Lattice gauge theory in the microcanonical ensemble

David J. E. Callaway
CERN, CH-1211, Genève 23, Switzerland and High Energy Physics Division, Argonne National Laboratory, Argonne, Illinois 60439*

Aneesur Rahman
Materials Science and Technology Division, Argonne National Laboratory, Argonne, Illinois 60439
(Received 2 May 1983)

The microcanonical-ensemble formulation of lattice gauge theory proposed recently is examined in detail. Expectation values in this new ensemble are determined by solving a large set of coupled ordinary differential equations, after the fashion of a molecular dynamics simulation. Following a brief review of the microcanonical ensemble, calculations are performed for the gauge groups U(1), SU(2), and SU(3). The results are compared and contrasted with standard methods of computation. Several advantages of the new formalism are noted. For example, no random numbers are required to update the system. Also, this update is performed in a simultaneous fashion. Thus the microcanonical method presumably adapts well to parallel processing techniques, especially when the action is highly nonlocal (such as when fermions are included).

I. PROLEGOMENA

The numerical analysis of lattice gauge theories[1] by Monte Carlo methods[2] has provided a new viewpoint for the examination of the nonperturbative structure of quantum field theories. Progress in this area will certainly continue. It is important to keep in mind, however, that the Monte Carlo approach is only one of many ways of studying a large dynamical system. One particular alternative to the Monte Carlo technique—the microcanonical-ensemble method—was recently proposed by us[3]; here we explore this new ensemble further.

In the microcanonical-ensemble formulation of lattice gauge theory, expectation values are calculated by solving numerically a large set of coupled *ordinary* (as opposed to partial) differential equations. One feature of this method is that it is *deterministic,* that is, no random numbers are required to update the system. The microcanonical-ensemble formulation may also be more appropriate when calculations are performed by parallel processing techniques (such as with an array processor), for differential equations can be integrated simultaneously. Typically in a Monte Carlo calculation the lattice cannot be updated in a single pass[4] because each link interacts with its neighbors, and so a simultaneous update of link and neighbor would violate detailed balance.

Probably the most intriguing possibility of such a direct deterministic algorithm is the possibility of avoiding situations where the system becomes "trapped" in metastable minima. The literature in the study of fluids contains several examples where microcanonical-ensemble techniques are far superior to Monte Carlo methods (water is a familiar case).[5] Of course, each lattice field theory must be examined separately in order to determine which method is more suitable. In keeping with this philosophy, microcanonical-ensemble methods are applied here to the gauge groups U(1), SU(2), and SU(3). A comparative analysis of the two methods is given in each case.

II. REVIEW OF THE MICROCANONICAL ENSEMBLE

Recall how expectation values are calculated in the usual[1] approach to quantum field theory on a lattice. The central assumption of this formalism is that expectation values of functionals $\mathcal{O}$ of a finite number N of $\{\phi\}$ on a Euclidean lattice can give physically meaningful results without having to go to the continuum limit. If the action is denoted by $S\{\phi\}$, such expectation values are defined by

$$\langle \mathcal{O} \rangle_{\text{lattice}} \equiv Z_{\text{lattice}}^{-1} \int \mathcal{D}\phi \mathcal{O}\{\phi\} e^{-S\{\phi\}} , \tag{1a}$$

where

$$Z_{\text{lattice}} \equiv \int \mathcal{D}\phi e^{-S\{\phi\}} \tag{1b}$$

and

$$\int \mathcal{D}\phi \equiv \prod_{n}^{N} \int d\phi_n . \tag{1c}$$

The aforementioned continuum limit is obtained by taking the simultaneous limit where the number N of lattice fields ϕ_n (and the volume of the system) increases without bound and the lattice spacing approaches zero. The standard formula for lattice field theory, Eq. (1), can be cast in a form which is obviously equivalent to a classical canonical ensemble. Recall that any quantity independent of all $\{\phi\}$ can be added to the action $S\{\phi\} \equiv \beta V\{\phi\}$ without affecting the expectation value of any functional of the $\{\phi\}$. We choose to add a quantity T, defined by

$$T \equiv \tfrac{1}{2} \sum_{n} p_n^2 , \tag{2}$$

so that

$$\langle \mathcal{O} \rangle_{\text{lattice}} = \langle \mathcal{O} \rangle_{\text{canon}} , \\ \equiv Z_{\text{canon}}^{-1} \int \mathcal{D}p \int \mathcal{D}\phi \mathcal{O}\{\phi\} e^{-\beta H} , \tag{3a}$$

where

$$Z_{\text{canon}} \equiv \int \mathcal{D}p \int \mathcal{D}\phi e^{-\beta H} \tag{3b}$$

and

$$H = T + V . \tag{3c}$$

The $\{p\}$ and $\{\phi\}$ are independent variables on the lattice.

An isomorphism with the canonical ensemble of statistical mechanics arises upon making the identification

$$p_n = \frac{d\phi_n}{d\tau} \tag{4}$$

provided that a new "artificial dimension" τ is introduced. A Lagrangian formulation of field theory in d dimensions is thus mapped to a "Hamiltonian" formalism embedded in d discrete dimensions and one continuous (artificial) dimension.

Equations (2) and (3) define a canonical ensemble with kinetic energy $T\{p\}$, potential energy $V\{\phi\}$ $(=\beta^{-1}S\{\phi\})$, and temperature β^{-1}. This ensemble represents the states of a system governed by a Hamiltonian H in contact with a heat bath at fixed temperature β^{-1}.

A *microcanonical* ensemble on the other hand describes a system in thermal isolation and therefore the "energy" $H[\{p\},\{\phi\}]$ of the system is constrained to be a fixed value E, while the "temperature" β^{-1} is defined to be the average kinetic energy (suitably normalized).[6] Of course the "kinetic energy" T is a fluctuating quantity. Each coordinate ϕ_n and momentum p_n evolves in τ according to Hamilton's equations,

$$\frac{dp_n}{d\tau} = \frac{-\partial V\{\phi\}}{\partial \phi_n} , \tag{5}$$

$$\frac{d\phi_n}{d\tau} = p_n . \tag{6}$$

Expectation values in the microcanonical ensemble are formally defined by

$$\langle \mathcal{O} \rangle_{\text{micro}} \equiv Z_{\text{micro}}{}^{-1}(E) \int \mathcal{D}p \int \mathcal{D}\phi \mathcal{O}[\{p\},\{\phi\}] \delta(H-E) , \tag{7a}$$

where

$$Z_{\text{micro}}(E) \equiv \int \mathcal{D}p \int \mathcal{D}\phi \delta(H-E) . \tag{7b}$$

The integrals in Eqs. (7) are over the $(2N-1)$-dimensional hypersurface of constant "energy" defined by $H[\{p\},\{\phi\}]=E$.

At this point it is useful to remind ourselves how the canonical ensemble arises in classical physics from the more fundamental microcanonical ensemble. Given the density of states $Z_{\text{micro}}(E)$ defined in Eqs. (7) it is possible to define the entropy of a system,

$$S(E,V) \equiv \ln Z_{\text{micro}}(E) , \tag{8a}$$

and the inverse temperature β,

$$\beta \equiv \frac{\partial S(E,V)}{\partial E} . \tag{8b}$$

Consider a very large system of energy E_t in thermal isolation (see Fig. 1). Imagine that this system is actually composed of two subsystems, a large one whose energy is given by E' and a much smaller one of energy H. Then

$$E_t = E' + H \tag{9}$$

is constant in time. Let us now calculate the probability $P(H)$ that the smaller system has energy H. This probability is proportional to the density of states $Z_{\text{micro}}(E')$ for the larger subsystem evaluated at energy E':

$$P(H) \propto \frac{Z_{\text{micro}}(E_t - H)}{Z_{\text{micro}}(E_t)} = \exp[S(E_t - H) - S(E_t)] . \tag{10}$$

Since the larger subsystem consists of many degrees of freedom, its temperature β^{-1} is essentially fixed. By Eq. (8b),

$$P(H) \propto \exp[-\beta H] \tag{11}$$

which is the standard formula for the probability distribution in the canonical ensemble. Thus the canonical ensemble is a consequence of the microcanonical ensemble.

Expectation values in the canonical ensemble are related to those in the microcanonical ensemble by[7,8]

$$\langle \mathcal{O} \rangle_{\text{canon}} = Z_{\text{canon}}{}^{-1} \int \langle \mathcal{O} \rangle_{\text{micro}} Z_{\text{micro}}(E) e^{-\beta E} dE . \tag{12}$$

In the (thermodynamic) limit of a large system, the probability distribution of energies in the canonical ensemble (proportional to the factor $e^{-\beta E}$) multiplied by the density of states $Z_{\text{micro}}(E)$ becomes sharply peaked about its mean value, which is denoted by $\bar{E}$. Thus it is logical to expand expectation values in the microcanonical ensemble about the value obtained by setting $E = \bar{E}$:

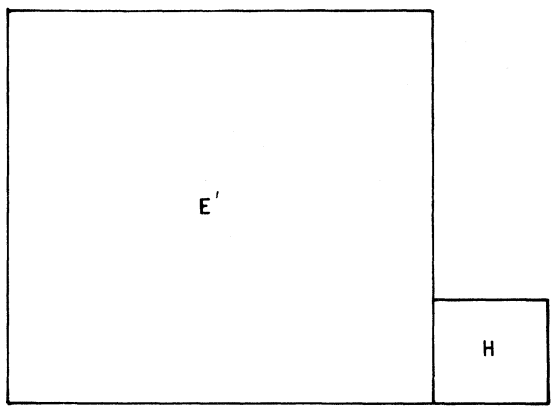

FIG. 1. Schematic diagram of a thermodynamic system of total energy $E_t = E' + H$ which consists of a large subsystem of energy E' in thermal contact with a much smaller subsystem of energy H.

$$\langle \mathcal{O} \rangle_{\text{micro}} = \langle \mathcal{O} \rangle_{\text{micro}} \Big|_{\bar{E}} + (E-\bar{E}) \frac{\partial}{\partial \bar{E}} \langle \mathcal{O} \rangle_{\text{micro}} \Big|_{\bar{E}} + \tfrac{1}{2}(E-\bar{E})^2 \frac{\partial^2}{\partial \bar{E}^2} \langle \mathcal{O} \rangle_{\text{micro}} \Big|_{\bar{E}} + \cdots . \tag{13}$$

Equation (13) is substituted into Eq. (12) to yield

$$\langle \mathcal{O} \rangle_{\text{canon}} = \langle \mathcal{O} \rangle_{\text{micro}} \Big|_{\bar{E}} + \tfrac{1}{2} \langle (E-\bar{E})^2 \rangle_{\text{canon}} \frac{\partial}{\partial \bar{E}^2} \langle \mathcal{O} \rangle_{\text{micro}} \Big|_{\bar{E}} + \cdots . \tag{14}$$

Equation (14) can be inverted to this order to give

$$\begin{aligned} \langle \mathcal{O} \rangle_{\text{micro}} &\cong \langle \mathcal{O} \rangle_{\text{canon}} - \langle (E-\bar{E})^2 \rangle_{\text{canon}} \frac{1}{2} \frac{\partial^2}{\partial \bar{E}^2} \langle \mathcal{O} \rangle_{\text{canon}} \\ &\cong \langle \mathcal{O} \rangle_{\text{canon}} + \frac{\partial \bar{E}}{\partial \beta} \frac{1}{2} \frac{\partial^2}{\partial \bar{E}^2} \langle \mathcal{O} \rangle_{\text{canon}} \\ &\cong \langle \mathcal{O} \rangle_{\text{canon}} + \frac{1}{2} \frac{\partial}{\partial \beta} \left[\left[\frac{\partial \beta}{\partial \bar{E}} \right] \frac{\partial}{\partial \beta} \langle \mathcal{O} \rangle_{\text{canon}} \right] . \end{aligned} \tag{15}$$

Therefore, in the thermodynamic limit expectation values in the two ensembles are related by

$$\langle \mathcal{O} \rangle_{\text{micro}} = \langle \mathcal{O} \rangle_{\text{canon}} + O(N^{-1} \langle \mathcal{O} \rangle_{\text{canon}}) . \tag{16}$$

It should be pointed out that in relating quantities which have the form of a variance [such as $\langle (T - \langle T \rangle)^2 \rangle$] calculated in the two ensembles, a more careful analysis[7] using the relation Eq. (15) is necessary, since the leading-order terms cancel in the first term in that relation.

In practice the expectation values Eq. (7) are calculated by integrating the Hamiltonian equations of motion, Eqs. (5) and (6) along a trajectory in the $2N$-dimensional phase space $[\{p\},\{\phi\}]$. Any expectation value in the microcanonical ensemble is thus given by

$$\langle \mathcal{O} \rangle_{\text{micro}} = \lim_{\tau \to \infty} \frac{1}{\tau} \int_0^\tau \mathcal{O}[\{p(\tau')\}, \{\phi(\tau')\}] d\tau' \tag{17}$$

and is a unique function of E.

The equivalence of Eqs. (7) and (17) follows from a vital assumption of ensemble theory known as the principle of equal weight.[9] This principle assures us that for a given E, the trajectories given by the solution of Eqs. (5) and (6) cover the $(2N-1)$-dimensional constant-energy surface in phase space $[\{p\},\{\phi\}]$ with equal density.

The average kinetic energy [which can be used to calculate β in U(1) and SU(2)] along a trajectory is given by

$$\text{U(1),SU(2):} \quad \beta^{-1} = \frac{2}{N_{\text{indep}}} \langle T \rangle_{\text{micro}} . \tag{18}$$

For the gauge group SU(3) a different method is used. The reason the divisor in Eq. (18) is N_{indep} and not N arises in part because of the local gauge symmetry present in lattice gauge theories.[3] This point is discussed further in the following sections. Another quantity of interest is the average action,

$$\langle S \rangle_{\text{micro}} = \langle V \rangle_{\text{micro}} . \tag{19}$$

III. GAUGE THEORIES IN THE MICROCANONICAL ENSEMBLE

A. Compact electrodynamics [U(1)]

The standard formulation[1] of compact electrodynamics [i.e., a lattice U(1) gauge theory] utilizes the action

$$\begin{aligned} &S = \beta V \text{ with } \beta = g_0^{-2} , \\ &V \equiv \sum_{\square} \text{Re}(1 - U_{n,\mu} U_{n+\mu,\nu} U^{-1}{}_{n+\nu,\mu} U^{-1}{}_{n,\nu}) , \\ &U_{n,\mu} = \exp(i\phi_{n,\mu}) , \end{aligned} \tag{20}$$

where g_0^2 is the bare lattice coupling constant and the sum defining V is over all elementary plaquettes in d dimensions. The $\{\phi\}$ are real gauge fields associated with each link of the lattice; the link $U_{n,\mu}$ connects the lattice point n to its nearest neighbor in the direction μ.

We now present a calculation of the average plaquette as a function of β for compact electrodynamics. The average plaquette is defined by the formula

$$P \equiv [2/d(d-1)] L^{-d} \langle V \rangle \tag{21}$$

on a hypercubic Euclidean lattice in d dimensions whose side is of length L lattice sites. A comparison is made between the average plaquette calculated in the microcanonical ensemble and in the canonical ensemble (i.e., by standard Monte Carlo methods). The average plaquette is calculated by integrating numerically the equations of motion Eqs. (5) and (6) and substituting them in Eqs. (17) and (21). The bare lattice coupling $\beta = g_0^{-2}$ is calculated in the same way, except Eq. (18) replaces Eq. (21) with

$$\begin{aligned} \text{U(1):} \quad N_{\text{indep}} &= \left[\frac{d-1}{d} \right] N \\ &= (d-1) L^d , \end{aligned} \tag{22}$$

where $N = dL^d$ is the number of links in the lattice.

As mentioned above, the reason N_{indep} is less than N is because there is a local gauge symmetry in the system which has no immediate physical relevance. General methods exist[10] for dealing with these redundant variables, familiar from classical mechanics as ignorable cyclic coordinates. One way to remove these ignorable coordinates is to impose constraints by choosing a gauge. Choices such as the axial gauge ($n_\mu A^\mu = 0$) or the Lorentz gauge ($\partial_\mu A^\mu = 0$) result in a set of N/d constraints [cf. Eq. (22)] on a lattice of N links in d dimensions. The following course seems more appropriate here.

By a suitable linear transformation from the variables $\phi_{n,\mu}$ to new variables ξ_i, the Hamiltonian equations of motion Eqs. (5) and (6) become

$$\ddot{\xi}_i \equiv \frac{d^2\xi_i}{d\tau^2} = f_i(\{\xi\}) , \quad i=1,2,\ldots,N \tag{23a}$$

with

$$f_i(\{\xi\})=0 , \quad N_{\text{indep}} < i \leq N . \tag{23b}$$

It then follows that in solving Eqs. (5) and (6) if the initial conditions $\dot{\phi}_{n,\mu}=0$ at $\tau=0$ are imposed, the $\dot{\xi}_i$ for $N_{\text{indep}} < i \leq N$ will equal zero for all τ. As a consequence in the average in Eq. (9) the correct divisor is N_{indep} and not N. This procedure is analogous to determining the temperature of a system of particles by measuring the average kinetic energy of each particle in the rest frame of the system (equivalent to imposing d constraints).

The results of the computation in each of the two ensembles are now presented. In the microcanonical-ensemble calculation the ordinary differential equations Eqs. (5) and (6) are solved by the Runge-Kutta method with step size $\Delta\tau=0.01$.[11] At $\tau=0$ each $\phi_{n,\mu}$ is chosen randomly between 0 and 2π, and each $\dot{\phi}_{n,\mu}$ is set equal to zero (for reasons described above). Expectation values are obtained as an average over $\tau=2000\ \Delta\tau$ (i.e., by using 2000 consecutive configurations).

After each calculation (or set of 2000 "measurements") is completed, the value of E is altered by multiplying all the $\{p\}$ by the same factor. This operation effectively heats (cools) the system if the factor is greater than (less than) unity. The Runge-Kutta algorithm is then applied to step the system forward from τ to $\tau+\Delta\tau$. By a proper choice of this heating/cooling factor, after each step the value of β can be easily adjusted to a given value with arbitrary accuracy after enough steps. The heating/cooling cycle is terminated after 1000 $\Delta\tau$ and the system is allowed to equilibrate for 3000 $\Delta\tau$ before more measurements are taken. On each such energy shell in a state of equilibrium the system has a unique value of $\langle V\rangle_{\text{micro}}$ and hence of $\langle T\rangle_{\text{micro}}$ and β. Note that the heating/cooling procedure conserves the vanishing of $\dot{\xi}_i$ for $i > N_{\text{indep}}$.

In the canonical-ensemble calculation the standard Metropolis *et al.*[12] Monte Carlo algorithm is used. At each β, equilibration for 200 iterations is allowed after which the average plaquette is measured for 2000 iterations. Further details of this calculation are given elsewhere.[13]

Figure 2 displays the value of the average plaquette as a function of β for both ensembles. The agreement between the two calculations is excellent, and the amount of computer time required for each calculation is roughly the same.

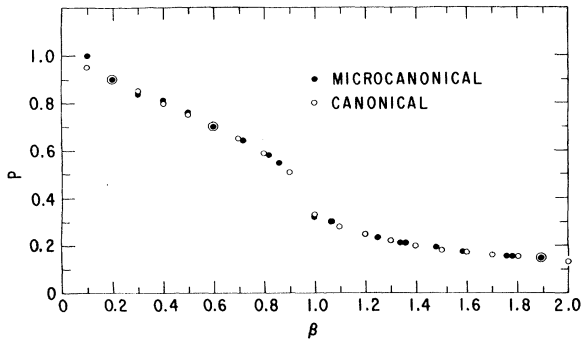

FIG. 2. Comparison of the results of a microcanonical-ensemble calculation of the average plaquette with a standard Monte Carlo calculation for compact electrodynamics [a U(1) lattice gauge theory]. The legend is filled circles, microcanonical; open circles, canonical (Monte Carlo).

B. SU(2)

The method of the microcanonical ensemble is next applied to the gauge group SU(2). For convenience we consider the fundamental representation only. The action in this representation can be written in the form

$$S=\beta V \text{ with } \beta=g_0{}^{-2} , \tag{24}$$
$$V=\sum_{\square}[1-\tfrac{1}{2}\mathrm{Tr}(U_{n,\mu}U_{n+\mu,\nu}U^{-1}{}_{n+\nu,\mu}U^{-1}{}_{n,\nu})] ,$$

where as usual the sum is over all ordered elementary plaquettes. Each of the $U_{n,\mu}$ can be parametrized[14] in terms of the 2×2 unit matrix 1 and the Pauli matrices σ^α:

$$U_{n,\mu}=a_0 1+ia_\alpha\sigma^\alpha , \tag{25}$$

where $\sigma^\alpha\sigma^\beta=i\epsilon^{\alpha\beta\gamma}\sigma^\gamma+\delta^{\alpha\beta}$.

The integration measure for SU(2) can then be written in the form

$$\int dU=\int \frac{d^4a}{2\pi^2}\delta(\underline{a}^2-1) , \tag{26}$$

where the square of quadruple $\underline{a}$ is $\underline{a}^2\equiv a_0{}^2+a_1{}^2+a_2{}^2+a_3{}^2$. If the four links comprising an elementary plaquette are consecutively parametrized (see Fig. 3) by quadruples $(a_0,\vec{a})$; $(b_0,\vec{b})$; $(c_0,\vec{c})$; and $(d_0,\vec{d})$ the potential for a single such plaquette can be written in the form

$$V_\square=1-\tfrac{1}{2}\mathrm{Tr}(U_aU_bU_c^\dagger U_d^\dagger)$$
$$=1-(A\cdot B)(C\cdot D)-(AB)_-\cdot(CD)_+ , \tag{27a}$$

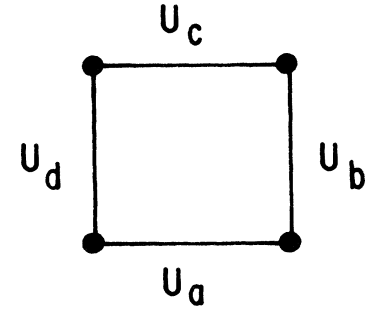

FIG. 3. Schematic diagram illustrating the notation used in Eqs. (27).

where

$$A\cdot B\equiv a_0b_0-\vec{a}\cdot\vec{b}\ , \tag{27b}$$

and

$$(AB)^{\gamma}_{\pm}\equiv a_0b_\gamma+b_0a_\gamma\pm a_\alpha b_\beta\epsilon_{\alpha\beta\gamma}\ . \tag{27c}$$

The microcanonical ensemble is then constructed via the Hamiltonian equations of motion. The force on a_α, a component of $\underline{a}$, is given by

$$\frac{dp_\alpha}{d\tau}\equiv\dot{p}_\alpha=-\frac{\partial V}{\partial a_\alpha}=-\sum_{\square}\frac{\partial V_\square}{\partial a_\alpha}\ ,$$
$$\dot{a}_\alpha=p_\alpha\ , \tag{28}$$

and similarly for $\underline{b}$, $\underline{c}$, and $\underline{d}$. These equations must be solved subject to the constraint $\underline{a}^2=1$. This constraint is enacted by means of the method of Lagrange multipliers.[15] Because the techniques of constraint dynamics are not well known to particle theorists, the method[15] for integrating Eqs. (28) is presented in some detail.

Consider the effect of adding the constraint potential V_c to V:

$$V_{\text{total}}=V+V_c\ ,$$
$$V_c=\tfrac{1}{2}\sum_{\underline{a}}\lambda(\underline{a})(1-\underline{a}^2)\ , \tag{29}$$

in order to generate (via the equations of motion) the total force on the system, including forces of constraint. Assume that the coordinates $\underline{a}$ and their velocities $\dot{\underline{a}}$ are known at τ. Call the force generated *only* by V (not including the force of constraint) $\underline{F}(\tau)$. Next construct the four-component quantity (which we call the incomplete new coordinate)

$$\underline{\xi}(\tau+h)=\underline{a}(\tau)+h\dot{\underline{a}}(\tau)+\tfrac{1}{2}h^2\underline{F}(\tau)\ , \tag{30}$$

where $h\equiv\Delta\tau$. Then, using the derivative of V_c given above, the relation

$$\underline{a}(\tau+h)=\underline{\xi}(\tau+h)+\tfrac{1}{2}h^2\lambda(\underline{a};\tau)\underline{a}(\tau) \tag{31}$$

is used to calculate the updated value $\underline{a}(\tau+h)$ once $\lambda(\underline{a};\tau)$ in Eq. (31) is determined by the requirement

$$\underline{a}(\tau+h)\cdot\underline{a}(\tau+h)=1\ . \tag{32}$$

The quadratic equation in $\lambda(\underline{a};\tau)$ resulting from Eq. (32) yields

$$\tfrac{1}{2}h^2\lambda(\underline{a};\tau)=-\underline{\xi}(\tau+h)\cdot\underline{a}(\tau)$$
$$+\{[\underline{a}(\tau)\cdot\underline{\xi}(\tau+h)]^2+1-\underline{\xi}^2(\tau+h)\}^{1/2}\ , \tag{33}$$

provided $\underline{a}(\tau)\cdot\underline{a}(\tau)=1$.

The update of the lattice is completed by calculating $\dot{\underline{a}}(\tau+h)$. First, a new set of quadruple forces $\underline{F}'(\tau+h)$ is calculated from $\underline{a}(\tau+h)$ using the potential V of Eq. (29). From these forces a new four-component incomplete velocity

$$\underline{\eta}(\tau+h)\equiv\dot{\underline{a}}(\tau)+\tfrac{1}{2}h[\underline{F}(\tau)+\lambda(\underline{a};\tau)\underline{a}(\tau)+\underline{F}'(\tau+h)] \tag{34}$$

is constructed. In terms of $\underline{\eta}(\tau+h)$, $\dot{\underline{a}}(\tau+h)$ is given by

$$\dot{\underline{a}}(\tau+h)=\underline{\eta}(\tau+h)+\mu h\underline{a}(\tau+h)\ , \tag{35}$$

where the new constraint parameter $\mu(\tau)$ is determined by setting

$$\underline{a}(\tau+h)\cdot\dot{\underline{a}}(\tau+h)=0 \tag{36}$$

[since $(\dot{\underline{a}}^2)=2\underline{a}\cdot\dot{\underline{a}}=0$], which yields

$$\mu h=-\underline{\eta}(\tau+h)\cdot\underline{a}(\tau+h)\ . \tag{37}$$

When Eq. (37) is substituted into Eq. (35), the update of the lattice coordinates and momenta can be completed.

The method of the microcanonical ensemble can be illustrated in the case of SU(2) [much as for U(1)] by a calculation of the average plaquette P as a function of β. The quantities P and β are determined as before from the relations Eqs. (21) and (18). It is important to remember however that there are *three* independent variables per link, and thus for the usual hypercubical lattice,

$$\text{SU(2):}\quad N_{\text{indep}}=3(d-1)L^d \tag{38}$$

[compare Eq. (22)].

Because the algorithm defined by Eqs. (30)–(37) updates coordinates and momenta separately, the same technique of heating/cooling the system can be used to adjust β to any desired value. The system is initially arranged with all the momenta equal to zero and the coordinates $\underline{a}$ set equal to random numbers satisfying the constraint $\underline{a}\cdot\underline{a}=1$. The system is then heated/cooled for 100 iterations, and allowed to equilibrate for 500 iterations before the average plaquette is measured for 1000 iterations. The step size h equals 0.1. The result is plotted as a function of β in Fig. 4. For comparison, a standard ("heat bath") calculation of $P(\beta)$ is plotted, following Creutz.[14] In the heat-bath calculation the system is allowed to equilibrate for 500 iterations before the plaquette is measured for 1000 iterations.

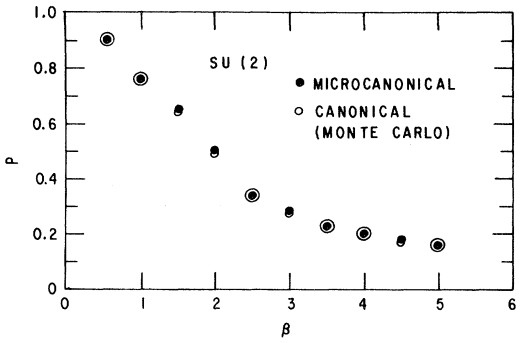

FIG. 4. Comparison of the results of a microcanonical-ensemble calculation of the average plaquette with a heat-bath calculation for the gauge group SU(2). The legend is as for Fig. 2.

The microcanonical and canonical (heat bath) methods required about the same amount of computer time to update the lattice. However, except for large values of β ($\beta \gtrsim 5$) the heat-bath method provided faster convergence to the plaquette value. For the local Wilson action in SU(2), the standard heat-bath algorithm[14] may thus be generally superior to our particular microcanonical method.

C. SU(3)

The same techniques can also be applied to SU(3). An element g of the fundamental representation can be written in the form

$$g=\begin{bmatrix} u_1^* & u_2^* & u_3^* \\ v_1^* & v_2^* & v_3^* \\ w_1 & w_2 & w_3 \end{bmatrix} \equiv \begin{bmatrix} \underline{u}^* \\ \underline{v}^* \\ \underline{w}^* \end{bmatrix} , \tag{39}$$

subject to the constraints

$$\underline{u}^* \cdot \underline{u}=1 , \tag{40a}$$

$$\underline{v}^* \cdot \underline{v}=1 , \tag{40b}$$

$$\underline{u}^* \cdot \underline{v}=0 , \tag{40c}$$

while

$$\underline{w}=\underline{u}\times\underline{v} . \tag{41}$$

The 12 components of the complex three-component vectors $\underline{u}$ and $\underline{v}$ are taken as dynamical variables subject to the four constraints Eqs. (40). The integration volume (or Haar measure) is given by[16]

$$\int dU=\int \frac{d^6u d^6v}{2\pi^2}\delta(|\underline{u}|^2-1)\delta(|\underline{v}|^2-1)\delta^2(\underline{u}^*\cdot\underline{v}) , \tag{42}$$

where $\delta^2(\underline{u}^*\cdot\underline{v})=\delta[\mathrm{Re}(\underline{u}^*\cdot\underline{v})]\delta[\mathrm{Im}(\underline{u}^*\cdot\underline{v})]$. The action is taken to be

$$S=\beta V , \tag{43}$$

with

$$V=\sum_{\square}[1-\tfrac{1}{3}\,\mathrm{Re\,Tr}(U_{n,\mu}U_{n+\mu,\nu}U^{-1}{}_{n+\nu,\mu}U^{-1}{}_{n,\nu})] .$$

A typical term in the sum over plaquettes can be written in the form

$$V=1-\tfrac{1}{3}\,\mathrm{Re\,Tr}(g^{\dagger}m) , \tag{44a}$$

where

$$m=\begin{bmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{bmatrix} , \tag{44b}$$

is the sum over all the $d(d-1)$ plaquettes interacting with the link denoted by g [Eq. (39)]. The equations of motion of the 12 components of the system are found in the usual fashion from Eqs. (44). Then the four constraints Eqs. (40) are introduced via the constraint potential

$$V_c\equiv\tfrac{1}{2}\lambda_u(1-\underline{u}^*\cdot\underline{u})+\tfrac{1}{2}\lambda_v(1-\underline{v}^*\cdot\underline{v}) +2\,\mathrm{Re}(\lambda_{uv}u^*\cdot v) , \tag{45}$$

where the real parameters λ_u and λ_v and the complex parameter λ_{uv} are determined using a generalization of Eqs. (32).

The coordinates $\underline{u}$ and $\underline{v}$ and their velocities $\dot{\underline{u}}$ and $\dot{\underline{v}}$ are taken as known at τ. Denote the complex three-component force on the complex triple $\underline{u}$ arising only from the potential V by $\underline{F}_u$. The incomplete new complex coordinate is then

$$\underline{\xi}_u=\underline{u}(\tau)+h\dot{\underline{u}}(\tau)+\tfrac{1}{2}h^2\underline{F}_u(\tau) \tag{46a}$$

to order h^2, while $\underline{u}(\tau+h)$ is given by

$$\underline{u}(\tau+h)=\underline{\xi}_u+\tfrac{1}{2}h^2\lambda_u\underline{u}(\tau)+\tfrac{1}{2}h^2 2\lambda_{uv}\underline{v}(\tau) , \tag{46b}$$

to the same order. The equations for $\underline{v}$ are similar:

$$\underline{\xi}_v=\underline{v}(\tau)+h\dot{\underline{v}}(\tau)+\tfrac{1}{2}h^2\underline{F}_v(\tau) , \tag{46c}$$

$$\underline{v}(\tau+h)=\underline{\xi}_v+\tfrac{1}{2}h^2\lambda_v\underline{v}(\tau)+\tfrac{1}{2}h^2 2\lambda^*_{uv}\underline{u}(\tau) . \tag{46d}$$

The two real multipliers λ_u and λ_v and the complex multiplier λ_{uv} are calculated by enforcing the constraints Eq. (40) at $\tau+h$, assuming they are satisfied at τ. To order h^2, the three multipliers are thus determined by the linear system

$$\begin{aligned} \tfrac{1}{2}h^2\lambda^*_{uv}(\tau)&=-\frac{1}{2}\frac{\underline{\xi}^*_u\cdot\underline{\xi}_v+\frac{1}{2}h^2\lambda_v\underline{\xi}^*_u\cdot\underline{v}+\frac{1}{2}h^2\lambda_u\underline{\xi}_v\cdot\underline{u}^*}{\underline{\xi}^*_u\cdot u+\underline{\xi}_v\cdot v^*} , \\ \tfrac{1}{2}h^2\lambda_u(\tau)&=\frac{1-|\underline{\xi}_u|^2-4\frac{1}{2}h^2\mathrm{Re}(\lambda^*_{uv}\underline{v}^*\cdot\underline{\xi}_u)}{2\,\mathrm{Re}(\underline{u}^*\cdot\underline{\xi}_u)} , \\ \tfrac{1}{2}h^2\lambda_v(\tau)&=\frac{1-|\underline{\xi}_v|^2-4\frac{1}{2}h^2\mathrm{Re}(\lambda_{uv}\underline{u}^*\cdot\underline{\xi}_v)}{2\,\mathrm{Re}(\underline{v}^*\cdot\underline{\xi}_v)} . \end{aligned} \tag{47}$$

The coordinates $\underline{u}(\tau+h)$ and $\underline{v}(\tau+h)$ are calculated according to Eqs. (46), and new forces $\underline{F}'_u$ and $\underline{F}'_v$ are calculated from these new coordinates. Then, generalizing Eq. (34), new incomplete velocities $\underline{\eta}$ are calculated:

$$\underline{\eta}_u(\tau)=\dot{\underline{u}}(\tau)+\tfrac{1}{2}h[\underline{F}_u(\tau)+\lambda_u\underline{u}(\tau)+2\lambda_{uv}\underline{v}(\tau) +\underline{F}'_u(\tau+h)] , \tag{48a}$$

$$\underline{\eta}_v(\tau)=\dot{\underline{v}}(\tau)+\tfrac{1}{2}h[\underline{F}_v(\tau)+\lambda_v\underline{v}(\tau)+2\lambda^*_{uv}\underline{u}(\tau) +\underline{F}'_v(\tau+h)] . \tag{48b}$$

The updated velocities of the $\underline{u}$ and $\underline{v}$ are then given by

$$\dot{\underline{u}}(\tau+h)=\underline{\eta}_u(\tau)+\frac{h}{2}\mu_u\underline{u}(\tau+h)+\frac{h}{2}2\mu_{uv}\underline{v}(\tau+h) , \tag{49a}$$

$$\dot{\underline{v}}(\tau+h)=\underline{\eta}_v(\tau)+\frac{h}{2}\mu_v\underline{v}(\tau+h)+\frac{h}{2}2\mu^*_{uv}\underline{u}(\tau+h) . \tag{49b}$$

The two real multipliers μ_u and μ_v and the complex multiplier μ_{uv} are determined by setting the derivatives (with respect to τ) of the constraints Eq. (40) equal to zero at $\tau+h$, assuming they were zero at τ. This operation yields

$$\frac{h}{2}\mu_u = -\mathrm{Re}(\underline{u}^* \cdot \underline{\eta}_u) ,$$
$$\frac{h}{2}\mu_v = -\mathrm{Re}(\underline{v}^* \cdot \underline{\eta}_v) , \tag{50}$$
$$\frac{h}{2}2\mu_{uv}^* = -\tfrac{1}{2}(\underline{\eta}_u^* \cdot \underline{v} + \underline{u}^* \cdot \underline{\eta}_v) .$$

With the use of Eqs. (49) and (50), $\dot{\underline{u}}(\tau+h)$ and $\dot{\underline{v}}(\tau+h)$ are calculated. This operation completes the update of the lattice. For the calculation of β (and hence the bare lattice coupling constant) a new technique must be applied. This difference is mandated by the fact that the parametrization, Eq. (39), of the SU(3) group element does not permit the separation of the gauge degrees of freedom by a linear transformation [cf. Eqs. (23)]. Thus the cyclic coordinates must be handled in a different fashion, described below.

A combination of the microcanonical and Monte Carlo methods can be formed as follows. The SU(3) lattice is updated as described above, except that instead of using Eqs. (48) the incomplete velocities $\underline{\eta}_u, \underline{\eta}_v$ of a small number of randomly chosen links are set equal to random numbers obeying a Gaussian distribution. This procedure is equivalent to establishing weak links between the microcanonical ensemble and a heat bath. The root mean square of each velocity in the system will then eventually approach the width of the Gaussian distribution. The bare lattice coupling constant is determined by

$$\text{SU(3):}\quad g_0^2 = \beta^{-1} = (8dL^d)^{-1}\left\langle \sum (|\dot{u}|^2 + |\dot{v}|^2) \right\rangle_{\text{micro}} \tag{51}$$

(recall there are eight independent variables per link) on the usual hypercubic lattice.

After the system is well equilibrated, the links between the microcanonical system and the heat bath can be cut, and expectation values such as the average plaquette can be calculated in the usual fashion. Alternatively, one can imagine storing the coordinates of well-equilibrated systems produced either by microcanonical or Monte Carlo methods. Later, when a calculation is to be performed, the coordinates can be recalled and the velocities generated according to an appropriate Gaussian distribution.

In Fig. 5 the results of microcanonical and Monte Carlo calculations for SU(3) on a 3^4 lattice are presented. In each case the coordinates are initially random [but satisfy the constraints Eqs. (40) and (41)]. For the microcanonical calculation, the velocities were initially set equal to an appropriately constrained Gaussian distribution for the desired β. Next the system was placed in weak contact with a heat bath (so that each link had a 1% chance of being touched during each update) for 250 τ steps, with $\Delta\tau = 0.05$.[17] The system was then allowed to equilibrate for 250 steps after which the average plaquette is measured for 350 steps.

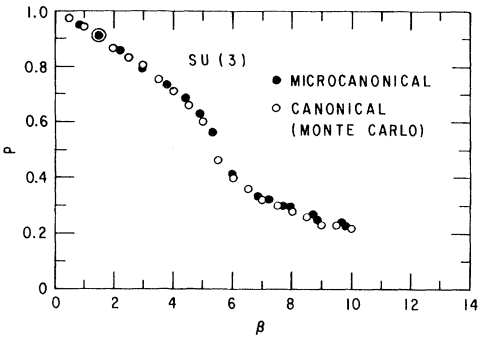

FIG. 5. Comparison of a microcanonical-ensemble calculation of the average plaquette with a Monte Carlo calculation for the gauge group SU(3). The legend is as for Fig. 2.

The Monte Carlo calculation used the standard Metropolis[12] procedure.[18] The system is first allowed to equilibrate for 100 iterations, after which the average plaquette is measured for 350 iterations.

As can be seen in Fig. 5, the agreement between the two calculations is excellent. Good agreement with Monte Carlo data was also obtained when the microcanonical calculation was performed without any contact with a heat bath, or when the contact was maintained throughout the calculation. The velocities were initially chosen in accordance with a Gaussian distribution as before.

Because of the relative importance of the SU(3) lattice gauge theory, we made an extensive comparative analysis of the microcanonical and canonical (Monte Carlo) methods. We could not find conclusive evidence of the superiority of either method. In both schemes, it appeared that long "time" correlations were present, so questions of convergence and metastability can be important in SU(3).

One quantity which gives some indication of how rapidly the system covers configuration space is

$$\chi_n^2 \equiv (4dL^d)^{-1}\sum (|u_{n+r}-u_r|^2 + |v_{n+r}-v_r|^2) , \tag{52}$$

where the subscripts r and $n+r$ refer to the iteration number (in the canonical ensemble) or τ step number (in the microcanonical ensemble). First an equilibrated system is constructed, and the values of the coordinates u_r and v_r are stored. Then for each following iteration (or τ step) the quantity χ_n^2 is evaluated. Initially χ_0^2 is zero; as n becomes large χ_n^2 approaches unity.

The value of n for which χ_n^2 is essentially unity (say to within 10%) gives a measure of how many iterations (or steps) are required to cover configuration space, i.e., to make a reasonable measurement of some quantity. On a 3^4 lattice with $\beta \sim 10$, we found $n \sim 70$ for iterations in the Monte Carlo calculation,[18] but only $n \sim 60$ steps ($h = 0.1$) were needed in the microcanonical simulation.[17] Thus once the system is in equilibrium the microcanonical method may be slightly more efficient in updating the system. As stated above, an equilibrium distribution for the coordinates can be generated by either the microcanonical or Monte Carlo method. The velocities can then be

chosen according to an appropriately constrained Gaussian distribution.

IV. CONCLUSIONS

We have shown that it is possible to cast U(1), SU(2), and SU(3) lattice gauge theories in a form recognizable as the microcanonical ensemble of classical statistical mechanics. Expectation values of operators can be calculated in this new ensemble by solving a large set of coupled ordinary differential equations. This procedure is to be contrasted to the standard method of calculation where Monte Carlo techniques are applied to Euclidean path integrals (which are formally equivalent to a classical *canonical* ensemble).

What are the advantages of this new formulation of lattice gauge theory? Besides the obvious benefits of a new way of looking at a problem, it allows the lattice to be updated in an entirely parallel fashion, such as occurs with an array processor. Typically in a Monte Carlo calculation the lattice cannot be updated in a single pass[4] because each link interacts with its neighbors, and thus a simultaneous update of link and neighbor would violate the principle of detailed balance. The microcanonical-ensemble formulation may thus be useful when the action governing the system is very nonlocal. Of course, when fermions are involved the action involves a determinant and is thus highly nonlocal.[19]

For the three lattice gauge theories we studied with the standard Wilson action, our microcanonical-ensemble method did not give a significant improvement in computation time. It must be pointed out however that the standard Wilson action is a fairly local object and both calculations were performed using a standard serial processor. Also, we did not perform calculations of large loops or other extended objects, where problems of metastability are important [particularly for SU(3)].

It is worth stressing our most interesting result. We have constructed a deterministic alternative to Monte Carlo calculations which is distinct from the usual methods of calculation. The method appears to be at least as generally applicable as Monte Carlo methods and its overall performance is roughly equivalent to standard techniques. For special machines (such as array processors) the microcanonical approach may in fact be better than Monte Carlo methods, due to the fact that it allows the lattice to be updated in an entirely parallel fashion.

Since this research was concluded, we have become aware of other lattice gauge calculations in the microcanonical ensemble.[20] A continuum formulation of quantum field theory in the microcanonical ensemble has also been constructed.[21] Finally it has been suggested[22] that a microcanonical simulation for a finite system may be useful in resolving problems of metastability[23] in the effective potential.

ACKNOWLEDGMENTS

It is a pleasure to thank E. L. Berger for his assistance in instigating this study and in obtaining the necessary computer time for the calculations, and M. Creutz for several useful discussions. This work was performed under the auspices of the United States Department of Energy.

---

*Present address.

[1]K. G. Wilson, Phys. Rev. D 14, 2455 (1974); A. M. Polyakov, Phys. Lett. 59B, 82 (1975); F. Wegner, J. Math. Phys. 12, 2259 (1971).

[2]The use of Monte Carlo methods in lattice gauge theory is reviewed in M. Creutz, Brookhaven National Laboratory Report No. BNL-29840, 1981 (unpublished). See also M. Creutz, L. Jacobs, and C. Rebbi, Phys. Rep. (to be published); J. B. Kogut, Urbana Report No. ILL-(TH)-82-46 (unpublished); R. C. Brower, in Les Houches Lectures, 1981 (unpublished); C. DeTar, Helsinki Report No. HU-TFT-82-28 (unpublished); B. Lautrup, CERN Report No. TH.3521-CERN, (unpublished) 1983; D. J. E. Callaway, Argonne Report No. ANL-HEP-82-57 (unpublished); and Argonne Report No. ANL-HEP-PR-83-04 (unpublished).

[3]D. J. E. Callaway and A. Rahman, Phys. Rev. Lett. 49, 613 (1982).

[4]Of course it is possible to change every link in the lattice and *then* apply the Metropolis (Ref. 12) algorithm to determine whether or not to keep the entire new system. A difficulty with this method is that if one new link is in an unfavorable region of configuration space it can cause the entire new system to be rejected. Thus the tentative updates made must be small and the system may not equilibrate quickly. To our knowledge such a procedure has not yet been analyzed in the literature.

[5]C. Pangali, M. Rao, and B. J. Berne, Chem. Phys. Lett. 55, 413 (1978); M. Rao, C. Pangali, and B. J. Berne, Mol. Phys. 37, 1773 (1979).

[6]The various ensembles are well discussed in, e.g., K. Huang, *Statistical Mechanics* (Wiley, New York, 1963); R. Kubo, *Statistical Mechanics* (North-Holland, New York, 1967); F. Reif, *Fundamentals of Statistical and Thermal Physics* (McGraw-Hill, New York, 1965).

[7]J. L. Lebowitz, J. K. Percus, and L. Verlet, Phys. Rev. 153, 250 (1967).

[8]M. E. Fisher, Arch. Ration. Mech. Anal. 17, 377 (1969); R. Griffith, J. Math. Phys. 6, 1447 (1965); J. Van der Linden, Physica (Utrecht) 32, 642 (1966).

[9]J. L. Lebowitz, in *Transport Phenomena,* edited by G. Kirzenow and J. Marro (Springer, Berlin, 1974); J. L. Lebowitz and O. Penrose, Phys. Today 26, No. 2, 23 (1973).

[10]See, for example, R. Abraham and J. Marsden, *Foundations of Mechanics* (Benjamin, Reading, Mass., 1978).

[11]With a predictor-corrector algorithm (using up to the fifth derivative of ϕ with respect to τ) and $\Delta\tau=0.01$ identical results were obtained on a 4^4-site problem. This choice of $\Delta\tau$ leads to a conservation of $H=T+V$ to better than one part in

10^7. We are now using $\Delta\tau=0.1$ which conserves H to one part in 10^4. The velocity correlation function $\langle p(0)p(\tau)\rangle$ decays to zero at about $\tau=3.0$. Another way of assessing the natural time scale is from the fact that for a 3^4 lattice the squares of the frequencies of the normal modes are 0, 3, 6, 9, and 12. Thus the shortest period is $2\pi/\sqrt{12}\approx 1.8$.

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

[13]D. J. E. Callaway and L. J. Carson, Phys. Rev. D 25, 531 (1982).

[14]M. Creutz, Phys. Rev. D 21, 2308 (1980).

[15]H. C. Andersen, J. Comp. Phys. (to be published).

[16]K. E. Eriksson, N. Svartholm, and B. S. Skagerstam, J. Math. Phys. 22, 2276 (1981); E. Pietarinen, Nucl. Phys. B190 [FS3], 349 (1981).

[17]As the step size h is increased, the errors in the numerical integration scheme first manifest themselves by the violation of the constraints Eqs. (40a) and (40b). We have found that by calculating the λ multipliers to order h^4 [rather than to order h^2 as in Eq. (47)] a larger step size of $h=0.1$ can safely be used.

[18]The Monte Carlo program we used for this comparison was written by M. Creutz. Tentative new links are generated by multiplying the old link by a matrix A. The matrix A is weighted toward the identity in an empirical fashion depending on β and the neighboring links. Each tentative new link is accepted or rejected according to the standard Metropolis (Ref. 12) algorithm. The update procedure is applied to each link 20 times before moving to the next link.

[19]The possibility that the microcanonical method may be advantageous when fermions are included was suggested in Ref. 3. Preliminary calculations are consistent with this hypothesis, however, a detailed analysis of this problem has not been carried out.

[20]M. Creutz, Phys. Rev. Lett. 50, 1411 (1983).

[21]A. Strominger, Ann. Phys. 146, 419 (1983).

[22]U. M. Heller and N. Seiberg, Phys. Rev. D 27, 2980 (1983).

[23]D. J. E. Callaway and D. J. Maloof, Phys. Rev. D 27, 406 (1983); D. J. E. Callaway, *ibid.* 27, 2974 (1983).