Physica 9D (1983) 189–208
North-Holland Publishing Company

MEASURING THE STRANGENESS OF STRANGE ATTRACTORS

Peter GRASSBERGER† and Itamar PROCACCIA
Department of Chemical Physics, Weizmann Institute of Science, Rehovot 76100, Israel

Received 16 November 1982
Revised 26 May 1983

We study the correlation exponent ν introduced recently as a characteristic measure of strange attractors which allows one to distinguish between deterministic chaos and random noise. The exponent ν is closely related to the fractal dimension and the information dimension, but its computation is considerably easier. Its usefulness in characterizing experimental data which stem from very high dimensional systems is stressed. Algorithms for extracting ν from the time series of a single variable are proposed. The relations between the various measures of strange attractors and between them and the Lyapunov exponents are discussed. It is shown that the conjecture of Kaplan and Yorke for the dimension gives an upper bound for ν. Various examples of finite and infinite dimensional systems are treated, both numerically and analytically.

1. Introduction

It is already an accepted notion that many nonlinear dissipative dynamical systems do not approach stationary or periodic states asymptotically. Instead, with appropriate values of their parameters, they tend towards strange attractors on which the motion is chaotic, i.e. not (multiply) periodic and unpredictable over long times, being extremely sensitive on the initial conditions [1–4].

A natural question is by which observables this situation is most efficiently characterized. Even more basically, when observing a seemingly strange behaviour, one would like to have clear-cut procedures which could exclude that the attractor is indeed multiply periodic, or that the irregularities are e.g. caused by external noise [5].

The first possibility can be ruled out by making a Fourier analysis, but for the second one has to turn to some other measures. These measures should be sensitive to the *local* structure, in order to distinguish the blurred tori of a noisy (multi-) periodic motion from the strictly deterministic motion on a fractal. Also, they should be able to distinguish between different strange attractors.

In this paper we shall propose such a measure. Before doing so we shall discuss however the existing approaches to the subject.

In a system with F degrees of freedom, an attractor is a subset of F-dimensional phase space towards which almost all sufficiently close trajectories get "attracted" asymptotically. Since volume is contracted in dissipative flows, the volume of an attractor is always zero, but this leaves still room for extremely complex structures.

Typically, a strange attractor arises when the flow does not contract a volume element in *all* directions, but stretches it in some. In order to remain confined to a bounded domain, the volume element gets folded at the same time, so that it has after some time a multisheeted structure. A closer study shows that it finally becomes (locally) Cantor-set like in some directions, and is accordingly a fractal in the sense of Mandelbrot [6].

Ever since the notion of strange attractors has been introduced, it has been clear that the Lyapunov exponents [7, 8] might be employed in studying them. Consider an infinitesimally small F-dimensional ball in phase space. During its

† Permanent address: Department of Physics, University of Wuppertal, W. Germany.

0167-2789/83/0000–0000/$03.00 © 1983 North-Holland

evolution it will become distorted, but being infinitesimal, it will remain an ellipsoid. Denote the principal axes of this ellipsoid by $\epsilon_i(t)$ $(i = 1, \ldots, F)$. The Lyapunov exponents λ_i are then determined by

$$\epsilon_i(t) \approx \epsilon_i(0)\, e^{\lambda_i t}. \tag{1.1}$$

The sum of the λ_i, describing the contraction of volume, has of course to be negative. But since a strange attractor results from a stretching and folding process, it requires at least one of the λ_i to be positive. Inversely, a positive Lyapunov exponent implies sensitive dependence on initial conditions and therefore chaotic behaviour.

One drawback of the λ_i's is that they are not easily measured in experimental situations. Another limitation is that while they describe the *stretching* needed to generate a strange attractor, they don't say much about the *folding*.

That these two are at least partially independent is best seen by looking at a horshoe-like map† embedded in 3-dimensional space (fig. 1). Assume that each step of the evolution consists of (i) stretching in the x-direction by a factor of 2, (ii) squeezing in the y- and z-direction by different factors $\mu_z < \mu_y < \frac{1}{2}$, and (iii) folding in the (x, y) plane (fig. 1a) or in the (x, z) plane (fig. 1b). From fig. 1 one realizes already that the attractor will in both cases be a Cantorian set of lines, being more "plane-filling" in the first case than in the second case. Indeed, using the results of Section 7, one finds easily that the fractal dimensions are $D_a = 1 + \ln 2/|\ln \mu_y|$ and $D_b = 1 + \ln 2/|\ln \mu_z|$, respectively.

It is this fractal (or Hausdorff–Besikovich) dimension which has until now attracted most attention [9–14] as a measure of the local structure of fractal attractors. In order to define it [5], one first covers the attractor by F-dimensional hypercubes of side length l and considers the limit $l \to 0$. If the minimal number of cubes needed for the covering grows like

$$M(l) \underset{l \to 0}{\simeq} l^{-D}, \tag{1.2}$$

the exponent D is called the Hansdorff dimension of the attractor [5].

Being a purely geometric measure, D is independent of the frequency with which a typical trajectory visits the various parts of the attractor.

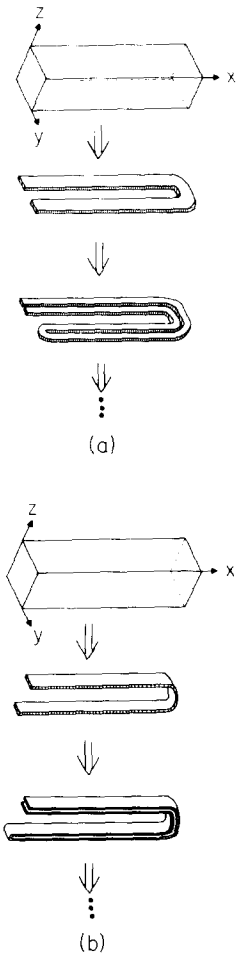

Fig. 1. Shape of an originally rectangular volume element after two iterations, each consisting of stretching, squeezing and folding. In fig. 1a (1b), the folding is in the $y(z)$-direction, which is the direction of lesser (stronger) squeezing.

† Notice that this is not a Smale's horseshoe. We also neglect in the following the bent parts of the horseshoe, in comparison to the parallel parts (i.e. we assume $L_x \gg L_y, L_z$; see fig. 1).

Even if these frequencies are very inequal, developing maybe even singularities somewhere, all parts contribute to D equally. It has been documented [12, 14] that the calculation of D is exceedingly hard and in fact impractical for higher dimensional systems.

Another measure which has been considered and which is sensitive to the frequency of visiting, is the information entropy of the attractor. By "information entropy" here we understand the information gained by an observer who measures the actual state $X(t)$ of the system with accuracy l, and who knows all properties of the system but not the initial condition $X(0)$. This is very similar to the entropy in statistical mechanics if we relate $X(t)$ to the microstate ($F \approx 10^{23}$), and the "system" to the macrostate. It is *not* the Kolmogorov entropy which is essentially the sum of all positive Lyapunov exponents.

Using the above partition of phase space into cells with length l, the information entropy can be written as

$$S(l) = -\sum_{i=1}^{M(l)} p_i \ln p_i , \tag{1.3}$$

where p_i is the probability for $X(t)$ to fall into the ith cell. For all attractors studied so far, $S(l)$ increases logarithmically with $1/l$ as $l \to 0$, and we shall accordingly make the ansatz

$$S(l) \simeq S_0 - \sigma \ln l . \tag{1.4}$$

The constant σ will be called, following ref. 8, the information dimension. It is always a lower bound to the Hausdorff dimension, and in most cases they are almost the same within numerical errors.

The measure on which we shall concentrate mostly in this paper, has been recently introduced by the present authors [15]. It is obtained from the correlations between random points on the attractor. Consider the set $\{X_i, i = 1 \cdots N\}$ of points on the attractor, obtained e.g. from a time series, i.e. $X_i \equiv X(t + i\tau)$ with a fixed time increment τ between successive measurements. Due to the exponential divergence of trajectories, most pairs (X_i, X_j) with $i \neq j$ will be *dynamically* uncorrelated pairs of essentially random points. The points lie however on the attractor. Therefore they will be spatially correlated. We measure this spatial correlation with the correlation integral $C(l)$, defined according to

$$C(l) = \lim_{N\to\infty} \frac{1}{N^2} \times \{\text{number of pairs } (i,j) \text{ whose distance } |X_i - X_j| \text{ is less than } l\} . \tag{1.5}$$

The correlation integral is related to the standard correlation function

$$c(\boldsymbol{r}) = \lim_{N\to\infty} \frac{1}{N^2} \sum_{\substack{i,j=1 \\ i\neq j}}^{N} \delta^F(X_i - X_j - \boldsymbol{r}) \tag{1.6}$$

by

$$C(l) = \int_0^l \mathrm{d}^F r c(\boldsymbol{r}) . \tag{1.7}$$

One of the central aims of this paper is to establish that for small l's $C(l)$ grows like a power

$$C(l) \sim l^{\nu} , \tag{1.8}$$

and that this "correlation exponent" can be taken as a most useful measure of the local structure of a strange attractor. It seems that ν is more relevant, in this respect, than D. In any case, its calculation yields also an estimate of σ and D, since we shall argue that in general one has

$$\nu \leqslant \sigma \leqslant D . \tag{1.9}$$

We found that the inequalities are rather tight in most cases, but not in all. Given an experimental signal, if one finds eq. (1.8) with $\nu < F$, one knows that the signal stems from deterministic chaos rather than random noise, since random noise will

always result in $C(l) \sim l^F$. Explicit algorithms will be proposed below.

One of the main advantages of ν is that it can easily be measured, at least more easily than either σ or D. This is particularly true for cases where the fractal dimension is large ($\gtrsim 3$) and a covering by small cells becomes virtually impossible. We thus expect that the measure ν will be used in experimental situations, where typically high dimensional systems exist.

In theoretical cases, when the evolution law is known analytically, the easiest quantities to evaluate are the Lyapunov exponents. General formulae expressing D in terms of the λ_i have been proposed by Mori [9] and by Kaplan and Yorke [10]. If they were correct, they would obviously be very useful. They have been verified in simple cases [11, 14]. But Mori's formula was shown to be wrong in one case by Farmer [8], and the above example shown in fig. 1 shows that also the Kaplan–Yorke formula

$$D = D_{\mathrm{KY}} \equiv j + \frac{\lambda_1 + \lambda_2 + \cdots + \lambda_j}{|\lambda_{j+1}|} \tag{1.10}$$

does not hold even in all those cases where $\nu = \sigma = D$. Here, the exponents are ordered in descending order $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_F$, and j is the largest integer for which $\lambda_1 + \lambda_1 + \ldots + \lambda_j \geq 0$.

In section 7 we shall take up this question again. We shall show that the counterexample in fig. 1b is not generic. We shall however claim that eq. (1.10) cannot generally be expected to be correct, and that in fact D_{KY} is an upper bound, if $\nu = \sigma = D$.

In the next section, we shall present numerical results for several simple models, for which the fractal dimensions are known from the literature. This will serve to illustrate the scaling law (1.8), and to verify the inequality $\nu \leqslant D$. This inequality and its stronger version, eq. (1.9), will be derived in section 3. The case of one-dimensional maps at infinite bifurcation (Feigenbaum [16]) points is special in that there the information dimension σ and the exponent ν can be calculated exactly, with the result $\nu \neq \sigma \neq D$. It is treated in section 4. Section 5 is dedicated to an important modification which allows to extract ν from a time series of one single variable, instead of from the series $\{\boldsymbol{X}_i\}$. This is of course most important for infinite-dimensional systems, but it is also very useful in low-dimensional cases where it diminishes systematic errors. Among others, we shall apply this method in section 6 to the Mackey–Glass [17] delay equation studied in great detail in ref. 8.

In section 7 we discuss the relation of ν to the Lyapunov exponents, and establish the result

$$\nu \leqslant D_{\mathrm{KY}} . \tag{1.11}$$

A summary and a discussion of the actual method of treating experimental signals is offered in section 8.

2. Case studies of low-dimensional systems

In this section we shall establish that $C(l)$ can be very well represented by a power law l^ν, by exhibiting numerical results for a number of low dimensional systems. These results are summarized in table I. In section 5 we shall show that this is the case also in high (and infinite) dimensional systems. Details of the numerical algorithms are discussed in appendix A.

2.1. *One-dimensional maps*

The simplest cases of chaotic system are represented by maps of some interval into itself, as e.g. the logistic map [2]

$$x_{n+1} = a x_n (1 - x_n) . \tag{2.1}$$

We shall study this map both at the point of onset of chaos via period doubling bifurcations, i.e. when $a = a_\infty = 3.5699456\ldots$ and for the case $a = 4.0$. In fig. 2 we show the result for the first case. It is well known [2, 16] that for this map the attractor* is

* Note that the term "attractor" would not be universally accepted here due to the fact that in any neighbourhood there exist trajectories which do not tend towards it asymptotically.

Table I

	ν	No. of iterations, time increment τ	D	σ
Hénon map	1.21 ± 0.01[d])	15000	1.26 (ref. 11)	-
$a = 1.4$, $b = 0.3$	1.25 ± 0.02[e])			
Kaplan–Yorke map	1.42 ± 0.02	15000	1.431(ref. 11)	-
$\alpha = 0.2$				
Logistic eq.,	0.500 ± 0.005	25000	0.538(ref. 13)	0.5170976
$b = 3.5699456\cdots$	$0.4926 < \nu < 0.5024$[f])			
Lorenz eq.[a])	2.05 ± 0.01	15000; $\tau = 0.25$	2.06 ± 0.01	-
Rabinovich–[b])	2.19 ± 0.01	15000; $\tau = 0.25$	-	-
Fabrikant eq.				
Zaslavskii map[c])	(≈ 1.5)	25000	1.39(ref. 11)	-

[a])Parameters as in refs. 7 and 11.
[b])Parameters as in section 3 of ref. 20.
[c])Parameters as in ref. 11.
[d])From eqs. (1.5) and (1.8).
[e])From single variable time series, with $f = 3$.
[f])Exact analytic bound.

Cantor-like with a fractal dimension satisfying the exact bound [13] $0.5376 < D < 0.5386$. In section 4 we shall prove exactly that $\sigma = 0.517097\ldots$, and that $0.4926 < \nu < 0.5024$ while from Fig. 2 we find $\nu = 0.500 \pm 0.005$. For very small distances, the data for $C(l)$ deviate from a power law, but that was to be expected: the behaviour at $a = a_\infty$ is not yet chaotic, and therefore the values x_n are strongly correlated. We verified that indeed the powerlaw holds down to smaller values of l if we increase N or use only values x_i, x_{i+p}, x_{i+2p}, x_{i+2p}, ... with p being a large odd number.

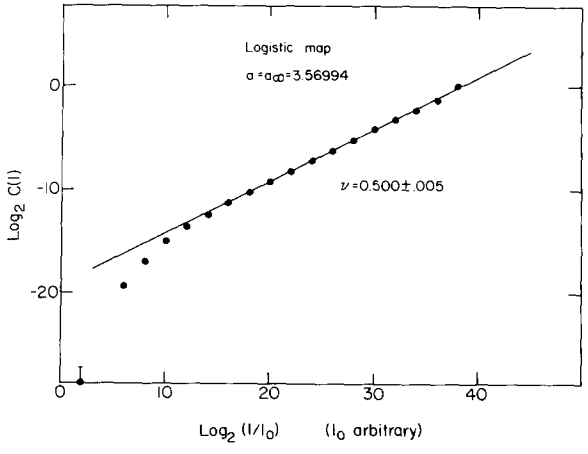

Fig. 2. Correlation integral for the logistic map (2.1) at the infinite bifurcation point $a = a_\infty = 3.699\ldots$ The starting point was $x_0 = \frac{1}{2}$, the number of points was $N = 30.000$.

The same map can be used also to introduce the important issue of corrections to scaling. These are found for the parameter value $a = 4$. It is well known that in this the attractor* consists of the interval [0, 1], and that the invariant probability density is equal to

$$p(x) \equiv \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \delta(x_i - x) \tag{2.2}$$

$$= \frac{1}{\pi} [x(1-x)]^{-1/2} . \tag{2.3}$$

From this, one finds easily

$$\nu = \sigma = D = 1 . \tag{2.4}$$

Notice, however, that while the scaling laws (1.2) and (1.4) are exact, the scaling law (1.8) for $C(l)$

* Again, the term is questionable, as no point outside the interval [0, 1] gets attracted towards it. We shall ignore this irrelevant point, which could be avoided by using $a = 4 - \epsilon$.

requires logarithmic corrections, due to the singular behaviour of $p(x)$:

$$C(l) = \int_0^1 \int dx\, dy p(x) p(y) \theta(|x-y| - l)$$
$$\underset{l \to 0}{\simeq} \frac{4}{\pi^2} l \ln 1/l . \tag{2.5}$$

Thus, a numerical calculation of ν is expected to converge very slowly. This problem and a remedy for it are discussed further in section 5.

2.2. *Maps of the plane*

Here we examined the Hénon [18] map

$$x_{n+1} = y_n + 1 - a x_n^2 ,$$
$$y_{n+1} = b x_n , \tag{2.6}$$

with $a = 1.4$ and $b = 0.3$, the Kaplan–Yorke [10] map

$$x_{n+1} = 2x_n \pmod 1 ,$$
$$y_{n+1} = \alpha y_n + \cos 4\pi x_n \tag{2.7}$$

with $\alpha = 0.2$, and the Zaslavskii [19] map

$$x_{n+1} = [x_n + \nu(1 + \mu y_n) + \epsilon \nu \mu \cos 2\pi x_n] \pmod 1,$$
$$y_{n+1} = \mathrm{e}^{-\Gamma}(y_n + \epsilon \cos 2\pi x_n) , \tag{2.8}$$

with the parameters

$$\mu = \frac{1 - \mathrm{e}^{-\Gamma}}{\Gamma} \tag{2.9}$$

and $\Gamma = 3.0$, $\nu = 400/3$, and $\epsilon = 0.3$ taken from ref. 11.

Figs. 3–5 exhibit the results for the correlation integrals. In the first two cases, we find excellent agreement with a power law; while for the Kaplan–Yorke map we find $\nu = 1.42 \pm .02$ in agreement with the published [11] value of D, a fit to the Hénon map yields $\nu_{\mathrm{eff}} = 1.21$, smaller than

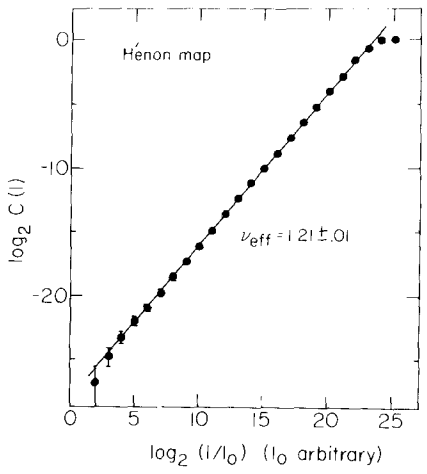

Fig. 3. Correlation integral for the Hénon map (2.6) with $a = 1.4$, $b = 0.03$ and $N = 15.000$.

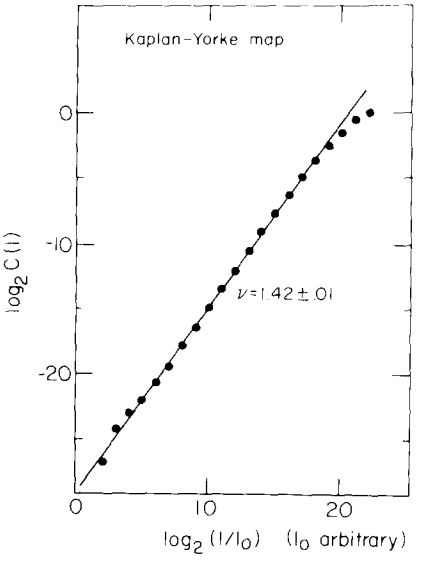

Fig. 4. Same as fig. 3, but for Kaplan–Yorke map (2.7) with $\alpha = 0.2$.

the value [11] $D = 1.261 \pm 0.003$. We shall argue in sectin 5 that actually the value of ν for the Hénon map is underestimated here, and that instead $\nu = 1.25 \pm 0.02 \approx D$.

The case of the Zaslavskii map is exceptional as it was the only system for which we did not find

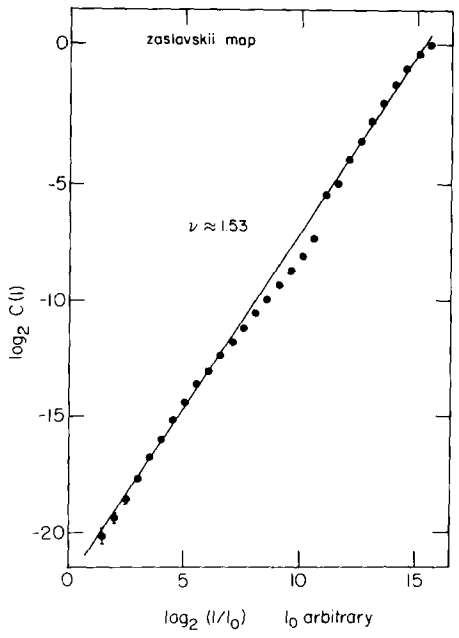

Fig. 5. Correlation integral for Zaslavskii map (eqs. (2.8), (2.9)); $N = 25.000$, parameters as in the text. For faster scaling, the y-coordinate was blown up by a factor of 25, rendering the attractor square-like at low resolution (see fig. 6; without this, the attractor would have looked effectively 1-dimensional for $l \gtrsim l_{max}/25$).

clear-cut power behaviour. Also, an (admittedly poor) fit would yield $\nu \approx 1.5$, in clear violation of the bound $\nu < D$. The reasons why our method has to fail for this map – with the parameters as quoted above – becomes clear when looking at fig. 6. Call l_0 the outer length scale. From fig. 6a one sees that the attractor looks 2-dimensional for $l \gtrsim l_0 \times 2^{-5}$ and $\approx$ 1-dimensional for $l_0 \times 2^{-5} \gtrsim l \gtrsim l_0 \times 2^{-9}$. From fig. 6b one sees that it looks $\approx$ 2-dimensional again down to $\approx l_0 \times 2^{-14}$, scaling behaviour setting in only at about that scale (which is beyond our resolution). It seems to us that the box-counting algorithm of ref. 11 in which D is evaluated, should confront the same problem†.

† Note added: Dr. Russel kindly provided us with the original data of $M(\epsilon)$ versus ϵ. From these, it seems that indeed a similar phenomenon occurs and that accordingly a value $D \approx 1.5$ cannot be excluded.

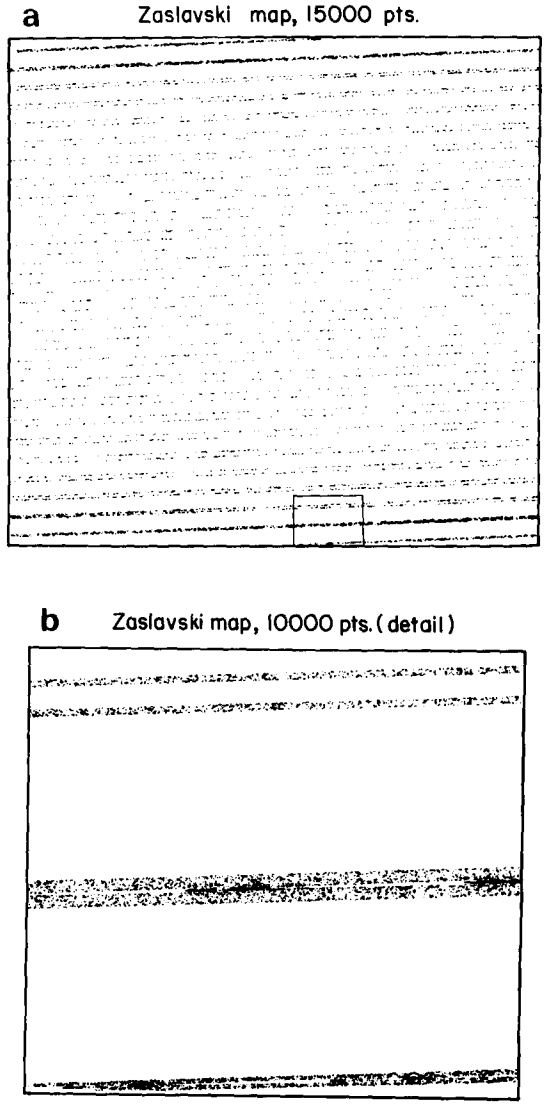

Fig. 6. Attractor of the Zaslavskii map. a) entire attractor (15.000 points plotted; y-scale blown up by factor 25); b) Blown up view of part indicated in part a (10.000 points plotted).

2.3. *Differential equations*

We have studied the Lorenz [1] model

$$\begin{aligned} \dot{x} &= \sigma(y - x), \\ \dot{y} &= -y - xz + Rx, \\ \dot{z} &= xy - bz, \end{aligned} \tag{2.10}$$

with $R = 28$, $\sigma = 10$, and $b = 8/3$, and the Rabinovich–Fabrikant [20] equations

$$\begin{aligned} \dot{x} &= y(z - 1 + x^2) + \gamma x, \\ \dot{y} &= x(3z + 1 - x^2) + \gamma y, \\ \dot{z} &= -2z(\alpha + xy), \end{aligned} \tag{2.11}$$

with $\gamma = 0.87$ and $\alpha = 1.1$.

As seen in fig. 7 we get adequate power laws for $C(l)$, and in the case of the Lorenz model, where D is known [11], we obtain $\nu \simeq D$.

Further examples will be studied in section 6, in the context of higher dimensional systems.

It should be stressed that the algorithm used to calculate ν converged quite rapidly. Although each entry in table I and figs. 2–7 were based on ≈ 15.000–25.000 points each, reasonable results (i.e. results for ν within $\pm 5\%$) were obtained in most cases already with only a few thousand points. This should be contrasted with the difficulties associated with estimating D in box-counting algorithms [11, 14].

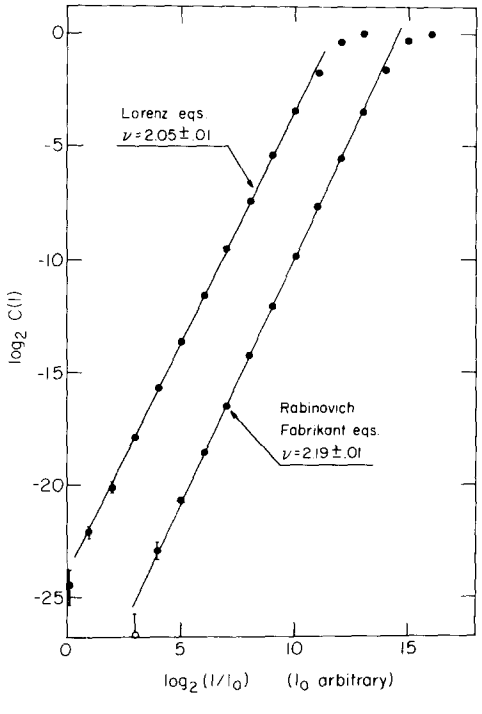

Fig. 7. Correlation integrals for the Lorenz equations (eq. (2.10); dots) and for the Rabinovich–Fabrikant equation (eq. (2.11); open circles). In both cases, $N = 15.000$ and $\tau = 0.25$.

Summarizing this section, we can say that except for the logistic map at $a = a_\infty$ ("Feigenbaum attractor") we found in all cases that $\nu \approx D$ within the limits of accuracy. We now turn to a theoretical analysis of the relations between ν, σ and D.

3. Relations between ν, σ and D

In this section we shall establish the inequalities (1.9). We shall do this in 3 steps.

a) The easiest inequality to prove is $\sigma \leqslant D$. Consider a covering of the attractor by hypercubes ("cells") of edge length l, and a time series $\{\boldsymbol{X}_k; k = 1, \ldots, N\}$. The probabilities p_i for an arbitrary $\boldsymbol{X}_k$ to fall into cell i are simply

$$p_i = \lim_{N\to\infty} \frac{1}{N} \mu_i . \tag{3.1}$$

where μ_i is the number of points $\boldsymbol{X}_k$ which fall into cell i.

If the coverage of the attractor is uniform, one has,

$$p_i = \frac{1}{M(l)}, \tag{3.2}$$

where $M(l)$ is the number of cells needed to cover the attractor, and one finds from eqs. (1.3) and (1.2)

$$S(l) = S^{(0)}(l) = \ln M(l) = \text{const} - D \ln l . \tag{3.3}$$

In the general case, one uses the convexity of $x \ln x$ in the usual way to prove that $S(l) \leqslant S^{(0)}(l)$. Invoking the ansatz $S(l) = \text{const} - \sigma \ln l$, we find $\sigma \leqslant D$.

b) Instead of showing immediately $\nu \leqslant \sigma$, let us proceed slowly and show first that $\nu \leqslant D$.

From the definition of $C(l)$, we get up to a factor

of order unity

$$C(l) \simeq \lim_{N\to\infty} \frac{1}{N^2} \sum_{i=1}^{M(l)} \mu_i^2 = \sum_{i=1}^{M(l)} p_i^2 . \tag{3.4}$$

Here, we have replaced the number of pairs with distance $< l$ by the number of pairs which fall into the same cell of length l. The error committed should be independent on l, and thus should not affect the estimation of ν. Using the Schwartz inequality we get

$$C(l) = M(l)\langle p_i^2 \rangle \geqslant M(l)\langle p_i \rangle^2 = \frac{1}{M(l)} \sim l^D . \tag{3.5}$$

In this equation square brackets denote average over all cells. Comparing eqs. (3.5) and (1.8) we find immediately $\nu \leqslant D$.

c) In order to derive $\nu \leqslant \sigma$, consider two nested coverings with cubes of lengths l and $2l$. The numbers of cubes that contain a piece of the attractor are then related by

$$M(l) = 2^D M(2l) . \tag{3.6}$$

Denote by p_i the probability to fall in cube i of the finer coverage, and by P_j the probability to fall in cube j of the coarser. Define $\omega_i (i = 1, \ldots M(l))$ by

$$p_i = \omega_i P_j \quad (i \in j) . \tag{3.7}$$

Evidently we have

$$P_j = \sum_{i\in j} p_i , \quad \sum_{i\in j} \omega_i = 1 . \tag{3.8}$$

We can then write the correlation integral as

$$C(l) \simeq \sum_{i=1}^{M(l)} p_i^2 = \sum_{j=1}^{M(2l)} P_j^2 \sum_{i\in j} \omega_i^2 . \tag{3.9}$$

Consider now the ratio

$$\frac{C(l)}{C(2l)} = \frac{\sum_j P_j^2 \sum_{i\in j} \omega_i^2}{\sum_j P_j^2} , \tag{3.10}$$

and compare it to the entropy difference

$$S(2l) - S(l) = \sum_{i=1}^{M(l)} p_i \ln p_i - \sum_{j=1}^{M(2l)} P_j \ln P_j = \sum_{j=1}^{M(2l)} P_j \sum_{i\in j} \omega_i \ln \omega_i . \tag{3.11}$$

In order to estimate eq. (3.10) in terms of eq. (3.11), we have to introduce a new assumption. We assume that the ω_i's are distributed independently of the P_j. This means essentially that locally the attractor looks the same in regions where it is rather dense (P_j large) as in regions where P_j is small. Although we cannot further justify this assumption, it seems to us very natural. It leads immediately to

$$\frac{C(l)}{C(2l)} = \frac{\langle \omega^2 \rangle}{\langle \omega \rangle} = 2^D \langle \omega^2 \rangle , \tag{3.12}$$

and to

$$S(2l) - S(l) = 2^D \langle \omega \ln \omega \rangle . \tag{3.13}$$

Define now a normalized variable W by

$$W = \frac{\omega}{\langle \omega \rangle} = 2^D \omega . \tag{3.14}$$

Using the inequality [21]

$$\langle W^2 \rangle > \exp \langle W \ln W \rangle , \tag{3.15}$$

we establish

$$\frac{C(l)}{C(2l)} \geqslant \exp[S(2l) - S(l)] \tag{3.16}$$

and thus

$$\nu \leqslant \sigma . \tag{3.17}$$

Remarks. From the proofs it is clear that if the attractor is uniformly covered, one has equalities

$$\nu = \sigma = D . \tag{3.18}$$

It is an interesting question how non-uniform the coverage must be in order to break them. With the exception of the Feigenbaum map (logistic map with $a = a_\infty$), which is however not generic, all examples of the last section were compatible with eq. (3.18).

In cases where $\nu \neq D$, we claim that indeed ν is the more relevant observable. In these cases, the neighbourhoods of certain points have higher "seniority" in the sense that they are visited more often than others. The fractal dimension is ignorant of seniority, being a purely geometric concept. But both the correlation integral and the entropy dimension weight regions according to their seniority.

Eqs. (1.9) and (3.18) have been used previously in the context of fully developed homogeneous turbulence [22]. The connection

$$c(l) \propto l^{D-F}, \quad l \in R^f$$

following from $\nu = D$ has been used previously also in percolation theory [23] and in a model for dendritic growth [24].

4. Information entropy and ν of the Feigenbaum attractor

In this section we shall compute exactly the information dimension and ν of one-dimensional maps

$$x_{n+1} = F(x_n) \tag{4.1}$$

at the onset of chaos. The method follows closely the one of ref. 13.

It is well known that such maps – provided they have a unique quadratic maximum – have universal scaling features, studied in most detail by Feigenbaum [16]. This behaviour is most easily described by observing that the iterations

$$F^{(2^n)}(x) = \underbrace{F(F(\ldots F(x)\ldots))}_{2^n \text{ times}} \tag{4.2}$$

tend after a suitable rescaling towards a universal function

$$g(x) = \lim_{n\to\infty} \frac{1}{F^{(2^n)}(0)} F^{(2^n)}(xF^{(2^n)}(0)) . \tag{4.3}$$

This "Feigenbaum function" $g(x)$ satisfies the exact scaling relation

$$g(g(x)) = -\frac{1}{\alpha} g(\alpha x) , \tag{4.4}$$

with $\alpha = 2.50290\ldots$, and the normalization condition $g(0) = 1$. We have here assumed that the maximum of $F(x)$ is at $x = 0$, which can always be achieved by a change of variables. In order to obtain the information dimension of the logistic map at $a = a_\infty = 3.5699345\ldots$, it is thus sufficient to compute σ for the Feigenbaum map.

The "attractor" (see the reservations in section 2) of $g(x)$ consists of the sequence $\{\xi_n, n = 0, 1, 2, \ldots\}$ with

$$\xi_0 = 0 \tag{4.5}$$

and

$$\xi_{n+1} = g(\xi_n) . \tag{4.6}$$

The first few ξ_k's are shown in fig. 8. There, it is also indicated how they build up the Cantorian structure of the attractor: the points ξ_1, ξ_2, $\xi_3, \ldots, \xi_{2^k+1}$ form the end-points of 2^k intervals, and the following ξ_k's fall all into these intervals. Furthermore, any sequence $\{\xi_n, \xi_{n+1} \ldots \xi_{n+2^k-1}\}$ of 2^k successive points visits each of these intervals exactly once. Thus, the a priori probabilities $p_i (i = 1, \ldots, 2^k)$ for an arbitrary x_n to fall into the ith interval are all equal to $p_i = 2^{-k}$.

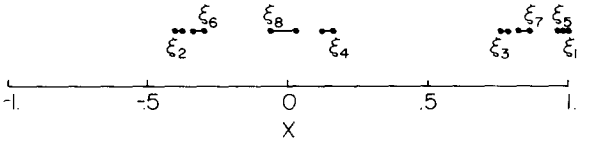

Fig. 8. First 16 points $\xi_1, \ldots, \xi_{16}$ of the attractor of the Feigenbaum equation describing the onset of chaos in 1-dimensional systems.

By the grouping axiom, we can first write the information entropy as

$$S(l) = \tfrac{1}{2}[S_{[2,4]}(l) + S_{[3,1]}(l)] + \ln 2\,, \tag{4.7}$$

where we denote by $S_{[i,j]}$ the information needed to specify the point on the interval $[\xi_i, \xi_j]$, and where we have used the fact that an arbitrary x_n has equal probability to be on $[\xi_2, \xi_4]$ or on $[\xi_3, \xi_1]$. From eq. (4.4) we find, however, that

$$\xi_{2n} = -\frac{1}{\alpha}\xi_n\,. \tag{4.8}$$

Thus, the interval $[\xi_2, \xi_4]$ is a down-scaled image of the whole attractor, and we have

$$S_{[2,4]}(l) = S(\alpha l) \approx S(l) - \sigma \ln \alpha\,, \tag{4.9}$$

where we have used the scaling ansatz (1.4).

In order to estimate $S_{[3,1]}(l)$, we decompose the interval [3, 1] into the 2^{k-1} subintervals discussed above, defined by the ξ_n with odd n's:

$$S_{[3,1]}(l) = (k-1)\ln 2 + 2^{-k+1}\sum_{i=1}^{2^{k-1}} S_i(l)\,.$$

Again, we have applied the grouping axiom, using that $p_i = 2^{-k}$. The $S_i(l)$ are the informations needed to pin down x_n provided one knows that it falls into the ith subinterval. Since each subinterval maps onto one on the left-hand piece $[\xi_2, \xi_4]$, each $S_i(l)$ is equal to the information $\tilde{S}_i(|g_i'|l)$ needed to pin x_{n+1} on the corresponding interval on the left-hand side. Here, g_i' is some average derivative of $g(x)$ in the ith subinterval. Using that $\tilde{S}_i(|g_i'|l) \simeq \tilde{S}_i(l) - \sigma \ln|g_i'|$, we obtain

$$\begin{aligned} S_{[3,1]}(l) &= (k-1)\ln 2 + 2^{-k+1}\sum_{i=1}^{2^{k-1}} \tilde{S}_i(l) \\ &\quad - \sigma \sum_{i=1}^{2^{k-1}} \ln|g_i'| \\ &= S_{[2,4]}(l) - \sigma \sum_{i=1}^{2^{k-1}} \ln|g_i'|\,. \end{aligned} \tag{4.10}$$

Inserting this and eq. (4.9) into eq. (4.7), we find after a few manipulations and after taking the limit $k \to \infty$

$$\sigma = \lim_{k\to\infty} \frac{\ln 2}{\ln\alpha + \dfrac{1}{2^{k+1}}\displaystyle\sum_{i=1}^{2^k} |g'(\xi_{2i-1})|}\,. \tag{4.11}$$

The limit converges very quickly, leading (for $k > 7$) to

$$\sigma = 0.5170976\,. \tag{4.12}$$

The calculation of the correlation exponent, or rather of the exponent of the Renyi entropy (see eq. (3.4))

$$R(l) = \sum_{i=1}^{M(l)} p_i^2 \tag{4.13}$$

follows even more closely the one in ref. 13.

As in that paper, we obtain a nested set of bounds. The first (and least stringent) is obtained by writing

$$R(l) = \tfrac{1}{4}\{R_{[2,4]}(l) + R_{[3,1]}(l)\} \tag{4.14}$$

and using $R_{[2,4]}(l) = R(\alpha.l)$ and $R_{[3,1]}(l) = R(\alpha g' l)$ with

$$|g'(\xi_3)| < g' < |g'(\xi_1)|\,. \tag{4.15}$$

Assuming $R(l) \sim l^\nu$, we obtain

$$1 + |g'(\xi_3)|^\nu < \frac{4}{\alpha^\nu} < 1 + |g'(\xi_1)|^\nu\,, \tag{4.16}$$

leading to $0.4857 < \nu < 0.5235$.

For the next more stringent bounds, we write further

$$R_{[3,1]}(l) = \tfrac{1}{4}\{R_{[3,7]}(l) + R_{[5,1]}(l)\}\,, \tag{4.17}$$

with

$$R_{[3,7]}(l) + R(\alpha^2 g^{(1)}.l), \quad |g'(\xi_3)| < g^{(1)} < |g'(\xi_7)| \tag{4.18}$$

and

$$R_{[5,1]}(l) + R_{[3,1]}(\alpha g^{(2)} l), \quad |g'(\xi_5)| < g^{(2)} < |g'(\xi_1)| . \tag{4.19}$$

Some algebra leads then to

$$|g'(\xi_5)|^\nu + \frac{\alpha^\nu}{4-\alpha^\nu}|g'(\xi_3)|^\nu < \frac{4}{\alpha^\nu} < |g'(\xi_1)|^\nu + \frac{\alpha^\nu}{4-\alpha^\nu}|g'(\xi_7)|^\nu , \tag{4.20}$$

with the result

$$0.4926 < \nu < 0.5024 , \tag{4.21}$$

in agreement with the numerical value $\nu = 0.500 \pm 0.005$.

5. Using a single-variable time series

Very often one does not have access to a time series $\{X_n\}$ of F-dimensional vectors. Instead one follows only one or at most a few components of X_n. This is particularly relevant for real (as opposed to computer) experiments where the number of degrees of freedom often is very high if not infinite. Such systems nevertheless can have low-dimensional attractors. It would be very desirable to have a reliable method which allows a characterization of this attractor from a single-variable time series. $\{x_i,\ i = 1, \ldots, N;\ x_i \in R\}$.

The essential idea [25, 26] consists in constructing d-dimensional vectors

$$\xi_i = (x_i, x_{i+1}, \ldots, x_{i+d-1}) \tag{5.1}$$

and using ξ-space instead of X-space. The correlation integral would e.g. be

$$C(l) = \lim_{N\to\infty} \frac{1}{N^2} \sum_{i,j=1}^{N} \theta(l - |\xi_i - \xi_j|) . \tag{5.2}$$

More generally, one can use

$$\xi_i = (x(t_i), x(t_i + \tau) \ldots x(t_i + (d-1)\tau)) , \tag{5.3}$$

with τ some fixed interval. The magnitude of τ should not be chosen too small since otherwise $x_i \approx x_{i+\tau} \approx x_{i+2\tau} \approx \cdots$ so that the attractor in ξ-space would be stretched along the diagonal and thus difficult to disentangle. On the other hand, τ should not be chosen too large since distant values in the time series are not strongly correlated (due to the exponential divergence of trajectories and unavoidable small errors).

A similar compromise must be chosen for the dimension d. Clearly, d must be larger than the Hausdorff dimension D of the attractor (otherwise, $C(l) \sim l^d$). If the attractor is Cantorian in more than one dimension, this might however not be sufficient. Also, it might be that, when looked at in d dimension, the density

$$p(\xi) = \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \delta(\xi_i - \xi) \tag{5.4}$$

develops singularities which are absent in more than d dimensions (such singularities occur e.g. when one projects a sphere with constant density, $p(\xi) = p\delta(x^2 + y^2 + z^2 - R^2)$, onto the x–y plane: the new density $\tilde{p}(x, y)$ is infinite at $x^2 + y^2 = R^2$).

On the other hand, one cannot make d too large without getting lost in experimental errors and lack of statistics.

In the next section, we shall study an infinite-dimensional system from this point of view. In the remainder of the present section, we shall apply these considerations to the logistic map with $a = 4$, and to the Hénon map.

In the logistic map, we have seen that there are logarithmic corrections to the power law $C(l) \sim l^\nu$. They result precisely from singularities of $p(x)$, at $x = 0$ and $x = 1$. While embedding the attractor in a higher dimensional space does not completely remove these singularities, it substantially reduces their influence. The reason is that embedding in higher dimensional space always results in stretching the attractor. However, the portions which are most strongly stretched are those which are most densely populated at the lower dimension. For example in the logistic map with $a = 4$ the "attrac-

tor" is the interval [0, 1] in $1d$ but is the parabola in $2d$. The parabola has highest slopes at the end points, exhibiting the stronger stretching associated with regions of singular distributions at a lower dimension. A similar effect appears when going from $d = 2$ to $d = 3$. We thus expect that the importance of the singularities in the distribution would be reduced in higher dimensions.

In order to check this, we have calculated for the logistic map at $a = 4$ the original correlation integral and the modified integral obtained by embedding in a 2- and 3-dimensional space. The results are shown in fig. 9. We observe indeed the expected decrease of systematic error when increasing d, accompanied by an increase of the statistical error.

Analogous results for the Hénon map are shown in fig. 10. There, we used as time series the series $\{x_n, x_{n+2}, x_{n+4}, \ldots\}$. While the 2-dimensional correlation integral gives an effective ν in agreement with the result of section 2, the 3-dimensional embedding gives a larger values $\nu = 1.25 \pm 0.02$ which agrees with the value of D found in refs. 11 and 14.

No such effects were observed in the Lorenz model, where both the originally defined $C(l)$ and

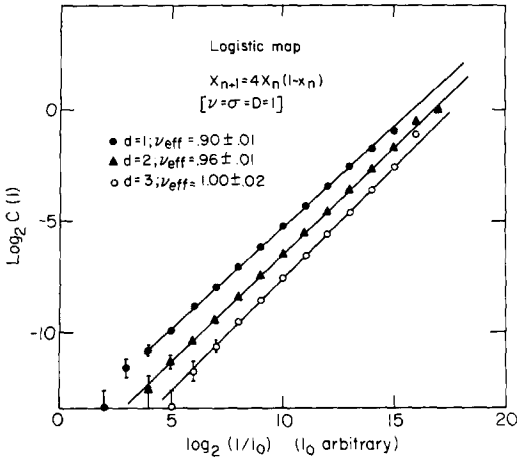

Fig. 9. Modified correlation integrals for the logistic map (2.1) with $a = 4$. The distance l between 2 points ξ_n and ξ_m on the attractor is defined as $l^2 = (\xi_n - \xi_m)^2 = (x_n - x_m)^2 + \ldots + (x_{n+d-1} - x_{m+d-1})^2$. For each value of d, we took $N = 15.000$.

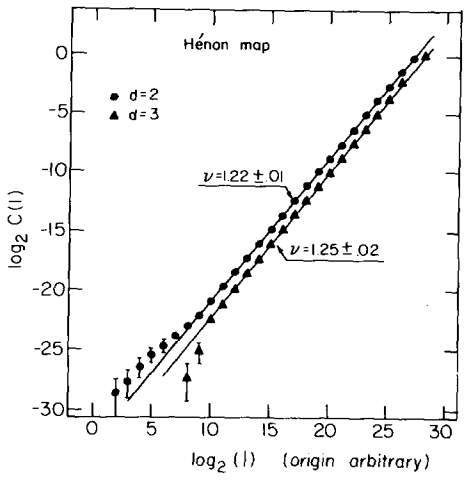

Fig. 10. Modified correlation integrals for the Hénon map (2.6). The time series consisted of coordinates x_n, x_{n+2}, $x_{n+4}, \ldots$, and $\xi = (x_n, x_{n+2}, \ldots, x_{n+2(d-1)})$ for each d. For $d = 2$, we took $N = 30.000$; for $d = 3$, we took $N = 20.000$.

the modified correlation integral using only a single coordinate time series gave values of ν which agreed with D [15].

The conclusion drawn from these examples is that it is often useful to represent the attractor in a higher dimensional space than absolutely necessary, in order to reduce systematic errors. These errors result from a strongly non-uniform coverage of the attractor, provided this non-uniformity is not so strong as to make $\nu \neq D$.

6. Infinite-dimensional systems: an example

An extremely convenient way of generating very high dimensional systems is to consider delay differential equations of the type

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t} = F(x(t), x(t-\tau)), \tag{6.1}$$

where τ is a given time delay. Such a delay equation is in fact infinite dimensional, as is most easily seen from the initial conditions necessary to solve eq. (6.1): they consist of the function $x(t)$ over a whole interval of length τ.

Following ref. 8, we shall study a particular example, introduced by Mackey and Glass [17] as a model for regeneration of blood cells in patients with leukemia. It is

$$\dot{x}(t) = \frac{ax(t-\tau)}{1+[x(t-\tau)]^{10}} - bx(t)\,. \tag{6.2}$$

As in ref. 8, we shall keep $a = 0.2$ and $b = 0.1$ fixed, and study the dependence on the delay time τ.

For the numerical investigation, eq. (6.2) is turned into an n-dimensional set of difference equations, with $n = 600$–1200. Details are described in the appendix. The time series was always chosen as $\{x(t),\ x(t+\tau),\ x(t+2\tau), \ldots\}$ except for some runs with $\tau = 100$, where we took points at times $t,\ t+\tau/2,\ t+2\tau/2, \ldots$.

The results for the correlation integral are shown in figs. 11–14. Estimated values of ν are given in table II, together with values of D obtained in ref. 8 by applying the defining eq. (1.2) to a Poincaré return map. Also shown in table II are the

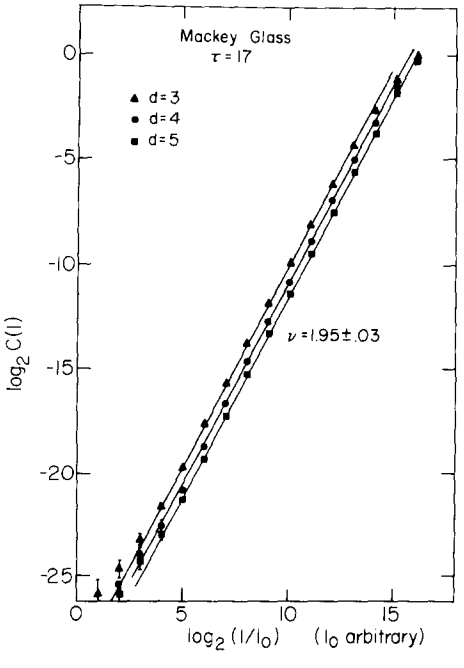

Fig. 11. Modified correlation integrals for the Mackey–Glass delay equation (6.2), with delay $\tau = 17$. The time series consisted of $\{X(t+i\tau);\ i = 1, \ldots, 25.000\}$.

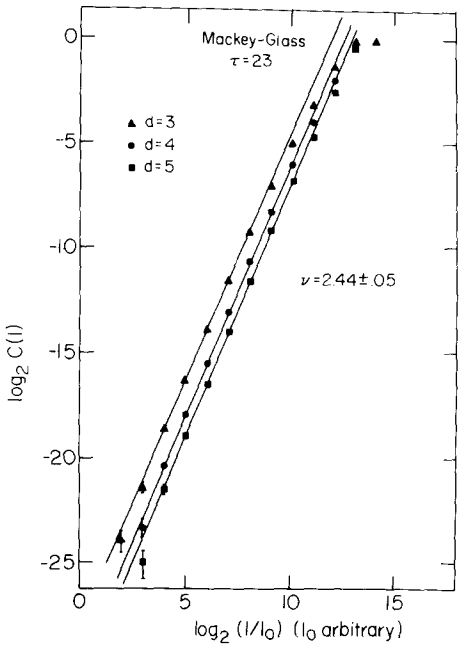

Fig. 12. Same as fig. 11, but for $\tau = 23$.

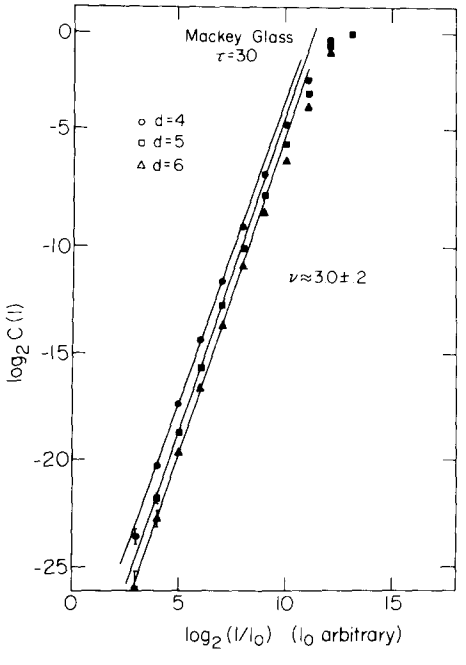

Fig. 13. Same as fig. 11, but for $\tau = 30$.

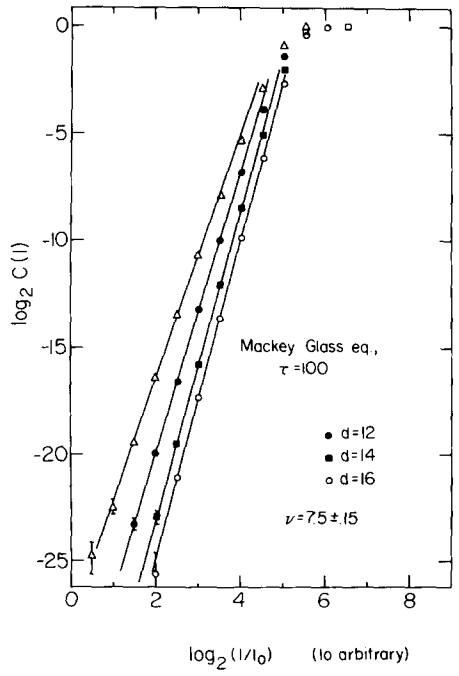

Fig. 14. Same as fig. 11, but for $\tau = 100$. For $d = 16$, the time series consisted of points $\{X(t + i\tau/2); i = 1, \ldots, 25.000\}$.

Kaplan–Yorke dimension D_{KY} (see eq. (1.10)) which will in the next section be shown to be an upper bound to ν, and the number of positive Lyapunov exponents, both taken from ref. 8. It is obvious that this latter number, called D_{LB}, is a lower bound to D. If the density of trajectories on the attractor is not too non-uniform, we expect that D_{LB} yields also a lower bound to ν.

From table II we see that indeed in all cases

$$D_{LB} \leqslant \nu \leqslant D \leqslant D_{KY}, \tag{6.3}$$

except for $\tau = 17$ where ν is slightly less than D_{LB}. However, for those small values of τ for which box-counting according to the definition of D had been feasible, our values of ν are considerably smaller than the values of D found in ref. 8, while the values of D were fairly close to D_{KY}.

In all cases, the linearity of the plot of $\log C(l)$ versus $\log l$ improved substantially when increasing d above its minimal required value. For increasing values of d, the effective exponent at first also

Table II
Estimates of the correlation exponent ν for the Mackey–Glass equation (6.2) with $a = 0.2$, $b = 0.1$. Values for D_{LB}, D and D_{KY} are from ref. 8. For $\tau = 100$ the value of ν saturated at $d = 16$

τ	D_{LB}	ν	D	D_{KY}
17.0	2	1.95 ± 0.03 ($d = 3$) 1.35 ± 0.03 ($d = 4$) 1.95 ± 0.03 ($d = 5$)	2.13 ± 0.03	2.10 ± 0.02
23.0	2	2.38 ± 0.15 ($d = 3$) 2.43 ± 0.05 ($d = 4$) 2.44 ± 0.05 ($d = 5$) 2.42 ± 0.1 ($d = 6$)	2.76 ± 0.06	2.92 ± 0.03
30.0	3	2.87 ± 0.3 ($d = 4$) 3.0 ± 0.2 ($d = 5$) 3.0 ± 0.2 ($d = 6$) 2.8 ± 0.3 ($d = 7$)	> 2.94	3.58 ± 0.04
100.0	6	5.8 ± 0.3 ($d = 10$) 6.6 ± 0.2 ($d = 12$) 7.2 ± 0.2 ($d = 14$) 7.5 ± 0.15 ($d = 16$)	-	≈ 10.0

increases, but settles at a value which we assume to be the true value of ν. We must stress that we have no *proof* that the values of ν obtained with the highest chosen d represent the "true" exponent. We feel however that they surely represent reasonable estimates even for attractors with dimensions as high as ≈ 7.

In real experiments, where Lyapunov exponents are not available and thus D_{LB} and D_{KY} not easily obtained, our method seems the only one which could distinguish such an attractor from a system where the stochasticity is due to random noise. In that case, one would expect $C(l) \sim l^d$ as the trajectory is space-filling, in clear distinction from what we observe.

7. Relation to Lyapunov exponents and the Kaplan–Yorke conjecture

As we already mentioned in the introduction, the Lyapunov exponents are related to the evolution of the shape of an infinitesimal F-dimensional ball in phase space: being infinitesimal, it depends only on the linearized part of the flow, and thus becomes an ellipsoid with exponentially shrinking or growing axes. Denoting the principal axes by $\epsilon_i(t)$, the Lyapunov exponents are given by

$$\lambda_i = \lim_{t\to\infty} \lim_{\epsilon_i(0)\to 0} \frac{1}{t} \ln \frac{\epsilon_i(t)}{\epsilon_i(0)}. \tag{7.1}$$

Directions associated with positive Lyapunov exponents are called "unstable", those associated with negative exponents are called "stable".

Originally [10], Kaplan and Yorke had conjectured that D_{KY} is equal to D. In a recent preprint [27], they claim that D_{KY} is generically equal to a "probabilistic dimension", which seems to be the same as σ.

This latter claim has been partially supported in ref. 28, where essentially D_{KY} is proven to be an upper bound to the probabilistic dimension.

As shown by the counter example mentioned in the introduction, there are (possibly exceptional) cases where this bound is not saturated. In this section, we shall elucidate this question by giving a heuristic proof for the inequality $\nu \leqslant D$. From this, we see necessary conditions for the Kaplan–Yorke conjecture to hold, and which do not seem to be met generally.

Consider two infinitesimally close-by trajectories $\boldsymbol{X}(t)$ and $\boldsymbol{X}'(t) = \boldsymbol{X}(t) + \boldsymbol{\Delta}(t)$, where the latter could indeed be $\boldsymbol{X}'(t) = \boldsymbol{X}(t+T)$, which for sufficiently large T is essentially independent of $\boldsymbol{X}(t)$. We assume that $\Delta_i(t)$ increase exponentially, without any fluctuations, as

$$\Delta_i(t) = \Delta_i(0)\, \mathrm{e}^{\lambda_i t}, \tag{7.2}$$

where the components are along the principal axes discussed above. This is of course a strong assumption which would imply, in particular, that $\nu = \sigma = D$. Corrections to it will be treated in a forthcoming paper, but our main conclusion will remain unchanged. Conservation of the number of trajectories implies that the correlation function increases like

$$c(\boldsymbol{\Delta}(t)) = \left|\frac{\partial(\boldsymbol{\Delta}(0))}{\partial(\boldsymbol{\Delta}(t))}\right| c(\boldsymbol{\Delta}(0)) = \mathrm{e}^{-t\sum_{i=1}^{F}\lambda_i}\, c(\boldsymbol{\Delta}(0)). \tag{7.3}$$

To proceed further, we need a scaling assumption which generalizes the scaling ansatz

$$c(|\boldsymbol{\Delta}|) \sim |\boldsymbol{\Delta}|^{\nu - F}. \tag{7.4}$$

Observing that the attractor is locally a topological product of an R^n with Cantor sets, and that the relevant axes are the principal axes, we associate with each axis an exponent ν_i, $0 < \nu_i < 1$, and make the ansatz

$$c(\boldsymbol{\Delta}) \approx \prod_{i=1}^{F} c_i(\Delta_i), \tag{7.5}$$

with

$$c_i(x) \propto \begin{cases} x^{\nu_i - 1}, & \text{if} \quad 0 < \nu_i \leqslant 1, \\ \delta(x), & \text{if} \quad \nu_i = 0. \end{cases} \tag{7.6}$$

If $\nu_i = 0$, this means that the motion along this axis dies asymptotically (example: directions normal to a limit cycle). Directions with $\nu_i = 1$ are the unstable directions, with the continuous density. Directions with $0 < \nu_i < 1$, finally, are either Cantorian or, in exceptional cases, directions along which the distribution is continuous but singular at $\Delta_i = 0$. Notice that $\nu_i > 1$ is impossible.

Substituting eq. (7.5) into (7.3), we find

$$\prod_i (\Delta_i(0)^{\nu_i - 1} e^{t\lambda_i(\nu_i - 1)}) = e^{-t\Sigma_i \lambda_i} \prod_i \Delta_i(0)^{\nu_i - 1}, \quad (7.7)$$

or

$$\sum_{i=1}^{F} \lambda_i \nu_i = 0. \quad (7.8)$$

In addition we have, from eqs. (7.5) and (7.4),

$$\sum_{i=1}^{F} \nu_i = \nu, \quad (7.9)$$

and

$$0 \leqslant \nu_i \leqslant 1. \quad (7.10)$$

It is now easy to find the maximum of ν subject to the constraints (7.8)–(7.10). It is obtained when

$$\nu_i = \begin{cases} 1, & \text{for } i \leqslant j, \\ 0, & \text{for } i \geqslant j+2, \end{cases} \quad (7.11a)$$

and

$$\nu_{j+1} = \frac{1}{|\lambda_{j+1}|} \sum_{i<j} \lambda_i. \quad (7.11b)$$

Here, we have used that $\lambda_1 \geqslant \lambda_2 \geqslant \ldots$, and that $\Sigma_j \lambda_i < 0$. Expressed in words, the distribution (7.11) means that the attractor is the most extended along the most unstable directions. Inserting it into eq. (7.9), we obtain

$$\nu \leq j + \frac{\sum_{i \leq j} \lambda_i}{|\lambda_{j+1}|} \equiv D_{KY}. \quad (7.12)$$

as we had claimed.

From the derivation it is clear that the Kaplan–Yorke conjectures $\sigma = D_{KY}$ or $D = D_{KY}$ cannot be expected to hold when *either the attractor is Cantorian in more than one dimension, or if the folding occurs in a direction which is not the minimally contracting one.* The latter was indeed the case for example b in fig. 1. But example b of fig. 1 is not generic, the generic case being the one where the folding is in a plane which encloses an arbitrary angle ϕ with the z-axis (see fig. 15). It is easy to convince oneself that $D = D_{KY}$ whenever $\phi \neq 0$, i.e. nearly always. A still more general case is obtained if we fold in each $(2n)$th iteration in a plane characterized by ϕ_1, and each $(2n+1)$st iteration in a different plane. Again, it seems that $D = D_{KY}$ is generic.

The examples might suggest that indeed $D = D_{KY}$ in all those generic cases in which $\nu = D$, but we consider it as not very likely in high-dimensional cases. For invertible two-dimensional maps, the above conditions are of course satisfied, and thus $\sigma = D_{KY}$ if $\nu = \sigma = D$ (see ref. 29).

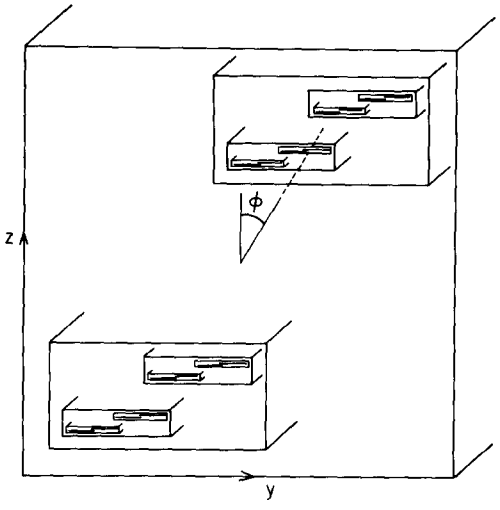

Fig. 15. Cross section through a rectangular volume element and its first 4 iterations under a map which stretches in x-direction, contracts in y- and z-directions (factors $\frac{1}{2}$ and $\frac{1}{4}$, respectively), and folds back under an angle ϕ with respect to the z-direction.

8. Conclusions

The theoretical arguments of section 3 and 6 of this paper have shown (though not with mathematical rigour) that the correlation exponent ν introduced in this paper is closely related to other quantities measuring the local structure of strange attractors.

The numerical results presented in section 2, 5 and 7 have yielded proof that ν can indeed be calculated with reasonable efforts. While all results presented in this paper were based on time series of 10.000–30.000 points, reasonable estimates of ν can already be obtained with series of a few thousand points, in most cases. Surely, for higher dimensional attractors one needs longer time series. However, rather than taking longer time series, we found it in general more important to embed the attractor in higher dimensional spaces, and to choose this embedding dimension judiciously. Compared to box-counting algorithms used previously by other authors, our method has two advantages: First, our storage requirements are drastically reduced. Secondly, in a box-counting algorithm one should iterate until *all* non-empty boxes of a given size l have been visited. This is clearly impractical, in particular if l is very small. Thus, one has systematic errors even if the number of iterations N is excessively large. In our method, there is no such problem. In particular, the finiteness of N induces no systematic errors beyond the corrections to the scaling law $C(l) \sim l^{\nu}$.

We found that in most cases ν was very close to the Hausdorff dimensions D and to the information dimension σ, with two notable exceptions. One was the Feigenbaum map, corresponding to the onset of chaos in 1 dimension. In that case, we were able to compute σ exactly in an analytic way, with the result $\sigma \neq D$, supporting the numerical evidence for $\nu < \sigma$.

The other exception was the Mackey–Glass delay equation, where we found numerically $\nu < D$. The information dimension has not been calculated directly in this case. Accepting the claim made in ref. 8 that the Kaplan–Yorke formula (1.10) predicts correctly σ, we would have $\nu < \sigma = D = D_{KY}$. This seems somewhat surprising, since we argued in section 7 that a rather direct connection (as an inequality $\nu \leqslant D_{KY}$) exists between ν and D_{KY}, while a connection between σ and D_{KY} seems less evident to us.

The main conclusion of this paper, as far as experiments are concerned, is that one can distinguish deterministic chaos from random noise. By analyzing the signal as explained in section 5, and embedding the attractor in an increasingly high dimensional space, one finds whether $C(l)$ scales like l^{ν} or l^{d}. With a random noise the slope of $\log C(l)$ vs. $\log l$ will increase indefinitely as d is increased. For a signal that comes from a strange attractor the slope will reach a value of ν and will then become d independent.

An issue of experimental importance is the effect of random noise *on top* of the deterministic chaos. The treatment of this question is beyond the scope of this paper and is treated elsewhere [30]. Here we just remark that when there is an external noise of a given mean square magnitude, a plot of $\log C(l)$ vs. $\log l$ has two regions. For length scales above those on which the random component blurrs the fractal structure, $C(l)$ continues to scale like l^{ν}. On length scales below those that are affected by the random jitter of the trajectory, $C(l)$ scales like l^{d}. The analysis of experimental signals along these lines can therefore yield simultaneously a characterization of the strange attractor *and* and estimate of the size of the random component. For more details see ref. 30.

It is thus our hope that the correlation exponent will indeed be measured in experiments whose dynamics is governed by strange attractors.

Acknowledgements

This work has been supported in part by the Israel Comission for Basic Research. P.G. thanks the Minerva Foundation for financial support. We thanks Drs. H.G.E. Hentschel and R.M. Mazo for a number of useful discussions.

Appendix A

All numerical calculations were performed in double precision arithmetic on an IBM 370/165 at the Weizmann Institute.

The integrations of the Lorenz and Rabinovich–Fabrikant equations were done using a standard Merson–Runge–Kutta subroutine of the NAG library.

In order to integrate the Mackey–Glass delay equation we approximated it by a N-dimensional set of difference equations by introducing a time step

$$\Delta t = \tau/n\,, \tag{A.1}$$

with n being some large integer, and writing

$$x(t+\Delta t) \approx x(t) + \frac{\Delta t}{2}(\dot{x}(t) + \dot{x}(t+\Delta t))\,. \tag{A.2}$$

Notice that this, being the optimal second-order approximation, is a very efficient algorithm – provided we can compute $\dot{x}(t+\Delta t)$. In the present case we can, due to the special form

$$\dot{x}(t) = f(x(t-\tau)) - bx(t)\,. \tag{A.3}$$

Inserting this in eq. (A.2) and rearranging terms, we arrive at

$$x(t+\Delta t) = \frac{2-b\Delta t}{2+b\Delta t}x(t) + \frac{\Delta t}{2+b\Delta t} \times \{f(x(t-\tau)) + f(x(t-\tau+\Delta t))\}\,. \tag{A.4}$$

In all runs shown in this paper, we used $n = 600$ (corresponding to $0.03 \lesssim \Delta t \lesssim 0.15$), except for the runs with $\tau = 100$, where we used $n = 1200$ and with $n = 600$, finding no appreciable differences.

We also performed control runs with a fourth-order approximation instead of eq. (A.2). The correlation integral was unchanged within statistical errors, and the stability of the solutions did not seem to improve much. This could result from the very large higher derivatives of x, resulting from the tenth power in eq. (6.2).

In order to ensure that all x_i are on the attractor, the first 100–200 iterations were discarded.

Generating the time series $\{X_i\}_{i=1}^{N}$ was indeed the less time-consuming part of our computation, the more important part consisting of calculating the $N(N-1)/2 \gtrsim 10^8$ pairs of distances $r_{ij} = |X_i - X_j|$ and summing them up to get the correlation integral.

In particular, we found that an efficient algorithm for the latter was instrumental in applying the method advocated in this paper.

Such a fast algorithm was found using the fact that floating-point numbers are stored in a computer in the form

$$r = \pm \text{ mantissa} \cdot \text{base}^{+\exp}\,. \tag{A.5}$$

with base = 16 in our case 1/base < mantissa < 1, and exp being an integer. If one can extract the exponent, one can bin the r_{ij}'s in bins of widths increasing geometrically. By extracting the exponent of an arbitrary power r^p of r, one can furthermore choose the width of this binning arbitrarily. Access to the exponent is made very easy and fast by using the shifting and masking operations available e.g. in extended IBM and in CDC Fortran. After having computed the numbers N_K of pairs (i, j) in the interval $2^{k-1} < r_{ij} < 2^k$, the correlation integrals are obtained by

$$c(r = 2^k) = \frac{1}{N^2}\sum_{k'=-\infty}^{k} N_{k'}\,. \tag{A.6}$$

We found this method to be nearly an order of magnitude faster than computing e.g. the logarithmics of r_{ij} directly, and binning by taking their integer parts. A typical run with 20.000 points took – depending on the model studied – between 15 and 30 minutes CPU time.

References

[1] E.N. Lorenz, J. Atmos. Sci. 20 (1963) 130.
[2] R.M. May, Nature 261 (1976) 459.
[3] D. Ruelle and F. Takens, Commun. Math. Phys. 20 (1971) 167.

[4] E. Ott, Rev. Mod. Phys. 53 (1981) 655.
[5] J. Guckenheimer, Nature 298 (1982) 358.
[6] B. Mandelbrot, *Fractals – Form, Chance and Dimension* (Freeman, San Francisco, 1977).
[7] V.I. Oseledec, Trans. Moscow Math. Soc. 19 (1968) 197. D. Ruelle, Proc. N.Y. Acad. Sci. 357 (1980) 1 (R.H.G. Helleman, ed.).
[8] J.D. Farmer, Physica 4D (1982) 366.
[9] H. Mori, Progr. Theor. Phys. 63 (1980) 1044.
[10] J.L. Kaplan and J.A. Yorke, in: Functional Differential Equations and Approximations of Fixed Points, H.-O. Peitgen and H.-O. Walther, eds. Lecture Notes in Math. 730 (Springer, Berlin, 1979) p. 204.
[11] D.A. Russel, J.D. Hanson and E. Ott, Phys. Rev. Lett. 45 (1980) 1175.
[12] H. Froehling, J.P. Crutchfield, D. Farmer, N.H. Packard and R. Shaw, Physica 3D (1981) 605.
[13] P. Grassberger, J. Stat. Phys. 26 (1981) 173.
[14] H.S. Greenside, A. Wolf, J. Swift and T. Pignataro, Phys. Rev. A25 (1982) 3453.
[15] P. Grassberger and I. Procaccia, Phys. Rev. Lett. 50 (1983) 346. Related discussions can be found in a preprint by F. Takens "Invariants Related to Dimensions and Entropy".
[16] M. Feigenbaum, J. Stat. Phys. 19 (1978) 25; 21 (1979) 669.
[17] M.C. Mackey and L. Glass, Science 197 (1977) 287.
[18] M. Hénon, Commun. Math. Phys. 50 (1976) 69.
[19] G.M. Zaslavskii, Phys. Lett. 69A (1978) 145.
[20] M.I. Rabinovich and A.L. Fabrikant, Sov. Phys. JETP 50 (1979) 311. (Zh. Exp. Theor. Fiz. 77 (1979) 617).
[21] W. Feller, An Introduction to Probability Theory and its Applications, vol. 2, 2nd ed. (Wiley, New York, 1971) p. 155.
[22] B.B. Mandelbrot, in: *Turbulence and the Navier–Stokes Equations*, R. Teman, ed., Lecture Notes in Math. 565 (Springer, Berlin, 1975). H.G.E. Hentschel and I. Procaccia, Phys. Rev. A., in press.
[23] D. Stauffer, Phys. Rep. 54C (1979) 1.
[24] T.A. Witten, Jr., and L.M. Sander, Phys. Rev. Lett. 47 (1981) 1400.
[25] N.H. Packard, J.P. Crutchfield, J.D. Farmer and R.S. Shaw, Phys. Rev. Lett. 45 (1980) 712.
[26] F. Takens, in: Proc. Warwick Symp. 1980, D. Rand and B.S. Young, eds, Lectures Notes in Math. 898 (Springer, Berlin, 1981).
[27] P. Frederickson, J.L. Kaplan, E.D. Yorke and J.A. Yorke, "The Lyapunov Dimension of Strange Attractors" (revised), to appear in J. Diff. Eq.
[28] F. Ledrappier, Commun. Math. Phys. 81 (1981) 229.
[29] L.S. Young, "Dimension, Entropy, and Lyapunov Exponents" preprint.
[30] A. Ben-Mizrachi, I. Procaccia and P. Grassberger, Phys. Rev. A, submitted.