Einführung in die Gittereichtheorien

Thomas Filk

Skript zur Vorlesung

Das Skript entstand im Zusammenhang folgender Vorlesungen und Veranstaltungen in der Vergangenheit und wird auch im Rahmen der aktuellen Vorlesung wieder erweitert und ergänzt:

Wintersemester 1992/93 an der Universität Freiburg

Wintersemester 1995/96 an der Universität Saarbrücken

Wintersemester 1996/97 an der Universität Freiburg

Wintersemester 1998/99 an der Universität Hannover

Wintersemester 2000/01 an der Universität Freiburg

Wintersemester 2006/07 - Mitteldeutsche Physik-Combo

Sommersemester 2010 - Universität Freiburg

(Letzte Änderung 4. August 2011)

Inhaltsverzeichnis

1 Einführung und Literaturhinweise **7**
- 1.1 Vorbemerkungen . . . 7
- 1.2 Historische Arbeiten . . . 8
- 1.3 Literaturhinweise . . . 9

2 Feldtheoretische Grundlagen **11**
- 2.1 Transformationsverhalten von Feldern . . . 11
- 2.2 Skalare Felder . . . 13
- 2.3 Spin 1/2 Felder . . . 13
- 2.4 Lagrange-Formalismus . . . 14
- 2.5 Abelsche Eichfelder . . . 15
- 2.6 Nicht-abelsche Eichtheorien . . . 16

3 Spontaneous symmetry breaking and the Higgs mechanism **21**
- 3.1 Spontaneous symmetry breaking . . . 21
- 3.2 The Higgs mechanism . . . 23

4 Why quantum fields? **27**
- 4.1 Particles as modes of oscillators . . . 27
- 4.2 The interpretation of the modes . . . 30
- 4.3 Second quantization of Schrödinger's equation . . . 32

5 Funktionalintegral-Formalismus **37**
- 5.1 Das Funktionalintegral in der Quantenmechanik . . . 37
- 5.2 Zeitgeordnete Vakuumerwartungswerte . . . 39
- 5.3 Euklidische Formulierung . . . 41
- 5.4 Störungstheorie . . . 42

6 Elementary axiomatic quantum field theory **45**
- 6.1 Wightman functions, τ-functions and Schwinger functions . . . 46
- 6.2 The Wightman axioms . . . 47
- 6.3 Schwinger functions and the Osterwalder-Schrader axioms . . . 49

7 Regularization and Renormalization **53**
7.1 Infinities in the integrals for Feynman diagrams . . . 54
7.2 Regularization methods . . . 56
7.2.1 Momentum cut-off . . . 56
7.2.2 Pauli-Villars . . . 56
7.2.3 Dimensional regularization . . . 57
7.2.4 Other analytical methods . . . 58
7.3 Renormalization . . . 59
7.3.1 The harmonic chain . . . 59
7.3.2 Renormalization in QFT . . . 61
7.4 Perturbative renormalizability - power counting . . . 63

8 Das freie Klein-Gordon Feld auf dem Gitter **67**
8.1 Allgemeine Vorbemerkungen . . . 67
8.1.1 Einige grundlegende und elementare Begriffsklärungen . . . 67
8.1.2 Randbedingungen . . . 68
8.2 Die gitterregularisierte Klein-Gordon Wirkung . . . 69
8.3 Der Grenzwert $L \to \infty$ . . . 71
8.4 Darstellung der 2-Punkt Funktion als Summation... . . . 71
8.5 Der Kontinuumslimes der 2-Punkt-Funktion . . . 73
8.6 Der Kontinuumslimes des freien Klein-Gordon Feldes . . . 75

9 Triviality of $(\phi^4)_4$ **79**

10 Reine Eichtheorien auf dem Gitter **81**
10.1 „Motivierende“ Vorbemerkungen . . . 81
10.2 Die Wirkung der reinen Eichtheorie... . . . 82
10.3 Das Integrationsmaß der Eichtheorie . . . 84
10.4 Observable der reinen Eichtheorie . . . 85
10.5 Die „Stringtension“ – ein Ordnungsparameter... . . . 86
10.6 Der Grenzwert großer Eichkopplung . . . 88
10.7 Experimentelle Bestimmung der Stringtension und des $q\bar{q}$-Potentials . . . 90

11 Ising-Dualities **93**
11.1 The 2-dimensional Ising model . . . 93
11.1.1 The high temperature expansion . . . 94
11.1.2 Low temperature expansion . . . 95
11.1.3 The self-duality of the 2-dimensional Ising model . . . 96
11.2 General types of $\mathbf{Z}_2$-models . . . 97
11.3 High- and low temperature expansion for generalized $\mathbf{Z}_2$ models . . . 99
11.4 Duality structure of generalized $\mathbf{Z}_2$ models . . . 102

12 Hamiltonsche Formulierung **105**
12.1 Der Transfermatrix-Formalismus . . . 105
12.1.1 Das 1-dimensionale Ising-Modell . . . 106
12.1.2 Der 1-dimensionale Anti-Ferromagnet . . . 108
12.1.3 Der allgemeine Transfermatrixformalismus . . . 109
12.2 Transfermatrix zur Quantenmechanik . . . 112
12.3 Skalare, reelle Feldtheorie . . . 113
12.4 Transfermatrix für Gittereichtheorien . . . 114
12.5 Zweidimensionale reine Gittereichtheorien . . . 115

13 Eichtheorie mit skalaren Materiefeldern **119**
13.1 Allgemeine Formulierung . . . 119
13.2 Grenzfälle im Phasendiagramm . . . 120
13.2.1 Die reine Materietheorie für $g \to 0$ . . . 120
13.2.2 Die reine Eichtheorie für $\lambda \to 0$ . . . 120
13.3 Beispiel: $\mathbf{Z}_2$-Eichtheorie mit $\mathbf{Z}_2$-Feldern . . . 121
13.4 Ordnungsparameter und Symmetriebrechung . . . 122
13.4.1 Ordnungsparameter in einer reinen Materietheorie... . . . 122
13.4.2 Ordnungsparameter in einer Eichtheorie mit Materiefeldern . . . 123
13.5 Das Phasendiagramm . . . 125

14 Das Monte-Carlo-Verfahren **127**
14.1 Vorbemerkungen . . . 127
14.2 (Diskrete) Markov-Prozesse . . . 128
14.2.1 Wahrscheinlichkeitsräume . . . 128
14.2.2 Wahrscheinlichkeitsräume von Folgen von Elementarereignissen . . . 129
14.2.3 Markov-Matrizen . . . 130
14.2.4 Detailed balance . . . 131
14.3 Monte-Carlo-Simulation . . . 132
14.3.1 Thermalisierung und Messung . . . 132
14.3.2 Das Problem der Autokorrelationen . . . 133
14.4 Spezielle Realisationen des Markov-Prozesses . . . 134
14.4.1 Das Metropolis-Verfahren . . . 134
14.4.2 Das „heat bath“– Verfahren . . . 135
14.4.3 Andere Verfahren . . . 136

15 Glueballs — Beispiel eines Massenspektrums **139**
15.1 Glueballs . . . 139
15.2 „Verschmierte“ Zustände . . . 142

16 Der Kontinuumslimes **145**
16.1 Allgemeiner Formalismus des Kontinuumslimes . . . 145
16.2 Die β-Funktion . . . 148

17 Fermionen auf dem Gitter - Teil I **153**
17.1 Die Grassmann-Algebra . . . 153
17.2 Euklidisches Funktionalintegral für Fermionen . . . 155
17.3 Die „naive“ Wirkung von Fermionen auf dem Gitter . . . 156
17.4 Chirale Fermionen . . . 159
17.5 Das No-Go-Theorem von Nielsen und Ninomiya . . . 160
17.6 Wilson-Fermionen und die Kopplung an Eichfelder . . . 161
17.7 “Staggert” Fermionen . . . 162
17.8 „Simulation“ von Fermionen . . . 163
17.9 Improved actions . . . 164

18 Fermions on the lattice - Part II **167**
18.1 Perfect Lattice Fermions - Free Case . . . 168
18.2 Domain wall and overlap fermions . . . 170
18.3 Twisted mass fermions . . . 171
18.4 Clover term . . . 171

19 Quantum field theories at finite temperature **173**
19.1 The functional integral for quantum systems at finite temperature . . . 173
19.2 Quantum fields at finite temperatures on the lattice . . . 175
19.3 Order parameters for phase transitions in lattice QCD . . . 176
19.4 The phase structure of Lattice QCD . . . 177

A1 Bündel, Zusammenhang, Krümmung **179**

A2 Gaußsche Integrale **183**
A2.1 Reelle Gaußsche Integrale . . . 183
A2.2 Komplexe Gaußsche Integrale . . . 185

A3 Summation über Wege auf beliebigen Graphen **187**

A4 Lösung des 2-dimensionalen Ising-Modells **191**
A4.1 Definition der Transfermatrix . . . 192
A4.2 Die Transfermatrix als Funktion von $\Sigma_3(n)$ und $\Sigma_1(n)$ . . . 193
A4.3 Die TfM als Funktion von Fermion-Operatoren . . . 194
A4.4 Die TfM als Funktion der Fourier-Moden... . . . 196
A4.5 Diagonalisierung von V_k . . . 197
A4.6 Das Spektrum der Transfermatrix . . . 199
A4.7 Das Ising-Modell als Theorie von Fermionen . . . 200

Literaturangaben **203**

Index **208**

Kapitel 1

Einführung und Literaturhinweise

1.1 Vorbemerkungen

In den 70er Jahren wurden Gittereichtheorien zu einem wesentlichen Bestandteil der theoretischen Hochenergiephysik. Nach Vorarbeiten von Wegner [99] formulierte Wilson [101] 1974 eine Gittertheorie, deren Kontinuumslimes einer euklidischen Version der Quantenchromodynamik entsprach. 1979 begannen Creuz, Jacobs und Rebbi [18] mit Monte-Carlo-Simulationen verschiedener Eichtheorien und untersuchten das Verhalten des Wilson-Loops als Ordnungsparameter für Confinement in einer reinen Eichtheorie. Innerhalb weniger Jahre kam es zu einer explosionsartigen Entwicklung und heute sind Monte-Carlo-Simulationen des Standardmodells neben der Störungstheorie zu einem der wichtigsten Hilfsmittel der phänomenologischen Hochenergiephysik geworden. Die Gitterformulierung einer Quantenfeldtheorie ist nicht-störungstheoretisch und erlaubt daher einen komplementären Zugang zu vielen beobachtbaren Größen, die oft störungstheoretisch nicht direkt zugänglich sind (Massenspektrum, Matrixelemente, $q\bar{q}$-Potential).

Diese Vorlesung soll eine Einführung in die theoretischen Grundlagen der Gittertheorien der Teilchenphysik geben. Insbesondere soll der Zusammenhang zwischen einer Quantenfeldtheorie (wie dem Standardmodell der Teilchenphysik) und bestimmten Gittertheorien erläutert werden. Auch wenn der Schwerpunkt der Vorlesung die theoretischen Aspekte betont, werden das Prinzip sowie die Probleme der Monte-Carlo-Simulationen behandelt.

Da diese Vorlesung bewusst keine Voraussetzungen über fundierte Kenntnisse der Quantenfeldtheorie und des Standardmodells macht, dienen die ersten zwei Kapitel einer kurzen Einführung in diese beiden Themen. Beweise wie auch eine ausführliche Behandlung der Themen fehlen. Es wird durch anschauliche Argumente, insbesondere durch Analogien zu Bekanntem, eine grobe Darstellung angestrebt. Der Rest der Vorlesung – die Formulierung der Gittertheorien etc. – ist im Wesentlich unabhängig vom Verständnis dieser ersten Kapitel. Sie sollen jedoch dazu beitragen, den Zusammenhang zu der üblichen Formulierung des Standardmodells als einer Quantenfeldtheorie besser zu verstehen.

In einem ersten vorbereitenden Kapitel (2) werden einige Grundlagen für das Verständnis des Standardmodells der Teilchenphysik behandelt. Dazu zählen folgende Begriffe: Skalare Felder,

fermionische Felder, Eichfelder (Abelsch und nicht-Abelsch), minimale Kopplung, Lagrange-Funktion einer Eichtheorie mit Materiefeldern.

Ein zweites vorbereitendes Kapitel (5.1) gibt eine Einführung in die Quantenfeldtheorie. Stichpunkte sind in diesem Zusammenhang: Funktionalintegral, zeitgeordnete Vakuumerwartungswerte, Euklidische Formulierung.

Die folgenden Themen umreißen den Stoff, der in dieser Vorlesung behandelt werden soll. Sie bilden den Kern der Theorie der Gittermodelle der Teilchenphysik:

Gittermodelle – Skalare Materiefelder auf dem Gitter – Eichfelder auf dem Gitter – Kopplung von Eichfeldern an Materiefelder – Kontinuumslimes: kritische Punkte, kritische Exponenten, Skalenlimes, Renormierungsgruppe – Observable: Wilson-Loop, $q\bar{q}$-Potential (Confinement), Massenspektrum, Matrixelemente – Methoden: Tief- und Hochtemperaturentwicklung, Monte-Carlo – Dirac-Feld auf dem Gitter. Weiterführende Spezialthemen sind: Quantenfeldtheorie bei endlichen Temperaturen - Quantengravitation auf dem Gitter.

Ziel dieser Vorlesung ist:

1. Den Zusammenhang zwischen der Teilchenphysik – allgemeiner: einer Quantenfeldtheorie – und Gittermodellen zu klären, und zwar in beide Richtungen:

   $$\text{Quantenfeldtheorie} \quad \longleftrightarrow \quad \text{Gittermodelle}$$

2. Verschiedene für die Teilchenphysik relevante Gittermodelle zu untersuchen (skalares Feld, Dirac-Feld, Eichfeld auf dem Gitter) und die physikalische Bedeutung der Erwartungswerte bestimmter Observablen herauszuarbeiten.

3. Methoden zur Berechnung von Erwartungswerten in Gittertheorien vorzustellen, insbesondere – neben systematischen Entwicklungen – auch die Monte-Carlo-Methode der Simulation einer Gittertheorie auf einem Computer.

Das Skript entstand Anfang der 1990er Jahre im Zusammenhang mit der Vorbereitung auf eine Vorlesung über dieses Thema. Die Vorlesung fand in den folgenden Jahren mehrfach statt und bei diesen Gelegenheiten wurde das Skript teilweise überarbeitet. Allerdings behandelt es den Stand der Forschung nur bis Mitte der 90er Jahre. Insbesondere auf dem Gebiet der (chiralen) Fermionen auf dem Gitter sind aber in den letzten fünfzehn Jahren wesentliche Fortschritte erzielt worden. Das Gleiche gilt für die Entwicklung auf dem Gebiet der Monte-Carlo Algorithmen. Außerdem sind aufgrund der leistungsstärkeren Computer die Daten heute wesentlich besser, und es lassen sich Simulationen auf größeren Gittern durchführen als noch vor wenigen Jahren.

Einige wichtige Bereiche der Gittereichtheorien werden überhaupt nicht behandelt. Dazu zählen: Gittereichtheorien bei endlichen Temperaturen, Gittermodelle der Gravitation, Berechnung von Formfaktoren, Verzweifungsverhältnissen, Matrixelementen etc. Aus diesen Gründen muss das Skript in mehrfacher Hinsicht als veraltet angesehen werden. Eine grundlegende Überarbeitung ist parallel zur Vorlesung im Sommersemester 2010 geplant.

1.2 Historische Arbeiten

Der folgende historische Werdegang der Gittereichtheorien dient nur einer groben Orientierung und ist alles andere als vollständig.

1925 Lenz formuliert ein einfaches 2-Zustands-Modell eines Ferromagneten, das er seinem Studenten Ising zur genaueren Untersuchung gibt [46].

1944 Lösung des 2-dimensionalen Ising-Modells ohne äußeres Magnetfeld durch Onsager [84].

1971 Wegner untersucht ($\mathbf{Z}_2$)-Eichtheorien (lokale Invarianz) auf dem Gitter sowie Ordnungsparameter, die in verallgemeinerter Form als Wilson-Loop bekannt sind [99].

1974 Formulierung einer nicht-Abelschen Eichtheorie auf dem Gitter. Neues Kriterium für Confinement im reinen Eichsektor: das Flächengesetz des Wilson-Loops [101].

1975 Hamiltonsche Formulierung einer Eichtheorie auf dem Gitter [58].

1979 Erste Monte-Carlo-Simulationen einer Eichtheorie (zunächst $\mathbf{Z_2}$, später SU(2) und SU(3) in 3 und 4 Dimensionen) auf dem Gitter durch Creutz, Jacobs und Rebbi [18]. Test verschiedener nicht-Abelscher Eichtheorien auf Confinement mit Monte-Carlo-Methoden.

1980 A. Hasenfratz und P. Hasenfratz berechnen den Zusammenhang zwischen dem so genannten Λ-Parameter in der gitterregularisierten Eichtheorie und in verschiedenen störungstheoretischen Regularisierungsschemata in der Kontinuumstheorie [42].

1981 Nicht-Abelsche Eichtheorien bei endlichen Temperaturen [72, 62, 25].

Fermionen auf dem Gitter. Problem der Spezies-Verdopplung [56].

Beweis eines No-Go-Theorems für chirale Fermionen auf dem Gitter von Nielsen und Ninomiya [82].

1981/2 Erste MC Simulationen zur Bestimmung des Glueball-Spektrums in reinen Eichtheorien [10, 8, 45].

Erste MC Simulationen zur Bestimmung von Mesonen- und Hadronenspektren in der reinen QCD [41, 71, 32].

1985/6 Erste MC-Simulationen zur Bestimmung von Matrix-Elementen [9, 70].

Gitterformulierung und MC-Simulationen der 2-dimensionalen Quantengravitation [57, 21, 3].

1998 Umgehung des Nielsen-Ninomiya Theorem und Formulierung exakter chiraler Symmetrie auf dem Gitter (siehe [81]).

1.3 Literaturhinweise

Eine immer noch gut lesbare Einführung zu den Themen „Nicht-Abelsche Eichtheorie“, „Standardmodell“, „Funktionalintegral“ etc. findet sich in Abers & Lee [1]. Eine ausführlichere Darstellung bietet jedes neuere Buch zur Quantenfeldtheorie. Ein (nicht ganz billiges) neueres Standardwerk sind die drei Bände von Weinberg [100]. Ein (nicht überall beliebtes) älteres Standardwerk ist Itzykson & Zuber [47], andere Beispiele sind das deutschsprachige Buch von Becher, Böhm & Joos [6], sowie das etwas unbekanntere, etwas anspruchsvolle Buch von Pokorski [89]. Leichter lesbar ist Ryder [92]. Weitere aktuelle Empfehlungen sind Peskin [88] und

Frampton [27]. Eine gute Einführung in die Theorie der Elementarteilchen und des Standardmodells (allerdings ohne Funktionalintegralformalismus) ist das Buch von Nachtmann [76].

Rebbi hat eine umfangreiche Sammlung von Originalarbeiten bis 1983 zum Thema „Gittertheorien“ herausgegeben [90]. Diese enthält auch eine kurze, aber gut lesbare Einführung in „Gittereichtheorien“ und „Monte-Carlo-Verfahren“. Zwei klassische Review-Artikel zum Thema „Gittertheorien“ existieren von Kogut [59, 60], ein elementarer Artikel zum gleichen Thema von Halliday [40]. Ein kurzer Review-Artikel mit Betonung des Monte-Carlo Verfahrens stammt von den „Gründern“ Creutz, Jacobs & Rebbi [19].

Ein mittlerweile in der dritten Auflage erschienenes neueres Buch mit einer guten Einführung in die verschiedenen Bereiche der Gittermodelle mit sehr ausführlicher Literaturangabe stammt von H. J. Rothe [91]. Auch Montvay und Münster haben ein gutes Buch zum aktuellen Stand der Gittereichtheorien geschrieben [74].

Artikel und Referenzen zu speziellen Themen aus dem Bereich „Gittertheorien“, „Monte-Carlo-Verfahren“ etc. finden sich in den Proceedings zu den jährlichen LATTICE-Konferenzen, erschienen in Nuclear Physics B – Proceedings Supplements.

Kapitel 2

Feldtheoretische Grundlagen des Standardmodells

In diesem Kapitel werden die wichtigsten Grundlagen zur Feldtheorie zusammengefasst. Dabei beschränken wir uns auf die Aspekte, die im Zusammenhang mit dem Standardmodell von Bedeutung sind. Dazu zählen die feldtheoretische Beschreibung von skalaren Feldern, von Spin-$\frac{1}{2}$-Feldern, von Eichfeldern sowie den Kopplungen dieser Felder untereinander. Außerdem wird der Lagrange-Formalismus eingeführt. Anhang A1 enthält eine kurze Einführung in die mathematischen Grundlagen von Eichtheorien. Diese Erläuterungen sollen das Verständnis vertiefen, sind aber für die weiteren Kapitel nicht wesentlich.

Im Folgenden werden wir (mit wenigen Ausnahmen) Einheiten benutzen, in denen das Wirkungsquantum und die Lichtgeschwindigkeit den Wert $\hbar = c = 1$ haben. Wenn nicht anders angegeben, nehmen griechische Indizes $\mu, \nu, \ldots$ die Werte 0,1,2,3 an (in der euklidischen Formulierung auch manchmal 1,2,3,4), lateinische Indizes $i, j, \ldots$ (sofern sie sich auf räumliche Komponenten beziehen) die Werte 1,2,3. Es gilt die Summenkonvention: Über doppelt auftretende Indizes (einmal unten, einmal oben) wird summiert. Die invariante Bilinearform des Minkowski-Raumes ist $g_{\mu\nu} = \mathrm{diag}\,(-1, 1, 1, 1)$. (Diese etwas ungewöhnliche Wahl wird hier vorgezogen, da der Übergang zu imaginärer Zeit direkt auf die euklidische Metrik führt.) Diese Bilinearform bzw. ihr Inverses erlaubt das „Runter- und Raufziehen“ von Indizes:

$$a_\mu \;=\; g_{\mu\nu} a^\nu \qquad\qquad a^\mu \;=\; g^{\mu\nu} a_\nu \,.$$

2.1 Transformationsverhalten von Feldern

Ohne Berücksichtigung der Gravitation sind die fundamentalen Naturgesetze invariant unter Poincaré-Transformationen

$$x \;\to\; x' \;=\; \Lambda x + a \,.$$

Hierbei beschreibt $a \in \mathbb{R}^4$ eine Raum-Zeit-Translation und Λ eine Lorentz-Transformation, d.h. eine lineare Abbildung, unter der die Minkowski-Metrik invariant ist:

$$g \;=\; \Lambda g \Lambda^T \qquad \text{bzw.} \qquad g_{\mu\nu} \;=\; \Lambda_\mu{}^\alpha \Lambda^\beta{}_\nu g_{\alpha\beta}\,.$$

Aus diesem Grunde fordern wir, dass die Lösungen der Feldgleichungen, die wir im Folgenden betrachten werden, invariant unter der Überlagerungsgruppe der Poincaré-Gruppe sind. Dies bedeutet, dass es eine geeignete (endlich-dimensionale) Darstellung $S_i^j(\Lambda)$ der Überlagerungsgruppe der Lorentz-Gruppe gibt, so dass mit jeder Lösung $\Psi_i,\ i = 1, \ldots, N$ der Feldgleichungen (formal $D[\Psi] = 0$) auch $\Psi_i^{\Lambda,a}$ eine Lösung der Feldgleichungen ist, d.h.

$$D[\Psi] \;=\; 0 \quad\longrightarrow\quad D[\Psi^{\Lambda,a}] \;=\; 0\,, \tag{2.1}$$

wobei

$$\Psi_i^{\Lambda,a}(x) \;:=\; S_i^j(\Lambda)\Psi_j(\Lambda^{-1}(x-a))\,. \tag{2.2}$$

– Wir haben hier einen „aktiven“ Standpunkt vertreten: Ein fester Beobachter beschreibt in seinem System die Felder wie auch die Feldgleichungen durch Koordinaten x. Die Felder Ψ und $\Psi^{\Lambda,a}$ sind physikalisch verschieden, d.h. nehmen bei denselben Raum-Zeit-Ereignissen unterschiedliche Werte an.

– Folgt die Feldgleichung aus einem Wirkungsprinzip (s.u.) so kann man eine verschärftere Version von Invarianz fordern. Eine Theorie ist invariant unter Poincaré-Transformationen, wenn für die Wirkung der Felder $S[\Psi] = S[\Psi^{\Lambda,a}]$ gilt. Eine noch schärfere Bedingung, nämlich die Invarianz der Feldgleichungen selber: $D[\Psi] = D[\Psi^{\Lambda,a}]$, ist im Allgemeinen nicht erfüllt.

– In manchen Fällen ist es sinnvoll, die Lorentz-Gruppe auch auf den Operator D wirken zu lassen, beispielsweise auf Koeffizientenfunktionen oder Quellterme (vgl. die inhomogenen Maxwell-Gleichungen (2.16). In diesem Fall muss auch der Strom j^μ mittransformiert werden, da die Poincaré-Invarianz durch die Quellterme bzw. Koeffizientenfunktionen gebrochen wird. Ein solcher Fall tritt auf, wenn die Gleichungen einen Teil der Materie nur als phänomenologischen Hintergrund beschreiben. Zur vollständingen Invarianz muss im Fall der Maxwell-Gleichungen der Strom j^μ durch die entsprechenden Skalar- oder Spinorfelder ausgedrückt werden.

– Wir haben oben nur die Invarianz unter der Überlagerungsgruppe der Poincaré-Gruppe verlangt, nicht die Invarianz unter der Poincaré-Gruppe selber. Der Grund ist derselbe wie in der Quantenmechanik, wo man weniger an den Darstellungen der Drehgruppe SO(3) interessiert ist als an den Darstellungen der Überlagerungsgruppe SU(2): Die physikalischen Zustände sind Strahlen im Hilbert-Raum, d.h., die Invarianz der Zustände verlangt nur eine projektive Darstellung der Gruppe. Diese erhält man gerade aus den Darstellungen der Überlagerungsgruppe im Hilbert-Raum.

Ist die Darstellung $S(\Lambda)$ irreduzibel, so entspricht ihr ein bestimmter Spin s. Die physikalisch relevanten, endlich-dimensionalen irreduziblen Darstellungen der Poincaré-Gruppe lassen sich durch zwei Größen charakterisieren: $m^2 \geq 0$ und s. Hierbei ist m^2 der Eigenwert des Casimir-Operators $H^2 - \boldsymbol{P}^2 = M^2$; er hat die Interpretation eines Massenquadrats. $s = 0, \frac{1}{2}, 1, \frac{3}{2}, \ldots$ ist der Spin. Damit das Feld ψ_i ein freies Teilchen mit einer festen Masse m beschreibt, muss jede Komponente der so genannten Klein-Gordon-Gleichung genügen:

$$(-\partial_\mu \partial^\mu + m^2)\psi_i(x) \;=\; 0\,. \tag{2.3}$$

Die Fourier-Transformierte des Feldes ψ_i hat also ihren Träger auf der so genannten Massenschale $p_\mu p^\mu = m^2$. In einer Quantenfeldtheorie wird man für die (quasi-freien) asymptotischen Teilchenzustände eine entsprechende Bedingung fordern.

Lösungen der Klein-Gordon-Gleichung werden im nicht-relativistischen Grenzfall durch Lösungen einer Schrödinger-Gleichung angenähert. Das folgende Argument ist jedoch eher

heuristisch und keine strenge Ableitung: Identifiziert man nach einer Fourier-Transformation $\omega = E/\hbar$ und $\boldsymbol{k} = \boldsymbol{p}/\hbar$, so folgt aus (2.3) die Beziehung:

$$\begin{aligned} 0 &= (E^2 - c^2p^2 - c^4m^2)\hat{\psi}(E,\boldsymbol{p}) \\ &= (E + \sqrt{c^2p^2 + c^4m^2}))(E - \sqrt{c^2p^2 + c^4m^2}))\hat{\psi}(E,\boldsymbol{p})\,. \end{aligned}$$

Die allgemeinste Lösung ist also durch

$$\hat{\psi}(E,\boldsymbol{p}) \;=\; g^+(E,\boldsymbol{p})\;\delta(E - \sqrt{c^2p^2 + c^4m^2}) \;+\; g^-(E,\boldsymbol{p})\;\delta(E + \sqrt{c^2p^2 + c^4m^2})$$

gegeben, für beliebige Funktionen $g^{\pm}(E,\boldsymbol{p})$. Diese Aufspaltung entspricht der Zerlegung der allgemeinsten Lösung nach positiven und negativen Energiebeiträgen. Haben die Funktionen $g^{\pm}$ ihren Träger bei $p/mc \ll 1$, so ergibt sich näherungsweise

$$\hat{\psi}(E,\boldsymbol{p}) \;\approx\; g^+(E,\boldsymbol{p})\;\delta(E - mc^2 - \frac{p^2}{2m}) \;+\; g^-(E,\boldsymbol{p})\;\delta(E + mc^2 + \frac{p^2}{2m})\,.$$

Die hyperbolischen Massenschalen lassen sich also für $p/mc \ll 1$ durch parabolische approximieren. Im nicht-relativistischen Grenzfall erhält man daher Lösungen der Schrödinger-Gleichung sowie der konjugierten Gleichung. Eine konsistente Behandlung der Beiträge mit negativer Energie wird allerdings erst im Rahmen einer Quantisierung möglich.

2.2 Skalare Felder

Ein freies, skalares Teilchen der Masse m wird durch die Lösungen der Klein-Gordon-Gleichung beschrieben:

$$(-\partial_\mu\partial^\mu + m^2)\varphi(x) \;=\; 0\,. \tag{2.4}$$

Ein reelles Feld beschreibt ein neutrales Teilchen, ein komplexes Feld ein geladenes Teilchen.

Die Klein-Gordon-Gleichung lässt sich um nichtlineare Terme erweitern, z.B.

$$(-\partial_\mu\partial^\mu + m^2)\varphi(x) \;+\; g\varphi^3(x) \;=\; 0\,. \tag{2.5}$$

Man kann sich leicht überzeugen, dass mit jeder Lösung $\varphi(x)$ dieser Gleichung auch $\varphi^{\Lambda,a}(x) = \varphi(\Lambda^{-1}(x-a))$ Lösung ist, φ also einem Spin-0-Feld entspricht. Nach einer Quantisierung dieser Gleichung beschreiben nicht-lineare Terme eine Selbstwechselwirkung der Teilchen. Im Standardmodell der Teilchenphysik ist nur das Higgs-Teilchen ein skalares Teilchen.

2.3 Spin 1/2 Felder

Ein freies Spin-$\frac{1}{2}$ Teilchen der Masse m wird relativistisch durch die Dirac-Gleichung beschrieben:

$$(\mathrm{i}\partial\!\!\!/ + m)\psi(x) \;=\; 0\,. \tag{2.6}$$

Dabei ist

$$\partial\!\!\!/ \;=\; \gamma_\mu\partial^\mu \tag{2.7}$$

$$\text{mit}\quad \{\gamma_\mu,\gamma_\nu\} \;=\; -2\,g_{\mu\nu} \qquad \text{Dirac}-\text{Matrizen}\;\;(4\times 4)\,, \tag{2.8}$$

ψ ist ein 4-komponentiges (für geladene Teilchen komplexes) Feld. Eine Lösung der Dirac-Gleichung erfüllt auch

$$\begin{aligned}(-\mathrm{i}\partial\!\!\!/ + m)(\mathrm{i}\partial\!\!\!/ + m)\psi(x) &= (\partial_\mu\partial_\nu\gamma^\mu\gamma^\nu + m^2)\psi(x) \\ &= (\tfrac{1}{2}\partial_\mu\partial_\nu\{\gamma^\mu,\gamma^\nu\} + m^2)\psi(x) \\ &= (-\partial_\mu\partial^\mu + m^2)\psi(x) \;=\; 0\,,\end{aligned}$$

d.h. beschreibt ein Feld zur Masse m. Außerdem erfüllen die Lösungen das Transformationsgesetzt (2.2), wobei $S_i^j(\Lambda)$ der Spin-$\frac{1}{2}$-Darstellung der Lorentz-Gruppe entspricht.

Im nicht-relativistischen Grenzfall werden die Lösungen der Dirac-Gleichung zu Lösungen der Schrödinger-Gleichung, d.h., man erhält vier entkoppelte Gleichungen, die jeweils das Teilchen (Elektron) und sein Antiteilchen (Positron) mit zwei Spinkomponenten beschreiben. Nach Ankopplung eines Eichpotentials (siehe Abschnitt 2.5) entkoppeln im nicht-relativistischen Grenzfall zwar immer noch die Elektron- und Positronanteile, ansonsten erhält man jedoch die Pauli-Gleichung für nicht-relativistische Spin-1/2-Teilchen (zwei-komponentige Spinoren) im Magnetfeld.

Die Spin-1/2-Teilchen des Standardmodells sind die Quarks (Up, Down, Strange, Charm, Top, Bottom mit jeweils drei Colour-Freiheitsgraden) und die Leptonen (Elektron, Muon, Tau mit den zugehörigen Neutrinos).

2.4 Lagrange-Formalismus

Die Feldgleichungen der Physik lassen sich im allgemeinen aus einem Wirkungsfunktional (d.h. einer Abbildung, die jedem Feld eine reelle Zahl zuordnet) durch ein Variationsprinzip erhalten. Die Lokalität der Wechselwirkungen äußert sich dadurch, dass sich die Wirkung als Raum-Zeit-Integral über eine Lagrange-Dichte schreiben lässt, wobei die Lagrange-Dichte im Allgemeinen eine Funktion der Felder und ihrer ersten Ableitungen ist:

$$S[\varphi] \;=\; \int \mathrm{d}^4x\; \mathcal{L}(\varphi, \partial_\mu\varphi)\,.$$

Die Bedingung, dass die erste Variation von S verschwindet, führt für jede Komponente des Feldes auf die Feldgleichungen:

$$\partial_\mu \frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi(x))} \;-\; \frac{\partial\mathcal{L}}{\partial\varphi(x)} \;=\; 0\,. \tag{2.9}$$

So folgt beispielsweise die Klein-Gordon-Gleichung aus folgendem Wirkungsfunktional:

$$S[\varphi] \;=\; \int \mathrm{d}^4x\; \tfrac{1}{2}[-(\partial_\mu\varphi)(\partial^\mu\varphi) - m^2\varphi^2] \;=\; \int \mathrm{d}^4x\; \mathcal{L}(\varphi(x), \partial_\mu\varphi(x)) \tag{2.10}$$

mit der Lagrange-Dichte

$$\mathcal{L}(\varphi(x), \partial_\mu\varphi(x)) \;=\; \tfrac{1}{2}[-(\partial_\mu\varphi(x))(\partial^\mu\varphi(x)) - m^2\varphi(x)^2]\,. \tag{2.11}$$

Wechselwirkungsterme erhält man durch Addition höherer Potenzen in φ, z.B.

$$\mathcal{L}_{\mathrm{int}} \;=\; \frac{g}{4}\varphi(x)^4\,. \tag{2.12}$$

Die Dirac-Gleichung ergibt sich aus einer Lagrange-Dichte

$$\mathcal{L} \;=\; \bar{\psi}(x)(\mathrm{i}\partial\!\!\!/ + m)\psi(x)\,, \tag{2.13}$$

mit $\bar{\psi} = \psi^{+}\gamma^{0}$. (Man beachte, dass $\partial\!\!\!/$ eine 4×4-Matrix ist, deren Elemente Differentialoperatoren sind. m steht eigentlich für $m \cdot \mathbb{1}$.)

2.5 Abelsche Eichfelder

Im Standardmodell werden die starke und elektroschwache Wechselwirkung durch eine Eichtheorie beschrieben. Vorbild ist die Elektrodynamik.

In der kovarianten Formulierung der Elektrodynamik fasst man die Ladungsdichte $\rho(x)$ und die Stromdichte $\boldsymbol{j}(x)$ zu einem *Vierer-Strom* zusammen

$$j^{\mu} \;\sim\; (\rho, \boldsymbol{j})\,. \tag{2.14}$$

Die elektrische und magnetische Feldstärke ($\boldsymbol{E}$ und $\boldsymbol{B}$) werden in einem *Feldstärketensor* zusammengefasst:

$$F_{\mu\nu} \;\sim\; \begin{pmatrix} 0 & -E_1 & -E_2 & -E_3 \\ E_1 & 0 & -B_3 & B_2 \\ E_2 & B_3 & 0 & -B_1 \\ E_3 & -B_2 & B_1 & 0 \end{pmatrix}\,. \tag{2.15}$$

Damit lassen sich die Maxwell-Gleichungen in kovarianter Form schreiben:

$$\partial_{\mu}F^{\mu\nu} \;=\; j^{\nu} \qquad \text{und} \qquad \epsilon_{\sigma\mu\nu\lambda}\partial^{\mu}F^{\nu\lambda} \;=\; 0\,. \tag{2.16}$$

Die rechte (homogene) Gleichung wird durch den Ansatz

$$F_{\mu\nu} \;=\; \partial_{\mu}A_{\nu} - \partial_{\nu}A_{\mu} \qquad\qquad A_{\mu} \;=\; (\phi, \boldsymbol{A}) \quad \text{Eichpotential} \tag{2.17}$$

gelöst. $A_{\mu}(x)$ ist durch den Feldstärketensor allerdings nur bis auf ein Gradientenfeld festgelegt, d.h., $A_{\mu}(x)$ und $A_{\mu}(x) + \partial_{\mu}\alpha(x)$ führen für beliebiges $\alpha(x)$ auf denselben Feldstärketensor. Dies erlaubt die Wahl einer Eichung (z.B. der Lorentz-Eichung: $\partial_{\mu}A^{\mu} = 0$) für das Eichpotential. In dieser Eichung wird aus der inhomogenen Maxwell-Gleichung: $\Box A^{\mu} \;=\; j^{\mu}$.

Die Dirac-Gleichung für Fermionen in Gegenwart eines Eichpotentials A_{μ} erhält man nach der Vorschrift der „minimalen Kopplung“ $\partial_{\mu} \to \mathrm{D}_{\mu} = \partial_{\mu} + \mathrm{i}eA_{\mu}$:

$$(\mathrm{i}\partial\!\!\!/ - eA_{\mu}\gamma^{\mu} + m)\psi(x) \;=\; (\mathrm{i}\not{\mathrm{D}} + m)\psi(x) \;=\; 0\,. \tag{2.18}$$

Aus der Invarianz unter Phasentransformationen $\psi(x) \to \mathrm{e}^{\mathrm{i}e\alpha}\psi(x)$ folgt die Ladungserhaltung und der zugehörige Strom

$$j^{\mu} \;=\; \bar{\psi}(x)\gamma^{\mu}\psi(x)\,. \tag{2.19}$$

Die Dirac-Gleichung sowie die inhomogene Maxwell-Gleichung lassen sich aus der Lagrange-Dichte

$$\mathcal{L}[A_{\mu}, \bar{\psi}, \psi] \;=\; -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} \;+\; \bar{\psi}(\mathrm{i}\not{\mathrm{D}} + m)\psi \tag{2.20}$$

herleiten. (Die homogene Maxwell-Gleichung ist in diesem Formalismus keine Bewegungsgleichung, sondern ergibt sich aus der Darstellung des elektromagnetischen Feldes durch das Eichpotential (Gl. 2.17)). Diese Theorie ist invariant unter lokalen Eichtransformationen

$$A_\mu(x) \;\to\; A_\mu(x) + \partial_\mu \alpha(x) \qquad\qquad \psi(x) \;\to\; \mathrm{e}^{\mathrm{i}e\alpha(x)}\psi(x) \tag{2.21}$$

mit einem beliebigen skalaren Feld α. Die Elektrodynamik ist eine abelsche Eichtheorie, da die Eichtransformationen eine abelsche Gruppe ($U(1)$) bilden. Das Teilchen, das nach einer Quantisierung durch das Feld A_μ beschrieben wird, ist das Photon.

Ein kovariant konstantes Feld erfüllt die Gleichungen

$$\mathrm{D}_\mu \varphi \;=\; 0 \qquad \text{oder} \qquad \partial_\mu \varphi \;=\; -\,\mathrm{i}eA_\mu \varphi\,. \tag{2.22}$$

Solch ein Gleichungssystem hat im Allgemeinen keine eindeutigen Lösungen. Eine notwendige (und lokal auch hinreichende) Bedingung für die Existenz eindeutiger Lösungen ist

$$\begin{aligned} \partial_\mu\partial_\nu\varphi &= \partial_\nu\partial_\mu\varphi \\ \Rightarrow \qquad (\partial_\mu A_\nu - \partial_\nu A_\mu)\varphi &= F_{\mu\nu}\varphi \;=\; 0\,. \end{aligned} \tag{2.23}$$

Ist diese Bedingung nicht erfüllt, kann man im Allgemeinen nur die kovariante Konstanz entlang eines Weges $\mathcal{C}$ fordern. Die Lösung dieses Problems ist

$$\varphi(y;\mathcal{C}) \;=\; \exp\left(-\mathrm{i}e \oint_x^y dx^\mu A_\mu\right)\varphi(x)\,. \tag{2.24}$$

Der „Paralleltransporter“

$$U(x,y;\mathcal{C}) \;=\; \exp\left(-\mathrm{i}e \oint_x^y dx^\mu A_\mu\right) \tag{2.25}$$

ist genau dann $\mathcal{C}$ unabhängig, wenn

$$U(x,x;\mathcal{C}) \;=\; 1 \tag{2.26}$$

für jeden geschlossenen Weg $\mathcal{C}$. Dies ist eine hinreichende globale Bedingung für eine eindeutige Lösung von (2.22). Man beachte, dass der Paralleltransporter ein Element der Invarianzgruppe ist: $U(x,y;\mathcal{C}) \in U(1)$.

2.6 Nicht-abelsche Eichtheorien

Sei $\psi(x)$ ein N-komponentiges (komplexes) Feld ($N = 3$ in der Quantenchromodynamik entspricht den Colour-Freiheitsgraden der Quarks)

$$\psi(x) \;=\; \begin{pmatrix} \psi_1(x) \\ \vdots \\ \psi_N(x) \end{pmatrix} \qquad \psi^+(x) \;=\; (\psi_1^*(x), \ldots, \psi_N^*(x))\,. \tag{2.27}$$

Das Produkt

$$\psi^+(x)\psi(x) \;=\; \sum_{i=1}^{N} \psi_i^*(x)\psi_i(x) \tag{2.28}$$

ist invariant unter den Transformationen

$$\psi(x) \to U\psi(x) \quad , \quad \psi^+(x) \to \psi^+(x)U^+ \qquad\qquad (U \in \mathrm{SU(N)})\,. \tag{2.29}$$

(Für reelle Felder erhält man entsprechend eine Invarianz unter orthogonalen Transformationen, d.h. U∈SO(N). Ist das Produkt allgemeiner von der Form $\psi_i^+(x)A_{ij}\psi_j(x)$, so erhält man als Invarianzgruppe die Menge der Transformationen, die A_{ij} invariant lassen, d.h. für die $U^+AU = A$.)

Bei einer lokalen Transformation $\psi(x) \to U(x)\psi(x)$ folgt für die partielle Ableitung des Feldes

$$\partial_\mu\psi(x) \;\;\to\;\; \partial_\mu(U(x)\psi(x)) \;=\; U(x)[\partial_\mu + U^{-1}(x)(\partial_\mu U(x))]\psi(x)\,. \tag{2.30}$$

Analog zur abelschen Eichtheorie definieren wir eine kovariante Ableitung

$$\mathrm{D}_\mu\psi(x) \;=\; (\partial_\mu + \mathrm{i}gA_\mu)\psi(x) \qquad\qquad A_\mu \simeq (N\times N) - \text{Matrix} \tag{2.31}$$

von der wir verlangen:

$$\mathrm{D}_\mu\psi(x) \;\;\to\;\; \mathrm{D}_\mu^U\psi^U(x) \;\stackrel{!}{=}\; U(x)\mathrm{D}_\mu\psi(x) \tag{2.32}$$

oder

$$\begin{aligned}(\partial_\mu + \mathrm{i}gA_\mu^U)U(x)\psi(x) \qquad\qquad & (2.33)\\ = \; U(x)[\partial_\mu + U^{-1}(x)\partial_\mu U(x) + \mathrm{i}gU^{-1}(x)A_\mu^U U(x)]\psi(x)&\\ \stackrel{!}{=} \; U(x)[\partial_\mu + \mathrm{i}gA_\mu]\psi(x)\,. & (2.34)\end{aligned}$$

Für das „Eichpotential“ folgt damit als Transformationsgesetz

$$\mathrm{i}gA_\mu \;=\; U^{-1}(x)\partial_\mu U(x) + \mathrm{i}gU^{-1}(x)A_\mu^U U(x)$$

bzw.

$$A_\mu^U(x) \;=\; U(x)A_\mu(x)U^{-1}(x) + \frac{\mathrm{i}}{g}(\partial_\mu U(x))U^{-1}(x)\,. \tag{2.35}$$

Man vergleiche dieses Transformationsgesetz für das nicht-abelsche Eichpotential mit dem entsprechenden Gesetz im abelschen Fall (2.21).

Außerdem ist aus (2.35) ersichtlich, dass A_μ ein Element der Lie-Algebra der Invarianzgruppe G ist. Zu jedem (unabhängigen) Generator einer Invarianzgruppe gibt es daher ein A-Feld, das in einer quantisierten Theorie wiederum ein Teilchen beschreibt. Den acht Generatoren der Gruppe $SU(3)$ entsprechen in der Quantenchromodynamik die Gluonen. Die vier Generatoren der Gruppe $SU(2)\times U(1)$ in der elektro-schwachen Theorie entsprechen (in einer bestimmten Basis) dem Photon (γ), dem Z^0-Teilchen und den geladenen $W^\pm$-Bosonen.

Eichinvariante Lagrange-Dichten für skalare (φ) oder Spin-1/2 (ψ) Felder sind also

$$\mathcal{L}_\psi \;=\; \bar\psi \not{\mathrm{D}}\psi \qquad \text{bzw.} \qquad \mathcal{L}_\varphi \;=\; (\mathrm{D}^\mu\varphi)^+(\mathrm{D}_\mu\varphi)\,. \tag{2.36}$$

Die Bedingung kovarianter Konstanz eines N-komponentigen Feldes ψ lautet

$$\mathrm{D}_\mu\psi \;=\; 0 \qquad \text{bzw.} \qquad \partial_\mu\psi \;=\; \mathrm{i}gA_\mu\psi\,, \tag{2.37}$$

wobei allerdings zu beachten ist, dass nun A_μ eine Matrix und ψ ein Vektor ist. Die Konsistenzbedingung einer (lokal) eindeutigen Lösung führt in diesem Fall auf die Gleichung

$$F_{\mu\nu} \;:=\; \partial_\mu A_\nu - \partial_\mu A_\nu + \mathrm{i}g[A_\mu, A_\nu] \;=\; 0\,. \tag{2.38}$$

Dabei haben wir den nicht-abelschen Feldstärketensor (in geometrischer Sprechweise sagt man auch „Krümmungstensor“, siehe Anhang A1) $F_{\mu\nu}$ definiert (ebenfalls eine $N \times N$-Matrix). Es lässt sich leicht zeigen, dass sich der Krümmungstensor unter einer lokalen Transformation der Eichfelder (2.35) wie

$$F^U_{\mu\nu} \;=\; U(x)F_{\mu\nu}U^{-1}(x) \tag{2.39}$$

transformiert. Eine eichinvariante Lagrange-Funktion für die reine Eichtheorie ist somit:

$$\mathcal{L}_{\text{gauge}} \;=\; -\frac{1}{4}\mathrm{tr}\; F_{\mu\nu}F^{\mu\nu}\,. \tag{2.40}$$

Die Lösung der Gleichung kovarianter Konstanz entlang eines Weges $\mathcal{C}$ ist nun schwieriger, da die Eichpotentiale $A_\mu(x)$ zu verschiedenen Raum-Zeit-Punkten nicht kommutieren müssen. Infinitesimal lässt sich schreiben

$$\psi(x + \mathrm{d}x) \;=\; \psi(x) + \mathrm{i}gA_\mu(x)\tfrac{\mathrm{d}x^\mu}{\mathrm{d}s}\mathrm{d}s\;\psi(x) \;+\; \ldots \tag{2.41}$$

$$\;=\; (\mathbb{1} + \mathrm{i}gA_\mu(x)\tfrac{\mathrm{d}x^\mu}{\mathrm{d}s}\mathrm{d}s)\;\psi(x)\,. \tag{2.42}$$

Sei x_i $(i = 1, \ldots N)$ eine Unterteilung des Weges $\mathcal{C}$ in $N+1$ Abschnitte, so erhält man formal:

$$\psi(y) \;=\; \lim_{N\to\infty}\prod_{i;x_i\in\mathcal{C}}^{N}\left(\mathbb{1} + \mathrm{i}gA_\mu(x_i)\;\frac{\mathrm{d}x^\mu}{\mathrm{d}s}\bigg|_{x_i}\Delta s\right)\;\psi(x) \tag{2.43}$$

$$\;\simeq\; \mathcal{P}\left(\exp\;\mathrm{i}g\oint_x^y A_\mu\mathrm{d}x^\mu\right)\;\psi(x)\,. \tag{2.44}$$

Dabei bezeichnet $\mathcal{P}(\ldots)$ das so genannte weggeordnete Produkt. Wiederum ist

$$U(x, y; \mathcal{C}) \;=\; \mathcal{P}\left(\exp\;\mathrm{i}g\oint_x^y A_\mu\mathrm{d}x^\mu\right) \;\in\; SU(N) \tag{2.45}$$

ein Element der Invarianzgruppe. Der Paralleltransporter erlaubt die kovariant konstante „Verschiebung“ eines Feldes entlang eines Weges $\mathcal{C}$. Insbesondere ist das Produkt

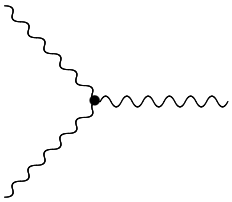
$$\psi^+(y)U(x, y; \mathcal{C})\psi(x) \tag{2.46}$$

invariant unter Eichtransformationen.

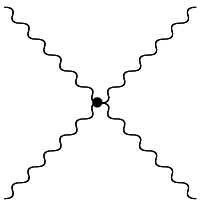
Die Lagrange-Dichte einer nicht-abelschen Eichtheorie mit skalaren und fermionischen Feldern enthält also typischerweise Terme der folgenden Art:

$$\begin{aligned}\mathcal{L} \;=\; & -\frac{1}{4}\mathrm{tr}\; F_{\mu\nu}F^{\mu\nu} \;+\; \psi^+\not{\mathrm{D}}\psi \;+\; m\psi^+\psi \\ & + \frac{1}{2}|\mathrm{D}_\mu\varphi|^2 \;+\; \frac{1}{2}M^2\varphi^+\varphi \;+\; \frac{\lambda}{4!}(\varphi^+\varphi)^2 \;+\ldots\,.\end{aligned}$$

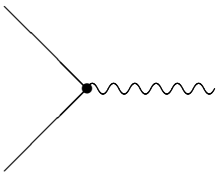
Damit werden im Rahmen einer quantisierten Theorie folgende Wechselwirkungen beschrieben:



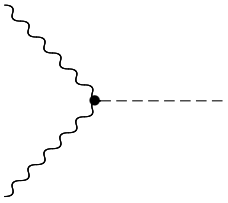
$\mathrm{i}g\mathrm{tr}\ (\partial_\mu A_\nu)[A_\mu, A_\nu]$
3 Pkt. W.W. von Eichteilchen



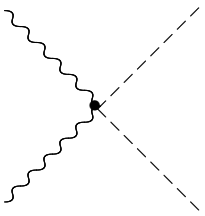
$g^2\mathrm{tr}\ ([A_\mu, A_\nu])^2$
4 Pkt. W.W. von Eichteilchen



$g\psi^+ A_\mu \gamma^\mu \psi$
W.W. von Eichteilchen mit Fermionen



$\mathrm{i}g(A_\mu\varphi_0)^+(A^\mu\varphi) + \mathrm{h.c.}$
3 Pkt. W.W. von Eichteilchen mit Skalaren
(bei spontaner Symmetriebrechung)

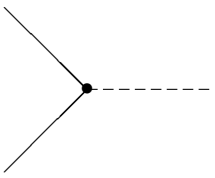


$g^2(A_\mu\varphi)^+(A^\mu\varphi)$
4 Pkt. W.W. von Eichteilchen mit Skalaren

Man beachte, dass die Kopplungskonstante in allen Fällen dieselbe ist. Dies folgt aus der Eichinvarianz der Theorie. Experimentell ergibt sich dadurch die Möglichkeit, die Wechselwirkung

der Elementarteilchen auf „Eichinvarianz“ zu prüfen.

Neben den oben aufgezählten Wechselwirkungen von Eichteilchen mit Materieteilchen gibt es im Standardmodell noch Wechselwirkungen von Fermionen mit Skalaren (Yukawa-Kopplung):



$g'(\psi^+\psi)\varphi$
Yukawa-Kopplung

In diesem Fall ist die Kopplungskonstante g' unabhängig von der Eichkopplung g. Nach einer Symmetriebrechung für das Higgs-Feld, d.h., wenn das Higgs-Feld einen nicht-verschwindenden Vakuumerwartungswert annimmt, liefert die Yukawa-Kopplung einen Massenterm für die Fermionen und die 4-Punkt-Wechselwirkung der Eichfelder mit dem Higgs-Teilchen einen Massenterm für die Eichteilchen $(Z^0, W^\pm)$.

Kapitel 3

Spontaneous symmetry breaking and the Higgs mechanism

An important ingredient of the standard model of particle physics is the Higgs mechanism. The main physical reason for the necessity of this mechanism is that some of the gauge particles (the Ws and the Z bosons) are massive. However, the naive gauge action requires massless gauge bosons, and the addition of an explicit mass term would break gauge invariance. In this case, not only the symmetry would be gone, but the QFT of an interacting massive spin-1 particle is not perturbatively renormalizable.

The Higgs mechanism - basically the interaction of a (massless) gauge particle with a scalar field which undergoes spontaneous symmetry breaking - allows to formulate a theory which is gauge invariant and for which (some of) the gauge bosons nevertheless aquire a mass. Furthermore, by coupling the Higgs field to fermions one can also give mass to initially massless fermionic particles. This mechanism is employed in the standard model. The Higgs mechanism predicts additional (massive) scalar particles, the Higgs bosons. To find the Higgs boson of the standard model is one of the goals of the experiments made at the LHC (mainly ATLAS and CMS).

3.1 Spontaneous symmetry breaking

Spontaneous symmetry breaking is a prerequisite for the Higgs mechanism. Therefore, we will first discuss the mechanism of symmetry breaking. In the next section we will explain what happens when a particle with spontaneous symmetry breaking is coupled to a gauge field.

As a simple example we will consider a theory which has a SO(2) invariance, i.e. a rotational invariance in a 2-dimensional plane. The rotational invariant Lagrange density of a 2-component scalar field is

$$\begin{aligned} L(\varphi(x), \partial_\mu \varphi(x)) &= \frac{1}{2}\Big((\partial_\mu \varphi_1(x))(\partial^\mu \varphi_1(x) + (\partial_\mu \varphi_2(x))(\partial^\mu \varphi_2(x))\Big) \\ &\quad -V(\varphi_1(x)^2 + \varphi_2(x)^2)\,. \end{aligned} \tag{3.1}$$

The "potential" $V(\varphi)$ usually contains a mass term as well as interaction terms. We now assume that the quadratic term has the "wrong" sign, i.e.:

$$V(x^2) = -\frac{\mu}{2}x^2 + \frac{1}{4}gx^4\,. \tag{3.2}$$

For the following we define

$$|\varphi(x)| = \sqrt{\varphi_1(x)^2 + \varphi_2(x)^2}\,. \tag{3.3}$$

The classical ground state, i.e. the field configuration with minimal energy, would be a static constant field which satisfies:

$$|\varphi_{\min}(x)| = \varphi_0 = \sqrt{\frac{\mu}{g}}\,. \tag{3.4}$$

Note that this condition does not specifiy the field configuration completely. Due to rotational invariance, this constant value can be distributed among the two components of the field. One particular choice is

$$\varphi_1(x) = \varphi_0 \quad , \quad \varphi_2(x) = 0\,. \tag{3.5}$$

The possible (classical) ground states are distributed around a circle of radius φ_0. Note, however, that whatever choice the system makes, the symmetry is broken by this state. Therefore, the ground state does not have the same symmetry as the theory, and in this case one speaks of spontaneous symmetry breaking. The mechanism is essentially the same as the one which is used in statistical mechanics to describe a phase transition in the framework of Landau's theory (apart from the additional feature that the symmetry which is broken is a continuous symmetry).

If we expand the action around the ground state we should be able to read off from the quadratic part the kind of particles which are described by field fluctuations around this ground state. So we write:

$$\varphi_1(x) = \varphi_0 + \phi_1(x) \quad \text{and} \quad \varphi_2(x) = \phi_2(x)\,, \tag{3.6}$$

and insert this into the Lagrange function. The partial derivatives are the same for the old and new fields, the interesting part is the potential:

$$\begin{aligned} V(\phi_1(x), \phi_2(x)) \quad &= \quad \text{const} + \mu\phi_1(x)^2 + \\ &\qquad \sqrt{\mu g}\,\phi_1(x)(\phi_1(x)^2 + \phi_2^2(x)) + \frac{g}{4}(\phi_1(x)^2 + \phi_2(x)^2)^2 \end{aligned} \tag{3.7}$$

There are several things to notice:

1. Of course, this potential is no longer rotational invariant. The symmetry is broken by the ground state and, therefore, small fluctuations around this ground state are not invariant either.

2. Closely related to the previous remark is the fact that we now find terms of third order in the fields:

   $$\phi_1(x)^3 \quad \text{and} \quad \phi_1(x)\phi_2(x)^2\,, \tag{3.8}$$

   i.e., in the language of perturbation theory we would expect 3-point interactions among these fields.

3. The most important point, however, is that the field $\phi_2(x)$ does not have a quadratic term in the potential (there is, of course, a quadratic term involving the first derivatives of the field). This means that the quantized fluctuations of this degree of freedom correspond to massless particles. These massless particles are called *Goldstone bosons*. According to a general theorem one will always find massless Goldstone bosons in a theory with a spontaneously broken continuous symmetry except in two space-time dimensions.

   The exclusion of two space-time dimensions is not obvious on this level. The (hand-waving) reason is that there are no massless particles in two space-time dimensions because the correlation function of a massless excitation in two dimensions is a logarithm which requires a scale.

   The field $\phi_1(x)$ which corresponds to the radial fluctuations with respect to the ground state has a mass term μ^2 which means that the actual mass of the corresponding particle is $m = \sqrt{2}\mu$.

4. If we had chosen a theory with a three-component field and an SO(3) invariance instead of an SO(2) invariance, there would have been two massless Goldstone bosons and one massive mode which corresponds to the Higgs boson. More general, if we start with a theory with a symmetry group G and if the ground state of the system is only invariant under a subgroup H of G, then the number of massless Goldstone bosons is equal to the difference between the dimension of G and the dimension of H, i.e. the dimension of G/H. To each generator in G/H corresponds one massless Goldstone boson.

3.2 The Higgs mechanism

We now couple the scalar field of the previous section to a gauge theory. The gauge group will be SO(2). This is an abelian group which, at first sight, may look too special. However, the essential mechanism does not depend on the nature of the gauge group. In general, what is important is the total symmetry group G and the subgroup $H \subset G$ for which the symmetry is not broken by the ground state of the scalar field. The Higgs mechanism then works for any gauge field which corresponds to a generator in G/H.

The particular case of a two-component (real) scalar field and an SO(2) symmetry group is equivalent to a one-component complex scalar field and an U(1) symmetry group. From this point of view it would be natural to switch to a complex valued field and simply work with "numbers". However, the SO(2) is slightly more instructive as it shows that the gauge field is a variable associated to the generator of a group, which is more difficult to recognize in the complex picture. For this reason we keep the SO(2) notion even though it may be more complicated.

SO(2) is the group of rotations in a 2-dimensional plane. Each element can be written as:

$$U(\alpha) = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix} . \tag{3.9}$$

So we can say that the scalar theory defined by eq. 3.1 is invariant under the transformation

$$\varphi(x) \longrightarrow U(\alpha)\varphi(x) \tag{3.10}$$

or, explicitly,

$$\begin{pmatrix} \varphi_1(x) \\ \varphi_2(x) \end{pmatrix} \longrightarrow \begin{pmatrix} \varphi_1'(x) \\ \varphi_2'(x) \end{pmatrix} = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} \varphi_1(x) \\ \varphi_2(x) \end{pmatrix} . \tag{3.11}$$

We now introduce a gauge field $A_\mu(x)$ in order to make this global invariance to a local one. We define the covariant derivative:

$$D_\mu\varphi(x) = (\partial_\mu + igA_\mu(x))\varphi(x) . \tag{3.12}$$

As $\varphi(x)$ is a 2-vector, the covariant derivative must be a 2×2 matrix. We know from the section on gauge fields, that $A_\mu(x)$ actually is an element of the Lie-algebra of the invariance group. In the simple case of SO(2) the Lie-algebra consists of only one generator, i.e. it corresponds to a one-dimensional linear space. The generator of SO(2) is well known, but it is also easily derived from the general expression:

$$\begin{aligned} U^{-1}(\alpha)\left(\frac{\partial}{\partial\alpha}U(\alpha)\right) &= \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} -\sin\alpha & \cos\alpha \\ -\cos\alpha & -\sin\alpha \end{pmatrix} \\ &= \begin{pmatrix} -\cos\alpha\sin\alpha + \sin\alpha\cos\alpha & \cos^2\alpha + \sin^2\alpha \\ -\sin^2\alpha - \cos^2\alpha & -\sin\alpha\cos\alpha + \cos\alpha\sin\alpha \end{pmatrix} \\ &= \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} =: \mathrm{J} . \end{aligned} \tag{3.13}$$

In the general case the group element depends on a set of "coordinates" $\{\alpha_a\}$ ($a = 1, ..., d$, d being the dimension of the group) and one can take the derivative with respect to each coordinate and obtains a generator T_a. The Lie-algebra is obtained by taking linear combinations of the these generators. Therefore, the general form of a gauge field is

$$A_\mu \sim \sum_a A_\mu^a T_a \tag{3.14}$$

or, to be even more explicit:

$$(A_\mu)_{ij} \sim \sum_a A_\mu^a (T_a)_{ij} \tag{3.15}$$

In the present case of SO(2) there is only one generator and we identify

$$A_\mu \sim \begin{pmatrix} 0 & A_\mu \\ -A_\mu & 0 \end{pmatrix} \tag{3.16}$$

According to the general formalism, the transformation law of the gauge field is (see eq. 2.6)

$$A_\mu \to A_\mu^U = U(x)^{-1}A_\mu(x)U(x) - \mathrm{i}U^{-1}(x)\partial_\mu U(x) , \tag{3.17}$$

(where now instead of writing $U(\alpha(x))$ we simply write $U(x)$).

As SO(2) is an abelian group, all group elements commute and they commute with the generators of the Lie-algebra. Therefore, we obtain the simple transformation law:

$$A_\mu^U(x) = A_\mu(x) - \partial_\mu\alpha(x) . \tag{3.18}$$

(To obtain the matrix notation we just have to multiply this equation with the only generator J, see eq. 3.13.)

The Lagrange function of the total system (gauge field and scalar field) is given by:

$$L = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{1}{2}(D_\mu\varphi)(D^\mu\varphi) - V(\varphi^2)\,. \tag{3.19}$$

Now let's see what happens, if the scalar field undergoes spontaneous symmetry breaking. We expand the scalar field around the minimal value (like in eq. 3.6) and express the Lagrange function in terms of the new field $\phi(x)$. Furthermore, we fix a gauge and eliminate the "spurious" gauge degree of freedom by setting

$$\text{unitarygauge} \qquad \varphi_2(x) \equiv 0\,. \tag{3.20}$$

As we have the freedom to make a rotation for the scalar field $\varphi(x)$ at every point independently, we can always achieve this condition. (There are problems, if the value of $\varphi(x)$ at some point x is 0, because in this case the above condition does not fix the gauge. This is a technically relevant point which, however, shall not concern us here. In the phase where the symmetry is spontaneously broken this case is not so relevant because the ground state is defined by a non-zero value of the scalar field.)

The relevant term in the new Lagrange function stems from the covariant derivatives. The field tensor of the gauge fields is not affected by the symmetry breaking, and the potential of the scalar field we have already calculated (eq. 3.7) where now, however, $\phi_2(x) \equiv 0$ due to our choice of gauge. So we only consider the terms involving the covariant derivative:

$$\begin{aligned}
\frac{1}{2}D_\mu\varphi(x)D^\mu\varphi(x) &= +\frac{1}{2}(\partial_\mu + \mathrm{i}gA_\mu(x))(\varphi_0 + \phi(x))(\partial^\mu + \mathrm{i}gA^\mu(x))(\varphi_0 + \phi(x)) \\
&= \frac{1}{2}\Big((\partial_\mu\phi)(\partial^\mu\phi)\Big) - g^2\varphi_0 A_\mu A^\mu\varphi_0\Big) \\
&\quad - \frac{1}{2}g^2\phi(x)A_\mu(x)A^\mu(x)\phi(x) - e^2\varphi_0 A_\mu A^\mu\phi(x)
\end{aligned}$$

All other terms (involving only one gauge field) vanish because the second component of the Higgs field vanishes (due to the choice of gauge) and, therefore, $\varphi\,\mathrm{J}\,\varphi = 0$. Notice, that now we find a "mass"-term for the gauge field, where the mass is given by $g^2\varphi_0^2$ and involves the coupling constant g as well as the vacuum expectation value φ_0 of the Higgs field.

At first sight it seems that gauge invariance is broken. However, we have only chosen a particular gauge. We still could have expressed everything in terms of both scalar fields. In this case the second component of the Higgs field still corresponds to the Goldstone Boson. In a hand-waving manner one can say that the Goldstone Boson of the gauge theory with broken symmetry corresponds to the additional degree of freedom which is needed in order to transform the originally massless gauge field into a massive gauge field. (Note that the massless photon only has two degrees of freedom corresponding to the two possible transversal oscillations of the classical field, while a massive spin-1 particle needs three degrees of freedom. The longitudinal degree of freedom comes from absorbing the Goldstone Boson into the gauge fields.

Kapitel 4

Why quantum fields?

Eventually, the reason why we use a particular model is determined by its success. So the question "Why quantum fields?" has the answer: "Because they are successful." However, there are certain ideas associated with quantum fields which make them a natural candidate for a theory of elementary particle physics. Therese ideas are centered around two aspects:

1. Particles are interpreted as the excitation modes of harmonic oscillators whose frequencies correspond to the possible energies of the particles. The paradigm of this idea is the linearly coupled chain of harmonic oscillators. Interactions are interpreted as "annihilation of certain modes and creation of other modes". It would correspond to non-linear interactions between the different oscillatory modes of the linear chain. This picture naturally leads to quantum field theories.

2. Starting with a family of Schrödinger equations for N particles ($N = 0, 1, 2, 3, ...$) and introducing operators which "create" particles or "delete" particles also leads to quantum fields. These quantum fields obey a non-linear Schrödinger equation. This way of introducing quantum fields is sometimes referred to as "second quantization".

In the following two sections we briefly explain these two aspects.

4.1 Particles as modes of oscillators

In this section we first consider briefly the linear chain of harmonically coupled oscillators where each massive particle is in an additional external harmonic oscillator potential. This systems describes "massive photons" along a linear chain and naturally extends to the higher dimensional cases as well as to the continuum case.

Consider the linear harmonic chain of N massive point-like particles defined by the Hamiltonian:

$$H = \frac{1}{2m}\sum_{i=1}^{N} P_i^2 + \frac{D}{2}\sum_{i=1}^{N}(Q_{i+1} - Q_i)^2 + \frac{M^2}{2}\sum_{i=1}^{N} Q_i^2 \tag{4.1}$$

where we use periodic boundary conditions, i.e. $Q_{N+1} = Q_1$. The unusual last term corresponds to an additional harmonic potential for each particle with a tension defined by M^2.

The equations of motion read:

$$\ddot{Q}_i = \frac{D}{m}(Q_{i+1} - 2Q_i + Q_{i-1}) - M^2 Q_i \,. \tag{4.2}$$

The unitary transformation from variables Q_i to variables q_k defined by

$$q_k = \frac{1}{\sqrt{N}} \sum_{n=1}^{N} \exp\left(\frac{2\pi \mathrm{i} k n}{N}\right) Q_n \tag{4.3}$$

with the inverse transformation

$$Q_n = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} \exp\left(\frac{-2\pi \mathrm{i} k n}{N}\right) q_k \tag{4.4}$$

leads to

$$\ddot{q}_k = -4\omega_0^2 \sin^2 \frac{\pi k}{N}\, q_k - M^2 q_k \quad \text{with} \;\; \omega_0^2 = \frac{D}{m} \,. \tag{4.5}$$

Therefore, we can write the equations of motion as

$$\ddot{q}_k = -(\omega(k)^2 + M^2)\, q_k \tag{4.6}$$

with

$$\omega(k) = 2\omega_0 \left| \sin \frac{\pi k}{N} \right| , \tag{4.7}$$

and the frequencies:

$$k = 0, \pm 1, \pm 2, ..., \pm \frac{N-2}{2}, \frac{N}{2} \qquad k \text{ even} \tag{4.8}$$

$$k = 0, \pm 1, \pm 2, ..., \pm \frac{N-1}{2} \qquad k \text{ odd} \,. \tag{4.9}$$

These are the equations of motion for N decoupled harmonic oscillators. The Hamiltonian of this system can be written as

$$H = \sum_k \epsilon(k) \left(a(k)^+ a(k) + \frac{1}{2} \right) \tag{4.10}$$

where

$$\epsilon(k) = +\hbar \sqrt{w(k)^2 + M^2} \,. \tag{4.11}$$

Remarks:

1. The operators q_k are not completely independent. The fact that the position operators Q_i are self-adjoint leads to the condition

$$q_{-k} = q_k^+ \,. \tag{4.12}$$

Notice that for the oscillating frequencies we have

$$\epsilon(k) = \epsilon(-k) \tag{4.13}$$

(otherwise the Hamiltonian would not be self-adjoint), i.e. the two eigenvalues are degenerate. Therefore, we can "decouple" these two modes into two real modes. This will be done in detail in the next section.

2. Notice that the equation of motion, eq. 4.2, is a discrete version of a one-dimensional "massive wave equation":

$$\frac{\partial^2}{\partial t^2} Q(t,x) = c^2 \frac{\partial^2}{\partial x^2} Q(t,x) - M^2 Q(t,x)\,, \tag{4.14}$$

where c is related to the propagation velocity of waves along the chain.

3. In a similar way, if we use the equation of motion to replace P by $m\dot{Q}$, we obtain for the Hamiltonian, eq. 4.1

$$H = \int \mathrm{d}x \left(\frac{m}{2} (\dot{Q}(x,t))^2 + \frac{D'}{2} (Q'(x,t))^2 + \frac{M^2}{2} Q(t,x)^2 \right) \tag{4.15}$$

which corresponds to a Lagrange function

$$L = \int \mathrm{d}x \left(\frac{m}{2} (\dot{Q}(x,t))^2 - \frac{D'}{2} (Q'(x,t))^2 - \frac{M^2}{2} Q(t,x)^2 \right) \tag{4.16}$$

This system describes N "massive phonons" which can propagate along the chain. It is a discretized version of a 1-dimensional quantum field theory and the prototype of a massive scalar field, where the mass of the field is given by the external harmonic potential characterized by M^2 and the different possible momenta are given by $w(k)$. Without the external potential ($M^2 = 0$) the excitations are massless.

We now consider briefly the more general case. For simplicity we consider one type of particles. A single particle may assume energies $\epsilon(k)$ where, for the moment, k is some running index. Lets assume that an operator $a^+(k)$ generates a particle (mode) of this energy and an operator $a(k)$ annihilates a particle (mode) of this energy. Then we would expect that the system of free particles is described by the Hamiltonian:

$$\hat{H} = \sum_k \epsilon(k)\, a^+(k) a(k)\,. \tag{4.17}$$

If $a(k)$ and $a^+(k)$ satisfy the usual commutation relations:

$$[a(k), a^+(k')] = \delta_{k\,k'} \qquad [a(k), a(k'))] = [a^+(k), a^+(k')] = 0\,,$$

the multi-particle states are automatically symmetric and will describe bosons. The eigenstates for the Hamiltonian, as well as for the particle number operator,

$$\hat{N} = \sum_k \hat{n}(k) \qquad \hat{n}(k) = a^+(k) a(k)$$

can be labeled by the occupation numbers $n(k)$:

$$|n(1), n(2), ...\rangle = \frac{1}{\sqrt{n_1! n_2! \cdots}} (a^+(1))^{n_1} (a^+(2))^{n_2} ... |0\rangle \tag{4.18}$$

where $|0\rangle$ denotes the ground state of the Hamiltonian (and the number operator) and is interpreted as the vacuum state:

$$a(k)|0\rangle = 0 \quad \Longrightarrow \quad \hat{H}|0\rangle = \hat{N}|0\rangle = 0\,.$$

For the states (eq. 4.18) we have:

$$\begin{aligned} \hat{H}|n_1, n_2, ...\rangle &= \left(\sum_i n(k_i)\epsilon(k_i)\right)|n_1, n_2, ...\rangle \\ \hat{N}|n_1, n_2, ...\rangle &= \left(\sum_i n(k_i)\right)|n_1, n_2, ...\rangle \end{aligned}$$

If the creation and annihilation operators anticommute,

$$\{a^+(k), a(k')\} = \delta_{k\,k'} \qquad \{a(k), a(k'))\} = \{a^+(k), a^+(k')\} = 0\,,$$

where

$$\{A, B\} = AB + BA\,,$$

the states will have the property that $n(k) \in \{0, 1\}$ and they are antisymmetric, i.e., they describe fermions.

The generalization of the linear chain discussed above is the Klein-Gordon equation which describes a neutral scalar particle. If the field is complex the particle would be interpreted as having a charge. If we apply canonical quantization to the Dirac field (for free fields, canonical quantization makes sense, see Sec. 6.1), there are some technical problems to be solved in connection with the negative energies which can be solved by "filling up the Dirac sea" of negative energy modes.

Slightly more difficult is the relation to interacting particles. These are described by terms like

$$\sum_{k_1,...,k_m,l_1,...,l_n} \lambda(k_1, ..., l_n)\delta\left(\sum_i \epsilon(k_i) - \sum_j \epsilon(l_j)\right) a^+(k_1)...a^+(k_n)a(l_1)...a(l_n)$$

in the Hamiltonian 4.17. This term represents an interaction: certain modes $\{l_i\}$ are annihilated while other modes $\{k_i\}$ are created and the overall energy is conserved. While in the context of solid state physics or quantum optics the coupling constants $\lambda(k_1, ..., l_n)$ are determined by experiment, the locality of interactions in particle physics puts strong constraints onto these couplings.

4.2 The interpretation of the modes

In the previous section we have defined the new modes q_k as functions of Q_i. All operators q_k commute, and as $q_k^+ = q_{-k}$, also the operator q_k^+ commutes with all q_k. In order to derive the correct correspondence between the creation and annihilation operators $a(k)$ and the operators Q_i and P_i, we start with the commutation relations for q_k and $\dot{q}_k^+$.

From eq. 4.3 we find:

$$\dot{q}_k = \frac{1}{\sqrt{N}} \sum_n \exp\left(\frac{2\pi \mathrm{i} kn}{N}\right) \dot{Q}_n \,, \tag{4.19}$$

and the corresponding inverse transformation. The canonical commutation relations for Q_i and P_i lead to:

$$\begin{aligned}
[q_k, \dot{q}_{k'}] &= \frac{1}{N} \sum_n \sum_{n'} \exp\left(\frac{2\pi \mathrm{i}}{N}(kn + k'n')\right) [Q_n, \dot{Q}_{n'}] \\
&= \frac{1}{N} \sum_{nn'} \exp\left(\frac{2\pi \mathrm{i}}{N}(kn + k'n')\right) \mathrm{i}\frac{\hbar}{m} \delta_{nn'} \\
&= \frac{1}{N} \sum_n \exp\left(\frac{2\pi \mathrm{i}}{N} n(k + k')\right) \frac{\mathrm{i}\hbar}{m} \\
&= \frac{\mathrm{i}\hbar}{m} \delta_{k+k'}
\end{aligned}$$

The non-trivial commutation relations are now between q_k and $\dot{q}_{-k}$.

Furthermore, when one inserts the expansion eq. 4.3 into the Hamilton operator eq. 4.1 (using also that $P_i = mQ_i$) one obtains:

$$H = \frac{m}{2} \sum_k \dot{q}_k \dot{q}_{-k} + \frac{M^2}{2} \sum_k q_k q_{-k} + \frac{D}{2} \sum_k \left(\exp\left(-\frac{2\pi \mathrm{i} k}{N}\right) - 1\right)^2 q_k q_{-k} \,. \tag{4.20}$$

This again confirms that the canonical conjugated variable to q_k is $\dot{q}_{-k}$. At first sight, this expression does not look like a self-adjoint operator. However, note that the sum runs over $k = 0, \pm 1, \pm 2, ...$ and the terms $q_k q_{-k}$ occur twice (apart from the cases $k = 0$ and - if N is even - $k = N/2$, for which the last term is real anyhow and for which $q_k = q_{-k}$). Therefore, we can rearrange the summation in the last term such that the factor in front of a term $q_k q_{-k}$ becomes real:

$$H = \frac{m}{2} \sum_k \dot{q}_k \dot{q}_{-k} + \frac{M^2}{2} \sum_k q_k q_{-k} + \frac{D}{2} \sum_k 4 \sin^2\left(\frac{\pi \mathrm{i} k}{N}\right) q_k q_{-k} \,. \tag{4.21}$$

In order to apply the methods of the harmonic oscillator it would be preferable to deal with self-adjoint operators. We define:

$$q'(k) = \frac{1}{\sqrt{2}}(q_k + q_{-k}) \quad \text{and} \quad q'(-k) = \frac{\mathrm{i}}{\sqrt{2}}(q_k - q_{-k}) \quad \text{for} \;\; k \neq 0, N/2 \tag{4.22}$$

and

$$q'(0) = q(0) \qquad q'(N/2) = q(N/2). \tag{4.23}$$

Obviously, $q'(k)$ are self-adjoint operators. We find the following commutation relations:

$$[q'(k), \dot{q}'(k)] = [q'(-k), \dot{q}'(-k)] = \frac{\mathrm{i}\hbar}{m} \tag{4.24}$$

and all other commutators vanish. Now, the commutation relations are "canonical". Furthermore, we find that

$$q'(k)^2 + q'(-k)^2 = q_k q_{-k} + q_{-k} q_k \,. \tag{4.25}$$

Therefore, the Hamiltonian now takes a canonical form:

$$H = \frac{m}{2}\sum_k \dot{q}'_k\dot{q}'_k + \frac{M^2}{2}\sum_k q'_k q'_k + \frac{D}{2}\sum_k 4\sin^2\left(\frac{\pi \mathrm{i} k}{N}\right) q'_k q'_k\,, \tag{4.26}$$

and due to the commutation relations (eq. 4.24), this corresponds to a sum of N independent harmonic oscillations. Now, we can make the step to creation and annihilation operators:

$$a(k) = \sqrt{\frac{m\sqrt{\omega(k)^2+M^2}}{2\hbar}}\left(q'(k) + \frac{\mathrm{i}}{\sqrt{\omega(k)^2+M^2}}\dot{q}'(k)\right) \tag{4.27}$$

$$a(k)^+ = \sqrt{\frac{m\sqrt{\omega(k)^2+M^2}}{2\hbar}}\left(q'(k) - \frac{\mathrm{i}}{\sqrt{\omega(k)^2+M^2}}\dot{q}'(k)\right)\,. \tag{4.28}$$

4.3 Second quantization of Schrödinger's equation

The following approach to Quantum Field Theory starts from a Schrödinger equation. We will briefly describe the non-relativistic case, but relativistic generalizations exist. The term "second quantization" refers to the following: We can start from a classical equation of motion (with conjugated variables x and p) and quantize this equation of motion by replacing position and momentum variables by operators Q and P which satisfy the canonical commutation relations. This is what is done in quantum mechanics and is called "first quantization". If we switch to the Schrödinger picture (time-dependent states) we obtain a classical field equation for the quantum theory. The second quantization consists in taking these classical fields (the Schrödinger fields $\Psi(x,t)$) and replacing them by operators which satisfy canonical commutation relations. This will lead, as we shall see, to a quantum field theory describing an indeterminate number of interacting particles. The following explanations are very rudimentary. For more details see, e.g., [7, 94].

The Schrödinger equation for N interacting particles (e.g., interacting via a potential $W(r)$) can be written as:

$$\begin{aligned}\mathrm{i}\hbar\frac{\partial}{\partial t}\psi(x_1,...,x_N,t) &= H(\partial_{x_1},...,\partial_{x_N},x_1,...,x_n)\psi(x_1,...,x_N,t)\\ &= \left(-\frac{\hbar^2}{2m}\sum_i \Delta_i + g\sum_{i\neq j} W(|x_i-x_j|)\right)\psi(x_1,...,x_N,t)\end{aligned}$$

Depending on whether one wants to describe bosons or fermions the wave function has to be symmetrized or antisymmetrized. Note that this equation can be derived from a Lagrange function

$$L[\{\psi\}] = \int \mathrm{d}^N x\ \bar{\psi}(x_1,...,x_N,t)\left(\mathrm{i}\hbar\frac{\partial}{\partial t} - H_x\right)\psi(x_1,...,x_N,t) \tag{4.29}$$

with

$$H_x = \left(-\frac{\hbar^2}{2m}\sum_i \Delta_i + g\sum_{i\neq j} W(|x_i-x_j|)\right)\,. \tag{4.30}$$

Actually, from this Lagrange function we obtain two equations, one for $\psi(x_1, ..., x_N, t)$ and the other (complex-conjugated equation) for $\bar{\psi}(x_1, ..., x_N, t)$. If we introduce also the trivial cases $N = 0$ and $N = 1$, we obtain an infinite family of Schrödinger equations:

$$\begin{aligned}
\mathrm{i}\hbar\frac{\partial}{\partial t}\psi(t) &= 0 \quad \Longrightarrow \quad \psi(t) = \psi_0 \in \mathbf{C} \\
\mathrm{i}\hbar\frac{\partial}{\partial t}\psi(x,t) &= \left(-\frac{\hbar^2}{2m}\right)\Delta\psi(x,t) \\
\mathrm{i}\hbar\frac{\partial}{\partial t}\psi(x_1,x_2,t) &= \left(-\frac{\hbar^2}{2m}(\Delta_1+\Delta_2) + gW(|x_1-x_2|)\right)\psi(x_1,x_2,t) \\
\cdots &\quad \cdots \\
\mathrm{i}\hbar\frac{\partial}{\partial t}\psi(x_1,...,x_N,t) &= \left(-\frac{\hbar^2}{2m}\sum_i \Delta_i + g\sum_{i\neq j} W(|x_i-x_j|)\right)\psi(x_1,...,x_N,t) \\
\cdots &\quad \cdots
\end{aligned} \tag{4.31}$$

A general state is now described in the Hilbert space $\mathcal{H} = \bigoplus_{N=0}^{\infty}\mathcal{H}_N$, where $\mathcal{H}_N$ is the standard N particle Hilbert space of (symmetrical or anti-symmetrical) square integrable functions in N variables. This space is sometimes referred to as a Fock space. Note that we take the vectorspace sum over N and not the tensor product. We want to describe superpositions of N-particle states, and not a decomposition of a large system into subsystems. In this space a state can be represented by:

$$|\Phi\rangle = |\psi_0, \psi_1, \psi_2, \psi_3, ...\rangle = \begin{pmatrix} \psi_0 \\ \psi_1 \\ \psi_2 \\ \psi_3 \\ \vdots \end{pmatrix}, \tag{4.32}$$

where ψ_N is a square integrable function of N variables (N particle state function). To obtain the matrix elements of Φ in position space we first define:

$$|x\rangle = \begin{pmatrix} 1 \\ |x_1\rangle \\ |x_1, x_2\rangle \\ \vdots \end{pmatrix}$$

and obtain:

$$\Phi(x) = \langle x|\Phi\rangle = (\psi_0, \psi_1(x_1), \psi_2(x_1, x_2), ...) \simeq \begin{pmatrix} \psi_0 \\ \psi_1(x_1) \\ \psi_2(x_1, x_2) \\ \vdots \end{pmatrix}. \tag{4.33}$$

For the following, all functions are symmetric in their arguments:

$$\psi_N(x_1, ..., x_N) = \psi_N(x_{\pi_1}, x_{\pi_2}, ..., x_{\pi_n})$$

where π denotes a permutation of the N arguments. For completely antisymmetric functions (describing fermions) a similar formalism exists.

The normalization condition reads:

$$\sum_{N=0}^{\infty} \int \mathrm{d}^N x \, |\psi_N(x_1, ..., x_N)|^2 = 1 \, .$$

Up to now we describe systems of N interacting particles, but we do not have the means to "create" or annihilate particles. In order to incorporate this possibility we introduce operators $\hat{\Psi}^+(x)$ and $\hat{\Psi}(x)$ on this Hilbert space $\mathcal{H}$ as well as "smeared operators"

$$\hat{\Psi}^+(f) = \int \mathrm{d}x \, \hat{\Psi}^+(x) f(x) \qquad \hat{\Psi}(g) = \int \mathrm{d}x \, \hat{\Psi}(x) g(x)^*$$

and define:

$$\hat{\Psi}^+(f) \begin{pmatrix} \psi_0 \\ \psi_1 \\ \psi_2 \\ \psi_3 \\ \vdots \end{pmatrix} = \begin{pmatrix} 0 \\ \frac{1}{\sqrt{1}} \psi_0 f \\ \frac{1}{\sqrt{2}} \psi_1 f \\ \frac{1}{\sqrt{3}} \psi_2 f \\ \vdots \end{pmatrix} , \tag{4.34}$$

where

$$\psi_N f = \sum_{i=1}^{N+1} \psi_N(x_1, ..., x_{i-1}, x_{i+1}, ..., x_{N+1}) f(x_i)$$

is a symmetric function in $N+1$ variables and belongs to the Hilbert space $\mathcal{H}_{N+1}$. This state describes $N+1$ particles (symmetrized) where one particle has the wave function $f(x)$. Without the smearing procedure one can say that $\hat{\Psi}^+(x)$ "creates" a particle at point x (plus automatic symmetrization).

Similarly, we define

$$\hat{\Psi}(f) \begin{pmatrix} \psi_0 \\ \psi_1 \\ \psi_2 \\ \vdots \end{pmatrix} = \begin{pmatrix} \sqrt{1} \int \psi_1 f^* \\ \sqrt{2} \int \psi_2 f^* \\ \sqrt{3} \int \psi_3 f^* \\ \vdots \end{pmatrix}$$

where

$$\int \psi_N f^* = \int \mathrm{d}x \, \psi_N(x_1, x_2, ..., x_{N-1}, x) f(x)^* \tag{4.35}$$

is a symmetric function of $N-1$ arguments. (Note that ψ_N is symmetric, i.e., it doesn't matter which of the arguments one integrates over.) This equation can be interpreted as a partial scalar product. If the wave function $f(x)$ is completely orthogonal to ψ_N, this product is zero. In a sense this product measures the component of $f(x)$ in Ψ_N and annihilates this component of a particle. Even more general one might say that $\hat{\Psi}(f)$ annihilates a particle with wave function f, or $\hat{\Psi}(x)$ annihilates a particle at point x.

One can easily show that

$$[\hat{\Psi}(g), \hat{\Psi}^+(f)] = \int \mathrm{d}x \, f(x) g(x)^* \qquad \text{or} \qquad [\hat{\Psi}(x), \hat{\Psi}^+(y)] = \delta(x-y) \tag{4.36}$$

and

$$[\hat{\Psi}(x), \hat{\Psi}(y)] = [\hat{\Psi}^+(x), \hat{\Psi}^+(y)] = 0\,.$$

The interesting aspect of these operators is that if we define the operator Hamiltonian:

$$\hat{H} = -\frac{\hbar^2}{2m}\int \mathrm{d}x\ \hat{\Psi}^+(x)\Delta\hat{\Psi}(x) + g\int\mathrm{d}x\int\mathrm{d}y\ \hat{\Psi}^+(y)\hat{\Psi}^+(x)W(x,y)\hat{\Psi}(x)\hat{\Psi}(y)\,,$$

then the Schrödinger equation

$$\mathrm{i}\hbar\frac{\partial}{\partial t}\Phi(x;t) = \hat{H}\Phi(x;t)$$

(with $\Phi(x)$ defined as in eq. 4.33) yields the infinite sequence of Schrödinger equations of eq. (4.31).

Instead of considering time-dependent states (the Schröding picture) we can consider time-dependent field operators (Heisenberg picture) where:

$$\hat{\Psi}_H(x,t) = \exp\left(\frac{\mathrm{i}}{\hbar}\hat{H}t\right)\hat{\Psi}(x)\exp\left(-\frac{\mathrm{i}}{\hbar}\hat{H}t\right)\,.$$

Of course, these field operators satisfy the Heisenberg equation of motion:

$$\mathrm{i}\hbar\frac{\partial}{\partial t}\hat{\Psi}(x,t) = [\hat{\Psi}(x,t), \hat{H}]\,.$$

Calculating the commutator on the right hand side using the canonical commutation relations (eq. 4.36) yields:

$$\mathrm{i}\hbar\frac{\partial}{\partial t}\hat{\Psi}(x) = -\frac{\hbar^2}{2m}\Delta\hat{\Psi}(x) + g\int \mathrm{d}y\ \hat{\Psi}^+(y)W(|x-y|)\hat{\Psi}(y)\hat{\Psi}(x)\,.$$

In particular, for a local interaction $W(|x-y|) = \delta(x-y)$ (hard-core point-like particles), we obtain the non-linear Schrödinger equation:

$$\mathrm{i}\hbar\frac{\partial}{\partial t}\hat{\Psi}(x) = -\frac{\hbar^2}{2m}\Delta\hat{\Psi}(x) + g\hat{\Psi}^+(x)\hat{\Psi}(x)\hat{\Psi}(x)$$

which now is an operator field equation.

Kapitel 5

Funktionalintegral-Formalismus

5.1 Das Funktionalintegral in der Quantenmechanik

Die Zeitentwicklung eines Zustandes wird in der Quantenmechanik durch die Schrödinger-Gleichung

$$\mathrm{i}\hbar\frac{\partial\psi}{\partial t} \;=\; H\psi \tag{5.1}$$

beschrieben. Die allgemeinste Lösung des Anfangswertproblems dieser Gleichung ist

$$\psi(x,t) \;=\; \int \mathrm{d}y\; K(x,y;t)\;\psi(y,0)\,, \tag{5.2}$$

wobei

$$K(x,y;t) \;=\; \left\langle x\left|\mathrm{e}^{-\frac{\mathrm{i}}{\hbar}Ht}\right|y\right\rangle \tag{5.3}$$

Lösung der Schrödinger-Gleichung ist mit der δ-Distribution als Anfangsbedingung:

$$\begin{aligned}\mathrm{i}\hbar\frac{\partial K(x,y;t)}{\partial t} &= \left(-\frac{\hbar^2}{2m}\Delta_x + V(x)\right)K(x,y;t)\\ K(x,y;0) &= \delta(x-y)\,.\end{aligned} \tag{5.4}$$

Der Propagator K erfüllt die Halbgruppengleichung

$$K(x,y;t_1+t_2) \;=\; \int \mathrm{d}z\; K(x,z;t_1)\;K(z,y;t_2) \tag{5.5}$$

oder

$$\left\langle x\left|\mathrm{e}^{-\frac{\mathrm{i}}{\hbar}H(t_1+t_2)}\right|y\right\rangle \;=\; \int \mathrm{d}z\;\left\langle x\left|\mathrm{e}^{-\frac{\mathrm{i}}{\hbar}Ht_1}\right|z\right\rangle\;\left\langle z\left|\mathrm{e}^{-\frac{\mathrm{i}}{\hbar}Ht_2}\right|y\right\rangle\,.$$

Durch wiederholte Anwendung dieser Identität erhält man

$$\begin{aligned} & K(x,y;t) \\ = & \int \mathrm{d}x_1 \ldots \mathrm{d}x_{N-1}\; K(x,x_1;t/N) K(x_1,x_2;t/N) \ldots K(x_{N-1},y;t/N)\,. \end{aligned}$$

Für sehr kurze Zeiten $t \to 0$ kann man den Propagator geeignet approximieren. Dazu verwenden wir die Baker-Campbell-Hausdorff-Formel:

$$e^X e^Y = e^{Z(X,Y)} \quad \text{mit} \quad Z(X,Y) = X + Y + \frac{1}{2}[X,Y] + \frac{1}{12}[X,[X,Y]] - \frac{1}{12}[Y,[X,Y]] + \ldots$$

Damit gilt

$$\mathrm{e}^{-\frac{\mathrm{i}}{\hbar}(H_0+V)t} = \mathrm{e}^{-\frac{\mathrm{i}}{\hbar}H_0 t}\, \mathrm{e}^{-\frac{\mathrm{i}}{\hbar}Vt + \mathcal{O}(t^2)}\,. \tag{5.6}$$

Sei

$$\begin{aligned} K_0(x,y;t) &= \left\langle x \left| \mathrm{e}^{-\frac{i}{\hbar}H_0 t} \right| y \right\rangle \\ &= \left(\frac{m}{2\pi \mathrm{i}\hbar t}\right)^{\frac{1}{2}} \exp\left(\frac{\mathrm{i}}{\hbar}\frac{m}{2}\frac{(x-y)^2}{t}\right) \end{aligned} \tag{5.7}$$

der bekannte Propagator der freien Schrödinger-Gleichung, dann erfüllt daher

$$K(x,y;t) = K_0(x,y;t)\, \exp\left(-\frac{\mathrm{i}}{\hbar}V(x)t + \mathcal{O}(t^2)\right) \tag{5.8}$$

näherungsweise die volle Gleichung (5.4), wobei der Fehler (bedingt durch die Anwendung des Laplace-Operators auf die potentialabhängige Phase) durch Terme höherer Ordnung in t korrigiert wird.

Man erhält so für den Propagator formal die Darstellung ($\mathcal{N}$ ist durch eine geeignete Normierung des Maßes zu absorbieren):

$$\begin{aligned} & K(x,y;t) = \\ & = \mathcal{N} \lim_{N\to\infty} \int \prod_{i=1}^{N-1} \mathrm{d}x_i\; \exp\frac{\mathrm{i}}{\hbar} \sum_{i=0}^{N-1} \left[\frac{1}{2}m\left(\frac{x_{i+1}-x_i}{\Delta t}\right)^2 - V(x_i)\right]\Delta t \\ & = \int_{y\to x} \mathcal{D}x(\tau)\; \exp\frac{\mathrm{i}}{\hbar}\int_0^t \left[\frac{1}{2}m\dot{x}(\tau)^2 - V(x(\tau))\right]\mathrm{d}\tau \\ & = \int_{y\to x} \mathcal{D}x(\tau)\; \exp\frac{\mathrm{i}}{\hbar}S[x(\tau)]\,. \end{aligned}$$

Dies nennt man die *Funktionalintegraldarstellung* des Propagators. Andere Bezeichnungen sind „Summation“ über alle Wege, „Summation“ über alle Möglichkeiten, „Summation“ über alle Geschichten.

Jeder Weg (Möglichkeit, Geschichte) wird mit einer Phase „gewichtet“, die sich aus der klassischen Wirkung ergibt.

Für viele theoretische Anwendungen und Umformungen des Funktionalintegrals ist oft die Darstellung als Integral über den Phasenraum wichtig. Dazu gehen wir nochmals von der Aufspaltung des Propagators für kurze Zeiten aus (5.6), fügen diesmal aber einen vollständingen Satz von p-Eigenzuständen ein:

$$\begin{aligned}\left\langle x \left| \mathrm{e}^{-\frac{\mathrm{i}}{\hbar}Ht} \right| y \right\rangle &\approx \int \mathrm{d}p \, \left\langle x \left| \mathrm{e}^{-\frac{\mathrm{i}}{\hbar}H_0(P)t} \right| p \right\rangle \left\langle p \left| \mathrm{e}^{-\frac{\mathrm{i}}{\hbar}V(Q)t} \right| y \right\rangle \\ &= \int \mathrm{d}p \, \mathrm{e}^{-\frac{\mathrm{i}}{\hbar}[H_0(p)+V(x)]t} \langle x|p\rangle\langle p|y\rangle \\ &= \int \mathrm{d}p \, \mathrm{e}^{-\frac{\mathrm{i}}{\hbar}[H_0(p)+V(x)]t} \, \mathrm{e}^{-\frac{\mathrm{i}}{\hbar}p(x-y)} .\end{aligned}$$

Mit dieser Darstellung der Propagatorfunktion für kurze Zeiten erhalten wir:

$$\begin{aligned}K(x,y;t) \; &= \\ &= \mathcal{N} \lim_{N\to\infty} \int \prod_{i=1}^{N-1} \mathrm{d}x_i \, \mathrm{d}p_i \; \exp \; \frac{\mathrm{i}}{\hbar} \sum_{i=0}^{N-1} \left[\frac{p_i(x_{i+1}-x_i)}{\Delta t} - H(p_i,x_i) \right] \Delta t \\ &= \int_{y\to x} \mathcal{D}x(\tau)\, \mathcal{D}p(\tau) \; \exp \; \frac{\mathrm{i}}{\hbar} \int_0^t \left[p(\tau)\dot{x}(\tau) - H(,p(\tau),x(\tau)) \right] \mathrm{d}\tau \\ &= \int_{y\to x} \mathcal{D}x(\tau)\, \mathcal{D}p(\tau) \; \exp \frac{\mathrm{i}}{\hbar} S_H[p(\tau),x(\tau)] \, ,\end{aligned}$$

wobei

$$S_H[p(\tau),x(\tau)] \; = \; \int_0^t \mathrm{d}\tau (p\dot{x} - H(p,x))$$

die (Hamiltonsche) Wirkung im Phasenraum ist.

5.2 Zeitgeordnete Vakuumerwartungswerte

Die typischen Experimente der Hochenergiephysik sind Streuexperimente, und die wesentlichen Größen, die den Streuquerschnitt bestimmen, sind die Elemente der Streumatrix:

$$\langle p_1,\ldots,p_n|S|k_1,\ldots,k_m\rangle \; = \; {}_f\langle p_1,\ldots,p_n|k_1,\ldots,k_m\rangle_i \, .$$

Hierbei ist $|k_1,\ldots,k_m\rangle_i$ ein Zustand von m einlaufenden quasi-freien Teilchen mit Impulsen k_i, (die anderen Quantenzahlen – Spin, Ladung, Teilchensorte, etc. – sind implizit), und $|p_1,\ldots,p_n\rangle_f$ ist ein Zustand von n auslaufenden quasi-freien Teilchen mit Impulsen p_i. Diese Matrix-Elemente lassen sich im Rahmen einer Quantenfeldtheorie nach der folgenden Formel berechnen (siehe beispielsweise [14]):

$$\begin{aligned}{}_f\langle p_1,\ldots,p_n|k_1,\ldots,k_m\rangle_i \; &= \; \left(\frac{i}{\sqrt{Z}}\right)^{m+n} \prod_{i=1}^{m} \int \mathrm{d}^4x_i \prod_{j=1}^{n} \int \mathrm{d}^4y_j \; f_{k_i}(x_i)\, f^*_{p_j}(y_j) \\ &(-\Box_{x_i}+m^2)(-\Box_{y_j}+m^2)\langle\Omega|\mathcal{T}(\phi(y_1)\ldots\phi(y_n)\phi(x_1)\ldots\phi(x_m))|\Omega\rangle \, .\end{aligned}$$

Der Einfachheit halber wurden hier nur ein reelles skalares Feld betrachtet. Die Bedeutung der einzelnen Terme ist:

- $\sqrt{Z}$ ist ein Normierungsfaktor zwischen den Zuständen in der wechselwirkenden Theorie und den Zuständen der quasi-freien asymptotischen Teilchen. $\sqrt{Z}$ hat die Bedeutung einer Wellenfunktionsrenormierung.

- $f_p(x)$ sind Lösungen der freien Klein-Gordon-Gleichung zum Impuls p, wobei $p^2 = m^2$ der Masse des quasi-freien asymptotischen Teilchens entspricht.

- $\langle\Omega|\mathcal{T}(\phi(y_1)\dots\phi(y_n)\phi(x_1)\dots\phi(x_m))|\Omega\rangle$ sind die so genannten zeitgeordneten Vakuumerwartungswerte der Feldoperatoren. $\mathcal{T}$ bezeichnet hierbei die Zeitordnung, beispielsweise

$$\mathcal{T}(\phi(t_2,\boldsymbol{x}_2)\phi(t_1,\boldsymbol{x}_1)) \;:=\; \begin{cases} \phi(t_2,\boldsymbol{x}_2)\phi(t_1,\boldsymbol{x}_1) & \text{für } t_2 \geq t_1 \\ \phi(t_1,\boldsymbol{x}_1)\phi(t_2,\boldsymbol{x}_2)) & \text{für } t_2 < t_1 \end{cases} \tag{5.9}$$

  $|\Omega\rangle$ ist der Vakuumzustand oder auch Grundzustand der Theorie, d.h. der Eigenzustand des Hamilton-Operators zum niedrigsten Energieeigenwert (üblicherweise auf 0 normiert).

Sind mehrere Teilchensorten vorhanden, insbesondere auch Spin-$\frac{1}{2}$-Teilchen, so verallgemeinert sich obige Formel in folgender Hinsicht:

- Jedes Feld $\phi, \psi, \dots$ hat seine eigene Wellenfunktionsrenormierung $Z_\phi, Z_\psi, \dots$.

- Jedes Fermion trägt einen Faktor (-1) bei.

- $f_k(x)$ ist eine Impuls-k-Lösung der zugehörigen freien Gleichung (beispielsweise der Dirac-Gleichung).

- Der freie Operator $(-\Box_x + m^2)$ bzw. $(-\Box_y + m^2)$ ist durch den entsprechenden freien Operator (beispielsweise $(i\partial\!\!\!/_x + m)$ bzw. $(-i\partial\!\!\!/_y + m)$) zu ersetzen.

- Der zeitgeordnete Vakuumerwartungswert enthält die Feldoperatoren zu den entsprechenden Teilchen.

Die Hauptaufgabe der Quantenfeldtheorie besteht darin, die zeitgeordneten Vakuumerwartungswerte zu berechnen. Außerdem müssen die Parameter der Theorie (Kopplungskonstanten und die so genannte 'bare-mass', die „nackte“ Masse) mit den beobachtbaren asymptotischen Größen $(Z, m, \dots)$ in Beziehung gebracht werden.

In der Quantenmechanik lassen sich die zeitgeordneten Erwartungswerte des Ortsoperators im Grundzustand formal durch das Funktionalintegral ausdrücken:

$$\langle\Omega|\mathcal{T}(Q(t_2)Q(t_1))|\Omega\rangle \;=\; \frac{1}{Z}\int \mathcal{D}x(\tau)\; x(t_1)x(t_2)\; \exp\frac{\mathrm{i}}{\hbar}S[x]\,. \tag{5.10}$$

Dabei ist über alle Wege zu integrieren und zwar von $t = -\infty$ bis $t = \infty$. Z ist die „Zustandssumme“, d.h. das entsprechende Funktionalintegral ohne Einschübe.

In der Feldtheorie ist der Ort $(\boldsymbol{x})$ ein (kontinuierlicher) „Index“, die dynamischen Freiheitsgrade sind die Werte des Feldes:

$$\boldsymbol{x}(t) \;\to\; \varphi(\boldsymbol{x},t)\,. \tag{5.11}$$

Das Funktionalintegral wird zu einer Integration über alle Felder:

$$Z \;=\; \int \mathcal{D}\varphi \; \mathrm{e}^{\frac{\mathrm{i}}{\hbar}\int \mathrm{d}^4x \; \mathcal{L}(\varphi(x),\partial_\mu\varphi(x))} \,, \tag{5.12}$$

und die Erwartungswerte zeitgeordneter Produkte lassen sich entsprechend schreiben:

$$\begin{aligned} \langle\Omega|\mathcal{T}(\phi(x_1)\dots\phi(x_n))|\Omega\rangle \;&=\; \\ &=\; \frac{1}{Z}\int \mathcal{D}\varphi \; \varphi(x_1)\dots\varphi(x_n) \; \mathrm{e}^{\frac{\mathrm{i}}{\hbar}\int \mathrm{d}^4x \; \mathcal{L}(\varphi(x),\partial_\mu\varphi(x))} \,. \end{aligned} \tag{5.13}$$

(Man beachte, dass auf der linken Seite dieser Gleichung der Erwartungswert von Operatoren, $\phi(x)$, steht, während im Funktionalintegral über klassische Feldkonfigurationen zu integrieren ist.)

Entsprechende Ausdrücke erhält man für Eichfelder (A_μ) und fermionische Felder (ψ), wobei im letzteren Fall jedoch antikommutierende Variable (so genannte Grassmann-Variable) zu integrieren sind. Auf die Funktionalintegraldarstellung von Fermionen werden wir in Kapitel 17 näher eingehen. Die Schwierigkeiten im Zusammenhang mit der Funktionalintegralquantisierung von Eichfeldern (Wahl einer Eichung, Einfügen einer so genannten Faddeev-Popov-Determinante, Geister) werden hier nicht behandelt, da bei der Gitterregularisierung auf die Wahl einer Eichung verzichtet werden kann.

5.3 Euklidische Formulierung

Betrachten wir einen zeitgeordneten Erwartungswert

$$\begin{aligned} &\langle\Omega|\mathcal{T}(\phi(x_1)\dots\phi(x_n))|\Omega\rangle \;= \\ &\quad= \; \langle\Omega|\phi(\boldsymbol{x}_1)\mathrm{e}^{-\mathrm{i}(t_1-t_2)H}\phi(\boldsymbol{x}_2)\mathrm{e}^{-\mathrm{i}(t_2-t_3)H}\dots\mathrm{e}^{-\mathrm{i}(t_{n-1}-t_n)H}\phi(\boldsymbol{x}_n)|\Omega\rangle \\ &\qquad\qquad\qquad \text{mit } t_i - t_{i+1} \geq 0 \,. \end{aligned}$$

Da H ein positiver Operator ist (d.h. das Spektrum von H ist positiv) und die Zeitargumente geordnet sind ($t_i - t_{i+1} \geq 0$), können wir diesen Erwartungswert in den Argumenten $\{t_i\}$ in die obere komplexe Halbebene analytisch fortsetzen. Oft ist es aus rein praktischen Gründen sinnvoll, Vakuumerwartungswerte bei imaginären Zeitargumenten zu betrachten:

$$t \;\rightarrow\; \tau = \mathrm{i}t \,. \tag{5.14}$$

Natürlich muss zur Berechnung physikalischer Größen diese *Wick-Rotation* rückgängig gemacht werden, was aber im Endergebnis oft einfacher ist.

Im Folgenden wird diese *euklidische Formulierung* immer als „praktischer Rechentrick“ angesehen. Es soll nicht darüber diskutiert werden, inwieweit die euklidische Zeit „Realität“ besitzt. Eine axiomatische Formulierung der euklidischen QFT stammt von Osterwalder und Schrader [85]. Näheres zum euklidischen Funktionalintegral findet man beispielsweise bei Glimm und Jaffe [36].

Die Zeitintegration der Wirkung ist nun durch

$$\int \mathrm{d}t \quad\rightarrow\quad -\,\mathrm{i}\int \mathrm{d}\tau \tag{5.15}$$

zu ersetzen, und für die Lagrangefunktion eines Punktteilchens erhält man

$$L(\dot{x},x) \;=\; \frac{m}{2}\dot{x}^2 - V(x) \quad\longrightarrow\quad L_{\rm E} \;=\; -\frac{m}{2}\dot{x}^2 - V(x) \;=\; -E(\dot{x},x)\,. \tag{5.16}$$

Aus der Zustandssumme (und analog den Erwartungswerten zeitgeordneter Produkte) wird:

$$Z \;\to\; Z_{\rm eukl.} \;=\; \int \mathcal{D}x(\tau)\; {\rm e}^{-\int {\rm d}\tau'\; E(\dot{x},x)}\,. \tag{5.17}$$

Entsprechend ist das euklidische Funktionalintegral in der Feldhtheorie definiert als:

$$Z \;=\; \int \mathcal{D}\varphi(x)\; {\rm e}^{-\int {\rm d}^4x\; \mathcal{L}_{\rm E}(\varphi,\partial_\mu\varphi)}\,. \tag{5.18}$$

Man beachte, dass in der euklidischen Formulierung der Feldtheorie die invariante Bilinearform $g_{\mu\nu} = \text{diag}\,(1,1,1,1)$ ist. Insbesondere folgt daraus auch, dass $\gamma_4 = i\gamma_0$, was zu dem entsprechenden Vorzeichenwechsel im kinetischen Term des Dirac-Lagrangians (2.13) führt. Für die euklidische Klein-Gordon-Lagrange-Dichte erhalten wir

$$\mathcal{L}_{\rm E} \;=\; \frac{1}{2}\left[(\partial_\mu\varphi)^2 + m^2\varphi^2\right] \;=\; \frac{1}{2}[\dot{\varphi}^2 + (\vec{\nabla}\varphi)^2 + m^2\varphi^2]\,,$$

und für die euklidische Lagrange-Dichte der Dirac-Theorie:

$$\mathcal{L}_{\rm E} \;=\; {\rm i}\psi^+\partial_t\psi + {\rm i}\psi^+\gamma_0\vec{\gamma}\cdot\vec{\nabla}\psi + m\bar{\psi}\psi\,.$$

5.4 Störungstheorie

Die Funktionalintegrale — sei es in der Minkowskischen oder euklidischen Formulierung — sind im Allgemeinen nur formale Ausdrücke, mathematisch nicht definiert. In gewisser Hinsich handelt es sich lediglich um "Absichtserklärungen". (Ähnlich verhält es sich mit anderen Verfahren zur Quantisierung von nicht-freien Feldtheorien.) Eine Möglichkeit, dem Funktionalintegral einen mathematisch definierten Sinn zu geben — abgesehen von einigen Spezialfällen sowie der in dieser Vorlesung zu behandelnden Gitterregularisierung —, ist die Störungstheorie. Dabei wird nach Potenzen des Wechselwirkungsterms in der Wirkung (bzw. nach einer Kopplungskonstanten) entwickelt:

$$Z \;=\; \int \mathcal{D}\varphi\; {\rm e}^{-S_0[\varphi] - gS_{\rm int}[\varphi]} \tag{5.19}$$

$$\;=\; \sum_{n=0}^{\infty} \frac{(-g)^n}{n!} \int \mathcal{D}\varphi\; (S_{\rm int}[\varphi])^n\; {\rm e}^{-S_0[\varphi]}\,. \tag{5.20}$$

Da $S_0[\varphi]$ quadratisch in den Feldern ist, hat man das Integral auf diese Weise formal auf die Berechnung von Momenten von Gaußschen Integralen umgeschrieben. Die einzelnen Terme lassen sich als Integrale über die Green'schen Funktionen der freien Wirkung schreiben (Feynman-Integrale). Diese sind mathematisch immer noch nicht wohl-definiert, da Green'sche Funktionen Distributionen sind und nicht multipliziert werden können. Man kann jedoch Regularisierungen

angeben (z.B. Cut-offs im Impulsraum, Subtraktion von Termen mit unphysikalischen Teilchen, etc.), sodass die einzelnen Integrale wohldefiniert sind. Allerdings wird dabei die Theorie abgeändert, was sich durch die Existenz eines neuen Parameters (z.B. den Cut-off) ausdrückt.

Es lässt sich zeigen, dass für gewisse Theorien (so genannte *störungstheoretisch renormierbare* Quantenfeldtheorien) die divergenten Terme, die auftreten, wenn man den unpyhsikalischen neuen Parameter zu entfernen versucht (Cut-off $\to \infty$), durch Einführen so genannter Counter-Terme – letztendlich durch *Renormierung* der physikalischen Parameter wie Kopplunskonstanten, Massen, Feldstärken etc. – absorbieren kann. In Kapitel 7 wird die Regularisierung und Renormierung einer QFT etwas ausführlicher, aber immer noch in schematischer Form, erläutert.

Vor der Einführung der Gitterregularisierung waren die störungstheoretisch renormierbaren Quantenfeldtheorien die einzigen Theorien, die mathematisch definiert werden konnten. Daher wurden Theorien, die im Sinne der Störungtheorie als nicht-renormierbar galten (wie z.B. die 4-Fermion Theorie der schwachen Wechselwirkung oder eine φ^6-Theorie) als nicht konsistent angesehen.

Das skizzierte Verfahren (Entwicklung nach der Kopplungskonstanten — Regularisierung der Feynman-Integrale — Absorption der Divergenzen durch Renormierung der physikalischen Parameter) erlaubt die Definition der Quantenfeldtheorie Term für Term in der Störungstheorie. Es sagt nichts aus über die Konvergenz der Störungsreihe und erlaubt auch keine Vorhersagen für nicht-störungstheoretische Größen, wie z.B. gebundene Zustände. (Tatsächlich hat die Störungsreihe in fast allen Fällen keinen nicht-verschwindenden Konvergenzradius in den Kopplungen (vgl. [24]). Sie läßt sich aber als asymptotische Reihe interpretieren, d.h., bei genügend kleinen Kopplungen wird die darzustellende Funktion durch die ersten Terme der Störungsreihe genähert. Die Ordnung, bis zu der diese Näherung qualitativ besser wird, hängt von der Größe der Kopplung ab.)

Kapitel 6

Elementary axiomatic quantum field theory

Any regularization of a classical theory (defined, e.g., by a classical Lagrange density) destroys - or, at least, changes - certain structures of the theory. In most cases one thinks of symmetries which are destroyed by the regularization procedure. As it is usually very difficult to prove that a certain explicitly broken symmetry is restored in the scaling limit where the regularization parameter is removed, one tries to find regularization schemes which do not break the relevant symmetries. If the regularized theory possesses a certain symmetry it is much easier to show that this symmetry also holds for the continuum theory.

Lattice regularization will break the symmetries of space-time: In Minkowski space this is the Poincaré group. In the euclidean version of a field theory this will become the symmetry generated by SO(d) and T^d (where T represents translations and d refers to the space-time dimension of the theory). This symmetry is expected to be restored in the continuum limit. Even though this has been shown explicitly in many examples, a general proof is still missing. Ansätze for such a proof are conceptually similar as what is known from the classical mechanics of rigid bodies: It suffices to show that the tensor of inertia is invariant under rotations of $90°$, then the eigenvalues are degenerated and the invariance holds actually for the full rotation group SO(3). Even though the actual shape of a body might not be rotationally invariant (like the cube or the tetrahedron) the corresponding Lagrange function is.

Of course, the lattice regularization also destroys any symmetries (or covariancies) related to scale transformations of the coordinates. From our experience with such symmetries (e.g., the massless ϕ^4 theory in 4 dimensions or pure gauge theory in 4 dimensions) we expect that the quantized theory does not realize this symmetry anyhow. We will come back to this point in the context of pure gauge theory.

However, apart from symmetries, the regularization procedures can destroy even more fundamental properties of the theory which are prerequisites in order to define a meaningful quantum theory. The most relevant properties of this kind are unitarity, locality and positivity (see below). It is already far from obvious that the euclidean formulation of a quantum field theory allows one to reconstruct a meaningful quantum theory. This problem was solved by Osterwalder and Schrader [85].

6.1 Wightman functions, τ-functions and Schwinger functions

In the 1950s and 1960s it became increasingly clear that some of the naive axioms of quantum mechanics (like the postulation of canonical commutation relations) do not hold for general quantum field theories. Haag's theorem [39, 37] explicitly states that any quantum field theory for which the canonical conjugated fields satisfy canonical commutation relations is unitarily equivalent to a free field theory. Furthermore, apart from free field theories not many examples existed of quantum field theories which could be solved and constructed exactly. This initiated axiomatic formulations of quantum field theories.

The most famous axiomatic formulation is due to Wightman (for more details see, e.g., [96]). Actually, there are two ways to formulate the Wightman axioms: (1) defining axiomatically the properties of the quantum fields (which are, technically speaking, operator valued distributions), (2) defining axiomatically the properties of the expectation values, in particular the Wightman distributions, of these quantum fields. We will take the second approach and only briefly indicate the relation.

Given a quantum field $\hat{\Phi}(x)$ (where x now refers to a point of space-time), the Wightman distributions can be defined as:

$$W(x_1, ..., x_n) = \langle \Omega | \hat{\Phi}(x_1) ... \hat{\Phi}(x_n) | \Omega \rangle \,. \tag{6.1}$$

Here x_i is a Minkowski 4-vector and $|\Omega\rangle$ is the vacuum state or, to be more precise, the state corresponding to the lowest eigenvalue of the Hamilton operator. More general, $\hat{\Phi}(x)$ can be a multi-component quantum field representing different types of fields, in which case the function W carries additional indices which we have suppressed.

As already noted, the quantum field $\hat{\Phi}(x)$ is an operator valued distribution, i.e., it is usually not defined for a definite value of x. What is defined are field operators smeared out by test functions f over $\mathbf{R}^4$ (where we do not specify here which of the many test function spaces is most appropriate):

$$\hat{\Phi}(f) = \int d^4x \, \hat{\Phi}(x) f(x) \,.$$

Therefore, also the Wightman distributions should not be considered as functions of their arguments but as mappings on test function spaces. It is sufficient to consider:

$$W(f_1, ..., f_n) = \int d^4x_1 ... d^4x_n \, W(x_1, ..., x_n) f_1(x_1) ... f_n(x_n) \,.$$

We should mention that there is a second important set of distributions which is closely related to the Wightman distributions. These are the τ-functions (which are actually distributions) which can be expressed as the time-ordered expectation values of quantum fields:

$$\tau(x_1, ..., x_n) = \langle \Omega | \mathcal{T}(\Phi(x_1) ... \Phi(x_n)) | \Omega \rangle \,. \tag{6.2}$$

The τ-functions are easily calculated from the Wightman distributions:

$$\tau(x_1, ..., x_n) = \sum_{\pi \in P_n} \theta(\{x^0_{\pi_i}\}) W(x_{\pi_1}, ..., x_{\pi_n}) \,,$$

where the sum extends over all permutations of the arguments and the θ-function is zero unless all arguments are arranged in increasing order (time-ordering). On the other hand, the Wightman distributions can be obtained from the τ-functions by analytic continuation which is usually achieved by a double Fourier transform.

Note that the (real time) functional integral approach yields time-ordered expectation values. Furthermore, the calculation of scattering amplitudes along the LSZ formalism (Lehman, Symanzik and Zimmermann [67], see any good book on quantum field theory) also requires the time-ordered τ-functions. However, the Wightman axioms are usually formulated in terms of the Wightman distributions.

6.2 The Wightman axioms

Given a set of distributions $\{W(x_1, ..., x_n)\}$ $(n = 1, 2, ...)$. (Again, we set aside all technicalities related to the nature of the test function spaces etc. For details see [96, 48].)

The Wightman axioms are:

(a) Relativistic Transformation Law

Let $(\Lambda, a) : x \to \Lambda x + a$ be a Poincaré transformation (L a Lorentz transformation and a a translation). This transformation acts on the Wightman functions:

$$W \longrightarrow W^{(\Lambda,a)} \quad \text{with} \quad W^{(\Lambda,a)}(x_1, ..., x_n) = S(\Lambda, a)W(\Lambda^{-1}x_1 - a, ..., \Lambda^{-1}x_n - a)$$

where $S(\Lambda, a)$ denotes a linear representation of $SL(2, C)$ (the covering group of the special Lorentz group) acting on "the index structure" of W. In particular, for a scalar field $S(\Lambda, a)$ is the trivial representation (the Wightman distributions are invariant).

This axiom is a consequence of the fact that the quantum fields should transform under an irreducible representation of (the covering group of) the Lorenz group and the expectation values transform under the induced tensor representation of the corresponding fields.

(b) Spectral condition

From the previous axiom is follows that $W^{(n)}(x_1, ..., x_n)$ only depends on the difference values $\xi_i = x_i - x_{i+1}$ $(i = 1, ..., n - 1)$, i.e., there are distributions $\hat{W}$ such that

$$W(x_1, ..., x_n) = \hat{W}(\xi_1, ..., \xi_{x-1})\,.$$

Define the Fourier transform of $\hat{W}$:

$$G(p_1, ..., p_{n-1}) = \int \exp\left(i \sum_{i=1}^{n-1} q_i \xi_i\right) \hat{W}(\xi_1, ..., \xi_{n-1})$$

Then the spectral condition reads:

$$G(p_1, ..., p_{n-1}) = 0$$

if any p is not in the energy momentum spectrum of the states, in particular if any p^0 (energy) is negative.

(c) Hermiticity

Let

$$W^*(x_1, ..., x_n) = \langle\Omega|\hat{\Phi}^*(x_1)...\hat{\Phi}^*(x_n)|\Omega\rangle$$

where $\hat{\Phi}^*$ refers to the hermitean conjugated field operator. Then we have:

$$W(x_1, ..., x_n) = W^*(x_n, ..., x_1)\,.$$

(d) Locality

Whenever $x_i - x_{i+1}$ is a space-like vector, we have

$$W(x_1, ..., x_i, x_{i+1}, ..., x_n) = (-1)^\epsilon W(x_1, ..., x_{i+1}, x_i, ...x_n)$$

where $\epsilon = 1$ if both arguments refer to Fermi fields and $\epsilon = 0$ otherwise.

This axiom expresses the commutation relations for fields for which the arguments are mutually in the causal complement (bosonic fields commute and fermionic fields anti-commute).

(e) Positive Definiteness Condition

Define

$$W_{jk}(x_1, ..., x_j, y_1, ..., y_k) = \langle\Omega|\hat{\Phi}^*(x_1)...\hat{\Phi}^*(x_j)\hat{\Phi}(y_1)...\hat{\Phi}(y_k)|\Omega\rangle\,.$$

For any sequence $\{f_j\}$ $(j = 1, ...)$ of test functions (where f_j is a test function on $\mathbf{R^{4j}}$) such that $f_j = 0$ except for a finite number of j's, we require:

$$\sum_{j,k=0}^{\infty} \int ... \int f_j^*(x_1, ..., x_j)\; W_{jk}(x_j, ..., x_1, y_1, ..., y_k) \times f_k(y_1, ..., y_k)\, \mathrm{d}x_1, ..., \mathrm{d}y_k \geq 0\,.$$

This condition arrises from the following consideration: Let

$$|\Psi\rangle = \sum_{k=0}^{\infty} \int \mathrm{d}x_1 ... \mathrm{d}x_k\, f_k(x_1, ..., x_k)\hat{\Phi}(x_1)...\hat{\Phi}(x_k)|\Omega\rangle \tag{6.3}$$

be an arbitrary state generated by acting with field operators $\hat{\Phi}(x)$ onto the vacuum state, then the expression above is the requirement that this state does not have negative norm. We have to allow for zero norm, because some combinations of field operators may annihilate the vacuum state. This will actually always be the case if the functions $f(x_1, ..., x_n)$ (which depend on 4-dimensional Minkowski coordinates) are not compatible with the time evolution of the system). It can be shown that a QFT is (in genral) already fixed by giving the field operators and their conjugated operators (usually time derivatives of the field operators) on one time slice.

This axiom is the most important one in reconstructing the Hilbert space of states. The Wightman functions are used to define a scalar product, and the condition above guarantees that this product is positive definite.

We shall briefly indicate the idea:
Starting point is the space of sequences of test functions (i.e. sequences $(f_0, f_1, f_2, ...)$ where each f_n is a test function on $\mathbf{R}^{4n}$) where all but finitely many functions are zero. This is an infinite dimensional vector space over $\mathbf{C}$. In order to define a hermitean scalar product on this space one uses the Wightman functions. On first defines a "pre-product":

$$\langle f_0, f_1, f_2, ... | g_0, g_1, g_2, ... \rangle = \sum_{j,k=0}^{\infty} \int ... \int f_j^*(x_1, ..., x_j) W_{jk}(x_j, ..., x_1, y_1, ..., y_k) \times g_k(y_1, ..., y_k)\, \mathrm{d}x_1, ..., \mathrm{d}y_k \,.$$

The rest are Hilber space technicalities: dividing out the states of norm zero (by taking equivalence classes) and completing the space with respect to the norm.

(f) Cluster Property

In particular for the formulation of scattering theory the following axiom is often included, even though it is not necessary for the reconstruction of the Hilbert space.

$$W(x_1, ..., x_j, x_{j+1} + \lambda a, ..., x_n + \lambda a) \longrightarrow W(x_1, ..., x_j) W(x_{j+1}, ..., x_n)$$

in the limit $\lambda \to \infty$ for any non-zero space-like vector a.

These axioms guarantee that one can reconstruct a positive definite Hilbert space, field operators acting on this Hilbert space, a unitary time evolution of these operators defined by a Hamiltonian with non-negative spectrum, a vacuum state corresponding to the lowest eigenvalue of the Hamiltonian, locality in accordance with Einstein causality, etc.

6.3 Schwinger functions and the Osterwalder-Schrader axioms

It was known since the beginning of axiomatic quantum field theory that the Wightman distributions possess far-reaching analytic properties, many of them a direct consequence of the spectral condition together with the relativistic transformation laws. The Wightman distributions allow for analytic continuations into the (multi-component) complex plane, in which case one sometimes speaks of Wightman functions. In particular, the Wightman function can be extended to the so-called Schwinger points $(it, \vec{x})$. In this way one obtains the *Schwinger functions* :

$$S((t_1, \vec{x}_1), ..., (t_n, \vec{x}_n)) = W((it_1, \vec{x}_1), ..., (it_1, \vec{x}_n)) \,. \tag{6.4}$$

The Schwinger functions are symmetric in their arguments (for the moment we talk about Schwinger functions for bosonic fields only; Schwinger functions for fermionic fields are antisymmetric). Therefore, it is a natural question whether the Schwinger functions can be represented as the moments of a probability measure for (classical) fields:

$$S(x_1, ..., x_n) = \int \phi(x_1)\phi(x_2) \cdots \phi(x_n)\, \mathrm{d}\mu[\{\phi(x)\}] \,. \tag{6.5}$$

(Here, x_i are 4-vectors in a euclidean space.)

The construction of such measures is what is done in constructive quantum field theory. The euclidean functional integral is the formal expression representing this formula and the task is to give this expression a well-defined mathematical meaning.

In order to reconstruct the Wightman distributions from the Schwinger functions, the Schwinger functions have to satisfy a certain set of conditions. These conditions are the axioms for euclidean quantum field theory and they have been formulated by Osterwalder and Schrader [85]. We will refer to them as the Osterwalder-Schrader axioms (OS-axioms).

Given a set of distributions $\{S_n = S(x_1, ..., x_n)\}$. We will write $S_n(f)$ for

$$S_n(f) = \int \mathrm{d}^4x_1 ... \mathrm{d}^4x_n \; S(x_1, ..., x_n) f(x_1, ..., x_n) \,.$$

OS0 Temperedness

The distributions S_n should be elements of the dual space of Schwartz's space of test functions which vanish with all their derivatives at coinciding points. Sometimes, for technical reasons, one requires further restrictions for the test functions.

OS1 Euclidean covariance

For a euclidean transformation (a, R), with $R \in SO(4)$, we define

$$f_{(a,R)} \sim f_{(a,R)}(x_1, ..., x_n) = f(Rx_1 + a, Rx_2 + a, ..., Rx_n + a) \,.$$

Euclidean invariance implies:

$$S_n(f) = S_n(f_{(a,R)}) \,.$$

(This is for scalar fields only. For non-scalar fields the Schwinger functions transform covariantly, i.e., like the corresponding tensor representations induced by the fields. These transformations act onto the "index strcuture" of the Schwinger functions.)

OS2 Positivity

Define:

$$\Theta f((t_1, \vec{x}_1), ..., (t_n, \vec{x}_n)) = f((-t_1, \vec{x}_1), ..., (-t_n, \vec{x}_n)) \,.$$

Then:

$$\sum_{j,k} S_{j+k}(\Theta f_j^* \times f_k) \geq 0$$

for all test functions f with support in the upper half space (positive time coordinates).

This is the most important and non-trivial condition. It is called Osterwalder-Schrader Positivity or Reflection Positivity.

OS3 Symmetry

Let π be a permutation and

$$f_\pi(x_1, ..., x_n) = f(x_{\pi_1}, ..., x_{\pi_n}) \,.$$

Then

$$S_n(f) = S_n(f_\pi) \tag{6.6}$$

So, the functions $S(x_1, ..., x_n)$ should be symmetric under permutations of their arguments. Again, this only holds for bosonic fields. For fermions the Schwinger functions should be antisymmetric.

OS4 Cluster property

$$S(x_1, ..., x_j, x_{j+1} + \lambda a, ..., x_n + \lambda a) \longrightarrow S(x_1, ..., x_j)S(x_{j+1}, ..., x_n)$$

for $\lambda \to \infty$ and a a non-zero (euclidean) vector.

Osterwalder and Schrader have shown that from these conditions one can derive a set of Wightman functions from the Schwinger functions such that the Wightman distributions satisfy the Wightman axioms. It may be interesting to note that there is no direct counterpart to the spectral condition. The reason is that the spectral condition follows automatically from OS0-OS2, in particular from the covariance condition. Locality for the Wightman functions essentially corresponds to the symmetry condition of Schwinger functions.

We now can come back to our original question whether lattice regularization does not destroy one of the fundamental requirements for a quantum field theory. Looking at the OS-axioms we see that only OS1 (euclidean invariance) and OS2 (positivity) are a problem. As mentioned before, euclidean invariance will be broken by the lattice approximation anyhow, but with strong indications that it will be restored in the continuum limit.

The central property is reflection positivity. Osterwald and Seiler [86] have shown that the usual lattice actions (which we will also use in this lecture) are reflection positive. In particular, this holds for Wilson's action for gauge fields. It has been shown that some other gauge actions (which in a naive continuum limit approach the classical gauge action) are not reflection positive (like Manton's action).

Kapitel 7

Regularization and Renormalization

As this lecture is not meant to replace a standard lecture on quantum field theory and, in particular, not the perturbative methods of quantum field theory, this chapter on regularization and renormalization can only be a brief introduction and a very sketchy overview.

Rougly speaking we can make the following brief descriptions:

Regularization replaces an ill-defined mathematical expression (e.g. the integrals corresponding to a Feynman graph) by a mathematically well-defined expression. This is achieved by introducing a new parameter a which has a physical dimension. In units where $\hbar = 1$ and $c = 1$ (which we will use in this chapter), this dimension can always be chosen to be a length.

The regularized theory is not the theory one is interested in. Strictly speaking it is a different theory. In many cases, it does not even correspond to a meaningful quantum theory, i.e. one (or several) of the axioms of QFT are violated.

Renormalization: Physically observable quantities should not depend on the arbitrary parameter a. The renormalization procedure tries to absorb all infinities which arise when a is taken to its physical limit (zero or infinity) into the (unobservable) "naked" parameters of the theory, like the coupling constants, the mass, the field normalization etc. These naked parameters in the Lagrange function have no direct physical meaning. A coupling constant is defined by the strength of a particular cross-section, a mass is defined by certain poles in the propagator etc. Renormalization adjusts certain observable quantities to the "real world" thereby eliminating the unphysical parameter a. All physical quantities are then defined in terms of these "renormalized" parameters.

Up to now this is only a program. The following sections will make this program slightly more explicit. In order to avoid certain technical problems related to the fact that the Green's function for the Klein-Gordon or Dirac equation are not uniquely defined, we will use the euclidean version of the functional integral. Furthermore, we will not care about factors of "2π" or combinatorical factors. This chapter just aims at giving an idea of how regularization and renormalization works.

7.1 Infinities in the integrals for Feynman diagrams

When one wants to calculate an n-point function (either in the Minkowski or in the euclidean version of a QFT) using the functional integral approach, the perturbation theory leads to expressions of the following type (we will use the notation of a φ^4-scalar field theory in this section, but this is no principle restriction):

$$\begin{aligned}\langle\varphi(z_1)...\varphi(z_n)\rangle &= \frac{1}{Z}\int\mathcal{D}\varphi(x)\,\varphi(z_1)...\varphi(z_n)\exp\left(-S[\varphi]\right) \qquad (7.1)\\ &= \frac{1}{Z}\sum_{k=0}^{\infty}\frac{1}{k!}\int\mathcal{D}\varphi(x)\,\varphi(z_1)...\varphi(z_n)S_{\rm int}[\varphi]^k\exp\left(-S_0[\varphi]\right)\\ &= \frac{1}{Z}\sum_{k=0}^{\infty}\frac{(-g)^k}{k!4^k}\int\prod_{j=1}^{k}\mathrm{d}^4x_j\;\int\mathcal{D}\varphi(x)\varphi(z_1)\varphi(z_n)\varphi(x_1)^4...\varphi(x_k)^4\exp\left(-S_0[\varphi]\right)\end{aligned}$$

Here $S_0[\varphi]$ corresponds to the part of the action which is at most quadratic in the fields, e.g.

$$S_0[\varphi]=\frac{1}{2}\int\mathrm{d}^4x\,[(\partial_\mu\varphi)^2+m^2\varphi(x)]=\frac{1}{2}\int\mathrm{d}^4x\,\varphi(x)(-\Delta+m^2)\varphi(x)$$

(where Δ refers to the 4-dimensional Laplacian) and

$$S_{\rm int}=\frac{g}{4}\int\mathrm{d}^4x\,\varphi(x)^4$$

is the interaction part of the action. A similar expansion has to be made for the functional integral in the denominator of eq. 7.1.

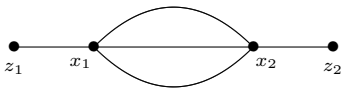
Formally, the evaluation of the n-point function has been reduced to the calculation of Gaussian integrals. As explained in the general section on Gaussian integrals, the terms in the perturbation expansion can be expressed as sums over all pairings of the set $\{z_1,...,z_n,x_1,x_1,x_1,x_1,x_2,...,x_k,x_k,x_k,x_k\}$ (notice that for the φ^4-interaction each argument x_i occurs four times). For each pair of arguments one has to insert the "inverse" operator of the quadratic form in the exponent, i.e. the Green's function corresponding to the operator $(-\Delta+m^2)$:

$$(-\Delta+m^2)^{-1}\simeq G(x,y)\quad\text{ with }\;(-\Delta+m^2)G(x,y)=\delta(x-y)$$

Let's look at a particular contribution to the 2-point function in second order of perturbation theory (neglecting any combinatorical and normalization factors):

$$C(z_1,z_2)=\int\mathrm{d}^4x_1\int\mathrm{d}^4x_2\,G(z_1,x_1)G(x_1,x_2)^3G(x_2,z_2)\,. \qquad (7.2)$$

This expression occurs in the following pairing: $\{(z_1,x_1),(x_1,x_2),(x_1,x_2),(x_1,x_2),(x_2,z_2)\}$, and it is graphically represented by the Feynman diagramm on the left hand side.



The Green's functions are very singular for coinciding arguments and this singularity makes the integral ill-defined. This is to be expected because Green's functions are actually distributions, and distributions cannot be multiplied. The singular behavior of this integral becomes more obvious in momentum space, where now

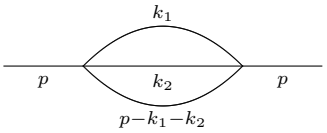
$$\tilde{G}(k) = \frac{1}{(2\pi)^2} \int \mathrm{d}^4x \, e^{ikx} G(x) \;\propto \frac{1}{k^2 + m^2} \,. \tag{7.3}$$

(Notice that $G(x, y)$ is actually only a function of the difference of the arguments, i.e. $G(x, y) \simeq G(x - y)$. Furthermore, k is, like x, a euclidean 4-vector and kx as well as k^2 denote euclidean products.)

Making the Fourier transformation in eq. 7.2 leads to an integral of the following type:

$$\tilde{C}(p) \;\propto\; \frac{1}{(p^2 + m^2)^2} \int \mathrm{d}^4k_1 \mathrm{d}^4k_2 \frac{1}{(k_1^2 + m^2)(k_2^2 + m^2)((k_1 + k_2 - p)^2 + m^2)} \tag{7.4}$$

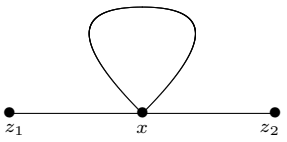
(The two integrations over x_1 and x_2 lead to momentum conservation at each of the two vertices. This makes the "initial" momentum p equal to the final momentum and eliminates k_3 in terms of the other momenta.)



In momentum space the corresponding Feynman diagramm indicates the momenta instead of the vertices.

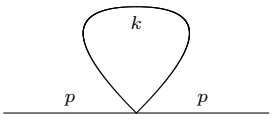
We now observe the following: There are two 4-dimensional integrals over momenta, i.e., formally this is an 8-dimensional k-integration. The product of propagators (which involve one of the k's) falls off for large arguments of k like k^{-6}. Therefore, this expression is formally "quadratically divergent".

Up to now we have considered only one particular example. There are "worse" examples, like the following contribution to the 2-point function in first order perturbation theory:



$$C'(z_1, z_2) = \int \mathrm{d}^4x \, G(z_1, x) G(x, x) G(x, z_2) \tag{7.5}$$

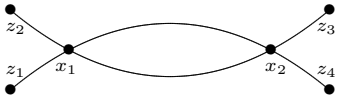
leading to an integral of the following form in Fourier space



$$\tilde{C}'(p) \;\propto\; \frac{1}{(p^2 + m^2)^2} \int \mathrm{d}^4k \, \frac{1}{k^2 + m^2} \,. \tag{7.6}$$

Again, this expression is quadratically divergent.

The following expression occurs in second order perturbation theory for the 4-point function (relevant e.g. for the cross-section of particle-particle scattering):

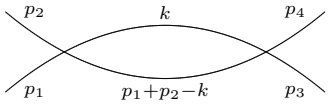


$$C(z_1, ..., z_4) = \int \mathrm{d}^4x_1 \mathrm{d}^4x_2 \, G(z_1, x_1) G(z_2, x_1) G(x_1, x_2) G(x_1, x_2) G(x_2, z_3) G(x_2, z_4) \,. \tag{7.7}$$

In Fourier space this integral leads to an expression of the following type:

$$\tilde{C}_4(p_1,p_2,p_3,p_4) \quad \propto \quad \frac{\delta(p_1+p_2-p_3-p_4)}{(p_1^2+m^2)(p_2^2+m^2)(p_3^2+m^2)(p_4^2+m^2)} \times \int \mathrm{d}^4k \frac{1}{(k^2+m^2)((p_1+p_2-k)^2+m^2)}\,. \tag{7.8}$$

This is a 4-dimensional k-integration and the integrant falls off like k^{-4} for large values of k. Hence, this integral is "logarithmically divergent".



7.2 Regularization methods

There are many regularization methods of which we will only mention a few.

7.2.1 Momentum cut-off

The most straightforward regularization of the divergent integrals which we have encountered in the last sections is the introduction of a cut-off Λ for momenta:

$$\int \mathrm{d}^4k \longrightarrow \int_{|k|<\Lambda} \mathrm{d}^4k\,.$$

This eliminates the divergencies stemming from large momenta. Λ has the dimension of a momentum, i.e. in units where $\hbar = c = 1$ it has the dimension of an inverse length. Our parameter a would correspond to $1/\Lambda$. The physical limit is $\Lambda \to \infty$. In this limit we would expect terms proportional to Λ^2 for quadratically divergent integrals and terms proportional to $\log(\Lambda/m)$ for logarithmically divergent integrals.

The cut-off method breaks Lorentz invariance.

7.2.2 Pauli-Villars

A regularization which does not break Lorenz invariance is the method of Pauli and Villars. Effectively it introduces an unphysical particle with the wrong sign of the propagator (hence violating the spectral condition). In the Feynman integrals the propagator of the scalar particle in momentum space is replaced by:

$$\frac{1}{k^2+m^2} \quad \longrightarrow \quad \frac{1}{k^2+m^2} - \frac{1}{k^2+M^2}\,. \tag{7.9}$$

The new propagator falls off for large momenta like k^{-4}, i.e. the integrals are much better behaved. (If this fall-off behavior is not sufficient, one can subtract further terms corresponding to derivatives of the propagator). The unphysical new regularization parameter is M. The physical limit is $M \to \infty$. The dimension of this parameter is that of an inverse length.

7.2.3 Dimensional regularization

A particular interesting regularization method is dimensional regularization. It's advantage is that it doesn't break gauge invariance and (at least formally) keeps the axiomatic structure of the theory intact.

All Feynman integrals can be expressed in terms of integrations over $|k|$ as well as integrations over angles. The dimension of the theory occurs only in the number of components of k. Lets consider for simplicity the following integral which occurs, e.g., in the 4-point function in the limit of vanishing external momenta:

$$C_4(m) = \int \mathrm{d}^4k \frac{1}{(k^2+m^2)^2} = \Omega(4) \int_0^\infty k^3 \mathrm{d}k \, \frac{1}{(k^2+m^2)^2} \,. \tag{7.10}$$

Here $\Omega(4)$ denotes the surface of a 4-dimensional unit sphere (the result of the integrating over the angles) and the remaining integral is over the absolute value of k only. The information about the dimension of the theory is only contained in Ω and in the power of k^3.

Notice that formally the right hand side is dimensionless. As m is the only parameter on which the result can depend and because m does have a dimension (inverse length), formally the left hand side cannot depend on m either and must, therefore, be a constant. However, the integral is not defined and, therefore, such formal considerations may not hold.

In d space-time dimensions the integral would read:

$$C_d(m) = \Omega(d) \int_0^\infty k^{d-1} \mathrm{d}k \, \frac{1}{(k^2+m^2)^2} \,. \tag{7.11}$$

$\Omega(d)$ now is the surface of a d-dimensional unit sphere,

$$\Omega(d) = 2\frac{\pi^{d/2}}{\Gamma(d/2)} \,.$$

Integrals of this type can be calculated by making the substitution

$$x = \frac{m^2}{k^2+m^2} \,,$$

and using the general integration formula:

$$\int_0^1 \mathrm{d}x \, x^{\alpha-1}(1-x)^{\beta-1} = B(\alpha,\beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)} \,.$$

Finally, we obtain:

$$\int \mathrm{d}^dk \frac{1}{(k^2+m^2)^2} = \pi^{d/2} m^{d-4} \Gamma(2-d/2) \,.$$

For $d = 4$ we formally obtain $\Gamma(0)$ on the right hand side which is not defined. This was to be expected. However, the Gamma-function is defined also for non-integer values of d. In this sense the integral has been regularized.

An important question arises at this point: Where is the regularization parameter which has a dimension? The parameter d which generalizes the dimension of the theory is dimensionless.

The problem is that the integral (eq. 7.11) in $d \neq 4$ dimensions has a dimension while it is dimensionless in 4 dimensions. In order to remain the correct physical dimension, we have to introduce a parameter μ with the dimension of a mass (or inverse length), such that the regularized integral reads:

$$C_d(m) = \mu^{4-d}\,\Omega(d)\int_0^\infty k^{d-1}\mathrm{d}k\,\frac{1}{(k^2+m^2)^2} \tag{7.12}$$

$$= \mu^{4-d}\pi^{(d/2)}(m)^{d-4}\Gamma(2-d/2)\,. \tag{7.13}$$

This integral now is formally dimensionless, like the integral in eq. 7.10.

The physical limit would be $d = 4$. So we introduce the parameter $\epsilon = 4 - d$ and expand the expression above in terms of this parameter. We have to use the Γ-function for small values of ϵ:

$$\Gamma(\epsilon) = \frac{1}{\epsilon} - \gamma_E + \mathcal{O}(\epsilon)$$

(where $\gamma_E \approx 0,577...$ is Euler's constant.) Hence, we obtain:

$$\int \mathrm{d}^d k \frac{1}{(k^2+m^2)^2} \stackrel{d\to 4}{=} \pi^2\left(\frac{1}{\epsilon} - \ln(m/\mu) - \gamma_E + \mathcal{O}(\epsilon)\right) \tag{7.14}$$

So the result consists of a divergent part $(1/\epsilon)$ and a finite part $(\ln(m/\mu))$ which depends on the regularization parameter μ.

7.2.4 Other analytical methods

There exist many regularization methods which make use of analytical continuations. We just mention one more where the integral (eq. 7.10) is replaced by:

$$C_4(m) = \int \mathrm{d}^4 k \frac{1}{(k^2+m^2)^2} \longrightarrow \mu^{4s-4}\int \mathrm{d}^4 k\,\frac{1}{(k^2+m^2)^{2s}}\,. \tag{7.15}$$

For s sufficiently large this integral is well-defined and can be calculated using the general formula

$$\int \frac{\mathrm{d}^d k}{(2\pi)^d}\frac{k^{2\alpha}}{(k^2+m^2)^\beta} = \frac{1}{(4\pi)^{d/2}\Gamma(d/2)}(m^2)^{d/2+\alpha-\beta}\frac{\Gamma(\beta-\alpha-d/2)\Gamma(d/2+\alpha)}{\Gamma(\beta)}\,.$$

The general prescription of this method replaces the scalar propagator in momentum space by the corresponding expression raised to the power s, multiplied with the regularization parameter μ^{2s} so that the physical dimension of the propagator remains unchanged:

$$\frac{1}{k^2+m^2} \longrightarrow \mu^{2s-2}\frac{1}{(k^2+m^2)^s}\,.$$

The physical value for s is $s = 1$. And like in the case of dimensional regularization the introduction of a general power s changes the dimension of the integral and one has to introduce the regularization parameter μ.

7.3 Renormalization

After having regularized the theory one has to make sure that the physically observable quantities do not depend on the arbitrary regularization parameter. This is done by using "renormalization".

There is a general impression that renormalization only occurs in the context of divergencies. This is not entirely true. The renormalization analysis has also been proven to be very successful in the context of statistical mechanics. In the next section, we will consider a trivial example (the one-dimensional harmonic chain) in order to explain the idea of renormalization.

The general idea is the following: Given a phenomenological theory with observable parameters. (One may think, e.g., of water which is described by a Navier-Stokes equation and the relevant macroscopic parameters are the density ρ, the pressure p, the viscosity η. For the more general case of a compressible fluid one may have to deal also with the compressibility κ.) One can make a microscopic model for this phenomenological theory which depends on certain microscopic (unobservable) parameters. Again, taking the example of water, one might think of a molecular model with a the average distance between molecules, m their mass, N their total number, and a set of parameters $\{\sigma_i\}$ characterizing the intermolecular forces (like a generalized Lennard-Jones potential).

In a specific model, the macroscopic parameters will be functions of the microscopic parameters, and in order to evaluate the microscopic parameters in terms of observable parameters, one has to invert these functions. However, it is not to be expected that the macroscopic parameters fix the microscopic parameters. Otherwise such a model would yield the atomic parameters a, m and N in terms of macroscopic observables. The determination of atomic distances, Avogadro's number or the mass of single atoms from ordinary macroscopic parameters is not possible. It turns out that a change of the parameter for the interatomic distance a, for example, can be compensated by a corresponding change in the other microscopic parameters, such that the macroscopic parameters remain unchanged.

We will now discuss a trivial example for a "renormalization" (the harmonic chain) and then briefly mention the main steps of the renormalization procedure in quantum field theory.

7.3.1 The harmonic chain

Suppose, we want to describe the oscillations of a solid rod of length L. The rod has a certain mass M, and the equation of motion of oscillations contain a parameter c (the velocity of waves along the rubber band):

$$\frac{\partial^2}{\partial t^2}\varphi(x,t) = \frac{1}{c^2}\frac{\partial^2}{\partial x^2}\varphi(x,t)\,. \tag{7.16}$$

$\varphi(x,t)$ is the deviation from the equilibrium position of the rod at point x at time t, and for simplicity we only take into account the longitudinal deviations. In addition to this equation which describes the propagation of waves along the rod (sound waves or, in a quantized version, phonons), we expect a harmonic oscillator potential (or, equivalently, Hooke's law) for the resistance against stretching (or pressing):

$$V = \frac{1}{2}Rx^2 \quad\Longrightarrow\quad F = Rx\,, \tag{7.17}$$

which defines the "spring constant" R. (Usually in material science one considers the compressibility $K = 1/R$, and in order to deal with a material constant, one would consider the compressibility per unit length: $\kappa = K/L = 1/(RL)$).

We have several macroscopic parameters: M (the mass of the rubber band), L (its length), c the propagation velocity of waves, and R the spring constant.

In a next step, we want to make a microscopic model for the rod in terms of a harmonically coupled linear chain. The microscopic equation of motion would be:

$$m\frac{\mathrm{d}^2}{\mathrm{d}t^2}\phi_i(t) = D(\phi_{i+1}(t) - 2\phi_i(t) + \phi_{i-1}(t))$$

$\phi_i(t)$ now is the deviation of the i-th particle from its equilibrium position (we only take into account longitudinal deviations), and $i = 1, ..., N$, where N is the total number of particles.

This model contains several microscopic paramters: m (the mass of the single particles along the chain), D (the spring constant for the harmonic force between neighboring particles), N (the total number of particles), and a (the distance between adjacent particles). The parameter a is only implicit, but we need this parameter in order to adjust the length of the chain to the length L of the rod.

The relation between the microscopic parameters and the macroscopic parameters is obvious:

$$L = aN \ , \quad M = Nm \ , \quad c = a\sqrt{\frac{D}{m}} \ , \quad \text{and} \quad R = \frac{D}{N} \, . \tag{7.18}$$

(For L and M it is trivial, for c we have to observe that the square root denotes a frequency (dimension 1/sec) and a a length. For R we have to convince ourselves that the same force which stretches a single spring by a distance x, stretches two springs "in series" by twice the amount, i.e. the combined string constant is half of the original ones, or, more general, string constants combine according to $1/D_{\mathrm{tot}} = 1/D_1 + 1/D_2$.)

In this case, there are as many microscopic parameters as macroscopic parameters. At first sight it seems that we can invert these equations and determine the microscopic parameters in terms of the macroscopic parameters. However, from physical considerations one would not expect this to be possible.

A closer inspection reveals that the macroscopic parameters are invariant under the following combined transformations of the microscopic parameters:

$$a \to \lambda a \ , \ N \to \frac{N}{\lambda} \ , \ m \to \lambda m \ , \ D \to \frac{D}{\lambda} \, . \tag{7.19}$$

This set of transformations can be called the renormalization group of the model. Formally, we cannot set $a = 0$, but we have to consider a limit, where $a \to 0$. In this limit the other microscopic quantitities become $N \to \infty$, $m \to 0$ and $D \to \infty$. So, all microscopic parameters assume "singular" values in the continuum limit. This does not imply, however, that there is anything "singular" or ill-defined with respect to the rod. It is related to the general problem of going from discrete variables to continuous variables (actually densities). Actually, what is important is not the limit $a \to 0$ (and the corresponding other limits for the microscopic quantities) but the fact that the macroscopic quantities do not depend on the scale λ. Therefore, one often speaks of a "scaling limit" instead of a "continuum limit".

The fact that we have as many macroscopic parameters as microscopic parameters on the one hand, but that there is a one-parameter group of transformations of the microscopic parameters which leaves the macroscopic parameters invariant on the other hand, leads to the conclusion that the macroscopic parameters cannot be independent. Indeed, the microscopic model makes the prediction that

$$R = c^2 \frac{M}{L^2} \,,$$

which is the observed macroscopic relation among these parameters.

In order to make contact with the formulas in the literature, we can express c in terms of the compressibility $\kappa = 1/(RL)$ and the linear mass density $\rho = M/L$ and obtain the well-known relation:

$$c = \frac{1}{\sqrt{\kappa\rho}} \,.$$

7.3.2 Renormalization in QFT

The regularized version of a quantum field theory can be compared to our microscopic model. There are certain microscopic parameters, like m (the "bare" mass of the theory), g (the coupling constant), the normalization Z of the field $\varphi(x)$ (about which we say more later) and the regularization parameter a (which can be $1/\Lambda$ for the cut-off method, $1/M$ for Pauli-Villars, or $1/\mu$ in the analytic regularization procedures). So we can say that the regularized theory depends on $\{m, g, Z, a\}$. These quantities are not observable.

On the other hand, one can observe (at least in principle) the n-point functions and their Fourier transforms, like the two-point function $\tilde{G}_2(p)$ or the four-point function $\tilde{G}_4(p_2 - p_1, p_3 - p_2, p_4 - p_3)$. The first step is to associate the physical parameters which correspond most closely to the parameters of the model with these observable functions. We shall indicate these renormalization conditions for the φ^4 theory. Each theory has its own parameters and its own observables, but the principle idea should become clear.

The field normalization: The classical theory is expressed in terms of a field $\varphi(x)$. The classical (euclidean) action is given by:

$$S[\varphi] = \int \mathrm{d}^4x \, \left(\frac{1}{2}(\partial_\mu \varphi(x))^2 + \frac{m^2}{2}\varphi(x)^2 + \frac{g}{4}\varphi(x)^4 \right) . \tag{7.20}$$

We have fixed the normalization of the field $\varphi(x)$ by the first term in this expression. For the second and the third term, a rescaling $\varphi \to \sqrt{Z}\varphi$ can be absorbed in a rescaled mass and a rescaled coupling constant, but for the first term we have chosen the factor to be 1.

For the full theory, the observable quantities are the n-point functions. A rescaling of the field, $\varphi \to \sqrt{Z}\varphi$ leads to a rescaling of the n-point function $G(z_1, ..., z_n) \to Z^{n/2} G(z_1, ..., z_n)$. (The square root is convention.) How can we fix this field normalization Z?

For the free theory ($g = 0$), the inverse 2-point function in Fourier space is of the form

$$(\tilde{G}_0(p))^{-1} = p^2 + m^2 \,.$$

The behavior of the full 2-point function $\tilde{G}(p)$ will be more complicated and, in particular, not only depend on the mass m but also on on the coupling constant g and on the regularization

parameter a. We now can use the same normalization condition for the field in the full theory, by requiring

$$\frac{\partial^2}{\partial p^2}(\tilde{G}_{\rm ren}(p))^{-1}\Big|_{p=0} = 1\,. \tag{7.21}$$

This condition fixes the field normalization.

The physical mass: We can also use the 2-point function to define the renormalized mass. After having fixed the field renormalization, we can require that

$$\tilde{G}_{\rm ren}(p)^{-1}\Big|_{p^2=-M^2} = 0\,. \tag{7.22}$$

So, we look for the closest pole of $\tilde{G}(p)$ in the complex p-plane, and this corresponds to the physical mass M of the particle. (Actually, in the Minkowski version of the quantum field theory, this pole will be on the real axis.) M is the mass which we observe in experiment, and the condition above defines a relation between the bare parameters m, g, and a and this physical mass M.

The physical coupling constant: A coupling constant characterizes the strength of an interaction. This interaction is expressed in terms of the "connected" 4-point function:

$$\begin{aligned} G_c(z_1,z_2,z_3,z_4) &= G_4(z_1,z_2,z_3,z_4) - G_2(z_1,z_2)G_2(z_3,z_4) - \\ &\qquad G_2(z_1,z_3)G_2(z_2,z_4) - G_2(z_1,z_4)G_2(z_2,z_3)\,. \end{aligned}$$

For a free theory, this 4-point function vanishes.

In lowest order (first order) perturbation theory, this function is given by:

$$G_c(z_1,z_2,z_3,z_4) = 6g\int {\rm d}^4x G_0(z_1,x)g_0(z_2,x)G_0(x.z_3)G_0(x,z_4)\,. \tag{7.23}$$

(The factor $6 = 3!$ is a combinatorical factor: There are 4! posibilities to connect the external points to the four edges coming from the internal vertex x, and we have defined the interaction term by a factor $g/4$.)

In momentum space this condition becomes:

$$\tilde{G}_c(p_1,...,p_4) = 6g\prod_{i=1}^4 \frac{1}{p_i^2+m^2}\delta(p_1+p_2-p_3-p_4)\,. \tag{7.24}$$

Defining the *truncated* 4-point function by:

$$\Gamma^4(p_1,...,p_4) = \left(\prod_{i=1}^4 \tilde{G}^2(p_i)^{-1}\right)\tilde{G}_c^4(p_1,...,p_4) \tag{7.25}$$

we can express this relation by

$$\Gamma^4(\{p_i\}) = 6g\delta(p_1+p_2-p_3-p_4)\,. \tag{7.26}$$

When we calculate this quantity in higher orders of perturbation theory, the dependence on the momenta (as well as on m and a) will be non-trivial. Therefore, one defines a particular renormalization point $(p_1^{\rm ren}, ..., p_4^{\rm ren})$ (e.g., all $p = 0$) and sets:

$$g_{\rm phys} = \frac{1}{6}\Gamma^4(\{p_i\})|_{p_i=p_i^{\rm ren}} \,. \tag{7.27}$$

Of course, for the higher orders this definition of the physical coupling constant will depend on the chosen renormalization point, i.e., on the particular momenta, for which the 4-point function is evaluated. A different choice for the momenta will lead to a different physical coupling. This dependency of the physical coupling constant on the momentum transfer of the 4-point function is essentially the so-called *running coupling constant*.

We started with four free "microscopic" parameters (m,g,Z,a) and we defined three renormalization conditions (eqs. 7.21, 7.22, 7.27). There remains a one-parameter family of renormalizations of the microscopic parameters such that these three physical conditions remain unchanged.

Of course, the big question remains: Can we express all physical observables in terms of the physical quantities, or, in other words, can we eliminate all microscopic parameters from the theory in terms of the physical quantities. If this is possible, we will call the theory renormalizable.

Up to now we have encountered divergencies in the 2-point function and in the 4-point function (and for the denominator contributions without any external legs). These divergencies can be absorbed in the renormalized quantities defined above. The question of renormalizability essentially leads to the question, whether there are divergencies in the 6-point function (or higher point functions) which are not due to 2-point or 4-point subgraphs. In this case one has to introduce a new coupling constant related to an effective 6-point interaction. One can easily show that if this happens, the theory is non-renormalizable. In this case there will also be effective 8-point, 10-point etc. interactions and one has to introduce an infinite number of additional coupling constants. In this case, the theory is (perturbatively) non renormalizable.

7.4 Perturbative renormalizability - power counting

In order to determine the degree of divergence of an integral, we made use of the asymptotic behavior of the Fourier transform of the Green's function of a scalar particle:

$$\text{scalar particle} \quad \tilde{G}(k) \propto \frac{1}{k^2 + m^2} \overset{k\to\infty}{\longrightarrow} \mathcal{O}(k^{-2}) \,. \tag{7.28}$$

One can easily show that for a spin-1/2 fermion (Dirac particle) one obtains

$$\text{spin } 1/2 \quad \tilde{S}(k) \propto \frac{\gamma_\mu k^\mu + m}{k^2 + m^2} \overset{k\to\infty}{\longrightarrow} \mathcal{O}(k^{-1}) \,, \tag{7.29}$$

and for a massless spin-1 particle

$$\text{spin 1, massless} \quad \tilde{G}_{\mu\nu}(k) \propto \delta_{\mu\nu} - \frac{k_\mu k_\mu}{k^2} \overset{k\to\infty}{\longrightarrow} \mathcal{O}(1) \,. \tag{7.30}$$

(The Green's function for a massive spin-1 particle has the same asymptotic behavior.) This asymptotic behavior of the Green's functions is independent of the dimension of the theory. By using simple power counting arguments, we find the following rules of thumb:

1. The degree of divergence of Feynman integrals gets worse with increasing spin.
2. The degree of divergence of Feynman integrals gets worse with increasing dimension (the asymptotic behavior of the propagators remains unchanged, but each momentum integration involves d variables).
3. The degree of divergence of Feynman integrals gets worse with increasing powers of interactions.

While the first two rules are obvious from the preceeding remarks, the last rule needs some additional comments.

Let us calculate the so-called *superficial degree of divergence* D_f of an arbitrary graph. For simplicity, we only deal with a single field (scalar, Fermi- or massless spin-1) and a single interaction of the form φ^p.

Let L be the number of external lines of a graph, V be the number of internal vertices, and p the degree of a single vertex. This means that there are pV internal "hooks". L of these hooks combine to the external lines, the rest leads to internal lines. Therefore, such a graph has $(pV - L)/2$ internal lines (each line has two hooks).

In momentum space, there is a k-integration for each internal line, and in d dimensions this leads to a total of $(pV - L)d/2$ integration variables. However, each vertex contributes a δ-function for momentum conservation, of which one δ-function takes care of the overall external momenta. This implies that $d(V - 1)$ integrations are eliminated by momentum conservation. This leaves us with

$$\text{independent integrations } = \frac{d}{2}(pV - L) - d(V - 1)$$

Now let us compare this quantity with the asymptotic fall-off of the integrand. Define k to be 2 (for scalars), 1(for fermions), and 0 (for spin 1 fields). The asymptotic behavior of the $(pV - L)/2$ internal propagators is

$$\text{asymptotic fall-off power } = \frac{k}{2}(pV - L)\,.$$

The superficial degree of divergence is defined as the difference between these two quantities:

$$D_f = \frac{d}{2}(pV - L) - d(V - 1) - \frac{k}{2}(pV - L) \tag{7.31}$$

$$= \left(\left(\frac{d-k}{2}\right)p - d\right)V + \left(d - \left(\frac{d-k}{2}\right)L\right)\,. \tag{7.32}$$

As V (the number of internal vertices) can become arbitrary large, we should require the factor in front of V to be 0 or negative, otherwise the superficial degree of divergence can become arbitrarily large with increasing orders of perturbation theory. This leads us to the following inequalities:

$$\left(\frac{d-k}{2}\right)p \le d \quad \text{or} \quad p \le \frac{2d}{d-k}\,. \tag{7.33}$$

For scalar particles ($k = 2$) and $d = 4$ we obtain $p \leq 4$. Any larger value of p would lead to a non-renormalizable theory (unless "magical" cancellations occur, e.g. due to symmetries). For spin-1/2 Fermions ($k = 1$) and $d = 4$ the situation is worse: even $p = 3$ would be excluded (which, however, would not be a meaningful theory, because the number of fermions has to be even in each interaction term). What we have not shown is that a vertex with 2 fermions and 1 boson (Yukawa coupling) leads to a renormalizable theory. For spin 1 particles we obtain $p = 2$, i.e., only the free theory is renormalizable.

This formula also shows that the situation gets worse for larger dimensions. For $d = 6$, only a φ^3 scalar theory would be perturbatively renormalizable. For $d = 3$, an interaction with $p = 6$ is allowed for scalar particles. In $d = 2$ dimensions, the scalar field theory is renormalizable for arbitrary powers of the interaction. Furthermore, the 4-Fermi interaction is renormalizable for $d = 2$.

These arguments are based on pure power counting. They do not include any combinatorics of the perturbation expansion of the theory under consideration. It can happen, that the leading singularities cancel in each order of perturbation theory and that a theory, which is formally non-renormalizable, actually can be renormalized. An example is Yang-Mills theory (which by simple power counting arguments appears to be non-renormalizable). In this case, the symmetry of the theory guarantees that certain singularities cancel each other. Supersymmetry leads to further cancellations which implies, that superficially non-renormalizable theories can actually be renormalizable.

Furthermore, all these arguments are based on perturbation theory. We will show in our discussion of scalar theories on the lattice, that certain perturbatively non-renormalizable theories (like φ^4 in $d = 6$) can actually be defined non-perturbatively and even correspond to free theories. Perturbative non-renormalizability does not exclude that a theory cannot be defined by non-perturbative methods.

Kapitel 8

Das freie Klein-Gordon Feld auf dem Gitter

8.1 Allgemeine Vorbemerkungen

8.1.1 Einige grundlegende und elementare Begriffsklärungen

Unter einem (d-dimensionalen) *Gitter* werde ich in dieser Vorlesung immer ein reguläres, (hyper-)kubisches Gitter verstehen. Es besteht aus einer Menge $P \subset \mathbf{Z}^d$ von *Punkten* (z.B. $P = \{(m_1, \ldots, m_d) \mid |m_i| < L/2\}$, dabei dient L als ein Volumen Cut-off), einer Menge E von Linien (d.h. Paare von Punkten, die sich nur in einer Koordinate um ± 1 unterscheiden), sowie Plaquetten etc.

Lokale Zustände werden durch eine Menge V indiziert. Beispielsweise ist $V = \{+1, -1\}$ beim Ising-Modell, $V = \{0, 1, 2\}$ beim Potts-Modell oder $V = \mathbb{R}$ für ein reelles skalares Feld.

Eine *Konfiguration* ist eine Festlegung von lokalen Zuständen auf den Gitterpunkten, manchmal auch Gitterlinien, d.h. eine Konfiguration ist eine Abbildung:

$$C \,:\, P \,\to\, V \qquad \text{oder} \qquad C \,:\, E \,\to\, V$$

Einer Konfiguration wird eine „Energie“ (bzw. in der Sprechweise der Quantenfeldtheorie, eine „Wirkung“) $S[C, \{g\}]$ zugeordnet, die von verschiedenen Kopplungskonstanten bzw. Parametern $\{g\}$ abhängen kann. Eine solche Energie definiert für jede Konfiguration einen „Boltzmann-Faktor“ $\simeq \exp(-S[C; \{g\}])$ (die in der statistischen Mechanik noch vorhandene Temperatur kann in die Kopplungen $\{g\}$ absorbiert werden).

Die *Zustandssumme* erhält man durch Summation der Boltzmann-Faktoren über alle Konfigurationen:

$$Z[\{g\}] \;=\; \sum_C \exp(-S[C; \{g\}]) \,.$$

Normiert — d.h. nach Division durch die Zustandssumme — erhält man aus den Boltzmann-Faktoren eine Wahrscheinlichkeitsverteilung auf der Menge der Konfigurationen:

$$\omega[C] \;=\; \frac{1}{Z} \exp(-S[C; \{g\}]) \,.$$

Unter den *Observablen* verstehen wir beliebige Funktionen über den Konfigurationen:

$$F \ : \ C \to \mathbb{R} \, .$$

(Manchmal ist es günstiger, beliebige komplexwertige Funktionen zuzulassen.) Wie in der statistischen Mechanik lassen sich Erwartungswerte von Observablen berechnen:

$$\langle F[C] \rangle \ = \ \sum_C \omega[C] \, F[C] \ = \ \frac{1}{Z} \sum_C F[C] \, \exp(-S[C;\{g\}]) \, .$$

Im Sinne der Wahrscheinlichkeitstheorie bilden die Konfigurationen die Menge der Elementarereignisse. Observable, d.h. Funktionen von C – insbesondere auch der Wert von C an bestimmten Gitterpunkten oder -linien – bilden die Menge der Zufallsvariablen. Die Funktion

$$P_F(x) \ = \ \sum_C \omega[C] \, \Theta(x - F[C])$$

ist die Verteilungsfunktion zu der Zufallsvariablen F. Sie gibt an, mit welcher Wahrscheinlichkeit $F[C]$ kleiner als x ist. Die Distribution

$$n_F(x) \ = \ \frac{\mathrm{d}P}{\mathrm{d}x}$$

ist die Wahrscheinlichkeitsverteilung von F ($n_F(x)\mathrm{d}x$ ist die Wahrscheinlichkeit, dass F zwischen x und $x + \mathrm{d}x$ liegt).

In einer algebraischen Sprechweise bilden die (komplexwertigen) Observablen eine C^*-Algebra („konfigurationsweise“ Addition und Multiplikation, Norm $\|F\| = \max_C |F(C)|$). Ein Zustand ist durch ein lineares, positives, normiertes Funktional auf dieser Algebra gegeben. Jede Wahrscheinlichkeitsverteilung $\omega[C]$ definiert in diesem Sinne einen Zustand oder auch Erwartungswertfunktional.

8.1.2 Randbedingungen

In der Gitterregularisierung des euklidischen Funktionalintegrals wird die d-dimensionale euklidische Raum-Zeit $\mathbb{R}^d$ durch ein d-dimensionales hyperkubisches Gitter

$$\mathbb{Z}^d \ = \ \{(m_1, \ldots, m_d) | m_i \in \mathbb{Z}\}$$

ersetzt. Dieses Gitter wird oft sowohl aus theoretischen als auch aus praktischen Gründen zunächst als „endlich” angenommen, d.h. man nimmt eine lineare Ausdehnung der „Länge” L an: $1 \leq m_i \leq L$. (Diese Länge kann für jede Gitterrichtung verschieden sein.) Da in der Wirkung $S[C;\{g\}]$ im Allgemeinen Wechselwirkungen von lokalen Zuständen an benachbarten (oder auch weiter entfernten) Gitterplätzen auftreten, müssen diese Wechselwirkungen an den Rändern besonders definiert werden. Das geschieht beispielsweise durch die Wahl von Randbedingungen. Die häufigsten Randbedingungen sind:

1. Periodische Randbedingungen:

$$\varphi(m_1, \ldots, m_i + L, \ldots, m_d) \ = \ \varphi(m_1, \ldots, m_i, \ldots, m_d) \quad \forall i = 1, \ldots, d$$

2. Feste Randbedingungen: $\varphi(\{m_i\})|_{m_j = 1 \text{ oder } L} \ = \ \text{const.}$

3. Offene Randbedingungen:
   In diesem Fall ist die Wechselwirkung der Felder mit den Nachbarfeldern am Rand des Gitters abzuändern, sodass nur Wechselwirkungen innerhalb des Gitters bestehen. Bei Festkörpern sind diese Randbedingungen die natürlichen, allerdings können spezielle Oberflächenphänomene auftreten.

4. Antiperiodische Randbedingungen:

$$\varphi(m_1, \ldots, m_i + L, \ldots, m_d) \;=\; -\,\varphi(m_1, \ldots, m_i, \ldots, m_d)$$

   Diese dienen oft dazu, unerwünschte langreichweitige Korrelationen zu unterbinden oder aber künstlich Grenzflächen zwischen korrelierten Gebieten zu erzeugen.

5. Helikale Randbedingungen:
   Während üblicherweise ein Punkt in einem d-dimensionalen hyperkubischen Gitter mit Gitterlänge L durch ein d-Tuppel von ganzen Zahlen gekennzeichnet wird, kann man diesen Punkt auch durch eine einzelne Zahl charakterisieren:

$$(m_1, \ldots, m_d) \;\;\rightarrow\;\; i \;=\; (m_1 + Lm_2 + L^2 m_3 + \ldots + L^{d-1} m_d)\,.$$

   Die nächsten Nachbarn des Punktes i sind nun:

$$i \pm 1\ ,\ i \pm L\ ,\ i \pm L^2\ , \ldots,\ i \pm L^{d-1}\ ,$$

   wobei noch die „Periodizitätsbedingung“

$$\varphi(i + L^d) \;=\; \varphi(i)$$

   zu berücksichtigen ist. Im Vergleich zu den periodischen Randbedingungen erhält man hier eine „schraubenförmige Versetzung“ an den Rändern. Diese Wahl der Randbedingungen hat meist numerische Gründe. Die DO-Schleifen durch das Gitter bei Monte-Carlo Simulationen werden auf eine einzelne DO-Schleife reduziert, was gerade bei vektorisierenden Maschinen bei geschickter Programmierung zu erheblichen Zeitersparnissen führen kann.

Ich werde im Folgenden nahezu ausschließlich periodische Randbedingungen benutzen. Diese sind mathematisch am einfachsten zu behandeln. Außerdem wird man erwarten, dass im thermodynamischen Limes ($L \rightarrow \infty$) die Randbedingungen für den thermodynamischen Zustand keine Rolle spielen.

Diese Aussage stimmt nicht ganz, wenn das System für bestimmte Parameterwerte (einschließlich der Temperatur) mehrere Grundzustände haben kann (in algebraischer Sprechweise würde man von mehreren KMS-Zuständen sprechen). Das ist beispielsweise beim Ising-Modell bei tiefen Temperaturen in der gebrochenen Phase der Fall. In diesem Fall hängt die Spinorientierung auch im thermodynamischen Grenzfall von den Randbedingungen ab, die man auf dem endlichen Gitter gewählt hat.

8.2 Die gitterregularisierte Klein-Gordon Wirkung

Die euklidische Wirkung des freien Klein-Gordon Feldes

$$S \;=\; \frac{1}{2}\int \mathrm{d}^d x\ \left(\sum_{\mu=1}^{d} (\partial_\mu \varphi(x))^2 \;+\; m^2 \varphi(x)^2\right)$$

soll nun durch ein Funktional der Felder auf dem Gitter ersetzt werden. Der naheliegendste Ansatz ist, die partiellen Ableitungen durch Differenzen der Felder an benachbarten Gitterpunkten zu ersetzen:

$$S \;=\; \frac{1}{2}\sum_{\{m_i\}}\left(\sum_{\mu=1}^{d}(\varphi(\boldsymbol{m}+\boldsymbol{e}_\mu)-\varphi(\boldsymbol{m}))^2 \;+\; M^2\varphi(\boldsymbol{m})^2\right) \tag{8.1}$$

$\boldsymbol{e}_\mu$ bezeichnet den Vektor, der in der μ-Komponente 1 ist, ansonsten 0. Somit ist $\boldsymbol{m}+\boldsymbol{e}_\mu$ der nächste Nachbar von $\boldsymbol{m}$ auf dem Gitter in Richtung μ.

Diese Wirkung lässt sich umschreiben zu

$$S \;=\; \frac{1}{2}\sum_{\{m_i\}}\left(\sum_{\mu=1}^{d}-2\varphi(\boldsymbol{m}+\boldsymbol{e}_\mu)\varphi(\boldsymbol{m}) \;+\; (M^2+2d)\varphi(\boldsymbol{m})^2\right) \tag{8.2}$$

$$\;=\; \frac{1}{2}\sum_{m,n}\varphi(\boldsymbol{m})K(\boldsymbol{m},\boldsymbol{n})\varphi(\boldsymbol{n})\,, \tag{8.3}$$

mit

$$K(\boldsymbol{m},\boldsymbol{n}) \;=\; M^2\,\delta_{m,n}-\sum_{\mu=1}^{d}\left(\delta_{m,n+e_\mu}-2\delta_{m,n}+\delta_{m,n-e_\mu}\right) \tag{8.4}$$

$$\;=\; (M^2+2d)\,\delta_{m,n}-\sum_{\mu=1}^{d}\left(\delta_{m,n+e_\mu}+\delta_{m,n-e_\mu}\right)\,. \tag{8.5}$$

(Im zweiten Teil der oberen Formel erkennt man den zweiten Differenzenquotienten wieder, der im diskreten Fall der zweiten Ableitung entspricht.) Man kann die Felder $\varphi(\boldsymbol{m})$ als Vektoren in einem L^d-dimensionalen Vektorraum auffassen. $K(\boldsymbol{m},\boldsymbol{n})$ ist eine lineare Abbildung, also eine Matrix, über diesem Vektorraum. Außerdem ist K symmetrisch und positiv (da die Wirkung S positiv ist für alle $\varphi\neq 0$ und $M\neq 0$).

Nach den allgemeinen Überlegungen zu Gaußschen Integralen folgt für die 2-Punkt-Funktion des Klein-Gordon Feldes auf dem Gitter:

$$\langle\varphi(\boldsymbol{m})\varphi(\boldsymbol{n})\rangle \;=\; (K^{-1})(\boldsymbol{m},\boldsymbol{n})\,. \tag{8.6}$$

Wir haben also die Matrix K zu invertieren. Ganz allgemein kann man das Inverse einer hermiteschen Matrix K in der Form:

$$(K^{-1})(\boldsymbol{m},\boldsymbol{n}) \;=\; \sum_k \frac{f_k^*(\boldsymbol{m})\,f_k(\boldsymbol{n})}{\lambda_k}$$

schreiben, wobei λ_k der k-te Eigenwert von K ist und $f_k(\boldsymbol{n})$ die $\boldsymbol{n}$-te Komponente des zugehörigen Eigenvektors ist. Zur Bestimmung der Eigenvektoren und Eigenwerte von K nutzen wir aus, dass K translationsinvariant ist:

$$K(\boldsymbol{m},\boldsymbol{n}) \;=\; K(\boldsymbol{m}+\boldsymbol{a},\boldsymbol{n}+\boldsymbol{a}) \qquad\qquad \boldsymbol{a}\in(\mathbf{Z}/L\mathbf{Z})^d$$

Diese Aussage ist äquivalent dazu, dass mit jedem Eigenvektor $f(\boldsymbol{n})$ zu K auch $f(\boldsymbol{n}+\boldsymbol{a})$ Eigenvektor ist, bzw. dass die Matrix K mit dem Translationsoperator P_a kommutiert. Beide

Operatoren haben also dieselben Eigenvektoren. Die Eigenvektoren des Translationsoperators erfüllen die algebraische Gleichung

$$f(\boldsymbol{n}+\boldsymbol{a}) \;=\; \kappa(\boldsymbol{a}) f(\boldsymbol{n})\,.$$

Die einzigen Lösungen sind Exponentialfunktionen. Wegen der Periodizitätsbedingung folgt, dass

$$f_k(\boldsymbol{n}) \;=\; \frac{1}{L^{(d/2)}}\;\exp\left(\frac{2\pi \mathrm{i}}{L}\boldsymbol{k}\cdot\boldsymbol{n}\right) \tag{8.7}$$

normierte Eigenvektoren von K sind für jeden Vektor $\boldsymbol{k}$ $(1 \le k_i \le L)$. Die zugehörigen Eigenwerte lassen sich leicht bestimmen:

$$\lambda(\boldsymbol{k}) \;=\; M^2 + 2d \;-\; \sum_\mu \left(\exp\left(\frac{2\pi\mathrm{i}}{L}k_\mu\right) \;+\; \exp\left(-\frac{2\pi\mathrm{i}}{L}k_\mu\right)\right) \tag{8.8}$$

$$\;=\; M^2 + 4\sum_\mu \sin^2\left(\frac{\pi}{L}k_\mu\right) \tag{8.9}$$

Damit folgt für die 2-Punkt-Funktion

$$\langle\varphi(\boldsymbol{m})\varphi(\boldsymbol{n})\rangle \;=\; \frac{1}{L^d}\sum_k \frac{\exp(\frac{2\pi\mathrm{i}}{L}\boldsymbol{k}(\boldsymbol{n}-\boldsymbol{m}))}{M^2 + 4\sum_\mu \sin^2(\frac{\pi}{L}k_\mu)}\,. \tag{8.10}$$

8.3 Der Grenzwert $L \to \infty$

Für den thermodynamischen Limes bietet es sich an, eine auf das Intervall $(0, 2\pi]$ beschränkte Variable

$$p_\mu \;=\; \frac{2\pi}{L}k_\mu \qquad\qquad \Delta p \;=\; \frac{2\pi}{L}$$

einzuführen. Damit wird aus (8.10)

$$\langle\varphi(\boldsymbol{m})\varphi(\boldsymbol{n})\rangle \quad = \quad \frac{1}{(2\pi)^d}\sum_{\{p_\mu\}}(\Delta p)^d\;\frac{\exp(\mathrm{i}\boldsymbol{p}(\boldsymbol{n}-\boldsymbol{m}))}{M^2 + 4\sum_\mu \sin^2(\frac{p_\mu}{2})} \tag{8.11}$$

$$\stackrel{\Delta p\to 0}{\longrightarrow} \quad \frac{1}{(2\pi)^d}\int_0^{2\pi} d^d p\;\frac{\exp(\mathrm{i}\boldsymbol{p}(\boldsymbol{n}-\boldsymbol{m}))}{M^2 + 4\sum_\mu \sin^2(\frac{p_\mu}{2})}\,. \tag{8.12}$$

8.4 Darstellung der 2-Punkt Funktion als Summation über Wege

Wir werden im Folgenden zeigen, dass sich die 2-Punkt Funktion des freien Klein-Gordon Feldes darstellen lässt als „Summe über alle Wege“ von Punkt $\boldsymbol{m}$ nach $\boldsymbol{n}$. Diese Darstellung ist nicht nur für praktische Zwecke wichtig, beispielsweise für Abschätzungen oder Näherungen, sondern bildet auch eine Form der Feynmanschen „Summation über Geschichten“ für ein relativistisches freies Teilchen. Wir betrachten zunächst ein unendliches Gitter.

Es soll berechnet werden

$$G(\boldsymbol{n}) \;:=\; \mathrm{e}^{-\mu} \sum_{\text{Wege von 0 nach } n} \mathrm{e}^{-\mu l} \;=\; \mathrm{e}^{-\mu} \sum_{l=0}^{\infty} P_l(\boldsymbol{n}) \mathrm{e}^{-\mu l}\,, \tag{8.13}$$

dabei ist l die Länge des Weges (in Einheiten des Gitterabstands) und μ ein noch zu spezifizierender Parameter. $P_l(\boldsymbol{n})$ ist die Anzahl der Wege der Länge l von $\mathbf{0}$ zum Punkt $\boldsymbol{n}$ auf dem hyperkubischen Gitter. Der Normierungsfaktor $\mathrm{e}^{-\mu}$ ist gerade so gewählt, dass wir am Ende die richtige Greensche Funktion erhalten.

Für P_l gibt es eine „erzeugende Funktion“:

$$\left(\mathrm{e}^{\mathrm{i}p_1} + \mathrm{e}^{-\mathrm{i}p_1} + \mathrm{e}^{\mathrm{i}p_2} + \mathrm{e}^{-\mathrm{i}p_2} \;\ldots\; \mathrm{e}^{\mathrm{i}p_d} + \mathrm{e}^{-\mathrm{i}p_d}\right)^l \;=\; \sum_n P_l(\boldsymbol{n})\; \mathrm{e}^{\mathrm{i}(p_1 n_1 + \ldots + p_d n_d)}\,. \tag{8.14}$$

Diese Darstellung rührt daher, dass die binomischen Koeffizienten der linken Seite gerade der Kombinatorik der Wege in einem d-dimensionalen hyperkubischen Gitter entsprechen. Man kann sich das folgendermaßen veranschaulichen: Ausmultiplikation der linken Seite ergibt eine Summe über alle möglichen Kombinationen von jeweils l Faktoren $\mathrm{e}^{\pm \mathrm{i}p_\mu}$. Jeden dieser Terme kann man in eindeutige Korrespondenz setzen zu einem Weg der Länge l auf dem Gitter, indem man $\mathrm{e}^{\pm \mathrm{i}p_\mu}$ als „Schritt“ in $\pm\mu$-Richtung interpretiert. Die Summe über alle $\boldsymbol{n}$ eines (in diesem Fall unendlichen) Gitters konvergiert, da $P_l(\boldsymbol{n}) = 0$ für Punkte $\boldsymbol{n}$, die sich nicht durch einen Weg der Länge l mit $\mathbf{0}$ verbinden lassen.

Integration der Exponentialfunktion über $p_\mu (\in [0, 2\pi])$ ergibt 0, sofern nicht der Exponent verschwindet. Damit können wir ein spezielles $P_l(\boldsymbol{n})$ „herausprojizieren“:

$$\frac{1}{(2\pi)^d} \int_0^{2\pi} \mathrm{d}^d p\; \exp(-\mathrm{i}\boldsymbol{p}\boldsymbol{n}) \left(\mathrm{e}^{\mathrm{i}p_1} + \mathrm{e}^{-\mathrm{i}p_1} + \mathrm{e}^{\mathrm{i}p_2} + \mathrm{e}^{-\mathrm{i}p_2} \;\ldots\; \mathrm{e}^{\mathrm{i}p_d} + \mathrm{e}^{-\mathrm{i}p_d}\right)^l \;=\; P_l(\boldsymbol{n})\,. \tag{8.15}$$

Einsetzen dieser Gleichung in (8.13) reduziert die Summe über alle Längen l auf eine geometrische Reihe, und man erhält:

$$G(\boldsymbol{n}) \;=\; \frac{\mathrm{e}^{-\mu}}{(2\pi)^d} \int_0^{2\pi} \mathrm{d}^d p\; \left(\frac{\exp(-\mathrm{i}\boldsymbol{p}\boldsymbol{n})}{1 - 2\mathrm{e}^{-\mu}\left(\sum_{\mu=1}^d \cos p_\mu\right)}\right) \tag{8.16}$$

$$\;=\; \frac{1}{(2\pi)^d} \int_0^{2\pi} \mathrm{d}^d p\; \left(\frac{\exp(-\mathrm{i}\boldsymbol{p}\boldsymbol{n})}{\mathrm{e}^{\mu} - 2d + 4\sum_\mu \sin^2(\frac{p_\mu}{2})}\right)\,. \tag{8.17}$$

Wir erhalten also die Integraldarstellung der 2-Punkt-Funktion des freien Klein-Gordon-Feldes (8.12), vorausgesetzt wir definieren

$$\mathrm{e}^{\mu} - 2d \;=\; M^2\,. \tag{8.18}$$

Wir wollen nun zeigen, dass man auch die Greensche Funktion auf einem endlichen periodischen Gitter durch Summation über alle Wege erhält. Dazu beweisen wir zunächst folgendes allgemeines Theorem:

Sei $f(\boldsymbol{n})$ eine Funktion auf einem unendlichen Gitter mit der Fourier-Darstellung:

$$f(\boldsymbol{n}) \;=\; \frac{1}{(2\pi)^d} \int_{-\pi}^{\pi} \mathrm{d}p\; \mathrm{e}^{i\boldsymbol{p}\,\cdot\,\boldsymbol{n}}\; \tilde{f}(\boldsymbol{p})\,,$$

dann gilt:

$$F(\boldsymbol{n}) \;:=\; \sum_m f(\boldsymbol{n}+\boldsymbol{m}L) \;=\; \frac{1}{L^d}\sum_k \mathrm{e}^{\frac{2\pi\mathrm{i}}{L}\boldsymbol{k}\cdot\boldsymbol{n}}\,\tilde{f}(\tfrac{2\pi}{L}\boldsymbol{k})\,.$$

Der Beweis ergibt sich sofort aus der Poissonschen Summenformel für die Exponentialfunktion:

$$\sum_{n=-\infty}^{\infty} \mathrm{e}^{inx} \;=\; 2\pi \sum_{k=-\infty}^{\infty} \delta(x-2\pi k)\,,$$

die hier in der Form

$$\sum_m \mathrm{e}^{\mathrm{i}\boldsymbol{p}\cdot\boldsymbol{m}L} \;=\; \left(\frac{2\pi}{L}\right)^d \sum_k \delta\left(\boldsymbol{p}-\frac{2\pi}{L}\boldsymbol{k}\right)$$

auftritt.

Für die Summation über alle Wege auf einem endlichen Gitter von $\mathbf{0}$ zum Punkt $\boldsymbol{n}$ ist nun auch über alle Möglichkeiten zu summieren, das periodische Gitter zu umrunden. Das ist aber effektiv dasselbe, als ob auf einem unendlichen Gitter alle Punkte identifiziert werden, die sich um Vielfache der Gitterlänge unterscheiden. Die Greensche Funktion auf dem endlichen Gitter $G_L(\boldsymbol{n})$ ergibt sich also formal aus der Greenschen Funktion auf dem unendlichen Gitter durch

$$G_L(\boldsymbol{n}) \;=\; \sum_m G(\boldsymbol{n}+\boldsymbol{m}L)\,. \tag{8.19}$$

Mit obigem Theorem erhält man damit aus Gl. (8.12) sofort wieder die 2-Punkt-Funktion aus Gl. (8.10). Allerdings sollte man beachten, dass der Ausdruck in Gl. (8.19) nur formal ist, da die Summe im Allgemeinen nicht konvergiert und daher geeignete Terme subtrahiert werden müssen.

In Anhang A3 wird gezeigt, dass die (euklidische) Greensche Funktion eines freien Klein-Gordon-Feldes auf einem beliebigen Graphen eine Darstellung als Summation über alle Wege besitzt. Obwohl bei obiger Herleitung von der hyperkubischen Gitterstruktur wesentlich Gebrauch gemacht wurde, gilt das Ergebnis für beliebige Graphen. Die „Massenrenormierung“ ändert sich allerdings.

8.5 Der Kontinuumslimes der 2-Punkt-Funktion

In diesem Abschnitt werden wir explizit eine Gitterkonstante a einführen, und für die 2-Punkt-Funktion den Kontinuumslimes berechnen. Ausgangspunkt ist die Integraldarstellung der Korrelationsfunktion für ein unendlich großes Gitter

$$G(\boldsymbol{n}) \;=\; \frac{1}{(2\pi)^d}\int_{-\pi}^{\pi} \mathrm{d}^d p\;\frac{\exp\mathrm{i}(\frac{p}{a})(a\boldsymbol{n})}{M^2+4\sum_\mu \sin^2\frac{p_\mu}{2}}\,, \tag{8.20}$$

wobei der trivial eingefügte Parameter a die Rolle des Gitterabstandes übernehmen wird. Ausgedrückt durch die Variablen

$$\boldsymbol{x} \;=\; a\boldsymbol{n} \qquad \text{und} \qquad \boldsymbol{k} \;=\; \frac{\boldsymbol{p}}{a}$$

erhalten wir:

$$G(\boldsymbol{x}) \;=\; \Big(\frac{a}{2\pi}\Big)^d \int_{-\frac{\pi}{a}}^{\frac{\pi}{a}} \mathrm{d}^d k \frac{\exp \mathrm{i}\boldsymbol{k}\boldsymbol{x}}{M^2 + 4\sum_\mu \sin^2 \frac{k_\mu a}{2}}\,.$$

Diese Funktion erfüllt die Gleichung:

$$\left(\sum_\mu\; G(\boldsymbol{x} + a\boldsymbol{e}_\mu) - 2G(\boldsymbol{x}) + G(\boldsymbol{x} - a\boldsymbol{e}_\mu)\right) \;+\; M^2 G(\boldsymbol{x}) \;=\; \delta_{x,0}\,.$$

Wir definieren nun

$$\tilde{G}_a(\boldsymbol{x}) \;=\; \frac{1}{a^{d-2}} G(\boldsymbol{x})\,. \tag{8.21}$$

Diese Funktion erfüllt

$$\left(\sum_\mu\; \frac{\tilde{G}_a(\boldsymbol{x} + a\boldsymbol{e}_\mu) - 2\tilde{G}_a(\boldsymbol{x}) + \tilde{G}_a(\boldsymbol{x} - a\boldsymbol{e}_\mu)}{a^2}\right) \;+\; \frac{M^2}{a^2}\,\tilde{G}_a(\boldsymbol{x}) \;=\; \frac{1}{a^d}\delta_{x,0}\,.$$

Für $a \to 0$ wird $\tilde{G}_a$ also Greensche Funktion zur Klein-Gordon-Gleichung (da es sich um die euklidische Formulierung handelt, spräche man besser von der Helmholtz-Gleichung), wobei $m = M/a$ ein Massenparameter ist und

$$\delta(\boldsymbol{x}) \;=\; \lim_{a\to 0}\; \frac{1}{a^d}\;\delta_{x,0}$$

die δ-Distribution wird. Dass der Grenzwert $a \to 0$ in der Integraldarstellung

$$\tilde{G}(\boldsymbol{x}) \;=\; \lim_{a\to 0}\;\; \frac{1}{(2\pi)^d} \int_{-\frac{\pi}{a}}^{\frac{\pi}{a}} \mathrm{d}^d k\;\; \frac{\exp \mathrm{i}\boldsymbol{k}\boldsymbol{x}}{m^2 + \frac{4}{a^2}\sum_\mu \sin^2 \frac{k_\mu a}{2}} \tag{8.22}$$

existiert und zu der bekannten Fourier-Darstellung der Greenschen Funktion zum (euklidischen) Klein-Gordon-Operator wird,

$$\tilde{G}(\boldsymbol{x}) \;=\; \frac{1}{(2\pi)^d} \int_{-\infty}^{\infty} \mathrm{d}^d k\;\; \frac{\exp \mathrm{i}\boldsymbol{k}\boldsymbol{x}}{m^2 + \sum_\mu (k_\mu)^2}\,, \tag{8.23}$$

ist allerdings nicht so offensichtlich. Da die Integrationsgrenzen für k in diesem Grenzwert immer größer werden, kann man die Sinus-Funktion nicht einfach entwickeln. Es ist zu zeigen, dass für $x \neq \mathbf{0}$ die großen k-Werte zum Integral wegen der raschen Oszillation der Exponentialfunktion nicht beitragen.

Die Problematik der obigen Argumentation wird offensichtlich, wenn man in (8.20) die ursprünglichen Integrationsgrenzen $[0, 2\pi]$ beibehält. In (8.22) sind die Integrationsgrenzen nun $[0, 2\pi/a]$, und im Grenzfall $a \to \infty$ werden diese Grenzen zu $[0, \infty]$. Das ist nur die Hälfte des Beitrags von (8.23). Der Fehler liegt darin, dass bei dieser Wahl der Integrationsgrenzen nicht nur der Bereich um $k = 0$ einen Beitrag zum Kontinuumsintegral liefert, sondern auch der Bereich um $k = 2\pi/a$. Diese beiden Beiträge zusammengenommen entsprechen dem Integral (8.23).

Abschließend noch einige Bemerkungen zu diesem Ergebnis:

1. Die Reskalierung (8.21) der Green'schen Funktion entspricht eigentlich einer Feldrenormierung. In der euklidischen Wirkung (8.1) sind sowohl der Parameter M als auch das Feld φ zunächst dimensionslos. In einer Kontinuumsversion wird die Summe über alle Gitterpunkte zu einem Integral über die Raum-Zeit und das Quadrat der Differenz zum Quadrat der ersten Ableitung. Damit hat das Feld φ im Kontinuum die Dimension $[\varphi] = [\text{Länge}]^{-(d-2)/2}$ (es handelt sich also strenggenommen um eine verallgemeinerte Dichte). Ein Kontinuumslimes in der Wirkung

$$\begin{aligned} S[\varphi] &= \frac{1}{2}\sum_{\{m_i\}} \frac{a^d}{a^{d-2}} \left(\sum_{\mu=1}^{d} \left(\frac{\varphi(\boldsymbol{x}+a\boldsymbol{e}_\mu)-\varphi(\boldsymbol{x})}{a} \right)^2 + \frac{M^2}{a^2}\varphi(\boldsymbol{x})^2 \right) \\ &\to \frac{1}{2}\int \mathrm{d}^d x \left(\sum_{\mu=1}^{d} \left(\frac{\partial\tilde{\varphi}}{\partial x^\mu} \right)^2 + m^2\tilde{\varphi}(\boldsymbol{x})^2 \right) \end{aligned}$$

verlangt die Renormierung der Felder

$$\varphi(\boldsymbol{m}) \longrightarrow \tilde{\varphi}(\boldsymbol{x})|_{x=am} = \frac{1}{a^{(d-2)/2}}\,\varphi(a\boldsymbol{m})\,.$$

Das überträgt sich entsprechend auf eine Renormierung der n-Punkt-Funktion

$$G(\boldsymbol{m}_1,\ldots,\boldsymbol{m}_n) \longrightarrow \tilde{G}(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n) = a^{-n(d-2)/2} G(a\boldsymbol{m}_1,\ldots,a\boldsymbol{m}_n)\,.$$

2. Für große Werte des Arguments fällt die 2-Punkt Funktion der Gittertheorie – ebenso wie die Greensche Funktion der Kontinuumstheorie – exponentiell ab:

$$\tilde{G}(\boldsymbol{x}) \overset{m|x|\gg 1}{\longrightarrow} \exp(-m|\boldsymbol{x}|) \qquad\qquad G(\boldsymbol{n}) \overset{|n|M\gg 1}{\longrightarrow} \exp(-M|\boldsymbol{n}|)\,.$$

3. Für große Werte des Arguments $|\boldsymbol{n}| \gg 1/M$ wird $G(\boldsymbol{n})$ näherungsweise "rotationssymmetrisch", d.h. $SO(d)$ invariant. Diese Eigenschaft ist wesentlich für eine rotationsinvariante (im Minkowski-Raum Lorentz-invariante) Korrelationsfunktion im Kontinuumslimes. Obwohl die Gitterregularisierung die Rotationsinvarianz bricht, wird diese im Kontinuumslimes wieder restauriert.

8.6 Der Kontinuumslimes des freien Klein-Gordon-Feldes

Wir werden später noch eine allgemeine Theorie des Kontinuumslimes behandeln. Das freie Klein-Gordon Feld ist in mehrfacher Hinsicht zu speziell. Trotzdem lassen sich bereits hier einige wesentliche Aspekte des Kontinuumslimes behandeln.

Implizit traten bei der Herleitung des Kontinuumslimes vier Größen auf:

1. Der Parameter M:
   Dieser Parameter steht in der Wirkung (8.1). Es handelt sich um einen so genannten „nackten" ('bare') Parameter, in diesem Fall um die nackte (Gitter)-Masse. M ist dimensionslos.

2. Die inverse Korrelationslänge $\bar{M}$:
   Sie ist definiert durch
   $$\bar{M} \;=\; \lim_{|n|\to\infty} -\frac{\ln G(\boldsymbol{n})}{|\boldsymbol{n}|}\,.$$
   Diese Größe ist ebenfalls dimensionslos (entspricht aber einer Distanz in Gittereinheiten) und ergibt sich aus der 2-Punkt-Funktion. Im Allgemeinen ist $\bar{M}$ eine Funktion von M und eventuell weiteren Kopplungskonstanten. Im vorliegenden Fall gilt $\bar{M}(M) = M$ (in dieser Hinsicht ist die freie Theorie zu speziell).

3. Die physikalische Masse m:
   m ist die physikalische Masse des Teilchens, das durch das Feld φ (zu dem $\bar{M}$ die inverse Korrelationslänge ist) beschrieben werden soll. m hat die Dimension einer Masse (inversen Länge) und wird von der „Realität“ (d.h. dem zu beschreibenden physikalischen Teilchen) vorgegeben (beispielsweise $m = 10^{13}\,\mathrm{cm}^{-1}$).

4. Der Gitterabstand a:
   Der Gitterabstand a ist definiert durch
   $$a \;=\; \frac{\bar{M}(M)}{m}\,. \tag{8.24}$$
   Durch die Wahl der nackten Masse M ergibt sich die inverse Korrelationslänge $\bar{M}(M)$ in Gittereinheiten. Diese soll die physikalische Masse m beschreiben. Durch diese Beziehung zwischen der inversen Korrelationslänge und der physikalischen Masse wird die Gitterkonstante festgelegt. Strenggenommen ist a also eine Funktion der physikalischen Masse m und des vorgegebenen Parameters M.

   Wenn wir a kleiner wählen, müssen wir auch für M einen entsprechend kleinere Wert wählen, sodass die physikalische Masse $m = M/a$ konstant bleibt. Der „Kontinuumslimes“ bedeutet also, dass M gegen 0 geht. An diesem Punkt wird die Korrelationslänge langreichweitig, d.h., für $M = 0$ beschreibt die Gittertheorie eine „kritische“ Theorie, wobei wir hier eine Theorie als „kritisch“ bezeichnen, wenn die Korrelationslänge unendlich ist. Der Kontinuumslimes bedeutet also, dass man sich einer kritischen Theorie nähert.

   Wir könnnen (8.24) auch unter einem anderen Gesichtspunkt interpretieren: Dem Abstand zwischen zwei benachbaten Gitterpunkten wird zunächst willkürlich ein physikalischer Gitterabstand a zugeordnet. Außerdem soll das Feld φ das physikalische Teilchen mit der Masse m beschreiben. Das dimensionslose Produkt, am, entspricht der Compton-Länge des Teilchens in Einheiten der willkürlich gewählten Gitterkonstanten. Diese Zahl soll identisch zur dimensionslosen inversen Korrelationslänge $\bar{M}(M)$ des Feldes φ auf dem Gitter sein[1]. Dadurch ist bei Vorgabe der Gitterkonstanten a und der physikalischen Masse m der nackte Parameter M festgelegt. Eine Änderung der (unbeobachtbaren) Gitterkonstanten a kann durch eine Änderung des (ebenfalls unbeobachtbaren) Parameters M kompensiert werden, sodass die physikalischen Größen konstant bleiben. Dies wird später der Ausgangspunkt der Renormierungsgruppengleichungen.

Der Kontinuumslimes besteht nun im wesentlichen aus zwei Schritten:

[1] Die Wellenfunktion des Teilchens im Kontinuum und das Feld φ auf dem Gitter unterscheiden sich im allgemeinen noch durch eine Normierungskonstante. Diese hat aber keinen Einfluss auf die Korrelationslänge.

1. Die freien Parameter der Theorie (hier M) müssen so adjustiert werden, dass die physikalischen Längen (z.B. die Korrelationslängen von 2-Punkt-Funktionen) im Vergleich zum Gitterabstand sehr groß werden.

2. Für jeden freien Parameter muss eine physikalische Größe des statistischen Modells (z.B. die Korrelationslänge der 2-Punkt-Funktion) an eine entsprechende Größe der „realen Welt“ (z.B. die Compton-Länge eines Teilchens) angepasst werden. Durch diesen Schritt wird dem Gitterabstand eine physikalische Länge zugeschrieben.

Der erste Schritt bedeutet für die freie Klein-Gordon Theorie, dass der Parameter M sehr klein gewählt werden muss. Die 2-Punkt-Funktion wird im Vergleich zum Gitterabstand also sehr langreichweitig. Im zweiten Schritt identifiziert man die Korrelationslänge der 2-Punkt-Funktion mit der physikalischen Länge (Compton-Länge) eines zu beschreibenden Systems. Dadurch wird bei Vorgabe der physikalischen Länge der nackte Parameter M mit dem Gitterabstand a in Beziehung gesetzt (Gl. (8.24). Will man z.B. Pionen beschreiben mit einer Compton-Länge von ungefähr 10^{-13} cm, und hat man dem Gitter willkürlich die Gitterkonstante 10^{-15} cm zugeordnet, so ist die Compton-Länge in Einheiten der Gitterkonstanten 100. Die Korrelationslänge sollte also ebenfalls 100 betragen, d.h., der nackte Parameter ist $M = 0.01$. In der Praxis, insbesondere bei MC-Simulationen, gibt man meist den Parameter M vor und bestimmt anschließend nach (8.24) die zugehörige Gitterkonstante.

Diese Darstellung klingt für das freie Klein-Gordon Feld teilweise übertrieben spitzfindig. Das rührt im wesentlichen daher, dass der einzige freie Parameter M gleichzeitig auch die Korrelationslänge $\bar{M}$ der 2-Punkt-Funktion ist, eine eigentliche „Renormierung“ also nicht notwendig ist. Außerdem gibt es kein nicht-triviales Massenspektrum, das die Theorie beschreiben kann. Alles, was die Theorie liefert, wird vorher hineindefiniert. Die folgende Überlegung soll jedoch eine Vorstellung davon geben, wie in einer wechselwirkenden Theorie durchaus interessante Vorhersagen gemacht werden können.

Wir betrachten eine spezielle 4-Punkt-Funktion des freien Klein-Gordon Feldes:

$$\langle \varphi(x)^2 \varphi(y)^2 \rangle \;=\; \langle \varphi(x)^2 \rangle \langle \varphi(y)^2 \rangle + 2 \langle \varphi(x)\varphi(y) \rangle^2 \quad \longrightarrow \quad c_0 + c_1 \mathrm{e}^{-2M|x-y|} \,. \tag{8.25}$$

Diese Korrelationsfunktion verschwindet nicht für große Distanzen $|x - y| \to \infty$. Der Zustand, den man durch zweifache Anwendung des Feldoperators $\phi(x)$ erhält, hat daher eine Komponente im Vakuum. Dies ist nicht erstaunlich, denn $\phi(x)$ ist eine Summe aus Teilchen-Erzeugungs- und -Vernichtungsoperatoren (man denke an den harmonischen Oszillator).

Außerdem wird aus dem asymptotischen Verhalten dieser Korrelationsfunktion deutlich, dass der kleinste Energieeigenwert (außer 0) im „2-Teilchen-Sektor“ durch $2M$ gegeben ist. Auch dies ist nicht überraschend, denn die kleinste Energie von zwei freien Teilchen der Masse M ist durch die Summe ihrer Ruheenergien gegeben. Man kann sich aber leicht vorstellen, dass in einer Theorie mit Wechselwirkungen zwei Teilchen einen gebundenen Zustand bilden können. In diesem Fall ist das asymptotische Verhalten der obigen Korrelationsfunktion durch $E = 2M - \epsilon$ charakterisiert, wobei ϵ die Bindungsenergie des Grundzustands dieses Zwei-Teilchensystems angibt.

Kapitel 9

Triviality of $(\phi^4)_4$

Let us denote by $(\phi^p)_d$ the scalar field theory in d dimensions (in the Minkowski formulation in $d-1$ spatial and one temporal dimension) with an interaction term of the form $\int \mathrm{d}^d x\, \phi(x)^p$. In the lattice regularized version (where we denote by Λ the lattice points), the measure is given by

$$\mathrm{d}\mu[\{\phi\}] = \prod_{i\in\Lambda} \mathrm{d}\phi_i \;\; \exp\left(J \sum_{\langle i,j\rangle} \phi_i\phi_j - M^2 \sum_{i\in\Lambda} \phi_i^2 - g \sum_{i\in\Lambda} \phi_i^p \right) , \tag{9.1}$$

where $\langle i,j\rangle$ denotes nearest neighbors on the lattice. Expectation values are

$$\tilde{S}(n_1,n_2,...,n_k) = \langle \phi_{n_1}\phi_{n_2}...\phi_{n_k}\rangle = \int \phi_{n_1}...\phi_{n_k} \;\mathrm{d}\mu[\{\phi\}]\,, \tag{9.2}$$

where n_i are lattice points. For a lattice spacing a define $x_i = an_i$ and consider the limit $a\to 0$ and $n_i\to\infty$ such that x_i fixed:

$$S(x_1,...,x_k) = \lim_{\substack{a\,\to\,0,\, n_i\,\to\,\infty \\ x_i\,=\,an_i}} Z^k \tilde{S}(n_1,n_2,...,n_k) \tag{9.3}$$

The continuum limit is achieved by fine-tuning the constants J, M and g as well as Z in the limit $a\to 0$ such that S is non-trivial, in particular the 2-point function $\tilde{S}(x_1-x_2) = S(x_1,x_2)$ should be locally integrable. A theory is called trivial if the connected part of the 4-point function

$$\begin{aligned} S^c(x_1,x_2,x_3,x_4) \;\; &= \;\; S(x_1,x_2,x_3,x_4) - \\ &-S(x_1,x_2)S(x_3,x_4) - S(x_1,x_3)S(x_2,x_4) - S(x_1,x_4)S(x_2,x_3) \end{aligned}$$

is zero in the continuum limit, in which case the measure corresponds essentially to a Gaussian measure and the expectation values of the fields factorize into 2-point functions.

On the other hand, if the constants J, g, M, Z can be tuned in such a way that S_2 remains locally integrable but "$S_4 - 3S_2$" is non-zero, then the theory is non-trivial and represents interacting fields.

It has been shown (see [2] and the methods used in [31]) that for $d > 4$ the theory $(\phi^4)_d$ is trivial in the sense defined above, i.e., it is a theory of free particles. This is particular interesting, as exactly these cases are perturbatively non-renormalizable, and it has often been said that these field theories do not exist. For $d = 2, 3$ is can be shown that the theories are non-trivial. There are strong believes (though no rigorous proof yet) that for $d = 4$ the theory is also trivial. This is particularly interesting for the standard model of particle theory, because the Higgs field is described by a ϕ^4 theory. If it is true that in the continuum limit this theory becomes trivial, this would be an indication that the Higgs field theory only exists with a cut-off (and the cut-off cannot be taken to infinity).

In the following we briefly sketch the proof of triviality of $(\phi^4)_d$ for $d > 4$. In section 8.16 we have shown that the 2-point function $S(x_1, x_2)$ can be written as a summation over all paths from x_1 to x_2 where each path it weighted by a factor whose logarithm is proportional to the length of the path (In the appendix A3 this result is generalized to arbitrary regular graphs). Symanzik has shown that the $2p$-point function $S(x_1, ..., x_{2p})$ can be expressed as a summation over p random walks that begin and end at disjoint pairs of lattice sites in $x_1, ..., x_{2p}$. Whenever two of these random walks meet at the same lattice site there is an additional factor g to the weight of this configuration. If the random walks meet at several sites there are the same combinatorical factors entering as in ordinary perturbation expansion. Indeed, as the 2-point function (the free propagator) can be expressed over sums of random walks, it is easy to imagine, that every Feynman graph which occurs in perturbation theory has a representation over random walks with additional weights entering whenever two walks meet.

What can be shown is that for $d > 4$ the probability that two random walks meet at the same site is practically zero (it becomes zero in the continuum limit). This effect cannot be balanced by a renormalization of g. Therefore, in the continuum limit, the connected part of the 4-point function in $d > 4$ is zero and the theory is essentially free.

For $d = 4$ there are logarithmic corrections in the estimates of the probability which are believed to become irrelevant in the continuum limit. A rigorous proof of that conjecture, however, seems still to be missing.

Kapitel 10

Reine Eichtheorien auf dem Gitter

10.1 „Motivierende“ Vorbemerkungen

In der Feldtheorie (vgl. Kap. 2.6, Gl.2.31) wurde die kovariante Ableitung $D_\mu = \partial_\mu + igA_\mu$ eingeführt, um Ausdrücke der Art

$$(\partial_\mu \varphi(x))^+ (\partial^\mu \varphi) \qquad \text{bzw.} \qquad \bar{\psi}(x)\partial_\mu \psi(x)$$

durch Ausdrücke zu ersetzen, die invariant unter lokalen Symmetrietransformationen zu einer kompakten Gruppe G sind. Das Eichpotential A_μ hat die Interpretation einer Zusammenhangsform. Es ist der Generator von parallelen Verschiebungen der Felder. Der aufintegrierte Paralleltransporter (Gl. (2.45), vgl. auch Anhang A1), der parallele Verschiebungen über endliche Distanzen beschreibt, ist ein Element der Symmetriegruppe.

Verlangt man lokale Invarianz unter Symmetrietransformationen

$$\varphi(\boldsymbol{m}) \;\rightarrow\; g(\boldsymbol{m})\varphi(\boldsymbol{m}) \qquad\qquad g(\boldsymbol{m}) \in G$$

auf dem Gitter, so sind es die Terme der Art

$$\varphi(\boldsymbol{m} + \boldsymbol{e}_\mu)^+ \; \varphi(\boldsymbol{m})\,, \tag{10.1}$$

die diese Invarianz nicht respektieren. Die Tatsache, dass man in diesen Termen das Skalarprodukt des Feldes an der Stelle $\boldsymbol{m}$ mit dem Feld an der Stelle $\boldsymbol{m} + \boldsymbol{e}_\mu$ bildet, legt nahe, das Feld zunächst „parallel zu verschieben“, bevor das Skalarprodukt gebildet wird, d.h., wir ersetzen den Ausdruck (10.1) durch

$$\varphi(\boldsymbol{m} + \boldsymbol{e}_\mu^+) \; U(\boldsymbol{m} + \boldsymbol{e}_\mu, \boldsymbol{m}) \; \varphi(\boldsymbol{m})\,. \tag{10.2}$$

Versteht man unter $U(\boldsymbol{m}+\boldsymbol{e}_\mu, \boldsymbol{m})$ den Paralleltransporter vom Punkte $\boldsymbol{m}$ zum Punkte $\boldsymbol{m}+\boldsymbol{e}_\mu$, so ist dieser Ausdruck invariant unter lokalen Invarianztransformationen.

Im Allgemeinen hängt der Paralleltransporter nicht nur vom Anfangs- und Endpunkt ab, sondern auch vom Weg der Verschiebung. Auf dem Gitter kann man sich unter $U(\boldsymbol{m} + \boldsymbol{e}_\mu, \boldsymbol{m})$

den Paralleltransport entlang der Gitterlinie vorstellen, obwohl eine solche Interpretation nicht von Bedeutung ist.

Diese Überlegungen legen nahe, als dynamische Variable einer Eichtheorie auf dem Gitter nicht die Eichfelder $A_\mu(x)$ zu wählen, sondern die „elementaren" Paralleltransporter zwischen benachbarten Punkten. Diese Elemente der Invarianzgruppe transformieren sich durch Konjugation am Anfangs- und Endpunkt:

$$U(\boldsymbol{m}+\boldsymbol{e}_\mu,\boldsymbol{m}) \quad \rightarrow \quad g(\boldsymbol{m}+\boldsymbol{e}_\mu)\; U(\boldsymbol{m}+\boldsymbol{e}_\mu,\boldsymbol{m})\; g(\boldsymbol{m})^{-1}\,. \tag{10.3}$$

10.2 Die Wirkung der reinen Eichtheorie auf dem Gitter

Die Vorbemerkungen gingen von einem Feld φ aus und versuchten heuristisch zu begründen, warum der Paralleltransporter entlang Gitterlinien eine naheliegende „Diskretisierung" der Eichpotentiale darstellt. Im Folgenden werden wir allerdings zunächst nicht die Kopplung des Eichsektors mit anderen Materiefeldern untersuchen, sondern die Formulierung des reinen Eichsektors begründen.

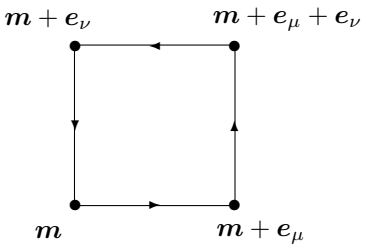
Paare von benachbarten Punkten $(\boldsymbol{m},\boldsymbol{m}\pm\boldsymbol{e}_\mu)$ werden im Folgenden immer kurz mit $(\boldsymbol{m},\pm\mu)$ bezeichnet. Die dynamischen Variable sind Elemente einer kompakten Gruppe G. Diese „leben" auf den Gitterlinien, d.h. Paaren benachbarter Punkte. Genauer definieren wir eine Konfiguration U als Abbildung von der Menge der (gerichteten) Linien E in die Gruppe G:

$$\begin{aligned} & & U\;:\;E &= \{(\boldsymbol{m},\mu)\} \;\rightarrow\; G \\ \text{mit} & & U(\boldsymbol{m},\mu) &= U(\boldsymbol{m}+\boldsymbol{e}_\mu,-\mu)^{-1}\,. \end{aligned}$$

Die letzte Bedingung bedeutet, dass der Paralleltransporter von x nach y das Inverse des Paralleltransporters von y nach x (entlang desselben Weges) ist.

Im Kontinuum ist die Wirkung durch das Quadrat der „Krümmung" $F_{\mu\nu}$ gegeben. Die Krümmung gibt an, um wieviel ein Vektorfeld sich (pro „Flächeninhalt") verändert, wenn man es infinitesimal um einen geschlossenen Weg parallel transportiert (vgl. Gl. (A1.2)). Dabei bezeichnet $\mu\nu$ die Ebene, in der diese Verschiebung stattgefunden hat. Auf dem Gitter gibt es keine „infinitesimale" Verschiebung. Die elementarste Verschiebung entlang eines geschlossenen Weges ergibt sich aus dem Produkt der Paralleltransporter um eine Plakette. Wir definieren:

$$\begin{aligned} U_p(m,\mu\nu) \;&:=\; U(m+e_\nu,-\nu)\;U(m+e_\mu+e_\nu,-\mu)\;U(m+e_\mu,\nu)\;U(m,\mu) \\ &=\; U(m,\nu)^{-1}\;U(m+e_\nu,\mu)^{-1}\;U(m+e_\mu,\nu)\;U(m,\mu)\,. \end{aligned}$$



$U_p(\boldsymbol{m},\mu\nu)$ bezeichnet also den Paralleltransporter um eine Plakette, definiert durch die vier Punkte $\boldsymbol{m}$, $\boldsymbol{m}+\boldsymbol{e}_\mu$, $\boldsymbol{m}+\boldsymbol{e}_\mu+\boldsymbol{e}_\nu$, $\boldsymbol{m}+\boldsymbol{e}_\nu$.

Eine Wirkung ist eine reelle Größe. Die Spur von U_p ist zwar eine Zahl, jedoch kann diese bei komplexen Gruppen auch komplex werden. Es bietet sich tr $(U_p+U_p^+)$ an, also der Realteil

dieser Zahl. Zur Motivation, dass diese Größe tatsächlich etwas mit der Wirkung zu tun hat, die wir im Kontinuum gefordert hatten, führen wir formal einen Gitterabstand a ein und schreiben:

$$U(\boldsymbol{m},\mu) \;=\; \exp(\mathrm{i}gA_\mu(\boldsymbol{m})a)\,. \tag{10.4}$$

$A_\mu(\boldsymbol{m}) = A^a_\mu(\boldsymbol{m})T_a$ ist hierbei ein Element der Lie-Algebra der Gruppe G und T_a ist ein geeigneter Satz von Generatoren. Bilden wir das Produkt der Gruppenelemente um eine Elementarplakette, so erhalten wir:

$$\begin{aligned}
U_p(\mu\nu) &= \mathrm{e}^{-\mathrm{i}gaA_\nu(m)}\,\mathrm{e}^{-\mathrm{i}gaA_\mu(m+ae_\nu)}\,\mathrm{e}^{\mathrm{i}gaA_\nu(m+ae_\mu)}\,\mathrm{e}^{\mathrm{i}gaA_\mu(m)} \\
&= \mathrm{e}^{-\mathrm{i}gaA_\nu(m)}\,\mathrm{e}^{-\mathrm{i}gaA_\mu(m)-\mathrm{i}ga^2\partial_\nu A_\mu(m)+\mathcal{O}(a^3)} \\
&\qquad\qquad \mathrm{e}^{\mathrm{i}gaA_\nu(m)+\mathrm{i}ga^2\partial_\mu A_\nu(m)+\mathcal{O}(a^3)}\,\mathrm{e}^{\mathrm{i}gaA_\mu(m)} \qquad (10.5)\\
&= \mathrm{e}^{\mathrm{i}a^2g\partial_\mu A_\nu(m)-\mathrm{i}a^2g\partial_\nu A_\mu(m)+g^2a^2[A_\mu(m),A_\nu(m)]+\mathcal{O}(a^3)} \qquad (10.6)\\
&= \mathrm{e}^{\mathrm{i}ga^2(\partial_\mu A_\nu(m)-\partial_\nu A_\mu(m)-\mathrm{i}g[A_\mu(m),A_\nu(m)])+\mathcal{O}(a^3)} \\
&= \mathrm{e}^{\mathrm{i}ga^2F_{\mu\nu}+\mathcal{O}(a^3)}\,. \qquad (10.7)
\end{aligned}$$

Dabei haben wir bei dem Schritt von (10.5) zu (10.6) die Baker-Hausdorff-Formel

$$\mathrm{e}^A\,\mathrm{e}^B \;=\; \mathrm{e}^{A+B+\frac{1}{2}[A,B]+\ldots}$$

angewandt und nur Terme bis zur Ordnung a^2 berücksichtigt. Insbesondere folgt:

$$U_p + U_p^+ \;\approx\; 2 - g^2a^4F_{\mu\nu}^2 + \mathcal{O}(a^5)\,. \tag{10.8}$$

Da in diesem Ausdruck der führende Term nach der Konstanten von der Ordnung a^4 ist könnte man meinen, dass wir von vorne herein die Entwicklung bis zu dieser Ordnung hätten ausführen müssen. Wir wissen aber, dass U_p eine unitäre Matrix ist, d.h. $U_p = \exp(\mathrm{i}T)$, wobei T eine hermitesche Matrix ist mit $T = ga^2F_{\mu\nu}+\mathcal{O}(a^3)$. Für eine unitäre Matrix gilt aber immer $U_p + U_p^+ = 2 - T^2 + \ldots$. Zur Ordnung a^4 in $U_p + U_p^+$ trägt also tatsächlich nur der führende Term der Ordnung a^2 in T bei.

Letztendlich haben wir in (10.7) nur nochmals gezeigt, dass die Krümmung $F_{\mu\nu}$ in einem gewissen Sinne als Generator von geschlossenen Paralleltransporten aufgefasst werden kann. Dabei sollte man jedoch beachten, dass obige Rechnung einem „naiven“ Kontinuumslimes entspricht. Das bedeutet folgendes: Wir nehmen eine stetige (und hinreichend oft differenzierbare) Kontinuumskonfiguration $A_\mu(\boldsymbol{x})$ an und überdecken den Raum mit einem Gitter der Gitterkonstanten a. Diese Kontinuumskonfiguration soll über Abstände der Ordnung der Gitterkonstanten a nur wenig variieren. Bezüglich dieser Kontinuumskonfiguration können wir die Paralleltransporter $U(\boldsymbol{m},\mu)$ entlang der Gitterlinien bestimmen. Für eine Konfiguration A_μ mit obigen Eigenschaften sind die Paralleltransporter nahe der Identität und daher in der Form (10.4) approximierbar. Im Prinzip treten allerdings Korrekturterme der Ordnung a^2 auf, die aber zu (10.7) nicht beitragen, wenn die Kontinuumskonfiguration die geforderten Eigenschaften hat.

Das Ergebnis dieses naiven Kontinuumslimes legt also folgende Definition für die Wirkung nahe:

$$S_{\text{gauge}} \;=\; -\frac{1}{g^2N}\sum_p \;\mathrm{tr}\,\left(\mathbb{1} - \frac{1}{2}(U_p + U_p^+)\right)\,. \tag{10.9}$$

Die Summe verläuft dabei über alle Plaketten im Gitter, d.h., wir erhalten eine Summation über alle Gitterpunkte $\boldsymbol{m}$ (im Kontinuumslimes wird diese Summation zur Raum-Zeit-Integration)

sowie über alle (linear unabhängigen) Richtungen μ, ν. Der Faktor $1/N$ ist ein Normierungsfaktor der Spur. Es soll damit ausgedrückt werden, dass U Element einer $N \times N$-dimensionalen Darstellung der Gruppe ist, und tr $\mathbb{1} = N$. In der obigen Summe erhält jede Plakette einen Beitrag $\frac{1}{2} F_{\mu\nu} F^{\mu\nu}$. Im Kontinuum gab es einen Faktor $\frac{1}{4}$, dort wurde aber in der Summation über alle $\mu\nu$ jede Plakette doppelt gezählt.

Während U_p sich unter Eichtransformationen (10.3) noch kovariant transformiert,

$$U_p(\boldsymbol{m}, \mu\nu) \;\longrightarrow\; g(\boldsymbol{m})\, U_p(\boldsymbol{m}, \mu\nu)\, g^{-1}(\boldsymbol{m})\,,$$

ist die Spur dieser Matrix und damit auch die Wirkung (10.9) unter Eichtransformationen invariant.

10.3 Das Integrationsmaß der Eichtheorie

Es muss noch geklärt werden, wodurch auf dem Gitter die formale Integration über alle Eichfelder zu ersetzen ist. Es ist naheliegend, die invariante Integration über die Gruppe zu wählen:

$$\int \mathcal{D} A_\mu(x) \quad \to \quad \int \prod_{(m,\mu)} \mathrm{d}\mu[U(\boldsymbol{m}, \mu)]\,. \tag{10.10}$$

Dabei ist

$$\mathrm{d}\mu[U] \;=\; \mathrm{d}\mu[gU] \;=\; \mathrm{d}\mu[Ug] \tag{10.11}$$

das (bei kompakten Gruppen bis auf eine multiplikative Konstante eindeutige) rechts- und linksinvariante Haar-Maß auf der Gruppe. Wenn eine Verwechslung ausgeschlossen ist, werden wir meist kurz $\mathrm{d}U$ anstatt $\mathrm{d}\mu[U]$ schreiben.

Jede lokalkompakte Gruppe besitzt ein (bis auf eine multiplikative Konstante eindeutiges) rechtsinvariantes Maß (d.h. $\mathrm{d}\mu_R[Ug] = \mathrm{d}\mu_R[U]$), sowie ein linksinvariantes Maß($\mathrm{d}\mu_L[gU] = \mathrm{d}\mu_L[U]$). Im Allgemeinen sind diese beiden Maße verschieden. Gruppen, für die das rechtsinvariante und linksinvariante Maß gleich sind, bezeichnet man als „unimodular“.

Es lässt sich zeigen, dass jede kompakte Gruppe unimodular ist. Es gibt aber auch nicht-kompakte Gruppen, die ebenfalls unimodular sind, beispielsweise die Gruppe $\mathbb{R}$ unter der Operation $(x, y) \to x + y$. Das invariante Haar-Maß ist in diesem Fall einfach das übliche Maß $\mathrm{d}x$.

Ein Beispiel für eine nicht-unimodulare Gruppe bildet die Menge der Matrizen

$$g(x, y) \;=\; \begin{pmatrix} x & y \\ 0 & 1 \end{pmatrix} \qquad\qquad x, y \in \mathbb{R}$$

mit der üblichen Matrixmultiplikation. Es lässt sich leicht zeigen, dass ein linksinvariantes Maß durch

$$\mathrm{d}\mu_L(x, y) \;=\; \frac{1}{x^2} \mathrm{d}x\, \mathrm{d}y$$

gegeben ist, wohingegen ein rechtsinvariantes Maß proportional zu

$$\mathrm{d}\mu_R(x, y) \;=\; \frac{1}{x} \mathrm{d}x\, \mathrm{d}y$$

ist.

Ist G diskret (auf dem Gitter kann man natürlich auch diskrete Eichgruppen betrachen, z.B. $\mathbf{Z}_2 = \{+1, -1\}$ mit der üblichen Multiplikation), so hat man über alle Gruppenelemente zu summieren.

Für die Gruppe $G = U(1)$ ist das Haar-Maß einfach: Parametrisiert man die Gruppenelemente durch $U(\alpha) = \exp(2\pi i\alpha)$, so ist das Haar-Maß $\mathrm{d}\mu(U) = \mathrm{d}\alpha$ mit $\alpha \in [0, 1)$. Dieses Maß ist offensichtlich invariant, da die Multiplikation mit einem Gruppenelement nur eine konstante Verschiebung von α bedeutet.

Für die Gruppe $G = SU(2)$ kann man die Darstellung

$$U(\{a_i\}) \;=\; \begin{pmatrix} a_0 + ia_3 & a_1 + ia_2 \\ -a_1 + ia_2 & a_0 - ia_3 \end{pmatrix} \qquad \text{mit} \quad a_i \in \mathbb{R}\,, \quad \sum_{i=0}^{3} a_i^2 \;=\; 1$$

benutzen. (Die Gruppe $SU(2)$ bildet als Mannigfaltigkeit eine 3-Sphäre.) Das invariante Maß für $SU(2)$ ist

$$\begin{aligned} \mathrm{d}U \;&=\; \delta(a_0^2 + a_1^2 + a_2^2 + a_3^2 - 1)\; \mathrm{d}a_0\mathrm{d}a_1\mathrm{d}a_2\mathrm{d}a_3 \\ &=\; [\delta(a_0 - \sqrt{1 - \boldsymbol{a}^2}) \;+\; \delta(a_0 + \sqrt{1 - \boldsymbol{a}^2})]\; \Theta(1 - \boldsymbol{a}^2) \;\times \\ &\qquad\qquad \times \frac{1}{\sqrt{1 - \boldsymbol{a}^2}}\; \mathrm{d}a_0\mathrm{d}a_1\mathrm{d}a_2\mathrm{d}a_3\,. \end{aligned}$$

Allgemein lässt sich das Maß fogendermaßen konstruieren. Jedes Element einer Lie-Gruppe besitzt in der Umgebung eines Elementes U_0 die Parametrisierung

$$U(\{A\}) \;=\; U_0\, \exp(\mathrm{i} \sum_a A^a T_a)\,,$$

wobei im Folgenden oft A^a als reell angenommen wird, für unitäre Gruppen ist somit T_a hermitesch. Die Matrizen T_a bilden die Generatoren der Lie-Algebra. Die Haar-Integration lässt sich dann immer als Integration über A^a schreiben, allerdings sind die Integrationsgrenzen oft nicht einfach anzugeben und zu berücksichtigen. Bei unitären Gruppen erhält man das Haar-Maß aus der Darstellung des invarianten Linienelements:

$$\mathrm{d}^2 s \;=\; \mathrm{tr}\;(\delta U^+ \delta U) \;=\; M[\{A\}]_{ab}\; \mathrm{d}A^a \mathrm{d}A^b\,.$$

Dabei bedeutet $\delta U = U(A + \mathrm{d}A) - U(A)$ und

$$M[\{A\}]_{ab} \;=\; \mathrm{tr}\; \frac{\partial U^+}{\partial A_a} \frac{\partial U}{\partial A_b}\,.$$

Für das Haar-Maß findet man somit

$$\mathrm{d}U \;=\; \sqrt{|\det\, M[\{A\}]_{ab}|}\; \prod_a \mathrm{d}A^a\,. \tag{10.12}$$

10.4 Observable der reinen Eichtheorie

Offensichtlich ist nicht nur die Wirkung eichinvariant, sondern auch das gewählte Maß. Damit folgt für Erwartungswerte von Funktionen der dynamischen Variablen $\{U(\boldsymbol{m}, \mu)\}$

$$\langle F[U(\boldsymbol{m}, \mu)]\rangle \;=\; \langle F[g(\boldsymbol{m} + \boldsymbol{e}_\mu) U(\boldsymbol{m}, \mu) g(\boldsymbol{m})^{-1}]\rangle \tag{10.13}$$

für beliebige $\{g(\boldsymbol{m})\}$. Es ist also nur sinnvoll, Erwartungswerte von eichinvarianten Größen zu betrachten. Eichinvarianten Funktionen der Variablen $\{U\}$ sind die Spuren von Produkten von $\{U\}$'s entlang geschlossener Wege. Wir definieren

$$W[\mathcal{C}] \;=\; \mathrm{tr}\;\prod_{i=1}^{L} U(\boldsymbol{m}_i,\mu_i)\Bigg|_{\mathcal{C}}\;. \tag{10.14}$$

Dabei beschreibt $(\boldsymbol{m}_i,\mu_i)$ mit

$$\boldsymbol{m}_{i+1}=\boldsymbol{m}_i+\boldsymbol{e}_{\mu_i} \qquad \text{und} \qquad \boldsymbol{m}_{L+1}=\boldsymbol{m}_1$$

eine Folge von gerichteten Gitterlinien zu einem geschlossenen Weg $\mathcal{C}$. Die Größen $W[\mathcal{C}]$ sind für geschlossene Wege $\mathcal{C}$ eichinvariant. Außerdem kann man sich leicht überlegen, dass die einzigen eichinvarianten Größen beliebige Funktionen dieser „Paralleltransporte entlang geschlossener Wege“ sind.

Die Bezeichnung der Größen $W[\mathcal{C}]$ ist leider nicht ganz eindeutig. Im Allgemeinen heißen sie „Wilson-Loops“. Manchmal versteht man aber unter einem Wilson-Loop das Produkt der $\{U\}$'s entlang eines geschlossenen Weges (also bevor die Spur in (10.14) genommen wird), manchmal versteht man unter dem Wilson-Loop auch den Erwartungswert von W. Ich werde im Folgenden immer W – definiert wie in (10.14) – als Wilson-Loop bezeichnen. (Wilson-Loops zu geschlossenen Wegen, die sich um ein periodisches Gitter herumwinden, heißen auch „Polyakov-Loops“. Diese spielen eine Rolle, wenn das Gitter im thermodynamischen Grenzfall in einer Richtung endlich gelassen wird, wie es z.B. bei Feldtheorien zu „endlicher Temperatur“ der Fall ist.)

10.5 Die „stringtension“ – ein Ordnungsparameter für Confinement

Sei $W[R,T]$ ein Wilson-Loop zu einem in einer Ebene gelegenen, rechteckigen, geschlossenen Weg mit Kantenlängen R und T. Die Funktion

$$V(R) \;=\; -\lim_{T\to\infty}\;\frac{1}{T}\;\ln\langle W[R,T]\rangle \tag{10.15}$$

wird als „statisches $q\bar{q}$-Potential“ interpretiert. Die Größe

$$\kappa \;=\; -\lim_{R\to\infty}\;\frac{1}{R}\left(\lim_{T\to\infty}\;\frac{1}{T}\;\ln\langle W[R,T]\rangle\right) \;=\; \lim_{R\to\infty}\;\frac{V(R)}{R} \tag{10.16}$$

heißt „stringtension“. κ ist ein Ordnungsparameter zur Unterscheidung zwischen Confinement bzw. Nicht-Confinement in einer reinen Eichtheorie.

Akzeptiert man für $V(R)$ die Interpretation als „Potential“ zwischen zwei geladenen Teilchen der Theorie (auf diesen Punkt gehen wir weiter unten ein), so bedeutet $\kappa\neq 0$, dass die potentielle Energie mit dem Abstand der Separation der Teilchen zunimmt, für asymptotisch entfernte Teilchen also unendlich groß würde, sodass dieser Zustand in einer dynamischen Theorie nicht auftritt. Außerdem kann man sich das lineare Anwachsen der potentiellen Energie dadurch

vorstellen, dass sich zwischen den Teilchen ein „string“ von Ladungsfluss ausbildet, dessen Energie proportional zu seiner Länge ist, der also eine konstante Energiedichte κ hat. Daher der Name „stringtension“.

$\kappa = 0$ bedeutet andererseits, dass die potentielle Energie für sehr große Abstände kaum noch zunimmt. Es sollte also möglich sein, dem System genügend viel Energie zuzuführen, um eine unendliche Separation der Teilchen zu erreichen.

Man beachte, dass in diesen Überlegungen die geladenen Materieteilchen nur als statische, „unendlich schwere“ Objekte auftreten, also nicht an der Dynamik teilnehmen. In der Realität wird die Energie zwischen zwei sich entfernenden Teilchen nur solange zunehmen, bis die potentielle Energie ausreicht, um aus dem Vakuum Teilchen-Antiteilchen-Paare zu erzeugen. Es entstehen somit lokal ladungsneutrale Zustände, die sich dann voneinander entfernen können.

Wilson hatte allgemeiner als Kriterium für Confinement bzw. Nicht-Confinement in einer reinen Eichtheorie das „Flächen-“ bzw. „Umfangs“-Gesetz für den Erwartungswert eines in einer Ebene liegenden Loops vorgeschlagen:

$$\langle W[\mathcal{C}]\rangle \propto \exp(-\text{`area'}) \qquad \text{Confinement} \tag{10.17}$$

$$\langle W[\mathcal{C}]\rangle \propto \exp(-\text{`perimeter'}) \qquad \text{kein Confinement}\,. \tag{10.18}$$

Geht man davon aus, dass $V(R)$ nicht stärker als $\propto R$ ansteigen kann, so erkennt man, dass das Flächengesetz äquivalent zu der Bedingung $\kappa \neq 0$ ist. Entsprechend bedeutet das Umfangsgesetz $\kappa = 0$.

Es fehlt noch eine Begründung, warum $V(R)$ als statisches $q\bar{q}$-Potential interpretiert werden kann.

Sei $H[R]$ der Hamilton-Operator der reinen Eichtheorie mit zwei statischen (entgegengesetzten) Ladungen im Abstand R. Dann ist $V(R)$ der niedrigste Eigenwert von $H[R]$. Diesen erhält man aus

$$V(R) = -\lim_{T\to\infty}\frac{1}{T}\ln\ \langle q\bar{q} \text{ in Distanz } R|\mathrm{e}^{-H[R]T}|q\bar{q} \text{ in Distanz } R\rangle$$

$|q\bar{q}$ in Distanz $R\rangle$ bezeichnet dabei irgend einen Zustand in dem Sektor der Theorie, in welchem sich zwei statische (unendlich schwere) Ladungsträger im Abstand R befinden. Dass der Erwartungswert im wesentlichen durch den Wilson-Loop gegeben ist, soll am Beispiel der abelschen Eichtheorie gezeigt werden. Die reine Eichtheorie in Gegenwart einer äußeren statischen Ladungsverteilung ist gegeben durch

$$Z[j] = \frac{1}{Z}\int[\mathcal{D}A_\mu(x)]\ \exp(\mathrm{i}S[A] + \mathrm{i}\int \mathrm{d}^4x A_\mu(x)j^\mu(x))\,.$$

Wählt man für j_μ einen Strom, der konstant auf einem Loop definiert ist, so folgt

$$\begin{aligned} Z[\mathcal{C}] &= \frac{1}{Z}\int[\mathcal{D}A_\mu(x)]\ \exp(\mathrm{i}S[A] + \mathrm{i}\oint \mathrm{d}x^\mu A_\mu(x)) \\ &= \langle\exp(\mathrm{i}\oint \mathrm{d}x^\mu A_\mu(x)) \\ &= \langle W[\mathcal{C}]\rangle\,. \end{aligned}$$

Der Erwartungswert des Wilson-Loops entspricht also genau der Zustandssumme in Gegenwart eines statischen Stroms.

Wilson hatte eine zweite Interpretation gegeben, die grundsätzlich dasselbe aussagt, aber einen anderen Gesichtspunkt betont:

Der Propagator zweier Quarks (die über Eichfelder miteinander wechselwirken) ist gegeben durch die Summe über alle Wege dieser beiden Quarks von einem Anfangs- zu einem Endzustand. Jeder Weg erhält dabei ein Gewicht, welches durch den Erwartungswert des entsprechenden Wilson-Loops gegeben ist. Erfüllt dieser Erwartungswert ein Flächengesetz, so sind Quark-Konfigurationen, bei denen die Quarks einen Abstand R haben, exponentiell unterdrückt, Quarks sind also „confined“. Erfüllt der Erwartungswert hingegen ein Umfangsgesetz, so können auch weit entfernte Quarks mit relativ hoher Wahrscheinlichkeit auftreten: Es gibt kein Confinement.

10.6 Der Grenzwert großer Eichkopplung – die „large coupling“-Entwicklung

Die Wirkung der reinen Eichtheorie

$$S_{\text{gauge}} \;=\; -\,\beta \sum_p \text{tr}\,\left(1 - \frac{1}{2}(U_p + U_p^+)\right) \qquad\qquad \text{mit} \quad \beta \;=\; \frac{1}{g^2 N}$$

hat nur einen freien Parameter: die Kopplung g, deren Bedeutung man in einem statistischen System mit einer Temperatur $\frac{1}{T} = \beta = \frac{1}{g^2 N}$ vergleichen könnte. Ich möchte im Folgenden zeigen, dass das führende Verhalten des Wilson-Loops für große Kopplung g (hohe Temperatur bzw. kleiner Wert für β) durch ein Flächengesetz bestimmt ist.

Für die weiteren Überlegungen benötigen wir im wesentlichen die folgenden Eigenschaften des Haar-Maßes:

$$\begin{aligned} \int \mathrm{d}U &= 1\,, \\ \int \mathrm{d}U\, U_{\alpha\beta} &= 0\,, \\ \int \mathrm{d}U\, U_{\alpha\beta} U^+_{\gamma\delta} &= c\;\delta_{\alpha\delta}\delta_{\beta\gamma}\,. \end{aligned} \tag{10.19}$$

Die erste Bedingung ist die Normierung des Haar-Maßes bei kompakten Gruppen: das Gruppenvolumen soll 1 sein. Die zweite Bedingung gilt in jeder nicht-trivialen Darstellung einer Gruppe und folgt aus der Invarianz des Haar-Maßes. Die genaue Form der dritten Gleichung werden wir nicht benötigen, wesentlich ist lediglich, dass es in diesem Fall nicht-verschwindende Matrixelemente nach der Haar-Integration gibt.

Wir entwickeln das Gewicht nach Potenzen von $1/g^2$:

$$\begin{aligned} \exp\left(-\frac{1}{g^2 N}\sum_p \text{tr}\;(1 - \tfrac{1}{2}(U_p + U_p^+))\right) \;&= \\ &= \;\prod_p \left(1 - \frac{1}{g^2 N}\text{tr}\;(1 - \tfrac{1}{2}(U_p + U_p^+) \;+\; \mathcal{O}(1/g^4)\right)\,. \end{aligned}$$

Der führende Term im Nenner des Erwartungswerts des Wilson-Loops (also der führende Term der Zustandssumme) ist durch den führenden Term dieser Entwicklung gegeben, d.h.

$$Z \;=\; 1 \;+\; \mathcal{O}(1/g^4)\,.$$

Für den Zähler erhalten wir:

$$\int \prod_{m,\mu} \mathrm{d}U(\boldsymbol{m},\mu)\ \mathrm{tr}\ [U(l_1)U(l_2)\cdots U(l_L)] \prod_p \left(1 - \frac{1}{g^2 N}\mathrm{tr}\ (1 - \tfrac{1}{2}(U_p + U_p^+)\ +\ \ldots\right).$$

Dabei bezeichnen $l_i = (\boldsymbol{m}_i, \mu_i)$ $(i = 1, \ldots, L)$ die (orientierten) Gitterlinien eines geschlossenen Weges $\mathcal{C}$ der Länge L. Wegen der zweiten Eigenschaft des Haar-Maßes (10.19) verschwindet im Allgemeinen der erste Term dieser Entwicklung, es sei denn, jedes Gruppenelement $U(l_i)$ entlang des Weges tritt zweimal auf, d.h., es handelt sich um einen Loop, der einen Weg vorwärts und zurück durchläuft. In diesem Fall verschwindet natürlich die eingeschlossene Fläche. Der führende nicht-verschwindende Term ist dadurch gegeben, dass jedes $\{U\}$ mindestens zweimal auftritt, und zwar jeweils als konjugiertes Element, d.h. bezüglich entgegengesetzter Richtung der Gitterlinien.

In diesem Fall muss von dem Gewicht ein Produkt von Plakett-Variablen beitragen, welches gerade den Wilson-Loop schließt, d.h. eine Fläche (auf dem hyperkubischen Gitter), welche den Wilson-Loop zum Rand hat, wird mit Plaketten ausgepflastert. Der führende Beitrag zum Zähler – und damit zum Erwartungswert – ist somit

$$\langle W[\mathcal{C}]\rangle \;\propto\; \left(c^2 \frac{1}{g^2 N}\right)^A \;=\; \exp -(\ln \tfrac{g^2 N}{c^2})A\,,$$

wobei A die Anzahl der Plaketten ist, welche zu der Minimalfläche mit dem Loop $\mathcal{C}$ als Rand gehört. Für einen rechteckigen Wilson-Loop ist $A = RT$. Der Beitrag c^{2A} ist der Beitrag der $\{U\}$-Integration, wobei $2A$ (in führender Ordnung) der Anzahl der Gitterlinien innerhalb dieser Minimalfläche entspricht. Der genaue Wert von c sowie der Proportionalitätsfaktor hängen von der Art der Gruppe sowie der Art des Loops ab.

Man kann das obige Argument noch verfeinern und sich überlegen, dass jeder nicht-verschwindende Beitrag zum Zähler des Wilson-Loop-Erwartungswertes ein Produkt von Plaketten darstellt, die gerade eine Fläche ausmachen mit dem Loop $\mathcal{C}$ als Berandung. Man erhält allerdings besondere Gewichte für den Fall der Selbstüberschneidung bzw. der Berührung solcher Flächen, da z.B. die Gruppenintegration über Terme der Art UU^+UU^+ nicht gleich dem Quadrat der Integration über UU^+ ist. Außerdem gibt es Beiträge von geschlossenen Flächen ohne Rand, allerdings werden diese durch den Nenner aufgehoben. Lediglich bei geschlossenen Flächen, welche die vorgegebene Fläche mit Rand $\mathcal{C}$ berühren oder schneiden, gibt es gesonderte Beiträge. Diese (kombinatorisch komplizierte) Darstellung des Erwartungswertes hat unter anderem dazu geführt, formal im Kontinuum den folgenden Ansatz zu erwägen:

$$\langle W[\mathcal{C}]\rangle \;=\; \sum_{\text{Flächen mit Rand } \mathcal{C}} \exp -(\gamma A)\,.$$

A ist der jeweilige Flächeninhalt und γ zunächst nur ein Parameter (ähnlich wie μ bei der entsprechenden Darstellung der 2-Punkt-Funktion als Summation über alle Wege, Gl. (8.13)). Man wird jedoch erwarten, dass γ durch die physikalische Stringtension ausgedrückt werden kann. Trotz vieler Ansätze ist es bisher jedoch noch nicht gelungen, diesem formalen Ausdruck einen mathematisch wohldefinierten Sinn zu geben.

Der führende Beitrag für große Werte der Kopplung g zum Erwartungswert des Wilson-Loops ist also mit der Fläche des Loops exponentiell unterdrückt. Diese Aussage wäre jedoch belanglos, könnte man nicht beweisen, dass die Entwicklung in $1/g^2$ einen nicht-verschwindenden Konvergenzradius besitzt. Dieser Beweis wurde von Osterwalder und Seiler [86] für reine Eichtheorien gegeben und in der Folge verallgemeinert (siehe z.B. Fradkin und Shenker [26] und die dortigen Referenzen). Für genügend große Werte der Kopplung besitzt somit jede Gittereichtheorie eine so genannte Confinement-Phase, in der für den Wilson-Loop das Flächengesetz gilt.

Dies wirft ein Problem auf: Sowohl die $U(1)$-Eichtheorie als auch die $SU(3)$-Eichtheorie zeigen für große Werte der Kopplung Confinement. Wenn die Gitterregularisierung sinnvoll ist, und die entsprechenden Modelle tatsächlich den reinen Eichsektor der Quantenelektrodynamik bzw. Quantenchromodynamik beschreiben, so müssen sich die beiden Theorien hinsichtlich ihrer Struktur als Funktion der Kopplung g unterscheiden. Wir werden noch sehen, dass der Kontinuumslimes für $g \to 0$ angenommen wird. Für die $U(1)$-Theorie erwartet man also einen Phasenübergang von einer „Confinement"- zu einer so genannten „Freie Ladung"-Phase, während ein solcher Phasenübergang in der $SU(3)$ Theorie nicht existieren sollte. Die numerischen Daten bestätigen diese Vermutung, allerdings existiert noch kein Beweis.

Auch eine Entwicklung für kleine Werte der Kopplung g hilft in diesem Fall nicht weiter. Für $g \to 0$ kann man zwar die führenden Terme der Störungsrechnung bestimmen (es handelt sich um ganz gewöhnliche Kontinuumsstörungstheorie) und erhält tatsächlich immer ein Coulomb-artiges Potential zwischen statischen Ladungen. Allerdings lässt sich in diesem Fall nicht beweisen, dass diese Störungsreihe einen nicht-verschwindenden Konvergenzradius besitzt. Aus allgemeinen Überlegung kann man sogar schließen, dass $g = 0$ einer wesentlichen Singularität der Theorie entspricht (vgl. Dyson [24]) und somit keine konvergente Entwicklung um diesen Punkt möglich ist.

10.7 Experimentelle Bestimmung der Stringtension und des $q\bar{q}$-Potentials

In diesem Abschnitt wollen wir kurz diskutieren, wie sich das $q\bar{q}$-Potential sowie die Stringtension κ experimentell bestimmen lassen. κ hat die Dimension [Energie]^2, bzw. $\sqrt{\kappa}^{-1}$ hat die Dimension einer Länge. Die Theorie erlaubt natürlich zunächst keine Vorhersagen für eine dimensionsbehaftete Größe, sondern κ kann lediglich dazu dienen, eine Beziehung zwischen der (dimensionslosen) Kopplungskonstanten g und einer physikalischen Längeneinheit (Gitterkonstanten a) herzustellen. Einzelheiten dazu werden in Kapitel 16 behandelt.

Experimentell lässt sich κ beispielsweise aus dem Spektrum massiver $q\bar{q}$-Zustände bestimmen. Es eignen sich besonders die gebundenen Charm-Anti-Charm- ($c\bar{c}$) und Bottom-Anti-Bottom-Zustände ($b\bar{b}$). Die Massen von Charm und Botton sind ausreichend schwer, sodass bei einer Diskussion ihrer gebundenen Zustände relativistische Effekte vernachlässigt werden können.

Bestimmt man das Energiespektrum eines Teilchens in einem Potential der Form $V(R) = \kappa R$ und vergleicht es mit den beobachteten Spektren der $c\bar{c}$- bzw. $b\bar{b}$-Zustände (den Ψ- (J/Ψ) bzw. Υ-Teilchen), so lässt sich für die Stringtension der Wert $\kappa \approx (440\,\mathrm{eV})^2$ bestimmen. Ähnliche Werte erhält man auch aus den Steigungen der so genannten Regge-Trajektorien.

Wie schon erwähnt, kann man diesen Wert nicht direkt mit den Gitterergebnissen vergleichen, sondern nur (analog zu der Diskussion der freien Klein-Gordon-Theorie) zur Anpassung einer Gitterkonstanten als Funktion der Kopplung g heranziehen. In Kapitel 15 werden wir weitere dimensionsbehaftete Observable für die reine Eichtheorie untersuchen (das Glueball-Massenspektrum). Nachdem über die Stringtension der Gitterkonstanten eine physikalische Länge zugeordnet wurde, sind die Resultate für das Glueball-Massenspektrum natürlich direkt mit dem Experiment vergleichbar.

Kapitel 11

Ising-Dualities

Duality transformations play a special role in many quantum field theories as well as in string theories. Usually they refer to the fact that there is a formulation of the theory involving a coupling constant λ (together with a perturbative formulation for $\lambda \to 0$) and a second formulation of the same theory where the relevant coupling constant is $1/\lambda$ (with a correpsonding perturbation expansion valid for $1/\lambda \to 0$). These two models are then referred to as dual to each other. They correspond to the same theory but the formulations as statistical models or quantum field theories my look completely different.

Duality transformations are well known from the Ising model. We will first briefly recall the high- and low-temperature expansion of the two dimensional Ising model. In this particular case the model is self-dual. Then we will introduce the generalized class of Ising models (or rather $\mathbf{Z}_2$ models) first formulated by Wegner [99], most of which are generalized gauge theories. Finally, we will show that any (generalized) $\mathbf{Z}_2$ model on a d-dimensional hypercubic lattice has a dual model which is again a $\mathbf{Z}_2$ model on the same lattice (actually on the dual lattice, but the hypercubic lattice is self-dual).

11.1 The 2-dimensional Ising model

The model has first been mentioned in a paper by Lenz in 1920. He asked his student E. Ising to investigate it in more detail [46]. In 1941, Kramers and Wannier calculated the exact value for the critical temperature of the phase transition using the duality transformation [61], and in 1944 Onsager found the first exact solution of the 2-dimensional Ising model (with vanishing external field) [84]. The 2-dimensional Ising model with non-vanishing field has not been solved yet, neither has been the Ising model in more than 2 dimensions.

We consider a 2-dimensional square lattice (if necessary, we define periodic boundary conditions on a toroidal lattice). To each lattice site $i = (n_1, n_2)$ is associated a classical spin $s_i \in \{+1, -1\}$. The energy of the Ising model is defined by

$$E[s] = -J \sum_{\langle i,j \rangle} s_i s_j$$

where the sum extends over all nearest neighbor sites on the lattice. The partition function is

given by

$$Z(\beta) = \sum_{\{s_i=\pm 1\}} \exp\left(\beta J \sum_{\langle i,j\rangle} s_i s_j\right) .$$

In the following we will omit the energy constant J which can be absorbed in a redefinition of β.

11.1.1 The high temperature expansion

We note that for any of the two values the product $s_i s_j$ can assume we have

$$\mathrm{e}^{\beta s_i s_j} \;=\; \cosh\beta \;+\; s_i s_j \sinh\,\beta \;=\; \cosh\beta\,(1 + s_i s_j \tanh\beta)\,.$$

This allows us to rewrite the partition function as:

$$\begin{aligned} Z &= (\cosh\beta)^{2L^2} \sum_{\{s_i=\pm 1\}} \prod_{\langle i,j\rangle} (1 + s_i s_j \tanh\beta) \\ &= (\cosh\beta)^{2L^2} \sum_{\{s_i=\pm 1\}} \sum_{\{\epsilon_{\langle i,j\rangle}=0,1\}} \left(\prod_{\langle i,j\rangle} (s_i s_j \tanh\beta)^{\epsilon_{\langle i,j\rangle}}\right) . \end{aligned}$$

($2L^2$ is the number of edges on a $L\times L$ square lattice.)

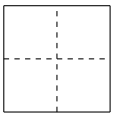
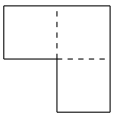
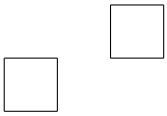
If we expand the product in the bracket we can represent every term by a "shaded" lattice configuration where to each lattice edge is associated a 0 or a 1 (thin line or thick line). The summation over $\epsilon_{\langle i,j\rangle} = \pm 1$ corresponds to a summation over all these configurations. Now we first sum over the spin variables. This yields either 0 (whenever one of the spin variables occurs with an odd power, i.e., whenever one of the thick lines has an end point. Otherwise, i.e. whenever all thick lines form closed loops, the contribution is the product of the hyperbolic tangent functions and a factor 2 for each sum over a spin. Obviously what one obtains is a summation over all closed loops (undirected) on the lattice. Each loop of length l is weighted by $(\tanh\beta)^l$. Hence, we obtain

$$Z \;=\; 2^{L^2} (\cosh\beta)^{2L^2} \sum_{\text{closed loops}} (\tanh\beta)^l \;=\; (2\cosh^2\beta)^{L^2} \sum_l P(l) (\tanh\beta)^l \quad . \tag{11.1}$$

$P(l)$ is the number of closed loops of length l. There can be multiple loops which even intersect, however, each edge of the lattice can only be part of one loop.

Note that this high temperature representation of the partition function of the Ising model holds for any dimension and any type of lattice (even for general graphs).

The following table summarizes the contributions to this sum up to order 8 (i.e., it includes all loops of total length 8):

length	number	shape	length	number	shape
0	1		8	L^2	
4	L^2			$2\ L^2$	
6	$2L^2$			$4\ L^2$	
				$\frac{L^2(L^2-5)}{2}$	

The the high temperature expansion of the partition function and the free energy starts with ($\tau = \tanh\beta$):

$$\begin{aligned} Z &= (2\cosh^2\beta)^{L^2}\ \left(1 + L^2\tau^4 + 2L^2\tau^6 + L^2(7 + \tfrac{1}{2}(L^2-5))\tau^8 + \cdots\right) \\ F_\infty &= -\lim_{L\to\infty}\frac{1}{\beta L^2}\ln Z \\ &= -\frac{1}{\beta}\ \left(\ln(2\cosh^2\beta)\ + \tau^4 + 2\tau^6 + \tfrac{9}{2}\tau^8 + \cdots\right) \quad . \end{aligned} \tag{11.2}$$

11.1.2 Low temperature expansion

The next step is the derivation of a low temperature expansion ($\beta \to \infty$). We extract a configuration independent factor in front of the sum:

$$Z\ =\ \mathrm{e}^{2\beta L^2} \sum_{\{s_i=\pm 1\}} \mathrm{e}^{\beta\sum_{\langle i,j\rangle}(s_i s_j - 1)} \quad . \tag{11.3}$$

Obviously, now the weight of each configuration is determined by the number of lattice lines which connect two points with opposite spin orientation. Each such line contributes a factor of $\mathrm{e}^{-2\beta}$ to this configuration.

We now switch to the so-called *dual lattice*: The sites of the dual lattice correspond to the plaquettes of the original lattice and vice versa. The edges of the dual lattice connect those

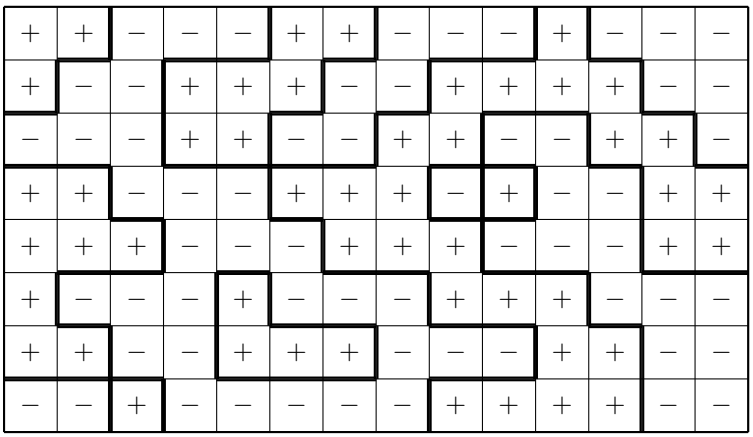


Abbildung 11.1: Boundary lines between regions of parallel spins on the dual lattice.

sites for which the corresponding dual plaquettes were neighbored. For a square lattice the dual edges are perpendicular to the original lattices.

Now the spin variables are defined on the plaquettes of the dual lattice and we can mark the boundary lines between plaquettes of different spin orientation (see fig. 11.1). These boundary lines form closed loops. As a total inversion of the spin orientation gives rise to the same boundary loops, we get an overall factor of 2. Otherwise, each configuration defines exactly one configuration of closed loops as boundaries, and vice versa, each configuration of closed loops defines (up to the above mentioned sign) a configuration of spins. The boundary lines obviously have the same properties as the loops of the high temperature expansion. We obtain for the partition function:

$$Z \;=\; 2\,\mathrm{e}^{2\beta L^2} \sum_{\text{closed loops}} \mathrm{e}^{-2\beta l} \quad . \tag{11.4}$$

The expansion is essentially the same as for high temperatures, where now, however, the weight becomes small for low temperatures (or high values of β). The low temperature expansion for the free energy is given by:

$$-\beta F_\infty \;=\; 2\beta \;+\; \mathrm{e}^{-8\beta} \;+\; 2\,\mathrm{e}^{-12\beta} \;+\; \tfrac{9}{2}\,\mathrm{e}^{-16\beta} + \ldots \quad . \tag{11.5}$$

The low temperature expansion depends on the dimension of the lattice. For a 3-dimensional cubic lattice, boundaries between different spin orientations are closed surfaces and, more general, in d dimensions these boundaries are closed objects of co-dimension 1.

11.1.3 The self-duality of the 2-dimensional Ising model

Up to now most considerations were largely independent of the dimension of the system. Now, however, we want to make use of the fact that the low temperature expansion and the high temperature expansion are expressed as summations over the same configurations (with the same combinatorics). They only differ in their respective weights.

We define the dual temperature by

$$\mathrm{e}^{-2\beta^*} \;=\; \tanh\beta \qquad \text{or} \qquad \beta^* \;=\; -\,\frac{1}{2}\ln(\tanh\beta) \tag{11.6}$$

This transformation is an involution:

$$\beta^{**} \;=\; -\,\frac{1}{2}\ln(\tanh\beta^*) \;=\; -\,\frac{1}{2}\ln[\tanh(-\tfrac{1}{2}\ln(\tanh\beta))] \;=\; \beta \quad .$$

If β is large, then β^* is small and vice versa.

Taking the low and high temperature expansions,

$$Z(\beta) \;=\; 2\,\mathrm{e}^{2L^2\beta}\sum_{l=0}^{\infty} P(l)\mathrm{e}^{-2\beta l} \qquad \text{low temperature} \tag{11.7}$$

$$Z(\beta) \;=\; (2\cosh^2\beta)^{L^2}\sum_{l=0}^{\infty} P(l)(\tanh\beta)^l \qquad \text{high temperature} \tag{11.8}$$

and using the definition of the dual temperature we obtain a relation between $Z(\beta)$ and $Z(\beta^*)$:

$$Z(\beta) \;=\; \frac{(2\cosh\beta\sinh\beta)^{L^2}}{2}\; Z(\beta^*) \quad . \tag{11.9}$$

If we now assume that the Ising model in two dimensions has only a single phase transition (characterized by a non-analyticity of the free energy), we can obtain an equation for the value of β where this phase transition occurs. It has to be the fixed point of the duality transformation:

$$\mathrm{e}^{-2\beta} \;=\; \frac{1-\mathrm{e}^{-2\beta}}{1+\mathrm{e}^{-2\beta}}$$

which yields

$$\beta_c \;=\; \frac{1}{2}\ln(1+\sqrt{2}) \;\approx\; 0.440687 \quad . \tag{11.10}$$

11.2 General types of $\mathbf{Z}_2$-models

Already in 1971, Wegner published a paper with a series of Ising-type models. They all were formulated with $\mathbf{Z}_2$ variables (i.e., the variables only assume the values $+1$ and -1), but the variables were attributed to different structures of the lattice, which in the following we will call cells. For the simplest model, the Ising model, the spin variables are assigned to the lattice sites. In the more general models they are attributed to the lattice lines, plaquettes, hyperplaquettes etc.

We now consider a d-dimensional hypercubic lattice. For each $0 \le D \le d$ we obtain certain types of cells:

- $D = 0$: lattice sites
- $D = 1$: lattice edges, equivalent to undirected pairs of nearest neighbor lattice sides.

- $D = 2$: plaquettes, equivalent to four lattice edges which form a closed loop (i.e. which have no boundary).
- $D = 3$: cubes, defined by eight lattice sites or six lattice plaquettes which form a closed cube

For general D we can define the D-dimensional hypercube as consisting of the points:

$$(n_1, ..., n_{i_1} + \epsilon_1, ..., n_{i_2} + \epsilon_2, ..., n_{i_3} + \epsilon_3, ...)$$

where $\epsilon_i = 0$ or 1. These are 2^D points which make up the D-dimensional hypercube (in this case labeled by one of its edges $n = (n_1, ..., n_d)$). Each D-dimensional hypercube has a boundary consisting of $D-1$-dimensional cells. If we denote a D-dimensional hypercube by α, we denote the set of $D-1$-dimensional cells making up the boundary of α by $\partial\alpha$.

Now we can formulate the general class of Ising models: The spin variables are associated to $D-1$ cells which we label by $\{l\}$. The D dimensional hypercubes are labeled by $\alpha, ...$ The energy is defined as

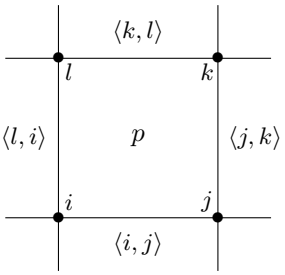
$$E = -J \sum_{\alpha} \prod_{l \in \partial\alpha} s_l \, .$$

The summation extends over all D-dimensional hypercubes and each term consists of a product of spin variables sitting on the boundary of this hypercube.

Obviously, the case $D = 1$ corresponds to the classical Ising model: the spins are attached to the 0-cells, i.e., the sites of the lattice, and the summation extends over all lattice edges with an energy contribution from the spins on the two sites which form the boundary of an edge.

The example $D = 2$ will be of particular importance. It corresponds to the Ising gauge theory. In this case the spin variables are attached to the edges of the lattice. For later purpose we label an edge by its two end points $l = \langle i, j \rangle$, and a plaquette by its four sites $p = \langle i, j, k, l \rangle$ (see fig.). The energy is given by:

$$E = -J \sum_{p = \langle i,j,k,l \rangle} s_{\langle i,j \rangle} s_{\langle j,k \rangle} s_{\langle k,l \rangle} s_{\langle l,i \rangle} \, . \tag{11.11}$$



The summation runs over all plaquettes of the hypercubic lattice and for each plaquette there is a contribution equal to the product of the four spins sitting on the edges surrounding the plaquette.

The generalized $\mathbf{Z}_2$-models possess a gauge invariance for $D \geq 2$. The case $D = 1$ is excluded: in this case the model has only a global $\mathbf{Z}_2$ invariance realized by switching the signs of all spins.

In order to obtain the gauge invariance for $D \geq 2$ (where the spin variables sit on $D-1$-cells), we consider the $D-2$-cells of the lattice and associate to each of these cells (labeled by $i, j, ...$) an element of the group $\mathbf{Z}_2$. So, for $g_i \in \mathbf{Z}_2$ we have $g_i = +1$ or $g_i = -1$. We now transform each spin variable s_l (siting on a $D-1$ cell l) by:

$$s_l \longrightarrow \prod_{i \in \partial l} g_i s_l$$

i.e., we multiply each spin variable by the product of the group elements on the boundary of the cell on which the spin variable is defined. We claim that the energy is invariant under this transformation of spins.

For a proof, notice that for a hypercube α the boundary cells in $\partial\alpha$ form a closed surface. This means that all the ($D-2$ dimensional) boundary elements of the $D-1$ dimensional cells in $\partial\alpha$ are pairwise matched. If we take the product of the spin variables on the boundary cells of α, each $D-2$ boundary of these cells occurs twice. This means that each g_i occurs twice, and as $g_i^2 = 1$ this has no effect.

Lets consider the Ising gauge theory as an example. At each lattice site i we can choose freely a group element $g_i = +1$ or $g_i = -1$. We now transform each spin variable $s_{\langle i,j\rangle}$ by:

$$s_{\langle i,j\rangle} \longrightarrow g_i s_{\langle i,j\rangle} g_j \, .$$

For the energy (eq. 11.11) one takes the product of the four spin variables siting on the edges around a plaquette. Each of these edges has two endpoints, and each of these endpoints occurs exactly twice. Therefore, this transformation does not change the energy. As the group elements can be chosen locally (at each lattice site independently), this is a local invariance group. The total invariance group is $\mathbf{Z}_2^{|V|}$, where $|V|$ is the number of lattice points.

Wegner was interested in order parameters in order to distinguish different phases. We will will come back to this point in the context of the Wilson loop. He noticed that the spins themselves are no good order parameters because they are not gauge invariant. The expectation value of any non gauge invariant quantity will be zero (or rather the group average of this variable). Therefore, Wegner defined products of spin variables sitting on closed cell structures. For the Ising gauge model, e.g., one can define the product of spin variables along an arbitrary closed loop. This will be gauge invariant. Similarly, for the higher order $\mathbf{Z}_2$-models one can define the product of spin variables along the boundaries of arbitrary assemblies of D-cells. These boundaries are sets of $D-1$-cells which have no boundary themselves and, therefore, the products of the spins on these boundaries are gauge invariant.

11.3 High- and low temperature expansion for generalized $\mathbf{Z}_2$ models

We will start with the high temperature expansion (small values of β) and proceed as before. The product of spin variables (attached on $D-1$-cells) along the boundary of a D-cell is again a $\mathbf{Z}_2$ variable, i.e., it assumes only the values $+1$ and -1. Therefore, we can write:

$$\begin{aligned} \exp\left(\beta \prod_{l\in\partial\alpha} s_l\right) &= \cosh\beta + \left(\prod_{l\in\partial\alpha} s_l\right)\sinh\beta \\ &= \cosh\beta\left(1 + \left(\prod_{l\in\partial\alpha} s_l\right)\tanh\beta\right) . \end{aligned}$$

(We have absorbed the coupling constant J into β.) As we did in the case of the Ising model we write

$$Z = (\cosh\beta)^{N_D} \sum_{\{s_i=\pm1\}} \prod_{\alpha} \left(1 + \left(\prod_{l\in\partial\alpha} s_l\right)\tanh\beta\right) , \tag{11.12}$$

where

$$N_D = \binom{d}{D} L^d$$

is the number of D-cells on the lattice. The expansion of the product will lead to a summation over all configurations over "shaded" D-cells (each D-cell "thin" or "thick"). The summation over spin configurations will give zero for each such configuration for which not every spin variable (sitting on $D-1$-cells) does occur with an even power. This is only the case for "closed" D-cell configurations, because there should be no boundary element of a D-cell left. So we can represent the partition function in the high temperature expansion by a summation over all closed D-cell configurations:

$$Z = 2^{N_{D-1}L^d} (\cosh\beta)^{N_D} \sum_{V=0}^{\infty} P_D(V_D) (\tanh\beta)_D^V \,, \tag{11.13}$$

where V_D is the "volume" of the configuration (= the number of D-cells it consists of) and $P_D(x)$ is the combinatorical factor of the number of closed D-cell configurations with volume x. The factor $2^{N_{D-1}L^d}$ comes from summing over all spins. Again, these configurations can have multiple components, but each D-cell can occur at most once.

Now we come to the low temperature expansion. As before we extract a constant energy contribution in front of the summation:

$$Z(\beta) = \mathrm{e}^{\beta N_D} \sum_{\{s_l=\pm 1\}} \exp\left(\beta \sum_{\alpha} \left(\prod_{l\in\partial\alpha} s_l - 1\right)\right) . \tag{11.14}$$

Each spin configuration gets a weight $\exp(-2\beta)N'$, where N' is the number of D-cells for which the product of spin variables along the boundary of the D-cell is -1.

Note that the product of spins around the boundary of a D-cell is a gauge invariant quantitiy, so the D-cell contribution to the energy is gauge invariant. In particular, N' is gauge invariant.

Replacing in a next step the summation over spin configurations by a summation over admissible configurations with ± 1's in D-cells (not every assignment of this type can be realized, as we shall see in a moment) means that we have to take into account separately the number of spin configurations which are gauge equivalent. In general, this is a problem for gauge theories, because this number is not the same as the "volume of the gauge group" (because some gauge transformations may not change a configuration or, even worse, the number of gauge equivalent configurations depends on the equivalence class). For the Ising gauge model ($D = 2$) the gauge group consists of all $\{g_i = \pm 1\}$ and the volume of the gauge group is $\mathcal{V} = 2^{L^d}$. However, apart from the identity transformation there is a second transformation which does not change the spins, which is $g_i \equiv -1$. Therefore, the number of gauge equivalent configurations is 2^{L^d-1}. For the higher order gauge models of Wegner this becomes (slightly) more complicated, because now the gauge transformations which do not change the spin configuration form itself a gauge group. In order not to become lost in details, we simply will write $\mathcal{V}$ for the number of gauge equivalent configurations. This number does not depend on any parameters of the theory (like β) and, therefore, leads to a constant in the free energy.

Now we want to characterize those configurations with ± 1 sitting on the D-cells which actually can be obtained by a spin configuration on $D-1$-cells. Not every assignment of ± 1 to

D-cells is possible. To give a simple example, consider the ordinary Ising model: It is impossible to realize a single -1 on one edge only and all other edges take the values $+1$. We have seen in the last section, that the edges with -1 form a kind of boundary (rather a co-boundary) between spin regions where $s_i = +1$ and regions with $s_i = -1$. This idea will now be generalized.

The statement is the following: Only those assignments of -1 to D-cells will occur as actual configurations stemming from the product of spins on $D-1$-cells, for which the lattice duals of the D-cells (these are $d-D$-cells) form a closed (hyper)-surface. So we can characterize any admissible configuration by closed (hyper)-surfaces of $d-D$ cells, and, vice versa, any closed hyper-surface gives rise to an admissible configuration. Furthermore, two distinct closed hyper-surfaces of $d-D$-cells will give rise to different configurations.

In this way we can write the low temperature expansion of the partition function of a generalized Ising gauge model as:

$$\begin{aligned} Z &= 2(\cosh\beta)^{N_D}\mathcal{V} \sum_{\text{closed } (d-D)\text{-surfaces}} \exp\left(-2\beta V_{d-D}\right) \\ &= 2(\cosh\beta)^{N_D}\mathcal{V} \sum_{V_{d-D}=0}^{\infty} P_{d-D}(V_{d-D})\, e^{-2\beta V_{d-D}}\,. \end{aligned}$$

where V_{d-D} is the number of $d-D$-surfaces in a configuration, and $P_{d-D}(x)$ is equal to the number of closed $d-D$ surfaces with x elements. (The factor of 2 will be explained below.)

We still have to give a proof that the not gauge equivalent configurations in a generalized Ising model with spins sitting on $D-1$-cells can be characterized in a one-to-one fashion by closed contours of $d-D$ cells. (These contours are sometimes referred to as Peierl's contours). We prove this "by induction". Lets suppose there is only one spin on a $D-1$ cell for which $s_l = -1$, all other spins take the value $+1$. In this case the D-cells for which the product of boundary spins is -1 are exactly those D cells which are in the co-boundary of the single $D-1$ cell, i.e. all D cells which have this $D-1$ cell as a boundary. We now switch to the dual lattice formulation, i.e. the spins are now assigned to $d-D+1$-cells and the energy contributions sit on the $d-D$-cells. If in the original formulation one single spin on a $D-1$-cell is negative, the dual energy contributions (on $d-D$-cells) which take the value -1 are the boundary of the $d-D+1$ cell which is dual to the original $D-1$-cell. As the boundary of a boundary is zero, these $d-D$-cells form a closed (hyper)-surface.

In a next step we show that the sum of two closed hyper-surfaces (modulo 2) is again a hyper-surface. If these two hyper-surfaces have no elements in common this is trivial. If they do have elements in common, these are taken away. Now the two hyper-surfaces are glued together at the boundary of the elements which have been taken away and become one close hyper-surface. So we have shown that any admissible spin configuration gives rise to a closed hyper-surface on the dual lattice.

The converse is also true: Given a closed hyper-surface of $d-D$-cells. Up to an overall sign (the factor of 2 in the partition function) this gives rise to a configuration of spins on $d-D+1$-cells: Assign to one of these cells arbitrarily a value of $s = +1$. Now proceed successively: Assign to a neighbor of a cell the same value if these two cells are separated by a boundary element which is not part of the closed hyper-surface, otherwise assign the opposite value. As this can be done consistently, this concludes the proof.

11.4 Duality structure of generalized $\mathbf{Z}_2$ models

We see that a generalized Ising gauge theory in d dimensions for which the spins are attached to $D-1$ cells has a high temperature expansion in terms of a summation over closed D-cells and a low temperature expansion in terms of closed $d-D$-cells. The dual temperatures are related in the same way as in the 2-dimensional Ising model (see eq. 11.6):

$$\tanh \beta^* = e^{-2\beta} .$$

The corresponding dual theory has a high temperature expansion over $d-D$ cells (therefore, the spins sit on $d-D-1$-cells) and a low temperature expansion over D cells. We see that there is a general duality between generalized Ising models with spins on $D-1$ cells and with spins on $d-D-1$-cells. For $d=2$ and $D=1$ this leads to the self-duality of the Ising model. Furthermore, we see that the Ising gauge theory (with $D=2$) in $d=3$ dimensions is dual to an ordinary Ising model in $d=3$ dimensions (and vice versa). The Ising gauge theory in $d=4$ dimensions is self-dual again.

These duality relations are far from trivial. Lets consider the $d=3$ case again. The ordinary Ising model is dual to the Ising gauge theory. These two models have "completely" different structures, as the following table shows:

Model	Spins attached to	number of configurations	Symmetry group
$d=3$ Ising	vertices (0-cells)	2^{L^3}	$\mathbf{Z}_2$
$d=3$ Ising gauge	edges (1-cells)	2^{3L^3}	$(\mathbf{Z}_2)^{L^3}$

On the perturbative level, once the model has been expressed in terms of closed D-cell hypersurfaces, the duality corresponds to Poincaré duality.

For all those generalized Ising models which are self-dual (assuming that there is only one phase transition), the critical temperature of the phase transition is given by eq. (11.10). Self dual models are: Ordinary Ising ($D=1$) in $d=2$. Ising gauge theory ($D=2$) in $d=4$, and, more general, $\mathbf{Z}_2$-models with D in $d=2D$. In all other cases the duality transformation is between different types of generalized Ising models. For these cases we only can say that the critical temperature of one model is equal to the dual critical temperature of the other model.

The following table summarizes the dualities for the dimensions up to $d=4$. Furthermore, If we interprete the summation over closed D-cells as summations over euclidean world sheets, we can associate to each expansion of the generalized Ising model a "continuum" quantum field theory of higher-dimensional objects. The simplest example is again the ordinary Ising model, for which the summation over "world-lines" leads to a particle picture, while the summation over closed surfaces for the Ising gauge theory can be interpreted as a summation of "world-sheets" of (euclidean) strings.

dimension	model	dual to	high temp. exp of QFT	low temp exp. of QFT
$d = 2$	Ising	self-dual	particle	particle
$d = 3$	Ising	Ising gauge	particle	string
	Ising gauge	Ising	string	particle
$d = 4$	Ising	2-cell gauge	particle	2-brane
	Ising gauge	self-dual	string	string
	2-cell gauge	Ising	2-brane	particle

In this sense one can say: Given an 11-dimensional theory for which the high temperature expansion describes strings (summation over closed surfaces) and the low temperature expansion describes 8-branes (summation over closed 9-surfaces), then the Ising gauge theory in 11 dimensions is a candidate for the underlying model.

Kapitel 12

Hamiltonsche Formulierung einer Gittereichtheorie

In Kapitel 5.1 sind wir von einer klassischen Feldtheorie ausgegangen, haben diese über den Funktionalintegralformalismus quantisiert und sind durch die analytische Fortsetzung der Vakuumerwartungswerte zur euklidischen Formulierung der Quantenfeldtheorie gelangt. Diese hatte die Interpretation eines klassischen statistischen Systems für 4-dimensionale Felder. Die Diskretisierung dieses Systems führte uns schließlich zu den Gittertheorien.

In diesem Kapitel soll dieser Weg teilweise in die umgekehrte Richtung gegangen werden. Ausgehend von der Gittertheorie soll ein Zustandsraum (Hilbert-Raum) und ein Hamilton-Operator hergeleitet werden, der im Kontinuumslimes zum Hamilton-Operator der ursprünglichen Quantenfeldtheorie wird. Diese Formulierung eines d-dimensionalen klassischen Gittermodells als $(d-1)$-dimensionales Quanten-Gittermodell ist oft hilfreich bei der physikalischen Interpretation mancher Erwartungswerte. Insbesondere zeigt dieser Formalismus den Übergang vom Gittermodell zur zugehörigen Quantenfeldtheorie.

In der Sprache der Gittermodelle bezeichnet man die Konstruktion des Zustandsraumes und des Hamilton-Operators als Transfermatrix-Formalismus. Unabhängig von der Interpretation als Quantensystem hat der Transfermatrix-Formalismus den Vorteil, ein kombinatorisches Problem – die Berechnung einer klassischen Zustandssumme – in ein algebraisches Problem – die Diagonalisierung einer Matrix, bzw. die Bestimmung des höchsten Eigenwertes einer Matrix – zu übertragen. Die Lösung des 2-dimensionalen Ising-Modells von Onsager erfolgte über den Transfermatrix-Formalismus. In Anhang A4 wird die Diagonalisierung der Transfermatrix des 2-dim. Ising-Modells vorgeführt.

12.1 Der Transfermatrix-Formalismus

Wir werden zunächst die Transfermatrix über ein einfaches Beispiel, das 1-dimensionale Ising-Modell, einführen. Anschließend wird der Transfermatrix-Formalismus allgemein behandelt. Die folgenden Abschnitte skizzieren, wie dieser Formalismus von bestimmten Gittermodellen auf die gewöhnliche Quantenmechanik, eine skalare Quantenfeldtheorie sowie die Quanteneichtheorie führt.

12.1.1 Das 1-dimensionale Ising-Modell

Im Ising-Modell werden den Gitterpunkten lokale Zustände zugeordnet, die nur zwei mögliche Werte annehmen können. Daher werden diese Freiheitsgrade oft als „Spin“ bezeichnet und mit $s_i = \pm 1$ benannt. Im Allgemeinen setzt sich die Energie einer Konfiguration aus einer Summe von Beiträgen zusammen, die eine Wechselwirkung dieser Spins zwischen nächsten Nachbarn auf einem Gitter beschreiben, sowie eines zusätzlichen Beitrags, der beispielsweise ein äußeres Magnetfeld h berücksichtigt:

$$E[\{s_i\}] \;=\; -\sum_{\langle i,j\rangle} s_i s_j + h\sum_i s_i\,.$$

Damit ergibt sich für die Zustandssumme

$$Z(\beta,h) \;=\; \sum_{\{s_i=\pm1\}} \exp\beta\left(\sum_{\langle i,j\rangle} s_i s_j - h\sum_i s_i\right)\,.$$

$\langle i,j\rangle$ bezeichnet Paare benachbarter Punkte in einem Gitter. Wir betrachten in diesem Abschnitt ein 1-dimensionales Modell, sodass wir für die Zustandssumme auch schreiben können:

$$\begin{aligned} Z(\beta,h) &= \sum_{\{s_i=\pm1\}} \exp\left(\beta\sum_{i=1}^{L}(s_i s_{i+1} - h s_i)\right) \\ &= \sum_{\{s_i=\pm1\}} \prod_{i=1}^{L} \exp\beta\left(s_i s_{i+1} - \frac{h}{2}(s_i + s_{i+1})\right)\,. \end{aligned} \tag{12.1}$$

Die hier angegebene Aufspaltung des h-abhängigen Anteils wird sich später als nützlich erweisen. Der Einfachheit wählen wir periodische Randbedingungen, d.h. $s_{L+1} = s_1$.

Wir stellen nun die Zustandssumme als Spur eines L-fachen Produkts einer Matrix dar. Die Zeilen und Spalten dieser Matrix werden durch die möglichen Zustände auf einem Gitterpunkt indiziert, d.h. in diesem Fall durch $s = \pm 1$. Wir definieren:

$$T_{s\,s'} \;\equiv\; \langle s|T|s'\rangle \;:=\; \exp\beta(s\,s' \,-\, \frac{h}{2}(s+s'))\,, \tag{12.2}$$

bzw.

$$T \;=\; \begin{pmatrix} \mathrm{e}^{\beta(1-h)} & \mathrm{e}^{-\beta} \\ \mathrm{e}^{-\beta} & \mathrm{e}^{\beta(1+h)} \end{pmatrix}\,.$$

Offensichtlich gilt für die Zustandssumme (12.1):

$$Z(\beta,h) \;=\; \sum_{\{s_i=\pm1\}} \prod_{i=1}^{L} T_{s_i\,s_{i+1}} \;=\; \mathrm{tr}\; T^L\,.$$

Die Matrix T bezeichnet man als Transfermatrix des 1-dimensionalen Ising-Modells. T ist symmetrisch, d.h. die Eigenwerte sind reell:

$$\lambda_\pm \;=\; \mathrm{e}^{\beta}\cosh\beta h \pm \sqrt{\mathrm{e}^{2\beta}\cosh^2(\beta h) - 2\sinh(2\beta h)}\,.$$

An dieser Stelle zeigt sich auch der Vorteil der symmetrischen Aufspaltung der Energie in (12.1) bzw. in (12.2), da anderenfalls T nicht symmetrisch wäre. Die Matrix

$$T' \;:=\; \exp\beta(s\,s' \;-\; hs)\,,$$

die sich zunächst als Transfermatrix anbieten würde, hat allerdings dieselben Eigenwerte wie T, da sich die beiden Matrizen nur um eine Ähnlichkeitstransformation unterscheiden. Grundsätzlich hätte sich somit am Transfermatrixformalismus wenig geändert; für viele allgemeine Überlegungen ist es jedoch hilfreich, die Transfermatrix in symmetrischer Form definiert zu haben.

Die Zustandssumme lässt sich nun durch die Eigenwerte der Transfermatrix ausdrücken:

$$Z(\beta,h) \;=\; \lambda_+^L + \lambda_+^L \;=\; \lambda_+^L \left[1 \;+\; \left(\frac{\lambda_-}{\lambda_+}\right)^L\right].$$

Mit Ausnahme des Falls $h = 0$ und $\beta = \infty$ gilt immer $\lambda_-/\lambda_+ < 1$, d.h. im thermodynamischen Grenzfall $L \to \infty$ trägt nur der höchste Eigenwert zur Zustandssumme bei. Für die freie Energie folgt

$$F \;=\; -\,\frac{1}{\beta}\ln Z \;=\; -\,L\frac{1}{\beta}\ln\lambda_+$$

und für die freie Energie „pro Volumen“:

$$f \;=\; -\lim_{L\to\infty}\frac{1}{L}F \;=\; -\,\frac{1}{\beta}\ln\lambda_+\,.$$

Die Berechnung der Freien Energie wurde somit auf die Bestimmung des größten Eigenwertes der Transfermatrix zurückgeführt.

Über die Beziehung

$$T \;=\; \mathrm{e}^{-H}$$

definieren wir nun eine Matrix H. Da die Eigenwerte von T positiv waren, ist die Matrix H wohldefiniert. Mit Hilfe der Formel

$$\exp\left(\alpha\sum_{i=1}^{3} n_i\sigma_i\right) \;=\; \cosh\alpha \;+\; \sum_{i=1}^{3} n_i\sigma_i \;\sinh\alpha\,,$$

wobei σ_i die Pauli-Matrizen sind, lässt sich die Berechnung von H geschlossen ausführen. Trotzdem betrachten wir der Einfachheit halber nur den Fall $h = 0$, für den die Transfermatrix die folgende Form annimmt:

$$T \;=\; \mathrm{e}^{\beta}(1 \;+\; \mathrm{e}^{-2\beta}\sigma_1) \;=\; \frac{\mathrm{e}^{\beta}}{\cosh\beta^*}\mathrm{e}^{\beta^*\sigma_1}\,.$$

Hierbei haben wir die so genannte „duale Temperatur“ β^* über die Relation

$$\tanh\beta^* \;=\; \mathrm{e}^{-2\beta}$$

definiert. Wir erhalten also für H:

$$\begin{aligned} H \;&=\; -\beta + \ln\cosh\beta^* \;-\; \beta^*\sigma_1 \\ &=\; E_0 \;+\; \beta^*(1-\sigma_1)\,. \end{aligned} \tag{12.3}$$

H bezeichnet man als den „Quantenspin-Operator“ zum 1-dimensionalen Ising-Modell.

Der Kontinuumslimes wird für das 1-dimensionale Ising-Modell für $\beta \to \infty$ angenommen, d.h. bei verschwindender Temperatur. Die 2-Punkt-Funktion im thermodynamischen Grenzfall $L \to \infty$,

$$G_k(\beta) \equiv \langle s_i s_{i+k} \rangle = (\tanh \beta)^k = \mathrm{e}^{-kM},$$

definiert die inverse Korrelationslänge $M = -\ln \tanh \beta = 2\beta^*$ in Gittereinheiten. Aus der Grundzustandsenergie in (12.3),

$$\begin{aligned} E_0 &= -\beta + \ln \cosh \beta^* - \beta^* = -\ln 2 \cosh \beta \\ &= -\frac{1}{L} \ln Z = \beta f, \end{aligned}$$

erhält man die freie Energie und somit die thermodynamischen Größen des Systems. Der „renormierte“ Hamilton-Operator $\hat{H}$ mit der Grundzustandsenergie 0 ist

$$\hat{H} = H - E_0 = \beta^*(1 - \sigma_1) = \frac{M}{2}(1 - \sigma_1).$$

Der (in diesem Fall einzige) angeregte Energieeigenwert ist daher M. Dies ist nicht überraschend, denn der erste angeregte Energieeigenwert bestimmt immer die Korrelationslänge der 2-Punkt-Funktion.

12.1.2 Der 1-dimensionale Anti-Ferromagnet

Im Ising-Modell werden Konfigurationen energetisch begünstigt, bei denen benachbarte Spins parallel sind. Es gibt jedoch auch Systeme, bei denen anti-parallele Spins energetisch günstiger sind. Solche Systeme nennt man “Anti-Ferromagneten”. Im Ising-Modell kann man diese Systeme auch durch eine negative Temperatur ($\beta < 0$) realisieren. (Für Systeme mit endlich vielen lokalen Freiheitsgraden sind negative Temperaturen – d.h. Abnahme der Entropie bei Zunahme der Energie – möglich.)

Für das 1-dimensionale anti-ferromagnetische Ising-Modell ohne äußeres Magnetfeld sind die Eigenwerte der Transfermatrix:

$$\lambda_{\max} = 2 \cosh \beta \qquad \lambda_{\min} = -2 \sinh \beta.$$

Dies ist ein Beispiel für ein System, bei dem die Eigenwerte der Transfermatrix nicht positiv sind. Die Freie Energie, durch den maximalen Eigenwert bestimmt, ist davon zwar nicht betroffen, allerdings ist die Korrelationsfunktion,

$$G_k \equiv \langle s_i s_{i+k} \rangle = (-1)^k (\tanh \beta)^k,$$

alternierend im Vorzeichen.

In diesem Fall kann man die Matrix H nicht mehr als hermitesche Matrix angeben, da T keine positive Matrix ist. Das langreichweitige Verhalten des Ising-Modells (hier im Grenzfall $\beta \to \infty$) lässt sich jedoch auch durch die Matrix T^2 beschreiben. Dies ist eine positive Matrix, und wir können definieren

$$H = -\frac{1}{2} \ln T^2.$$

H ist nun identisch mit dem Hamilton-Operator des 1-dim. gewöhnlichen Ising-Modells.

12.1.3 Der allgemeine Transfermatrixformalismus

Der Transfermatrixformalismus lässt sich allgemein auf ein Gitter der Form $\mathcal{G} \times \mathbf{Z}$ anwenden (vgl. Abb. 12.1), wobei $\mathcal{G}$ ein beliebiger Graph ist (d.h. eine Menge von Punkten und eine Menge von Linien, die zwei Punkte als benachbart kennzeichnen). Die durch $\mathbf{Z}$ durchnummerierte Koordinate heißt manchmal „Zeit"-Richtung, der Teil $\mathcal{G}$ „räumliches" Gitter. Sei $\hat{m}$ eine Durchnummerierung der Punkte von $\mathcal{G}$, so kennzeichnet $(\hat{m}, i)$ $(i \in \mathbf{Z})$ einen Punkt im gesamten Gitter. Bei einem d-dimensionalen hyperkubischen Gitter ist $\mathcal{G}$ ein $d-1$-dimensionales hyperkubischen Gitter. Auf diesen Fall wollen wir uns im Folgenden beschränken, obwohl die Verallgemeinerung auf beliebige Graphen keine besonderen Schwierigkeiten bereitet. Außerdem seien die lokalen Freiheitsgrade zunächst auf den Gitterpunkten definiert. Den Fall einer Eichtheorie, bei der die Freiheitsgrade den Gitterlinien zugeordnet sind, werden wir später gesondert untersuchen.

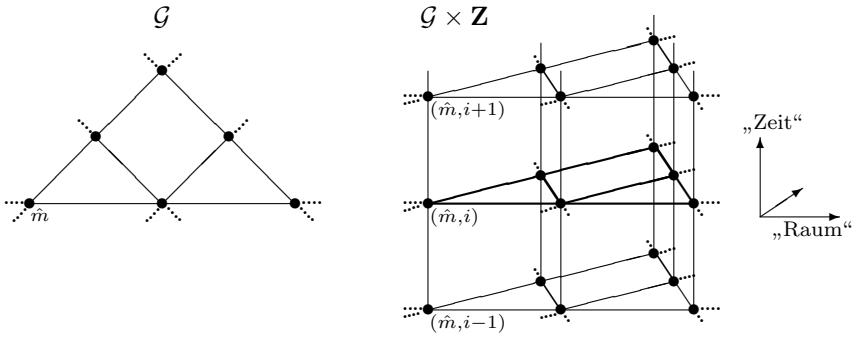


Abbildung 12.1: Gitter des Typs $\mathcal{G} \times \mathbf{Z}$ mit einer ausgezeichneten „Zeit"–Richtung.

Wir konstruieren zunächst den „Hilbert-Raum" der Zustände auf dem räumlichen Gitter. V sei der Raum der möglichen lokalen Zustände an einem Punkt. Beispielsweise ist beim Ising-Modell $V = \{+1, -1\}$ und bei einem skalaren reellen Feld $V = \mathbb{R}$. Dann definieren wir $\mathcal{H}_{\hat{m}}$ als den komplexen Vektorraum über den lokalen Zuständen am Punkte $\hat{m}$, d.h. der Vektorraum der komplexwertigen Funktionen über den lokalen Zuständen:

$$\mathcal{H}_{\hat{m}} \;=\; \{\psi | \psi : V \to \mathbb{C}\}\,.$$

Die lokalen Zustände, d.h. die Elemente $\varphi \in V$ zeichnen eine mögliche Basis von $\mathcal{H}_{\hat{m}}$ aus:

$$\{|\varphi\rangle \;\equiv\; \psi_\varphi\} \qquad \text{mit} \;\; \psi_\varphi(\varphi') \;=\; \delta_{\varphi\,\varphi'}\,.$$

Als Elemente des Funktionenraums über V besteht die Basis also aus den charakteristischen Funktionen zu den Zuständen. Auf dieser Basis definieren wir ein Skalarprodukt:

$$\langle\varphi|\varphi'\rangle \;=\; \delta_{\varphi\,\varphi'}\,.$$

Ist V eine unendliche Menge, dann müssen an ψ ähnliche Einschränkungen wie in der Quantenmechanik hinsichtlich der Quadratsummierbarkeit bzw. der Quadratintegrabilität gestellt

werden. Beim Ising-Modell handelt es sich lokal um Spin-Zustände $s(\hat{m}) = \pm 1$, d.h. $\mathcal{H}_{\hat{m}}$ ist ein (komplex) 2-dimensionaler Vektorraum. Für ein reelles Feld $\varphi(\hat{m}) \in \mathbb{R}$ ist $\mathcal{H}_{\hat{m}}$ der Vektorraum der quadratintegrablen Funktionen über der reellen Achse. Wir definieren zu einem räumlichen Gitter einen Zustandsraum

$$\mathcal{H} \;=\; \bigotimes_{\hat{m}\in\mathcal{G}} \mathcal{H}_{\hat{m}}\,.$$

Dies ist gleichzeitig der Raum aller (komplexwertigen) Funktionen über den Zuständen auf $\mathcal{G}$. Das Skalarprodukt in $\mathcal{H}$ ist das Produkt der Skalarprodukte in $\mathcal{H}_{\hat{m}}$. Eine mögliche Basis in $\mathcal{H}$ sind die Zustände

$$|\{\varphi\}\rangle \;=\; \bigotimes_{\hat{m}\in\mathcal{G}} |\varphi(\hat{m})\rangle\,.$$

Im diskreten Fall sind dies die charakteristischen Funktionen zu einer Konfiguration $\{\varphi(\hat{m})\}$. In Anlehnung an die „Ortsraumbasis“ in der Quantenmechanik bezeichnet man diese Basis die „Konfigurationsraumbasis“.

Entsprechend der Aufspaltung des Gitters in ein räumliches Hypergitter und eine Zeitrichtung spalten wir auch die Energie auf. Dabei wird vorausgesetzt, dass es nur Wechselwirkungsterme zwischen nächsten Nachbarn gibt:

$$\begin{aligned}\sum_{\langle(\hat{m},i),(\hat{n},j)\rangle} E(\varphi(\hat{m},i),\varphi(\hat{n},j)) \;&=\\ =\; \sum_i\sum_{\hat{m}} E(\varphi(\hat{m},i),\varphi(\hat{m},i+1)) \;&+\; \sum_i \sum_{\langle\hat{m},\hat{n}\rangle} E(\varphi(\hat{m},i),\varphi(\hat{n},i))\,.\end{aligned}$$

(Die Notation $\langle\ ,\ \rangle$ bezeichnet nächsten Nachbarn auf einem Gitter.)

Wir definieren nun die Transfermatrix T als linearen Operator auf $\mathcal{H}$ durch seine Matrixelemente in der Konfigurationsraumbasis

$$T_{\varphi,\psi} \;\equiv\; \langle\varphi|T|\psi\rangle \;=\; \mathrm{e}^{-\beta E[\varphi,\psi]}\,,$$

mit

$$E[\varphi,\psi] \;=\; \sum_{\hat{m}} E(\varphi(\hat{m}),\psi(\hat{m})) + \frac{1}{2}\sum_{\langle\hat{m},\hat{n}\rangle}\left[E(\varphi(\hat{m}),\varphi(\hat{n})) + E(\psi(\hat{m}),\psi(\hat{n}))\right].$$

Die so definierte Transfermatrix ist offensichtlich reell und symmetrisch, d.h. ihre Eigenwerte sind reell.

Da außerdem $T_{\varphi,\psi} > 0$, lässt sich zeigen, dass der höchste Eigenwert von T positiv und nicht entartet ist. Dies folgt aus dem so genannten Frobenius-Perron-Theorem über positive Matrizen (vgl. Kapitel 14). Wie wir am Beispiel des 1-dimensionalen Anti-Ferromagneten gesehen haben, sind jedoch nicht alle Eigenwerte von T notwendigerweise positiv. Die Eigenwerte von T^2 sind jedoch immer positiv. Wir definieren nun den Operator H durch

$$T^2 \;=\; \exp(-2H) \qquad\qquad H \text{ selbstadjungiert}\,.$$

H bezeichnet man als den Quantenoperator der Gittertheorie. Um jedoch den Hamilton-Operator der Gittertheorie zu erhalten, muss noch ein Kontinuumslimes in der Zeitrichtung ausgeführt werden. Dazu müssen die Kopplungskonstanten (bzw. β) in räumlicher und zeitlicher Richtung unterschiedlich gewählt werden. In zeitlicher Richtung muss dann ein Limes

der Kopplungskonstanten ausgeführt werden, so dass die Korrelationslängen in Zeitrichtung langreichweitig werden. Diesem Problem wollen wir uns hier jedoch nicht weiter widmen.

Für die Zustandssumme auf einem periodischen Raum-Zeit-Gitter, welches in Zeit-Richtung die Länge L hat, gilt offensichtlich

$$Z \;=\; \mathrm{tr}\; T^L \;=\; \sum_n \lambda_n^L \,, \tag{12.4}$$

wobei λ_n die Eigenwerte der Transfermatrix sind. Zum Beweis muss man sich wiederum davon überzeugen, dass die Definition der Spur bzw. der Matrixmultiplikation für T zur Summation über alle Konfigurationen führt.

In der Konfigurationsraumbasis können wir in $\mathcal{H}$ einen ausgezeichneten Satz von Operatoren definieren: Operatoren $\Phi(\hat{m})$, die in dieser Basis diagonal sind, und Operatoren $P(\hat{m})$, die Basisvektoren in andere Basisvektoren überführen, beispielsweise:

$$\begin{aligned} \Phi(\hat{m})|\varphi\rangle &= \varphi(\hat{m})|\varphi\rangle \\ P_\delta(\hat{m})|\ldots,\varphi(\hat{m}),\ldots\rangle &= |\ldots,\varphi(\hat{m})+\delta,\ldots\rangle\,. \end{aligned}$$

In diesem Fall entspricht P_δ dem exponenzierten Impulsoperator der Quantenmechanik, der Ortseigenzustände um δ translatiert. Die Notation ist im Allgemeinen nur symbolisch: P verändert den Zustand zu einer Konfiguration in bestimmter Weise. Das kann bei reellen Feldern die Addition einer Konstanten sein, das kann aber auch (z.B. bei Gruppen) in der Multiplikation mit einem generierenden Gruppenelement bestehen. Ebenso ist $\Phi(\hat{m})$ ein Operator für die Koordinaten einer Konfiguration, nicht für die Konfiguration selber. Man kann nun jeden Operator auf $\mathcal{H}$, insbesondere auch die Transfermatrix, durch die Operatoren $\Phi(\hat{m})$ und $P(\hat{m})$ ausdrücken.

Im 1-dimensionalen Ising-Modell sind beispielsweise in der Basis

$$|+\rangle \;\simeq\; \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad\qquad |-\rangle \;\simeq\; \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{12.5}$$

die Matrizen Φ und P durch

$$\Phi \;=\; \sigma_3 \;=\; \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \qquad \text{und} \qquad P \;=\; \sigma_1 \;=\; \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{12.6}$$

gegeben.

Auch Erwartungswerte lassen sich im Operatorformalismus ausdrücken, beispielsweise

$$\begin{aligned} \langle\varphi(\hat{m},i)\varphi(\hat{n},j)\rangle &= \lim_{L\to\infty} \frac{\mathrm{tr}\;\Phi(\hat{n})\;T^{(j-i)}\;\Phi(\hat{m})\;T^{L-(j-i)}}{\mathrm{tr}\;T^L} && (12.7)\\ &= \langle 0|\Phi(\hat{n})\;T^{(j-i)}\;\Phi(\hat{m})|0\rangle\,. && (12.8) \end{aligned}$$

Im zweiten Schritt haben wir ausgenutzt, dass im Grenzfall $L\to\infty$ von allen Zuständen nur der Grundzustand $|0\rangle$ von H (d.h. der Zustand zum höchsten Eigenwert der Transfermatrix) beiträgt.

12.2 Transfermatrix zur Quantenmechanik

Wir betrachten nun als Beispiel für den angegebenen Formalismus ein 1-dimensionales Gitter, dessen Punkten als lokale Freiheitsgrade Felder $\boldsymbol{x}_i \in \mathbb{R}^3$ zugeordnet sind. Die Energie einer Feldkonfiguration sei

$$\begin{aligned} E[\boldsymbol{x}_i] &= \sum_i \frac{D}{2}(\boldsymbol{x}_{i+1} - \boldsymbol{x}_i)^2 + \lambda \sum_i V(\boldsymbol{x}_i) \\ &= -\sum_i D(\boldsymbol{x}_{i+1}\boldsymbol{x}_i) + \sum_i (D\boldsymbol{x}_i^2 + \lambda V(\boldsymbol{x}_i)) , \end{aligned}$$

wobei V ein beliebiges Potential und λ eine Kopplungskonstante ist. Die Wechselwirkung zwischen nächsten Nachbarn entspricht einer harmonischen Kopplung, wie sie beispielsweise bei der schwingenden Saite vorliegt. D ist die „Federkonstante“ der Kopplung. Die Randbedingungen kann man zunächst wieder als periodisch annehmen, sie spielen aber für die Transfermatrix keine Rolle.

Ein raumartiges Gitter besteht nur aus einem Punkt, der Hilbert-Raum der raumartigen Freiheitsgrade ist somit der Raum der quadratintegrablen Funktionen im $\mathbb{R}^3$. Die Transfermatrix zu dieser linearen Kette ist:

$$\begin{aligned} \langle \boldsymbol{x}|T|\boldsymbol{y}\rangle &= T(\boldsymbol{x},\boldsymbol{y}) \\ &= \exp -\beta \left(\frac{D}{2}(\boldsymbol{x}-\boldsymbol{y})^2 + \frac{\lambda}{2}(V(\boldsymbol{x}) + V(\boldsymbol{y})) \right) . \end{aligned}$$

$T(\boldsymbol{x},\boldsymbol{y})$ ist ein Integralkern auf dem Hilbert-Raum. Die diagonalen Operatoren in der Konfigurationsraumbasis sind die drei Komponenten des „Orts“-Operators:

$$Q_i|\boldsymbol{x}\rangle = x_i|\boldsymbol{x}\rangle .$$

Die Verschiebeoperatoren

$$P_i(\alpha)|\boldsymbol{x}\rangle = |\boldsymbol{x} + \alpha \boldsymbol{e}_i\rangle$$

besitzen als Generator den „Impuls“-Operator P_i, der mit Q_i die Vertauschungsrelationen

$$[P_i, Q_j] = \mathrm{i}\delta_{ij}$$

erfüllt. Aus der Quantenmechanik ist folgende Relation bekannt (vgl. auch Gl. (5.7)):

$$\langle \boldsymbol{x}| \exp\left(-\frac{1}{2\beta D}\boldsymbol{P}^2 \right) |\boldsymbol{y}\rangle = \exp\left(-\frac{D\beta}{2}(\boldsymbol{x}-\boldsymbol{y})^2 \right) .$$

Damit können wir die Transfermatrix durch die Operatoren $\boldsymbol{P}$ und $\boldsymbol{Q}$ ausdrücken:

$$\langle \boldsymbol{x}|T|\boldsymbol{y}\rangle = \exp\left(-\beta\frac{\lambda}{2}(V(\boldsymbol{Q})) \right) \exp\left(-\frac{1}{2D\beta}\boldsymbol{P}^2 \right) \exp\left(-\beta\frac{\lambda}{2}(V(\boldsymbol{Q})) \right) .$$

Langreichweitige Korrelationen erhält man in diesem Fall für $\lambda \to 0$ und $D \to \infty$, die Korrelationslänge hängt jedoch von der Form von $V(\boldsymbol{Q})$ ab. β ist in dieser Theorie ein freier Parameter

und kann durch Reskalierung von $\boldsymbol{x}$ absorbiert werden. Der effektive Hamilton-Operator obiger Gittertheorie hat damit die Form:

$$H \;=\; \frac{1}{2}\boldsymbol{P}^2 \;+\; \lambda' V(\boldsymbol{Q})\,.$$

Die neue Kopplungskonstante λ' ist eine Funktion von β, D und λ und ist so zu adjustieren, dass die inverse Korrelationslänge $M(\beta,\lambda)$ gleich der Differenz zwischen den beiden niedrigsten Energieeigenwerten von H ist.

12.3 Skalare, reelle Feldtheorie

Wir betrachten nun eine freie, skalare Feldtheorie auf einem d-dimensionalen Gitter. Die lokalen Freiheitsgrade sind reelle Felder $\varphi(\boldsymbol{m}) \in \mathbb{R}$. Die Energie ist die gitterregularisierte Klein-Gordon-Wirkung (8.1):

$$E[\varphi] \;=\; \frac{1}{2}\sum_{\langle m,n\rangle}(\varphi(\boldsymbol{m})-\varphi(\boldsymbol{n}))^2 \;+\; \frac{1}{2}M^2\sum_{m}\varphi(\boldsymbol{m})^2\,.$$

Ein raumartiges Gitter besteht nun aus einem $(d-1)$-dimensionalen Untergitter, dessen Punkte wir durch $\{\hat{m}\}$ indizieren. Der Hilbert-Raum an jedem Punkt ist der Raum der quadratintegrablen Funktionen $\mathcal{L}_2 \simeq \{\psi(\varphi)\}$, der Hilbert-Raum $\mathcal{H}$ für das raumartige Gitter ist entsprechend das Tensorprodukt von $\mathcal{L}_2$ über alle Punkte des $(d-1)$-dimensionalen raumartigen Gitters. Jeder Feldkonfiguration $\{\varphi(\hat{m})\}$ auf dem raumartigen Gitter entspricht ein Basisvektor $|\{\varphi(\hat{m})\}\rangle$ in der Konfigurationsraumbasis von $\mathcal{H}$. Die Transfermatrix ist

$$\langle\{\varphi(\hat{m})\}|T|\{\varphi'(\hat{m})\}\rangle \;=\; \mathrm{e}^{-E[\varphi,\varphi']}\,,$$

mit

$$\begin{aligned} E[\varphi,\varphi'] \;&=\; \frac{1}{2}\sum_{\hat{m}}(\varphi(\hat{m})-\varphi'(\hat{m}))^2 \;+\; \frac{1}{4}M^2\sum_{m}[\varphi(\hat{m})^2+\varphi'(\hat{m})]^2 \\ &\quad+\; \frac{1}{4}\sum_{\langle\hat{m},\hat{n}\rangle}\left[(\varphi(\hat{m})-\varphi(\hat{n}))^2 \;+\; (\varphi'(\hat{m})-\varphi'(\hat{n}))^2\right]. \end{aligned}$$

Die diagonalen Operatoren sind die Feldoperatoren:

$$\Phi(\hat{m})|\{\varphi\}\rangle \;=\; \varphi(\hat{m})|\{\varphi\}\rangle\;.$$

Die Generatoren der Verschiebungen sind die konjugierten Operatoren:

$$\pi(\hat{m}) \;=\; -\,\mathrm{i}\frac{\delta}{\delta\varphi(\hat{m})} \qquad\qquad [\Phi(\hat{m}),\pi(\hat{n})] \;=\; -\,\mathrm{i}\delta(\hat{m},\hat{n})\,.$$

Ganz analog wie im letzten Abschnitt kann man die Transfermatrix nun durch $\Phi(\hat{m})$ und $\pi(\hat{m})$ ausdrücken und erhält nach einem geeigneten Kontinuumlimes ($M \to 0$) als Hamilton-Operator:

$$H \;=\; \frac{1}{2}\sum_{\hat{m}}\pi(\hat{m})^2 \;+\; \frac{1}{2}\sum_{\langle\hat{m},\hat{n}\rangle}(\Phi(\hat{m})-\Phi(\hat{n}))^2 \;+\; \frac{1}{2}M^2\sum_{\hat{m}}\Phi(\hat{m})^2\,.$$

Dies ist der diskretisierte Hamilton-Operator der Klein-Gordon-Theorie.

12.4 Transfermatrix für Gittereichtheorien

Bei den obigen Überlegungen wurde der Einfachheit halber angenommen, dass die lokalen Freiheitsgrade den Gitterpunkten zugeordnet sind, und dass es nur Wechselwirkungsterme zwischen direkt benachbarten Gitterpunkten gibt. Der Formalismus lässt sich auch auf Wechselwirkungen zwischen weiter entfernten Gitterpunkten übertragen sowie auf lokale Freiheitsgrade, die nicht den Gitterpunkten zugeordnet sind. Dies erfordert im Allgemeinen jedoch einen erheblichen Aufwand – der Hilbert-Raum der „räumlichen“ Konfigurationen muss dazu wesentlich erweitert werden. Im Fall der Gittereichtheorien vereinfacht sich der Formalismus jedoch wieder, sodass nur dieser Fall hier behandelt werden soll.

Der erste Schritt zum Transfermatrix-Formalismus einer Gittereichtheorie besteht in der Wahl einer speziellen Eichung. Wir wählen die so genannte temporale Eichung. Im Kontinuum bedeutet dies, dass alle zeitartigen Eichfelder $A_0(x)$ gleich Null gesetzt werden; für die Paralleltransporter einer Gittereichtheorie bedeutet diese Eichung:

$$U(\boldsymbol{m}, 0) \;=\; \mathbb{1}\,.$$

Die Gruppenvariablen auf allen „zeitartigen“ Gitterlinien werden also zur Identität geeicht. Das Problem, inwieweit bei gegebenen Randbedingungen eine solche Eichung global gewählt werden kann, soll hier nicht weiter diskutiert werden. Bei offenen Randbedingungen ist dies immer möglich. Bei periodischen Randbedingungen gelingt dies im Allgemeinen nicht vollständig.

Die dynamischen Freiheitsgrade sind also den raumartigen Gitterlinien zugeordnet. Solche raumartigen Gitterlinien kennzeichnen wir im Folgenden durch den „Hut“. So ist $U(\hat{l}, i)$ das Gruppenelement auf der raumartigen Gitterlinie $\hat{l}$ im i-ten räumlichen Hypergitter. Entsprechend sei

$$U_{\hat{p},i} \;=\; \prod_{\hat{l}\in\partial\hat{p}} U(\hat{l}, i)$$

das Produkt der vier Gruppenvariablen an den räumlichen Gitterlinien um eine räumliche Plakette $(\hat{p}, i)$. Für eine zeitartige Plakette sind zwei der Gruppenelemente gleich der Identität, und der Beitrag reduziert sich auf ein Produkt der Gruppenelemente auf den beiden gegenüberliegenden räumlichen Gitterlinien:

$$U_p \;\simeq\; U(\hat{l}, i+1)^+\, U(\hat{l}, i)\,.$$

Mit dieser Notation können wir nun die Wirkung der Gittereichtheorie in der temporalen Eichung wiederum in einen räumlichen und einen zeitartigen Anteil aufspalten:

$$\begin{aligned} S \;=\; & \frac{2}{g^2 N} \sum_{(\hat{p},i)} \mathrm{tr}\, \left(1 - \frac{1}{2}(U_{\hat{p},i} + U_{\hat{p},i}^+)\right) \\ & + \frac{2}{g^2 N} \sum_{(\hat{l},i)} \mathrm{tr}\, \left(1 - \frac{1}{2}(U(\hat{l}, i+1)^+ U(\hat{l}, i) + U(\hat{l}, i)^+ U(\hat{l}, i+1)\right)\,. \end{aligned}$$

Der lokale Hilbert-Raum ist nun den Linien $\{\hat{l}\}$ des raumartigen Gitters zugeordnet und besteht aus den integrablen Funktionen auf der Gruppenmannigfaltigkeit. Bei endlichen Gruppen ist $\mathcal{H}_{\hat{l}}$ der Raum der komplexen Funktionen über der Gruppe. Die lokale Konfigurationsraumbasis ist

$\{|U\rangle\}$, d.h. jedes Gruppenelement hat seinen eigenen Basisvektor. Der gesamte Hilbert-Raum zu einem raumartigen Gitter ist

$$\mathcal{H} \;=\; \bigotimes_{\hat{l}} \mathcal{H}_{\hat{l}}\,,$$

wobei das Produkt über Gitterlinien des raumartigen Untergitters zu nehmen ist. Die Transfermatrix wird wiederum durch ihre Matrixelemente in der Konfigurationsraumbasis definiert:

$$\langle U|T|V\rangle \;=\; \mathrm{e}^{-E[U,V]}$$

mit

$$\begin{aligned} E[U,V] \;&=\; \frac{2}{g^2 N}\sum_{(\hat{p})} \mathrm{tr}\,\left(1-\frac{1}{4}(U_{\hat{p}}+U_{\hat{p}}^{+}+V_{\hat{p}}+V_{\hat{p}}^{+})\right) \\ &\quad+\frac{2}{g^2 N}\sum_{(\hat{l})} \mathrm{tr}\,\left(1-\frac{1}{2}(U(\hat{l})^{+}V(\hat{l})\,+\,V(\hat{l})^{+}U(\hat{l}))\right)\,. \end{aligned}$$

Der nächste Schritt besteht in der Berechnung des Hamilton-Operators. Neben den in der Konfigurationsraumbasis diagonalen Operatoren $\hat{U}(\hat{l})$ gibt es noch Erzeugende für eine Verschiebung der Basis. Für diese Operatoren kann man bei einer kontinuierlichen Gruppe beispielsweise die Ableitungen $\delta/\delta A^a(\boldsymbol{l})$ wählen.

Die explizite Rechnung ist sehr aufwendig, aber im Rahmen eines Kontinuumslimes für die Zeitrichtung lässt sich zeigen, dass der Hamilton-Operator aus der Transfermatrix zum Hamilton-Operator einer Quanteneichtheorie auf einem räumlichen Gitter wird. Am Ende dieses Kapitels ist eine Tabelle angeführt, in der die wichtigsten Resultate zum Transfermatrixformalismus nochmals zusammengefasst werden.

12.5 Zweidimensionale reine Gittereichtheorien

Als eine Anwendung des Transfermatrix-Formalismus zeigen wir, dass 2-dimensionale reine Gittereichtheorien in demselben Sinn integrabel sind wie 1-dimensionale Spin-Ketten. Der Grund ist die Eichsymmetrie, die es im Prinzip ermöglicht, das 2-dimensionale Modell auf ein Produkt von (unabhängigen) 1-dimensionalen Modellen zurückzuführen. Wir werden uns im Folgenden auf die thermodynamischen Eigenschaften solcher Systeme konzentrieren und insbesondere Feinheiten, die mit Randbedingungen in Zusammenhang stehen, außer acht lassen.

Der erste Schritt zur Lösung einer 2-dimensionalen reinen Gittereichtheorie besteht in der Wahl einer Eichung, für die alle Gruppenelemente auf zeitartigen Linien zur Identität werden. Bei offenen Randbedingungen überzeugt man sich leicht, dass dies sukzessive möglich ist. Bei periodischen Randbedingungen kann es allerdings vorkommen, dass ein Gruppenelement auf einer zeitartigen Gitterlinie verbleibt, welches nicht durch eine Eichung zur Identität gebracht werden kann. Solche Randeffekte wollen wir hier vernachlässigen.

Die Energie einer Konfiguration in der so genannten “temporalen Eichung” ist

$$E[\{g\}] \;=\; -\,\frac{1}{2N}\sum_{\langle i_1,i_2\rangle,\langle i_3,i_4\rangle} \mathrm{tr}\,\left(g_{\langle i_1,i_2\rangle}g^{-1}_{\langle i_3,i_4\rangle}+\mathrm{h.c.}-2\right).$$

(Ihr Wert hat sich durch die Wahl der Eichung natürlich nicht geändert.) Die raumartigen Gitterlinien lassen sich durch die Gitterpunkte durchnummerieren. Bezeichnet $[x,t]$ im Folgenden die Linie zwischen Gitterpunkt (x,t) und $(x+1,t)$, so schreibt sich die Energie

$$E[\{g\}] \;=\; -\frac{1}{2N}\sum_{x,t}\mathrm{tr}\;(g_{[x,t]}g^{-1}_{[x,t+1]}+\mathrm{h.c.}-2)\,.$$

Man erkennt, dass die Gruppenelemente auf Gitterlinien zu verschiedenen Werten von x nicht mehr in Wechselwirkung stehen. Für die Zustandssumme bedeutet dies

$$\begin{aligned}
Z &= \int\prod_l dg_l\;\mathrm{e}^{\frac{\beta}{2N}\sum_{\langle l_1l_2l_3l_4\rangle}\mathrm{tr}\;(g_{l_1}g_{l_2}g_{l_3}g_{l_4}+\mathrm{h.c.}-2)}\\
&= \int\prod_{x,t} dg_{[x,t]}\;\mathrm{e}^{\frac{\beta}{2N}\sum_{x,t}\mathrm{tr}\;(g_{[x,t]}g^{-1}_{[x,t+1]}+\mathrm{h.c.}-2)}\\
&= \prod_x\left(\int\prod_t dg_{[x,t]}\;\mathrm{e}^{\frac{\beta}{2N}\sum_t\mathrm{tr}\;(g_{[x,t]}g^{-1}_{[x,t+1]}+\mathrm{h.c.}-2)}\right)\\
&= \left(\int\prod_t dg_{[t]}\;\mathrm{e}^{\frac{\beta}{2N}\sum_t\mathrm{tr}\;(g_{[t]}g^{-1}_{[t+1]}+\mathrm{h.c.}-2)}\right)^L.
\end{aligned}$$

Dabei ist L die Größe des Gitters in Raum-Richtung. Beim letzten Schritt wurde ausgenutzt, dass sich das Produkt von Zustandssummen zu L nicht miteinander wechselwirkenden gleichen Systemen als entsprechende Potenz einer Zustandssumme schreiben lässt.

Das Problem ist nun reduziert auf die Berechnung einer Zustandssumme einer linearen Kette. Dazu kann man wieder die Transfermatrix definieren:

$$\langle g|T|g'\rangle \;=\; \mathrm{e}^{\frac{\beta}{2N}\mathrm{tr}\;(g'g^{-1}+\mathrm{h.c.}-2)}\,.$$

Im Allgemeinen gibt es keine geschlossene Formel für die Eigenwerte dieser Matrix. Für abelsche Gruppen jedoch lassen sich die Eigenwerte der Transfermatrix als Summation über die Charaktere der Gruppe darstellen. Dies geschieht analog zum Fall von translationsinvarianten Transfermatrizen, welche sich durch die Charaktere der Translationsgruppe e^{ikx} ausdrücken lassen.

Zur Bestimmung der freien Energie hat man ein Integral der Form

$$\lambda \;=\; \int dg\;\mathrm{e}^{\frac{\beta}{2N}\mathrm{tr}\;(g+g^+-2)}$$

zu berechnen. Der Grenzfall $N\to\infty$ wurde von D.J. Gross und E. Witten [38] untersucht (siehe auch [15]). Sie fanden in diesem Fall einen Phasenübergang dritter Ordnung. Dies ist ein Beispiel für ein eindimensionales Gittermodell mit einem Phasenübergang. Die bekannten "No-Go"-Theoreme gelten hier nicht, weil es in diesem Grenzfall an jedem Gitterpunkt unendlich viele Freiheitsgrade gibt.

Tabelle zum Transfermatrixformalismus

	1-dim. Ising	1-dim. $\mathbf{R}^3$ (QM)	d-dim. Skalar	d-dim. Eichtheorie
lokale Zustände	$s \in \{+1, -1\}$	$x \in \mathbf{R}^3$	$\varphi_{\hat{m}} \in V$	$U_{\hat{l}} \in G$
$\mathcal{H}_{\hat{m}}$	$\mathbf{C}_2$	$\mathcal{L}_2(\mathbf{R}^3, \mathrm{d}^3x)$	$\mathcal{L}_2(V, \mathrm{d}v)$	$\mathcal{L}_2(G, \mathrm{d}\mu)$
$\mathcal{H}$ $\simeq \bigotimes_{\hat{m}} \mathcal{H}_{\hat{m}}$	$\mathbf{C}_2 \simeq$ $\{(\psi_+, \psi_-)\}$	$\mathcal{L}_2(\mathbf{R}^3, \mathrm{d}^3x)$ $\simeq \{\psi(x)\}$	$\bigotimes_{\hat{m}} \mathcal{L}_2(V_{\hat{m}}, \mathrm{d}v)$ $\simeq \{\Psi[\{\varphi_{\hat{m}}\}]\}$	$\bigotimes_{\hat{l}} \mathcal{L}_2(G_{\hat{l}}, \mathrm{d}\mu)$ $\simeq \{\Psi[\{U_{\hat{l}}\}]\}$
Konfig. Basis	$\lvert s\rangle \quad s = \pm 1$	$\lvert x\rangle$	$\lvert\{\varphi_{\hat{m}}\}\rangle$	$\lvert\{U_{\hat{l}}\}\rangle$
Diagon. Operator	$S\lvert s\rangle = s\lvert s\rangle$ $S \simeq \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$	$Q_i\lvert x\rangle = x_i\lvert x\rangle$	$\Phi_{\hat{n}}\lvert\{\varphi_{\hat{m}}\}\rangle = \varphi_{\hat{n}}\lvert\{\varphi_{\hat{m}}\}\rangle$	$\hat{U}_{\hat{k}}\lvert\{U_{\hat{l}}\}\rangle = U_{\hat{k}}\lvert\{U_{\hat{l}}\}\rangle$
„Shift“ Operator	$P\lvert s\rangle = \lvert -s\rangle$ $P \simeq \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$	$\mathrm{e}^{iaP}\lvert x\rangle = \lvert x + a\rangle$ $P_i \simeq -i\dfrac{\partial}{\partial x_i}$	$\mathrm{e}^{i\sum_{\hat{m}} \phi_{\hat{m}}\pi_{\hat{m}}}\lvert\{\varphi_{\hat{m}}\}\rangle = \lvert\{\varphi_{\hat{m}} + \phi_{\hat{m}}\}\rangle$ $\pi_{\hat{m}} \simeq -i\dfrac{\delta}{\delta\varphi_{\hat{m}}}$	$P_{g_{\hat{l}}}\lvert\{U_{\hat{l}}\}\rangle = \lvert\{g_{\hat{l}}U_{\hat{l}}\}\rangle$
[,]	$SP = -PS$	$Q_i\mathrm{e}^{iaP} = \mathrm{e}^{iaP}(Q_i + a_i)$ $[Q_i, P_j] = -i\delta_{ij}$	$[\Phi_{\hat{n}}, \pi_{\hat{m}}] = -i\delta_{\hat{m},\hat{n}}$	$\hat{U}_{\hat{k}}P_{g_{\hat{l}}} = g_{\hat{l}}P_{g_{\hat{l}}}\hat{U}_{\hat{k}}$

(ˆ bezieht sich auf Größen im raumartigen Gitter.)

Energie:

1-dim. Ising $\quad E = \sum_i s_i s_{i+1}$

QM im $\mathbf{R}^3$ $\quad E = \sum_i \frac{1}{2}(\boldsymbol{x}_i - \boldsymbol{x}_{i+1})^2 + V(\boldsymbol{x}_i)$

d-dim. Skalar $\quad E = \frac{1}{2}\sum_{\langle m,m'\rangle} (\varphi_m - \varphi_{m'})^2$

Eichth. $\quad E = -\frac{2}{g^2N}\mathrm{tr}\ \sum_p \left(\prod_{l\in\partial p} U_l + \mathrm{h.c.} - 2\right)$

$= -\frac{2}{g^2N}\mathrm{tr}\ \sum_i \sum_{\hat{l}} \left(U_{(\hat{l},i)}U^+_{(\hat{l},i+1)} + \mathrm{h.c.} - 2\right) - \frac{2}{g^2N}\mathrm{tr}\ \sum_i \sum_{\hat{p}} \left(\prod_{\hat{l}\in\partial\hat{p}} U_{(\hat{l},i)} + \mathrm{h.c.} - 2\right)$

Transfermatrixelemente:

$$\begin{array}{lrcl}
\text{1-dim. Ising} & \langle s|T|s'\rangle & = & \exp(\beta s s') \\
\text{QM im } \mathbf{R}^3 & \langle \boldsymbol{x}|T|\boldsymbol{y}\rangle & = & \exp\left(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{y})^2 - \frac{1}{2}(V(\boldsymbol{x})+V(\boldsymbol{y}))\right) \\
d\text{-dim. Skalar} & \langle\{\varphi_{\hat{m}}\}|T|\{\varphi'_{\hat{m}}\}\rangle & = & \exp\left(-\frac{1}{2}\sum_{\hat{m}}(\varphi_{\hat{m}}-\varphi'_{\hat{m}})^2 - \frac{1}{4}\sum_{\langle\hat{m},\hat{m}'\rangle}[(\varphi_{\hat{m}}-\varphi_{\hat{m}'})^2+(\varphi'_{\hat{m}}-\varphi'_{\hat{m}'})^2]\right) \\
\text{Eichth.} & \langle\{U_{\hat{l}}\}|T|\{U'_{\hat{l}}\}\rangle & = & \exp\left(-\frac{2}{g^2N}\mathrm{tr}\sum_{\hat{l}}\left(U_{\hat{l}}U'^{+}_{\hat{l}}+\text{h.c.}-2\right)\right) \\
 & & & \times\exp\left(-\frac{1}{g^2N}\mathrm{tr}\sum_{\hat{p}}\left(\prod_{\hat{l}\in\partial\hat{p}}(U_{\hat{l}}+U'_{\hat{l}}+\text{h.c.}-4)\right)\right)
\end{array}$$

Transfermatrix ausgedrückt durch „S“- und „P“-Operatoren:

$$\begin{array}{lrcl}
\text{1-dim. Ising} & T & = & \dfrac{\mathrm{e}^{\beta}}{\cosh\beta^*}\exp(\beta^*\sigma_1) \qquad \tanh\beta^* = \mathrm{e}^{-2\beta} \\
\text{QM im } \mathrm{R}^3 & T & \propto & \exp\left(-\frac{1}{2}V(\boldsymbol{Q})\right)\ \exp\left(-\frac{1}{2}\boldsymbol{P}^2\right)\ \exp\left(-\frac{1}{2}V(\boldsymbol{Q})\right) \\
d\text{-dim. Skalar} & T & \propto & \exp\left(-\frac{1}{4}\sum_{\langle\hat{m},\hat{m}'\rangle}(\Phi_{\hat{m}}-\Phi_{\hat{m}'})^2\right)\ \exp\left(-\frac{1}{2}\sum_{\hat{m}}\pi^2_{\hat{m}}\right)\ \exp\left(-\frac{1}{4}\sum_{\langle\hat{m},\hat{m}'\rangle}(\Phi_{\hat{m}}-\Phi_{\hat{m}'})^2\right)
\end{array}$$

Die wesentliche Relation beispielsweise beim 1-dim. $\mathbf{R}^3$-Modell ist

$$\exp\left(-\frac{1}{2}(x-y)^2\right) \ \propto\ \left\langle x\left|\exp\left(\frac{1}{2}\frac{\mathrm{d}^2}{\mathrm{d}x^2}\right)\right|y\right\rangle .$$

Führt man eine „Gitterkonstante“ Δt ein, so gilt

$$\exp\left(-\frac{1}{2}\frac{(x-y)^2}{\Delta t^2}\Delta t\right) \ \propto\ \left\langle x\left|\exp\left(\frac{1}{2}\Delta t\frac{\mathrm{d}^2}{\mathrm{d}x^2}\right)\right|y\right\rangle$$

und damit

$$T \ =\ \exp\left(-\frac{1}{2}V(Q)\Delta t\right)\exp\left(-\frac{1}{2}P^2\Delta t\right)\exp\left(-\frac{1}{2}V(Q)\Delta t\right) \ \stackrel{\Delta t\to 0}{\longrightarrow}\ \exp -\Delta t\left(\frac{1}{2}P^2+V(Q)\right) .$$

Als „Hamilton-Operator“ erhalten wir somit:

$$\begin{array}{lrcl}
\text{1-dim. Ising} & H & \propto & \sigma_1 \\
\text{1-dim. } \mathbf{R}^3 & H & \propto & \dfrac{1}{2}\boldsymbol{P}^2 + V(\boldsymbol{Q}) \\
d\text{-dim. Skalar} & H & \propto & \dfrac{1}{2}\sum_{\hat{m}}\pi^2_{\hat{m}} \ +\ \dfrac{1}{2}\sum_{\langle\hat{m},\hat{m}'\rangle}(\Phi_{\hat{m}}-\Phi_{\hat{m}'})^2
\end{array}$$

Kapitel 13

Eichtheorie mit skalaren Materiefeldern

13.1 Allgemeine Formulierung

Die motivierenden Vorbemerkungen zu Kapitel 10 deuteten bereits an, wie skalare Felder und Eichfelder auf dem Gitter zu koppeln sind. Sei $\varphi(\boldsymbol{m}) \in V$ ein Feld, definiert auf den Gitterpunkten, und V ein Raum, auf den die Gruppe G in einer Darstellung $\rho[U]$ $(U \in G)$ wirkt. Im Allgemeinen wird V ein Vektorraum sein, aber auch ein invarianter Unterraum bietet sich an, z.B. Felder mit $|\varphi| = 1$.

Wir definieren folgende Wirkung:

$$\begin{aligned} S[\{U,\varphi\}] &= \frac{2}{g^2 N}\sum_p \mathrm{tr}\,\left(1-\frac{1}{2}(U_p+U_p^+)\right) \; - \\ &\quad - \; \lambda\sum_{m,\mu}\varphi^+(\boldsymbol{m}+\boldsymbol{e}_\mu)\rho[U(\boldsymbol{m},\mu)]\varphi(\boldsymbol{m}) \; + \; \sum_m V(\varphi(\boldsymbol{m}))\,. \end{aligned} \tag{13.1}$$

Hierbei ist

$$V(\varphi(\boldsymbol{m})) \; = \; V(\rho[g(\boldsymbol{m})]\varphi(\boldsymbol{m}))$$

eine eichinvariante Funktion des Feldes, sehr oft eine Funktion von $\varphi^+\varphi$. $\rho[g]$ bzw. $\rho[U]$ bezeichnet eine Darstellung der Gruppe G, die auf die Felder φ wirkt. Diese muss nicht dieselbe Darstellung sein, die bisher einfach mit U bezeichnet wurde, und die im reinen Eichsektor der Wirkung steht. λ bezeichnet die Kopplungskonstante zwischen Eichfeldern und Materie. In Anlehnung an die entsprechenden Größen aus der Festkörperphysik bezeichnet man λ auch als „Hopping"-Parameter.

Offensichtlich ist diese Wirkung invariant unter den lokalen Eichtransformationen

$$\begin{aligned} \varphi(\boldsymbol{m}) &\to \rho[g(\boldsymbol{m})]\varphi(\boldsymbol{m}) \\ U(\boldsymbol{m},\mu) &\to g(\boldsymbol{m}+\boldsymbol{e}_\mu)U(\boldsymbol{m},\mu)g^{-1}(\boldsymbol{m})\,. \end{aligned}$$

13.2 Grenzfälle im Phasendiagramm

Die Aufgabe besteht nun in der Erstellung des Phasendiagramms als Funktion der Kopplungsparameter g und λ sowie des Potentials V. Allgemein ist dies nur mit Hilfe von Monte-Carlo-Simulationen möglich, allerdings erhält man erste Hinweise auf die Phasenstruktur durch die Betrachtung zweier Grenzfälle. Eine ausführliche Darstellung der Phasenstruktur von Eichtheorien mit skalaren Feldern stammt von Fradkin und Shenker [26].

13.2.1 Die reine Materietheorie für $g \to 0$

Die reine Materieheorie erhält man im Grenzfall $g \to 0$. In diesem Limes wird der Faktor zum reinen Eichanteil der Wirkung unendlich groß und Konfigurationen, für die nicht auf allen Plaketten $U_p \equiv \mathbb{1}$ gilt, tragen zur Zustandssumme nicht mehr bei. $U_p \equiv \mathbb{1}$ bedeutet aber, dass die Krümmung verschwindet und die Konfiguration eichäquivalent zur trivialen Konfiguration $U(\boldsymbol{m}, \mu) = \mathbb{1}$ ist. In diesem Fall gibt es ein Feld $\{g(\boldsymbol{m})\}$ von Gruppenelementen auf den Gitterpunkten, sodass sich jede Gruppenvariable auf einer Gitterlinie in der Form $U(\boldsymbol{m}, \mu) = g^{-1}(\boldsymbol{m}+\boldsymbol{e}_\mu)g(\boldsymbol{m})$ schreiben lässt. Dies zeigt man, indem man sukzessive die Variablen $U(\boldsymbol{m}, \mu)$ zur Identität eicht. Umgekehrt definiert jedes Feld $\{g(\boldsymbol{m})\}$ eine Konfiguration ohne Krümmung.

Die effektive Wirkung ist in diesem Fall

$$S[\{g, \varphi\}] \;=\; -\lambda \sum_{m,\mu} \varphi^+(\boldsymbol{m}+\boldsymbol{e}_\mu)\rho[g(\boldsymbol{m}+\boldsymbol{e}_\mu)]^{-1}\rho[g(\boldsymbol{m})]\varphi(\boldsymbol{m}) \;+\; \sum_m V(\varphi(\boldsymbol{m}))\,.$$

Für jede Konfiguration $\{g(\boldsymbol{m})\}$ führt die Variablentransformation $\varphi(\boldsymbol{m}) \to \varphi'(\boldsymbol{m}) = \rho[g(\boldsymbol{m})]\varphi(\boldsymbol{m})$ zu einer Wirkung, die nur noch von dem Feld φ' abhängt. Da das Maß $\mathrm{d}\varphi$ unter dieser Transformation invariant sein soll, hängt die Wirkung nicht von $\{g(\boldsymbol{m})\}$ ab. Wir erhalten in dem Grenzfall $g \to 0$ somit die reine Materietheorie für φ. Ob die reine Materietheorie einen Phasenübergang besitzt, beispielsweise zwischen einer Phase, in der die (globale) Invarianz unter G gebrochen ist, und einer ungebrochenen Phase, hängt von dem Potential V ab.

13.2.2 Die reine Eichtheorie für $\lambda \to 0$

Im Grenzfall $\lambda \to 0$ entkoppeln die Eichvariable $\{U\}$ und die Felder φ. Darüber hinaus entkoppeln die Felder φ an verschiedenen Gitterpunkten und die Integration über die Materiefelder führt in der Zustandssumme zu einem Faktor

$$\mathcal{N} = \left(\int \mathrm{d}\varphi \, \mathrm{e}^{-V(\varphi)} \right)^{L^d} ,$$

der sich bei der Bildung von Erwartungswerten weghebt.

Es verbleibt in dem Grenzfall $\lambda \to 0$ somit nur die reine Eichtheorie. Ob diese einen Phasenübergang zwischen einer Confinement-Phase (Stringtension ungleich 0) und einer „freien Ladungs“-Phase (Stringtension gleich 0) hat, hängt von der Eichgruppe und der Dimension des Gitters ab.

13.3 Beispiel: Z_2-Eichtheorie mit Z_2-Feldern

Das folgende Beispiel ist sehr einfach, zeigt aber alle typischen Charakteristika, die im Allgemeinen in Eichtheorien mit Materiefeldern auftreten, insbesondere bezüglich der Phasenstruktur.

Als Gruppe wählen wir $G = \mathbf{Z}_2 = \{+1, -1\}$ mit der üblichen Multiplikation. Wir definieren Variable $z_l = \pm 1$ ($l \simeq (\boldsymbol{m}, \mu)$) auf den Gitterlinien l. Da $z = z^{-1}$ benötigt man keine Orientierung der Linien. Außerdem definieren wir Variable $s_i = \pm 1$ auf den Gitterpunkten ($i \simeq \boldsymbol{m}$). Dann ist die Wirkung:

$$S \;=\; \beta \sum_p \left(1 - \prod_{l \in \partial p} z_l \right) \;-\; \lambda \sum_{\langle i,j \rangle} (s_i z_{ij} s_j - 1) \,.$$

Hierbei benutze ich einige vereinfachende, übliche und hoffentlich suggestive Notationen: $\prod_{l \in \partial p}$ soll andeuten, dass das Produkt über die Gitterlinien einer Plakette zu bilden ist. ($G = \mathbf{Z}_2$ ist eine Abelsche Gruppe, sodass die Reihenfolge keine Rolle spielt.) $\langle i, j \rangle$ ist eine Standardnotation für benachbarte Gitterpunkte i, j, dementsprechend bezeichnet z_{ij} das Gruppenelement zu der Gitterlinie, welche zu den benachbarten Punkten i und j gehört.

Die im vorherigen Abschnitt untersuchten Grenzfälle ergeben:

- für $\lambda \to 0$ die reine $\mathbf{Z}_2$-Eichtheorie. In drei Dimensionen gibt es einen Phasenübergang 1. Ordnung, in vier Dimensionen einen Phasenübergang 2. Ordnung.
- für $\beta \to \infty$ das reine Ising-Modell. Dieses besitzt in allen Dimensionen ≥ 2 einen Phasenübergang 2. Ordnung zwischen einer gebrochenen und einer ungebrochenen Phase.

Betrachten wir weiterhin noch den Grenzfall $\beta \to 0$. Der reine Eichanteil trägt nicht zur Wirkung bei, und wir erhalten für die Zustandssumme:

$$\begin{aligned} Z &= \sum_{z_l = \pm 1} \sum_{s_i = \pm 1} \prod_{\langle i,j \rangle} \exp(\lambda (s_i z_{ij} s_j - 1) \\ &= \mathrm{e}^{-\lambda d L^d} \sum_{s_i = \pm 1} \prod_{\langle i,j \rangle} 2 \cosh(\lambda s_i s_j) \\ &= 2^{dL^d} \left(1 + \mathrm{e}^{-2\lambda}\right)^{dL^d} . \end{aligned}$$

(Der hyperbolische Kosinus ist eine symmetrische Funktion, d.h. $\cosh(\lambda s_i s_j)$ hängt nicht von $s_i s_j$ ab. Außerdem wurde in dieser Rechnung, wie im Falle diskreter Gruppen üblich, einfach über die Gruppe summiert. Eine Normierung des Gruppenvolumens auf 1, wie sie für das Haar-Maß kontinuierlicher Gruppen üblich ist (Gl. 10.19), ergäbe einen Faktor $(1/2)^{dL^d}$.)

Untersuchen wir abschließend noch den Grenzfall $\lambda \to \infty$. Für das hier betrachtete $\mathbf{Z}_2$-Modell wird auch dieser Grenzfall trivial: Es tragen nur solche Konfigurationen bei, für die $s_i z_{ij} s_j = 1$. Bei beliebiger Vorgabe von $\{s_i\}$ sind die Werte für $\{z_{ij}\}$ damit festgelegt: $z_{ij} = s_i s_j$. Da dies einer „pure gauge“-Konfiguration entspricht, ist das Produkt der z-Variablen um eine Plakette immer gleich 1, und auch die Eichwirkung verschwindet.

Für allgemeinere Gittereichtheorien mit Materiefeldern ist die Situation komplexer. Betrachten wir als Beispiel ein reelles 3-komponentiges Feld $\varphi(\boldsymbol{m}) \in \mathbb{R}^3$, das durch ein geeignetes

Potential noch an jedem Punkt der Bedingung $|\varphi(\boldsymbol{m})| = 1$ unterliegen soll. Das Feld nimmt also Werte auf einer 2-dimensionalen Kugeloberfläche an. Die Symmetriegruppe sei $G = \mathrm{SO}(3)$. Zunächst können wir durch eine Eichtransformation das Feld an jedem Punkt in den Vektor $\hat{\varphi} \simeq (1, 0, 0)$ rotieren. Im Grenzfall $\lambda \to \infty$ tragen nur solche Konfigurationen bei, für die gilt

$$\varphi(\boldsymbol{m} + \mu)U(\boldsymbol{m}, \mu)\varphi(\boldsymbol{m}) \;=\; \hat{\varphi}\hat{U}(\boldsymbol{m}, \mu)\hat{\varphi} \;=\; 1\,. \tag{13.2}$$

Offensichtlich ist durch diese Bedingung das Gruppenelement noch nicht eindeutig festgelegt. Sei $H \subset G$ die so genannte „little group“, für die $H\hat{\varphi} = \hat{\varphi}$, die also den Vektor $\hat{\varphi}$ invariant lässt. Für die Elemente aus dieser Gruppe bleibt bei vorgegebenem $\hat{\varphi}$ die Bedingung (13.2) erhalten. Im Grenzfall $\lambda \to \infty$ wird die Eichtheorie daher nur bis auf die Untergruppe H eingefroren.

13.4 Ordnungsparameter und Symmetriebrechung

Wir haben gesehen, dass eine Gittereichtheorie mit Materiefeldern im Grenzfall $\lambda \to 0$ zur reinen Eichtheorie wird, im Grenzfall $\beta \to \infty$ zur reinen Materietheorie. Viele reine Gittereichtheorien zeigen einen Phasenübergang zwischen zwei Bereichen, die durch das asymptotische Verhalten des Wilson-Loops gekennzeichnet sind („Flächen“- bzw. „Umfangs“-Gesetz). Die String-tension κ dient dabei als Ordnungsparameter. Reine Materietheorien zeigen ebenfalls oft einen Phasenübergang zwischen einer geordneten und einer ungeordneten Phase. Als Ordnungsparameter dient beispielsweise im Ising-Modell die Magnetisierung bzw. der asymptotische Wert der Korrelationsfunktion (siehe unten). Von beiden Phasenübergängen wird man erwarten, dass sie bei einer Kopplung zwischen Eichtheorie und Materietheorie zumindest in einem gewissen Parameterbereich bestehen bleiben. Was aber sind geeignete Ordnungsparameter ? Die Korrelationsfunktion $\langle s_i s_j \rangle$ verschwindet wegen der Eichinvarianz identisch für $i \neq j$, und der Wilson-Loop zeigt – wie wir unten argumentieren werden – immer Umfangsgesetz.

13.4.1 Ordnungsparameter in einer reinen Materietheorie — Spontane Symmetriebrechung

Für das Ising-Modell ist bekannt, dass es in jeder Dimension $d \geq 2$ einen Phasenübergang gibt. Dieser Phasenübergang ist zweiter Ordnung und trennt die geordnete von der ungeordneten Phase. „Geordnet“ bedeutet dabei, dass die Spins bevorzugt in eine Richtung zeigen. Als Ordnungsparameter dient

$$\rho \;=\; \lim_{|i-j|\to\infty} \langle s_i s_j \rangle \;.$$

In der geordneten Phase ist $\rho \neq 0$, in der ungeordneten verschwindet ρ identisch.

Man würde zunächst vermuten, dass

$$\tilde{\rho} \;=\; \langle s_i \rangle \tag{13.3}$$

ein geeigneter Ordnungsparameter sei. Da jedoch

1. die Wirkung
2. das Maß

3. die Randbedingungen (periodisches Gitter)

die Symmetrie

$$s_i \longrightarrow -s_i \quad \forall i$$

besitzen, ist $\tilde{\rho} \equiv 0$ in jeder Phase. Trotzdem hat man den Eindruck, dass für tiefe Temperaturen, also in der geordneten Phase, eine entsprechende Größe die „spontane Symmetriebrechung“ anzeigen sollte. Der Ausdruck „spontane Symmetriebrechung“ bedeutet dabei, dass die drei oben genannten Anteile des Modells (Wirkung, Maß, Randbedingungen) eine Symmetrie besitzen, die vom Grundzustand der Theorie nicht respektiert wird. Diese spontane Symmetriebrechung wird erst im thermodynamischen Limes manifest.

Man kann einen der Größe $\tilde{\rho}$ (13.3) entsprechenden Ordnungsparameter definieren, wenn man für das Modell „explizit“ die Symmetrie bricht. Dies geschieht z.B. in der Wirkung durch Einführung eines konstanten äußeren „Feldes“:

$$S(\beta, h) = -\beta \sum_{\langle i,j \rangle} s_i s_j + h \sum_i s_i \,.$$

Nun definiert man Erwartungswerte bzgl. dieser Wirkung, z.B.

$$\langle s_i \rangle_h = \frac{\sum_{\{s_i\}=\pm 1} s_i \, \exp -S(\beta, h)}{\sum_{\{s_i\}=\pm 1} \exp -S(\beta, h)} \,,$$

nimmt zunächst den thermodynamischen Limes und definiert dann

$$\tilde{\rho} = \lim_{h \to 0} \langle s_i \rangle_h \,.$$

Eine andere Möglichkeit ist die Wahl von symmetriebrechenden Randbedingungen, z.B. feste Randbedingungen mit $s_i = +1$ auf dem Rand des Gitters. Im thermodynamischen Limes bildet $\tilde{\rho}$ wiederum einen geeigneten Ordnungsparameter. Allerdings hängt das Vorzeichen in diesem Fall von der Wahl der Randbedingungen ab (ebenso wie im vorherigen Fall das Vorzeichen von $\tilde{\rho}$ vom Vorzeichen von h abhing). In beiden Fällen gilt

$$\rho = \tilde{\rho}^2 \,.$$

Diese Diskussion der spontanen Symmetriebrechung lässt sich auf jede Art von globaler Symmetrie übertragen. Eine Besonderheit tritt jedoch auf, wenn die Symmetrie kontinuierlich ist, z.B. die Rotationssymmetrie der Magnetisierungsachse bei Modellen für Ferromagneten ($\varphi_i \in \mathbb{R}^3$). Ist im Grundzustand die Symmetrie gebrochen, so gibt es infinitesimale Veränderungen der Felder, die energetisch beliebig klein sind. Es gibt Wellen zu diesen Anregungen, die beliebig langreichweitig sind. Quantisiert man diese Wellen, so nennt man die zugehörigen Teilchen Goldstone-Bosonen. Dieser Zusammenhang zwischen „masselosen“ Teilchen und spontaner Symmetriebrechung gilt nur bei kontinuierlichen, globalen Symmetrien in mehr als 2 Dimensionen.

13.4.2 Ordnungsparameter in einer Eichtheorie mit Materiefeldern

Die Korrelationsfunktion $\langle s_i s_j \rangle$ verschwindet für $i \neq j$ identisch, sobald an die Spinvariable eine Eichvariable gekoppelt wird, da wegen der Eichinvarianz beispielsweise $\langle s_i s_j \rangle = \langle s_i (-s_j) \rangle$ gilt.

Auch die Stringtension, die in der reinen Eichtheorie als Ordnungsparameter dient, verschwindet immer, sobald Materiefelder an die Eichtheorie gekoppelt werden.

Statt eines Beweises für die letzte Behauptung wollen wir ein physikalisches Argument anführen. Wir hatten die Stringtension κ aus dem asymptotischen Verhalten des Potentials zwischen zwei statischen Quarks $q\bar{q}$ definiert. Bildet sich in einer Confinement-Phase zwischen den Quarks eine Flussröhre (fluxtube), so wächst die Energie des Strings proportional zu seiner Länge, nimmt also linear mit dem Abstand der Quarks zu. Dies ist in einer reinen Eichtheorie möglich und kennzeichnet die Phase mit einem „Flächengesetz“ für den Wilson-Loop.

Koppelt man jedoch ein Materiefeld an das Eichfeld, welches in einem gewissen Sinne eine quantenmechanische Beschreibung eines Vielteilchensystems darstellt, so wird die potentielle Energie zwischen zwei statischen Quarks nicht mehr beliebig groß werden können. Reicht die Energie des Strings aus, um aus dem Vakuum ein dynamisches $q\bar{q}$-Paar zu erzeugen, werden sich diese dynamischen Teilchen an ihren statischen Partner anlagern, und man erhält lokal neutrale gebundene Zustände. Diese können nun ohne weiteren Energiezuwachs voneinander getrennt werden.

Ein naiver Ansatz für eine Korrelationsfunktion zwischen zwei Materiefeldern, und damit für einen Ordnungsparameter, wäre:

$$\tilde{G}_{ij} \;=\; \left\langle s_i \prod_{l\in\mathcal{C}:j\to i} z_l \; s_j \right\rangle .$$

Es werden also die Spin-Variable an den Punkten i und j einfach durch einen direkten Paralleltransporter miteinander verbunden. Bezeichnet $\mathcal{C}$ jedoch eine gerade Verbindungslinie, so lässt sich zeigen, dass im Grenzfall $|i-j|\to\infty$ der zugehörige Zustand eine unendliche Energie besitzt und nicht mehr im physikalischen Hilbert-Raum liegt. Aus theoretischen Gründen ist $\rho = \lim_{|i-j|\to\infty} \tilde{G}_{ij}$ daher als Ordnungsparameter abzulehnen. Das Problem wurde von Fredenhagen und Marcu ([28]) gelöst. Sie postulieren folgende 2-Punkt-Funktion

$$G(R,T) \;:=\; \frac{\langle s_i \prod_{l\in\mathcal{C}(R,T)} z_l \; s_j\rangle}{\langle W(R,T)\rangle^{1/2}} ,$$

wobei $W(R,T)$ ein Wilson-Loop mit Kantenlänge R und T ist, und $\mathcal{C}(R,T)$ einen Weg bezeichnet, der i und j (im Abstand R) verbindet, allerdings entlang dreier Seiten eines Rechtecks mit Kantenlängen $T/2$ und R. Symbolisch schreibt man auch

$$G(R,T) \;=\; \frac{\left\langle \begin{array}{ccc} & R & \\ T/2 & \square & \\ i & & j \end{array} \right\rangle}{\left\langle \begin{array}{cc} T & \square \\ & R \end{array} \right\rangle^{1/2}} .$$

Die Größe

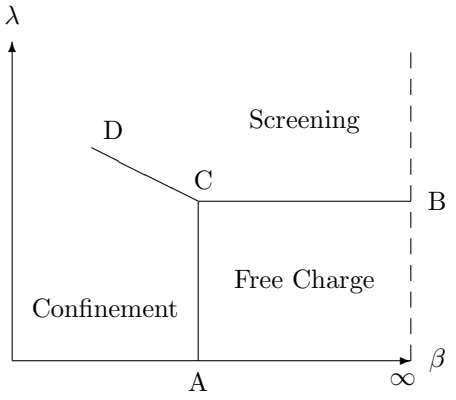
$$\rho \;=\; \lim_{R\to\infty} \lim_{T\to\infty} G(R,T) \tag{13.4}$$

(wobei die Reihenfolge der Grenzwertbildungen wesentlich ist) definiert einen Ordnungsparameter, der für $\beta \to \infty$ gegen den asymptotischen Erwartungswert der 2-Punkt-Korrelationsfunktion, d.h. das Quadrat der Magnetisierung, strebt.

13.5 Das Phasendiagramm

Numerische Rechnungen der $\mathbf{Z}_2$-Eichtheorie mit angekoppelten $\mathbf{Z}_2$- Spin-Variablen ergeben qualitativ folgendes Bild:

Man unterscheidet in diesem Diagramm drei Phasen: freie Ladungsphase, Confinement-Phase und screening-Phase. Dabei sind die Confinement-Phase und die screening-phase jedoch nicht immer durch einen Phasenübergang getrennt. (Dies ist beispielsweise vergleichbar mit der flüssigen und gasförmigen Phase von Wasser.)



Der Ordnungsparameter ρ (vgl. Gl. (13.4)) ist in der Confinement-Phase und der Screening-Phase von Null verschieden. In der „Freie-Ladung“-Phase verschwindet er. Der Phasenübergang entlang $C - D$ ist in 3 und 4 Dimensionen erster Ordnung, ρ macht dort einen Sprung. D ist ein kritischer Punkt (Phasenübergang zweiter Ordnung), bei dem die Korrelationsfunktion $G(R, T \to \infty)$ für große R algebraisch abfällt. Der Phasenübergang entlang $C - B$ ist zweiter Ordnung und zeigt die kritischen Exponenten des reinen Ising-Modells. Der Phasenübergang entlang $A - C$ ist in drei Dimensionen 2. Ordnung, in 4 Dimensionen 1. Ordnung. Für den Kontinuumslimes müssen sich die Parameter Punkt B nähern.

1. Freie Ladungs – Phase:

   In dieser Phase existieren freie, geladene Zustände. Das Potential zwischen Ladungen ist von der Form des Coulomb-Potentials.

2. Confinement – Phase:

   Freie geladene Zustände existieren nicht. Der Versuch, zwei geladene Teilchen zu trennen, erzeugt aus dem Vakuum entsprechende Paare entgegengesetzter Ladung, sodass zwei neutrale Teilchen entstehen. In der Nähe der reinen Eichtheorie ist das Flächengesetz des Wilson-Loops noch relativ gut nachweisbar.

3. Screening – Phase:

   Es existieren ebenfalls keine freien Ladungen. Hier führt der Versuch, zwei geladene Zustände zu trennen, jedoch zu einer Polarisation des Vakuums, welche die Ladungen abschirmt (screening). Im Gegensatz zum Confinement lässt sich kein eindeutiges „Anti-Paar“ ausmachen, welches die Neutralisation der Ladungen bewirkt, sondern es handelt sich um eine ganze Wolke von Teilchen.

Das gleiche Phasendiagramm ergibt sich beispielsweise auch für eine U(1)-Eichtheorie gekoppelt an Materie mit Symmetriebrechung, z.B. ein komplexes Feld vom Betrag 1. In einer Eichtheorie

ohne Freie-Ladung-Phase (z.B. SU(3)) hat man nur den Phasenübergang der reinen Materietheorie, der sich in das Diagramm erstreckt.

Kapitel 14

Das Monte-Carlo-Verfahren

14.1 Vorbemerkungen

In der gitterregularisierten (euklidischen) Quantenfeldtheorie haben wir das Problem, Erwartungswerte von Funktionen auf Konfigurationsräumen zu berechnen:

$$\langle f \rangle \;=\; \frac{\sum_C f[C]\;\exp(-S[C])}{\sum_C \exp(-S[C])}\,.$$

Selbst für diskrete lokale Zustände ist der Konfigurationsraum viel zu groß, als dass modernste Computer diese Summationen ausführen können. Das Ising-Modell auf einem 4^4 Gitter hat $2^{256} \approx 10^{77}$ Konfigurationen. Das ist mehr als ein idealer Computer (der 10^{44} Schritte in der Sekunde ausführt) im bisherigen Zeitalter der Universums auch nur hätte aufzählen können.

Die meisten Konfigurationen werden wegen ihres geringen Gewichtes zu diesen Erwartungswerten im Allgemeinen einen verschwindenden Beitrag liefern. Das Monte-Carlo-Verfahren geht von der Idee aus, dass die üblichen Erwartungswerte schon durch Auswertung und Mittelwertbildung der Observablen auf einem geeigneten Satz von unabhängigen, „typischen“ Konfigurationen $\{C_1, \ldots, C_N\}$ angenähert werden können:

$$\langle f \rangle \;\approx\; \frac{1}{N}\sum_{i=1}^{N} f[C_i]\,. \tag{14.1}$$

Jede Konfiguration sollte mit einer Wahrscheinlichkeit proportional zu ihrem Boltzmann-Faktor

$$w[C] \;\propto\; \exp(-S[C])$$

in dieser Auswahl („sample“) vertreten sein.

Dies kann sicherlich nicht für jede beliebige Funktion auf dem Konfigurationsraum gelten. Sei z.B.

$$f[C] \;=\; \delta_{C_0}(C) \;=\; \begin{cases} 1 & \text{falls } C = C_0 \\ 0 & \text{sonst} \end{cases},$$

wobei C_0 eine beliebig vorgegebene, sehr unwahrscheinliche Konfiguration ist. Dann ist

$$\langle f[C] \rangle \;=\; \frac{1}{Z}\mathrm{e}^{-S[C_0]}\,.$$

Dieses Ergebnis wird man nie erhalten, wenn man nur einen Satz von „wesentlichen“ Konfigurationen als „sample“ benutzt. Im Idealfall ist N sehr viel größer als die Anzahl der möglichen Elementarereignisse, also $N \gg |\Omega|$, sodass jede Konfiguration tatsächlich mehrmals — proportional zu ihrer Wahrscheinlichkeit — in diesem „sample“ vertreten ist. Dieser Idealfall lässt sich in der Praxis natürlich nicht erreichen.

Näheres zum Monte-Carlo-Verfahren im Allgemeinen findet man beispielsweise bei [49, 13]

14.2 (Diskrete) Markov-Prozesse

Die durch die Boltzmann-Gewichte vorgegebene Verteilung der Konfigurationen wird im Monte-Carlo-Verfahren als stationäre Verteilung eines Markov-Prozesses realisiert. Daher wollen wir kurz einige allgemeine Bemerkungen über Wahrscheinlichkeitsräume im Allgemeinen und zu Markov-Prozessen im Speziellen voranstellen.

14.2.1 Wahrscheinlichkeitsräume

Ein *Wahrscheinlichkeitsraum* $(\Omega, \mathcal{B}, w)$ besteht aus einer Menge Ω – die Menge der so genannten *Elementarereignisse* – einer σ-Algebra $\mathcal{B} \subset 2^\Omega$ und einem Wahrscheinlichkeitsmaß $w : \mathcal{B} \to [0, 1]$.

Eine σ-Algebra besteht dabei aus einer Menge von Teilmengen aus Ω, die bezüglich Komplementbildung, abzählbaren Vereinigungen und Durchschnittsbildungen abgeschlossen ist. (Außerdem ist $\Omega \in \mathcal{B}$ und $\emptyset \in \mathcal{B}$.) Ein Wahrscheinlichkeitsmaß erfüllt neben den Maßeigenschaften (Positivität sowie Additivität bei disjunkten Vereinigungen) noch die Normierungsbedingung $w(\Omega) = 1$.

Da wir im Folgenden jedoch nicht an maßtheoretischen Feinheiten interessiert sind, nehmen wir der Einfachheit Ω als endliche Menge an. $\mathcal{B}$ sei dann die Menge aller Teilmengen und für jedes Elementarereignis $C \in \Omega$ lässt sich seine Wahrscheinlichkeit $w(C)$ angeben.

Die Menge der *Observablen* (oder auch *Zufallsvariablen*) besteht aus den Funktionen über Ω: $F = \{f | f : \Omega \to \mathbb{R}\}$. Das Wahrscheinlichkeitsmaß w definiert dann einen so genannten *Zustand* $\omega : F \to \mathbb{R}$ (manchmal auch als *Erwartungswertfunktional* bezeichnet) über der Observablenmenge:

$$\omega(f) \;=\; \langle f \rangle \;=\; \sum_{C \in \Omega} f(C) w(C) \; .$$

Offensichtlich ist ω *linear*, *positiv* ($\omega(|f|) \geq 0$), und *normiert* ($\omega(1) = 1$).

In unserem Fall ist Ω der Raum aller Konfigurationen C, w ist durch den Boltzmann-Faktor gegeben,

$$w(C) \;=\; \frac{1}{Z} \mathrm{e}^{-S[C]} \qquad\qquad Z \;=\; \sum_{C \in \Omega} \mathrm{e}^{-S[C]} \, ,$$

und die Observablen sind allgemeine Funktionen der Konfigurationen (insbesondere beispielsweise die Spin-Variable an bestimmten Gitterpunkten oder die Wilson-Loops, d.h. das Produkt von Gruppenelementen entlang geschlossener Wege), von denen wir die Erwartungswerte bestimmen möchten.

14.2.2 Wahrscheinlichkeitsräume von Folgen von Elementarereignissen

Bei einem stochastischen Prozess lassen sich die Elementarereignisse als Funktionen von einer „Zeit“ in einen Raum von Elementarereignissen auffassen. Da bei den Markov-Prozessen diese „Zeit“ immer diskret ist, verzichten wir hier auf eine vollständige Behandlung stochastischer Prozesse in einer kontinuierlichen Zeit.

Zu einer gegebenen Menge von Elementarereignissen Ω können wir eine neue Menge von Elementarereignissen definieren, die aus Folgen von Elementen von Omega bestehen soll:

$$\tilde{\Omega}_N \;=\; \{(C_1, C_2, \ldots, C_N) | C_i \in \Omega\}\,.$$

Man kann diesen neuen Ereignisraum $\tilde{\Omega}$ auch als die Menge der Abbildungen von $\{1, \ldots, N\}$ in den Ereignisraum Ω auffassen. Observable über $\tilde{\Omega}_N$ sind nun Funktionen von diesen Folgen von Elementarereignissen: $\tilde{F}_N \simeq \{f(C_1, \ldots, C_N)\}$. Insbesondere können wir zu einer Oberablen f über Ω eine Observable $\tilde{f}$ definieren, die dem Mittelwert von f auf einer Folge von Elementarereignissen entspricht:

$$\tilde{f}(C_1, \ldots, C_N) \;=\; \frac{1}{N} \sum_{k=1}^{N} f(C_k)\;.$$

Im Allgemeinen werden Wahrscheinlichkeiten $\tilde{w}$ über $\tilde{\Omega}$ Funktionen von Folgen der Elementarereignisse sein: $\tilde{w} = \tilde{w}(C_1, \ldots, C_N)$. Wir wollen zwei besondere Wahrscheinlichkeiten $\tilde{w}$ hervorheben:

- $\tilde{w}(C_1, \ldots, C_N) = \prod_{k=1}^{N} w(C_k)$

  In diesem Fall ist die Gesamtwahrscheinlichkeit für die Folge der Elementarereignisse einfach das Produkt der Einzelwahrscheinlichkeiten. Die Ereignisse sind nicht korreliert, d.h.

  $$\langle C_i C_j \rangle \;-\; \langle C_i \rangle \langle C_j \rangle \;=\; 0 \qquad \text{für } i \neq j\,.$$

  Dieser Fall beschreibt idealisierte „Experimente“, die man zu einer Theorie mit Wahrscheinlichkeitsaussagen macht: Die experimentellen Ergebnisse sollten voneinander unabhängig sein, und für jedes einzelne Experiment sollte das Ereignis C mit Wahrscheinlichkeit $w(C)$ auftreten. Die relative Häufigkeit für das Ereignis C_0 in einer Folge

  $$\begin{aligned} P_{C_0}(C_1, \ldots, C_N) \;&=\; \frac{1}{N} \sum_{k=1}^{N} \delta_{C_0}(C_k) \\ &\text{mit } \delta_{C_0}(C) \;=\; \begin{cases} 1 & \text{falls } C_0 = C \\ 0 & \text{sonst} \end{cases} \end{aligned}$$

  ist eine Observable über $\tilde{\Omega}$, und es gilt

  $$\langle P_{C_0} \rangle \;=\; w(C_0)\,.$$

  Wichtiger ist jedoch die folgende Eigenschaft im Grenzfall $N \to \infty$:

  $$\langle (\Delta P_{C_0})^2 \rangle \;=\; \langle (P_{C_0} - w(C_0))^2 \rangle \;\stackrel{N\to\infty}{\longrightarrow}\; 0\,.$$

Mit „Wahrscheinlichkeit 1“ ist für $N \to \infty$ die relative Häufigkeit eines Ereignisses in Folgen aus Ω gleich seiner Wahrscheinlichkeit.

- $\tilde{w}(C_1, \ldots, C_N) = P(C_N, C_{N-1})P(C_{N-1}, C_{N-2}) \ldots P(C_2, C_1)$

  Die einzelnen Ereignisse in einer Folge sind also nun nicht mehr unabhängig, sondern in spezieller Weise untereinander gekoppelt. $P(C, C')$ ist eine so genannte *Markov-Matrix* (manchmal auch *stochastische Matrix* genannt) mit folgenden Eigenschaften:

$$P(C, C') \;\geq\; 0 \qquad \text{und} \qquad \sum_{C \in \Omega} P(C, C') \;=\; 1\,. \tag{14.2}$$

  Anschaulich wird also $(C_1, \ldots, C_N)$ nun zu einer Folge von Ereignissen, wobei die Wahrscheinlichkeit „von C_i nach C_{i+1} zu gelangen“, nur von diesen beiden Ereignissen selber abhängt, nicht aber von den vorherigen oder folgenden Ereignissen. Solch einen Wahrscheinlichkeitsraum bezeichnet man als diskreten *Markov-Prozess* (1. Ordnung). Den Index an den Einzelereignissen nennt man auch (diskrete) *Markov-Zeit*.

  Für $P(C, C')$ schreibt man auch manchmal $P(C' \to C)$ um anzudeuten, dass es sich um die Wahrscheinlichkeit handelt, von C' nach C zu gelangen. Im Allgemeinen ist die Markov-Matrix nicht symmetrisch, d.h. $P(C, C') \neq P(C', C)$. Der folgende Abschnitt fasst einige wichtige Eigenschaften von Markov-Matrizen zusammen.

14.2.3 Markov-Matrizen

Sei $P(C, C')$ eine Markov-Matrix (mit den Eigenschaften (14.2)). Außerdem verlangen wir noch, dass es ein n gibt, sodass

$$P^n(C \to C') \;=\; \sum_{C_1 \in \Omega} \sum_{C_2 \in \Omega} \ldots \sum_{C_{n-1} \in \Omega} P(C \to C_1)\; P(C_1 \to C_2)\; \ldots \tag{14.3}$$
$$\ldots P(C_{n-1} \to C') \;>\; 0\,,$$

für je zwei Konfigurationen C und C'. Physikalisch bezeichnet man diese Eigenschaft manchmal als *Ergodizität*. Sie bedeutet nämlich, dass es nach n (und mehr) Schritten möglich ist, von jeder Konfiguration C zu jeder anderen Konfiguration C' zu gelangen. In der Theorie positiver Matrizen wird diese Eigenschaft auch manchmal „Irreduzibilität“ bzw. “Primitivität“ genannt. Äquivalent kann man die Eigenschaft der Irreduzibilität auch dadurch charakterisieren, dass es nicht möglich ist, durch eine Permutation von Zeilen und Spalten die Matrix P auf eine Blockdiagonalgestalt zu bringen.

Für irreduzible Markov-Matrizen gilt folgender Satz (siehe z.B. [63]):

1. *Eine Markov-Matrix mit der Eigenschaft* (14.3) *hat einen Eigenvektor zum Eigenwert* 1. *Dieser Eigenwert ist nicht entartet, außerdem gilt für alle anderen Eigenwerte* λ *der Markov-Matrix die Bedingung* $|\lambda| < 1$.

2. *Seien* $\hat{w}[C]$ *die Komponenten eines Eigenvektors von* P *zum Eigenwert 1, so sind die Komponenten* $\hat{w}[C]$ *entweder alle nicht-negativ oder nicht-positiv. Nach Multiplikation mit einer geeigneten Konstanten können also alle Komponenten nicht-negativ und normiert gewählt werden:* $\hat{w}[C] > 0\ \forall C$ *, und* $\sum_C \hat{w}[C] = 1$. $\hat{w}$ *bezeichnet man auch als stationären Zustand des Markov-Prozesses* P.

3. Es existiert der Grenzwert

$$\begin{aligned} & \lim_{N\to\infty} P^N &=& P^\infty \\ \textit{und zwar ist} \qquad & P^\infty(C \to C') &=& \hat{w}[C']\,. \end{aligned}$$

Außerdem lässt sich zeigen, dass

$$\lim_{N\to\infty} \langle \tilde{f}(C_1,\ldots,C_N)\rangle \;=\; \lim_{N\to\infty} \frac{1}{N}\sum_{k=1}^{N} f(C_k) \;=\; \sum_C f(C)\hat{w}[C]\,.$$

Diese Eigenschaft, dass die Mittelung einer Funktion f über einer Folge von Ereignissen aus einer Markov-Kette gleich dem Erwartungswert dieser Funktion bezüglich der Verteilung $\hat{w}$ ist, gilt wiederum „mit Wahrscheinlichkeit 1“. Anders ausgedrückt, der Erwartungswert von $(\Delta\tilde{f})^2 = (\tilde{f} - \langle f\rangle)^2$ geht für $N \to \infty$ gegen Null.

Wir haben also folgende Situation: Gegeben eine Markov-Matrix P mit einem (eindeutigen) Eigenvektor $\hat{w}[C]$ zum Eigenwert 1. Dann generiert eine Realisation dieser Markov-Matrix eine Folge von Elementarereignissen $(C_1, C_2, \ldots, C_N)$, die für $N \to \infty$ einer Verteilung $\hat{w}[C]$ entspricht. Die Mittelung einer Funktion f über dieser Folge ist gleich dem Erwartungswert von f bezüglich der Verteilung $\hat{w}[C]$.

14.2.4 Detailed balance

Zunächst ist der Eigenzustand $\hat{w}[C]$ eine Eigenschaft der gewählten Markov-Matrix P und hat daher nichts mit einer vorgegebenen Verteilung $w[C]$ auf Ω zu tun. Für Monte-Carlo-Simulationen ist aber die Verteilung $\omega[C]$ vorgegeben (durch die Boltzmann-Faktoren). Wir suchen also eine Realisation einer Markov-Matrix P, sodass die Verteilung $\omega[C]$ gerade der Eigenvektor zum Eigenwert 1 von P ist.

Was garantiert aber, dass eine vorgegebene Wahrscheinlichkeitsverteilung $w[C]$ gerade Eigenvektor eines Markov-Prozesses $P(C \to C')$ zum Eigenwert 1 ist ? Eine *hinreichende* Bedingung ist die so genannte *detailed balance*:

$$w[C]\; P(C \to C') \;=\; w[C']\; P(C' \to C)\,. \tag{14.4}$$

Summation beider Seiten dieser Gleichung über alle C' ergibt wegen der Markov-Eigenschaft (14.2)

$$w[C] \;=\; \sum_{C'} w[C']\; P(C' \to C)\,,$$

d.h. $w[C]$ sind gerade die Komponenten des Eigenvektors der Matrix P mit Eigenwert 1.

Anschaulich besagt die Bedingung der detailed balance, dass in einem dynamischen System der Prozess $C \to C'$ ebenso häufig stattfindet wie der Prozess $C' \to C$, also eine Art „Fließgleichgewicht“ herrscht. Die Häufigkeit des Prozesses $C \to C'$ ist dabei das Produkt aus der Wahrscheinlichkeit, dass Konfiguration C vorliegt ($w[C]$), multipliziert mit der Wahrscheinlichkeit, bei der vorgegebenen Dynamik von C nach C' zu gelangen ($P(C \to C')$). Betrachtet man also eine (idealisierte) Folge $(C_1, C_2, \ldots)$ die von P generiert wurde und untersucht, wie oft ein bestimmtes Paar von Elementarereignissen (C, C') in dieser Folge auftritt — wie oft also

der Prozess $C \to C'$ stattgefunden hat —, so stellt man fest, dass die Paare (C, C') und (C', C) gleichhäufig vertreten sind.

Man beachte, dass in der Bedingung für detailed balance die Normierung von $w[C]$ herausfällt, es genügt somit die Kenntnis der Boltzmann-Faktoren $\exp(-S[C])$:

$$P(C \to C')\,\mathrm{e}^{-S[C]} \;=\; P(C' \to C)\,\mathrm{e}^{-S[C']}\,. \tag{14.5}$$

Wie bereits erwähnt wurde, ist die Bedingung der detailed balance nur eine hinreichende Bedingung dafür, dass eine vorgegebene Verteilung $w[C]$ gerade Eigenvektor einer Markov-Matrix P zum Eigenwert 1 ist. Allerdings ist sie sehr leicht zu erfüllen, und daher basieren nahezu alle Monte-Carlo-Algorithmen auf einem Markov-Prozess, für den die Bedingung der detailed balance gilt.

14.3 Monte-Carlo-Simulation

Wir werden später die gängigsten Verfahren zur Realisation von P in einer Monte-Carlo-Simulation auf dem Computer — insbesondere das so genannte Metropolis-Verfahren sowie das „heat bath"-Verfahren — kurz besprechen. An dieser Stellen wollen wir das allgemeine Vorgehen skizzieren. Da bei konkreten Realisationen die Anzahl der Konfigurationen N immer endlich ist, ergeben sich Abweichungen vom richtigen Resultat, die man möglichst klein halten möchte. (Siehe auch [54, 11].)

14.3.1 Thermalisierung und Messung

Bei einer Monte-Carlo-Simulation auf dem Computer wählt man zunächst eine Anfangs- oder Startkonfiguration C_0. Diese kann „geordnet" oder rein zufällig sein. Manchmal übernimmt man auch eine geeignete Konfiguration einer früheren Simulation.

Ausgehend von C_0 konstruiert man eine neue Konfiguration C_1 mit einer Wahrscheinlichkeit $P(C_0 \to C_1)$, die detailed balance bezüglich der durch $w[C] \propto \exp(-S[C])$ definierten Wahrscheinlichkeitsverteilung erfüllt. Diesen Prozess wiederholt man sukzessive, d.h. ausgehend von C_i konstruiert man C_{i+1} mit einer Wahrscheinlichkeit $P(C_i \to C_{i+1})$, so oft, bis man überzeugt ist, dass der Prozess zum Gleichgewicht konvergiert ist (also $P^N(C_0 \to C)$ nicht mehr von C_0 abhängt). Diesen ersten Teil der Simulation nennt man „Thermalisierung".

Anschließend beginnt man mit den so genannten „Messungen". Im Folgenden bezeichnen C_i $(i = 1, \ldots, N)$ die Konfigurationen, die man nach Abschluss der Thermalisierung (also in der durch die Boltzmann-Faktoren definierten Gleichgewichtsverteilung) generiert. Auf jeder Konfiguration C_i wertet man die Observablen $\{f\}$ (d.h. die Funktionen auf dem Konfigurationsraum) aus, deren Erwartungswerte man bestimmen möchte. Man addiert die so erhaltenen Werte $f[C_i]$ und dividiert am Ende durch die Gesamtzahl der Messungen (14.1).

Fast alle Algorithmen (mit Ausnahme der so genannten Cluster-Algorithmen) konstruieren C_{i+1} schrittweise durch lokale Änderungen von C_i. Dabei wird die Konfiguration C_i nur an einem Gitterpunkt $\vec{m}$ (bzw. einer Gitterlinie $(\vec{m}, \mu)$) verändert. Man geht meist systematisch durch das Gitter, da es sich gezeigt hat, dass eine zufällige Wahl der Gitterpunkte $\vec{m}$, an denen die Konfiguration verändert wird, keine Vorteile bringt.

14.3.2 Das Problem der Autokorrelationen

Die Konfigurationen, die man auf diese Weise konstruiert, sind sehr korreliert. Selbst wenn man die lokalen Änderungen an jedem Gitterpunkt vorgenommen hat, erhält man eine Konfiguration, die immer noch „sehr nahe“ an der alten Konfiguration ist. „Sehr nahe“ bedeutet dabei, dass die meisten Funktionen auf dem Konfigurationsraum, deren Erwartungswerte man berechnen möchte, ihren Wert nur wenig verändern. Gibt man sich z.B. eine feste Funktion $F : \{C\} \to \mathbb{R}$ vor, so erhält man eine Serie von Zahlen $F[C_1], F[C_2], \ldots, F[C_i], \ldots$, die korreliert sind. Die „Autokorrelationsfunktion“

$$\begin{aligned} G_F(k) &= \lim_{N\to\infty} \frac{1}{N-k} \sum_{i=1}^{N-k} \left(F[C_i] - \langle F\rangle\right)\left(F[C_{i+k}] - \langle F\rangle\right) \;\propto\; \mathrm{e}^{-\tau k} \\ \langle F\rangle &= \frac{1}{N}\sum_{i=1}^{N} F[C_i] \end{aligned}$$

definiert eine so genannte *Autokorrelationszeit* τ, bzw. eine *Autokorrelationslänge* $z = 1/\tau$. Diese hängt sowohl von dem zu simulierenden System, als auch von der Wahl des Algorithmus ab. Je näher ein System einem kritischen Punkt ist (d.h. je langreichweitiger Korrelationsfunktionen auf dem Gitter sind), desto größer ist auch die Autokorrelationslänge. Da für den Kontinuumslimes ein System in der Nähe eines kritischen Punktes sein muss, sind solche Monte-Carlo-Simulationen sehr aufwendig.

Die Existenz dieser Autokorrelationen muss bei der Berechnung des Fehlers bzw. der Varianz für einen Erwartungswert berücksichtigt werden. Entschließt man sich, die Observablen nur an unkorrelierten Konfigurationen auszuwerten, d.h. nur an Konfigurationen, die wesentlich mehr als z Monte-Carlo-Schritte auseinander liegen, so kann man den Fehler in der üblichen Weise durch die Varianz der „Messwerte“ bestimmen. Das bedeutet aber, dass weit mehr Konfigurationen generiert werden müssen, als letztendlich zur Auswertung kommen. Anderseits kann man natürlich die Observable auf jeder Konfiguration auswerten, dann muss bei der Bestimmung des Fehlers aber die Autokorrelation berücksichtigt werden.

Die Autokorrelationslänge steigt bei Systemen in der Nähe eines kritischen Punkts proportional zu einer Potenz der Korrelationslänge ξ an:

$$z \;\propto\; \xi^{\gamma}\,.$$

γ definiert einen Exponenten, welcher vom Algorithmus abhängt. Für lokale Algorithmen kann man zeigen, dass $\gamma \geq 2$. Dies hat dazu geführt, nach nicht-lokalen Algorithmen zu suchen, z.B. so genannte Cluster-Algorithmen, bei denen Konfigurationen bei einem Monte-Carlo-Schritt innerhalb großer Gebiete verändert werden, oder so genannte Multi-Grid-Verfahren, bei denen man durch einen Trick das System am kritischen Punkt durch ein anderes (äquivalentes) System, welches nicht an einem kritischen Punkt ist, ersetzt. Für nicht-lokale Algorithmen kann man theoretisch sogar $\gamma = 0$ erreichen. Allerdings hängt z nicht nur vom Exponenten γ ab, sondern natürlich auch von dem Proportionalitätsfaktor, und dieser ist bei nicht-lokalen Algorithmen im Allgemeinen sehr groß. Außerdem ist es bis jetzt noch nicht gelungen, einen Cluster-Algorithmus für nicht-abelsche Eichtheorien zu finden. Daher sind lokale Algorithmen trotz ihrer großen Autokorrelationslängen immer noch die Standardverfahren.

14.4 Spezielle Realisationen des Markov-Prozesses

14.4.1 Das Metropolis-Verfahren

Die Wahl einer neuen Konfiguration C_{i+1} geschieht beim Metropolis Verfahren in zwei Schritten:

1. Zunächst wählt man — ausgehend von C_i — eine neue Konfiguration C' mit einer Wahrscheinlichkeit p_1, welche die Bedingung $p_1(C_i \to C') = p_1(C' \to C_i)$ erfüllt.

2. In einem zweiten Schritt entscheidet man mit einer Wahrscheinlichkeit $p_2(C_i \to C')$, ob C' akzeptiert werden soll oder nicht. Diese Wahrscheinlichkeit erfüllt detailed balance. Wird C' akzeptiert, so ist $C_{i+1} = C'$, andernfalls ist $C_{i+1} = C_i$. Die Gesamtwahrscheinlichkeit $P = p_2 p_1$ für die Akzeptanz von C' erfüllt detailed balance.

Zum ersten Schritt:
Ist z.B. φ der Wert des skalaren Feldes in Konfiguration C_i an der Stelle $\vec{m}$, so kann man ein neues Feld $\varphi'(\vec{m})$ (und damit eine neue Konfiguration C') dadurch erhalten, dass man zu φ eine um 0 symmetrisch verteilte Zufallszahl addiert. Die symmetrische Verteilung bedeutet $p_1(\varphi \to \varphi') = p_1(\varphi' \to \varphi)$.

Die genaue Verteilungsfunktion dieser Zufallszahl spielt theoretisch zwar keine Rolle, sie wird im konkreten Fall jedoch so gewählt, dass die Akzeptanz des neuen Feldes möglichst hoch ist, anderseits aber auch die Veränderung möglichst groß. Diese beiden Bedingungen stehen in einem gewissen Gegensatz, sodass man ein geeignetes Mittelmaß finden muss. Der Gegensatz besteht darin, dass eine kleine Veränderung im zweiten Schritt sehr oft akzeptiert wird, während eine große Veränderung meist nicht akzeptiert wird. Möchte man eine große Akzeptanzrate, so sollte die Verteilungsfunktion der Zufallszahlen also eine geringe Breite haben. Anderseits wird das Feld dadurch nicht wesentlich geändert, d.h., die Autokorrelationen sind sehr groß.

Zum zweiten Schritt:
Beim Metropolis-Verfahren berechnet man im zweiten Schritt das Verhältnis der Gewichte von C und C'. Die Bedingung der detailed balance

$$\frac{P(C \to C')}{P(C' \to C)} \;=\; \frac{w[C']}{w[C]} \;=\; \exp(S[C] - S[C'])$$

wird dadurch erfüllt, dass man C' mit einer Wahrscheinlichkeit

$$p_2(C \to C') \;=\; \begin{cases} 1 & \text{falls } S[C] \geq S[C'] \\ \exp(S[C] - S[C']) & \text{falls } S[C] < S[C'] \end{cases}$$

akzeptiert. Konkret wird das dadurch realisiert, dass man mit einem Zufallszahlgenerator eine im Intervall $[0, 1]$ gleichverteile Zufallszahl r bestimmt, und festlegt:

$$\begin{array}{lcll} C_{i+1} & = & C' & \text{falls} \quad \exp(S[C] - S[C']) \geq r \\ C_{i+1} & = & C_i & \text{sonst} \end{array} \quad .$$

Ist $S[C'] < S[C]$ (die neue Konfiguration also „energetisch günstiger“ als die alte), so wird sie immer akzeptiert. Anderenfalls ist die Akzeptanz zufällig: je ungünstiger die neue Konfiguration, desto unwahrscheinlicher die Akzeptanz.

14.4.2 Das „heat bath“– Verfahren

Eine offensichtliche Lösung der detailed balance Bedingung ist

$$P(C \to C') \;=\; w[C']\,.$$

Man bestimmt also die neue Konfiguration C' direkt nach der Boltzmann-Verteilung $w[C']$ (daher der Name „heat bath“), unabhängig von der vorherigen Konfiguration C.

Sei φ der Wert des Feldes am Punkte $\vec{m}$, für welches der „up-date“ gemacht werden soll, und $\{\varphi_k\}$ seien die Werte des Feldes an allen anderen Gitterpunkten (bzw. Gitterlinien). Dann definiert

$$\tilde{w}(\varphi) \;=\; \frac{\exp(-S[\varphi,\{\varphi_k\}])}{\int d\varphi \exp(-S[\varphi,\{\varphi_k\}])}$$

eine Verteilungsfunktion für φ und man erhält

$$P[(\varphi,\{\varphi_k\}) \to (\varphi',\{\varphi_k\})] \;=\; \tilde{w}(\varphi')\,.$$

Man wählt $C_{i+1} = (r,\{\varphi_k\})$, wobei r eine Zufallszahl ist, welche nach der Verteilungsfunktion $\tilde{w}(r)$ generiert wurde.

$\tilde{w}(r)$ ist ein Funktional der anderen Felder $\{\varphi_k\}$. Im Allgemeinen werden diese Verteilungen schwer zu realisieren sein, selbst wenn die Wechselwirkung lokal ist, d.h. $\tilde{w}$ nur von den Feldern abhängt, die dem Punkte $\vec{m}$ unmittelbar benachbart sind. Eine Ausnahme bilden diskrete lokale Zustände, da in diesem Fall nur endlich viele Verteilungsfunktionen $\tilde{w}$ für diskrete Variable auftreten können, und diese lassen sich vor Beginn der Simulation festlegen. Dies soll am Beispiel des d-dimensionalen Ising-Modells erläutert werden:

Sei s der Wert des Spins am Punkte i, dann ist

$$\tilde{w}(s;X) \;=\; \frac{1}{Z}\mathrm{e}^{\beta s X} \;=\; \begin{cases} \dfrac{1}{Z}\exp\beta X & \text{falls } \; s=1 \\[2ex] \dfrac{1}{Z}\exp -\beta X & \text{falls } \; s=-1 \end{cases}$$

$$\text{mit} \quad Z \;=\; \mathrm{e}^{\beta X} \;+\; \mathrm{e}^{-\beta X}\,.$$

Dabei ist

$$X \;=\; \sum_{\langle j\rangle|_i} s_j \;\;\in\; \{-2d,-2d+2,\ldots,2d-2,2d\}$$

die Summe der Spins auf den zu i benachbarten Gitterpunkten. Man hat also (bei fester Kopplung β) insgesamt $2d+1$ verschiedene Verteilungsfunktionen für eine Zufallszahl, welche nur die Werte ± 1 annimmt. Sei r eine gleichverteilte Zufallszahl im Intervall $[0,1]$, so ist die neue Spinvariable

$$s' \;=\; \begin{cases} +1 & \text{falls } \; r \le \tilde{w}(1;X) \\ -1 & \text{sonst } \; . \end{cases}$$

Die $2d+1$ Zahlen $\tilde{w}(1;-2d), \tilde{w}(1;-2d+2), \ldots, \tilde{w}(1;2d)$ lassen sich vor Beginn der Simulation berechnen, sodass zum eigentlichen up-date nur die Summe der nächsten Nachbarspins X sowie der Vergleich mit einer gleichverteilten Zufallszahl nötig ist.

14.4.3 Andere Verfahren

Auf der Suche nach möglichst schnellen und effizienten Algorithmen wurden in den letzten Jahren viele Verfahren entwickelt, die je nach Modell unterschiedliche Anwendung finden. Im Folgenden sollen nur die häufigsten dieser Algorithmen erwähnt werden:

1. Der Langevin-Algorithmus

   Sei $\eta_i(\tau)$ ein Gauß-verteiltes, nicht τ-korreliertes stochastisches Rauschen mit

   $$\langle \eta_i(\tau)\eta_j(\tau')\rangle \;=\; \delta_{ij}\,\delta(\tau-\tau')\,.$$

   Eine zeitabhängige, stochastische Variable $\varphi_i(\tau)$, die Lösung einer Langevin-Gleichung

   $$\frac{\partial\varphi_i(\tau)}{\partial\tau} \;=\; -\,\frac{\partial S[\{\varphi\}]}{\partial\varphi_i} + \eta_i(\tau)$$

   ist, hat im Zeitmittel eine Verteilung proportional zu $\mathrm{e}^{-S[\{\varphi\}]}$. Die Stärke des Rauschens ist bei statistischen Systemen proportional zur Temperatur, die hier jedoch auf 1 gesetzt wurde.

   Diskretisiert man diese Gleichung ($\Delta\tau \simeq$ ein Monte-Carlo-Schritt), so kann man eine $\partial S/\partial\varphi$-abhängige Verteilungsfunktion für einen up-date definieren, die die Konfigurationen mit stochastischen Schwankungen auf den Gleichgewichtszustand „zutreibt“. Bei dieser Diskretisierung der Zeitschritte macht man allerdings einen systematischen Fehler, der zu korrigieren ist.

2. Die Methode der Molekulardynamik

   Diese Methode ist der Langevin-Methode sehr ähnlich. Man denkt sich $S[\{\varphi\}]$ als ein „Potential“ für die Felder $\{\varphi\}$, und die Verteilungsfunktion entspricht einem kanonischen Ensemble, bei welchem der impulsabhängige Anteil der (deterministischen) Dynamik schon ausintegriert wurde. Eine Diskretisierung der Bewegungsgleichung

   $$\begin{aligned}\frac{\partial^2\varphi_i(\tau)}{\partial\tau^2} \quad &= \quad -\frac{\partial S[\{\varphi\}]}{\partial\varphi_i}\\ \frac{\partial^2\varphi_i(\tau)}{\partial\tau^2} \quad &\longrightarrow \quad [\varphi_i(\tau_{n+1}) + \varphi_i(\tau_{n-1}) - 2\varphi_i(\tau)]/(\Delta\tau)^2\end{aligned}$$

   definiert einen Algorithmus, welcher im „Zeitmittel“ eine Boltzmann-Verteilung der Konfigurationen ergibt.

   Die Temperatur dieser Boltzmann-Verteilung wird durch die Energie der Startkonfiguration bestimmt. Die Gesamtenergie bleibt bei obiger Bewegungsgleichung erhalten, verteilt sich jedoch im Mittel gleichmäßig auf alle Freiheitsgrade. Nach dem Gleichverteilungssatz der statistischen Mechanik ist die mittlere Energie pro Freiheitsgrad proportional zur Temperatur.

   Oft addiert man zu obiger deterministischer Bewegungsgleichung noch ein stochastisches weißes Rauschen, das in diesem Fall jedoch nicht mit der Temperatur des Systems zusammenhängen muss. Die Temperatur wird über die Gesamtenergie gesteuert, d.h. gegebenenfalls muss man alle Geschwindigkeiten um einen geeigneten Faktor skalieren, um wieder ein System bei einer bestimmten Temperatur zu simulieren.

3. Hybrid Monte-Carlo

   Die beiden vorherigen Algorithmen lassen sich als Spezialfälle einer allgemeineren stochastischen Gleichung auffassen. Das Verfahren, welches auf dieser Verallgemeinerung beruht, heißt „hybrid“.

   Leider wird der Ausdruck „hybrid“ auch noch für ein anderes Vorgehen benutzt, nämlich wenn man verschiedene Algorithmen (z.B. Langevin und Metropolis) mischt.

4. Cluster-Algorithmen

   Bei einem Cluster Algorithmus wird mit Hilfe eines Verfahrens, welches detailed balance garantiert, ein zusammenhängendes Gebiet (Cluster) auf dem Gitter bestimmt, in welchem dann die Konfiguration gleichzeitig geändert wird (beim Ising Modell z.B. werden alle Spins innerhalb des Clusters „geflippt“), sodass die Änderung der Wirkung nur von dem Rand des Clusters abhängt.

   Der große Vorteil dieser Algorithmen ist, dass die Konfigurationen in einem Schritt wesentlich verändert werden (in einem guten Algorithmus ist die Größe eines Clusters von der Ordnung 20-50% der gesamten Gittergröße) und damit sehr kleine Autokorrelationen erziehlt werden. Der Nachteil ist, dass die Bestimmung des Clusters sehr aufwendig ist, d.h. „ein Schritt“ sehr lange dauert.

5. Multi-Grid-Verfahren

   Diese Verfahren sind die derzeit vermutlich anspruchvollsten. Das d-dimensionale Gitter wird dabei um eine zusätzliche Richtung erweitert, wobei allerdings das Gitter in dieser Zusatz-Richtung nicht konstant ist, sondern schrittweise gröber wird. Durch die Erweiterung des Gitters um eine Dimension erreicht man, dass das System nicht mehr „kritisch“ ist. Die Vergröberung des Gitters erlaubt es, Methoden der Renormierungsgruppe (diese werden zum Teil in einem späteren Kapitel noch erläutert) anzuwenden, um die gemessenen Observable mit den zu bestimmenden Größen in Beziehung zu setzen.

Kapitel 15

Glueballs — Beispiel eines Massenspektrums

Wir hatten schon mehrmals erwähnt bzw. benutzt, dass das asymptotische Abfallverhalten von 2-Punkt Funktionen eine Aussage über die niedrigsten Eigenwerte des Hamilton-Operators macht. Dies soll in diesem Abschnitt am Beispiel des Glueball-Spektrums in der reinen Eichtheorie diskutiert werden.

15.1 Glueballs

Glueballs sind gebundene Zustände von Gluonen, den Austauschteilchen der Quantenchromodynamik. In einer nicht-abelschen Eichtheorie tragen die Austauschteilchen ebenfalls eine Ladung, sodass sie untereinander wechselwirken. Daher erwartet man, dass sich diese gebundenen Zustände auch ohne Quarks bilden. Glueballzustände werden im Allgemeinen durch drei Quantenzahlen charakterisiert: J^{PC}. Dabei bezeichnet J den Gesamtspin des Zustands (ganzzahlig, da die beteiligten Partner ganzzahligen Spin haben und der Drehimpuls ebenfalls ganzzahlig ist), P und C stehen für das Transformationsverhalten des Zustandes unter Paritätstransformation bzw. Ladungskonjugation. Man kennt zwar einige Kandidaten für Glueballs, allerdings wäre ein Vergleich der beobachteten Massenspektren mit Vorhersagen aus der Theorie wichtig.

Bei den folgenden Überlegungen beschränken wir uns auf eine reine Eichtheorie, d.h., wir vernachlässigen die Wechselwirkung zwischen Gluonen und Quarks. Dies kann nur eine Näherung darstellen, denn zu den Bindungskräften zwischen Gluonen trägt auch der Austausch von Quarks bei.

Sei $G(t)$ die zeitgeordnete 2-Punkt-Funktion zu einem Operator $\hat{F}$:

$$\begin{aligned} G(t) &= \langle\Omega|\mathcal{T}[\hat{F}(t)\hat{F}(0)]|\Omega\rangle \;=\; \langle\Omega|\hat{F}\;\exp \mathrm{i}Ht\;\hat{F}|\Omega\rangle \\ &= \sum_n |\langle\Omega|\hat{F}|n\rangle|^2\;\exp \mathrm{i}E_n t \qquad (t>0)\,. \end{aligned}$$

Dabei bezeichnet $|n\rangle$ einen vollständigen Satz von Zuständen, die Eigenzustände zum Hamilton-Operator mit Eigenwerten E_n sind. Für große euklidische Zeiten $\mathrm{i}T = t$ ergibt sich

$$G_E(T) \;\to\; |\langle\Omega|\hat{F}|\Omega\rangle|^2 \;+ \tag{15.1}$$

$$
\begin{aligned}
&+ \ |\langle\Omega|\hat{F}|1\rangle|^2 \mathrm{e}^{-E_1 T} \left(1 \ + \ \mathcal{O}(\mathrm{e}^{-T(E_2 - E_1)})\right) \\
= \quad & c_0 + c_1 \, \mathrm{e}^{-TE_1} + \ldots \, . \qquad (15.2)
\end{aligned}
$$

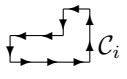
Diese Argumentation zeigt, dass der niedrigste Energieeigenwert über dem Grundzustand (dem Vakuum), für den das Matrixelement $\langle\Omega|\hat{F}|1\rangle$ ungleich Null ist, für das asymptotische Verhalten der 2-Punkt-Funktion verantwortlich ist. $\langle\Omega|\hat{F}|1\rangle \neq 0$ setzt voraus, dass der Zustand $\hat{F}|\Omega\rangle$ bezüglich aller Erhaltungsgrößen dieselben Quantenzahlen hat wie der Zustand $|1\rangle$. Durch geeignete Wahl des Operators $\hat{F}$ kann man so einen Sektor, der durch erhaltene Quantenzahlen charakterisiert ist, herausfiltern. Insbesondere können die Massen der stabilen Zustände mit den Quantenzahlen J, P und C bestimmt werden.

$\hat{F}$ sollte ein eichinvarianter Operator sein. Im Rahmen einer euklidischen Funktionalintegralquantisierung bzw. im Rahmen einer Gittereichtheorie ist $\hat{F}$ daher durch eine entsprechende Funktionen F von Wilson-Loops zu ersetzen. Sei F im Folgenden eine Summe von Wilson-Loops:

$$
F \ = \ \sum_i \alpha_i W[\mathcal{C}_i] \, . \qquad (15.3)
$$

Dies drückt man auch oft durch die symbolische Schreibweise

$$
F \ = \ \sum_i \alpha_i \ \mathcal{C}_i
$$



aus. Dann gilt

$$
\langle F(T)F(0)\rangle \ = \ \sum_i \sum_j \alpha_i \alpha_j \ \left\langle W[\mathcal{C}_i^T] \, W[\mathcal{C}_j]\right\rangle \, . \qquad (15.4)
$$

„$\mathcal{C}^T$“ bedeutet, dass der Wilson-Loop $\mathcal{C}$ um T Schritte in Zeitrichtung auf dem Gitter verschoben wird. Die Wilson-Loops selber können dabei auch Gitterlinien in zeitartiger Richtung besitzen. Oft wird jeder Loop noch über die gesamte raumartige Gitter(hyper)ebene verschoben, wodurch Zustände zum Impuls 0 heraus projiziert werden.

Der Paritätsoperator P sowie die Ladungskonjugation C haben nur ± 1 als möglichen Eigenwerte, sie lassen sich auf dem Gitter leicht konstruieren. Hingegen ergibt sich ein Problem bezüglich der Quantenzahlen, die den Darstellungen der Drehgruppe J entsprechen. Während die Kontinuumstheorie invariant unter der Lorentzgruppe ist, und damit Spin-Zustände im Ruhesystem, $\boldsymbol{p} = \mathbf{0}$, eines massiven Teilchens durch die Darstellungen der Drehgruppe SO(3) charakterisiert werden, hat das Gitter nur die verbleibende Invarianz unter der Gittersymmetrie. Für ein (hyper-) kubisches Gitter ist dies (im räumlichen 3-dimensionalen Gitter) die Untergruppe der Drehgruppe, die den Würfel invariant lässt. Dies ist die *oktaedrische* Gruppe. Sie hat 24 Elemente und wird oft mit O_{24} (oder auch einfach O) bezeichnet. Die irreduziblen, endlich dimensionalen Darstellungen der Drehgruppe lassen sich durch eine positive ganze Zahl l kennzeichnen; ihre Dimension ist $2l + 1$. Die Drehgruppe besitzt daher sehr viel mehr Darstellungen als diese endliche, diskrete Gruppe. Die irreduziblen Darstellungen der Gruppe SO(3) bilden im Allgemeinen für die Untergruppe O_{24} eine reduzible Darstellung.

Die 24 Elemente der Gruppe O_{24} zerfallen in fünf Klassen: Die Klasse $E(1)$ besteht aus dem Identitätselement. Die drei Elemente der Klasse $C_2(3)$ sind die Rotationen um 180° um die x-, y- bzw. z-Achse. Die Klasse $C_4(6)$ enthält 6 Elemente und zwar die Rotationen um $\pm 90°$ um die x-, y- bzw. z-Achse. Die Klasse

$C_2(6)$ enthält ebenfalls 6 Elemente, bestehend aus den Rotationen um 180° um Diagonale durch jeweils zwei gegenüberliegende Kanten des Würfels (d.h. um die 6 Diagonalen in der xy-, xz- und yz-Ebene). Die Klasse $C_3(8)$ schließlich enthält 8 Elemente der Ordnung 3 und zwar die Rotationen um $\pm 120°$ um die 4 Diagonalen durch jeweils gegenüberliegende Punkte des Würfels.

Die Gruppe O_{24} besitzt fünf verschiedene irreduzible Darstellungen, die manchmal mit A_1, A_2, E, F_1 und F_2 bezeichnet werden. A_1 und A_2 sind 1-dimensionale Darstellungen (A_1 die triviale Darstellung), E ist eine 2-dimensionale Darstellung und F_1, F_2 sind zwei nicht-isomorphe 3-dimensionale Darstellungen. Die Charaktere der Gruppe O_{24} zu den fünf Klassen und fünf irreduzieblen Darstellungen sind in Tabelle 15.1 zusammengefasst.

Tabelle 15.1: Charaktere χ der Gruppe O_{24}.

	E	$C_2(3)$	$C_4(6)$	$C_2(6)$	$C_3(8)$
A_1	1	1	1	1	1
A_2	1	1	−1	−1	1
E	2	2	0	0	−1
F_1	3	−1	−1	1	0
F_2	3	−1	1	−1	0

Unser nächstes Ziel besteht darin, aus einem gegebenen Wilson-Loop $W[\mathcal{C}]$ irreduzible Darstellungen zu P, C, O_{24} und Impuls $p = 0$ zu konstruieren. Sei $\mathcal{C}_{a,P,g}$ der Loop, den man aus $\mathcal{C}$ durch Translation um $a \in \mathbf{Z}_3$, Paritätstransformation $P \in \{\mathbb{1}, -\mathbb{1}\}$ und einer Gruppentransformation $g \in O_{24}$ erhält. Dann ist

$$F_{\Gamma^{PC}}[\mathcal{C}] \;=\; \sum_a \sum_{P,g} (-1)^P \chi_\Gamma(g)(W[\mathcal{C}_{a,P,g}] + (-1)^C W^+[\mathcal{C}_{a,P,g}])$$

ein Operator mit den Quantenzahlen Γ^{PC} zu verschwindendem Impuls. Hierbei bezeichnet $\Gamma = A_1, A_2, E, F_1, F_2$ die irreduziblen Darstellungen von O_{24}. Bei einfachen Loops $\mathcal{C}$ mit gewissen Symmetrieeigenschaften wird $F_{\Gamma^{PC}}[\mathcal{C}]$ für viele Darstellungen verschwinden. (Impulseigenzustände zum Impuls p erhielte man, indem man noch die Charaktere zur Translationsgruppe e^{ipa} einfügt.)

Die irreduziblen $2l$+1-dimensionalen Darstellungen der Gruppe SO(3) lassen sich nach Darstellungen von O_{24} wie folgt zerlegen (siehe z.B. [64, 66]):

$$\begin{aligned}
l = 0 &\;\to\; A_1 \\
1 &\;\to\; F_1 \\
2 &\;\to\; E + F_2 \\
3 &\;\to\; A_2 + F_1 + F_2 \\
4 &\;\to\; A_1 + E + F_1 + F_2 \\
5 &\;\to\; E + 2F_1 + F_2 \\
6 &\;\to\; A_1 + A_2 + E + F_1 + 2F_2 \\
l + 5 &\;\to\; A_1 + A_2 + 2E + 3F_1 + 3F_2 - [6 - l] \\
l + 12 &\;\to\; A_1 + A_2 + 2E + 3F_1 + 3F_2 + [l]
\end{aligned}$$

Hierbei bezeichnet $[6 - l]$ bzw. $[l]$ die Zerlegung der Darstellung von SO(3) zu den Werten $6 - l$ ($l = 0, ..., 6$) bzw. l. Umgekehrt können wir sagen, dass die Darstellung A_1 von O_{24} Anteile der irreduziblen Darstellungen von SO(3) zu $l = 0, 4, 6, \ldots$ enhält. Entsprechend hat F_1 Anteile zu $l = 1, 3, 4, \ldots$, E Anteile von $l = 2, 4, 5, \ldots$ und A_2 Anteile von $l = 3, 6, \ldots$. Will man also sicher sein, dass der tiefste Eigenwert zur Darstellung von F_1 beispielsweise auch zur Darstellung $l = 0$ gehört, so muss man prüfen, dass dieser Eigenwert nicht in der Darstellung A_2 auftritt, da der Zustand anderenfalls eher zu $l = 3$ gehört. Im Allgemeinen ist es schwer, Darstellungen zu $l > 3$ herauszukristallisieren. Im Folgenden ist $J = 0, 1, 2, 3$.

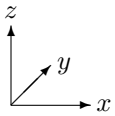
Streng genommen kann der Loop auch in „Zeit-Richtung“ orientiert sein, was dann auf die Darstellungstheorie der O_{24} als Untergruppe der 4-dimensionalen Würfelgruppe führt. Auf diese Details soll hier aber nicht weiter eingegangen werden, obwohl sie in der Praxis durchaus von Bedeutung sind.

Beispiele für Operatoren zu bestimmten Quantenzahlten J^{PC} sind die folgenden:

$F_{0^{++}}$ = Re + +

$F_{1^{--}}$ = Im

$F_{2^{++}}$ = Re −



15.2 „Verschmierte“ Zustände

Die Anwendbarkeit des obigen Verfahrens hängt bei numerischen Methoden von zwei weiteren Eigenschaften der Theorie ab. Einerseits sollte der nächst höhere Energieeigenwert (zu verschwindendem Impuls) E_2 viel größer sein als E_1, worauf man allerdings keinen Einfluss hat. Andererseits soll das Matrixelement $\langle\Omega|\hat{F}(0)|1\rangle$ möglichst groß sein. Diese zweite Bedingung lässt sich durch die Wahl geeigneter Operatoren $\hat{F}(0)$ erreichen.

Sei $F_{J^{PC}}$ eine Linearkombination von Operatoren zu festen Wilson-Loops:

$$F_{J^{PC}} = \sum_{\mathcal{C}} \alpha(\mathcal{C})\ F_{J^{PC}}(\mathcal{C}) \qquad\qquad \sum_{\mathcal{C}} |\alpha(\mathcal{C})|^2 = 1\,.$$

Man kann nun durch „probieren“ die $\{\alpha\}$'s variieren, bis der Koeffizient vor dem führenden Term (c_1 in Gleichung 15.2) maximal wird.

Mit einem vollständigen Satz von Operatoren zu festen Quantenzahlen J^{PC} (der allerdings auch Produkte von Wilson-Loops einschließt) kann man durch geeignete Wahl der Koeffizienten den Eigenzustand zu E_1 konstruieren. Kennt man den Eigenzustand zum niedrigsten Energieeigenwert E_1, so kann man durch Wahl eines Operators F' mit den gleichen Quantenzahlen, aber zu einem orthogonalen Zustand, auch die höheren Energieeigenwerte in dem entsprechenden Sektor bestimmen. Dieses Verfahren ist sehr aufwendig und bisher lediglich für den zweiten Energieeigenwert im 0^{++}-Sektor durchgeführt worden.

Abbildung 15.1 gibt das mit Monte Carlo Simulation (1989) gemessene Spektrum der J^{PC} Zustände der reinen SU(3)-Eichtheorie wieder. Die genauesten Massen sind für den 0^{++} und den 2^{++} Zustand bekannt. Unter der Annahme, dass die Stringtension den Wert $\sqrt{\kappa} = 440$ MeV hat (näheres dazu im Kapitel über den Kontinuumslimes), erhält man:

$$\begin{aligned} m_{0^{++}} &= 1370 \pm 90 \text{ MeV} \\ m_{2^{++}} &= 2115 \pm 125 \text{ MeV}\,. \end{aligned}$$

Näheres findet man auch in [73].

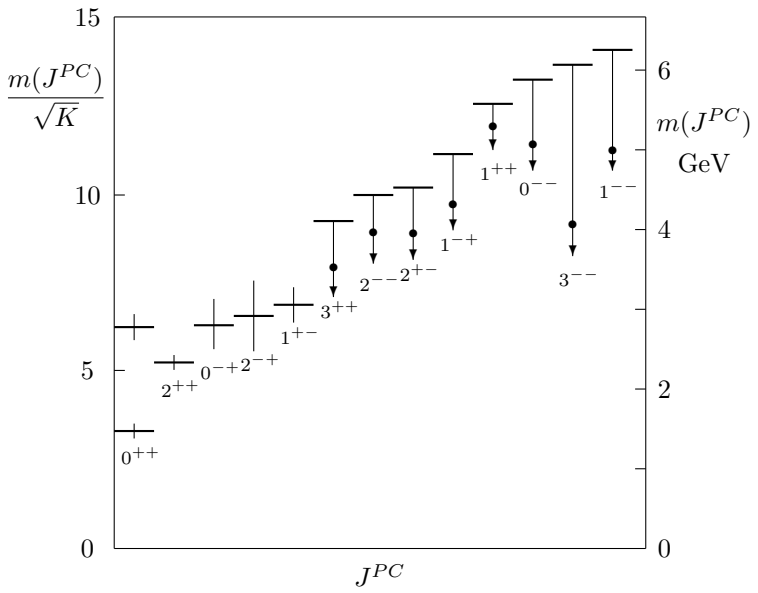


Abbildung 15.1: Massenspektrum einer reinen SU(3)-Eichtheorie (aus [73]).

Kapitel 16

Der Kontinuumslimes

Letztendlich möchten wir aus den Gittermodellen Vorhersagen über experimentell bestimmbare, physikalische Größen gewinnen. Wir wollen also Größen, die sich aus den Erwartungswerten der Gittertheorie ergeben, mit experimentellen Messgrößen vergleichen. Nun hängt die Gittertheorie aber von einem Satz von Parametern $\{g_i\}_{i=1,\ldots,N}$ ab, die zunächst keine physikalische Bedeutung haben. Im Gegensatz zur statistischen Mechanik der Festkörper, wo die Parameter $\{g_i\}$ in einem Energie- bzw. Wirkungsfunktional zumindest theoretisch aus der zugrundeliegenden Quantenmechanik der Atome bzw. Moleküle und ihren Wechselwirkungen ableitbar sind, gibt es hier keine zugrundeliegende Mikrotheorie. Wir müssen daher zunächst die Parameter $\{g_i\}$ durch physikalisch beobachtbare Größen ausdrücken. Da wir bis heute noch keine experimentelle Bestätigung einer diskreten Raum-Zeit haben (auch wenn im Rahmen einer Quantengravitation eine kontinuierliche Raum-Zeit im herkömmlichen Sinne aus Konsistenzgründen vermutlich auf Widersprüche führt), beziehen sich die beobachteten Größen auf Längenskalen, die sehr viel größer als der Gitterabstand sind. Unsere heutigen Energien reichen nicht aus, eine eventuell vorhandene Gitterstruktur aufzulösen. „Kontinuumslimes“ bedeutet nun, die Parameter der Gittertheorie so zu wählen, dass die Korrelationslängen im Vergleich zum Gitterabstand groß werden. Im Fall der reinen Eichtheorie ist das beispielsweise für $g \to 0$ der Fall. Die genaue Form, in der dieses Adjustieren der Parameter der Theorie vorgenommen werden muss, soll zunächst für den allgemeinen Fall geklärt werden.

16.1 Allgemeiner Formalismus des Kontinuumslimes

Wir nehmen an, dass die Wirkung der Gittertheorie von einen Satz von Parametern $\{g_i\}$ abhängt. Dies sind natürlich alles dimensionslose Zahlen. Der erste Schritt besteht darin, das Phasendiagramm der Theorie zu ermitteln. Im Kontinuumlimes müssen die Korrelationslängen der Theorie divergieren. Dies geschieht z.B. bei Phasenübergängen 2. Ordnung, allerdings auch bei manchen Randwerten (z.B. bei der reinen Eichtheorie für $g \to 0$ oder im 1-dimensionalen Ising-Modell für $\beta \to \infty (T \to 0)$, solche kritischen Punkte entsprechen im üblichen Sprachgebrauch keinen Phasenübergängen). In vielen Theorien gibt es ganze Hyperflächen kritischer Punkte, man denke z.B. an die kritische Linie (C-B) im Phasendiagramm einiger Eichtheo-

rien mit Materiefeldern (Kap. 13.1, Seite 125). Auf solchen kritischen Linien divergieren die Korrelationslängen von manchen Operatoren, während die Korrelationslängen anderer Operatoren endlich bleiben. Auf der Linie (C-B) divergieren beispielsweise die Korrelationslängen zu den Operatoren des Materiefeldes, wohingegen Korrelationslängen zu 2-Punkt-Funktionen von Wilson-Loops (die inversen Glueballmassen) im Allgemeinen endlich bleiben.

Die Korrelationslängen zu Operatoren, welche im Kontinuumslimes zu physikalischen Feldern werden sollen, müssen an dem kritischen Punkt, der den Kontinuumslimes definieren soll, divergieren. Im Allgemeinen geschieht das im Phasendiagramm nur an einem einzigen Punkt, der im Folgenden durch $\{g_i^c\}$ bezeichnet wird. In dem typischen Phasendiagramm einer Eichtheorie mit Materiefeldern (Diagramm auf Seite 125) ist dies der Punkt B. Dort lassen sich im Kontinuumslimes sowohl endliche Glueballmassen als auch endliche Massen für die Materiefelder definieren.

Analog zu unserer heuristischen Diskussion des Kontinuumslimes beim freien Klein-Gordon-Feld unterscheiden wir nun drei verschiedene Sätze von Größen:

- Der Satz von Parametern $\{g_i\}_{i=1,2,\ldots,N}$ im Energie- bzw. Wirkungsfunktional der Gittertheorie: Es handelt sich um dimensionslose Größen. Insbesondere sind diese Größen im Allgemeinen keine Observable der Gittertheorie (d.h. darstellbar als Erwartungswerte von Funktionen der Konfigurationsmenge).

- Ein „vollständiger und unabhängiger“ Satz von Größen, die sich aus den Erwartungswerten der Gittertheorie bestimmen lassen - also Observable sind - und direkt mit physikalischen Größen in Verbindung gebracht werden können. Diesen bezeichnen wir im Folgenden mit $\{\hat{M}, \mu_i\}_{i=2,3,\ldots,N}$.

  „Vollständig und unabhängig“ heißt hierbei, dass eine Vorgabe dieser Observablen $\{\hat{M}, \mu_i\}$ eindeutig einen Satz von Parametern $\{g_i\}$ festlegt (zumindest in der Nähe des kritischen Punktes $\{g_i^c\}$). Man denke beispielsweise an einen Satz von (inversen) Korrelationslängen zu verschiedenen Operatoren, oder an die Stringtension, oder auch an Kopplungskonstanten (z.B. die Vier-Vertex-Kopplung: $\tilde{g} = \frac{\langle\varphi^4\rangle - 3\langle\varphi^2\rangle^2}{\langle\varphi^2\rangle^2}$).

  „Direkt mit physikalischen Größen in Verbindung stehen“ bedeutet, dass diesen Observablen bestimmte experimentell messbare physikalische Größen entsprechen. So entsprechen beispielsweise die inversen Korrelationslängen den Massen von Teilchen, die Stringtension κ lässt sich aus dem $q\bar{q}$-Spektrum (zu schweren Quarks) bestimmen, und Kopplungskonstanten können mit Streuquerschnitten bei bestimmen Impulsüberträgen in Verbindung gebracht werden.

  Mindestens eine dieser Größen, die sich aus den Erwartungswerten in der Gittertheorie bestimmen lassen, sollte als physikalische Größe eine dimensionsbehaftete Größe gegenüberstehen, beispielsweise eine Masse. Gibt es mehrere dimensionsbehaftete Größen, so können wir geeignete Verhältnisse wählen, die dimensionslos sind. Wir fordern daher (ohne Einschränkung der Allgemeinheit) im Folgenden, dass der Größe $\hat{M}$ als physikalisch messbare Größe eine Masse entspricht, wohingegen die anderen Größen $\{\mu_i\}_{i=2,3,\ldots,N}$ dimensionslos sein sollen.

- Als dritten Satz von Größen haben wir die physikalisch (experimentell) messbaren Größen $\{m, \tilde{\mu}_i\}$, die den obengenannten Gittergrößen gegenüberstehen.

Jedem Punkt im Phasenraum $\{g_i\}$ entspricht ein Satz von Gittergrößen $\hat{M}[\{g_i\}]$ und $\mu_i[\{g_i\}]$. Umgekehrt kann man nach Voraussetzung (in einer Umgebung von $\{g_i^c\}$) die Parameter $\{g_i\}$ auch als Funktionen des einmal festgelegten Satzes von messbaren Gittergrößen auffassen:

$$g_i \;=\; g_i[\hat{M}, \{\mu_j\}]\,.$$

Die Kurven (parametrisiert durch $\hat{M}$) im Parameterraum der Theorie, die man für feste Werte von $\{\mu_j = \tilde{\mu}_i\}$ erhält, werden manchmal "lines of constant physics" genannt. Man wählt also die Parameter der Theorie $\{g_i\}$ so, dass die dimensionslosen Größen $\{\mu_i\}$ bereits den experimentellen Messwerten $\{\tilde{\mu}_i\}$ entsprechen. Für $\hat{M} \to 0$ enden die Linien „konstanter Physik" im kritischen Punkt $\{g_i^c\}$.

Durch die Identifikation der Gittergröße $\hat{M}$ mit der physikalischen Größe m wird dem Gitterabstand eine Skala $a = \hat{M}/m$ zugewiesen. Wir können somit auch die Parameter $\{g_i\}$ als Funktionen des Gitterabstands a, des physikalischen *Renormierungspunktes* m, sowie der Vorgabe der dimensionslosen Verhältnisse $\mu_i = \tilde{\mu}_i$ auffassen:

$$g_i \;=\; g_i[am, \{\tilde{\mu}_j\}]\,.$$

In dieser Form hängen die Parameter der Gittertheorie nur von physikalischen Größen der Kontinuumstheorie ab, welche man zur Festlegung der Theorie benutzt, sowie von dem willkürlichen Parameter a.

Kontinuumslimes bedeutet der Grenzfall $a \to 0$. Dies ist nicht zu verwechseln mit $a = 0$. In diesem Fall ist nämlich $\{g_i\} = \{g_i^c\}$ und alle Korrelationslängen sind unendlich, bzw. alle „Massen" verschwinden. $a \to 0$ bedeutet, dass dimensionslose Verhältnisse anderer physikalischer Größen, die nicht zur Festlegung der Parameter benutzt wurden, für genügend kleine Werte von a nicht mehr vom Gitterabstand abhängen. Für $a \to 0$ wird die Theorie somit invariant unter Skalentransformationen $a \to \lambda a$, daher nennt man dies auch den *Skalenlimes* der Theorie. Eine Transformation

$$a \to \lambda a \qquad g_i \to g_i^\lambda \;=\; g_i[\lambda am, \{\mu_j\}] \qquad (\mu_j = \text{const.})$$

heißt *Renormierungsgruppentransformation*. Diese Gruppe ist im strengen Sinne jedoch keine Invarianzgruppe der Theorie, da sich „die Physik" nur im Grenzfall $a \to 0$ nicht ändert.

In Monte-Carlo-Simulationen ist das Gitter endlich. Soll die größte Korrelationslänge noch erheblich kleiner als die lineare Ausdehnung des Systems sein, so kann man sich dem kritischen Punkt in einem endlichen Volumen nicht beliebig nähern. Andererseits darf die kürzeste Korrelationslänge, von der man noch erhoffen kann, dass die zugehörigen Operatoren noch etwas mit „Kontinuumsphysik" zu tun haben, nicht von der Größenordnung des Gitterabstandes werden. Daher spricht man bei endlichen Gittern auch manchmal von einem „physikalischen Fenster".

Werden die Ergebnisse von Gittersimulationen publiziert, ist es oft umständlich, die genauen Werte der Kopplungskonstanten $\{g_i\}$ anzugeben, bei denen die Größen bestimmt wurden. Statt dessen sagt man, welche physikalischen Größen angepasst wurden (das sind keine Vorhersagen sondern gefittete Größen) und man gibt den Gitterabstand a an, der den gewählten Parametern der Simulation entspricht. Auf diese Weise lassen sich auch Simulationen mit unterschiedlichen Parametern besser vergleichen. Typische Lattice Berechnungen beispielsweise von Mesonspektren verwenden Parameter, die einem Gitterabstand von $a \approx 0,1 - 0,2\,\text{fm}$ entsprechen. Das bedeutet, eine Masse von beispielsweise 600 MeV entspricht einer Korrelationslänge von 10-20

Gittereinheiten. Umgekehrt sollte die Gitterlänge zur Vermeidung von finite size Effekten größer als die doppelte Korrelationslïage der Teilchen sein. Auch die Größe des Gitters wird manchmal in physikalischen Einheiten von Femtometern angegeben.

16.2 Die β-Funktion

Die Tatsache, dass das angegebene Programm für den Kontinuumslimes überhaupt durchführbar ist, hängt von einigen höchst nicht-trivialen Eigenschaften der Gittertheorien ab. Es setzt nämlich voraus, dass alle Größen, die in der physikalischen Kontinuumstheorie dieselbe Skalendimension haben, in der gleichen Form divergieren, bzw. gegen 0 gehen im Grenzfall $\{g_i\} \to \{g_i^{\mathrm{c}}\}$. Die folgenden Überlegungen werden der Einfachheit halber in einer Theorie mit nur einem Parameter g (Beispiel: reine Eichtheorie) mit kritischem Punkt $g = 0$ erläutert, lassen sich aber ohne große Schwierigkeiten auf den allgemeinen Fall (entlang der Linien „konstanter Physik“) übertragen.

Wir betrachten den Fall, dass in der Theorie zwei Größen $\hat{M}_1$ und $\hat{M}_2$ definiert werden können, deren physikalische Gegenstücke die Dimension von Massen (z.B. inverse Korrelationslängen von zwei Operatoren) haben. In einer reinen Eichtheorie kann es sich beispielsweise um die inversen Korrelationslängen der 0^{++} und 2^{++} Zustände handeln, oder aber eine diese Größen und die Wurzel aus der Stringtension $\sqrt{\kappa}$. Im Grenzfall $g \to 0$ gehen beide Größen gegen 0, allerdings erwartet man, dass ihr Verhältnis gegen eine Konstante geht:

$$\hat{M}_i(g) \stackrel{g\to 0}{\longrightarrow} 0 \qquad\qquad \frac{\hat{M}_1(g)}{\hat{M}_2(g)} \stackrel{g\to 0}{\longrightarrow} \text{const.} = c\,.$$

Diese Konstante wird durch die Gittertheorie vorhergesagt. An ihr lässt sich durch die Wahl des Parameters g nichts ändern. Identifiziert man $\hat{M}_i$ mit physikalischen Massen m_i, so besagt die Theorie, dass $m_1/m_2 = c$. Das setzt aber voraus, dass beide Größen in der Gittertheorie dieselben kritischen Exponenten haben, im einfachsten Fall also

$$\hat{M}_1 \propto g^{\beta_1} \qquad \hat{M}_2 \propto g^{\beta_2} \qquad\qquad \beta_1 = \beta_2\,.$$

Es gibt aber im Allgemeinen keinen Grund, warum zwei Gittergrößen, für die die entsprechenden Kontinuumsgrößen dieselbe Skalendimension haben, im Limes $g \to 0$ dieselben kritischen Exponenten besitzen sollen. Angenommen $\beta_1 > \beta_2$. Dann wird im Kontinuumslimes $\hat{M}_1$ schneller gegen 0 gehen als $\hat{M}_2$. Man hat nun die Wahl, den Kontinuumslimes entweder durch $\hat{M}_1$ zu definieren, dann wird allerdings die zu $\hat{M}_2$ gehörige Masse unendlich, d.h. vom Spektrum verschwinden. Oder man definiert den Kontinuumslimes durch $\hat{M}_2$. Das führt zu der Vorhersage $m_1 = 0$. Eine Theorie, in der beide Massen endlich bleiben, gibt es nur falls $\beta_1 = \beta_2$.

Solche „No-Go“-Theoreme für den Kontinuumslimes von Gittertheorien haben oft interessante physikalische Konsequenzen. Beispiele sind skalare Feldtheorien in mehr als 4 Raum-Zeit Dimensionen. Hält man im Kontinuumslimes der Gittertheorie die Massen fest, so verschwinden die physikalischen Kopplungen, d.h., die Theorien werden zu freien Feldtheorien. Ein solches Verhalten wird man im Rahmen einer Störungstheorie nicht feststellen – im Gegenteil: dort gelten diese Theorien als nicht-renormierbar.

Ein weiteres interessantes Beispiel bilden skalare Theorien mit Symmetriebrechung (Beispiel: der Higgs-Sektor des Standardmodells). Hier stellt sich heraus, dass die Masse des Teilchens und der Erwartungswert (bzw. die Kopplung) ebenfalls nicht beliebig vorgegeben werden können. Dies erlaubt die Vorhersage einer so genannten „Trivialitätsschranke“ an die Higgsmasse. Die Theorie mit Masse und Erwartungswert gibt es nur

bei endlichem Gitterabstand. Die Schranke folgt aus der Forderung, dass der Gitterabstand kleiner sein muss als die inverse Masse.

Man sollte jedoch nicht vergessen, dass diese Überlegungen aus der gitterregularisierten Feldtheorie folgen. Es ist nicht ausgeschlossen, dass andere Regularisierungsverfahren zu konsistenten Kontinuumstheorien führen, allerdings sind bis heute keine Beispiele dafür bekannt.

Die Größen $\hat{M}_i$ müssen im Allgemeinen nicht mit einer Potenz von g im Grenzfall $g \to 0$ verschwinden. Allgemein kann man jedoch sagen, dass es eine Funktion $f(g)$ gibt, sodass

$$\hat{M}_i \;\propto\; f(g)(c_i + \mathcal{O}(g)) \qquad\qquad f(g) \;\stackrel{g\to 0}{\longrightarrow}\; 0\,.$$

Das führende Verhalten für $g \to 0$ dieser Funktion ist universell in einer Theorie. Jede Größe κ_i mit der Skalendimension d_i verhält sich im Kontinuumslimes wie

$$\kappa_i \;\propto\; f(g)^{-d_i}\,.$$

Es gibt somit eine physikalische Größe Λ mit der Dimension einer Masse, sodass

$$a\Lambda \;=\; f(g) \qquad \text{oder} \qquad a \;=\; \frac{1}{\Lambda} f(g)\,. \tag{16.1}$$

Diese Gleichung definiert die Beziehung zwischen der Kopplungskonstanten g und der Gitterkonstanten a. Wir definieren die *β-Funktion* durch:

$$\beta(g) \;=\; -\,a\frac{\partial g}{\partial a} \quad \Rightarrow \quad \beta(g) \;=\; -\,f(g)\left(\frac{\partial f}{\partial g}\right)^{-1} \tag{16.2}$$

(Im Allgemeinen Fall gibt es zu jedem Parameter g_i eine eigene β-Funktion.)

In der reinen $SU(N)$-Eichtheorie gilt

$$f(g) \;=\; \mathrm{e}^{-\frac{1}{2\beta_0 g^2}} \;\left(\beta_0 g^2\right)^{-\frac{\beta_1}{2\beta_0^2}} \;(1 + \mathcal{O}(g))$$

mit

$$\beta_0 \;=\; \frac{1}{16\pi^2}\left(\frac{11N}{3}\right) \qquad \beta_1 \;=\; \frac{1}{(16\pi^2)^2}\left(\frac{34N^2}{3}\right).$$

Die Koeffizienten der höheren Terme in g sind nicht universell. Sie hängen von der Wahl von Λ ab. Universell sind nur die Koeffizienten β_0 und β_1, die das führende Verhalten für $g \to 0$ bestimmen. Für die β–Funktion folgt

$$\beta(g) \;=\; -\,\frac{\beta_0 g^3}{1 - \frac{\beta_1}{\beta_0} g^2} \;\approx\; -\,\beta_0 g^3 - \beta_1 g^5 + \;\ldots\,.$$

Die β-Funktion ist auch in der herkömmlichen Störungstheorie einer Quantenfeldtheorie bestimmbar. Damit hat man ein geeignetes Verfahren, Gitterregularisierung und Störungstheorie miteinander zu vergleichen. Umgekehrt kann man auch prüfen, ob die Kopplungskonstanten der Gittertheorie nahe genug am kritischen Punkt sind und das richtige Skalenverhalten der Kontinuumstheorie (bestimmt durch die Störungstheorie) zeigen.

Aus der Kenntnis der β–Funktion lässt sich natürlich $f(g)$ durch Integration von Gl.(16.2) zurückgewinnen:

$$f(g) \;=\; \exp\left(\int_{g_0}^{g} \frac{1}{\beta(g')}\mathrm{d}g'\right)\; f(g_0)\,.$$

Kennt man von der β–Funktion eine Entwicklung nach Potenzen der Kopplung

$$\beta(g) \;=\; g^n \sum_{k=0}^{\infty} \beta_k g^k\,,$$

so folgt

$$\begin{aligned}\int_{g_0}^{g} \frac{1}{\beta(g')}\mathrm{d}g' \;&=\; \int_{g_0}^{g} \frac{1}{g'^n}\left(\sum_{k=0}^{\infty}\gamma_k g'^k\right)\mathrm{d}g' \\ &=\; -\frac{1}{g^{n-1}}\left(\sum_{k=0}^{n-2}\frac{\gamma_k g^k}{(n-1-k)}\right) \;+\; \gamma_{n-1}\ln g \;+\; \text{const.} \;+\; \mathcal{O}(g)\,,\end{aligned}$$

wobei die Koeffizienten sich durch formales Invertieren der Potenzreihe von $\beta(g)/g^n$ ergeben:

$$\left(\sum_k \beta_k g^k\right)\left(\sum_l \gamma_l g^l\right) \;=\; 1\,,$$

d.h. iterativ durch die Bedingungen:

$$\gamma_0 \;=\; \frac{1}{\beta_0} \qquad\qquad \gamma_l \;=\; -\frac{1}{\beta_0}\left(\sum_{k=0}^{l-1}\beta_{l-k}\gamma_k\right)\,.$$

Für die Skalenfunktion $f(g)$ folgt:

$$f(g) \;=\; \exp\left(-\sum_{k=0}^{n-2}\frac{\gamma_k}{(n-1-k)g^{n-1-k}}\right)\; (g)^{\gamma_{n-1}}\;(\text{const.} \;+\mathcal{O}(g))\,.$$

Da das Skalenverhalten für $g \to 0$ der Funktion $f(g)$ nicht vom Renormierungsverfahren (d.h. der Wahl der Größen (M, μ_i)) abhängt, müssen die Koeffizienten $\gamma_0, \ldots, \gamma_{n-1}$ universell sein. Dies impliziert entsprechend die Universalität der Koeffizienten $\beta_0, \ldots, \beta_{n-1}$. Wir erhalten also das

Theorem: Hängen die physikalischen Eigenschaften einer Gittertheorie im Kontinuumslimes nicht von der Wahl der Renormierungsparameter ab, und beginnt die β-Funktion in einer Entwicklung für Kopplungen nahe am kritischen Punkt mit der n.ten Potenz (d.h. mit $\beta_0(g - g^{\mathrm{c}})^n$, so sind die ersten n Koeffizienten der β-Funktion in dieser Entwicklung $(\beta_0, \ldots, \beta_{n-1})$ universell. Alle höheren Koeffizienten können vom Renormierungsverfahren abhängen.

In einer $SU(N)$ Eichtheorie beginnt die β-Funktion mit g^3, d.h., eigentlich sind die ersten 3 Koeffizienten universell. Da man aber im Rahmen einer Störungstheorie zeigen kann, dass die

geraden Potenzen in der β–Funktion verschwinden, erhält man nur zwei universelle Koeffizienten.

Im Allgemeinen verschwinden inverse Korrelationslängen am kritischen Punkt mit einer Potenz (dem kritischen Exponenten), in diesem Fall beginnt die β-Funktion linear in $(g - g^{\mathrm{c}})$, und nur der führende Koeffizient (der inverse kritische Exponent) ist universell.

Kapitel 17

Fermionen auf dem Gitter - Teil I

Wir wollen uns nun der Formulierung von Fermionen auf dem Gitter zuwenden. Mehr noch als schon bei den vergangenen Kapiteln wird es sich um eine elementare Darstellung einiger Grundbegriffe bzw. Stichworte handeln. Einige der Ideen können nur kurz angerissen werden. Für eine vertiefende Darstellung sei auf die Literatur verwiesen.

In Kapitel 8 hatten wir gesehen, dass es relativ problemlos ist, eine Theorie skalarer Felder auf dem Gitter zu formulieren. Die naheliegendste Idee – die Ersetzung der Ableitung durch eine Differenz der Felder an benachbarten Gitterpunkten – definiert eine Gittertheorie, die im Kontinuumslimes die Eigenschaften einer skalaren Quantenfeldtheorie hat. Für Spin-1 Eichfelder führt die – geometrisch naheliegende – Ersetzung der kovarianten Ableitung durch den Paralleltransporter, sowie die Ersetzung der Krümmung durch den Paralleltransport um elementare Plaketten, zu einer Eichtheorie auf dem Gitter, die ebenfalls im Kontinuumslimes (soweit bekannt) zur Quantenfeldtheorie des Eichsektors wird.

Für Fermionen (Spin-1/2) jedoch führt die „naheliegende“ Gitterregularisierung, nämlich die Ersetzung der Ableitung im Dirac-Operator durch eine symmetrisierte Differenz, nicht zu der bekannten Quantenfeldtheorie von Fermionen. Im Jahre 1982 bewiesen Nielsen und Ninomiya ein No-Go Theorem und für lange Zeit glaubte man, dass man schon sehr „abwegige“ Gitterformulierungen für Fermionen benötigt, um das so genannte Problem der „Fermionen-Verdopplung“ zu umgehen. Eng damit verbunden war der Glaube an die Unmöglichkeit, chirale Fermionen auf dem Gitter zu formulieren. Zu Beginn der neunziger Jahre wurde jedoch immer deutlicher, dass es durchaus akzeptable Verfahren gibt, chirale Fermionen auf dem Gitter zu formulieren. Hauptsächlich durch die bahnbrechenden Arbeiten von Lüscher [68, 69] wurden gegen Ende der 90er Jahre die grundlegenden Zusammenhänge geklärt.

Bevor wir im folgenden Kapitel zunächst auf das Theorem von Nielsen und Ninomiya und anschließend auf die gängigen Methoden, dieses Theorem zu umgehen, eingehen, soll kurz die Funktionalintegraldarstellung von Fermionen im euklidischen Kontinuum dargestellt weden.

17.1 Die Grassmann-Algebra

Für Fermionen kann die Funktionalintegraldarstellung nicht mit gewöhnlichen Feldern realisiert werden. Der Grund ist offensichtlich: Erwartungswerte fermionischer Felder müssen aufgrund

der fermionischen Statistik antisymmetrisch sein, beispielsweise:

$$\langle\bar{\psi}(x_1)\bar{\psi}(x_2)\psi(x_3)\psi(x_4)\rangle = -\langle\bar{\psi}(x_1)\bar{\psi}(x_2)\psi(x_4)\psi(x_3)\rangle\,. \tag{17.1}$$

Andererseits sind die Korrelationsfunktionen einer Wahrscheinlichkeitsverteilung für reelle oder komplexe Felder immer symmetrisch. In zwei Dimensionen lassen sich fermionische Felder durch Spin-Variable formulieren; dies funktioniert aber nicht mehr in höheren Dimensionen. Daher behilft man sich mit einem Trick.

Die Funktionalintegraldarstellung einer Quantenfeldtheorie für Fermionen geschieht mit Hilfe von antikommutierenden Variablen, so genannten Grassmann-Variablen. Daher soll auf die Algebra dieser Variablen kurz eingegangen werden.

Eine Grassmann-Algebra ist eine komplexe Algebra, die durch Objekte η_i ($i \in I \simeq$ Indexmenge) generiert wird, und in der die folgenden algebraischen Relationen gelten:

$$\{\eta_i, \eta_j\} \;\equiv\; \eta_i\eta_j + \eta_j\eta_i \;=\; 0\,.$$

Dies impliziert insbesondere: $\eta_i^2 = 0$. Ist die Indexmenge I endlich (hat man z.B. N Generatoren), so bildet diese Algebra einen 2^N-dimensionalen komplexen Vektorraum mit möglicher Basis:

$$|\epsilon_1, \epsilon_2, \ldots, \epsilon_N\rangle \;=\; \eta_1^{\epsilon_1}\eta_2^{\epsilon_2}\cdots\eta_N^{\epsilon_N} \qquad\qquad \epsilon_i \in \{0,1\}\,.$$

Wir definieren zwei lineare Operationen auf dieser Algebra, die Ableitung sowie die Integration:

$$\begin{aligned} \frac{\partial}{\partial\eta}1 &= 0 \qquad & \frac{\partial}{\partial\eta}\eta &= 1 \\ \int \mathrm{d}\eta 1 &= 0 \qquad & \int \mathrm{d}\eta\;\eta &= 1\,. \end{aligned} \tag{17.2}$$

Integral und Ableitung sind also identische Operationen auf der Algebra. Wir unterscheiden sie nur, um gewisse Analogien zu der üblichen Algebra von Funktionen zu verdeutlichen. Bei einem Produkt von mehreren Generatoren gilt:

$$\frac{\partial}{\partial\eta_i}\;\eta_1^{\epsilon_1}\cdots\eta_i^{\epsilon_i}\cdots\eta_N^{\epsilon_N} \;=\; \left[(-1)^{\sum_{j<i}\epsilon_j}\right]\;\epsilon_i\;\eta_1^{\epsilon_1}\cdots\eta_i^{\epsilon_i-1}\cdots\eta_N^{\epsilon_N}\,.$$

Entprechend für die Integration. Man muss also das entsprechende Element η_i zunächst nach links durchkommutieren, bevor die Ableitung bzw. das Integral genommen wird.

Ableitung und Integration antikommutieren ebenfalls:

$$\frac{\partial}{\partial\eta_i}\frac{\partial}{\partial\eta_j} \;+\; \frac{\partial}{\partial\eta_j}\frac{\partial}{\partial\eta_i} \;=\; 0 \qquad\qquad \mathrm{d}\eta_i\mathrm{d}\eta_j \;+\; \mathrm{d}\eta_j\mathrm{d}\eta_i \;=\; 0\,. \tag{17.3}$$

Bei einer Grassmann-Algebra mit einer geraden Anzahl von Generatoren können wir eine Paarung vornehmen und jeweils zwei Generatoren zu einem Paar $(\bar{\eta}_i, \eta_i), (i = 1, \ldots, N/2)$ zusammenfassen. Dies erlaubt die Definition einer "komplexen Konjugation", d.h. einer idempotenten *-Abbildung. Im Folgenden betrachten wir ausschließlich solche „komplexen" Grassmann-Variable.

17.2 Euklidisches Funktionalintegral für Fermionen

Warum gerade Grassmann-Variable in der Funktionalintegraldarstellung von fermionischen Quantensystemen auftreten, ist eine oft diskutierte Frage, auf die es nur unbefriedigende Antworten gibt. Man kann Grassmann-Variable als den formalen Limes $\hbar \to 0$ der Algebra von Fermion-Operatoren auffassen. Ableitung und Multiplikation von Grassmann-Variablen bilden eine Darstellung der fermionischen Operator-Algebra

$$\left\{ \eta_i, \frac{\partial}{\partial \eta_j} \right\} \;=\; \delta_{ij} \,,$$

ähnlich wie x und $\frac{\partial}{\partial x}$ im Schrödinger-Bild eine Darstellung der kanonischen Vertauschungsrelationen bilden. Die beste „Erklärung“ scheint mir die a posteriori Feststellung, dass die Kombinatorik der Störungstheorie, die sich aus Funktionalintegralen mit Grassmann-Variablen ergibt, mit derjenigen übereinstimmt, die man in einem Operator-Formalismus erhält. Beide Verfahren führen also zum selben Ergebnis.

Wir postulieren daher das Funktionalintegral für Fermionen

$$Z \;=\; \int \mathcal{D}\bar{\psi}(x)\mathcal{D}\psi(x)\; \mathrm{e}^{\mathrm{i}S[\bar{\psi},\psi]}$$

sowie die zeitgeordneten Vakuums–Erwartungswerte

$$\begin{aligned} S(x_1,\ldots,y_n) \;&=\; \langle\Omega|\bar{\psi}(x_1)\cdots\bar{\psi}(x_n)\psi(y_1)\ldots\psi(y_n)|\Omega\rangle \\ &=\; \frac{1}{Z}\int \mathcal{D}\bar{\psi}(x)\mathcal{D}\psi(x)\;\bar{\psi}(x_1)\cdots\bar{\psi}(x_n)\psi(y_1)\cdots\psi(y_n)\;\mathrm{e}^{\mathrm{i}S[\bar{\psi},\psi]}\,, \end{aligned}$$

mit Generatoren einer Grassmann-Algebra $\{\bar{\psi}(x),\psi(x)\}$. Wir werden im Folgenden wieder die euklidische Formulierung betrachten. Ist $S_E[\bar{\psi},\psi]$ die euklidische Wirkung, so ergeben sich die euklidischen Erwartungswerte aus einem erzeugenden Funktional

$$\begin{aligned} &Z[\bar{J}(x),J(x)] \;= \\ &\qquad=\; \frac{1}{Z}\int \mathcal{D}\bar{\psi}(x)\;\mathcal{D}\psi(x)\;\mathrm{e}^{-S_E[\bar{\psi},\psi]+\int \mathrm{d}x[\bar{\psi}(x)J(x)+\bar{J}(x)\psi(x)]}\,. \end{aligned}$$

$\bar{J}(x), J(x)$ sind wiederum Elemente einer Grassmann-Algebra, für welche die entsprechenden Ableitungsregeln gelten. Diese antikommutierenden Ableitungen (17.3) tragen der Statistik der Fermionen Rechnung.

Wir berechnen ein Gauß’ sches Integral für Grassmann-Variable. Sei A ein hermitescher Operator (nicht notwendigerweise positiv), dann gilt:

$$\begin{aligned} Z \;&=\; \int \prod_{i=1}^{N} \mathrm{d}\bar{\eta}_i \mathrm{d}\eta_i \;\exp\left(-\sum_{ij}\bar{\eta}_i A_{ij}\eta_j\right) \\ &=\; \int \prod_{i=1}^{N} \mathrm{d}\bar{\eta}_i \mathrm{d}\eta_i \;\sum_{n=0}^{\infty}\frac{(-1)^n}{n!}\;\left(\sum_{i_1 j_1}\cdots\sum_{i_n j_n}(\bar{\eta}_{i_1}A_{i_1 j_1}\eta_{j_1})\ldots(\bar{\eta}_{i_n}A_{i_n j_n}\eta_{j_n})\right) \end{aligned}$$

Dabei verschwinden alle Terme mit Quadraten oder höheren Potenzen desselben η_i. Andererseits folgt aus den Integrationsregeln (Gl.17.2), dass nur ein Term übrigbleibt, der alle η's und $\bar{\eta}$'s jeweils genau einmal enthält. Es gibt $n!$-mal die gleichen Paarungen, man kann somit die Reihenfolge z.B. der $\bar{\eta}$'s festlegen und damit dem Faktor $1/n!$ Rechnung tragen:

$$\begin{aligned} Z &= \sum_{j_1} \cdots \sum_{j_n} \epsilon_{j_1 j_2 \cdots j_n} \; A_{1 j_1} A_{2 j_2} \ldots A_{n j_n} \\ &= \det A \,. \end{aligned}$$

Der total antisymmmetrische Tensor $\epsilon_{i_1 \cdots i_n}$ ergibt sich aus den Anti-Vertauschungsregeln der η's. Das erzeugende Funktional der freien Theorie erhält man wiederum durch eine quadratische Ergänzung. Man benötigt lediglich die Invarianz der Grassmann-Integration unter Translationen:

$$\int \mathrm{d}\psi \; f(\psi) \;=\; \int \mathrm{d}\psi \; f(\psi + \eta) \,.$$

Das Ergebnis ist

$$Z[\bar{J}, J] \;=\; \exp\left(\sum_{ij} \bar{J}_i (A^{-1})_{ij} J_j \right) \,.$$

Die Kombinatorik für Erwartungswerte ist somit bis auf Vorzeichen, die sich bei Vertauschen von zwei Argumenten ergeben, die gleiche wie im bosonischen Fall.

17.3 Die „naive“ Wirkung von Fermionen auf dem Gitter

Unser Ziel ist es, die Dirac-Wirkung

$$S = \int \mathrm{d}^4 x \, \bar{\psi}(x) \left(\mathrm{i} \gamma^\mu \partial_\mu + m \right) \psi(x) \tag{17.4}$$

auf einem Gitter zu formulieren. Hierbei sind γ^μ die Dirac'schen Gamma-Matrizen, definiert durch die Vertauschungsrelationen

$$\{\gamma^\mu, \gamma^\nu\} = 2 g^{\mu\nu} \,. \tag{17.5}$$

Je nach der Darstellung, die man für die Gamma-Matrizen wählt, unterscheidet man verschiedene Arten von Fermionen. Bezeichnen wir mit

$$\sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma^2 = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix} \quad \sigma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{17.6}$$

die Pauli-Matrizen, so sind folgender Darstellungen der Gamma-Matrizen von besonderer Bedeutung:

- Dirac-Darstellung:

$$\gamma^0 = \begin{pmatrix} \mathbf{1} & 0 \\ 0 & -\mathbf{1} \end{pmatrix} \quad \gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix} \quad \gamma^5 = \begin{pmatrix} 0 & \mathbf{1} \\ \mathbf{1} & 0 \end{pmatrix} \,. \tag{17.7}$$

  In dieser Darstellung ist γ^0 diagonal.

- Weyl-Darstellung:

$$\gamma^0 = \begin{pmatrix} 0 & \mathbf{1} \\ \mathbf{1} & 0 \end{pmatrix} \quad \gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix} \quad \gamma^5 = \begin{pmatrix} -\mathbf{1} & 0 \\ 0 & \mathbf{1} \end{pmatrix}. \tag{17.8}$$

  In dieser Darstellung ist γ^5 diagonal.

- Majorana-Darstellung:

$$\gamma^0 = \begin{pmatrix} 0 & \sigma^2 \\ \sigma^2 & 0 \end{pmatrix} \qquad \gamma^1 = \begin{pmatrix} \mathrm{i}\sigma^3 & 0 \\ 0 & \mathrm{i}\sigma^3 \end{pmatrix} \quad \gamma^2 = \begin{pmatrix} 0 & -\sigma^2 \\ \sigma^2 & 0 \end{pmatrix}$$
$$\gamma^3 = \begin{pmatrix} -\mathrm{i}\sigma^1 & 0 \\ 0 & -\mathrm{i}\sigma^1 \end{pmatrix} \qquad \gamma^5 = \begin{pmatrix} \sigma^2 & 0 \\ 0 & -\sigma^2 \end{pmatrix}$$

  In dieser Darstellung sind sämtliche Dirac-Matrizen rein imaginär. Daher lassen sich die Lösungen als reelle Funktionen ansetzen.

Im Folgenden werden wir uns zunächst nicht weiter um die Darstellung der Dirac-Matrizen kümmern, allerdings hat in den vergangenen Jahren die Bedeutung der Majorana-Darstellung auch für die Fermionen auf dem Gitter zugenommen (siehe z.B. [75, 87, 34]).

Will man die euklidische Wirkung für Fermionen diskretisieren, so ergibt sich das Problem, wie man die Ableitung durch einen hermiteschen Ausdruck auf dem Gitter ersetzt. Bei skalaren Feldern tritt nur das Quadrat der Ableitung auf, welches auf dem Gitter durch

$$-\frac{\partial^2}{\partial x^2} \;\longrightarrow\; -\delta_{n+1,m} + 2\delta_{n,m} - \delta_{n-1,m}$$

diskretisiert wird (siehe Kap. 3). Dies ist ein symmetrischer Operator auf Feldern, die an Gitterpunkten definiert sind. Wählen wir für Fermionen ebenfalls Felder, die auf Gitterpunkten definiert sind, so ergibt die Ersetzung der Ableitung durch eine Differenz

$$i\frac{\partial}{\partial x} \;\longrightarrow\; i\,(\delta_{n+1,m} - \delta_{n,m}) \tag{17.9}$$

keinen hermitischen Operator. Man kann sich auch leicht davon überzeugen, dass die Matrix auf der rechten Seite nicht-reelle Eigenwerte hat. Allerdings ist die symmetrische Differenz

$$i\frac{\partial}{\partial x} \;\longrightarrow\; i\,\frac{1}{2}\,(\delta_{n+1,m} - \delta_{n-1,m}) \tag{17.10}$$

hermitisch. Im Kontinuum gilt natürlich (für stetige Funktionen)

$$\lim_{\epsilon\to 0} \frac{f(x+\epsilon) - f(x)}{\epsilon} \;=\; \lim_{\epsilon\to 0} \frac{1}{2}\,\frac{f(x+\epsilon) - f(x-\epsilon)}{\epsilon}\,.$$

Normale und symmetrische Differenz haben aber einen entscheidenden Unterschied: Es gibt nur eine Lösung der Gleichung

$$\psi(n+1) - \psi(n) \;=\; 0 \qquad \longrightarrow \qquad \psi(n) \;=\; \text{const.}\,,$$

genau wie im Kontinuum, jedoch hat die Gleichung

$$\psi(n+1) - \psi(n-1) \;=\; 0 \qquad \longrightarrow \qquad \psi(n) \;=\; c_1 + (-1)^n c_2$$

zwei linear unabhängige Lösungen. Im Kontinuum fällt wegen der Stetigkeit der alternierende Anteil weg. Die Existenz dieser zwei Lösungen für eine Ableitung auf dem Gitter wird sich als problematisch für die Gitterregularisierung von Fermionen herausstellen.

Ungeachtet der vorherigen Bemerkungen verfolgen wir zunächst den naiven Weg und wählen als Gitterwirkung für freie Fermionen

$$\begin{aligned} S &= \frac{\mathrm{i}}{2} \sum_{\vec{n},\mu} \bar{\psi}(\vec{n}) \gamma_\mu (\psi(\vec{n}+\vec{e}_\mu) - \psi(\vec{n}-\vec{e}_\mu)) \;+\; M \sum_{\vec{n}} \bar{\psi}(\vec{n}) \psi(\vec{n}) \\ &= \sum_{\vec{n},\vec{m}} \bar{\psi}(\vec{n}) \; K(\vec{n},\vec{m}) \; \psi(\vec{m}) \,, \end{aligned}$$

mit

$$K(\vec{n},\vec{m}) \;=\; \frac{\mathrm{i}}{2} \sum_{\mu} \; \gamma_\mu (\delta_{\vec{n}+\vec{e}_\mu,\vec{m}} - \delta_{\vec{n}-\vec{e}_\mu,\vec{m}}) \;+\; M \; \delta_{\vec{n},\vec{m}} \,.$$

$K(\vec{n},\vec{m})$ ist natürlich noch eine 4×4-Matrix bezüglich der Spinorindizes. (Wenn nicht anders erwähnt betrachten wir den Fall einer 4-dimensionalen euklidischen Raum-Zeit. Damit umgehen wir eine Diskussion der Darstellung von Dirac-Matrizen für beliebige Raum-Zeit Dimensionen. Die meisten Resultate, insbesondere das No-Go-Theorem von Nielsen und Ninomiya, gelten aber mit den entsprechenden Verallgemeinerungen für beliebige Dimensionen.)

Wegen der Translationsinvarianz von K führt eine Fourier-Zerlegung der Felder wiederum zu einer Diagonalisierung:

$$\begin{aligned} \tilde{K}(\vec{p}) &= \frac{\mathrm{i}}{2} \sum_{\mu} \; \gamma_\mu (\mathrm{e}^{-2\pi \mathrm{i} p_\mu} - \mathrm{e}^{2\pi \mathrm{i} p_\mu}) \;+\; M \\ &= \sum_{\mu} \gamma_\mu \sin 2\pi p_\mu \;+\; M \,. \end{aligned}$$

Das Inverse dieser Matrix ist der Propagator:

$$\begin{aligned} S(p) \;=\; \tilde{K}^{-1}(p) &= \left(\sum_{\mu} \gamma_\mu \sin 2\pi p_\mu \;+\; M \right)^{-1} \\ &= \frac{M \;-\; \sum_\mu \gamma_\mu \sin 2\pi p_\mu}{\sum_\mu \sin^2 2\pi p_\mu \;+\; M^2} \end{aligned}$$

Ganz entsprechend wie bei skalaren Feldern wird der Kontinuumslimes für $M \to 0$ erreicht. Das asymptotische Verhalten des Propagators im Ortsraum ist durch den Imaginärteil der Pole im Impulsraum gegeben. Man erkennt jedoch, dass der Propagator neben dem Pol im Ursprung weitere Pole hat, und zwar immer, wenn ein $p_i = 0$ oder $\frac{1}{2}$ ist. Es gibt insgesamt 16 Pole innerhalb des Impulsbereiches $0 \le p_i < 1$ in 4 Dimensionen.

Die Theorie beschreibt also nicht nur ein einzelnes Fermion, sondern insgesamt 16. Dies hängt mit der obigen Bemerkung zusammen: Pole des Propagators sind Nullmoden des Dirac-Operators. Da aber die diskretisierte Ableitung für jede Richtung zwei unabhängige Nullmoden besitzt, erhält man in d Raum-Zeit-Dimensionen auf dem Gitter 2^d Spezies von Fermionen. Dies bezeichnet man als das Problem der Fermionverdopplung.

17.4 Chirale Fermionen

Die Wirkung von masselosen Fermionen

$$S \;=\; \int \mathrm{d}^4x\; \bar\psi(x)\mathrm{i}\partial\!\!\!/\psi(x) \qquad\qquad \bar\psi \;=\; \psi^+\gamma_0$$

ist invariant unter 2 globalen Transformationen:

$$\begin{aligned} \psi(x) \;&\longrightarrow\; \mathrm{e}^{\mathrm{i}\alpha}\psi(x) & \bar\psi(x) \;&\longrightarrow\; \mathrm{e}^{-\mathrm{i}\alpha}\bar\psi(x) \\ \text{und}\qquad \psi(x) \;&\longrightarrow\; \mathrm{e}^{\mathrm{i}\alpha'\gamma_5}\psi(x) & \bar\psi(x) \;&\longrightarrow\; \bar\psi\,\mathrm{e}^{\mathrm{i}\alpha'\gamma_5} \end{aligned}$$

Die Invarianz unter Multiplikation mit einer Phase hängt mit der Ladungserhaltung zusammen. Die zweite Invarianz, die sich aus den Vertauschungrelationen der γ-Matrizen ergibt,

$$\begin{aligned} \{\gamma_5,\gamma_\mu\} \;=\; 0 \quad &\longrightarrow \quad \mathrm{e}^{\mathrm{i}\alpha\gamma_5}\gamma_\mu \;=\; \gamma_\mu\mathrm{e}^{-\mathrm{i}\alpha\gamma_5} \\ \Longrightarrow \qquad \mathrm{i}\gamma_\mu\partial^\mu\psi(x) \quad &\longrightarrow \quad \mathrm{i}\gamma_\mu\partial^\mu\mathrm{e}^{\mathrm{i}\alpha\gamma_5}\psi(x) \;=\; \mathrm{i}\mathrm{e}^{-\mathrm{i}\alpha\gamma_5}\gamma_\mu\partial^\mu\psi(x)\,, \end{aligned}$$

gilt also nicht für einen Masseterm $\bar\psi\psi$. Diese Transformation heißt *chirale Transformation*, der zugehörige Erhaltungssatz ist die *Chiralität*. Allgemein folgt sie aus der Relation:

$$\{\gamma_5, \not{D}\} \;=\; 0\,. \tag{17.11}$$

Definiert man links- und rechtshändige Wellenfunktionen

$$\psi_L \;=\; \frac{1}{2}(1-\gamma_5)\psi \qquad \text{und} \qquad \psi_R \;=\; \frac{1}{2}(1+\gamma_5)\psi\,,$$

so werden bei einer chiral invarianten Theorie diese Anteile der Wellenfunktion von Fermionen durch die Dynamik nicht gemischt, es ist also möglich eine ungleiche Anzahl von rechts- bzw. linkshändigen Teilchen im Universum zu haben.

Das Standardmodell behandelt zunächst alle Fermionen als chirale Teilchen, d.h. Elektronen oder Quarks sind anfänglich masselos, wobei links- und rechtshändige Anteile mit unterschiedlichen Termen in der Wirkung auftreten. Erst durch die Yukawa-Kopplung an das Higgsfeld und die spontane Symmetriebrechung des Higgsfeldes erhalten ein Teil der Fermionen ihre Massen. So sind z.B. das Elektron und das Neutrino zu einem linkshändigen $SU(2)$-Doublett $\begin{pmatrix} e \\ \nu \end{pmatrix}_L$ und einem rechtshändigen Singlett e_R (und falls sich das Neutrino als massiv herausstellen sollte auch einem rechtshändigen ν_R) zusammengefasst. Die Yukawa-Kopplung

$$g\,\bar\psi_L\psi_R\varphi + \text{c.c.} \;=\; g\,(\bar e\;\;\bar\nu)_L\cdot e_R\begin{pmatrix}\varphi_1\\ \varphi_2\end{pmatrix} + \text{c.c.}$$

ist natürlich invariant unter $SU(2)$–Transformationen. Ist der Grundzustand des Higgs-Feldes durch $\begin{pmatrix}\langle\varphi\rangle\\ 0\end{pmatrix}$ definiert, so erhält man einen Term der Art $g\langle\varphi\rangle\bar e_L e_R + \text{c.c.}$, der die Chiralität der Elektronen bricht (ohne ν_R allerdings nicht die der Neutrinos) und zu einem effektiven Massenterm führt.

17.5 Das No-Go-Theorem von Nielsen und Ninomiya

Im nächsten Abschnitt werden wir die gängiste Methode, das Problem der Fermionen-Verdopplung zu umgehen, kurz behandeln. Bei diesen so genannten „Wilson-Fermionen“ treten jedoch rechts- und linkshändige Anteile immer paarweise auf, auch im masselosen Fall. Da im Standardmodell rechts- und linkshändige Anteile der Fermionen jedoch unterschiedlich behandelt werden, möchte man natürlich auch chirale Fermionen auf dem Gitter formulieren können. Dass dies unter sehr allgemeinen Annahmen an die Theorie nicht möglich ist, besagt das Theorem von Nielsen und Ninomiya.

THEOREM (NIELSEN-NINOMIYA): *Sei der quadratische Anteil der Wirkung für Fermionen auf einem Gitter durch einen Ausdruck der Art*

$$S \;=\; \sum_{\vec{m},\vec{n}} \bar{\psi}(\vec{m}) K(\vec{m},\vec{n}) \psi(\vec{n})\,, \tag{17.12}$$

gegeben, dann lassen sich die folgenden Bedingungen nicht gleichzeitig erfüllen:

1. *Der „Dirac-Operator“ K anti-kommutiert mit γ_5*:

   $$\{\gamma_5, K(\vec{m},\vec{n})\} \;=\; 0\,.$$

   Diese Forderung entspricht der üblichen Formulierung der chiralen Invarianz auf dem Gitter.

2. *Die Wechselwirkung zwischen Fermionen ist lokal, d.h. $K(\vec{m},\vec{n}) \to 0$ für $|\vec{m}-\vec{n}| \to \infty$ derart, dass die Fourier-Transformierte von K einmal stetig ableitbar ist.*

3. *Die Wirkung auf dem Gitter ist translationsinvariant, d.h.*

   $$K(\vec{m},\vec{n}) \;=\; K(\vec{m}+\vec{a},\vec{n}+\vec{a})\,,$$

   für jeden Vektor $\vec{a}$ mit $a_i \in \mathbf{Z}$.

4. *Die Wirkung ist reell, d.h. die Matrix K ist hermitesch,*

   $$K(\vec{m},\vec{n}) \;=\; K(\vec{n},\vec{m})^*\,.$$

5. *Die Ladungen Q des Modells (z.B. Leptonenzahl etc.) sind erhalten, lokal definiert, quadratisch in den Feldern und quantisiert.*

Auf den exakten Beweis soll hier verzichtet werden (siehe z.B. [82] und [83]). Im Wesentlichen beruht der Beweis auf einer Verallgemeinerung der folgenden Beobachtung:

Eine einmal stetig differenzierbare, periodische Funktion von p mit Periode 1, die bei $p = 0$ verschwindet aber mit einem linearen Term beginnt, muss im Intervall $(0, 1)$ noch eine zweite Nullstelle haben.

Da der inverse Propagator $S^{-1}(p)$ für chirale Fermionen im Impulsraum in p bei $p = 0$ linear ist,

$$S^{-1}(p) \;\approx\; \gamma_\mu p^\mu\,,$$

erhält man eine weitere Nullstelle in jeder Richtung der Brillouin-Zone, in welcher das Gitter periodisch ist. Die Ableitung an dieser zweiten Nullstelle hat umgekehrtes Vorzeichen, sodass dieser Punkt ein Fermion mit entgegengesetzter Chiralität beschreibt. Somit treten Fermionen unterschiedlicher Chiralität immer paarweise auf.

An dieser Argumentation erkennt man auch die Möglichkeiten, das Theorem (unter Aufgabe obiger Annahmen) zu umgehen:

1. Wir verzichten auf die Relation
$$\{\gamma_5, K\} \;=\; 0\,.$$
Dies ist genau das, was sich in den letzten Jahren als die Lösung des Problems herauskristallisiert hat. Zunächst hat es den Anschein, als ob wir damit auf die chirale Invarianz verzichten. Es hat sich jedoch gezeigt, dass obige Bedingung zu restriktiv ist und durch eine etwas verallgemeinerte Bedingung ersetzt werden kann.

2. Aufgabe der Lokalität erlaubt einen nicht-stetigen inversen Propagator $S^{-1}(p)$. Mit einer Sprungstelle oder aber einem Pol lässt sich die zusätzliche Nullstelle umgehen (ein Beispiel hierfür sind MIT-Fermionen sowie auch die so genannten „exakten Wirkungen“).

3. Ein irreguläres (Zufalls-) Gitter hat keine Brillouin-Zone, d.h., die Funktion $S^{-1}(p)$ muss nicht periodisch sein.

4. Aufgabe der Hermitizität erlaubt der Funktion $S^{-1}(p)$ sich bei komplexen Impulsen „um die reelle Achse zu winden“, ohne eine erneute Nullstelle zu haben.

Es wurden noch weitere Möglichkeiten in der Literatur diskutiert, das Problem chiraler Fermionen auf dem Gitter zu behandeln (vgl. den Übersichtsartikel von J. Smit [97], der den Status der Fermionen auf dem Gitter Mitte der 80er Jahre gut zusammenfasst). Auf den aktuelleren Stand der Diskussion werden wir im nächsten Kapitel eingehen. Doch auch wenn man auf die chirale Invarianz auf dem Gitter verzichtet, ist das Standardmodell auf dem Gitter nicht von vorne herein zum Scheitern verurteilt. Für Wilson-Fermionen gibt es Modelle, für die sich zeigen lässt, dass die chirale Invarianz im Kontinuumslimes restauriert wird.

17.6 Wilson-Fermionen und die Kopplung an Eichfelder

Die wohl bekannteste klassische Methode, das Problem der Fermionen-Verdopplung zu umgehen, stammt von Wilson. In diesem Fall wird zur Dirac-Wirkung noch ein Term addiert, der zweite Ableitungen der Fermionen enthält:

$$\begin{aligned} S_W \;=\; & \frac{\mathrm{i}}{2}\sum_{\vec{n}}\sum_{\mu}\bar{\psi}(\vec{n})\gamma_\mu(\psi(\vec{n}+\vec{e}_\mu)-\psi(\vec{n}-\vec{e}_\mu)) \;- \\ & -\frac{r}{2}\sum_{\vec{n}}\sum_{\mu}\bar{\psi}(\vec{n})(\psi(\vec{n}+\vec{e}_\mu)+\psi(\vec{n}-\vec{e}_\mu)-2\psi(\vec{n}))\,. \end{aligned}$$

Die Addition dieses Terms verhindert, dass auf dem Rand der Brillouin-Zone weitere Nullstellen auftreten, allerdings bricht er auch für $r \neq 0$ explizit die chirale Invarianz. In einem formalen

Kontinuumslimes verschwindet der zusätzliche Term relativ zum führenden Dirac-Term proportional zur Gitterkonstanten.

Effektiv entsprich der Zusatzterm einer zweiten Ableitung und damit einer höreren Ordnung. Wir schreiben im Folgenden den Gitterparameter hinzu, damit die Ordnung für den Kontinuumslimes deutlicher wird. Die Wilson-Wirkung lautet in diesem Fall:

$$\begin{aligned} S &= a^4 \frac{\mathrm{i}}{2a} \sum_{\vec{n}} \sum_{\mu} \bar{\psi}(\vec{n}) \gamma_\mu (\psi(\vec{n}+\vec{e}_\mu) - \psi(\vec{n}-\vec{e}_\mu)) \; - \; a^4 m \sum_n \bar{\psi}(n)\psi(n) \\ &\quad -a^4 \frac{r}{2a} \sum_{\vec{n}} \sum_{\mu} \bar{\psi}(\vec{n}) (\psi(\vec{n}+\vec{e}_\mu) + \psi(\vec{n}-\vec{e}_\mu) - 2\psi(\vec{n}))\,. \end{aligned}$$

Der Propagator hat nun die Form:

$$\langle \bar{\psi}(p)\psi(-p)\rangle = \left(\sum_\mu \gamma_\mu \frac{1}{a} \sin p_\mu a + m + \sum_\mu \frac{2r}{a} \sin^2\left(\frac{p_\mu a}{2}\right) \right)^{-1} . \tag{17.13}$$

Die Kopplung von Eichfeldern an Fermionen erfolgt wie bei skalaren Feldern. Das Produkt von fermionischen Feldern an verschiedenen Gitterpunkten wird ersetzt durch das Produkt von Fermionen, welche durch den Paralleltransporter verbunden sind, z.B.:

$$\bar{\psi}(\vec{n})\psi(\vec{n}+\vec{e}_\mu) \quad \longrightarrow \quad \bar{\psi}(\vec{n}) U(\vec{n},\mu) \psi(\vec{n}+\vec{e}_\mu)\,.$$

In der entsprechenden Wilson-Wirkung

$$S_W[U,\bar{\psi},\psi] \; = \; \sum_{\vec{m},\vec{n}} \bar{\psi}(\vec{m}) K[U](\vec{m},\vec{n}) \psi(\vec{n})$$

ist die Matrix K nun ein Funktional der Eichkonfiguration $\{U\}$.

Die Kopplung von Fermionen an Eichfelder bringt ein neues Problem. Schon im Kontinuum lässt sich zeigen, dass einzelne chirale Fermionen, angekoppelt an Eichfelder, zu einer Anomalie führen (Adler-Bell-Jackiw-Anomalie). Diese lässt sich dadurch vermeiden, dass mehrere chirale Fermionen an Eichfelder gekoppelt werden, für welche die Summe der Ladungen verschwindet,

$$\sum_i Q_i \; = \; 0\,.$$

Diese Bedingung ist im Standardmodell für jede Generation von Fermionen getrennt erfüllt. Sie zeigt aber, dass Quarks oder Leptonen in der üblichen Form nicht alleine existieren würden, bzw. sie begründet (zumindest formal), warum die Ladungen der Quarks und der Leptonen relativ zueinander quantisiert sind.

17.7 "Staggered Fermionen"

Unter den vielen anderen Modellen, die vorgeschlagen wurden, um das Problem der Fermion-Verdopplung zu umgehen, sei hier nur noch kurz das Modell der "staggered fermions" erwähnt, da es immer noch praktische Anwendungen findet.

Die Grundidee ist in diesem Fall, aus der Not eine Tugend zu machen, und die 16 Pole im Fermi-Propagator verschiedenen Flavour-Freiheitsgraden zuzuordnen. Insgesamt beschreiben die 16 Pole vier Flavour-Freiheitsgrade, wobei jeder Flavour-Freiheitsgrad nochmals den vier Komponenten eines (massiven) Fermions entspricht. Der Nachteil dieses Verfahrens ist, dass die vier Flavour-Freiheitsgrade dieselbe Masse haben. Das entspricht nicht der Beobachtung (das Strange- und insbesondere das Charm-Quark sind sehr viel schwerer als die Up- und Down-Quarks). Es gibt zwar auch hier Tricks diese Nachteile zu umgehen, aber wie immer auf Kosten der Natürlichkeit.

Jedem Gitterpunkt wird ein Paar konjugierter Fermi-Felder $(\bar{\chi}, \chi)$ zugeschrieben. Das bedeutet, es gibt zunächst keine Gamma-Matrizen. Die Wirkung entspricht der herkömmlichen symmetrisierten Ableitung:

$$S = a^4 \sum_{m,\mu} \frac{1}{2a} \bar{\chi}(m) \eta_{m,\mu} (\chi(m + e_\mu) - \chi(m - e_\mu)) + a^4 m \sum_m \bar{\chi}(m) \chi(m) \,. \tag{17.14}$$

Außerdem definieren wir die folgenden Koeffizienten:

$$\eta_{m,1} = 1 \ , \ \eta_{m,2} = (-1)^{m_1} \ , \ \eta_{m,3} = (-1)^{(m_1+m_2)} \ , \ \eta_{m,4} = (-1)^{(m_1+m_2+m_3)} \,. \tag{17.15}$$

Diese alternierenden, schachbrettartig verteilten Vorzeichen für die Wirkung der Fermi-Felder führen effektiv zu den vier gewöhnlichen Fermionen mit jeweils vier Spinor-Komponenten.

Es gibt noch andere Modelle, die das Problem der Fermionen-Verdopplung auf dem Gitter zu umgehen versuchen, auf die soll hier jedoch nicht weiter eingegangen werden soll. Es gibt dazu die klassischen Übersichtsartikel von J. Smit [97] und E.H. de Groot [23] oder auch das (aktuellere) Buch von DeGrand und DeTar [22].

17.8 „Simulation“ von Fermionen

Das euklidische Funktionalintegral für Fermionen ist mit Grassmann-Variablen definiert. Diese bilden jedoch keine Zufallsvariable (d.h. reelle Größen mit einer Verteilungsfunktion), sodass man kein statistisches System erhält. Fermionen lassen sich somit nicht direkt in Monte-Carlo-Verfahren simulieren. (Auf 2-dimensionalen Gittern kann man sich die Äquivalenz von Fermionen mit dem Ising-Modell am kritischen Punkt zu Nutze machen. Solche oder ähnliche Ansätze sind jedoch für 4 Dimensionen nicht bekannt oder brauchbar.)

Das Standardmodell ist jedoch durch eine Wirkung definiert, welche quadratisch in den fermionischen Feldern ist, also Gaußsch bzgl. dieser Freiheitsgrade. Wie wir gesehen haben, sind Erwartungswerte in einer solchen Theorie (auch bei Grassmann-Variablen) durch die 2-Punkt-Funktion bzw. die Determinante — also $K^{-1}(\vec{m}, \vec{n})$ bzw. det K — ausdrückbar. So gilt z.B.

$$\begin{aligned}
\langle \bar{\psi}(\vec{m}) \psi(\vec{n}) \dots U(\vec{l}, \mu) \dots \rangle \quad &= \\
&= \frac{\int \mathrm{d}[U, \bar{\psi}, \psi] \ \bar{\psi}(\vec{m}) \psi(\vec{n}) \dots U(\vec{l}, \mu) \dots \ \mathrm{e}^{-S[U] - S_W[U, \bar{\psi}, \psi]}}{\int \mathrm{d}[U \bar{\psi} \psi] \ \mathrm{e}^{-S[U] - S_W[U, \bar{\psi}, \psi]}} \\
&= \frac{\int \mathrm{d}[U] \ \dots U(\vec{l}, \mu) \dots \ \left(\sum K[U]^{-1}(\vec{m}, \vec{n}) \right) \ \det K[U] \ \mathrm{e}^{-S[U]}}{\int \mathrm{d}[U] \ \det K[U] \ \ \mathrm{e}^{-S[U]}} \,.
\end{aligned}$$

(Hierbei bedeutet d$[U, \bar{\psi}, \psi]$ eine Integration über alle Felder U auf den Gitterlinien, bzw. $\bar{\psi}$ und ψ auf den Gitterpunkten. $S[U]$ ist der reine Eichanteil der Wirkung.) Das Problem, Erwartungswerte von fermionischen Feldern mit Eichfeldern zu berechnen, lässt sich also reduzieren auf das Problem, Erwartungswerte der Eichfelder sowie der inversen Matrix $K^{-1}[U]$ bezüglich einer effektiven Wirkung zu bstimmen. Diese effektive Wirkung

$$S_{\text{eff}} \;=\; S[U] - \ln \det\, K[U]$$

ist allerdings nicht mehr lokal in den Variablen $\{U\}$. Das macht ihre Berechnung schwierig: Für jede Konfiguration $\{U\}$ ist zunächst die Determinante von $K[U]$ zu bestimmen. Wird die neue Konfiguration im Monte Carlo Verfahren akzeptiert, so ist zur Bestimmung des Erwartungswertes noch $K^{-1}[U]$ zu berechnen. Das Verfahren ist also sehr aufwendig, allerdings im Prinzip durchführbar. Aus diesem Grunde sind die bisherigen Simulationen von Eichtheorien mit Fermionen auch auf sehr viel kleinere Gitter beschränkt.

Gerade die Berechnung der Determinante von $K[U]$ ist sehr rechenaufwendig, sodass man oft die Näherung det K=const. macht. Das vereinfacht die Rechnungen erheblich, man hat allerdings die Theorie der Fermionen mit Eichfeldern abgewandelt. Diese Näherung mit so genannten „quenched fermions“ erhält man formal aus der alten Theorie im Grenzfall $m \to \infty$, d.h. wenn die Fermionen unendlich schwer werden. In der Störungsrechnung drückt sich diese Näherung dadurch aus, dass man alle Graphen mit inneren Schleifen von Fermion-Propagatoren vernachlässigt. Das ist jedoch nicht gleichbedeutend mit den „statischen externen“Quellen, die wir im Zusammenhang mit reinen Eichtheorien betrachtet haben. Die Erwartungswerte der Fermi-Felder hängen immer noch von der gewählten Wirkung ab, da sie die Bestimmung des inversen Fermi-Propagators erfordern.

17.9 Improved actions

Zum Abschluss dieses Kapitels soll kurz auf ein Verfahren eingegangen werden, mit dem man den Kontinuumslimes einer Gittertheorie beschleunigen kann. Insbesondere wenn bei chiralen Fermionen die Gitterwirkung die chirale Symmetrie bricht, ist man an einem raschen Kontinuumslimes interessiert, in dem die chirale Symmetrie wieder sichtbar wird. Gewöhnlich muss man den Kontinuumslimes mit einem größeren Gitter und einer Simulation näher am kritischen Punkte erkaufen. Improved actions machen jedoch auch Simulationen auf kleineren Gittern möglich.

Die grundlegende Idee der „improved actions“ ist vergleichsweise einfach: Wenn wir beispielsweise die Ableitung einer Funktion im Kontinuum durch einen Differenzenquotienten auf dem Gitter ersetzen, so ist das nur eine Approximation, die höhere Korrekturen in der Gitterkonstanten a hat. Sei

$$\Delta f \;=\; \frac{f(x+a) - f(x-a)}{2a}$$

die symmetrisierte erste Ableitung einer Funktion auf dem Gitter. Falls f hinreichend differenzierbar ist, können wir die rechte Seite nach Potenzen von a entwickeln (das entspricht einem „naiven“ Kontinuumslimes, vgl. Kap. 16) und erhalten:

$$\Delta f \;=\; \frac{\partial f}{\partial x} + \frac{a^2}{6}\frac{\partial^3 f}{\partial x^3} + \mathcal{O}(a^4)\,.$$

Wir erhalten also einen verbesserten (improved) „naiven“ Kontinuumslimes, wenn wir den ersten Differentialquotienten auf dem Gitter durch

$$\Delta f - \frac{a^2}{6}\Delta^3 f \;=\; \frac{f(x+a)-f(x-a)}{2a} - \frac{a^2}{6}\frac{f(x+3a)-3f(x+a)+3f(x-a)-f(x-3a)}{8a^3}$$

ersetzen. Die Korrekturen zur 1. Ableitung im Kontinuum sind nun von der Ordnung a^4.

Die obigen Überlegungen beziehen sich jedoch auf einen „naiven“ Kontinuumslimes, bei dem die Funktion f als hinreichend „glatt“ angenommen wird, sodass man sie um einen Punkt herum entwickeln kann. Bei statistischen Modellen bzw. Monte Carlo Simulationen gibt es aber keine derartigen Einschränkungen an f; es macht eigentlich keinen Sinn, $f(x+a)$ durch $f(x)$ und ihre Ableitungen bei x auszudrücken. Die „improved actions“ der Gittereichtheorien sollten für relevante Observable, d.h. Erwartungswerte, eine raschere Konvergenz liefern. Meist geht man von einem allgemeinen Ansatz für die Wirkung aus, von dem zu erwarten steht, dass er zur selben Universalitätsklasse gehört, wie die ursprüngliche Wirkung. Für die Wirkung der reinen Eichtheorie kann das beispielsweise heißen:

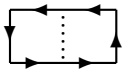
$$S \;=\; \beta\left(c_0\sum P_{\text{plaq}} + c_1\sum P_{\text{rect}} + c_2\sum P_{\text{chair}} + c_3\sum P_{\text{para}} + ...\right)\,,$$

wobei

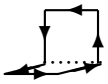
$$P_{\text{plaq}} \;=\; \prod_{l\in\text{plaquette}} U(l) \;\simeq$$

das übliche Produkt der Eichvariablen um eine elementare Plaquette ist. Die anderen drei Terme entsprechen den geschlossenen Wegen der Länge 6, dem Rechteck, dem abgeknickten Rechteck (chair) und einem Weg über den Würfel mit jeweils paarweise parallelen Linien:

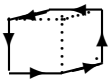
$$P_{\text{rect}} \;\simeq$$



$$P_{\text{chair}} \;\simeq$$



$$P_{\text{para}} \;\simeq$$



Im Prinzip lassen sich auch höhere Terme berücksichtigen. Damit $\beta = 1/2g^2N$ seine übliche Bedeutung behält, muss zwischen den Koeffizienten die folgende Beziehung bestehen (sofern wir nur die oben angegebenen Terme berücksichtigen [50]:

$$c_0 + 8c_1 + 16c_2 + 8c_3 \;=\; 1\,.$$

Theoretisch könnte man dieses Verfahren fortsetzung den Differenzenquotienten durch eine unedliche Reihe (und damit nicht-lokale Wirkung) von Differenzenquotienten ersetzen, die aber keine Korrekturen in der Gitterkonstanten mehr haben. Dies führ auf die exakten Gitterwirkungen, auf die wir im nächsten Kapitel kurz eingehen werden.

Kapitel 18

Fermions on the lattice - Part II

Up to now, we have only taken into account the development in the field during the 80th. The picture about fermions on the lattice, and in particular about chiral fermions on the lattice, changed drastically during the 90th. Many of the following explanations are based on a review article by Chandrasekharan and Wiese [16] (but see also [34, 69, 20]).

In 1993, Kaplan [51, 52] proposed a way to circumvent the theorem of Nielsen and Ninomiya. His idea was based on an extension of the 4-dimensional to a 5-dimensional model. By introducing special boundary conditions which effectively lead to two interfaces, so-called "domain walls", one part of the massless chiral fermions lived on one of these domain walls and the other part of these fermions lived on the other domain wall. In the limit of infinite separation of these domain walls one could prove that the effective 4-dimensional theory on the domain wall described chiral fermions. In this way, one obtains an exact chiral invariant theory on the lattice. Around the same time, Narayanan and Neuberger developed what is now known as "overlap lattice fermions" [77]. They used an infinite tower of fermions in order to obtain chirality. It turned out that this infinite tower of fermions in a 4-dimensional theory essentially corresponds to the fermions in a 5-dimensional theory with the fifth dimension being related to the infinite number of flavours.

Again a different approach was taken by Hasenfratz and Niedermayer [43]. They investigated so-called perfect lattice actions. By construction, these actions have the exact invariance of the continuum theory. They can be either obtained as the fixed points of renormalization group block transformations or by integrating over the continuum fermions within a cell around the lattice point and assigning this lattice point the "effective Fermi field" which results from this integration. In the context of these investigations, an old relation found by Ginsparg and Wilson [35] was rediscovered. Mainly through the work of M. Lüscher [68, 69] the relevance of this relation for chirality on the lattice was clarified. Perfect actions circumvent the Nielsen-Ninomiya theorem by being non-local. The propagators show poles in momentum space which allow for a sign-flip without crossing the $p = 0$-line.

Most of the ideas discussed so far (domain wall and/or overlap fermions, perfect actions) lead to exact chiral invariance on the lattice (at least in the infinite lattice limit). Even though these models are of great interest for theoretical reasons, they are numerically extremely demanding when it comes to practical calculations. A different approach, which also leads to chiral symmetry in the continuum limit but for which chiral invariance is broken on the lattice, was

taken by Frezzotti [30]. He defined what is called "twisted mass" lattice QCD. Effectively one adds an additional mass term (multiplied with γ_5) to the action. By properly choosing this extra mass one can show that exact chirality is restored in the continuum limit even though it is broken on the lattice. The advantage of this method is that it is numerically much less demanding. Overlap fermions and twisted mass fermions have been compared [17, 12] in order to get a better idea whether it is more meaningful to use the fast twisted mass algorithms at the cost of going close to the continuum limit (which also is numerically demanding) or to use the much more elaborated versions of overlap fermions with the advantage that they exhibit exact chiral invariance on the lattice.

18.1 Perfect Lattice Fermions - Free Case

We begin with perfect lattice fermions, because they are conceptually the most easiest approach (at least for the free case) to chiral fermions on the lattice. The general idea is to construct a lattice action which by definition yields the same results as the continuum theory. In particular, the (low energy) spectrum of the Hamiltonian is identical.

We begin by defining lattice Fermi fields which are block averages of continuum fields:

$$\Psi(m) = \frac{1}{a^4} \int \mathrm{d}^4x \, \psi(x) B(x-m) \,, \tag{18.1}$$

where $B(x-m)$ is the characteristic function of a box of linear size a which is centered around the (4-dim.) point labeled by $m \in \mathbf{Z}^4$. The field $\psi(x)$ is the continuum field with the continuum chiral action

$$S = \int \mathrm{d}^4x \, \bar{\psi}(x) \mathrm{i} \gamma_\mu \partial^\mu \psi(x) \,. \tag{18.2}$$

We now define the "blocked" 2-point function:

$$\langle \bar{\Psi}(m) \Psi(n) \rangle = \frac{1}{a^8} \int \mathrm{d}^4x \, \mathrm{d}^4y \, B(x-m) B(y-n) \langle \bar{\psi}(x) \psi(y) \rangle \,. \tag{18.3}$$

Notice that this is an "exact" expression, as the 2-point function on the right is obtained from the continuum theory. The blocking only eliminates some of the higher modes but keeps the low energy spectrum identical to the continuum expression.

We now use the exact blocked 2-point function to determine a lattice action which gives rise to this propagator. We define a matrix (indexed by lattice points and fermi-components) by

$$(\rho^{-1})(m,n) = \langle \bar{\Psi}(m) \Psi(n) \rangle \tag{18.4}$$

and use for the lattice action:

$$S = a^4 \sum_{m,n} \bar{\Psi}(m) \rho(m,n) \Psi(n) \,. \tag{18.5}$$

As the continuum action is quadratic, the continuum n-point functions factorize into 2-point functions. This factorization also holds for the blocked variables, i.e. the lattice action expressed in blocked Fermi-fields is also quadratic. By definition this action yields the continuum expression. Hence, for $m = 0$ it is chirally invariant, i.e. we should have

$$\{\gamma_5, \rho\} = 0 \,. \tag{18.6}$$

With this action, chiral symmetry is exact on the lattice. The propagator $\rho(mn,)$ can be calculated explicitly by using Fourier-transforms.

One may wonder, how the theorem of Nielsen and Ninomiya is circumvented. A detailed analysis shows that $\rho(m,n)$ is non-local, i.e. Fermi-fields which are arbitrarily far apart on the lattice have a direct interaction. Indeed, one can easily show that for $d=1$ (the 1-dimensional case) this lattice action is given by

$$\rho(m,n) = \frac{1}{a}(-1)^{(m-n)} \,. \tag{18.7}$$

The corresponding Fourier-transform has a pole in the Brillouin zone which allows for a change of sign without crossing the p-axis.

Formally, we may define the lattice action in the blocked variables also by the following construction:

$$\begin{aligned} \mathrm{e}^{-S[\bar\Psi,\Psi]} &= \int \mathcal{D}\bar\psi\mathcal{D}\psi \, \exp\left(-\int \mathrm{d}^4x\, \bar\psi(x)\mathrm{i}\gamma_\mu\partial^\mu + m)\psi(x)\right) \times \\ &\quad \prod_m \left[\delta\left(\bar\Psi(m) - \frac{1}{a^4}\int \mathrm{d}^4x \bar\psi(x) B(x-m)\right)\times \right. \\ &\quad \left. \delta\left(\Psi(m) - \frac{1}{a^4}\int \mathrm{d}^4x \psi(x) B(x-m)\right)\right] \end{aligned} \tag{18.8}$$

This method can even be applied when the action is not quadratic, i.e. when one cannot argue that the n-point functions factorize into 2-point functions. Furthermore, replacing the δ-functions by smoother functions we can make the effective action local and still keep a better control about the continuum limit.

We replace eq. 18.8 by

$$\begin{aligned} \mathrm{e}^{-S_c[\bar\Psi,\Psi]} &= \int \mathcal{D}\bar\psi\mathcal{D}\psi \, \exp\left(-\int \mathrm{d}^4x\, \bar\psi(x)(\mathrm{i}\gamma_\mu\partial^\mu + m)\psi(x)\right) \times \\ &\quad \prod_m \exp\left[-\frac{1}{c}\left(\bar\Psi(m) - \frac{1}{a^4}\int \mathrm{d}^4x \bar\psi(x) B(x-m)\right)\times \right. \\ &\quad \left. \left(\Psi(m) - \frac{1}{a^4}\int \mathrm{d}^4x \psi(x) B(x-m)\right)\right] \end{aligned} \tag{18.9}$$

This now is a Gaussian integral and in the limit $c \to 0$ we recover the δ-function from above. For $c \neq 0$ the locality properties of this action are much better. Furthermore, one can show that for

$$c = \frac{\exp(ma) - 1 - ma}{m^2 a} \longrightarrow \frac{a}{2} + \mathcal{O}(a^2) \tag{18.10}$$

one obtains for the anti-commutation relations of this propagator with γ_5 the famours Ginsparg-Wilson relation [35]:

$$\{\rho_c^{-1}, \gamma_5\} = 2c\gamma_5 = a\gamma_5 \,. \tag{18.11}$$

Expressed differently this relation reads:

$$\{D, \gamma_5\} = aD\gamma_5 D \,, \tag{18.12}$$

where D is the lattice Dirac operator. This relation becomes particularly important when gauge fields are coupled to the fermions. Ginsparg and Wilson derived it from general renormalization group calculations, but later Lüscher interpreted this relation as the fundamental lattice counterpart of the continuum chiral invariance.

From a conceptual perspective, actions like the one in eq. 18.8 or eq. 18.9 are very valuable, but when it comes to actual calculations and when the Fermi-fields are coupled to gauge fields, it is almost impossible to obtain exact expressions for these actions. Furthermore, they are numerically very expensive. Many groups, in particular Hasenfratz et al. [44] have calculated almost perfect lattice actions involving also gauge fields by determining fixed points of block transformations. Even though the starting point is a lattice theory, these fixed points do not "see" the lattice cut-off and should, therefore, possess the continuum properties.

18.2 Domain wall and overlap fermions

Fermions on the lattice and, in particular, the problem of chirality on the lattice, have become a huge field of research, and in a context like this it is almost impossible to elaborate on all approaches, even the most interesting ones. Therefore, the following remarks should only give a brief idea, and for more details I refer to the literature. The domain wall fermions have been intensively investigated by Kaplan et al. [53], and the formulation mainly used today is due to Shamir [95]. Overlap fermions have been investigated by Narayanan and Neuberger [77] (see also [78, 79]) who also showed the equivalence of overlap fermions with domain wall fermions.

The general idea of domain wall fermions consists in extending the model to a five dimensional lattice. The discretized Dirac operator in the four-dimensional layers remains the same as in the naive fermion action (symmetrized derivative, including the four-dimensional gauge group variables). Along the fifth direction the (symmetrized) derivative is multiplied with γ_5 (there are no gauge variables in this direction but instead clever arrangements of chiral projection operators). Most important, an additional mass term $M(s)$ is introduced which varies along the fifth direction (labeled by s). In particular, there are two four dimensional layers (because of periodic boundary conditions), where $M(s)$ vanishes and switches sign.

One can show that chiral fermions are trapped in these 4-dim. layers where $M(s)$ vanishes. The Nielsen-Ninomiya theorem is avoided, because there are both chiralities on the 5-dim. lattice, however, they are concentrated along different 4-dimensional hyperspaces. If these hyperspaces are infinitely separated (which, of course, is not possible on a finite lattice), chirality becomes exact along these hyperspaces.

Narayanan and Neuberger showed that the domain wall fermions are equivalent to a different, purely four dimensional formulation of chiral fermions on the lattice, which, however, uses an infinite tower of fermions. Effectively, the infinite tower of flavour degrees of freedom corresponds to the fifth dimension. One can derive an effective lattice action with chiral fermions on the four-dimensional lattice (see also [69]) and show that the Ginsparg-Wilson relation is satisfied for this effective action. This four-dimensional approach with an infinite tower of fermions has been named "overlap fermions".

18.3 Twisted mass fermions

All the "exact" chiral actions satisfying the Ginsparg-Wilson relation are numerically very expensive. Therefore, other approaches guaranteeing chirality in the continuum limit have been investigated. Even though the Wilson action does not satisfy the Ginsparg-Wilson relation one expects chirality to be recovered in the continuum limit. However, there can be small eigenvalues of the corresponding lattice Dirac operator which do not belong to the wanted spectrum of continuum fermions and which can be problematic in numerical applications.

In 2000, Roberto Frezzotti et al. [29] developed what has been called "twisted mass"-fermions. In order to eliminate the unwanted zero-modes from the propagator, they introduced an additional mass term to the Wilson action for a 2-flavour theory.

Lets consider the continuum theory for a moment. The twisted mass Dirac operator is given by:

$$D_{\mathrm{tm}} = D + m + \mathrm{i}\mu\gamma_5\tau_3\,, \tag{18.13}$$

where τ_3 is the third Pauli-matrix acting in the 2-dimensional flavour space.

Naively, in the continuum the physical theory should not depend on μ, because a chiral rotation

$$\psi' = \exp\left(\mathrm{i}\alpha\gamma_5\frac{\tau_3}{2}\right)\psi \tag{18.14}$$

$$\bar{\psi}' = \bar{\psi}\exp\left(\mathrm{i}\alpha\gamma_5\frac{\tau_3}{2}\right) \tag{18.15}$$

for $\tan\alpha = \mu/m$ leads to the action:

$$L = \bar{\psi}'(D+M)\psi' \tag{18.16}$$

with

$$M = \sqrt{m^2+\mu^2}\,. \tag{18.17}$$

On the lattice, this twisting of the mass term eliminates unwanted zero modes from the Wilson action, and in the continuum limit one expects the effect of the twisted mass to vanish. Very often one chooses a "maximal twist", defined by $\alpha = \pi/2$ or $\mu/m \to 0$.

18.4 Clover term

As pointed out in the last chapter, the difference between the naive lattice action of fermions (the symmetrized derivative) and the Wilson action is a term involving the second derivative of Fermi-fields. In the continuum, this term corresponds to an operator of fifth order (which vanishes proportional to the lattice constant in a naive continuum limit). When gauge fields are coupled to the fermions, there is a second operator of fifth order which can be coupled to the fermions. In the continuum, this operator correpsonds to a magnetic moment term:

$$S = \bar{\psi}(x)\sigma_{\mu\nu}F^{\mu\nu}\psi(x)\,, \tag{18.18}$$

where $\sigma_{\mu\nu} = -\mathrm{i}/2[\gamma_\mu,\gamma_\nu]$. In general, one can add this magnetic moment term as well as the second derivative of fermions, and in this case one speaks of the "Sheikholeslami-Wohlert"

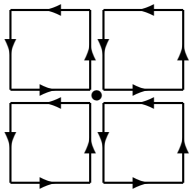
action or "clover action". On the lattice, this clover term is a sum of four oriented loops of links in the $\mu\nu$-plane with center point m:

$$C_{\mu\nu}(m) = \frac{1}{8\mathrm{i}}(P_{\mu\nu}(m) - P^{+}_{\mu\nu}(m))\,, \tag{18.19}$$

where

$$P_{\mu\nu}(m) = \sum_{i=1}^{4} U^{i}_{\mu\nu}(m)\,. \tag{18.20}$$

Each $U^{i}_{\mu\nu}(m_i)$ corresponds to a plaquette term (the product of four oriented group variables around a plaquette in $\mu\nu$ orientation) starting at the point m but going into the four different directions $(+\mu,+\nu),(+\nu,-\mu),(-\mu,-\nu),(-\nu,+\mu)$ (see figure).



clover term with center point •

These additional operators can be interpreted as contributions in the context of improved actions.

Kapitel 19

Quantum field theories at finite temperature

There are mainly three reasons why it is interesting to investigate the standard model at finite temperatures. The first reason is that one hopes to get a better inside about the state and the phases of the Universe immediately after the big bang when the temperatures were still high. It is believed that a new type of phase - the "quark-gluon plasma" (QGP) existed during this brief period where quarks rather behaved like charged particles in a plasma (they were not confined) and chiral symmetry was exact. (We shall see that there can be actually two transitions, a confinement-deconfinement transition as well as a transition from broken to restored chiral symmetry.) Secondly, heavy ion collisions are investigated at different laboratories, amongst others also at the ALICE detector at CERN (where the heavy ions mainly consist of lead nuclei). In heavy ion collisions one wants to investigate experimentally these phase transitions and the nature of the new phases. Lattice gauge theories at finite temperature can help to get a better idea of the observable signatures. Finally, a better understanding of quantum field theories at finite temperature should also yield a better insight into quantum field theories at zero temperature, e.g., concerning mechanisms involved in quark confinement and chiral symmetry breaking.

Many of the following sketchy remarks about quantum field theory at finite temperature on the lattice are taken from a review article by F. Karsch [54]. More details concerning quantum fields on the lattice can be found in the book of Rothe [91].

19.1 The functional integral for quantum systems at finite temperature

Lets consider the case of quantum mechanics and derive a functional integral expression for quantum mechanics at finite temperature in analogy to the functional integral expression for the time evolution operator.

The canonical ensemble for quantum theory at temperature β is defined by the density

matrix

$$\rho_c(\beta, V) = \frac{1}{Z_c} \exp\left(-\beta H(V)\right) , \tag{19.1}$$

where $H(V)$ is the quantum Hamiltonian of the system in a finite volume V, and Z_c is the canonical partition function

$$Z_c = \text{tr } \mathrm{e}^{-\beta H(V)} . \tag{19.2}$$

The trace of the density matrix may be defined in any basis. For simplicity, we take the position basis:

$$\text{tr } \mathrm{e}^{-\beta H} = \int \mathrm{d}x \langle x|\mathrm{e}^{-\beta H}|x\rangle . \tag{19.3}$$

As we did in the case of the unitary time evolution operator, we can write

$$\text{tr } \mathrm{e}^{-\beta H} = \int \mathrm{d}x \langle x| \left(\mathrm{e}^{-(\beta/n)H}\right)^n |x\rangle = \int \prod_{i=1}^{n} \mathrm{d}x_i \langle x_i|\mathrm{e}^{-(\beta/n)H}|x_{i+1}\rangle , \tag{19.4}$$

where $x_n = x$ and $x_{n+1} = x_1$. For n large enough we can approximate

$$\mathrm{e}^{-(\beta/n)H} = \exp\left(-\left(\frac{\beta}{n}\right)\frac{P^2}{2m}\right) \exp\left(-(\beta/n)V(Q)\right) + \mathcal{O}((\beta/n)^2 , \tag{19.5}$$

and express the matrix element by the classical Lagrange-density (compare Sec. 5.1):

$$\begin{aligned} &\text{tr } \mathrm{e}^{-\beta H} = \\ &\quad \lim_{n\to\infty} \mathcal{N} \int \prod_{i=1}^{n} \mathrm{d}x_i \exp\left(-\frac{1}{\hbar}\sum_{i=1}^{n}\left[\frac{1}{2}m\left(\frac{x_{i+1}-x_i}{\Delta\tau}\right)^2 - V(x_i)\right]\Delta\tau\right) \\ &= \int \mathcal{D}x(\tau)\ \exp\left(-\frac{1}{\hbar}\int_0^\beta \left[\frac{1}{2}m\dot{x}(\tau)^2 - V(x(\tau))\right]\mathrm{d}\tau\right) \\ &= \int \mathcal{D}x(\tau)\ \exp\left(-\frac{1}{\hbar}\int_0^\beta L(\dot{x}(\tau), x(\tau))\mathrm{d}\tau\right) . \end{aligned} \tag{19.6}$$

Here, $\Delta\tau = \beta/n$. Because of the trace involved in the partition function, this functional integral extends over all paths which are periodic in τ with period β: $x(0) = x(\beta)$. We can also think of a "torus" of perimeter β on which we consider the classical paths. Note that τ is a parameter which is rather related to a temperature than a time. There is no rotation to a euclidean time involved, but due to the fact that we are dealing with a thermal equilibrium and the canonical partition function, the "euclidean" functional integral is already the correct quantum mechanical expression.

This formalism can easily be extended to quantum fields. For the partition function of quantum fields (in thermal equilibrium) at finite temperature we therefore obtain the functional integral:

$$\text{tr } \mathrm{e}^{-\beta H} = \int \mathcal{D}\varphi(\tau)\ \exp\left(-\frac{1}{\hbar}\int_0^\beta \mathrm{d}\tau \int \mathrm{d}^3x\ \mathcal{L}(\dot{\varphi}(\tau, x), \varphi(\tau, x))\right) . \tag{19.7}$$

Again, the functional integral extends over all fields $\varphi(\tau, x)$ which are periodic in τ with period β. Even though the notation indicates a real scalar field, the same expressions holds for any

quantum field including the gauge degrees of freedom and the fermionic degrees of freedom. There is a slight complication concerning the boundary conditions of fermions: for fermions we have to consider anti-periodic boundary conditions due to the Grassmann nature of the variables.

19.2 Quantum fields at finite temperatures on the lattice

When we discretize the functional integral (eq. 19.7) we have to treat the β-direction (corresponding to the variable τ) separate from the other spacial directions. Therefore, one formulates the model on a $N_\beta \times (N_L)^3$ lattice. The thermodynamic limit corresponds to taking $N_L \to \infty$. If we have assigned a lattice spacing a to the lattice, we have the relation:

$$T = \frac{1}{\beta} = \frac{1}{N_\beta \cdot a}\,, \tag{19.8}$$

while the volume of the system is given by

$$V = (N_L)^3 a^3\,. \tag{19.9}$$

The limits $a \to 0$ and $N_\beta \to \infty$ have to be synchronized such that the temperature remains constant.

In a pure gauge theory one can take the string tension κ to define the lattice spacing and obtains for the temperature:

$$T = \sqrt{\sigma}\frac{1}{\sqrt{\kappa(g)}N_\beta} \tag{19.10}$$

Here, σ is the physical string tension (around $(425\,\mathrm{MeV})^2$) and $\kappa(g)$ is the string tension (eq. 10.16) in lattice units and measured for the lattice action (in lattice units) at coupling constant g. This yields T in units of MeV for a given coupling g and "temporal" lattice extension N_β.

For pure SU(3) gauge theory (or infinite quark masses) the square root of the string tension scales according to (compare Sec. 16)

$$\sqrt{\kappa(g)} \;\propto\; \mathrm{e}^{-\frac{1}{2\beta_0 g^2}} \;\left(\beta_0 g^2\right)^{-\frac{\beta_1}{2\beta_0^2}} \;(1+\mathcal{O}(g)) \tag{19.11}$$

with

$$\beta_0 \;=\; \frac{1}{16\pi^2}\,(11) \qquad \beta_1 \;=\; \frac{1}{(16\pi^2)^2}\,(102)\;. \tag{19.12}$$

If we couple the gauge theory to n_f dynamical quarks, these coefficients change to:

$$\begin{aligned}\beta_0 &= \frac{1}{16\pi^2}\left(11 - \frac{2}{3}n_f\right)\\ \beta_1 &= \frac{1}{(16\pi^2)^2}\left(102 - \left(10 + \frac{8}{3}\right)n_f\right)\;.\end{aligned}$$

The lattice action is the same as for the case $T = 0$, which corresponds to taking $N_\beta = N_L$ and the limit $N_L \to \infty$ before one takes the limit $a \to 0$ (or, to be more exact, $g \to 0$).

However, as many thermodynamic quantities decrease rapidly for large values of N_β (e.g., the pressure decreases as N_β^{-4}) it is very important to keep N_β small and still being close to the continuum limit. This makes improved lattice actions with only small lattice corrections particularly important for lattice calculations at finite temperatures. Many numerical simulations are performed for improved staggered lattice fermions in the quenched approximation.

19.3 Order parameters for phase transitions in lattice QCD

There are many order parameters one can think of which would indicate phase transitions. As we are interested in the deconfining phase transition as well as the chiral phase transition we mainly introduce two order parameters which are relevant for these two types of phase transitions. There seem to be cases where these two phase transitions coincide (amongst others the standard SU(3) QCD gauge theory with two light - up and down - and one medium light - strange - quark flavours), but there are also known cases where they obviously differ.

For the confining-deconfining phase transition a variant of the string tension lends itself as an order parameter. A *Polyakov-loop* is defined as a Wilson-loop which winds once around the torus in β-direction:

$$L(\vec{m}) = \mathrm{tr} \prod_{\tau=1}^{N_\beta} U((\vec{m}, \tau), 0) \tag{19.13}$$

So $\vec{m}$ is the "spacial" point where the Polykaov-loop is located, τ is the "temporal" or "β"-coordinate along the direction of N_β, and the orientation of the links $\mu = 0$ is always along this β-coordinate.

We now consider the correlations function of two Polyakov-loops (oriented in opposite directions, like a large Wilson-loop) in a spacial distance R:

$$G(R) = \langle L(\vec{m}) L^+(\vec{m} + R) \rangle = \exp\left(-\beta F(R)\right) . \tag{19.14}$$

$F(R)$ is the free energy of the system with two (infinite heavy) quarks in a distance R. It is not the ground state of the Hamiltonian, because we take β to be fixed and do not consider the limit $\beta \to \infty$ (or $T \to 0$, where now T is the temperature). Only in this limit the free energy conincides with the ground state of the Hamiltonian.

However, in analogy to the case of zero temperature, we would expect that in the confining phase the free energy tends to infinity in the limit $R \to \infty$, because the energy between the two quarks becomes infinite. In this case $G(R) \to 0$ for $R \to \infty$. If there is no confinement, i.e., if free quarks can exist, the free energy will tend to a constant in this limit and, therefore, also $G(R)$ will tend to a constant in the limit $R \to \infty$. So we come to the conclusion that

$$G(R) \overset{R\to\infty}{\longrightarrow} \begin{cases} 0 & \text{confining phase} \\ > 0 & \text{deconfining phase} \end{cases} \tag{19.15}$$

This defines an order parameter for the confining-deconfining phase transition. For the phase of the QGP we expect $G(R)$ to tend to a non-zero constant for $R \to \infty$ and for the confining phase $G(R)$ tends to zero.

The second phase transition is related to chiral symmetry. For a chirally invariant theory the expectation value

$$\rho = \langle \bar{\psi}(x)\psi(x) \rangle \tag{19.16}$$

should vanish (a chrial transformation can change this combination into its negative). ρ is called the expectation value of the chiral condensate. Therefore, we expect:

$$\langle \bar{\psi}(x)\psi(x) \rangle = \begin{cases} 0 & \text{chrially invariant phase} \\ > 0 & \text{phase of broken chirality} \end{cases} \tag{19.17}$$

Note that ρ can be obtained from the partition function by:

$$\rho = \langle \bar{\psi}\psi \rangle = \frac{1}{V} \frac{\partial Z(\beta, V, m)}{\partial m} \tag{19.18}$$

The mass m acts like a constant external field which couples to the chiral condensate.

19.4 The phase structure of Lattice QCD

As far as pure gauge theory is concerned, the phase structure is easily determined. The asymptotic behavior of the correlation function of Polyakov-loops is the only order parameter, and the only parameter to vary (after the lattice constant a has been fixed using, e.g., the string tension or one of the glueball masses) is the temperature.

Simulations indicate that for the pure SU(3) gauge theory there is a phase transition of first order from a confining to a deconfining phase at around $T = 270\,\mathrm{MeV}$.

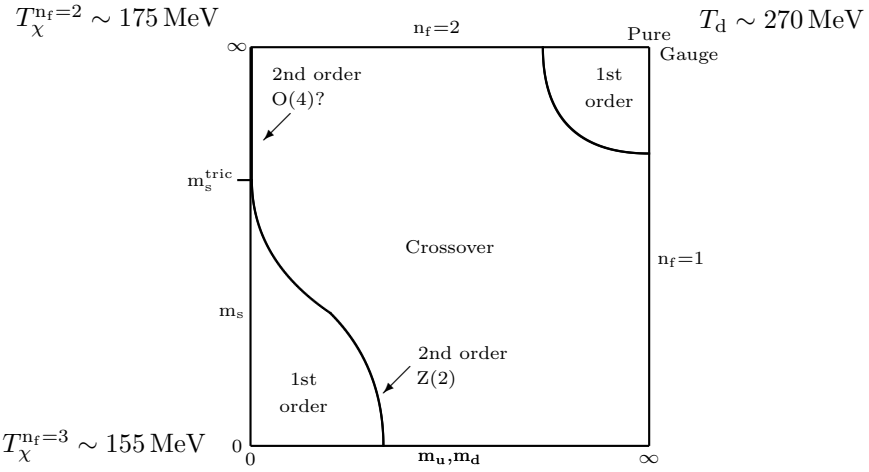


Abbildung 19.1: QCD phase diagram with 3-flavours. u and d-quark masses are degenerate, m_s a parameter. Taken from [55]

The situation gets much more complicated when quarks are coupled to the gauge theory. The main reason is that the phase diagramm depends sensitively on the number of (light) flavours as well as the chosen quark masses. Usually one takes the quark masses of the up and down quarks to be the same ($m_u = m_d$) and chooses a slightly higher mass m_s for the strange quark. This adds two parameters (apart from the number of flavours n_f) to the theory. Fig. 19.1 (taken from [55]) gives an idea of the phase diagram as a function of quark masses.

Kapitel A1

Bündel, Zusammenhang, Krümmung

Oft betrachtet man in der Physik bzw. Mathematik ein System, bei dem an jedem Punkt x einer differenzierbaren Mannigfaltigkeit M noch ein Vektorraum V_x gegeben ist, wobei die Vektorräume zu verschiedenen Punkten isomorph sind. Solch ein System bezeichnet man als „Vektorbündel“. Handelt es sich bei dem Vektorraum um die reellen oder komplexen Zahlen, so spricht man auch von einem (reellen bzw. komplexen) Linienbündel. Handelt es sich um den bei einer differenzierbaren Mannigfaltigkeit immer definierbaren Tangentialraum, so spricht man vom Tangentenbündel TM.

Ein *Vektorbündel* (P, M, V, π) besteht aus einer differenzierbaren Mannigfaltigkeit P, einer differenzierbaren Mannigfaltigkeit M, einem Vektorraum V und einer Abbildung (Projektion) $\pi : P \to M$ mit $\pi^{-1}(x) \sim V$ für alle $x \in M$. Lokal hat P damit die Form $U \times V$ für eine offene Umgebung $U \subset M$.

Zunächst lassen sich Vektoren an verschiedenen Punkten in M nicht miteinander vergleichen, d.h., die Frage, ob zwei Vektoren an verschiedenen Punkten von M „parallel“ oder „gleichlang“ sind, ist im Allgemeinen sinnlos. Für solche Vergleiche bedarf es einer Zusatzstruktur – den „Zusammenhang“. Entgegen der üblichen Vorgehensweise beginnen wir mit einer globalen Betrachtungsweise.

Seien x und y zwei Punkte in M und $\mathcal{C}$ ein Weg, der x mit y verbindet. Dann definiert man den *Paralleltransporter* $P(y, x; \mathcal{C})$ von x nach y entlang $\mathcal{C}$ als lineare Abbildung von V_x (d.h. $\pi^{-1}(x)$) nach V_y (d.h. $\pi^{-1}(y)$). Gilt

$$\boldsymbol{v}(y) \;=\; P(y, x; \mathcal{C})\boldsymbol{v}(x) \quad ,$$

dann bezeichnet man $\boldsymbol{v}(y)$ als den entlang $\mathcal{C}$ parallelverschobenen Vektor zu $\boldsymbol{v}(x)$. Von $P(y, x; \mathcal{C})$ verlangt man dabei die Pseudogruppen-Eigenschaft:

$$P(z, x; \mathcal{C}_1 + \mathcal{C}_2) \;=\; P(z, y; \mathcal{C}_2)\, P(y, x; \mathcal{C}_1) \quad , \tag{A1.1}$$

wobei $\mathcal{C}_1$ ein Weg von x nach y und $\mathcal{C}_2$ ein Weg von y nach z ist. Diese Pseudogruppen-Eigenschaft ermöglicht den Übergang zu einer infinitesimalen Transformation, indem man den Weg $\mathcal{C}$ immer feiner unterteilt. Sei nun $x(t) \sim \mathcal{C}$ ein parametrisierter Weg mit $x(0) = x$, dann

definieren wir:

$$\left.\frac{\mathrm{d}P(x(t),x(0);\mathcal{C})}{\mathrm{d}t}\right|_{t=0} \;=\; \mathrm{i}A_\mu(x)\;\left.\frac{\mathrm{d}x(t)^\mu}{\mathrm{d}t}\right|_{t=0} \quad .$$

$A_\mu(x)$ ist also der Generator des Paralleltransports am Punkte x in Richtung μ (das „i“ ist in der Physik Konvention).

Wählen wir in jedem Vektorraum V_x eine Basis, und sei $U(x)$ eine Basistransformation im Vektorraum V_x, dann transformiert sich P wie

$$P(y,x;\mathcal{C}) \to U(y)^{-1}\,P(y,x;\mathcal{C})\,U(x) \quad .$$

Diese Inhomogenität (P transformiert sich „von rechts“ anders als „von links“) überträgt sich auf die Generatoren A_μ:

$$\begin{aligned}\mathrm{i}A^U_\mu(x)\frac{\mathrm{d}x^\mu}{\mathrm{d}t} &= \frac{\mathrm{d}}{\mathrm{d}t}\,U(x(t))^{-1}\,P(x(t),x)\,U(x)\\ &= \left(\frac{\partial U(x)^{-1}}{\partial x^\mu}U(x)\;+\;U(x)^{-1}\mathrm{i}A_\mu(x)U(x)\right)\frac{\mathrm{d}x^\mu}{\mathrm{d}t}\end{aligned}$$

bzw.

$$A^U_\mu(x) \;=\; U(x)^{-1}A_\mu(x)U(x)\;+\;\mathrm{i}U^{-1}(x)\partial_\mu U(x) \quad .$$

Ausgehend von dem inhomogenen Transformationsverhalten des Paralleltransporters erkennt man die Herkunft des inhomogenen Terms im Transformationsverhalten des Zusammenhangs A_μ.

Seien $\boldsymbol{v}(x)$ und $\boldsymbol{v}(y)$ zwei Vektoren an infinitesimal benachbarten Punkten x (mit Koordinaten x^μ) und y (mit Koordinaten $x^\mu+\delta x^\mu$), dann ist $\boldsymbol{v}(y)$ „parallel“ zu $\boldsymbol{x}$, wenn für die Komponenten gilt:

$$\delta v(y)^a \;=\; \mathrm{i}A^a_{b\mu}v(x)^b\delta x^\mu \;.$$

Im Allgemeinen wird der Paralleltransport von x nach y vom Weg $\mathcal{C}$ abhängen. Insbesondere ist der Paralleltransport entlang eines geschlossenen Weges von x nach x im Allgemeinen nicht trivial. Sei $\mathcal{C}$ ein Weg in der μ-ν-Ebene, der eine infinitesimale Fläche $\delta\sigma^{\mu\nu}$ umschließt, dann gilt für die Komponenten eines Vektors:

$$\delta v(x)^a \;=\; F^a_{b\mu\nu}v(x)^b\delta\sigma^{\mu\nu} \quad . \tag{A1.2}$$

$F^a_{b\mu\nu}$ bezeichnet man als *Krümmungstensor* zum Zusammenhang A. Da sich der Paralleltransporter zu einem geschlossenen Weg bei Koordinatentransformationen homogen transformiert (rechts und links gleich), gilt dies auch für den Krümmungstensor:

$$F^U_{\mu\nu}(x) \;=\; U(x)^{-1}F_{\mu\nu}(x)U(x) \quad .$$

Das bisher Gesagte gilt ganz allgemein. Insbesondere beschreibt es auch den Fall der Eichtheorien. Der Vektorraum an jedem Punkt von M ist der Bildraum der Felder. Da im Allgemeinen in diesem Raum (z.B. dem $\mathbb{C}^n$) ein Skalarprodukt gegeben ist, wird man vom Paralleltransport verlangen, dass er das Skalarprodukt respektiert, d.h., das Skalarprodukt von parallelverschobenen Vektoren bleibt invariant. Damit wird P zu einem Element der Gruppe G, die

das Skalarprodukt invariant lässt, und A_μ zu einem Generator dieser Gruppe, d.h. zu einem Element der Lie-Algebra. Man spricht in diesem Zusammenhang auch von einem Prinzipalbündel mit Strukturgruppe G und assoziiertem Vektorbündel.

Haben wir als Vektorbündel nur ein komplexes Linienbündel (wie im Fall der Elektrodynamik), dann entfallen alle Indizes bezüglich der Vektorraumkoordinaten. Die lokalen Transformationen $U(x)$ (die das Skalarprodukt invariant lassen) werden zu komplexen Phasenmultiplikationen $\mathrm{e}^{\mathrm{i}\alpha(x)}$.

Im Fall der Riemannschen Geometrie gibt es einige Besonderheiten. Die Vektorräume an jedem Punkt x von M sind nun die Tangentialräume T_xM. Die Koordinatensysteme in T_xM sind aber an die Koordinatenwahl auf M geknüpft, d.h., der Unterschied zwischen Vektorraumkoordinaten und Koordinaten auf M entfällt. Außerdem ist bei einer Riemannschen Mannigfaltigkeit in jedem Tangentenraum ein Skalarprodukt $(g_{\mu\nu}(x))$ vorgegeben. Die folgenden zwei Bedingungen legen auf einer Riemannschen Mannigfaltigkeit einen Zusammenhang eindeutig fest:

1. Der Zusammenhang ist symmetrisch: $A^\alpha_{\beta\mu} = A^\alpha_{\mu\beta}$. Diese Bedingung bezeichnet man auch als „Torsionsfreiheit“ des Zusammenhangs. Sie ist bei allgemeinen Vektorbündeln nicht definierbar, da sich die Indizes μ und β auf verschiedene Räume beziehen.

2. Der Zusammenhang respektiert das Skalarprodukt.

Der so eindeutig definierte Zusammenhang ist der Levi-Civita-Zusammenhang $\Gamma^\alpha_{\beta\mu}$. Die zugehörige Krümmung bei einem Paralleltransport um einen infinitesimalen geschlossenen Weg in der μ-ν-Ebene ist der Riemannsche Krümmungstensor $R^\alpha_{\beta\mu\nu}$.

Kapitel A2

Gaußsche Integrale

Die Eigenschaften und Kombinatorik Gaußscher Integrale bilden einen wesentlichen Bestandteil der Störungsrechung in der Quantenfeldtheorie. Die wichtigsten Formeln sollen für den endlichdimensionalen Fall wiederholt werden. Wir beschränken uns in diesem Kapitel auf „bosonische" Integrationsvariable, wobei wir wegen der teilweise unterschiedlichen Kombinatorik den reellen und komplexen Fall getrennt behandeln. Gaußsche Integrale über Graßmann-Variable werden in Kapitel 17 behandelt.

A2.1 Reelle Gaußsche Integrale

Gaußsches Integral!reell

Sei A eine reelle, symmetrische, positive $N \times N$-Matrix, d.h.

$$A_{ij} \in \mathbb{R} \qquad A_{ij} = A_{ji} \qquad \sum_{i,j=1}^{N} x_i A_{ij} x_j \; > \; 0 \;\; \text{für } \boldsymbol{x} \neq 0 \;\; . \tag{A2.1}$$

In diesem Abschnitt benutze ich die Notation

$$(\boldsymbol{x}, \boldsymbol{y}) \;=\; \sum_i x_i y_i \qquad (\boldsymbol{x}, A\boldsymbol{y}) \;=\; \sum_{i,j} x_i A_{ij} y_j \;\; . \tag{A2.2}$$

Es soll zunächst das Gaußsche Integral

$$Z \;=\; \int \prod_{i=1}^{N} \mathrm{d}x_i \;\; \exp(-\tfrac{1}{2}(\boldsymbol{x}, A\boldsymbol{x})) \tag{A2.3}$$

berechnet werden.

Da A symmetrisch ist, lässt sich die Matrix durch eine orthogonale Transformation diagonalisieren. Die Positivität von A bedeutet, dass sämtliche Eigenwerte λ_i positiv sind. Sei

$$SAS^{+} \;=\; D \quad \text{mit} \quad D = \mathrm{diag}\,(\lambda_1, \ldots, \lambda_N) \qquad \text{und} \quad \boldsymbol{y} \;=\; S\boldsymbol{x} \;\; , \tag{A2.4}$$

so folgt:

$$Z \;=\; \int \prod_i \mathrm{d}y_i \;\exp(-\frac{1}{2}\sum_i \lambda_i y_i^2) \quad . \tag{A2.5}$$

(Man beachte, dass die Jacobi-Determinante der Transformation $\boldsymbol{x} \to \boldsymbol{y} = S\boldsymbol{x}$ eins ist, da S eine orthogonale Transformation ist.)

Das Integral faktorisiert nun in den Argumenten $\{y_i\}$, und man erhält das Produkt von 1-dimensionalen Gaußschen Integralen:

$$Z \;=\; \prod_{i=1}^{N} \left(\frac{2\pi}{\lambda_i}\right)^{1/2} \;=\; \frac{(2\pi)^{N/2}}{(\det\, A)^{1/2}} \quad . \tag{A2.6}$$

Erwartungswerte bzgl. des Gaußschen Maßes sind durch

$$\langle F[\{x_k\}]\rangle \;=\; \frac{1}{Z} \int \prod_i \mathrm{d}x_i \; F[\{x_k\}] \;\exp(-\tfrac{1}{2}(\boldsymbol{x}, A\boldsymbol{x})) \tag{A2.7}$$

definiert. Zur Berechnung solcher Erwartungswerte bedient man sich oft der sogenannten „erzeugenden Funktion“:

$$Z[\{b_i\}] \;:=\; \langle \exp(\boldsymbol{b}, \boldsymbol{x})\rangle \;=\; \frac{1}{Z} \int \prod_i \mathrm{d}x_i \;\exp\left(-\tfrac{1}{2}(\boldsymbol{x}, A\boldsymbol{x}) + (\boldsymbol{b}, \boldsymbol{x})\right) \quad . \tag{A2.8}$$

Durch die lineare Transformation

$$x_i \;=\; y_i + \sum_j (A^{-1})_{ij} b_j$$

entkoppeln im Exponenten von (A2.8) die Integrationsvariablen $\{y_i\}$ und der konstante Vektor $\{b_i\}$, und man erhält:

$$Z[\{b_i\}] \;=\; \exp\left(\tfrac{1}{2}(\boldsymbol{b}, A^{-1}\boldsymbol{b})\right) \quad . \tag{A2.9}$$

Dieses erzeugende Funktional lässt sich nun zur Berechnung von Erwartungswerten von Funktionen von x_i heranziehen. So gilt für die 2-Punkt-Funktion

$$\langle x_k x_l\rangle \;=\; \left.\frac{\partial^2 Z[\boldsymbol{b}]}{\partial b_k \partial b_l}\right|_{b=0} \;=\; (A^{-1})_{kl} \quad . \tag{A2.10}$$

Entsprechend erhält man für die 4-Punkt-Funktion:

$$\begin{aligned} \langle x_i x_j x_k x_l\rangle \;&=\; \left.\frac{\partial^4 Z[\boldsymbol{b}]}{\partial b_i \partial b_j \partial b_k \partial b_l}\right|_{b=0} \\ &=\; (A^{-1})_{ij}(A^{-1})_{kl} \;+\; (A^{-1})_{ik}(A^{-1})_{jl} \;+\; (A^{-1})_{il}(A^{-1})_{jk} \\ &=\; \langle x_i x_j\rangle\langle x_k x_l\rangle \;+\; \langle x_i x_k\rangle\langle x_j x_l\rangle \;+\; \langle x_i x_l\rangle\langle x_j x_k\rangle \,. \end{aligned} \tag{A2.11}$$

Es ist eine charakterisierende Eigenschaft der Gaußschen Verteilung, dass sich alle Momente von $\{x_i\}$ durch die 2-Punkt-Funktion darstellen lassen. Ganz allgemein verschwinden ungerade Momente, während für gerade Momente das sogenannte *Wicksche Theorem* gilt:

$$\langle x_{i_1} \ldots x_{i_{2n}}\rangle \;=\; \sum_{\text{Paarungen}} \prod^{n} \langle x_{i_k} x_{i_l}\rangle \quad , \tag{A2.12}$$

wobei über alle Möglichkeiten zu summieren ist, die Indexmenge $\{1, 2, \ldots, 2n\}$ in Paare (d.h. zwei-elementige Teilmengen) zu zerlegen, und jeweils das Produkt aller 2-Punkt-Funktionen zu diesen Paaren zu bilden ist.

Die inverse Matrix A^{-1} wird in Funktionenräumen auch oft die „Greensche Funktion“ eines Operators genannt.

A2.2 Komplexe Gaußsche Integrale

Sei nun $z_i \in \mathbb{C}$ mit der reellen Zerlegung

$$z_i \;=\; \frac{1}{\sqrt{2}}(x_i + iy_i) \quad \text{und} \quad \bar{z}_i \;=\; \frac{1}{\sqrt{2}}(x_i - iy_i) \;\; .$$

Außerdem sei A eine hermitesche, positive $N \times N$-Matrix, d.h.

$$A_{ij} = A^*_{ji} \qquad\qquad \sum_{i,j=1}^{N} \bar{z}_i A_{ij} z_j \;>\; 0 \;\; \text{für } \boldsymbol{z} \neq 0 \;\; . \tag{A2.13}$$

Für das Maß gilt

$$\mathrm{d}\bar{z}\,\mathrm{d}z \;=\; \mathrm{d}x\,\mathrm{d}y \;\; .$$

Zur Berechnung des Gaußschen Integrals

$$Z \;=\; \int \prod_{i=1}^{N} \mathrm{d}\bar{z}_i\,\mathrm{d}z_i \;\exp\left(-\sum_{i,j} \bar{z}_i A_{ij} z_j\right) \tag{A2.14}$$

wird A diesmal durch eine unitäre Transformation diagonalisiert, und man erhält:

$$\begin{aligned} Z \;&=\; \int \prod_i \mathrm{d}\bar{z}_i\,\mathrm{d}z_i \;\exp\left(-\sum_i \lambda_i\,\bar{z}_i z_i\right) \\ &=\; \int \prod_i \mathrm{d}x_i\,\mathrm{d}y_i \;\exp\left(-\tfrac{1}{2}\sum_i \lambda_i\,(x_i^2 + y_i^2)\right) \\ &=\; \frac{(2\pi)^N}{\det A} \;\; . \end{aligned} \tag{A2.15}$$

Zur Bestimmung von Erwartungswerten

$$\langle F[\{\bar{z}_k, z_k\}]\rangle \;=\; \frac{1}{Z} \int \prod_i \mathrm{d}z_i\,\mathrm{d}\bar{z}_i \;F[\{\bar{z}_k, z_k\}] \;\exp(-\textstyle\sum_{ij} \bar{z}_i A_{ij} z_j) \tag{A2.16}$$

(die sich natürlich ebenfalls in ihre reellen Bestandteile $\{x_k, y_k\}$ zerlegen lassen) berechnen wir wiederum die „erzeugende Funktion“:

$$\begin{aligned} Z[\{\bar{b}_k, b_k\}] \;&=\; \left\langle \exp \sum_i (\bar{z}_i b_i + \bar{b}_i z_i) \right\rangle \\ &=\; \frac{1}{Z} \int \prod_i \mathrm{d}\bar{z}_i\,\mathrm{d}z_i \;\exp\left(-\sum_{i,j} \bar{z}_i A_{ij} z_j \;+\; \sum_i (\bar{z}_i b_i + \bar{b}_i z_i)\right) \;\; . \end{aligned} \tag{A2.17}$$

Nach einer linearen Transformation

$$z_i \;=\; z_i' + \sum_j (A^{-1})_{ij} b_j \tag{A2.18}$$

entkoppeln die Integrationsvariablen wieder, und man erhält:

$$Z[\{\bar{b}_i, b_i\}] \;=\; Z[\bar{\boldsymbol{b}}, \boldsymbol{b}] \;=\; \exp\left(\sum_{i,j} \bar{b}_i (A^{-1})_{ij} b_j\right) \quad . \tag{A2.19}$$

Zur Berechnung von n-Punkt-Funktionen lassen sich $\bar{\boldsymbol{b}}$ und $\boldsymbol{b}$ wie unabhängige Variable behandeln, d.h. es gilt:

$$\langle F[\{\bar{z}_k, z_k\}]\rangle \;=\; F\left[\left\{\tfrac{\partial}{\partial b_k}, \tfrac{\partial}{\partial \bar{b}_k}\right\}\right] Z[\bar{\boldsymbol{b}}, \boldsymbol{b}]\Big|_{b=\bar{b}=0} \tag{A2.20}$$

Für die 2-Punkt-Funktion erhält man wieder

$$\langle \bar{z}_k z_l\rangle \;=\; (A^{-1})_{kl} \quad . \tag{A2.21}$$

Aus dieser Gleichung erklärt sich der unterschiedliche Faktor $\frac{1}{2}$ im Exponenten von (A2.3) und (A2.14) bzw. im Exponenten der erzeugenden Funktionen (A2.8) und (A2.17). Der Erwartungswert der 2-Punkt-Funktion ist in beiden Fällen einfach das Matrixelement der inversen Matrix. Auch an (A2.15) erkennt man, dass bei dieser Wahl der Normierung ein komplexes Feld zwei reellen Feldern entspricht.

Für die 4-Punkt-Funktion folgt nun:

$$\langle \bar{z}_i \bar{z}_j z_k z_l\rangle \;=\; \langle \bar{z}_i z_k\rangle\langle \bar{z}_j z_l\rangle \;+\; \langle \bar{z}_i z_l\rangle\langle \bar{z}_j z_k\rangle \quad . \tag{A2.22}$$

Ganz allgemein verschwinden Erwartungswerte, bei denen die Anzahl der $\bar{z}$'s ungleich der Anzahl der z's ist. (Beweis: Unter der Variablentransformation $z_i \to \mathrm{e}^{i\alpha} z_i$, $\bar{z}_i \to \mathrm{e}^{-i\alpha}\bar{z}_i$ ändert sich das Integral nicht, d.h. auch die Erwartungswerte müssen unter dieser Transformation invariant sein. Das gilt jedoch nur, wenn die Anzahl der z's und $\bar{z}$'s gleich ist.) Für Momente mit einer gleichen Anzahl von $\bar{z}$'s und z's gilt wieder das *Wicksche Theorem*:

$$\begin{aligned}\langle \bar{z}_{i_1} \ldots \bar{z}_{i_n} z_{j_1} \ldots z_{j_n}\rangle \;&=\; \sum_\sigma \prod_k^n \langle \bar{z}_{i_k} z_{j_{\sigma(k)}}\rangle \\ &=\; \sum_\sigma \langle \bar{z}_{i_1} z_{j_{\sigma(1)}}\rangle \ldots \langle \bar{z}_{i_n} z_{j_{\sigma(n)}}\rangle\end{aligned}$$

$\sum_\sigma$ bedeutet dabei eine Summation über alle Permutationen σ von n Elementen.

Kapitel A3

Summation über Wege auf beliebigen Graphen

In diesem Anhang soll gezeigt werden, dass die 2-Punkt-Korrelationsfunktion des (euklidischen) skalaren Feldes auf einem beliebigen Graphen eine Darstellung als Summation über Wege besitzt. Ein Graph ist im folgenden definiert als eine Menge V (die Vertizes) und eine symmetrische, nicht-reflexive Relation E auf V. Die Elemente $\{m,n\} \in E$ mit $m,n \in V$ bezeichnet man als die Linien des Graphen.

Jeder Graph lässt sich durch seine Adjazenzmatrix

$$A_{mn} \;=\; \begin{cases} 1 & m,n \in V \text{ sind benachbart, d.h. } \{m,n\} \in E \\ 0 & \text{sonst} \end{cases}$$

charakterisieren. Sei $\varphi : V \to \mathbb{R}$ ein reelles skalares Feld auf dem Graphen. Wir definieren die Wirkung:

$$\begin{aligned} S[\varphi] \;&=\; \frac{1}{4}\sum_{m,n} A_{mn}(\varphi_m - \varphi_n)^2 \;+\; \frac{1}{2}M^2 \sum_m \varphi_m^2 \\ &=\; \frac{1}{2}\sum_{m,n} \varphi_m (V_{mn} - A_{mn} + M^2\delta_{mn})\varphi_n \;. \end{aligned}$$

Hierbei ist

$$V_{mn} \;=\; v_m\,\delta_{mn} \qquad \left(\text{mit } \; v_m \;=\; \sum_n A_{mn}\right)$$

die Valenzmatrix zu A. V ist eine Diagonalmatrix, und v_m ist die Anzahl der Nachbarpunkte von m. Die Matrix

$$K_{mn} \;:=\; V_{mn} - A_{mn} + M^2\delta_{mn}$$

ist offensichtlich positiv definit, da die Wirkung für alle $\varphi \neq 0$ positiv ist. Die 2-Punkt-Korrelationsfunktion ist durch die Elemente der inversen Matrix gegeben:

$$\langle \varphi_m \varphi_n \rangle \;:\; = \;\; \frac{1}{Z}\int \prod_{i\in V} \mathrm{d}\varphi_i \;\varphi_m \varphi_n \;\mathrm{e}^{-S[\varphi]}$$

$$= \quad (K^{-1})_{mn} \ .$$

Wir leiten nun die „Summation über Wege"-Darstellung von K^{-1} ab. Ausgehend von

$$K \ = \ V \ + \ M^2 \ - \ A \ = \ (\mathbb{1} \ - \ A(V + M^2)^{-1})(V \ + \ M^2)$$

erhalten wir:

$$K^{-1} \ = \ (V + M^2)^{-1} \ \sum_{N=0}^{\infty} [A(V + M^2)^{-1}]^N \ . \tag{A3.1}$$

Der Einfachheit halber betrachten wir zunächst den Spezialfall eines regulären Graphen, d.h. $V = v\mathbb{1}$. In diesem Fall ist

$$K^{-1} \ = \ (v + M^2)^{-1} \ \sum_{N=0}^{\infty} A^N \mathrm{e}^{-\mu N}$$

bzw.

$$(K^{-1})_{mn} \ = \ (v + M^2)^{-1} \ \sum_{N=0}^{\infty} (A^N)_{mn} \ \mathrm{e}^{-\mu N}$$

mit

$$\mathrm{e}^{\mu} \ = \ v + M^2 \ .$$

(Man vergleiche diese Relation mit (8.18)). $(A^N)_{mn}$ ist gleich der Anzahl der Wege der Länge N von n nach m. Um dies einzusehen, betrachten wir das N-fache Matrixprodukt von A:

$$(A^N)_{mn} \ = \ \sum_{i_1,\ldots,i_{N-1}} A_{mi_1} A_{i_1 i_2} A_{i_2 i_3} \ldots A_{i_{N-1} n} \ .$$

Nun ist aber

$$A_{mi_1} A_{i_1 i_2} \ldots A_{i_{N-1} n} \ = \ \begin{cases} 1 & \text{falls } (m, i_1, i_2, \ldots, i_{N-1}, n) \text{ Weg auf Graph} \\ 0 & \text{sonst .} \end{cases}$$

Jeder Weg wird zusätzlich mit einem Faktor gewichtet, der nur von seiner Länge abhängt. Dies ist somit die Verallgemeinerung von Formel (8.13) für beliebige reguläre Graphen.

Wir betrachten nun den allgemeinen Fall. Es gilt:

$$\begin{aligned} \big((V + M^2)^{-1}[A(V + M^2)^{-1}]^N\big)_{mn} \ &= \\ &= \sum_{i_1,\ldots,i_{N-1}} \frac{1}{v_m + M^2} A_{mi_1} \frac{1}{v_{i_1} + M^2} A_{i_1 i_2} \ldots A_{i_{N-1} n} \frac{1}{v_n + M^2} \\ &= \sum_{i_1,\ldots,i_{N-1}} A_{mi_1} A_{i_1 i_2} \ldots A_{i_{N-1} n} \prod_{k=0}^{N} \frac{1}{v_{i_k} + M^2} \\ &\qquad\qquad \text{mit} \quad i_0 = m, i_N = n \\ &= \sum_{\text{Wege}} \omega(m, i_1, \ldots, i_{N-1}, n) \ . \end{aligned}$$

Jeder Weg erhält somit das Gewicht

$$\begin{aligned}\omega(m, i_1, \ldots, i_{N-1}, n) &= \prod_{k=0}^{N} \frac{1}{v_{i_k} + M^2} \\ &= \exp\left(- \sum_{k=0}^{N} \ln(v_{i_k} + M^2) \right) .\end{aligned}$$

Während (AV^{-1}) eine Interpretation als Markov-Matrix hat (vgl. Kap. 14) und die Summation über Wege für $M = 0$ somit auch als Markov-Prozess gedeutet werden kann (allerdings mit Konvergenzproblemen), ist $(A + M^2)^{-1}$ keine Markov-Matrix, da

$$\sum_m A_{mn} \frac{1}{v_n + M^2} = \frac{v_n}{v_n + M^2} < 1 .$$

Die Interpretation einer Wahrscheinlichkeit kann jedoch beibehalten werden, wenn wir an jedem Punkt eine gewisse „Stop-Wahrscheinlichkeit“ berücksichtigen: Der Zufallsweg kann an jedem Punkt n mit der Wahrscheinlichkeit $1/(v_n + M^2)$ zu einem der v_n Nachbarpunkte führen, d.h. mit der Wahrscheinlichkeit $v_n/(v_n + M^2)$ weiterlaufen, oder mit der Wahrscheinlichkeit $M^2/(v_n + M^2)$ enden. Das Matrixelement $(K^{-1})_{mn}$ lässt sich dann als Summe über alle solche Prozesse definieren, die bei n anfangen und bei m enden.

Kapitel A4

Lösung des 2-dimensionalen Ising-Modells

Das 2-dimensionale Ising-Modell (ohne äußeres Magnetfeld) gilt immer noch als der „Prototyp“ der nicht-trivialen, exakt lösbaren Modelle. „Exakt lösbar“ bedeutet in diesem Fall, dass die Eigenwerte der Transfermatrix (für eine beliebige Gittergröße) bestimmt werden können. Damit ist insbesondere die freie Energie im thermodynamischen Grenzfall bekannt. Sie zeigt einen Phasenübergang zweiter Ordnung mit logarithmischer Divergenz der spezifischen Wärme.

Die Erwartungswerte lassen sich jedoch im Allgemeinen nicht als exakten geschlossenen Ausdruck angeben. Am kritischen Punkt beschreibt das 2-dim. Ising-Modell masselose Fermionen, und man kann beispielsweise mit Methoden der konformen Feldtheorie die n-Punkt-Funktion bestimmen. Die kritischen Exponenten sind daher bekannt. In der Umgebung des kritischen Punktes reduziert sich das Ising-Modell im Rahmen eines geeigneten Kontinuumslimes zu einer Theorie massiver Fermionen, deren n-Punkt-Funktionen ebenfalls berechnet werden können (siehe die Arbeiten von Sato et. al [93] und Itzykson et. al [4, 102])). Außerdem kennt man einen geschlossenen Ausdruck für den Betrag der Magnetisierung (Wurzel aus dem asymptotischen Wert der 2-Punkt-Funktion).

Die erste exakte Bestimmung der Freien Energie des 2-dimensionalen Ising-Modells ohne äußeres Magnetfeld stammt von Onsager [84] (1944), der mit Hilfe geschickter algebraischer Manipulationen die Transfermatrix diagonalisieren konnte. Es folgten eine Fülle weiterer Arbeiten mit teilweise einfacheren bzw. alternativen Lösungswegen (siehe auch das Buch von Baxter [5]). Eine ältere Zusammenfassung (1953) findet man in dem Artikel von Newell und Montroll [80].

Der hier beschrittene Lösungsweg des 2-dimensionalen Ising-Modells folgt im wesentlichen dem Artikel von Lieb, Mattis und Schultz [65]. Für die Freie Energie pro Freiheitsgrad des Ising-Modells im Grenzfall unendlichen Volumens benötigt man nur den maximalen Eigenwert der Transfermatrix. Trotzdem ist es interessant, dass sich für das Ising-Modell sämtliche Eigenwerte (auch für endliches Volumen) bestimmen lassen.

A4.1 Definition der Transfermatrix

Wir betrachten das 2-dimensionale Ising-Modell ohne äußeres Magnetfeld auf einem $N \times M$ Gitter mit periodischen Randbedingungen. Die Kopplungen auf „raum-artigen“ und „zeit-artigen“ Gitterlinien werden zunächst als verschieden angenommen. Die Energie (einschließlich der reduzierten Temperaturen) ist

$$S[\{s\}] \;=\; -\;\beta_t \sum_{m,n} s(m,n)s(m+1,n) \;-\; \beta_x \sum_{m,n} s(m,n)s(m,n+1)\;.$$

Der 2^N dimensionale Vektorraum der raumartigen Konfigurationen hat die Konfigurationsraumbasis

$$|\{s\}\rangle \;=\; |s(1), s(2), \ldots, s(N)\rangle \;=\; |s\rangle_1 \otimes \ldots \otimes |s\rangle_N \qquad\qquad s(i) \;=\; \pm 1\;.$$

Wir wählen

$$|+\rangle \;\cong\; \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad \text{und} \qquad |-\rangle \;\cong\; \begin{pmatrix} 0 \\ 1 \end{pmatrix}\;.$$

In diesem Vektorraum definieren wir die Operatoren

$$\begin{aligned} \Sigma_3(n) \;&=\; \mathbb{1} \otimes \mathbb{1} \otimes \cdots \otimes \sigma_3 \otimes \cdots \otimes \mathbb{1} \qquad & \Sigma_3(N+1) \;&=\; \Sigma_3(1) \\ \Sigma_1(n) \;&=\; \mathbb{1} \otimes \mathbb{1} \otimes \cdots \otimes \sigma_1 \otimes \cdots \otimes \mathbb{1} \qquad & \Sigma_1(N+1) \;&=\; \Sigma_1(1) \end{aligned}$$

Dabei sind

$$\sigma_3 \;=\; \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \qquad \text{und} \qquad \sigma_1 \;=\; \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

die Pauli-Matrizen, die an der n.ten Stelle im Tensorprodukt stehen. $\Sigma_3(n)$ und $\Sigma_1(n)$ erfüllen die Vertauschungsrelationen:

$$\begin{aligned} &[\Sigma_3(m), \Sigma_1(n)] \;=\; [\Sigma_3(m), \Sigma_3(n)] \;=\; [\Sigma_1(m), \Sigma_1(n)] \;=\; 0 \quad \text{für } m \neq n \\ &\text{und} \qquad \{\Sigma_3(m), \Sigma_1(m)\} \;\equiv\; \Sigma_3(m)\Sigma_1(m) + \Sigma_3(m)\Sigma_1(m) \;=\; 0\;. \end{aligned} \tag{A4.1}$$

Diese Matrizen generieren die Algebra aller $2^N \times 2^N$–Matrizen.

Wir definieren die Transfermatrix durch Angabe ihrer Matrixelemente in der Konfigurationsraumbasis:

$$\langle\{s'\}|\hat{T}|\{s\}\rangle \;=\; \exp\left(\beta_t \sum_n s(n)s'(n) \;+\; \beta_x \sum_n s(n)s(n+1)\right)\;.$$

Im Gegensatz zu der allgemeinen Definition aus Kapitel 8 ist diese Matrix nicht symmetrisch. Wir werden jedoch sehen, dass sie durch eine Ähnlichkeitstransformation, die das Spektrum nicht ändert, zu einer symmetrischen Matrix gemacht werden kann.

A4.2 Die Transfermatrix als Funktion von $\Sigma_3(n)$ und $\Sigma_1(n)$

Wir wollen nun die Transfermatrix als Funktion der Operatoren $\Sigma_3(n)$ und $\Sigma_1(n)$ schreiben. Da $\Sigma_3(n)$ in der Konfigurationsraumbasis diagonal ist gilt

$$\exp\left(\beta_x \sum_n \Sigma_3(n)\Sigma_3(n+1)\right) |s(1),\ldots,s(N)\rangle \;=\; = \;\exp\left(\beta_x \sum_n s(n)s(n+1)\right) |s(1),\ldots,s(N)\rangle \;.$$

Für den nicht-diagonalen Anteil betrachten wir zunächst die 2×2-Matrix (die Transfermatrix des 1-dimensionalen Ising-Modells)

$$\langle s|T_1|s'\rangle \;=\; \mathrm{e}^{\beta_t s s'} \qquad \text{d.h.} \qquad T_1 \;=\; \begin{pmatrix} \mathrm{e}^{\beta_t} & \mathrm{e}^{-\beta_t} \\ \mathrm{e}^{-\beta_t} & \mathrm{e}^{\beta_t} \end{pmatrix} .$$

Ausgedrückt durch die σ-Matrizen gilt:

$$T_1 \;=\; \mathrm{e}^{\beta_t}\mathbb{1} \;+\; \mathrm{e}^{-\beta_t}\sigma_1 \;=\; \mathrm{e}^{\beta_t}\left(\mathbb{1} \;+\; \mathrm{e}^{-2\beta_t}\sigma_1\right) .$$

Diese Form können wir mit der folgenden allgemeinen Darstellung der Exponentialfunktion von σ_1 vergleichen:

$$\mathrm{e}^{\beta_t^*\sigma_1} \;=\; \cosh\beta_t^* \;+\; \sigma_1 \sinh\beta_t^* \;=\; \cosh\beta_t^*\,(\mathbb{1} \;+\; \tanh\beta_t^*\sigma_1) \;.$$

Wir definieren die duale Temperatur β^* durch

$$\tanh\beta^* \;=\; \mathrm{e}^{-2\beta} \qquad \text{bzw.} \qquad \beta^* \;=\; \frac{1}{2}\ln\tanh\beta \;. \tag{A4.2}$$

Diese Transformation hat die Eigenschaft $(\beta^*)^* = \beta$, was die Bezeichnung „dual“ rechtfertigt.

Die duale Transformation hat einen Fixpunkt β_c. Setzt man in (A4.2) $\beta = \beta^*$, so erhält man eine quadratische Gleichung für $\exp -2\beta$, die sich leicht lösen lässt:

$$\beta_c \;=\; \frac{1}{2}\ln(1+\sqrt{2}) \;\approx\; 0.44068679351 \;.$$

Diese Temperatur ist gleichzeitig die Temperatur des Phasenübergangs des Ising-Modells. Nimmt man an, dass das Ising-Modell nur einen Phasenübergang besitzt, so muss dies der Fixpunkt der Dualitätstransformation sein.

Wir erhalten also

$$T_1 \;=\; \mathrm{e}^{\beta_t}\,(\cosh\beta_t^*)^{-1}\;\mathrm{e}^{\beta_t^*\sigma_1} \;=\; \sqrt{2\sinh 2\beta_t}\;\mathrm{e}^{\beta_t^*\sigma_1} \;.$$

Der gemischte Anteil der Transfermatrix des 2-dimensionalen Ising-Modells ist offensichtlich das N-fache Tensorprodukt dieser Matrix. Da außerdem die $\Sigma_1(n)$'s an verschiedenen Punkten kommutieren, gilt

$$\hat{T} \;=\; (2\sinh 2\beta_t)^{N/2}\;\hat{V}_1\hat{V}_2$$

mit

$$\hat{V}_1 \;=\; \mathrm{e}^{\beta_t^* \sum_n \Sigma_1(n)} \qquad \text{und} \qquad \hat{V}_2 \;=\; \mathrm{e}^{\beta_x \sum_n \Sigma_3(n)\Sigma_3(n+1)} \;.$$

Die symmetrisierte Transfermatrix erhält man durch eine Ähnlichkeitstransformation mit $\hat{V}_2^{1/2}$.

Für die folgenden Überlegungen wird es einfacher sein, die Rollen von $\Sigma_3(n)$ und $\Sigma_1(n)$ (d.h σ_1 und σ_3) zu vertauschen. Dies können wir durch eine unitäre Transformation von T erreichen. Es gilt:

$$u\sigma_1 u^+ \;=\; \sigma_3 \qquad \text{und} \qquad u\sigma_3 u^+ \;=\; \sigma_1 \qquad \text{mit} \quad u \;=\; \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \;.$$

Wir definieren die transformierte Transfermatrix

$$T \;=\; U\hat{T}U^+ \;=\; V_1 V_2 \qquad\qquad U \;=\; u \otimes \ldots \otimes u$$

mit

$$V_1 \;=\; \mathrm{e}^{\beta_t^* \sum_n \Sigma_3(n)} \qquad \text{und} \qquad V_2 \;=\; \mathrm{e}^{\beta_x \sum_n \Sigma_1(n)\Sigma_1(n+1)} \;.$$

A4.3 Die Transfermatrix als Funktion von Fermion-Operatoren

Die Transfermatrix ist invariant unter der Ersetzung

$$\Sigma_3(n) \;\to\; \Sigma_3(n+1) \qquad \text{und} \qquad \Sigma_1(n) \;\to\; \Sigma_1(n+1) \;.$$

Man könnte daher auf die Idee kommen, die Matrizen $\Sigma_3(m)$ und $\Sigma_1(m)$ nach Fourier-Moden zu zerlegen, um die Kopplung zwischen verschiedenen Gitterpunkten in V_2 zu beseitigen und das Problem auf die Diagonalisierung von 2×2-Matrizen zu reduzieren. Allerdings sind die gemischten Kommutatorbeziehungen (Gl. (A4.1)) nicht invariant und würden für die Fouriertransformierten auch nicht-triviale Kommutatoren zu verschiedenen Moden ergeben.

Die sogenannte Jordan-Wigner-Transformation bietet eine Möglichkeit, von den Matrizen $\Sigma_3(n)$ und $\Sigma_1(n)$ zu neuen Matrizen $\Psi^+(n)$ und $\Psi(n)$ mit „standard“ Anti-Kommutatorrelationen zu gelangen. Wir definieren

$$\begin{aligned}
\Psi^+(n) &= (-\sigma_3) \otimes (-\sigma_3) \otimes \ldots \otimes \sigma^+ \otimes \mathbb{1} \otimes \ldots \otimes \mathbb{1} \\
&= \left(\prod_{i=1}^{n-1} \mathrm{e}^{i\pi \Sigma^+(i)\Sigma^-(i)} \right) \Sigma^+(n) \\
\Psi(n) &= (-\sigma_3) \otimes (-\sigma_3) \otimes \ldots \otimes \sigma^- \otimes \mathbb{1} \otimes \ldots \otimes \mathbb{1} \\
&= \left(\prod_{i=1}^{n-1} \mathrm{e}^{i\pi \Sigma^+(i)\Sigma^-(i)} \right) \Sigma^-(n) \;,
\end{aligned}$$

mit

$$\begin{aligned}
\Sigma^+(n) &= \mathbb{1} \otimes \ldots \otimes \mathbb{1} \otimes \sigma^+ \otimes \mathbb{1} \otimes \ldots \otimes \mathbb{1} \\
\Sigma^-(n) &= \mathbb{1} \otimes \ldots \otimes \mathbb{1} \otimes \sigma^- \otimes \mathbb{1} \otimes \ldots \otimes \mathbb{1} \;,
\end{aligned}$$

und

$$\sigma^+ \;=\; \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \qquad \text{und} \qquad \sigma^- \;=\; \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \;.$$

Die Vertauschungsrelationen der Sigma-Matrizen

$$\{\sigma^+,\sigma^-\} \;=\; 1 \qquad\qquad \{\sigma^+,\sigma_3\} \;=\; 0 \qquad\qquad \{\sigma^-,\sigma_3\} \;=\; 0$$

ergeben als Kommutatorrelationen für die Matrizen Ψ und Ψ^+:

$$\{\Psi^+(m),\Psi^+(n)\} \;=\; \{\Psi(m),\Psi(n)\} \;=\; 0 \qquad \{\Psi^+(m),\Psi(n)\} \;=\; \delta_{mn} \;.$$

Dies sind die Vertauschungsrelationen von fermionischen Operatoren. Die Matrizen Ψ^+ und Ψ bilden ebenfalls eine Basis für alle $2^N \times 2^N$-Matrizen. Insbesondere gilt:

$$\begin{aligned} & & \Psi^+(n)\Psi(n) - \Psi(n)\Psi^+(n) &= \Sigma_3(n) \\ \text{und} \qquad & (\Psi^+(n) - \Psi(n))\;(\Psi^+(n+1) + \Psi(n+1)) &= \Sigma_1(n)\Sigma_1(n+1) \qquad \text{für } n \neq N \;. \end{aligned}$$

Die Relation ist für $n = N$ leider etwas komplizierter. Sei

$$(-1)^{\mathcal{N}} \;=\; (-\sigma_3) \otimes \ldots \otimes (-\sigma_3) \;=\; \prod_{i=1}^{N} \mathrm{e}^{i\pi\Sigma^+(i)\Sigma^-(i)}$$

der Operator, welcher in der Konfigurationsraumbasis angibt, ob es eine gerade oder ungerade Anzahl von $+1$ Spins gibt. (Mit der Wahl von σ^+ als „Aufsteigeoperator“ haben wir „Spin-down“ als „Grundzustand“ definiert, d.h. $(-1)^{\mathcal{N}}$ zählt die Anzahl der angeregten Zustände.) Dann gilt

$$-(-1)^{\mathcal{N}}\;(\Psi^+(N) - \Psi(N))\;(\Psi^+(1) + \Psi(1)) \;=\; \Sigma_1(N)\Sigma_1(1) \;.$$

Somit erhalten wir

$$\begin{aligned} V_1 &= \mathrm{e}^{\beta_t^* \sum_n (\Psi^+(n)\Psi(n) - \Psi(n)\Psi^+(n))} \\ V_2 &= \mathrm{e}^{\beta_x \sum_{n=1}^{N-1} [(\Psi^+(n) - \Psi(n))(\Psi^+(n+1) + \Psi(n+1))] \;-\; (-1)^{\mathcal{N}}\;(\Psi^+(N) - \Psi(N))(\Psi^+(1) + \Psi(1))} \;. \end{aligned}$$

Die Sonderstellung des N.ten Terms rührt von den periodischen Randbedingungen her. Offene Randbedingungen bedürften dieser Sonderbehandlung nicht, allerdings verliert man dann die Translationsinvarianz. Das Verfahren zur Lösung des Problems ist in beiden Fällen das Gleiche. Der tiefere Grund für den Faktor $(-1)^{\mathcal{N}}$ wird später (Abschnitt A4.7) einsichtig.

Wir definieren

$$V_2^{\pm} \;=\; \mathrm{e}^{\beta_x \sum_{n=1}^{N} (\Psi^+(n) - \Psi(n))\;(\Psi^+(n+1) + \Psi(n+1))} \;,$$

wobei sich $\pm$ auf die Konvention der Randbedingungen bezieht: Für V^+ sollen $\Psi(n)$ und $\Psi^+(n)$ antiperiodische Randbedingungen haben,

$$\Psi(N+1) \;=\; -\,\Psi(1) \qquad\qquad \Psi^+(N) \;=\; -\,\Psi^+(1) \;,$$

bei V^- sollen die Randbedingungen periodisch sein:

$$\Psi(N+1) \;=\; \Psi(1) \qquad \Psi^+(N) \;=\; \Psi^+(1)\ .$$

Entsprechend erhalten wir auch zwei Transfermatizen $T^\pm$. Da $(-1)^{\mathcal{N}}$ mit allen relevanten Operatoren kommutiert

$$[(-1)^{\mathcal{N}}, V_1] \;=\; [(-1)^{\mathcal{N}}, V_2] \;=\; [(-1)^{\mathcal{N}}, T] \;=\; [(-1)^{\mathcal{N}}, V_2^\pm] \;=\; [(-1)^{\mathcal{N}}, T^\pm] \;=\; 0\ ,$$

lässt sich das Spektrum von $T^\pm$ sowie T jeweils in einen geraden und einen ungeraden Sektor einteilen. Das Spektrum von T ist im geraden Sektor identisch zu dem von T^+ im geraden Sektor und im ungeraden Sektor identisch zu dem von T^- im ungeraden Sektor. Wir diagonalisieren also sowohl T^+ als auch T^-, berücksichtigen aber von den jeweiligen Spektren nur den geraden bzw. ungeraden Sektor.

A4.4 Die Transfermatrix als Funktion der Fourier-Moden von Ψ und Ψ^+

Die Transfermatrix ist nun eine Funktion von quadratischen Ausdrücken von fermionischen Operatoren. Außerdem gilt nach wie vor (unter Berücksichtigung der Periodizität bzw. Antiperiodizität) die Translationsinvarianz der Transfermatrix. Da die fermionischen Kommutatorrelationen unter Fouriertransformation invariant bleiben, und quadratische Ausdrücke von antikommutierenden Operatoren selber kommutieren, können wir nun durch Zerlegung nach Fouriermoden nahezu eine Diagonalisierung erreichen.

Wir definieren:

$$\begin{aligned}
\Psi(n) &= \frac{1}{\sqrt{N}}\, \mathrm{e}^{\frac{-\mathrm{i}\pi}{4}} \sum_k \mathrm{e}^{\mathrm{i}kn}\, b(k) \\
\Psi^+(n) &= \frac{1}{\sqrt{N}}\, \mathrm{e}^{\frac{\mathrm{i}\pi}{4}} \sum_k \mathrm{e}^{-\mathrm{i}kn}\, b^+(k)
\end{aligned}$$

Die möglichen Werte für k hängen von den Randbedingungen ab:

$$\begin{aligned}
k &= 0, \pm\frac{2\pi}{N}, \pm\frac{4\pi}{N}, \ldots, \pm\frac{\pi(N-2)}{N}, \pi \qquad \text{periodische Randbedingungen} \\
k &= \pm\frac{\pi}{N}, \pm\frac{3\pi}{N}, \ldots, \pm\frac{(N-1)\pi}{N} \qquad \text{antiperiodische Randbedingungen}
\end{aligned}$$

(Dabei haben wir der Einfachheit halber N als gerade angenommen.) Für die Operatoren b und b^+

$$\begin{aligned}
b(k) &= \frac{1}{\sqrt{N}}\, \mathrm{e}^{\frac{\mathrm{i}\pi}{4}} \sum_n \mathrm{e}^{-\mathrm{i}kn}\, \Psi(n) \\
b^+(k) &= \frac{1}{\sqrt{N}}\, \mathrm{e}^{\frac{-\mathrm{i}\pi}{4}} \sum_n \mathrm{e}^{\mathrm{i}kn}\, \Psi^+(n)
\end{aligned}$$

folgen die Kommutatorrelationen

$$\{b(k), b(l)\} \;=\; \{b^+(k), b^+(l)\} \;=\; 0 \qquad\qquad \{b^+(k), b(l)\} \;=\; \delta_{kl} \;\;.$$

Wir können nun die Exponenten von $V_1, V_2^{\pm}$ durch die Operatoren b und b^+ ausdrücken:

$$\begin{aligned}
&\sum_n (\Psi^+(n) - \Psi(n))\,(\Psi^+(n+1) + \Psi(n+1)) \;\;=\\
&\sum_k \Big[\mathrm{i}\mathrm{e}^{\mathrm{i}k} b^+(k)b^+(-k) \;+\; \mathrm{e}^{\mathrm{i}k} b^+(k)b(k) \;-\; \mathrm{e}^{-\mathrm{i}k} b(k)b^+(k) \;+\; \mathrm{i}\mathrm{e}^{-\mathrm{i}k} b(k)b(-k) \Big]\\
=\; 2\sum_{k>0} &\big[\sin k\,[b^+(-k)b^+(k) + b(k)b(-k)] \;+\; \cos k\,[b^+(k)b(k) + b^+(-k)b(-k)]\big]\\
&+\; 2b^+(0)b(0) - 1 \;-\; 2b^+(\pi)b(\pi) + 1 \qquad \text{(bei period. Rdb.)}
\end{aligned}$$

(Die Summe $\sum_{k>0}$ erstreckt sich dabei über alle positiven Werte von k ungleich 0 oder π.)

$$\begin{aligned}
\sum_n (\Psi^+(n)\Psi(n) - \Psi(n)\Psi^+(n)) \;&=\; \sum_k (b^+(k)b(k) - b(k)b^+(k))\\
&=\; 2\sum_{k>0} [b^+(k)b(k) + b^+(-k)b(-k) - 1]\\
\text{bei period. Randbed.} \qquad &+\; 2b^+(0)b(0) - 1 \;+\; 2b^+(\pi)b(\pi) - 1
\end{aligned}$$

Da quadratische Ausdrücke in fermionischen Operatoren für verschiedene Werte von k kommutieren, können wir die Matrix T als Produkt von kommutierenden Matrizen T_k schreiben:

$$T^{\pm} \;=\; (2\sinh 2\beta_t)^{N/2} \prod_{0\le k\le\pi} T_k \qquad\qquad T_k = V_k^1 V_k^2 \;\;(k \neq 0, \pi)$$

mit

$$\begin{aligned}
V_k^1 \;&=\; \exp\big(2\beta_t^*[b^+(k)b(k) + b^+(-k)b(-k) - 1]\big)\\
V_k^2 \;&=\; \exp\big(2\beta_x[\cos k\,[b^+(k)b(k) + b^+(-k)b(-k)] \;+\; \sin k\,[b(k)b(-k) + b^+(-k)b^+(k)]]\big)\\
T_0 \;&=\; \exp\big(-2(\beta_t^* - \beta_x)\;(b^+(0)b(0) - \tfrac{1}{2})\big)\\
T_\pi \;&=\; \exp\big(-2(\beta_t^* + \beta_x)\;(b^+(\pi)b(\pi) - \tfrac{1}{2})\big)
\end{aligned}$$

T_0 und T_π sind effektiv 2×2-Matrizen (d.h., sie wirken nicht-trivial nur in einem 2-dimensionalen Unterraum, im Komplement dazu sind sie die Identitätsabbildung), bereits diagonal in der Besetzungszahldarstellung der Operatoren b und b^+ mit Eigenwerten

$$\lambda^{\pm}(0) \;=\; \mathrm{e}^{\pm(\beta_t^* - \beta_x)} \qquad \text{und} \qquad \lambda^{\pm}(\pi) \;=\; \mathrm{e}^{\pm(\beta_t^* + \beta_x)} \;\;. \tag{A4.3}$$

Die Matrizen T_k, die für verschiedene Werte von k kommutieren, sind effektiv 4×4-Matrizen.

A4.5 Diagonalisierung von V_k

Die Diagonalisierung der 4×4-Matrizen ist ein einfaches Problem, welches allerdings von allen Schritten zur Bestimmung des Spektrums der Transfermatrix rechnerisch am aufwendigsten ist. Daher soll nur die Idee sowie das Ergebnis angegeben werden.

Die Matrizen V_k^i sind (für festes k) Funktionen von quadratischen Ausdrücken in fermionischen Operatoren, d.h., sie erhalten die Fermionzahl modulo 2, was das Problem nochmals reduziert. Wir können eine explizite Darstellung (als 4×4-Matrizen) für die Fermionoperatoren angeben:

$$b(k) = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \bigotimes \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad b(-k) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \bigotimes \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$
$$b^+(k) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \bigotimes \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad b^+(-k) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \bigotimes \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

In dieser Darstellung ist

$$b^+(k)b(k) + b^+(-k)b(-k) = \text{diag}\ (2,1,1,0)$$

und

$$\cos k\,[b^+(k)b(k) + b^+(-k)b(-k)] + \sin k\,[b(k)b(-k) + b^+(-k)b^+(k)] = \begin{pmatrix} 2\cos k & 0 & 0 & \sin k \\ 0 & \cos k & 0 & 0 \\ 0 & 0 & \cos k & 0 \\ \sin k & 0 & 0 & 0 \end{pmatrix} .$$

Bezüglich der beiden inneren Einträge sind V_k^1 und V_k^2 also schon diagonal, und die Eigenwerte lassen sich sofort ablesen. Es bleibt also noch eine 2×2-Matrix zu diagonalisieren und zwar

$$\mathrm{e}^{2\beta_t^*\sigma_3}\ \exp\left(2\beta_x \begin{pmatrix} 2\cos k & \sin k \\ \sin k & 0 \end{pmatrix}\right) = \mathrm{e}^{2\beta_x \cos k}\ \mathrm{e}^{2\beta_t^*\sigma_3}\ \mathrm{e}^{2\beta_x(\cos k\ \sigma_3 + \sin k\ \sigma_1)} .$$

Dieses Produkt ist keine symmetrische Matrix, allerdings kann man durch eine Ähnlichkeitstransformation mit $(V_k^1)^{1/2}$ oder aber $(V_k^2)^{1/2}$ eine symmetrische Matrix mit demselben Spektrum erhalten. Die exponentierte Matrix lässt sich nach der bekannten Formel

$$\exp(\alpha\, \boldsymbol{n}\cdot\boldsymbol{\sigma}) = \cosh\alpha + \boldsymbol{n}\cdot\boldsymbol{\sigma}\ \sinh\alpha$$

in eine gewöhnliche Matrix umformen. Die quadratische Eigenwertgleichung führt auf die zwei verbleibenden Eigenwerte. Man erhält (nach einigen (hyper-)trigonometrischen Umformungen) folgendes Ergebnis:

Sei $\epsilon(k)$ die positive Lösung der Gleichung

$$\cosh\epsilon(k) = \cosh 2\beta_x\ \cosh 2\beta_t^* - \sinh 2\beta_x\ \sinh 2\beta_t^*\ \cos k\ , \tag{A4.4}$$

dann sind die Eigenwerte von V_k:

$$\lambda^{\pm\pm}(k) = \mathrm{e}^{2\beta_x\cos k - \epsilon(k)(\pm\frac{1}{2}\pm\frac{1}{2})} .$$

Der rechte Ausdruck enthält vier mögliche Kombinationen von Vorzeichen, welche den vier Eigenwerten entsprechen. Die Eigenwerte der Transfermatrizen $T^{\pm}$ lassen sich also in der Form schreiben:

$$\lambda^{\{\pm\}} = (2\sinh 2\beta_t)^{N/2}\ \exp\frac{1}{2}\sum_k \pm\epsilon(k)\ . \tag{A4.5}$$

(Hier wurde $\sum_k \cos k = 0$ ausgenutzt.) Die Summe erstreckt sich dabei über alle Werte von k (entsprechend den Randbedingungen für T^+ bzw. T^-), und wir haben $\epsilon(-k) = \epsilon(k)$ definiert. Setzen wir in Gl. (A4.4) $k = \pi$, so liefert das Additionstheorem des hyperbolischen Kosinus gerade die richtigen Werte $\epsilon(\pi) = \beta_x + \beta_t^*$. Für $k = 0$ ist $\epsilon(0) = \beta_t^* - \beta_x$. Dies folgt zwar ebenfalls aus dem Additionstheorem des hyperbolischen Kosinus, wenn man $k = 0$ in (A4.4) einsetzt, allerdings kann $\epsilon(0)$ je nach den Werten für β_x und β_t positiv oder negativ sein, wohingegen wir für $k \neq 0$ immer nur positive $\epsilon(k)$ haben. Es gibt insgesamt 2^N mögliche Vorzeichenkombinationen.

A4.6 Das Spektrum der Transfermatrix

In diesem Abschnitt soll das Spektrum der Transfermatrix etwas genauer untersucht werden. In Gl. (A4.4) hatten wir die möglichen Werte für $\epsilon(k)$ als positive Lösungen der Gleichung

$$\cosh \epsilon(k) \;=\; \cosh 2\beta_x \, \cosh 2\beta_t^* \;-\; \sinh 2\beta_x \, \sinh 2\beta_t^* \, \cos k$$

definiert. Diese Gleichung lässt sich nach $\epsilon(k)$ auflösen:

$$\begin{aligned} \epsilon(k) \;&=\; \ln\Big(a(k) + \sqrt{a(k)^2 - 1}\Big) \\ &\qquad \text{mit} \;\; a(k) \;=\; \cosh 2\beta_x \, \cosh 2\beta_t^* \;-\; \sinh 2\beta_x \, \sinh 2\beta_t^* \, \cos k \; . \end{aligned}$$

Für $k = \pi$ ist diese Gleichung ebenfalls korrekt, für $k = 0$ haben wir jedoch nicht die positive Lösung zu betrachten, sondern

$$\epsilon(0) \;=\; \beta_t^* \;-\; \beta_x \; .$$

Der kritische Punkt ist definiert durch $\beta_x = \beta_t^*$. Unterhalb — d.h. für $\beta_t^* < \beta_x$ — ist $\epsilon(0) < 0$. Ansonsten liefert das Additionstheorem der hyperbolischen Funktionen:

$$|\beta_x - \beta_t^*| \;<\; \epsilon(k) \;\leq\; \beta_x + \beta_t^* \qquad \text{für } k \neq 0 \; .$$

Für $\beta_x = \beta_t^*$ verschwindet also der „Massengap“, und die höheren Eigenwerte nähern sich beliebig dem höchsten Eigenwert.

Eine Entwicklung für kleine Werte von k zeigt am kritischen Punkt ($\beta_x = \beta_t^*$) das Verhalten:

$$\epsilon(k) \;\approx\; \text{const.}\, |k| \qquad (\beta_x = \beta_t^*) \; ,$$

also die Dispersionsrelation (Energie-Impuls-Relation) eines masselosen Teilchens. Weg vom kritischen Punkt gilt für kleine Werte von k:

$$\epsilon(k) \;\approx\; |\beta_x - \beta_t^*| \;+\; \text{const.}\, k^2 \qquad (\beta_x \neq \beta_t^*) \; .$$

Wir wollen nun die höchsten Eigenwerte der Transfermatrix untersuchen. Dazu bestimmen wir die jeweils höchsten Eigenwerte von T^+ und T^-. Die nächsthöheren Eigenwerte in den beiden Sektoren unterscheiden sich vom höchsten um mindestens $\epsilon(k)$, d.h. bei $\beta_x \neq \beta_t^*$ um einen nicht-verschwindenden Wert. Für antiperiodische Randbedingungen (T^+) ist der höchste Eigenwert:

$$\lambda_0^+ \;=\; (2 \sinh 2\beta_t) \;\exp \frac{1}{2} \sum_k \epsilon(k) \qquad k = \pm\frac{\pi}{N}, \pm\frac{3\pi}{N}, \ldots, \pm\frac{(N-1)\pi}{N} \; . \tag{A4.6}$$

Dieser Eigenwert gehört auch zum geraden Sektor von T^+, ist also gleichzeitig Eigenwert der Transfermatrix T.

Bei periodischen Randbedingungen (T^-) müssen wir den Fall $\beta_x < \beta_t^*$ und $\beta_x > \beta_t^*$ unterscheiden. In beiden Fällen ist der höchste Eigenwert zunächst:

$$\lambda_0^- \;=\; (2\sinh 2\beta_t)\;\exp\frac{1}{2}\sum_k |\epsilon(k)| \qquad k = 0, \pm\frac{2\pi}{N}, \ldots, \pm\frac{(N-2)\pi}{N}, \pi\;. \tag{A4.7}$$

Für $\beta_t^* > \beta_x$, d.h. bei hohen Temperaturen in der ungebrochenen Phase des Modells, ist $\epsilon(0) = \beta_t^* - \beta_x$ positiv. In diesem Fall können wir also die Absolutbetragsstriche in (A4.7) weglassen. Allerdings haben wir dann einen Eigenwert zu einer geraden Zahl von Anregungen (da sämtliche Transformationen von $\Sigma(n)$ zu $\Psi(n)$ zu $b(k)$ die Relation „gerade–ungerade“ erhalten, haben wir für eine gerade Anzahl von „$-$“-Zeichen in (A4.5) auch einen Zustand im geraden Sektor), dieser Eigenwert von T^- ist also kein Eigenwert der Transfermatrix.

Für $\beta_t^* < \beta_x$ (in der gebrochenen Tieftemperaturphase) ist $|\epsilon(0)| = -\epsilon(0)$. λ_0^- ist also ein Eigenwert von T^- zu einem ungeraden Zustand und daher ein Eigenwert der Transfermatrix. Im thermodynamischen Grenzfall wird $\lambda_0^+ = \lambda_0^-$, der höchste Eigenwert der Transfermatrix ist also bei tiefen Temperaturen zweifach entartet. Dies entspricht den beiden nicht-äquivalenten Gleichgewichtszuständen in der gebrochenen Phase des Modells.

Auf die Freie Energie pro Spin hat diese zweifache Entartung keinen wesentlichen Einfluss. Im thermodynamischen Grenzfall können wir die Summe über k-Werte durch eine Integration ersetzen (mit $\Delta k = 2\pi/N$) und erhalten:

$$\begin{aligned} \frac{1}{N}\ln Z \quad &= \quad \frac{1}{2}\ln(2\sinh 2\beta_t) \;+\; \frac{1}{2}\sum_k \epsilon(k) \\ (\text{für } N\to\infty) \qquad &\to \quad \frac{1}{2}\ln(2\sinh 2\beta_t) \;+\; \frac{1}{4\pi}\int_{-\pi}^{\pi}\epsilon(k)\,\mathrm{d}k\;. \end{aligned}$$

A4.7 Das Ising-Modell als Theorie von Fermionen

Das Ising-Modell lässt sich auch als eine Theorie von Fermionen interpretieren, wobei jedes Fermion in einem von N Zuständen (indiziert durch den „Impuls“ k) sein kann. Zur genaueren Diskussion dieser Interpretation unterscheiden wir wieder zwischen der Hoch- und Tieftemperaturphase.

Bei hohen Temperaturen ($\beta_t^* > \beta_x$) sind alle $\epsilon(k)$ positiv, der Grundzustand ist eindeutig und gehört zum höchsten Eigenwert von T^+. Alle Impulszustände sind unbesetzt. Der 1-Teilchen-Sektor wird von den Eigenwerten von T^- gebildet, bei denen ein Vorzeichen negativ ist, d.h. ein Zustand zu einem Impulswert mit $k = 0, \pm 2\pi/N, \ldots$ ist besetzt. Diese Impulse entsprechen periodischen Randbedingungen — die Wellenfunktion ist eindeutig. Die Eigenzustände im 2-Teilchen-Sektor stammen zu T^+. Zwei Impulszustände zu verschiedenen $k = \pm\pi/N, \ldots$ sind mit jeweils einem Fermion besetzt. Diese Impulswerte entsprechen antiperiodischen Randbedingungen und tragen so der Tatsache Rechnung, dass bei Vertauschung zweier Fermionen (was sich auf einem kreisförmigen Gebiet auch durch einen Umlauf um das Gitter erreichen lässt) ein relatives Minus-Zeichen auftritt. Auch in diesem Sinne ist die (antisymmetrische) 2-Teilchen-Wellenfunktion eindeutig.

Allgemein sind die n-Teilchenzustände zu geradem n auch Eigenzustände von T^+, die möglichen Impulse gehören zu antiperiodischen Randbedingungen da ein Umlauf durch das Gitter einer Vertauschung mit einer ungeraden Anzahl von Fermionen entspricht. Entsprechend sind die n-Teilchen-Zustände zu ungeradem n Eigenzustände von T^- mit Impulsen zu periodischen Randbedingungen. Schließlich ist im N-Teilchen-Sektor jeder Impulswert einmal besetzt. Da wir N als gerade angenommen haben, handelt es sich um einen Eigenzustand von T^+. Insgesamt erhält man 2^N Zustände.

Bei tiefen Temperaturen ist die Situation etwas komplizierter. Da nun $\epsilon(0)$ negativ ist, ist es energetisch günstiger, wenn der Zustand zu $k = 0$ besetzt ist. Der Grundzustand ist im Allgemeinen eine Linearkombination zwischen dem entsprechenden Eigenzustand aus T^+, bei dem alle Impulswerte unbesetzt sind, und dem Eigenzustand zu T^-, bei dem der $k = 0$ Zustand besetzt ist. Beide Zustände haben (im thermodynamischen Grenzfall) dieselbe Energie. Die 1-Teilchen-Zustände zu $k \neq 0$ haben eine höhere Energie als in der Hochtemperaturphase: $E(k) = -\epsilon(0) + \epsilon(k)$, da nun sowohl der Zustand zu $k = 0$ nicht besetzt ist (was wegen $\epsilon(0) < 0$ Energie kostet) als auch der Zustand zu k besetzt ist. Ansonsten ist die Diskussion ähnlich wie in der Hochtemperaturphase.

Wir können nun auch die Bedeutung des Vorzeichens $-(-1)^{\mathcal{N}}$ besser verstehen: Bei einer Translation eines einzelnen Fermions um N Gitterpunkte muss dieses Fermion an $\mathcal{N}-1$ anderen Fermionen vorbeigetauscht werden. Damit die Wellenfunktion trotzdem eindeutig ist, muss man für gerade $\mathcal{N}$ antiperiodische, für ungerade $\mathcal{N}$ periodische Randbedingungen fordern.

Literaturverzeichnis

[1] E. S. Abers & B. W. Lee; *Gauge Theories*; Phys. Rep. 9C (1973) 1.

[2] M. Aizenman; *Proof of the Triviality of φ_d^4 Field Theory and Some Mean-Field Features of Ising Models for $d > 4$*; Phys. Rev. Lett. 47 (1981) 1-4.

[3] J. Ambjørn, B. Durhuus, J. Fröhlich; Nucl. Phys. B 257 (1985) 433.

[4] M. Bander, C. Itzykson; Phys. Rev. D 15 (1977) 463.

[5] R.J. Baxter; *Exactly Solved Models in Statistical Mechanics*; Academic Press, 1982.

[6] P. Becher, M. Böhm & H. Joos; *Eichtheorien*; Teubner Studienbücher, Stuttgart (1981)

[7] F.A. Berezin; *The Method of Second Quantization*; Academic Press, New York and London, 1966.

[8] B. Berg, A. Billoire; Phys. Lett. 113 B (1982) 65.

[9] C. Bernard, T. Draper, G. Hockney, A. M. Rushton; Phys. Rev. Lett. 55 (1985) 2770.

[10] G. Bhanot, C. Rebbi; Nucl. Phys. B 180 (1981) 469.

[11] G. Bhanot; *The Metropolis Algorithm*; Rep. Prog. Phys. 51 (1988) 429.

[12] W. Bietenholz, S. Capitani, T. Chiarappa, N. Christian, M. Hasenbusch, K. Jansen, K.-I. Nagai, M. Papinutto, L. Scorzato, S. Shcheredin, A. Shindler, C. Urbach, U. Wenger, I. Wetzorke; *Going chiral: overlap versus twisted mass fermions*; arXiv:hep-lat/0411001v2 (2004).

[13] K. Binder; *Theory and Technical Aspects of Monte Carlo Simulation*; in *Monte Carlo Methods in Statistical Physics*; (Ed. K. Binder), Springer, Berlin, Heidelber (1987).
K. Binder; *Applications of the Monte-Carlo Method in Statistical Physics*; Springer, Berlin, Heidelberg (1995).

[14] J. D. Bjorken, S. D. Drell; *Relativistische Quantenfeldtheorie*; BI-Taschenbücher.

[15] Brezin, E., Gross, D.J.; Phys. Lett. B 97 (1980) 120.

[16] S. Chandrasekharan and U.-J. Wiese, *An Introduction to Chiral Symmetry on the Lattice*; arXiv:hep-lat/0405024v1 (2004); Prog. Part. Nucl. Phys., Vol. 53 (2004) 373-418.

[17] T. Chiarappa, K. Jansen, K.I. Nagai, M. Papinutto, L. Scorzato, A. Shindler, C. Urbach, U. Wenger and I. Wetzorke; *Iterative methods for overlap and twisted mass fermions*; Computational Science & Discovery 1 (2008) 015001 (29pp).

[18] M. Creutz, L. Jacobs, C. Rebbi; Phys. Rev. Lett. 42 (1979) 1390.
M. Creutz; Phys. Rev. Lett. 43 (1979) 553

[19] M. Creutz, L. Jacobs, C. Rebbi; *Monte Carlo Simulations in Lattice Gauge Theories*; Phys. Rep. 95 (1983) 201

[20] M. Creutz; *Chiral Symmetry on the Lattice*; Nucl. Phys. B (Proc. Suppl.) 42 (1995) 56.

[21] F. David; Nucl. Phys. B 257 (1985) 45.

[22] T. DeGrant and C. DeTar; *Lattice Methods For Quantum Chromodynamics*; World Scientific 2006.

[23] E.H. de Groot; *Fermions on a Lattice*; Acta Physica Polonica, Vol. 16 (1985) 627-633.

[24] F.J. Dyson, Phys. Rev. 85 (1952) 631.

[25] J. Engels, F. Karsch, H. Satz, I. Montvay; Phys. Lett. 101 B (1981) 89.

[26] E. Fradkin, S. Shenker; Phys. Rev. D 19 (1979) 3682.

[27] Paul H. Frampton; *Gauge Field Theories*; John Wiley & Sons, 2nd ed., Aug. 2000.

[28] K. Fredenhagen, M. Marcu; Commun. Math. Phys. 92 (1983) 81, und Phys. Rev. Lett. 56 (1986) 223.

[29] R. Frezzotti, P.A. Grassi, S. Sint, P. Weisz; *A local formulation of lattice QCD without unphysical fermion zero modes*; Nucl. Phys. Proc. Suppl. 83 (2000) 941-946.

[30] R. Frezzotti; em Twisted mass Lattice QCD; arXiv:hep-lat/0409138v2 (2004).

[31] D. Brydges, J. Fröhlich, T. Spencer; *The Random Walk Representation of Classical Spin Systems and Correlation Inequalities*; Commun. Math. Phys. 83 (1982) 123-150.

[32] F. Fucito, G. Martinelli, C. Omero, G. Parisi, R. Petronzio, R. Rapuano; Nucl. Phys. B 210 (1982) 407.

[33] V. Furman, Y. Shamir; Nucl. Phys. B 439 (1995) 54.

[34] K. Fujikawa; *Lattice Chiral Symmetry, CP-Violation and Majorana Fermions*; Ann. Henri Poincaré 4, Suppl. 2 (2003) S905-S912.

[35] P.H. Ginsparg and K.G. Wilson; Phys. Rev. D 25 (1982) 2649.

[36] Glimm, J., Jaffe A.; *Quantum Physics - A Functional Integral Point of View*, Springer-Verlag, New-York, 1981.

[37] O.W. Greenberg; *Haag's Theorem and Clothed Operators*; Phys. Rev. 115, No 3 (1959) 706-710.

[38] Gross, D.J., Witten, E.; Phys. Rev. D 21 (1980) 446.

[39] R. Haag, Kgl. Danske Videnskab. Selskab, Mat.-fys. Medd. **29**, No 12 (1955).

[40] I. G. Halliday; *Lattice Field Theories*; Rep. Prog. Phys. 47 (1984) 987

[41] H. Hamber, G. Parisi; Phys. Rev. Lett. 47 (1981) 1792.

[42] A. Hasenfratz, P. Hasenfratz; Phys. Lett. 93 B (1980) 165.

[43] P. Hasenfratz and F. Niedermayer; Nucl. Phys. B414 (1994) 785.

[44] P. Hasenfratz Nucl. Phys. B (Proc. Suppl.) 63 (1998) 2649.

[45] K. Ishikawa, M. Teper, G. Schierholz; Phys. Lett. 116 B (1982) 429.

[46] E. Ising; Z. Physik 31 (1925) 253.

[47] C. Itzykson & J. B. Zuber; *Quantum Field Theory*; McGraw-Hill Inc. (1980).

[48] J. Glimm and A. Jaffe; *Quantum Physics - A Functional Integral Point of View*; Springer-Verlag New York, Heidelberg, Berlin (1981).

[49] F. James; *Monte Carlo Theory and Practice*; Rep. Prog. Phys. 43 (1980) 1145.

[50] K. Kanaya; *An Introduction to Finite Temperature Quantum Chromodynamics on the Lattice*; Prog. Theor. Phys. Suppl., No. 131 (1998) 73.

[51] D.B. Kaplan; Phys. Lett. B 288 (1992) 342.

[52] D.B. Kaplan; Nucl. Phys. B 30 (Proc. Suppl.) (1993) 597.

[53] D.B. Kaplan and M. Schmaltz; *Domain wall fermions and the eta invariant*; Phys. Lett. B 368 (1996) 44-52.

[54] F. Karsch, E. Laermann; *Numerical Simulations in Particle Physics*; Rep. Prog. Phys. 56 (1993) 1347.

[55] F. Karsch; *Lattice QCD at High Temperature and Density*; arXiv hep-lat/0106019v2 (2001).

[56] L. H. Karsten, J. Smit; Nucl. Phys. B 185 (1981) 20.

[57] V. A. Kazakov, I. K. Kostov, A. A. Migdal; Phys. Lett. **B 157**, 295 (1985).

[58] J. Kogut & L. Susskind; Phys. Rev. D11 (1975) 395.

[59] J. B. Kogut; *An introduction to lattice gauge theory and spin systems*; Rev. Mod. Phys. 51 (1979) 659.

[60] J. B. Kogut; *The lattice gauge theory approach to quantum chomodynamics*; Rev. Mod. Phys. 55 (1983) 775.

[61] H.A. Kramers and G.H. Wannier; *Statistics of the two-dimensional ferromagnet*; Phys. Rev. **60** (1941) 252-262.

[62] J. Kuti, J. Polonyi, K. Szlachanyi; Phys. Lett. 98 B (1981) 199.

[63] P. Lancaster; *Theory of Matrices*; Academic Press, New York, London (1969).

[64] Melvin Lax; *Symmetry Principles in Solid State and Molecular Physics*; John Wiley & Sons, 1974 Inc.; S. 438.

[65] Lieb, Mattis und Schultz, *Rev. Mod. Phys.* (1964) 856.

[66] J. S. Lomont; *Applications of Finite Groups*; Academic Press, 1959; S. 146

[67] H. Lehmann, K. Symanzik, and W. Zimmerman, *On the formulation of quantized field theories*, Nuovo Cimento 1 (1955) 205.

[68] M. Lüscher; *Weyl fermions on the lattice and the non-abelian gauge anomaly*; Nuclear Physics B568 (2000) 162-179.

[69] M. Lüscher; *Chiral gauge theories revisited*; CERN-TH/2001-031; arXiv:hep-th/0102028v2.

[70] L. Maiani, G. Martinelli; Phys. Lett. 181 B (1986) 344.

[71] E. Marinari, G. Parisi, C. Rebbi; Phys. Rev. Lett. 47 (1981) 1795.

[72] L. D. McLerran, B. Svetitsky; Phys. Lett. 98 B (1981) 195.

[73] C. Michael; *Hadronic Physics from the Lattice*; in *Confinement, Duality, and Nonperturbative Aspects of QCD*; P. van Baal (Ed.); NATO ASI, Series B, Vol. 368, 1997.

[74] I. Montvay & G. Münster; *Quantum Fields on a lattice*; Cambridge University Press, 1994.

[75] I. Montvay; *Majorana fermions on the lattice*; arXiv:hep-lat/0108011v1 (2001).

[76] O. Nachtmann; *Elementarteilchenphysik; Phänomene und Konzepte*; Vieweg, Braunschweig (1986)

[77] R. Narayanan, H. Neuberger; Phys. Lett. B 302 (1993) 62.
Phys. Rev. Lett. 71 (1993) 3251.
Nucl. Phys. B 412 (1994) 574.
Nucl. Phys. B 443 (1995) 305.

[78] H. Heuberger; *Exactly massless quarks on the lattice*; Phys. Lett. B 417 (1998) 141-144.

[79] H. Neuberger; *A practical implementation of the overlap-Dirac operator*; Phys. Rev. Lett. 81 (1998) 4060-4062.

[80] Gordon, F. Newell und Elliott W. Montroll, Rev. Mod. Phys. 25 (1953) 353.

[81] Ferenc Niedermayer; *Exact chiral symmetry, topological charge and related topics*, Nucl. Phys. B (Proc. Suppl.) 73 (1999) 105.

[82] H. B. Nielsen, M. Ninomiya; Nucl. Phys. B 185 (1981) 20.

[83] H. B. Nielsen, M. Ninomiya; Nucl. Phys. B193 (1981) 173.

[84] L. Onsager; Phys. Rev. 65 (1944) 117.

[85] K. Osterwalder, R. Schrader; Comm. Math. Phys. 31 (1973) 83; und Comm. Math. Phys. 42 (1975) 281.

[86] K. Osterwalder, E. Seiler; Ann. of Phys. 110 (1978) 440.

[87] J.M. Pawlowski and Y. Igarashi; *On Majorana fermions on the lattice*; arXiv:hep-lat/0710.4215v1 (2007).

[88] Michael E. Peskin, Daniel V. Schroeder; *An Introduction to Quantum Field Theory*; Perseus Press, Sept. 1995.

[89] S. Pokorski; *Gauge Field Theories*; Cambridge University Press, 1999.
S. Pokorski; *Gauge Field Theories*; Cambridge Monographs on Mathematical Physics; Cambridge University Press (1987)

[90] C. Rebbi; *Lattice Gauge Theories and Monte Carlo Simulations*; World Scientific, Singapore (1983)

[91] H. J. Rothe; *Lattice Gauge Theories – An Introduction*; World Scientific Lecture Notes in Physics – Vol. 74; World Scientific, Singapur, Third Edition (2005).

[92] Lewis H. Ryder; *Quantum Field Theory*; Cambridge University Press; Mai 1996.

[93] M. Sato, T. Miwa, M. Jimbo; *Field Theory of the 2-dimensional Ising model in the scaling limit*; RIMS-preprint.

[94] S.S. Schweber; *An Introduction to Relativistic Quantum Field Theory*; Harper & Row, New York, Evanston, London, 1964.

[95] Y. Shamir; Nucl. Phys. B 406 (1993) 90.

[96] R.F. Streater and A.S. WIghtman, *PCT, Spin and Statistics and All That*, Princeton University Press, Landmarks in Mathematics and Physics, 2000.

[97] J. Smit; Acta Physica Polonica Vol. B 17 (1986), 531-551.

[98] K. Symanzik; *Euclidean quantum field theory*; in: *Local quantum theory.* R. Jost (ed.), Academic Press, New York, London (1969).

[99] F. J. Wegner; J. Math. Phys. 10 (1971) 2259.

[100] S. Weinberg; *The Quantum Theory of Fields: Foundations* (Juni 1995); *The Quantum Theory of Fields: Modern Applications* (August 1996); *The Quantum Theory of Fields: Supersymmetry* (März 2000); Cambridge University Press.

[101] K. G. Wilson; Phys. Rev. D10 (1974) 2445.

[102] J.B. Zuber, C. Itzykson; Phys. Rev. D 15 (1977) 2875.

Index

Abelsche Eichfelder, 15
Adjazenzmatrix, 115
Adler-Bell-Jackiw-Anomalie, 104
Akzeptanzrate, 80
Anti-Ferromagnet, 54
Autokorrelationsfunktion, 79
Autokorrelationslänge, 79
Autokorrelationszeit, 79

Baker-Hausdorff-Formel, 43
β-Funktion, 95
Boltzmann-Faktor, 29

Charaktere der oktaedrischen Gruppe, 87
Chirale Fermionen, 102, 103
Chirale Transformation, 102
Chiralität, 102
Cluster-Algorithmen, 83
Colour-Freiheitsgrade, 16
Compton-Länge, 38
Confinement, 46
Confinement-Phase, 71

Detailed balance, 77
Dirac-Gleichung, 13
Duale Temperatur, 53, 121

Eichfixierung, 59
Eichinvariante Funktionen, 45
Eichpotential, 15
Eichtransformation, 16, 42, 44, 65
    nicht-abelsch, 17, 108
Elementarereignis, 30, 74
Ergodizität, 76
Erwartungswerte, im Transfermatrixformalismus, 57
Erwartungswertfunktional, 30, 74
Erzeugende Funktion, 112, 113
Euklidische Formulierung, 25
Euklidische Lagrange-Dichte
    Dirac-Theorie, 26
    Klein-Gordon, 26
Euklidische Wirkung, 25
    Dirac-Feld, 100
    Klein-Gordon-Feld, 31
Euklidisches Funktionalintegral, 26
Explizite Symmetriebrechung, 69

Feldgleichungen, 14
Feldrenormierung, 37
Feldstärke
    elektrisch und magnetisch, 15
Feldstärketensor, 15
    nicht-abelsch, 18
Fermionische Operatoren, Gitterdarstellung, 123
Fermionverdopplung, 101
Flächengesetz, 47
Fließgleichgewicht, 77
Fluxtube, 70
Freie Energie, 53
Freie Ladungsphase, 71
Frobenius-Perron-Theorem, 56
Funktionalintegral, 24
    für Fermionen, 98
Funktionalintegraldarstellung, 22
    Phasenraumdarstellung, 22

Gaußsches Integral
    für Grassmann-Variable, 99
    komplex, 113
Gitter, 29
Gitterabstand, 35
Gittereichtheorie, 41
    2-dimensionale, reine, 61
Glueballs, 85

Goldstone-Bosonen, 69
Graph, 115
Grassmann-Algebra, 97
Grassmann-Variable, 97
Greensche Funktion, 113

Haar-Maß, 44, 48
Hamilton-Operator, 58
    einer Gittertheorie, 56
    zur Klein-Gordon-Theorie, 59
Heat bath-Verfahren, 81
Hilbert-Raum, 55
Hopping-Parameter, 65
Hybrid Monte-Carlo, 83

Invariantes Linienelement, 45
Invarianz, 12
Irreduzibilität, einer Markov-Matrix, 76
Irreduzible Darstellungen
    der oktaedrischen Gruppe, 87
Ising-Modell, 29
    1-dimensional, 52
    2-dimensionales, 119

Jordan-Wigner-Transformation, 122

Klein-Gordon-Gleichung, 12
Konfiguration, 29
    in reiner Gittereichtheorie, 42
Konfigurationsraumbasis, 55, 120
Kontinuumslimes, 91
    naiver, 43
Kontinuumslimes, des Klein-Gordon-Feldes, 35
Korrelationslänge, 38, 54
Kovariante Ableitung
    nicht-abelsch, 17
Kovariante Konstanz, 16
    nicht-abelsch, 18
Krümmung, 42
Krümmungstensor, 18, 108
    Riemannscher, 109

Ladungsdichte, 15
Ladungserhaltung, 15
Ladungskonjugation, 85
Lagrange-Dichte, 14, 17, 18
    Maxwell-Gleichungen, 15
Langevin-Algorithmus, 82
Large Coupling-Expansion, 48
Levi-Civita-Zusammenhang, 109
Lines of constant physics, 93
Linienbündel, komplexes, 109
Little Group, 68
Lokale Zustände, 29
Lorentz-Eichung, 15
Lorentz-Gruppe
    Irreduzible Darstellung, 12
Lorentz-Transformation, 12

Markov-Matrix, 76, 117
Markov-Prozess, 76
Markov-Zeit, 76
Maß, invariantes, 44
Massenschale, 12
Maxwell-Gleichungen, 15
Messung, in Monte-Carlo-Simulation, 78
Metropolis-Verfahren, 80
Minimale Kopplung, 15
MIT-Fermionen, 103
Molekulardynamik, 82
Monte-Carlo-Simulation, 78
Monte-Carlo-Verfahren, 73
Multi-Grid-Verfahren, 83

Nicht-abelsche Eichtheorie, 16
Nielsen-Ninomiya-Theorem, 103

Observable, 30, 74
    in reiner Gittereichtheorie, 45
Oktaedrische Gruppe, 86
Onsager, 119
Ordnungsparameter, 68
    in reiner Eichtheorie, 46
Ortsraumbasis, 55

Paralleltransporter, 16, 18, 41, 107
Parameter, nackte (bare), 37
Paritätstransformation, 85
Pauli-Matrizen, 53, 120
Phasenübergang, 62
Phasendiagramm, 66
Photon, 16
Physikalisches Fenster, 93

Plakettvariable, 42
Poincaré-Transformation, 11
Poissonsche Summenformel, 35
Polyakov-Loop, 46
Potts-Modell, 29
Primitivität, einer Markov-Matrix, 76
Prinzipalbündel, 109
Propagator, 21
    für Fermionen, 101
J/Ψ-Teilchen, 50
Pure Gauge-Konfiguration, 67

$q\bar{q}$-Zustände, 50
Quantenchromodynamik, 16
Quantenfeldtheorie bei endlicher Temperatur, 46
Quantenspin-Operator, 54
Quenched Fermions, 106

Randbedingungen, 30
    Antiperiodische, 31
    Feste, 30
    Helikale, 31
    Offene, 30
    Periodische, 30
Raum-Richtung, im Transfermatrixformalismus, 55
Regularisierung, 26
Reine Gittereichtheorie, 42
Renormierung, 27
Renormierungsgruppentransformation, 93
Renormierungspunkt, 93
Riemannsche Geometrie, 109
Rotationsinvarianz, Restauration, 37

Sample, 73
Schrödinger-Gleichung, 12, 21
Screening-Phase, 71
σ-Algebra, 74
Simulation von Fermionen, 105
Skalares Feld, 13
Spin 1/2 Feld, 13
Spontane Symmetriebrechung, 69
Störungstheoretische Renormierbarkeit, 26
Staggered Fermions, 104
Standardmodell, 13, 19, 102
Startkonfiguration, 78
Stationärer Zustand, 76
Statisches $q\bar{q}$-Potential, 46, 50
Stochastische Matrix, 76
Stochastischer Prozess, 75
Störungstheorie, 26
Streumatrix, 23
Stringtension, 46, 50
Stromdichte, 15
Strukturgruppe, 109
Summation über Wege, Darstellung der Zwei-Punkt-Funktion, 33

Temporale Eichung, 59
Thermalisierung, 78
Thermodynamischer Limes, 33
Torsionsfreiheit, 109
Transfermatrix, 52, 56, 60
    2-dimensionales Ising-Modell, 120
Transfermatrixformalismus, 51, 55
    für Gittereichtheorien, 59
    zur Klein-Gordon-Theorie, 58
    zur Quantenmechanik, 57
Transformation
    aktive, 12
Translationsoperator, 32
Trivialitätsschranke, 94

Überlagerungsgruppe, 12
Umfangsgesetz, 47
Unimodular, 44
Universelle Koeffizienten der β-Funktion, 96

Vakuumerwartungswerte, zeitgeordnet, 24
    in der Feldtheorie, 25
Vakuumzustand, 24
Valenzmatrix, 115
Vektorbündel, 107
Verschmierte Zustände, 88
Verteilungsfunktion, 30
Vier-Punkt-Funktion, 112
Vierer-Strom, 15

Wahrscheinlichkeitsmaß, 74
Wahrscheinlichkeitsraum, 74
Wahrscheinlichkeitsverteilung, 29
Weggeordnetes Produkt, 18
Wellenfunktion, links- bzw. rechtshändige, 102

Wellenfunktionsrenormierung, 24
Wick-Rotation, 25
Wicksches Theorem, 112, 114
Wilson-Fermionen, 104
Wilson-Kriterium, 47
Wilson-Loop, 46
Wirkung, 14, 29
    für Fermionen auf dem Gitter, 101
    Gittereichtheorie, 43
    Gittereichtheorie in temporaler Eichung, 60
Wirkungsfunktional, 14

Υ-Teilchen, 50
Yukawa-Kopplung, 19, 102

$\mathbf{Z}_2$-Eichtheorie mit $\mathbf{Z}_2$-Feldern, 67
Zeit-Richtung, im Transfermatrixformalismus, 55
Zeitordnung, 24
Zerlegung nach irreduziblen Darstellungen, 87
Zufallsvariable, 30, 74
Zusammenhang, 107
Zustand, 30, 74
Zustandssumme, 29
Zwei-Punkt-Funktion, 112
    Klein-Gordon-Feld, 33