Nuclear Physics B 380 (1992) 665–677
North-Holland

NUCLEAR
PHYSICS B

Hamiltonian evolution for the hybrid Monte Carlo algorithm

J.C. Sexton * and D.H. Weingarten
IBM Research, P.O. Box 218, Yorktown Heights, NY 10598, USA

Received 10 February 1992
Accepted for publication 16 March 1992

We discuss a class of reversible, discrete approximations to Hamilton's equations for use in the hybrid Monte Carlo algorithm and derive an asymptotic formula for the step-size-dependent errors arising from this family of approximations. For lattice QCD with Wilson fermions, we construct several different updates in which the effect of fermion vacuum polarization is given a longer time step than the gauge field's self-interaction. On a 4^4 lattice, one of these algorithms with an optimal choice of step size is 30% to 40% faster than the standard leapfrog update with an optimal step size.

1. Introduction

The hybrid Monte Carlo algorithm [1] is probably the most efficient method now available for evaluating the QCD path integral including the full effect of fermion vacuum polarization [2–5]. Nearly all of the numerical work required by hybrid Monte Carlo is spent solving Hamilton's equations for a classical mechanical system. In the present article, we consider a class of reversible, discrete approximations to Hamilton's equations for use in the hybrid Monte Carlo algorithm, and give a systematic procedure for evaluating the time-step-dependent errors in these algorithms. For some theories, but not QCD, we show that the effect of these errors is equivalent to replacing the true hamiltonian with an effective hamiltonian, differing by powers of the time step.

For lattice QCD, we then construct an algorithm with a long time step for an update stage which takes account of the force on the gauge fields arising from fermion vacuum polarization and a short time step for an update stage which handles the self-interaction of the gauge field. Since the number of arithmetic operations required to find the self-interaction of the gauge fields is much smaller than the number required to find the effect of fermion vacuum polarization, this algorithm requires almost the same amount of work as the simplest leapfrog

* Permanent address: Department of Mathematics, Trinity College, Dublin 2, Ireland.

algorithm [2,6] with the same time step for fermion vacuum polarization and the gauge field self-interaction. As a consequence of the smaller step size used for the gauge self-interaction, however, the error in the method we discuss is smaller than the error for leapfrog. We perform a numerical test of versions of this algorithm, leapfrog, and several others for QCD with two flavors of Wilson fermions on a 4^4 lattice. Our best strategy runs 30% to 40% faster than the other methods considered.

An algorithm related to ours, but lacking reversibility, was developed recently for use in molecular dynamics simulations [7]. A reversible version of the molecular dynamics calculation, incorporating the method described here, is discussed in ref. [8].

2. Hybrid Monte Carlo

We will begin by briefly summarizing the hybrid Monte Carlo algorithm. We wish to generate an ensemble of configurations of a set of coordinates $q = (q_1 \dots q_n)$ distributed according to

$$\mathrm{d}P(q) = Z^{-1} \exp[-S(q)]\, \mathrm{d}\mu(q), \tag{2.1}$$

$$Z = \int \exp[-S(q)]\, \mathrm{d}\mu(q), \tag{2.2}$$

where $\mathrm{d}\mu(q)$ is some measure on q. For this purpose we introduce a momentum vector $p = (p_1 \dots p_n)$ and a classical hamiltonian

$$H = \tfrac{1}{2} \sum_i p_i^2 + S(q). \tag{2.3}$$

A correctly distributed sequence of coordinates q is generated by choosing an arbitrary starting q and p, then successively updating these by a three-step process. First, each component p_i of p is chosen to be a new, independent, gaussian random variable with dispersion 1. Second, the momenta and coordinates (p, q) are evolved according to the hamiltonian H over a time interval t by taking $t/\Delta t$ repetitions of a discrete approximation to Hamilton's equations for a short time interval Δt

$$T(\Delta t): (p, q) \mapsto (p', q'). \tag{2.4}$$

This step is required to be reversible in the sense that

$$T(-\Delta t): (p', q') \mapsto (p, q), \tag{2.5}$$

and it must preserve volume

$$\mathrm{d}\mu(p')\, \mathrm{d}\mu(q') = \mathrm{d}\mu(p)\, \mathrm{d}\mu(q), \tag{2.6}$$

where $d\mu(p)$ is $\Pi_i \, dp_i$. Finally, in the third stage of an update, the new momenta and coordinates (p', q') are put in place of (p, q) with probability

$$\rho = \min\{1, \exp[H(p', q') - H(p, q)]\}. \tag{2.7}$$

To the extent that the update $T(\Delta t)$ is a good approximation to time evolution by the hamiltonian H, the value of H will be preserved and the acceptance rate ρ will be close to 1. Thus an efficient implementation of the hybrid Monte Carlo algorithm depends on an update $T(\Delta t)$ which minimizes the arithmetic work required to preserved H accurately over the interval t.

3. Hamiltonian time evolution

We will now describe a class of updates $T(\Delta t)$. For this purpose, define $\mathscr{E}$ to be the vector-space of functions $f(p, q)$ on phase space (p, q), and let the linear operator, $L(f)$, be given, by the Poisson bracket

$$\begin{aligned} L(f)g &= -\{f, g\} \\ &= \sum_i \left(\frac{\partial f}{\partial p_i} \partial_i g - \partial_i f \frac{\partial g}{\partial p_i} \right). \end{aligned} \tag{3.1}$$

If the coordinate space containing q is flat, the derivative ∂_i is

$$\partial_i = \frac{\partial}{\partial q_i}. \tag{3.2}$$

A slightly more complicated choice of ∂_i needed for QCD, in which the coordinate space is a product of (curved) SU(3) manifolds, will be reviewed later on.

For any $f(t) = f[p(t), q(t)]$, Hamilton's equations then become

$$\frac{df(t)}{dt} = L(H)f. \tag{3.3}$$

Thus we have

$$f(t) = \exp[tL(H)]f(0). \tag{3.4}$$

Now the exponential in eq. (3.4) can be expressed as $t/\Delta t$ iterations of

$$\exp[\Delta t \, L(H)]. \tag{3.5}$$

An approximation to this term can be expressed as a product

$$\exp[\Delta t\, L(H) + \epsilon] = \exp[\Delta t\, L(h_n)] \times \ldots \times \exp[\Delta t\, L(h_2)] \times \exp[\Delta t\, L(h_1)] \times \exp[\Delta t\, L(h_2)] \times \ldots \times \exp[\Delta t\, L(h_n)]. \quad (3.6)$$

for any decomposition of H

$$H = h_1 + 2h_2 + \ldots + 2h_n, \quad (3.7)$$

with an error ϵ which, we will show, is at most $O(\Delta t^3)$. As a consequence of the symmetric shape of the product in eq. (3.6), the reversibility condition, eq. (2.5), is fulfilled. It follows from eq. (3.1), that the volume preservation condition, eq. (2.6), is fulfilled separately by each step of the product in (3.6) and, therefore, by the product as a whole.

The utility of eq. (3.6) is that for convenient choices of the decomposition (3.7) the action of each term in eq. (3.6) has a simple closed form which can be implemented numerically. In particular, for any $f(p, q)$ in $\mathscr{E}$

$$\exp\left[\Delta t\, L\left(\tfrac{1}{2}\sum_i p_i^2\right)\right] f(p, q) = f(p, q + \Delta t\, p), \quad (3.8)$$

$$\exp\{\Delta t\, L[S(q)]\} f(p, q) = f(p_i - \Delta t\, \partial_i S, q). \quad (3.9)$$

The standard leapfrog [2,6,9] algorithm is given by splitting up eq. (2.3) as

$$h_1 = \tfrac{1}{2}\sum_i p_i^2, \quad (3.10)$$

$$h_2 = \tfrac{1}{2} S(q). \quad (3.11)$$

Applying eqs. (3.1), (3.8) and (3.9) to eqs. (3.10), (3.11), to find the effect of a leapfrog update on the momentum and cordinate vectors (p, q) yields a more familiar expression for a leapfrog update,

$$p_i' = p_i - \tfrac{1}{2}\Delta t\, \partial_i S(q) \quad (3.12)$$

$$q_i' = q_i + \Delta t\, p_i' \quad (3.13)$$

$$p_i'' = p_i' - \tfrac{1}{2}\Delta t\, \partial_i S(q'). \quad (3.14)$$

4. Error estimates

We now consider the error term on the left of eq. (3.6). For the simplest case, the Campbell–Baker–Hausdorff formula gives the expansion

$$\exp[\Delta t\, L(h_2)]\, \exp[\Delta t\, L(h_1)]\, \exp[\Delta t\, L(h_2)] = \exp[\Delta t\, L(h_1) + 2\Delta t\, L(h_2) + \epsilon]$$

$$\epsilon = \sum_{n=1}^{\infty} (\Delta t)^{2n+1} \epsilon_{2n+1}, \tag{4.1}$$

where the leading error term is

$$\epsilon_3 = -\tfrac{1}{6}[L(h_2), [L(h_2), L(h_1)]] - \tfrac{1}{6}[L(h_1), [L(h_2), L(h_1)]]. \tag{4.2}$$

By iteratively applying eqs. (4.1), (4.2), to eq. (3.6), working outward from the center of the product on the right-side of eq. (3.6), a general expression can be found for the error ϵ_3. We obtain

$$\epsilon_3 = -\tfrac{1}{6}(A + B),$$

$$A = \sum_{i>1} \{[L(h_i), [L(h_i), L(h_1)]] + [L(h_1), [L(h_i), L(h_1)]]\},$$

$$B = \sum_{i>j} \{2[L(h_i), [L(h_i), L(h_j)]] + 4[L(h_j), [L(h_i), L(h_j)]]\}. \tag{4.3}$$

Similar but more complicated expressions can be derived for the higher terms, ϵ_5, $\epsilon_7, \ldots$ contributing to ϵ in eq. (3.6).

Eqs. (4.2) and (4.3) have an interesting consequence for systems in which the derivative ∂_i takes the simple form given by eq. (3.2). It then follows that the algebra of commutators of $L(h)$ is isomorphic to the algebra of Poisson brackets

$$[L(g), L(h)] = L(\{g, h\}). \tag{4.4}$$

this equation, in turn, implies that discrete evolution by the product on the right side of eq. (3.6) is equivalent, to $O(\Delta t^3)$, to continues evolution with the effective hamiltonian

$$H_{\text{eff}} = H - \tfrac{1}{6}\Delta t^2 (a + b), \tag{4.5}$$

$$a = \sum_{i>1} [\{h_i, \{h_i, h_1\}\} + \{h_1, \{h_i, h_1\}\}] \tag{4.6}$$

$$b = \sum_{i>j} [2\{h_i, \{h_i, h_j\}\} + 4\{h_j, \{h_i, h_j\}\}]. \tag{4.7}$$

Corresponding expressions can be derived including the effects of the higher error terms, ϵ_5, ϵ_7, ... contributing to ϵ in eq. (3.6).

5. Multiple time scales

Eqs. (3.6) and (4.3) can be used to construct an improvement of leapfrog evolution in which different terms in the classical hamiltonian contribute to time evolution with time steps of different sizes. Consider a hamiltonian

$$H = \tfrac{1}{2}\sum_i p_i^2 + S_1(q) + S_2(q), \tag{5.1}$$

for which the number of arithmetic operations required to evaluate S_1 is much smaller than the number of operations required for S_2. Let the partial time evolution operators $T_1(\Delta t)$ and $T_2(\Delta t)$ be

$$T_1(\Delta t) = \exp\left[\tfrac{1}{2}\Delta t\, L(S_1)\right] \exp\left[\Delta t\, L\left(\tfrac{1}{2}\sum_i p_i^2\right)\right] \exp\left[\tfrac{1}{2}\Delta t\, L(S_1)\right], \tag{5.2}$$

$$T_2(\Delta t) = \exp\left[\Delta t\, L(S_2)\right]. \tag{5.3}$$

For any integer n we can then define a full time-evolution operator

$$T(\Delta t) = T_2\left(\frac{\Delta t}{2}\right)\left[T_1\left(\frac{\Delta t}{n}\right)\right]^n T_2\left(\frac{\Delta t}{2}\right) \tag{5.4}$$

$$= \exp(\Delta t\, H + \epsilon). \tag{5.5}$$

The coefficient of the leading contribution to ϵ in eq. (5.5) becomes

$$\begin{aligned}\epsilon_3 = &-\frac{1}{24n^2}\left[L(V_1), \left[L(V_1), L\left(\tfrac{1}{2}\sum_i p_i^2\right)\right]\right] \\ &-\frac{1}{12n^2}\left[L\left(\tfrac{1}{2}\sum_i p_i^2\right), \left[L(V_1), L\left(\tfrac{1}{2}\sum_i p_i^2\right)\right]\right] \\ &-\tfrac{1}{24}\left[L(V_2), \left[L(V_2), L\left(\tfrac{1}{2}\sum_i p_i^2\right)\right]\right] - \tfrac{1}{12}\left[L\left(\tfrac{1}{2}\sum_i p_i^2\right), \left[L(V_2), L\left(\tfrac{1}{2}\sum_i p_i^2\right)\right]\right].\end{aligned} \tag{5.6}$$

Higher contributions to ϵ which involve $L(V_1)$ are also suppressed by powers of n.

In the limit $n \to \infty$, the error in eq. (5.5) involves solely commutators of $L(V_2)$ and $L(\frac{1}{2}\Sigma_i p_i^2)$.

The leapfrog update of eqs. (3.12), (3.13), (3.14), for the hamiltonian (5.1) is the same as eq. (5.4) for $n = 1$. However since the main cost of carrying out a leapfrog update for this hamiltonian is the evaluation of the effect of V_2, eq. (5.4) with $n \gg 1$ is not appreciably more expensive than leapfrog. Since the error given by eq. (5.6) falls as n is increased, eq. (5.4) with $n \gg 1$ will produce a more efficient overall algorithm than leapfrog.

6. Error cancellation

More complicated versions of $T(\Delta t)$ can be defined with smaller error terms. Translated to the present framework, ref. [9] gives a method for choosing h_1, $h_2, \ldots, h_n$, in eq. (3.6) to eliminate any number of terms in the Δt expansion of ϵ. By eq. (3.6), we have

$$\begin{aligned}\exp\{\Delta t[L(h_1)+L(h_2)]+\epsilon\} &= \exp[\tfrac{1}{2}sL(h_2)]\,\exp[sL(h_1)]\,\exp[\tfrac{1}{2}(s+t)L(h_2)] \\ &\quad\times\exp[tL(h_1)] \\ &\quad\times\exp[\tfrac{1}{2}(s+t)L(h_2)]\,\exp[sL(h_1)]\,\exp[\tfrac{1}{2}sL(h_2)] \\ s &= \frac{\Delta t}{2-2^{1/3}}, \\ t &= \frac{-2^{1/3}\Delta t}{2-2^{1/3}}. \end{aligned} \tag{6.1}$$

Then according to eq. (4.3), the $O(\Delta t^3)$ contribution to ϵ vanishes. Adapting this results to $T(\Delta t)$, we obtain the time evolution

$$\begin{aligned} X(\Delta t) &= T_2\left(\frac{s}{2}\right)\left[T_1\left(\frac{s}{n}\right)\right]^n T_2\left(\frac{(s+t)}{2}\right) \\ &\quad\times\left[T_1\left(\frac{t}{n}\right)\right]^n \\ &\quad\times T_2\left(\frac{(s+t)}{2}\right)\left[T_1\left(\frac{s}{n}\right)\right]^n T_2\left(\frac{s}{2}\right), \end{aligned} \tag{6.2}$$

with s and t defined in eq. (6.1). By eq. (4.2) and definition (5.2), we have for the terms $T_1(u)$ in (6.2)

$$[T_1(u)]^n = \exp\left\{u\left[L\left(\tfrac{1}{2}\sum_i^i p_i^2\right) + L(V_1)\right] + O(n^{-2})\right\}. \tag{6.3}$$

Substituting eq. (6.3) in eq. (6.2), we obtain for $X(\Delta t)$ an expression of form (6.1). Thus the leading error in $X(\Delta t)$ is of order Δt^5 and, up to terms $O(n^{-2})$, involves only $L(V_2)$ and $L(\frac{1}{2}\Sigma_i p_i^2)$.

A slightly simpler alternative to $X(\Delta t)$ can be found which eliminates some, but not all, of the Δt^3 errors in $T(\Delta t)$. By eq.. (3.6), we have

$$\begin{aligned}\exp\{\Delta t[L(h_1) + L(h_2)] + \epsilon\} = {} & \exp[\tfrac{1}{6}\Delta t\, L(h_2)]\, \exp[\tfrac{1}{2}\Delta t\, L(h_1)] \\ & \times \exp[\tfrac{2}{3}\Delta t\, L(h_2)] \\ & \times \exp[\tfrac{1}{2}\Delta t\, L(h_1)]\, \exp[\tfrac{1}{6}\Delta t\, L(h_2)].\end{aligned} \tag{6.4}$$

Eq. (4.3) then gives for the leading contribution to ϵ a single term, in place of two terms for leapfrog,

$$\epsilon_3 = \tfrac{1}{72}\Delta t^3[L(h_2), [L(h_2), L(h_1)]]. \tag{6.5}$$

An evolution of this form in the absence of S_2 is

$$\begin{aligned}T_3(\Delta t) = {} & \exp[\tfrac{1}{6}\Delta t\, L(S_1)]\, \exp\left[\tfrac{1}{2}\Delta t\, L\left(\tfrac{1}{2}\sum_i p_i^2\right)\right] \\ & \times \exp[\tfrac{2}{3}\Delta t\, L(S_1)] \\ & \times \exp\left[\tfrac{1}{2}\Delta t\, L\left(\tfrac{1}{2}\sum_i p_i^2\right)\right] \exp[\tfrac{1}{2}\Delta t\, L(S_1)].\end{aligned} \tag{6.6}$$

Then one possible adaptation of eq. (6.4) to include S_2 gives the evolution

$$\begin{aligned}Y(\Delta t) = {} & T_2(\tfrac{1}{6}\Delta t)\left[T_3\left(\frac{\Delta t}{2n}\right)\right]^n \\ & \times T_2(\tfrac{2}{3}\Delta t) \\ & \times \left[T_3\left(\frac{\Delta t}{2n}\right)\right]^n T_2(\tfrac{1}{6}\Delta t).\end{aligned} \tag{6.7}$$

The leading error in $Y(\Delta t)$, up to terms $O(n^{-2})$, involves only a single commutator of $L(V_2)$ and $L(\frac{1}{2}\sum_i p_i^2)$.

7. QCD

We will now compare the efficiencies of the methods discussed in sects. 5 and 6 for QCD with two flavors of Wilson fermions. The theory is defined on a finite, 4-dimensional hypercubic lattice Λ. Each site x in Λ carriers the link matrices $U_\mu(x)$ in SU(3), the (canonically conjugate) real-valued link-momenta $P_{\mu a}(x)$, with link index $\mu = 1,\ldots,4$, and component index $a = 1,\ldots,8$, and a complex valued pseudofermion vector $\phi_{ic}(x)$, with spin index $i = 1,\ldots,4$, and color index $c = 1,\ldots,3$. The measure according to which these fields are distributed is

$$dP(U, P, \phi) = Z^{-1}\, d\mu(U)\, d\mu(P)\, d\mu(\phi)\, \exp[-H(U, P, \phi)], \tag{7.1}$$

$$Z = \int d\mu(U) \int d\mu(P) \int d\mu(\phi)\, \exp[-H(U, P, \phi)], \tag{7.2}$$

where $\int d\mu(U)\ldots$ is integration over one copy of SU(3) Haar measures for each link matrix, $\int d\mu(P)\ldots$ is integration over each component of $P_{\mu a}(x)$ at each site, and $\int d\mu(\phi)\ldots$ is integration over the real and imaginary coponents of each $\phi_{ia}(x)$ at each lattice site x. The hamiltonian $H(P, U, \phi)$ in eqs. (7.1) and (7.2) is a sum of the three terms [10–12]

$$K(P) = \tfrac{1}{2} \sum_{\mu a x} \left[P_{\mu a}(x)\right]^2, \tag{7.3}$$

$$S(U) = -\tfrac{1}{6}\beta \sum_p \left[\mathrm{Tr}\ U(p) + \mathrm{Tr}\ U^\dagger(p)\right], \tag{7.4}$$

$$S(\phi, U) = \tfrac{1}{2}\left(M_U^{-1}\phi,\ M_U^{-1}\phi\right), \tag{7.5}$$

where, as usual, $U(p)$ is the product of the link fields around a plaquette p, M_U is Wilson's fermion coupling matrix for the link field U, and the inner-product $(\ldots,\ldots)$ is defined in the natural way.

To generate fields distributed according to $dP(U, P, \phi)$ we use the algorithm of sect. 2 augumented with an additional step accompanying the update of the momentum field P. Each time a new field P is chosen, a new ϕ is chosen according to the gaussian differential probability

$$\exp[-S(\phi, U)]\, d\mu(\phi). \tag{7.6}$$

This is accomplished [13] by selecting the components of a field χ to be indepen-

dent gaussian random variables, then setting ϕ to be $M_U\chi$. For QCD the definition of the derivative $\partial_{a\mu x}$ used in eq. (3.1) becomes

$$\partial_{a\mu x} f\left[U_\mu(x)\right] = \frac{\partial}{\partial\alpha} f\left[(1 + i\alpha\lambda_a)U_\mu(x)\right]\Big|_{\alpha=0}, \tag{7.7}$$

where the λ_a are an orthonormal set of 3×3 traceless, hermitian matrices.

To solve Hamilton's equations for QCD, we will use the methods of sects. 5 and 6. Essentially all of the work required by any of these methods will go to the evaluation of the derivatives $\partial_{a\mu x} S(\phi, U)$ which requires solution of the linear equation

$$M_U^\dagger M_U \psi = M_U^\dagger \phi \tag{7.8}$$

for a specified ϕ. Thus we will treat $S(U)$ of eq. (7.4) as S_1 of eq. (5.1) and $S(\phi, U)$ of eq. (7.5) as S_2 of eq. (5.1). Then for a specified classical trajectory length t, if N is the number of times eq. (7.8) must be solved and A is the mean acceptance rate of the acceptance step governed by eq. (2.7), the relative efficiency of any algorithm is measured by N/A, the amount of work required to obtain an accepted classical trajectory.

To compare the various algorithms we chose, somewhat arbitrarily, to work on a $4 \times 4 \times 4 \times 4$ lattice at β of 5.4 with a Wilson hopping constant k of 0.162. For a trajectory length t of 1, Table 1 shows the values of N, A and N/A we obtained for various values of the step sizes Δt for methods using the same step size for the

TABLE 1
Work required for trajectories of length 1.0 for three algorithms A, B and C using the same step size for the effects fermion vacuum polarization and the gauge field's self-interaction

Method	Step	N	A	N/A
A	0.1667	7	0.000316 ± 0.000166	22120.0
	0.1250	9	0.111176 ± 0.033186	80.95303
	0.1000	11	0.264042 ± 0.039531	41.66000
	0.0833	13	0.433904 ± 0.040972	29.96056
	0.0714	15	0.565347 ± 0.038428	26.53239
	0.0625	17	0.659094 ± 0.034202	25.79297
	0.0556	19	0.727106 ± 0.029862	26.13099
	0.0500	21	0.777206 ± 0.025907	27.01985
B	0.2500	9	0.194586 ± 0.038466	46.25210
	0.2000	11	0.597320 ± 0.042293	18.41558
	0.1667	13	0.741039 ± 0.033409	17.54294
	0.1429	15	0.816117 ± 0.026132	18.37972
	0.1250	17	0.861291 ± 0.020742	19.73782
	0.1111	19	0.891467 ± 0.016783	21.31318
	0.1000	21	0.912652 ± 0.013807	23.00986
C	0.1667	19	0.192252 ± 0.034066	98.82845

TABLE 2
Work required for trajectories of length 1.0 for three algorithms A′, B′ and C′ using different time steps for the effects of fermion vacuum polarization and the gauge field's self-interaction

Method	Step	N	A	N/A
A′	0.1667	7	0.133879 ± 0.028851	52.28599
	0.1250	9	0.454113 ± 0.043459	19.81887
	0.1000	11	0.631497 ± 0.037635	17.41892
	0.0833	13	0.738657 ± 0.030215	17.59950
	0.0714	15	0.806899 ± 0.024085	18.58969
	0.0625	17	0.851782 ± 0.019345	19.95815
	0.0556	19	0.882743 ± 0.015755	21.52328
	0.0500	21	0.904950 ± 0.013023	23.20570
B′	0.2500	9	0.246264 ± 0.041398	36.54611
	0.2000	11	0.737808 ± 0.034058	14.90903
	0.1667	13	0.845624 ± 0.021004	15.37326
	0.1429	15	0.902610 ± 0.014145	16.61847
	0.1250	17	0.930868 ± 0.010574	18.26253
	0.1111	19	0.947835 ± 0.008237	20.04567
	0.1000	21	0.959044 ± 0.006622	21.89680
C′	0.1667	19	0.409866 ± 0.042922	46.35657

effects of fermion vacuum polarization and the gauge field's self-interaction. Method A here is the simple leapfrog algorithm specified by eqs. (3.10) and (3.11). Method B is the algorithm of eq. (6.4) in which some, but not all of the Δt^3 errors of leapfrog are cancelled, and method C is the algorithm of eq. (6.1) in which the entire Δt^3 error has been removed. Table 2 shows the values of N, A and N/A we obtained for various values of the step sizes Δt for methods using different step sizes for the effects of fermion vacuum polarization and the gauge field's self-interaction. Methods A′, B′ and C′ are our improved versions of algorithm A, B and C, specified respectively by eqs. (5.4), (6.7) and (6.2) with $n = 4$. For algorithms A, A′, B and B′, we attempted to find the step size Δt which produced the minimum value of N/A. For C and C′ we tried progressively smaller step sizes until it became clear that even if all trajectories were accepted, giving an A of 1.0, the minimum value of N/A would have to be greater than the minimum found for the other methods. The optimal algorithm, according to tables 1 and 2 is thus B′. The relative cost of a trajectory for this method appears is 40% lower than the cost of a trajectory using method A with an optimal step size. Table 3 shows the work for optimal choses of step size as a function of trajectory length. It is clear that in all cases either method A′ or B′ is optimal with A′ never more than a fraction of a percent ahead of B′. The faster of the two, on the other hand, remains consistently more than 30% faster than A.

Algorithms A′, B′ and C′ have also been tested with $n = 8$. The results do not differ significantly from those shown already in tables 1, 2 and 3.

TABLE 3
Minimal work required for a range of algorithms and trajectory lengths

Method	t	Step	N	A	N/A
A	0.25	0.0625	5	0.635451 ± 0.039878	7.868424
B	0.25	0.125	5	0.869170 ± 0.019644	5.752613
A′	0.25	0.0625	5	0.830576 ± 0.024124	6.019921
B′	0.25	0.125	5	0.930328 ± 0.009897	5.374449
A	0.50	0.0625	9	0.684592 ± 0.035808	13.14651
B	0.50	0.167	7	0.706477 ± 0.034592	9.908315
A′	0.50	0.0833	7	0.784387 ± 0.030967	8.924169
B′	0.50	0.167	7	0.841543 ± 0.019669	8.318054
A	1.00	0.0625	17	0.659094 ± 0.034202	25.79297
B	1.00	0.167	13	0.741039 ± 0.033409	17.54294
A′	1.00	0.100	11	0.631497 ± 0.037635	17.41892
B′	1.00	0.200	11	0.737808 ± 0.034058	14.90903
A	2.00	0.0714	29	0.652395 ± 0.038712	44.45163
B	2.00	0.167	25	0.734150 ± 0.034655	34.05300
A′	2.00	0.100	21	0.704238 ± 0.033423	29.81946
B′	2.00	0.182	23	0.766292 ± 0.026741	30.01466

All results shown in tables 1, 2 and 3 are, of course, for a rather small lattice with a hopping parameter some distance from critical. An unanswered question, however, is what gain might be obtained for larger lattices and hopping parameters closer to critical. This question is being investigated.

An interesting feature of table 3 is that the work required for long trajectories appears to grow no faster than linearly with the trajectory length. For a theory in which the effect of finite step size could be absorbed into an effective hamiltonian as discussed in sect. 4, this behaviour would not be too surprising. In such a theory, an effective energy, differing from the true energy by $O(\Delta t^2)$, is conserved for any trajectory length. Thus the change in the true energy over a trajectory will be primarily determined by the time step and not by the trajectory length. It follows that the rejection rate for trial steps at the end of a trajectory will also tend not to depend on the trajectory length. Therefore the optimal step size will not depend on trajectory length, and work will grow approximately linearly with trajectory length. Unfortunately, this argument can not be applied directly to QCD since eq. (4.4) does not hold, and an effective hamiltonian can not be found using eq. (4.5).

We are grateful to B.J. Berne for informing us of ref. [7] and for several interesting discussions.

References

[1] S. Duane, A.D. Kennedy, B.J. Pendleton and D. Roweth, Phys. Lett. B195 (1987) 216
[2] M. Creutz, Phys. Rev. D38 (1988) 1228
[3] R. Gupta, G.W. Kilcup and S.R. Sharpe, Phys. Rev. D38 (1988) 1278
[4] D. Weingarten, Nucl. Phys. B (Proc. Suppl.) 9 (1989) 447
[5] R. Gupta, A. Patel, C. Baillie, G. Guralnik, G. Kilcup and S. Sharpe, Phys. Rev. D40 (1989) 2072
[6] S. Gottlieb, W. Liu, D. Toussaint, R.L. Renken and R.L. Sugar, Phys. Rev. D35 (1987) 2531
[7] M. Tuckerman, G. Martyna and B. Berne, J. Chem. Phys. 93 (1990) 1287;
M. Tuckerman, B. Berne and A. Rossi, J. Chem. Phys. 94 (1990) 1465
[8] M. Tuckerman, B. Berne and G. Martyna, to appear in J. Chem. Phys.
[9] M. Creutz and A. Gocksch, Phys. Rev. Lett. 63 (1989) 9
[10] K. Wilson, Phys. Rev. D10 (1974) 2445
[11] D. Weingarten and D. Petcher, Phys. Lett. B99 (1981) 333
[12] D. Callaway and A. Rahman, Phys. Rev. Lett. 49 (1982) 613
[13] G.G. Batrouni, G.R. Katz, A.S. Kronfeld, G.P. Lepage, B. Svetitsky and K.G. Wilson, Phys. Rev. D32 (1985) 2736