Chapter 16
An Optimization Approach for Finding a Spectrum of Lyapunov Exponents

Panos M. Pardalos, Vitaliy A. Yatsenko, Alexandre Messo, Altannar Chinchuluun, and Petros Xanthopoulos

Abstract In this chapter, we consider an optimization technique for estimating the Lyapunov exponents from nonlinear chaotic systems. We then describe an algorithm for solving the optimization model and discuss the computational aspects of the proposed algorithm. To show the efficiency of the algorithm, we apply it to some well-known data sets. Numerical tests show that the algorithm is robust and quite effective, and its performance is comparable with that of other well-known algorithms.

16.1 Introduction

Brain electrical activity can be recorded from electrodes placed on the scalp or intracranial. It is now possible to record from relatively large areas with macro-electrodes (EEG) or from more localized regions, using microelectrodes. Such recordings, performed in awake, moving animals, and humans have advanced our understanding of

Panos M. Pardalos
Department of Industrial and Systems Engineering, Center for Applied Optimization, University of Florida, Gainesville, FL, USA, e-mail: pardalos@ufl.edu

Vitaliy A. Yatsenko
Department of Industrial and Systems Engineering, Center for Applied Optimization, University of Florida, Gainesville, FL, USA, e-mail: yatsenko@ufl.edu

Alexandre Messo
Department of Optimization, Kungliga Tekniska Högskolan, Stockholm, Sweden, e-mail: alex.messo@gmail.com

Altannar Chinchuluun
Department of Industrial and Systems Engineering, Center for Applied Optimization, University of Florida, Gainesville, FL, USA, e-mail: altannar@ufl.edu

Petros Xanthopoulos
Department of Industrial and Systems Engineering, Center for Applied Optimization, University of Florida, Gainesville, FL, USA, e-mail: petrosx@ufl.edu

This research is supported by NSF and Air Force grants.

W. Chaovalitwongse et al. (eds.), *Computational Neuroscience*,
Springer Optimization and Its Applications 38, DOI 10.1007/978-0-387-88630-5_16,

many normal physiological processes, such as the sleep–wake cycle and motor control, as well as pathological conditions such as epilepsy, Parkinson's disease and other movement disorders, and sleep disorders such as sleep apnea and narcolepsy. Traditionally, neurophysiologists analyze such signals using visual inspection or through statistical analysis of linear signal properties such as the spectrogram and coherence. More recently, investigators have begun to investigate the spatiotemporal dynamical features of neurological signals. However, many of these techniques have been applied to mathematical models or to dynamical systems that are much less complex than the brain. Therefore, there is need to develop and evaluate mathematical techniques that provide robust results in higher dimensional, nonstationary and noisy biological systems such as the brain. Although difficult or even impossible to prove, there has been little debate that brain activities should be modeled as nonlinear systems. As a result, during the last decade, a variety of nonlinear time series analysis techniques have been applied repeatedly to EEG recordings during physiologic and pathologic conditions. Among those, the algorithms based on the Lyapunov exponents appears promising for characterizing the spatiotemporal dynamics in electroencephalogram (EEGs) time series recorded from patients with temporal lobe epilepsy. Nevertheless, there are many improvements can be made in algorithms for finding Lyapunov exponents so that the estimation can be more robust, especially with respect to the presence of noise in the EEG. As the complexity of the algorithms for finding Lyapunov exponents with the noise and nonstationarity of EEG, this task requires development of novel techniques and numerous computational experiments.

This chapter is organized as follows. In the next section, we describe Lyapunov exponents and discuss some of the algorithms for finding them. In Section 16.3, an optimization model for estimating Lyapunov exponents is presented and a solution technique for the model is proposed. Brief descriptions of the models used for computational experiments and a comparison of performance, including sensitivity analysis, between the proposed algorithm and the two well-known algorithms are given in Sections 16.4 and 16.5, respectively. The details of the numerical computations are also given in Section 16.5.

16.2 Lyapunov Exponents

Chaos is one type of behavior exhibited by nonlinear dynamical systems, which are systems whose time evolution equations are nonlinear, that is, the dynamical variables describing the properties of the systems (for example, position, velocity, acceleration, pressure) appear in the equations in a nonlinear form. There are several techniques to measure chaos, depending on what one wants to characterize in the chaotic trajectory. Some of the techniques include: simple visual inspection of either the time series represented by a time plot of the trajectory, or the bounded strange attractor reconstructed from the time series; spectral analysis of the time series; Lyapunov exponents; and entropy analysis. Here, we focus

on *Lyapunov exponents*, as this metric has many important characteristics such as invariance to transformations and computability directly from data, without solving the differential or difference equations describing the corresponding dynamical system.

Let us consider any two nearby divergent trajectories originating from a 1D flow (i.e., trajectories in continuous time as described by differential equations). The growth of the difference δ_t between the two nearby trajectories over a time period $\Delta t = t_1 - t_0$ can be described by

$$\delta_t \sim \delta_0 e^{\lambda \Delta t}, \tag{16.1}$$

where λ denotes the systems Lyapunov exponent.

An important reason for using the Lyapunov exponent as a characteristic measure of a dynamical system is its invariance[1] to rescaling, shifts and other transformations of data such as the imprecise reconstruction of a strange attractor from a time series. The fact that trajectories diverge over the course of time would not in itself be very dramatic if it was only very slow, thus we speak of chaos only if this separation is exponentially fast. There are n different Lyapunov exponents for an n-dimensional system, defined as follows: Consider the evolution of an infinitesimal sphere of perturbed initial conditions. During its evolution along the reference trajectory, the sphere will become deformed into an infinitesimal ellipsoid. Let $\delta_k(t)$, $k = 1, 2, 3, \ldots, n$, denote the length of the kth principal axis of the ellipsoid. Thus, the deformation of the sphere corresponds to the stretching, contraction, and rotation of the principal directions. For large t, the diameter of the ellipsoid is controlled by the most positive λ_k. As we shall see later on, λ depends slightly on which trajectory we study, so we should average over many different points on the same trajectory to get the true value of λ (see Table 16.1).

In dissipative systems one can also find a negative maximal Lyapunov exponent which reflects the existence of a stable fixed point. Two trajectories which approach the fixed point also approach each other exponentially fast. If the motion settles down onto a limit cycle, two trajectories can only separate or approach each other slower than exponentially. In this case the maximal Lyapunov exponent is 0 and the motion is called marginally stable. If a deterministic system is perturbed by random noise, on the small scales it can be characterized by a diffusion process, with δ_t growing as $\sqrt{t}$. Thus the maximal Lyapunov exponent is infinite. According to the

Table 16.1: Possible types of motion and the corresponding Lyapunov exponents

Type of motion	Maximal Lyapunov exponent
Stable fixed point	$\lambda < 0$
Stable limit cycle	$\lambda = 0$
Chaos	$0 < \lambda < \infty$
Noise	$\lambda = \infty$

[1] See Oseledec's theorem in [12].

mathematical definition, this is true no matter how small the noise component is, however, we will show later on that Lyapunov exponents of an underlying deterministic system can in fact be measured.

When a system has a positive Lyapunov exponent, there is a time horizon beyond which prediction breaks down.

Wolf et al. [34] proposed the first algorithm for calculating the largest Lyapunov exponent. First, the phase space is reconstructed and the nearest neighbor is searched for one of the first embedding vectors. A restriction must be made when searching for the neighbor: it must be sufficiently separated in time in order not to compute as nearest neighbors successive vectors of the same trajectory. Without considering this correction, Lyapunov exponents could be spurious due to temporal correlation of the neighbors. Once the neighbor and the initial distance L is determined, the system is evolved forward some fixed time (evolution time) and the new distance L' is calculated. This evolution is repeated, calculating the successive distances, until the separation is greater than a certain threshold. Then a new vector (replacement vector) is searched as close as possible to the first one, having approximately the same orientation of the first neighbor. Finally, Lyapunov exponents can be estimated using the following formula:

$$\lambda_1 = \frac{1}{(t_M - t_0)} \sum_{k=1}^{M} \ln \frac{L'(t_k)}{L(t_k - 1)}, \tag{16.2}$$

where k is the number of time propagation steps.

The Wolf algorithm only estimates the largest Lyapunov exponent and not the whole spectrum of exponents. It is said to be sensitive to the number of observations as well as to the degree of measurement or system noise in the observations. This discovery motivated a search for new algorithm designs with improved finite-sample properties. Sano and Sawada [28], Eckmann et al. [5], Abarbanel et al. [1], Rosenstein et al. [27], and Pardalos and Yatsenko [24], among others, came up with improved algorithms for calculating the Lyapunov exponents from observed data.

16.3 An Optimization Approach

In the previous section, we mentioned a number of algorithms that have been proposed for estimating the Lyapunov exponents from a scalar time series. The problem of calculating these exponents can be reformulated as an optimization problem (see Pardalos and Yatsenko [24]), and in these following sections we present an algorithm for its solution which is globally and quadratically convergent. Here, we use well-established techniques from numerical methods for dealing with the optimization problem. We also discuss the computational aspects of this method and the difficulties which inevitably arises when estimating the Lyapunov exponents based on the use of time-delay embedding. Using numerically generated data sets, we consider the influence of the system parameters and the optimization algorithm on the quality of the estimates.

16.3.1 Theory

Let us consider a vector x in the phase space $\mathbb{R}^n$ which can be considered as a solution of a certain dynamical system

$$\dot{x} = f(x,t), \quad x(t_0) = x_0, \tag{16.3}$$

where $f(x,t)$ is a smooth vector field on a manifold $\mathcal{M}$. The vector field f yields a flow $\phi = \{\phi(t)\}$ on the phase space, where $\phi(t)$ is a map,

$$x \mapsto \phi(x,t), \quad t \in \mathbb{R}, \quad x \in \mathbb{R}^n. \tag{16.4}$$

The observed trajectory, starting at x_0, is

$$\{\phi(x_0,t) | t \in \mathbb{R}^+\}. \tag{16.5}$$

To get an information about the time evolution of arbitrarily small perturbed initial conditions, consider the evolution of tangent vectors in the tangent space $\mathcal{T}\mathcal{M}$. It is given by the linearization of Equation (16.3).

The Taylor expansion of $f(\phi(x_0,t))$ for small Δx is

$$f(\phi(x_0,t)) + Df(\phi(x_0,t))\Delta x + \cdots. \tag{16.6}$$

Here $Df(\phi(x_0,t))$ is the local Jacobian matrix of the vector field f at $\phi(x_0,t)$:

$$Df(\phi(x_0,t)) = J(x_0,t) = [(\partial f_i/\partial x_j)|_{\phi(x_0,t)}]. \tag{16.7}$$

For $\Delta x \to 0$, the following first-order approximation holds:

$$\dot{\delta} x = J(x_0,t)\delta x. \tag{16.8}$$

The solution of the linear nonautonomous variational equation (16.8) can be obtained as

$$\delta x(t) = D\phi(x_0,t)\delta x_0, \tag{16.9}$$

where $D\phi(x_0,t) = A(x_0,t) \in \mathbb{R}^{n\times n}$ is the linear operator which maps tangent vector δx_0 to $\delta x(t)$.

The spectrum of the Lyapunov exponents is the set of logarithms of the eigenvalues of the self-adjoint matrix

$$\Lambda_{x_0} = \lim_{t\to\infty} [A(x_0,t)^\top A(x_0,t)]^{1/2t}, \tag{16.10}$$

where $A(x_0,t)^\top$ is the transpose of the matrix $A(x_0,t)$. The existence of the limit in Equation (16.10) is proved by Oseledec's theorem in [12].

Let $E = (e_1,\ldots,e_n)$ be an $n \times n$ matrix, where the column vectors are a basis of the tangent space. If the following limit exists:

$$\lambda_i = \lim_{t\to\infty} \frac{1}{t} \log \|A(x_0,t)e_i\|, \tag{16.11}$$

then λ_i, $i = 1,\ldots,n$, are the Lyapunov exponents. They are ordered by their magnitudes $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$, and if they are independent of x_0, the system is called *ergodic*.

Therefore, one can write $A(x_0,t)$ as the product of $n \times n$ matrices $A(x_j,\Delta t)$, where each one maps $x_j = \phi(x_0, j\Delta t)$ to x_{j+1}:

$$A(x_0,k\Delta t) = \prod_{j=0}^{k-1} A(x_j,\Delta t), \tag{16.12}$$

with $k\Delta t = t$.

16.3.2 Implementation Details

We often have no knowledge of the nonlinear equations of the system which produce the observed time series. But there is a possibility of estimating the linearized flow map $A_{\Delta t} = D\phi(x_j,\Delta t)$ from a single trajectory by using the recurrent structure of strange attractors. Let $\{x_j\}$, $j = 1,2,\ldots$, denote a time series of some physical quantity measured at the discrete time interval Δt, i.e., $x_j = x(t_0 + (j-1)\Delta t)$. Consider a small ball of radius ε centered at the orbital point x_j, and find a set of N difference vectors included in this ball, i.e.,

$$\{y_i\} = \{x_j - x_i |\ \ \|x_j - x_i\| \leq \varepsilon\}, \quad i = 1,2,\ldots,N^2, \tag{16.13}$$

where y_i is the displacement vector between x_j and x_i. Here, $\|\cdots\|$ denotes a usual Euclidean norm defined as follows: $\|w\| = (w_1^2 + w_2^2 + \ldots + w_n^2)^{1/2}$ for some vector $w = (w_1, w_2, \ldots, w_n)$. After the evolution of a time interval $k\Delta t$, $y_i = x_j - x_i$ is mapped to the set

$$\{z_i\} = \{x_{j+k} - x_{i+k}\}, \quad i = 1,2,\ldots,N. \tag{16.14}$$

If the radius ε and the evolution time Δt are small enough for the displacement vectors $\{y_i\}$ and $\{z_i\}$ to be regarded as a good approximation of tangent vectors in the tangent space $\mathcal{TM}$, the evolution of y_i to z_i can be represented by some matrix A_j as

$$z_i = A_j y_i. \tag{16.15}$$

The matrix A_j should be a good approximation of the matrix of linearized flow in Equation (16.9). A plausible procedure for optimal estimation is the least-square

[2] In the implementation, among the N displacement vectors found inside the sphere of radius ε, only five to seven vectors with the smallest norm are chosen. N is often chosen as $d_E \leq N \leq 20$ [28] and is kept at a low value to optimize the efficiency of the algorithm.

algorithm, which minimizes the average of the squared error norm between z_i and $A_j y_i$ with respect to all components of A_j as follows:

$$\min_{A_j} S = \frac{1}{N} \sum_{i=1}^{N} \|z_i - A_j y_i\|^2, \tag{16.16}$$

$$\text{subject to} \quad a_{kl}^j \in \mathbb{R}, \tag{16.17}$$

where a_{kl}^j denotes the (k,l) component of matrix A_j. The evolution times Δt in the renormalization and the approximation process do not necessarily have to be the same, but are chosen equal for convenience.

Each invertible $n \times n$ matrix can be split uniquely into the product of an upper triangular matrix R and an orthogonal matrix Q, such that

$$A_j E_j = Q_j R_j = E_{j+1} R_j, \tag{16.18}$$

with $E_j = (e_j^1, \ldots, e_j^n)$. The matrix Q_j serves as the new basis E_{j+1} and the logarithms of the diagonal elements of R_j are local expanding coefficients, whose time-averaged values are the Lyapunov exponents. Using

$$A(x_0, k\Delta t)E_0 = \prod_{j=0}^{k-1} A(x_j, \Delta t)E_0 = Q_{k-1} \prod_{j=0}^{k-1} R_j \tag{16.19}$$

in Equation (16.10), we obtain

$$\lambda_i = \lim_{k\to\infty} \frac{1}{k\Delta t} \sum_{j=0}^{k-1} \log r_{ii}^j, \tag{16.20}$$

where r_{ii}^j are the diagonal elements of the matrix R_j.

In the numerical procedure, we let A_j operate on an arbitrary chosen set $\{e_i^j\}$, and then renormalize $A_j e_i^j$ to have unit length. Mutual orthogonality of the basis is maintained by using the Gram–Schmidt renormalization procedure. This is repeated for n iterations where Equation (16.20) is computed each time [25].

16.3.2.1 Phase Space Reconstruction

One important reason for using the above approach to computing λ is that the stability of the dynamical system can be determined without actually knowing and solving the underlying differential equations explicitly. This occurs when we obtain a chaotic time series from a dynamical system, reconstruct its strange attractor in the corresponding phase space, and then compute the Lyapunov exponents from the reconstructed strange attractor directly, without its explicit mathematical model.

The most important phase space reconstruction technique is the *method of delays*. The basic idea is very simple. We use the time series data of a single variable

to create a multidimensional reconstruction (embedding) space. If the embedding space is generated properly, the behavior of trajectories in this embedding space will have the same geometric and dynamical properties that characterize the actual trajectories in the full multidimensional phase space of the system. The method of delays was suggested by Packard et al. [17] in 1980 and was put on a firm theoretical basis by Takens [32] in 1981.

From the set of observations $x(t_0 + n\Delta t) = x(n)$, multivariate vectors in d-dimensional space

$$y(n) = (x(n), x(n+\tau), \ldots, x(n+(d-1)\tau)) \tag{16.21}$$

are used to trace out the orbit of the system. The observations, $x(n)$, are a projection of the multivariate phase space of the system onto the 1D axis of the $x(n)$'s. The purpose of time-delay embedding technique is to unfold the projection back to a multivariate phase space that is representative of the original system. In practice, the natural questions of what time delay τ and what embedding dimension d to use in this reconstruction have had a variety of answers. The following sections present the methods used in this chapter for determining τ and d.

The time-delay parameter τ: The choice of time delay is not a straightforward problem. If it is taken too small, there is almost no difference between the different elements of the delay vectors. If on the other hand τ is very large, the different coordinates may be almost uncorrelated. In this case the reconstructed attractor may become very complicated, even if the true underlying attractor is simple. This is typical of chaotic systems, where the autocorrelation function decays fast. Unfortunately, since τ has no relevance in the mathematical framework, there exists no rigorous way of determining its optimal value. At least a dozen different methods have been suggested for the estimation of τ, and since all these methods yield values of similar magnitude, we should estimate τ just by a single preferred tool and work with this estimate [11]. Past studies have made use of the autocorrelation function, but a quite reasonable objection to this procedure is that it is based on linear statistics, not taking into account nonlinear dynamical correlations. Therefore, it is sometimes recommended that one look for the first minimum of the time-delayed *mutual information*. This is the information we already possess about the value of $x(t+\tau)$ if we know $x(t)$.

On the interval explored by the data, we create a histogram for the probability distribution of the data. We denote by p_i the probability that the signal assumes a value inside the ith bin of the histogram, and let $p_{ij}(\tau)$ be the probability that $x(t)$ is in bin i and $x(t+\tau)$ is in bin j. Then the mutual information for time delay τ reads

$$I(\tau) = \sum_{i,j} p_{ij}(\tau) \ln p_{ij}(\tau) - 2\sum_i p_i \ln p_i. \tag{16.22}$$

The value of the mutual information is independent of the particular choice of histogram, as long as it is fine enough.[3] The first minimum of $I(\tau)$ marks the time

[3] Throughout this chapter 512 bins have been used.

delay where $x(t+\tau)$ adds maximal information to the knowledge we have from $x(t)$. This is the time delay used in this chapter.

Embedding dimension: What is the appropriate value of d to use as the embedding dimension? The procedure used here identifies the number of *false nearest neighbors*, points that appear to be nearest neighbors because the embedding space is too small, of every point on the attractor associated with the orbit $y(n)$, $n = 1,2,\ldots,N$. When the number of false nearest neighbors drops to 0, we have unfolded or embedded the attractor in $\mathbb{R}^d$, a d-dimensional Euclidean space.

If we are in d dimensions and we denote the rth nearest neighbor of $y(n)$ by $y^{(r)}(n)$, then from Equation (16.21), the square of the Euclidean distance between the point $y(n)$ and this neighbor is

$$R_d^2(n,r) = \sum_{k=0}^{d-1} [x(n+k\tau) - x^{(r)}(n+k\tau)]^2. \tag{16.23}$$

In going from dimension d to dimension $d+1$ by time-delay embedding we add a $(d+1)$th coordinate onto each of the vectors $y(n)$. This new coordinate is just $x(n+\tau d)$. The Euclidean distance, as measured in dimension $d+1$, between $y(n)$ and the same rth neighbor as determined in dimension d is given by

$$R_{d+1}^2(n,r) = R_d^2(n,r) + [x(n+\tau d) - x^{(r)}(n+\tau d)]^2. \tag{16.24}$$

A natural criterion for catching embedding errors is that the increase in distance between $y(n)$ and $y^{(r)}(n)$ is large when going from dimension d to $d+1$. The increase in distance can be stated quite simply from Equations (16.23) and (16.24). We state this criterion by designating as a false neighbor any neighbor for which

$$\sqrt{\frac{R_{d+1}^2(n,r) - R_d^2(n,r)}{R_d^2(n,r)}} = \frac{|x(n+\tau d) - x^{(r)}(n+\tau d)|}{R_d(n,r)} > R_{\text{tol}}, \tag{16.25}$$

where R_{tol} is some threshold. In practical settings the number of data points is often not large, and the following criterion handles the issue of limited data set size: If the nearest neighbor to $y(n)$ is not close ($R_d(n) \approx R_A$) and it is a false neighbor, then the distance $R_{d+1}(n)$ resulting from adding on a $(d+1)$th component to the data vectors will be $R_{d+1}(n) \approx 2R_A$ [13]. That is, even distant but nearest neighbors will be stretched to the extremities of the attractor when they are unfolded from each other, if they are false nearest neighbors. We write this second criterion as

$$\frac{R_{d+1}(n)}{R_A} > A_{\text{tol}}, \tag{16.26}$$

where R_A denotes the size of the attractor. Both criterions in Equations (16.25) and (16.26) are used jointly throughout the determination of d_E (also see Fig. 16.1). As a measure of R_A we have chosen the standard deviation $\sigma(\mathbf{x})$ of the observed data $\mathbf{x}$ according to [12]. This source also gives us the recommended values $R_{\text{tol}} = 15.0$ and $A_{\text{tol}} = 2.0$.

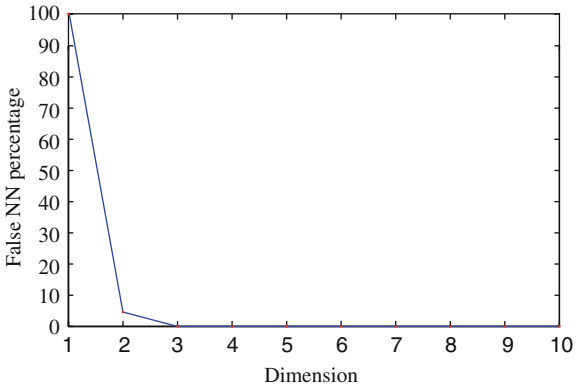

Fig. 16.1: The percentage of false nearest neighbors for 10,000 data points from the Lorenz equations (see Section 2.2.2). The data were output at $\Delta t = 0.01$ during the integration. A time lag $\tau = 12\Delta t = 0.12$, which is the location of the first minimum in the average mutual information for this system, was used in forming the time-delayed vectors.

From the point of view of the mathematics of the embedding process it does not matter whether one uses the minimum embedding dimension d_E or any $d \geq d_E$, since once the attractor is unfolded, the theorem's work is done. For a physicist the story is quite different. Working in any dimension larger than the minimum required by the data leads to excessive computation when investigating the Lyapunov exponents. It also enhances the problem of contamination by roundoff or instrumental error since this noise will populate and dominate the additional $d - d_E$ dimensions of the embedding space where no dynamics is operating. We should add that in going through the data set and determining which points are near neighbors of the point $y(n)$ we use the sorting method of a k-dimensional tree to reduce the computation time from $\mathcal{O}(n^2)$ to $\mathcal{O}(N \log_{10}(N))$.

16.4 Models Used in the Computational Experiments

We evaluate the algorithm performance using the signals simulated from four well-known dynamical mathematical models: Lorenz, Rössler, Hénon, and Hénon–Heilers. Brief descriptions of the models are given below.

16.4.1 Lorenz Attractor

We begin our study of Lyapunov exponents with the Lorenz equations:

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

Here $\sigma, R, b > 0$ are parameters. Edward Lorenz derived this 3D system from a simplified model of convection rolls in the atmosphere. Roughly one can say that the equations describe the flow of a fluid in a box which is heated along the bottom. This simple-looking deterministic system can have extremely erratic dynamics, the solutions oscillate irregularly over a wide range of parameters. In his original experiments he fixed the values of the parameters to $\sigma = 10$, $R = 28$, and $b = 8/3$ for which the system has chaotic behavior. These are the parameter values we will be using in the comparative study between the optimization approach described in [25] and the algorithms in [28, 34]. Figure 16.2a shows a 3D view of the Lorenz system.

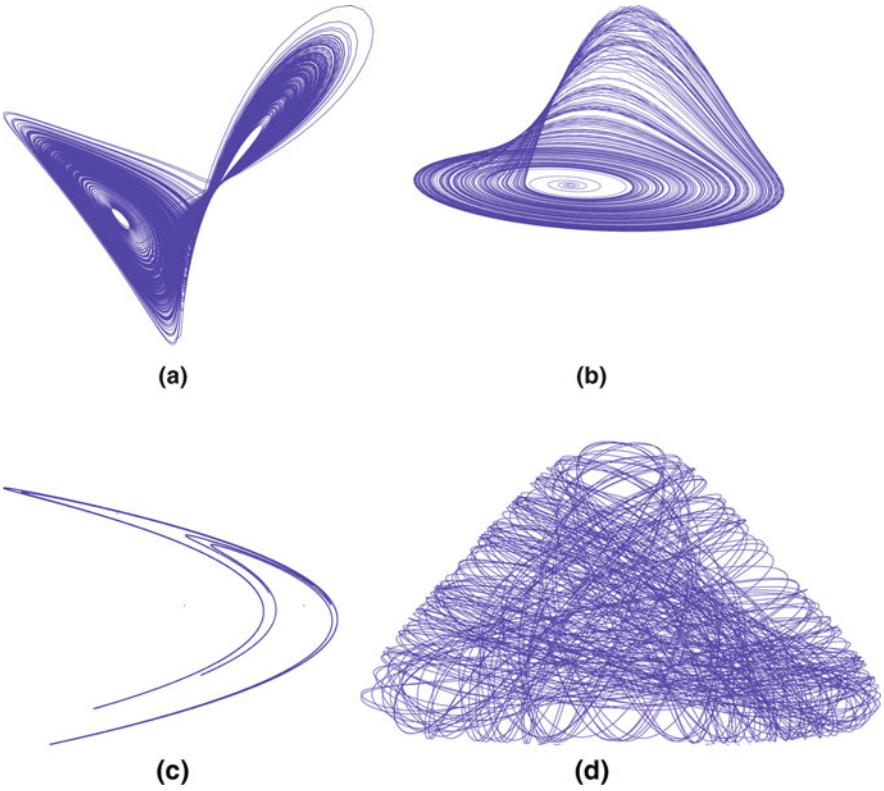

Fig. 16.2: The algorithm has been tested on the (**a**) Lorenz attractor, (**b**) Rössler attractor, (**c**) Hénon map and (**d**) Hénon–Heiles system (plotted in the x–y plane).

16.4.2 Rössler Attractor

In 1976, the Swiss mathematician Otto Rössler was studying oscillations in chemical reactions and discovered another set of equations with a chaotic attractor:

$$\dot{x} = -y - z$$
$$\dot{y} = x + ay$$
$$\dot{z} = b + z(x - c).$$

Both the Lorenz and Rössler equations are involved with the study of Navier–Stokes equations. Rössler is acclaimed to have used the parameter values $a = 0.2$, $b = 0.2$, and $c = 5.7$, which we will also use. This system of equations looks easier than the Lorenz system, with only one nonlinearity xz, but it is harder to analyze. Figure 16.2b shows the Rössler attractor.

16.4.3 Hénon Map

The Hénon map was devised by the theoretical astronomer Michel Hénon to illuminate the microstructure of strange attractors in 1976. Previous scientists had encountered numerical difficulties when tackling the Lorenz system, so instead Hénon sought a mapping that captured its essential features but which also had an adjustable amount of dissipation. Hénon chose to study mappings rather than differential equations because maps are faster to simulate and their solutions can be followed more accurately and for a longer time. The Hénon map is given by

$$x_{n+1} = y_n + 1 - ax_n^2$$
$$y_{n+1} = bx_n,$$

where a and b are adjustable parameters which are chosen as $a = 1.4$, $b = 0.3$.

16.4.4 The Hénon–Heiles Equations

The Hénon–Heiles model was introduced in 1964 by Michel Hénon and Carl Heiles as a model for the motion of a star inside a galaxy. With the Hamiltonian

$$H = \frac{1}{2}\left(p_1^2 + q_1^2 + p_2^2 + q_2^2\right) + q_1^2 q_2 - \frac{1}{3}q_2^3,$$

and if we let $q_1 = x$, $q_2 = y$, $p_1 = p_x$, and $p_2 = p_y$, then the Hamiltonian can be interpreted as a model for a single particle moving in two dimensions under the action of a force described by a potential energy function $V(x,y)$ [9]. Hamilton's equations for this system lead to the following equations for the dynamics of the system:

$$\dot{x} = \frac{\partial H}{\partial p_x} = p_x \qquad \dot{p}_x = -\frac{\partial H}{\partial x} = -x - 2xy$$
$$\dot{y} = \frac{\partial H}{\partial p_y} = p_y \qquad \dot{p}_y = -\frac{\partial H}{\partial y} = -y - x^2 + y^2.$$

It can be shown that the potential supports bounded motion for the particle for H$< 1/6$. Thus, to fulfill this condition, the initial values are chosen as $x_0 = 0$, $y_0 = -0.15$, $p_{x,0} = 0.50$, and $p_{y,0} = 0$.

16.5 Computational Experiments

This section presents the main results of the implemented algorithm and compares these with two other algorithms described in [28,34]. Some enhancements are made to the algorithm, which is finally tested for robustness against change in parameter values and noise.

16.5.1 Numerical Computations

The differential equations are integrated with the *Runge–Kutta (RK4)* method using a fixed step δt. This method is reasonably simple and robust, even without the adaptive step-size routine. The RK4 method is a fourth-order method, meaning that the error per step is $\mathcal{O}((\delta t)^5)$, while the total accumulated error has order $\mathcal{O}((\delta t)^4)$.

Table 16.2 summarizes the values for the computed Lyapunov exponents, which has been estimated using three different implemented algorithms described in

Table 16.2: Results of the preliminary computational experiments for $n = 2,000$. For the Lorenz attractor, the parameter values have been chosen as: $\tau = 9$, $\Delta t = 5$, $\delta t = 0.01$, $\varepsilon = 1.20$. The parameter values for the Rössler attractor are: $\tau = 6$, $\Delta t = 5$, $\delta t = 0.12$, $\varepsilon = 0.28$

System with initial condition		$n = 2,000$			$n = 4,000$		
		Pardalos	Sano	Wolf	Pardalos	Sano	Wolf
Lorenz							
$x_0 = 0$	λ_1	1.08547	1.08546	0.86241	1.15313	1.15312	0.92140
$y_0 = 1.0$	λ_2	−0.30421	−0.30421		−0.31226	−0.31232	
$z_0 = 0$	λ_3	−10.61753	−10.61754		−10.57328	−10.57319	
Rössler							
$x_0 = 0.1$	λ_1	0.06928	0.06924	0.06642	0.06891	0.06881	0.06593
$y_0 = 0.1$	λ_2	0.00132	0.00156		0.00122	0.00131	
$z_0 = 0.1$	λ_3	−1.27202	−1.27183		−1.28908	−1.28849	
Hénon							
$x_0 = 0.1$	λ_1	0.41672	0.41672	0.40445	0.41669	0.41672	0.41547
$y_0 = 0.1$	λ_2	−1.57647	−1.57647		−1.57765	−1.57773	
Hénon–Heiles							
$x_0 = 0$	λ_1	0.15207	0.15209	0.16012	0.14875	0.14874	0.15493
$y_0 = -0.15$	λ_2	0.01950	0.01949		0.01382	0.01385	
$p_{x,0} = 0.50$	λ_3	−0.04242	−0.04245		−0.04874	−0.04871	
$p_{y,0} = 0$	λ_4	−0.23309	−0.23395		−0.22142	−0.22141	

[24, 28, 34]. The rows show the Lyapunov exponents in a decreasing manner, i.e., $\lambda_1 > \lambda_2 > \ldots > \lambda_n$, for given initial conditions. The algorithm by Wolf et al. [34] only gives us an estimate of the largest exponent. As mentioned earlier, a positive Lyapunov exponent measures sensitive dependence on initial conditions, or how much our forecasts can diverge based upon different estimates of starting conditions. Another way to view Lyapunov exponents is the loss of predictability as we look forward in time. Thus, it is interesting to know a measure of information loss for avoiding possible misinterpretations.

If we assume that the true starting point x_0 of a time series is possibly displaced by an ε, we know only the information area I_0 about the starting point. After some steps the time series is in the information area at time t, I_t. The information about the true position of the data decreases due to the increase of the information area. Consequently, we get a bad predictability. The largest Lyapunov exponent can be used for the description of the average information loss; $\lambda_1 > 0$ leads to bad predictability. Therefore, the exponent values in Table 16.2 are given in units of nats/s.[4]

Of all the N displacement vectors found inside the sphere of radius ε, only five to seven vectors with the smallest norm are chosen. This has practically no noticeable effect on the exponent values, but speeds up the algorithm. It is further enhanced by introducing another constraint which enables us to search for displacement vectors close in phase space (Equation (16.13)), but far away in time

$$|t_j - t_i| > \frac{\varepsilon}{\delta t}, \quad \forall i, j, \quad i \neq j. \tag{16.27}$$

The Gauss–Newton algorithm is used to solve the nonlinear least-squares problem in Equation (16.16), while Sano [28] uses a linear approach to solve the same problem. By examining Table 16.2 we see that there is hardly any difference in the estimated exponent values between Pardalos's algorithm and the one described by Sano. This behavior is due to the small values of the evolution time Δt. During this short evolution, the mapping between t_j and $t_j + \Delta t$ does not show any stronger nonlinear properties, therefore, the results are similar. The value Δt should be kept small enough so that orbital divergence is monitored at least a few times[5] per (mean) orbit. A larger Δt has been shown to increase the difference between these two algorithms, as expected.

The displacement vectors y_i have been chosen to lie inside a sphere of radius $\varepsilon \lessapprox 0.02 L_A$, where L_A is the horizontal extent of the attractor. The choice of ε is good as long as we fulfill the condition of finding a minimum of five vectors inside the sphere. Though theory says this value should be infinitesimal, the optimization algorithm described in Section 16.2 is robust against small increase in ε. Figure 16.3 shows how the Lyapunov exponents for the examined systems converge.

The results from Pardalos's and Sano's algorithms, though different from the estimated values computed by the Wolf algorithm, are in good agreement with other numerical experiments performed in [30, 27, 4, 26, 8].

[4] 1 nat/s $\approx$ 1.44 bits/s.

[5] We have computed Equation (16.16) between 30 and 40 times per mean orbit.

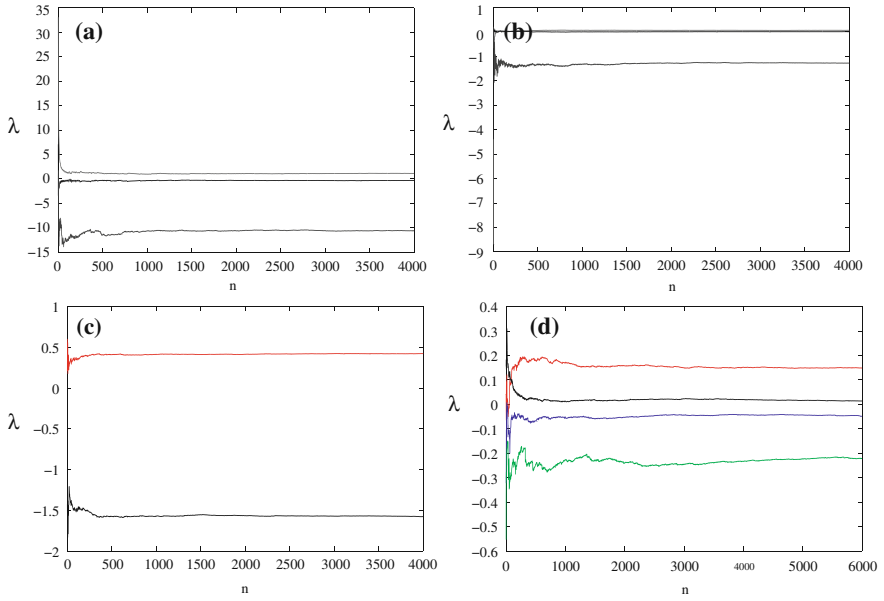

Fig. 16.3: Convergence of the spectrum of Lyapunov exponents for (**a**) the Lorenz attractor, (**b**) Rössler attractor, (**c**) Hénon map, and (**d**) the Hénon–Heiles system. The graphs show the results of the algorithm described in Section 16.2.

16.5.2 Sensitivity Analysis

Two parameters, namely the evolution time Δt and the time delay τ, have been chosen for further investigation for robustness. We mentioned in Section 2.2.2 that τ is determined as the lag which gives us the first minimum for the mutual average information for our observed data. Figure 16.4 shows how the spectrum of Lyapunov

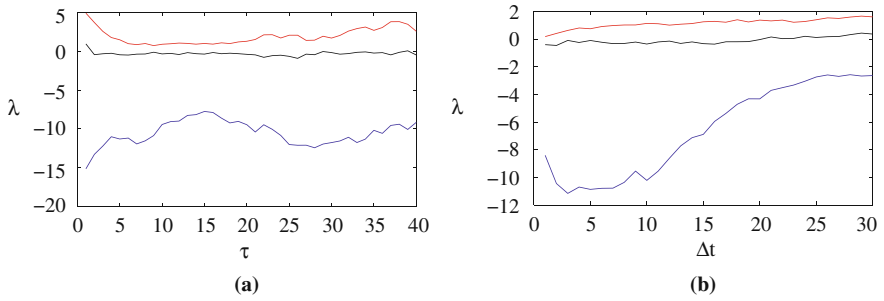

Fig. 16.4: The Lyapunov exponents as a function of (**a**) τ and (**b**) Δt for the Lorenz attractor.

exponent values depend on Δt and τ. The two largest Lyapunov exponents, λ_1 and λ_2, seem to be rather stable in the vicinity of the chosen values $\Delta t = 5$ and $\tau = 9$ while λ_3 is unstable for all values in the interval $1 < \Delta t < 30$, $1 < \tau < 40$. What is important to be reminded of here is that the systems are extremely sensitive due to their chaotic nature, and the observed "errors" due to perturbations of parameter values do not have to be entirely blamed on this specific algorithm. The science of choosing the right parameter values for these kind of problems is not general and depends on which system you are examining.

Figure 16.5 shows how the Lyapunov exponents for the Rössler attractor depend on τ and Δt. Again, λ_1 and λ_2 are stable in the neighborhood of the chosen values $\tau = 6$ and $\Delta t = 5$, while λ_3 is very irregular throughout the interval. This behavior is given a deeper theoretical explanation in [28].

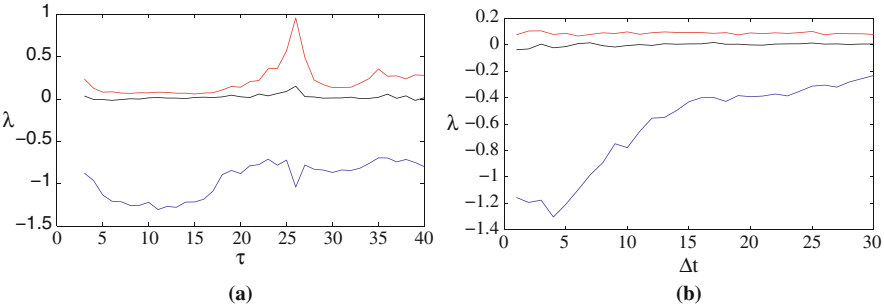

Fig. 16.5: The Lyapunov exponents as a function of (**a**) τ and (**b**) Δt for the Rössler attractor.

The algorithm has also been tested for noise contaminated data. We have added Gaussian white noise $w(t)$ to the solutions of the Lorenz and Rössler systems according to

$$\begin{aligned} x_{\text{noise}} &= x_{\text{clean}} + w(t)\sigma(x)s \\ y_{\text{noise}} &= y_{\text{clean}} + w(t)\sigma(y)s \\ z_{\text{noise}} &= z_{\text{clean}} + w(t)\sigma(z)s, \end{aligned}$$

where s is a scaling factor for the standard deviation σ.

Figure 16.6 shows the dependence between the Lyapunov exponents and the scaling factor within the interval $0 \leq s \leq 0.1$. We see that the exponents λ_1 and λ_2 for both systems are quite stable within the interval $0 < s < 0.01$, i.e., they are not sensitive to data contaminated with up to 0.01σ of Gaussian white noise. Once again we can confirm the sensitive nature of the smallest Lyapunov exponent λ_3 of the Lorenz attractor.

Many physical signals, including the time series studied in this chapter, are fundamentally different from linear time-invariant signals in that they are invariant to

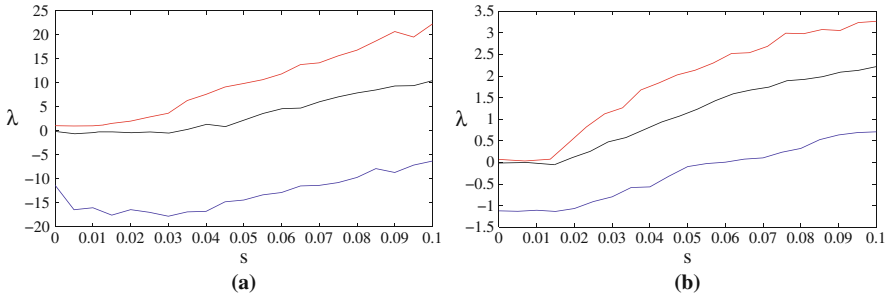

Fig. 16.6: The Lyapunov exponent values as a function of the scaling factor s for (**a**) the Lorenz attractor and (**b**) the Rössler attractor.

scale rather than to translation. The signals are often mixed with noise, and the separation may be very difficult if both the signal and the noise are broadband. The problem becomes inherently difficult when the signal is chaotic because its power spectrum is indistinguishable from a broadband noise, as in our case. Since there is a strong relationship between these *fractal* signals and the wavelet transform, the latter appears to be the natural signal processing technique, just as the Fourier transform is natural for the linear time-invariant signals [14]. Many new filtering techniques to handle these problems are still under development.

16.6 Summary and Conclusion

The Lyapunov exponents are conceptually the most basic indicators of deterministic chaos of dynamical systems. For the analysis of such dynamics, many numerical algorithms to determine the spectrum of the Lyapunov exponents have been proposed. In this chapter, we considered an optimization technique for calculating tangent maps with the aim of developing a robust algorithm. We have described a method which is shown to behave well in the perturbation of certain parameter values, but slightly sensitive in the presence of noise. This method uses the Gauss–Newton algorithm to solve the least-squares problem that arises, which is no more complicated to implement than the linear method. By using the new optimization method, we could obtain good estimates of the Lyapunov spectrum from the observed time series in a very systematic way.

References

1. Abarbanel, H., Brown, R., Kennel, M. Variation of Lyapunov exponents in chaotic systems: Their importance and their evaluation using observed data. J Nonlinear Sci **2**, 343–365 (1992)
2. Chen, G., Lai, D. Making a dynamical system chaotic: Feedback control of Lyapunov exponents for discrete-time dynamical systems. IEEE Trans Circuits Syst I Fundam Theory Appl **44**, 250–253 (1997)

3. Cvitanovic, P., Artuso, R., Mainieri, R., Tanner, G., Vattay, G. Classical and Quantum Chaos. 10th ed., ChaosBook.org Niels Bohr Institute, Copenhagen (2003) Webbook: http://chaosbook.org/.
4. Djamai, L., Coirault, P. Estimation of Lyapunov exponents by using the perceptron. Proceedings of the American Control Conference **6**, 5150–5155 (2002)
5. Eckmann, J.P., Kamphorst, S.O., Ruelle, D., Ciliberto, S. Lyapunov exponents from time series. Phys Rev A **34**, 4971–4979 (1986)
6. Eckmann, J.P., Ruelle, D. Ergodic theory of chaos and strange attractors. Rev Modern Phys **57**, 617–657 (1985)
7. Elger, C.E., Lehnertz, K. Seizure prediction by non-linear time series analysis of brain electrical activity. Eur J Neurosci **10**, 786–789 (1998)
8. Golovko, V. Estimation of the Lyapunov spectrum from one-dimensional observations using neural networks. Proceedings of the Second IEEE International Workshop on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications, 95–98 (2003)
9. Hilborn, R.C. Chaos and Nonlinear Dynamics. Oxford University Press, New York (2000)
10. Iasemidis, L.D., Pardalos, P.M., Sackellares, J.C., Yatsenko, V.A. Global optimization approaches to reconstruction of dynamical systems related to epileptic seizures. In: Fotiadis, D., Massalas, C.V. (eds.) Scattering and Biomedical Engineering: Modeling and Applications, World Scientific, Singapore, pp. 308–318 (2002)
11. Kantz, H., Schreiber, T. Nonlinear Time Series Analysis. Cambridge University Press, New York (1997).
12. Kelliher, J.: Lyapunov Exponents and Oseledec's Multiplicative Ergodic Theorem. http://www.ma.utexas.edu/users/kelliher/Geometry/Geometry.html (2005)
13. Kennel, M.B., Brown, R., Abarbanel, H.D.I. Determining embedding dimension for phase-space reconstruction using a geometrical construction. Phys Rev A **45**, 3403–3411 (1992)
14. Kinsner, W. Characterizing chaos through Lyapunov metrics. Second IEEE International Conference on Cognitive Informatics (ICCI'03), 189–201 (2003)
15. Lehnertz, K., Andrzejak, R.G., Arnhold, J., Kreuz, T., Mormann, F., Rieke, C., Widman, G., Elger, C. Nonlinear EEG analysis in epilepsy: Its possible use for interictal focus localization, seizure antipilation, and prevention. J Clin Neurophysiol **18**, 209–222 (2001)
16. Nair, S. Brain Dynamics and Control with Applications in Epilepsy. Dissertation at the University of Florida, Gainesville, FL (2006)
17. Packard, N.H., Crutchfield, J.P., Farmer, J.D., Shaw, R.S. Geometry from a time series. Phys Rev Lett **45**, 712–715 (1980)
18. Pardalos, P.M., Boginski, V., Vazakopoulos, A. (eds.) Data Mining in Biomedicine. Springer, New York (2007)
19. Pardalos, P.M., Chaovalitwongse, W., Iasemidis, L.D., Sackellares, J.C., Shiau, D-S., Carney, P.R., Prokopyev, O.A., Yatsenko, V.A. Seizure warning algorithm based on optimization and nonlinear dynamics. Math Program **101**, 365–385 (2004)
20. Pardalos, P.M., Principe, J. (eds.) Biocomputing. Kluwer Academic Publishers, Dordrecht (2002)
21. Pardalos, P.M., Sackellares, C., Carney, P., Iasemidis, L. (eds.) Quantitative Neuroscience. Kluwer Academic Publishers, Dordrecht (2004)
22. Pardalos, P.M., Sackellares, J.C., Iasemidis, L.D., Yatsenko, V.A., Yang, M., Shiau, D.-S., Chaovalitwongse,W. Statistical information approaches to modelling and detection of the epileptic human brain. Comput Stat Data Anal **43**(1), 79–108 (2003)
23. Pardalos, P.M., Sackellares, J.C., Yatsenko, V.A., Butenko, S.I. Nonlinear dynamical systems and adaptive filters in biomedicine. Ann Oper Res **119**, 119–142 (2003)
24. Pardalos, P.M., Yatsenko, V.A. Optimization approach to the estimation and control of Lyapunov exponents. J Optim Theory Appl **128**, 29–48 (2006)
25. Pardalos, P.M., Yatsenko, V.A., Sackellares, J.C., Shiau, D.-S., Chaovalitwongse, W., Iasemidis, L. Analysis of EEG data using optimization, statistics, and dynamical systems techniques. Comput Stat Data Anal **44**, 391–408 (2003)
26. Ramasubramanian, K., Sriram, M.S. A comparative study of computation of Lyapunov spectra with different algorithms. Physica D **139**, 72–86 (2000)

27. Rosenstein, M.T., Collins, J.J., De Luca, C.J. A practical method for calculating largest Lyapunov exponents from small data sets. Physica D **65**, 117–134 (1993)
28. Sano, M., Sawada, Y. Measurement of the Lyapunov spectrum from a chaotic time series. Phys Rev Lett **55**, 1082–1085 (1985)
29. Serfaty DeMarkus, A. Detection of the onset of numerical chaotic instabilities by Lyapunov exponents. Discrete Dyn Nat Soc, **6**, 121–128 (2001)
30. Shimada, I., Nagashima, T. A Numerical approach to ergodic problem of dissipative dynamical systems. Progr Theor Phys, **61**, 1605–1616 (1979)
31. Strogatz, S.H. Nonlinear Dynamics and Chaos. Perseus Books Publishing, Cambridge, MA (1994)
32. Takens, F. Detecting strange attractors in turbulence. In: Rand, D.A., Young, L.-S. (eds.) Dynamical Systems and Turbulence, *Lecture Notes in Mathematics*, Vol. **898**, pp. 366–381. Springer-Verlag, New York (1981)
33. Wiesel,W.E. Modal feedback control on chaotic trajectories. Phys Rev E **49**, 1990–1996 (1994)
34. Wolf, A., Swift, J.B., Swinney, H.L., Vastano, J.A. Determining Lyapunov exponents from a time series. Physica D **16**, 285–317 (1985)
35. Zeni, A.R., Gallas, J.A.C. Lyapunov exponents for a Duffing oscillator. Physica D **89**, 71–82 (1995)