RU–91–46
October 15, 1991

A possible barrier at $z \approx 1$ for local algorithms.

by

Georgios Bathas and Herbert Neuberger

Department of Physics and Astronomy
Rutgers University
Piscataway, NJ 08855–0849

Abstract

It is shown that a certain class of generalizations of over-relaxation algorithms is incapable to further reduce the dynamical exponent z below its standard over-relaxed value of $z \approx 1$. The $\approx$ sign means that the mean field value is one, while the true value can be somewhat different in certain dimensions. The generalizations are obtained by viewing over-relaxation as a slightly deformed deterministic algorithm and should, therefore, hold for Hybrid Monte Carlo algorithms as well.

Massively parallel computer architectures are best utilized by local Monte-Carlo algorithms [1,2]. By locality we mean the following: At each iteration an old site variable, generically denoted by $\phi(x)$, where x is the site, is replaced by a new site variable, $\phi\prime(x)$, and $\phi\prime(x) = F_x(\phi(y))$ with y restricted to a finite vicinity of x. By a finite vicinity we mean one that contains a finite, volume independent, number of sites. Such an algorithm is parallelizable because we can split the lattice into a finite number of equal sub-lattices (depending on the range of site interdependence at updates) such that updates of a variable in any given sub-lattice can depend only on the variables in the other sub-lattices and thus perform the update of an entire sub-lattice in parallel. The sub-lattices intertwine building up the complete lattice.

Let us denote by $\phi(x, n)$ the site variable at x during the n-th iteration. In most applications the system being simulated is translational invariant. This implies that the autocorrelation matrix for the site variables, asymptotically, decouples into finite blocks of size determined by the number of sub-lattices, N: By translational invariance we mean that, assuming that we started from a uniform configuration, the probability distribution of the variable $\phi(x, n)\phi(y, n+m)$ $m > 0$ remains unchanged under a simultaneous shift of x and y restricted to keep each site in its original sub-lattice. It's easy to see that ordinary algorithms respect this, as a one to one mapping between configurations can be established by simply shifting around the random numbers generated when updating the sites. As a consequence we obtain that the autocorrelation matrix can be block diagonalized for n large enough that the probability for the set $\phi(x, n)$ be invariant under shifts in x [3]. Each block is of size $N \times N$, corresponding to the N regions the Brillouin zone becomes devided into as a result of the restriction on the allowed translations.

Since the algorithm is local it is reasonable to accept the supposition that there exists a field theoretical representative of its dynamical universality class [4]. We are trying to find then the dynamical renormalization group trajectory connecting the ordinary dynamical GL fixed point to the true fixed point. Thus it makes some sense to start at the single place that we have under control calculationally, namely the free field case. One might be suspicious at having something interesting to say about a free field system; however the problem is not entirely trivial because finding a *local* algorithm (equivalently: dynamics) that has a critical exponent z smaller than 2 isn't that easy even for the free field case. Once we know the behavior for the free field case we ought to proceed to investigate the effect of interactions by the ordinary field theoretical tools, namely, characterizing the strength of perturbations, finding the dimension where the theory is renormalizable, etc. In the present note we shall only deal with the simplest first step. As we shall come up with a negative result there is no room for the above mentioned subsequent steps. However, our result does indicate that no new sort of dynamical fixed point, with a z smaller than 1, and continuously connected to the dynamical GL fixed point can exist, leaving us only with ordinary Langevin dynamics and with Over-relaxation. We don't have a rigorous proof of this assertion because we shall be making some additional assumptions on the way. We do suspect however that these assumptions are not restrictive in reality and that the assertion holds as stated.

Let us review first how over-relaxation works [5]. It is sufficient to have only two sub-lattices. The block that contains the slowest modes is the one coupling the two sub-lattice field averages, $\phi_{1,2}$. In the critical case the mass term is zero and the energy depends only on the difference of the modes. Hence the Hamiltonian, when restricted to the $\phi_{1,2}$ subspace, can be written as:

$$H_0 = (\phi_1 - \phi_2)^2 = \phi_i h_{ij} \phi_j; \qquad h_{11} = h_{22} = -h_{12} = 1 \tag{1}$$

A mass perturbation will change the Hamiltonian by

$$\delta H_0 = m^2 \phi_j \phi_j \tag{2}$$

The structure of the auto-correlation matrix is compatible with the assumption that the algorithm at each iteration only couples the two modes of the block (every mode couples only to the one whose crystal momentum is shifted by $(\pi, \pi, \ldots, \pi)$). Over-relaxation algorithms fulfill this and, moreover, have a limit (when an over-relaxation parameter is taken to an extreme value) where they become deterministic. In this limit the algorithm becomes a mapping under which H is preserved. This mapping should be viewed as some discrete generalization of Hamiltonian mechanics and hence is required to preserve the measure $d\phi_1 d\phi_2$ as an analogue of Liouville's theorem. This is needed for the algorithm to produce the right limiting distribution. The simplest way to ensure this is to make the mapping a linear operator T:

$$\phi_\alpha(n+1) = T_{\alpha,\beta} \phi_\beta(n) \tag{3}$$

The preservation of the measure means that T has unit determinant and, for convergence, we forbid any of the eigenvalues of T to be outside the unit circle. This implies that all of T's eigenvalues are on the unit circle.

When the mass is zero, we expect no convergence (we shall consider oscillatory behavior with a constant amplitude as "convergence" and the time scale governing this convergence as given by the related phase) and therefore one of the eigenvalues of T (the one corresponding to the total average) is real and equal to 1. When the mass term is turned on this eigenvalue of T changes, and for small m should go as $\lambda = \exp(i\theta)$ with $\theta \propto m^z$ where z is the dynamical exponent. Ordinarily, T would depend linearly on H and the order m^2 perturbation would translate into order m^2 changes in the eigenvalues hence $z = 2$. We assume that, from the point of view of matrix perturbation theory, the perturbation induced by the mass term in T is generic.* Essentially the same picture holds in all the blocks that contain a very low momentum (instead of strictly zero) mode. To achieve something else we need a situation where a generic order m^2 perturbation of T causes a *larger* change in the eigenvalues of T. This can happen only if, at $m = 0$, T is *nondiagonalizable*, because, otherwise, we can use ordinary perturbation theory. In our case this leaves us with the single possibility, that in some basis, T contains a sub-block of Jordan form with unity on the diagonal, i.e.:

* Even if this is not true for one particular block containing a soft mode it can't be true of all the blocks that contain soft modes.

$$T = \left(\begin{array}{cc} \underbrace{\begin{pmatrix} 1 & 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \ddots & & \vdots \\ & & & \ddots & \ddots & \\ & & & 0 & 1 & 1 \\ 0 & 0 & \dots & & 0 & 1 \end{pmatrix}}_{n\times n \text{ Jordan block}} & 0 \\ 0 & \begin{pmatrix} \dots \\ \dots \\ \dots \end{pmatrix} \end{array} \right) \tag{4}$$

For example, with ordinary over-relaxation applied to

$$S[\phi] = \frac{1}{2}\sum_{x,\mu}[\phi(x+\mu) - \phi(x)]^2 \tag{5}$$

where x are sites on a periodic d dimensional hypercubic lattice we have, in the deterministic limit,

$$T = \begin{pmatrix} 1 & -4 \\ 0 & 1 \end{pmatrix} \tag{6}$$

for the sub-block operating in the space spanned by $\phi_1 \pm \phi_2$. This matrix is similar to a 2×2 Jordan block.

When $m \neq 0$ there are two conjugate eigenvalues, $\exp(\pm i\theta)$ and θ is of order m. When the over-relaxation parameter is tuned down from the deterministic limit both eigenvalues decrease in absolute magnitude until, at some point, they become real. Throughout the range where they are complex they are at an order m distance from unity in the complex plane. The range the over-relaxation parameter is varying in is of the same order m also so its effect on the eigenvalues is of the same order as that of the direct mass perturbation. The main observation is that the variability of the over-relaxation parameter away from the deterministic limit is not the main mechanism by which the exponent z gets reduced, but rather, it is the deterministic limit that de "main job" of reducing z and the variability of the over-relaxation parameter only permits the introduction of dissipation (noise) without disrupting this "achievement" of the deterministic limit.

It is clear now how one might try to generalize and do even better. We wish to make the arc AO in Fig. 1 of length $O(m^z)$ with z now less than unity. When there are more sub-lattices one will have higher dimensional sub-blocks in the matrix T and to get a smaller z we would like to arrange for the sub-block of T containing the zero-mode to be similar to a matrix with a larger irreducible Jordan block. If the whole matrix is a single block of size $n \times n$ we could get $z = \frac{2}{n}$. A perturbation of order m^2 at every entry of an $n \times n$ Jordan block will change its characteristic equation, from $(\lambda - 1)^n = 0$ to $(\lambda - 1)^n = const. \cdot m^2 + O(m^2(1-\lambda))$ and hence $(\lambda - 1) \sim O(m^{2/n})$*. We shall show below that, regretfully, any acceptable T cannot contain a block with $n > 2$.

---

* It is easy to see that $const. \neq 0$ if a certain matrix element of the perturbation (the lower-left corner) doesn't vanish; this will happen in the generic case. To get a given $z = \frac{2}{n}$ the minimal needed size of the Jordan block is $n \times n$.

Since H is translationally invariant we take it as being block diagonal in the same basis that we assume T is and we again concentrate only on the block containing the slowest mode. In it we have:

$$H_0 = \phi_\alpha h_{\alpha,\beta} \phi_\beta \qquad \alpha, \beta = 1, 2, ..., n. \tag{7}$$

In the same block we assume that there exists a nonsingular matrix S such that $S^{-1}TS = J$ where $J_{\alpha,\beta} = \delta_{\alpha,\beta} + \delta_{\alpha+1,\beta}$ for $\alpha, \beta = 1, 2, ..., n$. Energy conservation, $T^T hT = h$, implies:

$$J^T kJ = k. \tag{8}$$

The above equations can be viewed as a set of equations for the matrix elements of S where h is given and $k = S^T hS$. Explicitly, they read:

$$k_{\alpha,\beta-1} + k_{\alpha-1,\beta} + k_{\alpha-1,\beta-1} = 0 \tag{9a}$$

$$k_{0,\alpha} \equiv k_{\alpha,0} \equiv k_{0,0} \equiv 0 \tag{9b}$$

The relation between the bilinear forms h and k means that they have equal numbers of positive, vanishing and negative eigenvalues. On the other hand we know that h has a single zero eigenvalue corresponding to the total field average and all other eigenvalues are strictly positive. Therefore this must also be true of k. The above equations in conjunction with the semi-definiteness of k lead to a contradiction if $n > 2$.

To see this let us agree that we shall call a line of elements in the matrix k defined by $\{k_{\alpha\beta};\ \alpha+\beta = m\}$ an m "off-diagonal". The equation on k implies that the m off-diagonals with $m < n$ vanish. The n off-diagonal also has to vanish because the determinant of k vanishes. The first nonvanishing off-diagonal has, as a result of the equation for k, entries that are equal in magnitude, but alternate in sign. Since k is symmetric it must be that the m of the first nonvanishing off-diagonal is odd. So k looks like:

$$
k = \begin{pmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0 & 0 & \gamma \\
0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0 & -\gamma & \star \\
0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \gamma & \star & \star \\
0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \star & \star & \star \\
0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \star & \star & \star \\
0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \star & \star & \star \\
0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0 & 0 & \gamma & \star & \star & \star & \star & \star & \star \\
0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0 & (-\gamma) & \star & \star & \star & \star & \star & \star & \star \\
0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \gamma & \star & \star & \star & \star & \star & \star & \star & \star \\
0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \star & \star & \star & \star & \star & \star & \star & \star & \star \\
0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \star & \star & \star & \star & \star & \star & \star & \star & \star \\
0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \star & \star & \star & \star & \star & \star & \star & \star & \star \\
0 & \cdot & \cdot & 0 & 0 & \gamma & \star & \star & \star & \star & \star & \star & \star & \star & \star & \star & \star & \star \\
0 & \cdot & \cdot & 0 & -\gamma & \star & \star & \star & \star & \star & \star & \star & \star & \star & \star & \star & \star & \star \\
0 & \cdot & \cdot & \gamma & \star & \star & \star & \star & \star & \star & \star & \star & \star & \star & \star & \star & \star & \star
\end{pmatrix} \tag{10}
$$

The determinants of the sub-matrices of size $l \times l$ (l odd), containing a piece of the original nonvanishing m off-diagonal as an l off-diagonal must all be positive. This cannot be true because their signs alternate with l. The single way out is that the largest l available be $l = 1$ and this happens in ordinary over-relaxation, when $n = 2$.

The previous argument also excludes the possibility that the sub-block of T containing the zero mode is not fully reducible but only has a further sub-block of Jordan form with $n > 2$. One simply applies the above logic to the appropriate principal sub-matrix of k (by "principal" we mean that it contains a piece of the original diagonal as its diagonal). This sub-matrix must also be positive semi-definite (it could no longer have a zero mode, but this doesn't help) and the above argument applies again.

We implicitly assumed in the above that S is real; we know that T is real but S will be complex if the eigenvalues of T have imaginary parts. It is easy to see that in this case the equations k has to satisfy are even more restrictive and again no solution exists.

In conclusion we have shown that, quite generally, one cannot come up with a deterministic local algorithm whose z would be lower than unity in the free field case. Intuitively, since a local algorithm must change just one spin in a vicinity at a time, we would expect that the time to reverse a long wavelength excitation would grow at least linearly with wavelength [6]. Therefore the result seems reasonable. The main new point of view presented in this note is that over-relaxation works just because it is a delicate deformation of a deterministic algorithm.* This puts over-relaxation closer conceptually to the Hybrid Monte Carlo algorithms [2,8] currently in use - the single difference is that here the system itself supplies also the conjugate momenta, while in Hybrid Monte Carlo, extra variables are introduced. For example, the zero mode of the field (Φ) in Hybrid Monte Carlo is coupled to the external conjugate variable (Π) by a 2×2 matrix, which, in the massless limit, with leap frog updating, takes the form:

$$T = \begin{pmatrix} 1 & \delta\tau \\ 0 & 1 \end{pmatrix} \tag{11}$$

Here the basis is Π, Φ and $\delta\tau$ is the time step. This equation is the direct analogue of the corresponding equation in the over-relaxed case (eq. (6)). We have little doubt that the main conclusion regarding the dynamical critical exponent z of the present note holds for all possible local generalizations of Hybrid Monte Carlo algorithms as well.

* As shown by Creutz [7], this deformation may not be needed at all in practice.

Acknowledgements. This research was supported in part by the DOE under grant # DE-FG05-90ER40559. We would like to thank Gyan Bhanot, Urs Heller and Sorin Solomon for useful comments.

References

[1] S. Adler, **Nucl. Phys. B. (Proc. Suppl.) 9** (1989) 437.

[2] A. Kennedy, **Nucl. Phys. B. (Proc. Suppl.) 4** (1988) 576.

[3] U. Heller, H. Neuberger, **Phys. Rev. D39** (1989) 616.

[4] P. C. Hohenberg, B. I. Halperin, **Rev. Mod. Phys. 49** (1977) 435.

[5] S. Adler, **Phys. Rev. D37** (1988) 458; H. Neuberger, **Phys. Rev. Lett. 59** (1987) 1877.

[6] A. Sokal, **Nucl. Phys. B. (Proc. Suppl.) 20** (1991) 55.

[7] M. Creutz, **Phys. Rev. D36** (1987) 515.

[8] A. Kennedy, B. Pendleton, **Nucl. Phys. B. (Proc. Suppl.) 20** (1991) 118.

Figure Caption

Fig. 1: *T-eigenvalue movement in the complex plane: On OA and OB the over-relaxation parameter is set to its deterministic limit and the mass is increased from $m = 0$ corresponding to O. On AC and BC the mass is fixed to a non-zero value and the over-relaxation parameter is varied away from the deterministic limit introducing dissipation. The whole line ACB is at a distance of order m from O, a property inherited from the point A.*