

LATTICE GAUGE THEORIES AND MONTE CARLO ALGORITHMS*

Michael Creutz, Physics Department, Brookhaven National Laboratory,
Upton, NY 11973

Lattice gauge theory has become the primary tool for non-perturbative calculations in quantum field theory. These lectures review some of the foundations of this subject. The first lecture reviews the basic definition of the theory in terms of invariant integrals over group elements on lattice bonds. The lattice represents an ultraviolet cutoff, and renormalization group arguments show how the bare coupling must be varied to obtain the continuum limit. Expansions in the inverse of the coupling constant demonstrate quark confinement in the strong coupling limit. The second lecture turns to numerical simulation, which has become an important approach to calculating hadronic properties. Here I discuss the basic algorithms for obtaining appropriately weighted gauge field configurations. The third lecture turns to algorithms for treating fermionic fields, which still require considerably more computer time than needed for purely bosonic simulations. Some particularly promising recent approaches are based on global accept-reject steps and should display a rather favorable dependence of computer time on the system volume.

LECTURE I: INTRODUCTION

Several speakers at this meeting are presenting recent numerical results from lattice gauge theory. Indeed, this technique has become the dominant tool for the study of non-perturbative phenomena in field theory. In the past, strong interaction physics dominated lattice simulations, with numerous calculations of the confining string tension, hadronic masses, and properties of the quark/gluon plasma. Recently weak interactions have begun to play an important role, with studies of corrections to weak matrix elements and studies of symmetry breaking and properties of the Higg's particle. You will hear much about all these topics at this meeting.

The goal of my lectures is to provide a pedagogical introduction to the subject, leaving the recent results to the other speakers. In this first lecture I give an overview of the general formulation and discuss how confinement arises in the strong coupling expansion. In my second lecture I discuss some of the algorithms used for numerical simulations of these systems, concentrating on the pure gauge theory. In my third lecture I will turn to the difficulties with simulating fermionic fields and discuss some promising recent approaches involving global trial changes to the dynamical fields. For a general reference on lattice gauge theory see ref. [1]. For a more specific review of Monte Carlo algorithms, see [2], and for a detailed discussion of the global algorithms for fermions, see [3].

* This manuscript has been authored under contract number DE-AC02-76CH00016 with the U.S. Department of Energy. Accordingly, the U.S. Government retains a non-exclusive, royalty-free license to publish or reproduce the published form of this contribution, or allow others to do so, for U.S. Government purposes.

I begin by defining Wilson's [4] lattice gauge theory. For simplicity I start with a four dimensional hypercubic lattice. For definiteness, assume that the lattice has N^4 sites and has periodic boundary conditions. For every ordered nearest neighbor pair of sites (i, j) on this lattice, I have a bond variable U_{ij} which is an element of the gauge group G . I will assume that G is a compact unitary group. To study the theory of the strong interactions, where the gauge fields are the gluons which bind quarks into hadrons, the gauge group is $SU(3)$. Thus I consider a configuration space which consists of $4 \times N^4$ group elements. When a bond is traversed in the reverse order, then the group element on the link is inverted

$$U_{ji} = U_{ij}^{-1}. \quad (1)$$

Associated with every configuration of this system is an action

$$S = \sum_p (1 - 1/n \text{Re Tr } U_p). \quad (2)$$

Here the trace is in some representation, usually the fundamental, of the group, and n is the dimension of the matrices in that representation. The sum is over all elementary squares, or "plaquettes," p , and U_p denotes an ordered product of the fundamental link variables around the given plaquette. Because of the trace it does not matter on which corner of the square the product starts. Because of the real part being taken in Eq. (2) and because the group is unitary, the direction taken around the square is unimportant.

The next step is to exponentiate the action and study the path integral over all gauge fields. This reveals a mathematical equivalence of the path integral with the statistical mechanics of this system of group elements. In this analogy, the exponentiated action plays the role of the Boltzmann weight. Correlation functions in this statistical system correspond to the field theoretical Green's functions continued to imaginary time. Thus I am led to consider the partition function

$$Z = \int (dU) e^{-\beta S}. \quad (3)$$

The parameter β is proportional to the inverse square of the bare gauge coupling g_0 used in conventional perturbative studies. For the gauge group $SU(n)$ the relation is

$$\beta = 2n/g_0^2. \quad (4)$$

The expectation value for some function F of the gauge variables is

$$\langle F \rangle = Z^{-1} \int (dU) F(U) e^{-\beta S}. \quad (5)$$

In these equations the integration is to be taken over all link variables using the group invariant measure. For compact groups this measure is unique and satisfies

$$\int dU f(U) = \int dU f(UU') = \int dU f(U'U) = \int dU f(U^{-1}) \quad (6)$$

where $f(U)$ is an arbitrary function over the group and U' is an arbitrary fixed group element. The measure is normalized such that

$$\int dU \, 1 = 1. \quad (7)$$

Eq. (6) can be schematically written $dU = dUU' = dU'U = dU^{-1}$.

In addition to the gauge fields on the links, one can readily add matter fields to the theory. These conventionally reside on the lattice sites. The interaction with the gauge fields is constructed to display the gauge symmetries to be discussed below. For simplicity, however, I will ignore such fields until my last lecture.

This completes the formal definition of a lattice gauge theory. The motivation for looking at this system can be found in the many reviews of the topic, for example reference [1]. Briefly, the classical continuum limit of the above action reproduces the Yang Mills [5] theory and the integration over the links gives a regularized Feynman path integral for the quantization of the system. The connection between the link variables and the vector potentials A_μ is

$$U_{ij} = \exp(ig_0 A_\mu a) \quad (8)$$

where a is the lattice spacing and the vector index μ lies along the direction of the bond. If I assume that A_μ is smooth, then as a goes to zero the Wilson action reduces to the standard Yang-Mills action

$$S = \int \frac{1}{4} F_{\mu\nu} F_{\mu\nu}. \quad (9)$$

The Wilson theory is formulated directly in Euclidian space. Thus all four dimensions are equivalent. In particular this means that the time evolution operator is e^{-Ht} rather than the more usual e^{-iHt} . As it is the same Hamiltonian in each case, no physical information is lost by this Wick rotation to imaginary time. The advantage of working in Euclidian space is that the Boltzmann weight is real and the analogy with a statistical mechanics problem is explicit.

For the quantum theory, I will take the point of view that the continuum limit of lattice gauge theory, if it exists, can be regarded as a definition of quantized gauge fields. The phenomenon of asymptotic freedom gives some useful information on how to take this limit. As the lattice spacing is reduced one will encounter the well known divergences of quantum field theory. In particular, this means that the bare coupling must be renormalized. The variation of the bare coupling with cutoff defines the renormalization group function, which can be calculated perturbatively

$$a \frac{\partial g_0}{\partial a} = \gamma(g_0) = \gamma_0 g_0^3 + \gamma_1 g_0^5 + \dots \quad (10)$$

For a non-Abelian gauge theory the famous phenomenon of asymptotic freedom [6] is manifested in the positive sign of the number γ_0 .

For our purposes it is useful to rewrite the renormalization group equation in integrated form and express the lattice spacing as a function of the bare coupling

$$a = \frac{1}{\Lambda} (g_0^2 \gamma_0)^{-\gamma_1/2\gamma_0^2} e^{-1/2\gamma_0 g_0^2} \times (1 + O(g_0^2)). \quad (11)$$

Here Λ is an integration constant and sets the overall scale of the theory.

Note that to take the lattice spacing to zero requires taking the bare coupling to zero. As the bare coupling can be regarded as an effective coupling at the scale of the cutoff, this is equivalent to the usual statement of asymptotic freedom that the effective coupling becomes small at short distances.

When working on the lattice it is quite natural to measure masses in units of the lattice spacing. A mass is extracted from the Yukawa law behavior of the correlation between two widely separated operators. If a particular particle has a finite physical mass m in the continuum limit, then Eq. (11) implies that the dimensionless combination ma will show an exponential decrease with increasing inverse coupling squared. The coefficient of this “scaling” behavior gives the particle mass in units of Λ . It is such a scaling that is looked for in essentially all numerical lattice gauge calculations attempting to extract physical observables.

Note that the factor of Λ will drop out of any mass ratios. Indeed, in the continuum limit the pure gauge theory should make parameter free predictions for all dimensionless observables. When quarks are added to the theory the only parameters are the quark masses (in units of Λ).

The action in Eq. (2) has an enormous symmetry. Suppose I associate an arbitrary group element g_i with every site i on our lattice. Using these, I can construct a new link element on each bond

$$U'_{ij} = g_i U_{ij} g_j^{-1}. \quad (12)$$

Since the dynamics involves the trace of link variables multiplied around closed loops, the factors of g will cancel in the calculation of the action for the new links. This exact local symmetry is the gauge symmetry of the model. On our N^4 site lattice, the symmetry group is the gauge group raised to the N^4 power.

The gauge symmetry implies that expectation values of gauge non-invariant quantities vanish. For example, consider the expectation of a single link variable

$$\begin{aligned} \langle U_{ij} \rangle &= Z^{-1} \int (dU) U_{ij} e^{-\beta S} \\ &= Z^{-1} \int (dU) g_i U_{ij} g_j^{-1} e^{-\beta S} \\ &= g_i \langle U_{ij} \rangle g_j^{-1}. \end{aligned} \quad (13)$$

Since g_i and g_j are arbitrary, $\langle U_{ij} \rangle$ must vanish (except for the trivial case where the gauge group has but one element). This result that gauge variant quantities have vanishing expectation values is known as Elitzur's theorem [7]. Because the symmetry is a local one,

only a finite number of degrees of freedom are involved in the above change of variables, and there is no possibility of a spontaneous breaking of this symmetry in the infinite volume limit.

In continuum formulations of gauge theory, one usually adds a gauge fixing term to the action. This can be done on the lattice as well. Consider the expectation of some observable F

$$\langle F \rangle = \frac{1}{Z} \int (dU) e^{-\beta S} F(U). \quad (14)$$

Assume that F is gauge invariant

$$F(U_{ij}) = F(g_i U_{ij} g_j^{-1}). \quad (15)$$

Now I introduce the gauge fixing function $h(U)$ which is an arbitrary function of the link variables. Then it is readily verified that the expectation of F can be written

$$\langle F \rangle = \frac{1}{Z} \int (dU) e^{-\beta S} F(U) \frac{h(U)}{\int (dg) h(gUg^{-1})}. \quad (16)$$

Note that the denominator inserted into this integral is just the integral of h over all gauges. If the gauge fixing involves a gaussian function, then this integral gives rise to the standard Faddeev-Popov [8] determinant. One important feature of the lattice formulation, however, is that the volume of the gauge orbits is finite and therefore this gauge fixing, while possible, is no longer necessary.

For observables one should look for gauge invariant quantities. One such is the trace of the product of link variables around a closed loop. The expectation of this is the famous Wilson loop. Confinement in the pure gauge theory is signaled by an exponential decrease with the minimal area enclosed by the loop. The coefficient of this area law, or "string tension," is a non-local order parameter which is useful for distinguishing certain phases of lattice gauge models. It is a physical observable with the dimensions of mass squared, and is the coefficient of a linear force between widely separated quarks.

The area law behavior of Wilson loops arises quite naturally in the strong coupling limit of the theory. Indeed, this was one of the major points of Wilson's original paper [4]. When the bare coupling becomes large, the parameter β is small and one can consider a power series expansion of the Boltzmann weight $e^{-\beta S}$. Any term in this expansion will involve various powers of the link variables, which must be integrated over. For $SU(n)$ the resulting integrals can be done using a set of graphical rules which I will not go into here [9]. Three simple integrals which are particularly useful are

$$\begin{aligned} \int dg g^{\alpha\beta} &= 0 \\ \int dg g^{\alpha\beta} g^{\dagger\gamma\delta} &= \frac{1}{n} \delta^{\alpha\delta} \delta^{\beta\gamma} \\ \int dg g^{\alpha_1\beta_1} \dots g^{\alpha_n\beta_n} &= \frac{1}{n} \epsilon_{\alpha_1, \dots, \alpha_n} \epsilon_{\beta_1, \dots, \beta_n}. \end{aligned} \quad (17)$$

where the $SU(n)$ matrix indices have been denoted by Greek superscripts. Here ϵ represents the totally antisymmetric tensor with $\epsilon_{1,\dots,n} = 1$.

Consider a large Wilson loop. If I ignore the Boltzmann weight in calculating its expectation, then the integrals over the links on the loop will give zero by the first part of Eq. (17). To get a non-vanishing result, one can extract plaquette factors from the exponentiated action to match these links with reversed links, enabling the use of the second relation in Eq. (17). This must be repeated until there are no unmatched links in the integral. This requires that enough powers of β be taken so that a minimal surface with the loop as boundary can be tiled with plaquettes taken from the Boltzmann factor. Evaluating all the integrals then gives the leading behavior for expectation value of the loop. For the gauge group $SU(n)$ with $n > 2$ this gives

$$\langle W \rangle = \left(\frac{\beta}{2n^2} \right)^A + \dots \quad (18)$$

where A is the area, in number of plaquettes, of the minimal surface tiling the loop, and the corrections are of higher order in β . For $SU(2)$, where the conjugate of the fundamental representation is not a different representation, $\frac{\beta}{2n^2}$ is replaced by $\frac{\beta}{4}$.

This factor of beta raised to the area of the loop is precisely the exponential suppression signaling confinement. Extracting the string tension from $\langle W \rangle \sim e^{-KA}$ gives the strong coupling behavior $K \sim \log\left(\frac{2n^2}{\beta}\right)$ for $SU(n)$, and $K \sim \log\left(\frac{4}{\beta}\right)$ for $SU(2)$.

Although confinement is natural in the strong coupling limit, the relevance of this result to the continuum is unclear. Indeed, the earlier discussion of asymptotic freedom showed that the continuum limit requires taking the bare coupling to zero. As an infinite statistical system can undergo phase transitions, so might lattice gauge theory, with the string tension vanishing below some finite value of the coupling. While there is no rigorous proof, Monte Carlo studies have indicated a rather smooth behavior of the string tension as a function of coupling for the $SU(2)$ and $SU(3)$ theories. The situation is rather different for $U(1)$, the gauge group of electrodynamics, which does appear to have a weak coupling phase with free photons.

Lattice gauge theory can also be formulated as an ordinary quantum mechanics problem in continuous time with the canonical coordinates being group elements on the links of a spatial lattice. This Hamiltonian formulation [10] can be related to the above Lagrangian formalism by fixing the temporal gauge ($A_0 = 0$) and then taking a continuous time limit [11]. In this approach, the link variables are operators \hat{U}_{ij} in a Hilbert space. Conjugate to them are conjugate "electric field" operators l_{ij}^α . The commutation relations amongst these objects are

$$\begin{aligned} [l_{ij}^\alpha, l_{ij}^\beta] &= i f^{\alpha\beta\gamma} l_{ij}^\gamma \\ [l_{ij}^\alpha, \hat{U}_{ij}] &= -\lambda^\alpha \hat{U}_{ij}. \end{aligned} \quad (19)$$

Here the group generators are denoted by λ^α , the group structure constants by $f^{\alpha\beta\gamma}$, and the Greek superscripts run up to the number of generators. In an explicit parameterization

of the group, the l_{ij}^α are represented by differential operators. In terms of these quantities the Hamiltonian takes the form

$$H = \frac{g^2}{2a} \sum_{\{ij\}} l_{ij}^2 + \frac{2}{g^2 a} \sum_p \text{Re Tr } \hat{U}_p. \quad (20)$$

Here g is the bare gauge coupling, a is the lattice spacing, the first sum is over all links and the second sum is over all spacelike plaquettes. Note that l^2 is just the quadratic Casimir operator on the respective link. The two terms in this equation can be regarded as the electric and magnetic contributions to the energy.

Note that the l operators do not commute among each other, unlike the case of the usual electric field operators in the continuum formulation of Yang Mills fields. This is because for convenience the l operators operate on group elements from the left side. Indeed there are other definitions for these operators; in particular the second part of Eq. (19) could be redefined to include the factor of λ on the right. The conventional commuting electric fields are neither of these but rather directly shift the vector potential which is exponentiated to give the link operators. In the continuum limit all these operators become the same, but on the lattice it is more convenient to work with the above construction.

In the strong coupling limit of large g the electric field term dominates the Hamiltonian and one can do standard perturbation theory in the second term. The natural basis for this expansion is in terms of definite representations of the gauge group on each link. The kinetic or electric term is minimized by placing all links into singlet states with $l_{ij}^2 = 0$. The potential or magnetic term then can excite links into intermediate states involving higher representations. For a review of this approach see ref. [12].

LECTURE II: NUMERICAL SIMULATION OF BOSONIC FIELDS

For the past several years research in lattice gauge theory has been dominated by Monte Carlo simulations. The basic idea is to approximate the integral in Eq. (5) by a sum over a finite number of configurations of the system with a weighting proportional to the Boltzmann factor $e^{-\beta S}$. Using the formal analogy with statistical mechanics, one should find a set of configurations typical of thermal equilibrium.

For simplicity in the following discussion, let me begin by considering just a single group element U . Consider the problem of stochastically picking group elements with a probability distribution

$$P_{eq}(U) \propto e^{-\beta S(U)} \quad (21)$$

where $S(U)$ is the action. The differential measure to be used here is the same invariant measure used for integration above.

For groups such as $SU(2)$ where the measure is rather simple it is not difficult to generate elements directly with a particular distribution. However, at first it might seem that for other groups with more complicated invariant measures it might be quite difficult

to generate elements with an arbitrary distribution. It turns out, however, that by doing a random walk over the group manifold this can be quite easy.

If one has an ensemble of group elements uniformly distributed over the group, the invariance of the measure insures that if all elements in this set are multiplied by some fixed group element, then the new ensemble will also be uniform. This fact provides a way to obtain such a uniform distribution. Suppose I have a set of factors which upon repeated multiplication can cover the group. Then construct a Markov chain of group elements by starting with an arbitrary element and repeatedly multiplying the current element U by factors chosen independent of U . This sequence will asymptotically uniformly cover the group. In a practical simulation the factors could be stored in a finite table.

Actually, rather than a uniform covering, Eq. (21) asks for a weighted distribution. At the simplest level this could be obtained by merely generating first a uniform distribution, and then accepting elements from this set with a probability proportional to the Boltzmann weight. This scheme will, however, have difficulties when the weight factor is highly peaked, in which case only a small number of elements will be accepted.

More efficient schemes follow by modifying the above Markov chain. This is usually done by incorporating an accept-reject procedure into the random walk, as originally suggested by Metropolis *et al.* [13]. To be explicit, construct U' as a tentative next element in our Markov chain by multiplying the current element U by a factor as indicated above. Thus this element is selected with some trial distribution $P_{T,U}(U')$. I place the subscript U on this distribution as a reminder that the distribution depends on the old element U . After U' is constructed, it is accepted only conditionally as the next member of the Markov chain. The acceptance criterion is not unique, but is conventionally taken to be

$$P_{acc} = \min \left[1, \frac{P_{T,U'}(U)}{P_{T,U}(U')} e^{-\beta(S' - S)} \right]. \quad (22)$$

Here S and S' denote the old and trial actions, respectively. Note that this requires knowledge of the reverse probability for choosing U as the trial had the old element been U' . If the new element is not accepted, the next term in the Markov chain is just a repeat the old element. It is readily verified that this process leads to a sequence of states where the asymptotic probability for finding a particular state is proportional to the Boltzmann weight. For a detailed discussion of this, see Ref. 2.

There are many variations on this approach depending on the choice of $P_{T,U}(U')$. Indeed, if the group measure is simple, one may be able to directly pick U' with the appropriate Boltzmann weight, in which case the factors in Eq. (22) cancel and P_{acc} is identically one. This is referred to as the “heat bath” algorithm because it effectively touches a thermal bath to the link variable being updated. While this approach works well for $SU(2)$, it appears not to be practical when the group manifold is more complicated.

The most common application of the Metropolis *et al.* [13] procedure is simply to form the trial element by multiplying the old link by a factor from a table. In this case it is

convenient to have this table contain the inverse of each of its elements. Then the forward and backward trial probabilities are equal and Eq. (22) reduces to

$$P_{acc} = \min \left[1, e^{-\beta(S' - S)} \right]. \quad (23)$$

This is the form conventionally used for lattice gauge simulations. Note that no knowledge of the detailed measure is needed; it is generated completely automatically.

Another potentially useful scheme for picking the trial elements was discussed in refs. [14–15]. This approach is a generalization of the overrelaxation ideas discussed in ref. [16]. Suppose there is some straightforward way to find a group element U_0 which approximately minimizes the action $S(U_0)$. Suppose further that U_0 is obtained with no explicit dependence on the current element U . Then a useful trial change for U is

$$U' = U_0 U^{-1} U_0. \quad (24)$$

In some sense, this element lies on the “opposite side” of U_0 from U . Note that this construction is symmetric under interchange of U' and U ; consequently, the ratio of trial probabilities drops out of the acceptance condition, just as with the standard procedure.

I will give three motivations for this choice of trial update. First, it produces a rather large change in U while, assuming the action is reasonably symmetric around U_0 , resulting in only a small change in the action. Thus one simultaneously obtains a high acceptance rate and a rapid flow through configuration space. Second, when considering a coupled system of many variables, the value of U_0 depends on neighbors which in previous steps have adjusted their values to accommodate U . If U is allowed to float, then these neighbors will tend to relax away and the optimum value for an updated element should lie beyond U_0 . Third, recent analyses by Adler [17] and Neuberger [18] indicate that overrelaxation can help reduce the increase of correlation times as a critical point is approached. Tests with this algorithm [19] indicate a possible savings on the order of a factor of three in computer time over the conventional Metropolis *et al.* [14,15,13] algorithm for pure $SU(3)$ gauge theory.

In the standard Metropolis *et al.* [13] approach mentioned earlier, the trial changes are made in an unbiased way. In the heat bath method one increases the acceptance by making trial changes with a carefully defined distribution. One might try something intermediate by combining a random change with a biased driving force. For example, consider updating some real degree of freedom A . For a trial change take

$$A' = A + \delta p + \delta^2 F(A). \quad (25)$$

Here p is an unbiased noise which, for the purposes of discussion, I take to be Gaussian

$$P(p) \propto e^{-\beta p^2/2}. \quad (26)$$

The driving force $F(A)$ is for the moment arbitrary, and δ^2 is a bookkeeping parameter which I will consider to be small. Solving Eq. (25) for p gives the trial probability

distribution

$$\begin{aligned} P_{T,A}(A') &\propto e^{-\beta p^2/2}, \\ p &= \frac{1}{\delta}(A' - A - \delta^2 F(A)). \end{aligned} \quad (27)$$

This can be inserted into the Metropolis *et al.* acceptance criterion, Eq. (22), to give

$$P_{acc} = \min \left[1, e^{-\beta(H' - H)} \right] \quad (28)$$

where I have defined

$$H(p, A) = S(A) + p^2/2 \quad (29)$$

and $H' = H(p', A')$. Here p' is the noise which would be required to move back from A' to A

$$\begin{aligned} p' &= \frac{1}{\delta}(A - A' - \delta^2 F(A')) \\ &= -(p + \delta(F(A) + F(A'))) \end{aligned} \quad (30)$$

Expanding in powers of δ , I find

$$H' - H = \left(p\delta + \frac{1}{2} \left(p^2 \frac{\partial}{\partial A} + 2F(A) \right) \delta^2 \right) \left(\frac{\partial S(A)}{\partial A} + 2F(A) \right) + O(\delta^3). \quad (31)$$

This implies that the choice

$$F(A) = -\frac{1}{2} \frac{\partial S}{\partial A} \quad (32)$$

would give

$$P_{acc} = 1 + O(\delta^3). \quad (33)$$

Thus for small changes one can indeed obtain a high acceptance.

If the changes are small enough, one might consider introducing a small error by ignoring the acceptance criterion altogether and simultaneously updating all fields with Eq. (25) using the driving force in Eq. (32). At this point it is instructive to change notation and write

$$\begin{aligned} \frac{A' - A}{\delta^2} &= \frac{dA}{d\tau} \\ \frac{1}{\delta} p &= \eta. \end{aligned} \quad (34)$$

Then the updating procedure reads

$$\frac{dA}{d\tau} = -\frac{1}{2} \frac{\partial S}{\partial A} + \eta \quad (35)$$

which is the standard Langevin equation.

Note that the combination Eq. (25) and Eq. (30) with the driving force in Eq. (33) and followed by an inversion of p to $-p$ represents a discretization of a microcanonical evolution under a dynamics generated by the Hamiltonian in Eq. (29). Iterating this procedure equations gives the microcanonical algorithm proposed for lattice gauge theory by Callaway and Rahman [20]. Adding an occasional refreshing of the momenta into new

gaussian random numbers gives the hybrid algorithm of Duane and Kogut [21]. I will return in the next lecture to this analysis of biased trial changes in the context of updating fermionic fields.

LECTURE III: FERMION ALGORITHMS

Fermionic fields provide several challenging problems for lattice gauge theory. One appears already at the level of formulating an appropriate action. Here one encounters the notorious doubling of species appearing in the simplest schemes incorporating chiral symmetry. In these lectures I ignore this issue and assume that one has an acceptable lattice transcription of the Dirac equation.

The problem I address here concerns the difficulties in simulating a coupled fermion-boson system. For this purpose I introduce the generic partition function

$$Z = \int (dA) (d\psi) (d\psi^*) \exp(-S_G(A) - \psi^* M(A) \psi). \quad (36)$$

As I will be concentrating on fermionic details, I will ignore the technicality that the gauge fields are group elements and write them as A . To further simplify the notation, I have absorbed the coupling parameter β into the action. The matrix $M(A)$ contains both the kinetic terms for the fermionic fields as well as the couplings to the gauge field.

The essence of the problem lies in the fact that the fermionic fields are not ordinary numbers, and therefore the exponentiated action cannot be regarded as a probability. Instead, the fields ψ and ψ^* are anticommuting variables, integration over which I now briefly review.

I begin by considering a set $\{\psi_i\}$ of anticommuting Grassmann variables

$$[\psi_i, \psi_j]_+ = \psi_i \psi_j + \psi_j \psi_i = 0. \quad (37)$$

Generalizing complex conjugation, I adopt the convention that corresponding to each ψ_i I have another independent Grassmann variable ψ_i^* . Furthermore, I postulate

$$\begin{aligned} (\psi_i^*)^* &= \psi_i \\ (\psi_1 \dots \psi_n)^* &= \psi_n^* \dots \psi_1^*. \end{aligned} \quad (38)$$

Considering just a single variable ψ , a general function $f(\psi)$ can be expanded with just two terms

$$f(\psi) = f_0 + \psi f_1. \quad (39)$$

To define integration over an anticommuting variable, I demand the properties of linearity and invariance under translation. These are summarized in the axioms

$$\int d\psi (f(\psi) \alpha + g(\psi) \beta) = \left(\int d\psi f(\psi) \right) \alpha + \left(\int d\psi g(\psi) \right) \beta \quad (40)$$

$$\int d\psi f(\psi) = \int d\psi f(\psi + \psi'). \quad (41)$$

This is sufficient to imply that for the function in Eq. (39)

$$\int d\psi f(\psi) = K f_1 \quad (42)$$

where the normalization K is undetermined. I adopt the convention $K = i$ so that

$$\begin{aligned} \int d\psi \psi &= i \\ \int d\psi 1 &= 0 \\ \int d\psi^* d\psi \psi^* \psi &= 1. \end{aligned} \quad (43)$$

Note that under multiplicative rescaling a Grassmann integral behaves as

$$\int d\psi f(\psi a) = \left(\int d\psi f(\psi) \right) a. \quad (44)$$

This can be written in the heuristic form $d(\psi a) = (d\psi) / a$.

For integration over several anticommuting variables, I have

$$\int d\psi_1 \dots d\psi_n \psi_1 \dots \psi_n = i^n (-1)^{n(n-1)/2}. \quad (45)$$

The analog of Eq. (44) in this case is

$$\int (d\psi) f(M\psi) = |M| \int d\psi f(\psi) \quad (46)$$

where M is an arbitrary invertible matrix, $|M|$ is its determinant, and $(d\psi)$ denotes $d\psi_1 \dots d\psi_n$. Note that Eq. (46) immediately implies the Matthews-Salam [17] formula for a fermionic Gaussian integral

$$\int (d\psi d\psi^*) e^{-\psi^* M \psi} = |M| \quad (47)$$

where $(d\psi d\psi^*) = d\psi_1 d\psi_1^* \dots d\psi_n d\psi_n^*$.

Eq. (47) provides an easy way out of the difficulty that our partition function is not an ordinary integral. Indeed, I can explicitly integrate out the fermions to convert Eq. (36) to

$$Z = \int (dA) |M| e^{-S_G}. \quad (48)$$

This is now an integral over numbers and therefore in principle amenable to Monte Carlo attack. For the remainder of these lectures I will assume that the fermions have been formulated such that M is a positive matrix and thus the integrand in Eq. (48) can be regarded as proportional to a probability measure. If this is not so, one can always double

the number of fermionic species, using M^\dagger for the extra ones, thus replacing $|M|$ by $|M||M^\dagger|$.

Direct Monte Carlo attack of the partition function in Eq. (48) is still not practical because of the large size of the matrix M . In our compact notation, this is a square matrix of dimension equal to the number of lattice sites times the number of Dirac components times the number of internal symmetry degrees of freedom. Thus, it is typically a tens of thousands by tens of thousands matrix, precluding any direct attempt to calculate its determinant. The matrix is, however, generally extremely sparse because popular fermion actions do not directly couple distant sites. All the Monte Carlo algorithms used in practice for simulation of this problem make essential use of this fact.

A simple but time consuming algorithm with the usual statistical but no systematic errors was presented by Weingarten and Petcher [22]. For simplicity in the following discussion let me assume that the matrix M is real. Weingarten and Petcher observe that by introducing a new set of real scalar fields ϕ one can rewrite Eq. (48) in the form

$$Z \propto \int (dA) (d\phi) \exp \left(-S_G - (M^{-1}\phi)^2 / 2 \right). \quad (49)$$

Thus, a successful fermionic simulation would be possible if one could obtain configurations of fields ϕ and A with probability distribution

$$P(A, \phi) \propto \exp \left(-S_G - (M^{-1}\phi)^2 / 2 \right). \quad (50)$$

Ref. 22 notes that while M^{-1} is the inverse of an enormous matrix, to calculate the effective action one really only needs M^{-1} applied to the single vector ϕ . Indeed, there exist reasonably efficient iterative schemes for finding the inverse of a large matrix applied to a vector. This amounts to solving a large system of linear equations. Thus Ref. 22 proposed to directly simulate the partition function in Eq. (49) using a Gauss-Seidel algorithm to calculate $M^{-1}\phi$. It now appears that the conjugate gradient algorithm may be somewhat preferable for this inversion.

In practice, at least when the correlation length is not large, the conjugate gradient inversion adequately converges in a number of iterations which does not grow with the lattice size. As each step involves a sum over the vector, which has length proportional to the lattice volume, this means that the conjugate gradient step takes a time which grows proportional to the volume of the system. Unfortunately this inversion must be repeated for the updating of each of the system variables. Thus the algorithm of Ref. 22 is expected to require computer time which grows as the square of the volume of the lattice. Such a severe growth has precluded use of this algorithm on any but the smallest lattices. Nevertheless, it does show the existence of an exact algorithm with less computational complexity than would be required for a repeated direct evaluation of the determinant of the fermionic matrix.

Actually it is only the A fields which are difficult to update in such simulation. Ref. [23] presented an efficient scheme for updating the field ϕ while holding A fixed. First,

generate a random real vector χ with a Gaussian probability distribution

$$P(\chi) \propto e^{-\chi^2/2}. \quad (51)$$

Then construct

$$\phi = M\chi. \quad (52)$$

This change of variables gives ϕ with the desired probability. The Jacobian of the transformation is irrelevant because A is temporarily being held fixed. Despite this trick for ϕ , updating the A field involves repeatedly changing the matrix M , thus requiring repeated inversions.

To circumvent the severe computer requirements of the Weingarten-Petcher algorithm, most of the fermionic algorithms used in practice make an additional extrapolation in a step size parameter. Here I include the original pseudofermionic [24] technique as well as the microcanonical [25], Langevin [26] and hybrid [27] approaches, which involve discretization of a differential evolution. All of the above approaches except the pseudofermion method utilize a linearization of the effective action in the exponent of Eq. (50). In particular, from this action

$$S = S_0 + (M^{-1}\phi)^2 / 2 \quad (53)$$

I obtain its derivative with respect to A

$$\frac{\partial S}{\partial A} = \frac{\partial S_0}{\partial A} - \left((MM^\dagger)^{-1} \phi, \frac{\partial M}{\partial A} M^{-1} \phi \right). \quad (54)$$

As the Metropolis *et al.* [23,13] acceptance condition only needs changes in the action, to lowest order in the step size these can be obtained from this derivative. Note that this requires knowledge of $(MM^\dagger)^{-1} \phi$ in addition to $M^{-1}\phi$. This is not a major complication; indeed, when M is not Hermitian the standard inversion algorithms calculate this quantity anyway. The approximate fermion algorithms use this linearization to calculate the action changes, and perform the required conjugate gradient inversions to update the derivative only once per sweep of all gauge variables. Thus they make an error of higher order in the size of the variable changes. A complete calculation would then require an extrapolation of physical results to zero step size.

The pseudofermion algorithm [24] differs from the others in that it uses an ensemble of auxiliary fields ϕ . The corresponding finite step errors are different, but still there. Ref. [28] presents a technique which interpolates between the Langevin and pseudofermion methods.

Although the systematic effects associated with this extrapolation are occasionally studied, the severe computational demands constrain the abilities to make as many checks as might be desired. While these difficulties will lessen with improved computational facilities, an efficient scheme for fermionic updating without this additional approximation would be welcome. I discuss one particularly promising scheme for the remainder of this lecture.

The difficulty with calculating the quantity $M^{-1}(A)\phi$ appearing in the exact action strongly encourages algorithms which perform this computation as rarely as possible. Indeed, one of the prime motivations for the approximate algorithms is to do this inversion only once per sweep of the lattice variables. There are, however, exact algorithms which also perform this inversion only once per sweep. In particular, one can perform a Metropolis *et al.* type accept/reject step to restore exact detailed balance only after making trial changes over the whole system. This forms the basis for the algorithms discussed in references [29–34] and [3].

The concept of applying an acceptance condition to a global change has the danger that one will have large increases in the action if many variables are changed and thus final acceptances may be unfeasibly small. Indeed, an arbitrary change in all variables will increase the action by an amount proportional to the system volume, leading to an exponential suppression of the acceptance with increasing system size. To counteract this, one must reduce the step size as the volume increases. By appropriately biasing the trial changes, one may be able to keep the acceptance at a reasonable level for step sizes which still allow a practical rate of exploration of new configurations. I now discuss this in more detail.

To begin, recall from the last lecture a possible trial change of a single variable A to

$$A' = A + p\delta + F(A)\delta^2. \quad (55)$$

Here δ is an adjustable step size parameter introduced for bookkeeping purposes. The “momentum” variable p represents a random noise, which for convenience I take as before to be Gaussianly distributed

$$P(p) \propto e^{-p^2/2}. \quad (56)$$

The function $F(A)$ represents a driving force or bias in the trial selection procedure and is for the moment arbitrary.

The Metropolis *et al.* [13] scheme accepts trial changes with a conditional probability chosen to maintain detailed balance when applied to an equilibrium ensemble. With an unbiased trial change this acceptance is determined entirely by the exponentiated action change. Here, however, the force term in the selection procedure must be corrected for in the acceptance condition. As discussed in the last lecture, I can fully restore detailed balance by accepting the new value A' with probability

$$P_{\text{acc}} = \min \left[1, e^{\{H(p,A) - H(p',A')\}} \right]. \quad (57)$$

Here H is a classical “Hamiltonian”

$$H(p, A) = p^2/2 + S(A). \quad (58)$$

In Eq. (57) I introduce p' as the reverse noise, *i.e.* the noise which would be required for the selection of A as the trial had A' been the initial value

$$p' = -p - (F(A) + F(A'))\delta. \quad (59)$$

After the accept/reject step, the momenta should be refreshed; thus, they should be modified in a Monte Carlo or other fashion that preserves the distribution in Eq. (56). Note that H is precisely the Hamiltonian used in the microcanonical algorithm [25] to describe evolution in "simulation time." Because of this analogy, I refer to H as the classical energy.

Had I set $F = -\frac{1}{2} \frac{\partial S}{\partial A}$, then Eq. (55) and Eq. (59) considered together with the inversion $p^i \rightarrow -p^i$ would represent a discretization of a microcanonical step of A and p under the Hamiltonian in Eq. (58). The microcanonical time step is δ . While a true microcanonical evolution conserves energy, the above discretization violates this slightly and gives less than unit acceptance in Eq. (57). I will shortly return to this choice for the driving force.

Note that with the second order terms in δ , the mapping defined by Eq. (55) and Eq. (59) exactly preserves areas in phase space

$$dA dp = dA' dp'. \quad (60)$$

Were this not so, the acceptance criterion would also need to depend on a ratio of measures.

The overall algorithm drives p and A towards equilibrium with the coupled probability

$$P(A, p) \propto e^{-S(A) - p^2/2}. \quad (61)$$

Some rather useful identities follow from expectation values over this distribution. Consider the partition function

$$Z = \int dA' dp' e^{-H'} = \int dA dp e^{-H} e^{H-H'} \quad (62)$$

where H and H' denote $H(p, A)$ and $H(p', A')$, respectively. Dividing by Z , I find

$$\langle e^{H-H'} \rangle = 1 \quad (63)$$

where the expectation value is over initial p and A distributed as in Eq. (61). By Jensen's inequality this immediately implies

$$\langle H' - H \rangle \geq 0. \quad (64)$$

with equality only possible if the algorithm exactly conserves energy. Below I will be considering small changes in the energy, in which case the expansion of Eq. (63) to second order

$$\langle H' - H \rangle = \frac{1}{2} \langle (H' - H)^2 \rangle + \mathcal{O} \left(\langle (H' - H)^3 \rangle \right) \quad (65)$$

will prove to be very useful.

Note that the above results are valid for any driving force whatsoever. In particular, there does not exist an area preserving mapping of phase space which will decrease the average value of the energy.

To proceed I use the parameter δ for an expansion of the energy change. In the last lecture I argued that

$$H' - H = \left(p\delta + \frac{1}{2} \left(p^2 \frac{\partial}{\partial A} + 2F(A) \right) \delta^2 \right) \left(\frac{\partial S(A)}{\partial A} + 2F(A) \right) + \mathcal{O}(\delta^3). \quad (66)$$

This implies that the choice

$$F_L(A) = -\frac{1}{2} \frac{\partial S}{\partial A} \quad (67)$$

leads to an energy change

$$H' - H = O(\delta^3) \quad (68)$$

Indeed, making this choice and ignoring the possibility of rejecting the trial change gives the usual Langevin algorithm [23,26], where the parameter δ is the square root of the step size used for discretization.

Let me first consider not making the Langevin choice for the driving force. Inserting Eq. (66) into Eq. (65) gives

$$\langle H' - H \rangle_{A,p} = \frac{\delta^2}{2} \left\langle \left(\frac{\partial S(A)}{\partial A} + 2F(A) \right)^2 \right\rangle + O(\delta^4). \quad (69)$$

Note that terms with odd powers of δ in the energy change expansion all involve odd powers of p and thus vanish on the average. If I now consider updating some large number V of variables together, the positive $O(\delta^2)$ quantities will coherently add and I expect to find a total energy change increasing linearly with V . By the central limit theorem, the fluctuations about this growth will become gaussian. Thus for large volumes I expect to find

$$H' - H \simeq C\delta^2 V + B\rho\delta V^{1/2} \quad (70)$$

where C and B are constants and ρ is a gaussian random variable which I normalize such that its probability distribution is

$$P(\rho) \sim e^{-\rho^2/2}. \quad (71)$$

If Eq. (70) were exact, then Eq. (63) would relate C and B

$$C = B^2/2. \quad (72)$$

With this explicit form for the energy change, I can obtain the expected acceptance in the large V limit

$$\langle P_{\text{acc}} \rangle = \langle \min [1, e^{H-H'}] \rangle = \frac{2}{\sqrt{\pi C V \delta^2}} e^{-C V \delta^2/4} \times \left(1 + O\left(\frac{1}{C V \delta^2}\right) \right). \quad (73)$$

The calculations required to derive Eq. (72) and Eq. (73), however, depend strongly on the tails of the distribution of the energy change and thus cannot be regarded as completely rigorous. For the following I will only assume that the expected acceptance is exponentially suppressed when $V\delta^2$ is large.

To avoid this exponential suppression and have a reasonable acceptance requires $\delta \sim V^{-1/2}$. However a small value for the step size raises the issue that the lattice will evolve only slowly from its original configuration. More precisely, consider taking N sweeps over

the lattice. As the motion of A has both random and driven terms, the overall change in any given variable should go as

$$\Delta A = O(\delta\sqrt{N}) + O(\delta^2 N) = O(\sqrt{N/V}) + O(N/V). \quad (74)$$

The final result is that the number of sweeps required to obtain a substantially new configuration should grow as V . If V is proportional to the system volume, then the overall algorithm requires time growing as volume squared, one factor of volume from the number of sweeps, and the other from the fact that each sweep takes time proportional to the volume.

For a bosonic simulations this growth would be a disaster. The standard algorithms only grow as the system volume, and thus should be preferred over updating many variables simultaneously. However, for exact fermionic algorithms one already is starting with a volume squared behavior, and thus there is no obvious penalty in going to global updates. Indeed, it might be possible to gain something by a judicious choice of F which will reduce the coefficient of this growth [2].

I now return to the Langevin choice of Eq. (67) for the driving force. To begin, consider again updating only a single variable. At first glance one might think that since the exact action is so difficult to calculate, the requisite derivative for this force would be intractable. Note, however, that this derivative is just what was needed in Eq. (54) to linearize the action for the approximate algorithms. Considering the action in Eq. (53), the Langevin force takes the form

$$F_L(A) = -\frac{1}{2} \frac{\partial S_0}{\partial A} + \frac{1}{2} \left((MM^\dagger)^{-1} \phi, \frac{\partial M}{\partial A} M^{-1} \phi \right) \quad (75)$$

As discussed below Eq. (54), calculation of this force is not significantly more difficult than finding the action.

To proceed I slightly generalize this force and take

$$F(A) = -\frac{1}{2} \frac{\partial S}{\partial A} + g(A) \delta^2. \quad (76)$$

The $g\delta^2$ piece is included for the purpose of discussing possible higher order improvements. Using this, I calculate the next term in the expansion for the energy change

$$H' - H = -\frac{\delta^3}{12} (S_3 p^3 - 3S_1 S_2 p - 24g p) + O(\delta^4). \quad (77)$$

Here I use the notation

$$S_n = \frac{\partial^n S}{\partial A^n}. \quad (78)$$

Note that if S_3 is non-vanishing, *i.e.* if the theory is not harmonic, then no choice of $g(A)$ can make the $O(\delta^3)$ term in this equation vanish for all p . Thus for any driving force whatsoever, a Metropolis *et al.* [13] correction to the Langevin evolution will always reject some changes if the step size is finite.

I wish to consider applying this procedure to a group of variables simultaneously. In particular, as in the earlier discussion of unbiased changes, I am interested in any coherent addition of action changes which could give an exponential suppression of the final acceptance. Because the $O(\delta^3)$ term in Eq. (77) contains only odd powers of p , it will vanish on the average. The expectation for the energy change can again be most easily found using Eq. (65). With a little algebra and explicitly doing the average over p , I find

$$\langle H' - H \rangle = \frac{\delta^6}{96} \langle 2S_3^2 + 3(8g - S_3 + S_1 S_2)^2 \rangle + O(\delta^8). \quad (79)$$

This can be written in many forms; this expression as the sum of two squares emphasizes positivity.

For non-harmonic interactions the quantity in Eq. (79) is positive for any g . Thus one cannot use the freedom in redefining the force to push the expected energy change to higher order. This result does, however, suggest that an improved driving force is

$$F(A) = -\frac{1}{2} \frac{\partial S}{\partial A} + (S_3 - S_1 S_2) \delta^2 / 8. \quad (80)$$

Note that this choice also eliminates the harmonic pieces in Eq. (77). Difficulties with calculating the higher derivatives of the action may preclude the utility of this form.

I now return to updating a large number V of independent variables simultaneously. The positive contributions indicated in Eq. (79) will add coherently. Similar arguments to those leading to Eq. (73) now give an expected acceptance falling as

$$P_{\text{acc}} \sim e^{-CV\delta^6} \quad (81)$$

To have a reasonable acceptance requires only $\delta \sim V^{-1/6}$. This changes Eq. (71) to

$$\Delta A = O(\delta\sqrt{N}) + O(\delta^2 N) = O\left(\sqrt{N/V^{1/3}}\right) + O(N/V^{1/3}). \quad (82)$$

Thus, the number of sweeps for an independent lattice grows as $V^{1/3}$ and the overall computer time for decorrelation increases as

$$T \sim V^{4/3}, \quad (83)$$

This behavior is only slightly worse than the linear growth of the pure bosonic theory.

This algorithm was proposed in Ref. 29 and tested further with somewhat discouraging results in Ref. 30. Ref. 31 presents a quite promising variation, which I now discuss. Recapitulating on the above treatment of biased updates, I constructed both the trial new A and the noise needed to return

$$A' = A + p\delta + F(A) \delta^2 \quad (84a)$$

$$p' = -p - (F(A) + F(A')) \delta. \quad (84b)$$

This is an area preserving map of the (A, p) plane onto itself. The scheme proposed in Ref. 31 is to iterate the combination of this mapping with an inversion $p' \rightarrow -p'$ several times before making the accept- reject decision. This iterated map remains reversible and area preserving. The second order terms in this equation make it equivalent to the leap frog procedure with an initial half step as used in Ref. 31.

The important point is that after each step the momentum remains exactly the negative of that which would be required to reverse the entire trajectory and return to the initial variables. Thus a final acceptance with the probability of Eq. (57) still makes the overall procedure exact. In this way the hybrid algorithm of Ref. 27 is made exact, just as the procedure with a single step removes the systematic errors of Langevin evolution. After each accept-reject step, the momenta p are refreshed, their values being replaced by new Gaussian random numbers. The fields ϕ could also be refreshed at this time, or less often, as turns out to be appropriate. The goal of the procedure is to use the microcanonical evolution as a way to restrict changes in the action so that the final acceptance will remain high for reasonable step sizes.

This procedure contains several parameters which can be adjusted for optimization. First is N_{mic} , the number of microcanonical iterations taken before the global accept/reject step and refreshing of the momenta p . Then there is the step size δ , which presumably should be set to give a reasonable acceptance. Finally, one can also vary the frequency with which the auxiliary scalar fields ϕ are updated.

The arguments for following a microcanonical trajectory for some distance before refreshing the momenta have been stressed in Ref. [27]. Refs. [33] and [3] show that this approach gives an algorithm where the computer time grows as $V^{5/4}$. I now review that argument.

The goal of the approach is to speed flow through phase space by replacing a random walk of the A field with a continued motion in the direction of p . As long as the total microcanonical time for a trajectory is smaller than some characteristic time for the system, the net change in A will grow linearly with both N_{mic} and δ ; thus Eq. (82) is replaced by

$$\Delta A \sim N_{mic} \delta. \quad (85)$$

which should be valid as long as

$$N_{mic} \delta < O(1). \quad (86)$$

With large N_{mic} , the change in the classical energy will also grow. In any given microcanonical step the energy changes by an amount of order δ^3 . For N_{mic} of order δ^{-1} , the total energy change will then be of order δ^2 . Because the evolution preserves areas in phase space, Eq. (65) still applies to the overall evolution and I have for the expected energy change

$$\langle H' - H \rangle = \frac{1}{2} \langle (H' - H)^2 \rangle + O \left(\langle (H' - H)^3 \rangle \right) = O(\delta^4). \quad (87)$$

If I update V independent variables together, these positive contributions can coherently add and earlier arguments now give an overall acceptance falling as

$$P_{\text{acc}} \sim e^{-CV\delta^4} \quad (88)$$

This means that δ should be taken to decrease with volume as $V^{-1/4}$. Correspondingly, N_{mic} should grow as $V^{1/4}$, the maximum allowed by Eq. (86). The final result is that the total time required to obtain a substantially changed lattice grows as

$$T \sim V^{5/4} \quad (89)$$

This may be only an asymptotic statement, valid for systems much larger than the correlation length. The main uncertainty lies in the unknown characteristic time scales that determine the $O(1)$ right hand side of Eq. (86). Nevertheless, the above growth is sufficiently slow that it compels further testing.

To summarize, I hope that these lectures along with those of the other speakers at this meeting will convince you that the lattice approach is a useful and interesting technique to investigate non-perturbative phenomena in field theory. As regards algorithms for numerical simulations, although conventional approaches work quite well for purely bosonic systems, overrelaxation methods may improve things by a factor of a few. With fermionic fields, algorithm development is very much an active field, with globally corrected hybrid algorithms appearing quite promising.

A difficult remaining question which I have not discussed is the simulation of fermionic systems when the corresponding determinant is not always positive. This situation is of considerable interest because it arises in the study of quark-gluon thermodynamics when a chemical potential is present. All known approaches to this problem are extremely demanding on computer resources. This is a particularly promising area for future algorithm development.

References

1. M. Creutz, *Quarks, Gluons and Lattices*, (Cambridge Univ. Press 1983).
2. M. Creutz, in *Lattice Gauge Theory Using Parallel Processors*, Xiaoyuan Li, Zhaoming Qiu, and Hai-Cang Ren, eds., p. 105 (Gordon and Breach, 1987). This book contains several useful reviews of various aspects of lattice gauge theory..
3. M. Creutz, *Phys. Rev.*D38, 1228 (1988).
4. K. Wilson, *Phys. Rev.*D10, 2445 (1974).
5. C. N. Yang and R. Mills, *Phys. Rev.*96, 191 (1954).
6. D. Gross and F. Wilczek, *Phys. Rev.* D8, 3633 (1973); H. D. Politzer, *Phys. Rev. Lett.*30, 1343 (1973).
7. S. Elitzur, *Phys. Rev.* D12, 3978.

8. L. Faddeev and V. Popov, Phys. Lett. 25B, 29 (1967).
9. M. Creutz, J. Math. Phys. 19, 2043 (1978).
10. J. Kogut and L. Susskind, Phys. Rev. D11, 395 (1975).
11. M. Creutz, Phys. Rev. D15, 1128 (1977).
12. J. Kogut, Rev. Mod. Phys. 51, 659 (1979).
13. N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, J. Chem. Phys. 21, 1087 (1953).
14. F. R. Brown and T. J. Woch, Phys. Rev. Lett. 58, 2394 (1987).
15. M. Creutz, Phys. Rev. D36, 515 (1987).
16. S. L. Adler, Phys. Rev. D23, 2901 (1981); C. Whitmer, Phys. Rev. D29, 306 (1984); Princeton University PhD thesis (1984).
17. S. L. Adler, Institute for Advanced Study preprint IASSNA-HEP-87/27 (1987).
18. H. Neuberger, Rutgers University preprint RU-35-87 (1987).
19. R. Gupta, G. Kilcup, A. Patel, S. Sharpe, and P. De Forcrand, Preprint LA-UR-88-824 (1988).
20. D. Callaway and A. Rahman, Phys. Rev. D28, 1506 (1983).
21. S. Duane and J. Kogut, Phys. Rev. Lett. 55, 2774 (1985); Nucl. Phys. B275, 398 (1986).
22. D. Weingarten and D. Petcher, Phys. Lett. 99B, 333 (1981).
23. G.G. Batrouni, G.F. Katz, A.S. Kronfeld, G.P. Lepage, B. Svetitsky and K.G. Wilson, Phys. Rev. D32, 2736 (1985).
24. F. Fucito, E. Marinari, G. Parisi and C. Rebbi, Nucl. Phys. B180[FS2], 369 (1981).
25. J. Polonyi and H.W. Wyld, Phys. Rev. Lett. 51, 2257 (1983);
J. Kogut, J. Polonyi, H.W. Wyld and D.K. Sinclair, Phys. Rev. Lett. 54, 1 475 (1985).
26. A. Ukawa and M. Fukugita, Phys. Rev. Lett. 55, 1854 (1985).
27. S. Duane and J. Kogut, Phys. Rev. Lett. 55, 2774 (1985);
Nucl. Phys. B275, 398 (1986).
28. M. Creutz and R. Gavaì, Nucl. Phys. B280 (1986).
29. R. T. Scalettar, D.J. Scalapino and R.L. Sugar, Phys. Rev. B34, 7911 (1986).
30. S. Gottlieb, W. Liu, D. Toussaint and R.L. Sugar, Phys. Rev. D35, 2611 (1986).
31. S. Duane, A. D. Kennedy, B.J. Pendleton and D. Roweth, Phys. Lett., 195, 2 (1987).
32. H. Gausterer and S. Sanielevici, Phys. Rev. D38, 1220 (1988).
33. R. Gupta, G. Kilcup, and S. Sharpe, Phys. Rev. D38, 1278 (1988).
34. K. Bitar, R. Horsley, A. Kennedy, S. Meyer and P. Rossi, preprint FSU-SCRI-88-14 (1988).