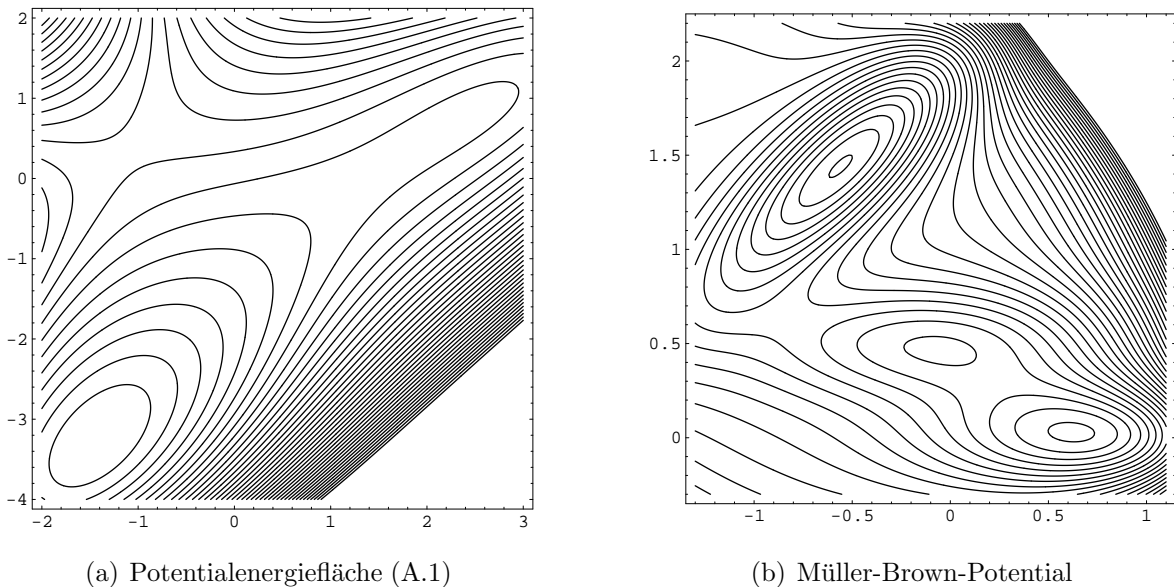


Abbildung 1.1: Zweidimensionale Potentialenergieflächen als Höhenlinienbilder

1.2. Reaktionswege auf Potentialenergieflächen

Durch die Trennung von Kern- und Elektronenbewegung nach der Born-Oppenheimer-Näherung können wir uns eine chemische Reaktion als Bewegung der Kerne auf der Potentialenergiefläche E vorstellen. Um den Verlauf einer Reaktion nachzuvollziehen, benötigen

1. Chemische Grundlagen

wir daher eine genaue Beschreibung der PES. Wie in Abschnitt 1.1.2 erläutert, ist es in der Realität nicht möglich, E vollständig zu berechnen; wir betrachten einen „chemisch relevanten“ Teilbereich der Fläche. Nach der Theorie des Übergangszustands stellt man sich unter einer Reaktion eine Trajektorie im Konfigurationsraum K vor, die, von den Reaktanten ausgehend, über eine Übergangsstruktur zu den Produkten führt. Mithilfe dieses Ansatzes genügen uns lokale Informationen über die PES entlang jener Trajektorie. Über Unzulänglichkeiten und Verbesserungen der Theorie des Übergangszustands wird an anderer Stelle diskutiert (etwa [50]).

An dieser Stelle sei angemerkt, daß wir im Folgenden häufig von „Tälern“, „Talrücken“ o.ä. im Zusammenhang mit Potentialenergieflächen sprechen werden. Im Fall der Dimensionalität $n = 2$ sind diese Begriffe sofort einleuchtend, im höherdimensionalen Fall allerdings absurd. Trotzdem verwenden wir die erwähnten Bezeichnungen sinnessprechend für beliebige n . Ein Tal ist dann ein Teilbereich der PES mit konvexen Äquipotentialflächen.

1.2.1. Theorie des Übergangszustands

Wir wollen kurz die Theorie des Übergangszustands (Transition State Theory, TST) nach Eyring und Polanyi [12] erläutern. Edukte und Produkte sind energetisch stabile Zustände (Gleichgewichtszustände) und bilden daher lokale Minima auf der PES. Beim Übergang der Ausgangsstoffe in die Reaktionsprodukte muß eine Potentialbarriere überwunden werden. In simplifizierender Darstellung wird die Reaktion durch jenen Pfad repräsentiert, der die beiden Minima miteinander verbindet und dabei die dazwischen liegenden „Energieberge“² an niedrigster Stelle überwindet. Das Energiemaximum des Reaktionspfades nennen wir Übergangsstruktur (Transition Structure, TS). In [28] sind die Mindestanforderungen an eine solche Struktur $\mathbf{x}^* \in K$ wie folgt zusammengefaßt:

- (i) \mathbf{x}^* ist ein stationärer Punkt. Das heißt, es gelten $\mathbf{g}(\mathbf{x}^*) = 0$ und $\det \mathbf{H}(\mathbf{x}^*) \neq 0$. Dies garantiert die Invarianz der TS bezüglich der speziellen Wahl des Koordinatensystems.
- (ii) $\mathbf{H}(\mathbf{x}^*)$ hat genau einen negativen Eigenwert. Hätte $\mathbf{H}(\mathbf{x}^*)$ keinen negativen Eigenwert, so handelte es sich bei \mathbf{x}^* um ein lokales Minimum, mithin um einen energetisch stabilen Zustand. Läge mehr als ein negativer Eigenwert vor, so gäbe es eine TS niedrigerer Energie mit genau einem negativen Eigenwert.

²„Der chemische Ausgangs- und Endzustand sind zwei Minima der Energie, die von einer Kette von Energiebergen getrennt sind. Der niedrigste Übergang, der über diese die Reaktionsträgheit verursachenden Energieberge führt, ergibt die Höhe der Aktivierungsenergie.“ [12]

- (iii) \mathbf{x}^* hat den höchsten Energiewert aller Punkte des stetigen Reaktionspfades zwischen Edukten und Produkten. Es ist zu beachten, daß hierbei keine Aussage über den Verlauf des Pfades getroffen wird; er ist in gewisser Weise beliebig.
- (iv) Für alle anderen Punkte $\tilde{\mathbf{x}}$, die (i)–(iii) erfüllen, gilt $E(\mathbf{x}^*) \leq E(\tilde{\mathbf{x}})$.

Definition

Für einen Punkt $\mathbf{x} \in K$ des Konfigurationsraums gelte $\det \mathbf{H}(\mathbf{x}) \neq 0$. Dann heißt \mathbf{x} *nicht entartet*, und die Anzahl aller negativen Eigenwerte von $\mathbf{H}(\mathbf{x})$ nennen wir *Index* von \mathbf{x} . Gilt zudem, daß $\mathbf{g}(\mathbf{x}) = 0$, dann ist \mathbf{x} *stationärer Punkt* in K .

Einen stationären Punkt mit Index Null bezeichnen wir als *Minimum*, einen stationären Punkt mit Index $i > 0$ als *Sattelpunkt vom Index i* .

Bemerkung

Gemäß der Forderung (ii) an eine TS $\mathbf{x}^* \in K$ handelt es sich bei \mathbf{x}^* um einen Sattelpunkt vom Index 1. Häufig wird ein Reaktant durch mehrere lokale Minima ähnlicher Energie auf der PES repräsentiert, die durch „niedrige“ Sattelpunkte voneinander getrennt sind. Jedes einzelne Minimum steht dabei für einen der beteiligten Konformere, Komplexe, etc. Die chemische Reaktion führt anschließend das gesamte System über höher liegende Sattelpunkte vom Index 1 [17].

Beispiel

Abbildung 1.2 zeigt die stationären Punkte des Müller-Brown-Potentials. Die genauen Koordinaten sind in (5.7) gegeben.

Aus offensichtlichen Gründen besteht ein großes Interesse, stationäre Punkte im Konfigurationsraum der PES zu finden. Vergleichsweise einfach gestaltet sich die Suche nach Minima mit der Methode des steilsten Abstiegs. Sattelpunkte vom Index 1 findet man etwa mit dem RGF-Verfahren (Reduced Gradient Following). Beide Methoden werden im weiteren Verlauf der Arbeit vorgestellt.

1.2.2. Minimum-Energie-Weg und Reaktionskoordinate

Gemäß der Transition State Theory ist eine chemische Reaktion eine Trajektorie im Konfigurationsraum K , die zwei Minima über einen Sattelpunkt vom Index 1 miteinander verbindet, wobei ihr exakter Verlauf nicht festgelegt ist. Wir wollen diese Trajektorie *Minimum-Energie-Weg* (Minimum Energy Path, MEP) oder auch *Reaktionspfad* (Reaction Path, RP) nennen, falls sie folgende Eigenschaften erfüllt:

- (i) Entlang des MEP wächst die potentielle Energie von den Reaktanten bis zur Übergangsstruktur streng monoton, vom Sattelpunkt zu den Produkten fällt sie monoton.

1. Chemische Grundlagen

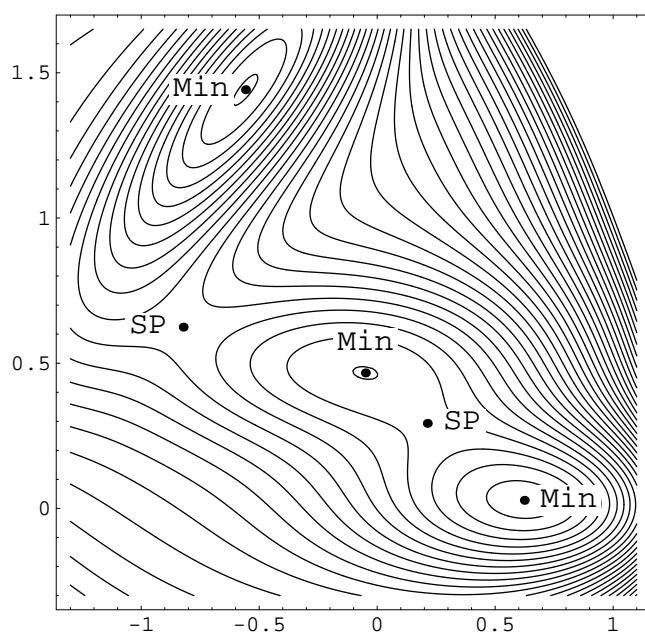


Abbildung 1.2: Stationäre Punkte des Müller-Brown-Potentials

- (ii) Bis zum Sattelpunkt verläuft der MEP durch ein Tal.
- (iii) Die stationären Punkte auf dem MEP sind entweder Minima oder Sattelpunkte vom Index 1.

Eine genauere mathematische Fixierung des Minimum-Energie-Wegs erweist sich aufgrund der Vielfältigkeit der PES als schwierig. So ist beispielsweise keinesfalls gesichert, daß zwischen Minimum und TS ein Talgrund verläuft. Daher müssen wir „vernünftig“ gewählte Reaktionspfade, die die Bedingungen (i)–(iii) erfüllen, als gleichwertig betrachten; es ist nicht zu entscheiden, welcher dieser gleichwertigen Pfade der tatsächliche Reaktionsweg ist. Einen eindeutigen MEP gibt es nicht.

Auch die verschiedenen Methoden zur Bestimmung des Minimum-Energie-Wegs, wie etwa

- (i) Methode des steilsten Abstiegs vom Sattelpunkt; Intrinsic Reaction Coordinate (IRC),
- (ii) Gradient Extremal (GE),
- (iii) kontinuierliches Newtonverfahren; Reduced Gradient Following (RGF)

würdigen zwar unterschiedliche Aspekte des MEP-Konzepts, ermitteln aber nicht *den* einen MEP. Wir werden die einzelnen Verfahren sowie ihre Vor- und Nachteile detailliert in den Abschnitten 2 und 3 vorstellen.

Bemerkungen

- (i) In der Literatur und auch in dieser Arbeit werden RP und MEP synonym verwendet.
- (ii) An dieser Stelle wollen wir ausdrücklich erwähnen, daß es sich bei dem Minimum-Energie-Weg um ein „künstliches chemisches Instrument“ [17] handelt, dessen größter Vorteil in der Überwindung des Dimensionalitätsproblems besteht. Einige unserer physikalischen Vorstellungen vom Verlauf chemischer Reaktionen werden durch das MEP-Konzept nicht bedient:
 - a) Im MEP werden die Terme der kinetischen Energie vernachlässigt. Wir gebrauchen trotzdem den Begriff „Trajektorie“ im Zusammenhang mit dem MEP, auch wenn dieser gewöhnlich für die Lösung von Bewegungsgleichungen verwendet wird.
 - b) Der quantenmechanische Tunneleffekt wird im MEP-Konzept nicht berücksichtigt.

Abschließend möchten wir noch den Begriff der Reaktionskoordinate (Reaction Coordinate, RC) einführen. Wie bereits erläutert ist das größte Problem bei der Analyse von Potentialenergieflächen deren hohe Dimensionalität. Mithilfe des Minimum-Energie-Pfads reduzieren wir dieses Problem auf die Untersuchung einer eindimensionalen Kurve. Dann heißt jede Parametrisierung des MEP gemäß

$$\mathbf{x}(t) = (x_1(t), \dots, x_n(t))^T$$

Reaktionskoordinate.

1.3. Verzweigungen auf Potentialenergieflächen

Bei gewissen chemischen Reaktionen treten Verzweigungen (Bifurkationen) auf. Geht etwa während einer Reaktion die C_s -Symmetrie³ verloren, dann muß sich der Reaktionsweg an einer Stelle (symmetrisch) verzweigen, denn es liegen einerseits zwei Produkte vor, die durch die C_s -Symmetrieoperation ineinander überführt werden können, aber andererseits nur ein Edukt, das durch diese Operation in sich selbst übergeht. Der MEP startet beim Edukt; die Symmetrie bleibt eine Weile erhalten, bevor eine Bifurkation des MEP auftritt [48]. Eine einzige TS kann zu mehreren Reaktionsprodukten führen, falls sich der MEP im Bereich zwischen TS und Produkten verzweigt. Anders formuliert teilen sich in dem

³Zur Punktgruppe C_s gehören Moleküle, die außer der Identität und einer Spiegelebene keine weiteren Symmetrieelemente besitzen, z.B. Stickstoffwasserstoffsäure, Chinolin.

1. Chemische Grundlagen

Fall mehrere Minimum-Energie-Pfade diesselbe TS. Diese Situation tritt ein, falls mehrere Sattelpunkte vom Index 1 direkt aufeinander folgen ohne dazwischenliegende Minima [11].

Wir wollen anmerken, daß offen bleibt, ob sich der Reaktionsweg vor, in der Nähe oder nach der Übergangsstruktur verzweigt. Im ersten Fall muß es allerdings zwei TS geben. Als Beispiel hierfür sei die Isomerisierung von Formaldehyd zu Hydroxycarben genannt [15]. Hingegen ereignet sich die Bifurkation bei der Ringöffnung von Cyclopropylen beinahe an der Übergangsstruktur [48]. Desweiteren können wir Verzweigungen bei der Bindungsverschiebung von Cyclooctatetraen [5], aber auch bei komplexeren, etwa perizyklischen organischen Reaktionen beobachten [11].

Punkte, an denen sich der Minimum-Energie-Pfad verzweigt, nennen wir *Bifurkationspunkte* (Bifurcation Points, BP). Die Lage eines BP hängt damit von der speziellen, nach Abschnitt 1.2.2 keinesfalls eindeutigen Wahl des MEP ab. Mittels folgender Überlegungen können wir jedoch unabhängig vom speziell gewählten Reaktionswegmodell Verzweigungen auf der PES lokalisieren: Nach der TS läuft der MEP abwärts entlang eines Tals in Richtung Edukte. Dieser Teil des Pfads ist in gewissem Sinne stabil; die Trajektorie wird sich im Talgrund befinden. Schlägt nun das Tal in einen Rücken um, so erhalten wir einen instabilen Pfad; kleinste Ablenkungen der Trajektorien führt zu immer größeren Ausschereungen. Das Tal wird in zwei Täler aufgespalten. Wir können daher Bifurkationen in der Nähe dieser *Tal-Rücken-Umschlagpunkte* (Valley-Ridge Inflection Points, VRI-Punkte) erwarten [48].

Während der MEP sich im Tal befindet, ist der Eigenwert desjenigen Eigenvektors der Hessematrix, der senkrecht zum Gradienten verläuft, positiv und wird im weiteren Verlauf kleiner. Auf dem Rücken ist dieser Eigenwert schließlich negativ. Damit erhalten wir folgende Definition für VRI-Punkte:

Definition

Sei $\mathbf{x}^{\text{vri}} \in K$ mit $\mathbf{g}(\mathbf{x}^{\text{vri}}) \neq \mathbf{0}$ gegeben und gelte, daß der Eigenwert desjenigen Eigenvektors von $\mathbf{H}(\mathbf{x}^{\text{vri}})$, der senkrecht zu $\mathbf{g}(\mathbf{x}^{\text{vri}})$ steht, gerade Null ist. Dann heißt \mathbf{x}^{vri} *Tal-Rücken-Umschlagpunkt* (Valley-Ridge Inflection Point, VRI-Punkt).

Bemerkungen

- (i) Bezeichne \mathbf{u}^{zero} den Eigenvektor der Hessematrix mit Eigenwert Null. Demnach gilt in VRI-Punkten die Beziehung

$$\mathbf{g}^T \mathbf{u}^{\text{zero}} = 0 . \quad (1.4)$$

- (ii) VRI-Punkte sind unabhängig von der speziellen Wahl des Reaktionswegmodells; in der Regel stimmen sie nicht mit den Bifurkationspunkten der verschiedenen Model-

le überein. Allerdings ist die Existenz eines VRI-Punktes hinreichende Bedingung dafür, daß in dessen Umgebung eine Verzweigung des Reaktionspfades auftritt. Man bezeichnet VRI-Punkt und BP als adjazent. Das heißt, zwischen beiden liegt kein mathematisch interessanter Punkt.

- (iii) Nach obiger Definition steht $\mathbf{g}(\mathbf{x}^{\text{vri}})$ senkrecht zu \mathbf{u}^{zero} . Damit liegt der Gradient nicht im Kern der Hessematrix, und es gilt nach [16] die Beziehung

$$\text{rk}(\mathbf{H}(\mathbf{x}^{\text{vri}})|\mathbf{g}(\mathbf{x}^{\text{vri}})) = \text{rk} \begin{pmatrix} \frac{\partial^2 E(\mathbf{x}^{\text{vri}})}{\partial x_1 \partial x_1} & \dots & \frac{\partial^2 E(\mathbf{x}^{\text{vri}})}{\partial x_1 \partial x_n} & \frac{\partial E(\mathbf{x}^{\text{vri}})}{\partial x_1} \\ \vdots & \ddots & \vdots & \vdots \\ \frac{\partial^2 E(\mathbf{x}^{\text{vri}})}{\partial x_n \partial x_1} & \dots & \frac{\partial^2 E(\mathbf{x}^{\text{vri}})}{\partial x_n \partial x_n} & \frac{\partial E(\mathbf{x}^{\text{vri}})}{\partial x_n} \end{pmatrix} < n . \quad (1.5)$$

Punkte, die die Ungleichung (1.5) erfüllen, sind die außerordentlichen Singularitäten des desingularisierten kontinuierlichen Newtonverfahrens, welches in Abschnitt 3.2 vorgestellt wird.

- (iv) Da lediglich zweidimensionale Potentialenergieflächen anschaulich darstellbar sind, werden VRI-Punkte in der Literatur häufig als singuläre (0-dimensionale) Punkte angegeben. Dies ist irreführend, denn für n -dimensionale PES können die VRI-Punkte eine $(n - 2)$ -dimensionale Mannigfaltigkeit in K bilden [27], was beispielsweise in [21] anhand der PES von Wasser bestätigt wird.
- (v) Spaltet sich das Tal am VRI-Punkt symmetrisch auf, so sprechen wir von einem *symmetrischen VRI-Punkt*. Symmetrische VRI-Punkte lassen sich mittels Branin-Verfahren [4] bestimmen. Problematischer ist die Berechnung asymmetrischer VRI-Punkte.

Beispiel

Die durch (A.1) gegebene Potentialenergiefläche enthält einen asymmetrischen VRI-Punkt in $(0, 0)$. Ändert man allerdings diese PES gemäß (A.3), so erhalten wir eine Potentialenergiefläche mit symmetrischem VRI-Punkt im Ursprung. Abbildung 1.3 zeigt beide Flächen, ihre stationären Punkte und insbesondere auch ihren VRI-Punkt.

An dieser Stelle wollen wir ein zweites Kriterium für VRI-Punkte beweisen. Es wird uns im weiteren Verlauf bei der numerischen VRI-Bestimmung helfen.

Lemma

Für VRI-Punkte \mathbf{x}^{vri} gilt die Beziehung

$$\mathbf{A}(\mathbf{x}^{\text{vri}})\mathbf{g}(\mathbf{x}^{\text{vri}}) = \mathbf{0} , \quad (1.6)$$

1. Chemische Grundlagen

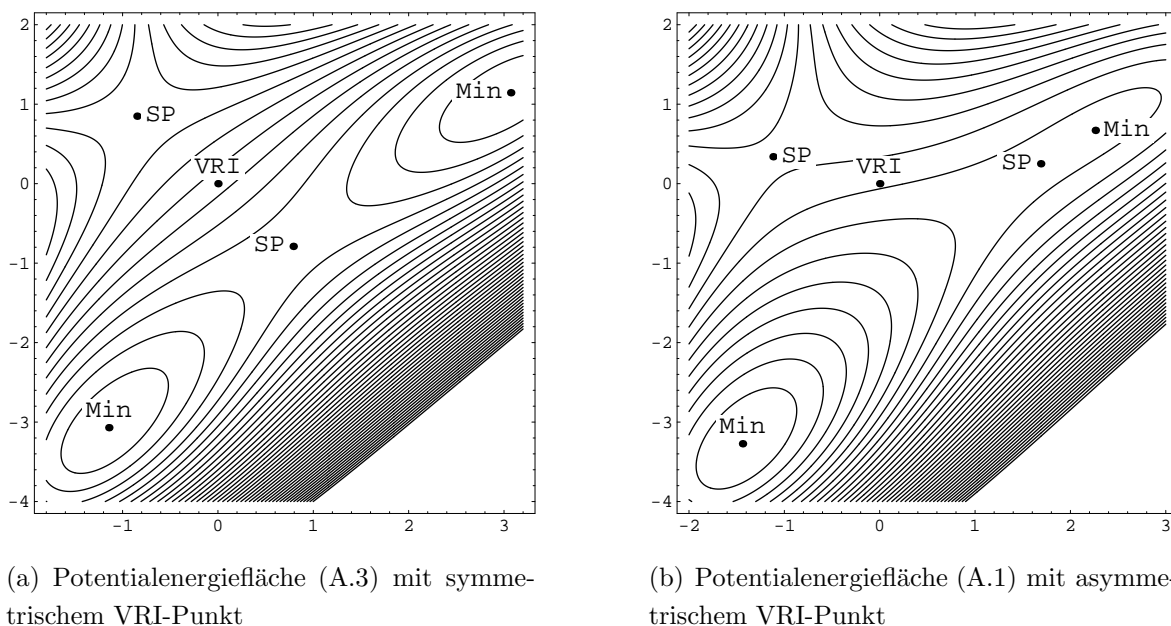


Abbildung 1.3: Symmetrische und asymmetrische VRI-Punkte

wobei wir mit \mathbf{A} die Adjunkte der Hessematrix bezeichnet hatten.

Beweis

Der Beweis folgt der Argumentation von [41]. Seien mit \mathbf{u}^i die Eigenvektoren von \mathbf{H} und mit λ_i die zugehörigen Eigenwerte bezeichnet. Aus der Eigenwertgleichung $\lambda_i \mathbf{u}^i = \mathbf{H} \mathbf{u}^i$ folgt nach Multiplikation mit \mathbf{A} :

$$\lambda_i \mathbf{A} \mathbf{u}^i = \mathbf{A} \mathbf{H} \mathbf{u}^i = \det(\mathbf{H}) \mathbf{u}^i = \left(\prod_{j=1}^n \lambda_j \right) \mathbf{u}^i .$$

Für $\lambda_i \neq 0$ hat \mathbf{A} damit die Eigenvektoren \mathbf{u}^i zu den Eigenwerten $(\prod_{j=1}^n \lambda_j) / \lambda_i$. Nach Definition ist aber ein Eigenwert von \mathbf{x}^{vri} gleich Null, o.B.d.A. $\lambda_1 = 0$. Dann haben auch die Eigenvektoren $\mathbf{u}^2, \dots, \mathbf{u}^n$ von \mathbf{A} den Eigenwert Null. Wir schreiben \mathbf{g} als Linearkombination der \mathbf{u}^i , also $\mathbf{g} = \sum_{j=1}^n \xi_j \mathbf{u}^j$. Da in \mathbf{x}^{vri} der Gradient \mathbf{g} senkrecht auf \mathbf{u}^1 steht, ist $\xi_1 = 0$ und wir erhalten die Beziehung

$$\mathbf{A}(\mathbf{x}^{\text{vri}}) \mathbf{g}(\mathbf{x}^{\text{vri}}) = \xi_1 \left(\prod_{j=2}^n \lambda_j \right) \mathbf{u}^1(\mathbf{x}^{\text{vri}}) + \sum_{i=2}^n \left[\xi_i \left(\frac{\prod_{j=1}^n \lambda_j}{\lambda_i} \right) \mathbf{u}^i(\mathbf{x}^{\text{vri}}) \right] = \mathbf{0} .$$

□

Bemerkung

Nach [43] indizieren Punkte \mathbf{x} , die der Bedingung

$$\mathbf{g}^T(\mathbf{x}) \mathbf{A}(\mathbf{x}) \mathbf{g}(\mathbf{x}) = 0 \tag{1.7}$$

genügen, gerade eine Grenzkurve zwischen Tal- und Talrücken-Regionen auf der PES. Wegen (1.6) liegen VRI-Punkte auf dieser Kurve.

Wir werden die Erforschung der VRI-Punkte in den folgenden Abschnitten vertiefen. Abschließen wollen wir unseren kurzen Einblick in die chemischen Grundlagen mit einer weiteren Klasse von entarteten Punkten auf der PES, die wir, ebenso wie die VRI-Punkte, unabhängig von der speziellen Wahl eines Reaktionspfades definieren: die Umkehrpunkte. Auch sie behandeln wir im Folgenden noch detaillierter.

Definition

Einen Punkt $\mathbf{x}^{\text{tp}} \in K$, für den $\det \mathbf{H}(\mathbf{x}^{\text{tp}}) = 0$ und

$$\text{rk}(\mathbf{H}(\mathbf{x}^{\text{tp}}) | \mathbf{g}(\mathbf{x}^{\text{tp}})) = \text{rk} \begin{pmatrix} \frac{\partial^2 E(\mathbf{x}^{\text{tp}})}{\partial x_1 \partial x_1} & \cdots & \frac{\partial^2 E(\mathbf{x}^{\text{tp}})}{\partial x_1 \partial x_n} & \frac{\partial E(\mathbf{x}^{\text{tp}})}{\partial x_1} \\ \vdots & \ddots & \vdots & \vdots \\ \frac{\partial^2 E(\mathbf{x}^{\text{tp}})}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 E(\mathbf{x}^{\text{tp}})}{\partial x_n \partial x_n} & \frac{\partial E(\mathbf{x}^{\text{tp}})}{\partial x_n} \end{pmatrix} = n$$

gilt, nennen wir *Umkehrpunkt* (Turning Point, TP).

2. Wegeverfolgung auf Potentialenergieflächen

Wie in Abschnitt 1 erläutert, interessieren wir uns zunächst vorrangig für die stationären Punkte auf einer PES. Insbesondere suchen wir nach Methoden, die uns Reaktionswege zwischen zwei Minima liefern und dabei einen Sattelpunkt vom Index 1 überqueren. Aus den Bedingungen, die wir an den MEP stellten, ergab sich kein eindeutiger Reaktionsweg. Mithin gibt es eine Vielzahl verschiedener Ansätze, die zwar jeweils Aspekte der MEP-Theorie erfüllen, aber zu unterschiedlichen Trajektorien führen.

In diesem Abschnitt wollen wir einige dieser Methoden vorstellen und ihre jeweiligen Vor- und Nachteile sowie ihr Verhalten beim Auftreten von Bifurkationen diskutieren. Dabei nehmen wir keine Rücksicht auf das dynamische Verhalten der Moleküle, sondern gehen von geometrisch definierten Reaktionswegen aus, die lediglich die Eigenschaften der PES berücksichtigen [36]. Das RGF-Verfahren sei an dieser Stelle lediglich genannt. Ihm wollen wir aufgrund seiner Bedeutung für diese Arbeit den gesamten Abschnitt 3 widmen.

2.1. Die Methode des steilsten Abstiegs

Die Methode des *steilsten Abstiegs* (Steepest Descent, SD) ist theoretisch die einfachste Möglichkeit, Minima auf der PES zu finden. Es ist leicht einzusehen, daß $-\mathbf{g}(\mathbf{x})$ ein Vektor in Richtung des steilsten Abstiegs ist. Denn betrachten wir für $\mathbf{y} \in \mathbb{R}^n$ mit $\|\mathbf{y}\|_2 = 1$ die differenzierbare Funktion $\varphi(t) := E(\mathbf{x} + t\mathbf{y})$, so gilt für deren Ableitung an der Stelle Null:

$$\varphi'(0) = \left(\frac{\partial E(\mathbf{x} + t\mathbf{y})}{\partial x_1}, \dots, \frac{\partial E(\mathbf{x} + t\mathbf{y})}{\partial x_n} \right) \Big|_{t=0} \cdot \mathbf{y} = \mathbf{g}(\mathbf{x})^T \mathbf{y}. \quad (2.1)$$

Setzen wir $\mathbf{g}(\mathbf{x}) \neq \mathbf{0}$ voraus, dann erhalten wir für $\mathbf{y} := -\mathbf{g}(\mathbf{x})/\|\mathbf{g}(\mathbf{x})\|_2$ ein Minimum in (2.1). Mit anderen Worten: $-\mathbf{g}(\mathbf{x})$ gibt die Richtung des steilsten Abstiegs von E in \mathbf{x} an [25]. Somit führt uns, von einem beliebigen Punkt $\mathbf{x}^0 \in K$ mit $\mathbf{g}(\mathbf{x}^0) \neq \mathbf{0}$ ausgehend, die zum Anfangswertproblem

$$\frac{d\mathbf{x}(t)}{dt} = -\mathbf{g}(\mathbf{x}(t)), \quad \mathbf{x}(0) = \mathbf{x}^0 \quad (2.2)$$

gehörige Kurve monoton fallend zu einem Minimum (falls existent auf der PES und von pathologischen Ausnahmefällen⁴ abgesehen), da die Energie sich verringert, falls t wächst:

$$0 \geq -\left\| \frac{d\mathbf{x}}{dt} \right\|^2 = -\left(\frac{d\mathbf{x}}{dt} \right)^T \left(\frac{d\mathbf{x}}{dt} \right) \stackrel{(2.2)}{=} (\mathbf{g}(\mathbf{x}))^T \left(\frac{d\mathbf{x}}{dt} \right) = \frac{dE(\mathbf{x}(t))}{dt}. \quad (2.3)$$

⁴In [25] sind als Beispiele solcher Ausnahmefälle unter anderem die Spiralfunktion und die Verallgemeinerte Rosenbrock-Funktion angegeben.

Das heißt, wir haben $E \xrightarrow{t \rightarrow \bar{t}} 0$ für ein $\bar{t} \leq \infty$, falls E nach unten beschränkt ist. Das Verfahren ist frei von Singularitäten und Verzweigungen, da die Tangente in jedem Punkt des Konfigurationsraums K eindeutig gegeben ist. Parametrisieren wir um⁵ und verwenden statt t die Bogenlänge s , dann erhalten wir in (2.2):

$$\bar{\mathbf{t}}(s) = \frac{d\mathbf{x}}{ds} = -\frac{\mathbf{g}(\mathbf{x})}{\|\mathbf{g}(\mathbf{x})\|_2} =: -\mathbf{w}(s) \quad (2.4)$$

mit $\bar{\mathbf{t}}(s)$ als Tangente und $\mathbf{w}(s)$ als normiertem Gradientenvektor.

2.1.1. Die intrinsische Reaktionskoordinate IRC

Als Startpunkt der Methode des steilsten Abstiegs wählten wir gemäß (2.2) einen nicht-stationären Punkt \mathbf{x}^0 und somit insbesondere keinen Sattelpunkt vom Index 1. Gemäß der Theorie des Übergangszustands muß ein solcher jedoch Teil des Reaktionsweges sein. Um dieses Problem zu lösen, gehen wir vom Sattelpunkt zunächst einen (infinitesimalen) Schritt weg in Richtung des zum negativen Eigenwert der Hessematrix gehörigen Eigenvektors, legen den so ermittelten Punkt als \mathbf{x}^0 fest und berechnen anschließend den weiteren Verlauf des Reaktionsweges mittels (2.4). Bezeichnen wir mit \mathbf{x}^{\min} das angesteuerte Minimum und mit \mathbf{y} die Richtung des zum kleinsten Eigenwert der Hessematrix $\mathbf{H}(\mathbf{x}^{\min})$ gehörigen Eigenvektors, dann trifft die durch (2.4) errechnete SD-Kurve aus \mathbf{y} kommend in \mathbf{x}^{\min} ein.

Ein solcher Reaktionsweg, der von der TS ausgehend abwärts zu Reaktanten und Produkten führt, heißt nach Fukui [13] *intrinsische Reaktionskoordinate* (Intrinsic Reaction Coordinate, IRC). Ein Algorithmus zur Bestimmung der IRC hätte folgende Gestalt:

- (i) Gehe vom Sattelpunkt \mathbf{x}^{sp} einen Schritt in Richtung des zum negativen Eigenwert der Hessematrix $\mathbf{H}(\mathbf{x}^{\text{sp}})$ gehörigen Eigenvektors zu einem Punkt \mathbf{x}^0 .
- (ii) Die IRC berechnet sich dann als Kurve des steilsten Abstiegs gemäß

$$\mathbf{x}^{i+1} = \mathbf{x}^i + \alpha \cdot \bar{\mathbf{t}} \stackrel{(2.4)}{=} \mathbf{x}^i - \alpha \cdot \frac{\mathbf{g}(\mathbf{x}^i)}{\|\mathbf{g}(\mathbf{x}^i)\|}$$

für $i = 0, 1, \dots$ und mit einer Schrittweite $\alpha > 0$. Der Abbruch erfolgt, wenn $\|\mathbf{g}(\mathbf{x}^k)\| < \varepsilon$ für ein $k < \infty$ bei beliebig kleinem $\varepsilon > 0$.

Der Algorithmus liefert uns als Reaktionspfad somit eine endliche Folge von Punkten $\{\mathbf{x}^i\}_{0 \leq i \leq k}$. Nach (2.3) konvergiert $\{\mathbf{g}(\mathbf{x}^i)\}_{0 \leq i \leq k}$ gegen Null. Daraus ist aber noch kein Konvergenzverhalten für $\{\mathbf{x}^i\}_{0 \leq i \leq k}$ abzuleiten.

⁵SD-Kurven sind unabhängig von der Parametrisierung und der Wahl des Koordinatensystems definierbar [18].

2. Wegeverfolgung auf Potentialenergieflächen

Wegen ihrer Simplität gilt die IRC nicht nur als bedeutendstes Konzept eines RP, sondern wird häufig auch als *der* Minimum-Energie-Weg bezeichnet. Den Grund hierfür gibt Fukui in [13] wie folgt an: Da die IRC senkrecht zur Tangentialebene der Äquipotentialfläche steht, erhalten wir in den Schnittpunkten der IRC mit der Tangentialebene stets minimale Werte für E . Dies rechtfertigt die Bezeichnung der IRC als „wahrer“ MEP. Da jedoch durch sämtliche Punkte der PES eine Gradientenabstiegskurve führt, wäre jede dieser Kurven in der Umgebung des Minimums ein MEP. Ihnen gegenüber besitzt die IRC lediglich die Besonderheit, ihren Startpunkt im Sattelpunkt vom Index 1 zu haben [19]. Von dieser Besonderheit abgesehen hat die IRC keine lokalen Eigenschaften, die sie von anderen SD-Kurven unterscheidet [23]. Mithin können wir für einen Punkt $\mathbf{x} \in K$ nicht lokal feststellen, ob er zur IRC gehört, ohne diese zu durchlaufen. Zudem können SD-Kurven (also insbesondere auch die IRC) sowohl durch Täler als auch über Talrücken verlaufen. Dies widerspricht der Vorstellung des MEP als Weg entlang eines Talgebiets. Der große Vorteil der IRC ist ihr geringer Rechenaufwand, da wir lediglich den Gradienten der PES benötigen.

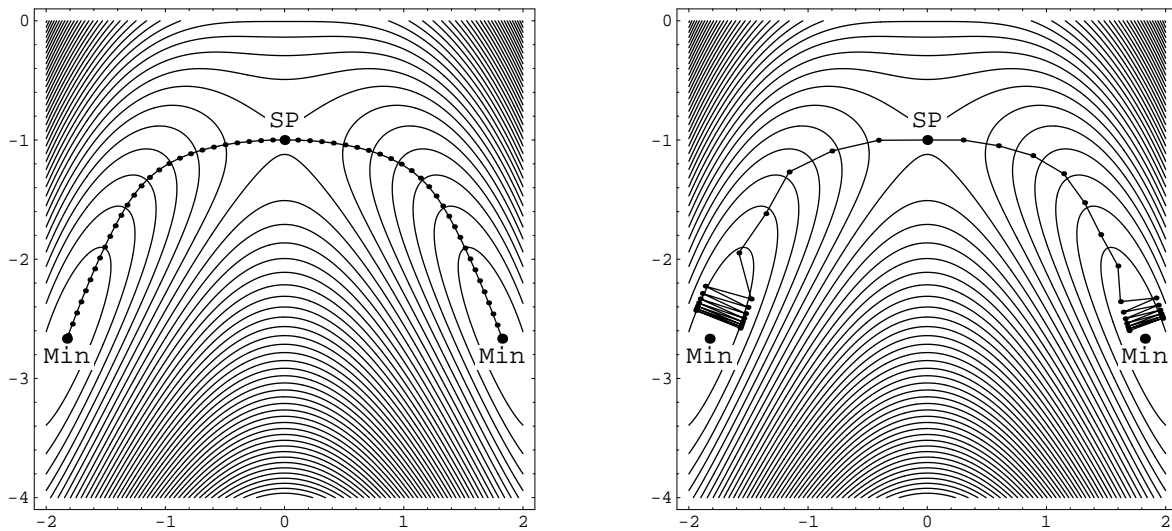


Abbildung 2.1: Die theoretische IRC (links) im Vergleich mit der numerischen IRC (rechts), die in Minimumnähe in Zickzack-Bewegungen verfällt.

Betrachten wir nun weitere Nachteile des IRC-Konzepts. Wie bereits erwähnt, ist die IRC eindeutig bestimmt. Sie verzweigt sich nicht und liefert damit keine Lösungsansätze für das Problem der RP-Bifurkation. Zudem leidet das Verfahren unter einem numerischen Dilemma: Die SD-Gleichung (2.4) ist eine steife Differentialgleichung. Löst man sie explizit wie der oben angegebene Algorithmus und wählt die Schrittweite α nicht klein genug in Bezug auf die gewünschte Fehlertoleranz ε , so bewegt sich die numerische Kurve in der

Nähe des Minimums im Zickzack-Kurs um die tatsächliche IRC. Wir wollen bemerken, daß dies kein Problem einer etwaig falschen Implementation ist, sondern direkt aus (2.4) folgt und sich damit nicht vermeiden läßt. Abbildung 2.1 veranschaulicht die Zickzack-Bewegungen. Ein weiterer Nachteil der IRC ist ihr Start im Sattelpunkt. Wir können die IRC aus numerischen Gründen nicht „aufwärts“ verfolgen, das heißt vom Minimum aus einen Schritt in Richtung des zum kleinsten Eigenwert der Hessematrix gehörenden Eigenvektors gehen und anschließend Schritt (ii) des Algorithmus mit $\mathbf{g}(\mathbf{x}^i)$ statt $-\mathbf{g}(\mathbf{x}^i)$ ausführen. Die so ermittelte Kurve bräche nach ihrem Start im Minimum seitlich vom gewünschten Aufstiegsfad weg und landete nicht im Sattelpunkt.

2.1.2. Die String-Methode und ihre Anwendung auf die IRC

Die in dieser Arbeit vorgestellten MEP-Ansätze (IRC, GE, RGF, etc.) lassen sich mithilfe verschiedener numerischer Methoden realisieren. Neben der in Abschnitt 2.1.1 behandelten punktweisen Ermittlung des Minimum-Energie-Wegs (in diesem Fall der IRC) wollen wir alternativ die String-Methode verwenden. Bei dieser wird ein gewisser Startpfad solange variiert bis er die Anforderungen des gewählten MEP-Ansatzes erfüllt. Zunächst aber wollen wir den Begriff des Pfads zwischen den Punkten \mathbf{x}^0 und \mathbf{x}^k des Konfigurationsraums erklären.

Definition

Sei $C : (a, b) \rightarrow K$ glatt und gelte, daß $C(a) = \mathbf{x}^0$ und $C(b) = \mathbf{x}^k$. Wir betrachten eine Zerlegung des Intervalls gemäß $a = \xi_0 < \xi_1 < \dots < \xi_{k-1} < \xi_k = b$ und setzen $\mathbf{x}^i := C(\xi_i)$ für $i = 1, \dots, k-1$. Dann definiert $\bigcup_{i=0}^{k-1} \overline{\mathbf{x}^i \mathbf{x}^{i+1}}$ einen *Pfad* im Konfigurationsraum der PES zwischen \mathbf{x}^0 und \mathbf{x}^k .

Betrachten wir nun insbesondere Reaktionspfade, so fordern wir, daß die \mathbf{x}^i den Bedingungen des gewählten MEP-Ansatzes (IRC, GE, RGF, etc.) genügen. Wir interpretieren dann den Reaktionspfad gemäß der Definition als Vereinigung der Verbindungsstrecken zwischen benachbarten Punkten der endlichen Folge $\{\mathbf{x}^i\}_{0 \leq i \leq k}$. Damit wird der Minimum-Energie-Weg durch $\{\mathbf{x}^i\}_{0 \leq i \leq k}$ repräsentiert.

Nach der in der Literatur gebräuchlichsten (und auch in Abschnitt 2.1.1 gewählten) Herangehensweise fixiert man einen Startpunkt $\mathbf{x}^0 \in K$. Üblicherweise ist \mathbf{x}^0 stationärer Punkt. Anschließend ermittelt man die weiteren Folgepunkte iterativ. Die genaue Abbildungsvorschrift für den Schritt $\mathbf{x}^i \mapsto \mathbf{x}^{i+1}$ hängt vom betrachteten MEP-Ansatz ab. Dank der Vorgabe von Schrittlänge $\alpha > 0$ und Fehlertoleranz $\varepsilon > 0$ endet die Iteration nach $k < \infty$ Schritten.

2. Wegeverfolgung auf Potentialenergieflächen

Eine Alternative stellt die in [37] vorgestellte String-Methode dar. Als *String* oder *Kette* bezeichnen wir einen Pfad in K , der durch die Folge $\{\mathbf{x}^i\}_{0 \leq i \leq k}$ repräsentiert wird. Typischerweise sind \mathbf{x}^0 und \mathbf{x}^k stationäre Punkte der PES. Projizieren wir die \mathbf{x}^i gemäß einer von der Wahl der RP-Methode abhängigen Vorschrift in Punkte $\tilde{\mathbf{x}}^i$, so erhalten wir einen neuen Pfad im Konfigurationsraum der PES, repräsentiert durch $\{\tilde{\mathbf{x}}^i\}_{0 \leq i \leq k}$. Mit vorgegebener Schrittweite $\alpha > 0$ und Fehlertoleranz $\varepsilon > 0$ führt endliche Wiederholung dieses Vorgehens zum gewünschten Reaktionspfad. An dieser Stelle sei der hierbei verwendete Orthogonalitätsprojektor definiert:

Definition

Für den normierten (Spalten-) Vektor $\mathbf{r} \in \mathbb{R}^n$ definiert

$$\mathbf{P}_{\mathbf{r}} = \mathbf{I}_n - \begin{pmatrix} r_1^2 & \cdots & r_1 r_n \\ \vdots & \ddots & \vdots \\ r_n r_1 & \cdots & r_n^2 \end{pmatrix} = \mathbf{I}_n - \mathbf{r} \mathbf{r}^T \quad (2.5)$$

den Orthogonalitätsprojektor, also den Projektor, der senkrecht zu \mathbf{r} abbildet. Dabei bezeichne \mathbf{I}_n die $(n \times n)$ -Einheitsmatrix.

Bemerkung

Wenden wir den Orthogonalitätsprojektor $\mathbf{P}_{\mathbf{r}}$ auf einen zu \mathbf{r} parallelen Vektor $\mu \mathbf{r}$ an, so erhalten wir

$$\mathbf{P}_{\mathbf{r}}(\mu \mathbf{r}) = (\mathbf{I}_n - \mathbf{r} \mathbf{r}^T) (\mu \mathbf{r}) = \mu [\mathbf{r} - \mathbf{r} (\mathbf{r}^T \mathbf{r})] = \mu (\mathbf{r} - \mathbf{r}) = \mathbf{0} . \quad (2.6)$$

Senkrecht auf \mathbf{r} stehende Vektoren werden durch $\mathbf{P}_{\mathbf{r}}$ nicht verändert.

Hat man die Projektionsvorschriften einmal ermittelt, so ist die String-Methode gerade bei komplizierteren Verfahren wie GE oder RGF deutlich leichter zu implementieren. Sie birgt jedoch Probleme, wie wir im weiteren Verlauf der Arbeit noch erfahren werden: Geht man von einer Startkette $\{\mathbf{x}^i\}_{0 \leq i \leq k}$ aus, bei der benachbarte Punkte den gleichen Abstand haben, so kann es unter der Projektion passieren, daß sich mit jeder Wiederholung mehr Punkte in Richtung eines Minimums zusammenziehen und somit Lücken⁶ in die Kette reißen. Abhilfe verschafft etwa Spline-Interpolation, mithin zusätzlicher numerischer Aufwand. Ein weiteres Problem der String-Methode ist die Tendenz unerwünschte Zickzack-Bewegungen auszuüben, nicht nur bei der IRC-Bestimmung, sondern auch beim GE-Verfahren [37].

⁶Der Begriff „Lücke“ soll an dieser Stelle anschaulich einen Bereich der Kette beschreiben, in dem der Abstand zwischen benachbarten Punkten sehr groß wird in Bezug auf den durchschnittlichen Abstand zwischen Nachbarpunkten.

Abschließend wollen wir die IRC mithilfe der String-Methode bestimmen. Wir wählen eine glatte Ausgangskurve $\mathbf{x}(s)$ auf der PES, die die beiden Minima verbindet. Numerisch betrachten wir $\mathbf{x}(s)$ als Pfad, der durch die Kette $\{\mathbf{x}^i\}_{0 \leq i \leq k}$ repräsentiert wird. Das heißt, \mathbf{x}^0 und \mathbf{x}^k sind gerade die beiden Minimalstellen. Typischerweise wählen wir die übrigen Kettenpunkte so, daß sie die Strecke zwischen \mathbf{x}^0 und \mathbf{x}^k äquidistant zerlegen. Ziel ist es, die Punkte \mathbf{x}^i derart zu variieren, daß die Kette einerseits der SD-Bedingung (2.4) genügt und andererseits durch den Sattelpunkt verläuft. Eine Kette, die diese beiden Eigenschaften erfüllt, ist eindeutig bestimmt, wie wir in Abschnitt 4.2.1 zeigen werden.

Nehmen wir zunächst an, daß es sich bei unserer Startkurve bereits um die IRC handelt. Bezeichnet $\mathbf{t}(s)$ die Tangente von $\mathbf{x}(s)$, so gilt nach (2.4) gerade $\mathbf{t}(s) = -\mathbf{w}(s)$. Wie wir aus den bisherigen Überlegungen bereits wissen, folgt die IRC dem negativen Gradienten. Mithin ist $\mathbf{t}(s)$ parallel zum Gradienten und orthogonal zur Äquipotentialfläche. Demnach gilt wegen (2.6):

$$\mathbf{P}_{\mathbf{t}}(-\mathbf{g}) = (\mathbf{I}_n - \mathbf{t}\mathbf{t}^T) (-\mathbf{g}) = -(\mathbf{I}_n - \mathbf{w}\mathbf{w}^T) (\|\mathbf{g}\|_2 \cdot \mathbf{w}) = \mathbf{0}. \quad (2.7)$$

Die Bedingung (2.7) stellt gewissermaßen das Abbruchkriterium der String-Methode dar: Liegen sämtliche Kettenpunkte $\{\mathbf{x}^i\}_{0 \leq i \leq k}$ auf der IRC, so verschwindet der auf den negativen Gradienten angewandte Projektionsoperator; die \mathbf{x}^i werden auf sich selbst abgebildet.

Betrachten wir nun den Fall, daß es sich bei $\mathbf{x}(s)$ noch nicht um die IRC handelt. Wir fordern insbesondere, daß wenigstens einer der Kettenpunkte $\mathbf{x}^0, \dots, \mathbf{x}^k$ nicht auf der IRC liegt. Dann ist

$$\mathbf{p}(\mathbf{x}^i) := \mathbf{P}_{\mathbf{t}(\mathbf{x}^i)}(-\mathbf{g}(\mathbf{x}^i)) = -(\mathbf{I}_n - \mathbf{t}(\mathbf{x}^i)\mathbf{t}(\mathbf{x}^i)^T) \mathbf{g}(\mathbf{x}^i)$$

für mindestens ein $i \in \{0, \dots, k\}$ verschieden vom Nullvektor. Die iterierten Kettenpunkte ergeben sich aus

$$\tilde{\mathbf{x}}^i = \mathbf{x}^i - \alpha \cdot \mathbf{p}(\mathbf{x}^i),$$

wobei $\alpha > 0$ wie üblich die Schrittweite bezeichnet. Wir variieren also die Kettenpunkte $\{\mathbf{x}^i\}_{0 \leq i \leq k}$ entlang der Vektoren $-\mathbf{p}(\mathbf{x}^i)$ und erhalten nach einer endlichen Anzahl von Iterationen die IRC. Wegen (2.7) wäre ein Test bezüglich der Summe der Abstände, etwa in der Form

$$\varepsilon > \sum_{i=0}^k \|\tilde{\mathbf{x}}^i - \mathbf{x}^i\|_2$$

2. Wegeverfolgung auf Potentialenergieflächen

für ein vorgegebenes $\varepsilon > 0$, ein typisches numerisches Abbruchkriterium. Aufgrund der Anfälligkeit der String-Methode in Zickzack-Bewegungen zu verfallen, ist hierbei allerdings Vorsicht geboten. Alternativ ließe sich die Veränderung am energetisch höchsten Punkt der Kette, also am mutmaßlichen Sattelpunkt, testen:

$$\varepsilon > \max_{0 \leq i \leq k} E(\tilde{\mathbf{x}}^i) - \max_{0 \leq i \leq k} E(\mathbf{x}^i) .$$

Bemerkung

Im Gegensatz zur „klassischen“ Berechnung der IRC mithilfe des in Abschnitt 2.1.1 vorgestellten Algorithmus benötigen wir bei Verwendung der String-Methode keine Kenntnis des Sattelpunktes.

Beispiel

In Abbildung 2.2 ist (ebenso wie in Abbildung 2.1) ein vereinfachtes zweidimensionales Modell für die PES von Malonaldehyd dargestellt. Sie ist gemäß (A.4) definiert. Wir sehen links die gesuchte IRC nach 53 Iterationen bei einer Schrittweite von $\alpha = 0.1$ sowie die gerade Ausgangskette der Länge $k = 30$. Als Abbruchkriterium diente ein Test nach dem Energiemaximum der Kette, dem Sattelpunkt. Die rechte Abbildung zeigt die Kette nach fünf Iterationen.

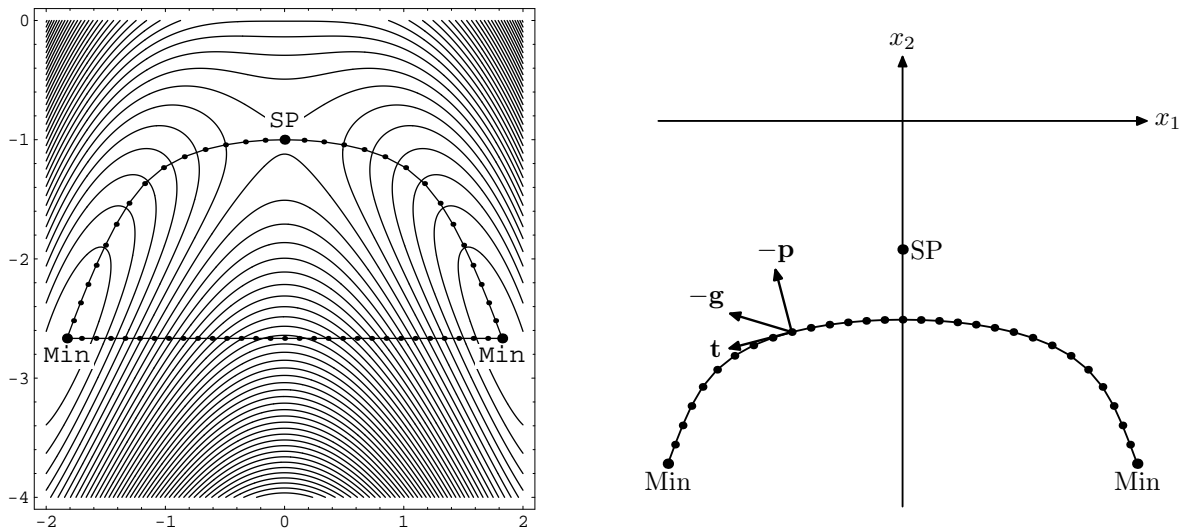


Abbildung 2.2: Ausgangskette und IRC auf der PES (A.4) (links) und die Kette nach fünf Iterationen (rechts).

2.1.3. Krümmung von Kurven des steilsten Abstiegs

Wir möchten diesen Abschnitt mit der Bestimmung des Krümmungsvektors \mathbf{k} von Kurven des steilsten Abstiegs (SD-Kurven) abschließen und die in [47] angegebene Formel für \mathbf{k} nachrechnen. Mithilfe der Krümmung werden wir einen Zusammenhang zu den im Anschluß vorgestellten Gradient Extremals herstellen. Wir betrachten eine in Bogenlänge parametrisierte SD-Kurve $\mathbf{x}(s)$. Dann ist \mathbf{k} wie folgt definiert:

$$\begin{aligned} \mathbf{k} &= \frac{d\bar{\mathbf{t}}}{ds} \stackrel{(2.4)}{=} \frac{d^2\mathbf{x}}{ds^2} = \frac{d}{ds} \left(-\frac{\mathbf{g}(\mathbf{x})}{\|\mathbf{g}(\mathbf{x})\|_2} \right) = - \left(\frac{d\mathbf{x}}{ds} \right) \cdot \nabla \left(\frac{\mathbf{g}(\mathbf{x})}{\|\mathbf{g}(\mathbf{x})\|_2} \right) \\ &= \frac{\mathbf{g}(\mathbf{x})}{\|\mathbf{g}(\mathbf{x})\|_2} \cdot \nabla \left(\frac{\mathbf{g}(\mathbf{x})}{\|\mathbf{g}(\mathbf{x})\|_2} \right). \end{aligned}$$

Damit gilt für die i -te Komponente des Krümmungsvektors:

$$\begin{aligned} k_i &= \frac{1}{\|\mathbf{g}\|_2} \sum_{j=1}^n g_j \cdot \frac{\partial}{\partial x_j} \left(\frac{g_i}{\|\mathbf{g}\|_2} \right) = \frac{1}{\|\mathbf{g}\|_2} \left(\sum_{j=1}^n g_j \cdot \frac{H_{ij}}{\|\mathbf{g}\|_2} - g_i \sum_{j=1}^n g_j \frac{\frac{\partial}{\partial x_j} \sqrt{\sum_{k=1}^n g_k^2}}{\|\mathbf{g}\|_2^2} \right) \\ &= \frac{1}{\|\mathbf{g}\|_2} \left(\sum_{j=1}^n H_{ij} w_j - w_i \sum_{j=1}^n w_j \frac{\sum_{k=1}^n H_{jk} g_k}{\|\mathbf{g}\|_2} \right) = \frac{(\mathbf{H}\mathbf{w})_i - w_i (\mathbf{w}^T \mathbf{H}\mathbf{w})}{\|\mathbf{g}\|_2}. \end{aligned}$$

Daraus folgt schließlich für den Krümmungsvektor:

$$\mathbf{k} = \frac{\mathbf{H}\mathbf{w} - \mathbf{w}\mathbf{w}^T \mathbf{H}\mathbf{w}}{\|\mathbf{g}\|_2} = \frac{(\mathbf{I}_n - \mathbf{w}\mathbf{w}^T)(\mathbf{H}\mathbf{w})}{\|\mathbf{g}\|_2} = \frac{\mathbf{P}_{\mathbf{w}}(\mathbf{H}\mathbf{w})}{\|\mathbf{g}\|_2}.$$

Durch die Anwendung des Projektors $\mathbf{P}_{\mathbf{w}}$ steht der Krümmungsvektor senkrecht zu \mathbf{w} und damit senkrecht zum Gradienten. Als Krümmung errechnen wir den Wert

$$\begin{aligned} k &= \sqrt{\mathbf{k}^T \mathbf{k}} = \frac{1}{\|\mathbf{g}\|_2} \left[(\mathbf{H}\mathbf{w} - \mathbf{w}(\mathbf{w}^T \mathbf{H}\mathbf{w}))^T (\mathbf{H}\mathbf{w} - \mathbf{w}(\mathbf{w}^T \mathbf{H}\mathbf{w})) \right]^{1/2} \\ &= \frac{1}{\|\mathbf{g}\|} \left[(\mathbf{w}^T \mathbf{H}) \mathbf{H}\mathbf{w} - (\mathbf{w}^T \mathbf{H}) \mathbf{w}\mathbf{w}^T \mathbf{H}\mathbf{w} - \underbrace{(\mathbf{w}^T \mathbf{H}\mathbf{w})^2 + \mathbf{w}^T \mathbf{w} (\mathbf{w}^T \mathbf{H}\mathbf{w})^2}_{=0} \right]^{1/2} \\ &= \frac{[\mathbf{w}^T \mathbf{H} \mathbf{P}_{\mathbf{w}}(\mathbf{H}\mathbf{w})]^{1/2}}{\|\mathbf{g}\|_2}. \end{aligned} \tag{2.8}$$

2.2. Das Gradient-Extremal-Verfahren

Mithilfe der IRC-Methode können wir vom Sattelpunkt ausgehend Minima finden. Aus praktischer Sicht ist allerdings die umgekehrte Herangehensweise interessanter. Das heißt, wir suchen Reaktionspfade, die im Minimum (Reaktant) starten und, einem Tal aufwärts

2. Wegeverfolgung auf Potentialenergieflächen

folgend, schließlich im Sattelpunkt (Übergangsstruktur) enden. Eine naheliegende Wahl sind die sogenannten Talgrundpfade. Das heißt, wir benötigen dann ein Kriterium, welches das Verbleiben des Pfads im Talgrund sicherstellt. Ein solches Kriterium sollte lokal sein, um punktweise feststellen zu können, ob sich der RP im Talgrund befindet. Auf diese Weise ließe sich der Reaktionspfad sowohl vom Minimum in Richtung Sattelpunkt verfolgen als auch umgekehrt [3].

Pančír lieferte für dieses Problem einen ersten Lösungsansatz. Seiner Ansicht nach müßten „offensichtlich“ die folgenden Bedingungen gelten [33]:

- (i) Die Energie wächst entlang jeder Richtung, die senkrecht zur Richtung des Talgrunds stehen.
- (ii) Die Krümmung der PES entlang der Richtung des Tales ist minimal bezüglich der Krümmung entlang einer beliebigen anderen Richtung.

Pančír behauptete, daß dann in Punkten, die den Bedingungen (i) und (ii) genügen, der Gradient ein Eigenvektor der Hessematrix sein müsse. Diese Behauptung wird durch den „Bergsteiger-Algorithmus“⁷ von Basilevsky und Shamov [3] sowie durch Hoffman, Nord und Ruedenberg [23] belegt. Auch wir werden Pančírs Vermutung im Folgenden nachweisen.

2.2.1. Pfade im Talgrund

Der Talgrundpfad ist ein zentraler Gegenstand dieses Abschnitts. Daher wollen wir zunächst eine Definition gemäß [23] angeben, um anschließend den Begriff des *Gradient Extremal* erklären zu können.

Definition

Für den Punkt $\mathbf{x}^i \in K$ bezeichne $M_{c_i} = \{\mathbf{x} \in K : E(\mathbf{x}) = c_i\}$ die zugehörige Äquipotentialfläche. Es gilt also $\mathbf{x}^i \in M_{c_i}$. Erfüllt ein Pfad $\{\mathbf{x}^i\}_{0 \leq i \leq k}$ die Bedingung

$$\|\mathbf{g}(\mathbf{x}^i)\|_2 = \min_{\mathbf{x} \in M_{c_i}} \|\mathbf{g}(\mathbf{x})\|_2$$

für alle $i = 0, \dots, k$, so nennen wir ihn *Talgrundpfad*.

Da der Gradient senkrecht zur Äquipotentialfläche steht, schneidet der Talgrundpfad die Fläche M_{c_i} gerade in dem Punkt \mathbf{x}^i , von dem aus ein Weitergehen senkrecht zu M_{c_i} den geringsten Höhenunterschied bedeutet, mithin dort, wo das Tal seinen geringsten Anstieg hat.

⁷Basilevsky und Shamov nannten ihr Verfahren “mountaineer’s algorithm“, da es durch das Verfolgen des Talgrundpfads bis zum Sattelpunkt das Verhalten eines Bergsteigers imitiert.

Wir haben damit die Aufgabe, Punkte \mathbf{x}^i zu bestimmen, für die der Betrag des Gradienten unter Variation entlang der zugehörigen Äquipotentialfläche minimal wird. Um die Rechnung zu vereinfachen, folgen wir der Notation in [45] und betrachten das äquivalente Problem:

$$\sigma(\mathbf{x}) := \frac{\|\mathbf{g}(\mathbf{x})\|_2^2}{2} \longrightarrow \underset{\mathbf{x} \in M_{c_i}}{\text{Min!}} .$$

Dann gilt:

$$\nabla \sigma = \nabla \left(\frac{\|\mathbf{g}\|_2^2}{2} \right) = \nabla \left(\frac{\mathbf{g}^T \mathbf{g}}{2} \right) = \nabla \left(\frac{(\nabla E)^T (\nabla E)}{2} \right) = (\nabla \nabla^T E) (\nabla E) = \mathbf{H} \mathbf{g} .$$

Eine notwendige Voraussetzung für Extremstellen \mathbf{x}^i (und Minimalstellen insbesondere) ist das Verschwinden der Orthogonalprojektion von $\nabla \sigma(\mathbf{x}^i)$ entlang des normierten Gradienten \mathbf{w} [23]. In den Extrema \mathbf{x}^i gilt demnach

$$\begin{aligned} \mathbf{0} &= \mathbf{P}_{\mathbf{w}}(\nabla \sigma) = \mathbf{P}_{\mathbf{g}/\|\mathbf{g}\|_2}(\mathbf{H} \mathbf{g}) = \left(\mathbf{I}_n - \frac{\mathbf{g} \mathbf{g}^T}{\|\mathbf{g}\|_2^2} \right) (\mathbf{H} \mathbf{g}) = \mathbf{H} \mathbf{g} - \frac{(\mathbf{g} \mathbf{g}^T) (\mathbf{H} \mathbf{g})}{\|\mathbf{g}\|_2^2} \\ &= \mathbf{H} \mathbf{g} - \lambda \mathbf{g} \end{aligned} \quad (2.9)$$

für $\lambda := \mathbf{g}^T \mathbf{H} \mathbf{g} / \|\mathbf{g}\|_2^2$. Damit ist Pančír's Behauptung gezeigt: In Punkten \mathbf{x}^i , für die $\|\mathbf{g}(\mathbf{x}^i)\|_2$ extremal (und insbesondere minimal) wird, ist $\mathbf{g}(\mathbf{x}^i)$ Eigenvektor von $\mathbf{H}(\mathbf{x}^i)$. Wir definieren somit nach [23]:

Definition

Punkte, deren Gradienten Eigenvektoren der Hessematrix sind, liegen auf einem *Gradient Extremal*(GE).

Bemerkungen

- (i) Gemäß der Herleitung obiger Definition können wir unter Verwendung des Funktionals σ auch schreiben: Ein Punkt $\mathbf{x} \in K$ liegt auf einer GE-Kurve, falls er das Kriterium

$$\mathbf{p}(\mathbf{x}) := \mathbf{P}_{\mathbf{w}(\mathbf{x})}(\mathbf{H}(\mathbf{x})\mathbf{g}(\mathbf{x})) = \mathbf{0} \quad (2.10)$$

erfüllt.

- (ii) Unter Berücksichtigung von (2.10) ist klar, daß sämtliche stationären Punkte $\mathbf{x}^* \in K$ auf GE-Kurven liegen.
- (iii) Die Eigenwertzerlegung der (symmetrischen) Hessematrix \mathbf{H} liefert die Diagonalmatrix $\mathbf{\Lambda} = \text{diag}(\lambda_1, \dots, \lambda_n) = \mathbf{U}^T \mathbf{H} \mathbf{U}$. Bezeichnet $[\mathbf{u}^j]^T := \{u_j\}$ die j -te Zeile von \mathbf{U} ,

2. Wegeverfolgung auf Potentialenergieflächen

so ist \mathbf{u}^j Eigenvektor von \mathbf{H} zum Eigenwert $\lambda := \lambda_j$. Dann fordert das GE-Kriterium (2.10), daß \mathbf{g} wegen (2.6) parallel zu \mathbf{u}^j ist. Dies wiederum gilt genau dann, wenn \mathbf{g} senkrecht auf allen anderen Eigenvektoren von \mathbf{H} steht:

$$\mathbf{g}^T \mathbf{u}^i = 0 \quad (2.11)$$

für alle $i \in \{1, \dots, n\} \setminus \{j\}$. Aus (2.11) erhalten wir $(n - 1)$ linear unabhängige Gleichungen, die uns $(n - 1)$ zugehörige GE-Kurven definieren.

Wir wollen uns ein wenig genauer mit den in Bemerkung (iii) erwähnten GE-Gleichungen befassen. Sei o.B.d.A. $j = 1$. Dann gilt (2.11) für $i = 2, \dots, n$. Aus der Extremalbedingung $\mathbf{P}_{\mathbf{w}}(\nabla\sigma) = 0$ folgt unmittelbar wegen (2.6), daß $\nabla\sigma$ parallel zum (normierten) Gradienten und damit auch parallel zu \mathbf{u}^1 liegt. Mit anderen Worten: $\nabla\sigma$ steht senkrecht auf allen anderen Eigenvektoren von \mathbf{H} . Auf diese Weise wollen wir die $(n - 1)$ GE-Gleichungen definieren:

$$\text{GE}_{i-1}(\mathbf{x}) := \nabla\sigma(\mathbf{x})^T \cdot \mathbf{u}^i(\mathbf{x}) = 0$$

für $i = 2, \dots, n$. Zusammengefaßt in Vektordarstellung ergibt sich daraus $\mathbf{GE} : \mathbb{R}^n \rightarrow \mathbb{R}^{n-1}$ mit $\mathbf{x} \mapsto \mathbf{P}_{\mathbf{w}(\mathbf{x})}(\nabla\sigma(\mathbf{x}))$. GE-Kurven erhalten wir aus der Bedingung $\mathbf{GE}(\mathbf{x}) = \mathbf{0}$.

Beispiel

Wir betrachten den Fall $n = 2$. Dann ist

$$\lambda = \frac{\mathbf{g}^T \mathbf{H} \mathbf{g}}{\|\mathbf{g}\|_2^2} = \frac{H_{11}g_1^2 + 2H_{12}g_1g_2 + H_{22}g_2^2}{g_1^2 + g_2^2}$$

und wir erhalten in (2.9):

$$\mathbf{0} = (g_1^2 + g_2^2) \begin{pmatrix} H_{11} & H_{12} \\ H_{12} & H_{22} \end{pmatrix} \begin{pmatrix} g_1 \\ g_2 \end{pmatrix} - (H_{11}g_1^2 + 2H_{12}g_1g_2 + H_{22}g_2^2) \begin{pmatrix} g_1 \\ g_2 \end{pmatrix} .$$

Durch Rechnung ergibt sich (für beide Komponenten) die GE-Gleichung

$$0 = H_{12}(g_2^2 - g_1^2) + g_1g_2(H_{11} - H_{22}) . \quad (2.12)$$

Entgegen der Vorstellung, daß es sich bei Gradient Extremals stets um Kurven handelt, sei die Modell-Potentialenergiefläche $E(x_1, x_2) = x_1^2 + x_2^2$ betrachtet [18]. Wegen $H_{12} = 0$ und $H_{11} = H_{22} = 2$ ist die Gleichung (2.12) für alle (x_1, x_2) der PES erfüllt.

Kehren wir nun zurück zu den eingangs betrachteten Talgrundpfaden. Offensichtlich werden diese durch GE-Kurven angezeigt, da für Punkte des Talgrunds σ minimal auf

der zugehörigen Äquipotentialfläche ist. Allerdings gilt die Umkehrung nicht: Ist $\sigma(\mathbf{x}^i)$ ein Minimum auf M_{c_i} , so kann \mathbf{x}^i ebenso auf der Kammlinie eines Talrückens liegen. Da sowohl Minima als auch Maxima und Stufenpunkte zu Gradient Extremals führen, werden durch GE-Kurven neben Talgründen auch andere Bereiche der PES indiziert. In [45] ist die Klassifikation einer zweidimensionalen PES mittels GE angegeben.

Nimmt allerdings σ einen Minimalwert an, so erhalten wir einen interessanten Zusammenhang zwischen SD-Kurven und Gradient Extremals. Nach (2.8) galt für die Krümmung von Kurven des steilsten Abstiegs der Wert $k = [\mathbf{w}^T \mathbf{H} \mathbf{P}_{\mathbf{w}}(\mathbf{H}\mathbf{w})]^{1/2} / \|\mathbf{g}\|_2$. Daraus folgt

$$k = 0 \iff \mathbf{w}^T \mathbf{H} \mathbf{P}_{\mathbf{w}}(\mathbf{H}\mathbf{w}) = 0 \iff \mathbf{P}_{\mathbf{w}}(\mathbf{H}\mathbf{w}) = \mathbf{0} \iff \mathbf{P}_{\mathbf{w}}(\mathbf{H}\mathbf{g}) = \mathbf{0} .$$

Nach (2.10) bestehen also Gradient Extremals gerade aus den Punkten, in denen SD-Kurven die Krümmung Null haben.

Mithilfe dieses Zusammenhangs erhalten wir folgende Eigenschaft: Im Talgrund verlaufen SD-Kurven, von links und rechts kommend, in einem Strom mit dem Gradient Extremal. Wir nennen das Verhalten der SD-Kurven im Talgrund *konfluent* [47] und bezeichnen den Talgrund selbst auch als „Flußbett“ (*Streambed*).

Beispiel

Abbildung 2.3(a) zeigt GE-Kurven (durchgehend) und einige konfluente Kurven des steilsten Abstiegs (gepunktet) auf der Potentialenergiefläche (A.1).

2.2.2. Numerische Verfolgung von GE-Kurven

Wie bei allen anderen Methoden zur Ermittlung des Minimum-Energie-Wegs gibt es auch beim GE-Verfahren unterschiedliche numerische Lösungsansätze. Wir wollen zunächst den schon erwähnten „Bergsteiger-Algorithmus“ zum Aufspüren des Talgrundpfads nach Basilevsky und Shamov [3] vorstellen, anschließend eine Predictor-Corrector-Methode nach Sun und Ruedenberg [47] und drittens die in Absatz 2.1.2 eingeführte RP-Bestimmung mithilfe von Ketten.

Der „Bergsteiger-Algorithmus“ [3] startet im Minimum und führt anschließend durch den Talgrund zum Sattelpunkt. Er liefert den Talgrundpfad und trägt somit lediglich teilweise zur Verfolgung von Gradient Extremals bei. Die Schritte des Algorithmus:

- (i) Starte im Minimum \mathbf{x}^{\min} mit einem (infinitesimalen) Schritt nach $\tilde{\mathbf{x}}^0$ in Richtung des Eigenvektors von $\mathbf{H}(\mathbf{x}^{\min})$ mit dem kleinsten Eigenwert.
- (ii) Der aktuelle Standort sei $\tilde{\mathbf{x}}^i$. Suche entlang der zu $\tilde{\mathbf{x}}^i$ gehörigen Äquipotentialfläche den Punkt \mathbf{x}^i , in dem σ minimal ist.

2. Wegeverfolgung auf Potentialenergieflächen

(iii) Gehe einen Schritt der Länge α aufwärts in Richtung des Gradienten:

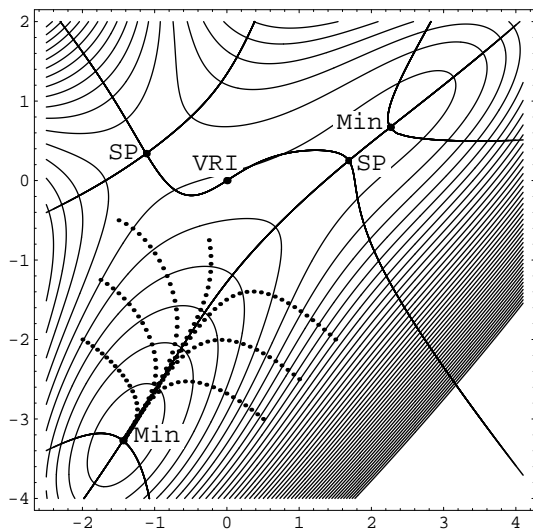
$$\tilde{\mathbf{x}}^{i+1} := \mathbf{x}^i + \alpha \cdot \mathbf{g}(\mathbf{x}^i) .$$

(iv) Wiederhole (ii) und (iii) bis der Sattelpunkt in gewünschter Genauigkeit ε erreicht ist.

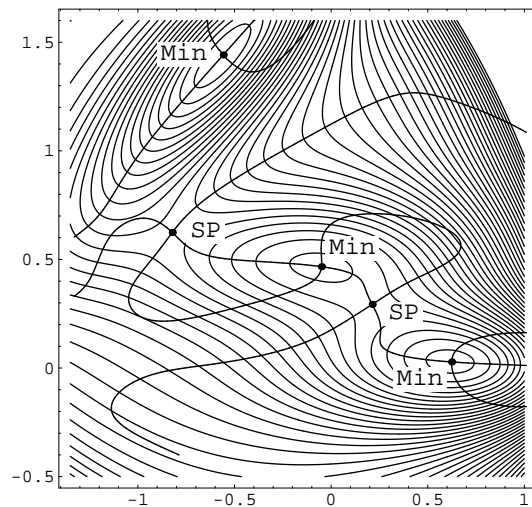
Auf diese Weise erhalten wir einen Pfad $\{\mathbf{x}^i\}_{0 \leq i \leq k}$, der die durch den Talgrund verlaufende GE-Kurve beschreibt. Häufig tritt jedoch der Fall ein, daß Täler einfach irgendwo auf der PES enden, ohne einen Sattelpunkt zu erreichen. Als Beispiel hierfür betrachten wir das Müller-Brown-Potential.

Beispiel

Abbildung 2.3(b) zeigt einige Gradient Extremals auf dem Müller-Brown-Potential. Während die beiden südlichen Minima zusammen mit den Sattelpunkten auf einer GE-Kurve liegen, gibt es keinen GE-Pfad, der das Minimum im Norden mit dem westlichen Sattelpunkt verbindet.



(a) Gradient Extremals (durchgehend) und konfluente SD-Kurven (gepunktet) auf der PES (A.1)



(b) Einige GE-Kurven auf dem Müller-Brown-Potential

Abbildung 2.3: Gradient Extremals und konfluente SD-Kurven

Kommen wir nun zur Predictor-Corrector-Methode nach Sun und Ruedenberg [47]. Mit ihr können wir GE-Kurven auch außerhalb des Talgrunds verfolgen. Wir wissen bereits, daß – im Gegensatz zu den Kurven des steilsten Abstiegs – nicht durch jeden Punkt $\mathbf{x} \in K$ ein GE verläuft, sondern \mathbf{x} der Bedingung (2.10) genügen muß. Wir müssen daher bei der

Iteration $\mathbf{x}^i \mapsto \mathbf{x}^{i+1}$ aufpassen, das GE nicht zu verlassen. Der Algorithmus erfolgt in zwei Schritten: Ausgehend von einem Punkt \mathbf{x}^i auf dem GE-Pfad bestimmen wir zuerst einen Übergangspunkt $\tilde{\mathbf{x}}^i$, der in der Regel abseits des GE liegt (Predictor-Schritt) und korrigieren diesen anschließend zu \mathbf{x}^{i+1} , welcher sich wiederum auf dem GE-Pfad befindet (Corrector-Schritt). Bevor wir aber beide Schritte erläutern, muß die Tangente an das Gradient Extremal bestimmt werden.

Nach (2.10) sind Gradient Extremals definiert gemäß $\mathbf{p} := \mathbf{P}_w(\mathbf{H}\mathbf{g}) = \mathbf{0}$. Damit ist die Tangente an das GE im Punkt \mathbf{x} gegeben durch

$$d\mathbf{p} = \sum_{i=1}^n \left(\frac{\partial \mathbf{p}}{\partial x_i} \right) dx_i =: l \cdot \sum_{i=1}^n \left(\frac{\partial \mathbf{p}}{\partial x_i} \right) \bar{t}_i = \mathbf{0}, \quad (2.13)$$

wobei wir $l\bar{\mathbf{t}} := d\mathbf{x}$ setzen und mit $\bar{\mathbf{t}}$ der normierte Tangentenvektor bezeichnet sei. Die ersten Ableitungen von \mathbf{p} sind im Anhang B.2 angegeben. Wir wählen die lokalen Koordinaten x_1, \dots, x_n so, daß die x_i parallel zu den Eigenvektoren von $\mathbf{H}(\mathbf{x})$ liegen und insbesondere x_1 parallel zum Gradienten. Setzen wir nun (B.8) in (2.13) ein, so gilt für $j = 2, \dots, n$ die Beziehung⁸

$$\begin{aligned} 0 &= \left(\frac{\partial p_j}{\partial x_1} \right) \bar{t}_1 - \sum_{i=2}^n \left(\frac{\partial p_j}{\partial x_i} \right) \bar{t}_i \\ &= \left(\frac{\partial H_{j1}}{\partial x_1} \right) \|\mathbf{g}(\mathbf{x})\|_2 \cdot \bar{t}_1 - \sum_{i=2}^n \left[\left(\frac{\partial H_{ji}}{\partial x_1} \right) \|\mathbf{g}(\mathbf{x})\|_2 + \lambda_j(\lambda_j - \lambda_1)\delta_{ji} \right] \bar{t}_i. \end{aligned}$$

Hierbei seien mit λ_i die Eigenwerte von $\mathbf{H}(\mathbf{x})$ bezeichnet. Definieren wir nun für $i, j \in \{2, \dots, n\}$ die Matrix $\tilde{\mathbf{J}} = (\tilde{J}_{ji})$ und den Vektor $\mathbf{b} = (b_j)$ gemäß

$$\tilde{J}_{ji} := \left(\frac{\partial H_{ji}}{\partial x_1} \right) + \frac{\lambda_j(\lambda_j - \lambda_1)\delta_{ji}}{\|\mathbf{g}(\mathbf{x})\|_2}, \quad b_j := \tilde{J}_{j1} = \frac{\partial H_{j1}}{\partial x_1}, \quad (2.14)$$

so erhalten wir durch

$$b_j \bar{t}_1 = \sum_{i=2}^n \tilde{J}_{ji} \bar{t}_i$$

ein Gleichungssystem aus $(n-1)$ Gleichungen für die n Unbekannten $\bar{t}_1, \dots, \bar{t}_n$. Ist $\tilde{\mathbf{J}}$ nichtsingulär, dann ergibt sich die Tangente an den GE-Punkt \mathbf{x} aus

$$\bar{\mathbf{t}} = \left(1, (\tilde{\mathbf{J}}^{-1}\mathbf{b})_2, \dots, (\tilde{\mathbf{J}}^{-1}\mathbf{b})_n \right)^T \bar{t}_1, \quad (2.15)$$

⁸Da $d\mathbf{p}$ stets senkrecht auf dem Gradienten steht, erhalten wir an dieser Stelle gemäß (B.8) lediglich $(n-1)$ Gleichungen für $j = 2, \dots, n$.

2. Wegeverfolgung auf Potentialenergieflächen

wobei sich wegen $\|\bar{\mathbf{t}}\|_2 = 1$ die übriggebliebene Unbekannte \bar{t}_1 als Normierungskonstante errechnen läßt.⁹

Mit diesem Vorwissen können wir in den eigentlichen Algorithmus starten. Wie bereits erwähnt, setzt er sich aus zwei Schritten zusammen:

- (i) Befinden wir uns im GE-Punkt \mathbf{x}^i , so bestimmen wir im Predictor-Schritt einen Übergangspunkt $\tilde{\mathbf{x}}^i$ nach der Vorschrift

$$\tilde{\mathbf{x}}^i = \mathbf{x}^i + \alpha \cdot \bar{\mathbf{t}}. \quad (2.16)$$

Hierbei steht α wie üblich für die Schrittweite und $\bar{\mathbf{t}}$ für den gemäß (2.15) definierten normalisierten Tangentenvektor. Wegen (2.14) benötigen wir in diesem Schritt die dritten Ableitungen von E , was einen enormen Rechenaufwand bedeutet.¹⁰

- (ii) Es ist davon auszugehen, daß $\tilde{\mathbf{x}}^i$ nicht auf der GE-Kurve liegt. Daher suchen wir im Corrector-Schritt denjenigen GE-Punkt \mathbf{x}^{i+1} mit dem geringsten Abstand zu $\tilde{\mathbf{x}}^i$. Wir fordern, daß die Schrittweite α in (2.16) so klein gewählt wurde, daß \mathbf{x}^{i+1} der linearen Entwicklung von \mathbf{p} in $\tilde{\mathbf{x}}^i$ genügt:

$$\mathbf{p}(\mathbf{x}^{i+1}) = \mathbf{p}(\tilde{\mathbf{x}}^i) + \sum_{l=1}^n \left. \frac{\partial \mathbf{p}}{\partial x_l} \right|_{\mathbf{x}=\tilde{\mathbf{x}}^i} \cdot [x_l^{i+1} - \tilde{x}_l^i]. \quad (2.17)$$

Da \mathbf{x}^{i+1} das GE-Kriterium (2.10) erfüllen soll, ist (2.17) gerade äquivalent zu:

$$\begin{aligned} -p_j(\tilde{\mathbf{x}}^i) &= \sum_{l=1}^n \left. \frac{\partial p_n}{\partial x_l} \right|_{\mathbf{x}=\tilde{\mathbf{x}}^i} \cdot [x_l^{i+1} - \tilde{x}_l^i] \\ &\stackrel{(B.9)}{=} \sum_{l=1}^n \underbrace{\left(\frac{\partial H_{jl}}{\partial x_1} \right) \|\mathbf{g}\|_2 + (\mathbf{H}^2)_{jl} - H_{11} H_{jl}}_{=: M_{jl}} \Big|_{\mathbf{x}=\tilde{\mathbf{x}}^i} \cdot [x_l^{i+1} - \tilde{x}_l^i] \\ &= \sum_{l=1}^n M_{jl} [x_l^{i+1} - \tilde{x}_l^i] \end{aligned}$$

für $j = 2, \dots, n$. Für $\mathbf{M} = (M_{jl})$ ergibt sich daraus \mathbf{x}^{i+1} nach (B.12) wie folgt:

$$x_l^{i+1} = \tilde{x}_l^i - \sum_{j=2}^n \left[\mathbf{M}^T (\mathbf{M}\mathbf{M}^T)^{-1} \right]_{lj} p_j(\tilde{\mathbf{x}}^i). \quad (2.18)$$

Iterativ wenden wir den nächsten Predictor-Schritt auf \mathbf{x}^{i+1} an.

⁹Ist $\tilde{\mathbf{J}}$ singulär, funktioniert das Verfahren im Allgemeinen nicht. In bestimmten Fällen kann allerdings eine Diagonalisierung von $\tilde{\mathbf{J}}$ weiterhelfen [47].

¹⁰Sun und Ruedenberg schlagen zur Minimierung des Aufwands vor, die dritten Ableitungen numerisch durch die Berechnung einer zweiten Hessematrix im Punkt $\mathbf{x}^i + \tilde{\alpha}\mathbf{v}$ zu bestimmen, wobei \mathbf{v} in Richtung $\mathbf{g}(\mathbf{x}^i)$ zeigt und $\tilde{\alpha}$ hinreichend klein ist [47].

Bemerkung

Sei $\mathbf{x} \in K$ erneut so parametrisiert, daß die Komponenten x_i parallel zu den Eigenvektoren der Hessematrix liegen und insbesondere x_1 parallel zum Gradienten. Dann gilt für die GE-Gleichungen

$$\text{GE}_{i-1}(\mathbf{x}) := \nabla\sigma(\mathbf{x})^T \cdot \mathbf{u}^i(\mathbf{x}) = \frac{\partial\sigma(\mathbf{x})}{\partial x_i} = 0.$$

Somit ergibt die in (2.14) definierte Matrix $\tilde{\mathbf{J}}$ zusammen mit dem Vektor \mathbf{b} die Jacobimatrix \mathbf{J} der Abbildung $\mathbf{GE} : \mathbb{R}^n \rightarrow \mathbb{R}^{n-1}$:

$$\mathbf{J}(\mathbf{x}) = \left(\frac{\partial \text{GE}_{j-1}}{\partial x_i} \right)_{\substack{j=2,\dots,n \\ i=1,\dots,n}} = \left(\mathbf{b} \mid \tilde{\mathbf{J}} \right) \stackrel{[47]}{=} \left[\left(\frac{\partial H_{ji}}{\partial x_1} \right) + \frac{\lambda_j(\lambda_j - \lambda_1)\delta_{ji}}{\|\mathbf{g}(\mathbf{x})\|_2} \right]_{\substack{j=2,\dots,n \\ i=1,\dots,n}}.$$

In Abschnitt 2.1.2 hatten wir die String-Methode vorgestellt und anschließend zur Ermittlung der IRC genutzt. Mit dem bereits gesammelten Wissen ist ihre Anwendung auf Gradient Extremals schnell erklärt. Mittels (2.10) ist eine Startkette $\{\mathbf{x}^i\}_{0 \leq i \leq k}$ solange gemäß

$$\tilde{\mathbf{x}}^i = \mathbf{x}^i - \alpha \cdot \mathbf{p}(\mathbf{x}^i) \stackrel{(2.10)}{=} \mathbf{x}^i - \alpha \cdot \mathbf{P}_{\mathbf{w}(\mathbf{x}^i)}(\mathbf{H}(\mathbf{x}^i)\mathbf{g}(\mathbf{x}^i))$$

zu iterieren bis die gewünschte Genauigkeit ε erreicht ist. Dabei bietet es sich an, als Konvergenzkriterium den Unterschied zwischen zwei Iterationen zu wählen. Doch bei der praktischen Umsetzung ergeben sich Probleme, wie das folgende Beispiel belegt.

Beispiel

Wir betrachten wieder die Potentialenergiefläche (A.4). Abbildung 2.4 zeigt Ergebnisse der String-Methode bei der Verwendung unterschiedlicher Startketten. Die tatsächlichen Gradient Extremals sind jeweils die durchgehenden Linien. Beginnen wir mit einer geraden Kette (dünne Punkte) zwischen beiden Minima, wie in Abbildung 2.4(a) dargestellt, so erhalten wir nach 18 Iterationen ($\alpha = 0.125$, $\varepsilon = 0.01$) sowohl Punkte auf dem Streambed als auch auf dem Talrücken unterhalb des Sattels. Allerdings konvergiert die Methode lediglich in der Umgebung des Sattelpunktes im Intervall $x_1 \in (-1.1, 1.1)$. Startpunkte in der Nähe der Minima liegen weit entfernt von der ursprünglich nächstliegenden GE-Kurve, da \mathbf{p} senkrecht zum (normierten) Gradienten und damit parallel zur Äquipotentialfläche abbildet. Somit werden Punkte nahe des Talgrunds nicht einfach abwärts verschoben, sondern „parallel“ zum GE [37]. Selbst bei einer bereits approximierten Startkurve (dünne Punkte), wie in Abbildung 2.4(b), erhalten wir nach 19 Iterationen ($\alpha = 0.1$, $\varepsilon = 0.01$) lediglich GE-Punkte in einer Umgebung des Sattelpunktes.

2. Wegeverfolgung auf Potentialenergieflächen

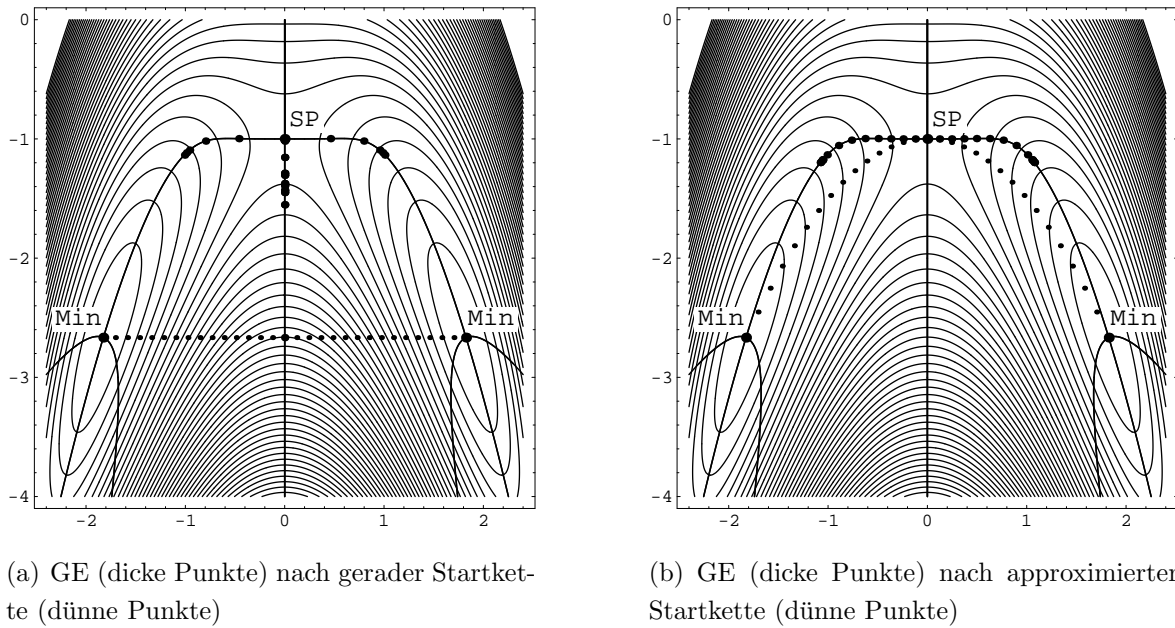


Abbildung 2.4: Gradient Extremals mittels String-Methode

2.2.3. Bifurkationen und Umkehrpunkte auf Gradient Extremals

Wie bereits erwähnt, können Täler auf der PES einfach enden, ohne einen Sattelpunkt erreicht zu haben. Folgt in solchen Fällen ein Gradient Extremal zunächst dem Streambed, so gibt es einen Punkt auf dem GE, in dem die Aufstiegsrichtung in einen Abstieg übergeht oder umgekehrt. Es sind dies die Punkte, in denen die Tangente senkrecht zum Gradienten steht; wir nennen sie *Umkehrpunkte* (Turning Points, TP).

Beispiel

Abbildung 2.5 zeigt eine gemäß (A.5) definierte PES [34], auf der kein Talgrundpfad zwischen Minimum und Sattelpunkt existiert. Vielmehr gibt es ein aufwärts verlaufendes Streambed vom Minimum bis zum ersten Umkehrpunkt bei $(-0.57, 0.06)$. Dort endet das Tal. Ein zweites Flußbett kann vom Sattelpunkt abwärts bis zum zweiten TP verfolgt werden. Zwischen den beiden Umkehrpunkten gibt es keinen Talgrund. Dennoch ist dieser Bereich durch die Lösung einer GE-Gleichung bestimmt [42].

Probleme bereiten uns die Umkehrpunkte aus numerischer Sicht: Für die Ermittlung des GE-Pfads $\{\mathbf{x}^i\}_{0 \leq i \leq k}$ gemäß des Sun-Ruedenberg-Algorithmus [47] ist die Tangente $\bar{\mathbf{t}}$ an die \mathbf{x}^i zu berechnen. Ist $\tilde{\mathbf{J}}$ nichtsingulär, so ist $\bar{\mathbf{t}}$ nach (2.15) eindeutig bestimmt, wobei die \mathbf{x}^i so parametrisiert sind, daß alle Komponenten parallel zu den Eigenvektoren der Hessematrix liegen und x_1^i parallel zum Gradienten. Da jedoch in Umkehrpunkten $\bar{\mathbf{t}}$ senkrecht zu \mathbf{g} steht, erhalten wir in diesem Fall aus (2.15) keine Suchrichtung für den Predictor-Schritt.

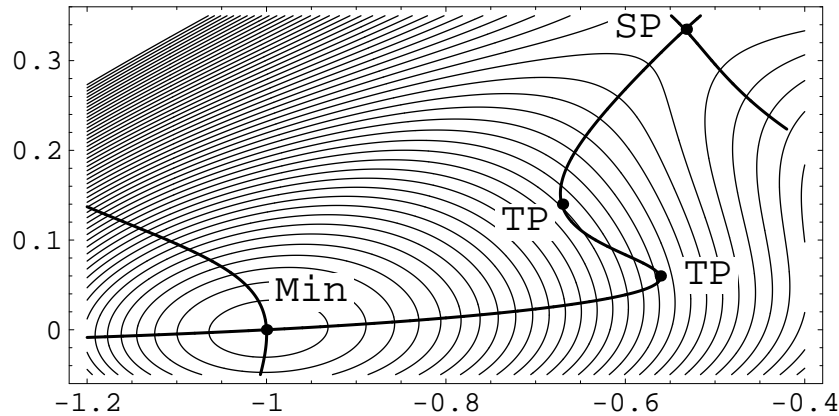


Abbildung 2.5: Umkehrpunkte auf einer Modell-Potentialenergiefläche [34]

Überspringt der Algorithmus etwa in der j -ten Iteration einen Umkehrpunkt, der TP liege also zwischen \mathbf{x}^{j-1} und \mathbf{x}^j , so sind die zugehörigen Tangenten $\bar{\mathbf{t}}^{j-1}$ und $\bar{\mathbf{t}}^j$ entweder beide talaufwärts oder beide talabwärts gerichtet. Mit anderen Worten: Der Algorithmus oszilliert nach (2.16) um den TP und bricht schließlich ab.

Einen Lösungsansatz bietet der TP-Test

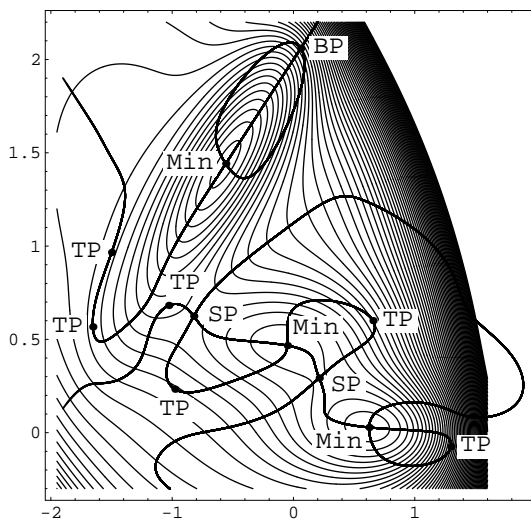
$$\left[\bar{\mathbf{t}}^{j-1}\right]^T \bar{\mathbf{t}}^j < 0$$

und ein Fortfahren des Algorithmus mit $-\alpha$ statt α in (2.16). Die genaue Umsetzung ist in [24] angegeben. Auch die in Abschnitt 3.5 vorgestellte TASC-Methode [42] zeigt, wie ein Streambed beim Auftreten von Umkehrpunkten verfolgt werden kann.

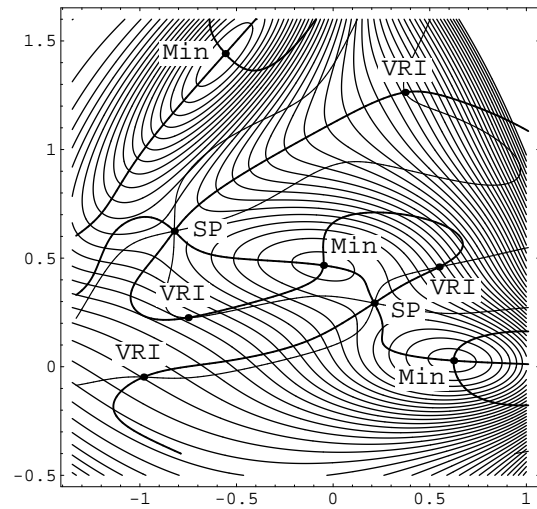
Abschließend wollen wir Bifurkationspunkte auf Gradient Extremals betrachten. In einem BP \mathbf{x}^{bp} schneiden sich zwei oder mehr GE-Kurven. Häufig verzweigt sich ein Tal in \mathbf{x}^{bp} in drei Täler [2], wobei wir darunter meist zwei „echte“ Täler an den Seiten und einen mittig liegenden Bergkessel verstehen. Dies weicht jedoch insbesondere von unserer Vorstellung von VRI-Punkten ab, in deren Nähe sich ein Tal in zwei Täler und einen Rücken aufspaltet. Tatsächlich treten GE-Bifurkationen auf, ohne daß ein VRI-Punkt in der Nähe liegt.

Per Definition verlaufen Gradient Extremals durch Punkte, in denen der Gradient ein Eigenvektor der Hessematrix (o.B.d.A. zum Eigenwert λ_1) ist. Hat ein solcher Punkt zudem einen vom Gradienten verschiedenen Eigenvektor der Hessematrix zum Eigenwert $\lambda_j = 0$ für $j \in \{2, \dots, n\}$, so handelt es sich nach (1.4) um einen VRI-Punkt. Verfolgen wir also GE-Kurven von einer konvexen in eine konkave Region der PES, so liegt im Schnittpunkt von GE und der durch (1.7) definierten Konvexitätsgrenze ein VRI-Punkt, da \mathbf{g} ein Eigenvektor von \mathbf{A} zum Eigenwert $\tilde{\lambda} = \prod_{i=2}^n \lambda_i = 0$ ist. Wie in Abbildung

2. Wegeverfolgung auf Potentialenergieflächen



(a) Wende- und Bifurkationspunkte



(b) VRI-Punkte und Konvexitätsgrenze gemäß (1.7) (dünn durchgehend)

Abbildung 2.6: Wende-, Bifurkations- und VRI-Punkte auf den Gradient Extremals des Müller-Brown-Potentials

2.6(b) zu sehen, liegen sämtliche VRI-Punkte des Müller-Brown-Potentials auf Gradient Extremals. Allerdings findet die GE-Bifurkation weit abseits dieser statt, was ein Vergleich mit Abbildung 2.6(a) belegt.

Dies spiegelt jedoch nicht den allgemeinen Fall wider. In der Definition der VRI-Punkte setzen wir nicht voraus, daß der Gradient ein Eigenvektor der Hessematrix sein muß. In der Regel werden wir nicht sämtliche VRI-Punkte einer PES mithilfe der GE-Verfolgung finden können. Insbesondere im höherdimensionalen Fall liegen die meisten Punkte einer VRI-Mannigfaltigkeit nicht auf Gradient Extremals.

3. Die Methode des reduzierten Gradienten

Die bisher vorgestellten Methoden zur Wegeverfolgung lieferten Lösungsansätze für die Bestimmung des sogenannten „Minimum-Energie-Wegs“, zeigten aber auch Schwächen in der Konstruktion und der numerischen Umsetzung.

So wird die IRC nach Fukui [13] zwar häufig als MEP bezeichnet und erfüllt auch die Monotonieforderung, unterscheidet aber nicht zwischen Tälern und Talrücken auf der vorliegenden Potentialenergiefläche. Zudem setzt sie die Kenntnis des Sattelpunktes voraus und neigt in Minimumnähe zum Zickzacken. Ihr großer Vorteil ist die Einfachheit.

Gradient Extremals verlaufen im Talgrund und zeigen in Bifurkationspunkten Talverzweigungen an. Sie dienen nur dann als Reaktionsweg, falls zwischen zwei stationären Punkten ein Streambed liegt. Dies ist jedoch aufgrund der Vielgestaltigkeit von Potentialenergieflächen keinesfalls vorauszusetzen. Zudem benötigen Verfahren zur GE-Bestimmung einen hohen Rechenaufwand.

Vorrangiges Ziel beider Verfahren ist die Ermittlung stationärer Punkte auf der PES. Während wir mithilfe der IRC Minima erhalten, liefern Gradient Extremals auch die Sattelpunkte, allerdings unter den oben genannten Einschränkungen. In diesem Abschnitt wollen wir die von Quapp, Hirsch, Imig und Heidrich [44] entwickelte Methode des reduzierten Gradienten (RGF-Verfahren) zur Lokalisierung von Sattelpunkten vorstellen. Auch sie wird keine universelle Antwort auf die Frage nach dem „echten“ MEP liefern, wobei an dieser Stelle nochmal angemerkt sei, daß es sich beim Minimum-Energie-Weg um ein theoretisches Konstrukt handelt und wir daher im Rahmen dieser Arbeit lediglich mögliche Modelle für Reaktionswege aufzeigen können, die gewissen Anforderungen der Reaktionstheorie genügen.

3.1. Die Distinguished-Coordinate-Methode

Beginnen möchten wir mit einem Vorläufer des RGF-Verfahrens: der Distinguished-Coordinate-Methode nach Rothman und Lohr [46]. Sie versucht, Sattelpunkte auf der PES nach folgender Vorgehensweise zu bestimmen:

- (i) Wähle in $\mathbf{x} \in K$ eine Koordinate, etwa x_n , aus (Distinguished Coordinate, DC).
- (ii) Minimiere $E(\mathbf{x})$ bei fixiertem x_n und erhalte daraus $\tilde{E}(x_n) = \min_{x_1, \dots, x_{n-1}} E(\mathbf{x})$, wobei $\tilde{E}(x_n) = E(x_1^*, \dots, x_{n-1}^*, x_n)$ mit $x_i^* = x_i^*(x_n)$ für $i = 1, \dots, n-1$ gelte. $\tilde{E}(x_n)$ wird als DC-Kurve bezeichnet.
- (iii) Variiere x_n bis $\tilde{E}(x_n)$ maximal ist. Wird das Maximum für x_n^* erreicht, so ist $\mathbf{x}^* = (x_1^*, \dots, x_n^*)^T$ der gesuchte Sattelpunkt.

3. Die Methode des reduzierten Gradienten

Rothman und Lohr behaupteten, daß die Stetigkeit des Vektors

$$\mathbf{p}(x_n) := (x_1^*(x_n), \dots, x_{n-1}^*(x_n))^T$$

im gewünschten Intervall $x_n \in I$ notwendiges und hinreichendes Kriterium für die Existenz des Maximums in (iii) (und damit des Sattelpunktes) ist.

Williams und Maggiora [49] zeigten jedoch, daß die Distinguished-Coordinate-Methode unter gewissen Umständen fehlschlägt. Sie unterschieden hierbei Potentialenergieflächen, bei denen lediglich ein Minimum zum Sattelpunkt führt (Typ S) und solche, bei denen der Sattelpunkt zwischen Reaktanten- und Produktminimum liegt (Typ D). Während die Methode für Flächen vom Typ S funktioniert, kann sie bei Flächen vom Typ D scheitern, da auf solchen $\mathbf{p}(x_n)$ unstetig ist und damit $\tilde{E}(x_n)$ kein echtes Maximum enthält. In diesen Fällen ist es möglich, daß die DC-Kurve $\tilde{E}(x_n)$ Sprünge enthält.

3.2. Das desingularisierte kontinuierliche Newtonverfahren

Bevor wir die Methode des reduzierten Gradienten erläutern, wollen wir uns mit einem Einschub dem Newtonverfahren widmen. Im Abschnitt 3.3 werden wir dann den Zusammenhang zwischen beiden erklären. Grundlage der folgenden Ausarbeitung ist [9].

In seiner klassischen Formulierung (Satz von Newton-Kantorovich) konvergiert das Verfahren

$$\mathbf{x}^{j+1} := \mathbf{x}^j - \alpha \cdot \mathbf{H}(\mathbf{x}^j)^{-1} \mathbf{g}(\mathbf{x}^j) \quad (3.1)$$

unter gewissen Voraussetzungen gegen eine Nullstelle des Gradienten \mathbf{g} . Die wichtigste Forderung hierbei ist, daß der Startpunkt \mathbf{x}^0 bereits hinreichend nahe am gesuchten stationären Punkt liegt. Der Parameter α bezeichnet wie üblich die Schrittweite. Im Fall $\alpha = 1$ nennen wir (3.1) *klassisches Newtonverfahren*.

Da man sich in diesem Fall jedoch manchmal durch einen Iterationsschritt zu weit von der erwünschten Nullstelle fortbewegt, läßt sich mittels Schrittweitensteuerung¹¹ der Konvergenzbereich des Verfahrens vergrößern. Im Allgemeinen ist hierfür α hinreichend klein zu wählen. Im Grenzfall infinitesimaler Schrittweiten, erhalten wir statt (3.1) die sogenannte kontinuierliche Newtonsche Differentialgleichung

$$\dot{\mathbf{x}} = -\mathbf{H}(\mathbf{x})^{-1} \mathbf{g}(\mathbf{x}) \quad (3.2)$$

¹¹Man verwendet etwa statt \mathbf{x}^{j+1} eine Konvexkombination aus \mathbf{x}^j und \mathbf{x}^{j+1} als nächsten Iterationspunkt.

3.2. Das desingularisierte kontinuierliche Newtonverfahren

und nennen das zugehörige Lösungsverfahren *kontinuierliches Newtonverfahren*. Fixieren wir einen Anfangswert $\mathbf{x}^0 := \mathbf{x}(0)$, so gilt entlang einer Trajektorie $\mathbf{x}(t)$ von (3.2) die Differentialgleichung

$$\dot{\mathbf{g}}(\mathbf{x}(t)) = \mathbf{H}(\mathbf{x}(t))\dot{\mathbf{x}} \stackrel{(3.2)}{=} -\mathbf{g}(\mathbf{x}(t))$$

mit der Lösung

$$\mathbf{g}(\mathbf{x}(t)) = \mathbf{g}(\mathbf{x}^0) \cdot e^{-t}. \quad (3.3)$$

Während also der Betrag des Gradienten im Verlauf der Trajektorie $\mathbf{x}(t)$ exponentiell abnimmt, bleibt die Richtung von \mathbf{g} konstant.

Wollen wir (3.2) numerisch lösen, so bereiten uns die Punkte, in denen die rechte Seite nicht definiert ist, also die Singularitäten von \mathbf{H} , die größten Probleme. Wir gehen daher mithilfe der Adjunkten \mathbf{A} in die desingularisierte Form über und nennen das Verfahren zur Lösung von

$$\dot{\mathbf{x}} = -\mathbf{A}(\mathbf{x}(t))\mathbf{g}(\mathbf{x}(t)) \quad (3.4)$$

das *desingularisierte kontinuierliche Newtonverfahren*. Wegen $\mathbf{A}\mathbf{H} = \det(\mathbf{H})\mathbf{I}_n$ stimmen die Phasenportraits von (3.2) und (3.4) abgesehen von der Orientierung überein. Analog zu (3.2) bleibt die Richtung des Gradienten entlang einer Trajektorie von (3.4) konstant. An dieser Stelle sind zwei Arten von Endpunkten des desingularisierten kontinuierlichen Newtonverfahrens zu unterscheiden:

- (i) Die Punkte, in denen für den Gradienten $\mathbf{g} = \mathbf{0}$ gilt und \mathbf{A} (und damit auch \mathbf{H}) nichtsingulär ist, heißen *wesentliche Singularitäten*. Per Definition sind dies gerade die stationären Punkte einer PES.
- (ii) Die Punkte, in denen $\mathbf{A}\mathbf{g} = \mathbf{0}$ gilt und $\mathbf{g} \neq \mathbf{0}$ ist, heißen *außerordentliche Singularitäten*. Gemäß (1.6) sind dies die VRI-Punkte einer PES.

Wir betrachten die Taylorentwicklung der rechten Seite von (3.4) in einer wesentlichen Singularität \mathbf{x}^* . Dann gilt

$$(-\mathbf{A}\mathbf{g})(\mathbf{x}) \approx -\mathbf{A}(\mathbf{x}^*) \underbrace{\mathbf{g}(\mathbf{x}^*)}_{=0} - \mathbf{A}(\mathbf{x}^*)\mathbf{H}(\mathbf{x}^*)(\mathbf{x} - \mathbf{x}^*) = -\det[\mathbf{H}(\mathbf{x}^*)](\mathbf{x} - \mathbf{x}^*). \quad (3.5)$$

Linearisieren wir also (3.4) in den wesentlichen Singularitäten, so sind diese entweder anziehende ($\det(\mathbf{H}) > 0$) oder abstoßende ($\det(\mathbf{H}) < 0$) Fixpunkte von (3.4).

3. Die Methode des reduzierten Gradienten

Diese Eigenschaft nutzte Branin aus, um numerisch mehrere Nullstellen von \mathbf{g} zu finden [4]. Er fixierte einen Startpunkt \mathbf{x}^0 und ermittelte die \mathbf{x}^0 enthaltende Trajektorie $\mathbf{x}(t)$ von (3.4). Auf dieser Trajektorie bleibt, wie oben erwähnt, die Richtung des Gradienten konstant und $\mathbf{x}(t)$ konvergiert¹² nach einer gewissen Zahl an Iterationen gegen eine Nullstelle \mathbf{x}^* von \mathbf{g} . Dann ist \mathbf{x}^* anziehender Fixpunkt von (3.4). Projiziert man nun die letzte Iteration in Richtung des Gradienten über \mathbf{x}^* hinweg, ändert das Vorzeichen der rechten Seite von (3.4) und fährt mit dem Verfahren fort, so besteht die Möglichkeit weitere Nullstellen von \mathbf{g} zu finden, da durch den Vorzeichenwechsel \mathbf{x}^* vom anziehenden zu einem abstoßenden Fixpunkt wird. Wir erhalten daraus die *Branin-Gleichung*

$$\dot{\mathbf{x}} = \pm \mathbf{A}(\mathbf{x}(t)) \mathbf{g}(\mathbf{x}(t)), \quad (3.6)$$

wobei wir den „+“-Fall für die Suche nach stationären Punkten mit ungeradem Index (etwa Sattelpunkte vom Index 1) verwenden und den „-“-Fall entsprechend für die Suche nach stationären Punkten mit geradem Index (etwa Minima). Offensichtlich sind die Endpunkte des desingularisierten kontinuierlichen Newtonverfahrens (3.4) (also die wesentlichen und außerordentlichen Singularitäten) auch Endpunkte der Braningleichung.

3.3. Das RGF-Verfahren und Newtontrajektorien

In der Literatur ist es üblich, Lösungen von (3.2) oder (3.4) als *Newtontrajektorien* zu bezeichnen. Wir wollen in dieser Arbeit den Begriff der Newtontrajektorie unter dem Aspekt definieren, daß die Anfangsrichtung des Gradienten im Verlauf konstant bleibt. Es wird sich zeigen, daß dies mit der herkömmlichen Verwendung des Begriffs im Einklang steht. Zunächst sei aber das RGF-Verfahren definiert.

Definition

Für die *Suchrichtung* $\mathbf{r} \in \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\|_2 = 1\}$ wird die Predictor-Corrector-Methode zur Lösung der Gleichung

$$\mathbf{P}_{\mathbf{r}}(\mathbf{g}) = \mathbf{0} \quad (3.7)$$

als *Methode des reduzierten Gradienten* (Reduced Gradient Following, RGF) bezeichnet. Die durch (3.7) definierte Kurve nennen wir *RGF-Kurve*.

Bemerkungen

- (i) Im zweidimensionalen Fall gibt man Suchrichtungen häufig als Winkel an, wobei $\mathbf{r} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ auf 0° gesetzt wird.

¹²Das Auftreten außerordentlicher Singularitäten sei vorübergehend ausgeschlossen.

- (ii) Statt des gemäß (2.5) definierten Projektionsoperators können wir in (3.7) auch diejenige Matrix $\overline{\mathbf{P}}_{\mathbf{r}} \in \mathbb{R}^{(n-1) \times n}$ verwenden, deren Zeilen sich aus den $(n-1)$ zu \mathbf{r} orthogonalen Einheitsvektoren des \mathbb{R}^n zusammensetzen. Da die Zeilen von $\overline{\mathbf{P}}_{\mathbf{r}}$ zusammen mit \mathbf{r} eine Orthonormalbasis des \mathbb{R}^n bilden, gilt

$$\ker(\overline{\mathbf{P}}_{\mathbf{r}}) = \{\mu\mathbf{r}\} = \text{lin}\{\mathbf{r}\}, \quad \text{im}(\overline{\mathbf{P}}_{\mathbf{r}}) = \text{lin}\{\mathbf{r}\}^\perp,$$

wodurch sich die äquivalente Verwendung von $\mathbf{P}_{\mathbf{r}}$ und $\overline{\mathbf{P}}_{\mathbf{r}}$ in (3.7) rechtfertigen läßt.

- (iii) Die hier angegebene Definition des RGF-Verfahrens stellt eine Verallgemeinerung der ursprünglichen Formulierung nach [44] dar. Gemäß dieser wird eine Kurve betrachtet, deren Punkte \mathbf{x} die $(n-1)$ Gleichungen

$$\frac{\partial E(\mathbf{x})}{\partial x_i} = 0 \tag{3.8}$$

für $i \in \{1, \dots, n\} \setminus \{j\}$ erfüllen. Aus (3.8) folgt, daß der Gradient (konstant) in Richtung der Koordinate x_j zeigt. Sie ist damit eine mathematische Neuformulierung der Distinguished-Coordinate-Methode und begründet zudem die Bezeichnung des *reduzierten Gradienten*.

Definition

Sei $\mathbf{r} \in \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\|_2 = 1\}$ eine Suchrichtung. Dann bezeichnen wir die Menge

$$\text{NT}_{\mathbf{r}} := \{\mathbf{x} \in K : \mathbf{g}(\mathbf{x}) = \mathbf{r}\|\mathbf{g}(\mathbf{x})\|_2\} \tag{3.9}$$

als *Newtontrajektorie* (Newton Trajectory, NT) in K entlang der Richtung \mathbf{r} .

Newtontrajektorien bestehen also aus genau den Punkten, bei denen der Gradient in Richtung \mathbf{r} zeigt. Wir sehen sofort, daß $\text{NT}_{\mathbf{r}}$ Lösungsmenge der Gleichung (3.7) ist:

$$\begin{aligned} \mathbf{P}_{\mathbf{r}}(\mathbf{g}) = \mathbf{0} & \stackrel{(2.6)}{\iff} \mathbf{0} = \mathbf{P}_{\mu\mathbf{g}}(\mathbf{g}) = (\mathbf{I}_n - \mu^2\mathbf{g}\mathbf{g}^T)\mathbf{g} = (1 - \mu^2\|\mathbf{g}\|_2^2)\mathbf{g} \\ & \iff \mathbf{r} = \mu\mathbf{g} = \mathbf{g}/\|\mathbf{g}\|_2. \end{aligned} \tag{3.10}$$

Mithin zeigt \mathbf{g} entlang einer Lösung von (3.7) (konstant) in Richtung \mathbf{r} . Damit sind wegen (3.3) die Trajektorien des (desingularisierten) kontinuierlichen Newtonverfahrens sowie der Braningleichung (3.6) auch (Zweige von)¹³ Newtontrajektorien in unserem Sinne.

Wegen der Eindeutigkeit des Gradienten ist in jedem nichtstationären Punkt $\mathbf{x} \in K$ durch die Richtung von $\mathbf{g}(\mathbf{x})$ genau eine NT gegeben. In stationären Punkten hingegen ist (3.9) für beliebiges \mathbf{r} erfüllt. Nur dort können sich Newtontrajektorien zu unterschiedlichen Suchrichtungen schneiden, wie Abbildung 3.1(b) verdeutlicht.

¹³Den Begriff es *Zweigs* einer Newtontrajektorie führen wir auf Seite 44 ein.

3. Die Methode des reduzierten Gradienten

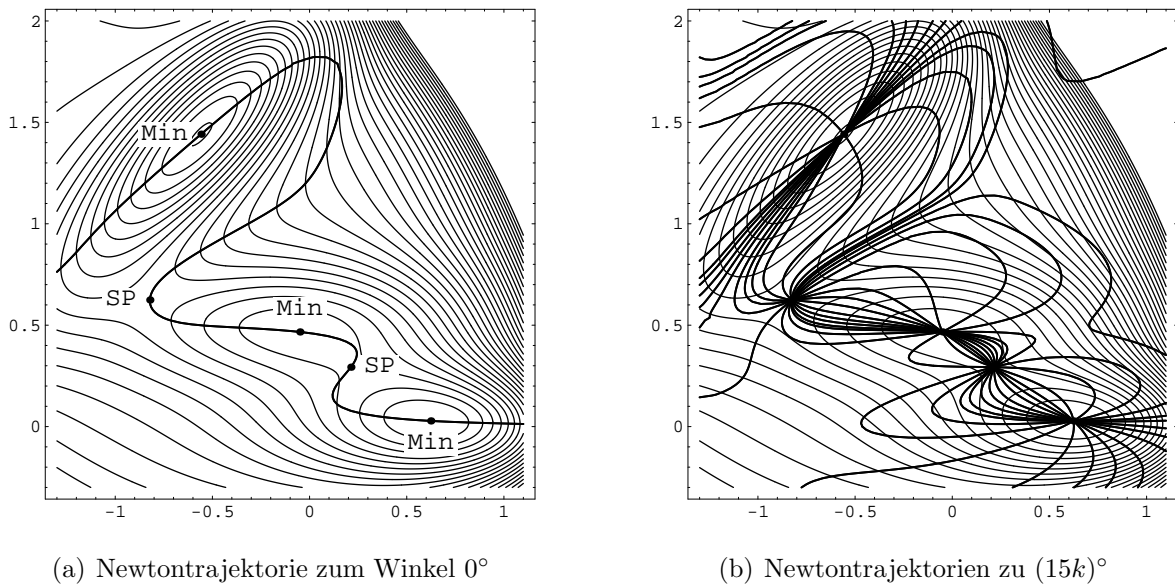


Abbildung 3.1: Newtontrajektorien auf dem Müller-Brown-Potential

Numerisch verfolgen wir die durch (3.7) gegebene RGF-Kurve¹⁴ entlang ihrer Tangente. Diese ist im Punkt $\mathbf{x}(t)$ bestimmt durch

$$\mathbf{0} = \frac{d\mathbf{P}_r(\mathbf{g}(\mathbf{x}))}{dt} = \mathbf{P}_r \left(\frac{d\mathbf{g}(\mathbf{x})}{dt} \right) = \mathbf{P}_r(\mathbf{H}(\mathbf{x})\dot{\mathbf{x}}) = \mathbf{P}_r(\mathbf{H}(\mathbf{x}))\dot{\mathbf{x}}. \quad (3.11)$$

Es ergeben sich zwei Arten von Singularitäten, die in der RGF-Tangentengleichung (3.11) auftreten können:

- (i) Es gilt $\mathbf{g}(\mathbf{x}) = \mathbf{0}$, das heißt, \mathbf{x} ist stationärer Punkt. Dies sind gerade die wesentlichen Singularitäten von (3.4) und (3.6).
- (ii) Der Rang von $\mathbf{P}_r(\mathbf{H}(\mathbf{x}))$ fällt ab. Dies geschieht genau dann, wenn $\mathbf{H}(\mathbf{x})$ einen senkrecht auf \mathbf{r} stehenden Eigenvektor \mathbf{u}^{zero} zum Eigenwert Null hat [19]. Wegen (3.9) steht dann \mathbf{u}^{zero} auch senkrecht auf \mathbf{g} , und es handelt sich bei \mathbf{x} um einen VRI-Punkt, mithin um eine außerordentliche Singularität von (3.4) und (3.6).

Damit sind die singulären Punkte der RGF-Tangentengleichung (3.11) identisch mit den Endpunkten der Braningleichung (3.6). Betrachten wir lediglich die nichtsingulären Punkte einer Newtontrajektorie, so erhalten wir eine Menge differenzierbarer und parametrisierbarer Kurvenstücke, an deren Enden singuläre Punkte (stationäre oder VRI) oder Punkte aus ∂K liegen. Diese Kurvenstücke wollen wir *Zweige* nennen und wie folgt charakterisieren:

¹⁴Die Äquivalenz (3.10) rechtfertigt die synonyme Verwendung der Begriffe „RGF-Kurve“ und „Newtontrajektorie“.

- (i) Verbindet ein NT-Zweig zwei stationäre Punkte miteinander, dann nennen wir ihn *regulär*. Wegen (3.5) ist einer der beiden Punkte anziehend, der andere abstoßend. Reguläre NT-Zweige verbinden somit stationäre Punkte, für deren Indizes i_1 und i_2 gilt, daß $i_1 \neq i_2 \pmod{2}$.
- (ii) Endet ein NT-Zweig in einer außerordentlichen Singularität, so nennen wir ihn *singular*.
- (iii) Anderenfalls enthält der Zweig keine Singularitäten oder endet auf ∂K .

Da Newtontrajektorien im allgemeinen in verschiedene *Zusammenhangskomponenten* zerfallen [9], müßte man auf jeder Komponente wenigstens einen Startpunkt für den numerischen Verfolgungsalgorithmus plazieren, um alle stationären Punkte der PES zu finden. Nach [9] ist die Anzahl der stationären Punkte auf einer kompakten und zu ∂K disjunkten Komponente einer NT restlos durch zwei teilbar. Starten wir also einen Algorithmus auf einer solchen Komponente in einem stationären Punkt, so finden wir mit Sicherheit einen zweiten. Entlehnt an [19] zeigt Abbildung 3.2 eine Newtontrajektorie mit ihren Zweigen und Komponenten.

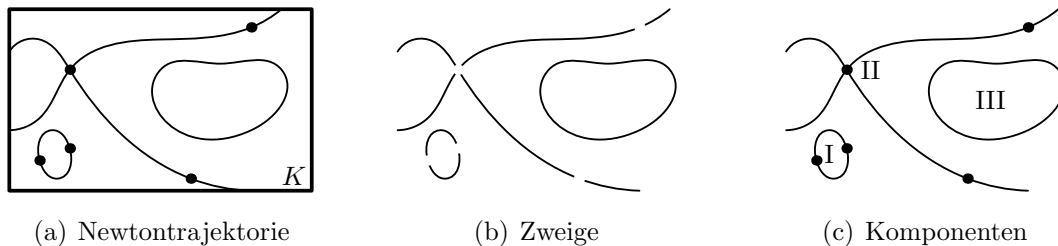


Abbildung 3.2: Newtontrajektorie mit ihren Zweigen und Komponenten

Seien nun $\mathbf{x}^* \in \text{NT}_{\mathbf{r}}$ ein nichtsingulärer Punkt und die Newtontrajektorie $\text{NT}_{\mathbf{r}}$ in einer Umgebung von \mathbf{x}^* nach t parametrisiert. Dann ist $\mathbf{x}^* = \mathbf{x}(t^*)$ ein lokales Extremum der Trajektorie, falls $E(\mathbf{x}^*)$ extremal ist in der Menge $\{E(\mathbf{x}(t)) : t \in (t^* - \varepsilon, t^* + \varepsilon)\}$. Die Tangente $\dot{\mathbf{x}}^*$ liegt tangential zur durch \mathbf{x}^* verlaufenden Äquipotentialfläche, senkrecht zum Gradienten $\mathbf{g}(\mathbf{x}^*)$ und nach (3.9) damit auch senkrecht zur Suchrichtung \mathbf{r} . Die lokalen Extrema sind ausgezeichnete Umkehrpunkte, die wir allgemein definieren wollen als Punkte \mathbf{x}^{tp} , in denen die Tangente senkrecht zur Suchrichtung steht. Anders als bei der Distinguished-Coordinate-Methode, die einen Pfad nicht über \mathbf{x}^{tp} hinweg verfolgen kann und springt, sind die Umkehrpunkte keine Singularitäten des RGF-Verfahrens (3.7). Die aus der DC-Methode resultierenden Kurvenstücke sind allerdings in einer Newtontrajektorie enthalten.

3. Die Methode des reduzierten Gradienten

3.3.1. VRI-Punkte als außerordentliche Singularitäten

Wir wissen bereits, daß sich wegen der Eindeutigkeit des Gradienten Newtontrajektorien zu verschiedenen Suchrichtungen nur in stationären Punkten schneiden können, da in diesen die Richtung des Gradienten unbestimmt ist. Andererseits kann es jedoch vorkommen, daß sich zwei Zweige derselben RGF-Kurve schneiden. Dies geschieht gerade in den VRI-Punkten:

Gemäß (3.11) handelt es sich bei VRI-Punkten um die Punkte $\mathbf{x}^{\text{vri}} \in K$, an denen die sogenannte *reduzierte Hessematrix* $\mathbf{P}_r(\mathbf{H}(\mathbf{x}^{\text{vri}}))$ einen Rangabfall bzw. einen nichttrivialen Kern hat. Betrachten wir zweidimensionale Potentialenergieflächen, dann ist in einem solchen Fall $\ker[\mathbf{P}_r(\mathbf{H}(\mathbf{x}^{\text{vri}}))]$ eindimensional¹⁵, und es gibt zwei Tangenten an die durch \mathbf{x}^{vri} verlaufende Newtontrajektorie. Mit anderen Worten: Die Newtontrajektorie verzweigt sich in \mathbf{x}^{vri} . Abbildung (3.3) zeigt uns die VRI-Punkte des Müller-Brown-Potentials als Verzweigungen von RGF-Kurven; die zugehörigen Suchrichtungen sind [19] entnommen.

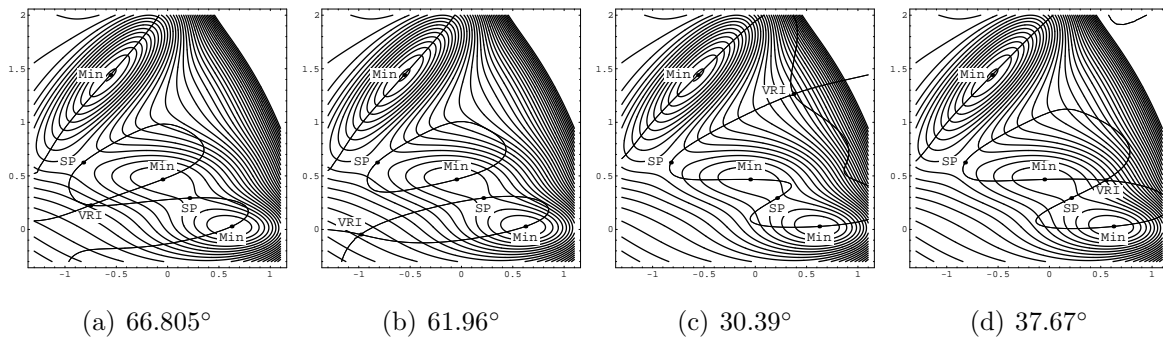


Abbildung 3.3: NT-Verzweigungen auf dem Müller-Brown-Potential

Der Begriff des *VRI-Punktes* suggeriert einen (nulldimensionalen) singulären Punkt auf der Potentialenergiefläche. Da zur Anschauung häufig auf zweidimensionale Modellpotentiale zurückgegriffen wird, verstärkt sich dieser Eindruck. Davon ist jedoch im allgemeinen Fall unbedingt Abstand zu nehmen. Wir wissen, daß VRI-Punkte vom gewählten Reaktionsweg unabhängige Punkte auf der PES sind, die wir zudem mit den außerordentlichen Singularitäten des desingularisierten kontinuierlichen Newtonverfahrens identifizieren können. Nach [27] ist die Menge der VRI-Punkte eine bis zu $(n - 2)$ -dimensionale abgeschlossene stratifizierte Teilmenge¹⁶ des Konfigurationsraum $K \subset \mathbb{R}^n$.

¹⁵Auf n -dimensionalen Potentialenergieflächen kann der Rang der reduzierten Hessematrix weiter abfallen. Gilt etwa $k := \dim \ker[\mathbf{P}_r(\mathbf{H}(\mathbf{x}^{\text{vri}}))]$, so nennen wir \mathbf{x}^{vri} einen k -fachen *VRI-Punkt*.

¹⁶Als stratifizierte Teilmenge von K bezeichnen wir eine disjunkte, lokal-endliche Vereinigung glatter Untermannigfaltigkeiten von K , wobei die Untermannigfaltigkeit höchster Dimension die Dimension der stratifizierten Teilmenge bestimmt.

Hierbei stellt sich natürlich die Frage nach der Bedeutung hochdimensionaler VRI-Punkt-Mannigfaltigkeiten in Bezug auf Reaktionswegverzweigungen und mithin auch nach der Rechtfertigung dafür, Reaktionswege als einzelne, ausgezeichnete Kurven auf der PES zu betrachten [19]. Im Rahmen dieser Arbeit kann hierzu keine weitere Diskussion erfolgen.

Reguläre NT-Zweige haben die Eigenschaft, stationäre Punkte mit den Indizes $i_1 \neq i_2 \pmod{2}$ miteinander zu verbinden. Existiert also beispielsweise eine Newtontrajektorie, die zwischen zwei Sattelpunkten (vom Index 1) keinen anderen stationären Punkt und insbesondere kein Minimum durchläuft, so muß auf dem Weg ein VRI-Punkt liegen.

Einen Ansatz für die Bestimmung von VRI-Punkten liefert die Braningleichung (3.6). Von einem beliebigen nichtstationären Punkt der PES ausgehend, enden ihre Trajektorien entweder in VRI- oder stationären Punkten. Da die Gradientenrichtung entlang der Trajektorie konstant bleibt, muß bei einer Suche nach \mathbf{x}^{vri} der Startpunkt \mathbf{x}^0 des Braninverfahrens

$$\mathbf{x}^{i+1} = \mathbf{x}^i \mp \alpha \mathbf{A}(\mathbf{x}^i) \mathbf{g}(\mathbf{x}^i)$$

so gewählt werden, daß die Gradientenrichtung von \mathbf{x}^0 exakt mit der von \mathbf{x}^{vri} übereinstimmt. Im Fall symmetrischer VRI-Punkte kann man sich durch die Betrachtung symmetrischer Unterräume von K systematisch an VRI-Punkt-Mannigfaltigkeiten herantasten [21, 41].

Das Problem bei der Suche nichtsymmetrischer VRI-Punkte¹⁷ besteht also genau darin, die Suchrichtung \mathbf{r} so zu bestimmen, daß sie der Gradientenrichtung des VRI-Punktes entspricht. Später werden wir hierfür einen Variationsansatz nutzen. Abbildung 3.4 zeigt eine Schar von Newtontrajektorien (grau) auf der PES (A.1). Lediglich die NT zur Suchrichtung $\mathbf{r} = (-0.25, 2)^T$ (schwarz) verläuft durch den VRI-Punkt $\mathbf{x}^{\text{vri}} = \mathbf{0}$.

3.3.2. Newtontrajektorien als Reaktionswege

Das RGF-Verfahren liefert uns eine Schar von Kurven, in denen die stationären Punkte sowie die VRI-Punkte einer Potentialenergiefläche enthalten sind. Offensichtlich sind nicht alle dieser Kurven auch Minimum-Energie-Wege. Allerdings erfüllen einige die qualitativen Forderungen des MEP-Konzeptes (Seite 13). Will man Newtontrajektorien als Reaktionsweg betrachten, ist daher eine geschickte Wahl der Suchrichtung \mathbf{r} vonnöten. Eine Newtontrajektorie, in deren Verlauf vom Minimum zum Sattelpunkt die potentielle Energie streng monoton wächst, verläuft mit Sicherheit durch ein Talgebiet [19]. Wir können jedoch nicht davon ausgehen, daß eine solche NT stets existiert.

¹⁷Vermutlich sind in der Theorie der Reaktionswegverzweigungen asymmetrische VRI-Punkte unbedeutender als symmetrische [36]. Natürlich besteht trotzdem ein Interesse an ihrer Ermittlung.

3. Die Methode des reduzierten Gradienten

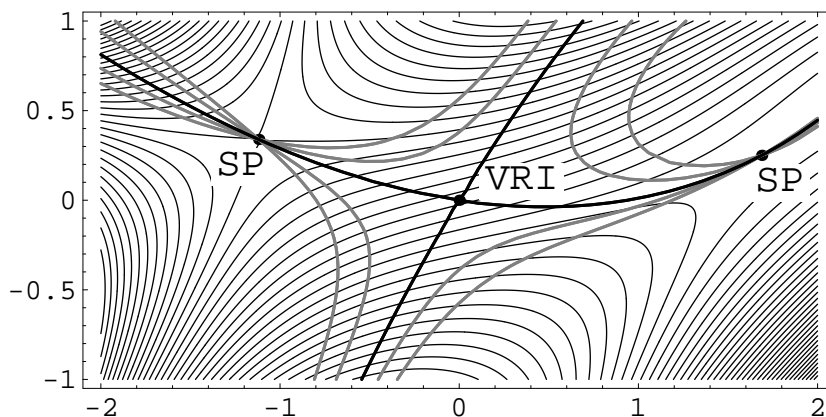


Abbildung 3.4: Newtontrajektorie durch asymmetrischen VRI-Punkt

Beispiel

Abbildung 3.5 zeigt die gemäß (A.6) definierte PES nach Neria, Fischer und Karplus [32] in modifizierter Form [19]. Sie dient als Modell zur Beschreibung der Verhältnisse im Alanindipeptid. Wir sehen, daß lediglich die GE-Kurve (schwarz) Minima und Sattelpunkt miteinander verbindet und dabei vollständig im Tal verläuft, was ihre Bedeutung als Index des Talgrunds unterstreicht. Eine Schar von Newtontrajektorien (grau) verbindet zwar die stationären Punkte, verbleibt aber dabei nicht ausschließlich in Talgebieten; es gibt keine NT, die dieser Anforderung genügt.

Wir sehen anhand von Abbildung 3.5, daß auch das RGF-Verfahren nicht universell geeignet ist Reaktionswege zu bestimmen. Es empfiehlt sich mithin eine Kombination der in dieser Arbeit vorgestellten Verfahren (IRC, GE, RGF) zur Untersuchung derjenigen Bereiche einer Potentialenergiefläche, durch die der Reaktionsweg verläuft.

3.4. Numerische Verfolgung von RGF-Kurven

Wie im Abschnitt 2.2 zum GE-Verfahren wollen wir auch hier eine Predictor-Corrector-Methode angeben, anschließend eine Verfeinerung derselben und zuletzt die RGF-Bestimmung mithilfe von Ketten (Strings) erläutern.

Die grundlegende Idee der Predictor-Corrector-Methode besteht darin, die Kurve zunächst entlang ihrer Tangente zu verfolgen und anschließend so zu korrigieren, daß (3.7) wieder erfüllt ist. In der Beschreibung des Algorithmus halten wir uns an [41] und verwenden statt des Projektionsoperators \mathbf{P}_r die in Bemerkung (ii) (Seite 43) eingeführte Matrix $\bar{\mathbf{P}}_r$. Dann sind folgende Schritte durchzuführen:

- (i) Wir bestimmen die Matrix $\bar{\mathbf{P}}_r$ unter Verwendung des Gram-Schmidtschen Orthonormalisierungsverfahrens mit \mathbf{r} als Startvektor.

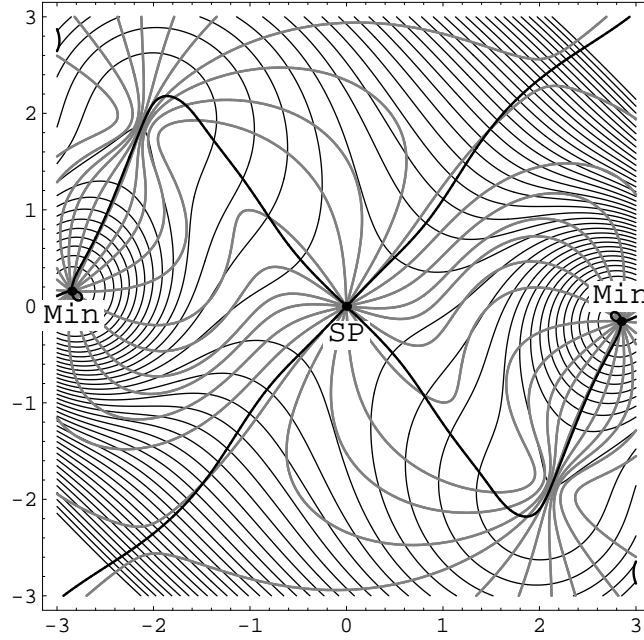


Abbildung 3.5: Gradient Extremals und Newtontrajektorien auf der modifizierten PES nach Neria, Fischer und Karplus

(ii) Predictor-Schritt. Wir bestimmen die nach (3.11) gemäß¹⁸

$$\mathbf{0} = \bar{\mathbf{P}}_r \mathbf{H} \dot{\mathbf{x}} = \bar{\mathbf{P}}_r \mathbf{H} \mathbf{t}$$

durch $(n - 1)$ Gleichungen gegebene Tangente \mathbf{t} mittels QR-Dekomposition und erhalten daraus den Zwischenschritt

$$\tilde{\mathbf{x}}^i = \mathbf{x}^i + \alpha \cdot \frac{\mathbf{t}(\mathbf{x}^i)}{\|\mathbf{t}(\mathbf{x}^i)\|_2} .$$

(iii) Ist $\|\bar{\mathbf{P}}_r \mathbf{g}(\tilde{\mathbf{x}}^i)\|_2 > \varepsilon$, dann führen wir einen Corrector-Schritt durch. Der Vektor \mathbf{c} , der von $\tilde{\mathbf{x}}^i$ nach \mathbf{x}^{i+1} führt, ist nach [1] gegeben durch

$$\bar{\mathbf{P}}_r \mathbf{H}(\tilde{\mathbf{x}}^i) \mathbf{c} = -\bar{\mathbf{P}}_r \mathbf{g}(\tilde{\mathbf{x}}^i) .$$

Für die Eindeutigkeit von \mathbf{c} fordert [1] zudem, daß die Korrektur senkrecht zur Tangente erfolgt, daß also $\mathbf{c}^T \mathbf{t}(\tilde{\mathbf{x}}^i) = 0$ gilt. Damit erhalten wir \mathbf{c} aus

$$\begin{bmatrix} \bar{\mathbf{P}}_r \mathbf{H}(\tilde{\mathbf{x}}^i) \\ \mathbf{t}(\tilde{\mathbf{x}}^i)^T \end{bmatrix} \mathbf{c} = \begin{bmatrix} -\bar{\mathbf{P}}_r \mathbf{g}(\tilde{\mathbf{x}}^i) \\ 0 \end{bmatrix} . \quad (3.12)$$

¹⁸Da sowohl \mathbf{P}_r als auch $\bar{\mathbf{P}}_r$ im zeitlichen Verlauf konstant bleibt, können wir ohne Einschränkung (3.11) mit $\bar{\mathbf{P}}_r$ verwenden.

3. Die Methode des reduzierten Gradienten

Die Predictor-Corrector-Methode benötigt die Hessematrix oder Updates von \mathbf{H} . Sie liegt damit in Bezug auf den Rechenaufwand zwischen IRC und GE.

Eine Verfeinerung des Verfahrens ergibt sich aus der Kombination von Predictor- und Corrector-Schritt nach Diener und Schaback [10]. Sie schlagen folgenden Algorithmus vor: Ermittle zunächst einen normierten Vektor $\bar{\mathbf{t}}(\mathbf{x}) \in \mathbb{R}^n$, der die Gleichung

$$\bar{\mathbf{P}}_r \mathbf{H}(\mathbf{x}) \bar{\mathbf{t}}(\mathbf{x}) = \mathbf{0} \quad (3.13)$$

erfüllt und bestimme anschließend $\mathbf{h}(\mathbf{x}, \alpha)$ aus dem System

$$\begin{bmatrix} \bar{\mathbf{P}}_r \mathbf{H}(\mathbf{x}) \\ \bar{\mathbf{t}}(\mathbf{x})^T \end{bmatrix} \mathbf{h}(\mathbf{x}, \alpha) = \begin{bmatrix} -\bar{\mathbf{P}}_r \mathbf{g}(\mathbf{x}) \\ \alpha \end{bmatrix}. \quad (3.14)$$

Dann gelten folgende Bedingungen:

- (i) Der Vektor $\bar{\mathbf{t}}(\mathbf{x})$ liegt nach (3.11) tangential zur Newtontrajektorie durch \mathbf{x} .
- (ii) $\mathbf{h}(\mathbf{x}, 0)$ entspricht \mathbf{c} im Corrector-Schritt (3.12) des obigen Algorithmus.
- (iii) Wegen

$$\bar{\mathbf{t}}(\mathbf{x})^T \mathbf{h}(\mathbf{x}, \alpha) \stackrel{(3.14)}{=} \alpha = 0 + \bar{\mathbf{t}}(\mathbf{x})^T \bar{\mathbf{t}}(\mathbf{x}) \alpha \stackrel{(3.14)}{=} \bar{\mathbf{t}}(\mathbf{x})^T \mathbf{h}(\mathbf{x}, 0) + \bar{\mathbf{t}}(\mathbf{x})^T \bar{\mathbf{t}}(\mathbf{x}) \alpha$$

gilt $\mathbf{h}(\mathbf{x}, \alpha) = \mathbf{h}(\mathbf{x}, 0) + \alpha \cdot \bar{\mathbf{t}}(\mathbf{x})$.

Liegt die aktuelle Iteration \mathbf{x}^i bereits auf der gewünschten Newtontrajektorie, so können wir $\mathbf{h}(\mathbf{x}^i, \alpha)$ wegen (i) und (iii) als Predictor-Schritt der Länge α auffassen. Liegt \mathbf{x}^i hingegen (mit geringer Entfernung) abseits der Trajektorie, so ist $\bar{\mathbf{t}}(\mathbf{x}^i)$ beinahe parallel zur Tangente des Trajektorienpunktes \mathbf{x}^{i+1} , der zu \mathbf{x}^i den geringsten Abstand aufweist. Da $\mathbf{h}(\mathbf{x}^i, 0)$ und $\bar{\mathbf{t}}(\mathbf{x}^i)$ orthogonal zueinander stehen, zeigt $\mathbf{h}(\mathbf{x}^i, 0)$ folglich in Richtung \mathbf{x}^{i+1} . Mithin fungiert $\mathbf{h}(\mathbf{x}^i, \alpha)$ für kleine α als Corrector-Schritt. In günstigen Fällen kommt dieses Verfahren ohne zusätzliche Corrector-Schritte (3.12) aus [20].

Eine Anwendung der String-Methode auf die RGF-Gleichung (3.7) führt uns zu folgendem Vorgehen: Setzen wir $\mathbf{p} := \mathbf{P}_r(\mathbf{g})$, so ist eine Startkette $\{\mathbf{x}^i\}_{0 \leq i \leq k}$ solange gemäß

$$\tilde{\mathbf{x}}^i = \mathbf{x}^i - \alpha \cdot \mathbf{p}(\mathbf{x}^i) = \mathbf{x}^i - \alpha (\mathbf{I}_n - \mathbf{r}\mathbf{r}^T) \mathbf{g}(\mathbf{x}^i) \quad (3.15)$$

zu iterieren bis sämtliche Kettenpunkte (3.7) mit gewünschter Genauigkeit erfüllen. Typisches Abbruchkriterium ist die Differenz ε zwischen j -ter und $(j+1)$ -ter Iteration. Als Suchrichtung bietet sich $\mathbf{r} = (\mathbf{x}^k - \mathbf{x}^0) / \|\mathbf{x}^k - \mathbf{x}^0\|_2$ an, also beispielsweise die Richtung zwischen Reaktanten- und Produktminimum.

Die Vorteile der String-Methode sind offensichtlich: Für die Bestimmung der iterierten Kettenpunkte sind weder Tangenten an die Kurve noch zweite Ableitungen zu berechnen. Desweiteren kann jeder Kettenpunkt einzeln mittels (3.15) verschoben werden – eine Eigenschaft, die wir für die Ermittlung der VRI-Punkte in Abschnitt 5 nutzen werden. Das größte Problem bei der numerischen Umsetzung besteht in der Wahl einer geeigneten Schrittweite α , insbesondere in sehr flachen Tälern der PES. Hier kann das oben beschriebene String-Verfahren sogar scheitern. Eine Verbesserung liefert die Growing-String-Methode (Growing String, GS) nach [38]. Wir wollen an dieser Stelle ihre neueste Formulierung gemäß [40] vorstellen, kehren aber später für die VRI-Bestimmung wieder zum gewöhnlichen String-Verfahren zurück.

Gehen wir zunächst von zwei fixierten Kettenendpunkten \mathbf{x}^0 und \mathbf{x}^k aus, so erklärt sich die GS-Methode wie folgt:

- (i) Predictor-Schritt. Für $i = 0, \dots, k - 2$ setzen wir $\xi_i := \frac{k-(i+1)}{k-i}$ und berechnen sukzessive die Hilfspunkte

$$\tilde{\mathbf{x}}^{i+1} := \xi_i \mathbf{x}^i + (1 - \xi_i) \mathbf{x}^k .$$

- (ii) Corrector-Schritt. Erfüllt der Hilfspunkt $\tilde{\mathbf{x}}^{i+1}$ die NT-Bedingung (3.7) mit einer vorgegebenen Genauigkeit ε , so setzen wir $\mathbf{x}^{i+1} := \tilde{\mathbf{x}}^{i+1}$ und kehren zu (i) zurück. Anderenfalls korrigieren wir $\tilde{\mathbf{x}}^{i+1}$ gemäß (3.12) zum neuen Kettenpunkt \mathbf{x}^{i+1} .

Da das GS-Verfahren im höherdimensionalen Fall ungleich schneller konvergiert als die gewöhnliche String-Methode und zudem deren Schwäche in sehr flachen Tälern behebt, wird der erhöhte Rechenaufwand in Kauf genommen.

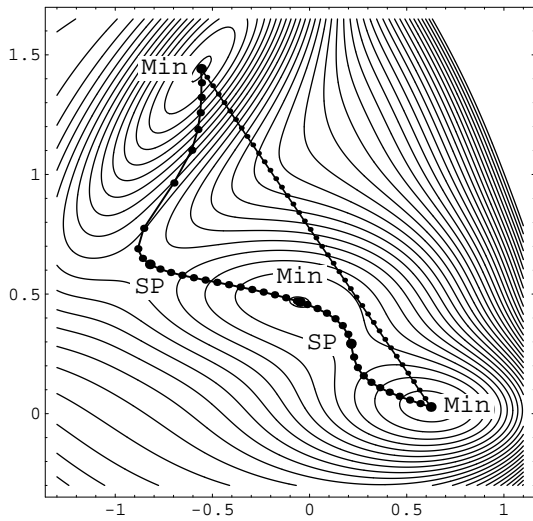
Beispiel

Abbildung 3.6 zeigt zwei Anwendungen der (gewöhnlichen) String-Methode auf dem Müller-Brown-Potential. In beiden Fällen iterieren wir die gerade Startkette (dünne Punkte) zwischen nördlichem und südöstlichem Minimum mit einer Schrittweite $\alpha = 0.5$ und einer kleineren Iterationsdifferenz als $\varepsilon = 0.01$. In Abbildung 3.6(a) wählen wir \mathbf{r} als Richtung zwischen \mathbf{x}^0 und \mathbf{x}^k und erhalten Konvergenz nach 44 Iterationen. In Abbildung 3.6(b) ist \mathbf{r} die Richtung zwischen nördlichem und mittlerem Minimum. Konvergenz ergibt sich nach 32 Iterationen.

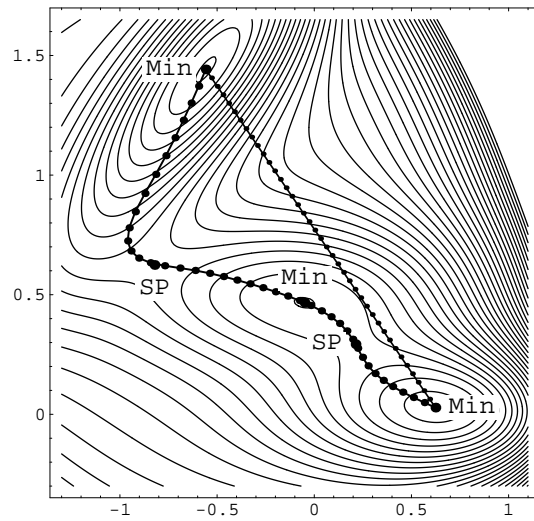
3.5. Das TASC-Verfahren

In Abschnitt 3.3.2 stellten wir fest, daß das RGF-Verfahren zwar ein probates Mittel zur Bestimmung aller stationären Punkte einer PES darstellt, es allerdings nur dann mögliche

3. Die Methode des reduzierten Gradienten



(a) \mathbf{r} ist Richtung zwischen nördlichem und südöstlichem Minimum.



(b) \mathbf{r} ist Richtung zwischen nördlichem und mittlerem Minimum.

Abbildung 3.6: String-Methode zur RGF-Bestimmung auf dem Müller-Brown-Potential

Reaktionspfade liefert, wenn man die Suchrichtung \mathbf{r} geschickt wählt. Das im Folgenden vorgestellte Tangent-Search-Concept (TASC) umgeht dieses Problem, denn es basiert zwar auf der RGF-Methode, konvergiert aber gegen die Talgrundkurve [35] und bedient somit die Forderungen des MEP-Konzeptes.

Der grundlegende Gedanke des RGF-Verfahrens bestand in der konstanten Suchrichtung und damit in der Konstanz des Projektionsoperators $\mathbf{P}_{\mathbf{r}}$ (oder $\overline{\mathbf{P}}_{\mathbf{r}}$). Im TASC-Verfahren wollen wir die Tangente der gesuchten Kurve als neue Gradientensuchrichtung nutzen. Das heißt, wir führen die Predictor-Corrector-Methode zur RGF-Bestimmung aus und nutzen in jedem Iterationsschritt die Tangente aus dem vorherigen Schritt als Suchrichtung¹⁹:

$$\mathbf{0} = \mathbf{P}_{\dot{\mathbf{x}}^{i-1}/\|\dot{\mathbf{x}}^{i-1}\|_2}(\mathbf{g}(\mathbf{x}^i)) . \quad (3.16)$$

Beim Übergang zu infinitesimalen Schrittweiten erhalten wir daraus mithilfe der Branin-gleichung:

$$\mathbf{0} = \mathbf{P}_{\dot{\mathbf{x}}(\mathbf{g})} \stackrel{(3.6)}{=} \mathbf{P}_{\mathbf{A}\mathbf{g}/\|\mathbf{A}\mathbf{g}\|_2}(\mathbf{g}) = \mathbf{g} - \mathbf{A}\mathbf{g} \frac{\mathbf{g}^T \mathbf{A}\mathbf{g}}{\|\mathbf{A}\mathbf{g}\|_2^2} . \quad (3.17)$$

¹⁹Üblicherweise verlassen wir dabei in einem Iterationsschritt die bisherige RGF-Kurve zugunsten einer neuen.

Punkte, die dieser Gleichung genügen, liegen aber gerade auf Gradient Extremals, denn multiplizieren wir beide Seiten von (3.17) mit \mathbf{H} , so erhalten wir

$$\begin{aligned}\mathbf{H}\mathbf{g} &= \mathbf{H}\mathbf{A}\mathbf{g} \frac{\mathbf{g}^T \mathbf{A}\mathbf{g}}{\|\mathbf{A}\mathbf{g}\|_2^2} = \det(\mathbf{H})\mathbf{I}_n \mathbf{g} \frac{\mathbf{g}^T \mathbf{A}\mathbf{g}}{\|\mathbf{A}\mathbf{g}\|_2^2} = \mathbf{g} \frac{\mathbf{g}^T \mathbf{A} \det(\mathbf{H})\mathbf{I}_n \mathbf{g}}{\|\mathbf{A}\mathbf{g}\|_2^2} = \mathbf{g} \frac{\mathbf{g}^T \mathbf{A}\mathbf{H}\mathbf{A}\mathbf{g}}{\|\mathbf{A}\mathbf{g}\|_2^2} \\ &= \mathbf{g} \frac{(\mathbf{A}\mathbf{g})^T \mathbf{H}(\mathbf{A}\mathbf{g})}{\|\mathbf{A}\mathbf{g}\|_2^2} = \lambda \mathbf{g}\end{aligned}$$

für $\lambda := (\mathbf{A}\mathbf{g})^T \mathbf{H}(\mathbf{A}\mathbf{g}) / \|\mathbf{A}\mathbf{g}\|_2^2$.

Wir wollen nun die Wirkung des TASC-Verfahrens (3.16) im Falle realer (also nichtinfinitesimaler) Schrittweiten beschreiben. Seien hierzu für $i = 1, \dots, n$ die Eigenvektoren von \mathbf{H} mit \mathbf{u}^i bezeichnet und die zugehörigen Eigenwerte mit λ_i . Aus den Beziehungen zwischen Hessematrix und ihrer Adjunkten geht hervor, daß \mathbf{A} dieselben Eigenvektoren wie \mathbf{H} hat, allerdings zu den Eigenwerten $\mu_i = \prod_{j \neq i} \lambda_j$. Die einzelnen Punkte der TASC-Kurve werden mithilfe des RGF-Verfahrens berechnet. Angenommen, wir befinden uns im Punkt \mathbf{x} der TASC-Kurve, dessen Gradient in Richtung \mathbf{r} zeigt. Ziel ist die Bestimmung der nächsten Suchrichtung, also gemäß (3.16) die Berechnung von $\dot{\mathbf{x}}$. Schreiben wir \mathbf{r} als Linearkombination der Eigenvektoren gemäß $\mathbf{r} = \sum_{i=1}^n \xi_i \mathbf{u}^i$, so gilt für die Tangentenrichtung:

$$\dot{\mathbf{x}} \stackrel{(3.6)}{=} \pm \mathbf{A}\mathbf{g} \stackrel{(3.9)}{=} c \cdot \mathbf{A}\mathbf{r} = c \sum_{i=1}^n \xi_i \mu_i \mathbf{u}^i = c \sum_{i=1}^n \left[\xi_i \left(\prod_{j \neq i} \lambda_j \right) \mathbf{u}^i \right].$$

Der Vorfaktor c ist wegen der anschließenden Normierung von $\dot{\mathbf{x}}$ zu vernachlässigen. Gelten $|\lambda_1| \leq |\lambda_i|$ für $i = 2, \dots, n$ und $\xi_1 \neq 0$, so rückt $\dot{\mathbf{x}}$ näher an die Richtung von \mathbf{u}^1 heran. Mit anderen Worten: Die in den Iterationsschritten (3.16) ermittelten Suchrichtungen konvergieren gegen die Richtung des Eigenvektors zum betragskleinsten Eigenwert; die TASC-Kurve nähert sich dem Talgrund an. Je größer die Differenzen $|\lambda_1 - \lambda_i|$ ausfallen, desto schneller tritt Konvergenz ein.

Ein Rückblick auf Abbildung 2.5 (Seite 37) zeigt, daß Täler irgendwo auf der PES enden können, ohne auf einen Sattelpunkt zu treffen. In einem solchen Fall sucht der TASC-Algorithmus durch eine Folge von Corrector-Schritten nach einem „benachbarten“ Talgrund [42].

Ein unlösbares Problem für das TASC-Verfahren ist die Bestimmung sogenannter „Don Quixote“-Sattelpunkte, in denen der negative Eigenwert der Hessematrix nicht den kleinsten Betrag hat. Trifft TASC in der Umgebung eines solchen Sattelpunktes auf die Stelle, an der ein positiver Eigenwert von \mathbf{H} kleiner wird als der Betrag des negativen Eigenwertes („Breakout Point“), so bricht das Verfahren die Sattelpunktsuche ab und sucht

3. Die Methode des reduzierten Gradienten

sich einen neuen Talgrund [42]. Indem man am Breakout Point mit dem gewöhnlichen RGF-Verfahren fortsetzt, sind „Don Quixote“-Sattelpunkte trotzdem ohne Probleme zu bestimmen. Mithin ist eine kombinierte Verwendung von RGF und TASC zur Wegeverfolgung zu empfehlen.

Beispiel

Das Modellpotential (A.7) hat einen „Don Quixote“-Sattelpunkt in $(0, 10)$ und ein Minimum im Nullpunkt [42]. Wegen (3.17) erhalten wir für die numerische Bestimmung der TASC-Kurve mittels String-Methode folgende Vorschrift: Iteriere für eine gewisse Schrittweite α die Startkette $\{\mathbf{x}^i\}_{0 \leq i \leq k}$ solange gemäß

$$\tilde{\mathbf{x}}^i = \mathbf{x}^i - \alpha \cdot \mathbf{P}_{\mathbf{A}(\mathbf{x}^i)\mathbf{g}(\mathbf{x}^i)/\|\mathbf{A}(\mathbf{x}^i)\mathbf{g}(\mathbf{x}^i)\|_2}(\mathbf{g}(\mathbf{x}^i))$$

bis Konvergenz eintritt [37]. Abbildung 3.7 zeigt die Anwendung dieser String-Methode für drei verschiedene Startketten. Die durchgehenden Kurven sind Gradient Extremals. Die drei Startketten verlaufen jeweils vom Minimum zu einem variablen Endpunkt. In Abbildung 3.7(a) stimmen End- und Sattelpunkt überein. Als TASC-Kurve ergibt sich in diesem Fall die direkte Verbindung zwischen Minimum und Sattelpunkt. Variieren wir jedoch das Kettenende minimal zu $(10^{-4}, 10)$, so zerreißt die TASC-Kurve und bricht vom Minimum ausgehend in das rechte Nachbaratal aus, wie Abbildung 3.7(b) belegt. Wir sehen, daß dieser Ausbruch nicht am GE-Bifurkationspunkt stattfindet. Grund hierfür ist die Definition des TASC-Verfahrens, daß in jedem Punkt eine (neue) Newtontrajektorie ermittelt. Abschließend zeigt Abbildung 3.7(c) die TASC-Kurve zwischen Minimum und dem variablen Endpunkt $(5.5, 10)$.

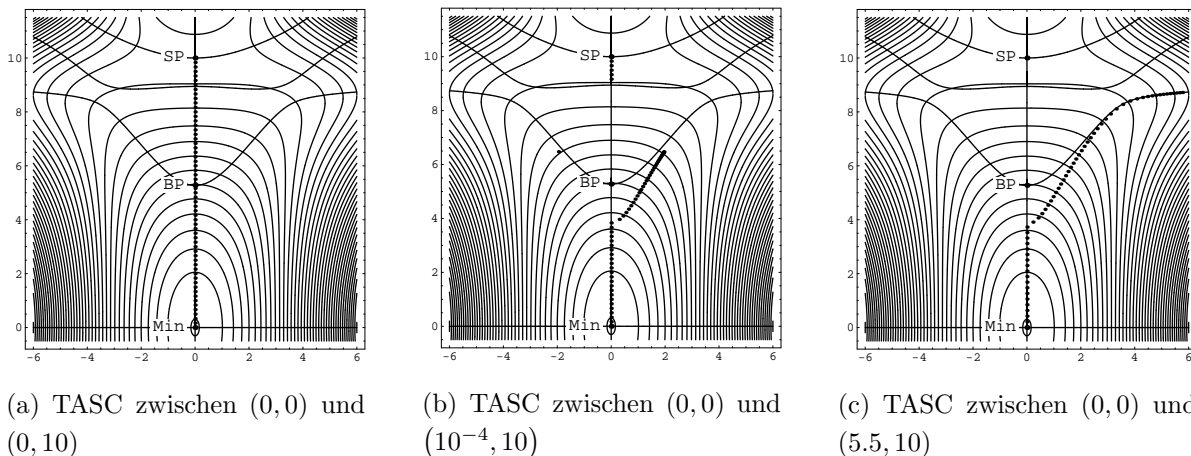


Abbildung 3.7: TASC-Kurven an einem „Don-Quixote“-Sattelpunkt

4. Variationsmethode

Nachdem nun mit der IRC, den Gradient Extremals und dem RGF-Verfahren einige bedeutende Reaktionswegmodelle vorgestellt wurden, wollen wir uns in diesem Abschnitt mit einem Variationsansatz beschäftigen. Mit seiner Hilfe lassen sich nicht nur Reaktionswege modellieren, sondern auch VRI-Punkte bestimmen, wie wir im weiteren Verlauf sehen werden. Ohne es bisher so bezeichnet zu haben, liegt bereits der in Abschnitt 2.1.2 eingeführten und seitdem mehrfach genutzten String-Methode ein Variationsansatz zugrunde.

4.1. Einführung in die Variationsrechnung

Zunächst wollen wir jedoch einige für unsere Bedürfnisse relevante Begriffe der Variationsrechnung einführen. Die Basis für diesen Abschnitt liefern [14] und [30]. Wir beginnen mit der Definition eines Variationsproblems und werden anschließend notwendige und hinreichende Bedingungen für dessen Lösung angeben. Da im Rahmen dieser Arbeit lediglich ein Einblick in das Themengebiet gewährt werden kann, sei für ausführlichere Erläuterungen und Beweise auf die angegebenen Referenzen verwiesen.

Definition

Für $t \in \mathbb{R}$ und $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ sei $F(t, \mathbf{x}, \mathbf{y})$ eine Funktion mit stetigen ersten und zweiten partiellen Ableitungen bezüglich ihrer Argumente. Weiter sei $\Omega \subset [t_0, t_1] \times \mathbb{R}^n$ ein einfach zusammenhängendes, offenes Gebiet, das die Punkte (t_0, \mathbf{x}^0) und (t_1, \mathbf{x}^1) beinhalte. Dann besteht

$$\Gamma := \left\{ \gamma : t \mapsto \mathbf{x}(t) \in \Omega \mid \mathbf{x} \in C^1[t_0, t_1], \mathbf{x}(t_0) = \mathbf{x}^0, \mathbf{x}(t_1) = \mathbf{x}^1 \right\}$$

aus allen stetig differenzierbaren Pfaden zwischen (t_0, \mathbf{x}^0) und (t_1, \mathbf{x}^1) . Wir nennen \mathbf{x} die zu $\gamma : t \mapsto \mathbf{x}(t)$ gehörige *Vektordarstellung*. Unter allen Trajektorien $\gamma \in \Gamma$ suchen wir diejenige, für die das Funktional

$$I(\gamma) := I(\mathbf{x}) = \int_{t_0}^{t_1} F(t, \mathbf{x}(t), \dot{\mathbf{x}}(t)) dt \quad (4.1)$$

einen Minimalwert annimmt. Hierbei bezeichnet $\dot{\mathbf{x}}$ die Ableitung von \mathbf{x} nach t . Die so formulierte Aufgabe nennen wir *Variationsproblem*. Zu $\gamma \in \Gamma$ gehörige Vektordarstellungen $\mathbf{x}(t)$ heißen *zulässig*.

Definition

Wir definieren für $\mathbf{x} \in C^1[t_0, t_1]$ die beiden Normen

$$\|\mathbf{x}\|_0 := \sup_{t \in [t_1, t_0]} |\mathbf{x}(t)|, \quad \|\mathbf{x}\|_1 := \sup_{t \in [t_1, t_0]} |\mathbf{x}(t)| + \sup_{t \in [t_1, t_0]} |\dot{\mathbf{x}}(t)|.$$

4. Variationsmethode

Dann heißt $\gamma^* \in \Gamma$ mit zugehöriger Vektordarstellung \mathbf{x}^*

- (i) *schwaches Minimum* von (4.1), falls ein $\varepsilon > 0$ existiert, sodaß für alle zulässigen \mathbf{x} mit $\|\mathbf{x}^* - \mathbf{x}\|_1 < \varepsilon$ die Beziehung $I(\mathbf{x}^*) \leq I(\mathbf{x})$ gilt.
- (ii) *starkes Minimum* von (4.1), falls ein $\varepsilon > 0$ existiert, sodaß für alle zulässigen \mathbf{x} mit $\|\mathbf{x}^* - \mathbf{x}\|_0 < \varepsilon$ die Beziehung $I(\mathbf{x}^*) \leq I(\mathbf{x})$ gilt.

Bemerkung

Starkes Minimum impliziert schwaches Minimum; die Umkehrung gilt im Allgemeinen nicht.

4.1.1. Hinreichende Bedingung für schwache Minima

Mithilfe zweier notwendiger Kriterien für schwache Minima von (4.1) und einer Aussage über konjugierte Punkte, wollen wir am Ende dieses Abschnitts eine hinreichende Bedingung formulieren. Doch zunächst sei der Begriff der zulässigen Variation definiert.

Definition

Sei $\mathbf{h} \in C^1([t_0, t_1], \mathbb{R}^n)$ und gelte $\mathbf{h}(t_0) = \mathbf{h}(t_1) = \mathbf{0}$, so heißt \mathbf{h} *zulässige Variation* von (4.1), falls es ein $\varepsilon_0 > 0$ gibt, sodaß für alle $\varepsilon \in (-\varepsilon_0, \varepsilon_0)$ und alle zulässigen \mathbf{x} die Funktion $\mathbf{x} + \varepsilon\mathbf{h}$ ebenfalls zulässig ist.

Definition

Für zulässige \mathbf{x} und zulässige Variationen \mathbf{h} ist die Funktion $\phi(\varepsilon) := I(\mathbf{x} + \varepsilon\mathbf{h})$ auf $(-\varepsilon_0, \varepsilon_0)$ wohldefiniert. Wir nennen

$$\begin{aligned} \delta I(\mathbf{h}) &:= \phi'(0) = \left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} \left(\int_{t_0}^{t_1} F(t, \mathbf{x} + \varepsilon\mathbf{h}, \dot{\mathbf{x}} + \varepsilon\dot{\mathbf{h}}) dt \right) \\ &= \int_{t_0}^{t_1} \left([F_{\mathbf{x}}(t, \mathbf{x}, \dot{\mathbf{x}})]^T \mathbf{h} + [F_{\dot{\mathbf{x}}}(t, \mathbf{x}, \dot{\mathbf{x}})]^T \dot{\mathbf{h}} \right) dt \end{aligned} \quad (4.2)$$

*erste Variation des Variationsproblems mit festen Enden.*²⁰

Wir wollen nun notwendige Bedingungen für ein (schwaches) Minimum²¹ von (4.1) finden und gehen hierfür davon aus, daß $I(\mathbf{x}^*)$ minimal ist. Offensichtlich gilt dann $\phi(0) \leq \phi(\varepsilon)$ für alle $\varepsilon \in (-\varepsilon_0, \varepsilon_0)$ und damit $\delta I(\mathbf{h}) = \phi'(0) = 0$. Auf diese Weise erhalten wir aus (4.2) unter Verwendung des Fundamentallemmas der Variationsrechnung²² eine erste notwendige Bedingung:

²⁰Hinweis zur Notation: Es gilt $F_{\mathbf{x}}^T \mathbf{h} = \sum_{i=1}^n F_{x_i} h_i$ (analog für $F_{\dot{\mathbf{x}}} \dot{\mathbf{h}}$).

²¹Da notwendige Bedingungen für schwache Minima auch solche für starke Minima sind, werden wir im Folgenden einfach von Minima ausgehen.

²²Für $f \in C([a, b])$ und $0 = \int_a^b f(x)\eta(x) dx$ mit $\eta \in C^2([a, b])$ und $\eta(a) = \eta(b) = 0$ gilt $f(x) = 0$.

Lemma (Euler-Gleichungen)

Ist $\gamma^* \in \Gamma$ mit zugehörigem \mathbf{x}^* ein Minimum von (4.1), so sind die sogenannten *Euler-Gleichungen*

$$0 = F_{x_i^*} - \frac{dF_{\dot{x}_i^*}}{dt} \quad (4.3)$$

für $i = 1, \dots, n$ erfüllt. Kurven $\gamma^* \in \Gamma$, für die die Euler-Gleichungen gelten, nennen wir auch *Extremale*.

Alternativ stellt die Gültigkeit der Legendre-Bedingung ein notwendiges Kriterium für die Minimalität von (4.1) dar. Hierfür benötigen wir die zweite Variation des Problems.

Definition

Wir nennen $\delta^2 I(\mathbf{h}) := \phi''(0)$ *zweite Variation des Variationsproblems mit festen Enden*. Für sie gilt:

$$\begin{aligned} \delta^2 I(\mathbf{h}) &= \frac{d}{d\varepsilon} \Big|_{\varepsilon=0} \int_{t_0}^{t_1} \left(\left[F_{\mathbf{x}}(t, \mathbf{x} + \varepsilon \mathbf{h}, \dot{\mathbf{x}} + \varepsilon \dot{\mathbf{h}}) \right]^T \mathbf{h} + \left[F_{\dot{\mathbf{x}}}(t, \mathbf{x} + \varepsilon \mathbf{h}, \dot{\mathbf{x}} + \varepsilon \dot{\mathbf{h}}) \right]^T \dot{\mathbf{h}} \right) dt \\ &= \int_{t_0}^{t_1} \left(\left[F_{\mathbf{xx}}(t, \mathbf{x}, \dot{\mathbf{x}}) \mathbf{h} \right]^T \mathbf{h} + 2 \left[F_{\mathbf{x}\dot{\mathbf{x}}}(t, \mathbf{x}, \dot{\mathbf{x}}) \mathbf{h} \right]^T \dot{\mathbf{h}} + \left[F_{\dot{\mathbf{x}}\dot{\mathbf{x}}}(t, \mathbf{x}, \dot{\mathbf{x}}) \dot{\mathbf{h}} \right]^T \dot{\mathbf{h}} \right) dt \\ &= \int_{t_0}^{t_1} \left(\left[F_{\dot{\mathbf{x}}\dot{\mathbf{x}}} \dot{\mathbf{h}} \right]^T \dot{\mathbf{h}} + \left[\left(F_{\mathbf{xx}} - \frac{dF_{\mathbf{x}\dot{\mathbf{x}}}}{dt} \right) \mathbf{h} \right]^T \mathbf{h} \right) dt, \end{aligned} \quad (4.4)$$

wobei im letzten Schritt partiell integriert wurde.²³

Damit \mathbf{x}^* das Problem (4.1) minimiert, muß $\delta^2 I(\mathbf{h}) \geq 0$ gelten. Die Nichtnegativität von (4.4) impliziert jedoch die Gültigkeit von $F_{\dot{\mathbf{x}}^* \dot{\mathbf{x}}^*} \geq 0$. Daraus formulieren wir die Legendre-Bedingung:

Lemma (Legendre-Bedingung)

Ist $\gamma^* \in \Gamma$ ein Minimum von (4.1), dann gilt die *Legendre-Bedingung* $F_{\dot{\mathbf{x}}^* \dot{\mathbf{x}}^*} \geq 0$.

Bemerkung

Legendre versuchte zu beweisen, daß aus der Gültigkeit der *verschärften Legendre-Bedingung*,

$$F_{\dot{\mathbf{x}}^* \dot{\mathbf{x}}^*} > 0 \quad (4.5)$$

für alle Punkte der Kurve γ^* , die schwache Minimalität von γ^* folgt. Er konnte später jedoch selbst zeigen, daß dies kein hinreichendes Kriterium darstellt.

²³Hinweis zur Notation: Es gilt $\left[F_{\dot{\mathbf{x}}\dot{\mathbf{x}}} \dot{\mathbf{h}} \right]^T \dot{\mathbf{h}} = \sum_{i,j=1}^n F_{\dot{x}_i \dot{x}_j} \dot{h}_i \dot{h}_j$ (analog in den restlichen Termen).

4. Variationsmethode

Bevor wir ein hinreichendes Kriterium für schwache Minima von (4.1) formulieren können, benötigen wir noch ein wenig Wissen über konjugierte Punkte:

Definition

Wir betrachten eine Extremale $\gamma : t \mapsto \mathbf{x}(t)$ für das Problem (4.1) sowie eine Schar von benachbarten Extremalen $\gamma^* : t \mapsto \mathbf{x}^*(t)$, die ebenfalls in \mathbf{x}^0 starten. Dann nennen wir den Punkt $(t_{\text{cp}}, \mathbf{x}(t_{\text{cp}})) \in \Omega$ einen zu (t_0, \mathbf{x}^0) *konjugierten Punkt*, wenn $\mathbf{x}(t_{\text{cp}})$ für $\|\mathbf{x}^*(t) - \mathbf{x}(t)\|_1 \rightarrow 0$ der Grenzwert der Schnittpunkte zwischen γ und γ^* ist. Wir sagen in dieser Situation auch: t_{cp} ist konjugiert zu t_0 .

Beispiel

Setzen wir voraus, daß Newtontrajektorien Extremale eines gewissen Variationsansatzes sind (wir werden das in Abschnitt 4.2.3 zeigen), dann sind adjazente stationäre Punkte für dieses Variationsproblem zueinander konjugiert, wie Abbildung 3.1(b) verdeutlicht.

Lemma (über konjugierte Punkte)

Für positiv definites $F_{\dot{\mathbf{x}}\dot{\mathbf{x}}}$ gilt: Die gemäß (4.4) definierte zweite Variation $\delta^2(\mathbf{h})$ ist genau dann für alle zulässigen Variationen \mathbf{h} positiv definit, wenn das Intervall $[t_0, t_1]$ keine zu t_0 konjugierten Punkte enthält.

Damit haben wir genügend Informationen gesammelt, um ein hinreichendes Kriterium für schwache Minima von (4.1) aufzustellen.

Satz (Hinreichende Bedingung für schwache Minima)

Eine für das Variationsproblem (4.1) zulässige Kurve $\gamma^* : t \mapsto \mathbf{x}^*(t)$ erfülle die folgenden Bedingungen:

- (i) γ^* ist Extremale, genügt also den Euler-Gleichungen (4.3).
- (ii) Entlang der Kurve γ^* gilt die verschärfte Legendre-Bedingung (4.5).
- (iii) γ^* enthält keine zu (t_0, \mathbf{x}^0) konjugierten Punkte.

Dann ist γ^* schwache Minimale für das Problem (4.1).

4.1.2. Hinreichende Bedingung für starke Minima

Im vorigen Abschnitt hatten wir konjugierte Punkte mithilfe benachbarter Extremale definiert. Die Untersuchung von Extremalen-Familien wird hier nun von entscheidender Bedeutung sein, um ein hinreichendes Kriterium zu entwickeln. Doch zunächst beschäftigen wir uns mit dem eng verwandten Begriff des Feldes.

Definition

Das in einer weiten Umgebung²⁴ U einer Extremalen definierte Vektorfeld

$$\dot{\mathbf{x}} = \boldsymbol{\psi}(t, \mathbf{x}) \quad (4.6)$$

mit $\boldsymbol{\psi} \in C^1(\Omega)$ nennen wir *Extremalenfeld* von Ω , wenn jede Lösung $\mathbf{x}(t)$ von (4.6) auch Lösung der Euler-Gleichungen (4.3) ist.

Bemerkungen

- (i) Ein Extremalenfeld in Ω ist eine Familie von Extremalen, von denen durch jeden vorgegebenen Punkt von Ω genau eine Extremale geht [6].
- (ii) Zu einer Extremalen γ^* von (4.1), die zudem Trajektorie des gemäß (4.6) definierten Feldes ist, sagen wir auch: γ^* läßt sich in das Feld (4.6) *einbetten*.

Lemma

Enthalte eine Extremale $\gamma^* : t \mapsto \mathbf{x}^*(t)$ von (4.1) keine zu (t_0, \mathbf{x}^0) konjugierten Punkte und gelte zudem $\det F_{\dot{\mathbf{x}}^* \dot{\mathbf{x}}^*} \neq 0$ in $[t_0, t_1]$. Dann läßt sich $\mathbf{x}(t)$ in ein Extremalenfeld von (4.1) einbetten.

Erinnern wir uns an die hinreichende Bedingung für schwache Minima (Seite 58), so sehen wir, daß die Kriterien (ii) und (iii) die Einbettung der gegebenen Extremalen $\mathbf{x}(t)$ implizieren. Mithilfe der Weierstraßschen E -Funktion können wir dann eine hinreichende Bedingung formulieren.

Satz (Hinreichende Bedingung für starke Minima)

Sei γ^* eine Extremale von (4.1), die in ein Extremalenfeld von Ω eingebettet werden kann. Gilt dann für die *Weierstraßsche E -Funktion*,

$$E(t, \mathbf{x}, \boldsymbol{\psi}, \mathbf{w}) := F(t, \mathbf{x}, \mathbf{w}) - F(t, \mathbf{x}, \boldsymbol{\psi}) - (\mathbf{w} - \boldsymbol{\psi})^T F_{\dot{\mathbf{x}}}(t, \mathbf{x}, \boldsymbol{\psi}), \quad (4.7)$$

die Beziehung $E(t, \mathbf{x}, \boldsymbol{\psi}, \mathbf{w}) \geq 0$ für alle Punkte $(t, \mathbf{x}) \in \Omega$ und alle finiten Vektoren \mathbf{w} , dann ist γ^* starke Extremale des Variationsproblems (4.1).

Bemerkung

Die Bedingung $E(t, \mathbf{x}, \boldsymbol{\psi}, \mathbf{w}) \geq 0$ für alle Punkte $(t, \mathbf{x}) \in \Omega$ und alle finiten Vektoren \mathbf{w} heißt auch *Weierstraßsche notwendige Bedingung*.

²⁴Seien $\gamma^* : t \mapsto \mathbf{x}^*(t)$ Extremale und $\gamma : t \mapsto \mathbf{x}(t)$. Dann definiert $U := U_\varepsilon = \{\gamma \in \Gamma : \|\mathbf{x}^* - \mathbf{x}\|_0 \leq \varepsilon\}$ eine weite Umgebung von γ^* . Verwendet man in der Definition von U die Norm $\|\cdot\|_1$, so sprechen wir von einer engen Umgebung.

4.2. Variationsansatz für Reaktionswege

Wir wollen an dieser Stelle das gesammelte Wissen zur Variationsrechnung auf die bereits vorgestellten Reaktionswegmodelle anwenden. Dabei wird sich in Abschnitt 4.2.3 zeigen, daß der Variationsansatz für Newtontrajektorien ein probates Mittel zur Bestimmung von VRI-Punkten darstellt, doch beginnen möchten wir mit dem einfachsten MEP-Konzept, der intrinsischen Reaktionskoordinate IRC.

4.2.1. Die intrinsische Reaktionskoordinate IRC

In Abschnitt 2.1.1 hatten wir die IRC nach Fukui [13] definiert als die Vereinigung der beiden Kurven des steilsten Abstiegs, die zwei adjazente Minima über einen gemeinsamen Sattelpunkt miteinander verbinden. Wegen ihrer Simplität ist die IRC trotz einiger Nachteile das bedeutendste Reaktionswegmodell. Ihre Bestimmung mittels Variationsansatz wurde von Crehuet und Bofill [8] bereits detailliert untersucht.

Zunächst möchten wir jedoch das Konzept der Catchment-Region vorstellen [29]. Als *Catchment-Region* $CR_{\mathbf{x}^*}$ eines vorgegebenen stationären Punktes \mathbf{x}^* bezeichnen wir die offene Menge aller Punkte $\mathbf{x} \in K$ des Konfigurationsraums, für die es eine Kurve des steilsten Abstiegs zu \mathbf{x}^* gibt. Auf diese Weise erhalten wir eine Partition des Konfigurationsraums K . Für gewöhnlich trennt ein $(n - 1)$ -dimensionales Talrückensystem die Catchment-Regionen zweier adjazenter Minima. Am unteren Ende eines jeden Rückens befindet sich dann ein Sattelpunkt [39].

Kurven des steilsten Abstiegs sind gemäß (2.2) definiert als Trajektorien des Differentialgleichungssystems

$$\dot{\mathbf{x}}(t) = -\mathbf{g}(\mathbf{x}(t)) . \quad (4.8)$$

Das zugehörige Variationsproblem sei wie folgt formuliert [8, 39]:

$$I_{\text{SD}}(\mathbf{x}) = \int_{t_0}^{t_1} \sqrt{[\mathbf{g}(\mathbf{x}(t))]^T \mathbf{g}(\mathbf{x}(t))} \sqrt{[\dot{\mathbf{x}}(t)]^T \dot{\mathbf{x}}(t)} dt =: \int_{t_0}^{t_1} F(t, \mathbf{x}, \dot{\mathbf{x}}) dt . \quad (4.9)$$

Wir wollen annehmen, daß es sich bei $\mathbf{x}^0 = \mathbf{x}(t_0)$ um ein Minimum auf der PES handelt und bei $\mathbf{x}^1 = \mathbf{x}(t_1)$ um einen variablen Punkt aus dem Abschluß von $CR_{\mathbf{x}^0}$, der Catchment-Region von \mathbf{x}^0 . Es ist schnell nachgerechnet, daß die durch (4.8) definierten

Kurven Extremalen von (4.9) sind, also den Euler-Gleichungen (4.3) genügen:

$$F_{x_i} = \frac{1}{\sqrt{\mathbf{g}^T \mathbf{g}}} \left(\sum_{j=1}^n g_j \frac{\partial g_j}{\partial x_i} \right) \sqrt{\dot{\mathbf{x}}^T \dot{\mathbf{x}}} \stackrel{(4.8)}{=} \sum_{j=1}^n g_j \frac{\partial g_j}{\partial x_i} = \sum_{j=1}^n H_{ji} g_j ,$$

$$-\frac{dF_{\dot{x}_i}}{dt} = -\frac{d}{dt} \left(\sqrt{\mathbf{g}^T \mathbf{g}} \frac{\dot{x}_i}{\sqrt{\dot{\mathbf{x}}^T \dot{\mathbf{x}}}} \right) \stackrel{(4.8)}{=} \frac{dg_i}{dt} = \sum_{j=1}^n \frac{\partial g_i}{\partial x_j} \dot{x}_j \stackrel{(4.8)}{=} - \sum_{j=1}^n H_{ij} g_j .$$

Die Gleichheit folgt aus der Symmetrie der Hessematrix. Demzufolge läßt sich jede SD-Kurve zwischen \mathbf{x}^0 und \mathbf{x}^1 in ein Extremalenfeld (hier: Gradientenfeld) von $\overline{\text{CR}}_{\mathbf{x}^0}$ einbetten.

Wegen der Eindeutigkeit des Gradienten können sich Kurven des steilsten Abstiegs nur in stationären Punkten schneiden. Mögliche konjugierte Punkte von \mathbf{x}^0 sind also Maxima und Sattelpunkte von beliebigem Index. Crehuet und Bofill rechneten jedoch nach, daß Sattelpunkte vom Index 1 keine zu \mathbf{x}^0 konjugierten Punkte sind [8]. Insbesondere gibt es dann zwischen \mathbf{x}^0 und einem solchen Sattelpunkt vom Index 1 nur genau einen steilsten Abstiegs Pfad. Desweiteren zeigten sie, daß von \mathbf{x}^0 ausgehende SD-Kurven auch die Weierstraßsche notwendige Bedingung in der Catchment-Region von \mathbf{x}^0 erfüllen.

Zwischen \mathbf{x}^0 und $\mathbf{x}^1 \in \overline{\text{CR}}_{\mathbf{x}^0}$ ist die eindeutig gegebene SD-Kurve γ_{SD} demnach eine starke Minimale von (4.9). Sei nun $\bar{\mathbf{x}}^0$ ein zu \mathbf{x}^0 adjazentes Minimum, \mathbf{x}^1 ein Sattelpunkt erster Ordnung, der auf dem gemeinsamen Rand der Abschlüsse von $\text{CR}_{\mathbf{x}^0}$ und $\text{CR}_{\bar{\mathbf{x}}^0}$ liegt und $\bar{\gamma}_{\text{SD}}$ die SD-Kurve zwischen $\bar{\mathbf{x}}^0$ und \mathbf{x}^1 . Dann setzt sich die IRC aus γ_{SD} und $\bar{\gamma}_{\text{SD}}$ zusammen und ist somit gerade die Vereinigung zweier starker Minimaler des gewählten Variationsansatzes.

Beispiel

In Abbildung 4.1(a) trennt der Talrücken $x_2 = 0$ die Catchment-Regionen der beiden Minima voneinander. Die nebenstehende Abbildung zeigt uns mehrere SD-Kurven (grau) im Extremalfeld (4.8). Lediglich die IRC (schwarz) verbindet beide Minima miteinander.

Bemerkungen

- (i) Wir fassen zusammen: Zwischen dem Minimum \mathbf{x}^0 und dem Sattelpunkt \mathbf{x}^1 vom Index 1 liegt wegen der Eindeutigkeit des Gradienten und der Nichtkonjugiertheit von \mathbf{x}^1 zu \mathbf{x}^0 genau eine Kurve des steilsten Abstiegs; und diese ist starke Minimale von (4.9) bezüglich aller in $\overline{\text{CR}}_{\mathbf{x}^0}$ liegenden, stetig differenzierbaren Kurven zwischen \mathbf{x}^0 und \mathbf{x}^1 .
- (ii) Dieses Ergebnis liefert Eindeutigkeit und Konvergenz für die in Abschnitt 2.1.2 eingeführte String-Methode: Unter allen Kurven zwischen zwei adjazenten Minima

4. Variationsmethode

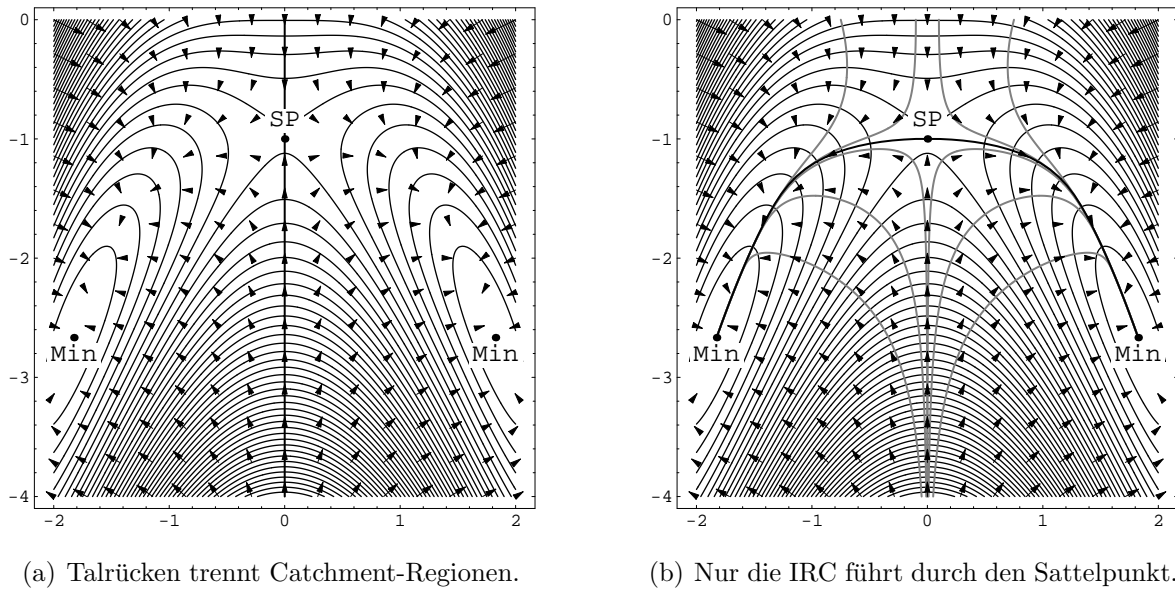


Abbildung 4.1: Kurven des steilsten Abstiegs im Vektorfeld $-\mathbf{g}$

gibt es genau eine durch den Sattelpunkt vom Index 1, die sich aus zwei SD-Kurven zusammensetzt und folglich dem MEP-Konzept genügt.

4.2.2. Gradient Extremals

Das grundlegende Ziel bei der Entwicklung des GE-Konzepts bestand in der Ermittlung von Kurven, die, ausgehend von einem gegebenen Minimum, den „leichtesten“ Anstieg zum Sattelpunkt wählen. In Abschnitt 2.2 nannten wir diese Kurven Talgrundpfade. Sie sollten Punkte \mathbf{x} enthalten, für die das Funktional

$$\sigma(\mathbf{x}) = \frac{\|\mathbf{g}(\mathbf{x})\|_2^2}{2} \quad (4.10)$$

entlang der zu \mathbf{x} gehörigen Äquipotentialfläche ein Minimum annimmt. Anhand des Ansatzes (4.10) stellten wir fest, daß dann entlang der Talgrundpfade der Gradient ein Eigenvektor der Hessematrix ist:

$$\mathbf{H}(\mathbf{x})\mathbf{g}(\mathbf{x}) = \lambda(\mathbf{x})\mathbf{g}(\mathbf{x}) . \quad (4.11)$$

Die durch (4.11) definierten Kurven nannten wir Gradient Extremals. Sie liegen isoliert im Konfigurationsraum und bilden keine Familien von benachbarten Kurven. Sie lassen sich insbesondere nicht durch stetige Variation ineinander überführen. Gradient Extremals ergeben sich damit nicht als Lösung eines zu (4.9) analogen Variationsansatzes.

4.2.3. Das RGF-Verfahren

Gemäß den Überlegungen in Abschnitt 3 faßten wir alle Punkte des Konfigurationsraums, deren Gradient parallel zu einer vorgegebenen Suchrichtung \mathbf{r} liegt, zu einer RGF-Kurve (auch Newtontrajektorie) zusammen. Wir zeigten, daß diese Kurven mit den Trajektorien der Braningleichung (3.6),

$$\dot{\mathbf{x}}(t) = \pm \mathbf{A}(\mathbf{x}(t)) \mathbf{g}(\mathbf{x}(t)) , \quad (4.12)$$

übereinstimmen. In diesem Abschnitt wollen wir einen Variationsansatz für Newtontrajektorien aufstellen und analog zur Behandlung der IRC nachweisen, daß es genau eine RGF-Kurve zwischen gewissen Punkten des Konfigurationsraums gibt, die durch den VRI-Punkt verläuft und dabei Minimale des gewählten Ansatzes ist. Dabei spielen die im Folgenden eingeführten Newtonkanäle in gewisser Weise die Rolle der für SD-Kurven betrachteten Catchment-Regionen.

Wir nannten NT-Zweige, die zwei adjazente stationäre Punkte mit Indizes $i_1 \neq i_2 \pmod{2}$ miteinander verbinden, regulär und solche, die in VRI-Punkten enden, singular. Als *Newtonkanal* $\text{NC}_{\mathbf{x}^*}$ eines stationären Punktes \mathbf{x}^* wollen wir die Menge aller regulären Zweige betrachten, die durch stetige Variation der Suchrichtung ineinander überführt werden können, ohne daß dabei singuläre Zweige auftreten [19]. Damit bilden singuläre Zweige Ränder zwischen zwei Newtonkanälen von \mathbf{x}^* .²⁵

Wir betrachten das Variationsproblem [39]

$$I_{NT}(\mathbf{x}) = \int_{t_0}^{t_1} \|\dot{\mathbf{x}} \mp \mathbf{A}(\mathbf{x}) \mathbf{g}(\mathbf{x})\|_2^2 dt =: \int_{t_0}^{t_1} F(t, \mathbf{x}, \dot{\mathbf{x}}) dt \quad (4.13)$$

für den stationären Punkt $\mathbf{x}^0 = \mathbf{x}(t_0)$ und einen variablen Punkt $\mathbf{x}^1 = \mathbf{x}(t_1)$ aus dem Abschluß eines durch einen singulären Zweig berandeten Newtonkanals von \mathbf{x}^0 . Wir sehen sofort ein, daß dann (4.12) Extremale von (4.13) ist:

$$\begin{aligned} F_{x_i} &= 2 (\dot{\mathbf{x}} \mp \mathbf{A} \mathbf{g})^T \left[\frac{\partial}{\partial x_i} (\dot{\mathbf{x}} \mp \mathbf{A} \mathbf{g}) \right] \stackrel{(4.12)}{=} 0 , \\ \frac{dF_{\dot{x}_i}}{dt} &= \frac{d}{dt} \left\{ 2 (\dot{\mathbf{x}} \mp \mathbf{A} \mathbf{g})^T \left[\frac{\partial}{\partial \dot{x}_i} (\dot{\mathbf{x}} \mp \mathbf{A} \mathbf{g}) \right] \right\} \stackrel{(4.12)}{=} 0 . \end{aligned}$$

Wir können also jede Newtontrajektorie zwischen \mathbf{x}^0 und \mathbf{x}^1 in ein durch (4.12) definiertes Extremalenfeld von $\overline{\text{NC}}_{\mathbf{x}^0}$ einbetten.

²⁵Ein von \mathbf{x}^* ausgehender singulärer Zweig stellt hierbei das Analogon zur SD-Kurve zwischen Minimum und Sattelpunkt vom Index 1 dar.

4. Variationsmethode

Beispiel

In Abbildung 4.2(a) trennt die Newtontrajektorie durch den VRI-Punkt vier Newtonkanäle voneinander. Die nebenstehende Abbildung zeigt uns mehrere Newtontrajektorien (grau) im Extremalfeld (4.12). Lediglich eine (schwarz) verbindet beide Sattelpunkte vom Index 1 miteinander.

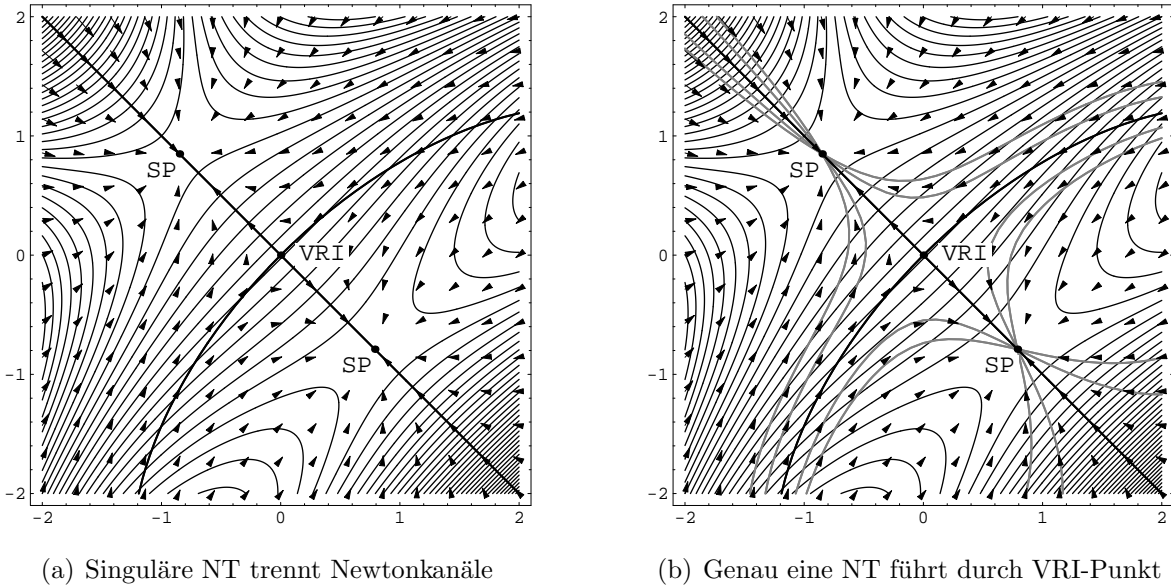


Abbildung 4.2: Newtontrajektorien im Vektorfeld \mathbf{Ag}

Bemerkung

Für den naheliegenderen, da zu (4.9) analogen Variationsansatz

$$I'_{\text{NT}} = \int_{t_0}^{t_1} \sqrt{(\mathbf{Ag})^T (\mathbf{Ag})} \sqrt{\dot{\mathbf{x}}^T \dot{\mathbf{x}}} dt = \int_{t_0}^{t_1} \sqrt{\mathbf{g}^T \mathbf{A}^2 \mathbf{g}} \sqrt{\dot{\mathbf{x}}^T \dot{\mathbf{x}}} dt$$

ist (4.12) nur dann Extremale, wenn $\partial A_{jk} / \partial x_i = \partial A_{ik} / \partial x_j$ erfüllt ist [39]. Diese zusätzliche Symmetriebedingung ist jedoch im Allgemeinen nicht vorauszusetzen.

Wie wir aus Abschnitt 3 wissen, schneiden sich Newtontrajektorien zu verschiedenen Suchrichtungen nur in stationären Punkten. Mithin sind alle zu \mathbf{x}^0 adjazenten stationären Punkte auch zu \mathbf{x}^0 konjugierte Punkte. In Abschnitt 3.3.1 hatten wir dargelegt, daß VRI-Punkte Verzweigungsstellen *einer* RGF-Kurve sind. Sei also \mathbf{x}^1 VRI-Punkt. Dann enthält die Newtontrajektorie zwischen \mathbf{x}^0 und \mathbf{x}^1 keine konjugierten Punkte.

Um zu zeigen, daß Newtontrajektorien starke Minimale von (4.13) sind, bleibt der Nachweis der Weierstraßschen notwendigen Bedingung. Hierfür berechnen wir zunächst

$F_{\dot{\mathbf{x}}}$. Wir lesen direkt aus (4.13) ab:

$$\begin{aligned} F_{\dot{x}_i} &= \frac{\partial}{\partial \dot{x}_i} \sum_{j=1}^n \left([\dot{x}_j \mp (\mathbf{A}\mathbf{g})_j]^2 \right) = 2 \sum_{j=1}^n \left([\dot{x}_j \mp (\mathbf{A}\mathbf{g})_j] \cdot \frac{\partial}{\partial \dot{x}_i} [\dot{x}_j \mp (\mathbf{A}\mathbf{g})_j] \right) \\ &= 2 \sum_{j=1}^n \left([\dot{x}_j \mp (\mathbf{A}\mathbf{g})_j] \delta_{ij} \right) = 2 [\dot{x}_i \mp (\mathbf{A}\mathbf{g})_i]. \end{aligned}$$

Somit erhalten wir $F_{\dot{\mathbf{x}}} = 2(\dot{\mathbf{x}} \mp \mathbf{A}\mathbf{g})$. Gemäß (4.12) ist ein Extremalenfeld von $\overline{\text{NC}}_{\mathbf{x}^0}$ durch $\boldsymbol{\psi}(t, \mathbf{x}) = \pm \mathbf{A}(\mathbf{x})\mathbf{g}(\mathbf{x})$ gegeben. Dann gilt in $\overline{\text{NC}}_{\mathbf{x}^0}$

$$\begin{aligned} E(t, \mathbf{x}, \boldsymbol{\psi}, \mathbf{w}) &= F(t, \mathbf{x}, \mathbf{w}) - F(t, \mathbf{x}, \boldsymbol{\psi}) - (\mathbf{w} - \boldsymbol{\psi})^T F_{\dot{\mathbf{x}}}(t, \mathbf{x}, \boldsymbol{\psi}) \\ &= \|\dot{\mathbf{w}} \mp \mathbf{A}\mathbf{g}\|_2^2 - \underbrace{\|\boldsymbol{\psi} \mp \mathbf{A}\mathbf{g}\|_2^2}_{=0} - 2(\mathbf{w} - \boldsymbol{\psi})^T \underbrace{(\boldsymbol{\psi} \mp \mathbf{A}\mathbf{g})}_{=0} \\ &= \|\dot{\mathbf{w}} \mp \mathbf{A}\mathbf{g}\|_2^2 \geq 0 \end{aligned}$$

für finite Vektoren \mathbf{w} . Demnach sind die RGF-Kurven zwischen \mathbf{x}^0 und einem variablen Endpunkt \mathbf{x}^1 in $\overline{\text{NC}}_{\mathbf{x}^0}$ starke Minimale von (4.13). Insbesondere liegt zwischen \mathbf{x}^0 und dem VRI-Punkt \mathbf{x}^1 genau eine Newtontrajektorie; und diese ist starke Minimale von (4.13) bezüglich aller in $\overline{\text{NC}}_{\mathbf{x}^0}$ liegenden, stetig differenzierbaren Kurven zwischen \mathbf{x}^0 und \mathbf{x}^1 .

Für einen weiteren stationären Punkt $\bar{\mathbf{x}}^0$ liege der VRI-Punkt \mathbf{x}^1 sowohl in $\overline{\text{NC}}_{\mathbf{x}^0}$ als auch in $\overline{\text{NC}}_{\bar{\mathbf{x}}^0}$. Dann sind die Newtontrajektorien zwischen \mathbf{x}^0 und \mathbf{x}^1 sowie zwischen $\bar{\mathbf{x}}^0$ und \mathbf{x}^1 jeweils starke Minimale von (4.13). Da die Suchrichtung in beiden Trajektorien jeweils konstant ist und somit insbesondere parallel zum Gradienten im gemeinsamen Punkt \mathbf{x}^1 , handelt es sich hierbei sogar um zwei Zweige derselben RGF-Kurve. Es gibt also genau eine RGF-Kurve, die \mathbf{x}^0 und $\bar{\mathbf{x}}^0$ über den VRI-Punkt \mathbf{x}^1 miteinander verbindet und dabei (4.13) minimiert.

5. Bestimmung von VRI-Punkten

Wie in Abschnitt 4.2.3 gezeigt, gibt es genau eine Newtontrajektorie, die gewisse (stationäre) Punkte einer PES über einen VRI-Punkt miteinander verbindet und dabei das zum RGF-Verfahren gehörige Variationsproblem minimiert. Diese spezielle NT sei im Folgenden kurz *VRI-NT* genannt. Es ist das Ziel des Abschnittes und der vorliegenden Arbeit, einen Algorithmus aufzustellen, der eine Startsuchrichtung solange variiert bis sie der Suchrichtung der VRI-NT entspricht. Wir wollen zunächst anhand des Modellpotentials (A.1) die Grundidee für einen solchen Algorithmus vorstellen und mit der Aufgabenstellung zusammenhängende Probleme benennen. Erinnerung sei noch einmal an die Beziehung (1.6): Für VRI-Punkte \mathbf{x}^{vri} gilt $\mathbf{A}(\mathbf{x}^{\text{vri}})\mathbf{g}(\mathbf{x}^{\text{vri}}) = \mathbf{0}$ mit $\mathbf{g}(\mathbf{x}^{\text{vri}}) \neq \mathbf{0}$.

5.1. Der Algorithmus

Für die korrekte Arbeitsweise des entstehenden Programms wollen wir voraussetzen, daß es einen VRI-Punkt auf der PES gibt und dieser in den Abschlüssen der Newtonkanäle $\overline{\text{NC}}_{\mathbf{x}^*}$ und $\overline{\text{NC}}_{\bar{\mathbf{x}}^*}$ zweier stationärer Punkte \mathbf{x}^* und $\bar{\mathbf{x}}^*$ liegt. Dann soll das Programm zur VRI-Bestimmung wie folgt vorgehen:

Wir legen eine gerade Kette $\{\mathbf{x}^i\}_{0 \leq i \leq k}$ zwischen $\mathbf{x}^0 \in \text{NC}_{\mathbf{x}^*}$ und $\mathbf{x}^k \in \text{NC}_{\bar{\mathbf{x}}^*}$ und bestimmen die Richtung zwischen \mathbf{x}^0 und \mathbf{x}^k als Startsuchrichtung \mathbf{r} .²⁶ Die beiden Randpunkte \mathbf{x}^0 und \mathbf{x}^k wollen wir im Folgenden als fixiert betrachten. Auf diese Weise stellen wir sicher, daß beide Newtonkanäle stets wenigstens einen Kettenpunkt beinhalten.

Da es sich bei \mathbf{r} typischerweise noch nicht um die Suchrichtung der VRI-NT handelt²⁷, wenden wir die String-Methode für das RGF-Verfahren auf die Kette $\{\mathbf{x}^i\}_{1 \leq i \leq k-1}$ an:

$$\tilde{\mathbf{x}}^i = \mathbf{x}^i - \alpha (\mathbf{I}_n - \mathbf{r}\mathbf{r}^T) \mathbf{g}(\mathbf{x}^i) \quad (5.1)$$

für $1 \leq i \leq k-1$ und eine gewisse Schrittweite α . Leider zieht jedoch die String-Methode häufig sämtliche Kettenpunkte (bis auf den fixierten Randpunkt) in einen der beiden Newtonkanäle, wie Abbildung 5.1 belegt.²⁸ Auf diese Weise erhalten wir Kettenpunkte in großer Entfernung von \mathbf{x}^0 und \mathbf{x}^k . Da wir jedoch den VRI-Punkt zwischen diesen beiden Punkten vermuten, wollen wir unsere Suche nach ihm auf einen gewissen Bereich zwischen \mathbf{x}^0 und \mathbf{x}^k einschränken. Zudem wäre für den weiteren Verlauf des Programms eine größere Anzahl an Kettenpunkten im zweiten Newtonkanal wünschenswert. Daher modifizieren wir die übliche String-Methode gemäß (5.1) wie folgt:

²⁶Typischerweise wählen wir $\mathbf{x}^0 := \mathbf{x}^*$ und $\mathbf{x}^k := \bar{\mathbf{x}}^*$.

²⁷Anderenfalls endet der Algorithmus natürlich an dieser Stelle.

²⁸Die zu \mathbf{r} gehörige Newtontrajektorie ist durchgehend gezeichnet, die mithilfe der String-Methode iterierte Kette gepunktet.

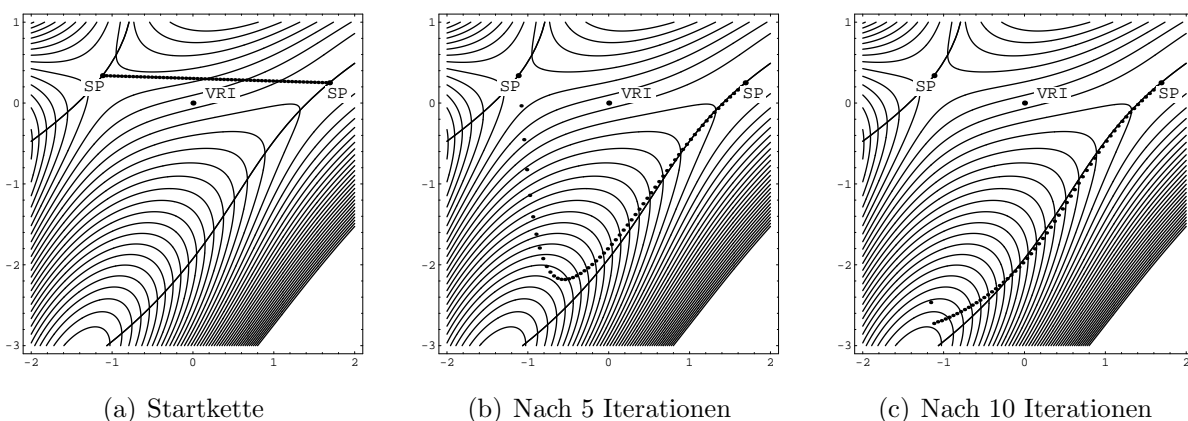


Abbildung 5.1: String-Methode für Newtontrajektorien in VRI-Punkt-Nähe

Wir iterieren den Kettenpunkt \mathbf{x}^i mit $1 \leq i \leq k - 1$ gemäß (5.1) solange bis eines der folgenden Szenarien eintritt:

- (i) Erfüllt nach einer gewissen Zahl an Iterationen der iterierte Kettenpunkt $\tilde{\mathbf{x}}^i$ nicht mehr²⁹ die beiden Ungleichungen

$$\|\tilde{\mathbf{x}}^i - \mathbf{x}^0\|_2 \leq \|\mathbf{x}^0 - \mathbf{x}^k\|_2, \quad (5.2a)$$

$$\|\tilde{\mathbf{x}}^i - \mathbf{x}^k\|_2 \leq \|\mathbf{x}^0 - \mathbf{x}^k\|_2, \quad (5.2b)$$

so setzen wir

- a) $\tilde{\mathbf{x}}^i$ auf den ursprünglichen Kettenpunkt \mathbf{x}^i zurück und führen (5.1) erneut aus, allerdings mit $-\alpha$ statt mit α .
 - b) $\tilde{\mathbf{x}}^i := \tilde{\mathbf{x}}^{i-1}$, falls das Szenario (ia) im Laufe der Iteration von \mathbf{x}^i bereits eingetreten sein sollte und wir mithin das Vorzeichen von α schon einmal gewechselt haben. Dieses Vorgehen ist durch die Fixierung von \mathbf{x}^0 , also die Gültigkeit von $\tilde{\mathbf{x}}^0 \equiv \mathbf{x}^0$ legitimiert. Im Extremfall setzen wir also $\tilde{\mathbf{x}}^i := \mathbf{x}^0$.
- (ii) Gelten für $\tilde{\mathbf{x}}^i$ beide Ungleichungen (5.2) und zudem

$$(\mathbf{L}_n - \mathbf{r}\mathbf{r}^T) \mathbf{g}(\tilde{\mathbf{x}}^i) < \varepsilon \quad (5.3)$$

für eine vorgegebene Genauigkeit ε , dann liegt $\tilde{\mathbf{x}}^i$ innerhalb des durch (5.2) definierten Bereichs auf einer Newtontrajektorie und wir fahren mit der Iteration von \mathbf{x}^{i+1} fort, falls $i < k - 1$ ist.

²⁹Für die ursprüngliche (gerade) Kette $\{\mathbf{x}^i\}_{0 \leq i \leq k}$ sind die Ungleichungen (5.2) stets in jedem Kettenpunkt erfüllt.

5. Bestimmung von VRI-Punkten

Wir wollen zunächst bemerken, daß jeder Kettenpunkt nach einer gewissen Zahl an Iterationen entweder den durch (5.2) definierten Bereich verläßt (ohne die zu \mathbf{r} gehörige RGF-Kurve erreicht zu haben) oder die durch \mathbf{r} bestimmte NT mit hinreichender Genauigkeit erreicht und damit (5.3) erfüllt. Mithilfe des Vorzeichenwechsels in (ia) besteht zudem die Möglichkeit mehrere Kettenpunkte in beiden Newtonkanälen zu plazieren, wie Abbildung 5.2 verdeutlicht.

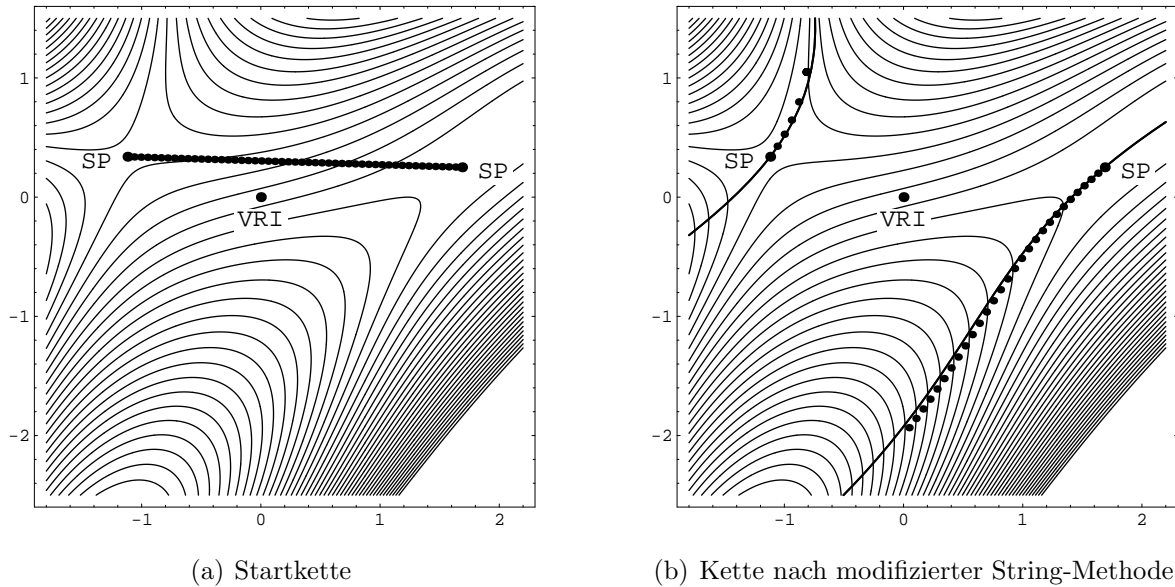


Abbildung 5.2: Modifizierte String-Methode

Damit sind die Voraussetzungen geschaffen, um eine verbesserte Suchrichtung zu finden. Zwischen alle Paare $(\tilde{\mathbf{x}}^i, \tilde{\mathbf{x}}^j)$ iterierter Punkte mit $i, j \in \{0, \dots, k\}$ und $i \neq j$ legen wir gerade Ketten $\{\mathbf{y}_{i,j}^l\}_{0 \leq l \leq k}$ mit $\mathbf{y}_{i,j}^0 := \tilde{\mathbf{x}}^i$ und $\mathbf{y}_{i,j}^k := \tilde{\mathbf{x}}^j$. Anschließend bestimmen wir bezüglich aller dieser Ketten das Minimum

$$\min_{\substack{i,j \in \{0, \dots, k\} \\ i \neq j}} \min_{\substack{0 \leq l \leq k \\ \|\mathbf{g}(\mathbf{y}_{i,j}^l)\| > 0.1 =: \delta}} \|\mathbf{A}(\mathbf{y}_{i,j}^l) \mathbf{g}(\mathbf{y}_{i,j}^l)\|_2. \quad (5.4)$$

Der Punkt, in dem das so ermittelte Minimum (5.4) angenommen wird, sei mit \mathbf{y} bezeichnet.

Bemerkung

Die Minimumbildung (5.4) ist durch die Beziehung (1.6) motiviert. Auf diese Weise suchen wir nach VRI-nahen Punkten. Durch die Bedingung $\|\mathbf{g}(\mathbf{y}_{i,j}^l)\| > 0.1 =: \delta$ ist gewährleistet, daß nicht versehentlich Kettenpunkte in einer kleinen Umgebung eines stationären Punktes als VRI-Punkte betrachtet werden. Den Grenzwert δ verstehen wir als Parameter; er ist dem vorliegenden Problem anzupassen.

Wird in \mathbf{y} das Minimum (5.4) angenommen, so setzen wir $\mathbf{r} := \mathbf{g}(\mathbf{y})$, legen eine gerade Kette der Länge k zwischen \mathbf{x}^0 und \mathbf{x}^k und starten erneut die modifizierte Stringmethode. In Abbildung 5.3(a) sehen wir den so ermittelten VRI-nahen Punkt \mathbf{y} , während Abbildung 5.3(b) die Kette nach erneuter Anwendung der modifizierten String-Methode zeigt.³⁰

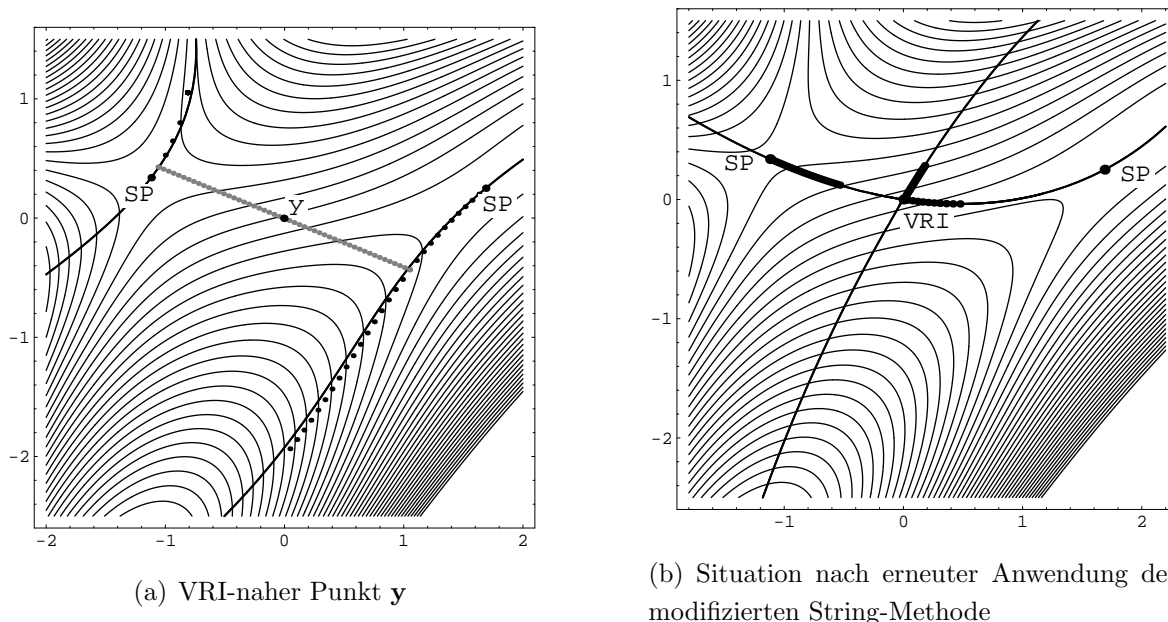


Abbildung 5.3: Etappen des Algorithmus

Aus diesen Überlegungen ergibt sich folgender Algorithmus zur Bestimmung von VRI-Punkten:

Algorithmus

Wir wollen voraussetzen, daß für zwei stationäre Punkte \mathbf{x}^* und $\bar{\mathbf{x}}^*$ ein VRI-Punkt in $\overline{NC_{\mathbf{x}^*}} \cap \overline{NC_{\bar{\mathbf{x}}^*}}$ liegt. Dann soll das Programm zur VRI-Bestimmung wie folgt vorgehen.

- (i) Bestimme zwei Punkte $\mathbf{x}^0 \in NC_{\mathbf{x}^*}$ und $\mathbf{x}^k \in NC_{\bar{\mathbf{x}}^*}$.
- (ii) Bestimme eine Anfangsrichtung \mathbf{r} . Üblicherweise³¹ wählen wir

$$\mathbf{r} := (\mathbf{x}^k - \mathbf{x}^0) / \|\mathbf{x}^k - \mathbf{x}^0\|_2 .$$

- (iii) Ist \mathbf{r} bereits die Suchrichtung der VRI-NT, so breche ab. Anderenfalls lege eine gerade Kette $\{\mathbf{x}^i\}_{0 \leq i \leq k}$ der Länge k zwischen \mathbf{x}^0 und \mathbf{x}^k .

³⁰In diesem Fall ist $\mathbf{y} = (-0.00582909859, -0.00289585264)^T$, während der tatsächliche VRI-Punkt im Ursprung liegt. Zudem erhalten wir $\mathbf{r} := \mathbf{g}(\mathbf{y}) = (-0.250012714, 0.999999888)^T$, wobei die zur VRI-NT gehörige Suchrichtung $(-0.25, 1)^T$ ist.

³¹Häufig bietet sich auch $\mathbf{r} = \mathbf{g}\left(\frac{\mathbf{x}^k + \mathbf{x}^0}{2}\right)$ als Anfangsrichtung an.

5. Bestimmung von VRI-Punkten

- (iv) Führe das oben erläuterte modifizierte Stringverfahren mit einer gewissen Schrittweite α für jeden der inneren Kettenpunkte \mathbf{x}^i mit $1 \leq i \leq k-1$ einzeln aus und erhalte somit iterierte Kettenpunkte $\tilde{\mathbf{x}}^i$, die mit Genauigkeit ε auf der zu \mathbf{r} gehörigen Newtontrajektorie innerhalb des durch (5.2) definierten Gebiets liegen.
- (v) Lege gerade Ketten $\{\mathbf{y}_{i,j}^l\}_{0 \leq l \leq k}$ der Länge k zwischen alle Paare iterierter Punkte $(\tilde{\mathbf{x}}^i, \tilde{\mathbf{x}}^j)$ mit $i, j \in \{0, \dots, k\}$ und $i \neq j$. Wir setzen dabei $\mathbf{y}_{i,j}^0 := \tilde{\mathbf{x}}^i$ und $\mathbf{y}_{i,j}^k := \tilde{\mathbf{x}}^j$.
- (vi) Bestimme das Minimum (5.4) und bezeichne den Punkt, in dem (5.4) angenommen wird, mit \mathbf{y} .
- (vii) Setze $\mathbf{r} := \mathbf{g}(\mathbf{y})$ als neue Suchrichtung fest und gehe zurück zu Schritt (iii).

Bemerkung

Als Abbruchkriterium in Schritt (iii) bietet sich etwa die Differenz der Suchrichtungen zweier Programmdurchläufe an. Alternativ kann man zur Verifizierung des Kriteriums eine der in Abschnitt 3.4 beschriebenen Predictor-Corrector-Methoden mit Suchrichtung \mathbf{r} und Anfangspunkt \mathbf{x}^* durchführen und überprüfen, ob diese einem VRI-Punkt hinreichend nahe kommt. Die in Schritt (vi) ermittelten Punkte \mathbf{y} werden sich dem VRI-Punkt nicht beliebig annähern, da sie als Kettenpunkte nicht zum VRI-Punkt konvergieren, sondern durch die Randpunkte der zugehörigen Ketten determiniert sind. Allerdings erreicht man beispielsweise mit der Erhöhung der Kettenlänge k verbesserte Genauigkeiten für \mathbf{y} .

Der beschriebene Algorithmus wurde in FORTRAN 77 implementiert. Der Quelltext (für die Anwendung auf dreidimensionale Potentialenergieflächen) befindet sich im Anhang C. Wir erhalten die folgenden Ergebnisse:

5.2. Anwendung auf zweidimensionale Modellpotentiale

Wir wollen nun die Wirkung des in Abschnitt 5.1 beschriebenen Algorithmus auf zweidimensionale Potentiale demonstrieren. Zunächst erfolgt die Anwendung auf die PES-Familie [43]

$$E_\mu(x_1, x_2) = \frac{1}{2} (x_1 x_2^2 - x_1^2 x_2 - \mu x_1 + 2x_2) + \frac{1}{30} (x_1^4 + x_2^4) . \quad (5.5)$$

5.2.1. Die PES-Familie (5.5)

Die Potentialenergieflächen der Familie (5.5) haben einen VRI-Punkt in $(0, 0)$; die zugehörige NT-Suchrichtung ist $\mathbf{r}^T = (-\frac{\mu}{2}, 1)$. Als feste Werte für Kettenlänge, Schrittweite und Genauigkeit seien

$$k = 50 , \quad \alpha = 0.125 , \quad \varepsilon = 10^{-8} \quad (5.6)$$

5.2. Anwendung auf zweidimensionale Modellpotentiale

festgelegt. Offensichtlich erhalten wir für $\mu = \frac{1}{2}$ gerade die Fläche (A.1). Die Wirkung des Programms auf dieses Potential läßt sich in Abbildung 5.3(b) ablesen.

In Abbildung 5.4 sehen wir die Anwendung des Algorithmus für verschiedene Belegungen von μ . Die Newtontrajektorie zur vom Programm errechneten VRI-Suchrichtung ist dabei als durchgehende Linie dargestellt. Für $\mu \in \{1.0, 1.5, 2.0, 2.5, 3.0\}$ wählten wir als Startkette die gerade Verbindung zwischen den beiden Sattelpunkten (im Folgenden mit SP_1 und SP_2 bezeichnet) und für $\mu \in \{\pm 5.0, \pm 10.0\}$ die Strecke zwischen den fixierten Punkten \mathbf{a} und \mathbf{b} . Als Anfangssuchrichtung \mathbf{r} diente die Richtung zwischen den jeweiligen Kettenendpunkten. Gezeigt werden die Ergebnisse des Programms nach jeweils fünf Durchläufen. Lediglich im symmetrischen Fall $\mu = 2$ bedurfte es keines Programmdurchlaufs, da hier die gewählte Startsuchrichtung bereits der VRI-NT entspricht. Mithin sehen wir in Abbildung 5.4(c) gerade die Startkette. Tabelle 5.1 faßt die numerischen Ergebnisse des Programms zusammen.

μ	Kettenanfang	Kettenende	VRI-Suchrichtung nach fünf Programmdurchläufen
0.5	SP_1	SP_2	$(-0.250007449, 1.00000258)^T$
1.0	SP_1	SP_2	$(-0.500000003, 1.00000001)^T$
1.5	SP_1	SP_2	$(-0.749999745, 0.999999543)^T$
2.0	SP_1	SP_2	$(-0.99999994, 0.99999994)^T$
2.5	SP_1	SP_2	$(-1.25000003, 1.00000002)^T$
3.0	SP_1	SP_2	$(-1.49999996, 0.999999949)^T$
-10.0	$\mathbf{a} = (-1.5, -1.5)^T$	$\mathbf{b} = (2.0, 0.5)^T$	$(5.00000146, 0.999999859)^T$
-5.0	$\mathbf{a} = (-1.5, -1.5)^T$	$\mathbf{b} = (2.0, 0.5)^T$	$(2.49997929, 1.00001437)^T$
5.0	$\mathbf{a} = (-1.5, 1.0)^T$	$\mathbf{b} = (2.0, 0.5)^T$	$(-2.49964773, 1.00197916)^T$
10.0	$\mathbf{a} = (-1.5, 1.0)^T$	$\mathbf{b} = (2.0, 0.5)^T$	$(-5.00048787, 1.00119552)^T$

Tabelle 5.1: Ergebnisse für die PES-Familie (5.5)

5. Bestimmung von VRI-Punkten

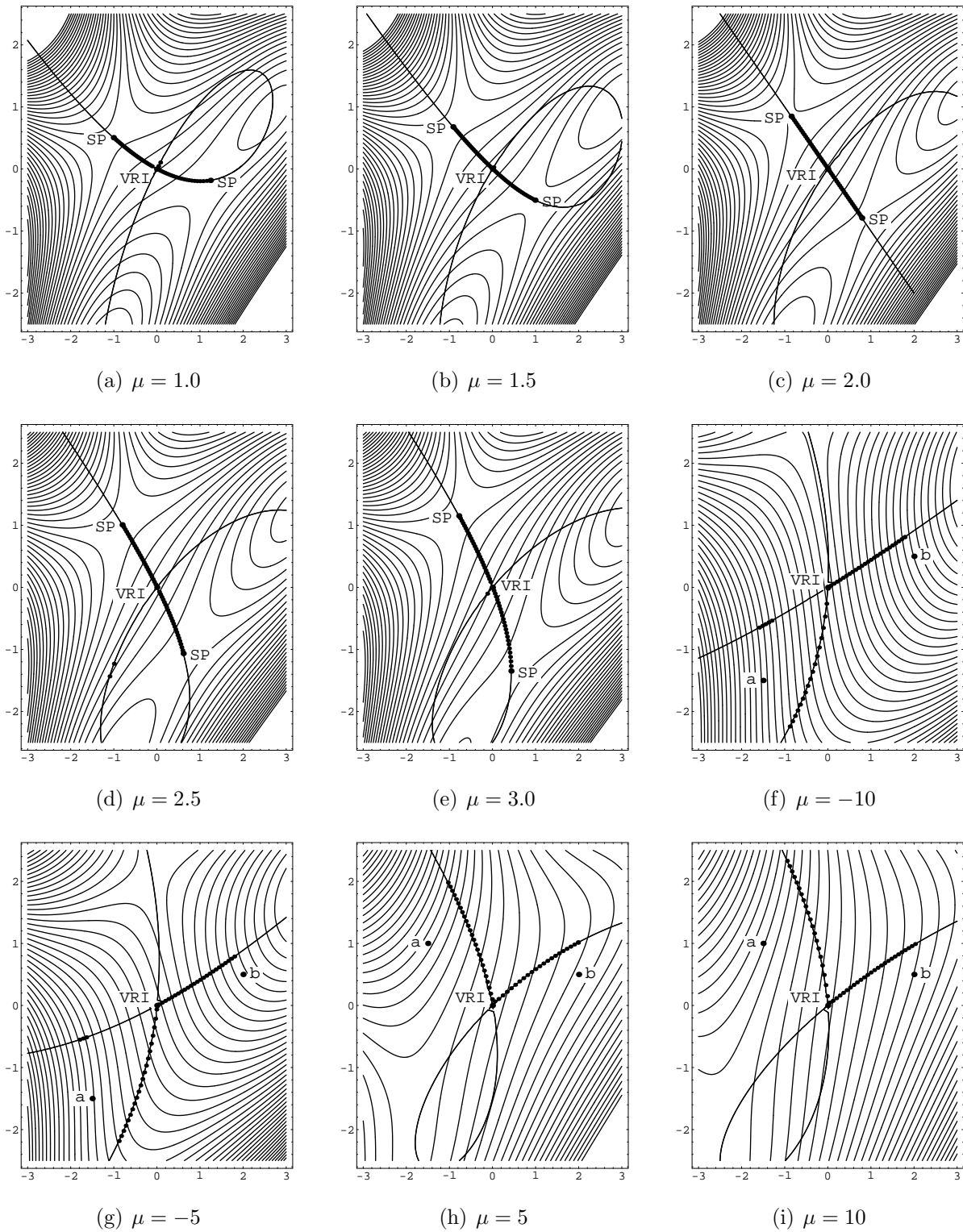


Abbildung 5.4: VRI-Bestimmung für verschiedene Wahlen von μ

5.2.2. Das Müller-Brown-Potential

Wir wollen zunächst die stationären Punkte des gemäß (A.2) definierten Müller-Brown-Potentials wie folgt benennen:

$$\text{Min}_1 = (-0.5582236346, 1.441725842)^T, \quad (5.7a)$$

$$\text{Min}_2 = (-0.050010823, 0.4666941049)^T, \quad (5.7b)$$

$$\text{Min}_3 = (0.6234994049, 0.02803775853)^T, \quad (5.7c)$$

$$\text{SP}_1 = (-0.8220015587, 0.6243128028)^T, \quad (5.7d)$$

$$\text{SP}_2 = (0.2124865820, 0.2929883251)^T. \quad (5.7e)$$

Wenden wir unseren Algorithmus auf das Müller-Brown-Potential an, so wählen wir als Ausgangskette die gerade Verbindung zwischen einem der stationären Punkte (5.7) und einem gewissen Punkt \mathbf{b} . Die Anfangssuchrichtung sei als Differenz der beiden Kettenrandpunkte gegeben. Alle übrigen Parameter belassen wir wie in (5.6). Dann zeigen uns die Abbildungen 5.5, 5.6 und 5.7 sowie Tabelle 5.3 die Ergebnisse des Programms nach jeweils fünf Durchläufen. Dabei sind die Newtontrajektorien zu den vom Programm ermittelten VRI-Suchrichtungen durchgehend gezeichnet. Es lohnt sich ein Vergleich zu den Werten aus Hirschs Dissertation [19]. Durch sukzessive Annäherung fand er folgende VRI-Punkte:

Bezeichnung	x_1 -Koordinate	x_2 -Koordinate	Suchrichtung [grad]
VRI ₁	0.37250	1.26315	30.39
VRI ₂	-0.75002	0.22586	66.805
VRI ₃	0.54859	0.45930	37.67
VRI ₄	-0.98072	-0.04753	61.96

Tabelle 5.2: VRI-Punkte auf dem Müller-Brown-Potential nach [19]

Bemerkung

Da VRI₄ durch den nahe gelegenen VRI₂ „verdeckt“ wird, gestaltet sich die Suche nach ihm als besonders schwierig. In Abbildung 5.8(a) sehen wir das Ergebnis nach fünf Durchläufen, wobei als Kettenrandpunkte Min₃ und $\mathbf{b} = (-1.3, -0.6)^T$, als Anfangssuchrichtung deren Differenz und die Parameter (5.6) gewählt wurden. Selbst bei einer geschickten Wahl der Anfangssuchrichtung erhalten wir keine brauchbaren Ergebnisse, wie Abbildung 5.8(b) verdeutlicht. An dieser Stelle sei jedoch angemerkt, daß es sich beim Müller-Brown-Potential um eine Modellfläche handelt. Mit derart skurrilen VRI-Punkten ist auf einer tatsächlichen PES allerdings nur selten zu rechnen.

5. Bestimmung von VRI-Punkten

Kettenanfang	Kettenende \mathbf{b}	VRI-Suchrichtung nach fünf Programmdurchläufen		VRI-Punkt
		als Vektor	als Winkel [grad]	
Min ₁	$(0.8, 1.2)^T$	$(0.156904778, 0.092078013)^T$	30.406	VRI ₁
Min ₂	$(0.5, 1.5)^T$	$(0.15342699, 0.0899789342)^T$	30.390	VRI ₁
SP ₁	$(0.7, 1.6)^T$	$(0.152326245, 0.0893388151)^T$	30.391	VRI ₁
SP ₁	$(0.8, 1.1)^T$	$(0.15351955, 0.0900331579)^T$	30.390	VRI ₁
SP ₁	$(-1.4, -0.2)^T$	$(-0.0598582422, -0.139756889)^T$	66.814	VRI ₂
Min ₂	$(-1.4, -0.2)^T$	$(-0.0603581858, -0.140857071)^T$	66.805	VRI ₂
SP ₁	$(-0.5, 0.0)^T$	$(-0.0613512782, -0.143180758)^T$	66.805	VRI ₂
SP ₂	$(-1.2, 0.3)^T$	$(-0.0613392838, -0.14315276)^T$	66.805	VRI ₂
SP ₂	$(0.7, 0.7)^T$	$(0.171035523, 0.132002749)^T$	37.660	VRI ₃
SP ₂	$(1.0, 1.0)^T$	$(0.170668075, 0.131716693)^T$	37.660	VRI ₃
Min ₂	$(1.0, 1.0)^T$	$(0.169014104, 0.130435936)^T$	37.659	VRI ₃
Min ₂	$(0.9, 0.2)^T$	$(0.171095051, 0.132048874)^T$	37.660	VRI ₃

Tabelle 5.3: Ergebnisse für das Müller-Brown-Potential

5.3. Ein dreidimensionales Beispiel

Wir wollen nun den Algorithmus auf eine dreidimensionale Fläche anwenden. Hierfür erweitern wir die durch (A.4) gegebene Modell-PES gemäß [41] wie folgt:

$$E(x_1, x_2, x_3) = 2x_2 + x_2^2 + (x_2 + 0.4x_1^2 + x_3^2)x_1^2 + 0.01x_3^2. \quad (5.8)$$

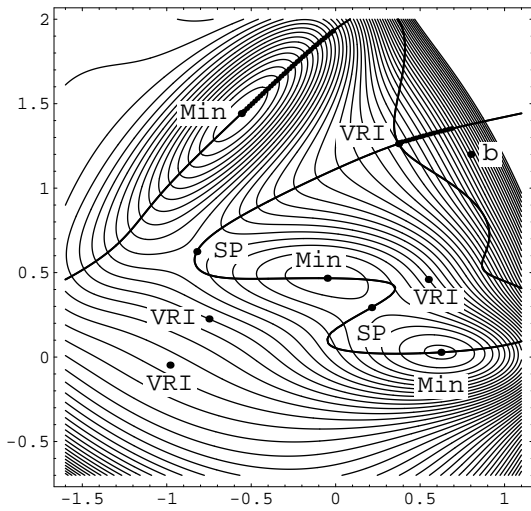
Die Fläche (5.8) stellt eine Vereinfachung der höherdimensionalen molekularen PES von Malonaldehyd-ähnlichen Systemen dar. Findet ein Protonentransfer im Malonaldehyd statt, so können wir x_1 als Verschiebung des H-Atoms, x_2 als symmetrische O-O-Ausdehnung und x_3 als Out-of-plane-Schwingung des gesamten Moleküls verstehen [41].

Die betrachtete Potentialenergiefläche (5.8) hat zwei Minima in $(\pm\sqrt{10/3}, -8/3, 0)^T$ und einen Sattelpunkt vom Index 1 in $(0, -1, 0)^T$. Ihre VRI-Punkte können wir nach Berechnung von Gradient und Hessematrix direkt ablesen:

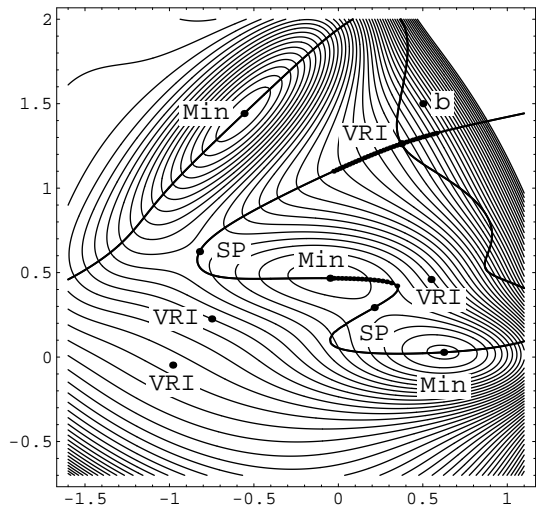
$$\mathbf{g}(0, x_2, x_3) = 2 \begin{pmatrix} x_1(0.8x_1^2 + x_2 + x_3^2) \\ 1 + x_2 + 0.5x_1^2 \\ x_3(x_1^2 + 0.01) \end{pmatrix} \Big|_{x_1=0} = 2 \begin{pmatrix} 0 \\ 1 + x_2 \\ 0.01x_3 \end{pmatrix},$$

$$\mathbf{H}(0, x_2, x_3) = 2 \begin{pmatrix} 2.4x_1^2 + x_2 + x_3^2 & x_1 & 2x_1x_3 \\ x_1 & 1 & 0 \\ 2x_1x_3 & 0 & x_1^2 + 0.01 \end{pmatrix} \Big|_{x_1=0} = 2 \begin{pmatrix} x_2 + x_3^2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0.01 \end{pmatrix}.$$

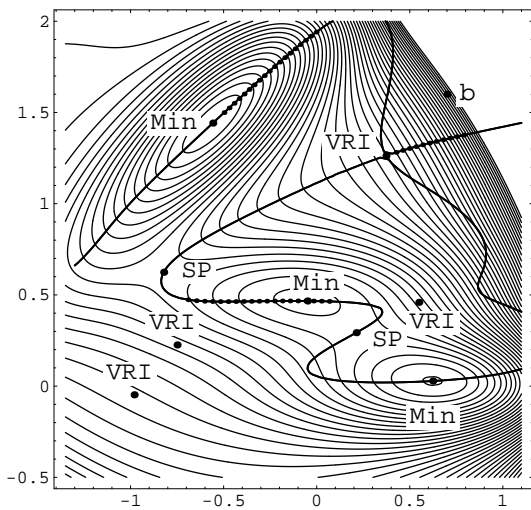
5.3. Ein dreidimensionales Beispiel



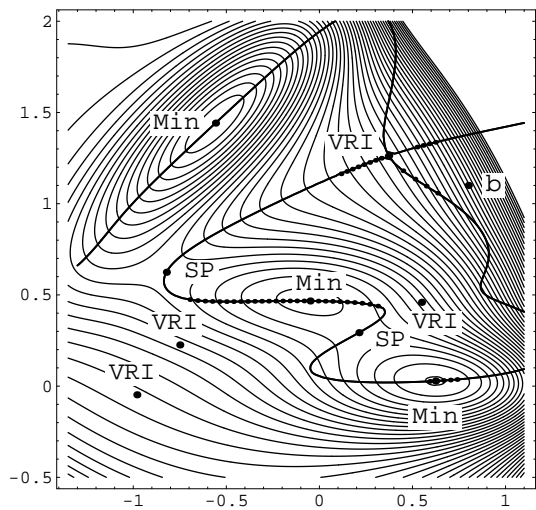
(a) Min_1 und $(0.8, 1.2)^T$



(b) Min_2 und $(0.5, 1.5)^T$



(c) SP_1 und $(0.7, 1.6)^T$



(d) SP_1 und $(0.8, 1.1)^T$

Abbildung 5.5: Anwendung des Programms auf das Müller-Brown-Potential mit den angegebenen Kettenrandpunkten für die Ermittlung von VRI_1

Wir sehen, daß für alle Punkte $(0, -x_3^2, x_3)^T$ die Hessematrix einen Rangabfall hat, während $\mathbf{g}(0, -x_3^2, x_3) \neq \mathbf{0}$ gilt. Da \mathbf{H} genau dann regulär ist, wenn ihre Adjunkte \mathbf{A} es auch ist, gilt für alle Punkte

$$\mathbf{x}^{\text{vri}} \in \{\mathbf{x} \in K : x_1 = 0, x_2 = -x_3^2\}$$

die Beziehung $\mathbf{A}(\mathbf{x}^{\text{vri}})\mathbf{g}(\mathbf{x}^{\text{vri}}) = \mathbf{0}$ mit $\mathbf{g}(\mathbf{x}^{\text{vri}}) \neq \mathbf{0}$. Die VRI-Punkte von (5.8) bilden also innerhalb des Konfigurationsraums eine Parabel in der Ebene $x_1 = 0$.

5. Bestimmung von VRI-Punkten

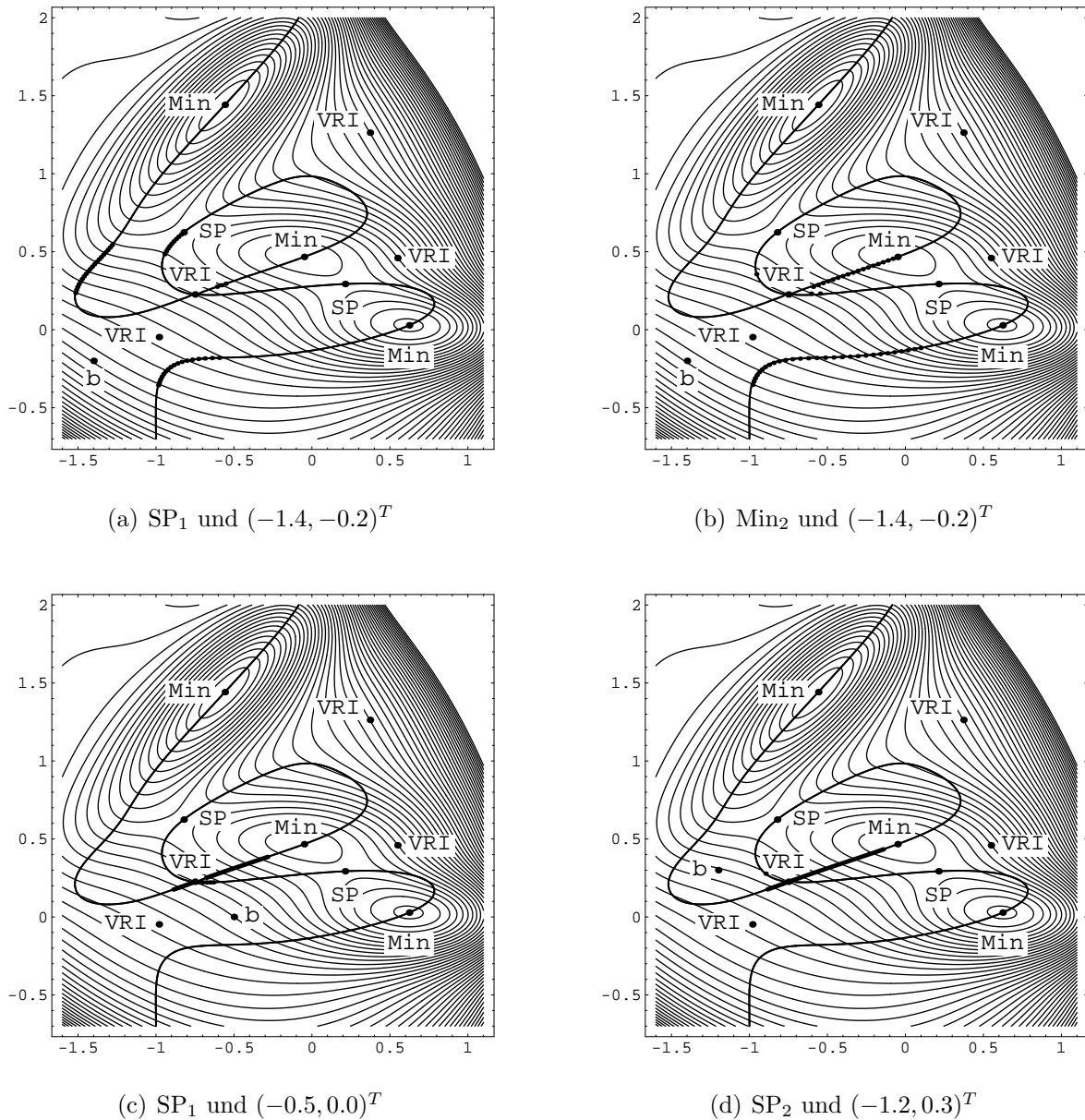
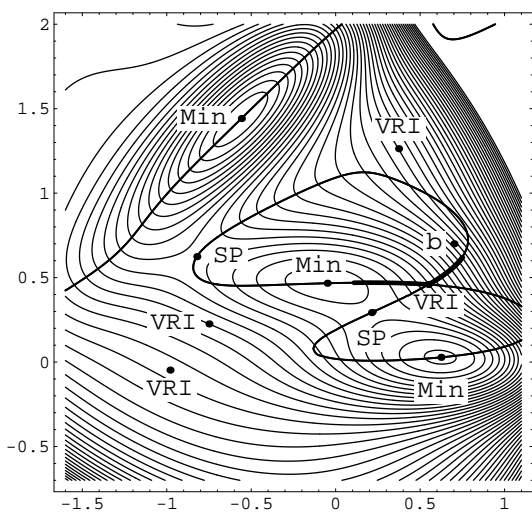


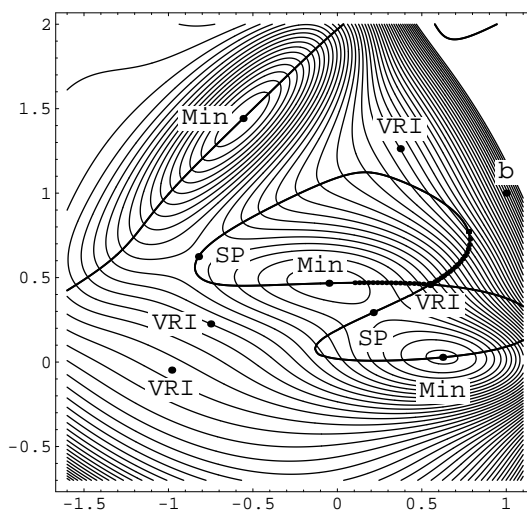
Abbildung 5.6: Anwendung des Programms auf das Müller-Brown-Potential mit den angegebenen Kettenrandpunkten für die Ermittlung von VRI_2

Bevor wir das Programm auf die gesamte Fläche anwenden, testen wir den Algorithmus an einigen zweidimensionalen Schnitten. Hierfür seien wie üblich die Parameter (5.6) gewählt. Wie bereits bemerkt, verstehen wir den Grenzwert $\delta = 0.1$ in der Minimumbildung (5.4) ebenfalls als Parameter. Wegen des (kleinen) Koeffizienten vor dem letzten Summanden in (5.8) gibt es VRI-Punkte, deren Gradient eine kleinere Norm als 0.1 hat. Demzufolge verwenden wir in diesem Fall (5.4) mit dem Parameter $\delta := 0.001$.

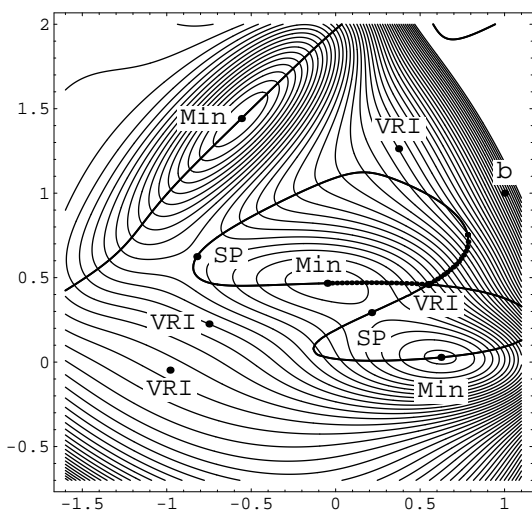
5.3. Ein dreidimensionales Beispiel



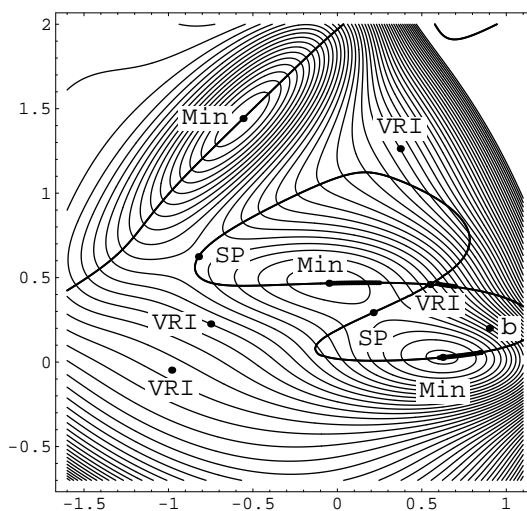
(a) SP_2 und $(0.7, 0.7)^T$



(b) SP_2 und $(1.0, 1.0)^T$



(c) Min_2 und $(1.0, 1.0)^T$



(d) Min_2 und $(0.9, 0.2)^T$

Abbildung 5.7: Anwendung des Programms auf das Müller-Brown-Potential mit den angegebenen Kettenrandpunkten für die Ermittlung von VRI_3

Abbildung 5.9 zeigt die Wirkung des Programms für drei Schnittflächen von (5.8) nach jeweils fünf Durchläufen. Als Kettenendpunkte wählten wir **a** und **b**. Die Newtontrajektorie zur vom Programm ermittelten VRI-Suchrichtung ist wie üblich durchgehend gezeichnet.

Wir wollen nun die Ergebnisse des Programms für die gesamte PES (5.8) festhalten. Wie üblich verwenden wir dabei die Parameter (5.6). Dann zeigt uns Tabelle 5.4 die vom

5. Bestimmung von VRI-Punkten

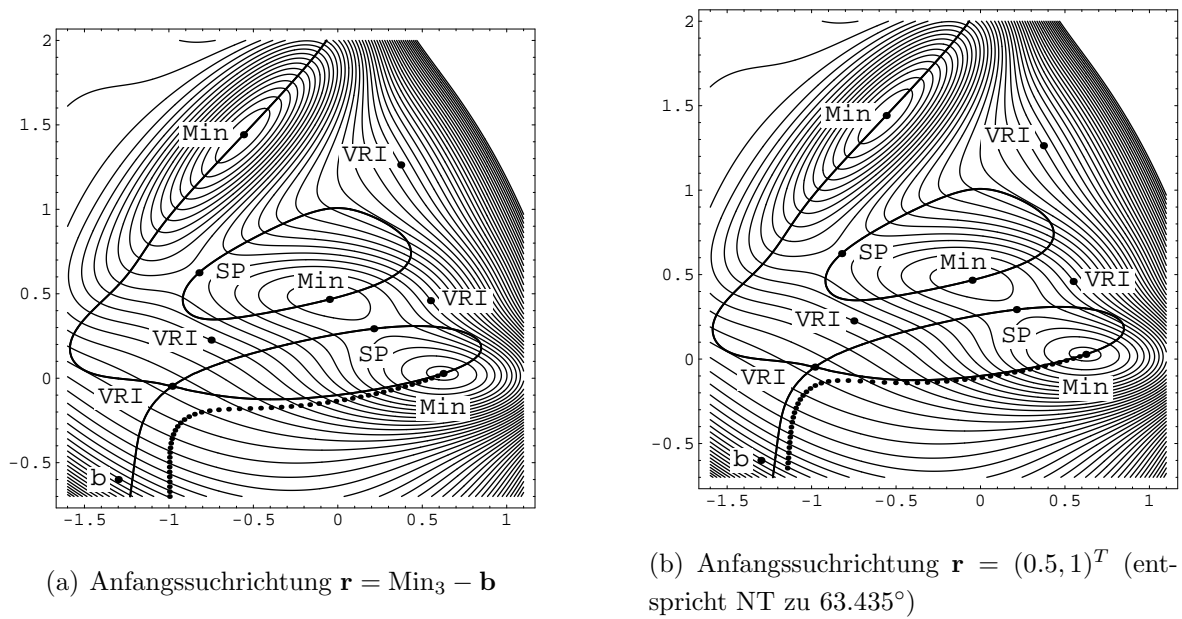


Abbildung 5.8: Verdeckter VRI-Punkt auf dem Müller-Brown-Potential

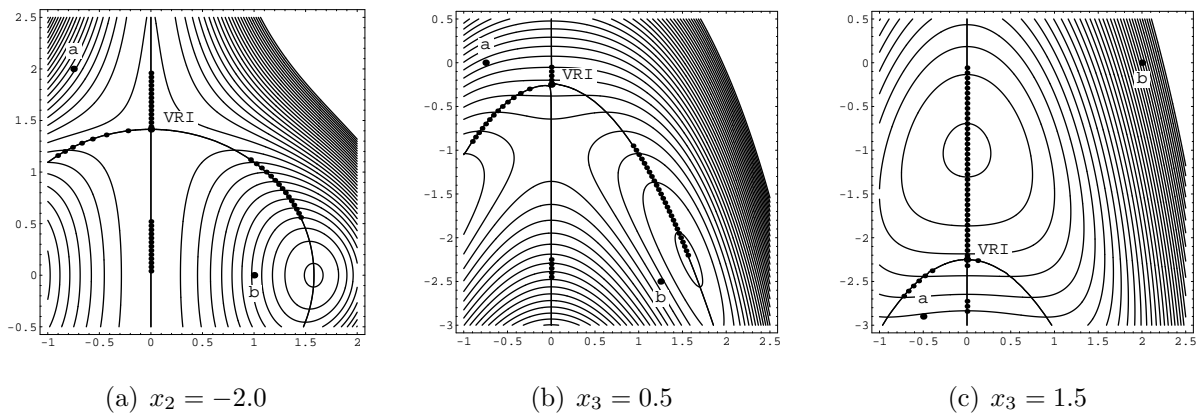


Abbildung 5.9: VRI-Bestimmung auf Schnittflächen einer dreidimensionalen PES

Programm ermittelten Suchrichtungen zur VRI-NT sowie den dabei berechneten VRI-Punkt für die jeweils angegebenen Kettenendpunkte. Fixieren wir die dritte Komponente der aus dem Programm hervorgegangenen Suchrichtung, so können wir zum Vergleich eine theoretische Suchrichtung angeben. Wir erinnern uns, daß die VRI-Punkte von (5.8) gegeben waren durch $(0, -x_3^2, x_3)^T$. Demzufolge ist die Suchrichtung der VRI-NT durch $\mathbf{g}(0, -x_3^2, x_3) = (0, 2 - 2x_3^2, 0.02x_3)$ gegeben. Die rechte Spalte von Tabelle 5.4 zeigt uns den zur theoretischen Suchrichtung gehörigen VRI-Punkt.

Abbildung 5.10 veranschaulicht die in Tabelle (5.4) aufgelisteten Ergebnisse. Wir sehen die Parabel der VRI-Punkte als durchgehende Linie dargestellt sowie die vom Programm

Kettenanfang	Kettenende	Ergebnisse nach 20 Programmdurchläufen		Theoretische Ergebnisse	
		VRI-Suchrichtung	VRI-Punkt	VRI-Suchrichtung	VRI-Punkt
$\begin{pmatrix} -0.01 \\ -2.5 \\ 1.5 \end{pmatrix}$	$\begin{pmatrix} 0.01 \\ -2 \\ 1.5 \end{pmatrix}$	$\begin{pmatrix} 0 \\ -2.5 \\ 0.03 \end{pmatrix}$	$\begin{pmatrix} 0 \\ -2.25 \\ 1.5 \end{pmatrix}$	$\begin{pmatrix} 0 \\ -2.5 \\ 0.03 \end{pmatrix}$	$\begin{pmatrix} 0 \\ -2.25 \\ 1.5 \end{pmatrix}$
$\begin{pmatrix} -0.01 \\ -2.2 \\ 1.4 \end{pmatrix}$	$\begin{pmatrix} 0.01 \\ -1.8 \\ 1.5 \end{pmatrix}$	$\begin{pmatrix} 2.69348172 \cdot 10^{-6} \\ -2.16030863 \\ 0.0288479536 \end{pmatrix}$	$\begin{pmatrix} -0.001672159 \\ -2.08015571 \\ 1.44199448 \end{pmatrix}$	$\begin{pmatrix} 0 \\ -2.16102214 \\ 0.0288479536 \end{pmatrix}$	$\begin{pmatrix} 0 \\ -2.08051107 \\ 1.44239768 \end{pmatrix}$
$\begin{pmatrix} -0.01 \\ -2.5 \\ 1 \end{pmatrix}$	$\begin{pmatrix} 0.01 \\ -1.5 \\ 1.6 \end{pmatrix}$	$\begin{pmatrix} 3.94557037 \cdot 10^{-8} \\ -1.84671358 \\ 0.0277368621 \end{pmatrix}$	$\begin{pmatrix} -0.0003184054 \\ -1.92335684 \\ 1.38682905 \end{pmatrix}$	$\begin{pmatrix} 0 \\ -1.8466676 \\ 0.0277368621 \end{pmatrix}$	$\begin{pmatrix} 0 \\ -1.9233338 \\ 1.386843105 \end{pmatrix}$
$\begin{pmatrix} -0.01 \\ -1.9 \\ 1.3 \end{pmatrix}$	$\begin{pmatrix} 0.01 \\ -1.7 \\ 1.4 \end{pmatrix}$	$\begin{pmatrix} 1.17886938 \cdot 10^{-6} \\ -1.61967903 \\ 0.0269042048 \end{pmatrix}$	$\begin{pmatrix} -0.000984 \\ -1.80984 \\ 1.34508 \end{pmatrix}$	$\begin{pmatrix} 0 \\ -1.61918117 \\ 0.0269042048 \end{pmatrix}$	$\begin{pmatrix} 0 \\ -1.80959059 \\ 1.34521024 \end{pmatrix}$
$\begin{pmatrix} -0.01 \\ -1.9 \\ 1.2 \end{pmatrix}$	$\begin{pmatrix} 0.01 \\ -1.4 \\ 1.4 \end{pmatrix}$	$\begin{pmatrix} -4.8399643 \cdot 10^{-8} \\ -1.34038201 \\ 0.0258489989 \end{pmatrix}$	$\begin{pmatrix} -0.0007836852 \\ -1.67019131 \\ 1.29237057 \end{pmatrix}$	$\begin{pmatrix} 0 \\ -1.34085373 \\ 0.0258489989 \end{pmatrix}$	$\begin{pmatrix} 0 \\ -1.67042686 \\ 1.292449945 \end{pmatrix}$
$\begin{pmatrix} -0.01 \\ -1.6 \\ 1.1 \end{pmatrix}$	$\begin{pmatrix} 0.01 \\ -1.3 \\ 1.3 \end{pmatrix}$	$\begin{pmatrix} -1.6141498 \cdot 10^{-6} \\ -0.923056496 \\ 0.0241890603 \end{pmatrix}$	$\begin{pmatrix} -0.0007422404 \\ -1.46152852 \\ 1.20938639 \end{pmatrix}$	$\begin{pmatrix} 0 \\ -0.925553188 \\ 0.0241890603 \end{pmatrix}$	$\begin{pmatrix} 0 \\ -1.4627766 \\ 1.209453015 \end{pmatrix}$
$\begin{pmatrix} -0.01 \\ -1.5 \\ 1 \end{pmatrix}$	$\begin{pmatrix} 0.01 \\ -1 \\ 1.3 \end{pmatrix}$	$\begin{pmatrix} 4.20367024 \cdot 10^{-7} \\ -0.627548326 \\ 0.0229259142 \end{pmatrix}$	$\begin{pmatrix} -0.0012217414 \\ -1.31377491 \\ 1.14612464 \end{pmatrix}$	$\begin{pmatrix} 0 \\ -0.627987717 \\ 0.0229259142 \end{pmatrix}$	$\begin{pmatrix} 0 \\ -1.31399385 \\ 1.146231571 \end{pmatrix}$
$\begin{pmatrix} -0.01 \\ -1 \\ 0.8 \end{pmatrix}$	$\begin{pmatrix} 0.01 \\ -0.5 \\ 1.2 \end{pmatrix}$	$\begin{pmatrix} 2.4056024 \cdot 10^{-7} \\ 0.00600914408 \\ 0.0199684322 \end{pmatrix}$	$\begin{pmatrix} -0.000564369 \\ -0.996995587 \\ 0.998389808 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 0.00630858468 \\ 0.0199684322 \end{pmatrix}$	$\begin{pmatrix} 0 \\ -0.9968457 \\ 0.99842161 \end{pmatrix}$
$\begin{pmatrix} 0.01 \\ -0.9 \\ 0.5 \end{pmatrix}$	$\begin{pmatrix} -0.01 \\ -0.7 \\ 1.0 \end{pmatrix}$	$\begin{pmatrix} -2.982594 \cdot 10^{-6} \\ 0.407254342 \\ 0.0178468242 \end{pmatrix}$	$\begin{pmatrix} 0.00201013685 \\ -0.796374849 \\ 0.891980789 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 0.407454339 \\ 0.0178468242 \end{pmatrix}$	$\begin{pmatrix} 0 \\ -0.79627284 \\ 0.89234121 \end{pmatrix}$
$\begin{pmatrix} -0.01 \\ -0.8 \\ 0.5 \end{pmatrix}$	$\begin{pmatrix} 0.01 \\ -0.1 \\ 1.1 \end{pmatrix}$	$\begin{pmatrix} 7.24632516 \cdot 10^{-7} \\ 0.830618933 \\ 0.0152921895 \end{pmatrix}$	$\begin{pmatrix} -0.0013363888 \\ -0.584691426 \\ 0.764472944 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 0.830744705 \\ 0.0152921895 \end{pmatrix}$	$\begin{pmatrix} 0 \\ -0.58462765 \\ 0.764609475 \end{pmatrix}$
$\begin{pmatrix} -0.01 \\ -0.7 \\ 0.4 \end{pmatrix}$	$\begin{pmatrix} 0.01 \\ -0.5 \\ 0.9 \end{pmatrix}$	$\begin{pmatrix} 6.2417347 \cdot 10^{-7} \\ 0.964585976 \\ 0.0143710363 \end{pmatrix}$	$\begin{pmatrix} -0.0002236459 \\ -0.517707037 \\ 0.718548222 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 0.967366576 \\ 0.0143710363 \end{pmatrix}$	$\begin{pmatrix} 0 \\ -0.51631671 \\ 0.718551815 \end{pmatrix}$
$\begin{pmatrix} 0.01 \\ -0.5 \\ 0.2 \end{pmatrix}$	$\begin{pmatrix} -0.01 \\ -0.3 \\ 0.7 \end{pmatrix}$	$\begin{pmatrix} 3.61683168 \cdot 10^{-9} \\ 1.38600344 \\ 0.0110979194 \end{pmatrix}$	$\begin{pmatrix} 1.98452 \cdot 10^{-6} \\ -0.306998278 \\ 0.554895971 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 1.38418092 \\ 0.0110979194 \end{pmatrix}$	$\begin{pmatrix} 0 \\ -0.30790954 \\ 0.55489597 \end{pmatrix}$
$\begin{pmatrix} -0.01 \\ -0.5 \\ 0.2 \end{pmatrix}$	$\begin{pmatrix} 0.01 \\ 0 \\ 0.7 \end{pmatrix}$	$\begin{pmatrix} -6.156486 \cdot 10^{-8} \\ 1.64098823 \\ 0.00848987186 \end{pmatrix}$	$\begin{pmatrix} -4.46863 \cdot 10^{-5} \\ -0.179505885 \\ 0.424493508 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 1.63961038 \\ 0.00848987186 \end{pmatrix}$	$\begin{pmatrix} 0 \\ -0.18019481 \\ 0.424493593 \end{pmatrix}$
$\begin{pmatrix} 0.01 \\ -0.5 \\ 0.1 \end{pmatrix}$	$\begin{pmatrix} -0.01 \\ -0.1 \\ 0.4 \end{pmatrix}$	$\begin{pmatrix} 3.0897083 \cdot 10^{-9} \\ 1.76836632 \\ 0.00680857469 \end{pmatrix}$	$\begin{pmatrix} 2.0633008 \cdot 10^{-5} \\ -0.115816841 \\ 0.34042872 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 1.76821655 \\ 0.00680857469 \end{pmatrix}$	$\begin{pmatrix} 0 \\ -0.11589172 \\ 0.340428735 \end{pmatrix}$
$\begin{pmatrix} -0.2 \\ -0.5 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 0.4 \\ 0.3 \\ 0 \end{pmatrix}$	$\begin{pmatrix} -1.1281 \cdot 10^{-12} \\ 1.99968 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 3.5252 \cdot 10^{-9} \\ -0.00016 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$

Tabelle 5.4: Ergebnisse für die dreidimensionale Potentialenergiefläche (5.8)

ermittelten VRI-Punkte (fett) und die zugehörigen Ketten (dünne Punkte).

Bemerkung

In unseren bisherigen Überlegungen sind wir meistens davon ausgegangen, daß sich VRI-Punkte zwischen zwei stationären Punkten befinden, die entweder beide einen geraden oder beide einen ungeraden Index haben. Eine solche Situation ist auf der hier betrachteten Fläche (5.8) nicht gegeben, wodurch die Wahl der Kettenendpunkte problematisch wird. Wir wollen bemerken, daß das Programm nur bei einer „geschickt“ gesetzten Startkette einwandfrei funktioniert. Dies zu spezifizieren ist aufgrund der Vielfalt möglicher Potentialenergieflächen müßig. Für die konkrete Anwendung muß der Benutzer mit dem

5. Bestimmung von VRI-Punkten

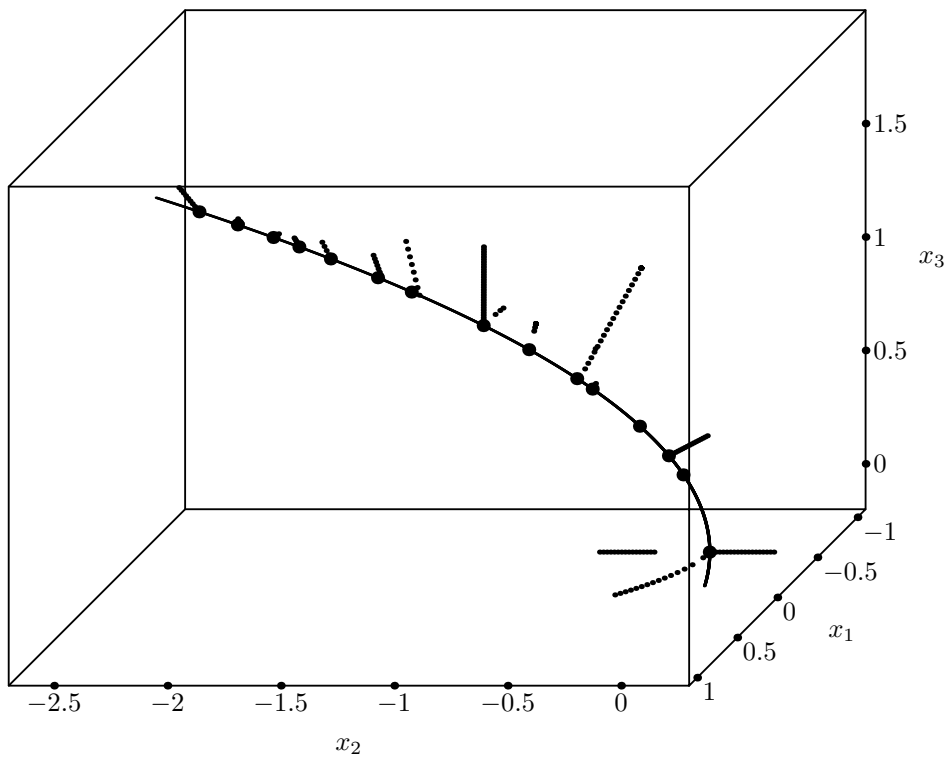


Abbildung 5.10: Ergebnisse für die dreidimensionale Potentialenergiefläche (5.8)

Programm interagieren, da es keine Universallösung für beliebige Flächen liefert. In jedem Fall ist eine theoretische Voruntersuchung der betrachteten PES unerlässlich.

Zusammenfassung

Der Minimum-Energie-Weg (MEP) ist das grundlegende Konzept, um den Ablauf chemischer Reaktionen zu verstehen. Er wird als Trajektorie im Konfigurationsraum der zur Reaktion gehörigen, typischerweise hochdimensionalen Potentialenergiefläche (PES) aufgefaßt. Da er mathematisch nicht eindeutig festgelegt ist, gibt es verschiedene Reaktionswegmodelle, die, abhängig von der betrachteten PES, den Kriterien des MEP genügen.

Die vorliegende Arbeit stellt mit der intrinsischen Reaktionskoordinate (IRC), den Gradient Extremals (GE) und der Methode des reduzierten Gradienten (RGF) drei bedeutende Modelle vor, diskutiert ihre Eignung als Minimum-Energie-Weg und beschreibt ihre numerische Umsetzung. Dabei stellte sich heraus, daß keines der drei als universelles Reaktionswegmodell aufgefaßt werden kann. Erfolgversprechend erscheint vielmehr eine Kombination der vorgestellten Verfahren, um die chemisch relevanten Bereiche der Potentialenergiefläche zu analysieren.

Ein weiterer Schwerpunkt der Arbeit liegt auf der Untersuchung von Reaktionswegverzweigungen, die, unabhängig vom gewählten Modell, in der Nähe von Tal-Rücken-Umschlagpunkten (VRI-Punkten) stattfinden. Für deren Ermittlung ist die RGF-Methode bzw. ihre mathematische Entsprechung, das desingularisierte kontinuierliche Newtonverfahren, von entscheidender Bedeutung, da sich ihre Trajektorien (auch Newtontrajektorien genannt) gerade in den VRI-Punkten verzweigen. Trotzdem gestaltet sich die numerische Bestimmung dieser Punkte aufgrund der Vielgestaltigkeit und Hochdimensionalität der PES als schwierig.

Einen vielversprechenden Lösungsansatz liefert die Variationsrechnung, derzufolge wir eine, in einem gewissen Bereich um den VRI-Punkt liegende Ausgangstrajektorie solange variieren können, bis sie durch den Tal-Rücken-Umschlagpunkt verläuft. Sie bietet damit die theoretische Grundlage für das in dieser Arbeit gewählte Verfahren zur VRI-Bestimmung.

Hauptziel der vorliegenden Arbeit war die Erstellung eines Programms zur Ermittlung von VRI-Punkten auf Potentialenergieflächen. Es zeigte sich, daß der in Abschnitt 5.1 entwickelte Algorithmus dieses Ziel für zweidimensionale Modellpotentiale und ein dreidimensionales, stark vereinfachtes Alanindipeptid-Potential erfüllt. Ihm zugrunde liegt ein Variationsansatz für Newtontrajektorien, der numerisch durch eine String-Methode realisiert wurde. Das Programm liefert statt der gesuchten VRI-Punkte deren Gradienten und damit die zum VRI-Punkt gehörige Newtontrajektorie. Auf diese Weise konnten zum Teil sehr hohe Genauigkeiten erreicht werden, wie die Ergebnisse in den Abschnitten 5.2 und 5.3 belegen.

Zusammenfassung

Eine VRI-Bestimmung auf weiteren, höherdimensionalen Potentialenergieflächen ist im Rahmen dieser Arbeit nicht erfolgt. Hierin liegt ein möglicher Anknüpfungspunkt für weitere Untersuchungen. Auch die Einbeziehung der in Abschnitt 3.4 kurz erläuterten Growing-String-Methode erscheint vielversprechend, da sie insbesondere in höheren Dimensionen den Rechenaufwand verringert und Nachteile der gewählten String-Methode beseitigt.

A. Verwendete Potentialenergieflächen

Abbildungen 1.1(a), 1.3(b), 2.3(a), 3.4, 5.1, 5.2, 5.3

Die verwendete PES ist gemäß [43] gegeben durch

$$E(x_1, x_2) = \frac{1}{2} (x_1 x_2^2 - x_1^2 x_2 - \frac{1}{2} x_1 + 2x_2) + \frac{1}{30} (x_1^4 + x_2^4) . \quad (\text{A.1})$$

Die stationären Punkte liegen bei

$$\begin{aligned} \text{Min}_1 &= (-11.21880242, 11.20660782) , & \text{SP}_1 &= (-1.118846247, 0.3389972869) , \\ \text{Min}_2 &= (-1.438595032, -3.272576948) , & \text{SP}_2 &= (1.688152221, 0.2504713827) , \\ \text{Min}_3 &= (2.260986845, 0.6704375189) . \end{aligned}$$

Abbildungen 1.1(b), 1.2, 2.3(b), 2.6, 3.1, 3.3, 3.6, 5.5, 5.6, 5.7, 5.8 (Müller-Brown-Potential)

Das Müller-Brown-Potential ist gemäß [31] gegeben durch

$$\begin{aligned} E(x_1, x_2) &= -200e^{-(x_1-1)^2-10x_2^2} - 100e^{-x_1^2-10(x_2-0.5)^2} \\ &\quad - 170e^{-6.5(x_1+0.5)^2+11(x_1+0.5)(x_2-1.5)-6.5(x_2-1.5)^2} \\ &\quad + 15e^{0.7(x_1+1)^2+0.6(x_1+1)(x_2-1)+0.7(x_2-1)^2} . \end{aligned} \quad (\text{A.2})$$

Abbildungen 1.3(a), 4.2

Die verwendete PES ist gemäß [43] gegeben durch

$$E(x_1, x_2) = \frac{1}{2} (x_1 x_2^2 - x_1^2 x_2 - 2x_1 + 2x_2) + \frac{1}{30} (x_1^4 + x_2^4) . \quad (\text{A.3})$$

Abbildungen 2.1, 2.2, 2.4, 4.1

Die verwendete PES ist gemäß [37] gegeben durch

$$E(x_1, x_2) = 2x_2 + x_2^2 + (x_2 + 0.4x_1^2) x_1^2 . \quad (\text{A.4})$$

Abbildung 2.5

Die verwendete PES ist gemäß [34] gegeben durch

$$\begin{aligned} E(x_1, x_2) &= 0.1553 (x_1^2 - 1)^2 + [x_2 + 0.7(x_1 + 1)]^2 [x_2 + 0.5(x_1 - 1)]^2 \\ &\quad + [x_2 - 0.7(x_1 - 1)]^2 [x_2 - 0.5(x_1 + 1)]^2 + 16x_2^2 (x_2 - 0.5)^2 . \end{aligned} \quad (\text{A.5})$$

Abbildung 3.5 (Neria-Fischer-Karplus-Potential)

Die verwendete PES ist gemäß [32] gegeben durch

$$E(x_1, x_2) = c (x_1^2 + x_2^2)^2 + x_1 x_2 - 9e^{-(x_1-3)^2-x_2^2} - 9e^{-(x_1+3)^2-x_2^2} . \quad (\text{A.6})$$

A. Verwendete Potentialenergieflächen

In [32] ist $c = 0.06$, während die in Abbildung 3.5 dargestellte modifizierte Form nach [19] den Parameter $c = 0.03$ verwendet.

Abbildung 3.7

Die verwendete PES ist gemäß [42] gegeben durch

$$E(x_1, x_2) = \frac{x_1^2 (80 - x_2^2)^2}{80} + 0.2x_1^4 + 0.1x_2^2 (200 - x_2^2) . \quad (\text{A.7})$$

B. Rechnungen

B.1. Eigenschaften der adjunkten Matrix

Wir wollen die Gültigkeit der auf den Seiten 10–11 genannten Beziehungen (i)–(iv) zwischen \mathbf{H} und ihrer Adjunkten \mathbf{A} beweisen.

- (i) Es ist zu zeigen, daß $\mathbf{AH} = \det(\mathbf{H})\mathbf{I}_n$ gilt. Bezeichne \mathbf{H}^{ij} diejenige Matrix, die aus \mathbf{H} durch Streichen der i -ten Zeile und j -ten Spalte entsteht. Dann gilt nach dem Entwicklungssatz von Laplace

$$\det(\mathbf{H}) = \sum_{i=1}^n (-1)^{i+j} H_{ij} \det(\mathbf{H}^{ij}) = \sum_{i=1}^n A_{ji} H_{ij}$$

für alle $j = 1, \dots, n$. Die Diagonaleinträge von \mathbf{AH} sind also gerade gleich $\det(\mathbf{H})$. Bleibt zu zeigen, daß sämtliche Nichtdiagonaleinträge $\sum_{i=1}^n A_{ji} H_{ik}$ mit $j \neq k$ gleich Null sind. Ersetzen wir die j -te Spalte von \mathbf{H} durch ihre k -te Spalte, bezeichnen die so entstandene Matrix mit $\overline{\mathbf{H}}$ und mit $\overline{\mathbf{H}}^{ij}$ diejenige Matrix die aus $\overline{\mathbf{H}}$ durch Streichen von i -ter Zeile und j -ter Spalte entsteht, so gilt offensichtlich $\mathbf{H}^{ij} = \overline{\mathbf{H}}^{ij}$. Wenden wir nun erneut den Laplaceschen Entwicklungssatz an, dann erhalten wir

$$\begin{aligned} 0 &= \det(\overline{\mathbf{H}}) = \sum_{i=1}^n (-1)^{i+j} \overline{H}_{ij} \det(\overline{\mathbf{H}}^{ij}) = \sum_{i=1}^n (-1)^{i+j} \overline{H}_{ij} \det(\mathbf{H}^{ij}) \\ &= \sum_{i=1}^n A_{ji} \overline{H}_{ij} = \sum_{i=1}^n A_{ji} H_{ik} . \end{aligned}$$

- (ii) Es ist zu zeigen, daß \mathbf{A} genau dann regulär ist, wenn \mathbf{H} es auch ist. Seien die Eigenvektoren von \mathbf{u}^i und die zugehörigen Eigenwerte mit λ_i bezeichnet. Sei o.B.d.A. $\lambda_i \neq 0$. Dann gilt

$$\begin{aligned} \mathbf{A}(\lambda_i \mathbf{u}^i) &= \mathbf{A}(\mathbf{H}\mathbf{u}^i) \stackrel{(i)}{=} \det(\mathbf{H})\mathbf{u}^i , \\ \mathbf{A}\mathbf{u}^i &= \frac{\det(\mathbf{H})}{\lambda_i} \mathbf{u}^i = \left(\prod_{j \neq i} \lambda_j \right) \mathbf{u}^i . \end{aligned} \tag{B.1}$$

Somit gelten folgende Äquivalenzen:

$$\begin{aligned} \mathbf{H} \text{ regulär} &\iff \det(\mathbf{H}) \neq 0 \iff \lambda_i \neq 0, \text{ für alle } i = 1, \dots, n \\ &\iff \det(\mathbf{A}) = \prod_{i=1}^n \frac{\det(\mathbf{H})}{\lambda_i} \neq 0 \iff \mathbf{A} \text{ regulär} . \end{aligned}$$

Die Aussagen (iii) und (iv) folgen direkt aus (B.1).

B.2. Sun-Ruedenberg-Algorithmus

Gradient Extremals sind gemäß (2.10) gegeben durch $\mathbf{p} := (\mathbf{I}_n - \mathbf{w}\mathbf{w}^T)(\mathbf{H}\mathbf{g}) = \mathbf{0}$, wobei \mathbf{w} den normierten Gradienten bezeichnet. An dieser Stelle werden Berechnungen durchgeführt, die wir für die Beschreibung des Sun-Ruedenberg-Algorithmus (Abschnitt 2.2.2) benötigen. Es sei bemerkt, daß wir uns nicht zwangsläufig auf einer Gradient Extremal befinden. Sollte dies der Fall sein, so wird ausdrücklich darauf hingewiesen. Wir wollen nun die in [47] angegebenen Formeln für die Ableitungen von \mathbf{p} nach x_i nachrechnen. Es gilt:

$$\frac{\partial \mathbf{p}}{\partial x_i} = (\mathbf{I}_n - \mathbf{w}\mathbf{w}^T) \left(\frac{\partial(\mathbf{H}\mathbf{g})}{\partial x_i} \right) - \left(\frac{\partial(\mathbf{w}\mathbf{w}^T)}{\partial x_i} \right) (\mathbf{H}\mathbf{g}). \quad (\text{B.2})$$

Dabei können wir die Ableitungsterme auf der rechten Seite von (B.2) wie folgt umschreiben:

$$\begin{aligned} \frac{\partial(\mathbf{H}\mathbf{g})_j}{\partial x_i} &= \frac{\partial(\sum_{k=1}^n H_{jk}g_k)}{\partial x_i} = \sum_{k=1}^n \left[\left(\frac{\partial H_{jk}}{\partial x_i} \right) g_k + H_{jk} \left(\frac{\partial g_k}{\partial x_i} \right) \right] \\ &= \sum_{k=1}^n \left[\left(\frac{\partial H_{ji}}{\partial x_k} \right) g_k + H_{jk} H_{ki} \right] = \sum_{k=1}^n \left(\frac{\partial H_{ji}}{\partial x_k} \right) g_k + (\mathbf{H}^2)_{ji}. \end{aligned} \quad (\text{B.3})$$

Um den zweiten Term der rechten Seite von (B.2) umzuformen, führen wir eine Zwischenrechnung aus. Es gilt für alle $i \in \{1, \dots, n\}$ die Beziehung

$$\begin{aligned} (\mathbf{I}_n - \mathbf{w}\mathbf{w}^T) \frac{\partial \mathbf{g}}{\partial x_i} &= (\mathbf{I}_n - \mathbf{w}\mathbf{w}^T) \frac{\partial(\|\mathbf{g}\|_2 \mathbf{w})}{\partial x_i} = (\mathbf{I}_n - \mathbf{w}\mathbf{w}^T) \left(\frac{\partial \|\mathbf{g}\|_2}{\partial x_i} \mathbf{w} + \|\mathbf{g}\|_2 \frac{\partial \mathbf{w}}{\partial x_i} \right) \\ &= \underbrace{\frac{\partial \|\mathbf{g}\|_2}{\partial x_i} \mathbf{P}_{\mathbf{w}}(\mathbf{w})}_{=0, \text{ nach (2.6)}} + \|\mathbf{g}\|_2 \frac{\partial \mathbf{w}}{\partial x_i} - \|\mathbf{g}\|_2 \mathbf{w} \left(\sum_{k=1}^n w_k \frac{\partial w_k}{\partial x_i} \right) \\ &= \|\mathbf{g}\|_2 \frac{\partial \mathbf{w}}{\partial x_i} - \frac{\|\mathbf{g}\|_2 \mathbf{w}}{2} \cdot \frac{\partial(\sum_{k=1}^n w_k^2)}{\partial x_i} \\ &= \|\mathbf{g}\|_2 \frac{\partial \mathbf{w}}{\partial x_i} - \frac{\|\mathbf{g}\|_2 \mathbf{w}}{2} \cdot \underbrace{\frac{\partial \|\mathbf{w}\|_2^2}{\partial x_i}}_{=0, \text{ da } \|\mathbf{w}\|_2 = 1} = \|\mathbf{g}\|_2 \frac{\partial \mathbf{w}}{\partial x_i}. \end{aligned}$$

Setzen wir dieses Ergebnis in den zweiten Term der rechten Seite von (B.2) ein, so erhalten wir

$$\begin{aligned}
 \frac{\partial \mathbf{w}\mathbf{w}^T}{\partial x_i}(\mathbf{H}\mathbf{g}) &= \left(\frac{\partial \mathbf{w}}{\partial x_i}\right) (\mathbf{w}^T \mathbf{H}\mathbf{g}) + \mathbf{w} \left(\frac{\partial \mathbf{w}^T}{\partial x_i}\right) (\mathbf{H}\mathbf{g}) \\
 &= \frac{(\mathbf{I}_n - \mathbf{w}\mathbf{w}^T)}{\|\mathbf{g}\|_2} \left(\frac{\partial \mathbf{g}}{\partial x_i}\right) [\mathbf{w}^T \mathbf{H}(\|\mathbf{g}\|_2 \mathbf{w})] \\
 &\quad + \mathbf{w} \left(\frac{\partial \mathbf{g}^T}{\partial x_i}\right) \frac{(\mathbf{I}_n - \mathbf{w}\mathbf{w}^T)}{\|\mathbf{g}\|_2} [\mathbf{H}(\|\mathbf{g}\|_2 \mathbf{w})] \\
 &= (\mathbf{I}_n - \mathbf{w}\mathbf{w}^T) \left(\frac{\partial \mathbf{g}}{\partial x_i}\right) (\mathbf{w}^T \mathbf{H}\mathbf{w}) + \mathbf{w} \left(\frac{\partial \mathbf{g}^T}{\partial x_i}\right) (\mathbf{I}_n - \mathbf{w}\mathbf{w}^T) (\mathbf{H}\mathbf{w}) .
 \end{aligned} \tag{B.4}$$

Die komponentenweise Betrachtung von (B.4) liefert uns

$$\begin{aligned}
 \left[\frac{\partial \mathbf{w}\mathbf{w}^T}{\partial x_i}(\mathbf{H}\mathbf{g})\right]_j &= [(\mathbf{I}_n - \mathbf{w}\mathbf{w}^T) \mathbf{H}]_{ji} (\mathbf{w}^T \mathbf{H}\mathbf{w}) + w_j [\mathbf{H}(\mathbf{I}_n - \mathbf{w}\mathbf{w}^T)(\mathbf{H}\mathbf{w})]_i \\
 &= [(\mathbf{w}^T \mathbf{H}\mathbf{w})(\mathbf{I}_n - \mathbf{w}\mathbf{w}^T) \mathbf{H}]_{ji} + w_j \{\mathbf{H}[\mathbf{H}\mathbf{w} - \mathbf{w}(\mathbf{w}^T \mathbf{H}\mathbf{w})]\}_i \\
 &= [(\mathbf{w}^T \mathbf{H}\mathbf{w})(\mathbf{I}_n - \mathbf{w}\mathbf{w}^T) \mathbf{H}]_{ji} + w_j [\mathbf{H}^2 \mathbf{w} - (\mathbf{w}^T \mathbf{H}\mathbf{w}) \mathbf{H}\mathbf{w}]_i \\
 &= \{(\mathbf{w}^T \mathbf{H}\mathbf{w})(\mathbf{I}_n - \mathbf{w}\mathbf{w}^T) \mathbf{H} + \mathbf{w}\mathbf{w}^T [\mathbf{H}^2 - (\mathbf{w}^T \mathbf{H}\mathbf{w}) \mathbf{H}]\}_{ji} .
 \end{aligned} \tag{B.5}$$

Bis auf weiteres betrachten wir Punkte auf Gradient Extremals. In diesem Fall gilt Bedingung (2.10) und wir vereinfachen wie folgt:

$$\{\mathbf{w}\mathbf{w}^T [\mathbf{H}^2 - (\mathbf{w}^T \mathbf{H}\mathbf{w}) \mathbf{H}]\}_{ji} = w_j \underbrace{[\mathbf{H}(\mathbf{I}_n - \mathbf{w}\mathbf{w}^T)(\mathbf{H}\mathbf{w})]_i}_{=0 \text{ nach (2.10) mit (2.6)}} = 0 . \tag{B.6}$$

Jetzt können wir (B.3) und (B.5) in (B.2) einsetzen und erhalten mithilfe von (B.6):

$$\begin{aligned}
 \frac{\partial p_j}{\partial x_i} &= \left\{ (\mathbf{I}_n - \mathbf{w}\mathbf{w}^T) \left[\sum_{k=1}^n \frac{\partial \mathbf{H}}{\partial x_k} \cdot g_k + \mathbf{H}^2 - (\mathbf{w}^T \mathbf{H}\mathbf{w}) \mathbf{H} \right] \right\}_{ji} \\
 &= \left\{ (\mathbf{I}_n - \mathbf{w}\mathbf{w}^T) \left(\sum_{k=1}^n \frac{\partial \mathbf{H}}{\partial x_k} \cdot g_k \right) + \underbrace{[\mathbf{H}^2 - (\mathbf{w}^T \mathbf{H}\mathbf{w}) \mathbf{H}]_{ji}}_{=\mathbf{H}^2 - \lambda \mathbf{H} \text{ nach (2.9)}} \right\} .
 \end{aligned} \tag{B.7}$$

Wegen $\mathbf{p} = \mathbf{0}$ gilt $\mathbf{w}^T \mathbf{p} = 0$ und somit auch $d\mathbf{w}^T \mathbf{p} + \mathbf{w}^T d\mathbf{p} = 0$. Da aber $\mathbf{p} = \mathbf{0}$ ist, gilt insbesondere, daß $\mathbf{w}^T d\mathbf{p} = 0$. Wählen wir nun die lokalen Koordinaten so, daß x_1, \dots, x_n parallel zu den Eigenvektoren von \mathbf{H} liegen und speziell x_1 parallel zu \mathbf{g} , dann erhalten wir für die partiellen Ableitungen von \mathbf{p} :

$$\frac{\partial p_j}{\partial x_i} = \begin{cases} 0 & \text{für } i = j = 1 \\ \frac{\partial H_{ji}}{\partial x_1} \|\mathbf{g}\|_2 + \lambda_j (\lambda_j - \lambda_1) \delta_{ji} & \text{für } j \geq 2 \end{cases} . \tag{B.8}$$

B. Rechnungen

Befinden wir uns hingegen nicht auf einer Gradient Extremal, dann erhalten wir aus (B.3) und (B.5) durch Einsetzen

$$\frac{\partial p_j}{\partial x_i} = \left\{ \left(\mathbf{I}_n - \mathbf{w}\mathbf{w}^T \right) \left[\sum_{k=1}^n \frac{\partial \mathbf{H}}{\partial x_k} \cdot g_k + \mathbf{H}^2 - (\mathbf{w}^T \mathbf{H} \mathbf{w}) \mathbf{H} \right] - \mathbf{w}\mathbf{w}^T \left[\mathbf{H}^2 - (\mathbf{w}^T \mathbf{H} \mathbf{w}) \mathbf{H} \right] \right\}_{ji} .$$

Wählen wir die lokalen Koordinaten x_1, \dots, x_n wie oben beschrieben, dann können wir wie folgt vereinfachen:

$$\frac{\partial p_j}{\partial x_i} = \begin{cases} -(\mathbf{H}^2)_{1i} + H_{11}H_{1i} & \text{für } j = 1 \\ \left(\frac{\partial H_{ji}}{\partial x_1} \right) \|\mathbf{g}\|_2 + (\mathbf{H}^2)_{ji} - H_{11}H_{ji} & \text{für } j \geq 2 \end{cases} . \quad (\text{B.9})$$

B.3. Minimaler Abstand zwischen Punkt und Gerade

Ziel ist es, den geringsten Abstand zwischen einem gegebenen Punkt \mathbf{x}' und einer durch $(n-1)$ linear unabhängige Gleichungen gegebenen Gerade zu ermitteln. Wir verwenden die Lagrangesche Multiplikatorenmethode. Das Problem sei wie folgt formuliert:

$$\|\mathbf{x} - \mathbf{x}'\|_2 \longrightarrow \text{Min!} \quad \sum_{i=1}^n M_{ji}(x_i - x'_i) = b_j$$

für $j = 1, \dots, n-1$. Wir setzen $\xi_i := x_i - x'_i$ und führen Lagrangemultiplikatoren λ_j ein. Dann gilt

$$0 = \frac{\partial}{\partial x_i} \left[\frac{1}{2} \sum_{i=1}^n \xi_i^2 - \sum_{j=1}^{n-1} \lambda_j \left(\sum_{i=1}^n M_{ji} \xi_i - b_j \right) \right] = \sum_{i=1}^n \xi_i - \sum_{j=1}^{n-1} \lambda_j \left(\sum_{i=1}^n M_{ji} \right) ,$$

$$\xi_i = \sum_{j=1}^{n-1} \lambda_j M_{ji} . \quad (\text{B.10})$$

Daraus erhalten wir

$$b_j = \sum_{i=1}^n M_{ji} (x_i - x'_i) = \sum_{i=1}^n M_{ji} \xi_i = \sum_{i=1}^n M_{ji} \left(\sum_{k=1}^{n-1} M_{ki} \lambda_k \right) = \sum_{k=1}^{n-1} (\mathbf{M}\mathbf{M}^T)_{jk} \lambda_k ,$$

$$\lambda_j = \sum_{k=1}^{n-1} (\mathbf{M}\mathbf{M}^T)_{jk}^{-1} b_k . \quad (\text{B.11})$$

Damit können wir den gesuchten Punkt \mathbf{x} ermitteln:

$$x_i = x'_i + \xi_i \stackrel{(\text{B.10})}{=} x'_i + \sum_{j=1}^{n-1} \lambda_j M_{ji} \stackrel{(\text{B.11})}{=} \sum_{j=1}^{n-1} \sum_{k=1}^{n-1} M_{ji} (\mathbf{M}\mathbf{M}^T)_{jk}^{-1} b_k ,$$

$$\mathbf{x} = \mathbf{x}' + \mathbf{M}^T (\mathbf{M}\mathbf{M}^T)^{-1} \mathbf{b} . \quad (\text{B.12})$$

C. Quelltext

```
1      PROGRAM vrisearch
2      !
3      ! Programm bestimmt auf der PES
4      !
5      !  $E(x, y, z) = 2y + y^2 + (y + 0.4x^2 + z^2)x^2 + 0.01z^2$ 
6      !
7      ! Suchrichtung der Newtontrajektorie, die den VRI-Punkt enthält
8      !
9      ! Benjamin Schmidt, Juli 2009
10     !
11     !     n = Dimension
12     ! length = Kettenlänge
13     !     stl = Schrittlänge
14     !     eps = Genauigkeit des RGF-Verfahrens
15     ! chain = Kette zwischen den Punkten xstart und xend
16     !     dir = Suchrichtung für RGF-Verfahren
17     !
18     ! Das Hauptprogramm geht in folgenden Schritten vor:
19     ! (1) Fixiere Start- und Endpunkt der Kette.
20     ! (2) Lege Suchrichtung "dir" als Differenz zwischen beiden fest.
21     ! (3) Lege gerade Kette der Länge "length" zwischen beide Punkte.
22     ! (4) Verteile Kettenpunkte auf der zu "dir" gehörigen NT.
23     ! (5) Bestimme VRI-nahen Punkt, setze dessen Gradient als neues "dir" fest.
24     ! (6) Wiederhole (3)-(6) hinreichend oft.
25     ! (7) Schreibe iterierte Kette in "chain.dat".
26     !
27     INTEGER n, length
28     DOUBLE PRECISION stl, eps
29
30     PARAMETER (n=3, length=50, stl=0.125d0, eps=0.00000001d0)
31
32     DOUBLE PRECISION chain(length+1, n)
33     DOUBLE PRECISION xstart(n), xend(n), dir(n)
34     DOUBLE PRECISION H(n, n), A(n, n), g(n), norm, Ag(n)
35
36     OPEN(7, FILE='chain.dat')
37
38     xstart(1)=-.2d0
39     xstart(2)=-.5d0
40     xstart(3)=.0d0
41     xend(1)=.4d0
42     xend(2)=.3d0
43     xend(3)=0.d0
44     DO 5 i=1, n
45         dir(i)=xend(i)-xstart(i)
46 5     CONTINUE
47
48     DO 10 i=1, 20
49         CALL straight_chain(chain, xstart, xend, n, length)
50         CALL rgf(chain, dir, stl, eps, n, length)
51         CALL getvridir(dir, chain, chain, n, length)
52 10    CONTINUE
53
```

C. Quelltext

```
54 100  FORMAT(A, f20.10, A, F20.10, A, F20.10, A)
55      DO 20 j=1, length+1
56          WRITE(7,100) 'draw dddpoint(' , chain(j,1) , ', ', chain(j,2) , ', ',
57      *          , chain(j,3) , ');'
58 20    CONTINUE
59      END
60 !
61 ! RGF-Verfahren, wie in der Arbeit beschrieben.
62 !
63      SUBROUTINE rgf(kette, r, stl, eps, nn, l)
64      DOUBLE PRECISION kette(l+1, nn), origkette(l+1, nn), stl, eps
65      DOUBLE PRECISION r(nn), g(nn), Pr(nn, nn), p(nn), sgn
66      DOUBLE PRECISION xrgf(nn), xanf(nn), xend(nn)
67      DOUBLE PRECISION distance, norm
68      sgn=-1.d0
69      CALL projector(Pr, r, nn)
70      DO 11 j=1, l+1
71          DO 10 i=1, nn
72              xanf(i)=kette(1, i)
73              xend(i)=kette(l+1, i)
74              origkette(j, i)=kette(j, i)
75 10      CONTINUE
76 11      CONTINUE
77          j=2
78 5      CONTINUE
79      CALL gradient(g, kette(j, 1), kette(j, 2), kette(j, 3), nn)
80      CALL multiply(p, Pr, g, nn)
81      DO 15 i=1, nn
82          xrgf(i)=kette(j, i)+sgn*stl*p(i)
83 15      CONTINUE
84      IF((distance(xrgf, xanf, nn).GT.1.d0*distance(xanf, xend, nn)).OR.
85 #      (distance(xrgf, xend, nn).GT.1.d0*distance(xanf, xend, nn))) THEN
86          IF(sgn.GT.0.d0) THEN
87              DO 30 i=1, nn
88                  kette(j, i)=kette(j-1, i)
89 30          CONTINUE
90              IF(j.LT.1) THEN
91                  sgn=-1.d0
92                  j=j+1
93                  GO TO 5
94              ENDIF
95          ELSE
96              sgn=1.d0
97              DO 40 i=1, nn
98                  kette(j, i)=origkette(j, i)
99 40          CONTINUE
100             GO TO 5
101         ENDF
102     ELSE
103         DO 20 i=1, nn
104             kette(j, i)=xrgf(i)
105 20     CONTINUE
106         IF (norm(p, nn).Ge. eps) THEN
107             GO TO 5
108         ELSE
109             IF(j.LT.1) THEN
```

```

110 |                                     sgn=-1.d0
111 |                                     j=j+1
112 |                                     GOTO 5
113 |                               ENDIF
114 |       ENDIF
115 |   ENDIF
116 |   END
117 | !
118 | ! Bestimmung des VRI-nahen Punktes, wie in Arbeit beschrieben. "Dir" ist
119 | ! Gradient des VRI-nahen Punktes, der durch Verbindungen zwischen allen
120 | ! Punkten der linken Kette "lkette" und der rechten Kette "rkette"
121 | ! ermittelt wird.
122 | !
123 |   SUBROUTINE getvridir(dir, lkette, rkette, nn, l)
124 |   DOUBLE PRECISION lkette(l+1,nn), rkette(l+1,nn), kette(l+1,nn)
125 |   DOUBLE PRECISION dir(nn), xstart(nn), xend(nn), vri(nn)
126 |   DOUBLE PRECISION minpt(nn), gminpt(nn), minimum
127 |   DOUBLE PRECISION Hminpt(nn,nn), Aminpt(nn,nn), Agminpt(nn)
128 |   DOUBLE PRECISION norm
129 |   INTEGER ind, min_Ag
130 |   minimum=10.d0
131 |   DO 50 k=1,l+1
132 |     DO 40 j=1,l+1
133 |       DO 10 i=1,nn
134 |         xstart(i)=lkette(k,i)
135 |         xend(i)=rkette(j,i)
136 | 10      CONTINUE
137 |       CALL straight_chain(kette, xstart, xend, nn, l)
138 |       ind=min_Ag(kette, nn, l)
139 |       DO 20 i=1,nn
140 |         minpt(i)=kette(ind, i)
141 | 20      CONTINUE
142 |       CALL gradient(gminpt, minpt(1), minpt(2), minpt(3), nn)
143 |       IF (norm(gminpt, nn).GT..001d0) THEN
144 |         CALL hessian(Hminpt, minpt(1), minpt(2),
145 | *           minpt(3), nn)
146 |         CALL adjoint(Aminpt, Hminpt, nn)
147 |         CALL multiply(Agminpt, Aminpt, gminpt, nn)
148 |         IF (norm(Agminpt, nn).LT.minimum) THEN
149 |           minimum=norm(Agminpt, nn)
150 |           DO 30 i=1,nn
151 |             vri(i)=minpt(i)
152 | 30          CONTINUE
153 |         ENDIF
154 |       ENDIF
155 | 40      CONTINUE
156 | 50      CONTINUE
157 |       WRITE(*,*) 'VRI-Punkt bei ', (vri(i), i=1,nn)
158 |       CALL gradient(dir, vri(1), vri(2), vri(3), nn)
159 |       WRITE(*,*) 'Suchrichtung der VRI-NT', (dir(i), i=1,nn)
160 |     END
161 | !
162 | ! Bestimmt den Kettenpunkt, für den |Ag| minimal wird, wobei |g|>0.001
163 | ! gelten soll.
164 | !
165 |   INTEGER FUNCTION min_Ag(kette, nn, l)

```

C. Quelltext

```

166     DOUBLE PRECISION g(nn),H(nn,nn),A(nn,nn),Ag(nn)
167     DOUBLE PRECISION kette(l+1,nn),norm,min_norm
168     min_Ag=1/2
169     CALL gradient(g,kette(1/2,1),kette(1/2,2),kette(1/2,3),nn)
170     CALL hessian(H,kette(1/2,1),kette(1/2,2),kette(1/2,3),nn)
171     CALL adjoint(A,H,nn)
172     CALL multiply(Ag,A,g,nn)
173     min_norm=norm(Ag,nn)
174     DO 10 j=1,l+1
175         CALL gradient(g,kette(j,1),kette(j,2),kette(j,3),nn)
176         IF (norm(g,nn).GT.0.001d0) THEN
177             CALL hessian(H,kette(j,1),kette(j,2),kette(j,3),nn)
178             CALL adjoint(A,H,nn)
179             CALL multiply(Ag,A,g,nn)
180             IF (norm(Ag,nn).LT.min_norm) THEN
181                 min_norm=norm(Ag,nn)
182                 min_Ag=j
183             ENDIF
184         ENDIF
185 10    CONTINUE
186     END
187     !
188     ! Berechnet den Projektionsoperator "pr" aus der Suchrichtung "r" gemäß
189     !   pr=-(Id-rr^T)
190     !
191     SUBROUTINE projector(pr,r,nn)
192     DOUBLE PRECISION r(nn),pr(nn,nn)
193     DOUBLE PRECISION nrm,norm
194     nrm=1.d0*norm(r,nn)
195     DO 5 i=1,nn
196         r(i)=r(i)/nrm
197 5     CONTINUE
198     DO 20 i=1,nn
199         DO 10 j=1,nn
200             IF (i.EQ.j) THEN
201                 pr(i,j)=-r(i)*r(j)+1.d0
202             ELSE
203                 pr(i,j)=-r(i)*r(j)
204             ENDIF
205 10    CONTINUE
206 20    CONTINUE
207     END
208     !
209     ! Legt gerade Kette "kette" der Länge "l" zwischen die Punkte "x" und
210     ! "y".
211     !
212     SUBROUTINE straight_chain(kette,x,y,nn,l)
213     DOUBLE PRECISION kette(l+1,nn),x(nn),y(nn)
214     DO 20 j=1,l+1
215         DO 10 i=1,nn
216             kette(j,i)=x(i)+((j-1)*(y(i)-x(i)))/l
217 10    CONTINUE
218 20    CONTINUE
219     END
220     !
221     ! Berechnet Gradient g im Punkt (x,y,z)

```

```

222 !
223     SUBROUTINE gradient (g,x,y,z,nn)
224     DOUBLE PRECISION g(nn),x,y,z
225     g(1)=1.6d0*x**3+2.d0*x*y+2.d0*x*z**2
226     g(2)=2.d0+2.d0*y+x**2
227     g(3)=2.d0*x**2*z+0.02d0*z
228     END
229 !
230 ! Berechnet Hessematrix H im Punkt (x,y,z)
231 !
232     SUBROUTINE hessian(H,x,y,z,nn)
233     DOUBLE PRECISION H(nn,nn),x,y,z
234     H(1,1)=4.8d0*x**2+2.d0*y+2.d0*z**2
235     H(1,2)=2.d0*x
236     H(1,3)=4.d0*x*z
237     H(2,1)=H(1,2)
238     H(2,2)=2.d0
239     H(2,3)=0.d0
240     H(3,1)=H(1,3)
241     H(3,2)=H(2,3)
242     H(3,3)=0.02d0+2.d0*x**2
243     END
244 !
245 ! Berechnet adjunkte Matrix A zu M
246 !
247     SUBROUTINE adjoint(A,M,nn)
248     DOUBLE PRECISION A(nn,nn),M(nn,nn)
249     A(1,1)=M(2,2)*M(3,3)-M(2,3)*M(3,2)
250     A(1,2)=M(1,3)*M(3,2)-M(1,2)*M(3,3)
251     A(1,3)=M(1,2)*M(2,3)-M(1,3)*M(2,2)
252     A(2,1)=M(2,3)*M(3,1)-M(2,1)*M(3,3)
253     A(2,2)=M(1,1)*M(3,3)-M(1,3)*M(3,1)
254     A(2,3)=M(1,3)*M(2,1)-M(1,1)*M(2,3)
255     A(3,1)=M(2,1)*M(3,2)-M(2,2)*M(3,1)
256     A(3,2)=M(1,2)*M(3,1)-M(1,1)*M(3,2)
257     A(3,3)=M(1,1)*M(2,2)-M(1,2)*M(2,1)
258     END
259 !
260 ! Multiplikation Ax=b
261 !
262     SUBROUTINE multiply(b,A,x,nn)
263     DOUBLE PRECISION b(nn),A(nn,nn),x(nn)
264     DO 20 i=1,nn
265         b(i)=0
266         DO 10 j=1,nn
267             b(i)=b(i)+A(i,j)*x(j)
268 10     CONTINUE
269 20     CONTINUE
270     END
271 !
272 ! Gibt die Norm von v aus
273 !
274     DOUBLE PRECISION FUNCTION norm(v,nn)
275     DOUBLE PRECISION v(nn),s
276     s=0.d0
277     DO 10 j=1,nn

```

C. Quelltext

```
278         s=s+v(j)**2
279 10    CONTINUE
280     IF (s.LT.1.D-15) THEN
281         norm=0.d0
282     ELSE
283         norm=DSQRT(s)
284     ENDIF
285     END
286 !
287 ! Gibt die Distanz zweier Punkte v und w aus
288 !
289     DOUBLE PRECISION FUNCTION distance(v,w,nn)
290     DOUBLE PRECISION v(nn),w(nn),s
291     s=0.d0
292     DO 10 j=1,nn
293         s=s+(v(j)-w(j))**2
294 10    CONTINUE
295     distance=DSQRT(s)
296     END
```

Abkürzungsverzeichnis

BP	Bifurcation Point (Bifurkationspunkt)
CR _x	Catchment-Region des Punktes x
DC	Distinguished Coordinate
GE	Gradient Extremal
GS	Growing String
IRC	Intrinsic Reaction Coordinate
MEP	Minimum Energy Path (Minimum-Energie-Weg)
Min	Minimum
NC _x	Newton Channel (Newtonkanal) des Punktes x
NT	Newtontrajektorie
PES	Potential Energy Surface (Potentialenergiefläche)
RC	Reaction Coordinate (Reaktionskoordinate)
RGF	Reduced Gradient Following
RP	Reaction Path (Reaktionspfad)
SD	Steepest Descent (steilster Abstieg)
SP	Sattelpunkt (vom Index 1)
TASC	Tangent Search Concept
TP	Turning Point (Umkehrpunkt)
TS	Transition Structure (Übergangsstruktur)
TST	Transition State Theory (Theorie des Übergangszustands)
VRI	Valley-Ridge Inflection Point (Tal-Rücken-Umschlagpunkt)
VRI-NT	Newtontrajektorie durch den VRI-Punkt

Abbildungsverzeichnis

1.1. Zweidimensionale Potentialenergieflächen als Höhenlinienbilder	11
1.2. Stationäre Punkte des Müller-Brown-Potentials	14
1.3. Symmetrische und asymmetrische VRI-Punkte	18
2.1. Vergleich von theoretischer und numerischer IRC	22
2.2. Bestimmung der IRC mithilfe der String-Methode	26
2.3. Gradient Extremals und konfluente SD-Kurven	32
2.4. Gradient Extremals mittels String-Methode	36
2.5. Umkehrpunkte auf einer Modell-Potentialenergiefläche [34]	37
2.6. Wende-, Bifurkations- und VRI-Punkte auf den Gradient Extremals des Müller-Brown-Potentials	38
3.1. Newtontrajektorien auf dem Müller-Brown-Potential	44
3.2. Newtontrajektorie mit ihren Zweigen und Komponenten	45
3.3. NT-Verzweigungen auf dem Müller-Brown-Potential	46
3.4. Newtontrajektorie durch asymmetrischen VRI-Punkt	48
3.5. Gradient Extremals und Newtontrajektorien auf der modifizierten PES nach Neria, Fischer und Karplus	49
3.6. String-Methode zur RGF-Bestimmung auf dem Müller-Brown-Potential . .	52
3.7. TASC-Kurven an einem „Don-Quixote“-Sattelpunkt	54
4.1. Kurven des steilsten Abstiegs im Vektorfeld $-\mathbf{g}$	62
4.2. Newtontrajektorien im Vektorfeld $\mathbf{A}\mathbf{g}$	64
5.1. String-Methode für Newtontrajektorien in VRI-Punkt-Nähe	67
5.2. Modifizierte String-Methode	68
5.3. Etappen des Algorithmus	69
5.4. Ergebnisse für die PES-Familie (5.5)	72
5.5. Ergebnisse für das Müller-Brown-Potential	75
5.6. Ergebnisse für das Müller-Brown-Potential (1. Fortsetzung)	76
5.7. Ergebnisse für das Müller-Brown-Potential (2. Fortsetzung)	77
5.8. Verdeckter VRI-Punkt auf dem Müller-Brown-Potential	78
5.9. VRI-Bestimmung auf Schnittflächen einer dreidimensionalen PES	78
5.10. Ergebnisse für die dreidimensionale Potentialenergiefläche (5.8)	80

Tabellenverzeichnis

5.1. Ergebnisse für die PES-Familie (5.5)	71
5.2. VRI-Punkte auf dem Müller-Brown-Potential nach [19]	73
5.3. Ergebnisse für das Müller-Brown-Potential	74
5.4. Ergebnisse für die dreidimensionale Potentialenergiefläche (5.8)	79

Literatur

- [1] Allgower, E.L. und K. Georg: *Numerical Continuation Methods – An Introduction*. Springer, 1990.
- [2] Basilevsky, M.V.: *The structural stability principle and branching points on multidimensional potential energy surfaces*. Theor Chim Acta, 72:63–67, 1987.
- [3] Basilevsky, M.V. und A.G. Shamov: *The local definition of the optimum ascent path on a multi-dimensional potential energy surface and its practical application for the location of saddle points*. Chemical Physics, 60:347–358, 1981.
- [4] Branin, F.H.: *Widely convergent methods for finding multiple solutions of simultaneous nonlinear equations*. IBM J Res Develop, 16:504–522, 1972.
- [5] Castaño, O., R. Palmeiro, L.M. Frutos und J. Luisandrés: *Role of Bifurcation in the Bond Shifting of Cyclooctatetraene*. J Comput Chem, 23:732–735, 2002.
- [6] Clegg, J.C.: *Variationsrechnung*. Teubner, Stuttgart, 1970.
- [7] Cramer, C.J.: *Essentials of Computational Chemistry. Theories and Models*. Wiley & Sons, 2. Auflage, 2004.
- [8] Crehuet, R. und J.M. Bofill: *The reaction path intrinsic reaction coordinate method and the Hamilton-Jacobi theory*. J Chem Phys, 122(234105), 2005.
- [9] Diener, I.: *Globale Aspekte des kontinuierlichen Newtonverfahrens*. Habilitation, Göttingen, 1991.
- [10] Diener, I. und R. Schaback: *An Extended Continuous Newton Method*. Journal of Optimization Theory and Applications, 67:57–77, 1990.
- [11] Ess, D.H., S.E. Wheeler, R.G. Iafe, L. Xu, N. Çelebi-Ölçüm und K.N. Houk: *Bifurcations on Potential Energy Surfaces of Organic Reactions*. Angew Chem Int Ed, 47:7592–7601, 2008.
- [12] Eyring, H. und M. Polanyi: *Über einfache Gasreaktionen*. Z physikal Chem, B12:279–311, 1931.
- [13] Fukui, K.: *A Formulation of the Reaction Coordinate*. J Phys Chem, 74:4161–4163, 1970.
- [14] Gelfand, I.M. und S.V. Fomin: *Calculus of Variations*. Dover, 2000.

- [15] Goddard, J.D. und H.F. Schaefer III: *The photodissociation of formaldehyde: Potential energy surface features*. J Chem Phys, 70:5117–5134, 1979.
- [16] Gomulka, J.: *Remarks on Branin's Method for Solving Nonlinear Equations*. In: Dixon, L.C.W. und G.P. Szegö (Herausgeber): *Towards Global Optimisation*, Seiten 96–106. North-Holland, 1974.
- [17] Heidrich, D.: *An Introduction to the Nomenclature and Usage of the Reaction Path Concept*. In: Heidrich, D. (Herausgeber): *The Reaction Path in Chemistry: Current Approaches and Perspectives*, Seiten 1–10. Kluwer Academic Publishers, 1995.
- [18] Heidrich, D., W.Kliesch und W.Quapp: *Properties of Chemically Interesting Potential Energy Surfaces*, Band 56 der Reihe *Lecture Notes in Chemistry*. Springer, 1991.
- [19] Hirsch, M.: *Zum Reaktionswegcharakter von Newtontrajektorien*. Dissertation, Universität Leipzig, 2004.
- [20] Hirsch, M. und W. Quapp: *Improved RGF Method to Find Saddle Points*. J Comput Chem, 23:887–894, 2002.
- [21] Hirsch, M., W. Quapp und D. Heidrich: *The set of valley-ridge inflection points on the potential energy surface of water*. Phys Chem Chem Phys, 1:5291–5299, 1999.
- [22] Hirst, D.M.: *A Computational Approach to Chemistry*. Blackwell Science Inc, 1990.
- [23] Hoffman, D.K., R.S. Nord und K. Ruedenberg: *Gradient Extremals*. Theor Chim Acta, 69:265–279, 1986.
- [24] Imig, O., D. Heidrich und W. Quapp: *Gradient Extremal Turning Points and Bifurcations*. unveröffentlicht.
- [25] Jarre, F. und J. Stoer: *Optimierung*. Springer, 2004.
- [26] Jensen, F.: *Introduction to Computational Chemistry*. Wiley & Sons, 2. Auflage, 2006.
- [27] Jongen, H.Th., P. Jonker und F. Twilt: *Nonlinear Optimization in Finite Dimensions*. Kluwer Academic, Dordrecht Boston London, 2000.
- [28] McIver, Jr., J.W. und A. Komornicki: *Structure of Transition States in Organic Reactions. General Theory and an Application to the Cyclobutene-Butadiene Isomerization Using a Semiempirical Molecular Orbital Method*. J Amer Chem Soc, 94:2625–2633, 1972.

Literatur

- [29] Mezey, P.G.: *Potential energy hypersurfaces*. Elsevier, Amsterdam, 1987.
- [30] Moser, J.: *Selected Chapters in the Calculus of Variations*. Lectures in Mathematics. ETH Zürich. Birkhäuser, 2003.
- [31] Müller, K. und L.D. Brown: *Location of Saddle Points and Minimum Energy Paths by a Constrained Simplex Optimization Procedure*. Theor Chim Acta, 53:75–93, 1979.
- [32] Neria, E., S. Fischer und M. Karplus: *Simulation of activation energies in molecular systems*. J Chem Phys, 105:1902–1921, 1996.
- [33] Pancíř, J.: *Calculation of the least energy path on the energy hypersurface*. Collection Czechoslov Chem Commun, 40:1112–1118, 1975.
- [34] Perković, S., E.M. Blokhuis, E. Tessler und B. Widom: *Boundary tension: From wetting transition to prewetting critical point*. J Chem Phys, 102:7584–7594, 1995.
- [35] Quapp, W.: *A Valley Following Method*. Optimization, 52:317–331, 2003.
- [36] Quapp, W.: *Reduced Gradient Methods and their Relation to Reaction Paths*. J Theor Comput Chem, 2:385–417, 2003.
- [37] Quapp, W.: *Reaction Pathways and Projection Operators: Application to String Methods*. J Comput Chem, 25:1277–1285, 2004.
- [38] Quapp, W.: *A growing string method for the reaction pathway defined by a Newton trajectory*. J Chem Phys, 122(174106), 2005.
- [39] Quapp, W.: *Chemical Reaction Paths and Calculus of Variations*. Theor Chem Acc, 121:227–237, 2008.
- [40] Quapp, W.: *The growing string method for flows of Newton trajectories by a second order method*. J Theor Comput Chem, 8:101–117, 2009.
- [41] Quapp, W., M. Hirsch und D. Heidrich: *Bifurcation of reaction pathways: the set of valley ridge inflection points of a simple three-dimensional potential energy surface*. Theor Chem Acc, 100:285–299, 1998.
- [42] Quapp, W., M. Hirsch und D. Heidrich: *Following the streambed reaction on potential-energy surfaces: a new robust method*. Theor Chem Acc, 105:145–155, 2000.
- [43] Quapp, W., M. Hirsch und D. Heidrich: *An approach to reaction path branching using valley-ridge inflection points of potential energy surfaces*. Theor Chem Acc, 112:40–51, 2004.

- [44] Quapp, W., M. Hirsch, O. Imig und D. Heidrich: *Searching for Saddle Points of Potential Energy Surfaces by Following a Reduced Gradient*. J Comput Chem, 19:1087–1100, 1998.
- [45] Quapp, W., O. Imig und D. Heidrich: *Gradient Extremals and their Relation to the Minimum Energy Path*. In: Heidrich, D. (Herausgeber): *The Reaction Path in Chemistry: Current Approaches and Perspectives*, Seiten 137–160. Kluwer Academic Publishers, 1995.
- [46] Rothman, M.J. und L.L. Lohr, Jr.: *Analysis of an energy minimization method for locating transition states on potential energy hypersurfaces*. Chem Phys Lett, 70:405–409, 1980.
- [47] Sun, J. Q. und K. Ruedenberg: *Gradient Extremals and steepest descent lines on potential energy surfaces*. J Chem Phys, 98:9707–9714, 1993.
- [48] Valtazanos, P. und K. Ruedenberg: *Bifurcations and transition states*. Theor Chim Acta, 69:281–307, 1986.
- [49] Williams, I.H. und G.M. Maggiora: *Use and abuse of the distinguished-coordinate method for transition-state structure searching*. J Mol Struct (Theochem), 89:365–378, 1982.
- [50] Zech, A.: *Modifikation der Transition State Theory durch die Tsallis-Statistik*. Diplomarbeit, Universität Leipzig, 2009.

Danksagung

Danksagung

Mein herzlicher Dank gilt Dr. Wolfgang Quapp für die zahlreichen fruchtbaren Gespräche, Ratschläge, Anregungen und Aufmunterungen. Dank vor allem auch für seine mir entgegengebrachte Geduld.

Erklärung

Die vorliegende Arbeit habe ich selbständig und nur unter Verwendung der angegebenen Quellen und Hilfsmittel angefertigt. Wörtliche und sinngemäße Zitate sind als solche gekennzeichnet.

Leipzig, am 15. Oktober 2009