International Journal of Theoretical Physics, Vol. 21, Nos. 6/7, 1982

Symbolic Dynamics of One-Dimensional Maps: Entropies, Finite Precision, and Noise

J. P. Crutchfield and N. H. Packard

Physics Board of Studies, University of California, Santa Cruz, California 95064

Received May 7, 1981

In the study of nonlinear physical systems, one encounters apparently random or chaotic behavior, although the systems may be completely deterministic. Applying techniques from symbolic dynamics to maps of the interval, we compute two measures of chaotic behavior commonly employed in dynamical systems theory: the topological and metric entropies. For the quadratic logistic equation, we find that the metric entropy converges very slowly in comparison to maps which are strictly hyperbolic. The effects of finite precision arithmetic and external noise on chaotic behavior are characterized with the symbolic dynamics entropies. Finally, we discuss the relationship of these measures of chaos to algorithmic complexity, and use algorithmic information theory as a framework to discuss the construction of models for chaotic dynamics.

1. INTRODUCTION

We will consider the role of computation in modeling the temporal behavior of (classical) physical systems from the vantage point of dynamical systems theory. The object of any such modeling attempt is to match the behavior of some physical system with the behavior of a model comprised of some other representative system. This model may be another, simpler physical system, an abstract mathematical system, or an algorithm executed on a digital computer. It is this latter example of a model that we will focus upon. Our discussion of this modeling process relies on ideas developed by Rob Shaw (1980), to whom we are indebted for the guidance of his perspective.

If a physical system executes "regular" behavior, i.e., fixed point or periodic motion in its state space, a "good" model may be constructed to give excellent correspondence between the physical behavior and simulated model behavior, so that the model can serve a *predictive* role in understanding the physical phenomena. If one measures the physical system's state, a

0020-7748/82/0600-0433$03.00/0

model simulation can then be used to determine the future state of the physical system to within some *acceptable* error.

A large class of systems are not so well behaved: they display "irregular" motion in their state space. We will define a *turbulent* physical system as one which is inherently unpredictable in the sense that the information obtained from making a measurement on the system is lost after some finite time $\hat{t}$. That is, an observer's knowledge of the system's state goes to zero at $\hat{t}$.[1] This definition of a turbulent physical system is very general, and highly dependent on the measurement process used by the observer. Observed turbulent behavior may be due to one of a number of different causes: observational uncertainty, inability to resolve many degrees of freedom (as in the case of observing Brownian motion), or deterministic chaos, the latter having been added to the list only recently (Lorenz, 1963).

The case where deterministic chaos underlies the observed turbulent behavior is of special interest, and the one with which we concern ourselves here. The term *strange attractor* was coined by Ruelle and Takens (1971) to refer to the geometric structure in the state space responsible for the chaotic behavior in dissipative dynamical systems.[2] A strange attractor is, first of all, an attractor because a wide range of initial conditions (all those in its *basin of attraction*) approach it asymptoticaly, and, second of all, strange because orbits on the attractor execute neither fixed point nor periodic motion, but instead wander erratically and randomly over the attractor.[3] Dynamical systems theory provides several characterizations of chaos in such systems, and so allows for the interpretation of deterministic models of the unpredictable behavior. For these systems, the unpredictability can be quantified and given an information theoretic interpretation: Nearby orbits on a strange attractor diverge exponentially in one or more directions. This process rapidly amplifies uncertainties in determining the system's state. The positive Lyapunov characteristic exponents quantify the average asymptotic rate of divergence, and can be interpreted as the strange attractor's information production rate, or alternatively, the rate at which the information contained in a measurement is lost (Shaw, 1980).

The modeling of turbulent physical systems with some underlying strange attractor presents special problems due to this inherent amplifica-

[1] Phase-coherent chaotic systems present a special case; see Farmer et al. (1979).

[2] We refer the reader to the many excellent reviews for the relevant technical definitions of dynamical systems and strange attractors: see, for example, Chillingsworth (1976), Guckenheimer et al. (1980), Collet and Eckmann (1980).

[3] Technical definitions of a strange attractor may be found in Ruelle and Takens (1971); in this paper we will use the physically motivated notion presented in the text, ignoring a few subtleties.

tion of measurement uncertainty. Since uncertainties are introduced because the physical system cannot be measured with infinite precision, there is no hope of using chaotic models to predict physical behavior for times longer than $\hat{t}$. The criteria for a "good" model must now be generalized: instead of requiring exact correspondence between the physical system and the model, we must be content with a correspondence between geometric and statistical characterizations of the dynamics, giving up the idea that our deterministic model will provide us with any long-range predictive capability.

A general procedure for constructing a model for a physical system from a series of observations has yet to be developed, but progress is being made (Packard et al., 1980; Froehling et al. 1981; Takens, 1980). In the present work we will not consider this problem, but instead address certain problems that occur once a model has been chosen. We will describe quantities which characterize the chaotic behavior of a strange attractor, with particular emphasis on how these quantities may be computed numerically. We will then show how these quantities are affected by the fact that the simulation is implemented on a machine with a finite number of states, and discuss the similarities and differences between roundoff errors and errors introduced by added noise fluctuations. Though we will analyze specific models, many of the numerical techniques may be applicable to the problem of directly analyzing the turbulent behavior of a physical system.

Interest in chaotic dynamical systems has prompted other work on the reliability and interpretation of numerical results obtained by simulating such systems (Benettin et al., 1978; Erber et al., 1979). We hope the present work will extend and complement this previous work.

2. CONSTRUCTION OF SYMBOLIC DYNAMICS

There are many motivations for using the techniques of symbolic dynamics in the study of chaotic dynamical systems.[4] The first, most physically compelling reason is that the (classical) measurement process can be viewed as producing an approximate representation of a physical system's evolution. This representation consists of a sequence of symbols, where each symbol corresponds to the output of a measuring instrument at discrete times. Any measuring instrument has finite resolution, hence the range of possible symbols is finite, with each symbol representing a different numerical value of the quantity being observed.

One of our original motivations for studying symbolic dynamics was to search for algorithms to compute an attractor's dimension and entropy from

[4]For a review of symbolic dynamical techniques, see Alekseyev and Yakobson (1981).

a time series. Of course, some algorithms already exist (Packard et al., 1980; Froehling et al., 1981; Takens, 1980; Farmer, 1981), but symbolic dynamics gives some hope of circumventing their severe precision and large data base requirements.

Symbolic dynamics also forms the cornerstone of the only rigorous treatments that exist for chaotic dynamical systems. And so, there are significant theoretical motivations for its study. Bowen (1975), Bowen and Ruelle (1975), and Sinai (1972) use symbolic dynamics to develop a fairly complete theoretical characterization of systems which satisfy Smale's axiom-A (Smale, 1967). One important aspect of this theory is the development of a "thermodynamic formalism" using the symbolic dynamics associated with an axiom-A system to prove the existence of an ergodic invariant measure which is an "equilibrium" or Gibb's state. It is not yet clear how the rigorous results concerning axiom-A systems will carry over to the many dynamical systems whose attractors are not in this class. This provides yet another reason to look at numerical results of symbolic dynamics computations of some simple non-axiom-A systems.

A dynamical system f: $M \to M$ can have many symbolic representations, each obtained by partitioning the state space M into a finite number of sets S_j, $j=1,\ldots,p$, and labeling each element of this partition $S=\{S_j\}$ with a symbol s_i. The time evolution of the dynamical system is translated into a sequence of symbols

$$s=\{\ldots,s_{-1},s_0,s_1,s_2,\ldots\}$$

and f itself is replaced by a *shift* operator σ, which reindexes a symbol sequence; that is,

$$\sigma(s)=s'$$

where for each symbol in the sequence s,

$$s_i'=(\sigma(s))_i=s_{i-1}$$

Thus the shift σ merely moves the time origin of a symbol sequence one place to the right.

In the space of all possible symbol sequences

$$\Sigma=\{s=(\ldots s_{-1},s_0,s_1,\ldots)\}$$

admissable sequences are those which satisfy $f^i(x_0)\in S_{s_i}$. We shall take $x_0\in M$ as a point on the attractor of the dynamical system. The set of admissable sequences is a closed invariant set in Σ, just as are the points on

the original system's attractor. The set of admissable sequences Σ_f, along with the shift σ_f restricted to this set, is called a *subshift*. This space with the shift σ_f is the symbolic dynamical system induced by f using the *measurement partition* S. We will be mostly concerned with the space of one-sided sequences

$$\Sigma_f = \{s = (s_0, s_1, \ldots)\}$$

obtained by observing which partition element is visited by the points of an f orbit $\{x_0, x_1, \ldots\}$, where $x_{n+1} = f(x_n)$. The action of f on M then induces a one-sided shift on the space of admissable symbols Σ_f.

The usefulness of the symbolic dynamics construction is that if the dynamical system associated with f is ergodic, then the induced symbolic dynamical system is a faithful representation, that is, the following diagram commutes:

$$\begin{array}{ccc} \Sigma_f & \xrightarrow{\sigma_f} & \Sigma_f \\ {\scriptstyle \pi}\downarrow & & \downarrow{\scriptstyle \pi} \\ M & \xrightarrow[f]{} & M \end{array}$$

with the projection operator

$$\pi(\ldots s_{-1}, s_0, s_1, \ldots) = \bigcap_{i=-\infty}^{\infty} f^{-i} S_{s_i}$$

One can then study the simpler, albeit abstract, symbolic dynamical system in order to answer various questions about the original dynamical system. Within this construction, every point on the attractor will have at least one symbol sequence representation. There are a few ambiguities in the labeling of orbits by symbol sequences that prevent π from being invertible, but they will not affect our numerical calculations.

Speaking now in the context on one-sided admissable sequences, a finite sequence of symbols $(s_0^n, \ldots, s_{n-1}^n)$ defines an *n-cylinder* $s^n = \{s\colon s_i = s_i^n, i = 0, \ldots, n-1\}$ which is a subset of Σ_f consisting of all sequences whose first n elements match with those of s_i^n. With the identification between symbol sequences and orbits above, we see that an n-cylinder corresponds to a set of orbits that are "close" to one another in the sense that their initial conditions and first $n-1$ iterates fall in the same respective partition elements. Since these orbits must follow each other for at least

$n-1$ iterations, they must all have initial conditions that are close, belonging to some set $U \subset M$. To a different n-cylinder will correspond a different set of orbits whose initial conditions are contained in some other set $U' \subset M$. Continuing with the set of all n-cylinders, M will become partitioned into as many elements as there are n-cylinders. As n is increased, the *n-cylinder-induced partition* of M will become increasingly refined.

3. SYMBOLIC DYNAMICS FOR ONE-DIMENSIONAL MAPS

In this paper, our numerical experiments will use very simple example systems, namely, maps of the unit interval onto itself (one-dimensional maps). Unless otherwise stated, we will consider the quadratic logistic equation $x_{n+1} = rx_n(1-x_n)$ as our prototypical chaotic dynamical system. As an example of the asymptotic behavior of this system, Figure 1 shows the probability density constructed from a histogram of 10^7 iterations of the logistic equation at $r = 3.7$.[5]

We will partition the interval $[0,1]$ into two subintervals $[0,d)$ and $[d,1]$. Although, we will consider only two-element partitions, others are clearly possible. The symbols we shall use will be 0 for $[0,d)$, and 1 for $[d,1]$. Changing the decision point d clearly generates a different set of admissable sequences, just as it generates different n-cylinder-induced partitions. Figure 2 illustrates the n-cylinder-induced partition with the decision point $d = 0.5$, and shows how the dividing points for the n-cylinder-induced partition are simply the collection

$$\{d, f^{-1}(d), f^{-2}(d), \ldots, f^{-(n-1)}(d), \ldots\}$$

whenever the specified inverse images exist. Whenever the map is not everywhere two onto one, some of the inverse images will not exist, corresponding to the fact that some n-cylinders are nonadmissable.

The space of one-sided symbol sequences can easily be metrized by mapping each symbol sequence to a power series

$$\phi(x) = \sum_{i=1}^{\infty} \frac{S(f^i x)}{2^i}$$

where $S(x)$ is 0 or 1 depending whether $x < d$ or $x \geq d$. This map identifies every sequence with a binary fraction whose value lies in $[0,1]$. We will

[5]All logarithms in this paper are computed to the base 2.

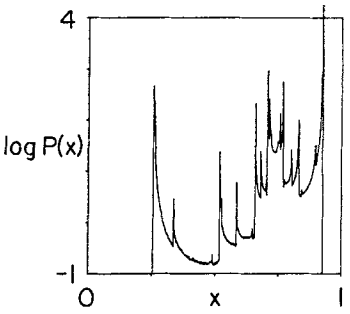

Fig. 1. Binary logarithm of the asymptotic probability distribution for the logistic equation $f(x)=rx(1-x)$ at $r=3.7$, using 2×10^7 iterations sorted into 10^3 bins.

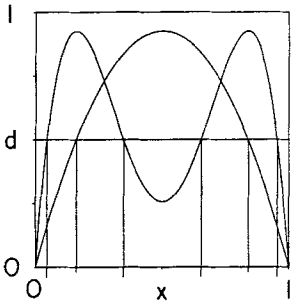

Fig. 2. Construction of the partition induced by taking n symbols (i.e., specifying an n cylinder) with a given decision point d. Illustrated is the case $d=0.5$, and the 1-cylinder-, 2-cylinder-, and 3-cylinder-induced partitions are shown with successively longer tic marks on the x-axis.

conveniently confuse s^n with its binary fraction representation unless the distinction is necessary.[6]

The Cantor set structure of the symbol sequences of the chaotic logistic equation is revealed in Figure 3 by a sequence of probability distributions for n-cylinder binary fractions: with the increase in length of n-cylinder the distributions show successively more, although narrower, peaks. An even more graphic demonstration of the Cantor set structure is the graph of the distribution of symbols s (truncated to a finite n-cylinder with $n=12$) versus position x, illustrated in Figure 4.

[6]Milnor and Thurston (1977) show how to form a slightly more sophisticated "invariant coordinate" which is monotonic. Our entropy calculations do not require this feature, so we use the computationally simpler binary fraction.

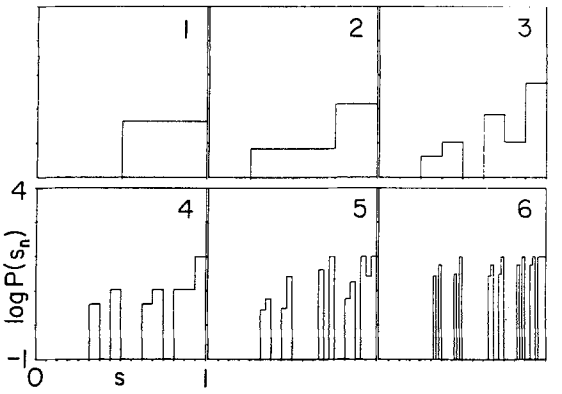

Fig. 3. The Cantor set structure of the subshift Σ_f, where f is the logistic equation with $r = 3.7$, is shown in this sequence of probability distributions for n cylinders where each n cylinder has been mapped onto the unit interval by using its binary fraction.

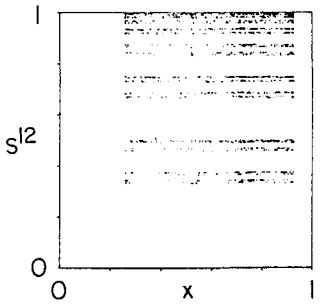

Fig. 4. Two-thousand iterations of the logistic equation with $r = 3.7$, showing in more detail the Cantor set structure of the distribution of binary fractions s^n (with $n = 12$) vs. the distribution of points x on the attractor.

We will now embark on the task of characterizing the chaotic behavior in a dynamical system using topological and metric entropies in that order. After giving their definitions, we will show various illustrative computations that will lay the groundwork for our discussions of noise and finite precision effects.

4. TOPOLOGICAL ENTROPY

Heuristically, the topological entropy of a dynamical system measures the asymptotic growth rate of the number of resolvable orbits whose initial

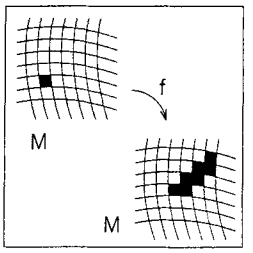

Fig. 5. Under the action of a chaotic dynamical system, points in a small region spread apart from each other exponentially, overlapping other regions, and eventually covering an attractor. The grid in each picture represents the measurement partition. The size of each element illustrates the limit of determining the system's state.

conditions are all close. Equivalently, the topological entropy quantifies the average time rate h of spreading a subset over nearby subsets (see Figure 5). This process is most easily illustrated by considering a collection of subsets which form a "cover" of the phase space, as shown schematically in Figure 5. In the figure, the dynamic f has spread the single cover element over other elements after some time t. The number of new cover elements $N(t)$ visited by points in the original cover element can be written

$$N(t) \sim e^{ht}$$

where $h > 0$ for chaotic dynamical systems. With this geometric motivation, we will now consider a more formal definition of the topological entropy h (Adler et al., 1965).

For a compact topological space M, with an open cover U, let $N(U)$ be the number of sets in a subcover of minimal cardinality. Two covers U and V may be "combined" to form a refinement W by

$$\begin{aligned} W &= U \vee V \\ &= \{A \cap B \mid A \in U \text{ and } B \in V\} \end{aligned}$$

Now if f: $M \to M$ is a continuous map, the *topological entropy* of f with respect to the cover U is defined as

$$h(f, U) = \lim_{n \to \infty} \log \frac{N(U \vee f^{-1}U \vee \cdots \vee f^{1-n}U)}{n}$$

The topological entropy $h(f)$ of the map itself is then simply the supremum of $h(f, U)$ over all open covers U.

In the space of symbol sequences Σ_f, each n cylinder s^n is an open set (in the discrete topology, or in the topology generated by the metric mentioned above), and the class of all n cylinders is an open cover. As n increases, the open cover becomes more refined, and

$$\lim_{n\to\infty} \frac{\log[N(n)]}{n} \to h(\sigma_f)$$

where $N(n)$ is the number of admissable n cylinders.[7] $N(n)$ is easily obtained numerically, so this formula presents us with an readily computable algorithm for the topological entropy.[8] Although h is the asymptotic slope of $\log[N(n)]$, it turns out, for reasons of convergence that will be discussed shortly, that the most practical way to compute $h(\sigma_f)$ from $N(n)$ is to fit a slope to $\log[N(n)]$ for $n=12,\ldots,16$, for example.

We have numerically computed the topological entropy of the shift induced by the binary partition $\{[0,0.05),[0.5,1]\}$ for the logistic equation $x_{n+1}=rx_n(1-x_n)$. Figure 6 shows the increase of $\log[N(n)]$ with n for two parameter values $r=3.7$ and 4.0, and Figure 7 shows the slope of this curve, $\log[N(n+1)]-\log[N(n)]$, for two parameter values, $r=3.9$ and 3.7. Figure 8 shows that this numerical estimate of the topological entropy converges when a sufficient number of iterations are used. In Figures 6–8 the quantities relating to the topological entropy are shown with dashed lines; solid lines refer to the metric entropy, which will be discussed below.

Figure 7 presents several questions about the convergence and rate of convergence of the above expression for the topological entropy. Figure 7 shows how the slope of $\log[N(n)]$ begins to oscillate as r is lowered toward the value where one chaotic band splits into two bands, $r=3.67857351\ldots$.[9] At the band joinings, the "two-point slope" $\log[N(n+1)]-\log[N(n)]$ oscillates indefinitely, causing $\log[N(n)]/n$ to converge rather slowly. This is illustrated in Figure 9. Convergence to $h=0$ at the period-doubling accumulation point also is especially problematic.[10]

[7]For the case of symbolic dynamics, this formula for the topological entropy was first introduced by Parry (1964).

[8]There are other algorithms to compute the topological entropy of a map f based on representing the dynamics as a branching process with a deterministic transition matrix. For certain cases, they allow one to analytically calculate the topological entropy and so to study the convergence of the topological entropy directly. These techniques are based on the kneading calculus of Milnor and Thurston (1977).

[9]For a review of the phenomenology of the quadratic one-dimensional map, see Collet and Eckmann (1980) and Crutchfield et al. (1981).

[10]We will discuss the convergence at the period-doubling accumulation parameter value r_c in terms of the structure of the symbol sequence at r_c in a future paper.

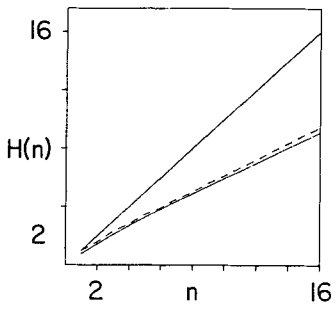

Fig. 6. Entropy convergence as a function of symbol length for the logistic equation. Top solid line is $H(n)$ and $\log[N(n)]$ for $r=4.0$ (2×10^6 iterations); the dashed line and lower solid line are $\log[N(n)]$ and $H(n)$, respectively, for $r=3.7$ (5×10^5 iterations). The topological entropy and metric entropy are the slopes of these two curves.

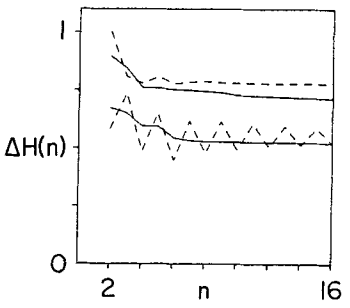

Fig. 7. Entropy convergence as a function of symbol length. Dashed lines are $\log[N(n)]-\log[N(n-1)]$ solid lines are $H(n)-H(n-1)$ for $r=3.9$ (upper set) and $r=3.7$ (lower set). 5×10^5 iterations.

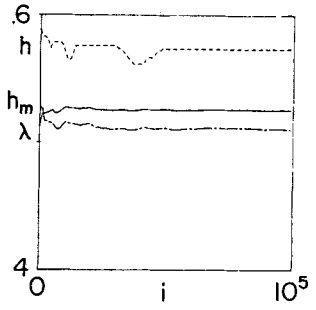

Fig. 8. Entropy convergence as a function of number of iterations i. Upper dashed curve shows topological entropy approximated by $\log[N(16)]-\log[N(15)]$, solid curve shows metric entropy $H(16)-H(15)$, and included for reference is the convergence of the Lyapunov exponent λ.

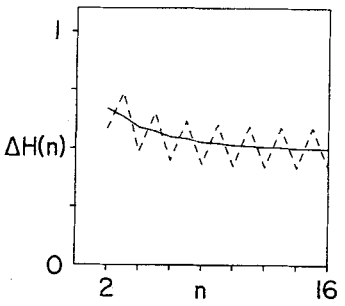

Fig. 9. Entropy convergence as a function of symbol length (as in Figure 7) at $r = 3.6785735\,1\ldots$, where two bands merge into one.

As mentioned before, changing the decision point will alter the set of admissable symbol sequences, and so the topological entropy of the shift σ_f. For example, if the decision point is on the far edge of the attractor, there will be a vast majority of one of the symbols. A graph of the topological entropy as a function of decision point is shown as the upper curve in Figure 10. From this figure we see that the measurement partition $\{[0,0.5),[0.5,1.0]\}$ yields a maximum value of $h(\sigma_f)$, indicating that the supremum over all partitions has been reached. In this case, we find that $h(\sigma_f)$ is identical to $h(f)$ computed using the kneading calculus of Milnor and Thurston (1977). Indeed, by the argument at the end of Section 2, each n cylinder corresponds to a particular element S_j of the cover of $[0,1]$ comprised of the n-cylinder induced partition. Thus, for the proper partition, $h(\sigma_f) = h(f|_{\text{attractor}})$. In effect, the symbolic dynamics computation

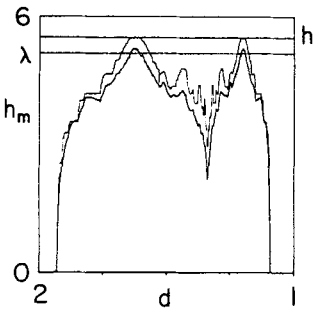

Fig. 10. Topological entropy (upper curve) and metric entropy (lower curve) of the shift induced by choosing different decision points d. The upper horizontal line is the topological entropy calculated to one part in 10^6 with the kneading determinant. The lower horizontal line is the Lyapunov characteristic exponent calculated to within 0.1%. The parameter r is 3.7.

provides a labeling scheme for the elements of the induced partition on $[0, 1]$.

5. METRIC ENTROPY

In presenting the topological entropy before metric entropy we have purposely reversed their historical order because there is a sense in which the metric entropy is a generalization of the topological entropy: the metric entropy also measures the asymptotic growth rate of the number of resolvable orbits having close initial conditions, but weighting each orbit with its probability of occurrence.

The definition of metric entropy for the dynamical system (M, f) requires an invariant measure μ and a σ algebra of measurable subsets of M: more structure than needed for the definition of topological entropy. For symbolic dynamical systems, this will not prove to be much of a problem, however, since the same n cylinders that formed elements of the open covers are also measurable subsets of Σ_f.

The structure of the attractors for the logistic equation can be quite complicated (Jonker and Rand, 1980), and the existence of many different asymptotic measures (i.e., measures whose averages of continuous functions coincide with time averages) raises serious questions concerning the validity of numerical calculations (Bennetin et al., 1978, 1979; Erber et al., 1980). Numerical evidence suggests, however, that a single asymptotic measure is selected by the simulated dynamics. The existence of such an "observable" asymptotic measure is also suggested by theoretical results that tell us when a small amount of noise is added to axiom-A systems a unique asymptotic measure is selected (Kifer, 1974).

If $P = \{P_i\}$ is a finite measurable partition of M with p elements, we define the entropy of P as

$$H_\mu(P) = \sum_{i=1}^{p} \mu(P_i) \log[\mu(P_i)]$$

Given two partitions P and Q, their refinement is

$$P \vee Q = \{P_i \cap Q_j \mid \text{for all } P_i \in P \text{ and } Q_j \in Q\}$$

The metric entropy of f with respect to the partition P is defined by

$$H_\mu(f, P) = \lim_{n \to \infty} H_\mu(P^n)$$

where

$$P^n = P \vee f^{-1}P \vee \cdots \vee f^{1-n}P$$

Finally the metric entropy of f itself is

$$h_\mu(f) = \sup_P H_\mu(f, P)$$

where the supremum is taken over all partitions P.

For the numerical computation of metric entropy, the latter is not a very useful definition, but there is a theorem due to Kolmogorov[11] that helps: if P is a *generator*, i.e., the partition P^n becomes arbitrarily fine as $n \to \infty$, then we have

$$h_\mu(f) = H_\mu(f, P)$$

Unfortunately, there is no way to determine in general whether any given partition is a generator for a dynamical system. This is not a problem for symbolic dynamical systems, however, because generating partitions abound: the simplest ones being the set of all 1-cylinders. We must emphasize here that even though an n-cylinder partition of Σ_f may be a generator for the shift σ_f obtained using some measurement partition on (f, M), this does *not* necessarily mean that the n-cylinder-induced partition on M (see Section 2 above) is a generator for f. The metric entropy computed for σ_f will be the same as the metric entropy of f *only* when this is so. In any case, the metric entropy of f is the supremum over all partitions, so it is at least as large as the entropy of the symbolic dynamical system induced by choosing some measurement partition. That is,

$$h_\mu(\sigma_f) \leqslant h_{\mu'}(f)$$

where μ and μ' are the asymptotic invariant measures on Σ_f and M, respectively.

We will now narrow the discussion to the special case of symbolic dynamical systems. Let $S = S^1 = \{s^1\}$ denote the partition of all 1-cylinders and S^n the partition of all n cylinders. Then, if we let $H_\mu(n) = H_\mu(\sigma_f, S^n)$,

$$h_\mu(\sigma_f) = \lim_{n \to \infty} \frac{H_\mu(n)}{n}$$

[11] This theorem as well as the original definition of metric entropy are presented in Kolmogorov (1958).

This formula allows the numerical computation of the metric entropy for the shift σ_f, where the invariant measure μ used to compute $H_\mu(n)$ is accumulated empirically with a frequency histogram.[12] Shimada (1979) was the first to compute the metric entropy using this formula, applying it to a shift induced by the Lorenz attractor.

As in the case of topological entropy, the metric entropy is the asymptotic slope of $H(n)$ as a function of n, but this asymptotic slope is obtained numerically most readily from the "local" slope $H_\mu(n+1)-H_\mu(n)$ [or even a fit of the slope of $H_\mu(n)$ for $n_1<n<n_2$] rather than $H_\mu(n)/n$. The quantity $H_\mu(n+1)-H_\mu(n)$ may also be written as a conditional entropy (Shimada, 1979; Billingsley, 1965), and it has a compelling informational interpretation: the metric entropy of the shift σ_f is exactly the average gain in information for each new symbol, obtained by an observer using the measurement partition S, as the number of symbols gets large.

From the above definition of the metric entropy, it is easy to see that $h \geqslant h_\mu$, since $H_\mu(f, P^n)$ is maximized when each element of P^n is equally probable [i.e., $\mu(P_i^n)$ is the same for all i]. In this case, the formula for metric entropy reduces to that for the topological entropy. This is also evident from Dinaburg's theorem, which states that

$$h = \sup_\mu h_\mu$$

where the supremum is taken over all invariant measures μ (Dinaburg, 1970).

We have numerically computed the metric entropy of the shift induced by the binary partition $\{[0,0.5),[0.5,1]\}$ for the logistic equation. Figure 6 shows $H_\mu(n)$ vs n, and Figure 7 is a graph of this curve's slope, which approaches the metric entropy as n gets large. Figure 8 shows that this expression converges after enough iterations. Figure 9 shows that there is no problem with the convergence at band joinings as for the topological entropy. Although the convergence of the metric entropy still presents a problem near the accumulation points r_c of the period-doubling bifurcation sequences, away from r_c, convergence to within 5% is obtained using 13 symbols and 2×10^5 iterations. Using more symbols requires more iterations to fill out the histograms sufficiently.

As in the case of the topological entropy calculations, the choice of measurement partition used for the symbolic dynamics can make a large

[12] We assume that there exists an asymptotic invariant measure whose averages are equal to time averages for continuous functions as well as the characteristic functions on the bins of the frequency histogram. This latter feature implies that the frequency histogram should converge to a "coarse-grained" approximation of the invariant measure.

difference in the entropy of the induced shift. Figure 10 shows how the metric entropy of the shift varies with the binary partition $\{[0, d), [d, 1]\}$ as d is varied. We see that the metric entropy is maximized when $d = 0.5$ and $d = 0.635\ldots$ (one of the inverse images of 0.5). This numerical evidence, in turn, suggests that these two decision points form generating partitions.

6. DIMENSION

When Σ_f is endowed with a metric (as described above) so that each one-sided symbol sequence is mapped to a binary fraction, the fractal dimension[13] D_c of the set of all admissable binary fractions is given by the same formula as the topological entropy. Since increasing the length of an n cylinder is equivalent to increasing the resolution with which one specifies an element in Σ_f, the fractal dimension is given by

$$D_c = \lim_{n \to \infty} \frac{\log[N(n)]}{n}$$

So we see that the topological entropy of the subshift σ_f, and so of the map f, is equal to the fractal dimension of the binary fraction representation of the Cantor set Σ_f.

If, instead, we weight the count of n cylinders with their respective probabilities, we can define another dimensionlike quantity, the *information dimension* (Rènyi, 1959; Farmer, 1981; Kaplan and Yorke, 1981),

$$D_I = \lim_{n \to \infty} \sum_{i=1}^{n} \mu(S_i^n) \log[\mu(S_i^n)]$$

which we see is identical to the metric entropy of σ_f. Fractal and information dimensions can also be defined for the attractors of the original dynamical system. To avoid confusion in this regard, we should emphasize that these are not the same quantities as the corresponding dimensional quantities computed for the associated symbolic dynamical system.[14]

[13] Properly speaking the defined quantity is the capacity of the Cantor set Σ_f. For a more detailed description of these notions of dimension, see Mandelbrot (1977).

[14] For a more complete overview of the information dimension concept, see Farmer (1981).

7. ENTROPIES AND LYAPUNOV CHARACTERISTIC EXPONENTS

We pause now to introduce Lyapunov characteristic exponents as another measure of chaos, and to discuss their relationship to the entropies described above. The Lyapunov characteristic exponents measure the average asymptotic divergence rate of nearby trajectories in different directions of a system's state space (Benettin et al., 1980; Shimada and Nagashima, 1979). For our one-dimensional examples, f: $I \to I$, there is only one characteristic exponent λ. It can be easily calculated since the divergence of nearby trajectories is simply proportional to the derivative of f (Shaw, 1980):

$$\lambda = \lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \log| f'(x_n)|$$

Or equivalently, if a continuous ergodic asymptotic invariant measure μ exists, then the characteristic exponent is given by

$$\lambda = \int_0^1 \log| f'(x)|\, d\mu$$

We shall assume that such a measure exists, and moreover that

$$\mu_N(x) = \frac{1}{N} \sum_{n=1}^{N} \delta_{f^n x}$$

converges to μ as $N \to \infty$ for almost every initial condition (Oono and Osikawa, 1980).

If M is an axiom-A attractor, there is a prescription for constructing a partition which is generating, and the equality of the metric entropy h_μ and the sum of the positive Lyapunov characteristic exponents can be proven (Bowen and Ruelle, 1975). In fact, whenever an absolutely continuous invariant measure exists, a theorem due to Piesin (1977) shows that the metric entropy of a diffeomorphism is equal to the sum of the positive exponents.[15] Ruelle (1978) has shown the same equality for any map that has an absolutely continuous invariant measure, and Ledrappier (1981) has

[15] In the general case, the exponents are a function of initial condition, so the sum must be integrated over the attractor, but we will consider only the case of an ergodic attractor where the exponents are constant almost everywhere with respect to the asymptotic invariant measure.

constructed a proof of this for the logistic equation. Shimada obtained good agreement between the characteristic exponent and the metric entropy for the Lorenz attractor and its induced symbolic dynamics using only nine symbols, and Curry (1981) has computed a metric entropy slightly lower than the positive Lyapunov characteristic exponent for a two dimensional diffeomorphism (Henon's map).

The surprising result of our numerical experiments for the logistic equation is that for most parameter values, the metric entropy of the shift induced by certain measurement partitions converges extremely slowly (from above) to the Lyapunov exponent. Naturally, the metric entropy is always less than the topological entropy. This slow convergence is clearly illustrated in Figures 8 and 10, where we see an apparent discrepancy of $\sim 5\%$ for $r = 3.7$ and $n = 16$. This slow convergence does not appear in the symbolic dynamics calculation of the topological entropy. Using 16 symbols we find it rapidly converges (to within 0.1%) to the topological entropy computed using the kneading calculus of Milnor and Thurston (1977).

The validity of the symbolic dynamics entropy calculations is supported by the fact that we find excellent agreement (to within 0.1%) between the topological entropy of the shift and the topological entropy of the map, computed using the kneading calculus of Milnor and Thurston.

We also note that, to within numerical accuracy ($< 0.1\%$), at $r = 4.0$, both entropies (computed using 16 symbols) and the Lyapunov characteristic exponent converge to 1.0. At $r = 3.6785735\ldots$ where two bands merge into one, Ruelle (1977) has shown that there exists an absolutely continuous invariant measure, so that the metric entropy must be equal to the Lyapunov exponent (Ruelle, 1978). Our calculations using 16 symbols and 5×10^6 iterations, yield a topological entropy of 0.5000 ± 0.0001, a metric entropy of 0.497 ± 0.001, and a Lyapunov characteristic exponent 0.491 ± 0.0005. Thus, after 16 symbols, the metric entropy has converged to within 1% of the Lyapunov exponent. Similar slow convergence was found at other parameter values (e.g. $r = 3.7$). These calculations are thus consistent with Ruelle's and Ledrappier's results mentioned above. The calculations also provide numerical evidence for the existence of an absolutely continuous invariant measure for a wide range of parameter values.

8. OTHER ONE-DIMENSIONAL MAPS

In light of the slow convergence of the metric entropy to the Lyapunov exponent, one natural question is how universal this result is. Shimada has already found good agreement between the symbolic dynamics entropy and the characteristic exponent for the cusp return map of the Lorenz attractor. To compare our results with Shimada's, we will now consider a map which

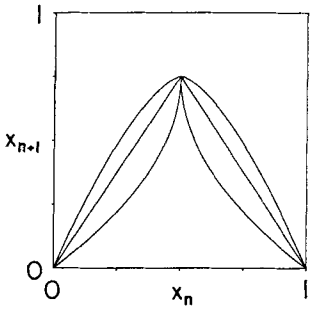

Fig. 11. Graph of $f(x)=a(1-|2x-1|^{1+\varepsilon})$ for $\varepsilon<0$ (cusp), $\varepsilon=0$ ("tent," or piecewise linear), and $\varepsilon>0$ (hump) and $a=3.7/4.0$.

changes under the continuous variation of a parameter from a map having a cusp, Lorenz-type maximum to one with a rounded, hump maximum. The map's functional form is

$$f(x)=a\left(1-|2x-1|^{1+\varepsilon}\right)$$

This map has a differentiable maximum ($f'=0$) for $\varepsilon>0$, a cusp maximum for $\varepsilon<0$, and is a piecewise linear tent map for $\varepsilon=0$; $a\in[0,1]$ determines the height of the map. Thus, $\varepsilon=1$ corresponds to the logistic equation studied in the previous section. Figure 11 shows three typical maps of this family with different values of ε. The results displayed in Figure 12 show that for sufficiently positive values of ε ($\sim>0.8$) there is the same discrepancy between the metric entropy and the characteristic exponent as seen in the logistic equation. Whereas for all $\varepsilon\sim<0.4$, we see good agreement

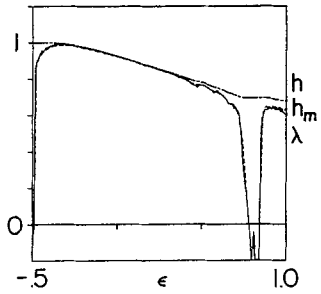

Fig. 12. Topological entropy (upper unevenly dashed line), metric entropy (dashed line), and Lyapunov exponent (solid line) for maps from cusps maps to smooth maximum maps, $\varepsilon=-0.5$ to 1.0. The entropies were calculated using the kneading calculus for the topological entropy and $H(16)-H(15)$ for the metric entropy, with 5×10^5 iterations.

between them. Shimada's results for the Lorenz attractor appear then to correspond to this latter regime, as one expects. As a reference, the upper curve in the figure is the topological entropy calculated from the kneading determinant of Milnor and Thurston (1977), to an accuracy of one part in 10^6. For ε in the approximate regime $[-0.3, 0.25]$ there is good (always $< 0.5\%$) agreement between all three quantities using 13 symbols. Below this regime, the topological entropy diverges from the other quantities which decrease rapidly to zero as $\varepsilon = 0.4736\ldots$ is approached. This particular parameter value corresponds to that at which the slope of a portion of the cusp map becomes less than one. Within this regime the origin is a stable attracting fixed point, although initial conditions may wander chaotically in portions of the map where the slope is greater than 1, before "decaying" into the attracting origin (Yorke and Yorke, 1979).

The tent map ($\varepsilon = 0$) displays the same band-merging bifurcation sequence as the logistic equation, but not period-doubling bifurcations. The symbolic dynamics entropy algorithm also has the same convergence problems near the critical parameter value where the slope is everywhere 1. As the slope of the tent map increases, both the topological and metric entropies converge quite readily to the characteristic exponent. One can show that these quantities are given by the logarithm of the tent map's slope and when calculated using symbolic dynamics they agree to within 0.1%.

The tent map limit $\varepsilon = 0$ is a symmetric map, for which the topological and metric entropies are equal. We also examined asymmetric piecewise linear maps for which the entropies are not equal. In these cases we find that there is good agreement ($\sim 0.1\%$) between the metric entropy and the Lyapunov exponent. Our numerical results indicate that any map that is *strictly hyperbolic*, i.e., one for which the absolute value of the first derivative is everywhere greater than 1, displays rapid convergence of the metric entropy the Lyapunov exponent. This may be related to the fact that these maps have no infinite singularities in their asymptotic invariant probability distributions (Lasota and Yorke, 1977).[16] In contrast to these systems, the logistic equation has a quadratic maximum that produces square root singularities in its asymptotic invariant probability distribution, and also displays the slow convergence.

From a practical perspective, these results indicate that a certain amount of caution must be used in the computation of the metric entropy from some experimental data set. If the underlying dynamics has singularities in its asymptotic distribution, then one might expect an overestimation of the metric entropy, and its slow convergence to the Lyapunov characteristic exponent.

[16] This fact may be seen from the Perron-Frobenius algorithm used to obtain the probability distribution; see, for example, Shaw (1980).

9. SYMBOLIC DYNAMICS IN THE PRESENCE OF FLUCTUATIONS

With these various theoretical results outlined, we now turn to more practical concerns: the effects of fluctuations and of finite precision on the application of symbolic dynamics to physical systems and their models. We shall address the first concern in this section and the latter in the next.

As discussed in the Introduction, symbolic dynamics appears to offer methods that circumvent the need for high-measurement resolution. Indeed, in the case of the one-dimensional maps discussed above, the measurement requirement was only 1 bit of resolution for the two-element partition $\{[0, d), [d, 1]\}$. We should point out, in contrast to this encouraging prospect, that the data base requirements can be quite substantial. For the calculation of the n-cylinder probability distributions the required memory increases as 2^{hn}, where h is the topological entropy. In comparison with other similar techniques, this is not an unusual requirement—in fact, it is typical. A second and more troubling feature of the application of symbolic dynamics to experimental systems is the requirement of a generating partition if one is to measure the "true" entropies. Such a partition is not given *a priori* to an experimentalist and, therefore, this requirement presents a significant theoretical limitation of experimental symbolic dynamics. On the practical side, however, if one keeps in mind the effects of varying the partition on the calculated entropies as shown in Figure 10, one could alter the partition to look for a maximum in the entropies. Such a procedure would be computationally intractable if the partition elements were of high dimension, but for simple low-dimensional chaotic behavior, at least, it is straightforward and feasible.

A further consideration is that the information acquisition rate must be greater than the metric entropy. That is, if the *a priori* information gain per measurement is I and the sample rate is R, then we must have $IR > h_\mu$ in order to resolve the deterministic chaotic dynamics. For example, using a binary partition at each iteration of a one-dimensional map, one cannot measure an entropy greater than one bit per iteration. We shall leave these problems, of data base, generating partition, and sample rate requirements, for future consideration and discuss the effect of fluctuations, or external noise, on the experimental application of symbolic dynamics.

Every experimental system is immersed in a "heat bath"; in other words, there will always be couplings between the system of interest and external degrees of freedom that are beyond observation by an experimentalist. To model this situation, we introduce the stochastic logistic equations which contains an explicit noise term,

$$x_{n+1} = rx_n(1 - x_n) + \tau_n$$

where τ_n is a uniform random variable with zero mean and standard deviation s. We shall call s the *noise level* of the fluctuations. The general effects of fluctuations on the logistic equation and the period-doubling bifurcation have been discussed elsewhere (Crutchfield and Huberman, 1980; Crutchfield et al., 1981; Haken and Mayer-Kress, 1981). We shall only briefly allude to these results and, instead, concentrate on the effects of fluctuations on the symbolic dynamics at a few parameter values of interest.

The values of the topological and metric entropies of the shift induced by the map in the presence of fluctuations depend on the manner in which the fluctuations alter the n-cylinder probability distributions. Changes in the characteristic exponent depend in similar way on the distribution $P(x)$ of iterates on the interval (recall Figure 1). Fluctuations smooth out the square root singularities in $P(x)$ and, for $r = 3.7$, affect the characteristic exponent very little ($<1\%$) over a wide range in noise level: $s \in [0.0, 10^{-3}]$ (Crutchfield et al., 1981). The sequence of n-cylinder probability distributions of Figure 13 shows a similar mild change in Cantor set structure up to $n = 6$. This figure should be compared to Figure 3. We should point out that the fluctuations do *not* truncate the Cantor set structure, but only produce more admissable symbols. In order to see any alteration in this sequence, it was necessary to use the relatively large noise level of $s = 10^{-2}$. At $n = 6$, new admissable symbols appear, which in turn should result in increased entropies. Indeed, Figure 14 shows that the fluctuations increase $H(n)$ and $\log[N(n)]$ monotonically at $r = 3.7$ over three noise levels: $s = 0.0, 10^{-3}, 10^{-2}$. The metric and topological entropies, which are the slopes of these curves, clearly increase with increasing noise level, but converges to values less than one.

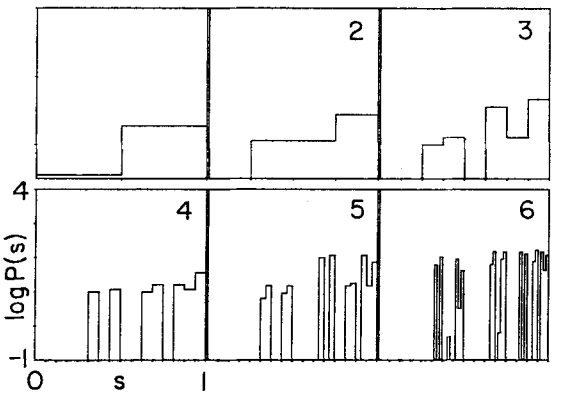

Fig. 13. n-Cylinder probability distribution (as in Figure 3) for the stochastic logistic equation with noise level $s = 10^{-2}$ at $r = 3.7$.

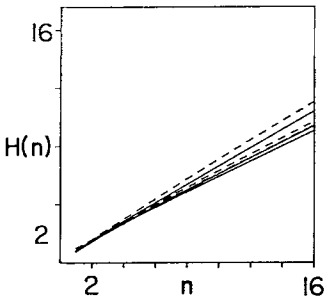

Fig. 14. Entropies $H(n)$ (solid line) and $\log[N(n)]$ (dashed) as a function of symbol length at $r = 3.7$ for three noise levels: $s = 10^{-2}$ (upper pair of dashed and solid lines), 10^{-3} (middle pair), and 0.0 (lowest pair). Note that $\log[N(n)]$ for $s = 0.0$ overlaps $H(n)$ for $s = 10^{-3}$. 2×10^5 iterations were used.

Figures 15 and 16 compare the effect of increased noise level at the parameters where four bands merge into two and at a period-4 orbit, respectively. For low noise levels $H(n)$, and so h and h_μ, differs substantially between these two parameters. At larger noise levels, however, the measured entropies are quite similar. In fact, at some noise level they would be completely indistinguishable: the period-4 orbit would appear as four bands (Crutchfield and Huberman, 1980).

The effects of external noise on the period-doubling bifurcation sequence can be described by a scaling theory and renormalization group approach (Crutchfield et al., 1981; Schraiman et al., 1981). These effects are reminiscent of an external magnetic field acting on a ferromagnet: noise acts as a disordering field on the chaotic dynamics. The characteristic exponent and the topological and metric entropies are the disorder parameters of the chaotic system, which scale in a manner similar to the magnetization order parameter for magnetic systems. In particular, the entropies in

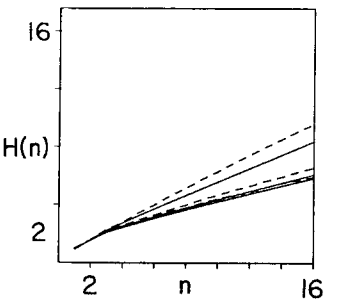

Fig. 15. Same details as Figure 14 except at the bifurcation from four bands to two, $r = 3.59257218\ldots$.

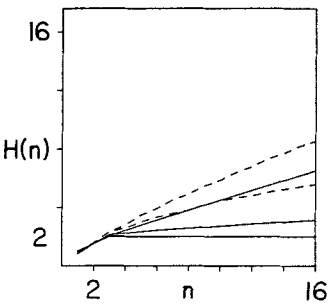

Fig. 16. Same details as Figure 14 except at period-4 orbit, $r=3.52$.

the presence of fluctuations provide a stronger analogy to the magnetization in the presence of an external field than does the Lyapunov exponent. Just as the magnetization has a nonzero tail above T_c at finite field, the entropies are nonzero in the periodic regime with noise added. The Lyapunov exponent, however, is negative in the periodic regime.

Adding external noise also increases the rate of convergence of the entropies. As an example of this, recall that at the merging of two bands into one, the topological entropy $h(\sigma_f)$ oscillates indefinitely, when calculated as the two-point slope $H_\mu(n)-H_\mu(n-1)$. When noise is added, the oscillation is "damped" and the topological entropy readily converges, albeit to a larger value than found with no noise added. A comparison of the zero noise case and that with $s=10^{-2}$ is shown in Figure 17. As the metric entropy decreases, the observer gains information about correlations between the observed symbols. When noise is added, these correlations decay, and so the metric entropy converges more rapidly.

In summary, the effects of additive fluctuations on symbolic dynamics calculations can be consistently described and so taken into account in the application of symbolic dynamics to experiments whose dynamics may be described by one dimensional maps. The entropies calculated from the n-cylinder distributions follow the deterministic values up to some n, above which the external noise causes $H(n)$ and $\log[N(n)]$ to increase at a rate greater than the zero noise case. Rather than considering this fact as compromising the usefulness of symbolic dynamics, one might suggest the use of symbolic dynamics entropies as measures of level of external fluctuations to which an experimental system is coupled.

One might conclude that, insofar as the effects of fluctuations can be described, turbulent physical systems subject to external fluctuations can be modeled by chaotic dynamical systems with added noise. This modeling process often involves the use of computer simulations which require the discretization of the model dynamics. In contrast to the apparently tractable case of external fluctuations, the effects of discretizing chaotic dynamics are

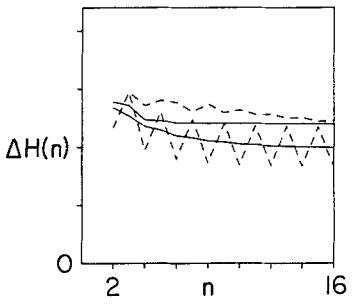

Fig. 17. Topological entropy approximated by $\log[N(16)]-\log[N(15)]$ (dashed line), and metric entropy approximated by $H(16)-H(15)$ (solid line), as a function of symbol length. The lower pair of curves corresponds to the deterministic case ($s=0$) and the upper pair to $s=10^{-2}$. 2×10^5 iterations were used.

not so well understood. We shall address questions related to finite-state simulations of chaotic dynamical systems and symbolic dynamics in the following section.

10. EFFECTS OF FINITE PRECISION

Simulations of chaotic systems are often carried out on finite-state machines. This raises questions about the relationship between the observation of turbulent physical behavior and the chaotic behavior observed in a finite precision simulation used for a system's representation. In particular, the finite number of states available for the simulation implies that any simulated orbit will be ultimately periodic. We will now address the question of how chaotic behavior may be successfully quantified in spite of this periodicity. The simulated orbits, in fact, retain many of the statistical properties of the "ideal" continuous system (Erber et al., 1980).

In most simulations of chaotic dynamical systems one assumes that the errors introduced by a finite-state computer's roundoff will play a role analogous to fluctuations (e.g., thermal fluctuations). When comparing model simulations to observed behavior, though, one must be careful to distinguish between the resolution of a measuring instrument, the noise level of external fluctuations, and the internal precision with which the simulations are carried out. We will be concerned in this section with the latter topic. We will investigate several deterministic rounding algorithms and present a model that allows a continuous transition between a deterministic algorithm and a totally random rounding algorithm which is equivalent to adding noise of a specified magnitude. With each of these models we will compute the entropies described above, examining the effects of finite precision on their values.

We begin with the simplest version of a roundoff algorithm: truncation. For these and all the calculations presented below we have reduced the unit interval state space to 2^k points by using k bits of precision. The iteration of the one-dimensional map was done with much greater precision (floating point), but after every iteration the state was rounded to one of the 2^k discrete states. For the case of truncation, this is accomplished by always rounding down to the next lower state.

Figure 18 displays $\log[N(n)]$ (dashed) and $H(n)$ vs n for a range of different precisions: 10, 15, 20, 25, and 30 bits. We naively might expect the graph to level off when the number of symbols n reaches the precision being used, since the orbit then necessarily becomes periodic with a period less than 2^n, and this is indeed what we see in the figure. The case $k = 10$ might seem a bit puzzling since $H(n)$ remains so low, while $\log[N(n)]$ increases, but this is due to the probability distribution being composed of two delta functions with skirts which correspond to infrequently occurring symbols.

The monotonicity of $H(n)$ with increasing precision is due to a fortuitous choice of precisions. Figure 19 shows $H(12)$ and $\log[N(12)]$ as a function of precision, and we see quite a bit of nonmonotonicity with increasing precision, due to the existence of stable, relatively low period orbits. We also see that the topological entropy converges sooner, and is less affected by the finite precision.

The second rounding procedure we try is deterministic round off: after the map is iterated once, the result is rounded down if the fractional part of $2^k f(x)$ is less than 0.5, and rounded up otherwise. Figure 20 shows that this procedure yields the same nonmonotonic results as illustrated for truncation in Figure 19, but with slightly different features.

To study the relationship between the deterministic rounding algorithms described above and the effects of random fluctuations as described in the previous section, we will now look at a "hybrid" rounding algorithm.

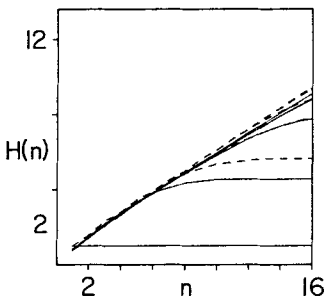

Fig. 18. Entropies $H(n)$ (solid) and $\log[N(n)]$ (broken) vs. n for truncation algorithm at various precisions, $k = 10$, 15, 20, 25, 30 bits; with the lowest line of each corresponding to 10 bits. $r = 3.7$, 10^5 iterations.

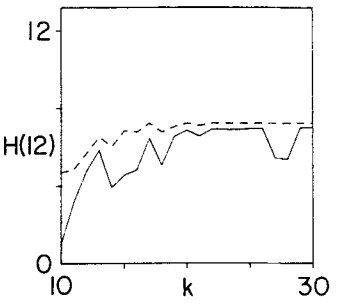

Fig. 19. $H(12)$ and $\log[N(12)]$ as a function of precision k, using truncation. $r=3.7$, 10^5 iterations.

Now the result of an iteration $f(x)$ will be rounded up or down exactly as in the deterministic round off algorithm unless the fractional part falls in a window about 0.5, in which case a random number is used to make the rounding decision. If the window width is zero, the result is equivalent to deterministic roundoff; if the window width is 1, the result is equivalent to adding a random fluctuation of magnitude 2^{-k}. The results of varying window width are illustrated in Figure 21, where we graph $H(n)$ and $\log[N(n)]$ vs n for different window widths, using a constant 20 bits of precision.

It is worth noting here that flipping the twentieth bit at random (i.e., using a window width of 1 at a precision of 20 bits) allows one to measure $H(16)$ as well as is possible using much higher precision (e.g., double-precision floating point). This point is further illustrated by Figure 22, where we see that $H(12)$ and $\log[N(12)]$ quickly converge to a constant value as the precision is increased for the window width of 1. These entropy values agree to within 1% of the double-precision floating point results.

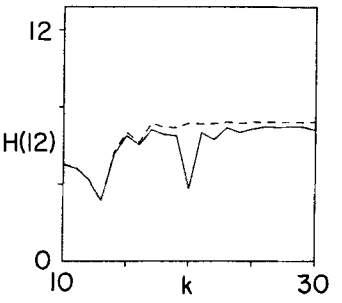

Fig. 20. Same as Figure 19 except for deterministic roundoff.

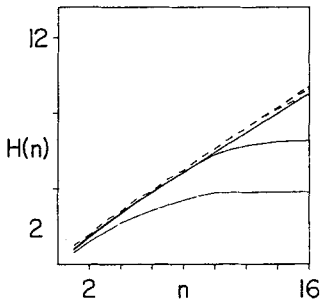

Fig. 21. $H(n)$ (solid) and $\log[N(n)]$ (dashed) vs. n using random roundoff algorithm at various window widths: 10^{-3}, 10^{-2}, 10^{-1}, 1.0, using 20 bits of precision. For each width, the $\log[N(n)]$ dashed curves remain close to each other, as do the $H(n)$ curves for widths 10^{-1} and 1.0, at the top of the figure. The lowest curve corresponds to $H(n)$ for a width of 10^{-3}, the next higher 10^{-2}, $r = 3.7$, 10^5 iterations.

Our results lead us to the practical suggestion that simulations of dynamical systems should include noise at the lowest bit of precision in order to effectively increase the precision and yield the physically relevant behavior. A simple example of this will illustrate its utility. Consider the tent map

$$x_{n+1} = \begin{cases} 2x_n, & x_n < 0.5 \\ 2(1-x_n), & x_n \geq 0.5 \end{cases}$$

which is equivalent to a binary shift of the initial condition x_0. If x_0 is specified to k bits, then after k iterations of the tent map all of the initial condition has been removed and the truncation (or deterministic roundoff) algorithm has replaced the lost bits with 0's. The net effect is that after only k iterations the orbit has trivially converged to 0, making the simulation remarkably unrepresentative. A numerically accumulated probability distribution of this map would be a delta function at the origin, although the map

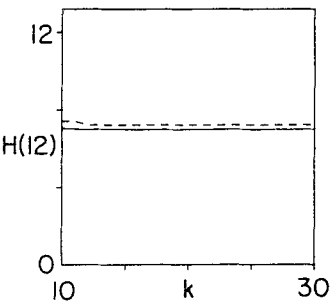

Fig. 22. Same as Figure 19 except for random roundoff algorithm with a window of 1.0.

is provably chaotic with positive entropies and characteristic exponent. Admittedly, this is an extreme example, but similar effects due to finite precision with deterministic roundoff should be expected in simulation of chaotic dynamical systems. To render such a finite-state simulation useful, we propose that the roundoff decision be made randomly. One would then observe the chaotic dynamics implicit in the mathematical specification of the map and also the behavior which most closely models real physical systems. The addition of this random degree of freedom to finite-state computers renders them effectively infinite-state machines.[17]

11. ENTROPY AND ALGORITHMIC COMPLEXITY

Both the topological and metric entropies described above characterize, in some sense, the "randomness" of a chaotic dynamical system. We shall now discuss the relationship between these concepts of randomness and another developed in the early 1960s independently by Solomonoff (1964),[18] Chaitin (1966), Kolmogorov (1965), and Martin-Lof (1966). Their essential idea is to define the *algorithmic complexity* $A(s^n)$ of a string of symbols $s^n = (s_1, \ldots, s_n)$ as the minimum size of a computer program required to generate the string. Without going into the subtleties of how to make this definition precise, we will present a heuristic discussion of how this concept may be related to chaotic dynamical systems. A summary of the technical details involved is given by Alekseyev and Yacobson (1981).

For a chaotic dynamical system f: $M \to M$ with a generating measurement partition S, we showed in Section 2 that to each orbit $(x_0, x_1, x_2, \ldots)$ there corresponds a symbol sequence $(s_0, s_1, s_2, \ldots)$, where

$$x \in \bigcap_{i=0}^{\infty} f^{-1}(S_{s_i})$$

The algorithmic complexity $A(x_0, f)$ of this orbit may then be defined as

$$A(x_0, f) = \lim_{N \to \infty} \frac{A(s^N)}{N}$$

We then have the results that the algorithmic complexity is bounded by the

[17] The source of this random decision could be any of a number of chaotic electronic circuits and so be easily implemented in the current generation of LSI arithmetic processors. In such a dynamical systems processor one would, of course, want to switch between deterministic roundoff and random roundoff to aid in software error checking.

[18] For the first presentation of Solomonoff's ideas see Minsky (1962).

topological entropy (Brudno, 1978; Zvonkin and Levin, 1970; Kamae, 1973):

$$A(x_0, f) \leqslant h(f)$$

and that, for any invariant measure μ, it equals the metric entropy,

$$A(x_0, f) = h_\mu(f)$$

for all x_0 (except some measure zero set). These results justify the use of algorithmic complexity as a measure of randomness for a chaotic dynamical system, and allows us to set $A(f) = h_\mu(f)$ if f is ergodic.

There may seem to be a paradox: even if the map f is chaotic, an n cylinder s^n may be obtained simply by iterating the map n times and observing the sequence of partition elements visited ... a very simple algorithm indeed! This algorithm for generating the n cylinder s^n is simple in the sense that to generate a longer n cylinder, only the parameter which specifies the number of iterations need be changed. Thus, the algorithm's length $A(s^n)$ grows like $\log(n)$ so that

$$\lim_{n \to \infty} \frac{A(s^n)}{n} \to 0$$

contradicting $A(f) = h_\mu(f) > 0$. We must realize, however, that the algorithm must also contain a specification of the initial condition, and if the map is chaotic, the amount of information contained in the observed symbol sequence is proportional to the amount of information contained in the specification of the initial condition. Thus, the size $A(s^n)$ must grow with n.

We may now use algorithmic information theory to formalize the notions of modeling mentioned in the Introduction. Again, we will consider an observer who makes a sequence of measurements $(s_0, s_1, \ldots)$ on a physical system with some instrument whose output is one of finitely many symbols. In this context, we may define a *predictive model* as an algorithm **A** which would produce the string $s^n = (s_0, \ldots, s_n)$ for any n. As a simple example, we see that if the system is executing periodic motion, the symbol sequence of successive measurements will also be periodic, so that a simple program could predict which symbol would be observed at any time in the future. The result that the algorithmic complexity $A(s^n)$ of n observations of a chaotic dynamical system grows like n (i.e., $A(f) > 0$) is then a concise statement of the inability of the observer to construct such a predictive model, since there is always some n for which $A(s^n)$ is larger than the size of any proposed predictive model **A**.

12. CONCLUDING REMARKS

One of our primary concerns here has been an elucidation of the *modeling process*, by which we mean a comparison between the observed temporal behavior of some physical system and the temporal behavior of a model comprised of some other representation system. If the physical system behaves periodically, a model can be constructed, in principle, that serves a *predictive* function.

In modeling turbulent physical systems with simple chaotic dynamical systems certain problems arise since one cannot, in principle, obtain exact correspondence between the observed physical behavior and the output of the model. Nevertheless, the chaos displayed by chaotic dynamical systems admits certain geometrical and statistical characterizations which must then be used in the modeling process as the new criteria for a model's validity.

The construction of symbolic dynamics, for example, enables a comparison between the statistical properties of a physical system and those of a model by providing measures of temporal complexity. At the same time, though, by projecting large portions of the state space onto a discrete set of symbols most, if not all, geometrical information is lost. The construction of symbolic dynamics reduces the dynamics of a physical system to a dynamical system comprised of a trivial dynamic (the shift) on a complicated state space. All the complexity of the temporal behavior of the physical system is contained in the structure of the symbol state space Σ_f. This complexity admits not only a dynamical quantification via the metric entropy[19] but also an *algorithmic* quantification.

The algorithmic interpretation considers the physical system as a finite-state machine, with the original dynamics corresponding to an algorithm for a Turing machine, say. If the observed dynamics are chaotic, the minimal algorithm needed to specify any subsequence in Σ_f grows with the length of the subsequence. And the rate of this growth is equal to the metric entropy of the induced shift on Σ_f. If the observed dynamics are not chaotic, however, Σ_f reduces to a periodic lattice which can be specified in its entirety with a finite algorithm.

Turing machines were invented as conceptual tools to make precise the notion of computability. Mathematical propositions were then recast as Turing machine algorithms and a proposition's decidability could be shown to be equivalent to its algorithm's computability. Thus an equivalence was established between the notions of computability and decidability. Via algorithmic information theory, then, symbolic dynamics appears as a link between turbulent physical dynamics and the notions of decidability, often

[19] With appropriate restrictions on the measurement partition, this will also be the metric entropy of the observed physical system.

associated with Godel's theorem. Shaw (1980) suggested a similar connection between the unpredictability arising from chaotic dynamics and the undecidability of propositions about a chaotic system's state after the information obtained from a measurement has been lost. Here we have attempted to point out that symbolic dynamics provides a common language for the discussion of these ideas.

As mentioned in the Introduction, the existence of deterministic, unpredictable behavior obviates the hope of "closed-form" descriptions, and so allows only for statistical and geometric characterizations. For chaotic dynamical systems, the explicit prediction of behavior is no longer possible, and we are led to reconsider the criteria for the appropriateness of models of turbulent phenomena. Chaos, in this sense, necessitates a generalization of the modeling process.

Underlying these conceptual motivations, there are many practical questions concerning the construction of symbolic dynamics. We addressed some of these by examining the conditions under which digital computations can successfully approximate chaotic dynamics of a continuous system driven with "thermal" fluctuations. We also discussed the applicability of symbolic dynamics techniques to experimental data analysis. Our research along these lines has also uncovered certain theoretical questions, such as why we observe a slow convergence of the metric entropy to the Lyapunov characteristic exponent for the logistic equation. We have been able to indicate, only in a preliminary way, the usefulness of dynamical systems and symbolic dynamics in describing observed unpredictable behavior.

In the early literature on the topics we have discussed here, there appears to have been a general appreciation of the connection between information theory, algorithms, and dynamics. Such a broad perspective, which flourished during the days of the first computers, seems to have motivated much of Shannon's information theory, and subsequently Kolmogorov's ideas on dynamical systems. In the present work we have discussed chaotic dynamics within this context in an effort to rekindle the apparently diffused understanding of the intimate connection between the physics of dynamics and computation.

ACKNOWLEDGMENTS

We are grateful to Doyne Farmer, Joe Ford, David Fried, John Guckenheimer, David Rand, and Rob Shaw for stimulating discussions and correspondence on these topics. The authors also wish to thank J.-P. Eckmann, F. Ledrappier, M. Misiurewicz, and others for useful discussions on the relationship between the metric entropy and Lyapunov characteristic exponents. JPC was supported by a UC Regents Fellowship. This work has been supported by NSF grant No. 443150-21299.

REFERENCES

Adler, R. L., Konheim, A. G., and McAndrew, M. H. (1965). "Topological Entropy," *Transactions of the American Mathematical Society*, **114**, 309.

Alekseyev, V. M., and Yakobson, M. V. (1981). "Symbolic Dynamics and Hyperbolic Dynamic Systems," *Physics Reports*, to be published.

Benettin, G., Casartelli, M., Galgani, L., Giorgilli, A., and Strelcyn, J.-M. (1978). "On the Reliability of Numerical Studies of Stochasticity," *Nuovo Cimento*, **44B**, 183; **50B**, 211.

Benettin, G., Galgani, L., Giorgilli, A., and Strelcyn, J.-M. (1980). "Lyapunov Characteristic Exponents for Smooth Dynamical Systems and for Hamiltonian Systems; A Method for Computing All of Them, Parts I and II," *Meccanica*, 21.

Billingsley, P. (1965). *Ergodic Theory and Information*. John Wiley and Sons, New York.

Bowen, R. (1975). "Equilibrium States and the Ergodic Theory of Anosov Diffeomorphisms," *Lecture Notes in Mathematics*, **470**, Springer Verlag.

Bowen, R., and Ruelle, D. (1975). "Ergodic Theory of Axiom-A," *Inventiones Mathematicas*, **29**, 181.

Brudno, A. A. (1978). *Uspekhi Matematicheskikh Nauk*, **33**, 207.

Chaitin, G. J. (1966). "On the Length of Programs for Computing Binary Sequences," *J. Assoc. Computing Machinary*. **13**, 547; See also "Randomness and Mathematical Proof," *Sci. Amer.* May 1975, 47.

Chillingworth, D. R. J. (1976). *Differential Topology With a View to Applications*. Pitman, San Fransisco, California.

Collet, P., and Eckmann, J.-P. (1980). *Iterated Maps of the Unit Interval as Dynamical Systems*. Birkhauser, Cambridge, Massachusetts.

Crutchfield, J. P., Nauenberg, M., and Rudnick, J. (1981). "Scaling for External Noise at the Onset of Chaos," *Physical Review Letters*, **46**, 933.

Crutchfield, J. P., and Huberman, B. A. (1980). "Fluctuations and the Onset of Chaos," *Phys. Lett.*, **77A**, 407.

Crutchfield, J. P., Farmer, J. D., and Huberman, B. A. (1981). "Fluctuations and Simple Chaotic Dynamics," *Physics Reports*, to appear.

Curry, J. H. (1981). "On Computing the Entropy of the Hénon Attractor," *Inst. des Hautes Études Sci.* preprint.

Dinaburg, E. I. (1970). "The Relation Between Topological Entropy and Metric Entropy," Sov. Math. **11**, 13.

Erber, T., Everett, P., and Johnson, P. W. (1980). "The Simulation of Random Processes on Digital Computers with Chebyshev Mixing Transformations," preprint.

Farmer, J. D. (1981). "Order Within Chaos," UCSC dissertation.

Froehling, H., Crutchfield, J. P., Farmer, J. D., Packard, N. H., and Shaw, R. S. (1981). "On Determining the Dimension of Chaotic Flows," *Physica* **3D**, 605.

Guckenheimer, J., Moser, J., and Newhouse, S. E., (1980). *Dynamical Systems*. Birkhauser, Cambridge, Massachusetts.

Haken, H., and Mayer-Kress, G. (1981). "The Effect of Noise on the Logistic Equation," *J. Stat. Phys.* to appear.

Jonker, L., and Rand, D. (1980). "Bifurcations in One Dimension," *Inventiones Mathematicas*, **62**, 347.

Kamae, T. (1973). "On Kolmogorov's Complexity and Information," *Osaka Journal of Mathematics*, **10**, 305.

Kaplan, J., and Yorke, J. (1981). Private communication.

Kifer, Ju. I. (1974). *USSR Izvestija*, **8**, 1083.

Kolmogorov, A. N. (1958). *Doklady Akademii Nauk*, **119**, 754.

Kolmogorov, A. N. (1965). "Three Approaches to the Quantitative Definition of Information," *Problems of Information Transmission (USSR)*, **1**, 1.

Lasota, L., and Yorke, J. (1976). "On the Existence of Invariant Measures for Transformations with Strictly Turbulent Trajectories," *Bull. Acad. Pol. Sci.*, **25**, 233.

Ledrappier, F. (1981). "Some Properties of Absolutely Continuous Invariant Measures on an Interval," *Ergod. Theo. Dyn. Sys.*, **1**, 77.

Lorenz, E. N. (1963). "Deterministic Non-Periodic Flow," *Journal of Atmospheric Science*, **20**, 130.

Mandelbrot, B. (1977). *Fractals: Form, Chance, and Dimension*. W. H.. Freeman, San Francisco, California.

Martin-Löf, P. (1966). "The Definition of Random Sequences," *Information Control*, **9**, 602.

Milnor, J., and Thurston, W. (1977). "On Iterated Maps of the Interval, I and II," Princeton University preprint.

Minsky, M. L. (1962). "Problems of Formulation for Artificial Intelligence," in *Mathematical Problems in the Biological Sciences, Proceedings of Symposia in Applied Mathematics XIV*, R. E. Bellman, ed. American Mathematical Society, Providence, Rhode Island.

Oono, Y., and Osikawa, M. (1980). "Chaos in Nonlinear Difference Equations I," *Progress in Theoretical Physics*, **64**, 54.

Packard, N. H., Crutchfield, J. P., Farmer, J. D., and Shaw, R. S. (1980). "Geometry from a Time Series," *Phys. Rev. Lett.*, **45**, 712.

Parry, W. (1964). "Intrinsic Markov Chains," *Transactions of the American Mathematical Society*, **122**, 55.

Piesin, Ya. B. (1977). "Characteristic Lyapunov Exponents and Smooth Ergodic Theory," *Uspeki Matematicheskikh Nauk*, **32**, 55.

Renyi, A. (1959). "On the Dimension and Entropy of Probability Distributions," *Acta Math. Hung.*, **10**, 193.

Ruelle, D. (1977). "Applications Conservant une Mesure Absolument Continue par rapport a dx sur [0, 1]," *Communications in Mathematical Physics*, **55**, 47.

Ruelle, D., and Takens, F. (1971). "On the Nature of Turbulence," *Communications in Mathematical Physics*, **20**, 167.

Ruelle, D. (1978). "An Inequality for the Entropy of Differentiable Maps," *Bull. Soc. Brasil Math.*, **9**, 331.

Schraiman, B., Wayne, C. E., and Martin, P. C. (1981). "Scaling Theory for Noisy Period-Doubling Transitions to Chaos," *Physical Review Letters*, **46**, 935.

Shaw, R. (1980). "On the Predictability of Mechanical Systems," UCSC dissertation.

Shaw, R. (1981). "Strange Attractors, Chaotic Behavior, and Information Flow," *Zeitschrift fuer Naturforschung*, **36a**, 80.

Shimada, I. (1979). "Gibbsian Distribution on the Lorenz Attractor," *Progress in Theoretical Physics*, **62**, 61.

Shimada, I., and Nagashima, T. (1979). "A Numerical Approach to Ergodic Problems of Dissipative Dynamical Systems," *Progress Theoretical Physics*, **61**, 1605.

Sinai, Ya. (1972)."Gibbsian Measures in Ergodic Theory," *Russian Mathematical Surveys*, **27**, 21.

Smale, S. (1967). "Differentiable Dynamical Systems," *Bulletin of the American Mathematical Society*, **13**, 747.

Solomonoff, R. J. (1964). "A Formal Theory of Inductive Control," *Information Control*, **7**, 224.

Takens, F. (1980). "Detecting Strange Attractors in Turbulence," preprint.

Yorke, J. A., and Yorke, E. D. (1979). "Metastable Chaos: The Transition to Sustained Chaotic Behavior in the Lorenz Model," *Journal of Statistical Physics*, **21**, 263.

Zvonkin, A. K., and Levin L. A. (1970). "The Complexity of Finite Objects and the Development of the Concepts of Information and Randomness by means of the Theory of Algorithms," *Russian Mathematical Survey*, **25**, 83.