Journal of Statistical Physics, Vol. 19, No. 3, 1978

Statistical Mechanics of a Dynamical System Based on Conway's Game of Life

L. S. Schulman[1,2] and P. E. Seiden[1,3]

Received October 27, 1977

We study a discrete dynamical system whose evolution is governed by rules similar to those of Conway's game of Life but also include a stochastic element (parametrized by a "temperature"). Statistical properties that are examined are density as a function of temperature and entropy (suitably defined). A phase transition and a certain "thermodynamic" constant of the motion are observed.

KEY WORDS: Discrete dynamical systems; stochastic games.

1. INTRODUCTION

In this paper we study the statistical mechanics of a discrete, stochastic dynamical system. The system is a two-dimensional array of squares each having two states designated "living" or "dead." The deterministic part of the dynamics is that of the game of "Life" invented by J. Conway.[1] Although this system is not related in any obvious way to a specific physical or biological system, there are various considerations which have led us to explore its properties. Our motivations fall into a number of different categories: (A) Physical—microdynamics of a nonequilibrium system; (B) biological—reproduction and evolution; (C) pattern formation and its relation to both physical and biological formalisms. We shall discuss each of these ideas in turn.

[1] Department of Physics, The Technion, Haifa, Israel, and IBM Thomas J. Watson Research Center, Yorktown Heights, New York.

[2] Permanent addresses: Department of Physics, the Technion, Haifa, Israel, and Physics Department, Indiana University, Bloomington, Indiana.

[3] Lady Davis Visiting Scientist at the Technion 1974–75. Permanent address: IBM Research Center, Yorktown Heights, New York.

0022-4715/78/0900-0293$05.00/0

In condensed matter physics, molecular dynamics has been successful in explaining the behavior of many phenomena and materials.[2] These systems are generally closed and have a Hamiltonian dynamics. We have sought to extend these techniques to other domains, such as the Life game, which is not a Hamiltonian system, has no energy function, and is completely nonlinear. The advantage of using molecular dynamics is that, in principle, the interaction can be treated exactly within the context of the model chosen. This is in contrast to the usual differential equations techniques, where increasing the complexity of the interaction greatly complicates the set of nonlinear differential equations usually involved and therefore makes their solution even more difficult. The price one pays for this is that the molecular dynamics technique is restricted to finite, and sometimes small, sample sizes. Whether or not this causes difficulties depends on the details of the problem being investigated and must be examined for each case.

That the system might have some biological significance is hinted at by the very name of the game: Life. However, one must not take this too literally, and even thinking biologically the interpretation of the living or dead squares can be as macromolecules, genes, cells, whole animals, or even ethnically distinct home buyers in an integrated neighborhood.

Much of the biological interest is sparked by the existence of self-reproducing forms. One might ask if there is anything to be learned about evolution from this system. The self-reproducing forms could be studied as to their robustness when perturbed, their likelihood of appearance under random or other conditions, or their survival when in competition with other "organisms" (i.e., self-reproducing forms). As reported in Ref. 1, these questions had been investigated by examination of the evolution of many individual systems. We wish to find ways to extract systematic information without exhaustively studying large numbers of arrays. In particular we seek ways of getting statistical indications of the emergence of ordered or otherwise interesting patterns.

This leads to what is perhaps the most ambitious of our goals, namely the use of physical techniques in fundamental biological questions. One can ask whether there is an entropy function which is a suitable measure of the information carried by a gene or even try to find physical or statistical indicators of a system's ability to reproduce itself. The Life game certainly seems to be a good place to begin asking these questions.

Dresden and Wong[3] have considered the game of Life from much the same motivation as we have expressed. However, their approach has been to look for analytic solutions from which one might be able to derive extensions and generalizations. Our approach is in a sense complementary in that we will be concerned with analytic solutions, but mainly as a check on understanding the important parameters of the evolving dynamic life system.

Our feeling is that the analytic approach, although very important for gaining initial insights, will become too complex to handle as we eventually complicate the system to make it more realistic.

Given the motivations just described, our first problem was that in the Life game one does not have available the usual physical quantities; in particular we have not found an analog of a conserved energy. By contrast, for a simple molecular dynamics problem one considers an assemblage of atoms or molecules with a well-defined force law acting between them and then one can write a computer program to keep track of the trajectories of all the particles as they scatter off one another due to the forces between them. After the computation proceeds for a sufficient time, the thermodynamic properties of the resulting array are calculated. These are the usual well-known properties such as energy, density, temperature, entropy, pressure, etc. The problem for the Life game is to know what to calculate. What are the parameters that determine the behavior of the system; what are the "constants of the motion"? Because "Life" does not model any particular natural system, we were further faced with a good deal of arbitrariness in introducing variations on the game. We could generalize "Life" in a number of ways,[3] but found that the basic Conway rules provide the richest structure when enhanced only by the introduction of a simple stochastic element, to be described below.

In Section 2 we describe formal details of the system. A traditional sort of statistical mechanics is the subject matter of Section 3, in which we derive equations of state, such as density as a function of temperature, for the Life game. We also describe experimental results, i.e., computer simulations which show a number of interesting properties, some of them accounted for by the theory. Section 4 deals with entropy and our use of this concept. Finally, in Section 5 we mention miscellaneous observations and give a general discussion of results.

2. THE MODEL SYSTEM

The Conway life game is based on a two-dimensional rectangular array of squares, each square in one of two states, "living" or "dead." At a time t let some subset of the squares in the array be living. The living squares at time $t + 1$ are determined by those at time t according to the following evolutionary rules:

1. If a live square has either two or three live neighbors, it will survive in the next generation; otherwise it will die.
2. If a dead square has exactly three live neighbors, there will be a "birth" in that square and it will be alive in the next generation.

All "births" and "deaths" take place simultaneously. Neighbors are defined to be the eight squares surrounding the square under consideration, or in other terms, the first and second crystallographic nearest neighbors in a square planar lattice.

In the actual system that we have implemented, we have included three generalizations. First, we may choose as we please the set of crystallographic nearest neighbors that we wish to consider as neighbors for evolution of the system. In fact, although we have explored some other choices, all the work reported here has been with the Conway choice.

Second, we allow a choice of boundary conditions. We can either let the board grow as the living areas expand, or impose fixed or periodic boundary conditions. The data reported here were taken with periodic boundary conditions on a square array.

Third, we can allow the birth and survival probabilities to have any value from zero to one, which causes the system to have a stochastic component. That is, we define $P_\alpha^{(k)}$ as the probability that a square having exactly k living neighbors (out of N) either survives ($\alpha = S$ and the square had previously been alive) or is born ($\alpha = B$ and the square had previously been dead). For the normal Conway game $P_S^{(k)} = 1$ for $k = 2$ or 3, $P_B^{(k)} = 1$ for $k = 3$, and they are both zero otherwise. In general, the $P_\alpha^{(k)}$ can have any value between zero and one.

As a specialization of this general stochastic facility, we introduce a particular stochastic component which we denote as "temperature." Temperature acts to provide a finite probability for birth or death no matter what the number of actual live neighbors might be. The temperature is defined in the following manner. Let

$$P_\alpha^{(k)}(T) = [P_\alpha^{(k)}(0) + \rho T]/(1 + T) \tag{1}$$

where ρ is the density of live squares on the entire array. This choice has two virtues: (1) the $T = 0$ limit is just the normal Conway game, and (2) the temperature part conserves density at equilibrium. Specifically, let the density at t be ρ and at $t + 1$ be ρ'. Suppose on the basis of the Conway step alone $\rho' = \rho$. Then with the above definition of temperature it is still true that $\langle \rho' \rangle = \rho$. To see this let $q(k, \sigma)$ ($\sigma = 0$ or 1) be the probability that a square has exactly k living neighbors if it is alive ($\sigma = 1$) or is dead ($\sigma = 0$). Then

$$\langle \rho' \rangle = \sum_{k=0}^{N} [\rho P_S^{(k)} q(k, 1) + (1 - \rho) P_B^{(k)} q(k, 0)] \tag{2}$$

Substituting Eq. (1) and observing that $\sum_k q(k, \sigma) = 1$, we obtain

$$\langle \rho' \rangle = \frac{1}{1 + T} \sum_{k=0}^{N} [\rho q(k, 1) P_S^{(k)}(0) + (1 - \rho) q(k, 0) P_B^{(k)}(0)] + \frac{\rho T}{1 + T} \tag{3}$$

The sum in Eq. (3) is just the new value of ρ based on the $P_\alpha^{(k)}(0)$ and if this yields ρ the entire expression yields ρ also. Therefore, temperature does not provide a direct driving force for the density. It serves merely to alter the pattern of live cells, thereby destroying some of the order. At the infinite-temperature limit all $P_\alpha^{(k)}$ equal ρ, so that the array becomes completely random with density ρ.

There is some similarity between our temperature and that used in the stochastic evolution of dynamical Ising models.[4] In those systems spin flips that lower energy are more likely to occur than those that raise energy. At zero temperature, energy-lowering flips always occur and energy-raising flips never occur. Nonzero temperature is a measure of the extent to which those rules are violated. From our definition, too, the extent to which the system can deviate from its zero-temperature evolution law (the "pure Conway rules") is also a monotonic function of temperature. But we are not able to extend the analogy beyond this, because we do not know of a conserved energy for the Conway dynamics. For the dynamical Ising model, the exact stochastic rule for flipping or not flipping depends on energy and temperature in such a way that by considering the equilibrium state toward which those dynamics tend, the temperature parameter in the stochastic dynamics can be identified with the usual thermodynamic temperature. The lack of an energy definition prevent us from making such a firm identification; monotonicity is the best we can do.

3. EQUATIONS OF STATE

In this section we establish theoretically and experimentally the simplest thermodynamic properties of the Conway game. Without looking at the "reproductive" properties of various special configurations, we ask only, what is the average density of living squares as a function of temperature? In later sections we shall deal with some of the qualitative phenomena encountered in the process of reaching an equilibrium density, such as the formation of a membrane between living and dead regions of the board, but for now we confine our attention to the global equilibrium properties. "Thermodynamic" is often taken to mean the limit that the size of the array goes to infinity. We feel that most of our conclusions are valid in that limit, too (i.e., we work with arrays large enough for finite size effects to be negligible), but we have made no systematic study of this point.

3.1. Zero-Temperature Limit

As an introduction to the kind of calculation that we shall later do with greater accuracy, we estimate the change of ρ on successive generations in the

pure Conway game (zero temperature) with the assumption that the initial distribution of living squares is purely random and has negligible correlations.

The exact evolution law for the Conway game is

$$\sigma_\alpha(t+1) = \delta\Big(3, \sum{}' \sigma(t)\Big) + \sigma_\alpha(t)\delta\Big(2, \sum{}' \sigma(t)\Big) \tag{4}$$

where $\sigma_\alpha(t)$ is a variable taking the value $+1$ if square α is alive at time t and zero if the square is dead at that time. The δ is the Kronecker δ and equals one if its arguments are equal and zero otherwise. The primed $\sum$ indicates a sum over the eight neighbors of the square α. The density of living squares is the average over α of $\sigma_\alpha(t)$ and this is denoted by brackets. Thus

$$\rho(t) = \langle \sigma_\alpha(t) \rangle = \frac{1}{M} \sum_\alpha \sigma_\alpha(t) \tag{5}$$

where M is the number of squares in the system. The evolution of $\rho(t)$ is found by averaging Eq. (4). The average of the first δ is the probability of square α having exactly three live neighbors. By translational invariance (true by virtue of our periodic boundary conditions) it is possible to look at any square α in this calculation—there are no corners or edges. Given that the overall density of live squares is $\rho(t)$, the probability that for any eight squares there are exactly three live ones among them (and in the absence of correlations) is

$$\binom{8}{3} \rho^3 (1-\rho)^5 \tag{6}$$

Reasoning similarly for the last term in Eq. (4), we obtain

$$\rho' = \binom{8}{3} \rho^3 (1-\rho)^5 + \binom{8}{2} \rho^3 (1-\rho)^6 = 28\rho^3 (1-\rho)^5 (3-\rho) \tag{7}$$

where $\rho' = \rho(t+1)$ and $\rho = \rho(t)$. If the system has a long-term steady state, then ρ' will equal ρ. Setting $\rho' = \rho$ in Eq. (7) yields an equation for ρ whose real solutions are

$$\rho = 0 \qquad \text{and} \qquad \rho \approx 0.37017 \tag{8}$$

Without correlations the system moves to the density given by the fixed points of Eq. (7). Figure 1 shows the behavior of ρ_{t+1}/ρ_t as a function of ρ_t, calculated from Eq. (7). Notice that for the curve marked "uncorrelated" the system evolves to the value $\rho = 0.37$ for $0.19 < \rho_t < 0.56$. Outside this range of initial densities $\rho = 0$ is the end result. The uncorrelated approximation, although not valid for the pure Conway game, is applicable to the case of temperatures sufficiently high ($T > 2$) to destroy the Conway correlations.

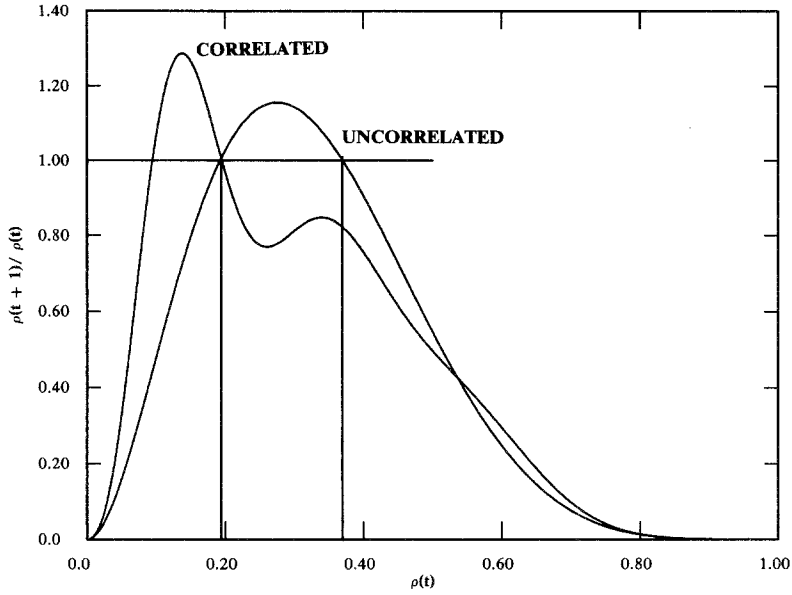

Fig. 1. Density at time $t + 1$ as a function of density at time t for the standard Conway rules.

Experiments carried out in this temperature range yield an equilibrium density of 0.370 ± 0.006, in excellent agreement with Eq. (8).

It is easy to generalize the foregoing to evolutionary rules described by the general vectors $P_\alpha^{(k)}$. In the absence of correlations we have, in place of Eq. (7),

$$\rho' = \sum_{k=0}^{N} \binom{N}{k} \rho^k (1 - \rho)^{N-k} [\rho P_S^{(k)} + (1 - \rho) P_B^{(k)}] \tag{9}$$

This formula allows us to explore various other possible rules. Figure 2 shows an example of fixed points for a rule where, at least at high temperature, two nonzero densities are possible.[4]

Improvement upon the estimate of Eq. (7) consists in taking into account correlations between squares. This can be done in a systematic manner by formally taking the expectation of Eq. (4), which we shall now do.

The δ-function can be written as a sum of products of σ's. To see this, we define a set L to be $L = \{1,..., l\}$ and K to be any subset of L of cardinality k. Thus K can be written $K = \{\alpha_1,..., \alpha_k\}$ with each α an integer between 1

[4] For this case $P_B^{(k)} = P_S^{(k)}$ for all k, and the values are $(0, 0, 1, 0, 0, 1, 1, 1, 0)$ for $(k = 0,..., 8)$, respectively.

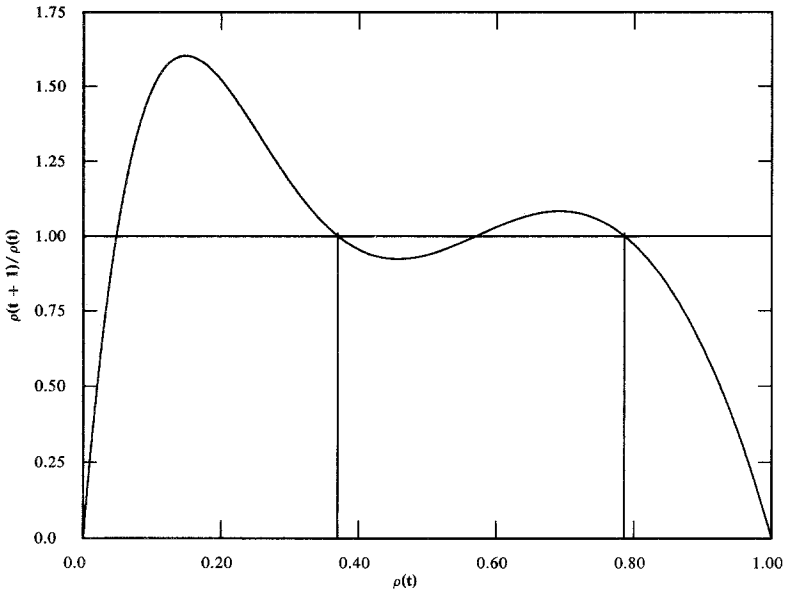

Fig. 2. Density at time $t + 1$ as a function of density at time t for a system with two nonzero stable points.

and l. Now suppose there is a Kronecker δ, $\delta(k, \sum' \sigma_\alpha)$, where the sum in the argument is over some l σ's. Then, letting $\tau_\alpha = 1 - \sigma_\alpha$, we have

$$\delta\left(k, \sum_{i=1}^{l} \sigma_i\right) = \sum_{K} \sigma_{\alpha_1} \cdots \sigma_{\alpha_k} \tau_{\alpha_{k+1}} \cdots \tau_{\alpha_l} \tag{10}$$

where the sum is over all subsets K having exactly k elements. There are $\binom{l}{k}$ terms in the sum.

The problem of evaluating the expectation of Eq. (4) has now been reduced to that of evaluating expectations of products. A convenient way to do this, while taking into account various correlations, is through the use of cumulants. The expectation of the δ-function can be expressed as a function of second-order cumulants as follows[5]:

$$\begin{aligned}\left\langle \delta\left(k, \sum' \sigma\right)\right\rangle &= \binom{l}{k}\rho^k\mu^{l-k} + \left[\sum_{r<s} \langle\sigma_r\sigma_s\rangle_C\right] \\ &\times \left[\binom{l-2}{k-2}\rho^{k-2}\mu^{l-k} - 2\binom{l-2}{k-1}\rho^{k-1}\mu^{l-k-1} + \binom{l-2}{k}\rho^k\mu^{l-k-2}\right]\end{aligned} \tag{11}$$

where

$$\sum_{r<s} \langle \sigma_r \sigma_s \rangle_C \tag{12}$$

is the sum of the cumulants for all relevant pairs, in this case all pairs from the eight neighbors of the central square. The quantity l is equal to 8, the number of neighbors of the central square. In Eq. (11) a combinatorial coefficient is interpreted to be zero if its lower argument exceeds its upper argument.

To obtain the expectation of Eq. (4) we also require the expectation of quantities of the form

$$\sigma_0 \delta\left(k, \sum{}' \sigma\right) \tag{13}$$

where σ_0 is associated with the central square. This can be shown to be[5]

$$\begin{aligned}\left\langle \sigma_0 \delta\left(k, \sum{}'\right) \right\rangle = {} & \binom{l}{k} \rho^{k+1} \mu^{l-k} \\ & + \sum_{r=1}^{8} \langle \sigma_0 \sigma_r \rangle_C \left[\binom{l-1}{k-1} \rho^{k-1} \mu^{l-k} - \binom{l-1}{k} \rho^k \mu^{l-k-1} \right] \\ & + \sum_{r<s} \langle \sigma_r \sigma_s \rangle_C \rho \left[\binom{l-2}{k} \rho^k \mu^{l-2-k} - 2\binom{l-2}{k-2} \rho^{k-1} \right. \\ & \left. \times \mu^{l-k-1} + \binom{l-2}{k-2} \rho^{k-2} \mu^{l-k} \right]\end{aligned} \tag{14}$$

with the same convention for combinatorial coefficients. The first sum over cumulants is simply the sum over the eight neighbors of square zero and the second sum is, as before, the sum over pairs of these neighbors.

To express sums of the form of Eq. (12) in a standard way we provide a notation for cumulants. Referring to Fig. 3, the cumulant L_i is the cumulant associated with the squares 0 and i of that figure. By symmetry, we shall lump

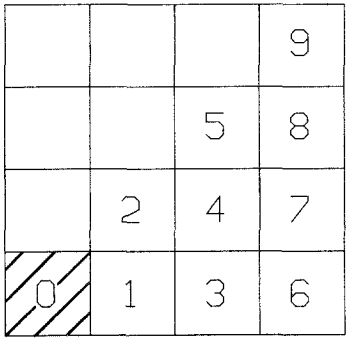

Fig. 3. Labeling convention for the neighbors of a given square (0).

together horizontal and vertical versions of similar cumulants; thus no separate definition is given for cumulants between pairs like 2–3 (counted together with 0–2) or 3–4 (counted together with 0–1), etc. With this notation it is found that

$$\sum_{r<s} \langle \sigma_r \sigma_s \rangle_C = 2(4L_1 + 2L_2 + 3L_3 + 4L_4 + L_5) \equiv 2A$$
$$\sum_{r=1}^{8} \langle \sigma_0 \sigma_r \rangle_C = 4(L_1 + L_2) \equiv 4B \tag{15}$$

We are now in a position to combine all our information on the expectation of Eq. (4). The result is

$$\begin{aligned} \rho' = {} & 56\rho^3\mu^5 + 28\rho^3\mu^6 \\ & + 2A(20\rho^3\mu^3 + 15\rho^3\mu^4 - 30\rho^2\mu^4 - 12\rho^2\mu^5 + 6\rho\mu^5 + \rho\mu^6) \\ & + 4B(7\rho\mu^6 - 21\rho^2\mu^5) \end{aligned} \tag{16}$$

where $\mu = 1 - \rho$.

Equation (16) does not provide adequate information to allow for iteration from generation to generation, since there is no way to iterate the cumulants that appear in B and A. To overcome this problem one can develop a formula for the cumulants in generation $t + 1$ in terms of the density in generation t. We here present an outline and the results only.[5] Let L be the cumulant associated with a pair of squares a and b. Thus

$$L = \langle \sigma_a \sigma_b \rangle_C = \langle \sigma_a \sigma_b \rangle - \langle \sigma_a \rangle \langle \sigma_b \rangle \tag{17}$$

Then its expectation in generation $t + 1$ is given by the product of two formulas of the form (4):

$$\begin{aligned} & L' + \langle \sigma_a' \rangle \langle \sigma_b' \rangle \\ & \quad = \langle \sigma_a' \sigma_b' \rangle \\ & \quad = \langle \delta(3, \textstyle\sum^{(a)})\, \delta(3, \sum^{(b)}) \rangle + [\langle \delta(3, \sum^{(a)}) \sigma_b\, \delta(2, \sum^{(b)}) \rangle \\ & \qquad + \langle \delta(3, \textstyle\sum^{(b)}) \sigma_a\, \delta(2, \sum^{(a)}) \rangle] + \langle \sigma_a\, \delta(2, \textstyle\sum^{(a)}) \sigma_b\, \delta(2, \sum^{(b)}) \rangle \end{aligned} \tag{18}$$

where superscripts a or b on sums indicate sums over the eight neighbors of the appropriate square. Primes refer to generation $t + 1$. Clearly, using Eq. (10), we can write quantities in Eq. (18) as sums of products of random variables, except that now the products may be somewhat more complicated. A novel feature which enters here is that squares of σ's for the same square may now enter and because $\sigma^2 = \sigma$, special treatment of such terms is required. Because of the complexity of the terms, we neglect cumulants in our evaluation of Eq. (18). Thinking of cumulants as second-order terms, we are just evaluating the second-order terms themselves to first order. This will

ultimately have to be justified numerically. In the language of statistical physics, our very first evaluation of Eq. (4), i.e., Eq. (7), was lowest order mean field theory, while our present estimates are the next correction to mean field theory.

We now present the results of the rather tedious calculation of cumulants L_1 through L_5:

$$L_1' + (\rho')^2 = \sum_{j=0}^{3} \binom{4}{j}\binom{4}{3-j}^2 \rho^{6-j}\mu^{6+j} + 2\sum_{j=0}^{2} \binom{4}{j}\binom{4}{2-j}\binom{3}{2-j}\rho^{5-j}\mu^{7+j} + \sum_{j=0}^{1} \binom{4}{j}\binom{3}{1-j}^2 \rho^{4-j}\mu^{8+j} \tag{19}$$

$$L_2' + (\rho')^2 = \sum_{j=0}^{2} \binom{2}{j}\binom{6}{3-j}^2 \rho^{6-j}\mu^{8+j} + 2\sum_{j=0}^{2} \binom{2}{j}\binom{5}{2-j}\binom{6}{2-j}\rho^{5-j}\mu^{9+j} + \sum_{j=0}^{1} \binom{2}{j}\binom{5}{1-j}^2 \rho^{4-j}\mu^{10+j} \tag{20}$$

$$L_3' + (\rho')^2 = \sum_{j=0}^{3} \binom{3}{j}\binom{5}{3-j}^2 \rho^{6-j}\mu^{7+j} + 2\sum_{j=0}^{2} \binom{3}{j}\binom{5}{3-j}\binom{5}{2-j}\rho^{6-j}\mu^{8+j} + \sum_{j=0}^{2} \binom{3}{j}\binom{5}{2-j}^2 \rho^{6-j}\mu^{9+j} \tag{21}$$

$$L_4' + (\rho')^2 = \sum_{j=0}^{2} \binom{2}{j}\binom{6}{3-j} \rho^{6-j}\mu^{8+j} + 2\sum_{j=0}^{2} \binom{2}{j}\binom{6}{3-j}\binom{6}{2-j}\rho^{6-j}\mu^{9+j} + \sum_{j=0}^{2} \binom{2}{j}\binom{6}{2-j}^2 \rho^{6-j}\mu^{10+j} \tag{22}$$

$$L_5' + (\rho')^2 = \sum_{j=0}^{1} \binom{1}{j}\left\{\binom{7}{3-j}^2 \rho^{6-j}\mu^{9+j} + 2\binom{7}{3-j}\binom{7}{2-j}\rho^{6-j}\mu^{10+j} + \binom{7}{2-j}^2 \rho^{6-j}\mu^{11+j}\right\} \tag{23}$$

In the foregoing expressions the quantity ρ' may be computed in terms of generation t quantities with neglect of cumulants. That is, maintaining first-order accuracy in the cumulants, ρ' can be expressed as a function of ρ alone, as in Eq. (7).

The curve marked "correlated" in Fig. 1 is the plot of Eq. (16). It is seen from this curve that $\langle\rho\rangle = 0.19$ for $0.10 < \rho_t < 0.67$. For ρ_t outside this range the density evolves to zero. In contrast to the uncorrelated case, this value is not in good agreement with experiment. As shown in Table I, the

Table I. Final Array Density as $t \to \infty$

	Experiment	Theory
Uncorrelated[a]		
ρ	0.37 ± 0.05[b]	0.370
$\rho_L^{(0)}$	0.377 ± 0.05	0.376
$\rho_L^{(B)}$	0.375 ± 0.04	—
Correlated[c]		
ρ	0.029 ± 0.009	0.189
$\rho_L^{(0)}$	0.229 ± 0.008	0.299
$\rho_L^{(B)}$	0.260 ± 0.30	—

[a] The uncorrelated experimental results are taken from high-temperature experiments.

[b] The ranges given on the experimental results are not errors, but are the variations observed over the number of experiments run.

[c] Correlated implies the $T = 0$ limit.

experimental value for ρ is far smaller than theory predicts. The experimental value represents the stable separated patterns that remain after evolution is complete. There exist a number of absolutely stable geometries in the pure Conway game[1] which, if formed far enough from other live squares, cannot be destroyed at zero temperature. The experimentally observed density is just the final density of these stable forms. As we shall see when we discuss the $T \neq 0$ case, the theory gives a poor account of the global behavior of the low-temperature system.

As the system evolves toward its final state the array will contain a large proportion of completely dead areas. Although the total array area may be large, the live cells will be concentrated in a number of small regions. For this reason, we introduce the concept of "local density." That is, we confine attention to those regions in which the live cells exist and disregard the totally dead areas.

We have considered two definitions of local density. In the first, which we call $\rho_L^{(0)}$, we exclude all squares that have no live nearest neighbors. We call these squares "vacuum" and denote their number by V. Redefining Eq. (5) for $\rho_L^{(0)}$, we have

$$\rho_L^{(0)} = \left(\sum_a \sigma_a\right) \Big/ (M - V) \tag{24}$$

Similarly, for the calculation of $\langle \rho \rangle$ given above, we find

$$\langle \rho_L^{(0)} \rangle = \frac{\langle \rho \rangle}{1 - \langle (1 - \sigma_0)\, \delta(0, \sum^{(0)} \sigma) \rangle} \tag{25}$$

where $\langle (1 - \sigma_0)\, \delta(0, \sum' \sigma) \rangle$ is the probability of finding a "vacuum" square.

The second form of the local density, $\rho_L^{(B)}$, is derived from an evaluation of the average number of "bonds" existing between live cells in the array. If the array were completely populated ($\rho = 1$), each cell would have N live nearest neighbors and there would be $N/2$ bonds per live cell. If we count the existing bonds between live cells (B) in a given array we can evaluate a local density in the form

$$\rho_L^{(B)} = 2B/Nn \tag{26}$$

where n is the total number of live cells ($=\rho M$).

Table I gives the experimental and theoretical values for all three densities for both the correlated and uncorrelated cases (we have not derived a theoretical expression for $\rho_L^{(B)}$). The agreement between theory and experiment is much better for $\rho_L^{(0)}$ than for ρ. For the uncorrelated case the experimental data (taken for $T > 2$) are in excellent agreement with theory for both local and global densities.

3.2. Finite Temperature

So far in this section we have treated the evolution as a deterministic process. We next take into account temperature. Referring to Eq. (3), we define

$$\langle \rho_0' \rangle = \sum_{k=0}^{n} [\rho q(k, 1) P_S^{(k)}(0) + (1 - \rho) q(k, 0) P_B^{(k)}(0)] \tag{27}$$

This is exactly the quantity we have been calculating in this section and for the pure Conway game, to second-order cumulants, is the ρ' of Eq. (16). Hence in this approximation

$$\langle \rho_0' \rangle = \Psi_0(\rho) + A\Psi_A(\rho) + B\Psi_B(\rho) \tag{28}$$

where $\Psi_\alpha(\rho)$, $\alpha = 0, A, B$, are polynomials in ρ, which can be read off from Eq. (16). It follows from Eq. (3) that

$$\langle \rho' \rangle = [1/(1 + T)](\langle \rho_0' \rangle + \rho T) \tag{29}$$

Our goal is to deduce an iteration law for the density. Equation (29) is as yet incomplete in this respect because the effect of temperature on the iteration of A and B (which are sums of order two cumulants) is still needed.

In order to evaluate cumulants $\langle \sigma_i \sigma_j \rangle_C$ we introduce a slightly different notation to describe the probabilistic evolution law. Let $X_{i\alpha}^{(k)}$ (i runs over squares of the array, $\alpha = B$ or S, and $k = 0,..., N$) be a collection of random variables taking the values 0 and 1, all independent of one another, and having the expectation values

$$\langle X_{i\alpha}^{(k)} \rangle = P_\alpha^{(k)} \tag{30}$$

independent of i. Then the evolution law (4) is generalized to

$$\sigma_i(t+1) = \sum_{k=0}^{n} \delta(k, \sum' \sigma(t))\{\sigma_i(t) X_{iS}^{(k)} + [1 - \sigma_i(t)] X_{iB}^{(k)}\} \tag{31}$$

[The expectation value of Eq. (31) is just Eq. (3).] Cumulants arise from products of σ's. For any two random variables σ_1 and σ_2 the second cumulant is in fact the second moment of $\sigma_1 - \langle\sigma_1\rangle$ and $\sigma_2 - \langle\sigma_2\rangle$. Thus

$$\langle\sigma_1\sigma_2\rangle_C = \langle\sigma_1\sigma_2\rangle - \langle\sigma_1\rangle\langle\sigma_2\rangle = \langle(\sigma_1 - \langle\sigma_1\rangle)(\sigma_2 - \langle\sigma_2\rangle)\rangle \tag{32}$$

We use Eq. (32) to write the cumulant for any two squares 1 and 2 in the presence of the temperature effect as

$$\begin{aligned}\langle\sigma_1(t+1)\sigma_2(t+1)\rangle_C &= \left\langle \left[\sum_{k=0}^{n} \delta(k, \sum^{(1)})\{\sigma_1(t) X_{1S}^{(k)} + [1 - \sigma_1(t)] X_{1B}^{(k)}\} - \langle\rho'\rangle\right] \right. \\ &\quad \left. \times \left[\sum_{j=0}^{n} \delta(j, \sum^{(2)})\{\sigma_2(t) X_{2S}^{(j)} + [1 - \sigma_2(t)] X_{2B}^{(j)}\} - \langle\rho'\rangle\right] \right\rangle \end{aligned} \tag{33}$$

The angular bracket in Eq. (33) involves both the expectation of the $X_{i\alpha}^{(k)}$ and the sums or expectation values over squares treated in Eq. (18) et seq. These expectations are entirely independent, the $X_{i\alpha}^{(k)}$ being independent of position and of each other. Consequently each $X_{i\alpha}^{(k)}$ in Eq. (33) can be replaced by $P_\alpha^{(k)}$ and the angular bracket takes on its more restricted meaning. The averaging that remains is therefore the same as done earlier and straightforward manipulation of Eq. (33) yields

$$\langle\sigma_1(t+1)\sigma_2(t+1)\rangle_C = \left(\frac{1}{1+T}\right)^2 \langle\sigma_1(t+1)\sigma_2(t+1)\rangle_C^0 \tag{34}$$

where the symbol $\langle\ \rangle_C^0$ refers to the cumulants computed in Eqs. (19)–(23) above, i.e., the zero-temperature cumulants based on deterministic Conway evolution.

Now A and B are linear combinations of cumulants. By Eqs. (19)–(23) zero-temperature iteration of A and B yields

$$\langle A(t+1)\rangle_{T=0} = Q_A(\rho), \qquad \langle B(t+1)\rangle_{T=0} = Q_B(\rho) \tag{35}$$

where Q_A and Q_B are deduced from Eqs. (19)–(23) and the definitions (15). It follows that at finite temperature

$$\langle A(t+1)\rangle = Q_A(\rho)/(1+T)^2, \qquad \langle B(t+1)\rangle = Q_B(\rho)/(1+T)^2 \tag{36}$$

The iteration law for ρ is obtained by combining Eqs. (27)–(29) to yield

$$\langle\rho(t+1)\rangle = \frac{1}{1+T}[\Psi_0(\rho) + A\Psi_A(\rho) + B\Psi_B(\rho) + \rho T] \tag{37}$$

Fixed points of the iteration law (36) and (37) provide equilibrium densities and correlations. A stable iteration provides a stable equilibrium since the iteration law is modeled on the evolution law of the system. The search for the fixed point can be further simplified by combining the two equations, eliminating A and B, and obtaining

$$\rho = \Psi_0(\rho) + (1 + T)^{-2}[Q_A(\rho)\Psi_A(\rho) + Q_B(\rho)\Psi_B(\rho)] \tag{38}$$

This equation, through its fixed points, provides a definite theoretical prediction for $\rho(T)$, the density as a function of temperature. This prediction is shown in Fig. 4 along with the experimental results. There are three things to be noted in this figure. First, as previously mentioned, the agreement between theory and experiment is very poor for the low-temperature range. Second, at high temperature the density approaches the value predicted for the uncorrelated case, Eq. (8), and is in good agreement with experiment. Third, a distinctly new phenomenon is seen to occur. A "phase transition" is observed in which the system moves from a state with $\langle\rho\rangle = 0$ to $\langle\rho\rangle = 0.37$. The transition is quite sharp and is apparent in both the theoretical curve and the experimental data, although the transition temperatures differ. Figure 5 presents similar data for the local density. In this case the agreement at low temperatures is considerably better, but the discrepancy in the position of the transition temperature remains.

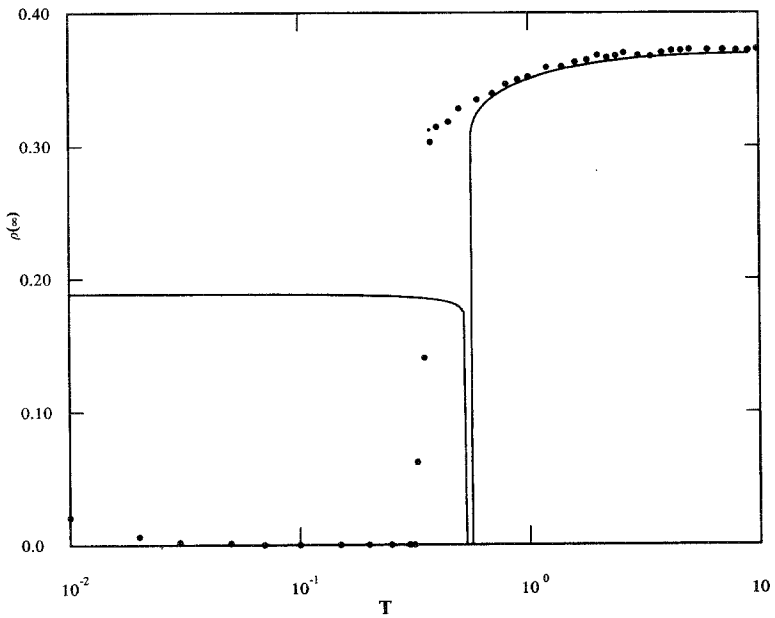

Fig. 4. Density at $t \to \infty$ as a function of temperature.

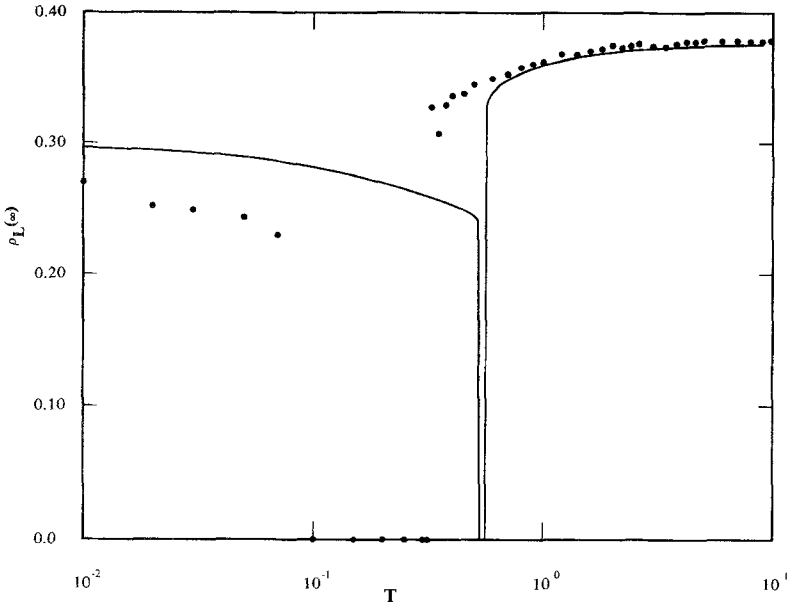

Fig. 5. Local density at $t \to \infty$ as a function of temperature.

4. ENTROPY

A system that begins randomly populated will, under the propagation rules of the Conway game, develop structure and correlations. One way to measure the increasing order in the system is by means of a quantity we shall call entropy and which is defined in close analogy with the correspondingly named quantity in statistical mechanics. The general idea is that we shall coarse-grain our system and measure the extent to which living squares do or do not cluster.

A $J \times J$ board is broken into grains of size $j \times j$. A coarse-grained description of the state of the system is provided by the sequence (n_i), $i = 1,\ldots, (J/j)^2$, where n_i is the number of live squares in the ith grain. The entropy associated with this coarse-grained description is given by the logarithm (taken to the base 2) of the total number of microscopic states that can be associated with the given sequence. Thus[5]

$$S = \frac{1}{J^2} \log \prod_{i=1}^{J(j)^2} \binom{j^2}{n_i} \tag{39}$$

[5] If the set of sequences $(\sigma_{11}, \sigma_{12},\ldots, \sigma_{NN})$, $\sigma_{ij} = \pm 1$, is considered the phase space of the dynamical system, then the definition (39) is precisely the usual entropy for the given coarse-graining. (See, for example, Reif.[6] His specification of the parameters "$y_1,\ldots, y_n$" is equivalent to some coarse-graining.)

This expression can be used directly to calculate the entropy of the arrays generated in the model calculation. The quantity S is the entropy per square and it is worth noting (although we shall not use this fact) that for large J and j it is approximately equal to

$$S = -\frac{j^2}{J^2}\sum_{i=1}^{(J/j)^2}\left[\frac{n_i}{j^2}\log\frac{n_i}{j^2} + \left(\frac{j^2 - n_i}{j^2}\right)\log\left(\frac{j^2 - n_i}{j^2}\right)\right] \tag{40}$$

which is of the general form for the entropy of a probability distribution.

We may estimate the entropy theoretically in the following fashion. The entropy associated with a single grain occupied by exactly k live squares is

$$\frac{1}{j^2}\log\binom{j^2}{k} \tag{41}$$

(this statement is not rigorously true, as entropy has only been associated with macroscopic states; however our forthcoming expression for the expectation of S will be approximately correct for large J^2). The entropy of a collection of grains each having probability $P(k)$ of being occupied by k live squares is then given by

$$\langle S\rangle = \frac{1}{j^2}\sum_{k=0}^{j^2} P(k)\log\binom{j^2}{k} \tag{42}$$

If there are n_l living squares on a board, the density of living squares is $\rho = n_l/J^2$. If these n_l living squares are randomly distributed on the board with no correlations among them, then the probability of finding a $j \times j$ grain populated by exactly k living squares is

$$P(k) = \binom{j^2}{k}\rho^k(1 - \rho)^{j^2-k} \tag{43}$$

Using Eqs. (42) and (43), we obtain for the expectation of the entropy for a distribution without correlations

$$\langle S\rangle = \frac{1}{j^2}\sum_{k=1}^{j^2}\binom{j^2}{k}\rho^k(1 - \rho)^{j^2-k}\log\binom{j^2}{k} \tag{44}$$

For the $j \times j$ grain it is possible to include in $P(k)$ the effect of correlations and to calculate the expected value of S with some given order of correlations. In general

$$P(k) = \langle\delta(k, \textstyle\sum \sigma_a)\rangle \tag{45}$$

the sum within the δ being taken over all squares in the grain. The δ can be written as a sum of products of σ's and τ's. Specifically

$$\delta(k, \textstyle\sum \sigma_a) = \sum_{\substack{\text{permutations}\\ \pi_i}}\prod \sigma_{\pi_1}\cdots\sigma_{\pi_1}\tau_{\pi_{k+1}}\cdots\tau_{\pi_j{}^2} \tag{46}$$

Evaluation of the expectation of the δ-function is performed exactly as in Section 3. Simply replace l by j^2 and reinterpret the sum appearing there (which was for 3 by 3 squares and excluded the center square) as the sum over all squares in a $j \times j$ grain. The result, up to correlations involving no more than two squares, is

$$P(k) = \langle \delta(k, \sum \sigma) \rangle = \binom{j^2}{k} \rho^k \mu^{j^2-k} + \left[\sum_{r<s} \langle \sigma_r \sigma_s \rangle_C \right] \left[\binom{j^2-2}{k-2} \rho^{k-2} \mu^{j^2-k} - 2\binom{j^2-2}{k-1} \rho^{k-1} \mu^{j^2-k-1} + \binom{j^2-2}{k} \rho^k \mu^{j^2-k-2} \right] \tag{47}$$

where as before we have the convention that if the lower index in a combinatorial coefficient is greater than the upper index, the coefficient is zero. The sum is defined to be

$$\sum_{r<s} = \sum_{s=2}^{j^2} \sum_{r=1}^{s} \tag{48}$$

We now list the relevant cumulants for various j. Cumulants involving three squares have been neglected. Thus, when the expectation of entropy is calculated using the forthcoming formulas, a deviation from measured entropy can be attributed to three or more square correlations.

For $j = 2$

$$\sum_{r<t} \langle \sigma_r \sigma_s \rangle_C = 2(2L_1 + L_2) \tag{49}$$

For $j = 3$

$$\sum_{r<s} \langle \sigma_r \sigma_s \rangle_C = 2(6L_1 + 4L_2 + 3L_3 + 4L_4 + L_5) \tag{50}$$

For $j = 4$

$$\sum_{r<s} \langle \sigma_r \sigma_s \rangle_C = 2(12L_1 + 9L_2 + 8L_3 + 12L_4 + 4L_5 + 4L_6 + 6L_7 + 4L_8 + L_9) \tag{51}$$

where cumulant labels are as defined in Fig. 3. The entropy, including correlations, is then obtained from Eqs. (42) and (47)–(51).

Figure 6 shows the experimental entropy as a function of generation (t) for the zero-temperature case determined by using Eq. (39). The entropy is a strong function of density since it measures the multiplicity of states available to the system, and the number of states in itself is a strong function of

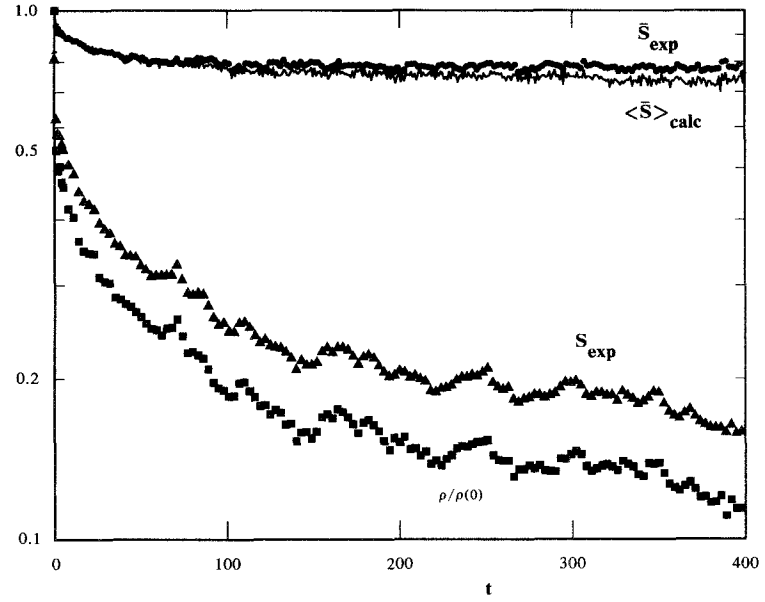

Fig. 6. Entropy and density as a function of time.

density. In order to determine if there exists a contribution to the entropy due to the ordering in the system, we normalize $\langle S\rangle$ by dividing by Eq. (44), the entropy of a random system of density ρ. If the entropy exhibited in Fig. 6 is simply due to a random distribution, the normalized entropy $\langle\bar{S}\rangle$ will be equal to unity.

The upper curves in Fig. 6 show $\langle\bar{S}\rangle$. Rather remarkably, the average normalized entropy decreases with time for the order of 40–50 generations, after which it remains constant even though the density continues to decrease by a factor of ten. The fluctuations in the entropy are roughly given by $n^{-1/2}$. The entropy calculated from Eqs. (42) and (47)–(51) is shown as the continuous line in Fig. 6. It exhibits the same behavior as the experimental data, although its value is slightly smaller.

It is tempting to ascribe this decrease in entropy to the order induced by the Conway rules. If that is true, the asymptotic value of the entropy could be used directly as a measure of order. Furthermore, it would indicate that after the order builds up during the initial phase of the evolution it remains constant throughout the rest of the evolutionary cycle, and is intrinsic to the evolutionary rules themselves. At present this is only a hypothesis. In any event, we have found an example of a "constant of the motion" for this system.

Another indication of the feasibility of using $\langle\bar{S}\rangle$ as a measure of order

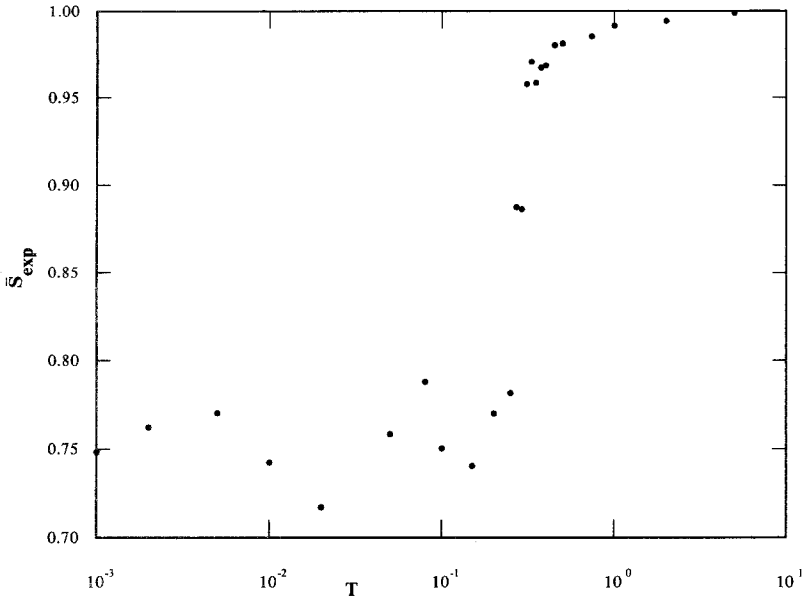

Fig. 7. Normalized entropy as a function of temperature.

is shown in Fig. 7. In this figure, we plot the asymptotic value of $\langle \bar{S} \rangle$ as a function of temperature. It is clear that as the temperature increases, the entropy also increases and approaches unity, indicating complete disorder. In addition, the entropy also exhibits a phase transition. At low temperatures the entropy is roughly constant and jumps sharply to a value close to unity at $T \simeq 0.28$. This value is slightly smaller than the transition temperature observed in the density data (Fig. 5), which is 0.32.

Entropy and normalized entropy as here defined may be expected to yield information about the evolution and ordering of other dynamical systems besides the Conway system. One could, for example, study the evolution of order and the distance scale of that ordering for the dynamical Ising model, in particular when the system is close to its critical temperature. Although we expect the ordering in the dynamical Ising model to be qualitatively different, especially in that nothing so exotic as self-reproducing forms should emerge, nevertheless the entropy results should not differ too much from those for the Conway game. This expectation is based on our successful theoretical calculation (above) of entropy based on two site correlations alone. The success of this calculation is also a bit disappointing, since it suggests that our entropy does not capture the richness (necessarily involving higher order correlations) of the self-reproducing property.

5. DISCUSSION

To some extent this paper represents an exercise in equilibrium and non-equilibrium statistical mechanics. Our main success in this vein is in the description of the equilibrium state at high temperature: on the one hand, we have a clean "experimental" setup, and on the other, a simple form of mean field theory that provides an adequate description of the experiment.

At lower temperatures there is a phase transition—here, too, theory and experiment are in qualitative agreement, even if they disagree on the value of the transition temperature. Below the transition the theoretically predicted global density is far too large, although we are able to find a "local" density whose theoretically predicted value (again, in mean field theory) is in better agreement with experiment.

The source of the discrepancies at low temperature is easily understood and were there a reason to improve the theory in that domain it should not be too hard to do so. Basically what is happening is that the system becomes extremely nonuniform. Great patches of vacuum develop; the progress of vacuum into living areas is steady and inexorable, while advances from living regions do not penetrate very well into vacuum. The multisquare correlations associated with vacuum are neglected in our theory and so this spreading phenomenon is not well described. An interesting feature of low-T patterns is the membrane that separates living from dead regions. This has distinctly different properties from other living regions: in particular, it is more densely populated. An easy estimate of its density turns out to be remarkably accurate. In particular, boundary or "membrane" squares have, on the average, three dead squares to one side. On the other hand, overall equilibrium (neglecting correlations) requires that the average square in a living region have $0.37 \times 8 \approx 3$ living neighbors. Thus one would expect the density of living squares in the *membrane* to be high enough to guarantee about three living neighbors to each square. This is $(8/5) \times 3 \approx 5$ neighbors, or a density of about 0.6. Such a figure indeed emerges when one studies various arrays, although it is hard to pin it down too precisely because of uncertainties inherent in the definition of membrane. The local density we define is designed to measure the density *within* the living region; hence the better agreement of theory and experiment.

Another goal of this paper was to see whether the Conway rules, which tend on the average to drive systems to self-reproducing forms of low periodicity (in time), would reveal, to statistical indicators, any reflection of their ordering abilities. To this end we measured the entropy of the evolving array. We found that from an initially random state, order developed and entropy decreased in this effectively open system. However, aside from this general trend, a far more striking effect was noted for the entropy when it

was normalized in a certain way. For a given density and a given coarse-graining (the latter involved in the definition of entropy) a random distribution of living squares in which states of neighboring squares are completely uncorrelated yields a certain expectation value for entropy. When the measured entropy is divided by this uncorrelated entropy the resulting ratio $\bar{S}$ is very nearly constant. For equilibrium states it remains relatively constant despite fluctuations in both ρ and S, and even more remarkably for non-equilibrium states, although ρ may change by more than a factor ten, $\bar{S}$ remains constant. Finally, as would be expected in the case of a first-order phase transition, the entropy is a strong function of temperature in the region of the phase transition and increases sharply as the system enters the disordered high-temperature state.

REFERENCES

1. M. Gardner, *Sci. Am.* **223**(Oct.):120 (1970); **223**(Nov.):116 (1970); **224**(Jan.):104 (1971); **224**(Feb.):112 (1971); **224**(Apr.):114 (1970); **226**(Jan.):104 (1972).
2. W. W. Wood and J. J. Erpenbeck, *Ann. Rev. Phys. Chem.* **27**:319 (1976); F. H. Ree, *Physical Chemistry. An Advanced Treatise*, D. Henderson, ed. (Academic Press, 1971).
3. M. Dresden and D. Wong, *Proc. Nat. Acad. Sci.* **72**:956 (1975).
4. R. J. Glauber, *J. Math. Phys.* **4**:294 (1963).
5. L. S. Schulman and P. E. Seiden, IBM Research Center Report RC 6802 (November 1977) (copies available from the author).
6. F. Reif, *Fundamentals of Statistical Physics and Thermal Physics* (McGraw-Hill, 1965), Chapter 3.