Comparison between the Langevin and the hybrid simulation techniques for a free field theory

1989 J. Phys. A: Math. Gen. 22 5065

(http://iopscience.iop.org/0305-4470/22/23/011)

J. Phys. A: Math. Gen. **22** (1989) 5065-5072. Printed in the UK

Comparison between the Langevin and the hybrid simulation techniques for a free field theory

Gloria M Buendía

Department of Physics, Universidad Simón Bolivar, Apartado 89000, Caracas 1080A, Venezuela† and Department of Physics, University of California at San Diego, La Jolla, CA 92093, USA

Received 2 May 1989

Abstract. Langevin and hybrid methods are studied on a free field theory. In this simple model it is easy to detect the nature of some of the problems associated with these algorithms. Systematic errors following from the discrete time step used in the numerical implementation of the algorithms are computed.

1. Introduction

At present, some of the most challenging problems in computer simulation techniques are those involving dynamical fermions. Systems with fermionic degrees of freedom are characterised by non-local actions, extremely hard to simulate with standard methods. Very recently several groups have devoted themselves to developing a global updating procedure where the action does not need to be recomputed as often as in the Metropolis algorithm [1]. Into this category fall the so-called dynamical techniques: microcanonical [2], Langevin [3, 4] and their combination, the hybrid algorithm [5].

The basic idea behind these dynamical methods is to introduce an extra degree of freedom, a fake time, and describe the evolution of the system by a set of differential equations, in such a way that for large times, the system will reach equilibrium. Configurational averages are replaced by time averages in the dynamical trajectories. The most attractive feature of these techniques is that the computer time required to perform a complete sweep (with a small step size) through the lattice grows linearly with the volume, compared with the computer time required to update a single link of the lattice with a Metropolis algorithm, that grows with the cube of the volume.

However, algorithms based on the solution of differential equations involve systematic errors due to the finite step size introduced to find their numerical solution. If, in order to get correct physical results, it is necessary to choose time steps that are too small, the evolution of the system will slow down considerably and, in the long run, these methods can be more expensive than the exact updating.

Several of these algorithms have already been employed to study QCD in the presence of dynamical quarks. The results are very encouraging, but they reinforce the need for a better understanding of the behaviour of the techniques, which can be clarified by the study of simple models. This is why we choose to perform our study with an extremely simple model, the free field theory.

† Present address.

0305-4470/89/235065+08$02.50

A system of non-interacting fields is described by an action of the form

$$S = \int_{-\infty}^{\infty} \mathrm{d}x\, \phi^*(x)(\partial^2 + m^2)\phi(x). \tag{1}$$

Performing a Fourier transformation of the fields

$$\phi(x) = \int_{-\infty}^{\infty} \frac{\mathrm{d}k}{\sqrt{2\pi}} \mathrm{e}^{\mathrm{i}kx} \tilde{\phi}(k) \qquad \phi^*(x) = \int_{-\infty}^{\infty} \frac{\mathrm{d}k}{\sqrt{2\pi}} \mathrm{e}^{-\mathrm{i}kx} \tilde{\phi}^*(k) \tag{2}$$

the action can be written in terms of the Fourier transformed fields as

$$S = \int_{-\infty}^{\infty} \mathrm{d}k\, \tilde{\phi}^*(k)(k^2 + m^2)\tilde{\phi}(k) \tag{3}$$

i.e., a collection of harmonic oscillators. Next, we are going to illustrate how the Langevin and the hybrid techniques can be applied to the simulation of this system. We hope that, in this simple model, it is going to be easier to detect and understand the nature of some of the problems of these techniques. This model is simple enough that it can be solved analytically, giving some insight into the performance of the algorithms and the dependence of the error on the step size.

2. Langevin technique [3, 4]

Given a statistical system

$$Z = \int \prod_n \mathrm{d}\phi_n\, \mathrm{e}^{-S(\phi)} \tag{4}$$

adding an extra coordinate, a fake time, we can give dynamics to the system by writing a Langevin equation for each of the fields:

$$\frac{\mathrm{d}\phi_n}{\mathrm{d}\tau} = -\frac{\partial S}{\partial \phi_n} + \eta(n, \tau) \tag{5}$$

where the last term is a Gaussian random field that satisfies

$$\overline{\eta(n_1, \tau_1)\eta(n_2, \tau_2)} = 2\delta_{n_1 n_2}\delta(\tau_1 - \tau_2) \qquad \overline{\eta(n, \tau)} = 0. \tag{6}$$

It can be proved that

$$\langle W[\phi] \rangle = \overline{W[\phi]} = \lim_{T \to \infty} \frac{1}{T} \int_0^T W[\phi(\tau)]\, \mathrm{d}\tau \tag{7}$$

i.e. ensemble averages can be calculated as averages over time trajectories.

Applying the Langevin equation to the system described by (3) we get

$$\frac{\mathrm{d}}{\mathrm{d}\tau} \tilde{\phi}(k, \tau) = -(k^2 + m^2)\tilde{\phi}(k, \tau) + \tilde{\eta}(k, \tau) \tag{8}$$

this equation can be solved exactly, given [6]:

$$\tilde{\phi}(k, \tau) = \int_{-\infty}^{\tau} \mathrm{d}s\, \exp[-(\tau - s)(k^2 + m^2)]\tilde{\eta}(k, s) \tag{9}$$

and the autocorrelation function can be calculated as a time average (indicated by an overbar)

$$C(k,\tau)=\overline{\tilde{\phi}^*(k,0)\phi(k,\tau)}$$
$$=\int_{-\infty}^{0}ds_1\int_{-\infty}^{\tau}ds_2\exp[s_1(k^2+m^2)][-(\tau_2-s_2)(k^2+m^2)]\overline{\tilde{\eta}^*(k,s_1)\tilde{\eta}(k,s_2)}$$
$$=\frac{\exp[-\tau(k^2+m^2)]}{k^2+m^2}. \tag{10}$$

In terms of the spatial fields, the autocorrelation function can be written as

$$\overline{\phi^*(x,0)\phi(x,\tau)}=\frac{1}{2\pi}\int_{-\infty}^{\infty}dk_1\int_{-\infty}^{\infty}dk_2\exp[-i(k_2-k_1)\tau]\overline{\tilde{\phi}^*(k_1,0)\tilde{\phi}(k_2,\tau)}$$
$$=\frac{1}{2m}[1-\Phi(m\sqrt{\tau})] \tag{11}$$

where $\Phi(m\sqrt{\tau})$ is the well known error function, defined as

$$\Phi(m\sqrt{\tau})=\frac{2}{\sqrt{\pi}}\int_0^{m\sqrt{\tau}}ds\exp(-s^2). \tag{12}$$

In figure 1 we show the results for the normalised autocorrelation function

$$C(\tau)=\overline{\phi^*(x,0)\phi(x,\tau)}/\overline{\phi^*(x,0)\phi(x,0)} \tag{13}$$

obtained by solving numerically the integral that appears in (12) for different values of the mass.

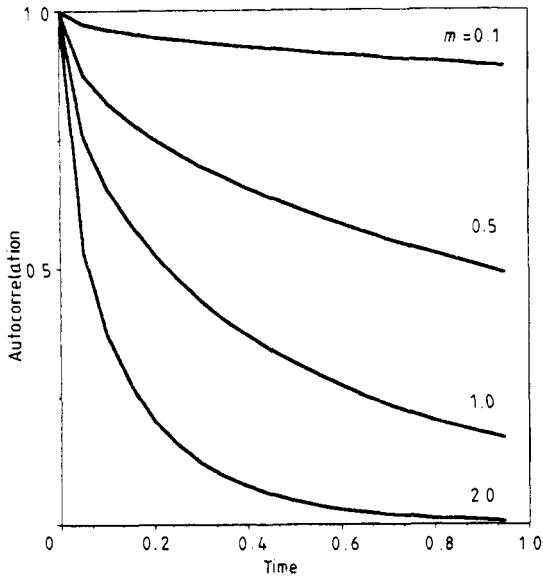

Figure 1. Autocorrelation function as a function of time at different masses, calculated with the Langevin algorithm.

Notice the dramatic increase in the autocorrelation function at small values of the mass (as predicted by (10)). This behaviour is due to the contribution of the long-wavelength (low-frequency) modes for which

$$\overline{\phi^*(k,0)\phi(k,\tau)} \underset{k \ll m}{\sim} \frac{\exp[-(k^2+m^2)\tau]}{m^2} \tag{14}$$

becomes large as the mass decreases.

Now, in order to evaluate the errors introduced when the Langevin equation is solved numerically (as happens in most cases) we are going to repeat the above procedure for the discrete version of (8)

$$\tilde{\phi}(k,n+1) = \tilde{\phi}(k,n) - \Delta(k^2+m^2)\tilde{\phi}(k,n) + \tilde{\eta}(k,n) \tag{15}$$

where Δ is the time step size. The solution of this equation can be easily calculated, given

$$\tilde{\phi}(k,n) = \sum_{s=-\infty}^{n-1} [1-\Delta(k^2+m^2)]^{n-1-s}\sqrt{\Delta}\,\tilde{\eta}(k,s) \tag{16}$$

from which we can calculate the autocorrelation function

$$\overline{\tilde{\phi}^*(k,0)\tilde{\phi}(k,n)} \simeq \frac{\exp[-\Delta(k^2+m^2)n]}{(k^2+m^2)[1-\frac{1}{2}\Delta(k^2+m^2)]}. \tag{17}$$

From (17) we can define the normalised correlation at time, $\tau = n$ as

$$C(\tau) = \overline{\phi^*(k,0)\phi(k,\tau)} / \overline{\phi^*(k,0)\phi(k,0)} = e^{-\tau/\gamma} \tag{18}$$

where the correlation time, γ, is defined as

$$\gamma = \frac{1}{\Delta(k^2+m^2)}. \tag{19}$$

Thus, in order to keep the stability of the algorithm, we must choose

$$\Delta < \frac{1}{k_{\max}^2+m^2} \tag{20}$$

where $k_{\max}$ is the maximum momentum allowed in the lattice.

Taking into account the restriction for the step size, the lower limit for the correlation time must be

$$\gamma_{\min} = \frac{k_{\max}^2+m^2}{k^2+m^2} \tag{21}$$

i.e. the high-momentum components will evolve much faster (shorter correlation time) than the low-momentum components.

For the low-momentum component, long-wavelength modes, the correlation time goes as

$$\gamma(k \to 0) \sim \frac{1}{m^2} \sim \zeta^2 \tag{22}$$

with ζ being the longest correlation length of the theory (in lattice units). As we approach the continuum limit $\zeta \to \infty$, the situation becomes worse.

3. The hybrid algorithm [5]

The hybrid method has its origin in the observation that the microcanonical and the Langevin algorithms have complementary advantages, so there is a chance that a method that combines both of them will offer a better performance.

The hybrid method works in the following way. An extra degree of freedom, a fake time, and a new set of coordinates p, conjugate to the ϕ, are introduced to the system. The evolution of the system is described by an algorithm such that

$$\tilde{\phi}(n+1) = \tilde{\phi}(n) + \Delta\tilde{p}(n) - \tfrac{1}{2}\Delta^2 \frac{\partial S}{\partial \tilde{\phi}(n)} \tag{23}$$

where the $p(n)$ are selected according to

$$\tilde{p}(n) = \begin{cases} (\tilde{\phi}(n+1) - \tilde{\phi}(n))/2\Delta & \text{probability } q\Delta \\ \zeta(n) & \text{otherwise.} \end{cases} \tag{24}$$

$\zeta(n)$ is a Gaussian random function satisfying $\langle\zeta^2\rangle = 1$ and Δ is the time step size.

The system evolves most of the time with the classical equations of motion (the microcanonical algorithm), and occasionally the momenta are refreshed; this means that all the momenta are substituted by new ones taken from a Gaussian random distribution, normalised in such a way that the average kinetic energy is correct. In this way, the system evolves most of the time with the fast algorithm, while the occasional refreshing of the momenta introduces the randomness that will assure ergodicity. This algorithm can be optimised by tuning the refreshing frequency (the parameter q) in order to minimise the autocorrelation. It is easy to see that the Boltzmann distribution is an invariant distribution of the system. In the limit $q = 0$ (never refreshing), the microcanonical algorithm is recovered. In contrast, if $q\Delta = 1$ (refreshing after each step), the algorithm is equivalent to the Langevin algorithm with the identification $\Delta(\text{Langevin}) = \Delta^2(\text{hybrid})/2$.

Following the approach of Duane [7], we are going to calculate the correlation function $\overline{\phi(0)\phi(\tau)}$, where the average is now over the initial momenta, the number of times of refreshings, and over the intermediate momenta obtained after the refreshing. For simplicity we are going to perform the calculation in the continuum limit of (23).

Without the refreshing step, the evolution would be determined by the classical equation of motion

$$\ddot{\tilde{\phi}}(k, \tau) = -(k^2 + m^2)\tilde{\phi}(k, \tau) \tag{25}$$

with solution

$$\tilde{\phi}_0(k, \tau) = \tilde{\phi}_0(0)\sin(\alpha + (k^2 + m^2)\tau) \tag{26}$$

and correlation function

$$\begin{aligned} \tilde{\phi}_0(k, 0)\tilde{\phi}_0(k, \tau) &= \frac{\tilde{\phi}_0^2(k, 0)}{2\pi}\int_0^{2\pi} \mathrm{d}\alpha \sin\alpha \sin(\alpha + (k^2 + m^2)\tau) \\ &= \tfrac{1}{2}\tilde{\phi}_0^2(k, 0)\cos(k^2 + m^2)\tau. \end{aligned} \tag{27}$$

The subscript 0 means that we are working with the microcanonical ensemble, no refreshing here.

The autocorrelation function $C(k, \tau)$ for the hybrid algorithm can be calculated as a series expansion in the number of refreshings occurring in the interval $(0, \tau)$. Each

scattering at time τ' introduces a factor of $q\,d\tau$, with q being the refreshing frequency. The series can be summed, giving an expression

$$C(k,\tau) = C_0(k,\tau) - q\int_0^\tau d\tau' \cos(\tau-\tau')C(\tau). \tag{28}$$

Then the $C(k,\tau)$ satisfy an equation for a damped harmonic motion

$$\ddot{C}(k,\tau) + q\dot{C}(k,\tau) + (k^2+m^2)C(k,\tau) = 0 \tag{29}$$

with solution of the form

$$C(k,\tau) \sim \frac{e^{-\tau/\gamma}}{k^2+m^2} \tag{30}$$

where the correlation time is

$$\gamma = \begin{cases} 2/q & q < 2\sqrt{k^2+m^2} \\ \dfrac{1}{(q/2)-\sqrt{(q^2/4)-(k^2+m^2)}} & q > 2\sqrt{k^2+m^2}. \end{cases} \tag{31}$$

The optimal refreshing frequency for the k mode, that which minimised the correlation function, is $q = 2\sqrt{(k^2+m^2)}$.

In terms of the spatial fields we can write the autocorrelation function as

$$\overline{\phi^*(x,0)\phi(x,\tau)} = \int_{-\infty}^{\infty} \frac{dk}{2\pi}\,\overline{\tilde{\phi}^*(k,0)\tilde{\phi}(k,\tau)} \tag{32}$$

$$= \frac{\exp[-(q/2)\tau]}{\pi} \times\left(\int_0^{\sqrt{(q^2/4)-m^2}} dk\,\frac{\exp[\sqrt{(q^2/4)-(m^2+k^2)}]}{k^2+m^2} + \int_{\sqrt{(q^2/4)-m^2}}^{\infty} dk\,\frac{1}{k^2+m^2}\right)$$

$$= \frac{\exp[-(q/2)\tau]}{m\pi}\left(\frac{\pi}{2} - \tan^{-1}\frac{\sqrt{(q^2/4)-m^2}}{2m}\right). \tag{33}$$

Figure 2 shows the normalised correlation function, defined by (13) obtained numerically from (33) for different values of the mass. We found the same effect as with the Langevin algorithm that the smaller the mass, the longer it takes the autocorrelation to decay.

An interesting characteristic of the figures is that the optimal refreshing frequency is always $q = 2m$, that is exactly the optimal value for the $k = 0$ mode, the slowest mode of the system. The autocorrelation is dominated by the low-frequency, long-wavelength modes.

It is evident that, for both algorithms, there is a strong relationship between the autocorrelation function and the mass. The autocorrelation time increases considerably when the mass decreases, the evolution of the system is dominated by the long-wavelength, low-frequency modes. This is a serious problem shared by almost all the Monte Carlo simulation methods and, in this study, we show that the Langevin and hybrid methods are not exempt from it.

The Langevin and the hybrid methods have serious problems for the simulation of systems with small masses. For these systems the step size must be reduced in such a way that the simulation becomes very slow, taking the long-wavelength modes too

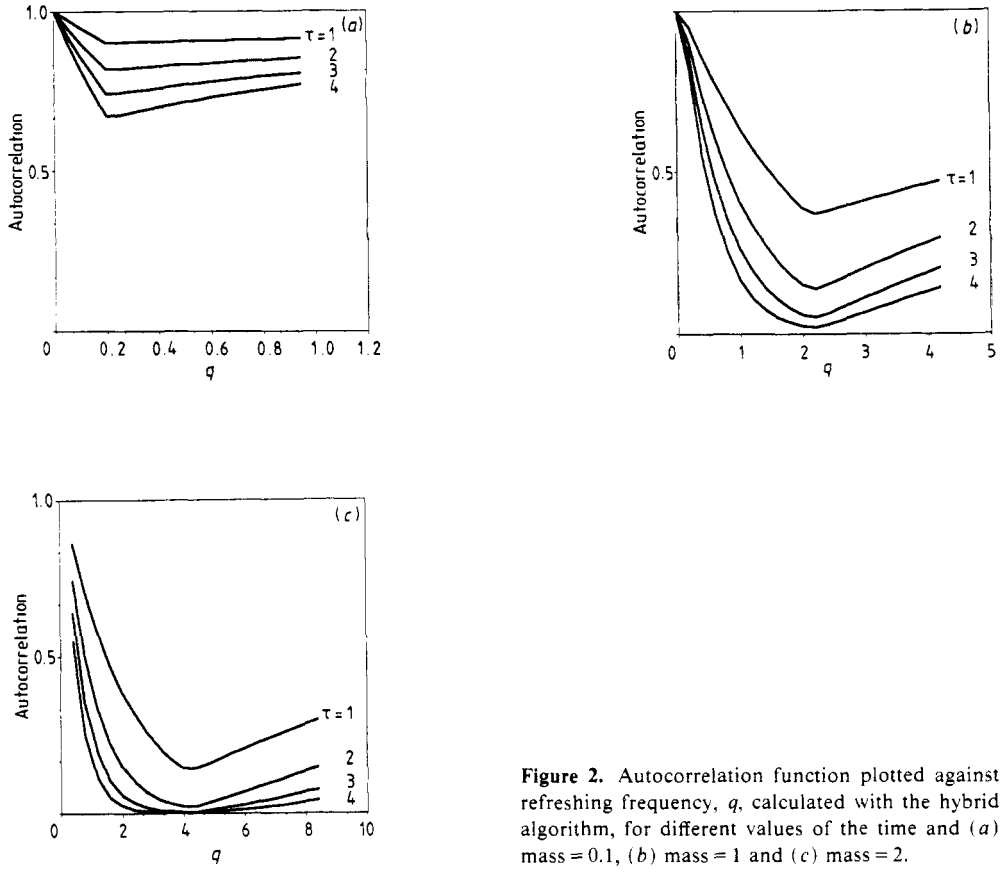

Figure 2. Autocorrelation function plotted against refreshing frequency, q, calculated with the hybrid algorithm, for different values of the time and (a) mass = 0.1, (b) mass = 1 and (c) mass = 2.

long to equilibrate. This seems to be a very serious handicap to the use of these algorithms for the simulation of systems near phase transitions, where all wavelengths are involved. This problem has already been observed in the simulation of QCD with dynamical fermions, where these algorithms have difficulty simulating systems with small quark masses [8].

There is hope that this problem can be avoided with the recently introduced fast Fourier acceleration techniques [4], which suggest the introduction of a non-local time step size, such that

$$\Delta(x) = \sum_k e^{ikx} \Delta(k)$$

where

$$\Delta(k) \sim \frac{1}{k^2 + m^2}$$

will give a minimum correlation time for all modes. However, it is not clear how these techniques can be applied to gauge theories and its implementation is under study.

Acknowledgment

I am indebted to Doug Toussaint for useful conversations and suggestions.

References

[1] Metropolis N, Rosenbluth A, Rosenbluth M, Teller A H and Teller E 1963 *J. Chem. Phys.* **21** 1987
[2] Callaway D and Rahman R 1982 *Phys. Rev. Lett.* **49** 613
Polonyi J and Wyld H 1983 *Phys. Rev. Lett.* **51** 2257
[3] Parisi G and Yongshi Wu 1981 *Sci. Sin.* **24** 483
[4] Batrouni G G, Katz G R, Kronfeld A S, Lepage G P, Svetitsky B and Wilson K G 1985 *Phys. Rev.* D **32** 2736
[5] Duane S 1985 *Nucl. Phys.* B **257 [FS14]** 652
[6] Martin O, Otto S and Flower J 1986 *Nucl. Phys.* B **264** 89
[7] Duane S and Kogut J B 1986 *Nucl. Phys.* B **275 [FS17]** 398
[8] Kogut J B 1986 *Phys. Rev. Lett.* **56** 2557
Ding H Q and Martin O 1987 *Nucl. Phys.* B **280 [FS18]** 497
Fukugita M and Ukawa A 1988 *Nucl. Phys.* B **300 [FS22]** 433