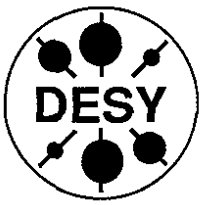
DEUTSCHES ELEKTRONEN – SYNCHROTRON



DESY 92-158
November 1992

Multigrid Methods for the Computation of Propagators in Gauge Fields

T. Kalkreuter
II. Institut für Theoretische Physik, Universität Hamburg

ISSN 0418-9833

NOTKESTRASSE 85 · D-2000 HAMBURG 52

To be sure that your preprints are promptly included in the
HIGH ENERGY PHYSICS INDEX,
send them to the following (if possible by air mail):

DESY Bibliothek Notkestraße 85 W-2000 Hamburg 52 Germany	DESY-IfH Bibliothek Platanenallee 6 O-1615 Zeuthen Germany

DESY 92– 158
November 1992

ISSN 0418 – 9833

Multigrid Methods for the Computation of Propagators in Gauge Fields *

Thomas Kalkreuter **

II. Institut für Theoretische Physik der Universität Hamburg,
Luruper Chaussee 149, W-2000 Hamburg 50, Germany

*Work supported by Deutsche Forschungsgemeinschaft.
**E-mail: I02KAL@DSYIBM.DESY.DE

Abstract

Multigrid methods were invented for the solution of discretized partial differential equations in order to overcome the slowness of traditional algorithms by updates on various length scales. In the present work generalizations of multigrid methods for propagators in gauge fields are investigated. Gauge fields are incorporated in algorithms in a covariant way. This avoids the necessity for gauge fixing in computations of propagators. The kernel C of the restriction operator which averages from one grid to the next coarser grid is defined by projection on the ground-state of a local Hamiltonian (e. g. a block-local approximation of the fermion matrix). The idea behind this definition is that the appropriate notion of smoothness depends on the dynamics. In traditional algorithms the lowest mode of the Hamiltonian is responsible for critical slowing down, and this mode should be represented as well as possible on coarser grids. The ground-state projection choice of C is usable in arbitrary space-time dimension d and for arbitrary gauge group. We discuss proper averaging operations for bosons and for staggered fermions. An efficient algorithm for computing C numerically is presented. The averaging kernels C can be used not only in deterministic multigrid computations, but also in multigrid Monte Carlo simulations, and for the definition of block spins and blocked gauge fields in Monte Carlo renormalization group studies of gauge theories. Actual numerical computations of kernels and propagators are performed in compact four-dimensional $SU(2)$ gauge fields. We prove that our proposals for block spins are "good", using renormalization group arguments. The argumentation uses an "optimal" interpolation kernel $\mathcal{A}$ which is associated with C. This $\mathcal{A}$ is a gauge covariant generalization of a kernel which was used successfully in rigorous works on constructive quantum field theory by Gawędzki and Kupiainen. A central result of the present work is that *the multigrid method works in arbitrarily disordered gauge fields, in principle*. It is proved that computations of propagators in gauge fields without critical slowing down are possible when one uses the ideal $\mathcal{A}$. Unfortunately, the idealized algorithm is not practical, but it was important to answer questions of principle. Practical multigrid methods are able to outperform the conjugate gradient algorithm in case of bosons. The case of staggered fermions is harder. Multigrid methods give considerable speed-ups compared to conventional relaxation algorithms, but on lattices up to 18^4 conjugate gradient is superior. A modification which can be thought of as "updating on a layer consisting of a single site" leads to an improvement of relaxation algorithms in case of bosons. However, we feel unable to predict whether the analogous method pays for staggered fermions on lattices of realizable sizes, because a volume effect remains with respect to how long it takes until errors decay exponentially.

Contents

1 Introduction

The lattice regularization provides presently the only known nonperturbative tool to investigate quantum field theories quantitatively. In the context of quantum chromodynamics (QCD) it was introduced by Wilson in 1974 [1].

Lattice calculations are performed in Euclidean space. Minkowski space field theories [2, 3] are transformed into Euclidean theories [4,5,6,7,8] by analytic continuation to imaginary time. Through this procedure field theory is mapped onto a system of classical statistical mechanics [9]. This mapping becomes manifest in the path integral representation of Green's functions; see e. g. Refs. [10,11] for reviews.

It is desirable to carry through analytic calculations on the lattice. Maybe the most prominent and important result in recent time is the solution of ϕ_4^4 theory by Lüscher and Weisz [12] which yielded a nonperturbative upper bound for the mass of the Higgs particle.

Besides analytic calculations the numerical investigation of lattice field theories is another important branch. Euclidean field theories can be simulated numerically by means of Monte Carlo methods [13,14,15,16]. A collection of "early-days-papers" is contained in Rebbi's book [17]. Further introductory references are [18,19,20,21].

In numerical simulations of lattice field theories one faces the problem of critical slowing down (CSD). Possible continuum limits can only be taken at those points in the phase diagram where phase transitions of second order arise. Expectation values of observables are determined through generating a Marcov chain in a Monte Carlo process. Subsequent field configurations in the Marcov chain are correlated. The correlation time determines, roughly speaking, the number of Monte Carlo sweeps which are necessary to obtain a statistically independent configuration. The problem of CSD consists of the phenomenon that correlation times diverge as critical points are approached.

Numerous algorithms have been tested and new ones developed in order to circumvent, or at least to mitigate, the problem of CSD. A very successful class of algorithms for spin models is the class of cluster algorithms. The first cluster simulation was performed by Swendsen and Wang [22] for the Ising model [23,9] in two and three dimensions. The generalization of their algorithm to $O(N)$ models is due to Wolff [24] and Hasenbusch [25]. Reviews of cluster algorithms can be found in [26].

Other candidates for fighting CSD are multigrid (MG) algorithms. Their basis is the following observation. The slowness of conventional algorithms is due to the fact that field configurations are updated only locally. The slow modes are the long-wavelength modes (i. e. smooth modes in ordered systems). These modes, however, can be treated more efficiently on coarser lattices. Measured in units of lattice spacings, long wavelengths are effectively shortened on coarser scales. Parisi [27] was the first one who pointed out that MG Monte Carlo methods would be promising candidates for tackling CSD.

MG proposals were made by H. Meyer-Ortmanns [28], and by Goodman and Sokal [29]. The latter authors presented the first MG Monte Carlo simulation of two-dimensional ϕ^4 theory. Their approach was to generalize existing deterministic methods to stochastic methods. Simulations in their spirit were performed in Refs. [30,31,32,33].

Mack [34] presented an MG approach to quantum field theory which is inspired by rigorous work in constructive field theory. This approach combines Monte Carlo simulations with analytical tools, namely with ideas from the renormalization group [35,36,37] and with the theory of polymer systems [38,39,40,34,7,8].

Mack's MG proposal has the additional advantage that it yields the effective action (in the sense of the renormalization group) and its derivatives for free [34,41]. Moreover, one can extract infinite volume results from a simulation in a finite volume as shown by Palma [42] for pure ϕ^4-theory. Palma found perfect agreement with the calculations of Lüscher and Weisz [12]. Furthermore, it was pointed out in Ref. [34] that a reformulation of a theory as a polymer system on an MG offers the possibility of performing simulations for continuum systems without imposing a UV-cutoff.

A central point in Mack's MG philosophy is the need to minimize couplings between low and high frequency fields in order to map the original (nearly) critical system onto a noncritical system on the MG. As a consequence interpolation kernels should be smooth in order to fight CSD. *The correctness of this contention was demonstrated in an impressive* way by Hasenbusch and Meyer [43,44].

The use of ideas from the theory of polymer systems in conventional Monte Carlo simulations was proposed by Mack and Pinn [45]. The present author [46] showed how their proposal can be used for more efficient simulations of theories with improved lattice actions [47]. Furthermore, the polymer approach can be viewed as a more general framework for cluster algorithms. A problem which one encounters in polymer algorithms is the possibility of a non-positive-definiteness of probability densities. The present author [48] made a proposal how to deal with systems with complex actions.

The problem of complex actions is also encountered in QCD simulations [49], and this brings us to the problems related with lattice fermions. Fermionic degrees of freedom cannot be simulated straightforwardly because they are represented by Grassmann variables. The common approach makes use of the fact that the fermionic path integrals in QCD are Gaussian, and can thus be performed analytically. The prize paid is the introduction of the fermionic determinant in the remaining bosonic path integral. The fermionic determinant can be rewritten as a path integral over fictitious bosonic "pseudofermion" fields. A review

is given in Ref. [50]. The current status of lattice QCD is summarized in Ref. [51].

The only exact and practical algorithm (up to now) for numerical simulations of field theories involving fermions is the hybrid Monte Carlo algorithm [52]. The most time-consuming part in this algorithm is the computation of propagators of fermions in background gauge fields. Hours of supercomputer time are spent for each matrix inversion [53, p. 1133]. It is thus obvious that improved methods for computing fermionic propagators are desirable. Deterministic MG methods are promising candidates for this task because they have proven efficiency in problems of classical physics [54]. The use of MG methods for quark propagators was proposed by Brower, Moriarty, Myers, and Rebbi [55].

Before we turn to embarking on MG methods, we must discuss the lattice transcription of fermionic fields. The fundamental conceptual problem is that of "species doubling" [56]. A naive discretization of the free fermionic action of one flavor leads to 2^d degenerate flavors in the classical continuum limit (in d space-time dimensions). Actually, the Nielsen-Ninomiya theorem [57] states that on a regular lattice there is no local fermion action without doublers which possesses chiral symmetry and a positive transfer matrix. Smit reviews the problems of lattice fermions in [58]. He characterizes them with the words, "Lattice fermions are like a many headed hydra monster".

Two versions of these monsters are in use today: Wilson fermions [56] and staggered (or Kogut-Susskind) fermions [59,60,61,62,63]. Wilson's method adds a term to the action which gives masses of the order of the cutoff to the unwanted doubler states. This procedure, however, breaks the chiral symmetry of the massless theory completely. Staggered fermions reduce the number of doublers of the naive lattice fermion action by a factor of $2^{d/2}$ while retaining some aspects of chiral symmetry. Therefore one can use staggered fermions to study the spontaneous breakdown of this remaining lattice symmetry. A drawback is, of course, that one always has $2^{d/2}$ degenerate flavors, whereas there is no restriction on the number of flavors in Wilson's formulation (although one has to simulate always an even number of flavors in the hybrid Monte Carlo algorithm for technical reasons). On the other hand, Becher and Joos [64] showed that the correct formal continuum limit of the staggered Dirac operator is equivalent to the Dirac-Kähler operator. The Dirac-Kähler equation is a differential geometric equation for $2^{d/2}$ degenerate flavors already in the continuum, so that there is no doubling when one goes to the lattice.

In this thesis we will be concerned with staggered fermions, mainly for two reasons. Firstly, for the merits of staggered fermions just mentioned. Secondly, for a computational reason. For Wilson fermions successful preconditioning techniques exist [65,66,67,68] which are able to accelerate the inversion of the fermionic matrix quite substantially. However, no such techniques have been proven useful for staggered fermions [69].

We now resume the discussion of computing quark propagators by MG methods. Traditional methods for inversion of the fermionic matrix suffer from CSD when the quark mass is small. CSD sets in already for masses larger than physically interesting values. In connection with deterministic algorithms the term CSD means divergence of relaxation times τ when (inverse) propagators approach criticality. Roughly speaking, τ measures the

amount of work which is necessary to reduce the error of an approximate solution to the propagator by a factor of e. [A kind of intermediate position between direct solvers and (infinite) iteration methods is taken by the conjugate gradient (CG) algorithm. Although the computational labor increases when the quark mass decreases, CG has no well-defined CSD behavior.]

The hope is that the problem of CSD can be avoided in an MG approach. If one had an efficient MG routine for computing quark propagators it could substitute for the conjugate gradient subroutine in existing hybrid Monte Carlo programs of full (unquenched) QCD and also in quenched approximations.

The application of MG methods to gauge theories is, however, not trivial, especially for staggered fermions. In particular the meaning of the central notion of smoothness, on which MG algorithms are based, needs to be clarified in gauge theories. It is not a priori clear if MG methods actually work in gauge theories. The potential obstacle is the disorder which is inherent in non-Abelian gauge fields. Edwards, Goodman, and Sokal [70,71] pointed out that it might be necessary to employ algebraic multigrid (AMG) for propagators in gauge fields. AMG is not based on a geometric conception of smoothness, but a drawback is its more expensive implementation which might make it prohibitive. Also, the existing theory of AMG does not extend immediately to complex matrices.

On the other hand, there is an argument that (geometric) MG has a chance for propagators in gauge fields. The MG procedure presented in this thesis is a gauge covariant generalization of a method which is known to be able to eliminate CSD in the absence of gauge fields. Because of gauge covariance the performance of MG will be the same in arbitrary pure gauges. However, arbitrary pure gauge field configurations appear to be very rough unless they are gauge-fixed. Hence, geometric roughness is not necessarily an obstacle for MG.

Indeed, it will be proved in the present work that *the MG method works in arbitrarily disordered gauge fields*, in principle. Unfortunately, the method implemented for this proof is not useful for production runs because of computational complexity and storage space requirements. But it was important to answer questions of principle.

The MG approach to the computation of propagators of staggered fermions is not trivial also with respect to another point. One of the first questions which has to be answered in MG algorithms is the question for the choice of block lattices. The answer to this question is obvious in case of bosons, but it is not straightforward in case of staggered fermions. Different strategies were pursued in the literature [72,73].

The blocking procedure proposed in the present work starts from the requirement that as much as possible is preserved of internal and space-time symmetries in the limiting case of vanishing gauge coupling. As a result we are forced to choose a scale factor of 3 (or any other odd number). Another consequence is the emergence of seemingly overlapping blocks. In the limiting case of a pure gauge the blocks have actually no sites in common. But in nontrivial gauge fields the symmetries of free staggered fermions are broken, and

for this reason one cannot a priori rule out the possibility that the blocks overlap in a nontrivial way.

In an MG approach one needs a restriction operator C which averages from one grid to the next coarser grid. Two qualitatively different proposals will be made for the choice of C in case of staggered fermions in nontrivial gauge fields. We will call them the "Laplacian choice" and the "Diracian choice". Both proposals reduce to the successful pure gauge construction in the limiting case of vanishing gauge coupling.

The Diracian choice is preferable for physical reasons, but it is computationally quite demanding. The reason is that blocks overlap in a nontrivial way, i. e. every fine grid site makes contributions to the block spin (average) at more than just one coarse grid site. The Laplacian choice is more convenient from the numerical point of view. It retains the property that also in nontrivial gauge fields the seemingly overlapping blocks have actually no sites in common, i. e. every fine grid site makes a contribution to the average at exactly one coarse grid site.

The definition of averaging maps C which will be used in the present work has been given the name "ground-state projection MG" [30,73,74,75]. In this method the kernel of (the adjoint of) C is a projector on the ground-state of a block-local approximation of a Hamiltonian (e. g. a local approximation of the fermion matrix). The idea behind this definition is that the appropriate notion of smoothness depends on the dynamics. In traditional algorithms the lowest mode of the Hamiltonian is responsible for CSD, and this mode should be represented as well as possible on coarser grids. The ground-state projection choice of C is usable in arbitrary space-time dimension d and for arbitrary gauge group. An efficient algorithm for the computation of C will be presented.

We will prove that the Laplacian proposal for block spins of staggered fermions is as "good" as the block spin definition for bosons, using renormalization group arguments. The argumentation uses an "optimal" interpolation kernel $\mathcal{A}$ which is associated with C. This $\mathcal{A}$ is a gauge covariant generalization of a kernel which was used successfully in rigorous works on constructive quantum field theory by Gawędzki and Kupiainen [76]. The criterion for a good choice of C is exponential decay of $\mathcal{A}$. [As a consequence, (exact) effective actions in the sense of the renormalization group will remain local.] The MG method, which proves that CSD in computations of propagators in gauge fields can be eliminated, uses also the generalized $\mathcal{A}$-kernel of Gawędzki and Kupiainen.

Since the Laplacian choice of C defines a good block spin for staggered fermions in nontrivial gauge fields, we will restrict ourselves to the numerical implementation of this proposal. The implementation of the Diracian proposal has to be postponed.

ORGANIZATION OF THE PRESENT PAPER

We will investigate the problems which one encounters in making MG methods usable for gauge theories in a general framework, and we will concentrate on four-dimensional $SU(2)$ gauge fields whenever algorithms are tested numerically. Generalization to $SU(3)$ would be tedious but not difficult.

We begin with a section on deterministic MG methods. Here the problem of CSD in computations of propagators is reviewed. MG algorithms which were invented to circumvent the phenomenon of CSD are introduced. The proper way for including gauge fields in MG algorithms is then discussed. A particularly attractive approach is the ground-state projection MG method. This method is applicable in arbitrary dimension and for arbitrary gauge group. It has the advantage that the procedure is gauge covariant, i. e. no gauge fixing is required. The ground-state projection MG method is based on "block-local Hamiltonians". However, the definition of a block-local Hamiltonian requires the specification of boundary conditions on block boundaries. The necessity of specifying boundary conditions is responsible for the fact that "ground-state projection MG" is not an a priori defined scheme. We will choose "Neumann boundary conditions", and we will discuss this choice in detail. The last three subsections of Sec. 2 will deal with the following. "Updating on a multigrid layer consisting of a single site" will be investigated, and gauge covariant related improvements of MG and conventional relaxation algorithms are considered. None of these improvements needs a tuning of additional parameters. We will argue that updating on single site can be viewed as a global rescaling of propagators. This procedure respects gauge covariance. The section ends with a cautionary remark concerning the term CSD.

In Sec. 3 gauge covariant block spins for bosons and for staggered fermions are discussed. Blocked gauge fields on coarser layers are defined. These block spins may not only be used in either deterministic or stochastic MG computations, but also in Monte Carlo renormalization group studies of gauge theories. An efficient algorithm is presented which allows to compute the integral kernels needed in ground-state projection MG. We prove that our proposals for block spins are "good", using renormalization group arguments. In a final paragraph the notion of "smoothness" in gauge theories is discussed.

Then we turn to actual computations of propagators. Sec. 4 is devoted to the simple case of free propagators (this means arbitrary pure gauge) where the bosonic and the fermionic problems are equivalent. Various MG methods are able to eliminate CSD in this trivial case.

Sec. 5 contains the central result that *the MG method works in disordered systems*. MG computations of propagators without CSD in nontrivial gauge fields are possible when an "optimal" MG scheme is used. This statement is true for *any* value of the gauge coupling, including the case of completely random gauge fields. The optimal MG scheme uses gauge covariant generalizations of integral kernels whose origin are rigorous works in constructive quantum field theory. The success of these computations gives ample evidence from the deterministic side that Mack's contention concerning the need for smooth interpolation kernels is correct.

In Secs. 6 and 7 various practical MG algorithms are tested for bosons and staggered fermions, respectively, in nontrivial gauge fields. The numerical studies reported here are done on lattices of sizes up to 18^4. MG methods give considerable speed-ups compared to conventional relaxation algorithms. In case of bosons the competitor algorithm (conjugate gradient) is outperformed. The case of staggered fermions is harder. The "updating on a layer consisting of a single site" leads to an improvement of relaxation algorithms in case of bosons. We will discuss with special emphasis on gauge covariance how the method can be generalized to the case of staggered fermions. However, on lattices up to 18^4 the method does not pay, and we feel unable to predict from numerical experiments whether the method pays for staggered fermions on larger lattices of realizable sizes, because a volume effect remains with respect to how long it takes until errors decay exponentially.

Sec. 8 gives a summary and an outlook. Two appendices comprise a survey of staggered fermions, and a review of the kernels in the optimal MG algorithm together with their Fourier representations (in the absence of gauge fields) and some tables giving numerical values for integral kernels.

Parts of this thesis have been published before in Refs. [105,112,136,137,106,153].

2 Deterministic Multigrid Methods

The content of this section is as follows. First the notations for propagators in lattice gauge theories are introduced. The computation of propagators amounts to solving a very large linear system of equations. This problem has to be solved frequently in numerical simulations of lattice field theories involving fermions (e. g. QCD). In the physically interesting region inverse propagators have an eigenvalue which is close to zero. The presence of such an eigenvalue makes the use of conventional algorithms unpractical, they suffer from critical slowing down. Multigrid methods were invented to circumvent this phenomenon. Variational multigrid and an "optimal" multigrid algorithm are discussed. The application of multigrid methods to gauge theories requires a proper inclusion of the gauge fields. A particularly attractive approach is the ground-state projection multigrid method. This method is applicable in arbitrary dimension and for arbitrary gauge group. It has the advantage that the procedure is gauge covariant, i. e. no gauge fixing is required. However, we will see that "ground-state projection MG" is not defined a priori, and we will discuss what supplements are needed. Finally we discuss the special case of "updating on a multigrid layer which consists of a single site" for bosons. (The case of staggered fermions will be discussed in Sec. 7.5.) Related improvements from a variational point of view will also be discussed. None of these improvements needs a tuning of additional parameters. This section ends with a cautionary remark concerning the term "critical slowing down".

2.1 Propagators on the Lattice

The notations for gauge covariant propagators of bosons and of staggered fermions on the lattice are introduced.

Notations.

A propagator in a lattice-regularized field theory is the solution of a linear equation

$$D\phi = f \tag{2.1}$$

on a d dimensional hypercubic lattice Λ of sites z and of lattice spacing a, for given f and non-singular matrix D. (Later on in the multigrid context, Λ will also be called Λ^0.) With respect to numerical simulations of (Euclidean) QCD one is interested in fermionic propagators where $D = -\not{D}^2 + m^2$ [52,50,77]. The r.h.s. f of Eq. (2.1) is for instance a

pseudofermion field in simulations with the hybrid Monte Carlo algorithm [52]. In case of *computations of quark correlation functions, f will be a δ-function; then $(-\not{D}+m)\phi$ is the* quark propagator. As a less complicated problem we shall also consider the case of bosons where $D = -\Delta + m^2$. The latter is a good toy for tests of new algorithms per se, and also because a lot of analytical information is available for bosonic propagators. Δ and $\not{D}$ are gauge covariant lattice Laplace or Dirac operators, respectively. They depend on an external gauge field U. In the absence of gauge fields, the Dirac operator is a square root *of the Laplacian, so that $(-\not{D}^2 + m^2) = (-\Delta + m^2)$ in this limiting case.*

In the following we will think of f as being a matter field which is defined on the sites of Λ (e. g. a pseudofermion field). Color indices will always be suppressed, and $\phi(z)$ is an $N_c \times N_c$ matrix where N_c is the number of colors. Hence, (2.1) is actually an equation for a $(|\Lambda| \cdot N_c \times N_c)$ matrix ϕ. Equivalently, one may think of (2.1) as a set of N_c systems for $(|\Lambda| \cdot N_c)$-component vectors.

Discretized covariant Laplace operator.

We use the standard discretization of the covariant Laplacian where Δ has kernel (a = lattice spacing)

$$a^{d+2}\,\Delta(z_1,z_2) = \begin{cases} -2d\,\mathbb{1} & \text{if } z_1 = z_2\ , \\ U(z_1,z_2) & \text{if } z_1 \text{ n. n. } z_2\ , \\ 0 & \text{otherwise}\ . \end{cases} \tag{2.2}$$

Here $\mathbb{1}$ denotes the $N_c \times N_c$ unit matrix, "n. n." *means nearest neighbor, and $U(z_1,z_2)$ is* the gauge field on the link (z_1,z_2). The oppositely orientated link (z_2,z_1) carries the gauge field $U(z_2,z_1) = U(z_1,z_2)^\dagger$, where † denotes the Hermitean conjugate of a matrix.

For a discussion about the lattice transcription of matter and gauge fields, and their naive continuum limit we refer to the textbooks of Creutz and Rothe [20,21].

Eq. (2.2) means that Δ acts on a field ϕ according to

$$\begin{aligned} (\Delta\phi)(z) &\equiv \int_{z'} \Delta(z,z')\,\phi(z') \equiv a^d \sum_{z'\in\Lambda} \Delta(z,z')\,\phi(z') \\ &= \frac{1}{a^2} \sum_{z' \text{ n.n. } z} \left[U(z,z')\,\phi(z') - \phi(z)\right]\ . \end{aligned} \tag{2.3}$$

Discretized covariant Dirac operator.

In case of fermions we shall use the Euclidean staggered lattice formulation [62]. In this case the discretized covariant Dirac operator has kernel[1)]

$$a^{d+1}\,\not{D}(z_1,z_2) = \begin{cases} \eta_\mu(z_1)\,U(z_1,z_2) & \text{if } z_2 = z_1 + \frac{1}{2}e_\mu,\ \mu = -d,\dots,d, \mu \neq 0, \\ 0 & \text{otherwise}\ , \end{cases} \tag{2.4}$$

which means

$$(\not{D}\phi)(z) = \frac{1}{a}\sum_{\mu=1}^{d} \eta_\mu(z)\left[U(z,z+\tfrac{1}{2}e_\mu)\,\phi(z+\tfrac{1}{2}e_\mu) - U(z,z-\tfrac{1}{2}e_\mu)\,\phi(z-\tfrac{1}{2}e_\mu)\right]\ . \tag{2.5}$$

[1)] We use the conventions $\eta_{-\mu} = -\eta_\mu$ and $e_{-\mu} = -e_\mu$.

η_μ are the lattice remnants of the γ-matrices. They are complex numbers of modulus 1, and may be chosen as $\eta_1(z) = 1$, $\eta_2(z) = (-1)^{n_1}$, $\eta_3(z) = (-1)^{n_1+n_2}$, $\eta_4(z) = (-1)^{n_1+n_2+n_3}$, for $z = \frac{a}{2}(n_1, n_2, n_3, n_4)$. Free staggered fermions enjoy discrete translation invariance under shifts by twice the separation of neighboring sites. Therefore we denote the lattice spacing by $a/2$ in this case. e_μ denotes a lattice vector of length a in μ-direction. The notations and conventions which we use in connection with staggered fermions are summarized in Appendix A, see also Sec. 3.2.

We note that $\not{D}$ is anti-Hermitean. Therefore $(\not{D} + m) \cdot (\not{D} + m)^\dagger = (-\not{D}^2 + m^2)$.

Gauge covariance.

An important notion in gauge theories is that of *gauge covariance*. A (local) gauge transformation g is specified on a lattice by a map $g : \Lambda \to G, z \mapsto g(z)$, where G denotes the unitary gauge group. Under a gauge transformation g a matter field ϕ transforms according to

$$\begin{aligned} \phi(z) &\mapsto \phi'(z) = g(z)\,\phi(z) \ , \\ \phi(z)^\dagger &\mapsto \phi'(z)^\dagger = \phi(z)^\dagger\, g(z)^{-1} \ . \end{aligned} \tag{2.6}$$

The transformation law for a parallel transporter $U(z, w)$ from site w to site z is

$$U(z, w) \mapsto U'(z, w) = g(z)\, U(z, w)\, g(w)^{-1} \ . \tag{2.7}$$

In particular, $U(z, w)$ may be a link variable $U(z_1, z_2)$. (Later on we will encounter general parallel transporters $U(z, w)$ which are linear combinations of path-ordered products of link variables along paths $C : w \to z$.)

The discretized partial differential equation (2.1) exhibits gauge covariance, i. e. if ϕ is the solution of (2.1) for given $\{U, f\}$, then $g\phi$ is the solution for the gauge-transformed configuration $\{U', f'\}$.

Boundary conditions.

Finally we have to specify boundary conditions for Δ and $\not{D}$, because in numerical computations all lattices have a finite extension. We shall use periodic boundary conditions in all d directions, for gauge fields U and for bosonic and fermionic propagators. This kind of boundary conditions is used in present day large scale QCD simulations [78].

2.2 Critical Slowing Down of Conventional Algorithms

Classical algorithms for the computation of propagators are discussed. Nearly critical inverse propagators are of practical interest. In nontrivial gauge fields one has to enforce criticality by hand (for bosons, and for staggered fermions on relatively small lattices), in order to study effects of critical slowing down. Criticality is enforced by subtracting the lowest eigenvalue $-m_{cr}^2$ from the discretized elliptic lattice operator D and adding a small mass term Δm^2, i. e. we consider bare masses $m^2 = m_{cr}^2 + Dm$. In case of bosonic propagators there is a well-known mathematical theory about the critical slowing down of

traditional relaxation methods. Due to periodic boundary conditions the critical slowing down depends only on the mass but not on the lattice size. This point must be especially emphasized: One-grid relaxation algorithms for nearly critical bosonic propagators are almost not sensitive to the disorder of the gauge field, and to the lattice size, as long as equal values of Δm^2 are compared. *A kind of intermediate position between direct solvers and (infinite) iteration methods is taken by the conjugate gradient algorithm which will also be described in this section. Although the computational labor increases when the quark mass decreases, the conjugate gradient algorithm has no well-defined critical slowing down behavior. In connection with the conjugate gradient method, we will introduce the gauge invariant energy functional which will play an important role when one considers algorithms from the variational point of view.*

Problems with the solution of Eq. (2.1).

The lattice operator D in Eq. (2.1) is a $(|\Lambda| \cdot N_c \times |\Lambda| \cdot N_c)$ matrix, which is huge in practice. Therefore direct solvers (e. g. Gaussian elimination) cannot be used. One has to employ iterative methods which take advantage of the fact that D is very sparse. However, the convergence of conventional iterative methods is hampered by critical slowing down (CSD). In connection with computations of propagators CSD means divergence of relaxation times when D approaches criticality, i. e. when it has an eigenvalue which is nearly zero. This happens in case of fermions for small quark masses. In the bosonic toy model we have to enforce criticality by hand.

We consider masses $m^2 = m_{cr}^2 + \Delta m^2$, i. e. we consider operators

$$D = (-\Delta + m_{cr}^2 + \Delta m^2) \quad \text{or} \quad D = (-\not{D}^2 + m_{cr}^2 + \Delta m^2) \ , \tag{2.8}$$

where $-m_{cr}^2$ denotes the lowest eigenvalue of the (semi)positive operator $-\Delta$ or $-\not{D}^2$. m_{cr}^2 depends on the particular gauge field configuration U. However, as all other eigenvalues, it is gauge invariant. In nontrivial gauge fields[2] $-m_{cr}^2$ is strictly positive. In case of staggered fermions m_{cr}^2 is a small number which tends to zero as the lattice becomes large, but this is not true for bosons. Small values of Δm^2 have to be chosen in order to make D nearly critical. Critical D are of interest because in practical applications $(-\not{D} + m^2)$ will be nearly critical near the continuum limit.

In this section it will be shown that classical relaxation methods for the computation of bosonic propagators exhibit CSD in any gauge field configuration. This CSD depends only on m^2 and *not* on the lattice size $|\Lambda|$.[3] This is contrarily to the Dirichlet problem, for instance, which is discussed in Ref. [79]. In our case there is only an implicit dependence on $|\Lambda|$ (and $\beta = 2N_c\, a^{d-4}/g^2$ where g is the gauge coupling) through the value of m_{cr}^2. The dimension d enters in the scaling relation for relaxation times only through the constant of proportionality.

[2] A gauge field U is called trivial (or pure gauge) if there is a gauge transformation g such that $g(z_1)U(z_1, z_2)g(z_2)^{-1} = \mathbb{1}$ for all link variables. Therefore this case is equivalent to the absence of gauge fields.

[3] For staggered fermions this result will be found numerically later on.

2.2.1 Classical Relaxation Algorithms

Classical iterative algorithms (see e. g. [80,81,82,83,84]) are (damped) Jacobi relaxation and SOR (successive over-relaxation). SOR includes Gauss-Seidel relaxation as a special case. These traditional algorithms generate a sequence $\{\phi^{(n)}\}$ of approximate solutions of Eq. (2.1). Every iteration essentially involves D only in the context of a matrix-vector multiplication.

Consider for illustration the bosonic problem. In this case damped Jacobi relaxation is defined by[4)]

$$\phi^{(n+1)}(z) = (1-\omega)\,\phi^{(n)}(z) + \frac{\omega}{2d+m^2}\,[f(z) + \sum_{z' \text{ n.n. } z} U(z,z')\,\phi^{(n)}(z')] \ . \tag{2.9}$$

ω is called the relaxation parameter. It is a real number, equal to 1 for undamped Jacobi iteration. SOR makes use of the fact that some updated values for ϕ at sites z' on the r. h. s. of (2.9) may have been computed before when $\phi^{(n+1)}(z)$ is updated. Whenever this is the case, $\phi^{(n)}(z')$ in (2.9) is replaced by $\phi^{(n+1)}(z')$. Whether or not $\phi^{(n+1)}(z')$ has already been computed when the propagator at site z is updated, depends on the ordering in which the sites are swept. A scheme which is particularly well-suited for implementations on vector computers is the checkerboard ordering. Here one employs that Δ couples only nearest neighbors, and that the neighbor of an even site is odd, and vice versa.[5)] Checkerboard SOR reads (modulo exchange even $\leftrightarrow$ odd)

(i) if z is even,

$$\phi^{(n+1)}(z) = (1-\omega)\phi^{(n)}(z) + \frac{\omega}{2d+m^2}\,[f(z) + \sum_{z' \text{ n.n. } z} U(z,z')\phi^{(n)}(z')] \ , \tag{2.10}$$

(ii) if z is odd,

$$\phi^{(n+1)}(z) = (1-\omega)\phi^{(n)}(z) + \frac{\omega}{2d+m^2}\,[f(z) + \sum_{z' \text{ n.n. } z} U(z,z')\phi^{(n+1)}(z')] \ .$$

SOR with $\omega = 1$ is Gauss-Seidel relaxation.

The above iterations can be written as affine fixed point iterations,

$$\phi^{(n+1)} = M\phi^{(n)} + (\mathbb{1} - M)D^{-1}f \ . \tag{2.11}$$

M is called the iteration matrix. Eq. (2.11) implies that the *error*

$$e^{(n)} \equiv \phi - \phi^{(n)} \tag{2.12}$$

gets (de)amplified by M: $e^{(n+1)} = Me^{(n)} = M^{n+1}e^{(0)}$. Hence, (2.11) converges if and only if the spectral radius ρ of M, i. e. the modulus of the largest eigenvalue of M, is bounded by unity. The asymptotic *relaxation time* τ is defined by

$$\tau = -\frac{1}{\ln \rho(M)} \ . \tag{2.13}$$

[4)] We are a little bit sloppy in writing $(2d+m^2)$ instead of $(2d+(ma)^2)$.

[5)] A lattice site z is called even/odd if the sum of its integer coordinates is even/odd.

Numerically τ is determined by monitoring ratios $\|r^{(n+1)}\|/\|r^{(n)}\|$, which approach $\rho(M)$ asymptotically. $r^{(n)}$ denotes the *residual*,

$$r^{(n)} \equiv f - D\phi^{(n)} \ . \tag{2.14}$$

For any positive definite operator D, SOR converges for arbitrary initial $\phi^{(0)}$ if and only if $0 < \omega < 2$. This is Ostrowski's Theorem [80, p. 77]. In our case D is positive definite, so we have an algorithm which solves Eq. (2.1). However, convergence is extremely slow when D is nearly critical. There is a scaling relation for τ which reads in arbitrary unitary gauge fields

$$\tau \propto (\Delta m^2)^{-z/2} \text{ for small } , \Delta m^2 = m^2 - m_{cr}^2 \ . \tag{2.15}$$

irrespective of the lattice size. Here z denotes the critical exponent. It equals 2 for SOR and damped Jacobi relaxation (with fixed ω). Below it will be shown that (2.15) is known analytically for bosons; its validity for staggered fermions will be found numerically later on.

Bosonic propagators.

In case of bosonic propagators we can apply a well-known, well-elaborated mathematical theory on classical relaxation algorithms. This is the theory of so called "consistently ordered, 2-cyclic" matrices [80,82]. A matrix A is called consistently ordered if the eigenvalues of the matrix $J(\alpha) \equiv \alpha\,\mathrm{diag}(A)^{-1}L + \alpha^{-1}\,\mathrm{diag}(A)^{-1}R$ are independent of α for all $\alpha \in \mathbb{C} \setminus \{0\}$; $\mathrm{diag}(A)$ denotes the diagonal of A, and L and R are the lower and upper triangular part, respectively. The notion of consistent ordering was introduced by Varga [80]. Criteria for consistent ordering can be found in the literature [80,82,84]. In order to define the term 2-cyclic, we have to introduce the (directed) graph $G(A)$, which is associated with A. If A is an $n \times n$ matrix, $G(A)$ consists of n vertices $P_1, \ldots, P_n$, and there is a directed link from P_i to P_j iff $a_{ij} \neq 0$. A is called irreducible if $G(A)$ is connected. An irreducible matrix A with nonvanishing diagonal elements is called 2-cyclic if the greatest common divisor of the lengths of all directed closed paths in $G(J(1))$ equals 2.

We note that $(-\Delta + m^2)$ is a consistently ordered, 2-cyclic operator in arbitrary unitary gauge fields U.[6)] Therefore the following conclusions hold [81,80,82].

(i) The undamped Jacobi iteration converges.

(ii) Let ρ_J be the spectral radius of the Jacobi iteration matrix, and ρ_{GS} that of the Gauss-Seidel iteration matrix. Then $\rho_J^2 = \rho_{GS}$.

(iii) For SOR there exists an optimal relaxation parameter ω_{opt} where the spectral radius ρ_ω of M is minimal:

$$\omega_{opt} = \frac{2}{1+\sqrt{1-\rho_J^2}} \quad , \quad \rho_{\omega_{opt}} = \omega_{opt} - 1 = \left(\frac{\rho_J}{1+\sqrt{1-\rho_J^2}}\right)^2 \ . \tag{2.16}$$

[6)] *Think of a block structure in terms of the $N_c \times N_c$ blocks $(-\Delta+m^2)(z_1,z_2)$. That $(-\Delta+m^2)$ is 2-cyclic* can be verified by using [80, Theorem 4.1] in connection with [80, Theorems 2.9, 1.6]; since $(-\Delta+m^2)$ is irreducible, the consistent ordering follows from [82, Theorem 8.3.11] in connection with [82, Theorem 8.3.9].

Generally one has

$$\rho_\omega = \begin{cases} \omega - 1 & \text{for } \omega_{opt} \le \omega \le 2 \,, \\ 1 - \omega + \frac{1}{2}\omega^2\rho_J^2 + \omega\rho_J\sqrt{1 - \omega + \frac{1}{4}\omega^2\rho_J^2} & \text{for } 0 \le \omega \le \omega_{opt} \,. \end{cases} \tag{2.17}$$

For $\rho_J = 1 - \triangle m^2/(2d + m^2)$, we obtain

$$\omega_{opt} = 2 - 2\sqrt{\frac{\triangle m^2}{d + \frac{1}{2}m_{cr}^2}} + O(\triangle m^2) \quad , \qquad \tau_{opt} = \frac{\frac{1}{2}\left[d + \frac{1}{2}m_{cr}^2\right]^{1/2}}{[\triangle m^2]^{1/2}} + O\left([\triangle m^2]^0\right) \,. \tag{2.18}$$

Hence, the critical exponent is 1 for optimal SOR. In practice, however, ω_{opt} cannot be determined accurately enough, because it depends on the critical mass squared of the particular gauge field U. One would rather choose a fixed $\tilde{\omega} > 1$. Eq. (2.17) implies that $\tau = -1/\ln(\tilde{\omega} - 1) =$ constant for all m^2 that are larger than the value $\widetilde{m}^2$ for which $\tilde{\omega}$ is optimal, and one has CSD with $z = 2$ as in Jacobi relaxation for $m^2 < \widetilde{m}^2$. The scaling relation of the Jacobi algorithm reads

$$\tau_J = \frac{\omega^{-1}\left[2d + m_{cr}^2\right]}{\triangle m^2} + (\omega^{-1} - \tfrac{1}{2}) + O\left(\triangle m^2\right) \,. \tag{2.19}$$

In anticipation of Secs. 4 and 6 we note that the validity of (2.19), including the constant $(\omega^{-1} - \frac{1}{2})$, can be determined numerically to a very high precision. In trivial gauge fields Jacobi iteration converges for $0 < \omega \le \omega_{max}$, where $\omega_{max} = (4d + 2m^2)/(4d + m^2)$ or $\omega_{max} = (4d + 2m^2)/(2d(1 + \cos\frac{\pi}{N}) + m^2)$ in case that Λ is an N^d lattice with even or odd N, respectively.

One result of the above considerations must be especially emphasized: *One-grid relaxation algorithms for nearly critical bosonic propagators are almost not sensitive to the disorder of the gauge field, and to the lattice size, as long as equal values of $\triangle m^2$ are compared.* At fixed β the dependence of $[2d + m_{cr}^2]$ on m_{cr}^2 can be neglected for practical purposes, because m_{cr}^2 is small compared to $2d$, and it does not fluctuate very much from one gauge field configuration to the next. The dependence of m_{cr}^2 on the volume of the lattice is also small. As a function of β or the gauge field correlation length the value of m_{cr}^2 varies only smoothly and slightly.

2.2.2 Conjugate Gradient Algorithms

In the rest of this section we will be concerned with the family of the conjugate gradient (CG) algorithms. CG algorithms are state of the art in lattice QCD computations [85, 66,77,86,87,88,89,78,90]. Hence, any new algorithm must be compared with CG. (The UKQCD group [91] uses an over-relaxed minimal residual algorithm, though.)

General CG methods are designed for finding stationary points of continuously differentiable functions [92,93]. The CG algorithm for propagators starts from the fact that solving Eq. (2.1) is equivalent to minimizing the energy functional

$$K[\phi] \;=\; \frac{1}{2} < \phi, D\phi > - < \phi, f > \tag{2.20}$$

where the scalar product is given by

$$< \phi, f > \;=\; \frac{1}{|\Lambda|}\sum_{z\in\Lambda}\frac{1}{N_c}\mathrm{Re\,Tr}\left[\phi(z)^\dagger f(z)\right] \,. \tag{2.21}$$

Gauge invariance property of the energy functional.

We note that the energy functional K is gauge invariant, i. e. $K[\phi] = K[g\,\phi]$. (Remember that we assumed that f transforms like a matter field under gauge transformations.) Later on we will encounter situations where $f(z) \mapsto g(z)\,f(z)\,g(w)^{-1}$ and $\phi^{(n)}(z) \mapsto g(z)\,\phi^{(n)}(z)\,g(w)^{-1}$ under gauge transformations, where w is fixed. *Note that K is also gauge invariant in these situations.*

The CG method is iterative in the sense that starting with an arbitrary initial $\phi^{(0)}$, it yields a sequence $\phi^{(0)} \to \phi^{(1)} \to \phi^{(2)} \cdots$ which converge to the solution of Eq. (2.1). But unlike the relaxation methods discussed above, CG arrives at the solution after at most $|\Lambda|$ steps (provided the arithmetic is exact).

The CG algorithm is a steepest descent method which performs a $(k+1)$ dimensional minimization in the step $\phi^{(k)} \to \phi^{(k+1)}$. $\phi^{(k+1)}$ is determined such that

$$K[\phi^{(k+1)}] = \min_{\alpha_0,\ldots,\alpha_k\in\mathbb{C}} K[\phi^{(k)} + \alpha_0 r^{(0)} + \cdots + \alpha_k r^{(k)}] \,. \tag{2.22}$$

The residuals $r^{(i)}$, $i = 0, \ldots, k$, are orthogonal, and thus independent, as long as $r^{(k)} \neq 0$. Since at most $|\Lambda|$ vectors[7] are independent, there must be an $l \le |\Lambda|$ with $r^{(l)} = 0$ and $\phi^{(l)}$ solves Eq. (2.1).

The cookbook recipe of the CG method is the following [82,84,83].

(1) Choose any $\phi^{(0)}$ and set $p^{(0)} = r^{(0)} = f - D\phi^{(0)}$, $\quad k = 0$.

(2) If $r^{(k)} = 0 \;\Leftrightarrow\; p^{(k)} = 0$: STOP, $\phi^{(k)}$ is the solution of $D\phi = f$.

(3) Else compute

$$\begin{aligned} a_k = \frac{< r^{(k)}, r^{(k)} >}{< p^{(k)}, Dp^{(k)} >} &\equiv \frac{< r^{(k)}, p^{(k)} >}{< p^{(k)}, Dp^{(k)} >} \,, \qquad \phi^{(k+1)} = \phi^{(k)} + a_k\, p^{(k)} \,, \\ r^{(k+1)} &= r^{(k)} - a_k\, Dp^{(k)} \,, \qquad b_k = \frac{< r^{(k+1)}, r^{(k+1)} >}{< r^{(k)}, r^{(k)} >} \,, \\ p^{(k+1)} &= r^{(k+1)} + b_k\, p^{(k)} \end{aligned} \tag{2.23}$$

(4) Increase k by 1 and go to (2).

[7] Remember that, strictly speaking, the vectors here are actually matrices.

See Refs. [82,84,83,94] for properties of the $\phi^{(k)}$, $p^{(k)}$, $r^{(k)}$, for a proof of convergence of the CG method and for a discussion about its numerical properties.

In practice one will inevitably have to deal with round-off errors, and one will stop the algorithm when $\|r^{(k)}\|$ is "small enough" (e. g. when $r^{(0)}$ is reduced by 10^{-5}). The same stopping criterion will of course be used in relaxation algorithms which never reach the solution otherwise. The reason for the success of CG is that the stopping criterion is reached in much less than $|\Lambda|$ iterations.

Sometimes one hears the statement that CG has a critical exponent $z \approx 1$ or even the statement $z = 1$ for CG. This must be taken cautiously. Strictly speaking, CG has no well defined z because it is not an infinite process. If one nevertheless wants to define an asymptotic relaxation time, τ would be zero, because after a finite number of iterations the error is zero. Regarded from that point of view, CG is a direct solver. What causes people to say "$z \approx 1$ for CG" is the following. One can prove the bound [84]

$$\|D^{1/2}\, e^{(k)}\| \;\leq\; \frac{2\, c^k}{1 + c^{2k}} \|D^{1/2}\, e^{(0)}\| \quad \text{with} \quad c = \frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1} \;. \tag{2.24}$$

Here κ denotes the condition number of D, i. e. the ratio of the largest to the smallest eigenvalue. If CG were an infinite iteration and if the bound were stringent, one would derive the scaling relation (2.15) from (2.24) with $z = 1$. But even if the iteration were infinite, the situation is more complicated, see the discussion in Ref. [84].

Finally we note that preconditioning techniques [83,84] exist which are very successful in accelerating the computation of propagators of Wilson fermions [65,66,67,68]. However, no such techniques have been proven useful for staggered fermions [69]. For this reason we will not discuss and use preconditioning in the present thesis.

2.3 Elimination of Critical Slowing Down by Multigrid Algorithms

The general deterministic multigrid method is explained in this subsection.

Multigrid (MG) algorithms were invented to circumvent the problem of CSD in the solution of discretized partial differential equations (in the absence of gauge fields). The slowness of traditional algorithms is overcome by updates on various length scales. Introductions to this subject can be found in the classical papers of Brandt, and Stüben and Trottenberg [95], in the textbook of Hackbusch [96], at a less advanced level in the books of McCormick [97] and of Hackbusch [84], and at a very elementary level in the book of Briggs [98].

The basic observation for MG methods is the following. Classical relaxation algorithms are effective in smoothing the error, but as soon as the error is smooth (on length scale a) it is reduced only very slowly because of CSD. However, a smooth function can be represented very well on a coarser lattice. Suppose for instance that the values of a lattice function are given only on every second site. Then, if one knows that the function is

smooth, one can reconstruct it to a good accuracy by interpolation. We will now explain these ideas in more detail. (The reader may prefer not to think of the presence of gauge fields in Eq. (2.1) until Sec. 2.6.)

In the MG approach for solving Eq. (2.1) we divide the original hypercubic lattice Λ of lattice spacing a into hypercubes ("blocks") x consisting of L_b^d sites $z \in \Lambda$, with $L_b \in \mathbb{N}$,[8] typically $L_b = 2, 3$. We identify each such hypercube x with the site $\hat{x}$ at its center, and we write $z \in x$ if z is a site in block x. (If L_b is even there is no such distinguished center. Then define arbitrarily an $\hat{x}$ in one block; this defines the other block centers by the requirement that the block lattice is regular.) The sites $\hat{x}$ form the first block lattice Λ^1 with lattice spacing $L_b a$, and so on. This yields a sequence of lattices $\Lambda = \Lambda^0$, Λ^1, Λ^2, ... of increasing lattice spacing a_i, viz. $a_{i+1} = L_b a_i$ with $a_0 = a$. (One may also use different blocking factors L_b on different layers of the MG).

After some relaxation sweeps on Λ^0 one gets an approximation $\phi^{(n)}$ to ϕ which differs from the exact solution by an error $e^{(n)} \equiv e_0$, Eq. (2.12). The error satisfies the *residual equation*

$$D_0 e_0 = r_0 \;, \tag{2.25}$$

where $r_0 \equiv r^{(n)}$ is the residual, (2.14), and we wrote D_0 for D. If e_0 is smooth, it is determined to a very good accuracy by a function e_1 on the next coarser lattice Λ^1, and can be represented in the form

$$e_0 = \mathcal{A} e_1 \tag{2.26}$$

with an interpolation map $\mathcal{A}$ which should be so chosen that it maps functions on Λ^1 into smooth functions on Λ^0. Conversely, e_1 can be obtained from e_0 with the help of an averaging map C which satisfies

$$C\mathcal{A} = \mathbb{1} \;. \tag{2.27}$$

It follows that $e_1 = C e_0$. Inserting Eq. (2.26) into Eq. (2.25) and acting on the result with C, we see that e_1 will satisfy the equation

$$D_1 e_1 = r_1 \tag{2.28a}$$

with

$$D_1 = C D_0 \mathcal{A} \;, \quad r_1 = C r_0 \;. \tag{2.28b}$$

The problem has been reduced to an equation on the coarser lattice Λ^1 which has fewer points. If there is still too much CSD at this level, one may repeat the procedure, going to coarser and coarser lattices. The procedure stops, because an equation on a "lattice" Λ^N with only a single point is easy to solve.

After solution of Eq. (2.28a) one replaces $\phi^{(n)} \mapsto \tilde{\phi}^{(n)} \equiv \phi^{(n)} + \mathcal{A} e_1$. Note that the residual of the corrected approximation $\phi^{(n)} + \mathcal{A} e_1$ vanishes when it is transfered back to Λ^1. If $\mathcal{A} e_1$ were equal to e_0, then $\tilde{\phi}^{(n)}$ would be the solution of Eq. (2.1). In practice, however, one has to repeat the procedure: do relaxation with $\tilde{\phi}^{(n)}$, solve the residual equation for the new error, etc.

[8] The requirement $L_b \in \mathbb{N}$ is not compulsory; see e. g. Sec. 7.7.

The reason for the efficiency of the MG method is that with a suitable choice of C, $\mathcal{A}$, D_1 etc. only a few iterations are needed to reduce the error to a small value, independent of the mass term m^2. In other words, CSD is completely eliminated by MG. This statement has been known to be true in the absence of gauge fields, and it is the subject of this thesis to address the question what the effects of the inclusion of gauge fields are. A crucial problem is how to define and exhibit smooth functions in a disordered context, i. e. when translation symmetry is strongly violated.

Another advantage of the MG method is that the computational labor for one MG iteration is comparable to that of conventional relaxation, irrespective of the total number of layers. For details of this work estimate see [95,96,98,79,84].

There are further terms which are relevant in MG algorithms for which we refer to the literature. These terms include the "cycle control parameter γ", the notion of "V-cycles" ($\gamma = 1$) and "W-cycles" ($\gamma = 2$), etc. We will not need these terms. We are interested in four space-time dimensions where the volumes of subsequent lattices differ by a factor of L_b^4. Throughout this thesis we will always fix $L_b = 3$ without further notice. Thus, Λ^1 has 81 times fewer points than Λ^0. Eq. (2.28a) on the coarser lattice was solved exactly by CG.[9] This suffices to test the power of the MG method.

Finally we note that the MG iteration is also an affine fixed point iteration like (2.11). For instance, the iteration matrix $\mathcal{M}$ of a two-grid algorithm is given by $\mathcal{M} = M^{\nu_2}\,[\mathbb{1} - \mathcal{A}(D_1)^{-1}CD_0]\,M^{\nu_1}$, where ν_1 and ν_2 denote the number of relaxation sweeps on Λ^0 before and after the MG correction step, respectively. In actual computations only the sum $\nu_1 + \nu_2$ matters, and we chose $\nu_1 + \nu_2 = 1$. By using the Fourier representations of integral kernels it is an easy exercise to show that $\mathcal{M}$ annihilates the lowest mode of D_0 with different choices of C and $\mathcal{A}$; see Appendix B.2.

2.4 Variational Multigrid

The variational choice of averaging and interpolation operators in MG algorithms is introduced.

In order to specify an MG algorithm, we have to make a specific choice for the restriction operator C and for the interpolation operator $\mathcal{A}$. These operators will be defined by their integral kernels $C(x,z)$ and $\mathcal{A}(z,x)$. ($z \in \Lambda^j$, $x \in \Lambda^{j+1}$, e. g. $z \in \Lambda^0$, $x \in \Lambda^1$) The notations in this thesis will always be such that z denotes sites in the finer lattice and x denotes sites in coarser layers.

For reasons of practicality one must require that

$$\mathcal{A}(z,x) = 0 \quad \text{unless } z \text{ is near } \hat{x}. \tag{2.29}$$

Two special choices for C and $\mathcal{A}$ will be discussed in this and the next section.

From the variational point of view one requires that the energy functional K (2.20) should be lowered as far as possible in every MG correction step $\phi^{(n)} \mapsto \phi^{(n)} + \mathcal{A}e_1$. It

[9] This two-grid algorithm is equivalent to an MG algorithm with $\gamma = \infty$ (or a large number).

follows that the averaging map C and the interpolation map $\mathcal{A}$ are adjoints of one another [95,96,79]:

$$C = \mathcal{A}^* \ , \tag{2.30}$$

and that the coarse grid operator D_1 is defined as

$$D_1 = CD_0C^* \equiv CDC^* \ . \tag{2.31}$$

This is called the Galerkin definition of D_1. The integral kernel of the adjoint averaging operator is $C^*(z,x) = C(x,z)^\dagger$.

Eq. (2.27) imposes the normalization condition $CC^* = \mathbb{1}$ on the averaging operator C in variational MG.

2.5 "Optimal" Multigrid

An "optimal" MG scheme is introduced. In this scheme there is complete decoupling between the layers of the MG. This decoupling will also hold in the presence of gauge fields.

Given the averaging kernel C, there exists an ideal choice of the interpolation kernel $\mathcal{A}$. It is determined as follows. For every function ("block spin") Φ on Λ^1, $\phi = \mathcal{A}\Phi$ minimizes the action $\mathcal{H} = <\phi, D\phi>$ subject to the constraint $C\phi = \Phi$. With this choice of $\mathcal{A}$, D_1 is guaranteed to be self-adjoint. A good "choice of block spin", i. e. of C, is characterized by the fact that the ideal kernel $\mathcal{A}(z,x)$ associated with it has good locality properties. This means that $\mathcal{A}(z,x)$ is big for $z \in x$, and decays exponentially in $|z - \hat{x}|$ with decay length one block lattice spacing.

The above characterization of $\mathcal{A}$ is equivalent to saying that with the ideal choice of $\mathcal{A}$, there is complete decoupling between layers. This means that the action $\mathcal{H}$ can be written as a sum of actions for the different layers of the MG, viz. $\mathcal{H} = \sum_j \mathcal{H}_j$, with $\mathcal{H}_j = <\zeta_j, D_j\zeta_j>$. ζ_j is a field which is defined on Λ^j, and D_j is the effective inverse propagator on Λ^j. In other words: Complete decoupling between layers means that the propagator is a sum of contributions from layers, and these contributions from different layers satisfy effective difference equations which do not couple. As a result, the convergence speed is determined by the convergence speed on the individual layers. We note that complete decoupling $\mathcal{H} = \sum_j \mathcal{H}_j$ does not hold in variational MG.

For the purpose of numerical computations, it is convenient to determine the optimal $\mathcal{A}$ as solution of the equation

$$([D + \kappa C^*C]\mathcal{A})(z,x) = \kappa C^*(z,x) \tag{2.32}$$

for large κ. The solution of Eq. (2.32) yields the same $\mathcal{A}$-kernel which was proposed in Mack's Cargèse lectures [34] for use in "optimal" MG Monte Carlo simulations. In Ref. [34], $\mathcal{A}$ was determined for 1-component ϕ^4 theory by a relaxation method. This is feasible, because in the absence of gauge fields and for the step function kernel C of Eq. (2.35) below, $(\Delta\mathcal{A})(z,x)$ are constants on blocks x as functions of z. Another possible

computation of $\mathcal{A}$ would be by standard optimization algorithms, making use of the above characterization of $\mathcal{A}$ as solution of an extremization problem.

In the presence of gauge fields, $\mathcal{A}$ is computed as the solution of the gauge covariant generalization of Eq. (2.32). We note that complete decoupling between the layers of the MG will continue to hold in the presence of gauge fields.

The origin of the optimal $\mathcal{A}$ lies in the works [76] of Gawędzki and Kupiainen in constructive quantum field theory, see also their Les Houches lectures [99]. Gawędzki and Kupiainen gave rigorous proofs of the existence of the continuum limit of some lattice field theories without gauge fields, using block spin renormalization group methods. The use of the Gawędzki-Kupiainen kernels as a starting point for numerical work was proposed by Mack [34]. He pointed out that it will be essential for beating CSD in interacting models that the layers of an MG decouple as much as possible. A necessary condition for this is smoothness of $\mathcal{A}$. The correctness of Mack's contention was confirmed in an impressive way by Hasenbusch and Meyer [43,44].

A derivation of the optimal $\mathcal{A}$, including Eq. (2.32), is summarized in Appendix B. In Secs. 3.5 and 5 further properties of the ideal $\mathcal{A}$ will be discussed. By its use it was possible to demonstrate that the MG method can cope with the frustration which is inherent in non-Abelian gauge fields (Sec. 5). Unfortunately, the optimal $\mathcal{A}$ does not fulfill the practicality condition (2.29), so that the idealized MG algorithm cannot be used for production runs. But the results of Sec. 5 settle questions of principal importance.

2.6 Inclusion of Gauge Fields

The requirements for making MG algorithms covariant in the presence of gauge fields are discussed. When an algorithm is gauge covariant, the necessity for fixing a gauge in computation of propagators is avoided.

Now we turn to the point of including gauge fields properly into the MG algorithm. This should be done in such a way that gauge covariance is ensured. Then no gauge fixing is required in computations of propagators. Note that Jacobi relaxation (2.9), SOR (2.10) and CG (2.23) are gauge covariant in the sense that all $\phi^{(n)}$ are gauge transformed by g if g is applied before relaxation is started. (We always begin with an initial $\phi^{(0)} = 0$.) The gauge covariance property is to be retained in multigrid algorithms.

The block spin $\Phi = C\phi$ should transform under gauge transformations (2.6), (2.7) like a matter field sitting at block centers $\widehat{x}$. Similarly, the "background field" $\psi = \mathcal{A}\Phi$ should transform like ϕ. This is achieved if under gauge transformations

$$\begin{aligned} \mathcal{A}(z,x) &\mapsto g(z)\,\mathcal{A}(z,x)\,g(\widehat{x})^{-1} , \\ C(x,z) &\mapsto g(\widehat{x})\,C(x,z)\,g(z)^{-1} . \end{aligned} \tag{2.33}$$

Then the equations of Secs. 2.3 – 2.5 become gauge covariant. Eq. (2.33) is consistent with $C\mathcal{A} = \mathbb{1}$. Note that the transformation law (2.33) is also consistent with the variational choice $\mathcal{A} = C^*$.

The most general expression of a kernel $\mathcal{A}$ with covariance property (2.33) is a weighted sum of parallel transporters $U(\mathcal{C})$ along paths $\mathcal{C}$ from $\widehat{x}$ to z, i. e.

$$\mathcal{A}(z,x) = \sum_{\mathcal{C}\,:\,\widehat{x}\to z} \varrho(\mathcal{C})\,U(\mathcal{C}) , \tag{2.34}$$

where $\varrho(\mathcal{C})$ are numbers. And analogously for C. One sees from (2.34) that the integral kernels of C and $\mathcal{A}$ are $N_c \times N_c$ matrices in gauge theories. However, they will not be elements of the gauge group, in general. From the fact that C and $\mathcal{A}$ are matrices, it follows especially that there is a color-mixing in averaging and in interpolation.

An MG approach which is based on the representation (2.34) is the "parallel transported multigrid" (PTMG) method, introduced by Ben-Av, Brandt, and Solomon [72] for computations in the Schwinger model (two-dimensional QED). The PTMG approach is also described in Ref. [100], in Ben-Av's thesis [101], and in [102].

We will prefer not to specify the weights ϱ in (2.34) explicitly, but to determine $\mathcal{A}$ and C as solutions of covariant equations. This alternative approach is the "ground-state projection MG", which will be discussed next.

2.7 Ground-State Projection Multigrid

The particularly attractive "ground-state projection multigrid" approach is defined. This method is applicable in arbitrary dimension and for arbitrary gauge group. It has the advantage that the procedure is gauge covariant, i. e. no gauge fixing is required. The ground-state projection MG method is based on "block-local Hamiltonians". However, the definition of a block-local Hamiltonian requires the specification of boundary conditions on block boundaries. The necessity of specifying boundary conditions is responsible for the fact that "ground-state projection MG" is not an a priori defined scheme. We will choose "Neumann boundary conditions", and we will discuss this choice in detail.

The central idea of the ground-state projection MG philosophy is that a local action should define the block spin (or C, respectively). The averaging operator C from a grid to the next coarser grid is a projector on the ground-state of a block-local Hamiltonian. The adjoint of C satisfies a gauge covariant eigenvalue equation, Eq. (2.37) below. The solution of the eigenvalue equation is made unique by imposing a normalization and a covariance condition (Eqs. (2.38) and (2.43)). The idea behind a ground-state projection definition is that the appropriate notion of smoothness depends on the dynamics, i. e. on D, in general. Results, which will be reported in later sections, confirm the insight that *smooth means little contributions from eigenfunctions to high eigenvalues of* D. This point is important in systems in gauge fields and for other disordered systems.

2.7.1 Averaging Map C for Free Bosons

Let us first discuss the averaging map C in the absence of gauge fields, and for bosons; the case of staggered fermions will be treated in Sec. 3.2. It is reasonable to demand that C is

local in the sense that the averaged field ("block spin") $\Phi(x) = (C\phi)(x)$ at a block lattice site x receives contributions only from $\phi(z)$ for $z \in x$, i. e. $C(x,z) = 0$ unless $z \in x$. The natural choice inside (non-overlapping) blocks is a constant. Hence,[10]

$$C(x,z) = \begin{cases} (L_b a)^{-d} \cdot 1 & \text{if } z \in x \ , \\ 0 & \text{otherwise} \ . \end{cases} \tag{2.35}$$

With this choice, Φ is the block average of ϕ,

$$\Phi(x) = \int_x C(x,z)\phi(z) = L_b^{-d} \sum_{z \in x} \phi(z) \equiv \underset{z \in x}{\text{av}}\, \phi(z) \ . \tag{2.36}$$

This block spin transformation was also used by Gawędzki and Kupiainen [76]. We will now see how it can be naturally generalized so that gauge fields are incorporated.

2.7.2 C as the Solution of an Eigenvalue Equation

We recall that we denoted by $C^*(z,x)$ the integral kernel of the adjoint of C. One has the equality $C^*(z,x) = C(x,z)^\dagger$, where † denotes Hermitean conjugation of a matrix, as usual. The adjoint of the averaging kernel (2.35) is a solution of the eigenvalue equation

$$(-\Delta_{N,x} C^*)(z,x) = \lambda_0(x)\, C^*(z,x) \ , \tag{2.37}$$

together with the subsidiary condition $C^*(z,x) = 0$ if $z \notin x$. $-\Delta_{N,x}$ is the negative lattice Laplacian with Neumann boundary conditions on the boundary of block x (see below), and $\lambda_0(x)$ is its lowest eigenvalue. $\Delta_{N,x}$ acts on argument z. In the absence of gauge fields $\lambda_0(x)$ equals zero for all blocks x. In this case solutions of (2.37) are constants on blocks. These constants can be determined by the normalization condition $CC^* = \mathbb{1}$, which reads for the kernel

$$(CC^*)(x_1,x_2) = \int_x C(x_1,z)C^*(z,x_2) = \delta(x_1 - x_2) \equiv (L_b a)^{-d} \delta_{x_1,x_2} \ . \tag{2.38}$$

Following Mack [75], this procedure can be re-interpreted as follows. Define a notion of block-local frequency which depends only on the behavior of the function which is to be decomposed into frequency components on the chosen block x. Thus, define frequency squared as eigenvalue of the negative Laplacian with Neumann boundary conditions. Define the block spin in two steps. First define the lowest frequency part $\psi(z)$ of $\phi(z)$ by projection

$$\psi(z) = \lim_{t \to \infty} (\exp\left[-t(-\Delta_{N,x} - \lambda_0(x))\right] \phi)(z) \ . \tag{2.39}$$

$\psi(z)$ will be a smooth function of z inside the block. In the second step the block spin is defined equal to the value of this smooth function at the block center $z = \hat{x}$.

[10] A remark about scale factors is in order here. The operator C is dimensionless, but its integral kernel has dimension of a length^{-d}. Almost always factors of $(L_b a)^{-d}$ are neglected in other works. But we stress that one has to keep track of them, otherwise there will appear inconsistencies. One can set the lattice spacing = 1 on each layer separately. But then one has to watch out for factors L_b when one goes from one layer to the next.

We note that Neumann boundary conditions on the Laplacian yield a local approximation of Δ which preserves the invariance of the action $\mathcal{H}_0 = <\phi, -\Delta\phi>$ under shifts by constant fields. One sees that the block spin $\Phi(x)$ retains the component of ϕ associated with the lowest eigenvalue of $-\Delta_{N,x}$.

For block spins in (non-Abelian) gauge theories one can proceed in exactly the same way, defining C^* as solution of Eq. (2.37). Now C depends on U, though we will not indicate that explicitly. $\Delta_{N,x}$ will be a discrete substitute for the Laplace operator with Neumann boundary conditions on the continuum.

The idea is that the definition of block spins involves dynamical information — think of $\int_{x \in x} \frac{1}{2} N_c^{-1} \operatorname{Re} \operatorname{Tr} [\phi(z)^\dagger (-\Delta_{N,x}\phi)(z)]$ as the part of the kinetic energy which is associated with the inside of block x. Thus, what is called "low frequency" is actually determined by kinetic energy. One could also admit dielectric gauge fields U [103], especially on coarser layers.

2.7.3 Neumann Boundary Conditions on the Lattice

Neumann boundary conditions (b. c.) specify the values of the normal gradients of a function ϕ on a boundary. In numerical analysis the term "Neumann boundary conditions" is used synonymously with "homogeneous Neumann boundary conditions" [93]. We follow this notation. That means $\nabla\phi = 0$ perpendicular to the boundary $\partial\Omega$ of a domain Ω where ϕ is defined as the solution of a partial differential equation $\mathcal{D}\phi = 0$. Derivatives are approximated on the lattice by difference quotients. The lattice transcription of "$\nabla\phi = 0$ on $\partial\Omega$" is done in such a way that derivative terms $a^{-1}[\phi(z) - \phi(z')]$ are omitted in the discretized version of the differential operator $\mathcal{D}$ when one site $z \in \Omega$ and the other $z' \notin \Omega$.

The gauge covariant lattice Laplacian $\Delta_{N,x}$ with Neumann b. c. on the boundary of a domain Ω which equals a block x, depends on the lattice gauge field U and is defined by

$$(\Delta_{N,x}\phi)(z) = \sum_{\substack{z' \text{ n.n. } z \\ z' \in x}} [U(z,z')\phi(z') - \phi(z)] \quad \text{for } z \in x. \tag{2.40}$$

This definition of $\Delta_{N,x}$ agrees with the one used by Balaban [104] in rigorous works on constructive gauge theories. Summation on the r. h. s. of Eq. (2.40) is over next neighbors z' of z which lie in the same block x.

The imposition of (homogeneous) Dirichlet b. c. instead of Neumann b. c. would result in the lattice operator

$$(\Delta_{D,x}\phi)(z) = -2d\,\phi(z) + \sum_{\substack{z' \text{ n.n. } z \\ z' \in x}} U(z,z')\phi(z') \quad \text{for } z \in x. \tag{2.41}$$

Note that for a blocking factor $L_b = 2$ the operators $\Delta_{N,x}$ and $\Delta_{D,x}$ differ by a multiple of the unit operator, namely $d\,\mathbb{1}$. In this case a solution of Eq. (2.37) is also a solution of the corresponding equation where $\Delta_{N,x}$ is replaced by $\Delta_{D,x}$, but this is not true when $L_b \neq 2$.

The reason why "ground-state projection MG" is not an a priori defined scheme requires further discussion: Ground-state projection MG is supposed to preserve block-locally the

lowest mode of D. In case of bosons, one has to impose b. c. for the covariant Laplace operator on the boundaries of blocks x. This gives us what we called "a block-local Hamiltonian". However, a priori there exists no distinguished choice of boundary conditions so that the meaning of "block-local Hamiltonian" is not unique. Therefore "ground-state projection MG" is a priori no completely defined scheme. The choice of proper boundary conditions is even more involved in case of staggered fermions. For this discussion we refer to Sec. 3.2.

In case of bosons we decided to impose Neumann b. c. for the following reasons. We noted already above that Neumann b. c. preserve the invariance of the free action $\mathcal{H}_0 = < \phi, -\Delta\phi >$ under shifts by constant fields when no gauge fields are present. In the absence of gauge fields, solutions of the eigenvalue equation (2.37) are constants on blocks if the subscript N stands for Neumann b. c. as defined in Eq. (2.40). These solutions are known to be "good" when one performs renormalization group calculations [76,99]. In the practical method of variational MG for propagators, "piecewise-constant" kernels are successful in eliminating CSD (Sec. 4). It is reasonable to start from such requirements on the averaging kernel C. If one used Dirichlet instead of Neumann b. c. in (2.37), variational MG would not be successful in the limiting case of vanishing gauge coupling.

In order to be precise, in what follows one should always read "ground-state projection MG with Neumann b. c. on the boundaries of blocks" instead of just "ground-state projection MG".

2.7.4 Averaging Map C for Bosons in Non-Abelian Gauge Fields

Eq. (2.37) is covariant when $\Delta_{N,x}$ is the operator (2.40), because if $C^*(z,x)$ is a solution in a given gauge field U, then $g(z)\,C^*(z,x)$ will satisfy (2.37) in the gauge transformed configuration (2.7). This is so because the eigenvalues of $\Delta_{N,x}$ are gauge invariant.

We noted earlier that $C(x,z)$ is an $N_c \times N_c$ matrix in gauge theories. Hence one might say that the eigenvalue equation (2.37) determines an "eigenmatrix" rather than a (1-column) eigenvector. In case of gauge group $SU(2)$, Eq. (2.37) has two degenerate solutions for any gauge field, when it is regarded as an equation for a 1-column vector in place of a matrix. To see this suppose that $(c_{11}, c_{12}, c_{21}, c_{22}, \ldots, c_{V1}, c_{V2})^T$, $V = L_b^d$, is an eigenvector of $-\Delta_{N,x}$ corresponding to an eigenvalue λ. By using the fact that the elements of (a multiple of) an $SU(2)$ matrix $(U_{ij})_{i,j=1,2}$ fulfill $U_{21} = -\overline{U_{12}}$, $U_{22} = \overline{U_{11}}$, one proves that $(-\overline{c_{12}}, \overline{c_{11}}, -\overline{c_{22}}, \overline{c_{21}}, \ldots, -\overline{c_{V2}}, \overline{c_{V1}})^T$ is also an eigenvector of $-\Delta_{N,x}$ with the same eigenvalue λ. The two independent 2-component solutions may be combined into a 2-column matrix $C(x,z)$. The freedom of taking linear combinations reflects itself in the freedom of performing dielectric gauge transformations on the block lattice, i. e. taking

$$C^*(z,x) \to C^*(z,x)\,\Upsilon(x) \ , \tag{2.42}$$

where $\Upsilon(x)$ is an arbitrary 2×2 matrix. The freedom of performing dielectric gauge transformations (i. e. transformations with matrices which are not elements of the gauge group) is eliminated by the normalization condition (2.38). But this leaves the freedom of

performing ordinary (unitary) gauge transformations on the block lattice. This remaining arbitrariness is eliminated by imposing the covariance condition

$$C(x,\hat{x}) = r(x)\,\mathbb{1} \ , \quad \text{with } r(x) > 0 \text{ real.} \tag{2.43}$$

For gauge groups different from $U(1)$ and $SU(2)$, the r. h. s. of (2.43) is to be replaced by a positive Hermitean matrix.

The condition (2.43) ensures the transformation property (2.33). In particular, if U is a pure gauge, then $(L_b a)^d C(x,z)$ equals the parallel transporter along an arbitrary path from z to $\hat{x}$. $C(x,z)$ is therefore known beforehand in trivial gauge fields.

Concluding remarks.

The first use of a ground-state projection definition of the restriction operator C was reported by Hulsebos, Smit and Vink [30] in MG Monte Carlo simulations of a two-dimensional scalar $U(1)$ model, and by Brower, Rebbi and Vicari [74] for propagators in two-dimensional $U(1)$ gauge fields. The present author pointed out that the ground-state projection method is applicable in arbitrary dimension d and for arbitrary gauge group [105]. In Ref. [105] an efficient algorithm for solving Eq. (2.37) was described, and it was also proposed to use ground-state projecting kernels in Monte Carlo renormalization group studies. The lowest eigenvalue $\lambda_0(x)$ of $-\Delta_{N,x}$ (or $-\Delta_{D,x}$) is a measure of disorder and its renormalization group flow is therefore instructive.

The numerical algorithm for solving Eq. (2.37) will be explained in Sec. 3.4. Computations of propagators will be reported in Secs. 4–7.

2.8 Updating on a Multigrid Layer Consisting of a Single Site

Updating on an MG layer consisting of a single site will be investigated. We will argue that such an updating can be viewed as a global rescaling of an approximate solution $\phi^{(n)}$ of Eq. (2.1) by an $N_c \times N_c$ matrix Ω: $\phi^{(n)} \mapsto \phi^{(n)}\Omega$. The matrix Ω is fixed by the requirement that the energy functional of the rescaled propagator gets minimized. The gauge covariance properties of this proposal will be discussed in detail.

Recall that the CSD of conventional relaxation algorithms for solving Eq. (2.1) depends only on m^2 and not on the lattice size, Eq. (2.15). Therefore one continues to have CSD on a lattice of only 2^d sites, and it seems necessary to go to a 1^d lattice (a single site) in order to eliminate the appearance of CSD. Of course, a $2^d N_c \times 2^d N_c$ matrix can be inverted with reasonable effort, but this is no objection against the investigation of "updating on a last single site".

The reader should have the bosonic case or Wilson fermions in mind when reading this section. The treatment of the case of staggered fermions requires some notations which have not been introduced yet. The discussion for staggered fermions will be deferred to Sec. 7.5.

General framework.

When we update on a 1^d sublattice $\Lambda^N = \{x_N\}$, interpolation leads to the replacement

$$\phi^{(n)}(z) \mapsto \phi^{(n)}(z) + \mathcal{A}(z, x_N)\,\tilde{\Omega}(x_N) \ . \tag{2.44}$$

Here we adopt the unigrid point of view, i. e. we consider the effect of the MG update on the field on Λ^0. $\mathcal{A}$ denotes a kernel which interpolates directly from a 1^d sublattice Λ^N to Λ^0. $\tilde{\Omega}(x_N)$ is the error of $\phi^{(n)}$ represented at the last site x_N. Note that $\tilde{\Omega}(x_N)$ is an $N_c \times N_c$ matrix in gauge theories.

Gauge covariance.

Let us discuss the covariance properties of (2.44), and let us assume that the r. h. s. f of Eq. (2.1) transforms like a matter field (2.6) under gauge transformations on Λ^0. For instance, this will be the case in applications with the hybrid Monte Carlo algorithm [52] in simulations of QCD with dynamical fermions where f is a pseudofermion field. [In the hybrid Monte Carlo algorithm pseudofermion fields f are generated by a Gaussian heatbath method. One chooses a field η from a Gaussian probability distribution, and one sets $f = (-\not{D} + m)\eta$. Since the Gaussian distribution of η is invariant under gauge transformations $\eta \mapsto g\eta$, it follows that pseudofermions fields transform according to (2.6) under gauge transformations.]

When f transforms like a matter field, the same will be true for the smeared propagator ϕ and for all the approximate solutions $\phi^{(n)}$. This is so because iterative algorithms are gauge covariant in the sense that all $\phi^{(n)}$ are gauge transformed by g if g is applied before relaxation is started. (We always begin with an initial $\phi^{(0)} = 0$.)

We required the interpolation kernel $\mathcal{A}$ to transform under gauge transformations on Λ^0 according to (2.33). However, in case that one considers a layer consisting of a single site, the choice of a "block center $\hat{x}_N$" is completely arbitrary because there is no distinguished site. Furthermore, we would like to have the possibility of performing updates of the form (2.44) *also in cases where the linear extension of the lattice Λ^0 is not a power of L_b.* For example, when $L_b = 3$, an 18^4 lattice can be blocked to a 6^4 lattice and that to a 2^4 lattice; but one cannot block any further if one insists in retaining $L_b = 3$. Of course, it is not compulsory in an MG algorithm to have the same scale factor L_b in every blocking step, but in case of staggered fermions one is restricted to odd L_b.[11] That follows from considerations concerning the symmetry of free staggered fermions and will be explained in Sec. 3.2.

Actually, one does not have to identify a "last site x_N" with a site $\hat{x}_N \in \Lambda^0$. This observation is similar to the point which Palma encountered in his computation of constraint effective potentials in gauge theories [42]. Palma pointed out that the constraint effective potential is not affected by gauge transformations on a last 1^d layer. An argument x_N is redundant, and therefore we can follow Palma and write $\mathcal{A}(z)$ instead of $\mathcal{A}(z, x_N)$ in (2.44). When we consider updates of the form (2.44) it is therefore legitimate to relax (2.33) by

[11] In case of staggered fermions the lattice in the above example is better replaced by a lattice where one ends up with a 4^4 lattice. The reason for this will become clear later on.

requiring that just

$$\mathcal{A}(z) \mapsto g(z)\,\mathcal{A}(z) \tag{2.45}$$

under gauge transformations g on Λ^0. Then $\tilde{\Omega}$ has no argument x_N and it must be gauge invariant, instead of transforming like a matter field sitting at $\hat{x}_N$.

Choice of the interpolation kernel $\mathcal{A}$.

Given an approximate solution $\phi^{(n)}$ of Eq. (2.1), the equation for the error $e^{(n)} = \phi - \phi^{(n)}$ is $De^{(n)} = r^{(n)}$, where the residual equals $r^{(n)} = f - D\phi^{(n)}$. We wish to have

$$(D\mathcal{A})(z)\,\tilde{\Omega} = r^{(n)}(z) \ , \tag{2.46}$$

but we cannot satisfy this equation, otherwise $\mathcal{A}\tilde{\Omega}$ would be the error and we were done. *Therefore we ask ourselves how Eq. (2.46) can be solved "as well as possible".*

Let us introduce the notation

$$\left(\phi^{(n)}, f\right) \equiv \frac{1}{|\Lambda^0|} \sum_{z \in \Lambda^0} \phi^{(n)}(z)^\dagger f(z) \ . \tag{2.47}$$

Note that (2.47) defines an $N_c \times N_c$ matrix in gauge theories. (Mind also the dagger †.)

By multiplying Eq. (2.46) from the left with the adjoint of some lattice field ψ which will be specified below, and summing the result over z, we obtain

$$\begin{aligned} \tilde{\Omega} &= \left(\psi, D\mathcal{A}\right)^{-1} \cdot \left(\psi, r^{(n)}\right) \\ &= \left(\psi, D\mathcal{A}\right)^{-1} \cdot \left(\psi, f\right) - \left(\psi, D\mathcal{A}\right)^{-1} \cdot \left(\psi, D\phi^{(n)}\right) \ . \end{aligned} \tag{2.48}$$

In the rest of this section and in Sec. 2.9 the dot "$\cdot$" shall emphasize that $N_c \times N_c$ matrices are multiplied. Inserting (2.48) into Eq. (2.46) yields

$$\int_{z'} D(z, z') \cdot \left[\mathcal{A}(z') \cdot \left(\psi, D\mathcal{A}\right)^{-1} \cdot \left(\psi, f\right) - \mathcal{A}(z') \cdot \left(\psi, D\mathcal{A}\right)^{-1} \cdot \left(\psi, D\phi^{(n)}\right)\right] = r^{(n)}(z) \ . \tag{2.49}$$

Compare Eq. (2.49) with the residual equation which reads $\int_{z'} D(z, z') \cdot \left[\phi(z') - \phi^{(n)}(z')\right] = r^{(n)}(z)$. The second term in the bracket of (2.49) equals $\phi^{(n)}$ if we set $\mathcal{A} = \phi^{(n)}$, irrespective of ψ. This is a legitimate choice which has the correct gauge covariance property discussed above.

By choosing $\mathcal{A} = \phi^{(n)}$, the first term in the bracket of (2.49) becomes $\phi^{(n)}(z') \cdot \left(\psi, D\phi^{(n)}\right)^{-1} \cdot \left(\psi, f\right)$. This term should equal $\phi(z')$ "as well as possible". Hence the problem has been reduced to the question how ψ should be chosen in order to approximate the residual equation in an optimal fashion.

We adopt the variational point of view and take $\phi^{(n)}(z') \cdot \left(\psi, D\phi^{(n)}\right)^{-1} \cdot \left(\psi, f\right)$ with that ψ as the best approximation to $\phi(z')$ where the energy functional (2.20) is lowest. For gauge groups $U(1)$ and $SU(2)$ the computation of this optimal ψ is straightforward. The result is $\psi = \phi^{(n)}$.

Note that with $\psi = \phi^{(n)} = \mathcal{A}$, the error $\tilde{\Omega}$ of $\phi^{(n)}$ represented at the last site can be written as $\tilde{\Omega} = (CD\mathcal{A})^{-1} \cdot (Cr^{(n)})$ with $C = \mathcal{A}^*$. This is exactly the same equation

which one obtains for $\tilde{\Omega}$ when one adopts a variational MG algorithm with $C = \mathcal{A}^*$; cf. Eqs. (2.28a) and (2.28b).

Re-interpretation of Eq. (2.44).

By putting everything together, we end up with the following. We noted that the argument x_N in (2.44) is redundant, and we chose $\mathcal{A}(z) = \phi^{(n)}(z)$. Then (2.44) reads $\phi^{(n)}(z) \mapsto \phi^{(n)}(z) \cdot (\tilde{\Omega} + \mathbb{1})$. For $\tilde{\Omega}$ we made the choice $\tilde{\Omega} = \left(\phi^{(n)}, D\phi^{(n)}\right)^{-1} \cdot \left(\phi^{(n)}, r^{(n)}\right)$. Since $r^{(n)} = f - D\phi^{(n)}$, one has $\tilde{\Omega} + \mathbb{1} = \left(\phi^{(n)}, D\phi^{(n)}\right)^{-1} \cdot \left(\phi^{(n)}, f\right) \equiv \Omega$.

Summarizing, we re-interpret the updating (2.44) as a global rescaling of $\phi^{(n)}$ with a matrix Ω from the right:

$$\phi^{(n)}(z) \mapsto \phi^{(n)}(z) \cdot \Omega \quad \text{with} \quad \Omega = \left(\phi^{(n)}, D\phi^{(n)}\right)^{-1} \cdot \left(\phi^{(n)}, f\right) \ . \tag{2.50}$$

We note that the global rescaling by multiplying the vector $\phi^{(n)}$ with a matrix Ω from the right is defined because $\phi^{(n)}(z)$ is an $N_c \times N_c$ matrix. We also note that the matrix Ω depends on the iteration number $^{(n)}$ but we refrain from indicating that in the notation.

Since we assumed that f transforms like a matter field under gauge transformations, the $N_c \times N_c$ matrix Ω is gauge invariant, and (2.50) is therefore gauge covariant.

In case of gauge group $SU(2)$ the matrix $\left(\phi^{(n)}, D\phi^{(n)}\right)$ is always a multiple of the unit matrix because D is Hermitean. In other cases, e. g. in case of gauge group $SU(3)$, the matrix $\left(\phi^{(n)}, D\phi^{(n)}\right)$ can be inverted easily.

The case that f is a δ-function.

In QCD simulations with the hybrid Monte Carlo algorithm and in quenched computations one will not only need the case that f behaves like a matter field under gauge transformations, but also the case that f is a δ-function. Let us discuss the case that $f(z) = \delta_{z,w}\,\mathbb{1}$. This f is gauge invariant.

Then all approximate solutions $\phi^{(n)}$ in iterative algorithms do not transform according to (2.6), but according to $\phi^{(n)}(z) \mapsto g(z) \cdot \phi^{(n)}(z) \cdot g(w)^{-1}$. Therefore we cannot proceed as above. We are forced to identify w with the "last site" when we choose $\mathcal{A} = \phi^{(n)}$, but this does not harm. The kernel $\mathcal{A} = \phi^{(n)}$ has the gauge covariance property (2.33) in this case. [We remark that here w is a distinguished site which was not present in the situation discussed above where f was assumed to transform like a matter field.]

The expression given for Ω in (2.50) will not be gauge invariant. Rather Ω will transform like a matter field sitting at site w in the adjoint representation, i. e. $\Omega \mapsto g(w) \cdot \Omega \cdot g(w)^{-1}$. However, this is exactly what is required to make the operation $\phi^{(n)} \mapsto \phi^{(n)} \cdot \Omega$ gauge covariant. So everything is o. k. also in the case that f is a δ-function.

Outlook to the case of staggered fermions.

It was noted above that (2.50) will have to be modified in case of staggered fermions. This modification will be discussed in Sec. 7.5. Here we remark that gauge covariance of the fermionic analog of (2.50) is ensured in applications with the hybrid Monte Carlo algorithm [52].

Suppose that the r. h. s. f in Eq. (2.1) is a pseudofermion field. The fermionic matrices Ω are given by similar expressions as in (2.50). These expressions are also gauge invariant.

Therefore a global rescaling $\phi^{(n)} \mapsto \phi^{(n)} \cdot \Omega$ is also gauge covariant in case of staggered fermions. When f is a δ-function, analogous remarks as above apply.

Concluding remarks.

Note that $\Omega \to \mathbb{1}$ and that the error represented at the last site $\tilde{\Omega} \to 0$, as $\phi^{(n)} \to \phi$. A finding of numerical experiments reported later is that in practice $\Omega = \mathbb{1}$ as soon as errors decay exponentially. Then the step (2.50) can be switched off.

The expression given for the matrix Ω in (2.50) is the same which one obtains when one starts from replacing $\phi^{(n)}$ by $(\phi^{(n)} \cdot \Omega)$, with Ω chosen such that the energy functional (2.20) of the rescaled approximation $\phi^{(n)} \cdot \Omega$ gets minimized.

The proposal (2.50) was made in a recent paper [106]. Numerical results of its implementation will be presented in Secs. 4.8, 6.6, and 7.5.

We add as a comment that inclusion of (2.50) in a CG iteration would have no effect. The CG iterates $\phi^{(n)}$ have the property that $K[\phi^{(n)} \cdot \Omega]$ is minimal for $\Omega = \mathbb{1}$.

2.9 Related Improvements from a Variational Point of View

Some gauge covariant modifications of (MG) relaxation algorithms which are related to the updating on an MG layer consisting of a single site as described in Sec. 2.8 will be discussed now. The new parameters are not tunable, they are all determined by the algorithms themselves. The modifications require additional computational work, though.

Modified MG correction updating step.

In conventional MG approaches one considers updates of the form $\phi^{(n)} \mapsto \phi^{(n)} + \varphi^{(n)}$ where $\varphi^{(n)}$ is obtained by interpolation of an approximate solution of a residual equation on a coarser lattice. We propose to generalize this to

$$\phi^{(n)}(z) \mapsto \tilde{\phi}^{(n)} \equiv \phi^{(n)}(z) \cdot \Omega + \varphi^{(n)}(z) \cdot \Theta \ . \tag{2.51}$$

The two $N_c \times N_c$ matrices Ω and Θ are chosen such that the energy functional $K[\tilde{\phi}^{(n)}]$ is minimized. In particular, this proposal may be an improvement in algorithms where the residual equation is only solved approximately, or in algorithms were coarse grid operators are not defined through the Galerkin prescription.

Modified checkerboard SOR.

Consider for illustration the bosonic problem. When we update at the even sites, we propose to modify checkerboard SOR (2.10) according to

$$\begin{aligned} &\phi^{(n)}(z) \mapsto \phi^{(n)}(z) \cdot \Omega \ + \ \varphi^{(n)}(z) \cdot \Theta \quad \text{if } z \text{ is even} \ , \\ &\phi^{(n)}(z) \mapsto \phi^{(n)}(z) \cdot \Xi \qquad\qquad\quad \text{if } z \text{ is odd} \ , \end{aligned} \tag{2.52}$$

where $\varphi^{(n)}(z) = (2d + m^2)^{-1}[f(z) + \sum_{z' \text{n.n.} z} U(z, z') \cdot \phi^{(n)}(z')]$. Again, the three $N_c \times N_c$ matrices Ω, Θ and Ξ should be chosen such that the functional K gets minimized.

The proposal (2.52) expresses the view that in gauge theories one should have gauge invariant/covariant relaxation matrices rather than relaxation parameters that are real numbers.

Modified damped Jacobi relaxation.

Damped Jacobi relaxation can be generalized according to (2.51) with $\varphi^{(n)} = (2d + m^2)^{-1} r^{(n)}$. If one fixes $\Omega = \mathbb{1}$, one recovers (in case of $U(1)$ and $SU(2)$ gauge fields) the minimal residual algorithm that was used by Hulsebos et al. [73,107,108,109,110].

Gauge covariance.

We remark that also the proposals (2.51) and (2.52) respect gauge covariance. The matrices introduced here are gauge invariant when f transforms like a matter field under gauge transformations (e. g. when f is a pseudofermion field). When f is a δ-function, the matrices transform like a matter field in the adjoint representation, as discussed in Sec. 2.8.

From now on we will refrain again from explicitly indicating the multiplication of matrices by a dot.

2.10 Cautionary Remarks Concerning the Term "Critical Slowing Down"

It will be pointed out that for some algorithms the asymptotic CSD behavior can be predicted from studies on lattices of fixed size with small values of $\triangle m^2$.

Numerical results which will be presented in subsequent sections show that in case of traditional relaxation algorithms for bosonic propagators the predicted CSD behavior of τ, discussed in Sec. 2.2, can be monitored extremely well in practice. An important point is that the constant of proportionality in the scaling relation (2.15) does practically *not* depend on the lattice size; e. g. in Jacobi relaxation it equals $(2d + m_{cr}^2)$,Eq. (2.19). The quantity m_{cr}^2 depends only slightly on the lattice size, and smoothly on β, but $(2d + m_{cr}^2)$ depends on $|\Lambda|$, or on β, only very weakly. (We are interested in $d = 4$.)

Therefore the study of the *asymptotic* behavior of τ under conditions where the linear extension of the lattice and $1/\triangle m$ are changed proportionally (scaling conditions), can be predicted from studies in a fixed volume. Certainly, in physical applications (Monte Carlo simulations) the inverse mass should be smaller than the extension of the lattice, but this is an aspect of finite size effects on physical observables. The existence of the $1/\triangle m^2$ divergence (with a constant independent of the lattice size) on a lattice of fixed size implies the appearance of CSD in computations where all quantities are scaled appropriately.

To the author's knowledge there exists no study in the literature where m_{cr}^2 is not disregarded in case of staggered fermions. This neglect is only justified by the smallness of m_{cr}^2 but it has never been checked whether the neglect is justified. A result of the present work is the validity of the relation $\tau = \text{const}/\triangle m^2$ in traditional one-grid and

variational MG relaxation,[12] with a constant which is also *independent of the lattice size*. Therefore also for fermions the CSD of τ in these algorithms under scaling conditions can be predicted equivalently from studies on small lattices.

By neglecting m_{cr}^2 and trying to determine z under scaling conditions on small lattices ($\lesssim 18^4$), one can at most obtain some effective critical exponent z_{eff}. This z_{eff} contains however a great deal of arbitrariness and cannot be defined uniquely. One can run into difficulties with this procedure [107]. The author admits that a z_{eff} is of more practical relevance as long as numerical simulations are limited to lattice sizes where m is not really small. But we look for algorithms which can be used in future large scale computations, and for these it will be z and not z_{eff} which governs CSD.

There is one weak point in this reasoning. It is the disregard of a possibly remaining volume effect with respect to how many iterations it takes until decays of errors are exponential. This volume effect has to be taken into consideration when one wants to estimate CPU times. The volume effect also shows up in CG where it must not be disregarded, otherwise CG would be an ideal algorithm because in practice its "asymptotic" convergence rate is very rapid. Therefore one might have to go back to a z_{eff}. We will see later that such a volume effect is only appreciable when the proposal (2.50) and its generalization for staggered fermions is included in algorithms. For this discussion we refer to Secs. 6.6 and 7.5.

[12] In order to be precise, one should say variational MG with the "Laplacian choice" of the averaging operator C. This choice will be introduced in Sec. 3.2.

3 Gauge Covariant Block Spins

In this section we will discuss gauge covariant block averages for bosons, for staggered fermions and for gauge fields. In case of staggered fermions the blocking procedure is not obvious. If one wants to preserve as much as possible of the lattice symmetry group of free staggered fermions, one is, in particular, forced to choose an odd scale factor L_b; a factor-of-2 coarsening is therefore not allowed. An efficient numerical algorithm for the computation of the ground-state projecting averaging kernels will be presented. After that, results are given for optimal interpolation kernels in $SU(2)$ gauge fields in four dimensions. They prove that ground-state projection is a good choice of block spin in the presence of gauge fields. As a supplement a short review of Bałaban's analytical work on decay of propagators in gauge fields is given. Finally the notion of "smoothness" in gauge theories is discussed.

3.1 Block Spin for Bosons

Block spins for bosons are defined by means of the ground-state projecting kernel C described in Sec. 2.7. Probability distributions of $\|C(x,z)\|$ in four-dimensional $SU(2)$ gauge fields are given.

For bosonic matter fields ϕ one can define block lattices in the "naive" straightforward way which was described in Sec. 2.3. This blocking is illustrated in Fig. 3.1 for a two-dimensional lattice. Block spins are defined as in Eq. (2.36), where the restriction operator C is defined by the ground-state projection definition (2.37), (2.38), and (2.43). The links, where both endpoints are in the same block and which yield contributions to the Laplacian with Neumann boundary conditions (2.40), are drawn with thick lines in Fig. 3.1. The links drawn with thin lines connect neighboring blocks. The gauge field on these links is not relevant for C.

In Sec. 3.5 it will be shown that ground-state projection defines a "good" block spin for bosons. This means that the associated optimal interpolation kernel $\mathcal{A}$ decays exponentially (see also Appendix B). It was mentioned in Sec. 2.6 that $C(x,z)$ will not be an element of the gauge group in general. Rather, $C(x,z)$ will be in the linear span[13] of

[13] i. e. $C(x,z)$ will be a real multiple of an element of $SU(2)$ if $G = SU(2)$, an arbitrary complex $N \times N$ matrix for $G = U(N)$, $N \geq 3$ etc; see Ref. [103].

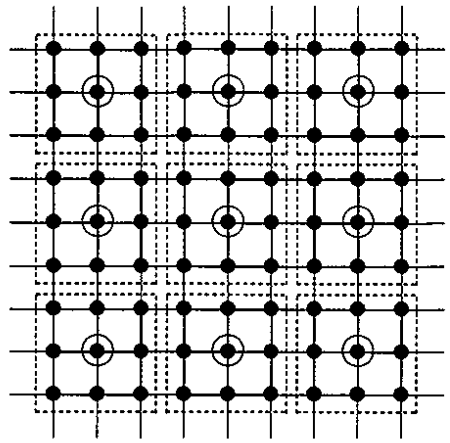


Fig. 3.1: *Division of a two-dimensional bosonic lattice Λ into blocks. Sites $z \in \Lambda$ are marked by $\bullet$. A scale factor of $L_b = 3$ is chosen. Blocks x are bordered by dashed lines and may be identified with the block centers $\hat{x}$ which are encircled. Links inside blocks are drawn with thick lines, while links which connect sites in different blocks are drawn with thin lines.*

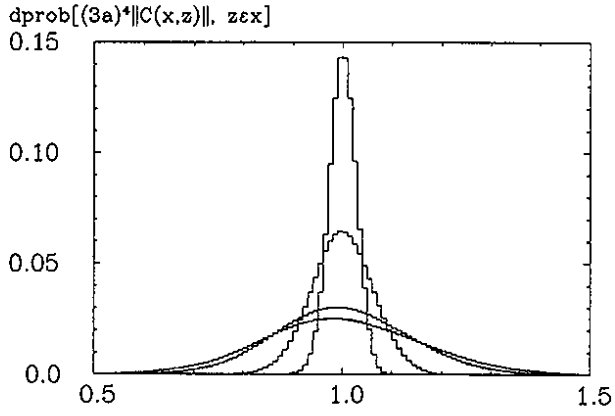


Fig. 3.2: *Averaging kernel C for bosons: Probability distributions of the dimensionless quantity $(L_b a)^d \|C(x,z)\|$, $z \in x$, in four-dimensional pure $SU(2)$ gauge theory. The norm equals $\|C(x,z)\| = [\frac{1}{2}\mathrm{Tr}\, C(x,z)^\dagger C(x,z)]^{1/2}$. The four curves shown correspond to $\beta = 10.0$, 5.0, 3.0, 2.7, with increasing width. In a pure gauge ($\beta = \infty$), $(L_b a)^d \|C(x,z)\| \equiv 1$ for $z \in x$.*

the gauge group. This fact is shown in Fig. 3.2 for four-dimensional $SU(2)$. The probability distributions were obtained from gauge field configurations on 18^4 lattices which were equilibrated with the Wilson action [1], $\mathcal{H}(U) = \beta \sum_P [1 - \frac{1}{2}\mathrm{Tr}\, U(P)]$. Brower, Rebbi and Vicari [111] proposed a variational method for approximating ground-state projection kernels. Their variational principle restricts $(L_b a)^d \, \|C(x,z)\|$ to 1. For finite β the solution of Eq. (2.37) is not well approximated by this procedure, but this does not imply that the performance of MG algorithms with this choice is necessarily bad.

3.2 Block Spins for Staggered Fermions

One of the first questions which has to be answered in MG algorithms is the question for the choice of block lattices. The answer to this question is obvious in case of bosons, but it is not straightforward in case of staggered fermions. Different strategies were pursued in the literature [72,73]. The blocking procedure proposed in this section starts from the requirement that as much as possible is preserved of internal and space-time symmetries in the limiting case of vanishing gauge coupling. As a result we are forced to choose a scale factor of 3 (or any other odd number). Another consequence is the emergence of seemingly overlapping blocks. In the limiting case of a pure gauge the blocks have actually no sites in common. But in nontrivial gauge fields the symmetries of free staggered fermions are broken, and for this reason one cannot a priori rule out the possibility that the blocks overlap in a nontrivial way. Two qualitatively different proposals will be made for the choice of the restriction operator C. It will also be defined as solution of a gauge covariant eigenvalue equation. We will call the two proposals the "Laplacian choice" and the "Diracian choice". Both proposals reduce to a successful construction in the limiting case of free staggered fermions. The Diracian choice is preferable for physical reasons, but it is computationally quite demanding. The reason is that blocks overlap in a nontrivial way, i. e. every fine grid site makes contributions to the block spin at more than just one coarse grid site. The Laplacian choice is more convenient from the numerical point of view. It retains the property that also in nontrivial gauge fields the seemingly overlapping blocks have actually no sites in common, i. e. every fine grid site makes a contribution to the average at exactly one coarse grid site.

This section is based on joint work with G. Mack and M. Speh [112].

Staggered fermions reduce the number of doublers of the naive lattice fermion action while retaining some aspects of chiral symmetry [61,62,63]. The idea is to reduce the number of spinor/flavor components at each lattice site to one. For simplicity, let the dimension d of space-time be even. ($d = 4$ is the case we are interested in.) The different fermionic degrees of freedom are distributed on the sites of hypercubes of volume 2^d sites.

In terms of crystallography as used in solid state physics [113], the lattice of staggered fermions is a hypercubic lattice Λ_a of unit cells of lattice spacing a with a 2^d-point basis. The basis is given by the sites of a hypercube of volume $(a/2)^d$. Sites within the basis are distinguished by their "pseudoflavor" H. Different pseudoflavors would correspond to different species of atoms or ions in the solid state analog. *If one does not distinguish*

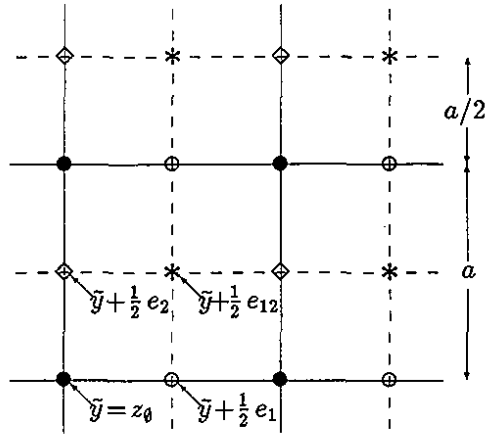


Fig. 3.3: *Distribution of pseudoflavor degrees of freedom in a two-dimensional lattice. Sites of pseudoflavor sublattices Λ_a^H, $H = \emptyset, \{1\}, \{2\}, \{1,2\}$, are marked with $\bullet$, $\circ$, $\diamond$, $*$, respectively.*

different pseudoflavors, one would have a hypercubic lattice $\Lambda_{a/2}$ of lattice spacing $a/2$. A summary about staggered fermions can be found in Appendix A, together with our notations and conventions.

3.2.1 Blocking Consistent with the Symmetries of Free Staggered Fermions

Pseudoflavor H is specified by an ordered set of distinct indices μ_i,

$$H = \{\mu_1, \ldots, \mu_h \mid 0 \le h < d,\ \mu_1 < \cdots \mu_h\} \ . \tag{3.1}$$

Sites $z = z_H \in \Lambda_{a/2}$ with pseudoflavor H are of the form

$$z_H = \tilde{y} + \tfrac{1}{2} e_H \ , \quad e_H = \sum_{\mu \in H} e_\mu \ , \tag{3.2}$$

where e_μ is a lattice vector of length a in μ-direction in Λ_a. $\tilde{y}$ is a site in Λ_a which is identified with site $z_\emptyset \in \Lambda_{a/2}$. The sites z_H with fixed H form a sublattice Λ_a^H of $\Lambda_{a/2}$, viz. $\Lambda_a^H = \Lambda_a + \frac{1}{2} e_H$. [In principle any sublattice Λ_a^H could be identified with Λ_a.] Fig. 3.3 illustrates this.

From now on we will write χ instead of ϕ when we explicitly deal with staggered fermion fields. *Let us first consider the free case.* Becher and Joos [64] showed that the formal continuum limit of the staggered Dirac operator is equivalent to the Dirac-Kähler operator. The Dirac-Kähler equation is a differential geometric equation for continuum fermions. There Dirac fields are described by differential forms. For this reason the lattice discretization is straightforward. *Joos and Schaefer [114] studied the symmetry group of*

staggered fermions in great detail, see also Ref. [115]. We call this group $\mathcal{G}_a$. Joos and Schaefer pointed out that $\mathcal{G}_a$ is a restriction of the continuum symmetry group of Dirac-Kähler fermions. We require that a blocking for staggered fermions is defined in such a way that as much as possible is preserved of the residual lattice symmetries. This means that the symmetry group $\mathcal{G}_{L_b a}$ on the staggered block lattice $\Lambda_{L_b a/2}$ should be a subgroup of $\mathcal{G}_a$.

The fine-to-coarse blocking step maps χ which lives on $\Lambda_{a/2}$ onto a field $\boldsymbol{\chi}$ on the block lattice $\Lambda_{L_b a/2}$ of lattice spacing $L_b a/2$, with an integer scale factor $L_b > 1$:

$$\boldsymbol{\chi}(x) = \int_{z \in \Lambda_{a/2}} C(x,z)\chi(z) \ . \tag{3.3}$$

The adjoint spinor $\overline{\chi}(z)$ is an independent Grassmannian integration variable. Its block spin can be defined by

$$\overline{\boldsymbol{\chi}}(x) = \int_{z \in \Lambda_{a/2}} \overline{\chi}(z) C^*(z,x) \ , \tag{3.4}$$

where $C^*(z,x)$ is the adjoint of the kernel C which appears in (3.3). This does not contradict the statement of G. Mai.in Ref. [116] that one needs different blocking procedures for adjoints. She considered blockings from the continuum to the lattice. The subtlety which she encountered is hidden in the fact that pseudoflavor H for χ and for $\overline{\chi}$ does not have the same physical meaning in terms of spin and flavor.

Emergence of odd L_b.

The first result which we derive from symmetry considerations is that L_b *must be odd*, i. e. $L_b = 3, 5, \ldots$, but *not* $L_b = 2$. This results arises as follows. On the lattice translations by integer multiples $\tilde{n}a$ of a (not $a/2$) survive as true translation symmetries. Pure flavor transformations do not survive as symmetries at all. What survives are certain combinations ϵd^H of flavor transformations and translations by $\frac{1}{2}e_H$, where $\epsilon = \pm 1$. We call them pseudoflavor transformations. The square $(d^H)^2$ of any such transformation is a true lattice translation by $-e_H$. To keep track of the dependence on the lattice spacing, we write $T_a(*)$ for the action of an element $*$ of $\mathcal{G}_a$. The action of translations is the obvious one. In particular, translations by the vectors e_H in Λ_a act as

$$(T_a(e_H)\chi)(z) = \chi(z - e_H) \ . \tag{3.5}$$

Pseudoflavor transformations ϵd^H act on the staggered fermion field as

$$(T_a(\epsilon d^K)\chi)(z) = \epsilon\, \check{\rho}_{H,K}\, \chi(z + \tfrac{1}{2}e_K) \ , \quad \text{if } z \in \Lambda_a^H \ . \tag{3.6}$$

$\check{\rho}_{H,K}$ is a sign factor which is the same as that which appears in the Clifford product $\vee$ of forms, defined by $dx^\mu \vee dx^\nu = dx^\mu \wedge dx^\nu + \delta^{\mu\nu}$, and

$$dx^H \vee dx^K = \check{\rho}_{H,K}\, dx^{H \triangle K} \ . \tag{3.7}$$

$H \triangle K$ denotes the symmetric difference $H \triangle K = (H \cup K) \setminus (H \cap K)$, and $dx^H = dx^{\mu_1} \wedge \cdots \wedge dx^{\mu_h}$. The sign factors η_μ in the Diracian (2.5) are special $\check{\rho}$-factors: $\eta_\mu(z) = \check{\rho}_{H,K}$ for $z \in \Lambda_a^K$ and $H = \{\mu\}$.

In the following, primed quantities refer to the block lattice. We write e'_μ for the lattice vectors of length $L_b a$ in $\Lambda_{L_b a}$, and similarly for e'_H, etc. Thus,

$$e'_H = L_b\, e_H \ . \tag{3.8}$$

The elementary pseudoflavor transformations d'^K on $\Lambda_{L_b a/2}$ must be linear combinations of pseudoflavor transformations d^K and translations e on $\Lambda_{a/2}$. From (3.6), it can be seen that d'^K contains an admixture of translation by $-\frac{1}{2}e'_K$. Since d^K contains an admixture of a translation by $\frac{1}{2}e_K$, we must have

$$-\frac{1}{2}e'_K = -\frac{1}{2}e_H + e \quad \text{with} \quad e = \sum_\mu n_\mu e_\mu \ , n_\mu \in \mathbb{N} \tag{3.9}$$

for some H. Because of (3.8), this is only possible if $H = K$, and $e = \frac{1}{2}(L_b - 1)e_K$. This requires that L_b is odd,

$$L_b = 3, 5, \ldots \tag{3.10}$$

The result (3.10) might have been expected, since only for odd L_b the pseudoflavors on the fine lattice $\Lambda_{a/2}$ and on the block lattice $\Lambda_{L_b a/2}$ match nicely when we regard $\Lambda_{L_b a/2}$ as a sublattice of $\Lambda_{a/2}$. (See Fig. 3.4 below).

3.2.2 The Averaging Kernel $C(x,z)$ in Trivial Gauge Fields

Now we turn to the averaging kernel $C(x,z)$ for free staggered fermions. Since the blocking should be consistent with the symmetries, we require that C commutes with the action of the surviving symmetries on the block lattice. For translations and pseudoflavor transformations this means

$$T_{L_b a}(e'_K)\, C \ = \ C\, T_a(L_b e_K) \ , \tag{3.11}$$

$$T_{L_b a}(\epsilon d'^K)\, C \ = \ C\, T_a(\epsilon d^K)\, T_a(\tfrac{1}{2}(L_b - 1)\, e_K) \ . \tag{3.12}$$

Eqs. (3.11) and (3.12) are fulfilled if C is translation invariant,

$$C(x + e'_H\,,\, z + L_b e_H) = C(x,z) \ , \tag{3.13}$$

and if in addition C has the support property

$$C(x,z) = 0 \quad \text{unless } x \text{ and } z \text{ carry the same pseudoflavor.} \tag{3.14}$$

In words, the support property states the following. If $x \in \Lambda^H_{L_b a}$, then the staggered block spin $\boldsymbol{\chi}(x)$ is a weighted sum of fields $\chi(z)$ at sites with the same pseudoflavor H. Blockings with such a property were also considered by Ben-Av et al. [72,100,117,101,102,118]. However, they chose a scale factor $L_b = 2$, and this forces them to treat different pseudoflavors on different footings. Another blocking procedure with non-overlapping blocks was

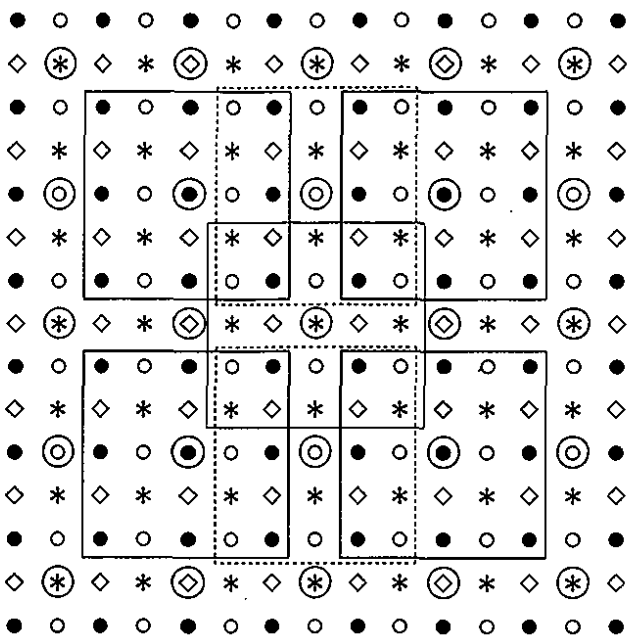


Fig. 3.4: *The good choice of blocks to average over in a two-dimensional pseudoflavor lattice. This blocking leads to staggered fermions on the coarse lattice again. Block centers $\widehat{x}$ are encircled (for $L_b = 3$). The boundaries of seven blocks are marked. The averaging kernel $C(x,z)$ for free staggered fermions (and in nontrivial gauge fields for the "Laplacian choice"; see below) is only nonvanishing if site z has the same pseudoflavor (symbol) as the block center $\widehat{x}$. Therefore the seemingly overlapping blocks have actually no sites in common in these cases.*

used by Hulsebos et al. [73,107,108,109,110]. But in their proposal one does not retain the successful pure gauge limit.

Blocking procedure for staggered fermions.

The blocking procedure which is consistent with the symmetries of free staggered fermions is as follows. With each site x of the block lattice $\Lambda_{L_b a/2}$ we associate a hypercube $X \subset \Lambda_{a/2}$ of sidelength $L_b a$, that is $2L_b$ lattice spacings $a/2$. When considered as a site in $\Lambda_{a/2}$, x will be the point in the center of this hypercube. These hypercubes X overlap. If $x \in \Lambda^H_{L_b a}$ then we call $z \in X$ an active site if $z \in \Lambda^H_a$, i. e. if it carries the same pseudoflavor as the central point x of the hypercube X. Every point z in $\Lambda_{a/2}$ is active site in exactly one hypercube X. Every hypercube X contains L_b^d active sites. Our choice of blocks is illustrated in Fig. 3.4. The block boundaries in Fig. 3.4 may be shifted away from the block centers by $a/2$ without any effect. Doing so would result in an agreement with the hypercubes X discussed above, but the figure were harder to read.

It is necessary to admit overlapping blocks x if one insists that $\chi(z)$ makes a contribution to some block spin $\chi(x)$ for any z, i. e. in order that every z is active site of some

block (no decimation).

"Piecewise-constant" kernels for free staggered fermions.

If $x \in \Lambda^H_{L_b a}$ then we set

$$C(x,z) = \begin{cases} (L_b a/2)^{-d} \cdot 1 & \text{if } z \in x \cap \Lambda^H_a \ , \\ 0 & \text{otherwise} \ . \end{cases} \tag{3.15}$$

$C(x,z)$ is only nonvanishing on active sites z of hypercube X, so that it has disjoint support in z for different x. Conditions (3.13) and (3.14) are fulfilled.

We remark that from the point of view of computation of propagators, the equation $(-\not{D}^2 + m^2)\chi = f$ is in the absence of gauge fields equivalent to computing 2^d uncoupled free bosonic propagators, one for each pseudoflavor. The choice (3.15) is the natural fermionic analog to the bosonic choice (2.35).

A variant of the blocking procedure described above was first introduced by G. Mai [116]. She considered blocking from the continuum to a lattice.

Eigenvalue equation for C and Neumann b. c. for staggered fermions.

The kernel (3.15) can also be defined as solution of an eigenvalue equation, analogously to Eq. (2.37),

$$(-\Delta_{N,x} C^*)(z,x) = \lambda_0(x)\, C^*(z,x) \ , \tag{3.16}$$

where $\lambda_0(x)$ is again the lowest eigenvalue of $-\Delta_{N,x}$. However, the choice for the fermionic analog of the bosonic $\Delta_{N,x}$ has to be made with some care. In the absence of gauge fields the square of the Diracian (2.5) equals a Laplacian (as in the continuum),

$$(\Delta\chi)(z) = \frac{1}{a^2} \sum_{\substack{\mu=-d \\ \mu\neq 0}}^{d} [\chi(z+e_\mu) - \chi(z)] \ . \tag{3.17}$$

This Δ couples only sites z with the same pseudoflavor. The kernel (3.15) satisfies Eq. (3.16) if we choose the Laplacian (3.17), and if we define Neumann boundary conditions as follows. If $x \in \Lambda^H_{L_b a}$, then $\Delta_{N,x}$ is defined by the same formula (3.17) as Δ, except that those terms in the sum are omitted where z and $z+e_\mu$ are both in Λ^H_a, but one is in block x and the other one is outside x. (Cf. the discussion in Sec. 2.7.) We also have to impose the subsidiary condition $C(x,z) = 0$ if $z \notin x$, and the normalization condition $CC^* = \mathbb{1}$.

So far we have imposed the requirement (3.14) as a separate condition. However, this condition could also be regarded as a consequence of the eigenvalue equation. If we restrict attention to sublattices Λ^K_a with $K \neq H$, then the modification of Δ through "Neumann boundary conditions" as stated above becomes inoperative, and the only solution of the eigenvalue equation (3.16) which vanishes outside x is $C(z,x) = 0$. This way of specifying the solution amounts to supplementing the Neumann boundary conditions on Λ^H_a by Dirichlet boundary conditions on Λ^K_a, $K \neq H$.

3.2.3 The Averaging Kernel $C(x,z)$ in Nontrivial Gauge Fields

We have to discuss now what happens when gauge fields are included.

Square of the covariant Dirac operator.

In this case the square of the Diracian is no longer just a Laplacian, but instead

$$(-\not{D}^2\chi)(z) = \left([-\Delta - \sum_{|\mu|\neq|\nu|} \sigma_{\mu\nu}(z) F^*_{\mu\nu}(z)]\chi \right)(z) \ . \tag{3.18}$$

Indices μ, ν are summed over positive and negative values, but $|\mu| \neq |\nu|$. The gauge covariant Laplacian for staggered fermions is given by

$$(\Delta\chi)(z) = \frac{1}{a^2} \sum_{\substack{\mu=-d \\ \mu\neq 0}}^{d} [U(z,2\mu)\,\chi(z+e_\mu) - U(z,\pm\mu)\,\chi(z)] \ , \tag{3.19}$$

where $U(z,2\mu) \equiv U(z,z+\frac{1}{2}e_\mu)\,U(z+\frac{1}{2}e_\mu,z+e_\mu)$ and $U(z,\pm\mu) \equiv U(z,z+\frac{1}{2}e_\mu)\,U(z+\frac{1}{2}e_\mu,z)$. The "forward-backward" parallel transport $U(z,\pm\mu)$ equals $\mathbb{1}$ for ordinary gauge fields, but this is not true when one admits dielectric gauge fields.
$F_{\mu\nu}$ in Eq. (3.18) is the natural definition of the field strength on the lattice [103],

$$\begin{aligned} a^2 F_{\mu\nu}(z) \ = \ & U(z+\tfrac{1}{2}e_\mu+\tfrac{1}{2}e_\nu\,,\,z+\tfrac{1}{2}e_\mu)\,U(z+\tfrac{1}{2}e_\mu\,,\,z) \\ & -U(z+\tfrac{1}{2}e_\mu+\tfrac{1}{2}e_\nu\,,\,z+\tfrac{1}{2}e_\nu)\,U(z+\tfrac{1}{2}e_\nu\,,\,z) \ . \end{aligned} \tag{3.20}$$

$\sigma_{\mu\nu}$ is a commutator of η-factors.

Eq. (3.18) is the lattice analog of the continuum equation $-\not{D}^2 = -D_\mu^2 - \frac{1}{2}g\sigma_{\mu\nu}F_{\mu\nu}$ with the usual definition of the covariant derivative $D_\mu = \partial_\mu + igA_\mu$ and the Yang-Mills field strength $F_{\mu\nu} = [D_\mu, D_\nu]/(ig)$, and $\sigma_{\mu\nu} = \frac{1}{2}i[\gamma_\mu,\gamma_\nu]$.

The Laplacian choice of $C(x,z)$.

For nonvanishing field strength (i. e. nontrivial gauge fields) it remains true that Δ couples only sites with the same pseudoflavor, but the same is not true of $\not{D}^2$. One possible generalization of (3.16) is now as follows. If we insist in retaining the support property (3.14) for the averaging kernel C, we could simply replace (3.17) by the covariant Laplacian (3.19). We call this the "Laplacian choice of C". [$-\Delta_{N,x}$ is a positive semidefinite operator, considered as acting on functions which vanish on sites $z \notin \Lambda_a^H$ when $x \in \Lambda^H_{L_b a}$.] In order make the definition of C unique and gauge covariant, we have also to impose the normalization condition $CC^* = \mathbb{1}$ and the covariance condition (2.43).

The kernel C of the Laplacian choice can be computed numerically with the same efficiency as the averaging kernel (2.37) for bosons (see Sec. 3.4). One can also use the same computer program, the only changes which have to be made are replacements of address lists and of parallel transporters in the Laplacian. In Sec. 3.5 we will see that the Laplacian choice of C defines a good block spin in the sense described in the previous section. Probability distributions of $\|C(x,z)\|$ in four-dimensional $SU(2)$ gauge fields are shown in Fig. 3.5.

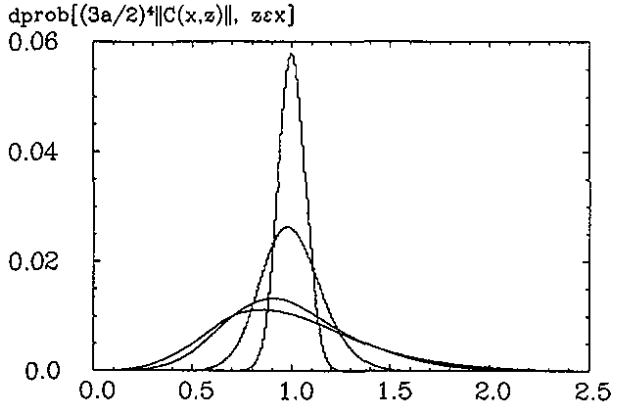


Fig. 3.5: *Averaging kernel C of the Laplacian choice for staggered fermions: Probability distributions of the dimensionless quantity $(L_b a/2)^d \,\|C(x,z)\|$, $z \in x \cap \Lambda^H_{a/2}$ for $x \in \Lambda^H_{a/2}$, in four-dimensional pure $SU(2)$ gauge theory. The norm equals $\|C(x,z)\| = [\frac{1}{2}\mathrm{Tr}\,C(x,z)^\dagger C(x,z)]^{1/2}$. The four curves shown correspond to β = 10.0, 5.0, 3.0, 2.7, with increasing width.*

The Diracian choice of $C(x,z)$.

The authors of Ref. [100] make the statement that it is compulsory for the averaging map C to have the support property (3.14) also in the presence of gauge fields. We do not agree with this. As an alternative to the above scheme one could think of using the square of the Dirac operator in place of the Laplacian in the eigenvalue equation (3.16). This may seem more natural and is closer to the philosophy of ground-state projection MG, but it leads to a much greater computational complexity and storage space requirements. Let us discuss these points in more detail.

We may define $\not{D}_{N,x}\chi(z)$ for $z \in x$ by (2.5) with with the modification that those terms in the sum over μ are omitted for which $z \pm \frac{1}{2}e_\mu$ are both in Λ_a^H, but one is in block x while the other is not in x. Setting

$$-\not{D}^2_{N,x} = \not{D}^*_{N,x}\not{D}_{N,x} \ , \tag{3.21}$$

one finds that $-\not{D}^2_{N,x}\chi(z)$ agrees with $-\Delta_{N,x}\chi(z)$ for $z \in \Lambda_a^H$ when $F_{\mu\nu}(z) = 0$. We might require that the averaging kernel C is a solution of the eigenvalue equation (3.16) with $-\not{D}^2_{N,x}$ substituted for $-\Delta_{N,x}$. But the square of the Dirac operator couples sites with different pseudoflavors. Therefore it is in general not consistent with the eigenvalue equation to require the support property (3.14). Instead we should supplement the boundary conditions. The requirement that $C(x,z)$ should vanish for sites outside x but not in Λ_a^H can be built into the eigenvalue equation as an additional boundary condition by setting

$$\not{D}^2_{N,D,x} = \Theta_x \not{D}^2_{N,x} \Theta_x \ , \tag{3.22}$$

where $\Theta_x\chi(z) = \chi(z)$ if $z \in x \cup \Lambda_a^H$, and $= 0$ otherwise.

The eigenvalue equation becomes

$$(-\not{D}^2_{N,D,x} C^*)(z,x) = \lambda_0(x)\, C^*(z,x) \ , \tag{3.23}$$

where $\lambda_0(x)$ is the lowest eigenvalue of the positive semidefinite operator $-\not{D}^2_{N,D,x}$ (considered as acting on functions which vanish on odd sites z if x is odd and on even sites otherwise). We call this the "Diracian choice of C".

We seek solutions which vanish outside x. In general the solution will not vanish when x and z carry different pseudoflavor. In the presence of a gauge field, the Dirac operator is *no longer translation invariant. Since pseudoflavor transformations contain an admixture* of translations, it is also not invariant under pseudoflavor transformations. For this reason we cannot rely on symmetry to rule out the possibility that Eq. (3.14) is violated. If we call z an active site of block x when $C(x,z) \neq 0$, then it is no longer true that any site z is an active site of exactly one block. In this sense, the blocks will overlap in a nontrivial way. In numerical applications to MG computations of nearly massless propagators for staggered *fermions in a gauge field U this leads to a great increase in computational complexity* and in storage space requirements.

The overlap problem could be ameliorated, but not eliminated completely, by introducing non-overlapping blocks $[x]$ and requiring that $C(x,z)$ vanishes outside $[x]$ when $z \notin \Lambda_a^H$. This is implemented by redefining Θ_x in Eq. (3.22): $\Theta_x \chi(z) = \chi(z)$ for $z \in [x] \cup \Lambda_a^H$, and $= 0$ otherwise. The redefinition amounts to imposing Neumann boundary conditions on larger blocks x in $\Lambda^H_{L_b a}$ and Dirichlet boundary conditions on smaller blocks $[x]$ in $\Lambda^K_{L_b a}$, $K \neq H$. (Remember that H is fixed by x). Proceeding in this way, the old solution for C is retained in the limit of a pure gauge, and therefore the blocking is consistent with the symmetries in the limit of a pure gauge.

The maximal non-overlapping hypercubes $[x]$ are given by[14] $z \in [x]$ if $\|z - [\hat{x}]\|_\infty \leq \frac{1}{2}(L_b - 1)\frac{a}{2}$. The overlap problem is eliminated if we choose $[x] = \emptyset$ instead. But this leads back to the Laplacian form of the eigenvalue equation which was considered before.

3.3 Blocked Gauge Fields

It is proposed to define blocked gauge fields by using interpolation or averaging kernels for matter fields. The blocked gauge fields are of "dielectric" type, but they may be brought back into the gauge group by performing a polar decomposition [103].

Remember that the kernels of the Laplacian and of the Dirac operator, Eqs. (2.2) and (2.4), contain the lattice gauge field U. A block spin definition for matter fields gives us the effective operators on coarser layers, $\Delta_{eff} = \mathcal{A}^* \Delta \mathcal{A}$ and $\not{D}_{eff} = \mathcal{A}^* \not{D} \mathcal{A}$. Given Δ_{eff} or $\not{D}_{eff}$, we have candidates for gauge fields $\mathcal{U}$ on the block lattice,

$$\begin{aligned} \mathcal{U}(x,y) &= (L_b a)^{d+2} \Delta_{eff}(x,y) \quad \text{for } x,y \text{ nearest neighbors,} \\ \text{and} \quad \mathcal{U}(x,y) &= (L_b a)^{d+1} \not{D}_{eff}(x,y) \,/\, \eta_\mu(x) \quad \text{if } y = x + L_b \tfrac{1}{2} e_\mu \ , \end{aligned} \tag{3.24}$$

[14] $\|\cdot\|_\infty$ denotes the Chebyshev norm, as usual, i. e. $\|w\|_\infty = \max_\mu |w_\mu|$.

respectively. These expressions for $\mathcal{U}(x,y)$ have the correct gauge covariance property.

We add some comments on the Dirac case. Since we use odd L_b, the staggered block lattice can be regarded as a sublattice of $\Lambda_{a/2}$ in which the pseudoflavor of a site does not depend on whether we regard it as a site of the fundamental or the block lattice, and the same is true of its property of being even or odd. The kernel $\mathcal{A}$ will always have the property that $\mathcal{A}(z,x) = 0$ unless z and x are both even or both odd. Therefore $\not{D}_{eff}$ connects even to odd sites. The factor $\eta_\mu(x)^{-1}$ can be regarded as defined by regarding x as a point in a sublattice of the fundamental lattice. This agrees with the natural definition on the block lattice. Note also that $\eta_\mu(z)$ is independent of the μ-component z_μ of z.

If one makes the variational choice $\mathcal{A} = C^*$, the above definition of $\mathcal{U}$, for instance for gauge fields interacting with bosons, reads explicitly:

$$\mathcal{U}(x,y) = L_b^{-d} \sum_{(z,z'): z \in x, z' \in y} C(x,z)\, U(z,z')\, C(y,z')^\dagger \ . \tag{3.25}$$

Summation is over those links (z,z') where z is in block x and z' is in the neighboring block y. (These are the "thin" links in Fig. 3.1.)

The blocked gauge fields $\mathcal{U}(x,y)$ are actually "dielectric" gauge fields [103]. In case of gauge group $SU(2)$ they are equal to an $SU(2)$ lattice gauge field multiplied with a fluctuating positive real factor ("fluctuating length"). [In renormalization group studies the occurrence of blocked fields which are not of the same kind as those in the original theory is not uncommon [39,34].] The dielectric gauge fields may be brought back into the gauge group by multiplying with a positive real number (for $U(1)$ and $SU(2)$) or with a positive matrix (for $U(N), N > 2$).

A different choice of block spin for gauge fields was introduced by Balaban and used in rigorous renormalization group work [104,119], see also [120], and [8].[15]

Another possible definition of dielectric block spins for gauge fields, which is similar to Balaban's choice, involves C and is as follows.

$$\mathcal{U}(x + L_b e_\mu, x) = (L_b a)^d \int_{z \in x} C(x + L_b e_\mu \,,\, z + L_b e_\mu)\, U(z + L_b e_\mu, z)\, C(x,z)^\dagger \ , \tag{3.26}$$

where $U(z + L_b e_\mu, z)$ denotes the path-ordered product of link variables along the straight line from z to $z + L e_\mu$. Actually, Eq. (3.26) defines a blocked gauge field interacting with bosons. In case of U coupled to staggered fermions, one has to replace e_μ by $\frac{1}{2} e_\mu$ (of course).

Finally, we note that the definitions given in this section for blocked gauge fields have the property that a pure gauge U on the finer lattice is mapped onto a pure gauge $\mathcal{U}$ on the coarser lattice. In that case, $\mathcal{U}(x,y)$ equals (up to a constant factor) the parallel transport along an arbitrary path from $\hat{y}$ to $\hat{x}$. [It is the path-independence of parallel transporters which characterizes a pure gauge in dielectric gauge fields, rather than the definition given in footnote 2), page 12.]

[15] Readers (including the present author) who find Balaban's works hard to read, may seek comfort in A.S. Wightman's words in Ref. [121]. An easier-to-read introduction to constructive gauge theories are Imbrie's Les Houches lectures [122]. There is also a recent text book on constructive field theory by Rivasseau [123].

3.4 Computation of Averaging Kernels

In numerical applications the averaging kernel C has to be recomputed whenever the gauge field U has been updated. An efficient algorithm is presented in this section which allows to compute the solution of the eigenvalue equation (2.37) numerically. This algorithm has also been described in Ref. [105]. The same algorithm can be used to compute the kernel C of the Laplacian choice for staggered fermions. There the algorithm has the same efficiency, and one can also use the same computer program; the only changes which have to be made are replacements of address lists and of parallel transporters in the Laplacian. The relevance of the eigenvalue spectrum of $-\Delta_{N,x}$ for the efficiency of the algorithm will be pointed out. Results of the numerical implementation will be reported.

3.4.1 An Algorithm for the Solution of the Eigenvalue Equation (2.37)

The eigenvalue problem (2.37) will be solved by means of Wielandt or inverse iteration. In each stage of this algorithm one has to solve a linear system of equations. Its solution will be determined by using the CG algorithm. The efficiency of the inverse iteration method depends on the quality of the estimate for the eigenvalue corresponding to the sought eigenvector. When determining the C-kernel, this estimate shall be close to the lowest eigenvalue $\lambda_0(x)$ of $-\Delta_{N,x}$, but we must demand that it is not greater than $\lambda_0(x)$. An efficient method for bounding eigenvalues from below is provided by Temple's Theorem.

Inverse iteration.

Inverse iteration is a method by which one can determine all eigenvalues and eigenvectors of a quadratic matrix[16] A, provided that sufficiently good approximate values for the eigenvalues are known.

Let $\lambda_0, \ldots, \lambda_m$ be the eigenvalues of A. Suppose that we know a good approximation λ to λ_j, say, i.e. $|\lambda_j - \lambda| \ll |\lambda_k - \lambda|$ for all $\lambda_k \neq \lambda_j$, but $\lambda \neq \lambda_j$. For finding an eigenvector x_j with eigenvalue λ_j, one starts with an arbitrary vector $x^{(0)}$ and computes a sequence of vectors $x^{(1)}, x^{(2)}, \ldots$, according to

$$(A - \lambda)x^{(i)} = x^{(i-1)} \quad , \quad i = 1, 2, 3, \ldots \quad . \tag{3.27}$$

It is easy to see by expanding the $x^{(i)}$ in terms of eigenvectors[17] of A that the iteration prescription (3.27) yields an eigenvector x_j. If there is a degeneracy in λ_j, one obtains a linear combination of eigenvectors with eigenvalue λ_j. If the condition $|\lambda_j - \lambda| \ll |\lambda_k - \lambda|$ for all $\lambda_k \neq \lambda_j$ cannot be fulfilled (i. e. if λ_j is nearly degenerate with one λ_k), the efficiency of inverse iteration is lessened. One would need more iterations in that case.
See Refs. [82,94,93] for a more comprehensive discussion on the method of inverse iteration.

Application of the method for solving Eq. (2.37).

We shall modify the rudimentary method (3.27) a little bit. In order to keep the components of the iterated vectors finite, it is advisable to normalize them. Also, the rate of convergence of the inverse iteration is faster the closer λ is to λ_j. One can therefore try to accelerate the method by updating the estimate λ in course of the procedure. To solve Eq. (2.37) we are thus led to the following form of inverse iteration: Choose a normalized $C_x^{*(0)} = C^{*(0)}(\cdot, x)$ for every block x and iterate

$$\left(-\Delta_{N,x} - \lambda^{(i)}(x)\right) V_x = C_x^{*(i-1)} \quad , \tag{3.28a}$$

$$C_x^{*(i)} = \frac{V_x}{\|V_x\|_x} \tag{3.28b}$$

for $i = 1, 2, 3, \ldots$. $\lambda^{(i)}(x)$ is an estimate of $\lambda_0(x)$ in the i-th iteration. The norm $\|\cdot\|_x$ is induced by the scalar product

$$< V_1, V_2 >_x = L_b^{-d} \sum_{x \in x} \frac{1}{N_c} \mathrm{Re}\,\mathrm{Tr}\,[V_1(x,z)^\dagger V_2(x,z)] \quad . \tag{3.29}$$

Recall that strictly speaking the vectors $C_x^{*(i-1)}$ and V_x are actually $(L_b^d \cdot N_c \times N_c)$ matrices. In our numerical implementation (Sec. 4.2), however, these matrices are represented by "true" (one-column) vectors.

Proper choices for $\lambda^{(1)}(x)$ and for $C_x^{*(0)}$, and the updating procedure $\lambda^{(i)}(x) \to \lambda^{(i+1)}(x)$ will be discussed below. Before that we discuss how the system (3.28a) is solved. The solution will be computed by means of the CG method (2.23) with the obvious replacements $D \to (-\Delta_{N,x} - \lambda^{(i)}(x))$, $f \to C_x^{*(i-1)}$, and $\phi \to V_x$. Also, the scalar product (2.21) is of course replaced by (3.29).

On a vector computer Eq. (3.28a) is solved for all blocks in parallel by vectorizing over the block index x. The solution will be obtained quite fast because CG solves Eq. (3.28a) in at most L_b^d iterations. In order that the operators $(-\Delta_{N,x} - \lambda^{(i)}(x))$ are positive definite, every $\lambda^{(i)}(x)$ must be less than the smallest eigenvalue $\lambda_0(x)$ of $-\Delta_{N,x}$. Since $-\Delta_{N,x}$ is positive semidefinite, we can for example choose $\lambda^{(i)}(x) = -\varepsilon$ with a small positive ε. But we can do better.

Estimation of eigenvalues by the Rayleigh quotient.

One starts inverse iteration with some lower bound $\lambda^{(1)}(x)$ for $\lambda_0(x)$ and gets a first approximation $C_x^{*(1)}$ of $C_x^* = C^*(\cdot, x)$. We wish to replace $\lambda^{(1)}(x)$ by an improved estimate $\lambda^{(2)}(x)$ which is closer to $\lambda_0(x)$, so that the components orthogonal to C_x^* become more strongly suppressed in going from $C_x^{*(1)}$ to $C_x^{*(2)}$ than they were in the step $C_x^{*(0)} \to C_x^{*(1)}$. Having solved Eq. (3.28a) for $i = 1$, we obtain an estimate for $\lambda_0(x)$ by computing the Rayleigh quotient of V_x which is defined through

$$R[V_x] = \frac{< V_x\,,\, -\Delta_{N,x} V_x >_x}{< V_x\,,\, V_x >_x} \quad . \tag{3.30}$$

In numerical analysis the Rayleigh quotient is the standard estimate for an eigenvalue from an eigenvector approximation [124]. R has the property that the eigenvalues of $-\Delta_{N,x}$ are its stationary points which are obtained when the corresponding eigenvectors

[16] The matrix need not be normal or normalizable.

[17] If A is defective one expands in terms of eigenvectors and principal vectors [82,94].

are plugged in. One has $R[V_x] = \lambda_0(x) + O(\delta^2)$, where $\delta = (\lambda_0(x) - \lambda^{(i)}(x))/(\lambda_1(x) - \lambda^{(i)}(x))$; $\lambda_1(x)$ is the second lowest eigenvalue of $-\Delta_{N,x}$.

Unfortunately, in order to have $(-\Delta_{N,x} - \lambda^{(i)}(x))$ positive to guarantee convergence of the CG algorithm, we must demand $\lambda^{(i)}(x) < \lambda_0(x)$. But $\lambda_0(x)$ is the absolute minimum of R, which means that always $R[V_x] \geq \lambda_0(x)$. In principle there is a simple way to circumvent this problem. Simply multiply both sides of Eq. (3.28a) by $(-\Delta_{N,x} - \lambda^{(i)}(x))$. But for this variant the CG method is substantially slower, because the computational labor is doubled in each iteration. Also, it is not a priori clear that a trial $\lambda^{(i+1)}(x) = R[V_x]$ is strictly smaller than $\frac{1}{2}(\lambda_0(x) + \lambda_1(x))$, which is a necessary condition for convergence of the inverse iteration towards C_x^*.

Bounding eigenvalues by Temple's Theorem.

A way out of the dilemma is furnished by a bounding theorem for eigenvalues which is known as Temple's Theorem.

TEMPLE'S THEOREM [125]:

Let T be a self-adjoint operator in a Hilbert space H with scalar product $< \cdot, \cdot >$. Let $f^{(0)}$ and $f^{(1)}$ be two elements of $H \setminus \{0\}$ with $Tf^{(1)} = f^{(0)}$. Compute the Schwarz constants

$$a_0 = < f^{(0)}, f^{(0)} > \quad , \quad a_1 = < f^{(0)}, f^{(1)} > \quad , \quad a_2 = < f^{(1)}, f^{(1)} > \quad , \tag{3.31}$$

and for a real number t the Temple quotient

$$\vartheta(t) = \frac{a_0 - t a_1}{a_1 - t a_2} \ . \tag{3.32}$$

If the interior of the interval (p, q) contains $\mu_2 \equiv a_1/a_2$ and exactly one isolated eigenvalue λ and no other element of the spectrum of T, then

$$\vartheta(q) \leq \lambda \leq \vartheta(p) \ . \tag{3.33}$$

See Ref. [125] for a proof.

A corollary for operators with a lowest eigenvalue.

A corollary of Temple's Theorem is the following. Suppose T has a lowest eigenvalue λ_0, which may also be degenerate, and one knows a lower bound l_1 for the second lowest eigenvalue λ_1 with $\mu_2 < l_1 \leq \lambda_1$. Then Temple's theorem is used with the interval $(-\infty, l_1)$ and one obtains

$$\mu_2 \frac{\mu_1 - l_1}{\mu_2 - l_1} = \mu_2 - \frac{\mu_1 - \mu_2}{\frac{l_1}{\mu_2} - 1} \leq \lambda_0 \leq \mu_2 \ , \tag{3.34}$$

where $\mu_1 \equiv a_0/a_1$.

Application of the corollary to (3.28a).

When we apply the corollary to the inverse iteration (3.28a) we obtain the bounds

$$\tau(l_1^{(i)}(x)) \equiv \lambda^{(i)}(x) + \mu_2^{(i)}(x) \frac{\mu_1^{(i)}(x) - l_1^{(i)}(x)}{\mu_2^{(i)}(x) - l_1^{(i)}(x)} \leq \lambda_0(x) \leq \lambda^{(i)}(x) + \mu_2^{(i)}(x) = R[V_x] \ , \tag{3.35}$$

Eigenvalue	0.	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	12.
Degeneracy	1	4	6	8	13	12	10	12	6	4	4	1

Table 3.1: *The spectrum of $-\Delta_{N,x}$ in $d = 4$ dimensions without gauge field for block size $L_b = 3$.*

where

$$\mu_1^{(i)}(x) = \frac{1}{< C_x^{*(i-1)}, V_x >_x} \quad , \quad \mu_2^{(i)}(x) = \frac{< C_x^{*(i-1)}, V_x >_x}{< V_x, V_x >_x} \tag{3.36}$$

$$l_1^{(i)}(x) = \tilde{\lambda}_1(x) - \lambda^{(i)}(x) \quad , \quad \tilde{\lambda}_1(x) \leq \lambda_1(x) \ .$$

Of course, the question arises how one obtains a lower bound $\tilde{\lambda}_1(x)$ for the second lowest eigenvalue $\lambda_1(x)$ of $-\Delta_{N,x}$. In Ref. [124] it is discussed how to overcome this difficulty in general eigenvalue problems either by comparison with a soluble problem exploiting Courant's maximum-minimum-principle [126] or by means of a Ritz method. For $-\Delta_{N,x}$ there is a simpler solution as will be shown below. Suppose for the moment that we know an appropriate $\tilde{\lambda}_1(x)$. Then we perform the updating of $\lambda^{(i)}(x)$ according to the following.

$$\begin{aligned} &(i) \quad \text{If } \mu_2^{(i)}(x) \geq l_1^{(i)}(x) \ , \quad \text{then} \quad \lambda^{(i+1)}(x) = \lambda^{(i)}(x) \ . \\ &(ii) \quad \text{Else compute } \tau(l_1^{(i)}(x)) \text{ and set } \lambda^{(i+1)}(x) = \max\left(\lambda^{(i)}(x), \tau(l_1^{(i)}(x))\right) \ . \end{aligned} \tag{3.37}$$

3.4.2 Relevance of the Eigenvalue Spectrum of $-\Delta_{N,x}$ for the Efficiency of the Algorithm

In order to make use of the updating prescription (3.37) we need a lower bound $\tilde{\lambda}_1(x)$ for the second lowest eigenvalue $\lambda_1(x)$ of $-\Delta_{N,x}$. How do we get it? Also, it was already mentioned that the performance of inverse iteration might be bad when there is a degeneracy in the eigenvalue spectrum. So let us look at the spectrum of low-lying eigenvalues of $-\Delta_{N,x}$.

Spectrum of eigenvalues of $-\Delta_{N,x}$ in pure gauges.

In case of pure gauges the spectrum of $-\Delta_{N,x}$ is known analytically. It is summarized in Table 3.1 for $L_b = 3$ in four dimensions. The spectrum is the same for any gauge group G, except for an (additional) N_c-fold degeneracy of each eigenvalue.

Spectrum of low-lying eigenvalues of $-\Delta_{N,x}$ in nontrivial gauge fields.

In nontrivial gauge field configurations the shape of the spectra will change, but the sum of all eigenvalues is invariant and equals always $N_c \cdot 2\,d\,L^{d-1}(L-1)$. (This is $\mathrm{Tr}\,[-\Delta_{N,x}]$ = twice the number of links within the block).

The distribution of the five lowest eigenvalues was determined numerically on a 3^4 lattice with periodic boundary conditions in four-dimensional pure $SU(2)$ gauge theory with Wilson action. The results are shown in Fig. 3.6. The eigenvalues $\lambda_2(x)$, $\lambda_3(x)$, and $\lambda_4(x)$ have no impact on the performance of our algorithm, but it is interesting to see how the degeneracy of $\lambda_1(x), \ldots, \lambda_4(x)$ is lifted in nontrivial gauge fields.

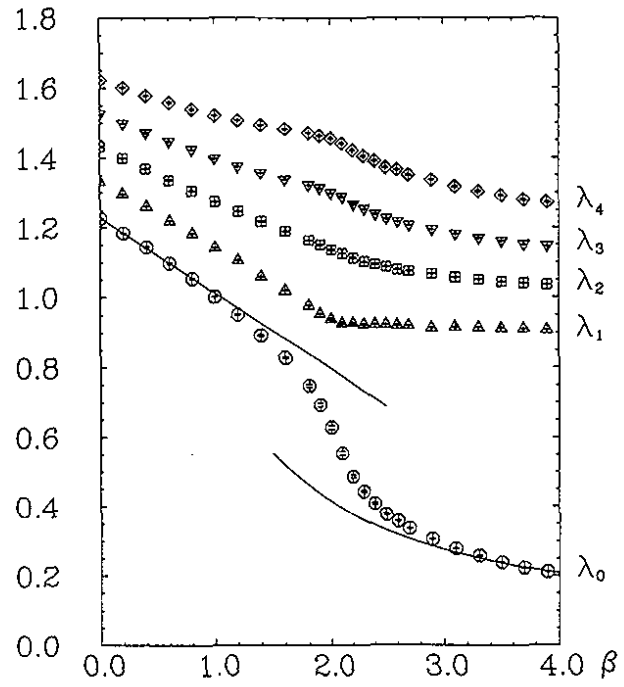


Fig. 3.6: *Expectation values of the five lowest eigenvalues (modulo two-fold degeneracy) of $-\Delta_{N,x}$ for gauge group $SU(2)$ in pure gauge theory on a 3^4 lattice as a function of β. In pure gauges ($\beta = \infty$), $\lambda_0(x) = 0$ and the other four eigenvalues are degenerate with value 1.0. The indicated strong and weak coupling expansion of $\langle\lambda_0(x)\rangle$ derived from the numerical data is $1.2284(7) - \beta \cdot 0.2163(15)$ and $\beta^{-1} \cdot 0.830(7)$, respectively.*

The Monte Carlo method implemented for updating the gauge field was Creutz's heat bath algorithm [127]. The eigenvalues were computed by means of existing libraries [128]. These computations were done on the computers of the University of Hamburg [129] and of DESY [130]. One thousand sweeps with measurements were performed after discarding 1000 sweeps to ensure thermalization.

It is interesting to note that $\lambda_0(x)$ behaves similar to the plaquette energy $\langle 1 - \frac{1}{2}\mathrm{Tr}\, U(P)\rangle$ which has strong coupling expansion $1 - \beta/4 + O(\beta^3)$ and weak coupling expansion $3/(4\beta)$ [131], and which exhibits a crossover from strong to weak coupling at β around 2.0. This crossover manifests itself also in the behavior of $\langle\lambda_0(x)\rangle$, and through a kind of blowing up of the band width of $\langle\lambda_1(x)\rangle, \ldots, \langle\lambda_4(x)\rangle$.

Implications for practical computations.

Fig. 3.6 is very promising, because we are mainly interested in the region of larger β-values where a distinct gap exists between $\lambda_0(x)$ and $\lambda_1(x)$. This permits to compute the

β	$\langle\lambda_0(x)\rangle$ on 3^4 lattice	$\langle\lambda_0(x)\rangle$ on 3^4 blocks in a 9^4 lattice
1.8	0.7469 ± 0.0034	0.7542 ± 0.0008
1.9	0.6924 ± 0.0044	0.7068 ± 0.0008
2.0	0.6284 ± 0.0062	0.6538 ± 0.0014
2.1	0.5537 ± 0.0055	0.5923 ± 0.0008
2.2	0.4868 ± 0.0037	0.5331 ± 0.0012
2.3	0.4435 ± 0.0025	0.4747 ± 0.0010
2.4	0.4082 ± 0.0025	0.4280 ± 0.0006
2.5	0.3780 ± 0.0016	0.3943 ± 0.0005
2.6	0.3592 ± 0.0016	0.3683 ± 0.0004
2.7	0.3372 ± 0.0016	0.3466 ± 0.0004

Table 3.2: *Finite size effect: Expectation values of the lowest eigenvalue $\lambda_0(x)$ of $-\Delta_{N,x}$ on a 3^4 lattice and on the 81 blocks of size 3^4 embedded in a 9^4 lattice.*

kernels with only two inverse iterations (at least for $\beta \geq 1.8$ where numerical investigations were performed, see below).

We note that the relatively large gap between $\lambda_0(x)$ and $\lambda_1(x)$ will also be present for other gauge groups at large β. This is so because the gap equals 1 for $\beta = \infty$, and $\langle\lambda_0(x)\rangle$ will always have a weak coupling expansion $const \cdot \beta^{-1}$, whereas $\langle\lambda_1(x)\rangle$ will stay close to 1.

One might speculate how the picture looks like when one implements the Diracian proposal for staggered fermions. In pure gauges the Diracian proposal reduces to the Laplacian choice, so that the gap between λ_0 and λ_1 will also equal 1 at $\beta = \infty$. The question is what happens at finite values of β.

Finite size effects.

When one determines $\langle\lambda_0(x)\rangle$ on the 81 blocks of a 9^4 lattice one observes some finite size effects as shown in Table 3.2. The values on the 9^4 lattice were determined by means of the Temple bounds.

3.4.3 Numerical Implementation of the Algorithm

The numerical implementation of the algorithm proposed above will now be explained.

Choice for lower bounds of eigenvalues.

The result for the low-lying eigenvalue spectrum of $-\Delta_{N,x}$ provides us with lower bounds $\lambda^{(1)}(x)$ and $\tilde{\lambda}_1(x)$. From the measured distribution of eigenvalues (not expectation values!) one can extract lower bounds for $\lambda_0(x)$ and $\lambda_1(x)$ which have a confidence level very close to 100%. These bounds are used for $\lambda^{(1)}(x)$ and $\tilde{\lambda}_1(x)$. In the course of the inverse iteration (3.28a), $\tilde{\lambda}_1(x)$ will not be changed and will retain its x-independent value[18], whereas $\lambda^{(i)}(x)$ is updated according to (3.37) and is only x-independent for $i = 1$.

[18] Note that $\tau(l_1^{(i)}(x))$ is a monotonically increasing function of $\tilde{\lambda}_1(x)$, but the dependence on $\tilde{\lambda}_1(x)$ is weak. It was checked that one can retain $\tau(l_1^{(i)}(x)) < \lambda_0(x)$, when $\tilde{\lambda}_1(x)$ is greater than a practically 100% c. l. upper bound for $\lambda_1(x)$. This means that $\tau(l_1^{(i)}(x))$ is a very stable lower bound for $\lambda_0(x)$.

Initialization of C.

It remains to fix a normalized $C_x^{*(0)}$. Two different choices were tested, a gauge variant and a gauge covariant choice. The first one is trivial initialization $(L_0 a)^d C^{*(0)}(z,x) = \mathbb{1}$ for all $z \in x$. The second choice is Bałaban's radial gauge [119], where one sums up with equal weights all parallel transporters along the taxi-driver paths (paths of shortest length) from $\hat{x}$ to z, for all $z \in x$; the result is brought back into the gauge group by polar decomposition [103].

By computing the Rayleigh quotient of $C_x^{*(0)}$ one finds that the component proportional to the sought eigenvector C_x^* is clearly dominating in Bałaban's gauge, because in this case $R[C_x^{*(0)}]$ lies between $\lambda_0(x)$ and $\lambda_1(x)$. For a trivially initialized $C_x^{*(0)}$ the components orthogonal to C_x^* dominate. A priori this does not imply anything for the rate of convergence of inverse iteration, but the finding is that it is more efficient to use Bałaban's gauge for $C_x^{*(0)}$.

Performance of the algorithm.

We conclude this section with some remarks on the performance of the algorithm described above.

The CG method was used with initialization of V_x to zero. In this case one needs less than 27 (29) CG iterations in the first (second) inverse iteration to solve Eq. (3.28a) sufficiently. The CG algorithm was stopped when the root mean square (RMS) residuals of all blocks were reduced by 10^{-4}. In general, a small residual does not mean that one is close to the solution, but the above criterion is good enough for the present case. This is demonstrated by explicitly computing $R[C_x^{*(i)}] = < C_x^{*(i)}, -\Delta_{N,x} C_x^{*(i)} >$ and comparing with $\lambda^{(i)}(x) + \mu_2^{(i)}(x)$. Both results agree.

The updating procedure (3.37) always yields an improved estimate of $\lambda_0(x)$ for the second inverse iteration if $\beta \gtrsim 2.15$. This is not the case if $C_x^{*(0)}$ is trivially initialized; either the condition $\mu_2^{(i)}(x) < l_1^{(i)}(x)$ is not fulfilled or if it is fulfilled $\tau(l_1^{(i)}(x))$ might not be larger than $\lambda^{(i)}(x)$.
Due to fluctuations, for $\beta \lesssim 2.15$ the lower bound $\tilde{\lambda}_1(x)$ of $\lambda_1(x)$ is not always greater than $\lambda_0(x)$, so that the necessary condition $\mu_2^{(i)}(x) \geq l_1^{(i)}(x)$ for getting an improved $\lambda^{(i+1)}(x)$ may be violated if $\lambda_0(x) \geq \tilde{\lambda}_1(x)$. Nevertheless, a second inverse iteration with a non-improved $\lambda^{(i)}(x)$ is also sufficient (at least down to $\beta = 1.8$).

After two inverse iterations the components orthogonal to C_x^* are practically eliminated. This is proved by inspecting the Temple bounds $\lambda^{(3)}(x)$ and $R[C_x^{*(2)}]$ for $\lambda_0(x)$, and by comparing $R[C_x^{*(2)}]$ with $|| - \Delta_{N,x} C_x^{*(2)}||$. The difference $R[C_x^{*(2)}] - \lambda^{(3)}(x)$ is typically less than 10^{-5} for larger values of β, while it becomes slightly greater for smaller β-values where $\lambda^{(1)}(x)$ or $\lambda^{(2)}(x)$ is not always improved. However, $R[C_x^{*(2)}]$ and $||-\Delta_{N,x} C_x^{*(2)}||$ practically coincide in the examined β-range from 1.8 to 2.7, which shows that $C_x^{*(2)} = C_x^*$ to a very high precision.

Finally we report the performance on a CRAY Y-MP [132]. The routines are completely vectorizable (vectorization over the block index). In case of a 9^4 lattice and block size $L_0^d = 3^4$ (vector length of 81) the CPU time needed to compute $C_x^{*(2)}$ is 5.7 *msec* per block. An average rate of 230 MFLOPS is achieved. The CPU time required to compute $C_x^{*(2)}$

for all blocks is comparable with that needed for one standard Monte Carlo sweep through the lattice. Thus, *the projective definition of block spins is numerically implementable for non-Abelian gauge theories.* It is not necessary to sacrifice the gauge covariance of the ground-state projection method by employing a particular gauge fixing condition (which also costs CPU time).

Concluding remarks.

In Ref. [105] the author reported some results of a Monte Carlo renormalization group (RG) study, using the averaging kernel (2.37) for gauge covariant blocking. He also considered a Laplacian with Dirichlet boundary conditions on block boundaries in [105]. The results are very promising. They indicate that the projective RG transformation might be fruitful when it is studied in the framework of color dielectric models [133], and also when it is used for RG investigations with block gauge field in the gauge group [134,135]. For details the reader is referred to Ref. [105].

3.5 Optimal Interpolation Kernels in the Presence of Gauge Fields

The gauge covariant generalization of the optimal Gawędzki-Kupiainen interpolation kernel $\mathcal{A}$ for bosons and for staggered fermions is studied in this subsection. The optimal kernel $\mathcal{A}$ decays exponentially for any *value of the gauge field coupling. Therefore* ground-state projection (with the Laplacian choice of the fermionic C) defines a good block spin in arbitrarily disordered gauge fields.

It was mentioned earlier that a good choice of block spin, i. e. of C, means that the associated optimal interpolation kernel $\mathcal{A}$ decays exponentially. We claimed already that ground-state projection (2.37), (3.16) is such a good choice in the presence of gauge fields.

Optimal $\mathcal{A}$ for free massless bosons and fermions.

When no gauge fields are present, the fall-off properties of $\mathcal{A}$ can be proven analytically. This is done by using the Fourier representation of $\mathcal{A}$, and shifting the path of the momentum integration into the complex plane [76].

In order to verify the decay property of $\mathcal{A}$ explicitly by giving numbers, $\mathcal{A}$ was computed as the solution of Eq. (2.32) for massless free bosons and staggered fermions. (Eq. (2.32) was solved by means of CG.) Since $\not{D}^2$ decomposes into 2^d independent Laplacians on the pseudoflavor sublattices when U is trivial, the optimal $\mathcal{A}$ fulfills $\mathcal{A}(z,x) = 0$ if z and x have different pseudoflavor. If z and x have the same pseudoflavor, then $\mathcal{A}(z,x)$ equals the bosonic $\mathcal{A}$-kernel on an N^d lattice if the staggered lattice has volume $(2N)^d$. This serves as a good check for consistency of the computer programs.

Bosonic $\mathcal{A}$-kernels for 9^4, 12^4, and 18^4 lattices (with periodic boundary conditions) are given in the first three tables of Appendix B.3. These tables contain values of $\mathcal{A}(z,x)$ for $U \equiv \mathbb{1}$. One sees an exponentially damped oscillating fall-off. In other trivial gauge fields (pure gauges), $\mathcal{A}(z,x)$ gets multiplied with the parallel transporter from $\hat{x}$ to z. Kernels on 18^4 lattices are visualized in Fig. 3.7.

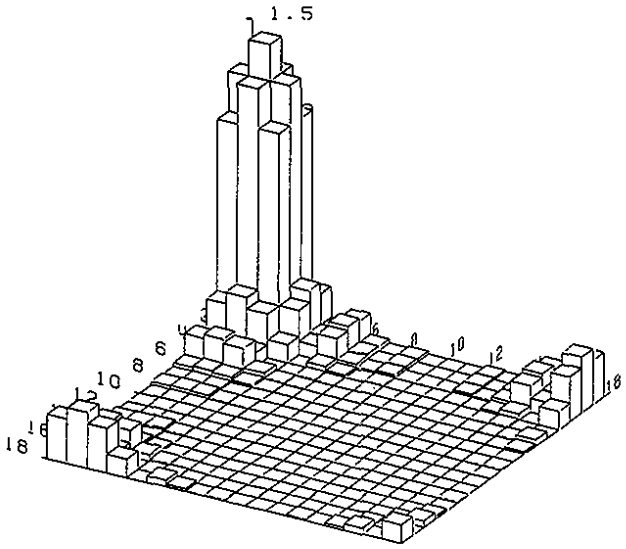
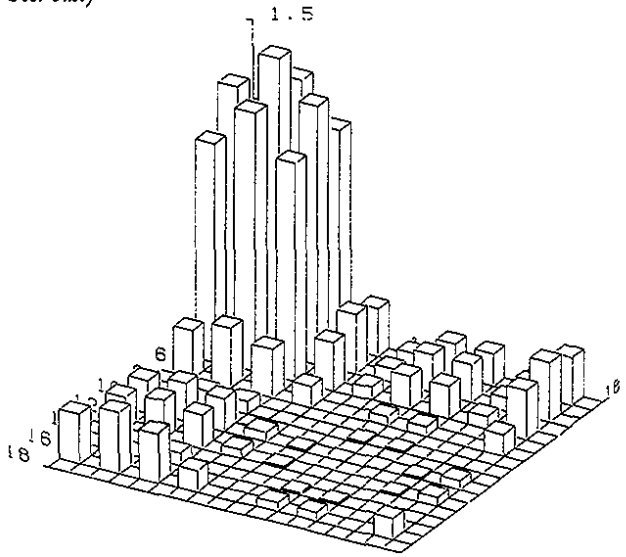


Fig. 3.7: *Optimal interpolation kernel $\mathcal{A}(z,0)$ in trivial gauge fields on 18^4 lattices ($L_b = 3$), for bosons (top) and staggered fermions (bottom). A two-dimensional cut through the block center $x = 0$ is shown, z_3 and z_4 are fixed. The vertical axis gives the norm of $\mathcal{A}(z,x)$. The fermionic kernel is the one which is associated with the Laplacian choice of C. The fall-off properties of this $\mathcal{A}$ are relevant for work on renormalization group transformations for fermi fields. Note the support properties of the fermionic $\mathcal{A}$. It has support only on one particular pseudoflavor sublattice, where it equals the bosonic $\mathcal{A}$ for a 9^4 lattice. (G. Mai [116] investigated a renormalization group transformation where she blocked free continuum fermions to staggered fermions; G. Mai showed analytically that the optimal $\mathcal{A}$ associated with her blocking procedure has exponential fall-off, cf. the remarks in Sec. 3.2.)*



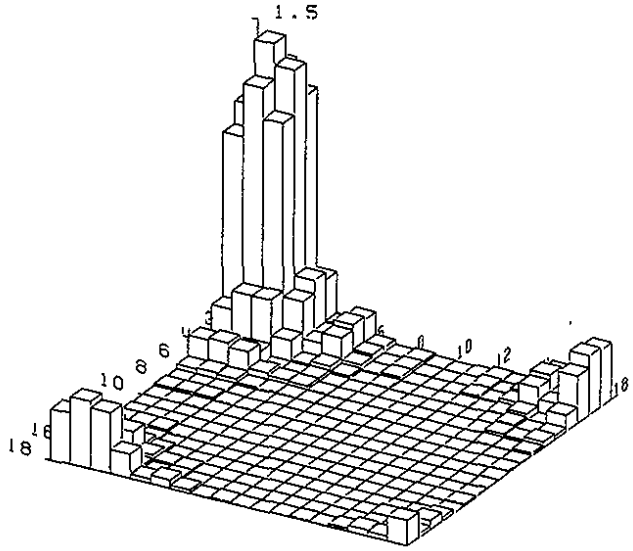
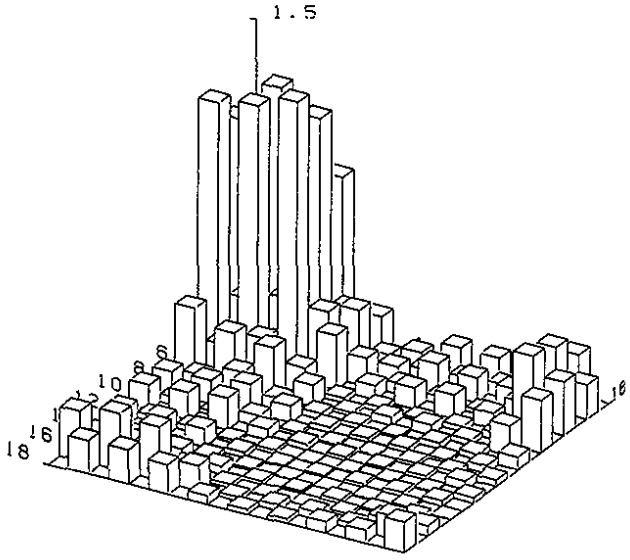


Fig. 3.8: *Optimal interpolation kernel $\mathcal{A}(z,0)$ in a quenched four-dimensional $SU(2)$ gauge field at $\beta = 2.7$ on 18^4 lattices ($L_b = 3$), for bosons (top) and staggered fermions (bottom). A two-dimensional cut through the block center $x = 0$ is shown, z_3 and z_4 are fixed. The vertical axis gives the trace norm of $\mathcal{A}(z,x)$. (Remember that $\mathcal{A}(z,x) \in \mathbb{R} \cdot SU(2)$.) The fermionic kernel is the one which is associated with the Laplacian choice of C. Note the support properties of this $\mathcal{A}$. It has support on all even lattice sites and is vanishing on all odd sites. The increased support compared to the case $\beta = \infty$ is due to the non-vanishing field strength at finite β.*



Optimal $\mathcal{A}$ for bosons and fermions in nontrivial gauge fields.

Next the $\mathcal{A}$-kernels were computed in nontrivial four-dimensional quenched $SU(2)$ gauge fields at various values of $\beta = 4/g^2$, including the case of a completely random gauge field ($\beta = 0$). The operator D in Eq. (2.32) was $D = -\Delta + m^2$ with $m^2 = m_{cr}^2$ in case of bosons,[19] while for fermions $D = -\not{D}^2 + m^2$ with $m^2 = 0$. It was verified that for *any* β-value the definition of the averaging kernel C as a solution of a gauge covariant Laplace eigenvalue equation (2.37), (3.16) yields a good choice of block spin in the sense described above. This is true both for bosons and for staggered fermions. Examples are in Fig. 3.8. $\mathcal{A}$ decays exponentially over distance $L_b a$ (nearly) as fast in the presence of gauge fields than in their absence. For this reason the more intricate Diracian choice of C *has not been implemented numerically yet.*

The effective coarse grid operators which are associated with the optimal $\mathcal{A}$-kernels decay also exponentially. Therefore they are to a good approximation also of (next) nearest neighbor form. Examples are given in Tables B.4–B.6, B.9, and B.10 of Appendix B.3.

3.6 Short Review of Bałaban's Analytical Work

Bałaban's analytical work on the decay properties of gauge covariant propagators is briefly outlined.

There exists analytical work on exponential decay of propagators in gauge fields. We repeat here an outline which was given in Ref. [112].

Bałaban presented a method to study fall-off properties of gauge covariant propagators [104], see also [120]. He considered gauge covariant bosonic "fluctuation field" propagators

$$\Gamma = \left(-\Delta + m^2 + \kappa C^* C\right)^{-1} , \tag{3.38}$$

cf. Eq. (B.10) in Appendix B. Bałaban showed exponential decay of Γ under general assumptions on C, and for sufficiently small gauge field coupling. He required C^*C to be a projector on constant functions in the absence of gauge fields; this requirement is fulfilled by our choice of C. It is instructive to see what difficulties arise when one tries to generalize Bałaban's proof to staggered fermions. Δ gets replaced by the square of the Dirac operator.

In the bosonic case Bałaban proceeds by writing down a generalized random walk expansion for Γ. Bałaban considers a covering of $\mathbb{R}^d$ by hypercubes $\square_j$ of side length $2L$, $L \in \mathbb{Z}$ which are centered at Lj, $j \in \mathbb{Z}^d$. Two hypercubes $\square_i$ and $\square_j$ have the property that they overlap iff $\|i-j\|_\infty \leq 1$. Every site $z \in \mathbb{Z}^d$ is contained in exactly 2^d hypercubes $\square_j$. Bałaban also introduces functions $h_j(z)$ whose supports are within $\square_j$ and whose squares are a partition of unity: $\sum_j h_j^2 = 1$. The leading term of Bałaban's generalized random walk expansion for Γ is

$$\sum_j h_j \Gamma_j^0 h_j \quad \text{where} \quad \Gamma_j^0 = \left(-\Delta_{N,j} + m^2 + \kappa C^* C\right)^{-1} , \tag{3.39}$$

[19] Values of m_{cr}^2 will be given in Secs. 3.7 and 5.

and $\Delta_{N,j}$ is the Laplacian with Neumann boundary conditions on boundary of the hypercube $\square_j$. It is defined as in Eq. (2.40) except that x has to be replaced by $\square_j$. The full expansion is of the form

$$\Gamma = \sum_j h_j \Gamma_j^0 h_j + \sum_{j_0, j_1, j_2} h_{j_0} \Gamma_{j_0}^0 R_{j_1, j_2} \Gamma_{j_2}^0 h_{j_2} + \sum_{j_0, \ldots, j_4} h_{j_0} \Gamma_{j_0}^0 R_{j_1, j_2} \Gamma_{j_2}^0 h_{j_2} R_{j_3, j_4} \Gamma_{j_4}^0 h_{j_4} + \cdots \tag{3.40}$$

where $R_{i,j}$ are operators which satisfy certain bounds on their norms, and whose kernels $R_{i,j}(z,w)$ are only nonvanishing when z, w are in neighboring hypercubes $\square_i$ and $\square_j$. [The operator $R_{i,j}$ contains h_i^2 and h_j.] Only random walks $\{j_0, j_1, \ldots, j_{2n}\}$ contribute to the expansion (3.40) where $\|j_{2i} - j_{2i+1}\|_\infty \leq 1$, $i = 0, 1, \ldots$

The Neumann boundary conditions lead to a decoupling such that

$$\Gamma_j^0(z,w) = 0 \quad \text{unless } z, w \text{ are in the same hypercube } \square_j. \tag{3.41}$$

It follows that the contributions to the series (3.40) up to n-th order make contributions to $\Gamma(z', w')$ only if z' and w' are no further than $n-1$ hypercubes away from each other, i. e. if $|z' - w'| < n\,L\,a$. Therefore the propagator Γ will have exponential decay if the Neumann series (3.40) converges.

Given the bounds on the norms of $R_{i,j}$, convergence will hold if the lowest eigenvalues of

$$-\Delta_{N,j} + \kappa C^* C$$

stays sufficiently far away from 0. $-\Delta_{N,j}$ has at most one eigenvalue close to zero, and if C^*C is the projector on this eigenmode and κ is not too small, then the lowest eigenvalue of $-\Delta_{N,j} + \kappa C^* C$ is equal to the second lowest eigenvalue of $-\Delta_{N,j}$ (which is at least of order $(\pi/L)^2$).

In conclusion, there are two essential ingredients. Neumann boundary conditions decouple between different hypercubes, and the lowest eigenvalues of $-\Delta_{N,j} + \kappa C^* C$ are not too close to zero.

In case of staggered fermions we want to replace the Laplacian Δ by the square of the Dirac operator $\not{D}^2$ on $\Lambda_{a/2}$. The trouble is that blocks x will overlap. This is innocuous in case of a pure gauge. In this case $\not{D}^2$ does not couple sites with different pseudoflavors H, and block spins $(C\chi)(x)$ receive contributions only from $\chi(z)$ when z carries the same pseudoflavor as x. *As a result, one can proceed as in the bosonic case.*

But this is no longer true when we have a nontrivial gauge field. One way out could be to retain $-\Delta_{N,j} + \kappa C^* C$ as 0-th order term, where $\Delta_{N,j}$ is still the Laplacian. But this means that the $F_{\mu\nu}$ term in $\not{D}^2$ has to be absorbed into $R_{i,j}$. If we use the Laplacian definition (3.16) of C, then the lowest eigenvalue of $-\Delta_{N,j} + \kappa C^* C$ is as big as we can hope to make it, and decoupling holds also, but a good enough bound on $R_{i,j}$ will require that $F_{\mu\nu}$ is small enough.

The results of Sec. 3.5 give a numerical proof that gauge covariant propagators decay exponentially for arbitrary gauge coupling, both in the bosonic and in the fermionic case. Hence the effect of the $F_{\mu\nu}$-term is reasonably mild. Thus the numerical results can be considered as an encouragement for analytical calculations in fermionic gauge theories.

3.7 Smoothness in Gauge Theories

The question is discussed what "smoothness" means in bosonic and fermionic gauge theories. The discussion supplies another argument why the generalized Gawędzki-Kupiainen interpolation kernel $\mathcal{A}$ is "optimal".

A central notion for MG methods is that of smoothness. Therefore it is crucial to answer the question: *What is a smooth function in the presence of gauge fields, or more generally in disordered systems?* This question is also discussed in Refs. [112,136,137,138].

We require to have a gauge invariant definition of smoothness, i. e. if a field configuration is smooth in a particular gauge it is also smooth in any other gauge. The standard definition of smoothness on length scale a of some function e would be that its derivative is small,

$$\|\nabla_\mu e\|^2 \equiv < \nabla_\mu e, \nabla_\mu e > = < e, -\Delta e > \ll < e, e > \equiv \|e\|^2 \qquad (3.42)$$

(in units $a = 1$).

Bosonic theories.

In bosonic theories (3.42) is equivalent to saying that the Rayleigh quotient $R[e]$ is small. When ∇_μ are ordinary derivatives then the definition (3.42) of smoothness is not gauge invariant. We could take covariant derivatives instead. In the bosonic case this gives indeed a natural gauge invariant notion of smoothness.

The lowest eigenvalue of the covariant negative Laplacian is a natural measure of disorder. Except for a pure gauge, the lowest eigenvalue of $-\Delta$ is strictly positive. Thus, in a disordered gauge field there is a limit to how smooth a function can be. When disorder is big, then there are no smooth functions at length scale a in the sense of (3.42). But in this case one does not really need them either in realistic problems, because propagators $(-\Delta + m^2)^{-1}(z, w)$ for $m^2 \geq 0$ fall off fast anyway and there will be no critical slowing down when one applies standard methods, and hence there will be no need for MG.

Some values of m_{cr}^2 (the lowest eigenvalue of $-\Delta$) in four-dimensional $SU(2)$ gauge fields are plotted in Fig. 3.9. These values were determined by means of inverse iteration.

Theories with staggered fermions.

In the case of staggered fermions the situation is different. There are now two seemingly reasonable definitions. We may take for Δ in (3.42) the covariant Laplacian as defined in Eq. (3.19), or substitute the square of the Dirac operator for it. We call these two possibilities the definitions of "Laplacian smoothness" and "Diracian smoothness", respectively.

Unfortunately, both definitions are qualitatively different, except in the case of a pure gauge. Laplacian smoothness of a function $\chi(z)$ implies only restrictions which relate the values of $\chi(z)$ at points $z \in \Lambda_{a/2}^H$ with the *same* pseudoflavor H. The same is *not* true of Diracian smoothness. Moreover, if the gauge field is fairly disordered, then there will be no functions with Laplacian smoothness. But it is *not* true in general that a propagator $(-\not{D}^2 + m^2)^{-1}(z, w)$ cannot have slow decay for $m \approx 0$.

Smoothness of the optimal interpolation kernel $\mathcal{A}$.

The foregoing discussion of smoothness in gauge theories supplies another argument

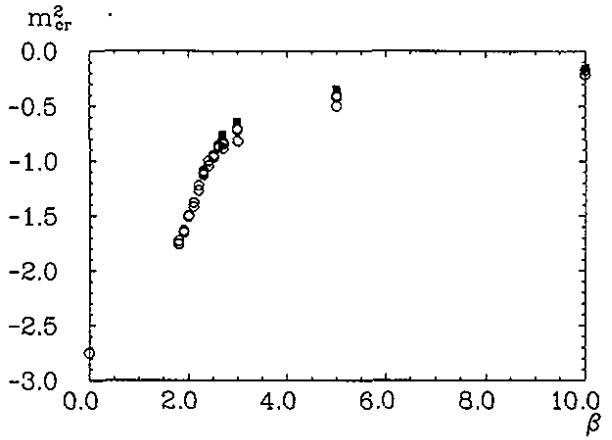


Fig. 3.9: *Critical bosonic masses squared m_{cr}^2 of $SU(2)$ gauge field configurations in $d = 4$ dimensions. Open circles represent values on 9^4 lattices, while filled symbols are values on 18^4 lattices. In a pure gauge ($\beta = \infty$), $m_{cr}^2 = 0$ for any lattice size.*

for using the optimal $\mathcal{A}$-kernel in MG computations. According to the MG philosophy $\mathcal{A}$ should map functions on Λ^1 into smooth functions on Λ^0. The optimal $\mathcal{A}$-kernel is characterized by requiring that for every function Φ on Λ^1, $\phi = \mathcal{A}\Phi$ minimizes the action $\mathcal{H} = < \phi, D_0 \phi >$ subject to the constraint $C\phi = \Phi$. With the definition of smoothness that (covariant) derivatives are small, this can be rephrased: the interpolation $\phi = \mathcal{A}\Phi$ yields the smoothest possible function with prescribed block average $\Phi = C\phi$.

When one follows the spirit of ground-state projection MG, then one should use the Diracian definition of smoothness for fermions, i. e. one should substitute $\not{D}^2$ for Δ in (3.42). In case of staggered fermions the gauge invariant norm $\|\not{D} e\| \equiv < e, -\not{D}^2 e >^{1/2}$ of a low mode of $\not{D}$ is always small compared with $\|e\|$, but the same is not true of $< e, -\Delta e >^{1/2}$.

4 Computation of Free Propagators

Now we turn to the numerical computation of propagators. We begin in this section with the simple free case (this means arbitrary pure gauge) where the bosonic and fermionic problems are equivalent. Computation of propagators in nontrivial gauge fields will be reported in the subsequent three sections.

Secs. 4.1 and 4.2 apply to all propagator computations.

Summary of results.

In trivial gauge fields the following results emerged: The convergence properties of the CG algorithm exhibit a substantial volume dependence for a given mass m^2. No such volume dependence is found for relaxation algorithms. The theoretical predictions for conventional one-grid relaxation algorithms could be verified to a very high precision. In particular, it was confirmed that the convergence behavior *depends only on* m^2 and *not on the lattice size*. Variational MG with "ground-state projection" $\equiv$ "piecewise-constant" interpolation, and the "optimal" MG scheme are both successful in eliminating the $1/\triangle m^2$ divergence of relaxation times τ on lattices of fixed size. No dramatic difference between variational MG and optimal MG is seen. It takes only a few iterations until errors decay exponentially. The variational MG method was investigated for various lattice sizes, and *no volume dependence was found. Hence, CSD is completely eliminated by MG for pure gauges.* It was found that a necessary condition for the elimination of CSD is an "almost" correct interpolation of the lowest mode (i. e. of a constant function). The inclusion of the "updating on a last site", as discussed in Sec. 2.8, cannot improve the performance of MG algorithms any further for pure gauges. The result of its inclusion in traditional one-grid relaxation algorithms is interesting. CSD of *asymptotic* relaxation times can be eliminated in this way. Surprisingly this is only the case when the matrix Ω in (2.50) is chosen such that the energy functional gets minimized. No effect is seen when Ω is chosen such that the norm of the residual is minimized. However, a volume effect remains with respect to how many iterations it takes until the asymptotic regime is reached. This volume effect is similar to the one which shows up in the CG algorithm.

4.1 Computational Details

All computations which will be reported in the sequel were performed on four-dimensional hypercubic lattices. The r.h.s. f of Eq. (2.1) was always a Kronecker-δ. (Other choices yield the same values for relaxation times.) The initial guess for the propagator is zero in case of CG, and $\delta_{x,0}\,\omega/(8+m^2)$ for the relaxation algorithms. The term "trivial gauge field" always means arbitrary pure gauge, and "nontrivial gauge fields" are always quenched $SU(2)$ gauge fields at finite gauge coupling. The gauge field configurations at finite β were generated with the Wilson action [1] by means of Creutz's heat bath algorithm [127].

As remarked at the end of Sec. 2.3, we take advantage of the facts that we work in four dimensions and that $L_b = 3$ is chosen, so that twogrid algorithms are sufficient to offer a big improvement if they are successful, because the second twogrid layer has 81 times fewer points.

Residual equations on the block lattice are always solved exactly (i. e. reducing the initial residual by 10^{-10}) by the CG algorithm. Hence, there is no relaxation parameter for the block lattice, and the value quoted for ω is always the relaxation parameter on the finer lattice.

Relaxation times τ in computations of propagators and number of iterations are always measured in units of cycles which involve only one sweep through the finest lattice. It was checked that values of τ practically do not depend on the numbers ν_1 and ν_2 of pre-smoothing and post-smoothing sweeps when τ's are properly rescaled. Relaxation times were determined as follows. The RMS (root mean square) residuals $\|r^{(n)}\| \equiv \left(|\Lambda^0|^{-1}\sum_{z\in\Lambda^0}\frac{1}{2}\mathrm{Tr}\left[r^{(n)}(z)^\dagger r^{(n)}(z)\right]\right)^{1/2}$ were computed, and τ was determined from an exponential fit after $\|r^{(n+1)}\|/\|r^{(n)}\|$ has become constant for several iterations with an accuracy of better than 10^{-4}.

If not stated otherwise, SOR was always done in lexicographic ordering, except for bosons on 12^4 and 18^4 lattices where updating was performed in checkerboard fashion.

Concerning the aspect of computational labor we note that the (non-optimized) implementations of the MG programs require a factor of 2.1/4.5 (bosons/staggered fermions) more arithmetic operations than CG. When updating on an MG layer consisting of a single site (Secs. 2.8,7.5) is included, the factor is 3.2/7.8. The amount of work for the averaging kernel C on the whole lattice is equivalent to less than 20 CG iterations.

Finally, we note that *all computer programs were checked for gauge covariance.*

4.2 Numerical Implementations

The numerical implementation used for storing $SU(2)$ gauge fields and performing multiplications of (multiples of) $SU(2)$ matrices is based on the well-known isomorphism between $SU(2)$ and the unit sphere S^3 in four dimensions. Let $\mathbb{1}$ denote the 2×2 unit matrix, and σ_k, $k = 1,2,3$, the three Pauli matrices. Any $SU(2)$ matrix U can be represented as

$$U = a_0\mathbb{1} + i\sum_{k=1}^{3} a_k\sigma_k \ , \qquad (4.1)$$

with real numbers a_0, a_k which fulfill $a_0^2 + \sum_k a_k^2 = 1$. There exists no constraint on a_0, a_k for matrices U in span$\{SU(2)\}$.

On the computer link variables, kernels, propagators etc. are stored by the four numbers a_0, a_k. Multiplication of two elements of span$\{SU(2)\}$ is implemented through statement functions. In the 4-vector representation the norm $\frac{1}{2}\mathrm{Tr}\,[U^\dagger U]^{1/2}$ of a multiple U of an $SU(2)$ matrix is simply given by $\sqrt{a_0^2 + \sum_k a_k^2}$.

4.3 Conjugate Gradient Algorithm

The convergence behavior of the (unpreconditioned) CG algorithm on an 18^4 lattice is shown in Fig. 4.1 below, where it is compared with variational MG ($\mathcal{A} = C^*$). The CG algorithm has a complicated convergence behavior. Note that the norm of the residual is not monotonically decreased for small values of m^2. The reason is that CG decreases the energy functional in every iteration, but this does not necessarily mean that $\|r^{(n)}\|$ also decreases.

It might be that the convergence behavior for small m^2 is affected by finite size effects, because the correlation length of the matter field (inverse mass) is not small compared to the linear extension of the lattice. However, this question could not be answered.

It should also be noted that CG exhibits a volume effect. At fixed m^2 CG needs more iterations to converge to a given precision the larger the lattice becomes. This is true for all values of m^2, not only for artificially small ones.

4.4 One-Grid Relaxation Algorithms

The theoretical predictions for conventional relaxation algorithms which were discussed in Sec. 2.2 could be confirmed to a very high precision. In particular, it was confirmed that the convergence behavior depends only on m^2 and *not* on the lattice size. For instance, in case of Jacobi iteration Eq. (2.19) predicts $\tau = 80\,000.5 + O(m^2)$ for $m^2 = 10^{-4}$, whilst the numerical result is $\tau = 80\,000.50 \pm 0.26$. Also the ω-dependence in Eq. (2.19) could be monitored very well. The predictions for SOR were also verified.

4.5 Variational Multigrid

An advantage of variational MG with $\mathcal{A} = C^*$ ("piecewise-constant interpolation") is that the effective Laplacian is of the same nearest neighbor form as Δ on the finest lattice. This simple MG algorithm is able to eliminate CSD completely, i. e. no mass dependence is observed for $m^2 \leq 10^{-2}$, and no volume dependence is observed at fixed m^2. Relaxation times τ are short. Hence, CG is clearly outperformed. Results are in Fig. 4.1 and Table 4.1.

Table 4.1 contains a parameter δ which is defined as follows. We wish to approximate the ideal MG algorithm as well as possible. Without knowing the exact kernel $\mathcal{A}$, one

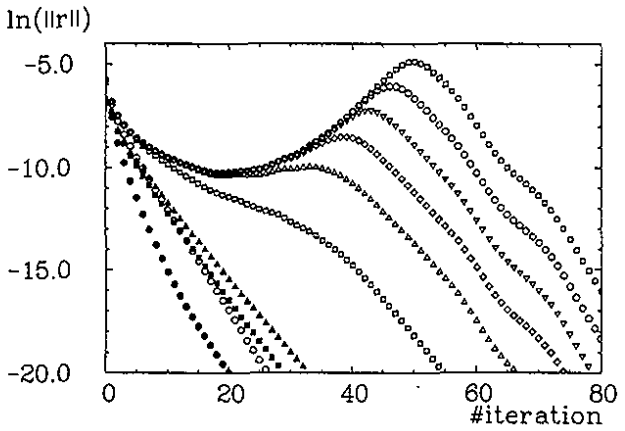


Fig. 4.1: *Convergence of computations of propagators $(-\Delta + m^2)^{-1}$ in trivial gauge fields on an 18^4 non-staggered lattice. The three (seven) curves of filled (open) symbols are results of variational MG (CG) computations. From left to right the curves belong to $m^2 = 1$, $m^2 = 0.1$ and $m^2 \leq 10^{-2}$ in case of MG, and $m^2 = 1, 0.01, 0.01, \ldots, 10^{-6}$ in case of CG. No volume dependence was found for MG, whereas at fixed m^2 CG needs more iterations to converge to a given precision the larger the lattice becomes.*

algorithm and lattice size	optimal δ	optimal ω	relaxation time τ for $m^2 =$ 1.0	0.1	0.01	$10^{-3}, \ldots, 10^{-10}$
MG Jacobi on 9^4	1.8	0.96	3.6	6.4	7.1	7.2
MG Jacobi on 18^4	1.8	0.94	4.0	6.4	6.9	7.0
MG SOR on 18^4	1.7	1.09	2.3	2.5	2.8	2.9

Table 4.1: *Variational MG: Optimum values of δ and ω in a four-dimensional pure gauge ($m_{cr}^2 = 0$) and demonstration of elimination of CSD.*

cannot compute the exact "effective difference operator" D_1. Knowing that D_1 is approximately of nearest neighbor form (see Appendix B), one can try to approximate it. The choice

$$D_1 = \delta^{-1} C D_0 C^* \tag{4.2}$$

yields another dielectric lattice gauge field $D_1(x,y)$ which also has the right gauge covariance properties. Its fluctuating length can be adjusted to fluctuate around the "right" value (of the ideal $D_1(x,y)$) by adjusting the real parameter δ. In trivial gauge fields the influence of δ on τ is very weak, but a tuning of δ in nontrivial gauge fields can accelerate convergence substantially. The same is true of the relaxation parameter ω. Dependences on δ and ω in the two-grid algorithm in trivial gauge fields are given in Table 4.2.

We anticipate (cf. Sec. 4.8) that the parameter δ need not be introduced when the

Dependence on δ								
δ	0.0	0.45	0.9	1.35	1.8	1.9	2.0	2.1
τ(Jacobi)	800.5	15.7	15.7	15.7	15.6	15.6	45.7	div.
τ(Gauss-Seidel)	400.2	6.2	4.4	3.8	4.3	8.5	46.6	div.

Dependence on ω for $\delta = 1.8$ with Jacobi relaxation			
ω	τ	ω	τ
1.0	15.6	0.9	8.0
0.99	11.8	0.8	9.1
0.98	9.4	0.7	10.5
0.97	7.7	0.6	12.4
0.96	7.4	0.5	14.9
0.95	7.6	0.4	18.5
0.94	7.6	0.3	24.1
0.93	7.7	0.2	33.8
0.92	7.8	0.1	54.8
0.91	7.9		

Dependence on ω for $\delta = 1.7$ with checkerboard SOR			
ω	τ	ω	τ
1.9	9.4	0.9	4.5
1.8	4.5	0.8	5.5
1.7	3.4	0.7	6.8
1.6	3.3	0.6	8.4
1.5	4.6	0.5	10.6
1.4	4.5	0.4	13.7
1.3	4.3	0.3	18.7
1.2	4.2	0.2	27.6
1.1	3.0	0.1	43.9
1.0	3.6		

Table 4.2: *Variational MG: Dependences on δ and ω in four-dimensional pure gauges for $m^2 = 0.01$. For $\delta = 0$ one has the traditional one-grid algorithms. "div." means that the algorithm diverges.*

improved correction scheme (2.51) is used. There the $N_c \times N_c$ matrix Θ takes the role of the real number δ. The scheme (2.51) has the advantage that Θ need not be chosen by hand, rather it is determined by the algorithm itself through the requirement of minimalization of the energy functional (2.20).

4.6 Optimal Multigrid

When no gauge fields are present, the lowest mode of $-\Delta$ is a constant. In this case the simple variational MG scheme is successful, because then C^* interpolates the lowest mode of $-\Delta$ correctly. The same is true for the optimal $\mathcal{A}$, since it fulfills $(L_b a)^4 \sum_{y \in \Lambda^1} \mathcal{A}(z,y) = 1$ for all $z \in \Lambda^0$ (Eq. (B.16) in Appendix B). In the trivial case $U = \mathbb{1}$ the interpolation of a constant function on Λ^1 with this $\mathcal{A}$ is unnecessarily complex, because there are nonvanishing contributions from all $y \in \Lambda^1$. However, it will be shown in the following sections that indeed more complicated choices of D_1 *and* of the interpolation map $\mathcal{A}$ are necessary to eliminate CSD in nontrivial gauge fields.

Numerical results with the optimal MG scheme are as follows. Compared with variational MG, the convergence rate in trivial gauge fields is not improved in case of Jacobi relaxation, but convergence speed is doubled in case of SOR. Results with lexicographic SOR are contained in Table 5.2 of the next section.

4.7 Importance of (Almost) Correct Interpolation of the Lowest Mode

We have just argued that variational MG with piecewise-constant interpolation works only because it interpolates the lowest mode of $-\Delta$ correctly. The relevance of the lowest eigenvector was also pointed out by Edwards [71]. From the marriage of MG with neural net methods [139] one obtains a similar conclusion [137,138,140,141]. It follows from neural net optimization criteria that the interpolation kernel $\mathcal{A}$ should have the property that the lowest eigenmode ϕ_0 of $D_0 = D$ should admit a representation of the form $\phi_0 = \mathcal{A}\Phi$ with a suitable Φ to a good accuracy.

The (adjoint of the) kernel (2.35) fulfills this requirement. The alternative proposal of [105], to define $\mathcal{A}$ through the eigenvalue equation (2.37) with $\Delta_{N,x}$ (2.40) replaced by $\Delta_{D,x}$ (2.41) does not meet this requirement. The lowest mode of the negative Laplacian with Dirichlet boundary conditions on boundaries of blocks x $-\Delta_{D,x}$ is a (lattice) sine wave, namely *const* $\prod_\mu \sin\left(\frac{\pi z_\mu/a}{L_b+1}\right)$, $z_\mu = 1, \ldots, L_b$. Only for block size $L_b = 2$ is this a constant on blocks, as in case of Neumann boundary conditions.

Numerical results confirm the theoretical considerations. If one uses the lowest mode of $-\Delta_{D,x}$ for interpolation, CSD in the computation of free propagators is not eliminated. (Remember that we chose $L_b = 3$.) The critical exponent z retains its value of 2.

In Sec. 6.5 numerical computations will be reported which show that CSD in computations of free propagators can be eliminated, even if the lowest mode of $-\Delta$ is *not* interpolated exactly. Therefore the above mentioned requirement that $\phi_0 = \mathcal{A}\Phi$ with a suitable Φ "to a good accuracy" does not mean "exactly".

4.8 Updating on a Multigrid Layer Consisting of a Single Site and Related Improvements

MG algorithms.

By including the rescaling (2.50) *in MG algorithms the performances are not improved* any further, neither in variational MG nor in optimal MG. The improved correction scheme (2.51), however, is able to halve τ and to bring it close to 1. Note that the parameter δ of (4.2) is eliminated when (2.51) is used.

Conventional one-grid relaxation algorithms.

By including (2.50) in conventional one-grid relaxation algorithms, CSD is eliminated in the sense that *asymptotic* relaxation times do not depend on m^2 (and on the lattice size). This is shown in Fig. 4.2. *The figure looks the same for bosons in nontrivial gauge fields when m^2 is replaced by Δm^2.*

The norm of the residual is not monotonically decreased for small m^2. This feature also shows up in CG which minimizes the energy functional K (i. e. the residual r in the norm induced by the scalar product $< \cdot, D^{-1} \cdot >$) rather than $||r||$ itself (cf. Fig. 4.1). This is an interesting point: Instead of (2.50), one could think of rescaling $\phi^{(n)}$ by another

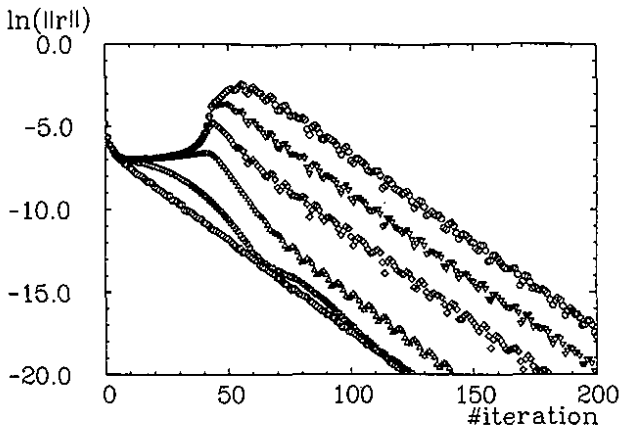


Fig. 4.2: *One-grid SOR plus (2.50) eliminates CSD of* asymptotic *relaxation times, shown here for a pure gauge. (ω = 1.90, checkerboard updating) The six curves correspond to $m^2 = 0.1, 0.01, \ldots, 10^{-6}$ on a 12^4 (non-staggered) lattice, with m decreasing from left to right. A small volume effect remains, not for τ (i. e. the asymptotic decay rate) but with respect to how long it takes until the asymptotic regime is reached: e. g. on an 18^4 lattice the number of iterations needed to obtain a given accuracy of $\|r\|$ for $m^2 = 10^{-6}$ is increased by ~ 20. The figure looks the same for bosons in nontrivial gauge fields when m^2 is replaced by $\triangle m^2$.*

matrix Ω' which is chosen such that not the energy functional but the norm of the residual $||r^{(n)}||$ is minimized. However, in this case there is no difference to one-grid relaxation. *This remains true in nontrivial gauge fields.*

When one increases the lattice size there is a small volume effect with respect to how long it takes until errors decay exponentially; see caption of Fig. 4.2. However, the asymptotic slope of the curves is independent of $|\Lambda|$. Therefore it seems that CSD of *asymptotic* relaxation times is eliminated.

Inefficiency of (2.50) as soon as errors decay exponentially.

Finally we note that practically $\Omega = \mathbb{1}$ as soon as $||r||$ decays exponentially. Then the step (2.50) could be switched off. *This statement holds also in nontrivial gauge fields and for MG algorithms.*

5 Proof that the Multigrid Method Works in Disordered Systems

This section contains the central result that the MG method works in disordered systems.

Summary of results.

MG computations of propagators in nontrivial gauge fields do not suffer from CSD, when the optimal interpolation kernel $\mathcal{A}$ and the optimal effective coarse grid operator are used. This statement is true for *any* value of the gauge coupling, including the case of completely random gauge fields ($\beta = 0$). These computations were reported for the first time in Ref. [136]. Unfortunately, the optimal MG scheme is not practical for production runs because of computational complexity and storage space requirements.

5.1 Unpracticality of the Optimal $\mathcal{A}$ for Production Runs

The optimal interpolation kernel $\mathcal{A}(z,y)$ decays exponentially, but it is nonvanishing for all arguments (z,y); cf. Figs. 3.7 and 3.8, and the tables in Appendix B. When no gauge fields are present $\mathcal{A}$ is translational invariant, but this is no longer true in nontrivial gauge fields. Therefore storage of this kernel in a computer requires $N_c^2 \cdot |\Lambda^0| \cdot |\Lambda^1|$ numbers.[20] Needless to say that because of this and computational complexity the use of the ideal MG algorithm is not practical for production runs. However, it is very important to answer questions of principle. Therefore the optimal MG scheme was implemented, and tested for bosonic propagators on 9^4 lattices.

5.2 Computation of Optimal Interpolation Kernels $\mathcal{A}$

The MG method is of interest near criticality, i. e. for slowly decaying propagators. For nontrivial gauge fields we have to enforce slow decay by choosing m^2 negative and very close to the negative of the lowest eigenvalue $-m_{cr}^2 > 0$ of $-\Delta$. Values of m_{cr}^2 were determined by inverse iteration to an accuracy of 10^{-7} or better (this can be seen from the data in Fig. 6.4

[20] In case of staggered fermions, one can halve this number because the even and odd sublattices are not coupled both for the Laplacian and for the Diracian choice of the averaging kernel C. Nevertheless, the storage space requirement is still much too large.

β	m_{cr}^2	optimal ω	τ for $\Delta m^2 \leq 10^{-3}$
∞	0	1.27	1.6
2.7	−0.8210607	1.38	1.9
2.5	−0.9477085	1.40	1.9
2.2	−1.2218471	1.45	1.9
1.8	−1.7567164	1.57	2.5
0	−2.7480401	1.69	5.2

Table 5.1: *Results of the idealized MG algorithm with lexicographic SOR for bosonic propagators in six different gauge field configurations on 9^4 lattices Λ^0.*

of Sec. 6). For each investigated gauge field configuration one $\mathcal{A}$-kernel was computed as the solution of the equation

$$\left([-\Delta + m_{cr}^2 + \kappa C^* C]\mathcal{A}\right)(z,y) = \kappa C^*(z,y) \tag{5.1}$$

for large κ; cf. Eqs. (2.32), (B.11). This equation was solved by means of CG. In practice κ-independence is reached for $\kappa \gtrsim O(10^4)$.

Actually, one should compute and use a new $\mathcal{A}$-kernel for each different value of m^2. i. e. one should have m^2 in place of m_{cr}^2 in Eq. (5.1). But it turned out that it is sufficient to use the "critical" $\mathcal{A}$-kernel (5.1) for any $m^2 > m_{cr}^2$. This $\mathcal{A}$ was computed for $SU(2)$ gauge fields on 9^4 lattices at $\beta = \infty$, 2.7, 2.5, 2.2, 1.8, and 0.

5.3 Numerical Results of the Idealized Multigrid Algorithm

Propagators $(-\Delta + m^2)^{-1}$ were computed for values of m^2 close to criticality, $m^2 = m_{cr}^2 + \Delta m^2$, $\Delta m^2 = 1, 0.1, 0.01, \ldots, 10^{-6}$.

It turned out that *CSD is completely eliminated at any value of β* (i. e. also in completely disordered gauge fields) when $D_1 = CD_0\mathcal{A} \equiv \mathcal{A}^* D_0 \mathcal{A}$ *and* e_1 is interpolated by $\mathcal{A}$. Results are in Table 5.1 and Fig. 5.1.

In order to be sure about the correct determination of m_{cr}^2, it was checked that simple one-grid relaxation algorithms exhibited CSD with $z = 2$, i. e. relaxation times diverged proportional to $1/\Delta m^2$. It was also checked that the coarse grid operator $CD_0\mathcal{A}$ with the optimal $\mathcal{A}$ is indeed Hermitean and equals $\mathcal{A}^* D_0 \mathcal{A}$. Numerical inaccuracies (round-off errors) are less than 10^{-9}.

The dependence of relaxation times on ω is shown in Table 5.2. This dependence is weaker than in variational MG, cf. Sec. 6.3.

In order to be precise, it must be pointed out that the results reported here prove only that the idealized MG algorithm is able to eliminate the $1/\Delta m^2$ divergence of relaxation times on lattices of a fixed size. However, we believe that no volume dependence would emerge if one were able to implement the optimal MG scheme on larger lattices. This belief is based on the fact that with the idealized MG algorithm the convergence behavior in nontrivial gauge fields is nearly the same as convergence in pure gauges. Exponential

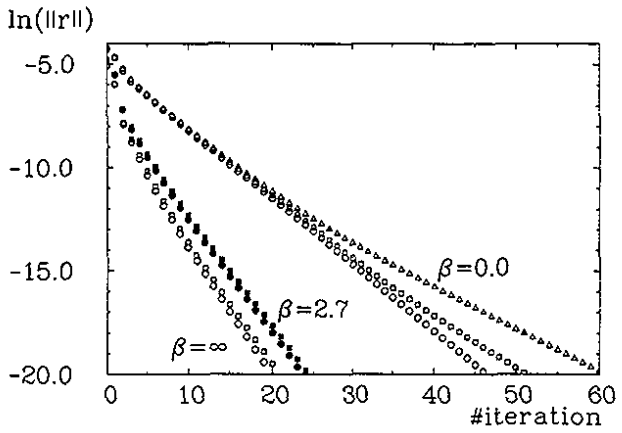


Fig. 5.1: *Convergence of the optimal MG algorithm for bosons on 9^4 lattices in quenched $SU(2)$ gauge fields at $\beta = \infty$, 2.7, and 0.0. The curves shown for each β-value correspond from left to right to $\Delta m^2 = 0.1$, $\Delta m^2 \leq 0.01$ for $\beta = \infty$ and $\beta = 2.7$, and to $\Delta m^2 = 0.1$, $\Delta m^2 = 0.01$, $\Delta m^2 \leq 0.001$ for $\beta = 0.0$. The relaxation parameter ω has the optimal value quoted in Table 5.1.*

ω	1.0	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9
$\tau(\beta = \infty)$	2.8	2.4	2.0	1.6	1.9	2.4	3.1	4.2	6.2	11.9
$\tau(\beta = 2.7)$	4.6	3.8	3.0	2.4	2.0	2.4	3.0	3.9	5.5	10.4
$\tau(\beta = 2.5)$	4.6	3.7	3.0	2.3	1.8	2.2	2.8	3.8	5.5	10.4
$\tau(\beta = 2.2)$	5.4	4.4	3.5	2.8	2.1	2.1	2.6	3.5	5.1	10.0
$\tau(\beta = 1.8)$	8.6	7.2	6.0	4.8	3.8	2.9	2.4	3.2	5.0	10.2
$\tau(\beta = 0.0)$	19.9	16.4	13.4	10.9	8.7	6.7	5.0	4.1	7.7	div.

Table 5.2: *Dependence of relaxation times τ on ω in the ideal MG algorithm with lexicographic SOR for computing bosonic propagators on 9^4 lattices with $\Delta m^2 = 0.01$. The algorithm does not converge for $\omega = 2.0$, and for $\omega = 1.9$ at $\beta = 0$.*

decay of errors sets in after only a few iterations. In pure gauges no volume dependence of convergence was found. Therefore it is reasonable to expect the same result for the idealized MG method.

5.4 Implications of the Numerical Results

The result that CSD is completely eliminated in the idealized MG algorithm proves three things:

(i) *The MG method can cope in principle with the frustration (disorder) which is inherent in non-Abelian gauge fields.*

(ii) *There exists an appropriate notion of smoothness in disordered cases as described in Sec. 3.7.* The ideal $\mathcal{A}$ is the smoothest kernel which obeys $C\mathcal{A} = \mathbb{1}$.

(iii) *Ground-state projection is a good choice of C (i. e. of a smooth block spin of the low frequency components) in gauge theories.* This statement is now confirmed directly from MG computations, not only from the point of view of the renormalization group as argued in Sec. 3.5.

The success of the idealized MG algorithm is especially remarkable in view of the fact that variational MG is not able to eliminate CSD in nontrivial gauge fields. This will be shown in Secs. 6.3 and 7.3.

5.5 Approximations to the Idealized Multigrid Algorithm

It was found that the convergence speed of MG is very sensitive to details of the interpolation. Typically CSD is not eliminated in nontrivial gauge fields when not all ingredients are optimal. Choosing either[21] $D_1 = \delta^{-1} C D_0 \mathcal{A}$ but using C^* instead of $\mathcal{A}$ for interpolation, or approximating $D_1 = \delta^{-1} C D_0 C^*$ but retaining $\mathcal{A}$ for interpolation, both fail to eliminate CSD at finite β. This is unlike the case $\beta = \infty$, where also these two modifications of the idealized algorithm succeed in fighting CSD.

For the question of approximations to the ideal $\mathcal{A}$ with a finite support (truncating the exponential tails) we refer to Sec. 6.5.

We remark that in our computations where we used only one $\mathcal{A}$-kernel for all values of m^2, the correctness of the value of m_{cr}^2 in Eq. (5.1) has influence on the performance. For instance, if one replaces m_{cr}^2 in (5.1) by $m_{cr}^2 \pm 0.01$, relaxation times stay constant and short for $\Delta m^2 \geq 10^{-3}$, but the algorithm diverges for $\Delta m^2 \leq 10^{-4}$. (In these investigations we used the coarse grid operator $\mathcal{A}^* D_0 \mathcal{A}$). Recall, however, the remark about the mass term after Eq. (5.1). A priori one can only hope for elimination of CSD when $\mathcal{A}$ is computed with the correct m^2. Yet, if one were able to employ the idealized MG algorithm for production runs one would have a fixed m^2, and therefore one would solve Eq. (5.1) with this m^2. As a result, there is no problem with the choice of the mass in $\mathcal{A}$ and CSD would always be eliminated in practice.

5.6 Idealized Multigrid Algorithm for Staggered Fermions

The idealized MG algorithm has not been implemented yet for staggered fermions. There the computational complexity and the storage space requirement are even larger than for bosons. We mentioned in footnote 20) that one can halve the required storage space in case of fermions compared with bosons. On the other hand, effectively fermionic lattices need to be larger by a factor of 16. Nevertheless, in order to answer questions of principle also for fermions, idealized MG computations are planned for the near future. We have little doubt that they will be as successful as the bosonic computations reported here.

[21] Cf. Eq. (4.2), no damping parameter δ is required in the optimal MG scheme.

6 Computation of Gauge Covariant Bosonic Propagators in Four-Dimensional SU(2) Gauge Fields

We turn now to practical algorithms for bosonic propagators. First we look at the CG algorithm, and at one-grid relaxation methods. The MG methods discussed thereafter are: (i) variational MG, (ii) algorithms with approximations to the ideal $\mathcal{A}$-kernel, (iii) inclusion of updating on an MG layer consisting of a single site, and (iv) the related improvements of Sec. 2.9. There is also a subsection with a remark concerning the use of SOR in MG computations.

Summary of results.

The CG algorithm is very sensitive to the disorder in the gauge field. For a given value of $\triangle m^2$, CG needs more iterations to converge to a given precision the more disordered the gauge field becomes.

The theoretical considerations of Sec. 2.2 predict that this is different in conventional one-grid relaxation algorithms. This prediction could be confirmed numerically. In one-grid relaxation methods convergence is almost the same, no matter how strong the disorder of the gauge field is, as long as equal values of $\triangle m^2$ are compared. There is also no volume dependence in practice, whereas at fixed β and fixed $\triangle m^2$ CG needs more iterations to converge the larger the lattice size is.

The variational MG algorithm is sensitive to the disorder of the gauge field, but not as strongly as the CG algorithm. Convergence of variational MG does not only depend on $\triangle m^2$, but also on β and on the lattice size. In contrast to the case of pure gauges, variational MG is not able to eliminate CSD in nontrivial gauge fields. However, the variational MG method considered here is able to outperform the CG algorithm in CPU time on lattices $\gtrsim 18^4$.

An MG algorithm with an "approximated optimal" interpolation kernel $\mathcal{A}$ was tested. This $\mathcal{A}$-kernel does not fulfill the requirement that it interpolates the lowest mode of the Laplacian correctly. Despite this, CSD is completely eliminated in trivial gauge fields. Hence, for the success of an MG algorithm it is not a necessary condition that the lowest mode is interpolated exactly.

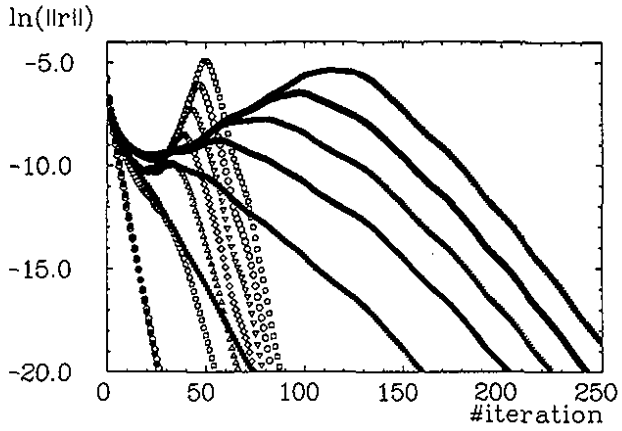


Fig. 6.1: *Convergence of CG for bosonic propagators $(-\Delta + m_{cr}^2 + \triangle m^2)^{-1}$ on an 18^4 lattice. Curves of open symbols are results of computations in a pure gauge field ($m_{cr}^2 = 0$), while curves of solid symbols are obtained in the background of a quenched SU(2) gauge field at $\beta = 2.7$. ($m_{cr}^2 = -0.7554339$.) Within each of the two families the curves are parameterized by $\triangle m^2 = 1, 10^{-1}, \ldots, 10^{-6}$, with $\triangle m^2$ decreasing from left to right The initial guess for the propagator is zero in all cases.*

The result might be taken as an indication that an overlap of blocks is more important. However, in nontrivial gauge fields CSD is not eliminated with the approximated $\mathcal{A}$-kernel.

The inclusion of "updating on a last site" as discussed in Sec. 2.8 leads to a method which eliminates CSD of *asymptotic* relaxation times. What remains, however, is a small volume effect with respect to how long it takes until errors decay exponentially.

The implementation of the further related improvements of Sec. 2.9 does not pay.

6.1 The Competitor: Conjugate Gradient Algorithm

The algorithm which is to be beaten is CG (2.23). Its behavior in case of bosonic propagators is shown in Fig. 6.1 where we compare convergence at $\beta = \infty$ with $\beta = 2.7$. No preconditioning is employed.

Recall what was said in Sec. 4.3 about the convergence for small values of m^2, or $\triangle m^2$ respectively. The norm of residuals is not necessarily decreased monotonically. The volume effect which was mentioned in Sec. 4.3 is larger the smaller β becomes. For fixed β and fixed $\triangle m^2$, CG needs longer to converge the larger the lattice size is.

Also, the CG algorithm is very sensitive to the disorder in the gauge field. For a

given value of Δm^2, CG needs more iterations to converge to a given precision the more disordered the gauge field becomes. It is worth to emphasize that this is not the case in conventional one-grid relaxation algorithms. There convergence is practically the same, no matter how strong the disorder of the gauge field is, as long as equal values of Δm^2 are compared.

6.2 One-Grid Relaxation Algorithms

Values of τ for relaxations on a single grid (Sec. 2.2) agree extremely well with our numerical results, also in nontrivial gauge fields at various values of β.

In particular, it must be emphasized again that one-grid relaxation algorithms are nearly insensitive to the disorder of the gauge field. Their convergence behavior depends only on the value of Δm^2. There is also practically no dependence on the lattice size. Cf. the discussion in Sec. 2.2.

As an example for the numerical verification of another theoretical prediction we quote the following. Remember the prediction for SOR (page 15): $\omega = 1.91$ is optimal for $\Delta m^2 \sim 0.009$. Hence, $\tau = -1.0/\ln(0.91) = 10.6$ for $\Delta m^2 \geq 0.009$, and one has CSD with $z = 2$ for smaller values of Δm^2. This behavior shows up in a distinct way in the data presented in Fig. 6.4 below.

6.3 Variational Multigrid

Existing literature.

The most practical MG method is variational coarsening (sometimes called "Galerkin definition"), because in this method coarse grid operators retain the property that at most nearest neighbors are coupled. For this reason variational MG is by far the most popular MG method. It is used exclusively in the literature [72,74,100,73,107,111,142,117,143,108, 144,109,145,101,110,102,118,146].

Ground-state projection MG computations for propagators were done before in two-dimensional $U(1)$ [74,73,107] and (approximate ground-state projections with gauge fixing) in two-dimensional and four-dimensional $U(1)$, and two-dimensional $SU(2)$ gauge fields [111,142,108,144,109,110]. The first ground-state projection MG computation of gauge covariant propagators in four-dimensional non-Abelian gauge fields (for gauge group $SU(2)$) was presented by the present author in Ref. [136]. Besides ground-state projection there are two other approaches for choosing the averaging kernel C. One alternative approach is called "parallel transported MG". It is based on the representation (2.34) and was studied in Refs. [72,100,117,102,118] by Ben-Av et al. The second alternative approach is pursued by Vyas [143,146]. He determines C by means of a Migdal-Kadanoff renormalization group transformation.

Gauge covariant procedure of the present work.

In all computations reported in this thesis exact ground-state projecting averaging kernels C were used exclusively. The advantage of this procedure is that no gauge fixing

Dependence on δ							
δ	0.45	0.9	1.35	1.8	2.25	2.7	3.15
$\tau(\beta = 2.7, \omega = 1.0)$	83.8	62.1	51.0	45.0	40.5	37.7	div.
Dependence on ω							
ω	1.0	1.1	1.2	1.3	1.4	1.5	1.6
$\tau(\beta = 2.7, \delta = 2.9)$	37.0	29.0	22.3	16.5	11.3	6.0	div.

Table 6.1: *Dependence of relaxation times τ on δ and on ω in variational MG SOR for computing bosonic propagators on 18^4 lattices with $\Delta m^2 = 0.01$. "div." means that the algorithm does not converge.*

is necessary, because the method is gauge covariant. The algorithm for computing C on the whole lattice (Sec. 3.4) is so efficient that it merely costs CPU time of the order of one heat bath sweep for the gauge field. [Cf. Sec. 3.4. It might well be that the eigenvalue equation is solved well enough with fewer CG iterations; however, this question has not been investigated.] Thus, it is not too expensive to solve (2.37), and it is not necessary to sacrifice gauge covariance by working with gauge field independent kernels in gauge-fixed U-configurations. (Gauge fixing costs also CPU time.)

Numerical work was done on 9^4, 12^4, and 18^4 lattices Λ with $SU(2)$ gauge fields at various values of β. The computational details mentioned in Sec. 4.1 apply here as well. The definition of the Galerkin operator contains the factor δ which was introduced in (4.2).

Numerical results.

Optimum values of ω and δ are: (i) $\beta = 10.0$: $\delta \sim 2.3$, $\omega \sim 0.99, 0.96, 1.25$ for MG Jacobi on a 9^4, on an 18^4, and for MG SOR on an 18^4 lattice, respectively. (ii) $\beta = 2.7$: $\delta \sim 2.8$, $\omega \sim 0.99, 0.99, 1.51$ for algorithms as above. These values were determined at $\Delta m^2 = 0.01$ and retained for all values of m^2.

For pure gauges, performance of the MG schemes depends little on δ (cf. Table 4.2), but for disordered gauge fields the relaxation time decreased with increasing δ up to a maximum value where the MG algorithms start to diverge suddenly. Also, in nontrivial gauge fields relaxation times in variational MG are stronger dependent on ω than in case of free propagators. Examples are in Table 6.1; see also the remarks in Sec. 6.4 below.

In Figs. 6.2 and 6.3 we compare the approach to convergence of three one-grid algorithms, namely damped Jacobi (Jac), SOR and CG, and two versions of twogrid cycles, MG Jacobi and MG SOR. Values of ω and δ in MG algorithms are the optimal ones as given above, and $\omega = 0.99$ in one-grid Jacobi, $\omega = 1.91$ in one-grid SOR. In the literature, convergence is considered achieved when the natural logarithm of the RMS residual has gone down by 10 [147] (dashed line in Figs. 6.2 and 6.3). It is seen that MG SOR outperforms CG, not only with respect to number of iterations, but also with respect to CPU time.

Small values of Δm^2 may seem to be of no practical importance on 18^4 lattices because the correlation length determined by the propagator becomes comparable to or larger than the lattice extension. But it is nevertheless important to study very small Δm^2 in order

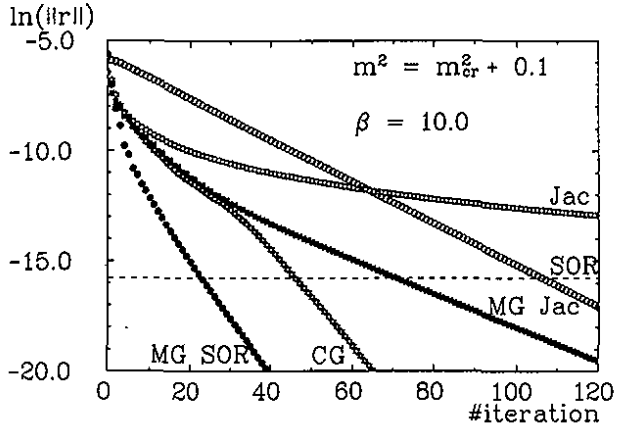
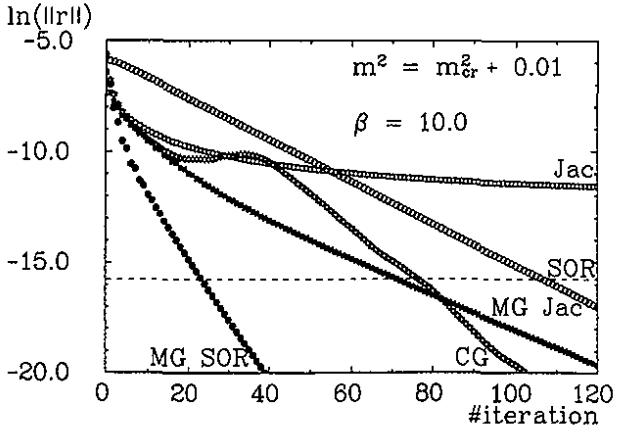


Fig. 6.2: *Computation of bosonic propagators* $(-\Delta + m_{cr}^2 + \Delta m^2)^{-1}$*: Convergence on an* 18^4 *lattice in a quenched* $SU(2)$ *gauge field at* $\beta = 10.0$*, with* $m_{cr}^2 = -0.1533739$*. (MG with* $A = C^*$*.)*



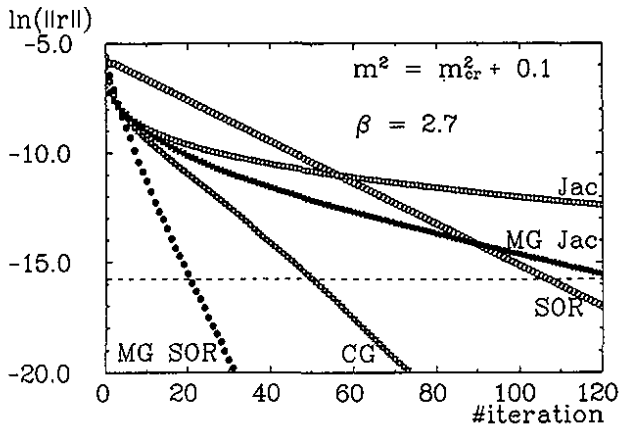
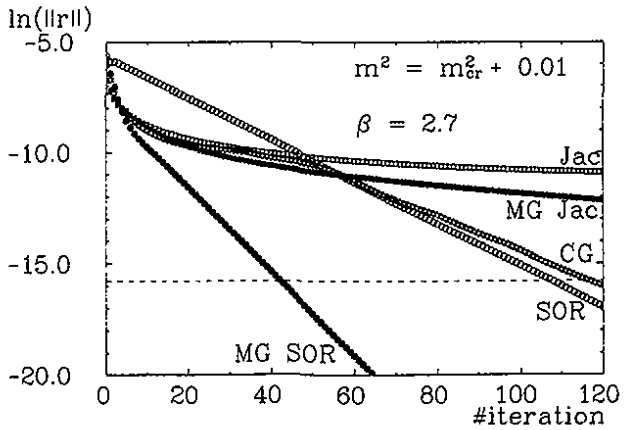


Fig. 6.3: *Computation of bosonic propagators* $(-\Delta + m_{cr}^2 + \Delta m^2)^{-1}$*: Convergence on an* 18^4 *lattice in a quenched* $SU(2)$ *gauge field at* $\beta = 2.7$*, with* $m_{cr}^2 = -0.7554339$*. (MG with* $A = C^*$*.)*



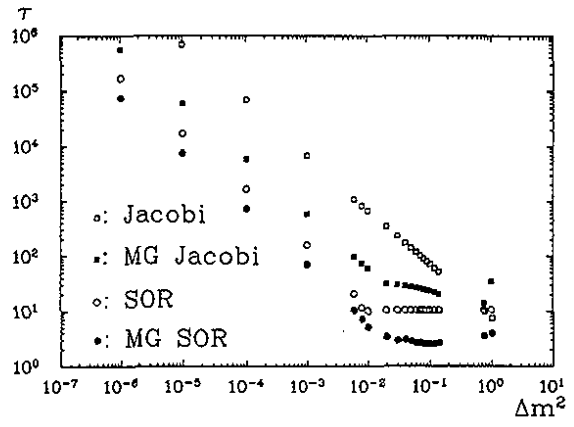


Fig. 6.4: *Computation of bosonic propagators* $(-\Delta + m_{cr}^2 + \Delta m^2)^{-1}$*: Relaxation times* τ *(in comparable work units) of iterative algorithms on an* 18^4 *lattice in a quenched* $SU(2)$ *gauge field at* $\beta = 2.7$ *with* $m_{cr}^2 = -0.7554339$*. (MG with* $A = C^*$*).*

to understand the operation and limits of the MG algorithms. We recall the remarks of Sec. 2.10. If one had a $1/\Delta m^2$ divergence of relaxation times on lattices of fixed size with a constant of proportionality which is independent of $|\Lambda|$, one would be able to definitely predict non-reduction of CSD under conditions where $1/\Delta m$ and the lattice size are changed proportionally.

In one-grid relaxation algorithms it was found that convergence at fixed Δm^2 does practically not depend on β and on the lattice size $|\Lambda|$. Whether this continues to hold for MG relaxation algorithms was investigated numerically. The result is that at fixed β convergence of MG is not independent of the lattice volume. However, no common trend was found how the convergence speed in number of iterations behaves at fixed values of Δm^2 and β when $|\Lambda|$ is increased.

In contrast to relaxation on one grid, it was also found numerically that the variational MG method is sensitive to the disorder of the gauge field. Convergence speed depends also on β, see below.

In Fig. 6.4 the relaxation times for various algorithms are shown as a function of Δm^2, at $\beta = 2.7$. One sees that CSD is not eliminated in the MG modes. This is unlike the pure gauge case, Sec. 4.5. The relaxation time behaves like $\tau \propto (\Delta m^2)^{-z/2}$ with $z = 2$ both for MG Jacobi and MG SOR. One could think of adjusting the mass term in D_1, but it turned out that this leads to no more than a $\sim 5\%$ acceleration of convergence.

At $\beta = 10.0$ the same qualitative behavior was found as in Fig. 6.4. Computations at other values of β down to 2.2 were also done. Surprisingly it was found that the value of Δm^2 at which CSD sets in (i. e. τ starts to increase) depends little on β. This contradicts

the scaling hypothesis of [144]. Basically the only dependence on β was found for the constant of proportionality in the scaling relation of τ. Compared with $\beta = 2.7$, its value is nearly doubled for $\beta = 2.2$, and roughly reduced by a factor of 3 at $\beta = 10.0$.

Conclusions.

A tentative interpretation of the permanence of CSD might be as follows. m_{cr}^2 is a natural gauge invariant measure of smoothness as discussed in Sec. 3.7. For $\beta \leq 10$ the lowest eigenvalue $-m_{cr}^2$ of $-\Delta$ appears to be too large to consider any function as smooth (cf. Fig. 3.9). Therefore the basic hypothesis of the MG method, namely smoothness of the slow modes, is not fulfilled, and therefore one cannot expect MG to work in an optimal fashion. There would be no problem if one had restricted attention to propagators $(-\Delta + m^2)^{-1}$ with $m^2 \geq 0$. This propagator is critical only when $-\Delta$ has a very low eigenvalue, and then smooth functions exist. One might believe that basically one is paying the prize for having introduced criticality artificially into a frustrated system. However, the investigations with the optimal interpolation kernel $\mathcal{A}$ which were reported in Sec. 5 showed that this is not the case. It seems that the variational MG method considered here is simply a too crude scheme in nontrivial gauge fields.

It would be interesting to investigate the question whether the situation improves when one uses overlapping blocks. Hasenbusch and Meyer found in their MG Monte Carlo simulations of two-dimensional asymptotically free models [44] that it is essential for the elimination/reduction of CSD to use overlapping blocks. However, in gauge theories one has to take into account that the storage space requirements and computer time grow when blocks overlap. For this reason, a systematic investigation of the simpler schemes without overlap was done and is reported here.

6.4 A Remark Concerning the Use of SOR in Multigrid Computations

We remark that the use of a relaxation parameter $\omega \neq 1$ in MG computations contradicts the conventional wisdom. According to this wisdom the only job of the relaxation procedure on Λ^0 is to smoothen the error, and this job is perfectly done by Gauss-Seidel or plain Jacobi iteration.

This conventional wisdom is confirmed by numerical results in trivial gauge fields. However, the finding of the present thesis is that the picture changes in nontrivial gauge fields. In case of Jacobi relaxation one does in fact gain nothing (or very little) when ω is chosen different from 1 or 0.99. But in case of SOR any over-relaxation in nontrivial gauge fields yields better performance than Gauss-Seidel, and much better performance than Jacobi or the minimal residual algorithm. A similar finding was reported in works on MG gauge fixing [148,149].

We refer to Tables 4.2, 5.2, and 6.1 for numerical results. Note that the idealized MG algorithm follows the conventional MG wisdom closer than variational MG.

n	≥ 4	3	2	1	0
# blocks where $\mathcal{A}(z,x) \neq 0$	81	65	33	9	1
$\max_{z \in \Lambda^0} \lvert 1 - \int_x \mathcal{A}(z,x) \rvert$	$3.3 \cdot 10^{-10}$	$4.3 \cdot 10^{-3}$	$2.1 \cdot 10^{-2}$	$6.2 \cdot 10^{-2}$	0.55

Table 6.2: *Violation of Eq. (B.16) on a 9^4 lattice in the absence of gauge fields, when the optimal $\mathcal{A}$-kernel is truncated at distance n, i. e. $\mathcal{A}(z,y) = 0$ is set by hand if $a^{-1}|z - \hat{y}| > n$. The finite value for $n \geq 4$ is due to round-off errors.*

6.5 Approximations to the Optimal Interpolation Kernel

The conclusion which can be drawn from the results of Secs. 5 and 6.3 is that one should find a compromise between the ideal MG algorithm and variational coarsening. One may try to approximate the optimal $\mathcal{A}$-kernel by a simpler choice with a finite support, i. e. one wants a kernel which fulfills $\mathcal{A}(z,x) = 0$ for $z \notin$ a certain neighborhood of block x. This neighborhood might be taken to be only the block x itself; cf. the practicality condition (2.29). The coarse grid operator would be defined through $D_1 = \mathcal{A}^* D_0 \mathcal{A}$.

Naive attempt.

A first very naive guess is to take simply the optimal $\mathcal{A}$ and to truncate it, i. e. setting $\mathcal{A}(z,y) = 0$ by hand for $a^{-1}|z - \hat{y}| > n$ where n is some small integer number. (The "taxi-driver" norm $|\cdot| = \|\cdot\|_1$ is used.) However, this procedure fails. The explanation for the failure appears simple. We argued that in the absence of gauge fields the validity of Eq. (B.16) is a necessary condition for a successful MG algorithm for propagators. But this equation is violated when the optimal $\mathcal{A}$ is truncated, see Table 6.2. Also, by rude truncation the kernel loses its smoothness.

The "Ohrenkern".

As an alternative one may think of the following proposal which was made by Mack [34]. We call this proposal "Ohrenkern".[22] It is defined as follows. Restrict the support of $\mathcal{A}(z,x)$ to block x and its eight nearest neighbors, i. e. $\mathcal{A}(z,x) = 0$ unless $z \in x$ or $z \in$ a nearest neighbor block of x. Compute an approximated interpolation kernel $\mathcal{A}$ as solution of Eq. (5.1) with Δ replaced by the covariant Laplacian with Dirichlet boundary conditions on the boundary of the desired support of $\mathcal{A}$.

Intuitively, this $\mathcal{A}$ should be "smoother" than the truncated one. On the other hand the Ohrenkern also violates Eq. (B.16). The violation is even larger than in case of the comparable truncated optimal $\mathcal{A}$-kernel ($n = 1$). One finds $\max_{z \in \Lambda^0} |1 - \int_x \mathcal{A}(z,x)| = 9.95 \cdot 10^{-2}$.

The numerical result of computations of propagators is somewhat surprising. Despite violation of Eq. (B.16), CSD is completely eliminated in trivial gauge fields. Hence, for the success of an MG algorithm it is not a necessary condition that the lowest mode is interpolated exactly. The result might be taken as an indication that an overlap of blocks is more important. However, CSD remains with the Ohrenkern in nontrivial gauge fields.

Numerical values for the Ohrenkern in the absence of gauge fields are given in Table B.7 of Appendix B.

[22] The author found no suitable translation for this German word.

The Ohrenkern was also used by Mack and Meyer [34,41] in MG simulations of ϕ^4 theory. One might wonder whether their MG Monte Carlo simulations would have been more successful if they also tried the optimal $\mathcal{A}$-kernel. The answer to this speculation is probably negative. The analysis of Grabenstein and Pinn [150] indicates that CSD in ϕ^4 theories cannot be beaten at all by MG algorithms (more precisely speaking, by the kind of unigrid algorithm they considered).

6.6 Inclusion of Updating on a Multigrid Layer Consisting of a Single Site

Conventional one-grid relaxation algorithms plus (2.50).

The remarks made in Sec. 4.8 about one-grid relaxation plus (2.50) apply in nontrivial gauge fields as well. Fig. 4.2 looks the same when m^2 is replaced by Δm^2. At fixed lattice size $|\Lambda|$ the algorithm retains the property that it is practically not sensitive to the disorder of the gauge field, i. e. convergence is only determined by Δm^2, no matter what the value of β is. However, there appears a volume effect, as stated in the caption of Fig. 4.2.

MG algorithms plus (2.50).

Variational MG plus "last point updating" (2.50) seems to beat CSD of *asymptotic* relaxation times, too. Compared with CG and one-grid relaxation plus (2.50), this modified MG performs better the more critical the system is. Convergence for $\Delta m^2 = 10^{-6}$ on 12^4 and 18^4 lattices is shown in Fig. 6.5.

In a fixed volume the results for the Δm^2-dependence of relaxation times in variational MG plus (2.50) look similar to those of one-grid plus (2.50), except that fewer iterations are required to reach a given accuracy. The *asymptotic* slope of the curves is independent of Δm^2 in a fixed volume. Fig. 6.5 shows that the asymptotic slope is also independent of the lattice size. From these findings we conclude, as in the case of trivial gauge fields, that *asymptotic* CSD seems to be eliminated. The volume effect with respect to how long it takes until errors decay exponentially is reasonably mild.

It must be emphasized that Ω in (2.50) really needs to be a matrix. Simple rescaling of $\phi^{(n)}$ with a real number does not accelerate convergence.[23]

[23] A remark about one-grid plus (2.50) is in order here: It can be proved by induction that in $SU(2)$ gauge fields $\Omega \propto \mathbb{1}$ always holds if one starts relaxation with $\phi^{(0)} = 0$. However, $\Omega \propto \mathbb{1}$ is not true in general.

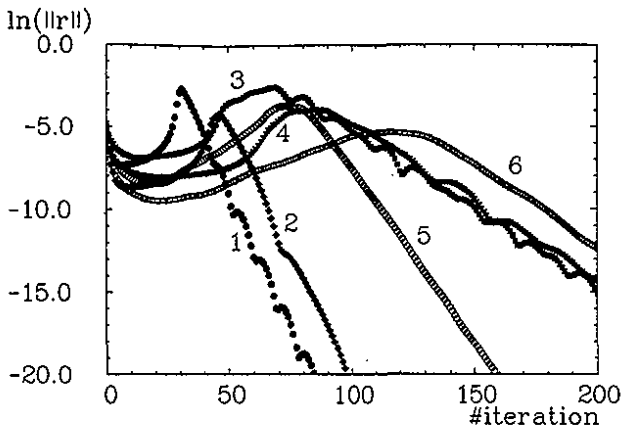


Fig. 6.5: *Convergence for bosonic propagators* $(-\Delta+m_{cr}^2+\Delta m^2)^{-1}$ *with* $\Delta m^2 = 10^{-6}$ *in quenched four-dimensional* $SU(2)$ *gauge fields equilibrated with Wilson's action at* $\beta = 2.7$. *The numbers refer to the following algorithms: 1/2: variational MG SOR* ($\omega = 1.50$) *plus (2.50) on a* $12^4/18^4$ *lattice; 3/4: one-grid SOR* ($\omega = 1.91$) *plus (2.50) on a* $12^4/18^4$ *lattice; 5/6: CG on a* $12^4/18^4$ *lattice. The critical masses are* $m_{cr}^2 = -0.7726281/-0.7554339$. *Without (2.50) MG SOR and one-grid SOR have relaxation times of* $O(10^5)$ *(Fig. 6.4).*

6.7 Inclusion of the Related Improvements of Section 2.9

The improved MG correction scheme (2.51) also beats CSD of asymptotic relaxation times. (Note that the parameter δ of (4.2) is eliminated when (2.51) is used.) However, at finite β (2.51) is not able to halve relaxation times, in contrast to the trivial case. Therefore the increased computational labor does not pay.

Another result with the improved coarse grid correction scheme (2.51) is the following. It was found that the performance of (2.51) with fixed $\Omega = \mathbb{1}$ plus (2.50) is equivalent to (2.51) with automatic determination of Ω.

Using the modified SOR version (2.52) does not pay. If one fixes $\Xi = \mathbb{1}$, one has nothing else but conventional Gauss-Seidel relaxation. Retaining Ξ yields little difference in performance, because after a few iteration $\Xi \approx \mathbb{1}$.

7 Computation of Gauge Covariant Propagators of Staggered Fermions in Four-Dimensional SU(2) Gauge Fields

This section contains reports about computations of propagators of staggered fermions. We begin again with looking at the CG algorithm, and discuss one-grid relaxation algorithms after that. The MG methods discussed then are: (i) variational MG, (ii) a slightly different scheme with a non-Galerkin coarse grid operator, (iii) inclusion of updating on an MG layer consisting of a single site, (iv) modified minimal residual algorithms, and finally, (v) we discuss the use of the $\sqrt{3}$ blocking scheme of Cordery, Gupta, and Novotny [151] in MG computations. In the section on point (iii) we put special emphasis on the gauge covariance properties of the procedure.

We recall that the computational details of Sec. 4.1 apply in this section as well.
If not stated explicitly, variational MG always means variational MG with the Laplacian choice of the averaging operator C.

Summary of results.

The CG algorithm for propagators of staggered fermions shares the characteristic features with the bosonic case qualitatively. Quantitatively CG appears to be even more sensitive to the disorder of the gauge field. For given mass the convergence speed slows down considerable when β is decreased, and when the lattice size is increased.

For staggered fermions it was verified numerically that the scaling relation (2.15) also holds for traditional one-grid relaxation algorithms, and for the variational MG algorithm with the Laplacian choice of C. The critical exponent z equals 2 (fixed ω). It was also verified numerically that the constant of proportionality in (2.15) is *independent of the lattice size*. Hence, the two algorithms mentioned will have CSD under conditions where the inverse mass and the linear extension of the lattice are changed proportionally.

The analytical predictions which can be made in case of traditional relaxation for bosonic propagators are not valid for staggered fermions (except Ostrowski's Theorem). In particular, the undamped Jacobi iteration does not converge in nontrivial gauge fields. This finding suggests that Jacobi is not a useful smoother in MG computations for fermionic propagators.

Despite non-elimination of CSD, MG with variational coarsening works better than traditional relaxation methods, but on lattices up to 18^4 this MG algorithm is not competitive with CG. We expect that the situation will be different on larger lattices. However, we cannot predict how large the lattices have to be. We speculate that $SU(2)$ gauge fields at realistic β-values are too disordered to make the Laplacian choice of the averaging operator C appropriate *in variational MG, i. e. when* $A = C^*$.

An alternative definition of a coarse grid operator D_1 was investigated. In this alternative a coarse grid Dirac operator is defined through a gauge field on the block lattice as proposed in Sec. 3.3. With the alternative D_1 the energy functional is not necessarily lowered in the coarse grid correction step. As a consequence we encounter the existence of problematic gauge field configurations where MG fails to converge. Similar results were also reported in Refs. [117,102,118]. An improved gauge covariant coarse grid correction scheme is able to cope with the problematic gauge field configurations, but it is not competitive in computer time with variational MG, and therefore it is also not competitive with CG.

In Sec. 7.5 a generalization of the "updating on a last site" as discussed for bosons in Sec. 2.8 to the case of staggered fermions is proposed. The gauge covariance properties of the generalization are discussed in detail. However, the numerical results are disappointing. On lattices up to 18^4 the method does not pay. The method is able to eliminate the asymptotic $1/\triangle m^2$ divergence of relaxation times in a fixed volume. But we feel unable to predict from numerical experiments whether the method will pay for staggered fermions on larger lattices of realizable sizes, because a volume effect which is larger than in case of bosons remains with respect to how long it takes until errors decay exponentially. Moreover, the asymptotic decay rate is not independent of the lattice size. Therefore it seems that not even CSD of *asymptotic* relaxation times is eliminated. A careful study of this question could not be done.

In Sec. 7.6 the performance of some modified MG minimal residual algorithms is reported. However, none of the tested modifications is able to improve the conventional minimal residual method in practice.

Finally the use of the $\sqrt{3}$ blocking scheme of Cordery, Gupta, and Novotny [151] in MG computations was investigated. In this scheme blocks are overlapping. This overlap leads to a complication of variational MG, because the coarse grid operator is denser than the original fine grid operator. An approximation where the gauge field degrees of freedom are reduced by 1/9 was implemented numerically. The results confirm the conjecture that an overlap of blocks accelerates the convergence speed. However, the costs of additional computational labor could not be outweighed.

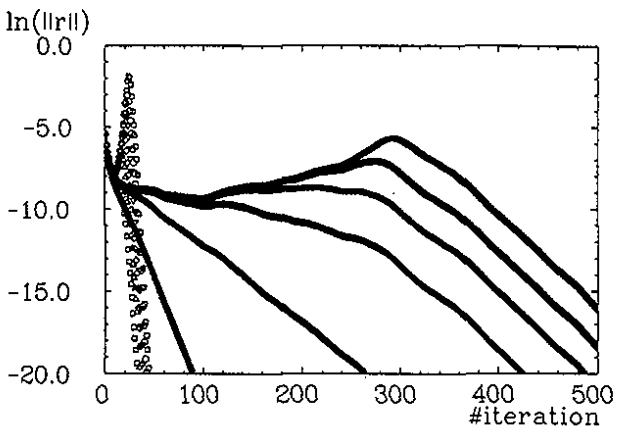


Fig. 7.1: *Convergence of CG for propagators of staggered fermions* $(-\not{D}^2 + m_{cr}^2 + \Delta m^2)^{-1}$ *on an* 18^4 *lattice. Curves of open symbols are results of computations in a pure gauge field* ($m_{cr}^2 = 0$), *while curves of solid symbols are obtained in the background of a quenched* $SU(2)$ *gauge field at* $\beta = 2.7$. ($m_{cr}^2 = -0.0096640$.) *Within each of the two families the curves are parameterized by* $\Delta m^2 = 0.1, 0.01, \ldots, 10^{-6}$; Δm^2 *decreasing from left to right. The initial guess for the propagator is zero in all cases.*

7.1 The Competitor: Conjugate Gradient Algorithm

We begin also this section with looking at the competitor algorithm. Convergence of CG in case of staggered fermions is shown in Fig. 7.1. Further results are in Table 7.2 below. No preconditioning of CG is employed because it turned out that preconditioning does not lead to an improvement in case of staggered fermions [69]. Hence it is the unpreconditioned CG algorithm which sets the standard.

Qualitatively CG has the same features for propagators of staggered fermions than it has for bosons. This concerns the not necessarily monotonic decrease of $\|r^{(n)}\|$ for small masses, and also the points how the lattice size and the disorder of the gauge field affect the convergence. Cf. Sec. 6.1. Quantitatively it appears that CG for staggered fermions is more dependent on β than it is for bosons.

7.2 One-Grid Relaxation Algorithms

Relevance of m_{cr}^2 in case of staggered fermions for a proper judging of CSD.

In case of staggered fermions m_{cr}^2 is much closer to zero than in case of bosons. Actually, it is often assumed that the finite value of m_{cr}^2 can be neglected completely in (2.15) so that $\tau \propto m^{-z}$. However, this neglect of m_{cr}^2 is not justified on lattices amenable in size to date.

For bosons, scaling (2.15) without violations is observed for $\Delta m^2 \lesssim 0.01$ (practically independent of $|\Lambda|$ and β), Sec. 6.3. For staggered fermions it was verified numerically that the scaling relation (2.15) also holds for traditional one-grid relaxation algorithms. The critical exponent z equals 2 (fixed ω). It was also verified numerically that the constant of proportionality in (2.15) is *independent of the lattice size*. Below it will be reported that this finding also holds for variational MG with the Laplacian choice of C.

Hence, one-grid relaxation algorithms and the variational MG algorithm with the Laplacian choice of C will have CSD under conditions where the inverse mass and the linear extension of the lattice are changed proportionally.

To obtain these results requires a careful analysis of data. The subtle point is that for fermions (2.15) is obeyed only for $\Delta m^2 \lesssim 0.001$, and asymptotic decay in the sense that only the slowest mode governs convergence does not set in before 400 – 500 iterations in four-dimensional $SU(2)$ gauge fields.

The lesson is that m_{cr}^2 must not be neglected, at least for lattices up to 18^4. Otherwise one cannot study the phenomenon of CSD thoroughly. In practice this might mean that on such relatively small lattices m^2 must also be given small negative values. This is of course artificial, but it is necessary when one wants to obtain results which are reliable for predictions how algorithms perform on large lattices. (For larger $|\Lambda|$ the system becomes critical already for small and positive m^2; then there is no need to test algorithms with negative values of m^2.) We recall the discussion of Sec. 2.10. If one does not investigate systems close enough to criticality and if one does not take care that decay rates are really asymptotic, there is the danger of extracting wrong values for z, even in case of pure gauges.

Analytical predictions for traditional relaxation methods.

The elaborated mathematical theory which is applicable in case of bosonic propagators cannot be applied to fermions. The only thing which remains valid is the implication of Ostrowski's Theorem: SOR on one grid converges for $0 < \omega < 2$, because $(-\not{D}^2 + m^2)$ is positive definite. However, $(-\not{D}^2 + m^2)$ is not a consistently ordered 2-cyclic operator when the field strength term does not vanish, i. e. as soon as the gauge field is not trivial. Therefore the conclusions, which were drawn for bosonic propagators in Sec. 2.2 and checked in Secs. 4.4 and 6.2, do not hold for fermions.

In particular, two points are worth mentioning.

- First, for given relaxation parameter ω (Jacobi or SOR) and given Δm^2, convergence is the same in trivial and in any nontrivial gauge fields in case of bosons. The same is not true for fermions.
- Second, the undamped Jacobi iteration does *not* converge for propagators of staggered fermions in nontrivial gauge fields. There is a threshold < 1 for the relaxation parameter ω below which damped Jacobi converges. The threshold coincides roughly with an optimal choice of ω. This value of ω is picked automatically by the minimal residual algorithm. The threshold is lowered the more disordered the gauge field becomes. This fact suggests that *(undamped) Jacobi relaxation is not a useful smoother in MG computations with fermions.*

7.3 Variational Multigrid

The introductory remarks of Sec. 6.3 should be recalled at this point.

Numerical implementation of the Laplacian choice of C.

The considerations concerning the choice of a good block spin in Sec. 3 gave the hint that the Laplacian choice (3.16) for C is sufficient. The optimal interpolation kernel $\mathcal{A}$ associated with this C has the desired fall-off properties in nontrivial gauge fields; recall Sec. 3.5. Therefore it does not seem to be necessary to implement the more intricate proposals discussed in Sec. 3.2. Hence, only the Laplacian form of the averaging kernel has been implemented on a computer yet. But the problem of $\mathcal{A}$ remains – if one puts $\mathcal{A} = C^*$ the problem is back.

For the computation of C we could use the same computer program as for bosons, except for a few marginal changes, namely replacements of address lists and of parallel transporters in the Laplacian. The efficiency of the computer program is retained for staggered fermions. [Perhaps some efficiency is lost when the algorithm will be used for the Diracian C. The question is how the gap between the lowest and the second lowest eigenvalue, which equals 1 in pure gauges, decreases at finite values of β; cf. the remarks in Sec. 3.4.]

The algorithm described in Sec. 3.4 depends essentially on the absence of degeneracies of the eigenvalues of $-\Delta_{N_t,x}$ (modulo the intrinsic N_c-fold degeneracy discussed in Sec. 2.7). This nondegeneracy condition is preserved in our blocking procedure, because the averaging is chosen such that it leads to staggered fermions on the coarse lattice again (unlike [73, 107,108,109,110]).

CSD of variational MG with the Laplacian choice of C.

As mentioned in Sec. 7.2, the scaling relation (2.15) also holds for variational MG with the Laplacian choice of C in case of staggered fermions, with $z = 2$. Therefore we conclude that this algorithm will not be able to eliminate CSD on large lattices.

If one is interested in the question how well algorithms perform on lattices of sizes that are used in present day simulations, the discussion in Sec. 7.2 might appear too academic. It might appear more natural to ask how many iterations are needed to obtain a given accuracy. From this viewpoint it turned out that MG is competitive or even superior to conventional algorithms in two-dimensional models [72,100,117,144,102,118]. However, it is more difficult to reach decisive conclusions in $d = 4$. Preliminary results on 16^4 [110] and 18^4 [152] lattices (both at $\beta = 2.7$) indicated that the MG methods tested so far will *not be able to outperform the CG algorithm.*

The current status is as follows [153]. Despite non-elimination of CSD, MG with variational coarsening works better than traditional relaxation methods, but on lattices up to 18^4 this MG algorithm is not competitive with CG. We expect that the situation will be different on larger lattices. However, we cannot predict how large the lattices have to be.

Fig. 7.2 shows convergence for propagators on 12^4 and 18^4 lattices with $m^2 = 0.01$ and $m^2 = 0$. Results for a propagator on an 18^4 lattice with $m^2 = 0$ using various algorithms

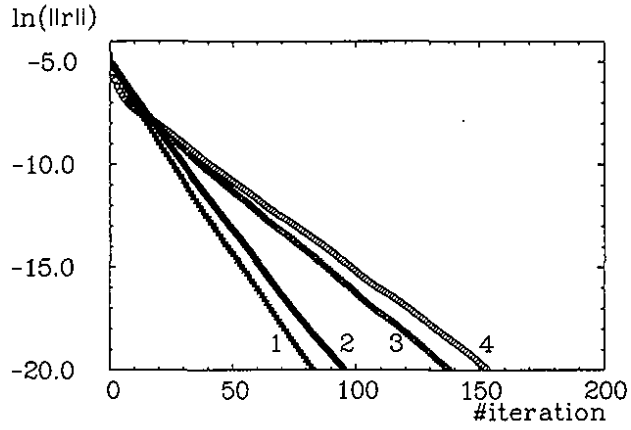
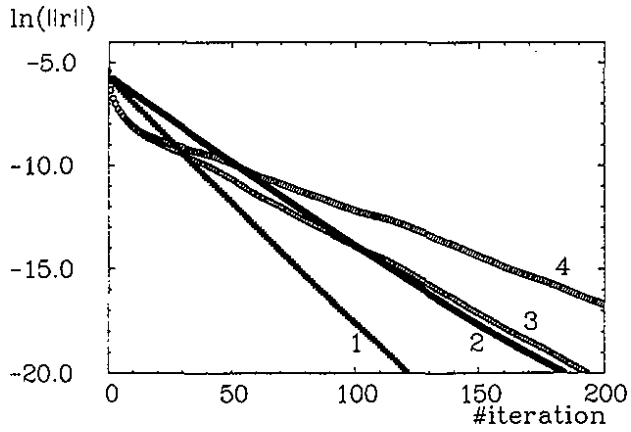


Fig. 7.2: *Computation of propagators of staggered fermions $(-\not{p}^2 + m^2)^{-1}$ for two different gauge fields on 12^4 (top) and 18^4 (bottom) lattices in quenched four-dimensional $SU(2)$ gauge fields at $\beta = 2.7$. The numbers refer to the following algorithms: 1/2: variational MG SOR (lexicographic ordering) with $m^2 = 0.01/0.0$, $\omega = 1.80/1.83$ (top), $\omega = 1.88/1.92$ (bottom); 3/4: CG with $m^2 = 0.01/0.0$. The critical masses of the gauge fields are $m^2_{cr} = -0.0368447$ (top) and $m^2_{cr} = -0.0096640$ (bottom).*



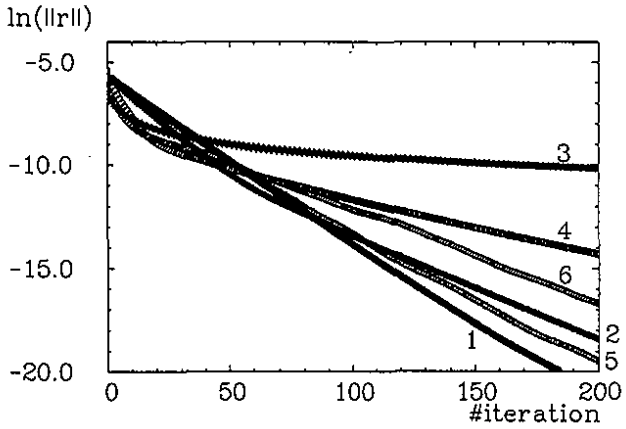


Fig. 7.3: *Computation of a propagator of staggered fermions $(-\not{D}^2+m^2)^{-1}$ on an 18^4 lattice with $m^2 = 0$ in a quenched four-dimensional $SU(2)$ gauge field at $\beta = 2.7$ ($m_{cr}^2 = -0.0096640$). The numbers refer to the following algorithms: 1/2: variational MG SOR with $\omega = 1.92/1.90$; 3: MG minimal residual; 4/5: one-grid SOR with $\omega = 1.80/1.91$; 6: CG. The SOR algorithms are swept in lexicographic ordering.*

are plotted in Fig. 7.3. No damping parameter δ (Eq. (4.2)) is used here.

In Fig. 7.3 results are shown for an MG minimal residual (MR) algorithm. In the MG MR algorithm, SOR on the fine lattice is replaced by the conventional MR algorithm. This algorithm works as follows. One replaces $\chi^{(n)}$ by $\chi^{(n+1)} = \chi^{(n)} + \alpha r^{(n)}$, where the number α is chosen such that the energy functional of $\chi^{(n+1)}$ is minimized. This algorithm is used by Hulsebos et al. [73,107,108,109,110].

Conclusions.

The conclusions which can be drawn from Figs. 7.2 and 7.3 are:

(*i*) The best value for the relaxation parameter ω in variational MG SOR and one-grid SOR increases the more critical the system becomes.

(*ii*) The (non-over-relaxed) MG minimal residual algorithm is clearly non-competitive with any other algorithm.

(*iii*) When ω is not chosen close enough to its optimal value, it might be that even the simple one-grid SOR algorithm performs better than variational MG with the Laplacian choice of C.

		lower bounds for $\lambda_0(x)$		average values $\langle\lambda_0(x)\rangle$		upper bounds for $\lambda_0(x)$	
β	$\lvert\Lambda\rvert$	bosons	fermions	bosons	fermions	bosons	fermions
∞	any	0	0	0	0	0	0
10.0	18^4	0.05	0.14	0.07	0.19	0.10	0.24
5.0	18^4	0.12	0.30	0.16	0.39	0.21	0.51
3.0	18^4	0.21	0.56	0.30	0.70	0.39	0.86
2.7	18^4	0.26	0.63	0.35	0.79	0.46	0.97
2.7	12^4	0.26	0.66	0.35	0.79	0.44	0.96

Table 7.1: *Lowest eigenvalues $\lambda_0(x)$ of $-\Delta_{N,x}$ for bosons and staggered fermions in the same four-dimensional $SU(2)$ gauge field configurations. The scale factor is $L_b = 3$ in both cases.*

The Laplacian choice of C in variational MG might be inappropriate at β-values of interest.

This last result (*iii*) might be a warning sign that an $SU(2)$ gauge field at $\beta = 2.7$ is still too disordered to make the Laplacian choice of the averaging operator C appropriate here, *i. e. when one chooses* $\mathcal{A} = C^*$. Indeed, it might well be that the Laplacian choice works satisfactorily *in variational MG* only on extremely large lattices. Since the parallel transporters in the fermionic Laplacian (3.19) are products of two link variables, disorder is in some sense enhanced compared to the bosonic case. This becomes obvious when we consider the lowest eigenvalues of $-\Delta_{N,x}$ which were introduced as measures of disorder and of smoothness. Comparison is done in Table 7.1. A similar comparison is provided by comparing Figs. 3.2 and 3.5. These results suggest that Laplacian smoothness works equivalently in variational MG for bosons at $\beta = 2.7$, and for fermions at $\beta \sim 5$, etc.

From the conventional MG point of view the large value of the relaxation parameter ω which was found optimal in variational MG SOR, indicates that the Laplacian choice of C is inappropriate here at β-values of interest. According to the conventional MG wisdom Gauss-Seidel relaxation ($\omega = 1$) should be a good smoother. We noted already in Sec. 6.4 that the picture changes in nontrivial gauge fields. In the idealized MG algorithm (Sec. 5) we found that the optimal value of ω increases the stronger the disorder in the gauge field becomes. In the variational MG method for bosonic propagators one needs even larger ω's than in the idealized algorithm in order to achieve peak performance. However, these values are by far not so large than those which were found optimal in case of variational MG for staggered fermions.

A tentative conclusion from Fig. 7.2 and Table 7.1 is that one nevertheless has to implement the Diracian choice of C in order to get a competitive variational MG algorithm at β-values of interest. At least the spectrum of low-lying eigenvalues of $-\not{D}^2$ is not preserved by ground-state projection with $-\Delta_{N,x}$, and this virtually contradicts our philosophy. However, the numerical implementation of the Diracian definition of C has to be left for a future study.

7.4 Coarse Grid Dirac Operator via Blocked Gauge Field

Definition of an alternative coarse grid operator.

In Sec. 3.3 it was proposed to define gauge fields on coarser lattices by Eq. (3.24). If we set $\mathcal{A} = C^*$ we obtain coarse grid link variables which can be used for an alternative definition of the coarse grid operator D_1 in a modified variational MG algorithm. With the Laplacian choice of C (3.24) gives a unique definition of $\mathcal{U}$;[24] there are $(1/L_\delta^d)$-times as many link variables on Λ^1 than on Λ^0.

We define $D_1 = \delta^{-1}(-\not{D}_{eff}^2 + m_{eff}^2)$ with $\not{D}_{eff} = C\not{D}C^*$. The effective mass term is unchanged compared to the previous section. The distinction to the Galerkin operator $-C\not{D}^2C^*$ is the product C^*C in $\not{D}_{eff}^2$. This product introduces a fluctuating length into the matrix elements of D_1. In the limiting case of a pure gauge this length makes essentially no difference because it is site-independent. However, in nontrivial gauge fields the fluctuating length is site-dependent.

Numerical results.

The results obtained with this modified algorithm are as follows. It is essential to introduce the parameter δ of (4.2) into D_1. Otherwise one cannot cope with all gauge field configurations. The existence of such problematic configurations where MG fails to converge was also reported in Refs. [117,102,118] in two-dimensional $SU(2)$ and $SU(3)$ gauge fields.

The authors of Refs. [72,100,117] report that in case of gauge group $U(1)$ their algorithm performs considerably better when they re-unitarize the blocked gauge field (matrix elements of the effective Dirac operator) such that it is again an element of the gauge group.

This point was checked in four-dimensional $SU(2)$ with our alternative definition of the coarse grid operator. We compared the case of $\not{D}_{eff}$ as described above with a definition of $\not{D}_{eff}$ where $\mathcal{U}(x,y)$ is replaced by $\mathcal{U}(x,y)/\|\mathcal{U}(x,y)\|$. The results are such that the difference between using a dielectric and a unitary gauge field on coarser layers can be compensated by adjusting δ in D_1. By tuning δ, one can re-unitarize the blocked gauge field "on the average".

Yet, compared with variational MG one monitors a worse convergence, even with optimized values of δ and ω. We ascribe this to the fact that the MG correction step does not necessarily lower the energy functional when D_1 is not the Galerkin operator.

Improved coarse grid correction scheme.

An idea of Sec. 2.9 was taken up. An improved MG coarse grid correction step was implemented according to (cf. (2.51))

$$\chi^{(n)}(z) \mapsto \chi'^{(n)} \equiv \chi^{(n)}(z) + \varphi^{(n)}(z)\,\Theta(H(z)) \;, \tag{7.1}$$

[24] If C were determined by Diracian ground-state projection, there would automatically be nonvanishing "link" variables $\mathcal{U}(x,y)$ with $y = x \pm \frac{1}{2}L_b(e_\mu \pm e_\nu)$, $\mu \neq \nu$. In case of the Galerkin operator $-C\not{D}^2C^*$ this happens already with the Laplacian choice of C.

where $\varphi^{(n)}$ is the interpolated solution of the residual equation on Λ^1. In (7.1) we introduced $2^d = 16$ $N_c \times N_c$ matrices $\Theta(H(z))$. These matrices depend on the pseudoflavor H of z. They are determined automatically by the algorithm in such a way that the energy functional of $\chi'^{(n)}$ is minimized. This eliminates the parameter δ.

The finding is that the performance of the algorithm with (7.1) is improved and becomes equivalent in numbers of iterations to variational MG, except that the optimal value of ω is a little bit lowered. Variational MG is however favored, because it needs less computational labor.

Remark on the pseudoflavor dependence of the matrices $\Theta(H)$ in (7.1).

We add a comment on the matrices $\Theta(H(z))$ in (7.1) concerning the pseudoflavor dependence.

In the limiting case of pure gauges the fermionic problem is equivalent to computing 2^d independent (uncoupled) bosonic propagators, one for each pseudoflavor. It is obvious that one should have different matrices Θ for the different pseudoflavors. Hence, we make Θ dependent on $H(z)$, and we propose to retain this dependence also in nontrivial gauge fields.

Compared with (2.51), there are no matrices $\Omega(H(z))$ in (7.1). The $\Omega(H(z))$'s were left in order to avoid unnecessarily complex programming. In case of bosons it was found that the performance of (2.51) with fixed $\Omega = \mathbb{1}$ plus (2.50) is equivalent to (2.51) with automatic determination of Ω. It is reasonable to expect a similar result for fermions. Therefore the full generalization of (2.51) could be studied in the framework of the next section.

Remark on the gauge covariance properties of (7.1).

The 2^d $N_c \times N_c$ matrices $\Theta(H)$ in (7.1) are fixed by the requirement of minimization of the energy functional K (2.20). The energy functional is gauge invariant. By computing the partial derivatives of $K[\chi'^{(n)}]$ with respect to the variables which parameterize the matrices $\Theta(H)$, and equating this gradient with zero, one obtains a linear system of equations whose solution yields the $\Theta(H)$'s.

Since K is gauge invariant, the solution of the linear system of equations will exhibit gauge covariance. For a more comprehensive discussion we refer to the analogous consideration in the next subsection.

7.5 Inclusion of Updating on a Multigrid Layer Consisting of a Single Site

As we saw in Sec. 6.6, CSD of asymptotic relaxation times in computations of bosonic propagators by variational MG could practically be eliminated, except for a reasonable volume effect, by including (2.50) in the algorithm. If a similar finding held also in case of staggered fermions, an efficient algorithm for propagators would have been found.

First of all, however, (2.50) should not be used straightforwardly for staggered fermions. The "last site" has to be chosen as a 2^d-point basis of the staggered lattice (cf. Sec. 3.2 for the 2^d-point basis).

7.5.1 Generalization of the Proposal of Section 2.8 for Fermions

We resume the discussion from the end of the preceding section. In case of staggered fermions we have to take pseudoflavor into account.

The case of pure gauges.

In the limiting case of vanishing gauge coupling, Eq. (2.1) is an equation for 2^d uncoupled bosonic propagators, one for each pseudoflavor H. This consideration leads us to introduce pseudoflavor dependent rescaling matrices Ω. We generalize the global rescaling in (2.50) according to

$$\chi^{(n)}(z) \mapsto \chi^{(n)}(z)\,\Omega(H(z)) \ . \tag{7.2}$$

Here the $N_c \times N_c$ matrices Ω depend on pseudoflavor H, and $\chi^{(n)}(z)$ is rescaled with $\Omega(H)$ when H is the pseudoflavor $H(z)$ of z.

The 2^d matrices $\Omega(H)$ are determined by the requirement that the energy functional K (2.20) of $[\chi^{(n)}\,\Omega]$ gets minimized. In the case of pure gauges there is no field strength term $F_{\mu\nu}$ in the square of the Dirac operator, and the energy functional decomposes into a sum of 2^d energy functionals. Each of these 2^d terms is an energy functional for a free bosonic propagator on a particular pseudoflavor sublattice. Hence, the $\Omega(H)$'s are given by expressions as in (2.50), but $\phi^{(n)}$ has to be replaced by $\chi^{(n)}$ restricted to the pseudoflavor sublattices Λ_a^H.

The case of nontrivial gauge fields.

In case of nontrivial gauge fields we propose to retain (7.2), and to call it "updating on a last site" for staggered fermions. As remarked above the "last site" is a 2^d-point basis of the lattice of staggered fermions.

We shall also retain the way by which the 2^d $N_c \times N_c$ matrices $\Omega(H)$ are determined. They are fixed automatically by the algorithm through the requirement of minimization of the energy functional K of the rescaled propagator.

In nontrivial gauge fields the expressions for the matrices $\Omega(H)$ are more complicated than the expression given in (2.50) for Ω in case of bosons. By computing the partial derivatives of $K[\chi^{(n)}\,\Omega]$ with respect to the variables which parameterize the matrices $\Omega(H)$, and equating this gradient with zero, one obtains a linear system of equations. Let us call this system Σ. The solution of Σ yields the $\Omega(H)_{i,j}$, $i,j = 1,\ldots,N_c$, for the 2^d pseudoflavors H.

In practice the $\Omega(H)_{i,j}$ can be computed by solving Σ by means of the CG algorithm. One could make use of the fact that the square of the Dirac operator couples pseudoflavors H and K only when the sublattices Λ_a^H and Λ_a^K are both subsets either of the even or of the odd sublattice of $\Lambda_{a/2}$. Therefore Σ can be decoupled into two systems.

With either choice, the expensive part in this calculation is not the solution of linear systems by CG, but the computation of the matrices of coefficients of these systems.

We remark that Σ decouples into 2^d systems in the limiting case of vanishing gauge coupling. Then the solution for $\Omega(H)$ is given by the expression in (2.50) where one has to take for $\phi^{(n)}$ the approximation $\chi^{(n)}$ restricted to the pseudoflavor lattice Λ_a^H, as discussed above.

Gauge covariance.

Concerning the aspect of gauge covariance of the proposal (7.2), one could almost repeat the discussion of Sec. 2.8. Basically, gauge covariance follows from the gauge invariance of the energy functional K (2.20) which is to be minimized.

When f is a pseudofermion field, the coefficients in the matrix of the linear system Σ and also the r.h.s. of Σ are gauge invariant. Therefore the solution of Σ, i. e. the 2^d matrices $\Omega(H)$, are gauge invariant. It follows that (7.2) is a gauge covariant operation. Analogously to the bosonic case one does not have to identify any reference sites $\hat{x}_N \in \Lambda_{a/2}$ with the "last site". [Cf. the discussion below for a δ-function f.]

We remark that gauge covariance can be exactly ensured in practice. It was proposed to use CG for the solution of Σ. When the initial guess in CG is fixed to zero, and because CG is a gauge covariant iteration prescription, it follows that all CG iterates $\Omega(H)^{(n)}$ and the iterates ${}^g\Omega(H)^{(n)}$ in a gauge transformed configuration are exactly connected by the 1 : 1-mapping g. If the CG iteration is stopped when the norm of the residual of $\Omega(H)^{(n)}$ has reached a given precision, then CG would be stopped at the same iteration number ${}^{(n)}$ in the gauge transformed configuration, because the norms of the residuals of all $\Omega(H)^{(n)}$ and ${}^g\Omega(H)^{(n)}$ are equal because of gauge covariance. Hence, the solutions $\Omega(H)$ and ${}^g\Omega(H)$ can indeed be mapped onto one another through the gauge transformation g. Of course this whole discussion is valid up to round-off errors, but these are not relevant in practice.

In case that $f(z) = \delta_{z,w}$, the coefficients in the matrix of the linear system Σ and also the r.h.s. of Σ transform under gauge transformations like a matter field in the adjoint representation sitting at site w. Therefore also the solution matrices $\Omega(H)$ will have this gauge covariance property, i. e. if $\Omega(H)$ is the solution of Σ in a particular gauge field configuration, then $g(w)\,\Omega(H)\,g(w)^{-1}$ is the solution in the g-transformed configuration. Analogously to the bosonic case it follows that (7.2) is a gauge covariant prescription. [When w is a site in the even sublattice of $\Lambda_{a/2}$, then $\chi^{(n)}(z)$ will be zero for odd z, and vice versa.]

We add a comment on semantics. Since we defined $D = (-\not{p}^2 + m^2)$ for staggered fermions and $(-\not{p}^2 + m^2) = (\not{p} + m)(-\not{p} + m)$, strictly speaking the propagator $(\not{p} + m)^{-1}(z,w)$ is not $\chi(z)$ for $f(z) = \delta_{z,w}$, but $([-\not{p}+m]\,\chi)(z)$. Nevertheless, gauge covariance is ensured.

We argued in the bosonic case that w is a distinguished site when $f(z) = \delta_{z,w}$, and we identified this site w with the reference site $\hat{x}_N$ of the "last layer" $\Lambda^N = \{x_n\}$. The fermionic analog of this identification in case of a δ-function f requires some discussion.

Our interpolation kernel is the approximate propagator $\chi^{(n)}$, and $\chi^{(n)}$ transforms under a gauge transformation g according to $\chi^{(n)}(z) \mapsto g(z)\,\chi^{(n)}(z)\,g(w)^{-1}$. Hence, $\chi^{(n)}$ regarded as a kernel has the desired property (2.33) with $\hat{x} = w$. So we have to identify the argument x which represents in the interpolation kernel $\mathcal{A}(z,x)$ "the" site in the last layer with w. It is only a subtlety but not an inconsistency that the last site appears here as a particular site in $\Lambda_{a/2}$ whereas it appeared before as a 2^d-point basis of the staggered lattice.

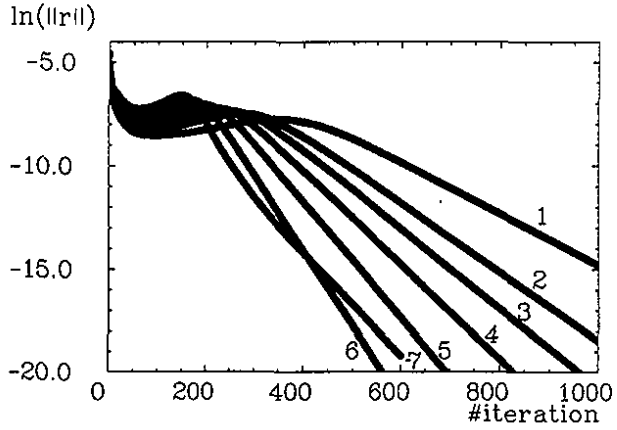


Fig. 7.4: *Dependence on ω in variational MG SOR (lexicographic ordering) plus (7.2): Computation of a propagator of staggered fermions $(-\not{D}^2 + m_{cr}^2 + \Delta m^2)^{-1}$ on a 12^4 lattice with $\Delta m^2 = 10^{-4}$ in a quenched four-dimensional $SU(2)$ gauge field at $\beta = 2.7$ ($m_{cr}^2 = -0.0368447$). The numbers 1–7 refer to the following values of ω: 1.83, 1.88, 1.90, 1.92, 1.94, 1.96, and 1.98.*

7.5.2 Numerical Results

Let us now look at the numerical implementation of (7.2).

Unfortunately, the inclusion of (7.2) has not such a striking effect as it has in case of bosons. This might be (at least partially) due to the fact that we use lexicographic ordering in SOR instead of checkerboard updating.

Global rescaling (7.2) does not pay for positive m^2 on small lattices ($\lesssim 18^4$).

The $1/\Delta m^2$ divergence of *asymptotic* relaxation times can be eliminated in a fixed volume. As in the bosonic case one finds also in case of staggered fermions that $\Omega(H) = \mathbb{1}$ (for all H) as soon as $||r||$ decays exponentially. Then the step (7.2) could be switched off. However, this finding is of no practical importance because of two points. Firstly, there remains a volume effect with respect to how long it takes until errors decay exponentially, and this volume effect is larger than it is for bosons. Secondly, the asymptotic decay rate does not stay constant when the lattice size is increased, in contrast to the bosonic case.

Fig. 7.4 displays the dependence on the relaxation parameter in variational MG SOR for $\Delta m^2 = 0.0001$ on a 12^4 lattice at $\beta \approx 2.7$. The optimal value $\omega = 1.96$ was fixed, and the dependence on Δm^2 was investigated. Results are in Fig. 7.5. For comparison, Fig. 7.6 gives results for a one-grid relaxation plus (7.2). Comparison with CG on 12^4 and 18^4 lattices is done in Fig. 7.7 for $\Delta m^2 = 10^{-6}$. Table 7.2 gives a survey of convergence on 12^4 and 18^4 lattices.

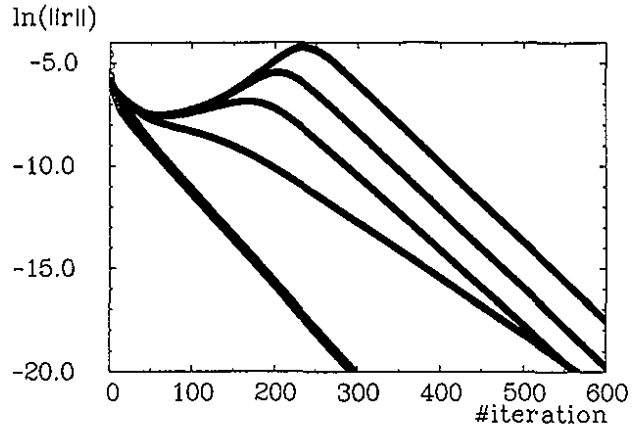


Fig. 7.5: *Dependence on Δm^2 in variational MG SOR (lexicographic ordering, $\omega = 1.96$) plus (7.2): Computation of propagators of staggered fermions $(-\not{D}^2 + m_{cr}^2 + \Delta m^2)^{-1}$ on a 12^4 lattice in a quenched four-dimensional $SU(2)$ gauge field at $\beta = 2.7$ ($m_{cr}^2 = -0.0368447$). The six curves correspond to $\Delta m^2 = 0.1, 0.01, \ldots, 10^{-6}$ with Δm^2 decreasing from left to right. (The curves for $\Delta m^2 = 0.1$ and $\Delta m^2 = 0.01$ lie on top of each other.)*

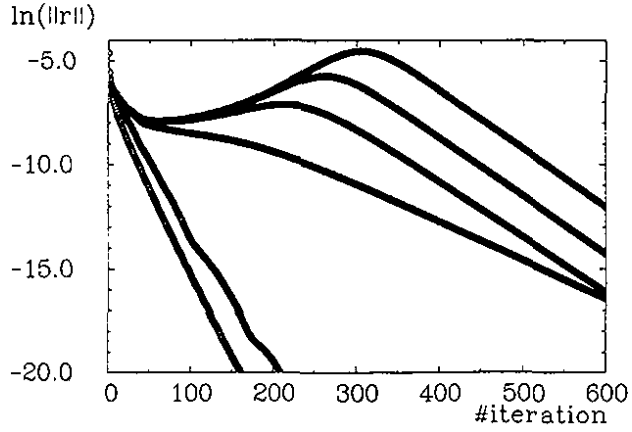


Fig. 7.6: *Dependence on Δm^2 in one-grid SOR (lexicographic ordering, $\omega = 1.90$) plus (7.2): Computation of propagators of staggered fermions $(-\not{D}^2 + m_{cr}^2 + \Delta m^2)^{-1}$ on a 12^4 lattice in a quenched four-dimensional $SU(2)$ gauge field at $\beta = 2.7$ ($m_{cr}^2 = -0.0368447$). The six curves correspond to $\Delta m^2 = 0.1, 0.01, \ldots, 10^{-6}$ with Δm^2 decreasing from left to right.*

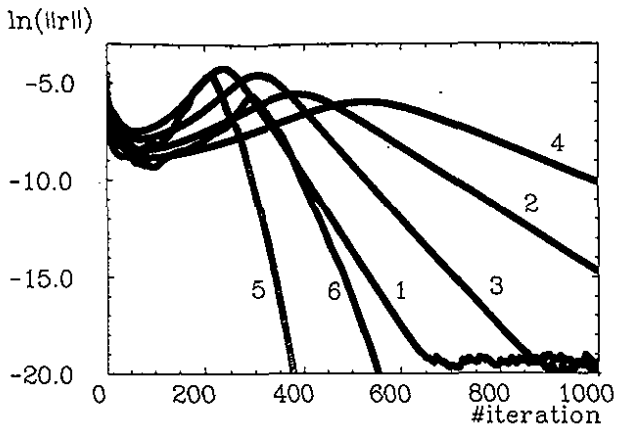


Fig. 7.7: *Convergence for propagators of staggered fermions* $(-\not{p}^2 + m_{cr}^2 + \triangle m^2)^{-1}$ *with* $\triangle m^2 = 10^{-6}$ *in quenched four-dimensional* $SU(2)$ *gauge fields at* $\beta = 2.7$. *The numbers refer to the following algorithms: 1/2: variational MG SOR* ($\omega = 1.96$) *plus (7.2) on a* $12^4/18^4$ *lattice; 3/4: one-grid SOR* ($\omega = 1.90$) *plus (7.2) on a* $12^4/18^4$ *lattice; 5/6: CG on a* $12^4/18^4$ *lattice. Relaxation algorithms are swept in lexicographic ordering. The critical masses are* $m_{cr}^2 = -0.0368447/-0.0096640$. *Without (7.2) MG SOR and one-grid SOR have* τ*'s of* $O(10^5)$.

7.6 Modified Minimal Residual Algorithms

The conventional minimal residual (MR) algorithm works as follows. One replaces $\chi^{(n)}$ by $\chi^{(n+1)} = \chi^{(n)} + \alpha r^{(n)}$, where the number α is chosen such that the energy functional of $\chi^{(n+1)}$ is minimized.

In Sec. 7.3 we have already ruled out the MG MR algorithm. But an improvement from the variational point of view was tested.

We proposed in Sec. 2.9 to generalize the MR method analogously to (2.51) by introducing two matrices Ω and Θ. In case of staggered fermions this proposal should of course be generalized to working with pseudoflavor dependent matrices $\Omega(H)$ and $\Theta(H)$.

However, the numerical result is disappointing. Working with pseudoflavor dependent matrices leads to no practical improvement. Also the inclusion of (7.2) does not make MG MR and the modified MG MR competitive to MG SOR.

Over-relaxed MR versions have not been investigated yet.

pure gauge configurations ($\beta = \infty$)						
	$m^2 =$					
algorithm and lattice size	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}
CG on 12^4	17	17	17	17	17	17
CG on 18^4	28	30	33	35	37	39
one-grid SOR plus (7.2), $\omega = 1.90$, on 12^4	105	125	155	195	230	270
one-grid SOR plus (7.2), $\omega = 1.90$, on 18^4	110	125	155	195	230	270
MG SOR without (7.2), $\omega = 1.09$, on 12^4	21	23	23	23	23	23
MG SOR without (7.2), $\omega = 1.09$, on 18^4	20	23	23	23	23	23
nontrivial gauge fields ($\beta = 2.7$)						
	$\triangle m^2 =$					
algorithm and lattice size	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}
CG on 12^4	65	175	265	300	320	340
CG on 18^4	65	180	350	415	455	495
one-grid SOR plus (7.2), $\omega = 1.90$, on 12^4	100	130	530	560	630	710
one-grid SOR plus (7.2), $\omega = 1.90$, on 18^4	95	125	710	1090	1320	1540
MG SOR plus (7.2), $\omega = 1.96$, on 12^4	180	185	385	425	475	535
MG SOR plus (7.2), $\omega = 1.96$, on 18^4	180	185	530	775	950	1070

Table 7.2: *Convergence in computations of propagators* $(-\not{p}^2 + m_{cr}^2 + \triangle m^2)^{-1}$ *of staggered fermions in four-dimensional* $SU(2)$ *gauge fields. Given is the number of iterations necessary for reducing* $\ln \|r^{(0)}\|$ *by 10. One-grid SOR and MG SOR are swept in lexicographic ordering. The critical mass of the nontrivial gauge field configuration on the* $12^4/18^4$ *lattice is* $m_{cr}^2 = -0.0368447/-0.0096640$.

7.7 Blocking and Interpolation with Overlap: $\sqrt{3}$ – Scheme

The contents of this section has been reported before on occasion of a DFG colloquium in Leipzig [152].

Ben-Av et al. [72,100,117,101,102,118] used a blocking procedure where only equal pseudoflavors are averaged, but where the blocks of one half of the pseudoflavors have an overlap. Overlap means that $\chi(z)$ at a site z makes a contribution to more than one block spin $(C\chi)(x)$. It is reasonable that an overlap of blocks makes the MG approach in some sense "smoother", since the arbitrary block structure appears then less forced upon.

The author tested an MG algorithm with an alternative way of implementing overlapping blocks where the scale factor L_b is $\sqrt{3}$. This blocking was introduced by Cordery, Gupta, and Novotny [151] in the context of Monte Carlo renormalization group studies of four-dimensional gauge theories.

Definition of the $\sqrt{3}$ blocking scheme.

In this blocking procedure the block lattice Λ^1 has (1/9)-times as many sites as Λ^0. Also in the $\sqrt{3}$–scheme the sites of Λ^1 may be considered as a sublattice of Λ^0. Block centers $\hat{x}$ are the sites considered in Sec. 3.2 (encircled symbols in Fig. 3.4), and all translates $\bar{x}$ of them, $\bar{x} = \hat{x} + a\,\tilde{e}_\mu$, $\mu = \pm 1, \ldots, \pm d$, by one of the following vectors

$$\tilde{e}_1 = (0,1,1,1)\ , \quad \tilde{e}_2 = (1,0,1,-1)\ , \quad \tilde{e}_3 = (1,-1,0,1)\ , \quad \tilde{e}_4 = (1,1,-1,0)\ ,$$

and $\tilde{e}_{-\mu} = -\tilde{e}_\mu$.

We retain an extension of the blocks of three lattice spacings in every direction along the axis of the fundamental lattice, so that every block spin $(C\chi)(x)$ receives contributions from 3^4 $\chi(z)$. In this way every $\chi(z)$ makes a contribution to the block spin at nine sites x. The reverse is true for interpolation in the variational method with $\mathcal{A} = C^*$.

Note that the support property (3.14) of averaging is preserved for the Laplacian choice of C, *i. e. a block spin at x receives contributions only from sites z with the same pseudoflavor.*

Problems with variational MG in $\sqrt{3}$ – blocking.

Due to the overlap of blocks, there is now a severe problem with "simple" MG with $\mathcal{A} = C^*$. The locality properties of the (square of the) Dirac operator are not preserved on the block lattice. Every block x is coupled to a large number of other blocks y. This leads to the situation that there are much more gauge field variables on the block lattice than on the original lattice. Also, the mass term Cm^2C^* becomes now a non-diagonal mass matrix.

A crude but simple approximation to variational MG with $L_b = \sqrt{3}$ was tested numerically. Averaging kernels C were computed according to the Laplacian choice (3.16). A coarse grid Dirac operator was defined as $C\not{D}C^*$, cf. Sec. 7.4. Then every block x is coupled to 216 other blocks y. Most of the 216 matrix elements of $C\not{D}C^*$ were neglected. Only those matrix elements $(C\not{D}C^*)(x,y)$ were retained where $y = x + \frac{3}{2}e_\mu$. This means that the gauge field degrees of freedom are reduced by the usual factor of $1/L_b^d = 1/9$. All but the diagonal elements of the effective mass matrix were neglected.

Numerical results.

Numerical results are as follows. On an 18^4 lattice at $\beta = 2.7$ convergence for $m^2 = 0$ and $m^2 = 0.01$ could be speeded up by a factor of about 4.5. This confirms the conjecture that an overlap of blocks accelerates the algorithm. However, this is of no practical relevance because the gain in convergence speed could not outweigh the costs of additional computational labor.

8 Summary and Outlook

In the present work gauge covariant generalizations of multigrid methods for propagators in lattice gauge fields were investigated. Covariant averaging operations for bosons, for staggered fermions, and for gauge fields were discussed. These averaging operations can be used not only in deterministic and stochastic MG computations, but also for Monte Carlo renormalization group studies of gauge theories.

In case of staggered fermions the blocking procedure is not obvious. We required the definition of averaging maps C for fermions to be such that as much as possible is preserved of the residual internal and space-time symmetries on the lattice in the limiting case of vanishing gauge coupling (free fermions). The block spins proposed in the present thesis lead back to staggered fermions on coarser MG layers.

For the purpose of numerical computations we proposed to use the ground-state projection MG method. This method is applicable in arbitrary space-time dimension d and for arbitrary gauge group. We discussed why "ground-state projection" is not an a priori defined scheme because of the boundary conditions which one has to impose on "block-local Hamiltonians". Reasons were given for the choice of Neumann boundary conditions. With this choice one knows analytically that CSD is completely eliminated by variational MG in the limiting case of vanishing gauge coupling.

In ground-state projection MG the (adjoint of the) averaging kernel C satisfies a gauge covariant eigenvalue equation. We presented an efficient numerical algorithm for the solution of this eigenvalue equation. This algorithm is so efficient that we can perform exact ground-state projecting computations. This has the advantage that we do not have to fix any gauge.

In case of staggered fermions two qualitatively different proposals were made for the averaging kernel C – the "Laplacian choice" and the "Diracian choice" – which both reduce to "piecewise-constant" kernels in the absence of gauge fields. Only the Laplacian choice of C has been implemented on a computer yet.

Actual numerical computations were performed in four-dimensional $SU(2)$ gauge fields. It was proved that ground-state projection defines a good block spin, both for bosons and for the Laplacian choice of C in case of staggered fermions. The proof uses arguments from the renormalization group which involve an optimal interpolation kernel $\mathcal{A}$ which is

associated with C. This $\mathcal{A}$ is a gauge covariant generalization of a kernel which was used successfully in rigorous works on constructive quantum field theory.

The same optimal interpolation kernel $\mathcal{A}$ was used in an idealized MG algorithm for bosonic propagators. By means of this idealized MG algorithm it was proved that *the MG method can cope in principle with systems with disorder*. MG computations of propagators without CSD in nontrivial gauge fields are possible, in principle. This statement is true for *any* value of the gauge coupling, including the case of completely random gauge fields. The success of these computations gives ample evidence from the deterministic side that Mack's contention concerning the need for smooth and overlapping interpolation kernels is correct.

Unfortunately, the idealized MG algorithm is not practical because of computational complexity and storage space requirements. But it was important to answer questions of principle.

The ultimate goal of the present thesis is the development of an efficient MG method for the computation of gauge covariant propagators of staggered fermions. The MG method has to be so efficient that it outperforms the best algorithm known, which is CG.

A practical MG method is variational coarsening. In the toy model of bosonic propagators the results of this method are as follows. Variational MG is able to outperform CG, also with respect to computer time. A practical modification which can be thought of as "updating on an MG layer consisting of a single site" succeeds in eliminating CSD of *asymptotic* relaxation times. A remaining volume effect with respect to many iterations it takes until the asymptotic regime is reached is reasonably mild. The method costs some additional computational labor, but as soon as errors decay exponentially, updating at the last site can be switched off. Therefore additional work must only be invested in the initial and in an intermediate stage of computations.

We have not succeeded yet in implementing a method which is competitive with the CG algorithm in case of staggered fermions, at least for lattice sizes up to 18^4.

Numerical results for staggered fermions can be summarized as follows. The scaling relation $\tau = \text{const}/\triangle m^2$ holds for one-grid relaxations and for variational MG with the Laplacian choice of C. const is independent of the lattice size. (For bosonic one-grid relaxations this result is known analytically.) Therefore we conclude that those variational MG algorithm will not be able to eliminate CSD on large lattices.

The fermionic generalization of the "updating at a last site" is successful only in so far that it also eliminates CSD of *asymptotic* relaxation times in a fixed volume. However, this is of no practical relevance because there remains not only a relatively large volume effect with respect to how long it takes until errors decay exponentially, but also the asymptotic decay rate is not independent of the lattice size (in contrast to the situation for bosons).

The implementation of the $\sqrt{3}$ blocking scheme showed that a nontrivial overlap of blocks can lead to an acceleration of convergence. Unfortunately, the acceleration in number of iterations could not outweigh the extra costs of more computational labor.

But other overlapping blocking procedures might be more successful. We recall that the Diracian choice of C has not been implemented and investigated yet. From the physical point of view this proposal, although computationally quite demanding, is preferable. It is also closer to the philosophy of ground-state projection MG. In the Diracian choice the averaging kernel $C(x,z)$ can be nonvanishing in nontrivial gauge fields when z and x carry different pseudoflavors. This might be a crucial point, because it is a qualitatively new kind of overlap of blocks compared with the $\sqrt{3}$ – scheme. It is planned to test the Diracian proposal. However, this has to be deferred to a future publication. One may also think of a further refinement where larger blocks (e. g. 5^4 blocks) overlap already in the limiting case of pure gauges and where $C(x,z)$ may be nonvanishing in nontrivial gauge fields when z and x carry different pseudoflavors.

Finally we note that fruitful results are expected when MG methods are combined with ideas from neural net computations [137,138,140,141].

ACKNOWLEDGMENTS

I would like to thank Professor Gerhard Mack for many stimulating discussions, and for constant support and encouragement. I appreciate being involved in the frontier of the exciting adventure of making MG methods usable for gauge theories.

I am also indebted to many others who contributed with helpful discussions and conversations, in particular the members of the "Hamburg Multigrid Group", and R.C. Brower, R.G. Edwards, H.G. Evertz, S. Meyer, K. Pinn, C. Rebbi, A.D. Sokal, and H.-J. Timme.

I wish to thank J. Smit and S. Solomon for encouraging me to continue the study of MG methods for fermionic propagators.

Financial support by Deutsche Forschungsgemeinschaft is gratefully acknowledged.

The computations reported in this article were performed on the Siemens 7882 of the University of Hamburg, on the IBM 3090 of DESY, and on the CRAY Y-MP of HLRZ Jülich. The staff of these three Computing Centers provided useful advices.

Appendix A: Staggered Fermions

This appendix comprises a summary of Euclidean staggered fermions. It consists of two parts. In the first one the better-known spin-diagonalization technique is summarized, while the second part shows the connection to the Dirac-Kähler formalism.

A.1. Staggered Fermions from Spin-Diagonalization

We will mainly follow the position space treatment of Kluberg-Stern et al. [63], see also the review given in Ref. [21]. A momentum space treatment can be found in Ref. [62]. The notations and conventions used in this thesis are also summarized here.

In order to be concrete, let the dimension d of space-time be 4. The Euclidean time z_4 and the Minkowski time t are connected through $z_4 = it$. Euclidean γ-matrices are defined by $\gamma_4 = \gamma_0^M$, $\gamma_k = i\gamma_k^M$, $k = 1,2,3$, where the superscript ^{M} refers to Minkowski space. The Euclidean γ-matrices are Hermitean, $\gamma_\mu^\dagger = \gamma_\mu$, and fulfill the Clifford algebra $\{\gamma_\mu, \gamma_\nu\} = 2\,\delta_{\mu\nu}\mathbb{1}$, $\mu,\nu = 1,2,3,4$. We also define $\gamma_5 = \gamma_1\gamma_2\gamma_3\gamma_4$, and $\gamma_{-\mu} = -\gamma_\mu$.

The naive action for free Dirac fermions Ψ on a lattice Λ of lattice spacing a is

$$\mathcal{H}(\overline{\Psi},\Psi) = \int_z \frac{1}{2a}\overline{\Psi}(z)\sum_\mu \gamma_\mu\left[\Psi(z+e_\mu) - \Psi(z-e_\mu)\right] \;+\; m\int_z \overline{\Psi}(z)\,\Psi(z)\ . \tag{A.1}$$

The continuum derivative has to be approximated by the central difference quotient in order that the (anti-)Hermiticity properties of the action are preserved on the lattice.

The lattice integration $\int_z$ means $a^d\sum_{z\in\Lambda}$, and e_μ denotes a lattice vector of length a in μ-direction. It is convenient to define also lattice vectors for negative directions, $e_{-\mu} = -e_\mu$, and $e_0 = 0$.

The key step for the staggered formulation is a "spin-diagonalization" of the naive action [62,63]: One performs a unitary transformation of the Dirac field,

$$\Psi(z) = T(z)\,\Xi(z)\ ,\quad \overline{\Psi}(z) = \overline{\Xi}(z)\,T(z)^{-1}\ , \tag{A.2}$$

where the $T(z)$ are unitary $2^{d/2}\times 2^{d/2}$ matrices, chosen such that

$$T(z)^{-1}\,\gamma_\mu\,T(z+e_\mu) = \eta_\mu(z)\mathbb{1}\ . \tag{A.3}$$

$\eta_\mu(z)$ are complex numbers of modulus 1. These are not completely arbitrary. For instance the product of η_μ's around a plaquette must always equal -1, but the various possible

choices are all equivalent [63]. Since we defined $\gamma_{-\mu} = -\gamma_\mu$, it follows that $\eta_{-\mu}(z) = -\eta_\mu(z-e_\mu)$. One realization of (A.3) is ($z = a(n_1,n_2,n_3,n_4)$)

$$T(z) = \gamma_1^{n_1}\,\gamma_2^{n_2}\,\gamma_3^{n_3}\,\gamma_4^{n_4}\ , \tag{A.4}$$

yielding for the phase factors

$$\eta_1(z) = 1,\quad \eta_2(z) = (-1)^{n_1},\quad \eta_3(z) = (-1)^{n_1+n_2},\quad \eta_4(z) = (-1)^{n_1+n_2+n_3}\ . \tag{A.5}$$

In this representation $\eta_\mu(z)$ is independent of the μ-component of z, and hence $\eta_{-\mu}(z) = -\eta_\mu(z)$.

Through the transformation (A.2) the action (A.1) has become diagonal in the Dirac indices α. Originally, $\alpha = 1,\ldots,2^{d/2}$. Staggered fermions are defined by retaining only the action for one of the Ξ^α components. An important point is that the decoupling of the Ξ^α components also holds when gauge fields are introduced. The arbitrarily chosen retained component Ξ^α will be called χ. Hence, the staggered fermion action is

$$\mathcal{H}(\overline{\chi},\chi) = \int_z \frac{1}{a}\sum_\mu \eta_\mu(z)\,\overline{\chi}(z)\,\left[\chi(z+\tfrac{1}{2}e_\mu) - \chi(z-\tfrac{1}{2}e_\mu)\right] \;+\; m\int_z \overline{\chi}(z)\,\chi(z)\ . \tag{A.6}$$

In the staggered fermion formulation χ is not a field with a physical continuum limit. A (naive) continuum limit which yields Dirac fermions can only be taken for a field ψ which is a superposition of $\chi(z)$ and which is defined on a lattice that has twice the lattice spacing of Λ. We have already considered this fact in Eq. (A.6), were we redefined a to half its value. (e_μ remains to be a vector of length a.)

The lattice on which χ is defined will be called $\Lambda_{a/2}$. $\int_z$ in (A.6) means $(a/2)^d\sum_{z\in\Lambda_{a/2}}$. For a construction of the ψ field, divide $\Lambda_{a/2}$ into hypercubes $\tilde{y}$ consisting of 2^d sites z. Sites within one hypercube are distinguished by their "pseudoflavor" H, cf. Eqs. (3.1), (3.2), and Fig. 3.3. Hence, $\Lambda_{a/2}$ is made up of 2^d pseudoflavor lattices Λ_a^H whose sites have all the same pseudoflavor H and whose lattice spacing is a. We write $z\in\tilde{y}$ if z is a site in hypercube $\tilde{y}$. The lattice $\{\tilde{y}\}$ will be denoted by Λ_a.

In terms of crystallography as used in solid state physics [113], the lattice of staggered fermions is a hypercubic lattice Λ_a of unit cells of lattice spacing a with a 2^d-point basis. The basis is given by the sites of a hypercube of volume $(a/2)^d$. Sites within the basis are distinguished by their "pseudoflavor" H. Different pseudoflavors would correspond to different species of atoms or ions in the solid state analog.

$\psi(\tilde{y})$ will be a linear combination of $\chi(z)$ with $z\in\tilde{y}$. It is defined by

$$\psi_\alpha^f(\tilde{y}) = 2^{-3d/2}\sum_{z\in\tilde{y}}\Gamma_{\alpha f}(z)\chi(z)\ ,\quad \overline{\psi}_\alpha^f(\tilde{y}) = 2^{-3d/2}\sum_{z\in\tilde{y}}\overline{\chi}(z)\Gamma_{\alpha f}^\dagger(z)\ . \tag{A.7}$$

The indices α and f take values $1,\ldots,2^{d/2}$. $\Gamma_{\alpha f}$ are unitary $2^{d/2}\times 2^{d/2}$ matrices which shall fulfill

$$\Gamma(z+e_\mu) = \Gamma(z)\ ,\quad \text{and}\quad \mathrm{Tr}\left[\Gamma(z_1)^\dagger\Gamma(z_2)\right] = \begin{cases} 2^{d/2} & \text{if } H(z_1) = H(z_2),\\ 0 & \text{otherwise.}\end{cases} \tag{A.8}$$

$H(z)$ denotes the pseudoflavor of z. Because of (A.8) the transformation (A.7) becomes invertable:

$$\chi(z) = 2^{d/4} \sum_{\alpha,f} \Gamma^\dagger_{f\alpha}(z)\, \psi^f_\alpha(\tilde{y}) \ , \quad \overline{\chi}(z) = 2^{d/4} \sum_{\alpha,f} \overline{\psi}^f_\alpha(\tilde{y})\, \Gamma_{\alpha f}(z) \ , \tag{A.9}$$

for $z \in \tilde{y}$. A possible representation for $\Gamma(z)$ is the same as for $T(z)$, Eq. (A.4). Inserting Eq. (A.9) into the action (A.6) yields

$$\begin{aligned} \mathcal{H}(\overline{\psi},\psi) \ &= \ \int_{\tilde{y}} \overline{\psi}(\tilde{y}) \sum_\mu \Big[(\gamma_\mu \otimes \mathbb{1}) \partial_\mu \psi(\tilde{y}) + \frac{a}{2} (\gamma_6 \otimes \gamma_\mu \gamma_6)(\partial\partial)_\mu \psi(\tilde{y}) \Big] \\ &\quad + m \int_{\tilde{y}} \overline{\psi}(\tilde{y}) (\mathbb{1} \otimes \mathbb{1}) \psi(\tilde{y}) \ . \end{aligned} \tag{A.10}$$

Here $\int_{\tilde{y}} = a^d \sum_{\tilde{y}}$, ∂_μ and $(\partial\partial)_\mu$ are first and second central derivatives on Λ_a: $\partial_\mu \psi(\tilde{y}) = a^{-1}[\psi(\tilde{y}+e_\mu) - \psi(\tilde{y}-e_\mu)]$, and $(\partial\partial)_\mu \psi(\tilde{y}) = a^{-2}[\psi(\tilde{y}+e_\mu) - 2\psi(\tilde{y}) + \psi(\tilde{y}-e_\mu)]$.

Formula (A.10) is the basis for the interpretation of staggered fermions [63]. The direct products of matrices are understood such that the first matrix acts on Greek indices and the second one on Latin indices. Greek indices are interpreted as spinorial indices, while Latin indices are regarded as labeling flavor degrees of freedom. In the naive continuum limit the action (A.10) describes $2^{d/2}$ degenerate flavors of Dirac fermions. The flavors are coupled by the second term in (A.10), even if this term vanishes for $a \to 0$.

The free propagator of ψ reads in momentum space:

$$S(p) = \frac{\frac{i}{a} \sum_\mu \Big[(\gamma_\mu \otimes \mathbb{1}) \sin(p_\mu a) + \frac{2}{a} \sin^2\Big(\frac{p_\mu a}{2}\Big) (\gamma_6 \otimes \gamma_\mu \gamma_6) \Big] \ + \ m(\mathbb{1} \otimes \mathbb{1})}{\frac{4}{a^2} \sum_\mu \sin^2\Big(\frac{p_\mu a}{2}\Big) + m^2} \ . \tag{A.11}$$

The denominator in (A.11) has only one pole in the first Brillouin zone. In the naive continuum limit $a \to 0$ one has

$$S(p) \underset{a\to 0}{\longrightarrow} \frac{i \sum_\mu (\gamma_\mu \otimes \mathbb{1}) p_\mu + \ m(\mathbb{1} \otimes \mathbb{1})}{p^2 + m^2} \ , \tag{A.12}$$

which is the correct propagator describing $2^{d/2}$ degenerate Dirac fermions.

It is instructive to compare with Wilson's formulation of lattice fermions [56]. The action for $2^{d/2}$ degenerate flavors of Wilson fermions reads as Eq. (A.10), except that the matrix $(\gamma_6 \otimes \gamma_\mu \gamma_6)$ in the second term in the sum over μ is replaced by $(\mathbb{1} \otimes \mathbb{1})$. Both terms vanish for $a \to 0$. But whereas for Wilson fermions axial symmetry is completely lost for finite a, staggered fermions preserve at least a nontrivial piece of the full chiral symmetry for $m = 0$. The action (A.10) with $m = 0$ is invariant under

$$\psi(\tilde{y}) \to e^{i\theta(\gamma_6 \otimes \gamma_6)} \psi(\tilde{y}) \ , \quad \overline{\psi}(\tilde{y}) \to \overline{\psi}(\tilde{y}) e^{i\theta(\gamma_6 \otimes \gamma_6)} \ ,$$

where θ is an arbitrary real parameter.

A.2. Staggered Fermions and the Dirac-Kähler Equation

The outline in this section follows Refs. [64,114]. In the Dirac-Kähler formalism the Dirac field (*in the continuum*) is described by a *differential form* Φ,

$$\Phi = \sum_H \varphi(y,H)\, dy^H \quad \text{with} \quad dy^H = dy^{\mu_1} \wedge \cdots \wedge dy^{\mu_h} \ , \tag{A.13}$$

where H is the ordered set of indices which specifies pseudoflavor, cf. Eq. (3.1). There are 2^d such sets H. It is understood that $dy^\emptyset = 1$. Φ may be decomposed into $2^{d/2}$ independent Dirac fields ψ^f, in such a way that iff Φ satisfies the Dirac-Kähler equation

$$(d - \delta + m)\Phi = 0 \ , \tag{A.14}$$

then the ψ^f are solutions of the Dirac equations

$$(\not{\partial} + m)\psi^f = 0 \ , \quad f = 1, \ldots, 2^{d/2} \ . \tag{A.15}$$

The symbols d and δ in Eq. (A.14) denote exterior differentiation and the codifferential operator, respectively.

In the multi-index notation one defines the matrices

$$\gamma^H \equiv \gamma^{\mu_1} \cdots \gamma^{\mu_h} \ , \quad \gamma^\emptyset = \mathbb{1} \ . \tag{A.16}$$

Note that these matrices are the same which were used in Appendix A.1 for a representation of the Γ matrices in (A.7)–(A.9). We have the equality $\Gamma(z) = \gamma^H$ when H is the pseudoflavor of z. One defines a matrix $Z = (Z_{f\alpha})$ of forms,

$$Z = 2^{d/4} \sum_H \left(\gamma^H\right)^\dagger dy^H \ . \tag{A.17}$$

The Dirac fields ψ^f are then defined by

$$\Phi = \sum_H \varphi(y,H)\, dy^H = \sum_{f,\alpha} \psi^f_\alpha(y) Z_{f\alpha} \ . \tag{A.18}$$

With the help of the definition of the Clifford product $\vee$ of forms, $dy^\mu \vee dy^\nu = dy^\mu \wedge dy^\nu + \delta^{\mu\nu}$, the Dirac-Kähler operator $(d - \delta)$ can be brought into the form $(d - \delta)\Phi = dy^\mu \vee \partial_\mu \Phi$. By explicit calculation one verifies that ψ^f, $f = 1, \ldots, 2^{d/2}$, satisfy the Dirac equation iff Φ solves the Dirac-Kähler equation.

Since the Dirac-Kähler equation is a differential geometric equation, its lattice transcription is straightforward [64,114]. The lattice Dirac-Kähler field becomes a staggered fermion field χ if one identifies $\chi(z) = \varphi(\tilde{y}, H)$, for $z = \tilde{y} + \frac{1}{2} e_H$, with e_H as in (3.2). (Compare Eq. (A.18) with (A.9).)

Appendix B: Kernels in "Optimal" Multigrid

This Appendix deals with kernels in MG algorithms. There are three parts: (i) a review of the kernels in the optimal MG algorithm, (ii) their Fourier representations (in the absence of gauge fields), and (iii) some tables giving numerical values.

B.1. Review of the Gawędzki-Kupiainen Kernels

In this appendix the choice of kernels in the "optimal" MG algorithm is reviewed. Equations for the ideal interpolation kernel $\mathcal{A}$, given an averaging kernel C, are derived. This $\mathcal{A}$ appeared for the first time in the works [76] of Gawędzki and Kupiainen. They gave rigorous proofs of the existence of the continuum limit of some lattice field theories without gauge fields, using block spin renormalization group methods. Mack [34] pointed out that the use of smooth interpolation kernels in numerical MG work will be essential for fighting CSD. He proposed the Gawędzki-Kupiainen kernels as a starting point. We will follow in our exposition mainly the script [154] of Timme.

Consider a bosonic matter field ϕ which is defined on a d-dimensional Euclidean lattice $\Lambda = \Lambda^0$ of lattice spacing a. Gauge fields are not yet included, although their inclusion is straightforward; cf. the remarks at the end of this section. Let $\mathcal{H}(\phi) = \frac{1}{2} < \phi, v^{-1}\phi > + \mathcal{V}(\phi)$ be the action. v is the free propagator. (In connection with the notations of Sec. 2, $v^{-1} = D$.) The partition function is

$$Z = \int d\mu_v(\phi)\, e^{-\mathcal{V}(\phi)} \ , \tag{B.1}$$

where $d\mu_v(\phi)$ denotes the normalized Gaussian measure [8] with covariance v,

$$d\mu_v(\phi) = \det(2\pi v)^{-1/2} \prod_{x \in \Lambda^0} [d\phi(x)]\ e^{-\frac{1}{2}<\phi, v^{-1}\phi>} \ . \tag{B.2}$$

A Gaussian block spin transformation is attained by inserting

$$1 = \int \prod_{x \in \Lambda^1} [d\Phi(x)\ \left(\tfrac{\kappa (L_b a)^d}{2\pi}\right)^{1/2}]\ \exp\left(-\frac{\kappa}{2}\|\Phi - C\phi\|^2\right) \tag{B.3}$$

into the partition function . Here Λ^1 is a block lattice of lattice spacing $L_b a$, Φ is called the block spin which is defined on Λ^1, and $\|\Phi - C\phi\|^2 = < \Phi - C\phi, \Phi - C\phi >$. C is a linear

averaging operator, with kernel (2.35), for instance. C associates a block spin $C\phi$ with every field ϕ. In the limit that the real parameter κ tends to infinity, (B.3) reduces to a δ block spin transformation,

$$1 = \int \prod_{x \in \Lambda^1} [d\Phi(x)\ \delta(\Phi(x) - (C\phi)(x))] \ . \tag{B.4}$$

The idea of the real space renormalization group approach to quantum field theory is to split the field ϕ on the original lattice into a low frequency part ψ and a high frequency part ζ and to integrate out the high frequency field ζ in order to obtain an effective action. ψ and ζ are called "background field" and "fluctuation field", respectively. ψ should be smooth and it is determined by the block spin Φ. The split of the field becomes

$$\phi = \mathcal{A}\Phi + \zeta \ . \tag{B.5}$$

$\mathcal{A}$ is an interpolation kernel which maps fields on Λ^1 into fields on Λ^0. By means of the split (B.5) the kinetic term of the action $\mathcal{H}$ becomes

$$\tfrac{1}{2} < \phi, v^{-1}\phi > = \tfrac{1}{2} < \mathcal{A}\Phi, v^{-1}\mathcal{A}\Phi > + \tfrac{1}{2} < \zeta, v^{-1}\zeta > + < v^{-1}\mathcal{A}\Phi, \zeta > \ . \tag{B.6}$$

It is an easy exercise to show that

$$\mathcal{Q}[\phi] = \tfrac{1}{2} < \phi, v^{-1}\phi > + \frac{\kappa}{2}\|\Phi - C\phi\|^2 \tag{B.7}$$

can be rewritten as

$$\mathcal{Q}[\phi] = \tfrac{1}{2}\|\Gamma^{-1/2}\phi - \kappa\Gamma^{1/2}C^*\Phi\|^2 + \tfrac{1}{2} < \Phi, u^{-1}\Phi > \ . \tag{B.8}$$

In Eq. (B.8) we defined the free propagator u of the block spin field Φ,

$$u = \left(\kappa\mathbb{1} - \kappa^2 C\,\Gamma\, C^*\right)^{-1} \ , \tag{B.9}$$

and the free propagator Γ of the fluctuation field ζ,

$$\Gamma = \left(v^{-1} + \kappa C^* C\right)^{-1} \ . \tag{B.10}$$

C^* denotes the adjoint of C.

Let us now require that $\mathcal{A}$ is determined such that $\mathcal{A}\Phi$ is the classical field configuration for the Hamiltonian $\mathcal{Q}$, i. e. $\phi = \mathcal{A}\Phi$ minimizes $\mathcal{Q}$. This immediately yields

$$\mathcal{A} = \kappa\,\Gamma\, C^* \ . \tag{B.11}$$

As a consequence

$$\mathcal{Q}[\phi = \mathcal{A}\Phi + \zeta] = \tfrac{1}{2} < \zeta, \Gamma^{-1}\zeta > + \tfrac{1}{2} < \Phi, u^{-1}\Phi > \ . \tag{B.12}$$

One also readily verifies,

$$\mathcal{A} = v\, C^* u^{-1} \ , \tag{B.13}$$

$$u = C\, v\, C^* + \kappa^{-1} \ , \tag{B.14}$$

$$v = \Gamma + \mathcal{A}\, u\, \mathcal{A}^* \ . \tag{B.15}$$

It follows that $C\mathcal{A} = \mathbb{1}$ and $u^{-1} = (CC^*)^{-1}Cv^{-1}\mathcal{A} = \mathcal{A}^*v^{-1}\mathcal{A}$ for $\kappa \to \infty$. In this limit the fluctuation field ζ has vanishing block averages, $C\zeta = 0$. Then, as a consequence of Eq. (B.13), the mixed term $< v^{-1}\mathcal{A}\Phi, \zeta >$ in (B.6) vanishes. Thus, there is no coupling between high and low frequencies in the free Hamiltonian.

Another property of the $\mathcal{A}$-kernel is

$$\int_x \mathcal{A}(z,x) = 1 \quad \text{for all } z \in \Lambda^0 \ , \tag{B.16}$$

where $\int_x = (L_b a)^d \sum_{x \in \Lambda^1}$. This equation can be proved by using the Fourier representation of $\mathcal{A}$, cf. Eq. (B.26).

Eq. (B.15) means that the split (B.5) of the field ϕ implies a split of the propagator v into a propagator Γ for the high frequency field ζ and a propagator $\mathcal{A}u\mathcal{A}^*$ for the low frequency field ψ. This is valid for any value of κ and it allows the application of the convolution formula for Gaussian measures,

$$d\mu_v(\phi) = d\mu_u(\Phi)\, d\mu_\Gamma(\zeta) \ . \tag{B.17}$$

This factorization may be checked with the help of the formula for the characteristic functional of a Gaussian measure [8], $\int d\mu_v(\phi)\ e^{i<J,\phi>} = e^{-\frac{1}{2}<J,vJ>}$.

In summary we have shown that the partition function (B.1) can be rewritten as

$$Z = \int d\mu_u(\Phi) \int d\mu_\Gamma(\zeta)\ e^{-\mathcal{V}(\mathcal{A}\Phi+\zeta)} \ . \tag{B.18}$$

This rewriting may be iterated [34]. It leads to a mapping of a (nearly) critical system on Λ^0 onto a noncritical system on the MG. Eq. (B.18) serves as a staring point for nonperturbative calculations in constructive quantum field theory [76,99,104,119]. For the success of these analytical works it is essential that effective Hamiltonians have good locality properties, to all orders in perturbation theory. Good locality properties follow from the decay properties of kernels. When C is the "piecewise-constant" kernel of Eq. (2.35), $\Gamma(z,w)$ decays exponentially with distance $|z-w|$, with decay length of order one block lattice spacing $L_b a$. This means that ζ is a field with an infrared cutoff (i. e. with a "mass"). The fall-off properties of Γ are shared by $\mathcal{A}(z,x)$ and by the effective Laplacian (plus mass term) u^{-1}. This can be proved by using their Fourier representations, and shifting the path of the momentum integration into the complex plane [76]. If the kernels decay exponentially, it will be said that C defines a good block spin. It is plausible that a good block spin in this sense is also essential in MG computations of gauge covariant propagators. In an interacting theory exponential decay of Γ and $\mathcal{A}$ is preserved to all orders of perturbation theory. In special cases, this has also been shown nonperturbatively [76,104].

We conclude with remarks on the consequences of introducing gauge fields. In $v^{-1} = (-\Delta + m^2)$ the Laplacian will become the gauge covariant Laplacian (2.2). The averaging kernel C may be given by the ground-state projection definition (2.37). It enjoys gauge covariance (2.33). This alone implies gauge covariance of Γ, u, and $\mathcal{A}$. All of the above equations remain valid and are gauge covariant. Translational invariance of the kernels is of course lost in the presence of gauge fields, but the fall-off properties are retained, see Sec. 3.5 and the tables in Appendix B.3.

B.2. Fourier Representations of Kernels

We will now quote the Fourier representations of integral kernels in the absence of gauge fields, for the δ block spin transformation ($\kappa = \infty$). In arbitrary pure gauges the kernels are given by the same formulas, multiplied with (arbitrary) parallel transporters from w_2 to w_1 for a kernel with arguments (w_1, w_2).

Notations.

We consider a d-dimensional hypercubic lattice $\Lambda = \Lambda^0$ of lattice spacing a and with sites z, with an extension of N sites in each direction. For functions on this N^d lattice we impose periodic boundary conditions. The momenta of the reciprocal lattice will be denoted by p. We superimpose a block lattice Λ^1 by blocking Λ^0 with an integer scale factor L_b. The momenta of the reciprocal block lattice are denoted by q.

We define lattice integrations

$$\int_z = a^d \sum_{z \in \Lambda^0} \qquad \text{and} \qquad \frac{1}{(2\pi)^d}\int_p = \frac{a^{-d}}{|\Lambda^0|} \sum_{p \in BZ} \ , \tag{B.19}$$

and analogously for the block lattice. The expressions in (B.19) reduce in the limit $a \to 0$ to the usual continuum integrals. Summation in the momentum integral is over all momenta in the first Brillouin zone (BZ).

The δ distributions are on the lattice

$$\delta(z_1 - z_2) = a^{-d}\, \delta_{z_1, z_2} \qquad \text{and} \qquad \delta(p_1 - p_2) = \frac{|\Lambda^0|}{(2\pi)^d}\, a^d\, \delta_{p_1, p_2} \ . \tag{B.20}$$

We will need the formulas

$$\frac{1}{(2\pi)^d}\int_z e^{ipz} = \delta(p) \qquad \text{and} \qquad \frac{1}{(2\pi)^d}\int_p e^{-ipz} = \delta(z) \ , \tag{B.21}$$

and Poisson's resummation formula [155] whose lattice version reads

$$\int_{z \in \Lambda^1} e^{iqz} \ = \ (2\pi)^d \sum_{n_1=0}^{L_b-1} \cdots \sum_{n_d=0}^{L_b-1} \delta(q + \frac{2\pi}{L_b a} n) \ , \tag{B.22}$$

where $\delta(q + \frac{2\pi}{L_b a} n) = \prod_{\mu=1}^d \delta(q_\mu + \frac{2\pi}{L_b a} n_\mu)$.

Integral Kernels.

With these preparations one finds the Fourier transformation of the kernel C of Eq. (2.35),

$$C(x,z) = \frac{1}{(2\pi)^d}\int_p \tilde{C}(p)\, e^{-ip(x-z)} \qquad \text{with} \qquad \tilde{C}(p) = \alpha(p) \prod_\mu \frac{\sin\left(\frac{p_\mu L_b a}{2}\right)}{L_b \sin\left(\frac{p_\mu a}{2}\right)} \ , \tag{B.23}$$

where $\alpha(p) = 1$ for odd L_b, and $\alpha(p) = \exp(-i\sum_\mu p_\mu a/2)$ otherwise.

The Klein-Gordan operator $a^{-2}D = -\Delta + m^2$ has the Fourier representation

$$D(z_1, z_2) = \frac{1}{(2\pi)^d}\int_p \widetilde{D}(p)\, e^{-ip(z_1 - z_2)} \qquad \text{with} \qquad \widetilde{D}(p) = \frac{4}{a^2}\sum_\mu \sin^2\left(\frac{p_\mu a}{2}\right) + m^2 \ . \tag{B.24}$$

The inverse of the optimal effective Klein-Gordan operator D_{eff} is given by (cf. (B.14), $u \equiv D_{\mathit{eff}}^{-1}$)

$$D_{\mathit{eff}}^{-1}(x,y) = \frac{1}{(2\pi)^d}\int_q \tilde{u}(q)\, e^{-iq(x-y)} \qquad \text{with} \qquad \tilde{u}(q) = \sum_{n_1=0}^{L_b-1}\cdots\sum_{n_d=0}^{L_b-1} \frac{|\tilde{C}(q+\frac{2\pi}{L_b a}n)|^2}{\widetilde{D}(q+\frac{2\pi}{L_b a}n)} \ . \tag{B.25}$$

For the optimal interpolation kernel $\mathcal{A}$ one finds

$$\begin{aligned} \mathcal{A}(z,x) &= \frac{1}{(2\pi)^d}\int_q \tilde{\mathcal{A}}(z,q)\, e^{-iq(z-x)} \\ &\text{with} \qquad\qquad (B.26) \\ \tilde{\mathcal{A}}(z,q) &= \frac{1}{\tilde{u}(q)} \sum_{n_1=0}^{L_b-1}\cdots\sum_{n_d=0}^{L_b-1} \frac{\tilde{C}(q+\frac{2\pi}{L_b a}n)}{\widetilde{D}(q+\frac{2\pi}{L_b a}n)}\, e^{-i2\pi nz/(L_b a)} \ . \end{aligned}$$

With the help of Eqs. (B.23) and (B.26) one can explicitly verify that $(C\mathcal{A})(x,y) = \delta(x-y)$. One also readily verifies Eq. (B.16).

Finally we quote the kernel of the Galerkin operator CDC^*:

$$\begin{aligned} (CDC^*)(x,y) &= \frac{1}{(2\pi)^d}\int_q \widetilde{CDC^*}(q)\, e^{-iq(x-y)} \\ &\text{with} \qquad\qquad (B.27) \\ \widetilde{CDC^*}(q) &= \sum_{n_1=0}^{L_b-1}\cdots\sum_{n_d=0}^{L_b-1} \widetilde{D}(q+\frac{2\pi}{L_b a}n)|\tilde{C}(q+\frac{2\pi}{L_b a}n)|^2 \ . \end{aligned}$$

Note the distinction between the Fourier transforms $\tilde{u}^{-1}$ and $\widetilde{CDC^*}$ of the optimal coarse grid operator and of the Galerkin operator.

By using the Fourier representations given here one can easily compute the two-grid iteration matrix $\mathcal{M}$. It is easy to check that that $\mathcal{M}$ annihilates the slowest mode of the one-grid iteration matrix M, both in the optimal MG scheme and in case of variational MG.

B.3. Numerical Values of Integral Kernels

This section comprises a summary in tabular form of numerical values for integral kernels. Given are the optimal $\mathcal{A}$-kernel and the optimal effective Laplacian $C\Delta\mathcal{A} = \mathcal{A}^*\Delta\mathcal{A}$ on 9^4, 12^4, and 18^4 lattices in the absence of gauge fields. We also give the values for the "Ohrenkern" (which has support on one block and its eight nearest neighbors; cf. Sec. 6.5), and for the resulting effective Laplacian on a 12^4 lattice. The last two tables contain examples of an optimal effective gauge covariant Laplacian and a square of the Dirac operator in nontrivial gauge fields. The corresponding $\mathcal{A}$-kernels are shown in Fig. 3.8. All $\mathcal{A}$-kernels are "critical" kernels. They were computed as solution of Eq. (2.32) with $m^2 = m_{\mathrm{cr}}^2$, $\kappa = 100\,000$.

Optimal interpolation kernel $\mathcal{A}$ in the absence of gauge fields, 9^4 lattice.

−0.1638	−0.1910	−0.1638	−0.0441	−0.0103	−0.0038
−0.0625	−0.0688	−0.0625	0.0020	−0.0010	−0.0103
0.2432	0.3014	0.2432	0.1063	0.0020	−0.0441
1.1018	1.3063	1.1018	0.2432	−0.0625	−0.1638
1.3063	**1.5543**	1.3063	0.3014	−0.0688	−0.1910
1.1018	1.3063	1.1018	0.2432	−0.0625	−0.1638

Table B.1: *Optimal interpolation kernel $(L_b a)^4\mathcal{A}(z,0)$ in the absence of gauge fields on a 9^4 lattice for block size $L_b = 3$, at $z = (z_1, z_2, 1, 1)$ for $z_1, z_2 = 0, \ldots, 5$. The largest value is the framed one at $z = (1,1,1,1)$. The non-given values of $\mathcal{A}(z,x)$ follow from translational invariance and periodicity.*

Optimal interpolation kernel $\mathcal{A}$ in the absence of gauge fields, 12^4 lattice.

−0.0113	−0.0152	−0.0113	−0.0069	−0.0006	0.0017	0.0002	0.0000	0.0002
0.0179	0.0192	0.0179	0.0032	0.0026	0.0029	0.0000	−0.0005	0.0000
−0.0113	−0.0152	−0.0113	−0.0069	−0.0006	0.0017	0.0002	0.0000	0.0002
−0.1303	−0.1498	−0.1303	−0.0313	−0.0076	−0.0020	0.0017	0.0029	0.0017
−0.0718	−0.0777	−0.0718	−0.0028	−0.0038	−0.0076	−0.0006	0.0026	−0.0006
0.2246	0.2824	0.2246	0.0937	−0.0028	−0.0313	−0.0069	0.0032	−0.0069
1.1036	1.3124	1.1036	0.2246	−0.0718	−0.1303	−0.0113	0.0179	−0.0113
1.3124	**1.5660**	1.3124	0.2824	−0.0777	−0.1498	−0.0152	0.0192	−0.0152
1.1036	1.3124	1.1036	0.2246	−0.0718	−0.1303	−0.0113	0.0179	−0.0113

Table B.2: *Optimal interpolation kernel $(L_b a)^4\mathcal{A}(z,0)$ in the absence of gauge fields on a 12^4 lattice for block size $L_b = 3$, at $z = (z_1, z_2, 1, 1)$ for $z_1, z_2 = 0, \ldots, 8$. The largest value is the framed one at $z = (1,1,1,1)$.*

OPTIMAL INTERPOLATION KERNEL $\mathcal{A}$ IN THE ABSENCE OF GAUGE FIELDS, 18^4 LATTICE.

0.2267	0.2845	0.2267	0.0955	−0.0019	−0.0330	−0.0127	0.0016	0.0057
−0.0708	−0.0768	−0.0708	−0.0019	−0.0033	−0.0082	−0.0034	0.0014	0.0028
−0.1342	−0.1545	−0.1342	−0.0330	−0.0082	−0.0027	−0.0002	0.0015	0.0018
−0.0281	−0.0343	−0.0281	−0.0127	−0.0034	−0.0002	−0.0002	0.0002	0.0003
0.0091	0.0098	0.0091	0.0016	0.0014	0.0015	0.0002	−0.0001	−0.0002
0.0166	0.0188	0.0166	0.0057	0.0028	0.0018	0.0003	−0.0002	−0.0002
0.0015	0.0019	0.0015	0.0010	0.0003	−0.0001	0.0000	0.0000	0.0000
−0.0024	−0.0026	−0.0024	−0.0006	−0.0006	−0.0006	−0.0001	0.0000	0.0000
0.0015	0.0019	0.0015	0.0010	0.0003	−0.0001	0.0000	0.0000	0.0000
0.0166	0.0188	0.0166	0.0057	0.0028	0.0018	0.0003	−0.0002	−0.0002
0.0091	0.0098	0.0091	0.0016	0.0014	0.0015	0.0002	−0.0001	−0.0002
−0.0281	−0.0343	−0.0281	−0.0127	−0.0034	−0.0002	−0.0002	0.0002	0.0003
−0.1342	−0.1545	−0.1342	−0.0330	−0.0082	−0.0027	−0.0002	0.0015	0.0018
−0.0708	−0.0768	−0.0708	−0.0019	−0.0033	−0.0082	−0.0034	0.0014	0.0028
0.2267	0.2845	0.2267	0.0955	−0.0019	−0.0330	−0.0127	0.0016	0.0057
1.1034	1.3116	1.1034	0.2267	−0.0708	−0.1342	−0.0281	0.0091	0.0166
1.3116	1.5645	1.3116	0.2845	−0.0768	−0.1545	−0.0343	0.0098	0.0188
1.1034	1.3116	1.1034	0.2267	−0.0708	−0.1342	−0.0281	0.0091	0.0166

Table B.3: *Optimal interpolation kernel $(L_b a)^4 \mathcal{A}(z,0)$ in the absence of gauge fields on an 18^4 lattice for block size $L_b = 3$, at $z = (z_1, z_2, 1, 1)$ for $z_1 = 0, \ldots, 8$, and $z_2 = 0, \ldots, 17$. The largest value is the framed one at $z = (1,1,1,1)$.*

OPTIMAL EFFECTIVE LAPLACIAN Δ_{eff} IN THE ABSENCE OF GAUGE FIELDS, 3^4 BLOCK LATTICE.

−2.5691	0.0841	0.0841
−2.5691	0.0841	0.0841
17.8535	−2.5691	−2.5691

Table B.4: *Kernel $(L_b a)^6(-\Delta_{eff})(0,y)$, $y = (y_1, y_2, 0, 0)$ of the optimal effective Laplacian $\mathcal{A}^* \Delta \mathcal{A} = C \Delta \mathcal{A}$ in the absence of gauge fields on a 3^4 block lattice, obtained from blocking a 9^4 lattice with $L_b = 3$. $y_1, y_2 = 0, 1, 2$. The value of $(L_b a)^6(-\Delta_{eff})(0,0)$ is framed.*

OPTIMAL EFFECTIVE LAPLACIAN Δ_{eff} IN THE ABSENCE OF GAUGE FIELDS, 4^4 BLOCK LATTICE.

−2.9089	0.0192	0.0601	0.0192
0.6911	0.0601	−0.0214	0.0601
−2.9089	0.0192	0.0601	0.0192
18.2590	−2.9089	0.6911	−2.9089

Table B.5: *Kernel $(L_b a)^6(-\Delta_{eff})(0,y)$, $y = (y_1, y_2, 0, 0)$ of the optimal effective Laplacian $\mathcal{A}^* \Delta \mathcal{A} = C \Delta \mathcal{A}$ in the absence of gauge fields on a 4^4 block lattice, obtained from blocking a 12^4 lattice with $L_b = 3$. $y_1, y_2 = 0, 1, 2, 3$. The value of $(L_b a)^6(-\Delta_{eff})(0,0)$ is framed.*

OPTIMAL EFFECTIVE LAPLACIAN Δ_{eff} IN THE ABSENCE OF GAUGE FIELDS, 6^4 BLOCK LATTICE.

−2.8742	0.0336	0.0312	−0.0165	0.0312	0.0336
0.3515	0.0312	−0.0053	0.0010	−0.0053	0.0312
−0.0893	−0.0165	0.0010	0.0000	0.0010	−0.0165
0.3515	0.0312	−0.0053	0.0010	−0.0053	0.0312
−2.8742	0.0336	0.0312	−0.0165	0.0312	0.0336
18.2109	−2.8742	0.3515	−0.0893	0.3515	−2.8742

Table B.6: *Kernel $(L_b a)^6(-\Delta_{eff})(0,y)$, $y = (y_1, y_2, 0, 0)$ of the optimal effective Laplacian $\mathcal{A}^* \Delta \mathcal{A} = C \Delta \mathcal{A}$ in the absence of gauge fields on a 6^4 block lattice, obtained from blocking an 18^4 lattice with $L_b = 3$. $y_1, y_2 = 0, 1, 2, 3, 4, 5$. The value of $(L_b a)^6(-\Delta_{eff})(0,0)$ is framed.*

"OHRENKERN" IN THE ABSENCE OF GAUGE FIELDS.

−0.1070	−0.1262	−0.1070	0.0000	0.0000	0.0000
−0.0644	−0.0678	−0.0644	0.0000	0.0000	0.0000
0.2011	0.2706	0.2011	0.0000	0.0000	0.0000
1.1081	1.3256	1.1081	0.2011	−0.0644	−0.1070
1.3256	1.5877	1.3256	0.2706	−0.0678	−0.1262
1.1081	1.3256	1.1081	0.2011	−0.0644	−0.1070

Table B.7: *"Ohrenkern" $(L_b a)^4 \mathcal{A}(z,0)$ in the absence of gauge fields for block size $L_b = 3$, at $z = (z_1, z_2, 1, 1)$ for $z_1, z_2 = 0, \ldots, 5$. This interpolation kernel has support only on block $z = 0$ and its eight nearest neighbors. It is the solution of Eq. (5.1) with Δ replaced by the Laplacian with Dirichlet boundary conditions on the boundary of the desired support of $\mathcal{A}$.*

EFFECTIVE LAPLACIAN ASSOCIATED WITH THE "OHRENKERN" IN THE ABSENCE OF GAUGE FIELDS ON A 4^4 BLOCK LATTICE.

−2.8416	0.0342	0.0356	0.0342
0.5313	0.0356	0.0000	0.0356
−2.8416	0.0342	0.0356	0.0342
18.8755	−2.8416	0.5313	−2.8416

Table B.8: *Kernel $(L_b a)^6(-\Delta_{\mathit{eff}})(0,y)$, $y = (y_1, y_2, 0, 0)$ of the effective Laplacian $\mathcal{A}^*\Delta\mathcal{A}$ in the absence of gauge fields on a 4^4 block lattice, obtained from blocking a 12^4 lattice with $L_b = 3$. Here $\mathcal{A}$ is the kernel of Table B.7.*

AN OPTIMAL EFFECTIVE LAPLACIAN Δ_{eff} IN A FOUR-DIMENSIONAL $SU(2)$ GAUGE FIELD, $\beta = 2.7$, 6^4 BLOCK LATTICE.

2.2616	0.3173	0.0304	0.0171	0.0536	0.2360
0.2722	0.0495	0.0069	0.0017	0.0050	0.0288
0.0501	0.0187	0.0013	0.0004	0.0014	0.0116
0.3084	0.0272	0.0045	0.0014	0.0040	0.0353
2.4617	0.2733	0.0509	0.0105	0.0410	0.1610
14.2018	2.1237	0.2655	0.0387	0.2726	2.3101

Table B.9: *Kernel $(L_b a)^6 \|(-\Delta_{\mathit{eff}})(0,y)\|$, $y = (y_1, y_2, 0, 0)$ of an optimal effective Laplacian $\mathcal{A}^*\Delta\mathcal{A} = C\Delta\mathcal{A}$ in a four-dimensional $SU(2)$ gauge field at $\beta = 2.7$ on a 6^4 block lattice, obtained from blocking an 18^4 lattice with $L_b = 3$. $y_1, y_2 = 0, 1, 2, 3, 4, 5$.*

AN OPTIMAL EFFECTIVE DIRACIAN $\not{D}^2_{\mathit{eff}}$ IN A FOUR-DIMENSIONAL $SU(2)$ GAUGE FIELD, $\beta = 2.7$, 6^4 BLOCK LATTICE.

0.0000	0.6121	0.0000	0.1195	0.0000	0.4612
1.1497	0.0000	0.1096	0.0000	0.1348	0.0000
0.0000	0.1565	0.0000	0.0425	0.0000	0.1134
1.4623	0.0000	0.1722	0.0000	0.0711	0.0000
0.0000	0.2583	0.0000	0.1256	0.0000	0.2931
9.7798	0.0000	2.1071	0.0000	1.0994	0.0000

Table B.10: *Kernel $(L_b a)^6 \|(-\not{D}^2_{\mathit{eff}})(0,y)\|$, $y = (y_1, y_2, 0, 0)$ of an optimal effective squared Dirac operator $\mathcal{A}^*\not{D}^2\mathcal{A} = C\not{D}^2\mathcal{A}$ in a four-dimensional $SU(2)$ gauge field at $\beta = 2.7$ on a 6^4 block lattice, obtained from blocking an 18^4 lattice with $L_b = 3$. $y_1, y_2 = 0, 1, 2, 3, 4, 5$. Note the support property: the kernel is only nonvanishing on even lattice sites.*

References

[1] K.G. Wilson, *Confinement of quarks*, Phys. Rev. D10 (1974) 2445

[2] C. Itzykson and J.-B. Zuber, *Quantum field theory*, second printing (McGraw-Hill, Singapore, 1986)

[3] L.H. Ryder, *Quantum field theory* (Cambridge University Press, Cambridge, 1985)

[4] K. Symanzik, *Euclidean quantum field theory*, Varenna 1968, in: Local quantum theory, ed. R. Jost (Academic Press, New York, 1969)

[5] K. Osterwalder and R. Schrader, *Euclidean Fermi fields and a Feynman-Kac formula for boson-fermion models*, Helv. Phys. Acta 46 (1973) 277

[6] K. Osterwalder and E. Seiler, *Gauge field theories on a lattice*, Ann. Phys. 110 (1978) 440

[7] E. Seiler, *Gauge theories as a problem of constructive quantum field theory and statistical mechanics*, Lecture Notes in Physics 159 (Springer-Verlag, Berlin, 1982)

[8] J. Glimm and A. Jaffe, *Quantum physics*, second edition (Springer-Verlag, New York, 1987)

[9] K. Huang, *Statistical mechanics*, second edition (John Wiley & Sons, New York, 1987)

[10] J. Jersák, *Path integral method in quantum theory*, in: Path integral method, lattice gauge theory and critical phenomena, Proceedings of the Autumn College on Techniques in Many Body Problems, Lahore, Pakistan, 1987 (World Scientific, Singapore, 1989)

[11] G. Roepstorff, *Pfadintegrale in der Quantenphysik*, (F. Vieweg & Sohn, Braunschweig, 1991)

[12] M. Lüscher and P. Weisz, *Scaling laws and triviality bounds in the lattice ϕ^4 theory. I. One-component model in the symmetric phase*, Nucl. Phys. B290 [FS20] (1987) 25;

M. Lüscher and P. Weisz, *Scaling laws and triviality bounds in the lattice ϕ^4 theory. II. One-component model in the phase with spontaneous symmetry breaking*, Nucl. Phys. B295 [FS21] (1988) 65;

M. Lüscher and P. Weisz, *Scaling laws and triviality bounds in the lattice ϕ^4 theory. III. n-component model*, Nucl. Phys. B318 (1989) 705

[13] J.M. Hammersley and D.C. Handscomb, *Monte Carlo methods* (Methuen & Co Ltd, London 1965)

[14] K. Binder, *Monte Carlo investigations of phase transitions and critical phenomena*, in: Phase transitions and critical phenomena, Vol.5b, eds. C. Domb, M.S. Green (Academic Press, New York, 1976);

K. Binder (ed.), *Monte Carlo methods in statistical physics*, Topics in Current Physics 7, second edition (Springer-Verlag, Berlin, 1986);

K. Binder (ed.), *Applications of the Monte Carlo method in statistical physics*, Topics in Current Physics 36, second edition (Springer-Verlag, Berlin, 1987)

[15] M. Creutz, L. Jacobs, and C. Rebbi, *Monte Carlo computations in lattice gauge theories*, Phys. Rep. 95 (1983) 201

[16] D.W. Heermann, *Computer simulation methods in theoretical physics* (Springer-Verlag, Berlin, 1986)

[17] C. Rebbi, *Lattice gauge theories and Monte Carlo simulations* (World Scientific, Singapore, 1983)

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

[19] J.B. Kogut, *The lattice gauge theory approach to quantum chromodynamics*, Rev. Mod. Phys. 55 (1983) 775

[20] M. Creutz, *Quarks, gluons and lattices* (Cambridge University Press, Cambridge, 1983)

[21] H.J. Rothe, *Lattice gauge theories* (World Scientific, Singapore, 1992)

[22] R.H. Swendsen and J.-S. Wang, *Nonuniversal critical dynamics in Monte Carlo simulations*, Phys. Rev. Lett. 58 (1987) 86

[23] E. Ising, *Beitrag zur Theorie des Ferromagnetismus*, Z. Phys. 31 (1925) 253

[24] U. Wolff, *Collective Monte Carlo Updating for Spin Systems*, Phys. Rev. Lett. 62 (1989) 361

[25] M. Hasenbusch, *Improved estimators for a cluster updating of $O(N)$ spin models*, Nucl. Phys. B333 (1990) 581

[26] U. Wolff, *Critical slowing down*, Nucl. Phys. B (Proc. Suppl.) 17 (1990) 93;

A.D. Sokal, *How to beat critical slowing down: 1990 update*, Nucl. Phys. B (Proc. Suppl.) 20 (1991) 55

[27] G. Parisi, *Prolegomena to any future computer evaluation of the QCD mass spectrum*, Cargèse 1983, in: Progress in gauge field theory, eds. G.'t Hooft et al. (Plenum Press, New York, 1984)

[28] H. Meyer-Ortmanns, *Proposal of a new upgrading procedure for Monte Carlo experiments*, Z. Phys. C27 (1985) 553

[29] J. Goodman and A.D. Sokal, *Multigrid Monte Carlo for lattice field theories*, Phys. Rev. Lett. 56 (1986) 1015

[30] A. Hulsebos, J. Smit, and J.C. Vink, *Multigrid Monte Carlo in an external $U(1)$ gauge field*, Nucl. Phys. B (Proc. Suppl.) 9 (1989) 512;

A. Hulsebos, J. Smit, and J.C. Vink, *Multigrid Monte Carlo for a Bose field in an external gauge field*, Nucl. Phys. B331 (1990) 531

[31] A. Hulsebos, J. Smit, and J.C. Vink, *Multigrid method for the 2-D XY model*, Nucl. Phys. B (Proc. Suppl.) 17 (1990) 354;

A. Hulsebos, J. Smit, and J.C. Vink, *Multigrid simulation of the XY model*, Nucl. Phys. B356 (1991) 775

[32] R.G. Edwards, J. Goodman, and A.D. Sokal, *Multigrid Monte Carlo. 2. Two-dimensional XY model*, Nucl. Phys. B354 (1991) 289;

R.G. Edwards, S.J. Ferreira, J. Goodman, and A.D. Sokal, *Multigrid Monte Carlo. 3. Two-dimensional $O(4)$ symmetric nonlinear σ-model*, Nucl. Phys. B380 (1992) 621

[33] M.L. Laursen, J. Smit, and J.C. Vink, *Multigrid updating of $U(1)$ gauge fields*, Phys. Lett. B262 (1991) 467;

M.L. Laursen and J.C. Vink, *Multigrid updating of compact $U(1)$ gauge fields in 4 dimensions*, KFA Jülich preprint HLRZ-92-14 (March 1992)

[34] G. Mack, *Multigrid methods in quantum field theory*, Cargèse 1987, in: Nonperturbative quantum field theory, eds. G. 't Hooft et al. (Plenum Press, New York, 1988)

[35] K.G. Wilson, *Renormalization group and strong interactions*, Phys. Rev. D3 (1971) 1818;

K.G. Wilson and J. Kogut, *The renormalization group and the ϵ expansion*, Phys. Rep. 12C (1974) 75;

K.G. Wilson, *The renormalization group and critical phenomena*, Rev. Mod. Phys. 55 (1983) 583

[36] D.J. Amit, *Field theory, the renormalization group and critical phenomena* (World Scientific, Singapore, 1984)

[37] J.J. Binney, N.J. Dowrick, A.J. Fisher, and M.E.J. Newman, *The theory of critical phenomena. An introduction to the renormalization group* (Clarendon Press, Oxford, 1992)

[38] C. Gruber and H. Kunz, *General properties of polymer systems*, Commun. Math. Phys. 22 (1971) 133

[39] G. Mack, *Renormalization group and Mayer expansions*, Cargèse 1983, in: Progress in gauge field theory, eds. G. 't Hooft et al. (Plenum Press, New York, 1984)

[40] G. Mack and A. Pordt, *Convergent perturbation expansions for Euclidean quantum field theory*, Commun. Math. Phys. 97 (1985) 267

[41] G. Mack and S. Meyer, *The effective action from multigrid Monte Carlo*, Nucl. Phys. B (Proc. Suppl.) 17 (1990) 293

[42] G. Palma, *Finite size scaling analysis of the constraint effective potential computed from multigrid Monte Carlo*, preprint DESY 90-083 (July 1990);

G. Palma, *Renormalized loop expansion to compute finite size effects of the constraint effective potential*, Z. Phys. C54 (1992) 679;

G. Palma, *Renormalized loop expansion for the gauge covariant constraint effective Higgs potential on the lattice*, Ph. D. thesis and preprint DESY 92-031 (February 1990)

[43] M. Hasenbusch, S. Meyer, and G. Mack, *Noncritical multigrid Monte Carlo: $O(3)$ non-linear σ model*, Nucl. Phys. B (Proc. Suppl.) 20 (1991) 110

[44] M. Hasenbusch and S. Meyer, *Multigrid acceleration for asymptotically free theories*, Phys. Rev. Lett. 68 (1992) 435;

M. Hasenbusch and S. Meyer, *Multigrid acceleration for asymptotically free theories*, Nucl. Phys. B (Proc. Suppl.) 26 (1992) 610;

M. Hasenbusch and S. Meyer, *Testing accelerated algorithms in the lattice CP^3 model*, Phys. Rev. D45 (1992) 4376

[45] G. Mack and K. Pinn: *A Monte Carlo procedure for Hamiltonians with small nonlocal correction terms*, Phys. Lett. B173 (1986) 434

[46] T. Kalkreuter, *Simulation von Gitterfeldtheorien mit Hilfe der Polymermethode*, Diploma thesis, Universität Hamburg (April 1989)

[47] K. Symanzik, *Continuum limit and improved action in lattice theories, (I.) Principles and ϕ^4-theory*, Nucl. Phys. B226 (1983) 187;

K. Symanzik, *Continuum limit and improved action in lattice theories, (II.) O(N) non-linear sigma model in perturbation theory*, Nucl. Phys. B226 (1983) 205

[48] T. Kalkreuter, *Simulation of statistical systems with not necessarily real and positive probabilities*, Nucl. Phys. B352 (1991) 251

[49] I. Montvay, *Simulation of staggered fermions by polymer averaging*, Phys. Lett. B227 (1989) 260;

I. Montvay, *Simulation of staggered fermions by polymer algorithms*, Cargèse 1989, in: P.H. Damgaard et al. (eds.), Probabilistic methods in quantum field theory and quantum gravity (Plenum Press, New York, 1990)

[50] D. Weingarten, *Monte Carlo algorithms for QCD*, Nucl. Phys. B (Proc. Suppl.) 9 (1989) 447

[51] D. Touissaint, *The QCD spectrum and phase diagram on the lattice - 1991*, Nucl. Phys. B (Proc. Suppl.) 26 (1992) 3

[52] S. Duane, A.D. Kennedy, B.J. Pendleton, and D. Roweth, *Hybrid Monte Carlo*, Phys. Lett. B195 (1987) 216;

A.D. Kennedy, *Higher order Hybrid Monte Carlo*, Nucl. Phys. B (Proc. Suppl.) 9 (1989) 457;

A.D. Kennedy, *The theory of hybrid stochastic algorithms*, Cargèse 1989, in: Probabilistic methods in quantum field theory and quantum gravity, eds. P.H. Damgaard et al. (Plenum Press, New York, 1990)

[53] S.R. Sharpe, *Lattice QCD: Going beyond the mass spectrum*, J. Stat. Phys. 43 (1986) 1129

[54] A. Brandt, *Multigrid Methods in Lattice Field Computations*, Nucl. Phys. B (Proc. Suppl.) 26 (1992) 137

[55] R.C. Brower, K.J.M. Moriarty, E. Myers, and C. Rebbi, *The multigrid method for fermion calculations in quantum chromodynamics*, in: Multigrid Methods, ed. S.F. McCormick (Marcel Dekker, New York, 1988)

[56] K.G. Wilson, *Quarks and strings on a lattice*, Erice 1975, in: New phenomena in subnuclear physics, ed. A. Zichichi (Plenum Press, New York, 1977)

[57] H.B. Nielsen and M. Ninomiya, *Absence of neutrinos on a lattice. (I). Proof by homotopy theory*, Nucl. Phys. B185 (1981) 20, Errata ibid. B195 (1982) 541;

H.B. Nielsen and M. Ninomiya, *Absence of neutrinos on a lattice. (II). Intuitive topological proof*, Nucl. Phys. B193 (1981) 173;

H.B. Nielsen and M. Ninomiya, *A no-go theorem for regularizing chiral fermions*, Phys. Lett. B105 (1981) 219;

L.H. Karsten, *Lattice fermions in Euclidean space-time*, Phys. Lett. B104 (1981) 315

[58] J. Smit, *Fermions on a lattice*, Acta Phys. Pol. B17 (1986) 531;

J. Smit, *Attempts to put the standard model on the lattice*, Nucl. Phys. B (Proc. Suppl.) 4 (1988) 451;

J. Smit, *Standard model and chiral gauge theories on the lattice*, Nucl. Phys. B (Proc. Suppl.) *17 (1990) 3*

[59] J. Kogut and L. Susskind, *Hamiltonian formulation of Wilson's lattice gauge theory*, Phys. Rev. D11 (1975) 395

[60] T. Banks, L. Susskind, and J. Kogut, *Strong-coupling calculations of lattice gauge theories: (1+1) dimensional exercise*, Phys. Rev. D13 (1976) 1043

[61] L. Susskind, *Lattice fermions*, Phys. Rev. D16 (1977) 3031

[62] H. Sharatchandra, H. Thun, and P. Weisz, *Susskind fermions on a Euclidean lattice*, Nucl. Phys. B192 (1981) 205

[63] H. Kluberg-Stern, A. Morel, O. Napoly, and P. Petersson, *Flavours of Lagrangian Susskind fermions*, Nucl. Phys. B220 [FS8] (1983) 447

[64] P. Becher and H. Joos, *The Dirac-Kähler equation and fermions on the lattice*, Z. Phys. C15 (1982) 343

[65] Y. Oyanagi, *An incomplete LDU decomposition of lattice fermions and its application to conjugate reidual methods*, Comput. Phys. Commun. 42 (1986) 333

[66] P. Rossi, C.T.H. Davies, and G.P. Lepage, *A comparison of a variety of matrix inversion algorithms for Wilson fermions on the lattice*, Nucl. Phys. B297 (1988) 287

[67] T.A. DeGrand, *A conditioning technique for matrix inversion for Wilson fermions*, Comput. Phys. Commun. 52 (1988) 161

[68] G. Katz, G. Batrouni, C. Davies, A. Kronfeld, P. Lepage, P. Rossi, B. Svetitsky, and K. Wilson, *Fourier acceleration in lattice gauge theories. II. Matrix inversion and the quark propagator*, Phys. Rev. D37 (1988) 1589

[69] P. Rossi and C.T.H. Davies, *A comment on an attempt to Fourier accelerate Kogut-Susskind fermions*, Phys. Lett. B202 (1988) 547

[70] R.G. Edwards, J. Goodman and A.D. Sokal, *Multigrid method for the random resistor problem*, Phys. Rev. Lett. 61 (1988) 1333;

R.G. Edwards, J. Goodman and A.D. Sokal, *Algebraic Multi-grid method for disordered systems*, Nucl. Phys. B (Proc. Suppl.) 9 (1989) 521

[71] R.G. Edwards, *Numerical simulations in lattice gauge theories and statistical mechanics*, Ph. D. thesis, New York University (June 1989)

[72] R. Ben-Av, A. Brandt, and S. Solomon, *The fermionic matrix, instantons, zero modes and multigrid*, Nucl. Phys. B329 (1990) 193

[73] A. Hulsebos, J. Smit, and J.C. Vink, *Multigrid inversion of the staggered fermion matrix*, Nucl. Phys. B (Proc. Suppl.) 20 (1991) 94

[74] R.C. Brower, C. Rebbi, and E. Vicari, *Projective multigrid method for propagators in lattice gauge theory*, Phys. Rev. D43 (1991) 1965

[75] G. Mack, *unpublished*

[76] K. Gawędzki and A. Kupiainen, *A rigorous block spin approach to massless lattice theories*, Commun. Math. Phys. 77 (1980) 31;

K. Gawędzki and A. Kupiainen, *Massless lattice ϕ^4_4 theory: Rigorous control of a renormalizable asymptotically free model*, Commun. Math. Phys. 99 (1985) 197

[77] P.B. Mackenzie, *Critique of practical methods for computer evaluation of the hadron spectrum*, Nucl. Phys. B (Proc. Suppl.) 17 (1990) 103

[78] K.M. Bitar, R.G. Edwards, U.M. Heller, A.D. Kennedy, T.A. DeGrand, S. Gottlieb, A. Krasnitz, J.B. Kogut, W. Liu, P. Rossi, M.C. Ogilvie, R.L. Renken, D.K. Sinclair, R.L. Sugar, D. Toussaint, and K.C. Wang, *Hadron spectrum in QCD with valence Wilson fermions and dynamical staggered fermionsat* $6/g^2 = 5.6$, Phys. Rev. D46 (1992) 2169

[79] J. Goodman and A.D. Sokal, *Multigrid Monte Carlo method. Conceptual foundations*, Phys. Rev. D40 (1989) 2035

[80] R.S. Varga, *Matrix iterative analysis*, (Prentice-Hall, Inc., Englewood Cliffs, NJ, 1962)

[81] D.M. Young, *Iterative solution of large linear systems* (Academic Press, New York, 1971)

[82] J. Stoer and R. Bulirsch, *Einführung in die Numerische Mathematik II* (Springer-Verlag, Berlin, 1978)

[83] G.H. Golub and C.F. v. Loan, *Matrix computations*, second edition, (The Johns Hopkins University Press, Baltimore, 1990)

[84] W. Hackbusch, *Iterative Lösung großer schwachbesetzter Gleichungssysteme* (B. G. Teubner, Stuttgart, 1991)

[85] C.B. Chalmers, R.D. Kenway, and D. Roweth, *Algorithms for calculating quark propagators on large lattices*, J. Comp. Phys. 70 (1987) 500

[86] D. Henty, R. Setoodeh, and C.T.H. Davies, *A study of block algorithms for fermions matrix inversion*, Nucl. Phys. B337 (1990) 487

[87] A.N. Burkitt and A.C. Irving, *Inversion of the fermion matrix and the equivalence of the conjugate gradient and Lanczos algorithms*, Comput. Phys. Commun. 59 (1990) 447

[88] A.D. Kennedy, *Hybrid Monte Carlo algorithm on the connection machine*, Int. J. Mod. Phys. C3 (1992) 1

[89] S. Aoki, R. Shrock, B. Berg, C. Bernard, M. Ogilvie, D. Petcher, G. Bhanot, P. Rossi, K. Bitar, R. Edwards, U.M. Heller, A. Kennedy, S. Sanielevici, R. Brower, J. Potvin, C. Rebbi, F.R. Brown, N. Christ, R. Mawhinney, C. Detar, T. Draper, K.-F. Liu, S. Gottlieb, H. Hamber, G. Kilcup, J. Shigemitsu, J. Kogut, A. Kronfeld, I.-H. Lee, J. Negele, S. Ohta, J.C. Sexton, E. Shuryak, D.K. Sinclair, A. Soni, and W. Wilcox, *Physics goals of the QCD teraflop project*, Int. J. Mod. Phys. C2 (1991) 829

[90] F.R. Brown, H. Chen, N.H. Christ, Z. Dong, R.D. Mawhinney, W. Schaffer, and A. Vaccarino, *Lattice QCD with 8 light quark flavors*, Columbia University preprint CU-TP-541 (May 1992)

[91] C.R. Allton, H. Duong, C.T. Sachrajda, I.M. Barbour, A. Bell, C.T.H. Davies, S.J. Hands, E. Klepfish, R.M. Baxter, S.P. Booth, K.C. Bowler, S. Collins, D.S. Henty, R.D. Kenway, C. McNeile, B.J. Pendleton, D.G. Richards, J.N. Simone, A.D. Simpson, B.E. Wilkes, J.E.J. Bottomley, S.M. Catterall, F.R. Devlin, I.T. Drummond, R.R. Horgan, K. Decker, A. Hulsebos, A.C. Irving, A. McKerrell, C. Michael, M. Prisznyak, P.W. Stephenson, G. Martinelli, M. Teper, and J.F. Wheater, *First results from the UKQCD collaboration*, Nucl. Phys. B (Proc. Suppl.) 26 (1992) 211

[92] E. Polak, *Computational methods in optimization* (Academic Press, New York, 1971)

[93] W.H. Press, B.P. Flannery, S.A. Teukolsky, and W.T. Vetterling, *Numerical recipes* (Cambridge University Press, Cambridge, 1986)

[94] J.H. Wilkinson and C. Reinsch, *Linear algebra* (Springer-Verlag, Berlin, 1971)

[95] A. Brandt, *Multi-level adaptive solutions to boundary-value problems*, Math. Comput. 31 (1977) 333;

K. Stüben and U. Trottenberg, *Multigrid methods: fundamental algorithms, model problem analysis and applications*, in: Multigrid methods, Lecture Notes in Mathematics 960, eds. W. Hackbusch and U. Trottenberg (Springer-Verlag, Berlin, 1982);

A. Brandt, *Multigrid techniques: 1984 guide with applications to Fluid Dynamics*, GMD-Studie No. 85

[96] W. Hackbusch, *Multi-grid methods and applications*, Springer Series in Computational Mathematics 4 (Springer-Verlag, Berlin, 1985)

[97] S.F. McCormick, *Multilevel adaptive methods for partial differential equations* (SIAM, Philadelphia, 1989)

[98] W.L. Briggs, *A multigrid tutorial* (SIAM, Philadelphia, 1987)

[99] K. Gawędzki and A. Kupiainen, *Asymptotic freedom beyond perturbation theory*, in: Critical phenomena, random systems, gauge theories, Les Houches 1984, eds. K. Osterwalder and R. Stora (North Holland, Amsterdam, 1986)

[100] M. Harmatz, P.G. Lauwers, R. Ben-Av, A. Brandt, E. Katznelson, S. Solomon, and K. Wolowesky, *Parallel-transported multigrid and its application to the Schwinger model*, Nucl. Phys. B (Proc. Suppl.) 20 (1991) 102

[101] R. Ben-Av, *Acceleration methods in MC simulations*, Ph. D. thesis, The Weizmann Institute of Science (October 1991)

[102] R. Ben-Av, M. Harmatz, P.G. Lauwers, and S. Solomon, *Parallel-transported multigrid for inverting the Dirac-operator: Variants of the method and their efficiency*, Arbeitspapiere der GMD 674 (August 1992)

[103] G. Mack, *Dielectric lattice gauge theory*, Nucl. Phys. B235 [FS11] (1984) 197

[104] T. Balaban, *Regularity and decay of lattice Green's functions*, Commun. Math. Phys. 89 (1983) 571

[105] T. Kalkreuter, *Projective block spin transformations in lattice gauge theory*, Nucl. Phys. B376 (1992) 637

[106] T. Kalkreuter, *Improving multigrid and conventional relaxation algorithms for propagators*, preprint DESY 92-108 (July 1992), to appear in Int. J. Mod. Phys. C

[107] A. Hulsebos, J. Smit, and J.C. Vink, *Multigrid inversion of lattice fermion operators*, Nucl. Phys. B368 (1992) 379

[108] A. Hulsebos, J. Smit, and J.C. Vink, *Multigrid inversion of the staggered fermion matrix with $U(1)$ and $SU(2)$ gauge fields*, Int. J. Mod. Phys. C3 (1992) 161

[109] J.C. Vink, *Multigrid inversion of staggered and Wilson fermion operators with $SU(2)$ gauge fields in two dimensions*, Phys. Lett. B272 (1991) 81

[110] J.C. Vink, *Multigrid inversion of fermion operators with $SU(2)$ gauge fields in two and four dimensions*, Nucl. Phys. B (Proc. Suppl.) 26 (1992) 607

[111] R.C. Brower, C. Rebbi, and E. Vicari, *Non-Abelian projective multigrid for lattice gauge theory*, Phys. Rev. Lett. 66 (1991) 1263;

R.C. Brower, K.J.M. Moriarty, C. Rebbi, and E. Vicari, *Variational multigrid for non-Abelian gauge theory*, Nucl. Phys. B (Proc. Suppl.) 20 (1991) 89

[112] T. Kalkreuter, G. Mack, and M. Speh, *Blockspin and multigrid for staggered fermions in non-Abelian gauge fields*, Int. J. Mod. Phys. C3 (1992) 121

[113] N.W. Ashcroft and N.D. Mermin, *Solid state physics* (Holt-Saunders Int. Ed., Tokyo, 1976)

[114] H. Joos and M. Schaefer, *The representation theory of the symmetry group of lattice fermions as a basis for kinematics in lattice QCD*, Z. Phys. C34 (1987) 465

[115] M.F.L. Golterman, *Irreducible representations of the staggered fermion symmetry group*, Nucl. Phys. B278 (1986) 417

[116] G. Mai, *Ein Blockspin für $2^{d/2}$ Fermionen*, Diploma thesis, Universität Hamburg (October 1989)

[117] R. Ben-Av, A. Brandt, M. Harmatz, E. Katznelson, P.G. Lauwers, S. Solomon, and K. Wolowesky, *Fermion simulations using parallel transported multigrid*, Phys. Lett. B253 (1991) 185;

P.G. Lauwers, R. Ben-Av, and S. Solomon, *Inverting the Dirac matrix for $SU(2)$ lattice gauge theory by means of parallel-transported multigrid*, Nucl. Phys. B374 (1992) 249;

S. Solomon and P.G. Lauwers, *Parallel transported multigrid beats conjugate gradient*, Int. J. Mod. Phys. C3 (1992) 149

[118] P.G. Lauwers and T. Wittlich, *Inversion of the fermion matrix in lattice QCD by means of parallel-transported multigrid (PTMG)*, preprint BONN 92–29 (September 1992);

P.G. Lauwers and T. Wittlich, *Parallel-transported multigrid (PTMG) for inverting the Dirac-operator in $SU(3)$ lattice gauge theory*, to appear in Nucl. Phys. B (Proc. Suppl.) (LATTICE 92)

[119] T. Bałaban, *Propagators and renormalization transformations for lattice gauge theories. I*, Commun. Math. Phys. 95 (1984) 17;

T. Bałaban, *Averaging operations for lattice gauge theories*, Commun. Math. Phys. 98 (1985) 17;

T. Bałaban, *Renormalization group approach to lattice gauge field theories: I. Generation of effective actions in a small field approximation and a coupling constant renormalization in four dimensions*, Commun. Math. Phys. 109 (1987) 249

[120] T. Bałaban and A. Jaffe, *Constructive gauge theory*, Erice 1985, in: Fundamental problems of gauge field theory, eds. G. Velo and A.S. Wightman (Plenum Press, New York, 1986)

[121] A.S. Wightman, *Constructive quantum field theory from I to II. Remarks on what we have learned*, in: Constructive quantum field theory II, eds. G. Velo and A.S. Wightman, Erice 1988 (Plenum Press, New York, 1990)

[122] J. Imbrie, *Renormalization group methods in gauge field theories*, in: Critical phenomena, random systems, gauge theories, Les Houches 1984, eds. K. Osterwalder and R. Stora (North Holland, Amsterdam, 1986)

[123] V. Rivasseau, *From perturbative to constructive renormalization*, Princeton Series in Physics (Princeton University Press, Princeton, 1991)

[124] L. Collatz, *Eigenwertaufgaben mit technischen Anwendungen* (Akademische Verlagsgesellschaft, Leipzig, 1963)

[125] L. Collatz, *Funktionalanalysis und Numerische Mathematik* (Springer-Verlag, Berlin, 1964)

[126] R. Courant and D. Hilbert, *Methoden der Mathematischen Physik I* (Springer-Verlag, Berlin, 1968)

[127] M. Creutz, *Monte Carlo study of quantized $SU(2)$ gauge theory*, Phys. Rev. D21 (1980) 2308

[128] NAGLIB, Numerical Algorithms Group, NAG Fortran (Vector) Library, Mark 13

[129] SIEMENS 7.882, Regionales Rechenzentrum der Universität Hamburg

[130] IBM 3084 Q, IBM 3090–150E, IBM ES 9000/720 VF (IBM 9021), DESY-Rechenzentrum

[131] M. Creutz, *Confinement and the critical dimensionality of space-time*, Phys. Rev. Lett. 43 (1979) 553

[132] CRAY Y-MP8/832, HLRZ at KFA Jülich

[133] H.J. Pirner, J. Wroldsen, and M. Ilgenfritz, *Construction of an effective colour dielectric lattice action from QCD by blocking*, Nucl. Phys. B294 (1987) 905;

H.J. Pirner, *Confinement in the color dielectric model*, Nucl. Phys. B (Proc. Suppl.) 15 (1990) 207;

B. Großmann, H.J. Pirner, A.I. Signal, R. Baier, and J. Wroldsen, *Searching for an optimal lattice color dielectric theory*, Int. J. Mod. Phys. A6 (1991) 2649;

B. Großmann, H.J. Pirner, and J. Wroldsen, *Phase structure of $SU(2)$ lattice gauge theory with fluctuating coupling*, Nucl. Phys. B377 (1992) 424;

H.J. Pirner, *The color dielectric model of QCD*, Prog. Part. Nucl. Phys. 29 (1992) 33

[134] D.J.E. Callaway and R. Petronzio, *Monte Carlo renormalization group trajectories in lattice gauge theories*, Phys. Lett. B145 (1984) 381

[135] A. Patel and R. Gupta, *Monte Carlo renormalization group investigations of the $SU(2)$ lattice gauge theory*, Nucl. Phys. B251 [FS13] (1985) 789

[136] T. Kalkreuter, *Ground-state projection multigrid for propagators in four-dimensional $SU(2)$ gauge fields*, Phys. Lett. B276 (1992) 485

[137] G. Mack, T. Kalkreuter, G. Palma, and M. Speh, *Effective Field Theories*, preprint DESY 92–070 (May 1992), to appear in the proceedings of the 31st IUKT, Schladming, February 1992

[138] M. Bäker, T. Kalkreuter, G. Mack, and M. Speh, *Neural multigrid for gauge theories and other disordered systems*, preprint DESY 92–126 (September 1992), to appear in the proceedings of the 4th International Conference on Computational Physics PC92, Prague, August 1992

[139] J. Hertz, A. Krogh, and R.G. Palmer, *Introduction to the theory of neural computation* (Addison-Wesley, Redwood City CA, 1991)

[140] G. Mack and M. Speh, *in preparation*

[141] M. Bäker, Diploma thesis, *in preparation*

[142] R.C. Brower, K.J.M. Moriarty, C. Rebbi, and E. Vicari, *Multigrid propagators in the presence of disordered $U(1)$ gauge fields*, Phys. Rev. D43 (1991) 1974

[143] V. Vyas, *A multigrid algorithm for calculating fermionic propagators using Migdal-Kadanoff transformation*, Wuppertal preprint WUB 91–10 (February 1991);

V. Vyas, *Calculating the quark propagators using the Migdal-Kadanoff transformation*, Int. J. Mod. Phys. C3 (1992) 169

[144] R.C. Brower, R.G. Edwards, C. Rebbi, and E. Vicari, *Projective multigrid for Wilson fermions*, Nucl. Phys. B366 (1991) 689

[145] A. Hulsebos, *Multigrid methods in quantum field theories*, Ph.D. thesis, Universiteit van Amsterdam (September 1991)

[146] V. Vyas, *An efficient algorithm for calculating the quark propagators using Migdal-Kadanoff transformation*, Wuppertal preprint WUB 92–30 (September 1992)

[147] K.M. Bitar, R.G. Edwards, T.A. DeGrand, S. Gottlieb, U.M. Heller, A.D. Kennedy, J.B. Kogut, A. Krasnitz, W. Liu, M.C. Ogilvie, R.L. Renken, D.K. Sinclair, R.L. Sugar, D. Toussaint, and K.C. Wang, *QCD hadron spectroscopy with staggered dynamical quarks at $\beta = 5.6$*, Nucl. Phys. B (Proc. Suppl.) 26 (1992) 259

[148] Ph. de Forcrand and R. Gupta, *Multigrid techniques for quark propagators*, Nucl. Phys. B (Proc. Suppl.) 9 (1989) 516

[149] A. Hulsebos, M.L. Laursen, and J. Smit, *$SU(N)$ multigrid Landau gauge fixing*, Phys. Lett. B291 (1992) 431

[150] M. Grabenstein and K. Pinn, *Acceptance rates in multigrid Monte Carlo*, Phys. Rev. D45 (1992) 4372;

M. Grabenstein and K. Pinn, *Kinematics of multigrid Monte Carlo*, preprint DESY 92–094 (July 1992)

[151] R. Cordery, R. Gupta, and M.A. Novotny, *An improved renormalization group transformation in four dimensions*, Phys. Lett. B128 (1983) 425

[152] T. Kalkreuter, *Multigrid for staggered fermions in four-dimensional $SU(2)$ gauge fields*, talk presented at the DFG-Colloquium "Dynamical Fermions" held in Leipzig, March 1992

[153] T. Kalkreuter, *Multigrid for propagators of staggered fermions in four-dimensional $SU(2)$ gauge fields*, to appear in Nucl. Phys. B (Proc. Suppl.) (LATTICE 92)

[154] H.-J. Timme, *Blockspintransformationen*, unpublished (1987)

[155] L.M. Gelfand and G.E. Schilow, *Verallgemeinerte Funktionen (Distributionen) I* (VEB Deutscher Verlag der Wissenschaften, Berlin, 1967)