Изв.вузов «ПНД», т.3, № 4, 1995

# TOOLS FOR ANALYZING OBSERVED CHAOTIC DATA*

*Henry D.I. Abarbanel*

## 3. Invariant Characteristics of the Dynamics

Classifying the dynamical systems that one observes is a critical part of the analysis of measured signals. When the source of the signals is linear, we are used to the idea of identifying the spectral peaks in the Fourier spectrum of the signal. When we see a spectral peak at some frequency, we know that if we were to stimulate the same system at a different time with forcing of a different strength, we would see the spectral peak in the same location with possibly a different integrated power under the peak. The Fourier frequency is an ***invariant*** of the system motion. The phase associated with that frequency depends on the time at which the measurements begin, and the power under the peak depends on the strength of the forcing. Since chaotic motion produces continuous, broadband Fourier spectra, we clearly have to replace narrowband Fourier signatures with other characteristics of the system for purposes of identification and classification.

The two major features which have emerged as classifiers are ***fractal dimensions*** and ***Lyapunov exponents***. Fractal dimensions are characteristic of the geometric figure of the attractor and relate to the way points on the attractor are distributed in $R^{d_E}$. Lyapunov exponents tell how orbits on the attractor move apart (or together) under the evolution of the dynamics. Both are invariant under the evolution operator of the system, and thus are independent of changes in the initial conditions of the orbit, and both are independent of the coordinate system in which the attractor is observed. This means we can evaluate them reliably in the reconstructed phase space made out of time delay vectors $\mathbf{y}(n)$ as described above; thus, we can evaluate them from experimental data. Each is connected with an ergodic theorem [3] which allows them to be seen as ***statistical quantities characteristic of a deterministic system***. If this seems contradictory, it is only semantic. Once one has a distribution of points in $R^d$, as we do on strange attractors, then using the natural distribution of these points

$$\rho(\mathbf{x}) = \lim_{N\to\infty} 1/N \, \Sigma_{k=1}^{N} \delta^d (\mathbf{x} - \mathbf{y}(k)), \quad (30)$$

we can define statistical quantities with this $\rho(\mathbf{x})$ acting as a density of probability or, better, frequency of occurrence. This $\rho(\mathbf{x})$ is called the natural distribution or measure

* *Продолжение.* Начало см. Изв. вузов «ПНД», т.3, № 2, 3, 1995

since its connection to the number of points in a volume is simple. The number of points within a volume $V$ is phase space is

$$\int_V d^dx\, \rho(\mathbf{x}). \tag{31}$$

In nonlinear systems without any noise, there are many measures. All of them except this one are associated with unstable motions, so this is the only one which survives in any real experiment [3].

Now using this density we can easily see that any function on the phase space, call it $f(\mathbf{x})$ can be used to define an *invariant* under the evolution $\mathbf{y}(k) \to \mathbf{F}(\mathbf{y}(k)) = \mathbf{y}(k+1)$. Indeed, just integrate the function with $\rho(\mathbf{x})$ to form ($N \to \infty$ is implicit)

$$\bar{f} = \int d^dx\, \rho(\mathbf{x}) f(\mathbf{x}) = 1/N \sum_{k=1}^{N} f(\mathbf{y}(k)) = 1/N \sum_{k=1}^{N} f(\mathbf{F}^{\,k-1}(\mathbf{y}(1))). \tag{32}$$

Now it is easy to see that if we evaluate the average of $f\,(\mathbf{F}(\mathbf{x}))$, namely the function evaluated at the point to which $\mathbf{x}$ evolves $\mathbf{x} \to \mathbf{F}(\mathbf{x})$, we find

$$\int d^dx\, \rho(\mathbf{x}) f(\mathbf{F}(\mathbf{x})) = \bar{f} + 1/N\, [f(\mathbf{y}(N+1)) - f(\mathbf{y}(1))], \tag{33}$$

which in the limit of $N \to \infty$, or many data points in practice, is just $\bar{f}$. While an interesting observation, this becomes useful only when we select the function $f(\mathbf{x})$ well. We now turn to interesting choices of $f(\mathbf{x})$.

### 3.1. Fractal Dimensions

Perhaps the most interesting function $f\,(\mathbf{x})$ which is considered in looking for quantities invariant under the action of the dynamics is one which tells us the way in which the number of points within a sphere of radius $r$ scales as the radius shrinks to zero. The relevance of this is that the volume occupied by a volume of radius $r$ in dimension $d$ behaves as $r^d$, so we might expect to achieve a sense of dimension by seeing how the density of points on an attractor scales when we examine it on small scales in phase space.

To motivate this a bit more, imagine a set of points in some Euclidean space, namely the embedding space constructed from the data $\mathbf{y}(n)$; $n=1,2, ..., N$. Go to some point $\mathbf{x}$ on or anyway near the attractor and ask how the number of points on the orbit within a distance $r$ of $\mathbf{x}$ changes as we make $r$ small. Well, not too small, since we have only a finite number of data, and soon the radius will be so small that no points fall within the sphere. So, small enough that there are a lot of points within the sphere and not so large that all the data is within the sphere. The latter would happen if $r \approx R_A$. Then we expect the number of points $n\,(\mathbf{x},r)$ within $r$ at $\mathbf{x}$ to scale as

$$n(\mathbf{x},r) \approx r^{d(\mathbf{x})}, \tag{34}$$

for small $r$. If the attractor were a regular geometric figure of dimension $D$, then in each such ball of radius $r$ around $\mathbf{x}$ we would find approximately $r^D$ points - times some geometric factor of no special relevance here. This would have us set $d(\mathbf{x})=D$ for all $\mathbf{x}$. On an attractor, which is a geometric figure which is not as regular as a sphere or a torus, we do not expect to get the same value for $D$ everywhere, so having $d(\mathbf{x})$ vary with $\mathbf{x}$ seems natural [34,35]. We might call $d(\mathbf{x})$ a local dimension, but since it refers to some specific point on the attractor, we have no particular reason to think it would be the same for all $\mathbf{x}$, and thus under the action of the dynamics $\mathbf{x} \to \mathbf{F}(\mathbf{x})$, it would change. It is sensitive to initial conditions.

To find something invariant under $\mathbf{F}(\mathbf{x})$ we need to define

$$n(\mathbf{x},r) = 1/N \sum_{k=1}^{N} \theta(r - |\mathbf{y}(k) - \mathbf{x}|), \tag{35}$$

where $\theta(u)$ is the Heaviside function

$$\theta(u) = \begin{cases} 1; & \text{if } u > 0, \\ 0; & \text{if } u < 0. \end{cases} \quad (36)$$

This clearly counts all the points on the orbit $\mathbf{y}(k)$ within a radius of $r$ from the point $\mathbf{x}$ and normalizes that number by the total number of data points. Now we recognize that the density of points on an attractor need not be uniform on the figure of the attractor, so it may be quite revealing to look at the moments of the function $n(\mathbf{x},r)$. We choose for our function $f(\mathbf{x}) = n(\mathbf{x},r)^{(q-1)}$, and then we define the function $C(q,r)$ of two variables $q$ and $r$ by the mean of $f(\mathbf{x})$ over the attractor weighted with the natural density $\rho(\mathbf{x})$:

$$C(q,r) = \begin{cases} \int d^d x \rho(\mathbf{x}) n(\mathbf{x},r)^{(q-1)}, \\ 1/M \sum_{k=1}^{M} [\, 1/K \sum_{n=1}^{K} \theta(r - |\mathbf{y}(n) - \mathbf{y}(k)|)]^{(q-1)}. \end{cases} \quad (37)$$

This is often called the «correlation function» on the attractor, and while such quantities were known in statistics from the work of Renyi [27], the idea to examine this kind of quantity to characterize strange attractors is due to Grassberger and Procaccia [28] who originally discussed the case $q$=2.

Now this function of two variables is an invariant on the attractor, but it has become conventional to look only at the variation of this quantity when $r$ is small. In that limit it is *assumed* that

$$C(q,r) \approx r^{(q-1)D_q}, \quad (38)$$

defining the fractal dimension $D_q$ when it exists. Clearly there is a geometric appeal to seeking a single number for each moment of the density of points in the sphere of radius $r$ for small $r$, but we really need not think of this as fundamental. Now we shall see in just a moment that $D_q$ is defined by a limit, and because of this one can show it is invariant under changes in coordinate systems, so it takes on a particularly special geometric meaning. However, if one is working in a given coordinate system, defined by the time delay vectors $\mathbf{y}(n)$ and in a fixed embedding dimension $d_E$, then the whole curve $C(q,r)$ becomes of interest. Since this is the situation one imagines when analyzing data from a given observed source, it seems wise not to throw away the full information in $C(q,r)$ just to focus on the slope of a hoped for linear segment of it when $\lg[C(q,r)]$ is plotted against $\lg[r]$.

One thinks of the dimension $D_q$ as defined by a limit

$$D_q = \lim_{r} \lg[C(q,r)]/((q-1)\lg[r]), \quad (39)$$

and from this definition we can see that the normalization doesn't matter for the evaluation of $D_q$. Also from this definition we can learn that $D_{q-1} \geq D_q$.

In practice we need to compute $C(q,r)$ for a range of $r$ over which we can imagine that the function $\lg[C(q,r)]$ is linear in $\lg[r]$ and then pick off the slope over that range. This is not as easy as it sounds, and numerous papers have been written on how one does this with a finite amount of noisy data and what to do about $r \neq 0$, etc. Essentially all the work has concentrated on the quantity $D_2$ because its evaluation is numerically rather simple and reliable. The papers by Theiler [29], Smith [30], Ruelle [31], and Essex and Nerenberg [32] are critical of how one makes these calculations, what one believes about these calculations, how much data is required for these calculations, and other issues as well. It would be basically out the question to discuss all that has been said on this subject, but a flavor of the discussion can be gotten from those papers. One of the interesting points is the rule of thumb that, if one is to evaluate $D_2$ with some confidence, then a decade of dynamic range in $\lg[r]$ is required. This suggests that at least $10^{D_2/2}$ data points are needed to believe a report of a fractal dimension $D_2$. While only a rule of thumb, it is a useful one.

While as a matter of course I will report some values of $D_2$ evaluated by the method suggested above, it is well worth examining what one might learn by evaluating $D_2$. First of all, since we would have established by now using false nearest neighbours that we have a low dimensional system to work with, and we would know from the same method what dimension we require to unfold the attractor, we would then be seeking a single number to characterize the attractor. If $D_2$ is not integer, this is a very interesting statement about the dynamics at the source of our observed signal, but certainly not a complete characterization. Indeed, at this time no one knows what is a complete set of invariants to characterize an attractor. So, as long as we do not think of this as a way to determine whether the signal comes from a low dimensional source or whether the signal comes from a deterministic source but simply as a characteristic number for that source, we will be on stable ground. It is important to ask just what we will have learned about the source of the chaotic signal from this single number, especially with the remarkable uncertainty that enters its determination from experimental data.

Now this has been a rather long introduction and absent the fact that so much importance has been placed on the dimensions $D_q$, especially $D_2$, it would not have been warranted. Given the huge effort placed on the evaluation of this single quantity and the importance placed on its interpretation, it seems worth making this alert ahead of any numerical display of its values. Personally I am inclined to place more interest in the entire function $C(q,r)$ and its values for a wide range of $r$. When the assumption that $C(q,r)$ behaves as a power of $r$ fails, the function itself is still quite interesting.

Just as a side comment the quantity $d_A$ which we used as «the» dimension of the attractor in discussing the embedding process can now be identified as $d_A = D_0$[7].

**3.1.1.** ***Lorenz Model.*** We choose to display the function $C(2,r)$ for the Lorenz model using quite a bit of well sampled, clean data. Each computation used 50.000 points from the Lorentz systems with $\tau_s = 0.01$; in these dimensionless units an approximate time to go around the attractor is 0.5. 50000 points means circulating about the attractor nearly 1000 times. This is not what one would usually find in observed data, so this is a departure from the tone of this article, but it may be useful to see what a clean example would produce. In Figures 36-38 we display the curves lg$[C(2,r)]$ versus lg$[r]$ evaluated from each of the three variables $x(n)$, $y(n)$, and $z(n)$. The

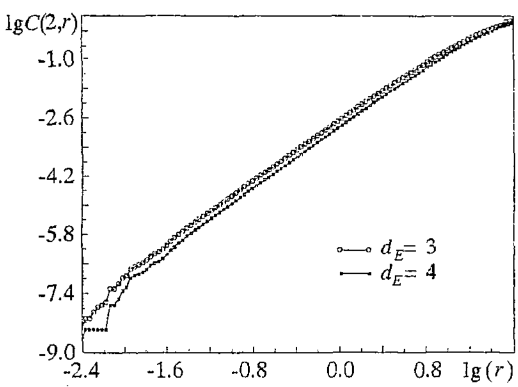


Figure 36. lg $[C(2,r)]$ versus lg $[r]$ for $x(n)$ data from the Lorenz attractor. The global embedding dimension was taken as both $d_E=3$ and $d_E=4$. Note the region with a linear look to it

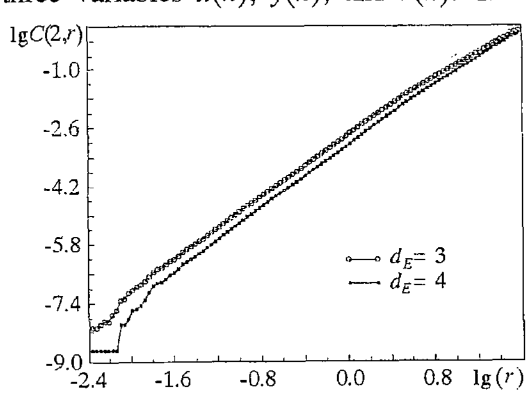


Figure 37. lg $[C(2,r)]$ versus lg $[r]$ for $y(n)$ data from the Lorenz attractor. The global embedding dimension was taken as both $d_E=3$ and $d_E=4$. Note the region with a linear look to it

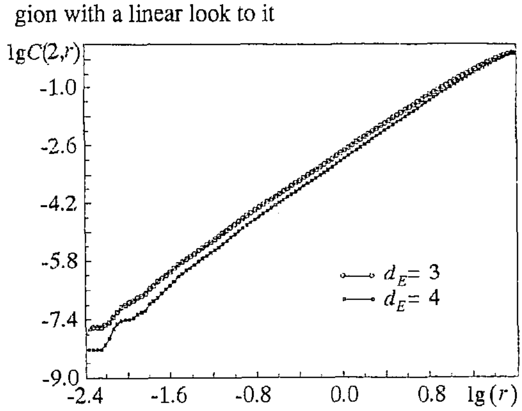


Figure 38. lg $[C(2,r)]$ versus lg $[r]$ for $z(n)$ data from the Lorenz attractor. The global embedding dimension was taken as both $d_E=3$ and $d_E=4$. Note the region with a linear look to it

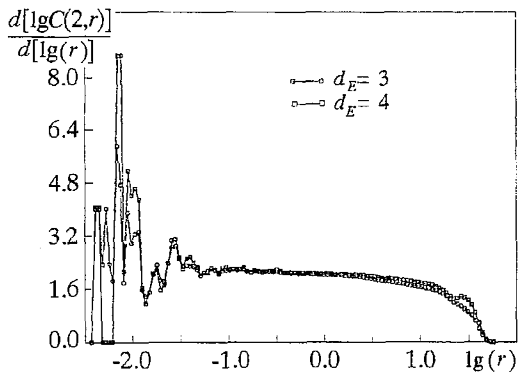


Figure 39. The derivative of lg $[C(2,r)]$ with respect to lg$[r]$ made from $x(n)$ data in the Lorenz model using three neighboring poins in Figure 36 to estimate the derivative. Note that there is only an approximate region where the derivative is constant

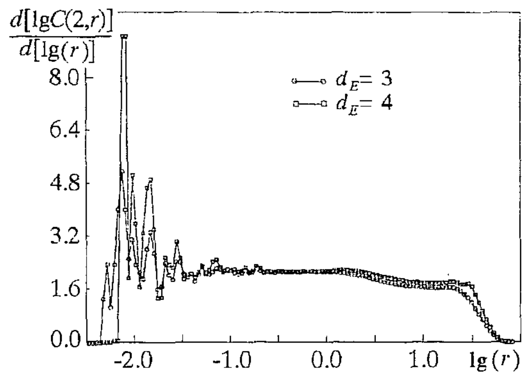


Figure 40. The derivative of lg $[C(2,r)]$ with respect to lg$[r]$ made from $y(n)$ data in the Lorenz model using three neighboring poins in Figure 37 to estimate the derivative. Note that there is only an approximate region where the derivative is constant

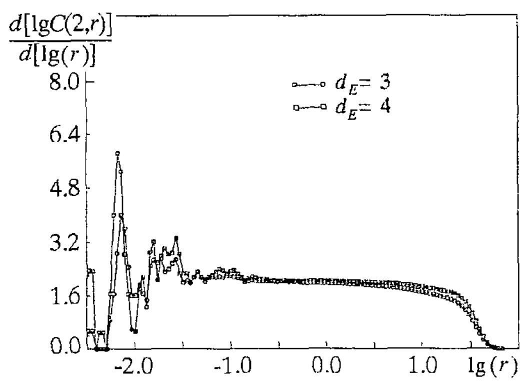


Figure 41. The derivative of lg $[C(2,r)]$ with respect to lg$[r]$ made from $z(n)$ data in the Lorenz model using three neighboring poins in Figure 38 to estimate the derivative. Note that there is only an approximate region where the derivative is constant

curves are shown using vectors in both $d_E$=3 and $d_E$=4. The normalization of the function $C(q,r)$ depends on embedding dimension, but, accounting for that, one sees that the various curves are the same within numerical precision. This is as it should be for invariants and emphasizes the utility both of the whole function $C(q,r)$ and the strict invariance of its logarithm with respect to lg$[r]$. The Figures 39-41 are a plot of

$$d\lg[C(2,r)]/d\lg[r], \quad (40)$$

versus lg$[r]$ for each of the data sets $x(n)$, $y(n)$ and $z(n)$ from the Lorenz system. Derivatives are defined as the local average over three neighbouring points. While one can see a «middle» region in lg$[r]$ where these derivatives are all just above 2, the difficulties of establishing a clean unsullied value for $D_2$ should be clear. With real data it doesn't get better.

### 3.2. Global Lyapunov Exponents

The stability of an observed orbit $\mathbf{y}(k)$ of the dynamical system $\mathbf{y}(k)\to\mathbf{F}(\mathbf{y}(k))=$ $=\mathbf{y}(k+1)$ to small perturbations $\Delta(k)$ is revealed by looking at

$$\mathbf{y}(k+1) + \Delta(k+1) = \mathbf{F}(\mathbf{y}(k) + \Delta(k)) \approx \mathbf{DF}(\mathbf{y}(k))\cdot\Delta(k) + \mathbf{F}(\mathbf{y}(k)), \quad (41)$$

or

$$\Delta(k+1) = \mathbf{DF}(\mathbf{y}(k))\cdot\Delta(k), \quad (42)$$

as long as $\Delta(\bullet)$ remains small. In this the Jacobian matrix

$$DF(\mathbf{x})_{ab} = \partial F_a(\mathbf{x})/\partial x_b \quad (43)$$

enters. The stability of the orbit is determined by the fate of $\Delta(k)$ as the number of evolution steps from the starting time «$k$» grows large. Suppose we move ahead from time $k$ to time $k+L$ using the linearized evolution for $\Delta(k)$. Then we find

$$\Delta(k+L) = \mathbf{DF}(\mathbf{y}(k+L-1))\cdot\mathbf{DF}(\mathbf{y}(K+L-2))\cdots\mathbf{DF}(\mathbf{y}(k)\cdot\Delta(k)) \equiv \mathbf{DF}^L(\mathbf{y}(k))\cdot\Delta(k)), \quad (44)$$

defining the composition of $L$ Jacobians $\mathbf{DF}^L(\mathbf{x})$. Now in an intuitive sense we see that if the eigenvalues of $\mathbf{DF}^L(\mathbf{y}(k))$ behave as $\exp[L\lambda]$ with $\lambda>0$, then the orbit $\mathbf{y}(k)$ at which the Jacobians are evaluated is unstable. This is the generalization from the study we all undertook in school of the linear stability under small perturbations of a fixed point ($\mathbf{y}(k)$ is independent of time).

There is a theorem associated with this problem which is due to the Russian mathematician Oseledec [33]. It is called the ***multiplicative ergodic theorem*** and states that if we look at the length of the perturbation $\Delta(k+L)$, then the essential quantity determining this scalar is

$$[\mathbf{DF}^L(\mathbf{x})]^T\mathbf{DF}^L(\mathbf{x}), \quad (45)$$

where the superscript $T$ means transpose. The length of $\Delta(k+L)$ is

$$|\Delta(k+L)|^2=\Delta^T(k)[\mathbf{DF}^L(\mathbf{x})]^T\,\mathbf{DF}^L(\mathbf{x})\Delta(k). \quad (46)$$

The multiplicative ergodic theorem states that if we form the Oseledec matrix

$$\mathbf{OSL}(\mathbf{x},L) = ([\mathbf{DF}^L(\mathbf{x})]^T\mathbf{DF}^L(\mathbf{x}))^{1/2L}, \quad (47)$$

then the limit of this as $L \rightarrow \infty$ exists and is independent of $\mathbf{x}$ for all $\mathbf{x}$ (well, for almost all $\mathbf{x}$) in the basin of attraction of the attractor to which the orbit belongs. The logarithm of the eigenvalues of this orthogonal matrix when $L \rightarrow \infty$ are denoted $\lambda_1\geq\lambda_2\geq\ldots\geq\lambda_d$ and the notion indicates we order them as shown. The $\lambda_a$ are the ***global Lyapunov exponents*** of the dynamical system $\mathbf{x}\rightarrow\mathbf{F}(\mathbf{x})$. If one or more of the $\lambda_a>0$, then we have chaos. The sum of the Lyapunov exponents $\lambda_1+\lambda_2+\ldots+\lambda_d<0$ by the dissipative nature of the systems we consider.

The $\lambda_a$ are unchanged under smooth changes of coordinate system, so one may evaluate them in any coordinate system one chooses. In particular, the value for the $\lambda_a$ as evaluated in the original coordinates for the system or in the coordinates provided by time delays of any measured quantity are the same. The $\lambda_a$ are also unchanged under the dynamics $\mathbf{x}\rightarrow\mathbf{F}(\mathbf{x})$ when the vector field $\mathbf{F}(\bullet)$ is smooth, as this is essentially a change of coordinates.

The determination of the eigenvalues $\lambda_a$ of the Oseledec matrix is not numerically trivial even though the dimension of the matrix may be small. The point is that $[\mathbf{DF}^L(\mathbf{x})]^T\cdot\mathbf{DF}^L(\mathbf{x})$ is quite ill conditioned as $L$ becomes large. The condition number is approximately $e^{[L(\lambda_1-\lambda_d)]}$. The evaluation if the eigenvalues rests on the idea of a recursive QR decomposition which was described by Eckmann, *et al* [36] and works for large $L$.

To find the Lyapunov exponents from observed scalar data we need some method for accurately determining the Jacobian matrix $\mathbf{DF}(\mathbf{y}(k))$ at locations on the attractor which are visited by the orbit $\mathbf{y}(k)$. For this we require some way to acquire a sense of the variation of the dynamics $\mathbf{y}(k+1)=\mathbf{F}(\mathbf{y}(k))$ in the neighborhood of the observed orbit. The main idea on how we can do this [36,37] is to recognize that attractors are compact objects in their phase space and that any orbit will come back into the neighborhood of any given point on the attractor given a long enough orbit. Thus from one orbit we can acquire information about the phase space behavior of quantities such as $\mathbf{F}(\bullet)$ by looking at the phase space neighbors of $\mathbf{y}(k)$.

Suppose we look at this point $\mathbf{y}(k)$ and find its $N_B$ nearest neighbors: $\mathbf{y}^{(r)}(k)$; $r=1,2\ldots, N_B$. Then each of these neighbors evolves into a known point $\mathbf{y}^{(r)}(k)\rightarrow\mathbf{y}(r,k+1)$ which is in the neighborhood of $\mathbf{y}(k+1)$. The point of the notation is that $\mathbf{y}(r,k+1)$ may not be the $r^{th}$ nearest neighbor of $\mathbf{y}(k+1)$. If we make a ***local*** map from neighborhood to neighborhood:

$$y(r,k+1) = \Sigma_{m=1}^{M} \mathbf{c}(m,k)\phi_m(\mathbf{y}^{(r)}(k)), \tag{48}$$

where the functions $\phi_m(\mathbf{x})$ are some basis set we choose *a priori*, and the $\mathbf{c}(m,k)$ are local coefficients we will determine in a moment, then the components of the Jacobian matrix evaluated at $\mathbf{y}(k)$ are

$$DF_{ab}(\mathbf{y}(k)) = \Sigma_{m=1}^{M} c_a(m,k)\partial\phi_m(\mathbf{x})/\partial x_b|_{\mathbf{x}=\mathbf{y}(k)}. \tag{49}$$

If the local functions are polynomials [38,39], then the linear term in the mapping determines $\mathbf{DF}(\mathbf{y}(k))$, and the rest of the terms make the local mapping accurate. If there are inaccuracies in the evaluation of $\mathbf{DF}(\mathbf{y}(k))$, they tend to be exponentially magnified by the ill conditioned nature of the numerical problem here. Polynomials are only one choice [24,2]. The coefficients $\mathbf{c}(m,k)$ are determined by requiring the residuals

$$\Sigma_{r=1}^{N_B} |\mathbf{y}(r,k+1) - \Sigma_{m=1}^{M} \mathbf{c}(m,k)\phi_m(\mathbf{y}^{(r)}(k))|^2, \tag{50}$$

to be minimized.

We will look at applying this to data after we say a few words about local Lyapunov exponents.

### 3.3. Local Lyapunov Exponents

The global Lyapunov exponents we have just discussed tell us how a perturbation to an orbit $\mathbf{y}(k)$ will behave over a long time averaged over the whole attractor. This may not be particularly relevant information for an engineering approach to this subject since what happens as time gets very large may be of no importance to what happens in the next few nanoseconds. Predicting weather five days from now is certainly more interesting than $10^5$ days (274 years) from now.

The eigenvalues of the Oseledec matrix $\mathbf{OSL}(\mathbf{x},L)$, which we call $e^{[2L\lambda_a(\mathbf{x},L)]}$, tell us how rapidly ($e^{[L\lambda_a(\mathbf{x},L)]}$) perturbations to the orbit at point $\mathbf{x}$ in phase space grow or shrink in $L$ time steps away from the time of the perturbation. These $\lambda_a(\mathbf{x},L)$ are called local Lyapunov exponents [40,41]. They certainly satisfy

$$\lambda_a(\mathbf{x},L) \rightarrow \lambda_a, \tag{51}$$

as $L\rightarrow\infty$. The variations around the limit or global Lyapunov exponent are also of importance. The $\lambda_a(\mathbf{x},L)$ vary quite significantly with the location $\mathbf{x}$ on the attractor, especially for small $L$ which is the main interest. The moments of $\lambda_a(\mathbf{x},L) - \lambda_a$, with moments defined by integrals with the natural density $\rho(\mathbf{x})$, are invariants of the dynamics. The mean

$$\bar{\lambda}_a(L) = \int \rho(\mathbf{x})[\lambda_a(\mathbf{x},L) - \lambda_a] \tag{52}$$

satisfies

$$\bar{\lambda}(L) \approx \lambda_a + K_a/L^{\nu_a} + K_a'/L, \tag{53}$$

where $K_a$, $K_a'$ and $\nu_a<1$ are constants. The last term comes from the geometric dependence of the $\lambda_a(\mathbf{x},L)$ on the coordinate system in which it is evaluated. Other moments behave similarly.

There is quite a bit of information in these local exponents. First of all they tell us on the average around the attractor how well we can predict the evolution of the system $L$ steps ahead of wherever we are. If the second and higher moments are substantial, then this predictability will vary substantially as we move about the attractor. Second, if we are observing data from a flow, that is, a set of differential equations, then one of the $\lambda_a$

*must* be zero [3]. The reason is simple since if we choose to make a perturbation to an orbit exactly along the direction the orbit is going, then that perturbation will be another orbit point and move precisely along the same orbit. Thus the divergence of the new orbit from the old will be absent; this gives $\lambda_a=0$ for that particular direction. In the case of a mapping underlying our data, there is no flow direction. So if we observe an average local Lyapunov exponent $\bar{\lambda}_a(L)$ going to zero, we can be confident that we have a flow. Third, the values of the $\bar{\lambda}(L)$ are each dynamical invariants characterizing the source of the measurements. Their limit, the $\bar{\lambda}_a$, also are invariant under changes in the coordinate system, these having been seen to contribute a term of order $1/L$ to the $\bar{\lambda}_a(L)$.

Finally, the $\lambda_a$ give us a sense of dimension [34,42]. When we have a chaotic system, there is at least one positive Lyapunov exponent. This is the signal of the intrinsic instability we call chaos. The largest of the exponents $\lambda_1$ determines how line segments grow under the dynamics. Areas grow according to $e^{[L(\lambda_1+\lambda_2)]}$. Three dimensional volumes according to $e^{[L(\lambda_1+\lambda_2+\lambda_3)]}$, etc. The sum of all exponents is negative, so somewhere there must be a combination of exponents which can be associated with a volume is phase space which neither grows nor shrinks. Kaplan and Yorke [42] suggested that this be used to define a Lyapunov dimension

$$D_L = K + \Sigma_{a=1}^{K}\lambda_a/|\lambda_{K+1}|, \tag{54}$$

where $\Sigma_{a=1}^{K}\lambda_a>0$ and $\Sigma_{a=1}^{K+1}\lambda_a<0$. This dimension has been associated with the information dimension $D_1$, as defined above, but no general connection seems to have been made to the satisfaction of mathematicians. This is not a big problem since $D_L$ is generally about the same size as most of the $D_q$ we discussed earlier.

**3.3.1.** ***Lorenz Model.*** We take our usual data set from the variable $x(n)$ and form a three dimensional state space. From this we evaluate the local Jacobians as indicated above and using the recursive QR decomposition of the required product of Jacobians we have computed the three $\bar{\lambda}_a(L)$ for the Lorenz system. These are displayed in Figure 42. It is clear that one of the exponents is zero, one is positive, and the third is negative. The values we read off the Figure are $\lambda_1=1.51$, $\lambda_2=0$, and $\lambda_3=-19.0$. With these values the Lyapunov dimension is found to be $D_L=2.08$. In Figure 43 we examine these results in an expanded format. In the calculations just shown we used $d_E=4$ for unfolding the attractor and a local dimension $d_L=3$ for the dynamics. If we had used $d_E=3$ there would have been little change. In Figure 44 we look at another aspect of local Lyapunov exponents by evaluating the exponents in $d_E=4$ and $d_L=4$. In the Figure is shown both the four forward exponents and minus the four backward exponents. True exponents of the dynamics will change sign because of time reversal; however, false exponents associated with our having chosen too large a dynamical dimension $d_E$ will behave otherwise. Indeed, we clearly see that three exponents change under reversing the way we read the data in time, and one does not. Unfortunately this way of determining $d_L$ is not robust against contamination.

**3.3.2.** ***Chaotic Lakes.*** Here we take the data from the Great Salt Lake volume and using the procedure just described arrive at the results shown in Figure 45. We see that there is again a single zero exponent indicating that we have a differential equation describing the dynamics. The Lyapunov dimension is determined by the values of the $\lambda_a$: $\lambda_1=0.17$, $\lambda_2=0.0$, $\lambda_3=-0.14$ and $\lambda_4=-0.65$ in inverse units of $\tau_s=15$ days. This yields $D_L\approx3.05$. Since the predictability time is about $1/\lambda_1$ this means that models for the processes producing fluctuations in the volume of the Great Salt Lake should allow prediction for about three months from any given time before the intrinsic instabilities of the system mask any such ability.

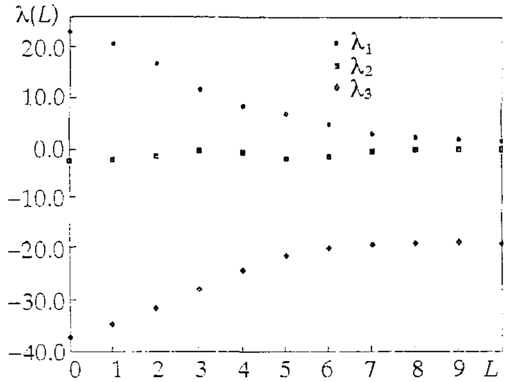


Figure 42. Average Lyapunov exponents for the Lorenz attractor. Data comes from the $x(n)$ variable of the Lorenz model. 50,000 data points were used and cubic neighborhood to neighborhood maps were made with the linear term giving the required local Jacobian to use in the Oseledec matrix. An embedding dimension $d_E$=4 was used in the calculations with a local dimension $d_L$=3 as determined by local false nearest neighbors. Using $d_E$=3 would change little. There is one positive exponent, one zero exponent - indicating the data comes from a flow instead of a mapping - and one negative exponent. The Lyapunov dimension of the system is 2.08

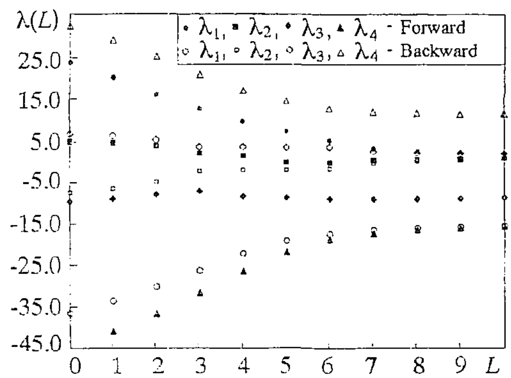


Figure 44. Average Lyapunov exponents for the Lorenz attractor. Data comes from the $x(n)$ variable of the Lorenz model. 50,000 data points were used and cubic neighborhood to neighborhood maps were made with the linear term giving the required local Jacobian to use in the Oseledec matrix. An embedding dimension $d_E$=4 was used in the calculations with a local dimension $d_L$=4. Four Lyapunov exponents are evaluated both forward and then backward in time. True exponents change sign under this operation, and we see that three exponents are true

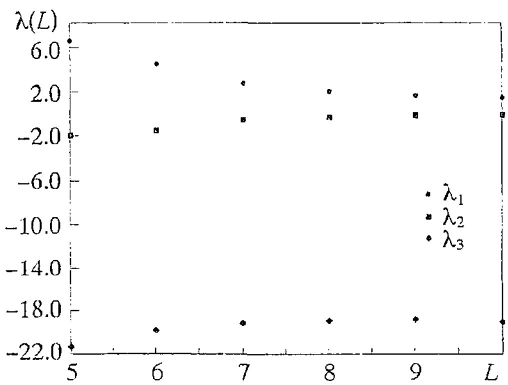


Figure 43. An enlargement of the previous graph showing how the local Lyapunov exponents approach their global values as $L \rightarrow \infty$

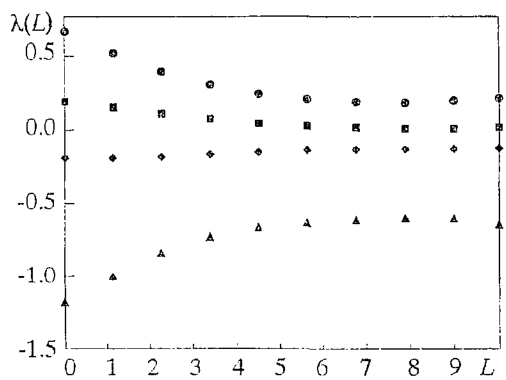


Figure 45. Average local Lyapunov exponents for data from the Great Salt Lake volume. 3463 data points were used and $d_E$=$d_L$=4 as indicated by local false nearest neighbors. The Lyapunov dimension of the attractor is $D_L$≈3.05

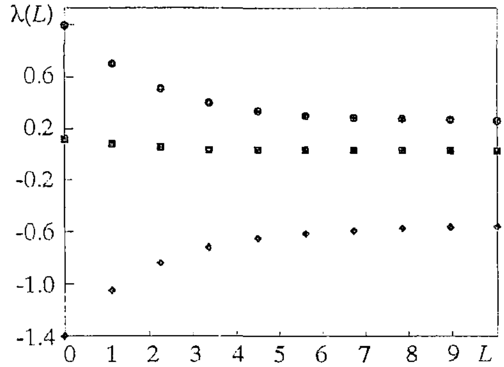


Figure 46. Average local Lyapunov exponents for the hysteretic circuit using data from $V_B(n)$. $d_E$=$d_L$=3 as indicated by local false nearest neighbors. The Lyapunov dimension for this circuit is $D_L$=2.51

**3.3.3. *Chaotic Circuits.*** In analyzing the local Lyapunov exponents from this data source we have our choice of two measured quantities. If we use $V_A$ we must working in a global embedding space $d_E$=5 and make local three dimensional mappings from neighbourhood to neighbourhood. This works, but not as accurately as using $V_B$ and making these same mappings with a global three dimensional phase space. Basically when we have any contamination on our data the evaluation of Lyapunov exponents can be difficult. If our attractor has $d_L$=3 and can be embedded in $d_E$=3, then we do not populate the other two dimensions with contamination and can more accurately determine each of the required local jacobians. One can demonstrate with simulated data where as much uncontaminated data can be made available, that working in $d_E>d_L$ can give $d_L$ accurate exponents.

So we use data from $V_B$ confident that the same numerical values would be achieved by using $d_E$=5 and $d_L$=3 for $V_A$ data. The local Lyapunov exponents which results are displayed in Figure 46. The appearance of a zero exponent tells us we have differential equations underlying this data, and that is what we expect. The positive exponent is $\lambda_1$=0.26 and the third exponent is $\lambda_3$=-0.56. From these values we determine a Lyapunov dimension $D_L$=2.51. The units of these exponents is about 10 per millisecond, so we should be able to predict this circuit forward in time about 0.4 ms.

### 3.4. A Few Remarks About Lyapunov Exponents

As classifiers of the source of a chaotic signal Lyapunov exponents carry a substantial advantage over the use of fractal dimensions. Lyapunov exponents, especially local exponents, as their limit for a large number of time steps gives the global exponents as well, bring not only a set of numbers for classifying the system but also tells us the limits to predictability of the chaotic system and through the Lyapunov dimension, which is conjectured to be $D_1$ among the many $D_q$ fractal dimensions, also gives one a sense of dimension which corresponds accurately to other fractal dimension estimates. It seems that the combination of fractal dimensions, when they can be accurately estimated, or better whole curves of the correlation functions $C(q,r)$ nicely compliment the Lyapunov exponents as classifiers for the dynamics.

Estimating all $d_L$ Lyapunov exponents for a system is quite time consuming if the local exponents are desired as well, but this is a limitation which is likely to be overcome by parallelizing the computations. The point is that doing these computations in serial mode is not natural. Really one wants to estimate the local Jacobian matrices $\mathbf{DF}(\mathbf{y}(k))$ which enter the Oseledec matrix, and this can be done at many phase space locations at the same time. One can expect a speed up in this kind of computation which is linear in the number of processors. In the near future one should be able to handle large data sets and large $d_L$ with this kind of improvement. Indeed, the reason we have not reported on the Lyapunov exponents for the chaotic boundary layer flow or the laser intensity fluctuations is not that we cannot do the computations, but that with the large $d_E$ required (dimension eight to ten) we would have had to use an amount of data which is presently prohibitive to have confidence enough in the answers. The methods described here and in the literature are quite adequate for the task.

A part of this general subject which is of some importance in applications and requires additional attention is the description of local exponents $\lambda_a(\mathbf{x},L)$ as a function of phase space location $\mathbf{x}$ and a nice means for locating those regions where all the $\lambda_a(\mathbf{x},L)$ are negative. Such regions stand out as the phase space places to which one might want to control a system because the local stability is so marked. Another important topic is that of exploring the idea of regional Lyapunov exponents which involve averaging over real space locations in the underlying system. We saw the role these might play when we looked at data from the Great Salt Lake and from the boundary layer chaos. In each case we are seeing data with small scale motions suppressed yet very interesting dynamics on the larger scales sensed in the observations. How the Lyapunov exponents of a system vary with spatial averaging is the issue. Another application area of some engineering interest is the behavior of continuum systems such as elastic beams when the high frequency structural modes are suppressed by the sensor.

## References

[2] Abarbanel, H.D.I., R.Brown, J.J.(«SID») Sidorowich, and L.Sh.Tsimring, «The Analysis of Observed Chaotic Data in Physical Systems», to appear in Rev.Mod.Phys October, 1993.

[3] Eckmann, J.-P. and D.Ruelle «Ergodic Theory of Chaos and Strange Attractors» Rev. Mod. Phys. **57**, 617 (1985).

[7] Casdagli, M.,T.Sauer, and J.A.Yorke, «Embedology», J.Stat.Phys. **65**, 579-616 (1991)

[24] Parlitz, U., «Identification of True and Spurious Lyapunov Exponents from Time Series», Int. J.Bif.Chaos 2,155-165(1992).

[27] Renyi, A.Probability Theory, North-Holland, Amsterdam, 1970; J.Balatoni and A. Renyi, Pub. Math. Inst. Hungarian Acad.Sci. **1**, 9 (1956)

[28] Grassberger, P. and I.Procaccia, «Characterization of Strange Attractors», Phys. Rev. Lett., **50**, 346-349, (1983); «Measuring the Strange Attractors», Physica D**9**, 189-208 (1983).

[29] Theiler, J. «Estimating Fractal Dimension» J.Opt. Soc. Am.A **7**, 1055-1073 (1990) and other papers referred to therein.

[30] Smith, L.A. «Intrinsic Limits on Dimension Calculations» Phys. Lett. A **133**, 283-288 (1988)

[31] Ruelle, D., «Deterministic Chaos: the Science and the Fiction», Proc.Roy.Soc. London A **427**, 241-248, (1990).

[32] Essex, C. and Nerenberg, M.A.H., «Comments on Deterministic Chaos: the Science and the Fiction by D.Ruelle», Proc.Roy.Soc.London A **435**, 287-292 (1991).

[33] Oseledec, V.I., «A Multiplicative Ergodic Theorem. Lyapunov Characteristic Numbers for Dynamical Systems», Trudy Mosk.Mat. Obsc **19**, 197 (1968); Moscow Math.Soc. **19**, 197 (1968).

[34] Ott,E. Chaos in Dynamical Systems, Cambridge University Press, New York (1993).

[35] Paladin, G. and A.Vulpiani, «Anomalous Scaling Laws in Multifractal Objects», Phys. Repts. **156**, 147-225 (1987). This is a good discussion of the inhomogeneity of points distributed on an attractor. It provides a clear statement of how one determines quantitative measures of this inhomogeneity and how one looks for physical effects of this.

[36] Eckmann, J.-P.,S.O. Kamphorst, D.Ruelle, and S.Ciliberto, «Liapunov Exponents from Time Series», Phys.Rev. A**34**, 4971-4979 (1986). A similar recursive method was shown to me independently by E.N.Lorenz, private communication (1990).

[37] Sano,M. and Y.Sawada, «Measurement of Lyapunov Spectrum from a Chaotic Time Series», Phys. Rev. Lett. **55**, 1082 (1985).

[38] Brown,R., Bryant,P. and H.D.I. Abarbanel, «Computing the Lyapunov Spectrum of a Dynamical System from Observed Time Series» Phys. Rev. A **43**, 2787 (19910; Phys. Rev. Lett. **65**, 1523 (1990).

[39] Briggs,K., «An Improved Method for Estimating Liapunov Exponents of Chaotic Time Series», Phys. Lett., **151A**, 27-32 (1990).

[40] Abarbanel, H.D.I., Brown,R., and Matthew B.Kennel, «Variation of Lyapunov Exponents on a Strange Attractor», Journal of Nonlinear Science, **1**, 175-199 (1991).

[41] Abarbanel, H.D.I., R.Brown, and Matthew B.Kennel, «Local Lyapunov Exponents from Observed Data», J.Nonlin. Sci. **2**, 343-365 (1992).

[42] Kaplan, J.L. and Yorke, J.A., «Chaotic Behavior in Multidimensional Difference Equations», Lecture Notes in Mathematics, **730**, 228 (1979).

*Institute for Nonlinear Science*

*Поступила в редакцию 18.12.94*