DEUTSCHES ELEKTRONEN – SYNCHROTRON

DESY 92-185
December 1992

Towards a Multigrid Scheme in SU(2) Lattice Gauge Theory

F. Gutbrod

Deutsches Elektronen-Synchrotron DESY, 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 (if possible by air mail):

DESY
Bibliothek
Notkestraße 85
W-2000 Hamburg 52
Germany

DESY-IfH
Bibliothek
Platanenallee 6
O-1615 Zeuthen
Germany

DESY 92-185
December 1992

ISSN 0418-9833

Towards a Multigrid Scheme in SU(2) Lattice Gauge Theory

F. Gutbrod
Deutsches Elektronen-Synchrotron DESY
2000 Hamburg 52, Notkestr. 85, Germany

Abstract

The task of constructing a viable updating multigrid scheme for SU(2) lattice gauge theory is discussed in connection with the classical eigenvalue problem. For a nonlocal overrelaxation Monte Carlo update step, the central numerical problem is the search for the minimum of a quadratic approximation to the action under nonlocal constraints. Here approximate eigenfunctions are essential to reduce the numerical work, and these eigenfunctions are to be constructed with multigrid techniques. A simple implementation on asymmetric lattices is described, where the grids are restricted to 3-dimensional hyperplanes. The scheme is shown to be moderately successful in the early stages of the updating history (starting from a cold configuration). The main results of another, less asymmetric scheme are presented briefly.

1 Introduction

Excellent discussions on the elimination of critical slowing down (CSD) in lattice field theories have been given several years ago [1,2], but there seems to be not a single successful application for nonabelian lattice gauge theories (NALGT) in four dimensions. The problem is certainly technically difficult. It also has been shown recently [3] that the standard multigrid technique has little chance to be as efficient as e.g. in the field of linear differential equations. In this paper I first want to reformulate the problem in a language differing slightly from the one given previously. Two special aspects are the connections of the multigrid blocking scheme with the eigenvalue problem and with nonlinear properties of the action. These considerations suggest to mix the standard blocking either with more refined diagonalization schemes or with a very simple overrelaxation step. The proposed method is applied to asymmetric lattices with large correlation lengths, where it is shown to lead to a speedup. It is necessary to include many links into basic blocks, and if the blocking is made too fine, one encounters difficulties with acceptance.

There are two distinct phenomena in NALGT. Since the theory is asymptotically free, its short range behaviour should be describable in terms of almost plane

waves with a cut off (for admissible wavelengths) in the order of the confinement scale or somewhat smaller. It is here where the standard multigrid techniques are promising tools to reduce CSD. On the other hand, there are topologically nontrivial excitations in the vacuum, which also show CSD, and there no method for improvement seems to be known. The paper is thus concerned with a treatment of the former problem only. Now, within the framework of multigrid (MG) there are many options, and several do not work in the case of NALGT. For an attempt to understand the reasons, it is useful to start with a quadratic approximation to the full action and to consider the associated eigenvalue problem, together with the numerical techniques employed there [1].

It is proposed to take the MG method as a sacrifice of two closely related properties of exact eigenfunction of a quadratic system, namely their noncompactness (in physical space) and their orthogonality, in favour of very large gains in numerical speed. Figures of merit will be how close the lowest expectation values of the action for the truncated functions come to the true eigenvalues, and how quickly they can be found. In the case of updating via overrelaxation techniques, the speed to find the action minimum is decisive, and eigenfunctions need not to be known in in great detail.

In spite of using quadratic approximations, it is not claimed that the nonlinear properties of the action of NALGT can be swept under the rug in any way. For geometrical considerations, they will determine the size of certain basic blocks in x-space and eventually in colour space. Numerically, they show up in the matrix elements of the quadratic approximation and in the surface terms. This will be discussed in section 3, after the notations and the Monte Carlo updating strategy has been explained in section 2. The applications are in sections 4 and 5, with conclusions in section 6.

2 Definition of Multigrid for Almost Quadratic Systems

The standard multigrid techniques as a tool for solving linear equations close to criticality and for Monte Carlo simulations have been well described in [2], and some familiarity with them is assumed here. This also holds for the local iteration techniques for field theories on the lattice. In the following, only the possibilities for accelerating the modes with longer wavelenths are discussed from the viewpoint of the eigenfunction method.

2.1 The approximate eigenfunctions

To fix the notation, I consider a finite volume[2] Ω, where restrictions in the motion of field variables are introduced by a blocking scheme, i.e. by allowing only certain

[1] To use a quadratic approximation to the full action has been suggested several years ago [4,5]

[2] The volume Ω will normally be a small subvolume of the full volume, with an extension in the order of the correlalion length. It will not always have a smooth surface Ω. In gauge theories, schemes known presently provide surfaces $\partial\Omega$ comparable in volume to Ω itself

collective changes in the fields. These collective changes will be denoted by $\vec{r}$. The components of this vector, r_i, describe nonlocal multiplicative changes or rotations of the fields in subvolumes $\omega_i \in \Omega$, called the basic blocks. Both $\vec{r}$ and the ω_i should be chosen such as to minimize the dependence of the action on $\vec{r}$. In that sense they correspond to approximate eigenmodes of the action on ω_i with the lowest eigenvalues. For nonlinear problems, it will not always be possible to keep the action S on the ω_i constant for finite rotations. It is also not to be expected that the action can be expressed effectively in closed form in terms of the $\vec{r}$, although this would be highly desirable for an iterative scheme. Especially for the search for the minimum of S (to be motivated immediately) it suffices to have an implicit definition in terms of the original fields, depending on $\vec{r}$.

We now assume that S has just one minimum in Ω under variations of the collective variables. For later use, let us denote the minimal point of S by $\vec{r_0}$. Since the $\vec{r}$ are continuous variables, we can expand the action around the minimum and keep only terms up to second order. Thus I describe, on the lattice, the quadratic approximation to S by a $N \times N$ matrix M and by a linear term, $\vec{y}$,

$$S(\vec{r}) \approx S_q(\vec{r}) = \frac{1}{2}(\vec{r}, M\vec{r}) - \vec{r} \cdot \vec{y} + const. \qquad (1)$$

The linear term is due to the action contributed by the surface of Ω. N includes, in our case, the coordinates of the basic blocks, spatial directions and colour degrees of freedom. The dependence of S on the field variables is indicated only by its dependence on the rotations $\vec{r}$. Thus the matrix elements of M still depend on the fields, and $\vec{r} = \mathbf{0}$ describes the current field configuration.

For the solution of an almost linear equation, one has to find the minimum of the full action S w.r.t. $\vec{r}$, and, for stochastic systems, to construct the Boltzmann distribution in $\vec{r}$. In these processes, CSD is due to the fact that if the volume Ω goes to infinity, some of the N eigenvalues of M will approach zero. A direct update of these eigenmodes would eliminate CSD, but now it is crucial to note that in general the support of all eigenfunctions is given by Ω. If analytical tools are not available for the given action, numerical methods are required. Especially, M will be known only numerically and the eigenmodes have to be found by standard routines. These require computer work of order N^3. Of course, this is prohibitive for large N, so the idea looks like a blind alley.

This dilemma is due to the properties of the exact eigenmodes, and this exactness may be regarded as a luxury for numerical work. For all iterative methods, approximate eigenfunctions (AE), which even are not orthogonal, work almost as well. This can be understood in the case of the search for a minimum. If we expand the starting vector in terms of eigenfunctions, the minimization of all eigenmodes will bring us to the minimum in N steps. For approximate eigenfunctions, a single minimization will lead to perhaps nonoptimal, but slight changes in other eigenmodes. This may require just a few more iteration steps, and it is irrelevant anyhow in a slightly nonlinear system.

The basic cure is given by the MG technique, which is a prescription for the truncation of eigenmodes:

1. For the shortest wave lengths admitted temporarily for relaxation in the system, the eigenfunctions will be identified with the r_i. For those, piecewise constancy on ω_i is the natural, but not mandatory choice.
2. AE with larger wavelength will be built recursively out of those of smaller wave lengths only over a limited region of neighbourhood both in x-space and in modes with slightly higher eigenvalues. Here "building" means diagonalization in a restricted subspace.
3. For nonlinear quantum systems, a restriction on the ω_i has to be added: The basic blocks have to be chosen such that the current configuration (given by $\vec{r} = \mathbf{0}$) is on the average not too far from the restricted minimum, in order to keep the influence of nonlinearities small. In detail, we need

$$S(\vec{r} = \mathbf{0}) = S_q(\vec{r} = \mathbf{0}) + \delta, \qquad (2)$$

where an error of average size $< \delta > \leq 0.1$ is required [3] for SU(2) with $\beta \approx 3.0$. This accuracy can also partly be achieved by limiting the volume Ω where the field variables are allowed to vary.

The implementation of these principles has to compromise between quality of eigenmodes (most importantly the size of the smallest eigenvalues, then approximate orthogonality and completeness), between numerical speed and programmability, and between the size of the nonlinearities for the current configuration.

2.2 Monte Carlo simulations

Given the AE, there are at least two strategies for the numerical simulation of a quantum system. First, one can exactly orthogonalize a few of the AE with the lowest eigenvalues and propose to change the corresponding modes by a standard heatbath excitation or by an overrelaxation step. Detailed balance (DB) is then assured by the well known Metropolis filtering step [7], namely accepting the proposal $\vec{r}$ with probability

$$P_A = min\{1, \frac{P(\vec{r}^{-1})}{P(\vec{r})} \frac{exp(-\beta S(\vec{r}))}{exp(-\beta S(\vec{r} = \mathbf{0}))}\}, \qquad (3)$$

provided sufficient care is exercised in the construction of the AE (any error must not depend whether one starts from the current configuration or from the one generated by the present algorithm). Here $P(\vec{r}^{-1})$ denotes the probability for going from the rotated configuration to the original one. This and $P(\vec{r})$ are determined by the quadratic action S_q in the case of a heatbath step, e.g.

$$P(\vec{r}) \propto d\vec{r}\, exp(-\frac{\beta}{2} \sum_k \lambda_k ((\vec{r} - \vec{r_0}) \cdot \vec{e_k})^2), \qquad (4)$$

where λ_k and $\vec{e_k}$ are the eigenvalues and -vectors resp. For overrelaxation the $P(\vec{r})$ are given by unity. This procedure will solve the problem of slow motion through phase space of the lowest eigenmodes provided the acceptance P_A is close to 1.

[3] This will become quantitatively clear in the next subsection

It is well known [6], however, that overrelaxation steps are much more efficient in numerical simulations than heatbath steps [4]. The simplest overrelaxation step is a reflection of the whole vector $\vec{r}$ in the point of minimal action, $\vec{r}_0$,

$$\vec{r} \Rightarrow 2\vec{r}_0 - \vec{r}, \tag{5}$$

which is equivalent to a reflection of all eigenmodes. Again, for a nonlinear theory, the full action S will not be preserved under 5, and a Metropolis step is necessary. In order that the acceptance is close to one, we need the product of all exponentials in 3 to be larger than roughly $exp(-0.1\beta)$ which quantifies the principle 3 above.

Now the actual approximate eigenmodes need not to be known, one 'just' has to find $\vec{r}_0$. Here the use of the AE will a very efficient tool, e.g. superior to a conjugate gradient algorithm because of the restricted support of the AE. Since we do not need the exact eigenvalues as in the case of a heatbath step, idealized AE are sufficient, which e.g. can be simply found by solving the eigenfunction problem with the first matrix M which shows up in the course of the iterations. The search for the minimum has thus to occur along the directions of N optimally localized AE, and it should be reached in $O(N)$ steps with an average work load of $logN$.

It should be repeated that the specification of Ω and of the basic blocks ω_i is tightly connected with the nonlinear properties of the action, with respect both for the acceptance and for the uniqueness of the minimum. If we take e.g. as basic blocks just one lattice site, the fields have enormous freedom in the search for the minimum and the quadratic approximation will be very poor for the current configuration, resulting in a disastrous Metropolis acceptance. This will be discussed in more detail in the next section.

3 Implementations for NALGT

There are a few technical problems and a major fundamental one in NALGT[5]. Among the former, one realizes the need for linearization, for the respectation of DB in the search for the minimum, for an efficient definition of the quadratic action and for a fast evaluation of the full action. The main difficulty is in the prescription for building the basic blocks ω_i.

3.1 Linearization of the gauge fields

The above discussion was based on the concept of a linear space, and mainly here the specialization to SU(2) is helpful in making the transition to linear algebra easy. The collective changes are global rotations $R_i \in SU(2)$ which multiply gauge variables U_j, $j = 1, ..., N_i$ which live on a set of links[6] $\omega_i \in \Omega$. Actually,

[4]The quadratic approximation also offers the possibility of a cluster algorithm [8], with the reflection in the minimum as basic operation

[5]The remarks in this section will be qualitative, as detailed technical realizations will be given in sections 4 and 5

[6]It is hoped that the definition of the ω_i as a set of sites and their use as a set of gauge variables living on links attached to the sites will not cause confusion to a careful reader

the R_i will act only on gauge transformed fields U'_j (with the gauge to be described below),

$$U'_j \Rightarrow R_i U'_j, \tag{6}$$

or, equivalently, the R_i will be parallel transported [2,3] within ω_i.

The quadratic approximation is obtained by working with the 'small' components $\vec{r}_i$

$$R_i = r_{i,0}\mathbf{1} + i\vec{\sigma}\vec{r}_i, \quad r_{i,0} \approx 1 - \vec{r}_i^{\,2}/2, \tag{7}$$

where the $\vec{\sigma}$ are the Pauli matrices. In principle, performing linear operations with the $\vec{r}$ may lead to inconsistencies. This can be avoided by using a kind of stereographic projection [5], but the necessity to keep the cubic terms small is restriction enough in practice to avoid this problem.

3.2 The blocking

Blocking schemes for gauge theories have been given in [2]. A simple one is to consider the sites on a 2^4 hypercube, Λ_0, which are being connected either by parallel transport or by gauging, and then use a common rotation on the links emanating from the sites in a given positive direction. This rotates 8 parallel links simultaneously, and many plaquettes have all their four links changed. The complete set of ω_i is obtained by combining all 4 directions. The scheme has been dubbed "wildly impractical" in [2], because there is no hope to get the action in terms of the $\vec{r}$ in closed form, but this is not necessary here.

There is indirect evidence that this blocking is too finegrained and that it will introduce too strong nonlinearities. This suspicion comes from another blocking scheme, described in [5]. There the links on the **surface** of the 2^4 hypercube are gauged (the gauge being optimized w.r.t. staples $\notin \Lambda_0$ to ensure DB) and changed. This scheme can be modified easily such as to gauge several links to 1 and rotate the remaining ones such that the action is preserved exactly for plaquettes $\in \Lambda_0$. It is thus easily possible to rotate 12 link variables simultaneously. We shall see in section 5 that the acceptance for this scheme is dangerously low, and it is to be expected that the same is true for the scheme of [2].

From the programming point of view, a convenient scheme is to combine all the links pointing e.g. into the x-direction and being attached to sites in a hyperplane with fixed x. Such a blocking has been employed in [3] and it will be studied also here, being called **the x-slice blocking** . Apart from simplicity and from a special physical situation, this blocking scheme in general is not a good candidate. The various blocks shifted along the x-direction are not coupled and a spread of the change of the action along this direction is not possible. Thus we expect the minimal eigenvalues to be considerably too large as to be optimal. A natural extension would be to consider blocks of links orthogonal to x and shifted in x-direction by one unit. It remains to be seen, whether this is feasible. A straightforward, but insufficient correction is to add a standard local Monte Carlo update step, changing the orthogonal links immediately after the global update step.

For the extension of the x-slice method mentioned just before and for the first scheme [2], there are plaquettes where more than two links will change under variation of the $\vec{r}$. Thus the action is genuinely nonlinear, as it should be in Yang-Mills theory. The quadratic approximation obviously cannot handle this feature convincingly. Either one restricts the basic blocks such that only two links of a plaquette are changed[7], in which case the special Yang-Mills couplings are created on the spatial surface of Ω, or one realizes the couplings through the Metropolis acceptance step. This is probably not efficient. It should be noted, however, that the continuum limit of the Yang-Mills coupling due to one plaquette is quadratic in all the colour indices individually. It is therefore possible to consider the restriction to one colour component as part of the definition of Ω, of course at the expense of tripling many of the numerical operations. To find the proper compromise at this point is an unsolved task.

A significant part of the blocking prescription lies in the definition of a smooth gauge. This affects the structure of the quadratic action considerably, and it will be discussed in the next section.

3.3 The quadratic action

It is wasteful to calculate the matrix M already at the beginning of the iterative search for the minimum. The first few relaxation steps can conveniently be done by semilocal steps in the ω_i individually, until the higher than quadratic terms have been damped away. Only then the expansion 7 will be used on all relevant plaquettes to obtain M. The matrix elements are sums over the plaquettes connected to the various r_i with the structure of 3×3 blocks in colour space and with a band structure with nearest neighbour couplings in x-space. To make use of the latter is especially transparent in the linear blocking geometry of the next section. The approximate eigenfunctions of small wavelength can be obtained easily by switching off certain nearest neighbour couplings, and calculations with AE can be speeded up by simple restrictions in dot products.

In a global smooth gauge (on Ω) and at large β, the diagonal elements in colour space are dominant. This follows from the fact that in colour space parallel rotation r_i should leave the action in the continuum limit almost invariant. This allows to take as AE functions with excitations along one direction in colour space only, which saves considerable work. On the other hand, if one searches for a smooth gauge only within the ω_i, the gauging process will be speeded up. The price is that the nondiagonal elements in x-space will have no colour-diagonal dominance. Whether it is possible to diagonalize these efficiently by semi-global gauge transformations between the ω_i - without ruining the gauge within the ω_i - remains to be seen. For programming simplicity only global gauges have been used in the present work.

It is advantegeous to try to write the full action with the sum over the plaquettes performed. In this way, the acceptance test 3 can be done quickly. For coupling schemes with at most two variable links per plaquette, this leads to a 4×4 tensor in colour space which is easy to manage, whereas with three links

[7] We assume that the nonlinearities due to the compactification of the gauge group are less important

per plaquette it may become tedious. It should be repeated, however, that the pessimistic statements of [2] do not apply here because no further blocking is necessary. The latter is done in the quadratic approximation.

4 The x-Slice Blocking

4.1 The physics

The overrelaxation technique (reflection in the action minimum) is applied to the case of spatially small lattices with large correlation length along the (long) time direction. The sizes are taken as

$$Volume = L_s^3 \times L_t, \tag{8}$$

where the sizes L_s and L_t compare to the correlation length λ as

$$\begin{aligned} L_s &< \lambda, \\ L_t &\gg \lambda. \end{aligned} \tag{9}$$

The physical situation in these small volumes [9,10,11,12] is that the gauge fields along the short directions fluctuate around classical vacua. The vacua are separated by potential valleys along which tunneling can take place. This manifests itself in domains along the t-direction, in which the Polyakov loops

$$P_x(y,z,t) = \frac{tr}{2} \prod_{1 \le x \le L_s} U_x(x,y,z,t), \tag{10}$$

(and similarly for the y- and z - directions) are clustering close to the SU(2) center elements ± 1. The tunneling of these domains occurs quite slowly and may require hundreds or thousands of Monte Carlo iterations.

A successful application of MG requires that the correlation length is large compared to the lattice unit. Here I increase the correlation length in t-direction by reducing the bare coupling constant in t-direction. Since the lattice is asymmetric anyhow, this is no loss. In detail, the plaquettes with one link along the t-direction will be multiplied by

$$\beta_t = \kappa\beta, \tag{11}$$

and in the numerical simulation I take $\kappa = 1.5$. This will generate correlation lengths in the order of 20 to 30 lattice spacings.

4.2 The blocking

The principal volumes Ω are taken as simple subsets of the links sitting on 3-dimensional hyperplanes, pointing into the fourth direction (see fig.1 for a 3-dimensional simplification).

Because of the alignment of the $P_x(y,z,t)$ in the spatial direction, the size of the basic blocks can be taken as L_s, i.e. a finer separation in the spatial directions

does not seem to be neccessary at first. It remains to define the basic blocks in the t-direction[8]. There we perform no blocking at all but allow all time slices to vary independently in the search for the action minimum. In order to keep the computer load managable, the volume Ω will be restricted to an interval of size T. Thus, the basic blocks ω_i are defined by the links $U_\mu(\vec{x})$ with direction $\mu = x$ and position $\vec{x}$ of sites

$$\vec{x} = \{x = x_i, 1 \leq y, z \leq L_s, t = t_i\} \tag{12}$$

with

$$1 \leq x_i \leq L_s, \;\; t_0 \leq t_i < t_0 + T. \tag{13}$$

In order to define meaningful nonlocal rotations, a smooth gauge has to be found. In the hyperplane given by $x = x_i$, all the links with $\mu \neq x$ and with $t_i \leq t < t_i + T$ will be iteratively subject to a Landau gauge, i.e. the small components of the U_μ are minimized. In an initial step, the links are gauged axially along a maximal tree. Since these links are not changed under the global rotations, there is no difficulty with DB. This gauge will be used simultaneously for the links pointing into negative x-direction, starting from the same sites as above.

The nonlocal rotations R_i of the gauged links U'_j are now defined by

$$U'_j \Rightarrow R_i U'_j, \quad U'_j \in \omega_i. \tag{14}$$

If $R_{i,0}$ denotes the set for which the action is minimal, the overrelaxion step is given by

$$U_j \Rightarrow R_{i,0} U'^{\dagger}_j R_{i,0}, \quad U_j \in \omega_i, \quad \omega_i \in \Omega. \tag{15}$$

The principal shortcomings of this blocking have been mentioned in the previous section, together with the preliminary improvement to include single link updates in the appropriate hyperplanes after the global rotations.

4.3 The search for the minimum

In the present simulations, block sizes up to $T = 16$ have been considered. In the linear approximation, this amounts to 48 variables for the global rotations R_i which of course pose no problem in the search of the minimum. The purpose here is to learn the techniques, which include a correct choice of the starting point for the R_i, judicious use of local relaxation steps and of standard minimization techniques from linear algebra.

4.3.1 The starting point

Setting the R_i to unity at the beginning will violate DB unless the minimum is unique and is found with high accuracy. This is because the starting point is just the present configuration so the resulting point $R_{i,0}$ will depend on the initial condition. To cure this, we need a starting point which is the same for

[8]It is here where the present approach differs essentially from the calculations of ref. [3]

all semilocally rotated configurations. We first observe that in the calculation of the action, the summation over the plaquettes and the rotation by R_i can be interchanged. Thus, e.g. the contribution to S from the left boundary can be written in terms of the closed plaquettes P_k (see left side of fig.2), which are attached to the links of ω_1, as

$$S_l = \frac{tr}{2} R_1^{\dagger} \sum_k P_k = p_0 \frac{tr}{2} R_1^{\dagger} P_{k,0}, \tag{16}$$

where p_0 denotes the determinant of the sum, and $P_{k,0}$ is a SU(2) matrix. The block rotation is thus a global multiplication also in the action. If we rotate the matrix $P_{k,0}$ to a direction independent of the current configuration (insofar as it may be changed by the present update) and start with $R_i = 1$ by gauging with a matrix G_1, the minimum will be reached in an configuration independent way. An appealing choice for this direction is the T-th root of a line product leading from the lower boundary to the upper one which is independant of the R_i. The path corresponding to such a product is depicted in fig.2. This method allows a successive determination of SU(2) elements G_i, $i = 1, ..., T$ for all the elements making up the action (the internal elements are only slightly more complicated). The current configuration is then given by the $R_{i,c} = G_i^{\dagger}$, and the formula for the overrelaxation 15 has to be modified accordingly. For alternatives in choosing the starting point, cf. [1].

4.3.2 Relaxation towards the minimum

Conveniently the relaxation can be done in two or three different groups of steps. The first group should consist of local steps in the R_i directly, because a large part of the deviation from the minimum is due to the quantum noise on the coarsest level. Whether these steps should be done with the full action or already with the quadratic approximation eq. 1 obtained via 7, depends on the average rotation angle given by the R_i. In our case this angle is of the order of 10° which certainly allows to use 1 immediately. Since we start from $R_i = 1$ or, equivalently, from $r_i = 0$, the first two steps are (the colour index has been combined with the spatial index)

$$\begin{aligned} r_i &= y_i / M_{i,i}, \;\; i = 1, ..., 3T, \\ r_i &\Rightarrow r_i - \textstyle\sum_{j \neq i} M_{i,j} r_j / M_{i,i}. \end{aligned} \tag{17}$$

Given the $\vec{r}$, the full R_i, $i = 1, ..., T$ are reconstructed by unitarization, the potentials (boundary and internal) are gauged as in the last section to achieve $R_i = 1$ as above, and the $R_{i,c}$ describing the current configuration have to be gauged accordingly. This group of steps can be repeated a few times.

Since the matrix M can be obtained from the action in a form where all the sums over the plaquettes contributing to S have been performed, these steps require work of order $N = 3T$ with a small coefficient (small as compared with the work necessary to calculate all the plaquettes attached to the links $\in \Omega$).

As a preparation for the next step, AE are generated as follows. At the beginning of the calculation, from the first matrix M encountered, a truncated matrix

M' is obtained by setting those elements to zero which are nondiagonal in colour [9]. From M' a complete set of eigenvectors $\vec{e}'_k$, $k = 1, N$ is obtained and stored for the future. These eigenvectors come in three groups having support only for one colour component each.

As the next improvement, the linear piece is expanded in terms of the e'_k and we have

$$\vec{r} = \sum_k \vec{e}'_k \; \vec{e}'_k \cdot \vec{y}/\lambda'_k, \tag{18}$$

where the approximate eigenvalues λ'_k are given by the expectation value of M,

$$\lambda'_k = (\vec{e}'_k, M\vec{e}'_k). \tag{19}$$

Whereas the step 18 obviously requires work of order $3T^2$, one has to take into account the band structure of M to recognize the same for 19. Here the truncation of the colour-nondiagonal terms saves, for larger T, already a factor 3 in computer time.

A possible refinement in the relaxation as a third step will not be discussed here.

4.4 How to monitor acceleration

The interesting physical quantity here is the lowest mass which can be derived from the correlation function in t-direction of spatially averaged Polyakov loops, $P_x(t)$,

$$\begin{aligned} P_x(t) &= < P_x(y,z,t) >_{y,z}, \\ C_x(\tau) &= < P_x(t) P_x(t+\tau) >_t, \\ C(\tau) &= < C_w(\tau) >_{w=x,y,z}. \end{aligned} \tag{20}$$

It has been shown in [13] that a good way to cope with the strong correlation of statistical errors in $C(\tau)$ is to consider its Fourier transforms with frequencies ν_n,

$$\tilde{C}_n = \tilde{P}^2(\nu_n), \quad \nu_n = 2\pi n/L_t, \quad n = 0, ..., L_t - 1. \tag{21}$$

Especially important are the transforms for small n, which are most sensitive to the smallest mass. Their autocorrelation function in computer time (number of iterations) is difficult to measure as it is, in the case of tunneling, characterized by large fluctuation. It turned out that a fast and, for not too distant measurements, reasonably well reproducible quantity is the increase of the variance Σ_n as function of the binwidth i_{max},

$$\Sigma_n(i_{max}) = \sqrt{< \tilde{C}_n^2 >_i - < \tilde{C}_n >_i^2}, \quad i = i_0, ..., i_0 + i_{max}. \tag{22}$$

Independent of the algorithm, this quantity will eventually approach the same limit for $i_{max} \to \infty$, but the speed of approach will depend on the algorithm. As a simple measure of acceleration, I take the number of iterations needed to reach a

[9] See last section for a justification

given Σ_n. For a final refinement, the variance can be averaged over a few Fourier coefficients with the smallest frequencies, as these are quite similar in magnitude. This reduces the statistical noice. As we will see, there is a marked dependence of the resulting Σ on the starting point of the bin, i_0 in the sense that for 'young' configurations (ca 1000 iterations after a cold start) the increase of Σ as function of i_{max} is much steeper than for 'older' configurations (after ca 5000 iterations). Taking this into account, the assignment of an error to Σ is quite problematic.

4.5 Results and computer timing

Lattices of the size $L_s = 8$ and $L_t = 64$, 192 and 256 have been investigated on workstations, on a cluster of workstations and on the parallel computer iPSC/860. For the cluster, the PVM message passing package has been used which could be tested on a single workstation. The change between this and the iPSC parallel software was quite easy. Parallelization was done by domain decomposition along the t-direction. The Wilson form of the single plaquette action was taken. The values of β used were 2.65 and 3.0, with the modification for timelike plaquettes 11. At $\beta = 3.0$, this raises the expectation value of spacelike plaquettes from $< P >= 0.723$ to $< P >= 0.755$, which is not a dramatic change.

All initial lattices configurations were cold, such that all Polyakov loops were equal to unity. The tunneling to negativ values sets in after a few hundred iterations, and after a few thousands sweeps large domains with negative values are observed. From this it can be concluded that the special conditions -large and asymmetric β- do not make the system absolutely exotic. Measurements of the Σ_n were performed repeatedly around $i_0 = 1000$ sweeps and around 3000 - 6000 sweeps. The results are shown in figs. 4 and 5. Obviously, for the 'young' configurations the gain ranges between factors 3 and 5, if the number of sweeps necessary to reach the same variance is taken as a measure. For the older configurations, this gain reduces to a factor 1.5 to 3. This phenomenon will be called 'aging' in the following[10].

There are various ways how to rate this gain. If the number of links which have been updated is taken, the gain is reduced by a factor 1.5, and, if the average of the ratios of computer times is taken, it is reduced by a factor around 3. Thus the net gain is dubious, but it is necessary to look into the various parts of the code and estimate the chances for tuning. The biggest individual contribution comes from gauging (45 %) which divides equally between the axial gauge and the Landau gauge. The former has excessive index calculation, whereas the latter is also a candidate for MG which could reduce the number of iterations considerably. The single link updates consume 17 % of the time and the search for the minimum 18 %. The search can also be speeded up by the use of spatially truncated AE (we only use truncation in colour space).

Other figures of interest are:

- The average angle of rotation of the blocks which is initially around 13°, dropping by 2° during aging.

[10] Of course, it is a phenomenon of equilibration

- The smallest eigenvalues, which correspond to

$$\lambda_{min} = 2\partial S/\partial\theta^2 = 0.25 \qquad (23)$$

per link, where θ is the global link rotation angle. This is 5% of the value of a local change.

- The ratio of the smallest eigenvalue (found by diagonalizing M) to the expectation value of M for spatially constant (t-independent) rotation. This ration is around 2.0, which is a significant improvement.
- The trend of the eigenvalues if the volume of Ω is increased. By increasing T from 8 to 16 the smallest eigenvalues decrease by 40%
- the Metropolis acceptance rate, which is around 96 %.

To find the source of aging will require special investigations. The fact that it sets in relatively late indicates that it is not connected with short range fluctuations (the 'quantum noise'). More likely is a connection with tunneling, due to the appearance of several complicated transition regions between different domains. It is not clear how extended these transition regions are and whether the Polyakov loops in all three directions are affected by the necessary disorder. This problem is under study. A possible improvement could be to split the ω_i along the y,z plane into e.g. 2×2 subvolumes.

5 The Partial Gauge Blocking Scheme

If one considers as a basic block all the links on the surface of the hypercube $\Lambda_0 = 2^4$ and the corresponding eigenmodes with minimal action, those with global rotation of 8 parallel links are a natural choice. These rotations require either gauging in a 3-dimensional subcube or parallel transportation along various paths on this subcube. For nonabelian theories the rotations will change the action on the plaquettes $\in \Lambda_0$, since the rotation does not commute with the parallel transporters. This, at least, requires to calculate many more plaquettes in the search for the minimum and in the check for Metropolis acceptance than in the scheme described in the following.

If one rotates all the 4 links leaving a given corner of Λ_0,

$$U_j \Rightarrow U_j R, \quad j = 1, ..., 4, \quad U_j \in \Lambda_0, \qquad (24)$$

then the plaquettes on Λ_0 will not change. Since it is half of the multiplications of a full local gauge transformation, I call it a partial gauge. As stated in [13], one can directly find the local action minimum in analogy to the single link technique. Four such eigenmodes can easily be put on Λ_0, without overlap of links. The coupling to modes defined on other subvolumes occurs at the corners where two neighbouring cubes meet.

In order to check wether this scheme can be successful it is not necessary to determine the quadratic approximation, as the action minimum can also be found by repeated semilocal relaxations of the form 24, requiring O(10) steps. It turns

out that the admissible rotations on Λ_0 are so large that nonlinearities will spoil the acceptance P_A completely.

Nonlinearities can be reduced by combining more links into the basic blocks. For this purpose, one starts at one corner on Λ_0 and gauges all four links to unity. Applying then the rotation 24 at that corner and at the endpoints of the 4 links, one effectively can change 12 links simultaneously (see fig. 3 for a projection of the 4-dimensional case. The links gauged to unity are shown by broken lines). A coupling to rotations on other cubes is introduced if on those cubes which are connected on one preselected 4-dimensional diagonal, the start-corners are reflected in the center of the cubes (see fig. 3 for more details). It is tedious to find out that the plaquettes which carry the coupling contain only 2 varying links, such that the conditions for the quadratic approximation are met. The minimum is again found by semilocal relaxation.

This blocking scheme has been applied to the same lattice geometry as in the last section, with the volume size T varying between 4 and 16. It was found out, on the one hand, that the average rotation angle is almost twice as large as in the x-slice blocking scheme, and that the Metropolis acceptance varies between 75% and 55% for the quoted volumes, which seems to be tolerable. On the other hand, the effect on Polyakov loops is too small by a factor 2 (expressed in the variances Σ_n, eq. 22) to make this scheme competitive with the x-slice blocking and even with the local algorithms. A natural interpretation is that the coupling between the basic blocks is too weak as to generate really long range correlations.

6 Discussion and Conclusion

The results of section 4 can be summarized as follows: The search for the action minimum under the constraints of a blocking scheme is technically feasible, and many chances for improvements are apparent, such that more than 16 basic blocks can be used without excessive demands in computer time. With the 64 links present in a basic block, useful rotation angles in the order of 10° have been obtained, and the smallest eigenvalues (which need not be calculated in practice) have been reduced with respect to the single link case by a factor 20. The nonlinearities associated with this blocking scheme are so small that the Metropolis acceptance is very good. Especially the last point will allow a more refined split of the volume Ω with a further lowering of the eigenvalues. For 'young' configurations, the gain in speed (measured in numbers of sweeps) is large enough that it can be converted into a gain in computer time after a few tuning steps. This is marginally true for well equilibrated configurations.

The partial gauge blocking scheme investigated in section 4 indicates that basic blocks with 12 links most likely are too small for a successful quadratic approximation, and that the coupling between the basic blocks has to be made stronger than in that scheme in order to get a significant speedup. On the other hand, coupling schemes where more than two links are changed simultaneously within a plaquette, have genuine nonlinearities which need special treatment of the colour degrees of freedom.

The general framework for MG schemes given in sections 2 and 3 does not differ deeply from the ones given previously [1,2]. It emphasizes more strongly the need to simultaneously

- define the the basic blocks such that nonlinearities are absorbed in the action matrix M and into the linear term y (see eq. 1),
- lower the expectation values of approximate eigenfunctions,
- find a fast way to get to the action minimum under the constraints of the basic blocking.

The quadratic approximation serves as a useful tool for heatbath and overrelaxation proposals.

It remains to be seen whether the problem of of aging can first of all be understood in detail and then be dealt with by moderate improvements.

Acknowledgement

Calculations have been performed on a RS/6000 workstation, on a HP730 workstation, on a cluster of 4 RS/6000 at the IBM Scientific Center in Heidelberg and on 8 nodes of the iPSC/860 of the KFA Jülich and the HLRZ. Special thank is due to members of R2 at DESY, of ZAM at KFA and of IBM for their help with the various machines. Usefull conversations with G. Mack and M. Grabenstein are gratefully acknowledged. I also want to thank M. Grabenstein for communicating his unpublished results on the acceleration in the blocking scheme of [3]. The help of P. Altevogt and U. Scharffenberger (IBM Heidelberg) with the PVM program, with problems of parallelization and of linear algebra has been very valuable.

References

[1] G. Mack in G. 't Hooft et al. (ed.), Nonperturbative Quantum Field Theory, Plenum Press, N.Y. 1988 (Cargèse Lectures 1987)

[2] J. Goodman and A. Sokal, Phys. Rev. D **40** (1989) 2035

[3] M. Grabenstein and K. Pinn, Kinematics of Multigrid Monte Carlo DESY-094, July 1992

[4] S.L. Adler, Phys. Rev. **D38** (1988) 1349

[5] F. Gutbrod in A.S. Kronfeld et al. (ed.), Lattice 88 (Fermi National Laboratory Symposium 1988), Nuclear Physics (Proc. Suppl.) 9 (1989), p. 503

[6] S.L. Adler in A.S. Kronfeld et al.(ed.), Lattice 88 (Fermi National Laboratory Symposium 1988), Nuclear Physics (Proc. Suppl.) 9 (1989), p. 503

[7] Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller and E. Teller, J. Chem. Phys. **21** (1953) 1087

[8] U. Wolff, Phys. Rev. Lett. **60** (1988) 1461, Phys. Rev. Lett. **62** (1989) 361

[9] M. Lüscher, Nucl. Phys. B **219** (1983) 233

[10] J. Koller and P. van Baal, Ann. Phys. (NY) **174** (1987) 299

[11] J. Koller and P. van Baal, Nucl. Phys. **B302** (1988) 1

[12] P. van Baal, Phys. Lett. **224B** (1989) 397

[13] F. Gutbrod, Z. Phys. C, **55** (1992) 463

Figure Captions

Fig. 1: Geometry for the x-slice blocking in 3 dimensions with $T = 4$. The links sitting on one vertical line form the basic blocks.

Fig. 2 Update invariant path to define the starting gauge. The plaquettes on the left and right contribute to the surface action.

Fig. 3 Geometry for the partial gauge blocking scheme. The full links will vary such that the enclosed plaquettes do not change.

Fig. 4 Averaged variance Σ for $\beta = 2.65$ on a lattice of size $8^3 \times 192$, as function of bin width i_{max}. Units for Σ are arbitrary. The full circles are due to the a mixture of 4 overrelaxation steps with 1 heatbath step.

Fig. 5 Averaged variance Σ for $\beta = 3.0$ on a lattice of size $8^3 \times 256$, as function of bin width i_{max} (see also fig. 4).

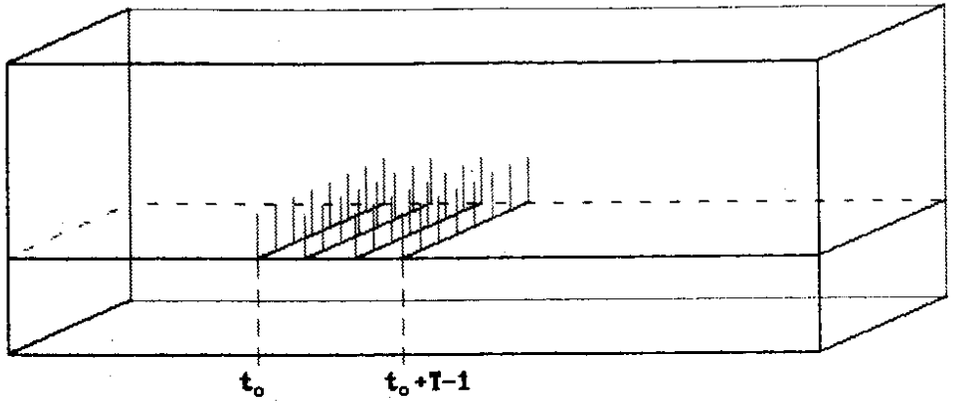

Fig. 1

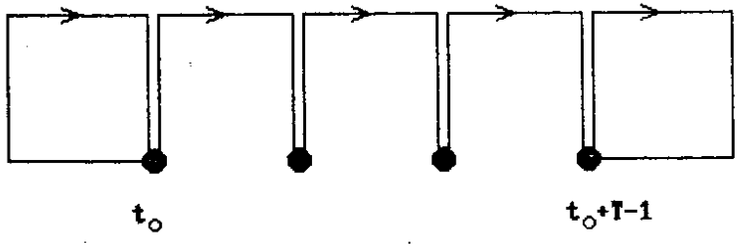

Fig. 2

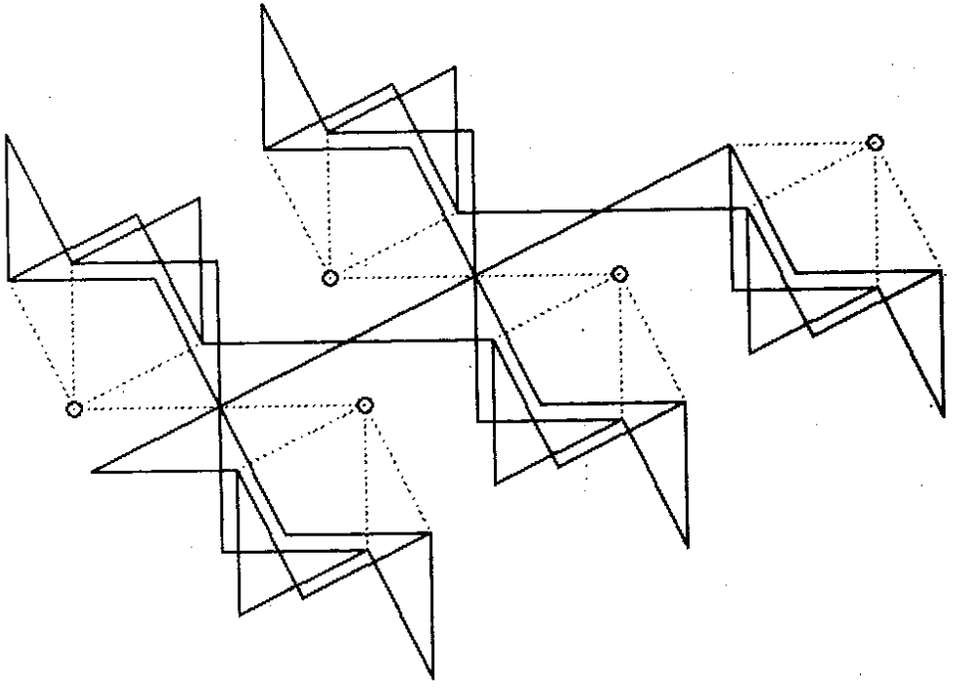

Fig. 3

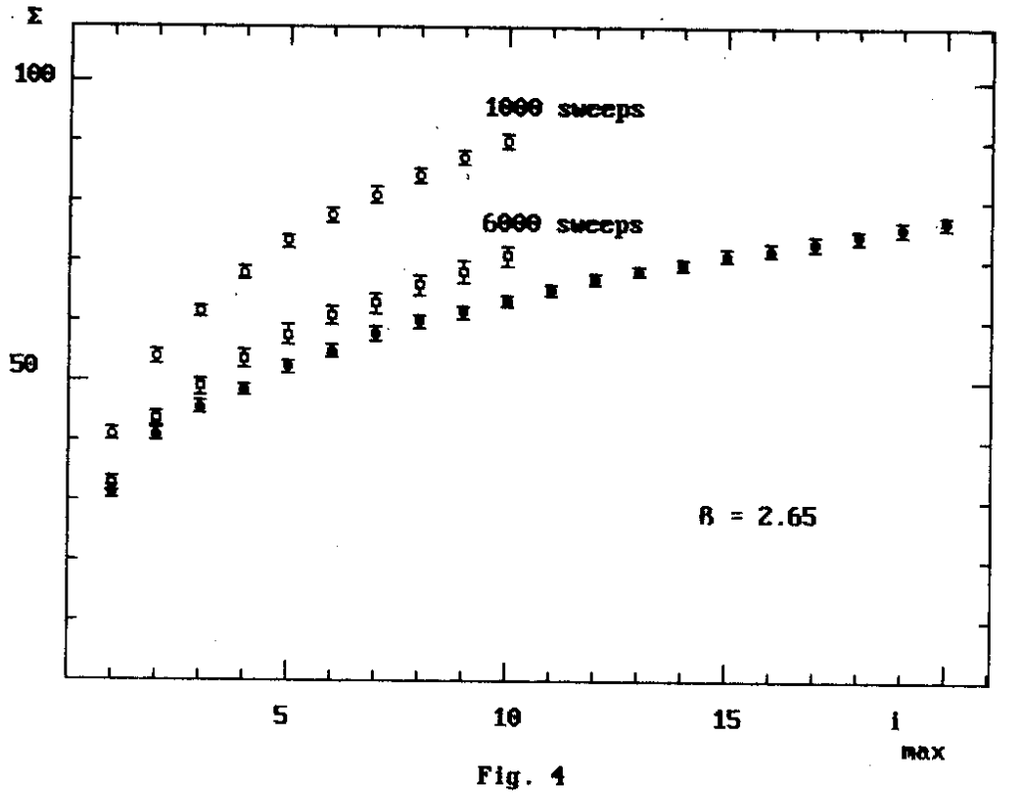
Σ
100
50
1000 sweeps
6000 sweeps
ß = 2.65
5
10
15
i
max

Fig. 4

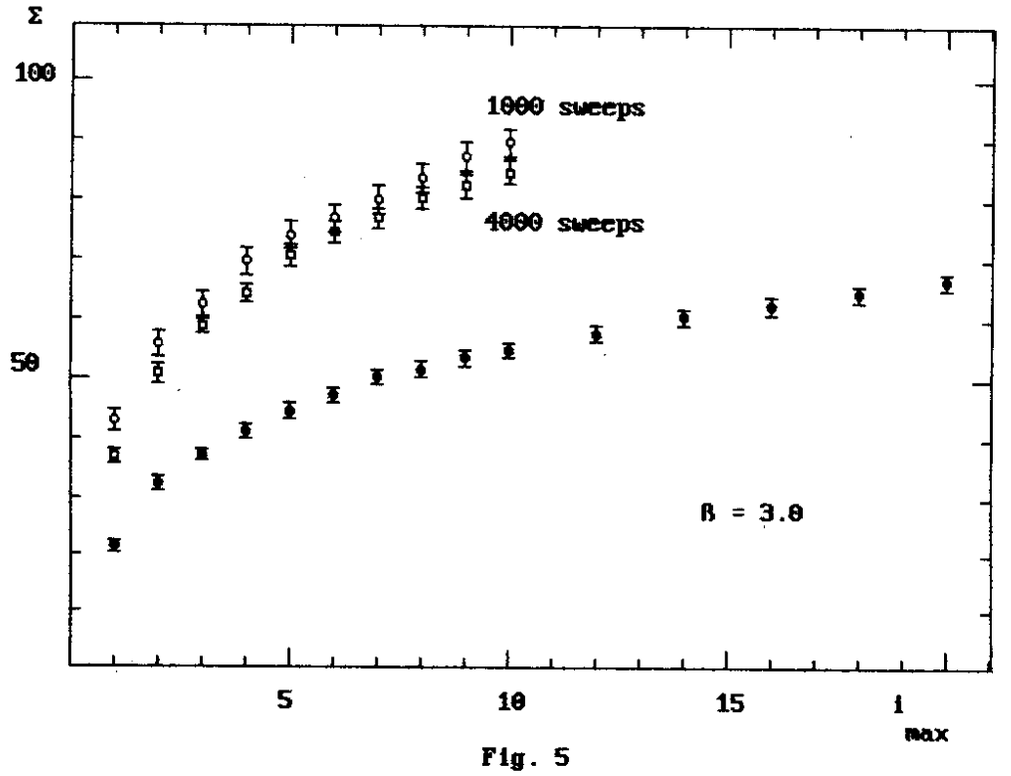
Σ
100
50
1000 sweeps
4000 sweeps
ß = 3.0
5
10
15
i
max

Fig. 5