Dynamical Chirally Improved Fermions in Lattice QCD

Dissertation

zur Erlangung des akademischen Grades
Doktor der Naturwissenschaften.

Angefertigt am Institut für Physik,
Karl-Franzens-Universität Graz.

Eingereicht von:
Wolfgang Ortner

Betreuung:
Univ.Prof. Dr. C.B. Lang

Präambel

Ich, Wolfgang Ortner, bestätige, dass es sich bei der hier vorgelegten Dissertation um eine Originalarbeit handelt, die von mir selbstständig angefertigt und abgefasst wurde.

Summary

This thesis resulted from work on a project with the goal to simulate lattice QCD with light dynamical quarks. To achieve this goal we employ so-called chirally improved fermions.

Quantum Chromodynamics (QCD) is the quantum field theory of strong interactions, formulated in terms of quarks and gluons. Lattice QCD is QCD formulated on a four-dimensional euclidian space-time lattice. It provides a non-perturbative regularization scheme of QCD. Apart from regularizing QCD the lattice also provides a way of putting the theory on a computer and simulating it.

Chiral symmetry is an approximate symmetry of the light quark flavors of QCD. It would be exact for massless quarks. Despite this small explicit breaking by the quark masses its effects can clearly be seen in the hadron spectrum. The implementation of chiral symmetry has been a longstanding problem in lattice QCD. It has been partly resolved in a negative way by Nielsen and Ninomiya in 1981. In their famous no-go theorem they proved that it is impossible to have exact chiral symmetry in a formulation of QCD on a finite lattice that is local and which has the correct number of flavors. To be more precise in this last point, in such a local and chiral formulation of lattice QCD the quarks would come in groups of 16 for every different quark mass one wants to introduce. This is known as the *fermion doubling problem*, of course it is clearly unacceptable if one wants to describe nature.

In 1982 Ginsparg and Wilson suggested a relation which is now known as the Ginsparg-Wilson relation. This relation defines theories, in which the chiral symmetry is broken in a local way only. At the time of its finding the Ginsparg-Wilson relation did not have a great impact because nobody knew how to construct a theory which fulfills the relation. Only during the 1990-s such theories started to appear. The *chirally improved* action is a lattice discretization of the Dirac operator which fulfills the Ginsparg-Wilson relation approximately. This action has been successfully applied in various simulations over the last six years. These simulations, however, have been neglecting the vacuum polarization effects from dynamical quark loops.

Simulating lattice QCD means solving high-dimensional integrals. Therefore Markov Chain Monte Carlo methods are employed. Dynamical quark loops are often neglected because including them renders simulations much more expensive. This is especially true if one wants to consider light quarks. Simulations with small quark masses always are more expensive than simulations with large quark masses, no matter what lattice action is used. Apart from other problems, however, simulations with light quarks that use a non Ginsparg-Wilson action are plagued by the breaking of chiral symmetry. This can be at least partly resolved using Ginsparg-Wilson actions. At quark masses smaller than half the strange quark mass, or 40 MeV, it can be expected that simulations will substantially profit from using Ginsparg-Wilson actions.

Until know there exist only few simulations using Ginsparg-Wilson actions which include these quark loop effects, and currently this is a topic of active research by different groups in the field. Among these efforts is a project of C.B. Lang, Pushan Majumdar and myself, which has been started in 2003 and to which I have been contributing from the beginning. This project is the topic of my thesis.

Chapter 1 of the thesis introduces the main ideas behind the project. Chapters 2 to 5 of this thesis contain a review of some properties of lattice QCD and Markov Chain Monte Carlo simulations which are important in the context of the project. Chapter 2, titled “Lattice QCD”, contains some basic facts of lattice QCD in general. Since chiral symmetry in lattice QCD is a main topic of this work, it is discussed separately in Chapter 3, “Chiral Symmetry in Lattice QCD”. The next chapter, Chapter 4, “Markov Chain Monte Carlo”, describes besides some standard techniques of Markov Chain Monte Carlo the Hybrid Monte Carlo algorithm. In Chapter 5, “Lattice QCD Simulations”, simulations of lattice QCD, using the methods described in Chapter 4, are discussed. The two main chapters of the thesis, describing the project itself, are Chapters 6 and 7.

In Chapter 6, titled “Hybrid Monte Carlo for Generalized Dirac Operators” the techniques of employing the Hybrid Monte Carlo algorithm for Dirac operators like the chirally improved operator are developed. The Hybrid Monte Carlo algorithm is the most successful algorithm in dynamical simulations using various actions. However, employing this for an action like the chirally improved action poses serious technical problems. These problems are serious enough that in order to avoid them a a special algorithm called Partial-Global algorithm has been suggested. However, the performance we were able to achieve with that approach was very dissatisfying. For that reason we decided to employ the Hybrid-Monte-Carlo algorithm. Therefore we had to solve the technical problems mentioned above.

One important conclusion we can draw from this work is that it is possible to employ the HMC algorithm for an action like the chirally improved action. Chapter 6 of my thesis describes the various aspects of implementing HMC for an action like that. We found a relatively elegant way of solving the technical problems involved, which we describe in detail Chapter 6 of the thesis. We think that the methods developed there could be useful for for simulations not only with the chirally improved action, but with other, similar, actions too.

Chapter 7 is called "Simulating QCD with two Dynamical Flavors of CI Fermions", and discusses results from using them methods of Chapter 6 in practical simulations. After some initial studies on smaller lattices, we simulated lattice QCD with two dynamical flavors of mass-degenerate chirally improved quarks on lattices with $12^3 \times 24$ lattice sites (12 lattice points in the three space dimensions, and 24 in the time direction). These early calculations have been restricted due computer resources. Again, our big goal was to simulate light dynamical quarks. Therefore our discussion in Chapter 7. of the thesis put a strong emphasis on quark and pion masses. On this lattice size and with the computer resources at hand we reached quark masses down to about 30 MeV which corresponded to a pion mass of about 460 MeV in our calculations. Our calculations seem to indicate that quark masses of about 20 MeV can be reached on lattices of this size. It can be expected that the quark masses can be further reduced in simulations on larger lattices.

Contents

Chapter 1

Introduction

Quantum Chromodynamics (QCD) is the quantum field theory of strong interactions, formulated in terms of quarks and gluons [1]. Together with perturbative methods it has been very successful in predicting phenomena at small distances, where the coupling constant of QCD is small. These perturbative methods start from a free theory and treat the coupling as a small perturbation to this. Therefore they can only work at small coupling. At the scale of the hadronic world (about 1 fm), the coupling constant of QCD is too large for such perturbative expansions to work.

Lattice QCD (LQCD) was introduced by Wilson in 1974 [2]. LQCD is QCD formulated on a four-dimensional euclidian space-time lattice. It provides a non-perturbative regularization scheme of QCD: On the finite grid the infinities, which occur in the continuum, are removed. Furthermore, physical quantities appear to have a finite, well behaved limit when the lattice spacing is taken to 0, the continuum limit. Until today it is the only known gauge invariant regularization method which allows non perturbative solution approaches to QCD. All other known regularization schemes are tied closely to the perturbative expansions mentioned above.

Apart from regularizing QCD the lattice also provides a way of putting the theory on a computer and simulating it. However, such simulations turned out to be computationally very demanding. Therefore in the last thirty years a lot of effort has been invested into improvements of simulation algorithms and the way QCD is discretized.

Whether a lattice discretization is good or bad depends crucially on the way it deals with the symmetries of QCD. The breaking of symmetries when discretizing is unavoidable in many cases. This can be readily seen if one considers continuous rotations, which simply cannot exist on a discrete grid. Breaking of symmetries is acceptable, as long as the symmetries are restored in the continuum limit. In this case the breaking can be seen as a discretization error which can be controlled by making the lattice fine enough. How fine it has to be depends on the discretization. From a pragmatical point of view we could say that one discretization is better than another one, if it allows for coarser lattices.

Wilson's proposal for a lattice discretization [2] was formulated in terms of very short range interactions. This provides a simple and elegant formulation of lattice QCD. This discretization preserves gauge invariance. On the other hand it breaks the chiral symmetry of lattice QCD.

Chiral symmetry is an approximate symmetry of the light quark flavors of QCD. It would be exact for massless quarks. Despite this small explicit breaking by the quark masses its effects can clearly be seen in the hadron spectrum.

The implementation of chiral symmetry has been a longstanding problem in lattice QCD. It has been partly resolved in a negative way by Nielsen and Ninomiya in 1981 [3]. In their famous no-go theorem they proved that it is impossible to have exact chiral symmetry in a formulation of QCD on a finite lattice that is local and which has the correct number of flavors. To be more precise in this last point, in such a local and chiral formulation of lattice QCD the quarks would come in groups of 16 for every different quark mass one wants to introduce. This is known as the *fermion doubling problem*, of course it is clearly unacceptable if one wants to describe nature.

In 1982 Ginsparg and Wilson suggested a relation which is now known as the Ginsparg-Wilson relation [4]. This relation defines theories, in which the chiral symmetry is broken in a local way only. At the time of its finding the Ginsparg-Wilson relation did not have a great impact because nobody knew how to construct a theory which fulfills the relation. Only during the 1990-s such theories started to appear [5, 6, 7]. The *chirally improved* [8, 9] action is a lattice discretization of the Dirac operator which fulfills the Ginsparg-

Wilson relation approximately. This action has been successfully applied in various simulations over the last six years. These simulations, however, have been neglecting the vacuum polarization effects from dynamical quark loops.

Dynamical quark loops are often neglected because including them renders simulations much more expensive. This is especially true if one wants to consider light quarks. Simulations with small quark masses always are more expensive than simulations with large quark masses, no matter what lattice action is used. Apart from other problems, however, simulations with light quarks that use a non Ginsparg-Wilson action are plagued by the breaking of chiral symmetry. This can be at least partly resolved using Ginsparg-Wilson actions. At quark masses smaller than half the strange quark mass, or 40 MeV, it can be expected that simulations will substantially profit from using Ginsparg-Wilson actions.

Until know there exist only very few simulations using Ginsparg-Wilson actions which include these quark loop effects, and currently this is a topic of active research by different groups in the field. Among these efforts is a project of C.B. Lang, Pushan Majumdar and myself [10, 11, 12, 13], which has been started in 2003 and to which I have been contributing from the beginning. This project is the topic of my thesis.

Chapter 2

Lattice QCD

2.1 Euclidian QCD at the Classical Level

Definitions

We will use the euclidian formulation of quantum field theory exclusively. The action of QCD in its euclidian formulation is defined on 4-dimensional euclidian space:

$$S[A,\overline{\psi},\psi] = \int d^4x \; \mathcal{L} \; , \tag{2.1.1}$$

where the *Lagrangian Density* is defined as

$$\mathcal{L} = \frac{1}{2g^2} \; \mathrm{tr} \; F_{\mu\nu} F_{\mu\nu} + \overline{\psi}(\not{D} + m)\psi \; , \tag{2.1.2}$$

with

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu] \; , \tag{2.1.3}$$

$$\not{D} = \gamma_\mu(\partial_\mu + igA_\mu) \; . \tag{2.1.4}$$

The A field is a non-abelian gauge field $A_\mu \in su(3)$ (i.e., the algebra of the gauge group $SU(3)$). The fermionic Dirac spinors $\overline{\psi}$ and ψ, called anti-quark field and quark field respectively. They carry three indices, the 3-component color index, the 4-component Dirac index and a flavor index. In nature there are 6 different flavors called up, down, strange, charm, bottom and top (lattice QCD calculations are often restricted to less flavors, however). The mass matrix m is a diagonal matrix with the masses of the different flavors appearing in the diagonal. The euclidian γ matrices are hermitian and obey

$$\begin{aligned} \{\gamma_\mu, \gamma_\nu\} &= 2\delta_{\mu,\nu} \ , \\ \{\gamma_\mu, \gamma_5\} &= 0 \ , \\ \gamma_5 &= \gamma_1\gamma_2\gamma_3\gamma_4 \ . \end{aligned} \tag{2.1.5}$$

Symmetries

The action defined above is invariant under 4-dimensional translations. It is invariant under $O(4)$, the euclidian equivalent of the full Lorentz group in Minkowski space. There is also an euclidian version of charge conjugation which leaves the action invariant. We will not discuss the transformation properties of the gauge and quark fields under these transformations here, they can of course be found in text books, e.g., [14].

What we want to discuss in a bit more detail are two other symmetries: The action of QCD is invariant under a local $SU(3)$ transformation. This is called a *gauge transformation*; we denote it by $\Omega(x) \in SU(3)$. Under this the gauge fields transform as

$$A'(x)_\mu = \Omega(x)(A(x)_\mu + \partial_\mu)\Omega(x)^\dagger \ , \tag{2.1.6}$$

and the anti-quark and quark fields transform as

$$\overline{\psi}'(x) = \overline{\psi}(x)\Omega(x)^\dagger, \ \psi'(x) = \Omega(x)\psi(x) \ . \tag{2.1.7}$$

We define the chiral projectors as

$$P_L = \tfrac{1}{2}(1-\gamma_5) \ , \ P_R = \tfrac{1}{2}(1+\gamma_5) \ , \tag{2.1.8}$$

and use this to define left-handed and right-handed quark and anti-quark fields as

$$\begin{aligned} \psi_L &= P_L\psi \ , \ \psi_R = P_R\psi \ , \\ \overline{\psi}_L &= \overline{\psi}P_R \ , \ \overline{\psi}_R = \overline{\psi}P_L \ . \end{aligned} \tag{2.1.9}$$

Inserting these into the QCD action with $m = 0$ we get

$$S[A, \overline{\psi}, \psi] = \int d^4x \ \frac{1}{2g^2} \ \mathrm{tr} \ F_{\mu\nu}F_{\mu\nu} + \overline{\psi}_L \not{D} \psi_L + \overline{\psi}_R \not{D} \psi_R \ . \tag{2.1.10}$$

In the massless case therefore the left and right handed anti-quark and quark fields decouple, and as a result the action possesses an $U(N_f)_L \otimes U(N_f)_R$ symmetry, called *chiral symmetry.* Chiral symmetry manifests itself in the Dirac matrix $\not{D}$ as

$$\{\not{D}\,, \gamma_5\} = 0 \ , \tag{2.1.11}$$

which is a necessary condition for the symmetry. The masses of the up and down quarks are quite small compared to the scale of spontaneous chirally symmetry breaking in QCD of about 1 GeV. That is why one sees strong effects of a spontaneously broken $U(2)_L \otimes U(2)_R$.

2.2 Quantization via Path Integrals

To quantize the classical theory discussed in the last Section the method of *Feynman path integration* can be used. This framework allows to formulate the quantized theory in terms of a system of classical statistical mechanics. In the rest of this Section a few basic facts of this method are summarized.

We are interested in vacuum expectation values of time ordered products of certain operators

$$\langle 0|T\{\hat{O}_1(x_1), \hat{O}_2(x_2), \dots\}|0\rangle \ . \tag{2.2.1}$$

Using Feynman path integrals we can formally express our vacuum expectation values in terms of expectation values in the framework of classical statistical mechanics

$$\begin{aligned} &\langle 0|T\{\hat{O}_1(x_1), \hat{O}_2(x_2), \dots\}|0\rangle = \\ &\frac{1}{Z}\int \mathrm{d}[A, \overline{\psi}, \psi]\ O_1(x_1)O_2(x_2)\dots\exp(-S) \ , \end{aligned} \tag{2.2.2}$$

where

$$Z = \int \mathrm{d}[A, \bar{\psi}, \psi]\ \ \exp(-S) \ . \tag{2.2.3}$$

This function Z we call *partition function* in analogy to statistical mechanics. The definitions above are only formal at this point since we have not yet defined the integration measure.

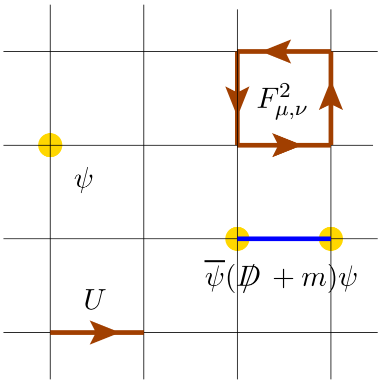

Figure 2.1: Illustration of the lattice discretization of QCD

2.3 Wilson's Formulation of Lattice QCD

To give the path integrals of the last Section a proper definition the procedure of lattice regularization can be used: The action gets discretized on a four dimensional hypercubic lattice. We set the lattice spacing $a = 1$, and use the notation $\hat{\mu}$ for the identity vector in the direction μ. On this lattice the lattice gauge field is defined as a set of $SU(3)$ matrices connecting two neighboring lattice sites, which are called link variables. The link variable between two sites x and $x + \hat{\mu}$ we denote as $U_{x,\mu}$. The discretization of the quark and anti-quark fields introduces Grassmann variables, which live on the lattice sites. We will denote them as $\psi(x)$ and $\overline{\psi}(x)$ just like their continuous counterparts. This discretization procedure is illustrated in Fig. 2.1 The historically first lattice discretization of the QCD action, and one of the simplest, is called Wilson's action [2, 15], named after its inventor.

We split the Lagrangian (2.1.2) into one part containing only the gauge fields $\mathcal{L}_g$ and another part containing gauge fields and quark (fermionic) fields $\mathcal{L}_f$.

$$\begin{aligned} \mathcal{L} &= \mathcal{L}_g + \mathcal{L}_f \ , \\ \mathcal{L}_g &= \frac{1}{2g^2} \operatorname{tr} F_{\mu\nu} F_{\mu\nu} \ , \\ \mathcal{L}_f &= \overline{\psi}(\not{D} + m)\psi \ . \end{aligned} \tag{2.3.1}$$

Using this the we split the action into

$$\begin{aligned} S &= S_g + S_f \ , \\ S_g &= \int d^4x \ \mathcal{L}_g \, , \ S_f = \int d^4x \ \mathcal{L}_f \ . \end{aligned} \tag{2.3.2}$$

The Gauge Part

In Wilson's approach the discretization of the gauge part of the action S_g is done using only the smallest closed loops on the lattice, the so-called *plaquettes*. They consist of four link variables between the four neighboring points connecting them, e.g., in the following way:

$$x \rightarrow x + \hat{\mu} \rightarrow x + \hat{\mu} + \hat{\nu} \rightarrow x + \hat{\nu} \rightarrow x \ . \tag{2.3.3}$$

The parallel transporter around this plaquette is

$$U^{pl}_{x,\mu\nu} = U_{x,\mu} U_{x+\hat{\mu},\nu} U^{\dagger}_{x+\hat{\nu},\mu} U^{\dagger}_{x,\nu} \ . \tag{2.3.4}$$

Wilson's gauge action (discretization of S_g) is defined as

$$\begin{aligned} S_g[U] &= \beta_{\mathrm{W}} \sum_x \sum_{\mu<\nu} \left(1 - \tfrac{1}{3} \mathrm{Re} \, \mathrm{tr}(U^{pl}_{x,\mu\nu})\right) \ , \\ \beta_{\mathrm{W}} &= \frac{1}{g^2 a^2} \ , \end{aligned} \tag{2.3.5}$$

where we have introduced Wilson's beta β_{W}.

The Fermionic Part

A simple symmetric lattice version of the covariant derivative $A_\mu + \partial_\mu$ is given by

$$D_\mu(x) = \tfrac{1}{2}(U_{x,\mu}\delta_{x+\hat{\mu},y} - U^{\dagger}_{x-\hat{\mu},\mu}\delta_{x-\hat{\mu},y}) \ . \tag{2.3.6}$$

Using this the so-called naive lattice fermion matrix M_{N} can be derived:

$$M_{\mathrm{N}}(x,y) = m \ \delta_{x,y} + \tfrac{1}{2} \sum_{\mu=1}^{4} \gamma_\mu \left[U_{x,\mu}\delta_{x+\hat{\mu},y} - U^{\dagger}_{x-\hat{\mu},\mu}\delta_{x-\hat{\mu},y} \right] \ . \tag{2.3.7}$$

It is called naive for the following reason: In the free case, i.e., all $U_{x,\mu} = 1$, and for one fermion flavor, the resulting propagator in momentum-space is:

$$\langle \overline{\psi} \psi(p) \rangle = [i \sum_\mu \gamma_\mu \sin(p_\mu a) + m]^{-1} \ . \tag{2.3.8}$$

In the continuum limit and for $m = 0$ this should have one pole at $p = (0, 0)$, corresponding to the one fermion species. What we get instead are 16 poles, one at every corner of the Brioullin zone

$$p_1 = (0, 0, 0, 0),\ p_2 = (0, 0, 0, \pi),\ \ldots,\ p_{16} = (\pi, \pi, \pi, \pi)\ , \tag{2.3.9}$$

which all survive in the continuum limit. This is known as the *fermion doubling problem.*

To overcome the fermion doubling problem Wilson added a second derivative term to the fermion matrix:

$$\tfrac{r}{2}[U_{x,\mu}\delta_{x+\hat{\mu},y} - 2\delta_{x,y} + U^{\dagger}_{x-\hat{\mu},\mu}\delta_{x-\hat{\mu},y}]\ , \tag{2.3.10}$$

which gives the unwanted extra 15 fermion a mass proportional to $1/a$. This way the unwanted fermions get infinitely massive in the continuum limit and ought to decouple. The Wilson fermion matrix can be written as

$$M_{\mathrm{W}}(x, y) = \delta_{x,y} - \kappa \sum_{\mu=1}^{4} \left[(r - \gamma_\mu)U_{x,\mu}\delta_{x+\hat{\mu},y} + (r + \gamma_\mu)U^{\dagger}_{x-\hat{\mu},\mu}\delta_{x-\hat{\mu},y}\right]\ . \tag{2.3.11}$$

Wilson's fermionic action is given by

$$S_f(U, \overline{\psi}, \psi) = \overline{\psi} M_{\mathrm{W}} \psi\ . \tag{2.3.12}$$

2.3.1 Symmetries of Wilson's Action

The different symmetries of the continuum QCD Lagrangian (Section 2.1) are reflected by Wilson's lattice action in different ways:

- Gauge symmetry and charge conjugation are preserved exactly as they appear in the continuum.
- The translation and $O(n)$ rotation symmetries from the continuum manifest themselves on the lattice as discrete translations and hypercubic rotations and reflections. In the continuum limit $a \to 0$ these symmetries become their continuum counterpart again.

- The introduction of the second derivative term 2.3.10 violates the γ_5 anti-commutation relation

  $$\{M_{\mathrm{W}}, \gamma_5\} \neq 0 \ ,$$

  even at vanishing quark mass. Therefore chiral symmetry is not maintained like in the continuum. Furthermore chiral symmetry is broken in a hard way in $\mathcal{O}(a)$. Also there is no exact lattice symmetry corresponding to continuum chiral symmetry (like in the case of $O(n)$ and translation symmetries which have discrete counterparts, see above).

2.4 Path Integrals on the Lattice

For the integration over the lattice gauge fields U Wilson proposed the Haar group measure, we denote it as $d\,U_{x,\mu}$. This measure is invariant under the gauge group transformations ($SU(3)$ in our case). For the integration over the quark fields the rules of integration over Grassmann variables are applied. The lattice counterpart of the vacuum expectation value (2.2.1) is

$$\begin{aligned}&\langle 0|T\{\hat{O}_1(x_1), \hat{O}_2(x_2), \ldots\}|0\rangle = \\ &\frac{1}{Z}\int \prod_{x,\mu} dU_{x,\mu} \prod_x d\overline{\psi}(x) \prod_x d\psi(x)\ O_1(x_1)O_2(x_2)\ldots e^{-S(U,\overline{\psi},\psi)}\ ,\end{aligned} \tag{2.4.1}$$

$$Z = \int \prod_{x,\mu} dU_{x,\mu} \prod_x d\overline{\psi}(x) \prod_x d\psi(x)\ e^{-S(U,\overline{\psi},\psi)}\ . \tag{2.4.2}$$

The continuum vacuum expectation values are defined as the *continuum limit* of the lattice vacuum expectation values, i.e., the limit of $a \to 0$. To perform the continuum limit the free parameters of the theory, the gauge coupling and the quark mass, have to be replaced by measurable quantities. This procedure is referred to as *renormalization*.

2.5 Improved Gauge Actions

The Taylor expansion of Wilson's gauge action gives $F_{\mu,\nu}F_{\mu,\nu}$ in leading order, corrections are $\mathcal{O}(a^2)$. There are various approaches that try to improve that by adding gauge loops that are longer than the plaquette, for instance of length six. These longer loops ought to cancel the $\mathcal{O}(a^2)$ corrections coming from the plaquette. These longer loops get a certain coupling strength. To choose this coupling strength (such that the wanted cancellation takes place), different techniques have been applied: Methods from perturbation theory, mean field improvement or renormalization group inspired so-called perfect actions.

2.6 Improved Fermionic Actions and Smearing

For Wilson's fermionic action, lattice corrections are starting at $\mathcal{O}(a)$. As in the case of the gauge action various ways to improving this have been suggested, perturbative and non-perturbative ones. One class of non perturbatively improved actions are the *lattice chiral* or *Ginsparg-Wilson actions*, which we will discuss in Chapter 3.

Another way of to effectively improve a fermionic action is *link smearing.* The idea is to define new link variables called *smeared links* that consist of some local combination of the usual link variables (which are then often called *thin links* for distinguishing). The process of smearing ought to reduce short distance (ultra-violet) fluctuations and leave the long distance properties of the system unchanged. It has been shown for various fermionic actions that some of their properties, including scaling, improve significantly when some sort of smearing is employed (see for instance [16, 17, 18]).

Chapter 3

Chiral Symmetry in Lattice QCD

3.1 Nielsen-Ninomiya No-Go Theorem

There is a theorem called *Nielsen-Ninomiya No-Go theorem* [3, 19] for a lattice Dirac operator D (a lattice discretization of the continuum Dirac operator $\not{D}$). This theorem states that such a D which has no doubler modes cannot obey the following conditions simultaneously:

- D is (exponentially) local
- D is translational invariant
- D is chirally symmetric (i.e. $\{D, \gamma_5\} = 0$).
- $D(p) = i\gamma_\mu p_\mu + (O)(ap^2)$, where $D(p)$ is the Fourier transform of D.

3.2 The Ginsparg-Wilson Relation

Ginsparg and Wilson [4] suggested to replace chiral symmetry in continuum, $\{D, \gamma_5\} = 0$, which is equivalent to $\{D^{-1}, \gamma_5\} = 0$, by the milder condition

$$\{D^{-1}, \gamma_5\} = 2aR\gamma_5 \ , \tag{3.2.1}$$

where a is the lattice spacing and R is some local operator.

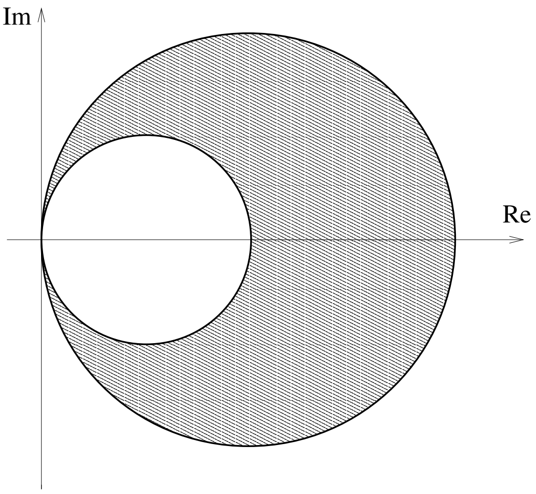

Figure 3.1: The spectrum of a lattice chiral Dirac operator is bounded by two circles.

It has been shown that this implies an exact symmetry of the lattice action, which becomes the usual chiral symmetry in the continuum limit [20], and is often referred to as *lattice chiral symmetry.* This symmetry implies some very favorable properties for actions satisfying it: Lattice chiral actions possess no additive mass correction. At finite masses they are always invertible (there are no exceptional configurations). They are automatically $\mathcal{O}(a)$ improved. There is a lattice version of the Atiyah-Singer index theorem, relating topology to the eigenmodes of the Dirac operator. For the renormalization constants the relation $Z_A = Z_V$ holds.

3.3 Neuberger's Overlap Fermions

Neuberger's overlap fermion matrix [7, 21] with zero mass fulfills the Ginsparg-Wilson relation (3.2.1) with $R = \frac{1}{2}I$:

$$\{D^{-1}, \gamma_5\} = a\gamma_5 \ , \tag{3.3.1}$$

I being the identity matrix. One starts with a kernel Dirac matrix, for instance the Wilson fermion matrix M_{W} with a negative mass, $m \approx -1/2$.

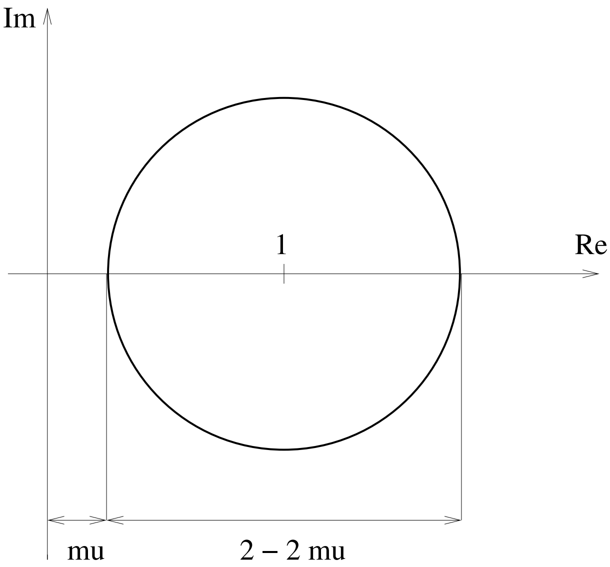

Figure 3.2: The spectrum of the massive overlap operator lies on the circle sketched above.

Using this one defines the overlap Dirac operator as

$$D_{\text{ov}} = 1 + \gamma_5 \, \text{sign}(H_{\text{W}}) \ , \tag{3.3.2}$$

where

$$H_{\text{W}} = \gamma_5 M_{\text{W}} \ . \tag{3.3.3}$$

From the massless case it is possible to construct massive fermions using the relation [22, 23]

$$M_{\text{ov}}(\mu) = (1-\mu) D_{\text{ov}} + \mu \ . \tag{3.3.4}$$

The mass parameter μ must be in the interval $[0,1]$, the corresponding fermion mass is

$$m = \mu Z_m^{-1} (1 + \mathcal{O}(a^2)) \ , \tag{3.3.5}$$

Z_m is the mass renormalization constant for which

$$Z_\psi Z_m = 1 \ , \tag{3.3.6}$$

and Z_ψ is the wave function renormalization constant.

3.4 Generalized Dirac Operators

Some parameterizations of Ginsparg-Wilson fermions start by writing down a most general Dirac operator that satisfies the basic symmetries of gauge invariance, translation invariance, hypercubic rotations and reflections, charge conjugation plus γ_5-hermiticity. These conditions fix some of the properties of the ansatz

$$D_{x,y} = \sum_n c_n(U) \delta_{x,y} \otimes \Gamma_n \otimes u_n \ , \tag{3.4.1}$$

where $c_n(U)$ is a real coefficient, Γ_n is an element of the Clifford algebra and u_n is a gauge path connecting x and y. Some freedom in the choice of the coefficients remains, this can be used to give the operator certain desired properties, for instance lattice chiral symmetry. For detailed descriptions of this approach we want to refer to [24] and to [8].

3.4.1 Fixed-point Fermions

The *fixed-point* (also referred to as *classically perfect*) fermions [25] follow a concept which is based on renormalization-group techniques. There one aims to simulate the theory close to the renormalized trajectory. From this one expects very good scaling properties. It has been shown that the fixed-point action satisfies the Ginsparg-Wilson relation [26].

This concept can be used to fix the free parameters of a generalized Dirac operator. Doing this one can retain a practical implementation or the theoretical concept [27]. For practical reasons these constructions can contain only a certain number of terms of the most general Dirac operator. An exact fixed-point action would have to contain infinitely many terms. Therefore such constructions can only be approximate fixed-point actions.

3.4.2 Chirally Improved Fermions

Another possibility fix the coefficients of a generalized Dirac operator was suggested in [8]. The idea is to plug the generalized Dirac operator into the Ginsparg-Wilson relation. Then a system of algebraic equations for the coefficients is retained. Solving these equations one can fix the coefficients of the generalized Dirac operator and retains a lattice chiral operator. Again in practice one is always restricted to a certain number of terms. For an exact Ginsparg-Wilson operator one would need infinitely many terms, therefore such an operator can only be approximately lattice chiral.

In practice it was found that one would need very many terms in order to get an operator with no additive mass correction. Therefore the original concept has been modified to allow for an additional parameter which ensures that there is no additive mass correction [9]. This parameter is dependent of the underlying gauge background. In this modified version the concept was successfully used in quenched QCD calculations, for instance [28].

Chapter 4

Markov Chain Monte Carlo

4.1 Introduction

The purpose of this chapter is to discuss some mathematical aspects of of the simulation techniques we want to use. In the first part of this chapter (Sections 4.2 to 4.7) a quick overview over some standard methods of Markov Chain Monte Carlo methods is given. These can be found in various textbooks, for instance [29, 30]. In the second part of the chapter (Sections 4.8 to 4.9) the Hybrid Monte Carlo algorithm is discussed. This is done with a special emphasis on how it fits into the framework of the methods of the first part of this chapter. Many of the results in this chapter can be proved in an elementary and instructive way, therefore we want to discuss some of these proofs. For the notation throughout this chapter we will closely follow [29]. In the last Section of this chapter we will briefly discuss autocorrelation and error estimation for data with autocorrelation.

4.2 General Definitions

For our lattice QCD calculations we are interested in integrals of the general form

$$E(m) = \int dx\, \pi(x) m(x) \ . \tag{4.2.1}$$

Here $x \in R$ stands for some field, taking values from a measure-space R, and dx for an appropriate measure. The function $m(x)$ is assumed to be complex valued, the function $\pi(x)$ is assumed to be positive real valued, and normalized, i.e.,

$$\int dx\, \pi(x) = 1 \ . \tag{4.2.2}$$

We can use this this to define a new measure Π on R as $\Pi(dx) = \pi(x)dx$, $\Pi(A) = \int_A \Pi(dx)$, which we call *distribution*, $\pi(x)$ we call *distribution density.* Using this the integral can be written as

$$E(m) = \int \Pi(dx)\, m(x) \ . \tag{4.2.3}$$

The method of *Markov Chain Monte Carlo* integration will allow us to generate samples $\{ X_i : i = 1, 2, \dots \}$ according to the distribution Π. Using these we can then express our integral as

$$E(m) = \lim_{n\to\infty} \frac{1}{n+1} \sum_{i=0}^{n} m(X_i) \ . \tag{4.2.4}$$

This will allow us to estimate our integral as a finite sum over a certain sample.

4.3 Markov Chains

Let $\{ X_i : i = 0, 1, \dots \}$ be a set of random variables taking values in a set R. Let further X_n be related to $X_{n-1}, \dots, X_0$ by a conditional probability $P(X_n \in B | X_0 = x_0, \dots, X_{n-1} = x_{n-1})$ (the probability that $X_n \in B$ under the condition $X_0 = x_0, \dots, X_{n-1} = x_{n-1}$). The set $\{ X_i : i = 0, 1, \dots \}$ is then called a *Markov chain* if

$$P(X_n \in B | X_0 = x_0, \dots, X_{n-1} = x_{n-1}) = P(X_n \in B | X_{n-1} = x_{n-1}) \tag{4.3.1}$$

for every $B \subset R$ and $n \in 0, 1, \dots, N$. The chain is called *time homogeneous* if

$$P(X_n \in B | X_{n-1} = x) = P(X_1 \in B | X_0 = x) \ . \tag{4.3.2}$$

We will restrict ourselves to that case from now on. The function

$$K(B|x) = P(X_1 \in B | X_0 = x) \tag{4.3.3}$$

we call the *transition probability*, and

$$K^n(B|x) = P(X_n \in B|X_0 = x) \tag{4.3.4}$$

the *n-step transition probability.* The Markov chain is determined by K and some initial distribution P_0.

A distribution Π on R is called a *stationary distribution* if

$$\Pi(B) = \int_R K(B|x)\Pi(dx)\ ,\ \forall B \subset R\ . \tag{4.3.5}$$

The Markov chain $\{\, X_i : i = 0, 1, \ldots \}$ is called *reversible* with respect to Π if

$$\int_A K(B|x)\Pi(dx) = \int_B K(A|x)\Pi(dx) \tag{4.3.6}$$

for all $A, B \subset R$. The (4.3.6) is also called *detailed balance condition.* From (4.3.6) it follows immediately that

$$\int_R K(B|x)\Pi(dx) = \int_B K(R|x)\Pi(dx) = \Pi(B)\ , \tag{4.3.7}$$

therefore Π is a stationary distribution for the chain. A chain is called *ϕ-irreducible* if there exists a measure ϕ on R such that for all $A \subset R$ with $\phi(A) > 0$ and for all $x \in R$ there exists a positive integer n such that $K^n(A|x) > 0$.

If there is an integer $d \geq 2$ and disjoint subsets of R ,

$$A_0, \ldots, A_d \subset R\ , \tag{4.3.8}$$

with $\Pi(A_i) > 0$ such that

$$K(A_i|x) = 1 \text{ for all } x \in A_i and\ 0 \leq i \leq d-1\ , \tag{4.3.9}$$

and

$$K(A_1|x) = 1 \text{ for all } x \in A_d\ , \tag{4.3.10}$$

then the chain is said to be *periodic.* The chain is called *a-periodic* otherwise.

If a chain has a stationary distribution π and it is ϕ-irreducible and a-periodic then it converges to π starting from almost every $x \in R$ (i.e., every $x \in R$ except maybe for a subset $A \subset R$ of measure $\Pi(A) = 0$).

4.4 The Metropolis-Hastings Algorithm

Metropolis et al. [31] proposed in 1953 an algorithm to construct a chain satisfying detailed balance (4.3.6) which was later generalized by Hastings [32]. We will discuss this generalized version, also known as *Metropolis-Hastings* algorithm. The chain is constructed recursively: To start we choose an $x_0 \in R$ and set $X_0 = x_0$. To generate X_{n+1} from X_n where $X_n = x_n$ we

- Choose a state y from a conditional distribution $Q(dy|x) = q(y|x)dy$. We refer to Q as proposal distribution, and to q as proposal density.
- Evaluate
$$\alpha(y|x) = \min\{1, \frac{q(x|y)\pi(y)}{q(y|x)\pi(x)}\} \ . \tag{4.4.1}$$
This α we call acceptance probability function.
- Generate $r \in [0, 1]$ with uniform distribution
- Set
$$X_{n+1} = \begin{cases} y \text{ if } \ \alpha(y|x) \geq r \\ x \text{ otherwise} \end{cases} . \tag{4.4.2}$$

This procedure defines a transition probability K:

$$K(B|x) = \int_B \alpha(y|x)q(y|x)dy + r(x)I_B(x) \ , \tag{4.4.3}$$

where

$$I_B(x) = \begin{cases} 1 & \text{if } x \in B \\ 0 & \text{otherwise} \end{cases} , \tag{4.4.4}$$

and $r(x)$ is the probability that a proposal is chosen which is then rejected, i.e., the probability that the new element of the chain is equal to the old one, x. We will not need an explicit form of this this function r for the proof of detailed balance below. For the matter of completeness only we write it down:

$$r(x) = \frac{1}{\pi(x)} \int_{\{y:\pi(y)q(x|y)<\pi(x)q(y|x)\}} [\pi(x)q(y|x) - \pi(y)q(x|y)]dy \ . \tag{4.4.5}$$

The transition probability K satisfies detailed balance with respect to π. To show that, we first derive a relation for α:

$$\begin{aligned} \frac{\alpha(y|x)}{\alpha(x|y)} &= \frac{\min\{1, \frac{q(x|y)\pi(x)}{q(y|x)\pi(y)}\}}{\min\{1, \frac{q(x|y)\pi(y)}{q(y|x)\pi(x)}\}} \\ &= \frac{q(x|y)\pi(x)}{q(y|x)\pi(y)} \ . \end{aligned} \tag{4.4.6}$$

Using this we get

$$\begin{aligned} &\int_A K(B|x)\pi(x)dx \\ &\quad = \int_A \left[\int_B \alpha(y|x)q(y|x)dy + r(x)I_B(x) \right] \pi(x)dx \\ &\quad = \int_A \int_B \alpha(y|x)q(y|x)\pi(x)dy + \int_{A\cap B} r(x)\pi(x)dx \\ &\quad = \int_B \int_A \alpha(y|x)q(y|x)\pi(x)dy + \int_B r(x)I_A(x)\pi(x)dx \\ &\quad = \int_B K(A|y)\pi(y)dy \ . \end{aligned} \tag{4.4.7}$$

The original Metropolis algorithm of [31] is a special case of the above algorithm with a symmetric proposal distribution, i.e., $q(x|y) = q(y|x)$.

4.5 Sampling in Subspaces

Let now be R a Cartesian product of a number n of sets $S^1, \dots, S^n$

$$R = S^1 \times S^2 \times \cdots \times S^n \ . \tag{4.5.1}$$

The elements of R can be written as n-tuples, for instance

$$x = (x^1, x^2, \dots, x^n) \ , \tag{4.5.2}$$

and the m-th element X_m of a Markov chain as

$$X_m = (X_m^1, X_m^2, \dots, X_m^n) \ . \tag{4.5.3}$$

Often it is impracticable to update this whole vector at once.

We are now going to construct a Markov chain recursively, altering only one of the X_m^k at once. Cycling through all the variables we will then be able to get a Markov chain converging to some distribution π for the whole vector. For every $1 \leq l \leq n$ we have a certain probability of getting from $X_m^l = x^l$ to $X_{m+1}^l = y^l$ in a single step, which we denote by

$$P(X_{m+1}^l \in B^l | X_{m+1}^1 = y^1, \dots, X_{m+1}^{l-1} = y^{l-1}, X_m^l = x^l, \dots, X_m^n = x^n) \ . \tag{4.5.4}$$

We want to construct a completely new state y out of a given state x by cycling through all the components of our state vector. The transition matrix of this chain is the product of all these probabilities

$$\begin{aligned} P(y \in B|x) =& P(X_{m+1}^1 \in B^1 | X_m^1 = x^1, \dots) \\ & P(X_{m+1}^2 \in B^2 | \dots, X_m^2 = x^2, \dots) \cdots \\ & P(X_{m+1}^n \in B^n | \dots, X_m^n = x^n) \ . \end{aligned} \tag{4.5.5}$$

If we choose them such that all leave our specific distribution Π invariant, then Π is a stationary distribution for the chain.

4.6 Gibbs Sampling

Ideally we would always like to draw independent samples from our distribution $\Pi(dx^1, \dots, dx^n)$ right from the start of our Markov chain (no equilibration time). This is not possible usually. However, in some cases it is possible to sample in that way in some small subspace from a corresponding conditional distribution. For instance one would sample in the l-th subspace (variable x^l), with the values of the other $x^k, k \neq l$ held constant.

$$\Pi(dx^l | x^k, k \neq l) \ , \tag{4.6.1}$$

and set the probabilities of (4.5.5):

$$\begin{aligned} P(X_{m+1}^l \in B^l | X_{m+1}^1 = y^1, \dots, X_{m+1}^{l-1} = y^{l-1}, X_m^l = x^l, \dots, X_m^n = i^n) = \\ \Pi(B^l | y^1, \dots, y^{l-1}, x^{l+1}, \dots, i^n) \ . \end{aligned} \tag{4.6.2}$$

Then Π is a stationary distribution for our chain. This so called *Gibbs Sampling* was introduced by Geman and Geman in 1984 [33].

4.7 Mixed Chains

In the last Sections we discussed two possibilities of generating a distribution which is stationary with respect to a chosen distribution, the Metropolis-Hastings algorithm and Gibbs-sampling. There are many others. If we combine them such that we generate samples in different subspaces with different of these algorithms, the resulting chain will again be stationary with respect to our chosen distribution. Such an approach we call a *mixed chain.*

4.8 Hybrid Algorithm

Let the distribution be of the form

$$\Pi(dx) = e^{-V(x)} dx \ , \tag{4.8.1}$$

where $x = (x_1, \ldots, x_n) \in \mathbb{R}^N$. For every x_i we add an auxiliary variable $p_i \in \mathbb{R}$, $p = (p_1, \ldots, p_n)$, and define a Hamiltonian

$$\mathcal{H} = \sum_{i=1}^{N} \frac{p_i^2}{2} + V(x) \ . \tag{4.8.2}$$

We now define a distribution

$$\tilde{\Pi}(dx, dp) = e^{-H(x,p)} \ dx \ dp \ . \tag{4.8.3}$$

Integrating over p we get back our original distribution Π

$$\int_p \tilde{\Pi}(\cdot, dp) = \Pi(\cdot) \ . \tag{4.8.4}$$

Thus if we sample $\{(x,p)\}$from $\tilde{\Pi}$, x will be distributed according to Π.

Integrating Hamilton's equations,

$$\begin{aligned} \frac{dx_i}{dt} &= \frac{\partial \mathcal{H}}{\partial p_i} \ , \\ \frac{dp_i}{dt} &= -\frac{\partial \mathcal{H}}{\partial x_i} \ , \end{aligned} \tag{4.8.5}$$

for a time t one gets an evolution of x and p:

$$(x,p) \to (x', p') = g^t(x,p) \ , \tag{4.8.6}$$

which is

- Deterministic,
- Energy conserving: $\mathcal{H}(x', p') = \mathcal{H}(x, p)$,
- Reversible: $g^t(g^{-t}(x, p)) = (x, p)$,
- Area preserving: $dx'\ dp' = dx\ dp$.

This evolution can be used to construct an updating scheme which satisfies detailed balance with respect to our distribution $\tilde{\Pi}(dx, dp) = e^{-\mathcal{H}(x,p)}\ dx\ dp$ and which is ergodic:

- Choose p from the gaussian distribution $e^{\sum_{i=1}^{N} \frac{p_i^2}{2}}$.
- Evolve the system for a certain time t, going from (x, p) to $(x', p') = g^t(x, p)$.

To see how this satisfies detailed balance we choose some $A \in R^{2n}$ and define $B = g(A)$. We call K^t the transition probability which results from evolving for a certain time t:

$$K^t(B|(x, p)) = I_B(g^t(x, p)) = \begin{cases} 1 & \text{if } g^t(x, p) \in B \\ 0 & \text{otherwise} \end{cases} \tag{4.8.7}$$

(I_B defined as in (4.4.4)). For this we get the relation:

$$\begin{aligned} \int_A K^t(B|(x, p))e^{-\mathcal{H}(x,p)}\ dx\ dp &= \int_A I_B(g^t(x, p))e^{-\mathcal{H}(x,p)}\ dx\ dp \\ &= \int_B I_A(g^{-t}(x', p'))e^{-\mathcal{H}(x',p')}\ dx'\ dp' \\ &= \int_B K^{-t}(A|(x', p'))e^{-\mathcal{H}(x',p')}\ dx'\ dp'\ . \end{aligned} \tag{4.8.8}$$

This would be detailed balance in the usual form (4.3.6) if we had $K^t = K^{-t}$. To restore detailed balance we could randomly choose the sign of t for the evolution. However, we do not have to do that: Before starting the evolution we refresh our p, and the chance of getting p and $-p$ is the same ($e^{-p^2/2} = e^{-(-p)^2/2}$). Reversing the sign of t has the same effect as reversing the sign of p.

4.9 Hybrid Monte Carlo Algorithm

Assume we have an evolution which is reversible and area preserving, but not energy conserving. This case occurs often in practice: Numerical techniques for the integration of differential equations make use of a time discretization leading to discretization errors. In our case this would mean a violation of energy conservation. Using this for the hybrid algorithm described above would violate detailed balance. Therefore in the hybrid algorithm one has to make sure that the violation of energy conservation is kept small. This can be achieved by making the time discretization fine. Choosing a fine discretization, however, makes the integration expensive. On the other hand, there exist numerical integration techniques which satisfy the requirement of area preservation and reversibility at arbitrarily coarse time discretizations.

For such an integrator detailed balance can be restored. The idea of the *Hybrid Monte Carlo* algorithm [34] is to add a Metropolis-Hastings like accept-reject at the end of the hybrid algorithm, which corrects for this change of energy. The result is a three step algorithm (where the first two steps are the same as in the hybrid algorithm):

- Choose p from the gaussian distribution $e^{\sum_{i=1}^{N}\frac{p_i^2}{2}}$.
- Evolve the system for a certain time t, going from (x,p) to $(x',p') = g^t(x,p)$.
- Accept the (x',p') with a probability $\min\{1, e^{-\delta\mathcal{H}}\}$ where $\delta\mathcal{H} = \mathcal{H}(x',p') - \mathcal{H}(x,p)$.

The transition probability in this case is

$$K^t(B|(x,p)) = I_B(g^t(x,p))\min\{1, e^{-\delta\mathcal{H}}\} + r(x)I_B(x,p) \ . \tag{4.9.1}$$

By $r(x)$ we denote the probability that (x',p') is rejected, explicitly this is

$$r(x) = \begin{cases} 0 & \text{if } e^{-\delta\mathcal{H}} > 1 \\ 1 - e^{-\delta\mathcal{H}} & \text{otherwise} \end{cases} \ . \tag{4.9.2}$$

The explicit form of $r(x)$ is given only for the sake of completeness, it is not needed in the following relation:

$$\begin{aligned}
&\int_A K^t(B|(x,p))e^{-\mathcal{H}(x,p)}d^nxd^np \\
&= \int_A [I_B(g^t(x,p))e^{-\mathcal{H}(x,p)}\min\{1,e^{-\delta\mathcal{H}}\} + r(x)I_B(x,p)]d^nxd^np \\
&= \int_B I_A(g^{-t}(x',p'))e^{-\mathcal{H}(x',p')}e^{\delta\mathcal{H}}\min\{1,e^{-\delta\mathcal{H}}\}d^nx'd^np' + \int_{A\cap B} r(x)d^nxd^np \\
&= \int_B I_A(g^{-t}(x',p'))e^{-\mathcal{H}(x',p')}\min\{e^{\delta\mathcal{H}},1\}d^nx'd^np' + \int_B r(x',p')I_A(x',p')d^nx'd^np' \\
&= \int_B K^{-t}(A|(x',p'))e^{-\mathcal{H}(x',p')}d^nx'd^np' \ .
\end{aligned} \tag{4.9.3}$$

The result is the same as in (4.8.8), detailed balance follows from this (see argumentation following (4.8.8)).

In principle the time evolution does not need have anything to do with $\mathcal{H}$ as long at it is area preserving and reversible. However the acceptance will be very poor if the energy non-conservation is too large.

4.10 Autocorrelation and Error Estimation

Error Bars for Uncorrelated Samples

Let$\{X_1,\ldots,X_N\}$ be uncorrelated random numbers of the same type. Being of the same type these random numbers of course have the same expectation value

$$\mu_X = \langle X_i\rangle \ , \ i=1,\ldots,N \tag{4.10.1}$$

and variance

$$\sigma_X^2 = \langle (X_i-\mu)^2\rangle \ , \ i=1,\ldots,N \ . \tag{4.10.2}$$

We define the sample average as

$$\hat{X} = \frac{1}{N}\sum_{i=1}^{N} X_i \tag{4.10.3}$$

The variance of $\hat{X}$ is then

$$\sigma^2_{\hat{X}} = \frac{1}{N}\sigma_X \ . \tag{4.10.4}$$

Using a sample $\{x_1, \ldots, x_N\}$ of $\{X_1, \ldots, X_N\}$ to estimate μ_X, as one usually does in a Markov Chain Monte Carlo calculation, one can take

$$\sigma_{\hat{X}} = \sqrt{\sigma^2_{\hat{X}}} \tag{4.10.5}$$

as a measure of the error of this estimate, also called error bar. Of course in such a calculation $\sigma_{\hat{X}}$ has to be estimated from the sample $\{x_1, \ldots, x_N\}$ too. This can be done in various, for instance using the jackknife method.

Error Bars for Correlated Samples

If the random variables $\{X_1, \ldots, X_N\}$ are correlated the situation is different. To define the integrated autocorrelation time τ_{int} we consider an infinite set of random variables of the same type, $\{X_1, X_2, \ldots\}$ and let $\{X_1, \ldots, X_N\}$ be the first N of these. We define the autocorrelation function ρ as

$$\begin{aligned}\rho(i) &= \frac{\Gamma(i)}{\Gamma(0)} \\ \Gamma(i) &= \langle (X_k - \langle X \rangle)(X_{k+i} - \langle X \rangle) \rangle \ .\end{aligned} \tag{4.10.6}$$

The k appearing in the definition of Γ stands for any value from $\{1, 2, \ldots\}$. Because we assumed all X_i to be of the same type this gives a unique definition of $\Gamma(i)$ independent of the value of k.

The integrated autocorrelation time is defined as

$$\tau_{int} = \frac{1}{2}\sum_{i=1}^{\infty} \rho(i) \tag{4.10.7}$$

When estimating the statistical error, the autocorrelation time has to be taken into account. The variance of the sample average for autocorrelated data is

$$\sigma^2_{\hat{X}} = \frac{2\tau_{int}}{N}\sigma^2_X \tag{4.10.8}$$

Binning

To calculate $\sigma^2_{\hat{X}}$ using a sample $\{x_1, \ldots, x_N\}$ from $\{X_1, \ldots, X_N\}$ the method of binning can be used. For this method one introduces bins of size M. We assume M to be an integer fraction of N. If $N = L \times N$ then one divides $\{X_1, \ldots, X_N\}$ into L bins. For these bins one can calculate a sample average which we we simply call $\hat{M}$, and a corresponding variance which we call $\sigma^2_{\hat{M}}$. For uncorrelated data one gets

$$\sigma^2_{\hat{M}} = \frac{1}{M}\sigma^2_{\mu} \ . \tag{4.10.9}$$

The expectation value of the average over all samples takes the value μ (correlated and uncorrelated data). Again for uncorrelated samples the variance of this is simply σ^2_{μ}. For correlated data this variance depends on the bin size M, it approaches a constant in the limit $M \to \infty$. In practice this variance will be more or less constant if $M > \tau_{int}$. Therefore error bars for autocorrelated data can be estimated using bins of various size, and taking the variances for a bin size where the variance is maximal. This procedure only makes sense for $M \ll N$, however.

Like in the case of un-binned data one can apply the method of jackknife to the binned data to estimate the wanted variances. This we refer to as *binned jackknife.*

Chapter 5

Lattice QCD Simulations

5.1 Prerequisites

When talking about "Lattice QCD simulations" what we mean is to calculate path integrals of the form (2.4.1). We are doing this using Markov Chain Monte Carlo techniques. Before we can start doing that, we integrate out the Grassmann fields $\overline{\psi}$ and ψ. In the simplest case, where except in the action S_f there are no Grassmann variables in the path integral the analytic integration yields

$$\int d\overline{\psi} d\psi \; e^{\overline{\psi} M \psi} = \det(M) \; . \tag{5.1.1}$$

If there are fermionic operators in the path integral the integration yields some fermionic propagators, for instance (simplest case)

$$\int d\overline{\psi} d\psi \; \overline{\psi}_i \psi_j e^{\overline{\psi} M \psi} = M^{-1}_{i,j} \det(M) \; . \tag{5.1.2}$$

The path integrals of (2.4.1) can thus be expressed as integrals over the gauge fields U only, what we have to deal with is (in a most general form)

$$\frac{1}{Z} \int \prod_{x,\mu} dU_{x,\mu} \; f(U) e^{-S_g(U)} \det(M) \; , \tag{5.1.3}$$

$$Z = \int \prod_{x,\mu} dU_{x,\mu} \; e^{-S_g(U)} \det(M) \; . \tag{5.1.4}$$

If we take the $\det(M)^n$ instead of $\det(M)$ in the above formulae we get n mass-degenerate flavors for every flavor of our original action. In the rest of this chapter we will discuss how to sample this Z using Markov Chain Monte Carlo methods.

5.2 Quenched Simulations

For a so-called *quenched simulation* we assume the fermionic determinant to be constant. We set

$$\det(M(U)) = 1 \ . \tag{5.2.1}$$

Of course this is an uncontrolled approximation, doing this one ignores all fermionic pair-creation and - annihilation processes. This yields a weight function which is merely dependent on the gauge action

$$Z = \int \prod_{x,\mu} dU_{x,\mu} \; e^{-S_g(U)} \ . \tag{5.2.2}$$

Exploiting the local nature of the gauge action it allows for an efficient updating. In a *local updating scheme* one goes through the lattice changing one link variable after the other. The locality of the action means that the cost of one such step is independent of the lattice volume V. An update all the link variables is called *sweep*. For this we have to do V local updates. The cost of one sweep thus increases linearly with the lattice volume.

5.3 Locality of Lattice QCD and the Fermionic Determinant

A priori, the fermionic as well as the gauge part of our action is local. However, integrating out the Grassmann variables we get the fermionic determinant $det(M)$ as part of our action. This determinant is not only non-local but the cost of evaluating it increases fastly with the lattice volume. The cost of evaluating a determinant of a general $n \times n$ matrix with standard methods (gaussian elimination, usually done as a LU decomposition) increases as n^3. The size n of the fermionic matrix is related to the number of lattice sites N_s by $n = 12N_s$. For any simulation using such a brute-force approach the cost would therefore increase at least proportional to N_s^3. Fortunately there exist much better ways to do simulations with dynamical fermions which will be discussed in the subsequent Sections of this chapter.

5.4 Pseudofermions

In 1981 Petcher and Weingarten proposed a way to deal with the fermionic determinant by introducing an auxiliary bosonic field which is often called a *pseudofermion field* [35]. They used the relation

$$\det(A^\dagger A) \propto \int d\mathrm{Re}(\phi) d\mathrm{Im}(\phi) e^{-|A^{-1}\phi|^2} \ . \tag{5.4.1}$$

If we set $A = M$ we can rewrite a $2k$ flavor partition function (assuming M is an k-flavor fermion matrix) as

$$\begin{aligned} Z &= \int \prod_{x,\mu} dU_{x,\mu} \ e^{-S_g(U)} \det(M)^2 \\ &= \int \prod_{x,\mu} dU_{x,\mu} \ e^{-S_g(U)} d\mathrm{Re}(\phi) d\mathrm{Im}(\phi) e^{-|M^{-1}\phi|^2} \ . \end{aligned} \tag{5.4.2}$$

If we want to avoid this doubling of flavors (which has of course nothing to do with the "fermion doubling problem" discussed in Section 2.3) for any M, with $\det(M)$ positive and real, we can set $A = \sqrt{M}$, retaining the determinant of M as $\det(M) = \det(A^\dagger A)$.

A Simple Updating Scheme with Pseudofermions

We want to discuss here a simple updating scheme for the partition function (5.4.2). In this two-step updating scheme we update U in step one and ϕ in the other.

- For fixed U we can generate ϕ directly from the distribution $e^{-|M^{-1}\phi|^2} d\phi$ (Gibbs sampling, see Section 4.6): This can be done by generating an auxiliary vector ξ according to the gaussian density $e^{-|\xi|^2}$ and setting $\phi = M\xi$.
- To update U we use the Metropolis-Hastings algorithm 4.4: For fixed ϕ generate a proposal gauge field U' from U with a transition probability

function $Q(U'|U) = q(U'|U)dU'$ and accept it with

$$\min\left(1, \frac{q(U|U')e^{-S_g(U')-|M^{-1}(U')\phi|^2}}{q(U'|U)e^{-S_g(U)-|M^{-1}(U)\phi|^2}}\right) . \tag{5.4.3}$$

Note that in every of these two steps we have to re-calculate $M^{-1}\phi$, either because ϕ changes or because U and therefore $M = M(U)$ changes. This calculation of $M^{-1}\phi$ is the most time consuming part of this procedure. It can be done efficiently using conjugate gradient algorithms (see for instance [36]) which can solve the problem in a time proportional to the size of the matrix n for sparse matrices, like the Dirac Matrices we are concerned about.

A sweep through the whole lattice thus takes a time proportional to N_s^2 (N_s being the number of lattice sites) for the updating scheme discussed here.

5.5 Hybrid Monte Carlo for Lattice QCD

We will now discuss how the second step of the algorithm of the last Section can be replaced by the Hybrid Monte Carlo procedure (see also Section 4.9).

For this we need a generalization of the Hamiltonian evolution for a system of classical mechanics to our system of U fields (the pseudofermions ϕ we want to hold constant in this step). This generalization can be found in [37]: For every $U \in SU(3)$ a conjugate momentum $p \in su(3)$ is introduced which is used to define a time derivative of U as

$$\dot{U} = ipU , \tag{5.5.1}$$

which is the equation of motion of U. These momenta are then used to define a Hamiltonian as

$$\mathcal{H} = \frac{1}{2}\sum_{i,\mu} \mathrm{tr}(p_{i,\mu}^2) + S_g + \phi^\dagger (M^\dagger M)^{-1}\phi . \tag{5.5.2}$$

The equation of motion for p is retained by writing down $\dot{\mathcal{H}}$ in terms of $U, p, \dot{U}$ and $\dot{p}$ and demanding that $\dot{\mathcal{H}} = 0$. This gives $\dot{p} = f(U, p, \dot{U})$. The function f of course depends on the $\mathcal{H}$ which in turn depends on p and the action used. The exact form of f for Wilson's gauge action and staggered as well as Wilson's fermion action can be found in [37]. For the Lüscher-Weisz action and the chirally improved fermionic action we will derive it in Chapter 6.

To do the evolution we still need a way to integrate the equations of motion. The integration scheme has to be reversible and area preserving. It should conserve energy as good as possible, since a large $\delta\mathcal{H}$ leeds to poor acceptance. A particularly simple integration scheme that is area preserving and reversible is the *leapfrog* integration. We call the evolution time t again, and divide it into n intervals of $\delta t = t/n$:

```
$U(\frac{\delta t}{2}) = e^{i\frac{\delta t}{2}\,p(0)}\,U(0)$
for $s = 0$ to $t - 2\delta t$ step $\delta t$ do
    $p(s+\delta t) = p(s) + \dot{p}(s)\delta t$
    $U(s+\frac{3\delta t}{2}) = e^{i\delta t\,p(s+\frac{\delta t}{2})}\,U(s+\frac{\delta t}{2})$
end for
$p(t) = p(t-\delta t) + \dot{p}(t-\delta t)\delta t$
$U(t) = e^{i\frac{\delta t}{2}\,p(t-\frac{\delta t}{2})}\,U(t-\frac{\delta t}{2})$
```

How does the cost of this HMC updating grow as a function of the volume N_s? Since the acceptance probability falls exponentially like $e^{\langle\delta\mathcal{H}\rangle}$ we have to keep $\delta\mathcal{H}$ small, otherwise we will get very poor performance. We can keep it small by adjusting δt. It has been shown [38, 39] that with this integration scheme in order to keep $\langle\delta H\rangle$ constant the integration step size has to grow like $\frac{1}{\delta t} \propto N_s^{1/4}$. Therefore, in order to keep t constant the number of steps n has to grow like $n \propto N_s^{1/4}$. Since the cost of evaluating $\mathcal{H}$ grows like N_s the cost of HMC with leapfrog integration grows like $N_s^{5/4}$.

5.6 The Cost of Dynamical Fermion Simulations

In this Section we want to discuss the cost of dynamical fermion simulations using the Hybrid Monte Carlo algorithm depending on the (not independent) quantities M_π/M_ρ (determined by the quark mass m_q) the number of lattice sites N_s the physical lattice volume V and the lattice spacing a. For traditional HMC an empirical formula for the scaling behavior can be found in [40]. It was found that the cost grows proportional to :

$$\left(\frac{m_\pi}{m_\rho}\right)^{-6} V^{5/4} a^{-7} \ . \tag{5.6.1}$$

This behavior was found empirically. We want to discuss some theoretical arguments supporting that it should be like that.

- The V dependence: At constant a and quark mass parameter the physical volume $V \propto N_s$. The cost of evaluating our action grows like N_s. Another factor of $N_s^{1/4}$ is obtained because going to larger lattices the integration step size during the molecular dynamics trajectory has to be decreased [38, 39].

- The a dependence: To keep the physical volume V constant when decreasing a, N_s has to grow proportional to a^{-4}, therefore from the $N_s^{5/4}$ growth of the cost discussed above we get a factor proportional to a^{-5}. One also expects a growth of the cost related to the fact that the condition number of the fermion matrix grows like $1/(am_q)$, or like $1/a$ for m_q held fixed. This is expected to make the inversion of the fermion matrix more expensive by a factor of $1/a^2$. Also the integration step size has to be decreased when the condition number is increased.

- The M_π/M_ρ dependence: The cost of the inversion is expected to grow like $1/(am_q)^2$, in addition to this the integration step size has to be shrunk. This provides a factor of m_q^{-3} for fixed a. Using the relation $M_\pi^2 \propto m_q$ (leading order chiral perturbation theory) and the fact that M_ρ is approaching a constant when m_q is taken to zero this gives a factor $(M_\pi/M_\rho)^{-6}$ in leading order of m_q.

Dynamical fermion actions using Wilson's action and traditional HMC proved to be impractical at $M_\pi/M_\rho < 0.5$, or $M_\pi < 400 MeV$ on the computers available today or in the near future (see [41, 42, 43, 44, 45] for some more recent results).

Recent algorithmic developments now allow to simulate Wilson's theory at much smaller masses than with traditional algorithms. One approach

[46] is to use multiple pseudofermions [47] together with multiple timescale integration [48]. The other approach is to use domain decomposition methods [49].

These new methods may overcome the $(M_\pi/M_\rho)^{-6}$ cost increase of dynamical fermion simulations and promise to allow to simulate at pion masses as low as 200 MeV in the near future [50].

5.7 Improved Actions

Apart from improvement of the algorithm there is much room for improving the performance of dynamical fermion simulations by using improved actions. This can be seen from the discussion in Section 5.6. Improved actions allow to work with larger lattice spacings a and the cost of the simulations grows as a^{-7}.

Traditional perturbatively improved fermionic actions like clover-improved Wilson and improved staggered as well as improved gauge actions like the Lüscher-Weisz action have long been used in dynamical simulations. They can be used with HMC straightforwardly.

It has been shown that link-smearing can improve fermionic actions substantially. Factors of two in lattice spacing could be gained in tests with HYP smeared link variables [16]. Therefore such techniques bear a large potential for improvement in quenched as well as in dynamical simulations. Unfortunately traditional smeared link variables are not well suited for being used together with the HMC algorithm since they are difficult to differentiate. To circumvent that problem different ways have been suggested: The first one was to develop methods other than HMC for the simulation of quantum field theories with dynamical fermions [51, 52, 53]. These are very substantially improved versions of the algorithm discussed in Section 5.4. The big downturn of these methods is that they scale unfavorably with the lattice volume, like V^2, or maybe even worse.

Another approach to do dynamical fermion simulations with smeared link variables is to develop new methods of smearing, which can be used together with HMC [54, 55].

5.8 Ginsparg-Wilson Type Actions

Having an action which fulfills the Ginsparg-Wilson relation certainly is very desirable for every quenched and dynamical simulation. Besides other advantages such actions are expected to allow for larger lattice spacings which is a definite advantage when it comes to performance of dynamical fermion simulations. Of course such actions are much more expensive to evaluate than Wilson's action at large quark masses and a given lattice size. However, when going to small quark masses and exploiting the fact that one can use coarser lattices it might well be that simulations with Ginsparg-Wilson actions turn out to be even cheaper at some point.

Chapter 6

Hybrid Monte Carlo for Generalized Dirac Operators

6.1 Outline

This chapter describes the technical aspects of a Hybrid Monte Carlo applied to actions containing a generalized Dirac operators. Though we have implemented and tested it only in case of the chirally improved operator, the methods described here can be applied more generally.

6.2 The Basic Hamiltonian Evolution

We start describing the evolution of a basic HMC Hamiltonian:

$$\mathcal{H} = \frac{1}{2} \sum_{i,\mu} \mathrm{tr}(p_{i,\mu}^2) + S_g + \phi^\dagger (M^\dagger M)^{-1} \phi \ , \tag{6.2.1}$$

where the $p_{i,\mu}$ are traceless hermitian matrices acting as conjugate momenta to $U_{i,\mu}$, ϕ are complex vectors (pseudofermions), $M = M(U)$ is the fermion matrix and $S_g = S_g(U)$ is the gauge action. For the moment we assume that the Dirac matrix acts directly on the thin link variables, the case of smearing will be discussed later.

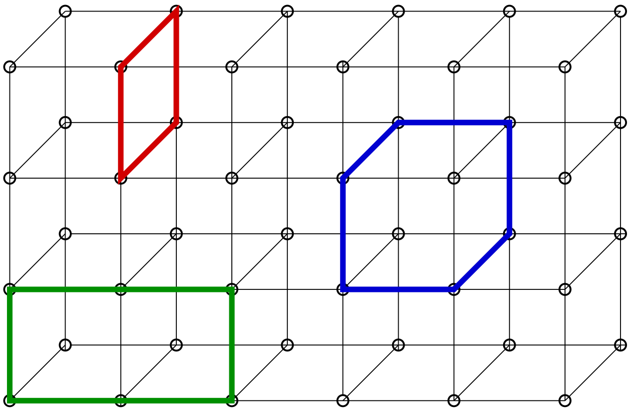

Figure 6.1: The types of gauge loops of the Lüscher-Weisz gauge action. Top: The plaquette; Bottom: The rectangle; Right: The twisted bent

6.2.1 The Equations of Motion

In this Section we derive the equations of motion for our Hamiltonian (6.2.1). We will do this by generalizing the methods of [37] for our action. The equation of motion for the link variables U is defined as

$$\dot{U}_{j,\mu} = i p_{j,\mu} U_{j,\mu} \ . \tag{6.2.2}$$

We now formally write down the time derivative of our Hamiltonian

$$\dot{\mathcal{H}} = \sum_{i,\mu} \mathrm{tr}(\dot{p_{i,\mu}} p_{i,\mu}) + \dot{S}_g + \tfrac{d}{dt}(\phi^\dagger (M^\dagger M)^{-1} \phi) \ . \tag{6.2.3}$$

To derive the equation of motion for p is we demand that the time evolution leaves $\mathcal{H}$ invariant.

Contributions from the Gauge Action

As our gauge action we choose the Lüscher-Weisz gauge action [56]. In addition to the plaquette term from Wilson's gauge action it contains a sum over all plane 2×1 loops, which we call rectangle, and all closed loops of length 6 along the edges of all 3-cubes, which we call twisted bent (see also Fig. 6.1).

$$\begin{aligned}S_g =&\beta_1 \sum_{pl} \tfrac{1}{3}\mathrm{Re}\,\mathrm{tr}[1-U^{pl}] + \beta_2 \sum_{re} \tfrac{1}{3}\mathrm{Re}\,\mathrm{tr}[1-U^{re}] + \\ &\beta_3 \sum_{par} \tfrac{1}{3}\mathrm{Re}\,\mathrm{tr}[1-U^{par}]\ ,\end{aligned} \tag{6.2.4}$$

with three real coefficients β_1, β_2 and β_3. To take the derivative we first rewrite this as

$$\begin{aligned}S_g =&\beta_1 \sum_{pl} \tfrac{1}{3}\,\mathrm{tr}[1-U^{pl}-(U^{pl})^\dagger] + \beta_2 \sum_{re} \tfrac{1}{3}\,\mathrm{tr}[1-U^{re}-(U^{re})^\dagger] + \\ &\beta_3 \sum_{par} \tfrac{1}{3}\,\mathrm{tr}[1-U^{par}-(U^{par})^\dagger]\ .\end{aligned} \tag{6.2.5}$$

For the time derivative $\dot{S}_g$ we have to generalize the notion of a staple for the plaquette to the longer loops: The usual staples $V^{pl}_{x,\mu}$ (for a link variable $U_{x,\mu}$) are defined as the sum over all open 3-loops starting at $x+\hat{\mu}$ and ending at x. We define $V^{re}_{x,\mu}$ as sum over all open 5-loops starting at $x+\hat{\mu}$ and ending at x which one gets by removing $U_{x,\mu}$ from the corresponding rectangle, and $V^{par}_{x,\mu}$ analogous. Using these definitions we get

$$\begin{aligned}\dot{S}_g =&\beta_1 \sum_{x,\mu} \tfrac{1}{3}\,\mathrm{tr}[\dot{U}_{x,\mu}V^{pl}_{x,\mu} - \dot{U}^\dagger_{x,\mu}(V^{pl}_{x,\mu})^\dagger] + \\ &\beta_2 \sum_{x,\mu} \tfrac{1}{3}\,\mathrm{tr}[\dot{U}_{x,\mu}V^{re}_{x,\mu} - \dot{U}^\dagger_{x,\mu}(V^{re}_{x,\mu})^\dagger] + \\ &\beta_3 \sum_{x,\mu} \tfrac{1}{3}\,\mathrm{tr}[\dot{U}_{x,\mu}V^{par}_{x,\mu} - \dot{U}^\dagger_{x,\mu}(V^{par}_{x,\mu})^\dagger]\ .\end{aligned} \tag{6.2.6}$$

Using the relation (6.2.2), hermiticity of $p_{j,\mu}$, and reordering we get

$$\begin{aligned}\dot{S}_g = \sum_{x,\mu} \tfrac{1}{3}\{\, &\mathrm{tr}[ip_{j,\mu}U_{j,\mu}(\beta_1 V^{pl}_{x,\mu} + \beta_2 V^{re}_{x,\mu} + \beta_3 V^{par}_{x,\mu})] + \\ &\mathrm{tr}[(-i)U^\dagger_{j,\mu}p_{j,\mu}(\beta_1(V^{pl}_{x,\mu})^\dagger + \beta_2(V^{re}_{x,\mu})^\dagger) + \beta_3(V^{par}_{x,\mu})^\dagger)]\}\ .\end{aligned} \tag{6.2.7}$$

Using the cyclicity of the trace we further rewrite this as

$$\begin{aligned}\dot{S}_g = \sum_{j,\mu} \tfrac{1}{3}\,\mathrm{tr}\, p_{j,\mu}[&iU_{j,\mu}(\beta_1 V^{pl}_{j,\mu} + \beta_2 V^{re}_{j,\mu} + \beta_3 V^{par}_{j,\mu}) + \\ &(\beta_1(V^{pl}_{j,\mu})^\dagger + \beta_2(V^{re}_{j,\mu})^\dagger + \beta_3(V^{par}_{j,\mu})^\dagger)(-i)U^\dagger_{j,\mu}]\ .\end{aligned} \tag{6.2.8}$$

Contributions from the Fermionic Action

We want to get the time derivative of the fermionic part of our Hamiltonian (6.2.1), which is

$$S_f = \phi^\dagger (M^\dagger M)^{-1} \phi \; . \tag{6.2.9}$$

We employ a standard trick to rewrite $\dot{M}^{-1}$ in terms of $\dot{M}$:

$$\frac{d}{dt}(M^{-1}M) = 0 \; , \tag{6.2.10}$$

and therefore

$$\dot{M}^{-1} = -M^{-1}\dot{M}M^{-1} \; . \tag{6.2.11}$$

Using this (and the analog for $M^\dagger$) we get

$$\begin{aligned} \phi^\dagger \frac{d}{dt}(M^\dagger M)^{-1}\phi &= -[X^\dagger M^\dagger \dot{M} X + X^\dagger \dot{M}^\dagger M X] \\ &= -[X^\dagger M^\dagger \dot{M} X + \text{h.c.}] \; , \end{aligned} \tag{6.2.12}$$

where $X = (M^\dagger M)^{-1}\phi$, and h.c. stands for hermitian conjugate (of the term left of it). Defining $P = XX^\dagger$ (exterior product) and using the cyclicity of the trace we can further rewrite this as

$$\phi^\dagger \frac{d}{dt}(M^\dagger M)^{-1}\phi = -\text{tr}[P M^\dagger \dot{M} + \text{h.c.}] \; . \tag{6.2.13}$$

Next we split off the mass term m from the Dirac matrix M

$$M(U) = D(U) + m \; . \tag{6.2.14}$$

Of course the constant term gives no contribution to the derivative, therefore

$$\dot{M}(U) = \dot{D}(U) \; . \tag{6.2.15}$$

The gauge dependent part $D(U)$ of our generalized Dirac matrix can be written as

$$D = \sum_n c_n \delta_{x_n,y_n} \otimes d_n \otimes u_n \; . \tag{6.2.16}$$

Every of these blocks connects just two sites x_n and y_n. The $d_n \in \{ \Gamma_1, \ldots, \Gamma_{16} \}$ are Clifford algebra elements (acting on the Dirac components of our vectors only), u_n is a certain gauge path between x_n and y_n, and c_n is a real coefficient (see also Fig. 6.2). We now write down the equations of motion for one

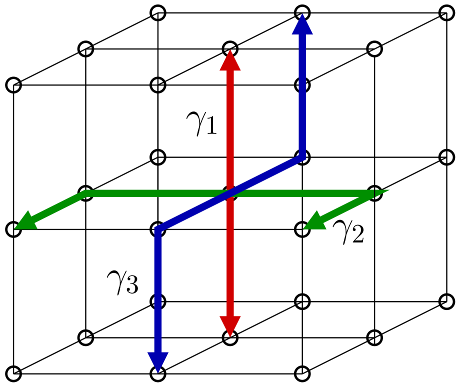

Figure 6.2: A schematic representation of three arbitrary terms of a generalized Dirac operator.

such element in the most general form. For the matter of illustration we will for a moment assume that $u = U_{j_1,\mu_1} U^\dagger_{j_2,\mu_2} U_{j_3,\mu_3}$. To get a time derivative for our example u we use (6.2.2) obtaining

$$\begin{aligned}\dot{u} =& i p_{j_1,\mu_1} U_{j_1,\mu_1} U^\dagger_{j_2,\mu_2} U_{j_3,\mu_3} \\ &+ U_{j_1,\mu_1} U^\dagger_{j_2,\mu_2} (-i p_{j_2,\mu_2}) U_{j_3,\mu_3} \\ &+ U_{j_1,\mu_1} U^\dagger_{j_2,\mu_2} (i p_{j_3,\mu_3} U_{j_3,\mu_3}) \ .\end{aligned} \tag{6.2.17}$$

We generalize this to an arbitrary u from (6.2.16), consisting of arbitrarily many link variables. Its time derivative can always be written it the following form:

$$\dot{u} = \sum_{j,\mu,k} W_{j,\mu,k,1} (\pm i p_{j,\mu}) W_{j,\mu,k,2} \ , \tag{6.2.18}$$

where $W_{j,\mu,k,1}$ and $W_{j,\mu,k,2}$ are again products of link variables. Using this we get an expressions for $\dot{M}$:

$$\dot{M} = \sum_{j,\mu,k} c_{j,\mu,k} \delta_{x_{j,\mu,k}, y_{j,\mu,k}} \otimes d_{j,\mu,k} \otimes W_{j,\mu,k,1} (\pm i p_{j,\mu}) W_{j,\mu,k,2} \ . \tag{6.2.19}$$

We have replaced the index n from (6.2.16= by the multi-index j, μ, k where j and μ refer to the site and direction of p and k is an additional index which we need because in general there are more than one term corresponding to one p. For $\dot{S}_f$ we now get

$$\begin{aligned}\dot{S}_f &= \phi^\dagger \tfrac{d}{dt}(M^\dagger M)^{-1}\phi = \\ &- \operatorname{tr}\left[PM^\dagger \sum_{j,\mu,k} c_{j,\mu,k}\delta_{x_{j,\mu,k},y_{j,\mu,k}} \otimes d_{j,\mu,k} \otimes W_{j,\mu,k,1}(\pm i p_{j,\mu})W_{j,\mu,k,2} + \text{h.c.}\right] \ .\end{aligned} \tag{6.2.20}$$

Using the cyclicity of the trace we can collect the momenta $(p_{j,\mu})$ on the left side:

$$\dot{S}_f = -\sum_{j,\mu} \operatorname{tr}_c \left[\pm i p_{j,\mu} \sum_k [W_{j,\mu,k,2}\ \operatorname{tr}_{d,s}(PM^\dagger \delta_{x_{j,\mu,k},y_{j,\mu,k}} \otimes d_{j,\mu,k} \otimes W_{j,\mu,k,1}) - \text{h.c.}]\right] \ . \tag{6.2.21}$$

Here we have introduced the notation tr_c for a trace over the color indices only and $\operatorname{tr}_{d,s}$ for a trace over Dirac and site indices.

Putting the Contributions Together

We are now ready to derive our second equation of motion, which, together with the first equation of motion (6.2.2) will define our Hamiltonian evolution. We have prepared all our derivative terms such that the momenta $(p_{j,\mu})$ appear on the left side, we will now pull them out:

$$\begin{aligned}\dot{\mathcal{H}} &= \sum_{j,\mu} \operatorname{tr}_c\ (p_{j,\mu}\dot{p}_{j,\mu}) + \dot{S}_g + \dot{S}_f = \\ &= \sum_{j,\mu} \operatorname{tr}_c\{p_{j,\mu}[\dot{p}_{j,\mu} - f^g_{j,\mu} - f^f_{j,\mu}]\} \ ,\end{aligned} \tag{6.2.22}$$

where we define the *gauge force* $f^g_{j,\mu}$ as

$$\begin{aligned}f^g_{j,\mu} = -\tfrac{1}{3}[&iU_{j,\mu}(\beta_1 V^{pl}_{j,\mu} + \beta_2 V^{re}_{j,\mu} + \beta_3 V^{par}_{j,\mu}) + \\ &(\beta_1 (V^{pl}_{j,\mu})^\dagger + \beta_2 (V^{re}_{j,\mu})^\dagger + \beta_3 (V^{par}_{j,\mu})^\dagger)(-i)U^\dagger_{j,\mu}] \ ,\end{aligned} \tag{6.2.23}$$

and the *fermionic force* $f^f_{j,\mu}$ as

$$f^f_{j,\mu} = \pm i \sum_k [W_{j,\mu,k,2}\ \operatorname{tr}_{d,s}(PM^\dagger \delta_{x_{j,\mu,k},y_{j,\mu,k}} \otimes d_{j,\mu,k} \otimes W_{j,\mu,k,1}) - \text{h.c.}] \ . \tag{6.2.24}$$

To get our second equation of motion we demand that the evolution conserves energy, i.e., $\dot{\mathcal{H}} = 0$. Looking at (6.2.22) we see that for this condition to hold it is sufficient that

$$\dot{p}_{j,\mu} = f^g_{j,\mu} + f^f_{j,\mu} \ . \tag{6.2.25}$$

To ensure that our evolution of the U field leaves it within the group of $SU(3)$ we have to make sure that p is traceless and hermitian during the whole evolution. In order to ensure this we have to ensure that $\dot{p}$ is always hermitian and traceless. By construction our force term $f^g_{j,\mu} + f^f_{j,\mu}$ is hermitian, but in general it is not traceless. We have to make it traceless explicitly. We have some freedom to do that. One simple choice doing this leads us to our *second equation of motion*

$$\dot{p}_{j,\mu} = f^g_{j,\mu} + f^f_{j,\mu} - \tfrac{1}{3}\,\mathrm{tr}(f^g_{j,\mu} + f^f_{j,\mu}) \ . \tag{6.2.26}$$

To see that this will still conserve energy we substitute $\dot{p}_{j,\mu}$ for the right hand side of the above equation into (6.2.22):

$$\begin{aligned}\dot{\mathcal{H}} &= \sum_{j,\mu} \mathrm{tr}_c\{p_{j,\mu}[f^g_{j,\mu} + f^f_{j,\mu} - \tfrac{1}{3}\,\mathrm{tr}(f^g_{j,\mu} + f^f_{j,\mu}) - f^g_{j,\mu} - f^f_{j,\mu}]\} \\ &= \sum_{j,\mu} \mathrm{tr}_c\{p_{j,\mu}[-\tfrac{1}{3}\,\mathrm{tr}(f^g_{j,\mu} + f^f_{j,\mu})]\} = 0 \ .\end{aligned} \tag{6.2.27}$$

That this expression is zero holds because $p_{j,\mu}$ is a traceless matrix and $-\frac{1}{3}\,\mathrm{tr}(f^g_{j,\mu} + f^f_{j,\mu})$ is just a constant. Therefore $\mathrm{tr}_c\{p_{j,\mu}[-\frac{1}{3}\,\mathrm{tr}(f^g_{j,\mu} + f^f_{j,\mu})]\} = 0$.

We summarize what we have achieved:

- We have derived our two equations of motion (equations (6.2.2) and (6.2.26)), which we need for the Hamiltonian evolution. This evolution conserves the energy ($\dot{\mathcal{H}} = 0$) and it evolves the U-field such that its elements remain within the group of $SU(3)$.
- The second equation of motion we have split into a part coming from the gauge action $f^g_{j,\mu}$ and another coming from the fermionic action $f^f_{j,\mu}$. These terms we call fermionic force and gauge force, respectively.
- The relatively simple gauge force $f^g_{j,\mu}$ we have given explicitly.

- The much more complicated fermionic force $f^f_{j,\mu}$ (which comes from our generalized Dirac operator) we have not derived explicitly. Instead we have split the Dirac operator into simple fragments (6.2.16) and written down the derivative of these fragments in a standard form (6.2.18). We can tabularize all the derivative terms in this standard form (using a computer program). Equation (6.2.24) tells us how every one of these terms contributes to the fermionic force $f^f_{j,\mu}$, and therefore to our second equation of motion (6.2.26).

6.3 Mass Preconditioning

In this Section we want to discuss a method introduced in [47] called *Mass Preconditioning*, which is also referred to as *Hasenbusch Trick*. It has been shown that this method can significantly speed up dynamical fermion simulations using HMC [47, 46, 57].

Here we want to adapt this method for M_{CI} and discuss technical issues on how to employ this in an HMC simulation. To do that we need to have at least approximate bounds for the eigenvalues of the operator. We will assume the ideal case of the spectrum lying on the Ginsparg-Wilson circle (see Fig. 3.2), with the mass parameter μ defined as in (3.3.4). If $M_{\mathrm{CI}}(0)$ denotes the massless chirally improved operator then the massive one $M_{\mathrm{CI}}(\mu)$ is:

$$M_{\mathrm{CI}}(\mu) = (1-\mu)\left[M_{\mathrm{CI}}(0) + \frac{\mu}{1-\mu}\right] \ , \tag{6.3.1}$$

with a mass parameter $0 < \mu < 1$. The eigenvalues of $M_{\mathrm{CI}}(\mu)$ are bounded by

$$\begin{aligned} \lambda_{min} &= \mu \ , \\ \lambda_{max} &= 2-\mu \ . \end{aligned} \tag{6.3.2}$$

Therefore the condition number is

$$\kappa = \frac{\lambda_{max}}{\lambda_{min}} = \frac{2-\mu}{\mu} \ . \tag{6.3.3}$$

In the partition function Z (5.1.4) only the determinant of the fermionic matrix appears. Therefore we can safely replace the fermionic matrix, $M = M_{\mathrm{CI}}(\mu)$ in our case, by any matrix with the same determinant.

For instance one could replace it by a matrix of $n \times n$ blocks with $\sqrt[n]{M}$ in the diagonal and zeros everywhere else:

$$A = \begin{pmatrix} \sqrt[n]{M} & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \sqrt[n]{M} \end{pmatrix} . \tag{6.3.4}$$

The condition number for this matrix A is the same as for one diagonal block, in our case:

$$\kappa(A) = \sqrt[n]{\kappa(M)} = \sqrt[n]{\frac{2-\mu}{\mu}} . \tag{6.3.5}$$

In [57] it was proposed to approximate the n-th roots of M by a rational approximation, theoretical arguments are given that using that in an HMC setup can lead to a performance increase, the optimal choice of n, for the two flavor case, was derived to be $n_{opt} = \log(\kappa(M)^2)$.

The original proposal of [47] avoids the need for taking (approximating) roots of the fermionic matrix. The method is laid out for Wilson's fermions in [47]. Here we will work out an analog for a Ginsparg-Wilson operator, like the CI operator. The case of $n = 2$ blocks will be discussed first. Define a matrix B:

$$B = \begin{pmatrix} \tilde{M} & 0 \\ 0 & M\tilde{M}^{-1} \end{pmatrix} , \tag{6.3.6}$$

with

$$\tilde{M} = M + a . \tag{6.3.7}$$

This a is a real constant, its choice will be discussed next. The spectral bounds and condition number for $\tilde{M}$ are

$$\begin{aligned} \lambda_{min}(\tilde{M}) &= \mu + a , \\ \lambda_{max}(\tilde{M}) &= 2 - \mu + a , \\ \kappa(\tilde{M}) &= \frac{2-\mu+a}{\mu+a} . \end{aligned} \tag{6.3.8}$$

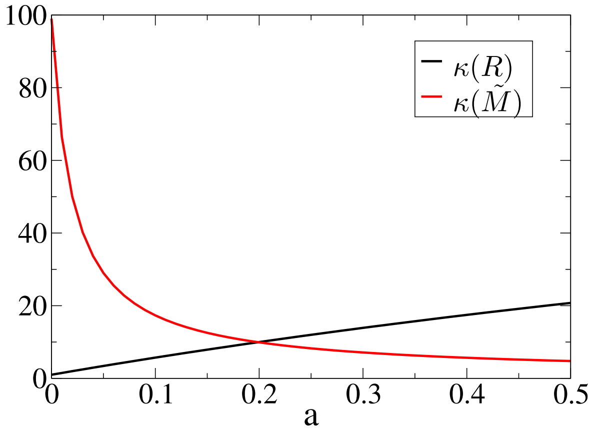

Figure 6.3: The condition number $\kappa(B) = \max\{\kappa(\tilde{M}), \kappa(R)\}$ (6.3.10) takes a minimum where $\kappa(\tilde{M}) = \kappa(R)$. This holds for any $0 < \mu < 1$, for this plot we have set $\mu = 0.02$.

the spectral bounds for $R = M\tilde{M}^{-1}$:

$$\begin{aligned} \lambda_{min}(R) &= \frac{\mu}{\mu + a} \ , \\ \lambda_{max}(R) &= \frac{2-\mu}{2-\mu+a} \ , \\ \kappa(R) &= \frac{(2-\mu)(\mu+a)}{(2-\mu+a)\mu} \ . \end{aligned} \tag{6.3.9}$$

We restrict the calculations to positive mass parameters, $0 < \mu < 1$. For $a + \mu = 0$ the matrix $\tilde{M}$ is singular, therefore we choose $a + \mu > 0$. Under these conditions $\kappa(\tilde{M})$ is monotonically decreasing with a and $\kappa(R)$ is monotonically increasing with a. Therefore the condition number for B which is

$$\kappa(B) = \max\{\kappa(\tilde{M}), \kappa(R)\} \tag{6.3.10}$$

will take a minimum when $\kappa(\tilde{M}) = \kappa(R)$ (see also Fig. 6.3). This equation has one solution for $a + \mu > 0$ at

$$a_{opt} = \frac{\sqrt{-\mu^3 + 3\mu^2 - 2\mu}}{\sqrt{\mu - 1}} = \sqrt{2\mu} + \mathcal{O}(\mu^{\frac{3}{2}}) \ . \tag{6.3.11}$$

For the condition numbers of R and $\tilde{M}$ we get

$$\kappa_{opt}(\tilde{M}) = \kappa_{opt}(R) = \sqrt{\frac{2}{\mu}} + \mathcal{O}(\mu^{\frac{1}{2}}) \ . \tag{6.3.12}$$

In leading order in μ this is the same as $\kappa(A)$ (6.3.4) with $n = 2$.

To generalize this for $n \neq 2$ we define a matrix which we call B again as

$$B = \begin{pmatrix} \tilde{M}_1 & 0 & \cdots & 0 \\ 0 & \tilde{M}_2\tilde{M}_1^{-1} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & M\tilde{M}_{n-1}^{-1} \end{pmatrix} , \tag{6.3.13}$$

with

$$\tilde{M}_i = M + a_i \ . \tag{6.3.14}$$

For the condition numbers of the different non-zero blocks we get

$$\begin{aligned} \kappa_1 = \kappa(\tilde{M}_1) &= \frac{2-\mu+a_1}{\mu+a_1} \ , \\ \kappa_i = \kappa(\tilde{M}_i\tilde{M}_{i-1}^{-1}) &= \frac{(2-\mu)(a_i+a_{i-1})}{(2-\mu+a_{i-1})a_i} \ \ (2 \leq i \leq n-1) \ , \\ \kappa_n = \kappa(M\tilde{M}_{n-1}^{-1}) &= \frac{(2-\mu)(\mu+a_{n-1})}{(2-\mu+a_{n-1})\mu} \ . \end{aligned} \tag{6.3.15}$$

Next we will show that for small μ the a_i can be chosen such that the condition number of B, $\kappa(B) = \sqrt[n]{\kappa(M)}$, the same as for A of (6.3.4). For the a_i we make an educated guess: From our prior experience with $n = 2$ we assume that for small μ : $\mu \ll a_i \ll 2$. Then $\kappa_1 \approx \frac{2}{a_1}$. Setting $\frac{2}{a_1} = \sqrt[n]{\frac{2}{\mu}}$ and solving for a_1 we get $a_1 = 2^{\frac{n-1}{n}}\mu^{\frac{1}{n}}$. We do the same for κ_n : $\kappa_n \approx \frac{a_{n-1}}{\mu}$, set $\frac{a_{n-1}}{\mu} = \sqrt[n]{\frac{2}{\mu}}$ and get $a_{n-1} = 2^{\frac{1}{n}}\mu^{\frac{n-1}{n}}$.

For general $1 \leq i \leq n-1$ we try the ansatz:

$$a_i = 2^{\frac{n-i}{n}}\mu^{\frac{i}{n}} \ . \tag{6.3.16}$$

For all different κ_i we then get

$$\kappa_i = \sqrt[n]{\frac{2}{\mu}} + \mathcal{O}(\mu^0) \tag{6.3.17}$$

(which has been tested using a computer algebra program). Thus we get

$$\kappa(B) = \sqrt[n]{\frac{2}{\mu}} + \mathcal{O}(\mu^0) = \sqrt[n]{\kappa(M)} + \mathcal{O}(\mu^0) \ . \tag{6.3.18}$$

Mass Preconditioning and HMC

Because $\det(B) = \det(M)$ we can use B instead of M in our Hamiltonian (6.2.1). The fermionic part of this Hamiltonian would then be

$$S_f = \phi^\dagger (B^\dagger B)^{-1} \phi \ . \tag{6.3.19}$$

This can be written as a sum:

$$\begin{aligned} S_f = & \phi_1^\dagger (\tilde{M}_1^\dagger \tilde{M}_1)^{-1} \phi_1 \\ & + \sum_{i=2}^{n-1} \phi_i^\dagger [(\tilde{M}_i \tilde{M}_{i-1}^{-1})^\dagger \tilde{M}_i \tilde{M}_{i-1}^{-1}]^{-1} \phi_i \\ & + \phi_n^\dagger [(M \tilde{M}_{n-1}^{-1})^\dagger M \tilde{M}_{n-1}^{-1}]^{-1} \phi_n \ . \end{aligned} \tag{6.3.20}$$

Defining $\tilde{M}_n = M$ this can be written more compactly as

$$S_f = \phi_1^\dagger (\tilde{M}_1^\dagger \tilde{M}_1)^{-1} \phi_1 + \sum_{i=2}^{n} \phi_i^\dagger [(\tilde{M}_i \tilde{M}_{i-1}^{-1})^\dagger \tilde{M}_i \tilde{M}_{i-1}^{-1}]^{-1} \phi_i \ . \tag{6.3.21}$$

We want to take the time derivative of this. For the first term we get, in complete analogy to (6.2.12) :

$$\phi_1^\dagger \tfrac{d}{dt} (\tilde{M}_1^\dagger \tilde{M})^{-1} \phi_1 = -[X_1^\dagger \tilde{M}_1^\dagger \dot{\tilde{M}}_1 X_1 + \text{h.c.}] \ , \tag{6.3.22}$$

where

$$X_1 = (\tilde{M}_1^\dagger \tilde{M}_1)^{-1} \phi_1 \ , \tag{6.3.23}$$

and h.c. stands for hermitian conjugate (of the term left of it). The other terms are more complicated :

$$\begin{aligned} & \phi_i^\dagger \tfrac{d}{dt} [(\tilde{M}_i \tilde{M}_{i-1}^{-1})^\dagger \tilde{M}_i \tilde{M}_{i-1}^{-1}]^{-1} \phi_i = \\ & \phi_i^\dagger \tfrac{d}{dt} [\tilde{M}_{i-1} \tilde{M}_i^{-1} (\tilde{M}_i^\dagger)^{-1} \tilde{M}_{i-1}^\dagger] \phi_i = \\ & - [\phi_i^\dagger \tilde{M}_{i-1} \tilde{M}_i^{-1} \dot{\tilde{M}} \tilde{M}_i^{-1} (\tilde{M}_i^\dagger)^{-1} \tilde{M}_{i-1}^\dagger \phi_i + h.c.] + \\ & + [\phi_i^\dagger (\dot{\tilde{M}}_{i-1} \tilde{M}_i^{-1} (\tilde{M}_i^\dagger)^{-1} \tilde{M}_{i-1}^\dagger) \phi_i + h.c.] \ . \end{aligned} \tag{6.3.24}$$

Because a_i is just a constant, it does not contribute to the time derivative, thus:

$$\dot{\tilde{M}}_i = \tfrac{d}{dt}(M + a_i) = \dot{M} \ . \tag{6.3.25}$$

Using this we can put the result of (6.3.24) into a much more compact form:

$$\begin{aligned}
&\phi_i^\dagger \tfrac{d}{dt}[(\tilde{M}_i\tilde{M}_{i-1}^{-1})^\dagger \tilde{M}_i\tilde{M}_{i-1}^{-1}]^{-1}\phi_i = \\
&-[\phi_i^\dagger \tilde{M}_{i-1}\tilde{M}_i^{-1}\dot{M}\tilde{M}_i^{-1}(\tilde{M}_i^\dagger)^{-1}\tilde{M}_{i-1}^\dagger\phi_i + h.c.]+ \\
&+[\phi_i^\dagger \dot{M}\tilde{M}_i^{-1}(\tilde{M}_i^\dagger)^{-1}\tilde{M}_{i-1}^\dagger\phi_i + h.c.] = \\
&-[X_i^\dagger \tilde{M}_i^\dagger \dot{M} X_i + h.c.]+ \\
&+[\phi_i^\dagger \dot{M} X_i + h.c.] = \\
&-[(X_i^\dagger \tilde{M}_i^\dagger - \phi_i^\dagger)\dot{M} X_i + h.c.] \ ,
\end{aligned} \tag{6.3.26}$$

where

$$X_i = \tilde{M}_i^{-1}(\tilde{M}_i^\dagger)^{-1}\tilde{M}_{i-1}^\dagger\phi_i \ . \tag{6.3.27}$$

6.4 Smearing

Some sort of smearing is part of almost every non-perturbatively improved fermionic action. In all mayor quenched calculations using the chirally improved action the HYP smearing [58] has been used, i.e., the chirally improved operator has been applied to HYP smeared link variables instead of thin link variables. HYP smearing can be seen as extension of APE smearing [59]. We want to discuss the case of isotropic four dimensional APE smearing first. There, as a first step to every link variable U the sum of the associated staples V times some factor ϵ is added:

$$\tilde{U}'_{x,\mu} = U_{x,\mu} + \epsilon V_{x,\mu} \ , \tag{6.4.1}$$

$$V_{x,\mu} = \sum_{\nu\neq\mu}[U_{x,\nu}U_{x+\hat{\nu},\mu}U_{x+\hat{\mu},\nu}^\dagger + U_{x-\hat{\nu},\nu}^\dagger U_{x-\hat{\nu},\mu}U_{x-\hat{\nu}+\hat{\mu},\nu}] \ . \tag{6.4.2}$$

In a second step this is projected into SU(3) (suppress indeces x and μ) $\tilde{U}' \to \tilde{U}$. In traditional APE smearing this projection is defined as $\tilde{U} \in SU(3)$ which maximizes Re tr $\tilde{U}\tilde{U}'^\dagger$. Technically this is done as an iterative process. This iterative procedure renders the differentiation with respect to the thin link variables, which is needed for HMC, difficult if not impossible. HYP smearing consists of three steps of modified APE smearing, which are such that every smeared link variable gets contributions only from link variables within the same hypercube. It is therefore more local than three steps of usual APE smearing. In every of these three steps again the above described projection procedure is used. HYP smearing therefore has the same problems with differentiation as APE smearing.

The problems with the differentiation can be circumvented by replacing the iterative projection procedure by something else. The first version of an almost differentiable projection was proposed and successful used in an HMC setup in [54].

In our project we used the method of stout smearing [55]. More precisely we used four dimensional isotropic stout smearing: The first step is done as described in (6.4.1). To get our smeared link variable $\tilde{U}$ (suppress indices again)

$$\tilde{U} = \exp(iQ)U \tag{6.4.3}$$

where

$$Q = \tfrac{1}{2}(\Omega^\dagger - \Omega) - \tfrac{1}{2\times 3}\mathrm{tr}(\Omega^\dagger - \Omega) \tag{6.4.4}$$

which is a traceless anti hermitian matrix by construction, and

$$\Omega = \epsilon V \tag{6.4.5}$$

The value of ϵ should be chosen such that the plaquette of the resulting smeared link variables becomes a maximum (in this case the smearing takes the maximal effect). After some testing we have chosen a value of $\epsilon = 0.165$ for all our runs.

We now redefine our fermionic action to act on stout-smeared link variables instead of thin link variables. For our Hamiltonian evolution we have then to reinterpret the derivative terms in (6.2.24) as derivatives terms of the action with respect to the smeared link variables. We then need to get the derivatives of these smeared link variables with respect to the thin link variables (since the thin link variables are the ones we want to evolve). For this we want to refer to [55] where these derivatives are worked out in detail for a more general case than ours.

Chapter 7

Simulating QCD with two Dynamical Flavors of CI Fermions

7.1 Introduction

In the last chapter the theoretical framework of dynamical chirally improved fermion simulations with the Hybrid Monte Carlo algorithm has been discussed. After developing these techniques we have implemented them and started to do simulations. The aim of this chapter is to discuss various aspects of these simulations. Some results have also been published in Ref.s [12, 11, 10].

7.2 Overview

We want to start this chapter with an overview over the simulations performed. Table 7.1 contains the simulation parameters and some basic measured quantities of all these simulations. The parameter determining each simulation are the lattice size, the parameters $\beta_1, \beta_2, \beta_3$ for the Lüscher-Weisz gauge action, and the quark-mass parameter (in lattice units) am (for the exact definition of the mass parameter see Section 7.4.1). Lattice sizes are $8^3 \times 16$ and $12^3 \times 24$.

lattice size $12^3 \times 24$

#	β_1	β_2	β_3	$a\,m$	$a\,m_{\mathrm{AWI}}$	HMC time	$\langle plaq \rangle$	a_S[fm]
a	5.2	-0.408	-0.00552	0.02	0.025(1)	463	0.490	0.115(6)
b	5.2	-0.408	-0.00552	0.03	0.037(1)	363	0.489	0.125(6)
c	5.3	-0.418	-0.00575	0.04	0.037(2)	438	0.496	0.120(4)
d	5.3	-0.419	-0.00576	0.05	0.050(1)	302	0.493	0.129(1)
e	4.7	-0.389	-0.00583	-0.05	0.023(2)	386	0.462	0.147(18)

lattice size $8^3 \times 24$

#	β_1	β_2	β_3	$a\,m$	HMC time	$\langle plaq \rangle$	a_S[fm]
aa	5.3	-0.419	-0.00579	0.05	1245	0.493	0.135(3)
bb	5.4	-0.423	-0.00574	0.05	649	0.502	0.114(3)
cc	5.4	-0.427	-0.00427	0.08	776	0.497	0.138(3)

Table 7.1: Summary of the simulations performed

The statistics gathered is expressed via the total HMC time. The basic measured quantities in Table 7.1 are the expectation value of the plaquette, the lattice spacing (determined using the Sommer parameter [60]) and the AWI quark mass (see also Section 7.4.1).

It is expected that, depending on the quark action used, dynamical quarks renormalize the gauge, leading to an effectively smaller gauge coupling g (see for instance [61]). The lattice spacing is expected to decrease when going to smaller quark masses (keeping the parameters of the gauge coupling constant). The dependence of the lattice spacing on gauge coupling and quark mass for our runs is summarized as a diagram in Fig. 7.1.

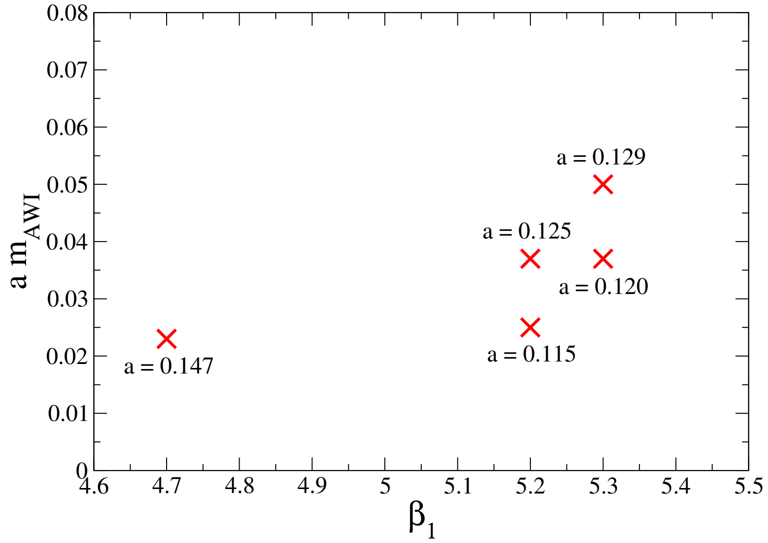

Figure 7.1: Change of the lattice spacing a (given in fm), as a function of the AWI quark mass $a\, m_{\mathrm{AWI}}$ and the LW gauge action parameter β_1, for lattices of size $12^3 \times 24$.

7.3 Fixing the Parameters for the Tadpole Improved Lüscher-Weisz Gauge Action

The chirally improved operator, as we are using it, has been optimized and tested with quenched gauge configurations, generated with the tadpole improved Lüscher-Weisz gauge action [62]. We decided to use this action for our dynamical simulations too.

The Lüscher-Weisz action is discussed in Section 6.2.1, where also the molecular dynamics equations of motion are derived. What is not discussed there is how the coefficients β_1, β_2 and β_3 of (6.2.4) are determined. For this we follow the method of [62]: The coefficient β_1 is chosen to be the independent gauge coupling, and the other two, β_2 and β_3, are determined from tadpole-improved perturbation theory. The idea is to calculate β_2 and β_3 self-consistently from

$$u_0 = \left(\frac{1}{3} \mathrm{Re}\ \mathrm{tr} \langle U_{\mathrm{pl}} \rangle \right)^{\frac{1}{4}} \;, \quad \alpha = - \frac{1}{3.06839} \log \left(u_0^4 \right) \;, \tag{7.3.1}$$

as

$$\beta_2 = \frac{\beta_1}{20\,u_0^2}\,(1 + 0.4805\,\alpha)\ , \quad \beta_3 = \frac{\beta_1}{u_0^2}\,0.03325\,\alpha\ . \tag{7.3.2}$$

At this point we want to distinguish between the *assumed plaquette* and the *measured plaquette* : The assumed plaquette (the term plaquette is used as an abbreviation for $\mathrm{Re}\,\mathrm{tr}\langle U_{\mathrm{pl}}\rangle$) is what enters (7.3.1). The measured plaquette is what is measured in an actual simulation. Self consistency is achieved if these two match. For our further discussion it is very practical to parameterize the LW action via β_1 and the assumed plaquette. The demanded matching can be achieved using an iterative procedure: One starts to simulate at some β_1 and assumed plaquette and measures the plaquette for this simulation. Then one does the simulation with the same β_1 but with the assumed plaquette set to the measured one of the previous simulation, and measures the plaquette again. This procedure has to be redone until assumed and measured plaquette converge. In quenched simulations this has been done with high precision for a range of β_1 values [63].

For dynamical simulations such exact self-consistency is much more difficult to achieve than in the quenched case. In dynamical simulations the re-determination of the assumed plaquette has to be done not only for every β_1 but also for every different $a\,m$. This would mean that we would have to redo the iterative procedure described above for each of our runs. Of course the dynamical simulations are expensive and an iterative procedure to determine the values with a precision comparable to that achieved in the quenched case ([63]) is not feasible for us: we can simply not afford to generate several ensembles of dynamical configurations just for that.

It should also be mentioned here, that the amount of improvement one expects from an improved gauge action is less than for a quenched simulation anyway. Improved gauge actions like the tadpole improved Lüscher-Weisz action start from Wilson's gauge action which is formulated in terms of 1 × 1 loops and longer loops to remove $\mathcal{O}(a^2)$ artefacts (see also Section 2.5). For these improvements, it is assumed that there are no other gauge loops than the ones coming from the gauge action, which is the case in quenched simulations. On the other hand, any fermionic action used in a dynamical simulation will introduce gauge loops of various lengths. These longer loops

lattice size $12^3 \times 24$

#	β_1	β_2	β_3	assumed plaquette	measured plaquette	difference [% m. pl.]
a	5.2	-0.408	-0.00552	0.500	0.490	2.0 %
b	5.2	-0.408	-0.00552	0.500	0.489	2.2 %
c	5.3	-0.418	-0.00575	0.495	0.496	0.2 %
d	5.3	-0.419	-0.00576	0.494	0.493	0.2 %
e	4.7	-0.389	-0.00583	0.460	0.462	0.4 %

lattice size $8^3 \times 24$

#	β_1	β_2	β_3	assumed plaquette	measured plaquette	difference [% m. pl.]
e	5.3	-0.419	-0.00579	0.493	0.493	0.0 %
f	5.4	-0.423	-0.00574	0.500	0.502	0.4 %
g	5.4	-0.427	-0.00427	0.493	0.497	0.8 %

Table 7.2: Adjustment of the parameters of the tadpole improved Lüscher-Weisz gauge action. Ideally the values of the assumed plaquette and measured plaquette should match.

will certainly modify some of the improvement effects from the carefully for the quenched case adjusted improved action. Therefore, in dynamical simulations, one cannot expect as much from improved gauge actions, as one would expect in quenched simulations.

Because of the points discussed above, we decided to adjust the assumed plaquette only approximately, typically readjusting the assumed plaquette once for each run. In Table 7.2 values of the assumed and measured plaquette, and the values of β_2 and β_3 resulting from the quoted assumed plaquette, for the various runs can be found. The assumed plaquette deviates from the measured plaquette by maximally 2.2 % of the measured plaquette in all the different runs.

7.4 Quark Masses

7.4.1 Quark Mass Parameters

We want to start this Section by stressing a similarity between the mass parameters of the chirally improved Dirac operator and Wilson's Dirac operator. For Wilson's operator (2.3.11) one has to adjust the κ parameter according to the underlying gauge configurations. For every ensemble of gauge configurations (generated according to a quenched or full QCD partition function) there is a critical κ_c at which the quarks get massless. If one knows κ_c one can introduce a quark mass using the κ parameter via the relation $a\, m_q = \frac{1}{2\kappa} - \frac{1}{2\kappa_c}$.

In the case of the chirally improved Dirac operator there exist two parameters called z_s and z_v (acting in the scalar and vector sectors of the operator) which have to be adjusted according to the gauge ensemble in order to get massless quarks [9]. If the z-parameters are properly adjusted the operator becomes massless. One can then introduce a finite quark mass via the relation

$$M_{\text{CI}}(\mu) = (1-\mu)\left[M_{\text{CI}}(0) + \frac{\mu}{1-\mu}\right] . \tag{7.4.1}$$

For this work, however, we use a slightly different mass parameter which we call $a\, m$ and which is introduced via the relation

$$M_{\text{CI}}(a\, m) = M_{\text{CI}}(0) + a\, m. \tag{7.4.2}$$

To see how this corresponds to μ of (7.4.1) we scale $M(\mu)$ by a factor $1/(1-\mu)$. Such a scaling is always allowed, since this introduces factor to the fermionic determinant which is the same for all gauge configurations and thus cancels out of the partition function. We call this scaled operator $M'_{\text{CI}}(\mu)$, and write it down explicitly:

$$M_{\text{CI}}(\mu) = M_{\text{CI}}(0) + \frac{\mu}{1-\mu} . \tag{7.4.3}$$

Comparing equations (7.4.3) and (7.4.2) we find the relation

$$a\, m = \frac{\mu}{1-\mu} = \mu + \mathcal{O}(\mu^2) . \tag{7.4.4}$$

Through this relation the parameter $a\, m$ can always strictly be related to μ. We also want to note that in the range of mass parameters we want to use the difference between $a\,m$ and μ is relatively small, for instance the smallest $a\,m = 0.02$ corresponds to $\mu = 0.0196$ and the largest mass parameter used in our calculations on 12×24 lattices, $a\,m = 0.05$ corresponds to $\mu = 0.047$.

Let us return to the z parameters. Practically, the z-parameters can be adjusted by inspecting the low-lying eigenvalues of the massless Dirac matrix. These eigenvalues lie, more or less exactly, on a circle. The z-parameters are adjusted such that this circle goes through the point 0. This can be done relatively easily in the quenched case. Given an ensemble of configurations the spectrum of $M_{\mathrm{CI}}(0)$ for different z-parameters is calculated. Then the z-parameters for which the circle goes most exactly through zero is used for physical calculations. In the dynamical case, however, the adjustment becomes a time consuming iterative procedure (like the self-consistent determination of the parameters for the tadpole-improved LW gauge action, discussed in Section 7.3). In this case the ensemble of gauge configurations is not only dependent on the gauge action but also on the fermionic action and thus implicitly on the z-parameters themselves. This is why it is not sufficient to generate gauge configurations and adjust z afterwards.

Like in the case of the assumed plaquette which goes into the gauge action we did this only approximately. This of course means that we accept to have some additive quark mass correction. In all our dynamical simulations we used $z_s = z_v = 0.93$. This has been adjusted such that there is only very little additive mass correction for the runs c and d, which have been the first runs on 12×24 lattices we performed. The additive mass correction can be easily read off from Table 7.1, calculating the difference $a\,m_{\mathrm{AWI}} - a\ m$ (the definition of m_{AWI} can be found in Subsection 7.4.2). This turned out to be relatively small for runs $a - d$, between -0.03 and 0.05. For run e we find a significantly larger additive mass correction of about 0.07. We are compensating this by using a negative mass parameter $m = -0.05$ to get an AWI mass in the range of about 0.02 (it turned out to be 0.023(2) see Table 7.1).

Empirically we found a relatively simple, almost linear relation between the plaquette of the stout smeared gauge configuration $\langle P\rangle_{smear}$ and the additive mass correction. This is useful for choosing the mass parameter such that it compensates this additive mass correction, therfore it is plotted

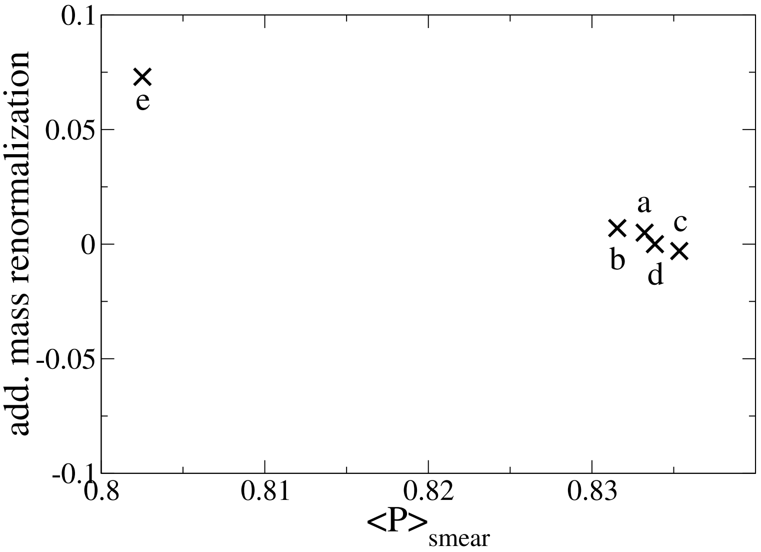

Figure 7.2: The additive mass correction for the CI operator with $z_s = z_v = 0.93$ for our runs $a - e$, plotted again the plaquette measured after stout smearing the underlying configurations.

in Fig. 7.2.

7.4.2 AWI Quark Mass

The axial Ward identity (AWI) allows for a definition of the quark mass via the asymptotic behavior of the relation

$$\frac{Z_A}{Z_P}\,\frac{\langle\,\partial_t A_4(\vec{p}=\vec{0},t)\,X(0)\,\rangle}{\langle\,P(\vec{p}=\vec{0},t)\,X(0)\,\rangle} = 2\,Z_m\,m = 2\,m^{(r)}\;, \tag{7.4.5}$$

where X is any interpolator coupling to the pion and Z_A, Z_P and Z_m denote the renormalization factors relating the $\overline{MS}$-scheme at a scale of 2 GeV. The meson interpolators P and A_4 are defined in (7.5.1).

The renormalization factors have been calculated for the quenched case at several values of the gauge coupling and came out close to unit value [64]. We do not know the values for the dynamical case but for the mass values presented here we expect them to be close to unit value, like in the quenched case. We therefore compute the ratio

$$\frac{\langle\, \partial_t A_4(\vec{p}=\vec{0},t)\; P(0)\,\rangle}{\langle\, P(\vec{p}=\vec{0},t)\; P(0)\,\rangle} \equiv 2\, m_{\mathrm{AWI}}\ , \tag{7.4.6}$$

defining the so-called AWI-mass. What we do in practice is to calculate

$$\langle A_4(\vec{p}=0,t)P(0)\rangle\ , \tag{7.4.7}$$

in order to construct

$$\langle \partial_t A_4(\vec{p}=0,t)P(0)\rangle\ . \tag{7.4.8}$$

Ratios involving the lattice derivative $\partial_t A_4$ depend on the way the derivative is taken. Numerical derivatives are always based on assumptions on the interpolating function. Usual simple 2- or 3-point formulas assume polynomials as interpolating functions. We can do better by utilizing the information on the expected sinh-dependence. In fact, we may use this function for local 3-point ($t-1$, t, $t+1$) interpolation and get the derivative at t therefrom. We cannot use correlators like $\langle X(t)\,\partial_t A_4(0)\rangle$ since the source is fixed to the time slice $t=0$ and thus we cannot construct the lattice derivative there. To improve the signal for our correlators we used the method of Jacobi Smearing, which we discuss in Section 7.5. However, (7.4.6) assumes interpolating fields with point quark sources and sinks. The smearing of the sinks introduces a normalization factor relative to the point sinks. These factors have therefore got to be calculated and taken into account. The factors can be obtained from the ratios of, e.g.,

$$c_P = \frac{\langle P(t)P\rangle}{\langle P_s(t)P\rangle}\ , \quad c_A = \frac{\langle A_0(t)P\rangle}{\langle A_{0,s}(t)P\rangle}\ , \tag{7.4.9}$$

where the index s denotes the interpolator built from smeared sources. Taking this into account we find the AWI-mass from plateau values like that shown in Fig. 7.3. The final average was taken in the same interval as was used for the mass analysis and the error was computed using binned jack-knife (see also Section 4.10).

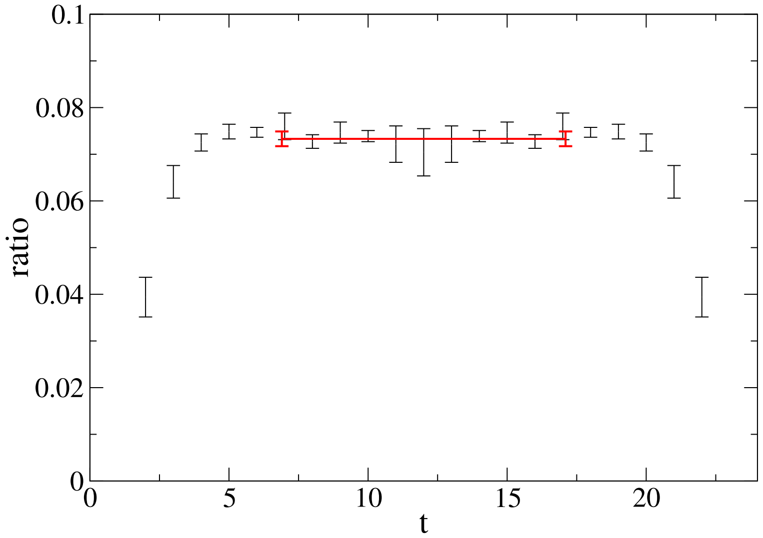

Figure 7.3: The ratio (7.4.6) for $\beta = 5.2$, $a\,m = .03$. The plateau fit range is indicated.

7.5 Mesons

We start this Section by defining the meson interpolators we want to use:

$$\begin{aligned} \text{Pseudoscalar:} \quad P &= \overline{d}\,\gamma_5\,u \;, \\ A_4 &= \overline{d}\,\gamma_5\gamma_4\,u \;, \\ \text{Vector:} \quad V_k &= \overline{d}\,\gamma_k\,u \;, \end{aligned} \tag{7.5.1}$$

where u and d denote the quark mass fields. A_4 denotes the temporal component of the axial vector, which couples to the pion, like P.

We computed the correlation functions

$$\langle P(\vec{p} = 0, t) P(0) \rangle \;, \tag{7.5.2}$$

$$\langle A_4(\vec{p} = 0, t) A_4(0) \rangle \;, \tag{7.5.3}$$

$$\langle V_i(\vec{p} = 0, t) V_i(0) \rangle \;. \tag{7.5.4}$$

The results (see, e.g., Fig. 7.4) were then fitted to

$$C(t) = D(M)\,\left(\mathrm{e}^{-M\,t} \pm \mathrm{e}^{-M\,(T-t)}\right) \;. \tag{7.5.5}$$

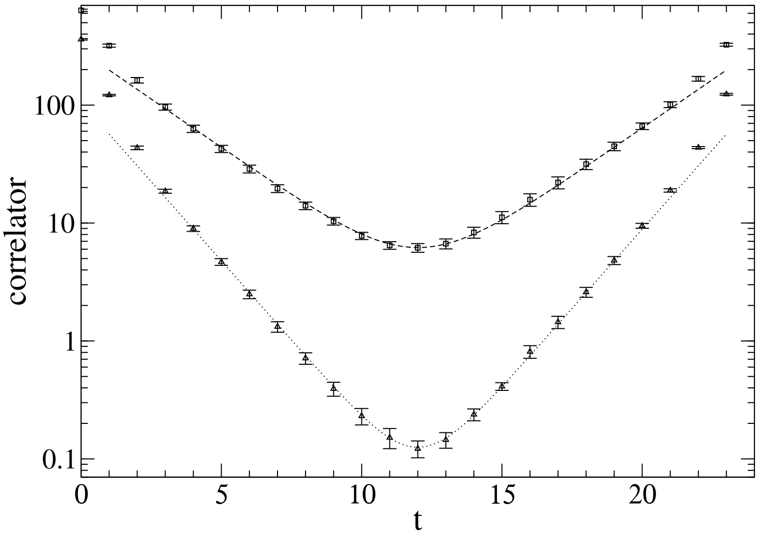

Figure 7.4: Correlation function for run (b), for the pseudoscalar (top) and the vector (bottom) meson. The curves give the results of fits to the cosh-behavior as discussed in the text.

To improve the signal we used the method of Jacobi smearing. This is a method for creating extended quark sources which maintains gauge invariance. Using extended quark sources for the interpolating functions above can lead to improved overlap of these interpolators with wave function of a particle one is interested in, in our case the π and ρ meson. Even though one expects best results when smearing the source and the sink, one can get significant improvement also when smearing only one of them. In our studies we smeared only the sinks. Doing so is less expensive than smearing sources and sinks in our case, because for some of our calculations we need point correlators too (see subsection on the pion decay constant, later in this section and Section 7.4.2). If one wants to use smeared sources and point sources, one has to calculate quark propagators (i.e. inverses of the Dirac matrix) for every different type of source one wants to use. By smearing only the sinks we were able to use quark propagators calculated with point sources for all our calculations. For the definitions and notation for the Jacobi smearing we followed the (quenched) studies in [65, 64]. For the results presented here the narrow smearing distribution was used.

#	aM_π	aM_ρ	M_π/M_ρ	M_π[MeV]	M_ρ[MeV]
a	0.292(10)	0.535(35)	0.55(5)	501(44)	918(109)
b	0.378(8)	0.619(30)	0.61(4)	597(42)	977(96)
c	0.326(18)	0.502(21)	0.65(6)	534(48)	823(62)
d	0.431(8)	0.626(18)	0.69(3)	657(16)	954(33)
e	0.349(3)	0.755(12)	0.46(1)	468(61)	1011(139)

Table 7.3: Meson masses

Meson Masses

The masses of the π and ρ are derived from the exponential decay of $\langle P(\vec{p}=0,t)P(0)\rangle$ and $\langle V_i(\vec{p}=0,t)V_i(0)\rangle$, the corresponding masses appearing as M in the fit function (7.5.5).

Pion Decay Constant

The pion decay constant f_π is related to the coefficient of the $\langle A_4 A_4\rangle$ correlator with point source and sink via

$$Z_A^2\langle A_4(\vec{p}=\vec{0},t)A_4(0)\rangle \sim M_\pi f_\pi^2 e^{-M_\pi t} \ , \qquad (7.5.6)$$

and thus may be extracted from the asymptotic behavior.

Like we did for the AWI masses, again we assume $Z_A \approx 1$. In Fig. 7.5 we show the results, which are compatible with the experimental values. Also like in the case of the AWI mass in this case it is not possible to use smeared sources directly (like we did in the case of the meson masses). If one wants to use smeared sources, one would have to calculate factors similar to the ones we had to calculate for the AWI masses (see (7.4.9)). For this measurement we used the point - point correlation function only, thus avoiding the need to calculate such factors. Table 7.4 summarizes our results in units of the Sommer parameter.

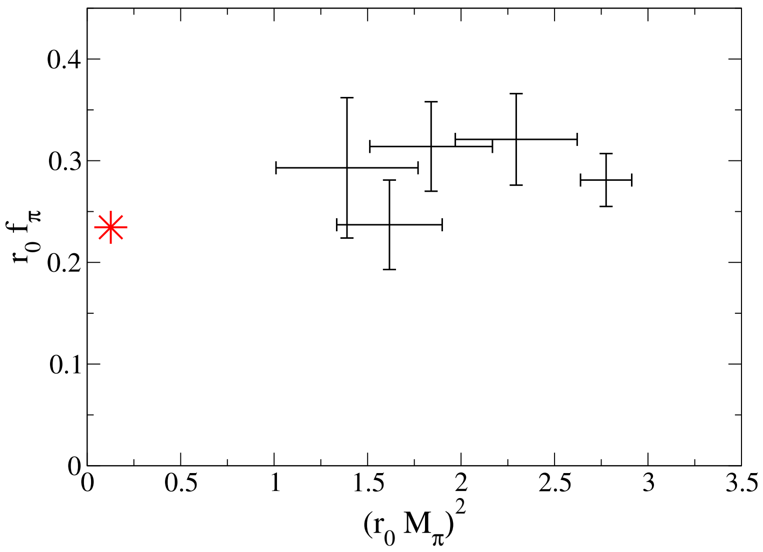

Figure 7.5: f_π vs pion mass squared in units of the Sommer parameter r_0 for all 4 data sets (a–e). The $*$ denotes the experimental value.

The Gell-Mann-Oakes-Renner Relation

In full, renormalized QCD the Gell-Mann-Oakes-Renner (GMOR) relation relates the pion and quark masses:

$$f_\pi^2 \, M_\pi^2 = -2 \, m_q \, \Sigma \, . \tag{7.5.7}$$

Here two flavor of degenerate-mass quarks are assumed. The quark mass and the condensate (contribution per flavor d.f.) are renormalization scheme dependent and have to be given in, e.g., the $\overline{\text{MS}}$-scheme. Since the AWI-mass m_{AWI} is proportional to the renormalized quark mass m_q, a linear relationship like (7.5.7) between the M_π^2 and m_{AWI} may hold. Indeed, in lattice calculations surprisingly linear behavior has been found.

In Fig. 7.6 we plot our results for M_π^2 and m_{AWI} for all four runs. Within the errors the results are compatible with the expected linear dependence. Neglecting the renormalization factors, and taking the experimental value of 92 MeV for the pion decay constant, the slope, via (7.5.7) corresponds to a value for the condensate of $\Sigma = -(288(8)\text{ MeV})^3$. The errors are purely statistical, from the fit to the straight line, neglecting possible higher order

#	am_{AWI}	$r_0^2 M_\pi^2$	$r_0\, m_{AWI}$	$r_0\, f_\pi$
a	0.025(1)	1.62(28)	0.103(9)	0.237(44)
b	0.037(1)	2.29(33)	0.147(10)	0.321(45)
c	0.037(2)	1.84(33)	0.154(13)	0.314(44)
d	0.050(1)	2.78(14)	0.195(6)	0.281(26)
e	0.023(2)	1.39(38)	0.080(17)	0.293(69)

Table 7.4: Results for the AWI-mass in lattice units, and the AWI-mass, M_π and the pion decay constant in units of the Sommer-parameter r_0 (which is usually assumed to be 0.5 fm).

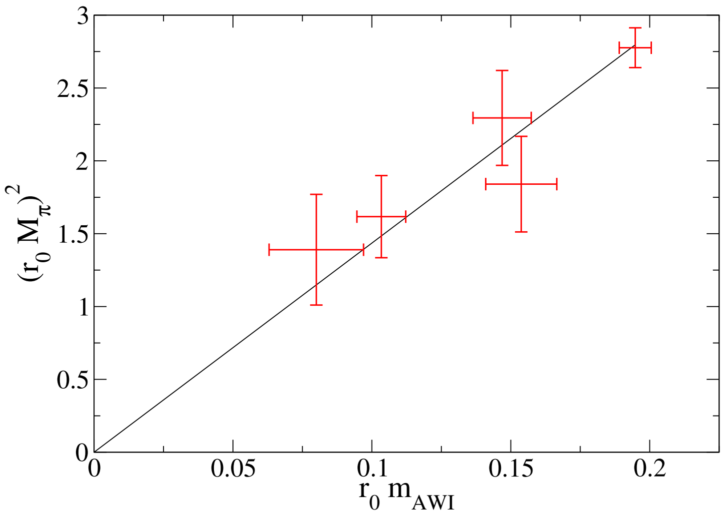

Figure 7.6: The pion mass squared vs. the AWI-mass in units of the Sommer parameter r_0 for all 4 data sets (a–e). The fitted line corresponds to the lowest order chiral perturbation theory behavior, i.e., to the GMOR relation (7.5.7).

chiral perturbation theory contributions.

The GMOR relation also allows for a simple definition of an average light quark mass m_{ud} using the experimental values for f_π and M_π:

$$m_{ud} = \frac{(92\,\mathrm{MeV})^2(140\,\mathrm{MeV})^2}{-2\Sigma}\ . \qquad (7.5.8)$$

Using our value for Σ in this formula gives $m_{ud} = 3.47(29)$ MeV.

7.6 Eigenvalues of the Dirac Matrix

The spectrum of an exact lattice chiral Dirac operator fulfills certain properties, which have been discussed in Chapter 3. The chirally improved operator is an approximate solution to the GW relation in the form of (3.3.1). Because it is only approximate, we have to expect the spectrum of the chirally improved operator to lie on the Ginsparg-Wilson circle, as in Fig. 3.2, only approximately. From the opposite side, we can learn about how well the operator fulfills the GW relation by looking at its spectrum and compare it to the ideal case of the exact GW circle of Fig. 3.2.

In Fig. 7.7 and 7.8 we plot the lowest eigenvalues, computed on 7 configurations each, from runs a and e. We plot them together with the GW circle of Fig.3.2, where for the mass parameter μ we take the measured AWI mass. Runs a and e have been chosen for this comparison because they have similar AWI masses, but quite different lattice spacings. One can see, that for run a, which has the smaller lattice spacing, the spectrum lies much closer to the Ginsparg-Wilson circle.

There is one very direct and important consequence resulting from the spectral properties of a GW operator which we want to stress in some more detail. At any finite mass, a GW operator can never become singular, there are no eigenvalues lower then μ (see again Fig. 3.2). On the other hand, using lattice Dirac operators like Wilson's operator or the chirally improved operator one always has to deal with such singularities at finite masses. These singularities can of course be avoided by choosing large enough quark masses. On the other hand they also disappear in the continuum limit. In this respect the singularities set a certain limit to how small masses can be simulated with a certain action, lattice size and gauge parameters. The situation for dynamical fermion simulations with Wilson's operator and clover improved Wilson's operator at various lattice sizes and gauge couplings was recently discussed in [66]. We want to do a similar discussion for our simulations. We define a *spectral gap* for each gauge configuration and quark mass as the smallest eigenvalue of $M_{\mathrm{CI}}(a\,m)$ (smallest in the sense of the absolute value):

$$\gamma = min\{|\lambda| : \lambda \text{ eigenvalue of } M_{\mathrm{CI}}(a\,m)\} \tag{7.6.1}$$

We can use random matrix theory arguments to sketch the shape of this

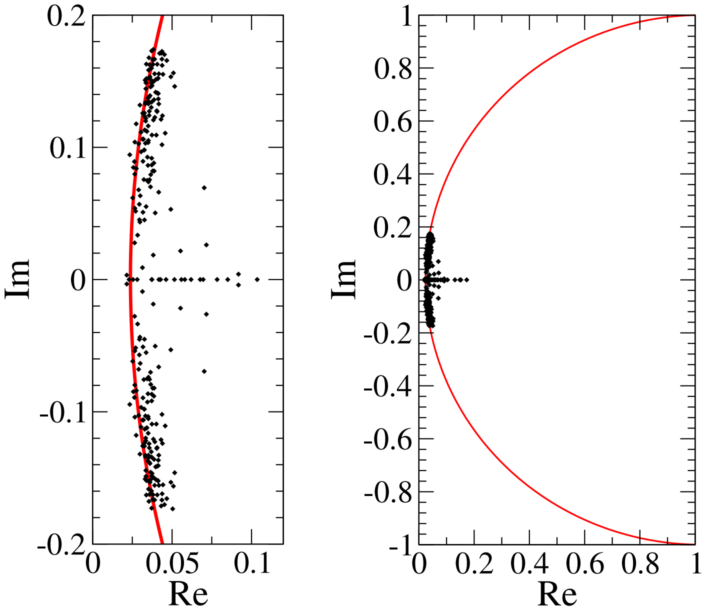

Figure 7.7: The lowest 40 eigenvalues of 7 configurations from run a, together with the Ginsparg-Wilson circle corresponding to the AWI mass. The right picture shows the same as the left picture, but on a larger scale.

distribution for the case of an exact Ginsparg-Wilson operator: For every topologically nontrivial configuration one would get an eigenvalue on the real axis at the mass parameter $a\,m$. In our distribution this would give a delta function at $a\,m$ (if we want to normalize it, we have to multiply this by the fraction of topologically nontrivial configurations in our ensemble). For every topologically trivial configuration γ would be larger than $a\,m$. Together this would give a distribution with a peak at $a\,m$ and a certain tail on the right side of this peak.

In Fig. 7.9 and 7.10 we have plotted histograms and HMC-time evolutions of the lowest eigenvalue for different configurations from our runs a – e. The shape of the distribution shows a form like the one expected for an exact Ginsparg-Wilson operator (see discussion above), but with some distortions. Probably the most interesting aspect of this figure are the eigenvalues lying on the left side of the respective AWI mass. One can see that for the runs a, b and d they are not extending too far to the left of the AWI mass in

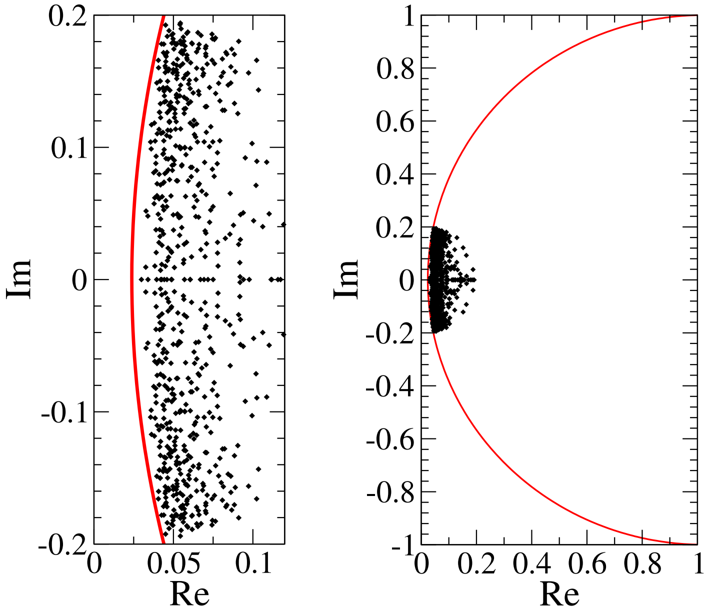

Figure 7.8: The lowest 100 eigenvalues of 7 configurations from run e, together with the Ginsparg-Wilson circle corresponding to the AWI mass. The right picture shows the same as the left picture, but on a larger scale.

all cases. Thus it can be expected that there is still some room for going to smaller quark masses at similar lattice size and lattice spacing without running into troubles with singular configurations. The situation looks a bit different for run e which has a significantly larger lattice spacing than the other runs. There one of the evaluated configurations gives an eigenvalue close to the imaginary axis.

7.7 Topology

In continuum QCD the Atiyah-Singer Index theorem [67] relates the topological charge of a gauge configuration to the null-space of the Dirac operator. Operators obeying the Ginsparg-Wilson relation fulfill a lattice version of the index theorem [68, 69]. A massless Ginsparg-Wilson operator possesses a number of exact zero eigenmodes which have a chirality of either $+1$ or -1. This means that for an eigenvector v_λ to an eigenvalue $\lambda = 0$ the relation

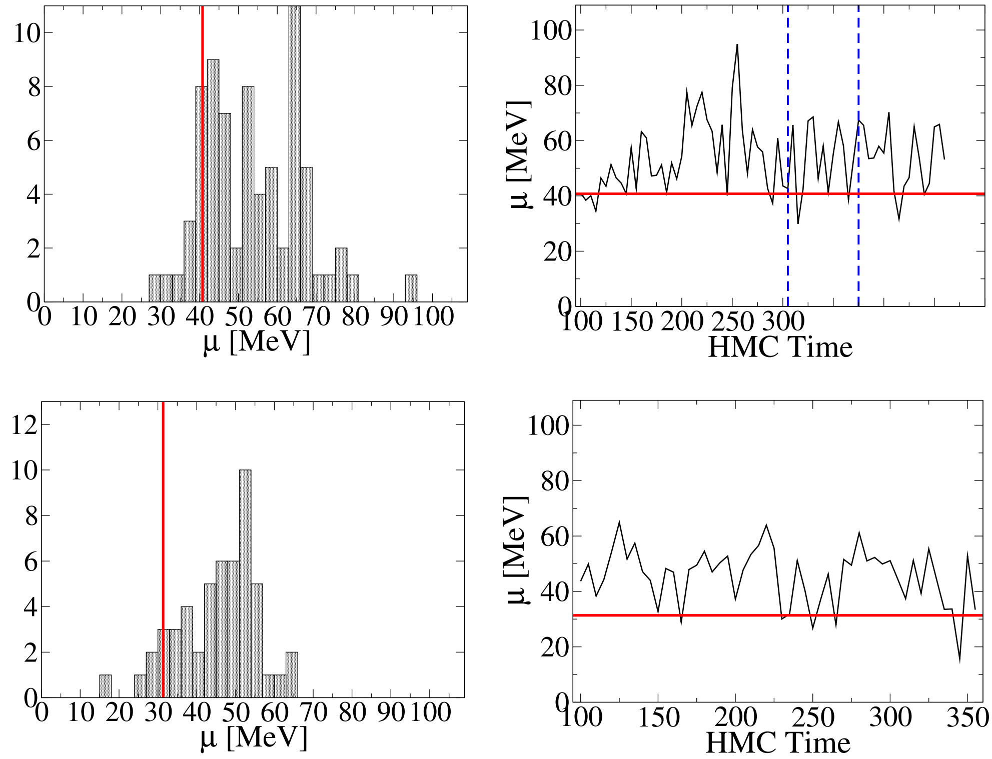

Figure 7.9: Histogram and evolution of the lowest eigenvalue, for runs a, and e. The solid lines in the histograms and evolutions mark the measured AWI masses, the dashed lines in the evolution separate different chains.

$\gamma_5 v_\lambda = \pm v_\lambda$ holds. We call the number of zero modes with positive chirality n_+ and the number of zero modes with negative chirality n_-. Using that, it is possible to define a lattice topological charge ν as

$$\nu = n_- - n_+ \tag{7.7.1}$$

which in the continuum limit converges to the continuum topological charge of

$$\nu = \frac{1}{32\pi^2} \int d^4x \epsilon_{\mu\nu\rho\sigma} (F_{\mu\nu} F_{\rho\sigma}) \ . \tag{7.7.2}$$

For an approximate GW operator like the chirally improved operator, there are no exact zero modes. There are exact real modes, though, which are close to 0. These modes are the only ones with a non-vanishing chirality, which, however, is not exactly ±1. We use these modes to define a topological

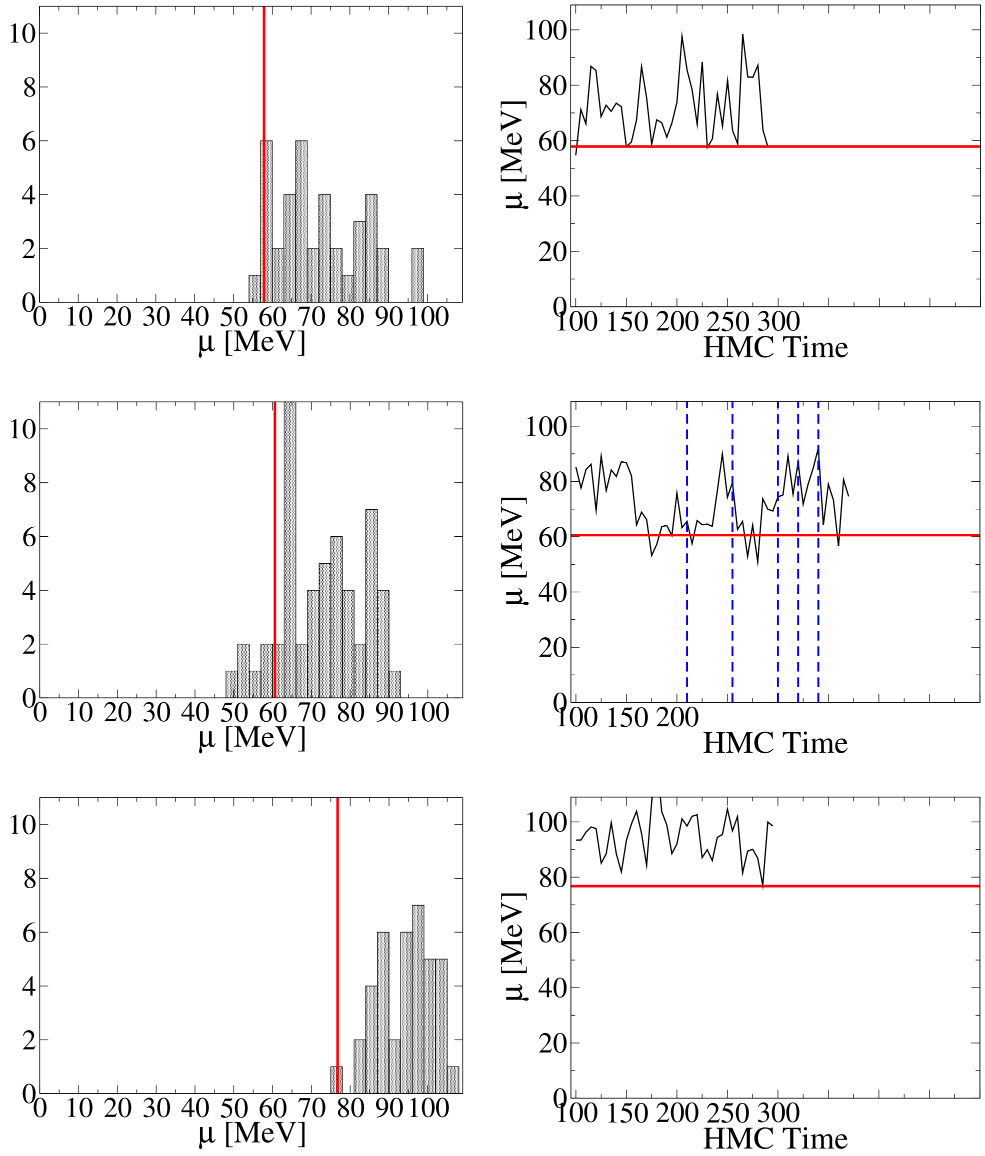

Figure 7.10: Histogram and evolution of the lowest eigenvalue, for runs b, c and d. The solid lines in the histograms and evolutions mark the measured AWI masses, the dashed lines in the evolution separate different chains.

charge as in (7.7.1), where now n_+ counts the number of such modes with positive chirality and n_- the ones with negative chirality. The distribution of the topological charge is expected to be gaussian. In Fig. 7.11 normalized

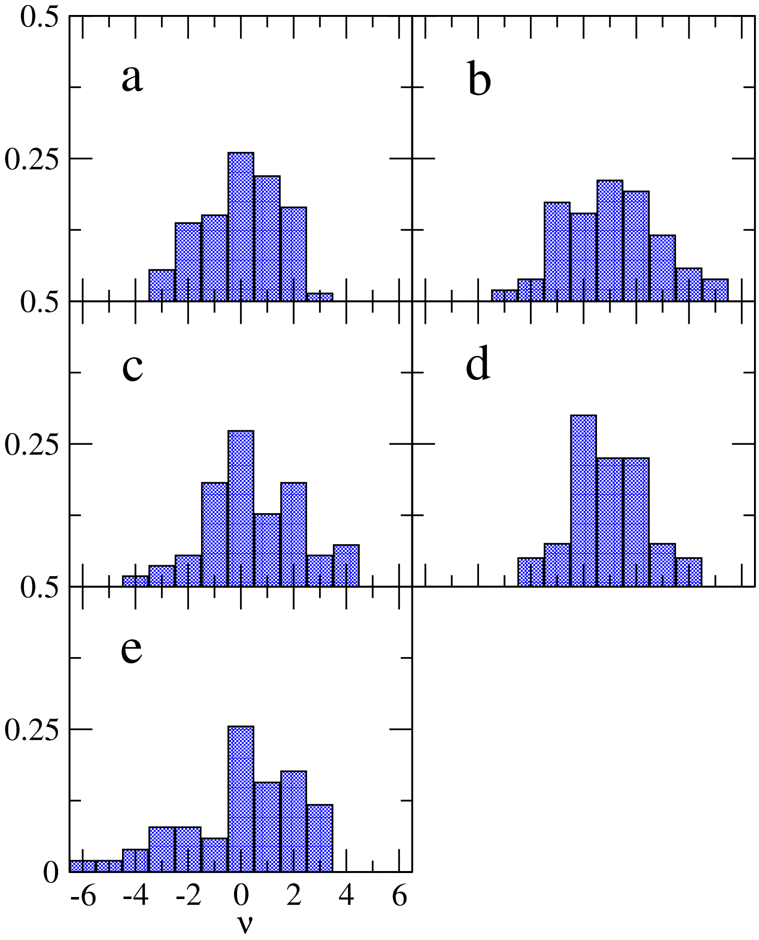

Figure 7.11: Normalized histograms of the topological charge for runs a – e.

histograms of the topological charge for runs a – e are shown.

The topological susceptibility is defined by

$$\chi_{\text{top}} = \left\langle \nu^2 \right\rangle / V \tag{7.7.3}$$

where V is the physical volume of the system. This χ_{top} is related to the chiral condensate Σ, and number of dynamical quark flavors N_f of mass m via [70, 71]

$$\chi_{\text{top}} = -\frac{m\Sigma}{N_f} + \mathcal{O}(m^2) \ . \tag{7.7.4}$$

The topological susceptibility for runs a – e can be found in Table 7.5. According to () the topological charge should depend linearly on the quark mass

#	m_{AWI}[MeV]	χ_{top}
a	40.8(35)	(147 MeV)^4
b	57.9(41)	(150 MeV)^4
c	60.6(50)	(156 MeV)^4
d	76.8(23)	(128 MeV)^4
e	31.4(69)	(147 MeV)^4

Table 7.5: Topological susceptibilities and AWI masses for our runs a – e. The topological susceptibility is expected to be linearly dependent on the AWI mass (7.7).

up to order m^2 corrections. Our data for χ_{top} is too poor to verify such a behavior. Due to the very large autocorrelation length of the topological charge (see Fig. **??**) it is not even possible to estimate error bars. The only thing that we can say that for all our runs the topological susceptibility seems to be significantly smaller than the $\chi_{\mathrm{top}} \approx (190\ \mathrm{MeV})^4$ which was measured in quenched simulations [72, 73, 74].

7.8 Equilibration and Autocorrelation

In Markov Chain Monte Carlo simulations like our the data obtained is always autocorrelated. A certain number of configurations which are correlated to each other obviously contain less information than the same number of uncorrelated configurations would contain. How to quantify this is discussed in Section 4.10. A quantity which is closely related to the autocorrelation time is the Monte Carlo time which is needed to drive the system into equilibrium, starting at a certain point in the configuration space. This is of course not only dependent on the autocorrelation time, but also on the point in configuration space where the Markov Chain is started. We want to look at autocorrelation and equilibration for our simulation in this Section.

In Fig. 7.12 the time histories for different observables from run *b* are plotted. This run has been started from a a gauge configuration with all link variables set to 1. Such a configuration is clearly far out of equilibrium.

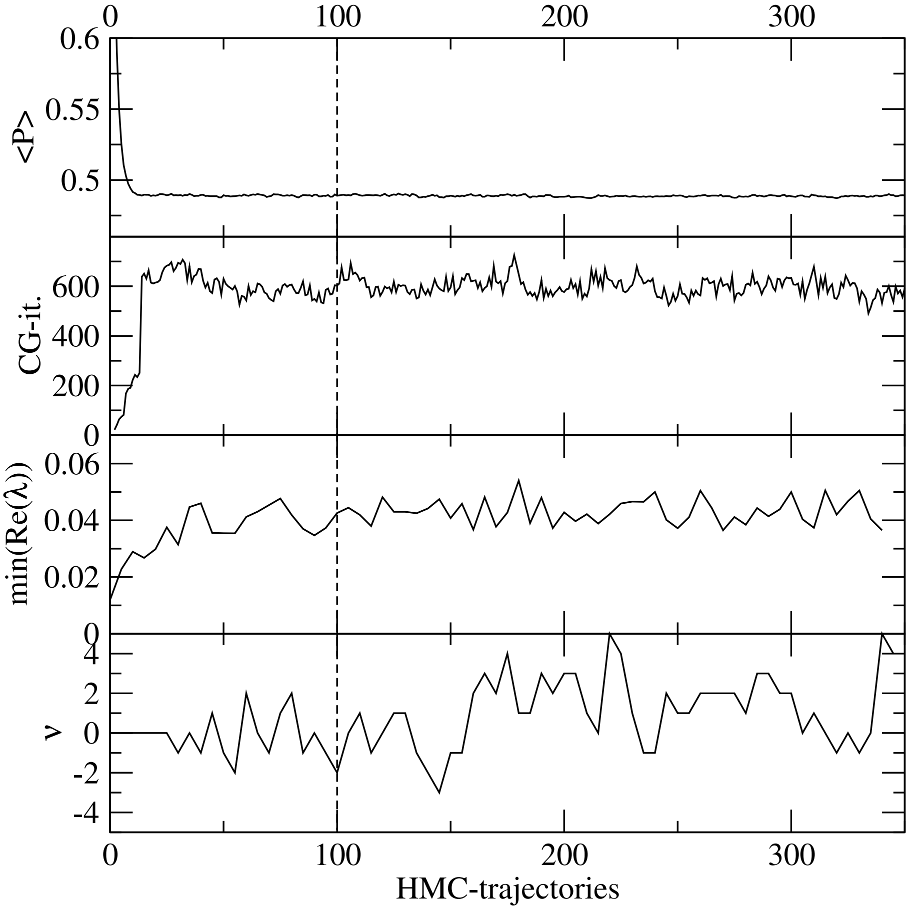

Figure 7.12: Time histories for different observables from run b: Mean plaquette $\langle P \rangle$, conjugate gradient iterations for the accept/reject step, the smallest real part of any eigenvalue min(Re(λ)), the topological charge ν. The quantities min(Re(λ) and ν have been calculated on every fifth configuration only.

Thus one can very well see how the system approaches equilibrium. In Fig. 7.12 we have chosen different observables which show significantly different equilibration behavior and autocorrelation times (and of course also different fluctuations). The plaquette, which is a very local quantity, equilibrates quickly and shows the least autocorrelation of all. The other quantities are all derived from the lower bound of the spectrum of the Dirac matrix which reflects long-distance properties of the system. The number of iterations needed to get convergence of the conjugate gradient procedure in the accept-reject step at the end of each trajectory depends on the condition number of $M^\dagger M$, and thus on the smallest eigenvalue of $M^\dagger M$. By $\min(\mathrm{Re}(\lambda))$ we denote the smallest real value of any eigenvalue for each configuration (the distance of the spectrum from the imaginary axis). The topological charge ν is defined as in (7.7.1).

Unfortunately we have not got enough statistics to reliably calculate autocorrelation times. For the observables of Fig. 7.12 we can estimate the order of magnitude by visual inspection. Looking at the periodicity of the quantities plotted there we would estimate autocorrelation times between 10 and 50 for any of them, where the autocorrelation of the topological charge seems to be the longest.

This analysis gives us a rough idea on the autocorrelation we have to expect for the different physical quantities we want to measure. Based on that we decided to measure on every 5-th configuration. Of course when doing so one has still got to expect some autocorrelation, which has to be taken into account when estimating statistical errors. For doing that we used a binned jackknife method (see also Section 4.10).

Also looking at plots like Fig. 7.12 we decided to discard the first 100 configurations from every run as not equilibrated. This is indicated by a dashed line in this figure.

We want to take a closer look at the topological charge. This quantity is known to show long autocorrelation in dynamical fermion simulations with various actions [75]. The problem is most severe for an exact Ginsparg-Wilson operator, since for such an operator a change of the topological sector means a discontinuity in the Dirac spectrum, and therefore in the fermionic action. Such a discontinuity poses a severe problem for the HMC algorithm, which in its usual form needs an action which is differentiable. For simulations with the overlap operator a method has been developed to circumvent the problem [76]. However, the need for dealing with this problem makes simulations more expensive. For non Ginsparg-Wilson operators like the Wilson operator there is no such discontinuity, but changing a topological sector corresponds to a rather abrupt change of the spectrum of these operators too [77]. For every action, however, the change of topological charge becomes more difficult as the quark mass decreases, since lighter sea quark masses mean stronger contributions from dynamical fermions.

The chirally improved operator is an approximate Ginsparg-Wilson operator. It can be seen as an intermediate between the overlap operator and Wilson's operator. For the problem of changing topological sectors in an HMC simulation there is of course no strict discontinuity like in the case of the overlap operator. On the other hand, it has to be expected, that the change of the spectrum when changing topological sectors is more abrupt than in the case of the Wilson Dirac operator. In Fig. **??** the history of the topological charge is plotted again for run b, together with run a and run e. For both runs b and a the same parameters for the gauge action, but a different mass parameter has been used ($a\,m = 0.03$ for run b and $a\,m = 0.02$ for run a, see also Table 7.1). From these two runs we can thus study how the frequency of the change of topological sectors changes when the quark mass is decreased. It has to be expected, that the tunneling frequency decreases when going to smaller quark masses. In this plot, however, one does not observe a significant decrease in the frequency of the change of topological sectors. In both case such a change occurs about once in 4 units of HMC time on average.

For the two runs a and e the AWI mass in lattice units is very similar ($a\,m = 0.025$ for run a and $a\,m = 0.023$ for run e), but the lattice spacing is significancy larger for run e. While the topological susceptibility is significantly higher for run e, the tunneling frequency seems to be very similar in both cases.

7.9 Performance

Most of the computer time needed for the program, like in other HMC codes for dynamical fermions, is spent for the conjugate gradient solver that is needed in every integration step during the molecular dynamics trajectory. The time needed for this scales linearly with the conjugate gradient steps. The cost for one of these steps is largely determined by the cost of the matrix-vector multiplication. All the other operations that one has to do for the conjugate gradient, are negligible in our case. For this reason we will take the number of matrix-vector multiplications needed as a practical and

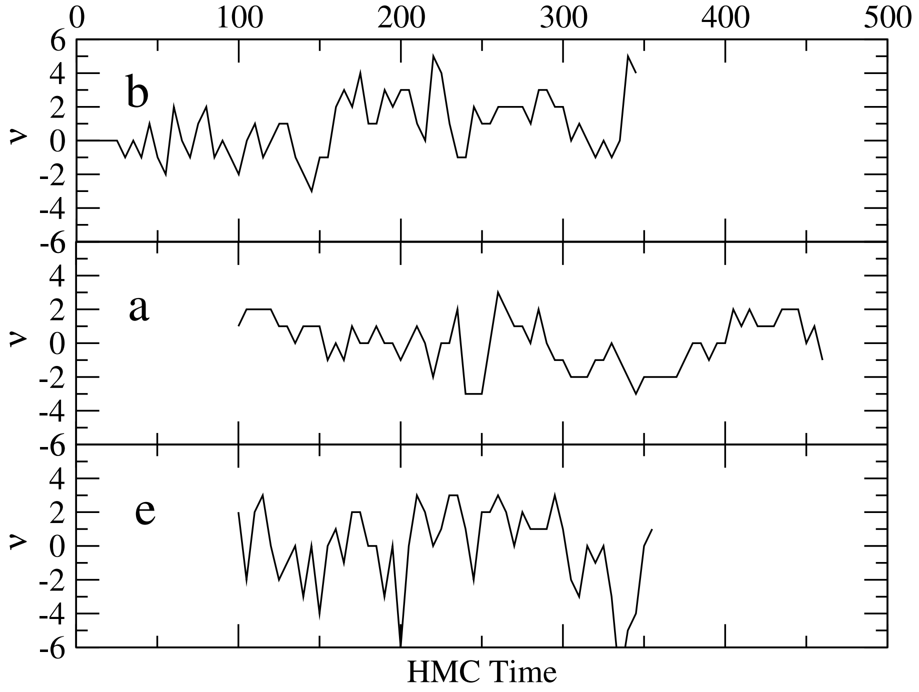

Figure 7.13: Histories of the topological charge for runs a, b and e

machine independent measure for the cost of our algorithm. We tune the parameters of our algorithm such that this number gets small.

The conjugate gradient solver for the molecular dynamics evolution takes most of the computer time, but there are other parts of the program which take a non negligible part of the computer time too. In our code a "Chronological Inverter by Minimal Residual Extrapolation" [78] is used, where 12 previous solutions are taken into account (see also discussion later in this Section). This takes 12 matrix-vector multiplications per step, in addition to the number of matrix-verctor multiplications needed for the conjugate gradient inversion. There is one more conjugate gradient inversion in the accept-reject step at the end of each trajectory. Because it is done only once per trajectory it contributes little, typically about 2 % of the whole computer time was needed for this in our runs on $12^3 \times 24$ lattices. A bit more expensive is the calculation of the force term according to (6.2.24) and (6.2.23), where the cost of the fermionic force term dominates, and the smearing (see

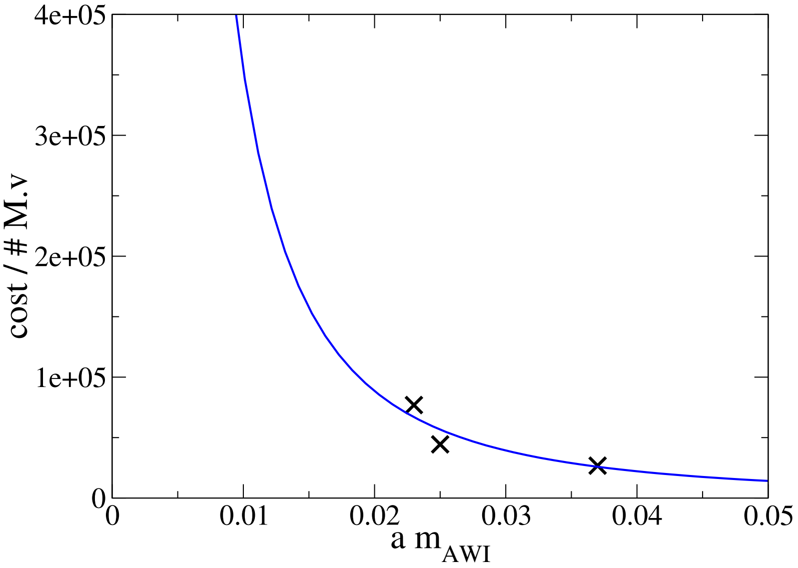

Figure 7.14: The cost of of the conjugate gradient inversion in the molecular dynamics evolution plotted against the AWI mass for our runs a,b and e. The blue line is a fitted curve proportional to $1/(a\,m_{\mathrm{AWI}})^2$.

Section 6.4). In addition to that we need some time to initialize the chirally improved Dirac operator. Initialization of the operator, force calculation and smearing has to be done in every integration step, and it takes the amount of time which about 90 matrix-vector multiplications would take, depending slightly on the machines used and the level of parallelization.

How does the cost of the individual contributions depend on the simulation parameters? The time needed for initialization, the force terms and smearing is linearly dependent on the number of steps per trajectory. The conjugate gradient solver in the molecular dynamics evolution is dependent on the quark mass, the residual and the number of steps per trajectory. It is supposed to grow with the quark mass like $1/(a\,m_{\mathrm{AWI}})^2$. For our runs such a behavior can be seen in Fig. 7.14. This figure contains only runs a,b and e since the other runs use a different conjugate gradient residual and therefore the cost is not directly comparable in this way. For a given quark mass one tries to minimize it choosing the conjugate gradient residual and the number of steps appropriately. In [78] one can find a detailed discussion on how to

#	steps/traj	δt	cg res	$a\, m_{\mathrm{AWI}}$	acc. rate
a	120	0.008	10^{-7}	0.025(1)	93 %
b	100	0.1	10^{-7}	0.037(1)	94 %
c	100	0.1	10^{-8}	0.037(2)	92 %
d	50	0.015	10^{-8}	0.050(1)	89 %
e	120	0.008	10^{-7}	0.023(2)	89 %

Table 7.6: Parameters for the molecular dynamics evolution, AWI mass and acceptance rate for our runs $a-e$. The horizontal lines separate runs with different parameters for the gauge action (see Table 7.1).

choose time step and residual. It is recommended to choose the time step such that one gets acceptance rates close to one. The authors also discuss that one should not choose the residual too large, in order to avoid violations of the reversibility of the molecular dynamics evolution. In Table 7.6 we have summarized the number of steps per trajectory, cg residual, AWI mass and acceptance rate for our runs $a-e$. The conjugate gradient residual has been 10^{-8} for our first run performed on lattice size $12^3 \times 24$, run d. We have reduced this down to 10^{-7} to reduce the cost for our later runs. This is the only, historical, reason for the different choice of residuals.

In Table 7.7 the average cost per unit of HMC time for our runs $a-e$ is given in terms of matrix vector multiplications equivalent and CPU hours, for a 2.2 GHz Opteron processor. The CPU hours are determined running the program on a cluster of 16 Sun v20z nodes with 2 Opteron processors of 2.2 GHz, multiplying the wall-clock time needed by 32. These CPU hours therefore include also the time for the MPI communication.

As a last point in this Section we want to discuss the choice of the number of previous solutions used for the chronological inverter. For this one has to find out how the number of conjugate gradient iterations falls with increasing number of previous solutions. We have used 12 previous solutions in all our runs. Looking at the figures 7.15 and 7.16 one does not expect any improvement from using more than 12 previous solutions in any of the runs. It actually seems that the optimal number would lie a bit below 12 for our

#	tot. cost # Mv	tot. cost CPU h	cg solver evol.	min. res. extrapol.	cg-solver acc/rej	init., force, smearing
a	57928	224	76.7 %	2.5 %	2.3 %	18.6 %
b	38210	148	70.2 %	3.1 %	3.1 %	23.6 %
c	45166	174	75.1 %	2.7 %	2.3 %	19.9 %
d	33034	128	66.5 %	3.7 %	2.5 %	27.2 %
e	76904	297	81.7 %	1.9 %	2.4 %	14.0 %

Table 7.7: Average cost per unit of HMC time for our runs $a-e$. The total cost is given in number of matrix-vector multiplications equivalent and CPU hours, for a 2.2 GHz Opteron processor. The horizontal lines separate runs with different parameters for the gauge action (see Table 7.1) and different cg residuals (see Table 7.6).

runs (the cost of the chronological inverter increases with the number of previous solutions, taking less than 12 previous solutions one could maybe save a small amount of computer time).

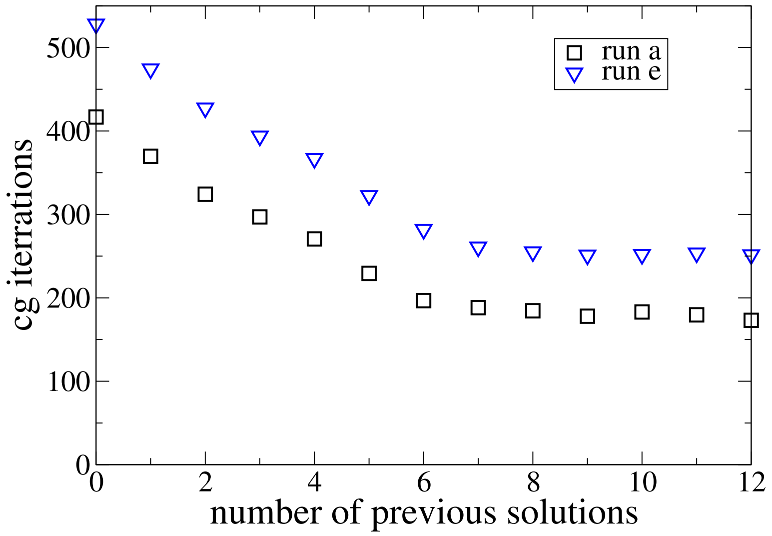

Figure 7.15: The effect of the chronological inverter. The number of average number of conjugate gradient iterations is plotted against the number of previous solutions used. The cg residual taken was 10^{-7} for both runs taken for this figure.

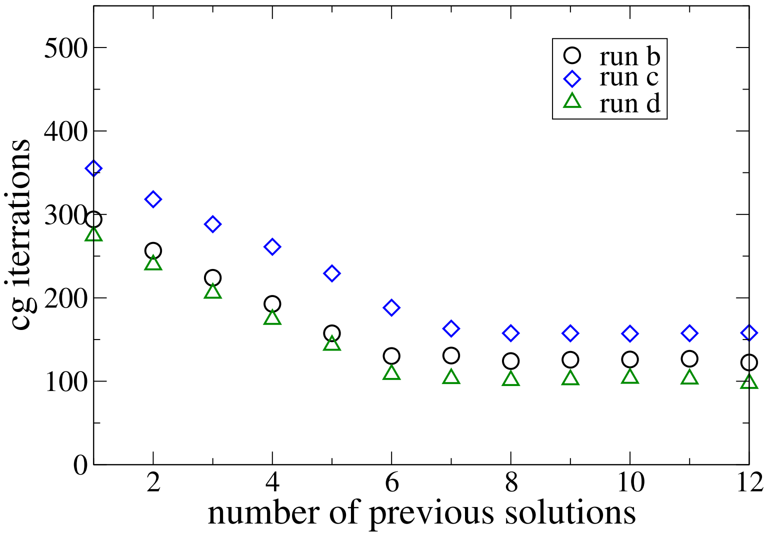

Figure 7.16: The effect of the chronological inverter. The number of average number of conjugate gradient iterations is plotted against the number of previous solutions used. The cg residual taken was 10^{-8} for all three runs taken for this figure.

Chapter 8

Conclusions

This thesis resulted from work on a project with the goal to simulate lattice QCD with light dynamical quarks. To achieve this goal we employ so-called chirally improved fermions. We have discussed various aspects of such simulations in the Chapters 6 and 7. Here we want to summarize what we consider the main conclusions from this work.

At the beginning of this project we had to decide for an algorithm to be used for these simulations. The most successful algorithm in dynamical simulations using various actions is the Hybrid Monte Carlo algorithm. However, employing this for an action like the chirally improved action poses serious technical problems. These problems are serious enough that in order to avoid them a a special algorithm called Partial-Global Algorithm has been suggested. Our experience with that is described in [79]. However, the performance we were able to achieve with that approach was very dissatisfying. For that reason we decided to employ the Hybrid-Monte-Carlo algorithm. Therefore we had to solve the technical problems mentioned above.

One important conclusion we want to draw here therfore is that it is possible to employ the HMC algorithm for an action like the chirally improved action. Chapter 6 describes the various aspects of implementing HMC for an action like that. We found a relatively elegant way of solving the technical problems involved, which we describe in detail Chapter 6. We think that the methods developed there could be useful for for simulations not only with the chirally improved action, but with other, similar, actions too.

After some initial studies on smaller lattices, we simulated lattice QCD with two dynamical flavors of mass-degenerate chirally improved quarks on lattices with $12^3 \times 24$ lattice sites (12 lattice points in the three space dimensions, and 24 in the time direction). These early calculations have been restricted due computer resources. Again, our big goal was to simulate light dynamical quarks. Therefore our discussion in Chapter 7 put a strong emphasis on quark and pion masses. On this lattice size and with the computer resources at hand we reached quark masses down to about 30 MeV which corresponded to a pion mass of about 460 MeV in our calculations. Our calculations seem to indicate that quark masses of about 20 MeV can be reached on lattices of this size. It can be expected that the quark masses can be further reduced in simulations on larger lattices.

These numbers are promising and encouraging. Whether we will really be able to reach these promised quark masses has to be seen in the future.

Acknowledgments

I would like express my special gratitude to Univ.Prof. Dr. C.B. Lang, my supervisor, for teaching my so many aspects of physics, not only, but especially during the time of my doctoral studies.

Very specially I would also like to thank Dr. Pushan Majumdar, for helping me climb not only the mountains of physics. I want to thank him at this point in particular for guiding me at the beginning of my thesis project and afterwards, and for the inspiring work we did together.

I would like to particularly thank Univ.Prof. Dr. Christof Gattringer for supporting this work in many ways.

Certainly I would like to thank all my colleagues and friends at the physics institute of the University of Graz, for a very pleasant time in general, and help not only concerning issues of physics.

Of course I would like to thank my parents, Balder and Brigitte Ortner, who most of all supported me in my entire life, and therefore made this work possible.

Last but not least I want to thank Marion, for all your love and understanding.

Bibliography

[1] H. Fritzsch, M. Gell-Mann, H. Leutwyler, Advantages of the color octet gluon picture, Phys. Lett. B47 (1973) 365–368.

[2] K. G. Wilson, Confinement of quarks, Phys. Rev. D10 (1974) 2445–2459.

[3] H. B. Nielsen, M. Ninomiya, Absence of neutrinos on a lattice. 1. Proof by homotopy theory, Nucl. Phys. B185 (1981) 20.

[4] P. H. Ginsparg, K. G. Wilson, A remnant of chiral symmetry on the lattice, Phys. Rev. D25 (1982) 2649.

[5] D. B. Kaplan, A Method for simulating chiral fermions on the lattice, Phys. Lett. B288 (1992) 342–347.

[6] T. A. DeGrand, A. Hasenfratz, P. Hasenfratz, P. Kunszt, F. Niedermayer, Fixed-point action for fermions in QCD, Nucl. Phys. Proc. Suppl. 53 (1997) 942–944.

[7] H. Neuberger, Exactly massless quarks on the lattice, Phys. Lett. B417 (1998) 141–144.

[8] C. Gattringer, A new approach to Ginsparg-Wilson fermions, Phys. Rev. D63 (2001) 114501.

[9] C. Gattringer, I. Hip, C. B. Lang, Approximate Ginsparg-Wilson fermions: A first test, Nucl. Phys. B597 (2001) 451–474.

[10] C. B. Lang, P. Majumdar, W. Ortner, Dirac eigenmodes in an environment of dynamical fermions, Contributed to Sense of Beauty in Physics:

Miniconference in Honor of Adriano Di Giacomo on his 70th Birthday, Pisa, Italy, 26-27 Jan 2006, hep-lat/0512045 .

[11] C. B. Lang, P. Majumdar, W. Ortner, QCD with two dynamical flavors of chirally improved quarks, Phys. Rev. D73 (2006) 034507.

[12] C. B. Lang, P. Majumdar, W. Ortner, First results from dynamical chirally improved fermions, PoS LAT2005 (2005) 131.

[13] C. B. Lang, P. Majumdar, W. Ortner, Implementing hybrid Monte Carlo with stout-smeared chirally improved Dirac operators, PoS LAT2005 (2005) 124.

[14] J. Zinn-Justin, Quantum field theory and critical phenomena, Oxford university press, 2002, international series of monographs on physics.

[15] K. G. Wilson, Quarks and strings on a lattice New Phenomena In Subnuclear Physics, Part A, ed. A.Zichichi, Plenum Press, New York, 1977, p.69.

[16] A. Hasenfratz, F. Knechtli, Flavor symmetry and the static potential with hypercubic blocking, Phys. Rev. D64 (2001) 034504.

[17] T. A. DeGrand, Tests of hypercubic fermion actions, Phys. Rev. D58 (1998) 094503.

[18] S. Durr, C. Hoelbling, U. Wenger, Filtered overlap: Speedup, locality, kernel non-normality and Z(A) approx. 1, JHEP 09 (2005) 030.

[19] H. B. Nielsen, M. Ninomiya, Absence of neutrinos on a lattice. 2. Intuitive topological proof, Nucl. Phys. B193 (1981) 173.

[20] M. Lüscher, Exact chiral symmetry on the lattice and the Ginsparg-Wilson relation, Phys. Lett. B428 (1998) 342–345.

[21] H. Neuberger, More about exactly massless quarks on the lattice, Phys. Lett. B427 (1998) 353–355.

[22] R. G. Edwards, U. M. Heller, R. Narayanan, A study of chiral symmetry in quenched QCD using the overlap-Dirac operator, Phys. Rev. D59 (1999) 094510.

[23] H. Neuberger, Vector like gauge theories with almost massless fermions on the lattice, Phys. Rev. D57 (1998) 5417–5433.

[24] P. Hasenfratz, et al., The construction of generalized Dirac operators on the lattice, Int. J. Mod. Phys. C12 (2001) 691–708.

[25] P. Hasenfratz, F. Niedermayer, Perfect lattice action for asymptotically free theories, Nucl. Phys. B414 (1994) 785–814.

[26] P. Hasenfratz, Prospects for perfect actions, Nucl. Phys. Proc. Suppl. 63 (1998) 53–58.

[27] P. Hasenfratz, et al., Progress using generalized lattice Dirac operators to parametrize the fixed-point QCD action, Nucl. Phys. Proc. Suppl. 94 (2001) 627–635.

[28] C. Gattringer, et al., Quenched spectroscopy with fixed-point and chirally improved fermions, Nucl. Phys. B677 (2004) 3–51.

[29] M. Evans, T. Swartz, Approximating Integrals via Monte Carlo and Deterministic Methods, Oxford University Press, 2000.

[30] J. S. Liu, Monte Carlo Strategies in Scientific Computing, Springer-Verlag New York Inc., 2002.

[31] N. Metropolis, A. Rosenbluth, M. Rosenbluth, A. Teller, E. Teller, Equations of State Calculations by Fast Computing Machines, J.Chemical Physics 21 (1953) 1087.

[32] W. Hastings, Monte Carlo sampling methods using Markov chains and their applications, Biometrika 57,97-109 57 (1970) 97–109.

[33] S. Geman, D. Geman, Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images, IEEE Trans. on Pattern Analysis and Machine Intelligence 6 (1984) 721.

[34] S. Duane, A. D. Kennedy, B. J. Pendleton, D. Roweth, Hybrid Monte Carlo, Phys. Lett. B195 (1987) 216–222.

[35] D. H. Weingarten, D. N. Petcher, Monte Carlo integration for lattice gauge theories with fermions, Phys. Lett. B99 (1981) 333.

[36] R. Barrett, H. van der Vorst, R. Pozo, J. Dongarra, V. Eijkhout, C. Romine, Templates for the Solution of Linear Systems: Building Blocks for Iterative Methods (Miscellaneous Titles in Applied Mathematics Series No 43), Society for Industrial and Applied Mathematics, 1993.

[37] S. A. Gottlieb, W. Liu, D. Toussaint, R. L. Renken, R. L. Sugar, Hybrid Molecular Dynamics algorithms for the numerical simulation of quantum chromodynamics, Phys. Rev. D35 (1987) 2531–2542.

[38] R. Gupta, G. W. Kilcup, S. R. Sharpe, Tuning the Hybrid Monte Carlo algorithm, Phys. Rev. D38 (1988) 1278.

[39] M. Creutz, Global Monte Carlo algorithms for many-fermion systems, Phys. Rev. D38 (1988) 1228–1238.

[40] A. Ukawa, Computational cost of full QCD simulations experienced by CP-PACS and JLQCD Collaborations, Nucl. Phys. Proc. Suppl. 106 (2002) 195–196.

[41] Y. Namekawa, et al., Light hadron spectroscopy in two-flavor QCD with small sea quark masses, Phys. Rev. D70 (2004) 074503.

[42] S. Aoki, et al., Light hadron spectroscopy with two flavors of O(a)-improved dynamical quarks, Phys. Rev. D68 (2003) 054502.

[43] C. R. Allton, et al., Improved Wilson QCD simulations with light quark masses, Phys. Rev. D70 (2004) 014501.

[44] B. Orth, T. Lippert, K. Schilling, Finite-size effects in lattice QCD with dynamical Wilson fermions, Phys. Rev. D72 (2005) 014503.

[45] G. S. Bali, et al., Static potentials and glueball masses from QCD simulations with Wilson sea quarks, Phys. Rev. D62 (2000) 054503.

[46] C. Urbach, K. Jansen, A. Shindler, U. Wenger, HMC algorithm with multiple time scale integration and mass preconditioning, Comput. Phys. Commun. 174 (2006) 87–98.

[47] M. Hasenbusch, Speeding up the Hybrid-Monte-Carlo algorithm for dynamical fermions, Phys. Lett. B519 (2001) 177–182.

[48] D. H. Weingarten, J. C. Sexton, Hamiltonian evolution for the hybrid Monte Carlo algorithm, Nucl. Phys. Proc. Suppl. 26 (1992) 613–616.

[49] M. Lüscher, Schwarz-preconditioned HMC algorithm for two-flavour lattice QCD, Comput. Phys. Commun. 165 (2005) 199–220.

[50] M. Lüscher, Lattice QCD with light Wilson quarks, PoS LAT2005 (2006) 002.

[51] A. Hasenfratz, F. Knechtli, Simulation of dynamical fermions with smeared links, Comput. Phys. Commun. 148 (2002) 81–86.

[52] A. Alexandru, A. Hasenfratz, Partial-global stochastic Metropolis update for dynamical smeared link fermions, Phys. Rev. D66 (2002) 094502.

[53] A. Hasenfratz, P. Hasenfratz, F. Niedermayer, Simulating full QCD with the fixed point action, Phys. Rev. D72 (2005) 114508.

[54] W. Kamleh, D. B. Leinweber, A. G. Williams, Hybrid Monte Carlo with fat link fermion actions, Phys. Rev. D70 (2004) 014502.

[55] C. Morningstar, M. J. Peardon, Analytic smearing of SU(3) link variables in lattice QCD, Phys. Rev. D69 (2004) 054501.

[56] M. Lüscher, P. Weisz, On-shell improved lattice gauge theories, Commun. Math. Phys. 97 (1985) 59.

[57] M. A. Clark, A. D. Kennedy, Accelerating fermionic molecular dynamics, Nucl. Phys. Proc. Suppl. 140 (2005) 838–840.

[58] A. Hasenfratz, F. Knechtli, Flavor symmetry and the static potential with hypercubic blocking, Phys. Rev. D64 (2001) 034504.

[59] M. Albanese, et al., Glueball masses and string tension in lattice QCD, Phys. Lett. B192 (1987) 163.

[60] R. Sommer, A New way to set the energy scale in lattice gauge theories and its applications to the static force and alpha-s in SU(2) Yang-Mills theory, Nucl. Phys. B411 (1994) 839–854.

[61] A. Hasenfratz, T. A. DeGrand, Heavy dynamical fermions in lattice QCD, Phys. Rev. D49 (1994) 466–473.

[62] M. G. Alford, W. Dimm, G. P. Lepage, G. Hockney, P. B. Mackenzie, Lattice QCD on small computers, Phys. Lett. B361 (1995) 87–94.

[63] C. Gattringer, R. Hoffmann, S. Schaefer, Setting the scale for the Luescher-Weisz action, Phys. Rev. D65 (2002) 094503.

[64] C. Gattringer, M. Göckeler, P. Huber, C. B. Lang, Renormalization of bilinear quark operators for the chirally improved lattice Dirac operator, Nucl. Phys. B694 (2004) 170–186.

[65] T. Burch, et al., Excited hadrons from improved interpolating fields, Nucl. Phys. Proc. Suppl. 140 (2005) 284–286.

[66] L. Del Debbio, L. Giusti, M. Lüscher, R. Petronzio, N. Tantalo, Stability of lattice QCD simulations and the thermodynamic limit, JHEP 02 (2006) 011.

[67] M. F. Atiyah, I. M. Singer, The Index of elliptic operators. 5, Annals Math. 93 (1971) 139–149.

[68] P. Hasenfratz, V. Laliena, F. Niedermayer, The index theorem in QCD with a finite cut-off, Phys. Lett. B427 (1998) 125–131.

[69] F. Niedermayer, Exact chiral symmetry, topological charge and related topics, Nucl. Phys. Proc. Suppl. 73 (1999) 105–119.

[70] R. J. Crewther, Chirality selection rules and the U(1) problem, Phys. Lett. B70 (1977) 349.

[71] H. Leutwyler, A. Smilga, Spectrum of Dirac operator and role of winding number in QCD, Phys. Rev. D46 (1992) 5607–5632.

[72] C. Gattringer, R. Hoffmann, S. Schaefer, The topological susceptibility of SU(3) gauge theory near T(c), Phys. Lett. B535 (2002) 358–362.

[73] L. Del Debbio, C. Pica, Topological susceptibility from the overlap, JHEP 02 (2004) 003.

[74] L. Del Debbio, L. Giusti, C. Pica, Topological susceptibility in the SU(3) gauge theory, Phys. Rev. Lett. 94 (2005) 032003.

[75] B. Alles, et al., Scanning the topological sectors of the QCD vacuum with hybrid Monte Carlo, Phys. Rev. D58 (1998) 071503.

[76] Z. Fodor, S. D. Katz, K. K. Szabo, Dynamical overlap fermions, results with HMC algorithm, Nucl. Phys. Proc. Suppl. 140 (2005) 704–706.

[77] B. Alles, G. Cella, M. Dilaver, Y. Gunduc, Testing fixed points in the 2D O(3) non-linear sigma- model, Phys. Rev. D59 (1999) 067703.

[78] R. C. Brower, T. Ivanenko, A. R. Levi, K. N. Orginos, Chronological inversion method for the Dirac matrix in hybrid Monte Carlo, Nucl. Phys. B484 (1997) 353–374.

[79] C. B. Lang, P. Majumdar, W. Ortner, Experiences with dynamical chirally improved fermions, hep-lat/0412016 (2004) .