NOISE REDUCTION: FINDING THE SIMPLEST DYNAMICAL SYSTEM CONSISTENT WITH THE DATA

Eric J. KOSTELICH^{a,b,1} and James A. YORKE^{a,c}

^a Institute for Physical Science and Technology, University of Maryland, College Park, MD 20742, USA ^b Center for Nonlinear Dynamics, Department of Physics, University of Texas, Austin, TX 78712, USA ^c Department of Mathematics, University of Maryland, College Park, MD 20742, USA

Received 3 March 1989 Revised manuscript received 11 October 1989 Accepted 18 October 1989 Communicated by R. Westervelt

A novel method is described for noise reduction in chaotic experimental data whose dynamics are low dimensional. In addition, we show how the approach allows experimentalists to use many of the same techniques that have been essential for the analysis of nonlinear systems of ordinary differential equations and difference equations.

1. Introduction

Numerical computation and computer graphics have been essential tools for investigating the behavior of nonlinear maps and differential equations. The pioneering work of Lorenz [25] was made possible by numerical integration on a computer, allowing him to take nearby pairs of initial conditions and compare the trajectories. Hénon [19] discovered the complex dynamics of his celebrated quadratic map with the aid of a programmable calculator. A variety of classical and modern techniques has been exploited to find periodic orbits, their stable and unstable manifolds [14], basins of attraction [26], fractal dimension [27], and Lyapunov exponents [10, 31, 37]. In some cases, numerical methods can establish rigorously the existence of initial conditions whose trajectories have essentially the same intricate structure that one sees on a computer screen [18].

Until recently, experimentalists have not been able to apply most of these methods to the analysis of experimental data, since they do not in general have explicit equations to model the behavior of their apparatus. In cases where it is possible to find accurate models of the physical system, quantitative predictions about the behavior of actual experiments are possible [17]. However, all that is available in a typical experiment is the time-dependent output (e.g. voltage) from one or more probes, which is a function of the dynamics.

One fundamental problem in the analysis of experimental data concerns the correspondence between the dynamics that governs the behavior of the apparatus and the discretely sampled time series that comprises the data. Another question is how to minimize the effect of noise. In this paper, we show how the *time delay embedding method*, now commonly used to reconstruct an attractor from experimental data, yields a novel procedure for reducing noise in data whose dynamics can be characterized as low dimensional. Moreover, we

0167-2789/90/\$03.50 © Elsevier Science Publishers B.V. (North-Holland)

¹Current address: Department of Mathematics, Arizona State University, Tempe, AZ 85287, USA.

show how the approach can be extended to allow experimentalists access to many of the analytical tools mentioned above.

Section 2 reviews the time delay embedding method and some of its applications. Section 3 introduces some of the problems associated with traditional filters and outlines our noise reduction method.

2. The time delay embedding method

As stated in section 1, one problem in analyzing experimental data is how to relate the measurements with the dynamics. Before the early 1980's, power spectra were the principal method for analyzing such data. For instance, Fenstermacher et al. [13] relied heavily on power spectra to detect transitions from periodic to weakly turbulent flow between concentric rotating cylinders. However, Fourier analysis alone is inadequate for describing the dynamics.

Other methods also have been used to analyze time series output from dynamical systems. Lorenz [25] used next amplitude maps to describe some features of the dynamics; that is, he plotted z_{n+1} against z_n where z_n is the *n*th relative maximum of the third coordinate of the numerically calculated solution. Such maps are often useful, not only for investigating features of the Lorenz attractor [32], but also for instance in experiments on intermittency in oscillating chemical reactions [30].

In the past decade, the time delay embedding method has come into common use as a way of reconstructing an attractor from a time series of experimental data. In this approach, one supposes that the dynamical behavior is governed by a solution traveling along an attractor^{#1} (which is not observable directly). However, one assumes there is a smooth function that maps points on the attractor to real numbers (the experimental measurements). In the embedding method, one generates a set of *m*-dimensional points whose coordinates are values in the time series separated by a constant delay [11]. For example, when m = 3, the reconstructed attractor is the set of points $\{x_i = (s_i, s_{i+\tau}, s_{i+2\tau})\}$ where τ is the time delay. Takens [34] has shown that under suitable hypotheses, this procedure yields a set whose properties are equivalent to those of the original attractor provided that the embedding dimension *m* is large enough.

In principle, the embedding method allows one to study the dynamics in detail. The earliest applications may be called *static* in that the analysis focuses on the geometric properties of the set of points on the reconstructed attractor. For example, phase portraits and Poincaré sections are used in ref. [5] to help determine the transition between quasiperiodic and chaotic flow in a Couette-Taylor experiment. Another important application is the estimation of attractor dimension from experimental data, for which there is a large literature [27]. In addition, various information theoretic notions can be used to find good choices of embedding dimension and time delay [15].

More recent applications of the embedding method are quite different in nature and can be called *dynamic* in that information about the dynamics is stored in the computer for analysis. With each data vector x_i , one stores the "next" vector, for example, $x_{i+\delta}$ for some $\delta > 0$. This makes it possible to compute a linear approximation of the dynamics in a neighborhood of x_i , assuming that there is a low-dimensional dynamical system underlying that data^{#2}. In particular, a linear approximation provides an estimate of the Jacobian of the map at x_i [11]. Eckmann et al. [10] use linear maps computed in this way to integrate a set of variational equations and find the positive Lyapunov exponents^{#3}.

^{#1}Existing numerical methods require the attractor to be low dimensional.

^{#2}This material was first presented by D. Ruelle at a Nobel symposium in 1984.

^{#3}Wolf et al. [37] have proposed a different method in which nearby pairs of points are followed to estimate the largest Lyapunov exponent.

In fact, the time delay embedding method provides a powerful set of tools for analyzing the dynamics, the breadth of which may not have been realized by Eckmann and Ruelle. In the remainder of this paper, we discuss two novel applications that are possible, specifically:

(1) Noise reduction. Since one can approximate the dynamics at each point, it becomes possible to identify and correct inaccuracies in trajectories arising from random errors in the original time series. Numerical evidence suggests that the noise reduction procedure described below improves the accuracy of other analyses, such as Lyapunov exponents and dimension calculations.

(2) Simplicial approximations. Linear approximations can be computed at each point on a grid in a neighborhood of the attractor to form a simplicial approximation of the dynamical system. This can be used to locate unstable periodic orbits near the attractor.

We consider noise reduction in section 3.

3. Noise reduction

The ability to extract information from timevarying signals is limited by the presence of noise. Recent experiments to study the transition to turbulence in systems far from equilibrium, like those by Fenstermacher et al. [13], Behringer and Ahlers [2], and Libchaber et al. [24], succeeded largely because of instrumentation that enabled them to quantify and reduce the noise. However, it is often expensive and time consuming to redesign experimental apparatus to improve the signal to noise ratio.

An important question, therefore, is how the experimental data can be filtered or otherwise preprocessed before it is analyzed further. One common approach is to use Fourier analysis: one might model the noise as a collection of highfrequency components and subtract them from a power spectrum (or Fourier transform) of the input data. The transform can be inverted to yield a new time series with some of the high-frequency components removed. This is the basic idea behind Wiener and other bandpass filters [29].

However, as noted previously, power spectral analysis is insufficient to characterize the dynamics when the data are chaotic. Since the power spectrum of a low-dimensional chaotic signal resembles that of a noisy one, the suppression of certain frequencies can alter the dynamics of the filtered output signal. Badii et al. [1] have shown that a simple low-pass filter effectively introduces an extra Lyapunov exponent that depends on the cutoff frequency. If the cutoff frequency is sufficiently low, then the filter can increase the fractal dimension of the reconstructed attractor. This result also has been confirmed by Mitschke et al. [28] with data from an electronic circuit.

We now consider a different approach and show how the time delay embedding method can be exploited to reduce the noise, at least in cases where the time series can be viewed as a dynamical system with a low-dimensional attractor. Our objective is to use the dynamics to detect and correct errors in trajectories that result from noise. This is done in two steps once an embedding dimension m and a time delay τ have been fixed.

In the first step, we consider the motion of an *ensemble* of points in a small neighborhood of each point on the attractor in order to compute a linear approximation of the dynamics there. In the second step, we use these approximations to consider how well an *individual* trajectory obeys them. That is, we ask how the observed trajectory can be perturbed slightly to yield a new trajectory that satisfies the linear maps better. The trajectory adjustment is done in such a way that a new time series is output whose dynamics are more consistent with those on the phase space attractor.

This approach is fundamentally different from traditional noise reduction methods. Because we consider the motion of points on a phase space attractor, we are using information in the original signal that is not localized in a time or frequency domain. Points that are close in phase space correspond to data that in general are widely and irregularly spaced in time, due to the sensitive dependence on initial conditions on chaotic attractors. In contrast, Kalman [4] and similar filters examine data that are closely spaced in time; bandpass filters operate in the frequency domain.

4. Eckmann-Ruelle linearization

The discrete sampling of the original signal means that the points on the reconstructed attractor can be treated as iterates of a nonlinear map f whose exact form is unknown. We assume that f is nearly linear in a small neighborhood of each attractor point x and write

$$f(\mathbf{x}) \approx A\mathbf{x} + \mathbf{b} \equiv L(\mathbf{x})$$

for some $m \times m$ matrix A and m-vector **b**. (The matrix A is the Jacobian of f at x.)

This approximation, which we call the Eckmann-Ruelle linearization at x, can be computed with least-squares methods similar to those described in refs. [11, 10]. Given a reference point x_{ref} , let $\{x_i\}_{i=1}^n$ be a collection of the *n* points which are closest to x_{ref} . With each point x_i we store the next point (i.e., the image of x_i), denoted $y_i^{\#4}$. The *k* th row a_k of *A* and the *k* th component b_k of **b** are given by the least-squares solution of the equation

$$y_k = b_k + a_k \cdot x, \tag{1}$$

where y_k is the kth component of y and the dot denotes the dot product. Fig. 1 illustrates the idea^{#5}.



Fig. 1. Schematic diagram for the first stage of the noise reduction method. A collection of points in an ϵ -ball about the reference point x_{ref} is used to find a linear approximation of the dynamics there.

We mention three difficulties in computing the local linear approximations in the subsections below.

4.1. Ill conditioned least squares

There is a particular problem when one tries to compute solutions to eq. (1) with a finite data set of limited accuracy that has not been addressed in previous papers [10, 31]. Suppose for example that all the points in a neighborhood of x_{ref} lie nearly along a single line, i.e., the attractor appears one dimensional within the available resolution. Although it is possible to measure the expansion along the unstable manifold at x_{ref} , there are not enough points in other directions to measure the contraction. Hence it is not possible to compute a 2×2 Jacobian matrix accurately. Any attempt to do so will result in an estimate of the Jacobian whose elements have large relative errors, This kind of least-squares problem is *ill conditioned*.

The ill conditioning can be avoided by changing coordinates so that the first vector in the new basis points in the unstable direction^{#6}. A one-dimensional approximation of the dynamics is computed using the new coordinates; that is, we approximate the dynamics only along the unstable manifold. We recover the matrix A by changing coordinates back to the original basis.

For example, if we are working in the plane and the unstable direction is the line y = x, then we rotate the coordinate axes by 45°. The dynamics are approximated by a one-dimensional linear map

^{#4}The points x_i are points on the attractor which are *not* consecutive in time. The subscript *i* merely enumerates all the points on the attractor contained within a small distance ϵ of x_{ref} . In this notation, x_i and y_i are consecutive in time. ^{#5}Farmer and Sidorowich [12] observe that the Eckmann-

^{*} Farmer and Sidorowich [12] observe that the Eckmann-Ruelle linearization can be used for prediction. Given a reference point x_i , find the Eckmann-Ruelle linearization $A_i x + b_i$, compute $x_{i+1} = A_i x_i + b_i$, and repeat the process to get the predicted trajectory.

^{#6}This is done by computing the right singular vectors [9] of the $n \times m$ matrix whose *j*th row is x_j . The procedure is called *principal component analysis* in the statistical literature.

computed along the line y = x. Then we rotate back to the original coordinates. (The resulting matrix A has rank 1 in this example.) This approach substantially enhances the robustness of the numerical procedure.

4.2. Finding nearest neighbors

A second problem is finding an efficient way to locate all of the points closest to a given reference point. The dynamical embedding method imposes stringent requirements on any nearest-neighbor algorithm. The storage overhead for the corresponding data structures must be small, because there are tens of thousands of attractor points. The algorithm must be fast, since there is one nearest-neighbor problem for each linear map to be computed.

We solve this problem by partitioning the phase space into a grid of boxes that is parallel to the coordinate axes. Each coordinate axis is divided into B intervals. (Fig. 2 illustrates the grid in two dimensions.) Each point on the attractor is assigned a box number according to its coordinates. For example, a point on the plane whose first coordinate falls in the *i*th interval (counting from 0) along the x axis and whose second coordinate falls in the kth interval along the y axis is assigned to box number kB + j. The list of box numbers is sorted, carrying along a pointer to the original data point. Given a reference point x_{ref} , its box number is found using the above formula. A binary search in the list of box numbers then locates the address of x_{ref} and all the other points

| $B^2 - B$ | $B^2 - B + 1$ | $B^2 - B + 2$ | | $B^2 - 1$ |
|-----------|---------------|---------------|---|----------------|
| : | : | : | : | ; |
| В | <i>B</i> + 1 | <i>B</i> + 2 | | 2 <i>B</i> ~ 1 |
| 0 | 1 | 2 | | B - 1 |

Fig. 2. Box numbering scheme in two dimensions. The attractor is normalized to fit in the unit square. The bottom row of boxes rests against the x axis and the leftmost row of boxes against the y axis.

in the same box number. The search is extended if necessary to adjacent boxes.

Only a crude partition is needed for this algorithm to work efficiently (typically we choose B = 40), and the grid is extended only to the first three coordinate axes. When the embedding dimension is larger than three, a preliminary list of nearest neighbors is obtained using only the first three coordinates of each attractor point. The final list is extracted by computing the distances from x_{ref} to each point in the preliminary list.

Although there are circumstances where this algorithm can perform poorly (e.g., when most of the attractor points are concentrated in a handful of boxes), the distribution of points on typical attractors is sufficiently uniform that the running time is very fast. Memory use is also efficient: a set of N attractor points requires 3N storage locations. In contrast, the tree-search algorithm advocated in ref. [12] requires several times more storage (although the lookup time is probably slightly less). Because $N \approx 10^5$ in typical applications, we believe that the box-grid approach (or some variant) is the most practical. A survey of other nearest-neighbor algorithms is given in ref. [3].

4.3. Errors in variables

There is a potential difficulty in the use of ordinary least squares to compute the linear maps. In the usual statistical problem of fitting a straight line, one has observations (x_i, y_i) where x_i is known exactly and y_i is measured. One assumes that $y_i = a_0 + a_1 x_i + \epsilon_i$, where the ϵ_i are independent errors drawn from the same normal distribution. (Analogous assumptions hold in the multivariate case.) In the present situation, however, both x_i and y_i are measured with error. It can be shown that the ordinary least-squares method produces biased estimates of the parameters a_0 and a_1 in this case [16, 23]. In practice this does not seem to be a serious problem, but statistical procedures to handle this situation (the socalled "errors in variables" methods) may provide



Fig. 3. Schematic diagram of the trajectory adjustment procedure. The trajectory defined by the sequence $\{x_i\}$ is perturbed to a new trajectory given by $\{\hat{x}_i\}$ which is more consistent with the dynamics. In this example we show what the perturbed trajectory might look like if the dynamics were approximately horizontal translation to the right.

an alternative approach to noise reduction. We consider this question in the appendix.

5. Trajectory adjustment by minimizing self-inconsistency

The Eckmann-Ruelle linearization procedure described above is computed and the resulting maps are stored for a sequence of reference points along a given trajectory (for the results quoted here, the sequence usually contains 24 points). We now consider how to perturb this trajectory so that it is more consistent with the dynamics. The objective is to choose a new sequence of points \hat{x}_i to minimize the sum of squares

$$\sum w \|\hat{x}_{i} - x_{i}\|^{2} + \|\hat{x}_{i} - L_{i-1}(\hat{x}_{i-1})\|^{2} + \|\hat{x}_{i+1} - L_{i}(\hat{x}_{i})\|^{2}, \quad (2)$$

where $L(x_i) = A_i x_i + b_i$, w is a weighting factor, and the sum runs over all the points along the trajectory^{#7}. Eq. (2) can be solved using least squares. Heuristically, eq. (2) measures the selfinconsistency of the data, assuming that the linear approximations of the dynamics are accurate. See fig. 3. We say the new sequence $\{x_i\}$ is more *self-consistent*. The trajectory adjustment can be iterated. That is, once a new trajectory \hat{x}_i has been found, one can replace each x_i in eq. (2) by \hat{x}_i and compute a new sequence $\{\hat{x}_i\}$.

We place an upper limit on the distance a point can move. Points which seem to require especially large adjustments can be flagged and output unchanged. (This may be necessary if the input time series contains large "glitches" or if nonlinearities are significant over small distances in certain regions of the attractor.)

When the input is a time series, we modify the above procedure slightly since we require a time series as output. The trajectory adjustment is done so that changes to the coordinates of x_i (corresponding to particular time series values) are made consistently for all subsequent points whose coordinates are the same time series values. For example, suppose the time delay is 1 and the embedding dimension is 2. Then trajectories are perturbed so that the second coordinate of the *i*th point is the same as the first coordinate of the (i + 1)st point. That is, when $x_i = (s_i, s_{i+1})$ is moved to the point $\hat{x}_i = (\hat{s}_i, \hat{s}_{i+1})$, we require that the first coordinate of \hat{x}_{i+1} be \hat{s}_{i+1} .

6. Results using experimental data

We note that the attractor need not be chaotic for this noise reduction procedure to be effective. Fig. 4a shows a phase portrait of noisy measurements of wavy vortex flow in a Couette-Taylor experiment [20]. This flow is periodic, so the attractor is a limit cycle (widened into a band because of the noise) and the power spectrum consists of one fundamental frequency and its harmonics above a noise floor. See fig. 4b. Figs. 4c, 4d show the same data after noise reduction. The noise reduction procedure makes the limit cycle much narrower, and the noise floor in the power spectrum is reduced by almost two orders of magnitude. However, no power is subtracted from any of the fundamental frequencies, and in fact some harmonics are revealed which previously were obscured by the noise.

 $^{*^7}$ In the results described in this paper, the Eckmann-Ruelle linearization procedure is done using a collection of points within a radius of 1-6% of each reference point, depending on the embedding dimension, the dimension of the attractor, and the number of attractor points. This results in collections of 50-200 points per ball, which gives reasonably accurate map approximations without making the computer program too slow. The weighting factor w is set to 1.



Fig. 4. Phase portraits and power spectra for measurements of wavy vortex flow in a Couette-Taylor experiment. (a), (b) Phase portrait and power spectrum before noise reduction is applied; (c), (d) after noise reduction; (e), (f) after a low-pass filter is applied to the original data. The vertical axis in (b), (d) and (f) is the base-10 logarithm of the power spectral density; the horizontal axis is in multiples of the Nyquist frequency.

These results are significantly different from those obtained by low-pass filtering. Figs. 4e, 4f show the phase portrait and power spectrum when the original data are passed through a 12th-order Butterworth filter with a cutoff frequency of 0.35. The dynamical noise reduction procedure is more effective than low-pass filtering since the noise appears to have a broad spectrum.

However, the dynamical noise reduction method appears to subtract power from a mode whose fundamental frequency is approximately 0.3 times the Nyquist frequency. We do not know exactly



Fig. 5. Phase portraits and power spectra for measurements of weakly chaotic flow in a Couette-Taylor experiment. (a), (b) Phase portrait and power spectrum before noise reduction is applied; (c), (d) after noise reduction. The units for the power spectrum plots are the same as those in ref. [5].

why this occurs. However, this peak corresponds to the rotation frequency of the inner cylinder and may result from a defect in the Couette-Taylor apparatus [33]. We do not consider this to be a serious problem, because the power associated with this mode is several orders of magnitude smaller than that of the wavy vortex flow.

We emphasize that our objective is to find a simple dynamical system that is consistent with the data. It is possible for this method to eliminate certain dynamical behavior from an attractor if those dynamics have very small amplitude, as fig. 4f shows. This situation is most likely to arise when there are not enough data to distinguish such dynamics from random noise. In the present example, the noise reduction procedure reveals the limit cycle behavior quite well^{#8}.

The results obtained by applying the method to chaotic data from the Couette-Taylor fluid flow experiment described in ref. [5] are shown in fig. 5. Fig. 5a shows a two-dimensional phase portrait of the raw time series at a Reynolds number $R/R_c =$ 12.9, which corresponds to weakly chaotic flow [5]. The corresponding phase portrait from the filtered time series is shown in fig. 5b. Figs. 5c, 5d show

^{#8}We have not attempted to find the smallest amplitude at which the noise reduction procedure can distinguish quasiperiodic from periodic flow. In general this will depend on the amount of data, the sampling rate, the embedding dimension, and other factors.

the power spectra for the corresponding time series^{#9}.

It is difficult to estimate how much noise is removed from the data in this example on the basis of power spectra. One problem is that the transition from quasiperiodic to weakly chaotic fluid flow is marked by a sudden rise in the noise floor in the power spectrum (cf. fig. 3 in ref. [5]). Hence one cannot determine how much of the noise floor is due to deterministic chaos and how much results from broad-band noise. The noise reduction procedure described here has the effect of reducing the power in the high-frequency components of the signal. One question therefore is whether reducing the high-frequency noise corresponds to discovering the true dynamics which have been masked by noise. We believe that the answer is yes, based on those cases where there is an underlying low-dimensional dynamical system. However, in chaotic processes some high-frequency components remain, because they are appropriate to the dynamics.

7. Numerical experiments on noise reduction

One important question is how much noise this method removes from the data. The power spectra above suggest that the method eliminates most of the noise, but it is impossible to give a precise estimate for typical chaotic experimental data.

However, the Hénon map [19] provides a convenient way to quantify the noise reduction, because it can be written as a time delay map of the form

$$x_{i+1} = f(x_i, x_{i-1}) = 1 - \alpha x_i^2 + \beta x_{i-1}.$$
 (3)

We use eq. (3) to generate a time series as follows (with the standard parameter values $\alpha = 1.4$, $\beta = 0.3$). We choose an initial condition and discard the first 100 iterates. The next 32 768 iterates are stored, and a time series is generated by adding a uniformly distributed random number to each iterate. This simulates a time series with *measurement noise*, i.e., a time series where noise results from errors in measuring the signal, not from perturbations of the dynamics.

We measure the improvement in the signal after processing by considering the *pointwise error*

$$e_i = ||x_{i+1} - f(x_i, x_{i-1})||_{t}$$

i.e., the distance between the observed image and the predicted one. Let the *mean error* be

$$E = \left(\frac{\sum e_i^2}{N}\right)^{1/2},$$

the rms value of the pointwise error over all N points on the attractor. We define the *noise reduction* as

$$R = 1 - E_{\text{fitted}} / E_{\text{noisy}}$$

where the mean errors are computed for the adjusted and original noisy time series, respectively. The quantity R is a measure of the self-consistency of the time series. (In other words, Rmeasures how much better on the average the output attractor obeys eq. (3) as one hops from point to point.)

When 1% noise is added to the input as described above, the noise reduction (measured with the actual map) is $79\%^{\pm 10}$. Nearly identical results are obtained when the input contains only 0.1% noise. In addition, noise levels can be reduced almost as much in cases where the noise is added to the dynamics, i.e., where the input is of the form $\{x_{i+1}: x_{i+1} = f(x_i + \eta_i, x_{i-1} + \eta_{i-1}), \eta_i, \eta_{i-1} \text{ random}\}$. When the program is run on noise-less input, the mean error in the output is 0.025% of the attractor extent, which suggests that errors

^{#9}The time series consists of 32768 values, from which an attractor is reconstructed in four dimensions. Linear maps are computed using 50–100 points in each ball. Trajectories are fitted using sequences of 24 points.

 $^{^{\}pm 10}$ The pointwise error is measured using eq. (3). However, the attractor can be embedded in more than two dimensions when performing the noise reduction.

arising from small nonlinearities are negligible when the input contains enough points.

8. Simplicial approximations of dynamical systems

Recent work has shown that simplicial approximations of dynamical systems can reproduce the behavior of the original system to high accuracy [36]. (See also ref. [35] for a bilinear approach.) In particular, the fractal structure of the original attractors and basin boundaries is preserved over many scales. Such approximations can yield significant computational savings, especially when the original system consists of ordinary differential equations.

This approach can be extended in a natural way to generate simplicial approximations of the dynamics on attractors reconstructed from experimental data. Our objective here is to find an approximate dynamical system in a neighborhood of the attractor as follows.

A simplex in an *m*-dimensional space is a triangle with m + 1 vertices. Suppose the map is known at each point on a grid. Then there is a unique way to extend the map linearly to the interior of the simplex S whose vertices are grid points. Given a point P in the interior of S, let $\{b_i\}_{i=0}^m$ be its corresponding *barycentric coordinates* (see ref. [36] for an algorithm to compute them). Let $f(v_i)$ be the map at the *i*th vertex. The dynamical system at P is iterated by computing

$$\Phi(P) = \sum_{i=0}^{m} b_i f(v_i).$$
(4)

We apply this method to experimental data by finding a linear approximation of the dynamics at each vertex v_i with the least-squares method described above, using a collection of points in a small ball around v_i . The maps are stored and retrieved using a hashing algorithm similar to that described in ref. [36]. This yields a piecewise linear approximation of the dynamics from a set of experimental data which can be analyzed with the methods that previously were available only to theorists^{#11}.

We illustrate the approach using a time series of 32768 values from the Hénon map with $\alpha = 1.2$, $\beta = 0.3$ using eq. (3) and adding 0.1% noise as described above. The original attractor is shown in fig. 6a. We take a grid of points which are spaced at 1% intervals (this and subsequent distances are expressed as a fraction of the original attractor extent). The time series is embedded in two dimensions, and a linear approximation of the dynàmics is computed at each grid point for which 50 or more attractor points can be collected with a ball of radius 0.03; the set of such grid points is shown in fig. 6b. We take an initial condition near the original attractor and show the first 3000 iterates using eq. (4) in fig. 6c. Although some defects are visible, the attractor produced by the approximate dynamical system looks almost identical to the original one.

One application of simplicial approximations is the location of periodic saddles and the estimation of the largest eigenvalue of the corresponding Jacobian. That is, if x is a periodic point of period p, then we find the eigenvalue of $Df^{p}(x)$ of largest modulus, where $Df^{p}(x)$ refers to the matrix of partial derivatives of the pth iterate of the map f evaluated at x.

Given an initial guess for x, one can apply Newton's method using the maps computed at the grid points and eq. (4) to locate the saddle using the simplicial approximations. Likewise, eq. (3) can be used to locate the corresponding "exact" saddle. Saddle orbits up to period 8 have been computed in this way. In all cases, the saddle point for the simplicial approximation is within 2% of the corresponding saddle point for the Hénon map. Table 1 shows the largest eigenvalues of the saddle orbits. (The columns labeled m = 2and m = 3 refer to the embedding dimension used to reconstruct the attractor.) In most cases, the

^{#11}This approach is less ambitious than that of Crutchfield [8], who attempts to find a single set of nonlinear difference equations that creates the observed attractor.



relative error is only a few percent, and in no case exceeds 25%. (The largest relative error is for the period 8 saddles, where one finds the eigenvalue of the product of 8 Jacobians computed from the least squares.)

This method can be extended to experimental data sets. However, there are relatively stringent requirements on the data that can be handled: the time series must be long enough to trace out many trajectories near the principal unstable saddle orbits, and the noise level must be low. (Presumably, noisy data can be preprocessed using the approach



Fig. 6. (a) Hénon attractor computed from eq. (3) with $\alpha = 1.2$, $\beta = 0.3$. (b) 1% grid on which linear approximations of the dynamics are computed from the available attractor points. (c) Attractor produced by the simplicial approximations.

described in section 4.) The current computer implementation uses a large amount of disk space to store the linear map approximations at the grid points.

We have constructed a simplicial approximation for an attractor obtained from a Belousov-Zhabotinskii chemical reaction [7, 30]. The attractor is reconstructed in three dimensions from a set of 32 768 measurements of bromide ion concentration. The phase portrait is shown in fig. 7a.

Linear approximations of the dynamics are computed at each point of a grid consisting of 50

Table 1 The largest eigenvalues of the Jacobian of the periodic orbits located using the simplicial approximation of the Hénon attractor.

| Period | m = 2 | Exact | <i>m</i> = 3 |
|--------|-------|-------|--------------|
| 1 | 1.793 | 1.695 | 1.757 |
| 2 | 2.178 | 2.199 | 2.183 |
| 4 | 4.226 | 4.329 | 4.051 |
| 6 | 10.38 | 10.70 | 9.626 |
| 6 | 10.38 | 11.32 | 12.12 |
| 8 | 25.80 | 24.88 | 30.25 |
| 8 | 20.02 | 20.60 | 20.38 |
| 8 | 17.70 | 24.32 | 21.70 |

intervals along each coordinate axis for which 50 or more attractor points can be located within an 8% radius of the grid point. This produces a database of 59 550 maps. We observe from graphical evidence that many trajectories approach what appears to be a period-3 saddle in the middle of



Fig. 7. (a) The attractor reconstructed from a time series of bromide ion concentrations in a Belousov-Zhabotinskii chemical reaction. (b) The period-3 saddle orbit.

the attractor. Using initial guesses from some of the trajectories, we apply Newton's method to locate the saddle orbit shown in fig. 7b. Moreover, we obtain estimates of the Jacobian Df of the map evaluated at a point on the saddle orbit. The eigenvalues of Df are estimated as $\lambda_1 = 1.14$, $\lambda_2 = 0.102$, and $\lambda_3 = -1.53$. These quantitative results confirm that the orbit is a saddle since $\lambda_1 >$ $0 > \lambda_3$. (Note that one expects $\lambda_2 = 0$ for a flow generated from a set of differential equations.)

9. Conclusion

Methods for approximating the dynamics of attractors reconstructed from experimental data provide powerful tools. Most of the same procedures that have been so important for theoretical insight, such as Poincaré maps, unstable fixed points and their manifolds, basin boundaries, and the like, are now available to experimenters, at least in cases where the dynamics are low dimensional. There is little doubt that these tools will lead to breakthroughs in the understanding of a wide variety of physical systems. However, considerable effort is needed before we learn which kinds of systems will benefit most from these types of analyses. Significant improvements in technique will certainly extend the applicability of dynamical embedding methods, for example to higherdimensional attractors.

Appendix

In this appendix we outline a possible alternative noise reduction method based on the theory of least squares when all the quantities in the regression are measured with error.

In ordinary least squares, the variables in the problem fall into two classes: the *independent* variables, which are known exactly, and the *dependent* variables, which are observations assumed to be functions of the independent variables. The dependent variables are subject to random errors that are assumed independent and identically distributed (i.i.d.). On an attractor reconstructed from experimental data, we assume that the mapping which takes points in a sufficiently small ball to their images is approximately linear. However, the locations of all the points are subject to small random errors because of the noise. Hence one cannot describe the points as independent variables and their images as dependent variables. The usual leastsquares method produces a biased estimate of the linear map, and this bias does not decrease if more observations are added [16, 23].

The so-called "errors in variables" least-squares methods can be used to handle the latter problem. This approach can be used to obtain both an estimate of the linear map as well as estimates of the "true" values of each of the observations.

At first this appears to be an underdetermined problem: from n pairs of observations one wants to compute the parameters of the functional relation between them as well as estimates of the nactual pairs^{#12}. However, it is possible to solve this problem by making some assumptions about the errors [16, 23].

In our case, we *assume* that the errors in the location of each point and its image are i.i.d. In particular, we let the covariance matrix of the errors in the variables be the identity matrix. This assumption is valid whenever the noise is independent of the dynamics^{#13}.

We illustrate the procedure for the case where we are given a collection of n points (in \mathbb{R}^{m}) and their images. Following Jefferys [21], we form a set of n equations of condition given by

$$\boldsymbol{f}_i(\boldsymbol{x}_i) = \boldsymbol{x}_{n+i} - A\boldsymbol{x}_i - \boldsymbol{b}_i \equiv \boldsymbol{x}_{n+i} - L(\boldsymbol{x}_i), \qquad (5)$$

where x_i is the *i*th point, x_{n+i} is its observed image, A is an $m \times m$ matrix, and **b** is an *m*-vector. The goal is to find estimates of L (i.e., A and

b), together with perturbations
$$\hat{v}$$
, such that

$$\boldsymbol{f}_i(\boldsymbol{x}_i + \hat{\boldsymbol{v}}_i) = (\boldsymbol{x}_{n+1} + \hat{\boldsymbol{v}}_{n+i}) - L(\boldsymbol{x}_i + \hat{\boldsymbol{v}}_i) = 0$$

and such that the quadratic form

$$S_0 = \frac{1}{2} \hat{\boldsymbol{v}}^{\mathrm{i}} \boldsymbol{\sigma}^{-1} \hat{\boldsymbol{v}} \tag{6}$$

is minimized. The superscript t denotes transpose and σ is the covariance matrix of the observations (which we assume is the identity matrix here).

This minimization problem can be solved using Lagrange multipliers (see refs. [21, 22] for a numerical algorithm). The solution gives A and btogether with estimates $x_i + \hat{v}_i$ of the "true" observations. It can be shown [16] under fairly mild hypotheses that the estimates of L and the observations are the best in the class of linear estimators.

One way to approach noise reduction is to extend eq. (5) to include several iterations of the observed points. Given a collection of points in a ball, together with the next p iterates of each point, the method above is used to find a collection of linear maps L_1, L_2, \ldots, L_p approximating the dynamics. The method also finds estimates of the actual observations. In this approach, therefore, the calculation of the maps and the adjustment of the trajectories is done in one step. Moreover, each point and its image exactly satisfy a linear relationship.

Of course, p cannot be too large, because nonlinear effects eventually will become significant when the dynamics are chaotic. On the other hand, eq. (5) provides a natural way to include quadratic or other nonlinear terms.

We have written a computer program to implement this alternative noise reduction algorithm. So far, the results of this approach have not been as good as those from the method described in the main part of the paper, but further refinement should improve them.

Acknowledgements

Dan Lathrop provided invaluable assistance in finding periodic orbits in the Hénon and BZ attractors. We thank Bill Jefferys for useful discus-

 $^{^{\#12}}$ In the statistical literature, the problem is said to be unidentified.

^{#13}Dynamical noise (i.e., each point is perturbed slightly before iterating) yields a covariance matrix which depends on the point. However, as long as the dynamical noise is small, our assumptions about the covariance matrix of the errors should not compromise the accuracy of the method.

sions and computer software for the errors in variables least-squares problem. Andy Fraser, Randy Tagg and Harry Swinney all provided helpful suggestions. This research is supported by the Applied and Computational Mathematics Program of the Defense Advanced Research Projects Agency (DARPA-ACMP) and by the Department of Energy Office of Basic Energy Sciences.

References

- R. Badii, G. Broggi, D. Derighetti, M. Ravani, S. Ciliberto and A. Politi, Phys. Rev. Lett. 60 (1988) 979.
- [2] R.P. Behringer and G. Ahlers, J. Fluid Mech. 125 (1982) 219;

G. Ahlers and R.P. Behringer, Phys. Rev. Lett. 40 (1978) 712.

- [3] J.L. Bentley and J.H. Friedman, ACM Comput. Surv. 11 (1979) 397.
- [4] S.M. Bozic, Digital and Kalman Filtering (Edward Arnold Publishers Ltd., London, 1979).
- [5] A. Brandstäter and H.L. Swinney, Phys. Rev. A 35 (1987) 2207.
- [6] M. Casdagli, Nonlinear prediction of chaotic time series, preprint (December 1987).
- [7] K.C. Coffman, Ph.D. Thesis, University of Texas at Austin (1987).
- [8] J.P. Crutchfield and B. McNamara, Complex Systems 1 (1987) 417;

H.D. Abarbanel, R. Brown and J.B. Kadtke, Prediction and system identification in chaotic nonlinear systems: time series with broadband spectra, preprint (January 1989).

- [9] J.J. Dongarra, C.B. Moler, J.R. Bunch and G.W. Stewart, LIN-PACK User's Guide (Society for Industrial and Applied Mathematics, Philadelphia, 1979).
- [10] J.-P. Eckmann, S.O. Kamphorst, D. Ruelle and S. Ciliberto, Phys. Rev. A 34 (1986) 4971.
- [11] J.-P. Eckmann and D. Ruelle, Rev. Mod. Phys. 57 (1985) 617.
- [12] J.D. Farmer and J.J. Sidorowich, Phys. Rev. Lett. 59 (1987) 845.

- [13] P.R. Fenstermacher, H.L. Swinney and J.P. Gollub, J. Fluid Mech. 94 (1979) 103.
- [14] W. Franceschini and L. Russo, J. Stat. Phys. 25 (1981) 757.
- [15] A. Fraser and H.L. Swinney, Phys. Rev. A 34 (1986) 1134.
- [16] W.A. Fuller, Measurement Error Models (Wiley, New York, 1987).
- [17] E.G. Gwinn and R.M. Westervelt, Phys. Rev. A 33 (1986) 4143.
- [18] S.M. Hammel, J.A. Yorke and C. Grebogi, J. Complexity 3 (1987) 136; Bull. Am. Math. Soc. 19 (1988) 465.
- [19] M. Hénon, Comm. Math. Phys. 50 (1976) 69.
- [20] D. Hirst, Ph.D. Dissertation, University of Texas (December 1987).
- [21] W.H. Jefferys, Astron. J. 85 (1980) 177.
- [22] W.H. Jefferys, Astron. J. 86 (1981) 149.
- [23] M.G. Kendall and A. Stuart, The Advanced Theory of Statistics, Vol. 2 (Griffin, London, 1961) p. 375.
- [24] A. Libchaber, S. Fauve and C. Laroche, Physica D 7 (1983) 73.
- [25] E.N. Lorenz, J. Atmos. Sci. 20 (1963) 130.
- [26] S.W. MacDonald, C. Grebogi, E. Ott and J.A. Yorke, Physica D 17 (1985) 125.
- [27] G. Mayer-Kress, ed., Dimensions and Entropies in Chaotic Systems (Springer, Berlin, 1986), and references therein.
- [28] F. Mitschke, M. Möller and W. Lange, Phys. Rev. A 37 (1988) 4518.
- [29] L.R. Rabiner and B. Gold, Theory and Application of Digital Signal Processing (Prentice-Hall, Englewood Cliffs, NJ, 1975).
- [30] J.-C. Roux, Physica D 7 (1983) 57;
 J.-C. Roux, R.H. Simoyi and H.L. Swinney, Physica D 8 (1983) 257.
- [31] M. Sano and Y. Sawada, Phys. Rev. Lett. 55 (1985) 1082.
- [32] C. Sparrow, The Lorenz Equations: Bifurcations, Chaos, and Strange Attractors (Springer, New York, 1982).
- [33] R. Tagg, private communication.
- [34] F. Takens, in: Dynamical Systems and Turbulence, Springer Lecture Notes in Mathematics, Vol. 898, eds. D.A. Rand and L.-S. Young (Springer, Berlin, 1980) p. 366.
- [35] B.H. Tongue, Physica D 28 (1987) 401.
- [36] F. Varosi, C. Grebogi and J.A. Yorke, Phys. Lett. A 124 (1987) 59.
- [37] A. Wolf, J.B. Swift, H.L. Swinney and J.A. Vastano, Physica D 16 (1985) 285.