Physica D 64 (1993) 431–454
North-Holland

PHYSICA D

Coarse-grained embeddings of time series: random walks, Gaussian random processes, and deterministic chaos

Daniel T. Kaplan and Leon Glass
Department of Physiology, McGill University, 3655 Drummond St., Montreal, Quebec, Canada H3G 1Y6

Received 25 April 1992
Revised manuscript received 1 September 1992
Accepted 20 September 1992
Communicated by E. Jen

A new method for studying time series is described based on a statistic indicating the degree to which trajectories passing through a small region of an embedding space are parallel. The method is particularly suited to time series with a significant correlation time. Analytic results are presented for Brownian motion and Gaussian random processes. These are generally different from the results for chaotic systems, allowing a test for deterministic dynamics in a time series. A variety of examples are presented of the application of the method to low- and high-dimensional systems.

1. Introduction

Inferring the dynamical mechanism of a system from a measured time series is a general problem in science. For complicated, aperiodic data, a decision needs to be made: is irregularity in the data to be ascribed to noise or to deterministic chaotic dynamics? Fig. 1 shows a sampler of time series, some of which arise from stochastic processes, some from deterministic chaotic equations, and some from experimental or natural systems. It is not always easy, by eye, to distinguish the deterministic time series from their stochastic counterparts.

There exists an extensive literature describing techniques for relating physical systems to linear filters and for interpreting time series as linearly filtered random noise. Even for systems known to be nonlinear, this linear "systems analysis" approach often provides technologically satisfactory results; it is the basis for most of engineering control theory.

It is only in the past 15 years or so that there has been a general awareness of the possibility that irregular-looking fluctuations may be caused by deterministic chaotic dynamics [1]. A number of tools have been developed for estimating from time series characteristics of chaotic systems such as dimension [2,3] or the Lyapunov numbers that measure the exponential divergence of nearby chaotic trajectories [4–6]. Initially, it was hoped that these characteristics could be used to distinguish between deterministic chaotic systems and stochastic systems such as those dealt with in linear systems analysis. More recently, emphasis has shifted to testing to see if a time series is generated by a deterministic process and the construction of deterministic nonlinear models from time series [7–9]. One test for determinism versus randomness consists of evaluating whether a deterministic nonlinear model or a linear model allows better predictions to be made of future values of the time series given past values. Another test is whether the error in

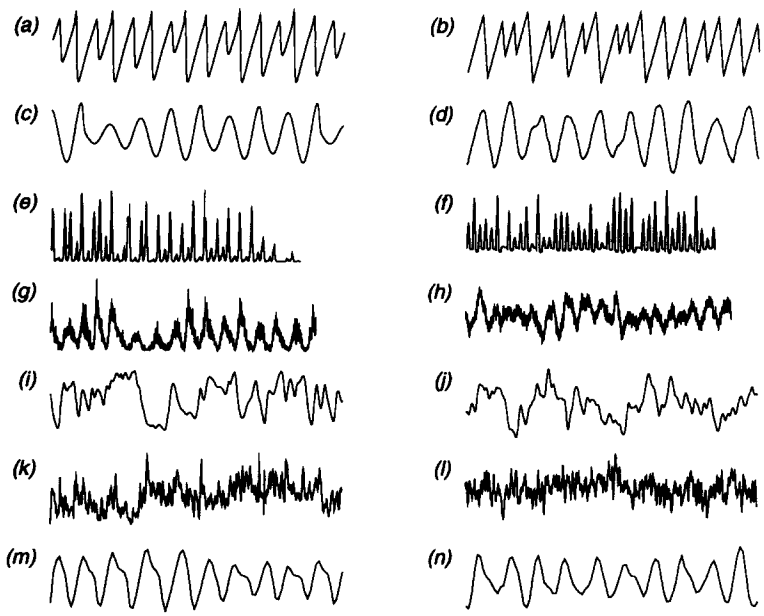

Fig. 1. Segments of various time series that are studied in section 9. The time series in the left column are from experimental, natural, or mathematically deterministic systems; the time series on the right are synthesized stochastic time series. (a) Bromide ion potential from the Belousov–Zhabotinskii reaction, (b) a time series constructed by connecting randomly chosen peaks in (a) with line segments of constant slope. Data provided by H. Swinney. (c) The x component of the Rossler system and (d) a Gaussian random process with the same autocovariance function. (e) Numbers of measles cases reported each month in New York City, and (f) a yearly peak with random amplitude. The whole time series is shown. Data provided by G. Sugihara through the Santa Fe Institute. (g) The Wolper sunspot numbers and (h) a Gaussian random process with the same autocovariance function. The whole time series is shown. (i) A time series from a delay-differential equation model of blood cell production (eq. (36)) and (j) a Gaussian random process with the same autocovariance function. (k) Intervals between heart beats over a 1/2 hour period and (l) a Gaussian random processes with the same autocovariance function. The whole time series is shown. (m) Experimental recording from a system of 8 coupled nonlinear oscillators and (n) a Gaussian random process with the same autocovariance function. Data provided by P. Linsay.

the predictions increases with increasing prediction time in the manner expected for chaotic systems [9].

An alternative approach involves investigating whether the trajectory constructed by embedding a time series is consistent with a dynamical rule that is a single-valued function of position in phase space. In section 2 we present in detail a new method for assessing the possible deterministic origins of a time series measured from a continuous-time dynamical system. The method, introduced in ref. [10], involves embedding a time series in a phase space and examining the extent to which the tangents to the trajectory in a finite sized region of the phase space point in similar directions. The local statistic used to quantify this alignment is denoted by the symbol $|V_j|$, where the index j refers to the region of the phase space. In section 3 we discuss the application of the method to deterministic systems and some of the reasons why, even for a deterministic system, the tangents to the trajectory at nearby points in the phase space may not point in identical directions.

In section 4 we derive the expected values of $|V_j|$ for Brownian motion, i.e., a random walk in m-dimensional space. A random walk is not a very good model for many time series; a more generally applicable model is a Gaussian random process (GRP), which results from passing white noise through a linear filter. In section 5 we show that the expected value for $|V_j|$ for a GRP de-

pends only on the autocovariance function of the time series, and we derive an approximate expression for $|V_j|$ as a function of position in phase space. One result from this analysis is that for appropriately chosen embedding lags τ, a Gaussian random process often appears the same as Brownian motion in terms of $|V_j|$.

The statistic $|V_j|$ describes the alignment of the trajectory tangent vectors in region j of phase space. In section 6 we describe two ways of averaging $|V_j|$ over all of phase space, and a method to derive confidence intervals on the resulting statistics. One of these statistics, $\bar{\Lambda}$, takes on the value 1 for a perfectly deterministic system and 0 for Brownian motion or an appropriately embedded GRP.

A useful test for the possibility of deterministic chaotic dynamics is to compare $\bar{\Lambda}$ for the time series under analysis with $\bar{\Lambda}$ from a GRP time series with a similar autocovariance function. A more general type of model than a GRP is a GRP passed through a static nonlinear transformation (e.g., squaring each term in the time series). Section 7 discusses the influence of static nonlinear transformations on $\bar{\Lambda}$. In section 8, we discuss briefly the ability of $\bar{\Lambda}$ to discern deterministic nonlinear dynamics in the presence of noise.

Finally, in section 9, we present as examples of the application of the method an analysis of the time series shown in fig. 1.

2. The method: coarse-grained flow averages

Consider a time series $x(t)$ generated from a deterministic dynamical system

$$\frac{dz}{dt} = f(z)\,; \quad x(t) = g(z(t))\,, \tag{1}$$

where the dynamical function $f(\cdot)$: $\mathcal{R}^p \to \mathcal{R}^p$ and the measurement function $g(\cdot)$: $\mathcal{R}^p \to \mathcal{R}$ are smooth and single-valued. As is well known, $x(t)$ in $\mathcal{R}$ can be put in the form of a trajectory $\boldsymbol{x}(t)$ in $\mathcal{R}^m$ by the simple device of time-lag embedding

$$\boldsymbol{x}^{\mathrm{T}}(t) = (x(t), x(t-\tau), \ldots, x(t-(m-1)\tau))\,, \tag{2}$$

where $\boldsymbol{x}^{\mathrm{T}}$ is the transpose of the vector $\boldsymbol{x}$. It has been shown [11] that for $m \geq 2p+1$ there is a smooth function $\mathcal{R}^m \to \mathcal{R}^p$ that transforms the trajectory $\boldsymbol{x}(t)$ into $z(t)$ so that nearby points in $\mathcal{R}^m$ are also nearby in $\mathcal{R}^p$. (A method for assessing whether this is the case when embedding a time series has been proposed by Kennel et al. [12].) For the present method, we are concerned with the evidence that $\boldsymbol{x}(t)$ can provide about the single valuedness of $f(\cdot)$, and hence the determinism of the system $dz/dt = f(z)$.

The derivative dz/dt is a vector tangent to the trajectory $z(t)$. Insofar as $f(\cdot)$ is a smooth function of z, we expect that two tangent vectors $dz_1/dt = f(z_1)$ and $dz_2/dt = f(z_2)$ will point in similar directions when z_1 is close to z_2. Since the map from z to $\boldsymbol{x}$ is smooth, the tangent vectors $d\boldsymbol{x}_1/dt$ and $d\boldsymbol{x}_2/dt$ should also point in similar directions.

Now, assume that a time series $\boldsymbol{x}(t)$ is measured from a stochastic system. For the stochastic system, the tangent to the trajectory $d\boldsymbol{x}/dt$ is not a single valued function of position $\boldsymbol{x}$. For a particle undergoing random Brownian motion, for instance, the vector $z(t)$ refers to the position in space, while dz/dt is a random variable with a uniform distribution of directions.

The method we present here is based on collecting the tangent vectors $d\boldsymbol{x}(t)/dt$ in a small region $\mathcal{R}^m$, and assessing the degree to which the vectors point in similar directions or point randomly.

The first step in the method is to embed the time series $x(t)$ in $\mathcal{R}^m$ using eq. (2). Next, the space $\mathcal{R}^m$ is divided into "boxes", non-overlapping m-dimensional hypercubes with edge length ϵ. For finite ϵ, a finite number of boxes covers the entire trajectory, and the set of covering boxes can be indexed with an integer j.

The trajectory passes through box j one or more times. Index the individual passes through box j with index k, where a pass consists of the

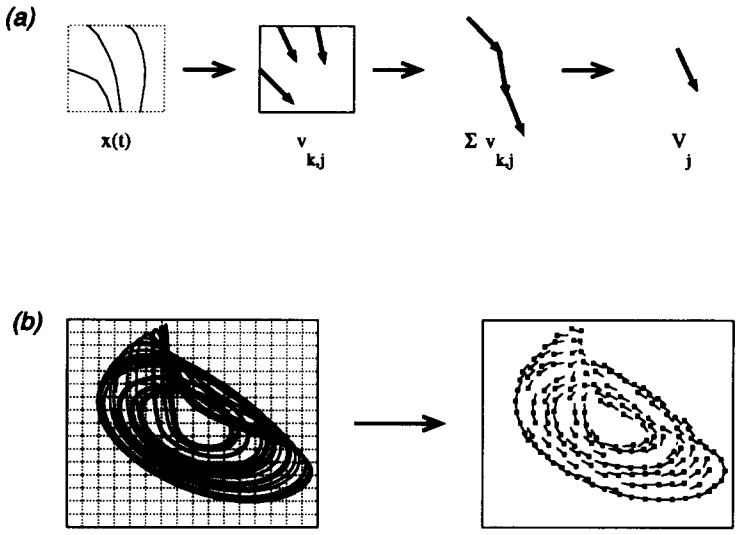

Fig. 2. An illustration of the method in a 2-dimensional embedding. (a) The steps in the calculation of V_j for one box j of the coarse grained phase space. There are three passes of the trajectory through this box. The corresponding unit length trajectory vectors $v_{k,j}$ $k = 1, 2, 3$ are averaged to produce the resultant vector V_j. (b) An embedding of the Rossler x-component (see section 9.1.1) and the V_j for all boxes of the coarse grained phase space.

segment of the trajectory from the point in $\mathcal{R}^m$ where the trajectory enters the box j to the point where it next leaves the box j. Construct a vector of unit length, $v_{k,j}$ whose direction is given by the displacement between the point where pass k enters box j to the point where the trajectory next exits from the box. The *trajectory vector* $v_{k,j}$ points in the direction of a secant to the trajectory, i.e., the direction of a finite difference approximation to the tangent vector $\mathrm{d}x/\mathrm{d}t$.

The set of trajectory vectors in box j can be summarized by their mean, the *resultant vector*

$$V_j \equiv \sum_k v_{k,j}/n_j \,, \tag{3}$$

where n_j is the number of trajectory vectors in box j. The transformation from the embedded time series to V_j is shown in fig. 2.

The calculation of the finite-difference tangent to the trajectory within a box implicitly assumes that the embedded time series contains time-consecutive points that fall into the same box. For continuous-time systems, this can be accomplished by sampling the time series at a high rate. The method is not applicable to discrete-time maps such as the logistic or Henon maps where consecutive points may be widely separated in the phase space. In cases where the time series from a continuous-time system contains small amounts of high frequency noise, it may be appropriate to disregard in eq. (3) those resultant vectors whose box passage time b is below a threshold value, or to low-pass filter the time series. (See sections 8 and 9.1.4.)

3. Deterministic systems

As argued above, for a deterministic system the trajectory vectors in a small enough region of the phase space point in the same direction. This can occur only when the embedding dimension m of the time series is large enough to resolve the system dynamics; a necessary condition is that $m \geq p$, a sufficient condition is $m \geq 2p + 1$. When m is too small to represent the dynamics (e.g., $m < p$), different parts of the trajectory

$z(t)$ in $\mathcal{R}^p$ are mapped by the embedding to the same region of $\mathcal{R}^m$; the trajectory vectors will not be aligned.

When m is large enough, in the limit of vanishingly small box size the trajectory vectors are exactly aligned and

$$\lim_{\epsilon \to 0} |V_j| = 1 . \tag{4}$$

In practice, this limit cannot be realized for several reasons.

When a box has finite size, the flow field in the phase space may vary over the domain of the box. For example, trajectories going through two nearby points in phase space may be oriented slightly towards or away from one another, corresponding to the action of positive or negative Lyapunov exponents. The large scale curvature of attractors is not always of consequence in this regard, since the trajectory may follow a similar path on each pass through the box – for instance, for periodic systems the identical path is followed on every pass despite curvature of the trajectory.

When time series are short, boxes may need to be quite large in order to contain more than one pass of the trajectory through several boxes, particularly when the dimension of the system is high. (See section 9.2.) Big boxes may also be needed to help reduce the effects of measurement noise. In both cases, the spatial variation of the vector field in the box causes $|V_j| < 1$.

The method of eq. (2) of reconstructing dynamics using delays introduces another cause for $|V_j| < 1$. Chaotic systems have at least one positive Lyapunov exponent. This causes points that are initially close in the original phase space $\mathcal{R}^p$ to be pulled apart in time. Over the time $(m-1)\tau$ involved in the delay embedding, different components of the embedded vector $x(t)$ can become causally disconnected with one another. The resulting decrease in $|V_j|$ with increasing τ can be seen in the examples of sections 9.1.1, 9.1.2, 9.2.1. The precise form of the fall-off in those examples is not completely understood.

4. Brownian motion – random flights in m dimensions

An important limiting case is Brownian motion, corresponding to a random walk in phase space. In this case, the $\boldsymbol{v}_{k,j}$ are random variables whose direction is uniformly distributed in $\mathcal{R}^m$. In this section we analyze the expected value of the resultant vector assuming that the orientations in each pass through any given box are random.

The history of this problem goes back to Rayleigh and Chandrasekhar who studied random flights in m dimensions [13,14]. The general problem can be formulated as follows. Consider a flight consisting of n steps of unit length in m dimensions where the angle from each step to the next is chosen randomly. The average displacement per step is $\boldsymbol{R}_n^m$.

The surface of the d-dimensional hypersphere of radius a is defined to be the locus of points satisfying $\Sigma_{i=1}^m x_i^2 = a^2$. The extent of the surface of the 1-sphere is 2 and of the 2-sphere is $2\pi a$. From fig. 3 we see that the extent of the surface of the m-sphere of radius a, denoted $|S^m(a)|$, can be determined recursively from the integral

$$|S^m(a)| = \int_0^{\pi} \mathrm{d}\theta \, a |S^{m-1}(a \sin\theta)| , \quad m \geq 2 , \tag{5}$$

to obtain

$$|S^m(a)| = \frac{2\pi^{m/2} a^{m-1}}{\Gamma(m/2)} , \tag{6}$$

which is valid for all m.

Now consider the summation of two vectors of unit length in m dimensions. Let the angle between the two vectors be θ and consider the average length integrated over surface of the $(m-1)$-sphere. Since the distance between two vectors of unit length is given by $(2-2\cos\theta)^{1/2}$ (see fig. 3) we find that the mean length of two vectors in m dimensions, $\bar{R}_2^m$ is

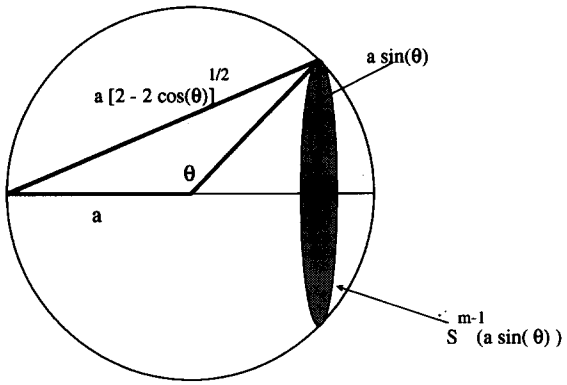

Fig. 3. Schematic diagram showing the geometry underlying eq. (5).

$$\bar{R}_2^m = \frac{1}{2|S^m(1)|}\int_0^\pi \mathrm{d}\theta \times (2-2\cos\theta)^{1/2}|S^{m-1}(\sin\theta)|\,. \tag{7}$$

Carrying out the integral leads to the result

$$\bar{R}_2^m = \frac{2^{m-2}\Gamma(m/2)\Gamma(m/2)}{\pi^{1/2}\Gamma(m-1/2)}\,. \tag{8}$$

The average length of 2 unit vectors for several values of m is given in table 1. Using Stirling's approximation, we find $\lim_{m\to\infty} \bar{R}_2^m = 1/\sqrt{2}$, a result found previously by Hammersley and Lord using other techniques [15,16].

We now consider the computation of $\bar{R}_n^m$ when n is large. Rayleigh [13] and Chandrasekhar [14] computed the density of the mean displacement for $m = 2$, 3 and Feller [17], analyzed the problem for any m. The probability density, W_n^m, that $\bar{R}_n^m$ lies between R and $R + \mathrm{d}R$ is

$$W_n^m = \frac{1}{n}\,|S^m(R)|\,\frac{\exp(-R^2/2s)}{(2\pi s)^{m/2}}\,, \tag{9}$$

where $s = n/m$. From this we compute the mean displacement per step

$$\bar{R}_n^m = c_m/\sqrt{n}\,, \tag{10}$$

where

$$c_m = \left(\frac{2}{m}\right)^{1/2}\frac{\Gamma[(m+1)/2]}{\Gamma(m/2)}\,. \tag{11}$$

Several values of c_m are given in table 1. Although, the derivation of eq. 10 is only valid for the asymptotic limit when n is large, it agrees within a few percent of the analytical values computed in eq. 8 for $n = 2$, and with computer simulations (not shown) for $n \geq 2$.

5. Gaussian random processes

As a starting point in studying stochastic systems other than Brownian motion, we consider Gaussian random processes (GRP) [18,19]. GRPs form the basis for much of the technology surrounding system estimation and control [20]. The comparison of a time series to a similar GRP has been proposed as a bootstrapping test for the significance of dimension and Lyapunov

Table 1
Vector statistics for random flights.

	m				
	2	3	4	5	∞
$\lvert S^d(1)\rvert$	2π	4π	$2\pi^2$	$\frac{8}{3}\pi^2$	...
$\bar{R}_2^d$(eq. 8, exact)	$\frac{2}{\pi} \approx 0.637$	$\frac{2}{3} \approx 0.667$	$\frac{32}{15\pi} \approx 0.679$	$\frac{24}{35} \approx 0.686$	$\frac{1}{\sqrt{2}} \approx 0.707$
$\bar{R}_2^d$ (eq. 10, approx.)	$\sqrt{\frac{\pi}{8}} \approx 0.627$	$\frac{4}{\sqrt{12\pi}} \approx 0.651$	$\frac{3\sqrt{\pi}}{8} \approx 0.665$	$\frac{8}{3}\sqrt{\frac{1}{5\pi}} \approx 0.673$	$\frac{1}{\sqrt{2}} \approx 0.707$
c_m (eq. 11)	$\frac{\sqrt{\pi}}{2} \approx 0.886$	$\frac{4}{\sqrt{6\pi}} \approx 0.921$	$\frac{3\sqrt{\pi}}{\sqrt{32}} \approx 0.940$	$\frac{8}{3}\sqrt{\frac{2}{5\pi}} \approx 0.952$	1

exponent estimates [21]. Some examples of GRPs are given in figs. 1d,h,j,l and n. After a brief introduction to GRPs, we derive the properties of coarse-grained embeddings of GRPs in sections 5.1 and 5.2, and discuss ways of calculating $|V_j|$ in section 5.3.

A GRP can be represented by a randomly forced linear dynamical system such as

$$\frac{\mathrm{d}z}{\mathrm{d}t} = \mathbf{A}\cdot z + \mathbf{B}\cdot w\,, \qquad x = c\cdot z \tag{12}$$

where $\mathbf{A}$ and $\mathbf{B}$ are matrices of constants, c is a constant vector, and w is a vector of independently and identically distributed Gaussian white noise (GWN) processes. When the real parts of the eigenvalues of $\mathbf{A}$ are negative, the homogeneous system is stable and the stochastic system is stationary [18,22]. Such systems are often studied in terms of their autocovariance functions.

When $x(t)$ is a stationary GRP the probability density function (pdf) for x at time t is written as a Gaussian pdf,

$$p_t(x) = \frac{1}{\sqrt{2\pi\Psi(0)}}\exp[-x^2/2\Psi(0)]\,, \tag{13}$$

where $\Psi(0) = \langle x(t)\,x(t)\rangle$. (Here and in the following it is assumed that the mean of a measured signal, in this case $\langle x(t)\rangle$, is zero.)

For a GRP, the multivariate pdf for any vector that is a linear function of $x(t)$ can be written as a multivariate Gaussian pdf [18,23]. In particular, the pdf for the embedded vector

$$x^{\mathrm{T}} = (x(t), x(t-\tau), \ldots, x(t-(m-1)\tau)) \tag{14}$$

is

$$p_{t,t-\tau,\ldots,t-(m-1)\tau}(\dot{x}) = \frac{1}{(2\pi)^{m/2}|\mathbf{X}|^{1/2}}\exp(-\tfrac{1}{2}x^{\mathrm{T}}\cdot\mathbf{X}^{-1}\cdot x)\,, \tag{15}$$

where $\mathbf{X}$ is the covariance matrix which is the outer product $\langle xx^{\mathrm{T}}\rangle$ having terms

$$\mathbf{X}_{i,j} = \langle x(t-i\tau)\,x(t-j\tau)\rangle = \Psi((i-j)\tau)\,. \tag{16}$$

When the off-diagonal terms of $\mathbf{X}$ are zero, the components of x are uncorrelated with one another. For a GRP, this means that there is no statistical dependence between the components of x. For a system that is not a GRP, e.g., a nonlinear chaotic system, it is possible to have statistical dependence between the components of x even if the off-diagonal terms of $\mathbf{X}$ are zero. Thus, eq. (15) cannot be used to describe probability densities for chaotic systems.

5.1. Flows of Gaussian random processes

The matrix $\mathbf{X}$ describes the probability density of the cloud of points that results from the embedding. We are interested here in the structure of the trajectory vectors in the embedding space: whether they point in random directions independently of position in the embedding space, or whether they tend to point in similar directions as a function of position. Consider the finite difference approximation to the tangent to the embedded trajectory

$$\begin{aligned}\Delta x(t)^{\mathrm{T}} &\equiv (\Delta x(t), \Delta x(t-\tau), \Delta x(t-2\tau)\,,\\ &\quad \ldots, \Delta x(t-(m-1)\tau)\\ &= (x(t+b)-x(t), x(t-\tau+b)-x(t-\tau)\,,\\ &\quad \ldots, x(t-(m-1)\tau+b)\\ &\quad - x(t-(m-1)\tau))\,,\end{aligned} \tag{17}$$

where b represents the *time* it takes the trajectory to pass through a typical box of edge length ϵ. Although $\Delta x(t)$ is not normalized to unit length as are the trajectory vectors, it does point in the same direction as some trajectory vector for some choice of box gridding.

Let box j be centered on position x. The resultant vector V_j is, roughly, the average of Δx over a small region surrounding x. To examine

the coarse-grained spatial average V_j, we need to consider the pdf of the $2m$-dimensional vector

$$\xi^{\mathrm{T}} = (x^{\mathrm{T}}, \Delta x^{\mathrm{T}}) . \tag{18}$$

Since ξ is a linear transformation of x, the pdf for ξ is in the form of eq. (15) with covariance matrix $\mathbf{M} = \langle \xi \xi^{\mathrm{T}} \rangle$. This can be written

$$\mathbf{M} = \begin{pmatrix} \mathbf{X} & \mathbf{N} \\ \mathbf{N}^{\mathrm{T}} & \mathbf{W} \end{pmatrix} , \tag{19}$$

where $\mathbf{X}$, $\mathbf{N}$, and $\mathbf{W}$ are $m \times m$ matrices and $\mathbf{X}$ is given in eq. (16). The elements of these matrices are:

$$\begin{aligned} \mathbf{N}_{i,k} &= \langle \Delta x(t - i\tau)\, x(t - k\tau) \rangle \\ &= \Psi((i-k)\tau - b) - \Psi((i-k)\tau) , \end{aligned} \tag{20}$$

which describes the correlations between the ith component of the trajectory vector and the kth component of position;

$$\begin{aligned} \mathbf{W}_{i,k} &= \langle \Delta x(t - i\tau)\, \Delta x(t - k\tau) \rangle \\ &= -\Psi((i-k)\tau + b) + 2\Psi((i-k)\tau) \\ &\quad - \Psi((i-k)\tau - b) , \end{aligned} \tag{21}$$

which describes the correlations ith and kth components of the trajectory vector; and the $\mathbf{X}_{i,k} = \Psi((i-k)\tau)$ which describes correlations between the ith and kth components of the position.

When the off-diagonal terms involving Δx are zero, there are no correlations between the components of the trajectory vector and the position in phase space. Since the pdf of the trajectory of a GRP is fully described by the covariance matrix $\mathbf{M}$, the trajectory vectors are equally likely to point in any direction and are independent of position, i.e., $p(\Delta x|x) = p(\Delta x)\; p(x)/p(x) = p(\Delta x)$. This case corresponds to a random walk, leading to

$$|V_j| = \bar{R}^m_{n_j} . \tag{22}$$

When the off-diagonal elements of $\mathbf{M}$ are non-zero, we need to examine the conditional probability

$$p(\Delta x|x) = p(x, \Delta x)/p(x) . \tag{23}$$

Calculating $p(x) = \int_{-\infty}^{\infty} p(x, \Delta x)\, \mathrm{d}\Delta x$ and writing

$$\mathbf{M}^{-1} = \begin{pmatrix} \mathbf{X} & \mathbf{N} \\ \mathbf{N}^{\mathrm{T}} & \mathbf{W} \end{pmatrix}^{-1} = 2 \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^{\mathrm{T}} & \mathbf{C} \end{pmatrix} , \tag{24}$$

we find

$$\begin{aligned} p(\Delta x|x) = \frac{|2\mathbf{C}|}{(2\pi)^{m/2}} \exp[&-(\Delta x + \mathbf{H}^{\mathrm{T}} \cdot x) \cdot \mathbf{C} \\ &\cdot (\Delta x + \mathbf{H} \cdot x)] , \end{aligned} \tag{25}$$

where $\mathbf{H} = \mathbf{C}^{-1} \cdot \mathbf{B}^{\mathrm{T}}$. The conditional pdf $p(\Delta x|x)$ is itself a Gaussian probability density with covariance matrix $\mathbf{C}^{-1}/2$ and a mean that depends on position x in phase space

$$\langle \Delta x \rangle = -\mathbf{H} \cdot x = -\mathbf{C}^{-1} \cdot \mathbf{B}^{\mathrm{T}} \cdot x . \tag{26}$$

Each trajectory vector Δx consists of a random component governed by the covariance matrix $\mathbf{C}^{-1}/2$ and a "deterministic" component $-\mathbf{H} \cdot x$ which is a function of position x in phase space.

5.2. *A universal form of V_j for Gaussian random processes*

We can approximate the expectation value of $|V_j|$ as it depends on $\mathbf{B}$ by assuming that the off-diagonal elements of $\mathbf{C}$ are negligible. When $\mathbf{C}$ is diagonal, the directions of the random components are uniformly distributed: we write the mean squared length of the random component as

$$\nu^2 = \mathrm{trace}(\mathbf{C}^{-1}/2) . \tag{27}$$

In a given small region of phase space, write the kth velocity vector as $\Delta x_k = d + w_k$, i.e., the sum of the deterministic part $d = -\mathbf{H} \cdot x$ that depends on position in phase space and a random part w_k whose probability density is independent of posi-

tion in phase space. The sum of n_j of these vectors is $n_j\boldsymbol{d} + \Sigma_{k=1}^{n_j} \boldsymbol{w}_i$. Since $\Sigma_{k=1}^{n_j} \boldsymbol{w}_i$ points randomly with respect to $\boldsymbol{d}$, we can make the approximation that the square modulus of the sum of vectors is $|\Sigma_{k=1}^{n_j}\Delta\boldsymbol{x}|^2 = |n_j\boldsymbol{d}|^2 + |\Sigma_{k=1}^{n_j}\boldsymbol{w}_i|^2$. The expectation value of this is

$$n_j^2|\boldsymbol{d}|^2 + n_j^2\nu^2(\bar{R}_{n_j}^m)^2 , \tag{28}$$

where we use the fact that the expectation value of the length of the sum of n_j randomly oriented unit length vectors in an m-dimensional space is $n\bar{R}_{n_j}^m$. For each of the vectors $|\Delta\boldsymbol{x}_k|^2 \approx |\boldsymbol{d}|^2 + |\boldsymbol{w}_k|^2$, with expectation value $|\boldsymbol{d}|^2 + |\nu|^2$. Recalling the definition of $\boldsymbol{V}_j$ in terms of the velocity vectors

$$\boldsymbol{V}_j = \sum_{k=1}^{n_j} \left(\frac{\Delta\boldsymbol{x}_k}{|\Delta\boldsymbol{x}_k|}\right) \approx \frac{1}{\langle|\Delta\boldsymbol{x}_k|\rangle_k} \sum_{k=1}^{n_j} \Delta\boldsymbol{x}_k , \tag{29}$$

gives

$$|\boldsymbol{V}_j|^2 \approx \frac{(|\boldsymbol{d}|^2 + \nu^2(\bar{R}_{n_j}^m)^2)}{(|\boldsymbol{d}|^2 + \nu^2)} . \tag{30}$$

Writing this in terms of the elements of the covariance matrix $\mathbf{M}$ we have for the case of a GRP

$$|\boldsymbol{V}_j| \approx \left(\frac{\boldsymbol{x}\cdot\mathbf{H}^{\mathrm{T}}\cdot\mathbf{H}\cdot\boldsymbol{x} + \nu^2(\bar{R}_n^m)^2}{\boldsymbol{x}\cdot\mathbf{H}^{\mathrm{T}}\cdot\mathbf{H}\cdot\boldsymbol{x} + \nu^2}\right)^{1/2} . \tag{31}$$

The expectation value of the direction of $\boldsymbol{V}_j$ is $\boldsymbol{d} = -\mathbf{H}\cdot\boldsymbol{x}$.

Eq. (31) represents the key result for the analysis of a GRP. The elements of $\mathbf{H}$ and ν are derived from the autocovariance function $\Psi(T)$. The following features of coarse-grained flow averages are thus universal for GRPs:

(i) The expectation value $\langle\boldsymbol{V}_j\rangle$ depends only on $\Psi(T)$, the embedding parameters τ and m, and the gridding size ϵ.

(ii) $|\boldsymbol{V}_j|$ is constant on ellipsoidal surfaces of constant $\boldsymbol{x}\cdot\mathbf{H}^{\mathrm{T}}\mathbf{H}\cdot\boldsymbol{x}$. (The quadratic equation $\boldsymbol{x}\cdot\mathbf{H}^{\mathrm{T}}\cdot\mathbf{H}\cdot\boldsymbol{x} = a \geq 0$ describes an ellipsoid, because the matrix $\mathbf{H}^{\mathrm{T}}\cdot\mathbf{H}$ has nonnegative eigenvalues.)

(iii) As $\boldsymbol{x}$ increases from the origin, $|\boldsymbol{V}_j|$ also increases, saturating at $|\boldsymbol{V}_j| = 1$ for very large $\boldsymbol{x}$. However, for large $\boldsymbol{x}$, $p(\boldsymbol{x}) \to 0$, so the probability of actually finding $|\boldsymbol{V}_j| \approx 1$ may be small in a finite time series depending on the respective sizes of $\boldsymbol{d}$ and ν.

(iv) $|\boldsymbol{V}_j| \approx \bar{R}_n^m$ at the origin, $\boldsymbol{x} = \mathbf{0}$. When $\boldsymbol{H} = \mathbf{0}$ (corresponding to the case of off-diagonal elements of $\mathbf{M}$ being zero), we recover the result $|\boldsymbol{V}_j| = \bar{R}_{n_j}^m$.

5.3. Computation of $\boldsymbol{V}_j$ for Gaussian random processes

To generate $\boldsymbol{V}_j$ for a GRP at least three different procedures could be used.

(i) Given $\Psi(T)$ and a desired embedding lag τ and coarse graining b, calculate $\mathbf{H}$ and ν, and use eq. (31) to generate $\boldsymbol{V}_j$. (An appropriate n_j for any $\boldsymbol{x}$ can be calculated from the probability density of eqs. (15) and (16).)

(ii) Calculate $\mathbf{M}$ and generate vectors ξ using a Monte Carlo method. $\boldsymbol{V}_j$ is then calculated directly from these randomly generated vectors. This avoids the approximations that lead to eq. 31, however, it still involves the approximation that b is the same for all boxes. (See fig. 4b.)

(iii) Generate a GRP with the given $\Psi(T)$. In a typical case, one has a time series $x(t)$ and wants to generate an instance of a GRP with the same $\Psi(T)$ as $x(t)$. An operationally effective way of doing this [24,21] is to take the Fourier transform of $x(t)$, randomize the phases in $[0, 2\pi)$, and then take the inverse Fourier transform. This procedure leaves the power spectrum unaltered, hence preserving $\Psi(T)$ which is the inverse Fourier transform of the power spectrum. (See fig. 4c.)

Eqs. (16)–(31) show that for a GRP, given τ and b, the expectation value of $\boldsymbol{V}_j$ depends only on $\Psi(T)$. In this sense, the properties of an embedded GRP of a given $\Psi(T)$ are universal –

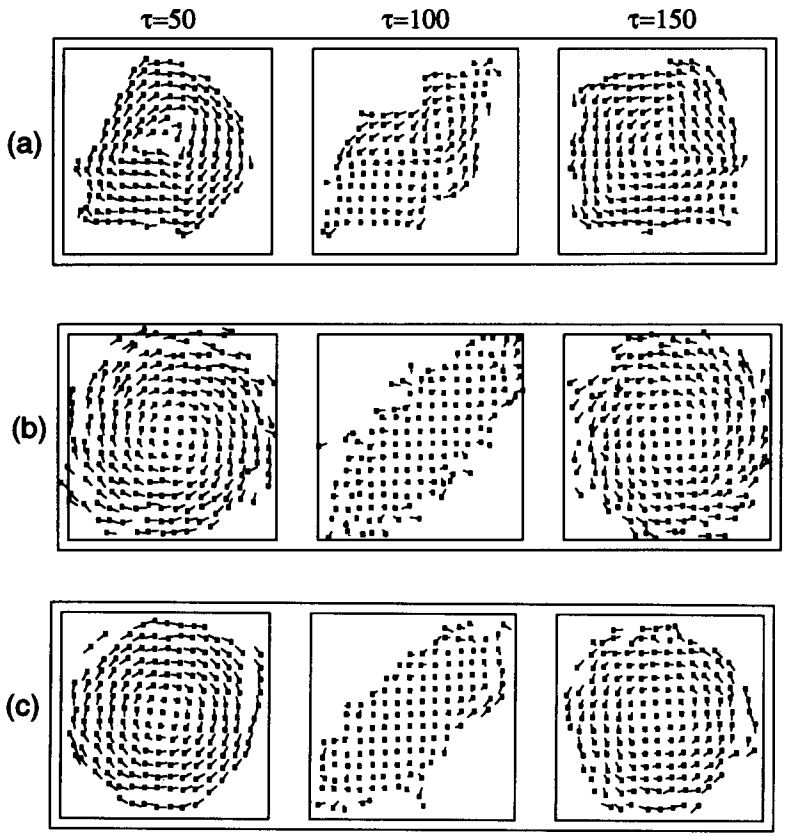

Fig. 4. V_j at $\tau = 50$, 100, 150 for (a) the time series of fig. 1m; (b) V_j for a GRP derived from the covariance matrix **M** using a Monte Carlo method; and (c) V_j calculated directly from a phase-randomized time series. See fig. 15 for $\Psi(\tau)$. A box size of $\epsilon = 1/16$ (scaled to the range of the time series in 1m) has been used. Note the ellipsoidal symmetry in (b) and (c) corresponding to the linear nature of a GRP, while this symmetry is violated in (a). (The arrow length in the figure has been scaled so that when $|V_j| = 1$, the arrow has length ϵ.)

the particular set of random Fourier phases used to construct the GRP are not important. Clearly, however, there are some sets of phases for which the V_j are substantially different than those indicated by eq. (31). For example, the phases implicit in the construction of the deterministic time series $x(t)$ in fig. 4 do not, for the given $\Psi(T)$, produce a GRP. However, these same phases, when applied to another $\Psi(T)$ do produce a GRP time series – the "randomness" of a set of phases depends on the context $\Psi(T)$ in which it is applied. Our experience is that the phases produced by a standard random number generator [25] are in practice random in all contexts.

6. Test statistics on $|V_j|$

A picture of the spatial distribution of V_j of the form of figs. 2 or 4 is useful for visualizing the flow field in a 2-dimensional embedding, but does not allow the ready comparison of V_j for different time series, particularly in higher embedding dimensions. To allow such comparisons to be made, a statistical test is needed.

In many statistical tests, it is appropriate to specify a *null hypothesis* that may or may not be rejected on the basis of the available data. Two possible null hypotheses are suggested by the analysis of sections 4 and 5: the dynamics represented by the time series correspond to Brown-

ian motion; or the dynamics correspond to a Gaussian random process. The $|V_j|$ in both these cases are related to $\bar{R}_n^m$, so in constructing an average of $|V_j|$ it is reasonable to do so by reference to $\bar{R}_n^m$.

We can test whether the dynamics of the embedded time series are distinct from Brownian motion by constructing the mean value of $|V_j|$ for all the boxes that have n elements

$$\bar{L}_n^m \equiv \langle |V_j| \rangle_{j \text{ such that } n_j = n} . \tag{32}$$

The uncertainty in this estimate of $\bar{L}_n^m$ can itself be estimated by the standard error of the mean

$$\sigma_n^m \equiv \frac{[\Sigma_{j \text{ s.t. } n_j = n}(|V_j| - \bar{L}_n^m)^2]^{1/2}}{\Sigma_{n_j = n} 1} . \tag{33}$$

σ_n^m provides the means to decide whether a given $\bar{L}_n^m$ is significantly different from $\bar{R}_n^m$. For example, when the $|V_j|$ are approximately normally distributed, the Student's t test [26] may be used to set confidence intervals on $\bar{L}_n^m$. Note that $\bar{L}_n^m$ has no intrinsic dependence on the time series length. However, σ_n^m depends in a complicated way on the time series length, since it depends on the number of boxes with n passes which may go up or down as the length of the time series increases.

In general, not all of the boxes covering the trajectory of a time series will have the same number of passes. Thus, it is appropriate to calculate $\bar{L}_n^m$ for $n = 2, 3, \ldots$ and display all the $\bar{L}_n^m$ and the corresponding σ_n^m on a single graph. Fig. 5 shows $\bar{L}_n^m$ plotted as a function of n for the Rossler system and a random system with the same $\Psi(T)$ for various values of m. The null hypothesis values of $\bar{R}_n^m$ are plotted as continuous curves connecting the values they take on at the discrete n. It is evident that the $\bar{L}_n^m$ for the Rossler system are significantly displaced from $\bar{R}_n^m$, while for the random system $\bar{L}_n^m$ and $\bar{R}_n^m$ are statistically indistinguishable. In addition, for the Rossler system $\bar{L}_n^3$ are closer to 1 than $\bar{L}_n^2$ – the Rossler system cannot adequately be embedded in $\mathcal{R}^2$, although it can in $\mathcal{R}^3$. For the random system, $\bar{L}_n^2 \approx \bar{L}_n^3$.

In some cases it is convenient to combine the $\bar{L}_n^m$ for a given embedding of a given time series into a single number. We consider the weighted average of $\bar{L}_n^m$, i.e.

$$\bar{\Lambda} = \frac{1}{\Sigma_j n_j} \sum_j n_j \frac{(\bar{L}_{n_j}^m)^2 - (\bar{R}_{n_j}^m)^2}{1 - (\bar{R}_{n_j}^m)^2} . \tag{34}$$

For Brownian motion, $\bar{\Lambda}_m = 0$, while for ideal deterministic systems $\bar{\Lambda}_m = 1$.

The number of passes n_j through box j de-

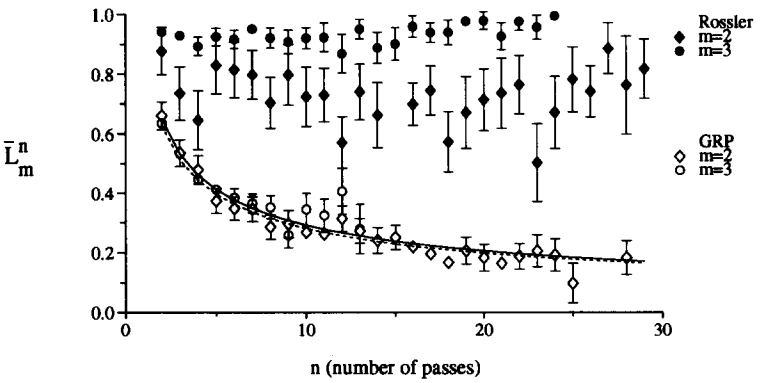

Fig. 5. $\bar{L}_n^m$ versus n for the Rossler system (filled symbols) and a GRP (open symbols) with the same $\Psi(T)$. $\tau = 585$, $m = 3$, $\epsilon = 1/32$, time series length 655 time units. Error bars show $\pm 2\sigma_n^m$. The curves show the theoretical values of $\bar{R}_n^m = c_m/\sqrt{n}$ for Brownian motion.

pends on the length of the time series under analysis. Since $\bar{R}^m_{n_j}$ obviously depends on n_j, a summary statistic involving $\bar{R}^m_{n_j}$ might itself be a function of n_j and hence of the time series length; this is an undesirable feature of a test statistic. $\bar{\Lambda}$ has been constructed so that for the cases of ideal deterministic systems, Brownian motion, and GRPs it is independent of n_j. In these three cases, $\langle |V_j| \rangle$ is respectively 1, $\mathcal{R}^m_{n_j}$, and as specified by eq. (30). Substitution of these terms into eq. (34) shows that for these three cases, $\bar{\Lambda}$ does not depend on $R^m_{n_j}$. (See table 2.)

Confidence intervals for $\bar{\Lambda}$ can be derived by propagating σ^m_n (eq. (33)) through eq. (34)

$$\sigma_m = \frac{1}{\Sigma_j n_j} \left(\sum_j n_j [2\sigma^m_{n_j} \bar{L}^m_{n_j} / (1 - (\bar{R}^m_{n_j})^2)]^2 \right)^{1/2}. \tag{35}$$

Table 2
Expectation values of $|V_j|$ and $\bar{\Lambda}$ for different types of dynamical systems. The vector $\boldsymbol{d} = -\mathbf{H} \cdot \boldsymbol{x}$ depends on the position $\boldsymbol{x}$ of the box j, while $\nu^2 = \text{trace}(\mathbf{C}^{-1}/2)$ is constant for all j.

	$\langle \lvert V_j \rvert \rangle_{n_j = n}$	$\bar{\Lambda}$
Ideal determinism	1	1
Brownian motion	$\bar{R}^m_n$	0
GRP	$\dfrac{[\lvert \boldsymbol{d} \rvert^2 + \nu^2 (R^m_{n_j})^2]^{1/2}}{\lvert \boldsymbol{d} \rvert^2 + \nu^2}$	$\left(\dfrac{\lvert \boldsymbol{d} \rvert^2}{\lvert \boldsymbol{d} \rvert^2 + \nu^2} \right)^{1/2}$

For longer time series, $\Sigma_j n_j$ is larger, tending to reduce the size of the confidence intervals.

Such confidence intervals suggest the range of $\bar{\Lambda}$ that might be encountered if a different set of multiply occupied boxes was selected randomly for constructing the average implicit in eq. (34). As such, confidence intervals of this type are

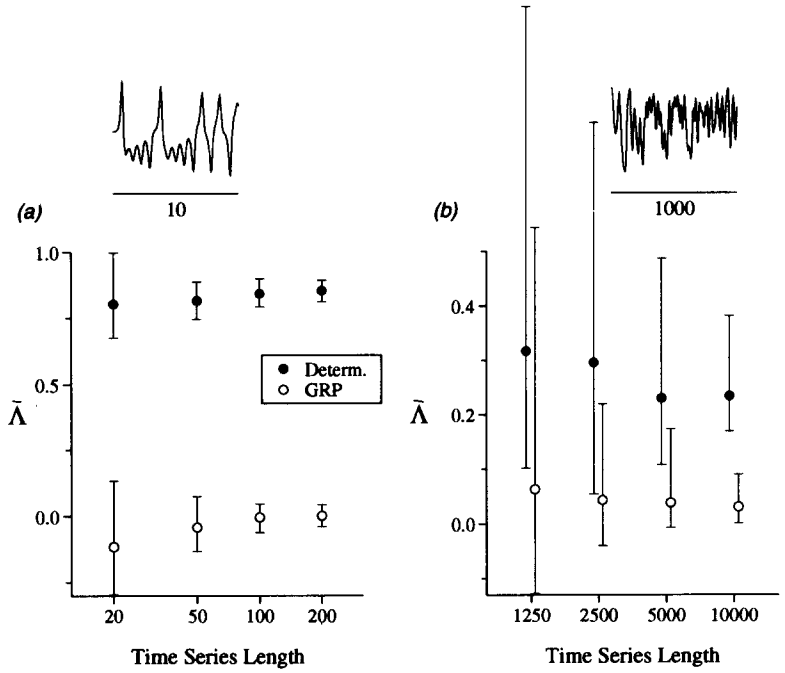

Fig. 6. $\bar{\Lambda}$ versus the length of the time series for the x-component of the Lorenz equations ($\dot{x} = 10(y - x)$; $\dot{y} = 28x - y - x$; $\dot{z} = xy - 8z/3$) and for a GRP. The analysis was repeated for several non-overlapping segments. The error bar plotted above the symbol is twice the population standard deviation of $\bar{\Lambda}$ for the different segments. The error bar plotted below the symbols is 2σ (eq. (35)) calculated within each segment (averaged over the different segments), $m = 3$, $\tau = 0.75$. The inset shows a segment of the deterministic time series of length 10 time units. (b) The same as (a), but for the system of eq. (36) in section 9.2.1 with the parameters given in the text. $m = 10$, $\tau = 75$. The inset shows a segment of the deterministic time series of length 1000 time units.

generally interpreted as indicating the range of values that will be encountered if, from a given dynamical process, a time series is collected from a random initial time. However, for deterministic systems the multiply filled boxes cannot be said to be selected randomly. The relevance of the confidence intervals given by eq. (35) can only be established empirically. Fig. 6 compares the confidence intervals calculated from eq. (35) and from a population of non-overlapping time series from a low- and high-dimensional dynamical system and corresponding GRPs. The two forms of confidence intervals are similar, although for the high-dimensional system and its GRP, the population-based confidence interval is approximately twice as large as the confidence interval from eq. (35). This suggests that the confidence intervals from eq. (35) should be used only as a guide in interpreting results, rather than as a precise delimiter of when nonlinear determinism has been detected.

The test for whether a time series is inconsistent with a GRP generally involves calculating $\bar{\Lambda}$ or $\bar{L}_n^m$ for an appropriate GRP (see section 5.3) and making a comparison to $\bar{\Lambda}$ or $\bar{L}_n^m$ for the time series itself using the confidence interval σ (eqs. (35) or (33)) as a guide to the statistical significance of any difference found. We note that often, when the embedding lag τ is selected appropriately, an embedded GRP resembles Brownian motion, and so $\bar{\Lambda} \approx 0$ and $\bar{L}_n^m \approx \bar{R}_n^m$. This is illustrated in figs. 5, 9, 10 and 15. (If the time series has a histogram that has much longer tails than a Gaussian distribution, the box size ϵ must be chosen with care. See sections 7 and 9.1.3.)

7. Non-Gaussian processes

In many cases, there is a trivial statistical test to distinguish a measured time series from a GRP. A time series $x(t)$ has a one-time probability density $p(x)$ which is often termed a "histogram". The histogram of many time series is not Gaussian, while a GRP does have a Gaussian histogram. This fact does not have much significance from a dynamical point of view, since the histogram of a measured time series can be reshaped by a nonlinear transformation of the time series itself, which does not effect the underlying dynamics. Thus, a time series from a GRP may be passed through a static nonlinear filter (e.g., $x(t) = y^2(t)$) which can produce a non-Gaussian histogram.

The irrelevance of the histogram to the system's dynamics[#1] should be reflected in the analysis technique – the results should not depend strongly on a dynamically irrelevant static nonlinear transformation. Theiler et al. [21] have suggested rescaling measured time series to give them a Gaussian density before comparing them to GRPs.

In the case of coarse-grained flow averages, the shape of the histogram is not critical to the results obtained. The reason for this can be understood as follows: any static nonlinear transformation applied to the time series $x(t)$ applies similarly to all components of the embedded vector $\boldsymbol{x}(t)$. At any point in the phase space, this static nonlinear transformation of the time series amounts to a locally linear, diagonal transformation of the phase space. Although such transformations may change the angles between vectors, acute angles are kept acute and obtuse angles are kept obtuse. Thus, trajectory vectors that originally had similar orientations will not, under the transformation, end up having opposing orientations.

The effect of a variety of static nonlinear transformations on $\bar{\Lambda}$ from time series from the Lorenz equations and the corresponding GRP are shown in fig. 7. To a great extent, $\bar{\Lambda}$ from all of the transformed random time series behave like the GRP, and the $\bar{\Lambda}$ from the transformed deterministic time series behave much like the original Lorenz system. However, for time series

[#1] Second- and higher-dimension probability densities, e.g. $p(x(t), x(t-\tau))$ are not generally irrelevant to the dynamics.

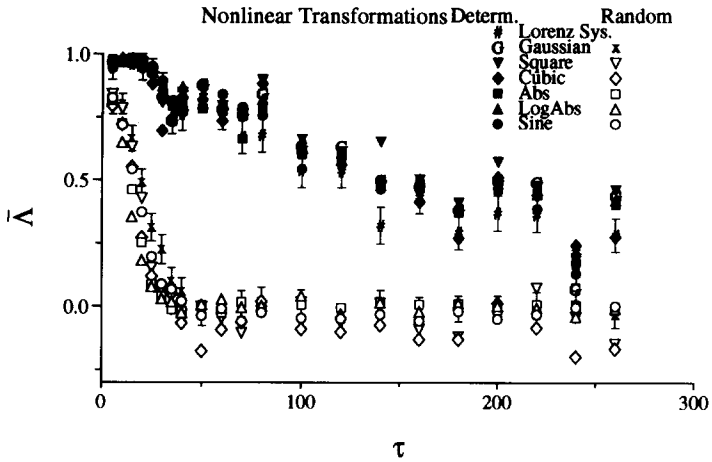

Fig. 7. $\bar{\Lambda}$ versus τ for a time series from the x-component of the deterministic Lorenz system and for a GRP with the same $\Psi(T)$. Other time series were derived from the Lorenz and the GRP by application of the indicated static nonlinear transformations. Square $x^2(t)$; Cubic $x^3(t)$; Abs $|x(t)|$; LogAbs $\log(|x(t)|)$; Sine $\sin x(t)/8$. Time series length = 10000, $\tau = 0.75$, $m = 3$, box resolution 1/16 (for Gaussian and square, 1/32; for cubic 1/128).

with long tails, the box size ϵ must be small compared to the overall range of the histogram in order to prevent the majority of points from falling into just a few boxes.

8. Noise

The division of dynamical systems into the two categories of deterministic and stochastic is largely a matter of notational convenience. Noise is almost always involved in the measurement of time series. This implies the addition of a stochastic component to the embedding reconstruction of otherwise deterministic systems. In addition, the dynamics of all real-world systems are influenced by an environment: the complexities of the environment usually dictate that they be treated as random forcing of the system.

The distinction we have made up to now has been between strictly deterministic systems and globally linear dynamical systems forced by random noise (and perhaps with a static nonlinear transformation applied to the measured time series). This distinction is useful partly because linear systems cannot produce chaos.

What about nonlinear dynamical systems with a stochastic element? The case that is often of interest is when the stochastic element is "small", so that much of the dynamics of interest would arise even in the absence of the stochastic element. An important issue is the extent to which such underlying nonlinear deterministic dynamics can be seen through the haze of the stochastic influence.

A situation that is relatively easy to analyze and of wide relevance is additive measurement noise. Numerical experiments conducted by adding noise to various deterministic signals suggest that the method is fairly robust to the addition of noise: although $\bar{\Lambda} < 1$ in the presence of noise, it is still often possible to distinguish between the noisy deterministic time series and a GRP with the same $\Psi(T)$ even when the signal-to-noise ratio is 2:1. (See fig. 8.)

The power spectrum of the added noise is of critical importance here. Adding white noise quickly degrades the value of $\bar{\Lambda}$ unless a box traversal time threshold is used in eq. 3 or the time series is low-pass filtered so that it becomes strongly correlated over the box traversal time. Certain types of "optimal" linear filters attempt

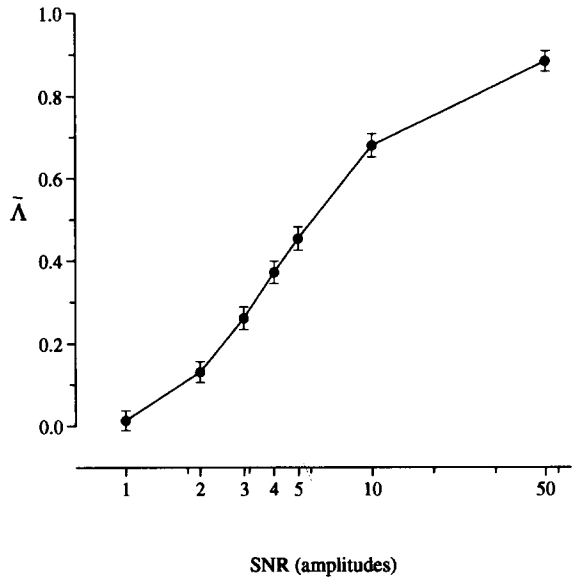

Fig. 8. $\bar{\Lambda}$ versus signal-to-noise ratio for the Lorenz system $m = 3$, $\tau = 0.75$, $\epsilon = 1/64$. To construct the noisy time series, a GRP having the same $\Psi(T)$ as the deterministic time series was added at the indicated amplitude ratios. The value of $\bar{\Lambda}$ for the pure GRP is 0.00 ± 0.02.

to shape the noise to the power spectrum of the underlying signal [27]; this is appropriate here. However, a simple low-pass filter can be quite effective. For pseudo-white noise (where the highest frequency is 50 cycles per time unit) added to the Lorenz system, low-pass filtering allows the distinction to be made between the deterministic system and a GRP even when the signal-to-noise ratio is 1:2 (before filtering). The behavior of the noise's autocovariance function $\Psi_w(T)$ near $T = 0$ determines whether it is best to use large or small box size ϵ: large box size is often useful for dealing with exponentially correlated noise, while small boxes can be used when the noise has a flat $\Psi_w(T)$ near $T = 0$.

The case of random influences on the dynamics themselves is much harder to analyze and remains a subject for further investigation. For chaotic systems, random noise added to dz/dt can be amplified by the positive Lyapunov exponents. One effect of this is that the fall-off of $\bar{\Lambda}$ with τ (see section 9.1.2) occurs at smaller τ as the random forcing is increased in amplitude.

9. Examples

This section presents an analysis of time series from a variety of sources. Some of the examples are known to be deterministic, while in others the question of determinism must be answered by analysis of the time series itself. We first examine some time series where determinism is evident in a three-dimensional embedding – the lowest possible effect embedding dimension for a continuous-time chaotic system. Then, systems where this is not the case are studied. $\bar{\Lambda}$ is used throughout as a test statistic.

9.1. *Low-dimensional systems* ($m = 3$)

9.1.1. *The Rossler system*

The Rossler system [28]

$$\dot{x} = -(y + z)\,; \quad \dot{y} = x + 0.15y\,;$$

$$\dot{z} = 0.2 + z(x - 10)$$

has been widely studied as an example of low-dimensional chaos. Its autocovariance function oscillates with an envelope that decays very slowly.

Consistent with the system's determinism, $\bar{\Lambda} \approx 1$ for a wide range of τ. (See fig. 9.) An oscillation in $\bar{\Lambda}$ is evident whose minima occur when $\Psi(\tau)$ is at an extremum. In these cases, the embedded time series has the shape of a narrow cigar (since $\Psi(\tau) \approx \pm 1$) and a box may cover both the ascending and descending limbs of the trajectory. When τ is large compared to the pseudo-period of the time series, $\bar{\Lambda}$ at these minima falls slowly to zero.

A Gaussian random process (GRP) generated with the same $\Psi(\tau)$ has a smaller $\bar{\Lambda}(\tau)$ than the deterministic system. As derived theoretically in Section 5. $\bar{\Lambda} \to 0$ when $\Psi(\tau)$ is at a maximum or minimum, but at other τ, $\bar{\Lambda} > 0$. This is due to the high predictability of the spectrally narrow-band random signal. This effect can be made small by making the box size ϵ small.

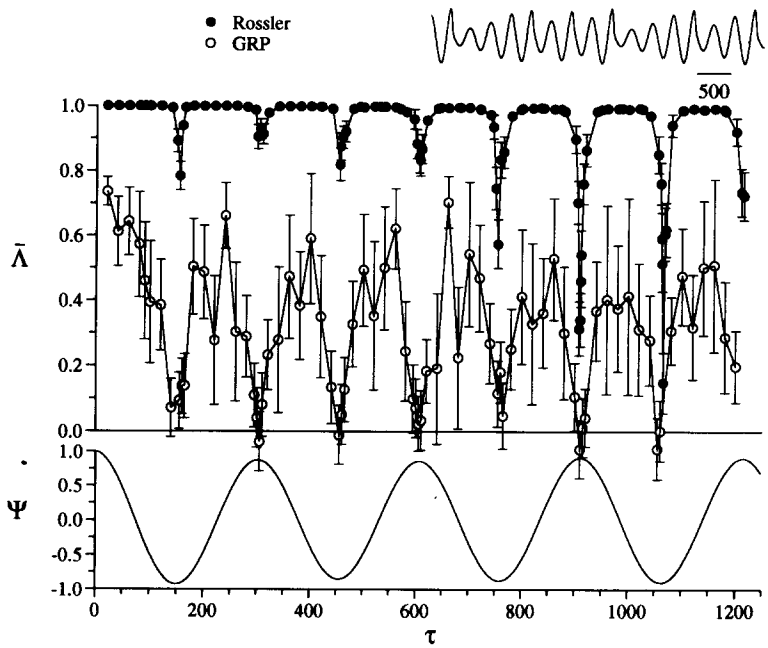

Fig. 9. $\bar{\Lambda}$ versus τ for the Rossler x-component and a GRP with the same $\Psi(T)$. The inset shows a short segment of the Rossler time series; the bar indicates the time scale, $m = 3$, $\epsilon = 1/128$.

An analysis of another low-dimensional deterministic system, the Lorenz system, can be found in ref. [10].

9.1.2. Belousov–Zhabotinskii reaction

The Belousov–Zhabotinskii reaction is an oscillating chemical reaction. The time series studied here comes from an experiment involving a continuously stirred reactor into which the various chemicals were fed at a constant rate [30,31]. The time series consists of measurements of the bromide ion potential with approximately 14-bit resolution. The oscillations in the bromide ion potential are superficially like a saw-tooth wave of varying height.

For a wide range of τ, $\bar{\Lambda} \approx 1$, indicating the system's determinism. (See fig. 10.) The corresponding GRP produces a $\bar{\Lambda}$ that falls off rapidly to near 0, particularly at extrema of $\Psi(\tau)$. The fall-off of the deterministic $\bar{\Lambda}$ with τ suggests either a Lyapunov exponent (bits/pseudoperiod) that is somewhat larger than that for the Rossler system, or noise affecting the dynamics.

A GRP is not satisfactory as a surrogate for the BZ time series: it does not capture the obvious time asymmetry and sawtooth nature of the BZ data. To assess the extent to which the non-zero $\bar{\Lambda}$ in the BZ arises simply from the sawtooth appearance, we generated a time series consisting of constant slope ascenders and descenders, whose amplitude endpoints were randomly selected from the segment endpoints found in the BZ time series. (See fig. 1b.) Consistent with the deterministic mechanism of this synthesized signal over short times, $\bar{\Lambda}$ is close to 1 for short τ. Within approximately 2 pseudocycles, however, $\bar{\Lambda}$ has fallen to near zero – much faster than for the BZ data itself. The observed determinism of the BZ time series therefore goes well beyond that due to its sawtooth shape. Note that if an embedding lag $\tau < 50$ were used, it would be hard to distinguish the determinism of the BZ time series from the piecewise linear random time series. Such short

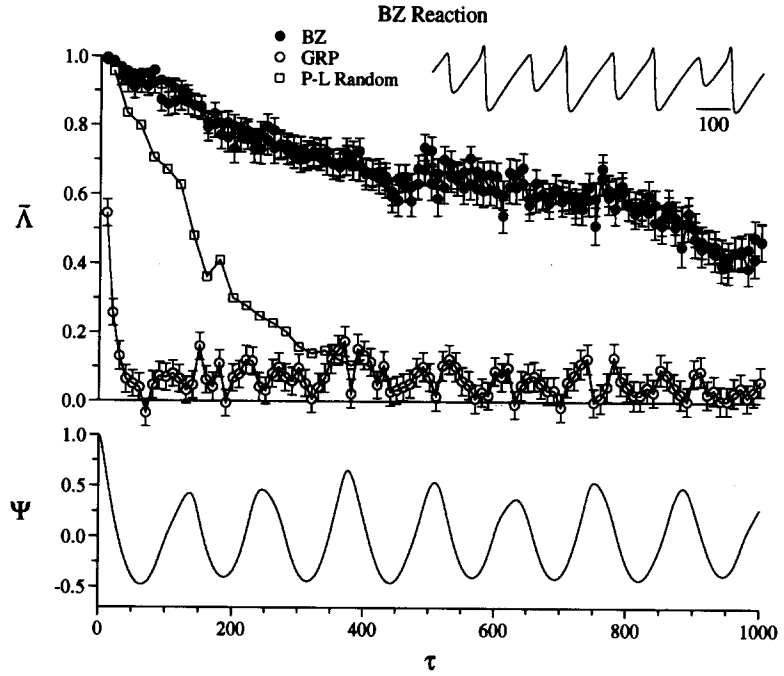

Fig. 10. $\bar{\Lambda}$ versus τ for the Belousov–Zhabotinskii time series (fig. 1a), a Gaussian random process, and the synthesized piece-wise linear random time series (P–L Random). B–Z data provided by H. Swinney.

embedding lags are commonly used in dimension calculations [29].

9.1.3. Measles

The number of measles cases reported each month in New York City has been recorded since 1928. Sugihara and May [9] analyzed the time series for 1928–1963 by assessing its predictability using linear and nonlinear models, and concluded that it shows signs of deterministic chaos. The dimension and Lyapunov exponent of the time series have been examined by Schaffer et al. [32].

$\bar{\Lambda}$ for the measles time series (fig. 11) is clearly above that for a corresponding GRP, although even for the measles, $\bar{\Lambda} \to 0$ when $\Psi(\tau)$ is at an extremum, which occurs at 6 month intervals (corresponding to oscillations with a pseudo-period of 1/year). There is, however, an obvious dissimilarity between the GRP and the measles time series. A histogram of the measles time series shows a sharp peak near zero and long tails that correspond to the yearly winter spikes in the number of measles cases. This is very different from the Gaussian histogram of the GRP. This difference can be reduced by statically transforming the measles data $x(t)$ to generate a new time series $y(t) = g(x(t))$.

The transformation $g(\cdot)$ can be chosen in a number of ways; we used $g(x) = x^{1/5}$ (where the real root is taken), which produces a histogram of $y(t)$ without long tails. $\bar{\Lambda}$ for $x(t)$ and $\bar{\Lambda}$ for $y(t)$ are very similar, consistent with section 7. However, the GRP based on $y(t)$ has a more slowly decaying autocovariance function than $x(t)$. This produces a $\bar{\Lambda}$ for the $y(t)$ GRP that shows periodic peaks, reducing the difference between $\bar{\Lambda}$ for $y(T)$ and the GRP.

An interesting aspect of $\bar{\Lambda}$ for both $x(t)$ and $y(t)$ is that $\bar{\Lambda} \to 0$ for $\tau = 6, 12, 18, \ldots$ months. This suggests that points in $x(t)$ separated by multiples of 6 months are dynamically unrelated

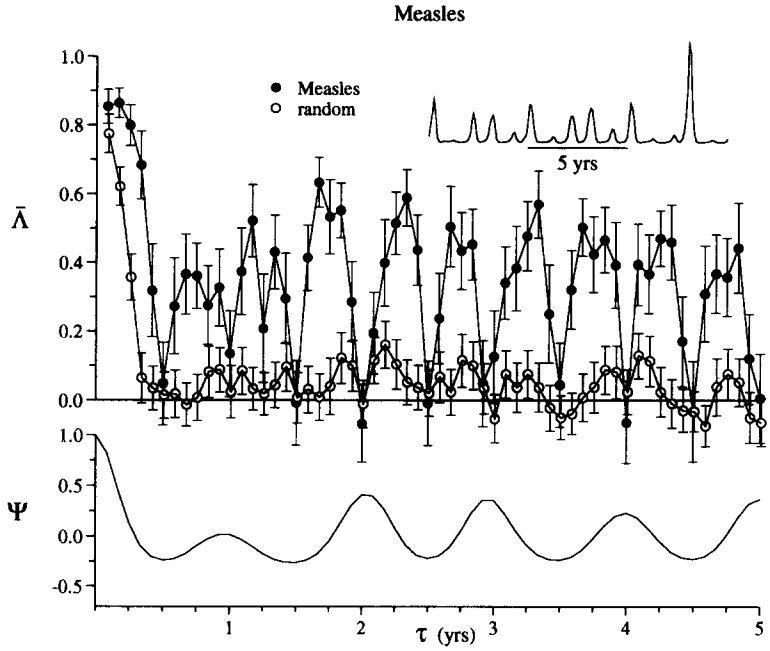

Fig. 11. $\bar{\Lambda}$ versus τ for the Measles time series. Data provided by G. Sugihara through the Santa Fe Institute.

to one another; in particular one year's level of measles is not predictable from the previous year's. (This is consistent with the numerical results of Sugihara and May (fig. 4d in [9]).)

The yearly winter peaks in measles cases are due to an increase in the contact rate between infectives and susceptibles during the winter [33]. A possible model for the measles time series is yearly winter peaks of random amplitudes (see fig. 1f). This model is a mixture of deterministic dynamics (the strictly regular phase the yearly oscillation) and randomness (the amplitudes of the peaks). $\bar{\Lambda}$ for this model shows a pattern similar to that observed in the measles time series, suggesting that much of the observed determinism of the measles time series is due to the seasonal nature of measles.

The second half of the time series shows alternating high and low peaks, suggesting a deterministic mechanism for the amplitudes of the peaks. Such observations are perhaps better studied by Poincaré return maps than the methods described in this paper.

9.1.4. Sunspots

A record of the number of visible sunspots each month has been kept since 1610 [34]. This record shows somewhat irregular peaks at intervals of approximately 11 years together with month-to-month variability.

No significant difference was found between $\bar{\Lambda}$ for the sunspots and $\bar{\Lambda}$ for the random signal. To avoid the influence of high-frequency noise in the time series (see sections 3 and 8), both were low-pass filtered using a 31-month cosine filter. After filtering, $\bar{\Lambda}$ (fig. 12) clearly shows signs of nonlinear determinism in the >3 year variability in sunspots numbers. Similar values for $\bar{\Lambda}$ are found if the filtered time series are rescaled to produce a uniform histogram. These results are consistent with previous studies that have found the time series to be better modelled with nonlinear rather than linear models [35,21].

9.2. High-dimensional systems

High-dimensional systems pose a difficult

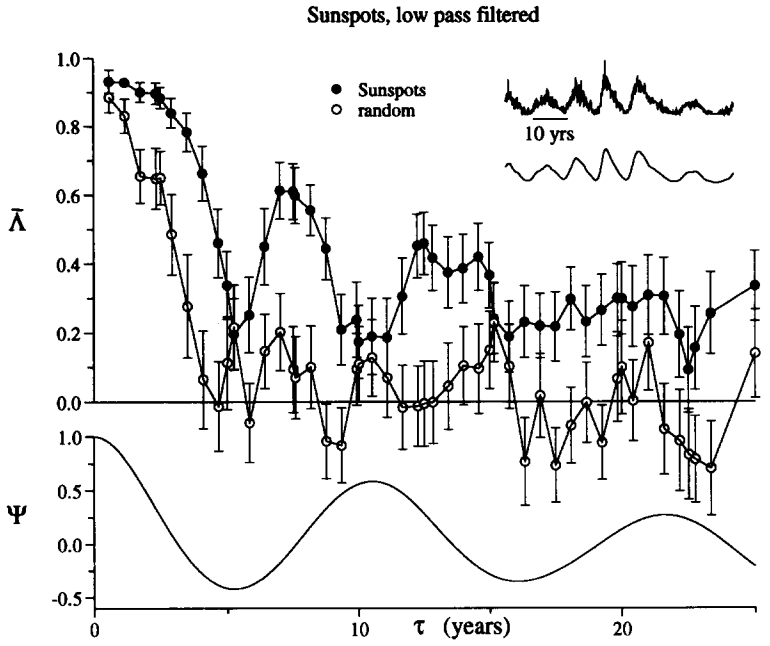

Fig. 12. $\bar{\Lambda}$ versus τ for the low-pass filtered sunspots data (the lower signal on the inset).

problem for analysis. If the linear extent of the region of phase space occupied by the trajectory is 1, and the box size is ϵ, then ϵ^{-m} boxes are needed to cover the relevant region of the phase space. For example, for a 10-dimensional embedding and boxes of size 1/32, $2^{50} \approx 10^{15}$ boxes are needed. Quite aside from the computational problem of keeping track of this number of boxes, the length of the time series needed to generate multiple passes through many boxes is immense, and unlikely to be encountered in practice.

9.2.1. A delay-differential equation

Delay-differential equations are, in principal, infinite-dimensional systems, but in some cases the attractors for such systems are finite-dimensional. The delay equation

$$\frac{dx}{dt} = \frac{ax(t-\delta)}{1+x(t-\delta)^c} - bx(t) \qquad (36)$$

has been proposed as a model for nonlinear feedback control in physiology [36]. It has been widely studied in the context of nonlinear dynamics [37,3] in part because attractors of different dimension are found for different values of the parameter δ. For $a = 0.2$, $b = 0.1$, $c = 10$ and $\delta = 100$, eq. (36) has an attractor whose estimated dimension is ~7.5 [3]. This suggests that an appropriate embedding dimension is $m \geq 8$.

For an embedding dimension of 10 and a time series of length 1.32×10^5 time units, the smallest useful box resolution is 1/4. The relationship between $\bar{\Lambda}$ and τ (fig. 13) shows that $\bar{\Lambda}$ is bigger for the deterministic time series than the random one – this shows that $\bar{\Lambda}$ distinguishes between the high-dimensional delay-differential system and a GRP, even if the time series are very similar to the eye (fig. 1i and j).

Casdagli [8] investigated the system of eq. (36) by constructing an ad hoc predictor model from the time series. He found that it was possible to make informative predictions using $\tau = 6$. Fig. 13 suggests that it is possible to distinguish between the deterministic system and a GRP even at τ much larger than this.

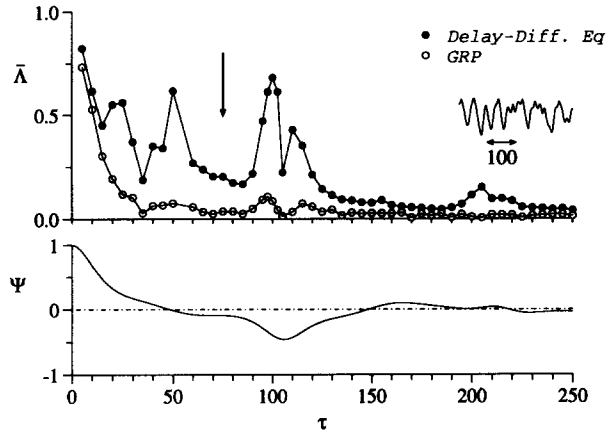

Fig. 13. $\bar{\Lambda}$ versus τ for a time series of length 5×10^4 time units from equation 36 and for a random signal with the same $\Psi(T)$. $m = 10$ and the box edge length is 1/4 the amplitude range of the time series.

An interesting aspect of eq. (36) is that dx/dt at time t is a simple function of $x(t - \delta)$ and $x(t)$. This means that when the time series is embedded with a delay $\tau = \delta/k$, $k = 1, 2, 3, \ldots$ one component of the tangent vector to the trajectory is a simple function of another component of the position in phase space. At other embedding lags, these components are only indirectly related to one another; the relatively fast oscillations ($<\delta$) of the time series suggest that this relationship is complicated. At one extreme, when $\tau = \delta$, component k of the tangent to the trajectory depends only on component $k - 1$ of the position in phase space – this creates a strong dependence between the trajectory vector and phase space position even when the embedding dimension is not adequate to represent the attractor. Indeed, $\bar{\Lambda}$ shows peaks at $\tau = 100$ and 50 (and perhaps at 33 and 25 as well). When $\tau = \delta$, an embedding dimension of $m = 3$ is adequate to demonstrate determinism, although at other τ this is not the case. We can offer no explanation for the small peak observed at $\tau = 200$.

The coarse box resolution used in the analysis prevents $\bar{\Lambda} \to 1$ at a theoretically satisfactory embedding dimension. There are hints of deterministic structure in the time series for embedding dimensions <6, but the box resolution does not appear to be the limiting factor in $\bar{\Lambda}$. (See fig. 14.) At higher embedding dimension, however, smaller boxes lead to lager $\bar{\Lambda}$. For the random time series, $\bar{\Lambda}$ is more or less independent of box size or embedding dimension.

For the length of time series analysed, it is possible to use boxes as small as 1/16 with $m = 10$ and still have several boxes with two

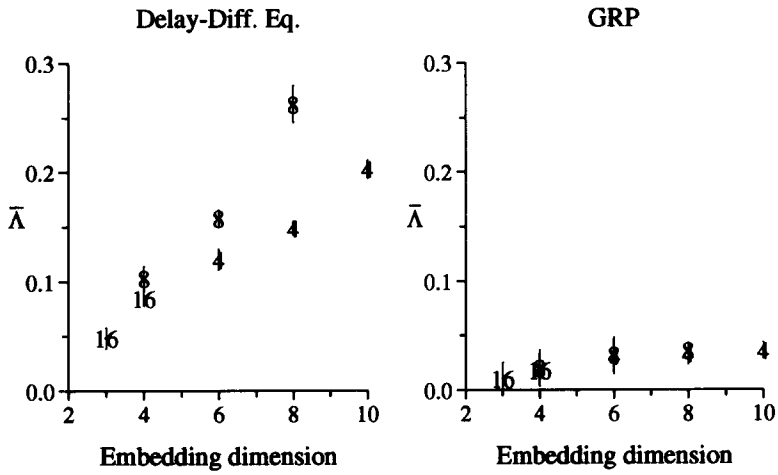

Fig. 14. $\bar{\Lambda}$ versus embedding dimension m for various values of box resolution. The symbols indicate the box resolution used, e.g., 8 indicates a resolution of 1/8 the amplitude range of the time series. We selected $\tau = 75$ as marked by the arrow in fig. 13 in order to minimize the effects of the peaks shown in that figure. (Time series length is 1.32×10^5.)

passes of the trajectory. However, there may be a "selection bias" in these boxes towards larger $\bar{\Lambda}$. This possibility can be assessed by reference to the random time series – when boxes are small so that a very small fraction of boxes have multiple passes of the trajectory, $\bar{\Lambda}$ for the random time series can increase dramatically from its value with larger boxes.

9.2.2. *Coupled nonlinear oscillators*

The time series shown in fig. 1m originates in an experimental electronic nonlinear circuit consisting of a harmonic oscillator driving 8 inductively coupled diodes [38]. The harmonic forcing of the circuit gives the power spectrum of the time series a sharp peak at this frequency. The embedded time series is – at large length scales – grossly like a limit cycle. When a coarse box resolution of 1/8 is used to calculate $\bar{\Lambda}$, the results are consistent with a noisy limit cycle, and the nonlinearity of the system can be seen even in a 2-dimensional embedding. (See fig. 4.)

Because the trajectory is far from filling the phase space uniformly (unlike the delay-differential system), it is possible to use small boxes (1/32) in high-dimensional embeddings. Using small boxes allows details of the dynamics to be resolved beyond those of the coarse limit cycle. $\bar{\Lambda}$ indicates that in a 4-dimensional embedding, the time series is only distinguishable from a GRP at $\tau < 200$. (fig. 15.) When $m = 6$, however, $\bar{\Lambda}$ shows strong signs of nonlinear determinism, which is further increased when $m = 8$. This suggests that higher-dimensional embeddings are untangling the dynamics. $\bar{\Lambda}$ for the GRP is similar for $m = 4.6$. ($\bar{\Lambda}$ for the GRP embedding in 8-dimensions cannot be calculated at this box size, since so few boxes have multiple passes of the trajectory.) We note without explanation the dip in $\bar{\Lambda}$ for the time series at $\tau = 400$.

9.2.3. *Heart rate*

The time interval between heart beats varies from beat to beat. It has been suggested that in

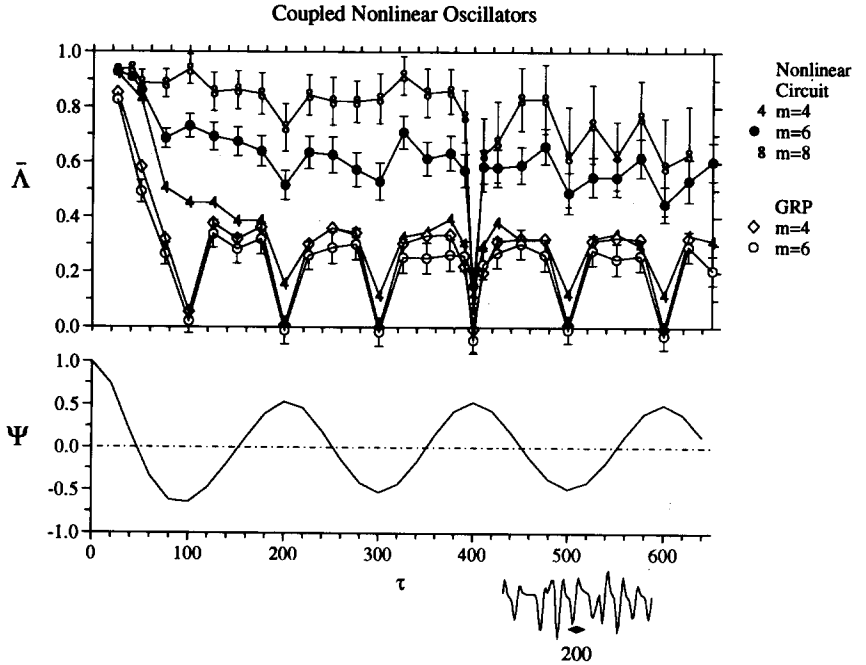

Fig. 15. $\bar{\Lambda}$ versus τ for the nonlinear electronic circuit time series and a corresponding GRP. $\epsilon = 1/32$. The inset shows a short segment of the time series, which had a total length of 200000 points (1000 pseudo-periods). Data provided by P. Linsay.

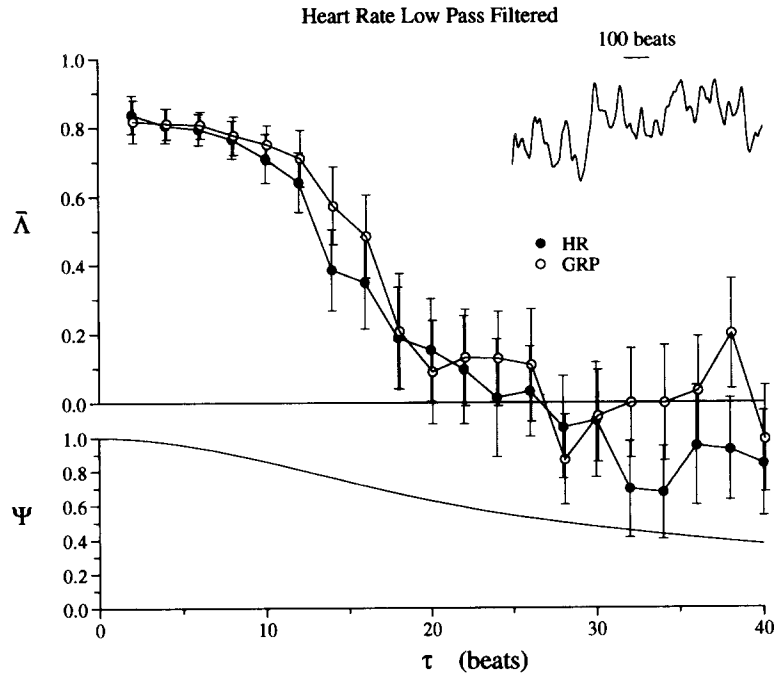

Fig. 16. $\bar{\Lambda}$ versus τ for low-pass filtered heart rate data and a corresponding GRP. $m = 6$, $\epsilon = 1/16$. The inset shows a short segment of the time series, which had a total length of 2048 beats. Data provided by M. Talajic.

healthy individuals, heart rate is the output of a deterministically chaotic system [39]. Figure 1k shows beat-to-beat intervals between heart beat over 30 minutes recorded from a healthy, adult subject. The corresponding GRP appears similar by eye. No claim can be made that this one record is representative of heart rate in general.

The relationship between $\bar{\Lambda}$ and τ for this time series shows no difference between the original time series and the GRP. To some extent, this may be due to the very rapid fluctuations in heart rate. To test this possibility, we filtered the heart rate signal and the corresponding GRP in two ways: (i) a low-pass filter below 0.1 Hz and (ii) a band-pass filter between 0.1 and 0.4 Hz. These two bands correspond to frequency regions believed to be of physiological significance in heart rate [40]. Neither of the filtered signals showed signs of nonlinear determinism for $m = 4$ or $m = 6$. Use of higher embedding dimensions is limited by the time series length.

10. Summary

In this paper, we address the issue of deciding whether a time series is from a deterministic or a stochastic system. The basis we suggest here for this decision is whether a delay-embedding of the time series is consistent with a flow that is a single-valued function of position in phase space – whether the flow through nearby points in the phase space goes in similar directions. The method of local coarse-grained flow averages and the statistic $\bar{\Lambda}$ offer a practical means of assessing the single-valuedness of the flow for an embedded time series.

A limiting case for stochastic systems occurs when the trajectory is equally likely to point in any direction in phase space – this corresponds to Brownian motion. $\bar{\Lambda} = 0$ for this case, in contrast with the situation for ideal determinism where $\bar{\Lambda} = 1$.

However, even stochastic systems can have

trajectories that may be locally parallel. For Gaussian random processes (GRP), the coarse-grained flow average depends only on the embedding parameters and the autocovariance function of the time series, and has certain universal features. Because of the generality of GRPs and the fact that they model linear dynamical systems, a GRP is a good "null hypothesis" for a time series. A straightforward test for chaotic dynamics in a time series is to compare $\bar{\Lambda}$ of the time series and of a corresponding GRP. If there is no embedding dimension where $\bar{\Lambda}$ of the time series is greater than that of the GRP, no evidence is provided for chaos. If $\bar{\Lambda}$ is larger in a statistically significant way for the time series than for the GRP, then there is evidence for nonlinear determinism – but the most that can definitely be said is that the time series is inconsistent with a GRP or a static nonlinear transformation of a GRP.

If a deterministic system generating a time series is low-dimensional (e.g., $m = 3$), $\bar{\Lambda}$ may approach 1 (the value for ideal determinism), even when a corresponding GRP has $\bar{\Lambda}$ near 0. In the presence of additive noise, the coarse-grained flow averages are degraded towards the random value, but numerical experiments indicate that a signal-to-noise ratio as poor as 2:1 (in amplitudes) does not prevent the distinction being made between a deterministic system and a corresponding GRP.

For time series from high-dimensional systems (e.g., $m = 8$), the test can in practice provide only an indication of nonlinear determinism. The primary limitation here is the length of the time series. For short time series and the consequent large box size needed for averaging, the $\bar{\Lambda}$ statistic may fall substantially below 1, although it may be easily distinguishable from the $\bar{\Lambda}$ of a corresponding GRP, even when the signal-to-noise ratio is 5:1. An increase in $\bar{\Lambda}$ with m may also provide an indication of high-dimensional determinism.

We believe that the method of coarse-grained flow averages should prove useful in a variety of contexts. For the calculation of dimension and Lyapunov exponents from deterministic time series, the method should be able to provide an independent means of selecting an appropriate embedding dimension. For modelling of time series, the method can indicate whether deterministic nonlinear dynamics should be emphasized over stochastic dynamics.

Acknowledgements

We thank Paul Linsay for providing time series from his experiments with coupled oscillators, Daniel Lathrop and Harry Swinney for the BZ reaction time series, and Neil Gershenfeld for showing us some of the resources from the time series competition of the Santa Fe Institute. We also thank Hiroyuki Ito and Jérome Losson for helpful discussions. This research is supported by a grant from the Natural Sciences and Engineering Research Council of Canada. D. Kaplan is a Postdoctoral Fellow of the North American Society of Pacing and Electrophysiology.

References

[1] P. Cvitanovic, Universality in Chaos, 2nd edition (Adam Hilger, Bristol, 1989).
[2] P. Grassberger and I. Procaccia, Characterization of strange attractors, Phys. Rev. Lett. 50 (1983) 346–349.
[3] P. Grassberger and I. Procaccia, Measuring the strangeness of strange attractors, Physica D 9 (1983) 189–208.
[4] J. Guckenheimer, Noise in chaotic systems, Nature 298 (1982) 358–361.
[5] A. Wolf, J.B. Swift, H.L. Swinney and J.A. Vastano, Determining Lyapunov exponents from a time series, Physica D 16 (1985) 285–317.
[6] P. Bryant, R. Brown and H.D.I. Abarbanel, Lyapunov exponents from observed time series, Phys. Rev. Lett. 65(13) (1990) 1523–1526.
[7] J.D. Farmer and J.J. Sidorowich, Predicting chaotic time series, Phys. Rev. Lett. 59(8) (1987) 845–848.
[8] M. Casdagli, Nonlinear prediction of chaotic time series, Physica D 35 (1989) 335–356.
[9] G. Sugihara and R.M. May, Nonlinear forecasting as a way of distinguishing chaos from measurement error in time series, Nature 344 (1990) 734–741.

[10] D.T. Kaplan and L. Glass, A direct test for determinism in a time series, Phys. Rev. Lett. 68(4) (1992) 427–430.

[11] F. Takens, in: Dynamical Systems and Turbulence, Warwick, 1980, eds. D.A. Rand and L.S. Young (Springer, Berlin, 1981) p. 366.
N. Gershenfeld, An experimentalist's introduction to the observation of dynamical systems, in: Directions in Chaos, ed. Hao Bai-lin (World Scientific, Singapore, 1988) vol. 2, p. 310.
T. Sauer, J.A. Yorke and M. Casdagli, Embedology, J. Stat. Phys. 65(3/4) (1991) 579–616.

[12] M.B. Kennel, R. Brown and H.D.I. Abarbanel, Determining embedding dimension for phase space reconstruction using the method of false nearest neighbors, Phys. Rev. A, in press.

[13] Lord Rayleigh, On the problem of random vibrations, and of random flights in one, two, or three dimensions, Phil. Mag. 37 (1919) 321.

[14] S. Chandrasekhar, Stochastic problems in physics and astronomy, Rev. Mod. Phys. 15 (1943) 1, 89.

[15] J.M. Hammersley, The distribution of distance in a hypersphere, Ann. Math. Stat. 21 (1950) 447–452.

[16] R.D. Lord, The distribution of distance in a hypersphere, Ann. Math. Stat. 25 (1954) 794–795.

[17] W. Feller, An Introduction to Probability Theory and Its Applications (Wiley, New York, 1966) vol. 2.

[18] G.M. Jenkins and D.G. Watts, Spectral Analysis and its Applications (Holden-Day, San Francisco, 1968), ch. 5.

[19] B.B. Mandelbrot, The fractal geometry of nature (W.H. Freeman, New York, 1983) p. 247.

[20] Technical Staff of the Analytic Sciences Corporation, Applied Optimal Estimation, ed. A. Gelb. (MIT Press, Cambridge, 1974).

[21] J. Theiler, B. Galdrikian, A. Longtin, S. Eubank and J.D. Farmer, Using surrogate data to detect nonlinearity in time series, eds. M. Casdagli and S. Eubank, in: Nonlinear Prediction and Modeling (Addison-Wesley, Reading, 1991).

[22] H. Tong, Non-linear Time Series: A Dynamical System Approach (Oxford Univ. Press, Oxford, 1990) pp. 4–12.

[23] P.Z. Peebles, Jr., Probability, Random Variables, and Random Signal Principles 2nd ed. (McGraw Hill, New York, 1987), chap. 5.

[24] D.T. Kaplan and R.J. Cohen, Is fibrillation chaos?, Circulation Research 67 (1990) 886.

[25] drand48(3) in the Sun Microsystems UNIX reference manual, vol. 2.

[26] Jenkins and Watts, op cit. pp. 83, 85.

[27] W.H. Press, B.P. Flannery, S.A. Teukolsky, W.T. Vetterling, Numerical Recipes in C: The Art of Scientific Computing (Cambridge Univ. Press, Cambridge, 1988), pp. 434–437.

[28] O.E. Rossler, An equation for continuous chaos, Phys. Lett. A 57 (1976) 397–398.

[29] A.I. Mees, P.E. Rapp and L.S. Jennings, Singular-value decomposition and embedding dimension, Phys. Rev. A 36 (1987) 340–346.
A.M. Albano, J. Muench, C. Schwartz, A.I. Mees and P.E. Rapp, Singular-value decomposition and the Grassberger-Procaccia Algorithm, Phys. Rev. A 38 (1988) 3017–3026.

[30] K.G. Coffman, W.D. McCormick, Z. Noszticzius, R.H. Simoyi and H.L. Swinney, Universality, multiplicity, and the effect of iron impurities in the Belousov–Zhabotinskii reaction, J. Chem. Phys. 86(1) (1986) 119–129.

[31] D.P. Lathrop and E.J. Kostelich, Characterization of an experimental strange attractor by periodic orbits, Phys. Rev. A 40(7) (1989) 4028–4031.

[32] W.M. Schaffer, L.F. Oslen, G.L. Truty and S.L. Fulmer, The case for chaos in childhood epidemics, in: ed. S. Krasner. The Ubiquity of Chaos (Amer. Assoc. for the Advancement of Science, Washington, D.C., 1990) pp. 138–166.

[33] W.P. London and J.A. Yorke, Recurrent outbreaks of measles, chickenpox and mumps: I. Seasonal variation in contact rates. Am. J. Epidem. 98 (1973) 453–468.

[34] D.F. Andrews and A.M. Herzberg, Data: A collection of problems from many fields for the student and research worker (Springer-Verlag, New York, 1985).
M. Waldmeier, The Sunspot Activity in the Years 1610–1960 (Schulthess, Zurich, 1961).

[35] J. Kurths and A.A. Ruzmaikin, On forecasting the sunspot numbers, Solar Physics 126 (1990) 407–410.
H. Tong and K.S. Lim, Threshold autoregression, limit cycles and cyclical data, J. Royal Stat. Soc. B 42 (1980) 245–292.
M. Mundt, W.B. Maguire and R.P. Chase, Chaos in the sunspot cycles: Analysis and prediction, J. Geophys. Res. 96 (1991) 1705–1716.

[36] M.C. Mackey and L. Glass, Oscillation and chaos in physiological control systems, Science 197 (1977) 287–289.
L. Glass and M.C. Mackey, Pathological conditions resulting from instabilities in physiological control systems, Ann. N.Y. Acad. Sci. 316 (1979) 214–235.

[37] J.D. Farmer, Chaotic attractors of an infinite-dimensional dynamical system, Physica D 4 (1982) 366–393.

[38] G.H. Gunaratne, P.S. Linsay and M.J. Vinson, Chaos beyond onset: A comparison of theory and experiment, Phys. Rev. Lett. 63(1) (1989) 1–4.

[39] A.L. Goldberger, D.R. Rigney and B.J. West, Chaos and fractals in human physiology, Scientific American 262 (1990) 42–49.

[40] S. Akselrod, D. Gordon, F.A. Ubel, D.C. Shannon, A.C. Barger and R.J. Cohen, Power spectrum analysis of heart rate fluctuations: A quantitative probe of beat-to-beat cardiovascular control, Science 213 (1981) 220–222.