Nuclear Physics B 376 (1992) 637–660
North-Holland

NUCLEAR
PHYSICS B

Projective block spin transformations in lattice gauge theories *

Thomas Kalkreuter

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

Received 7 January 1991
(Revised 10 October 1991)
Accepted for publication 21 November 1991

A block spin definition for Higgs fields and non-abelian lattice gauge fields is examined which is in the spirit of the projective multigrid procedure of Hulsebos et al., and Brower et al. Their procedure is generalizable to a systematic multigrid method for non-abelian gauge theories in any number of space-time dimensions. The block spin transformation involves a gauge-covariant kernel C which makes the fine-to-coarse transition, and a kernel $\mathcal{A}$ which makes the coarse-to-fine interpolation. These kernels could be used in a projective multigrid computation of propagators, for instance. Vectorizable algorithms for the computation of C and $\mathcal{A}$ are presented. For SU(2) lattice gauge theory in four dimensions, the required CPU time for computing C or $\mathcal{A}$ on the whole lattice is comparable to that for one standard Monte Carlo sweep through the lattice. Numerical results obtained after performing one blocking transformation from a 9^4 to a 3^4 lattice are presented. The block spin computation yields auxiliary quantities of interest, such as the lowest eigenvalues λ_0, λ_0' of the negative gauge-covariant laplacian with Neumann and Dirichlet boundary conditions on block boundaries. λ_0 and λ_0' are measures of disorder and their renormalization group flow is therefore instructive.

1. Introduction

The question of the proper definition of block spins is posed as soon as one wants to investigate a physical theory by means of renormalization group (RG) methods [1]. The multigrid (MG) approach to quantum field theory [2,3] amounts to simultaneous consideration of a whole sequence of renormalization group transformations (RGTs). An important ingredient is the choice of maps C and $\mathcal{A}$ which make the fine-to-coarse and coarse-to-fine transitions. Let us briefly pause to introduce them.

Through a sequence of RGTs a theory with fields $\phi(z) = \Phi^N(z)$ on a fundamental lattice Λ_N of lattice spacing $a = a_N$ gets mapped step by step

*Work supported by Deutsche Forschungsgemeinschaft.

into theories with fields $\Phi^{N-j}(x)$, called block spins, which live on lattices Λ_{N-j} of increasing lattice spacing $a_{N-j} = L^j a$. Typically, $L = 2$ or 3 is chosen and a_0 is of the order of the physical length scale. A block spin transformation amounts to a specification of maps

$$C : \Phi^k \to \Phi^{k-1} = C\Phi^k . \tag{1.1}$$

In the language of the MG approach [4] to partial differential equations, this is called the fine-to-coarse transition. In an MG approach one needs also a map effecting a coarse to fine interpolation

$$\mathcal{A} : \Phi^{k-1} \to \psi^k . \tag{1.2}$$

ψ^k is called the background field. It lives on Λ_k and is required to have block spin $C\psi^k = \Phi^{k-1}$, i.e.

$$C\mathcal{A} = \mathbf{1} . \tag{1.3}$$

For matter fields ϕ (Higgs fields or Fermi fields) which take their values in a linear space, one may choose C and $\mathcal{A}$ as linear maps, so that they are given by kernels $C(x, z)$ and $\mathcal{A}(z, x)$. Points $x \in \Lambda_{N-j}$ may be identified with cubes of side length L lattice spacings in the lattice Λ_{N-j+1} with points z, etc. In practice C is always chosen local, so that $C(x, z) = 0$ unless z is in cube x ($z \in x$ for short).

In a gauge theory the block spins are required to transform in a gauge-covariant way when the matter field on the fundamental lattice is subject to a gauge transformation. This implies that kernels C and $\mathcal{A}$ are gauge covariant, too, and must therefore depend on the lattice gauge field. This means that they have to be adjusted whenever the gauge field is updated. Therefore, efficient algorithms to compute them are important. It will be seen in sect. 2 that a blocked gauge field can be defined by using kernels C, $\mathcal{A}$ for the Higgs field.

The background field $\psi^k = \mathcal{A}\Phi^{k-1}$ is supposed to be the sum of the low frequency components of Φ^k, the information about these must therefore be contained in Φ^{k-1}. Traditionally [5,6] the kernel C is somehow chosen a priori, and $\mathcal{A}$ is then determined so that $\psi = \mathcal{A}\Phi$ minimizes some approximation to the (effective) action (on Λ_k), subject to the constraint that the block spin $\Phi = C\psi$ is fixed. However, it seems reasonable that the notion of low frequency, and therefore also the proper definition of block spin, i.e. of C, should depend on the local action. This point of view is taken by Hulsebos et al. [7] (stochastic MG) and by Brower et al. [8] (deterministic MG) and we adopt it. They take $\mathcal{A} = C^*$, but this is not crucial. The proposal for the construction of block spins presented in the present paper aims at use in

multigrid Monte Carlo (MGMC) simulations without critical slowing down of non-abelian gauge theories. A correct MGMC programme for euclidean field theories — without truncation of effective hamiltonians — was initiated by Mack [3] some time ago. This MG proposal has the additional advantage that it yields the effective action and its derivatives for free [9]. Moreover, one can extract infinite volume results from a simulation in a finite volume as shown by Palma [10] for pure ϕ^4 theory. Furthermore, it was pointed out in ref. [3] 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. The present author [11] made a proposal how to deal with polymer systems with indefinite activities.

Previous works or proposals [12–17] attempt to construct parallel transporters on a block lattice by restricting attention to preselected paths on the finer lattice. The weights of the different paths are either predetermined or one tries to optimize them by means of operator matching techniques [16]. Mack [18], Hulsebos et al. [7] and recently Brower et al. [8], made a proposal for an RGT in gauge theories which differs from previous approaches. This new idea is meant to separate high- and low-momentum parts and to construct block spins which contain only the lowest frequency part of the fields on the finer lattice.

Apart from their intended use in MGMC simulations the kernels $\mathcal{A}$ and C could be used also in MG computations of gauge-covariant propagators, as in the works of Brower et al. [8], and Ben-Av et al. [12]. This is of interest for fermions primarily, and will be studied elsewhere. In addition the kernels $\mathcal{A}$ and C could be useful in ordinary Monte Carlo RG investigations and in studying dielectric lattice gauge theories [19] numerically. It may supplement the work of Pirner and collaborators [15,16,20] who have studied the properties of dielectric gauge models extensively. The block spin computation yields auxiliary quantities of interest, such as the lowest eigenvalues λ_0, λ_0' of the negative gauge-covariant laplacian with Neumann and Dirichlet boundary conditions on block boundaries. λ_0 and λ_0' are measures of disorder and their renormalization group flow is therefore instructive.

In sect. 2 a definition of kernels $\mathcal{A}$ and C for Higgs fields which is in the spirit of refs. [7,8], and definitions of blocked gauge fields which use these kernels are given. This provides a systematic method for non-abelian gauge theories in any number of space-time dimension. An analytical expression for C as given in ref. [8] for two-dimensional U(1) theories cannot be found for non-abelian gauge theories in higher dimensions. A numerical solution of this problem is given by presenting an algorithm which permits to compute the gauge field dependent kernels C and $\mathcal{A}$ very efficiently (sect. 3). The actual computations are done for SU(2) gauge fields on a four-dimensional lattice. A discussion of the performance of the algorithm (which is interesting for

further development of projective MG methods) and physics results which are relevant for colour dielectric models and ordinary RG studies are given in sect. 4. The algorithm vectorizes very well. To compute the kernel C or A on the whole fundamental lattice costs CPU time of the order of one standard Monte Carlo sweep for the gauge field. The kernels on the coarser lattices are computed in the same way, using block gauge fields. The cost in CPU for this is negligible in comparison, because there are much fewer points on block lattices.

2. Renormalization group transformations for gauge theories

Let us consider real space RGTs with some integer scale factor $L > 1$ for a euclidean gauge theory on a fundamental d-dimensional space-time lattice Λ with lattice spacing a. As a concrete model we will take the SU(2) Higgs model in $d = 4$ into consideration; for recent reviews see ref. [21]. In the first RG step the fundamental lattice of sites z is covered by a block lattice Λ' consisting of block sites x with side length $a' = La$ and containing L^d sites z. We write $z \in x$, if z is in block x. One site z in each block x is selected as the block centre $\hat{x}$. The procedure is iterated in order to go to coarser lattices.

2.1. GAUGE-COVARIANT BLOCK AVERAGES

The theory on the fundamental lattice exhibits local gauge invariance, i.e. the action $\mathcal{H}(\phi, u)$ of a Higgs field ϕ on the sites of Λ and a gauge field u on the links of Λ is invariant under

$$\phi(z) \to \phi'(z) = g_z \phi(z), \tag{2.1}$$

$$u(z + \hat{\mu}, z) \to u'(z + \hat{\mu}, z) = g_{z+\hat{\mu}} u(z + \hat{\mu}, z) g_z^{-1}, \tag{2.2}$$

where g_z are elements of some unitary gauge group G; $\hat{\mu}$ denotes a vector of length a in μ-direction, and $(z + \hat{\mu}, z)$ is the directed link emerging from site z in μ-direction.

In order to define block spins $\Phi = \Phi(\phi, u)$ and $U = U(u) \in \mathrm{G}$ which live on the sites resp. links of Λ', we must specify a block transformation in a gauge-covariant way. One demands that the block spins transform under the gauge transformations (2.1), (2.2) on Λ according to

$$\Phi(x) \to \Phi'(x) = g_{\hat{x}} \Phi(x), \tag{2.3}$$

$$U(x + L\hat{\mu}, x) \to U'(x + L\hat{\mu}, x) = g_{\hat{x}+L\hat{\mu}} U(x + L\hat{\mu}, x) g_{\hat{x}}^{-1}. \tag{2.4}$$

A discussion about exact RGTs for gauge theories and some proposals of different authors can be found in ref. [22].

The block spin Φ shall be a linear function of ϕ. It is obtained by

$$\Phi(x) = (C\phi)(x) = \int_{z\in\Lambda} C(x,z)\phi(z)\,, \tag{2.5}$$

where $C(x,z)$ is the integral kernel of the operator C which maps a field on Λ into a field on Λ'. The lattice notation $\int_{z\in\Lambda}$ means $a^d\sum_{z\in\Lambda}$. Gauge covariance demands that $C(x,z)$ transforms under a gauge transformation like a parallel transporter from z to $\hat{x}$; i.e.

$$C(x,z) \to g_{\hat{x}}\, C(x,z)\, g_z^{-1}\,. \tag{2.6}$$

Similarly, the background field ψ shall be a linear function of the block spin Φ,

$$\psi(z) = \int_{x\in\Lambda'} \mathcal{A}(z,x)\Phi(x) \equiv (a')^d \sum_{x\in\Lambda'} \mathcal{A}(z,x)\Phi(x)\,. \tag{2.7}$$

The background field should transform in the same way as the field ϕ under gauge transformations, therefore

$$\mathcal{A}(z,x) \to g_z\, \mathcal{A}(z,x)\, g_{\hat{x}}^{-1}\,. \tag{2.8}$$

Given a path $\mathcal{C}$ on Λ from z_1 to z_2, the parallel transporter $u(\mathcal{C})$ is defined as usual as path-ordered product of the lattice gauge fields $u(b)$ on the links b of which the path $\mathcal{C}$ is composed. We seek kernels C, $\mathcal{A}$ which take their values in the linear span of the gauge group, i.e. which are real multiples of SU(2) matrices when G = SU(2). The most general kernel C which enjoys gauge covariance (2.6) is a weighted average of parallel transporters $u(\mathcal{C})$ over paths $\mathcal{C}$ from z to $\hat{x}$, with arbitrary real weight factors $\rho(\mathcal{C})$, and similarly for $\mathcal{A}$. Balaban's block spin choice amounts to computing C by summing over all taxi driver paths $\mathcal{C}$ which stay inside the block x and proceed in each direction at most once. This is equivalent to averaging over the block in the lattice version of the Schwinger–Fock (= radial) gauge.

2.2. DEFINITION OF BLOCK SPINS FOR HIGGS AND GAUGE FIELDS

It is appropriate to review briefly the block spin transformation which was used by Kupiainen and Gawędzki [5] to give rigorous proofs of the existence of the continuum limit of some lattice field theories without gauge fields. The proposal made below is the natural gauge-covariant generalization of this transformation. The kernel C is given by $C(x,z) = (a')^{-d}\chi_x(z)$, where χ_x is the characteristic function of x, i.e. $\chi_x(z) = 1$ if $z \in x$ and $\chi_x(z) = 0$

otherwise. This kernel implies a local definition of Φ as block average of ϕ and fulfills the normalization condition

$$CC^*(x_1, x_2) = \int_z C(x_1, z) C^*(z, x_2) = (a')^{-d} \delta_{x_1 x_2} = \delta(x_1 - x_2) . \quad (2.9)$$

The kernel of C resp. its adjoint C^* — which is the same as C for pure scalar theory, generally $C^*(z, x) = C(x, z)^*$ where star denotes hermitian conjugation — can be defined as solution of the eigenvalue equation

$$(-\Delta_{\mathrm{N},x} C^*)(z, x) = \lambda_0(x) C^*(z, x) , \quad (2.10)$$

where $-\Delta_{\mathrm{N},x}$ is the negative lattice laplacian with Neumann boundary conditions on the boundary of block x, and $\lambda_0(x)$ is the lowest eigenvalue, equal to zero for all blocks in pure scalar theory. $\Delta_{\mathrm{N},x}$ acts on argument z. Solutions of eq. (2.10) with $\lambda_0(x) = 0$ are constants on the blocks, these are determined by the normalization condition (2.9).

Following Mack [18], this procedure can be re-interpreted as follows. Define a notion of (block-local) frequency which depends only on the behaviour of the function which is to be decomposed into frequency components on the chosen block x. Thus, define frequency (squared if one wants to) as eigenvalue of the 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

$$\psi(z) = \lim_{t \to \infty} \exp\left[-t(-\Delta_{\mathrm{N},x} - \lambda_0(x))\right] \phi(z) . \quad (2.11)$$

$\psi(z)$ 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 centre $z = \hat{x}$.

The euclidean action of ϕ^4 theory is of the form $\mathcal{H} = \mathcal{H}_0 + V$, and $\mathcal{H}_0 = (\phi, h\phi)$ with "first quantized hamiltonian" h in the space of wave functions ϕ equal to $h = -\Delta$. A local approximation of this which preserves the invariance under shifts by constant fields would be $h_x = -\Delta_{\mathrm{N},x}$. One sees that the block spin $\Phi(x)$ retains the component of ϕ associated with the lowest eigenvalue of h_x, in agreement with the philosophy of refs. [7,8].

Mack [18] proposed to proceed for block spins of Higgs fields in non-abelian gauge theories in exactly the same way, defining C^* as solution of eq. (2.10). $C^*(z, x)$ will be in the linear span * of the gauge group G, and $\Delta_{\mathrm{N},x}$ will be

*That is, $C^*(z, x)$ 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), $N \geqslant 3$ etc, see ref. [19].

the covariant laplacian with Neumann boundary conditions. It depends on the lattice gauge field u and is defined by

$$(\Delta_{\mathrm{N},x}\phi)(z) = \sum_{\substack{z' \text{n.n. } z \\ z' \in x}} [u(z,z')\phi(z') - \phi(z)] \quad \text{for} \quad z \in x\,. \tag{2.12}$$

Summation is over next neighbours z' of z which lie in block x, and $u(z,z')$ is the gauge field attached to the link from z' to z. The idea is again that the definition of block spins involves dynamical information — think of $\int_{z\in x} \phi^*(z)(-\Delta_{\mathrm{N},x}\phi)(z)$ 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 [19] u, especially on coarser layers *.

Some properties of $\Delta_{\mathrm{N},x}$ are: $-\Delta_{\mathrm{N},x}$ is positive semidefinite so that its lowest eigenvalue $\lambda_0(x)$ is always non-negative; $\lambda_0(x)$ is only zero iff u is a pure gauge within x; under gauge transformations (2.1), (2.2) the kernel of the covariant laplacian Δ transforms according to $\Delta(z,z') \to g_z \Delta(z,z') g_{z'}^{-1}$; eigenvectors of $-\Delta_{\mathrm{N},x}$ transform as (2.1) (modulo arbitrariness stated below), and the eigenvalues of $-\Delta_{\mathrm{N},x}$ are gauge invariant.

The normalization condition (2.9) does not specify C uniquely. One retains the freedom of a gauge transformation per block $C^*(z,x) \to C^*(z,x) g_x$, $g_x \in \mathrm{G}$, under which eq. (2.9) is invariant. This freedom is eliminated by demanding that for $\mathrm{G} = \mathrm{U}(1), \mathrm{SU}(2)$

$$C(x,\hat{x}) = C^*(\hat{x},x) = r(x)\,\mathbf{1}, \tag{2.13}$$

with $r(x)$ a positive real number. This ensures the right gauge-covariance property, viz. $C^*(z,x) \to g_z\, C^*(z,x)\, g_{\hat{x}}^{-1}$. For other G, $\mathbf{1}$ on the r.h.s. of eq. (2.13) is to be replaced by a positive hermitian matrix. If u is pure gauge within a block x, $(a')^d C(x,z)$ is an element of G for all $z \in x$, equal to the path-ordered product of link variables along an arbitrary path $\mathcal{C}_{\hat{x}z}$ which stays entirely within x.

Now we turn to the kernel $\mathcal{A}$. In the work of Kupiainen and Gawędzki, the kernel $\mathcal{A}$ is defined so that $\psi = \mathcal{A}\Phi$ minimizes the kinetic energy $\mathcal{H}_0$ subject to the constraint that the block spin is prescribed. One defines the fluctuation field ζ, which is supposed to contain the high frequency components of ϕ, by $\phi = \psi + \zeta$. The above choice of $\mathcal{A}$ amounts to requiring that the free hamiltonian $\mathcal{H}_0$ does not couple high and low frequency,

$$\begin{aligned}(\phi, -\Delta\phi) &= (\psi, -\Delta\psi) + (\zeta, -\Delta\zeta) \\ &= (\Phi, -\Delta_{\mathrm{eff}}\Phi) + (\zeta, -\Delta\zeta)\,,\end{aligned} \tag{2.14}$$

*Although we do not indicate notationally the explicit dependence of $\Delta_{\mathrm{N},x}$, $\lambda_0(x)$, C and C^* on the gauge field u, it should be remembered in the following.

where Δ_{eff} is an operator which acts on functions on Λ' and has kernel

$$\Delta_{\text{eff}}(x_1, x_2) = (\mathcal{A}^* \Delta \mathcal{A})(x_1, x_2) = \int_z \mathcal{A}(z, x_1)^* (\Delta \mathcal{A})(z, x_2) . \quad (2.15)$$

When defined in this way, the kernel $\mathcal{A}$ will have exponential tails. The decay length of these tails is about one block lattice spacing a' [5]. For numerical work this is inconvenient. Therefore, kernels $\mathcal{A}$ were used in the numerical multigrid work of Mack and Meyer [9] where $\mathcal{A}(z, x)$ was approximately equal to the kernel defined above, but supported as a function of z in a neighbourhood of block x.

As argued in ref. [9], it is essential for the fighting of critical slowing down in MGMC that kernel $\mathcal{A}$ should be as smooth as possible. This contention is supported by the recent work of Hasenbusch et al. [23]. In theories with gauge fields one may have to constrain the support of $\mathcal{A}(z, x)$ to one block, viz. $z \in x$, for reasons of computational complexity. A natural choice is then to compute $\mathcal{A}$ as solution of the eigenvalue equation

$$(-\Delta_{\mathrm{D},x} \mathcal{A})(z, x) = \lambda_0'(x) \mathcal{A}(z, x) , \quad (2.16)$$

where $-\Delta_{\mathrm{D},x}$ is the negative gauge-covariant laplacian with Dirichlet boundary conditions on the boundary of x, and $\lambda_0'(x)$ is the lowest eigenvalue. In d dimensions

$$(\Delta_{\mathrm{D},x} \phi)(z) = -2d\phi(z) + \sum_{\substack{z' \text{ n.n. } z \\ z' \in x}} u(z, z') \phi(z') \quad \text{for } z \in x . \quad (2.17)$$

This leaves the same freedom as for $C^*(z, x)$; this freedom is eliminated by the requirement $C\mathcal{A} = \mathbf{1}$. If $\mathcal{A}$ is supported on one block, this requirement reduces to

$$\int_{z \in x} C(x, z) \mathcal{A}(z, x) = (a')^{-d} \mathbf{1} . \quad (2.18)$$

The kernel $\mathcal{A}$ can be computed in the same way as C and with the same efficiency. The change of boundary conditions amounts to changing one line in the computer programme.

Reading Δ as covariant laplacian, eq. (2.15) defines a gauge-covariant "effective laplacian" on the block lattice. Given the support properties of $\mathcal{A}$, it is nonzero only if either $x_1 = x_2$ or x_2 is a nearest neighbour of x_1. (Because of disorder in the gauge field, the "effective laplacian" may behave more like a laplacian minus mass squared, multiplied with a constant.) A blocked gauge field can now be defined by

$$\tilde{U}(x_1, x_2) = \Delta_{\text{eff}}(x_1, x_2) , \quad (2.19)$$

x_2 nearest neighbour of x_1. Actually, this is still a "dielectric" gauge field. It may be brought back into SU(2) by multiplying with a suitable positive number. (More generally one would have to perform a polar decomposition [19].)

The choice (2.16) of $\mathcal{A}$ minimizes the local approximation $(\psi, -\Delta_{D,x}\psi)$ to the kinetic energy of the Higgs field, subject to the constraint that the block spin is prescribed.

The reason why one wants a smooth kernel $\mathcal{A}$ is simple. Consider a cube $x \in \Lambda_{N-k}$ of side length la, $l = L^k$. Using a step function $\mathcal{A}$, updating $\Phi^k(x)$ costs kinetic energy proportional to the surface l^{d-1} of the cube, compared to the optimum l^{d-2} for a smooth kernel $\mathcal{A}$ which interpolates between Λ_N and Λ_{N-k} so that $\nabla\mathcal{A} = O(1/la)$. This increase by a factor l will lead to unacceptable low acceptance rates on coarse lattices where l is large.

Another possible definition of (dielectric) block spins for gauge fields, which is similar (and approximately equal) to Balaban's choice, involves C and reads

$$\tilde{U}(x + L\hat{\mu}, x) = L^{-d} \sum_{z \in x} \hat{C}(x + L\hat{\mu}, z + L\hat{\mu})\, u(z + L\hat{\mu}, z)\, \hat{C}^*(z, x), \tag{2.20}$$

where $\hat{C}(x, z)$ is $C(x, z)$ without the factor $(a')^{-d}\chi_x(z)$, and $u(z + L\hat{\mu}, z)$ is the path-ordered product of link variables on Λ along the straight line from z to $z + L\hat{\mu}$.

For G = SU(2) the polar decomposition of $\tilde{U}$ reads

$$\tilde{U}(x + L\hat{\mu}, x) = \sigma_\mu(x)\, U(x + L\hat{\mu}, x), \tag{2.21}$$

where

$$\begin{aligned} \sigma_\mu(x) &= \left[\tfrac{1}{2}\mathrm{Tr}\, \tilde{U}(x + L\hat{\mu}, x)^* \tilde{U}(x + L\hat{\mu}, x)\right]^{1/2}, \\ U(x + L\hat{\mu}, x) &\in \mathrm{SU}(2). \end{aligned} \tag{2.22}$$

$\sigma_\mu(x)$ is a gauge-invariant quantity.

3. Numerical algorithms for the computation of block spins

In this section the algorithms used for computing the kernels C and $\mathcal{A}$ numerically are described. The algorithms will be explained for the C-kernel, but it is obvious which replacements and modifications have to be done for the $\mathcal{A}$-kernel.

The eigenvalue problem (2.10) 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 a conjugate gradient method. This method has proven to be well suited for large but sparse linear

systems with a positive matrix of coefficients. For instance it is generally accepted to be the most effective way for inversion of the fermionic matrix in hitherto simulations [24]. 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 $-\Delta_{\mathrm{N},x}$, but we must demand that it is not greater than $\lambda_0(x)$. An efficient method for bounding eigenvalues from below is explained in the last part of this section.

3.1. DETERMINATION OF EIGENVECTORS BY INVERSE ITERATION

Inverse iteration is a method by which one can determine all eigenvalues and eigenvectors of a quadratic matrix* 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, fulfilling $Ax_j = \lambda_j x_j$, one starts with an arbitrary vector $x^{(0)}$ and computes a sequence of vectors $x^{(1)}, x^{(2)}, \ldots$, according to

$$(A - \lambda)x^{(i)} = x^{(i-1)}, \; i = 1, 2, 3, \ldots . \tag{3.1}$$

The iteration prescription (3.1) yields an eigenvector x_j in the limit of large i. This can be easily seen by expanding the $x^{(i)}$ as linear combinations of the eigenvectors** of A. See refs. [25–27] for a more comprehensive discussion on the method of inverse iteration.

We shall modify the rudimentary method (3.1) 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. (2.10) 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

$$\left(-\Delta_{\mathrm{N},x} - \lambda^{(i)}(x)\right) V_x = C_x^{*(i-1)}, \tag{3.2}$$

$$C_x^{*(i)} = \frac{V_x}{\|V_x\|_x} \tag{3.3}$$

*The matrix need not be normal or normalizable.

**If A is defective one expands in terms of eigenvectors and principal vectors [25,26].

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

$$\langle V_1, V_2\rangle_x = L^{-d}\sum_{z\in x}\frac{1}{N}\mathrm{Tr}\, V_1(x,z)^* V_2(x,z)\,, \tag{3.4}$$

if $V_k(x,z)$, $k = 1,2$, are $N \times N$ matrices. Proper choices for $\lambda^{(1)}(x)$ and for $C_x^{*(0)}$ are discussed in subsects. 3.2 and 4.2. The updating procedure $\lambda^{(i)}(x) \to \lambda^{(i+1)}(x)$ is explained in subsect. 3.3. The solution of the linear system of equations (3.2) is determined by means of a conjugate gradient method which is outlined next.

3.2. CONJUGATE GRADIENT METHOD

The conjugate gradient (CG) method, due to Hestenes and Stiefel, is an efficient way of solving a system of linear equations

$$Ty = b\,, \tag{3.5}$$

where T is a positive definite hermitian $n \times n$ matrix in a Hilbert space with scalar product $\langle\cdot,\cdot\rangle$. The CG method is iterative in the sense that starting with an arbitrary initial vector $x^{(0)}$, it yields a sequence of vectors $x^{(0)} \to x^{(1)} \to x^{(2)}\cdots$ which converge to the solution of (3.5). But unlike other iterative methods — e.g. Jacobi, Gauss–Seidel or over-relaxation — it arrives at the solution after at most n steps (provided the arithmetic is exact).

General CG methods are designed for finding stationary points of continuously differentiable functions [28,27]. Hestenes's and Stiefel's method minimizes the quadratic form

$$F[y] = \tfrac{1}{2}\langle Ty - b\,,\, T^{-1}(Ty-b)\rangle = \tfrac{1}{2}\langle y\,,\, Ty\rangle - \mathrm{Re}\,\langle b\,,\, y\rangle + \tfrac{1}{2}\langle b\,,\, T^{-1}b\rangle\,. \tag{3.6}$$

The solution of eq. (3.5) is the minimum of F, because T^{-1} exists and is also positive definite.

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

$$F[x^{(k+1)}] = \min_{v_0,\ldots,v_k} F[x^{(k)} + v_0 r^{(0)} + \cdots + v_k r^{(k)}]\,, \tag{3.7}$$

where $r^{(i)} = b - Tx^{(i)}$ for $i \leqslant k$. The $r^{(i)}$ are orthogonal and thus independent, as long as $r^{(k)} \neq 0$. Since at most n vectors are independent, there must be an $l \leqslant n$ with $r^{(l)} = 0$ and $x^{(l)}$ solves eq. (3.5).

The cookbook recipe of the CG method is the following [25].

(1) Choose any vector $x^{(0)}$ and set $p^{(0)} = r^{(0)} = b - Tx^{(0)}$, $k = 0$.

(2) If $r^{(k)} = 0 \Leftrightarrow p^{(k)} = 0$: STOP, $x^{(k)}$ is the solution of $Tx = b$.

(3) Else compute

$$a_k = \frac{\langle r^{(k)}, r^{(k)} \rangle}{\langle p^{(k)}, Tp^{(k)} \rangle} \equiv \frac{\langle r^{(k)}, p^{(k)} \rangle}{\langle p^{(k)}, Tp^{(k)} \rangle}, \qquad x^{(k+1)} = x^{(k)} + a_k p^{(k)},$$

$$r^{(k+1)} = r^{(k)} - a_k T p^{(k)}, \qquad b_k = \frac{\langle r^{(k+1)}, r^{(k+1)} \rangle}{\langle r^{(k)}, r^{(k)} \rangle},$$

$$p^{(k+1)} = r^{(k+1)} + b_k p^{(k)}, \tag{3.8}$$

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

See refs. [25,26] for properties of the $x^{(k)}, p^{(k)}, r^{(k)}$, for a proof of convergence of the CG method and for a discussion about its numerical properties.

When applying the CG algorithm for solving eq. (3.2), we have $T = T_x = -\Delta_{\mathrm{N},x} - \lambda^{(i)}(x)$; $y = y_x = V_x$, and $b = b_x = C_x^{*(i-1)}$ are $N \cdot L^d \times N$ matrices (for a gauge group represented by $N \times N$ matrices), and the scalar product $\langle \cdot, \cdot \rangle = \langle \cdot, \cdot \rangle_x$ is given by (3.4). On a vector computer eq. (3.2) is solved for all blocks in parallel by vectorizing over the block index x.

In order that T_x is positive definite, all $\lambda^{(i)}(x)$ must be less than the smallest eigenvalue $\lambda_0(x)$ of $-\Delta_{\mathrm{N},x}$. Since $-\Delta_{\mathrm{N},x}$ is positive semidefinite, we can for example choose $\lambda^{(i)}(x) = -\varepsilon$ with a small positive ε. But below we will see that we can do better.

3.3. BOUNDING AND UPDATING EIGENVALUES

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. (3.2) for $i = 1$, we obtain an estimate for $\lambda_0(x)$ by computing the Rayleigh quotient of V_x which is defined through

$$R[V_x] = \frac{\langle V_x, -\Delta_{\mathrm{N},x} V_x \rangle_x}{\langle V_x, V_x \rangle_x}. \tag{3.9}$$

In numerical analysis the Rayleigh quotient is the standard estimate for an eigenvalue from an eigenvector approximation [29]. R has the property that the eigenvalues of $-\Delta_{\mathrm{N},x}$ are its stationary points which are obtained when the corresponding eigenvectors are plugged in. One has $R[V_x] = \lambda_0(x) + \mathrm{O}(\delta^2)$,

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 $R[V_x] \geqslant \lambda_0(x)$.

In principle there is a simple way to circumvent this problem. One could take a variant of the CG method [25] which is also applicable in cases where the matrix T is not hermitian and which solves $T^*Ty = T^*b$ rather than eq. (3.5). The matrix T^*T is positive if T is nonsingular, which means for $T_x = -\Delta_{N,x} - \lambda^{(i)}(x)$ that one must only demand $\lambda^{(i)}(x) \neq \lambda_0(x)$ (more generally $\lambda^{(i)}(x) \notin$ spectrum of $-\Delta_{N,x}$). But the variant of the CG method is substantially slower, because in each iteration two matrix multiplications instead of one have to be performed. Also, it is not a priori clear that a trial $\lambda^{(i+1)}(x) = R[V_x] < \frac{1}{2}(\lambda_0(x) + \lambda_1(x))$, which is a necessary condition for convergence of the inverse iteration towards C_x^*.

A way out of the dilemma is furnished by a bounding theorem for eigenvalues which is known as

Temple's Theorem [30]. Let T be a self-adjoint operator in a Hilbert space H with scalar product $\langle \cdot, \cdot \rangle$. Let $f^{(0)}$ and $f^{(1)}$ be two elements of $H \setminus \{0\}$ with $Tf^{(1)} = f^{(0)}$. Compute the Schwarz constants

$$a_0 = \langle f^{(0)}, f^{(0)} \rangle, \quad a_1 = \langle f^{(0)}, f^{(1)} \rangle, \quad a_2 = \langle f^{(1)}, f^{(1)} \rangle \tag{3.10}$$

and for a real number t the Temple quotient

$$\vartheta(t) = \frac{a_0 - ta_1}{a_1 - ta_2} . \tag{3.11}$$

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

$$\vartheta(q) \leqslant \lambda \leqslant \vartheta(p) . \tag{3.12}$$

See ref. [30] for a proof.

A corollary 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 $\mu_2 < l_1 \leqslant \lambda_1$. Then Temple's theorem is used with the interval $(-\infty, l_1)$ and one obtains

$$\mu_2 \frac{\mu_1 - l_1}{\mu_2 - l_1} = \mu_2 - \frac{\mu_1 - \mu_2}{(l_1/\mu_2) - 1} \leqslant \lambda_0 \leqslant \mu_2 , \tag{3.13}$$

where $\mu_1 \equiv a_0/a_1$.

When we apply Temple's theorem to the inverse iteration (3.2), we get the bounds

$$\begin{aligned}\tau(l_1^{(i)}(x)) &\equiv \lambda^{(i)}(x) + \mu_2^{(i)}(x)\frac{\mu_1^{(i)}(x) - l_1^{(i)}(x)}{\mu_2^{(i)}(x) - l_1^{(i)}(x)} \\ &\leqslant \lambda_0(x) \leqslant \lambda^{(i)}(x) + \mu_2^{(i)}(x) = R[V_x], \end{aligned} \tag{3.14}$$

where

$$\begin{aligned}&\mu_1^{(i)}(x) = \frac{1}{\langle C_x^{*(i-1)}, V_x\rangle_x}, \quad \mu_2^{(i)}(x) = \frac{\langle C_x^{*(i-1)}, V_x\rangle_x}{\langle V_x, V_x\rangle_x}, \\ &l_1^{(i)}(x) = \tilde{\lambda}_1(x) - \lambda^{(i)}(x), \quad \tilde{\lambda}_1(x) \leqslant \lambda_1(x). \end{aligned} \tag{3.15}$$

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. [29] 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 [31] or by means of a Ritz method. For $-\Delta_{N,x}$ and $-\Delta_{D,x}$ (and gauge group $G = SU(2)$) there is a simpler solution as will be shown in sect. 4. 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.

(i) If $\mu_2^{(i)}(x) \geqslant l_1^{(i)}(x)$, then $\lambda^{(i+1)}(x) = \lambda^{(i)}(x)$.

(ii) Else compute $\tau(l_1^{(i)}(x))$ and set $\lambda^{(i+1)}(x) = \max\left(\lambda^{(i)}(x), \tau(l_1^{(i)}(x))\right)$. (3.16)

4. Results for pure SU(2) gauge theory

The performance of the numerical algorithms for computing the kernels C and $\mathcal{A}$, and some properties of, and relevant for, block gauge fields (2.19), (2.20) were investigated for pure SU(2) gauge theory in $d = 4$ dimensions with Wilson action

$$\mathcal{H}(u) = \beta \sum_p [1 - \tfrac{1}{2}\mathrm{Tr}\, u(p)], \quad \beta = 4/g^2. \tag{4.1}$$

The Monte Carlo method implemented for updating the gauge field was Creutz's heat bath algorithm. A scale factor of $L = 3$ was chosen.

4.1. LOW-LYING SPECTRA OF $-\Delta_{N,x}$ AND $-\Delta_{D,x}$ ON A FUNDAMENTAL 3^4 LATTICE

The low-lying spectra of $-\Delta_{N,x}$ and $-\Delta_{D,x}$ are of importance for the performance of inverse iteration. As a rule of thumb one can say that the fewer inverse iterations will be needed the better the second lowest eigenvalues $\lambda_1(x)$, $\lambda_1'(x)$ are separated from the lowest ones $\lambda_0(x)$, $\lambda_0'(x)$.

In case of pure gauges the spectra of $-\Delta_{N,x}$ and $-\Delta_{D,x}$ are the same for any gauge group G, except an (additional) N-fold degeneracy of each eigenvalue if G is represented by $N \times N$ matrices. The spectra of $-\Delta_{N,x}$ and $-\Delta_{D,x}$ without gauge field are summarized in tables 1 and 2 for block size $L = 3$ in $d = 4$ dimensions. For arbitrary gauge field configurations the shape of the spectra will change, but the sums of all eigenvalues are invariant and equal always $N \cdot 2d\,L^{d-1}(L-1)$ and $N \cdot 2d\,L^d$, respectively.

The distribution of the five lowest eigenvalues of $-\Delta_{N,x}$ and $-\Delta_{D,x}$, modulo two-fold degeneracy for G = SU(2), was determined on a fundamental 3^4 lattice with periodic boundary conditions for β between 0 and 3.9. This investigation had to be done only once, so that no special routines were written for this purpose but existing libraries [32] were used. One thousand sweeps with measurements were performed after discarding 1000 sweeps to ensure thermalization. The results are shown in figs. 1 and 2. It is interesting to note that $\lambda_0(x)$ and $\lambda_0'(x)$ behave similar to the plaquette energy $\langle 1 -$

TABLE 1

The spectrum of $-\Delta_{N,x}$ in $d = 4$ dimensions without gauge field for block size $L^d = 3^4$. The sum of all eigenvalues equals $\mathrm{Tr}(-\Delta_{N,x})$ which is $2d\,L^{d-1}(L-1) = 432$ [twice the number of links within the block]

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 2

The spectrum of $-\Delta_{D,x}$ in $d = 4$ dimensions without gauge field for block size $L^d = 3^4$. The sum of all eigenvalues equals $\mathrm{Tr}(-\Delta_{D,x})$ which is $2d\,L^d = 648$

Eigenvalue	Degeneracy
$8 - 8\cos(\pi/4) = 2.3431$	1
$8 - 6\cos(\pi/4) = 3.7574$	4
$8 - 4\cos(\pi/4) = 5.1716$	10
$8 - 2\cos(\pi/4) = 6.5858$	16
8.0000	19
$8 - 2\cos(3\pi/4) = 9.4142$	16
$8 - 4\cos(3\pi/4) = 10.8284$	10
$8 - 6\cos(3\pi/4) = 12.2426$	4
$8 - 8\cos(3\pi/4) = 13.6569$	1

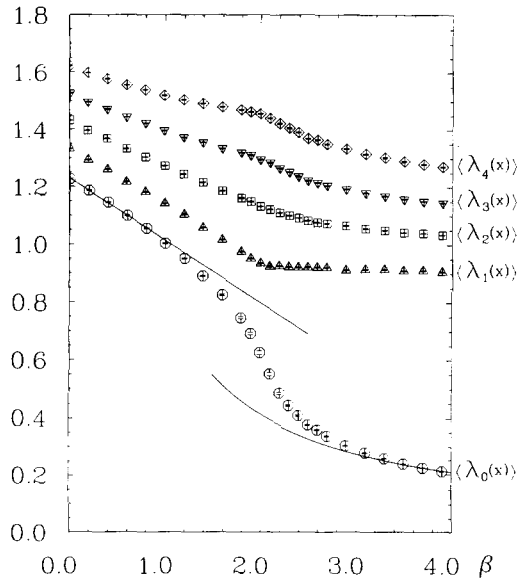

Fig. 1. Expectation values of the five lowest eigenvalues (modulo two-fold degeneracy) of $-\Delta_{N,x}$ for gauge group SU(2) in pure gauge theory on a fundamental 3^4 lattice as a function of β. In pure gauge, $\lambda_0(x) = 0$. The indicated strong and weak coupling expansion of $\langle\lambda_0(x)\rangle$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

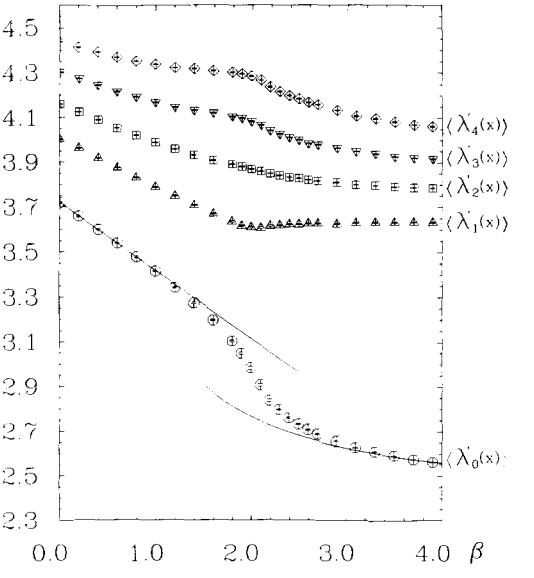

Fig. 2. Expectation values of the five lowest eigenvalues (modulo two-fold degeneracy) of $-\Delta_{D,x}$ for gauge group SU(2) in pure gauge theory on a fundamental 3^4 lattice as a function of β. In pure gauge, $\lambda'_0(x) = 2d(1 - \cos(\pi/(L+1))) = 2.3431$. The indicated strong and weak coupling expansion of $\langle\lambda'_0(x)\rangle$ derived from the numerical data is $3.7199(6) - \beta \cdot 0.3013(15)$ and $2.3431 + \beta^{-1} \cdot 0.853(7)$, respectively. (Note that the ordinate does not start with zero.)

TABLE 3

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

β	$\langle\lambda_0(x)\rangle$ on fund. 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 4

Finite-size effect: Expectation value of the lowest eigenvalue $\lambda'_0(x)$ of $-\Delta_{D,x}$ on a fundamental 3^4 lattice and on the 81 blocks of size 3^4 embedded in a 9^4 lattice

β	$\langle\lambda'_0(x)\rangle$ on fund. 3^4 lattice	$\langle\lambda'_0(x)\rangle$ on 3^4 blocks in a 9^4 lattice
1.8	3.1064 ± 0.0040	3.1066 ± 0.0014
1.9	3.0505 ± 0.0048	3.0591 ± 0.0011
2.0	2.9850 ± 0.0068	3.0024 ± 0.0012
2.1	2.9083 ± 0.0058	2.9463 ± 0.0016
2.2	2.8414 ± 0.0048	2.8843 ± 0.0010
2.3	2.7996 ± 0.0025	2.8254 ± 0.0010
2.4	2.7624 ± 0.0031	2.7797 ± 0.0009
2.5	2.7337 ± 0.0018	2.7452 ± 0.0007
2.6	2.7105 ± 0.0016	2.7186 ± 0.0006
2.7	2.6906 ± 0.0016	2.6977 ± 0.0007

$\frac{1}{2}\mathrm{Tr}\,u(p)\rangle$ which has strong coupling expansion $1 - \beta/4 + O(\beta^3)$ and weak coupling expansion $3/(4\beta)$ [33], and which exhibits a crossover from strong to weak coupling at β around 2.0. This crossover manifests itself also in the behaviour of $\langle\lambda_0(x)\rangle$, $\langle\lambda'_0(x)\rangle$ and through a kind of blowing up of the band width of $\lambda_1(x),\ldots,\lambda_4(x)$ and $\lambda'_1(x),\ldots,\lambda'_4(x)$, respectively. For larger values of β the situation of pure gauges is approached and $\lambda_0(x)$, $[\lambda'_0(x)]$ tends to zero $[2d(1 - \cos(\pi/(L+1))) = 2.3431]$, while $\lambda_1(x),\ldots,\lambda_4(x)$ $[\lambda'_1(x),\ldots,\lambda'_4(x)]$ become degenerate with value $2(1-\cos(\pi/L)) = 1$ $[2d - 2(d-1)\cos(\pi/(L+1)) - 2\cos(2\pi/(L+1)) = 3.7574]$.

When one determines $\langle\lambda_0(x)\rangle$ and $\langle\lambda'_0(x)\rangle$ on the 81 blocks of a 9^4 lattice one observes some finite-size effects as shown in tables 3 and 4.

4.2. PERFORMANCE OF THE NUMERICAL ALGORITHMS

Figs. 1 and 2 are very promising, because we are mainly interested in the region of larger β-values where a distinct gap exists between $\lambda_0(x)$ and $\lambda_1(x)$, and $\lambda_0'(x)$ and $\lambda_1'(x)$, respectively. This permits to compute the kernels with only two inverse iterations as will be shown, at least for $\beta \geqslant 1.8$ where numerical investigations were performed.

From the measured distribution of eigenvalues one can extract lower bounds for $\lambda_0(x)$, $\lambda_0'(x)$ and $\lambda_1(x)$, $\lambda_1'(x)$ which have practically a confidence level of 100%. These bounds are used for $\tilde{\lambda}_1(x)$ and $\lambda^{(1)}(x)$. In course of the inverse iteration (3.2), $\tilde{\lambda}_1(x)$ will not be changed and will retain its x-independent value *, whereas $\lambda^{(i)}(x)$ is updated according to (3.16) and is only x-independent for $i = 1$.

In order to be specific in the sequel, we will again discuss items for kernel C. To start inverse iteration we must choose some normalized $C_x^{*(0)}$. Two different choices were tested. The first one is a trivial initialization $(a')^d C^{*(0)}(z,x) = \mathbf{1}$ for all $z \in x$. The second choice is Balaban's radial gauge, 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 projected into SU(2) (cf. eqs. (2.21) and (2.22)). (In case of $\mathcal{A}$ one could think of multiplying with the pure gauge value $(2L/(L+1))^{d/2} \prod_{\mu=1}^{d} \sin(\pi z_\mu/(L+1))$, but it turns out that the efficiency of the algorithm is not enhanced significantly by this.)

By computing the Rayleigh quotient of $C_x^{*(0)}$ one sees that the component proportional to the sought eigenvector C_x^* is clearly dominating in Balaban'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 indeed that it is more efficient to use Balaban's gauge for $C_x^{*(0)}$.

The CG method for solving (3.2) was used with initializing V_x with zero. In this case one needs less than 27 (29) CG iterations in the first (second) inverse iteration to solve eq. (3.2) sufficiently. The CG algorithm was stopped when the residual of all blocks fulfilled $||(-\Delta_{\mathrm{N},x} - \lambda^{(i)}(x))V_x - C_x^{*(i-1)}||^2 < L^{-d} \cdot 10^{-6}$. 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 $R[C_x^{*(i)}] = \langle C_x^{*(i)}, -\Delta_{\mathrm{N},x} C_x^{*(i)} \rangle$ and comparing with $\lambda^{(i)}(x) + \mu_2^{(i)}(x)$. Both results agree.

*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 has been checked that one can have $\tau(l_1^{(i)}(x)) < \lambda_0(x)$, even if $\tilde{\lambda}_1(x)$ 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 $\lambda_0(x)$.

The updating procedure (3.16) 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 $\mu_2^{(i)}(x) \geqslant l_1^{(i)}(x)$ for getting an improved $\lambda^{(i+1)}(x)$ may be violated if $\lambda_0(x) \geqslant \tilde{\lambda}_1(x)$. 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 bounds $\lambda^{(3)}(x)$ and $R[C_x^{*(2)}]$ for $\lambda_0(x)$, and by comparing $R[C_x^{*(2)}]$ with $\| - \Delta_{\mathrm{N},x} C_x^{*(2)} \|$. 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 $\| - \Delta_{\mathrm{N},x} C_x^{*(2)} \|$ 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 *.

To conclude this paragraph the performance of the numerical algorithms on a CRAY Y-MP is reported. The routines are completely vectorizable (vectorization over the block index). In case of a 9^4 lattice and block size $L^d = 3^4$ (vector length of 81) the CPU time needed to compute $C_x^{*(2)}$ is 5.7 ms 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. Since for larger β the autocorrelation times are much greater than 1, kernels C and $\mathcal{A}$ will not be computed very frequently, which means that the new block spin construction is not more expensive from the point of view of CPU time than previous constructions.

4.3. DIELECTRIC BLOCK SPINS

In the last two paragraphs of this section a short account is given of some physical properties. First we consider the dielectric field $\sigma_\mu(x)$, and afterwards we turn to renormalization effects.

Dielectric gauge theory models as introduced in ref. [19] are candidates for effective actions for Yang–Mills theories, see also ref. [20]. For gauge group SU(2) the scalar field $\sigma_\mu(x)$ (2.22) is identified with the colour dielectric field. The effective action confines on the classical level if $\varepsilon = \langle \sigma_\mu(x)^4 \rangle$ (no

*Even if the eigenvalue equation were only solved approximately, an MGMC algorithm is set up in such a way that the equilibrium distribution is not affected by the choice of the C-kernel, assuming that certain normalization and orthogonality relations are valid [18].

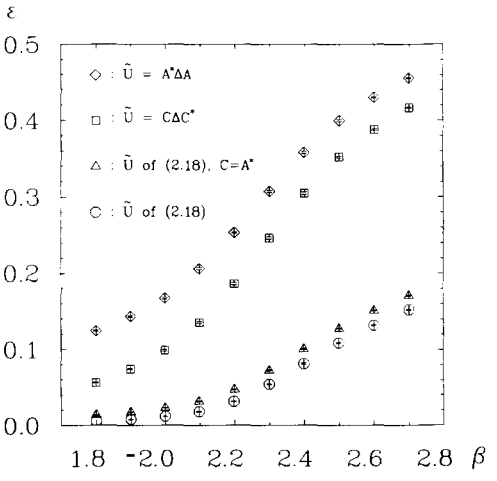

Fig. 3. Dielectric constant $\varepsilon = \langle \sigma_\mu(x)^4 \rangle$ for various definitions of block gauge fields $\tilde{U}$ relative to values in a pure gauge as a function of β, obtained when blocking a 9^4 lattice with scale factor $L = 3$.

summation over μ) approaches zero for large distances [15].

One blocking transformation from a 9^4 to a 3^4 lattice was performed, using eqs. (2.19) and (2.20) for the construction of block gauge fields, where the variational choices $\mathcal{A} = C^*$ and $C = \mathcal{A}^*$ were also considered. Periodic boundary conditions were used. Results for ε relative to values in pure gauges are plotted in fig. 3. The sample size comprised data of 100 configurations, measurements were only performed every fifth sweep, 1000–2000 sweeps had been discarded for thermalization.

The qualitative behaviour of ε is similar to the results of ref. [15] where another RGT was used. One observes also a crossover from strong to weak coupling at β around 2.0. This raises the hope that block spin definitions (2.19) and (2.20) might be fruitful for further investigations in the framework of the colour dielectric model.

4.4. EFFECTIVE COUPLING AFTER ONE BLOCKING TRANSFORMATION

The aim of the present paper is not to present a comprehensive Monte Carlo renormalization group investigation with the new block spin definitions. Only some preliminaries will be given.

While in subsect. 4.3 the field $\sigma_\mu(x)$ was examined, we pay attention to the unitary gauge field U on the block lattice in this section. The block lattice Λ' consists of 3^4 sites. This suggests to compare observables on Λ' and on

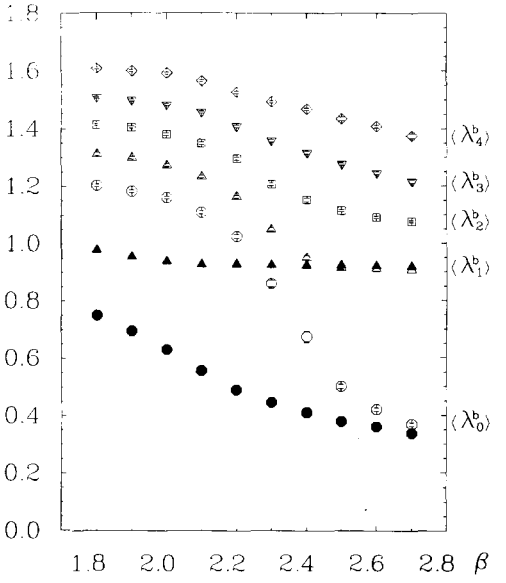

Fig. 4. Expectation values of the five lowest eigenvalues λ_i^b, $i = 0,\ldots,4$ (modulo two-fold degeneracy) of $-\Delta_{N,x}$ with the block gauge field U, using block spin definition (2.19) with $\Delta_{\text{eff}} = \mathcal{A}^*\Delta\mathcal{A}$, on a 3^4 block lattice versus β on the fundamental 9^4 lattice (open symbols). For comparison the results for $\langle\lambda_0(x)\rangle$ and $\langle\lambda_1(x)\rangle$ on a fundamental 3^4 lattice are also shown (filled symbols).

a fundamental 3^4 lattice. Firstly, the plaquette energy $\langle 1 - \frac{1}{2}\mathrm{Tr}\,U(P)\rangle$ on Λ' was measured. The results are collected in table 5. Secondly, the five lowest eigenvalues $\lambda^b{}_i$, $\lambda'{}^b_i$, $i = 0,\ldots,4$, (modulo two-fold degeneracy) of $-\Delta_{N,x}$ and $-\Delta_{D,x}$ with the block gauge field U were determined (library routines [32] were used again). The values $\langle\lambda^b{}_i\rangle$ obtained with block spin definition (2.19) are shown in fig. 4, where the results for $\langle\lambda_0(x)\rangle$ and $\langle\lambda_1(x)\rangle$ on the fundamental 3^4 lattice are also displayed for comparison. Clear renormalization effects are visible. The crossover behaviour of the lowest eigenvalue has become sharper. It is also shifted to β between 2.4 and 2.5. This shift is close to the two-loop perturbative result of $\Delta\beta = 0.44$ for scale factor 3. The crossover on the block lattice is also signaled by $\langle 1 - \frac{1}{2}\mathrm{Tr}\,U(P)\rangle$ (table 5) and ε (fig. 3); the errors of both quantities peak at β around 2.4, 2.5. It is interesting to note that the results for $\langle\lambda^b{}_i\rangle$, $\langle\lambda'{}^b_i\rangle$, $i = 0,\ldots,4$, are independent within errors of whether U is defined via (2.19) or (2.20), with either C or $\mathcal{A}$ (resp. their adjoints).

One can define an "effective coupling" β_{eff} on the block lattice which equals that value of β on the fundamental lattice where corresponding expectation values match. It is satisfactory that the values of β_{eff} derived from the plaquette

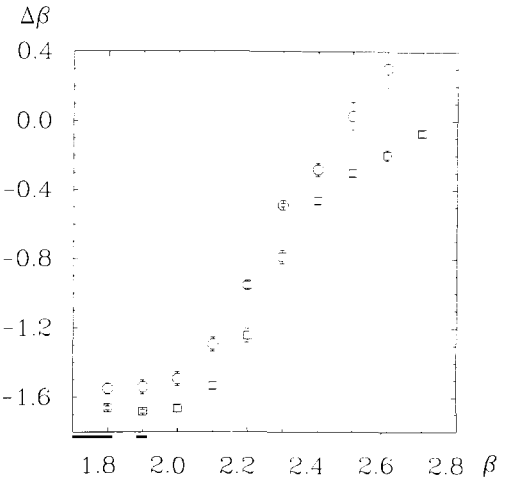

Fig. 5. The difference $\Delta\beta = \beta_{\text{eff}} - \beta$ versus β, derived from observables defined with kernels C (open circles) and $\mathcal{A}$ (open squares). β_{eff} on the 3^4 block lattice is defined to equal that value of β on the fundamental 3^4 lattice where corresponding expectation values match.

TABLE 5

Expectation value of the plaquette energy on a 3^4 block lattice as a function of β, obtained when blocking a 9^4 lattice with scale factor $L = 3$. (Recall that U is a unitary gauge field, while $\tilde{U}$ is dielectric.)

	$\langle 1 - \frac{1}{2}\mathrm{Tr}\, U(P)\rangle$			
	$\tilde{U}$ via (2.19)		$\tilde{U}$ via (2.20)	
β	with $\Delta_{\text{eff}} = C\Delta C^*$	with $\Delta_{\text{eff}} = \mathcal{A}^*\Delta\mathcal{A}$	with C of (2.10)	with $C = \mathcal{A}^*$
1.8	0.9330 ± 0.0027	0.9623 ± 0.0022	0.9535 ± 0.0025	0.9676 ± 0.0030
1.9	0.9081 ± 0.0035	0.9468 ± 0.0028	0.9218 ± 0.0032	0.9525 ± 0.0022
2.0	0.8676 ± 0.0032	0.9135 ± 0.0022	0.8778 ± 0.0023	0.9222 ± 0.0022
2.1	0.8018 ± 0.0026	0.8621 ± 0.0028	0.8098 ± 0.0040	0.8705 ± 0.0025
2.2	0.6925 ± 0.0031	0.7690 ± 0.0033	0.7046 ± 0.0033	0.7827 ± 0.0034
2.3	0.5509 ± 0.0055	0.6391 ± 0.0051	0.5616 ± 0.0033	0.6523 ± 0.0046
2.4	0.4215 ± 0.0098	0.5120 ± 0.0083	0.4177 ± 0.0033	0.5251 ± 0.0084
2.5	0.3169 ± 0.0034	0.4054 ± 0.0045	0.3229 ± 0.0080	0.4174 ± 0.0043
2.6	0.2687 ± 0.0017	0.3536 ± 0.0018	0.2734 ± 0.0028	0.3641 ± 0.0019
2.7	0.2381 ± 0.0016	0.3154 ± 0.0021	0.2363 ± 0.0020	0.3236 ± 0.0022

energies are consistent with those obtained from the lowest eigenvalues of $-\Delta_{\mathrm{N},x}$, $-\Delta_{\mathrm{D},x}$ for various definitions of U, except for U's defined via kernel C when $\beta \geqslant 2.6$, but here finite-size effects become severe. The difference $\Delta\beta = \beta_{\text{eff}} - \beta$ is shown in fig. 5. Of course, the β-function cannot be immediately related with this $\Delta\beta$, because no other couplings than the marginal gauge coupling are taken into account. An extensive investigation along the line carried out by Patel and Gupta [14] would be mandatory for that purpose.

5. Summary and outlook

An RGT for non-abelian gauge theories due to Mack, and in the spirit of the ground state projection method of Hulsebos et al. [7], and of Brower et al. [8], was explained. This projective definition of block spins involves dynamical information. Only the lowest frequency parts of the fields are used for constructing block spins. The systematic method presented in this paper provides an approach to deterministic and stochastic MG computations of non-abelian gauge theories in any number of space-time dimension.

Numerical algorithms for computing block spins resp. gauge-covariant MG restriction and interpolation operators were proposed. In case of pure SU(2) lattice gauge theory on a four-dimensional lattice the algorithms work very well. The required CPU time for computing the kernels C and $\mathcal{A}$ is comparable with that needed for one heat-bath sweep through the fundamental lattice.

First numerical physics results are very promising. They indicate that the projective RGT might be fruitful when studied in the framework of colour dielectric models, and also when it is used for RG investigations with block gauge field in the gauge group.

The new RGT is intended to be used in MGMC simulations as initiated by Mack [3]. With respect to this application it is important to record that the projective definition of block spins is numerically implementable for non-abelian gauge theories.

Further research will consider MG computation of gauge-covariant bosonic propagators in four dimensions, MG simulations of the SU(2) Higgs model, and analogous treatment of gauge fields coupled to fermions (replace $-\Delta$ by the Dirac operator $\not{D}$ resp. $-\not{D}^2$ ...) to compute gauge-covariant fermion propagators by means of projective MG methods.

I would like to thank Professor G. Mack for many stimulating and helpful discussions and for a reading of the manuscript. I am also indebted to K. Pinn and H.G. Evertz for useful discussions and hints. Professor L. Collatz taught me Temple's theorem in a very stimulating lecture on eigenvalue problems in the winter term of 1986/87. With sorrow I learnt of his death in September 1990. 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.

References

[1] K.G. Wilson and J. Kogut, Phys. Rep. 12C (1974) 75;
K.G. Wilson, Rev. Mod. Phys. 55 (1983) 583

[2] G. Mack and A. Pordt, Commun. Math. Phys. 97 (1985) 267
[3] G. Mack, Cargèse Lectures July 1987, *in* Nonperturbative quantum field theory, ed. G. 't Hooft et al. (Plenum Press, New York, 1988)
[4] W. Hackbusch, Multi-grid methods and applications, Springer Series in Computational Mathematics 4 (Springer-Verlag, Berlin, 1985)
[5] K. Gawędzki and A. Kupiainen, Commun. Math. Phys. 77 (1980) 31;
K. Gawędzki and A. Kupiainen, Commun. Math. Phys. 99 (1985) 197
[6] T. Balaban, Commun. Math. Phys. 109 (1987) 249
[7] A. Hulsebos, J. Smit and J.C. Vink, Nucl. Phys. B (Proc. Suppl.) 9 (1989) 512; Nucl. Phys. B331 (1990) 531
[8] R.C. Brower, C. Rebbi and E. Vicari, Phys. Rev. D43 (1991) 1965
[9] G. Mack and S. Meyer, Nucl. Phys. B (Proc. Suppl.) 17 (1990) 293
[10] G. Palma, DESY 90-083, July 1990
[11] T. Kalkreuter, Nucl. Phys. B352 (1991) 251
[12] R. Ben-Av, A. Brandt and S. Solomon, Nucl. Phys. B329 (1990) 193;
R. Ben-Av, A. Brandt, M. Harmatz, E. Katznelson, P.G. Lauwers, S. Solomon and K. Wolowesky, Phys. Lett. B253 (1991) 185
[13] D.J.E. Callaway and R. Petronzio, Phys. Lett. B145 (1984) 381
[14] A. Patel and R. Gupta, Nucl. Phys. B251 [FS13] (1985) 789
[15] H.J. Pirner, J. Wroldsen and M. Ilgenfritz, Nucl. Phys. B294 (1987) 905
[16] B. Großmann, H.J. Pirner, A.I. Signal, R. Baier and J. Wroldsen, Int. J. Mod. Phys. A6 (1991) 2649
[17] J. Goodman and A.D. Sokal, Phys. Rev. D40 (1989) 2035
[18] G. Mack, unpublished
[19] G. Mack, Nucl. Phys. B235 [FS11] (1984) 197
[20] H.J. Pirner, Nucl. Phys. B (Proc. Suppl.) 15 (1990) 207
[21] J. Jersák, *in* Higgs particle(s), ed. A. Ali (Plenum Press, New York, 1990);
H. Neuberger, Rutgers University preprint RU-89-44
[22] J. Glimm and A. Jaffe, Quantum physics, second edition (Springer-Verlag, New York, 1987)
[23] M. Hasenbusch, G. Mack and S. Meyer, Nucl. Phys. B (Proc. Suppl.) 20 (1991) 110
[24] A.D. Kennedy, Comput. Phys. Commun. 57 (1989) 57; Cargèse Lectures, August 1989, *in* Probabilistic methods in quantum field theory and quantum gravity, ed. P.H. Damgaard et al. (Plenum Press, New York, 1990)
[25] J. Stoer and R. Bulirsch, Einführung in die Numerische Mathematik II (Springer-Verlag, Berlin, 1978)
[26] J.H. Wilkinson and C. Reinsch, Linear Algebra (Springer-Verlag, Berlin, 1971)
[27] W.H. Press, B.P. Flannery, S.A. Teukolsky and W.T. Vetterling, Numerical Recipes (Cambridge Univ. Press, Cambridge, 1986)
[28] E. Polak, Computational methods in optimization (Academic Press, New York, 1971)
[29] L. Collatz, Eigenwertaufgaben mit technischen Anwendungen (Akademische Verlagsgesellschaft, Leipzig, 1963)
[30] L. Collatz, Funktionalanalysis und Numerische Mathematik (Springer-Verlag, Berlin, 1964)
[31] R. Courant and D. Hilbert, Methoden der Mathematischen Physik I (Springer-Verlag, Berlin, 1968)
[32] NAGLIB, Numerical Algorithms Group, NAG Fortran (Vector) Library, Mark 13
[33] M. Creutz, Phys. Rev. Lett. 43 (1979) 553