Overrelaxation algorithms for lattice field theories

Stephen L. Adler
Institute for Scientific Interchange, Villa Gualino-Viale Settimio Severo 65, 10131 Torino, Italy
*and The Institute for Advanced Study, Princeton, New Jersey 08540**
(Received 1 June 1987)

We study overrelaxation algorithms for the thermalization of lattice field theories with multiquadratic and more general actions. Overrelaxation algorithms are one-parameter generalizations of the heat-bath algorithm which satisfy the detailed-balance condition; the parameter is the relaxation parameter ω, $0<\omega<2$, with $\omega=1$ corresponding to the heat bath. First, we show that the $\omega\to 0$ (extreme underrelaxation) limit of the overrelaxation algorithm is equivalent to the Langevin equation approach. We analyze the thermalization of a free-field action, and show that for $\omega\sim 2$ an overrelaxed Gauss-Seidel algorithm yields a critical slowing down which is independent of wavelength, and has a correlation time which is a factor N smaller than that for an unaccelerated Jacobi iteration, with N the linear dimension of the lattice in lattice units. For a general nonmultiquadratic action, we give a generalized overrelaxation algorithm which satisfies detailed balance with respect to an effective action which is explicitly computable in terms of the original action. In the case of SU(n) lattice gauge theory we use this construction to formulate an overrelaxed algorithm which has exact lattice gauge invariance, and which satisfies detailed balance with respect to an effective action differing from the Wilson action only by terms of relative order a^2 in the continuum limit, with a the lattice spacing.

I. OVERRELAXATION AND ITS RELATION TO THE LANGEVIN APPROACH

The generic lattice field-theory problem is that of evaluating the Euclidean partition function

$$Z=\int d[\phi]e^{-\beta S[\{\phi\}]}, \tag{1}$$

with S the Euclidean action on a lattice and with $\int d[\phi]$ an integration over discretized lattice variables. In the Monte Carlo method for evaluating this integral, one generates a Markov chain of configurations $\{\phi_i\}$, $i=1,2,\ldots$ by application of a transition probability $W[\{\phi\}\to\{\phi'\}]$, which is chosen to satisfy the detailed-balance condition

$$e^{-\beta S[\{\phi\}]}W[\{\phi\}\to\{\phi'\}]=e^{-\beta S[\{\phi'\}]}W[\{\phi'\}\to\{\phi\}], \tag{2}$$

as well as normalization and erogodicity conditions.[1] These conditions guarantee that in the limit $i\to\infty$, the ensemble of configurations $\{\phi_i\}$ is distributed according to the equilibrium probability density $\exp(-\beta S[\{\phi\}])$. In what follows, I will refer to the problem of generating such an equilibrium distribution of configurations as the *thermalization problem*. If we now consider the $\beta\to\infty$ (zero-temperature) limit, only the configuration which minimizes S contributes to Eq. (1). Hence the zero-temperature limit of a thermalization algorithm will be an algorithm for the *minimization problem* of finding configurations $\{\bar{\phi}\}$ which satisfy

$$\frac{\delta}{\delta\phi}S[\{\phi\}]\bigg|_{\{\bar{\phi}\}}=0. \tag{3}$$

Conversely, we may expect that by generalizing methods which have been useful in solving the minimization problem, we can get useful algorithms for the thermalization problem.

Following this line of reasoning, a number of years ago I showed[2] that for the special case of multiquadratic actions (which includes[3] the classical Yang-Mills action), the standard Gauss-Seidel overrelaxation algorithm for the minimization problem can be generalized to an overrelaxation algorithm for the thermalization problem. Since this earlier work forms the starting point for the analysis of the present paper, I proceed now to briefly summarize it. A multiquadratic action is one which, for any node variable ϕ_k, can be decomposed as

$$S[\{\phi\}]=S[\{\phi\}_{\neq k},\phi_k]=A_k(\phi_k-C_k)^2+B_k,\quad A_k>0, \tag{4}$$

with A_k, B_k, and C_k functions of the remaining node variables

$$\{\phi\}_{\neq k}\equiv\{\phi_i, i\neq k\}. \tag{5}$$

A Gauss-Seidel iteration for the minimization problem consists of the successive replacement of each node variable ϕ_k by the value C_k which minimize the action as a function of that single variable, with the other variables $\{\phi\}_{\neq k}$ held fixed. Although this procedure gives the largest single step reduction in S, it is in fact not the most efficient procedure when coherent effects over the entire lattice are taken into account. A better minimization algorithm, which is no more demanding computationally, is the *overrelaxed* Gauss-Seidel algorithm

$$\phi_k\to\phi_k'=\omega C_k+(1-\omega)\phi_k, \tag{6}$$

with ω the "relaxation parameter." Convergence is guaranteed provided that S remains nonincreasing at each step, which requires

$$0 \le S[\{\phi\}_{\neq k}, \phi_k] - S[\{\phi\}_{\neq k}, \phi_k'] = A_k(\phi_k - \phi_k')^2 \left[\frac{2}{\omega} - 1\right], \tag{7}$$

giving the restriction

$$0 < \omega < 2 . \tag{8}$$

When $\omega = 1$, Eq. (6) reduces to the Gauss-Seidel prescription $\phi_k' = C_k$, in which the new value ϕ_k' has no memory of the old value ϕ_k. In practice, optimum convergence is obtained by doing several iterations with $\omega = 1$, and then doing many iterations with a value of ω close to 2.

Let us now turn to the thermalization problem for the action of Eq. (4). The thermalization analog of the Gauss-Seidel iteration is the heat-bath algorithm, in which a heat bath of temperature β^{-1} is touched in succession to each node variable ϕ_k, with the other variables $\{\phi\}_{\neq k}$ held fixed. In Ref. 2, I showed that the heat-bath algorithm for Eq. (4) admits a one-parameter generalization, analogous to Eq. (6), in which the normalized transition probability W is given by

$$W[\{\phi\}_{\neq k}, \phi_k \to \phi_k'] = \left[\frac{\beta A_k}{\pi\omega(2-\omega)}\right]^{1/2} \exp\left[-\left[\frac{\beta A_k}{\omega(2-\omega)}\right][\phi_k' - \omega C_k - (1-\omega)\phi_k]^2\right] . \tag{9}$$

When $\omega = 1$, Eq. (9) reduces to the heat-bath algorithm, since the new values ϕ_k' are distributed according to the equilibrium action and are independent of the old values ϕ_k. To see that Eq. (9) satisfies detailed balance for general ω, let us introduce a hyperbolic angle θ defined by

$$\omega - 1 = \tanh\theta, \quad \frac{1}{[\omega(2-\omega)]^{1/2}} = \cosh\theta, \quad \frac{1-\omega}{[\omega(2-\omega)]^{1/2}} = -\sinh\theta , \tag{10}$$

in terms of which Eq. (9) takes the form

$$W[\{\phi\}_{\neq k}, \phi_k \to \phi_k'] = (\beta A_k \cosh^2\theta/\pi)^{1/2} \exp\{-\beta A_k[\cosh\theta(\phi_k' - C_k) + \sinh\theta(\phi_k - C_k)]^2\} . \tag{11}$$

Detailed balance now immediately follows from the fact that

$$\begin{aligned}(\phi_k - C_k)^2 + [\cosh\theta(\phi_k' - C_k) + \sinh\theta(\phi_k - C_k)]^2 &= \cosh^2\theta[(\phi_k - C_k)^2 + (\phi_k' - C_k)^2] \\ &\quad + 2\cosh\theta\sinh\theta(\phi_k - C_k)(\phi_k' - C_k) \\ &= \text{symmetric in } \phi_k, \phi_k' .\end{aligned} \tag{12}$$

Since the transition probability of Eq. (9) is a Gaussian, it can be conveniently represented as a stochastic difference equation. Let n by a fictitious "time" index which increases by one for each update of the entire lattice, and let $\eta_{n,k}$ be a set of Gaussian noise variables distributed according to

$$W[\{\eta\}] = \prod_{n,k}\left[\frac{1}{\sqrt{4\pi}} e^{-\eta_{n,k}^2/4}\right], \tag{13}$$

and hence which obey

$$\langle \eta_{n,k}\eta_{n',k'}\rangle_\eta = 2\delta_{n,n'}\delta_{k,k'} . \tag{14}$$

Writing $\phi_k \equiv \phi_k^n$, $\phi_k' \equiv \phi_k^{n+1}$, Eq. (9) is evidently equivalent to

$$\phi_k^{n+1} - \phi_k^n = -\omega(\phi_k^n - C_k) - \left[\frac{\omega(2-\omega)}{4\beta A_k}\right]^{1/2} \eta_{n,k} , \tag{15}$$

with C_k and A_k functions of the ϕ_i^{n+1} for those nodes which precede ϕ_k, and of ϕ_i^n for those nodes which follow ϕ_k, in the sweep of the lattice. [This just corresponds to the fact that an updating of the whole lattice is accomplished by the *successive* application of the transition probability of Eq. (9) to each node of the lattice, in some specified sweep order.] Let us now rewrite Eq. (15) by using the fact that $\phi_k - C_k$ is proportional to $\partial S/\partial\phi_k$,

$$2\beta A_k(\phi_k^n - C_k) = \beta\frac{\partial S}{\partial\phi_k}[\{\phi_{i<k}^{n+1}\}, \{\phi_{i\ge k}^n\}] \equiv \beta\frac{\partial S}{\partial\phi_k} , \tag{16}$$

and by defining ϵ_k according to

$$\frac{\omega}{2\beta A_k} = \epsilon_k[\{\phi_{i<k}^{n+1}\}, \{\phi_{i\ge k}^n\}] \equiv \epsilon_k , \tag{17}$$

giving

$$\phi_k^{n+1} - \phi_k^n = -\epsilon_k\beta\frac{\partial S}{\partial\phi_k} - \left[1 - \frac{\omega}{2}\right]^{1/2} \epsilon_k^{1/2}\eta_{n,k} . \tag{18}$$

Apart from the extra factor of $(1-\omega/2)^{1/2}$, which approaches unity as $\omega \to 0$, Eq. (18) is just the discrete form of the Langevin equation with variable step size ϵ_k and a Gauss-Seidel interpretation of $\partial S/\partial\phi$, and approaches the corresponding Langevin stochastic differential equa-

tion as $\epsilon_k \propto \omega \to 0$. Hence the Langevin equation approach corresponds to the *extreme underrelaxation* limit[4] of the overrelaxation algorithm of Eq. (9).

II. CRITICAL SLOWING DOWN FOR A FREE-FIELD ACTION

In this section we give a detailed theoretical analysis of the performance of the overrelaxed minimization and thermalization algorithms, motivated by the fact that numerical studies by Whitmer,[5] Creutz,[4] and Brown and Woch[6] suggest that overrelaxation can improve the correlation time, as well as the speed of thermalization, in Monte Carlo simulations. We consider for simplicity the case of a single massless scalar free field ϕ in d dimensions; the inclusion of interaction and mass terms is not expected[7] to change the qualitative conclusions reached below. The node variable is thus

$$\phi_{i_1,\ldots,i_d} \tag{19}$$

and the action is taken as

$$S=\sum_{i_1,\ldots,i_d=1}^{N} \tfrac{1}{2}[(\phi_{i_1+1,i_2,\ldots,i_d}-\phi_{i_1,i_2,\ldots,i_d})^2 + (\phi_{i_1,i_2+1,\ldots,i_d}-\phi_{i_1,i_2,\ldots,i_d})^2 + \cdots + (\phi_{i_1,\ldots,i_d+1}-\phi_{i_1,\ldots,i_d})^2], \tag{20}$$

with homogeneous (Dirichlet or Neumann) boundary conditions applied at the edges of the lattice.[8] Introducing the notation

$$\phi(i_\mu \pm 1) \equiv \phi_{i_1,\ldots,i_d} \mid_{i_{l\neq\mu} \text{ fixed},\ i_\mu \to i_\mu \pm 1}, \tag{21}$$

we can now write the dependence of S on a given node $\phi_I \equiv \phi_{i_1,\ldots,i_d}$ as

$$S=\tfrac{1}{2}\sum_{\mu=1}^{d}\{[\phi(i_\mu+1)-\phi_I]^2+[\phi(i_\mu-1)-\phi_I]^2\}+\tilde{S}, \quad \tilde{S} \text{ independent of } \phi_I . \tag{22}$$

From Eqs. (4) and (11) of Sec. I, we see that an overrelaxed transition probability for the update $\phi_I \to \phi_I'$ can be constructed as

$$W[\phi_I \to \phi_I'] = \mathcal{N} \exp\left[-\tfrac{1}{2}\beta \sum_{\mu=1}^{d} (\{\cosh\theta[\phi(i_\mu+1)-\phi_I'] + \sinh\theta[\phi(i_\mu+1)-\phi_I]\}^2 + \{\cosh\theta[\phi(i_\mu-1)-\phi_I'] + \sinh\theta[\phi(i_\mu-1)-\phi_I]\}^2)\right], \tag{23}$$

with the normalization constant $\mathcal{N}$ independent of ϕ_I and ϕ_I'. Rewriting the ϕ_I' dependence by completing the square, and then substituting Eq. (10), Eq. (23) can be reexpressed as

$$\begin{aligned} W[\phi_I \to \phi_I'] &= \mathcal{N} \exp\left\{-\frac{\beta}{4d}\left[\cosh\theta\left[2d\phi_I' - \sum_{\mu=1}^{d}[\phi(i_\mu+1)+\phi(i_\mu-1)]\right] + \sinh\theta\left[2d\phi_I - \sum_{\mu=1}^{d}[\phi(i_\mu+1)+\phi(i_\mu-1)]\right]\right]^2 + \phi_I,\phi_I'\text{-independent}\right\} \\ &= \mathcal{N} \exp\left[-\frac{1}{4}\left[\frac{4\beta}{d\omega(2-\omega)}\right]\left[d\phi_I' - (1-\omega)d\phi_I - \tfrac{1}{2}\omega\sum_{\mu=1}^{d}[\phi(i_\mu+1)+\phi(i_\mu-1)]\right]^2 + \phi_I,\phi_I'\text{-independent}\right]. \end{aligned} \tag{24}$$

Applying the procedure of Eqs. (13)–(18), Eq. (24) can be rewritten as a Gauss-Seidel stochastic difference equation

$$\begin{aligned} &d\phi_I^{n+1} - (1-\omega)d\phi_I^n - \tfrac{1}{2}\omega\sum_{\mu=1}^{d}[\phi^n(i_\mu+1)+\phi^{n+1}(i_\mu-1)] = -\sigma\eta_{I,n}, \\ &\langle \eta_{I,n}\eta_{I',n'}\rangle_\eta = 2\delta_{I,I'}\delta_{n,n'}, \quad \sigma = \left[\frac{d\omega(2-\omega)}{4\beta}\right]^{1/2}, \quad \delta_{I,I'} \equiv \delta_{i_1,i_1'}\cdots\delta_{i_d,i_d'}. \end{aligned} \tag{25}$$

It will be informative, in what follows, to also analyze the corresponding Jacobi stochastic difference equation, in which the old values $\phi^n(i_\mu - 1)$ are used for the earlier nodes in the sweep, instead of the updated values $\phi^{n+1}(i_\mu - 1)$,

$$d\phi_I^{n+1} - (1-\omega)d\phi_I^n - \tfrac{1}{2}\omega\sum_{\mu=1}^{d}[\phi^n(i_\mu+1)+\phi^{n+1}(i_\mu-1)] = -\sigma\eta_{I,n}, \tag{26}$$

Although Eq. (26) is not equivalent to the iteration of an algorithm which satisfies detailed balance with the action of Eq. (20), it has been extensively studied by Batrouni *et al.*,[9] and so furnishes a useful point of comparison.

To solve Eqs. (25) and (26), we proceed by introducing a Green's function

$$G^{n,n'}_{i_1,\ldots,i_d;i'_1,\ldots,i'_d} \equiv G^{n,n'}_{I,I'} , \tag{27}$$

which satisfies the stochastic difference equation,

Gauss-Seidel case:

$$dG^{n+1,n'}_{I,I'} - (1-\omega) dG^{n,n'}_{I,I'} - \tfrac{1}{2}\omega \sum_{\mu=1}^{d} [G^{n,n'}_{I'}(i_\mu+1) + G^{n+1,n'}_{I'}(i_\mu-1)] = \delta_{I,I'}\delta_{n,n'} , \tag{28a}$$

Jacobi case:

$$dG^{n+1,n'}_{I,I'} - (1-\omega) dG^{n,n'}_{I,I'} - \tfrac{1}{2}\omega \sum_{\mu=1}^{d} [G^{n,n'}_{I'}(i_\mu+1) + G^{n,n'}_{I'}(i_\mu-1)] = \delta_{I,I'}\delta_{n,n'} \tag{28b}$$

and the boundary condition

$$G^{0,n'}_{I,I'} = 0 . \tag{29}$$

Then the solution of Eqs. (25) and (26) can be written as

$$\phi^n_I = - \sum_{I',n'} G^{n,n'}_{I,I'} \sigma \eta_{I',n'} + \bar{\phi}^{\,n}_I , \tag{30}$$

where $\bar{\phi}^{\,n}_I$ is the solution of the $\sigma = 0$ (noise-free, or zero temperature) iteration,

Gauss-Seidel case:

$$d\bar{\phi}^{\,n+1}_I - (1-\omega) d\bar{\phi}^{\,n}_I - \tfrac{1}{2}\omega \sum_{\mu=1}^{d} [\bar{\phi}^{\,n}(i_\mu+1) + \bar{\phi}^{\,n+1}(i_\mu-1)] = 0 , \tag{31a}$$

Jacobi case:

$$d\bar{\phi}^{\,n+1}_I - (1-\omega) d\bar{\phi}^{\,n}_I - \tfrac{1}{2}\omega \sum_{\mu=1}^{d} [\bar{\phi}^{\,n}(i_\mu+1) + \bar{\phi}^{\,n}(i_\mu-1)] = 0 \tag{31b}$$

with the initial condition $\bar{\phi}^{\,0}_I = \phi^0_I$. Since Eq. (30) implies that

$$\langle \phi^n_I \rangle_\eta = \bar{\phi}^{\,n}_I , \tag{32}$$

we see that introduction of the Green's function has permitted us to separate ϕ^n_I into a mean value term and individual noise contributions. The rapidity of thermalization is determined by the rate of decay of $\bar{\phi}^{\,n}_I$ with n, while the correlation time (the number of updates required to evolve from one thermalized configuration to an independent one) is determined by the rate of decay of $G^{n,n'}_{I,I'}$ with n. For a general Monte Carlo calculation the thermalization and the correlation times are different, but they will turn out to be equal for the overrelaxed quadratic action case studied in this section.

Since we are really interested only in the asymptotic limit of small mesh spacings or large lattices, we do not attempt to solve the difference equations (28) and (31) directly. (An alternative method, working directly from the iteration matrix for the difference equations, and yielding similar conclusions, has been given by Goodman and Sokal.[10]) Instead we follow the method of Garabedian[11] and convert the discrete equations to an equivalent continuum problem, for which the corresponding partial differential equations can be solved by standard methods. Let us denote the mesh spacing by a and introduce continuum variables x_μ, t by the correspondence

$$\begin{aligned} & x_\mu \leftrightarrow a i_\mu , \quad \int dx_\mu \leftrightarrow a \sum_{i_\mu} , \quad d/dx_\mu \leftrightarrow a^{-1}\Delta_{i_\mu} , \\ & t \leftrightarrow a n , \quad \int dt \leftrightarrow a \sum_{n} , \quad d/dt \leftrightarrow a^{-1}\Delta_t , \\ & a^{d+1}\delta(x_1 - x'_1)\cdots\delta(x_d - x'_d)\delta(t-t') \leftrightarrow \delta_{I,I'}\delta_{n,n'} , \end{aligned} \tag{33}$$

with Δ the finite difference operator. Treating first the Jacobi iteration case, we rewrite Eqs. (28b) and (31b) as

$$\begin{aligned} & a^{-1} d a^{-1} (G^{n+1,n'}_{I,I'} - G^{n,n'}_{I,I'}) - \tfrac{1}{2}\omega a^{-2} \sum_{\mu=1}^{d} [G^{n,n'}_{I'}(i_\mu+1) + G^{n,n'}_{I'}(i_\mu-1) - 2G^{n,n'}_{I,I'}] = a^{d-1} a^{-d-1} \delta_{I,I'}\delta_{n,n'} , \\ & a^{-1} d a^{-1} (\bar{\phi}^{\,n+1}_I - \bar{\phi}^{\,n}_I) - \tfrac{1}{2}\omega a^{-2} \sum_{\mu=1}^{d} [\bar{\phi}^{\,n}(i_\mu+1) + \bar{\phi}^{\,n}(i_\mu-1) - 2\bar{\phi}^{\,n}_I] = 0 . \end{aligned} \tag{34}$$

Making the correspondence

$$G^{n,n'}_{I,I'} \leftrightarrow G(x,x';t,t') , \quad \bar{\phi}^{\,n}_I \leftrightarrow \bar{\phi}(x,t) \tag{35}$$

and referring to Eq. (33), we see that Eqs. (34) are the discrete analogs of the continuum parabolic partial differential equations

$$\begin{aligned} & \frac{2d}{\omega a}\frac{\partial}{\partial t} G(x,x';t,t') - \sum_{\mu=1}^{d} \frac{\partial^2}{\partial x_\mu^{\,2}} G(x,x';t,t') = \frac{2}{\omega} a^{d-1}\delta(x_1 - x'_1)\cdots\delta(x_d - x'_d)\delta(t-t') , \\ & \frac{2d}{\omega a}\frac{\partial}{\partial t}\bar{\phi}(x,t) - \sum_{\mu=1}^{d} \frac{\partial^2}{\partial x_\mu^{\,2}}\bar{\phi}(x,t) = 0 , \end{aligned} \tag{36}$$

with the initial conditions

$$G(x,x';0,t')=0, \quad t'>0, \quad \bar{\phi}(x,0)=\text{smooth interpolation of } \phi_I^0 \ . \tag{37}$$

Turning next to the Gauss-Seidel iteration case, we follow the Garabedian analysis and anticipate the fact that the optimum ω is related to the mesh spacing a by

$$\omega=\frac{2}{1+Ca} \ , \tag{38}$$

with C a constant of order unity. Substituting Eq. (38) into Eqs. (28a) and (31a), these can be rewritten in the form

$$\begin{aligned} dCa^{-1}(G_{I,I'}^{n+1,n'}-G_{I,I'}^{n,n'})-a^{-2}\sum_{\mu=1}^{d}[G_{I'}^{n,n'}(i_\mu+1)+G_{I'}^{n,n'}(i_\mu-1)-2G_{I,I'}^{n,n'}] \\ +a^{-2}\sum_{\mu=1}^{d}\{G_{I'}^{n+1,n'}(i_\mu)-G_{I'}^{n+1,n'}(i_\mu-1)-[G_{I'}^{n,n'}(i_\mu)-G_{I'}^{n,n'}(i_\mu-1)]\}=\frac{2}{\omega}a^{d-1}a^{-d-1}\delta_{I,I'}\delta_{n,n'} \ , \end{aligned} \tag{39}$$

$$\begin{aligned} dCa^{-1}(\bar{\phi}_I^{n+1}-\bar{\phi}_I^{n})-a^{-2}\sum_{\mu=1}^{d}[\bar{\phi}^{n}(i_\mu+1)+\bar{\phi}^{n}(i_\mu-1)-2\bar{\phi}_I^{n}]+a^{-2}\sum_{\mu=1}^{d}\{\bar{\phi}^{n+1}(i_\mu)-\bar{\phi}^{n+1}(i_\mu-1) \\ -[\bar{\phi}^{n}(i_\mu)-\bar{\phi}^{n}(i_\mu-1)]\}=0 \ . \end{aligned}$$

Again making the correspondence of Eq. (35), we see that Eqs. (39) are the discrete analogs of the continuum hyperbolic partial differential equations:

$$\begin{aligned} dC\frac{\partial}{\partial t}G(x,x';t,t')-\sum_{\mu=1}^{d}\frac{\partial^2}{\partial x_\mu^{\,2}}G(x,x';t,t')+\sum_{\mu=1}^{d}\frac{\partial^2}{\partial t\,\partial x_\mu}G(x,x';t,t')=\frac{2}{\omega}a^{d-1}\delta(x_1-x_1')\cdots\delta(x_d-x_d')\delta(t-t') \ , \\ dC\frac{\partial}{\partial t}\bar{\phi}(x,t)-\sum_{\mu=1}^{d}\frac{\partial^2}{\partial x_\mu^{\,2}}\bar{\phi}(x,t)+\sum_{\mu=1}^{d}\frac{\partial^2}{\partial t\,\partial x_\mu}\bar{\phi}(x,t)=0 \ , \end{aligned} \tag{40}$$

with initial conditions[12] as in Eq. (37). Now making the change of variable

$$s=t+\tfrac{1}{2}\sum_{\mu=1}^{d}x_\mu \ , \tag{41}$$

some straightforward algebra shows that Eq. (40) is transformed into the canonical hyperbolic form

$$\begin{aligned} dC\frac{\partial}{\partial s}G(x,x';s,s')+\frac{d}{4}\frac{\partial^2}{\partial s^2}G(x,x';s,s')-\sum_{\mu=1}^{d}\frac{\partial^2}{\partial x_\mu^{\,2}}G(x,x';s,s')=\frac{2}{\omega}a^{d-1}\delta(x_1-x_1')\cdots\delta(x_d-x_d')\delta(s-s') \ , \\ dC\frac{\partial}{\partial s}\bar{\phi}(x,s)+\frac{d}{4}\frac{\partial^2}{\partial s^2}\bar{\phi}(x,s)-\sum_{\mu=1}^{d}\frac{\partial^2}{\partial x_\mu^{\,2}}\bar{\phi}(x,s)=0 \ , \end{aligned} \tag{42}$$

with the boundary conditions

$$\begin{aligned} &G(x,x';s,s')=0 \ , \\ &\bar{\phi}(x,s)=\text{smooth interpolation of } \phi_I^0 \end{aligned} \tag{43a}$$

imposed on the surface

$$s=\tfrac{1}{2}\sum_{\mu=1}^{d}x_\mu \ . \tag{43b}$$

Let us proceed now to solve Eqs. (36) and (42) by separation of variables. Let $\psi_m(x)$ be a complete set of eigenfunctions of the d-dimensional Laplace operator

$$\nabla^2=\sum_{\mu=1}^{d}\frac{\partial^2}{\partial x_\mu^{\,2}} \ , \tag{44}$$

subject to homogeneous boundary conditions on the edge of the cube $0\le x_\mu\le L$ which bounds the lattice. We assume that the boundary conditions are such that there are no normalizable zero modes. (This implies that the corresponding Laplace equation with inhomogeneous boundary conditions has a unique solution.) Then we have

$$\begin{aligned} &\nabla^2\psi_m=-k_m^{\,2}\psi_m \ , \\ &\delta(x_1-x_1')\cdots\delta(x_d-x_d')=\sum_m\psi_m(x)\psi_m^*(x') \ , \end{aligned} \tag{45}$$

with the minimum eigenvalue k_1 of order L^{-1}. In the Jacobi iteration case, we expand

$$\begin{aligned} &G(x,x';t,t')=\sum_m G_m(x';t,t')\psi_m(x) \ , \\ &\bar{\phi}(x,t)=\sum_m\bar{\phi}_m(t)\psi_m(x) \ , \end{aligned} \tag{46}$$

with the coefficients G_m and $\bar{\phi}_m$ obeying the differential equations

$$\frac{2d}{\omega a}\frac{d}{dt}G_m(x';t,t')+k_m{}^2G_m(x';t,t') = \frac{2}{\omega}a^{d-1}\psi_m^*(x')\delta(t-t') , \tag{47}$$

$$\frac{2d}{\omega a}\frac{d}{dt}\bar{\phi}_m(t)+k_m{}^2\bar{\phi}_m(t)=0 .$$

Expanding the initial condition on $\bar{\phi}$ as

$$\bar{\phi}(x,0)=\sum_m \phi_m^{(0)}\psi_m(x) , \tag{48}$$

Eqs. (47) are readily integrated to give

$$G(x,x';t,t')=\frac{a^d}{d}\sum_m \psi_m(x)\psi_m^*(x')e^{-\lambda_m(t-t')}\theta(t-t') , \tag{49}$$

$$\bar{\phi}(x,t)=\sum_m \phi_m^{(0)}\psi_m(x)e^{-\lambda_m t}, \quad \lambda_m=\frac{\omega a k_m{}^2}{2d} .$$

We turn next to the Gauss-Seidel iteration case. We now expand

$$G(x,x';s,s')=\sum_m G_m(x';s,s')\psi_m(x) , \qquad \bar{\phi}(x,s)=\sum_m \bar{\phi}_m(s)\psi_m(x) , \tag{50}$$

with the coefficients G_m and $\bar{\phi}_m$ obeying the differential equations

$$dC\frac{d}{ds}G_m(x';s,s')+\frac{d}{4}\frac{d^2}{ds^2}G_m(x';s,s') +k_m{}^2G_m(x';s,s')=\frac{2}{\omega}a^{d-1}\psi_m^*(x')\delta(s-s') , \tag{51}$$

$$dC\frac{d}{ds}\bar{\phi}_m(s)+\frac{d}{4}\frac{d^2}{ds^2}\bar{\phi}_m(s)+k_m{}^2\bar{\phi}_m(s)=0 .$$

The general solution for $\bar{\phi}_m(s)$ has the form first given by Garabedian:[11]

$$\bar{\phi}_m(s)=a_m e^{-p_m s}+b_m e^{-q_m s} , \qquad p_m=2[C-(C^2-k_m{}^2/d)^{1/2}] , \qquad q_m=2[C+(C^2-k_m{}^2/d)^{1/2}] , \tag{52}$$

with the coefficients a_m, b_m implicitly (but not explicitly) determined by matching to the initial condition on $\bar{\phi}$ of Eq. (43). The values of a_m, b_m in fact do not matter; all we need for what follows is that $\bar{\phi}$ decays as a function of time at least as fast as

$$\exp[-t\ \min_m(\mathrm{Re}p_m, \mathrm{Re}q_m)] . \tag{53}$$

To solve the equation for G_m, we write

$$G_m(x';s,s')=\begin{cases} G_m^>(x';s,s'), & s>s' , \\ G_m^<(x';s,s'), & s<s' , \end{cases} \tag{54}$$

with

$$G_m^{>,<}(x';s,s')=a_m^{>,<}(x';s')e^{-p_m s}+b_m^{>,<}(x';s')e^{-q_m s} . \tag{55}$$

Continuity across $s=s'$ requires

$$G_m^>(x';s',s')=G_m^<(x';s',s') , \tag{56}$$

while the δ function on the right-hand side of Eq. (51) requires the first derivative discontinuity to be

$$\frac{d}{4}\frac{d}{ds}[G_m^>(x';s,s')-G_m^<(x';s,s')]\Big|_{s=s'}=\frac{2}{\omega}a^{d-1}\psi_m^*(x') . \tag{57}$$

To solve these, let us make the ansatz

$$G_m^<=0 \tag{58}$$

and then show that this does in fact satisfy the boundary condition of Eq. (43). Assuming Eq. (58), a little algebra shows that Eqs. (56) and (57) are satisfied by

$$G_m^>(x';s,s')=\frac{4}{d}\frac{2}{\omega}a^{d-1}\psi_m^*(x')\frac{e^{p_m(s'-s)}-e^{q_m(s'-s)}}{q_m-p_m} , \tag{59}$$

as can be verified by inspection. Hence $G(x,x';s,s')$ is given by

$$G(x,x';s,s')=\frac{4}{d}\frac{2}{\omega}a^{d-1} \times\sum_m \psi_m(x)\psi_m^*(x') \times\frac{e^{p_m(s'-s)}-e^{q_m(s'-s)}}{q_m-p_m}\theta(s-s') \tag{60}$$

and corresponds to taking a solution to the hyperbolic equation Eq. (42) which has support only inside the forward light cone:

$$\left[\sum_{\mu=1}^d (x_\mu-x'_\mu)^2\right]^{1/2}\le\frac{2}{d^{1/2}}(s-s') . \tag{61}$$

Now the Schwartz inequality implies

$$\pm\sum_{\mu=1}^d (x_\mu-x'_\mu)\le\left[\sum_{\mu=1}^d 1\right]^{1/2}\left[\sum_{\mu=1}^d (x_\mu-x'_\mu)^2\right]^{1/2} =d^{1/2}\left[\sum_{\mu=1}^d (x_\mu-x'_\mu)^2\right]^{1/2} \tag{62}$$

and combining the inequalities of Eqs. (61) and (62), we see that inside the forward light cone we have the inequalities[13]

$$0\le s-s'\mp\tfrac{1}{2}\sum_{\mu=1}^d (x_\mu-x'_\mu)=\begin{cases} t-t' \\ t-t'+\sum_{\mu=1}^d (x_\mu-x'_\mu) . \end{cases} \tag{63}$$

Thus for $t'>0$, the surface $t=0$ lies entirely *outside* the forward light cone, and hence the initial condition that G vanish at $t=0$ is satisfied by Eq. (60), even though the differential equation is not separable in the x,t coordinate system. From Eq. (60) we learn that G also decays

as a function of time at least as fast as Eq. (53).

Let us now determine (following again Ref. 7) the value of C which maximizes the decay exponent

$$\lambda_{\rm GS}=\min_m(\mathrm{Re}p_m,\mathrm{Re}q_m)\ , \tag{64}$$

where GS denotes Gauss-Seidel. For $k_m{}^2$ large enough so that $k_m{}^2/d\geq C^2$, we have

$$\mathrm{Re}p_m=\mathrm{Re}q_m=2C\ . \tag{65}$$

On the other hand, for values of $k_m{}^2$ small enough so that $(C^2-k_m{}^2/d)^{1/2}$ is real, we have

$$\min(\mathrm{Re}p_m,\mathrm{Re}q_m)=p_m\geq p_1=2[C-(C^2-k_1{}^2/d)^{1/2}]\ . \tag{66}$$

Hence

$$\min_m(\mathrm{Re}p_m,\mathrm{Re}q_m)=\min\{2C,2[C-\mathrm{Re}(C^2-k_1{}^2/d)^{1/2}]\}\ , \tag{67}$$

and this expression is maximized for

$$C_{\rm opt}=k_1/d^{1/2}\ . \tag{68}$$

At the optimum C we have

$$\lambda_{\rm GS}=\mathrm{Re}p_m=\mathrm{Re}q_m=2C_{\rm opt}=2k_1/d^{1/2} \tag{69}$$

and all modes have the *same* time decay exponent. By contrast, for the Jacobi iteration the decay exponent [Eq. (49)] varies quadratically with wave number

$$\lambda_m=\omega_J a k_m{}^2/d \tag{70}$$

and becomes very small at the largest wavelengths, giving for the most slowly decaying mode

$$\lambda_J=\omega_J a k_1{}^2/d\ . \tag{71}$$

Comparing Eqs. (69) and (71) we have

$$\frac{\lambda_{\rm GS}}{\lambda_J}=\frac{2}{\omega_J}d^{1/2}\frac{1}{k_1 a}\ . \tag{72a}$$

Since $k_1\sim L^{-1}$ and $L/a=N$, with N the dimension of the lattice in lattice units, we get our fundamental result

$$\frac{\lambda_{\rm GS}}{\lambda_J}\sim\frac{2}{\omega_J}d^{1/2}N\ . \tag{72b}$$

Hence the overrelaxed Gauss-Seidel algorithm dramatically improves both the rapidity of thermalization and the correlation time as compared with the Jacobi algorithm, and makes critical slowing down independent of wave length. This improvement becomes even more pronounced when compared with Langevin-Jacobi procedures, for which (as shown in Sec. I) one has $\omega_J\ll 1$.

We conclude this section with two checks on the analysis given above. First, the Jacobi case analyzed above is just a continuum version of the model for the correlation length studied by Batrouni *et al.*[9] In units with $a=1$, they find

$$N_c\sim\frac{1}{\bar{\epsilon}(p^2+m^2)}\ , \tag{73}$$

which with the correspondences $\bar{\epsilon}\sim\omega_J$ (cf. Sec. I) and $p^2+m^2\sim k_m{}^2$ becomes

$$N_c\sim\frac{1}{\omega_J k_m^2}\sim\lambda_m{}^{-1}\ , \tag{74}$$

and so our result agrees with theirs. Second, as a check on the reasoning leading to Eq. (60), we have explicitly evaluated the time dependence of the Gauss-Seidel Green's function for the $L\to\infty$ limit in which the ψ_m are infinite-space mode functions. Details of this calculation are given in the Appendix; the result is

$$G(x,x';t,t')=\frac{4}{d}\frac{2}{\omega}\frac{a^{d-1}}{(2\pi)^d}\int d^dl\, e^{il\cdot(\mathbf{x}-\mathbf{x}')}g(l,t-t')\ , \tag{75}$$

$$g(l^{\parallel},l^{\perp},t)=\frac{1}{4(C+il^{\parallel}/d^{1/2})}\times\exp\left[-t\frac{(l^{\parallel})^2+(l^{\perp})^2}{(l^{\parallel})^2+C^2d}\times(C-il^{\parallel}/d^{1/2})\right]\theta(t)\ ,$$

with $l^{\parallel}$ and $l^{\perp}$ the components of l parallel and perpendicular to the fixed vector $(1,1,\ldots,1)$. The presence of the factor $\theta(t-t')$ implies that $G(x,x';t,t')$ vanishes at $t=0$ for $t'>0$, and so Eq. (60) does indeed satisfy the initial condition of Eq. (37). For $C=k_1/d^{1/2}$, Eq. (75) implies that the decay exponent is $\sim k_1$ for wave numbers $|l|$ larger than k_1, in agreement with Eq. (69). [For wave numbers $|l|$ smaller than k_1 Eq. (75) is no longer relevant, since the difference between infinite space and finite box mode functions becomes significant.]

III. A GENERALIZED OVERRELAXATION ALGORITHM, AND APPLICATION TO SU(n) LATTICE FIELD AND GAUGE THEORY

The results of Sec. II indicate that overrelaxation should be of computational value for the thermalization problem, and so we proceed next to construct overrelaxation algorithms for the Yang-Mills action (which, as noted above, is multiquadratic in the components of the gauge potential), and for the Wilson lattice gauge action (which is not multiquadratic). The construction employs the following generalization of the overrelaxation algorithm of Sec. I: Consider a field theory with field variables which can be divided into two disjoint classes $\{\phi\},\{\psi\}$, with functional integration measure

$$d\mu=\prod_1^{N_\phi}\int d\phi\prod_1^{N_\psi}\int d\psi \tag{76a}$$

and with the general (nonmultiquadratic) action

$$S=S_1[\{\phi\},\{\psi\}]+S_2[\{\psi\}]\ . \tag{76b}$$

For this theory, consider an updating $\{\phi\}\to\{\phi'\}$ in which only the $\{\phi\}$ variables (or some subset of them) is changed, and let $\tilde{S}[\{\phi\},\{\phi'\},\{\psi\};\theta]$ be any auxiliary functional of the indicated field variables and the relaxation parameter θ which is symmetrical under the interchange $\{\phi\}\leftrightarrow\{\phi'\}$. For this updating, we take the transition probability to be

$$W[\{\phi\}\to\{\phi'\}]=\mathcal{N}[\{\phi\},\{\psi\};\theta]\exp(-\beta\cosh^2\theta S_1[\{\phi'\},\{\psi\}] \\ -\beta\sinh^2\theta S_1[\{\phi\},\{\psi\}]-\beta\tilde{S}[\{\phi\},\{\phi'\},\{\psi\};\theta]) \tag{77a}$$

with the normalization $\mathcal{N}[\{\phi\},\{\psi\};\theta]$ given by

$$\mathcal{N}^{-1}[\{\phi\},\{\psi\};\theta]=\prod_1^{N_\phi}\int d\phi'\exp(-\beta\cosh^2\theta S_1[\{\phi'\},\{\psi\}]-\beta\sinh^2\theta S_1[\{\phi\},\{\psi\}]-\beta\tilde{S}[\{\phi\},\{\phi'\},\{\psi\};\theta])\ . \tag{77b}$$

Then W satisfies detailed balance with respect to the effective action

$$S_{\text{eff}}[\{\phi\},\{\psi\};\theta]=S[\{\phi\},\{\psi\}]+\beta^{-1}\ln(\mathcal{N}[\{\phi\},\{\psi\};\theta]/\bar{\mathcal{N}}[\{\psi\};\theta])\ , \\ \bar{\mathcal{N}}[\{\psi\};\theta]=\prod_1^{N_\phi}\int d\phi\,\mathcal{N}[\{\phi\},\{\psi\};\theta]\Big/\prod_1^{N}\int d\phi\ . \tag{78}$$

The proof follows directly from the fact that

$$W[\{\phi\}\to\{\phi'\}]\exp(-\beta S[\{\phi\},\{\psi\}]-\ln(\mathcal{N}[\{\phi\},\{\psi\};\theta]/\bar{\mathcal{N}}[\{\psi\};\theta])) \\ =\bar{\mathcal{N}}[\{\psi\};\theta]\exp(-\beta\cosh^2\theta(S_1[\{\phi'\},\{\psi\}]+S_1[\{\phi\},\{\psi\}])-\beta\tilde{S}[\{\phi\},\{\phi'\},\{\psi\};\theta]-\beta S_2[\{\psi\}]) \\ =\text{symmetrical in }\{\phi\},\{\phi'\}\ . \tag{79}$$

The multiquadratic case of the generalized algorithm is recovered by taking

$$S_1[\{\phi\},\{\psi\}]=\sum_{ij}L_i[\{\phi\},\{\psi\}]A_{ij}[\{\psi\}]L_j[\{\phi\},\{\psi\}] \tag{80}$$

with the L_i linear functionals of the subset of variables $\{\phi\}$. The overrelaxation algorithm of Secs. I and II then corresponds to the choice of auxiliary functional

$$\tilde{S}[\{\phi\},\{\phi'\},\{\psi\};\theta]=\sinh\theta\cosh\theta\sum_{ij}(L_i[\{\phi'\},\{\psi\}]A_{ij}[\{\psi\}]L_j[\{\phi\},\{\psi\}] \\ +L_i[\{\phi\},\{\psi\}]A_{ij}[\{\psi\}]L_j[\{\phi'\},\{\psi\}])\ , \tag{81}$$

for which the transition probability W of Eq. (77a) becomes

$$W[\{\phi\}\to\{\phi'\}]=\mathcal{N}\exp\left[-\beta\sum_{ij}(\cosh\theta L_i[\{\phi'\},\{\psi\}]+\sinh\theta L_i[\{\phi\},\{\psi\}]) \right. \\ \left. \times A_{ij}[\{\psi\}](\cosh\theta L_j[\{\phi'\},\{\psi\}]+\sinh\theta L_j[\{\phi\},\{\psi\}])\right]\ ; \tag{82a}$$

in terms of a relaxation parameter ω related to θ as in Eq. (10), this can also be written as

$$W[\{\phi\}\to\{\phi'\}]=\mathcal{N}\exp\left[-\frac{\beta}{\omega(2-\omega)}\sum_{ij}(L_i[\{\phi'\},\{\psi\}]-(1-\omega)L_i[\{\phi\},\{\psi\}]) \right. \\ \left. \times A_{ij}[\{\psi\}](L_j[\{\phi'\},\{\psi\}]-(1-\omega)L_j[\{\phi\},\{\psi\}])\right]\ . \tag{82b}$$

Since by the linearity of L we have

$$\cosh\theta L_i[\{\phi'\},\{\psi\}]+\sinh\theta L_i[\{\phi\},\{\psi\}]=L_i[\{\cosh\theta\phi'+\sinh\theta\phi\},\{\psi\}]\ , \tag{83}$$

the normalization is now given by

$$\mathcal{N}^{-1}=\prod_1^{N_\phi}\int d\phi'\exp(-\beta L_i[\{\cosh\theta\phi'+\sinh\theta\phi\},\{\psi\}]A_{ij}[\{\psi\}]L_j[\{\cosh\theta\phi'+\sinh\theta\phi\},\{\psi\}]) \\ =(\cosh\theta)^{-N_\phi}\prod_1^{N_\phi}\int d\phi'\exp(-\beta L_i[\{\phi'\},\{\psi\}]A_{ij}[\{\psi\}]L_j[\{\phi'\},\{\psi\}])=\text{independent of }\{\phi\}\ , \tag{84}$$

and so the second term on the right-hand side of Eq. (78) vanishes, giving in the multiquadratic case $S_{\text{eff}}[\{\phi\},\{\psi\};\theta]=S[\{\phi\},\{\psi\}]$. In the application of the generalized algorithm to the Wilson lattice gauge theory given below, the second term on the right-hand side of Eq. (78) will be nonzero, but is arranged to be a higher-order correction in the continuum limit as compared with the original action S.

Let us now apply this algorithm to the Yang-Mills action

$$\beta S = \int d^4x \tfrac{1}{2}\mathrm{Tr}(F_{\mu\nu}F_{\mu\nu}), \quad F_{\mu\nu}=F^j_{\mu\nu}T^j, \quad \mathrm{Tr}(T^iT^j)=\tfrac{1}{2}\delta_{ij} , \tag{85}$$

with the field-strength $F_{\mu\nu}$ related to the potential A_μ by

$$A_\mu = A^j_\mu T^j, \quad F_{\mu\nu}=\partial_\mu A_\nu - \partial_\nu A_\mu + ig_0[A_\mu, A_\nu] . \tag{86}$$

To formulate a discrete version of Eq. (85), we set up a cubic lattice with unit cell of side a, and associate the potential variables with the centers of the links. Then for a plaquette in the x_μ-x_ν plane with center $(x_{c\mu},x_{c\nu})$, as shown in Fig. 1, we have, for the field-strength component $F_{\mu\nu}$ at the center of the plaquette,

$$\begin{aligned} F_P \equiv F_{\mu\nu}(x_{c\mu},x_{c\nu}) = a^{-1}[&A_\nu(x_{c\mu}+\tfrac{1}{2}a,x_{c\nu}) - A_\nu(x_{c\mu}-\tfrac{1}{2}a,x_{c\nu}) - A_\mu(x_{c\mu},x_{c\nu}+\tfrac{1}{2}a) + A_\mu(x_{c\mu},x_{c\nu}-\tfrac{1}{2}a)] \\ &+ ig_0\tfrac{1}{2}[A_\mu(x_{c\mu},x_{c\nu}-\tfrac{1}{2}a), A_\nu(x_{c\mu}+\tfrac{1}{2}a,x_{c\nu})] \\ &+ ig_0\tfrac{1}{2}[A_\mu(x_{c\mu},x_{c\nu}+\tfrac{1}{2}a), A_\nu(x_{c\mu}-\tfrac{1}{2}a,x_{c\nu})] + O(a^2) , \end{aligned} \tag{87}$$

where the dependence on coordinates other than x_μ and x_ν is not shown explicitly. Summing over plaquettes, and noting that each plaquette is shared between two unit cells, we have for the discretized action

$$\beta S = \sum_P a^4 \mathrm{Tr}(F_P{}^2) . \tag{88}$$

Consider now an update in which the potential A_l on a single link l is changed. By Eq. (87), for each plaquette $P \supset l$, the field strength F_P is a linear functional of A_l, while for all other plaquettes the field strength has no dependence on A_l. Hence if we let $\{\phi\}$ be the set of potential components A^j_l, and $\{\psi\}$ be all other potential components, then in terms of these variables the action of Eq. (88) has precisely the form of Eqs. (76b) and (80). Moreover, if we choose any canonical gauge fixing (or if we do not gauge fix), then the integration measure has the form

$$d\mu = \prod \int dA^j_l \prod \int d\psi \tag{89}$$

required by Eq. (76a). Thus the conditions for validity of the algorithm of Eq. (82) are satisfied, and so an overrelaxed algorithm for the update $A_l \to A_l'$ is

$$\begin{aligned} &W[A_l \to A_l'] = \mathcal{N}\exp\left[-\sum_{P\supset l} a^4 \mathrm{Tr}(\cosh\theta F'_P + \sinh\theta F_P)^2\right] = \mathcal{N}\exp\left[\frac{-1}{\omega(2-\omega)}\sum_{P\supset l} a^4 \mathrm{Tr}[F'_P - (1-\omega)F_P]^2\right] , \\ &F'_P = F_P \mid_{A_l \to A_l'} . \end{aligned} \tag{90}$$

Although Eq. (90) exactly satisfies detailed balance with respect to the discretized action of Eq. (88), it is not exactly gauge invariant, and this limits its usefulness in simulations where maintaining exact gauge invariance is important. To get a computationally useful algorithm, we must construct an analog of Eq. (90) within the framework of Wilson's lattice gauge theory.[14] Because the lattice gauge theory action is not a multiquadratic form, it is not possible to construct an overrelaxed algorithm which exactly satisfies detailed balance with respect to the Wilson lattice action.

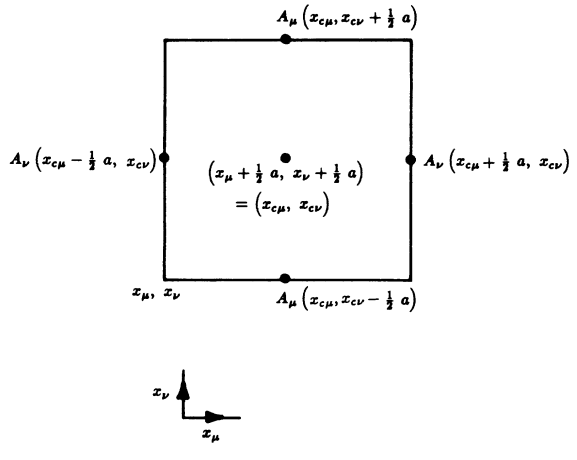

FIG. 1. Plaquette and potential variables in the x_μ-x_ν plane used to formulate Yang-Mills lattice field theory.

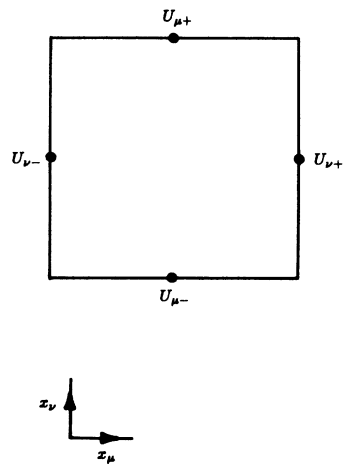

FIG. 2. Unitary matrices associated with the links of the plaquette of Fig. 1, which are used to formulate SU(n) lattice gauge theory.

However, this is a stronger requirement than is needed, since the lattice action is in any case only an order-a^2 approximation to the continuum action, and any member of the equivalence class of local, gauge-invariant lattice actions which differ from the Wilson action by relative order-a^2 terms in the continuum limit is equally suitable as a lattice action. We will show that it *is* possible to construct an exactly gauge-invariant lattice gauge theory transition probability by the procedure of Eqs. (76)–(78) above, which satisfies detailed balance with respect to an explicitly computable effective action differing from the Wilson action only by terms of relative order a^2 in the continuum limit.

To carry out this construction we rewrite Eq. (90) as

$$W[A_l \to A_l'] = \mathcal{N} \exp\left[-\sinh\theta\cosh\theta \sum_{P \supset l} a^4 \mathrm{Tr}(F_P' + F_P)^2 - (\cosh^2\theta - \sinh\theta\cosh\theta) \sum_{P \supset l} a^4 \mathrm{Tr}(F_P')^2 \right.$$
$$\left. - (\sinh^2\theta - \sinh\theta\cosh\theta) \sum_{P \supset l} a^4 \mathrm{Tr}(F_P)^2 \right] \tag{91}$$

and look for lattice gauge theory realizations of $a^4\mathrm{Tr}(F_P)^2$, $a^4\mathrm{Tr}(F_{P'})^2$, and $a^4\mathrm{Tr}(F_{P'} + F_P)^2$. Consider the plaquette P drawn in Fig. 1; in Fig. 2 we have redrawn this plaquette with the links labeled by the SU(n) matrices to which they correspond in lattice gauge theory. Let us assume that the link potential being updated is $A_l = A_\mu(x_{c\mu}, x_{c\nu} - \frac{1}{2}a)$, or in terms of lattice gauge theory variables, $U_l = U_{\mu -}$. We define

$$U_P = U_{\mu -} U_{\nu +} U_{\mu +} U_{\nu -}, \quad U_P' = U_{\mu -}' U_{\nu +} U_{\mu +} U_{\nu -} = U_P \mid_{U_l \to U_l'} . \tag{92}$$

Then the lattice gauge theory analog of Eq. (91) is

$$W[U_l \to U_l'] = \mathcal{N} \exp\left[-\sinh\theta\cosh\theta \sum_{P \supset U_l} \beta_0 \left[1 - \frac{1}{n} \mathrm{Re\,Tr}(U_P U_P') \right] \right.$$
$$- (\cosh^2\theta - \sinh\theta\cosh\theta) \sum_{P \supset U_l} \beta_0 \left[1 - \frac{1}{n} \mathrm{Re\,Tr} U_P' \right]$$
$$\left. - (\sinh^2\theta - \sinh\theta\cosh\theta) \sum_{P \supset U_l} \beta_0 \left[1 - \frac{1}{n} \mathrm{Re\,Tr} U_P \right] \right]$$
$$= \mathcal{N} \exp\left[\frac{1-\omega}{\omega(2-\omega)} \sum_{P \supset U_l} \beta_0 \left[1 - \frac{1}{n} \mathrm{Re\,Tr}(U_P U_P') \right] - \frac{1}{\omega} \sum_{P \supset U_l} \beta_0 \left[1 - \frac{1}{n} \mathrm{Re\,Tr} U_{P'} \right] \right.$$
$$\left. - \frac{1-\omega}{\omega} \sum_{P \supset U_l} \beta_0 \left[1 - \frac{1}{n} \mathrm{Re\,Tr} U_P \right] \right], \tag{93}$$

with Re denoting the real part, with $\mathcal{N}$ fixed by the requirement

$$\int d[U_l'] W[U_l \to U_l'] = 1 , \tag{94}$$

and with the parameter β_0 fixed in terms of n and the bare coupling g_0 by the usual relation[15]

$$\frac{\beta_0 {g_0}^2}{2n} = 1 . \tag{95}$$

We note that Eq. (93) is independent of the cyclic ordering of the link factors in U_P, as long as U_P' and U_P are ordered in the same way; in other words, by cyclic invariance of the trace we have

$$\mathrm{Tr}[(U_{\mu -} U_{\nu +} U_{\mu +} U_{\nu -})(U_{\mu -}' U_{\nu +} U_{\mu +} U_{\nu -})]$$
$$= \mathrm{Tr}[(U_{\nu -} U_{\mu -} U_{\nu +} U_{\mu +})(U_{\nu -} U_{\mu -}' U_{\nu +} U_{\mu +})]$$
$$= \mathrm{Tr}[(U_{\mu +} U_{\nu -} U_{\mu -} U_{\nu +})(U_{\mu +} U_{\nu -} U_{\mu -}' U_{\nu +})]$$
$$= \cdots . \tag{96}$$

The gauge invariance of Eq. (93) follows from the fact that since U_l and U_l' have the same behavior under gauge transformation, so do U_P and U_P':

$$U_P \to u_g U_P u_g^{-1}, \quad U_P' \to u_g U_P' u_g^{-1} , \tag{97}$$

and hence again by cyclic invariance of the trace the quantities $\mathrm{Tr}(U_P U_P')$, $\mathrm{Tr}U_P$, and $\mathrm{Tr}U_P'$ are exactly gauge invariant. Finally, since the U_l' dependence of Eq. (93) is of the form $\mathrm{Tr}(U_l' \tilde{U})$, with $\tilde{U}$ a linear combination of SU(n) matrices, in the case $n=2$ the efficient SU(2) algorithm of Creutz[16] can be applied to the generation of links U_l' distributed according to $W[U_l \to U_l']$.

The argument that Eq. (93) is an acceptable algorithm now runs as follows: Comparing with Eqs. (76)–(78), we see that Eq. (93) has precisely the form of the generalized algorithm, with $\{\phi\}$ corresponding to U_l, with $\{\psi\}$ corresponding to the other links in the plaquettes $P \supset U_l$, and with S_1 corresponding to those terms in the Wilson action involving plaquettes $P \supset U_l$. Hence Eq. (93) exactly satisfies detailed balance with respect to an

effective action, which differs from the Wilson action by a term proportional to $\ln(\mathcal{N}[U_l,\{\psi\};\theta]/\bar{\mathcal{N}}[\{\psi\},\theta])$, with $\bar{\mathcal{N}}$ the average of $\mathcal{N}[U_l,\{\psi\};\theta]$ over U_l. Since Eq. (93) only involves couplings of the link l to links in plaquettes P containing l, and since the entire construction is manifestly lattice gauge invariant, the effective action is local and lattice gauge invariant. Suppose that we can show that Eq. (93) differs from Eq. (91) by terms of relative order a^2 (absolute order a^6) in the continuum limit; then to leading order (absolute order a^4) the normalization factor $\mathcal{N}[U_l,\{\psi\};\theta]$ is independent of U_l, since by translation invariance $\mathcal{N}$ in Eq. (91) is independent of A_l. It then follows that $\ln(\mathcal{N}[U_l,\{\psi\};\theta]/\bar{\mathcal{N}}[\{\psi\};\theta])$ is of order a^6 in the continuum limit, and the equilibrium effective action for the algorithm of Eq. (93) is a member of the equivalence class of acceptable lattice actions.

To verify that in the continuum limit Eq. (93) reduces to Eq. (91) up to an error of order a^2, we start from the continuum limit of the individual link variables,

$$\begin{aligned} U_{\mu-} &= \exp[ig_0 a A_\mu(x_{c\mu},x_{cv}-\tfrac{1}{2}a)+O(a^3)] , \\ U_{v+} &= \exp[ig_0 a A_v(x_{c\mu}+\tfrac{1}{2}a,x_{cv})+O(a^3)] , \\ U_{\mu+} &= \exp[-ig_0 a A_\mu(x_{c\mu},x_{cv}+\tfrac{1}{2}a)+O(a^3)] , \\ U_{v-} &= \exp[-ig_0 a A_v(x_{c\mu}-\tfrac{1}{2}a,x_{cv})+O(a^3)] , \end{aligned} \tag{98}$$

with $\mathrm{Tr}O(a^3)=0$ since the U's are all SU(n) matrices and hence have unit determinant. For the products of adjacent links which appear in U_P, we have

$$\begin{aligned} &\exp[ig_0 a A_\mu(x_{c\mu},x_{cv}-\tfrac{1}{2}a)]\exp[ig_0 a A_v(x_{c\mu}+\tfrac{1}{2}a,x_{cv})] = e^{\Phi_+ +\delta_+} , \\ &\exp[-ig_0 a A_\mu(x_{c\mu},x_{cv}+\tfrac{1}{2}a)]\exp[-ig_0 a A_v(x_{c\mu}-\tfrac{1}{2}a,x_{cv})] = e^{\Phi_- +\delta_-} , \end{aligned} \tag{99}$$

with

$$\begin{aligned} \Phi_+ &= ig_0 a A_\mu(x_{c\mu},x_{cv}-\tfrac{1}{2}a)+ig_0 a A_v(x_{c\mu}+\tfrac{1}{2}a,x_{cv}) \\ &\quad -\tfrac{1}{2}g_0{}^2a^2[A_\mu(x_{c\mu},x_{cv}-\tfrac{1}{2}a),A_v(x_{c\mu}+\tfrac{1}{2}a,x_{cv})] , \end{aligned} \tag{100}$$

$$\begin{aligned} \Phi_- &= -ig_0 a A_\mu(x_{c\mu},x_{cv}+\tfrac{1}{2}a)-ig_0 a A_v(x_{c\mu}-\tfrac{1}{2}a,x_{cv}) \\ &\quad -\tfrac{1}{2}g_0{}^2a^2[A_\mu(x_{c\mu},x_{cv}+\tfrac{1}{2}a),A_v(x_{c\mu}-\tfrac{1}{2}a,x_{cv})] . \end{aligned}$$

The errors δ_+ and δ_- satisfy $\delta_+=O(a^3)$, $\delta_-=O(a^3)$, $\mathrm{Tr}\delta_+=\mathrm{Tr}\delta_-=0$. Moreover, since $\delta_-=\delta_+(a\to -a)$ and $\Phi_-=\Phi_+(a\to -a)$, we have that $\delta_++\delta_-=O(a^4)$ and that the commutator $[\Phi_+,\Phi_-]$ is odd in a. Hence for the plaquette product U_P we have

$$\begin{aligned} U_P &= U_{\mu-}U_{v+}U_{\mu+}U_{v-} \\ &= \exp\{\Phi_+ +\Phi_- +\tfrac{1}{2}[\Phi_+,\Phi_-]+O(a^4)\} , \end{aligned} \tag{101}$$

with $\frac{1}{2}[\Phi_+,\Phi_-]=O(a^3)$, with $\mathrm{Tr}O(a^4)=0$ and [referring to Eq. (87)] with

$$\Phi_+ +\Phi_- = ig_0 a^2 F_P . \tag{102}$$

For a general altered set of potentials A' we have

$$\begin{aligned} U'_P &= U'_{\mu-}U'_{v+}U'_{\mu+}U'_{v-} \\ &= \exp\{\Phi'_+ +\Phi'_- +\tfrac{1}{2}[\Phi'_+,\Phi'_-]+O(a^4)\} , \end{aligned} \tag{103}$$

again with $\frac{1}{2}[\Phi'_+,\Phi'_-]=O(a^3)$, with $\mathrm{Tr}O(a^4)=0$, and with

$$\Phi'_+ +\Phi'_- = ig_0 a^2 F'_P . \tag{104}$$

From these equations we find

$$\begin{aligned} \mathrm{Tr}U_P &= n-\tfrac{1}{2}g_0{}^2a^4\mathrm{Tr}(F_P)^2+O(a^6) , \\ \mathrm{Tr}U'_P &= n-\tfrac{1}{2}g_0{}^2a^4\mathrm{Tr}(F'_P)^2+O(a^6) , \\ \mathrm{Tr}(U_PU'_P) &= n-\tfrac{1}{2}g_0{}^2a^4\mathrm{Tr}(F_P+F'_P)^2+\Delta+O(a^6) , \\ \Delta &= \mathrm{Tr}\{(\Phi_+ +\Phi_- +\Phi'_+ +\Phi'_-)(\tfrac{1}{2}[\Phi_+,\Phi_-] \\ &\qquad +\tfrac{1}{2}[\Phi'_+,\Phi'_-])\} , \end{aligned} \tag{105}$$

which when $\Delta=0$ can be combined with Eqs. (95) and (93) to give Eq. (91). (Because of the identity $\mathrm{Tr}(\alpha[\alpha,\gamma])=0$, terms analogous to Δ do not appear in $\mathrm{Tr}U_P$ and $\mathrm{Tr}U'_P$.) Since the error term Δ is potentially of order a^5, to complete the derivation we must show that $\Delta=0$. Now repeated use of the identity

$$\mathrm{Tr}(\alpha[\beta,\gamma])=\mathrm{Tr}([\gamma,\alpha]\beta) , \tag{106}$$

which follows from cyclic invariance of the trace, shows that Δ can be reduced to the form

$$\Delta=\tfrac{1}{2}\mathrm{Tr}\{[\Phi'_+-\Phi_+,\Phi'_--\Phi_-](\Phi_++\Phi_-)\} . \tag{107}$$

In general $\Delta\neq 0$, but for the special case in which A' differs from A by the change of only the single link variable $A_\mu(x_{c\mu},x_{cv}-\frac{1}{2}a)$ or equivalently $U_{\mu-}$, we have $\Phi'_-=\Phi_-$ and Δ vanishes. Hence we have verified that Eq. (93) is a suitable overrelaxed algorithm for lattice gauge theory, for the case in which a single link at a time is updated.

To conclude this section, let us compare the small-ω continuum limit of the overrelaxation algorithm of Eq. (93) with the continuum limit of the lattice Langevin algorithm of Batrouni *et al.*; according to our analysis of Sec. I, these should correspond. Taking the continuum limit of the overrelaxation algorithm from Eq. (90), we have

$$\mathrm{Tr}(F_P'-F_P+\omega F_P)^2=\mathrm{Tr}\left[\frac{1}{a}(A_l'-A_l)+\tfrac{1}{2}ig_0[(A_l'-A_l),A_{\text{adjacent}}]+\omega F_P\right]^2, \tag{108a}$$

with A_{adjacent} the potential on the leg of the plaquette adjacent to and following A_l. Referring to Eq. (98), we recall that in the continuum limit $g_0 a A$ is the effective expansion parameter; approximating Eq. (108a) to leading-order accuracy in this expansion gives

$$\mathrm{Tr}(F_P'-F_P+\omega F_P)^2\approx\mathrm{Tr}\left[\frac{1}{a}(A_l'-A_l)+\omega F_P\right]^2=\frac{1}{2}\frac{1}{a^2}(A_l'^j-A_l^j)^2+\frac{1}{a}(A_l'^j-A_l^j)\omega F_P^j+A_l'\text{-independent}. \tag{108b}$$

In four dimensions there are six plaquettes P containing l, and so

$$\begin{aligned}\sum_{P\supset l}\mathrm{Tr}(F_P'-F_P+\omega F_P)^2&\approx\frac{3}{a^2}(A_l'^j-A_l^j)^2+\frac{1}{a}(A_l'^j-A_l^j)\omega\sum_{P\supset l}F_P^j+A_l'\text{-independent}\\&=3\left[\frac{1}{a}(A_l'^j-A_l^j)+\frac{\omega}{6}\sum_{P\supset l}F_P^j\right]^2+A_l'\text{-independent}.\end{aligned} \tag{109}$$

Substituting Eq. (109) into Eq. (90), dropping A_l'-independent terms and approximating $2-\omega\approx 2$, we have

$$W[A_l\to A_l']\approx\mathcal{N}\exp\left[-\tfrac{1}{4}a^4\frac{6}{\omega}\left[\frac{1}{a}(A_l'^j-A_l^j)+\frac{\omega}{6}\sum_{P\supset l}F_P^j\right]^2\right], \tag{110}$$

which can be rewritten as the stochastic difference equation

$$\frac{1}{a}(A_l'^j-A_l^j)\approx-\frac{\omega}{6}\sum_{P\supset l}F_P^j-\left[\frac{\omega}{6}\right]^{1/2}\frac{1}{a^2}\eta_j. \tag{111}$$

Now the lattice Langevin algorithm of Batrouni *et al.*, in the notation used above, takes the form

$$U_l'=e^{-F_l}U_l, \tag{112}$$

$$F_l=iT^j\left[\bar\epsilon\sum_{P\supset l}\left[\frac{-i\beta_0}{2n}\right]\mathrm{Tr}[T^j(U_P-U_P^\dagger)]+\bar\epsilon^{1/2}\eta_j\right].$$

Substituting

$$U_l=e^{ig_0aA_l},\quad U_P=e^{ig_0a^2F_P} \tag{113}$$

into Eq. (112) and working to leading order in the expansion in powers of $g_0 a A$, we get, for the continuum limit,

$$ig_0a(A_l'^j-A_l^j)T^j\approx-iT^j\left[\bar\epsilon\left[\frac{-i\beta_0}{2n}\right]\sum_{P\supset l}ig_0a^2F_P^j+(\bar\epsilon)^{1/2}\eta_j\right], \tag{114}$$

which on substituting $\beta_0 g_0^2/(2n)=1$ and factoring away the generators T^j becomes

$$\frac{1}{a}(A_l'^j-A_l^j)=-\frac{\bar\epsilon}{g_0^2}\sum_{P\supset l}F_P^j-\left[\frac{\bar\epsilon}{g_0^2}\right]^{1/2}\frac{1}{a^2}\eta_j. \tag{115}$$

Equations (111) and (115) have precisely the same structure, and give the identification

$$\omega=\frac{6}{g_0^2}\bar\epsilon, \tag{116}$$

again showing that the Langevin approach corresponds to the small-ω limit of the overrelaxation algorithm.

IV. DISCUSSION

In closing I comment briefly on the comparison between the acceleration strategy pursued above and that proposed by Batrouni *et al.*[9] Let us adopt as the "figure of merit" for an acceleration scheme the ratio of its inverse correlation time λ to that for an $\omega=1$ Jacobi iteration. As we have seen in Sec. II, for an optimally overrelaxed Gauss-Seidel iteration, the figure of merit is then N, the length of a side of the lattice in lattice units. By contrast, Batrouni *et al.* employ a Langevin method based on the Jacobi algorithm, and propose a method of Fourier acceleration in which the Langevin step size is taken to have a momentum dependence which compensates the critical slowing down at long wavelengths. In principle, their method can yield (up to logarithms) an inverse correlation time of $\lambda\sim\bar\epsilon a^{-1}$, with a the lattice spacing and $\bar\epsilon$ the small parameter which governs the Langevin step size. Thus, recalling Eq. (71), for the method of Batrouni *et al.*, the figure of merit can be as large as

$$\frac{\bar\epsilon a^{-1}}{ak_1^2}\sim\bar\epsilon(L/a)^2=\bar\epsilon N^2. \tag{117}$$

For lattices of moderate size, where $\bar\epsilon N\sim 1$, the overrelaxation method should be competitive with Fourier acceleration, but for very large lattices the Fourier method wins out, irrespective of the step size $\bar\epsilon$. Clearly, an optimal algorithm would combine the advantages of both, by permitting a step size of unity, as in the overrelaxed Gauss-Seidel approach, while replacing the factor $k_1\sim L^{-1}$ in Eq. (69) by a wave number of order a^{-1}. One possible way to try to achieve an improved algorithm is to combine overrelaxation with a mesh-doubling

lattice refinement scheme, as is done in the case of the minimization problem by the "hyper-overrelaxation" algorithm of Press[17] or the mesh-refinement-interpolation scheme of Adler and Piran.[7] A closely related approach is the "multigrid" Monte Carlo method advocated by Goodman and Sokal.[10] I hope to pursue these issues in future work.

Note added

In Sec. II we determined an optimum value of ω—let us call it ω_b—defined as the value of ω which minimizes the correlation time τ. By definition, this value of ω maximizes the asymptotic rate of decay of the correlation between two lattice configurations, as the "time" separation $\Delta T=\Delta M a$ between the two configurations becomes infinite (a=lattice spacing, ΔM=number of iterations separating the two configurations). However, in an actual Monte Carlo calculation this asymptotic decay rate is not the quantity which directly governs errors. What one does in a Monte Carlo calculation is to perform some total number M of iterations, but to only take every mth iterate as a member of the ensemble of configurations used for measurements, where $m\sim\tau/a$. Taking more configurations than this increases the amount of effort spent in measurement without improving the statistics, since the additional configurations are not statistically independent, while taking fewer configurations than this needlessly dilutes the statistics. Hence the quantity to be optimized is the absolute correlation between two configurations separated by m iterations, not the asymptotic rate of correlation decay.

This optimization problem also arises in the overrelaxation solution of differential equations, and the solution is as follows: For the iterations $i=0,1,\ldots$ one uses overrelaxation with a sequence of relaxation parameters ω_i with $\omega_0=1$ and with $\omega_i\to\omega_b$ for large i. In the case of iterations based on "odd/even" or "checkerboard" ordering, as opposed to the "typewriter" ordering used in Sec. II, the optimum ω_i's can be computed explicitly in terms of i and ω_b using Chebyshev polynomials. In the Monte Carlo application, one would use a "sawtooth" pattern of ω's, returning ω to 1 for the initial iteration after each configuration selected for measurement, and then stepping through the first m members of the Chebyshev or other optimal sequence. Taking $\omega\equiv\omega_b$ for all iterations can actually make the correlations *worse* after a finite number of iterations than simply using $\omega=1$, while the simple expedient of taking $\omega_0=1$ and $\omega_{i>1}=\omega_b$ already guarantees monotonically decreasing correlations. For a brief and lucid discussion of these issues see Hockney and Eastwood,[18] while for a detailed theoretical analysis see Vargas.[19] A simple, explicit, "checkerboard" iteration version of the calculation of Sec. II has recently been given by Neuberger,[20] and the Chebyshev method is directly applicable to Neuberger's scheme.

ACKNOWLEDGMENTS

I wish to thank F. Brown, M. Creutz, and C. Whitmer for stimulating conversations about lattice algorithms, and for informing me before publication of their numerical work suggesting that overrelaxation improves the correlation time. I also wish to thank G. P. Lepage for helpful conversations about the Langevin algorithm and the work of the Cornell group of Batrouni *et al.* I am grateful to M. Rassetti and T. Regge for the hospitality of the Institute for Scientific Interchange in Torino, and to E. Predazzi for the hospitality of the University of Torino and the use of its physics library. Partial support for this publication was provided by the U.S. Department of Energy under Grant No. DE-AC02-76ERO2220.

APPENDIX: EVALUATION OF THE INFINITE-SPACE GREEN'S FUNCTION

We evaluate here the time dependence of the Gauss-Seidel Green's function of Eq. (60), in the infinite-space limit in which the mode functions are

$$\psi_m(x)=\frac{1}{(2\pi)^{d/2}}e^{i\mathbf{k}\cdot\mathbf{x}},\quad \mathbf{k}\cdot\mathbf{x}=\sum_{\mu=1}^{d}k_\mu x_\mu\ . \tag{A1}$$

Substituting into Eq. (60), and noting that because of translation invariance there is no loss of generality in setting $x'=t'=0$, we have

$$G(x,0;t,0)=\frac{4}{d}\frac{2}{\omega}\frac{a^{d-1}}{(2\pi)^d}\int d^dk\, e^{i\mathbf{k}\cdot\mathbf{x}}\frac{e^{-pt-p\mathbf{x}\cdot\mathbf{n}/2}-e^{-qt-q\mathbf{x}\cdot\mathbf{n}/2}}{q-p}\theta(t+\tfrac{1}{2}\mathbf{x}\cdot\mathbf{n})\ ,$$

$$\mathbf{n}=(1,1,\ldots,1),\quad \mathbf{n}\cdot\mathbf{x}=\sum_{\mu=1}^{d}x_\mu,\quad p=2[C-(C^2-k^2/d)^{1/2}]\ , \tag{A2}$$

$$q=2[C+(C^2-k^2/d)^{1/2}],\quad k^2=\sum_{\mu=1}^{d}k_\mu{}^2\ .$$

We wish to evaluate the Fourier transform $g(l,t)$ defined by

$$G(x,0;t,0)=\frac{4}{d}\frac{2}{\omega}\frac{a^{d-1}}{(2\pi)^d}\int d^dl\, e^{il\cdot x}g(l,t)\ . \tag{A3}$$

Taking the inverse Fourier transform of Eq. (A2), the x and k integrations in the $d-1$ directions perpendicular to $\mathbf{n}$ can be done immediately, leaving [with $\mathbf{n}=d^{1/2}\hat{\mathbf{n}}$, $x=\mathbf{x}\cdot\hat{\mathbf{n}}$, $l=l\cdot\hat{\mathbf{n}}$, $k=\mathbf{k}\cdot\hat{\mathbf{n}}$, $(l^\perp)^2=l^2-l^2$]

$$g(l,l^{\perp},t)=\int\frac{dx}{2\pi}e^{-ilx}\int dk\,e^{ikx}\frac{e^{-pt-pd^{1/2}x/2}-e^{-qt-qd^{1/2}x/2}}{q-p}\theta(t+\tfrac{1}{2}d^{1/2}x)\ ,$$
$$\left.\begin{matrix}p\\q\end{matrix}\right\}=2(C\mp\{C^2-[k^2+(l^{\perp})^2]/d\}^{1/2})\ . \tag{A4}$$

To do the x integration, we make the change of variables

$$t+\tfrac{1}{2}d^{1/2}x=u\ , \tag{A5}$$

giving

$$g(l,l^{\perp},t)=\frac{1}{\pi d^{1/2}}\int_{-\infty}^{\infty}dk\,e^{-i(k-l)2t/d^{1/2}}I_u\ ,$$
$$I_u=\int_0^{\infty}du\,e^{i(k-l)2u/d^{1/2}}\frac{e^{-pu}-e^{-qu}}{q-p}=\frac{1}{q-p}\left[\frac{1}{p-i(k-l)2/d^{1/2}}-\frac{1}{q-i(k-l)2/d^{1/2}}\right] \tag{A6}$$
$$=\frac{1}{pq-i(k-l)2(p+q)/d^{1/2}-4(k-l)^2/d}\ .$$

Substituting p and q from Eq. (A4), and setting $k-l=w$, we are left with the single integral

$$g(l,l^{\perp},t)=\frac{1}{\pi d^{1/2}}\int_{-\infty}^{\infty}dw\,e^{-iw2t/d^{1/2}}\frac{1}{4[l^2+(l^{\perp})^2+2lw]/d-iw8C/d^{1/2}}\ . \tag{A7}$$

Since the denominator has a single zero in the lower half of the w complex plane, for $t<0$ we can close the contour up to get $g=0$, while for $t>0$ we can close the contour down to get the answer quoted in Eq. (75) of the text.

---

*Permanent address.

[1]See K. Binder, in *Phase Transitions and Critical Phenomena,* edited by C. Domb and M. S. Green (Academic, New York, 1976), Vol. 5b; M. Creutz, *Quarks, Gluons and Lattices* (Cambridge University Press, Cambridge, England, 1983), Chap. 18.

[2]S. L. Adler, Phys. Rev. D **23**, 2901 (1981). This earlier paper contains an inadvertent normalization error which is corrected in Eqs. (9) and (11) below.

[3]S. L. Adler and T. Piran, in *High-Energy Physics—1980,* proceedings of the XXth International Conference, Madison, Wisconsin, edited by L. Durand and L. G. Pondrom (AIP Conf. Proc. No. 68) (AIP, New York, 1981).

[4]A connection between underrelaxation and the Langevin approach has been independently noted, in a qualitative way, by M. Creutz, Phys. Rev. D **36**, 515 (1987).

[5]C. Whitmer, Phys. Rev. D **29**, 306 (1984); thesis, Princeton University, 1984.

[6]F. R. Brown and T. J. Woch, Phys. Rev. Lett. **26**, 2394 (1987).

[7]This is in general the case for the minimization problem; see S. L. Adler and T. Piran, Rev. Mod. Phys. **56**, 1 (1984). However, the *equality* of the thermalization and the correlation exponents found below is a feature only of the quadratic action, or free field, case.

[8]Although the choice of boundary conditions should not matter for sufficiently large lattices, the use of periodic boundary conditions leads to noncausal "wave propagation" in the continuum analog of the Gauss-Seidel iteration, and hence does not permit the use of the causality argument of Eqs. (61)–(63) below to solve the $t=0$ boundary condition.

[9]G. G. Batrouni, G. R. Katz, A. S. Kronfeld, G. P. Lepage, B. Svetitsky, and K. G. Wilson, Phys. Rev. D **32**, 2736 (1985).

[10]J. Goodman and A. D. Sokal, Phys. Rev. Lett. **56**, 1015 (1986); work in preparation; A. D. Sokal (private communication).

[11]P. R. Garabedian, Math. Tables Aids Comp. **10**, 183 (1956); see also W. F. Ames, *Numerical Methods for Partial Differential Equations* (Academic, New York, 1977), Chap. 3.

[12]A hyperbolic equation normally requires two initial conditions on an initial-value surface which is not a characteristic. However, the surface $t=0$ is a characteristic of Eq. (40) and so this general rule does not apply. In fact, since $dC+\sum_{\mu=1}^{d}\partial/\partial x_\mu$ is an invertible operator, the time development of G and $\bar{\phi}$ can be integrated forward from the values of G and $\bar{\phi}$ for all x at $t=0$.

[13]An alternative argument is to use the discrete form of the iteration for G given in Eqs. (28a) and (29) to infer that $G^{n,n'}_{l,l'}=0$ unless $n\geq n'$ and $n+i_1+\cdots+i_\mu\geq n'+i'_1+\cdots+i'_\mu$. The continuum limit of these inequalities bounding the region of support is $t-t'\geq 0$ and $t-t'+\sum_{\mu=1}^{d}(x_\mu-x'_\mu)\geq 0$. Adding the two inequalities we learn that G has support only in $s-s'\geq 0$, and so $G^{<}_{m}=0$ is the correct boundary condition.

[14]K. Wilson, Phys. Rev. D **10**, 2445 (1974).

[15]Creutz, *Quarks, Gluons and Lattices* (Ref. 1), p. 36.

[16]Creutz, *Quarks, Gluons and Lattices* (Ref. 1), pp. 132 and 133.

[17]W. Press, Harvard-Smithsonian Center for Astrophysics Report No. 971, 1978 (unpublished).

[18]R. W. Hockney and J. W. Eastwood, *Computer Simulation Using Particles* (McGraw-Hill, New York, 1981), pp. 174–181.

[19]R. S. Varga, *Matrix Iterative Analysis* (Prentice-Hall, Englewood Cliffs, NJ, 1962), Chap. 5.

[20]H. Neuberger, Phys. Rev. Lett. **59**, 1877 (1987).