PHYSICAL REVIEW D VOLUME 23, NUMBER 12 15 JUNE 1981

Over-relaxation method for the Monte Carlo evaluation of the partition function for multiquadratic actions

Stephen L. Adler
The Institute for Advanced Study, Princeton, New Jersey 08540
(Received 2 March 1981)

I formulate a successive over-relaxation (SOR) procedure for the Monte Carlo evaluation of the Euclidean partition function for multiquadratic actions (such as the Yang-Mills action with canonical gauge fixing). A convergence analysis for the quadratic-action (Abelian) case shows that as thermalization proceeds the mean nodal fields relax according to the difference equation arising from the standard SOR analysis of the associated classical Euclidean field equation. Hence, SOR should accelerate the thermalization process, just as it accelerates convergence in the numerical solution of second-order elliptic differential equations.

As has been much emphasized,[1] the Euclidean partition function is a fundamental tool for studying quantum field theories. For the case of a boson field theory[2] containing spin-0 scalar and spin-1 gauge fields, denoted collectively by ϕ, the partition function at inverse temperature β is given by the functional integral[3]

$$Z=\int d\phi_i \int_{\phi_i}^{\phi_i}[d\phi]\exp(-S)\,,$$
$$S=\int_0^\beta dt\int d^3x\,\mathcal{L}_E\,. \tag{1}$$

In Eq. (1) $\mathcal{L}_E$ is the Euclidean action density, including source terms, and the path integral extends over periodic paths, with $\phi(0)=\phi(\beta)=\phi_i$. I will restrict my attention in the following discussion to the case where $\mathcal{L}_E$ is a multiquadratic form (that is, it is at most quadratic in each individual field component), and will assume that the Euclidean action S is bounded from below. This restriction excludes interacting spin-0 fields from consideration (renormalizability for scalars requires a ϕ^4 term in the action), but allows ϕ to contain any number of non-Abelian spin-1 gauge fields, since the outer-product form of the gauge-field self-interaction is easily seen to imply a multiquadratic action.[4,5] Of course, when gauge fields are present, the partition function as written in Eq. (1) is formally infinite, as a result of integrations over gauge transformations which leave the action invariant. In reducing Eq. (1) to a discrete form for Monte Carlo evaluation, there are two natural strategies for dealing with the gauge infinities. The first, introduced by Wilson[6] and extensively studied[7] over the past few years, consists of using a discrete procedure in such a way that an exact, but compact gauge-invariance group remains, which can then be safely included in the Monte Carlo integration.[7] While this approach has many interesting features, it suffers from the drawbacks that (1) it is expressed in terms of unitary-matrix link variables, and has no natural discrete analogs of the gauge potentials and gauge fields, and (2) the multiquadratic form of the action is lost. A second natural strategy, which I will pursue in this paper, is to use the Faddeev-Popov method[1] to break the gauge invariance. In particular, if one chooses the canonical gauge fixing[8]

$$\begin{aligned} &b^1=0 \text{ in } R_4\text{: } -\infty<x_1,\ldots,x_4<\infty\,,\\ &b^2=0 \text{ in } R_3\text{: } x_1=0,-\infty<x_2,x_3,x_4<\infty\,,\\ &b^3=0 \text{ in } R_2\text{: } x_1=x_2=0,-\infty<x_3,x_4<\infty\,,\\ &b^4=0 \text{ in } R_1\text{: } x_1=x_2=x_3=0,-\infty<x_4<\infty \end{aligned} \tag{2}$$

for each gauge potential b^μ in ϕ, the gauge degeneracy is completely broken, with a Faddeev-Popov determinant which is constant. The functional integral can then be made discrete by taking the nodal values of the gauge potentials as the variables, and applying the standard replacement[9] of derivatives by finite differences to the action S. Denoting the set of node variables which are integrated over by $\{\phi\}=\{\phi(i),\ i=1,\ldots,N\}$, this procedure yields a multiple integral of the form

$$Z=\left[\prod_{i=1}^{N}\int_{-\infty}^{\infty}d\phi(i)\right]e^{-S[\{\phi\}]}\,, \tag{3}$$

with S a multiquadratic form which is bounded from below. Thus, for any node variable $\phi(k)$, S can be decomposed as

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

with A_k, B_k, and C_k functions of the subset of node variables $\{\phi\}_k\equiv\{\phi(i),\ i=1,\ldots,k-1,\ k+1,\ldots,N\}$.

Since in typical applications the dimensionality N of the multiple integral is very large, the

numerical estimation of Eq. (3) requires use of the Monte Carlo method.[7,10] Starting from any initial configuration $\{\phi_0\}$, one generates a sequence of successive configurations, or Markov chain, $\{\phi_1\}, \{\phi_2\}, \ldots, \{\phi_M\}, \ldots$ by repeated application of a transition probability $W[\{\phi\}\to\{\phi'\}]$. The transition probability W is chosen so that in the limit as M becomes infinite, the configurations in the chain are distributed according to the equilibrium probability density $P_{eq}[\{\phi\}]$,

$$P_{eq}[\{\phi\}] = e^{-S[\{\phi\}]}. \tag{5}$$

Sufficient conditions[11] on W to guarantee an asymptotic equilibrium probability distribution are the normalization condition

$$\left[\prod_{i=1}^{N}\int d\phi(i)'\right] W[\{\phi\}\to\{\phi'\}] = 1 \text{ for all } \{\phi\}, \tag{6a}$$

the ergodicity condition

$$P_{eq}[\{\phi\}]>0,\quad P_{eq}[\{\phi'\}]>0 \Rightarrow W[\{\phi\}\to\{\phi'\}]>0, \tag{6b}$$

and the detailed-balance condition

$$P_{eq}[\{\phi\}]W[\{\phi\}\to\{\phi'\}] = P_{eq}[\{\phi'\}]W[\{\phi'\}\to\{\phi\}]. \tag{6c}$$

In numerical work it is generally most convenient to change only a single node variable at a time. When specialized to this case, the form of W, for a step in which the node variable $\phi(k)$ is changed, is

$$W = w[\{\phi\}_k; \phi(k)\to\phi(k)'], \tag{7}$$

with w required to be ergodic and to satisfy the normalization and detailed-balance conditions

$$\int_{-\infty}^{\infty} d\phi(k)' w[\{\phi\}_k; \phi(k)\to\phi(k)'] = 1, \tag{8a}$$

$$P_{eq}[\{\phi\}_k, \phi(k)] w[\{\phi\}_k; \phi(k)\to\phi(k)'] = P_{eq}[\{\phi\}_k, \phi(k)'] w[\{\phi\}_k; \phi(k)'\to\phi(k)]. \tag{8b}$$

As is well known, the conditions of Eq. (8) do not fix w uniquely. The choice used in most Monte Carlo studies of gauge theories, motivated by the intuitive idea[7] of successively thermalizing the individual node variables, is

$$w[\{\phi\}_k; \phi(k)\to\phi(k)'] = N[\{\phi\}_k]^{-1} e^{-S[\{\phi\}_k, \phi(k)']},$$
$$N[\{\phi\}_k] = \int_{-\infty}^{\infty} d\phi(k)' e^{-S[\{\phi\}_k, \phi(k)']}, \tag{9}$$

which makes the distribution of new values $\phi(k)'$ completely independent of the old value $\phi(k)$ being replaced. For a multiquadratic action, where the dependence of S on $\phi(k)'$ is known explicitly from Eq. (4), the transition probability of Eq. (9) becomes

$$w[\{\phi\}_k; \phi(k)\to\phi(k)'] = (\pi A_k)^{-1/2} e^{-A_k[\phi(k)'-C_k]^2}. \tag{10}$$

This evidently corresponds to choosing a Gaussian distribution of the new kth-node value around a central value C_k, where C_k is the value of $\phi(k)$ which minimizes $S[\{\phi\}_k, \phi(k)]$.

As motivation for the generalization of Eq. (10) which I am about to discuss, let us briefly consider the problem of minimizing the discrete action functional $S[\{\phi\}]$. This can also be accomplished by an iterative procedure, the simplest form of which consists of starting from an initial configuration $\{\phi_0\}$, and then successively replacing each node value $\phi(k)$, $k=1,\ldots,N$ by the value $C_k[\{\phi\}_k]$ which minimizes S. Since S is nonincreasing under this relaxation procedure, in the limit of an infinite number of steps the minimum of S (assuming it exists and is unique[12]) will be attained. However, it is well known[9] that the procedure just outlined is not the optimal point-iterative algorithm for minimizing S; much more rapid convergence to the minimum can be obtained by using the *successive over-relaxation* (SOR) method in which $\phi(k)'$ is given by

$$\phi(k)' = \omega C_k + (1-\omega)\phi(k)$$
$$= C_k + (1-\omega)[\phi(k) - C_k], \tag{11}$$

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

$$0 \le S[\{\phi\}_k, \phi(k)] - S[\{\phi\}_k, \phi(k)']$$
$$= A_k[\phi(k) - \phi(k)'][\phi(k) + \phi(k)' - 2C_k]$$
$$= A_k[\phi(k) - \phi(k)']^2\left(\frac{2}{\omega} - 1\right), \tag{12}$$

giving the restriction

$$0<\omega<2. \tag{13}$$

When $\omega = 1$, Eq. (11) reduces to $\phi(k)' = C_k$, corresponding to the simple minimization procedure in which the new value $\phi(k)'$ is independent of the old value $\phi(k)$. When $\omega \neq 1$, the new value $\phi(k)'$ clearly retains a memory of the old value $\phi(k)$. In practice, optimum convergence is obtained by doing several iterations with $\omega = 1$, and then doing many iterations with a value $\omega = \omega_{opt}$ close to 2, adjusted to maximize the rate of final approach of S to its minimum.

Let us now return to the problem of evaluating the partition function of Eq. (3), and ask whether there is a parametrized, over-relaxation generalization of the Gaussian transition probability of Eq. (10). A simple investigation shows that such a generalization does exist, and is given by

$$w[\{\phi\}_k;\phi(k)\to\phi(k)']=\left[\frac{\omega(2-\omega)}{\pi A_k}\right]^{1/2}\exp\left\{-\left[\frac{A_k}{\omega(2-\omega)}\right][\phi(k)'-\omega C_k-(1-\omega)\phi(k)]^2\right\}, \tag{14a}$$

which can be rewritten as

$$w[\{\phi\}_k;\phi(k)\to\phi(k)']=(\pi A_k\cosh^2\theta)^{-1/2}\exp\{-A_k[\cosh\theta(\phi(k)'-C_k)+\sinh\theta(\phi(k)-C_k)]^2\},\ \omega-1=\tanh\theta. \tag{14b}$$

To verify Eq. (14), we note that it obviously satisfies the normalization condition of Eq. (8a), while since

$$\begin{aligned}A_k[\phi(k)-C_k]^2+A_k[\cosh\theta(\phi(k)'-C_k)+\sinh\theta(\phi(k)-C_k)]^2&\\ =A_k\cosh^2\theta\{[\phi(k)-C_k]^2+[\phi(k)'-C_k]^2\}+2A_k\cosh\theta\sinh\theta[\phi(k)-C_k][\phi(k)'-C_k]&\\ =A_k[\phi(k)'-C_k]^2+A_k[\cosh\theta(\phi(k)-C_k)+\sinh\theta(\phi(k)'-C_k)]^2,&\end{aligned} \tag{15}$$

it also satisfies the detailed balance condition of Eq. (8b). Hence, the transition probability of Eq. (14) provides an SOR method for the Monte Carlo evaluation of the Euclidean partition function for multiquadratic actions.

To determine whether SOR accelerates the thermalization process, let us analyze in detail the case where the action S is a quadratic (as opposed to a multiquadratic) form, corresponding to an Abelian gauge theory with external sources. Let $\phi(k)^M$ denote the value of the kth-node variable after M complete iterations, let

$$\{\phi\}_k^M=\{\phi(1)^M,\ldots,\phi(k-1)^M,\phi(k+1)^{M-1},\ldots,\phi(N)^{M-1}\}$$

denote the set of node variables which are passive when the kth-node variable is being altered during the Mth-iteration sweep, and let $P[\{\phi\};N(M-1)+k-1]$ be the joint probability distribution of the node variables after $N(M-1)+k-1$ individual node replacements. Then we evidently have

$$P[\{\phi\}_k^M,\phi(k)^M;N(M-1)+k]=\int_{-\infty}^{\infty}d\phi(k)^{M-1}w[\{\phi\}_k^M;\phi(k)^{M-1}\to\phi(k)^M]P[\{\phi\}_k^M,\phi(k)^{M-1};N(M-1)+k-1], \tag{16}$$

which tells us how the joint probability distribution evolves from step to step. Integrating Eq. (16) with respect to $\phi(k)^M$, and using the normalization condition of Eq. (8a), we learn that

$$\int_{-\infty}^{\infty}d\phi(k)^M P[\{\phi\}_k^M,\phi(k)^M;N(M-1)+k]=\int_{-\infty}^{\infty}d\phi(k)^{M-1}P[\{\phi\}_k^M,\phi(k)^{M-1};N(M-1)+k-1], \tag{17}$$

which means that the joint probability distribution for the subset of node variables $\{\phi\}_k$ [with $\phi(k)$ integrated out] is unchanged during the iterative step in which $\phi(k)$ is altered. This in turn implies that the mean value of any node variable $\bar{\phi}(k)$, defined by

$$\bar{\phi}(k)=\left[\prod_{i\neq k}\int_{-\infty}^{\infty}d\phi(i)\right]\int_{-\infty}^{\infty}d\phi(k)\phi(k)P[\{\phi\};\ldots], \tag{18}$$

changes only during an iteration step in which $\phi(k)$ is altered, and so is uniquely specified by the notation $\bar{\phi}(k)^M$, which gives its value after M complete iterations. To study the evolution of the mean values, we multiply Eq. (16) by $\phi(k)^M$ and integrate, giving

$$\begin{aligned}\bar{\phi}(k)^M&=\left[\prod_{i=1}^{k-1}\int_{-\infty}^{\infty}d\phi(i)^M\right]\int_{-\infty}^{\infty}d\phi(k)^M\phi(k)^M\left[\prod_{i=k+1}^{N}\int_{-\infty}^{\infty}d\phi(i)^{M-1}\right]P[\{\phi\}_k^M,\phi(k)^M;N(M-1)+k]\\ &=\left[\prod_{i=1}^{k-1}\int_{-\infty}^{\infty}d\phi(i)^M\right]\left[\prod_{i=k}^{N}\int_{-\infty}^{\infty}d\phi(i)^{M-1}\right]\\ &\quad\times\int_{-\infty}^{\infty}d\phi(k)^M\{[\phi(k)^M-\omega C_k-(1-\omega)\phi(k)^{M-1}]_{(1)}+[\omega C_k+(1-\omega)\phi(k)^{M-1}]_{(2)}\}\\ &\quad\times w[\{\phi\}_k^M;\phi(k)^{M-1}\to\phi(k)^M]P[\{\phi\}_k^M,\phi(k)^{M-1};N(M-1)+k-1].\end{aligned} \tag{19}$$

Comparing with Eq. (14a), we see that the contribution of the term labeled $[\ ]_{(1)}$ vanishes, while using Eq. (8a) the contribution of the term labeled $[\ ]_{(2)}$ simplifies to give

$$\bar{\phi}(k)^M=\left[\prod_{i=1}^{k-1}\int_{-\infty}^{\infty}d\phi(i)^M\right]\left[\prod_{i=k}^{N}d\phi(i)^{M-1}\right][\omega C_k[\{\phi\}_k^M]+(1-\omega)\phi(k)^{M-1}]P[\{\phi\}_k^M,\phi(k)^{M-1};N(M-1)+k-1]. \quad (20)$$

Up to this point the analysis is completely general, and applies to multiquadratic as well as quadratic actions. Specializing now to the case of quadratic actions, for which C_k is a linear functional of its arguments, Eq. (20) becomes

$$\bar{\phi}(k)^M=\omega C_k[\{\bar{\phi}\}_k^M]+(1-\omega)\bar{\phi}(k)^{M-1},$$
$$\{\bar{\phi}\}_k^M=\{\bar{\phi}(1)^M,\ldots,\bar{\phi}(k-1)^M,\bar{\phi}(k+1)^{M-1},\ldots,\bar{\phi}(N)^{M-1}\}. \quad (21)$$

Thus, under SOR thermalization for a quadratic action, the mean nodal values evolve according to Eq. (21), which is just the difference equation encountered in the SOR minimization of the action S. Since SOR is known to accelerate the minimization process, Eq. (21) implies that it will accelerate convergence of the thermalization process as well. Although the precise statement of Eq. (21) can be made only for quadratic actions, the general conclusion reached here, that SOR accelerates thermalization, is very likely to carry over to the general multiquadratic case as well, much as SOR accelerates the minimization[4] of multiquadratic as well as quadratic actions.

As compared with the conventional[6,7] lattice gauge theory approach, the strategy for evaluating the partition function outlined above may have several advantages. First, since the potentials remain as the variables, there are natural discrete analogs of the gauge potentials and gauge fields, which should permit the study of such questions[13] as the behavior of the effective action for weak fields. Second, since the mean node variables for the Abelian theory thermalize according to the SOR equation encountered in minimizing the discrete action S, and since this equation is just the conventional[9] discrete version of the classical Euclidean field equation derived from the continuum S, the Abelian theory will never give a confining potential for static sources. Thus, if confinement is found in the non-Abelian case, it should not be as an artifact of the discrete procedure. Finally, the SOR method outlined above may well be computationally faster than the lattice gauge theory method, both because of the possibility of acceleration of the thermalization process, and because the Gaussian distribution of Eq. (14a) can be obtained from an array of prestored, normally distributed random numbers by the calculation of a single square root and a relatively small number of arithmetic operations. Detailed numerical experiments will, of course, be needed to see if these conjectured gains are realized in practice.

I wish to thank R. F. Dashen, T. Piran, and S. Samuel for interesting conversations. This work was supported by the U. S. Department of Energy under Grant No. DE-AC02-76ER02220.

[1]For a review, see E. S. Abers and B. W. Lee, Phys. Rep. 9C, 1 (1973).

[2]The Monte Carlo evaluation of fermion functional integrals involves special problems not present in the boson case. For a discussion, see F. Fucito *et al.*, Report No. CERN TH-2960, 1980 (unpublished) and G. De Angelis, D. de Falco, and F. Guerra, Phys. Rev. Rev. D 23, 1747 (1981).

[3]D. Gross, R. Pisarski, and L. Yaffe, Rev. Mod. Phys. 53, 43 (1981).

[4]S. L. Adler and T. Piran, in *High Energy Physics—1980*, proceedings of the XX International Conference, Madison, Wisconsin, edited by L. Durand and L. G. Pondrom (AIP, New York, 1981).

[5]A ϕ^4 term can be formally expressed in multiquadratic form by replacing it with $-\sigma^2+2\sigma\phi^2$ with σ an auxiliary variable, but this expression is unbounded from below. (The restriction to multiquadratic actions which are bounded from below, together with the requirement that the action contain no dimensional parameters, automatically selects theories with dynamical scale-invariance breaking. I suspect that this is more than a coincidence.)

[6]K. Wilson, Phys. Rev. D 10, 2445 (1974); for a review, see M. Creutz, Rev. Mod. Phys. 50, 561 (1978). See also J. Kogut and L. Susskind, Phys. Rev. D 11, 395 (1975).

[7]M. Creutz, L. Jacobs, and C. Rebbi, Phys. Rev. D 20, 1915 (1979); M. Creutz, *ibid.* 21, 2308 (1980).

[8]I. M. Singer (unpublished); M. B. Halpern, Phys. Rev. D 19, 517 (1979).

[9]W. F. Ames, *Numerical Methods for Partial Differential Equations* (Academic, New York, 1977), Chap. 3; E. Bauer, O. Betancourt, and P. Garabedian, *A Computational Method in Plasma Physics* (Springer, New York, 1978), Chap. 3.

[10]N. Metropolis, A. W. Rosenbluth, A. H. Teller, and E. Teller, J. Chem. Phys. 21, 1087 (1953). For a detailed review, see K. Binder, in *Phase Transitions and Critical Phenomena*, edited by C. Domb and M. S. Green (Academic, New York, 1976), Vol. 5b.

[11]The detailed-balance condition is the simplest way of satisfying a more general requirement called the homogeneous-state condition. See K. Binder, Ref. 10, for a discussion.

[12]A minimum will exist, in general, if S is bounded from below. If there are several relative minima, one which can be reached by a downhill path from the starting configuration is selected.

[13]S. L. Adler, this issue, Phys. Rev. D 23, 2905 (1981).