# Construction of Differential Equations from Experimental Data

J. Cremers and A. Hübler *

Physik-Department, Technische Universität München, D-8046 Garching

Z. Naturforsch. **42 a**, 797–802 (1987); received September 15, 1986

A new algorithm to determine the number of degrees of freedom of dynamic systems is presented. To obtain a concise description of an observed chaotic time sequence, an approximation of the flow in a state space representation by series is shown to be useful.

## 1. Introduction

Oscillators with marked nonlinearity and chaotic solutions provide good mathematical models in various fields of physics, as in mechanics [1], electrodynamics [2], medical physics [3], chemical thermodynamics [4], geophysics [5], etc. [6]. Haken and others [7–9] have shown that problems of the classical field theory can often be well described by low dimensional systems of ordinary differential equations.

Generalized dimensions [10, 11], Kolmogorow entropy [12], Lyapunov exponents and diffusion coefficients [1] are used to describe state space representations of periodic and chaotic solutions resulting from numerical simulations and measurements [13–15]. These quantities do not give information, though, about the global structure of the flow vector field, which possesses in many cases a high symmetry and can be described by a series expansion with few terms only. Examples for this are Duffing's oscillator [15–18] and the overturning pendulum [2].

The purpose of this paper is to estimate the parameters (for example the coefficient of a power expansion of the flow vector field) of the differential equations leading to an observed chaotic time sequence and, in order to get smooth flow vector fields, to estimate an embedding dimension of strange attractors in continuous time flows. This embedding dimension can be substantially smaller than the embedding dimension estimated from the Hausdorff dimension [19–21].

* Part of Ph.D. thesis.

Reprint requests to Prof. Dr. E. Lüscher, Physik-Department, Technische Universität München, D-8046 Garching.

## 2. Geometric Reconstruction of the Flow Vector Field

If an observed time sequence originates from a system whose dynamics can be described by an $n$-dimensional system of differential equations of first order $\dot{x} := \mathrm{d}/\mathrm{d}t\, \boldsymbol{x} = \boldsymbol{f}(\boldsymbol{x})$ ($\boldsymbol{x} \in S = n$-dimensional state space, $t$ = time) with a continuous, time-independent flow vector field $\boldsymbol{f}$, the difference of the two flow vectors $\boldsymbol{f}(\boldsymbol{r}_1)$, $\boldsymbol{f}(\boldsymbol{r}_2)$ at neighbouring points $\boldsymbol{r}_1, \boldsymbol{r}_2 \in S$ must vanish, if the distance of $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$ vanishes:

$$\lim_{\boldsymbol{r}_1 \to \boldsymbol{r}_2} |\boldsymbol{f}(\boldsymbol{r}_1) - \boldsymbol{f}(\boldsymbol{r}_2)| = 0 . \quad (1)$$

The usage of a time-independent flow vector field is no essential restriction because methods exist to change time-dependent flow vector fields into time-independent ones [22].

We proceed as follows in order to obtain a continuous flow vector field:

The data set is geometrically represented by a trajectory $\boldsymbol{x}_D(t)$ in a state space of dimension $D$ [11]. The flow vectors $\boldsymbol{f}_i$, $i = 1, \ldots, K$ are numerically calculated by differentiation $\boldsymbol{f}_i = \dot{\boldsymbol{x}}_D(t_i)$ at $K$ positions $\boldsymbol{r}_i = \boldsymbol{x}_D(t_i)$ $(t_1 < t_2 < \cdots < t_K)$ at the trajectory. If possible, the positions $\boldsymbol{r}_i$ should be equally distributed in the state space, i.e. the region of interest. In order to check whether the flow vectors $\boldsymbol{f}_i$ are consistant with a continuous flow vector field, it is investigated whether (1) holds. As (1) is very sensitive to numeric errors it is investigated whether (1) holds on an average. The average is taken over all flow vectors $\boldsymbol{f}_i$, $\boldsymbol{f}_j$ with

$$|\boldsymbol{r}_i - \boldsymbol{r}_j| \in \left[ r - \frac{\Delta r}{2}, r + \frac{\Delta r}{2} \right], \quad \begin{array}{l} i \neq j, \quad i = 1, \ldots, K, \\ j = 1, \ldots, K, \end{array}$$

$$W(D, r) = \lim_{\Delta r \to 0} \overline{|\boldsymbol{f}_i - \boldsymbol{f}_j|} = r^* \,\mathrm{const} + O(r^2) . \quad (2)$$

0932-0784 / 87 / 0800-0797 $ 01.30/0. – Please order a reprint rather than making your own copy.

$W(D, r)$ is called crossover probability. If (2) does not hold, the dimension of the state space will be enlarged. The minimal dimension $D$ where (2) holds is called crossover dimension $D_c$.

Due to averaging, this does not give exactly Whitney's embedding dimension [19]. If a flow vector field is continuous, $W(D, r)$ for $r \to 0$ vanishes in contrast to statistical data where $W(D, r=0) > 0$, as the flow vectors are not correlated at neighbouring points in the state space.

In Fig. 1 the crossover probability $W$ is plotted versus $r$ for the chaotic dynamics of the Lorenz attractor [23, 11]. In order to calculate the crossover probability $W$, the attractor represented by $N$ states and $N$ flow vectors was examined at $K$ positions. A linear dependence on the radius $r$ results for all dimensions $D > 2$ (Figure 1). As the extrapolation of $W(D, r)$ for $r \to 0$ is not zero for $D = 1$ and $D = 2$ but tends to zero within the numeric error bars for $D > 2$, the crossover dimension of the Lorenz attractor becomes $D_c = 3$. This result is consistent, because the Lorenz attractor is described by a system of differential equations with three variables. The embedding dimension estimated from the Hausdorff dimension of the Lorenz attractor [10] is 6 [11, 20, 21].

A harmonic oscillator

$$\dot{x}(t) = \begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix} = \begin{pmatrix} x_2 \\ -w^2 x_1 \end{pmatrix}, \quad x(t=0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

$w = 0.04$ represented in a state space with the coordinates

$$y_i(t) = x_1(t + (i-1)\,\mathrm{d}t), \tag{3}$$

$$i = 0, \ldots, D-1, \quad \mathrm{d}t = \frac{\pi}{2w}$$

($D$ is the dimension of the state space)

should be free of any crossover in 2 dimensions. The numerical estimates are shown in Figure 2. The numerical value of the inclination $I = 0.043 \pm 0.003$ of the resulting straight line agrees with the analytic result $W(2, r) = w \cdot r$ for small $r$.

The error bars of $W(D, r)$ are estimated by the standard error function. An effect leading to errors takes place when $r$ is smaller than the average distance between the trajectories. In this case, all the flow vectors in the neighbourhood of a state come from the same trajectory. The correlation can lead to an incorrect result and to error bars which are too small. This effect can be observed in Fig. 1 at $r = 0.05$ ($D = 2$). The breakdown of $W$ moves to smaller $r$ as the density of the trajectories is increased. In order to obtain reproducible results, at

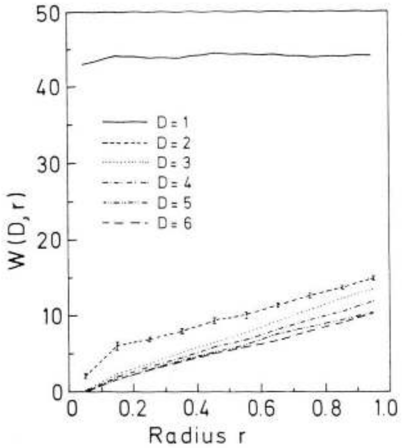


Fig. 1. The crossover probability versus the distance $r$ for the Lorenz attractor. If the dimension $D$ of the state space is three or higher, $W(D, r=0) = 0$ results by extrapolation ($\mathrm{d}t = 0.3$, $te = 0.005$, $N = 15\,000$, $K = 450$).

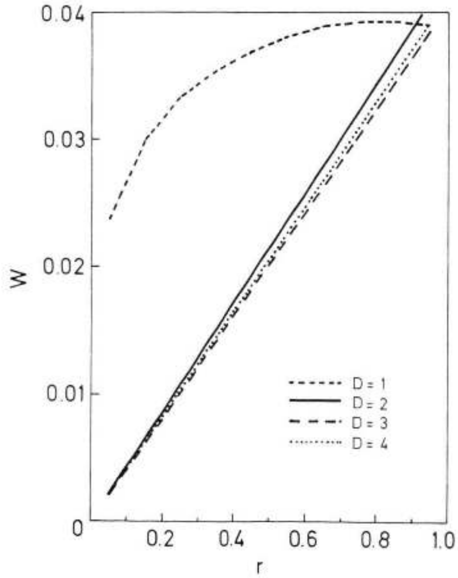


Fig. 2. The crossover probability $W$ of a harmonic oscillator (amplitude 1, frequency 0.04) versus the distance $r$. If the dimension $D$ of the state space is two or higher, $W(r=0) = 0$ results by extrapolation ($\mathrm{d}t = 30$, $te = 1.2$, $N = 10\,000$, $K = 50$).

least $10^4$ points in the state-space are required for the Lorenz attractor. For larger dimensions of the attractors (the fractal dimension of the Lorenz attractor [10] is 2.05), even more data are necessary. For the reconstruction of the attractor there are 3 typical time scales (Figure 3) for a typical oscillation period $T$: (i) the time delay $\mathrm{d}t = 0.25\,T$, if the coordinates of the state space are constructed by a delay [11], (ii) the small delay $te = 0.01\,T$ in order to calculate the flow vectors by numerical differentiation and (iii) $tt \gg T$, the time between two data points where the neighbourhood of the trajectory is examined. In order to obtain a good survey of the attractor, $tt$ must be chosen as large as possible. The large sets of data necessary for the reconstruction of the flow vector field of the attractor do not have to be measured continuously.

In a representation without any crossover the variation rate of the flow vectors as a function of the distance $r$ is given by $I = \mathrm{d}W/\mathrm{d}r$. $I$ is therefore correlated with the average curvature of the trajectories. The distinction between the average curvature and $I$ is given by the way in which the averaging process is performed. For the curvature only neighbouring positions on the same trajectory are included in the average. For $I$, however, the average is taken over the complete neighbourhood of the trajectory. Figure 4 represents the crossover probabilities $W(r)$ of a harmonic oscillator for different $\mathrm{d}t$ (3). A small average curvature implies that the attractor is unfolded. By a systematic variation of the geometrical representation of the attractor an optimal, i.e. unfolded representation can be found. When the elliptic trajectory of a harmonic oscillator in the state-space changes to a circular path, the slope of $I$ is minimal. For all $\mathrm{d}t = 40 + k^* 80$, $k = 0, 1, 2, \ldots$ the state space is completely unfolded (Figure 5). Representations taking $\mathrm{d}t = k^* 80$ are not continuous, because the harmonic system is represented as a straight double line in the state space.

## 3. The Influence of Noise

It is of interest to study the change of the crossover probability $W(D, r)$ when a time dependent random noise contribution $\boldsymbol{p}(t)$ is added to the deterministic system,

$$\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}) + \boldsymbol{p}(t)\,. \tag{4}$$

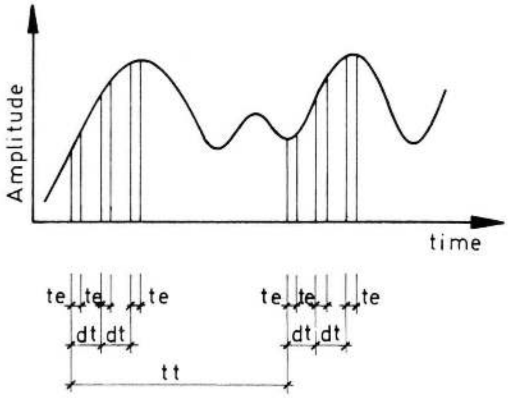


Fig. 3. Reconstruction of the flow vectors in a $D$-dimensional state space by a delay $\mathrm{d}t$. $a(t)$ is the experimental signal.

$$\boldsymbol{r}_i = \begin{pmatrix} a(t_0 + i \cdot tt) \\ a(t_0 + i \cdot tt + \mathrm{d}t) \\ \vdots \\ a(t_0 + i \cdot tt + (D-1)\,\mathrm{d}t) \end{pmatrix}$$

$i = 1, 2, \ldots, N$, $N$ is the number of points in the state space. The corresponding flow vectors are in a first order approximation.

$$\boldsymbol{f}_i = \frac{1}{t_e} \begin{pmatrix} a(t_0 + i \cdot tt + te) & - a(t_0 + i \cdot tt) \\ a(t_0 + i \cdot tt + \mathrm{d}t + te) & - a(t_0 + i \cdot tt + \mathrm{d}t) \\ \vdots & \\ a(t_0 + i \cdot tt + (D-1)\,\mathrm{d}t + te) - a(t_0 + i \cdot tt & \\ & + (D-1)\,\mathrm{d}t) \end{pmatrix}$$

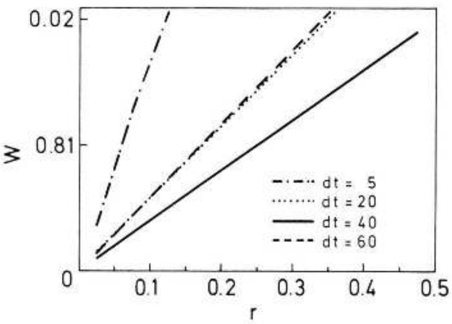


Fig. 4. The crossover probability $W$ of a harmonic oscillator (amplitude 1, frequency 0.04) versus $r$ and the delay $\mathrm{d}t$ for (3).

When for a deterministic system the crossover probability of the flow vectors $\boldsymbol{f}_i$ vanishes for $r \to 0$, for the same data the crossover probability is not equal to zero when a random noise contribution is added at each flow vector $\boldsymbol{f}_i \to \boldsymbol{f}_i + \boldsymbol{p}_i$ according to

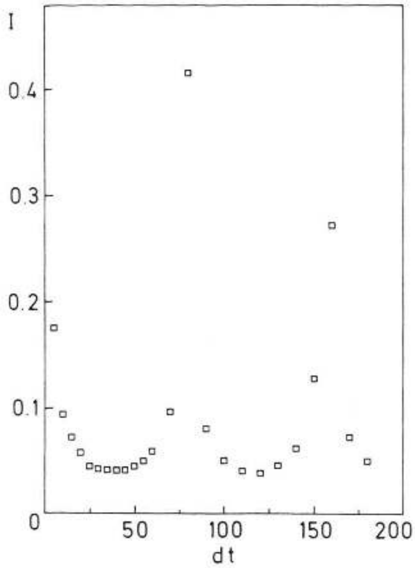


Fig. 5. $I = \mathrm{d}W(2, r)/\mathrm{d}r$ of a harmonic oscillator (amplitude 1, frequency 0.04) versus the delay $\mathrm{d}t$. For $\mathrm{d}t = 40 + k * 80$, i.e. a phaseshift of $\pi/2 + k\pi$, $k = 0, 1, 2, \ldots$, the geometrical reconstruction of the trajectories is unfolded for (3).

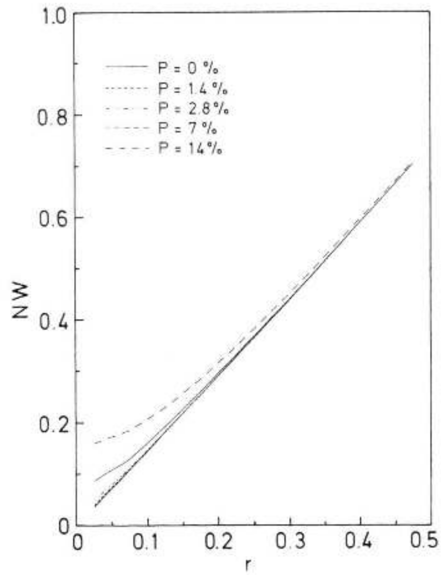


Fig. 6. The normalized crossover probability of a noisy harmonic oscillator ((4), amplitude 1, frequency 0.04). The length of the noise contributions $p_i$ is randomly chosen from the interval $[0, t_n p/2]$. The directions of $p_i$ are randomly chosen. For $W(r \to 0)$ the noise contribution $p$ results by extrapolation. $NW := W/t_n$.

(4) ($\boldsymbol{p}_i$ = noise contribution of the flow vector $\boldsymbol{f}_i$)

$$W(D, r = 0) = p^* t_N, \quad t_N = \frac{1}{N} \sum_{i=1}^{N} |\boldsymbol{f}_i|, \tag{5}$$

with

$$p := \lim_{\Delta r \to 0} \overline{|\boldsymbol{p}_i - \boldsymbol{p}_j|} / t_N .$$

The average is taken over all the noise contributions $\boldsymbol{p}_i, \boldsymbol{p}_j$ of the flow vectors $\boldsymbol{f}_i, \boldsymbol{f}_j$ with

$$|\boldsymbol{r}_i - \boldsymbol{r}_j| \in \left[0, \frac{\Delta r}{2}\right[, \quad \begin{array}{l} i \neq j, \quad i = 1, \ldots, K, \\ j = 1, \ldots, K. \end{array}$$

$t_N$ is the average length of the flow vectors. For small noise contributions, $p \ll 1$, and large $r$, results

$$W_{p \neq 0}(D, r) \approx W_{p=0}(D, r) \quad \text{for} \quad D \geqq D_c . \tag{6}$$

This means that for a small noise contribution and large $r$ the crossover probabilities with and without noise are nearly equal. For $D = 2$ the probability $W$ as a function of the noise $p$ for a harmonic oscillation is shown in Figure 6. In this case the following scalar product is used:

$$\boldsymbol{a} \cdot \boldsymbol{b} = \frac{1}{D} \sum_{i=1}^{D} a_i \cdot b_i , \quad \boldsymbol{a} = \begin{pmatrix} a_1 \\ \vdots \\ a_D \end{pmatrix}, \quad \boldsymbol{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_D \end{pmatrix},$$

$$a_i, b_i \in \mathbb{R} \quad \text{for} \quad i = 1, \ldots, -D. \tag{7}$$

If all coordinates of the flow have the same random noise contribution (e.g. due to the sampling method) the modified scalar product (7) should be used in order to make the effects of the noise independent of the dimension of the state space.

The deterministic terms in the differential equation are responsible for the extrapolated straight line passing through the origin. For $r \to 0$, the crossover-probability represents the random noise contribution.

## 4. Estimation of the Parameters of the Differential Equation

The complete reconstruction of the differential equation is based on a fit of the flow vector field by an ansatz. If nothing is known about the properties of the flow vector field, a fit by a polynom series is frequently favourable. The flow vector field of nearly all investigated chaotic oscillators, like e.g. the

Table 1. Reconstruction of the differential equation of a van der Pol oscillator (8) after the coordinate transformation (9). – An approximation of the flow vector field by a series of Legendre polynomials $P^k(x)$ of order $k$ in $x$

$$\dot{x}_j = \sum_{l,m=0}^{l+m<4} c_{j,l,m} P^l(x_1) P^m(x_2), \quad j=1,2,$$

by a least squares fit yields to the following coefficients:

| $j$ | $l$ | $m$ | $c_{j,l,m}$ | $j$ | $l$ | $m$ | $c_{j,l,m}$ |
|---|---|---|---|---|---|---|---|
| 1 | 0 | 0 | −.00 | 2 | 0 | 0 | .00 |
| 1 | 0 | 1 | −.69 | 2 | 0 | 1 | −.40 |
| 1 | 0 | 2 | −.00 | 2 | 0 | 2 | −.00 |
| 1 | 0 | 3 | .18 | 2 | 0 | 3 | −.17 |
| 1 | 1 | 0 | −.23 | 2 | 1 | 0 | 1.52 |
| 1 | 1 | 1 | −.00 | 2 | 1 | 1 | −.00 |
| 1 | 1 | 2 | −.23 | 2 | 1 | 2 | .51 |
| 1 | 2 | 0 | −.00 | 2 | 2 | 0 | .00 |
| 1 | 2 | 1 | .33 | 2 | 2 | 1 | −.39 |
| 1 | 3 | 0 | −.15 | 2 | 3 | 0 | .30 |

Error bars ± .005.

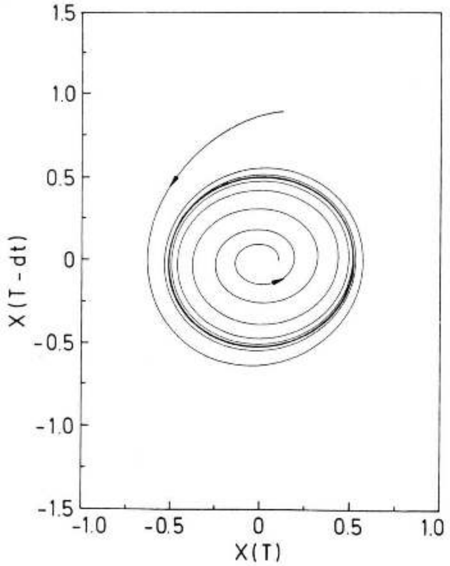


Fig. 7. A state space representation of the trajectories of the reconstructed differential equation. Again a limit cycle with radius 0.5 results.

Lorenz attractor, the driven Duffing oscillator or the Henon-Heiles system, can be described by a power series with only a few non zero coefficients. The mathematical condition for such an ansatz is that the flow vector field is an analytic function. To verify this, the $m$-fold differentiability ($m=1, 2, \ldots$) has to be examined in addition to the check for continuity (Chapter 2). Differentiability is difficult to check with experimentally investigated data. Anyway, in many cases the flow vector field is already analytic when the representation is continuous, and an ansatz for the flow vector field is justified (within the experimental uncertainties) if the statistical controls of a fit allow for it. The parameters of a successfully fitted ansatz are the parameters of the differential equation. The obtained differential equations possess naturally only validity in the range of values of the experimental data. The flow vector field of the Lorenz attractor

$$\dot{x}_1 = -10x_1 + 10x_2,$$
$$\dot{x}_2 = -x_1 x_3 + 28x_3 - x_2,$$
$$\dot{x}_3 = x_1 x_2 - \tfrac{8}{3} x_3$$

and of a van der Pol oscillator

$$\dot{x}_1 = x_2,$$
$$\dot{x}_2 = m c x_2 - \omega^2 x_1 - m\omega^2 x_1^2 x_2 - m x_2^3$$

with

$$m = 0.1, \quad \omega = 3, \quad c = 1,$$

has been fitted with a polynom of 4-th order. All coefficients of the original differential equation are reproduced to better than 1‰. The other coefficients are zero within $0 \pm 0.0005$. In order to investigate the efficiency of the algorithm we apply the following test:

$x$- and $y$-values are generated by using numerical integration of the differential equations

$$\dot{x} = y,$$
$$\dot{y} = m(c^2 - y^2 - w^2 x^2)\, y - w^2 x \tag{8}$$

with $m = 0.8$, $w = 1$, $c = 0.25$.

The radius of the resulting limit cycle is 0.5. The next step is a transformation of the coordinate-system:

$$x_1(t) = x(t),$$
$$x_2(t) = x(t - \mathrm{d}t), \tag{9}$$

where $\mathrm{d}t$ is approximately $\pi/2w$.

These new variables $x_1$ and $x_2$ serve for a new differential equation, for which the coefficients of a power expansion up to the 3rd order are given in Table 1. Legendre polynoms were used for fitting. When the flow vectors are fitted by a polynom up to the 4th order, all coefficients of the 4th order vanish and the coefficients determined by an approximation up to the 3rd order only vary within their error bars. This new differential equation shall now be numerically integrated. Again a limit cycle with radius 0.5 results (Figure 7).

We thank E. Lüscher and H. Haken for their continuous support and stimulating discussions. We should also like to express our gratitude to Dr. Kroy, Dr. Altmann, Dr. Wohofsky, Mr. Deisz, Dr. Berding and to Messerschmitt-Bölkow-Blohm for their very much appreciated support.

[1] For a review see: A J. Lichtenberg and M. A. Lieberman, Regular and Stochastic Motion. Springer-Verlag, New York 1982, Chapter 1.
[2] D. D. Humieres, M. R. Beasley, B. A. Huberman, and A. Libchaber, Phys. Rev. **A 26,** 3483 (1982).
[3] A. Babloyantz, J. M. Salazar, and C. Nicolis, Phys. Lett. **111 A,** 152 (1985).
[4] D. A. Weitz, J. S. Huang, M. Y. Lin, and J. Sung, Phys. Rev. Lett. **54,** 1416 (1985).
[5] J. Guckenheimer and G. Buzyna, Phys. Rev. Lett. **51,** 1438 (1983).
[6] H. L. Swinney, Physica **7 D,** 3 (1983).
[7] H. Haken, Synergetics, Springer, New York 1983, Chapter 8.
[8] D. Ruelle and F. Takens, Commun, Math. Phys. **20,** 167 (1971).
[9] H. Schuster, Deterministic Chaos, Physik-Verlag, Weinheim 1984, p. 175.
[10] P. Grassberger and I. Procaccia, Phys. Rev. Let. **50,** 346 (1983). – P. Grassberger and I. Procaccia, Physica **13 D,** 34 (1984).
[11] For a review see: J. P. Eckmann and D. Ruelle, Rev. Mod. Phys. **57,** 617 (1985).
[12] Y. Termonia, Phys. Rev. A **29,** 1612 (1984).
[13] B. Malraison, P. Atten, P. Berge, and M. Dubois, J. Physique **44,** L 897 (1983).
[14] H. G. Schuster, S. Martin, and W. Martienssen, Phys. Rev. A **33,** 3547 (1986).
[15] T. Klinker, W. Meyer-Ilse, and W. Lauterborn, Phys. Lett. **101 A,** 371 (1984) and a lot of other references in the review articles [1, 6, 7–9, 11].
[16] B. A. Huberman and J. P. Crutchfield, Phys. Rev. Lett. **43,** 1743 (1979).
[17] J. P. Crutchfield and B. A. Huberman, Phys. Lett. **77 A,** 407 (1980).
[18] U. Parlitz and W. Lauterborn, Phys. Lett. **107 A,** 351 (1985).
[19] H. Whitney, Ann. Math. **37,** 645 (1936).
[20] F. Takens, in: Dynamical Systems and Turbulence. Warwick 1980, Springer Lecture Notes in Mathematics **898,** 366 (1981).
[21] R. Mane, in: Dynamic Systems and Turbulence, Warwick 1980, Springer Lecture Notes in Mathematics **898,** 230 (1981).
[22] J. Guckenheimer and P. Holmes, Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields, Springer, New York 1983, Chapter 2.
[23] E. N. Lorenz, J. Atmos. Sci. **20,** 130 (1963).