Evolving Numerical Enzymes: Accelerating Relaxation in the Frenkel-Kontorova Model

Shelly L. Shumway and James P. Sethna
Laboratory of Atomic and Solid State Physics, Cornell University, Ithaca, New York 14853-2501
(Received 22 February 1991)

We have investigated the finite-temperature behavior of the Frenkel-Kontorova model, and have found it unable to reach the ground state when cooled at a finite rate, freezing instead into some metastable configuration. The correlation length of the final state grows as the logarithm of the logarithm of the cooling time. By adding "numerical enzymes," or long-range Monte Carlo moves which precisely eliminate certain barriers to relaxation, we can equilibrate to significantly lower temperatures. Our numerical method for developing these enzymes is motivated by Darwinian evolution.

PACS numbers: 64.70.Pf, 02.70.+d

There are three themes to this paper. The first is a technical discussion of the finite-temperature dynamics of the Frenkel-Kontorova (FK) model, which is the simplest we know of that includes incommensurability and frustration in a natural way. We show that when cooled at a finite rate, the model falls out of equilibrium and freezes into some metastable configuration, and that the correlation length diverges incredibly slowly with slower cooling. The second is a general approach for equilibrating models of configurational glasses and other models with slow relaxation time scales. We use "numerical enzymes," or complicated multiple-atom Monte Carlo moves fine tuned to bypass the barriers to relaxation. The third is a practical implementation of principles of Darwinian evolution, which we use for finding the enzymes; we discuss insight gained from our experience about what is essential for such a system to thrive in an ever-changing environment.

The FK model [1] has been studied by many groups and applied to a variety of problems [2]. It consists of a one-dimensional chain of atoms, each connected to its nearest neighbors by springs, with an externally applied sinusoidal potential:

$$H = \tfrac{1}{2} K \sum_j (x_{j+1} - x_j - a)^2 - V \sum_j \cos x_j . \qquad (1)$$

Frustration occurs when the periodicity of the applied potential competes with the tendency of the springs to keep the atoms evenly spaced. We have studied the dynamical behavior of finite temperatures in the pinned limit [3], where $K \ll V$, using finite chains with periodic boundary conditions for the numerics, with the average number of atoms per well equal to rational approximants of the "golden mean," $(\sqrt{5}-1)/2$. We have investigated the behavior on cooling from a finite temperature, where the atoms are mobile, to zero temperature, where the chain is frozen. No finite cooling rate is slow enough for the system to equilibrate all the way into the ground state, but cooling more slowly results in fewer defects.

In the ground state, atoms occupy the same wells they would if uniformly distributed. An atom in the wrong well causes a local compression or stretching of the chain, which costs energy. The single-atom excitation energies fall into a clear hierarchy, as illustrated in Fig. 1; the energy cost of ϵ of a defect type is related to how delocalized it is, or to its characteristic length scale L, and goes as $(K/V)^L$ to leading order in K/V. Arbitrary defects may be represented as combinations of these elementary excitations, and although the precise energy depends on the exact location of every atom in the chain, to leading order in K/V it is simply the sum of the elementary excitation energies. This energy spectrum may be derived by a straightforward but tedious renormalization-group analysis which will be discussed elsewhere; it also follows from the work of Vallet, Schilling, and Aubry [4].

In equilibrium at temperature T, defects of energy ϵ will be populated with probability $e^{-\epsilon/kT}/(1+e^{-\epsilon/kT})$, so "shorter" defects, with $\epsilon \gg kT$, will be essentially nonexistent at temperature T, and "longer" defects, with $\epsilon \ll kT$, will be randomly populated. The length scale of

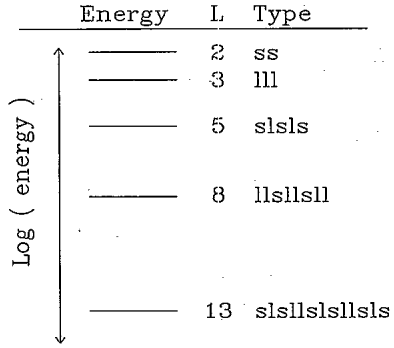

FIG. 1. The elementary excitation spectrum, computed numerically for a chain of 89 atoms with $K/V = 1/3$. An s ("short") represents a spring connecting atoms in adjacent wells and an l ("long"), a spring between atoms separated by an empty well. Each defect shown is due to one atom in the wrong well causing a segment of L springs to span one more or one fewer well than a corresponding segment of the ground state. The extra energy may be considered as arising from the interaction between the atom preceding the segment and the atom following, screened through L springs in between. The defect energy is $V(K/V)^L$ to leading order in K/V; corrections due to environment are small and are not resolvable here. (One displaced atom actually causes two defects, but the energy of the longer one is of higher order in K/V.)

the highest-energy defect type that is significantly populated determines the shortest distance between defects and sets the scale for any reasonable definition of correlation length. Since the length of the highest-energy defect in the chain scales as $-\ln(kT)$, the equilibrium correlation length must scale as $-\ln(kT)$.

Because hopping from well to well is a thermally activated process, the time scale for hopping is some geometry-dependent constant times $e^{E/kT}$, where E is the barrier height (~ 2 V). If the time to cool from $kT \approx 2\epsilon$ to $\sim \frac{1}{2}\epsilon$ is shorter than the time required for the necessary fraction of atoms to cross the barrier, the chain will end up stuck with an overpopulation of this defect type. If cooled at a rate γ, the system will fall out of equilibrium near temperature T if $\epsilon/\gamma \sim e^{E/kT}$. The lowest temperature to which the chain can equilibrate therefore goes asymptotically as the reciprocal of the logarithm of the cooling time $1/\gamma$, so the correlation length at the end of a run at rate γ to zero temperature scales as

$$\xi \sim \ln[\ln(1/\gamma)]\,. \tag{2}$$

We have stimulated this using a Metropolis Monte Carlo algorithm, and the divergence in correlation length is indeed extremely weak, as seen in Fig. 2. If this kind of simulation had been the only way to study the model, we would probably not even have suspected that the limiting behavior was an ordered ground state. The model is glassy in the sense that diverging time scales cause the "melt" to fall out of equilibrium at some history-dependent temperature and get stuck in one of many possible metastable configurations, generally exhibiting no long-range order. Like a glass, cooling a lot slower allows equilibration to only a slightly lower temperature.

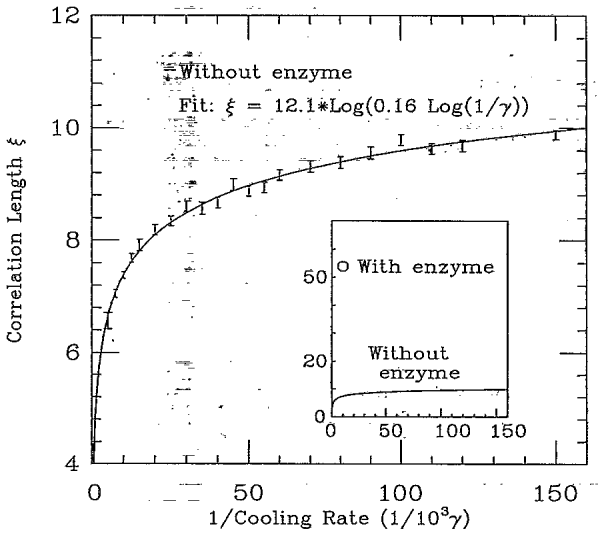

FIG. 2. The divergence of correlation length upon slower cooling, using an ordinary Monte Carlo algorithm and linear cooling schedule. The cooling rate γ is given as ΔkT per Monte Carlo step; each run was started at $kT=10$, so the longest run plotted was cooled for 1.5×10^6 time steps. The correlation length is the average length, measured in number of springs, of segments whose occupation patterns (or "*sl*" patterns as in Fig. 1) are indistinguishable from the ground state. By adding the 15-atom enzyme shown in Fig. 3, we were able to equilibrate rapidly to a correlation length of 54, as shown in the inset. Each point is an average over fifty runs of 610-atom chains. Here and throughout the paper, the unit of energy is the spring constant K, and $K/V=1/3$.

Many models exhibit time scales slow enough to make numerical study difficult. Various groups have been able to speed the dynamics of slow systems by adding special large moves to Monte Carlo simulations. Widom, Strandburg, and Swendsen [5] discovered that by adding long-range hops and a particular three-atom move to the normal small relaxations in simulating a two-component system of Lennard-Jones disks, they were able to equilibrate to tilable configurations where otherwise they got stuck in glassy states. Swendsen and Wang found in studying dynamical critical properties of Potts models that they could significantly reduce the exponent for size scaling at criticality and thus obtain good statistics for larger systems. Their special move involved mapping the Potts model onto a percolation model and changing an entire percolation cluster at once, resulting in a new Potts configuration, where each step could differ substantially from the previous one [6].

We expect that some kind of "numerical enzymes" like these may be useful in accelerating the dynamics of a wide variety of models with slow relaxation time scales. In the FK model, it is not difficult to figure out the appropriate enzymes since we already know what all the metastable states look like. For more complicated problems, however, it may be too difficult to figure out such moves analytically or intuitively. We have developed a technique that enables the computer to systematically discover good enzymes without much guidance, and have found that it works well in the FK model.

Our approach is similar in spirit to the genetic algorithms introduced by Holland [7] during the 1970s. Most applications have involved optimization problems in engineering [8]; the method has also been used for modeling experimental data in physics [9]. The basic idea is that the solution to a problem is discovered through evolution of a population of some kind, in which an appropriate formula for health or fitness determines each individual member's survival and reproduction. Those that do poorly die off, and are replaced most often by the offspring of exceptionally healthy individuals. In our algorithm, we have a population of moves: Each individual is represented by an n-dimensional vector telling how far to attempt to move each of n atoms. We use a modified Monte Carlo–like code in which we select one individual at random from this finite population of discrete moves for every update of the FK chain, keeping track of each one's performance in promoting equilibration. Detailed balance is preserved by multiplying the move by a random

sign whenever it is used. This algorithm is not a Markov process, and therefore not genuine Monte Carlo, but it is able to discover precise and complicated enzymes that can be used to accelerate a legitimate Monte Carlo code.

The formula for health must be based on success in lowering the energy of the chain, but any reasonable implementation of this should work fine. We have chosen to multiply an individual's health by a factor $\lambda < 1$ whenever it is tried, and add an amount proportional to the energy decrease of the chain if the new configuration is accepted. The population is updated periodically, and those that are unhealthy die off. We maintain a population of constant size; whenever one dies, two moves are selected randomly with probability proportional to their health, and a new individual is created as their vector sum or difference. One parent is occasionally zero padded so the child can move more atoms.

We begin each run at a high temperature, with the moves distributed randomly from 0 to 0.3 lattice spacing in length, then let the population evolve as the FK chain slowly cools. It quickly adjusts to an appropriate length scale for the prevailing temperature, but with an excess of long moves for hopping atoms between wells, as shown in Fig. 3. Upon cooling, the thermal distribution sharpens and the enzymes become separate, no longer just an anomalously long tail to the ordinary thermal peak. By the time the chain is cold enough that *ss* defects have equilibrated out, the population discovers that the enzymes must move at least three atoms at once, hopping one atom a large distance across the barrier and relaxing two neighbors. This is the needed move for equilibrating out *lll* defects. At still colder temperatures, the enzymes find that they need to move more and more atoms at once. If they fail to discovery some necessary improvement, they will be unable to continue equilibrating defects out as the chain cools, and will become extinct, leaving only thermal moves in the population.

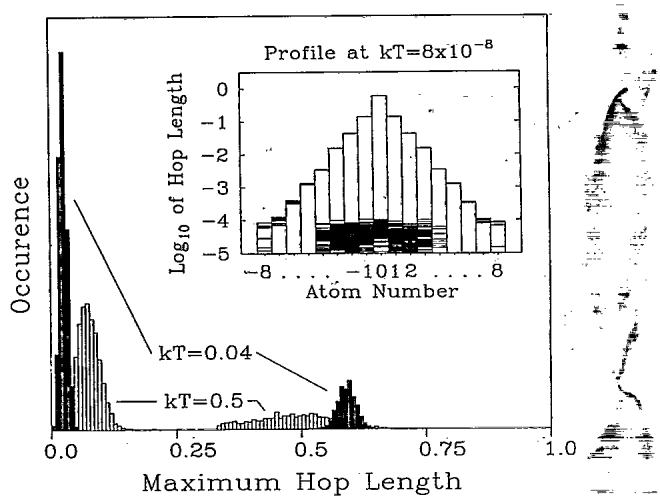

FIG. 3. As illustrated by the histograms, enzymes first develop as an extension to the thermal move distribution, but thrive and adapt as the system cools and specific moves between wells become important. They gradually improve, moving more and more atoms simultaneously and ever more precisely. Inset: A plot of an entire population at low temperature with all moves superimposed. Each of the enzymes here moves sixteen atoms, moving one atom (plotted as number 0) a distance of 0.5949, and moving the neighbor to the right (atom number 1) and to the left (number −1), 0.1385 in the same direction, etc. The thermal moves involve six to eight atoms each, with distances distributed randomly up to about 8×10^{-5}. The unit of length here is the periodicity of the applied potential.

So long as we cool slowly enough and set all the parameters to suitable values, then long-range, fine-tuned enzymes usually develop that are capable of equilibrating out even very-low-energy defects. The inset of Fig. 3 shows the 15-atom enzymes discovered by a run with $K/V=1/3$. When we added this enzyme to a proper Monte Carlo code, we were able to equilibrate to lower temperature and longer correlation length in half an hour of computer time than we could have in years or even centuries using the unassisted algorithm. We have had similar success with lower values of K/V.

While this method is quite robust and not particularly sensitive to details in the algorithm, some thought must be given to ensure that it generally encourages survival of the enzymes. For example, the relative scarcity of low-energy defects means that enzymes succeed infrequently at low temperatures; they simply cannot compete. We eliminate this bias by dividing the population into two classes, those that move at least one atom at least a third of a lattice spacing and those that do not, and multiplying the health of each individual by a normalization factor to promote equality between the classes. Of course, no amount of special treatment will help the enzymes if they fail to adapt to the changing environment and become unable to ever lower the energy of the chain.

Mutations are necessary to counter the otherwise monotonic loss of diversity, so we occasionally multiply one hop by a random number when a new individual is born. If the mutation rate is too low, the population may collapse and be taken over by a few strong moves, who can no longer adapt because they have only copies and multiples of themselves to reproduce with. Too high a mutation rate, however, randomizes the birth process and keeps healthy traits from propagating.

Various input parameters need to be reasonable, although fine tuning is unnecessary. The death rate must be fast enough for the population to adapt to the changing environment, but not so high that good moves get killed off along with the weak. Also, since health is determined dynamically, it is important to allow sufficient time between deaths to give newborns a chance to prove their worth.

The rate at which *successful* moves can be created determines how fast the environment may change without devastating the population. By intelligently biasing the

reproduction scheme to reduce the birth rate of moves that are likely to be poor, we can save on computer time. For the FK model, we know that the offspring of two enzymes is almost always a failure, so we have explicitly required enzymes to mate only with thermal moves. For other models, guidance of a different kind might be given.

In conclusion, we have found that diverging relaxation time scales prevent the FK model from equilibrating all the way to its zero-temperature ground state given a finite cooling rate, and that much slower cooling results in equilibration to only slightly lower temperature. We have developed a method for accelerating relaxation in such a system in order to study the equilibrium behavior numerically. To do this, we use a modified Monte Carlo–like algorithm in which we keep track of the success of various moves, retaining and refining those that are especially useful for reducing the energy of the chain. We use the enzymes, or precisely coordinated motion of many atoms at once, that are discovered by the modified program to dramatically accelerate equilibration in a legitimate Monte Carlo code. The method should be especially useful in studying complicated models where the ground states and equilibrium properties cannot be deduced analytically.

This work was supported by Grant No. NSF-DMR-88-18558-A02 from the Cornell University Materials Science Center and an NSF graduate fellowship (S.L.S.), and computer time was provided through the Cornell-IBM Joint Study on Computing for Scientific Research. S.L.S. thanks Steve Girvin and Allan MacDonald for their hospitality and Arthur Smith for helpful discussions.

[1] Y. I. Frenkel and T. Kontorova, Zh. Eksp. Teor. Fiz. **8**, 1340 (1938).

[2] We cite here only a few recent papers: M. Mansfield and R. J. Needs, J. Phys. Condens. Matter **2**, 2361 (1990); Y. Braiman, J. Baumgarten, Joshua Jortner, and J. Klafter, Phys. Rev. Lett. **65**, 2398 (1990); Ofer Biham and David Mukamel, Phys. Rev. A **39**, 5326 (1989); F. Vallet and S. Aubry, J. Phys. C **21**, 67 (1988).

[3] The pinning transition is discussed in, e.g., S. N. Coppersmith and D. S. Fisher, Phys. Rev. B **28**, 2566 (1983), and Braiman *et al.* [Ref. [2]).

[4] F. Vallet, R. Schilling, and S. Aubry, Europhys. Lett. **2**, 815 (1986).

[5] Michael Widom, Katherine J. Strandburg, and Robert H. Swendsen, Phys. Rev. Lett. **58**, 706 (1987).

[6] Robert H. Swendsen and Jian-Sheng Wang, Phys. Rev. Lett. **58**, 86 (1987).

[7] J. Holland, *Adaptation in Natural and Artificial Systems* (Univ. Michigan Press, Ann Arbor, 1975).

[8] See, e.g., David E. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning* (Addison-Wesley, Reading, MA, 1989).

[9] Thomas P. Meyer, Fred C. Richards, and Norman H. Packard, Phys. Rev. Lett. **63**, 1735 (1989).