ACCELERATED DYNAMICS IN SIMULATIONS OF FIRST ORDER PHASE TRANSITIONS

Ghassan George Batrouni

Physics Department
Boston University
Boston, MA 02215

Benjamin Svetitsky

Center for Theoretical Physics
Laboratory for Nuclear Science
and
Department of Physics
Massachusetts Institute of Technology
Cambridge, MA 02139

Submitted to Physical Review Letters

PACS: 0260.+y, 05.50+q, 64.60.Cn, 64.60.My

CTP #1432
December 1986

ABSTRACT. We show that the method of Fourier acceleration can be used to speed up tunnelling between degenerate and near-degenerate minima of the effective action in Langevin simulations of statistical systems. Acceleration factors of two orders of magnitude are attained. Such improvement is important for the study of weak first-order phase transitions, for simulations of spin glasses, and for the method of simulated annealing in optimization problems.

Fourier acceleration is the name given to a method for speeding up various iterative numerical methods. It was introduced[1,2] as a means to defeat "critical slowing down" in the simulation of Langevin diffusion processes[3], that is, the divergence of autocorrelation times as a critical point is approached. Since then, it has been applied to matrix inversion in the simulation of fermionic effects in quantum chromodynamics[2,4], to the minimization problem posed by gauge fixing[4], to relaxation on a fractal lattice[5], and to microcanonical simulations[6]. In this letter, we demonstrate the usefulness of the method for speeding up tunnelling[7] between degenerate and near-degenerate ground states. While our focus is on the behavior of a statistical system near a first-order transition, it will be apparent that studies of a variety of complex statistical systems will benefit from adaptations of the technique.

Two related problems are frequently faced in the simulation of such systems. One is, how does one evade local minima of the effective Hamiltonian and find the global minimum? This, of course, is the central issue when one employs simulated annealing in the solution of optimization problems[8]. The other problem is, how does one efficiently sample the states close in energy to that global minimum, *including* near-degenerate local minima separated from it by potential barriers? This is the problem plaguing simulations of disordered systems such as spin glasses[9]; it is closely connected to the large dynamical exponents characteristic of these systems, which cause critical slowing down to be particularly severe.

Both problems appear in the numerical study of first-order phase transitions[10]. Consider the two methods often used to pin down the location of weak transitions. In the hysteresis method, one alternately heats and cools the system past the transition point and watches for tunnelling from the metastable phase to the stable one; the more easily such tunnelling may be brought about, the smaller the hysteresis loop, and the higher the precision of the determination of the transition temperature. In cases where the gap is too small to observe directly in the fluctuations, one studies instead the distribution of some field as the system evolves at constant temperature. A histogram of this distribution will show two peaks, one for each phase, and the relative population of these peaks indicates which side of the transition the system is in. Here it is essential to have rapid tunnelling between the two phases, since only thus does one build up the statistics necessary for reliable comparison of the populations[11]. The solution to both these problems lies in the acceleration of tunnelling through potential barriers.

We now describe the method of Fourier acceleration as applied to the Langevin equation. In the simplest Langevin process, the evolution of a scalar field ϕ in "simulation time" τ is described by

$$\frac{\partial\phi(x,\tau)}{\partial\tau} = -\frac{\delta S[\phi]}{\delta\phi(x,\tau)} + \eta(x,\tau). \tag{1}$$

S is the action of a d-dimensional field theory, and the Gaussian noise η satisfies

$$\langle\eta(x,\tau)\eta(x',\tau')\rangle_\eta = 2\,\delta^d(x'-x)\,\delta(\tau'-\tau). \tag{2}$$

The probability distribution $P(\{\phi\},\tau)$ approaches as $\tau \to \infty$ the equilibrium distribution $P_0 \propto e^{-S}$. This approach is governed by the Fokker-Planck equation[12], the solution to which is in general an eigenfunction expansion

$$P(\{\phi\},\tau) \propto P_0(\{\phi\}) + \sum_i C_i P_i(\{\phi\}) e^{-\epsilon_i\tau} \tag{3}$$

where $\epsilon_i > 0$, and the coefficients C_i depend on the initial conditions. The eigenvalues ϵ_i determine both the rate of approach to equilibrium and the autocorrelation times of the equilibrium fluctuations. It was argued in Ref. 2 that the lowest eigenvalues correspond to fluctuations in the infrared modes of the theory, and hence these modes have the longest autocorrelation times. Near a critical point, the lowest ϵ_i approach zero and cause critical slowing down.

There is an infinite number of stochastic processes which lead to the same stationary distribution $P_0 \propto e^{-S}$. In general, these processes will possess different Fokker-Planck spectra. In particular, equation (1) may be generalized to

$$\frac{\partial\phi(x,\tau)}{\partial\tau} = \int dy \left\{ -Q_{xy}\frac{\delta S[\phi]}{\delta\phi(y,\tau)} + Q_{xy}^{1/2}\eta(y,\tau) \right\}, \tag{4}$$

where Q_{xy} is a positive definite matrix. The eigenvalues and eigenfunctions in (3) are now dependent on the choice one makes for Q_{xy} (with the exception of P_0, corresponding to $\epsilon_0 = 0$, which is guaranteed to be unchanged). Fourier acceleration consists of setting $Q_{xy} \propto \Delta(x-y)$, where Δ is the propagator corresponding to the lowest-mass particle in the theory. (The kernel Q is local in momentum space, so the convolutions in (4) are evaluated using Fourier transforms.) It is easy to show in perturbation theory that the autocorrelation time then becomes independent of the momentum scale; effectively, the Fokker-Planck spectrum has been modified to speed up the evolution of the infrared modes. Since all features of the time evolution of the theory are governed by (3), it follows that equilibration times, autocorrelation times, and tunnelling times for the infrared modes will all be reduced.

The theory we have studied is the scalar field theory in two dimensions with the Euclidean action

$$S[\phi] = \int d^2x \left[\frac{1}{2}(\partial_\mu \phi)^2 + \frac{\lambda}{4!}(\phi^2 - \phi_0^2)^2 + h\phi \right]. \tag{5}$$

For ϕ_0 greater than some critical value $\phi_0^{\mathrm{cr}}(\lambda)$, the symmetry $\phi \to -\phi$ is spontaneously broken. In this regime, there is a first order transition as h is varied through zero at constant λ and ϕ_0, with $\lim_{h\to 0^{\pm}} \langle\phi\rangle = \mp\bar{\phi}(\lambda, \phi_0)$. At $h = 0$, the curvature at either minimum of the classical potential gives the tree level mass via $m^2 = \lambda\phi_0^2/6$, and this is the mass we use in $\Delta(x)$ for Fourier acceleration.

We discretize the continuous process (4) via the simple Euler method, with the index i replacing τ:

$$\phi_{i+1}(x) - \phi_i(x) = \sum_y \left\{ -d\tau\, Q_{xy} \frac{\delta S}{\delta\phi_i(y)} + \sqrt{d\tau}\, Q_{xy}^{1/2} \eta_i(y) \right\}, \tag{6}$$

where the Gaussian random variables $\eta_i(x)$ are normalized according to

$$\langle \eta_i(x)\eta_j(y) \rangle = 2\,\delta_{ij}\,\delta_{xy}\,. \tag{7}$$

The kernel for Fourier acceleration is

$$Q_{xy} = \frac{1}{N^2} \sum_{p_x,p_y=-N/2}^{N/2-1} e^{\frac{2\pi i}{N}\vec{\mathrm{p}}\cdot(\vec{\mathrm{x}}-\vec{\mathrm{y}})} \frac{8+m^2}{\left(4 - 2\cos\frac{2\pi}{N}p_x - 2\cos\frac{2\pi}{N}p_y\right) + m^2}\,, \tag{8}$$

i. e., the lattice propagator on an $N \times N$ lattice, normalized so that $\lim_{m\to\infty} Q_{xy} = \delta_{xy}$ and so that the minimum value of the summand is 1. For both the accelerated and the unaccelerated algorithm, a finite $d\tau$ introduces a systematic error[2] which may be characterized as a correction to the effective action, so that $P_0 \propto \exp(-S_{\text{eff}})$, with $S_{\text{eff}} = S + d\tau\, S_1 + (d\tau)^2 S_2 + \cdots$. In order to compare the two algorithms, we shifted ϕ_0 in order to eliminate[2] the $O(d\tau)$ correction. Thus runs with equal parameters should represent comparable physics. (Detailed comparison of observables requires rescaling of ϕ as well, but this is irrelevant to our purposes.)

We simulated the theory on a lattice with 16×16 sites, setting $d\tau = 0.1$. We show in Fig. 1a the time evolution of the volume-averaged field when $h = 0$, in the accelerated method. Note the frequent tunnellings between the $\phi < 0$ and the $\phi > 0$ vacua, which allow rapid accumulation of the histogram in Fig. 1b. The tunnelling rate in this run of 3×10^4 iterations is seen to be dramatically large when juxtaposed with the fact that the corresponding unaccelerated run tunneled only once in 1.8×10^5 iterations.

We show in Fig. 2a the time evolution for the unaccelerated method in a case where it did tunnel in a reasonable length of time. The contrast with the accelerated run in Fig. 2b is striking. The accelerated run, extended to 3×10^4 iterations, allows accumulation of a balanced histogram like that shown in Fig. 1b, which makes it clear that $h = 0$ is indeed the site of the transition; the unaccelerated run gives a messy, unbalanced histogram which would leave considerable room for doubt. We have found that the mean time between tunnelling events can be estimated consistently by taking half the autocorrelation time of $\langle\phi\rangle$. By this measure, Fourier acceleration speeds up tunnelling by two orders of magnitude for the cases we have studied.

Finally, we show in Fig. 3 the evolution for the theory with non-degenerate minima, simulated with Fourier acceleration, and the corresponding histogram. The initial condition for the run was $\phi = +\phi_0$, which is near the false minimum. Two features should be noted: (1) The initial tunnelling to the true vacuum was immediate. A hysteresis loop calculated with this algorithm would give the location of the transition at $h = 0$ with high precision. By contrast, for this choice of parameters the unaccelerated simulation kept the system in the metastable state for 3×10^4 iterations without tunnelling; the corresponding hysteresis loop would be large. (2) Tunnelling back up to the false minimum was fairly frequent, so that one could begin to accumulate a realistic histogram. This is clearly impractical with the unaccelerated algorithm.

The application of this method to some simulated annealing problems is straightforward. One multiplies the Gaussian noise in (4) by a parameter T. The equilibrium probability is then $P_0 \propto e^{-S/T}$, showing that T acts as a temperature. In simulated annealing, one seeks the global minimum of S by repeatedly lowering T, hoping that a sufficiently slow variation of T will allow the system to tunnel out of "false" minima. The speed with which the accelerated algorithm finds the true minimum means that a greatly shortened annealing schedule can be adopted.

A complete comparison of the accelerated with the unaccelerated algorithm would include variation of $d\tau$. One condition that should be imposed is that systematic errors in the two algorithms be the same. As mentioned above, we have shifted ϕ_0 to eliminate the $O(d\tau)$ term in S_{eff}, but $O(d\tau^2)$ terms remain. In some problems, moreover, it may be impossible even to eliminate the $O(d\tau)$ term. The correction terms all explicitly depend on Q_{xy}, so that it is conceivable that to reach a given accuracy, the accelerated algorithm would require a smaller $d\tau$ than the unaccelerated one. However, we have not seen a case in which this would outweigh the factor of ~ 100 gained in tunneling time. It should also be mentioned that Fourier acceleration (in our model) increases the computer time per iteration by 50–100%. For more complicated models, evaluation of $\delta S/\delta\phi$ will be considerably more time-consuming, and the time required for Fourier transforms will be insignificant. Even for our model, the price of acceleration is far from cancelling the advantages of the method.

ACKNOWLEDGMENTS

We thank Roscoe Giles and Richard Brower for arranging use of the Encore

Multimax, a research computer at the Boston University College of Engineering, and we thank Peter Lepage and Paul Mackenzie for stimulating conversations. This work was supported in part by the U.S. Department of Energy under Grant No. DE-AC02-76ER03069.

REFERENCES

1. G. Parisi, in *Progress in Gauge Field Theory*, edited by G. 't Hooft *et al.* (Plenum, New York, 1984), p. 531.
2. G. G. Batrouni, G. R. Katz, A. S. Kronfeld, G. P. Lepage, B. Svetitsky, and K. G. Wilson, Phys. Rev. D**32**, 2736 (1985).
3. P. C. Hohenberg and B. Halperin, Rev. Mod. Phys. **49**, 435 (1977); J. Zinn-Justin, Nucl. Phys. **B275**, 135 (1986).
4. G. R. Katz, Ph. D. Thesis, Cornell University (June 1986); C. Davies, G. Batrouni, G. Katz, A. Kronfeld, G. P. Lepage, P. Rossi, B. Svetitsky, and K. G. Wilson, J. Stat. Phys **43**, 1073 (1986); C. Davies, talk given at Conference on Lattice Gauge Theory, Brookhaven National Laboratory, September, 1986, to appear in proceedings.
5. G. G. Batrouni, A. Hansen, and M. Nelkin, Phys. Rev. Lett. **57**, 1336 (1986).
6. J. Kogut, Nucl. Phys. **B275**, 1 (1986); E. Dagotto and J. Kogut, Illinois preprint ILL-TH-86-57 (unpublished, 1986).
7. To avoid confusion, we point out that the tunnelling of which we speak occurs in Langevin simulation time, and has nothing to do with true physical tunnelling in equivalent quantum problems.
8. S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, Science **220**, 671 (1983); S. Kirkpatrick, J. Stat. Phys. **34**, 975 (1984).
9. See for example A. T. Ogielski, Phys. Rev. B**32**, 7384 (1985).
10. K. Binder, ed., *Monte Carlo Methods in Statistical Physics* (Springer-Verlag, New York, 1979).
11. See for example A. D. Kennedy, J. Kuti, S. Meyer, B. J. Pendleton, Phys. Rev. Lett. **54**, 87 (1985).
12. N. G. Van Kampen, *Stochastic Processes in Physics and Chemistry* (North Holland, Amsterdam, 1981); H. Risken, *The Fokker-Planck Equation* (Springer-Verlag, New York, 1986).

FIGURE CAPTIONS

Fig. 1 (a) Fourier accelerated time evolution of the volume-averaged magnetization $\langle \phi \rangle$ for $\lambda = 0.4$, $\phi_0 = 1.6$, $h = 0$, with $dt = 0.1$. Initial conditions were a "cold start," with $\phi = +1.6$. Each point represents an average over 10 successive iterations. (b) A histogram of the iteration-by-iteration values.

Fig. 2 (a) Unaccelerated evolution for $\lambda = 0.4$, $\phi_0 = 1.4$, $h = 0$, with $dt = 0.1$. Initial conditions as in Fig. 1. (b) Accelerated evolution for the same case. Note the change in horizontal scale between (a) and (b).

Fig. 3 (a) Accelerated evolution for $\lambda = 0.4$, $\phi_0 = 1.6$, $h = 0.005$ (asymmetric minima), with $dt = 0.1$. Initial conditions as in Fig. 1. (b) Histogram of the iteration-by-iteration values.