

Untersuchung der Reversibilitätsverletzung beim Hybrid-Monte-Carlo-Algorithmus

Carsten Urbach



Diplomarbeit
Fachbereich Physik
Freie Universität Berlin
September 2002

Für meine Eltern.

Diese Arbeit ist entstanden unter Betreuung von:

**Dr. Karl Jansen (NIC/DESY Zeuthen) und
Prof. Dr. Volkard Linke (Freie Universität Berlin)**

Inhaltsverzeichnis

| | |
|--|-----------|
| Einleitung | V |
| 1 Theoretische Grundlagen | 1 |
| 1.1 Quantenfeldtheorie | 1 |
| 1.1.1 Pfadintegral-Formalismus | 2 |
| 1.1.2 Euklidisierung | 2 |
| 1.1.3 Observablen | 3 |
| 1.2 Feldtheorie auf dem Gitter | 4 |
| 1.3 Anharmonischer Oszillator | 6 |
| 1.3.1 Anharmonischer Oszillator auf dem Gitter | 6 |
| 1.4 Nichtlineares $O(n)$ - σ -Modell | 8 |
| 1.4.1 Nichtlineares $O(n)$ - σ -Modell auf dem Gitter | 9 |
| 1.4.2 Erwartungswert der Energie | 9 |
| 1.4.3 Asymptotische Freiheit | 10 |
| 1.5 Lyapunov-Exponent | 12 |
| 2 Simulationsverfahren | 15 |
| 2.1 HMC-Algorithmus | 15 |
| 2.1.1 Markov-Ketten | 15 |
| 2.1.2 Metropolis-Algorithmus | 17 |
| 2.1.3 Molekular-Dynamik-Methode | 18 |
| 2.1.4 HMC-Algorithmus | 20 |
| 2.1.5 Observablen | 24 |
| 2.1.6 Einstell-Parameter | 24 |
| 2.2 Reversibilitatsverletzung | 26 |
| 2.3 Korrelationslange | 27 |
| 2.4 Fehlerabschatzung | 29 |

| | | |
|----------|--|-----------|
| 3 | Anwendung des HMC | 33 |
| 3.1 | HMC-Algorithmus für den anharmonischen Oszillator | 33 |
| 3.2 | HMC-Algorithmus für das nichtlineare $O(4)$ - σ -Modell | 34 |
| 3.2.1 | Lagrangesche Multiplikatoren | 34 |
| 3.2.2 | Darstellung mit $SU(2, \mathbb{C})$ | 36 |
| 4 | Ergebnisse und Diskussion | 41 |
| 4.1 | Verifikation der Implementierung | 44 |
| 4.2 | Bestimmung des <i>massgap</i> | 46 |
| 4.2.1 | Asymptotisches Skalierungsverhalten | 48 |
| 4.3 | Lyapunov-Exponent | 50 |
| 4.3.1 | Nichtlineares σ -Modell | 50 |
| 4.3.2 | Anharmonischer Oszillator | 54 |
| 4.4 | Auswirkung auf Observablen | 57 |
| 4.4.1 | Wirkung | 58 |
| 4.4.2 | Erwartungswert der Energie | 62 |
| 4.4.3 | Hamilton-Funktion | 64 |
| 4.4.4 | Korrelationsfunktion | 65 |
| 5 | Zusammenfassung und Ausblick | 69 |
| | Danksagung | 72 |
| A | Mathematischer Anhang | 73 |
| A.1 | Zwei-Punkt-Funktion | 73 |
| A.2 | Ergänzungen zu $SU(2, \mathbb{C})$ | 74 |
| A.3 | Haar-Maß | 77 |
| A.4 | <i>strong coupling expansion</i> | 78 |
| B | Numerischer Anhang | 80 |
| B.1 | <i>expression templates</i> in C++ | 80 |
| B.2 | Zufallszahlen-Generator | 81 |
| | Abbildungsverzeichnis | 83 |
| | Tabellenverzeichnis | 84 |
| | Literaturverzeichnis | 87 |

Einleitung

Seit der richtungweisenden Arbeit von K.G. Wilson aus dem Jahr 1974 [1] stellt die Regularisierung von Feldtheorien mit Hilfe eines Raum-Zeit-Gitters ein wichtiges nicht-störungstheoretisches Werkzeug der Quantenfeldtheorie dar, das mittlerweile auf viele verschiedene Modelle angewandt wurde, insbesondere auf die Quantenchromodynamik (QCD), für deren Behandlung die Gitterfeldtheorie entwickelt wurde. Ziel ist es, Resultate für physikalische Observablen zu erzeugen, um sie mit Meßwerten aus Experimenten vergleichen zu können und damit zu überprüfen, ob die Theorie, beispielsweise in der Hochenergiephysik das Standardmodell, die Natur richtig beschreibt.

Diese Art der Regularisierung macht es zugleich erst möglich, die in das Euklidische fortgesetzte Feldtheorie auf dem Computer zu bearbeiten, da der Computer nur mit endlichen diskreten Systemen umzugehen weiß. In Verbindung mit statistischen Methoden können auf dieser Grundlage Simulationen von Quantenfeldtheorien am Computer durchgeführt werden.

Um aus dem regularisierten System wieder die Physik extrahieren zu können, ist nach der Regularisierung durch einen Impuls-*cut-off* Λ eine Renormierung der Größen notwendig, um sie im Grenzwert $cut-off \rightarrow \infty$ zu erhalten. Das bedeutet für die Gittertheorie den Limes für den Gitterabstand $a \propto 1/\Lambda$ gegen Null bei konstanter physikalischer Skala. Um die physikalische Skala festzuhalten, muß bei diesem Grenzwert zugleich das Gitter vergrößert werden, was zu einer Zunahme der Rechenzeit führt.

In bisherigen Computer-Simulationen der QCD wurden Fermionen im allgemeinen als statische Teilchen („statische Fermionen“) behandelt, um einerseits eine annehmbar feine Diskretisierung nutzen zu können, andererseits aber die Rechenzeit zu begrenzen. Trotz dieser zunächst problematisch erscheinenden Näherung wurden damit Ergebnisse erzielt, die sehr nahe an den experimentell bestimmten Werten liegen - ein Resultat, das bis heute noch nicht wirklich verstanden ist.

Mittlerweile ist die Leistungsfähigkeit moderner Computer so gestiegen, daß es realistisch ist, mit dynamischen Fermionen zu arbeiten. Dafür müssen auch die verwendeten Algorithmen angepaßt werden: Der sogenannte Hybrid-Monte-Carlo (HMC) Algorithmus kommt zur Anwendung [2].

Beim HMC-Algorithmus handelt es sich um eine Kombination aus dem Metropolis-Monte-Carlo-Algorithmus und einem Molekular-Dynamik-Teil, in dem Bewegungsgleichungen

integriert werden. Er stellte die erste exakte¹ Methode zur Simulation dynamischer Fermionen dar, die mit akzeptablen Kosten umsetzbar ist.

Die Exaktheit des Algorithmus ist allerdings nur in dem Umfang gewährleistet, in dem die numerische Integration der Bewegungsgleichungen im Molekular-Dynamik-Teil reversibel durchgeführt werden kann. Dies ist wegen der unvermeidbaren Rundungsfehler im Computer allerdings keineswegs gewährleistet. Hinzu kommt, daß aufgrund der chaotischen Natur der zugrundeliegenden klassischen Bewegungsgleichungen der Rundungsfehler und damit die Verletzung der Reversibilität exponentiell mit der Integrationslänge zunimmt, was durch einen Lyapunov-Exponenten gekennzeichnet ist.

Daraus ergibt sich die Frage, ob der Lyapunov-Exponent eine charakteristische Eigenschaft der zugrundeliegenden Feldtheorie ist, und wie stark die Verletzung der Reversibilität die mit dem HMC-Algorithmus berechneten Ergebnisse beeinflusst. Der letzte Teil der Frage ist zumindest für kompliziertere Modelle, wie zum Beispiel die QCD, nicht einfach zu beantworten, da es nicht möglich ist, mit exakten Ergebnissen zu vergleichen.

Zum ersten Teil der Frage stellen Edwards et. al. in [3] die Hypothese auf, daß der Lyapunov-Exponent in Abhängigkeit von der Kopplungskonstanten des Modells proportional zur inversen Korrelationslänge ist, was bedeuten würde, daß der Exponent wie eine physikalische Größe skaliert. Allerdings reichen die in [3] angeführten Daten nicht zur Verifikation der Hypothese aus.

Eine sorgfältige Überprüfung dieser Hypothese sowie die Untersuchung der Auswirkung der Reversibilitätsverletzung auf die Ergebnisse einer Simulation ist Inhalt dieser Arbeit. Dafür wurde als Modell das zweidimensionale nichtlineare $O(4)$ - σ -Modell gewählt, weil es große Ähnlichkeiten mit der QCD aufweist, aber wegen der niedrigeren Dimension deutlich weniger Rechenaufwand benötigt. Die Übereinstimmungen bestehen insbesondere in der Eigenschaft der asymptotischen Freiheit und der nicht-pertubativen Erzeugung einer „Masse“, so daß in beiden Modellen ein ähnliches Skalierungsverhalten des Exponenten erwartet werden kann.

Als weiteres Modell wird der quantenmechanische anharmonische Oszillator untersucht, das aufgrund seiner Eindimensionalität mit wenig Aufwand behandelt werden kann, aber eine zum σ -Modell und zur QCD grundsätzlich unterschiedliche Wirkung aufweist. Es bietet sich damit die Möglichkeit, herauszufinden, ob ein vom Modell unabhängiges Verhalten des Exponenten beobachtet werden kann.

Abschließend soll, falls der Einfluß der Reversibilitätsverletzung nicht vernachlässigbar ist, wenn möglich ein Kriterium formuliert werden, mit dessen Hilfe in Abhängigkeit von den Parametern des Modells bereits aus der Reversibilitätsverletzung abgeschätzt werden kann, wie stark dieser Einfluß ist.

¹ exakt in dem Sinne, daß der Algorithmus keine systematischen Fehler enthält.

Aufbau der Arbeit

Die Arbeit ist wie folgt gegliedert:

In Kapitel 1 werden die für diese Arbeit wichtigen theoretischen Grundlagen der Quantenfeldtheorie, des nichtlinearen $O(4)$ - σ -Modells und des anharmonischen Oszillators beschrieben. Anschließend werden in Kapitel 2 die verwendeten Simulationsverfahren dargestellt. Dabei wird insbesondere der HMC-Algorithmus eingeführt. In Kapitel 3 wird dieser dann auf die beiden konkreten Modelle übertragen.

Im darauffolgenden Kapitel 4 werden die Ergebnisse dargestellt und diskutiert. Im Anschluß daran wird eine Zusammenfassung und ein Ausblick gegeben. Im Anhang werden Themen behandelt, die im eigentlichen Text keinen Platz mehr gefunden haben, aber trotzdem von Interesse sind. Der Quelltext der verwendeten Programme hätte allerdings den Rahmen des Anhangs gesprengt, und steht deshalb wie dieses Dokument unter

www.physik.fu-berlin.de/~urbach/diplom

zur Verfügung.

Kapitel 1

Theoretische Grundlagen

In der vorliegenden Arbeit sollen das nichtlineare zweidimensionale $O(4)$ -symmetrische σ -Modell und der anharmonische Oszillator mit Hilfe von nichtstörungstheoretischen Methoden untersucht werden. Dies wird mit Methoden der Gitter-Physik geschehen, die neben einer Diskretisierung des Modells eine Darstellung verwenden, die es ermöglicht, das Modell mit Methoden der statistischen Mechanik zu untersuchen. Letzteres wird durch den Pfadintegral-Formalismus möglich, der im ersten Abschnitt dieses Kapitels erläutert wird. In den darauf folgenden Abschnitten wird die diskrete Quantenfeldtheorie eingeführt und auf das nichtlineare $O(4)$ - σ -Modell und den anharmonischen Oszillator konkretisiert.

1.1 Quantenfeldtheorie

Im Unterschied zur klassischen Mechanik ist es in der Quantenmechanik lediglich möglich, Wahrscheinlichkeiten anzugeben, mit denen sich ein System in seinen möglichen Zuständen befindet. Beispielsweise ist der Weg von Punkt A nach Punkt B in der klassischen Mechanik durch das Hamiltonsche Wirkungsprinzip eindeutig festgelegt, wohingegen in der Quantenmechanik prinzipiell alle Pfade von A nach B möglich sind, wenn auch die meisten nur mit sehr geringer Wahrscheinlichkeit. Dieses Bild führt unmittelbar zum Pfadintegral-Formalismus, der von R. P. Feynman eingeführt wurde [4].

In den folgenden Abschnitten werden nur die für diese Arbeit wichtigen Ergebnisse rekapituliert¹.

¹ Einführungen in den Pfadintegral-Formalismus und die Einbettung in den Kontext der Quantenfeldtheorie finden sich beispielsweise bei H. Kleinert [5] und J. Glimm und A. Jaffe [6].

1.1.1 Pfadintegral-Formalismus

Die Tatsache, daß alle möglichen Zustände eines Systems auch zu den Erwartungswerten beitragen, drückt sich am besten durch den Pfadintegral-Formalismus aus, der ein eigenständiges, zur kanonischen Quantisierung äquivalentes Quantisierungsverfahren darstellt. Um die Übergangsamplitude von einem Anfangszustand i zum Zeitpunkt t zu einem Endzustand f zum Zeitpunkt t' zu berechnen, wird über alle möglichen Pfade zwischen diesen beiden Zuständen summiert:

$$\langle f, t' | i, t \rangle \propto \int \mathcal{D}q \exp \left(\frac{i}{\hbar} \int_t^{t'} d\tau L(q, \dot{q}) \right). \quad (1-1)$$

Das Maß $\mathcal{D}q$ wird an dieser Stelle nicht genauer definiert, sondern soll intuitiv die Integration über alle Pfade darstellen. Zu einem späteren Zeitpunkt wird eine präzise Definition möglich. Bemerkenswert ist, daß offensichtlich auch die quantenmechanische Übergangsamplitude in dieser Formulierung vollständig durch die klassische Wirkung

$$S(q, \dot{q}, t, t') = \int_t^{t'} d\tau L(q, \dot{q})$$

bestimmt ist.

Diese Formulierung kann auf skalare Felder ϕ übertragen werden (siehe z.B. [7]), so daß die zu (1-1) analoge Formel lautet:

$$W := \langle \phi', t' | \phi, t \rangle \propto \int \mathcal{D}\phi \exp \left(\frac{i}{\hbar} S(\phi, \dot{\phi}) \right), \quad (1-2)$$

wobei die Wirkung wie folgt definiert ist:

$$S(\phi, \dot{\phi}) = \int d^4x \mathcal{L}(\phi, \dot{\phi}). \quad (1-3)$$

1.1.2 Euklidisierung

Unter bestimmten Voraussetzungen ist es möglich, in der Zeit-Komponente analytisch zu rein imaginären Werten fortzusetzen („Wick-Rotation“). Die Voraussetzungen werden im Rekonstruktionstheorem von K. Osterwalder und R. Schrader [8] angegeben. Das bedeutet eine analytische Fortsetzung in der komplexen t -Ebene, so daß t in $i\tau$ übergeht und damit eine euklidische Metrik entsteht. Das hat den Vorteil, daß der Phasenfaktor im Pfadintegral (1-2) in ein Boltzmanngewicht übergeht:

$$W_E \propto \int \mathcal{D}\phi e^{-\frac{1}{\hbar} S_E}. \quad (1-4)$$

Dabei ist S_E die euklidisierte Version der ursprünglichen Wirkung. An dieser Stelle wird eine Analogie zur statistischen Physik deutlich: $\exp(-S_E/\hbar)$ ist ein Boltzmann-Faktor, so daß die Übergangsamplitude W_E formal als Zustandssumme interpretiert werden kann, die aus der statistischen Physik als

$$Z = \sum_n e^{-\beta E_n} \quad (1-5)$$

bekannt ist, wobei $\beta = (kT)^{-1}$ die inverse Temperatur ist. Diese Analogie eröffnet die Möglichkeit, Probleme der QFT mit Methoden der statistischen Physik zu bearbeiten. Im speziellen heißt das, daß beispielsweise Monte-Carlo-Methoden verwendet werden können, um Systeme der QFT zu simulieren.

Für den weiteren Gebrauch kann die Zustandssumme für die QFT definiert werden:

Definition 1.1.1 (Zustandssumme)

$$\mathcal{Z}_E := \int \mathcal{D}\phi e^{-\frac{1}{\hbar} S_E}.$$

\hbar kann als Temperatur interpretiert werden. Mit $\beta = 1/\hbar$ kann man schreiben:

$$\mathcal{Z}_E := \int \mathcal{D}\phi e^{-\beta S_E}. \quad (1-6)$$

Da in der folgenden Arbeit ausschließlich die euklidische Metrik verwendet wird, wird im folgenden auf den Subskript E verzichtet.

1.1.3 Observablen

Observablen werden in der Quantenmechanik durch hermitesche Operatoren repräsentiert. Ihr Erwartungswert berechnet sich in der Quantenfeldtheorie wiederum durch Summe über alle Zustände, gewichtet mit der Exponentialfunktion der Wirkung:

$$\langle \mathcal{O} \rangle = \frac{1}{\mathcal{Z}} \int \mathcal{D}\phi \mathcal{O}([\phi]) e^{-\beta S([\phi])}. \quad (1-7)$$

Entsprechend der Analogie zwischen statistischer Mechanik und euklidischer Quantenfeldtheorie können Observablen über die Zustandssumme und eventuell die Einführung äußerer Quellen zum Beispiel für die Korrelationsfunktion definiert werden. Beispielsweise kann die freie Energie F definiert werden:

$$F = \frac{1}{V} \ln \mathcal{Z}. \quad (1-8)$$

V ist das Volumen. Daraus bestimmt sich der Erwartungswert der inneren Energie mittels Differenzieren:

$$\langle E \rangle = -\frac{\partial}{\partial \beta} F = -\frac{1}{V} \frac{\partial}{\partial \beta} \ln \mathcal{Z} = \frac{1}{V} \langle S \rangle. \quad (1-9)$$

1.2 Feldtheorie auf dem Gitter

Wie bereits in der Einleitung erwähnt, hat sich die Regularisierung mit Hilfe eines Raum-Zeit-Gitters zu einem wichtigen Werkzeug in der Quantenfeldtheorie entwickelt². Der endliche Gitterabstand spielt dabei die Rolle des Regulators bzw. des Energie *cut-off*.

Zugleich macht es die Diskretisierung der Raum-Zeit erst möglich, die euklidisierte Quantenfeldtheorie auf einem Computer zu simulieren, wobei es zusätzlich nötig ist, ein endliches, nicht nur diskretes Volumen zu betrachten.

Es wird ein diskretes Raum-Zeit-Gitter \mathcal{L} in d Dimensionen eingeführt:

Definition 1.2.1 (Diskretes Raum-Zeit-Gitter)

$$\mathcal{L} := \{\vec{x} \in \mathbb{R}^d \mid x_\mu/a \in \mathbb{Z}, \mu = 1, 2, \dots, d\}. \quad (1-10)$$

a wird Gitterkonstante genannt.

Bemerkung:

Die Gitterkonstante a ist die einzige dimensionsbehaftete Größe der Theorie.

Dieses diskrete Gitter \mathcal{L} ist immer noch unendlich ausgedehnt, so daß auch die Beschränkung auf einen Ausschnitt des Gitters nötig ist, um den Unzulänglichkeiten eines Computers gerecht zu werden. Das heißt, es wird die Gitterlänge $L \cdot a$ eingeführt und nur eine Ausschnitt von \mathcal{L} betrachtet:

$$\mathcal{L}_L := \{\vec{x} \in \mathbb{R}^d \mid L > x_\mu/a \in \mathbb{N}^0, \mu = 1, 2, \dots, d\}, \quad (1-11)$$

wobei der Einfachheit halber nur positive Werte zugelassen sind und auf das Subskript L in Zukunft verzichtet wird. Die Gitterlänge $L \cdot a$ muß so gewählt werden, daß sie deutlich größer als die für das Modell relevante Skala ist, und somit die Auswirkungen des endlichen Volumens auf Ergebnisse möglichst klein sind. Die Gitterkonstante a hingegen muß deutlich kleiner als die relevante Skala sein, um Diskretisierungsfehler möglichst zu vermeiden.

Mit der Entscheidung für die endliche Ausdehnung des Gitters müssen auch die Randbedingungen definiert werden: Es gibt verschiedene Möglichkeiten, das zu tun. In dieser Arbeit werden ausschließlich periodische Randbedingungen verwendet. Außerdem wird im folgenden meist die Dimension $d = 2$ betrachtet, wobei dann jeder Punkt des Gitters als $\mathbf{x} \in \mathcal{L}$ notiert wird und die Komponenten mit $\mathbf{x} = (x, t)$ angegeben werden.

Bemerkung:

Man bezeichnet die Abweichungen von der kontinuierlichen Theorie, die durch die Diskretisierung entstehen, als Diskretisierungs-Effekte, wohingegen die durch das endliche Volumen hervorgerufenen Fehler als *finite size*-Effekte bezeichnet werden.

² Siehe dazu auch [9] und [10].

Um wieder zur Kontinuums-Physik zurückzugelangen, müssen somit zwei Grenzwerte gebildet werden, wobei jeweils die physikalische Skala $1/La$ festgehalten werden muß:

- Der Kontinuums-limes: $a \rightarrow 0$.
- Der Limes zum unendlichen Volumen: $L \rightarrow \infty$.

Die physikalische Skala hängt dabei vom betrachteten System ab, es kann sich um eine Massenskala oder auch eine perturbative Skala handeln. Die Durchführung des Kontinuums-limes ist oft sehr schwierig; sie ist aber zwingend: Die in der kontinuierlichen Theorie auftretenden Unendlichkeiten werden zwar durch das Gitter regularisiert. Die physikalischen Größen erhält man aber erst durch anschließendes Renormalisieren.

Im Anschluß an die Diskretisierung der Raum-Zeit werden alle Größen, die von \vec{x} abhängen, auf das diskrete Gitter übertragen, in dem sie nur noch an den Gitter-Punkten definiert sind:

$$f(\vec{x}) \longrightarrow f_{\mathbf{x}}, \quad \mathbf{x} \in \mathcal{L}. \quad (1-12)$$

Das bedeutet, daß der „kontinuierliche Index“ x in den diskreten Index $\mathbf{x} \in \mathcal{L}$ übergeht. Des weiteren muß die Integration und Differentiation auf dem Gitter erklärt werden. Beides bedient sich intuitiv der Tatsache, daß Integration und Differentiation als Grenzwertprozesse definiert sind. Die einfachste Möglichkeit, die Ableitung zu definieren, ist, den Differenzen-Quotienten zu benutzen:

Definition 1.2.2 (Vorwärts-Ableitung in Richtung μ)

$$\Delta_{\mu} f_{\mathbf{x}} := \frac{1}{a} (f_{\mathbf{x}+a\hat{\mu}} - f_{\mathbf{x}}), \quad (1-13)$$

mit $\hat{\mu}$ als Einheitsvektor in Richtung μ .

Die Rückwärts-Ableitung kann unter Benutzung des Gitter-Punktes $\mathbf{x} - a\hat{\mu}$ analog zu (1-13) definiert werden.

Die Integration über die Raum-Zeit geht auf dem Gitter in eine endliche Summe über:

$$\int d^d x \longrightarrow \sum_{\mathbf{x} \in \mathcal{L}} a^d \quad . \quad (1-14)$$

Mit Hilfe von (1-12) ist es möglich, ein Feld $[\phi]$ auf das Gitter zu übertragen und damit auch eine präzise Definition des Maßes $\mathcal{D}\phi$ (auf dem Gitter) zu geben:

$$\int \mathcal{D}\phi \longrightarrow \int \prod_{\mathbf{x} \in \mathcal{L}} d\phi_{\mathbf{x}}, \quad (1-15)$$

so daß die Zustandssumme (1-6) bzw. der Erwartungswert (1-7) sich wie folgt ergeben:

$$\mathcal{Z} = \int \prod_{\mathbf{x} \in \mathcal{L}} d\phi_{\mathbf{x}} e^{-\beta S([\phi])} \quad (1-16)$$

und

$$\langle \mathcal{O} \rangle = \frac{1}{\mathcal{Z}} \int \prod_{\mathbf{x} \in \mathcal{L}} d\phi_{\mathbf{x}} \mathcal{O}([\phi]) e^{-\beta S([\phi])}. \quad (1-17)$$

Damit sind alle Werkzeuge bei der Hand, um eine Feldtheorie auf einem Computer zu simulieren.

1.3 Anharmonischer Oszillator

In diesem Abschnitt wird der quantenmechanische anharmonische Oszillator betrachtet. Es wird die Wirkung und deren Diskretisierung dargestellt, woran anschließend eine Neu-Parametrisierung durchgeführt wird, um Vergleiche mit dem später betrachteten σ -Modell einfacher möglich zu machen.

Eine erste Darstellung und Untersuchung findet sich dazu bei M. Creutz und B. Freedman in [11].

Die Wirkung des anharmonischen Oszillators in einer Dimension hat folgende Form:

$$S = \int dt \left(\frac{1}{2} \dot{\varphi}^2 + V(\varphi) \right), \quad \varphi = \varphi(t) \in \mathbb{R}. \quad (1-18)$$

Dabei ist V das Potential:

$$V(\varphi) = \frac{1}{2} \omega^2 \varphi^2 + \lambda \varphi^4, \quad \lambda \geq 0. \quad (1-19)$$

Der Fall $\lambda = 0$ beschreibt den harmonischen Oszillator. $\lambda > 0$ dagegen beschreibt den anharmonischen Fall. Der harmonische Oszillator ist noch geschlossen lösbar, was für den anharmonischen nicht mehr gilt.

1.3.1 Anharmonischer Oszillator auf dem Gitter

Um den anharmonischen Oszillator auf das Gitter zu übertragen, also zu diskretisieren, wird kein quadratisches Gitter benötigt, sondern lediglich ein eindimensionales Zeit-Gitter der Ausdehnung L , so daß $T = La$ ist. Benutzt werden periodische Randbedingungen.

Der kontinuierliche Index t wird wie (1-12) aus dem vorigen Abschnitt mit

$$\varphi(t) \longrightarrow \varphi(t_i) =: \varphi_i, \quad i = 1, \dots, L,$$

1.3. ANHARMONISCHER OSZILLATOR

diskretisiert, wobei anstelle von \mathbf{x} in diesem eindimensionalen Fall der Index i verwendet wird. Entsprechend (1-13) wird der kinetische Term mit

$$\dot{\varphi}(t) \longrightarrow \frac{\varphi_{i+1} - \varphi_i}{a}$$

auf das Gitter übertragen. Ebenso gilt für das Potential:

$$V(\varphi(t)) \longrightarrow V_i = \frac{1}{2}\omega^2\varphi_i^2 + \lambda\varphi_i^4.$$

Damit kann die Wirkung für das diskrete Modell angegeben werden:

$$S = \sum_{i=1}^L a \left(\frac{1}{2} \frac{(\varphi_{i+1} - \varphi_i)^2}{a^2} + V_i \right). \quad (1-20)$$

Um das Modell besser mit dem später erläuterten σ -Modell vergleichen zu können, ist es sinnvoll, es neu zu parametrisieren: Folgende Ersetzungen führen zu dimensionslosen Größen:

$$\varphi_i \rightarrow \sqrt{2a\kappa}\phi_i, \quad \omega^2 \rightarrow \frac{1}{a^2\kappa}(1 - 2\lambda_L), \quad \lambda \rightarrow \frac{\lambda_L}{4a^3\kappa^2}. \quad (1-21)$$

Damit kann die Wirkung wie folgt ausgedrückt werden:

$$S = \sum_{i=1}^L [\kappa(\phi_{i+1} - \phi_i)^2 + V_i], \quad (1-22)$$

mit

$$V_i = \lambda_L(\phi_i^2 - 1)^2 + \phi_i^2 - \lambda_L.$$

Es wird ersichtlich, daß für den Fall $\lambda_L \rightarrow \infty$ alle Konfigurationen unterdrückt werden, für die nicht $(\phi_i^2 - 1)^2 = 0$ gilt; das heißt dieser Fall führt in einer Dimension zum Ising-Modell mit folgender Wirkung:

$$S = -2\kappa \sum_{\langle ij \rangle} \phi_i \phi_j, \quad \phi_i^2 = 1 \quad \forall i. \quad (1-23)$$

Diese Wirkung ist im Vorgriff auf den folgenden Abschnitt identisch zu der Wirkung des nichtlinearen σ -Modells in einer Dimension. Das nichtlineare σ -Modell ergibt sich also als ein möglicher Grenzwert aus dem anharmonischen Oszillator.

1.4 Nichtlineares $O(n)$ - σ -Modell in $d = 2$ Dimensionen

Im folgenden Abschnitt wird allgemein das $O(n)$ - σ -Modell betrachtet, weil eine Einschränkung auf ein konkretes Modell keine zusätzlichen Ergebnisse liefert. In der Arbeit wurde das Modell für $n = 4$ untersucht, weil dann die Abbildung auf die Gruppe $SU(2, \mathbb{C})$ möglich ist und sich damit Ähnlichkeiten zur QCD auftun, in der mit der Gruppe $SU(3, \mathbb{C})$ gearbeitet wird. Die Abbildung auf die Gruppe $SU(2, \mathbb{C})$ wird im Abschnitt 3.2.2 behandelt.

Beim nichtlinearen $O(n)$ - σ -Modell handelt es sich um ein kontinuierliches Spin-Modell: An jedem Punkt x der zweidimensionalen Raum-Zeit befindet sich ein n -komponentiger Spinvektor $\phi(x) = (\phi^1(x), \dots, \phi^n(x)) \in \mathbb{R}^n$ der Länge eins. Gekoppelt sind die „Spins“ an den Raum-Zeit-Punkten über den Kopplungsparameter g , der auch als Temperatur interpretiert werden kann.

Die euklidische Standard-Lagrange-Funktion für das nichtlineare $O(n)$ - σ -Modell ist wie folgt definiert:

$$\mathcal{L}_\sigma(x) = \frac{1}{g} \sum_{\mu=1}^2 [\partial_\mu \phi(x)]^2. \quad (1-24)$$

$\mu = 1, 2$ indiziert die Normal-Richtungen, hier in Raum- und Zeit-Richtung. Das Skalarprodukt ist wie üblich für den \mathbb{R}^n definiert:

$$\phi\psi = \langle \phi, \psi \rangle = \sum_{k=1}^n \phi^k \psi^k.$$

Die Norm wird über das Skalarprodukt definiert, da der \mathbb{R}^n ein euklidischer Vektorraum (bzw. ein Hilbertraum) ist:

$$\|\phi(x)\|^2 = \langle \phi(x), \phi(x) \rangle = 1. \quad (1-25)$$

Deswegen ist $\phi(x)$ ein Element der Oberfläche der n -dimensionalen Einheitskugel $S_{(n-1)}$. Die Einschränkung (1-25) der Norm kann als holonome Zwangsbedingung interpretiert werden.

Die inverse Kopplungskonstante β wird wie folgt definiert:

$$\beta = \frac{1}{g}. \quad (1-26)$$

Im folgenden ist β gemeint, wenn von Kopplungskonstante die Rede ist, außer, es wird ausdrücklich darauf hingewiesen³.

³ In diesem Zusammenhang heißt der Limes β gegen 0 *strong coupling expansion*, weil dann g gegen ∞ geht und umgekehrt.

Mit (1-24) kann die euklidische Wirkung des nichtlinearen $O(n)$ - σ -Modells definiert werden:

$$S_\sigma = \int d^2x \quad \beta \sum_\mu [\partial_\mu \phi(x)]^2, \quad \text{mit} \quad \phi(x)^2 = 1 \quad \forall x. \quad (1-27)$$

Damit ist es möglich, einige Bemerkungen zum Verhalten des Modells zu machen: Allgemein ist für den Grenzwert g gegen 0 zu erwarten, daß sich alle Spins relativ zueinander parallel stellen, für $g \rightarrow \infty$ dagegen, daß die Spins sich relativ zueinander nicht ausrichten, also die Richtung des Nachbarn zufällig zu der eigenen ist. Das physikalisch interessante Verhalten findet meist an kritischen Punkten der Theorie statt. Dieser kritische Punkt liegt im vorliegenden Modell bei $g = 0$.

Anders gesagt: Bei hohen Temperaturen g ist der Abstand, über den zwei Spins an verschiedenen Punkten in der Raum-Zeit einander noch beeinflussen klein. Bei niedrigen Temperaturen ist dieser Abstand groß: Er divergiert für den Grenzfall, weil dort der Phasenübergang stattfindet. Diese Tatsache findet sich im Verhalten der Korrelationslänge wieder, die ein Maß für diesen Abstand ist.

1.4.1 Nichtlineares $O(n)$ - σ -Modell auf dem Gitter

Mit den Werkzeugen aus Abschnitt 1.2 ist es nun möglich, die Wirkung des nichtlinearen $O(n)$ - σ -Modells in zwei Dimensionen für ein diskretes Gitter mit periodischen Randbedingungen anzugeben:

$$S_\sigma^{\mathcal{L}} = \sum_{\mathbf{x}} a^2 \beta \sum_{\mu} \frac{1}{a^2} (\phi_{\mathbf{x}+\hat{\mu}} - \phi_{\mathbf{x}})^2.$$

Benutzt man die Zwangsbedingung $\phi_{\mathbf{x}}^2 = 1$ an jedem Punkt des Gitters, so erhält man

$$S_\sigma^{\mathcal{L}} = \sum_{\mathbf{x}} 2\beta \left(1 - \sum_{\mu} \phi_{\mathbf{x}} \phi_{\mathbf{x}+\hat{\mu}} \right) = -2\beta \sum_{\mathbf{x},\mu} \phi_{\mathbf{x}} \phi_{\mathbf{x}+\hat{\mu}} + \text{const.}$$

Nach Reskalieren ergibt sich folgender Ausdruck für die diskrete Wirkung des Modells und für die Zwangsbedingung entsprechend Gleichung (1-23):

$$S_\sigma^{\mathcal{L}} = -2\beta \sum_{\mathbf{x},\mu} \phi_{\mathbf{x}} \phi_{\mathbf{x}+\hat{\mu}}, \quad \phi_{\mathbf{x}} \phi_{\mathbf{x}} = 1. \quad (1-28)$$

1.4.2 Erwartungswert der Energie

In Abschnitt 1.1.3 wurde der Zusammenhang zwischen euklidischer QFT und statistischer Mechanik beschrieben, der sich jetzt in Form von Gleichung (1-9) ausnutzen läßt, um einen Ausdruck für den Erwartungswert der inneren Energie auf dem Gitter anzugeben:

$$\langle E \rangle = - \left\langle \frac{1}{V} \sum_{\mathbf{x},\mu} \phi_{\mathbf{x}} \phi_{\mathbf{x}+\hat{\mu}} \right\rangle, \quad V = L^2. \quad (1-29)$$

KAPITEL 1. THEORETISCHE GRUNDLAGEN

Da das Modell auch auf dem Gitter eine Rotations-Symmetrie für Drehungen um $\pi/2$ Grad aufweist, gilt damit auch⁴

$$\langle E \rangle = -\left\langle \frac{2}{V} \sum_{\mathbf{x}} \phi_{\mathbf{x}} \phi_{\mathbf{x}+\hat{i}} \right\rangle =: 2\langle E_W \rangle. \quad (1-30)$$

Entsprechend der Bemerkung zum allgemeinen Verhalten des Modells kann man aus dieser Definition bereits ablesen, wie sich der Erwartungswert von E_W in Abhängigkeit von dem Kopplungsparameter β verhalten wird: Da für kleine Werte in β zwei benachbarte Felder vorzugsweise ungleich ausgerichtet sind, nimmt $\langle E_W \rangle$ bei kleinen Werten von β einen Wert an, der nah bei 0 liegt. Genauere Erkenntnisse für kleine Werte von β liefert die *strong coupling expansion*, die in Anhang A.4 dargestellt ist.

Im anderen Extremfall großer Werte für β werden zwei benachbarte Felder vorzugsweise parallel ausgerichtet sein. Folglich wird sich der Erwartungswert von E_W mit wachsenden Werten von β der 1 annähern.

Für spätere Zwecke wird außerdem der Erwartungswert der Wirkung benötigt, der sich jedoch nur um den Faktor βL^2 von dem der Energie unterscheidet:

$$\langle S \rangle = -2\beta \left\langle \sum_{\mathbf{x}, \mu} \phi_{\mathbf{x}} \phi_{\mathbf{x}+\hat{\mu}} \right\rangle. \quad (1-31)$$

1.4.3 Asymptotische Freiheit

Eine wichtige Eigenschaft des nichtlinearen $O(n)$ - σ -Modells für $n > 3$ ist seine asymptotische Freiheit. Das bedeutet, daß die Theorie bei sehr kleinen Abständen wechselwirkungsfrei wird, Felder einander also nicht beeinflussen. Diese Eigenschaft teilt das Modell mit der QCD. Sie war neben anderen ausschlaggebend dafür, das Modell in der vorliegenden Arbeit zu untersuchen.

Die folgenden Untersuchungen behandeln das diskrete Modell, obwohl ähnliche Überlegungen auch im Kontinuum angestellt werden können. Im zweidimensionalen, nichtlinearen $O(n)$ - σ -Modell nimmt die Korrelation zwischen Feldern exponentiell mit dem Abstand ab: Die Felder erhalten, wie auch in der QCD, eine dynamisch generierte Masse, welche aber für $g \rightarrow 0$ ebenfalls gegen Null geht. Das Verhalten dieser Masse – des sogenannten *mass-gap* $m = 1/\xi$ – unter Skalen-Transformationen ist durch die Callan-Symanzik- β -Funktion⁵ festgelegt, die darüber hinaus der Tatsache Ausdruck verleiht, daß die nackte Kopplung g und der Gitterabstand a nicht unabhängig voneinander sind [12]:

⁴ Die Definition von E_W wurde für den Vergleich mit Ergebnissen anderer Autoren eingeführt.

⁵ Um Verwirrung zu vermeiden, wird darauf hingewiesen, daß aus historischen Gründen der inverse Kopplungsparameter (1-26) mit β bezeichnet wird, ebenso wie die β -Funktion.

Definition 1.4.1 (β -Funktion) Die β -Funktion ist definiert durch:

$$\beta(g) = -a \frac{\partial}{\partial a} g = -\beta_0 g^2 - \beta_1 g^3 - \beta_2 g^4 + O(g^5). \quad (1-32)$$

Für die σ -Modelle ist die β -Funktion analytisch gut bekannt. Bei β_0 und β_1 handelt es sich um universelle, vom Regularisierungsschema unabhängige Koeffizienten (ein- und zwei-loop Beiträge):

$$\beta_0 = \frac{n-2}{2\pi}, \quad \beta_1 = \frac{n-2}{(2\pi)^2},$$

mit der Dimension n der „Spins“. Für das nichtlineare $O(n)$ - σ -Modell in verschiedenen Regularisierungsschemata ist auch β_2 bekannt (Falcioni und Treves, [13]). Für die vorliegende Arbeit sind die folgenden Koeffizienten zu verwenden:

$$\beta_2 = \frac{n-2}{(2\pi)^3} [0.514 - (n-2)0.089].$$

Über die Renormierungs-Gruppen-Gleichung⁶ (RG-Gleichung)

$$\left[\beta(g) \frac{\partial}{\partial g} - a \frac{\partial}{\partial a} \right] \Lambda(g, a) = 0$$

definiert sich der Λ -Parameter als eine RG-Invariante, und ist eine Größe von der Dimension einer Masse. Da die Theorie keinen Massenterm in der Wirkung enthält, müssen alle spontan erzeugten Massen proportional zu Λ sein:

$$a\Lambda(g) = e^{-1/(\beta_0 g)} (\beta_0 g)^{-\beta_1/\beta_0^2} \left[1 + \frac{\beta_1^2 - \beta_0 \beta_2}{\beta_0^3} g + O(g^2) \right]. \quad (1-33)$$

Damit ist das asymptotische Skalierungsverhalten von $m = 1/\xi$ für g gegen 0 festgelegt. Da es sich bei m um eine dynamisch generierte Größe handelt, gibt es keine Möglichkeit, sie störungstheoretisch zu berechnen.

Oft ist es nützlich, die Kopplung zu reparametrisieren, um beispielsweise das Verhalten der Störungsreihe zu verbessern: Bei Übergang von der Kopplung g zu einer anderen Kopplung

$$g' = f(g) = g \left(1 + \sum_{i=1}^{\infty} C_i g^i \right),$$

folgt aus (1-33) und einem einfachen Argument [9], daß folgende Gleichheit gilt:

$$\frac{\Lambda(g')}{\Lambda(g)} = \exp \left(\frac{C_1}{\beta_0} \right). \quad (1-34)$$

⁶ Siehe zum Thema Renormierung [12] und [14]

Da m wie Λ skaliert, folgt aus (1-33), daß die nackte Kopplung mit am abnimmt. Das bedeutet für den Kontinuums-Limes:

$$g \rightarrow 0 \quad \Leftrightarrow \quad \beta \rightarrow \infty. \quad (1-35)$$

Damit gilt für Größen \mathbf{m} mit der Dimension einer Masse, daß im Kontinuums-Limes bzw. für den Grenzwert β gegen ∞ das Verhältnis \mathbf{m}/Λ konstant werden sollte. Man spricht davon, daß die Größe „skaliert“, wenn man den Bereich erreicht hat bzw. wenn es einen Bereich gibt, in dem das Verhältnis konstant wird. Ist für eine Größe eine störungstheoretische Näherungsformel bekannt, so spricht man von „asymptotischem Skalieren“ der Größe, falls das Verhältnis zum störungstheoretisch berechneten Wert konstant wird.

1.5 Lyapunov-Exponent

Im folgenden Abschnitt soll kurz erläutert werden, was deterministisches Chaos bedeutet und woher der Begriff „Lyapunov-Exponent“ stammt⁷.

Chaotische Systeme sind Systeme, die auf minimale Änderungen der Anfangsbedingungen sehr empfindlich reagieren. Es wird somit praktisch unmöglich, aus einem bekannten Zustand eines Systems auf dessen weiteren Verlauf zu schließen, weil bereits eine sehr kleine Änderung in den Anfangsbedingungen zu deutlich oder völlig anderem Verhalten zu einem späteren Zeitpunkt führt.

Kleine Unterschiede in den Anfangsbedingungen kommen schon deswegen zustande, weil es prinzipiell unmöglich ist, den Anfangszustand exakt anzugeben.

Chaotisches Verhalten wird immer dann beobachtet, wenn sich zwei benachbarte Trajektorien im Mittel exponentiell schnell voneinander entfernen. Dieses exponentielle Wachstum der Abweichung kann durch einen charakteristischen Exponenten spezifiziert werden, den Lyapunov-Exponenten [16].

Das Konzept des Lyapunov-Exponenten läßt sich unter sehr allgemeinen Bedingungen ableiten. Es soll an dieser Stelle unter etwas eingeschränkten Voraussetzungen dargestellt werden.

Um ein Maß für die Geschwindigkeit des Auseinanderlaufens zweier Trajektorien zu erhalten, betrachtet man eine Lösung $\Phi(t, \phi_0)$ zu den Anfangsbedingungen ϕ_0 der Klasse von Bewegungsgleichungen

$$\frac{\partial}{\partial t} \phi = F(\phi, t),$$

die eine große Anzahl von Systemen beschreibt. Dabei sind Φ , ϕ und F jeweils mehrkomponentige Größen und außerdem ist

$$\Phi(0, \phi_0) = \phi_0.$$

⁷ Siehe [15].

Wird die Anfangsbedingung ϕ_0 minimal um $\delta\phi_0$ verändert, stellt sich die Frage, wie sich die Differenz $\delta\Phi(t)$ aus $\Phi(t, \phi_0)$ und $\Phi(t, \phi_0 + \delta\phi_0)$ in der Zeit verhält. Es gilt in linearer Näherung:

$$\frac{\partial}{\partial t}\delta\Phi \equiv \dot{\Phi}(t, \phi_0 + \delta\phi_0) - \dot{\Phi}(t, \phi_0) = L(t)(\Phi(t, \phi_0 + \delta\phi_0) - \Phi(t, \phi_0)) \equiv L(t)\delta\Phi. \quad (1-36)$$

Die Matrix L ist dabei durch die partiellen Ableitungen gegeben:

$$L(t) = \left(\frac{\partial F_i}{\partial \phi_k} \right)_{\Phi}. \quad (1-37)$$

In der Regel läßt sich die Differentialgleichung (1-36) nicht analytisch integrieren. Nimmt man an, es gibt eine Lösung. Bei $t = 0$ ist dann $\delta\Phi(t = 0) = \delta\phi_0$ und für $t > 0$ sei die zeitliche Entwicklung von $\delta\Phi(t)$ wie folgt gegeben:

$$\delta\Phi(t) = \mathbf{A}(t)\delta\phi_0.$$

Die Matrix \mathbf{A} genügt, wie aus (1-36) folgt, selbst der Differentialgleichung

$$\frac{\partial}{\partial t}\mathbf{A}(t) = L(t)\mathbf{A}(t), \quad (1-38)$$

die sich formal integrieren läßt:

$$\mathbf{A}(t) = \exp\left(\int_0^t dt' L(t')\right) \mathbf{A}(0), \quad (1-39)$$

wobei $\mathbf{A}(0) = \mathbb{I}$ ist. Überdies angenommen, die Matrix L hänge in erster Näherung im betrachteten Zeitraum nicht von der Zeit ab. Seien λ_k ihre Eigenwerte. Unter Verwendung der dazugehörigen Eigenvektoren wird $\mathbf{A}(t)$ diagonalisiert, und es ist:

$$\mathbf{A}(t) = \{\exp(\lambda_k t)\}.$$

Daraus folgt unmittelbar, daß benachbarte Lösungen genau dann auseinanderlaufen, wenn mindestens einer der Eigenwerte einen positiven Realteil $\text{Re}(\lambda_k) = (\lambda_k + \lambda_k^*)/2 > 0$ besitzt. Dies kann festgestellt werden, indem bei

$$\frac{1}{2t} \ln \text{Sp}(\mathbf{A}^\dagger(t)\mathbf{A}(t)) = \frac{1}{2t} \ln \text{Sp}[\exp((\lambda_k + \lambda_k^*)t)]$$

der Grenzwert $t \rightarrow \infty$ gebildet wird. Dabei überlebt nur der Eigenwert mit dem größten positiven Realteil. Daher wird der Lyapunov-Exponent ν wie folgt definiert:

Definition 1.5.1 (Lyapunov-Exponent)

$$\nu := \lim_{t \rightarrow \infty} \frac{1}{2t} \ln \text{Sp}(\mathbf{A}^\dagger(t)\mathbf{A}(t)). \quad (1-40)$$

KAPITEL 1. THEORETISCHE GRUNDLAGEN

Diese Definition eröffnet die Möglichkeit, zu entscheiden, ob es zumindest eine Richtung gibt, in der sich benachbarte Trajektorien im Mittel exponentiell voneinander entfernen, und zwar mit $\exp(\nu t)$. Das ist genau dann der Fall, wenn der Lyapunov-Exponent positiv ist. In diesem Fall reagiert die Lösung der Bewegungsgleichung sehr empfindlich auf minimale Veränderungen der Anfangsbedingungen.

Bemerkung:

Es kann gezeigt werden, daß der Lyapunov-Exponent für integrable Systeme verschwindet [17]. Integrable Systeme sind Systeme, bei denen die Untersuchung der Bewegung auf die Berechnung eindimensionaler Integrale zurückgeführt werden kann. Eine hinreichende Bedingung für Integrabilität gibt der „Satz von Liouville“ [17].

Kapitel 2

Simulationsverfahren

Im folgenden Kapitel soll das Simulationsverfahren erläutert werden, das für die Untersuchungen in der vorliegenden Arbeit verwendet wurde: Der HMC-Algorithmus. Ferner wird dargestellt, wo und warum es zu Verletzungen der Reversibilität kommt und wie diese untersucht werden können. Schließlich werden einige allgemeine Betrachtungen zu statistischen Fehlern und der Zuverlässigkeit der Simulationen angestellt.

Der HMC-Algorithmus wurde zuerst von S. Duane et al. in [2] vorgestellt. Die Überlegungen in diesem Kapitel folgen aber in Teilen dem Gedankengang von H.J. Rothe in [10].

2.1 Der Hybrid-Monte-Carlo-Algorithmus

Um die Zustandssumme (1-6) oder den Erwartungswert einer Observablen (1-7) bzw. deren diskrete Versionen (1-16) und (1-17) zu berechnen, ist es notwendig, ein hochdimensionales Integral über den Raum der Feld-Konfigurationen zu berechnen. Dabei tragen fast alle Konfigurationen nur mit einem sehr geringen Gewicht bei, so daß eine ungewichtete stochastische Integration („*simple sampling*“) äußerst ineffektiv ist.

Die Effektivität kann deutlich gesteigert werden, indem das sogenannte *importance sampling* eingesetzt wird. Bei dieser Methode werden bevorzugt solche Konfigurationen ausgewählt, die mit großer Wahrscheinlichkeit auftreten. Dabei wird eine Stichprobe erzeugt, für die der einfache Mittelwert über die Stichprobe den Ensemble-Mittelwert abschätzt. Zu diesem Zweck werden Markov-Ketten eingesetzt:

2.1.1 Markov-Ketten

Zunächst wird der Begriff der stochastischen Prozesse benötigt. Eine Folge von Zufallsgrößen wird als stochastischer Prozeß bezeichnet, wenn die Folge von der Zeit abhängt.

KAPITEL 2. SIMULATIONSVERFAHREN

Das heißt, wenn die k -te Zufallszahl einerseits von der $(k-1)$ -ten, $(k-2)$ -ten usw. abhängt und andererseits vom „Zeitpunkt“ k selbst. Dieser Begriff kann auf Markov-Prozesse spezialisiert werden: Ein stochastischer Prozeß heißt dann Markov-Prozeß (erster Stufe), wenn das k -te Element nur vom $(k-1)$ -ten Element abhängt und nicht vom Zeitpunkt k selbst. Für die QFT bedeutet das: Sei $\{[\phi_1], [\phi_2], \dots\}$ die Menge von Konfigurationen des Systems. Ein Markov-Prozeß über diese Menge wird mit Hilfe einer Übergangs-Wahrscheinlichkeit $T_{ij} := T([\phi_i] \rightarrow [\phi_j])$ wie folgt erzeugt:

$$[\phi_{t_1}] \xrightarrow{T_{t_1 t_2}} [\phi_{t_2}] \xrightarrow{T_{t_2 t_3}} [\phi_{t_3}] \dots ,$$

wobei die Folge $(t_n)_n$ über der Indexmenge der Konfigurationen definiert ist. Wird die Größe $p_i^{(k)}$ als die Wahrscheinlichkeit definiert, die Konfiguration $[\phi_i]$ zur Zeit k zu finden, für die folgende Normierung gilt:

$$\sum_j p_j^{(k)} = 1, \quad \forall k, \quad (2-1)$$

so kann eine präzisere Definition von Markov-Prozessen gegeben werden:

Definition 2.1.1 (Markov-Prozeß (erster Ordnung)) Sei $\{[\phi_i]\}$ eine abzählbare Menge von Konfigurationen. Ein stochastischer Prozeß heißt Markov-Prozeß erster Ordnung, falls die Wahrscheinlichkeit $p_i^{(k)}$, Konfiguration $[\phi_i]$ zum Zeitpunkt k zu finden, nur von der unmittelbar vorhergehenden Konfiguration zum Zeitpunkt $(k-1)$ abhängt, und nicht vom Zeitpunkt k selbst.

Die Folge der so erzeugten Konfigurationen heißt Markov-Kette.

In der Ensemble-Interpretation ergibt sich für die Wahrscheinlichkeiten:

$$p_j^{(k+1)} = \sum_i p_i^{(k)} T_{ij}, \quad (2-2)$$

oder in Matrix-Schreibweise mit den Zeilenvektoren $p^{(k)} = (p_1^{(k)}, p_2^{(k)}, \dots)$ und $T = (T_{ij})$

$$p^{(k+1)} = p^{(k)} T. \quad (2-3)$$

Nun ist es interessant, das Spektrum von T als linearen Operator zu betrachten. Insbesondere ist es wichtig zu wissen, ob T einen stabilen Fixpunkt hat

$$\lim_{k \rightarrow \infty} p^{(k)} = p \quad \text{mit} \quad pT = p, \quad (2-4)$$

unabhängig von der Startkonfiguration. Dies gilt dann, wenn T eine stochastische Matrix ist [18]:

Definition 2.1.2 (Stochastische Matrix) Eine Matrix $T = (T_{ij})$ heißt stochastische Matrix, falls gilt:

1. Die Übergangswahrscheinlichkeiten T_{ij} sind alle größer gleich 0:

$$T_{ij} \geq 0 \quad \forall i, j. \quad (2-5)$$

2. Normierung:

$$\sum_j T_{ij} = 1 \quad \forall i. \quad (2-6)$$

Denn dann gilt der folgende Satz:

Satz 2.1.1 (Perron-Frobenius) Sei T eine stochastische Matrix. Dann existiert ein $p \geq 0$, mit $pT = p$. Das heißt, p ist Eigenvektor von T zum Eigenwert 1. Ist zusätzlich $T_{ij} > 0$ für alle i, j , so ist der Eigenwert 1 nicht entartet.

Konstruktion der Übergangsmatrix

Aus (2-4) ergibt sich unmittelbar das **Stabilitätskriterium**, das die Übergangsmatrix erfüllen muß:

$$\sum_i p_i T_{ij} = p_j, \forall j. \quad (2-7)$$

In der Praxis wird die Übergangsmatrix allerdings so erstellt, daß an sie anstelle der Bedingung der Stabilität die stärkere Bedingung der *detailed balance* gestellt wird:

$$p_i T_{ij} = p_j T_{ji} \quad \forall i, j. \quad (2-8)$$

Aus der *detailed balance*-Bedingung folgt die Stabilität, wenn man in (2-8) auf beiden Seiten über i summiert und benutzt, daß $\sum_i T_{ji} = 1$ ist. Es handelt sich um eine stärkere Bedingung als die Stabilität, die hinreichend, aber nicht notwendig ist, die sich aber bei der Konstruktion einer Übergangsmatrix in der Praxis leichter handhaben läßt.

2.1.2 Metropolis-Algorithmus

Eine Möglichkeit, die Übergangsmatrix zu konstruieren, wird durch die Übergangsvorschrift des Metropolis-Algorithmus definiert. Der Fixpunkt des Algorithmus soll eine Boltzmann-Verteilung bezüglich der Wirkung des betrachteten Modells sein, also:

$$p_i = \exp(-S([\phi_i])) , \quad \forall i.$$

Die Markov-Kette wird beim Metropolis-Monte-Carlo in Abhängigkeit von der Wirkung des Modells ausgehend von der Konfiguration $[\phi_i]$ durch Wiederholung folgender Vorschrift erzeugt:

1. Zufällige Wahl einer beliebigen Test-Konfiguration $[\phi_j]$.
2. Akzeptanz von $[\phi_j]$ als neue Konfiguration mit der Wahrscheinlichkeit

$$T_{ij} = \begin{cases} \frac{p_j}{p_i} & \text{falls } p_j < p_i \\ 1 & \text{sonst.} \end{cases} \quad (2-9)$$

also immer, wenn die Wirkung kleiner geworden ist und sonst mit der Wahrscheinlichkeit $\exp(S([\phi_i]) - S([\phi_j]))$.

Schritt 2 wird auch Akzeptanzschritt genannt und die Erzeugung von $[\phi_j]$ *update*. Von einem lokalen *update* spricht man, wenn $[\phi_j]$ aus $[\phi_i]$ durch Änderungen an einem Gitter-Platz hervorgeht. Werden die Werte an alle Gitter-Plätzen in einem Schritt aufgefrischt, ist von einem globalen *update* die Rede.

Daß mit (2-9) die Bedingung der *detailed balance* (2-8) erfüllt wird, ist leicht zu zeigen:

Beweis:

$$\begin{aligned} p_i T_{ij} &= p_i \cdot \min\{1, p_j/p_i\} \\ &= \begin{cases} p_j & \text{falls } p_j < p_i \\ p_i & \text{sonst} \end{cases} \\ &= \begin{cases} p_j & \text{falls } p_j < p_i \\ p_j \frac{p_i}{p_j} & \text{falls } p_i \leq p_j \end{cases} \\ &= p_j \cdot \min\{1, p_i/p_j\} \\ &= p_j T_{ji}. \end{aligned}$$

□

Als problematisch stellt sich beim Metropolis-Monte-Carlo-Algorithmus heraus, daß die Akzeptanzrate zwar bei lokalen *updates* annehmbar hoch ist, bei globalen *updates* bei gleicher Autokorrelationszeit jedoch zu klein wird. Globale *updates* werden dann interessant, wenn der Akzeptanzschritt bei einem lokalen *update* genauso „teuer“ ist, wie bei einem globalen. Dies ist beispielsweise in der QCD der Fall, da in jedem, gleich ob lokalen oder globalen *update* eine große Matrix invertiert werden muß.

2.1.3 Molekular-Dynamik-Methode

Die Molekular-Dynamik-Methode (MD-Methode) wird üblicherweise dazu verwendet, das Verhalten klassischer Vielteilchensysteme zu untersuchen. Zu diesem Zweck werden die Hamiltonschen Bewegungsgleichungen numerisch integriert. Da es sich um konservative Systeme handelt, sollte die Energie erhalten bleiben.

Aus den Bewegungsgleichungen

$$\dot{\phi} = \frac{\partial \mathcal{H}([\phi, \pi])}{\partial \pi}, \quad (2-10a)$$

$$\dot{\pi} = -\frac{\partial \mathcal{H}([\phi, \pi])}{\partial \phi}, \quad (2-10b)$$

ergibt sich unmittelbar die Energieerhaltung, falls die Hamilton-Funktion nicht explizit von der Zeit abhängt:

$$\frac{d\mathcal{H}([\phi, \pi])}{dt} = 0.$$

Die Bewegungsgleichungen (2-10) können benutzt werden, um die Entwicklung des Systems in der Zeit zu modellieren: Aus einem Anfangszustand erhält man mittels Integration der Bewegungsgleichungen einen Zustand, der dem Zustand des Systems zu einem späteren Zeitpunkt entspricht. Der Endzustand ist vollständig durch den Anfangszustand bestimmt.

Bemerkung:

An dieser Stelle ist zu bemerken, daß es sich bei der Integrationszeit im ursprünglichen MD-Algorithmus um die tatsächliche physikalische Zeit des Systems handelt. Es wird also die reale Dynamik des Systems betrachtet.

Wie diese klassische Methode für statistische Simulationen von Quantensystemen verwendet werden kann, wird im folgenden geschildert: Beim Metropolis-Monte-Carlo-Algorithmus wird die Akzeptanz einer Konfiguration durch die Differenz zweier Wirkungen bestimmt, die im Mittel groß sein wird, wenn globale *updates* durchgeführt werden. In der QCD kann auf globale *updates* aber nicht verzichtet werden, weil für jeden Akzeptanzschritt die Invertierung einer großen Matrix nötig ist.

Um globale *updates* und hohe Akzeptanzraten zu verbinden, wird die Idee umgesetzt, eine Modell-Hamilton-Funktion an Stelle der Wirkung im Akzeptanzschritt zu verwenden und die *updates* mit Hilfe der Integration der Bewegungsgleichungen durchzuführen, da wegen der Energieerhaltung der Wert der Hamilton-Funktion bis auf Rundungs- und Diskretisierungs-Fehler erhalten ist, so daß eine hohe Akzeptanz erwartet werden kann.

Dabei muß es sich nicht a priori um die oben beschriebene physikalische MD-Hamilton-Funktion handeln, und bei den *updates* auch nicht um die Integration der physikalischen Bewegungsgleichungen. Theoretisch reicht eine Erhaltungsgröße und die dazugehörige Symmetrie-Transformation aus, um eine hohe Akzeptanz zu erhalten, wenn man anschließend zeigen kann, daß das Kriterium der Stabilität weiterhin bezüglich der Wirkung gilt und die Transformation hinreichend ergodisch ist.

Eine MD-Hamilton-Funktion bietet sich deswegen an, weil sich die Wirkung ohne großen Aufwand einfügen läßt und Grund zu der Annahme besteht, daß die Transformation ergodisch ist.

Die Ausführung dieser Idee führt zum HMC-Algorithmus, der im Anschluß eingeführt wird:

2.1.4 HMC-Algorithmus

Der *Hybrid-Monte-Carlo*-Algorithmus, entwickelt in [2], stellt eine Kombination aus der MD-Methode und dem Metropolis-Algorithmus im oben dargestellten Sinne dar. Ziel ist es, wie am Ende des vorigen Abschnittes beschrieben, ein globales *update* der Konfiguration durchzuführen und gleichzeitig eine große Akzeptanzrate zu erhalten, um damit den Rechenaufwand für eine unabhängige Konfiguration möglichst klein zu halten.

Darüber hinaus sorgt der Monte-Carlo-Akzeptanz-Teil dafür, daß die systematischen Fehler der MD-Methode korrigiert werden, so daß der Algorithmus im Prinzip exakt ist.

Kombination von Metropolis- und MD-Methode

Bei der MD-Methode wurde ausgenutzt, daß die Hamilton-Funktion eines Systems eine Konstante der Bewegung ist. Eine solche Konstante der Bewegung wird nun für die vorliegende Problematik konstruiert, indem eine Modell-Hamilton-Funktion erzeugt wird.

Jene muß offenbar die Wirkung $S([\phi])$ des Modells enthalten, darüber hinaus aber auch einen Anteil „kinetischer Energie“, der eine Veränderung des Systems bewirken kann. Es liegt nahe, die Feldvariablen $[\phi]$ mit den Ortsvariablen des MD-Systems zu identifizieren und Zufallsvariablen $[\pi]$ als dazu kanonisch konjugierte Impulse künstlich einzuführen.

Der Metropolis-Akzeptanzschritt soll dann nicht mehr durch die Wirkung bestimmt werden, sondern durch die Modell-Hamilton-Funktion:

$$\mathcal{H}([\phi, \pi]) = \sum_x \frac{1}{2} \pi_x^2 + S([\phi]). \quad (2-11)$$

Durch die Integration der Bewegungsgleichungen über die Trajektorienlänge τ im MD-Teil wird aus der Konfiguration $[\phi] = [\phi(0)]$ die Konfiguration $[\phi'] = [\phi(\tau)]$ beziehungsweise werden aus den Impulsen $[\pi] = [\pi(0)]$ die Impulse $[\pi'] = [\pi(\tau)]$ deterministisch erzeugt. Das Paar $[\phi', \pi']$ wird dann mit der Wahrscheinlichkeit

$$P_A([\phi, \pi] \rightarrow [\phi', \pi']) = \min\{1, \exp(\mathcal{H}([\phi, \pi]) - \mathcal{H}([\phi', \pi']))\} \quad (2-12)$$

akzeptiert. Kann die Integration exakt ausgeführt werden, so ist der Wert dieser Wahrscheinlichkeit immer gleich 1. Wird die Integration aber numerisch ausgeführt, so ist der Wert der Hamilton-Funktion nur bis auf Diskretisierungsfehler erhalten, und P_A kann auch Werte annehmen, die kleiner als 1 sind. Für das weitere Vorgehen ist es zunächst nicht von Bedeutung, ob die Integration exakt oder numerisch ausgeführt wird, so daß keine Unterscheidung vorgenommen wird.

Wie später ersichtlich wird, muß die Wahrscheinlichkeitsverteilung der Impulse eine Gaussverteilung sein:

$$P_G([\pi]) = \mathcal{N} e^{-\frac{1}{2} \Sigma \pi_x^2}, \quad (2-13)$$

mit einer Normalisierungskonstanten \mathcal{N} . Denn dann kann gezeigt werden, daß die *detailed balance*-Bedingung weiterhin bezüglich der Wirkung erfüllt ist. Der Beweis (siehe dazu auch [10]) wird hier angegeben, weil er für das weitere Verständnis dieser Arbeit von Bedeutung ist:

Beweis: Zunächst muß die Übergangswahrscheinlichkeit von einem Zustand $[\phi]$ in einen Zustand $[\phi']$ definiert werden. Sie setzt sich aus dem Produkt mehrerer Wahrscheinlichkeiten zusammen:

Die Wahrscheinlichkeitsdichte P_G der gaussverteilten Impulse ist bereits mit (2-13) gegeben. Die Bewegungsgleichungen (2-10) für $[\phi]$ enthalten nur die zweite Ableitung nach der Zeit und somit kann die Bewegung durch Vorzeichenwechsel in der Zeit umgekehrt werden. Deswegen gilt für die Wahrscheinlichkeit P_M , durch die Integration der Bewegungsgleichung von der Konfiguration $[\phi, \pi]$ zur Konfiguration $[\phi', \pi']$ zu gelangen, folgende Gleichheit:

$$P_M([\phi, \pi] \rightarrow [\phi', \pi']) = P_M([\phi', -\pi'] \rightarrow [\phi, -\pi]). \quad (2-14)$$

Das bedeutet, daß die Integration reversibel ist, bzw. sein muß, um diesen Beweis führen zu können.

Die neue Konfiguration $[\phi', \pi']$ wird mit der Wahrscheinlichkeit P_A entsprechend (2-12) akzeptiert. Aus P_G, P_M, P_A kann nun durch Produktbildung und Integration über die Impulse die Wahrscheinlichkeit $P([\phi] \rightarrow [\phi'])$ gewonnen werden, von der Konfiguration $[\phi]$ zur Konfiguration $[\phi']$ zu gelangen:

$$P([\phi] \rightarrow [\phi']) = \int \mathcal{D}\pi \mathcal{D}\pi' P_G([\pi]) P_M([\phi, \pi] \rightarrow [\phi', \pi']) P_A([\phi, \pi] \rightarrow [\phi', \pi']). \quad (2-15)$$

Für den Beweis der *detailed balance*-Bedingung ist entsprechen (2-8) folgendes zu zeigen:

$$e^{-S([\phi])} P([\phi] \rightarrow [\phi']) = e^{-S([\phi'])} P([\phi'] \rightarrow [\phi]).$$

Dafür ergibt sich mit (2-13) aus (2-15):

$$\begin{aligned} e^{-S([\phi])} P([\phi] \rightarrow [\phi']) &= \int \mathcal{D}\pi \mathcal{D}\pi' \\ &\times \mathcal{N} e^{-\mathcal{H}([\phi, \pi])} P_M([\phi, \pi] \rightarrow [\phi', \pi']) P_A([\phi, \pi] \rightarrow [\phi', \pi']). \end{aligned} \quad (2-16)$$

Analog zum Beweis für die *detailed balance*-Bedingung im Monte-Carlo-Algorithmus auf Seite 18 kann man wegen der Form von P_A zeigen, daß gilt:

$$e^{-\mathcal{H}([\phi, \pi])} P_A([\phi, \pi] \rightarrow [\phi', \pi']) = e^{-\mathcal{H}([\phi', \pi'])} P_A([\phi', \pi'] \rightarrow [\phi, \pi]). \quad (2-17)$$

KAPITEL 2. SIMULATIONSVERFAHREN

Verwendet man diese Beziehung in (2-16) und macht sich zunutze, daß aufgrund der quadratischen Form der Hamilton-Funktion in den Impulsen außerdem $\mathcal{H}([\phi, \pi]) = \mathcal{H}([\phi, -\pi])$ gilt, so erhält man, nachdem die Substitutionen $\pi \rightarrow \pi'$ und $-\pi \rightarrow -\pi'$ durchgeführt worden sind:

$$\begin{aligned} e^{-S([\phi])} P([\phi] \rightarrow [\phi']) &= e^{-S([\phi'])} \int \mathcal{D}\pi \mathcal{D}\pi' \\ &\times P_G([\pi']) P_M([\phi', \pi'] \rightarrow [\phi, \pi]) P_A([\phi', \pi'] \rightarrow [\phi, \pi]) \\ &= e^{-S([\phi'])} P([\phi'] \rightarrow [\phi]). \end{aligned} \quad (2-18)$$

Damit ist gezeigt, daß die *detailed balance*-Bedingung bezüglich der Wirkung gilt. \square

Nach Abschluß des Beweises ist nun auch zu erkennen, daß die Wahl der Wahrscheinlichkeitsverteilung der Zufallsvariablen P_G nicht beliebig ist, sondern eine wichtige Rolle im Beweis spielt: Es muß gelten, daß $P_G \cdot \exp(-S([\phi]))$ bis auf Faktoren gleich $\exp(-\mathcal{H}([\phi, \pi]))$ ist. Wegen der Form der Hamilton-Funktion müssen die Impulse also gaussverteilt sein.

Bemerkung:

Im Gegensatz zu Abschnitt 2.1.3 handelt es sich bei den „Zeitschritten“ des HMC-Algorithmus nicht um die physikalische Zeit des Systems, sondern um eine fiktive Computer-Zeit. Es gibt dafür keine Entsprechung in der realen Welt. Um diesen Unterschied deutlich zu machen, wird die Computerzeit mit τ notiert, im Gegensatz zur physikalischen Zeit t .

Aus Gleichung (2-14) im Beweis für die *detailed balance*-Bedingung folgt, daß die numerische Integration der Bewegungsgleichung reversibel durchgeführt werden muß. Ein Algorithmus, der dies gewährleistet, ist der „*leap frog*“ Integrations-Algorithmus. Dabei werden zunächst die Impulse einen halben Zeitschritt integriert. Dies stellt die Reversibilität sicher. Dann werden jeweils abwechselnd die Impulse und die Felder integriert, um abschließend noch einmal einen Halbschritt in den Impulsen zu berechnen.

Nach (1-13) und (2-10) lauten die diskretisierten Bewegungsgleichungen:

$$\frac{\phi(\tau + \Delta\tau) - \phi(\tau)}{\Delta\tau} = \frac{\partial \mathcal{H}([\phi(\tau), \pi(\tau)])}{\partial \pi(\tau)}, \quad (2-19a)$$

$$\frac{\pi(\tau + \Delta\tau) - \pi(\tau)}{\Delta\tau} = -\frac{\partial \mathcal{H}([\phi(\tau), \pi(\tau)])}{\partial \phi(\tau)} = -\frac{\partial S([\phi(\tau)])}{\partial \phi(\tau)}. \quad (2-19b)$$

Damit ist es möglich, den Hybrid-Monte-Carlo-Algorithmus vollständig anzugeben:

Der HMC-Algorithmus in Kurzform

Die Integration soll über eine Trajektorienlänge $\tau = N\Delta\tau$ ausgeführt werden. Beginnend bei den Startkonfigurationen $[\phi(0), \pi(0)]$ wird $[\phi(0 + n\Delta\tau)]$ bzw. $[\pi((n + 1/2)\Delta\tau)]$ als $[\phi(n)]$ bzw. $[\pi(n + 1/2)]$ notiert.

1. Sinnvolle Wahl der Startwerte für die Konfiguration $[\phi]$, zum Beispiel zufällig-gleichverteilt.
2. Erzeugung der kanonisch konjugierten Felder $[\pi]$, so daß sie nach (??) mit Varianz 1 und Mittelwert 0 gaussverteilt sind.
3. Berechnung eines Halbschrittes für $[\pi]$:

$$\pi_{\mathbf{x}}(1/2) = \pi_{\mathbf{x}}(0) - \frac{\Delta\tau}{2} \frac{\partial S([\phi(0)])}{\partial \phi_{\mathbf{x}}(0)} . \quad (2-20)$$

4. Iteration der beiden folgenden Schritte:

(a)

$$\phi_{\mathbf{x}}(k) = \phi_{\mathbf{x}}(k-1) + \Delta\tau \pi_{\mathbf{x}}(k-1/2) , \quad k = 1, \dots, N . \quad (2-21)$$

(b)

$$\pi_{\mathbf{x}}(k+1/2) = \pi_{\mathbf{x}}(k-1/2) - \Delta\tau \frac{\partial S([\phi(k)])}{\partial \phi_{\mathbf{x}}(k)} , \quad k = 1, \dots, (N-1) . \quad (2-22)$$

5. Berechnung eines abschließenden Halb-Schrittes:

$$\pi_{\mathbf{x}}(N) = \pi_{\mathbf{x}}(N-1/2) - \frac{\Delta\tau}{2} \frac{\partial S([\phi(N)])}{\partial \phi_{\mathbf{x}}(N)} . \quad (2-23)$$

6. Akzeptanz der neuen Konfiguration $[\phi', \pi']$ mit der Wahrscheinlichkeit (2-12).
7. Wiederbeginn entweder mit der neuen oder der alten Konfiguration bei Schritt zwei.

Könnte die Integration exakt ausgeführt werden, so wäre $-\mathcal{H}([\phi', \pi']) + \mathcal{H}([\phi, \pi]) =: -\Delta\mathcal{H}$ identisch Null und die neue Konfiguration würde immer akzeptiert. Doch für numerische Integrationen mit endlichem Zeitschritt $\Delta\tau \neq 0$ ist $\Delta\mathcal{H}$ im allgemeinen ungleich Null. Die systematischen Fehler, die durch den endlichen Zeitschritt entstehen, werden durch den Metropolis-Akzeptanzschritt behoben.

Der Algorithmus ist also im Prinzip exakt. Allerdings ist es wegen der Rundungsfehler, die bei jeder Rechnung mit endlicher Genauigkeit entstehen, nicht möglich, die Reversibilität exakt zu gewährleisten. Damit ist die Exaktheit des gesamten Verfahrens wieder in Frage gestellt.

Erwartungswert von $\exp(-\Delta\mathcal{H})$

Da die Abbildung $[\phi, \pi] \rightarrow [\phi', \pi']$ Maß-erhaltend ist, kann der Erwartungswert von $\exp(-\Delta\mathcal{H})$ analytisch angegeben werden, was sehr gut zur Überprüfung der Richtigkeit einer Implementierung herangezogen werden kann: Es gilt

$$\int \mathcal{D}\phi' \mathcal{D}\pi' e^{-\mathcal{H}([\phi', \pi'])} = \int \mathcal{D}\phi \mathcal{D}\pi e^{-\mathcal{H}([\phi, \pi]) - \Delta\mathcal{H}([\phi, \pi])}.$$

Daraus folgt unmittelbar:

$$1 = \langle e^{-\Delta\mathcal{H}} \rangle \geq e^{-\langle \Delta\mathcal{H} \rangle}, \quad (2-24)$$

wobei die Ungleichung aus der Konvexität der Exponentialfunktion folgt. Über die Bestimmung des Erwartungswertes von $\exp(-\Delta\mathcal{H})$ kann also überprüft werden, ob eine konkrete Implementation des HMC-Algorithmus korrekt ist.

2.1.5 Berechnung von Observablen

Da mit Hilfe des Monte-Carlo-Verfahrens die Konfigurationen so erzeugt werden, daß sie dem thermischen Gleichgewicht des Systems entsprechen, kann man den Erwartungswert für Observablen durch Mitteln über eine Stichprobe erhalten:

$$\langle \mathcal{O} \rangle \approx \frac{1}{M} \sum_{i=1}^M \mathcal{O}_i. \quad (2-25)$$

Die Erwartungswerte sind dann mit entsprechenden statistischen Fehlern der Größenordnung $O(1/\sqrt{M})$ behaftet. Der statistische Fehler und seine Abschätzung wird in Abschnitt 2.4 dargestellt.

2.1.6 Einstell-Parameter

Beim HMC gibt es zwei Parameter, mit deren Hilfe sich die Effizienz des Algorithmus bei gegebenem Volumen beeinflussen läßt: Dies sind die Trajektorienlänge $\tau = N\Delta\tau$ und die Schrittweite $\Delta\tau$.

In einem *leap frog*-Integrationsschritt wird die Energieerhaltung wegen der Diskretisierungsfehler in der Ordnung $O(\Delta\tau^3)$ verletzt [19]:

$$\Delta\mathcal{H} \propto \Delta\tau^3 + O(\Delta\tau^4). \quad (2-26)$$

Darüber hinaus zeigen Gupta et al. in [19], daß der Erwartungswert von $\Delta\mathcal{H}$ bei konstanter Trajektorienlänge und $N > 1$ folgende Abhängigkeit vom Volumen V und der MD-Schrittweite $\Delta\tau$ zeigt:

$$\langle \Delta\mathcal{H} \rangle \propto V\Delta\tau^4. \quad (2-27)$$

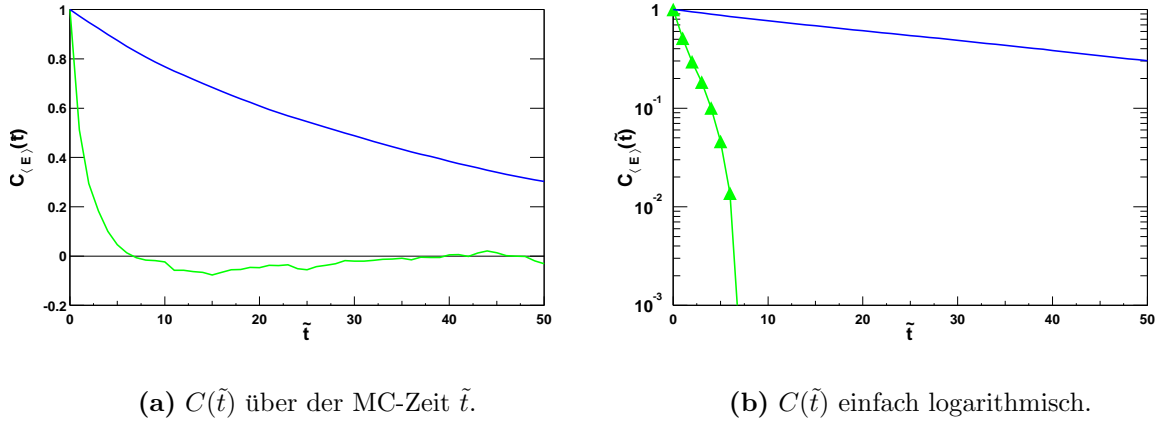


Abbildung 2.1: Autokorrelationsfunktion $C_E(\tilde{t})$ des Energieerwartungswertes für $\tau = 0.5$ und $\tau = 0.1$ berechnet auf einem 12^2 Gitter bei $\beta = 1$ und $\Delta\tau = 0.05$.

Über den Parameter $\Delta\tau$ läßt sich also bei konstantem Volumen die Akzeptanzrate des HMC so einstellen, wie sie benötigt wird, und zwar für beliebige Trajektorienlängen τ . Welche Rolle spielt dabei die Trajektorienlänge? Je größer die Trajektorienlänge gewählt wird, desto weniger korreliert sind die erzeugten Konfigurationen. Gleichzeitig steigt aber auch die benötigte Rechenzeit für die Erstellung einer Konfiguration. Es gilt also, bei der Wahl für τ zwischen diesen beiden Faktoren auszugleichen.

Als Maß für die Abhängigkeit verschiedener Konfigurationen kann die Autokorrelationsfunktion verwendet werden:

Definition 2.1.3 (Autokorrelationsfunktion) Sei $M \in \mathbb{N}$ und $\mathcal{O}(\tilde{t}), \tilde{t} \in \{0, 1, \dots, M\}$ eine Folge von Meßwerten für den Operator \mathcal{O} . Dann nennt man

$$C_{\mathcal{O}}(\tilde{t}) = \frac{\langle \mathcal{O}(\tilde{t}_0 + \tilde{t}) \mathcal{O}(\tilde{t}_0) \rangle - \langle \mathcal{O} \rangle^2}{\langle \mathcal{O}^2 \rangle - \langle \mathcal{O} \rangle^2} \quad (2-28)$$

die Autokorrelationsfunktion zum Operator \mathcal{O} . Bei \tilde{t} handelt es sich weder um die MD-Zeit τ , noch um die physikalische Zeit t , sondern um die sogenannte Monte-Carlo-Zeit, die im wesentlichen die Konfigurationen in der Reihenfolge ihrer Erzeugung indiziert. M ist die Anzahl der Konfigurationen.

In Abbildung 2.1(a) ist die Autokorrelationsfunktion der inneren Energie für verschiedene Trajektorienlängen beim nichtlinearen $O(4)$ - σ -Modell bei $\beta = 1$ auf einem 12^2 Gitter aufgetragen, in 2.1(b) ebenfalls, aber einfach logarithmisch. Es wird ersichtlich, daß die Autokorrelationsfunktion für unterschiedliche Trajektorienlängen exponentiell mit verschiedenen Exponenten abfällt.

Die Autokorrelationszeit nimmt zwar mit steigender Trajektorienlänge ab, die Akzeptanzrate allerdings ebenfalls, was wiederum bedeutet, daß die Autokorrelationszeit steigt. Empirische Untersuchungen zeigen, daß die Wahl der Trajektorienlänge zu etwa 1 dieses Problem gut optimiert.

2.2 Reversibilitätsverletzung

Wie aus dem Beweis für die *detailed balance* auf Seite 21, Gleichung (2-14) hervorgeht, ist der HMC nur dann exakt, wenn die Integration der Bewegungsgleichungen reversibel erfolgt; sie kann aber in der Praxis nicht exakt reversibel ausgeführt werden, da Computer nur eine endliche Zahlendarstellung zulassen. In [3, 20, 21] wird gezeigt, daß diese Reversibilitätsverletzung mit der Trajektorienlänge exponentiell zunimmt. Das exponentielle Wachstum der Verletzung kann durch einen Lyapunov-Exponenten charakterisiert werden.

Bemerkung:

Generell kann die Reversibilitätsverletzung zu einem vernachlässigbaren Problem gemacht werden, indem die Integrationslänge sehr klein gewählt wird. Entsprechend der Überlegungen am Ende des Abschnittes 2.1.6 führt dies aber zu sehr großen Autokorrelationszeiten.

Aufgrund der Tatsache, daß durch die Verletzung der Reversibilität der gesamte Algorithmus nicht mehr exakt ist, stellt sich die Frage, inwiefern dies auch die Ergebnisse der Simulation beeinflußt. Oder anders gefragt: Kann der Fehler, der durch die Reversibilitätsverletzung entsteht, gegenüber dem statistischen Fehler vernachlässigt werden? Diese Frage ist an sich schwer zu beantworten, da kein Vergleich mit Ergebnissen eines exakten Algorithmus zur Verfügung steht. Trotzdem kann versucht werden, den Fehler abzuschätzen.

Dies kann beispielsweise geschehen, indem künstlich Fehler verschiedener Größe hinzugefügt oder indem Rechnungen verschiedener Genauigkeit verglichen werden. Für letzteres bietet sich an, zwischen Rechnung in einfacher (32 Bit) und in doppelter (64 Bit) Gleitkomma-Genauigkeit zu vergleichen, wobei die doppelte Genauigkeit als exakt angenommen wird.

Es folgen einige allgemeine Überlegungen dazu:

Zunächst wird man annehmen, daß im Mittel genauso oft nach oben wie nach unten gerundet wird und sich damit der Rundungsfehler nicht auf die Meßwerte auswirkt. Doch dieser einfache Ansatz gilt nur für Größen, die linear in den mit Fehlern behafteten Werten sind. Schon eine einfache Rechnung zeigt, daß bei einer Größe, die quadratisch von den mit Fehlern behafteten Werten abhängt, ein Mittelwert der Abweichung zu erwarten ist, der ungleich 0 ist, wenn zwar der Mittelwert des Fehlers gleich 0 ist, der des Quadrates

aber nicht:

$$\overline{(x \pm \delta x)^2} - \overline{x^2} = \overline{x^2} \pm 2\overline{x\delta x} + \overline{(\delta x)^2} - \overline{x^2} = \overline{(\delta x)^2} > 0,$$

mit $\overline{2x\delta x} = 2\overline{x}\overline{\delta x} = 0$. Ebenfalls kann man eine Abweichung des Mittelwertes von Null erwarten, wenn eine Größe von zwei verschiedenen Variablen abhängt, die beide fehlerbehaftet sind, deren Fehler aber korreliert sind. Das heißt, daß zum Beispiel $\overline{\delta x_i \delta x_j} \neq 0$ ist, auch wenn wiederum die Mittelwerte der Einzelfehler verschwinden.

Dabei ist es nicht so interessant, die Abweichung in Abhängigkeit von der Integrationslänge zu betrachten, da in Simulationen in der Regel nur die Trajektorienlänge l verwendet wird. Aber die Abhängigkeit von der Größe des Systems gewinnt an Bedeutung.

2.3 Korrelationslänge

Um auf dem Gitter Ergebnisse zu erhalten, die nur wenig mit Diskretisierungsfehlern behaftet sind, ist es notwendig, das Gitter so groß zu wählen, daß die Korrelationslänge deutlich kleiner ist als die Gitter-Ausdehnung. Anschaulich ausgedrückt, stellt die Korrelationslänge die Entfernung auf dem Gitter dar, über die sich zwei Gitter-Punkte noch beeinflussen. Aufgrund der periodischen Randbedingungen leuchtet ein, daß diese Länge der Gitter-Ausdehnung nicht zu nah sein darf.

Die Korrelationslänge bestimmt sich aus dem Kehrwert des *massgap* m , das wiederum über die Zweipunktfunktion:

$$C(t) = \langle \phi(0)\phi(t) \rangle - \langle \phi(0) \rangle \langle \phi(t) \rangle \quad (2-29)$$

bestimmt wird. Diese ist nämlich asymptotisch proportional zu $\exp(-mt)$. Wegen der periodischen Randbedingungen gilt folgende Proportionalität:

$$C(t) \propto \exp(-mt) + \exp(m(L-t)) \propto \cosh(m(t-L/2)), \quad (2-30)$$

wobei L und t in Einheiten der Gitterkonstanten angegeben sind. Eine Herleitung dazu findet sich in Anhang [A.1](#).

Für endliche t trägt nicht nur ein Zustand zur Zweipunktfunktion bei. Da man am *massgap* des tiefsten Zustandes interessiert ist, können nicht alle Meßwerte verwendet werden: Für kleine t (bzw. $(L-t)$) wird die gemessene Zweipunktfunktion eine Überlagerung mehrerer Zustände sein, deren Beiträge aber schneller aussterben als die des Grundzustands. Somit kann man erst ab einem bestimmten t_0 anfitten. Dieses t_0 ist erreicht, falls das Verhältnis

$$\ln \frac{C(t_n)}{C(t_{n+1})} \quad (2-31)$$

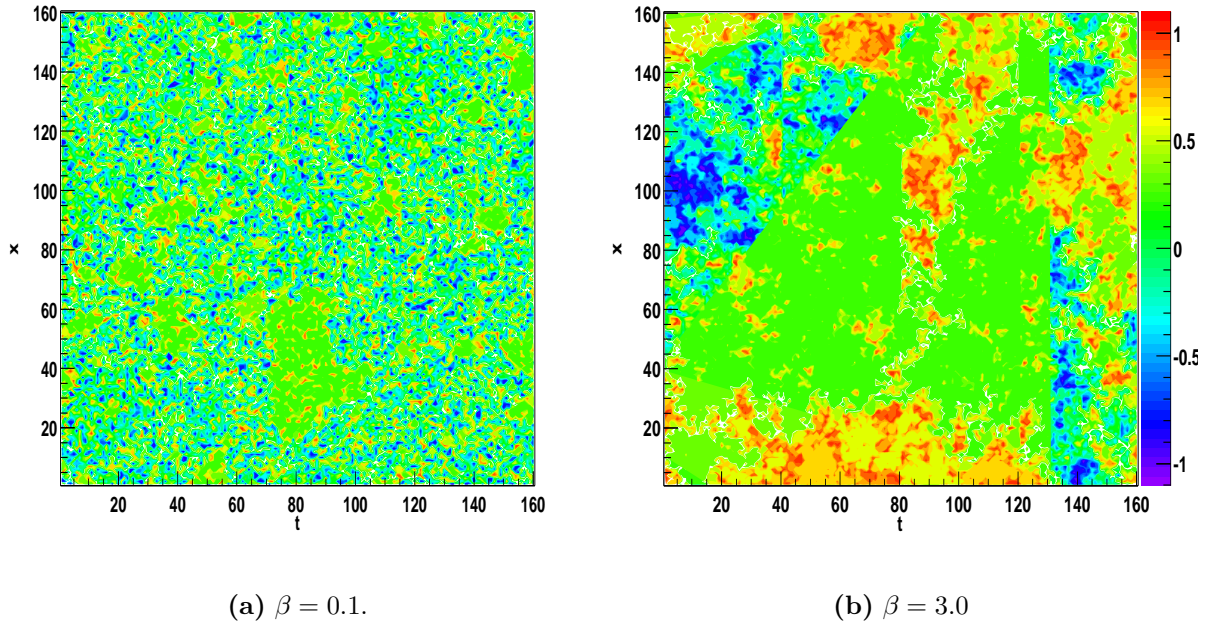


Abbildung 2.2: Korrelation des Feldes jeweils mit dem Feld im Koordinatenursprung $\langle \phi_{0,0}, \phi_{x,t} \rangle$ für zwei thermalisierte Konfigurationen eines 160^2 Gitters bei $\beta = 0.1$ (←LINKS) und $\beta = 3$ (→RECHTS). Die Werteskala geht von -1 über 0 bis 1 .

konstant wird. Um Diskretisierungs-Effekte möglichst zu vermeiden, wird versucht $1/m < L/5$ für die Simulationen zu gewährleisten.

Um die Bedeutung der Korrelationslänge deutlich zu machen, ist in Abbildung 2.2 für zwei jeweils thermalisierte Konfigurationen ($\beta = 0.1$ und $\beta = 3$.) eines 160^2 Gitters für jeden Gitter-Punkt das Skalarprodukt des Feldes am Punkt (x, t) mit dem des Feldes im Koordinatenursprung $(0, 0)$ für zwei verschiedene Werte des Kopplungsparameters aufgetragen. Das Skalarprodukt ist gleich 1 , falls die Felder an beiden Punkten übereinstimmen, und gleich -1 , falls sie entgegengesetzt ausgerichtet sind. Der Wert des Skalarproduktes ist in einer Farbskala von -1 bis 1 dargestellt. Gebiete mit gleicher Farbe sind also etwa gleich ausgerichtet.

Für $\beta = 0.1$ sind nur sehr kleine Gebiete auszumachen, in denen das Feld annähernd gleich ausgerichtet ist, wohingegen für $\beta = 3$ bereits große Gebiete entstanden sind, in denen das Feld ähnlich ausgerichtet ist. Die Korrelationslänge gibt eine Abschätzung für die Größe solcher Gebiete, in denen das Feld gleich ausgerichtet ist.

2.4 Fehlerabschätzung

Bei Computersimulationen mit Monte-Carlo-Methoden müssen verschiedene Fehlerquellen berücksichtigt werden: Neben den bereits in Abschnitt 1.2 beschriebenen Diskretisierungs- und *finite size*-Effekten und den daraus resultierenden Fehlern sind noch zwei weitere Fehler-Arten in Betracht zu ziehen: Der statistischen Fehler und der Rundungsfehler.

Letzterer erwächst aus der endlichen Zahlendarstellung auf dem Computer und ist prinzipiell nicht vermeidbar. Er wird zwar im allgemeinen vernachlässigbar sein, doch wenn er sich, wie in Abschnitt 2.2 diskutiert, stark potenziert, müssen seine Auswirkungen zumindest untersucht werden.

Der statistische Fehler rührt daher, daß die zur Verfügung stehende Rechenzeit endlich ist. Der Mittelwert (2-25) wäre nur dann exakt, wenn die Stichprobe unendlich groß wird. Der statistische Fehler kann abgeschätzt werden, was Inhalt des folgenden Abschnitts ist.

Statistischer Fehler

Zur Abschätzung des statistischen Fehlers kann die mittlere quadratische Abweichung vom exakten Mittelwert verwendet werden. Es ist zu beachten, daß strenggenommen zwei Mittelungsprozesse stattfinden. Erstens wird über das Gitter gemittelt und zweitens über die Konfigurationen. In diesem Abschnitt soll es um den Mittelungsprozeß über die Konfigurationen gehen. Der Meßwert \mathcal{O}_i ist also bereits ein Mittelwert über eine gesamte Konfiguration.

Bei der Mittelung über M Konfigurationen für die Bestimmung der Observablen \mathcal{O} berechnet sich die mittlere quadratische Abweichung¹ zu:

$$\sigma_{\mathcal{O},M}^2 = \frac{1}{M-1} \left(\frac{1}{M} \sum_{i=1}^M \mathcal{O}_i^2 - \left[\frac{1}{M} \sum_{i=1}^M \mathcal{O}_i \right]^2 \right) = \frac{1}{M-1} (\langle \mathcal{O}^2 \rangle - \langle \mathcal{O} \rangle^2). \quad (2-32)$$

Da bei dieser Abschätzung vorausgesetzt wird, daß die Meßwerte unkorreliert sind, ist der Fehler (2-32) im allgemeinen zu klein, wenn wegen der sukzessiven Erzeugung der Konfigurationen jeweils aus der vorhergehenden die Meßwerte nicht unabhängig sind. Er wird deshalb auch als der „naive“ Fehler bezeichnet. Um den Fehler besser abzuschätzen, ist es hilfreich, die integrierte Autokorrelationszeit zu betrachten. Sie kann ähnlich der Autokorrelationsfunktion $C(t)$ (2-28) definiert werden:

¹ Der Nenner $M-1$ ergibt sich daraus, daß auch der exakte Mittelwert $\bar{\mathcal{O}}$ durch die Schätzung $\langle \mathcal{O} \rangle$ ersetzt wurde, da der exakte Mittelwert nicht bekannt ist.

Definition 2.4.1 (Integrierte Autokorrelationszeit) *Unter der integrierten Autokorrelationszeit versteht man die Größe*

$$\tau_{int} = \frac{1}{2} \sum_{\tilde{t}=-\infty}^{\infty} \frac{\Gamma_A(\tilde{t})}{\Gamma_A(0)}. \quad (2-33)$$

Dabei wurde die Autokorrelation $\Gamma_A(|i - j|)$

$$\Gamma_A(|i - j|) = \langle (\mathcal{O}_i - \langle \mathcal{O} \rangle)(\mathcal{O}_j - \langle \mathcal{O} \rangle) \rangle$$

eingeführt, die nur noch vom Abstand $|i - j|$ abhängt. Die Autokorrelation klingt typischerweise exponentiell mit der sogenannten Autokorrelationszeit τ_c ab:

$$\Gamma_A(\tilde{t}) \propto \exp(-\tilde{t}/\tau_c).$$

Es gibt zwei Möglichkeiten, die Korrelation mit Hilfe von τ_{int} bei der Fehlerabschätzung zu berücksichtigen:

Zum einen können bei der Erzeugung der Konfigurationen jeweils der Autokorrelationszeit entsprechend viele Konfigurationen übersprungen werden, bis wieder eine unabhängige Konfiguration zur Messung benutzt wird. In diesem Fall kann der naive Fehler (2-32) als Abschätzung des tatsächlichen Fehlers dienen. Dies ist unter Umständen sehr zeitaufwendig, da $O(\tau_{int})$ viele Konfigurationen erzeugt werden müssen, um eine unkorrelierte zu erhalten.

Zum anderen kann versucht werden, die Autokorrelationszeit bei der Berechnung der Fehler zu berücksichtigen: Es läßt sich zeigen [9], daß sich der Fehler korrelierter Meßwerte wie folgt auf den naiven Fehler zurückführen läßt:

$$\sigma_{\mathcal{O},M}^2 = 2\tau_{int} (\sigma_{\mathcal{O},M}^{naiv})^2. \quad (2-34)$$

Das bedeutet de facto eine Reduktion der Anzahl der Meßwerte auf $M/2\tau_{int}$.

Eine weitere einfache Methode, die Fehler zu analysieren, besteht in der Blockbildung (*binning*). Dabei werden die ursprünglichen Messungen blockweise vorgemittelt:

$$\mathcal{O}_{l,B} = \frac{1}{B} \sum_{i=1+(l-1)B}^{lB} \mathcal{O}_i, \quad l = 1, \dots, M_B = \frac{M}{B}. \quad (2-35)$$

Auch diese Werte sind noch korreliert. Wenn die Blöcke B allerdings genügend lang sind (in Einheiten von τ), so wird dies zu einem vernachlässigbaren Effekt. Mit den „neuen“ Meßwerten kann der Fehler durch

$$\sigma^2 = \frac{1}{M_B(M_B - 1)} \sum_l \left(\mathcal{O}_{l,B} - \frac{1}{M_B} \sum_{l'} \mathcal{O}_{l',B} \right)^2 \quad (2-36)$$

bestimmt werden. Dieser steigt zunächst mit wachsendem B an, um schließlich beim wahren Wert konstant zu werden. Letzteres geschieht allerdings nur dann, wenn die Stichprobe genügend groß ist. Stehen nur Größenordnung τ oder weniger Meßwerte zur Verfügung, so wird das Plateau nicht erreicht. Allerdings ist in diesem Fall auch die Bestimmung der integrierten Autokorrelationszeit unmöglich.

Eine weitere Methode ist das sogenannte *Jackknife binning*, welches dem *binning* sehr ähnlich ist, und lediglich die zu den *binning*-Blöcken komplementären Blöcke verwendet. Die Blöcke sind bei diesem Vorgehen deutlich größer:

$$\mathcal{O}_{l,\bar{B}} = \frac{1}{M-B} \left(\sum_{i=1}^{(l-1)B} \mathcal{O}_i + \sum_{i=lB+1}^M \mathcal{O}_i \right).$$

Es läßt sich mit einer einfachen Rechnung zeigen, daß der Fehler in folgender Weise abgeschätzt werden kann:

$$\sigma^2 = \frac{M_B - 1}{M_B} \sum_l \left(\mathcal{O}_{l,\bar{B}} - \frac{1}{M_B} \sum_{l'} \mathcal{O}_{l',\bar{B}} \right)^2.$$

Das *Jackknife binning* ist insbesondere für sekundäre Größen von Bedeutung [9].

Kapitel 3

Anwendung des HMC-Algorithmus auf die konkreten Modelle

In diesem Kapitel wird der Hybrid-Monte-Carlo-Algorithmus aus Kapitel 2 auf den anharmonischen Oszillator und das nichtlineare $O(4)$ - σ -Modell angewendet.

Für den anharmonischen Oszillator wird nicht der gesamte HMC-Algorithmus dargestellt. Lediglich die Bewegungsgleichungen sollen zwecks Bestimmung des Lyapunov-Exponenten mit Hilfe des Programmpakets MATHEMATICA® integriert werden. Deswegen werden im ersten Abschnitt des folgenden Kapitels nur die Bewegungsgleichungen angegeben.

Ziel der darauffolgenden Abschnitte ist, die Bewegungsgleichungen in verschiedenen Darstellungen für das nichtlineare σ -Modell herzuleiten, denn eine triviale Übertragung ist aufgrund der an die Felder gestellten Zwangsbedingungen nicht möglich. Dabei wird auch eine Formulierung der Modell-Hamilton-Funktion in der jeweiligen Darstellung präsentiert, die den Akzeptanz-Teil des Algorithmus bestimmt.

3.1 HMC für den anharmonischen Oszillator

Die Übertragung der Methode des HMC-Algorithmus auf den anharmonischen Oszillator stellt kein großes Problem dar, da es sich um ein eindimensionales, skalares Modell handelt. Um die Modell-Hamilton-Funktion definieren zu können, werden reelle Impulse π_i eingeführt, die die Rolle der konjugierten Größen zu den Feldern ϕ_i spielen. Sie werden zu Beginn der Integration zufällig-gaussverteilt gewählt. Damit kann die Modell-Hamilton-Funktion zu

$$\mathcal{H} = \sum_{i=1}^L \frac{1}{2} \pi_i^2 + S([\varphi]) \quad (3-1)$$

definiert werden. Mit der Definition (1-20) für die diskrete Wirkung erhält man aus den Bewegungsgleichungen (2-10) folgende Gleichungen für den anharmonischen Oszillator:

$$\dot{\varphi}_i = \frac{\partial \mathcal{H}}{\partial \pi_i} = \pi_i, \quad (3-2a)$$

$$\dot{\pi}_i = -\frac{\partial S}{\partial \varphi_i} = -[-\varphi_{i-1} - \varphi_{i+1} + 2\varphi_i + \omega^2 \varphi_i + 4\varphi_i^3]. \quad (3-2b)$$

Dabei wurde $\lambda = 1$ und $a = 1$ gesetzt. Der Parameter ω wird rein imaginär gewählt und das Verhalten der Gleichungen in Abhängigkeit von ω^2 untersucht.

Da insbesondere der Lyapunov-Exponent von Interesse ist, und nicht so sehr eine Simulation des Modells, sollen obige Bewegungsgleichungen mit Hilfe von MATHEMATICA® auf ihr chaotisches Verhalten untersucht werden. Die Benutzung von MATHEMATICA® hat den Vorteil, daß einerseits die Genauigkeit beliebig einstellbar ist, und es sich andererseits um einen vom für das σ -Modell verwendeten Algorithmus verschiedenen Algorithmus für die Integration handelt.

3.2 HMC für das nichtlineare $O(4)$ - σ -Modell

In den folgenden Unterabschnitten werden zwei verschiedene Möglichkeiten vorgestellt, die an die Felder gestellten Zwangsbedingungen (1-25) während der numerischen Integration aufrechtzuerhalten.

Es werden zwei grundsätzlich verschiedene Möglichkeiten betrachtet, dies umzusetzen: Erstens mit Hilfe der sogenannten Lagrangeschen Multiplikatoren, also mit Hilfe eines Zusatz-Terms in der Modell-Hamilton-Funktion. Und zweitens, indem man die Felder mit Hilfe der Gruppe $SU(2, \mathbb{C})$ darstellt, so daß die Zwangsbedingungen aufgrund der Gruppen-Eigenschaften eingehalten werden.

In der vorliegenden Arbeit wurden zunächst beide Möglichkeiten implementiert, um Vor- und Nachteile testen zu können. Aufgrund der größeren Ähnlichkeit der zweiten Methode mit den Methoden, die in der QCD zum Einsatz kommen, wurden dann aber die eigentlichen Untersuchungen mit Hilfe dieser Methode durchgeführt. In der Praxis stellte sie sich auch als die brauchbarere Methode (oder Implementierung) heraus.

3.2.1 Lagrangesche Multiplikatoren

Diese Methode greift auf Erkenntnisse der theoretischen Mechanik zurück: Um Zwangsbedingungen $F(\phi_{\mathbf{x}}) = 0$ in der Mechanik aufrechtzuerhalten, werden sogenannte Lagrangesche Multiplikatoren $\lambda_{\mathbf{x}}$ eingeführt:

$$S_{\sigma}^L[\phi] = S[\phi] + \sum_{\mathbf{x}} \lambda_{\mathbf{x}} F(\phi_{\mathbf{x}}). \quad (3-3)$$

3.2. HMC FÜR DAS NICHTLINEARE $O(4)$ - σ -MODELL

Im konkreten Fall mit der Zwangsbedingung $F(\phi_{\mathbf{x}})$:

$$F(\phi_{\mathbf{x}}) = \phi_{\mathbf{x}}^2 - 1 = 0 \quad (3-4)$$

und der Wirkung $S([\phi])$:

$$S([\phi]) = -2\beta \sum_{\mathbf{x}, \mu} \phi_{\mathbf{x}} \phi_{\mathbf{x}+\mu}. \quad (3-5)$$

Dies kann wie in der klassischen Mechanik als die Einführung von Zwangskräften interpretiert werden, die genau so konstruiert werden, daß sie für die Einhaltung der Zwangsbedingung sorgen. Die Modell-Hamilton-Funktion für den Hybrid-Monte-Carlo ergibt sich wie in (2-11) vorgegeben zu:

$$\mathcal{H}[\phi, \pi] = \sum_{\mathbf{x}} \frac{1}{2} \pi_{\mathbf{x}}^2 + S_{\sigma}^L[\phi], \quad (3-6)$$

diesmal mit der veränderten Wirkung (3-3). Daraus ergeben sich die Bewegungsgleichungen zu:

$$\dot{\phi}_{\mathbf{x}}^{\alpha} = \frac{\partial \mathcal{H}[\phi, \pi]}{\partial \pi_{\mathbf{x}}^{\alpha}} = \pi_{\mathbf{x}}^{\alpha} \quad (3-7a)$$

$$\dot{\pi}_{\mathbf{x}}^{\alpha} = -\frac{\partial \mathcal{H}[\phi, \pi]}{\partial \phi_{\mathbf{x}}^{\alpha}} = -\frac{\partial S[\phi]}{\partial \phi_{\mathbf{x}}^{\alpha}} - \lambda_{\mathbf{x}} \frac{\partial F(\phi_{\mathbf{x}})}{\partial \phi_{\mathbf{x}}^{\alpha}}. \quad (3-7b)$$

Bildet man die zweite Ableitung nach der Zeit von $F(\phi_{\mathbf{x}})$ und setzt $\dot{\pi}$ aus der Bewegungsgleichung ein

$$\begin{aligned} 0 &= \frac{\partial^2 F(\phi_{\mathbf{x}})}{\partial t^2} = 2\phi_{\mathbf{x}} \ddot{\phi}_{\mathbf{x}} + 2\dot{\phi}_{\mathbf{x}} \dot{\phi}_{\mathbf{x}} \\ &= \sum_{\alpha} \left(-2\phi_{\mathbf{x}}^{\alpha} \frac{\partial S}{\partial \phi_{\mathbf{x}}^{\alpha}} - 2\phi_{\mathbf{x}}^{\alpha} \lambda_{\mathbf{x}} \frac{\partial F}{\partial \phi_{\mathbf{x}}^{\alpha}} + 2(\pi_{\mathbf{x}}^{\alpha})^2 \right), \end{aligned} \quad (3-8)$$

so bestimmt man $\lambda_{\mathbf{x}}$ zu:

$$\lambda_{\mathbf{x}} = \frac{\sum_{\alpha} \left((\pi_{\mathbf{x}}^{\alpha})^2 - \phi_{\mathbf{x}}^{\alpha} \frac{\partial S[\phi]}{\partial \phi_{\mathbf{x}}^{\alpha}} \right)}{\sum_{\alpha} \phi_{\mathbf{x}}^{\alpha} \frac{\partial F(\phi_{\mathbf{x}})}{\partial \phi_{\mathbf{x}}^{\alpha}}} = \frac{1}{2} \sum_{\alpha} \left((\pi_{\mathbf{x}}^{\alpha})^2 - \phi_{\mathbf{x}}^{\alpha} \frac{\partial S[\phi]}{\partial \phi_{\mathbf{x}}^{\alpha}} \right). \quad (3-9)$$

Damit ergibt sich die Bewegungsgleichung für π unter Ausnutzung der Zwangsbedingung:

$$\dot{\pi}_{\mathbf{x}}^{\alpha} = -\frac{\partial S}{\partial \phi_{\mathbf{x}}^{\alpha}} - \sum_{\beta} \left((\pi_{\mathbf{x}}^{\beta})^2 - \phi_{\mathbf{x}}^{\beta} \frac{\partial S[\phi]}{\partial \phi_{\mathbf{x}}^{\beta}} \right) \phi_{\mathbf{x}}^{\alpha}. \quad (3-10)$$

Die Gleichungen des HMC-Algorithmus

Die Gleichungen des Hybrid-Monte-Carlo-Algorithmus ergeben sich somit entsprechend (2-20, 2-21, 2-22):

$$\phi_{\mathbf{x}}^{\alpha}(n+1) = \phi_{\mathbf{x}}^{\alpha}(n) + \Delta\tau\pi_{\mathbf{x}}^{\alpha}(n+1/2) \quad (3-11a)$$

$$\pi_{\mathbf{x}}^{\alpha}(n+1/2) = \pi_{\mathbf{x}}^{\alpha}(n-1/2) - \Delta\tau \left[\frac{\partial S[\phi(n+1)]}{\partial \phi_{\mathbf{x}}^{\alpha}(n+1)} - 2\lambda_{\mathbf{x}}(n+1)\phi_{\mathbf{x}}^{\alpha}(n+1) \right]. \quad (3-11b)$$

Die Zwangsbedingung wird natürlich nur für $\lim \Delta\tau \rightarrow 0$ exakt erfüllt. Die Einhaltung kann verbessert werden, indem man die ϕ nach jedem Simulationsschritt wieder auf 1 normiert.

3.2.2 Darstellung mit $SU(2, \mathbb{C})$

Die Idee, die Sphäre S_3 auf die $SU(2, \mathbb{C})$ abzubilden und damit generalisierte Größen einzuführen, rührt daher, daß die Elemente jeweils drei reelle Freiheitsgrade haben, die Handhabung der $SU(2, \mathbb{C})$ aber einfacher ist, da es sich um eine multiplikative Gruppe handelt. Das bedeutet, daß keine Lagrangeschen Multiplikatoren mehr nötig sind, weil die Gruppen-Eigenschaften garantieren, daß man die Gruppe nicht verläßt.

Dafür ist eine Abbildung zu finden, die S_3 auf die Gruppe $SU(2, \mathbb{C})$ abbildet und dabei alle Informationen erhält. Dafür muß insbesondere das Skalarprodukt - und damit die Norm - erhalten bleiben. In diesem Fall lassen sich die Gruppen-Eigenschaften ausnutzen, um die Dynamik des Modells zu beschreiben.

Zunächst die Definition der Gruppe:

Definition 3.2.1 ($SU(2, \mathbb{C})$) *Die Menge aller unitären 2×2 Matrizen der Determinante 1 nennt man $SU(2, \mathbb{C})$. Sie bilden eine Untergruppe von $U(2, \mathbb{C})$, der Menge aller unitären 2×2 Matrizen.*

Das Skalarprodukt der komplexen 2×2 -Matrizen wird folgendermaßen durch

$$\langle U, V \rangle = \frac{1}{2} \operatorname{Re}(\operatorname{Sp}(UV^{\dagger})) \quad (3-12)$$

definiert. Die Erzeugenden der Gruppe $SU(2, \mathbb{C})$ sind die Pauli-Matrizen, allerdings mit einem Vorfaktor i und gemeinsam mit der Identität \mathbb{I} der 2×2 -Matrizen. Sie bilden ebenfalls mit der Identität \mathbb{I} die Basis der vierdimensionalen Algebra über \mathbb{R} der Quaternionen und sind über folgende Multiplikationstafel definiert [22]:

$$\mathbb{I}\sigma_i = \sigma_i\mathbb{I} = \sigma_i, \quad \sigma_i^2 = \mathbb{I}, \quad \sigma_i\sigma_j = -\sigma_j\sigma_i = \sigma_k, \quad (3-13)$$

wobei $\{i, j, k\}$ eine gerade Permutation über $\{1, 2, 3\}$ bilden.

3.2. HMC FÜR DAS NICHTLINEARE $O(4)$ - σ -MODELL

Eine mögliche Darstellung der Pauli-Matrizen ist die folgende:

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (3-14)$$

Damit ist eine Darstellung des Modells mit Hilfe von $SU(2, \mathbb{C})$ möglich:

Darstellung mit $SU(2, \mathbb{C})$

Zuerst ist zu zeigen, daß eine Abbildung von S_3 auf $SU(2, \mathbb{C})$ ohne Informationsverlust möglich ist: Sei S_3 die dreidimensionale Sphäre

$$S_3 = \{\vec{x} \in \mathbb{R}^4 \mid |\vec{x}| = 1\} \subset \mathbb{R}^4.$$

Definiert man die Abbildung $T : S_3 \mapsto SU(2, \mathbb{C})$ mit

$$T(\vec{x}) = x_4 \mathbb{I} + i(x_1 \sigma_1 + x_2 \sigma_2 + x_3 \sigma_3) \quad \forall \vec{x} \in S_3, \quad (3-15)$$

so gilt mit Lemma A.2.1:

$$\langle \vec{x}, \vec{y} \rangle = \langle T(\vec{x}), T(\vec{y}) \rangle, \quad (3-16)$$

Die Norm, die über das Skalarprodukt definiert ist, bleibt also unter der Abbildung T erhalten und sowie T als auch T^{-1} sind damit in natürlicher Weise stetig. Zusätzlich ist T ein Homöomorphismus von S_3 nach $SU(2, \mathbb{C})$, wie der folgende Satz zeigt:

Satz 3.2.1 *Sei S_3 die dreidimensionale Sphäre und T definiert wie in Gleichung (3-15). Dann ist T ein (Gruppen-)Homöomorphismus von S_3 auf $SU(2, \mathbb{C})$.*

Ein Beweis ist in Anhang A.2 angegeben. Ebenfalls in Anhang A.2 wird eine konkrete Darstellung vorgestellt, um der Tatsache Rechnung zu tragen, daß eine $SU(2, \mathbb{C})$ -Matrix lediglich drei Freiheitsgrade besitzt, so daß bei der Implementierung ein Faktor Zwei an Rechenoperationen gespart werden kann.

Die Wirkung

Um die Darstellung mit $SU(2, \mathbb{C})$ zu vervollständigen, ist als nächstes die Wirkung zu übertragen: Dabei wird der Übergang von den Vektoren $\phi_{\mathbf{x}}$ zu den $SU(2, \mathbb{C})$ -Matrizen $U_{\mathbf{x}}$ vollzogen. Mit Hilfe des Skalarproduktes (3-12) kann die diskrete Wirkung des $O(4)$ - σ -Modells (1-28) in folgender Weise nach $SU(2, \mathbb{C})$ übertragen werden:

$$-2\beta \sum_{\mathbf{x}, \mu, \alpha} \phi_{\mathbf{x}}^{\alpha} \phi_{\mathbf{x}+\mu}^{\alpha} \quad \longrightarrow \quad -2\beta \sum_{\mathbf{x}, \mu} \langle U_{\mathbf{x}}, U_{\mathbf{x}+\mu} \rangle \in \mathbb{R}.$$

Also ergibt sich die Wirkung zu:

$$S[U] = -\beta \operatorname{Re} \left(\sum_{\mathbf{x}, \mu} \operatorname{Sp}(U_{\mathbf{x}} U_{\mathbf{x}+\mu}^{\dagger}) \right). \quad (3-17)$$

Um auch die Gleichungen des HMC-Algorithmus in dieser Notation ausdrücken zu können, muß die Ableitung der Wirkung nach $U_{\mathbf{x}} \in SU(2, \mathbb{C})$ definiert werden:

Definition 3.2.2 (Ableitung nach $U_{\mathbf{x}} \in SU(2, \mathbb{C})$ in Richtung j)

$$D_{\mathbf{x}j} f(U_{\mathbf{x}}) = \frac{\partial}{\partial \alpha} f(e^{i\alpha \sigma_j} U_{\mathbf{x}})_{\alpha=0}. \quad (3-18)$$

Damit gilt für die Ableitung der Wirkung an der Stelle \mathbf{x} in Richtung j :

$$\begin{aligned} D_{\mathbf{x}j} S[U] &= D_{\mathbf{x}j} \left(-\frac{1}{2} \beta \sum_{\mathbf{x}, \mu} \operatorname{Sp}[U_{\mathbf{x}} U_{\mathbf{x}+\mu}^{\dagger} + U_{\mathbf{x}}^* (U_{\mathbf{x}+\mu}^{\dagger})^*] \right) \\ &= -\frac{\beta}{2} \sum_{\mu} \operatorname{Sp}[i\sigma_j U_{\mathbf{x}} U_{\mathbf{x}+\mu}^{\dagger} - i\sigma_j^* U_{\mathbf{x}}^* (U_{\mathbf{x}+\mu}^{\dagger})^* \\ &\quad - (i\sigma_j U_{\mathbf{x}-\mu} U_{\mathbf{x}}^{\dagger} - i\sigma_j^* U_{\mathbf{x}-\mu}^* (U_{\mathbf{x}}^{\dagger})^*)]. \end{aligned}$$

Dabei wurde der Realteil als Summe mit dem komplex Konjugierten ausgedrückt. j läuft von 1 bis 3. Zusammenfassend kann man schreiben:

$$D_{\mathbf{x}j} S[U] = \frac{\beta}{2} \operatorname{Im} \left(\sum_{\mu} \operatorname{Sp} \left[\sigma_j (U_{\mathbf{x}} U_{\mathbf{x}+\mu}^{\dagger} - U_{\mathbf{x}-\mu} U_{\mathbf{x}}^{\dagger}) \right] \right). \quad (3-19)$$

Darstellung der Impulse

Wie sind die zu den $SU(2, \mathbb{C})$ -Feldern konjugierten Impulse zu wählen? Es müssen verschiedene Bedingungen erfüllt werden: Veränderungen an den Feldern können nur Drehungen sein, da die Felder sich auf der drei-dimensionalen Sphäre befinden. Das heißt, daß es sich bei den Impulsen $P_{\mathbf{x}}$ um Drehungen handeln muß. Ferner muß die Hamilton-Funktion ein Skalar und \dot{U} wieder aus $SU(2, \mathbb{C})$ (der Tangentialraum zu $SU(2, \mathbb{C})$ ist wieder $SU(2, \mathbb{C})$) sein, damit sich das System nicht von der S_3 wegbewegt.

Die Pauli-Matrizen sind die Erzeuger infinitesimaler Drehungen. Endliche Drehungen können mit $\exp(i \sum_j \alpha_j \sigma_j)$, $\alpha_j \in \mathbb{R}$ erzeugt werden. Dabei kann auch noch $\Delta\tau$ eingefügt werden: $\exp(i \Delta\tau \sum_j \alpha_j \sigma_j)$. Wählt man nun für jeden Gitter-Punkt \mathbf{x} drei standard-gaussverteilte Zufallszahlen p_i und erzeugt Matrizen $P_{\mathbf{x}}$ daraus zu:

$$P_{\mathbf{x}} = \sum_{i=1}^3 p_i \sigma_i, \quad (3-20)$$

so ist

$$\sum_{\mathbf{x}} \text{Sp}(P_{\mathbf{x}} P_{\mathbf{x}}^{\dagger}) = \sum_{\mathbf{x}} \text{Sp}(P_{\mathbf{x}}^2) = \sum_i 2p_i^2. \quad (3-21)$$

Damit ergibt sich für die Modell-Hamilton-Funktion folgender skalarer Ausdruck:

$$\mathcal{H} = \sum_{\mathbf{x}} \frac{1}{4} \text{Sp}(P_{\mathbf{x}} P_{\mathbf{x}}^{\dagger}) + S = \sum_i \frac{1}{2} p_i^2 + S. \quad (3-22)$$

Mit der Wahl (3-20) ist $\exp(i\Delta\tau P_{\mathbf{x}}) \in SU(2, \mathbb{C})$, weil $P_{\mathbf{x}}$ hermitesch ist, denn dann gilt:

$$\exp(iP_{\mathbf{x}})^{\dagger} = \exp(-iP_{\mathbf{x}}^{\dagger}) = \exp(-iP_{\mathbf{x}}).$$

Darüber hinaus ist $P_{\mathbf{x}}$ spurlos, womit gilt:

$$\det(\exp(P_{\mathbf{x}})) = \exp(\text{Sp}(P_{\mathbf{x}})) = \exp(0) = 1.$$

Gleichungen des Hybrid-Monte-Carlo-Algorithmus

Somit wurde neben den Impulsen auch die Modell-Hamilton-Funktion in die Darstellung mit $SU(2, \mathbb{C})$ übertragen, und es können nun die Gleichungen des HMC vollständig in diesem Formalismus angegeben werden:

1. Für den ersten Halbschritt in den Impulsen:

$$\begin{aligned} P_{\mathbf{x}}(\Delta\tau/2) &= P_{\mathbf{x}}(0) - \sum_j \frac{\Delta\tau}{2} D_{\mathbf{x}j} S(U_{\mathbf{x}}(0)) \sigma_j \\ &= P_{\mathbf{x}}(0) - \frac{\Delta\tau}{2} \vec{D}_{\mathbf{x}} S(U_{\mathbf{x}}(0)) \vec{\sigma}. \end{aligned} \quad (3-23)$$

2. Für $k = 0, \dots, N - 1$ bzw. $k = 1, \dots, N - 1$:

$$\begin{aligned} U_{\mathbf{x}}(k\Delta\tau + \Delta\tau) &= \exp(i\Delta\tau P_{\mathbf{x}}((k + 1/2)\Delta\tau)) U_{\mathbf{x}}(k\Delta\tau) \\ P_{\mathbf{x}}(k\Delta\tau + \Delta\tau/2) &= P_{\mathbf{x}}(k\Delta\tau) - \Delta\tau \vec{D}_{\mathbf{x}} S(U_{\mathbf{x}}(k\Delta\tau)) \vec{\sigma}. \end{aligned} \quad (3-24)$$

3. Abschließend ein weiterer Halbschritt in den Impulsen

$$P_{\mathbf{x}}(N\Delta\tau) = P_{\mathbf{x}}(N\Delta\tau - \Delta\tau/2) - \frac{\Delta\tau}{2} \vec{D}_{\mathbf{x}} S(U_{\mathbf{x}}(N\Delta\tau)) \vec{\sigma}. \quad (3-25)$$

Die Exponentialfunktion in Gleichung (3-24) wird in der Implementierung mit Hilfe einer endlichen Summe gehandhabt, was bei kleinen „Zeitschritten“ hinreichend genau ist:

$$\exp(i\Delta\tau P_{\mathbf{x}}) \approx \mathbb{I} + \sum_{j=1}^m \frac{1}{j!} (i\Delta\tau P_{\mathbf{x}})^j, \quad m > 1.$$

Zusätzlich werden die Determinanten der Matrizen U nach jedem „Zeitschritt“ $\Delta\tau$ wieder auf 1 normiert, um zu vermeiden, daß sich die Felder von der Gruppe $SU(2, \mathbb{C})$ wegbegeben. Es ist nur die Normierung der Determinante notwendig, da die Darstellung der Matrizen (siehe Anhang A.2) die Eigenschaft der Unitarität garantiert.

Kapitel 4

Ergebnisse und Diskussion

Im folgenden Kapitel sollen zunächst Ergebnisse präsentiert werden, die sicherstellen, daß die verwendete Implementierung für das nichtlineare $O(4)$ - σ -Modell innerhalb der statistischen Fehler richtige Ergebnisse liefert. Die Bestimmung des *massgaps* m im darauffolgenden Abschnitt dient der Abschätzung der *finite size*-Effekte und der Bestimmung des Skalierungsverhaltens.

In den anschließenden Abschnitten werden die Untersuchungen zum Lyapunov-Exponenten und dessen Skalierungsverhalten sowie Vergleichsdaten zwischen einfacher und doppelter Rechengenauigkeit und der Auswirkung der Reversibilitätsverletzung auf Observablen präsentiert.

Alle in diesem Kapitel vorgestellten Ergebnisse sind für das nichtlineare $O(4)$ - σ -Modell gewonnen worden, mit Ausnahme der Ergebnisse im Abschnitt 4.3.2: Das Skalierungsverhalten des Lyapunov-Exponenten wurde nicht nur für das nichtlineare $O(4)$ - σ -Modell untersucht, sondern auch für den anharmonischen Oszillator.

Die Ergebnisse für das nichtlineare $O(4)$ - σ -Modell wurden mit einem Programm erzeugt, das auf dem $SU(2, \mathbb{C})$ -Formalismus entsprechend Abschnitt 3.2.2 aufbaut. Die Verwendung von Lagrangeschen Multiplikatoren wurde zu Vergleichszwecken ebenfalls getestet, und das entsprechende Programm lieferte korrekte Resultate: Allerdings zeigte es sich, daß die Akzeptanzrate nur von der Trajektorienlänge $N\Delta\tau$ abhängig war und damit Akzeptanzrate und Autokorrelation nur sehr schlecht aufeinander abgestimmt werden konnten. Diese Tatsache und die größere Ähnlichkeit mit Implementierungen der QCD gab den Ausschlag für die Benutzung des $SU(2, \mathbb{C})$ -Formalismus.

Doch bevor die Ergebnisse in den folgenden Abschnitten dargestellt werden, werden an dieser Stelle einige Eigenschaften und Parameter der verwendeten Implementierung vorgestellt.

Die Parameter für das nichtlineare $O(4)$ - σ -Modell sind in Tabelle 4.1 zusammengestellt: Angegeben sind für verschiedene Gittergrößen und Kopplungsparameter die MD-Schritt-

KAPITEL 4. ERGEBNISSE UND DISKUSSION

| L | β | $\Delta\tau$ | N | A | R | $R/(NL^2 10^{-7})$ |
|-----|---------|--------------|------|--------|-------|--------------------|
| 160 | 1. | 0.008 | 125 | 99.08% | 1.38 | 4.31 |
| | 2. | 0.008 | 125 | 98.1% | 1.37 | 4.28 |
| | 3. | 0.00625 | 160 | 97.98% | 1.71 | 5.34 |
| 320 | 2. | 0.004 | 250 | 99.1% | 10.95 | 4.27 |
| | 3. | 0.0025 | 400 | 99.4% | 16.49 | 4.03 |
| | 3.5 | 0.0025 | 400 | 99.05% | 16.45 | 4.02 |
| | 4. | 0.002 | 500 | 99.42% | 22.38 | 4.37 |
| 500 | 2.5 | 0.002 | 500 | 99.6% | 49.82 | 3.99 |
| | 3. | 0.002 | 500 | 99.25% | 49.31 | 3.94 |
| | 3.5 | 0.002 | 500 | 99.1% | 49.88 | 3.99 |
| | 4. | 0.001 | 1000 | 99.60% | 99.37 | 3.97 |

Tabelle 4.1: Parameter des HMC-Algorithmus für das σ -Modell für verschiedene Gittergrößen L . Es ist der Kopplungsparameter β , die Schrittweite $\Delta\tau$, die Schrittzahl N sowie die Akzeptanzrate A und die Laufzeit R in Sekunden des Programms pro Konfiguration und Prozessor in Sekunden auf 16 Prozessoren des PC-Clusters des DESY Zeuthen angegeben.

weite $\Delta\tau$, die MD-Schrittzahl N und die damit erzielte Akzeptanzrate A . Außerdem ist die Rechenzeit R in Sekunden pro Konfiguration bzw. pro Konfiguration normiert mit NL^2 auf 16 Prozessoren des PC-Clusters des DESY Zeuthen angegeben. Mit „Rechenzeit“ ist an dieser Stelle die Zeit gemeint, die benötigt wird, eine Konfiguration zu erzeugen, ohne Beachtung der Autokorrelation.

Die Rechenzeit R sollte von der Gittergröße L quadratisch und von der Anzahl der MD-Schritte N linear abhängen, so daß gilt:

$$R \propto NL^2.$$

Dieser Zusammenhang kann aufgrund der Daten in Tabelle 4.1 bestätigt werden.

Bei steigender Gittergröße L muß $\Delta\tau$ verkleinert werden, wenn die Akzeptanzrate gleich bleiben soll, weil $\Delta\mathcal{H}$ sowohl von der Gittergröße als auch von der Schrittweite abhängt. Der Zusammenhang von Schrittweite und Gittergröße bei konstanter Akzeptanzrate (also konstantem $\langle\Delta\mathcal{H}\rangle$) ist entsprechend (2-27):

$$\Delta\tau \propto L^{-1/2}.$$

Auch dieser Zusammenhang ist – obwohl der Tabelle 4.1 nicht direkt zu entnehmen – mit den angegebenen Daten verträglich. Die Schrittweite $\Delta\tau$ bei konstanter Akzeptanzrate hängt aber nicht nur von der Gittergröße, sondern auch von der Kopplungskonstante β ab, da die Modell-Hamilton-Funktion mit steigendem β betraglich ebenfalls größer wird.

Da es sich, wie später erläutert wird, als notwendig erweist, sehr große Gitter zu untersuchen, sind wegen der gerade geschilderten Zusammenhänge lange Rechenzeiten zu erwarten. Deswegen stellt es sich als sinnvoll heraus, die Implementierung mit Hilfe des *Message Passing Interface* (MPI) zu parallelisieren, um die effektive Rechenzeit zu verkürzen. Gerechnet wurde auf den PC-Clustern des DESY Zeuthen und am Zentralinstitut für Angewandte Mathematik (ZAM) am FZ Jülich. Die in Tabelle 4.1 angegebenen Rechenzeiten R/NL^2 sind nahezu unabhängig von der Gittergröße. Das zeigt, daß für diese konkrete Implementierung eines zweidimensionalen Problems die Kommunikation zwischen den Prozessoren keine wesentliche Rolle für die Rechenzeit spielt, denn es ist zu erwarten, daß die Kommunikation bei festgehaltener Prozessorenzahl stark von der Gittergröße abhängt. Die leichte Abnahme von $R/(NL^2)$ ist folgendermaßen zu erklären: Bei steigender Gittergröße ist die Anzahl der Randpunkte gegenüber der Anzahl aller Gitterpunkte immer stärker vernachlässigbar. Da die Kommunikation von der Anzahl der Randpunkte abhängt, ist auch die Kommunikation gegenüber der restlichen Rechenzeit, die von der Anzahl aller Gitterpunkte abhängt vernachlässigbar, wenn das Gitter nur groß genug ist. Die in Tabelle 4.1 angegebenen Akzeptanzraten sind relativ hoch, auch Werte um 80% wären akzeptabel. Das bedeutet, daß die Konfigurationen stärker korreliert sind, als bei Akzeptanzraten um 80%. Allerdings reagierte die vorliegende Implementierung sehr empfindlich auf Veränderungen in $\Delta\tau$: Es war unproblematisch, ein $\Delta\tau$ zu finden, so daß $N\Delta\tau = 1$ gilt und die Akzeptanzrate hoch ist. Letztere war dann meist bereits im Bereich um 98% und blieb auch so hoch, wenn man $\Delta\tau$ moderat vergrößerte. Der Übergang hin zu kleinen Akzeptanzraten fand dann allerdings innerhalb eines sehr schmalen Intervalls von $\Delta\tau$ statt, so daß es erheblichen Mehraufwand bedeutet hätte, die Werte auf Akzeptanzraten um 80% einzustellen.

Für die jeweiligen Messungen wurden nur die Konfigurationen benutzt, die auch akzeptiert wurden, was keine notwendige Einschränkung ist, sondern lediglich die Autokorrelationszeit verkürzt. Einige Implementierungsdetails sind in Anhang B dargestellt.

Ferner sind in allen Abbildungen dieses Kapitels jeweils Fehlerbalken miteingezeichnet. Oft sind die Fehler allerdings so klein, daß die Fehlerbalken nicht sichtbar sind.

| T | L | β | E_W | E_W aus [23] |
|-----|-----|---------|-------------|----------------|
| 128 | 128 | 2.1 | 0.60102(13) | 0.60081(5) |
| 128 | 128 | 2.2 | 0.62248(23) | 0.62261(4) |
| 128 | 128 | 2.3 | 0.64198(16) | 0.64198(2) |
| 256 | 256 | 2.4 | 0.65945(10) | 0.65946(4) |
| 256 | 256 | 2.5 | 0.67517(9) | 0.67518(3) |

Tabelle 4.2: Vergleich der Ergebnisse mit Resultaten aus [23].

4.1 Verifikation der Implementierung

Um sicherzustellen, daß die benutzte Implementierung innerhalb der statistischen Fehler die richtigen Ergebnisse liefert, wurde unter anderem mit Resultaten für E_W von U. Wolff aus [23] verglichen, die mit einem *Cluster*-Algorithmus berechnet wurden.

In Tabelle 4.2 sind die Ergebnisse aus [23] neben den Daten aus der vorliegenden Arbeit dargestellt. In [23] wird zwischen Gitterausdehnung in Zeitrichtung T und in Raumrichtung L unterschieden. Deshalb ist in Tabelle 4.2 neben L auch T angegeben.

Die Übereinstimmung für den Energieerwartungswert ist innerhalb der statistischen Fehler sehr gut. Dabei ist zu bemerken, daß sich der Energieerwartungswert als numerisch sehr stabile Observable herausgestellt hat. Unabhängig von der gewählten Gittergröße und den Systemparametern wurden nur geringe Fluktuationen des Mittelwertes um den wahren Wert beobachtet.

Die Fehler in Tabelle 4.2 sind mit Hilfe des *binning*-Verfahrens entsprechend Abschnitt 2.4 berechnet worden, nachdem jeweils jede 20. akzeptierte Konfiguration für die Berechnung des Fehlers benutzt wurde.

Energieerwartungswert und *strong coupling expansion*

In Abbildung 4.1 ist der Erwartungswert für die Energie in Abhängigkeit von β aufgetragen. Wie erwartet geht der Erwartungswert für den Grenzwert $\beta \rightarrow 0$ ebenfalls gegen 0 und nähert sich im Grenzwert β gegen ∞ dem Wert 1 an.

Wie in Anhang A.4 dargestellt, kann der Erwartungswert für die Energie außerdem für kleine Werte des Kopplungsparameters β störungstheoretisch berechnet werden. Vergrößert ist in Abbildung 4.1 der Abschnitt der x-Achse für kleine Werte in β dargestellt. Die numerischen Ergebnisse sind in rot aufgetragen. Daneben ist in blau ein linearer zweiparametrischer Fit an die Daten eingezeichnet: Es ergibt sich in Übereinstimmung mit A.4 die Abhängigkeit $-\langle E_W \rangle = 0.25\beta$ für kleine Werte von β .

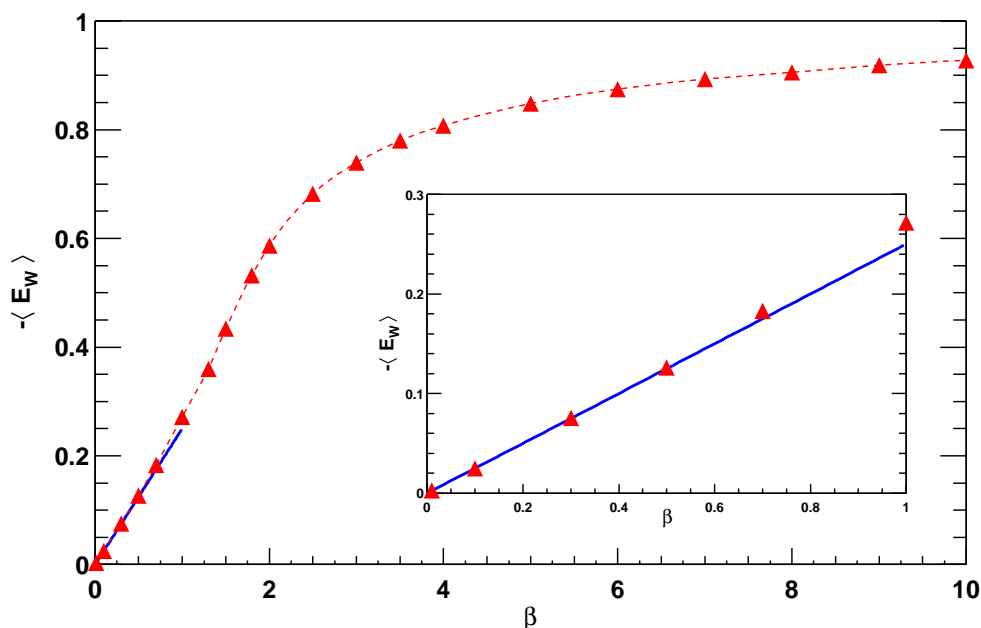


Abbildung 4.1: Energieerwartungswert E_W in Abhängigkeit von β und die *Strong coupling expansion* für den Energieerwartungswert: Es ist $-\langle E_W \rangle(\beta)$ und außerdem die Fitfunktion $0.254(1)\beta + 0.0003(3)$ aufgetragen. Außerdem ist ein vergrößerter Ausschnitt abgebildet.

Die Übereinstimmung zwischen *strong coupling expansion* und gemessenen Werten ist für ein großes Intervall in β sehr gut. Dies ist nicht unbedingt zu erwarten: Beispielsweise ist in der QCD die Übereinstimmung zwischen *strong coupling expansion* und gemessenen Daten nur für ein wesentlich kleineres Intervall sehr gut.

Insgesamt bestätigen die im obigen Abschnitt präsentierten Resultate, daß das verwendete Programm insofern korrekt ist, als die Überprüfung möglich und sinnvoll ist.

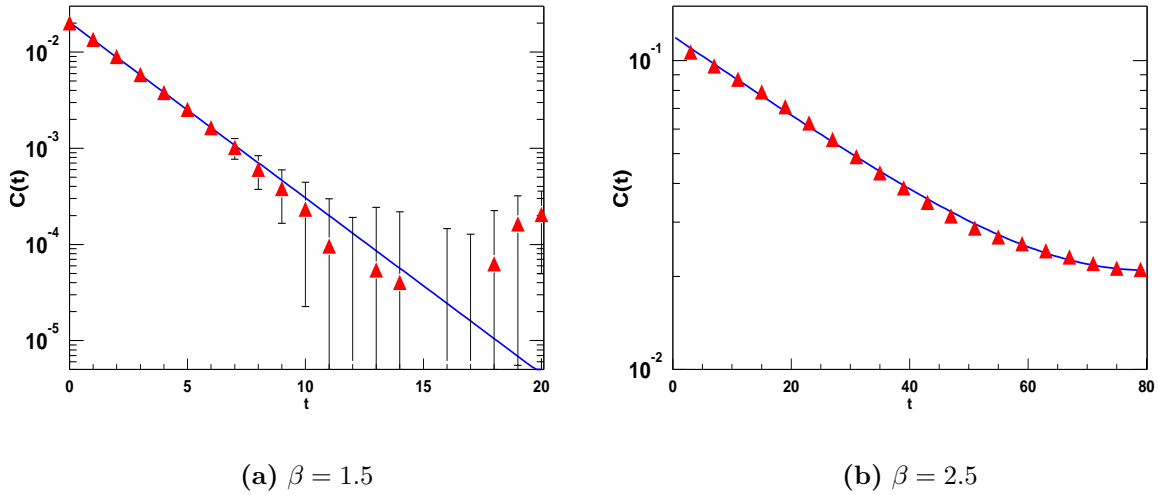


Abbildung 4.2: Erwartungswert der Zwei-Punkt-Funktion $C(t)$ (\blacktriangle) für $L = 160$ einfach logarithmisch aufgetragen. Die blaue durchgezogene Linie stellt jeweils einen Fit von $h_{a,m}(t) = a \cosh(mt - 80)$ an die Daten dar.

4.2 Bestimmung des *massgap*

Um das Skalierungsverhalten des Lyapunov-Exponenten studieren zu können, ist es nötig, den Exponenten für ein möglichst großes Intervall in der Kopplungskonstanten zu bestimmen. Dafür muß die Gitter-Größe so gewählt werden, daß die *finite size*-Effekte auch bei großen Werten für den Kopplungsparameter vernachlässigbar bleiben. Die nötige Gitter-Größe kann mit Hilfe der Korrelationslänge abgeschätzt werden.

Mit dem Ziel, die Korrelationslänge zu bestimmen, wurde zunächst der Erwartungswert der Zwei-Punkt-Funktion $\langle C(t) \rangle$ (4-11) für verschiedene t gemessen. In Abbildung 4.2 ist dieser Erwartungswert für ein Gitter mit $L = 160$ und $\beta = 1.5$ bzw. $\beta = 2.5$ einfach logarithmisch aufgetragen. Wegen der periodischen Randbedingungen geht die Darstellung in 4.2(b) nur bis $t = 80$; in Abbildung 4.2(a) sind die Daten sogar nur bis $t = 20$ aufgetragen, weil das Signal für $C(t)$ ab dort im Rauschen untergeht. Außerdem ist ein Fit von $h_{a,m}(t) = a \cosh(mt - 80)$ an die Daten eingezeichnet. Über einen solchen Fit für jedes β wurde jeweils das m in Abhängigkeit von β bestimmt.

Abbildung 4.2 kann auch ohne die entsprechenden Fits entnommen werden, daß m für $\beta = 1.5$ deutlich größer ist, als für $\beta = 2.5$. Außerdem demonstriert Abbildung 4.2(a) eine Schwierigkeit, welche bei der Bestimmung von m bei kleinen Kopplungsparametern auftritt: Da der Wert der Zwei-Punkt-Funktion sehr schnell auf ein Niveau fällt, das nur noch um Null fluktuiert (Rauschen), bleiben effektiv nur wenige Punkte zum Fit. Damit ist die Bestimmung in diesem Parameterbereich entsprechend ungenau, auch wenn der

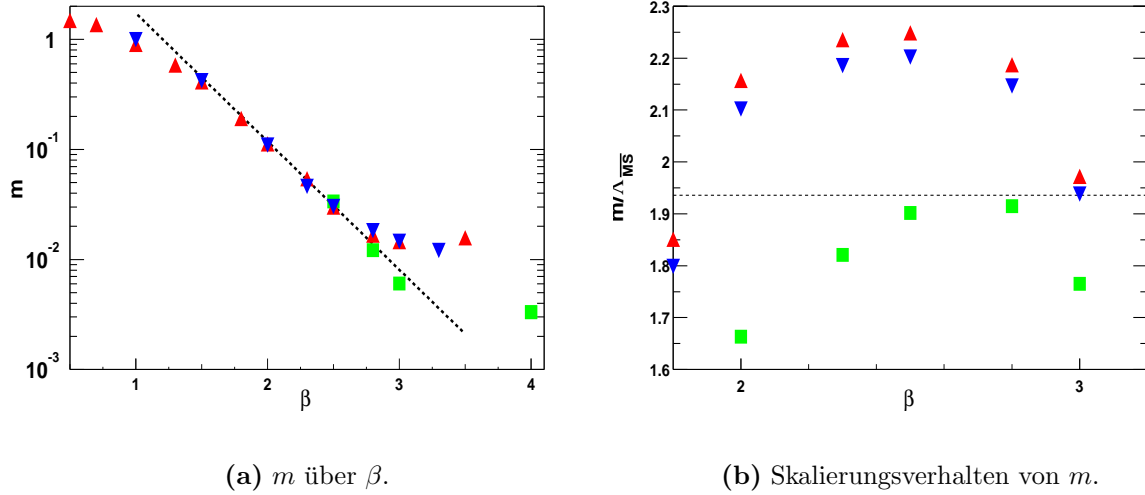


Abbildung 4.3: ←LINKS: *massgap* über dem Kopplungsparameter β für die Gitter-Größen 100^2 (\blacktriangle), 160^2 (\blacktriangledown) und 500^2 (\blacksquare). Die gestrichelte schwarze Linie stellt einen exponentiellen Fit dar. →RECHTS: Verhältnis von m und $\Lambda_{\overline{MS}}$ in Abhängigkeit von β , $\Lambda_{\overline{MS}}$ mit Λ_L (4-1) (\blacktriangle) und mit Λ (4-2) (\blacktriangledown) und Λ_E (4-7) (\blacksquare) berechnet. Die gestrichelte schwarze Linie ist die analytische Voraussage (4-8) für das Verhältnis.

Fehler, der beim Fit berechnet wird, dies nicht reflektiert.

Der Fit der cosh-Funktion an die Daten wurde jeweils nicht über den gesamten verfügbaren Bereich in t angewendet, sondern es wurden am Anfang entsprechend (2-31) einige Punkte ausgelassen, um zu vermeiden, Zustände einzubeziehen, die höherer Ordnung sind (siehe Anhang A.1).

m ist in Abbildung 4.3(a) in Abhängigkeit vom Kopplungsparameter β einfach logarithmisch für ein 100^2 , ein 160^2 und ein 500^2 Gitter abgebildet. Neben den gemessenen Werten ist ein exponentieller Fit eingezeichnet.

Der Fit in 4.3(a) läßt erkennen, daß m exponentiell mit steigendem β abnimmt. Der Fit soll nur als Hilfe dienen, die Proportionalität zur β -Funktion setzt erst für größere Werte von β ein. Das exponentielle Verhalten geht aber nicht über die ganze β -Skala für jede Gitter-Größe: Etwa bei $\beta = 2.5$ bricht es zum Beispiel für das 100^2 Gitter ab. Das liegt an der Tatsache, daß auch für die Bestimmung von m die Gitterlänge L des Systems groß gegenüber der Korrelationslänge sein muß.

Bei $\beta = 2.5$ ist also die Korrelationslänge $1/m$ so groß geworden, daß die Bestimmung von m auf einem 100^2 Gitter nicht mehr ohne Einflüsse von *finite size*-Effekten möglich ist. Deswegen muß man für weitergehende Untersuchungen auf größere Gitter zurückgreifen, was zur Folge hat, daß aufgrund des exponentiellen Anstiegs der Korrelationslänge zur Vermeidung von *finite size*-Effekten auch die Gitter-Größe L exponentiell vergrößert

werden muß. Damit steigt aber auch der Rechenaufwand exponentiell, der benötigt wird, um Ergebnisse für größere Kopplungsparameter zu erhalten.

Das exponentielle Verhalten von m für große Werte von β stammt bereits von der später einsetzenden Proportionalität von m mit der β -Funktion (1-32) her. Bei jener überwiegt für große β der exponentielle Abfall, für kleine Werte von β aber müssen höhere Ordnungen im Kopplungsparameter mit in die Näherung einbezogen werden. Dies erklärt, warum m mit abnehmenden Werten von β vom exponentiellen Verhalten abweicht, wie in Abbildung 4.3(a) ersichtlich wird.

In Abbildung 4.3 kann darüber hinaus abgelesen werden, daß $1/m$ bei $\beta = 3$ etwa 100 ist. Unter der Annahme, daß die Volumina groß genug sind, so lange $mL > 5$ ist, muß L bei dieser Kopplung also bereits 500 betragen, um Resultate zu erhalten, die frei von *finite size*-Effekten sind.

Eine weitere Feststellung läßt sich beim Vergleich zu den Ergebnissen aus [23] machen: Der HMC-Algorithmus für das nichtlineare σ -Modell ist nicht annähernd so effektiv wie der *Cluster*-Algorithmus. Das liegt erstens daran, daß die Autokorrelationszeiten sehr groß sind, und zweitens, daß das Phänomen des *critical slowing down*¹ auftritt: Mit steigendem β nimmt nicht nur die Korrelationslänge exponentiell zu, sondern auch die Autokorrelationszeit.

Damit muß für den HMC-Algorithmus bei steigendem Kopplungsparameter nicht nur das Gitter vergrößert werden, sondern auch die Anzahl der erzeugten Konfigurationen pro unabhängige Konfiguration steigt stark an - mit entsprechenden Auswirkungen auf die Rechenzeit.

4.2.1 Asymptotisches Skalierungsverhalten

Die zu Anfang des nun folgenden Unterabschnittes dargestellten Überlegungen folgen [23]. Wie bereits in Abschnitt 1.4.3 dargestellt, ist das nichtlineare $O(4)$ - σ -Modell asymptotisch frei. Das bedeutet, daß das *massgap* für große Werte in β proportional zu Λ_L wird. Λ_L läßt sich aus den universellen 1- und 2-*loop*-Koeffizienten der β -Funktion entsprechend Gleichung (1-32) (Seite 11), konstruieren:

$$\Lambda_L(\beta) = \exp(-2\pi\beta)(\pi\beta)^{1/2}. \quad (4-1)$$

Da für das σ -Modell auch der 3-*loop*-Term der β -Funktion bekannt ist [13], kann man anstelle von Λ_L die genauere Skala Λ benutzen, die entsprechend Gleichung (1-33) folgendes Aussehen hat:

$$\Lambda(\beta) = \Lambda_L(\beta) \left(1 + \frac{0.664}{4\pi\beta} + O(\beta^{-2}) \right). \quad (4-2)$$

¹ siehe dazu [24].

Meistens werden die Ergebnisse im \overline{MS} -Schema² angegeben; $\Lambda_{\overline{MS}}$ berechnet sich wie folgt aus Λ_L :

$$\frac{\Lambda_{\overline{MS}}}{\Lambda_L} = \sqrt{32} \exp(\pi/4). \quad (4-3)$$

Eine weitere Möglichkeit, das Skalierungsverhalten zu untersuchen, ist mit Hilfe einer Resummierungstechnik entsprechend der Idee in [25]. Ausgangspunkt ist eine störungstheoretische Entwicklung für die Energie E_W :

$$E_W = 1 - \frac{3}{4\beta} - \frac{3}{32\beta^2} + O(\beta^{-3}). \quad (4-4)$$

Benutzt man die ersten zwei Terme aus (4-4), so kann eine neue Kopplung β_E definiert werden:

$$\beta_E = \frac{3}{4(1 - E_W)}. \quad (4-5)$$

β_E geht für $\beta \rightarrow \infty$ in β über. Sie kann zur Definition von Λ_E benutzt werden, wenn β_E alternativ zu β als nackte Kopplung zur Berechnung von Λ_L in (4-1) herangezogen wird: $\Lambda_E = \Lambda_L(\beta_E)$. Mit (4-4) kann $g_E = 1/\beta_E$ in $g = 1/\beta$ entwickelt werden:

$$g_E = g \left(1 + \frac{1}{8}g + O(g^2) \right). \quad (4-6)$$

Daraus folgt mit Gleichung (1-34) das Verhältnis der beiden Λ -Parameter:

$$\frac{\Lambda_E}{\Lambda_L} = e^{\pi/8}. \quad (4-7)$$

Das exakte Verhältnis von $m/\Lambda_{\overline{MS}}$ wurde von P. Hasenfratz, M. Maggiore und F. Niedermayer in [26] bzw. in [27] analytisch bestimmt:

$$\frac{m}{\Lambda_{\overline{MS}}} = \left(\frac{8}{e} \right)^{1/2} \frac{1}{\Gamma(1.5)}. \quad (4-8)$$

In Abbildung 4.3(b) ist das Verhältnis von m zu $\Lambda_{\overline{MS}}$ über der Korrelationslänge abgebildet. Dabei wurde für die roten Dreiecke (\blacktriangle) Λ_L nach (4-1) zur Berechnung von $\Lambda_{\overline{MS}}$ verwendet. Für die blauen Dreiecke (\blacktriangledown) dagegen Λ entsprechend Gleichung 4-2 und für die grünen Quadrate (\blacksquare) Λ_E aus Gleichung 4-7 mit E_W aus den Daten.

Die Ergebnisse in Abbildung 4.3(b) sind vergleichbar mit den Ergebnissen, die in [23] veröffentlicht sind. Es ist zu erkennen, daß mit Λ das Skalierungsverhalten besser dargestellt wird, als mit Λ_L , und genauso ergeben sich bessere Ergebnisse mit Λ_E als mit Λ . Die gute Übereinstimmung von den mit Λ_L und Λ erzeugten Werten bei $\beta = 3$ ist wahrscheinlich Zufall. Das schlechtere Ergebnis für Λ_E bei $\beta = 3$ hingegen ist der Tatsache zuzuschreiben, daß die Bestimmung des Energieerwartungswertes in diesem Fall nicht

² siehe [9].

genau genug vorgenommen werden konnte. Für die Resultate bei kleineren Werten für β konnten die sehr genauen Werte aus [23] benutzt werden.

Ferner wird ersichtlich, daß der Bereich des asymptotischen Skalierungsverhaltens noch nicht ganz erreicht wird. Allerdings ist erstens in Abbildung 4.3(a) deutlich das exponentielle Verhalten von m zu erkennen und zweitens in Abbildung 4.3(b), daß immerhin eine Annäherung an den analytisch berechneten Wert (4-8) stattfindet. Allgemein wird der Beginn des asymptotischen Skalierungsverhalten bei etwa $\beta = 3$ angegeben [23].

Das im nächsten Abschnitt zu betrachtende Skalierungsverhalten des Lyapunov-Exponenten wird auf den Ergebnissen dieses Abschnittes aufbauend ebenfalls über das Verhältnis zu $\Lambda_{\overline{MS}}$ untersucht werden.

4.3 Lyapunov-Exponent

In diesem Abschnitt wird das Skalierungsverhalten des Lyapunov-Exponenten für große Werte von β untersucht. Dies geschieht zunächst für das nichtlineare $O(4)$ - σ -Modell und anschließend für den anharmonischen Oszillator.

Anlaß für diese Untersuchung ist folgende Hypothese, die von R. G. Edwards, I. Horváth und A. D. Kennedy in [3] formuliert wurde: Falls das chaotische Verhalten nicht nur eine Eigenschaft der dem HMC-Algorithmus zugrunde liegenden (fiktiven) Bewegungsgleichungen ist, sondern auch eine Eigenschaft der zugrunde liegenden Kontinuums-Feldtheorie, dann wäre der Lyapunov-Exponent ν gemessen in „physikalischen“ Einheiten konstant, also ν/m konstant im Grenzwert $\beta \rightarrow \infty$.

Dies hätte nicht nur zur Folge, daß es eine weitere charakteristische physikalische Größe der Feldtheorie gäbe, deren Bedeutung und Interpretation zu klären wäre. Es würde auch bedeuten, daß es erstens möglich wäre, die Reversibilitätsverletzung in Abhängigkeit von β vorherzusagen und somit zu kontrollieren, ohne direkt zu messen. Und zweitens, daß die Reversibilitätsverletzung im Grenzwert $\beta \rightarrow \infty$ exponentiell auf Null abfiele und damit den Kontinuumslimit nur noch unwesentlich beeinträchtigen würde.

Die Hypothese wird in [3] mit Daten belegt, die - nach Aussage der Autoren selbst - nicht ausreichen, um endgültige Sicherheit zu erhalten, weil große *finite size*-Effekte nicht ausgeschlossen werden konnten und deswegen nur ein kleines Intervall in β für die Überprüfung zur Verfügung stand.

4.3.1 Nichtlineares σ -Modell

Zunächst wird der Lyapunov-Exponent in Abhängigkeit von der Kopplungskonstanten gemessen, um dessen Skalierungsverhalten bestimmen zu können. Dazu kann die Reversibilitätsverletzung direkt gemessen werden, indem die Integration umgekehrt wird: Be-

ginnend mit einer thermalisierten Konfiguration U und dazu konjugierten Impulsen P , werden die Bewegungsgleichungen über die Trajektorienlänge τ integriert. Anschließend werden durch Vorzeichenwechsel in der Zeit τ die Bewegungsgleichungen wiederum über die Trajektorienlänge τ integriert, diesmal rückwärts in der Zeit bis zur Konfiguration U' . Wäre der Integrationsalgorithmus exakt reversibel, so wäre $U = U'$, was - wie bereits erläutert - nicht der Fall ist.

In Abhängigkeit von τ kann daraus die Reversibilitätsverletzung für die Felder über die Differenz der Norm bestimmt werden:

$$\Delta\delta U(\tau) = \|U'(\tau) - U(\tau)\|, \quad \|U\| = \frac{1}{V} \sum_x \|U_x\|. \quad (4-9)$$

Die oben beschriebene Prozedur muß mit möglichst vielen unabhängigen Anfangskonfigurationen wiederholt werden, um den statistischen Fehler abschätzen zu können. Dies stellt sich als sehr rechenzeitaufwendig heraus, da bis zu Trajektorienlängen von 10 integriert werden muß, um den Exponenten gut bestimmen zu können.

Um die Rechenzeit zu verkürzen, kann ausgenutzt werden, daß nicht für jeden Wert von τ neu hin- und zurückintegriert werden muß: Vielmehr wird einmal die Trajektorienlänge $\tau = 10$ hin- und zurückintegriert. Dabei wird in Schritten von 0.5 die jeweilige Konfiguration sowohl auf dem Hin- als auch auf dem Rückweg zwischengespeichert. Dies erlaubt eine Auswertung für $\Delta\delta U$ für alle Zwischen-Werte, wenn die Konfiguration nach Abschluß der Hin-Integration als die für $\tau = 0$ angenommen wird.

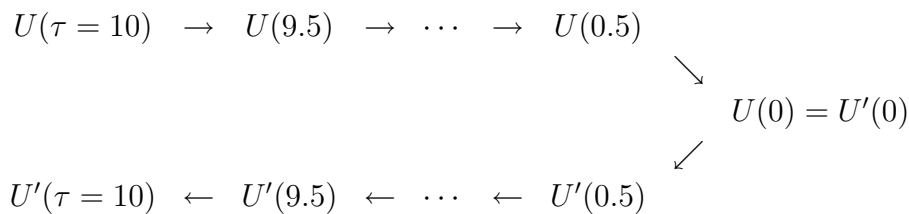


Abbildung 4.4: Zur Bestimmung von $\Delta\delta U(\tau)$

Wie in [Abbildung 4.4](#) dargestellt, beginnt man also praktisch mit der Konfiguration $U(\tau = 10)$, geht dann über alle Zwischenstationen zu $U(\tau = 0) = U'(\tau = 0)$ und integriert anschließend zu $U'(\tau = 10)$ zurück, wiederum über alle Zwischenstationen.

Die Bestimmung des Lyapunov-Exponenten ν geschieht dann über einen Fit einer zweiparametrischen Exponentialfunktion:

$$f_{\nu,b}(\tau) = b \exp(\nu\tau) \quad (4-10)$$

an die Meßwerte für $\Delta\delta U(\tau)$.

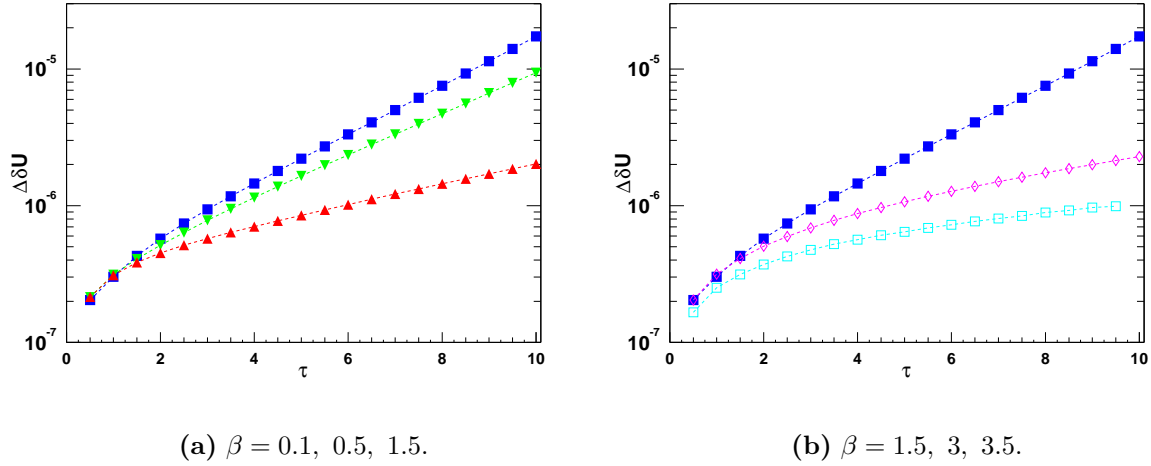


Abbildung 4.5: Die Verletzung der Reversibilität $\Delta\delta U$ über der Integrationslänge τ einfach logarithmisch aufgetragen für ein 160^2 Gitter bei verschiedenen Kopplungskonstanten $\beta = 0.1$ (\blacktriangle), $\beta = 0.5$ (\blacktriangledown), $\beta = 1.5$ (\blacksquare), $\beta = 3$ (\blacklozenge) und $\beta = 3.5$ (\square).

In [Abbildung 4.5](#) ist $\Delta\delta U(\tau)$ für ein 160^2 Gitter bei verschiedenen Kopplungsparametern einfach logarithmisch dargestellt. Es ist zu erkennen, daß $\Delta\delta U$ sich wirklich exponentiell entwickelt, und zwar mit unterschiedlichen Exponenten für verschiedene Kopplungsparameter. Das exponentielle Verhalten setzt allerdings nicht sofort bei $\tau = 0$ ein, sondern beginnt erst später. Dies ist konsistent mit den Beobachtungen, die in [\[3\]](#) gemacht werden. C. Liu, A. Jaster und K. Jansen zeigen in [\[20\]](#), daß es sich für sehr kleine Trajektorienlängen um eine Art Random-Walk-Verhalten handelt.

Es ist auch deutlich zu erkennen, daß die Steigung bis etwa $\beta = 1.5$ zunimmt und daraufhin wieder abnimmt.

Analoge Daten wurden in der vorliegenden Arbeit für verschiedene Gitter-Größen erstellt, um den Einfluß von *finite size*-Effekten ausschließen bzw. beobachten zu können.

In [Abbildung 4.6\(a\)](#) ist der Lyapunov-Exponent über dem Kopplungsparameter aufgetragen. Dabei wurden die Ergebnisse für unterschiedliche Gitter-Größen benutzt, was durch die verschiedenen Farben kenntlich gemacht ist. Durch Vergleich mit größeren Gittern ist zu erkennen, daß die Ergebnisse für ein 12^2 Gitter bei größeren β -Werten beginnen, *finite size*-Effekte zu zeigen, wie es zu erwarten ist. Die [Abbildung](#) legt aber auch nahe, daß die *finite size*-Effekte eher klein sind: Bis $\beta = 3$ sind die Werte für ν für alle abgebildeten Gitter-Größen nahezu gleich.

Zum allgemeinen Verlauf von ν ist [Abbildung 4.6\(a\)](#) zu entnehmen, daß der Exponent für β gegen 0 ebenfalls gegen 0 geht. Dies war zu erwarten, da sich die einzelnen Spins im Grenzfall $\beta = 0$ nicht mehr beeinflussen. D.h., daß das Gesamtsystem in viele einzelne Systeme zerfällt, die jeweils für sich integrierbar sind. Für integrierbare Systeme ist der

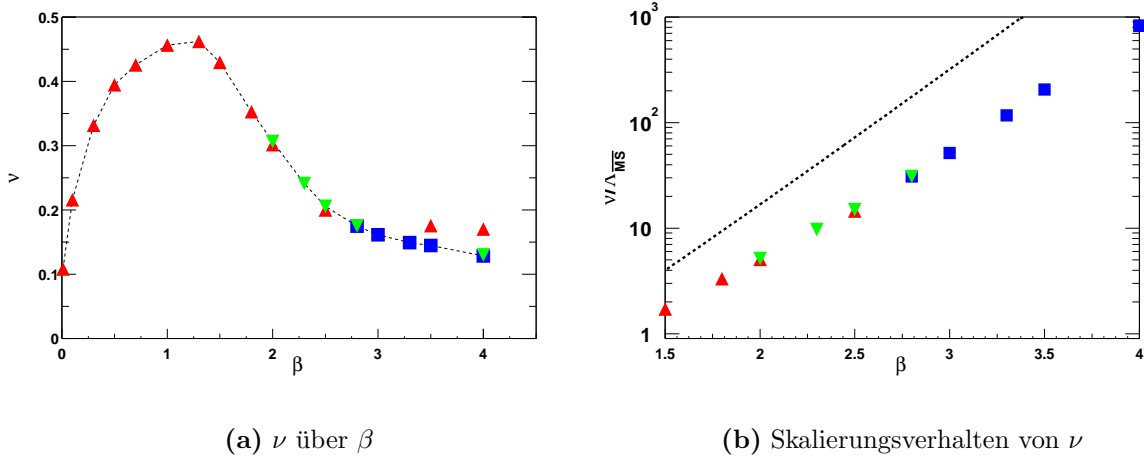


Abbildung 4.6: ◀LINKS: Der Lyapunov-Exponent ν über dem Kopplungsparameter β . Die gestrichelte Linie diene als Hilfe für das Auge. ▶RECHTS: Verhältnis von ν und $\Lambda_{\overline{MS}}$ in Abhängigkeit von β einfach logarithmisch. Zur Berechnung von $\Lambda_{\overline{MS}}$ wurde $\Lambda(\beta)$ (Gleichung 4-2) verwendet. Die gestrichelte Linie ist die Auftragung von $1/\Lambda_{\overline{MS}}(\beta)$. Gitter-Größen links wie rechts: 12^2 (\blacktriangle), 160^2 (\blacktriangledown), 500^2 (\blacksquare).

Exponent identisch 0.

Als nächstes ist zu bemerken, daß der Verlauf der Abhängigkeit des Exponenten von β und der Verlauf der entsprechenden Kurve in der QCD, wie er den Veröffentlichungen [3] und [20] zu entnehmen ist, übereinstimmt. Es liegt also nahe anzunehmen, daß dieses Verhalten des Lyapunov-Exponenten unabhängig von der betrachteten Feldtheorie ist. Dies wird im Abschnitt 4.3.2 genauer ausgeführt.

Können die vorliegenden Daten aber zur Verifikation oder Widerlegung der in [3] formulierten Hypothese, der Lyapunov-Exponent skalieren wie eine physikalische Größe, dienen? Selbst wenn die Daten nur bis zu $\beta = 3$ verlässlich sind, lassen sie nicht erwarten, daß ν exponentiell auf Null abfällt. Vielmehr scheint sich der Exponent für $\beta \rightarrow \infty$ einem Wert anzunähern, der von Null verschieden ist.

Das bedeutet für den Kontinuumslimit, daß auch für sehr große Werte von β Reversibilitätsverletzungen nicht ausgeschlossen werden können. Sollten also Auswirkungen auf Erwartungswerte nachgewiesen werden, die nicht vernachlässigbar sind, dann werden diese auch im Kontinuumslimit die Resultate beeinflussen.

Abbildung 4.6(b) kann ebenfalls in diesem Sinne interpretiert werden: Es ist das Verhältnis von ν und der universellen Skala $\Lambda_{\overline{MS}}$ über dem Kopplungsparameter einfach logarithmisch aufgetragen. Zur Berechnung von $\Lambda_{\overline{MS}}$ (4-3) wurde $\Lambda(\beta)$ (4-2) verwendet. Sollte der Lyapunov-Exponent wie eine physikalische Größe skalieren, so hieße das, daß er proportional zu m sein müßte. Damit müßte auch beim Verhältnis von ν zu $\Lambda_{\overline{MS}}$ eine Tendenz zu erkennen sein, sich einer Konstante anzunähern.

Dies wird durch die vorliegenden Daten nicht gestützt. Ganz im Gegenteil kann 4.6(b) entnommen werden, daß $\ln(\nu/\Lambda_{\overline{MS}})$ proportional zu $a\beta+b$ ist. Dabei ist als ein Hauptresultat der vorliegenden Arbeit festzuhalten, daß offensichtlich $a \neq 0$ ist.

Dies läßt einen der folgenden drei Schlüsse zu:

- Der Exponent skaliert nicht asymptotisch.
- Der Bereich in β ist noch nicht erreicht, in dem das asymptotische Skalierungsverhalten beginnt.
- Die Daten sind für $\beta \gtrsim 2.5$ nicht frei von Effekten des endlichen Volumens.

Vergleicht man die Abbildungen 4.3(b) und 4.6(b), so kommt man zu dem Schluß, daß der Exponent nicht asymptotisch skaliert. Es kann zwar nicht garantiert werden, daß die Daten völlig frei von *finite size*-Effekten sind, doch können jene nicht zu stark in das Ergebnis eingehen, da selbst für $\beta = 3$ und $\beta = 4$ die Ergebnisse für verschiedene Volumina in Abbildung 4.6(a) sehr dicht beieinander liegen.

Somit kann die Hypothese von Edwards et al. aufgrund der in der vorliegenden Arbeit präsentierten Ergebnisse nicht bestätigt werden.

4.3.2 Anharmonischer Oszillator

Für ein vom Modell unabhängiges Verhalten, und damit ebenfalls gegen die Hypothese des physikalischen Skalierens des Exponenten aus [3], sprechen die Daten, die im Rahmen dieser Arbeit für den anharmonischen Oszillator gewonnen wurden und im folgenden vorgestellt werden.

Für die Bestimmung des Lyapunov-Exponenten beim anharmonischen Oszillator werden die Bewegungsgleichungen (3-2) mit Hilfe von MATHEMATICA® und dessen Funktion `NDSolve` numerisch integriert. Die Anfangswerte für die Felder werden dabei zufällig gleichverteilt, die für die Impulse zufällig-gaussverteilt gewählt. Die Integration wird zweimal durchgeführt, wobei beim zweiten Mal die Anfangswerte der Felder nach der 8. Stelle gerundet werden. Die nicht-gerundeten Felder werden mit $[\varphi]$ bezeichnet, die zu gerundeten Anfangsbedingungen mit $[\varphi']$.

Mit `NDSolve` kann jeweils für $[\varphi]$ und $[\varphi']$ eine numerische Lösung in Abhängigkeit von τ bestimmt werden, aus der die Differenz D_ω in Abhängigkeit von τ berechnet wird:

$$D_\omega(\tau) = \sum_x |\varphi_x(\tau) - \varphi'_x(\tau)|. \quad (4-11)$$

Aus D_ω wird dann mittels einer zweiparametrischen Exponentialfunktion, die an die Daten gefittet wird, der Lyapunov-Exponent bestimmt. Dieses Vorgehen wird für viele Anfangswerte wiederholt, um den statistischen Fehler abschätzen zu können. Außerdem wird die Abhängigkeit von ω untersucht, wobei ω rein imaginär gewählt wird.

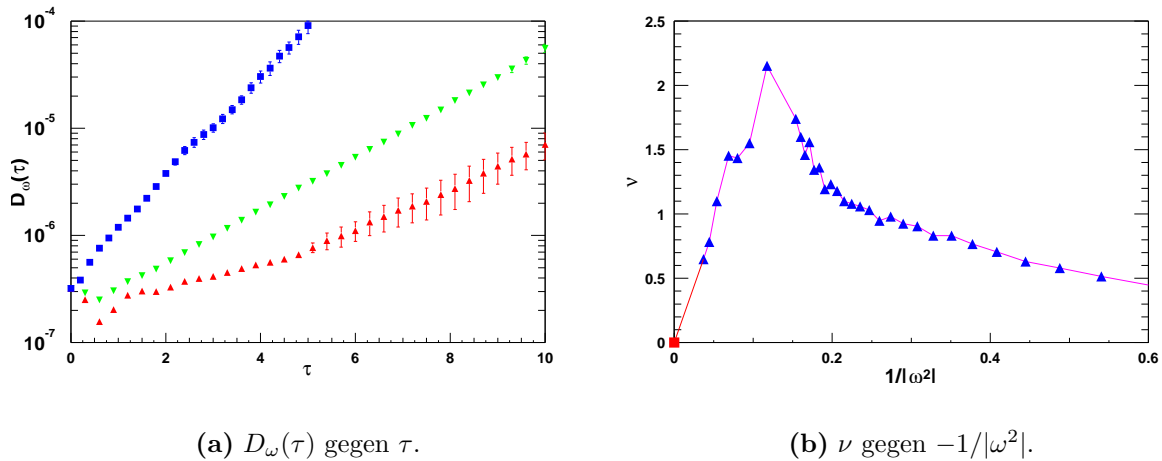


Abbildung 4.7: Der Lyapunov-Exponent für den anharmonischen Oszillator. ◀LINKS: Zur Bestimmung des Exponenten ist $D_\omega(\tau)$ gegen τ für $\omega^2 = -1$ (▲), $\omega^2 = -4$ (▼) und $\omega^2 = -16$ (■) aufgetragen. ▶RECHTS: ν gegen $1/|\omega^2|$ mit $\lambda = 1$ und $L = 128$ (▲). Die rote Linie diene als Hilfe für das Auge und ■ ist der theoretische Wert bei $1/|\omega^2| = 0$.

In Abbildung 4.7(a) ist $D_\omega(\tau)$ für drei unterschiedliche ω^2 ($\omega^2 = -1$ (▲), $\omega^2 = -4$ (▼) und $\omega^2 = -16$ (■)) einfach logarithmisch aufgetragen. Es ist zu erkennen, daß die Entwicklung exponentiell verläuft und außerdem mit verschiedenen Exponenten in Abhängigkeit von ω .

Der Lyapunov-Exponent über $-1/|\omega^2|$ ist in Abbildung 4.7(b) dargestellt. Die Auftragung erfolgt deswegen über $-1/|\omega^2|$, weil dann der Vergleich mit dem σ -Modell möglich ist: Aufgrund der in Abschnitt 1.3 vorgestellten zweiten Parametrisierung der Wirkung (1-22) ist $1/\omega^2 \propto \kappa$ für große $|\omega^2|$, wenn κ aus dem Intervall $[-1/\sqrt{8}, 0]$ gewählt wird. In diesem Fall entspricht der Grenzwert $-|\omega^2|$ gegen 0 dem Grenzwert κ gegen $-1/\sqrt{8}$.

Der Vergleich der beiden Abbildungen 4.7(b) und 4.6(a) ergibt eine große Ähnlichkeit für den Verlauf des Lyapunov-Exponenten im σ -Modell und beim anharmonischen Oszillator. Ihnen ist gemeinsam, daß sie im Grenzwert der verschwindenden Kopplung 0 sind, anschließend für kleine Kopplungsparameter ein Maximum haben, um dann für größere Kopplungsparameter langsam wieder zu fallen. Es kann nicht mit letzter Sicherheit beantwortet werden, ob sich der Exponent einem von 0 verschiedenen Wert annähert.

Das Verhalten im Grenzwert Kopplung $-1/|\omega^2|$ gegen 0 hängt davon ab, ob das durch die Bewegungsgleichungen beschriebene klassische Modell im Grenzwert integrabel ist. Für $-1/|\omega^2|$ nah bei 0 (also ω^2 nahe bei ∞) dominiert der quadratische Term in der Wirkung (1-20) so stark, daß alle weiteren Terme vernachlässigbar werden, so daß das Modell integrabel ist.

Die große Ähnlichkeit des Verhaltens des Exponenten zwischen den verschiedenen be-

trachteten Modellen, der QCD, des nichtlinearen σ -Modells und des anharmonischen Oszillators, welche sich sowohl in der Struktur der Wirkung als auch in der Dimension unterscheiden, deutet darauf hin, daß es sich beim Verhalten des Lyapunov-Exponenten um ein generisches Verhalten handelt, daß gleich für alle betrachteten Feldtheorien ist.

Die vorliegenden Daten lassen also den Schluß zu, daß das Verhalten des Lyapunov-Exponenten in Abhängigkeit von der Kopplung nicht durch die ursprüngliche Feldtheorie bestimmt, sondern eine generelle Eigenschaft des HMC-Algorithmus, oder genauer, der im Algorithmus verwendeten Bewegungsgleichungen ist.

Zusammenfassend kann also auf der Grundlage der in diesem Abschnitt 4.3 präsentierten Resultate für den Lyapunov-Exponenten ν folgendes festgehalten werden: Die Hypothese aus [3] kann durch die vorliegenden Daten für ν nicht gestützt werden. Durch Vergleich drei verschiedener Modelle, des nichtlinearen $O(n)$ - σ -Modells und des anharmonischen Oszillators mit Daten aus dieser Arbeit und der QCD mit Daten aus der Literatur, wird nahegelegt, daß das Verhalten des Exponenten eine allgemeine Eigenschaft des HMC-Algorithmus darstellt, die nicht vom simulierten Modell abhängt.

4.4 Auswirkung auf Observablen

Es bleibt noch die Frage offen, inwieweit der Fehler, der durch die Reversibilitätsverletzung entsteht, Auswirkungen auf die Berechnung von Erwartungswerten hat. Diese Frage und eventuell die Formulierung eines Kriteriums ist Inhalt des nächsten Abschnittes.

Um die Auswirkung der Reversibilitätsverletzung auf Erwartungswerte zu bestimmen, können Rechnungen verschiedener Genauigkeit verglichen werden. Es bietet sich dabei an, Rechnungen in einfacher (32-Bit) und in doppelter (64-Bit) Genauigkeit miteinander zu vergleichen.

Verschiedene Observablen sind von Interesse:

- Die Wirkung, weil sie das Verhalten des Systems bestimmt. Aus ihr ergibt sich zusätzlich und unmittelbar der Energieerwartungswert.
- Die Hamilton-Funktion, weil sie für die Dynamik in der Computerzeit verantwortlich ist und die Akzeptanzrate bestimmt.
- Der Erwartungswert von $\exp(-\Delta\mathcal{H})$, weil dieser für den exakten Algorithmus analytisch bekannt ist und somit vielleicht als Kriterium für die Exaktheit des Algorithmus herangezogen werden kann.
- Die Korrelationsfunktion, weil ihre Werte für größere Abstände exponentiell schnell gegen Null gehen und deswegen der Fehler durch die Reversibilitätsverletzung relativ zum Erwartungswert größer sein könnte, als beispielsweise bei der Wirkung.

Mit Ausnahme der Korrelationsfunktion wird die Berechnung folgendermaßen durchgeführt: Ausgehend von der gleichen thermalisierten Ausgangskonfiguration und dem gleichen Satz von Impulsen werden die Bewegungsgleichungen mit dem Integrationsverfahren aus dem HMC-Algorithmus integriert. Die Schrittweite $\Delta\tau$ wird so gewählt, daß die Akzeptanzrate im Fall einer Simulation groß genug wäre. β wurde zu 2 gewählt, weil dies der interessanten Stelle in der QCD entspricht, wenn man das Verhalten des Exponenten vergleicht.

Nach einer Trajektorienlänge τ wird die jeweilige Observable in beiden Genauigkeiten berechnet. Hier ist zu bemerken, daß sich die verschiedenen Genauigkeiten nur auf die Speicherung der Konfigurationen und der Impulse bezieht; die Erwartungswerte werden jeweils in der höheren Genauigkeit berechnet.

Anschließend wird die Differenz der Werte zwischen doppelter und einfacher Präzision berechnet. Nach entsprechender Wiederholung kann ein Histogramm der Differenzen erstellt und damit auch der Mittelwert der Abweichung berechnet werden.

Als Notation für die Differenz des Erwartungswertes einer Observablen \mathcal{O} in einfacher und in doppelter Genauigkeit wird $\Delta_{s,d}\mathcal{O}$ verwendet. Diese Untersuchung wurde auf vier

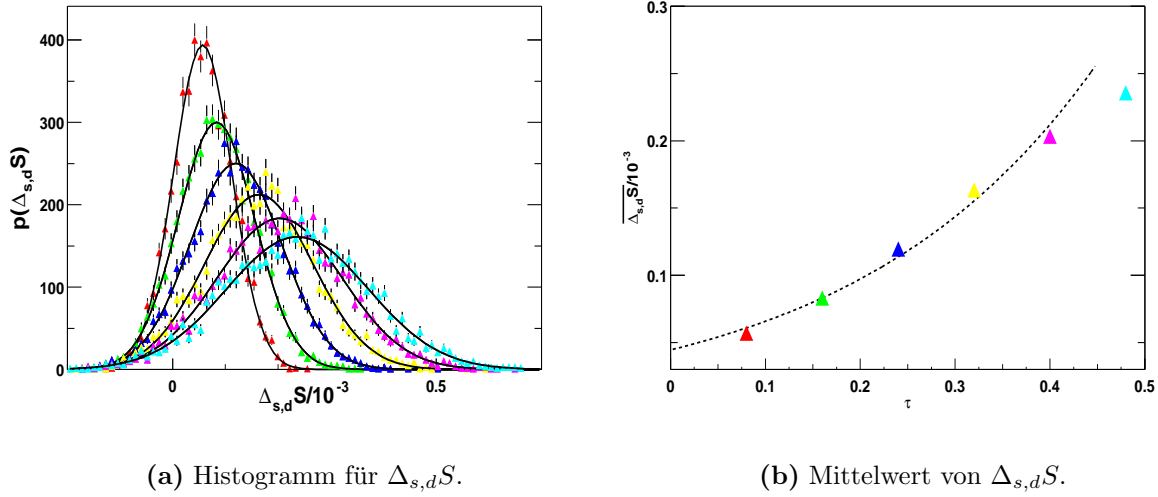


Abbildung 4.8: ◀LINKS: Histogramm für $\Delta_{s,d}S$ auf einem 160^2 Gitter für verschiedene Integrationslängen. Die durchgezogenen Linien sind jeweils Fits der Daten an eine Gaussfunktion. ▶RECHTS: Mittelwert der Verteilung aus der linken Abbildung aufgetragen über der Integrationslänge. Die gestrichelte Linie stellt einen zweiparametrischen exponentiellen Fit an die Daten dar. Die Farben entsprechen jeweils der Integrationslänge ($\tau = 0.08, 0.16, 0.24, 0.32, 0.40, 0.48$). Die Daten wurden bei $\beta = 2$ aufgenommen.

Gittern verschiedener Größe durchgeführt: Ein 160^2 -Gitter mit $\Delta\tau = 0.008$ und $N = 125$, ein 120^2 -Gitter mit $\Delta\tau = 0.01$ und $N = 100$, ein 80^2 -Gitter mit $\Delta\tau = 0.01$ und $N = 100$ und ein 40^2 -Gitter mit $\Delta\tau = 0.025$ und $N = 25$. Die drei größten Gitter wurden auf einem massiv parallelen Rechner bearbeitet, wohingegen das 40^2 Gitter auf einer Einzel-Prozessor-Maschine gerechnet wurde.

4.4.1 Wirkung

Zunächst sind in Abbildung 4.8(a) Histogramme für die Differenz in der Wirkung zwischen einfacher und doppelter Präzision nach den Integrationslängen $\tau = 0.08, 0.16, 0.24, 0.32, 0.40, 0.48$ aufgetragen. Verwendet wurde für diese Abbildung ein 160^2 Gitter. Die Histogramme wurden jeweils über 4001 Konfigurationen erstellt.

Die durchgezogenen schwarzen Linien stellen jeweils Gauss-Funktionen dar:

$$f_{\alpha, \bar{x}, \sigma_x}(x) = \alpha \exp - \left[\frac{1}{2} \frac{(x - \bar{x})^2}{\sigma_x} \right], \quad (4-12)$$

die an die Daten gefittet wurden. Dabei ist x in diesem Fall durch $\Delta_{s,d}S$ zu ersetzen. \bar{x} ist der Mittelwert, σ_x die Breite und α die Höhe der Gaussverteilung. In Abbildung

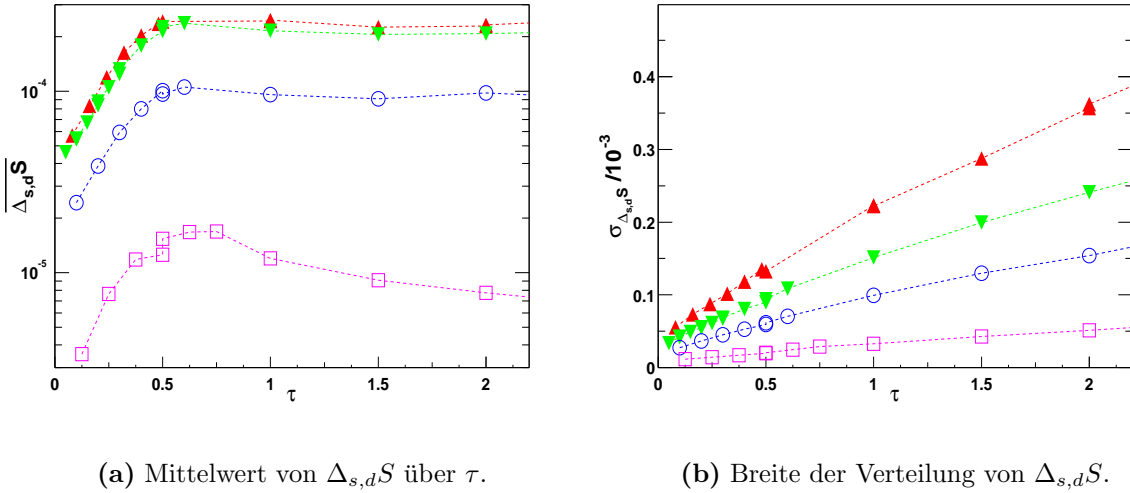


Abbildung 4.9: ◀LINKS: Die Entwicklung des Mittelwertes der Verteilung von $\Delta_{s,d}S$ über der Integrationslänge τ einfach logarithmisch aufgetragen. ▶RECHTS: Entwicklung der Breite der Verteilung von $\Delta_{s,d}S$ über der Integrationslänge. Aufgetragen wurden die Werte jeweils für ein 160^2 (\blacktriangle), ein 120^2 (\blacktriangledown), ein 80^2 (\circ) und ein 40^2 (\square) Gitter für $\beta = 2$.

4.8(b) ist jeweils der Mittelwert des Histogramms aus 4.8(a) über der Trajektorienlänge aufgetragen. Die Farben entsprechen links und rechts der Trajektorienlänge.

Es ist zu erkennen, daß der Mittelwert der Histogramme von 0 verschieden ist und mit zunehmender Trajektorienlänge größer wird. Abbildung 4.8(b) ist am exponentiellen Fit zu entnehmen, daß das Wachstum des Mittelwertes exponentiell ist.

Dieses Ergebnis läßt sich für andere Gittergrößen reproduzieren: In Abbildung 4.9(a) ist die Entwicklung des Mittelwertes von $\Delta_{s,d}S$ für verschiedene Gittergrößen in der Trajektorienlänge einfach logarithmisch aufgetragen. Ein weiteres Merkmal wird erkennbar: Im Gegensatz zur Breite der Histogramme, die fortlaufend zunimmt, sättigt der Mittelwert der Verteilung bei einer Konstanten, die von der Gittergröße abhängt.

Der Effekt der Sättigung ist damit zu erklären, daß es sich bei der Gruppe $SU(2, \mathbb{C})$ um eine kompakte Gruppe handelt und damit der Abstand zweier Elemente nach oben beschränkt ist.

Zusammenfassend sind vier Merkmale der Histogramme von Bedeutung:

1. Sie bilden jeweils eine Gauss-Verteilung, wie an den in Abbildung 4.8(a) eingetragenen Fits zu erkennen ist.
2. Der Mittelwert der Histogramme ist nicht Null: Der Mittelwert wächst mit steigender Trajektorienlänge zunächst exponentiell an, wie am exponentiellen Fit in

KAPITEL 4. ERGEBNISSE UND DISKUSSION

Abbildung 4.8(b) und in Abbildung 4.9(a) zu erkennen ist. Die Abweichung des Mittelwertes von Null sättigt nach einer Trajektorienlänge von etwa 0.7.

3. Die Breite der Verteilung nimmt mit wachsender Trajektorienlänge linear zu, wie in Abbildung 4.9(b) zu sehen ist.
4. Während der allgemeine Verlauf der Kurven für alle betrachteten Gittergrößen gleich ist, hängt beim Mittelwert die absolute Größe und bei der Breite die Steigung von der Gittergröße und damit auch von den gewählten Parametern ab.

Beim zweiten Punkt ist ebenfalls bemerkenswert, daß die exponentielle Entwicklung für Werte von τ kleiner als etwa 0.7 auf allen Gittern mit nahezu der gleichen Steigung geschieht.

Der Erwartungswert der Wirkung enthält das Produkt benachbarter Felder, die die fehlerbehafteten Größen sind. Es stellt sich die Frage, wie Fehler in den Feldern sich auf die Wirkung auswirken. Um den Fehler in der Wirkung abzuschätzen, kann wie folgt vorgegangen werden: Man betrachte zwei benachbarte Felder $\phi_x(t)$ und $\phi_y(t)$: Angenommen, die zu den Feldern konjugierten Impulse $P_x(t)$ und $P_y(t)$ enthalten zur Zeit $t = 0$ keine Fehler, die Felder jeweils den Fehler $\delta\phi_\alpha$:

$$\begin{aligned} P_\alpha(\Delta\tau/2) &= P_\alpha(0) + 0.5\Delta\tau\phi_\gamma(0) , \\ \phi_\alpha(0) &= \phi_\alpha + \delta\phi_\alpha , \end{aligned} \quad (4-13)$$

wobei $\alpha, \gamma \in \{x, y\}$, $\alpha \neq \gamma$ ist. Wird ferner die lineare Näherung der Bewegungsgleichungen benutzt:

$$\begin{aligned} P_\alpha((k+1/2)\Delta\tau) &= P_\alpha((k-1/2)\Delta\tau) + \Delta\tau\phi_\gamma((k-1)\Delta\tau) , \\ \phi_\alpha(k\Delta\tau) &= \phi_\alpha((k-1)\Delta\tau) - \Delta\tau P_\alpha((k+1/2)\Delta\tau) , \end{aligned} \quad (4-14)$$

so ist es möglich, die Fehler \mathfrak{F} im Produkt der beiden Felder $\phi_x(k)\phi_y(k)$ in Abhängigkeit von k zu betrachten: Dabei sind die Terme von Interesse, die proportional zu $\delta\phi_x\delta\phi_y$ und $[\delta\phi_\alpha]^2$ sind, weil die Terme linear in den Fehlern verschwinden, da der Mittelwert der Fehler selbst 0 ist. Die Form des Terms proportional zu $\delta\phi_x\delta\phi_y$ im Produkt $\phi_x(k)\phi_y(k)$ wurde mit Hilfe von MATHEMATICA® zu:

$$\mathfrak{F}_{xy}(k) = \delta\phi_x\delta\phi_y \left(1 + \sum_{i=1}^k \mathfrak{C}_{i,k}\Delta\tau^{4i} \right) \quad (4-15)$$

bestimmt, mit positiven reellen Koeffizienten $\mathfrak{C}_{i,k}$. Der führende Term ist in diesem Fall also $1 + \mathfrak{C}_{1,k}\Delta\tau^4$. Der Term proportional zu $[\delta\phi_\alpha]^2$ hat dagegen folgende Form:

$$\mathfrak{F}_{\alpha\alpha} = -[\delta\phi_\alpha]^2 \sum_{i=1}^k \mathfrak{D}_{i,k}\Delta\tau^{4i-2}, \quad (4-16)$$

| k | $\mathfrak{C}_{1,k}$ | $\mathfrak{D}_{1,k}$ |
|-----|----------------------|----------------------|
| 1 | 0.25 | 0.5 |
| 2 | 5.00 | 2.0 |
| 3 | 26.25 | 4.5 |
| 4 | 84.00 | 8.0 |
| 5 | 206.25 | 12.5 |
| 6 | 429.00 | 18.0 |
| 7 | 796.25 | 24.5 |
| 8 | 1360.00 | 32.0 |
| 9 | 2180.25 | 40.5 |
| 10 | 3325.00 | 50.0 |

Tabelle 4.3: Koeffizienten der Fehlerentwicklung für die Wirkung.

mit ebenfalls positiven reellen Koeffizienten $\mathfrak{D}_{i,k}$. Im Gegensatz zu \mathfrak{F}_{xy} ist der führende Term bei $\mathfrak{F}_{\alpha\alpha}$ also $\mathfrak{D}_{1,k}\Delta\tau^2$. Die Koeffizienten $\mathfrak{C}_{1,k}$ und $\mathfrak{D}_{1,k}$ wurden ebenfalls mit MATHEMATICA® bestimmt und sind Tabelle 4.3 zu entnehmen. Für die Koeffizienten $\mathfrak{D}_{1,k}$ kann aus Tabelle 4.3 folgende Formel abgelesen werden:

$$\mathfrak{D}_{1,k} = \sum_{i=1}^k (i - 0.5) = \frac{k^2}{2} .$$

Unter der Annahme, δ_{xx} und δ_{yy} sind von der gleichen Größenordnung, kann der gesamte Fehler wie folgt abgeschätzt werden:

$$\mathfrak{F}(k\Delta\tau) \propto \delta\phi_x\delta\phi_y - \delta_{xx}k^2\Delta\tau^2 + \delta\phi_x\delta\phi_y\mathfrak{C}_{1,k}\Delta\tau^4. \tag{4-17}$$

Es ist zu erkennen, daß diese einfache Näherung nicht in der Lage ist, das exponentielle Wachstum des Fehlers zu erklären oder zu quantifizieren. Das liegt wahrscheinlich vor allem daran, daß nur zwei Gitter-Punkte miteinbezogen wurden: Bereits bei $k \geq 2$ kommen zum Fehler Terme hinzu, die von Gitter-Punkten stammen, welche zum betrachteten einen Abstand von k haben. Alle diese Terme wurden bei obiger Abschätzung nicht miteinbezogen.

An Gleichung (4-17) können aber trotz dieser Einschränkungen einige Eigenschaften abgelesen werden: Das Fehlerquadrat an einem Ort $[\delta\phi_\alpha]^2$, von dem der Fehler einerseits abhängt, ist, wie in Abschnitt 2.2 dargestellt, im Mittel ungleich Null. Außerdem ist der Fehler von $\Delta\tau^2$, von der Schrittzahl k und offensichtlich auch vom Volumen des Gitters abhängig. Die Koeffizienten \mathfrak{D} erklären somit für den Fall von nur zwei benachbarten Feldern kein exponentielles Wachstum, es ist aber vorstellbar, daß unter Einbeziehung aller Gitter-Punkte ein solches entsteht. Und diese Näherung kann, wie gerade erläutert,

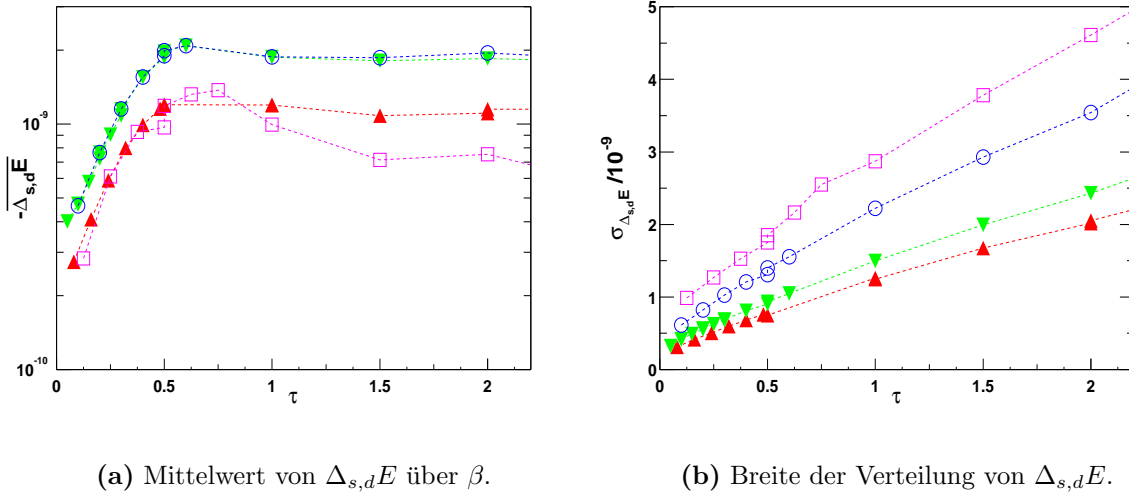


Abbildung 4.10: ◀LINKS: Die Entwicklung des Mittelwertes der Verteilung von $\Delta_{s,d}E$ über der Integrationslänge τ einfach logarithmisch aufgetragen. ▶RECHTS: Entwicklung der Breite von $\Delta_{s,d}E$ über der Integrationslänge. Aufgetragen wurden die Werte jeweils für ein 160^2 - (▲), ein 120^2 - (▼), ein 80^2 - (○) und ein 40^2 -Gitter (□) für $\beta = 2$.

erklären, daß der Fehler vom Volumen, von der Schrittweite sowie von der Anzahl der Schritte abhängt.

Dies ist bereits ohne die Annahme gültig, daß auch die Fehler benachbarter Felder korreliert sind und damit auch $\overline{\delta\phi_x\delta\phi_y} \neq 0$ ist. Bei hinreichend kleinem $\Delta\tau$ würde durch $\overline{\delta\phi_x\delta\phi_y} \neq 0$ in erster Näherung auch lediglich ein konstanter Term zum Fehler hinzukommen und könnte damit nicht für die Erklärung der „zeitlichen“ Entwicklung (in k) des Fehlermittelwertes herangezogen werden.

4.4.2 Erwartungswert der Energie

Wie bereits erwähnt, hängt der Energieerwartungswert sehr direkt vom Erwartungswert der Wirkung ab. Es handelt sich allerdings um eine intrinsische Größe, das heißt, der Erwartungswert ist weitgehend unabhängig von der betrachteten Gittergröße. Ansonsten sind keine großen Unterschiede zu den Ergebnissen bei der Wirkung zu erwarten.

In Abbildung 4.10(a) ist die Entwicklung des Mittelwertes der Differenzen des Energieerwartungswertes bei zunehmender Trajektorienlänge dargestellt. Für die Gittergrößen 120^2 und 80^2 bzw. 160^2 und 40^2 kommen die Kurven jeweils zur Deckung. Auch der Unterschied zwischen diesen beiden Teilen ist nur gering. Der Mittelwert für die Energiedifferenz, der sozusagen den Mittelwert der Wirkungsdifferenz geteilt durch das Volumen darstellt, ist also nahezu unabhängig von der betrachteten Gitter-Größe.

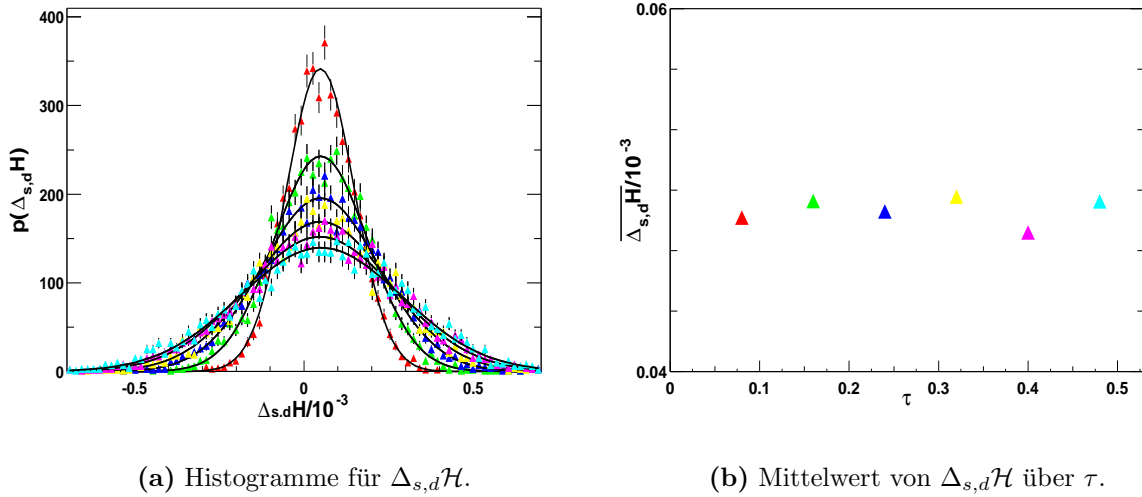


Abbildung 4.11: ◀LINKS: Histogramm für $\Delta_{s,d}\mathcal{H}$ auf einem 160^2 Gitter für verschiedene Integrationslängen. Die durchgezogenen Linien sind jeweils Fits an eine Gaussfunktion. ▶RECHTS: Mittelwert der Verteilung aus der linken Abbildung aufgetragen über der Integrationslänge. Die Farben korrespondieren jeweils mit der Integrationslänge ($\tau = 0.08, 0.16, 0.24, 0.32, 0.40, 0.48$).

Abbildung 4.10(a) legt nahe, daß neben der Gitter-Größe auch die Parameter $\Delta\tau$ und N eine Rolle für die Größe der Differenz spielen: Für das 80^2 - und das 120^2 -Gitter wurden die gleichen Parameter verwendet und die Kurven liegen in 4.10(a) nahezu perfekt übereinander, so daß in diesem Fall die Differenzen intrinsischer Größen vollständig unabhängig von der Gitter-Größe sind. Dies ist konsistent mit der Abschätzung des Fehlers in der Wirkung 4-17: Auch dort wird eine Abhängigkeit des Fehlers von $\Delta\tau$ und N vorhergesagt.

An dieser Stelle lohnt sich ein Vergleich der Größenordnungen von Energieerwartungswert und Mittelwert der Energieerwartungswertdifferenzen: Für $\beta = 2$ liegt der Energieerwartungswert bei etwa 0.6. Der Mittelwert der Differenzen liegt bei gleichem β und $\tau = 1$ bei ungefähr $2 \cdot 10^{-9}$. Es liegen also acht Größenordnungen zwischen dem Fehler durch die Rechnung in einfacher Präzision und dem Mittelwert. Die gleiche Beobachtung machen Liu, Jaster und Jansen bei ihrer Untersuchung zur Polyakov-Linie in [20].

Diesen Fehler neben dem statistischen Fehler nachzuweisen, ist praktisch unmöglich. Es kann an Hand dieser Ergebnisse davon ausgegangen werden, daß für Observablen dieser Art die Reversibilitätsverletzung die Ergebnisse nicht beeinflusst.

Für der Breite der Verteilung, die über der Trajektorienlänge in Abbildung 4.10(b) für verschiedene Gittergrößen aufgetragen ist, fallen die Kurven nicht übereinander. Für dieses Verhalten konnte ebenso wie für das lineare Wachstum der Breite keine Erklärung gefunden werden.

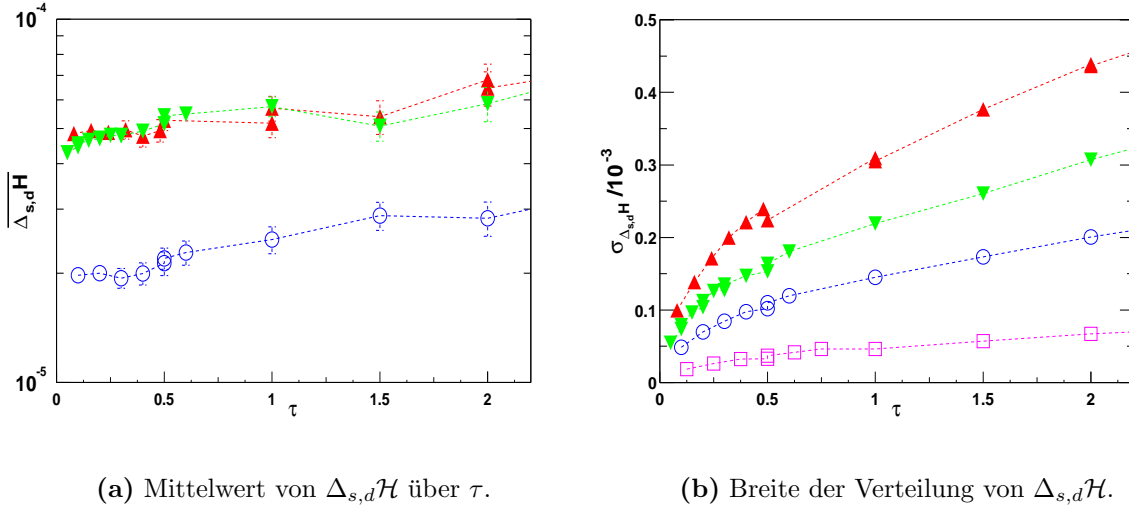


Abbildung 4.12: Einfach logarithmische Auftragung des Mittelwertes der Verteilung von $\Delta_{s,d}\mathcal{H}$ (←LINKS) und die Entwicklung der Breite der Verteilung von $\Delta_{s,d}\mathcal{H}$ (→RECHTS) für ein 160^2 - (▲), ein 120^2 - (▼), ein 80^2 - (○) und (nur rechts) ein 40^2 -Gitter (□).

4.4.3 Hamilton-Funktion

In Abbildung 4.11(a) sind entsprechend der Abbildung 4.8(a) Histogramme für die Differenz der Hamilton-Funktion $\Delta_{s,d}\mathcal{H}$ neben der Entwicklung des Mittelwertes $\overline{\Delta_{s,d}\mathcal{H}}$ über der Trajektorienlänge in 4.11(b) aufgetragen. Die Daten wurden wiederum mit einem 160^2 Gitter aufgenommen.

Im Gegensatz zu den Differenzen der Wirkung ist hier keine wesentliche Entwicklung des Mittelwertes zu erkennen, er bleibt konstant. Lediglich die Breite der Verteilung nimmt mit steigender Trajektorienlänge wie bei der Wirkung zu. Da die Hamilton-Funktion aus der Differenz von der Summe Quadrate der Impulse einerseits und der Wirkung andererseits besteht, heben sich also die Fehler in den Impulsen und die Fehler in den Feldern nahezu auf, wie in Abbildung 4.11(b) zu sehen ist: Es bleibt nur ein konstanter, aber relativ zum Fehler in der Wirkung sehr kleiner *bias* in der Hamilton-Funktion bestehen.

Das bedeutet, daß der kleine *bias* zwar einen Einfluß auf die Akzeptanzrate haben wird, dieser Einfluß aber immerhin nicht mit steigender Trajektorienlänge zunimmt.

Auch dieses Ergebnis läßt sich auf allen anderen untersuchten Gittern reproduzieren: In Abbildung 4.12(a) ist die Entwicklung des Mittelwertes von $\Delta_{s,d}\mathcal{H}$ über der Trajektorienlänge einfach logarithmisch für verschiedene Gittergrößen abgebildet, in 4.12(b) die Entwicklung der Breite. In Abbildung 4.12(a) wurde aus Gründen der Übersichtlichkeit das 40^2 -Gitter nicht dargestellt. Während der Mittelwert nahezu konstant bleibt, steigt die Breite wie bei der Wirkung linear an, was in Abbildung 4.12(b) zu sehen ist.

| L | Akzeptanzrate | | $\langle \exp(-\Delta\mathcal{H}) \rangle$ | |
|-----|---------------|---------------|--|---------------|
| | <i>single</i> | <i>double</i> | <i>single</i> | <i>double</i> |
| 160 | 98.48% | 98.20% | 1.00121(61) | 1.00142(61) |
| 120 | 98.23% | 97.90% | 1.00218(68) | 1.00237(68) |
| 80 | 98.85% | 98.85% | 1.00021(45) | 1.00024(45) |
| 40 | 98.75% | 98.60% | 1.0064(14) | 1.0069(13) |

Tabelle 4.4: Akzeptanzrate für verschiedene Gittergrößen bei $\tau = 1$ und $\langle \exp(-\Delta\mathcal{H}) \rangle$ jeweils gemittelt über 4001 Konfigurationen.

Die Modell-Hamilton-Funktion bestimmt das Akzeptanzverhalten des Algorithmus. Wie also wirkt sich der sehr kleine *bias* auf die Akzeptanz aus?

Es können nur minimale Unterschiede sowohl in der Akzeptanzrate als auch im Erwartungswert für $\langle \exp(-\Delta\mathcal{H}) \rangle$ festgestellt werden, wie man an Tabelle 4.4 erkennen kann.

Die in Tabelle 4.4 aufgeführten Zahlen lassen es ebenfalls schwierig erscheinen, die Abweichung des Erwartungswertes von $\langle \exp(-\Delta\mathcal{H}) \rangle$ vom analytisch bestimmten Wert 1 als ein Kriterium für die Exaktheit des Algorithmus heranzuziehen. Die Abweichungen von 1 sind klein und außerdem sind die Erwartungswerte in beiden Genauigkeiten innerhalb der statistischen Fehler gleich, so daß daraus nicht auf starke Verletzungen der Exaktheit des Algorithmus zu schließen ist.

4.4.4 Korrelationsfunktion

Für die bisherigen Vergleiche zwischen einfacher und doppelter Genauigkeit wurden stets nur Werte am Ende einer MD-Trajektorie herangezogen, wobei bei gleichen Anfangsbedingungen gestartet wurde. Dies stellt insofern kein realistisches Szenario dar, als in der Regel die Erzeugung der Konfigurationen entweder nur in einfacher (32-Bit) oder nur in doppelter (64-Bit) Genauigkeit durchgeführt wird.

Darüber hinaus ist es interessant, Observablen zu betrachten, deren Erwartungswert betragsmäßig klein ist, so daß unter Umständen der Reversibilitätsfehler relativ stärker zum Tragen kommt.

Um diesen Tatsachen Rechnung zu tragen, wurden zwei Simulationen - eine in einfacher, eine in doppelter Genauigkeit - durchgeführt, die lediglich zu Beginn mit den gleichen Ausgangsbedingungen gestartet wurden und alle Parameter gemeinsam hatten. Die Simulationen wurden bei $\beta = 2$ auf einem 120^2 -Gitter ausgeführt und dabei die Korrelationsfunktion bestimmt.

| | Einfache Genauigkeit | | |
|--|----------------------|-----------------------|--------------|
| | EW | Fehler | τ_{int} |
| $\langle \exp(-\Delta\mathcal{H}) \rangle$ | 1.00005 | $0.453 \cdot 10^{-6}$ | - |
| m | 0.125292 | $0.519 \cdot 10^{-3}$ | - |
| Rate | 98.203% | - | - |
| R[s] | $1.17 \cdot 10^6$ | - | - |
| Anzahl Konf. | 9821 | - | - |
| | EW/ 10^{-3} | Fehler/ 10^{-3} | τ_{int} |
| $C(0)$ | 52.95 | 0.363 | 0.497 |
| $C(5)$ | 27.78 | 0.257 | 0.396 |
| $C(10)$ | 14.89 | 0.214 | 0.395 |
| $C(15)$ | 7.966 | 0.1921 | 0.391 |
| $C(20)$ | 4.234 | 0.1783 | 0.444 |
| $C(25)$ | 2.357 | 0.1706 | 0.511 |
| $C(30)$ | 1.415 | 0.1730 | 0.576 |
| $C(35)$ | 0.8917 | 0.1646 | 0.636 |
| $C(40)$ | 0.5209 | 0.1622 | 0.586 |

Tabelle 4.5: Der Erwartungswert für $C(t)$ auf einem 120^2 -Gitter bei $\beta = 2$ wurde in einfacher Genauigkeit berechnet. Jeweils 1000000 Konfigurationen wurden erzeugt, jede 100. akzeptierte wurde für den Mittelwert der Korrelationsfunktion verwendet. Gerechnet wurde auf acht Prozessoren. Die Akzeptanzrate, die Gesamtrechnzeit R sowie der Erwartungswert von $\exp(-\Delta\mathcal{H})$ sind ebenfalls angegeben.

Gestartet wurde bei der gleichen thermalisierten Konfiguration auf jeweils acht Prozessoren des PC-Clusters am DESY Zeuthen. Anschließend wurden jeweils eine Millionen Konfigurationen erzeugt, von denen jeweils jede 100. akzeptierte für die Messung gespeichert wurde. Durch Mitteln über die gespeicherten Konfigurationen wurde die Korrelationsfunktion in Zeitrichtung bestimmt.

Die Ergebnisse für die Simulation in einfacher Genauigkeit sind in Tabelle 4.5, die für die Simulation in doppelter Genauigkeit in Tabelle 4.6 zusammengestellt.

Ein Vergleich der beiden Tabellen ergibt, daß die Akzeptanzrate sowie der Erwartungswert von $\exp(-\Delta\mathcal{H})$ in beiden Genauigkeiten praktisch identisch ist. Die Akzeptanzrate ist im Fall der einfachen Genauigkeit leicht höher als bei der doppelten Genauigkeit.

Die mit Hilfe der ebenfalls eingetragenen Autokorrelationszeit bestimmten Fehler (siehe Abschnitt 2.4, Gleichung 2-34) zeigen, daß ab $C(20)$ der Erwartungswert von $C(t)$ im statistischen Rauschen untergeht, so daß also maximal bis $C(15)$ ein Vergleich von einfacher und doppelter Genauigkeit sinnvoll ist.

4.4. AUSWIRKUNG AUF OBSERVABLEN

| | Doppelte Genauigkeit | | |
|--|----------------------|-----------------------|--------------|
| | EW | Fehler | τ_{int} |
| $\langle \exp(-\Delta\mathcal{H}) \rangle$ | 1.00006 | $0.453 \cdot 10^{-6}$ | - |
| m | 0.125751 | $0.509 \cdot 10^{-3}$ | - |
| Rate | 98.198% | - | - |
| R[s] | $1.60 \cdot 10^6$ | - | - |
| Anzahl Konf. | 9820 | - | - |
| | EW/ 10^{-3} | Fehler/ 10^{-3} | τ_{int} |
| $C(0)$ | 53.67 | 0.367 | 0.329 |
| $C(5)$ | 28.66 | 0.262 | 0.314 |
| $C(10)$ | 15.24 | 0.218 | 0.411 |
| $C(15)$ | 8.152 | 0.194 | 0.477 |
| $C(20)$ | 4.347 | 0.179 | 0.477 |
| $C(25)$ | 2.300 | 0.169 | 0.453 |
| $C(30)$ | 1.222 | 0.161 | 0.476 |
| $C(35)$ | 0.578 | 0.155 | 0.471 |
| $C(40)$ | 0.357 | 0.151 | 0.507 |

Tabelle 4.6: Auf einem 120^2 -Gitter bei $\beta = 2$ wurde in doppelter Genauigkeit der Erwartungswert der Korrelationsfunktion $C(t)$ und daraus m bestimmt. Jeweils 1000000 Konfigurationen wurden erzeugt, jede 100. akzeptierte wurde für den Mittelwert der Korrelationsfunktion verwendet. Gerechnet wurde auf acht Prozessoren. Die Akzeptanzrate, die Gesamtrechnzeit R sowie der Erwartungswert von $\exp(-\Delta\mathcal{H})$ sind ebenfalls angegeben.

Ebenfalls ist in den beiden Tabellen das durch einen Fit an die Korrelationsfunktion bestimmte *massgap* eingetragen.

Die Differenzen zwischen Tabelle 4.5 und 4.6 sind in Tabelle 4.7 aufgelistet. Außerdem sind in der gleichen Tabelle zum besseren Vergleich auch die Daten aus der Simulation mit einfacher Genauigkeit aus Tabelle 4.5 eingetragen. Es ist festzustellen, daß im Gegensatz zu der Akzeptanzrate und $\exp(\Delta\mathcal{H})$ die Differenzen in $C(t)$ signifikant sind: teilweise bis zu drei mal so groß (bei $C(0)$ und $C(5)$), teilweise etwa gleich groß (bei $C(10)$ und $C(15)$) liegen sie in der Größenordnung der statistischen Fehler.

Damit ist gezeigt, daß die durch die Rundungsfehler entstehenden Fehler in Observablen nicht unbedingt im statistischen Fehler untergehen, sondern, natürlich abhängig von der vorliegenden Stichprobe, sogar größer als der statistische Fehler sein können, wenn man die gesamte Simulation vergleicht.

Dies liegt offensichtlich daran, daß der durch Rundungsfehler entstehende Effekt zwar am Ende einer MD-Trajektorie vernachlässigbar ist, wie die weiter oben dargestellten

| | <i>single</i> | Differenzen | |
|---------|-------------------|------------------------|-------------------|
| | Fehler/ 10^{-3} | $\Delta_{s,d}/10^{-3}$ | Fehler/ 10^{-6} |
| $C(0)$ | 0.363 | -0.7210 | -3.178 |
| $C(5)$ | 0.257 | -0.8737 | -5.597 |
| $C(10)$ | 0.214 | -0.3524 | -4.297 |
| $C(15)$ | 0.1921 | -0.1862 | -2.351 |
| $C(20)$ | 0.1783 | -0.1133 | -0.7433 |
| m | 0.518796 | -0.459 | 9.472 |

Tabelle 4.7: Differenzen zwischen einfacher und doppelter Genauigkeit aus den Tabellen 4.5 und 4.6 und zum Vergleich der statistische Fehler aus der *single precision*-Simulation.

Ergebnisse vermuten lassen. Für die wiederholte Aneinanderreihung der MD-Integration wird der Fehler aber fortgepflanzt, so daß er zu durchaus signifikanter Größe heranwächst. Interessanterweise wirken sich die Differenzen in der Korrelationsfunktion nicht genauso stark auf eine daraus bestimmte Größe aus, wie hier das *massgap*. Der aus den Daten der einfachen Genauigkeit bestimmte Wert für m stimmt innerhalb der Fit-Fehler mit jenem aus den Daten der doppelten Genauigkeit überein. Trotzdem ist $\Delta_{s,d}m$ in der Größenordnung des statistischen Fehlers.

Zusammenfassend kann also folgendes festgehalten werden: Es ist eindeutig ein Einfluß der Rundungsfehler bzw. der Reversibilitätsverletzung auf physikalische Observablen festzustellen. Dieser Einfluß hängt von der Gitter-Größe, von $\Delta\tau$ sowie der Trajektorienlänge ab.

Wird dieser Einfluß lediglich über eine MD-Integration betrachtet, so ist er so klein, daß er scheinbar vernachlässigbar ist. Wird allerdings eine gesamte Simulation durchgeführt, so nimmt der Fehler in dem Maß zu, daß der Fehler aus der Reversibilitätsverletzung größer wird als der statistische Fehler und somit unter Umständen keinen vernachlässigbaren Effekt mehr darstellt.

Der Erwartungswert von $\exp(-\Delta\mathcal{H})$ erweist sich für die vorliegenden Daten nicht als sensibler Indikator für Reversibilitätsverletzungen im HMC-Algorithmus. Daraus ein Kriterium zu entwickeln, das es erlaubt, ohne praktische Tests die Größe der Reversibilitätsverletzungen abzuschätzen, ist daher aufgrund der in dieser Arbeit dargestellten Resultate nicht möglich.

Kapitel 5

Zusammenfassung und Ausblick

In der vorliegenden Arbeit wurde zunächst das Skalierungsverhalten des Lyapunov-Exponenten ν untersucht, der die Reversibilitätsverletzung im HMC-Algorithmus und in verwandten Algorithmen kennzeichnet.

Ausgangspunkt war einerseits die Hypothese von Edwards et al. aus [3], der Exponent skaliere wie eine physikalische Größe. Dies hätte weitreichende Konsequenzen gehabt: Erstens hätte eine weitere charakteristische physikalische Größe zur Beschreibung des Modells zur Verfügung gestanden. Zweitens hätte es bedeutet, daß die Größe der Reversibilitätsverletzung vorausgesagt werden kann, wenn das Skalierungsverhalten bekannt ist. Im zweidimensionalen nichtlinearen $O(4)$ - σ -Modell und beim anharmonischen Oszillator wurde zur Überprüfung der Hypothese der Lyapunov-Exponent für einen großen Bereich in der Kopplungskonstanten bestimmt.

Andererseits sollte die Frage beantwortet werden, inwiefern die Reversibilitätsverletzung sich auf Resultate von Simulationen auswirkt, die mit dem HMC-Algorithmus durchgeführt werden. Die Reversibilitätsverletzung bedeutet für den HMC-Algorithmus, daß er seine Exaktheit verliert, also die Resultate (im Prinzip) mit systematischen Fehlern behaftet sind. Der dadurch entstehende Fehler kann vernachlässigbar klein sein, aber auch von signifikanter Größe: Dies war zu untersuchen.

Zur Beantwortung der ersten Frage wurde insbesondere beim nichtlinearen $O(4)$ - σ -Modell auch für große Werte in $\beta = 1/g$ darauf geachtet, *finite size*-Effekte auszuschließen. Zu diesem Zweck wurde die Korrelationslänge $1/m$ bestimmt. Eine Verifizierung oder Widerlegung der Hypothese war den Autoren in [3] gerade deswegen nicht möglich, weil sie *finite size*-Effekte wegen des sehr hohen Rechenzeitaufwands in der QCD nicht ausschließen konnten.

Die für den Lyapunov-Exponenten in der vorliegenden Arbeit erzielten Ergebnisse können nicht als Bestätigung der oben beschriebenen Hypothese interpretiert werden: Das Verhalten des Exponenten bei großen Werten von β ist im Vergleich zum Verhalten von m

nicht durch die universelle Skala Λ bestimmt: Während das asymptotische Skalierungsverhalten von m gut nachgewiesen werden konnte (Abbildung 4.3, Seite 47), war selbiges für den Exponenten nicht möglich, wie in Abbildung 4.6, Seite 53 zu erkennen ist.

Dies hat auch Auswirkungen auf den Kontinuumslimites: Es konnte ebenfalls nicht nachgewiesen werden, daß der Exponent innerhalb des Bereiches, in dem mit dem Beginn des asymptotischen Skalierungsverhalten gerechnet werden kann, auf Null abfällt. Somit sind auch Werte für den Kontinuumslimites mit Fehlern aus der Reversibilitätsverletzung behaftet.

Ganz im Gegenteil stellt sich das Verhalten des Lyapunov-Exponenten ν in Abhängigkeit von der Kopplung als allgemeine Eigenschaft nicht des zugrunde liegenden Modells, sondern der im HMC-Algorithmus verwendeten klassischen chaotischen Bewegungsgleichungen heraus. Dies konnte durch Vergleich des Verhaltens für verschiedene Modelle, das nichtlineare $O(4)$ - σ -Modell, den anharmonischen Oszillator und die QCD (durch Vergleich mit [20] und [3]), nachgewiesen werden.

Ferner wurde zur Beantwortung der zweiten Frage beim nichtlinearen $O(4)$ - σ -Modell untersucht, inwiefern die Reversibilitätsverletzung Auswirkung auf die Ergebnisse einer Simulation mit dem HMC-Algorithmus hat. Dazu wurden Erwartungswerte für mehrere Observablen, die mit verschiedenen Genauigkeiten (32- und 64-Bit) berechnet wurden, jeweils nach einer Molekular-Dynamik-Trajektorie verglichen: Deutlich konnte eine Abweichung des Mittelwertes der Differenz zwischen einfacher und doppelter Rechengenauigkeit in der Wirkung (Abbildung 4.9, Seite 59) und im Erwartungswert der Energie (Abbildung 4.10, Seite 62) nachgewiesen werden, die sich zunächst exponentiell mit zunehmender Trajektorienlänge vergrößert, um nach einer bestimmten Trajektorienlänge aufgrund der Kompaktheit der Gruppe $SU(2, \mathbb{C})$ zu sättigen.

Bei der Modell-Hamilton-Funktion hingegen war zu beobachten (Abbildung 4.12, Seite 64), daß sich die Fehler in den Feldern und den Impulsen nahezu aufheben, so daß relativ zum Erwartungswert nur noch ein sehr kleiner *bias* bestehen bleibt.

Die Abweichungen hängen bei nicht intrinsischen Observablen vom Volumen des Gitters, von der Integrationslänge und von der Schrittweite $\Delta\tau$ ab, während bei intrinsischen Observablen nur eine Abhängigkeit von der Integrationslänge festgestellt werden konnte.

Die Differenzen sind allerdings nach nur einer MD-Trajektorie jeweils relativ zum Erwartungswert der Observablen selbst so klein, daß daraus nicht geschlossen werden konnte, daß sie einen Einfluß auf Ergebnisse haben, der nicht in den statistischen Fehlern untergeht.

Darüber hinaus wurden zwei vollständige Simulationen in beiden Rechengenauigkeiten bei der gleichen Kopplungskonstanten $\beta = 2$ durchgeführt: Beide wurden bei den gleichen Anfangsbedingungen gestartet und die Ergebnisse zum Ende der gesamten Simulation verglichen. Dieses muß als das realistischere Szenario angesehen werden, als der Vergleich

nach nur einer MD-Trajektorie, denn in der Praxis wird die gesamte Rechnung in einer Genauigkeit durchgeführt. Für diese Untersuchung wurde die Korrelationsfunktion $C(t)$ ausgewählt, weil es ebenfalls von Interesse ist, die relativen Auswirkungen auf betragslich kleine Erwartungswerte zu untersuchen. Da bei der Korrelationsfunktion trotz großer Stichprobe der statistische Fehler bereits für kleine Werte von t die Größe des Erwartungswerts erreicht hat, sind genaue Aussagen zu diesem Punkt aber nicht möglich.

Bei der Untersuchung der Korrelationsfunktion zeigte sich (Tabelle 4.7, Seite 68), daß die Differenzen zwischen den Werten, die mit verschiedener Genauigkeit simuliert wurden, in der Größenordnung des statistischen Fehlers liegen. Teilweise überschritten die Abweichungen den statistischen Fehlers sogar bis zum Dreifachen. Aus letzterem folgt, daß die Abweichungen aufgrund der vorliegenden Daten als signifikant angesehen werden müssen. Dies bedeutet für die Ergebnisse einer Simulation, daß sie mit nicht vernachlässigbaren systematischen Fehlern behaftet sind.

Die Idee, mit Hilfe des Erwartungswertes von $\exp(-\Delta\mathcal{H})$ ein Kriterium zu entwickeln, mit dem die Größenordnung des Einflusses der Reversibilitätsverletzung auf die Ergebnisse bestimmt werden kann, stellte sich als unpraktikabel heraus, weil der Erwartungswert keine signifikante Abweichung vom analytisch bestimmten Wert 1 zeigt (Tabellen 4.4, Seite 65).

Ausblick

Aufgrund der in der vorliegenden Arbeit erzielten Resultate ist es für die Zukunft von Interesse, einen Test auf die Reversibilitätsverletzung auch in der QCD durchzuführen. Dazu wurde in Fortsetzung dieser Arbeit bereits damit begonnen [28], eine solche Funktionalität in vorhandenem Quelltext einzubauen, der für Simulationen in der QCD verwendet wird. Dabei wird es allerdings wegen des enormen Rechenaufwands und der damit verbundenen Kosten nicht möglich sein, verschiedene Genauigkeiten zu vergleichen.

Um trotzdem für die Praxis relevante Aussagen machen zu können, ist beispielsweise folgende Vorgehensweise denkbar: Es soll ein Diagramm erstellt werden, in dem in Abhängigkeit von der Reversibilitätsverletzung in den Feldern $|\delta\Delta\phi|$ – eine Größe, die ohne Vergleich von verschiedenen Genauigkeiten durch Hin- und Zurückintegrieren gewonnen werden kann – die Differenz des Erwartungswertes einer „Referenzobservablen“ (z.B. der Korrelationsfunktion bei $t = 5$) zwischen unterschiedlichen Genauigkeiten dargestellt wird.

Mit Hilfe eines solchen Diagrammes könnte es dann möglich sein, nur über die Bestimmung der Reversibilitätsverletzung in den Feldern eine Aussage darüber zu machen, ob der Einfluß der Verletzung der Reversibilität auf die Resultate einer konkreten Simulation signifikant ist oder im statistischen Fehler untergeht.

Danksagung

An dieser Stelle möchte ich die Gelegenheit nutzen, allen zu danken, die direkt oder indirekt zum Gelingen dieser Arbeit beigetragen haben.

Zuallererst möchte ich Karl Jansen danken, von dem auch die Idee zu dieser Arbeit stammt: Trotz seines oft sehr engen Zeitplans hat er sich immer ausführlich Zeit genommen, mit mir Probleme zu besprechen und Fragen zu klären, mir Ideen und Anregungen zu geben. Außerdem hat er eine Atmosphäre geschaffen, die immer dazu beigetragen hat, mich neu zu motivieren, nicht zuletzt, weil ich stets die Möglichkeit hatte, auch an Physik außerhalb meiner Diplomarbeit teilzuhaben.

Außerdem hat er mir einen Forschungsaufenthalt in Mailand (Dank dafür an das INFN in Mailand) ermöglicht, bei dem ich die Möglichkeit hatte, mich mit Roberto Frezzotti auszutauschen, dem ich ebenfalls für wertvolle Anregungen danken möchte.

Des weiteren gilt mein Dank Volkard Linke, der ebenfalls immer Zeit für mich hatte und mich vor allem zu Präzision und Genauigkeit in dieser Arbeit angeleitet hat. Die gute Atmosphäre in der von ihm geleiteten Arbeitsgruppe hat viel zum Gelingen dieser Arbeit beigetragen.

Dank auch an Frank Hofheinz, der bei Problemen immer zugehört hat und sogar seine Freundin für die Korrektur dieser Arbeit eingespannt hat, und Jens Dreger, der mir wertvolle Hilfestellung für \LaTeX und \MATHEMATICA gegeben hat.

Nicht zuletzt gilt mein Dank meiner Familie und meiner Freundin: Meiner Familie, die immer an mich geglaubt hat und damit dazu beigetragen hat, daß ich bis zu dieser Danksagung gekommen bin. Meiner Freundin Katja Schiffers, nicht nur, weil sie diese Arbeit mehrmals gelesen und korrigiert hat, sondern auch, weil sie mir in den letzten drei Jahren immer zur Seite gestanden hat.

Auch die Kollegen vom DESY Zeuthen und vom Zentralinstitut für angewandte Mathematik (ZAM) des Forschungszentrums Jülich dürfen nicht unerwähnt bleiben: Ich hatte die Möglichkeit, die jeweiligen PC-Cluster zu nutzen und die damit verbundene Unterstützung in Anspruch zu nehmen. Am ZAM hatte ich außerdem im Rahmen eines Gaststudentenprogramms die Möglichkeit, paralleles Programmieren zu lernen und einzusetzen.

Anhang A

Mathematischer Anhang

A.1 Herleitung der Zwei-Punkt-Funktion

Sei $\hat{S}(t) = \exp(-Hta)$ der Zeittranslationsoperator in der euklidisierten Gitterfeldtheorie. Dann gilt auf einem Gitter mit periodischen Randbedingungen für einen Energieeigenzustand ϕ :

$$\hat{S}(t)|\phi\rangle = e^{-E_\phi t}|\phi\rangle.$$

Damit kann nun zunächst folgender Ausdruck umgeformt werden, indem zweimal ein vollständiger Satz von Energieeigenzuständen eingeschoben wird:

$$\begin{aligned}\langle \mathcal{O}(t)\mathcal{O}^\dagger(0) \rangle &= \frac{1}{\mathcal{Z}} \text{Sp} \hat{S}(L-t)\mathcal{O}\hat{S}(t)\mathcal{O}^\dagger \\ &= \frac{1}{\mathcal{Z}} \sum_{\phi, \phi'} e^{-E_\phi(L-t)} \langle \phi | \mathcal{O} | \phi' \rangle e^{-E_{\phi'} t} \langle \phi' | \mathcal{O}^\dagger | \phi \rangle.\end{aligned}$$

L ist dabei die Ausdehnung des Gitters in Zeitrichtung. Läßt man nun $L \rightarrow \infty$ gehen und hält t oder $L-t$ fest, so folgt:

$$\begin{aligned}\mathcal{Z} \langle \mathcal{O}(t)\mathcal{O}^\dagger(0) \rangle &= \langle 0 | \mathcal{O} | 0 \rangle \langle 0 | \mathcal{O}^\dagger | 0 \rangle \\ &+ \sum_{|\phi\rangle \neq |0\rangle} \langle 0 | \mathcal{O} | \phi \rangle \langle \phi | \mathcal{O}^\dagger | 0 \rangle e^{-E_\phi t} \\ &+ \sum_{|\phi\rangle \neq |0\rangle} \langle \phi | \mathcal{O} | 0 \rangle \langle 0 | \mathcal{O}^\dagger | \phi \rangle e^{-E_\phi(L-t)}.\end{aligned}$$

$|0\rangle$ ist dabei der Grundzustand. Damit läßt sich nun die Zwei-Punkt-Funktion ausdrücken:

$$\begin{aligned}
 C(t) &= \langle \mathcal{O} \mathcal{O}^\dagger \rangle - |\langle \mathcal{O} \rangle|^2 \\
 &= \frac{1}{\mathcal{Z}} \sum_{|\phi\rangle \neq |0\rangle} |\langle \phi | \mathcal{O} | 0 \rangle|^2 (e^{-E_\phi t} + e^{-E_\phi(L-t)}) \\
 &= \frac{2}{\mathcal{Z}} \sum_{|\phi\rangle \neq |0\rangle} |\langle \phi | \mathcal{O} | 0 \rangle|^2 e^{-E_\phi L/2} \cosh[E_\phi(t - L/2)].
 \end{aligned} \tag{A-1}$$

Ist t bzw $L - t$ groß genug, so überlebt nur der Zustand $|1\rangle$ mit der kleinsten Energie, so daß die Korrelationsfunktion wie folgt abgeschätzt werden kann:

$$C(t) \approx \frac{2}{\mathcal{Z}} |\langle 1 | \mathcal{O} | 0 \rangle|^2 e^{-E_1 L/2} \cosh[E_1(t - L/2)], \quad t \text{ groß gegen } 1. \tag{A-2}$$

Wird auch der Nenner $\mathcal{Z} = \text{Sp}(\hat{S}(L))$ berücksichtigt, so ist die Zwei-Punkt-Funktion für große t abhängig von der Energiedifferenz (oder dem *massgap*) zwischen Grundzustand und erstem angeregten Zustand:

$$C(t) \propto \cosh[(E_1 - E_0)(t - L/2)], \quad t \text{ groß gegen } 1. \tag{A-3}$$

A.2 Ergänzungen zu $SU(2, \mathbb{C})$

In diesem Abschnitt werden zwei Beweise angegeben, die die Abbildung von der dreidimensionalen Sphäre S_3 auf die Gruppe $SU(2, \mathbb{C})$ möglich machen. Außerdem wird dabei eine konkrete Darstellung vorgestellt, die bei der Implementierung des HMC für das nicht-lineare $O(4)$ σ -Modell verwendet wurde.

Zunächst ein Lemma:

Lemma A.2.1 *Seien $\sigma_x, \sigma_y, \sigma_z$ die Pauli-Matrizen, $\vec{x}, \vec{y} \in S_3$ und sei T definiert wie in (3-15). Dann gilt*

$$\begin{aligned}
 \vec{x}\vec{y} &= \frac{1}{2} \text{Sp} [(x_4 \mathbb{I} + i(x_1 \sigma_1 + x_2 \sigma_2 + x_3 \sigma_3))(y_4 \mathbb{I} - i(y_1 \sigma_1 + y_2 \sigma_2 + y_3 \sigma_3))] \\
 &= \langle T(\vec{x}), T(\vec{y}) \rangle.
 \end{aligned} \tag{A-4}$$

Beweis:

$$\begin{aligned}
 &\text{Sp} [(x_4 \mathbb{I} - i(x_1 \sigma_1 + x_2 \sigma_2 + x_3 \sigma_3))(y_4 \mathbb{I} + i(y_1 \sigma_1 + y_2 \sigma_2 + y_3 \sigma_3))] = \\
 &= \text{Sp} [(x_1 y_1 + x_2 y_2 + x_3 y_3 + x_4 y_4) \mathbb{I}] \\
 &= 2(\vec{x}\vec{y}).
 \end{aligned}$$

□

Als nächstes wird ein Beweis für den Satz 3.2.1 angegeben:

Die Abbildung T war in (3-15) wie folgt definiert:

$$T : S_3 \mapsto SU(2, \mathbb{C}) : T(\vec{x}) = x_4\mathbb{I} + i(x_1\sigma_1 + x_2\sigma_2 + x_3\sigma_3).$$

Da die Stetigkeit von T mit (3-16) schon bewiesen ist, bleibt noch zu zeigen, daß T bijektiv ist. Beim Beweis wird folgendermaßen vorgegangen: Erstens wird gezeigt, daß jedes U der Form $U = x_4\mathbb{I} + i(x_1\sigma_1 + x_2\sigma_2 + x_3\sigma_3)$ ein Element der Gruppe $SU(2, \mathbb{C})$ ist, falls $(x_1, x_2, x_3, x_4) \in S_3$. Zweitens wird dann gezeigt, daß T bijektiv ist.

Lemma A.2.2 Sei $\vec{x} \in S_3$ und $U = x_4\mathbb{I} + i(x_1\sigma_1 + x_2\sigma_2 + x_3\sigma_3)$. Dann ist $U \in SU(2, \mathbb{C})$.

Beweis: Dafür ist zu zeigen, daß 1. $UU^\dagger = U^\dagger U = \mathbb{I}$ und 2. $\det(U) = 1$ ist.

1. Da mit (3-13) $\sigma_i^2 = \mathbb{I}$ und $\sigma_i\sigma_j = -\sigma_j\sigma_i$ für $i \neq j$ gilt, kann gezeigt werden:

$$\begin{aligned} UU^\dagger &= (x_4\mathbb{I} + i(x_1\sigma_1 + x_2\sigma_2 + x_3\sigma_3))(x_4\mathbb{I} - i(x_1\sigma_1 + x_2\sigma_2 + x_3\sigma_3)) \\ &= (x_4^2 + x_1^2 + x_2^2 + x_3^2)\mathbb{I} = \mathbb{I}. \end{aligned}$$

Der Beweis geht analog für $U^\dagger U = \mathbb{I}$.

2. Da $\vec{x} \in S_3$ ist, gilt $\|\vec{x}\| = 1$, so daß folgt:

$$\begin{aligned} \det(U) &= \det \begin{pmatrix} x_4 + ix_3 & x_2 + ix_1 \\ -x_2 + ix_1 & x_4 - ix_3 \end{pmatrix} \\ &= x_4^2 + x_1^2 + x_2^2 + x_3^2 = \|\vec{x}\|^2 = 1. \end{aligned}$$

Also ist $U \in SU(2, \mathbb{C})$. □

Damit ist gezeigt, daß T die S_3 nach $SU(2, \mathbb{C})$ abbildet. Als nächstes folgt wiederum ein Lemma:

Lemma A.2.3 $\{\mathbb{I}, i\sigma_1, i\sigma_2, i\sigma_3\}$ sind linear unabhängig über den reellen Zahlen.

Beweis: Sei $d\mathbb{I} + i(a\sigma_1 + b\sigma_2 + c\sigma_3) = 0$ mit $a, b, c, d \in \mathbb{R}$, dann ist auch das dazu adjungierte gleich 0 und es gilt:

$$0 = [d\mathbb{I} + i(a\sigma_1 + b\sigma_2 + c\sigma_3)][d\mathbb{I} - i(a\sigma_1 + b\sigma_2 + c\sigma_3)] = a^2 + b^2 + c^2 + d^2.$$

Daraus folgt, daß $a = b = c = d = 0$ ist, was zu zeigen war. □

Mit Hilfe dieses Lemmas kann nun gezeigt werden, daß T injektiv ist:

Anhang A Mathematischer Anhang

Lemma A.2.4 Sei T definiert wie in (3-15). Dann ist T injektiv.

Beweis: Sei $\vec{x}, \vec{y} \in S_3$ mit $\vec{x} \neq \vec{y}$ und $x_4\mathbb{I} + i(x_1\sigma_1 + x_2\sigma_2 + x_3\sigma_3) = y_4\mathbb{I} + i(y_1\sigma_1 + y_2\sigma_2 + y_3\sigma_3)$. Dann folgt, da die $\{\mathbb{I}, i\sigma_1, i\sigma_2, i\sigma_3\}$ linear unabhängig sind:

$$\vec{x} = \vec{y}$$

und damit ein Widerspruch. □

Damit bleibt nur noch die Sujektivität zu zeigen. Auch dieser Teil ist wieder in zwei Hilfssätze unterteilt:

Lemma A.2.5 Ist $U = d\mathbb{I} + i(a\sigma_1 + b\sigma_2 + c\sigma_3)$ mit $U \in SU(2, \mathbb{C})$. Dann ist $(a, b, c, d) \in S_3$.

Beweis: Sei $a, b, c, d \in \mathbb{R}$ und $U = d\mathbb{I} + i(a\sigma_1 + b\sigma_2 + c\sigma_3) \in SU(2, \mathbb{C})$. Dann ist

$$1 = \det(U) = a^2 + b^2 + c^2 + d^2.$$

Also ist $(a, b, c, d) \in S_3$. □

Im zweiten Lemma wird gezeigt, daß sich jedes Element von $SU(2, \mathbb{C})$ in der Form $U = d\mathbb{I} + i(a\sigma_1 + b\sigma_2 + c\sigma_3)$ darstellen läßt:

Lemma A.2.6 Ist U ein Element von $SU(2, \mathbb{C})$, so gibt es $a, b, c, d \in \mathbb{R}$ mit $U = d\mathbb{I} + i(a\sigma_1 + b\sigma_2 + c\sigma_3)$.

Beweis: Sei $U \in SU(2, \mathbb{C})$, dann gilt:

$$U^\dagger U = UU^\dagger = 1, \quad \det(U) = 1, \tag{A-5}$$

mit $U^\dagger = (U^*)^T$ und somit folgt:

$$U^\dagger = U^{-1}. \tag{A-6}$$

Sei $u, v, w, z \in \mathbb{C}$ und

$$U = \begin{pmatrix} u & v \\ w & z \end{pmatrix}, \tag{A-7}$$

dann ist

$$U^{-1} = \frac{1}{\det(U)} \begin{pmatrix} z & -v \\ -w & u \end{pmatrix}. \tag{A-8}$$

Damit folgt aus (A-6), da $\det(U) = 1$ ist:

$$w = -v^*, \quad z = u^*. \tag{A-9}$$

Also kann U in folgender Weise dargestellt werden:

$$U = \begin{pmatrix} u & v \\ -v^* & u^* \end{pmatrix} \quad \text{mit } uu^* + vv^* = 1. \quad (\text{A-10})$$

Wählt man $u = d + ic$ und $v = b + ia$, so folgt mit der Darstellung der Pauli-Matrizen (3-14): $U = d\mathbb{I} + i(a\sigma_1 + b\sigma_2 + c\sigma_3)$. \square

Somit ist gezeigt, daß T bijektiv ist und Satz 3.2.1 ist bewiesen. Die Abbildung von S_3 auf $SU(2, \mathbb{C})$ ist also ohne Informationsverlust möglich.

Außerdem stellt (A-10) eine konkrete Darstellung von $SU(2, \mathbb{C})$ mit zwei komplexen Zahlen dar, die es ermöglicht, viel Rechenzeit zu sparen.

A.3 Haar-Maß

Sei \mathcal{U} eine topologische Gruppe, d.h. die Abbildung von $\mathcal{U} \times \mathcal{U} \rightarrow \mathcal{U} : U, V \mapsto UV^{-1}$ ist stetig. Außerdem sei die Gruppe \mathcal{U} als topologischer Raum ein kompakter metrischer Raum. Sei \mathcal{B} die Klasse der Borelmengen von \mathcal{U} . Ein Maß auf \mathcal{U} ist ein nichtnegatives Maß auf $(\mathcal{U}, \mathcal{B})$.

Definition A.3.1 (Invariantes Haar-Maß auf \mathcal{U}) *Ein invariantes Maß auf der Gruppe \mathcal{U} ist ein Maß λ für das gilt:*

$$\lambda(U^{-1}\mathcal{A}) = \lambda(\mathcal{A}) \quad \text{und} \quad \lambda(\mathcal{A}U^{-1}) = \lambda(\mathcal{A}), \quad \forall \mathcal{A} \in \mathcal{B}.$$

Dabei ist die Menge $U^{-1}\mathcal{A}$ gleich der Menge aller Elemente der Form $U^{-1}V$ mit $V \in \mathcal{A}$.

Im folgenden wird $d\lambda(U)$ als dU geschrieben. Für die Gruppe $SU(2, \mathbb{C})$ kann das Haar-Maß explizit konstruiert werden, wenn man die Parametrisierung

$$U \sim x_4\mathbb{I} + i\vec{x}\vec{\sigma}, \quad x_4^2 + \vec{x}\vec{x} = 1, \quad x = (\vec{x}, x_4)$$

benutzt. Damit kann schnell gezeigt werden, daß

$$dU = \frac{1}{\pi^2} \delta(x^2 - 1) d^4x \quad (\text{A-11})$$

gilt, was im übrigen mit dem einheitlichen Maß von S_3 übereinstimmt. Mit dieser expliziten Darstellung schließlich ist es möglich, unter anderem die folgenden Gleichungen herzuleiten: Die Normierung:

$$\int dU = 1 \quad (\text{A-12})$$

und folgende wichtige Eigenschaft:

$$\int dUU = 0. \quad (\text{A-13})$$

Es kann außerdem gezeigt werden (siehe zum Beispiel [29]):

$$\int dUU_{ij}U_{kl}^{-1} = \frac{1}{2}\delta_{jk}\delta_{il}, \quad U = (U_{ij}), \quad (\text{A-14})$$

so daß folgender Ausdruck integriert werden kann:

$$\begin{aligned} & \int dU \text{Sp}(UV) \text{Sp}(U^{-1}V') \\ &= \int dU (U_{00}V_{00} + U_{01}V_{10} + U_{10}V_{01} + U_{11}V_{11}) \times \\ & \quad (U_{00}^{-1}V'_{00} + U_{01}^{-1}V'_{10} + U_{10}^{-1}V'_{01} + U_{11}^{-1}V'_{11}) \\ &= \int dU (U_{00}V_{00}U_{00}^{-1}V'_{00} + U_{01}V_{10}U_{10}^{-1}V'_{01} + U_{10}V_{01}U_{01}^{-1}V'_{10} + U_{11}V_{11}U_{11}^{-1}V'_{11}) \\ &= \frac{1}{2}(V_{00}V'_{00} + V_{01}V'_{10} + V_{10}V'_{01} + V_{11}V'_{11}) \\ &= \frac{1}{2}\text{Sp}(VV'). \end{aligned} \quad (\text{A-15})$$

A.4 *strong coupling expansion*

Zur Verifikation der Implementation ist es nützlich, Erwartungswerte analytisch berechnen zu können. Dies ist auf dem Gitter approximativ möglich für den Fall, daß die Werte der Kopplungskonstanten β klein sind und alle Terme proportional zu β^n , $n \geq 2$ vernachlässigt werden können. Dann ist es möglich, die Integration über die Felder auszuführen und man erhält den Erwartungswert in Abhängigkeit von β . Das so berechnete Ergebnis kann dann mit den numerischen Resultaten verglichen werden.

Entsprechend der im vorigen Abschnitt dargestellten Eigenschaften des Haar-Maßes kann der Erwartungswert der Energie für $\beta \rightarrow 0$ näherungsweise berechnet werden. Zunächst wird dafür die Zustandssumme in erster Ordnung behandelt: Man entwickelt $\exp(-S)$ in einer Potenzreihe in β :

$$\exp(-S) = \exp\left(\frac{\beta}{2} \sum_{\mathbf{x}, \mu} \text{Sp}(U_{\mathbf{x}}U_{\mathbf{x}+\mu}^{\dagger})\right) = 1 + \frac{\beta}{2} \sum_{\mathbf{x}, \mu} \text{Sp}(U_{\mathbf{x}}U_{\mathbf{x}+\mu}^{\dagger}) + O(\beta^2). \quad (\text{A-16})$$

Dies in die Zustandssumme eingesetzt ergibt:

$$\begin{aligned}
 \mathcal{Z} &= \int \prod_{\mathbf{x}} dU_{\mathbf{x}} e^{-S} \\
 &= \int \prod_{\mathbf{x}} dU_{\mathbf{x}} \prod_{\mathbf{y}, \mu} \left(1 + \frac{\beta}{2} \text{Sp}(U_{\mathbf{y}} U_{\mathbf{y}+\mu}^{\dagger}) + O(\beta^2)\right) \\
 &= \int \prod_{\mathbf{x}} dU_{\mathbf{x}} + O(\beta^2) \\
 &= 1 + O(\beta^2).
 \end{aligned} \tag{A-17}$$

Anschließend wird die Entwicklung für den Erwartungswert der Energie in Abhängigkeit von der Kopplungskonstanten β ebenfalls in erster Ordnung durchgeführt, wobei die Entwicklung (A-17) für \mathcal{Z} und die Gleichung (A-15) benutzt wird:

$$\begin{aligned}
 -\langle E \rangle(\beta) &= \frac{1}{\mathcal{Z}} \int \prod_{\mathbf{x}} dU_{\mathbf{x}} \frac{1}{N} \sum_{\mathbf{y}, \mu} \frac{1}{2} \text{Sp}(U_{\mathbf{y}} U_{\mathbf{y}+\mu}^{\dagger}) \prod_{z, \nu} \left(1 + \frac{\beta}{2} \text{Sp}(U_z U_{z+\nu}^{\dagger}) + O(\beta^2)\right) \\
 &= \int \prod_{\mathbf{x}} dU_{\mathbf{x}} \frac{1}{2N} \sum_{\mathbf{y}, \mu} \text{Sp}(U_{\mathbf{y}} U_{\mathbf{y}+\mu}^{\dagger}) \left(\sum_{z, \nu} \frac{\beta}{2} \text{Sp}(U_z U_{z+\nu}^{\dagger}) \right) + O(\beta^2) \\
 &= \int \prod_{\mathbf{x}} dU_{\mathbf{x}} \frac{\beta}{4N} \sum_{\mathbf{y}, \mu} \text{Sp}(U_{\mathbf{y}} U_{\mathbf{y}+\mu}^{\dagger}) \text{Sp}(U_{\mathbf{y}} U_{\mathbf{y}+\mu}^{\dagger}) + O(\beta^2) \\
 &= \int \prod_{\mathbf{x}} dU_{\mathbf{x}} \frac{\beta}{4N} \sum_{\mathbf{y}, \mu} \text{Sp}(U_{\mathbf{y}} U_{\mathbf{y}+\mu}^{\dagger}) \text{Sp}(U_{\mathbf{y}}^{\dagger} U_{\mathbf{y}+\mu}) + O(\beta^2) \\
 &= \frac{\beta}{4N} 2N \int dU \text{Sp}(U_{\mathbf{x}} U_{\mathbf{y}}^{\dagger}) \text{Sp}(U_{\mathbf{x}}^{\dagger} U_{\mathbf{y}}) + O(\beta^2) \\
 &= \frac{\beta}{2} \left(\frac{1}{2} \text{Sp}(U_{\mathbf{y}} U_{\mathbf{y}}^{\dagger}) \right) + O(\beta^2) \\
 &= \frac{\beta}{2} + O(\beta^2).
 \end{aligned}$$

Der Energieerwartungswert zeigt also für kleine β ein lineares Verhalten:

$$\langle E \rangle(\beta) \approx -\frac{\beta}{2}. \tag{A-18}$$

Daraus folgt mit Definition (1-30) unmittelbar die Entwicklung für E_W :

$$\langle E_W \rangle(\beta) = \frac{1}{2} \langle E \rangle(\beta) \approx -\frac{\beta}{4}. \tag{A-19}$$

Anhang B

Numerischer Anhang

Alle für diese Arbeit verwendeten Programme wurden in C++ geschrieben. Das hat den Vorteil, daß die Programme letztendlich sehr gut lesbar sind und die mathematischen Formeln sehr intuitiv im Programm-Text auftauchen. Allerdings erfordert die Benutzung von Klassen auch einen etwas größeren Programmier-Aufwand, um Programme zu erzeugen, die möglichst wenig Rechenzeit verbrauchen.

Zur Performance-Steigerung wurden *expression templates* verwendet:

B.1 *expression templates* in C++

Das Problem von Klassen in C++ ist, daß es dem Compiler nahezu unmöglich gemacht wird, performanten Code zu erzeugen. Das liegt daran, daß beispielsweise folgender einfacher Ausdruck von drei Repräsentationen einer Klasse von Matrizen

$$A = B + C;$$

nicht direkt ausgeführt werden kann; das Ergebnis von $B + C$ wird zwischengespeichert und das Zwischenergebnis dann A zugewiesen. Das bedeutet, daß für diese Operation einmal zusätzlich Speicher reserviert werden muß, was sehr zeitaufwendig ist.

Dieses Problem kann man umgehen, wenn man sogenannte *expression templates* nach einer Idee von T. Veldhuizen verwendet, die er in [30] darlegt. Die Funktionalität der Templates in C++ wird dabei ausgenutzt, um dem Compiler die Ableitung der Ausdrücke abzunehmen und damit die Erzeugung temporärer Objekte zu vermeiden.

Zu diesem Zweck werden Template-Klassen eingeführt (die sogenannten *expression templates*), die Ausdrücken entsprechen, also z.B. $A + B$, die aber nur Zeiger auf die beiden Objekte und die Operation „+“ enthalten. Diese Zeiger werden erst bei der Auswertung der Operation „=“ aufgelöst, so daß der Compiler, wenn er entsprechende Optimierungsfunktionen hat, die Ergebnisse wieder direkt, ohne Zwischenspeicherung, zuweisen kann.

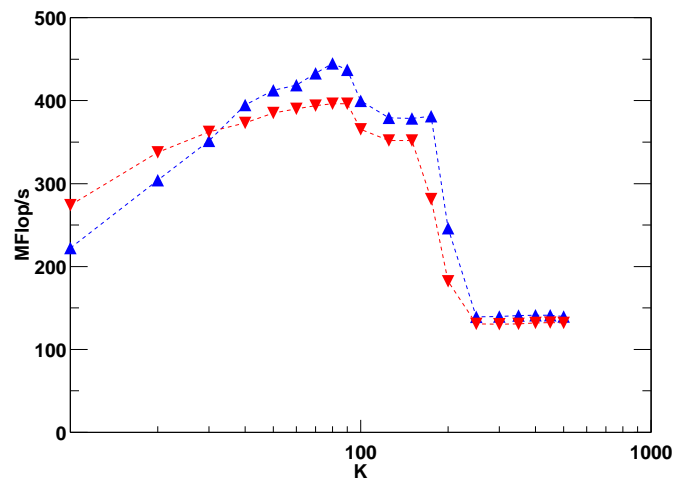


Abbildung B.1: MFlop/s über der Matrixgröße K für ein C-Programm ▼ und ein C++-Programm mit *expression templates* ▲.

Mit Hilfe dieser *expression templates* kann die Performance eines C++- Programms bis zu der entsprechender C- oder Fortran-Programme gesteigert werden. Für das Beispiel der Matrix-Vektor-Operation

$$\vec{d} = \mathbf{A}\vec{b} + \vec{c}$$

für Matrizen und Vektoren der Dimension K wurden eine C-Implementierung und eine C++-Implementierung mit *expression templates* verglichen. Die Ergebnisse für einen Athlon XP 1500+ und den GNU C++-Compiler finden sich in Abbildung B.1.

Das Plateau ab ca. $K = 200$ erklärt sich mit der Bandbreite, die der Prozessor zum Hauptspeicher hat, ab dort passen die Daten also nicht mehr in den *secondary level-cache*. Beim ersten Plateau bei ca. $K = 100$ wachsen die Datengrößen über den *first level-cache*.

B.2 Zufallszahlen-Generator

Für die in der vorliegenden Arbeit benutzten Monte-Carlo-Methoden müssen Zufallszahlen auf dem Computer generiert werden. Damit das Programm auf verschiedenen Betriebssystemen gleich arbeitet, es auch im parallelisierten Code reproduzierbare Ergebnisse liefert und die Zufallszahlen schnell erzeugt werden, wurde für diese Arbeit ein Zufallszahlen-Generator implementiert, der diese Eigenschaften erfüllt.

Es gibt verschiedene Möglichkeiten, einen solchen Zufallszahlen-Generator zu definieren. Im Anschluß wird die Möglichkeit kurz vorgestellt, die für die vorliegende Arbeit implementiert wurde.

Für die Implementierung des Zufallszahlen-Generators wurden die Überlegungen zu einem portablen hochwertigen Generator von M. Lüscher verwendet, die er in [31] darstellt. Sie gehen auf einen Algorithmus von Marsaglia und Zaman zurück.

Sei b eine beliebige natürliche Zahl und X die Menge aller natürlichen Zahlen x mit $1 \leq x < b$. Der Algorithmus erzeugt eine Folge von Zufallszahlen $x_0, x_1, x_2, \dots \in X$ gemeinsam mit einer Folge von *carry-bits* $c_0, c_1, c_2, \dots \in \{0, 1\}$ rekursiv aus einer Anfangskonfiguration nach folgender Vorschrift:

Seien $r, s \in \mathbb{N}$, so daß $r > s \geq 1$. Für $n \geq r$ wird zunächst

$$\Delta_n = x_{n-s} - x_{n-r} - c_{n-1}$$

berechnet und dann x_n und c_n aus

$$x_n = \Delta_n, \quad c_n = 0, \quad \text{falls } \Delta_n \geq 0,$$

$$x_n = \Delta_n + b, \quad c_n = 1, \quad \text{falls } \Delta_n < 0$$

bestimmt. Um die Rekursion starten zu können, müssen die ersten r Werte als Anfangskonfiguration vorgegeben werden, wobei einige triviale Initialisierungen¹ vermieden werden sollten.

Wählt man

$$b = 2^{24}, \quad r = 24, \quad s = 10,$$

so sind alle Zahlen in X und b ohne Rundungsfehler auf dem Computer darstellbar und der Algorithmus somit exakt, weil auch bei der Berechnung der Zufallszahlen nicht gerundet werden muß. Auch x/b ist dann für $x \in X$ exakt in einfacher Präzision (32 Bit) darstellbar.

¹ siehe dazu [31]

Abbildungsverzeichnis

| | | |
|------|---|----|
| 2.1 | Autokorrelationsfunktion für verschiedene τ | 25 |
| 2.2 | Korrelationslänge | 28 |
| 4.1 | Energieerwartungswert über β und <i>Strong coupling expansion</i> | 45 |
| 4.2 | Erwartungswert der Zweipunktfunktion | 46 |
| 4.3 | <i>massgap</i> und $m/\Lambda_{\overline{MS}}$ über β | 47 |
| 4.4 | Zur Bestimmung von $\Delta\delta U(\tau)$ | 51 |
| 4.5 | Reversibilitätsverletzung für verschiedene Integrationslängen. | 52 |
| 4.6 | Lyapunov-Exponent in Abhängigkeit von der Kopplungskonstante | 53 |
| 4.7 | Der Lyapunov-Exponent für den anharmonischen Oszillator. | 55 |
| 4.8 | <i>bias</i> in der Wirkung | 58 |
| 4.9 | Entwicklung des <i>bias</i> und von σ der Wirkung | 59 |
| 4.10 | Entwicklung des <i>bias</i> und von σ des Energieerwartungswertes | 62 |
| 4.11 | <i>bias</i> in der Hamilton-Funktion | 63 |
| 4.12 | Entwicklung des <i>bias</i> der Hamilton-Funktion und des Energieerwartungswertes | 64 |
| B.1 | Performancevergleich C und C++ (<i>expression templates</i>) | 81 |

Tabellenverzeichnis

| | | |
|-----|--|----|
| 4.1 | Parameter des HMC-Algorithmus. | 42 |
| 4.2 | Verifikation der Implementierung. | 44 |
| 4.3 | Koeffizienten der Fehlerentwicklung für die Wirkung. | 61 |
| 4.4 | Vergleich von Akzeptanzrate und $\langle \exp(-\Delta\mathcal{H}) \rangle$ in einfacher und doppelter Genauigkeit. | 65 |
| 4.5 | $C(t)$ für eine Simulation in 32-Bit-Genauigkeit. | 66 |
| 4.6 | $C(t)$ für eine Simulation in 64-Bit-Genauigkeit. | 67 |
| 4.7 | Differenzen von $C(t)$ zwischen einfacher und doppelter Genauigkeit. | 68 |

Literaturverzeichnis

- [1] WILSON, K. G.: Confinement of quarks. In: *Phys. Rev.* D10 (1974), S. 2445–2459
- [2] DUANE, S. ; KENNEDY, A. D. ; PENDLETON, B. J. ; ROWETH, D.: Hybrid Monte Carlo. In: *Phys. Lett.* B195 (1987), S. 216–222
- [3] EDWARDS, R. G. ; HORVATH, I. ; KENNEDY, A. D.: Instabilities and non-reversibility of Molecular Dynamics trajectories. In: *Nucl. Phys.* B484 (1997), S. 375–402. hep-lat/9606004
- [4] FEYNMAN, R. P.: Space-time approach to non-relativistic quantum mechanics. In: *Rev. Mod. Phys.* 20 (1948), S. 367–387
- [5] KLEINERT, H.: *Path integrals in quantum mechanics, statistics and polymer physics.* 2nd Edition. World Scientific, Singapore, 1995
- [6] GLIMM, J. ; JAFFE, A.: *Quantum physics: A functional integral point of view.* 2nd Edition. Springer Verlag, 1987
- [7] GREINER, W.: *Theoretische Physik.* Bd. 7a: *Feldquantisierung.* Verlag Harri Deutsch, 1993
- [8] OSTERWALDER, K. ; SCHRADER, R.: Axioms for euclidean Green’s functions. 2. In: *Commun. Math. Phys.* 42 (1975), S. 281
- [9] MONTVAY, I. ; MÜNSTER, G.: *Quantum fields on a lattice.* Cambridge University Press, 1994 (Cambridge Monographs on Mathematical Physics)
- [10] ROTHE, H.J.: *Lattice gauge theories.* World Scientific, Singapore, 1992. – 528 S
- [11] CREUTZ, M. ; FREEDMAN, B.: A statistical approach to quantum mechanics. In: *Ann. Phys.* 132 (1981), S. 427
- [12] CHENG, T. P. ; LI, L. F.: *Gauge theory of elementary particle physics: Problems and solutions.* Oxford, UK: Clarendon, 2000. – 306 S

LITERATURVERZEICHNIS

- [13] FALCIONI, M. ; TREVES, A.: Corrections to universal scaling for the two-dimensional nonlinear sigma model on the lattice. In: *Phys. Lett.* B159 (1985), S. 140
- [14] COLLINS, J.C.: *Renormalisation*. Cambridge University Press, 1994 (Cambridge Monographs on Mathematical Physics)
- [15] SCHECK, F.: *Mechanics, from Newton's laws to deterministic chaos*. 3rd Edition. Springer Heidelberg, 1999
- [16] LYAPUNOV, A. M.: In: *Ann. Math Studies* 17 (1907), S. 1947
- [17] KUYPERS, F.: *Klassische Mechanik*. Wiley-VCH, Weinheim, 1997
- [18] FRITZ, F.-J ; HUPPERT, B. ; WILLEMS, W.: *Stochastische Matrizen*. Springer Verlag, Berlin, 1979
- [19] GUPTA, S ; IRBACK, A. ; KARSCH, F. ; PETERSSON, B.: The acceptance probability in the Hybrid Monte Carlo Method. In: *Phys. Lett.* B242 (1990), S. 437–443
- [20] LIU, C. ; JASTER, A. ; JANSEN, K.: Liapunov exponents and the reversibility of Molecular Dynamics algorithms. In: *Nucl. Phys.* B524 (1998), S. 603–617. hep-lat/9708017
- [21] JANSEN, K. ; LIU, C.: Study of Liapunov exponents and the reversibility of Molecular Dynamics algorithms. In: *Nucl. Phys. Proc. Suppl.* 53 (1997), S. 974–976. hep-lat/9607057
- [22] CHEVALLEY, C.: *Theory of lie groups*. Princeton University Press, 1946 (Princeton Landmarks in Mathematics)
- [23] WOLFF, U.: Asymptotic scaling in 2-D $O(n)$ nonlinear sigma models. In: *Phys. Lett.* B248 (1990), S. 335–339
- [24] WOLFF, U.: Critical slowing down. In: *Nucl. Phys. Proc. Suppl.* 17 (1990), S. 93–102
- [25] SAMUEL, S. ; MARTIN, O. ; MORIARTY, K.: Seeing asymptotic freedom in $SU(3)$ lattice gauge theory. In: *Phys. Lett.* B153 (1985), S. 87
- [26] HASENFRATZ, P. ; MAGGIORE, M. ; NIEDERMAYER, F.: The exact mass gap of the $O(3)$ and $O(4)$ nonlinear sigma models in $d = 2$. In: *Phys. Lett.* B245 (1990), S. 522–528
- [27] HASENFRATZ, P. ; NIEDERMAYER, F.: The exact mass gap of the $O(N)$ sigma model for arbitrary N greater than 2 in $d = 2$. In: *Phys. Lett.* B245 (1990), S. 529–532
- [28] FREZZOTTI, R. ; URBACH, C. *Work in progress*. 2002

- [29] BROECKER, T. ; TOM DIECK, T.: *Representations of compact Lie groups*. Springer Verlag, 1985 (Graduate Texts in Mathematics 98)
- [30] VELDHUIZEN, T.: Expression templates. In: *C++ Report* Vol. 7 No. 5 (1995), S. 26
- [31] LÜSCHER, M.: A Portable high quality random number generator for lattice field theory simulations. In: *Comput. Phys. Commun.* 79 (1994), S. 100–110. hep-lat/9309020