Pergamon

Chaos, Solitons & Fractals, Vol. 8, No. 12, pp. 1911–1920, 1997

0960-0779/97 $17.00 + 0.00

PII: S0960-0779(97)00083-0

Quantification of Transients Using Empirical Orthogonal Functions

JÖRG ARNDT, HANSPETER HERZEL, SUMIT BOSE, MARTIN FALCKE and ECKEHARD SCHÖLL

Institut für Theoretische Physik, Technische Universität Berlin, D-10623 Berlin, Germany

(Accepted 14 March 1997)

Abstract—The method of empirical orthogonal functions (EOFs) is used to analyse spatio-temporal data obtained from simulations of heterogeneous catalysis. It is found that EOFs are well suited for the quantitative determination of transients in the data. The effectivity of the technique is demonstrated using synthetical data with an analytically known transient component. © 1997 Elsevier Science Ltd

1. INTRODUCTION

The typical problems that arise in the analysis of spatio-temporal data are different from those in the analysis of scalar data: long transients are common for such systems, more probable if a larger number of 'characteristic wavelengths' fit into the system length. For some systems, the transient time may be so long that one cannot expect ever to observe stationary dynamics [1]. If the transient time is unknown, it may be difficult or impossible to decide whether the examined data are part of the transient or not.

In a sense, the investigation of spatio-temporal data has experimental character, since partial differential equations (PDEs) exhibit potentially an infinite number of degrees of freedom. Consequently, techniques which have been developed for ordinary differential equations have to be used with caution. The analysis of transients using empirical orthogonal functions (EOFs) can be a useful first step.

The focus of this paper is the analysis of simulations of PDEs describing the heterogeneous catalytic oxidation of carbon monoxide, using EOFs. It turns out that the EOFs allow us to quantify the amount of transient behavior effectively. We illustrate the technique using synthetic spatio-temporal data with a well-defined superimposed transient.

2. THE METHOD OF EMPIRICAL ORTHOGONAL FUNCTIONS

The method of EOFs, also called the method of empirical eigenfunctions, proper orthogonal decomposition (POD), principal component analysis, singular spectrum analysis, Karhunen–Loéve expansion or bi-orthogonal decomposition, is a pattern recognition algorithm that uses linear correlations in space to find coherent structures, the EOFs (empirical eigenfunctions, topos or empirical modes) in spatio-temporal data [2–9].

By a 'photo', we mean a vector holding the spatial data samples at a certain time. Let a 'movie' (space-time data) be written as an $X \times T$ matrix, either as a column of T photos $\vec{p}_i$ or as a row of X columns $\vec{a}_k$, each describing the dynamics at one location:

$$\mathbf{A} := (\vec{a}_1 \quad \vec{a}_2 \quad \cdots \quad \vec{a}_X) = \begin{pmatrix} \vec{p}_1 \\ \vec{p}_2 \\ \vdots \\ \vec{p}_T \end{pmatrix}. \tag{1}$$

For data sets with two or more spatial dimensions, the notation need not be changed, photos of more dimensions can be represented as row vectors by simply identifying the vector components with the locations of the photos in any arbitrary order. For the definition of the empirical modes, one first has to define the covariance matrix, which is a measure for the linear interdependence of the dynamics at different locations. Each entry i,j of the covariance matrix (two-point correlation matrix) $\mathbf{R}$ is the covariance of location i with location j. Subject to the condition that the time average at every location is zero, it may be written as

$$\mathbf{R} := \begin{pmatrix} \langle \vec{a}_1 \vec{a}_1 \rangle & \langle \vec{a}_1 \vec{a}_2 \rangle & \cdots & \langle \vec{a}_1 \vec{a}_X \rangle \\ \langle \vec{a}_2 \vec{a}_1 \rangle & \langle \vec{a}_2 \vec{a}_2 \rangle & \cdots & \langle \vec{a}_2 \vec{a}_X \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle \vec{a}_X \vec{a}_1 \rangle & \langle \vec{a}_X \vec{a}_2 \rangle & \cdots & \langle \vec{a}_X \vec{a}_X \rangle \end{pmatrix} = \langle \mathbf{A}^{\mathrm{T}} \mathbf{A} \rangle. \tag{2}$$

The brackets $\langle\ \rangle$ denote time averages.

The EOFs $\Phi_i(x)$ of the movie $\mathbf{A}$ are defined as the eigenvectors $\vec{e}_i$ of the symmetric matrix $\mathbf{R}$, sorted with respect to the size of the eigenvalues ω_i (largest first).

The sequence of scalar products of the photos with one particular mode is called the mode amplitude (or chronos) $c_i(t)$ of that mode. This is a time series-like data set. For the spatio-temporal data obtained from a linear system (e.g. coupled linear oscillators), this method finds the normal modes of the system as empirical eigenfunctions [3].

The EOFs are often used for the reduction of data since the output of a simulation of partial differential equations with many spatial sampling points quickly leads to huge amounts of data. The empirical modes may be used to characterise the 'patterns' of motion; each mode amplitude can then be examined separately with the methods of time series analysis to characterize the type of the dynamics.

In [5] for example, mode amplitudes are used to characterize regular and chaotic vocal fold vibrations. An example for a semiconductor laser model is provided in [6]. An application to semiconductor transport is given in [7]. In principle, EOFs can be exploited to project PDEs on low-dimensional systems of ODEs. This is exemplified in [8] for a hydrodynamic flow of complex geometry. Examples of data reduction using EOFs are given in [5] and [9].

3. DATA FROM A MODEL OF HETEROGENEOUS CATALYSIS

The system examined is that of catalytic oxidation of CO on a Pt(110) single crystal surface in a reactor at low pressures ($p < 10^{-3}$ torr).

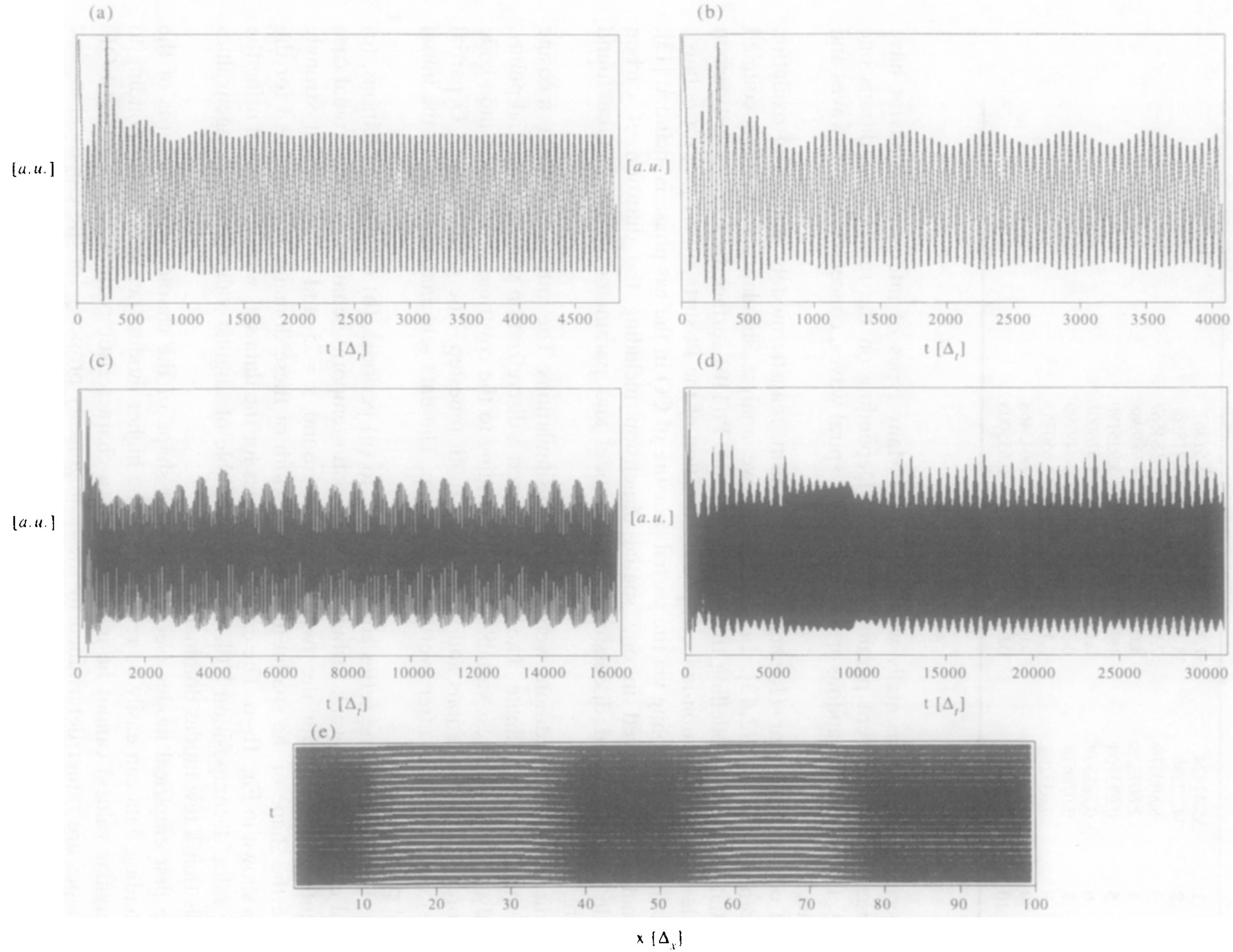

Fig. 1. Time series of the partial pressure of carbon monoxide in the heterogeneous catalysis including the initial transient for different control parameters: (a) periodic dynamics, (b) toroidal dynamics, (c) temporal chaos, (d) weak turbulence. (e) The spatio-temporal dynamics of the CO coverage for the data set of (b). Time and space are plotted in units of the sampling intervals Δ_t and Δ_x, respectively.

Table 1. List of the first ten singular values for the data sets of Fig. 1. They are normalized to sum up to 100%. Each value quantifies the relative contribution of the corresponding empirical mode to the original data

#	(a) Periodic	(b) Torus	(c) Temporal chaos	(d) Weak turbulence
1	55.47424	51.72563	43.57567	35.9689
2	34.27286	33.14202	29.31886	27.5575
3	6.609278	7.599817	7.684667	7.82502
4	2.607328	4.024955	3.989968	5.80164
5	0.623500	1.905375	3.925190	4.50114
6	0.268399	0.755899	2.918001	4.09786
7	0.076638	0.337746	2.849799	2.95492
8	0.033294	0.197319	1.650753	2.22570
9	0.011695	0.082795	1.042468	1.61913
10	0.005611	0.053195	0.655008	1.05982

This system is experimentally well examined. Many types of interesting behavior have been observed for different parameter values. Depending on the reaction conditions one observes, for example, standing or moving waves, spiral waves, concentric elliptic waves and turbulence [11].

Based on the Langmuir–Hinshelwood mechanism, realistic models of the CO oxidation have been developed [12, 13]. In this paper, we analyse simulations of three coupled differential equations modelling the coverage of the Pt(110) surface by oxygen and carbon monoxide and an absorption-induced phase transition of the Pt surface (see [12] for details). In particular, global coupling via the partial pressure of CO in the gas phase is included [13]. The equations are solved in one spatial dimension including the diffusion of carbon monoxide only. Detailed discussions of the model and parameter values can be found elsewhere [12, 13].

Periodic boundary conditions were used in all simulations. The initial condition is a cosine wave in the spatial coordinate. The simulation used a discretization with 1000 spatial points. The analysed dimensionless variable u corresponding to the coverage with carbon monoxide was recorded at 100 locations and after every 160th timestep. The global data (CO partial pressure) were recorded after every 16th timestep. The data sets analysed include the initial transient phase.

Figure 1 shows the time series corresponding to (a) periodic, (b) toroidal oscillations, (c) temporal chaos and (d) weak turbulence. The spatio-temporal behavior of the toroidal case is presented in (e). Note the two 'domains' around $x = 25$ and $x = 65$ which strongly influence the shape of the spatial modes. The width of these domains is modulated for the data sets shown in Fig. 1b–d. The empirical modes for the data sets were computed using the singular value decomposition [10]. Table 1, the table of singular values, i.e. the eigenvalues ω_i, shows that a few modes dominate.

In the first empirical modes (see Fig. 2), the shape of the characteristic patterns of the corresponding data can easily be recognized. The higher modes (i.e. those corresponding to small singular values) cannot be interpreted as a characteristic pattern found by the EOF. Their shapes are rather determined by the orthogonality property for the empirical modes: they generally tend to become more oscillatory in the regions of the borders between the patterns.

The 'x-shaped' amplitude of the fifth empirical mode (see Fig. 3) looks very different from the first four mode amplitudes. Such x-shaped mode amplitudes were found several times when a POD was performed on different data sets. In order to rule out that this was an artefact of the specific implementation of the POD, the spatio-temporal data were cut into

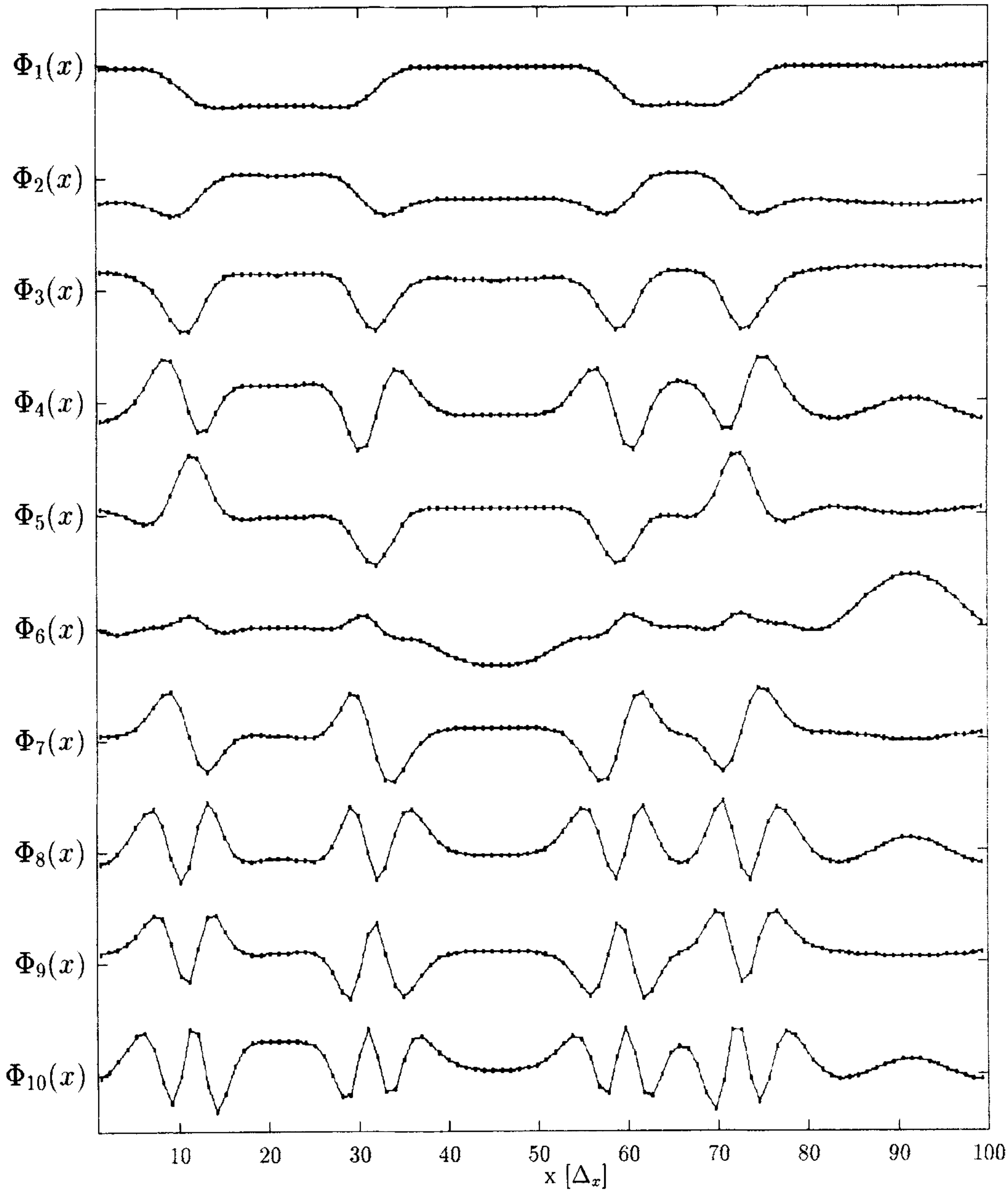

Fig. 2. The first ten empirical modes $\Phi_k(x)$ for the data set of Fig. 1c. Each mode is a vector of norm one.

blocks of equal length in time. The blocks were swapped and the resulting data set was used as a test input for the POD. A flaw in the implementation should have been detected if the new mode amplitude was of a similar shape as the original one. The result, however, was exactly as predicted by the definition of the POD (see Fig. 4): as it is insensitive to a permutation of the photos, the POD of any permutation of the input photos should lead to the same modes and corresponding amplitudes which are permutated in the same way as the photos.

For data without transients, such x-shaped amplitudes never occurred. This is a strong hint that empirical modes with modulated amplitudes should be considered as effects of

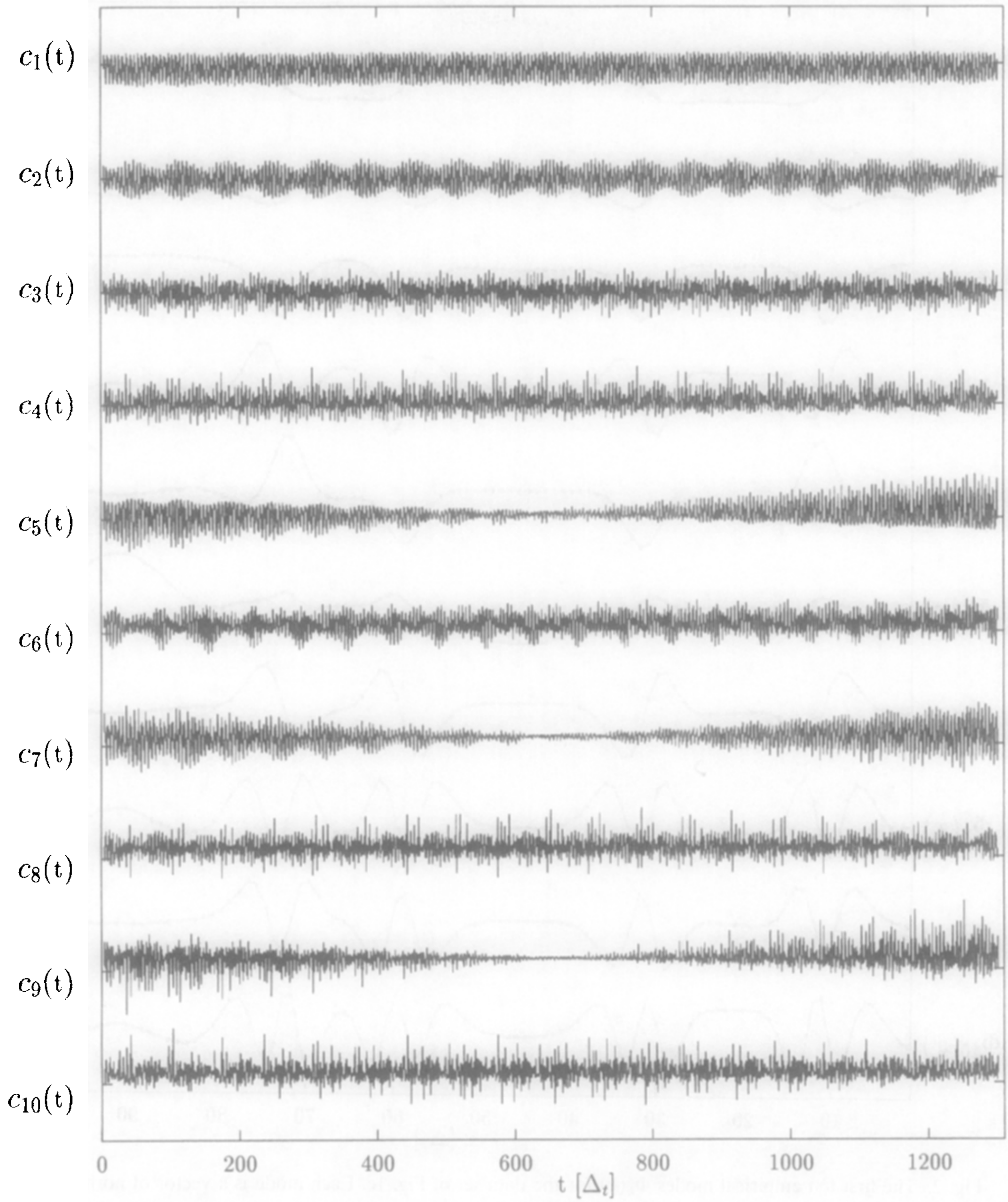

Fig. 3. The first ten mode amplitudes $c_k(t)$ for the data set of Fig. 1c. Each mode amplitude is a vector of norm one.

transients in the data. The specific forms of the mode amplitudes are due to the orthogonality property of the amplitudes from a POD: even if only a part of the data set is used as input for the POD, the mode amplitude describing the transient always keeps its x-shaped form. The POD adjusts the form of the modes in a way that the amplitudes describing transients always seem to suffer a phase rotation of 180 degrees. If short parts of a data set are used, the number of the first x-shaped mode amplitude shifts towards higher values. This can be expected because shorter data sets only contain a smaller portion of

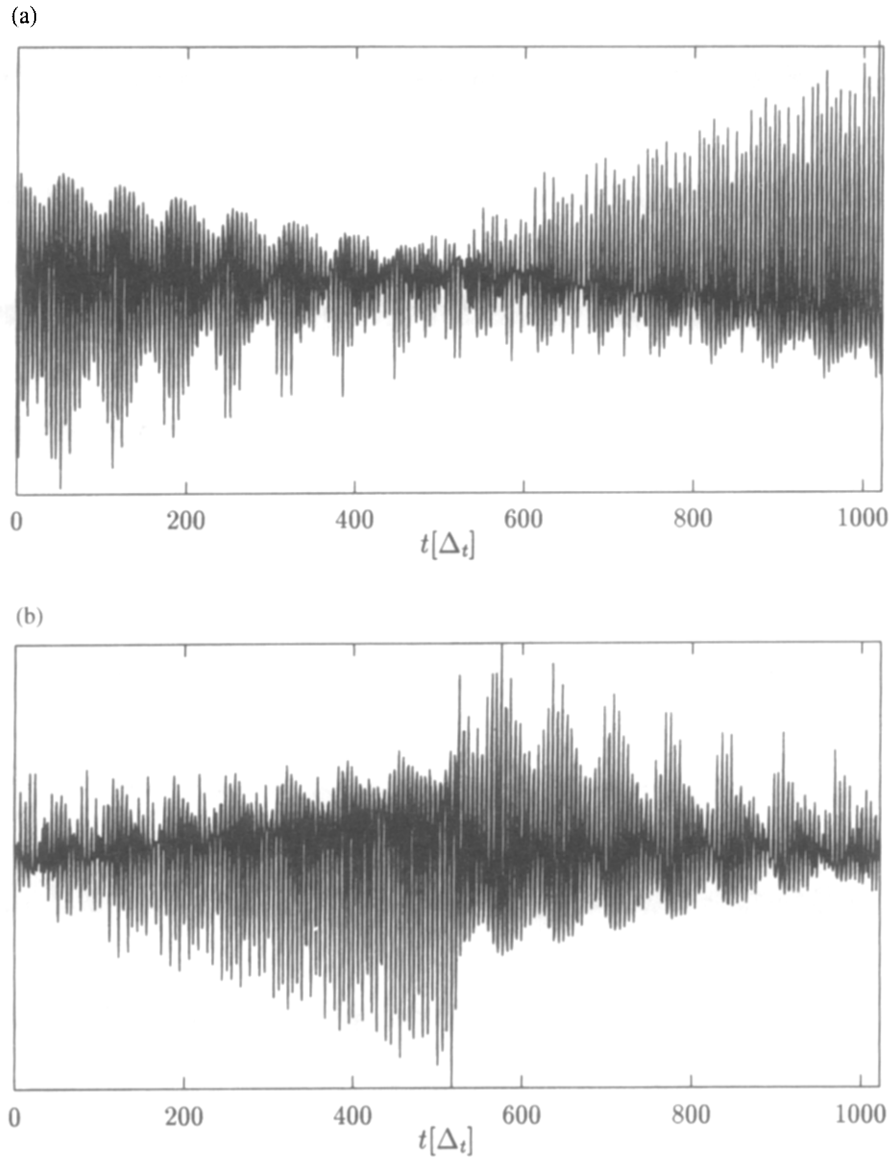

Fig. 4. Typical 'x-shaped' mode amplitude. (a) Fifth mode amplitude for the data set of Fig. 1c; the corresponding singular value is 3.9%. (b) Mode amplitude of the same data with the first and the second half of the block swapped.

transient behavior. The corresponding singular values give a quantitative estimate of the percentage of transient behavior.

Among the mode amplitudes of the data set of Fig. 1d (see Fig. 5) there are also 'irregular' ones (numbers 4, 6 and 9) which seem to suffer a phase rotation. But their form is not due to transients; they reflect slow fluctuations of the borders of the stripe patterns in the data. Again the corresponding singular values give a measure for the amount of these weak and slow modulations.

This means that the POD can supply valuable information for data sets where it strictly speaking cannot be applied. If the transient or turbulent behavior of a system is not too strong, the POD produces modes that can be assigned to these instationary parts of the

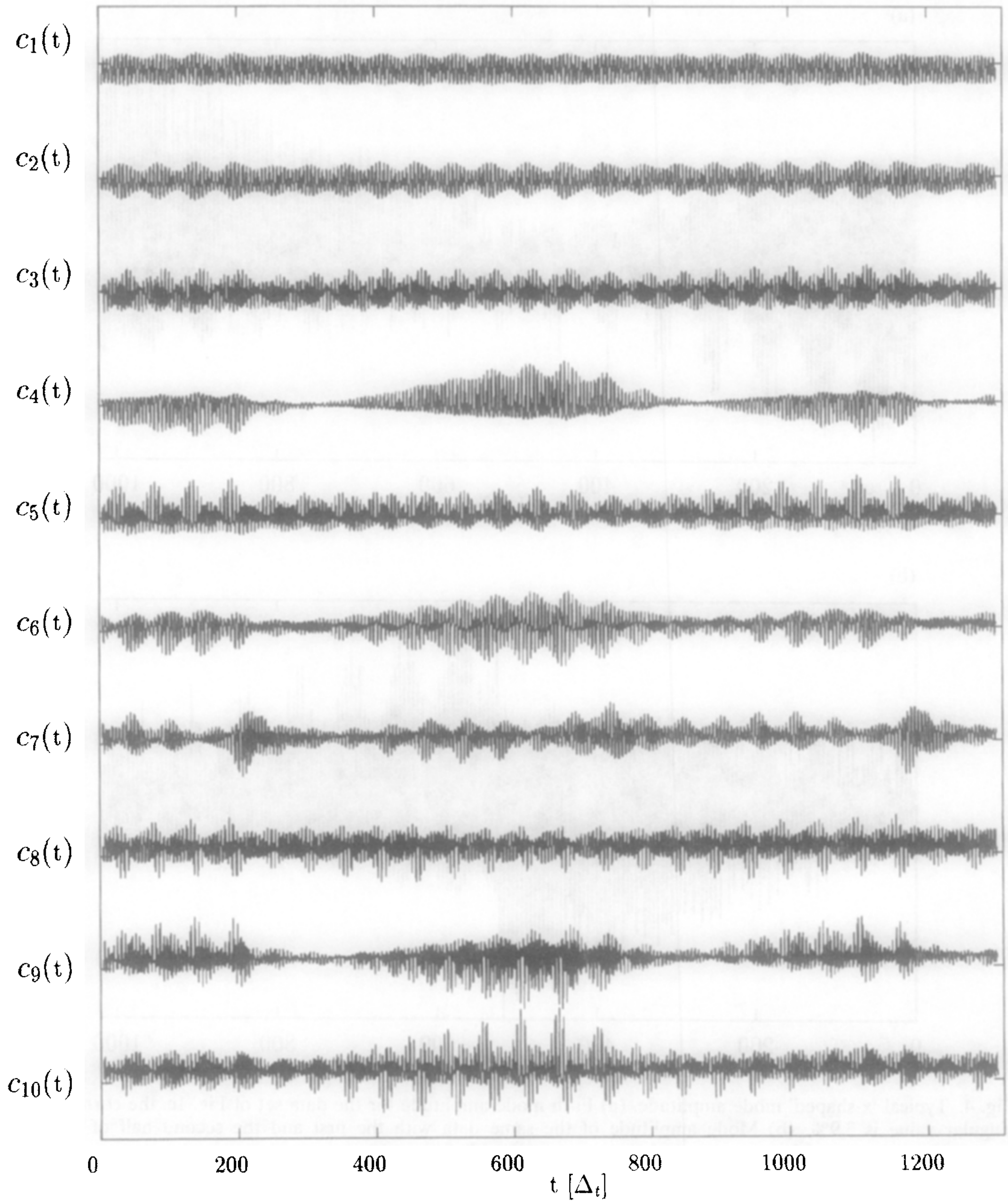

Fig. 5. The first ten mode amplitudes $c_k(t)$ for the data set of Fig. 1d.

dynamics. These modes and their amplitudes indicate where and when that behavior occurs; the singular values give a measure for its strength.

4. SYNTHETICAL DATA

The applicability of EOFs to the detection of transients is now corroborated by means of artificially generated data. These data exhibit spatial and temporal modulations with three

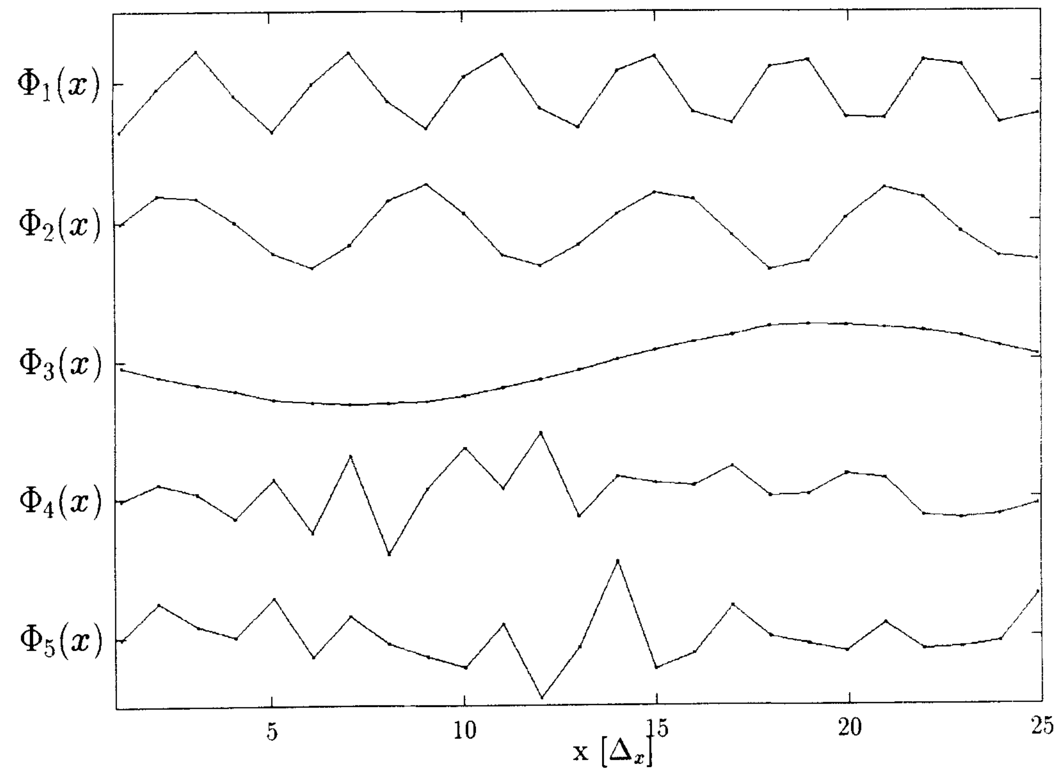

Fig. 6. The first five empirical modes $\Phi_k(x)$ for the synthetical data set with $X = 25$, $T = 200$, $A_c = 2$, $\Omega_c = 1.61$, $\omega_c = 0.61$, $A_s = 1$, $\Omega_s = 1$, $\omega_s = 1$, $A_t = 0.1$, $\Omega_t = 2\pi/X$, $\omega_t = 0.5$, $\tau = 100$, $A_n = 0.01$.

independent frequencies, white noise, and a relatively weak exponential decay which mimics a transient component.

The synthetical data set is defined by

$$\phi(x,t) = A_c \cos(\Omega_c x)\cos(\omega_c t) + A_s \sin(\Omega_s x)\sin(\omega_s t)$$
$$+ A_t \sin(\Omega_t x)\cos(\omega_t t)\exp(-t/\tau) + A_n e(x,t),$$
$$\text{for } x \in [0,1,\ldots,X-1],\ t \in [0,1,\ldots,T-1], \quad (3)$$

where $e(x,t)$ is white noise with unity variance.

These data were examined by the method of the EOFs. The resulting modes (see Fig. 6) and amplitudes (see Fig. 7) can be interpreted easily: the empirical modes correspond to the different spatial contributions in the trigonometric terms, and the mode amplitudes correspond to the time-dependent terms. Thus $c_1(t)$, $c_2(t)$ and $c_3(t)$ precisely map the $\cos(\omega_c t)$, $\cos(\omega_s t)$ and $\cos(\omega_t t)$ terms, respectively. The singular values coincide (within the error bounds) with the variance of the individual terms (64.1% for 'cos cos' and 31.6% for 'sin sin'). In particular, the third mode amplitude (1.61%), which represents the weak exponential modulation, describes indeed the transient component. The higher-order singular values are below 1% and reflect the noise contribution.

5. CONCLUSIONS

This work is devoted to the analysis of spatio-temporal data with empirical orthogonal functions. For the simulation of heterogeneous catalysis, complex regimes have been

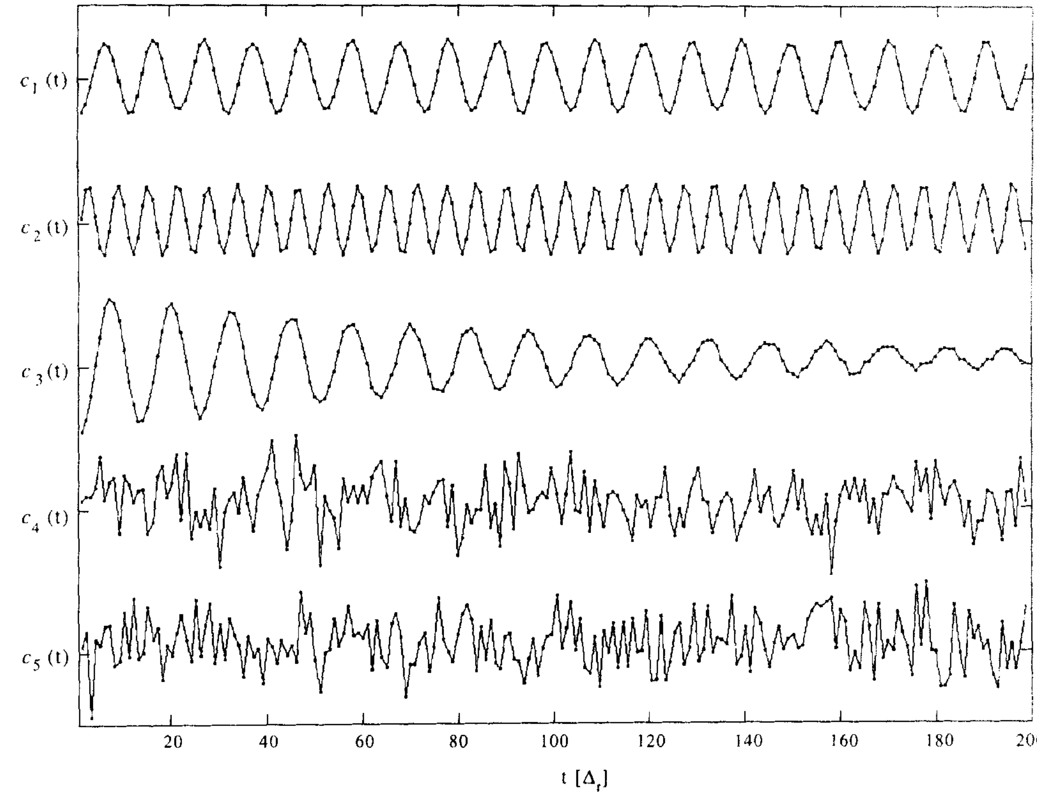

Fig. 7. The first five mode amplitudes $c_k(x)$ for the synthetical data set.

studied. Toroidal dynamics is covered by the first modes. In chaotic data, transients and low frequency modulations occur. These appear as x-shaped higher-order modes.

In order to verify the capability of empirical orthogonal functions to detect weak transients, we have generated synthetical data. These simulations show that even the quantitative amount of the transient can be estimated from the corresponding singular value.

In summary, we emphasize that, in addition to the previously known features of empirical orthogonal functions (such as data reduction, filtering and Galerkin projection), another important aspect has been discovered: higher-order modes can visualize (and even quantify) weak transients and modulations of spatiotemporal data.

REFERENCES

1. Crutchfield, J. P. and Kaneko, K., *Phys. Rev. Lett.*, 1988, **60,** 2715.
2. Lumley, J. L., *Stochastic Tools in Turbulence*. Academic Press, New York, 1970.
3. Sirovich, L., *Quart. Appl. Math.*, 1987, **XLV,** 561.
4. Holden, A. V., Poole, M. J., Tucker, J. V. and Zhang, H., *Chaos, Solitons & Fractals*, 1994, **4,** 2248.
5. Berry, D. A., Titze, I. R., Herzel, H. and Krischer, K., *J. Acoust. Soc. Am.*, 1994, **95,** 3595.
6. Hess, O. and Schöll, E., *Phys. Rev. A*, 1994, **50,** 787.
7. Meixner, M., Bose, S. and Schöll, E., *Phys. D* 1997, *in press*.
8. Deane, A. E., Kevrekidis, I. G., Karniadakis, G. E. and Orszag, S. A., *Phys. Fluids A*, 1991, **3,** 2337.
9. Graham, M. D., Kevrekidis, I. G., Hudson, J. L., Veser, G., Krischer, K. and Imbihl, R., *Chaos, Solitons & Fractals*, 1995, **5,** 1817.
10. Press, W. H., Teukolsky, S. A., Vetterling, W. T. and Flannery, B. P., *Numerical Recipes in C*. Cambridge University Press, Cambridge, 1992.
11. Ertl, G., *Science*, 1991, **254,** 1750.
12. Krischer, K., Eiswirth, M. and Ertl, G., *J. Chem. Phys.*, 1992, **96,** 9161.
13. Falcke, M., Engel, H. and Neufeld, M., *Phys. Rev. E*, 1995, **52,** 763.