# Nonstandard Cutoff Effects in $O(N)$ Nonlinear Sigma Models

D I P L O M A R B E I T

zur Erlangung des akademischen Grades
Diplom Physiker
(Dipl.-Phys.)

eingereicht an der
Mathematisch-Naturwissenschaftlichen Fakultät I
Humboldt-Universität zu Berlin


von
Björn Leder
geboren am 22.10.1978 in Altenburg

Präsident der Humboldt-Universität zu Berlin:
Prof. Dr. Jürgen Mlynek

Dekan der Mathematisch-Naturwissenschaftlichen Fakultät I:
Prof. Dr. Michael Linscheid

Gutachter:

1. Prof. Dr. Ulrich Wolff
2. Prof. Dr. Michael Müller-Preussker

| | |
|---|---|
| eingereicht am: | 10. Oktober 2003 |
| Tag der mündlichen Prüfung: | 30. Oktober 2001 |

**Zusammenfassung**

Die Regularisierung mit Hilfe eines Raum-Zeit-Gitters ist eine mathematisch wohl definierte, nicht störungstheoretische Formulierung einer Quantenfeldtheorie. Theoretische Physiker auf dem Gebiet der Gitter-QCD widmen große Mengen Rechenleistung der Frage, ob die QCD die Physik der leichten Hadronen bei kleinen Energien beschreibt. Da die Diskretisierung der Raum-Zeit systematische Fehler mit sich bringt, muß ein Kontinuumsgrenzwert bestimmt werden.

Aufgrund seiner Ähnlichkeit zu den physikalisch relevanteren vierdimensionalen Eichtheorien wird das nichtlineare O(N)-Sigma-Modell benutzt, um störungstheoretische Vorhersagen zu testen. Im nichtlinearen O(3)-Sigma-Modell wurden nicht der Erwartung entsprechende Diskretisierungseffekte gefunden.

Das Verhalten der Diskretisierungseffekte wird für kleine und mittlere N bis hin zu N gegen unendlich untersucht. Für N gegen unendlich ist das nichtlineare Sigma-Modell exakt lösbar. Das Verhalten der Diskretisierungseffekte wird im O(4)- und O(8)-Modell mit Hilfe von Monte-Carlo-Methoden bestimmt. Die Gitter-Artefakte werden mit verschiedenen Ansätzen verglichen. Neuen theoretische Vorhersagen für den Kontinuumswert der Step-Scaling-Funktion werden MC-Daten gegenübergestellt.

Für die Simulationen wurde ein effizienter Cluster-Algorithmus und eine varianzreduzierende Schätzfunktion implementiert.

Auch im nichtlinearen O(4)- und O(8)-Sigma-Modell werden nicht der Erwartung entsprechende Diskretisierungseffekte beobachtet. Aber die Gitter-Artefakte sind kleiner und die Abweichung ist nicht so deutlich wie im nichtlinearen O(3)-Sigma-Modell.

**Schlagwörter:**
Gitter-QFT, nichtlineares Sigma-Modell, Monte-Carlo-Simulation, Gitter-Artefakte

**Abstract**

Lattice regularization is a mathematically well defined, nonperturbative approach to quantum field theory. The lattice QCD community dedicates a huge amount of computing power to verify that the QCD Lagrangian describes physics of light hadrons at low energy. But the discretization of space-time involves systematical errors. Thus a continuum limit should be taken.

Because of its similarity to the physically more relevant four dimensional gauge theories, the two dimensional nonlinear O(N) sigma model is used as a testing ground for perturbation theory predictions. In the nonlinear O(3) sigma model nonstandard cutoff effects were found.

The behavior of the cutoff effects is analyzed as N changes from small over intermediate values towards the large N limit, where the model is exactly solvable. The cutoff dependence in the O(4) and O(8) model is determined using Monte Carlo methods. The lattice artifacts are fitted to several forms. Recently presented theoretical predictions for the continuum value of the step scaling function are confronted with the MC data.

For the simulations an efficient cluster algorithm and an improved estimator are implemented.

Nonstandard cutoff effects are observed in the nonlinear O(4) and O(8) sigma model too, but the lattice artifacts are smaller and the discrepancy is not as distinct as in the O(3) sigma model.

**Keywords:**
lattice QFT, nonlinear sigma model, Monte Carlo simulation, lattice artifacts

# Danksagung

Ich danke allen Mitgliedern der Arbeitsgruppe Computational Physics für die Unterstützung, die hilfreichen Diskussionen und die hohe Integration in die Gruppe. Ulli Wolff und Francesco Knechtli gilt besonderer Dank für die intensive Zusammenarbeit.

Ich danke Hendrik Hache, Jörg Reichardt, Tomasz Korcez, Sylvia Schikora und Ralf Tönjes für das Korrigieren des Manuskripts und die vielen Verbesserungsvorschläge.

Ich danke meinen Eltern, die mich immer unterstützen und mir alles gegeben haben.

# Contents


**1 Introduction** **1**

**2 Lattice Quantum Field Theory** **5**
2.1 Functional Integral Quantization . . . . . 6
2.1.1 Path Integral in Quantum Mechanics . . . . . 6
2.1.2 Quantization of Field Theories . . . . . 7
2.1.3 Lattice Regularization . . . . . 8
2.1.4 The Nonlinear Sigma Model . . . . . 8
2.2 The Transfer Matrix Formalism . . . . . 9
2.3 Field Theory and Statistical Mechanics . . . . . 11

**3 Nonstandard Cutoff Effects** **13**
3.1 Cutoff Effects in Lattice Simulations . . . . . 14
3.1.1 Lattice Regularization and Continuum Limit . . . . . 14
3.1.2 Perturbative Understanding of Lattice Artifacts . . . . 14
3.2 Cutoff Effects in the Nonlinear Sigma Models . . . . . 16
3.2.1 XY Model in 2d - A Pathological Case . . . . . 16
3.2.2 Numerical Results for $N = 3$ . . . . . 17
3.2.3 Large $N$ Predictions . . . . . 17
3.2.4 Summary of the Situation for $N \geq 3$ . . . . . 18
3.3 Testing Predictions . . . . . 18
3.3.1 Field Theory in a Finite Volume . . . . . 18
3.3.2 Step Scaling Function . . . . . 19
3.3.3 Lattice Effects . . . . . 21
3.3.4 Strategy . . . . . 21

**4 Numerical Methods** **23**
4.1 Computation of the Step Scaling Function . . . . . 24
4.1.1 Time Slice Correlation Function . . . . . 25
4.1.2 Finite Volume Mass Gap . . . . . 28
4.1.3 Step Scaling Function and Continuum Limit . . . . . 30
4.2 Improved Estimator . . . . . 31
4.2.1 General Remarks . . . . . 31

4.2.2 $O(N)$ Vector Models . . . 32
4.2.3 Time Slice Correlation Function . . . 34
4.2.4 Monte Carlo Algorithm . . . 35
4.2.5 Testing the Estimator . . . 41

**5 Large $N$ Expansion** **45**
5.1 $1/N$ Expansion of the $O(N)$ Model . . . 46
5.1.1 $1/N$ Expansion in the Continuum . . . 46
5.1.2 $1/N$ Expansion on the Lattice . . . 47
5.2 Step Scaling Function in the Large $N$ Limit . . . 51
5.2.1 The Small $z$ Expansion . . . 52
5.2.2 Step Scaling Function . . . 56
5.2.3 Checking the Expansion . . . 59

**6 Results** **63**
6.1 Comparison of the Cutoff Effects: $N = 3, 4, 8, \infty$ . . . 64
6.2 Fit of the Lattice Artifacts . . . 68
6.2.1 Simple Linear or Quadratic . . . 68
6.2.2 Rational Exponent . . . 73
6.2.3 More Than One Term . . . 77
6.3 Fit of the $1/N$ Expansion . . . 80
6.4 Conclusion . . . 83

**A Lattice Notation** **87**
A.1 Basic Definitions . . . 88
A.2 The $2d$ Nonlinear Sigma Model . . . 90

**B Monte Carlo** **92**
B.1 Critical Slowing Down . . . 93
B.2 Multi Cluster Algorithm . . . 94
B.3 Single Cluster Algorithm . . . 96
B.3.1 Single Cluster Algorithm for Nonlinear Sigma Models 98

**C Tables** **100**

# List of Figures


3.1 Continuum extrapolation in lattice QCD . . . . . 15
3.2 Continuum limit in the XY model . . . . . 16
3.3 Misleading power-like fit in the XY model . . . . . 16

4.1 Translation invariance when using free boundary conditions . 28
4.2 Time slice correlation function in the $O(8)$ model . . . . . 29
4.3 Decay rate of the time slice correlation function . . . . . 29
4.4 Acceptance of the Metropolis update in the improved estimator 40
4.5 Time Slice Correlation function measured with different estimators . . . . . 42
4.6 Relative error of the time slice correlation function . . . . . 42
4.7 Relative error of the finite volume mass . . . . . 42
4.8 Comparism of improved and standard estimator . . . . . 43
4.9 Time consumption of the improved measuring routine . . . . 44
4.10 Relative error vs. number of updates for the improved estimator 44

5.1 The function $K(a/L)$ and the coefficient $\tilde{a}_2(a/L)$ . . . . . 56
5.2 Continuum step scaling function in the limit $N \to \infty$ . . . . . 57
5.3 The continuum limit of the large $N$ step scaling function at $z = 0.4, 0.47, 0.53$ . . . . . 61

6.1 Lattice artifacts of the step scaling function, $N = 3$ . . . . . 66
6.2 Lattice artifacts of the step scaling function, $N = 4$ . . . . . 66
6.3 Lattice artifacts of the step scaling function, $N = 8$ . . . . . 67
6.4 Lattice artifacts of the step scaling function, $N = \infty$ . . . . . 67
6.5 Linear and quadratic fit of former $O(3)$ MC data . . . . . 70
6.6 Linear and quadratic fit of the $O(4)$ MC data . . . . . 71
6.7 Linear and quadratic fit of the $O(8)$ MC data . . . . . 72
6.8 Rational exponent fit of former $O(3)$ MC data . . . . . 74
6.9 Rational exponent fit of the $O(4)$ MC data . . . . . 75
6.10 Rational exponent fit of the $O(8)$ MC data . . . . . 76
6.11 Polynomial and PT fit of former $O(3)$ MC data . . . . . 78
6.12 Polynomial and PT fit of the $O(4)$ MC data . . . . . 79
6.13 Polynomial fit of the continuum values in $1/N$ . . . . . 82

6.14 Comparison of different estimates for the $N = 8$ continuum value . . . . . 82

# List of Tables


5.1 Exact values of the large $N$ lattice step scaling function . . . 62
5.2 Result of continuum limit fits of the large $N$ step scaling function . . . 62

6.1 MC data for $N = 4, 8$, former $N = 3$ MC data and $N = \infty$ values . . . 65
6.2 Linear and quadratic fits for $N = 3, 4, 8$ . . . 69
6.3 Rational exponent fits for $N = 3, 4, 8$ . . . 71
6.4 Polynomial and PT fits for $N = 3, 4$ . . . 77
6.5 Form independent lattice artifacts mutually fitted to the whole MC data . . . 81

C.1 Time slice correlation function and mass measured with three different estimators . . . 101
C.2 Numerical costs of different estimators . . . 102
C.3 Functions $K(a/L)$ and $\tilde{a}_2(a/L)$ . . . 103
C.4 Continuum step scaling function . . . 104

# Chapter 1

# Introduction

In the Standard Model of the known fundamental interactions of elementary particles quantum chromodynamics (QCD) is supposed to describe the phenomena of strong interactions. Whether this is really the case has to be verified by comparing theoretical predictions to experiments. But obtaining such predictions turns out to be rather difficult. Perturbative QCD is only applicable at high energy ($\geq$ 1 GeV) hadronic processes. This is due to the renormalization group behavior of the running coupling of QCD: it becomes large at low energies (100 MeV - 1 GeV) and small at high energy. The property of a vanishing running coupling at high energy (short distances) is called *asymptotic freedom*.

At the other end of the scale, at low energy, QCD is nonperturbative since the expansion parameter is not small. Also, noninteracting quarks (zeroth order of PT) are not a good approximation to explain the composite states which make up the spectrum of light hadrons. Consequently *nonperturbative* methods are needed to verify that the QCD Lagrangian describes physics of hadrons at low energy (mass spectrum, decay constants, low energy scattering).

The lattice regularization of a quantum field theory yields a formulation which is mathematically well defined, also at the nonperturbative level. It provides a momentum cutoff inversely proportional to the lattice spacing $a$. But it involves systematical errors due to discretization of space-time. Thus the regulator has to be removed before results are compared to the real world. This actually means a *continuum limit* should be taken.

In lattice QCD quantitative results are almost exclusively obtained using numerical simulations. In the course of this approach the continuum limit is taken by computing the quantities of interest for several values of $a$ and extrapolating the results to $a = 0$. Since simulation programs slow down proportionally to at least $a^{-7}$ for QCD, one cannot go to arbitrarily small lattice spacings. Therefore one usually fits a set of few data points and strongly relies on the theoretically expected behavior of the *lattice artifacts*.

To obtain meaningful results from a lattice regularized quantum field theory like lattice QCD, a detailed theoretical understanding of the approach to the continuum limit is required and extensive numerical studies are needed to confirm (or disprove) the expected behavior.

In this thesis the continuum limit of the lattice regularized *nonlinear sigma model* in two dimensions is investigated. The theory consists of $N$-component vectors that are constrained to the $(N-1)$ - sphere. The global $O(N)$ symmetry restricts the possible terms in the Lagrangian to a product of two derivatives. Although the theory looks very simple the nonlinear constraint leads to complex interactions and a rich phenomenology. The theory's two dimensional version shares with four dimensional gauge theories (like $SU(3)$-Yang-Mills of QCD) the property of being asymptotically free, at least in the weak coupling perturbative expansion. There is no mass term in the Lagrangian, but a mass gap is dynamically generated and determines

the large distance behavior of the two-point function.

Because of its similarity to the physically more relevant four dimensional gauge theories the $O(N)$ sigma model is used as a testing ground for perturbation theory predictions and for developing new methods. The $O(N)$ vector models can be simulated on the lattice very efficiently and for very small lattice spacings due to collective Monte Carlo updating methods [1] and improved estimators [2].

The continuum limit of the $O(3)$ model has been investigated for a long time. Thereby nonstandard (in the meaning of: not as expected from PT) cutoff effects have been discovered. The standard assumption for a scalar field theory are $\mathcal{O}\left(a^2\right)$ [3] lattice artifacts. In [4] Lüscher, Weisz and Wolff fitted their data to such an form, but the sign of the artifacts was opposite to the prediction. Later, in [5] and [6], the quadratic fit had to be rejected. The data is better fitted, when the artifacts are assumed to decay only linearly. At this point it was not clear whether this is a specialty to the $O(3)$ model or a general feature of the nonlinear sigma model.

For large $N$ the $O(N)$ sigma model can be solved exactly and the lattice artifacts can be studied analytically [7],[8],[9]. Such analysis suggests a cutoff dependence similar to PT. Both predict leading lattice artifacts proportional to $a^2$ [10],[11].

Starting from the numerical evidence for $N = 3$, which seems to be contradictory to PT and $N \to \infty$, the aim of this thesis is to measure the cutoff effects for $N = 4, 8$. The main interest is, how the behavior of the lattice artifacts transfers from small $N$ over intermediate towards the large $N$ limit.

Recently presented theoretical predictions for the low energy spectrum of the continuum $O(3)$ and $O(4)$ model yield continuum values for the measured step scaling function [12]. These prediction will be compared to the MC data and used to increase the significance when fitting the lattice artifacts.

In Chapter 2 a brief overview of lattice quantum field theory is given. First the functional integral quantization of field theories is introduced, concentrating on its application in the lattice regularization. Then the important transfer matrix formalism is outlined in the context of the nonlinear sigma model. Finally the correspondence between quantum field theory and statistical mechanics is pointed out.

The cutoff effects generated by the discretization in lattice simulations are discussed in Chapter 3. A summary of the situation in the nonlinear sigma model is given and the strategy to determine the lattice artifacts is explained.

The numerical methods used to implement this strategy are presented in Chapter 4. It is explained in detail how the *step scaling function* is measured on the lattice and how its approach to the continuum is computed. Furthermore the improved estimator, used to reduce the variance of the primary observable, is discussed.

In Chapter 5 the large $1/N$ expansion of the nonlinear sigma model on the lattice is introduced. The leading order is used to compute the step scaling function in the large $N$ limit and to examine the cutoff effects in this limit.

Finally, in Chapter 6, the results of the Monte Carlo simulations are presented and analyzed. The lattice artifacts for $N = 3, 4, 8, \infty$ are compared and the data is fitted to different forms.

In this thesis the undefined expression $N = \infty$ always refers to $N \to \infty$. Furthermore the Einstein summation convention and $\hbar = c = 1$ are used if not stated otherwise. Against the common habit of setting $a = 1$, in this thesis the dependence on the lattice spacing $a$ is mostly made explicit.

# Chapter 2

# Lattice Quantum Field Theory

## 2.1 Functional Integral Quantization

### 2.1.1 Path Integral in Quantum Mechanics

The objects of quantum mechanics are transition amplitudes or, equivalent, probability amplitudes. For example the probability amplitudes for a particle to move from $a$ to $b$ within the time $t$ is

$$\langle b|\mathrm{e}^{-\frac{i}{\hbar}Ht}|a\rangle\,, \tag{2.1}$$

where $H$ is the Hamiltonian of the particle

$$H = \frac{\hat{p}^2}{2m} + V(\hat{x})\,. \tag{2.2}$$

If the particle moves in a potential $V(\hat{x}) \neq 0$ the amplitude can in general not be calculated explicitly. But for the amplitude (2.1) a so called path integral representation exists that does not need such difficult concepts as non-commuting operators:

$$\langle b|\mathrm{e}^{-\frac{i}{\hbar}Ht}|a\rangle = \int \mathrm{D}x\ \mathrm{e}^{\frac{i}{\hbar}S[x]}\,, \tag{2.3}$$

where

$$S[x] = \int_0^t dt'\,\left[\frac{m}{2}\dot{x}^2 - V(x)\right], \qquad x(0) = a\,,\ x(t) = b\,, \tag{2.4}$$

is the action of the particle moving along the path $x(t)$ and $\int \mathrm{D}x$ means "integral over all possible paths". This is of course no mathematical definition and therefore eq. (2.3) has to be regularized, i.e. to give it a mathematically well defined meaning.

Since all paths occurring in eq. (2.3) are weighted by an exponential function with imaginary phase, oscillations and interference will occur. In the classical limit $\hbar \to 0$ the transition amplitude (2.1) should get contributions only from the classical path (defined by $\delta S = 0$). How can one see this? The phase $\varphi$ of the exponent is

$$\varphi = \frac{S}{\hbar} = 2\pi n \quad \to \quad n = \frac{S}{h}\,. \tag{2.5}$$

Thus the behavior of the phase depends on the ratio $S/h$:

> $S \gg h$:
> The exponential factor strongly oscillates with $S$. Since a small change of $S$ (a small change of the path) causes a $\Delta n > 1$ the phase $\varphi$ changes about several periods. Therefore these paths will give no contribution to the total amplitude. Only the classical path $x_c(t)$ satisfying $\delta S = 0$ and paths in a narrow tube around it will survive.

$S \approx h$:
A lot of paths with $\delta S \neq 0$ also contribute to the amplitude because $\Delta n < 1$, $\varphi$ does not change over a period and there will be no destructive interference for them.

The path integral representation of quantum mechanics leads to the same results as the canonical representation [13],[14]. Because of this equivalence one also speaks of the *path integral quantization*. This concept can be extended to field theories.

### 2.1.2 Quantization of Field Theories

The path integral formula eq. (2.3) holds for any quantum system, so it should be applicable in the case of quantum fields as well. The formalism is also called *functional integral*, for one is integrating over a set (or even space) of functions. But because of the fact that the integral is complex and strongly oscillating it is difficult to give it a satisfactory mathematical meaning. By introducing imaginary time this problem can be overcome. If the time coordinate is purely imaginary

$$x^0 = -ix^4\,, \qquad x^4 \in \mathbb{R}\,, \tag{2.6}$$

the Minkowski space-time metric for the coordinates $x^0, \ldots, x^3$ can be replaced by an Euclidean one using the coordinates $x^1, \ldots, x^4$. Then one speaks of *Euclidean quantum field theory*. The Euclidean formulation is the starting point of lattice field theory and that is the framework for the considerations in this thesis. For in quantum field theories all information is contained in the $n$-point correlation functions, one has to assure that they can be analytically extended to imaginary time [14].

For a real scalar field theory in $d$ dimensions and field amplitudes $\phi(x)$ the classical action reads

$$S_{\mathrm{cl}}[\phi] = \int \mathrm{d}^d x\, \mathcal{L}\,, \tag{2.7}$$

with the Lagrangian density

$$\mathcal{L} = \frac{1}{2}\partial_\mu \phi \cdot \partial^\mu \phi - V[\phi]\,. \tag{2.8}$$

In Euclidean space-time the action becomes

$$S[\phi] = \int \mathrm{d}^d x \left\{ \partial_\mu \phi \cdot \partial^\mu \phi + V(\phi) \right\}. \tag{2.9}$$

Then the $n$-point correlation functions of the Euclidean quantum field theory can be expressed as the moments of a measure

$$\langle \phi(x_1) \ldots \phi(x_n) \rangle = \int \mathrm{d}\mu\, \phi(x_1) \ldots \phi(x_n)\,, \tag{2.10}$$

where $\mathrm{d}\mu$ is formally written as

$$\mathrm{d}\mu = \frac{1}{Z}\mathrm{e}^{-S[\phi]} \prod_x \mathrm{d}\phi(x)\,, \tag{2.11}$$

with the Euclidean action in the weight factor. The normalization factor $Z$ is called (in analogy to statistical systems) *partition function*

$$Z = \int \prod_x \mathrm{d}\phi(x)\, \mathrm{e}^{-S[\phi]}\,. \tag{2.12}$$

The concept of functional integrals together with the concept of functional derivatives are important tools in modern quantum field theory.

The equations above have to be understood as formal expressions, for they consist of infinite dimensional integrals over an infinite number of degrees of freedom. In the majority of cases they have to be regularized before evaluating physical quantities, for example by thinking of the functional integral in eq. (2.12) as the limit of a well defined integral over a finite and discretized Euclidean space-time.

### 2.1.3 Lattice Regularization

Consider a quantum field theory living on a $d$ dimensional finite Euclidean lattice $\Lambda$ with lattice spacing $a$ and extensions $L_\mu$

$$\Lambda = \{x \,|\, x_\mu/a \in \mathbb{Z},\, x_\mu < L_\mu\}, \quad \mu = 1, \ldots, d\,. \tag{2.13}$$

Then a discretized version of the action eq. (2.9) has to be used:

$$S[\phi] = a^d \sum_x \left\{ \Delta^f_\mu \phi(x) \cdot \Delta^f_\mu \phi(x) + V(\phi) \right\}. \tag{2.14}$$

For details of the lattice notation, e.g. the definition of the forward derivative $\Delta^f_\mu$, the reader is referred to Appendix A. Then the product in eq. (2.12) is discrete and finite. But now two limits have to be considered: towards infinite volume and towards the continuum. The infinite volume limit is rather trivial and sometimes it is not even taken, but the finite size effects are rather used to study aspects of the theory (finite size scaling). One has to be more concerned about the continuum limit. The discretization eq. (2.14) is just one possible choice and if other versions are used, it has to be verified whether they reproduce the continuum action correctly.

### 2.1.4 The Nonlinear Sigma Model

In this section a scalar field theory in two dimensions will be studied. In $d = 2$ a scalar field is dimensionless; thus the Lagrangian has a dimensionless coupling and is renormalizable. The nonlinear $O(N)$ sigma model is made up

by a scalar $N$-component vector field $s = s(x) = (s_1(x), ..., s_N(x))$ subject to the constraint

$$s^2 = s(x) \cdot s(x) = \sum_{i=1}^{N} (s_i(x))^2 = 1\,. \tag{2.15}$$

Then the most general (up to a multiplicative constant) $O(N)$ symmetric Lagrangian with at most two derivatives is

$$\mathcal{L} = \partial_\mu s \cdot \partial^\mu s\,. \tag{2.16}$$

Therefore the action of the nonlinear sigma model is written as follows:

$$S = \frac{1}{2f} \int \mathrm{d}^2x\; \partial_\mu s(x) \cdot \partial^\mu s(x)\,, \tag{2.17}$$

where $f$ is called the *bare coupling*. The lattice version of this action is derived in Appendix A. Using the model's symmetry and the constraint one ends up with

$$S_{Lat} = -\beta \sum_{x,\mu} s(x)\, s(x + a\hat{\mu})\,, \tag{2.18}$$

where $\beta = 1/f$ is introduced in analogy to the inverse temperature in statistical systems and is referred to as bare coupling, too.

## 2.2 The Transfer Matrix Formalism

In quantum mechanics one also can introduce imaginary times, define an *Euclidean path integral* and discretize the time interval. Then the time evolution operator that shifts the states by one spacing in time is called *transfer matrix*. This concept leads to a very appealing description of the path integral (see e.g. [14]).

The transfer matrix for field theories will be introduced within the context of the nonlinear sigma model. Let the theory be defined on a 1+1 dimensional lattice $\Lambda$ with lattice spacing $a$. The starting point is the partition function

$$Z = \int \mathrm{D}s\, \mathrm{e}^{-S[s]}\,, \tag{2.19}$$

where

$$\mathrm{D}s = \prod_{x \in \Lambda} \mathrm{d}^N s(x) \tag{2.20}$$

and

$$S = -\beta \sum_{x=(x_1,x_2)\in\Lambda} \left[ s(x_1, x_2)\, s(x_1, x_2 + a) + s(x_1, x_2)\, s(x_1 + a, x_2) \right]\,. \tag{2.21}$$

To make the expressions more readable one defines the time slice field $s_t$ through

$$s_t = \{s(x_1, x_2) \,:\, x_2 = t\}\,, \qquad s_t(y) \equiv s(y,t)\,. \tag{2.22}$$

The action can be rewritten such that it is symmetric in adjacent time slices $s_t$ and $s_{t+a}$

$$S = -\beta \sum_t L[s_{t+a}, s_t]\,, \tag{2.23}$$

with

$$\begin{aligned} L[s_{t+a}, s_t] \;=\; & \sum_{x_1} \Big[ s_t(x_1)\, s_{t+a}(x_1) + \\ & + \frac{1}{2}\, s_t(x_1)\, s_t(x_1+a) + \frac{1}{2}\, s_{t+a}(x_1)\, s_{t+a}(x_1+a) \Big]\,. \end{aligned} \tag{2.24}$$

Now the transfer matrix $\Pi$ is defined by its matrix elements

$$\langle s_{t+a}|\Pi|s_t\rangle = \exp(-\beta\, L[s_{t+a}, s_t])\,. \tag{2.25}$$

The lattice Hamiltonian of the field theory is tightly connected to the transfer matrix, for the transfer matrix is the Euclidean time evolution operator, which shifts the state $|s_t\rangle$ by one $a$-unit

$$\Pi = \exp(-a\, H)\,. \tag{2.26}$$

Using the transfer matrix, the partition function becomes

$$Z = \int \mathrm{D}s \prod_t \langle s_{t+a}|\Pi|s_t\rangle\,. \tag{2.27}$$

The states $|s_t\rangle$ form a complete basis of the theory's Hilbert space

$$\int \mathrm{D}s\, |s_t\rangle\langle s_t| = \int \prod_{y,t} \mathrm{d}s(y,t)\, \Big|s(x)\,, x_2 = t\Big\rangle\Big\langle s(x)\,, x_2 = t\Big| = 1 \tag{2.28}$$

and therefore eq. (2.27) reduces to

$$Z = \lim_{N_t \to \infty} \mathrm{tr}\, \Pi^{N_t}\,. \tag{2.29}$$

Here $N_t \equiv T/a$ the number of the lattice sites in time direction. For finite time sites $N_t$ the partition function is that of a theory with finite time extension $T$ and periodic boundary conditions in time direction.

In order to compute the two-point correlation function, the multiplicative operator $\hat{s}_y$ is introduced

$$\hat{s}_y|s_t\rangle = \int \prod_{t'} \mathrm{d}s(y,t')\; |s_{t'}\rangle\langle s_{t'}|s_t\rangle = s(y,t)\,|s_t\rangle\,. \tag{2.30}$$

The functional integral formula for the two-point function

$$\langle s(x)s(y)\rangle = \int \mathrm{D}s\, \mathrm{e}^{-S[s]}\; s(x)s(y)\,, \tag{2.31}$$

then can be written in terms of the transfer matrix

$$\begin{aligned}\langle s(x)s(y)\rangle &= \int \mathrm{D}s\; s(x)s(y) \prod_t \langle s_{t+a}|\Pi|s_t\rangle \\ &= \lim_{N_t\to\infty} \frac{\mathrm{tr}\left(\Pi^{N_t-\tau/a}\;\hat{s}_{x_1}\;\Pi^{\tau/a}\;\hat{s}_{y_1}\right)}{\mathrm{tr}\,\Pi^{N_t}}\,,\end{aligned} \tag{2.32}$$

where $x_2 - y_2 = \tau > 0$ is assumed.

It is now assumed, that the lattice Hamiltonian $H$, defined in eq. (2.26), has a discrete spectrum with a non degenerate ground state $|0\rangle$ belonging to the eigenvalue $E_0$. Then the transfer matrix has a spectral decomposition

$$\Pi = \sum_i \mathrm{e}^{-a\,E_i}\;|i\rangle\langle i| \tag{2.33}$$

and since $E_0 < E_i\,,\ \forall\, i > 0$ the operator $\Pi^{N_t}$ becomes a projector on the ground state as $N$ goes to infinity

$$\Pi^{N_t} \overset{N_t\to\infty}{\approx} \mathrm{e}^{-a\,E_0\,N_t}\;|0\rangle\langle 0|\,. \tag{2.34}$$

This can be used to study the large distance behavior of the two-point function. For the sake of simplicity, let in eq. (2.32) the spatial coordinates coincide: $x_1 = y_1 = 0$. Then, plugging in the asymptotic behavior of $\Pi^{N_t}$ yields

$$\langle s(x)s(y)\rangle = \sum_i \left|\langle 0|\hat{s}_0|i\rangle\right|^2\; \mathrm{e}^{-(E_i-E_0)\,\tau}\,. \tag{2.35}$$

Consequently the two-point function decays exponentially for large distances $\tau$ like

$$\langle s(x)s(y)\rangle \sim \mathrm{e}^{-(E_1-E_0)\,\tau}\,. \tag{2.36}$$

The *correlation length* $\xi$, determining the decay rate, is defined through

$$\xi = \frac{1}{a\,M}\,, \qquad M = E_1 - E_0\,, \tag{2.37}$$

as the inverse of the *mass gap*.

## 2.3 Field Theory and Statistical Mechanics

The expression eq. (2.19) reminds one of a statistical mechanics problem's partition function. The integrand $\mathrm{e}^{-S}$ corresponds to the Boltzmann factor $e^{-\beta\mathcal{H}}$ and the action $S$ corresponds to $\beta\mathcal{H}$, where $\mathcal{H}$ can be interpreted

as the classical Hamiltonian of a 2 dimensional spin system. The inverse temperature of the spin system $\beta$ corresponds to the bare coupling in eq. (2.21). For example the $O(2)$ nonlinear sigma model corresponds to the XY-model, the $O(3)$ model to the Heisenberg model and so on.

The relation between field theory and statistical mechanics can be extended to many other quantities [15],[13]. For example the vacuum expectation value of the field corresponds to the mean magnetization per site and the two-point function corresponds to the spin-spin correlation function. From eq. (2.37) one sees that if the theory has a finite mass in the continuum limit then it should be possible to approach a limit where $a$ goes to zero and $m$ stays finite by simply tuning the inverse temperature of the underlying spin system (bare coupling of the field theory). In this limit the correlation length has to diverge. In statistical mechanics a point in the phase diagram where the correlation length diverges is called *critical point* and the system undergoes a second order *phase transition* when passing this point.

# Chapter 3

# Nonstandard Cutoff Effects

## 3.1 Cutoff Effects in Lattice Simulations

### 3.1.1 Lattice Regularization and Continuum Limit

The lattice regularization of quantum field theory is a mathematically well defined formulation. The measure in the partition function becomes a discrete product and for a finite lattice it becomes a finite product. The theory then has a finite number of degrees of freedom. The lattice also provides a momentum cutoff inversely proportional to the lattice spacing $a$.

But the discretization involves systematical errors. In general, all quantities may now depend on the lattice spacing and, if the lattice is finite, on the lattice size. Thus a continuum limit should be taken before the results are compared with the real world or continuum regularization schemes.

Quantitative results in the lattice regularization are obtained using numerical simulation techniques like Monte Carlo. Only discrete changes of the lattice spacing are possible in this simulations. In practice, the continuum limit is taken by computing the quantities of interest for several values of $a$ and extrapolating the results to $a = 0$. Since the simulation programs slow down proportionally to at least $a^{-(d+x)}$, where $d$ is the dimension of the lattice and $x$ is a theory dependent constant (for QCD $d = 4$, $x = 3$), one cannot go to arbitrarily small lattice spacings.

Therefore a detailed theoretical understanding of the approach to the continuum limit is required. This understanding is crucial when extrapolating a few measurements at small lattices to the continuum: the continuum value strongly depends on the form of the fit. Every theoretical prediction for the so called *lattice artifacts* has to be checked and extensive numerical studies are needed to confirm (or disprove) the expected behavior.

### 3.1.2 Perturbative Understanding of Lattice Artifacts

Almost 25 years ago Symanzik has done a seminal work studying the cutoff effects in perturbation theory [3]. His idea was to describe the lattice theory through an effective continuum theory: the cutoff dependence is made explicit through additional terms in the continuum Lagrangian proportional to $a^k$, $k \geq 1$. Thereby the lattice is assumed as a substructure that is only visible at energies $\sim 1/a$. In a sense, the continuum theory with the added terms is then a low energy effective theory of some underlying more complete theory.

The possible combinations of operators entering the additional terms are determined by the symmetries of the lattice theory and by dimensional considerations. For scalar field theories this analysis reveals that the leading lattice corrections vanish like $a^2 (\ln a)^l$. In the case of fermionic theories like QCD the leading term is proportional to $a (\ln a)^l$, where $l$ is the loop order in perturbation theory.

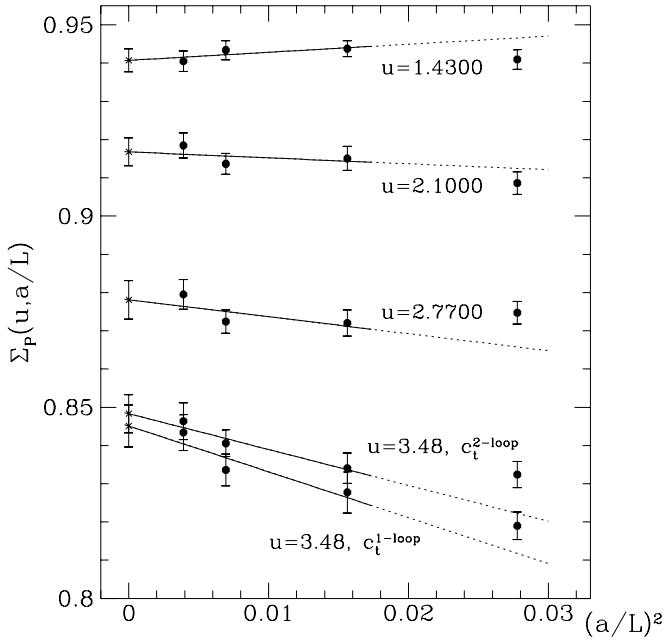



FIGURE 3.1: The plot shows continuum extrapolations of a step scaling function in quenched lattice QCD for different parameter values. The extrapolations were done excluding the point of smallest lattice at $L/a = 6$. Since the action used in the simulations is $\mathcal{O}(a)$ improved the leading artifacts are $\sim a^2$. See [17] for details.

These statements hold in every order of perturbation theory. Later Symanzik extended his concept and proposed a method to accelerate the approach to the continuum, which is today known as the Symanzik improvement programme [16]. The idea is to add irrelevant operators to the lattice action that cancel the leading lattice artifacts ($\mathcal{O}(a)$ improvement in lattice QCD).

The results of Symanzik are assumed to hold beyond perturbation theory and extrapolations are done accordingly. In QCD extrapolations proportional to $a^2$ (for the improved action) seem to work well. See for example Fig. 3.1 taken from [17]. However, one has to care about the limitations of these predictions. Firstly, the stated lattice artifacts are the leading order of an asymptotic expansion: it is not clear what lattice spacing is "sufficiently small". Secondly, there could be nonperturbative terms in addition to the cutoff effects.

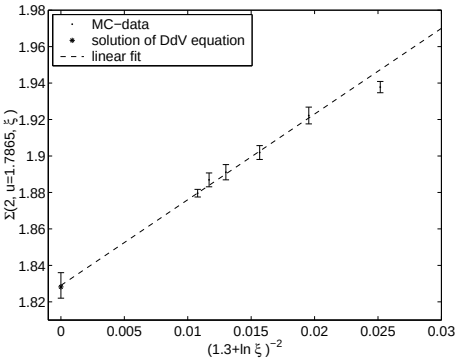



FIGURE 3.2: Continuum limit of the step scaling function in the XY model. The MC data plotted over the squared inverse logarithm of the infinite volume correlation length. A linear fit to this term reproduces the analytically known continuum value. The plot is taken from[19].

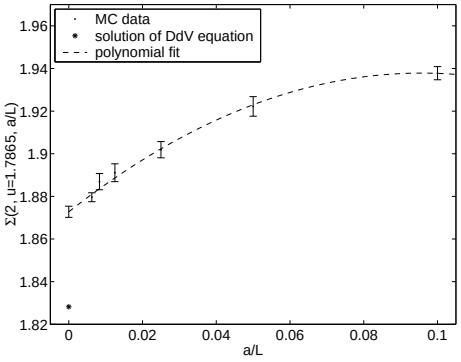



FIGURE 3.3: The same data as in the left figure but plotted over $a/L$. The predicted continuum value could not have been confirmed, if power-like lattice artifacts $\propto (a/L)^2 + c(a/L) + d$ were assumed. The plot is taken from[19].

## 3.2 Cutoff Effects in the Nonlinear Sigma Models

### 3.2.1 XY Model in 2d - A Pathological Case

The nonlinear sigma model with two components per vector at each lattice site corresponds to the XY model of statistical mechanics. This model is some sort of outstanding and differs a lot from the models with $N \geq 3$. The symmetry group is the abelian $O(2)$. Furthermore the theory is not asymptotically free in the conventional sense and rather undergoes a special phase transition at finite temperature. There exist exact theoretical predictions not only for the continuum value of the finite volume mass gap, but also for the lattice artifacts [18]. These theoretical predictions have recently been compared to numerical results by Tomasz Korzec in his diploma thesis [19], [20]. The Monte Carlo data is in good agreement with the analytical results. The theoretical considerations lead to an expansion for the cutoff effects in the inverse logarithm of the infinite volume correlation length [18] rather then the usual powers of the lattice spacing. If the data is fitted by power-like lattice artifacts the extrapolation misses the exact continuum value, see Fig. 3.2 and Fig. 3.3.

Since the $O(N)$ nonlinear sigma model with $N = 2$ is a special case in this family of models, this result does not have direct consequences for the models with $N \geq 3$. It is not asymptotically free in the conventional sense and has an abelian symmetry. Nevertheless this example teaches us to be very careful about extrapolations to the continuum: they may be misleading.

### 3.2.2 Numerical Results for $N = 3$

In the late 80s, early 90s collective Monte Carlo updating algorithms were proposed with reduced or eliminated critical slowing down (see Appendix B). The so called *cluster algorithms* made MC simulation of spin systems deep inside the critical region (at small lattice spacing) possible. Since the $O(N)$ models are equivalent to spin systems the efficient algorithms can be used to simulate these models as well. Thereby a detailed investigation of the continuum limit is possible and perturbative predictions can be tested. The outcome of such studies in the $O(3)$ model was surprising.

In 1991 Lüscher, Weisz and Wolff proposed a finite size scaling technique to numerically compute the LWW running coupling $\overline{g}^2(L)$ in a finite volume from large volumes (nonperturbative) down to very small scales (perturbative) [4]. They measured the so called *step scaling function* at different lattice spacings and extrapolated their data to the continuum. The data's relative error is of $\mathcal{O}\left(10^{-3}\right)$. For the form of the lattice artifacts they used the expected $a^2$ term and found no disagreement. But the sign of the lattice artifacts they observed at $\overline{g}^2(L) = \mathcal{O}(1)$ is opposite to that obtained in PT.

Then, in 2000 Hasenfratz and Niedermayer studied the behavior of asymptotically free spin and gauge models when their continuous symmetry is replaced by discrete subgroups [5]. They measured another running coupling, the renormalized zero momentum 4-point coupling $g_{\mathrm{R}}$, with a relative error of $\mathcal{O}\left(10^{-4}\right)$. The higher precision revealed nonstandard cutoff effects. A fit to different forms suggests $\mathcal{O}(a)$ lattice artifacts rather then the Symanzik $\mathcal{O}\left(a^2\right)$.

Finally, in 2001 Hasenbusch, Hasenfratz, Niedermayer, Seefeld and Wolff did high precision measurements of the renormalized zero momentum 4-point coupling $g_{\mathrm{R}}$ and of the LWW running coupling $\overline{g}^2(L)$ in order to examine this unexpected behavior [6]. Their data confirms the anomalous linear behavior of the lattice artifacts.

### 3.2.3 Large $N$ Predictions

The $O(N)$ nonlinear sigma model can be solved exactly in the limit $N \to \infty$ and a systematical $1/N$ expansion can be derived (see [9] for a review of the topic). Caracciolo et. al. analytically studied the corrections to finite size scaling in the lattice model for $N = \infty$ [10], [11]. For the step scaling function they found artifacts decaying like

$$(a/L)^2 \left(\ln(a/L)\right)^{-q}, \qquad q = -1, 0, 1, 2, \ldots . \tag{3.1}$$

They point out that the negative powers are unexpected from the point of view of perturbation theory.

In the first section of Chapter 5 the derivation of the leading order of the $1/N$ expansion in the lattice model is presented. Then in Section 5.2.2

the exact equations for $N = \infty$ are used to numerically compute the step scaling function of the running coupling in the continuum and for finite lattice spacings. In order to extract the leading lattice artifacts, first a weak coupling expansion is implemented and then asymptotic expansions of the momentum sums involved are used. This calculation tightly follows unpublished notes by Peter Weisz [21]. The approach to the continuum is found to be from below for small couplings and from above for couplings of $\mathcal{O}(1)$. The artifacts seem to decrease like $(a/L)^2 \ln(a/L)$ for small $a/L$.

### 3.2.4 Summary of the Situation for $N \geq 3$

Summarizing the last sections the following can be said about the leading lattice artifacts suggested by different methods tackling the problem in the nonlinear sigma model:

$$\begin{aligned} \text{Perturbation theory} &\sim (a/L)^2 \ln(a/L) \\ \text{MC data for } N = 3 &\sim (a/L) \\ \text{Large } N \text{ limit} &\sim (a/L)^2 \ln(a/L) \end{aligned}$$

It would be interesting to see how the behavior of the lattice artifacts transfers from small $N$ over intermediate towards the $N = \infty$ limit.

## 3.3 Testing Predictions

The nonlinear sigma model is a scalar $N$-component vector field subject to a nonlinear constraint (see Section 2.1.4). Its two dimensional version shares with four dimensional gauge theories the property of being asymptotically free, at least in the weak coupling perturbative expansion. With the efficient cluster algorithms introduced in Appendix B.2 and B.3 it is possible to perform MC simulations of large statistics close to the continuum limit. Also an improved estimator used to reduce the variance of observables, hence to reduce the numerical costs for the envisaged relative error is presented.

Therefore the nonlinear sigma model in $2d$ is a perfect testing ground for PT predictions. High precision determination of the approach to the continuum limit is possible. In the following sections the quantity for which the approach to the continuum is studied, will be discussed. First an introduction to finite volume effects is given. Then the step scaling function is defined in the continuum theory and finally the effects of finite lattice spacing are considered.

### 3.3.1 Field Theory in a Finite Volume

Suppose a field theory in compact space, i.e living in a periodic box with finite spatial size $L$ and infinite extent in time direction. The natural exter-

nal scale in such a world is the box size $L$. Large $L$ correspond to the low energies and small $L$ correspond to high energies. Therefore one may define a coupling running with the box size.

The two dimensional $O(N)$ nonlinear sigma model will be studied in such an environment. The continuum action is

$$S = \frac{1}{2f} \int \mathrm{d}^2x \; \partial_\mu s(x) \cdot \partial^\mu s(x) \tag{3.2}$$

and the vectors $s(x)$ are subject to the constraint $s(x)^2 = 1$. Since the space is compact the energy spectrum is discrete. Lüscher pioneered in computing the low-lying states in perturbation theory to one-loop order [22]. The ground state is found to be unique and the lowest excited states make up an $O(N)$ vector multiplet. Their energy is

$$M(L) = \frac{N-1}{2L} \left\{ f_{\mathrm{MS}} + \alpha \, f_{\mathrm{MS}}^2 + O(f_{\mathrm{MS}}^3) \right\} , \tag{3.3}$$

with the coefficient

$$\alpha = \frac{N-2}{4\pi} \left[ \ln \mu^2 L^2 - \ln 4\pi - \Gamma'(1) \right] , \tag{3.4}$$

where $f_{\mathrm{MS}}$ is the renormalized coupling constant in the dimensional regularization scheme with minimal subtraction (MS) and $\mu$ is the normalization mass. In leading order $f_{\mathrm{MS}}$ coincides with the bare coupling $f$. Today the mass gap is known on the 3-loop level [23].

The mass gap eq. (3.3) is in leading order proportional to $f_{\mathrm{MS}}$ and therefore Lüscher, Weisz, Wolff [4] defined a running coupling through

$$\overline{g}^2(L) = \frac{2}{N-1} \, M(L) \, L \, , \tag{3.5}$$

called LWW running coupling.

### 3.3.2 Step Scaling Function

The Callan-Symanzik $\beta$-function for the running coupling $\overline{g}^2(L)$ is defined by

$$\beta(\overline{g}^2) = -L \frac{\partial \overline{g}^2}{\partial L} \, . \tag{3.6}$$

In perturbation theory it can be expanded in powers of the coupling

$$\beta(u) \sim -u^2 \sum_{l=0}^{\infty} b_l \, u^l \, . \tag{3.7}$$

The $\beta$-function describes the variation of $\overline{g}^2$ when the external scale (the box size $L$) is changed infinitesimally. The first two coefficients are independent

of the regularization scheme and of the definition of the running coupling, i.e. are universal

$$b_0 = \frac{N-2}{2\pi}\,, \qquad b_1 = \frac{N-2}{(2\pi)^2}\,. \tag{3.8}$$

The next two coefficients are also known for the MS scheme [23]. Since $b_0$ is positive, $\overline{g}^2$ vanishes logarithmically as $L$ goes to zero, i.e. for high energies

$$\begin{aligned} -L\frac{\partial u}{\partial L} &\overset{L\to 0}{=} -b_0 u^2 \\ \to \quad \frac{du}{u^2} &= b_0 \frac{dL}{L} \\ \to \quad \overline{g}^2(L) &\overset{L\to 0}{=} \frac{1}{b_0 \ln(c/L)}\,. \end{aligned} \tag{3.9}$$

A theory with a running coupling that vanishes for high energies is called *asymptotically free*.

The $\beta$-function was introduced to describe infinitesimal changes of the box size $L$. Now the *step scaling function* $\sigma(s,u)$ describes the effect of a finite change of the external scale

$$\sigma(s, \overline{g}^2(L)) = \overline{g}^2(sL)\,, \tag{3.10}$$

where $s$ is the positive number by which the box size $L$ is scaled $L' = sL$, e.g. $s = 2$. The step scaling function can be regarded as an integrated version of the $\beta$-function

$$\ln s = \int_u^{\sigma(s,u)} \frac{\mathrm{d}v}{\beta(v)}\,. \tag{3.11}$$

Using this relation and eq. (3.7) an expansion of $\sigma(s,u)$ can be derived

$$\sigma(s,u) = u + \sigma_0(s)u^2 + \sigma_1(s)u^3 + \sigma_2(s)u^4 + \dots\,, \tag{3.12}$$

with coefficients

$$\sigma_0(s) = b_0 \ln s\,, \tag{3.13}$$

$$\sigma_1(s) = b_0^2 (\ln s)^2 + b_1 \ln s\,, \tag{3.14}$$

$$\sigma_2(s) = b_0^3 (\ln s)^3 + \frac{5}{2} b_0 b_1 (\ln s)^2 + b_2 \ln s\,. \tag{3.15}$$

In [4] the step scaling function was proposed and defined in order to relate the coupling at small scales $L$ (perturbative) to the coupling in large volumes (nonperturbative).

The step scaling function can easily be measured on the lattice. One has to tune the bare coupling until the renormalized coupling $\overline{g}^2(L) = u$ is measured at lattice size $L/a$. Then $L$ is scaled by a factor, say $s = 2$, while the bare coupling is kept fixed. Now $\overline{g}^2(2L)$ is measured and the result is equal to $\sigma(2,u)$ up to lattice artifacts.

### 3.3.3 Lattice Effects

The step scaling function $\sigma(2,u)$ is defined in the continuum theory. When measured on the lattice one expects lattice artifacts. One defines a *lattice step scaling function* to care about such effects. This means on the lattice one does not measure $\sigma(2,u)$ but rather

$$\Sigma(s,\overline{g}^2(L),a/L)=\overline{g}^2(sL)\,. \tag{3.16}$$

It is expected, of course, that the lattice step scaling function $\Sigma(s,u,a/L)$ approximates $\sigma(2,u)$ well, if $a/L$ is small enough. In perturbation theory one expects an expansion in analogy to eq. (3.12)

$$\Sigma(s,u,a/L)=u+\Sigma_0(s,a/L)u^2+\Sigma_1(s,a/L)u^3+\dots\,, \tag{3.17}$$

but here the coefficients will additionally depend on the lattice spacing. From the analysis of Symanzik (Section 3.1.2) one expects

$$\Sigma_l(s,a/L)-\sigma_l(s)=\mathcal{O}\left(a^2(\ln a)^{l+1}\right)\,. \tag{3.18}$$

Thus the lattice artifacts of $\Sigma(s,u,a/L)$ decay roughly like $a^2$.

### 3.3.4 Strategy

In Section 3.2.4 it was pointed out that in the $O(N\geq 3)$ models the behavior of the lattice artifacts is inconsistent. There are predictions available from PT and the large $N$ limit. They are compatible at leading order. But the MC data for the $N=3$ model does not follow these predictions. Since the two dimensional nonlinear sigma model has crucial properties in common with the physically more important four dimensional Yang-Mills theories (like asymptotic freedom, nonabelian symmetry) it is of general interest where these inconsistencies come from or how the different behavior fits together.

The aim of this thesis is to investigate the cutoff dependence for intermediate $N$, for example $N=4,8$. The question is whether the nonstandard linear behavior in the $O(3)$ model is also found for higher $N$ or if this is a singularity and a specialty of $O(3)$.

Since the cutoff effects are studied as $N$ changes the measured quantities will also depend on $N$. So from now on this dependence will be explicitly indicated, i.e. the lattice step scaling function reads

$$\Sigma(s,u,N,a/L)\,. \tag{3.19}$$

This quantity will be measured at a fixed value of the renormalized coupling for different lattice spacings, i.e. for different lattice sizes (size in the computer memory, the physical size is kept fixed) and different $N$. To fix the renormalized coupling one has to tune the bare coupling till the desired

value is measured. The result will be plotted over $a/L$ and fits to several forms will be tested.

The analysis will also profit from a very recently published result concerning an analytical approach to the mass gap, and thus the step scaling function, in the $O(3)$ and $O(4)$ model [12]. The continuum value will be used to constrain the fits and increase the $\chi^2$ values.

# Chapter 4

# Numerical Methods

## 4.1 Computation of the Step Scaling Function

The continuum limit of the step scaling function in the nonlinear $O(N)$ invariant sigma model for intermediate $N$ is to be studied using MC simulation. In the last chapter the step scaling function was introduced in the context of the LWW coupling (Section 3.3.2). Since the definition of this coupling

$$\overline{g}^2(L) = \frac{2}{N-1}\, M(L)\, L\,, \tag{4.1}$$

is not suited for a large $N$ expansion (which is considered in Chapter 5), it is substituted by the renormalization group invariant variable

$$z = M(L)\, L\,. \tag{4.2}$$

The coupling $\overline{g}^2(L)$ could have been made well defined in the limit $N \to \infty$ by rescaling the bare coupling $f \to N/f$. But using the variable $z$ the results for different $N$ are easier to compare. Note that for $N = 3$ $z$ equals $\overline{g}^2$ by coincidence.

Now, the lattice step scaling function for the variable $z$ is defined through

$$\Sigma_z(2, u, N, a/L) = M(2L)\, 2L\,, \qquad u = M(L)\, L\,, \tag{4.3}$$

where the scale factor $s$ is set to two. This is also the value used in the MC simulations. Eq. (4.3) means one has to measure the finite volume mass gap $M$ two times in order to determine $\Sigma_z$ at a certain value $M(L)\, L = u_0$. First one has to tune the bare coupling till $u = u_0$ is measured on the "small" lattice with $L/a$ spatial sites. Then one goes to the "large" lattice ($2L/a$ spatial sites) and measures $M(2L)\, 2L$ using the bare coupling determined before. This is repeated for several values of $L/a$, say $L/a = 8, 10, 12, 16, \ldots$, giving $\Sigma_z$ at $z = u_0$ for different lattice spacing $a/L$. In Section 4.1.3 it is explained how the mentioned tuning is done and how the involved statistical and systematical errors are treated.

Until now it was not mentioned how the finite volume mass gap is measured. This topic is postponed to the following sections. First some general characteristics of the MC runs are stated:

*Lattice geometry* The simulations were carried out on simple square lattices

$$L \times T\,, \qquad T = 5L\,, \tag{4.4}$$

where $L/a$ and $T/a$ are the number of lattice sites in the spatial and temporal direction respectively. This strip geometry is dictated by the determination of the mass gap, which is defined through the large distance behavior of the time slice correlation function (see Section 4.1.1).

*Action* The action used is the standard nearest-neighbor action (see Appendix A.2)

$$S = -\beta \sum_{x,\mu} s(x)\, s(x + a\hat{\mu})\,, \qquad \mu = 1, 2\,. \tag{4.5}$$

*MC updating* For the $O(N)$ vector model very efficient cluster algorithms are known, that reduce or even eliminate critical slowing down (see Appendix B for details). Throughout the simulations the single cluster algorithm proposed by Wolff [1] is used.

*Improved estimator* The concept of improved estimators is illuminated in Section 4.2. There an estimator proposed by Hasenbusch [2] is presented and a detailed explanation how the ideas are implemented is given. Furthermore the efficiency and the accuracy of the improved estimator is tested.

The above-mentioned improved estimator is derived for the time slice correlation function. This is the primary observable in the MC simulations. Now this quantity will be defined and then in Section 4.1.2 follows an explanation how the mass gap is extracted from this correlator.

### 4.1.1 Time Slice Correlation Function

In Section 2.2 the mass gap was introduced as the difference between the ground state and the first excited state. It was found to govern the large distance behavior of the two-point function eq. (2.36). There the effect of non vanishing momenta was treated a little bit sloppy. Therefore the space averaged multiplicative field operators are now considered:

$$\hat{s} = \sum_{y} \hat{s}_y\,, \tag{4.6}$$

where $\hat{s}_y$ was defined as (recall eq. (2.30))

$$\hat{s}_y |s_t\rangle = \int \prod_{t'} \mathrm{d}s(y, t')\, |s_{t'}\rangle\langle s_{t'}|s_t\rangle = s(y, t)\, |s_t\rangle\,. \tag{4.7}$$

Now the zero momentum (nothing else means space averaged) or time slice (because $\hat{s}$ is some sort of time slice spin) correlation function can be defined

$$C(\tau) = \langle \hat{s}\, \hat{s} \rangle_\tau = \sum_{x_1, y_1} \langle s(x) s(y) \rangle_{x_2 - y_2 = \tau}\,, \tag{4.8}$$

whose large distance behavior is now studied a bit more seriously. Using the transfer matrix formalism the vacuum expectation value in eq. (4.8) can be written as (see eq. (2.32), $x_2 - y_2 = \tau > 0$)

$$C(\tau) = \lim_{N_t \to \infty} \frac{\mathrm{tr}\left(\Pi^{N_t - x_2/a}\, \hat{s}\, \Pi^{\tau/a}\, \hat{s}\, \Pi^{y_2/a}\right)}{\mathrm{tr}\, \Pi^{N_t}}\,. \tag{4.9}$$

For a finite number of sites in time direction $N_t \equiv T/a$ this equation holds only if *periodic boundary conditions (pbc)* are imposed. Taking the trace then means summing over all eigenstates $|n\rangle$ of the transfer matrix, i.e. of the theory's Hamiltonian (eq. (2.26))

$$C(\tau)_{\text{pbc}} = \frac{1}{Z} \sum_n \langle n| \, \Pi^{N_t - x_2/a} \, \hat{s} \, \Pi^{\tau/a} \, \hat{s} \, \Pi^{y_2/a} \, |n\rangle \, . \tag{4.10}$$

Since the theory is $O(N)$ invariant the Hamiltonian commutes with the Casimir operator of the Lie algebra of $O(N)$ and the eigenstates can be characterized by a spin quantum number. The ground state $|0\rangle$ is not degenerate and has spin 0 [24]. Let us assume the ground state energy to vanish: $E_0 = 0$. Then all excited $O(N)$ invariant states have energies of at least $4\pi/L$, apart from the lowest excited state. The mass gap, like all other excited states, has in leading order PT an energy of $\mathcal{O}\,(1/\beta L)$. In the sum in eq. (4.10) all these states occur. To use the asymptotic behavior eq. (2.34) of $\Pi^x$ and thus to single out the mass gap one would have to go to very large lattices (in time direction) and would still fight with a very noisy signal.

Therefore one rather considers *free boundary conditions (fbc)* in time direction. Then the two-point function reads

$$C(\tau)_{\text{fbc}} = \lim_{N_t \to \infty} \frac{1}{Z} \langle \psi| \, \Pi^{N_t - x_2/a} \, \hat{s} \, \Pi^{\tau/a} \, \hat{s} \, \Pi^{y_2/a} \, |\phi\rangle \, . \tag{4.11}$$

Free boundary conditions mean constant fields $\psi$ and $\phi$ at $t = 0$ and $t = T - a$. Constant fields have a constant wave function, are $O(N)$ invariant and spin 0. Plugging in the spectral decomposition eq. (2.33) of $\Pi$ one gets

$$\begin{aligned} C(\tau)_{\text{fbc}} \quad = \quad & \lim_{N_t \to \infty} \frac{1}{Z} \sum_{i,k} \langle \psi|i\rangle \, \langle k|\phi\rangle \, \mathrm{e}^{-E_i \, (a \, N_t - x_2) - E_k \, y_2} \times \\ & \times \langle i| \, \hat{s} \, \Pi^{\tau/a} \, \hat{s} \, |k\rangle \, . \end{aligned} \tag{4.12}$$

For $\langle \varphi|n\rangle = 0$ if $\varphi$ and $n$ have different spin only those states $|i\rangle$ contribute that have spin 0. Since the excited states with spin 0 have energies of at least $4\pi/L$ they are exponentially suppressed. So for large $N_t$ and $y_2/a$, while $\tau = x_2 - y_2$ is kept fixed, only the ground state will survive (remember $E_0 = 0$)

$$C(\tau)_{\text{fbc}} \propto \langle 0| \, \hat{s} \, \Pi^{\tau/a} \, \hat{s} \, |0\rangle \, . \tag{4.13}$$

Using again the spectral decomposition of the transfer matrix this reduces to

$$C(\tau)_{\text{fbc}} \propto \mathrm{e}^{-E_i \, \tau} \sum_i |\langle 0| \, \hat{s} \, |i\rangle|^2 \, . \tag{4.14}$$

The operators $\hat{s}_x$, and so $\hat{s}$, are spin 1 operators. Therefore only spin 1 states contribute in eq. (4.14). The lowest excited state, giving the mass gap $M$, has spin 1 and $M$ is of $\mathcal{O}\left(1/2\beta L\right)$. All other excited $O(N)$ invariant states have energies of at least $4\pi/L$. So, for $\tau \geq L$, they are exponentially suppressed. Hence the finite volume mass gap may be evaluated from (omitting the subscript fbc from now on)

$$M(L) = -\lim_{\tau\to\infty} \frac{\partial}{\partial\tau} \ln C(\tau)\,. \tag{4.15}$$

In the calculations above the normalization factor $1/Z$ was dropped at some point. The partition function $Z$ can be defined imposing free boundary conditions along similar steps as for the time slice correlation function. At the end it is a constant factor that can be omitted.

Also the calculations were performed under the tacit assumption of translation invariance in time direction (e.g. in the step from eq. (4.12) to eq. (4.13)). Clearly, this assumption is not satisfied at the boundaries. But far away from the boundaries, where the boundary fields have already decayed to the ground state, there should be a region of approximately valid translation invariance. The invariance is used in the MC simulations to improve the variance of $C(\tau)$ by scanning $x_2$ over this region. Thus

$$D(\tau) = \frac{a}{T_{\mathrm{inv}} - \tau + a} \sum_{x_2=t_1}^{t_2-\tau} \frac{a^2}{L^2} \sum_{x_1,y_1} \langle s(x_1,x_2)\, s(y_1,x_2+\tau)\rangle_{\mathrm{MC}}\,, \tag{4.16}$$

is measured, where it is indicated that the expectation value now refers to the average over a set of configurations produced by a MC algorithm. The first sum is over the region were translation invariance is assumed

$$T_{\mathrm{inv}} = t_2 - t_1\,. \tag{4.17}$$

However, it depends on $\tau$ how much of $T_{\mathrm{inv}}$ can be used to average over (note the upper limit of the sum). Since the excited states in the boundary fields decay with a rate of $\mathcal{O}\left(4\pi/L\right)$ a sensible choice would be (remember $T = 5L$)

$$t_1 = L\,, \quad t_2 = T - L \quad \rightarrow \quad T_{\mathrm{inv}} = 3L\,. \tag{4.18}$$

In order to verify these assumptions the time slice correlation function $C(\tau)$ is measured for a constant time separation $\tau$ but varying position $x_2$ of the first time slice

$$C(\tau, x_2) = \frac{a^2}{L^2} \sum_{x_1,y_1} \langle s(x_1,x_2)\, s(y_1,x_2+\tau)\rangle_{\mathrm{MC}}\,, \tag{4.19}$$

Thus a profile of the whole time extension is drawn. In Fig. 4.1 this is done for a $10 \times 50$ lattice. From the plot one concludes that "boundary effects" decay rapidly and that in the region $[L, T-L]$ translation invariance is a justified assumption. Thus it is justified to use eq. (4.16) to determine the time slice correlation function.

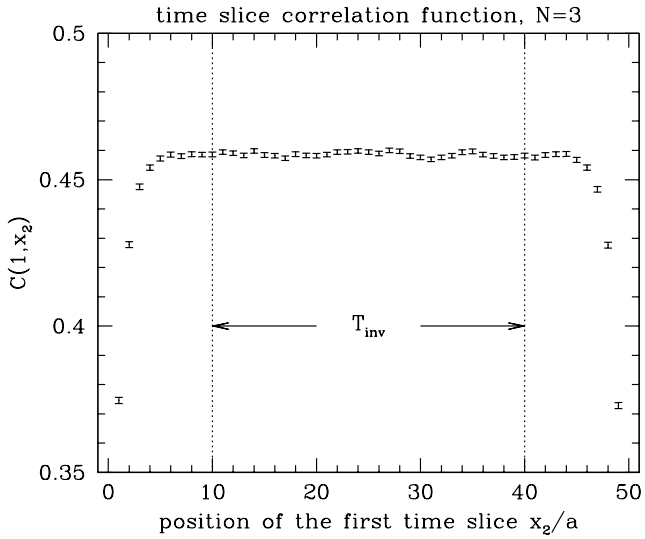



FIGURE 4.1: The time slice correlation function measured in the $O(3)$ model at $\beta = 1.5699$, $M(L)\,L = 1.292(2)$ on a $10\times 50$ lattice. The separation of the time slices is $a$ and the whole time extension $T/a = 50$ of the lattice is scanned. In the interval $[L, T-L]$ translation invariance can be assumed: $T_{\rm inv} = 3L$.

### 4.1.2 Finite Volume Mass Gap

Now eq. (4.15) is to be applied in order to evaluate the mass gap from the measurements eq. (4.16). A typical set of data is plotted in Fig. 4.2. Since the higher states are suppressed by a factor of $\exp(-4\pi a/L)$ they should have vanished for $\tau \approx L$ and the mass gap should govern the exponential decay of $D(\tau)$, i.e.

$$M(L) \stackrel{\tau\to L}{=} \frac{1}{a}\,\ln\left[\frac{D(\tau)}{D(\tau+a)}\right]. \tag{4.20}$$

If the l.h.s. is plotted (Fig. 4.3), it may be called the decay rate of $D(\tau)$, a plateau is observed starting from somewhere at $\tau = L/2$.

Since the relative error does not increase with the time separation (this is due to the use of the improved estimator discussed in Section 4.2) one could use the plateau and average the mass gap

$$M(L) = \frac{1}{L - t_p + a}\sum_{\tau=t_p}^{L}\ln\left[\frac{D(\tau)}{D(\tau+a)}\right]. \tag{4.21}$$

The time separation $t_p$ where this plateau average can start is defined through the claim, that the contribution from higher states is beneath the envisaged error of the mass gap. To calculate $t_p$, suppose the exponential decay of the time slice correlation function as a superposition of two terms:

$$D(\tau) \approx a_1\,\mathrm{e}^{-M\tau} + a_2\,\mathrm{e}^{-m_1\tau} = a_1\,\mathrm{e}^{-M\tau}\left(1 + \frac{a_2}{a_1}\mathrm{e}^{-\Delta m\,\tau}\right), \tag{4.22}$$

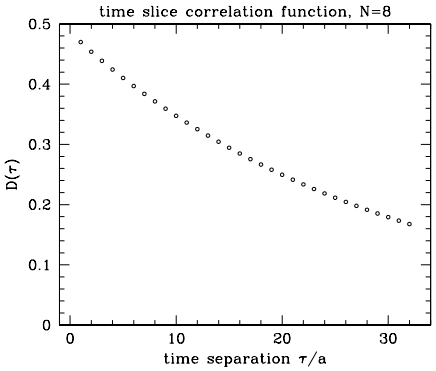



FIGURE 4.2: Time slice correlation function measured for time separations ranging from 1 to 32. The errors are too small to be displayed on this scale. ($O(8)$ model at $\beta = 6.8070$ on a $32 \times 160$ lattice)

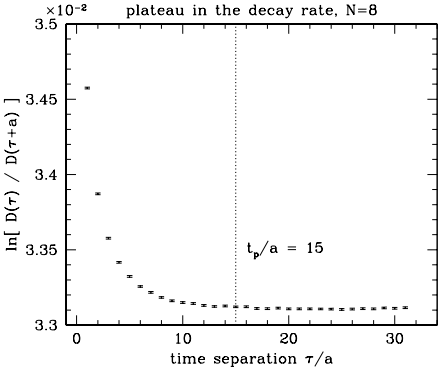



FIGURE 4.3: The decay rate of the time slice correlation function shown in Fig. 4.2 has a plateau. There one can read off the mass gap. The higher states have died out.

the mass gap $M$ and one higher state $m_1$ ($\Delta m = m_1 - M > 0$). Then, assuming $\exp(-\Delta m\,\tau) \ll 1$, the mass gap is

$$M \approx \frac{1}{a}\,\ln\left[\frac{D(\tau)}{D(\tau+a)}\right] + \frac{1}{a}\,\frac{a_2}{a_1}\mathrm{e}^{-\Delta m\,\tau}\,(1-\mathrm{e}^{a\,\Delta m}) + \mathcal{O}\left(\mathrm{e}^{-2\,\Delta m\,\tau}\right)\,. \quad (4.23)$$

If the expected error of $M$ is of $\mathcal{O}(\delta M)$, then the contribution from higher states is smaller for

$$\tau \geq t_p = \frac{1}{\Delta m}\,\ln\left[\frac{1-\mathrm{e}^{-a\,\Delta m}}{a\,\delta M}\,\frac{a_2}{a_1}\right]\,. \quad (4.24)$$

The improvement coming from the plateau average will not be immense, because the data (one is averaging) is strongly correlated. This is because successive points in the plateau have one measurement of the time slice correlation function in common. So, one has to care about autocorrelation and use techniques like jackknife [25] and/or directly analyze the autocorrelation [26]. In practice, in order to evaluate the parameters, the measured time slice correlation function is fitted to the form eq. (4.22) using jackknife. Then the start of the plateau is determined through eq. (4.24). Finally the plateau average of the mass gap eq. (4.21) and its error are computed using the method proposed in [26].

In the case of the data presented in Fig. 4.2 and Fig. 4.3 the plateau average was started at $t_p = 15$ and the error could be reduced by 25% compared to an arbitrary point in the plateau.

In practice it turned out that the value of $t_p/L$ depends on $L$. So, to be safe, $D(\tau)$ is always measured for $\tau/a = 1, 2, \ldots, L/a$ and $M(L)$ is determined like discussed.

### 4.1.3 Step Scaling Function and Continuum Limit

Now the computation of the step scaling function is considered. As described in the beginning of Section 4.1, to determine $\Sigma_z(2, u, N, a/L)$ for a certain value $z = u_0$ one has to tune the bare coupling $\beta$ till one measures $M(L)\,L = u_0$ on the small lattice. Since MC simulation is a statistical process, one never gets to this point. One rather stops when the measured $u = M(L)\,L$ is in the vicinity of $u_0$. With the so fixed $\beta$ one then measures the mass gap $M(2L)\,2L$ on the large lattice. Finally PT is used to extrapolate the measured $\Sigma_z(2, u, N, a/L)$ to the wanted $\Sigma_z(2, u_0, N, a/L)$

$$\Sigma_z(2, u_0, N, a/L) = \Sigma_z(2, u, N, a/L) + \left.\frac{\partial \Sigma_z}{\partial u}\right|_u (u_0 - u) + \mathcal{O}\left((u_0 - u)^2\right) . \tag{4.25}$$

For the gradient $\partial\Sigma_z/\partial u$ it is sufficient to compute the perturbative expansion eq. (3.12) up to two-loop order. The third order term can be used to control the systematical error, i.e. to assure that it is smaller than the statistical one. In this way the error of $u = M(L)\,L$ propagates into that of $\Sigma_z(2, u_0, N, a/L)$

$$\delta\Sigma_z^0 = \sqrt{(\delta\Sigma_z)^2 + \left(-\left.\frac{\partial \Sigma_z}{\partial u}\right|_u\right)^2 (\delta u)^2}\,. \tag{4.26}$$

The explained procedure is repeated for different sizes $L$ of the small lattice giving the step scaling function for a range of lattice spacings $a/L$.

## 4.2 Improved Estimator for 2d $O(N)$ Invariant Vector Models

First the basic ideas behind improved estimators will be sketched. Then the case of two dimensional $O(N)$ invariant vector models (Section 4.2.2), especially the estimator of time-slice correlation functions in these models (Section 4.2.3), will be considered. Constructing an improved estimator for these observables will lead to an integral over the rotation group $SO(N)$. In Section 4.2.4 the properties of this integral will be explored and a few hints towards implementation will be given. of an algorithm solving the problem by MC integration. The statements in Sections 4.2.1 - 4.2.3 follow more or less the original paper (Chapter 2 in [2]) and Section 4.2.4 is inspired by an existing code of Martin Hasenbusch. The discussion ends with numerical tests of the estimator and compare it to the standard estimator.

### 4.2.1 General Remarks

Consider a scalar quantum field theory with action $S[\phi]$. The fields $\phi_i(x)$ label the degrees of freedom and take values from minus to plus infinity. The expectation value of an observable $A$ is then

$$\langle A\rangle = \frac{\int \mathrm{D}\phi \exp(-S[\phi])A(\phi)}{\int \mathrm{D}\phi \exp(-S[\phi])}\,, \tag{4.27}$$

where $\phi$ denotes a field configuration. Now split $\phi$ into two parts $\phi = (\chi, \psi)$ such that the integral over $\psi$ can be done exactly for fixed $\chi$. This simply means explicitly integrating out some degrees of freedom. The improved estimator is then defined as

$$A_{\mathrm{imp}}(\chi) = \langle A\rangle_{\mathrm{cond}} = \frac{\int \mathrm{D}\psi_1 \exp(-S(\chi,\psi_1))A(\chi,\psi_1)}{\int \mathrm{D}\psi_2 \exp(-S[\chi,\psi_2])}\,. \tag{4.28}$$

The denominator is called *conditional* partition function because it still depends on some degrees of freedom, namely $\chi$. The expectation value of this quantity is then

$$\begin{aligned}
\langle A_{\mathrm{imp}}\rangle &= \frac{1}{Z}\int \mathrm{D}\chi\mathrm{D}\psi\; A_{\mathrm{imp}}(\chi)\; \mathrm{e}^{-S[\chi,\psi]} \\
&= \frac{1}{Z}\int \mathrm{D}\chi\mathrm{D}\psi\; \mathrm{e}^{-S[\chi,\psi]}\frac{\int \mathrm{D}\psi_1 A(\chi,\psi_1)\mathrm{e}^{-S[\chi,\psi_1]}}{\int \mathrm{D}\psi_2 \mathrm{e}^{-S[\chi,\psi_2]}} \\
&= \langle A\rangle\,,
\end{aligned}$$

where in the last step the term in front of the fraction has been canceled by the denominator. Proceeding with the variance one yields

$$\begin{aligned}\langle A_{\text{imp}}^2\rangle &= \frac{1}{Z}\int \mathrm{D}\chi\mathrm{D}\psi\ \mathrm{e}^{-S}\frac{\left(\int \mathrm{D}\psi_1 A\mathrm{e}^{-S}\right)^2}{\left(\int \mathrm{D}\psi_2\mathrm{e}^{-S}\right)^2}\\ &\leq \frac{1}{Z}\int \mathrm{D}\chi\mathrm{D}\psi\ \mathrm{e}^{-S}\frac{\left(\int \mathrm{D}\psi_1 A^2\mathrm{e}^{-S}\right)\cdot\left(\int \mathrm{D}\psi_3\mathrm{e}^{-S}\right)}{\left(\int \mathrm{D}\psi_2\mathrm{e}^{-S}\right)^2}\\ &= \langle A^2\rangle\,,\end{aligned}$$

where the Schwarz inequality

$$\left|\int f(x)g(x)\mathrm{d}x\right|^2 \leq \left(\int |f|^2\mathrm{d}x\right)\left(\int |g|^2\mathrm{d}x\right)$$

was used. Hence the variance of the improved estimator $A_{\text{imp}}$ is reduced compared to that of $A$

$$\langle A^2\rangle - \langle A\rangle^2 \geq \langle A_{\text{imp}}^2\rangle - \langle A_{\text{imp}}\rangle^2\,. \tag{4.29}$$

### 4.2.2 $O(N)$ Vector Models

The action for an $O(N)$ invariant vector model on a 2d lattice with time extension $T$, spatial extension $L$ and lattice spacing $a = 1$ reads (A.28)

$$S = -\beta\sum_{x=1}^{L}\sum_{t=1}^{T}(s(x,t)\ s(x+1,t) + s(x,t)\ s(x,t+1))\,, \tag{4.30}$$

where $s(x,t) = (s_1,\ldots,s_N)$ denotes the $O(N)$ vector attached to the lattice site $(x,t)$. Assuming periodic boundary conditions in space and free b.c. in time direction implies $s(x,T+1) = 0\ \forall x$ and $s(L+1,t) = s(1,t)\ \forall t$ in eq. (4.30). Think of the lattice as being built up of time-slices, i.e. embedded 1d models. For every given configuration of such a time-slice at $t$ multiplication of all spins in this slice with the same $O(N)$ matrix $R(t)$ yields a new valid configuration. By integrating over all possible $O(N)$ rotations one gains the desired improved estimator. To show this one starts with the partition function of action (4.30)

$$Z = \int \mathrm{D}s\ \mathrm{e}^{-S[s]} = \int\prod_{x,t}\mathrm{d}s(x,t)\ \mathrm{e}^{-S[s]}\,. \tag{4.31}$$

A transformation of variables

$$s(x,t)\ \rightarrow\ s'(x,t) = R(t)^{-1}s(x,t) \tag{4.32}$$

leaves $Z$ invariant because the measure is invariant ($\det(R(t)) = 1$)

$$\mathrm{d}(R(t)^{-1}s(x,t)) \equiv \mathrm{d}s'(x,t) = \mathrm{d}s(x,t)$$

$$Z = \int \prod_{x,t} \mathrm{d}(R(t)^{-1}s(x,t))\ \mathrm{e}^{-S[R^{-1}s]} = \int \prod_{x,t} \mathrm{d}s(x,t)\ \mathrm{e}^{-S[R^{-1}s]}\,. \tag{4.33}$$

Hence $Z$ does not depend on $R(t)$ and one could integrate over the group introducing the normalized Haar measure $\mathrm{d}R(t)$ of $O(N)$

$$Z = \int_{O(N)} \prod_t \mathrm{d}R(t) \int \prod_{x,t} \mathrm{d}s(x,t)\ \mathrm{e}^{-S[R^{-1}s]}\,. \tag{4.34}$$

The order of the integrations can be interchanged and the group integral will give the conditional partition function. But first consider the transformed action. Since the rotations do not affect the spatial bonds the action can be split up into a $R(t)$-dependent and a $R(t)$-independent part

$$\begin{aligned} S &= -\beta \sum_{x,t} \Big\{ R(t)^{-1}s(x,t) \cdot R(t)^{-1}s(x+1,t) \\ &\quad + R(t)^{-1}s(x,t) \cdot R(t+1)^{-1}s(x,t+1) \Big\} \\ &= -\beta \sum_{x,t} \Big\{ s(x,t)s(x+1,t) + R(t)^{-1}s(x,t) \cdot R(t+1)^{-1}s(x,t+1) \Big\} \\ &= S_{\mathrm{spatial}} + S_{\mathrm{temp}}(R)\,. \end{aligned}$$

Therefore the spatial bonds factor out and in analogy with eq. (4.28) the conditional partition function and the improved estimator of an observable $A$ may be defined as

$$A^{O(N)}_{\mathrm{imp}}(s) = \frac{1}{Z_{\mathrm{cond}}} \int_{O(N)} \mathrm{D}R\ \mathrm{e}^{-S_{\mathrm{temp}}(R)}\ A(R) \tag{4.35}$$

$$Z_{\mathrm{cond}} = Z_{\mathrm{cond}}(s) = \int_{O(N)} \mathrm{D}R\ \mathrm{e}^{-S_{\mathrm{temp}}(R)}\,. \tag{4.36}$$

So far it could be shown that integration over all possible rotations of all vectors in a time slice of a given configuration yields an improved estimator for $O(N)$ models. But the structure of expression (4.36) has to be further investigated. To this end a componentwise description of the temporal part of the action is introduced

$$\begin{aligned} S_{\mathrm{temp}}(R) &= -\beta \sum_{x,t} \sum_{i,j,k} \Big( R_{ij}(t)s_j(x,t) \Big) \Big( R_{ik}(t+1)s_k(x,t+1) \Big) \\ &= -\beta \sum_{x,t} \sum_{i,j,k} R_{ij}(t)R_{ik}(t+1)s_j(x,t)s_k(x,t+1) \\ &= -\sum_t \sum_{j,k} X_{jk}(t,t+1)\ Q_{jk}(t,t+1) \\ &= -\sum_t \mathrm{tr}\left( X(t,t+1)\ Q^T(t,t+1) \right)\,, \end{aligned} \tag{4.37}$$

where the relative rotations $X$ and coupling matrices of adjacent time-slices $Q$ are defined by

$$X_{jk}(t,t+1) = \sum_i R_{ij}(t)R_{ik}(t+1) = \left(R^{-1}(t)R(t+1)\right)_{jk} \quad (4.38)$$

$$Q_{ij}(t,t+1) = \beta \sum_x s_i(x,t)s_j(x,t+1)\,. \quad (4.39)$$

The relative rotation $X(t,t+1)$ has the property

$$R(t) = R(1)X(1,2)\ldots X(t-2,t-1)X(t-1,t), \qquad t>1 \quad (4.40)$$

(Note: Periodic boundary conditions would imply $\prod_t X(t,t+1) = 1$).

The conditional partition function can be rewritten in terms of the relative rotations

$$\begin{aligned} Z_{\text{cond}} &= \int_{O(N)} \prod_{t=1}^{T} \mathrm{d}R(t)\ \mathrm{e}^{-S_{\text{temp}}(R(t))} \\ &= \int_{O(N)} \mathrm{d}R(1) \int_{O(N)} \prod_{t=1}^{T-1} \mathrm{d}X(t,t+1)\ \mathrm{e}^{-S_{\text{temp}}(X(t,t+1))}\,. \end{aligned}$$

In the last line properties of the Haar measure were used. The integral over $R(1)$ is one since the measure is normalized. From eq. (4.37) one easily sees that the conditional partition function factorizes in partition functions of adjacent time-slices

$$Z_{\text{cond}} = \prod_{t=1}^{T-1} z(t,t+1) \quad (4.41)$$

$$z(t,t+1) = \int \mathrm{d}X \exp\left(\mathrm{tr}\left(X\ Q^T\right)\right)\,. \quad (4.42)$$

This is the final result for an observable $A$. The concept is now applied to the special case of time-slice correlation functions. From now on the name of the group at the integral sign is omitted, whenever it is clear what is meant.

### 4.2.3 Time Slice Correlation Function

In Section 4.2.2 all the ingredients were gathered to construct an improved estimator for time-slice correlation functions defined by

$$G(\tau) = \sum_i S_i(t)S_i(t+\tau)\,, \quad (4.43)$$

where $S_i(t)$ are the space averaged time-slice vector components

$$S_i(t) = \sum_x s_i(x,t)\,. \quad (4.44)$$

Introducing the transformation of variables (4.32) one can immediately write down the improved estimator

$$\begin{aligned} G^{O(N)}_{\rm imp}(s,\tau) &= \Big\langle \sum_{i,j,k} R_{ij}(t) S_j(t) R_{ik}(t+\tau) S_k(t+\tau) \Big\rangle_{\rm cond} \\ &= \sum_{j,k} \Big\langle S_j(t) X_{jk}(t,t+\tau) S_k(t+\tau) \Big\rangle_{\rm cond} \\ &= \sum_{j,k} S_j(t) \Big\langle X_{jk}(t,t+\tau) \Big\rangle_{\rm cond} S_k(t+\tau)\,, \end{aligned} \tag{4.45}$$

where

$$\langle A \rangle_{\rm cond} = \frac{1}{Z_{\rm cond}} \int_{O(N)} {\rm D}X {\rm e}^{-S_{\rm temp}(X)}\ A(X) \tag{4.46}$$

and

$$\begin{aligned} X(t,t+\tau) &= R^{-1}(t) R(t+\tau) \\ &= \Big( R(1) X(1,2) \ldots X(t-1,t) \Big)^{-1} \\ &\quad \times R(1) X(1,2) \ldots X(t+\tau-1,t+\tau) \\ &= X(t,t+1) \ldots X(t+\tau-1,t+\tau)\,. \end{aligned} \tag{4.47}$$

The conditional expectation value in eq. (4.45) can be further evaluated using properties (4.41), (4.42) and (4.47)

$$\begin{aligned} \langle X(t,t+\tau) \rangle_{\rm cond} &= \frac{1}{Z_{\rm cond}} \int {\rm D}X\ {\rm e}^{-S_{\rm temp}(X)} \\ &\quad \times X(t,t+1) \ldots X(t+\tau-1,t+\tau) \\ &= \frac{1}{\prod_{t'} z(t',t'+1)} \int \prod_{t=1}^{T-1} {\rm d}X(t,t+1)\ {\rm e}^{{\rm tr}\left(XQ^T\right)} \\ &\quad \times X(t,t+1) \ldots X(t+\tau-1,t+\tau) \\ &= \langle X(t,t+1) \rangle^*_{\rm cond} \ldots \langle X(t,t+\tau) \rangle^*_{\rm cond}\,. \end{aligned}$$

So one is left with one integral to be computed

$$\langle X(t,t+1) \rangle^*_{\rm cond} = \frac{\int {\rm d}X \exp({\rm tr}\left(XQ^T\right))\ X}{\int {\rm d}X \exp({\rm tr}\left(XQ^T\right))}\,. \tag{4.48}$$

### 4.2.4 Monte Carlo Algorithm

For $N = 2$ the integral (4.48) can be solved analytically [2]. For $N = 3$ [2] suggests a parameterization of $SO(3)$ in terms of the Euler angles. Then one of three integrals can be performed analytically whereas the other two have to be solved numerically.

Since it seems impossible to solve the integral analytically for arbitrary $N$, a numerical access to (4.48) valid for $N \geq 2$ is needed. First the integration is also restricted to the $SO(N)$ subgroup of $O(N)$

$$\langle X \rangle = \frac{1}{Z} \int_{SO(N)} \mathrm{d}X \; \mathrm{e}^{\mathrm{tr}(XQ^T)} \; X \,. \tag{4.49}$$

Recall the Singular Value Decomposition of a matrix

$$A = U_A W_A V_A \,, \tag{4.50}$$

where $U$ and $V$ are orthogonal and $W$ is diagonal and positive. This decomposition always exists [27]. The measure of $SO(N)$ is invariant under left-right multiplications

$$\int \mathrm{d}X f(X) = \int \mathrm{d}X f(UXV) \,. \tag{4.51}$$

Using this property and the SVD of $Q^T = V^T \widetilde{Q} U^T$ the divisor in eq. (4.48) can be transformed like

$$\int \mathrm{d}X \; \mathrm{e}^{\mathrm{tr}(UXV\,Q^T)} = \int \mathrm{d}X \; \mathrm{e}^{\mathrm{tr}(X\,\widetilde{Q})} \,. \tag{4.52}$$

Since $\widetilde{Q}$ is diagonal $\widetilde{Q}_{ij} = \widetilde{Q}_i \delta_{ij}$ only diagonal elements contribute to the sum in the exponent

$$Z = \int_{SO(N)} \mathrm{d}X \exp\big(\sum_i X_{ii} \widetilde{Q}_i\big) \,. \tag{4.53}$$

With the substitutions

$$Q^T = V^T \widetilde{Q} U^T \quad \text{and} \quad Y = U^T X V^T$$

the expectation value eq. (4.48) finally reads

$$\langle X \rangle = U \langle Y \rangle V = U \left( \frac{1}{Z} \int_{SO(N)} \mathrm{d}Y \exp\big(\sum_i Y_{ii} \widetilde{Q}_i\big) \; Y \right) V \,. \tag{4.54}$$

It is the structure of the average $\langle Y \rangle$ what will be investigated in the following.

**Structure of $\langle Y \rangle$**

Since the Haar measure has the property

$$\mathrm{d}(Y^{-1}) = \mathrm{d}Y \tag{4.55}$$

and the trace is invariant under transposition one easily sees that $\langle Y \rangle$ has to be symmetric. The aim is to show that furthermore all off-diagonal elements vanish.

Consider first the case where $\widetilde{Q}_i = q, \ \forall \ i$

$$\langle Y \rangle = \frac{1}{Z} \int_{SO(N)} \mathrm{d}Y \ \mathrm{e}^{\,q \,\mathrm{tr}(Y)} \ Y \,. \tag{4.56}$$

Then $\langle Y \rangle$ is invariant under left-right-multiplication with an arbitrary element $O$ of $SO(N)$

$$\begin{aligned} O\langle Y \rangle O^{-1} &= \frac{1}{Z} \int_{SO(N)} \mathrm{d}Y \ \mathrm{e}^{\,q \,\mathrm{tr}(Y)} \ OYO^{-1} \\ Y \to Y' = OYO^{-1} \quad \to &= \frac{1}{Z} \int_{SO(N)} \mathrm{d}Y' \ \mathrm{e}^{\,q \,\mathrm{tr}(Y')} \ Y' \\ &= \langle Y \rangle \,, \end{aligned}$$

where the cyclic properties of the trace were used. Therefore $\langle Y \rangle$ must be a multiple of the unity (diagonal with all diagonal elements equal).

Now, assume only two elements of $\widetilde{Q}$ coincide: $\widetilde{Q}_a = \widetilde{Q}_b = q$ $(a \neq b)$. In order to be able to use again the cyclic properties of the trace, the rotation $O$ has to be restricted to a one parameter subgroup of $SO(N)$, i.e. an $SO(2)$ rotation affecting only the rows and columns $a$ and $b$ (see eq. (4.59) as an example of $a = 1$, $b = 2$). The so constructed matrix $O$ commutes with $\widetilde{Q}$ and due to the same argument as above the off-diagonal elements in the rows and columns $a$ and $b$ have to vanish and the diagonal elements $\langle Y \rangle_{aa} = \langle Y \rangle_{bb}$ are equal.

One might tend to say not much is gained from the arguments above for the condition $\widetilde{Q}_a = \widetilde{Q}_b$ is rather strict. But at the end of the following calculation its consequences will become very useful.

The general case of arbitrary $\widetilde{Q}$ is tackled by introducing again a rotation $R(\theta)$ of the subgroup $SO(2)$ using the invariance of the measure

$$\langle Y \rangle = \frac{1}{Z} \int_{SO(N)} \mathrm{d}Y \ \mathrm{e}^{\mathrm{tr}\left(YR(\theta)\widetilde{Q}\right)} \ YR(\theta) \,. \tag{4.57}$$

Since $\langle Y \rangle$ does not depend on $\theta$, integration over it gives just a factor of $2\pi$ (this is just the explicit version of the Haar measure for SO(2))

$$\begin{aligned} \langle Y \rangle &= \frac{1}{2\pi Z} \int_0^{2\pi} \mathrm{d}\theta \int_{SO(N)} \mathrm{d}Y \ \mathrm{e}^{\mathrm{tr}\left(YR(\theta)\widetilde{Q}\right)} \ YR(\theta) \\ &= \frac{1}{2\pi Z} \int_{SO(N)} \mathrm{d}Y \int_0^{2\pi} \mathrm{d}\theta \ \mathrm{e}^{\mathrm{tr}\left(YR(\theta)\widetilde{Q}\right)} \ YR(\theta) \,, \end{aligned} \tag{4.58}$$

where in the second line the order of the integrations was interchanged. Without loss of generality, suppose $R(\theta)$ affects columns 1 and 2

$$R(\theta) = \begin{pmatrix} \cos(\theta) & \sin(\theta) & 0 & \dots \\ -\sin(\theta) & \cos(\theta) & 0 & \dots \\ 0 & 0 & 1 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \tag{4.59}$$

$$YR(\theta) = \begin{pmatrix} y_{11}\cos(\theta) - y_{12}\sin(\theta) & y_{11}\sin(\theta) + y_{12}\cos(\theta) & y_{13} & \dots \\ y_{21}\cos(\theta) - y_{22}\sin(\theta) & y_{21}\sin(\theta) + y_{22}\cos(\theta) & y_{23} & \dots \\ y_{31}\cos(\theta) - y_{32}\sin(\theta) & y_{31}\sin(\theta) + y_{32}\cos(\theta) & y_{33} & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

Grouping all constant (i.e. independent of $\theta$) factors in the weight in eq. (4.58) gives the following expression for the first off-diagonal element in column 1

$$\langle Y\rangle_{21} = \frac{1}{Z}\int_{SO(N)} \mathrm{d}Y \exp\left(\sum_{i>2}(y_{ii}\widetilde{Q}_i)\right) I_{21} \tag{4.60}$$

$$\begin{aligned} I_{21} &= \int_0^{2\pi} \mathrm{d}\theta\, \mathrm{e}^{\widetilde{Q}_1(y_{11}\cos(\theta) - y_{12}\sin(\theta))}\mathrm{e}^{\widetilde{Q}_2(y_{21}\sin(\theta)+y_{22}\cos(\theta))} \\ &\quad \times \{y_{21}\cos(\theta) - y_{22}\sin(\theta)\} \\ &= \int_0^{2\pi} \mathrm{d}\theta\, \mathrm{e}^{\widetilde{Q}_1(y_{11}\cos(\theta) - y_{12}\sin(\theta))}\frac{1}{\widetilde{Q}_2}\frac{\partial}{\partial\theta}\left(\mathrm{e}^{\widetilde{Q}_2(y_{21}\sin(\theta)+y_{22}\cos(\theta))}\right) \\ &= -\int_0^{2\pi} \mathrm{d}\theta\, \frac{1}{\widetilde{Q}_2}\frac{\partial}{\partial\theta}\left(\mathrm{e}^{\widetilde{Q}_1(y_{11}\cos(\theta) - y_{12}\sin(\theta))}\right)\mathrm{e}^{\widetilde{Q}_2(y_{21}\sin(\theta)+y_{22}\cos(\theta))} \\ &= \frac{\widetilde{Q}_1}{\widetilde{Q}_2}\int_0^{2\pi} \mathrm{d}\theta\, \mathrm{e}^{\widetilde{Q}_1(y_{11}\cos(\theta) - y_{12}\sin(\theta))}\mathrm{e}^{\widetilde{Q}_2(y_{21}\sin(\theta)+y_{22}\cos(\theta))} \\ &\quad \times \{y_{11}\sin(\theta) + y_{12}\cos(\theta)\} \\ &= \frac{\widetilde{Q}_1}{\widetilde{Q}_2} I_{12}\,. \end{aligned}$$

where in the second step a partial integration was performed. This result means

$$\langle Y\rangle_{21} = \frac{\widetilde{Q}_1}{\widetilde{Q}_2}\langle Y\rangle_{21} \tag{4.61}$$

and thus, for $\widetilde{Q}_1 \neq \widetilde{Q}_2$, that these off-diagonal elements must vanish since it was shown earlier that $\langle Y\rangle$ has to be symmetric (eq. (4.55)). But for $\widetilde{Q}_1 = \widetilde{Q}_2$ it was also already shown that these matrix elements must vanish. Since the matrix $R(\theta)$ could have affected any two columns or rows, this argument is valid for all off-diagonal elements. So $\langle Y\rangle$ is truly a diagonal matrix and this can be used when solving eq. (4.54) by MC integration.

### MC Algorithm for $\langle Y \rangle$

The key point in a MC integration of an expectation value like eq. (4.54) is how to derive a sample of "configurations" (here $SO(N)$ matrices) distributed like the weight in the integrand. This procedure has to obey two demands: ergodicity and detailed balance. Ergodicity means the algorithm must assure that every element of $SO(N)$ is reachable from every other element in a finite number of steps.

*Ergodicity* A natural way to get a new group element $Y'$ from an old group element $Y$ is to multiply it by an arbitrary element $R$

$$Y \to Y' = YR\,.$$

By general arguments $R$ can be represented as a product of $n = N(N-1)/2$ elements of $SO(2)$-subgroups of $SO(N)$

$$R = R_1 R_2 \dots R_n\,.$$

So the whole transition $Y \to Y'$ can be split up into $n$ $SO(2)$-rotations with angles $\theta_i \in [-\pi, \pi]$, and it is mutually assured that the whole group is covered.

*Detailed balance* Assume a sequence of Metropolis steps. For this kind of algorithm, detailed balance can be shown [28]. A $SO(N)$-rotation is achieved by $n$ Metropolis steps of independent $SO(2)$-rotations. To be concrete, consider again rotation eq. (4.59)

$$Y \to Y' = YR(\theta)\,.$$

The acceptance probability for such a transition would read [28]

$$A(Y', Y) = \min\left\{1, \exp\left(\sum_i (Y'_{ii} - Y_{ii})\widetilde{Q}_i\right)\right\}\,, \tag{4.62}$$

where only two terms in the sum survive, say $i = a, b$ (in the case of rotation eq. (4.59) it would be $i = 1, 2$). From this expression one derives a first guess for a reasonable range of the angle $\theta$ by taking the values where $A(Y', Y)$ has decreased to $1/\mathrm{e}$ as limits, thus solving

$$(y'_{aa} - y_{aa})\widetilde{Q}_a + (y'_{bb} - y_{bb})\widetilde{Q}_b = -1\,. \tag{4.63}$$

Since $YY^T = 1$ the mean amplitude of $Y$'s diagonal elements should be $1/\sqrt{N}$ (neglecting the off-diagonal elements). For this simplification one gets

$$\sin\theta = \pm\sqrt{\frac{2\sqrt{N} - \frac{N}{\widetilde{Q}_a + \widetilde{Q}_b}}{\widetilde{Q}_1 + \widetilde{Q}_2}}\,. \tag{4.64}$$

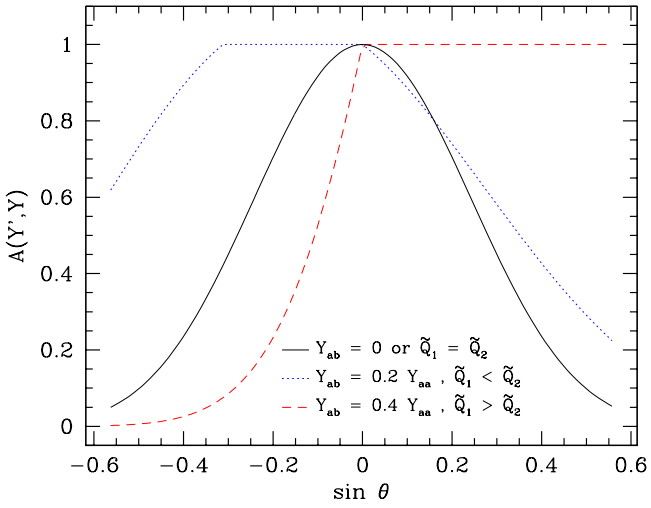



FIGURE 4.4: The acceptance of the Metropolis algorithm proposed in the text. The three curves belong to different ratios between diagonal and off-diagonal elements (see text).

In Fig. 4.4 the acceptance eq. (4.62) is plotted in this range. The three curves belong to different ratios between diagonal and off-diagonal elements. Vanishing off-diagonal elements or $\widetilde{Q}_a = \widetilde{Q}_b$ lead to a symmetric acceptance in the interval (solid line). Non-vanishing off-diagonal elements shift the maximum of the acceptance towards positive or negative rotations depending on the ratio $\widetilde{Q}_a/\widetilde{Q}_b$ (dotted and dashed lines). In the case of the plotted curves the integrated acceptance is about 75%. For $\widetilde{Q}_a \ll \widetilde{Q}_b$ (or $\widetilde{Q}_a \gg \widetilde{Q}_b$) the acceptance becomes a step function in the vicinity of $\sin\theta = 0$ and thus the smallest integrated acceptance is 50%. The same is true when the off-diagonal elements are much larger than the diagonal elements.

*Equilibration* By taking $X_0 = 1$ as starting value in the MC algorithm one does not have to equilibrate. This is because $X = 1$ could be obtained from an equilibrated configuration. Just keeping $X = 1$ would give the standard estimator. In terms of $Y$ this means (see eq. (4.54))

$$Y_0 = U^T V^T \,.$$

Note that if $Y_0 = 1$ one would have to equilibrate first.

*Over-relaxation* Since $\langle Y \rangle$ has to be symmetric, i.e $\langle Y \rangle^T = \langle Y \rangle$, an over-relaxation step would be replacing $Y$ by its transpose. But one does not have to do this in reality. Remembering that $\langle Y \rangle$ is a diagonal matrix one can simply throw away all off-diagonal elements after the Metropolis updates.

Nevertheless the transpose of $Y$ can be used for a very simple and effective estimator. Simply taking

$$\langle Y \rangle = \frac{1}{2}(Y_0 + Y_0^T) = \frac{1}{2}(U^T V^T + VU)$$

gives a reasonable improvement (see below). In terms of $\langle X \rangle$ this means

$$\langle X \rangle = \frac{1}{2}(1 + (UV)^2)\,.$$

Thus no Metropolis steps are needed, two matrix multiplications will do.

Whether the proposed improved estimators are really able to reduce the variance of observables, it has to be verified by implementing them in MC simulations. Furthermore, for the estimator including a small MC integration, it may be called *Metropolis improved estimator*, the dependence on the number of updates has to be investigated. For the algorithm proposed in the last paragraph, it may be called *simple improved estimator*, it is interesting to see how much improvement it already brings.

### 4.2.5 Testing the Estimator

#### Bias and Large Distance Behavior

First the primary observable, the time slice correlation function, measured with the standard is compared to measurements with the improved estimators. For the number of updates in the Metropolis improved estimator the ad hoc value of $M = 10$ is used. In Fig. 4.5 the result of simulations in the $O(3)$ model at $\beta = 1.6982$ on a $20 \times 100$ lattice is presented (Table C.1). The joint plot of the correlation functions from the standard and the improved estimators shows no bias. This supports the correctness of the new estimators.

Before turning to the estimator's costs, attention should be drawn to the behavior of the error when increasing the time separation of the time slices. For the standard estimator the error of the mass grows already on the scale where it will be measured, i.e. at $t \leq L/2$ where the higher states are expected to be exponentially suppressed. In Fig. 4.6 and Fig. 4.7 the relative error of the time slice correlation function and the mass is plotted. For all three estimators and for both quantities the relative error grows exponentially. But in the case of the improved estimators the increase happens on a much smaller scale. For the finite volume mass it is even almost constant for $t \in [0, L]$ (Fig. 4.7). This feature of the new estimators is used in the computation of the step scaling function (Section 4.1). It will allow to average the mass over a plateau region in order to reduce the final error.

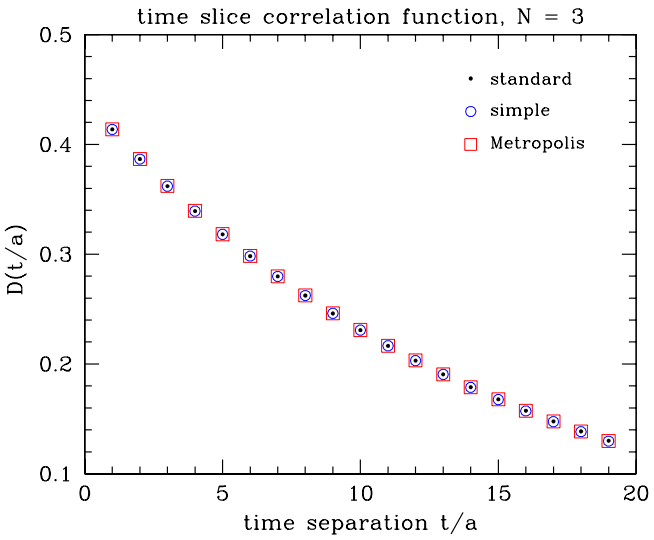



FIGURE 4.5: The improved estimators seem unbiased. The errors are too small to be plotted on this scale.

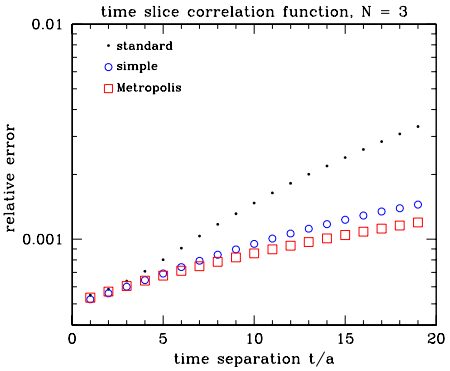



FIGURE 4.6: The relative error of time slice correlation function grows exponentially with the time separation of the slices. The scale of this behavior is much smaller in the case of the simple and the Metropolis improved estimators.

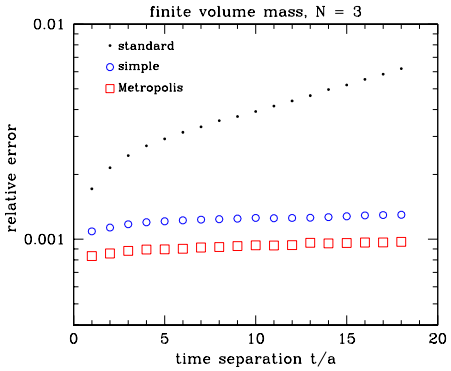



FIGURE 4.7: The relative error of the mass grows exponentially as well. For the improved estimators this happens to be a very small effect.

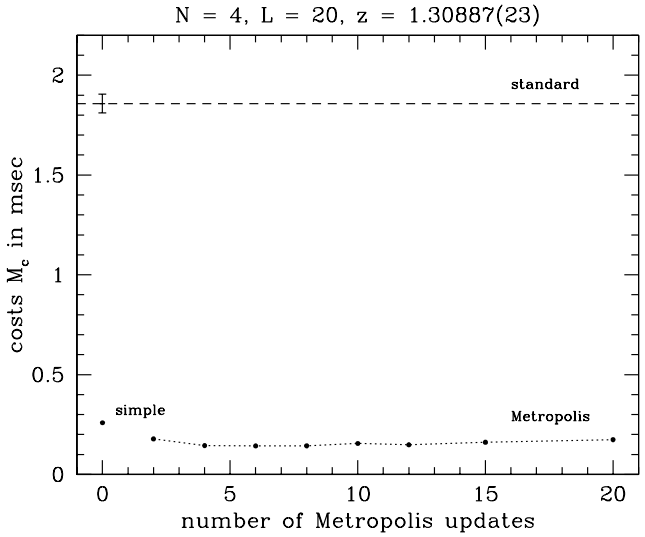



FIGURE 4.8: The Metropolis improved estimator reduces the costs by a factor of 12 and shows almost no dependence on the number of updates. The simple improved estimator already reduces the costs by a factor of 7.

### Numerical Costs and Optimal Number of Metropolis Updates

Measurements in the $O(4)$ model for $\beta = 2.4765$ on a $20 \times 100$ lattice ($z = 1.30887(23)$) were done in order to determine the new estimator's efficiency and how much the variance of the observable is reduced. Runs with $N_m = 10000$ measurements of the time slice correlation function were done: with the standard estimator, the Metropolis improved estimator including $M = 2, 4, \ldots, 20]$ updates and the simple improved estimator (Table C.2). These runs were repeated 10 times to get a reasonable statistic. The quantity determined for each estimator are the numerical costs

$$M_c = T_{run} \cdot \left( \frac{\Delta x}{x} \right)^2 , \tag{4.65}$$

the product of the runtime and the squared relative error. The runtime consists of two terms

$$T_{run} = N_m \cdot \left( f_{up} \cdot \left[ \frac{V}{\langle C \rangle} \right] \cdot t_{up} + t_m \right) ,$$

where $t_{up}$ and $t_m$ are the times for one single cluster update and for one measurement respectively. The expression $[x]$ means greatest integer $\leq x$ and $V/\langle C \rangle$ is the ratio of lattice volume to mean cluster size in the single cluster algorithm (Section B.3) used for updating the system. In Fig. 4.8 the costs are plotted vs. the number of Metropolis updates. The Metropolis improved estimator reduces the costs by a factor of 12 and shows almost

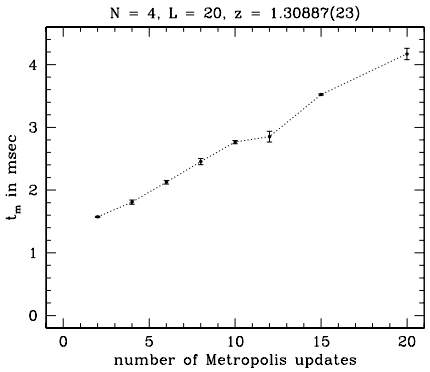



FIGURE 4.9: The time spent in the improved routine measuring the time slice correlation function $t_m$ grows linear with the number of updates.

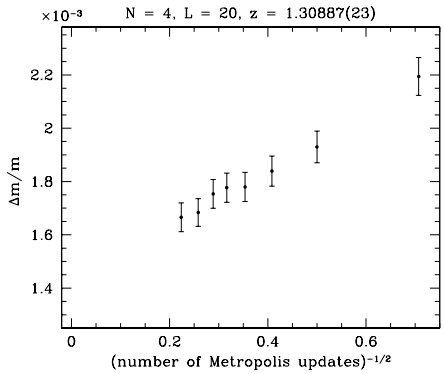



FIGURE 4.10: The relative error of the mass, calculated from the time slice correlation function is proportional to $\sqrt{1/M}$ ($M$: number of Metropolis updates).

no dependence on the number of updates. The simple improved estimator already reduces the costs by a factor of 7. This is a great enhancement of efficiency in both cases. Since the Metropolis improved estimator cuts down the costs by a factor of 2 compared to the simple one, it should be preferred.

Why is the dependence of the Metropolis improved estimator on the number of updates $M$ so weak? Now, two quantities enter $M_c$, which depend on $M$: the time spent in the measuring routine $t_m$ and the relative error of the observable. The first one should grow linear with $M$ and that is observed in Fig. 4.9. The relative error of the derived quantities depend on the variance of the primary observables. This variance should be reduced due to the improved estimator eq. (4.29). For the largest possible reduction one has to solve the integral eq. (4.54) exactly. In an MC integration the error is inversely proportional to the square root of the number of samples. This means the relative error in eq. (4.65) is proportional to $\sqrt{1/M}$ and exactly that is observed in the data displayed in Fig. 4.10. Therefore the two dependencies cancel and $M_c$ is expected to be independent of $M$; at least for $M \geq 10$. This is because one needs the central limit theorem to derive the stated behavior of the error in MC integration. In Fig. 4.8 one observes that for $M > 10$ the costs begin to rise. This probably due to other effects, e.g. autocorrelation. Hence for the simulations quoted in this thesis $M = 10$ is used.

# Chapter 5

# Large $N$ Expansion

## 5.1 $1/N$ Expansion of the $O(N)$ Model

More variables, thus greater complexity - an obvious statement one might think. But there exist families of field theories with symmetry group $SO(N)$ that become simpler as $N$ becomes lager [13],[29].

The $1/N$ expansion of the model described in Section 2.1.4 will be studied. The discussion starts in the continuum theory and a brief overview is given of the steps included. The technical details are explained in the context of the lattice regularization of the theory (Section 5.1.2).

### 5.1.1 $1/N$ Expansion in the Continuum

In the path integral description of the model the $1/N$ expandability becomes evident when solving the constraint (2.15) by a Lagrange multiplier field $\alpha(x)$. This is done by substituting in the naive first version of the path integral

$$\int \mathrm{D}s \prod_x \delta(s^2(x)-1) \, \exp\left[-\frac{1}{2f}\int \mathrm{d}^2x \, (\partial_\mu s(x))^2\right] , \tag{5.1}$$

the product of delta functions by an integral representation

$$\int \mathrm{D}s\mathrm{D}\alpha \, \exp\left[-\frac{1}{2f}\int \mathrm{d}^2x \, (\partial_\mu s(x))^2 - \frac{i}{2f}\int \mathrm{d}^2x \, \alpha(x)(s^2(x)-1)\right] . \tag{5.2}$$

Since

$$0 = (\partial^2)\, s^2 = \partial_\mu \left(\partial_\mu s \cdot s + s \cdot \partial_\mu s\right) = 2\,\partial_\mu(\partial_\mu s)\cdot s + 2\,\partial_\mu s \cdot \partial_\mu s \tag{5.3}$$

which leads to

$$(\partial_\mu s)^2 = -s \cdot \partial^2 s \,, \tag{5.4}$$

the $s$ integration is Gaussian and can easily be performed. One obtains the effective action for the $\alpha$ field

$$S_{\text{eff}} = \frac{N}{2}\left\{\mathrm{Tr}\ln(-\partial^2 + i\alpha(x)) - \frac{i}{Nf}\int \mathrm{d}^2x \, \alpha(x)\right\} . \tag{5.5}$$

In the limit $N \to \infty$ and simultaneously keeping the product $Nf$ fixed, the path integral is dominated by the function $\alpha(x)$ that minimizes the exponent and thus $S_{\text{eff}}$

$$\frac{\delta}{\delta\alpha(x)} S_{\text{eff}} = \frac{iN}{2}\left\{\langle x|(-\partial^2 + i\alpha(x))^{-1}|x\rangle - \frac{1}{Nf}\right\} = 0 \,. \tag{5.6}$$

The matrix element must be constant and real. Thus, one should look for a solution in which $\alpha(x)$ is translation invariant and purely imaginary

$$\alpha(x) = -im_0^2 \,, \tag{5.7}$$

where $m_0$ is defined by the so called gap equation (momentum space version of (5.6))

$$\frac{1}{Nf} = \int \frac{\mathrm{d}^2k}{(2\pi)^2} \frac{1}{k^2 + m_0^2} \,. \tag{5.8}$$

A $1/N$ expansion then means taking into account the fluctuations around this saddle point.

### 5.1.2 $1/N$ Expansion on the Lattice

In order to introduce a lattice regularization of the path integral (5.1) one has to choose a lattice counterpart of the action (2.17). The simplest version includes only nearest-neighbor interaction and no terms higher than quadratic in the fields (standard n.n. action). The matrix notation of (A.24) is already in a suitable form in order to perform the integration over the $s$ fields. Setting $a = 1$ it reads

$$S = \frac{1}{2f} \sum_{x,y} s(x) \, \Box_{xy} \, s(y) \,. \tag{5.9}$$

A well defined path integral can be written down

$$Z = \frac{1}{Z_0} \int \prod_{x'} \mathrm{d}^N s(x') \; \delta(s^2(x') - 1) \; \exp\left[ -\frac{1}{2f} \sum_{x,y} s(x) \, \Box_{xy} \, s(y) \right] . \tag{5.10}$$

The factor $1/Z_0$ is determined by the normalization condition. Again the constraint (2.15) is solved using an auxiliary field $\alpha(x)$ and the integral representation of the delta function

$$\delta(s^2(x) - 1) = \frac{1}{2f} \int_{-i\infty}^{+i\infty} \frac{\mathrm{d}\alpha(x)}{2\pi i} \; \exp\left[ -\frac{1}{2f} \alpha(x)(s^2(x) - 1) \right] . \tag{5.11}$$

Substituting this into (5.10) and grouping terms including products of the fields $s$ together the exponent becomes

$$-\frac{1}{2f} \sum_{x,y} s(x) \, \mathcal{M}_{xy} \, s(y) + \frac{1}{2f} \sum_{x} \alpha(x) \,, \tag{5.12}$$

where the matrix

$$\mathcal{M}_{xy} = \Box_{xy} + \alpha(x) \, \delta_{x,y} \,, \tag{5.13}$$

has been introduced. Now rescaling $s_i \to \sqrt{f} s_i$ results in a version of the path integral where the integrations over the $s$ fields is Gaussian

$$\begin{aligned} Z \;\; = \;\; & \frac{1}{Z_0'} \int \prod_{x'} \mathrm{d}^N s(x') \int_{-i\infty}^{+i\infty} \frac{\mathrm{d}\alpha(x')}{2\pi i} \times \\ & \times \exp\left[ -\frac{1}{2} \sum_{x,y} s(x) \, \mathcal{M}_{xy} \, s(y) + \frac{1}{2f} \sum_{x} \alpha(x) \right] . \end{aligned} \tag{5.14}$$

Any constant factors have been absorbed into $Z_0'$. The result of such an integration is well known

$$\int \mathrm{d}^k\phi \exp\left\{(\phi, A\phi)\right\} = (2\pi)^{k/2}(\det A)^{-1/2}\,, \tag{5.15}$$

where $(\phi, A\phi) = \sum_{ij}\phi_i A_{ij}\phi_j$ and $A$ is a real, symmetric and positive matrix. Hence the $s$ integrations can easily be performed resulting in an effective action for the $\alpha$ field

$$S_{\text{eff}} = \frac{N}{2}\left\{\mathrm{Tr}\ln\mathcal{M} - \frac{1}{Nf}\sum_x \alpha(x)\right\}\,. \tag{5.16}$$

To arrive at the last equation one also needs the relation between the trace and the determinant of a matrix $A$ with eigenvalues $a_i$

$$\det A = \prod_i a_i = \exp\left[\sum_i \ln a_i\right] = \exp\left[\mathrm{Tr}\ \ln\ A\right]\,. \tag{5.17}$$

At this point the model's dependence on $N$ becomes apparent. In the limit $N \to \infty$ while keeping $\gamma = Nf$ fixed the path integral (remaining factors again absorbed into $Z_0'$)

$$Z = \frac{1}{Z_0'}\int \mathrm{D}\alpha\ \mathrm{e}^{-S_{\text{eff}}}\,, \tag{5.18}$$

$$\int \mathrm{D}\alpha = \int_{-i\infty}^{+i\infty}\prod_{x'}\frac{\mathrm{d}\alpha(x')}{2\pi i}\,, \tag{5.19}$$

can be approximated by a saddle point method. Writing the exponent as

$$-S_{\text{eff}} \equiv N\cdot\sum_x F(\alpha(x)) = N\sum_x \frac{1}{2}\left\{\frac{1}{\gamma}\alpha(x) - \ln\mathcal{M}_{xx}\right\}\,, \tag{5.20}$$

$Z$ obviously is a product of contour integrals in the complex plane. Let us concentrate on one such integral

$$I = \int_C \frac{\mathrm{d}\alpha}{2\pi i}\,\mathrm{e}^{NF(\alpha)}\,. \tag{5.21}$$

If the contour $C$ can be altered such that there is at some point $\alpha_0$ a maximum of $\mathrm{Re}\,F$ and simultaneously $\mathrm{Im}\,F$ is constant in a neighborhood of $\alpha_0$ the main contribution to the integral would come from the maximum. And the value of $I$ could be approximated by expanding $F$ around $\alpha_0$. The point $\alpha_0$ is called saddle point. In the present case this leads to the large $N$ expansion of the nonlinear $\sigma$-model. A first approximation would include just this saddle point and corrections can be obtained from fluctuations around

$\alpha_0$. A maximum of all $F(\alpha(x))$ compares to a minimum of $S_{\rm eff}$. It can be found by differentiation with respect to $\alpha(x)$

$$\left.\frac{\partial S_{\rm eff}}{\partial\alpha(x)}\right|_{\alpha(x)=\alpha_0(x)} = \frac{N}{2}\left\{(\mathcal{M}^{-1})_{xx} - \frac{1}{\gamma}\right\} = 0\,. \tag{5.22}$$

Thus the matrix element must be constant and real. Hence the configuration $\alpha_0(x) = \alpha_0$ obeying this condition is real and constant. For translation invariant $\alpha$ the matrix elements $\mathcal{M}_{xy}$ are functions of $x-y$ only and one can write down the effective action in momentum space

$$S_{\rm eff} = \frac{N}{2}\frac{1}{LT}\sum_p\left\{\ln\left[\sum_\mu 4\sin^2\frac{p_\mu}{2} + \alpha\right] - \frac{1}{\gamma}\alpha\right\}. \tag{5.23}$$

A solution to (5.22) would be

$$\alpha_0 = m_0^2\,, \tag{5.24}$$

where $m_0^2$ is implicitly defined by the gap equation

$$\frac{1}{\gamma} = \frac{1}{LT}\sum_p \frac{1}{\sum_\mu 4\sin^2\frac{p_\mu}{2} + m_0^2} \equiv \frac{1}{LT}\sum_p \frac{1}{\hat{p}^2 + m_0^2}\,. \tag{5.25}$$

The following standard notations are used

$$\hat{p}^2 \equiv \sum_\mu \hat{p}_\mu^2 \quad , \quad \hat{p}_\mu \equiv 2\sin\frac{p_\mu}{2}\,. \tag{5.26}$$

Introducing an external source $J(x) = (J_1(x), \ldots, J_N(x))$ coupled to the spin field $s$ the path integral (5.10) becomes the generating functional for correlation functions ($Z \equiv Z[J=0]$)

$$Z[J] = \int \mathrm{D}s\, \exp\left[-\frac{1}{2f}\sum_{x,y} s(x)\Box_{xy}s(y) + \sum_x J(x)\,s(x)\right], \tag{5.27}$$

where the measure is abbreviated by:

$$\mathrm{D}s = \frac{1}{Z}\prod_{x'} \mathrm{d}^N s(x')\;\delta(s^2(x') - 1)\,. \tag{5.28}$$

The $n$-point correlation functions are obtained by differentiating $Z[J]$

$$\langle s(x_1)\ldots s(x_n)\rangle = \frac{\partial}{\partial J(x_1)}\ldots\frac{\partial}{\partial J(x_n)} Z[J]\bigg|_{J=0}\,. \tag{5.29}$$

Solving the constraint in (5.27) as before one gets a similar expression as in (5.14) but with an extra term $\sum_x J(x)\, s(x)$ linear in the $s$ fields. Therefore the Gaussian integration is performed slightly different [14] and the result is

$$Z[J] = \frac{1}{Z} \int \mathrm{D}\alpha \ \exp\left[ -\frac{N}{2} \left( \frac{\gamma}{N} \mathbf{J}\, \mathcal{M}^{-1}\, \mathbf{J} + \mathrm{Tr} \ln \mathcal{M} + \frac{1}{\gamma} \sum_x \alpha(x) \right) \right] . \tag{5.30}$$

The $\gamma$ factor in front of the first term is due to the rescaling of the $s$ fields ($s_i \to \sqrt{f} s_i$). In the leading order $1/N$ expansion the two-point function is then given by [7],[8]

$$\langle s(x) s(y) \rangle = \gamma\, (\mathcal{M}^{-1})_{xy} + \mathcal{O}\,(1/N) \ , \tag{5.31}$$

and this can be rewritten using the Fourier transform of $\mathcal{M}$

$$\langle s(x) s(y) \rangle = \frac{\gamma}{LT} \sum_p \frac{\mathrm{e}^{ip\cdot(x-y)}}{\hat{p}^2 + m_0^2} + \mathcal{O}\,(1/N) \ . \tag{5.32}$$

## 5.2 Step Scaling Function in the Large $N$ Limit

The step scaling function can also be studied in the context of the $1/N$ expansion of the nonlinear sigma model. In the following sections the leading order of that expansion is derived using the lattice-regularized model in order to examine the cutoff effects. These sections follow unpublished notes by Peter Weisz [21]. All steps in these notes are verified and the derived expansions are extended in most cases.

Let us start with a finite square lattice of extension $L_1 \times L_2$. For simplicity, to ensure translation invariance, periodic boundary conditions are imposed. The two-point function in the leading order $1/N$ expansion was derived in the last section

$$\langle s(x)s(y)\rangle = \frac{1}{L_1 L_2}\sum_p \widetilde{G}(p)\,\mathrm{e}^{ip\cdot(x-y)} + \mathcal{O}\left(1/N\right)\,, \tag{5.33}$$

$$\widetilde{G}(p) = \gamma\left(\hat{p}^2 + m_0^2\right)^{-1}\,, \tag{5.34}$$

where the sum over momenta $p$ runs over the first Brillouin zone eq. (A.4) and $m_0$ is given as solution to the gap equation

$$\frac{1}{\gamma} = \frac{1}{L_1 L_2}\sum_p \frac{1}{\hat{p}^2 + m_0^2}\,. \tag{5.35}$$

The step scaling function is to be studied in a finite volume, i.e. finite space and infinite time extension. Therefore the limit $L_2 = T \to \infty$ should be taken in the equations above. Calling $L_1 = L$ and $p_1 = p_0$ eq. (5.35) then becomes

$$\frac{1}{\gamma} = \frac{1}{L}\sum_{n=0}^{L/a-1}\int_{-\pi/a}^{+\pi/a}\frac{\mathrm{d}p_0}{2\pi}\,\frac{1}{\frac{1}{a^2}(4\sin^2\frac{ap_0}{2} + 4\sin^2\frac{a\pi n}{L}) + m_0^2}\,. \tag{5.36}$$

The $p_0$ integral can be evaluated explicitly [30]

$$\int\frac{\mathrm{d}x}{b^2 + c^2\sin^2 ex} = \frac{1}{e\,b\,\sqrt{b^2+c^2}}\arctan\frac{\sqrt{b^2+c^2}\,\tan ex}{b}\,, \tag{5.37}$$

and one ends up with

$$\frac{1}{\gamma} = \frac{a}{L}\sum_{n=0}^{L/a-1}\frac{1}{2\,\epsilon(n)}\,, \tag{5.38}$$

with abbreviations

$$\epsilon(n) = \sqrt{w(n)\cdot\left(1 + \frac{w(n)}{4}\right)} \tag{5.39}$$

$$w(n) = 4\,\sin^2\frac{a\pi n}{L} + a^2 m_0^2\,. \tag{5.40}$$

Equation (5.38) is still an implicit expression defining the solution $m_0$ of the saddle point condition eq. (5.22). But in the definition of the LWW coupling eq. (3.5) the physical mass in a finite volume $M(L)$ is needed. But the physical mass $M$ is defined as the first pole on the imaginary axis of the two-point function at zero spatial momentum. So from eq. (5.34) one gains an expression relating $m_0$ and $M$:

$$\widetilde{G}(p_1 = 0, p_2 = iM) = 0\,, \tag{5.41}$$

which leads to

$$a^2 m_0^2 = 4\,\sinh^2\frac{aM}{2}\,. \tag{5.42}$$

This can be inverted

$$aM = \pm \ln\left(\frac{a^2 m_0^2}{2} + 1 \pm \sqrt{\frac{a^4 m_0^4}{4} + a^2 m_0^2}\right) \tag{5.43}$$

and expanded for small $m_0$

$$aM = m_0\left(1 - \frac{a^2 m_0^2}{24} + \mathrm{O}(a^4 m_0^4)\right)\,. \tag{5.44}$$

Finally the dimensionless variable [22]

$$z = M(L)\cdot L \tag{5.45}$$

is defined and a function $F(z, a/L)$ describing the scaling of $M(L)$ and deviation due to lattice artifacts:

$$\frac{M}{\Lambda_{\overline{\mathrm{MS}}}} = F(z, a/L)\,, \tag{5.46}$$

where $\Lambda_{\overline{\mathrm{MS}}}$ is the $\Lambda$-parameter in the minimal subtraction scheme [22]. Taking the continuum limit ($a \to 0$) of eq. (5.46) would mean sending the lattice size $L/a$ to infinity while keeping $z$ fixed.

### 5.2.1 The Small $z$ Expansion

In [22] Lüscher gives a small $z$ expansion of the r.h.s of eq. (5.46) in the continuum and for arbitrary $N$:

$$\frac{M}{\Lambda_{\overline{\mathrm{MS}}}} = K(N)\cdot z\,\mathrm{e}^{\pi/z\cdot(N-1)/(N-2)}\cdot\left(1 + \sum_{\nu=1}^{\infty} a_\nu\left(\frac{z}{2\pi}\right)^\nu\right)\,, \tag{5.47}$$

$$K(N) = (4\pi)^{-1}\,\exp\left[-\Gamma'(1)\right]\cdot\left(\pi^{-1}(N-2)/(N-1)\right)^{1/(N-2)}\,. \tag{5.48}$$

The coefficients $a_\nu$ can be computed using an expansion of $M(L)$ in the renormalized coupling constant defined in the dimensional regularization

scheme with minimal subtraction (MS). This expansion is also given in [22] (eq. (6)). Furthermore the author states an analytic expression for the case $N = \infty$, equivalent to a continuum limit of eq. (5.46)

$$\lim_{a\to 0} F(z, a/L) \equiv f(z) = \mathrm{e}^{H(z)}\,, \tag{5.49}$$

$$H(z) = \int_0^\infty \frac{\mathrm{d}t}{t} \exp(-tz^2) \sum_{k=1}^{\infty} \exp(-k^2/4t)\,, \tag{5.50}$$

determined from the large-$N$ saddle point condition. $H(z)$ also has a small $z$ expansion

$$H(z) \overset{z\to 0}{\sim} \frac{\pi}{z} + \ln z - \Gamma'(1) - \ln 4\pi + \sum_{\nu=1}^{\infty} (-1)^\nu \binom{2\nu}{\nu} \zeta_R(2\nu+1) \left(\frac{z}{4\pi}\right)^{2\nu}\,, \tag{5.51}$$

where $\zeta_R$ denotes the Riemann zeta function.

In the continuum the coefficients in eq. (5.47) are pure numbers independent of any scale, e.g $L$. However, on the lattice there will be corrections due to the cutoff. Hence in the limit $N \to \infty$ on the lattice one expects an expansion of the form

$$\frac{M}{\Lambda_{\overline{\mathrm{MS}}}} = K(a/L) \cdot z\, \mathrm{e}^{\pi/z} \cdot \left(1 + \sum_{\nu=1}^{\infty} \tilde{a}_\nu(a/L) \left(\frac{z}{2\pi}\right)^\nu\right)\,, \tag{5.52}$$

where the function $K(a/L)$ is different from the function $K(N)$ in eq. (5.47) (in the following it is exclusively referred to the first one). Comparing the last expression to eq. (5.49) and eq. (5.51) one obtains the function $K$ and the coefficients $\tilde{a}_\nu$ in the continuum limit ($a/L \to 0$):

$$\begin{aligned}
K(0) &= \exp[-\Gamma'(1) - \ln 4\pi]\,, && (5.53)\\
\tilde{a}_\nu(0) &= 0 \quad , \quad \nu = \mathrm{odd}\,, && (5.54)\\
\tilde{a}_2(0) &= -1/2\, \zeta_R(3)\,, && (5.55)\\
\tilde{a}_4(0) &= 1/8\, [\, 3\, \zeta_R(5) + \zeta_R(3)^2\,]\,, && (5.56)\\
&\ldots \quad .
\end{aligned}$$

In his letter Lüscher points out that the series eq. (5.47) and eq. (5.51) are absolutely convergent for $|z| < 2\pi$.

But the expansion eq. (5.52) is wanted for finite $a/L$. So one proceeds solving the gap equation for $m_0^2$. First one expands in power series of the bare coupling $\gamma$

$$a^2 m_0^2 = \gamma^2 w_1 \cdot (1 + \gamma\, \overline{w}_2 + \gamma^2\, \overline{w}_3 + \ldots)\,, \tag{5.57}$$

and then compares to the gap equation eq. (5.38) order by order to determine the coefficients. Yielding

$$w_1 = \frac{a^2}{4L^2}\,, \tag{5.58}$$

$$\overline{w}_2 = \frac{a}{L}\sum_{n=1}^{L/a-1} R(n)^{-1/2}\,, \tag{5.59}$$

$$\overline{w}_3 = -\frac{1}{4}w_1 + \frac{3}{4}\overline{w}_2^2\,, \tag{5.60}$$

$$\overline{w}_4 = -\frac{1}{2}w_1\overline{w}_2 + \frac{1}{2}\overline{w}_2^3 - w_1\cdot\frac{a}{2L}\sum_{n=1}^{L/a-1} R(n)^{-3/2}\cdot\left(1+2\sin^2\frac{a\pi n}{L}\right)\,, \tag{5.61}$$

where

$$R(n) = 4\sin^2\frac{a\pi n}{L}\cdot\left[1+\sin^2\frac{a\pi n}{L}\right]\,. \tag{5.62}$$

In order to get an expansion for $z$ this result has first to be inserted in eq. (5.44) and then multiplied by $L$. Using eqs. (5.58) - (5.61) the physical mass is

$$aM = \sqrt{w_1}\gamma\cdot(1+m_2\,\gamma+m_3\,\gamma^2+\ldots)\,, \tag{5.63}$$

with

$$m_2 = \frac{1}{2}\overline{w}_2\,, \tag{5.64}$$

$$m_3 = \frac{1}{4}\overline{w}_2^2 - \frac{1}{6}w_1\,, \tag{5.65}$$

$$m_4 = \frac{1}{8}\overline{w}_2^3 - \frac{1}{4}w_1\overline{w}_2 - w_1\cdot\frac{a}{4L}\sum_{n=1}^{L/a-1} R(n)^{-3/2}\cdot\left(1+2\sin^2\frac{a\pi n}{L}\right)\,. \tag{5.66}$$

Since $\sqrt{w_1} = a/2L$:

$$z = \frac{1}{2}\gamma\cdot(1+m_2\,\gamma+m_3\,\gamma^2+\ldots)\,. \tag{5.67}$$

Inverting the last line finally yields an expansion for the bare coupling $\gamma$ in the dimensionless, renormalization group invariant variable $z$

$$\gamma = 2\,z\cdot(1+\gamma_2\,z+\gamma_3\,z^2+\gamma_4\,z^3+\ldots)\,, \tag{5.68}$$

with

$$\gamma_2 = -2\,m_2\,, \tag{5.69}$$

$$\gamma_3 = 4\,(2\,m_2^2 - m_3)\,, \tag{5.70}$$

$$\gamma_4 = -8\,(5\,m_2^3 - 5\,m_2\,m_3 + m_4)\,. \tag{5.71}$$

This can be used to express the Λ-parameter on the l.h.s. of eq. (5.52) in terms of $z$. To this end one needs the lattice Λ-parameter for $N \to \infty$ [22]

$$\Lambda_{\rm Lat} = \frac{1}{a}\,\mathrm{e}^{-2\pi/\gamma}\cdot\left[1+\mathcal{O}\left(\gamma\right)\right], \tag{5.72}$$

and the relation between the Λ-parameters

$$\frac{\Lambda_{\rm Lat}}{\Lambda_{\overline{\rm MS}}} = \frac{1}{\sqrt{32}}\,. \tag{5.73}$$

For small $z$ one can use eq. (5.68) to rewrite eq. (5.72)

$$\begin{aligned}
\Lambda_{\rm Lat} &= \frac{1}{a}\,\mathrm{e}^{-\frac{\pi}{z}\,(1+\gamma_2\,z+\gamma_3\,z^2+\gamma_4\,z^3+\ldots)^{-1}}\,, && (5.74)\\
&= \frac{1}{a}\,\mathrm{e}^{-\frac{\pi}{z}}\,\mathrm{e}^{\pi\gamma_2}\cdot\Big[1+\pi\,z\,(\gamma_3-\gamma_2^2) && (5.75)\\
&\quad +z^2\left(\pi\{\gamma_4-2\gamma_2\gamma_3+\gamma_2^3\}+\pi^2/2\,\{\gamma_3-\gamma_2^2\}^2\right)+\ldots\Big]\,.
\end{aligned}$$

Using eq. (5.75) and eqs. (5.63) - (5.66) the l.h.s. of eq. (5.52) may be written as a power series in $z$. Then, by comparing equal powers, one obtains the coefficients of the r.h.s.

$$\begin{aligned}
K(a/L) &= \frac{1}{\sqrt{32}}\,\mathrm{e}^{\pi\widetilde{w}_2}\,\frac{a}{L}\,, && (5.76)\\
\tilde{a}_1(a/L) &= -\frac{\pi^2}{3}\,\left(\frac{a}{L}\right)^2\,, && (5.77)\\
\tilde{a}_2(a/L) &= \frac{1}{2}\,\tilde{a}_1(a/L)^2 && (5.78)\\
&\quad -2\pi^3\,\left(\frac{a}{L}\right)^3\sum_{n=1}^{L/a-1} R(n)^{-3/2}\cdot\left(1+2\sin^2\frac{a\pi n}{L}\right)\,,\\
&\ldots\quad .
\end{aligned}$$

The lattice sums appearing in these expressions are too complicated to reveal the $a/L$ dependence. Therefore an asymptotic expansion of these sums is derived using the Euler-Maclaurin formula [31]. The function $K(a/L)$ can be approximated by

$$K(a/L) = K(0)\cdot\left[1+\frac{\pi^2}{36}\,(a/L)^2+\mathcal{O}\left((a/L)^4\right)\right]\,, \tag{5.79}$$

whereas the coefficient $\tilde{a}_2(a/L)$ converges slower due to logarithmic terms (see also Fig. 5.1)

$$\tilde{a}_2(a/L) = \tilde{a}_2(0)+\frac{\pi^2}{4}\,(a/L)^2\cdot\left(2\ln(a/L)+2\ln\pi-\ln 2-\frac{1}{2}\right)+\ldots\,. \tag{5.80}$$

The relative deviation of the asymptotic expansions from the exact value at

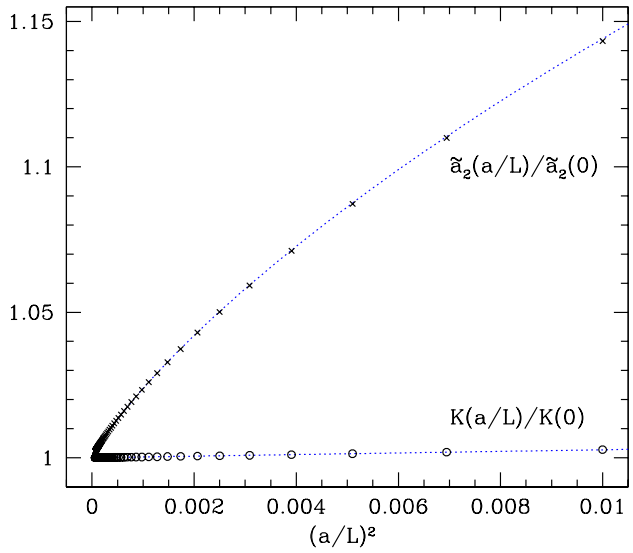



FIGURE 5.1: The function $K(a/L)$ (circles) and the coefficient $\tilde{a}_2(a/L)$ (crosses) normalized by their continuum value together with the corresponding asymptotic expansions eq. (5.79) and eq. (5.80) (dotted lines). $\tilde{a}_2(a/L)$ converges slower due to logarithmic terms.

$L/a = 10$ (100) is $\sim 10^{-5}$ ($10^{-9}$) for $K(a/L)$ and $\sim 10^{-4}$ ($10^{-6}$) for $\tilde{a}_2(a/L)$ (see Table C.3). These asymptotic expansions will be used to get an idea of the step scaling function's leading lattice artifacts in the large $N$ limit. Taking into account eq. (5.79) and eq. (5.80) one expects $\mathcal{O}\left((a/L)^2\right)$ and $\mathcal{O}\left((a/L)^2 \ln(a/L)\right)$ terms.

### 5.2.2 Step Scaling Function

Now, the series derived in the last section will be used to calculate the step scaling function for the variable $z$ on the lattice and in the continuum. The variable $z$ is related to the LWW coupling as

$$\overline{g}^2(L) = \frac{2}{N-1}\, M(L)\, L = \frac{2}{N-1}\, z\,. \tag{5.81}$$

In the continuum the step scaling function for $z$ may be defined as

$$\sigma_z(s, u, N) = M(sL)\, sL\,, \tag{5.82}$$

where the arguments are the number $s > 0$ by which the box size $L$ is scaled, $u = M(L)\, L$ the value of $z$ at the scale $L$ and, of course, $N$. The step scaling function can be understood as an integrated form of the $\beta$-function and a perturbative expansion can be derived (see Section 3.3.2). The step scaling function is to be studied in the limit $N \to \infty$. To this end one uses eq.

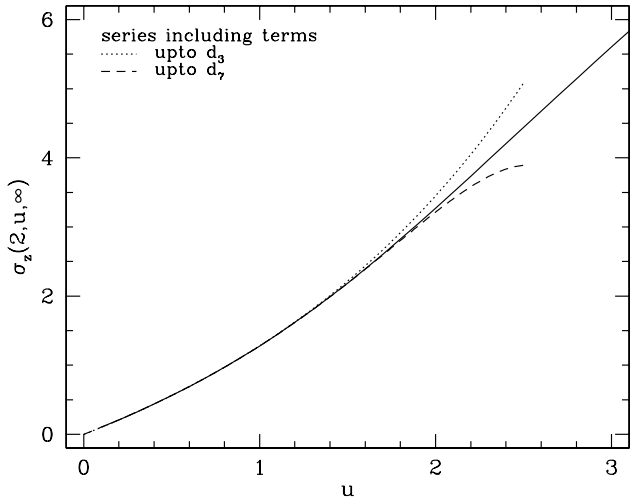



FIGURE 5.2: The continuum step scaling function in the limit $N \to \infty$. For the straight line eq. (5.83) was solved numerically. The dotted and the dashed line are obtained from the series eq. (5.98) in the limit $a/L \to 0$ including coefficients up to $d_3$ and $d_7$ respectively.

(5.49) to write the ratio $u/\sigma_z$ as

$$\frac{u}{\sigma_z(s,u,\infty)}\, s = \frac{f(u)}{f(\sigma_z)}\,. \tag{5.83}$$

This is an implicit equation for $\sigma_z$ that is accessible to numerical evaluation. In such a calculation the sum in the integral of the exponent in eq. (5.50) has to be done only up to $k = 100$ (see Fig. 5.2).

On the lattice, however, one expects cutoff effects, and the lattice step scaling function will in general also depend on the lattice spacing:

$$\Sigma_z(s,u,N,a/L) = M(sL)\, sL\,, \tag{5.84}$$

where $M$ is measured on the lattice by extracting it from an appropriate correlator. Again one is interested in the special case of large $N$. The lattice equivalent of eq. (5.83) is

$$\frac{u}{\Sigma_z(s,u,\infty,a/L)}\, s = \frac{F(u,a/L)}{F(\Sigma_z,a/L)} \tag{5.85}$$

and this will be the starting point of a small $u$ (i.e. $z$) expansion of $\Sigma_z$.

First the following abbreviations are introduced

$$L' = sL \qquad \text{and} \qquad r = \ln s\,. \tag{5.86}$$

Then, just plugging in the ansatz eq. (5.52), one gets

$$\mathrm{e}^r = \frac{K(a/L)\cdot \mathrm{e}^{\pi/u}\cdot\left(1+\sum_{\nu=1}^{\infty}\tilde{a}_\nu(a/L)\left(\frac{u}{2\pi}\right)^\nu\right)}{K(a/L')\cdot \mathrm{e}^{\pi/\Sigma_z}\cdot\left(1+\sum_{\nu=1}^{\infty}\tilde{a}_\nu(a/L')\left(\frac{\Sigma_z}{2\pi}\right)^\nu\right)}\,, \tag{5.87}$$

and by reorganizing the terms

$$\exp\left[\frac{\pi}{\Sigma_z}\right] = \exp\left[\frac{\pi}{u} - r - \ln\frac{K(a/L')}{K(a/L)}\right] \times \times \frac{(1+\tilde{a}_1(a/L)\cdot u/2\pi + \ldots)}{(1+\tilde{a}_1(a/L')\cdot \Sigma_z/2\pi + \ldots)}\,. \tag{5.88}$$

Since $\sigma_z$ is normalized such that in PT the leading order term is equal to $u$ eq. (3.12), a reasonable ansatz for $1/\Sigma_z$ would be

$$\frac{1}{\Sigma_z(s,u,\infty,a/L)} = \frac{1}{u} + c_0(a/L) + c_1(a/L)\,u \;+\; \ldots\,, \tag{5.89}$$

where the first coefficient can easily be read off eq. (5.88). And since evaluating the succeeding coefficients become more and more involved at finite $a/L$ only the next two are given here

$$c_0(a/L) = -\frac{1}{\pi}\left[r + \ln\frac{K(a/L')}{K(a/L)}\right], \tag{5.90}$$

$$c_1(a/L) = \frac{1}{2\pi^2}\left[\tilde{a}_1(a/L) - \tilde{a}_1(a/L')\right], \tag{5.91}$$

$$c_2(a/L) = -\frac{1}{8\pi^3}\left[\tilde{a}_1(a/L)^2 - \tilde{a}_1(a/L')^2 - 4\pi\,\tilde{a}_1(a/L')\,c_0(a/L) - 2\left(\tilde{a}_2(a/L) - \tilde{a}_2(a/L')\right)\right]. \tag{5.92}$$

This will be enough to see how the cutoff effects behave for $u \lesssim 1$ in the large $N$ limit. To reach a satisfying accuracy for the continuum value it will be sufficient to evaluate the higher order coefficients for $a/L \to 0$ ($\tilde{a}_i \equiv \tilde{a}_i(0)$)

$$c_3(0) = -\frac{1}{2\pi^4}\,r\,\tilde{a}_2\,, \tag{5.93}$$

$$c_4(0) = -\frac{3}{4\pi^5}\,r^2\,\tilde{a}_2\,, \tag{5.94}$$

$$c_5(0) = -\frac{1}{\pi^6}\,r\left[r^2\tilde{a}_2 + \frac{1}{8}\left(2\,\tilde{a}_4 - \tilde{a}_2^2\right)\right], \tag{5.95}$$

$$c_6(0) = -\frac{1}{\pi^7}\,r\left[\frac{1}{4}\left(2\,\tilde{a}_2^2 + 5r^2\tilde{a}_2\right) + \frac{5}{16}\left(2\,\tilde{a}_4 - \tilde{a}_2^2\right)\right]. \tag{5.96}$$

$$\tag{5.97}$$

Now inverting the series eq. (5.89) one has

$$\Sigma_z(s,u,\infty,a/L) = u\cdot\left(1 + d_0(a/L)\,u + d_1(a/L)\,u^2 \;+\; \ldots\right), \tag{5.98}$$

with ($c_i \equiv c_i(0)$)

$$d_0(a/L) = -c_0(a/L)\,, \tag{5.99}$$

$$
\begin{aligned}
d_1(a/L) &= c_0(a/L)^2 - c_1(a/L)\,, && (5.100)\\
d_2(a/L) &= -c_0(a/L)^3 + 2\,c_0(a/L)\,c_1(a/L) - c_2(a/L)\,, && (5.101)\\
d_3(0) &= c_0^4 - c_3\,, && (5.102)\\
d_4(0) &= -c_0^5 + 2\,c_0\,c_3 - c_4\,, && (5.103)\\
d_5(0) &= c_0^6 - 3\,c_0^2\,c_3 + 2\,c_0\,c_4 - c_5\,, && (5.104)\\
d_6(0) &= -c_0^7 + 4\,c_0^3\,c_3 - 3\,c_0^2\,c_4 + 2\,c_0\,c_5 - c_6\,, && (5.105)\\
&\dots\ .
\end{aligned}
$$

Again from the third coefficient on the continuum values are used ($a/L = 0$). One has to check whether this is a good approximation. The MC data in this thesis is gained by doubling the lattice size, i.e. $s = 2$. Then, plugging in the asymptotic expansions eq. (5.79) and eq. (5.80), the final result reads

$$
\begin{aligned}
d_0(a/L) &= 0.2206356 - 0.0654499\,(a/L)^2 + \mathcal{O}\left((a/L)^4\right)\,, && (5.106)\\
d_1(a/L) &= 0.0486801 + 0.0961189\,(a/L)^2 + \mathcal{O}\left((a/L)^4\right)\,, && (5.107)\\
d_2(a/L) &= 0.0107406 + 0.0131547\,(a/L)^2 + \\
&\qquad + 0.0298416\,(a/L)^2\,\ln(a/L) + \dots\,, && (5.108)\\
d_3(0) &= 0.0002313\,, && (5.109)\\
d_4(0) &= -0.0011285\,, && (5.110)\\
d_5(0) &= -0.0006473\,, && (5.111)\\
d_6(0) &= -0.0002288\,, && (5.112)\\
d_7(0) &= -0.0000445\,, && (5.113)\\
&\dots\ .
\end{aligned}
$$

### 5.2.3 Checking the Expansion

The series eq. (5.89) together with the coefficients (5.106) - (5.113) is truncated in three ways. The first step was to expanded in $z$ including terms up to $d_l\,z^{l+2}$ with $l \leq 7$. Furthermore only the first three coefficients have been calculated at finite lattice spacing. The remaining ones are continuum values. This is the second step. Finally, for the lattice sums in the lower order coefficients, representing the cutoff dependence, asymptotic expansions in $a/L$ including quadratic terms and logarithms have been derived. One has to check every step and answer the question whether they lead to an easy to handle but still meaningful approximation for the parameter region of interest. At the end it should provide us with a clue to the leading lattice artifacts.

*Small z expansion* As mentioned before it is possible to numerically solve the implicit relation eq. (5.83) for the continuum step scaling function. By setting $a/L = 0$ one can compare the series to the exact value in the continuum, see Fig. 5.2 and Table C.4. The series including terms up to $d_3$ and

up to $d_7$ has been evaluated. From the plot and the numbers one concludes that the expansion can be trusted for $u \lesssim 1$ and for $u \lesssim 1.5$ respectively. It is understood that this, so far, counts only for $a/L = 0$.

*Coefficients at finite lattice spacing* Looking at the factors belonging to the $\mathcal{O}\left((a/L)^2\right)$ terms in (5.106) - (5.108) one recognizes that they have different signs. Therefore the approach to the continuum should change from monotonic increasing to monotonic decreasing as $u$ goes from zero to one. To see whether this is not only a feature of the expansion but also true for the real approach one needs to know the exact lattice step scaling function. To this end one reapplies the gap equation

$$\frac{1}{\gamma} = \frac{a}{L} \sum_{n=0}^{L/a-1} \frac{1}{2\,\epsilon(n)} \equiv \mathcal{G}(L/a, M)\,, \tag{5.114}$$

where $\epsilon(n)$ is defined in eq. (5.39) and one should recall the relation between the expectation value $m_0^2$ of the auxiliary field $\alpha$ and the finite volume mass $M(L)$ eq. (5.43). Since a point of the step scaling function $\Sigma_z(2, u, \infty, a/L)$ is understood as the value of the quantity $z$ when $L$ is doubled but the bare coupling is kept fixed one can use eq. (5.114) to write

$$\mathcal{G}(L/a, M(L)) = \mathcal{G}(2L/a, M(2L))\,. \tag{5.115}$$

So by fixing $L$ and $u = M(L)\,L$ the only unknown quantity is $M(2L)$. The equation can be treated with numerical root finding methods giving the exact value of the lattice step scaling function for arbitrary lattice spacings. Thus the approach to the continuum can be studied for $u \in [0, 1]$ and the expected picture really shows up. In Fig. 5.3 the exact values for $L/a = 10, 11, \ldots, 500$ and the series are plotted for three values of $u$ covering the region where the change in the behavior of the lattice artifacts occurs. The approach to the continuum is from below for $u \lesssim 0.4$ and from above for $u \gtrsim 0.53$. In-between these values lies a transition region where the approach's slope changes sign at finite lattice spacing. Coming back to the question of the series' reliability, one has to admit that only for the smallest $u$ value the series looks quite good also at coarser lattices. Already at $u = 0.47$ the series quantitatively only reproduces the continuum value. Nevertheless, at finite lattice spacing, the series seems to show the right qualitative behavior. This supports the assumption of quadratic and logarithmic terms for the leading lattice artifacts in the large $N$ limit. One could probably reduce the deviation for larger $u$ by evaluating more coefficients at finite lattice spacing. But a clue to the cutoff dependence is already gained.

*Asymptotic expansion* To check if such a combination of terms is not only able to reproduce the lattice artifacts correctly but is also better than other

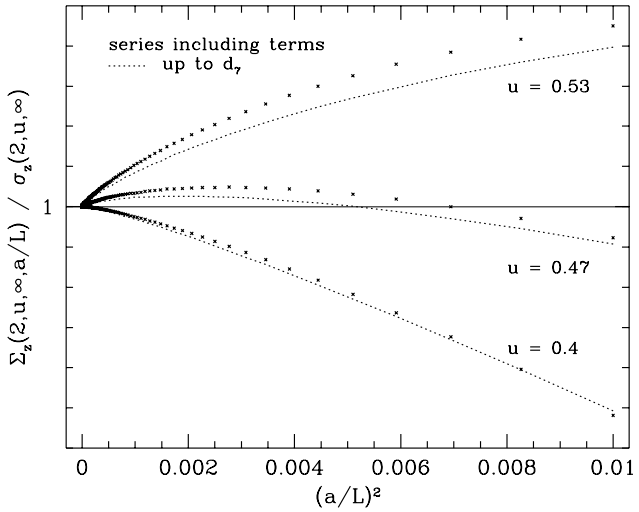



FIGURE 5.3: The approach to the continuum of lattice step scaling function in the limit $N \to \infty$ normalized by its continuum value. The three couples of points with curve correspond to $u = 0.4$, 0.47, 0.53 covering the region where the approach changes qualitatively. The points are obtained by solving the gap equation numerically (see text). The dotted lines are obtained from the series eq. (5.98) including coefficients up to $d_7$.

obvious choices; the exact values for $u = 1.0595$ have been tried to fit to several forms

$$\text{A}: \quad k_0 + k_1 (a/L)^2 , \tag{5.116}$$

$$\text{B}: \quad k_0 + k_1 (a/L)^{k_2} , \tag{5.117}$$

$$\text{C}: \quad k_0 + k_1 (a/L) + k_2 (a/L)^2 , \tag{5.118}$$

$$\text{D}: \quad k_0 + k_1 (a/L)^2 + k_2 (a/L)^2 \ln(a/L) . \tag{5.119}$$

The fits were done by minimizing the normalized quadratic deviations

$$\chi^2 = \min_{[k_0,k_1,k_2]} \sum_{\{a/L\}} \frac{(\Sigma_z - \Sigma_z^{\text{fit}})^2}{\Sigma_z^2} \tag{5.120}$$

and included the points in Table 5.1. The $\chi^2$ values in the last column of Table 5.2 suggest that the cutoff dependence is really of the kind B or D rather than purely quadratic or linear and quadratic.

| $L/a$ | $\Sigma_z(2, u, \infty, a/L)$ |
|---|---|
| 100 | 1.3759882 |
| 150 | 1.3759742 |
| 200 | 1.3759689 |
| 250 | 1.3759664 |
| 300 | 1.3759649 |
| 350 | 1.3759641 |
| 400 | 1.3759635 |
| 450 | 1.3759631 |
| 500 | 1.3759628 |

TABLE 5.1: Points included in the fits at $u = 1.0595$.

| | $k_0$ | $k_1$ | $k_2$ | $\chi^2$ |
|---|---|---|---|---|
| A | 1.3759619627 | 0.2653 | - | $10^{-13}$ |
| B | 1.3759613293 | 0.1201 | 1.8248 | $10^{-16}$ |
| C | 1.3759608340 | 0.0005 | 0.2246 | $10^{-15}$ |
| D | 1.3759613805 | -0.0478 | 0.0483 | $10^{-21}$ |

TABLE 5.2: Fits of the forms (5.116) - (5.119). Exact continuum value: $\sigma_z = 1.3759613806$.

# Chapter 6

# Results

## 6.1 Comparison of the Cutoff Effects: $N = 3, 4, 8, \infty$

In this Section the results of the MC simulations in the nonlinear $O(4)$ and $O(8)$ sigma model will be presented. The MC data will be compared to former results in the $O(3)$ model and to the exact large $N$ limit.

The $O(3)$ data is taken from [6], wherein Seefeld et. al. study the cutoff effects measuring the lattice step scaling function $\Sigma_z(s, u, N, a/L)$ at $u = 1.0595$. They confirmed the nonstandard cutoff effects (see Section 3.2.2). The value of $u = M(L)\,L$ is some sort of "canonical" and appears in many papers. I decided to continue this tradition. This makes it possible to directly compare my own results for $N = 4, 8$ to the $N = 3$ case. In the limit $N \to \infty$ the step scaling function can be computed by numerically solving analytical expressions in virtually no time on a modern PC for any value of $u$ (see Section 5.2.3). So, there is data for four values of $N$ available (Table 6.1).

The standard assumption for the lattice artifacts, originating from the pertubative analysis by Symanzik (Section 3.1.2), would be a leading term quadratic in the lattice spacing. Therefore all data is firstly plotted against $(a/L)^2$. On the next pages the lattice step scaling function at $u = 1.0595$ for $N = 3, 4, 8, \infty$ is display Fig. 6.1 - Fig. 6.4. The smallest lattice included in the plots has $L/a = 10$. The x-axis are the same in all four plots and the y-axis are equally scaled. Also the continuum values are marked where available.

By comparing the amplitude of the lattice artifacts the first observation is that the cutoff effects are weaker for larger $N$. Looking at the values at the smallest lattice, the deviation from the continuum value for $N = 4$ is roughly half of that for $N = 3$. The same can be said comparing $N = 4$ and $N = 8$. The latter seems to be already similar to $N = \infty$.

From the plot in Fig. 6.1 one can immediately tell that these points cannot be linearly fitted to $(a/L)^2$: the dependence rather looks like a square root. For $N = 4$ and $N = 8$ the cutoff dependence is to weak and it is not visible to the naked eye whether the leading term is quadratic or something else.

In the next section fits to several forms are compared in order to see what kind of dependence is able to reproduce the data best.

| $L/a$ | $N = 3$ | $N = 4$ | $N = 8$ | $N = \infty$ |
|---|---|---|---|---|
| 5 | 1.29379(8) | | - | - |
| 6 | - | 1.31256(47) | 1.35065(40) | 1.37949129 |
| 8 | - | 1.31068(43) | 1.34891(35) | 1.37821466 |
| 10 | 1.27994(9) | 1.30825(39) | 1.34835(34) | 1.37752387 |
| 12 | 1.27668(9) | 1.30608(37) | 1.34721(32) | 1.37711185 |
| 16 | 1.27228(12) | 1.30477(34) | 1.34725(30) | 1.37666489 |
| 24 | 1.26817(9) | 1.30379(32) | 1.34649(28) | 1.37630852 |
| 32 | 1.26591(9) | 1.30263(31) | 1.34635(27) | 1.37606266 |
| 64 | 1.26306(16) | 1.30225(32) | 1.34570(25) | 1.37602171 |
| 96 | - | 1.30138(33) | 1.34577(24) | 1.37599030 |
| 128 | - | 1.30149(34) | 1.34563(24) | 1.37597848 |

TABLE 6.1: Monte Carlo data and large $N$ values for the lattice step scaling function $\Sigma_z(2, u, N, a/L)$. The $N = 3$ data is taken from [6].

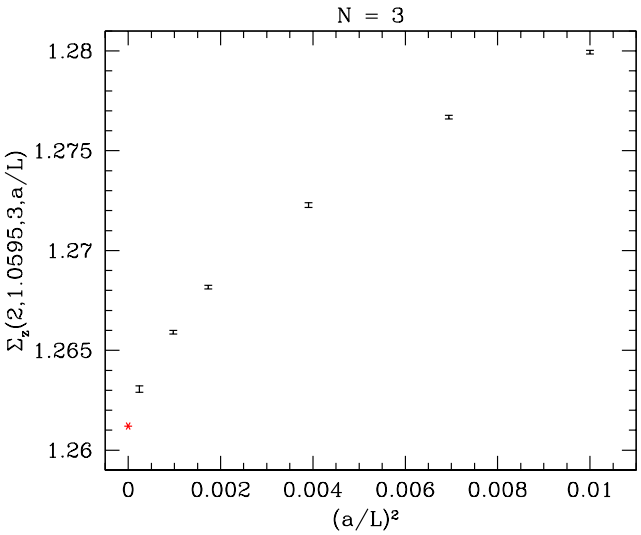



FIGURE 6.1: Lattice artifacts of the step scaling function in the $O(3)$ model. The MC data is taken from [6]. The continuum value (star) is published in [12].

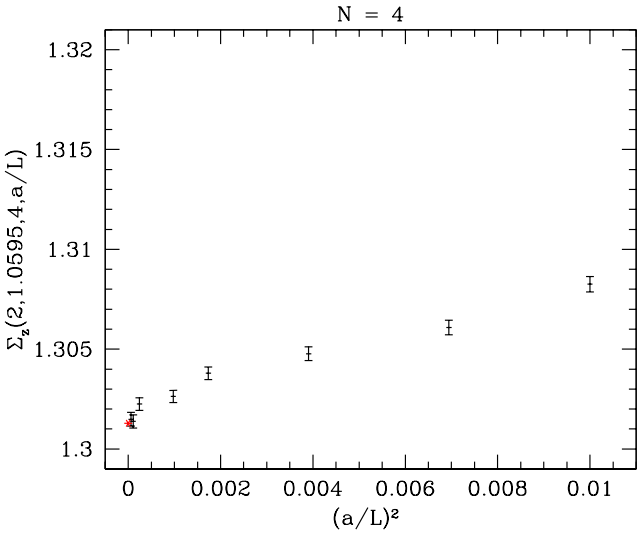



FIGURE 6.2: Lattice artifacts of the step scaling function in the $O(4)$ model. The plot contains the MC data and the continuum value (star) provided by Balog on the base of [12].

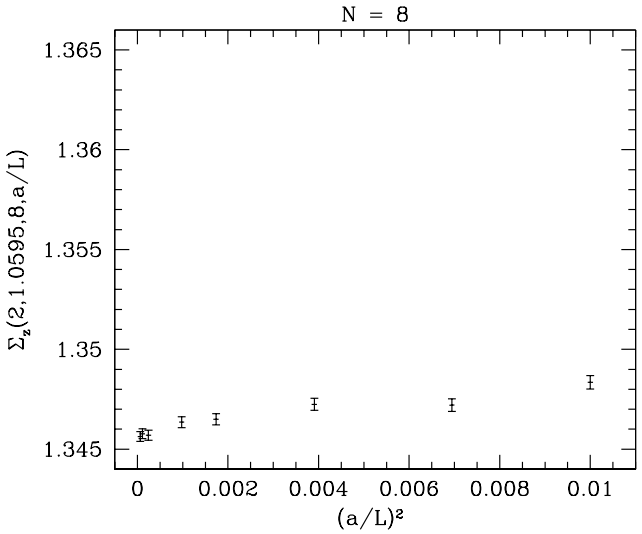



FIGURE 6.3: Lattice artifacts of the step scaling function in the $O(8)$ model. The plot contains the MC data. So far, an analytically computed continuum value is not available.

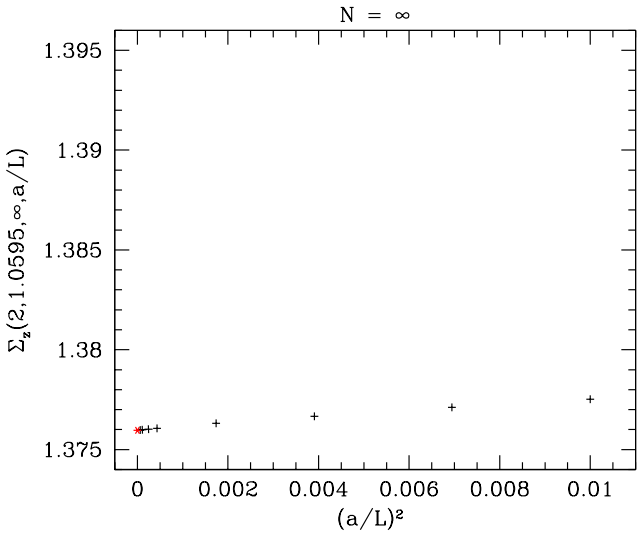



FIGURE 6.4: Lattice artifacts of the step scaling function in the large $N$ limit. The crosses are computed as described in Section 5.2.3.

## 6.2 Fit of the Lattice Artifacts

In Fig. 6.1 and Fig. 6.2 continuum values are marked. These values are provided by an analytical approach to compute the mass gap in the $O(3)$ and $O(4)$ model [12]. The derivation is not rigorous and the numbers are therefore predictions that have to be compared to other approaches. Since the value for $N = 3$ seems to be consistent with MC data [12],[6] the value for $N = 4$ is assumed to be correct as well. Accordingly, in the following sections the continuum values are used to constrain the fits and increase the significance.

### 6.2.1 Simple Linear or Quadratic

Considering the discussion of Section 3.2.4 the simplest and most obvious forms to fit to are just a linear or a quadratic term

$$\text{A}: \quad c_0 + c_1 \, (a/L) \,, \tag{6.1}$$

$$\text{B}: \quad c_0 + c_1 \, (a/L)^2 \,. \tag{6.2}$$

The quadratic fit corresponds to the PT prediction and the linear fit corresponds to the nonstandard behavior in the $O(3)$ model. The two forms are fitted in two ways: with $c_0$ as free parameter, referred to as *unconstrained* fit, and with $c_0$ fixed by the analytical predictions [12]

$$\sigma_z(2, 1.0595, N = 3) = 1.261208(1) \,, \tag{6.3}$$

$$\sigma_z(2, 1.0595, N = 4) = 1.3012876(1) \,, \tag{6.4}$$

referred to as *constrained* fit. The smallest lattice included in the fits is $L/a = 10$. In Fig. 6.5(a), Fig. 6.6(a) and Fig. 6.7 the results of the unconstrained fits are displayed for $N = 3, 4, 8$. For $N = 3, 4$ the continuum value is marked.

In the case of the unconstrained fits the deviation of the extrapolation from the analytic prediction normalized by the error of the extrapolation can be evaluated

$$|\sigma - c_0| / \delta c_0 \,. \tag{6.5}$$

This quantity should takes values between 1 and 2 for a fit of statistical data; if the analytic prediction is assumed to be true. Another value characterizing the fits is *"chi-squared per degrees of freedom"* $\chi^2/dof$. This means the sum of the squared deviations of the fit from the data normalized by the squared errors of the data. This value should be roughly 1 for a fit of statistical data. One can rate the quality of a fit by asking for the probability to find a $\chi^2/dof$ larger than present one. This should be larger than 10% [32].

Both quantities have been calculated for the linear and the quadratic fits of the $O(3)$ and $O(4)$ model. Furthermore the dependence on the smaller

| $N$ | Fit | $c_0$ | $c_1$ | $\chi^2/dof$ | $\lvert\sigma - c_0\rvert/\delta c_0$ |
|---|---|---|---|---|---|
| 3 | A | 1.259675(96) | 0.2030(14) | 5.0/4 | 15.9 |
| | | - | 0.18229(61) | 256/5 | - |
| | B | 1.264845(66) | 1.593(11) | 515/4 | 55.5 |
| | | - | 2.0891(71) | 3597/5 | - |
| 4 | A | 1.30082(20) | 0.0681(39) | 7.6/6 | 2.4 |
| | | - | 0.0607(23) | 13.2/7 | - |
| | B | 1.30192(15) | 0.645(37) | 13.1/6 | 4.2 |
| | | - | 0.743(29) | 30.4/7 | - |
| 8 | A | 1.34542(15) | 0.0266(32) | 3.3/6 | - |
| | B | 1.34583(12) | 0.253(31) | 6.7/7 | - |

TABLE 6.2: Fits to the forms eq. (6.1) and eq. (6.2) for $L_{\min}/a = 10$. For $N = 3, 4$ results of unconstrained (first line) and constrained (second lined) fit are listed.

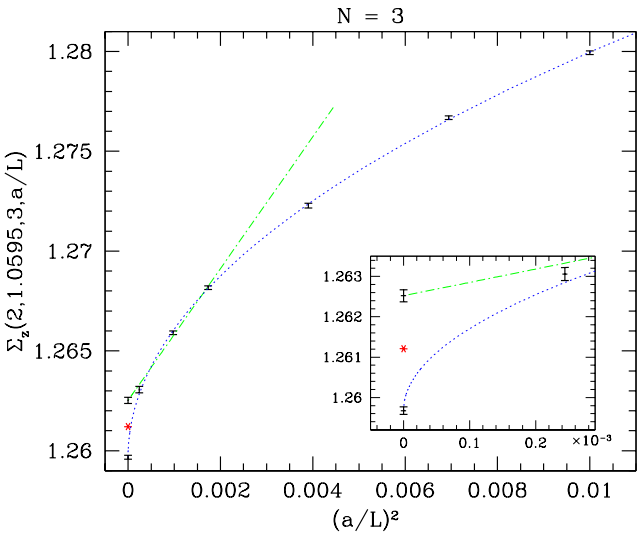
lattices was investigated by successively omitting these. The resulting curves are plotted in Fig. 6.5(b,c) and Fig. 6.6(b,c) respectively. The points in these plots correspond to fits including lattices with $L \geq L_{\min}$.

For $N = 3$ one sees, just by the scale of the plot, that the quadratic form must be rejected, even if only the three largest lattices are left. The constrained linear fit has a much too large $\chi^2/dof$ and is therefore eliminated as well. The unconstrained linear fit has a acceptable $\chi^2/dof$, but it has to be refused if the continuum value eq. (6.3) is assumed to be correct.

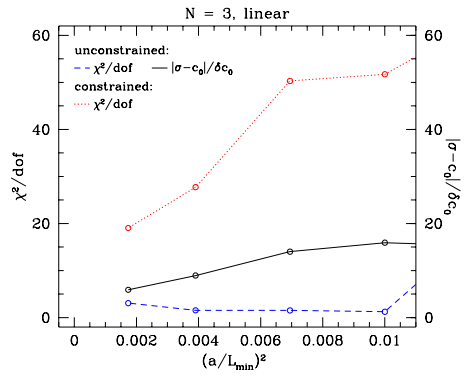
Since the ratio: amplitude of lattice artifacts to error of the data, is worser for $N = 4$, the information from a similar analysis is not so distinct. If only enough small lattices are dropped, then all fits become acceptable. But, comparing linear and quadratic fit, one has to admit that the linear fit is already good for $L_{\min}/a = 10$, whereas the quadratic fit becomes tolerable not until $L_{\min}/a \geq 16$.

The situation becomes even worse for $N = 8$. The lattice artifacts, compared to the error of the data, are to small to eliminate one of the two forms or to at least prefer one. Also, no analytical prediction is available so far. So the analysis of the small lattice dependence is omitted for $O(8)$. The parameters of the fits in Fig. 6.7 are listed in Table 6.2.

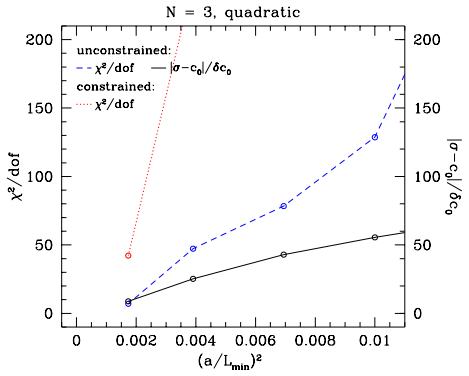
Neither the quadratic nor the linear form is able to reproduce the lattice artifacts correctly in the $O(3)$ model. For $N = 4$ the quadratic artifacts seem to be unlikely, at least for $L_{\min}/a = 10$. There is no preferred dependence for $N = 8$.




(a) Linear (dotted) and quadratic (dashed) unconstrained fit. The insert shows a close up of the extrapolation to the continuum. The star marks the continuum value. The linear fit includes lattices $L/a \geq 10$ and the quadratic fit includes only the three largest lattices ($L/a = 24, 32, 64$). Clearly, a quadratic fit of all data points can be refused by the naked eye.

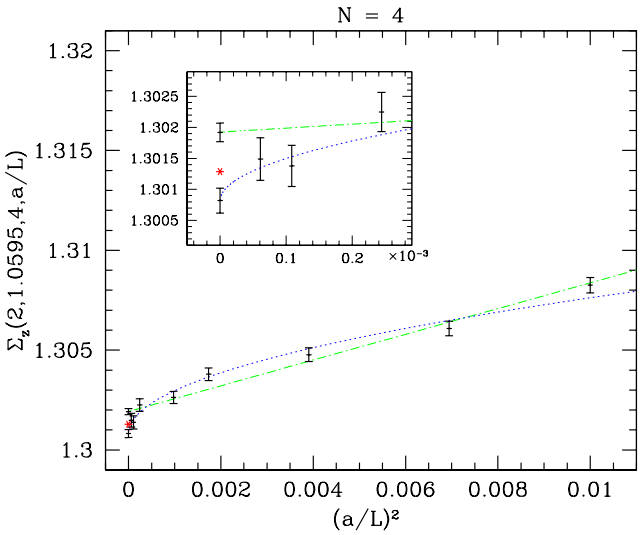



(b)

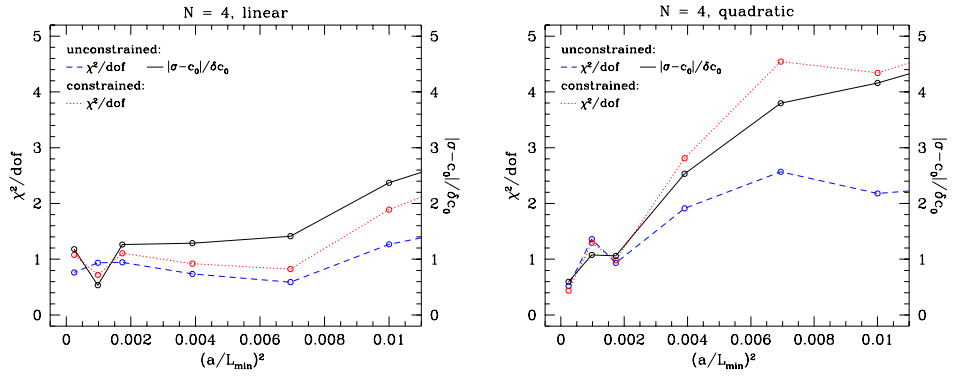



(c)

FIGURE 6.5: Simple linear and quadratic fit of former $O(3)$ MC data. Fig. (b) and (c) show the behavior of the fits when the smaller lattices are successively omitted: $L_{\min}/a$ is the smallest lattice included. The last point in these two plots refers to a fit including $L/a = 24, 32, 64$.




(a) Linear (dotted) and quadratic (dashed) unconstrained fit including lattices $L/a \geq 10$. The insert shows a close up of the extrapolation to the continuum. The star marks the continuum value.




(b) (c)

FIGURE 6.6: Simple linear and quadratic fit of the $O(4)$ MC data. Since the lattice artifacts are smaller than in the $N = 3$ case, the plots are not so predictive.

| $N$ | $c_0$ | $c_1$ | $\alpha_{\rm opt}$ | $\chi^2/dof$ | $\lvert\sigma - c_0\rvert/\delta c_0$ |
|---|---|---|---|---|---|
| 3 | 1.259914(95) | 0.2120(15) | 1.02(4) | 4.6/3 | 13.7 |
| | - | 0.27337(91) | 1.16(2) | 19.4/4 | - |
| 4 | 1.30130(18) | 0.1275(73) | 1.29(22) | 5.8/5 | 0.09 |
| | - | 0.1251(48) | 1.28(12) | 5.8/6 | - |
| 8 | 1.34546(15) | 0.0294(35) | 1.05(44) | 3.3/5 | - |

TABLE 6.3: Fits to the form eq. (6.6) for $L_{\rm min}/a = 10$. For $N = 3, 4$ results of unconstrained (first line) and constrained (second lined) fit are listed.

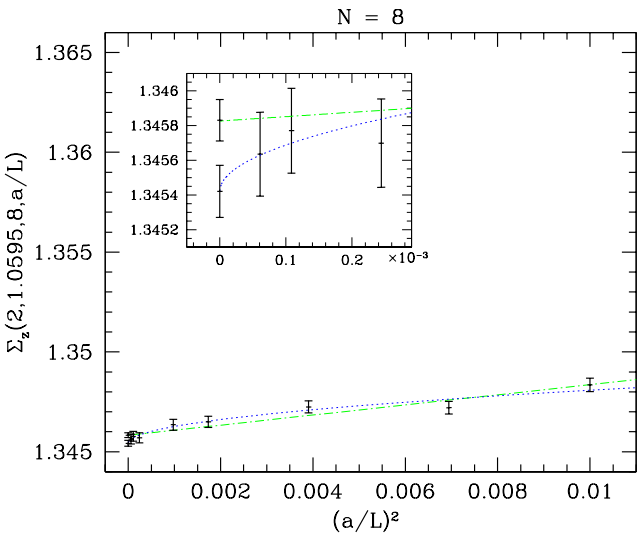



FIGURE 6.7: Linear (dotted) and quadratic (dashed) unconstrained fit including lattices $L/a \geq 10$. The insert shows a close up of the extrapolation to the continuum. The lattice artifacts are too small to rule out the quadratic or the linear fit.

### 6.2.2 Rational Exponent

In Fig. 6.5(a) one observes that the quadratic form overestimates the continuum value, whereas the linear form underestimates it. Therefore an exponent between 1 and 2 may meet this point, i.e. a fit to the form

$$c_0 + c_1(a/L)^\alpha \,, \tag{6.6}$$

with the exponent $\alpha$ as a free parameter. This is an ad hoc ansatz, which has no theoretical motivation. But one can consider it as a formal possibility. For the $O(3)$ data this was also done in [6].

Fitting a form like eq. (6.6) is a nonlinear problem. It is solved by minimizing the $\chi^2$ of a linear fit problem when $\alpha$ ia fixed. For a given $\alpha$ the fit of eq. (6.6) is solved by linear algebra yielding $\chi^2(\alpha)$. Then this function is minimized with respect to $\alpha$. This technique is called least squares method [32].

The error of the so determined optimal exponent $\alpha_{\text{opt}}$ is approximated by the amount $\delta\alpha$, by which $\alpha_{\text{opt}}$ can be changed so that $\chi^2$ increases by 1

$$\chi^2(\alpha_{\text{opt}} + \delta\alpha) - \chi^2(\alpha_{\text{opt}}) \stackrel{!}{=} 1 \,. \tag{6.7}$$

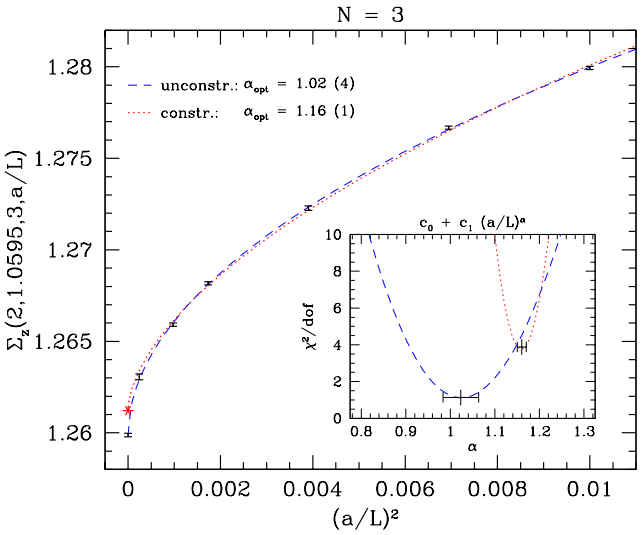
Since the function $\chi^2(\alpha)$ is found to be symmetric in the vicinity of $\alpha_{\text{opt}}$ only the positive solution to eq. (6.7) is stated.

Again, for $N = 3, 4$, the continuum values eq. (6.3) and eq. (6.4) are assumed to be valid and used to constrain the fits. For the $O(3)$ MC data (Fig. 6.8) the unconstrained fit says $\alpha_{\text{opt}} \approx 1$. But already in the last section it was observed that this is not compatible with the continuum value. Therefore it is not surprising to see that the constrained fit leads to a very different optimal exponent. By dropping the small lattices, it seems to stabilize at $\alpha_{\text{opt}} \approx 1.2$. Note that the fit for $L_{\text{min}}/a = 10$ must be refused, because of the large chi-squared per d.o.f. (Fig. 6.8(c) and Table 6.3).

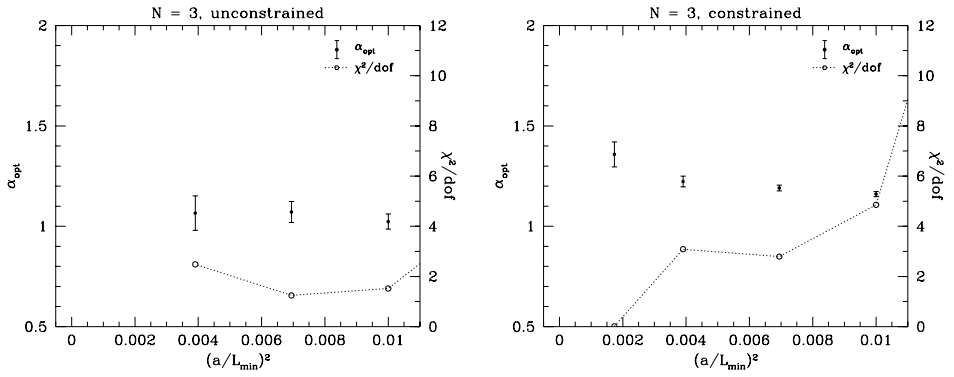
In contrast to $N = 3$, the unconstrained and the constrained fit for $N = 4$ lead to the same optimal exponent, though with a different error (Fig. 6.9). Its value fluctuates between 1 and 1.5, but for $L_{\text{min}}/a = 24$ even the constrained fit cannot exclude an $\alpha = 2$.

Unfortunately the lattice artifacts are such small for $N = 8$, that, with now three free parameters, the fit contains virtually no information about the form (Fig. 6.10). Only a continuum extrapolation can be obtained.

The fit to the ad hoc ansatz eq. (6.6) is very unlikely for $N = 3$ when considering the continuum value eq. (6.3) as correct. For $N = 4$ the tendency is again more towards linear then towards quadratic artifacts, but neither can be refused.

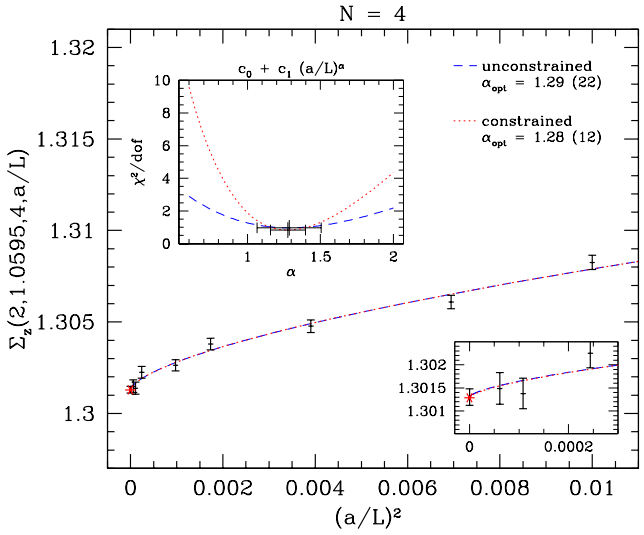



(a) The unconstrained (dashed) and constrained (dotted) fit including lattices $L/a \geq 10$. The star marks the continuum value. The insert shows the chi-squared per d.o.f. for both fits as function of the exponent $\alpha$. The minimum indicates $\alpha_{\rm opt}$. The unconstrained fit misses the continuum value, therefore the constrained fit leads to a different $\alpha_{\rm opt}$

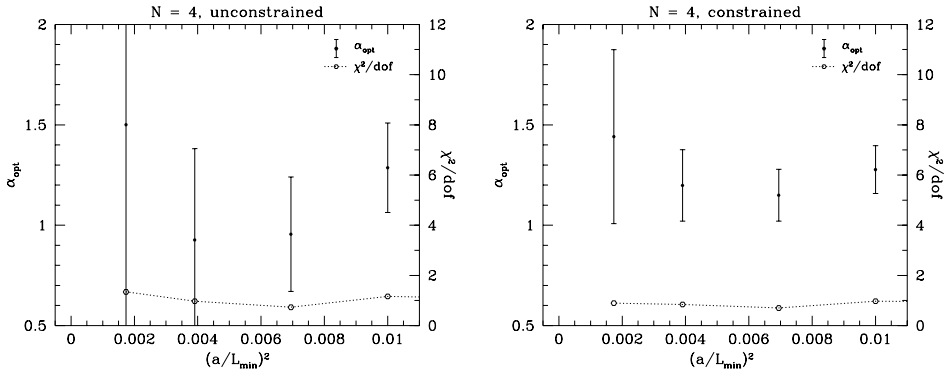



(b) (c)

FIGURE 6.8: Rational exponent fit of former $O(3)$ MC data. Fig. (b) and (c) show the behavior of the fits when the smaller lattices are successively omitted. The plots include $\alpha_{\rm opt}$ and the chi-squared per d.o.f.. Note that the two quantities belong to differently scaled y-axis.

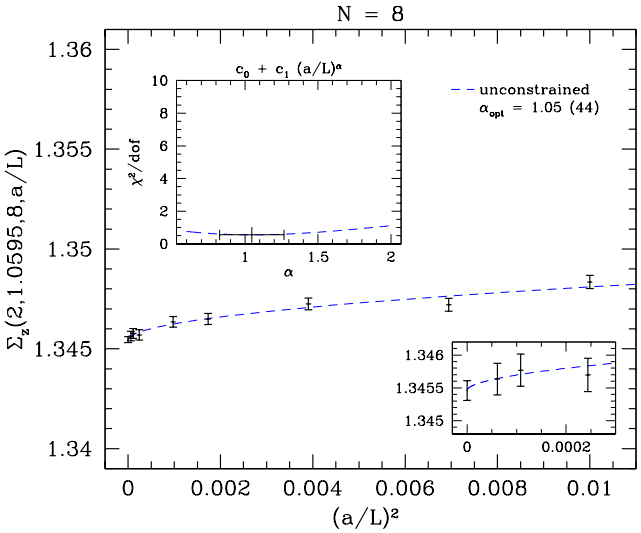



(a) The unconstrained (dashed) and constrained (dotted) fit including lattices $L/a \geq 10$. The inserts shows a close up of the extrapolation to the continuum and the chi-squared per d.o.f. as function of the exponent $\alpha$. The star marks the continuum value. Both fits lead to the same $\alpha_{\rm opt}$, but with a different error.

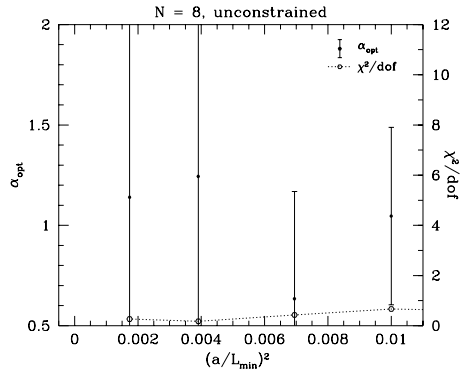



(b) (c)

FIGURE 6.9: Rational exponent fit of the $O(4)$ MC data. By constraining the fit the error of the optimal exponent becomes much smaller. For $L_{\rm min}/a = 24$ an $\alpha = 2$ cannot be excluded even in the constrained fit.




(a) The unconstrained (dashed) and constrained (dotted) fit including lattices $L/a \geq 10$. The inserts shows a close up of the extrapolation to the continuum and the chi-squared per d.o.f. as function of the exponent $\alpha$. The minimum is too flat to really restrict the exponent.




(b)

FIGURE 6.10: Rational exponent fit of the $O(8)$ MC data. The lattice artifacts are too small and the error too large, no information here, besides the continuum extrapolation. The form is not important, all results are compatible.

| $N$ | Fit | $c_0$ | $c_1$ | $c_2$ | $\chi^2/dof$ | $\lvert\sigma - c_0\rvert/\delta c_0$ |
|---|---|---|---|---|---|---|
| 3 | C | 1.25976(23) | 0.1996(88) | 0.028(70) | 4.8/3 | 6.2 |
| | | - | 0.1472(25) | 0.423(29) | 43.0/4 | - |
| | D | 1.26208(14) | -1.846(82) | -2.47(18) | 11/3 | 6.2 |
| | | - | -2.297(39) | -3.436(93) | 49.5/4 | - |
| 4 | C | 1.30117(31) | 0.045(16) | 0.23(15) | 5.3/5 | 2.4 |
| | | - | 0.0396(79) | 0.274(98) | 3.5/6 | - |
| | D | 1.30155(22) | -0.46(19) | -0.43(44) | 7.0/5 | 1.2 |
| | | - | -0.62(13) | -0.77(32) | 8.4/6 | - |

TABLE 6.4: Fits to the forms eq. (6.9) and eq. (6.10) for $L_{\min}/a = 10$. For $N = 3, 4$ results of unconstrained (first line) and constrained (second lined) fit are listed.

### 6.2.3 More Than One Term

Both, Symanzik's perturbative analysis and the large $N$ limit suggest lattice artifacts that are quadratic in $a$ up to logarithms:

$$a^2 \left(\ln a\right)^l , \tag{6.8}$$

where for PT $l \geq 0$ is the loop order and for large $N$ the exponent is $l = 1, 0, -1, -2, \ldots$ (see Sections 3.1.2 and 3.2.3). The two leading terms coincide and therefore a fit to a form containing these two terms, referred to as PT fit, will be tested. Also a polynomial fit of order two is presented:

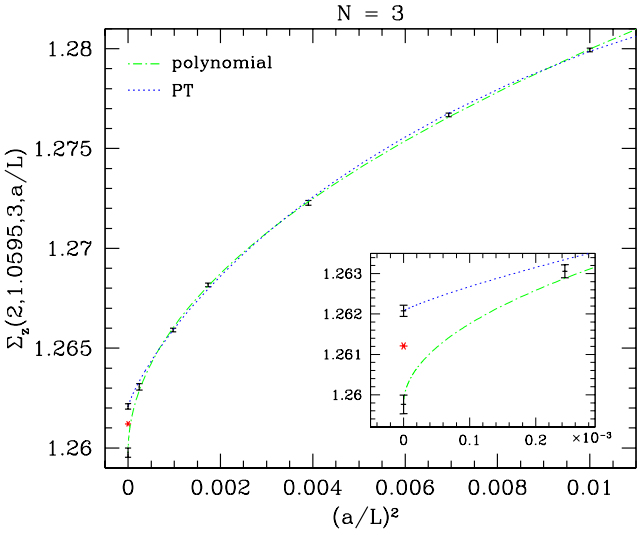
$$\text{C}: \quad c_0 + c_1 \left(a/L\right) \ + c_2 \left(a/L\right)^2 , \tag{6.9}$$

$$\text{D}: \quad c_0 + c_1 \left(a/L\right)^2 + c_2 \left(a/L\right)^2 \ln(a/L) . \tag{6.10}$$

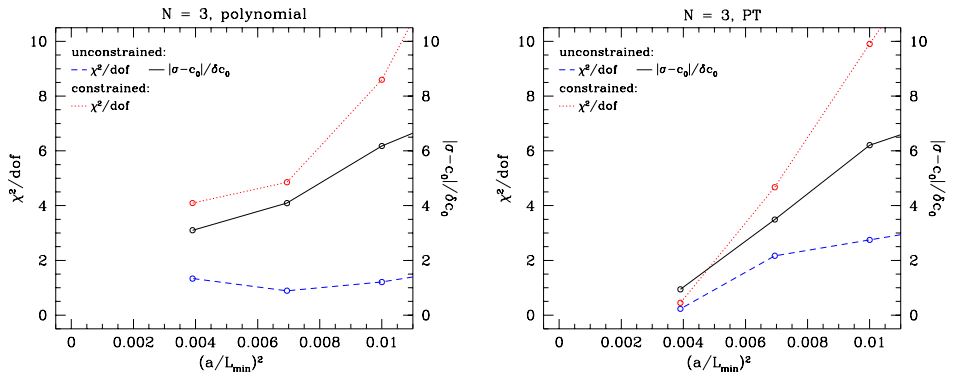
Because of the additional term the number of free parameters is three again. For $N = 8$ one expects no new information, because the lattice artifacts are too small. Therefore only $N = 3, 4$ are fitted to the forms above.

In the case $N = 3$ and $L_{\min}/a = 10$ both fits have to be refused (Fig. 6.11). At $L_{\min}/a = 16$ the polynomial is still quite unlikely, it had to be refused on the 4% level. The PT fit becomes better for smaller $L_{\min}/a$ and is perfect for $L_{\min}/a = 16$.

Both fits are absolutely acceptable for $N = 4$ already at small $L_{\min}/a$. Therefore one cannot obtain information about the cutoff dependency, besides that these two forms are not rejected at this precision.

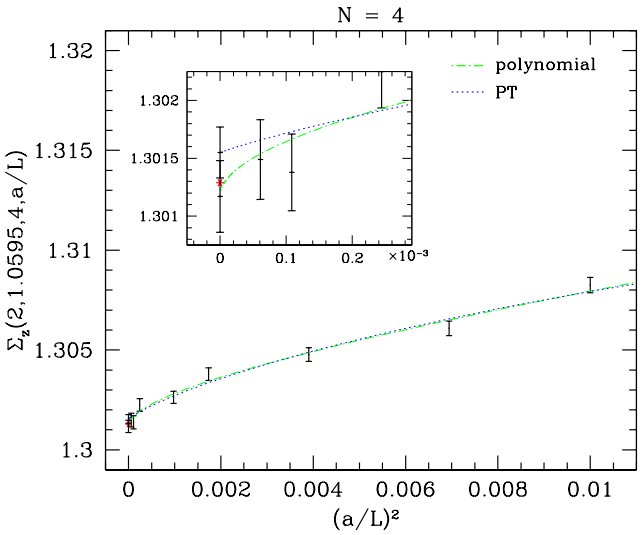



(a) Polynomial (dotted) and PT (dashed) unconstrained fit including lattices $L/a \geq 10$.. The insert shows a close up of the extrapolation to the continuum. The star marks the continuum value. Both fits miss the continuum value by 6 standard deviations.

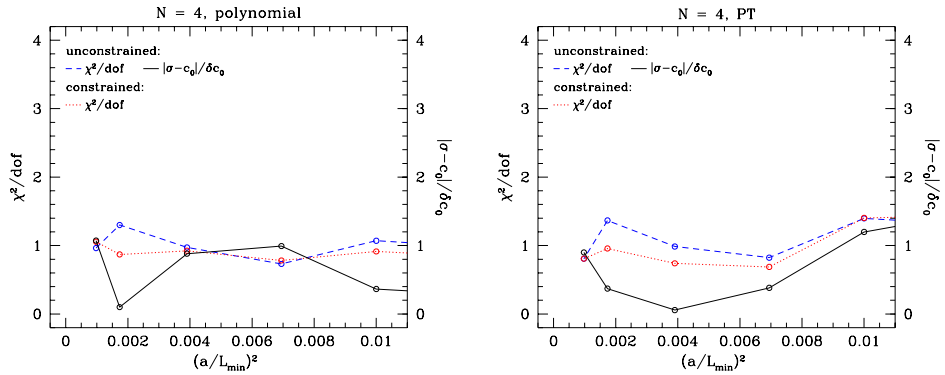



(b) (c)

FIGURE 6.11: Polynomial and PT fit of former $O(3)$ MC data. Fig. (b) and (c) show the behavior of the fits when the smaller lattices are successively omitted. For $L_{\min}/a = 10$ both fits have to be refused. The polynomial fit is also for $L_{\min}/a = 16$ (one degree of freedom) quite unlikely, whereas the PT fit is perfect.




(a) Polynomial (dotted) and PT (dashed) unconstrained fit including lattices $L/a \geq 10$. The insert shows a close up of the extrapolation to the continuum. The star marks the continuum value. Both fits are compatible with the continuum value.




(b) (c)

FIGURE 6.12: Polynomial and PT fit of the $O(4)$ MC data. Both fits work fine.

## 6.3 Fit of the $1/N$ Expansion

If the continuum values eq. (6.3) and eq. (6.4) are considered to be correct, then, together with the exactly known continuum for $N = \infty$, a prediction for the $N = 8$ continuum value can be gained. The three points are simply fitted to a second order polynomial in $1/N$.

$$c_0 + c_1 (1/N) + c_2 (1/N)^2 , \tag{6.11}$$

This ansatz is based on the fact, that the nonlinear sigma model can be expanded in $1/N$ (Section 5.1). The continuum value in the large $N$ limit for $z = 1.0595$ is

$$\sigma_z(2, 1.0595, \infty) \;=\; 1.37596138\ldots . \tag{6.12}$$

The computation of the continuum step scaling function is explained in Section 5.2.2.

Since the points that are fitted are exact values and do not have an statistical error, the parameter will not have an error neither. An estimate for the error of an interpolation to $N = 8$ would be the next order in $1/N$. It should be of $\mathcal{O}\left(1/N^3\right)$. So the result is shown in Fig. 6.13. The parameter are

$$c_0 = 1.375961 \qquad c_1 = -0.16193 \qquad c_2 = -0.54794 \tag{6.13}$$

and for the interpolation one gets

$$\sigma_z^{\text{fit}}(2, 1.0595, 8) = 1.347158 + \mathcal{O}\left(2 \cdot 10^{-3}\right) . \tag{6.14}$$

The error is one order of magnitude larger then the error of the extrapolation of the MC data (see e.g. Table 6.2). When compared to these values, the deviation is really found to be $\sim 2 \cdot 10^{-3}$. Therefore one cannot use this value to constrain the fits, like in the case of $N = 3, 4$.

Another approach would be to mutually fit all available data to an ansatz with form independent lattice artifacts, but the coefficient expanded in $1/N$:

$$c_0^{(N)} + c_1(1/N) \cdot f(a/L) , \tag{6.15}$$

with

$$c_1(1/N) = b_0 + b_1 \cdot (1/N) + b_2 \cdot (1/N)^2 + b_3 \cdot (1/N)^3 . \tag{6.16}$$

This assumes, of course, that the functional form of, at least the leading order, cutoff effects is the same for every order of $1/N$.

For every lattice spacing $i = a/L$ there will be a parameter $f_i = f(a/L)$. The parameter $b_0$ is fixed by the exact values for $N = \infty$

$$b_0 = \frac{1}{f(a/L_i)} \Big( \Sigma_z(2, 1.0595, \infty, a/L_i) - \sigma_z(2, 1.0595, \infty) \Big) \tag{6.17}$$

| | | | |
|---|---|---|---|
| $c_0^{(8)}$ | 1.34570(14) | $f_{10}$ | 2.598(12) |
| $b_1$ | 0.0148(33) | $f_{12}$ | 2.165(12) |
| $b_2$ | -0.160(26) | $f_{16}$ | 1.568(16) |
| $b_3$ | 0.525(49) | $f_{24}$ | 1.002(13) |
| $\chi^2/dof$ | 1.04 | $f_{32}$ | 0.696(13) |
| | | $f_{64}$ | 0.275(24) |
| | | $f_{96}$ | 0.04(17) |
| | | $f_{128}$ | 0.07(18) |

TABLE 6.5: Fit to the forms eq. (6.15) for $L/a = 10$.

and the $c_0^{(N)}$ are likewise known through the exact continuum values for $N = 3, 4, \infty$. The fit will yield an estimate for the continuum value $c_0^{(8)}$. Then, counting the lattice spacings with $L_i/a \geq 10$ in Table 6.1, there are 12 free parameters

$$c_0^{(8)}\,, b_1\,, b_2\,, b_3\,, f_i\,, \quad i = 10, 12, 16, 24, 32, 64, 96, 128 \tag{6.18}$$

and 22 sets of data, i.e. 10 degrees of freedom.

The fit can be split up in a linear and a nonlinear part, where the nonlinear part is solved by the least squares method. The result is listed in Table 6.5.

The two estimates for the continuum value can be plotted together with the MC data and other fits, in order to compare the results. In Fig. 6.14 one sees the large discrepancy of the continuum value obtained by a second order polynomial fit. The next order term cannot be neglected. The form independent fit takes this into account and yields a continuum value compatible to the MC data. Unfortunately the error is not improved.

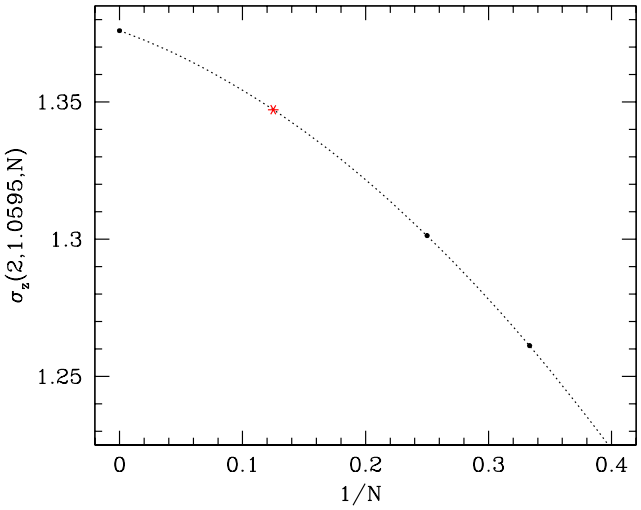



FIGURE 6.13: Second order polynomial fit of the continuum values at $N = 3, 4, \infty$ in $1/N$. The interpolation for $N = 8$ (star) looks good, but is not compatible to MC data.

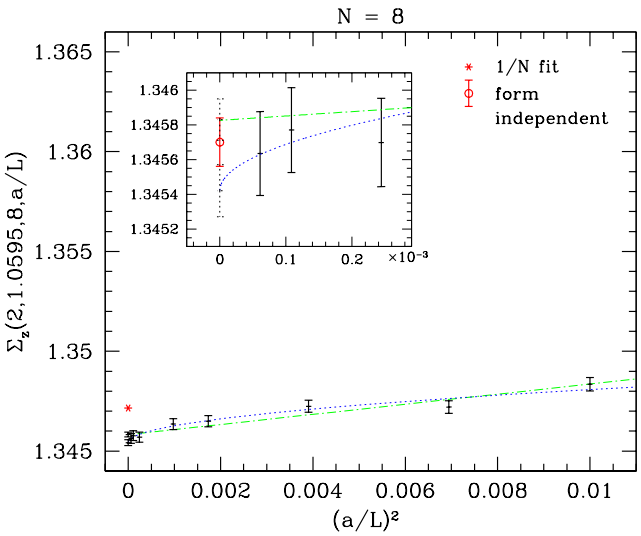



FIGURE 6.14: The $N = 8$ plot from Section 6.2.1 supplemented by the two new estimates for the continuum value. The second order polynomial fit in $1/N$ is definitely far away from the true value.

## 6.4 Conclusion

In the case $N = 3$ all proposed forms have to be rejected if one includes lattices $L/a \geq 10$ and assumes the theoretical predicted continuum value as correct. Of course, if one excludes enough of the small lattices, all forms become tolerable at some point. The other way round, one could say, that the theoretical prediction is not correctly, because the data cannot be fitted with this constraint.

Due to the smaller lattice artifacts at comparable statistical errors, the information from the same analysis, but for the $N = 4, 8$ MC data, is not so distinct. Since a linear fit is preferred for $N = 4$, there may be similar nonstandard cutoff effects as for $N = 3$. The fits nicely reproduce the theoretical prediction.

For $N = 8$ the lattice artifacts are even smaller. There is almost no information about the cutoff dependence. But one can determine a continuum extrapolation.

It remains unclear whether the $O(3)$ model is a specialty in this family of models. The lattice artifacts for $N = 4$ show a tendency similar to $N = 3$. It may be interesting to measure the step scaling function with a modified action, which has larger lattice artifacts. Then a form independent fit due to universality would be possible.

Also the next to leading order in the $1/N$ expansion could yield useful information about the approach to the continuum.

# Bibliography


[1] U. Wolff, "Collective Monte Carlo updating for spin systems," *Phys. Rev. Lett.* **62** (1989) 361.

[2] M. Hasenbusch, "An improved estimator for the correlation function of 2-D nonlinear sigma models," *Nucl. Phys. Proc. Suppl.* **42** (1995) 764–766, hep-lat/9408019.

[3] K. Symanzik, "Cutoff dependence in lattice $\phi_4^4$ theory," in *Recent developments in gauge theories (Cargése, 1979)*, G. 't Hooft et. al., ed. Plenum Press, New York, 1980.

[4] M. Lüscher, P. Weisz, and U. Wolff, "A Numerical method to compute the running coupling in asymptotically free theories," *Nucl. Phys.* **B359** (1991) 221–243.

[5] P. Hasenfratz and F. Niedermayer, "Unexpected results in asymptotically free quantum field theories," *Nucl. Phys.* **B596** (2001) 481–494, hep-lat/0006021.

[6] M. Hasenbusch, P. Hasenfratz, F. Niedermayer, B. Seefeld, and U. Wolff, "Nonstandard cutoff effects in the nonlinear sigma model," *Nucl. Phys. Proc. Suppl.* **106** (2002) 911–913, hep-lat/0110202.

[7] V. F. Müller, T. Raddatz, and W. Rühl, "The scaling behavior of the two-dimensional o(n) symmetric heisenberg model to the order n**-1," *Nucl. Phys.* **B251** (1985) 212.

[8] H. Flyvbjerg and S. Varsted, "1/n expansion of the nonlinear sigma model: The first three orders," *Nucl. Phys.* **B344** (1990) 646–664.

[9] M. Campostrini and P. Rossi, "The 1/n expansion of two-dimensional spin models," *Nucl. Phys. Proc. Suppl.* **34** (1994) 680–682, hep-lat/9305007.

[10] S. Caracciolo and A. Pelissetto, "Corrections to finite-size scaling in the lattice $N$-vector model for $N = \infty$," *Phys. Rev.* **D58** (1998) 105007, hep-lat/9804001.

[11] S. Caracciolo, A. Montanari, and A. Pelissetto, "Testing the efficiency of different improvement programs," *Nucl. Phys.* **B556** (1999) 295–326, hep-lat/9812014.

[12] J. Balog and A. Hegedus, "Tba equations for excited states in the o(3) and o(4) nonlinear sigma-model," hep-th/0309009.

[13] M. E. Peskin and D. V. Schroeder, *An Introduction to Quantum Field Theory*. Addison-Wesley Publishing Company, 1995.

[14] I. Montvay and G. Münster, *Quantum Fields on a Lattice*. Cambridge University Press, 1994.

[15] J. B. Kogut, "An introduction to lattice gauge theory and spin systems," *Rev. Mod. Phys.* **51** (1979) 659.

[16] K. Symanzik, "Continuum limit and improved action in lattice theories," *Nucl. Phys.* **B226** (1983) 187 and 205.

[17] **ALPHA** Collaboration, S. Capitani, M. Lüscher, R. Sommer, and H. Wittig, "Non-perturbative quark mass renormalization in quenched lattice QCD," *Nucl. Phys.* **B544** (1999) 669–698, hep-lat/9810063.

[18] J. Balog, "Kosterlitz-thouless theory and lattice artifacts," *J. Phys.* **A34** (2001) 5237–5250, hep-lat/0011078.

[19] T. Korzec, "Comparison of exact and numerical results in the XY model," Master's thesis, Institut für Physik, Humboldt-Universität zu Berlin, 2003. http://dochost.rz.hu-berlin.de.

[20] J. Balog, F. Knechtli, T. Korzec, and U. Wolff, "Numerical confirmation of analytic predictions for the finite volume mass gap of the XY-model," hep-lat/0309028.

[21] P. Weisz, "The universal z-z' curve in the $O(N)$ nonlinear $\sigma$-model in the limit $N \to \infty$." unpublished notes.

[22] M. Lüscher, "A new method to compute the spectrum of low lying states in massless asymptotically free field theories," *Phys. Lett.* **B118** (1982) 391–394.

[23] D.-S. Shin, "A determination of the mass gap in the o(n) sigma-model," *Nucl. Phys.* **B496** (1997) 408–434, hep-lat/9611006.

[24] B. Aebischer, "Berechnung der Massenlücke in asymptotisch freien Spinmodellen in zwei Dimensionen," Master's thesis, Institut für Theoretische Physik, Univerität Bern, 1984.

[25] U. Wolff and B. Bunk, "Computational physics II," 2002. Lecture given at Humboldt University Berlin.

[26] **ALPHA** Collaboration, U. Wolff, "Monte Carlo errors with less errors," hep-lat/0306017.

[27] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C*. Cambridge University Press, 1988.

[28] M. Hasenbusch, "Monte Carlo Simulationen in der statistischen Physik." Lecture given at Humboldt University Berlin.

[29] J. Zinn-Justin, *Quantum Field Theory and Critical Phenomena*. Oxford University Press Inc., New York, 1996.

[30] I. N. Bronstein and K. A. Semendjajew, *Taschenbuch der Physik*. B. G. Teubner Verlagsgesellschaft Leipzig, 1996.

[31] M. Abramowitz and I. A. Stegun, eds., *Handbook of mathematical functions*. Dover Publications, Inc., New York, 1965.

[32] J. Rolf, "Statistische und numerische Methoden der Datenanalyse I+II," 2003. Lecture given at Humboldt University Berlin http://www-com.physik.hu-berlin.de.

[33] F. Niedermayer, "Cluster algorithms," 1996. Lecture given at the summer school on 'Advances in Computer Simulations', Budapest, July 1996.

# Appendix A

# Lattice Notation

## A.1 Basic Definitions

All functions in a $d$-dimensional Euclidean space-time box are defined on a lattice $\Lambda$. The extension of the box in the $d-1$ spatial directions is labeled by $L_1, \ldots, L_{d-1}$ and in temporal direction by $L_d = T$. The lattice spacing is $a$. Then sites of the lattice are defined as

$$\Lambda = \{x \,|\, x_\mu/a \in \mathbb{Z},\, x_\mu < L_\mu\}, \quad \mu = 1, \ldots, d\,. \tag{A.1}$$

For simplicity assume the extension in all spatial directions equal $L$, then the total number of sites is

$$\Omega = \frac{1}{a^d}\, V \cdot T = \frac{1}{a^d}\, L^{d-1} \cdot T\,. \tag{A.2}$$

The space-time integral now is a sum over all lattice sites

$$\int \mathrm{d}^d x \to\ a^d \sum_{x\in\Lambda}\,. \tag{A.3}$$

Imposing periodic boundary conditions the allowed lattice momenta are those of the first Brillouin zone $\mathcal{B}$ defined by

$$p_\mu = \frac{2\pi}{L_\mu} n_\mu \quad , \quad n_\mu = 0, \ldots, L_\mu/a - 1 \tag{A.4}$$

and for finite lattices momentum integrals become summations, too,

$$\int \frac{\mathrm{d}^d p}{(2\pi)^d}\ \to\ \frac{1}{a^d\Omega} \sum_{p\in\mathcal{B}}, \tag{A.5}$$

whereas in the limit $L_1, \ldots, L_d \to \infty$ this sum becomes an integral over the Brillouin zone

$$\frac{1}{a^d\Omega} \sum_{p\in\mathcal{B}}\ \to\ \int_{-\pi/a}^{+\pi/a} \frac{\mathrm{d}^d p}{(2\pi)^d}\,. \tag{A.6}$$

The symbol $\int_{\mathcal{B}}$ is used to cover both cases and to indicate that in any explicit computation (A.5) or (A.6) has to be substituted.

*Derivatives:* There is an infinite number of possible definitions of lattice derivatives. Here the standard nearest neighbor ones are used ($\hat{\mu}$ is the unit vector of direction $\mu$):

- forward derivative

$$\Delta_\mu^f f(x) = \frac{1}{a}(f(x + a\hat{\mu}) - f(x)) \tag{A.7}$$

- backward derivative

$$\Delta^b_\mu f(x) = \frac{1}{a}(f(x) - f(x - a\hat{\mu})) \tag{A.8}$$

These operators possess the useful properties

$$a^d \sum_x \left(\Delta^f_\mu f(x)\right) g(x) = -a^d \sum_x f(x) \left(\Delta^b_\mu g(x)\right) \tag{A.9}$$

$$a^d \sum_{x,\mu} \Delta^f_\mu f(x)\, \Delta^f_\mu f(x) = -a^d \sum_{x,\mu} f(x) \left(\Delta^b_\mu \Delta^f_\mu f(x)\right) \equiv a^d \sum_x f(x)\, (\Box f(x)) \ , \tag{A.10}$$

where the symbol $\Box$ refers to the negative Laplace operator (sometimes called lattice d'Alembert operator) defined as

$$\Box = -\sum_\mu \Delta^b_\mu \Delta^f_\mu = -\Delta \tag{A.11}$$

$$\Box f(x) = \frac{1}{a^2} \sum_\mu (2f(x) - f(x + a\hat{\mu}) - f(x - a\hat{\mu}))\,. \tag{A.12}$$

*Matrix notation:* For some computations a matrix notation of the derivatives and the other operators is desirable. Forward and backward derivative can be written as

$$\Delta^f_{\mu,xy} = \frac{1}{a}(\delta_{x,y-a\hat{\mu}} - \delta_{x,y}) \tag{A.13}$$

$$\Delta^b_{\mu,xy} = \frac{1}{a}(\delta_{x,y} - \delta_{x,y+a\hat{\mu}})\,. \tag{A.14}$$

Now the property (A.9) becomes

$$\Delta^f_\mu = -(\Delta^b_\mu)^+ \tag{A.15}$$

and the Laplace operator has the matrix notation

$$-\Delta_{xy} = \Box_{xy} = \frac{1}{a^2} \sum_\mu (2\delta_{x,y} - \delta_{x,y+a\hat{\mu}} - \delta_{x,y-a\hat{\mu}})\,. \tag{A.16}$$

*Fourier transformations:* On a finite lattice the Fourier transformations read

$$F(x) = \frac{1}{\Omega} \sum_{p \in \mathcal{B}} \mathrm{e}^{ipx} \widetilde{F}(p) \tag{A.17}$$

$$\widetilde{F}(p) = \sum_x \mathrm{e}^{-ipx} F(x) \tag{A.18}$$

and instead of delta functions one has Kronecker symbols

$$\frac{1}{\Omega}\sum_x \mathrm{e}^{-i(p-p')x} = \delta_{p,p'} \tag{A.19}$$

$$\frac{1}{\Omega}\sum_{p\in\mathcal{B}} \mathrm{e}^{ip(x-x')} = \delta_{x,x'} \tag{A.20}$$

The matrix element of the Laplace operator (A.16) is only a function of $x-y$. Thus Fourier transformation into momentum space can be done

$$\begin{aligned}
\Box(p) &= \sum_{x-y} \mathrm{e}^{-ip(x-y)}\Box_{xy} \\
&= \frac{1}{a^2}\sum_\mu (2 - \mathrm{e}^{-ia\,p\cdot\hat{\mu}} - \mathrm{e}^{ia\,p\cdot\hat{\mu}}) \\
&= \frac{1}{a^2}\sum_\mu 2(1-\cos a\,p_\mu) \\
&= \frac{1}{a^2}\sum_\mu 4\sin^2\frac{a\,p_\mu}{2}\,.
\end{aligned} \tag{A.21}$$

## A.2 The $2d$ Nonlinear Sigma Model

There are many ways of writing down a lattice version of the continuum action

$$S = \frac{1}{2f}\int d^2x\;\partial_\mu s(x)\cdot\partial_\mu s(x)\,, \tag{A.22}$$

that has the proper naive continuum limit (Einstein summing conventions are used). Here only the standard nearest-neighbor action is used. That is, the above mentioned versions of the forward and backward derivatives are used to write:

$$\frac{1}{2f}\int d^2x\;\partial_\mu s(x)\cdot\partial_\mu s(x) \;\rightarrow\; \frac{a^2}{2f}\sum_x \Delta^f_\mu s(x)\,\Delta^f_\mu s(x)\,. \tag{A.23}$$

In matrix notation the lattice action then reads

$$S_{Lat} = \frac{a^2}{2f}\; s\cdot\Box s = \frac{a^2}{2f}\sum_{x,y} s(x)\Box_{xy}s(y)\,. \tag{A.24}$$

For the nonlinear sigma model this simplifies due to the symmetry

$$S_{Lat} = \frac{a^2}{2f}\sum_{x,y} s(x)\,\Box_{xy}s(y) \tag{A.25}$$

$$= \frac{1}{2f}\sum_{x,\mu} s(x)\;(2s(x) - s(x+a\hat{\mu}) - s(x-a\hat{\mu})) \tag{A.26}$$

$$= -\frac{1}{f}\sum_{x,\mu} s(x)\,s(x+a\hat{\mu}) + const.\,, \tag{A.27}$$

where $s^2(x) = 1$ and the translation invariance of the lattice sum was used. The expression in the last line is especially useful for lattice simulations. The constant term is irrelevant for the path integral. In analogy to the inverse temperature in thermodynamics the bare coupling $f$ is often replaced by $\beta = 1/f$. Finally the standard n.n. lattice action is

$$S_{Lat} = -\beta \sum_{x,\mu} s(x)\, s(x + a\hat{\mu})\,. \tag{A.28}$$

# Appendix B

# Monte Carlo

## B.1 Critical Slowing Down

- dynamical evolution of the system in MC time slows down considerably
- also occurs in real systems
- evolution is characterized by the *autocorrelation time*, i.e. the number of MC steps needed to generate a statistically independent new configuration
- autocorrelation ($A_t$ is the value of an observable $A$ at a given MC time $t$):
$$C_A(t) = \langle A_s A_{s+t} \rangle - \langle A \rangle^2$$
- for large $t$ autocorrelation decays exponentially, $C_A(t) \sim \exp(-t/\tau_{exp,A})$, where $\tau_{exp,A}$, the *exponential autocorrelation time*, corresponds to the slowest mode in the MC dynamics
- statistical error in MC:
usually $\langle A \rangle$ is estimated from the average of $n$ subsequent measurements, $\bar{A} = \frac{1}{n}\sum_{t=1}^{n} A_t$, and its statistical error is given by ($n \gg \tau_{int,A}$)
$$\delta\bar{A} = \sqrt{2\tau_{int,A}} \cdot \delta\bar{A}_{native}$$
where
$$\tau_{int,A} = \frac{1}{2}\sum_{t=-\infty}^{+\infty} C_A(t)/C_A(0)$$
is the *integrated autocorrelation time* and
$$\delta\bar{A}_{native} = \sqrt{\frac{1}{n}C_A(0)}$$
- this means that in order to reach a given accuracy one has to spend computer time $\sim \tau_{int}$
- autocorrelation time depends on the actual MC algorithm
- in general, near critical point ($\xi \gg 1$) the autocorrelation time diverges as
$$\tau \approx c\xi^z$$
where $z$ is the *dynamical critical exponent*
- Metropolis algorithm: $z \approx 2$, over-relaxation algorithm: down to $z \approx 1$
- collective, i.e. non-local updates: optimal cases $z \approx 0$, i.e. $\tau$ independent of $\xi$

## B.2 Multi Cluster Algorithm

This section is inspired by [33], an excellent review of cluster algorithms for classical and quantum spin systems.

The considerations below are restricted to the still quite general case of lattice actions containing interactions between variables $\phi_n$ belonging to only two different sites $n, n'$. These two sites are connected by a link $l$. For a nearest-neighbor interaction $l$ would label all nearest neighbors. This means models with a partition function

$$Z = \int D\phi \, \mathrm{e}^{-S[\phi]} \,, \tag{B.1}$$

are considered, where the lattice version of the action reads

$$S[\phi] = \sum_l S_l[\phi] = \sum_l S_l[\phi_n \in l] \,. \tag{B.2}$$

*Global symmetry* Suppose the interaction term $S_l[\phi]$ of the system is globally invariant to the symmetry group $G$

$$S_l[g\phi_n \in l] = S_l[\phi_n \in l], \qquad g \in G \,. \tag{B.3}$$

*Building the clusters* From the point of view of critical slowing down, an algorithm allowing to update large domains of the system in one step is desired. But with an a priori shape of such clusters (sites that are simultaneously updated) one introduces a bias and does not in general get the right weight function. Therefore the cluster are built dynamically activating some of the links. The probability $p_l[\phi]$ for a "bond" between sites belonging to a link $l$ is assumed to be globally invariant as well

$$p_l[g\phi_n \in l] = p_l[\phi_n \in l], \qquad g \in G \,. \tag{B.4}$$

The probability for a configuration of bonds $B$ given a configuration of the system $\phi$ on a lattice $\Lambda$ is then

$$\begin{aligned} w(B|\phi) &= (\text{prop. of bonds in } \Lambda) \times (\text{prob. of 'anti'-bonds in } \Lambda) \\ &= \prod_{l \in B} p_l[\phi] \prod_{l \notin B} (1 - p_l[\phi]) \,. \end{aligned} \tag{B.5}$$

Now a cluster is called a set of sites in $\Lambda$ which can be visited by going through bonds. So the whole lattice $\Lambda$ will split up in $N_C$ clusters of different shape and size and in the interspace, i.e. sites belonging to no cluster. This terminology is known from percolation theory [25].

*Updating the configuration - Ergodicity* Variables within the clusters are transformed simultaneously by stochastically chosen $g_i \in G,\ i = 1, \ldots, N_C$. The probability for the $g_i$'s should be independent of the cluster distribution and for detailed balance it should be symmetric under inversion

$$p(g_i) = p(g_i^{-1})\,. \tag{B.6}$$

Note that no further constraint is necessary at this point. Particularly $p(g_i)$ has not to be flat. But one important property is essential. Namely the transformations $g_i \in G$ have to assure the ergodicity of the algorithm. This means that every configuration $\phi'$ must be in principal reachable from any other configuration $\phi$ in a finite number of steps, i.e successive building of clusters and updating with sets of transformation $\{g_i\}$.

*Detailed balance* Let $w(\phi \to \phi')$ be the probability to reach configuration $\phi'$ from $\phi$ by one update. Then detailed balance means

$$\frac{w(\phi \to \phi')}{w(\phi' \to \phi)} = \frac{w(B|\phi) \cdot P(\{g_i\})}{w(B|\phi') \cdot P(\{g_i^{-1}\})} \stackrel{!}{=} \frac{\mathrm{e}^{-S[\phi]}}{\mathrm{e}^{-S[\phi']}}\,. \tag{B.7}$$

The probabilities for the set of transformation $P(\{g_i\}) = \prod_i p(g_i)$ cancel because of eq. (B.6). And because of the global symmetry of eqs. (B.3) and (B.4) all contributions from a link $l$ that is a bond, i.e all sites belong to the same cluster (same transformation), cancel. Then using eq. (B.5) the detailed balance condition eq. (B.7) becomes

$$\prod_{l \notin B} \frac{1 - p_l[\phi]}{1 - p_l[\phi']} = \prod_{l \notin B} \mathrm{e}^{-(S_l[\phi] - S_l[\phi'])}\,. \tag{B.8}$$

But still more factors cancel. Only those links survive which connect different domains of the lattice $\Lambda$, i.e. connecting different clusters or a cluster and the interspace of the clusters. Summarized, only those links give non-canceling contributions that are on the "surface" $\{\partial C\}$ of the clusters. This is because sites of a link lying completely inside the interspace of the clusters or completely in a cluster are equally transformed during the update $\phi \to \phi'$. There is no interaction within the clusters anymore. The original system has been mapped onto a system of $N_C$ "sites" represented by the clusters surfaces

$$\prod_{l \in \{\partial C\}} \frac{1 - p_l[\phi]}{1 - p_l[\phi']} = \prod_{l \in \{\partial C\}} \mathrm{e}^{-(S_l[\phi] - S_l[\phi'])}\,. \tag{B.9}$$

In order to arrive at a condition for the bond-probabilities $p_l[\phi]$ the following parameterization is assumed

$$p_l[\phi] = 1 - e^{-(Q_l[\phi] - S_l[\phi])}, \qquad p_l[\phi] \in [0, 1) \ \to\ Q_l[\phi] \geq S_l[\phi]\,. \tag{B.10}$$

Note that for $Q_l[\phi] = S_l[\phi]$ ($p_l[\phi] \equiv 0$) each site becomes a cluster and local MC updating is recovered. Inserting eq. (B.10) in eq. (B.9) leads to

$$\prod_{l\in\{\partial C\}} e^{-Q_l[\phi]} = \prod_{l\in\{\partial C\}} \mathrm{e}^{-Q_l[\phi']}\,. \tag{B.11}$$

*The choice of $Q_l$* Clearly, the structure of $Q_l[\phi]$ is not arbitrary. In order to provide a meaningful and efficient algorithm it should obey some conditions:

- eq. (B.4) and eq. (B.10) demand $Q_l$ to be globally invariant
- from eq. (B.11) detailed balance is valid for $Q_l[\phi] = Q_l[\phi']$, $l \in \{\partial C\}$
- $Q_l$ should be sensitive to the bare coupling and the structure of the action
- $Q_l - S_l$ should not be too large, otherwise the bond probability becomes too large as well and the largest cluster will usually occupy almost the whole lattice: but applying a global transformation on the whole lattice does not change anything

Now a possible choice of $Q_l$ is presented. The transformations $g_i$ for the update are restricted to a subgroup $H \subset G$. Then $Q_l[\phi]$ is defined as the maximum of the action term $S_l[\phi]$ when all elements $g \in H$ are independently applied to the sites of $l$

$$Q_l(\phi_n, \phi_{n'}) = \max_{g_n, g_{n'}\in H} S_l(g_n\phi_n, g_{n'}\phi_{n'}), \qquad H \subset G\,. \tag{B.12}$$

For this choice $Q_l$ is independent of arbitrary transformations at different sites belonging to $l$. This is because $Q_l(g_n\phi_n, g_{n'}\phi_{n'}) = Q_l(\phi_n, \phi_{n'})$ for any $g_n, g_{n'} \in H$ due to group properties. Thus detailed balance is valid. Updating the clusters this way automatically reproduces the wanted weights for the configurations. The clusters are built dynamically, i.e they are sensitive to the interaction and to the configuration.

## B.3 Single Cluster Algorithm

The idea of cluster algorithms is to update large domains of the system in one step. With an a priori shape of such clusters (sites that are simultaneously updated) one introduces a bias and does not in general get the right weight function. Therefore the cluster are built dynamically activating some of the links $l$.

With the multi cluster algorithm in Section B.2 one is still updating small clusters too often compared to large ones. By growing only one cluster from a randomly chosen "seed" one enhances the probability of updating large

clusters since a cluster of size $|C|$ in a "bond" distribution of a multi cluster algorithm is hit with the probability $|C|/V$.

Like in Section B.2 a model with a global symmetry $G$ and an action containing interactions over one link $l$ is considered (see eqs. (B.1)-(B.3)).

*Growing the cluster* First one initial site is randomly chosen and added to the cluster. Then all links originating form this site are activated with a probability $p_l[\phi]$. All sites reached by the activated links, i.e. "bonds", are added to the cluster as well. All links originating from the new sites are activated through the same procedure. This is repeated until no new sites are added to the cluster, i.e. all links originating from sites in the cluster were probed.

Thus the probability for a cluster $C$ with surface $\partial C$ given a configuration of the system $\phi$ is then

$$\begin{aligned} w(C|\phi) &= (\text{prop. of bonds in } C) \times (\text{prob. of 'anti'-bonds in } \partial C) \\ &= \prod_{l \in C} p_l[\phi] \prod_{l \in \partial C} (1 - p_l[\phi]) \,. \end{aligned} \tag{B.13}$$

*Updating $C$ - Ergodicity* Variables within the cluster are transformed simultaneously by stochastically chosen $g \in G$. The probability for $g$ should be symmetric under inversion (detailed balance)

$$p(g) = p(g^{-1}) \,. \tag{B.14}$$

Note that no further constraint is necessary at this point. Particularly $p(g)$ has not to be flat. Again the transformations $g \in G$ have to assure the ergodicity of the algorithm (see Section B.2).

*Detailed balance* Repeating the steps in Section B.2 one again ends up with a condition for the $Q_l$'s in the parameterization of the bond-probability $p_l[\phi]$ eq. (B.10)

$$\prod_{l \in \partial C} e^{-Q_l[\phi]} = \prod_{l \in \partial C} \mathrm{e}^{-Q_l[\phi']} \,. \tag{B.15}$$

*The choice of $Q_l$* For the single cluster algorithm a rather simple choice in line with the conditions at the end of Section B.2 is possible. Suppose the cluster is to be updated by a transformation $g$ and one has to probe whether the link $l$ should be activated. Without loss of generality let $l = (n_1, n_2)$ and $n_1 \in C$, $n_2 \notin C$, then define

$$Q_l[\phi] \equiv Q_l(\phi_{n_1}, \phi_{n_2}) = S_l(g \cdot \phi_{n_1}, \phi_{n_2}) \,. \tag{B.16}$$

With this the bond-probability eq. (B.10) becomes

$$p_l[\phi] = 1 - e^{-(S_l(g \cdot \phi_{n_1}, \phi_{n_2}) - S_l(\phi_{n_1}, \phi_{n_2})} \,. \tag{B.17}$$

Sites connected by the link $l$ being in a state where they are invariant under $g \in G$ are therefore preferably put together in a cluster. Now, again without loss of generality, let this link remain inactivated and hence show up in eq. (B.15). Thus for the inverse transition $\phi' \to \phi$ one gets (remember $n_1 \in C$, $n_2 \notin C$)

$$Q_l[\phi'] \equiv Q_l(g^{-1} \cdot \phi'_{n_1}, \phi'_{n_2}) = Q_l(\phi_{n_1}, \phi_{n_2})\,, \tag{B.18}$$

and detailed balance is achieved.

### B.3.1 Single Cluster Algorithm for Nonlinear Sigma Models

In the Ising model with its simple $Z(2)$ symmetry the only possible transformation is the spin-flip. There, an implementation of the single cluster algorithm described above would be straightforward. Maintaining the idea of spin-flip when moving to an $O(N)$ symmetry requires a generalization. In [1] this is achieved by embedding an Ising model into the $N$-vector model. One randomly chooses a direction that defines a $N-1$ dimensional hyperplane dividing the $N-1$ sphere into two hemispheres. Vectors lying in the one hemisphere are regarded as spin up and the others as spin down. A spin-flip is then a reflection with respect to the hyperplane.

To make the calculation explicit the variables at each lattice site are $O(N)$-vectors $\vec{s}$ and the action consists of nearest-neighbor products

$$S = -\beta \sum_{\langle xy \rangle} \vec{s}_x \cdot \vec{s}_y\,. \tag{B.19}$$

The set of reflections with respect to a hyperplane orthogonal to a unit vector $\vec{r} \in S_{N-1}$ is a subgroup of $O(N)$ containing all rotations by an angle $\pi$. The axis of this rotation is defined by the projection of the vector to be reflected into the hyperplane. Ergodicity is guaranteed since there is always a reflection connecting any two $O(N)$-vectors. The reflections are defined as

$$R(\vec{r})\,\vec{s} = \vec{s} - 2(\vec{s}\cdot\vec{r})\,\vec{r}, \qquad R(\vec{r})^2 = 1\,. \tag{B.20}$$

Summarizing everything the algorithm finally reads [1]:

1. Choose a random reflection $\vec{r} \in S_{N-1}$ and a random lattice site $x$ as the starting point of the cluster $C$

2. Grow the cluster $C$ as described above; the bond-probability is

$$Q_{\langle xy \rangle} = -\beta\, R(\vec{r})\,\vec{s}_x \cdot \vec{s}_y = -\beta\,\vec{s}_x \cdot R(\vec{r})\,\vec{s}_y \tag{B.21}$$

$$\to \quad p_{\langle xy \rangle} = 1 - \exp\{\min(0, \beta\,\vec{s}_x \cdot [1 - R(\vec{r})]\,\vec{s}_y)\} \tag{B.22}$$

3. Reflect all vectors in $C$

How has one to choose the random vector $\vec{r} \in S_{N-1}$ in order to guarantee ergodicity and detailed balance? Since the reflections $R(\vec{r})$ are idempotent the restriction (B.14) is not important here ($p(R^{-1}) = p(R)$). The only requirement is ergodicity, well, and efficiency. One can intuitively say that a uniformly over $S_{N-1}$ distributed $\vec{r}$ is the best choice. But it would not destroy the validity of the algorithm if the distribution were not uniform. The main thing is that it covers the whole $S_{N-1}$ in order to achieve ergodicity.

# Appendix C

# Tables

TABLE C.1: Time slice correlation function $D(t/a)$ and mass $m(t/a) \equiv \ln[D(t/a)/D(t/a+1)]$ measured with three different estimators in the $O(3)$ model at $\beta = 1.6982$ on a $20 \times 100$ lattice.

| $t/a$ | standard | | simple | | Metropolis, $M = 10$ | |
|---|---|---|---|---|---|---|
| | $D(t/a) \cdot 10^{-1}$ | $m(t/a)$ | $D(t/a) \cdot 10^{-1}$ | $m(t/a)$ | $D(t/a) \cdot 10^{-1}$ | $m(t/a)$ |
| 1 | 4.1365(22) | 1.3521(23) | 4.1364(21) | 1.3514(14) | 4.1382(22) | 1.3514(11) |
| 2 | 3.8661(22) | 1.3174(28) | 3.8662(21) | 1.3151(14) | 3.8678(22) | 1.3146(11) |
| 3 | 3.6196(23) | 1.2980(31) | 3.6201(21) | 1.2983(15) | 3.6217(21) | 1.2981(11) |
| 4 | 3.3922(24) | 1.2883(34) | 3.3926(21) | 1.2895(15) | 3.3941(21) | 1.2893(11) |
| 5 | 3.1805(25) | 1.2857(37) | 3.1807(22) | 1.2848(15) | 3.1822(21) | 1.2859(11) |
| 6 | 2.9825(27) | 1.2827(40) | 2.9828(22) | 1.2829(15) | 2.9840(21) | 1.2832(11) |
| 7 | 2.7972(28) | 1.2827(42) | 2.7975(22) | 1.2815(15) | 2.7986(20) | 1.2810(11) |
| 8 | 2.6234(30) | 1.2834(45) | 2.6238(22) | 1.2804(15) | 2.6249(20) | 1.2804(11) |
| 9 | 2.4604(32) | 1.2822(47) | 2.4611(22) | 1.2813(15) | 2.4622(20) | 1.2793(11) |
| 10 | 2.3076(33) | 1.2800(50) | 2.3084(21) | 1.2815(16) | 2.3096(19) | 1.2795(11) |
| 11 | 2.1645(35) | 1.2781(53) | 2.1651(21) | 1.2798(16) | 2.1664(19) | 1.2789(11) |
| 12 | 2.0305(36) | 1.2711(55) | 2.0309(21) | 1.2793(16) | 2.0322(18) | 1.2789(12) |
| 13 | 1.9055(38) | 1.2690(58) | 1.9050(21) | 1.2802(16) | 1.9063(18) | 1.2774(12) |
| 14 | 1.7883(39) | 1.2727(63) | 1.7869(20) | 1.2795(16) | 1.7884(18) | 1.2789(12) |
| 15 | 1.6780(40) | 1.2744(66) | 1.6762(20) | 1.2797(16) | 1.6776(17) | 1.2784(12) |
| 16 | 1.5744(41) | 1.2739(70) | 1.5723(20) | 1.2805(16) | 1.5737(17) | 1.2777(12) |
| 17 | 1.4773(42) | 1.2686(74) | 1.4747(19) | 1.2801(16) | 1.4763(16) | 1.2786(12) |
| 18 | 1.3865(42) | 1.2762(79) | 1.3833(19) | 1.2804(16) | 1.3849(16) | 1.2786(12) |
| 19 | 1.3008(43) | - | 1.2975(18) | - | 1.2991(15) | - |

TABLE C.2: Numerical costs $M_c$, time spent in the measuring routine $t_m$ and error of the running coupling $\Delta z$ ($z = m(L)\,L$) measured in the $O(4)$ model for $\beta = 2.4765$ on a $20 \times 100$ lattice ($z = 1.30887(23)$) for the Metropolis improved estimator with different number of updates, the simple improved estimator and the standard one.

| $M$ | $M_c$ in msec | $t_m$ in msec | $\Delta z$ |
|---|---|---|---|
| 2 | 0.1775(11) | 1.5734(94) | 2.872(93) |
| 4 | 0.1444(21) | 1.808(34) | 2.525(77) |
| 6 | 0.1429(17) | 2.124(28) | 2.407(74) |
| 8 | 0.1430(22) | 2.454(48) | 2.329(71) |
| 10 | 0.1550(14) | 2.765(26) | 2.325(71) |
| 12 | 0.1484(32) | 2.852(85) | 2.295(70) |
| 15 | 0.1608(90) | 3.523(10) | 2.203(67) |
| 20 | 0.1740(29) | 4.169(89) | 2.180(70) |
| simple | 0.2588(69) | 1.124(35) | 3.77(12) |
| standard | 1.858(47) | 0.0490(31) | 12.35(29) |

TABLE C.3: Functions $K(a/L)$ and $\tilde{a}_2(a/L)$

| L | $K(a/L)$ | | $\tilde{a}_2(a/L)$ | |
|---|---|---|---|---|
| | exact | expansion | exact | expansion |
| 10 | 0.14212817 | 0.14212181 | -0.68713606 | -0.68760605 |
| 15 | 0.14190715 | 0.14190594 | -0.64823196 | -0.64840023 |
| 20 | 0.14183076 | 0.14183038 | -0.63113786 | -0.63122422 |
| 25 | 0.14179557 | 0.14179541 | -0.62206283 | -0.62211562 |
| 30 | 0.14177649 | 0.14177641 | -0.61663629 | -0.61667201 |
| 35 | 0.14176500 | 0.14176496 | -0.61311684 | -0.61314266 |
| 40 | 0.14175755 | 0.14175753 | -0.61069568 | -0.61071524 |
| 45 | 0.14175244 | 0.14175243 | -0.60895390 | -0.60896924 |
| 50 | 0.14174879 | 0.14174878 | -0.60765610 | -0.60766847 |
| 55 | 0.14174609 | 0.14174608 | -0.60666137 | -0.60667155 |
| 60 | 0.14174404 | 0.14174403 | -0.60588097 | -0.60588950 |
| 65 | 0.14174244 | 0.14174244 | -0.60525665 | -0.60526390 |
| 70 | 0.14174117 | 0.14174117 | -0.60474884 | -0.60475508 |
| 75 | 0.14174015 | 0.14174015 | -0.60432986 | -0.60433529 |
| 80 | 0.14173931 | 0.14173931 | -0.60397985 | -0.60398461 |
| 85 | 0.14173862 | 0.14173862 | -0.60368425 | -0.60368847 |
| 90 | 0.14173804 | 0.14173804 | -0.60343218 | -0.60343594 |
| 95 | 0.14173755 | 0.14173755 | -0.60321538 | -0.60321876 |
| 100 | 0.14173713 | 0.14173712 | -0.60302747 | -0.60303052 |

TABLE C.4: Continuum step scaling function

| $u$ | $\sigma_z(2,u,\infty)$ | | |
|---|---|---|---|
| | exact | up to $d_3$ | up to $d_7$ |
| 0.1 | 0.10225611 | 0.10225611 | 0.10225611 |
| 0.2 | 0.20923204 | 0.20923212 | 0.20923204 |
| 0.3 | 0.32125815 | 0.32125913 | 0.32125815 |
| 0.4 | 0.43868870 | 0.43869454 | 0.43868870 |
| 0.5 | 0.56189877 | 0.56192242 | 0.56189875 |
| 0.6 | 0.69127872 | 0.69135368 | 0.69127861 |
| 0.7 | 0.82722591 | 0.82742640 | 0.82722534 |
| 0.8 | 0.97013262 | 0.97060612 | 0.97013018 |
| 0.9 | 1.12036990 | 1.12138609 | 1.12036101 |
| 1.0 | 1.27826684 | 1.28028757 | 1.27823842 |
| 1.1 | 1.44408573 | 1.44786007 | 1.44400396 |
| 1.2 | 1.61799409 | 1.62468169 | 1.61777894 |
| 1.3 | 1.80003619 | 1.81135933 | 1.79951171 |
| 1.4 | 1.99010722 | 2.00852901 | 1.98891136 |
| 1.5 | 2.18793492 | 2.21685614 | 2.18536532 |
| 1.6 | 2.39307281 | 2.43703579 | 2.38783805 |
| 1.7 | 2.60490853 | 2.66979297 | 2.59474792 |
| 1.8 | 2.82268795 | 2.91588290 | 2.80381875 |
| 1.9 | 3.04555307 | 3.17609133 | 3.01190246 |
| 2.0 | 3.27258847 | 3.45123474 | 3.21476857 |
| 2.1 | 3.50286978 | 3.74216069 | 3.40685643 |
| 2.2 | 3.73550781 | 4.04974807 | 3.58098500 |
| 2.3 | 3.96968350 | 4.37490736 | 3.72801537 |
| 2.4 | 4.20467159 | 4.71858093 | 3.83646004 |
| 2.5 | 4.43985275 | 5.08174335 | 3.89203314 |

# Selbstständigkeitserklärung

Hiermit erkläre ich, die vorliegende Arbeit selbständig ohne fremde Hilfe verfaßt und nur die angegebene Literatur und Hilfsmittel verwendet zu haben.

Björn Leder
10. Oktober 2003