BROOKHAVEN NATIONAL LABORATORY

June, 1989 BNL-42968

Lattice Fields and Strong Interactions

Michael Creutz

Physics Department
Brookhaven National Laboratory
Upton, NY 11973

ABSTRACT

I review the lattice formulation of gauge theories and the use of numerical methods to investigate nonperturbative phenomena. These methods are directly applicable to studying hadronic matter at high temperatures. Considerable recent progress has been made in numerical algorithms for including dynamical fermions in such calculations. Dealing with a non-vanishing baryon density adds new unsolved challenges.

Lectures at the NATO International Advanced Course on the Nuclear Equation of State, Peñiscola, Spain, 21 May - 3 June, 1989.

I. INTRODUCTION

The theme of this conference concerns large scale effects of strong interaction physics. The dynamics underlying the nuclear equation of state is generally believed to be of quarks interacting via the exchange of gauge gluons. In this picture the effective couplings are large, particularly at high temperatures or densities. Among particle theorists, the lattice approach has become the primary tool for the study of the nonperturbative aspects of this theory. In principle, this technique can provide information on the structure of hot dense hadronic matter. Indeed, in he case where the average baryon density vanishes, lattice gauge simulations have provided the most accurate information available on the temperature of the transition to a quark gluon plasma.

Historically quantum field theory has had its greatest successes with the perturbative calculations of electrodynamics. The renormalization program successfully removed the divergences encountered on the way. For strong interaction phenomena, however, we do not have the benefit of a small parameter in which to expand. Thus we need a formalism which controls divergences but does not rely on the standard Feynman diagrammatic expansion. This is the underlying purpose of the lattice formulation of gauge field theories. The lattice spacing represents an ultraviolet cutoff, which must be removed by a limiting procedure before physical results are inferred.

Once the lattice cutoff is in place, a gauge field theory is a well defined mathematical system quite suited for numerical study. The theory has a precise mathematical analogy with classical statistical mechanics, for which Monte Carlo simulation methods have become an important tool. The last decade has seen many productive applications of these techniques to gauge theories.

The most successful results of lattice gauge theory have concerned phenomena involving gluons alone or at most a small number of quarks. For the nuclear equation of state, however, one is interested in a large background baryonic density. This raises some interesting new challenges, although lattice methods should still provide a powerful tool.

The problems with dynamical quarks occur at several levels. The first is to find a simple lattice transcription of the Dirac equation which behaves properly in the continuum limit. Subtleties with chiral symmetry and anomalies make this not totally trivial, and

several schemes have been formulated. Next is the problem of numerical simulation with variables which anticommute. Considerable progress has been made in the last few years on algorithms to treat the situation with zero net baryon density. Finally there is the full problem of hadronic matter at high density. In this case numerical methods can in principle be applied to small systems, but the computer demands are high and appear to grow exponentially with the system size. To improve this situation is an area where algorithm development is badly needed and will hopefully come with time.

In these lectures I will first review the general lattice gauge formalism and the primary results for pure gauge fields. I will then turn to the difficulties with fermions and summarize the recent advances. Finally I will comment on why the problem with a finite baryon density is so hard. I hope that some of you will find clues to circumventing these difficulties.

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].

II. THE LATTICE FORMULATION

I begin with a mathematical description of Wilson's [4] theory, and defer the physical motivation until later. The theory is usually formulated on a four dimensional hypercubic lattice. For definiteness, assume that the lattice has N^4 sites and has periodic boundary conditions. The basic gauge variables are group elements located on the sites of the lattice. Thus, for any ordered nearest neighbor pair of sites (i, j), I consider a bond variable U_{ij} which is an element of the gauge group G. Here i and j denote the sites labelling the ends of the link, while the group matrix indices are understood. 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 $SU(3)$, i.e. the group of three by three unitary matrices with unit determinant. 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

$$U_{ji} = U_{ij}^{-1}. \tag{1}$$

To obtain a dynamics for these variables, I use the Wilson [4] action

$$S = \sum_p (1 - 1/n \mathrm{Re}\,\mathrm{Tr}\, U_p)\,. \tag{2}$$

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 a group element obtained as 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.

To quantize the theory, I exponentiate this action and study the path integral over all gauge fields. This reveals a mathematical equivalence with the statistical mechanics of this system of group elements. In this analogy, the exponentiated action plays the role of a 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

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

The parameter β is proportional to the inverse square of the bare gauge coupling g_0. For the gauge group $SU(3)$ the relation is

$$\beta = 6/g_0^2. \tag{4}$$

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

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

In these relations, the integration over group elements is to be taken using the group invariant measure. For compact groups this measure is unique and satisfies

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

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

$$\int dU\ 1 = 1. \tag{7}$$

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

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. Their 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 later.

This completes the formal definition of lattice gauge theory. The motivations 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

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

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

$$S = \int \frac{1}{4} F_{\mu\nu} F_{\mu\nu}. \tag{9}$$

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 integrand becomes real and the analogy with a statistical mechanics problem is explicit. Indeed, this is the motivation for the notation β for the coefficient of the action in Eq. (3). Note that β is not directly related to any physical temperature in the field theory. I will shortly discuss how the real temperature can be adjusted by making the lattice finite in the temporal direction.

For continuum physics, we must take the lattice spacing to zero. The renormalization group gives important information on this limit. As the lattice spacing is reduced, one encounters 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

$$a\frac{\partial g_0}{\partial a} = \gamma\left(g_0\right) = \gamma_0 g_0^3 + \gamma_1 g_0^5 + \ldots \tag{10}$$

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

$$a = e^{-1/2\gamma_0 g_0^2}\ \frac{1}{\Lambda}\left(g_0^2\gamma_0\right)^{-\gamma_1/2\gamma_0^2}\left(1+\mathcal{O}\left(g_0^2\right)\right). \tag{11}$$

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 dominantly 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 lattice gauge action 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

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

Since the dynamics involves the trace of link variables multiplied around closed loops, the factors of g cancel in the calculation of the action for the new links. This exact local symmetry is the gauge symmetry of the model.

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 of this expectation 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, one can consider a power series expansion in the small parameter β. The first nonvanishing term in such an expansion for a Wilson loop of area A occurs at order β^A. This factor of beta raised to the area of the loop is precisely the exponential suppression signaling confinement.

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 distinct 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 [7] can be obtained from the above Lagrangian formalism by first fixing the temporal gauge ($A_0 = 0$) and then taking a continuous time limit [8]. If we order the path integral into successive integrals over time

slices, the integration over any given time slice can be regarded as a sum over intermediate states in the quantum Hilbert space. This is the space of square integrable functions of the space-like links. In this approach, the spatial link variables are operators $\hat{U}_{ij}$ in the quantum space. Conjugate to them are conjugate "electric field" operators l^{α}_{ij}, which generate group rotations of the corresponding links. The commutation relations amongst these objects are

$$\begin{aligned} \left[l^{\alpha}_{ij}, l^{\beta}_{ij}\right] &= i f^{\alpha\beta\gamma} l^{\gamma}_{ij} \\ \left[l^{\alpha}_{ij}, \hat{U}_{ij}\right] &= -\lambda^{\alpha} \hat{U}_{ij}. \end{aligned} \tag{13}$$

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^{α}_{ij} can be represented by differential operators. In terms of these quantities the Hamiltonian takes the form

$$H = \frac{g^2}{2a} \sum_{\{ij\}} l^2_{ij} + \frac{2}{g^2 a} \sum_{p} \mathrm{Re}\,\mathrm{Tr}\,\hat{U}_p. \tag{14}$$

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 space-like plaquettes. Note that l^2 is just the quadratic Casimir operator on the respective link. The two terms in this equation represent the electric and magnetic contributions to the energy.

In the strong coupling limit of large g the electric field term dominates the Hamiltonian and one can do standard perturbation theory in the plaquette term in Eq. (14). 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 $l^2_{ij} = 0$. The potential or magnetic term then can excite links into intermediate states involving higher representations. For a review of this approach see ref. [9].

In the transfer matrix formalism with periodic boundary conditions in the time direction, the path integral of Eq. (3) becomes

$$Z = \mathrm{Tr} e^{-TH} \tag{15}$$

where T is the total temporal size of the lattice. For conventional hadronic physics, one wishes to take the infinite volume limit where T goes to infinity. In this case the leading

behavior is dominated by the lowest energy state, i.e. the vacuum. However, this equation makes it clear that if we do not take this limit, then we are merely studying the system at a finite physical temperature.

With the spatial extent of the lattice considered as finite, one can consider the special Wilson loops consisting of straight lines of links in the temporal direction closing due to the periodicity. These are referred to as "Wilson lines" or "Polyakov loops" [10]. As the temporal extent is increased, the expectation of such a line should decrease exponentially with the free energy of an isolated quark. With the pure gauge theory exhibiting confinement, this expectation will vanish. At higher physical temperatures, however, Ref. 10 has argued for the appearance of a phase transition to plasma phase with liberated quarks. This transition has been observed in numerous simulations. In the high temperature phase, the Wilson line acquires a non-vanishing expectation value. This observable forms a particularly clear order parameter for observing this transition.

III. NUMERICAL RESULTS

In the path integral formulation, lattice gauge theory is equivalent to a classical statistical mechanics problem. As such, it is directly amenable to numerical simulation using standard techniques. The general procedure begins by storing in a computer memory some initial values for the gauge variables on a finite lattice. These are updated by successive pseudo-random changes on the link variables. The changes are made in a biased manner such that the ultimate probability of obtaining any configuration C is proportional to the Boltzmann weight

$$P(C) \sim e^{-\beta S}. \tag{16}$$

As this Markov chain of configurations becomes longer, expectation values for the true theory are increasingly accurately approximated by averages within the chain.

The details of the Monte Carlo procedure are highly non-unique and many variations have been applied to lattice gauge simulations. For a review, see Ref. 2. The most intuitive approach is to loop over the variables and replace each in turn with a new value chosen randomly with a weighting proportional to the Boltzmann factor calculated with all neighboring elements temporarily held fixed. A procedure which is generally simpler to implement was devised some time ago by Metropolis *et al.* [11]. This scheme begins with

a random trial change to the system and then conditionally accepting this change with a probability constructed to ensure that the proper Boltzmann weighting is stable under the algorithm. For lattice gauge theory the usual implementation of the Metropolis *et al.* [11] algorithm considers the trial change for an element U to

$$U' = U\delta \tag{17}$$

where δ is a group element taken randomly from a table of elements biased near the identity and containing the inverse of each of its elements. This trial change is then accepted with probability

$$P_{acc} = \min\left[1, e^{-\beta(S'-S)}\right] \tag{18}$$

where S' is the action calculated using the trial variable. Variations on this scheme abound; indeed, later I will discuss some in the context of fermionic simulations.

Once the system is in an approximate thermal equilibrium, then one can measure any desired function of the fields. The first physical result to come from such numerical lattice gauge calculations was the relation between the scales of confinement and asymptotic freedom. For large separations R a quark and an antiquark will display a linear confining potential

$$V(R) \sim KR \tag{19}$$

where K is a numerical constant of dimension mass squared. On the other hand, at small separations the effective gauge coupling shows the logarithmic decrease characteristic of asymptotic freedom

$$g^2 \sim \frac{1}{\log\left(\frac{1}{R\Lambda}\right)} \tag{20}$$

where Λ is numerical scale of dimension mass. Monte Carlo simulation allows us to interpolate between the long and short distance regimes and relate the scales K and Λ. A recent compendium of the results [12] gives

$$\Lambda_{\overline{MS}} = (0.32 \pm 0.06)\sqrt{K} \tag{21}$$

Here $\Lambda_{\overline{MS}}$ denotes a particular convention for normalizing the asymptotic freedom scale. From hadronic phenomenology of either heavy quark bound states or Regge trajectories, K is believed to be approximately $(400\ \text{MeV})^2$. This gives $\Lambda_{\overline{MS}}$ of order 130 Mev, a number

not inconsistent with phenomenological values. Indeed, the error in this calculation is less than the current experimental uncertainty in Λ.

The second physical number to come from these simulations was the temperature of the transition to the quark gluon plasma. As mentioned above, when the physical temperature exceeds a critical value T_c, the vacuum undergoes a phase transition to a plasma in which a single quark has only finite energy. This temperature is obtained by monitoring the Wilson or Polyakov line, defined in the previous section. Ref. 12 has compiled these results to obtain

$$T_c/\Lambda_{\overline{MS}} \sim 1.6 - 1.7 \tag{22}$$

With the above value for Λ, this give a transition temperature of order 200 Mev, a rather reasonable number which may be attainable in relativistic heavy ion collisions.

Recently there has been some controversy about the order of this transition. While it is clear that there is a very rapid change in behaviour at the above temperature, some simulations [13] suggest that there is no actual discontinuity in expectation values, while others [14] indicate that the transition is first order. This issue has only limited experimental consequences because (1) these simulations are of the gluonic sector of the theory and do not include dynamical quark loops, and (2) it will be rather difficult in foreseeable experiments to distinguish the difference between a second order transition and one with a small nonvanishing latent heat.

Considerable effort has gone into extracting hadronic spectra from lattice simulations. In the valence approximation of neglecting dynamical quark loops, the technique begins by solving the Dirac equation on equilibrated lattices. The propagation of various quark and antiquark combinations is then studied, and masses for the bound states extracted. As the underlying dynamics is based on a quark model, and the naive quark model works well in many cases, it is expected that these calculations must do reasonably well. The most interesting result here is that a light pion comes out. Indeed, the chiral symmetry prediction of a pion mass going to zero with the square of its constituent quark masses appears to be born out, even in this approximation neglecting dynamical quark feedback on the gauge fields.

Some artifacts of the valence approximation do appear in these spectrum calculations. In particular, even after an extrapolation to the physical pion mass, the ratio of the rho

mass to the proton mass appears to be coming out somewhat too low [15]. It is hoped that this problem is only an effect of the valence approximation, but it may also be due to the lattice spacing still being too large.

In these spectrum calculations, one difficult point has been the masses of states containing no quarks, the so called "glueballs." What complicates the calculation of these states is the subtraction of large connected pieces to obtain the required correlation functions, and after these subtractions statistical errors remain quite large. Nevertheless, the results are consistent with glueball states of order 1.4 GeV, consistent with several experimental candidates [16].

One of the more promising areas for the near future is the study of hadronic matrix elements of operators relevant to weak decays. In particular, there now exists considerable evidence that strong interaction effects can indeed give rise to the substantial enhancement of $\Delta I = 1/2$ transitions observed in kaon decays [17]. The relevant matrix elements involve strong interaction corrections in an essential way, and the lattice approach is perhaps the most promising way to evaluate the required quantities. Furthermore, these numbers are sufficiently unknown and important that even relatively large statistical errors can be tolerated.

IV. DYNAMICAL FERMIONS

The results mentioned in the last section all involve the valence approximation of neglecting the effects of dynamical quark loops. While many results are now appearing with the quark effects included, they should still be regarded as preliminary. The deconfinement transition is the most studied quantity, and appears to persist in the presence of dynamical quarks. Most studies suggest that the transition is first order, but one should remember the subtleties of this transition in the pure gauge case.

There is reason to believe that progress with dynamical fermions will increase greatly in the near future. This is because of considerable recent algorithmic progress. It is these improvements that I will concentrate on in this section.

The problems with fermions begin at the level of the lattice formulation. If we simply put a Dirac field ψ on the sites there is a tendency to have extra flavors of fermion in the continuum limit. In the path integral ψ and $\bar{\psi}$ are independent anticommuting quantities.

In fermionic generalization of the transfer matrix formalism, these each correspond to a fermionic space, and we tend to have twice as many particles as intended. This doubling is associated with the temporal motion of the particles, and additional doublings come from the spatial motion. The extra species can be projected out by only allowing some components to propagate. One scheme for doing this was proposed by Wilson some time ago [18]. Whenever a fermion moves along a link in direction with unit vector e_μ, the corresponding wave function picks up a factor of the projection matrix $(1 \pm e_\mu \gamma_\mu)/2$. The doubling problem is deeply entwined with chiral symmetry [19]. Indeed, the Wilson prescription breaks chiral symmetry even when the bare fermion mass is zero.

Assuming that one has chosen a reasonable fermionic action, I now turn to the problems encountered for numerical simulations. For this discussion I will be quite generic and consider the partition function

$$Z = \int (dA)\,(d\psi)\,(d\bar{\psi})\ \exp\left(-S_0(A) - \bar{\psi} M(A)\,\psi\right). \tag{23}$$

where S_0 is the pure gauge part of the action.

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 essential difficulty is that the fermionic fields are not ordinary numbers, and therefore the exponentiated action cannot be regarded as a probability. They can, however, be formally eliminated from the problem using the Matthews-Salam [20] formula for a fermionic Gaussian integral

$$\int (d\psi d\bar{\psi})\ e^{-\bar{\psi} M \psi} = |M|. \tag{24}$$

Eq. (24) provides a way around of the difficulty that our partition function is not an ordinary integral. Indeed, I can explicitly integrate out the fermions to convert Eq. (23) to

$$Z = \int (dA)\ |M|\ e^{-S_0}. \tag{25}$$

This is now an integral over numbers and therefore in principle amenable to Monte Carlo attack.

To proceed it would be simplest if the determinant of M were a positive quantity, in which case the integrand in Eq. (25) could be used as a statistical weighting. For the Wilson action with no chemical potential $|M|$ is real. This can be seen conjugating the action and restoring the projection operators by a γ_5 rotation of ψ. The determinant is, however, not necessarily positive. It can be made so by an extra doubling of the number of species, using $M^\dagger$ for the extra ones and replacing $|M|$ by $|M||M^\dagger|$. As will be discussed later, this trick does not work when a net baryon density is present. For now, however, assume that $|M|$ is indeed positive.

Direct Monte Carlo attack of the partition function in Eq. (25) 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 make essential use of this fact.

To proceed I now replace the determinant of M by an integral over a set of auxiliary commuting fields, as suggested by Weingarten and Petcher [21]. 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. (25) in the form

$$Z \propto \int (dA)\,(d\phi) \exp\left(-S_0 - \left(M^{-1}\phi\right)^2/2\right). \tag{26}$$

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

$$P(A,\phi) \propto \exp\left(-S_0 - \left(M^{-1}\phi\right)^2/2\right). \tag{27}$$

Ref. 21 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, such as the conjugate gradient algorithm, for finding the inverse of a large matrix applied to a vector.

In practice, at least when the correlation length is not large, the conjugate gradient method 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 proportionally to the system volume. Unfortunately this inversion must be repeated for each trial change of the system variables. A sequential updating of all variables would thus be expected to require computer time growing as the square of the volume of the lattice. Most recent algorithm developments have concerned methods for eliminating this severe growth.

Actually it is only the A fields which are difficult to update in such a simulation. Ref. [22] 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

$$P(\chi) \propto e^{-\chi^2/2}. \tag{28}$$

Then construct

$$\phi = M\chi. \tag{29}$$

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.

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, several approximate algorithms gain substantially in speed by doing 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 after making a global trial change of the whole system. This forms the basis for the algorithms discussed in references [23-28] and [3].

Applying an acceptance condition to a global change has the danger that large increases in the action may give unfeasibly small final acceptances. 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. I will now discuss how, by appropriately biasing the trial changes, one can minimize this reduction. To further simplify notation, I define the full action

$$S(A) = S_0(A) + \frac{1}{2}\left(M^{-1}(A)\,\phi\right)^2 \tag{30}$$

To begin, consider a possible trial change of a single variable A to

$$A' = A + p\delta + F(A)\,\delta^2. \tag{31}$$

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 to be Gaussianly distributed

$$P(p) \propto e^{-p^2/2}. \tag{32}$$

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.* [11] 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. I can fully restore detailed balance by accepting the new value A' with probability

$$P_{\rm acc} = \min\left[1, e^{\{H(p,A) - H(p',A')\}}\right]. \tag{33}$$

Here H is a classical "Hamiltonian"

$$H(p,A) = p^2/2 + S(A). \tag{34}$$

In Eq. (33) 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

$$p' = -p - \left(F(A) + F\left(A'\right)\right)\delta. \tag{35}$$

After the accept/reject step, the momenta should be refreshed; thus, they should be modified in a Monte Carlo or other fashion that the preserves the distribution in Eq. (32).

Note that H is precisely the Hamiltonian used in the microcanonical algorithm [29] to describe evolution in "simulation time." Because of this analogy, I refer to H as the classical energy. Up to the sign of p', the change of variables indicated in Eq. (31) and Eq. (35) is an approximation to a Newtonian evolution by time δ under this Hamiltonian.

Note that with the second order terms in δ, the mapping defined by Eq. (31) and Eq. (35) exactly preserves areas in phase space

$$dA\ dp = dA'\ dp'. \tag{36}$$

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

This preservation of phase space volumes gives rise to a useful identity on the corresponding energy change. Consider the partition function

$$Z = \int dA'\ dp'\ e^{-H(A',p')}. \tag{37}$$

Changing variables to A and p and adding and subtracting $H(A,p)$ in the exponent gives

$$Z = \int dA\ dp\ e^{-H}\ e^{-(H'-H)} \tag{38}$$

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

$$\langle e^{-(H'-H)} \rangle = 1 \tag{39}$$

where the expectation is over all A and p with the equilibrium distribution $P(A,p) \propto e^{-H(A,p)}$. Because the exponential function is convex, Eq. (39) immediately implies

$$\langle H' - H \rangle \geq 0 \tag{40}$$

with equality only possible if the dynamics is exactly energy conserving. If we consider small changes in energy, a useful consequence of Eq. (39) follows by expanding the exponential

$$\langle H' - H \rangle = \frac{1}{2} \langle (H' - H)^2 \rangle + \mathcal{O}\left((H' - H)^3 \right). \tag{41}$$

To proceed, I use the parameter δ for an expansion of the energy change. It is readily verified that

$$H' - H = \left(p\delta + \frac{1}{2} \left(p^2 \frac{\partial}{\partial A} + 2F(A) \right) \delta^2 \right) \left(\frac{\partial S(A)}{\partial A} + 2F(A) \right) + \mathcal{O}\left(\delta^3\right). \tag{42}$$

This implies that the choice

$$F(A) = -\frac{1}{2}\frac{\partial S}{\partial A} \tag{43}$$

leads to an energy change

$$H' - H = \mathcal{O}\left(\delta^3\right) \tag{44}$$

and, by Eq. (41), that

$$\langle H' - H \rangle = \mathcal{O}\left(\delta^6\right). \tag{45}$$

Indeed, making this choice and ignoring the possibility of rejecting the trial change gives the usual Langevin algorithm [30], where the parameter δ is the square root of the step size used for discretization.

At first one might think that since the exact action is so difficult to calculate, the requisite derivative for this force would be intractable. Considering the action in Eq. (30), the Langevin force takes the form

$$F(A) = -\frac{1}{2}\frac{\partial S_0}{\partial A} + \frac{1}{2}\left(\left(MM^\dagger\right)^{-1}\phi, \quad \frac{\partial M}{\partial A}M^{-1}\phi\right) \tag{46}$$

Note that this requires knowledge of $\left(MM^\dagger\right)^{-1}\phi$ 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.

Let me now consider making the Langevin approach exact by adding a final Metropolis *et al.* [22,11] acceptance step. If I now consider updating some large number V of variables together, the positive $\mathcal{O}\left(\delta^6\right)$ quantities will coherently add and I expect to find a total energy change increasing linearly with V. This leads me to expect the acceptance to be exponentially suppressed when $V\delta^6$ is large.

To avoid this exponential suppression and have a reasonable acceptance requires $\delta \sim V^{-1/6}$. 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 is has both random and driven terms, the overall change in any given variable should go as

$$\Delta A = \mathcal{O}\left(\delta\sqrt{N}\right) + \mathcal{O}\left(\delta^2 N\right) = \mathcal{O}\left(\sqrt{N/V^{1/3}}\right) + \mathcal{O}\left(N/V^{1/3}\right). \tag{47}$$

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

$$T \sim V^{4/3}, \tag{48}$$

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

This algorithm was proposed in Ref. 23 and tested further with somewhat discouraging results in Ref. 24. Ref. 25 presents a quite promising variation, which I now discuss. Recapitulating on the above treatment of biased updatings, I constructed both the trial new A and the noise needed to return

$$A' = A + p\delta + F(A)\,\delta^2 \tag{49a}$$

$$p' = -p - \left(F(A) + F\left(A'\right)\right)\delta. \tag{49b}$$

This is an area preserving map of the (A, p) plane onto itself. The scheme proposed in Ref. 25 is to iterate the combination of this mapping with an inversion $p' \to -p'$ 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. 25.

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. (33) still makes the overall procedure exact. In this way the hybrid algorithm of ref. [31] becomes 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.

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. [31]. Refs. [27] 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. (47) is replaced by

$$\Delta A \sim N_{mic}\delta. \tag{50}$$

which should be valid as long as

$$N_{mic}\delta < \mathcal{O}(1). \tag{51}$$

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. (41) still applies to the overall evolution and I have for the expected energy change

$$\langle H' - H \rangle = \mathcal{O}\left(\delta^4\right). \tag{52}$$

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

$$P_{\text{acc}} \sim e^{-CV\delta^4} \tag{53}$$

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. (51). The final result is that the total time required to obtain a substantially changed lattice grows as

$$T \sim V^{5/4} \tag{54}$$

It has recently been realized that by using higher order discretization schemes it is possible to yet further reduce the asymptotic growth of this algorithm [32], although it is not clear that present systems are large enough to make such higher order approaches effective. One simple way to formulate the higher order schemes is in a recursive manner, starting with the above leapfrog.

Consider a discretization $T_n(\delta) : (A,p) \mapsto (A',p')$ of Hamilton's equations which is accurate to order n. By this I mean that A' and p' differ from the values obtained by evolving A and p for a time δ under the exact Hamiltonian by an amount of order δ^n. Furthermore, assume that this mapping is reversible in the sense

$$T_n^{-1}(\delta) = T_n(-\delta). \tag{55}$$

This reversibility is sufficient for detailed balance to be satisfied after a trajectory is accepted or rejected with the Metropolis criterion and the momenta are refreshed. For example, the discretization in Eq. (49) followed by an inversion of the momentum satisfies these conditions with $n = 3$.

It is now easy to combine these transformations to give a new transformation which is accurate to a higher order. First note that we don't need to worry about the case where n is even. Indeed, with n even the energy change in going a step δ would be, to this order, the same as the energy change going a step $-\delta$. Then the combination $T_n(\delta)\,T_n(-\delta)$ would change the energy by twice this amount. This, however, is inconsistent with the reversibility condition Eq. (55), which requires this combination to change nothing.

Thus, without loss of generality, assume n is odd. As we know of a scheme which gives $n = 3$, assume also that $n > 1$. We can then use the nonlinearity of the error to cancel the energy changes to order $n+2$. This construction is not unique, one combination that works is to sandwich a backward step between several forward ones

$$T_{n+2}((2i-s)\,\delta) = (T_n(\delta))^i \; T_n(-s\delta) \; (T_n(\delta))^i. \tag{56}$$

Here i is an arbitrary integer and $s = (2i)^{1/n}$ is a magnification factor for the central backward step. The size of s is determined by the requirement that the order n errors cancel.

Repeating the above arguments on volume dependence, the penalty paid for fermions is reduced to less than any power of the volume. The higher order algorithms do, however, pay a price in complexity, and the optimum order grows as the square root of the logarithm of the volume [32]. The overall theoretical computer time grows with volume as $T \sim Ve^{C\sqrt{\log V}}$.

One possible problem with the above hybrid microcanonical Monte Carlo schemes comes from the approximate microcanonical nature of the trajectories. If phase space is

separated into sectors by regions with a high potential energy, then it will be difficult to travel across the barriers between these sectors. One way such a division could occur is if the determinant of the fermionic matrix has zeros when considered as a function of the gauge field. Very little is known about the nature of these zeros. In principle this problem can be circumvented by including some Monte Carlo steps involving large trial changes of the gauge fields. For this purpose it would be convenient to have more intuition on the nature of the zeros of the fermionic determinant.

So far I have assumed that the fermion determinant is a positive quantity so that the integrand in Eq. (42) can be used as a probability weighting. Unfortunately, for the nuclear equation of state we do not want to be bound by this assumption. In particular, when the chemical potential is non-vanishing, that is whenever we have a net baryon density, then this is simply not true. Without a chemical potential we can always enforce positivity by artificially by adding an extra species of fermions interacting with the gauge potential with $M^\dagger$ rather than M. Then the full determinant will be a square. However with a baryonic potential added to M, these extra species are antibaryons, and the net baryon number will vanish.

If the determinant is not positive, one can in principle move its phase into observables and simulate with the absolute value of the determinant. This will technically give the correct answer if we divide the results by the average of the phase of the determinant. The problem is that for large systems this average phase is expected to be small numerically, and will have substantial statistical error. The average phase is expected to be strongly volume dependent. As it obtains contributions from the entire lattice volume, the average is naively expected to go exponentially to zero as the volume increases. This suggests that the net computer time to obtain a statistically significant result for a typical observable will grow exponentially with the system volume. Simulations of the Hubbard model indicate that this these problems can be overcome on small systems [33], but little has yet been done on the full theory of the strong interactions. The importance of this problem makes it particularly open for new algorithmic developments.

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. J. Kogut and L. Susskind, Phys. Rev. D11, 395 (1975).

8. M. Creutz, Phys. Rev. D15, 1128 (1977).

9. J. Kogut, Rev. Mod. Phys. 51, 659 (1979).

10. A. Polyakov, Phys. Lett. 72B, 477 (1978); L. Susskind, Phys. Rev. D20, 2610 (1979).

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

12. M. Fukugita, contribution in *Lattice Gauge Theory Using Parallel Processors*, Xiaoyuan Li, Zhaoming Qiu, and Hai-Cang Ren, eds. (Gordon and Breach, 1987).

13. P. Bacilieri et al., preprint ROM2F-88-020 (1988).

14. N. Christ, proceedings of the 1988 Symposium on Lattice Field Theory, Fermilab, 22-25 Sept. 1988 (to be published).

15. P. Bacilieri et al., preprint ROM2F-88-021.

16. See talks by P. van Baal and A. Kronfeld, proceedings of the 1988 Symposium on Lattice Field Theory, Fermilab, 22-25 Sept. 1988 (to be published).

17. See the reviews by C. Bernard and G. Martinelli in the proceedings of the 1988 Symposium on Lattice Field Theory, Fermilab, 22-25 Sept. 1988 (to be published).

18. K. Wilson, contribution in *New Phenomena in Subnuclear Physics*, ed. by A. Zichichi (Plenum Press, N.Y., 1977).

19. L. Karsten and J. Smit, Nucl. Phys. B183, 103 (1981).

20. P.T. Matthews and A. Salam, Nuovo Cim. 12, 563 (1954).

21. D. Weingarten and D. Petcher, Phys. Lett. 99B, 333 (1981).

22. G.G. Batrouni, G.R. Katz, A.S. Kronfeld, G.P. Lepage, B. Svetitsky and K.G. Wilson, Phys. Rev. D32, 2736 (1985).

23. R. T. Scalettar, D.J. Scalapino and R.L. Sugar, Phys. Rev. B34, 7911 (1986).

24. S. Gottlieb, W. Liu, D. Toussaint and R.L. Sugar, Phys. Rev. D35, 2611 (1986).

25. S. Duane, A. D. Kennedy, B.J. Pendleton and D. Roweth, Phys. Lett., 195, 2 (1987).

26. H. Gausterer and S. Sanielevici, Phys. Rev.D38, 1220 (1988).

27. R. Gupta, G. Kilcup, and S. Sharpe, Phys. Rev. D38, 1278 (1988).

28. K. Bitar, R. Horsley, A. Kennedy, S. Meyer and P. Rossi, preprint FSU-SCRI-88-14 (1988).

29. 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).

30. A. Ukawa and M. Fukugita, Phys. Rev. Lett. 55, 1854 (1985).

31. S. Duane and J. Kogut, Phys. Rev. Lett. 55, 2774 (1985).

32. M. Creutz and A. Gocksch, preprint BNL-42601 (1989).

33. S. White, R. Sugar, and R. Scalettar, U.C. Santa Barbara preprint (1988).