36. K. W. Madison *et al.*, preprint available at http://xxx.lanl.gov/abs/cond-mat/9912015.
37. S. Burger *et al.*, preprint available at http://xxx.lanl.gov/abs/cond-mat/9910487.

38. We thank K. Burnett, M. A. Edwards, and K. Jones for discussions. Supported in part by the U.S. Office of Naval Research and NASA. J.D., J.E.S., and W.P.R. acknowledge support from the Alexander von Humboldt Foundation, a NIST–National Research Council fellowship, and NSF, respectively.

15 October 1999; accepted 2 December 1999

REPORTS

Equilibrium Regained: From Nonequilibrium Chaos to Statistical Mechanics

David A. Egolf

Far-from-equilibrium, spatially extended chaotic systems have generally eluded analytical solution, leading researchers to consider theories based on a statistical rather than a detailed knowledge of the microscopic length scales. Building on the recent discovery of a separation of length scales between macroscopic behavior and microscopic chaos, a simple far-from-equilibrium spatially extended chaotic system has been studied computationally at intermediate, coarse-grained scales. Equilibrium properties such as Gibbs distributions and detailed balance are recovered at these scales, which suggests that the macroscopic behavior of some far-from-equilibrium systems might be understood in terms of equilibrium statistical mechanics.

Statistical mechanics describes the macroscopic physical properties of matter through a probabilistic, rather than a detailed, knowledge of the microscopic dynamics and has been applied successfully to a wide variety of equilibrium systems, from simple molecular gases to white dwarf stars. It has provided a theoretical understanding of the phases of matter, the transitions between phases, and the deep property of universality that unifies the descriptions of continuous transitions in systems that are physically quite distinct (for example, magnets and gases). In nature, however, many systems are not in equilibrium, including, for example, large-scale flows in the atmosphere, the evolution of ecological systems, and the transport of energy in cells. None of these situations can be understood with equilibrium statistical mechanics.

Although theory has been developed to extend equilibrium statistical mechanics to systems only slightly perturbed away from equilibrium (for which the evolution of the system is well-approximated with only linear terms), in deterministic systems driven far from equilibrium (where nonlinearities are important) theoretical progress has been limited to "simple" situations, such as the onset of symmetry breaking, the stability of perfect patterns, and the motions of single topological defects in perfect patterns (*1*). Theorists have not yet developed an understanding of the intriguing phenomenon of "spatiotemporal chaos" (or spatially extended chaos) that is typically characterized by disordered arrays of defects, patches of uncorrelated regions, and a chaotic dynamics that persists indefinitely (*2*). This remarkable behavior has been found in large, deterministic, far-from-equilibrium systems as varied as convecting horizontal fluid layers (*3*), chemical reaction-diffusion systems (*4*), colonies of microorganisms (*5*), and fibrillating heart tissue (*6*). These disparate systems often display strikingly similar macroscopic features (such as locally ordered striped or hexagonal patterns and dislocation, spiral, and target defects) and behavior (for example, dramatic qualitative changes in response to modifications of experimental parameters reminiscent of phase transitions in equilibrium systems). Such behavior within a system and the similarities between different systems beg the question of whether one can construct a statistical, predictive theory of phases and transitions in these chaotic, far-from-equilibrium systems.

At first glance, far-from-equilibrium, strongly dissipative, deterministic systems may appear to have little in common with equilibrium systems; for example, at the detailed level, these systems do not have the benefit of tending toward the minimum of a free-energy functional, do not have a Gibbsian distribution of states, and do not allow the calculational technique of averages over noise terms. However, several experimental and computational studies have explored the similarities in the behaviors of these systems and the behaviors of equilibrium systems. A particular focus has been the possibility of phase transition–like behavior in these systems (*4*, *7*–*11*). The data reported here uncover a deeper level of similarity and suggest the possibility of salvaging much of the framework of equilibrium statistical mechanics. In particular, large-scale computational studies of a simple, large, chaotic, far-from-equilibrium system demonstrate that several cornerstones of equilibrium statistical mechanics—ergodicity, detailed balance, Gibbs distributions, partition functions, and renormalization group flows of coupling constants—are recovered at a coarse-grained scale.

In analogy to the simple explorations of equilibrium statistical mechanics with the Ising model, one of the simplest spatially extended chaotic systems was used as a test bed (*12*). This system, a coupled map lattice (CML) first studied by Miller and Huse (*8*), consists of a set of scalar variables $u^t_{\vec{x}}$ at integer time t on a square two-dimensional spatially periodic $L \times L$ grid with positions indicated by $\vec{x} = a\hat{x} + b\hat{y}$, where a and b are integers and $\hat{x}$ and $\hat{y}$ are the unit vectors of the two-dimensional lattice. The rule for updating the variables from time t to $t + 1$ is

$$u^{t+1}_{\vec{x}} = \phi(u^t_{\vec{x}}) + g\sum_{\vec{y}(\vec{x})}[\phi(u^t_{\vec{y}}) - \phi(u^t_{\vec{x}})] \tag{1a}$$

where g indicates the strength of the spatial coupling, and $\vec{y}(\vec{x})$ denotes nearest neighbors of site $\vec{x}$. The chaotic local map $\phi(u)$ is given by

$$\phi(u) = \begin{cases} -3u - 2 & -1 \le u \le -\frac{1}{3} \\ 3u & -\frac{1}{3} < u < \frac{1}{3} \\ -3u + 2 & \frac{1}{3} \le u \le 1 \end{cases} \tag{1b}$$

This CML exhibits chaotic, spatially disordered dynamics for values of g at least within the range [0, 0.25]. Miller and Huse (*8*) reported that at $g_c \approx 0.2054$, this system undergoes a paramagnetic-to-ferromagnetic transition exhibiting a number of features in common with the equilibrium transition in the Ising ferromagnet (*13*).

To study the statistical bulk properties of spatially extended chaotic states ["extensive chaos" (*1*, *14*)], the "thermodynamic limit" of systems approaching infinite size was taken. O'Hern *et al.* (*15*) demonstrated that the behavior of Eq. 1 can be considered extensive for system sizes as small as $L \approx 9$. Results reported here were obtained for system sizes ranging from 1×1 to 1024×1024 over times as large as 10^{10} iterations (after typically 10^6 iterations of transient), often averaged over ensembles of up to 256 systems with identical parameters but differing initial conditions [with each site $u^{t=0}_{\vec{x}}$ initialized to a

Center for Nonlinear Studies (MS B258), Theoretical Division and Condensed Matter and Thermal Physics, Los Alamos National Laboratory, Los Alamos, NM 87545, USA. E-mail: egolf@cnls.lanl.gov

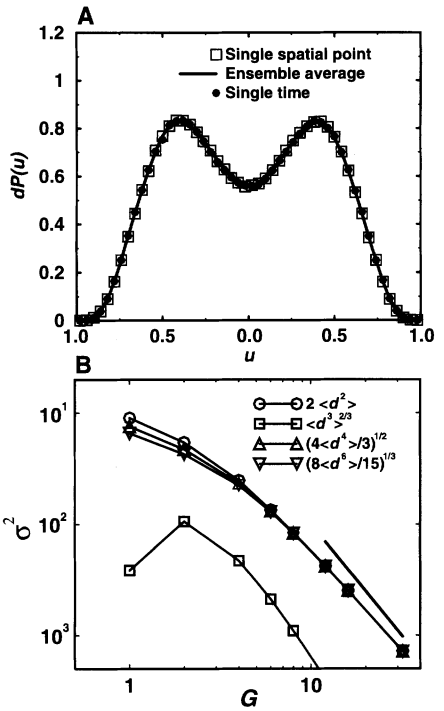

Fig. 1. (**A**) Demonstration of ergodicity and self-averaging. Data were obtained from simulations of Eq. 1 with $g = 0.195$. Open squares and the solid line are probability distributions of time-dependent measurements of $u^t_{\vec{x}_0}$, with $\vec{x}_0 = (300, 500)$ on a 512×512 over 4,000,000 iterations after a transient of 100,000 iterations, for a single system (squares) and an ensemble of eight systems (line). Solid circles show the probability distribution of all site values $u^{t_0}_{\vec{x}}$ for a 1024×1024 lattice at time $t_0 = 200{,}000$ iterations. (**B**) Dependence on coarse-grain size G of the 2nd (circles), 3rd (squares), 4th (triangles), and 6th (upside-down triangles) moments of the fluctuations of $\langle|u|\rangle_g$ about its mean. The even moments are scaled such that if the fluctuations are Gaussian, the moments will be equal. (The odd moments should vanish.) The solid line is a guide to the G^{-2} fall-off. Data for larger coarse-grain sizes (not shown on this plot) continue to fall off as G^{-2}.

random number chosen from $(-1, 1)$].

One property of equilibrium systems that is perhaps not surprising to find in spatiotemporal chaotic systems that are far from equilibrium is ergodicity. A system is ergodic if the infinite-time average of an observable is independent of the initial condition (except for a set of initial states of measure zero). Indeed, a certain class of chaotic CMLs with weak coupling ($g \ll 1$ for Eq. 1a) has been proven to exhibit ergodicity (*16*). Unfortunately, this proof has not been extended to stronger coupling. Figure 1A suggests that Eq. 1 is ergodic even for strong couplings. Rather than reporting simple time-averages, Fig. 1A shows the distribution of values of the observable $u^t_{\vec{x}_0}$ obtained at a single site $\vec{x}_0$ over a large number of iterations for a single system (open squares) and an ensemble of systems (solid line). The overlapping distributions imply that the system is ergodic, at least for observables that depend only on averages over individual site values (rather than on more complicated combinations of site values).

Figure 1A illustrates another feature of Eq. 1. The distribution indicated by the filled circles, collected from all sites of a large system at a single instant in time, suggests that the system is self-averaging—for a large enough lattice, the single system is equivalent to an ensemble of independent systems (or a single system over a long period of time). The distributions in Fig. 1A are statistically stationary in that after a transient period, the distributions do not depend on time. The properties of ergodicity and self-averaging hold at least for $0.18 < g < g_c = 0.2054$. Above g_c, Eq. 1 exhibits broken ergodicity, such that any system remains in either of two regions of phase space (at least in the infinite system size limit), depending on the initial state—a positive "magnetization" or a negative "magnetization." This same broken ergodicity is observed in equilibrium continuous-phase transitions.

Equilibrium systems are often assumed to obey the principle of detailed balance:

$$P_r\, R_{r\to s} = P_{-s}\, R_{-s\to -r} \qquad (2)$$

where P_r is the (stationary) probability of a particular system (out of a large ensemble of systems) being in state r, $R_{r\to s}$ is the rate at which a system in state r makes a transition to state s, and $-r$ represents the time-reversed state corresponding to state r (for example, in particle systems this would indicate the state with all of the momenta reversed). For certain cases, detailed balance can be proven to be a necessary and sufficient condition for equilibrium (*17*). Often a proof of detailed balance can be provided by using the underlying reversibility of the microscopic processes involved. However, this microscopic reversibility is not a necessary condition for the observation of detailed balance at scales coarser than that of the microscopic dynamics. In particular, the microscopic irreversibility of dissipative, far-from-equilibrium systems does not preclude the observation of detailed balance, at least in states coarser than the microscopic scales.

The spatiotemporal dynamics of Eq. 1 suggests a natural scale for studying properties such as detailed balance. A recent series of papers has demonstrated the existence of a separation of length scales in several far-from-equilibrium spatiotemporal chaotic systems (*9*, *14*, *15*, *18*, *19*). For these systems, a chaotic length scale was found to be much shorter than the typical macroscopic length scale. This separation of scales suggests the consideration of a statistical description of a "coarse-grained" intermediate-scale dynamics, with the fine-scale chaos acting as an effective noise or temperature bath. [This scenario is similar to the assumption of molecular chaos (*20*) in gases.] O'Hern *et al.* (*15*) found two length scales within large systems described by Eq. 1—a macroscopic scale ξ_u and a microscopic scale ξ_c. The length scale ξ_u describes the decay of the two-point equal-time correlations in sign(u) and diverges at the apparent critical point $g_c \approx 0.2054$. The microscopic length scale ξ_c is on the order of one to two lattice spacings over a wide range of coupling strengths g (including g_c), describes the decay of correlations in $|u|$, and agrees with a length scale characterizing the chaotic dynamics (*1*, *9*, *15*).

Figure 1B shows that spatial averages $\langle|u|\rangle_G$ over sublattices of size $G \times G$ approximate Gaussian-distributed white noise for coarse-grain regions as small as 6×6 lattice sites. For a Gaussian distribution, the even moments of the temporal and spatial fluctuations in $\langle|u|\rangle_G$ about its mean $\langle|u|\rangle_{x,t}$ (scaled by the factors used in Fig. 1B) should be equal, and the odd moments should be zero. Figure 1B also shows that $\langle|u|\rangle_G$ is delta-function–correlated in space because the mean-square fluctuations fall off as G^{-2} for large values of G (equivalent to averaging independent Gaussian samples). Similar plots for temporal coarse-graining show that $\langle|u|\rangle_G$ is delta-function–correlated in time for large enough temporal coarse-grains. For the data shown below at $g = 0.204$, a coarse-grain size $G = 16$ was used. This grain size is smaller than the macroscopic length scale $\xi_u = 24$, but large enough that the effective noise is well described by Gaussian-distributed delta-function–correlated white noise. For the tests described below, a variety of coarse-grain sizes were tested; sizes $G \geq 8$ all yielded similar results (*21*).

To test detailed balance at coarse-grained scales, the dynamics of a $GN \times GN$ lattice described by Eq. 1 was coarse-grained into a new regular, periodic $N \times N$ lattice with site variables $\tilde{u}^T_{\vec{x}} = \pm 1$ and an internal hidden chaotic dynamics characterized by fluctuations in $|u|$ at each site. Each site on the coarse-grained lattice represents one of the N^2 distinct $G \times G$ regions of the original lattice. The values of the coarse-grained variables are given by $\tilde{u}^T_{\vec{x}} = \text{sign}\,(\langle u^T_{\vec{y}}\rangle_{CG(\vec{x}),\Delta T})$, where $CG(\vec{x})$ is the coarse-grain region corresponding to position $\vec{x}$, and $\langle\rangle_{\Delta T}$ represents an average over the times $[T, T + \Delta T)$ with $\Delta T = 100$, typically (*22*). An example of the coarse-graining (with $N = 16$ and $G = 16$) is shown in Fig. 2. Each of the $2^{(N^2)}$ possible coarse-grained states is identified by a particular arrangement of pluses and minuses on the $N \times N$ coarse-grained lattice. In this report, the coarse-grained dynamics was studied with $N = 4$, yielding 2^{16} possible states (*23*).

As described above, the coarse-grained dynamics of Eq. 1 is said to obey detailed balance if, over a large number of iterations (or a large ensemble, or both), the number of transitions from state i to state j, $T_{i\to j} = P_i R_{i\to j}$, is balanced by the number of transitions from state j to state i, $T_{j\to i} = P_j R_{j\to i}$ (for this system, a state i and its time-reversal $-i$ are the same). However, a measurement of $\Delta = T_{i\to j} - T_{j\to i}$ performed over only a finite period of time (or on a finite ensemble of systems) will be only approximate-

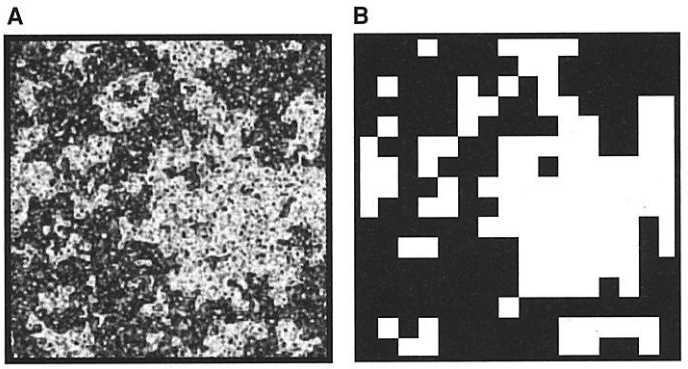

Fig. 2. (**A**) Spatial field $u_{\vec{x}}^t$ of Eq. 1 and (**B**) coarse-grained field $\tilde{u}_{\vec{x}}^t$ at $t = 50{,}000$ and $g = 0.204$ for a lattice of size 256×256 and coarse-grain size 16×16 ($G = 16$). The color range from black to white represents field values from -1 to $+1$.

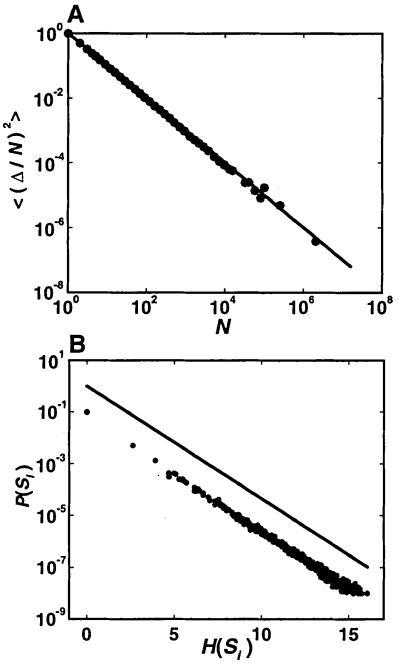

Fig. 3. (**A**) Relation between the number of transitions $N = T_{i\to j} + T_{j\to i}$ and the square relative error $\langle(\Delta/N)^2\rangle$ (where $\Delta = T_{i\to j} - T_{j\to i}$) for 10^6 different pairs of states i and j using Eq. 1 with $g = 0.204$ on a square periodic lattice of size 64×64 with coarse-grain size $G = 16$ obtained over an interval $T = 3.5 \times 10^8$ from an ensemble of 256 systems. The averages are taken over all pairs of states with transition counts within bins that increase in size exponentially with the number of transitions. The solid line indicates the expected relation for perfect detailed balance. (**B**) Relation between the probability of a given state S_i and its effective energy for the 65,160 (out of 65,536) states with probabilities above 10^{-8} for the same conditions as in (A). The effective energy used only the first four terms of Eq. 4, with parameters $\alpha(d_{i,j})$ determined through a least-squares fit: $\alpha(1) = 0.685$, $\alpha(\sqrt{2}) = 0.100$, $\alpha(2) = -0.075$, and $\alpha(\sqrt{5}) = -0.049$. The solid line is a guide to the eye for the expected power law for a Gibbs distribution.

ly zero for a finite number of transitions $N = T_{i\to j} + T_{j\to i}$. If each of the transitions $i \to j$ and $j \to i$ are independent, then we expect the root-mean-square measurement error $\langle(\Delta/N)^2\rangle^{1/2}$ (averaged over many measurements of Δ) to shrink as $N^{-1/2}$:

$$\left\langle \left(\frac{\Delta}{N}\right)^2 \right\rangle = N^{-1} \qquad (3)$$

where the average is taken over measurements of Δ for many transitions $i \to j$, each with approximately the same value of N. Figure 3A shows that over six decades in the number of transitions N, Δ approaches zero in the manner described by Eq. 3. Thus, coarse-grained detailed balance is present for the far-from-equilibrium system characterized by Eq. 1 despite the underlying lack of true microscopic detailed balance.

The presence of detailed balance at coarse-grained scales suggests that Eq. 1 might also obey a coarse-grained Hamiltonian. The space of possible Hamiltonians is large; however, the underlying reflection symmetry $f(u) = -f(-u)$ in Eq. 1 suggests a Hamiltonian with Ising symmetry:

$$\mathcal{H} = \sum_{\vec{x},\vec{y}} \alpha(|\vec{x} - \vec{y}|)\tilde{u}_{\vec{x}}^T\tilde{u}_{\vec{y}}^T \qquad (4)$$

where the sum is over all possible pairs of spins and $\alpha(d)$ are the (reduced) couplings between spins at $\vec{x}$ and $\vec{y}$, separated by a distance $d = |\vec{x} - \vec{y}|$. (Only pairwise interactions of "spins" $\tilde{u}_{\vec{x}}^T$ were included.) The couplings $\alpha(d)$ were determined by measuring the probability of each of the 2^{16} states and performing a multidimensional fit to a Gibbs distribution $P(S_i) = \exp[-\mathcal{H}(S_i)]/Z$, where $Z = \Sigma_{S_i} \exp[-\mathcal{H}(S_i)]$ is the partition function. The validity of the extracted couplings and of the Gibbs distribution itself is shown in Fig. 3B demonstrating a Gibbsian relation between the effective energy $\mathcal{H}(S_i)$ of a coarse-grained state S_i and its probability $P(S_i)$ extending over six decades. Thus, the coarse-grained system is obeying an underlying Hamiltonian of the form Eq. 4 and the system behaves as if it were in contact with a heat reservoir of some temperature T (that is, as a canonical ensemble). The presence of a coarse-grained Hamiltonian also allows the calculation of the partition function Z and the host of thermodynamic quantities that depend on Z. Detailed balance, effective Hamiltonians, and partition functions were present on both sides of the critical point $g_c \approx 0.2054$ for a variety of values g.

The behavior of the reduced couplings $\alpha(d)$ as a function of the coarse-grain size G is at least qualitatively consistent with the behavior of couplings in equilibrium systems in the Ising universality class. As the size G is increased, the non-nearest-neighbor couplings quickly approach zero (become irrelevant), with the couplings of farther neighbors approaching zero more quickly. The nearest-neighbor coupling $\alpha(1)$ also behaves as expected—on the paramagnetic side of the transition $g < g_c$, $\alpha(1)$ flows toward a fixed point of $\alpha(1) = 0$ (equivalent to temperature $T \to \infty$), whereas on the ferromagnetic side $g > g_c$, $\alpha(1)$ flows toward the zero-temperature fixed point $\alpha(1) \to \infty$. Future large-scale computational studies are necessary to determine whether this behavior is also quantitatively consistent with the equilibrium behavior.

Simple, far-from-equilibrium, dissipative, extensively chaotic systems can recover the equilibrium properties of ergodicity, detailed balance, Gibbs distributions, partition functions, and renormalization group flow at coarse-grained scales with the underlying chaotic dynamics serving as a temperature bath. This remarkable result suggests that the long-wavelength behavior of some far-from-equilibrium systems can be understood by using the powerful tools of equilibrium statistical mechanics. Studies are needed to explore the underlying mechanism for the recovery of equilibrium and the range of far-from-equilibrium systems that are susceptible to such analysis. Systems with multiple length scales and that recover symmetries at the coarse-grained scales might exhibit such behavior. Additional systems might be understood through familiar techniques of statistical mechanics that describe small excursions away from equilibrium. The system studied here possesses some important differences from true equilibrium systems. Perhaps the most intriguing is that the effective noise strength (or temperature) is internally generated and dependent on the state of the system, rather than imposed by an external temperature bath. This difference poses a challenge for explorations of the second law of thermodynamics in these systems.

References and Notes

1. M. C. Cross and P. C. Hohenberg, *Rev. Mod. Phys.* **65**, 851 (1993).
2. ———, *Science* **263**, 1569 (1994).
3. See, for example, S. W. Morris, E. Bodenschatz, D. S. Cannell, G. Ahlers, *Phys. Rev. Lett.* **71**, 2026 (1993).
4. See, for example, Q. Ouyang and J.-M. Flesselles, *Nature* **379**, 143 (1996).
5. See, for example, K. J. Lee, E. C. Cox, R. E. Goldstein, *Phys. Rev. Lett.* **76**, 1174 (1996).
6. See, for example, A. Garfinkel, M. L. Spano, W. L. Ditto, *Science* **257**, 1230 (1992); R. A. Gray *et al.*, *Science* **270**, 1222 (1995); Z. L. Qu, J. N. Weiss, A. Garfinkel, *Phys. Rev. Lett.* **78**, 1387 (1997).
7. B. I. Shraiman *et al.*, *Physica D* **57**, 241 (1992).
8. J. Miller and D. A. Huse, *Phys. Rev. E* **48**, 2528 (1993).
9. D. A. Egolf and H. S. Greenside, *Phys. Rev. Lett.* **74**, 1751 (1995).

10. See, for example, R. E. Ecke, Y. Hu, R. Mainieri, G. Ahlers, *Science* **269**, 1704 (1995); R. V. Cakmur, D. A. Egolf, B. B. Plapp, E. Bodenschatz, *Phys. Rev. Lett.* **79**, 1853 (1997); K. M. S. Bajaj, D. S. Cannell, G. Ahlers, *Phys. Rev. E* **55**, R4869 (1997); H.-W. Xi, X. J. Li, J. D. Gunton, *Phys. Rev. Lett.* **78**, 1046 (1997).
11. P. Marcq, H. Chaté, P. Manneville, *Phys. Rev. Lett.* **77**, 4003 (1996); *Phys. Rev. E* **55**, 3228 (1997).
12. The particular map lattice studied here (Eq. 1) was chosen because of extensive previous work (*15*) uncovering the separation of length scales in the system and because of its symmetry properties. Further work and insight (concerning properties such as the separation of variables into coarse- and fine-grain contributions and the symmetries of the Hamiltonian terms) will be needed to expand the techniques in this report to other popular systems such as a couple lattice of simpler single tent maps.
13. Marcq and co-workers (*11*) have reported that all of the critical exponents of the transition do not agree with the Ising universality class. Rather, some exponents depend on the update rule (for example, Eq. 1a). The results presented here were not tested with the alternate update rule.
14. D. A. Egolf and H. S. Greenside, *Nature* **369**, 129 (1994).
15. C. S. O'Hern, D. A. Egolf, H. S. Greenside, *Phys. Rev. E* **53**, 3374 (1996).
16. L. A. Bunimovich and Y. G. Sinai, *Nonlinearity* **1**, 491 (1988).
17. R. H. Fowler, *Statistical Mechanics: The Theory of the Properties of Matter in Equilibrium* (Cambridge Univ. Press, London, ed. 2, 1955), pp. 659–660.
18. T. Bohr, E. Bosch, W. van de Water, *Nature* **372**, 48 (1994).
19. D. A. Egolf, I. V. Melnikov, E. Bodenschatz, *Phys. Rev. Lett.* **80**, 3228 (1998).
20. D. Durr and H. Spohn, *Nature* **394**, 831 (1998); P. Gespard *et al.*, *Nature* **394**, 865 (1998).
21. Detailed balance and effective Hamiltonians appear for $G \geq 8$. At $G = 8$, finite-size effects are noticeable in the tails of plots such as those in Fig. 3. For $G < 8$, it is not clear whether the breakdown of detailed balance is due to finite-size effects or the non-Gaussian nature of the noise. The requirement $G < \xi_2$ does not appear to be necessary.
22. For values of $\Delta T < 100$, the probabilities of states is still as shown in Fig. 3B; however, the transition rates leading to plots such as Fig. 3A actually appear to give better agreement than expected simply due to the extra crossings as the system wiggles from one state to the next. Temporal coarse-grainings of $\Delta T \geq 100$ create a smoother (less "noisy") transition between states.
23. The 2^{16} states is a small enough number such that a large computational effort yields a large fraction of the states being visited often enough to give statistically significant results. Significantly larger lattices have too many states for meaningful simulations on today's computers. A coarse-grained lattice size of 4×4 is also just large enough such that the dynamics is not overwhelmed by finite-size effects for a wide range of coupling values g.
24. I thank R. Ecke, C. Henley, R. Mainieri, I. Melnikov, and J. Socolar for valuable discussions. This research was supported by the United States Department of Energy under contract W-7405-EN9-36, and the computations were performed on the Nirvana machines of the Advanced Computing Laboratory at Los Alamos National Laboratory and on the Avalon cluster at the Center for Nonlinear Studies.

9 August 1999; accepted 18 November 1999

A Nucleation Site and Mechanism Leading to Epitaxial Growth of Diamond Films

S. T. Lee,* H. Y. Peng, X. T. Zhou, N. Wang, C. S. Lee, I. Bello, Y. Lifshitz†

A diamond nucleation site responsible for epitaxial growth of diamond on silicon by chemical vapor deposition (CVD) is identified in high-resolution transmission electron microscopic images. Other sites in the same sample leading to polycrystalline growth, but deleterious to epitaxial CVD growth, are also described. A mechanism for the heteroepitaxial growth of diamond is suggested, in which etching of the nondiamond carbon binder exposes and removes nonadherent nanodiamond nuclei, leaving intact only those directly nucleated on the silicon substrate. This work enhances our understanding of diamond nucleation and heteroepitaxial growth and its potential applications.

The quest for artificial methods of diamond production is motivated not only by its gemstone quality, but also by its unique set of properties, which make it an excellent candidate for numerous important applications (*1–3*). Diamond was successfully produced in the 1950s by the high-pressure, high-temperature (HPHT) method (*1*, *2*). An alternative method, CVD of diamond at low pressure (typically with the use of an excited CH_4/H_2 mixture on substrates held at ~700° to 800°C), has also been applied successfully over the last 15 to 20 years (*1–3*). The homoepitaxial growth of diamond on a diamond substrate by CVD methods is relatively well understood (*1–3*). Experimental methods for diamond nucleation on nondiamond substrates (the first necessary step of a heteroepitaxial growth process) have also been developed (*4–7*). The most effective method uses bias-enhanced nucleation (BEN) (*4*, *6–10*), in which the target is biased, with a relatively methane-rich CH_4/H_2 mixture (several percent methane) as a first step, followed by a conventional CVD step (typically 1% methane or less). However, the nucleation mechanism of diamond on nondiamond substrates remains poorly understood (*11*), largely because of the tremendous difficulty of locating and identifying the nucleation sites. This is a major obstacle to further advances in diamond science and technology.

Here we present direct high-resolution transmission electron microscopy (HRTEM) evidence that a step on a single crystalline Si surface serves as a nucleation site for heteroepitaxial diamond growth, and we propose a scheme for the growth of epitaxial diamond films on Si wafers.

A BEN treatment with a double-bias–assisted hot filament CVD was used (*12*). In this process, a negative-bias voltage is applied to the Si substrate and a positive-bias voltage is applied to a grid that is placed on top of the hot filament (*13*). The samples were analyzed by HRTEM (*13*), high-resolution scanning electron microscopy (HRSEM), and micro-Raman spectroscopy. Transmission electron microscopy (TEM) observations were made along the [110] direction of the Si substrate.

HRSEM images of the samples show a rough, granular morphology, with an average grain size of about 200 nm. Raman spectroscopy shows a dominant graphitic structure that has no detectable 1330 cm^{-1} diamond peak, but that does have a small peak indicative of "nanodiamond precursors" at ~1100 cm^{-1}. Cross-sectional HRTEM (Fig. 1) shows a grooved Si morphology onto which a predominantly amorphous C (a-C) film is deposited, which explains the Raman data. Diamond crystallites with diameters of about 2 to 6 nm (small enough to identify the nucleation sites) are either embedded in the a-C matrix (white arrows in Fig. 1) or attached to different sites of the Si (black arrows). Selected-area electron diffraction patterns confirmed the diamond structure of the nanocrystallites. No SiC crystallites were found. The diamond crystallites grew randomly (Fig. 2), partially epitaxially (Fig. 3), or perfectly heteroepitaxially (Fig. 4) with respect to the Si surface. In Fig. 3, a set of diamond {111} planes is parallel to the Si {111} planes, which indicates a partially oriented diamond nucleus on Si. The interface between the nucleus and Si near the (001) plane of Si could not be resolved. In Fig. 4, two nuclei have grown epitaxially on stepped areas of the Si substrate. Five to 10 such nuclei have been observed in each of the six samples studied. The diamond crystallites were identified by measuring the spacings of the lattice fringes and the angles of the intersecting lattice planes. This measurement is very precise because the Si (111) lattice in the same image can be used as an internal reference. In Fig. 4A, the interfaces between the diamond crystallite

Department of Physics and Materials Science and Center of Super-Diamond and Advanced Films, City University of Hong Kong, Hong Kong, China.

*To whom correspondence should be addressed. E-mail: apannale@cityu.edu.hk

†On leave from Soreq Nuclear Research Center, Yavne 81800, Israel.