



Reconstructing a dynamics from a scalar time series

G. Gouesbet, L. Le Sceller, Christophe Letellier, R. Brown, J.R. Buckler, Z. Kolláth

► To cite this version:

G. Gouesbet, L. Le Sceller, Christophe Letellier, R. Brown, J.R. Buckler, et al.. Reconstructing a dynamics from a scalar time series. *Annals of the New York Academy of Sciences*, 1997, 808, pp.25-50. 10.1111/j.1749-6632.1997.tb51652.x . hal-01596857

HAL Id: hal-01596857

<https://hal.science/hal-01596857>

Submitted on 18 May 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Reconstructing a Dynamics from a Scalar Time Series

G. GOUESBET,^a L. LE SCHELLER,^a C. LETELLIER,^a
R. BROWN,^b J. R. BUCHLER,^{c,d} AND Z. KOLLÁTH^e

^a *LESP/URA CNRS 230/INSA de Rouen/BP 08
Place Emile Blondel*

76131 Mont Saint-Aignan Cédex, France

^b *Institute for Nonlinear Science
University of California
La Jolla, California 92093-0402*

^c *Department of Physics
University of Florida
Gainesville, Florida 32611*

INTRODUCTION

Chaotic time series data are now commonly observed in a large variety of fields. The analysis of such time series and the extraction of physical information is therefore a topic of growing interest. The techniques have matured to the point where they have become of practical use. It is the purpose of this review to discuss some such applications. For excellent general recent reviews of the available techniques we refer the reader to references 1 and 2.

This type of analysis presumes that the data have been generated by a low dimensional physical dynamics of *a priori* unknown nature, and that the signal may have been contaminated by noise of various sorts. In most applications we have at our disposal only a single measured quantity, i.e., a scalar time series. Fortunately one can exploit a very powerful redundancy principle that is discussed in the pioneering papers by Packard *et al.*³ and by Takens⁴ (cf. also reference 5) which allow one to reconstruct an image of the *whole* dynamics in an embedding space. These embedding theorems allow one then to extract useful information about the unknown dynamics from this single measured variable.

The first step in the analysis then consists of a phase-space reconstruction in an embedding space. In order to achieve it, we have to determine the embedding dimension d_E of the reconstructed phase space (also called state space). Of particular interest is of course the minimum embedding dimension, i.e., the lowest dimension for which the reconstruction gives an unambiguous image of the attractor. The reconstruction may be carried out in a number of ways, but it is safest to use several techniques because none of them alone is foolproof (for excellent reviews, cf. references 1 and 2). All these methods rely on delay coordinates, i.e., constructing the

^dSupported by the National Science Foundation.

vectors of length (dimension) d_E

$$\mathbf{Y}^n = (s(t_n), s(t_{n-\tau}), s(t_{n-2\tau}), \dots, s(t_{n-(d_E-1)\tau})) \quad (1)$$

from a scalar time series $\{s_n\} = \{s(t_n)\}$ that is sampled stroboscopically, i.e., at equal time intervals $t_n = t_o + n\delta t$, where the *delay* $\Delta = \tau\delta t$. We note that in some applications, such as astronomical observations, such a regular sampling may not be possible, and it is generally necessary to interpolate the data prior to the analysis.

Once the dimension d_E is known, we have the freedom to build a coordinate set to span the reconstructed state space. There exists an infinity of linear combinations of these delay vectors that are mathematically equivalent (e.g., reference 6), but because of unavoidable noise, in practice a suitable choice is important. A prolific literature is devoted to this issue. Many applications use the bare \mathbf{Y}_n because this way the d_E coordinates all have the same noise properties, but other linear combinations have also been used, such as derivative coordinates, either direct differences, discrete Legendre polynomials⁶ or Krawchuck polynomials (Auvergne, priv. comm.), etc. We note also that some care must be exercised because certain types of filtering of the data can augment the apparent dimension of the dynamics (e.g., reference 1). In whatever form delay coordinates are used, a crucial step is to determine an optimal value of the time delay Δ . However, this optimal value depends on the application. One popular choice is to use the value for which the mutual information⁷ has its first zero. When polynomial maps are constructed from real data we shall see that the optimal value is a compromise that limits the nonlinearity of the map and minimizes the effect of noise. The method of *false nearest neighbors*¹ provides an efficient general guideline both for the dimensionality of the attractor and for the delay.

Once we have reconstructed the attractor in an embedding phase space, the next step is to obtain a set of equations which models the evolution of the system in the phase space. More particularly, we are interested in *global models* here, i.e., models which describe the dynamics on the whole state space, rather than models which approximate the dynamics in local balls. These methods then allow what we call a *global vector field or flow reconstruction*.

The chapter is organized as follows. First, we briefly review the general methods based on delay coordinates⁸⁻¹¹ and on derivative coordinates¹² for global vector field reconstructions. Secondly, a section is devoted to applications of these methods to real, i.e., experimental and observational data. Thirdly, a recent extension of global vector field reconstruction methods, which provides a set of equations incorporating a control parameter with an explicit physical meaning, is examined. In particular, preliminary results with data arising from an electrodisolution experiment are discussed. Finally, the conclusions are presented.

THE RECONSTRUCTION METHODS

Let us consider an *a priori* unknown nonlinear dynamical system defined by a set of autonomous ODEs:

$$\dot{\mathbf{x}} = \mathbf{g}(\mathbf{x}; \boldsymbol{\mu}) \quad (2)$$

in which $\mathbf{x} \in \mathbb{R}^d$ and \mathbf{g} is the unknown true vector field associated with the underlying dynamics of the physical system, and we call the solution vector $\mathbf{x}(t; \boldsymbol{\mu})$ the *state vector* which describes a *trajectory* in phase space. The quantity $\boldsymbol{\mu} \in \mathbb{R}^p$ is the parameter vector with p components, which for a given time series is assumed to be constant in this section. We note that in many situations we do not have an *a priori* knowledge of the dimension d of the dynamics. The dynamics can equivalently be described by a *map*

$$\mathbf{X}^{n+1} = \mathbf{G}(\mathbf{X}^n; \boldsymbol{\mu}), \quad (3)$$

where $\mathbf{X}^n \equiv \mathbf{x}(t_n) \in \mathbb{R}^d$, with $t_n = n\delta t$, represents a stroboscopic sampling of the trajectory at the equal time intervals δt . The measured time series s_n is assumed to be a generic function of the trajectory, $s_n = s(\mathbf{x}(t_n))$.

Depending on the applications, our goal is either to reconstruct a corresponding map \mathbf{F} in a d_E dimensional embedding space with the help of the given delay vectors $\mathbf{Y}^n(1)$

$$\mathbf{Y}^{n+1} = \mathbf{F}(\mathbf{Y}^n; \boldsymbol{\mu}) \quad (4)$$

or to reconstruct a flow \mathbf{f}

$$\dot{\mathbf{y}} = \mathbf{f}(\mathbf{y}; \boldsymbol{\mu}) \quad (5)$$

where $\mathbf{y}(t_n) = \mathbf{Y}^n$. The embedding theorems tell us that there is a one-to-one correspondence between the embedded flow \mathbf{f} of (4) and the physical flow \mathbf{g} (2). We refer to the latter as *physical* because it represents the number of variables that are required to fully describe the dynamics. It is however important to distinguish between the dimensions of the two flows. Because the embedding can cause spurious cusps or intersections, in principle d_E is greater or equal to d ; in practice it is often found, or assumed, that $d = d_E$ (perhaps because the cusps and intersection points are rare and do not appear to cause havoc when the map is iterated to create synthetic signals).

Maps

The reconstruction can be very conveniently made with polynomial nonlinearities in such a way that the components of $\mathbf{F}(\mathbf{Y})$ take the form of multivariate polynomials. The Weierstrass theorem¹³ guarantees the convergence of the approximation of any smooth function by a polynomial expansion. Consequently, using polynomial expansions to build global models of dynamical systems comes into mind in a natural way. Polynomial techniques for example have been extensively used for *local* modeling, in which the vector fields $\mathbf{F}(\mathbf{Y})$ (or $\mathbf{f}(\mathbf{y})$) are locally approximated by polynomials whose coefficients are determined by least-squares (LS) fits. In contrast, when a global model is desired, the situation becomes more difficult due to the large number of required data points and to the need for the use of higher-order polynomials. In order to overcome such difficulties, several groups have independently proposed using orthogonal, or better, orthonormal polynomials generated by the invariant density on the attractor.^{8,9,12,14,15} There are two equivalent ways to

construct these polynomials, namely, by a Gram-Schmidt orthonormalization or by a LS fitting with monomials, both of which give complementary insight into the data requirements.

Gram-Schmidt Construction of Orthogonal Polynomials

Interestingly the determination of the coefficients of the polynomials and therefore of the coefficients of the reconstructed vector field, $\tilde{\mathbf{F}}(\mathbf{y})$, only requires the computation of moments of the data. (Henceforth we shall drop the tilde from the reconstructed maps and vector fields when there is no confusion possible.) More specifically, a basis of orthonormal multivariate polynomials $\{\phi_k(\mathbf{y}) = \phi_k(y_1, y_2, \dots, y_{d_E})\}$ on \mathbb{R}^{d_E} can be built with a conventional Gram-Schmidt (GS) procedure. According to this procedure, the polynomials $\phi_k(\mathbf{y})$ read as:

$$\phi_k(\mathbf{y}) = \frac{\phi_k^*(\mathbf{y})}{(\phi_k^*, \phi_k^*)}, \quad (6)$$

in which $\phi_k^*(\mathbf{y})$ are defined as:

$$\begin{cases} \phi_1^* = P_1 \\ \phi_k^* = P_k - \sum_{\alpha=1}^{k-1} (P_k, \phi_\alpha) \phi_\alpha, \quad k > 1, \end{cases} \quad (7)$$

in which P_k 's designate the leading multivariate monomials and $(,)$ is a scalar product. The orthonormality condition is $(\phi_i, \phi_j) = \delta_{ij}$. The polynomials of the orthonormal basis are then found to be given by expansions on the monomials P_k with expansion coefficients defined as:¹⁶

$$A_\alpha^k = A_\alpha^{*k} \left/ \left[\sum_{\beta=1}^k \sum_{\gamma=1}^k A_\beta^{*k} A_\gamma^{*k} (P_\beta, P_\gamma) \right]^{1/2} \right., \quad (8)$$

where

$$A_\alpha^{*k} = - \sum_{\beta=\alpha}^{k-1} A_\alpha^\beta \left[\sum_{\gamma=1}^\beta A_\gamma^\beta (P_k, P_\gamma) \right]. \quad (9)$$

Each component F_i of the vector field $\mathbf{F}(\mathbf{y})$ which describes the dynamics is approximated as:

$$F_i = \sum_{j=1}^P c_{ij} \phi_j, \quad (10)$$

in which the c_{ij} are the Fourier coefficients for the i th component F_i of the vector field \mathbf{F} , which are:

$$c_{ij} = (F_i, \phi_j). \quad (11)$$

In practice, the expansion is terminated at some maximum polynomial order P .

Least-Squares Construction of Orthogonal Polynomials

A mathematically equivalent alternative to constructing the orthogonal polynomials and the expansion coefficients is to perform a linear LS fit with all possible monomials up to order P , i.e., to minimize the expression

$$S = \sum_{n=1}^{N-1} |\mathbf{Y}^{n+1} - \mathbf{F}(\mathbf{Y}^n)|^2, \quad (12)$$

where N is the number of delay vectors. The vector function \mathbf{F} is given by

$$\mathbf{F}(\mathbf{Y}) = \sum_{\mathbf{k}} \mathbf{A}^{\mathbf{k}} Y_1^{k_1} Y_2^{k_2} Y_3^{k_3} \dots, \quad (13)$$

where the Y_i are the components of \mathbf{Y} , and the vector index $\mathbf{k} = (k_1, k_2, k_3, \dots)$ runs over all powers with the constraint $\sum_j k_j \leq P$, the maximum order of the expansion.

The LS minimization¹⁷ is most straightforwardly carried out with a QR algorithm. It can be shown that the expansion of \mathbf{F} in monomials (14) can be regrouped into the form obtained with the GS procedure (11). The two approaches, mathematically, are therefore strictly equivalent, although their numerical stability properties are different. In fact, both the GS and the QR methods run into numerical difficulties when either the embedding dimension or the order of the expansion P gets large, because the number of coefficients gets large and the fit becomes ill-conditioned when used on short data sets.

For that reason Serre *et al.*⁹ have found it useful to use another standard method for solving the LS problem, namely, a singular value decomposition (SVD).¹⁷ In the LS problem, which determines a vector of parameters \mathbf{a} by minimization of (12),

$$S = \|\mathbf{A} \cdot \mathbf{a} - \mathbf{b}\|^2, \quad (14)$$

the matrix \mathbf{A} is decomposed into a product of orthogonal and diagonal matrices as follows:

$$\mathbf{A} = \mathbf{U} \cdot \text{diag}(w_i) \cdot \mathbf{V}^T \quad (15)$$

and the solution appears in the form

$$a_i = \sum_{j,k} V_{ij} \frac{1}{w_j} U_{kj} b_k. \quad (16)$$

When the problem is well conditioned, all eigenvalues w_i are above machine precision and the SVD gives an answer equivalent to the other solutions of the LS problem. The power of the method resides in the fact that it possesses a good behavior even when the LS design matrix \mathbf{A} is poorly conditioned, i.e., when the parameters are close to being linearly dependent, whether because of a poor choice of parameters or because their number exceeds the number of data points. It suffices then to restrict the sum i over the eigenvalues w_j to the ones above some (numerical) noise threshold. With SVD the number of coefficients can be taken arbitrarily large,

in fact even larger than the number of available data. The SVD method thus automatically selects the most important linear combinations of the coefficients and discards the rest.

Flows

Standard Formulation

The preceding methods are not only useful for constructing global polynomial maps, but they can also be extended to reconstruct flows globally. In fact since the data are given at discrete times one needs to replace the standard flow (4) by a discrete approximation anyway and the difference between a map and a flow is not very big. As previously, one uses delay coordinates generating a vector \mathbf{y}^n in the phase space, namely,

$$\mathbf{y}^n = (s_n, s_{n-\tau}, \dots, s_{n-(d_E-1)\tau}). \quad (17)$$

The practical implementation of the vector field obviously also depends on the integration scheme. For instance, an explicit Euler integration may be used:

$$\mathbf{y}^{n+1} = \mathbf{y}^n + \mathbf{f}(\mathbf{y}^n)\delta t, \quad (18)$$

where now $\mathbf{y}_n = \mathbf{y}(n\delta t)$ with the time step δt . Given this equation, the model works well when the time interval δt between measurements is small. In many cases, δt must be a few hundred times smaller than the characteristic time of the oscillations that occur in \mathbf{y} .

In many experimental situations, it is not feasible to use a small enough sampling interval. It is therefore worthwhile to seek a method that is capable of accurately modeling \mathbf{f} when δt , the time step between two successive measurements, is much larger than the Euler method would allow. This issue has been addressed by Brown *et al.*⁸ who advocate the use of an Adams predictor–corrector method. Under this method, one formally integrates an initial condition \mathbf{y}^n to the next point \mathbf{y}^{n+1} in one step of size δt via

$$\mathbf{y}^{n+1} = \mathbf{y}^n + \delta t \sum_{j=0}^M a_j^{(M)} \mathbf{f}(\mathbf{y}^{n+1-j}), \quad (19)$$

where δt is the time step of the integration and $a_j^{(M)}$ are the implicit Adams predictor–corrector coefficients.⁸ M indicates the order of the corrector portion of the integration. The numerical values of the $a_j^{(M)}$ are given by the method described in reference 17. In their study, Brown *et al.*⁸ use a time delay Δ in terms of δt that is estimated by using the mutual information method.¹

The model flow \mathbf{f} is optimized with the help of a minimum description length technique in which an error function

$$\chi^2 = \frac{1}{2N\sigma^2} \sum_{n=1}^N \left| \mathbf{y}^{n+1} - \mathbf{y}^n - \delta t \sum_{j=0}^M a_j^{(M)} \sum_{i=0}^{N_p} \mathbf{K}_i \phi_i(\mathbf{y}^{n+1-j}) \right|^2 \quad (20)$$

is introduced. This error function is quadratic with respect to the fitting coefficients \mathbf{K}_i giving a LS problem. The full error function involves a length criterion based on the minimum description length principle for truncating a model (references 8 and,

more recently, 18 and references therein). It is indeed very elaborate and provides an efficient automatic criterion to determine the model coefficients. The quality of the model depends on the time delay Δ , on the order M of the correction in the integration and on the number N_p of retained polynomials. Unfortunately, the interest and efficiency of the error function are reduced by the need for a serious coding effort and large computing time requirements. However, the use of an SVD approach which would alleviate these problems has not yet been tried.

The Derivative Approach

In many experimental situations, a dynamical system may be described by a variable, say s , and its successive derivatives \dot{s} , \ddot{s} , The first derivative can be called a speed and the second derivative an acceleration, as in the case of mechanical oscillators, i.e., derivatives may receive a direct physical interpretation. For this reason (and others), it is of interest to develop models which use an embedding based on derivative coordinates in which the phase space vector is:

$$\mathbf{z}_n = (s_n, s_n^{(1)}, s_n^{(2)}, \dots, s_n^{(d_E-1)}), \quad (21)$$

where $s_n^{(i)}$ designates the i th derivative of the time series variable s_n at time $n\delta t$.¹² The successive derivatives may be estimated either by using a linear discrete filter built on a basis of discrete Legendre polynomials⁶ or by using a local fit with an m th-order polynomial (where m is greater than the largest order of derivative required in the model). In both cases, a window size Δ has to be determined and plays a role rather similar to the time delay Δ used in the time delay method. Nevertheless, from our experience, it appears that the quality of the derivative embedding is less sensitive to the value of Δ than the time delay Δ in the case of a time delay embedding. For a good evaluation of the derivatives, the time interval δt must be a few hundred times smaller than the typical characteristic time of oscillations that occur in the recorded time series $\{s_n\}$.

We thus search for a model in the canonical form

$$\begin{cases} \dot{z}_1 = z_2 \\ \dot{z}_2 = z_3 \\ \vdots \\ \dot{z}_{d_E} = h_{d_E}(\mathbf{z}). \end{cases} \quad (22)$$

One may then remark that the reconstructed vector field $\mathbf{f}(\mathbf{z})$ is composed of $(d_E - 1)$ explicitly known components and by a single unknown component h_{d_E} which may be given a polynomial form. Consequently, this method only requires the determination of a single function h_{d_E} in contrast with the previous method which requires the determination of the d_E functions. This advantage is however balanced by the need of a higher-order polynomial which is required to express the scalar h_{d_E} .

The function h_{d_E} is taken to have the form

$$h_{d_E}(\mathbf{z}) = \sum_{i=1}^{N_p} c_i \phi_i(\mathbf{z}) = \sum_{i=1}^{N_p} K_i P_i(\mathbf{z}), \quad (23)$$

where $P_i(\mathbf{z})$ are the monomials up to order N_p that can be formed from the d_E components of the vector \mathbf{z} according to a relationship defined in reference 12, namely, $P_i(\mathbf{z}) = (z_1^{i_1} z_2^{i_2} \cdots z_{d_E}^{i_{d_E}})$. The biunivocal relationship used between the subscript i and superscript set $\{i_1, \dots, i_{d_E}\}$ is extensively described in reference 12 when $d_E = 3$. In their applications¹² $\{\phi_i(\mathbf{z})\}$ is a basis of orthonormal polynomials constructed via a GS procedure starting from $\phi_1(\mathbf{z}) = 1$ and using a scalar product defined on a data set, and the vector field (22) \mathbf{F} is integrated by using a fourth-order Runge-Kutta scheme with an adaptive stepsize control.¹⁷

The model is optimized by using an error function

$$E_r = \sum_{i=1}^{N_q} \left| y_i^{(d_E)} - \sum_{p=1}^{N_p} K_p P_p \right| \left/ \sum_{i=1}^{N_q} |y_i^{(d_E)}| \right. \quad (24)$$

This error function is calculated by using absolute values for computational efficiency. It is found that the quality of the model depends on reconstruction parameters, namely, the number N_q of vectors \mathbf{y}_n , the number N_s of \mathbf{y}_n 's sampled per pseudo-period and the number N_p of retained multivariate polynomials. The error function E_r is used as a guideline to determine the values of the reconstruction parameters as follows. First, optimal values of N_s and N_p are found by minimizing the error function. However, for a given N_q , the obtained approximated system is not necessarily successful, in particular it cannot necessarily be successfully integrated, although the error function passes through a (possibly local) minimum. Consequently, the search of a successful global modeling needs systematical trials on N_q 's which can nevertheless be automatically done with computational help. More generally, finding an optimal modeling of nonlinear data series remains a tricky problem for which Brown *et al.*⁸ proposed a more elaborated error function than the one of Gouesbet and Letellier, which however is computationally faster.

It is the ultimate goal of the analysis and the nature of the available data set which determine whether one prefers to use a map or a flow, and in the second alternative which of the alternate forms (5) or (23). Generally it seems that maps give more stable reconstructions than flows (e.g., references 10 and 11).

As far as the choice of GS approach versus the LS approach with SVD is concerned, both approaches work for large data sets, but the first method can be ill-conditioned for high embedding dimensions and high-order polynomials. Furthermore the SVD LS approach is faster.

Noise

Measured time series are of course always contaminated with noise, and the question arises how these methods hold up under realistic situations. We note that the noise problem can be particularly bad in astronomy. Not only is the signal to noise ratio often small, the data generally are not or cannot be obtained at the equal space time intervals that the methods require. The necessary interpolations introduce additional noise.

References 1, 9–11 and 21 address the problems introduced by noise. It should be noted though that noise can also have a beneficial effect, namely, in moderate

amounts it stabilizes the maps without affecting their properties, and in fact it has been found useful to add low-level noise to the data for this reason.^{9,11,21}

APPLICATIONS TO PHYSICAL SYSTEMS

The global reconstruction techniques discussed above have been successfully applied to experimental and observational data, as well as to data generated by realistic numerical models of complex physical systems as reviewed below. In each case, a brief discussion concerning the validation of the models will be given. We note that not all variants of the reconstruction method have been applied to the same data. It is therefore not possible to make direct comparisons between them.

Nonlinear Vibrations in a Thin Wire

Brown *et al.*²¹ have applied their techniques to experimental data obtained from an apparatus used to study nonlinear vibrations in a thin wire.¹⁹ The apparatus basically consists of a mount holding a tensioned wire fed by an alternating current. The frequency of the current is close to the fundamental frequency of the free oscillations of the wire. This current excites forced vibrations when the wire is placed in a permanent magnetic field. As the amplitude and frequency of the current are varied, the system can undergo a torus-doubling route to chaos.²⁰ Optical detectors are used to measure the transverse amplitude of the wire. The transverse displacements of the wire generate the time series to be analyzed. Control parameters of the experimental set-up may be found in reference 19.

The amplitude oscillations have a pseudo-period equal to about 0.1 s. The time step δt between two successive measurements is equal to 7.7×10^{-4} s. Thus, 130 points per cycle are available. A three-dimensional embedding (FIG. 1) of a chaotic trajectory is created out of the scalar amplitude measurements made by a single detector by using delay coordinates. The time delay is fixed to $\Delta = 39\delta t$, based on the mutual information criterion. The embedding dimension d_E is found to be equal to 3.

The model (FIG. 2) is validated by comparing Lyapunov exponents. From the experimental data, the positive Lyapunov exponent is $\lambda_1 = 4.97 \times 10^{-2}$ and the negative one is $\lambda_3 = -0.702$, while, from the synthetic data, they are $\lambda_1 = 4.31 \times 10^{-2}$ and $\lambda_3 = -0.576$. These values differ by approximately 15% for the positive exponent and 20% for the negative exponent.¹⁹

Another method for checking the model is synchronization.⁸ The modified Fuyisaka and Yamasada method of synchronization is given by

$$\begin{cases} \dot{\mathbf{x}} = \mathbf{g}(\mathbf{x}) \\ \dot{\mathbf{y}} = \mathbf{f}(\mathbf{y}) - \mathbf{e}(\mathbf{y} - \mathbf{x}) \end{cases}, \quad (25)$$

where \mathbf{e} is the coupling matrix between the two dynamical systems involved in (26). The $\mathbf{x}(t) \in \mathbb{R}^{d_x}$ denotes the trajectory of the experimental system in its phase space, and $\mathbf{y}(t) \in \mathbb{R}^{d_y}$ the trajectory generated by the model in the embedding phase space, namely the response system in the jargon of synchronization.²¹ The driving

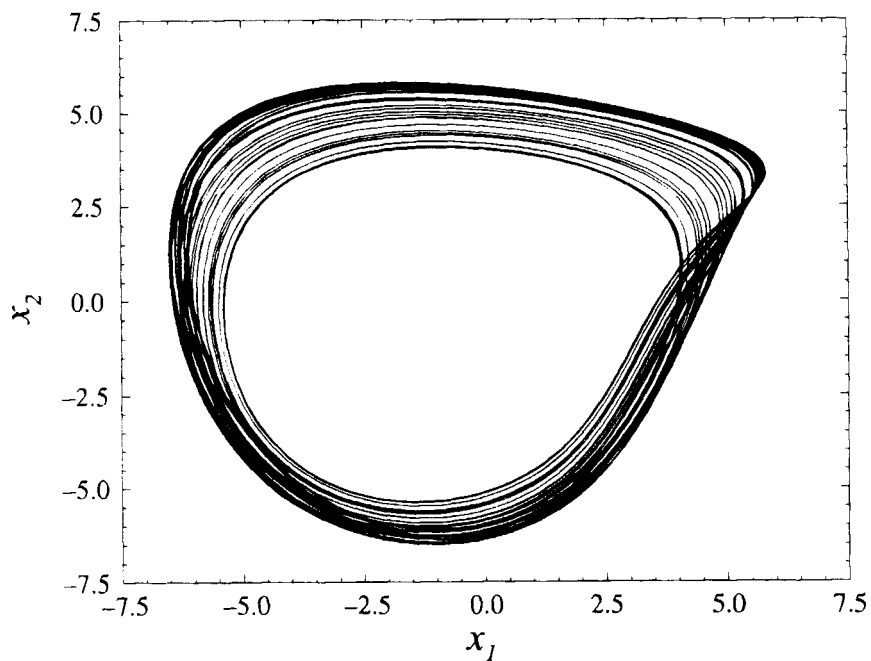


FIGURE 1. Plane projection of the three embedded time series from the string experiment.

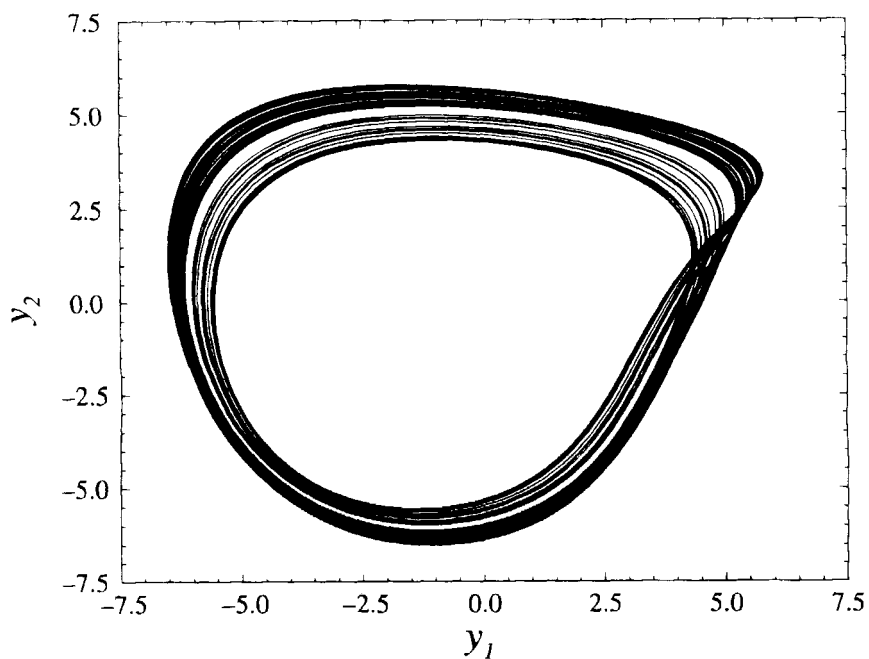


FIGURE 2. Projection of the reconstructed attractor generated by integrating the model.

and the response systems are said to be synchronized if there exists a coupling matrix with a few nonzero components ensuring that the response trajectory is slaved to the driving trajectory. Brown *et al.*²¹ used only one nonzero component lying on the leading diagonal which may be written as $e_{\beta\beta} = \varepsilon$. They showed that, depending on the value of β , setting ε to a too small or to a too large value does not produce any synchronous motion. More important, synchronization could only be obtained if the reconstructed vector field \mathbf{f} was a good enough approximation to the true dynamics \mathbf{f} .⁸ By using the third component of \mathbf{x} for coupling, they found that the model is synchronized for $\varepsilon \geq \varepsilon_c \approx 0.5$. Tests were given for $\varepsilon = 0.75$ and 3.0 .

The topological analysis has been also used as a third test to check the model. A review concerning topological characterization may be found in reference 24. Tufillaro *et al.*¹⁹ showed that the model induces the same template as the experimental data, i.e., the relative organization of the periodic orbits is the same for both the experimental data and the model data. Such a topological analysis guarantees that the original attractor and the reconstructed attractor are topologically equivalent. We think that topological analyses may provide the best criterion for checking a model, particularly if the populations of periodic orbits are also compared.

Brown *et al.*⁸ have also obtained models for experimental data arising from an electronic circuit for two values of a control parameter α . The corresponding attractors are displayed in FIGURES 3 and 5. They form a so-called single attractor and a symmetric attractor with two wings, respectively. The models contain $N_p = 4$ and

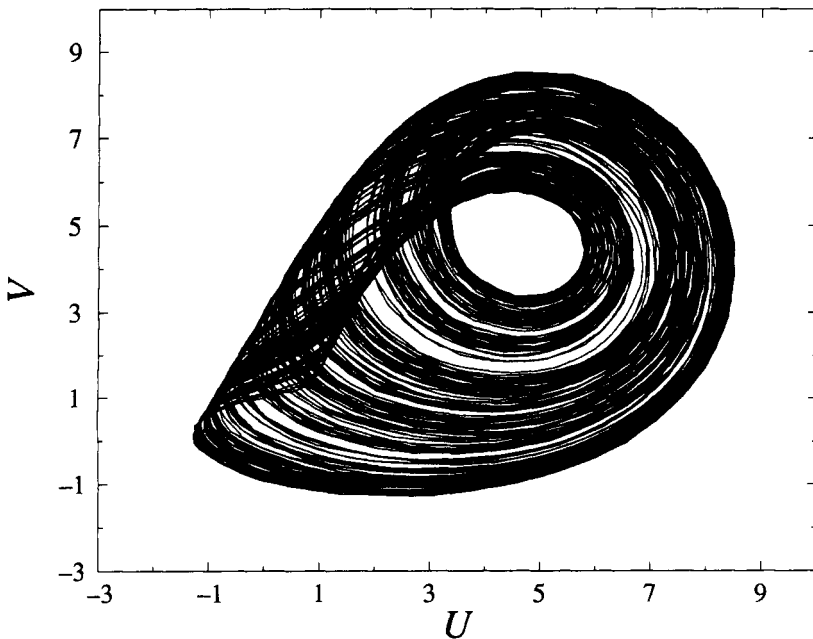


FIGURE 3. Plane projection of the single attractor.

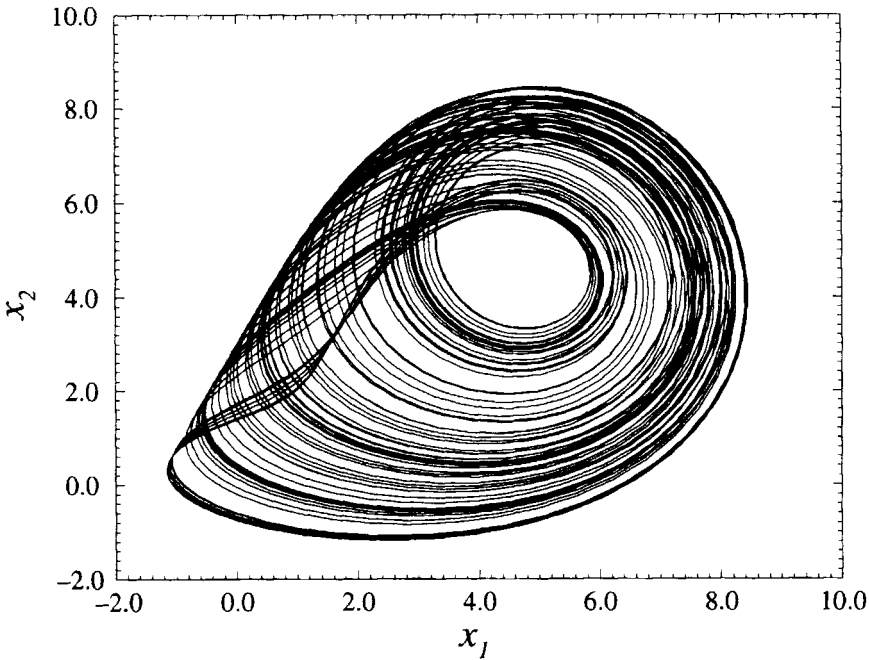


FIGURE 4. Model of the single attractor.

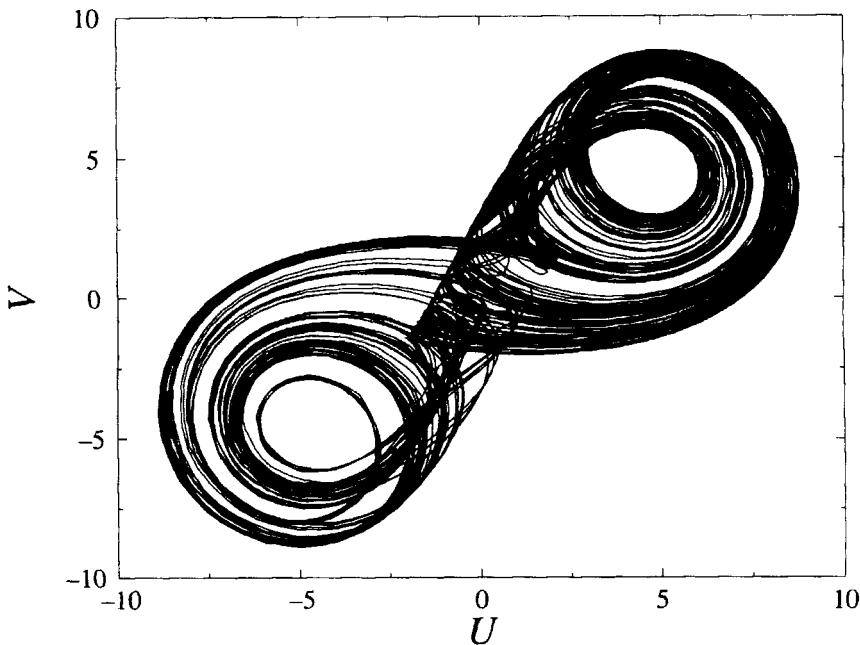


FIGURE 5. Plane projection of the double scroll attractor.

$N_p = 5$ polynomials for $\alpha = 17.4$ and $\alpha = 18.9$, respectively, i.e., chaotic trajectories have been successfully reconstructed with rather simple mathematical expressions (FIGS. 4 & 6).

Finally, we mention the reconstruction of a model of the Belousov–Zhabotinskii reaction⁸ whose topological analysis has been given by Mindlin *et al.*²³ A projection of the corresponding attractor (cf. reference 23 for a specific embedding procedure) is displayed in FIGURE 7. Here again, the model is found to be rather simple, containing $N_p = 7$ polynomials with a 3D-embedding (FIG. 8).

Electrodissolution of a Rotating Cu Electrode

We start with a time series which has been obtained from current measurements during the potentiostatic electrodisolution of a rotating Cu electrode in phosphoric acid. The experiments have been described in detail in references 25 and 26. Briefly, the experimental set-up consisted of a rotating disc electrode which had a copper rod, 8.26 mm in diameter, embedded in a 2 cm diameter Teflon cylinder. The rotating speed was maintained at 4400 rpm. The cell was a 500 ml flask with a side neck in which the capillary probe was fixed. The cell contained 250 ml of 85% phosphoric acid and a water bath was used to maintain its temperature at 20°C. A Potentiostat (Princeton Applied Research model 273) was used to regulate the potential of the

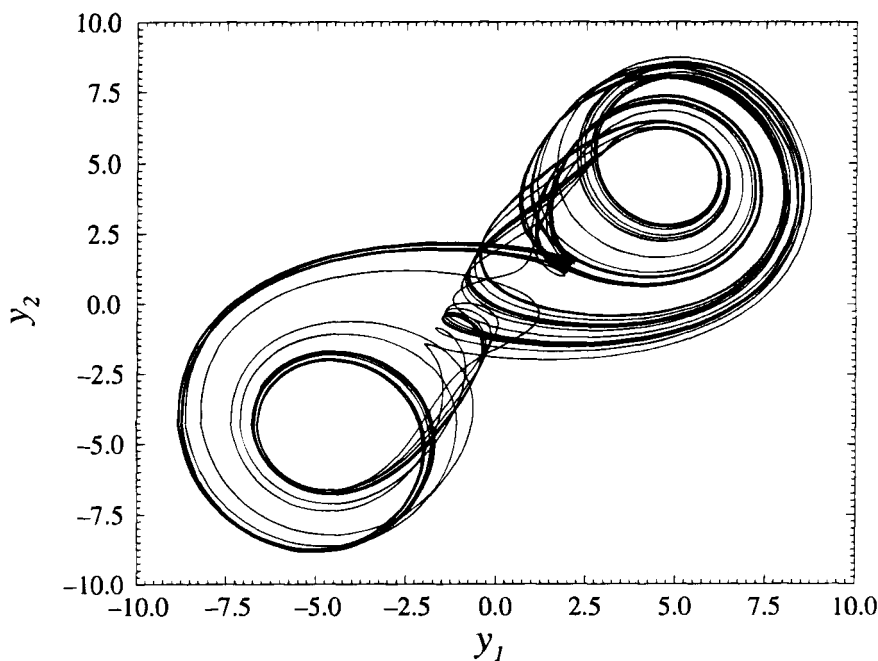


FIGURE 6. Model of the double scroll attractor.

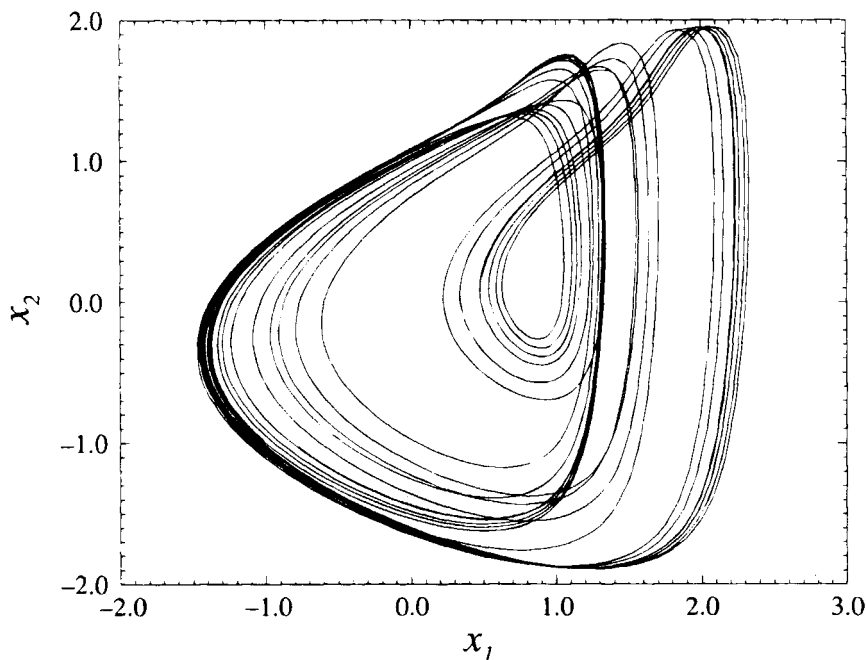


FIGURE 7. A differential-integral phase space embedding of the Belousov-Zhabotinskii data.

working disc electrode with respect to the reference electrode and to monitor the current. The data were recorded at a frequency f_e of 1500 Hz.

The state space of the electrodisolution has been reconstructed from the current time series $I(t)$ by using successive time derivatives. The embedding dimension d_E is found²⁵ to be equal to 3 from the estimation of the correlation dimension by using the Grassberger-Proccacia algorithm.³² In the reconstructed phase space, the asymptotic motion settles down onto a chaotic attractor displayed in FIGURE 9. The so-called *copper* attractor is found to be characterized by a horseshoe template (cf. references 25, 26 or 24). The window size Δ used to evaluate the derivatives is taken to be equal to $21\delta t$, where $\delta t = 6.67 \times 10^{-4}$ s. The pseudo-period of the oscillations is about 0.102 s. Thus, we have approximately 150 points per pseudo-period. The number of retained polynomials in the model is found to be equal to 26 which is rather large in comparison with N_p 's obtained in the cases discussed in the previous section. As explained earlier, it is inherent to the form of the model, all the nonlinear dynamics being stored in a single standard function.

By numerically integrating the reconstructed model, a reconstructed attractor (FIGURE 10) is obtained which looks rather similar to the original copper attractor. The model is checked by using topological analysis. Both attractors indeed induce the same template. Moreover, a slight difference which has been found between the corresponding two orbit spectra has been explained in terms of contamination of the experimental data by noise.²⁶

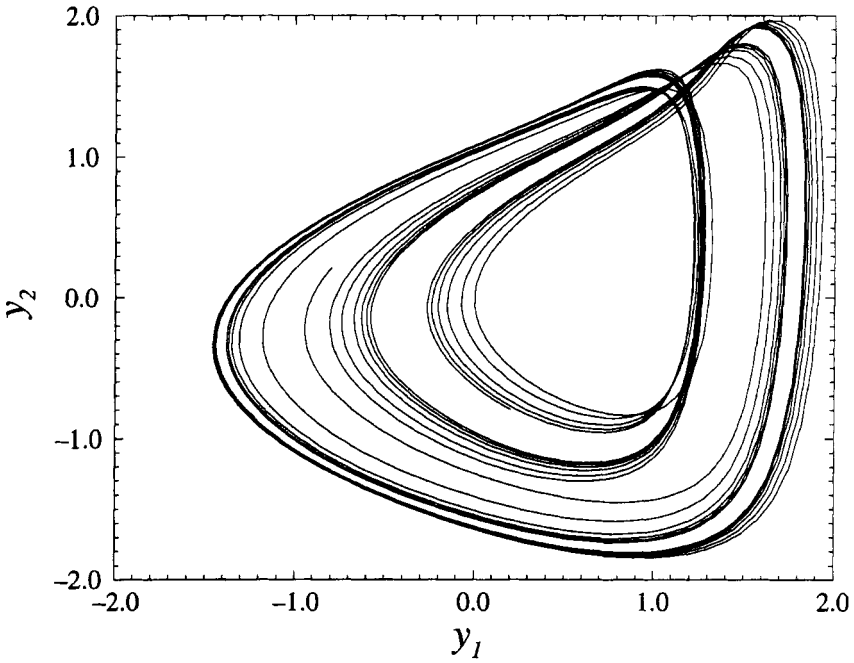


FIGURE 8. Reconstructed model of the Belousov-Zhabotinskii reaction.

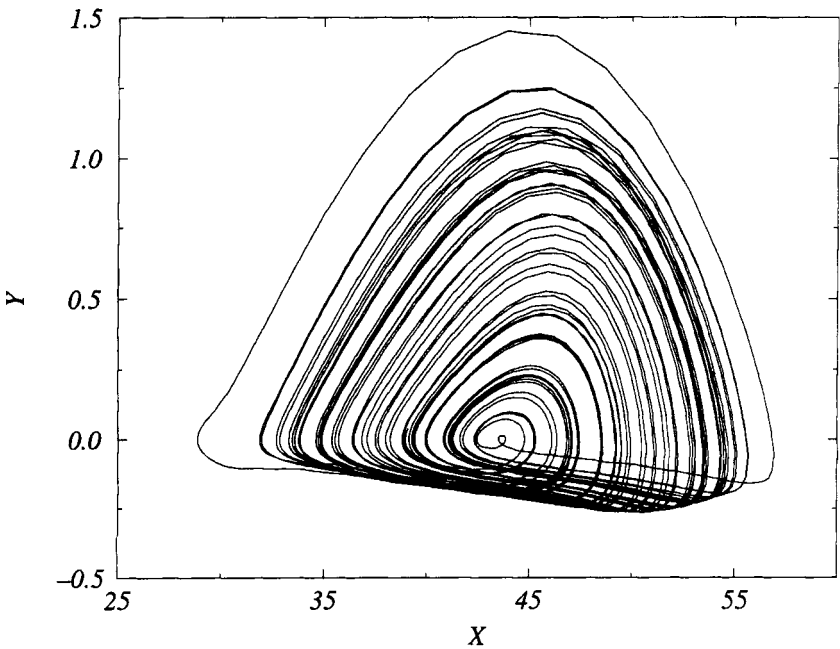


FIGURE 9. Attractor generated by the copper electrodissoolution (copper attractor).

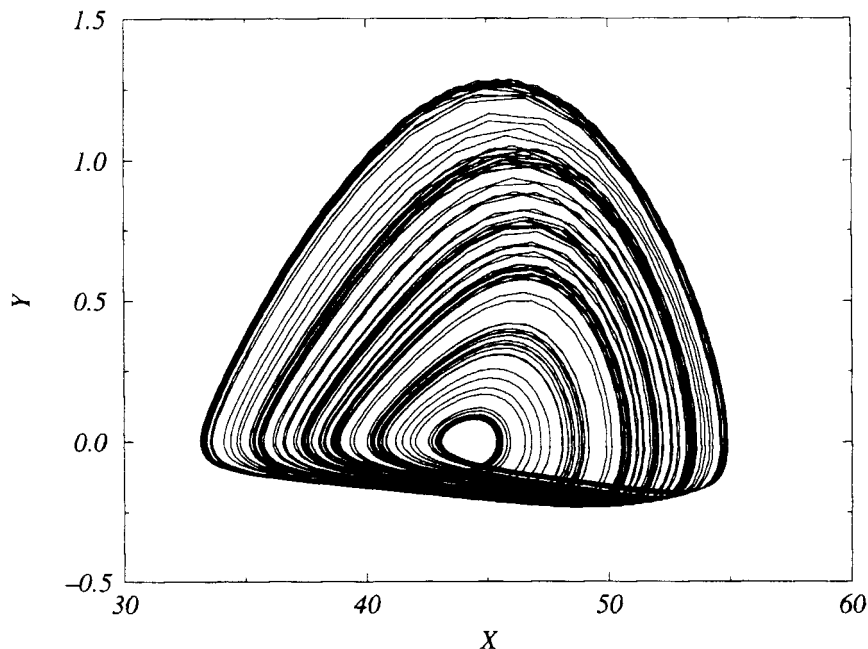


FIGURE 10. Reconstructed attractor A_R of the copper electrodisso- lution by integrating the reconstructed vector field.

A second electrodisso- lution has also been studied subsequently. The experimen- tal set-up is very similar to the previously described one except that it concerns an iron electrodisso- lution in sulfuric acid. The embedding dimension d_E has been found to be equal to 3 by using the mutual information criterion analysis (kindly per- formed by R. Brown). The window size Δ is equal to $35\delta t$. The attractor generated by the experimental data is displayed in FIGURE 11.

In a model for this dynamics, the number N_p of polynomials has been found to be equal to 78. The attractor obtained by integrating the model (FIGURE 12) looks very similar to, but also slightly different from the original attractor. The model has then been validated, but only partially²⁷ insofar as the topological analysis could not be safely performed. In particular, the absence of a hole in the middle of the attractor prevents a safe enough definition of a Poincaré section inducing many difficulties in performing the topological characterization. For these reasons, syn- chronization has also been used. The model has been successfully synchronized with the experimental time series by using $E_{11} = 0.05$ and $E_{33} = 0.05$ as nonzero ele- ments in the coupling matrix of (25). We may then state that the model dynamics is close to the experimental dynamics.

Astrophysical Applications

Two applications of the global flow reconstruction method have been made to problems of astrophysical interest; the first to the irregular pulsations of a model of

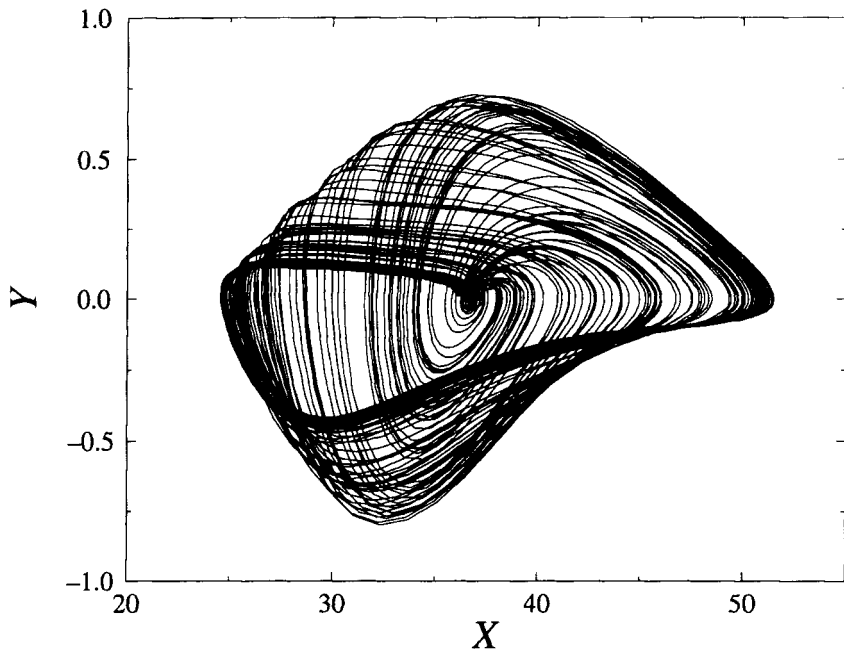


FIGURE 11. Projection in the XY -plane of the attractor generated by the iron electro-dissolution.

a star of the W Vir type that were obtained with a state-of-the-art numerical hydrodynamical code, and the second to the observed irregular light-curve data of two stars of the RV Tau type.

W Vir Model Pulsations

The reconstruction of Serre *et al.*¹¹ confirmed that the irregular pulsations of W Vir models²⁸ are indeed chaotic, and they furthermore could show that the dynamics could be embedded in a three-dimensional space ($d_E \geq 3$). Lyapunov exponents were also computed and led to a fractal Lyapunov dimension $d_L \approx 2.02$ for the reconstructed attractor, with d_L independent of the embedding dimension. With the aid of the embedding theorem they thus concluded that the physical attractor also has a fractal dimension of ≈ 2.02 . From these two bounds $d_L < d \leq d_E$ it follows that the *physical dimension* d is also 3.

These results were robust with respect to changes in the delay and the maximum order P of the polynomials of the map. They also found that as d_E is increased, the highest polynomial degree P can be decreased. A convenient and in some sense optimal representation of the attractor can be made in Broomhead–King coordinates ξ_k (projections onto the eigenvectors of the correlation matrix³⁴). The projections onto the planes spanned by the first three principal components are shown in

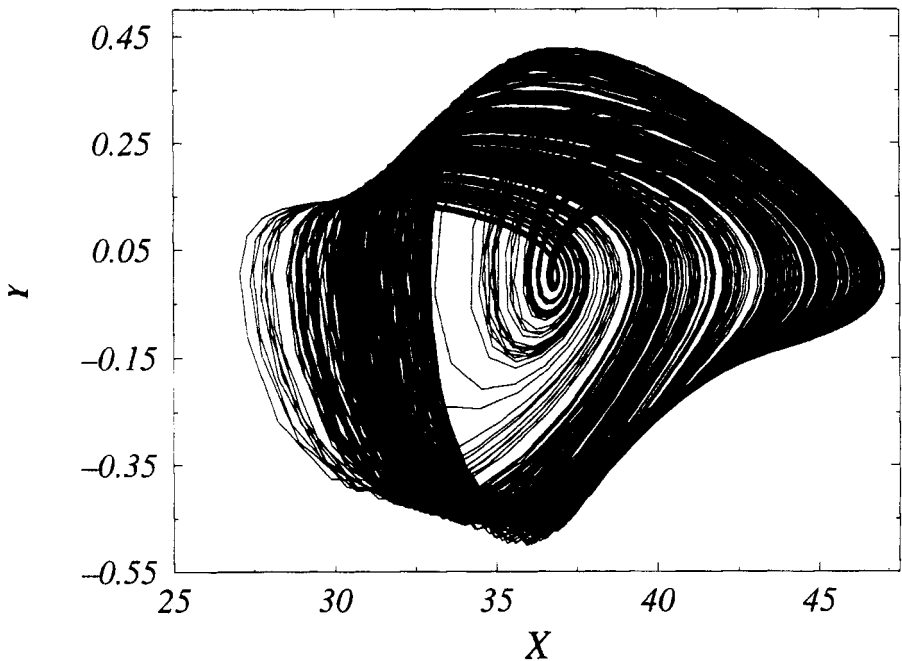


FIGURE 12. Reconstructed attractor of the iron electrodissoolution obtained by integrating the reconstructed vector field.

FIGURE 13. Once the map has been determined it can be iterated to produce synthetic signals. FIGURE 13 shows that the synthetic signals are very similar to the original data set, independently of the embedding dimension. For further details we refer to reference 11.

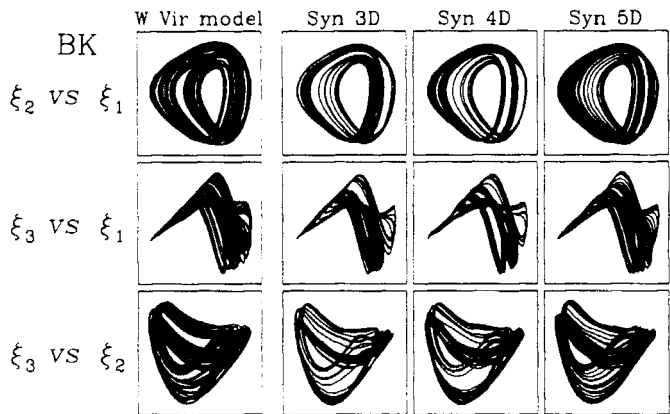


FIGURE 13. W Vir model, lowest three principal component projections. Column 1: data set; 2-4: synthetic signals from noisy data set in 3D, 4D and 5D.

One notes that the numerical hydrodynamical code replaces the PDEs of fluid dynamics and heat flow by a discrete approximation consisting of N mass shells. In the above-mentioned work a set of $3N = 180$ coupled nonlinear ODEs were thus solved. (These ODEs are very stiff in the mathematical sense because of the sharp temperature sensitivity of the equation of state and opacity law.) It is therefore nontrivial that the resultant pulsation can be fully described by a mere three ODEs. This dimensional shrinkage from 180D (or any other number of zones N that sufficiently resolves the spatial structure of the star) to 3D shows that there is an *inertial manifold* in which the dynamics takes place. This implies that the time-dependence of all the physical variables throughout the star (e.g., velocity, density and temperature fields) can be expressed in terms of only three basic variables.

While the *physical* picture is not entirely clear yet, these results taken together with a Floquet analysis of the hydrodynamical models³⁵ suggest that the linearly unstable fundamental mode of oscillation of the star interacts nonlinearly with a single overtone through a parametric resonance to produce the complex behavior. How the additional dimensional shrinkage from 4D (two complex modes) to 3D comes about is still under investigation.

R Scuti Observational Data

The recent availability of a suitably large and densely sampled observational data set has made it possible to apply the global flow reconstruction method to astronomical data^{10,15} and to probe the properties of the irregular pulsation cycles of an actual star, namely star R Scuti. The light-curve is shown in FIGURE 14 on top. The authors demonstrated first that the observational light-curve data of R Scuti, a star of the RV Tau type, is not multi-periodic, and that it cannot have been generated by a linear stochastic (AR) process. The nonlinear reconstruction analysis shows that this star's complicated light-curve is captured by a simple 4D polynomial map or flow (four first-order ODEs) and that the bulk of the signal consists of low dimensional chaos. FIGURE 14, at the bottom, shows a typical synthetic light-curve that has been generated from the reconstructed 4D map.

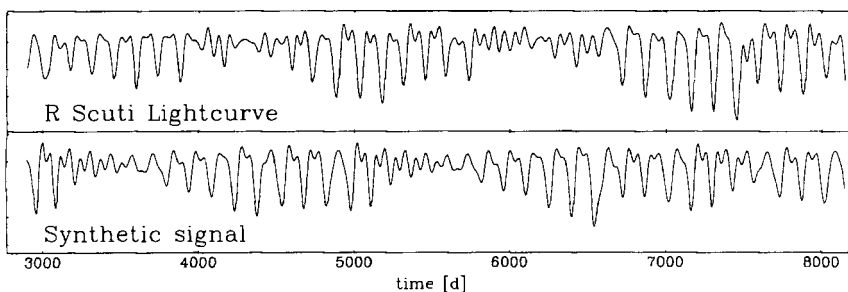


FIGURE 14. Top: R Scuti light-curve; bottom: synthetic light-curve.

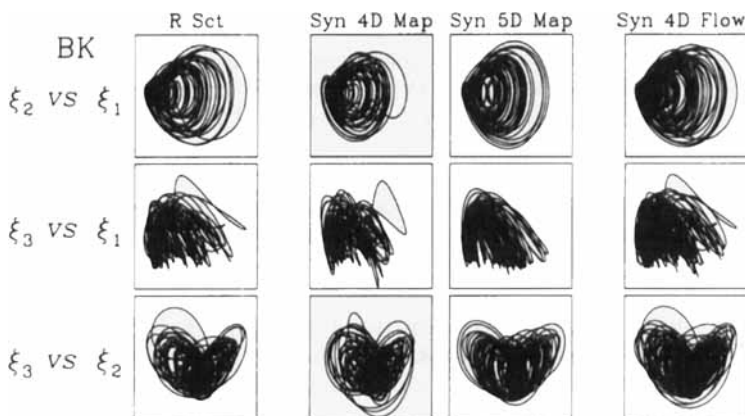


FIGURE 15. Lowest three Broomhead–King coordinate projections. *Column 1:* data set; *2 & 3:* synthetic signals from 4D and 5D maps; *4:* synthetic signal from 4D flow.

FIGURE 15 shows the lowest Broomhead–King projections of the R Scuti light-curve together with synthetic light-curves in 4D and 5D. The rightmost figures show a synthetic signal from a reconstructed flow (using the Adams integration scheme). The synthetic signals are not only very similar in appearance to the R Scuti light-curve as FIGURE 14 shows, but the more discriminating BK projections in FIGURE 15 show that they are really very similar, and robust with respect to dimension. They also show that a flow really underlies the dynamics.

For astrophysical purposes it is important that the method allows us to *quantify an irregular signal*, in terms of Lyapunov exponents and dimension, which has the potential novel benefit for extracting novel stellar constraints from irregular light-curves.

The low dimensionality 4 of the flow again suggests a simple physical picture of the pulsations, namely that the pulsations of R Scuti are the result of the nonlinear interaction of two vibrational “normal” modes of the star¹⁰ (see also Kolláth & Buchler in this volume for additional evidence).

A preliminary study of another star of the same type, but not quite as irregular, AC Her, again indicates a low dimensional dynamics of 4.³³

Following the earlier discussion, we also note that in a first study¹⁵ the orthonormal polynomials were constructed with the GS procedure. This led to difficulties, in particular concerning the highest polynomial degree as their figure 3 and associated comments indicate. (Numerical cancellations cause negative norms.) A subsequent more thorough study resorted to the LS procedure with the help of an SVD algorithm. This procedure is not only a factor of 10 faster, but it remains numerically stable for all values of d_E and P (cf. reference 10).

RECONSTRUCTION WITH A CONTROL PARAMETER DEPENDENCE

We have discussed some applications of global vector field reconstruction with *given* control parameters to experimental systems. An additional step can be taken

in the direction of reconstructing models as a function of a control parameter dependence. Actually such a goal has already been reached with neural networks,²⁹ but a study of control parameter dependence with a global vector field reconstruction technique such as those discussed in the present paper is still very new.

Principle

We assume that the dynamical system of (2) is studied under the variation of a control parameter α taken from the control parameter vector μ . With the method described earlier, the standard system reconstructed from the recorded variable x with α -dependence reads as

$$\begin{cases} \dot{z}_1 = z_2 \\ \dot{z}_2 = z_3 \\ \dot{z}_3 = f_s(z_1, z_2, z_3, \alpha), \end{cases} \quad (26)$$

where the standard function f_s now depends on the control parameter α . The basis polynomials therefore involve monomials of the form $(z_1^i z_2^j z_3^k \alpha^l)$, generalizing the previous discussion. The error function is furthermore modified to account for the control parameter α (for more details, see reference 30).

Numerical Check

A numerical check of the validity of this approach is now given in the case of the Rössler system

$$\begin{cases} \dot{x} = -y - z \\ \dot{y} = x + ay \\ \dot{z} = b + z(x - c). \end{cases} \quad (27)$$

This system has been extensively studied along a line in the control parameter space defined by $(a \in [0.33, 0.557], b = 2, c = 4)$.²² Moreover, we have shown that there exists a diffeomorphism between the y -induced attractor and the original Rössler attractor.³¹ Consequently, the y -variable for the Rössler system provides a good checkpoint for the present extension. The exact standard function, which may be analytically derived, then has a polynomial form involving 51 monomials.³⁰

The learning set of data is composed of four time series $\{y_i, i \in [1, N_c = 1000]\}$, generated for four different values of the control parameter a taken to be equal to 0.2, 0.2625, 0.325 and 0.3875. Let us remark that these a -values correspond to three different limit cycles of the period-doubling cascade and to a chaotic behavior just beyond the accumulation point at $a_\infty = 0.386$. The sampling rate of the time series is taken to be equal to $\delta t = 10^{-2}$ s. A good approximation of the standard function is easily found with $N_p = 51$ and $\Delta = 7\delta t$, for which the generalized error function is $E_r = 3.8 \times 10^{-7}$. The estimated coefficients of the reconstructed standard function are reported in Table II of reference 30.

The integration of the reconstructed model generates attractors which may be shown to be topologically equivalent to the original Rössler attractors. In order to check our model with respect to the control parameter dependence, let us now compare the bifurcation diagrams of the reconstructed and of the standard exact systems along the line $a \in [0.33, 0.557]$ (Figs. 16 and 17). The agreement between both diagrams is very satisfactory indeed.

Preliminary Results from Experimental Data

We now turn to a discussion of preliminary results concerning the use of the previously described technique with control parameter dependence to the case of experimental data. The data arise from copper electrodisolution experiments in which the potential between the reference electrode and the copper rod is varied. These data (from J. Hudson and Z. Fei, University of Virginia) are made of 18 time series which have been recorded for different values of the potential, taken as the control parameter. The pseudo-period of the current oscillations is about 0.1 s and the time step δt is equal to 5×10^{-4} s. Thus, 200 points per pseudo-period are available. Each time series contains 5000 points and is smoothed.

The experimental and reconstructed bifurcation diagrams are displayed in FIGURES 18 and 19, respectively. Although the agreement is not perfect, there is a general resemblance which is certainly encouraging. In particular, the model gener-

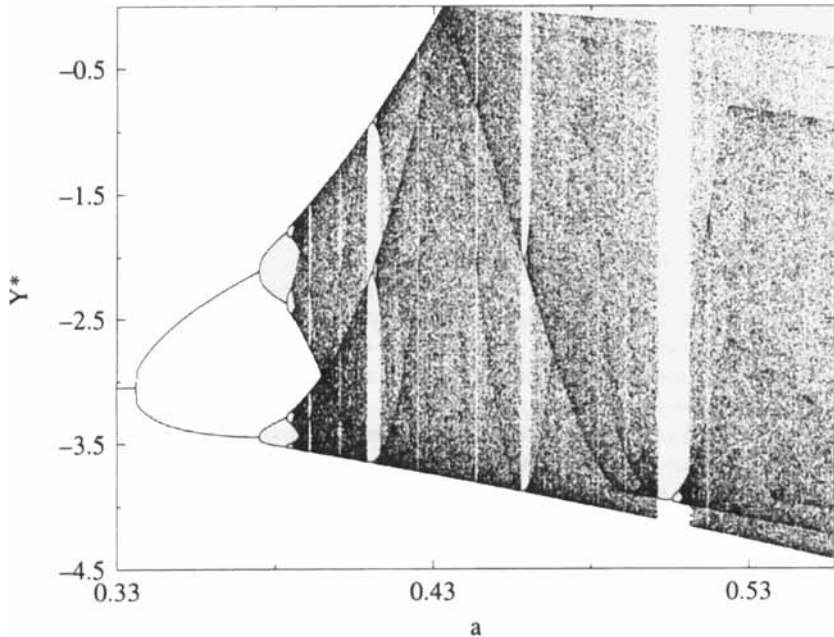


FIGURE 16. Bifurcation diagram for the original Rössler system with y -coordinate.

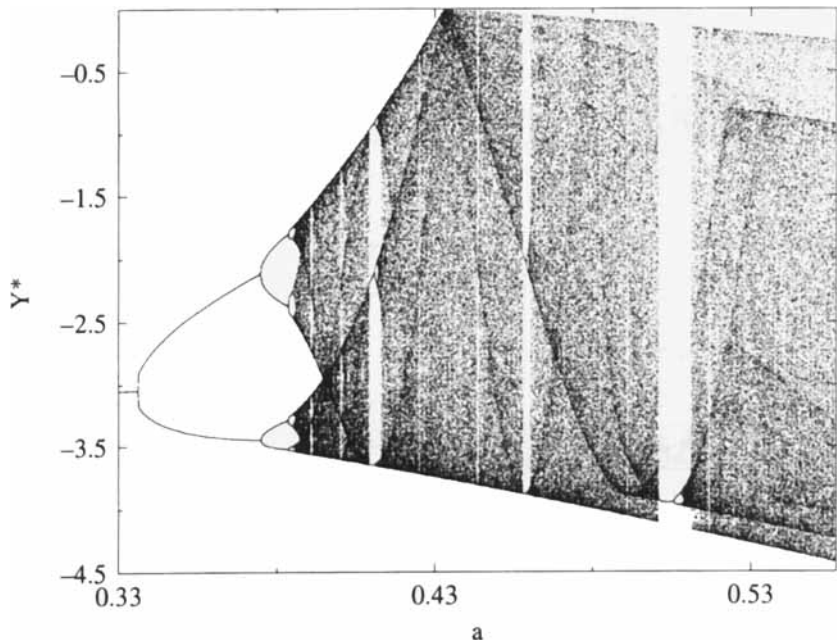


FIGURE 17. Bifurcation diagram from the reconstructed model, with y -coordinate.

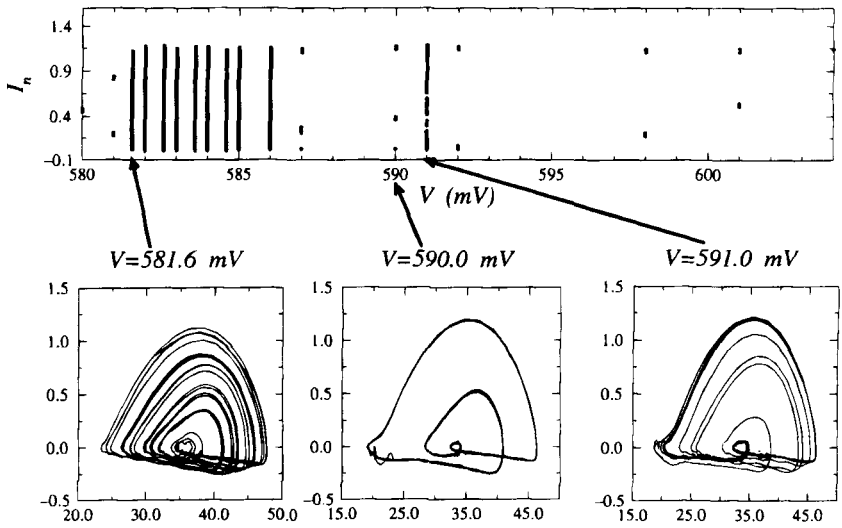


FIGURE 18. Experimental bifurcation diagram.

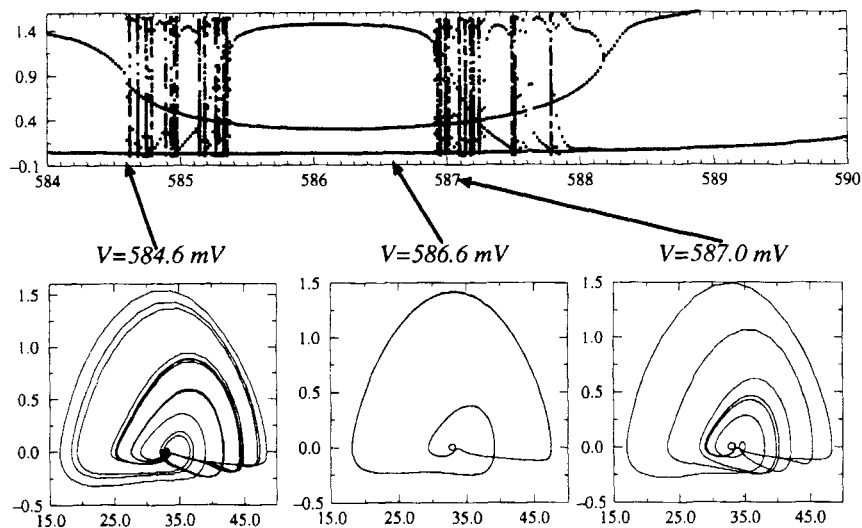


FIGURE 19. Reconstructed bifurcation diagram.

ates a sequence of behavior which is similar to the experimental one. We hope that a further effort will significantly improve these preliminary results.

CONCLUSIONS

Global vector field reconstructions are a topic of growing interest. They allow one to obtain a global model from measurements carried out on a *single* variable scalar time series. In other words, they provide relevant information on variables from which no measurements have been taken. Several variants of this technique have been reviewed and applications have been discussed. It is important to remark that global vector field reconstruction is nowadays developed enough to allow applications to (short, noisy) experimental data. We have also demonstrated that a control parameter dependence can be included in the technique. Then, measurements on one variable for a few values of the control parameter may allow us to get information on variables and for control parameter values which were not experimentally investigated. We may imagine many uses of the phenomenological models so obtained, such as for prediction, encoding of long time series, and control, to only quote a few key words, and it is expected that the interest for this topic will continue to grow.

ACKNOWLEDGMENTS

We wish to thank J. Hudson and Z. Fei for their beautiful data from the copper electrodisolution and N. Rul'kov for the data from the electronic circuit. We would

also like to thank these coworkers for their helpful discussions, and Janet Mattei for the AAVSO observational data.

REFERENCES

1. ABARBANEL, H. D. I., R. BROWN, J. J. SIDOROWICH & L. SH. TSMIRING. 1993. *Rev. Mod. Phys.* **65**: 1331.
2. WEIGEND, A. S. & N. A. GERSHENFELD. 1994. *Time Series Prediction*. Addison-Wesley, Reading, MA.
3. PACKARD, N. H., J. P. CRUTCHFIELD, J. D. FARMER & R. S. SHAW. 1980. *Phys. Rev. Lett.* **45**: 712.
4. TAKENS, F. 1981. *In* *Dynamical Systems and Turbulence*, Warwick 1980, Lecture Notes in Mathematics. D. A. Rand & L. S. Young, Eds. Vol. 898: 366. Springer-Verlag, New York.
5. SAUER, T., J. A. YORKE & M. CASDAGLI. 1991. *J. Stat. Phys.* **65**: 5.
6. GIBSON, J. F., J. D. FARMER, M. CASDAGLI & S. EUBANK. 1992. *Physica D* **57**: 1.
7. FRASER, A. M. 1989. *IEEE Trans. Info. Theory* **35**: 245.
8. BROWN, R., N. F. RUL'KOV & E. R. TRACY. 1994. *Phys. Rev. E* **49**: 3784.
9. SERRE, T., Z. KOLLÁTH & J. R. BUCHLER. 1995. *Astron. Astrophys.* In press.
10. BUCHLER, J. R., Z. KOLLÁTH, T. SERRE & J. MATTEI. 1996. *Astrophys. J.* In press.
11. SERRE, T., Z. KOLLÁTH & J. R. BUCHLER. 1996. *Astron. Astrophys.* In press.
12. GOUESBET, G. & C. LETELLIER. 1994. *Phys. Rev. E* **49**: 4955. See also references therein discussing the use of models based on rational functions.
13. RICE, J. R. 1964. *The Approximation of Functions*, Vol. 1. Addison-Wesley, Reading, MA. Also: 1969. Vol. 2.
14. GIONA, M., F. LENTINI & V. CIMAGALLI. 1991. *Phys. Rev. A* **44**: 3496.
15. BUCHLER, J. R., T. SERRE, Z. KOLLÁTH & J. MATTEI. 1995. *Phys. Rev. Lett.* **74**: 842.
16. LETELLIER, C., E. RINGUET, B. MAHEU, J. MAQUET & G. GOUESBET. 1995. 4^e Journées Européennes de Thermodynamique Contemporaine, Nancy, to be reprinted in *Entropie*.
17. PRESS, W. H., B. P. FLANNERY, S. A. TEUKOLSKY & W. T. VETTERLING. 1988. *Numerical Recipes*. Cambridge University Press, Cambridge, U.K.
18. JUDD, K. & A. MEES. 1995. *Physica D* **82**: 426.
19. TUFILLARO, N. B., P. WYCKOFF, R. BROWN, T. SCHREIBER & T. MOLTEO. 1995. *Phys. Rev. E* **51**: 164.
20. TUFILLARO, N. B., T. ABBOTT & J. REILLY. 1992. *An Experimental Approach to Nonlinear Dynamics and Chaos*. Addison-Wesley, Reading, MA.
21. BROWN, R., N. RUL'KOV & N. B. TUFILLARO. 1994. *Phys. Rev. E* **50**: 4488.
22. LETELLIER, C., P. DUTERTRE & B. MAHEU. 1995. *Chaos* **5**: 271.
23. MINDLIN, G. B., H. G. SOLARI, M. A. NATIELLO, R. GILMORE & X. J. HOU. 1991. *J. Nonlin. Sci.* **1**: 147.
24. LETELLIER, C. & G. GOUESBET. 1997. Topological structure of chaotic systems. This issue.
25. LETELLIER, C., L. LE SCHELLER, E. MARÉCHAL, P. DUTERTRE, B. MAHEU, G. GOUESBET, Z. FEI & J. L. HUDSON. 1995. *Phys. Rev. E* **51**: 4262.
26. LETELLIER, C., L. LE SCHELLER, P. DUTERTRE, G. GOUESBET, Z. FEI & J. L. HUDSON. 1995. *J. Phys. Chem.* **99**: 2016.
27. LETELLIER, C., L. LE SCHELLER, G. GOUESBET, Z. FEI & J. L. HUDSON. 1996. Topological analysis and global vector field reconstruction of multi-branched chaos: the case of iron electrodisolution. In preparation.
28. KOVÁCS, G. & J. R. BUCHLER. 1988. *Ap. J.* **334**: 971.
29. RICO-MARTÍNEZ, R., K. KRISCHER, I. G. KEVREKIDIS, M. C. KUBE & J. L. HUDSON. 1992. *Chem. Eng. Comm.* **118**: 25.
30. LE SCHELLER, L., C. LETELLIER & G. GOUESBET. 1996. *Phys. Lett. A* **211**: 211.

31. LETELLIER, C. & G. GOUESBET. 1996. *J. Phys. II (France)* **6**: 1615.
32. GRASSBERGER, P. & I. PROCCACIA. 1983. *Physica D* **9**: 189.
33. BUCHLER, J. R., Z. KOLLÁTH & T. SERRE. 1995. *Ann. N.Y. Acad. Sci.* **773**: 1.
34. BROOMHEAD, D. S. & G. P. KING. 1986. *Physica D* **20**: 217.
35. MOSKALIK, P. & J. R. BUCHLER. 1990. *Ap. J.* **355**: 590.